Problems and Solutions in Mathematical Finance

Stochastic Calculus

Eric Chin, Dian Nel and Sverrir Ólafsson
Problems and Solutions
in Mathematical Finance
For other titles in the Wiley Finance Series
please see www.wiley.com/finance
Problems and Solutions in Mathematical Finance

Volume 1: Stochastic Calculus

Eric Chin, Dian Nel and Sverrir Ólafsson

WILEY
"the beginning of a task is the biggest step"

Plato, *The Republic*
## Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>ix</td>
</tr>
<tr>
<td>Prologue</td>
<td>xi</td>
</tr>
<tr>
<td>About the Authors</td>
<td>xv</td>
</tr>
<tr>
<td><strong>1 General Probability Theory</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Problems and Solutions</td>
<td>4</td>
</tr>
<tr>
<td>1.2.1 Probability Spaces</td>
<td>4</td>
</tr>
<tr>
<td>1.2.2 Discrete and Continuous Random Variables</td>
<td>11</td>
</tr>
<tr>
<td>1.2.3 Properties of Expectations</td>
<td>41</td>
</tr>
<tr>
<td><strong>2 Wiener Process</strong></td>
<td>51</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>51</td>
</tr>
<tr>
<td>2.2 Problems and Solutions</td>
<td>55</td>
</tr>
<tr>
<td>2.2.1 Basic Properties</td>
<td>55</td>
</tr>
<tr>
<td>2.2.2 Markov Property</td>
<td>68</td>
</tr>
<tr>
<td>2.2.3 Martingale Property</td>
<td>71</td>
</tr>
<tr>
<td>2.2.4 First Passage Time</td>
<td>76</td>
</tr>
<tr>
<td>2.2.5 Reflection Principle</td>
<td>84</td>
</tr>
<tr>
<td>2.2.6 Quadratic Variation</td>
<td>89</td>
</tr>
<tr>
<td><strong>3 Stochastic Differential Equations</strong></td>
<td>95</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>95</td>
</tr>
<tr>
<td>3.2 Problems and Solutions</td>
<td>102</td>
</tr>
<tr>
<td>3.2.1 Itô Calculus</td>
<td>102</td>
</tr>
<tr>
<td>3.2.2 One-Dimensional Diffusion Process</td>
<td>123</td>
</tr>
<tr>
<td>3.2.3 Multi-Dimensional Diffusion Process</td>
<td>155</td>
</tr>
<tr>
<td><strong>4 Change of Measure</strong></td>
<td>185</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>185</td>
</tr>
<tr>
<td>4.2 Problems and Solutions</td>
<td>192</td>
</tr>
<tr>
<td>4.2.1 Martingale Representation Theorem</td>
<td>192</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.2 Girsanov’s Theorem</td>
<td>194</td>
</tr>
<tr>
<td>4.2.3 Risk-Neutral Measure</td>
<td>221</td>
</tr>
<tr>
<td>5 Poisson Process</td>
<td>243</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>243</td>
</tr>
<tr>
<td>5.2 Problems and Solutions</td>
<td>251</td>
</tr>
<tr>
<td>5.2.1 Properties of Poisson Process</td>
<td>251</td>
</tr>
<tr>
<td>5.2.2 Jump Diffusion Process</td>
<td>281</td>
</tr>
<tr>
<td>5.2.3 Girsanov’s Theorem for Jump Processes</td>
<td>298</td>
</tr>
<tr>
<td>5.2.4 Risk-Neutral Measure for Jump Processes</td>
<td>322</td>
</tr>
</tbody>
</table>

Appendix A Mathematics Formulae | 331 |
Appendix B Probability Theory Formulae | 341 |
Appendix C Differential Equations Formulae | 357 |
Bibliography | 365 |
Notation | 369 |
Index | 373 |
Preface

Mathematical finance is based on highly rigorous and, on occasions, abstract mathematical structures that need to be mastered by anyone who wants to be successful in this field, be it working as a quant in a trading environment or as an academic researcher in the field. It may appear strange, but it is true, that mathematical finance has turned into one of the most advanced and sophisticated field in applied mathematics. This development has had considerable impact on financial engineering with its extensive applications to the pricing of contingent claims and synthetic cash flows as analysed both within financial institutions (investment banks) and corporations. Successful understanding and application of financial engineering techniques to highly relevant and practical situations requires the mastering of basic financial mathematics. It is precisely for this purpose that this book series has been written.

In Volume I, the first of a four volume work, we develop briefly all the major mathematical concepts and theorems required for modern mathematical finance. The text starts with probability theory and works across stochastic processes, with main focus on Wiener and Poisson processes. It then moves to stochastic differential equations including change of measure and martingale representation theorems. However, the main focus of the book remains practical. After being introduced to the fundamental concepts the reader is invited to test his/her knowledge on a whole range of different practical problems. Whereas most texts on mathematical finance focus on an extensive development of the theoretical foundations with only occasional concrete problems, our focus is a compact and self-contained presentation of the theoretical foundations followed by extensive applications of the theory. We advocate a more balanced approach enabling the reader to develop his/her understanding through a step-by-step collection of questions and answers. The necessary foundation to solve these problems is provided in a compact form at the beginning of each chapter. In our view that is the most successful way to master this very technical field.

No one can write a book on mathematical finance today, not to mention four volumes, without being influenced, both in approach and presentation, by some excellent text books in the field. The texts we have mostly drawn upon in our research and teaching are (in no particular order of preference), Tomas Björk, Arbitrage Theory in Continuous Time; Steven Shreve, Stochastic Calculus for Finance; Marek Musiela and Marek Rutkowski, Martingale Methods in Financial Modelling and for the more practical aspects of derivatives John Hull, Options, Futures and
Other Derivatives. For the more mathematical treatment of stochastic calculus a very influential text is that of Bernt Øksendal, *Stochastic Differential Equations*. Other important texts are listed in the bibliography.

**Note to the student/reader.** Please try hard to solve the problems on your own before you look at the solutions!
Prologue

IN THE BEGINNING WAS THE MOTION...

The development of modern mathematical techniques for financial applications can be traced back to Bachelier’s work, Theory of Speculation, first published as his PhD Thesis in 1900. At that time Bachelier was studying the highly irregular movements in stock prices on the French stock market. He was aware of the earlier work of the Scottish botanist Robert Brown, in the year 1827, on the irregular movements of plant pollen when suspended in a fluid. Bachelier worked out the first mathematical model for the irregular pollen movements reported by Brown, with the intention to apply it to the analysis of irregular asset prices. This was a highly original and revolutionary approach to phenomena in finance. Since the publication of Bachelier’s PhD thesis, there has been a steady progress in the modelling of financial asset prices.

Few years later, in 1905, Albert Einstein formulated a more extensive theory of irregular molecular processes, already then called Brownian motion. That work was continued and extended in the 1920s by the mathematical physicist Norbert Wiener who developed a fully rigorous framework for Brownian motion processes, now generally called Wiener processes.

Other major steps that paved the way for further development of mathematical finance included the works by Kolmogorov on stochastic differential equations, Fama on efficient-market hypothesis and Samuelson on randomly fluctuating forward prices.

Further important developments in mathematical finance were fuelled by the realisation of the importance of Itô’s lemma in stochastic calculus and the Feynman-Kac formula, originally drawn from particle physics, in linking stochastic processes to partial differential equations of parabolic type. The Feynman-Kac formula provides an immensely important tool for the solution of partial differential equations “extracted” from stochastic processes via Itô’s lemma. The real relevance of Itô’s lemma and Feynman-Kac formula in finance were only realised after some further substantial developments had taken place.

The year 1973 saw the most important breakthrough in financial theory when Black and Scholes and subsequently Merton derived a model that enabled the pricing of European call and put options. Their work had immense practical implications and lead to an explosive increase in the trading of derivative securities on some major stock and commodity exchanges. However, the philosophical foundation of that approach, which is based on the construction of risk-neutral portfolios enables an elegant and practical way of pricing of derivative contracts, has had a lasting and revolutionary impact on the whole of mathematical finance. The development initiated by Black, Scholes and Merton was continued by various researchers, notably Harrison, Kreps and Pliska in 1980s. These authors established the hugely important role of
martingales and arbitrage theory for the pricing of a large class of derivative securities or, as they are generally called, contingent claims. Already in the Black, Scholes and Merton model the risk-neutral measure had been informally introduced as a consequence of the construction of risk-neutral portfolios. Harrison, Kreps and Pliska took this development further and turned it into a powerful and the most general tool presently available for the pricing of contingent claims.

Within the Harrison, Kreps and Pliska framework the change of numéraire technique plays a fundamental role. Essentially the price of any asset, including contingent claims, can be expressed in terms of units of any other asset. The unit asset plays the role of a numéraire. For a given asset and a selected numéraire we can construct a probability measure that turns the asset price, in units of the numéraire, into a martingale whose existence is equivalent to the absence of an arbitrage opportunity. These results amount to the deepest and most fundamental in modern financial theory and are therefore a core construct in mathematical finance.

In the wake of the recent financial crisis, which started in the second half of 2007, some commentators and academics have voiced their opinion that financial mathematicians and their techniques are to be blamed for what happened. The authors do not subscribe to this view. On the contrary, they believe that to improve the robustness and the soundness of financial contracts, an even better mathematical training for quants is required. This encompasses a better comprehension of all tools in the quant’s technical toolbox such as optimisation, probability, statistics, stochastic calculus and partial differential equations, just to name a few.

Financial market innovation is here to stay and not going anywhere, instead tighter regulations and validations will be the only way forward with deeper understanding of models. Therefore, new developments and market instruments requires more scrutiny, testing and validation. Any inadequacies and weaknesses of model assumptions identified during the validation process should be addressed with appropriate reserve methodologies to offset sudden changes in the market direction. The reserve methodologies can be subdivided into model (e.g., Black-Scholes or Dupire model), implementation (e.g., tree-based or Monte Carlo simulation technique to price the contingent claim), calibration (e.g., types of algorithms to solve optimization problems, interpolation and extrapolation methods when constructing volatility surface), market parameters (e.g., confidence interval of correlation marking between underlyings) and market risk (e.g., when market price of a stock is close to the option’s strike price at expiry time). These are the empirical aspects of mathematical finance that need to be a core part in the further development of financial engineering.

One should keep in mind that mathematical finance is not, and must never become, an esoteric subject to be left to ivory tower academics alone, but a powerful tool for the analysis of real financial scenarios, as faced by corporations and financial institutions alike. Mathematical finance needs to be practiced in the real world for it to have sustainable benefits. Practitioners must realise that mathematical analysis needs to be built on a clear formulation of financial realities, followed by solid quantitative modelling, and then stress testing the model. It is our view that the recent turmoil in financial markets calls for more careful application of quantitative techniques but not their abolishement. Intuition alone or behavioural models have their role to play but do not suffice when dealing with concrete financial realities such as, risk quantification and risk management, asset and liability management, pricing insurance contracts or complex financial instruments. These tasks require better and more relevant education for quants and risk managers.

Financial mathematics is still a young and fast developing discipline. On the other hand, markets present an extremely complex and distributed system where a huge number of interrelated
financial instruments are priced and traded. Financial mathematics is very powerful in pricing and managing a limited number of instruments bundled into a portfolio. However, modern financial mathematics is still rather poor at capturing the extremely intricate contractual inter-relationship that exists between large numbers of traded securities. In other words, it is only to a very limited extent able to capture the complex dynamics of the whole markets, which is driven by a large number of unpredictable processes which possess varying degrees of correlation. The emergent behaviour of the market is to an extent driven by these varying degrees of correlations. It is perhaps one of the major present day challenges for financial mathematics to join forces with modern theory of complexity with the aim of being able to capture the macroscopic properties of the market, that emerge from the microscopic interrelations between a large number of individual securities. That this goal has not been reached yet is no criticism of financial mathematics. It only bears witness to its juvenile nature and the huge complexity of its subject.

Solid training of financial mathematicians in a whole range of quantitative disciplines, including probability theory and stochastic calculus, is an important milestone in the further development of the field. In the process, it is important to realise that financial engineering needs more than just mathematics. It also needs a judgement where the quant should constantly be reminded that no two market situations or two market instruments are exactly the same. Applying the same mathematical tools to different situations reminds us of the fact that we are always dealing with an approximation, which reflects the fact that we are modelling stochastic processes i.e. uncertainties. Students and practitioners should always bear this in mind.
About the Authors

**Eric Chin** is a quantitative analyst at an investment bank in the City of London where he is involved in providing guidance on price testing methodologies and their implementation, formulating model calibration and model appropriateness on commodity and credit products. Prior to joining the banking industry he worked as a senior researcher at British Telecom investigating radio spectrum trading and risk management within the telecommunications sector. He holds an MSc in Applied Statistics and an MSc in Mathematical Finance both from University of Oxford. He also holds a PhD in Mathematics from University of Dundee.

**Dian Nel** has more than 10 years of experience in the commodities sector. He currently works in the City of London where he specialises in oil and gas markets. He holds a BEng in Electrical and Electronic Engineering from Stellenbosch University and an MSc in Mathematical Finance from Christ Church, Oxford University. He is a Chartered Engineer registered with the Engineering Council UK.

**Sverrir Ólafsson** is Professor of Financial Mathematics at Reykjavik University; a Visiting Professor at Queen Mary University, London and a director of Riskcon Ltd, a UK based risk management consultancy. Previously he was a Chief Researcher at BT Research and held academic positions at The Mathematical Departments of Kings College, London; UMIST Manchester and The University of Southampton. Dr Ólafsson is the author of over 95 refereed academic papers and has been a key note speaker at numerous international conferences and seminars. He is on the editorial board of three international journals. He has provided an extensive consultancy on financial risk management and given numerous specialist seminars to finance specialists. In the last five years his main teaching has been MSc courses on Risk Management, Fixed Income, and Mathematical Finance. He has an MSc and PhD in mathematical physics from the Universities of Tübingen and Karlsruhe respectively.
General Probability Theory

Probability theory is a branch of mathematics that deals with mathematical models of trials whose outcomes depend on chance. Within the context of mathematical finance, we will review some basic concepts of probability theory that are needed to begin solving stochastic calculus problems. The topics covered in this chapter are by no means exhaustive but are sufficient to be utilised in the following chapters and in later volumes. However, in order to fully grasp the concepts, an undergraduate level of mathematics and probability theory is generally required from the reader (see Appendices A and B for a quick review of some basic mathematics and probability theory). In addition, the reader is also advised to refer to the notation section (pages 369–372) on set theory, mathematical and probability symbols used in this book.

1.1 INTRODUCTION

We consider an experiment or a trial whose result (outcome) is not predictable with certainty. The set of all possible outcomes of an experiment is called the sample space and we denote it by \( \Omega \). Any subset \( A \) of the sample space is known as an event, where an event is a set consisting of possible outcomes of the experiment.

The collection of events can be defined as a subcollection \( \mathcal{F} \) of the set of all subsets of \( \Omega \) and we define any collection \( \mathcal{F} \) of subsets of \( \Omega \) as a field if it satisfies the following.

**Definition 1.1** The sample space \( \Omega \) is the set of all possible outcomes of an experiment or random trial. A field is a collection (or family) \( \mathcal{F} \) of subsets of \( \Omega \) with the following conditions:

(a) \( \emptyset \in \mathcal{F} \) where \( \emptyset \) is the empty set;
(b) if \( A \in \mathcal{F} \) then \( A^c \in \mathcal{F} \) where \( A^c \) is the complement of \( A \) in \( \Omega \);
(c) if \( A_1, A_2, \ldots, A_n \in \mathcal{F}, n \geq 2 \) then \( \bigcup_{i=1}^{n} A_i \in \mathcal{F} \) – that is to say, \( \mathcal{F} \) is closed under finite unions.

It should be noted in the definition of a field that \( \mathcal{F} \) is closed under finite unions (as well as under finite intersections). As for the case of a collection of events closed under countable unions (as well as under countable intersections), any collection of subsets of \( \Omega \) with such properties is called a \( \sigma \)-algebra.

**Definition 1.2** If \( \Omega \) is a given sample space, then a \( \sigma \)-algebra (or \( \sigma \)-field) \( \mathcal{F} \) on \( \Omega \) is a family (or collection) \( \mathcal{F} \) of subsets of \( \Omega \) with the following properties:

(a) \( \emptyset \in \mathcal{F} \);
(b) if \( A \in \mathcal{F} \) then \( A^c \in \mathcal{F} \) where \( A^c \) is the complement of \( A \) in \( \Omega \);
(c) if \( A_1, A_2, \ldots \in \mathcal{F} \) then \( \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \) – that is to say, \( \mathcal{F} \) is closed under countable unions.
We next outline an approach to probability which is a branch of measure theory. The reason for taking a measure-theoretic path is that it leads to a unified treatment of both discrete and continuous random variables, as well as a general definition of conditional expectation.

**Definition 1.3** The pair \((\Omega, \mathcal{F})\) is called a measurable space. A probability measure \(\mathbb{P}\) on a measurable space \((\Omega, \mathcal{F})\) is a function \(\mathbb{P} : \mathcal{F} \mapsto [0, 1]\) such that:

(a) \(\mathbb{P}(\emptyset) = 0\);
(b) \(\mathbb{P}(\Omega) = 1\);
(c) if \(A_1, A_2, \ldots \in \mathcal{F}\) and \(\{A_i\}_{i=1}^{\infty}\) is disjoint such that \(A_i \cap A_j = \emptyset, i \neq j\) then \(\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)\).

The triple \((\Omega, \mathcal{F}, \mathbb{P})\) is called a probability space. It is called a complete probability space if \(\mathcal{F}\) also contains subsets \(B\) of \(\Omega\) with \(\mathbb{P}\)-outer measure zero, that is \(\mathbb{P}^*(B) = \inf\{\mathbb{P}(A) : A \in \mathcal{F}, B \subset A\} = 0\).

By treating \(\sigma\)-algebras as a record of information, we have the following definition of a filtration.

**Definition 1.4** Let \(\Omega\) be a non-empty sample space and let \(T\) be a fixed positive number, and assume for each \(t \in [0, T]\) there is a \(\sigma\)-algebra \(\mathcal{F}_t\). In addition, we assume that if \(s \leq t\), then every set in \(\mathcal{F}_s\) is also in \(\mathcal{F}_t\). We call the collection of \(\sigma\)-algebras \(\mathcal{F}_t, 0 \leq t \leq T\), a filtration.

Below we look into the definition of a real-valued random variable, which is a function that maps a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) to a measurable space \(\mathbb{R}\).

**Definition 1.5** Let \(\Omega\) be a non-empty sample space and let \(\mathcal{F}\) be a \(\sigma\)-algebra of subsets of \(\Omega\). A real-valued random variable \(X\) is a function \(X : \Omega \mapsto \mathbb{R}\) such that \(\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}\) for each \(x \in \mathbb{R}\) and we say \(X\) is \(\mathcal{F}\) measurable.

In the study of stochastic processes, an adapted stochastic process is one that cannot “see into the future” and in mathematical finance we assume that asset prices and portfolio positions taken at time \(t\) are all adapted to a filtration \(\mathcal{F}_t\), which we regard as the flow of information up to time \(t\). Therefore, these values must be \(\mathcal{F}_t\) measurable (i.e., depend only on information available to investors at time \(t\)). The following is the precise definition of an adapted stochastic process.

**Definition 1.6** Let \(\Omega\) be a non-empty sample space with a filtration \(\mathcal{F}_t, t \in [0, T]\) and let \(X_t\) be a collection of random variables indexed by \(t \in [0, T]\). We therefore say that this collection of random variables is an adapted stochastic process if, for each \(t\), the random variable \(X_t\) is \(\mathcal{F}_t\) measurable.

Finally, we consider the concept of conditional expectation, which is extremely important in probability theory and also for its wide application in mathematical finance such as pricing options and other derivative products. Conceptually, we consider a random variable \(X\) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a sub-\(\sigma\)-algebra \(\mathcal{G}\) of \(\mathcal{F}\) (i.e., sets in \(\mathcal{G}\) are also in \(\mathcal{F}\)). Here \(X\) can represent a quantity we want to estimate, say the price of a stock in the future, while
1.1 INTRODUCTION

\( \mathcal{G} \) contains limited information about \( X \) such as the stock price up to and including the current time. Thus, \( \mathbb{E}(X|\mathcal{G}) \) constitutes the best estimation we can make about \( X \) given the limited knowledge \( \mathcal{G} \). The following is a formal definition of a conditional expectation.

**Definition 1.7 (Conditional Expectation)** Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \mathcal{G} \) be a sub-\( \sigma \)-algebra of \( \mathcal{F} \) (i.e., sets in \( \mathcal{G} \) are also in \( \mathcal{F} \)). Let \( X \) be an integrable (i.e., \( \mathbb{E}(|X|) < \infty \)) and non-negative random variable. Then the conditional expectation of \( X \) given \( \mathcal{G} \), denoted \( \mathbb{E}(X|\mathcal{G}) \), is any random variable that satisfies:

(a) \( \mathbb{E}(X|\mathcal{G}) \) is \( \mathcal{G} \) measurable;
(b) for every set \( A \in \mathcal{G} \), we have the partial averaging property

\[
\int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P} = \int_A X \, d\mathbb{P}.
\]

From the above definition, we can list the following properties of conditional expectation. Here \( (\Omega, \mathcal{F}, \mathbb{P}) \) is a probability space, \( \mathcal{G} \) is a sub-\( \sigma \)-algebra of \( \mathcal{F} \) and \( X \) is an integrable random variable.

- **Conditional probability.** If \( I_A \) is an indicator random variable for an event \( A \) then

  \[ \mathbb{E}(I_A|\mathcal{G}) = \mathbb{P}(A|\mathcal{G}). \]

- **Linearity.** If \( X_1, X_2, \ldots, X_n \) are integrable random variables and \( c_1, c_2, \ldots, c_n \) are constants then

  \[ \mathbb{E}(c_1X_1 + c_2X_2 + \ldots + c_nX_n|\mathcal{G}) = c_1\mathbb{E}(X_1|\mathcal{G}) + c_2\mathbb{E}(X_2|\mathcal{G}) + \ldots + c_n\mathbb{E}(X_n|\mathcal{G}). \]

- **Positivity.** If \( X \geq 0 \) almost surely then \( \mathbb{E}(X|\mathcal{G}) \geq 0 \) almost surely.

- **Monotonicity.** If \( X \) and \( Y \) are integrable random variables and \( X \leq Y \) almost surely then

  \[ \mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G}). \]

- **Computing expectations by conditioning.** \( \mathbb{E}[\mathbb{E}(X|\mathcal{G})] = \mathbb{E}(X) \).

- **Taking out what is known.** If \( X \) and \( Y \) are integrable random variables and \( X \) is \( \mathcal{G} \) measurable then

  \[ \mathbb{E}(XY|\mathcal{G}) = X \cdot \mathbb{E}(Y|\mathcal{G}). \]

- **Tower property.** If \( \mathcal{H} \) is a sub-\( \sigma \)-algebra of \( \mathcal{G} \) then

  \[ \mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}] = \mathbb{E}(X|\mathcal{H}). \]

- **Measurability.** If \( X \) is \( \mathcal{G} \) measurable then \( \mathbb{E}(X|\mathcal{G}) = X \).

- **Independence.** If \( X \) is independent of \( \mathcal{G} \) then \( \mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X) \).

- **Conditional Jensen’s inequality.** If \( \varphi : \mathbb{R} \mapsto \mathbb{R} \) is a convex function then

  \[ \mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi[\mathbb{E}(X|\mathcal{G})]. \]
1.2 PROBLEMS AND SOLUTIONS

1.2.1 Probability Spaces

1. **De Morgan’s Law.** Let $A_i$, $i \in I$ where $I$ is some, possibly uncountable, indexing set. Show that
   
   \[(a) \quad (\bigcup_{i \in I} A_i)^c = \bigcap_{i \in I} A_i^c ; \]
   
   \[(b) \quad (\bigcap_{i \in I} A_i)^c = \bigcup_{i \in I} A_i^c ; \]

   **Solution:**
   
   (a) Let $a \in (\bigcup_{i \in I} A_i)^c$ which implies $a \notin \bigcup_{i \in I} A_i$, so that $a \in A_i^c$ for all $i \in I$. Therefore,
   
   \[\left( \bigcup_{i \in I} A_i \right)^c \subseteq \bigcap_{i \in I} A_i^c,\]
   
   On the contrary, if we let $a \in \bigcap_{i \in I} A_i^c$ then $a \notin A_i$ for all $i \in I$ or $a \in (\bigcup_{i \in I} A_i)^c$ and hence
   
   \[\bigcap_{i \in I} A_i^c \subseteq \left( \bigcup_{i \in I} A_i \right)^c,\]

   Therefore, $(\bigcup_{i \in I} A_i)^c = \bigcap_{i \in I} A_i^c$.

   (b) From (a), we can write
   
   \[\left( \bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} (A_i^c)^c = \bigcap_{i \in I} A_i.\]

   Taking complements on both sides gives
   
   \[\left( \bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c.\]

2. Let $\mathcal{F}$ be a $\sigma$-algebra of subsets of the sample space $\Omega$. Show that if $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

   **Solution:** Given that $\mathcal{F}$ is a $\sigma$-algebra then $A_i^c, A_i^c, \ldots \in \mathcal{F}$ and $\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$. Furthermore, the complement of $\bigcup_{i=1}^{\infty} A_i^c$ is $\left( \bigcup_{i=1}^{\infty} A_i^c \right)^c \in \mathcal{F}$. Thus, from De Morgan’s law (see Problem 1.2.1.1, page 4) we have $(\bigcup_{i=1}^{\infty} A_i^c)^c = \bigcap_{i=1}^{\infty} (A_i^c)^c = \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

3. Show that if $\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$ then $\{\emptyset, \Omega\} \in \mathcal{F}$.

   **Solution:** $\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$, hence if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$. Since $\emptyset \in \mathcal{F}$ then $\emptyset^c = \Omega \in \mathcal{F}$. Thus, $\{\emptyset, \Omega\} \in \mathcal{F}$.
4. Show that if $A \subseteq \Omega$ then $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$ is a $\sigma$-algebra of subsets of $\Omega$.

**Solution:** $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$ is a $\sigma$-algebra of subsets of $\Omega$ since

(i) $\emptyset \in \mathcal{F}$.

(ii) For $\emptyset \in \mathcal{F}$ then $\emptyset^c = \Omega \in \mathcal{F}$. For $\Omega \in \mathcal{F}$ then $\Omega^c = \emptyset \in \mathcal{F}$. In addition, for $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$. Finally, for $A^c \in \mathcal{F}$ then $(A^c)^c = A \in \mathcal{F}$.

(iii) $\emptyset \cup \Omega = \Omega \in \mathcal{F}$, $\emptyset \cup A = A \in \mathcal{F}$, $\emptyset \cup A^c = A^c \in \mathcal{F}$, $\Omega \cup A = \Omega \in \mathcal{F}$, $\Omega \cup A^c = \Omega \in \mathcal{F}$ and $\Omega \cup A \cup A^c = \Omega \in \mathcal{F}$.

$\blacksquare$

5. Let $\{\mathcal{F}_i\}_{i \in I}$, $I \neq \emptyset$ be a family of $\sigma$-algebras of subsets of the sample space $\Omega$. Show that $\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$ is also a $\sigma$-algebra of subsets of $\Omega$.

**Solution:** $\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$ is a $\sigma$-algebra by taking note that

(a) Since $\emptyset \in \mathcal{F}_i$, $i \in I$ therefore $\emptyset \in \mathcal{F}$ as well.

(b) If $A \in \mathcal{F}_i$ for all $i \in I$ then $A^c \in \mathcal{F}_i$, $i \in I$. Therefore, $A \in \mathcal{F}$ and hence $A^c \in \mathcal{F}$.

(c) If $A_1, A_2, \ldots \in \mathcal{F}_i$ for all $i \in I$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_i$, $i \in I$ and hence $A_1, A_2, \ldots \in \mathcal{F}$ and $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

From the results of (a)–(c) we have shown $\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$ is a $\sigma$-algebra of $\Omega$.

$\blacksquare$

6. Let $\Omega = \{\alpha, \beta, \gamma\}$ and let

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{\alpha\}, \{\beta, \gamma\}\} \quad \text{and} \quad \mathcal{F}_2 = \{\emptyset, \Omega, \{\alpha, \beta\}, \{\gamma\}\}. $$

Show that $\mathcal{F}_1$ and $\mathcal{F}_2$ are $\sigma$-algebras of subsets of $\Omega$.

Is $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ also a $\sigma$-algebra of subsets of $\Omega$?

**Solution:** Following the steps given in Problem 1.2.1.4 (page 5) we can easily show $\mathcal{F}_1$ and $\mathcal{F}_2$ are $\sigma$-algebras of subsets of $\Omega$.

By setting $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 = \{\emptyset, \Omega, \{\alpha\}, \{\beta, \gamma\}, \{\alpha, \beta\}, \{\beta, \gamma\}\}$, and since $\{\alpha\} \in \mathcal{F}$ and $\{\gamma\} \in \mathcal{F}$, but $\{\alpha\} \cup \{\gamma\} = \{\alpha, \gamma\} \notin \mathcal{F}$, then $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ is not a $\sigma$-algebra of subsets of $\Omega$.

$\blacksquare$

7. Let $\mathcal{F}$ be a $\sigma$-algebra of subsets of $\Omega$ and suppose $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$ so that $\mathbb{P}(\Omega) = 1$. Show that $\mathbb{P}(\emptyset) = 0$.

**Solution:** Given that $\emptyset$ and $\Omega$ are mutually exclusive we therefore have

$$\emptyset \cap \Omega = \emptyset \quad \text{and} \quad \emptyset \cup \Omega = \Omega.$$ 

Thus, we can express

$$\mathbb{P}(\emptyset \cup \Omega) = \mathbb{P}(\emptyset) + \mathbb{P}(\Omega) - \mathbb{P}(\emptyset \cap \Omega) = 1.$$ 

Since $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\emptyset \cap \Omega) = 0$ therefore $\mathbb{P}(\emptyset) = 0$.

$\blacksquare$
8. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\mathbb{Q} : \mathcal{F} \rightarrow [0, 1]\) be defined by \(\mathbb{Q}(A) = \mathbb{P}(A|B)\) where \(B \in \mathcal{F}\) such that \(\mathbb{P}(B) > 0\). Show that \((\Omega, \mathcal{F}, \mathbb{Q})\) is also a probability space.

**Solution:** To show that \((\Omega, \mathcal{F}, \mathbb{Q})\) is a probability space we note that

(a) \(\mathbb{Q}(\emptyset) = \mathbb{P}(\emptyset|B) = \frac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = 0\).

(b) \(\mathbb{Q}(\Omega) = \mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1\).

(c) Let \(A_1, A_2, \ldots\) be disjoint members of \(\mathcal{F}\) and hence we can imply \(A_1 \cap B, A_2 \cap B, \ldots\) are also disjoint members of \(\mathcal{F}\). Therefore,

\[
\mathbb{Q}\left( \bigcup_{i=1}^{\infty} A_i \right) = \mathbb{P}\left( \bigcup_{i=1}^{\infty} A_i \Big| B \right) = \frac{\mathbb{P}\left( \bigcup_{i=1}^{\infty} (A_i \cap B) \right)}{\mathbb{P}(B)} = \sum_{i=1}^{\infty} \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_{i=1}^{\infty} \mathbb{Q}(A_i).
\]

Based on the results of (a)–(c), we have shown that \((\Omega, \mathcal{F}, \mathbb{Q})\) is also a probability space.

\(\square\)

9. **Boole’s Inequality.** Suppose \(\{A_i\}_{i \in I}\) is a countable collection of events. Show that

\[
\mathbb{P}\left( \bigcup_{i \in I} A_i \right) \leq \sum_{i \in I} \mathbb{P}(A_i).
\]

**Solution:** Without loss of generality we assume that \(I = \{1, 2, \ldots\}\) and define \(B_1 = A_1, B_i = A_i \setminus (A_1 \cup A_2 \cup \ldots \cup A_{i-1}), i \in \{2, 3, \ldots\}\) such that \(\{B_1, B_2, \ldots\}\) are pairwise disjoint and

\[
\bigcup_{i \in I} A_i = \bigcup_{i \in I} B_i.
\]

Because \(B_i \cap B_j = \emptyset, \ i \neq j, i, j \in I\) we have

\[
\mathbb{P}\left( \bigcup_{i \in I} A_i \right) = \mathbb{P}\left( \bigcup_{i \in I} B_i \right) = \sum_{i \in I} \mathbb{P}(B_i) = \sum_{i \in I} \mathbb{P}(A_i \setminus (A_1 \cup A_2 \cup \ldots \cup A_{i-1})) = \sum_{i \in I} \left\{ \mathbb{P}(A_i) - \mathbb{P}\left( A_i \cap (A_1 \cup A_2 \cup \ldots \cup A_{i-1}) \right) \right\} \leq \sum_{i \in I} \mathbb{P}(A_i).
\]

\(\square\)
1.2.1 Probability Spaces

10. **Bonferroni’s Inequality**. Suppose \( \{ A_i \}_{i \in I} \) is a countable collection of events. Show that

\[
P \left( \bigcap_{i \in I} A_i \right) \geq 1 - \sum_{i \in I} P \left( A_i^c \right).
\]

**Solution:** From De Morgan’s law (see Problem 1.2.1.1, page 4) we can write

\[
P \left( \bigcap_{i \in I} A_i \right) = P \left( \left( \bigcup_{i \in I} A_i^c \right)^c \right) = 1 - P \left( \bigcup_{i \in I} A_i^c \right).
\]

By applying Boole’s inequality (see Problem 1.2.1.9, page 6) we will have

\[
P \left( \bigcap_{i \in I} A_i \right) \geq 1 - \sum_{i \in I} P \left( A_i^c \right)
\]

since \( P \left( \bigcup_{i \in I} A_i^c \right) \leq \sum_{i \in I} P \left( A_i^c \right) \).

\[\blacklozenge\]

11. **Bayes’ Formula.** Let \( A_1, A_2, \ldots, A_n \) be a partition of \( \Omega \), where \( \bigcup_{i=1}^{n} A_i = \Omega \), \( A_i \cap A_j = \emptyset \), \( i \neq j \) and each \( A_i, i, j = 1, 2, \ldots, n \) has positive probability. Show that

\[
P (A_i | B) = \frac{P (B | A_i) P (A_i)}{\sum_{j=1}^{n} P (B | A_j) P (A_j)}.
\]

**Solution:** From the definition of conditional probability, for \( i = 1, 2, \ldots, n \)

\[
P (A_i | B) = \frac{P (A_i \cap B)}{P (B)} = \frac{P (B | A_i) P (A_i)}{P \left( \bigcup_{j=1}^{n} (B \cap A_j) \right)} = \frac{P (B | A_i) P (A_i)}{\sum_{j=1}^{n} P (B \cap A_j)} = \frac{P (B | A_i) P (A_i)}{\sum_{j=1}^{n} P (B | A_j) P (A_j)}.
\]

\[\blacklozenge\]

12. **Principle of Inclusion and Exclusion for Probability.** Let \( A_1, A_2, \ldots, A_n, n \geq 2 \) be a collection of events. Show that

\[
P (A_1 \cup A_2) = P (A_1) + P (A_2) - P (A_1 \cap A_2).
\]

From the above result show that

\[
P (A_1 \cup A_2 \cup A_3) = P (A_1) + P (A_2) + P (A_3) - P (A_1 \cap A_2) - P (A_1 \cap A_3) - P (A_2 \cap A_3) + P (A_1 \cap A_2 \cap A_3).
\]

Hence, using mathematical induction show that

\[
P \left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P (A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P (A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P (A_i \cap A_j \cap A_k).
\]
1.2.1 Probability Spaces

Finally, deduce that
\[
P\left(\bigcap_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i \cup A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i \cup A_j \cup A_k) \\
- \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n).
\]

Solution: For \( n = 2 \), \( A_1 \cup A_2 \) can be written as a union of two disjoint sets
\[
A_1 \cup A_2 = A_1 \cup (A_2 \setminus A_1) = A_1 \cup (A_2 \cap A_1^c).
\]

Therefore,
\[
P(A_1 \cup A_2) = P(A_1) + P(A_2 \cap A_1^c)
\]
and since \( P(A_2) = P(A_2 \cap A_1) + P(A_2 \cap A_1^c) \) we will have
\[
P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).
\]

For \( n = 3 \), and using the above results, we can write
\[
P(A_1 \cup A_2 \cup A_3) = P(A_1 \cup A_2) + P(A_3) - P[(A_1 \cup A_2) \cap A_3] \\
= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3).
\]

Since \( (A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3) \) therefore
\[
P[(A_1 \cup A_2) \cap A_3] = P[(A_1 \cap A_3) \cup (A_2 \cap A_3)] \\
= P(A_1 \cap A_3) + P(A_2 \cap A_3) - P[(A_1 \cap A_3) \cap (A_2 \cap A_3)] \\
= P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3).
\]

Thus,
\[
P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) \\
- P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3).
\]

Suppose the result is true for \( n = m \), where \( m \geq 2 \). For \( n = m + 1 \), we have
\[
P\left(\bigcup_{i=1}^{m} A_i \cup A_{m+1}\right) = P\left(\bigcup_{i=1}^{m} A_i\right) + P\left(A_{m+1}\right) - P\left(\bigcup_{i=1}^{m} (A_i \cap A_{m+1})\right) \\
= P\left(\bigcup_{i=1}^{m} A_i\right) + P\left(A_{m+1}\right) - P\left(\bigcup_{i=1}^{m} (A_i \cap A_{m+1})\right).
\]
By expanding the terms we have

\[
P\left( \bigcup_{i=1}^{m+1} A_i \right) = \sum_{i=1}^{m} P(A_i) - \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} P(A_i \cap A_j) + \sum_{i=1}^{m-2} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^{m} P(A_i \cap A_j \cap A_k)
\]

\[\ldots \, + (-1)^{m+1} P(A_1 \cap A_2 \cap \ldots \cap A_m)
\]

\[+ P(A_{m+1}) - P \left( (A_1 \cap A_{m+1}) \cup (A_2 \cap A_{m+1}) \ldots \cup (A_m \cap A_{m+1}) \right)
\]

\[= \sum_{i=1}^{m+1} P(A_i) - \sum_{i=1}^{m} \sum_{j=i+1}^{m} P(A_i \cap A_j) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^{m} P(A_i \cap A_j \cap A_k)
\]

\[\ldots \, + (-1)^{m+1} P(A_1 \cap A_2 \cap \ldots \cap A_m)
\]

\[+ \ldots \, + (-1)^{m+1} P(A_1 \cap A_2 \cap \ldots \cap A_m)\]

\[= \sum_{i=1}^{m+1} P(A_i) - \sum_{i=1}^{m} \sum_{j=i+1}^{m} P(A_i \cap A_j) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m-1} \sum_{k=j+1}^{m} P(A_i \cap A_j \cap A_k)
\]

\[\ldots \, + (-1)^{m+2} P(A_1 \cap A_2 \cap \ldots \cap A_{m+1}).\]

Therefore, the result is also true for \( n = m + 1 \). Thus, from mathematical induction we have shown for \( n \geq 2 \),

\[
P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k)
\]

\[\ldots \, + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n).
\]

From Problem 1.2.1.1 (page 4) we can write

\[
P\left( \bigcap_{i=1}^{n} A_i \right) = P\left( \left( \bigcup_{i=1}^{n} A_i^c \right)^c \right) = 1 - P\left( \bigcup_{i=1}^{n} A_i^c \right).
\]

Thus, we can expand

\[
P\left( \bigcap_{i=1}^{n} A_i \right) = 1 - \sum_{i=1}^{n} P(A_i^c) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i^c \cap A_j^c) - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i^c \cap A_j^c \cap A_k^c)
\]

\[\ldots \, + (-1)^{n+1} P(A_1^c \cap A_2^c \cap \ldots \cap A_n^c)\]

\[= 1 - \sum_{i=1}^{n} (1 - P(A_i)) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (1 - P(A_i \cup A_j))
\]

\[\ldots \, + (-1)^{n-1} \sum_{i=1}^{n} (1 - P(A_i \cup A_j \cup A_k))\].
1.2.1 Probability Spaces

\[ + \ldots - (-1)^{n+1}(1 - \mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n)) \]

\[ = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbb{P}(A_i \cup A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n} \sum_{k=j+1}^{n} \mathbb{P}(A_i \cup A_j \cup A_k) \]

\[ - \ldots + (-1)^{n+1} \mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n). \]

\[ \Box \]

13. **Borel–Cantelli Lemma.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(A_1, A_2, \ldots\) be sets in \(\mathcal{F}\). Show that

\[ \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \subseteq \bigcup_{k=m}^{\infty} A_k \]

and hence prove that

\[ \mathbb{P}\left( \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \right) = \begin{cases} 1 & \text{if } A_i \cap A_j = \emptyset, i \neq j \text{ and } \sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty \\ 0 & \text{if } \sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty. \end{cases} \]

**Solution:** Let \(B_m = \bigcup_{k=m}^{\infty} A_k\) and since \(\bigcap_{m=1}^{\infty} B_m \subseteq B_m\) therefore we have

\[ \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \subseteq \bigcup_{k=m}^{\infty} A_k. \]

From Problem 1.2.1.9 (page 6) we can deduce that

\[ \mathbb{P}\left( \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \right) \leq \mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k \right) \leq \sum_{k=m}^{\infty} \mathbb{P}(A_k). \]

For the case \(\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty\) and given it is a convergent series, then \(\lim_{m \to \infty} \sum_{k=m}^{\infty} \mathbb{P}(A_k) = 0\) and hence

\[ \mathbb{P}\left( \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \right) = 0 \]

if \(\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty.\)

For the case \(A_i \cap A_j = \emptyset, i \neq j\) and \(\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty,\) since \(\mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k \right) + \mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k^c \right) = 1\) therefore from Problem 1.2.1.1 (page 4)

\[ \mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k \right) = 1 - \mathbb{P}\left( \left( \bigcup_{k=m}^{\infty} A_k \right)^c \right) = 1 - \mathbb{P}\left( \bigcap_{k=m}^{\infty} A_k^c \right). \]
1.2.2 Discrete and Continuous Random Variables

From independence and because \( \sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty \) we can express

\[
\mathbb{P}\left( \bigcap_{k=m}^{\infty} A_k^c \right) = \prod_{k=m}^{\infty} \mathbb{P}(A_k^c) = \prod_{k=m}^{\infty} (1 - \mathbb{P}(A_k)) \leq \prod_{k=m}^{\infty} e^{-\mathbb{P}(A_k)} = e^{-\sum_{k=m}^{\infty} \mathbb{P}(A_k)} = 0
\]

for all \( m \) and hence

\[
\mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k \right) = 1 - \mathbb{P}\left( \bigcap_{k=m}^{\infty} A_k^c \right) = 1
\]

for all \( m \). Taking the limit \( m \to \infty \),

\[
\mathbb{P}\left( \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k \right) = \lim_{m \to \infty} \mathbb{P}\left( \bigcup_{k=m}^{\infty} A_k \right) = 1
\]

for the case \( A_i \cap A_j = \emptyset, \ i \neq j \) and \( \sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty \).

\[\square\]

1.2.2 Discrete and Continuous Random Variables

1. **Bernoulli Distribution.** Let \( X \) be a Bernoulli random variable, \( X \sim \text{Bernoulli}(p), p \in [0, 1] \) with probability mass function

\[
\mathbb{P}(X = 1) = p, \quad \mathbb{P}(X = 0) = 1 - p.
\]

Show that \( \mathbb{E}(X) = p \) and \( \text{Var}(X) = p(1 - p) \).

**Solution:** If \( X \sim \text{Bernoulli}(p) \) then we can write

\[
\mathbb{P}(X = x) = p^x(1 - p)^{1-x}, \quad x \in \{0, 1\}
\]

and by definition

\[
\mathbb{E}(X) = \sum_{x=0}^{1} x \mathbb{P}(X = x) = 0 \cdot (1 - p) + 1 \cdot p = p
\]

\[
\mathbb{E}(X^2) = \sum_{x=0}^{1} x^2 \mathbb{P}(X = x) = 0 \cdot (1 - p) + 1 \cdot p = p
\]

and hence

\[
\mathbb{E}(X) = p, \quad \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p(1 - p).
\]

\[\square\]
2. **Binomial Distribution.** Let \( \{X_i\}_{i=1}^n \) be a sequence of independent Bernoulli random variables each with probability mass function

\[
\mathbb{P}(X_i = 1) = p, \quad \mathbb{P}(X_i = 0) = 1 - p, \quad p \in [0, 1]
\]

and let

\[
X = \sum_{i=1}^n X_i.
\]

Show that \( X \) follows a binomial distribution, \( X \sim \text{Binomial}(n, p) \) with probability mass function

\[
\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n
\]

such that \( \mathbb{E}(X) = np \) and \( \text{Var}(X) = np(1-p) \).

Using the central limit theorem show that \( X \) is approximately normally distributed, \( X \sim \mathcal{N}(np, np(1-p)) \) as \( n \to \infty \).

**Solution:** The random variable \( X \) counts the number of Bernoulli variables \( X_1, \ldots, X_n \) that are equal to 1, i.e., the number of successes in the \( n \) independent trials. Clearly \( X \) takes values in the set \( N = \{0, 1, 2, \ldots, n\} \). To calculate the probability that \( X = k \), where \( k \in N \) is the number of successes we let \( E \) be the event such that \( X_{i_1} = X_{i_2} = \ldots = X_{i_k} = 1 \) and \( X_j = 0 \) for all \( j \in N \setminus S \) where \( S = \{i_1, i_2, \ldots, i_k\} \). Then, because the Bernoulli variables are independent and identically distributed,

\[
\mathbb{P}(E) = \prod_{j \in S} \mathbb{P}(X_j = 1) \prod_{j \in N \setminus S} \mathbb{P}(X_j = 0) = p^k (1-p)^{n-k}.
\]

However, as there are \( \binom{n}{k} \) combinations to select sets of indices \( i_1, \ldots, i_k \) from \( N \), which are mutually exclusive events, so

\[
\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.
\]

From the definition of the moment generating function of discrete random variables (see Appendix B),

\[
M_X(t) = \mathbb{E}(e^{tx}) = \sum_x e^{tx} \mathbb{P}(X = x)
\]

where \( t \in \mathbb{R} \) and substituting \( \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \) we have

\[
M_X(t) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^t)^x(1-p)^{n-x} = (1 - p + pe^t)^n.
\]
1.2.2 Discrete and Continuous Random Variables

By differentiating $M_X(t)$ with respect to $t$ twice we have

\[ M'_X(t) = np e^t (1 - p + pe^t)^{n-1} \]
\[ M''_X(t) = np e^t (1 - p + pe^t)^{n-1} + n(n-1)p^2 e^{2t} (1 - p + pe^t)^{n-2} \]

and hence

\[ \mathbb{E}(X) = M'_X(0) = np \]
\[ \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = M''_X(0) - M'_X(0)^2 = np(1-p). \]

Given the sequence $X_i \sim \text{Bernoulli}(p), i = 1, 2, \ldots, n$ are independent and identically distributed, each having expectation $\mu = p$ and variance $\sigma^2 = p(1-p)$, then as $n \to \infty$, from the central limit theorem

\[ \frac{\sum_{i=1}^n X_i - n\mu}{\sigma \sqrt{n}} \xrightarrow{D} \mathcal{N}(0,1) \]

or

\[ \frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{D} \mathcal{N}(0,1). \]

Thus, as $n \to \infty$, $X \approx \mathcal{N}(np, np(1-p))$. 

\[ \square \]

3. Poisson Distribution. A discrete Poisson distribution, Poisson($\lambda$) with parameter $\lambda > 0$ has the following probability mass function:

\[ \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots \]

Show that $\mathbb{E}(X) = \lambda$ and $\text{Var}(X) = \lambda$.

For a random variable following a binomial distribution, Binomial($n, p$), $0 \leq p \leq 1$ show that as $n \to \infty$ and with $p = \lambda/n$, the binomial distribution tends to the Poisson distribution with parameter $\lambda$.

\textbf{Solution:} From the definition of the moment generating function

\[ M_X(t) = \mathbb{E} \left( e^{tX} \right) = \sum_x e^{tx} \mathbb{P}(X = x) \]

where $t \in \mathbb{R}$ and substituting $\mathbb{P}(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$ we have

\[ M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{\lambda(e^t-1)}. \]

By differentiating $M_X(t)$ with respect to $t$ twice

\[ M'_X(t) = \lambda e^t e^{\lambda(e^t-1)}, \quad M''_X(t) = (\lambda e^t + 1)\lambda e^t e^{\lambda(e^t-1)} \]
1.2.2 Discrete and Continuous Random Variables

we have
\[ \mathbb{E}(X) = M'_X(0) = \lambda \]
\[ \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = M''_X(0) - M'_X(0)^2 = \lambda. \]

If \( X \sim \text{Binomial}(n, p) \) then we can write
\[
\mathbb{P}(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}
= \frac{n(n-1) \cdots (n-k+1)(n-k)!}{k!(n-k)!} \cdot p^k \left( 1 - (n-k)p + \frac{(n-k)(n-k-1)}{2!} p^2 + \ldots \right)
= \frac{n(n-1) \cdots (n-k+1)}{k!} p^k \left( 1 - (n-k)p + \frac{(n-k)(n-k-1)}{2!} p^2 + \ldots \right).
\]

For the case when \( n \to \infty \) so that \( n \gg k \) we have
\[
\mathbb{P}(X = k) \approx \frac{n^k}{k!} p^k \left( 1 - np + \frac{(np)^2}{2!} + \ldots \right)
= \frac{n^k}{k!} p^k e^{-np}.
\]

By setting \( p = \lambda/n \), we have \( \mathbb{P}(X = k) \approx \frac{\lambda^k}{k!} e^{-\lambda} \).

\[ \square \]

4. Exponential Distribution. Consider a continuous random variable \( X \) following an exponential distribution, \( X \sim \text{Exp}(\lambda) \) with probability density function
\[ f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0 \]
where the parameter \( \lambda > 0 \). Show that \( \mathbb{E}(X) = \frac{1}{\lambda} \) and \( \text{Var}(X) = \frac{1}{\lambda^2} \).

Prove that \( X \sim \text{Exp}(\lambda) \) has a memory less property given as
\[ \mathbb{P}(X > s + x | X > s) = \mathbb{P}(X > x) = e^{-\lambda x}, \quad x, s \geq 0. \]

For a sequence of Bernoulli trials drawn from a Bernoulli distribution, Bernoulli(\( p \)), \( 0 \leq p \leq 1 \) performed at time \( \Delta t, 2\Delta t, \ldots \) where \( \Delta t > 0 \) and if \( Y \) is the waiting time for the first success, show that as \( \Delta t \to 0 \) and \( p \to 0 \) such that \( p/\Delta t \) approaches a constant \( \lambda > 0 \), then \( Y \sim \text{Exp}(\lambda) \).

**Solution:** For \( t < \lambda \), the moment generating function for a random variable \( X \sim \text{Exp}(\lambda) \) is
\[
M_X(t) = \mathbb{E}(e^{tX}) = \int_0^\infty e^{tu} \lambda e^{-\lambda u} du = \lambda \int_0^\infty e^{-(\lambda-t)u} du = \frac{\lambda}{\lambda - t}.
\]
Differentiation of $M_X(t)$ with respect to $t$ twice yields

$$M_X'(t) = \frac{\lambda}{(\lambda - t)^2}, \quad M_X''(t) = \frac{2\lambda}{(\lambda - t)^3}.$$  

Therefore,

$$E(X) = M_X'(0) = \frac{1}{\lambda}, \quad E(X^2) = M_X''(0) = \frac{2}{\lambda^2},$$

and the variance of $X$ is

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{1}{\lambda^2}.$$

By definition

$$P(X > x) = \int_x^{\infty} \lambda e^{-\lambda u} du = e^{-\lambda x}$$

and

$$P(X > s + x \mid X > s) = \frac{P(X > s + x, X > s)}{P(X > s)} = \frac{P(X > s + x)}{P(X > s)} = \frac{\int_{s+x}^{\infty} \lambda e^{-\lambda u} du}{\int_x^{\infty} \lambda e^{-\lambda v} dv} = e^{-\lambda x}.$$

Thus,

$$P(X > s + x \mid X > s) = P(X > x).$$

If $Y$ is the waiting time for the first success then for $k = 1, 2, \ldots$

$$P(Y > k\Delta t) = (1 - p)^k.$$

By setting $y = k\Delta t$, and in the limit $\Delta t \to 0$ and assuming that $p \to 0$ so that $p/\Delta t \to \lambda$, for some positive constant $\lambda$,

$$P(Y > y) = P\left( Y > \left( \frac{x}{\Delta t} \right) \Delta t \right)$$

$$\approx (1 - \lambda \Delta t)^y/\Delta t$$

$$= 1 - \lambda y + \frac{\left( \frac{y}{\Delta t} \right)}{2!} (\lambda \Delta t)^2 + \ldots$$

$$\approx 1 - \lambda y + \frac{(\lambda y)^2}{2!} + \ldots$$

$$= e^{-\lambda x}.$$

In the limit $\Delta t \to 0$ and $p \to 0$,

$$P(Y \leq y) = 1 - P(Y > y) \approx 1 - e^{-\lambda y}$$

and the probability density function is therefore

$$f_Y(y) = \frac{d}{dy} P(Y \leq y) \approx \lambda e^{-\lambda y}, y \geq 0.$$
5. **Gamma Distribution.** Let $U$ and $V$ be continuous independent random variables and let $W = U + V$. Show that the probability density function of $W$ can be written as

\[
f_W(w) = \int_{-\infty}^{\infty} f_V(w-u)f_U(u) \, du = \int_{-\infty}^{\infty} f_U(w-v)f_V(v) \, dv
\]

where $f_U(u)$ and $f_V(v)$ are the density functions of $U$ and $V$, respectively.

Let $X_1, X_2, \ldots, X_n \sim \text{Exp}(\lambda)$ be a sequence of independent and identically distributed random variables, each following an exponential distribution with common parameter $\lambda > 0$. Prove that if

\[Y = \sum_{i=1}^{n} X_i\]

then $Y$ follows a gamma distribution, $Y \sim \text{Gamma}(n, \lambda)$ with the following probability density function:

\[f_Y(y) = \frac{(\lambda y)^{n-1}}{(n-1)!} e^{-\lambda y}, \quad y \geq 0.\]

Show also that $\mathbb{E}(Y) = \frac{n}{\lambda}$ and $\text{Var}(Y) = \frac{n}{\lambda^2}$.

**Solution:** From the definition of the cumulative distribution function of $W = U + V$ we obtain

\[F_W(w) = \mathbb{P}(W \leq w) = \mathbb{P}(U + V \leq w) = \int_{u+v \leq w} f_{UV}(u, v) \, dudv\]

where $f_{UV}(u, v)$ is the joint probability density function of $(U, V)$. Since $U \perp V$ therefore $f_{UV}(u, v) = f_U(u)f_V(v)$ and hence

\[
F_W(w) = \int_{u+v \leq w} f_U(u)f_V(v) \, dudv
\]

\[
= \int_{u+v \leq w} \left\{ \int_{-\infty}^{u-w} f_V(v) \, dv \right\} f_U(u) \, du
\]

\[
= \int_{-\infty}^{\infty} f_V(w-u)f_U(u) \, du.
\]

By differentiating $F_W(w)$ with respect to $w$, we have the probability density function $f_W(w)$ given as

\[
f_W(w) = \frac{d}{dw} \int_{-\infty}^{\infty} f_V(w-u)f_U(u) \, du = \int_{-\infty}^{\infty} f_Y(w-u)f_U(u) \, du.
\]
1.2.2 Discrete and Continuous Random Variables

Using the same steps we can also obtain
\[ f_W(w) = \int_{-\infty}^{\infty} f_U(w-v)f_V(v) \, dv. \]

To show that \( Y = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \lambda) \) where \( X_1, X_2, \ldots, X_n \sim \text{Exp}(\lambda) \), we will prove the result via mathematical induction.

For \( n = 1 \), we have \( Y = X_1 \sim \text{Exp}(\lambda) \) and the gamma density \( f_Y(y) \) becomes
\[ f_Y(y) = \lambda e^{-\lambda y}, \quad y \geq 0. \]

Therefore, the result is true for \( n = 1 \).

Let us assume that the result holds for \( n = k \) and we now wish to compute the density for the case \( n = k+1 \). Since \( X_1, X_2, \ldots, X_{k+1} \) are all mutually independent and identically distributed, by setting \( U = \sum_{i=1}^{k} X_i \) and \( V = X_{k+1} \), and since \( U \geq 0, V \geq 0 \), the density of \( Y = \sum_{i=1}^{k+1} X_i \) can be expressed as
\[
    f_Y(y) = \int_{0}^{y} f_V(y-u)f_U(u) \, du \\
    = \int_{0}^{y} \lambda e^{-\lambda(y-u)} \cdot \frac{(\lambda u)^{k-1}}{(k-1)!} \lambda e^{-\lambda u} \, du \\
    = \frac{\lambda^{k+1} e^{-\lambda y}}{(k-1)!} \int_{0}^{y} u^{k-1} \, du \\
    = \frac{(\lambda y)^k}{k!} \lambda e^{-\lambda y}
\]

which shows the result is also true for \( n = k + 1 \). Thus, \( Y = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \lambda) \).

Given that \( X_1, X_2, \ldots, X_n \sim \text{Exp}(\lambda) \) are independent and identically distributed with common mean \( \frac{1}{\lambda} \) and variance \( \frac{1}{\lambda^2} \), therefore
\[
    \mathbb{E}(Y) = \mathbb{E}\left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{n}{\lambda}
\]

and
\[
    \text{Var}(Y) = \text{Var}\left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i) = \frac{n}{\lambda^2}.
\]

\[ \square \]

6. Normal Distribution Property I. Show that for constants \( a, L \) and \( U \) such that \( t > 0 \) and \( L < U \),
\[
    \frac{1}{\sqrt{2\pi t}} \int_{L}^{U} e^{\frac{-(w-L)^2}{2t}} \, dw = e^{\frac{1}{2at^2}} \left[ \Phi\left( \frac{U-at}{\sqrt{t}} \right) - \Phi\left( \frac{L-at}{\sqrt{t}} \right) \right]
\]

where \( \Phi(\cdot) \) is the cumulative distribution function of a standard normal.
Solution: Simplifying the integrand we have

\[
\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w-at}{\sqrt{t}} \right)^2} \, dw = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w-2at}{\sqrt{t}} \right)^2} \, dw \\
= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w-at}{\sqrt{t}} \right)^2} \, dw \\
= \frac{e^{\pi^2 t}}{\sqrt{2\pi t}} \int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w-at}{\sqrt{t}} \right)^2} \, dw.
\]

By setting \( x = \frac{w-at}{\sqrt{t}} \), we can write

\[
\int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w-at}{\sqrt{t}} \right)^2} \, dw = \int_{-\infty}^{U \sqrt{t}} e^{-\frac{1}{2} x^2} \, dx = \Phi \left( \frac{U - at}{\sqrt{t}} \right) - \Phi \left( \frac{L - at}{\sqrt{t}} \right).
\]

Thus, \( \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{U} e^{-\frac{1}{2} \left( \frac{w}{\sqrt{t}} \right)^2} \, dw = e^{\frac{\pi^2 t}{2}} \left[ \Phi \left( \frac{U - at}{\sqrt{t}} \right) - \Phi \left( \frac{L - at}{\sqrt{t}} \right) \right]. \)

\[\square\]

7. Normal Distribution Property II. Show that if \( X \sim \mathcal{N}(\mu, \sigma^2) \) then for \( \delta \in \{ -1, 1 \}, \)

\[
\mathbb{E}[\max\{\delta(e^X - K), 0\}] = \delta e^{\mu + \frac{1}{2} \sigma^2} \Phi \left( \frac{\delta(\mu + \sigma^2 - \log K)}{\sigma} \right) - \delta K \Phi \left( \frac{\delta(\mu - \log K)}{\sigma} \right)
\]

where \( K > 0 \) and \( \Phi(\cdot) \) denotes the cumulative standard normal distribution function.

Solution: We first let \( \delta = 1, \)

\[
\mathbb{E}[\max\{e^X - K, 0\}] = \int_{\log K}^{\infty} (e^x - K) \, f_X(x) \, dx \\
= \int_{\log K}^{\infty} (e^x - K) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx \\
= \int_{\log K}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx - K \int_{\log K}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx.
\]

By setting \( w = \frac{x-\mu}{\sigma} \) and \( z = w - \sigma \) we have

\[
\mathbb{E}[\max\{e^X - K, 0\}] = \int_{\log K - \mu}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} w^2} \, dw - K \int_{\log K - \mu}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} w^2} \, dw
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{\log K - \mu}^{\infty} e^{-\frac{1}{2} w^2} \, dw - K \frac{1}{\sqrt{2\pi}} \int_{\log K - \mu}^{\infty} e^{-\frac{1}{2} w^2} \, dw.
\]
1.2.2 Discrete and Continuous Random Variables

\[ = e^{\mu + \frac{1}{2} \sigma^2} \int_{\log K - \mu}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(u - \sigma^2)^2} du - K \int_{\log K - \mu}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \]

\[ = e^{\mu + \frac{1}{2} \sigma^2} \int_{\log K - \mu - \sigma^2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz - K \int_{\log K - \mu}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \]

\[ = e^{\mu + \frac{1}{2} \sigma^2} \Phi \left( \frac{\mu + \sigma^2 - \log K}{\sigma} \right) - K \Phi \left( \frac{\mu - \log K}{\sigma} \right). \]

Using similar steps for the case \( \delta = -1 \) we can also show that

\[ \mathbb{E}[\max\{K - \mathcal{X}, 0\}] = K \Phi \left( \frac{\log K - \mu}{\sigma} \right) - e^{\mu + \frac{1}{2} \sigma^2} \Phi \left( \frac{\log K - (\mu + \sigma^2)}{\sigma} \right). \]

\[ \square \]

8. For \( x > 0 \) show that

\[ \frac{1}{x} - e^{-\frac{1}{2}x^2} \leq \int_{x}^{\infty} e^{-\frac{1}{2}z^2} dz \leq \frac{1}{x \sqrt{2\pi}} \]

**Solution:** Solving \( \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \) using integration by parts, we let \( u = \frac{z}{\sqrt{2\pi}} \), \( du = \frac{1}{\sqrt{2\pi}} dz \) and \( v = -e^{-\frac{1}{2}z^2} \). Therefore,

\[ \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \left[ 1 - \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \right] \]

\[ = \frac{1}{x \sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \]

\[ \leq \frac{1}{x \sqrt{2\pi}} e^{-\frac{1}{2}x^2} \]

since \( \int_{x}^{\infty} e^{-\frac{1}{2}z^2} dz > 0. \)

To obtain the lower bound, we integrate \( \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \) by parts where we let \( u = \frac{1}{\sqrt{2\pi}} \), \( du = \frac{z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \) so that \( du = -\frac{3}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \) and \( v = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \) and hence

\[ \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \left[ 1 - \int_{x}^{\infty} \frac{3}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \right] \]

\[ = \frac{1}{x^3 \sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \int_{x}^{\infty} \frac{3}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \]
Therefore,
\[
\int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz = \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{x^3\sqrt{2\pi}} e^{-\frac{1}{2}x^2} + \int_{x}^{\infty} \frac{3}{z^4\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz
\geq \left( \frac{1}{x} - \frac{1}{x^3} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}
\]

since \( \int_{x}^{\infty} \frac{3}{z^4\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz > 0 \). By taking into account both the lower and upper bounds we have
\[
\left( \frac{1}{x} - \frac{1}{x^3} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \leq \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz \leq \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.
\]
\[\Box\]

9. Lognormal Distribution I. Let \( Z \sim \mathcal{N}(0, 1) \), show that the moment generating function of a standard normal distribution is
\[
\mathbb{E}(e^{\theta Z}) = e^{\frac{1}{2}\theta^2}
\]
for a constant \( \theta \).
Show that if \( X \sim \mathcal{N}(\mu, \sigma^2) \) then \( Y = e^X \) follows a lognormal distribution, \( Y = e^X \sim \log\mathcal{N}(\mu, \sigma^2) \) with probability density function
\[
f_Y(y) = \frac{1}{y\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log y - \mu}{\sigma} \right)^2}
\]
with mean \( \mathbb{E}(Y) = e^{\mu + \frac{1}{2}\sigma^2} \) and variance \( \text{Var}(Y) = \left( e^{\sigma^2} - 1 \right) e^{2\mu + \sigma^2} \).

Solution: By definition
\[
\mathbb{E}(e^{\theta Z}) = \int_{-\infty}^{\infty} e^{\theta z} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz
= \left[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\theta)^2 + \frac{1}{2}\theta^2} \, dz \right]
= e^{\frac{1}{2}\theta^2} \left[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\theta)^2} \, dz \right]
= e^{\frac{1}{2}\theta^2}.
\]
For \( \gamma > 0 \), by definition
\[
\mathbb{P}(e^X < \gamma) = \mathbb{P}(X < \log \gamma) = \int_{-\infty}^{\log \gamma} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx
\]
and hence
\[ f_Y(y) = \frac{d}{dy} \mathbb{P}(e^X < y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log y - \mu}{\sigma} \right)^2}. \]

Given that \( \log Y \sim \mathcal{N}(\mu, \sigma^2) \) we can write \( \log Y = \mu + \sigma Z, Z \sim \mathcal{N}(0, 1) \) and hence
\[ \mathbb{E}(Y) = \mathbb{E}(e^{\mu + \sigma Z}) = e^\mu \mathbb{E}(e^{\sigma Z}) = e^{\mu + \frac{1}{2}\sigma^2} \]
since \( \mathbb{E}(e^{\theta Z}) = e^{\frac{1}{2}\theta^2} \) is the moment generating function of a standard normal distribution.

Taking second moments,
\[ \mathbb{E}(Y^2) = \mathbb{E}(e^{2\mu + 2\sigma Z}) = e^{2\mu} \mathbb{E}(e^{2\sigma Z}) = e^{2\mu + 2\sigma^2}. \]

Therefore,
\[ \text{Var}(Y) = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2 = e^{2\mu + 2\sigma^2} - \left( e^{\mu + \frac{1}{2}\sigma^2} \right)^2 = \left( e^{\sigma^2} - 1 \right) e^{2\mu + \sigma^2}. \]

\[ \square \]

10. Lognormal Distribution II. Let \( X \sim \log-N(\mu, \sigma^2) \), show that for \( n \in \mathbb{N} \)
\[ \mathbb{E}(X^n) = e^{\mu + \frac{1}{2}n^2\sigma^2}. \]

**Solution:** Given \( X \sim \log-N(\mu, \sigma^2) \) the density function is
\[ f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log x - \mu}{\sigma} \right)^2}, \quad x > 0 \]
and for \( n \in \mathbb{N} \),
\[ \mathbb{E}(X^n) = \int_0^\infty x^n f_X(x) \, dx \]
\[ = \int_0^\infty \frac{1}{x\sigma\sqrt{2\pi}} x^n e^{-\frac{1}{2} \left( \frac{\log x - \mu}{\sigma} \right)^2} \, dx. \]

By substituting \( z = \frac{\log x - \mu}{\sigma} \) so that \( x = e^{\sigma z + \mu} \) and \( \frac{dx}{dz} = \frac{1}{x\sigma} \),
\[ \mathbb{E}(X^n) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{(\sigma z + \mu)} e^{-\frac{1}{2} z^2} \frac{1}{\sigma} \, dz \]
\[ = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2 + n(\sigma z + \mu)} \, dz \]
\[ = e^{\mu + \frac{1}{2}n^2\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} (z-n\mu)^2} \, dz. \]
1.2.2 Discrete and Continuous Random Variables

\[ e^{\mu} = e^{\mu + 1/2 n^2} \]

since \[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-1/2 (z-\mu)^2} \, dz = 1. \]

\[ \square \]

11. Folded Normal Distribution. Show that if \( X \sim \mathcal{N}(\mu, \sigma^2) \) then \( Y = |X| \) follows a folded normal distribution, \( Y = |X| \sim \mathcal{N}_f(\mu, \sigma^2) \) with probability density function

\[ f_Y(y) = \frac{\sqrt{2}}{\pi} \frac{e^{-1/2 \left( \frac{y^2 + \mu^2}{\sigma^2} \right)}}{\cosh \left( \frac{\mu y}{\sigma^2} \right)} \]

with mean

\[ \mathbb{E}(Y) = \sigma \sqrt{\frac{2}{\pi}} \frac{e^{-1/2 \left( \frac{y^2}{\sigma^2} \right)} + \mu \left[ 1 - 2 \Phi \left( -\frac{\mu}{\sigma} \right) \right]}{\cosh \left( \frac{\mu y}{\sigma^2} \right)} \]

and variance

\[ \text{Var}(Y) = \mu^2 + \sigma^2 - \left\{ \sigma \sqrt{\frac{2}{\pi}} \frac{e^{-1/2 \left( \frac{y^2}{\sigma^2} \right)} + \mu \left[ 1 - 2 \Phi \left( -\frac{\mu}{\sigma} \right) \right]}{\cosh \left( \frac{\mu y}{\sigma^2} \right)} \right\}^2 \]

where \( \Phi(\cdot) \) is the cumulative distribution function of a standard normal.

\textbf{Solution:} For \( y > 0 \), by definition

\[ P(|X| < y) = P(-y < X < y) = \int_{-y}^{y} \frac{1}{\sqrt{2\pi}} e^{-1/2 \left( \frac{x^2}{\sigma^2} \right)} \, dx \]

and hence

\[ f_Y(y) = \frac{d}{dy} P(|X| < y) = \frac{1}{\sigma \sqrt{2\pi}} \left[ e^{-1/2 \left( \frac{y^2 + \mu^2}{\sigma^2} \right)} + e^{-1/2 \left( \frac{y^2}{\sigma^2} \right)} \right] \]

\[ = \frac{2}{\pi \sigma^2} e^{-1/2 \left( \frac{\mu^2}{\sigma^2} \right)} \cosh \left( \frac{\mu y}{\sigma^2} \right). \]

By definition

\[ \mathbb{E}(Y) = \int_{-\infty}^{\infty} y f_Y(y) \, dy \]

\[ = \int_{0}^{\infty} \frac{y}{\sigma \sqrt{2\pi}} e^{-1/2 \left( \frac{y^2 + \mu^2}{\sigma^2} \right)} \, dy + \int_{0}^{\infty} \frac{y}{\sigma \sqrt{2\pi}} e^{-1/2 \left( \frac{y^2}{\sigma^2} \right)} \, dy. \]

By setting \( z = (y + \mu)/\sigma \) and \( w = (y - \mu)/\sigma \) we have

\[ \mathbb{E}(Y) = \int_{\mu/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} (\sigma z - \mu) e^{-1/2 \left( \frac{z^2}{\sigma^2} \right)} \, dz + \int_{-\mu/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} (\sigma w + \mu) e^{-1/2 w^2} \, dw \]
1.2.2 Discrete and Continuous Random Variables

\[\begin{align*}
\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2} &= \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2} - \mu \left[1 - \Phi \left(\frac{\mu}{\sigma}\right)\right] + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{-y}{\sigma}\right)^2} - \mu \left[1 - \Phi \left(-\frac{\mu}{\sigma}\right)\right] \\
&= \frac{2\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2} + \mu \left[\Phi \left(\frac{\mu}{\sigma}\right) - \Phi \left(-\frac{\mu}{\sigma}\right)\right] \\
&= \sigma \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2} + \mu \left[1 - 2\Phi \left(-\frac{\mu}{\sigma}\right)\right].
\end{align*}\]

To evaluate \(\mathbb{E}(Y^2)\) by definition,

\[\begin{align*}
\mathbb{E}(Y^2) &= \int_{-\infty}^{\infty} y^2 f_Y(y) \, dy \\
&= \int_{0}^{\infty} \frac{y^2}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2} \, dy + \int_{0}^{\infty} \frac{y^2}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{-y}{\sigma}\right)^2} \, dy.
\end{align*}\]

By setting \(z = (y + \mu)/\sigma\) and \(w = (y - \mu)/\sigma\) we have

\[\begin{align*}
\mathbb{E}(Y^2) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sigma^2 z^2 e^{-\frac{1}{2} z^2} \, dz + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \sigma^2 w^2 e^{-\frac{1}{2} w^2} \, dw \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\mu/\sigma}^{\infty} z^2 e^{-\frac{1}{2} z^2} \, dz - \frac{2\mu \sigma}{\sqrt{2\pi}} \int_{-\mu/\sigma}^{\infty} ze^{-\frac{1}{2} z^2} \, dz + \mu^2 \int_{-\mu/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} \, dz \\
&\quad + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\mu/\sigma}^{\infty} w^2 e^{-\frac{1}{2} w^2} \, dw + \frac{2\mu \sigma}{\sqrt{2\pi}} \int_{-\mu/\sigma}^{\infty} we^{-\frac{1}{2} w^2} \, dw \\
&\quad + \mu^2 \int_{-\mu/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} w^2} \, dw \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left[ \left(\frac{\mu}{\sigma}\right) e^{-\frac{1}{2} \left(\frac{\mu}{\sigma}\right)^2} + \sqrt{2\pi} \left(1 - \Phi \left(\frac{\mu}{\sigma}\right)\right)\right] - \frac{2\mu \sigma}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\mu}{\sigma}\right)^2} \\
&\quad + \mu^2 \left[1 - \Phi \left(\frac{\mu}{\sigma}\right)\right] \\
&\quad + \frac{\sigma^2}{\sqrt{2\pi}} \left[ \left(-\frac{\mu}{\sigma}\right) e^{-\frac{1}{2} \left(\frac{-\mu}{\sigma}\right)^2} + \sqrt{2\pi} \left(1 - \Phi \left(-\frac{\mu}{\sigma}\right)\right)\right] + \frac{2\mu \sigma}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{-\mu}{\sigma}\right)^2} \\
&\quad + \mu^2 \left[1 - \Phi \left(-\frac{\mu}{\sigma}\right)\right] \\
&= (\mu^2 + \sigma^2) \left[2 - \Phi \left(\frac{\mu}{\sigma}\right) - \Phi \left(-\frac{\mu}{\sigma}\right)\right] \\
&= \mu^2 + \sigma^2.
\end{align*}\]
Therefore,

\[
\text{Var}(Y) = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2 = \mu^2 + \sigma^2 - \left\{ \sigma \sqrt{\frac{2}{\pi}} \ e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} + \mu \left[ 1 - 2\Phi\left(-\frac{\mu}{\sigma}\right) \right] \right\}^2.
\]

\[\square\]

12. **Chi-Square Distribution.** Show that if \( X \sim \mathcal{N}(\mu, \sigma^2) \) then \( Y = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1) \).

Let \( Z_1, Z_2, \ldots, Z_n \sim \mathcal{N}(0, 1) \) be a sequence of independent and identically distributed random variables each following a standard normal distribution. Using mathematical induction show that

\[
Z = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \chi^2(n)
\]

where \( Z \) has a probability density function

\[
f_Z(z) = \frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} z^{n/2 - 1} e^{-z/2}, \quad z \geq 0
\]

such that

\[
\Gamma\left(\frac{n}{2}\right) = \int_0^\infty e^{-x} x^{\frac{n}{2} - 1} dx.
\]

Finally, show that \( \mathbb{E}(Z) = n \) and \( \text{Var}(Z) = 2n \).

**Solution:** By setting \( Y = \frac{X - \mu}{\sigma} \) then

\[
\mathbb{P}(Y \leq y) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq y\right) = \mathbb{P}(X \leq \mu + \sigma y).
\]

Differentiating with respect to \( y \),

\[
f_Y(y) = \frac{d}{dy} \mathbb{P}(Y \leq y)
\]

\[
= \frac{d}{dy} \int_{-\infty}^{\mu + \sigma y} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2} dx
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} \sigma
\]

\[
= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}
\]

which is a probability density function of \( \mathcal{N}(0, 1) \).

For \( Z = Z_1^2 \) and given \( Z \geq 0 \), by definition

\[
\mathbb{P}(Z \leq z) = \mathbb{P}(Z_1^2 \leq z) = \mathbb{P}\left(-\sqrt{z} < Z_1 < \sqrt{z}\right) = 2 \int_0^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.
\]
1.2.2 Discrete and Continuous Random Variables

and hence

\[ f_Z(z) = \frac{d}{dz} \left[ 2 \int_0^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \, dz_1 \right] = \frac{1}{\sqrt{2\pi}} z^{-\frac{1}{2}} e^{-\frac{1}{2}z}, \quad z \geq 0 \]

which is the probability density function of \( \chi^2(1) \).

For \( Z = Z_1^2 + Z_2^2 \) such that \( Z \geq 0 \) we have

\[ P(Z \leq z) = P(Z_1^2 + Z_2^2 \leq z) = \int \int_{z_1^2 + z_2^2 \leq z} \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)} \, d(z_1, z_2). \]

Changing to polar coordinates \((r, \theta)\) such that \( z_1 = r \cos \theta \) and \( z_2 = r \sin \theta \) with the Jacobian determinant

\[ |J| = \left| \begin{array}{cc} \frac{\partial (z_1, z_2)}{\partial (r, \theta)} & \frac{\partial (z_1)}{\partial r} \frac{\partial z_1}{\partial \theta} \\ \frac{\partial (z_2)}{\partial r} \frac{\partial z_2}{\partial \theta} & \end{array} \right| = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \]

then

\[ \int \int_{z_1^2 + z_2^2 \leq z} \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)} \, d(z_1, z_2) = \int_0^{2\pi} \int_0^{\sqrt{z}} \frac{1}{2\pi} e^{-\frac{1}{2}r^2} \, r \, dr \, d\theta = 1 - e^{-\frac{1}{2}z}. \]

Thus,

\[ f_Z(z) = \frac{d}{dz} \left[ \int \int_{z_1^2 + z_2^2 \leq z} \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)} \, d(z_1, z_2) \right] \]

\[ = \frac{1}{2} e^{-\frac{1}{2}z}, \quad z \geq 0 \]

which is the probability density function of \( \chi^2(2) \).

Assume the result is true for \( n = k \) such that

\[ U = Z_1^2 + Z_2^2 + \ldots + Z_k^2 \sim \chi^2(k) \]

and knowing that

\[ V = Z_{k+1}^2 \sim \chi^2(1) \]

then, because \( U \geq 0 \) and \( V \geq 0 \) are independent, using the convolution formula the density of \( Z = U + V = \sum_{i=1}^{k+1} Z_i^2 \) can be written as

\[ f_Z(z) = \int_0^z f_U(z-u) f_V(u) \, du \]

\[ = \int_0^z \left\{ \frac{1}{\sqrt{2\pi}} (z-u)^{-\frac{1}{2}} e^{-\frac{1}{2}(z-u)} \right\} \cdot \left\{ \frac{1}{2^\frac{1}{2} \Gamma\left(\frac{1}{2}\right)} u^{\frac{1}{2}-1} e^{-\frac{1}{2}u} \right\} \, du \]
\[ f_Z(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi z^2}} \frac{1}{\Gamma \left( \frac{k+1}{2} \right)} \int_0^z (z-u)^{-\frac{1}{2}} u^{k-1} \, du. \]

By setting \( v = \frac{u}{z} \) we have
\[
\int_0^z (z-u)^{-\frac{1}{2}} u^{k-1} \, du = \int_0^1 (z-vz)^{-\frac{1}{2}} (vz)^{\frac{k+1}{2}} \, dv
\]
\[
= z^{\frac{k+1}{2}} \int_0^1 (1-v)^{-\frac{1}{2}} v^{\frac{k-1}{2}} \, dv
\]
and because
\[
\int_0^1 (1-v)^{-\frac{1}{2}} v^{\frac{k-1}{2}} \, dv = B \left( \frac{1}{2}, \frac{k}{2} \right) = \frac{\sqrt{\pi} \Gamma \left( \frac{k}{2} \right)}{\Gamma \left( \frac{k+1}{2} \right)} \text{ (see Appendix A)}
\]
therefore
\[
f_Z(z) = \frac{1}{2 \sqrt{\pi} \Gamma \left( \frac{k+1}{2} \right)} z^{\frac{k+1}{2}-1} e^{-\frac{1}{2}z^2}, \quad z \geq 0
\]
which is the probability density function of \( \chi^2(k+1) \) and hence the result is also true for \( n = k + 1 \). By mathematical induction we have shown
\[
Z = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \chi^2(n).
\]

By computing the moment generation of \( Z \),
\[
M_Z(t) = \mathbb{E} \left( e^{tZ} \right) = \int_0^\infty e^{tz} f_Z(z) \, dz = \frac{1}{2 \sqrt{\pi} \Gamma \left( \frac{n}{2} \right)} \int_0^\infty z^{\frac{n}{2}-1} e^{-\frac{1}{2}(1-2t)z} \, dz
\]
and by setting \( w = \frac{1}{2} (1-2t)z \) we have
\[
M_Z(t) = \frac{1}{2 \sqrt{\pi} \Gamma \left( \frac{n}{2} \right)} \left( \frac{2}{1-2t} \right)^{\frac{n}{2}} \int_0^\infty w^{\frac{n}{2}-1} e^{-w} \, dw = (1-2t)^{-\frac{n}{2}}, \quad t \in \left( -\frac{1}{2}, \frac{1}{2} \right)
\]
Thus,
\[
M_Z'(t) = n(1-2t)^{-\frac{n+1}{2}}, \quad M_Z''(t) = 2n \left( \frac{n}{2} + 1 \right) (1-2t)^{-\frac{n+2}{2}}
\]
such that
\[
\mathbb{E}(Z) = M_Z'(0) = n, \quad \mathbb{E}(Z^2) = M_Z''(0) = 2n \left( \frac{n}{2} + 1 \right)
\]
and
\[
\text{Var}(Z) = \mathbb{E}(Z^2) - [\mathbb{E}(Z)]^2 = 2n.
\]
1.2.2 Discrete and Continuous Random Variables

13. Marginal Distributions of Bivariate Normal Distribution. Let $X$ and $Y$ be jointly normally distributed with means $\mu_x$, $\mu_y$, variances $\sigma_x^2$, $\sigma_y^2$ and correlation coefficient $\rho_{xy} \in (-1, 1)$ such that the joint density function is

$$f_{XY}(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y - \mu_y}{\sigma_y} \right) \left( \frac{x - \mu_x}{\sigma_x} \right) + \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right]}.$$ 

Show that $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$.

Solution: By definition,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y - \mu_y}{\sigma_y} \right) \left( \frac{x - \mu_x}{\sigma_x} \right) + \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right]} \, dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_x \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y - \mu_y}{\sigma_y} \right) \left( \frac{x - \mu_x}{\sigma_x} \right) + \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right]} \, dy$$

$$= \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2} \int_{-\infty}^{\infty} g(x, y) \, dy$$

where

$$g(x, y) = \frac{1}{\sqrt{1 - \rho_{xy}^2} \sigma_x \sqrt{2\pi}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{y - \mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y - \mu_y}{\sigma_y} \right) \left( \frac{x - \mu_x}{\sigma_x} \right) + \left( \frac{x - \mu_x}{\sigma_x} \right)^2 \right]}$$

is the probability density function for $\mathcal{N}(\mu_y + \rho_{xy} \sigma_y (x - \mu_x) / \sigma_x, (1 - \rho_{xy}^2)\sigma_y^2)$. Therefore,

$$\int_{-\infty}^{\infty} g(x, y) \, dy = 1$$

and hence

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2}.$$ 

Thus, $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$. Using the same steps we can also show that $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$. 

\[
\square
\]
14. **Covariance of Bivariate Normal Distribution.** Let $X$ and $Y$ be jointly normally distributed with means $\mu_x, \mu_y$, variances $\sigma_x^2, \sigma_y^2$ and correlation coefficient $\rho_{xy} \in (-1, 1)$ such that the joint density function is

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{x-\mu_x}{\sigma_x} \right)^2 - 2\rho_{xy} \left( \frac{x-\mu_x}{\sigma_x} \right) \left( \frac{y-\mu_y}{\sigma_y} \right) + \left( \frac{y-\mu_y}{\sigma_y} \right)^2 \right]}.$$ 

Show that the covariance of $X$ and $Y$, $\text{Cov}(X, Y) = \rho_{xy} \sigma_x \sigma_y$, and hence show that $X$ and $Y$ are independent if and only if $\rho_{xy} = 0$.

**Solution:** By definition, the covariance of $X$ and $Y$ is

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_x)(Y - \mu_y)]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f_{XY}(x, y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy - \mu_x \int_{-\infty}^{\infty} x f_{XY}(x, y) \, dx \, dy$$

$$- \mu_y \int_{-\infty}^{\infty} y f_{XY}(x, y) \, dx \, dy + \mu_x \mu_y \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy - \mu_x \int_{-\infty}^{\infty} x f_{XY}(x, y) \, dx \, dy$$

$$- \mu_y \int_{-\infty}^{\infty} y f_{XY}(x, y) \, dx \, dy + \mu_x \mu_y \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy - \mu_x \mu_y$$

where $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx$. Using the result of Problem 1.2.13 (page 27) we can deduce that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy = \int_{-\infty}^{\infty} \frac{x}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu_x}{\sigma_x} \right)^2} \left( \int_{-\infty}^{\infty} y g(x, y) \, dy \right) \, dx$$

where

$$g(x, y) = \frac{1}{\sqrt{1 - \rho_{xy}^2} \sqrt{2\pi}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left( y - \rho_{xy} \sigma_y \left( \frac{\mu_y}{\sigma_y} \right) \right)^2}$$

and

$$\int_{-\infty}^{\infty} g(x, y) \, dx = \int_{-\infty}^{\infty} \frac{x}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu_x}{\sigma_x} \right)^2} \left( \int_{-\infty}^{\infty} y g(x, y) \, dy \right) \, dx.$$
1.2.2 Discrete and Continuous Random Variables

is the probability density function for \( N\left(\mu_x + \rho_{xy} \sigma_y (x - \mu_x) / \sigma_x, (1 - \rho_{xy}^2) \sigma_y^2\right) \).

Therefore,
\[
\int_{-\infty}^{\infty} yg(x, y) \, dy = \mu_y + \rho_{xy} \sigma_y \left( \frac{x - \mu_x}{\sigma_x} \right).
\]

Thus,
\[
\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \frac{X}{\sigma_Y} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \left[ \mu_y + \rho_{xy} \sigma_y \left( \frac{x - \mu_x}{\sigma_x} \right) \right] \, dx - \mu_x \mu_y
\]
\[
= \mu_x \mu_y + \rho_{xy} \sigma_y \int_{-\infty}^{\infty} \frac{X^2}{\sigma_X} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \, dx - \frac{\rho_{xy} \sigma_y \mu_x^2}{\sigma_x} - \mu_x \mu_y
\]
\[
= \rho_{xy} \sigma_x \sigma_y
\]
where
\[
\mathbb{E}(X^2) = \int_{-\infty}^{\infty} \frac{X^2}{\sigma_X} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \, dx = \sigma_x^2 + \mu_x^2.
\]

To show that \( X \) and \( Y \) are independent if and only if \( \rho_{xy} = 0 \) we note that if \( X \perp Y \) then \( \text{Cov}(X, Y) = 0 \), which implies \( \rho_{xy} = 0 \). On the contrary, if \( \rho_{xy} = 0 \) then from the joint density of \( (X, Y) \) we can express it as
\[
f_{XY}(x, y) = f_X(x) f_Y(y)
\]
where \( f_X(x) = \int_{-\infty}^{\infty} \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu_x}{\sigma_x} \right)^2} \, dx \) and \( f_Y(y) = \int_{-\infty}^{\infty} \frac{1}{\sigma_Y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \, dy \) and so
\( X \perp Y \).

Thus, if the pair \( X \) and \( Y \) has a bivariate normal distribution with means \( \mu_x, \mu_y \), variances \( \sigma_x^2, \sigma_y^2 \) and correlation \( \rho_{xy} \) then \( X \perp Y \) if and only if \( \rho_{xy} = 0 \).

\[\square\]

15. Minimum and Maximum of Two Correlated Normal Distributions. Let \( X \) and \( Y \) be jointly normally distributed with means \( \mu_x, \mu_y \), variances \( \sigma_x^2, \sigma_y^2 \) and correlation \( \rho_{xy} \in (-1, 1) \) such that the joint density function is
\[
f_{XY}(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \frac{(x - \mu_x)}{\sigma_x} \right]^2 - 2\rho_{xy} \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) - \frac{(y - \mu_y)}{\sigma_y}^2].
\]

Show that the distribution of \( U = \min\{X, Y\} \) is
\[
f_{U}(u) = \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_y (u - \mu_x)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \right) f_X(u) + \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_y (u - \mu_x)}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) f_Y(u).
\]
and deduce that the distribution of \( V = \max \{ X, Y \} \) is

\[
f_X(v) = \Phi \left( \frac{v - \mu_y - \rho_{xy} \sigma_y (v - \mu_x)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \right) f_X(v) + \Phi \left( \frac{v - \mu_x - \rho_{xy} \sigma_x (u - \mu_x)}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) f_Y(v)
\]

where \( f_X(z) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{z - \mu_x}{\sigma_x} \right)^2} \), \( f_Y(z) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{z - \mu_y}{\sigma_y} \right)^2} \) and \( \Phi(\cdot) \) denotes the cumulative standard normal distribution function (cdf).

**Solution:** From Problems 1.2.2.13 (page 27) and 1.2.2.14 (page 28) we can show that \( X \sim \mathcal{N}(\mu_x, \sigma_x^2), Y \sim \mathcal{N}(\mu_y, \sigma_y^2), \text{Cov}(X, Y) = \rho_{xy} \sigma_x \sigma_y \) such that \( \rho_{xy} \in (-1, 1) \).

For \( U = \min \{ X, Y \} \) then by definition the cumulative distribution function (cdf) of \( U \) is

\[
\mathbb{P}(U \leq u) = \mathbb{P}(\min \{ X, Y \} \leq u) = 1 - \mathbb{P}(\min \{ X, Y \} > u) = 1 - \mathbb{P}(X > u, Y > u).
\]

To derive the probability density function (pdf) of \( U \) we have

\[
f_U(u) = \frac{d}{du} \mathbb{P}(U \leq u) = -\frac{d}{du} \mathbb{P}(X > u, Y > u)
\]

\[
= -\frac{d}{du} \int_{x=u}^{\infty} \int_{y=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}}
\]

\[
\times e^{-\frac{1}{2(1-\rho_{xy})^2} \left[ \frac{(x-\mu_x)}{\sigma_x} + \frac{(y-\mu_y)}{\sigma_y} \right]^2 - 2 \rho_{xy} \left( \frac{(x-\mu_x)}{\sigma_x} \right) \left( \frac{(y-\mu_y)}{\sigma_y} \right) + \left( \frac{(x-\mu_x)}{\sigma_x} \right)^2 + \left( \frac{(y-\mu_y)}{\sigma_y} \right)^2 } \) \, dy \, dx
\]

\[
= g(u) + h(u)
\]

where

\[
g(u) = \int_{y=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy})^2} \left[ \frac{(y-\mu_y)}{\sigma_y} \right]^2 - 2 \rho_{xy} \left( \frac{(y-\mu_y)}{\sigma_y} \right) + \left( \frac{(y-\mu_y)}{\sigma_y} \right)^2 } \, dy
\]

and

\[
h(u) = \int_{x=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy})^2} \left[ \frac{(x-\mu_x)}{\sigma_x} \right]^2 - 2 \rho_{xy} \left( \frac{(x-\mu_x)}{\sigma_x} \right) + \left( \frac{(x-\mu_x)}{\sigma_x} \right)^2 } \, dx.
\]
1.2.2 Discrete and Continuous Random Variables

By focusing on

\[
g(u) = \int_{y=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{u - \mu_x}{\sigma_x} \right)^2 - 2\rho_{xy} \left( \frac{u - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \, dy
\]

\[
= \int_{y=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \left( \frac{\sigma_y}{\sigma_x} \right)^2 \left( \frac{\sigma_y}{\sigma_x} \right)^2 \left( \frac{u - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) \, dy
\]

\[
= \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{u - \mu_x}{\sigma_x} \right)^2} \int_{y=u}^{\infty} \frac{1}{\sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \, dy
\]

and letting \( w = \frac{y - \mu_y}{\sigma_y} - \rho_{xy} \left( \frac{u - \mu_x}{\sigma_x} \right) \) we have

\[
g(u) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{u - \mu_x}{\sigma_x} \right)^2} \frac{1}{\sqrt{1 - \rho_{xy}^2}} \int_{w=u}^{\infty} e^{-\frac{1}{2} w^2} \, dw
\]

\[
= \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_x (u - \mu_x)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \right) f_X(u).
\]

In a similar vein we can also show

\[
h(u) = \int_{x=u}^{\infty} \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{u - \mu_x}{\sigma_x} \right)^2 - 2\rho_{xy} \left( \frac{u - \mu_x}{\sigma_x} \right) \left( \frac{x - \mu_y}{\sigma_y} \right) + \left( \frac{x - \mu_y}{\sigma_y} \right)^2} \, dx
\]

\[
= \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{u - \mu_x}{\sigma_x} \right)^2} \int_{x=u}^{\infty} \frac{1}{\sigma_x \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{x - \mu_y}{\sigma_y} \right)^2} \, dx
\]

\[
= \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_x (u - \mu_x)}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) f_Y(u).
\]

Therefore,

\[
f_{1,\rho}(u) = \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_x (u - \mu_x)}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) f_X(u) + \Phi \left( \frac{-u + \mu_x + \rho_{xy} \sigma_x (u - \mu_x)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \right) f_Y(u).
\]
As for the case \( V = \max\{X, Y\} \), then by definition the cdf of \( V \) is
\[
P(V \leq v) = P(\max\{X, Y\} \leq v) = P(X \leq v, Y \leq v).
\]
The pdf of \( V \) is
\[
f_V(v) = \frac{d}{dv} P(V \leq v) = \frac{d}{dv} P(X \leq v, Y \leq v)
\]
\[
= \frac{d}{dv} \int_{-\infty}^{x=v} \int_{-\infty}^{y=v} \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} e^{-\frac{1}{2\rho_{xy}^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho_{xy}\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right] + \left(\frac{y-\mu_y}{\sigma_y}\right)^2} dy dx
\]
\[
= \int_{-\infty}^{x=v} \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} e^{-\frac{1}{2\rho_{xy}^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho_{xy}\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right] + \left(\frac{y-\mu_y}{\sigma_y}\right)^2} dy
\]
\[
+ \int_{-\infty}^{x=v} \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} e^{-\frac{1}{2\rho_{xy}^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho_{xy}\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right] + \left(\frac{y-\mu_y}{\sigma_y}\right)^2} dx.
\]
Following the same steps as described above we can write
\[
f_V(v) = \Phi\left(\frac{v - \mu_x - \frac{\rho_{xy}\sigma_x}{\sigma_y}(v - \mu_x)}{\sigma_y\sqrt{1-\rho_{xy}^2}}\right) f_X(v) + \Phi\left(\frac{v - \mu_x - \frac{\rho_{xy}\sigma_x}{\sigma_y}(v - \mu_x)}{\sigma_x\sqrt{1-\rho_{xy}^2}}\right) f_Y(v).
\]
\[\square\]

16. **Bivariate Standard Normal Distribution.** Let \( X \sim \mathcal{N}(0, 1) \) and \( Y \sim \mathcal{N}(0, 1) \) be jointly normally distributed with correlation coefficient \( \rho_{xy} \in (-1, 1) \) where the joint cumulative distribution function is
\[
\Phi(\alpha, \beta, \rho_{xy}) = \int_{-\infty}^{\alpha} \int_{-\infty}^{\beta} \frac{1}{2\pi\sqrt{1-\rho_{xy}^2}} e^{-\frac{1}{2}\left(\frac{\xi^2 - 2\rho_{xy}\xi\eta + \eta^2}{1-\rho_{xy}^2}\right)} d\xi d\eta.
\]
By using the change of variables \( Y = \rho_{xy}X + \sqrt{1-\rho_{xy}^2}Z, \ Z \sim \mathcal{N}(0, 1), \ X \perp Z \) show that
\[
\Phi(\alpha, \beta, \rho_{xy}) = \int_{-\infty}^{\beta} f_X(x) \Phi\left(\frac{\beta - \rho_{xy}x}{\sqrt{1-\rho_{xy}^2}}\right) dx
\]
where \( f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \) and \( \Phi(\cdot) \) is the cumulative distribution function of a standard normal.

Finally, deduce that \( \Phi(\alpha, \beta, \rho_{xy}) + \Phi(\alpha, -\beta, -\rho_{xy}) = \Phi(\alpha) \).

**Solution:** Let \( y = \rho_{xy}x + \sqrt{1 - \rho_{xy}^2}z \). Differentiating \( y \) with respect to \( z \) we have \( \frac{dy}{dz} = \sqrt{1 - \rho_{xy}^2} \), and hence

\[
\Phi(\alpha, \beta, \rho_{xy}) = \int_{-\infty}^{\beta} \int_{-\infty}^{a} \frac{1}{2\pi \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{x^2 - 2\rho_{xy}xy + y^2}{1 - \rho_{xy}^2} \right)} \, dx \, dy
\]

\[
= \int_{-\infty}^{\beta} \int_{-\infty}^{a} \frac{1}{2\pi \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2} \left( \frac{x^2 - 2\rho_{xy}xy + y^2}{1 - \rho_{xy}^2} \right)} \times e^{-\frac{1}{2} \left( \frac{y^2 - 2\rho_{xy}xy + 1 - \rho_{xy}^2}{1 - \rho_{xy}^2} \right)} \sqrt{1 - \rho_{xy}^2} \, dx \, dz
\]

\[
= \int_{-\infty}^{\beta} \int_{-\infty}^{a} \frac{1}{2\pi} e^{-\frac{1}{2} (x^2 + z^2)} \, dx \, dz
\]

\[
= \int_{-\infty}^{\beta} \Phi \left( \beta - \rho_{xy}x \sqrt{1 - \rho_{xy}^2} \right) \, dx
\]

Finally,

\[
\Phi(\alpha, \beta, \rho_{xy}) + \Phi(\alpha, -\beta, -\rho_{xy}) = \int_{-\infty}^{\alpha} f_X(x) \Phi \left( \frac{\beta - \rho_{xy}x}{\sqrt{1 - \rho_{xy}^2}} \right) \, dx
\]

\[
+ \int_{-\infty}^{\alpha} f_X(x) \Phi \left( \frac{-\beta + \rho_{xy}x}{\sqrt{1 - \rho_{xy}^2}} \right) \, dx
\]

\[
= \int_{-\infty}^{\alpha} f_X(x) \left[ \Phi \left( \frac{\beta - \rho_{xy}x}{\sqrt{1 - \rho_{xy}^2}} \right) + \Phi \left( \frac{-\beta + \rho_{xy}x}{\sqrt{1 - \rho_{xy}^2}} \right) \right] \, dx
\]
34 1.2.2 Discrete and Continuous Random Variables

\[
= \int_{-\infty}^{\alpha} f_X(x) \, dx \\
= \Phi(\alpha)
\]

since \( \Phi \left( \frac{\beta - \rho_{xy} x}{\sqrt{1 - \rho_{xy}^2}} \right) + \Phi \left( \frac{-\beta + \rho_{xy} x}{\sqrt{1 - \rho_{xy}^2}} \right) = 1. \)

N.B. Similarly we can also show that

\[
\Phi(\alpha, \beta, \rho_{xy}) = \int_{-\infty}^{\beta} f_Y(y) \Phi \left( \frac{\alpha - \rho_{xy} y}{\sqrt{1 - \rho_{xy}^2}} \right) \, dy
\]

where \( f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y^2} \) and \( \Phi(\alpha, \beta, \rho_{xy}) + \Phi(-\alpha, \beta, -\rho_{xy}) = \Phi(\beta). \)

\[
\square
\]

17. Bivariate Normal Distribution Property. Let \( X \) and \( Y \) be jointly normally distributed with means \( \mu_x, \mu_y \), variances \( \sigma_x^2, \sigma_y^2 \) and correlation coefficient \( \rho_{xy} \in (-1, 1) \) such that the joint density function is

\[
f_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho_{xy} \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right]}
\]

Show that

\[
\mathbb{E} \left[ \max\{e^X - e^Y, 0\} \right] = e^{\mu_x + \frac{1}{2}\sigma_x^2} \Phi \left( \frac{\mu_x - \mu_y + \sigma_x(\sigma_x - \rho_{xy}\sigma_y)}{\sqrt{\sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2}} \right) \\
- e^{\mu_y + \frac{1}{2}\sigma_y^2} \Phi \left( \frac{\mu_x - \mu_y - \sigma_x(\sigma_x - \rho_{xy}\sigma_y)}{\sqrt{\sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2}} \right)
\]

Solution: By definition

\[
\mathbb{E} \left[ \max\{e^X - e^Y, 0\} \right] = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{x=\infty} (e^x - e^y) f_{XY}(x, y) \, dx \, dy \\
= I_1 - I_2
\]

where \( I_1 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} e^x f_{XY}(x, y) \, dx \, dy \) and \( I_2 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} e^y f_{XY}(x, y) \, dx \, dy. \)
1.2.2 Discrete and Continuous Random Variables

For the case \( I_1 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} e^{xf_{XY}(x,y)} \, dx \, dy \) we have

\[
I_1 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} \frac{e^{x}}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho_{xy}^2}} \left[ \left( \frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y-\mu_y}{\sigma_y} \right) \left( \frac{x-\mu_x}{\sigma_x} \right) + \left( \frac{x-\mu_x}{\sigma_x} \right)^2 \right] \, dx \, dy
\]

\[
= \int_{y=-\infty}^{y=\infty} \frac{1}{\sigma_y \sqrt{2\pi}} \, e^{-\frac{1}{2} \left( \frac{y-\mu_y}{\sigma_y} \right)^2} \left[ \int_{x=y}^{x=\infty} e^{x} \, dx \right] \, dy
\]

where \( g(x,y) = \frac{1}{\sigma_x \sqrt{2\pi}} \, e^{-\frac{1}{2} \left( \frac{y-\mu_y}{\sigma_y} \right)^2} \) which is the probability density function of \( \mathcal{N} \left[ \mu_x + \rho_{xy} \sigma_x \left( \frac{y-\mu_y}{\sigma_y} \right), (1-\rho_{xy})^2 \sigma_x^2 \right] \). Thus, from Problem 1.2.2.7 (page 18) we can deduce

\[
\int_{x=y}^{x=\infty} e^{x} g(x,y) \, dx = e^{\mu_x + \rho_{xy} \sigma_x \left( \frac{y-\mu_y}{\sigma_y} \right) + \frac{1}{2} (1-\rho_{xy})^2 \sigma_x^2} \times \Phi \left( \frac{\mu_x + \rho_{xy} \sigma_x \left( \frac{y-\mu_y}{\sigma_y} \right) + (1-\rho_{xy})^2 \sigma_x^2 - y}{\sigma_x \sqrt{1-\rho_{xy}}} \right).
\]

Thus, we can write

\[
I_1 = e^{\mu_x + \frac{1}{2} (1-\rho_{xy})^2 \sigma_x^2} \times \int_{y=-\infty}^{y=\infty} \frac{1}{\sigma_y \sqrt{2\pi}} \, e^{-\frac{1}{2} \left( \frac{y-\mu_y}{\sigma_y} \right)^2} \left[ \left( \frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho_{xy} \left( \frac{y-\mu_y}{\sigma_y} \right) \left( \frac{x-\mu_x}{\sigma_x} \right) + \left( \frac{x-\mu_x}{\sigma_x} \right)^2 \right] \, dx \, dy
\]

\[
\times \Phi \left( \frac{\mu_x + \rho_{xy} \sigma_x \left( \frac{y-\mu_y}{\sigma_y} \right) + (1-\rho_{xy})^2 \sigma_x^2 - y}{\sigma_x \sqrt{1-\rho_{xy}}} \right).
\]
1.2.2 Discrete and Continuous Random Variables

\[
e^{\mu_u + \frac{1}{2} \sigma^2_{u}} \int_{y=-\infty}^{y=\infty} \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y - \mu_y - \rho_{xy} \sigma_x \sigma_y}{\sigma_y} \right)^2} \times \Phi \left( \frac{\mu_x + \rho_{xy} \sigma_x \left( \frac{y - \mu_y}{\sigma_y} \right) + (1 - \rho_{xy})^2 \sigma_x^2 - y}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) dy.
\]

Let \( u = \frac{y - \mu_y - \rho_{xy} \sigma_x \sigma_y}{\sigma_y} \) then from the change of variables

\[
I_1 = e^{\mu_u + \frac{1}{2} \sigma^2_{u}} \int_{u=-\infty}^{u=\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} \Phi \left( \frac{\mu_x - \mu_y + \sigma_x (\sigma_x - \rho_{xy} \sigma_y) - u(\sigma_y - \rho_{xy} \sigma_y)}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) du
\]

\[
e^{\mu_u + \frac{1}{2} \sigma^2_{u}} \int_{u=-\infty}^{u=\infty} \int_{u=-\infty}^{u=\infty} \frac{1}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \frac{1}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \frac{1}{2\pi} e^{-\frac{1}{2} \left( \frac{u - \mu_x - \rho_{xy} \sigma_x \sigma_y}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right)^2} e^{-\frac{1}{2} \left( \frac{w - \mu_y - \sigma_y \sigma_x \sigma_y - u(\sigma_y - \rho_{xy} \sigma_y)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \right)^2} dwdu.
\]

By setting \( w = \frac{u(\sigma_y - \rho_{xy} \sigma_y)}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \) and \( \sigma^2 = \sigma_x^2 - 2\rho_{xy} \sigma_x \sigma_y + \sigma_y^2 \),

\[
I_1 = e^{\mu_u + \frac{1}{2} \sigma^2_{u}} \times \int_{u=-\infty}^{u=\infty} \int_{u=-\infty}^{u=\infty} \frac{1}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \frac{1}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \frac{1}{2\pi} \times e^{-\frac{1}{2} \left( \frac{u^2 - 2uw}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \right) \left( \frac{u^2}{\sigma_x^2 (1 - \rho_{xy}^2)} \right) \left( \frac{w^2}{\sigma_y^2 (1 - \rho_{xy}^2)} \right)} dwdu
\]

\[
e^{\mu_u + \frac{1}{2} \sigma^2_{u}} \times \int_{u=-\infty}^{u=\infty} \int_{u=-\infty}^{u=\infty} \frac{1}{\sigma_x \sqrt{1 - \rho_{xy}^2}} \frac{1}{\sigma_y \sqrt{1 - \rho_{xy}^2}} \frac{1}{2\pi} \times e^{-\frac{1}{2} \left( \frac{u^2}{\sigma_x^2 (1 - \rho_{xy}^2)} \right) \left( \frac{w^2}{\sigma_y^2 (1 - \rho_{xy}^2)} \right) \left( \frac{2uw}{\sigma_x^2 (1 - \rho_{xy}^2)} + \frac{u^2}{\sigma_y^2 (1 - \rho_{xy}^2)} \right)} dwdu.
\]
1.2.2 Discrete and Continuous Random Variables

Finally, by setting $\bar{w} = \begin{pmatrix} w \\ \sigma \\ \sigma \sqrt{1 - \rho_x^2} \end{pmatrix}$,

$$I_1 = e^{\mu_x + \frac{1}{2} \sigma_x^2} \int_{u=-\infty}^{\mu_y + \sigma_y (\sigma_x - \rho_{xy} \sigma_y)} \int_{w=-\infty}^{\mu_x - \rho_{xy} \sigma_x (\sigma_y - \rho_{xy} \sigma_y)} \frac{1}{2\pi \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} (u^2 - 2\rho_{xy} uw + w^2)} \, dw \, du$$

where $\rho_{xy} = \frac{\sigma_y - \rho_{xy} \sigma_x}{\sigma}$.

Therefore,

$$I_1 = e^{\mu_x + \frac{1}{2} \sigma_x^2} \Phi \left( \frac{\mu_x - \mu_y + \sigma_x (\sigma_x - \rho_{xy} \sigma_y)}{\sqrt{\sigma_x^2 - 2\rho_{xy} \sigma_x \sigma_y + \sigma_y^2}}, \infty, \rho_{xy} \right)$$

where $\Phi(a, b, \rho) = \int_{-\infty}^{\beta} \int_{-\infty}^{\alpha} \frac{1}{2\pi \sqrt{1 - \rho^2}} e^{-\frac{1}{2} \left( \frac{u^2}{1 - \rho^2} \right)} \, dx \, dy$ is the cumulative distribution function of a standard bivariate normal.

Using similar steps for the case $I_2 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} e^y f_{XY}(x, y) \, dx \, dy$ we have

$$I_2 = \int_{y=-\infty}^{y=\infty} \int_{x=y}^{x=\infty} e^y \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left( \frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho_{xy} \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right)} \, dx \, dy$$

$$= \int_{y=-\infty}^{y=\infty} \frac{e^y}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y-\mu_y}{\sigma_y} \right)^2} \left[ \int_{x=y}^{x=\infty} \frac{1}{\sigma_x \sqrt{2\pi(1 - \rho_{xy}^2)}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left( \mu_x + \rho_{xy} \sigma_x \left( \frac{y-\mu_y}{\sigma_y} \right) \right)^2} \, dx \right] \, dy$$

$$= \int_{y=-\infty}^{y=\infty} \frac{e^y}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y-\mu_y}{\sigma_y} \right)^2} \left[ \int_{x=y}^{x=\infty} g(x, y) \, dx \right] \, dy$$
38

1.2.2 Discrete and Continuous Random Variables

where \( g(x,y) = \frac{1}{\sigma_x \sqrt{2\pi(1-\rho_{xy}^2)}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \frac{1}{\sigma_y^2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2 + \frac{1}{\sigma_x^2} \left( \frac{x - \mu_x - \rho_{xy} \sigma_y}{\sigma_x} \right)^2 \right]} \) which is the probability density function of \( X \left[ \mu_x + \rho_{xy} \sigma_x \left( \frac{y - \mu_y}{\sigma_y} \right), (1 - \rho_{xy}^2) \sigma_x^2 \right] \).

Thus,

\[
I_2 = \int_{y=-\infty}^{y=\infty} \frac{e^y}{\sigma_y \sqrt{2\pi}} \left[ \int_{x=\infty}^{x=z} \frac{1}{\sigma_x \sqrt{2\pi(1-\rho_{xy}^2)}} e^{-\frac{1}{2} \left( \frac{x - \mu_x - \rho_{xy} \sigma_y}{\sigma_x} \right)^2} dx \right] dy
\]

and by setting \( z = \frac{x - \mu_x - \rho_{xy} \sigma_x}{\sigma_x \sqrt{1-\rho_{xy}^2}} \),

\[
I_2 = \int_{y=-\infty}^{y=\infty} \frac{e^y}{\sigma_y \sqrt{2\pi}} \left[ \int_{z=-\infty}^{z=\infty} \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2} \Phi \left( \frac{\mu_x - \mu_y - \frac{\sigma_y}{\sigma_x} (\sigma_x - \rho_{xy} \sigma_x)}{\sigma_x \sqrt{1-\rho_{xy}^2}} \right) dz \right] dy
\]

By setting \( \tilde{z} = \frac{y - \mu_y}{\sigma_y} \) therefore

\[
I_2 = e^{\mu_y + \frac{1}{2} \sigma_y^2} \int_{\tilde{z}=-\infty}^{\tilde{z}=\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\tilde{z} - \sigma_y)^2} \Phi \left( \frac{\mu_x - \mu_y - \frac{\sigma_y}{\sigma_x} (\sigma_x - \rho_{xy} \sigma_x)}{\sigma_x \sqrt{1-\rho_{xy}^2}} \right) d\tilde{z}
\]

and substituting \( u = \tilde{z} - \sigma_y \),

\[
I_2 = e^{\mu_y + \frac{1}{2} \sigma_y^2} \int_{u=-\infty}^{u=\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} \Phi \left( \frac{\mu_x - \mu_y - (u + \sigma_y)(\sigma_y - \rho_{xy} \sigma_x)}{\sigma_x \sqrt{1-\rho_{xy}^2}} \right) du
\]

\[
= e^{\mu_y + \frac{1}{2} \sigma_y^2} \int_{u=-\infty}^{u=\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} \Phi \left( \frac{\mu_x - \mu_y - \sigma_y \sigma_x - \rho_{xy} \sigma_x - \sigma_y \rho_{xy} \sigma_x - \sigma_y \rho_{xy} \sigma_x}{\sigma_x \sqrt{1-\rho_{xy}^2}} \right) du
\]
1.2.2 Discrete and Continuous Random Variables

Using the same steps as described before, we let \( w = v + \frac{u(\sigma_y - \rho_{xy}\sigma_x)}{\sigma_x\sqrt{1 - \rho_{xy}^2}} \) and \( \sigma^2 = \sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2 \) so that

\[
I_2 = e^{\mu_y + \frac{1}{2} \sigma_y^2} \Phi \left( \frac{\mu_x - \mu_y - \sigma_y(\sigma_y - \rho_{xy}\sigma_x)}{\sqrt{\sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2}} , \infty, \bar{\rho}_{xy} \right)
\]

where \( \bar{\rho}_{xy} = \frac{\sigma_y - \rho_{xy}\sigma_x}{\sigma} \), thus

\[
I_2 = e^{\mu_y + \frac{1}{2} \sigma_y^2} \Phi \left( \frac{\mu_x - \mu_y - \sigma_y(\sigma_y - \rho_{xy}\sigma_x)}{\sqrt{\sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2}} , \infty, \bar{\rho}_{xy} \right)
\]

\[
= e^{\mu_y + \frac{1}{2} \sigma_y^2} \Phi \left( \frac{\mu_x - \mu_y - \sigma_y(\sigma_y - \rho_{xy}\sigma_x)}{\sqrt{\sigma_x^2 - 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2}} \right).
\]
By substituting $I_1$ and $I_2$ back into $\mathbb{E}[\max\{e^X - e^Y, 0\}]$ we have

$$\mathbb{E}[\max\{e^X - e^Y, 0\}] = e^{\mu_x + \frac{1}{2}\sigma_x^2} \Phi \left( \frac{\mu_x - \mu_y + \sigma_y \rho_{xy} \sigma_x}{\sqrt{\sigma_x^2 - 2\rho_{xy} \sigma_x \sigma_y + \sigma_y^2}} \right) - e^{\mu_y + \frac{1}{2}\sigma_y^2} \Phi \left( \frac{\mu_y - \mu_x + \sigma_y \rho_{xy} \sigma_x}{\sqrt{\sigma_x^2 - 2\rho_{xy} \sigma_x \sigma_y + \sigma_y^2}} \right).$$

\[\square\]

18. **Markov’s Inequality.** Let $X$ be a non-negative random variable with mean $\mu$. For $\alpha > 0$, show that

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mu}{\alpha}.$$  

**Solution:** Since $\alpha > 0$, we can write

$$1_{\{X \geq \alpha\}} = \begin{cases} 1 & X \geq \alpha \\ 0 & \text{otherwise} \end{cases}$$

and since $X \geq 0$, we can deduce that

$$1_{\{X \geq \alpha\}} \leq \frac{X}{\alpha}.$$  

Taking expectations

$$\mathbb{E}(1_{\{X \geq \alpha\}}) \leq \frac{\mathbb{E}(X)}{\alpha}$$

and since $\mathbb{E}(1_{\{X \geq \alpha\}}) = 1 \cdot \mathbb{P}(X \geq \alpha) + 0 \cdot \mathbb{P}(X \leq \alpha) = \mathbb{P}(X \geq \alpha)$ and $\mathbb{E}(X) = \mu$, we have

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mu}{\alpha}.$$  

**N.B.** Alternatively, we can also show the result as

$$\mathbb{E}(X) = \int_0^\infty (1 - F_x(u)) \, du \geq \int_0^\alpha (1 - F_x(u)) \, du \geq \alpha (1 - F_x(\alpha))$$

and hence it follows that $\mathbb{P}(X \geq \alpha) = 1 - F_x(\alpha) \leq \frac{\mathbb{E}(X)}{\alpha}$.  

\[\square\]

19. **Chebyshev’s Inequality.** Let $X$ be a random variable with mean $\mu$ and variance $\sigma^2$. Then for $k > 0$, show that

$$\mathbb{P}(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$  

**Solution:** Take note that $|X - \mu| \geq k$ if and only if $(X - \mu)^2 \geq k^2$. Because $(X - \mu)^2 \geq 0$, and by applying Markov’s inequality (see Problem 1.2.2.18, page 40) we have

$$\mathbb{P}((X - \mu)^2 \geq k^2) \leq \frac{\mathbb{E}((X - \mu)^2)}{k^2} = \frac{\sigma^2}{k^2}$$
1.2.3 Properties of Expectations

and hence the above inequality is equivalent to

\[ \mathbb{P}(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}. \]

\[ \square \]

1.2.3 Properties of Expectations

1. Show that if \( X \) is a random variable taking non-negative values then

\[
\mathbb{E}(X) = \begin{cases} 
\sum_{x=0}^{\infty} \mathbb{P}(X > x) & \text{if } X \text{ is a discrete random variable} \\
\int_{0}^{\infty} \mathbb{P}(X \geq x) \, dx & \text{if } X \text{ is a continuous random variable.} 
\end{cases}
\]

Solution: We first show the result when \( X \) takes non-negative integer values only. By definition

\[
\mathbb{E}(X) = \sum_{y=0}^{\infty} y \mathbb{P}(X = y)
\]

\[
= \sum_{y=0}^{\infty} \sum_{x=0}^{y} \mathbb{P}(X = y)
\]

\[
= \sum_{x=0}^{\infty} \sum_{y=x+1}^{\infty} \mathbb{P}(X = y)
\]

\[
= \sum_{x=0}^{\infty} \mathbb{P}(X > x).
\]

For the case when \( X \) is a continuous random variable taking non-negative values we have

\[
\mathbb{E}(X) = \int_{0}^{\infty} y f_{X}(y) \, dy
\]

\[
= \int_{0}^{\infty} \left\{ \int_{0}^{y} f_{X}(y) \, dx \right\} dy
\]

\[
= \int_{0}^{\infty} \left\{ \int_{x}^{\infty} f_{X}(y) \, dy \right\} dx
\]

\[
= \int_{0}^{\infty} \mathbb{P}(X \geq x) \, dx.
\]

\[ \square \]

2. Hölder’s Inequality. Let \( \alpha, \beta \geq 0 \) and for \( p, q > 1 \) such that \( \frac{1}{p} + \frac{1}{q} = 1 \) show that the following inequality:

\[ \alpha \beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q} \]
holds. Finally, if $X$ and $Y$ are a pair of jointly continuous variables, show that

\[
E(|XY|) \leq (E(|X|^p))^{1/p} (E(|Y|^q))^{1/q}.
\]

**Solution:** The inequality certainly holds for $\alpha = 0$ and $\beta = 0$. Let $\alpha = e^{x/p}$ and $\beta = e^{y/q}$ where $x, y \in \mathbb{R}$. By substituting $\lambda = 1/p$ and $1 - \lambda = 1/q$,

\[
e^{\lambda x + (1-\lambda)y} \leq \lambda e^x + (1-\lambda)e^y
\]

holds true since the exponential function is a convex function and hence

\[
\alpha \beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.
\]

By setting

\[
\alpha = \frac{|X|}{E(|X|^p)}^{1/p}, \quad \beta = \frac{|Y|}{E(|Y|^q)}^{1/q}
\]

hence

\[
\frac{|XY|}{E(|X|^p)}^{1/p} (E(|Y|^q))^{1/q} \leq \frac{|X|^p}{p E(|X|^p)} + \frac{|Y|^q}{q E(|Y|^q)}.
\]

Taking expectations we obtain

\[
E(|XY|) \leq (E(|X|^p))^{1/p} (E(|Y|^q))^{1/q}.
\]

\[\square\]

3. **Minkowski’s Inequality.** Let $X$ and $Y$ be a pair of jointly continuous variables, show that if $p \geq 1$ then

\[
\{E(|X + Y|^p)\}^{1/p} \leq \{E(|X|^p)\}^{1/p} + \{E(|Y|^p)\}^{1/p}.
\]

**Solution:** Since $E(|X + Y|) \leq E(|X|) + E(|Y|)$ and using Hölder’s inequality we can write

\[
E(|X + Y|^p) = E\left(|X + Y||X + Y|^{p-1}\right)
\]

\[
\leq E\left(|X||X + Y|^{p-1}\right) + E\left(|Y||X + Y|^{p-1}\right)
\]

\[
\leq \{E(|X|^p)\}^{1/p} \{E(|X + Y|^{p-1})\}^{1/q} + \{E(|Y|^p)\}^{1/p} \{E(|X + Y|^{p-1})\}^{1/q}
\]

\[
= \{E(|X|^p)\}^{1/p} \{E(|X + Y|^p)\}^{1/q} + \{E(|Y|^p)\}^{1/p} \{E(|X + Y|^p)\}^{1/q}
\]

since $\frac{1}{p} + \frac{1}{q} = 1$.

Dividing the inequality by $\{E(|X + Y|^p)\}^{1/q}$ we get

\[
\{E(|X + Y|^p)\}^{1/p} \leq \{E(|X|^p)\}^{1/p} + \{E(|Y|^p)\}^{1/p}.
\]

\[\square\]
4. Change of Measure. Let $\Omega$ be a probability space and let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on $\Omega$. Let $Z(\omega)$ be the Radon–Nikodým derivative defined as

$$Z(\omega) = \frac{\mathbb{Q}(\omega)}{\mathbb{P}(\omega)}$$

such that $\mathbb{P}(Z > 0) = 1$. By denoting $\mathbb{E}^\mathbb{P}$ and $\mathbb{E}^\mathbb{Q}$ as expectations under the measure $\mathbb{P}$ and $\mathbb{Q}$, respectively, show that for any random variable $X$,

$$\mathbb{E}^\mathbb{Q}(X) = \mathbb{E}^\mathbb{P}(XZ), \quad \mathbb{E}^\mathbb{P}(X) = \mathbb{E}^\mathbb{Q}\left(\frac{X}{Z}\right).$$

Solution: By definition

$$\mathbb{E}^\mathbb{Q}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{Q}(\omega) = \sum_{\omega \in \Omega} X(\omega)Z(\omega)\mathbb{P}(\omega) = \mathbb{E}^\mathbb{P}(XZ).$$

Similarly

$$\mathbb{E}^\mathbb{P}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) = \sum_{\omega \in \Omega} \frac{X(\omega)}{Z(\omega)}\mathbb{Q}(\omega) = \mathbb{E}^\mathbb{Q}\left(\frac{X}{Z}\right).$$

5. Conditional Probability. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $\mathbb{1}_A$ is an indicator random variable for an event $A$ defined as

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

show that

$$\mathbb{E}(\mathbb{1}_A|\mathcal{G}) = \mathbb{P}(A|\mathcal{G}).$$

Solution: Since $\mathbb{E}(\mathbb{1}_A|\mathcal{G})$ is $\mathcal{G}$ measurable we need to show that the following partial averaging property:

$$\int_B \mathbb{E}(\mathbb{1}_A|\mathcal{G}) \, d\mathbb{P} = \int_B \mathbb{1}_A \, d\mathbb{P} = \int_B \mathbb{P}(A|\mathcal{G}) \, d\mathbb{P}$$

is satisfied for $B \in \mathcal{G}$. Setting

$$\mathbb{1}_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \mathbb{1}_{A \cap B}(\omega) = \begin{cases} 1 & \text{if } \omega \in A \cap B \\ 0 & \text{otherwise} \end{cases}$$

and expanding $\int_B \mathbb{P}(A|\mathcal{G}) \, d\mathbb{P}$ we have

$$\int_B \mathbb{P}(A|\mathcal{G}) \, d\mathbb{P} = \mathbb{P}(A \cap B) = \int_{\Omega} \mathbb{1}_{A \cap B} \, d\mathbb{P} = \int_{\Omega} \mathbb{1}_A \, d\mathbb{P} \cdot \mathbb{1}_B \, d\mathbb{P} = \int_B \mathbb{1}_A \, d\mathbb{P}.$$
1.2.3 Properties of Expectations

Since $\mathbb{E}(\mathbb{1}_A|\mathcal{G})$ is $\mathcal{G}$ measurable we have

$$\int_B \mathbb{E}(\mathbb{1}_A|\mathcal{G}) \ d\mathbb{P} = \int_B \mathbb{1}_A \ d\mathbb{P}$$

and hence $\mathbb{E}(\mathbb{1}_A|\mathcal{G}) = \mathbb{P}(A|\mathcal{G})$.

\[\square\]

6. Linearity. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $X_1, X_2, \ldots, X_n$ are integrable random variables and $c_1, c_2, \ldots, c_n$ are constants, show that

$$\mathbb{E} \left( c_1 X_1 + c_2 X_2 + \ldots + c_n X_n | \mathcal{G} \right) = c_1 \mathbb{E}(X_1|\mathcal{G}) + c_2 \mathbb{E}(X_2|\mathcal{G}) + \ldots + c_n \mathbb{E}(X_n|\mathcal{G}).$$

Solution: Given $\mathbb{E} \left( c_1 X_1 + c_2 X_2 + \ldots + c_n X_n | \mathcal{G} \right)$ is $\mathcal{G}$ measurable, and for any $A \in \mathcal{G}$,

$$\int_A \mathbb{E} \left( c_1 X_1 + c_2 X_2 + \ldots + c_n X_n | \mathcal{G} \right) \ d\mathbb{P} = \int_A (c_1 X_1 + c_2 X_2 + \ldots + c_n X_n) \ d\mathbb{P}$$

$$= c_1 \int_A X_1 \ d\mathbb{P} + c_2 \int_A X_2 \ d\mathbb{P} + \ldots + c_n \int_A X_n \ d\mathbb{P}.$$

Since $\int_A X_i \ d\mathbb{P} = \int_A \mathbb{E}(X_i|\mathcal{G}) \ d\mathbb{P}$ for $i = 1, 2, \ldots, n$ therefore $\mathbb{E} \left( c_1 X_1 + c_2 X_2 + \ldots + c_n X_n | \mathcal{G} \right) = c_1 \mathbb{E}(X_1|\mathcal{G}) + c_2 \mathbb{E}(X_2|\mathcal{G}) + \ldots + c_n \mathbb{E}(X_n|\mathcal{G}).$

\[\square\]

7. Positivity. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $X$ is an integrable random variable such that $X \geq 0$ almost surely then show that $\mathbb{E}(X|\mathcal{G}) \geq 0$ almost surely.

Solution: Let $A = \{ w \in \Omega : \mathbb{E}(X|\mathcal{G}) < 0 \}$ and since $\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$ measurable therefore $A \in \mathcal{G}$. Thus, from the partial averaging property we have

$$\int_A \mathbb{E}(X|\mathcal{G}) \ d\mathbb{P} = \int_A X \ d\mathbb{P}.$$

Since $X \geq 0$ almost surely therefore $\int_A X \ d\mathbb{P} \geq 0$ but $\int_A \mathbb{E}(X|\mathcal{G}) \ d\mathbb{P} < 0$, which is a contradiction. Thus, $\mathbb{P}(A) = 0$, which implies $\mathbb{E}(X|\mathcal{G}) \geq 0$ almost surely.

\[\square\]

8. Monotonicity. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $X$ and $Y$ are integrable random variables such that $X \leq Y$ almost surely then show that

$$\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G}).$$
1.2.3 Properties of Expectations

Solution: Since $\mathbb{E}(X - Y | \mathcal{G})$ is $\mathcal{G}$ measurable, for $A \in \mathcal{G}$ we can write

$$\int_A \mathbb{E}(X - Y | \mathcal{G}) \, d\mathbb{P} = \int_A (X - Y) \, d\mathbb{P}$$

and since $X \leq Y$, from Problem 1.2.3.7 (page 44) we can deduce that

$$\int_A \mathbb{E}(X - Y | \mathcal{G}) \, d\mathbb{P} \leq 0$$

and hence

$$\mathbb{E}(X - Y | \mathcal{G}) \leq 0.$$

Using the linearity of conditional expectation (see Problem 1.2.3.6, page 44)

$$\mathbb{E}(X - Y | \mathcal{G}) = \mathbb{E}(X | \mathcal{G}) - \mathbb{E}(Y | \mathcal{G}) \leq 0$$

and therefore $\mathbb{E}(X | \mathcal{G}) \leq \mathbb{E}(Y | \mathcal{G})$.

\[\square\]

9. Computing Expectations by Conditioning. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). Show that

$$\mathbb{E}[\mathbb{E}(X | \mathcal{G})] = \mathbb{E}(X).$$

Solution: From the partial averaging property we have, for $A \in \mathcal{G}$,

$$\int_A \mathbb{E}(X | \mathcal{G}) \, d\mathbb{P} = \int_A X \, d\mathbb{P}$$

or

$$\mathbb{E} \left[ \mathbb{1}_A \cdot \mathbb{E}(X | \mathcal{G}) \right] = \mathbb{E}(\mathbb{1}_A \cdot X)$$

where

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

is a $\mathcal{G}$ measurable random variable. By setting $A = \Omega$ we obtain $\mathbb{E}[\mathbb{E}(X | \mathcal{G})] = \mathbb{E}(X)$.

\[\square\]

10. Taking Out What is Known. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $X$ and $Y$ are integrable random variables and if $X$ is $\mathcal{G}$ measurable show that

$$\mathbb{E}(XY | \mathcal{G}) = X \cdot \mathbb{E}(Y | \mathcal{G}).$$
1.2.3 Properties of Expectations

Solution: Since $X$ and $\mathbb{E}(Y|\mathcal{G})$ are $\mathcal{G}$ measurable therefore $X \cdot \mathbb{E}(Y|\mathcal{G})$ is also $\mathcal{G}$ measurable and it satisfies the first property of conditional expectation. By calculating the partial averaging of $X \cdot \mathbb{E}(Y|\mathcal{G})$ over a set $A \in \mathcal{G}$ and by defining

$$
\mathbb{I}_A(\omega) = \begin{cases} 
1 & \text{if } \omega \in A \\
0 & \text{otherwise}
\end{cases}
$$

such that $\mathbb{I}_A$ is a $\mathcal{G}$ measurable random variable we have

$$
\int_A X \cdot \mathbb{E}(Y|\mathcal{G}) \, d\mathbb{P} = \mathbb{E}[\mathbb{I}_A \cdot X \mathbb{E}(Y|\mathcal{G})] = \mathbb{E}[\mathbb{I}_A \cdot XY] = \int_A XY \, d\mathbb{P}.
$$

Thus, $X \cdot \mathbb{E}(Y|\mathcal{G})$ satisfies the partial averaging property by setting $\int_A XY \, d\mathbb{P} = \int_A \mathbb{E}(XY|\mathcal{G}) \, d\mathbb{P}$. Therefore, $\mathbb{E}(XY|\mathcal{G}) = X \cdot \mathbb{E}(Y|\mathcal{G})$.

11. Tower Property. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $\mathcal{H}$ is a sub-$\sigma$-algebra of $\mathcal{G}$ (i.e., sets in $\mathcal{H}$ are also in $\mathcal{G}$) and $X$ is an integrable random variable, show that

$$
\mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}] = \mathbb{E}(X|\mathcal{H}).
$$

Solution: For an integrable random variable $Y$, by definition we know that $\mathbb{E}(Y|\mathcal{H})$ is $\mathcal{H}$ measurable, and hence by setting $Y = \mathbb{E}(X|\mathcal{G})$, and for $A \in \mathcal{H}$, the partial averaging property of $\mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}]$ is

$$
\int_A \mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}] \, d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P}.
$$

Since $A \in \mathcal{H}$ and $\mathcal{H}$ is a sub-$\sigma$-algebra of $\mathcal{G}$, $A \in \mathcal{G}$. Therefore,

$$
\int_A \mathbb{E}(X|\mathcal{H}) \, d\mathbb{P} = \int_A X \, d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P}.
$$

This shows that $\mathbb{E}(X|\mathcal{H})$ satisfies the partial averaging property of $\mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}]$, and hence $\mathbb{E}[\mathbb{E}(X|\mathcal{G})|\mathcal{H}] = \mathbb{E}(X|\mathcal{H})$.

12. Measurability. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If the random variable $X$ is $\mathcal{G}$ measurable then show that

$$
\mathbb{E}(X|\mathcal{G}) = X.
$$
1.2.3 Properties of Expectations

**Solution:** From the partial averaging property, for $A \in \Omega$,

$$\int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P} = \int_A X \, d\mathbb{P}$$

and if $X$ is $\mathcal{G}$ measurable then it satisfies

$$\mathbb{E}(X|\mathcal{G}) = X.$$ 

\[\square\]

13. **Independence.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ (i.e., sets in $\mathcal{G}$ are also in $\mathcal{F}$). If $X = \mathbb{1}_B$ such that

$$\mathbb{1}_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{otherwise} \end{cases}$$

and $\mathbb{1}_B$ is independent of $\mathcal{G}$ show that

$$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X).$$

**Solution:** Since $\mathbb{E}(X)$ is non-random then $\mathbb{E}(X)$ is $\mathcal{G}$ measurable. Therefore, we now need to check that the following partial averaging property:

$$\int_A \mathbb{E}(X) \, d\mathbb{P} = \int_A X \, d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P}$$

is satisfied for $A \in \mathcal{G}$.

Let $X = \mathbb{1}_B$ such that

$$\mathbb{1}_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{otherwise} \end{cases}$$

and the random variable $\mathbb{1}_B$ is independent of $\mathcal{G}$. In addition, we also define

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbb{1}_A$ is $\mathcal{G}$ measurable. For all $A \in \mathcal{G}$ we have

$$\int_A X \, d\mathbb{P} = \int_A \mathbb{1}_B \, d\mathbb{P} = \int_A \mathbb{P}(B) \, d\mathbb{P} = \mathbb{P}(A) \mathbb{P}(B).$$

Furthermore, since the sets $A$ and $B$ are independent we can also write

$$\int_A X \, d\mathbb{P} = \int_A \mathbb{1}_B \, d\mathbb{P} = \int_\Omega \mathbb{1}_A \mathbb{1}_B \, d\mathbb{P} = \int_\Omega \mathbb{1}_{A \cap B} \, d\mathbb{P} = \mathbb{P}(A \cap B).$$
where
\[ 1_{A \cap B}(\omega) = \begin{cases} 
1 & \text{if } \omega \in A \cap B \\
0 & \text{otherwise}
\end{cases} \]
and hence
\[ \int_A X \, d\mathbb{P} = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{E}(X) = \int_A \mathbb{E}(X) \, d\mathbb{P}. \]

Thus, we have \( \mathbb{E}(X) = \mathbb{E}(X). \)

\[ \Box \]

14. **Conditional Jensen’s Inequality.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\mathcal{G}\) be a sub-\(\sigma\)-algebra of \(\mathcal{F}\) (i.e., sets in \(\mathcal{G}\) are also in \(\mathcal{F}\)). If \(\varphi : \mathbb{R} \rightarrow \mathbb{R}\) is a convex function and \(X\) is an integrable random variable show that
\[ \mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}(X)|\mathcal{G}). \]

Deduce that if \(X\) is independent of \(\mathcal{G}\) then the above inequality becomes
\[ \mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}(X)). \]

**Solution:** Given that \(\varphi\) is a convex function,
\[ \varphi(x) \geq \varphi(y) + \varphi'(y)(y - x). \]

By setting \(x = X\) and \(y = \mathbb{E}(X)|\mathcal{G}\) we have
\[ \varphi(X) \geq \varphi(\mathbb{E}(X)|\mathcal{G}) + \varphi'[\mathbb{E}(X)|\mathcal{G}][\mathbb{E}(X)|\mathcal{G}] - X \]
and taking conditional expectations,
\[ \mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}(X)|\mathcal{G}). \]

If \(X\) is independent of \(\mathcal{G}\) then from Problem 1.2.3.13 (page 47) we can set \(y = \mathbb{E}(X)|\mathcal{G} = \mathbb{E}(X). Using the same steps as described above we have
\[ \varphi(X) \geq \varphi(\mathbb{E}(X)) + \varphi'[\mathbb{E}(X)][\mathbb{E}(X) - X] \]
and taking expectations we finally have
\[ \mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}(X)). \]

\[ \Box \]
15. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\mathcal{G}\) be a sub-\(\sigma\)-algebra of \(\mathcal{F}\) (i.e., sets in \(\mathcal{G}\) are also in \(\mathcal{F}\)). If \(X\) is an integrable random variable and \(\mathbb{E}(X^2) < \infty\) show that
\[
\mathbb{E}\left[\mathbb{E}(X^2|\mathcal{G})\right] \leq \mathbb{E}(X^2).
\]

**Solution:** From the conditional Jensen’s inequality (see Problem 1.2.3.14, page 48) we set \(\psi(x) = x^2\) which is a convex function. By substituting \(x = \mathbb{E}(X|\mathcal{G})\) we have
\[
\mathbb{E}(X^2|\mathcal{G}) \leq \mathbb{E}(X^2|\mathcal{G})\).
\]

Taking expectations
\[
\mathbb{E}\left[\mathbb{E}(X^2|\mathcal{G})\right] \leq \mathbb{E}\left[\mathbb{E}(X^2|\mathcal{G})\right]
\]
and from the tower property (see Problem 1.2.3.11, page 46)
\[
\mathbb{E}\left[\mathbb{E}(X^2|\mathcal{G})\right] = \mathbb{E}(X^2).
\]
Thus, \(\mathbb{E}\left[\mathbb{E}(X^2|\mathcal{G})\right] \leq \mathbb{E}(X^2)\).
2 Wiener Process

In mathematics, a Wiener process is a stochastic process sharing the same behaviour as Brownian motion, which is a physical phenomenon of random movement of particles suspended in a fluid. Generally, the terms “Brownian motion” and “Wiener process” are the same, although the former emphasises the physical aspects whilst the latter emphasises the mathematical aspects. In quantitative analysis, by drawing on the mathematical properties of Wiener processes to explain economic phenomena, financial information such as stock prices, commodity prices, interest rates, foreign exchange rates, etc. are treated as random quantities and then mathematical models are constructed to capture the randomness. Given these financial models are stochastic and continuous in nature, the Wiener process is usually employed to express the random component of the model. Before we discuss the models in depth, in this chapter we first look at the definition and basic properties of a Wiener process.

2.1 INTRODUCTION

By definition, a random walk is a mathematical formalisation of a trajectory that consists of taking successive random steps at every point in time. To construct a Wiener process in continuous time, we begin by setting up a symmetric random walk – such as tossing a fair coin infinitely many times where the probability of getting a head (H) or a tail (T) in each toss is \(\frac{1}{2}\). By defining the \(i\)-th toss as

\[
Z_i = \begin{cases} 
1 & \text{if toss is } H \\
-1 & \text{if toss is } T 
\end{cases}
\]

and setting \(M_0 = 0\), the process

\[
M_k = \sum_{i=1}^{k} Z_i, \quad k = 1, 2, \ldots
\]

is a symmetric random walk. In a continuous time setting, to approximate a Wiener process for \(n \in \mathbb{Z}^+\) we define the scaled symmetric random walk as

\[
W^{(n)}_t = \frac{1}{\sqrt{n}} M_{\lfloor nt \rfloor} = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} Z_i
\]

such that in the limit of \(n \to \infty\) we can obtain the Wiener process where

\[
\lim_{n \to \infty} W^{(n)}_t = \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} Z_i \xrightarrow{D} \mathcal{N}(0, t).
\]
Given that a Wiener process is a limiting distribution of a scaled symmetric random walk, the following is the definition of a standard Wiener process.

**Definition 2.1 (Standard Wiener Process)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A stochastic process \(\{W_t : t \geq 0\}\) is defined to be a standard Wiener process (or \(\mathbb{P}\)-standard Wiener process) if:

(a) \(W_0 = 0\) and has continuous sample paths;
(b) for each \(t > 0\) and \(s > 0\), \(W_{t+s} - W_t \sim \mathcal{N}(0, s)\) (stationary increment);
(c) for each \(t > 0\) and \(s > 0\), \(W_{t+s} - W_t \perp \perp W_t\) (independent increment).

A standard Wiener process is a standardised version of a Wiener process, which need not begin at \(W_0 = 0\), and may have a non-zero drift term \(\mu \neq 0\) and a variance term not necessarily equal to \(t\).

**Definition 2.2 (Wiener Process)** A process \(\{\hat{W}_t : t \geq 0\}\) is called a Wiener process if it can be written as

\[ \hat{W}_t = \nu + \mu t + \sigma W_t \]

where \(\nu, \mu \in \mathbb{R}, \sigma > 0\) and \(W_t\) is a standard Wiener process.

Almost all financial models have the Markov property and without exception the Wiener process also has this important property, where it is used to relate stochastic calculus to partial differential equations and ultimately to the pricing of options. The following is an important result concerning the Markov property of a standard Wiener process.

**Theorem 2.3 (Markov Property)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. The standard Wiener process \(\{W_t : t \geq 0\}\) is a Markov process such that the conditional distribution of \(W_t\) given the filtration \(\mathcal{F}_s, s < t\) depends only on \(W_s\).

Another generalisation from the Markov property is the strong Markov property, which is an important result in establishing many other properties of Wiener processes such as martingales. Clearly, the strong Markov property implies the Markov property but not vice versa.

**Theorem 2.4 (Strong Markov Property)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. If \(\{W_t : t \geq 0\}\) is a standard Wiener process and given \(\mathcal{F}_t\), is the filtration up to time \(t\), then for \(s > 0\), \(W_{t+s} - W_t \perp \perp \mathcal{F}_t\).

Once we have established the Markov properties, we can use them to show that a Wiener process is a martingale. Basically, a stochastic process is a martingale when its conditional expected value of an observation at some future time \(t\), given all the observations up to some earlier time \(s\), is equal to the observation at that earlier time \(s\). In formal terms the definition of this property is given as follows.
Definition 2.5 (Martingale Property for Continuous Process) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A stochastic process \(\{X_t : t \geq 0\}\) is a continuous-time martingale if:

(a) \(\mathbb{E}(X_t | \mathcal{F}_s) = X_s\) for all \(0 \leq s \leq t\);
(b) \(\mathbb{E}(|X_t|) < \infty\);
(c) \(X_t\) is \(\mathcal{F}_t\)-adapted.

In addition to properties (b) and (c), the process is a submartingale if \(\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s\) and a supermartingale if \(\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s\) for all \(0 \leq s \leq t\).

In contrast, we can also define the martingale property for a discrete process.

Definition 2.6 (Martingale Property for Discrete Process) A discrete process \(X = \{X_n : n = 0, 1, 2, \ldots\}\) is a martingale relative to \((\Omega, \mathcal{F}, \mathbb{P})\) if for all \(n:\)

(a) \(\mathbb{E}(X_{n+1} | \mathcal{F}_n) = X_n\);
(b) \(\mathbb{E}(|X_n|) < \infty\);
(c) \(X_n\) is \(\mathcal{F}_n\)-adapted.

Together with properties (b) and (c), the process is a submartingale if \(\mathbb{E}(X_{n+1} | \mathcal{F}_n) \geq X_n\) and a supermartingale if \(\mathbb{E}(X_{n+1} | \mathcal{F}_n) \leq X_n\) for all \(n\).

A most important application of the martingale property of a stochastic process is in the area of derivatives pricing where under the risk-neutral measure, to avoid any arbitrage opportunities, all asset prices have the same expected rate of return that is the risk-free rate. As we shall see in later chapters, by modelling an asset price with a Wiener process to represent the random component, under the risk-neutral measure, the expected future price of the asset discounted at a risk-free rate given its past information is a martingale. In addition, martingales do have certain features even when they are stopped at random times, as given in the following theorem.

Theorem 2.7 (Optional Stopping (Sampling)) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A random time \(T \in [0, \infty)\) is called a stopping time if \(\{T \leq t\} \in \mathcal{F}_t\) for all \(t \geq 0\). If \(T < \infty\) is a stopping time and \(\{X_t\}\) is a martingale, then \(\mathbb{E}(X_T) = \mathbb{E}(X_0)\) if any of the following are true:

(a) \(\mathbb{E}(T) < \infty\);
(b) there exists a constant \(K\) such that \(\mathbb{E}(|X_{t+\Delta t} - X_t|) \leq K\) for \(\Delta t > 0\).

Take note that an important application of the optional stopping theorem is the first passage time of a standard Wiener process hitting a level. Here one can utilise it to analyse American options, where in this case the exercise time is a stopping time. In tandem with the stopping time, the following reflection principle result allows us to find the joint distribution of \((\max_{0 \leq t \leq T} W_t, W_T)\) and the distribution of \(\max_{0 \leq t \leq T} W_t\), which in turn can be used to price exotic options such as barrier and lookback options.
Theorem 2.8 (Reflection Principle) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. By setting $T$ as a stopping time and defining

$$
\tilde{W}_t = \begin{cases} 
W_t & \text{if } t \leq T \\
2W_T - W_t & \text{if } t > T 
\end{cases}
$$

then $\{\tilde{W}_t : t \geq 0\}$ is also a standard Wiener process.

Finally, another useful property of the Wiener process is the quadratic variation, where if $\{W_t : t \geq 0\}$ is a standard Wiener process then by expressing $dW_t$ as the infinitesimal increment of $W_t$ and setting $t_i = it/n$, $i = 0, 1, 2, \ldots, n - 1$, $n > 0$ such that $0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t$, the quadratic variation of $W_t$ is defined as

$$
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = \int_0^t dW_u^2 = t
$$

which has a finite value. In addition, the cross-variation of $W_t$ and $t$, the quadratic variation of $t$ and the $p$-order variation, $p \geq 3$ of $W_t$ can be expressed as

$$
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})(t_{i+1} - t_i) = \int_0^t dW_u dt = 0,
$$

$$
\lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = \int_0^t du^2 = 0
$$

and

$$
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^p = \int_0^t dW_u^p = 0, \quad p \geq 3.
$$

Informally, we can therefore write

$$
(dW_t)^2 = dt, \quad dW_t dt = 0, \quad (dt)^2 = 0, \quad (dW_t)^p = 0, \quad p \geq 3
$$

where $dW_t$ and $dt$ are the infinitesimal increment of $W_t$ and $t$, respectively. The significance of the above results constitutes the key ingredients in Itô’s formula to find the differential of a stochastic function and also in deriving the Black–Scholes equation to price options.

In essence, the many properties of the Wiener process made it a very suitable choice to express the random component when modelling stock prices, interest rates, foreign currency exchange rates, etc. One notable example is when pricing European-style options, where, due to its inherent properties we can obtain closed-form solutions which would not be possible had it been modelled using other processes. In addition, we could easily extend the one-dimensional Wiener process to a multi-dimensional Wiener process to model stock prices that are correlated with each other as well as stock prices under stochastic volatility. However, owing to the normal distribution of the standard Wiener process it tends to provide stock price returns that are symmetric and short-tailed. In practical situations, stock price returns are skewed and have heavy tails and hence they do not follow a normal distribution. Nevertheless,
2.2.1 Basic Properties

even if there exist complex models to capture such attributes (for example, using a Lévy process), a Wiener process is still a vital component to model sources of uncertainty in finance.

2.2 PROBLEMS AND SOLUTIONS

2.2.1 Basic Properties

1. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. We consider a symmetric random walk such that the \(j\)-th step is defined as

\[
Z_j = \begin{cases} 
1 & \text{with probability } \frac{1}{2} \\ 
-1 & \text{with probability } \frac{1}{2} 
\end{cases}
\]

where \(Z_i \perp Z_j, i \neq j\). By setting \(0 = k_0 < k_1 < k_2 < \ldots < k_t\), we let

\[
M_{k_i} = \sum_{j=1}^{k_i} Z_j, \quad i = 1, 2, \ldots, t
\]

where \(M_0 = 0\).

Show that the symmetric random walk has independent increments such that the random variables

\[
M_{k_{i+1}} - M_{k_i}, M_{k_{i+2}} - M_{k_{i+1}}, \ldots, M_{k_t} - M_{k_{i-1}}
\]

are independent.

Finally, show that \(\mathbb{E}(M_{k_{i+1}} - M_{k_i}) = 0\) and \(\text{Var}(M_{k_{i+1}} - M_{k_i}) = k_{i+1} - k_i\).

Solution: By definition

\[
M_{k_i} - M_{k_{i-1}} = \sum_{j=k_{i-1}+1}^{k_i} Z_j = s_i
\]

and for \(m < n, m, n = 1, 2, \ldots, t\),

\[
\mathbb{P}(M_{k_m} - M_{k_{m-1}} = s_m | M_{k_n} - M_{k_{n-1}} = s_n) = \mathbb{P}(M_{k_n} - M_{k_{n-1}} = s_n) / \mathbb{P}(M_{k_{m-1}} - M_{k_{m-1}} = s_m)
\]

\[
= \mathbb{P}(\text{sum of walks in } [k_{m-1} + 1, k_m] \cap \text{sum of walks in } [k_{m-1} + 1, k_m]) / \mathbb{P}(\text{sum of walks in } [k_{m-1} + 1, k_m]).
\]

Because \(m < n\) and since \(Z_i\) is independent of \(Z_j, i \neq j\), there are no overlapping events between the intervals \([k_{n-1} + 1, k_n]\) and \([k_{m-1} + 1, k_m]\) and hence for \(m < n, m, n = 1, 2, \ldots, t\),

\[
\mathbb{P}(M_{k_m} - M_{k_{m-1}} = s_m | M_{k_n} - M_{k_{n-1}} = s_n) = \mathbb{P}(\text{sum of walks in } [k_{n-1} + 1, k_n]) = \mathbb{P}(M_{k_m} - M_{k_{m-1}} = s_m).
\]
Thus, we can deduce that $M_{k_1} - M_{k_0}, M_{k_2} - M_{k_1}, \ldots, M_{k_t} - M_{k_{t-1}}$ are independent. Since $\mathbb{E}(Z_j) = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0$ and $\text{Var}(Z_j) = \mathbb{E}(Z_j^2) = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1$, and because $Z_j, j = 1, 2, \ldots$ are independent, we have

$$\mathbb{E}(M_{k_i} - M_{k_{i-1}}) = \sum_{j=k_{i-1}+1}^{k_i} \mathbb{E}(Z_j) = 0$$

and

$$\text{Var}(M_{k_i} - M_{k_{i-1}}) = \sum_{j=k_{i-1}+1}^{k_i} \text{Var}(Z_j) = \sum_{j=k_{i-1}+1}^{k_i} 1 = k_i - k_{i-1}.$$

\[\Box\]

2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For a symmetric random walk

$$M_k = \sum_{i=1}^{k} Z_i$$

with starting point at $M_0 = 0$, $\mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = -1) = \frac{1}{2}$, show that $M_k$ is a martingale.

**Solution:** To show that $M_k$ is a martingale we note that

(a) For $j < k, j, k \in \mathbb{Z}^+$, using the independent increment property and $\mathbb{E}(M_k - M_j) = 0$ as shown in Problem 2.2.1.1 (page 55),

$$\mathbb{E}(M_k | \mathcal{F}_j) = \mathbb{E}(M_k - M_j + M_j | \mathcal{F}_j)$$

$$= \mathbb{E}(M_k - M_j | \mathcal{F}_j) + \mathbb{E}(M_j | \mathcal{F}_j)$$

$$= \mathbb{E}(M_k - M_j) + M_j$$

$$= M_j.$$

(b) Given $M_k = \sum_{i=1}^{k} Z_i$ it follows that

$$|M_k| = \left| \sum_{i=1}^{k} Z_i \right| \leq \sum_{i=1}^{k} |Z_i| = \sum_{i=1}^{k} 1 = k < \infty.$$

(c) $M_k$ is clearly $\mathcal{F}_k$-adapted.

From the results of (a)–(c), we have shown that $M_k$ is a martingale.

\[\Box\]

3. **Donsker Theorem.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For a symmetric random walk

$$M_k = \sum_{i=1}^{k} Z_i$$
2.2.1 Basic Properties

where \( M_0 = 0, \Pr(Z_1 = 1) = \Pr(Z_1 = -1) = \frac{1}{2} \) and by defining

\[
W_{t}^{(n)} = \frac{1}{\sqrt{n}} M_{[nt]} = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Z_i
\]

for a fixed time \( t \), show that

\[
limit_{n \to \infty} W_{t}^{(n)} = \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Z_i \xrightarrow{D} \mathcal{N}(0, t).
\]

**Solution:** Since for all \( i = 1, 2, \ldots \),

\[
\mathbb{E}(Z_i) = 0 \quad \text{and} \quad \text{Var}(Z_i) = 1
\]

and given \( Z_i \perp Z_j, i \neq j \),

\[
\mathbb{E}\left(W_{t}^{(n)}\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathbb{E}(Z_i) = 0
\]

and

\[
\text{Var}\left(W_{t}^{(n)}\right) = \frac{1}{n} \sum_{i=1}^{[nt]} \text{Var}(Z_i) = \frac{1}{n} \sum_{i=1}^{[nt]} 1 = \frac{[nt]}{n}.
\]

For \( n \to \infty \) we can deduce that

\[
\lim_{n \to \infty} \mathbb{E}\left(W_{t}^{(n)}\right) = 0
\]

and

\[
\lim_{n \to \infty} \text{Var}\left(W_{t}^{(n)}\right) = \lim_{n \to \infty} \frac{[nt]}{n} = t.
\]

Therefore, from the central limit theorem,

\[
limit_{n \to \infty} W_{t}^{(n)} = \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Z_i \xrightarrow{D} \mathcal{N}(0, t).
\]

\( \square \)

4. **Covariance of Two Standard Wiener Processes.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Show that

\[
\text{Cov}(W_s, W_t) = \min\{s, t\}
\]

and deduce that the correlation coefficient of \(W_s\) and \(W_t\) is

\[
\rho = \sqrt{\frac{\min\{s, t\}}{\max\{s, t\}}}.
\]
2.2.1 Basic Properties

Solution: Since $W_t \sim \mathcal{N}(0, t)$, $W_s \sim \mathcal{N}(0, s)$ and by definition

$$\text{Cov}(W_s, W_t) = \mathbb{E}(W_s W_t) - \mathbb{E}(W_s)\mathbb{E}(W_t) = \mathbb{E}(W_s W_t).$$

Let $s \leq t$ and because $W_s \perp \perp W_t - W_s$,

$$\mathbb{E}(W_s W_t) = \mathbb{E}(W_s(W_t - W_s) + W_s^2)$$
$$= \mathbb{E}(W_s(W_t - W_s)) + \mathbb{E}(W_s^2)$$
$$= \mathbb{E}(W_t - W_s) + \mathbb{E}(W_s^2)$$
$$= s.$$

For $s > t$ and because $W_t \perp \perp W_s - W_t$,

$$\mathbb{E}(W_s W_t) = \mathbb{E}(W_t(W_s - W_t) + W_t^2)$$
$$= \mathbb{E}(W_t(W_s - W_t)) + \mathbb{E}(W_t^2)$$
$$= \mathbb{E}(W_s - W_t) + \mathbb{E}(W_t^2)$$
$$= t.$$

Therefore, $\text{Cov}(W_s, W_t) = s \wedge t = \min\{s, t\}$.

By definition, the correlation coefficient of $W_s$ and $W_t$ is defined as

$$\rho = \frac{\text{Cov}(W_s, W_t)}{\sqrt{\text{Var}(W_s)\text{Var}(W_t)}}$$
$$= \frac{s}{\sqrt{st}}.$$

For $s \leq t$,

$$\rho = \frac{s}{\sqrt{st}} = \sqrt{\frac{s}{t}},$$

whilst for $s > t$,

$$\rho = \frac{t}{\sqrt{st}} = \sqrt{\frac{t}{s}}.$$

Therefore, $\rho = \sqrt{\frac{\min\{s, t\}}{\max\{s, t\}}}$. \qed

5. Joint Distribution of Standard Wiener Processes. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_{t_k} : t_k \geq 0\}, k = 0, 1, 2, \ldots, n$ be a standard Wiener process where $0 = t_0 < t_1 < t_2 < \ldots < t_n$.

Find the moment generating function of the joint distribution $(W_{t_1}, W_{t_2}, \ldots, W_{t_n})$ and its corresponding probability density function.
2.2.1 Basic Properties

Show that for $t < T$, the conditional distributions

$$W_t | W_T = y \sim \mathcal{N}\left(\frac{yt}{T}, \frac{T-t}{T}\right)$$

and

$$W_T | W_t = x \sim \mathcal{N}(x, T-t).$$

**Solution:** By definition, the moment generating function for the joint distribution $(W_t, W_{t_2}, \ldots, W_{t_n})$ is given as

$$M_{W_t, W_{t_2}, \ldots, W_{t_n}}(\theta_1, \theta_2, \ldots, \theta_n) = \mathbb{E}\left(e^{\theta_1 W_t + \theta_2 W_{t_2} + \cdots + \theta_n W_{t_n}}\right), \quad \theta_1, \theta_2, \ldots, \theta_n \in \mathbb{R}.$$ 

Given that $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent and normally distributed, we can write

$$\begin{align*}
\theta_1 W_t + \theta_2 W_{t_2} + \cdots + \theta_n W_{t_n} &= (\theta_1 + \theta_2 + \cdots + \theta_n)W_t + (\theta_2 + \cdots + \theta_n)(W_{t_2} - W_t) \\
&\quad + \cdots + \theta_n(W_{t_n} - W_{t_{n-1}})
\end{align*}$$

and since for any $\theta, s < t$, $e^{\theta(W_t - W_s)} \sim \log\mathcal{N}(0, \theta^2(t-s))$ therefore

$$\begin{align*}
\mathbb{E}\left(e^{\theta_1 W_t + \theta_2 W_{t_2} + \cdots + \theta_n W_{t_n}}\right) &= \mathbb{E}\left[e^{(\theta_1 + \theta_2 + \cdots + \theta_n)W_t + (\theta_2 + \cdots + \theta_n)(W_{t_2} - W_t) + \cdots + \theta_n W_{t_n}}\right] \\
&= \mathbb{E}\left[e^{(\theta_2 + \cdots + \theta_n)W_{t_2}}\mathbb{E}[e^{	heta_1 W_t + (\theta_2 + \cdots + \theta_n)(W_{t_2} - W_t) + \cdots + \theta_n W_{t_n}}|W_{t_2}]\right] \\
&\quad \times \cdots \times \mathbb{E}\left[e^{\theta_n(W_{t_n} - W_{t_{n-1}})}\right] \\
&= e^{\frac{1}{2}(\theta_1 + \cdots + \theta_n)^2 + \frac{1}{2}(\theta_2 + \cdots + \theta_n)^2_{t_2-t_1} + \cdots + \frac{1}{2}\theta_n^2_{t_n-t_{n-1}}} \\
&= e^{\frac{1}{2}\theta^T\Sigma\theta}
\end{align*}$$

where $\theta^T = (\theta_1, \theta_2, \ldots, \theta_n)$ and $\Sigma$ is the covariance matrix for the Wiener process. From Problem 2.2.1.4 (page 57) we can express

$$\Sigma = \begin{bmatrix}
\mathbb{E}(W_{t_1}^2) & \mathbb{E}(W_{t_1}W_{t_2}) & \cdots & \mathbb{E}(W_{t_1}W_{t_n}) \\
\mathbb{E}(W_{t_2}W_{t_1}) & \mathbb{E}(W_{t_2}^2) & \cdots & \mathbb{E}(W_{t_2}W_{t_n}) \\
\vdots & \vdots & \ddots & \vdots \\
\mathbb{E}(W_{t_n}W_{t_1}) & \mathbb{E}(W_{t_n}W_{t_2}) & \cdots & \mathbb{E}(W_{t_n}^2)
\end{bmatrix} = \begin{bmatrix} t_1 & t_1 & \cdots & t_1 \\
t_1 & t_1 & \cdots & t_1 \\
\vdots & \vdots & \ddots & \vdots \\
t_1 & t_2 & \cdots & t_n
\end{bmatrix} - \begin{bmatrix} t_1 & t_2 & \cdots & t_n \\
t_2 & t_2 & \cdots & t_n \\
\vdots & \vdots & \ddots & \vdots \\
t_n & t_2 & \cdots & t_n
\end{bmatrix}.$$

Using elementary column operations we can easily show that the determinant $|\Sigma| = \prod_{i=1}^n (t_i - t_{i-1}) \neq 0$. Therefore, $\Sigma^{-1}$ exists and the probability density function for the joint distribution $(W_{t_1}, W_{t_2}, \ldots, W_{t_n})$ is given as

$$f_{W_{t_1}, W_{t_2}, \ldots, W_{t_n}}(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}x^T\Sigma^{-1}x}$$

where $x^T = (x_1, x_2, \ldots, x_n)$. 
For the case of joint distribution of \((W_t, W_T), t < T\) we can deduce that

\[
f_{W_t, W_T}(x, y) = \frac{1}{(2\pi)^{\frac{1}{2}}(t(T-t))^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} \left( \frac{Tx^2 - 2txy + Ty^2}{t(T-t)} \right) \right]
\]

with conditional density function of \(W_t\) given \(W_T = y\)

\[
f_{W_t|W_T}(x|y) = \frac{f_{W_t, W_T}(x, y)}{f_{W_T}(y)} = \frac{1}{\sqrt{2\pi(T-T)}} \exp \left[ -\frac{1}{2} \left( \frac{x - \frac{yt}{T}}{T} \right)^2 \right]
\]

and conditional density function of \(W_T\) given \(W_t = x\)

\[
f_{W_T|W_t}(y|x) = \frac{f_{W_t, W_T}(x, y)}{f_{W_t}(x)} = \frac{1}{\sqrt{2\pi(T-t)}} \exp \left[ -\frac{1}{2} \left( \frac{y - x}{T-t} \right)^2 \right].
\]

Thus,

\[
W_t|W_T = y \sim \mathcal{N}\left(\frac{yt}{T}, \frac{t(T-t)}{T}\right) \quad \text{and} \quad W_T|W_t = x \sim \mathcal{N}(x, T-t).
\]

6. Reflection. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \{\(W_t : t \geq 0\)\} be a standard Wiener process. Show that under reflection, \(B_t = -W_t\) is also a standard Wiener process.

**Solution:**
(a) \(B_0 = -W_0 = 0\), and given that \(W_t\) has continuous sample paths we can deduce that \(B_t\) has continuous sample paths as well.
(b) For \(t > 0, s > 0\)

\[
B_{t+s} - B_t = -W_{t+s} + W_t = -(W_{t+s} - W_t) \sim \mathcal{N}(0, s).
\]

(c) Since \(W_{t+s} - W_t \perp W_t\) therefore

\[
\mathbb{E}[(B_{t+s} - B_t)|B_t] = \text{Cov}(B_{t+s} + W_t, -W_t) = \text{Cov}(W_{t+s} - W_t, W_t) = 0.
\]

Given \(B_t \sim \mathcal{N}(0, t)\) and \(B_{t+s} - B_t \sim \mathcal{N}(0, s)\) and the joint distribution of \(B_t\) and \(B_{t+s} - B_t\) is a bivariate normal (see Problem 2.2.1.5, page 58), then if \(\text{Cov}(B_{t+s} - B_t, B_t) = 0\) so \(B_{t+s} - B_t \perp B_t\).

From the results of (a)–(c) we have shown that \(B_t = -W_t\) is also a standard Wiener process.
7. **Time Shifting.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Show that under time shifting, \(B_t = W_{t+u} - W_u, \ u > 0\) is also a standard Wiener process.

**Solution:** By setting \(\tilde{t} = t + u\).

(a) \(B_0 = W_u - W_u = 0\), and given \(W_t\) has continuous sample paths therefore we can deduce \(B_t\) has continuous sample paths as well.

(b) For \(t > 0, u > 0, s > 0\)

\[
B_{t+s} - B_t = W_{t+s} - W_{t+u} \sim \mathcal{N}(0, s).
\]

(c) Since \(W_{t+s} - W_t \perp W_t\) therefore

\[
\mathbb{E}[(B_{t+s} - B_t)B_t] = \text{Cov}(W_{t+s} - W_{t+u}, W_{t+u}) = \text{Cov}(W_{t+s} - W_t, W_t) = 0.
\]

Given \(B_t \sim \mathcal{N}(0, t)\) and \(B_{t+s} - B_t \sim \mathcal{N}(0, s)\) and the joint distribution of \(B_t\) and \(B_{t+s} - B_t\) is a bivariate normal (see Problem 2.2.1.5, page 58), then if \(\text{Cov}(B_{t+s} - B_t, B_t) = 0\) so \(B_{t+s} - B_t \perp B_t\).

From the results of (a)–(c) we have shown that \(B_t = W_{t+u} - W_u\) is also a standard Wiener process.

8. **Normal Scaling.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Show that under normal scaling, \(B_t = cW_{\frac{t+s}{c^2}}, c \neq 0\) is also a standard Wiener process.

**Solution:**

(a) \(B_0 = cW_0 = 0\) and it is clear that \(B_t\) has continuous sample paths for \(t \geq 0\) and \(c \neq 0\).

(b) For \(t > 0, s > 0\)

\[
B_{t+s} - B_t = c \left( W_{\frac{t+s}{c^2}} - W_{\frac{t}{c^2}} \right)
\]

with mean

\[
\mathbb{E}(B_{t+s} - B_t) = 0
\]

and variance

\[
\text{Var}(B_{t+s} - B_t) = t + s + t - 2\text{min}\{t+s,t\} = s.
\]

Since both \(W_{\frac{t+s}{c^2}} \sim \mathcal{N}\left(0, \frac{t+s}{c^2}\right)\) and \(W_{\frac{t}{c^2}} \sim \mathcal{N}\left(0, \frac{t}{c^2}\right)\), then \(B_{t+s} - B_t \sim \mathcal{N}(0, s)\).
(c) Finally, to show that $B_{t+s} - B_t \perp B_t$, we note that since $W_t$ has the independent increment property, so

$$
\mathbb{E} \left[ (B_{t+s} - B_t) B_t \right] = \mathbb{E} (B_{t+s} B_t) - \mathbb{E} (B_t^2) \\
= \text{Cov} (B_{t+s}, B_t) - \mathbb{E} (B_t^2) \\
= \min\{t + s, t\} - t \\
= 0.
$$

Since $B_t \sim \mathcal{N}(0, t)$ and $B_{t+s} - B_t \sim \mathcal{N}(0, s)$ and the joint distribution of $B_t$ and $B_{t+s} - B_t$ is a bivariate normal (see Problem 2.2.1.5, page 58), then if $\text{Cov} (B_{t+s} - B_t, B_t) = 0$ so $B_{t+s} - B_t \perp B_t$.

From the results of (a)-(c) we have shown that $B_t = c W_t$, $c \neq 0$ is also a standard Wiener process.

\[\Box\]

9. **Time Inversion.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Show that under time inversion,

$$
B_t = \begin{cases} 
0 & \text{if } t = 0 \\
tW_{\frac{1}{t}} & \text{if } t \neq 0
\end{cases}
$$

is also a standard Wiener process.

**Solution:**

(a) $B_0 = 0$ and it is clear that $B_t$ has continuous sample paths for $t > 0$. From continuity at $t = 0$ we can deduce that $tW_{\frac{1}{t}} \to 0$ as $t \to 0$.

(b) Since $W_t \sim \mathcal{N}(0, t)$ we can deduce $tW_{\frac{1}{t}} \sim \mathcal{N}(0, t), t > 0$. Therefore,

$$
B_{t+s} - B_t = (t + s)W_{\frac{1}{t+s}} - tW_{\frac{1}{t}}
$$

with mean

$$
\mathbb{E}(B_{t+s} - B_t) = \mathbb{E}\left((t + s)W_{\frac{1}{t+s}}\right) - \mathbb{E}\left(tW_{\frac{1}{t}}\right) = 0
$$

and variance

$$
\text{Var}(B_{t+s} - B_t) = \text{Var}\left((t + s)W_{\frac{1}{t+s}}\right) + \text{Var}\left(tW_{\frac{1}{t}}\right) - 2\text{Cov}\left((t + s)W_{\frac{1}{t+s}}, tW_{\frac{1}{t}}\right) \\
= t + s + t - 2\min\{t + s, s\} \\
= s.
$$

Since the sum of two normal distributions is also a normal distribution, so $B_{t+s} - B_t \sim \mathcal{N}(0, s)$. 

---

(c) Finally, to show that $B_{t+s} - B_t \perp B_t$, we note that since $W_t$ has the independent increment property, so

$$
\mathbb{E} \left[ (B_{t+s} - B_t) B_t \right] = \mathbb{E} (B_{t+s} B_t) - \mathbb{E} (B_t^2) \\
= \text{Cov} (B_{t+s}, B_t) - \mathbb{E} (B_t^2) \\
= \min\{t + s, t\} - t \\
= 0.
$$

Since $B_t \sim \mathcal{N}(0, t)$ and $B_{t+s} - B_t \sim \mathcal{N}(0, s)$ and the joint distribution of $B_t$ and $B_{t+s} - B_t$ is a bivariate normal (see Problem 2.2.1.5, page 58), then if $\text{Cov} (B_{t+s} - B_t, B_t) = 0$ so $B_{t+s} - B_t \perp B_t$.

From the results of (a)-(c) we have shown that $B_t = c W_t$, $c \neq 0$ is also a standard Wiener process.

\[\Box\]

9. **Time Inversion.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Show that under time inversion,

$$
B_t = \begin{cases} 
0 & \text{if } t = 0 \\
tW_{\frac{1}{t}} & \text{if } t \neq 0
\end{cases}
$$

is also a standard Wiener process.

**Solution:**

(a) $B_0 = 0$ and it is clear that $B_t$ has continuous sample paths for $t > 0$. From continuity at $t = 0$ we can deduce that $tW_{\frac{1}{t}} \to 0$ as $t \to 0$.

(b) Since $W_t \sim \mathcal{N}(0, t)$ we can deduce $tW_{\frac{1}{t}} \sim \mathcal{N}(0, t), t > 0$. Therefore,

$$
B_{t+s} - B_t = (t + s)W_{\frac{1}{t+s}} - tW_{\frac{1}{t}}
$$

with mean

$$
\mathbb{E}(B_{t+s} - B_t) = \mathbb{E}\left((t + s)W_{\frac{1}{t+s}}\right) - \mathbb{E}\left(tW_{\frac{1}{t}}\right) = 0
$$

and variance

$$
\text{Var}(B_{t+s} - B_t) = \text{Var}\left((t + s)W_{\frac{1}{t+s}}\right) + \text{Var}\left(tW_{\frac{1}{t}}\right) - 2\text{Cov}\left((t + s)W_{\frac{1}{t+s}}, tW_{\frac{1}{t}}\right) \\
= t + s + t - 2\min\{t + s, s\} \\
= s.
$$

Since the sum of two normal distributions is also a normal distribution, so $B_{t+s} - B_t \sim \mathcal{N}(0, s)$. 

---

(c) Finally, to show that $B_{t+s} - B_t \perp B_t$, we note that since $W_t$ has the independent increment property, so

$$
\mathbb{E} \left[ (B_{t+s} - B_t) B_t \right] = \mathbb{E} (B_{t+s} B_t) - \mathbb{E} (B_t^2) \\
= \text{Cov} (B_{t+s}, B_t) - \mathbb{E} (B_t^2) \\
= \min\{t + s, t\} - t \\
= 0.
$$

Since $B_t \sim \mathcal{N}(0, t)$ and $B_{t+s} - B_t \sim \mathcal{N}(0, s)$ and the joint distribution of $B_t$ and $B_{t+s} - B_t$ is a bivariate normal (see Problem 2.2.1.5, page 58), then if $\text{Cov} (B_{t+s} - B_t, B_t) = 0$ so $B_{t+s} - B_t \perp B_t$.

From the results of (a)-(c) we have shown that $B_t = c W_t$, $c \neq 0$ is also a standard Wiener process.

\[\Box\]
2.2.1 Basic Properties

(c) To show that \( B_{t+s} - B_t \perp B_t \) we note that since \( W_t \) has the independent increment property, so

\[
\mathbb{E}((B_{t+s} - B_t)B_t) = \mathbb{E}(B_{t+s}B_t) - \mathbb{E}(B_t^2) = \text{Cov}(B_{t+s}, B_t) - \mathbb{E}(B_t^2) = \min(t + s, t) - t = 0.
\]

Since \( B_t \sim \mathcal{N}(0, t) \) and \( B_{t+s} - B_t \sim \mathcal{N}(0, s) \) and the joint distribution of \( B_t \) and \( B_{t+s} - B_t \) is a bivariate normal (see Problem 2.2.1.5, page 58), then if \( \text{Cov}(B_{t+s} - B_t, B_t) = 0 \) so \( B_{t+s} - B_t \perp B_t \).

From the results of (a)–(c) we have shown that \( B_t = \begin{cases} 0 & \text{if } t = 0 \\ tW_1 & \text{if } t \neq 0 \end{cases} \) is also a standard Wiener process.

\[\square\]

10. Time Reversal. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Show that under time reversal, \( B_t = W_1 - W_{1-t} \) is also a standard Wiener process.

Solution:

(a) \( B_0 = W_1 - W_1 = 0 \) and it is clear that \( B_t \) has continuous sample paths for \( t \geq 0 \).

(b) Since \( W_t \) is a standard Wiener process, we have \( W_1 \sim \mathcal{N}(0, 1) \) and \( W_{1-t} \sim \mathcal{N}(0, 1-t) \).

Therefore,
\[
B_{t+s} - B_t = W_1 - W_{1-(t+s)} - W_1 + W_{1-t} = W_{1-t} - W_{1-(t+s)}
\]

with mean
\[
\mathbb{E}(B_{t+s} - B_t) = \mathbb{E}(W_{1-t}) - \mathbb{E}(W_{1-(t+s)}) = 0
\]

and variance
\[
\text{Var}(B_{t+s} - B_t) = \text{Var}(W_{1-t}) + \text{Var}(W_{1-(t+s)}) - 2\text{Cov}(W_{1-t}, W_{1-(t+s)})
\]
\[
= 1 - t + 1 - (t + s) - 2\min\{1 - t, 1 - (t + s)\} = s.
\]

Since the sum of two normal distributions is also a normal distribution, so \( B_{t+s} - B_t \sim \mathcal{N}(0, s) \).

(c) To show that \( B_{t+s} - B_t \perp B_t \) we note that since \( W_t \) has the independent increment property, so

\[
\mathbb{E}((B_{t+s} - B_t)B_t) = \mathbb{E}(B_{t+s}B_t) - \mathbb{E}(B_t^2) = \text{Cov}(B_{t+s}, B_t) - \mathbb{E}(B_t^2)
\]
12. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let \( \{W_t : t \geq 0\}, i = 1, 2, \ldots, n \) be a sequence of independent standard Wiener processes. Show that the vector \( W_t = \left( W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)} \right)^T \) is an \( n \)-dimensional standard Wiener process with the following properties:

(a) \( W_0 = 0 \) and \( W_t \) is a vector of continuous sample paths;
(b) for each \( t > 0 \) and \( s > 0, W_{t+s} - W_t \sim \mathcal{N}(0, s) \) where \( I \) is an \( n \times n \) identity matrix;
(c) for each \( t > 0 \) and \( s > 0, W_{t+s} - W_t \perp W_t \).

Solution:

(a) Since \( W_0^{(i)} = 0 \) for all \( i = 1, 2, \ldots, n \) therefore \( W_0 = \left( W_0^{(1)}, W_0^{(2)}, \ldots, W_0^{(n)} \right)^T = 0 \)

where \( 0 = (0, 0, \ldots, 0)^T \) is an \( n \)-vector of zeroes. In addition, \( W_t \) is a vector of continuous sample paths due to the fact that \( W_t^{(i)}, t \geq 0 \) has continuous sample paths.
(b) For \( s, t > 0 \) and for \( i, j = 1, 2, \ldots, n \) we have

\[
\mathbb{E}(W_{t+s}^{(i)} - W_t^{(i)}) = 0, \quad i = 1, 2, \ldots, n
\]

and

\[
\mathbb{E}\left[ (W_{t+s}^{(i)} - W_t^{(i)}) (W_{t+s}^{(j)} - W_t^{(j)}) \right] = \begin{cases} s & i = j \\ 0 & i \neq j. \end{cases}
\]

Therefore, \( W_{t+s} - W_t \sim \mathcal{N}(0, s I) \) where \( I \) is an \( n \times n \) identity matrix.
(c) For \( s, t > 0 \) and since \( W_{t+s}^{(i)} - W_t^{(i)} \perp W_t^{(i)} \), \( W_{t+s}^{(i)} \perp W_t^{(j)}, i \neq j, i, j = 1, 2, \ldots, n \) we can deduce that \( W_{t+s}^{(i)} - W_t^{(i)} \perp W_t^{(j)} \). Therefore, \( W_{t+s} - W_t \perp W_t \).

From the results of (a)–(c) we have shown that \( W_t = \left( W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)} \right)^T \) is an \( n \)-dimensional standard Wiener process.

\[ \square \]

12. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Show that the pair of random variables \( \left( W_t, \int_0^t W_s \, ds \right) \) has the following covariance matrix

\[
\Sigma = \begin{bmatrix}
\frac{1}{2} t^2 & \frac{1}{2} t^2 \\
\frac{1}{2} t^2 & \frac{1}{3} t^3
\end{bmatrix}
\]

with correlation coefficient \( \frac{\sqrt{3}}{2} \).
2.2.1 Basic Properties

Solution: By definition, the covariance matrix for the pair \( (W_t, \int_0^t W_s \, ds) \) is

\[
\Sigma = \begin{bmatrix}
\text{Var}(W_t) & \text{Cov}(W_t, \int_0^t W_s \, ds) \\
\text{Cov}(W_t, \int_0^t W_s \, ds) & \text{Var}(\int_0^t W_s \, ds)
\end{bmatrix}.
\]

Given that \( W_t \sim \mathcal{N}(0, t) \) we know \( \text{Var}(W_t) = t \).

For the case \( \text{Cov}(W_t, \int_0^t W_s \, ds) \) we can write

\[
\text{Cov}(W_t, \int_0^t W_s \, ds) = \mathbb{E} \left( W_t \int_0^t W_s \, ds \right) - \mathbb{E}(W_t) \mathbb{E} \left( \int_0^t W_s \, ds \right)
\]

\[
= \mathbb{E} \left( W_t \int_0^t W_s \, ds \right)
\]

\[
= \mathbb{E} \left( \int_0^t W_t W_s \, ds \right)
\]

\[
= \int_0^t \mathbb{E}(W_t W_s) \, ds
\]

\[
= \int_0^t \mathbb{E} \left[ W_t(W_t - W_s) + W_s^2 \right] \, ds.
\]

From the independent increment property of a Wiener process we have

\[
\text{Cov}(W_t, \int_0^t W_s \, ds) = \int_0^t \mathbb{E}(W_t) \mathbb{E}(W_t - W_s) \, ds + \int_0^t \mathbb{E}(W_s^2) \, ds
\]

\[
= \int_0^t \mathbb{E}(W_s^2) \, ds
\]

\[
= \int_0^t s \, ds
\]

\[
= \frac{t^2}{2}.
\]

Finally,

\[
\text{Var} \left( \int_0^t W_s \, ds \right) = \mathbb{E} \left[ \left( \int_0^t W_s \, ds \right)^2 \right] - \mathbb{E} \left( \int_0^t W_s \, ds \right)^2
\]

\[
= \mathbb{E} \left[ \left( \int_0^t W_s \, ds \right) \left( \int_0^t W_u \, du \right) \right] - \mathbb{E}(W_t)^2
\]
2.2.1 Basic Properties

\[ 
= \mathbb{E} \left[ \int_{s=0}^{t=0} \int_{u=0}^{w=t} W_s W_u \, du \, ds \right] \\
= \int_{s=0}^{t=0} \int_{u=0}^{w=t} \mathbb{E}(W_s W_u) \, du \, ds. 
\]

From Problem 2.2.1.4 (page 57) we have

\[ 
\mathbb{E}(W_s W_u) = \min\{s, u\} 
\]

and hence

\[ 
\text{Var} \left( \int_{0}^{t} W_s \, ds \right) = \int_{s=0}^{t} \int_{u=0}^{u} \min\{s, u\} \, du \, ds \\
= \int_{s=0}^{t} \int_{u=0}^{u} u \, du \, ds + \int_{s=0}^{t} \int_{u=s}^{u} s \, du \, ds \\
= \int_{0}^{t} \frac{1}{2} s^2 \, ds + \int_{0}^{t} s(t-s) \, ds \\
= \frac{1}{3} t^3. 
\]

Therefore,

\[ 
\Sigma = \begin{bmatrix} t & \frac{1}{2} t^2 \\ \frac{1}{2} t^2 & \frac{1}{3} t^3 \end{bmatrix} 
\]

with correlation coefficient

\[ 
\rho = \frac{\text{Cov} \left( W_t, \int_{0}^{t} W_s \, ds \right)}{\sqrt{\text{Var}(W_t) \text{Var} \left( \int_{0}^{t} W_s \, ds \right)}} = \frac{t^2/2}{t^3/3} = \frac{\sqrt{3}}{2}. 
\]

13. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Show that the covariance of \(\int_{0}^{s} W_u \, du\) and \(\int_{0}^{t} W_v \, dv, s, t > 0\) is

\[ 
\text{Cov} \left( \int_{0}^{s} W_u \, du, \int_{0}^{t} W_v \, dv \right) = \frac{1}{3} \min\{s^3, t^3\} + \frac{1}{2} |t-s| \min\{s^2, t^2\} 
\]

with correlation coefficient

\[ 
\sqrt{\min\{s^3, t^3\} / \max\{s^3, t^3\}} + \frac{3}{2} |t-s| \sqrt{\min\{s, t\} / \max\{s, t\}}. 
\]
Solution: By definition

\[
\text{Cov} \left( \int_0^s W_u \, du, \int_0^t W_v \, dv \right) = \mathbb{E} \left[ \left( \int_0^s W_u \, du \right) \left( \int_0^t W_v \, dv \right) \right] - \mathbb{E} \left[ \int_0^s W_u \, du \right] \mathbb{E} \left[ \int_0^t W_v \, dv \right]
\]

\[
= \int_0^s \int_0^t \mathbb{E} (W_u W_v) \, du \, dv - \left[ \int_0^s \mathbb{E} (W_u) \, du \right] \left[ \int_0^t \mathbb{E} (W_v) \, dv \right]
\]

\[
= \int_0^s \int_0^t \text{min} \{u, v\} \, du \, dv
\]

since \( \mathbb{E}(W_u W_v) = \text{min} \{u, v\} \) and \( \mathbb{E}(W_u) = \mathbb{E}(W_v) = 0 \).

For the case \( s \leq t \), using the results of Problem 2.2.1.12 (page 64),

\[
\text{Cov} \left( \int_0^s W_u \, du, \int_0^t W_v \, dv \right) = \int_0^s \int_0^t \text{min} \{u, v\} \, du \, dv
\]

\[
= \int_0^s \int_0^t \text{min} \{u, v\} \, du \, dv + \int_0^t \int_s^t \text{min} \{u, v\} \, du \, dv
\]

\[
= \frac{1}{3} s^3 + \int_0^s \int_s^t v \, du \, dv
\]

\[
= \frac{1}{3} s^3 + \int_0^t v(t - s) \, dv
\]

\[
= \frac{1}{3} s^3 + \frac{1}{2} (t - s) t^2.
\]

Following the same steps as described above, for the case \( s > t \) we can also show

\[
\text{Cov} \left( \int_0^t W_u \, du, \int_0^s W_v \, dv \right) = \frac{1}{3} t^3 + \frac{1}{2} (s - t)^2.
\]

Thus,

\[
\text{Cov} \left( \int_0^s W_u \, du, \int_0^t W_v \, dv \right) = \frac{1}{3} \text{min} \{s^3, t^3\} + \frac{1}{2} |t - s| \text{min} \{s^2, t^2\}.
\]

From the definition of the correlation coefficient between \( \int_0^s W_u \, du \) and \( \int_0^t W_v \, dv \),

\[
\rho = \frac{\text{Cov} \left( \int_0^s W_u \, du, \int_0^t W_v \, dv \right)}{\sqrt{\text{Var} \left( \int_0^s W_u \, du \right) \text{Var} \left( \int_0^t W_v \, dv \right)}}
\]
2.2.2 Markov Property

\[
\begin{align*}
\rho &= \frac{1}{3} \min\{s^3, t^3\} + \frac{1}{2} |t-s| \min\{s^2, t^2\} \\
&= \frac{1}{3} \sqrt{s^3 t^3} + \frac{3}{2} \frac{|t-s| \min\{s^2, t^2\}}{\sqrt{s^3 t^3}}.
\end{align*}
\]

For \(s \leq t\) we have

\[
\rho = \sqrt{\frac{s^3}{t^3} + \frac{3}{2} (t-s)} \frac{s}{t^3}
\]

and for \(s > t\)

\[
\rho = \sqrt{\frac{t^3}{s^3} + \frac{3}{2} (s-t)} \frac{t}{s^3}.
\]

Therefore, we can deduce

\[
\rho = \sqrt{\frac{\min\{s^3, t^3\}}{\max\{s^3, t^3\}} + \frac{3}{2} |t-s| \sqrt{\frac{\min\{s, t\}}{\max\{s^3, t^3\}}}.
\]

\[\square\]

2.2.2 Markov Property

1. The Markov Property of a Standard Wiener Process. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process with respect to the filtration \(\mathcal{F}_t, t \geq 0\). Show that if \(f\) is a continuous function then there exists another continuous function \(g\) such that

\[
\mathbb{E}[f(W_t) | \mathcal{F}_u] = g(W_u)
\]

for \(0 \leq u \leq t\).

**Solution:** For \(0 \leq u \leq t\) we can write

\[
\mathbb{E}[f(W_t) | \mathcal{F}_u] = \mathbb{E}[f(W_t - W_u + W_u) | \mathcal{F}_u].
\]

Since \(W_t - W_u \perp \mathcal{F}_u\) and \(W_u\) is \(\mathcal{F}_u\) measurable, by setting \(W_u = x\) where \(x\) is a constant value

\[
\mathbb{E}[f(W_t - W_u + x) | \mathcal{F}_u] = \mathbb{E}[f(W_t - W_u + x)].
\]

Because \(W_t - W_u \sim \mathcal{N}(0, t-u)\) we can write \(\mathbb{E}[f(W_t - W_u + x)]\) as

\[
\mathbb{E}[f(W_t - W_u + x)] = \frac{1}{\sqrt{2\pi(t-u)}} \int_{-\infty}^{\infty} f(w+x)e^{-\frac{w^2}{2(t-u)}} dw.
\]
2.2.2 Markov Property

By setting $\tau = t - u$ and $y = w + x$, we can rewrite $\mathbb{E}[f(W_t - W_u + W_u)]$ as

$$\mathbb{E}[f(W_t - W_u + W_u)] = \frac{1}{\sqrt{2\pi \tau}} \int_{-\infty}^{\infty} f(y) e^{-\frac{(y-x)^2}{2\tau}} dy = \int_{-\infty}^{\infty} f(y) p(\tau, W_u, y) dy$$

where the transition density

$$p(\tau, W_u, y) = \frac{1}{\sqrt{2\pi \tau}} e^{-\frac{(y-W_u)^2}{2\tau}}$$

is the density of $Y \sim \mathcal{N}(W_u, \tau)$. Since the only information from the filtration $\mathcal{F}_u$ is $W_u$, therefore

$$\mathbb{E}[f(W_t)|\mathcal{F}_u] = g(W_u)$$

where

$$g(W_u) = \int_{-\infty}^{\infty} f(y) p(\tau, W_u, y) dy.$$

\[ \square \]

2. The Markov Property of a Wiener Process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process with respect to the filtration $\mathcal{F}_t$, $t \geq 0$. By considering the Wiener process

$$\hat{W}_t = a + bt + cW_t, \quad a, b \in \mathbb{R}, \quad c > 0$$

show that if $f$ is a continuous function then there exists another continuous function $g$ such that

$$\mathbb{E}[f(\hat{W}_t)|\mathcal{F}_u] = g(\hat{W}_u)$$

for $0 \leq u \leq t$.

Solution: For $0 \leq u \leq t$ we can write

$$\mathbb{E} \left[ f(\hat{W}_t) | \mathcal{F}_u \right] = \mathbb{E} \left[ f(\hat{W}_t - \hat{W}_u + \hat{W}_u) | \mathcal{F}_u \right].$$

Since $W_t - W_u \perp \mathcal{F}_u$, we can deduce that $\hat{W}_t - \hat{W}_u \perp \mathcal{F}_u$. In addition, because $W_u$ is $\mathcal{F}_u$ measurable, hence $\hat{W}_u$ is also $\mathcal{F}_u$ measurable. By setting $\hat{W}_u = x$ where $x$ is a constant value,

$$\mathbb{E} \left[ f(\hat{W}_t - \hat{W}_u + \hat{W}_u) | \mathcal{F}_u \right] = \mathbb{E} \left[ f(\hat{W}_t - \hat{W}_u + x) \right].$$
2.2.2 Markov Property

Taking note that \( \hat{W}_t - \hat{W}_u \sim \mathcal{N}(b(t - u), c^2(t - u)) \), we can write

\[
\mathbb{E}\left[ f(\hat{W}_t - \hat{W}_u + x) \right] = \frac{1}{c\sqrt{2\pi(t-u)}} \int_{-\infty}^{\infty} f(w + x)e^{-\frac{1}{2}\left(\frac{(w-b(t-u))^2}{c^2(t-u)}\right)} \, dw.
\]

By setting \( \tau = t - u \) and \( y = w + x \), we can rewrite

\[
\mathbb{E}\left[ f(\hat{W}_t - \hat{W}_u + \hat{W}_u) \right] = \mathbb{E}\left[ f(\hat{W}_t - \hat{W}_u + \hat{W}_u) \right] = \frac{1}{c\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} f(y)p(\tau, \hat{W}_u, y) \, dy
\]

where the transition density

\[
p(\tau, W_u, y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{1}{2}\left(\frac{(y-b\tau-\hat{W}_u)^2}{c^2\tau}\right)}
\]

is the density of \( Y \sim \mathcal{N}(b\tau + \hat{W}_u, c^2\tau) \). Since the only information from the filtration \( \mathcal{F}_u \) is \( W_u \), therefore

\[
\mathbb{E}\left[ f(\hat{W}_t) \bigg| \mathcal{F}_u \right] = g(\hat{W}_u)
\]

where

\[
g(\hat{W}_u) = \int_{-\infty}^{\infty} f(y)p(\tau, \hat{W}_u, y) \, dy.
\]

\( \square \)

3. The Markov Property of a Geometric Brownian Motion. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process with respect to the filtration \( \mathcal{F}_t \), \( t \geq 0 \). By considering the geometric Wiener process

\[
S_t = S_0e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t}, \quad \mu \in \mathbb{R}, \quad \sigma > 0
\]

where \( S_0 > 0 \) show that if \( f \) is a continuous function then there exists another continuous function \( g \) such that

\[
\mathbb{E}[f(S_t) | \mathcal{F}_u] = g(S_u)
\]

for \( 0 \leq u \leq t \).

**Solution:** For \( 0 \leq u \leq t \) we can write

\[
\mathbb{E}[f(S_t) | \mathcal{F}_u] = \mathbb{E}\left[ f\left( \frac{S_t}{S_u} \cdot S_u \right) \bigg| \mathcal{F}_u \right]
\]
2.2.3 Martingale Property

where
\[
\log \left( \frac{S_t}{S_u} \right) \sim \mathcal{N} \left( \left( \mu - \frac{1}{2} \sigma^2 \right) (t - u), \sigma^2 (t - u) \right).
\]

Since \( S_t / S_u \perp \mathcal{F}_u \) and \( S_u \) is \( \mathcal{F}_u \) measurable, by setting \( S_u = x \) where \( x \) is a constant value
\[
\mathbb{E} \left[ f \left( \frac{S_t}{S_u} \cdot S_u \right) \bigg| \mathcal{F}_u \right] = \mathbb{E} \left[ f \left( \frac{S_t}{S_u} \cdot x \right) \right].
\]

By setting \( \nu = \mu - \frac{1}{2} \sigma^2 \) so that \( S_t / S_u \sim \log\mathcal{N}(\nu(t - u), \sigma^2 (t - u)) \), we can write
\[
\mathbb{E} \left[ f \left( \frac{S_t}{S_u} \cdot x \right) \right] = \mathbb{E} \left[ f \left( \frac{S_t}{S_u} \right) \right] = \mathbb{E} \left[ f \left( \frac{S_t}{S_u} \right) \right] = \frac{1}{\sigma \sqrt{2\pi (t - u)}} \int_{-\infty}^{\infty} f(w \cdot x) e^{-\frac{1}{2} \left( \frac{\log w - \nu (t - u)}{\sigma^2 (t - u)} \right)^2} dw.
\]

By setting \( \tau = t - u \) and \( y = w \cdot x \), we can rewrite \( \mathbb{E} \left[ f \left( \frac{S_t}{S_u} \cdot x \right) \right] = \mathbb{E} \left[ f \left( \frac{S_t}{S_u} \right) \right] \) as
\[
\mathbb{E} \left[ f \left( \frac{S_t}{S_u} \cdot x \right) \right] = \frac{1}{\sigma y \sqrt{2\pi \tau}} \int_{-\infty}^{\infty} f(y) e^{-\frac{1}{2} \left( \frac{\log y - \frac{\nu}{\sigma^2 (t - u)}}{\sigma^2 (t - u)} \right)^2} dy
\]
\[
= \int_{-\infty}^{\infty} f(y) p(\tau, W_u, y) dy
\]

where the transition density
\[
p(\tau, W_u, y) = \frac{1}{\sigma y \sqrt{2\pi \tau}} e^{-\frac{1}{2} \left( \frac{\log y - \log S_u + y \tau}{\sigma^2 \tau} \right)^2}
\]
is the density of \( Y \sim \log\mathcal{N}(\log S_u + y \tau, \sigma^2 \tau) \). Since the only information from the filtration \( \mathcal{F}_u \) is \( S_u \), therefore
\[
\mathbb{E}[f(S_t) | \mathcal{F}_u] = g(S_u)
\]
such that
\[
g(S_u) = \int_{-\infty}^{\infty} f(y) p(\tau, W_u, y) dy.
\]

\[\square\]

2.2.3 Martingale Property

1. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Show that \( W_t \) is a martingale.

Solution: Given \( W_t \sim \mathcal{N}(0, t) \).
2.2.3 Martingale Property

(a) For $s \leq t$ and since $W_t - W_s \perp \mathcal{F}_s$, we have
\[ E(W_t|\mathcal{F}_s) = E(W_t - W_s + W_s|\mathcal{F}_s) = E(W_t - W_s|\mathcal{F}_s) + E(W_s|\mathcal{F}_s) = W_s. \]

(b) Since $W_t \sim \mathcal{N}(0,t)$, $|W_s|$ follows a folded normal distribution such that $|W_s| \sim \mathcal{N}_f(0,t)$. From Problem 1.2.2.11 (page 22), we can deduce $E(|W_t|) = \sqrt{2t/\pi} < \infty$. In contrast, we can also utilise Hölder’s inequality (see Problem 1.2.3.2, page 41) to deduce that $E(|W_t|) \leq \sqrt{E(W_t^2)} = \sqrt{t} < \infty$.

(c) $W_t$ is clearly $\mathcal{F}_t$-adapted.

From the results of (a)–(c) we have shown that $W_t$ is a martingale.

\[ \square \]

2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Show that $X_t = W_t^2 - t$ is a martingale.

**Solution:** Given $W_t \sim \mathcal{N}(0,t)$

(a) For $s \leq t$ and since $W_t - W_s \perp \mathcal{F}_s$, we have
\[ E(W_t^2 - t|\mathcal{F}_s) = E[(W_t - W_s + W_s)^2|\mathcal{F}_s] - t \]
\[ = E[(W_t - W_s)^2|\mathcal{F}_s] + 2E[W_s(W_t - W_s)|\mathcal{F}_s] + E(W_s^2|\mathcal{F}_s) - t \]
\[ = t - s + 0 + W_s^2 - t \]
\[ = W_s^2 - s. \]

(b) Since $|X_t| = |W_t^2 - t| \leq W_t^2 + t$ we can therefore write
\[ E\left(\left|W_t^2 - t\right|\right) \leq E(W_t^2 + t) = E(W_t^2) + t = 2t < \infty. \]

(c) Since $X_t = W_t^2 - t$ is a function of $W_t$, hence it is $\mathcal{F}_t$-adapted.

From the results of (a)–(c) we have shown that $X_t = W_t^2 - t$ is a martingale.

\[ \square \]

3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. For $\lambda \in \mathbb{R}$ show that $X_t = e^{iW_t - \frac{1}{2}\lambda^2t}$ is a martingale.

**Solution:** Given $W_t \sim \mathcal{N}(0,t)$ we can write $\log X_t = \lambda W_t - \frac{1}{2}\lambda^2t \sim \mathcal{N}\left(-\frac{1}{2}\lambda^2t, \lambda^2t\right)$ and hence $e^{X_t} \sim \log\mathcal{N}\left(-\frac{1}{2}\lambda^2t, \lambda^2t\right)$.

(a) For $s \leq t$ and since $W_t - W_s \perp \mathcal{F}_s$, we have
\[ E\left(e^{iW_t - \frac{1}{2}\lambda^2t}|\mathcal{F}_s\right) = e^{-\frac{1}{2}\lambda^2t}E\left(e^{iW_t}|\mathcal{F}_s\right) \]
\[ = e^{-\frac{1}{2}\lambda^2t}E\left[e^{(W_t-W_s)+iW_s}|\mathcal{F}_s\right] \]
\[ = e^{-\frac{1}{2}\lambda^2t}E\left[e^{iW_s}|\mathcal{F}_s\right] = e^{-\frac{1}{2}\lambda^2t}. \]
2.2.3 Martingale Property

\[ e^{-\frac{1}{2} \lambda^2 t} \mathbb{E} \left[ e^{\lambda (W_t - W_s)} \mid \mathcal{F}_s \right] \mathbb{E} \left[ e^{\lambda W_s} \mid \mathcal{F}_s \right] \]

\[ = e^{-\frac{1}{2} \lambda^2 t} \cdot e^{\frac{1}{2} \lambda^2 (t-s)} \cdot e^{\lambda W_s} \]

\[ = e^{\lambda W_s - \frac{1}{2} \lambda^2 s}. \]

(b) By setting \( |X_t| = \left| e^{\lambda (W_t - W_s)} \right| = e^{\lambda W_t - \frac{1}{2} \lambda^2 t}, \)

\[ \mathbb{E}(X_t) = \mathbb{E} \left( e^{\lambda W_t - \frac{1}{2} \lambda^2 t} \right) = e^{-\frac{1}{2} \lambda^2 t} \mathbb{E}(e^{\lambda W_s}) = e^{-\frac{1}{2} \lambda^2 t} \cdot e^{\frac{1}{2} \lambda^2 t} = 1 < \infty. \]

(c) Since \( X_t \) is a function of \( W_t \), hence it is \( \mathcal{F}_t \)-adapted.

From the results of (a)–(c) we have shown that \( X_t = e^{\lambda W_t - \frac{1}{2} \lambda^2 t} \) is a martingale.

\[ \square \]

4. Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Show that \( X_t = W_t^3 - 3tW_t \) is a martingale.

Solution:

(a) For \( s \leq t \) and since \( W_t - W_s \perp \perp \mathcal{F}_s \), we have

\[ \mathbb{E} \left( X_t \mid \mathcal{F}_s \right) = \mathbb{E} \left( W_t^3 - 3tW_t \mid \mathcal{F}_s \right) \]

\[ = \mathbb{E} \left[ W_t \left( (W_t - W_s + W_s)^3 - 3t \right) \mid \mathcal{F}_s \right] \]

\[ = \mathbb{E} \left[ W_t (W_t - W_s)^3 \mid \mathcal{F}_s \right] + 2W_s \mathbb{E} \left[ W_t (W_t - W_s) \mid \mathcal{F}_s \right] \]

\[ + W_s^2 \mathbb{E} \left( W_t \mid \mathcal{F}_s \right) - 3t \mathbb{E} \left( W_t \mid \mathcal{F}_s \right) \]

\[ = \mathbb{E} \left[ (W_t - W_s + W_s) (W_t - W_s)^2 \mid \mathcal{F}_s \right] + W_s^3 - 3tW_s \]

\[ = \mathbb{E} \left[ (W_t - W_s)^3 \mid \mathcal{F}_s \right] + \mathbb{E} \left[ W_s (W_t - W_s)^2 \mid \mathcal{F}_s \right] \]

\[ + 2W_s \mathbb{E} \left[ (W_t - W_s)^2 \mid \mathcal{F}_s \right] + 2W_s \mathbb{E} \left[ W_s (W_t - W_s) \mid \mathcal{F}_s \right] + W_s^3 - 3tW_s \]

\[ = W_s (t-s) + 2W_s (t-s) + W_s^3 - 3tW_s \]

\[ = W_s^3 - 3sW_s \]

where, from Problem 2.2.1.5 (page 58), we can deduce that the moment generating function of \( W_t \) is \( M_{W_t}(\theta) = e^{\frac{1}{2} \theta^2 t^2} \) and \( \mathbb{E}(W_t^3) = \frac{d^3}{d\theta^3} M_{W_t}(\theta) \bigg|_{\theta=0} = 0. \)
(b) Since $|X_t| = |W_t^3 - 3tW_t| ≤ |W_t^3| + 3t|W_t|$ and from Hölder’s inequality (see Problem 1.2.3.2, page 41),

$$E(|X_t|) ≤ \sqrt{E \left( |W_t^3|^2 \right) E \left( |W_t|^2 \right) + 3tE \left( |W_t| \right)}$$

Since $W_t/t \sim \chi (1)$ we have $E(W_t) = 3t$ and utilising Hölder’s inequality again, we have $E(|X_t|) ≤ \sqrt{t} < \infty$. Therefore, $E(|X_t|) ≤ t\sqrt{3} + 3t\sqrt{t} < \infty$.

(c) Since $X_t$ is a function of $W_t$, hence it is $\mathcal{F}_t$-adapted.

From the results of (a)–(c) we have shown that $X_t = W_t^3 - 3tW_t$ is a martingale.

\[\square\]

5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t ≥ 0\}$ be a standard Wiener process. For $\lambda \in \mathbb{R}$ show that the following hyperbolic processes:

$$X_t = e^{-\frac{1}{2} - \frac{1}{2}} \cosh(\lambda W_t)$$

$$Y_t = e^{-\frac{1}{2} - \frac{1}{2}} \sinh(\lambda W_t)$$

are martingales.

Solution: By definition we can write

$$X_t = e^{-\frac{1}{2} - \frac{1}{2}} \cosh(\lambda W_t) = \frac{1}{2} \left( e^{\lambda W_t - \frac{1}{2} t \lambda} + e^{-\lambda W_t - \frac{1}{2} t \lambda} \right).$$

Since for $\lambda \in \mathbb{R}$, $X_t^{(1)} = \frac{1}{2} e^{\lambda W_t - \frac{1}{2} t \lambda}$ and $X_t^{(2)} = \frac{1}{2} e^{-\lambda W_t - \frac{1}{2} t \lambda}$ are martingales we have the following properties:

(a) For $s \leq t$, $E \left( X_s^{(1)} + X_t^{(2)} \right| \mathcal{F}_s) = X_s^{(1)} + X_t^{(2)}$.

(b) Because $E \left( |X_t^{(1)}| \right) < \infty$ and $E \left( |X_t^{(2)}| \right) < \infty$, so $E \left( \left| \left( X_t^{(1)} + X_t^{(2)} \right) \right| \right) < \infty$.

(c) Since $X_t$ is a function of $W_t$, hence it is $\mathcal{F}_t$-adapted.

From the results of (a)–(c) we have shown that $X_t = e^{-\frac{1}{2} - \frac{1}{2}} \cosh(\lambda W_t)$ is a martingale.

For $Y_t = e^{-\frac{1}{2} - \frac{1}{2}} \sinh(\lambda W_t)$ we note that

$$Y_t = e^{-\frac{1}{2} - \frac{1}{2}} \sinh(\lambda W_t) = \frac{1}{2} \left( e^{\lambda W_t - \frac{1}{2} t \lambda} - e^{-\lambda W_t - \frac{1}{2} t \lambda} \right)$$

and similar steps can be applied to show that $Y_t$ is also a martingale.

\[\square\]

6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Show that the sample paths of a standard Wiener process $\{W_t : t ≥ 0\}$ are continuous but not differentiable.
2.2.3 Martingale Property

Solution: The path $W_t \sim \mathcal{N}(0, t)$ is continuous in probability if and only if, for every $\delta > 0$ and $t \geq 0$,

$$\lim_{\Delta t \to 0} P(|W_{t+\Delta t} - W_t| \geq \delta) = 0.$$ 

Given that $W_t$ is a martingale and from Chebyshev’s inequality (see Problem 1.2.2.19, page 40),

$$P(|W_{t+\Delta t} - W_t| \geq \delta) = P(|W_{t+\Delta t} - \mathbb{E}(W_{t+\Delta t} | \mathcal{F}_t)| \geq \delta) \leq \frac{\text{Var}(W_{t+\Delta t} | \mathcal{F}_t)}{\delta^2} = \frac{\text{Var}(W_{t+\Delta t} - W_t + W_t | \mathcal{F}_t)}{\delta^2} = \frac{\text{Var}(W_{t+\Delta t} - W_t)}{\delta^2} + \frac{\text{Var}(W_t | \mathcal{F}_t)}{\delta^2} = \frac{\Delta t}{\delta^2}$$

since $W_{t+\Delta t} - W_t \perp \mathcal{F}_t$ and $W_t$ is $\mathcal{F}_t$ measurable and hence $\text{Var}(W_t | \mathcal{F}_t) = 0$. Taking the limit $\Delta t \to 0$, we have

$$P(|W_{t+\Delta t} - W_t| \geq \delta) \to 0$$

and therefore we conclude that the sample path is continuous. By setting

$$\Delta W_t = W_{t+\Delta t} - W_t = \phi \sqrt{\Delta t}$$

where $\phi \sim \mathcal{N}(0, 1)$ and taking limit $\Delta t \to 0$, we have

$$\lim_{\Delta t \to 0} \frac{\Delta W_t}{\Delta t} = \lim_{\Delta t \to 0} \frac{\phi}{\sqrt{\Delta t}} = \pm \infty$$

dependning on the sign of $\phi$. Therefore, $W_t$ is not differentiable.

$\square$

7. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process with respect to the filtration $\mathcal{F}_t$, $t \geq 0$.

Show that if $\varphi$ is a convex function and $\mathbb{E}[|\varphi(W_t)|] < \infty$ for all $t \geq 0$, then $\varphi(W_t)$ is a submartingale.

Finally, deduce that $|W_t|$, $W_t^2$, $e^{aw_t}$, $\alpha \in \mathbb{R}$ and $W_t^+ = \max\{0, W_t\}$ are non-negative submartingales.

Solution: Let $s < t$ and because $W_t$ is a martingale we therefore have $\mathbb{E}(W_t | \mathcal{F}_s) = W_s$ and hence $\varphi[\mathbb{E}(W_t | \mathcal{F}_s)] = \varphi(W_s)$. From the conditional Jensen’s inequality (see Problem 1.2.3.14, page 48),

$$\mathbb{E}[\varphi(W_t) | \mathcal{F}_s] \geq \varphi[\mathbb{E}(W_t | \mathcal{F}_s)] = \varphi(W_s).$$

In addition, since $\mathbb{E}[|\varphi(W_t)|] < \infty$ and $\varphi(W_t)$ is clearly $\mathcal{F}_t$-adapted we can conclude that $\varphi(W_t)$ is a submartingale.
By setting $\varphi(x) = |x|$, $x \in \mathbb{R}$ and for $\theta \in (0, 1)$ and $x_1, x_2 \in \mathbb{R}$ such that $x_1 \neq x_2$, we have
\[
|\theta x_1 + (1 - \theta)x_2| \leq \theta|x_1| + (1 - \theta)|x_2|.
\]
Therefore, $|x|$ is a non-negative convex function. Because $\mathbb{E}(|W_t|) < \infty$ for all $t \geq 0$, so $|W_t|$ is a non-negative submartingale.

On the contrary, by setting $\varphi(x) = x^2$, $x \in \mathbb{R}$ and since $\varphi''(x) = 2 \geq 0$ for all $x$, so $x^2$ is a non-negative convex function. Since $\mathbb{E}(|W_t^2|) < \infty$ for all $t \geq 0$, so $W_t^2$ is a non-negative submartingale.

Using the same steps we define $\varphi(x) = e^{ax}$ where $a, x \in \mathbb{R}$ and since $\varphi''(x) = a^2 e^{ax} \geq 0$ for all $x$, so $e^{ax}$ is a non-negative convex function. Since $\mathbb{E}(|e^{aW_t}|) < \infty$ for all $t \geq 0$ we can conclude that $e^{aW_t}$ is a non-negative submartingale.

Finally, by setting $\varphi(x) = \max\{0, x\}$, $x \in \mathbb{R}$ and for $\theta \in (0, 1)$ and $x_1, x_2 \in \mathbb{R}$ such that $x_1 \neq x_2$, we have
\[
\max\{0, \theta x_1 + (1 - \theta)x_2\} \leq \max\{0, \theta x_1\} + \max\{0, (1 - \theta)x_2\}
\]
\[
= \theta \max\{0, x_1\} + (1 - \theta)\max\{0, x_2\}.
\]
Therefore, $x^+ = \max\{0, x\}$ is a non-negative convex function and since $\mathbb{E}(|W_t^x|) < \infty$ for all $t \geq 0$, so $W_t^x = \max\{0, W_t\}$ is a non-negative submartingale.

2.2.4 First Passage Time

1. **Doob’s Maximal Inequality.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{X_t : 0 \leq t \leq T\}$ be a continuous non-negative submartingale with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Given $\lambda > 0$ and $\tau = \min\{t : X_t \geq \lambda\}$, show that
\[
\mathbb{E}(X_0) \leq \mathbb{E}\left(X_{\min\{\tau, T\}}\right) \leq \mathbb{E}(X_T).
\]

By writing
\[
X_{\min\{\tau, T\}} = X_{\tau} I_{\{\tau \leq T\}} + X_T I_{\{\tau > T\}}
\]
show that
\[
\mathbb{P}\left(\sup_{0 \leq s \leq \tau} X_s \geq \lambda\right) \leq \frac{\mathbb{E}(X_T)}{\lambda}.
\]

Deduce that if $\{Y_t : 0 \leq t \leq T\}$ is a continuous non-negative supermartingale with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$ then
\[
\mathbb{P}\left(\sup_{0 \leq s \leq \tau} Y_s \geq \lambda\right) \leq \frac{\mathbb{E}(Y_0)}{\lambda}.
\]

**Solution:** For $\lambda > 0$ we let $\tau = \min\{t : X_t \geq \lambda\}$ so that $0 \leq \min\{\tau, T\} \leq T$. Because $X_t$ is a non-negative submartingale we have
\[
\mathbb{E}\left(X_T \mid \mathcal{F}_{\min\{\tau, T\}}\right) \geq X_{\min\{\tau, T\}}.
\]
2.2.4 First Passage Time

or

\[ \mathbb{E}(X_{\min \{\tau, T\}}) \leq \mathbb{E}\left[ \mathbb{E}(X_T | \mathcal{F}_{\min \{\tau, T\}}) \right] = \mathbb{E}(X_T). \]

Using the same steps we can deduce

\[ \mathbb{E}(X_0) \leq \mathbb{E}(X_{\min \{\tau, T\}}) \leq \mathbb{E}(X_T). \]

By definition

\[ X_{\min \{\tau, T\}} = X_\tau I_{\{\tau \leq T\}} + X_T I_{\{\tau > T\}} \]

where

\[ I_{\{\tau \leq T\}} = \begin{cases} 1 & \tau \leq T \\ 0 & \tau > T \end{cases}, \quad I_{\{\tau > T\}} = \begin{cases} 1 & \tau > T \\ 0 & \tau \leq T \end{cases}. \]

Therefore,

\[ \mathbb{E}(X_{\min \{\tau, T\}}) = \mathbb{E}(X_\tau I_{\{\tau \leq T\}}) + \mathbb{E}(X_T I_{\{\tau > T\}}) \geq \lambda \mathbb{P}(\tau \leq T) + \mathbb{E}(X_T I_{\{\tau > T\}}). \]

Taking note that \( \mathbb{E}(X_{\min \{\tau, T\}}) \leq \mathbb{E}(X_T) \), we can write

\[ \lambda \mathbb{P}(\tau \leq T) \leq \mathbb{E}(X_{\min \{\tau, T\}}) - \mathbb{E}(X_T I_{\{\tau > T\}}) \leq \mathbb{E}(X_T) - \mathbb{E}(X_{\min \{\tau, T\}}) \leq \mathbb{E}(X_T). \]

Since

\[ \{\tau \leq T\} \iff \left\{ \sup_{0 \leq t \leq T} X_t \geq \lambda \right\} \]

therefore

\[ \mathbb{P}\left( \sup_{0 \leq t \leq T} X_t \geq \lambda \right) \leq \frac{\mathbb{E}(X_T)}{\lambda}. \]

If \( \{Y_t\}_{0 \leq t \leq T} \) is a supermartingale then

\[ \mathbb{E}(Y_T) \leq \mathbb{E}(Y_{\min \{\tau, T\}}) \leq \mathbb{E}(Y_0) \]

where in this case \( \tau = \min\{t : Y_t \geq \lambda\} \) and by analogy with the above steps, we have

\[ \mathbb{P}\left( \sup_{0 \leq t \leq T} Y_t \geq \lambda \right) \leq \frac{\mathbb{E}(Y_0)}{\lambda}. \]

\[ \square \]

2. Doob’s \( L^p \) Maximal Inequality. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( Z \) be a continuous non-negative random variable where for \( m > 0, \mathbb{E}(Z^m) < \infty \). Show that

\[ \mathbb{E}(Z^m) = m \int_0^\infty a^{m-1} \mathbb{P}(Z > a) \, da. \]
2.2.4 First Passage Time

Let \( \{X_t : 0 \leq t \leq T\} \) be a continuous non-negative submartingale with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T \). Using the above result, for \( p > 1 \) and if \( \mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) < \infty \) show that

\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) \leq \left( \frac{p}{p - 1} \right)^p \mathbb{E} (X_T^p).
\]

Deduce that if \( \{Y_t\}_{0 \leq t \leq T} \) is a continuous non-negative supermartingale with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T \) then

\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} Y_t^p \right) \leq \left( \frac{p}{p - 1} \right)^p \mathbb{E} (Y_0^p).
\]

**Solution:** By defining the indicator function

\[
\mathbb{1}_{\{Z > \alpha\}} = \begin{cases} 1 & Z > \alpha \\ 0 & Z \leq \alpha \end{cases}
\]

we can prove

\[
\int_0^\infty ma^{m-1} \mathbb{P}(Z > \alpha) \, d\alpha = \int_0^\infty ma^{m-1} \mathbb{E}(\mathbb{1}_{\{Z > \alpha\}}) \, d\alpha
\]

\[
= \mathbb{E} \left[ \int_0^\infty ma^{m-1} \mathbb{1}_{\{Z > \alpha\}} \, d\alpha \right]
\]

\[
= \mathbb{E} \left[ \int_0^\infty ma^{m-1} \, d\alpha \right]
\]

\[
= \mathbb{E}(Z^m).
\]

Let \( \tau = \min\{t : X_t \geq \lambda\}, \lambda > 0 \) so that \( 0 \leq \min\{\tau, T\} \leq T \). Therefore, we can deduce

\[
\{\tau \leq T\} \iff \left\{ \sup_{0 \leq t \leq T} X_t \geq \lambda \right\}
\]

and from Problem 2.2.4.1 (page 76) we can show that

\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) = \int_0^\infty p \lambda^{p-1} \mathbb{P} \left( \sup_{0 \leq t \leq T} X_t > \lambda \right) \, d\lambda
\]

\[
\leq \int_0^\infty p \lambda^{p-2} \mathbb{E} \left( X_T \mathbb{1}_{\{\sup_{0 \leq t \leq T} X_t \geq \lambda\}} \right) \, d\lambda
\]

\[
= p \mathbb{E} \left( X_T \int_0^\infty \lambda^{p-2} \mathbb{1}_{\{\sup_{0 \leq t \leq T} X_t \geq \lambda\}} \, d\lambda \right)
\]

\[
= p \mathbb{E} \left( X_T \int_0^{\sup_{0 \leq t \leq T} X_t} \lambda^{p-2} \, d\lambda \right)
\]

\[
= \frac{p}{p - 1} \mathbb{E} \left( X_T \cdot \left\{ \sup_{0 \leq t \leq T} X_t^{p-1} \right\} \right).
\]
Using Hölder’s inequality and taking note that $\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) < \infty$, we can write

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) \leq \frac{p}{p-1} \mathbb{E} \left( X_T^p \right)^{\frac{1}{p}} \mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^{p-1} \right)^{\frac{1}{p}}$$

or

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right)^{\frac{1}{p}} \leq \frac{p}{p-1} \mathbb{E} \left( X_T^p \right)^{\frac{1}{p}}$$

and hence

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} X_t^p \right) \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} \left( X_T^p \right).$$

From Problem 2.2.4.1 (page 76), if $\{Y_t\}_{0 \leq t \leq T}$ is a supermartingale then

$$\mathbb{E} (Y_T) \leq \mathbb{E} (Y_{\min\{\tau, T\}}) \leq \mathbb{E} (Y_0)$$

where in this case $\tau = \min\{t : Y_t \geq \lambda\}$. Following the same steps as discussed before, we have

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} Y_t^p \right) \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} \left( Y_0^p \right).$$

3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Using Doob’s maximal inequality show that for every $T > 0$ and $\lambda, \theta > 0$, we can express

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} e^{\theta W_t} \geq e^{\theta \lambda} \right) \leq e^{\frac{\lambda}{2} T - \theta \lambda}$$

and hence prove that

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} W_t \geq \lambda \right) \leq e^{-\frac{\lambda^2}{2T}}.$$

Finally, deduce that

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} |W_t| \geq \lambda \right) \leq 2e^{-\frac{\lambda^2}{2T}}.$$

**Solution:** Since $\lambda > 0$ and for $\theta > 0$ we can write

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} W_t \geq \lambda \right) = \mathbb{P} \left( \sup_{0 \leq t \leq T} \theta W_t \geq \theta \lambda \right) = \mathbb{P} \left( \sup_{0 \leq t \leq T} e^{\theta W_t} \geq e^{\theta \lambda} \right).$$

From Problem 2.2.3.7 (page 75) we have shown that $e^{\theta W_t}$ is a submartingale and it is also non-negative. Thus, from Doob’s maximal inequality theorem,

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} e^{\theta W_t} \geq e^{\theta \lambda} \right) \leq \frac{\mathbb{E} \left( e^{\theta W_T} \right)}{e^{\theta \lambda}} = e^{\frac{1}{2} \theta^2 T - \theta \lambda}.$$
By minimising $e^{\frac{1}{2}θ^2T−θλ}$ we obtain $θ = λ/T$ and substituting it into the inequality we have

$$\mathbb{P}\left(\sup_{0≤t≤T} W_t ≥ λ\right) ≤ e^{−\frac{λ^2}{2T}}.$$ 

Given that $W_t \sim \mathcal{N}(0, t)$ we can deduce that $\mathbb{P}(|W_t| ≥ λ) = \mathbb{P}(W_t ≥ λ) + \mathbb{P}(W_t ≤ −λ) = 2\mathbb{P}(W_t ≥ λ)$ and hence $\mathbb{P}\left(\sup_{0≤t≤T} |W_t| ≥ λ\right) ≤ 2e^{−\frac{λ^2}{2T}}$. □

4. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $W_t$ be the standard Wiener process with respect to the filtration $\mathcal{F}_t$, $t ≥ 0$. By writing $X_t = x_0 + W_t$ where $x_0 \in \mathbb{R}$ show that $\{X_t : t ≥ 0\}$ is a continuous martingale. Let $T_a = \inf\{t ≥ 0 : X_t = a\}$ and $T_b = \inf\{t ≥ 0 : X_t = b\}$ where $a < b$. Show, using the optional stopping theorem, that

$$\mathbb{P}(T_a < T_b) = \frac{b−x_0}{b−a}.$$ 

**Solution:** Let $X_t = x_0 + W_t$ where $W_t$ is a standard Wiener process. Because $W_t$ is a martingale (see Problem 2.2.3.1, page 71) and because $x_0$ is a constant value, following the same steps we can easily prove that $X_t$ is also a martingale.

By writing $T = \inf\{t ≥ 0 : X_t \notin (a, b)\}$ as the first exit time from the interval $(a, b)$, then from the optional stopping theorem

$$\mathbb{E}(X_T | X_0 = x_0) = \mathbb{E}(x_0 + W_T | X_0 = x_0) = x_0$$

and by definition

$$\mathbb{E}(X_T | X_0 = x_0) = a\mathbb{P}(X_T = a | X_0 = x_0) + b\mathbb{P}(X_T = b | X_0 = x_0)$$

$$x_0 = a\mathbb{P}(T_a < T_b) + b\mathbb{P}(T_a ≥ T_b)$$

$$x_0 = a\mathbb{P}(T_a < T_b) + b\left[1 - \mathbb{P}(T_a < T_b)\right].$$

Therefore,

$$\mathbb{P}(T_a < T_b) = \frac{b−x_0}{b−a}. □$$

5. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $W_t$ be the standard Wiener process with respect to the filtration $\mathcal{F}_t$, $t ≥ 0$. By writing $X_t = x + W_t$ where $x \in \mathbb{R}$ show that $X_t$ and $(X_t − x)^2 − t$ are continuous martingales. Let $T = \inf\{t ≥ 0 : X_t ∉ (a, b)\}$ be the first exit time from the interval $(a, b)$, $a < x < b$ and assuming $T < ∞$ almost surely show, using the optional stopping theorem, that

$$\mathbb{P}(X_T = a | X_0 = x) = \frac{b−x}{b−a} \quad \text{and} \quad \mathbb{P}(X_T = b | X_0 = x) = \frac{x−a}{b−a}$$

with

$$\mathbb{E}(T) = (b−x)(x−a).$$
Solution: Let \( X_t = x + W_t \) where \( W_t \) is a standard Wiener process. Because \( W_t \) and \( W_t^2 - t \) are martingales (see Problems 2.2.3.1, page 71 and 2.2.3.2, page 72) and because \( x_0 \) is a constant value, following the same steps we can easily prove that \( X_t \) and \((X_t - x)^2 - t\) are also martingales.

For \( x \in (a, b) \) and because \( X_t \) is a martingale, from the optional stopping theorem
\[
\mathbb{E}(X_T|X_0 = x) = \mathbb{E}(x + W_T|X_0 = x) = x
\]

and by definition
\[
\mathbb{E}(X_T|X_0 = x) = a\mathbb{P}(X_T = a|X_0 = x) + b\mathbb{P}(X_T = b|X_0 = x)
\]

\[
x = a\mathbb{P}(X_T = a|X_0 = x) + b\left[1 - \mathbb{P}(X_T = a|X_0 = x)\right].
\]

Therefore,
\[
\mathbb{P}(X_T = a|X_0 = x) = \frac{b - x}{b - a}
\]

and
\[
\mathbb{P}(X_T = b|X_0 = x) = 1 - \mathbb{P}(X_T = a|X_0 = x) = \frac{x - a}{b - a}.
\]

Since \( Y_t = (X_t - x)^2 - t \) is a martingale, from the optional stopping theorem we have
\[
\mathbb{E}(Y_T|X_0 = x) = \mathbb{E}(Y_0|X_0 = x) = 0
\]

or
\[
\mathbb{E}\left[(X_T - x)^2 - T|X_0 = x\right] = \mathbb{E}\left[(X_T - x)^2\right|X_0 = x\right] - \mathbb{E}(T|X_0 = x) = 0.
\]

Therefore,
\[
\mathbb{E}(T|X_0 = x) = \mathbb{E}\left[(X_T - x)^2\right|X_0 = x\right]
\]
\[
= (a - x)^2\mathbb{P}(X_T = a|X_0 = x) + (b - x)^2\mathbb{P}(X_T = b|X_0 = x)
\]
\[
= (a - x)^2\left(\frac{b - x}{b - a}\right) + (b - x)^2\left(\frac{x - a}{b - a}\right)
\]
\[
= (b - x)(x - a).
\]

6. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{X_t : t \geq 0\}\) be a continuous martingale on \(\mathbb{R}\).

Show that if \(T = \inf\{t \geq 0 : X_t \notin (a, b)\}\) is the first exit time from the interval \((a, b)\) and assuming \(T < \infty\) almost surely then for \(\theta > 0\),
\[
Z_t = (e^{\theta b} - e^{-\theta a})e^{\theta X_t - \frac{1}{2}\theta^2 t} + (e^{-\theta a} - e^{\theta b})e^{-\theta X_t - \frac{1}{2}\theta^2 t}
\]
is a martingale. Using the optional stopping theorem and if \(a < x < b\), show that
\[
\mathbb{E}\left(e^{-\frac{1}{2}\theta^2 T}\right|X_0 = x) = \frac{\cosh\left[\theta\left(x - \frac{1}{2}(a + b)\right)\right]}{\cosh\left[\frac{1}{2}\theta(b - a)\right]}, \quad \theta > 0.
\]
Solution: Given that $X_t$ is a martingale and both $e^{\theta b} - e^{-\theta a}$ and $e^{-\theta a} - e^{\theta b}$ are independent of $X_t$, then using the results of Problem 2.2.3.3 (page 72) we can easily show that 
\{Z_t : t \geq 0\} is a martingale.

From the optional stopping theorem

$$
\mathbb{E}(Z_T | X_0 = x) = \mathbb{E}(Z_0 | X_0 = x)
$$

and using the identity \( \sinh(x) + \sinh(y) = 2 \sinh \left( \frac{x + y}{2} \right) \cosh \left( \frac{x - y}{2} \right) \) we have

$$
\mathbb{E}(Z_0 | X_0 = x) = e^{\theta(b-x)} - e^{-\theta(b-x)} + e^{-\theta(a+x)} - e^{\theta(a-x)}
$$

$$
= 2 \left[ \sinh[\theta(b-x)] + \sinh[-\theta(a-x)] \right]
$$

$$
= 4 \sinh \left[ \frac{1}{2} \theta(b-a) \right] \cosh \left[ \theta \left( x - \frac{1}{2}(a+b) \right) \right]
$$

and

$$
\mathbb{E}(Z_T | X_0 = x) = \mathbb{E}(Z_T | X_T = a, X_0 = x) \mathbb{P}(X_T = a | X_0 = x)
$$

$$
+ \mathbb{E}(Z_T | X_T = b, X_0 = x) \mathbb{P}(X_T = b | X_0 = x)
$$

$$
= \mathbb{E}(Z_T | X_T = a, X_0 = x) \mathbb{P}(X_T = a | X_0 = x)
$$

$$
+ \mathbb{E}(Z_T | X_T = b, X_0 = x) (1 - \mathbb{P}(X_T = a | X_0 = x))
$$

$$
= [\mathbb{E}(Z_T | X_T = a, X_0 = x) - \mathbb{E}(Z_T | X_T = b, X_0 = x)] \mathbb{P}(X_T = a | X_0 = x)
$$

$$
+ \mathbb{E}(Z_T | X_T = b, X_0 = x)
$$

$$
= \left[ e^{\theta(b-a)} - e^{-\theta(b-a)} \right] \mathbb{E} \left( e^{-\frac{1}{2} \theta^2 T} \bigg| X_0 = x \right)
$$

$$
= 2 \sinh[\theta(b-a)] \mathbb{E} \left( e^{-\frac{1}{2} \theta^2 T} \bigg| X_0 = x \right)
$$

since \( \mathbb{E}(Z_T | X_T = a, X_0 = x) = \mathbb{E}(Z_T | X_T = b, X_0 = x) \). Finally,

$$
\mathbb{E} \left( e^{-\frac{1}{2} \theta^2 T} \bigg| X_0 = x \right) = \frac{4 \sinh \left[ \frac{1}{2} \theta(b-a) \right] \cosh \left[ \theta \left( x - \frac{1}{2}(a+b) \right) \right]}{2 \sinh[\theta(b-a)]}
$$

$$
= \frac{\cosh \left[ \theta \left( x - \frac{1}{2}(a+b) \right) \right]}{\cosh \left[ \frac{1}{2} \theta(b-a) \right]}.
$$

\[\square\]
2.2.4 First Passage Time

7. Laplace Transform of First Passage Time. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process.

Show that for \(\lambda \in \mathbb{R}, X_t = e^{iW_t - \frac{1}{2} \lambda^2 t}\) is a martingale.

By setting \(T_{a,b} = \inf\{t \geq 0 : W_t = a + bt\}\) as the first passage time of hitting the sloping line \(a + bt\), where \(a\) and \(b\) are constants, show that the Laplace transform of its distribution is given by

\[
\mathbb{E} \left( e^{-\theta T_{a,b}} \right) = e^{-a(b + \sqrt{b^2 + 2\theta})}, \quad \theta > 0
\]

and hence show that

\[
\mathbb{E}(T_{a,b}) = \left( \frac{a}{b} \right) e^{-2ab} \quad \text{and} \quad \mathbb{P}(T_{a,b} < \infty) = e^{-2ab}.
\]

**Solution:** To show that \(X_t = e^{iW_t - \frac{1}{2} \lambda^2 t}\) is a martingale, see Problem 2.2.3.3 (page 72).

From the optional stopping theorem

\[
\mathbb{E}(X_{T_{a,b}}) = \mathbb{E}(X_0) = 1
\]

and because at \(t = T_{a,b}, W_{T_{a,b}} = a + bT_{a,b}\), we have

\[
\mathbb{E}\left[ e^{i(a+bT_{a,b}) - \frac{1}{2} \lambda^2 T_{a,b}} \right] = 1
\]

or

\[
\mathbb{E}\left[ e^{(i\lambda - \frac{1}{2} \lambda^2)T_{a,b}} \right] = e^{-i\lambda a}.
\]

By setting \(\theta = -\left( \frac{\lambda b - \frac{1}{2} \lambda^2}{a} \right)\) we have \(\lambda = b \pm \sqrt{b^2 + 2\theta}\). Since \(\theta > 0\) we must have \(\mathbb{E}(e^{-\theta T_{a,b}}) \leq 1\) and therefore \(\lambda = b + \sqrt{b^2 + 2\theta}\). Thus,

\[
\mathbb{E} \left( e^{-\theta T_{a,b}} \right) = e^{-a(b + \sqrt{b^2 + 2\theta})}.
\]

To find \(\mathbb{E}(T_{a,b})\) we differentiate \(\mathbb{E}(e^{-\theta T_{a,b}})\) with respect to \(\theta\),

\[
\mathbb{E} \left( T_{a,b} e^{-\theta T_{a,b}} \right) = \frac{a}{\sqrt{b^2 + 2\theta}} e^{-a(b + \sqrt{b^2 + 2\theta})}
\]

and setting \(\theta = 0,\)

\[
\mathbb{E}(T_{a,b}) = \left( \frac{a}{b} \right) e^{-2ab}.
\]

Finally, to find \(\mathbb{P}(T_{a,b} < \infty)\) by definition

\[
\mathbb{E} \left( e^{-\theta T_{a,b}} \right) = \int_0^\infty e^{-\theta f_{T_{a,b}}(t)} dt
\]
where $f_{Ta,b}(t)$ is the probability density function of $T_{a,b}$. By setting $\theta = 0$,

$$\mathbb{P}(T_{a,b} < \infty) = \mathbb{E}(e^{-\theta T_{a,b}})\bigg|_{\theta=0} = e^{-2ab}.$$  

N.B. Take note that if $b = 0$ we will have $\mathbb{E}(T_{a,b}) = \infty$ and $\mathbb{P}(T_{a,b} < \infty) = 1$ (we say $T_{a,b}$ is finite almost surely).  

2.2.5 Reflection Principle

1. **Reflection Principle.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. By setting $T$ as a stopping time and defining

$$\tilde{W}_t = \begin{cases} 
W_t & \text{if } t \leq T \\
2W_T - W_t & \text{if } t > T 
\end{cases}$$

show that $\{\tilde{W}_t : t \geq 0\}$ is also a standard Wiener process.

**Solution:** If $T$ is finite then from the strong Markov property both the paths $Y_t = \{W_{t+T} - W_T : t \geq 0\}$ and $-Y_t = \{-(W_{t+T} - W_T) : t \geq 0\}$ are standard Wiener processes and independent of $X_t = \{W_t : 0 \leq t \leq T\}$, and hence both $(X_t, Y_t)$ and $(X_t, -Y_t)$ have the same distribution. Given the two processes defined on $[0, T]$ and $[0, \infty)$, respectively, we can paste them together as follows:

$$\phi : (X_t, Y_t) \mapsto (X_t1_{t \leq T} + (Y_t - W_T)1_{t \geq T} : t \geq 0).$$

Thus, the process arising from pasting $X_t = \{W_t : 0 \leq t \leq T\}$ to $Y_t = \{W_{t+T} - W_T : t \geq 0\}$ has the same distribution, which is $\{W_t : t \geq 0\}$. In contrast, the process arising from pasting $X_t = \{W_t : 0 \leq t \leq T\}$ to $-Y_t = \{-(W_{t+T} - W_T) : t \geq 0\}$ is $\{\tilde{W}_t : t \geq 0\}$. Thus, $\{\tilde{W}_t : t \geq 0\}$ is also a standard Wiener process.

2. **Reflection Equality.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. By defining $T_m = \inf\{t \geq 0 : W_t = m\}$, $m > 0$ as the first passage time, then for $w \leq m$, $m > 0$, show that

$$\mathbb{P}(T_m \leq t, W_t \leq w) = \mathbb{P}(W_t \geq 2m - w).$$

**Solution:** From the reflection principle in Problem 2.2.5.1 (page 84), since $W_{T_m} = m$,

$$\mathbb{P}(T_m \leq t, W_t \leq w) = \mathbb{P}(T_m \leq t, 2W_{T_m} - W_t \leq w) = \mathbb{P}(W_t \geq 2m - w).$$
3. **First Passage Time Density Function.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. By setting \(T_w\) as a stopping time such that \(T_w = \inf\{t \geq 0 : W_t = w\}\), \(w \neq 0\) show using the reflection principle, that

\[
\mathbb{P}(T_w \leq t) = 1 + \Phi\left(\frac{|w|}{\sqrt{t}}\right) - \Phi\left(\frac{|w|}{\sqrt{t}}\right)
\]

and the probability density function of \(T_w\) is given as

\[
f_{T_w}(t) = \frac{|w|}{\sqrt{2\pi t}} e^{-\frac{w^2}{2t}}.
\]

**Solution:** We first consider the case when \(w > 0\). By definition

\[
\mathbb{P}(T_w \leq t) = \mathbb{P}(T_w \leq t, W_t \leq w) + \mathbb{P}(T_w \leq t, W_t \geq w).
\]

For the case when \(W_t \leq w\), by the reflection equality

\[
\mathbb{P}(T_w \leq t, W_t \leq w) = \mathbb{P}(W_t \geq w).
\]

On the contrary, if \(W_t \geq w\) then we are guaranteed \(T_w \leq t\). Therefore,

\[
\mathbb{P}(T_w \leq t, W_t \geq w) = \mathbb{P}(W_t \geq w)
\]

and hence

\[
\mathbb{P}(T_w \leq t) = 2\mathbb{P}(W_t \geq w) = \int_{w}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} dy.
\]

By setting \(x = y/\sqrt{2t}\) we have

\[
\mathbb{P}(T_w \leq t) = 2 \int_{\frac{w}{\sqrt{2t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 + \Phi\left(\frac{-w}{\sqrt{t}}\right) - \Phi\left(\frac{w}{\sqrt{t}}\right).
\]

For the case when \(w \leq 0\), using the reflection principle

\[
\mathbb{P}(T_w \leq t) = \mathbb{P}(T_w \leq t, W_t \leq w) + \mathbb{P}(T_w \leq t, W_t \geq w)
\]

\[
= \mathbb{P}(T_w \leq t, -W_t \geq -w) + \mathbb{P}(T_w \leq t, -W_t \leq -w)
\]

\[
= \mathbb{P}(T_w \leq t, -W_t \geq -w) + \mathbb{P}(T_w \leq t, -W_t \leq -w)
\]

\[
= \mathbb{P}(T_w \leq t, \tilde{W}_t \geq -w) + \mathbb{P}(T_w \leq t, \tilde{W}_t \leq -w)
\]

where \(\tilde{W}_t = -W_t\) is also a standard Wiener process. Therefore,

\[
\mathbb{P}(T_w \leq t) = 2\mathbb{P}(\tilde{W}_t \geq -w) = 1 + \Phi\left(\frac{w}{\sqrt{t}}\right) - \Phi\left(\frac{-w}{\sqrt{t}}\right).
\]
and hence
\[ \mathbb{P}(T_w \leq t) = 1 + \Phi \left( -\frac{|w|}{\sqrt{t}} \right) - \Phi \left( \frac{|w|}{\sqrt{t}} \right). \]

To find the density of \( T_w \) we note that
\[ f_{T_w}(t) = \frac{d}{dt} \mathbb{P}(T_w \leq t) = \frac{d}{dt} \left[ 1 + \Phi \left( -\frac{|w|}{\sqrt{t}} \right) - \Phi \left( \frac{|w|}{\sqrt{t}} \right) \right] = \frac{|w|}{t\sqrt{2\pi t}} e^{-\frac{w^2}{2t}}. \]

\[ \square \]

4. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( M_t = \max_{0 \leq s \leq t} W_s \) be the maximum level reached by a standard Wiener process \( \{ W_t : t \geq 0 \} \) in the time interval \([0, t] \). Then for \( a \geq 0, x \leq a \) and for all \( t \geq 0 \), show using the reflection principle that
\[ \mathbb{P}(M_t \leq a, W_t \leq x) = \begin{cases} \Phi \left( \frac{x}{\sqrt{t}} \right) - \Phi \left( \frac{x-2a}{\sqrt{t}} \right) & a \geq 0, x \leq a \\ \Phi \left( \frac{a}{\sqrt{t}} \right) - \Phi \left( \frac{-a}{\sqrt{t}} \right) & a \geq 0, x \geq a \\ 0 & a \leq 0 \end{cases} \]

and the joint probability density function of \((M_t, W_t)\) is
\[ f_{M_t, W_t}(a, x) = \begin{cases} \frac{2(2a-x)}{\sqrt{2\pi t}} e^{-\frac{1}{2} \left( \frac{2a-x}{\sqrt{t}} \right)^2} & a \geq 0, x \leq a \\ 0 & \text{otherwise} \end{cases} \]

**Solution:** For \( a \geq 0, x \leq a \), let \( T_a = \inf \{ t \geq 0 : W_t = a \} \) then
\[ \{ M_t \geq a \} \iff \{ T_a \leq t \} . \]

Taking \( T = T_a \) in the reflection principle,
\[
\mathbb{P}(M_t \geq a, W_t \leq x) = \mathbb{P}(T_a \leq t, W_t \leq x) = \mathbb{P}(T_a \leq t, W_t \geq 2a - x) = \mathbb{P}(W_t \geq 2a - x) = 1 - \Phi \left( \frac{2a-x}{\sqrt{t}} \right) = \Phi \left( \frac{x-2a}{\sqrt{t}} \right). \]
2.2.5 Reflection Principle

Because

\[
\mathbb{P}(W_t \leq x) = \mathbb{P}(M_t \leq a, W_t \leq x) + \mathbb{P}(M_t \geq a, W_t \leq x)
\]

then for \( a \geq 0, x \leq a \),

\[
\mathbb{P}(M_t \leq a, W_t \leq x) = \mathbb{P}(W_t \leq x) - \mathbb{P}(M_t \geq a, W_t \leq x)
\]

\[
= \Phi\left( \frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{x - 2a}{\sqrt{t}} \right).
\]

For the case when \( a \geq 0, x \geq a \) and because \( M_t \geq W_t \), we have

\[
\mathbb{P}(M_t \leq a, W_t \leq x) = \mathbb{P}(M_t \leq a, W_t \leq a) = \mathbb{P}(M_t \leq a)
\]

and by substituting \( x = a \) into \( \Phi\left( \frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{x - 2a}{\sqrt{t}} \right) \) we have

\[
\mathbb{P}(M_t \leq a, W_t \leq x) = \Phi\left( \frac{a}{\sqrt{t}} \right) - \Phi\left( -\frac{a}{\sqrt{t}} \right).
\]

Finally, for the case when \( a \leq 0 \), and since \( M_t \geq W_0 = 0 \), then \( \mathbb{P}(M_t \leq a, W_t \leq x) = 0 \).

To obtain the joint probability density function of \((M_t, W_t)\), by definition

\[
f_{M_t, W_t}(a, x) = \frac{\partial^2}{\partial a \partial x} \mathbb{P}(M_t \leq a, W_t \leq x)
\]

and since \( \mathbb{P}(M_t \leq a, W_t \leq x) \) is a function in both \( a \) and \( x \) only for the case when \( a \geq 0, x \leq a \), then

\[
f_{M_t, W_t}(a, x) = \frac{\partial^2}{\partial a \partial x} \left[ \Phi\left( \frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{x - 2a}{\sqrt{t}} \right) \right]
\]

\[
= \frac{\partial}{\partial a} \left[ \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} - \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-2a)^2}{2t}} \right]
\]

\[
= \frac{-2(x - 2a)}{t\sqrt{2\pi t}} e^{-\frac{(x-2a)^2}{2t}}
\]

\[
= \frac{2(2a - x)}{t\sqrt{2\pi t}} e^{-\frac{(2a-x)^2}{2t}}.
\]

\[\square\]
5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(m_t = \min_{0 \leq s \leq t} W_s\) be the minimum level reached by a standard Wiener process \(\{W_t : t \geq 0\}\) in the time interval \([0, t]\). Then for all \(t \geq 0\) show that

\[
\mathbb{P}(m_t \geq b, W_t \geq x) = \begin{cases} 
\Phi \left( \frac{-x}{\sqrt{t}} \right) - \Phi \left( \frac{2b - x}{\sqrt{t}} \right) & b \leq 0, b \leq x \\
\Phi \left( \frac{-b}{\sqrt{t}} \right) - \Phi \left( \frac{b}{\sqrt{t}} \right) & b \leq 0, b \geq x \\
0 & b \geq 0
\end{cases}
\]

and the joint probability density function of \((m_t, W_t)\) is

\[
f_{m_t, W_t}(b, x) = \begin{cases} 
\frac{-2(2b - x) - e^{\frac{-1}{2} \left( \frac{2b - x}{\sqrt{t}} \right)^2}}{t\sqrt{2\pi t}} & b \leq 0, b \leq x \\
0 & \text{otherwise}
\end{cases}
\]

**Solution:** Given \(m_t = -\max_{0 \leq s \leq t} \{-W_s\}\) we have for \(b \leq 0, b \leq x,\)

\[
\mathbb{P}(m_t \geq b, W_t \geq x) = \mathbb{P}\left( -\max_{0 \leq s \leq t} \{-W_s\} \geq b, W_t \geq x \right)
\]

\[
= \mathbb{P}\left( -\max_{0 \leq s \leq t} \{-W_s\} \geq b, -W_t \leq -x \right)
\]

\[
= \mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \leq -b, \widetilde{W}_t \leq -x \right),
\]

where the last equality comes from the symmetric property of the standard Wiener process such that \(\widetilde{W}_t = -W_t \sim \mathcal{N}(0, t)\) is also a standard Wiener process.

Since

\[
\mathbb{P}(\widetilde{W}_t \leq -x) = \mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \leq -b, \widetilde{W}_t \leq -x \right) + \mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \geq -b, \widetilde{W}_t \leq -x \right)
\]

and from Problem 2.2.5.4 (page 86), we can write

\[
\mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \geq -b, \widetilde{W}_t \leq -x \right) = \Phi \left( \frac{2b - x}{\sqrt{t}} \right)
\]

so

\[
\mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \leq -b, \widetilde{W}_t \leq -x \right) = \mathbb{P}(\widetilde{W}_t \leq -x) - \mathbb{P}\left( \max_{0 \leq s \leq t} \widetilde{W}_t \geq -b, \widetilde{W}_t \leq -x \right)
\]

\[
= \Phi \left( \frac{-x}{\sqrt{t}} \right) - \Phi \left( \frac{2b - x}{\sqrt{t}} \right) .
\]
2.2.6 Quadratic Variation

Consequently, if \( b \leq 0, b \leq x \) we have

\[
\mathbb{P}(m_t \geq b, W_t \geq x) = \Phi\left( -\frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{2b - x}{\sqrt{t}} \right).
\]

For the case when \( b \leq 0, b \geq x \) and since \( m_t \leq W_t \),

\[
\mathbb{P}(m_t \geq b, W_t \geq x) = \mathbb{P}(m_t \geq b, W_t \geq b) = \mathbb{P}(m_t \geq b).
\]

Setting \( x = b \) in \( \Phi\left( -\frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{2b - x}{\sqrt{t}} \right) \) would therefore yield

\[
\mathbb{P}(m_t \geq b, W_t \geq x) = \Phi\left( -\frac{b}{\sqrt{t}} \right) - \Phi\left( \frac{b}{\sqrt{t}} \right).
\]

Finally, for \( b \geq 0 \), since \( m_t \leq W_0 = 0 \) so \( \mathbb{P}(m_t \geq b, W_t \geq x) = 0 \).

To obtain the joint probability density function by definition,

\[
f_{m_t, W_t}(b, x) = \frac{\partial^2}{\partial b \partial x} \mathbb{P}(m_t \leq b, W_t \leq x)
= \frac{\partial^2}{\partial b \partial x} \mathbb{P}(m_t \geq b, W_t \geq x)
\]

and since \( \mathbb{P}(m_t \geq b, W_t \geq x) \) is a function in both \( b \) and \( x \) only for the case when \( b \leq 0, b \leq x \), so

\[
f_{m_t, W_t}(b, x) = \frac{\partial^2}{\partial b \partial x} \left[ \Phi\left( -\frac{x}{\sqrt{t}} \right) - \Phi\left( \frac{2b - x}{\sqrt{t}} \right) \right]
= \frac{\partial}{\partial b} \left[ -\frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}(\frac{x}{\sqrt{t}})^2} + \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}(\frac{2b-x}{\sqrt{t}})^2} \right]
= -\frac{2(2b-x)}{t\sqrt{2\pi t}} e^{-\frac{1}{2}(\frac{2b-x}{\sqrt{t}})^2}.
\]

\[\square\]

2.2.6 Quadratic Variation

1. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Show that it has finite quadratic variation such that

\[
\langle W, W \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \).

Finally, deduce that \( dW_t \cdot dW_t = dt \).
Solution: Since the quadratic variation is a sum of random variables, we need to show that its expected value is \( t \) and its variance converges to zero as \( n \to \infty \).

Let \( \Delta W_{i_j} = W_{i + 1} - W_i \sim \mathcal{N}(0, t/n) \) where \( \mathbb{E}(\Delta W_{i_j}^2) = t/n \) then, by taking expectations we have

\[
\mathbb{E} \left( \lim_{n \to \infty} \sum_{i=0}^{n-1} \Delta W_{i_j}^2 \right) = \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E}(\Delta W_{i_j}^2) = t.
\]

Because \( \Delta W_{i_j}^2/(t/n) \sim \chi^2(1) \) we have \( \mathbb{E}(\Delta W_{i_j}^4) = 3(t/n)^2 \). Therefore, by independence of increments

\[
\text{Var} \left( \lim_{n \to \infty} \sum_{i=0}^{n-1} \Delta W_{i_j}^2 \right) = \mathbb{E} \left[ \left( \lim_{n \to \infty} \sum_{i=0}^{n-1} \Delta W_{i_j}^2 - \frac{t^2}{n} \right)^2 \right]
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ \left( \Delta W_{i_j}^2 - \frac{t^2}{n} \right)^2 \right]
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( \frac{3t^2}{n^2} - 2 \frac{t^2}{n^2} + \frac{t^2}{n^2} \right)
\]

\[
= \lim_{n \to \infty} \frac{2t^2}{n} = 0.
\]

Since \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{i+1} - W_i)^2 = \int_0^t dW_t \cdot dW_t = t \) and \( \int_0^t ds = t \), then by differentiating both sides with respect to \( t \) we have \( dW_t \cdot dW_t = dt \).

\( \square \)

2. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Show that the following cross-variation between \( W_t \) and \( t \), and the quadratic variation of \( t \), are

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{i+1} - W_i) (t_{i+1} - t_i) = 0
\]

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = 0
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \).

Finally, deduce that \( dW_t \cdot dt = 0 \) and \( dt \cdot dt = 0 \).

Solution: Since \( W_{i+1} - W_i \sim \mathcal{N}(0, t/n) \) then taking expectation and variance we have

\[
\mathbb{E} \left[ \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{i+1} - W_i) (t_{i+1} - t_i) \right] = \lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i) \mathbb{E}(W_{i+1} - W_i) = 0
\]
2.2.6 Quadratic Variation

and

$$\text{Var} \left[ \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (t_{i+1} - t_i) \right] = \lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 \text{Var} \left( W_{t_{i+1}} - W_{t_i} \right)$$

$$= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( \frac{1}{n} t_i \right)^2 = \lim_{n \to \infty} \frac{t^3}{n} = 0.$$

Thus, \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (t_{i+1} - t_i) = 0. \)

In addition,

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( \frac{1}{n} \right)^2 = \lim_{n \to \infty} \frac{t^2}{n} = 0.$$

Finally, because

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (t_{i+1} - t_i) = \int_0^t dW_s \cdot ds = 0$$

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = \int_0^t ds \cdot ds = 0$$

then, by differentiating both sides with respect to \( t \), we can deduce that \( dW_t \cdot dt = 0 \) and \( dt \cdot dt = 0. \)

\( \square \)

3. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Show that it has unbounded first variation such that

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} |W_{t_{i+1}} - W_{t_i}| = \infty$$

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}. \)

Solution: Since \( \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 \leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \cdot \sum_{i=0}^{n-1} |W_{t_{i+1}} - W_{t_i}| \), then

$$\sum_{i=0}^{n-1} |W_{t_{i+1}} - W_{t_i}| \geq \frac{\sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2}{\max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}|}. $$
As $W_t$ is continuous, \( \lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| = 0 \) and \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t < \infty \), we can conclude that \( \lim_{n \to \infty} \sum_{i=0}^{n-1} |W_{t_{i+1}} - W_{t_i}| = \infty. \)

\[ \square \]

4. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Show that for \(p \geq 3\),
\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^p = 0
\]
where \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\).
Finally, deduce that \((dW_t)^p = 0, p \geq 3\).

**Solution:** We prove this result via mathematical induction.

For \(p = 3\), we can express
\[
\left| \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) \right| \leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \cdot \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2
\]
\[
= \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \cdot \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2
\]

Because $W_t$ is continuous, we have
\[
\lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| = 0.
\]

In addition, because \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t < \infty \) (see Problem 2.2.6.1, page 89), so
\[
\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) \right| \leq 0.
\]

or
\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^3 = 0.
\]

Thus, the result is true for \(p = 3\).

Assume the result is true for \(p = m, m > 3\) so that \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^m = 0 \).

Then for \(p = m + 1\),
\[
\left| \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^{m+1} \right| \leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \cdot \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^m
\]

and because $W_t$ is continuous such that $\lim_{n \to \infty} \max_{0 \leq k \leq n-1} (W_{t_{k+1}} - W_{t_k}) = 0$ and

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^m = 0,$$

so

$$\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^{m+1} \right| \leq 0$$

or

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^{m+1} = 0.$$

Thus, the result is also true for $p = m + 1$. From mathematical induction we have shown

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^p = 0, \ p \geq 3.$$

Since for $p \geq 3$, $\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^p = \int_0^t (dW_s)^p = 0$, by differentiating both sides with respect to $t$ we have $(dW_t)^p = 0, \ p \geq 3.$
A stochastic differential equation (SDE) is a differential equation in which one or more of the terms has a random component. Within the context of mathematical finance, SDEs are frequently used to model diverse phenomena such as stock prices, interest rates or volatilities to name but a few. Typically, SDEs have continuous paths with both random and non-random components and to drive the random component of the model they usually incorporate a Wiener process. To enrich the model further, other types of random fluctuations are also employed in conjunction with the Wiener process, such as the Poisson process when modelling discontinuous jumps. In this chapter we will concentrate solely on SDEs having only a Wiener process, whilst in Chapter 5 we will discuss SDEs incorporating both Poisson and Wiener processes.

### 3.1 INTRODUCTION

To begin with, a one-dimensional stochastic differential equation can be described as

$$dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t$$

where $W_t$ is a standard Wiener process, $\mu(X_t, t)$ is defined as the drift and $\sigma(X_t, t)$ the volatility. Many financial models can be written in this form, such as the lognormal asset random walk, common spot interest rate and stochastic volatility models.

From an initial state $X_0 = x_0$, we can write in integrated form

$$X_t = X_0 + \int_0^t \mu(X_s, s) \, ds + \int_0^t \sigma(X_s, s) \, dW_s$$

with $\int_0^t [\vert \mu(X_s, s) \vert + \sigma(X_s, s)^2] \, ds < \infty$ and the solution of the integral equation is called an \textit{Itô process} or an \textit{Itô diffusion}.

In Chapter 2 we described the properties of a Wiener process and, given that it is non-differentiable, there is a crucial difference between stochastic calculus and integral calculus. We consider an integral with respect to a Wiener process and write

$$I_t = \int_0^t f(W_s, s) \, dW_s$$

where the integrand $f(W_t, t)$ is $\mathcal{F}_t$ measurable (i.e., $f(W_t, t)$ is a stochastic process adapted to the filtration of a Wiener process) and is square-integrable

$$\mathbb{E} \left( \int_0^t f(W_s, s)^2 \, ds \right) < \infty$$

for all $t \geq 0$. 
3.1 INTRODUCTION

Let $t > 0$ be a positive constant and assume $f(W_{t_i}, t_i)$ is constant over the interval $[t_i, t_{i+1})$, where $t_i = it/n$, $0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t$ for $n \in \mathbb{N}$. Here we call such a process $f$ an elementary process or a simple process, and for such a process the Itô integral $I_t$ can be defined as

$$I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i}).$$

We now give a formal result of the Itô integral as follows.

**Theorem 3.1** Let \( \{W_t : t \geq 0\} \) be a standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and let \( \mathcal{F}_t, t \geq 0 \) be the associated filtration. The Itô integral defined by

$$I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i})$$

where $f$ is a simple process and $t_i = it/n$, $0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t$, $n \in \mathbb{N}$ has the following properties:

- The paths of $I_t$ are continuous.
- For each $t$, $I_t$ is $\mathcal{F}_t$ measurable.
- If $I_t = \int_0^t f(W_s, s) \, dW_s$ and $J_t = \int_0^t g(W_s, s) \, dW_s$ where $g$ is a simple process, then

$$I_t \pm J_t = \int_0^t \left[ f(W_s, s) \pm g(W_s, s) \right] \, dW_s$$

and for a constant $c$

$$cI_t = \int_0^t cf(W_s, s) \, dW_s.$$

- $I_t$ is a martingale.
- $\mathbb{E}[I_t^2] = \mathbb{E} \left( \int_0^t f(W_s, s)^2 \, ds \right)$.
- The quadratic variation $\langle I, I \rangle_t = \int_0^t f(W_s, s)^2 \, ds$.

In mathematics, Itô's formula (or lemma) is used in stochastic calculus to find the differential of a function of a particular type of stochastic process. In essence, it is the stochastic calculus counterpart of the chain rule in ordinary calculus via a Taylor series expansion. The formula is widely employed in mathematical finance and its best-known application is in the derivation of the Black–Scholes equation used to value options. The following is a formal result of Itô's formula.

**Theorem 3.2 (One-Dimensional Itô’s Formula)** Let \( \{W_t : t \geq 0\} \) be a standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and let \( \mathcal{F}_t, t \geq 0 \) be the associated filtration. Consider a stochastic process $X_t$ satisfying the following SDE

$$dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t.$$
3.1 INTRODUCTION

or in integrated form,

\[ X_t = X_0 + \int_0^t \mu(X_s, s) \, ds + \int_0^t \sigma(X_s, s) \, dW_s \]

with \( \int_0^t \left[ |\mu(X_s, s)| + \sigma(X_s, s)^2 \right] \, ds < \infty \). Then for any twice differentiable function \( g(X_t, t) \), the stochastic process \( Y_t = g(X_t, t) \) satisfies

\[ dY_t = \frac{\partial g}{\partial t}(X_t, t) \, dt + \frac{\partial g}{\partial X_t}(X_t, t) \, dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial^2 X_t^2}(X_t, t)(dX_t)^2 \]

where \((dX_t)^2\) is computed according to the rule

\[ (dW_t)^2 = dt, \quad (dt)^2 = dW_t \, dt = dt \, dW_t = 0. \]

In integrated form,

\[ Y_t = Y_0 + \int_0^t \frac{\partial g}{\partial t}(X_s, s) \, ds + \int_0^t \frac{\partial g}{\partial X_t}(X_s, s) \, dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 g}{\partial^2 X_t^2}(X_s, s) \, d\langle X, X \rangle_s \]

\[ = Y_0 + \int_0^t \left[ \frac{\partial g}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial g}{\partial X_t}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 g}{\partial^2 X_t^2}(X_s, s) \right] \, ds \]

\[ + \int_0^t \sigma(X_s, s) \frac{\partial g}{\partial X_t}(X_s, s) \, dW_s \]

where \( \langle X, X \rangle_t = \int_0^t \sigma(X_s, s)^2 \, ds \).

The theory of SDEs is a framework for expressing the dynamical models that include both the random and non-random components. Many solutions to SDEs are Markov processes, where the future depends on the past only through the present. For this reason, the solutions can be studied using backward and forward Kolmogorov equations, which turn out to be linear parabolic partial differential equations of diffusion type. But before we discuss them, the following Feynman–Kac theorem describes an important link between stochastic differential equations and partial differential equations.

**Theorem 3.3 (Feynman–Kac Formula for One-Dimensional Diffusion Process)** Let \( \{W_t : t \geq 0\} \) be a standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and let \( \mathcal{F}_t \), \( t \geq 0 \) be the associated filtration. Let \( X_t \) be the solution of the following SDE:

\[ dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t \]
and define \( r \) as a function of \( t \). For \( t \in [0, T] \) where \( T > 0 \) and if \( V(X_t, t) \) satisfies the partial differential equation (PDE)

\[
\frac{\partial V}{\partial t}(X_t, t) + \frac{1}{2} \sigma^2(X_t, t) \frac{\partial^2 V}{\partial X_t^2}(X_t, t) + \mu(X_t, t) \frac{\partial V}{\partial X_t}(X_t, t) - r(t) V(X_t, t) = 0
\]

subject to the boundary condition \( V(X_T, T) = \Psi(X_T) \), then under the filtration \( \mathcal{F}_t \) the solution of the PDE is given by

\[
V(X_t, t) = \mathbb{E} \left[ e^{-\int_t^T r(u) du} \Psi(X_T) \mid \mathcal{F}_t \right].
\]

The Feynman–Kac formula can be used to study the transition probability density function of the one-dimensional stochastic differential equation

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t.
\]

For \( t \in [0, T] \), \( T > 0 \) and conditioning \( X_t = x \) we can write

\[
V(x, t) = \mathbb{E} \left[ e^{-\int_t^T r(u) du} \Psi(X_T) \mid \mathcal{F}_t \right] = e^{-\int_t^T r(u) du} \int_{-\infty}^{\infty} \Psi(y) p(x, t; y, T) \, dy
\]

where the random variable \( X_T \) has transition probability density \( p(x, t; y, T) \) in the \( y \) variable.

The transition probability density function \( p(x, t; y, T) \) satisfies two parabolic partial differential equations, the forward Kolmogorov equation (or Fokker–Planck equation) and the backward Kolmogorov equation. In the forward Kolmogorov function, it involves derivatives with respect to the future state \( y \) and time \( T \), whilst in the backward Kolmogorov function, it involves derivatives with respect to the current state \( x \) and time \( t \).

**Theorem 3.4 (Forward Kolmogorov Equation for One-Dimensional Diffusion Process)** Let \( \{W_t : t \geq 0\} \) be the standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). Consider the stochastic differential equation

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t
\]

and for \( t \in [0, T] \), \( T > 0 \) and by conditioning \( X_t = x \), the random variable \( X_T \) having a transition probability density \( p(x, t; y, T) \) in the \( y \) variable satisfies

\[
\frac{\partial}{\partial T} p(x, t; y, T) = \frac{1}{2} \sigma^2(y, T) \frac{\partial^2}{\partial y^2} p(x, t; y, T) - \frac{\partial}{\partial y} (\mu(y, T) p(x, t; y, T)).
\]

**Theorem 3.5 (Backward Kolmogorov Equation for One-Dimensional Diffusion Process)** Let \( \{W_t : t \geq 0\} \) be the standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). Consider the stochastic differential equation

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t
\]
3.1 INTRODUCTION

and for \( t \in [0, T] \), \( T > 0 \) and by conditioning \( X_t = x \), the random variable \( X_T \) having a transition probability density \( p(x, t; y, T) \) in the \( x \) variable satisfies

\[
\frac{\partial}{\partial t} p(x, t; y, T) + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2}{\partial x^2} p(x, t; y, T) + \mu(x, t) \frac{\partial}{\partial x} p(x, t; y, T) = 0.
\]

In contrast, a multi-dimensional diffusion process can be described by a set of stochastic differential equations

\[
dX_t^{(i)} = \mu^{(i)}(X_t^{(i)}, t) \, dt + \sum_{j=1}^{n} \sigma^{(i,j)}(X_t^{(i)}, t) \, dW_t^{(j)}, \quad i = 1, 2, \ldots, m
\]

where \( W_t = \begin{bmatrix} W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)} \end{bmatrix}^T \) is the \( n \)-dimensional standard Wiener process such that \( dW_t^{(i)} dW_t^{(j)} = \rho_{ij} \, dt \), \( \rho_{ij} \in (-1, 1) \) for \( i \neq j \), \( i, j = 1, 2, \ldots, n \), \( \mu(X, t) = \begin{bmatrix} \mu^{(1)}(X_t^{(1)}, t), \mu^{(2)}(X_t^{(2)}, t), \ldots, \mu^{(m)}(X_t^{(m)}, t) \end{bmatrix}^T \) is the \( m \times n \) volatility matrix. Given the initial state \( X_0^{(i)} \), we can write in integrated form

\[
X_t^{(i)} = X_0^{(i)} + \int_0^t \mu^{(i)}(X_s^{(i)}, s) \, ds + \sum_{j=1}^{n} \int_0^t \sigma^{(i,j)}(X_s^{(i)}, s) \, dW_s^{(j)}
\]

with \( \int_0^t \left[ |\mu^{(i)}(X_s^{(i)}, s)| + |\sigma^{(i,p)}(X_s^{(i)}, s)\sigma^{(i,q)}(X_s^{(i)}, s)| \right] \, ds < \infty \) for \( i = 1, 2, \ldots, m \) and \( p, q = 1, 2, \ldots, n \).

**Definition 3.6 (Multi-Dimensional Itô’s Formula)** Let \( \{W_t^{(i)} : t \geq 0\}, \, i = 1, 2, \ldots, n \) be a sequence of standard Wiener processes on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and let \( \mathcal{F}_t, \, t \geq 0 \) be the associated filtration. Consider a stochastic process \( X_t^{(i)} \) satisfying the following SDE

\[
dX_t^{(i)} = \mu^{(i)}(X_t^{(i)}, t) \, dt + \sum_{j=1}^{n} \sigma^{(i,j)}(X_t^{(i)}, t) \, dW_t^{(j)}, \quad i = 1, 2, \ldots, m
\]

or in integrated form

\[
X_t^{(i)} = X_0^{(i)} + \int_0^t \mu^{(i)}(X_s^{(i)}, s) \, ds + \sum_{j=1}^{n} \int_0^t \sigma^{(i,j)}(X_s^{(i)}, s) \, dW_s^{(j)}
\]

with \( \int_0^t [\mu^{(i)}(X_s^{(i)}, s)] + [\sigma^{(i,p)}(X_s^{(i)}, s)\sigma^{(i,q)}(X_s^{(i)}, s)] \, ds < \infty \) for \( i = 1, 2, \ldots, m \) and \( p, q = 1, 2, \ldots, n \). Then, for any twice differentiable function \( h(X, t) \), where \( X = \begin{bmatrix} X_1^{(1)}, X_2^{(2)}, \ldots, X_m^{(m)} \end{bmatrix} \),
\( X_i^{(m)} \) is the stochastic process \( Z_t = h(X_i^{(1)}, X_i^{(2)}, \ldots, X_i^{(m)}, t) \) satisfies

\[
dZ_t = \frac{\partial h}{\partial t}(X_i, t) \, dt + \sum_{i=1}^{m} \frac{\partial h}{\partial X_i^{(i)}}(X_i, t) \, dX_i^{(i)} + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 h}{\partial X_i^{(i)} \partial X_j^{(j)}}(X_i, t) \, dX_i^{(i)} dX_j^{(j)}
\]

\[
= \left[ \frac{\partial h}{\partial t}(X_i, t) + \sum_{i=1}^{m} \mu_i^{(i)}(X_i^{(i)}, t) \, \frac{\partial h}{\partial X_i^{(i)}}(X_i, t) \right. \left. + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \rho_{ij} \left( \sum_{k=1}^{n} \sum_{l=1}^{n} \sigma_{i,k}(X_i^{(i)}, t) \sigma_{j,l}(X_j^{(j)}, t) \right) \frac{\partial^2 h}{\partial X_i^{(i)} \partial X_j^{(j)}}(X_i, t) \right] \, dt
\]

\[
+ \sum_{i=1}^{m} \left( \sum_{j=1}^{n} \sigma_{i,j}^{(i)}(X_i^{(i)}, t) \right) \frac{\partial h}{\partial X_i^{(i)}}(X_i, t) \, dW_i^{(i)}
\]

where \( dX_i^{(i)} dX_j^{(j)} \) is computed according to the rule

\[
dW_i^{(i)} dW_j^{(j)} = \rho_{ij} dt, \quad (dt)^2 = dW_i^{(i)} dt = dt dW_i^{(i)} = 0
\]

such that \( \rho_{ij} \in (-1, 1) \) and \( \rho_{ii} = 1 \). In integrated form

\[
Z_t = Z_0 + \int_0^t \frac{\partial h}{\partial t}(X_i, s) \, ds + \int_0^t \sum_{i=1}^{m} \frac{\partial h}{\partial X_i^{(i)}}(X_i, s) \, dX_i^{(i)}
\]

\[
+ \int_0^t \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 h}{\partial X_i^{(i)} \partial X_j^{(j)}}(X_i, s) \, d(X_i^{(i)}, X_j^{(j)})
\]

\[
= Z_0 + \int_0^t \left[ \frac{\partial h}{\partial t}(X_i, s) + \sum_{i=1}^{m} \mu_i^{(i)}(X_i^{(i)}, s) \, \frac{\partial h}{\partial X_i^{(i)}}(X_i, s) \right. \left. + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \rho_{ij} \left( \sum_{k=1}^{n} \sum_{l=1}^{n} \sigma_{i,k}(X_i^{(i)}, s) \sigma_{j,l}(X_j^{(j)}, s) \right) \frac{\partial^2 h}{\partial X_i^{(i)} \partial X_j^{(j)}}(X_i, s) \right] \, ds
\]

\[
+ \int_0^t \sum_{i=1}^{m} \left( \sum_{j=1}^{n} \sigma_{i,j}^{(i)}(X_i^{(i)}, s) \right) \frac{\partial h}{\partial X_i^{(i)}}(X_i, s) \, dW_i^{(i)}
\]

where \( \langle X_i^{(i)}, X_j^{(j)} \rangle_t = \int_0^t \rho_{ij} \sum_{k=1}^{m} \sigma_{i,k}^{(i)}(X_i^{(i)}, s) \sigma_{j,k}^{(j)}(X_j^{(j)}, s) \, ds \).

The Feynman–Kac theorem for a one-dimensional diffusion process also extends to a multi-dimensional version.

**Theorem 3.7 (Feynman–Kac Formula for Multi-Dimensional Diffusion Process)** Let \( \{ W_i^{(i)} : t \geq 0 \}, i = 1, 2, \ldots, n \) be a sequence of standard Wiener processes on the probability
3.1 INTRODUCTION

Theorem 3.8 (Forward Kolmogorov Equation for Multi-Dimensional Diffusion Process) Let \( \mathcal{F}_t \) be the solution of the following SDE

\[
dX_t^{(i)} = \mu^{(i)}(X_t^{(i)}, t) \, dt + \sum_{j=1}^{m} \sigma^{(ij)}(X_t^{(i)}, t) \, dW_t^{(j)}, \quad i = 1, 2, \ldots, m
\]

where \( dW_t^{(i)} dW_t^{(j)} = \rho_{ij} dt, \) \( \rho_{ij} \in (-1, 1) \) for \( i \neq j, \rho_{ii} = 1, \) \( i, j = 1, 2, \ldots, n \) and define \( r \) to be a function of \( t \). By denoting \( X_t = \left[ X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(m)} \right]^T \), for \( t \in [0, T] \) where \( T > 0 \) and if \( V(X_t, t) \) satisfies the PDE

\[
\frac{\partial V}{\partial t}(X_t, t) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \left( \sum_{k=1}^{n} \sum_{l=1}^{n} \sigma^{(ik)}(X_t^{(i)}, t) \sigma^{(lj)}(X_t^{(l)}, t) \right) \frac{\partial^2 V}{\partial X_t^{(i)} \partial X_t^{(j)}}(X_t, t) + \sum_{i=1}^{m} \mu^{(i)}(X_t^{(i)}, t) \frac{\partial V}{\partial X_t^{(i)}}(X_t, t) - r(t)V(X_t, t) = 0
\]

subject to the boundary condition \( V(X_T, T) = \Psi(X_T) \), then under the filtration \( \mathcal{F}_t \), the solution of the PDE is given by

\[
V(X_t, t) = \mathbb{E} \left[ e^{-\int_0^T r(u) \, du} \Psi(X_T) \middle| \mathcal{F}_t \right].
\]

Similarly, we have the Kolmogorov forward and backward equations for multi-dimensional diffusion processes as well.

Theorem 3.8 (Forward Kolmogorov Equation for Multi-Dimensional Diffusion Process) Let \( \{W_t^{(i)} : t \geq 0\}, i = 1, 2, \ldots, n \) be a sequence of standard Wiener processes on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). Consider the stochastic differential equations

\[
dX_t^{(i)} = \mu^{(i)}(X_t^{(i)}, t) \, dt + \sum_{j=1}^{m} \sigma^{(ij)}(X_t^{(i)}, t) \, dW_t^{(j)}, \quad i = 1, 2, \ldots, m
\]

where \( dW_t^{(i)} dW_t^{(j)} = \rho_{ij} dt, \) \( \rho_{ij} \in (-1, 1) \) for \( i \neq j, \rho_{ii} = 1, \) \( i, j = 1, 2, \ldots, n \). By denoting \( X_t = \left[ X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(m)} \right]^T \), and for \( t \in [0, T], T > 0 \) and by conditioning \( X_t = x \) where \( x = \left[ x^{(1)}, x^{(2)}, \ldots, x^{(m)} \right]^T \), the m-dimensional random variable \( X_T \) having a transition probability density \( p(x, t ; y, T) \) in the m-dimensional variable \( y = \left[ y^{(1)}, y^{(2)}, \ldots, y^{(m)} \right]^T \) satisfies

\[
\frac{\partial}{\partial T} p(x, t ; y, T) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial y^{(i)} \partial y^{(j)}} \left[ \sum_{k=1}^{n} \sum_{l=1}^{n} (\rho_{kl} \sigma^{(ik)}(y^{(i)}, T) \sigma^{(lj)}(y^{(j)}, T)) p(x, t ; y, T) \right] - \sum_{i=1}^{m} \frac{\partial}{\partial y^{(i)}} (\mu^{(i)}(y^{(i)}, T) p(x, t ; y, T)).
\]
Theorem 3.9 (Backward Kolmogorov Equation for Multi-Dimensional Diffusion Process) Let \( \{W_i^j : t \geq 0\}, i = 1, 2, \ldots, n \) be a sequence of independent standard Wiener processes on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Consider the stochastic differential equations

\[
dX_i^j = \mu_i^j(X_i^j, t) \, dt + \sum_{j=1}^{n} \sigma_i^{(ij)}(X_i^j, t) \, dW_i^j, \quad i = 1, 2, \ldots, m
\]

where \( dW_i^j dW_i^j = \rho_{ij} dt, \rho_{ij} \in (-1, 1) \) for \( i \neq j \), \( \rho_{ii} = 1, i, j = 1, 2, \ldots, n \). By denoting \( X_i = [X_i^1, X_i^2, \ldots, X_i^m]^T \), and for \( t \in [0, T] \), \( T > 0 \) and by conditioning \( X_i = x \) where \( x = [x^1, x^2, \ldots, x^m]^T \), the m-dimensional random variable \( X_T \) having a transition probability density \( p(x, t; y, T) \) in the m-dimensional variable \( y = [y^1, y^2, \ldots, y^m]^T \) satisfies

\[
\frac{\partial}{\partial t} p(x, t; y, T) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \left( \sum_{k=1}^{n} \sum_{l=1}^{n} \rho_{kl} \sigma_i^{(kl)}(x_i^k, t) \sigma_i^{(jl)}(x_i^l, t) \right) \frac{\partial^2}{\partial x_i^k \partial x_i^l} p(x, t; y, T) + \sum_{i=1}^{m} \mu_i^0(x_i^0, t) \frac{\partial}{\partial x_i^0} p(x, t; y, T) = 0.
\]

### 3.2 PROBLEMS AND SOLUTIONS

#### 3.2.1 Itô Calculus

1. **Itô Integral.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_i : t \geq 0\} \) be a standard Wiener process. Let the Itô integral of \( W_i \, dW_i \) be defined as the following limit

\[
I(t) = \int_0^t W_i \, dW_i = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} - W_{t_i}
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t \) for \( n \in \mathbb{N} \).

Show that the quadratic variation of \( W_i \) is

\[
\langle W, W \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t
\]

and hence

\[
I(t) = \frac{1}{2} (W_t^2 - t).
\]

Finally, show that the Itô integral is a martingale.

**Solution:** For the first part of the solution, see Problem 2.2.6.1 (page 89).

Given \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t \) and by expanding,

\[
I(t) = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i} (W_{t_{i+1}} - W_{t_i})
\]
3.2.1 Itô Calculus

\[ = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left\{ \frac{1}{2} \left( W_{i+1}^2 - W_i^2 \right) - \frac{1}{2} \left( W_{i+1} - W_i \right)^2 \right\} \]

\[ = \frac{1}{2} \left[ \lim_{n \to \infty} \left( W_{n+1}^2 - W_n^2 \right) - n \right] \]

\[ = \frac{1}{2} \left( W_t^2 - t \right). \]

To show that \( I(t) \) is a martingale, see Problem 2.2.3.2 (page 72).

\[ \textit{N.B.} \quad \text{Without going through first principles, we can also show that} \]

\[ \int_0^t W_s \, dW_s = \frac{1}{2} \left( W_t^2 - t \right) \text{ by using Itô's formula on} \]

\[ X_t = \frac{1}{2} W_t^2, \text{ where} \]

\[ dX_t = \frac{\partial X_t}{\partial t} \, dt + \frac{\partial X_t}{\partial W_t} \, dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \, dW_t^2 + \ldots \]

\[ = W_t \, dW_t + \frac{1}{2} \, dt. \]

Taking integrals,

\[ \int_0^t dX_s = \int_0^t W_s \, dW_s + \frac{1}{2} \int_0^t ds \]

\[ X_t - X_0 = \int_0^t W_s \, dW_s + \frac{1}{2} \, t \]

and since \( W_0 = 0 \), so \( \int_0^t W_s \, dW_s = \frac{1}{2} \left( W_t^2 - t \right) \).

\[ \square \]

2. \textit{Stratonovich Integral.} Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Let the Stratonovich integral of \( W_t \circ dW_t \) be defined by the following limit

\[ S(t) = \int_0^t W_s \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \left( W_{i+1} + W_i \right) \left( W_{i+1} - W_i \right) \]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \).

Show that \( S(t) = \frac{1}{2} W_t^2 \) and show also that the Stratonovich integral is not a martingale.

\textbf{Solution:} Expanding,

\[ S(t) = \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \left( W_{i+1} + W_i \right) \left( W_{i+1} - W_i \right) \]

\[ = \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \left( W_{i+1}^2 - W_i^2 \right) \]

\[ = \lim_{n \to \infty} \frac{1}{2} \left( W_{n+1}^2 - W_0^2 \right) \]

\[ = \frac{1}{2} W_t^2. \]
Let $S(u) = \int_0^u W_s \circ dW_s$, $u < t$ and under the filtration $\mathcal{F}_u$ and because $W_t - W_u \perp \mathcal{F}_u$, we have

$$
\mathbb{E}(S(t)|\mathcal{F}_u) = \frac{1}{2} \mathbb{E} \left( W_t^2 | \mathcal{F}_u \right) - \frac{1}{2} \mathbb{E} \left( W_u^2 | \mathcal{F}_u \right) + \frac{1}{2} \mathbb{E} \left( \int_0^t (W_s - W_u)^2 ds | \mathcal{F}_u \right)
$$

$$
= \frac{1}{2} E \left( (W_t - W_u)^2 | \mathcal{F}_u \right) + \frac{1}{2} E \left( W_u (W_t - W_u) | \mathcal{F}_u \right) + \frac{1}{2} E \left( W_u^2 | \mathcal{F}_u \right)
$$

$$
= \frac{1}{2} (t - u) + \frac{1}{2} W_u^2
$$

Therefore, $S(t)$ is not a martingale.

\[ \square \]

3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Let the integral of $W_t \circ dW_t$ be defined by the following limit

$$
J(t) = \int_0^t W_s \circ dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} - W_{t_i}
$$

where $t_i = it/n$, $0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t$, $n \in \mathbb{N}$.

Show that the quadratic variation of $W_t$ is

$$
\langle W, W \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t
$$

and hence

$$
J(t) = \frac{1}{2} \left( W_t^2 + t \right).
$$

Finally, show also that the integral is not a martingale.

Solution: To show that $\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t$, see Problem 2.2.6.1 (page 89).

By expanding,

$$
J(t) = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} - W_{t_i}
$$

$$
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( \frac{1}{2} (W_{t_{i+1}}^2 - W_i^2) + \frac{1}{2} (W_{t_{i+1}} - W_i)^2 \right)
$$

$$
= \frac{1}{2} \left[ \lim_{n \to \infty} (W_{t_n}^2 - W_0^2) + t \right]
$$

$$
= \frac{1}{2} (W_t^2 + t).
$$
3.2.1 Itô Calculus

To show that \( J(t) \) is not a martingale, we note from Problem 3.2.1.2 (page 103) that under the filtration \( \mathcal{F}_u, u < t \),

\[
\mathbb{E}(J(t)|\mathcal{F}_u) = \mathbb{E} \left( \frac{1}{2} (W_t^2 + t) \bigg| \mathcal{F}_u \right)
\]
\[
= \frac{1}{2} t + \frac{1}{2} \mathbb{E} \left( W_u^2 \bigg| \mathcal{F}_u \right)
\]
\[
= \frac{1}{2} t + \frac{1}{2} (t - u) + \frac{1}{2} W_u^2
\]
\[
\neq \frac{1}{2} (W_u^2 + u).
\]

Therefore, \( J(t) \) is not a martingale. \( \square \)

4. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. We define the Stratonovich integral of \( f(W_t, t) \circ dW_t \) as the following limit

\[
\int_0^t f(W_s, s) \circ dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f \left( \frac{W_{t_i} + W_{t_{i+1}}}{2}, t_i \right) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

and the Itô integral of \( f(W_t, t) dW_t \) as

\[
\int_0^t f(W_s, s) dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) (W_{t_{i+1}} - W_{t_i})
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \).

Show that

\[
\int_0^t f(W_s, s) \circ dW_s = \frac{1}{2} \int_0^t \frac{\partial f(W_s, s)}{\partial W_s} dW_s + \int_0^t f(W_s, s) dW_s.
\]

Solution: To prove this result we consider the difference between the two stochastic integrals and using the mean value theorem,

\[
\int_0^t f(W_s, s) \circ dW_s - \int_0^t f(W_s, s) dW_s
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} f \left( \frac{W_{t_i} + W_{t_{i+1}}}{2}, t_i \right) \left( W_{t_{i+1}} - W_{t_i} \right) - \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ f \left( \frac{W_{t_i} + W_{t_{i+1}}}{2}, t_i \right) - f(W_{t_i}, t_i) \right] \left( W_{t_{i+1}} - W_{t_i} \right)
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ \frac{1}{2} \frac{\partial f(W_{t_i}, t_i)}{\partial W_{t_i}} \left( W_{t_{i+1}} - W_{t_i} \right)^2 + \frac{1}{4} \frac{\partial^2 f(W_{t_i}, t_i)}{\partial W_{t_i}^2} \left( W_{t_{i+1}} - W_{t_i} \right)^3 + \ldots \right]
\]
3.2.1 Itô Calculus

since \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{i+1} - W_i)^p = 0 \) for \( p \geq 3 \) and hence, for a simple process \( g(W_t, t) \),

\[
\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} g(W_{i+1}, t_i)(W_{i+1} - W_i)^p \right|
\leq \lim_{n \to \infty} \max_{0 \leq k \leq n-1} \left| g(W_{i+k}, t_k) \right| \left| \sum_{i=0}^{n-1} (W_{i+1} - W_i)^p \right|
= 0, \quad p \geq 3
\]
or

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} g(W_{i+1}, t_i)(W_{i+1} - W_i)^p = 0, \quad p \geq 3.
\]

Therefore,

\[
\int_0^t f(W_s, s) \circ dW_s - \int_0^t f(W_s, s) dW_s = \int_0^t f(W_s, s) \circ dW_s \circ dW_s - \int_0^t f(W_s, s) dW_s
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \frac{\partial f(W_{i+1}, t_i)}{\partial W_{i+1}} (W_{i+1} - W_i)^2
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \frac{\partial f(W_{i+1}, t_i)}{\partial W_{i+1}} (t_{i+1} - t_i)
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \frac{\partial f(W_{i+1}, t_i)}{\partial W_{i+1}} (t_{i+1} - t_i)
= \frac{1}{2} \int_0^t \frac{\partial f(W_s, s)}{\partial W_s} ds
\]

since, from Problem 2.2.6.1 (page 89),

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{2} \frac{\partial f(W_{i+1}, t_i)}{\partial W_{i+1}} \left[ (W_{i+1} - W_i)^2 - (t_{i+1} - t_i) \right]
\leq \lim_{n \to \infty} \max_{0 \leq k \leq n-1} \left| \frac{1}{2} \frac{\partial f(W_{i+k}, t_k)}{\partial W_{i+k}} \left[ \sum_{i=0}^{n-1} (W_{i+1} - W_i)^2 - \sum_{i=0}^{n-1} (t_{i+1} - t_i) \right] \right|
= 0.
\]

Thus,

\[
\int_0^t f(W_s, s) \circ dW_s = \frac{1}{2} \int_0^t \frac{\partial f(W_s, s)}{\partial W_s} ds + \int_0^t f(W_s, s) dW_s.
\]
5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process such that \(\mathcal{F}_t\) is the associated filtration. The Itô integral with respect to the standard Wiener process can be defined as

\[
I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

where \(f\) is a simple process and \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\).

Show that the path of \(I_t\) is continuous and that it is also \(\mathcal{F}_t\) measurable for all \(t\).

**Solution:** Given that \(W_{t_i}, t_i = it/n, n \in \mathbb{N}\) is both continuous and \(\mathcal{F}_{t_i}\) measurable, then for a simple process, \(f\),

\[
I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right).
\]

The path of the Itô integral is also continuous and \(\mathcal{F}_t\) measurable for all \(t\).

\(\square\)

6. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process such that \(\mathcal{F}_t\) is the associated filtration. The Itô integrals with respect to the standard Wiener process can be defined as

\[
I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

and

\[
J_t = \int_0^t g(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} g(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

where \(f\) and \(g\) are simple processes and \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\). Show that

\[
I_t \pm J_t = \int_0^t \left[ f(W_s, s) \pm g(W_s, s) \right] \, dW_s
\]

and for constant \(c\),

\[
cl_t = \int_0^t cf(W_s, s) \, dW_s, \quad cJ_t = \int_0^t cg(W_s, s) \, dW_s.
\]

**Solution:** Using the sum rule in integration,

\[
I_t \pm J_t = \int_0^t f(W_s, s) \, dW_s \pm \int_0^t g(W_s, s) \, dW_s
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \pm \lim_{n \to \infty} \sum_{i=0}^{n-1} g(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \pm g(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right]
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \pm g(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right]
\]
3.2.1 Itô Calculus

\[ I_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ f(W_{t_i}, t_i) \pm g(W_{t_i}, t_i) \right] \left( W_{t_{i+1}} - W_{t_i} \right) \]

\[ = \int_0^t \left[ f(W_s, s) \pm g(W_s, s) \right] dW_s. \]

For constant \( c \),

\[ cI_t = c \int_0^t f(W_s, s) dW_s \]

\[ = c \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \]

\[ = \int_0^t cf(W_s, s) dW_s. \]

The same steps can be applied to show that \( cJ_t = \int_0^t cg(W_s, s) dW_s. \)

\[ \square \]

7. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process such that \( \mathcal{F}_t \) is the associated filtration. The stochastic Itô integral with respect to the standard Wiener process can be defined as

\[ I_t = \int_0^t f(W_s, s) dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \]

where \( f \) is a simple function and \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}. \)

Using the properties of the standard Wiener process, show that \( I_t \) is a martingale.

**Solution:** Given that \( W_t \) is a martingale, we note the following:

(a) Under the filtration \( \mathcal{F}_u, u < t \), by definition

\[ \int_0^t f(W_s, s) dW_s = \int_0^u f(W_s, s) dW_s + \int_u^t f(W_s, s) dW_s \]

\[ = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) + \lim_{n \to \infty} \left[ \sum_{i=m}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right] \]

where \( I_u = \int_0^u f(W_s, s) dW_s = \lim_{n \to \infty} \sum_{i=0}^{m-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right), \quad m < n - 1 \) and

\[ \mathbb{E}(I_u | \mathcal{F}_u) = I_u. \]
Finally, because \( \{ W_t : t \geq 0 \} \) is a martingale we have

\[
\mathbb{E}(I_t | \mathcal{F}_u) = \mathbb{E} \left[ \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_u \right]
\]

\[+ \mathbb{E} \left[ \lim_{n \to \infty} \sum_{i=m}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_u \right]
\]

\[= I_u + \lim_{n \to \infty} \mathbb{E} \left[ \sum_{i=m}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_u \right]
\]

\[= I_u + \lim_{n \to \infty} \sum_{i=m}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_u \right]
\]

\[= I_u + \lim_{n \to \infty} \sum_{i=m}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_u \right]
\]

\[= I_u.
\]

(b) Assuming \( |f(W_t, t)| < \infty \), we have

\[
|I_t| = \lim_{n \to \infty} \left| \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right|
\]

\[\leq \lim_{n \to \infty} \left| \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right|
\]

\[\leq \lim_{n \to \infty} \left[ \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \sum_{i=0}^{n-1} \left| f(W_{t_i}, t_i) \right| \right].
\]

Because \( W_t \) is continuous we have \( \lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| = 0 \), therefore we can deduce that \( \mathbb{E}(|I_t|) < \infty \).

(c) Since \( I_t \) is a function of \( W_t \), hence it is \( \mathcal{F}_t \)-adapted.

From the results of (a)–(c) we have shown that \( I_t \) is a martingale.

\( N.B. \) From the above result we can easily deduce that if \( \{ M_t \}_{t \geq 0} \) is a martingale with continuous sample paths and by defining the following stochastic Itô integral:

\[
I_t = \int_0^t f(M_s, s) \, dM_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(M_{t_i}, t_i) \left( M_{t_{i+1}} - M_{t_i} \right)
\]

where \( f \) is a simple process and \( t_i = it/n \), \( 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \), then \( I_t \) is a martingale.

\( \square \)
8. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. The stochastic Itô integral with respect to a standard Wiener process can be defined as

\[
\int_0^t s \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} t_i (W_{t_{i+1}} - W_{t_i})
\]

where \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\). Prove that

\[
\int_0^t s \, dW_s = tW_t - \int_0^t W_s \, ds.
\]

**Solution:** By definition, \(W_0 = 0\) and we can expand

\[
\int_0^t s \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} t_i (W_{t_{i+1}} - W_{t_i})
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( t_i (W_{t_{i+1}} - W_{t_i}) + t_{i+1}W_{t_{i+1}} - t_{i+1}W_{t_{i+1}} \right)
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( t_{i+1}W_{t_{i+1}} - t_iW_i - (t_{i+1} - t_i) W_{t_{i+1}} \right)
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( t_{i+1}W_{t_{i+1}} - t_iW_i \right) - \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} (t_{i+1} - t_i)
\]

\[
= rW_t - \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} (t_{i+1} - t_i) .
\]

By definition, \(\int_0^t W_s \, ds = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_i (t_{i+1} - t_i)\)

and to show

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} (t_{i+1} - t_i) - \lim_{n \to \infty} \sum_{i=0}^{n-1} W_i (t_{i+1} - t_i) = 0
\]

we note from Problem 2.2.6.2 (page 90) that

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_i) (t_{i+1} - t_i) = 0.
\]

Therefore, we can deduce that

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_{i+1}} (t_{i+1} - t_i) = \int_0^t W_s \, ds .
\]
3.2.1 Itô Calculus

and hence
\[ \int_0^t s \, dW_s = tW_t - \int_0^t W_s \, ds. \]
\[ \square \]

9. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. The stochastic Itô integral with respect to a standard Wiener process can be defined as
\[ \int_0^t W_s^2 \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i}^2 \left( W_{t_{i+1}} - W_{t_i} \right) \]
where \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\). Prove that
\[ \int_0^t W_s^2 \, dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s \, ds. \]

Solution: By definition,
\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i}^2 \left( W_{t_{i+1}} - W_{t_i} \right) = \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{3} \left( W_{t_{i+1}}^3 - W_{t_i}^3 \right) - \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i} \left( W_{t_{i+1}} - W_{t_i} \right)^2 \\
- \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{3} \left( W_{t_{i+1}} - W_{t_i} \right)^3 \\
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{3} \left( W_{t_{i+1}}^3 - W_{t_i}^3 \right) - \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i} \left( t_{i+1} - t_i \right) \\
- \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{1}{3} \left( W_{t_{i+1}} - W_{t_i} \right)^3 \\
= \frac{1}{3} W_t^3 - \int_0^t W_s \, ds
\]

since from Problems 2.2.6.1 (page 89) and 2.2.6.4 (page 92),
\[
\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} W_{t_i} \left( W_{t_{i+1}} - W_{t_i} \right)^2 \right| \\
\leq \lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_k}| \left| \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right)^2 - \sum_{i=0}^{n-1} (t_{i+1} - t_i) \right| \\
= \lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_k}| (t - t_k) \\
= 0
\]
3.2.1 Itô Calculus

and

$$\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^3 = 0.$$ 

Therefore, \( \int_0^t W_s^2 \, dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s \, ds \).

N.B. Instead of going through first principles, we can also show

\( \int_0^t W_s^2 \, dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s \, ds \) by applying Itô’s formula on \( X_t = \frac{1}{3} W_t^3 \), where

\[ dX_t = \frac{\partial X_t}{\partial t} \, dt + \frac{\partial X_t}{\partial W_t} \, dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \, dW_t^2 + \ldots \]

\[ = W_t^2 \, dW_t + W_t \, dt. \]

Taking integrals,

\[ \int_0^t dX_s = \int_0^t W_s^2 \, dW_s + \int_0^t W_s \, ds \]

\[ X_t - X_0 = \int_0^t W_s^2 \, dW_s + \int_0^t W_s \, ds \]

and since \( W_0 = 0 \), therefore \( \int_0^t W_s^2 \, dW_s = \frac{1}{3} W_t^3 - \int_0^t W_s \, ds \). \[\Box\]

10. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. The stochastic Itô integral with respect to a standard Wiener process can be defined as

\[ I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \]

where \(f\) is a simple process and \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n \leq t, n \in \mathbb{N}\). Using the properties of a standard Wiener process, show that

\[ \mathbb{E} \left[ \int_0^t f(W_s, s) \, dW_s \right] = 0 \]

and deduce that if \(\{M_t : t \geq 0\}\) is a martingale then

\[ \mathbb{E} \left[ \int_0^t g(M_s, s) \, dM_s \right] = 0 \]

where \(g\) is a simple process.
3.2.1 Itô Calculus

Solution: By definition,
\[
\mathbb{E} \left[ \int_0^t f(W_s, s) \, dW_s \right] = \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E}\left[f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \right]
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E}\left[f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right) \bigg| \mathcal{F}_{t_i} \right]
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E}\left[f(W_{t_i}, t_i) (W_{t_{i+1}} - W_{t_i}) \right]
\]
\[
= 0
\]

since \( \{ W_t : t \geq 0 \} \) is a martingale and therefore \( \mathbb{E} \left( W_{t_{i+1}} \big| \mathcal{F}_{t_i} \right) = W_{t_i} \).

Using the same steps as described above, if \( \{ M_t : t \geq 0 \} \) is a martingale then
\[
\mathbb{E} \left[ \int_0^t g(M_s, s) \, dM_s \right] = 0
\]
where \( g \) is a simple process.

\[\square\]

11. Itô Isometry. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. The stochastic Itô integral with respect to a standard Wiener process can be defined as
\[
I_t = \int_0^t f(W_s, s) \, dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) \left( W_{t_{i+1}} - W_{t_i} \right)
\]

where \( f \) is a simple process and \( t_i = t \cdot n \), \( 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t \), \( n \in \mathbb{N} \).

Using the properties of a standard Wiener process, show that
\[
\mathbb{E} \left[ \left( \int_0^t f(W_s, s) \, dW_s \right)^2 \right] = \mathbb{E} \left[ \int_0^t f(W_s, s)^2 \, ds \right].
\]

Solution: Using the property of independent increments of a standard Wiener process as well as the martingale properties of \( W_t \) and \( W_t^2 - t \), we note that
\[
\mathbb{E} \left[ \left( \int_0^t f(W_s, s) \, dW_s \right)^2 \right] - \mathbb{E} \left[ \int_0^t f(W_s, s)^2 \, ds \right]
\]
\[
= \mathbb{E} \left\{ \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i) (W_{t_{i+1}} - W_{t_i}) \right\}^2 - \mathbb{E} \left[ \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i)^2 (t_{i+1} - t_i) \right]
\]
\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i)^2 (W_{t_{i+1}} - W_{t_i})^2 \right] - \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i)^2 (t_{i+1} - t_i) \right]
\]
3.2.1 Itô Calculus

\[
\begin{align*}
&= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( (W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \right] \\
&= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( (W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right) \right] \\
&= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( W_{t_{i+1}}^2 - 2W_{t_{i+1}}W_{t_i} + W_{t_i}^2 - t_{i+1} + t_i \right) \right] \\
&= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(W_{t_i}, t_i) \left( W_{t_i}^2 - t_i - 2W_{t_i}^2 + W_{t_i}^2 + t_i \right) \right] \\
&= 0.
\end{align*}
\]

Therefore, \( \mathbb{E} \left[ \left( \int_0^t f(W_s, s) \, dW_s \right)^2 \right] = \mathbb{E} \left[ \int_0^t f(W_s, s)^2 \, ds \right] \).

\[
\square
\]

12. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. The stochastic Itô integral with respect to a standard Wiener process can be defined as

\[
I_t = \int_0^t f(W_s, s) \, dW_s
\]

where \(f\) is a simple process. Show that the Itô integral has quadratic variation process \(\langle I, I \rangle_t\) given by

\[
\langle I, I \rangle_t = \int_0^t f(W_s, s)^2 \, ds.
\]

Solution: By definition,

\[
\langle I, I \rangle_t = \lim_{m \to \infty} \sum_{k=0}^{m-1} (t_{k+1} - t_k)^2
\]

where \(t_k = kt/m, 0 = t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m = t, m \in \mathbb{N}\). We first concentrate on one of the subintervals \([I_k, I_k+1)\) on which \(f(W_s, s) = f(W_{t_k}, t_k)\) is a constant value for all \(s \in [I_k, I_{k+1})\). Partitioning

\[
I_k = s_0 < s_1 < \ldots < s_n = I_{k+1},
\]

we can write

\[
I_{s_{i+1}} - I_{s_i} = \int_{s_i}^{s_{i+1}} f(W_{t_k}, t_k) \, dW_s = f(W_{t_k}, t_k) \left( W_{s_{i+1}} - W_{s_i} \right).
\]
3.2.1 Ito Calculus

Therefore,

\[
(I_{k+1} - I_k)^2 = \lim_{n \to \infty} \sum_{i=0}^{n-1} (I_{s_{i+1}} - I_{s_i})^2
\]

\[
= \lim_{n \to \infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_k) \left( W_{s_{i+1}} - W_{s_i} \right)^2
\]

\[
= (W_{t_k}, t_k)^2 \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{s_{i+1}} - W_{s_i})^2
\]

\[
= (W_{t_k}, t_k)^2 \left( t_{k+1} - t_k \right)
\]

Since \( \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{s_{i+1}} - W_{s_i})^2 \) converges to the quadratic variation of a standard Wiener process over \([t_k, t_{k+1})\), which is \( t_{k+1} - t_k \). Therefore,

\[
(I_{k+1} - I_k)^2 = f(W_{t_k}, t_k)^2 \left( t_{k+1} - t_k \right) = \int_{t_k}^{t_{k+1}} f(W_s, s)^2 \, ds
\]

where \( f(W_s, s) \) is a constant value for all \( s \in [t_k, t_{k+1}) \).

Finally, to obtain the quadratic variation of the Itô integral \( I_t \),

\[
\langle I, I \rangle_t = \lim_{m \to \infty} \sum_{k=0}^{m-1} (I_{s_{k+1}} - I_{s_k})^2 = \lim_{m \to \infty} \sum_{k=0}^{m-1} f(W_{t_k}, t_k)^2 \left( t_{k+1} - t_k \right) = \int_0^t f(W_s, s)^2 \, ds
\]

N.B. In differential form we can write \( d\langle I, I \rangle_t = dI_t dI_t = f(W_t, t)^2 \, dt \). By comparing the results of the quadratic variation \( \langle I, I \rangle_t \) and \( \mathbb{E}(I_t^2) \), we can see that the former is computed path-by-path and hence the result is random. In contrast, the variance of the Itô integral, \( \text{Var}(I_t) = \mathbb{E}(I_t^2) = \mathbb{E}(\langle I, I \rangle_t) \), is the mean value of all possible paths of the quadratic variation and hence is non-random.

\[ \square \]

13. Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process.

Using integration by parts, show that

\[
\int_0^t W_s \, ds = \int_0^t (t - s) \, dW_s
\]

and prove that \( \int_0^t W_s \, ds \sim \mathcal{N} \left( 0, \frac{t^3}{3} \right) \).

Is

\[
B_t = \begin{cases} 
0 & t = 0 \\
\frac{\sqrt{3}}{t} \int_0^t W_s \, ds & t > 0
\end{cases}
\]

a standard Wiener process?
3.2.1 Ito Calculus

**Solution:** Using integration by parts,
\[
\int u \left( \frac{dv}{ds} \right) ds = uv - \int v \left( \frac{du}{ds} \right) ds.
\]
We set \( u = W_s \) and \( dv/ ds = 1 \). Therefore,
\[
\int_0^t W_s \, ds = \int_0^t W_s \, ds - \int_0^t s \, dW_s = tW_t - \int_0^t s \, dW_s = \int_0^t (t - s) \, dW_s.
\]
Taking expectations,
\[
\mathbb{E} \left( \int_0^t W_s \, ds \right) = \mathbb{E} \left( \int_0^t (t - s) \, dW_s \right) = 0
\]
\[
\mathbb{E} \left[ \left( \int_0^t W_s \, ds \right)^2 \right] = \mathbb{E} \left( \int_0^t (t - s)^2 \, ds \right) = \frac{t^3}{3}.
\]
To show that \( \int_0^t W_s \, ds \) follows a normal distribution, let
\[
M_t = \int_0^t W_s \, ds = \int_0^t (t - s) \, dW_s
\]
and using the properties of the Ito integral (see Problems 3.2.1.7, page 108 and 3.2.1.12, page 114), we can deduce that
\[
M_t \text{ is a martingale and } \langle M, M \rangle_t = \int_0^t (t - s)^2 \, ds = \frac{t^3}{3}
\]
so that \( dM_t \cdot dM_t = t^2 \, dt \).
By defining
\[
Z_t = e^{\theta M_t - \frac{1}{2} \theta^2 \langle M \rangle_t}, \quad \theta \in \mathbb{R}
\]
then expanding \( dZ_t \) using Taylor’s theorem and applying Ito’s lemma, we have
\[
dZ_t = \frac{\partial Z_t}{\partial t} \, dt + \frac{\partial Z_t}{\partial M_t} \, dM_t + \frac{1}{2} \frac{\partial^2 Z_t}{\partial M_t^2} (dM_t)^2 + \ldots
\]
\[
= -\frac{1}{2} \theta^2 t^2 Z_t \, dt + \theta Z_t \, dM_t + \frac{1}{2} \theta^2 Z_t \, (dM_t)^2 + \ldots
\]
\[
= -\frac{1}{2} \theta^2 t^2 Z_t \, dt + \theta Z_t \, dM_t + \frac{1}{2} \theta^2 \frac{t^3}{3} \, dt
\]
\[
= \theta Z_t \, dM_t.
\]
3.2.1 Itô Calculus

Taking integrals, we can express

\[
\int_0^t dZ_u = \int_0^t \theta Z_u dM_u
\]

\[
Z_t - Z_0 = \theta \int_0^t Z_u dM_u
\]

\[
Z_t = 1 + \theta \int_0^t Z_u dM_u.
\]

Finally, by taking expectations and knowing that \(M_t\) is a martingale, we have

\[
\mathbb{E}(Z_t) = 1 + \theta \mathbb{E} \left( \int_0^t Z_u dM_u \right) = 1
\]

and hence

\[
\mathbb{E}(e^{\theta M_t}) = e^{\frac{1}{2} \theta^2 \left( \frac{t}{3} \right)}
\]

which is the moment generating function of a normal distribution with mean zero and variance \(\frac{t}{3}\). Thus, \(\int_0^t W_s ds \sim \mathcal{N}(0, \frac{t^3}{3})\).

Even though \(B_t \sim \mathcal{N}(0, t)\) for \(t > 0\), \(B_t\) is not a standard Wiener process since for \(t, u > 0\),

\[
\mathbb{E}(B_{t+u} - B_t) = \mathbb{E}(B_{t+u}) - \mathbb{E}(B_t)
\]

\[
= \mathbb{E} \left( \frac{\sqrt{3}}{t+u} \int_0^{t+u} W_s ds \right) - \mathbb{E} \left( \frac{\sqrt{3}}{t} \int_0^t W_s ds \right)
\]

\[
= 0
\]

and using the result of Problem 2.2.1.13 (page 66),

\[
\text{Var}(B_{t+u} - B_t) = \text{Var}(B_{t+u}) + \text{Var}(B_t) - 2\text{Cov}(B_{t+u}, B_t)
\]

\[
= \text{Var} \left( \frac{\sqrt{3}}{t+u} \int_0^{t+u} W_s ds \right) + \text{Var} \left( \frac{\sqrt{3}}{t} \int_0^t W_s ds \right)
\]

\[
-2\text{Cov} \left( \frac{\sqrt{3}}{t+u} \int_0^{t+u} W_s ds, \frac{\sqrt{3}}{t} \int_0^t W_s ds \right)
\]

\[
= t + u + t - \frac{6}{rt(t+u)} \left[ \frac{1}{3} t^3 + \frac{1}{2} u^2 \right]
\]

\[
= 2t + u - \frac{t(2t+3u)}{t+u}
\]

\[
= \frac{u^2}{t+u}
\]

\[
\neq u
\]
which shows that $B_t$ does not have the stationary increment property. Therefore, $B_t$ is not a standard Wiener process.

\[ \Box \]

14. Generalised Itô Integral. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Given that $f$ is a simple process, show

\[
\int_0^t f(W_s, s) \, dW_s = W_t f(W_t, t) - \int_0^t \left[ W_s \frac{\partial f}{\partial t}(W_s, s) + \frac{\partial f}{\partial W_t}(W_s, s) + \frac{1}{2} W_s \frac{\partial^2 f}{\partial W_t^2}(W_s, s) \right] \, ds
\]

and

\[
\int_0^t s \, dW_s
\]

\[
\int_0^t f(W_s, s) \, ds = tf(W_t, t) - \int_0^t s \left[ \frac{\partial f}{\partial t}(W_s, s) + \frac{1}{2} \frac{\partial^2 f}{\partial W_t^2}(W_s, s) \right] \, ds
\]

and

\[
\int_0^t \frac{\partial f}{\partial W_t}(W_s, s) \, dW_s
\]

Solution: For the first result, using Taylor’s theorem on $d(W_t f(W_t, t))$ and subsequently applying Itô’s formula we have

\[
d(W_t f(W_t, t)) = W_t \frac{\partial f}{\partial t}(W_t, t) \, dt + \left[ f(W_t, t) + W_t \frac{\partial f}{\partial W_t}(W_t, t) \right] \, dW_t
\]

\[
+ \frac{1}{2} \left[ \frac{\partial f}{\partial W_t}(W_t, t) + \frac{\partial f}{\partial W_t}(W_t, t) + W_t \frac{\partial^2 f}{\partial W_t^2}(W_t, t) \right] \, dW_t^2 + \ldots
\]

\[
= W_t \frac{\partial f}{\partial t}(W_t, t) \, dt + \left[ f(W_t, t) + W_t \frac{\partial f}{\partial W_t}(W_t, t) \right] \, dW_t
\]

\[
+ \frac{1}{2} \left[ 2 \frac{\partial f}{\partial W_t}(W_t, t) + W_t \frac{\partial^2 f}{\partial W_t^2}(W_t, t) \right] \, dt
\]

\[
= \left[ W_t \frac{\partial f}{\partial t}(W_t, t) + \frac{\partial f}{\partial W_t}(W_t, t) + \frac{1}{2} W_t \frac{\partial^2 f}{\partial W_t^2}(W_t, t) \right] \, dt
\]

\[
+ \left[ f(W_t, t) + W_t \frac{\partial f}{\partial W_t}(W_t, t) \right] \, dW_t
\]

Taking integrals from 0 to $t$,

\[
\int_0^t d(W_s f(W_s, s)) = \int_0^t \left[ W_s \frac{\partial f}{\partial t}(W_s, s) + \frac{\partial f}{\partial W_t}(W_s, s) + \frac{1}{2} W_s \frac{\partial^2 f}{\partial W_t^2}(W_s, s) \right] \, ds
\]

\[
+ \int_0^t \left[ f(W_s, s) + W_s \frac{\partial f}{\partial W_t}(W_s, s) \right] \, dW_s
\]
3.2.1 Itô Calculus

and rearranging the terms, finally

\[
\int_0^t f(W_s, s) \, dW_s = W_t f(W_t, t) - \int_0^t \left[ W_s \frac{\partial f}{\partial t}(W_s, s) + \frac{\partial f}{\partial W_s}(W_s, s) + \frac{1}{2} W_s \frac{\partial^2 f}{\partial W_s^2}(W_s, s) \right] \, ds
\]

\[
- \int_0^t W_s \frac{\partial f}{\partial W_s}(W_s, s) \, dW_s
\]

since \(W_0 = 0\).

As for the second result, from Taylor’s theorem and Itô’s formula

\[
d(f(W_t, t)) = f(W_t, t) \, dt + t \frac{\partial f}{\partial W_t}(W_t, t) \, dW_t + \frac{1}{2} t^2 \frac{\partial^2 f}{\partial W_t^2}(W_t, t) \, dW_t^2 + \ldots
\]

\[
= f(W_t, t) \, dt + t \frac{\partial f}{\partial W_t}(W_t, t) \, dW_t + \frac{1}{2} t^2 \frac{\partial^2 f}{\partial W_t^2}(W_t, t) \, dt
\]

Taking integrals from 0 to \(t\) we have

\[
\int_0^t d(f(W_s, s)) = \int_0^t \left[ f(W_s, s) + s \frac{\partial f}{\partial t}(W_s, s) + \frac{1}{2} s^2 \frac{\partial^2 f}{\partial W_s^2}(W_s, s) \right] \, ds
\]

\[
+ \int_0^t s \frac{\partial f}{\partial W_s}(W_s, s) \, dW_s
\]

and hence

\[
\int_0^t f(W_s, s) \, ds = f(W_t, t) - \int_0^t \left[ \frac{\partial f}{\partial t}(W_s, s) + \frac{1}{2} s \frac{\partial^2 f}{\partial W_s^2}(W_s, s) \right] \, ds
\]

\[
- \int_0^t s \frac{\partial f}{\partial W_s}(W_s, s) \, dW_s.
\]

\(\square\)

15. **One-Dimensional Lévy Characterisation Theorem.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{M_t : t \geq 0\}\) be a martingale with respect to the filtration \(\mathcal{F}_t, t \geq 0\). By assuming \(M_0 = 0, M_t\) has continuous sample paths whose quadratic variation

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})^2 = t
\]

where \(t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N}\) show that \(M_t\) is a standard Wiener process.
Solution: We first need to show that $M_t \sim \mathcal{N}(0, t)$ or, using the moment generating function approach, we need to show that $\mathbb{E}(e^{\theta M_t}) = e^{\frac{\theta^2}{2} t}$ for a constant $\theta$.

Let $f(M_t, t) = e^{\theta M_t - \frac{1}{2} \theta^2 t}$ and since $dM_t \cdot dM_t = dt$ and $(dt)^\nu = 0$, $\nu \geq 2$ from Itô’s formula,

$$df(M_t, t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial M_t} dM_t + \frac{1}{2} \frac{\partial^2 f}{\partial M_t^2} (dM_t)^2 + \ldots$$

$$= \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial M_t^2} \right) dt + \frac{\partial f}{\partial M_t} dM_t$$

$$= \left( -\frac{1}{2} \theta^2 t + \frac{1}{2} \theta^2 t \right) f(M_t, t) dt + \theta f(M_t, t) dM_t$$

Taking integrals from 0 to $t$, and then taking expectations, we have

$$\int_0^t df(M_s, s) = \theta \int_0^t f(M_s, s) dM_s$$

$$f(M_t, t) - f(M_0, 0) = \theta \int_0^t f(M_s, s) dM_s$$

$$\mathbb{E}(f(M_t, t)) = 1 + \theta \mathbb{E} \left[ \int_0^t f(M_s, s) dM_s \right].$$

By definition of the stochastic Itô integral we can write

$$\mathbb{E} \left[ \int_0^t f(M_s, s) dM_s \right] = \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(M_{t_i}, t_i) \left( M_{t_{i+1}} - M_{t_i} \right) \right]$$

$$= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ \mathbb{E} \left[ f(M_{t_i}, t_i) \left( M_{t_{i+1}} - M_{t_i} \right) \bigg| \mathcal{F}_{t_i} \right] \right]$$

$$= \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathbb{E} \left[ f(M_{t_i}, t_i) \left( M_{t_i} - M_{t_i} \right) \right]$$

$$= 0$$

since $\{M_t\}_{t \geq 0}$ is a martingale and hence

$$\mathbb{E}(f(M_t, t)) = 1$$

or

$$\mathbb{E}(e^{\theta M_t}) = e^{\frac{\theta^2}{2} t}$$

which is the moment generating function for the normal distribution with mean zero and variance $t$. Therefore, $M_t \sim \mathcal{N}(0, t)$. 
3.2.1 Itô Calculus

Since \( M_t \) is a martingale, for \( s < t \)

\[
\mathbb{E}(M_t | \mathcal{F}_s) = \mathbb{E}(M_t - M_s + M_s | \mathcal{F}_s)
\]

\[
= \mathbb{E}(M_t - M_s | \mathcal{F}_s) + \mathbb{E}(M_s | \mathcal{F}_s)
\]

\[
= \mathbb{E}(M_t - M_s | \mathcal{F}_s) + M_s
\]

Therefore, \( \mathbb{E}(M_t - M_s | \mathcal{F}_s) = \mathbb{E}(M_t - M_s) = 0 \) and hence \( M_t - M_s \perp \mathcal{F}_s \). So, we have shown that \( M_t \) has the independent increment property.

Finally, to show that \( M_t \) has the stationary increment, for \( t > 0 \) and \( s > 0 \) we have

\[
\mathbb{E}(M_{t+s} - M_t) = \mathbb{E}(M_{t+s}) - \mathbb{E}(M_t) = 0
\]

and using the independent increment property of \( M_t \)

\[
\text{Var}(M_{t+s} - M_t) = \text{Var}(M_{t+s}) + \text{Var}(M_t) - 2 \text{Cov}(M_{t+s}, M_t)
\]

\[
= t + s + t - 2 \left[ \mathbb{E}(M_{t+s} M_t) - \mathbb{E}(M_{t+s}) \mathbb{E}(M_t) \right]
\]

\[
= 2t + s - 2 \mathbb{E}(M_{t+s} M_t)
\]

\[
= 2t + s - 2 \mathbb{E}(M_t (M_{t+s} - M_t) + M_t^2)
\]

\[
= 2t + s - 2 \mathbb{E}(M_t) \mathbb{E}(M_{t+s} - M_t) - 2 \mathbb{E}(M_t^2)
\]

\[
= 2t + s - 2t
\]

\[
= s.
\]

Therefore, \( M_{t+s} - M_t \sim \mathcal{N}(0, s) \).

Because \( M_0 = 0 \) and also \( M_t \) has continuous sample paths with independent and stationary increments, so \( M_t \) is a standard Wiener process.

\[
\square
\]

16. Multi-Dimensional Lévy Characterisation Theorem. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \{M^{(1)}_t : t \geq 0\}, \{M^{(2)}_t : t \geq 0\}, \ldots, \{M^{(n)}_t : t \geq 0\} \) be martingales with respect to the filtration \( \mathcal{F}_t, t \geq 0 \). By assuming \( M^{(0)}_0 = 0 \), \( M^{(i)}_t \) has continuous sample paths whose quadratic variation

\[
\lim_{n \to \infty} \sum_{k=0}^{m} \left( M^{(i)}_{t_{k+1}} - M^{(i)}_{t_k} \right)^2 = t
\]

and cross-variation between \( M^{(i)}_t \) and \( M^{(j)}_t \), \( i \neq j, i, j = 1, 2, \ldots, n \)

\[
\lim_{m \to \infty} \sum_{k=0}^{m} \left( M^{(i)}_{t_{k+1}} - M^{(i)}_{t_k} \right) \left( M^{(j)}_{t_{k+1}} - M^{(j)}_{t_k} \right) = 0
\]

where \( t_k = kt/m, \ 0 = t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m = t, \ m \in \mathbb{N}, \) show that \( M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t \) are independent standard Wiener processes.
Solution: Following Problem 3.2.1.15 (page 119), we can easily prove that
$M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$ are standard Wiener processes. In order to show $M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$ are mutually independent, we need to show the joint moment generating function of $M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$

$$M^{(n)}_t$$ is

$$\mathbb{E}\left(e^{\theta^{(1)}M^{(1)}_t+\theta^{(2)}M^{(2)}_t+\ldots+\theta^{(n)}M^{(n)}_t}\right) = e^{\frac{1}{2}(\theta^{(1)})^2t} \cdot e^{\frac{1}{2}(\theta^{(2)})^2t} \cdot \ldots \cdot e^{\frac{1}{2}(\theta^{(n)})^2t}$$

for constants $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(n)}$.

Let $f(M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t, t) = \prod_{i=1}^{n} e^{\theta^{(i)}M^{(i)}_t - \frac{1}{2}(\theta^{(i)})^2t}$ and since $dM^{(i)}_t \cdot dM^{(j)}_t = dt$, $dM^{(i)}_t \cdot dM^{(j)}_t = 0, i, j = 1, 2, \ldots, n, i \neq j$ and $(dt)^v = 0, v \geq 2$, from Itô’s formula

$$df(M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t, t) = \frac{\partial f}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial f}{\partial M^{(i)}_t} dM^{(i)}_t$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial M^{(i)}_t \partial M^{(j)}_t} dM^{(i)}_t dM^{(j)}_t$$

$$= \frac{\partial f}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial f}{\partial M^{(i)}_t} dM^{(i)}_t$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial M^{(i)}_t \partial M^{(j)}_t} \cdot \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial (M^{(i)}_t)^2} (dM^{(i)}_t)^2$$

$$= \left( \frac{\partial f}{\partial t} + \sum_{i=1}^{n} \frac{\partial f}{\partial M^{(i)}_t} \frac{\partial^2 f}{\partial (M^{(i)}_t)^2} \right) dt + \sum_{i=1}^{n} \frac{\partial f}{\partial M^{(i)}_t} dM^{(i)}_t$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial M^{(i)}_t} dM^{(i)}_t$$

$$= \sum_{i=1}^{n} \theta^{(i)} f(M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t, t) dM^{(i)}_t$$

since

$$\frac{\partial f}{\partial t} + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial (M^{(i)}_t)^2} = \left[ -\frac{1}{2} \sum_{i=1}^{n} (\theta^{(i)})^2 + \frac{1}{2} \sum_{i=1}^{n} (\theta^{(i)})^2 \right] f(M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t, t) = 0.$$
3.2.2 One-Dimensional Diffusion Process

Because $M^{(i)}_t$ is a martingale, we can easily show (see Problem 3.2.1.15, page 119) that

$$
\mathbb{E} \left[ \int_0^t f \left( M^{(1)}_s, M^{(2)}_s, \ldots, M^{(n)}_s, t \right) \, dM^{(i)}_s \right] = 0
$$

for $i = 1, 2, \ldots, n$ and hence

$$
\mathbb{E} \left( e^{\theta^{(1)} M^{(1)}_t + \theta^{(2)} M^{(2)}_t + \ldots + \theta^{(n)} M^{(n)}_t} \right) = e^{\frac{1}{2} \theta^{(1)}^2 t \cdot e^{\frac{1}{2} \theta^{(2)}^2 t} \cdot \ldots \cdot e^{\frac{1}{2} \theta^{(n)}^2 t}}
$$

where the joint moment generating function of $M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$ is a product of moment generating functions of $M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$. Therefore, $M^{(1)}_t, M^{(2)}_t, \ldots, M^{(n)}_t$ are independent standard Wiener processes.

\[\square\]

3.2.2 One-Dimensional Diffusion Process

1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Find the SDE for the random process $X_t = W^n_t, n \in \mathbb{Z}^+$. 

   Show that

   $$
   \mathbb{E} \left( W^n_t \right) = \frac{1}{2} n(n-1) \int_0^t \mathbb{E} \left( W^{(n-2)}_s \right) \, ds
   $$

   and using mathematical induction prove that

   $$
   \mathbb{E} \left( W^n_t \right) = \begin{cases} 
   \frac{n! \left( \frac{n}{2} \right)^n}{2^n} & n = 2, 4, 6, \ldots \\
   0 & n = 1, 3, 5, \ldots
   \end{cases}
   $$

   **Solution:** By expanding $dX_t$ using Taylor's formula and applying Itô's formula,

   $$
   dX_t = \frac{\partial X_t}{\partial t} \, dt + \frac{\partial X_t}{\partial W_t} \, dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \, dW_t^2 + \ldots
   $$

   $$
   dW^n_t = nW_t^{(n-1)} \, dW_t + \frac{1}{2} n(n-1)W_t^{(n-2)} \, dt.
   $$

   Taking integrals,

   $$
   \int_0^t dW^n_t = \int_0^t nW_s^{(n-1)} \, dW_s + \frac{1}{2} n(n-1) \int_0^t W_s^{(n-2)} \, ds
   $$

   $$
   W^n_t = \int_0^t nW_s^{(n-1)} \, dW_s + \frac{1}{2} n(n-1) \int_0^t W_s^{(n-2)} \, ds.
   $$

   Finally, by taking expectations,

   $$
   \mathbb{E} \left( W^n_t \right) = \frac{1}{2} n(n-1) \int_0^t \mathbb{E} \left( W_s^{(n-2)} \right) \, ds.
   $$
To prove the final result, we will divide it into two sections, one for even numbers \( n = 2k, \ k \in \mathbb{Z}^+ \) and another for odd numbers \( n = 2k + 1, \ k \in \mathbb{Z}^+ \). We note that for \( n = 2 \) we have
\[
\mathbb{E} \left( W_2^2 \right) = \frac{2!}{2} = t
\]
and because \( W_t \sim \mathcal{N}(0, t) \), the result is true for \( n = 2 \).

We assume that the result is true for \( n = 2k, \ k \in \mathbb{Z}^+ \). That is
\[
\mathbb{E} \left( W_{2k}^2 \right) = \frac{(2k)!k}{2^kk!}.
\]

For \( n = 2(k + 1), \ k \in \mathbb{Z}^+ \) we have
\[
\mathbb{E} \left( W_{2(k+1)}^2 \right) = \frac{1}{2}(2k + 2)(2k + 1) \int_0^t \mathbb{E} \left( W_{2k}^2 \right) \, ds
\]
\[
= \frac{1}{2}(2k + 2)(2k + 1) \int_0^t \frac{(2k)!s}{2^kk!} \, ds
\]
\[
= \frac{(2k + 2)!}{2^kk!(k+1)!} \int_0^t s^k \, ds
\]
\[
= \frac{(2k + 2)!k^{k+1}}{2^kk!(k+1)!}
\]
\[
= \frac{(2k + 2)!}{2^kk!(k+1)!} \frac{2(k+1)}{2} = \frac{(2k + 2)!}{2^kk!(k+1)!} \frac{2(k+1)}{2}.
\]
Thus, the result is also true for \( n = 2(k + 1), \ k \in \mathbb{Z}^+ \).

For \( n = 1 \), we have
\[
\mathbb{E} \left( W_1 \right) = 0
\]
and because \( W_t \sim \mathcal{N}(0, t) \), the result is true for \( n = 1 \).

We assume the result is true for \( n = 2k + 1, \ k \in \mathbb{Z}^+ \) such that
\[
\mathbb{E} \left( W_{2k+1}^2 \right) = 0.
\]

For \( n = 2(k + 1) + 1, \ k \in \mathbb{Z}^+ \)
\[
\mathbb{E} \left( W_{2(k+1)+1}^2 \right) = \frac{1}{2}(2k + 3)(2k + 2) \int_0^t \mathbb{E} \left( W_{2k+1}^2 \right) \, ds = 0
\]
and hence the result is also true for \( n = 2(k + 1) + 1, \ k \in \mathbb{Z}^+ \). Therefore, by mathematical induction
\[
\mathbb{E} \left( W_n^2 \right) = \begin{cases} 
\frac{n!t^{\frac{n}{2}}}{2^\frac{n}{2} \left( \frac{n}{2} \right)!} & n = 2, 4, 6, \ldots \\
0 & n = 1, 3, 5, \ldots 
\end{cases}
\]
\(\square\)
2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let \{\(W_t : t \geq 0\)\} be a standard Wiener process. For a constant \(\theta\) find the SDE for the random process \(X_t = e^{\theta W_t - \frac{1}{2} \theta^2 t}\).

By writing the SDE in integral form calculate \(\mathbb{E}(e^{\theta W_t})\), the moment generating function of a standard Wiener process.

**Solution:** Expanding \(dX_t\) using Taylor’s theorem and applying Itô’s formula,

\[
dX_t = \frac{\partial X_t}{\partial t} dt + \frac{\partial X_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} dW^2_t + \ldots
\]

\[
= \left( \frac{\partial X_t}{\partial t} + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \right) dt + \frac{\partial X_t}{\partial W_t} dW_t
\]

\[
= (\frac{1}{2} \theta^2 X_t + \frac{1}{2} \theta^2 X_t) dt + \theta X_t dW_t
\]

\[
= \theta X_t dW_t.
\]

Taking integrals, we have

\[
\int_0^t dX_s = \int_0^t \theta X_s dW_s
\]

\[
X_t - X_0 = \int_0^t \theta X_s dW_s.
\]

Since \(X_0 = 1\) and taking expectations,

\[
\mathbb{E} \left( X_t \right) - 1 = \mathbb{E} \left( \int_0^t \theta X_s dW_s \right) = 0
\]

so

\[
\mathbb{E} \left( X_t \right) = 1 \quad \text{or} \quad \mathbb{E} \left( e^{\theta W_t - \frac{1}{2} \theta^2 t} \right) = 1.
\]

Therefore, \(\mathbb{E}(e^{\theta W_t}) = e^{\frac{1}{2} \theta^2 t}\). \(\square\)

3. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \{\(W_t : t \geq 0\)\} be a standard Wiener process. Consider the process

\[
Z_t = e^{\theta W_t}
\]

where \(\theta\) is a constant parameter.

Using Itô’s formula, find an SDE for \(Z_t\).

By setting \(m_t = \mathbb{E}(e^{\theta W_t})\) show that the integrated SDE can be expressed as

\[
\frac{dm_t}{dt} - \frac{1}{2} \theta^2 m_t = 0.
\]

Given \(W_0 = 0\), solve the first-order ordinary differential equation to find \(m_t\).
Solution: Using Itô’s formula we expand $Z_t$ as

$$dZ_t = \frac{\partial Z_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 Z_t}{\partial W_t^2} (dW_t)^2 + \ldots$$

$$= \theta e^{\theta W_t} dW_t + \frac{1}{2} \theta^2 e^{\theta W_t} dt$$

$$= \theta Z_t dW_t + \frac{1}{2} \theta^2 Z_t dt.$$

Taking integrals and then expectations,

$$\int_0^t dZ_s = \int_0^t \theta Z_s dW_s + \int_0^t \frac{1}{2} \theta^2 Z_s ds$$

$$Z_t - Z_0 = \int_0^t \theta Z_s dW_s + \int_0^t \frac{1}{2} \theta^2 Z_s ds$$

$$\mathbb{E}(Z_t) - \mathbb{E}(Z_0) = \mathbb{E} \left( \int_0^t \theta Z_s dW_s \right) + \mathbb{E} \left( \int_0^t \frac{1}{2} \theta^2 Z_s ds \right)$$

$$\mathbb{E}(Z_t) - 1 = \int_0^t \frac{1}{2} \theta^2 \mathbb{E}(Z_s) ds$$

where $Z_0 = 1$ and $\mathbb{E} \left( \int_0^t \theta Z_s dW_s \right) = 0$. By differentiating the integral equation we have

$$\frac{d\mathbb{E}(Z_t)}{dt} = \frac{d}{dt} \int_0^t \frac{1}{2} \theta^2 \mathbb{E}(Z_s) ds$$

$$\frac{d\mathbb{E}(Z_t)}{dt} = \frac{1}{2} \theta^2 \mathbb{E}(Z_t)$$

or

$$\frac{dm_t}{dt} - \frac{1}{2} \theta^2 m_t = 0.$$

Setting the integrating factor as $I = e^{-\int^t_0 \frac{1}{2} \theta^2 dt} = e^{-\frac{1}{2} \theta^2 t}$ and multiplying the differential equation with $I$, we have

$$\frac{d}{dt} \left( m_t e^{-\frac{1}{2} \theta^2 t} \right) = 0 \quad \text{or} \quad e^{-\frac{1}{2} \theta^2 t} \mathbb{E} \left( e^{\theta W_t} \right) = C$$

where $C$ is a constant. Since $\mathbb{E} \left( e^{\theta W_0} \right) = 1$, so $C = 1$ and hence we finally obtain $\mathbb{E} \left( e^{\theta W_t} \right) = e^{\frac{1}{2} \theta^2 t}$. \hfill \Box

4. Let $M_t = \int_0^t f(s) dW_s$ and show that the SDE satisfied by

$$X_t = \exp \left\{ \theta M_t - \frac{1}{2} \theta^2 \int_0^t f(s)^2 ds \right\}$$

is

$$dX_t = \theta f(t) X_t dW_t$$
and show also that \( M_t \sim \mathcal{N} \left( 0, \int_0^t f(s)^2 \, ds \right) \).

**Solution:** From Itô’s formula,

\[
\begin{align*}
\frac{dX_t}{dt} & = \frac{\partial X_t}{\partial t} \, dt + \frac{\partial X_t}{\partial W_t} \, dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \, dW_t^2 + \ldots \\
& = \frac{\partial X_t}{\partial t} \, dt + \left( \frac{\partial X_t}{\partial M_t} \cdot \frac{\partial M_t}{\partial W_t} \right) \, dW_t + \frac{1}{2} \frac{\partial}{\partial W_t} \left( \frac{\partial X_t}{\partial M_t} \cdot \frac{\partial M_t}{\partial W_t} \right) \, dt \\
& = \left[ \frac{\partial X_t}{\partial t} + \frac{1}{2} \frac{\partial}{\partial W_t} \left( \frac{\partial X_t}{\partial M_t} \cdot \frac{\partial M_t}{\partial W_t} \right) \right] \, dt + \left( \frac{\partial X_t}{\partial M_t} \cdot \frac{\partial M_t}{\partial W_t} \right) \, dW_t.
\end{align*}
\]

Since

\[
\frac{\partial X_t}{\partial M_t} = \theta \exp \left\{ \theta M_t - \frac{1}{2} \int_0^t f(s)^2 \, ds \right\} = \theta X_t, \quad \frac{\partial M_t}{\partial W_t} = f(t)
\]

and

\[
\frac{\partial}{\partial W_t} \left( \frac{\partial X_t}{\partial M_t} \cdot \frac{\partial M_t}{\partial W_t} \right) = \theta^2 f(t)^2 X_t, \quad \frac{\partial X_t}{\partial t} = -\frac{1}{2} \theta^2 f(t)^2 X_t
\]

so

\[
dX_t = \theta f(t) X_t \, dW_t.
\]

Writing in integral form and then taking expectations,

\[
\begin{align*}
\int_0^t dX_s &= \int_0^t \theta f(s) X_s \, dW_s \\
X_t - X_0 &= \int_0^t \theta f(s) X_s \, dW_s \\
\mathbb{E} \left( X_t \right) - \mathbb{E} \left( X_0 \right) &= \mathbb{E} \left( \int_0^t \theta f(s) X_s \, dW_s \right) = 0 \\
\mathbb{E} \left( X_t \right) &= \mathbb{E} \left( X_0 \right) = 1.
\end{align*}
\]

Therefore,

\[
\mathbb{E} \left( e^{\theta M_t} \right) = \exp \left( \frac{1}{2} \theta^2 \int_0^t f(s)^2 \, ds \right)
\]

which is the moment generating function for a normal distribution with mean zero and variance \( \int_0^t f(s)^2 \, ds \). Hence \( M_t \sim \mathcal{N} \left( 0, \int_0^t f(s)^2 \, ds \right) \). \( \square \)

5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Suppose \( X_t \) follows the generalised SDE

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t
\]

where \( \mu \) and \( \sigma \) are functions of \( X_t \) and \( t \). Show that \( X_t \) is a martingale if \( \mu(X_t, t) = 0 \).
3.2.2 One-Dimensional Diffusion Process

Solution: It suffices to show that under the filtration $\mathcal{F}_s$ where $s < t$, for $X_t$ to be a martingale

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s.$$ 

By taking integrals of the SDE,

$$\int_t^s dX_v = \int_s^t \mu(X_v, t) \, dv + \int_s^t \sigma(X_v, v) \, dW_v \quad \text{or} \quad X_t - X_s = \int_s^t \mu(X_v, t) \, dv + \int_s^t \sigma(X_v, v) \, dW_v.$$ 

Taking expectations,

$$\mathbb{E}(X_t - X_s) = \mathbb{E}\left[ \int_s^t \mu(X_v, t) \, dv \right]$$

or

$$\mathbb{E}(X_t) = \mathbb{E}(X_s) + \mathbb{E}\left[ \int_s^t \mu(X_v, t) \, dv \right].$$

Under the filtration $\mathcal{F}_s, s < t$

$$\mathbb{E}(X_t | \mathcal{F}_s) = \mathbb{E}(X_s | \mathcal{F}_s) + \mathbb{E}\left[ \int_s^t \mu(X_v, t) \, dv \big| \mathcal{F}_s \right]$$

$$= X_s + \mathbb{E}\left[ \int_s^t \mu(X_v, t) \, dv \big| \mathcal{F}_s \right].$$

Therefore, if $\mu(X_v, t) = 0$ then $X_t$ is a martingale. \qed

6. Bachelier Model (Arithmetic Brownian Motion). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{W_t : t \geq 0\}$ be a standard Wiener process. Suppose $X_t$ follows the arithmetic Brownian motion with SDE

$$dX_t = \mu dt + \sigma dW_t,$$

where $\mu$ and $\sigma$ are constants. Taking integrals show that for $t < T$,

$$X_T = X_t + \mu(T - t) + \sigma W_{T-t}$$

where $W_{T-t} = W_T - W_t \sim \mathcal{N}(0, T-t)$. Deduce that $X_T$, given $X_t = x$, follows a normal distribution with mean

$$\mathbb{E}(X_T | X_t = x) = x + \mu(T - t)$$

and variance

$$\text{Var}(X_T | X_t = x) = \sigma^2(T - t).$$
3.2.2 One-Dimensional Diffusion Process

**Solution:** Taking integrals of \( dX_t = \mu dt + \sigma dW_t \) we have

\[
\int_t^T dX_s = \int_t^T \mu \, ds + \int_t^T \sigma \, dW_s
\]

or

\[
X_T - X_t = \mu(T - t) + \sigma(W_T - W_t)
\]

where \( W_{T-t} \sim \mathcal{N}(0, T - t) \). Since \( X_t, \mu \) and \( \sigma \) are deterministic components therefore \( X_T \), given \( X_t = x \) follows a normal distribution with mean \( x + \mu(T - t) \) and variance \( \sigma^2(T - t) \).

\( \square \)

7. **Black–Scholes Model (Geometric Brownian Motion).** Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Suppose \( X_t \) follows the geometric Brownian motion with SDE

\[
dX_t = \mu X_t dt + \sigma X_t dW_t
\]

where \( \mu \) and \( \sigma \) are constants. By applying Itô’s formula to \( Y_t = \log X_t \) and taking integrals show that for \( t < T \),

\[
X_T = X_t e^{(\mu - \frac{1}{2} \sigma^2)(T-t) + \sigma W_{T-t}}
\]

where \( W_{T-t} \sim \mathcal{N}(0, T - t) \). Deduce that \( X_T \), given \( X_t = x \) follows a lognormal distribution with mean

\[
\mathbb{E}(X_T | X_t = x) = xe^{\mu(T-t)}
\]

and variance

\[
\text{Var}(X_T | X_t = x) = x^2 \left( e^{\sigma^2(T-t)} - 1 \right) e^{2 \mu(T-t)}.
\]

**Solution:** From Taylor’s expansion and subsequently using Itô’s formula,

\[
d(\log X_t) = \frac{1}{X_t} dX_t - \frac{1}{2X_t^2} (dX_t)^2 + \ldots
\]

\[
= \mu dt + \sigma dW_t - \frac{1}{2X_t^2} (\sigma^2 X_t^2 dt)
\]

\[
= \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t.
\]

Taking integrals,

\[
\int_t^T d(\log X_u) = \int_t^T \left( \mu - \frac{1}{2} \sigma^2 \right) du + \int_t^T \sigma \, dW_u
\]

\[
\log X_T - \log X_t = \left( \mu - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma(W_T - W_t)
\]

\[
X_T = X_t e^{\left( \mu - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma W_{T-t}}
\]
3.2.2 One-Dimensional Diffusion Process

where \( W_T - W_t = W_{T-t} \sim \mathcal{N}(0, T-t) \). Therefore,

\[
X_T \sim \log\mathcal{N} \left[ \log X_t + \left( \mu - \frac{1}{2} \sigma^2 \right) (T-t), \sigma^2 (T-t) \right]
\]

and from Problem 1.2.2.9 (page 20) the mean and variance of \( X_T \), given \( X_t = x \) are

\[
\mathbb{E} (X_T | X_t = x) = x e^{\int_t^T \mu_s ds} \\
\text{Var} (X_T | X_t = x) = x^2 \left( e^{\int_t^T \sigma^2_s ds} - 1 \right) e^{2 \int_t^T \mu_s ds}
\]

respectively.

8. **Generalised Geometric Brownian Motion.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Suppose \( X_t \) follows the generalised geometric Brownian motion with SDE

\[
dX_t = \mu_t X_t dt + \sigma_t X_t dW_t
\]

where \( \mu_t \) and \( \sigma_t \) are time dependent. By applying Itô’s formula to \( Y_t = \log X_t \) and taking integrals show that for \( t < T \),

\[
X_T = X_t \exp \left\{ \int_t^T \left( \mu_s - \frac{1}{2} \sigma^2_s \right) ds + \int_t^T \sigma_s dW_s \right\}
\]

Deduce that \( X_T \), given \( X_t = x \) follows a lognormal distribution with mean

\[
\mathbb{E} (X_T | X_t = x) = x e^{\int_t^T \mu_s ds}
\]

and variance

\[
\text{Var} (X_T | X_t = x) = x^2 \left( e^{\int_t^T \sigma^2_s ds} - 1 \right) e^{2 \int_t^T \mu_s ds}
\]

**Solution:** From Taylor’s expansion and using Itô’s formula,

\[
d(\log X_t) = \frac{1}{X_t} dX_t - \frac{1}{2X_t^2} (dX_t)^2 + \ldots
\]

\[
= \mu_t dt + \sigma_t dW_t - \frac{1}{2X_t^2} \left( \sigma_t^2 X_t^2 dt \right)
\]

\[
= \left( \mu_t - \frac{1}{2} \sigma_t^2 \right) dt + \sigma_t dW_t.
\]
3.2.2 One-Dimensional Diffusion Process

Taking integrals,
\[
\int_t^T d(\log X_s) = \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_t^T \sigma_s dW_s
\]
\[
\log X_T - \log X_t = \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_t^T \sigma_s dW_s,
\]
\[
X_T = X_t \exp \left\{ \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_t^T \sigma dW_s \right\}.
\]

Thus,
\[
X_T \sim \log \mathcal{N} \left[ \log X_t + \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds, \int_t^T \sigma_s^2 ds \right]
\]
and from Problem 1.2.2.9 (page 20) we have mean
\[
\mathbb{E} (X_T | X_t = x) = e^{\log x \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \frac{1}{2} \int_t^T \sigma_s^2 ds} = xe^{\int_t^T \mu_s ds}
\]
and variance
\[
\text{Var} (X_T | X_t = x) = \left( e^{\int_t^T \sigma_s^2 ds} - 1 \right) e^{2 \left( \log x \int_t^T \left( \mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \frac{1}{2} \int_t^T \sigma_s^2 ds \right)} = x^2 \left( e^{\int_t^T \sigma_s^2 ds} - 1 \right) e^{2 \int_t^T \mu_s ds}.
\]

9. Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Suppose \(X_t\) follows the geometric Brownian motion with SDE
\[
dX_t = \mu X_t dt + \sigma X_t dW_t
\]
where \(\mu\) and \(\sigma\) are constants. Show that if \(Y_t = X^n_t\), for some constant \(n\), then \(Y_t\) follows the geometric Brownian process of the form
\[
\frac{dY_t}{Y_t} = n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) dt + n \sigma dW_t,
\]
Deduce that given \(Y_t, t < T\), \(Y_T\) follows a lognormal distribution with the form
\[
Y_T = Y_t e^{n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t) + n \sigma W_{T-t}}
\]
where \(W_{T-t} \sim \mathcal{N}(0, T-t)\) with mean
\[
\mathbb{E} \left( Y_T | Y_t = y \right) = ye^{n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t)}
\]
3.2.2 One-Dimensional Diffusion Process

and variance

\[
\text{Var} \left( Y_T | Y_t = y \right) = y^2 \left( e^{\mu^2 \sigma^2 (T-t)} - 1 \right) e^{2n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t)}.
\]

**Solution:** From Itô’s formula,

\[
dY_t = nX_t^\sigma dX_t + \frac{1}{2} n(n-1)X_t^{\sigma^2} dt
\]

\[
= nX_t^\sigma (\mu dt + \sigma dW_t) + \frac{1}{2} n(n-1)\sigma^2 dt
\]

\[
= n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) X_t^\sigma dt + n\sigma X_t^\sigma dW_t.
\]

By substituting \( X_t^\sigma = Y_t \) we have

\[
dY_t / Y_t = n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) dt + n\sigma dW_t.
\]

Since \( Y_t \) follows a geometric Brownian motion, by analogy with Problem 3.2.2.7 (page 129) and by setting \( \mu \leftarrow n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) \) and \( \sigma \leftarrow n\sigma \), we can easily show that

\[
Y_T = y e^{n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t) + n\sigma W_{T-t}}
\]

follows a lognormal distribution where \( W_{T-t} \sim \mathcal{N}(0, T-t) \). In addition, we can also deduce

\[
\mathbb{E} \left( Y_T | Y_t = y \right) = ye^{n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t)}
\]

and

\[
\text{Var} \left( Y_T | Y_t = y \right) = y^2 \left( e^{\mu^2 \sigma^2 (T-t)} - 1 \right) e^{2n \left( \mu + \frac{1}{2} (n-1) \sigma^2 \right) (T-t)}.
\]

\[\square\]

10. **Ornstein–Uhlenbeck Process.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Suppose \(X_t\) follows the Ornstein–Uhlenbeck process with SDE

\[
dX_t = \kappa (\theta - X_t) \, dt + \sigma dW_t
\]

where \(\kappa, \theta\) and \(\sigma\) are constants. By applying Itô’s formula to \(Y_t = e^{\kappa t} X_t\) and taking integrals show that for \(t < T\),

\[
X_T = X_t e^{-\kappa (T-t)} + \theta \left[ 1 - e^{-\kappa (T-t)} \right] + \int_t^T \sigma e^{-\kappa (T-s)} dW_s.
\]

Using the properties of stochastic integrals on the above expression, find the mean and variance of \(X_T\), given \(X_t = x\). Deduce that \(X_T\) follows a normal distribution.
3.2.2 One-Dimensional Diffusion Process

Solution: Expanding \( Y_t = e^{\kappa t}X_t \) using Taylor’s formula and applying Itô’s formula, we have

\[
\begin{align*}
  d(e^{\kappa t}X_t) &= \kappa e^{\kappa t}X_t dt + e^{\kappa t}dX_t + \frac{1}{2} e^{\kappa t}(dX_t)^2 + \ldots \\
  &= \kappa e^{\kappa t}X_t dt + e^{\kappa t}(\kappa(\theta - X_t) dt + \sigma dW_t) \\
  &= \kappa \theta e^{\kappa t} dt + \sigma e^{\kappa t}dW_t.
\end{align*}
\]

Integrating the above expression,

\[
\int_t^T d(e^{\kappa s}X_s) = \int_t^T \kappa \theta e^{\kappa s} ds + \int_t^T \sigma e^{\kappa s}dW_s
\]

\[
X_T = X_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \int_t^T \sigma e^{-\kappa(T-s)}dW_s.
\]

Given the fact that

\[
\mathbb{E} \left( \int_t^T \sigma e^{-\kappa(T-s)}dW_s \right) = 0
\]

and

\[
\mathbb{E} \left[ \left( \int_t^T \sigma e^{-\kappa(T-s)}dW_s \right)^2 \right] = \mathbb{E} \left( \int_t^T \sigma^2 e^{-2\kappa(T-s)} ds \right) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right]
\]

the mean and variance of \( X_T \), given \( X_t = x \) are

\[
\mathbb{E} \left( X_T | X_t = x \right) = xe^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right]
\]

and

\[
\text{Var} \left( X_T | X_t = x \right) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right]
\]

respectively.

Since \( \int_t^T \sigma e^{-\kappa(T-s)}dW_s \) can be written in the form

\[
\int_t^T \sigma e^{-\kappa(T-s)}dW_s = \lim_{n \to \infty} \sum_{i=0}^{n-1} \sigma e^{-\kappa(T-t_i)} \left( W_{t_{i+1}} - W_{t_i} \right)
\]

where \( t_i = t + i(T-t)/n, t = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = T, n \in \mathbb{N} \) then due to the stationary increment of a standard Wiener process, we can see that each term of \( W_{t_{i+1}} - W_{t_i} \sim \mathcal{N} \left( 0, \frac{T-t}{n} \right) \) is normal multiplied by a deterministic exponential term. Thus, the product is normal and given that the sum of normal variables is normal we can deduce

\[
X_T \sim \mathcal{N} \left( xe^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right], \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right).
\]
11. **Geometric Mean-Reverting Process.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Suppose \(X_t\) follows the geometric mean-reverting process with SDE

\[
dX_t = \kappa (\theta - \log X_t)X_t dt + \sigma X_t dW_t, \quad X_0 > 0
\]

where \(\kappa, \theta\) and \(\sigma\) are constants. By applying Itô’s formula to \(Y_t = \log X_t\) show that the diffusion process can be reduced to an Ornstein–Uhlenbeck process of the form

\[
dY_t = \left[\kappa (\theta - Y_t) - \frac{1}{2} \sigma^2\right] dt + \sigma dW_t.
\]

Show also that for \(t < T\),

\[
\log X_T = (\log X_t) e^{-\kappa (T-t)} + \left( \theta - \frac{\sigma^2}{2\kappa} \right) (1 - e^{-\kappa (T-t)}) + \int_t^T \sigma e^{-\kappa (T-s)} dW_s.
\]

Using the properties of stochastic integrals on the above expression, find the mean and variance of \(X_T\), given \(X_t = x\) and deduce that \(X_T\) follows a lognormal distribution.

**Solution:** By expanding \(Y_t = \log X_t\) using Taylor’s formula and subsequently applying Itô’s formula, we have

\[
d(\log X_t) = \frac{1}{X_t} dX_t - \frac{1}{2X_t^2} (dX_t)^2 + \ldots
\]

\[
= \kappa (\theta - \log X_t) dt + \sigma dW_t - \frac{1}{2} \sigma^2 dt
\]

\[
= \left( \kappa (\theta - \log X_t) - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t
\]

and hence

\[
dY_t = \left( \kappa (\theta - Y_t) - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t.
\]

Using the same steps in solving the Ornstein–Uhlenbeck process, we apply Itô’s formula on \(Z_t = e^{\kappa Y_t}\) such that

\[
d(e^{\kappa Y_t}) = \kappa e^{\kappa Y_t} dY_t + e^{\kappa Y_t} dY_t + \frac{1}{2} \kappa^2 e^{\kappa Y_t} (dY_t)^2 + \ldots
\]

\[
= \kappa e^{\kappa Y_t} dt + e^{\kappa Y_t} \left[ \left( \kappa (\theta - Y_t) - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t \right]
\]

\[
= \left( \kappa \theta e^{\kappa Y_t} - \frac{1}{2} \sigma^2 e^{\kappa Y_t} \right) dt + \sigma e^{\kappa Y_t} dW_t.
\]

Taking integrals from \(t\) to \(T\), we have

\[
\int_t^T d(e^{\kappa Y_s}) = \int_t^T \left( \kappa \theta e^{\kappa Y_s} - \frac{1}{2} \sigma^2 e^{\kappa Y_s} \right) ds + \int_t^T \sigma e^{\kappa Y_s} dW_s
\]
3.2.2 One-Dimensional Diffusion Process

or
\[ Y_T = Y_t e^{-\kappa(T-t)} + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right) + \int_t^T \sigma e^{-\kappa(T-s)} dW_s. \]

By analogy with Problem 3.2.2.10 (page 132), we can deduce that \( Y_T = \log X_T \) follows a normal distribution and hence
\[ \log X_T \sim \mathcal{N} \left( \log X_t \left( e^{-\kappa(T-t)} \right) + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right), \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) \right) \]

or
\[ X_T \sim \mathcal{N} \left( \log X_t \left( e^{-\kappa(T-t)} \right) + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right), \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) \right) \]

with mean
\[ \mathbb{E} \left( X_T | X_t = x \right) = \exp \left\{ e^{-\kappa(T-t)} \log x + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right) + \frac{\sigma^2}{4\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) \right\} \]

and variance
\[ \text{Var} \left( X_T | X_t = x \right) = \exp \left\{ \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) - 1 \right\} \times \exp \left\{ 2e^{-\kappa(T-t)} \log x + 2 \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right) + \frac{\sigma^2}{2\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) \right\}. \]

\[ \square \]

12. Cox–Ingersoll–Ross (CIR) Model. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \{ \( W_t \) : \( t \geq 0 \) \} be a standard Wiener process. Suppose \( X_t \) follows the CIR model with SDE
\[ dX_t = \kappa \left( \theta - X_t \right) \, dt + \sigma \sqrt{X_t} \, dW_t, \quad X_0 > 0 \]

where \( \kappa, \theta \) and \( \sigma \) are constants. By applying Itô’s formula to \( Z_t = e^{\kappa t} X_t \) and \( Z_t^2 = e^{2\kappa t} X_t^2 \) and taking integrals show that for \( t < T \),
\[ X_T = X_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \int_t^T \sigma e^{-\kappa(T-s)} \sqrt{X_s} \, dW_s \]

and
\[ X_T^2 = X_t^2 e^{-2\kappa(T-t)} + \left( 2\kappa \theta + \sigma^2 \right) \int_t^T e^{-2\kappa(T-s)} X_s \, ds + 2\sigma \int_t^T e^{-2\kappa(T-s)} X_s^2 \, dW_s. \]

Using the properties of stochastic integrals on the above two expressions, find the mean and variance of \( X_T \), given \( X_t = x \).
Solution: Using Taylor’s formula and applying Itô’s formula on \( Z_t = e^\xi X_t \), we have

\[
d (e^{\xi t} X_t) = \kappa e^{\xi t} X_t dt + e^{\xi t} dX_t + \frac{1}{2} \kappa^2 e^{\xi t} (dX_t)^2 + \ldots
\]

\[
= \kappa e^{\xi t} X_t dt + e^{\xi t} (\kappa (\theta - X_t) dt + \sigma \sqrt{X_t} dW_t)
\]

\[
= \kappa \theta e^{\xi t} dt + \sigma e^{\xi t} \sqrt{X_t} dW_t.
\]

Integrating the above expression,

\[
\int_t^T d (e^{\xi s} X_s) = \int_t^T \kappa \theta e^{\xi s} ds + \int_t^T \sigma e^{\xi s} \sqrt{X_s} dW_s
\]

and therefore

\[
X_T = X_t e^{-\kappa (T - t)} + \theta \left[ 1 - e^{-\kappa (T - t)} \right] + \int_t^T \sigma e^{-\kappa (T - s)} \sqrt{X_s} dW_s.
\]

For the case of \( Z_t^2 = e^{2\xi t} X_t^2 \), by the application of Taylor’s expansion and Itô’s formula we have

\[
d (e^{2\xi t} X_t^2) = 2\kappa e^{2\xi t} X_t^2 dt + 2e^{2\xi t} X_t dX_t + \frac{1}{2} \left( 4\kappa^2 \right) e^{2\xi t} X_t^2 (dX_t)^2 + \ldots
\]

\[
= 2\kappa e^{2\xi t} X_t^2 dt + 2e^{2\xi t} X_t (\kappa (\theta - X_t) dt + \sigma \sqrt{X_t} dW_t) + e^{2\xi t} (\sigma^2 X_t dt)
\]

\[
= e^{2\xi t} (2\kappa \theta + \sigma^2) X_t dt + 2\sigma e^{2\xi t} X_t^3/2 dW_t.
\]

By taking integrals,

\[
\int_t^T d (e^{x s} X_s^2) = (2\kappa \theta + \sigma^2) \int_t^T e^{x s} X_s ds + 2\sigma \int_t^T e^{x s} X_s^3/2 dW_s
\]

and we eventually obtain the following expression:

\[
X_T^2 = e^{-2\kappa (T - t)} X_t^2 + (2\kappa \theta + \sigma^2) \int_t^T e^{-2\kappa (T - s)} X_s ds + 2\sigma \int_t^T e^{-2\kappa (T - s)} X_s^3/2 dW_s.
\]

Given \( X_t = x \), and by taking the expectation of the expression \( X_T \), we have

\[
\mathbb{E} (X_T | X_t = x) = xe^{-\kappa (T - t)} + \theta \left( 1 - e^{-\kappa (T - t)} \right).
\]

To find the variance, \( \text{Var} (X_T | X_t = x) \) we first take the expectation of \( X_T^2 \),

\[
\mathbb{E} (X_T^2 | X_t = x) = x^2 e^{-2\kappa (T - t)} + (2\kappa \theta + \sigma^2) \int_t^T e^{-2\kappa (T - s)} \mathbb{E} (X_s | X_t = x) ds
\]

\[
= x^2 e^{-2\kappa (T - t)}
\]
3.2.2 One-Dimensional Diffusion Process

\[ + (2x\theta + \sigma^2) \int_t^T e^{-2x(T-s)} \left[ xe^{-\kappa(s-t)} + \theta \left( 1 - e^{-\kappa(s-t)} \right) \right] ds \]

\[ = x^2 e^{-2x(T-t)} + \left( \frac{2x\theta + \sigma^2}{\kappa} \right) (x - \theta) \left( e^{-\kappa(T-t)} - e^{-2x(T-t)} \right) \]

\[ + \frac{\theta(2x\theta + \sigma^2)}{2\kappa} \left( 1 - e^{-2x(T-t)} \right). \]

Therefore,

\[ \text{Var} (X_T | X_t = x) = \mathbb{E} (X_T^2 | X_t = x) - \left[ \mathbb{E} (X_T | X_t = x) \right]^2 \]

\[ = \frac{x\sigma^2}{\kappa} \left( e^{-\kappa(T-t)} - e^{-2x(T-t)} \right) + \frac{\theta\sigma^2}{2\kappa} \left( 1 - 2e^{-\kappa(T-t)} + e^{-2x(T-t)} \right). \]

\[ \square \]

13. Brownian Bridge Process. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process. Suppose \(X_t\) follows the Brownian bridge process with SDE

\[ dX_t = \frac{y - X_t}{1 - t} dt + dW_t, \quad X_1 = y \]

where the diffusion is conditioned to be at \(y\) at time \(t = 1\). By applying Itô’s formula to \(Y_t = \frac{y - X_t}{1 - t}\) and taking integrals show that under an initial condition \(X_0 = x\) and for \(0 \leq t < 1\),

\[ X_t = yt + (1-t) \left( x + \int_0^t \frac{1}{1-s} dW_s \right). \]

Using the properties of stochastic integrals on the above expression, find the mean and variance of \(X_t\), given \(X_0 = x\) and show that \(X_t\) follows a normal distribution.

Solution: By expanding \(Y_t = \frac{y - X_t}{1 - t}\) using Taylor’s formula and subsequently applying Itô’s formula, we have

\[ dY_t = \frac{\partial Y_t}{\partial t} dt + \frac{\partial Y_t}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 Y_t}{\partial X_t^2} (dX_t)^2 + \ldots \]

\[ = \frac{y - X_t}{(1 - t)^2} dt - \left[ \left( \frac{1}{1 - t} \right) \left( \frac{y - X_t}{1 - t} dt + dW_t \right) \right] \]

\[ = - \left( \frac{1}{1 - t} \right) dW_t. \]

Taking integrals,

\[ \int_0^t dY_s = - \int_0^t \frac{1}{1 - s} dW_s \]

\[ Y_t - Y_0 = - \int_0^t \frac{1}{1 - s} dW_s \]
3.2.2 One-Dimensional Diffusion Process

\[ \frac{y - X_t}{1 - t} = y - x - \int_0^t \frac{1}{1 - s} \, dW_s. \]

Therefore,

\[ X_t = yt + (1 - t) \left( x + \int_0^t \frac{1}{1 - s} \, dW_s \right). \]

Using the properties of stochastic integrals,

\[ \mathbb{E} \left( \int_0^t \frac{1}{1 - s} \, dW_s \right) = 0 \]

\[ \mathbb{E} \left[ \left( \int_0^t \frac{1}{1 - s} \, dW_s \right)^2 \right] = \mathbb{E} \left( \int_0^t \frac{1}{(1 - s)^2} \, ds \right) = \frac{1}{1 - t} \]

therefore the mean and variance of \( X_t \) are

\[ \mathbb{E} (X_t | X_0 = x) = yt + x(1 - t) \]

and

\[ \text{Var} (X_t | X_0 = x) = \frac{1}{1 - t} \]

respectively.

Since \( \int_0^t \frac{1}{1 - s} \, dW_s \) is in the form \( \int_0^t f(s) \, dW_s \), from Problem 3.2.2.4 (page 126) we can easily prove that \( \int_0^t \frac{1}{1 - s} \, dW_s \) follows a normal distribution,

\[ \int_0^t \frac{1}{1 - s} \, dW_s \sim \mathcal{N} \left( yt + x(1 - t), \frac{1}{1 - t} \right). \]

\[ \square \]

14. **Forward Curve from an Asset Price Following a Geometric Brownian Motion.** Let \( (\Omega, \mathcal{F}, P) \) be a probability space and let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process. Suppose the asset price \( S_t \) at time \( t \) follows a geometric Brownian motion

\[ \frac{dS_t}{S_t} = \mu dt + \sigma dW_t \]

where \( \mu \) is the drift parameter and \( \sigma \) is the volatility. We define the forward price \( F(t, T) \) as an agreed-upon price set at time \( t \) to be paid or received at time \( T, t \leq T \) and is given by the relationship

\[ F(t, T) = \mathbb{E} \left( S_T | \mathcal{F}_t \right). \]

Show that the forward curve follows

\[ \frac{dF(t, T)}{F(t, T)} = \sigma dW_t. \]
Solution: Given that \( S_t \) follows a geometric Brownian motion then, following Problem 3.2.2.7 (page 129), we can easily show

\[
\mathbb{E} \left( S_T \mid \mathcal{F}_t \right) = S_t e^{\mu(T-t)}.
\]

Because \( F(t, T) = \mathbb{E} \left( S_T \mid \mathcal{F}_t \right) = S_t e^{\mu(T-t)} \), and by expanding \( F(t, T) \) using Taylor’s theorem and applying Itô’s lemma,

\[
dF(t, T) = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} (dS_t)^2 + \ldots
\]

\[
= -\mu S_t e^{\mu(T-t)} dt + e^{\mu(T-t)} dS_t.
\]

Since \( dS_t = \mu S_t dt + \sigma dW_t \) we have

\[
dF(t, T) = -\mu S_t e^{\mu(T-t)} dt + e^{\mu(T-t)} \left( \mu S_t dt + \sigma S_t dW_t \right)
\]

\[
= \sigma S_t e^{\mu(T-t)} dW_t
\]

\[
= \sigma F(t, T) dW_t.
\]

\[\square\]

15. **Forward Curve from an Asset Price Following a Geometric Mean-Reverting Process.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Suppose the asset price \( S_t \) at time \( t \) follows a geometric mean-reverting process

\[
\frac{dS_t}{S_t} = \kappa \left( \theta - \log S_t \right) dt + \sigma dW_t
\]

where \( \kappa \) is the mean-reversion rate, \( \theta \) is the long-term mean and \( \sigma \) is the volatility parameter. We define the forward price \( F(t, T) \) as an agreed-upon price set at time \( t \) to be paid or received at time \( T, t \leq T \) and is given by the relationship

\[
F(t, T) = \mathbb{E} \left( S_T \mid \mathcal{F}_t \right).
\]

Show that the forward curve follows

\[
\frac{dF(t, T)}{F(t, T)} = \sigma e^{-\kappa(T-t)} dW_t.
\]

Solution: Using the steps described in Problem 3.2.2.11 (page 134), the mean of \( S_T \) given \( S_t \) is

\[
\mathbb{E} \left( S_T \mid \mathcal{F}_t \right) = \exp \left\{ e^{-\kappa(T-t)} \log S_t + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left( 1 - e^{-\kappa(T-t)} \right) + \frac{\sigma^2}{4\kappa} \left( 1 - e^{-2\kappa(T-t)} \right) \right\}.
\]

Because \( F(t, T) = \mathbb{E} \left( S_T \mid \mathcal{F}_t \right) \) and expanding \( F(t, T) \) using Taylor’s theorem, we have

\[
dF(t, T) = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} (dS_t)^2 + \ldots
\]
3.2.2 One-Dimensional Diffusion Process

By substituting \( dS_t = \kappa \left( \theta - \log S_t \right) dt + \sigma S_t dW_t \) and applying Itô’s lemma,

\[
\frac{dF(t, T)}{F(t, T)} = \left[ \kappa e^{-\kappa (T-t)} \log S_t - \left( \theta - \frac{\sigma^2}{2\kappa} \right) \kappa e^{-\kappa (T-t)} - \frac{\sigma^2}{2} e^{-2\kappa (T-t)} \right] dt + \frac{e^{-\kappa (T-t)}}{S_t} F(t, T) dS_t + \frac{1}{2} \left[ - \frac{e^{-\kappa (T-t)}}{S_t^2} + \frac{e^{-2\kappa (T-t)}}{S_t^2} \right] F(t, T)(dS_t)^2 + \ldots
\]

By expanding \( \log F(t, T) \) using Taylor’s theorem and then applying Itô’s lemma,

\[
dF(t, T) = \left[ \kappa \left( \theta - \log S_t \right) \right] dS_t + \sigma S_t dW_t - \frac{\sigma^2}{2} \left[ e^{-\kappa (T-t)} - 2 e^{-2\kappa (T-t)} \right] dt + \frac{\sigma^2}{2} e^{-2\kappa (T-t)} dt
\]

\[
= \sigma e^{-\kappa (T-t)} dW_t.
\]

\[
\square
\]

16. Forward–Spot Price Relationship I. Let \( \{ W_t : t \geq 0 \} \) be the standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). We define the forward curve \( F(t, T) \) following the SDE

\[
\frac{dF(t, T)}{F(t, T)} = \sigma(t, T) dW_t
\]

as an agreed-upon price of an asset with current spot price \( S_t \) to be paid or received at time \( T, t \leq T \) where \( F(t, T) = \mathbb{E} \left( S_T \mid \mathcal{F}_t \right) \) such that \( S_T \) is the spot price at time \( T \) and \( \sigma(t, T) > 0 \) is a time-dependent volatility.

Show that the spot price has the following SDE

\[
\frac{dS_t}{S_t} = \left[ \frac{\partial \log F(0, t)}{\partial t} - \int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} \ dW_u + \int_0^t \frac{\partial \sigma(u, t)}{\partial t} \ dW_u \right] dt + \sigma(t, t) dW_t.
\]

**Solution:** By expanding \( \log F(t, T) \) using Taylor’s theorem and then applying Itô’s lemma,

\[
d \log F(t, T) = \frac{1}{F(t, T)} dF(t, T) - \frac{1}{2 F(t, T)^2} dF(t, T)^2 + \ldots
\]

\[
= \sigma(t, T) dW_t - \frac{1}{2} \sigma(t, T)^2 dt
\]

and taking integrals,

\[
\int_0^t d \log F(u, T) = \int_0^t \sigma(u, T) \ dW_u - \frac{1}{2} \int_0^t \sigma(u, T)^2 \ du
\]

so we have

\[
F(t, T) = F(0, T)e^{-\frac{1}{2} \int_0^t \sigma(u, T)^2 \ du + \int_0^t \sigma(u, T) dW_u}.
\]

By setting \( T = t \), the spot price \( S_t = F(t, t) \) can be expressed as

\[
S_t = F(0, t)e^{-\frac{1}{2} \int_0^t \sigma(u, t)^2 \ du + \int_0^t \sigma(u, t) dW_u}.
\]
3.2.2 One-Dimensional Diffusion Process

Expanding using Taylor’s theorem,

\[
dS_t = \frac{\partial S_t}{\partial t} dt + \frac{\partial S_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 S_t}{\partial t^2} dt^2 + \frac{1}{2} \frac{\partial^2 S_t}{\partial W_t^2} dW_t^2 + \frac{\partial^2 S_t}{\partial t \partial W_t} dt dW_t + \ldots
\]

and applying Itô’s lemma again,

\[
dS_t = \left( \frac{\partial S_t}{\partial t} + \frac{1}{2} \frac{\partial^2 S_t}{\partial W_t^2} \right) dt + \frac{\partial S_t}{\partial W_t} dW_t.
\]

From the spot price equation we have

\[
\frac{\partial S_t}{\partial t} = \frac{\partial F(0, t)}{\partial t} F(0, t)^{-1} S_t - \frac{1}{2} \left[ \sigma(t, t)^2 + \int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} du - 2 \int_0^t \frac{\partial \sigma(u, t)}{\partial t} dW_u \right] S_t,
\]

\[
\frac{\partial S_t}{\partial W_t} = \sigma(u, t) dW_u | S_t = \sigma(t, t) S_t,
\]

and

\[
\frac{\partial^2 S_t}{\partial W_t^2} = \sigma(t, t)^2 S_t.
\]

By substituting the values of \( \frac{\partial S_t}{\partial t} \), \( \frac{\partial S_t}{\partial W_t} \) and \( \frac{\partial^2 S_t}{\partial W_t^2} \) into \( dS_t = \left( \frac{\partial S_t}{\partial t} + \frac{1}{2} \frac{\partial^2 S_t}{\partial W_t^2} \right) dt + \frac{\partial S_t}{\partial W_t} dW_t \), we finally have

\[
\frac{dS_t}{S_t} = \left[ \frac{\partial \log F(0, t)}{\partial t} - \int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} du + \int_0^t \frac{\partial \sigma(u, t)}{\partial t} dW_u \right] dt + \sigma(t, t) dW_t.
\]

17. Clewlow–Strickland 1-Factor Model. Let \( \{W_t : t \geq 0\} \) be the standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Suppose the forward curve \( F(t, T) \) follows the process

\[
\frac{dF(t, T)}{F(t, T)} = \sigma e^{-\alpha(T-t)} dW_t,
\]

where \( t \leq T \), \( \alpha \) is the mean-reversion parameter and \( \sigma \) is the volatility. From the forward–spot price relationship \( F(t, T) = \mathbb{E} \left( S_T \mid \mathcal{F}_t \right) \) where \( S_T \) is the spot price at time \( T \), show that

\[
\frac{dS_t}{S_t} = \left[ \frac{\partial \log F(0, t)}{\partial t} + \alpha \left( \log F(0, t) - \log S_t \right) + \frac{\sigma^2}{4} \left( 1 - e^{-2\alpha t} \right) \right] dt + \sigma dW_t
\]
and the forward curve at time $t$ is given by

$$F(t, T) = F(0, T) \left[ \frac{S_t}{F(0, t)} \right]^{e^{-\alpha(T-t)}} e^{-\frac{\sigma^2}{2 \alpha}(e^{2\alpha t} - 1)(e^{-\alpha t} - e^{-\alpha T})}.$$ 

Finally, show that conditional on $F(0, T)$, $F(t, T)$ follows a lognormal distribution with mean

$$\mathbb{E}[F(t, T) | F(0, T)] = F(0, T)$$ 

and variance

$$\text{Var}[F(t, T) | F(0, T)] = F(0, T)^2 \exp \left\{ \frac{\sigma^2}{2 \alpha} \left[ e^{-2\alpha(T-t)} - e^{-2\alpha T} \right] - 1 \right\}.$$ 

**Solution:** For the case when

$$\frac{dF(t, T)}{F(t, T)} = \sigma(t, T) \, dW_t$$

with $\sigma(t, T)$ a time-dependent volatility, from Problem 3.2.2.16 (page 140) the corresponding spot price SDE is given by

$$\frac{dS_t}{S_t} = \left[ \frac{\partial \log F(0, t)}{\partial t} - \int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} \, du + \int_0^t \frac{\partial \sigma(u, t)}{\partial t} \, dW_u \right] \, dt + \sigma(t, t) \, dW_t.$$ 

Let $\sigma(t, T) = \sigma e^{-\alpha(T-t)}$ and taking partial differentiation with respect to $T$, we have

$$\frac{\partial \sigma(t, T)}{\partial T} = -\alpha \sigma e^{-\alpha(T-t)}.$$ 

Thus,

$$\int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} \, du = -\int_0^t \alpha \sigma^2 e^{-2\alpha(T-t)} \, du = -\frac{\sigma^2}{2} \left( 1 - e^{-2\alpha t} \right)$$

and

$$\int_0^t \frac{\partial \sigma(u, t)}{\partial t} \, dW_u = -\int_0^t \alpha \sigma e^{-\alpha(T-u)} \, dW_u.$$ 

Using Itô’s lemma,

$$d \log F(t, T) = \frac{1}{F(t, T)} dF(t, T) - \frac{1}{2 F(t, T)^2} dF(t, T)^2 + \ldots$$

$$= \sigma e^{-\alpha(T-t)} \, dW_t - \frac{1}{2} \sigma^2 e^{-2\alpha(T-t)} \, dt.$$
and taking integrals,

\[
\log F(t, T) = \log F(0, T) - \frac{1}{2} \int_0^t \sigma^2 e^{-2a(t-u)} \, du + \int_0^t \sigma e^{-a(T-u)} \, dW_u.
\]

By setting \( T = t \) such that \( F(t, t) = S_t \), we can write

\[
\log S_t = \log F(0, t) - \frac{1}{2} \int_0^t \sigma^2 e^{-2a(t-u)} \, du + \int_0^t \sigma e^{-a(t-u)} \, dW_u
\]

therefore

\[
\int_0^t \frac{\partial \sigma(u, t)}{\partial t} \, dW_u = \alpha \left( \log F(0, t) - \log S_t \right) - \frac{1}{2} \int_0^t \sigma^2 e^{-2a(t-u)} \, du
\]

\[
= \alpha \left( \log F(0, t) - \log S_t \right) - \frac{\sigma^2}{4} \left( 1 - e^{-2at} \right).
\]

By substituting the values of \( \int_0^t \sigma(u, t) \frac{\partial \sigma(u, t)}{\partial t} \, du \) and \( \int_0^t \frac{\partial \sigma(u, t)}{\partial t} \, dW_u \) into the spot price SDE and taking note that \( \sigma(t, t) = \sigma \), we eventually have

\[
\frac{dS_t}{S_t} = \left[ \frac{\partial \log F(0, t)}{\partial t} + \alpha \left( \log F(0, t) - \log S_t \right) + \frac{\sigma^2}{4} \left( 1 - e^{-2at} \right) \right] \, dt + \sigma dW_t.
\]

Since

\[
F(t, T) = F(0, T)e^{-\frac{1}{2} \int_0^t \sigma^2 e^{-2a(T-u)} \, du + \int_0^t \sigma e^{-a(T-u)} \, dW_u}
\]

with

\[
\int_0^t \sigma^2 e^{-2a(T-u)} \, du = \frac{\sigma^2}{2a} e^{-2aT} (e^{2at} - 1)
\]

and using the spot price equation

\[
\int_0^t \sigma e^{-a(T-u)} \, dW_u = e^{-a(T-t)} \int_0^t \sigma e^{-a(t-u)} \, dW_u
\]

\[
= e^{-a(T-t)} \left\{ \log \left( \frac{S_t}{F(0, t)} \right) + \frac{1}{2} \int_0^t \sigma^2 e^{-2a(T-u)} \, du \right\}
\]

\[
= e^{-a(T-t)} \left\{ \log \left( \frac{S_t}{F(0, t)} \right) + \frac{\sigma^2}{4a} \left( 1 - e^{-2at} \right) \right\}
\]

therefore

\[
F(t, T) = F(0, T) \left[ \frac{S_t}{F(0, t)} \right] e^{-a(T-t)} \left( e^{\frac{\sigma^2}{4a} e^{-2a(T-t)} (1 - e^{-2at})} \right).
\]

Finally, using

\[
d \log F(t, T) = \sigma e^{-a(T-t)} \, dW_t - \frac{1}{2} \sigma^2 e^{-2a(T-t)} \, dt
\]
and taking integrals we have

\[
\log F(t, T) = \log F(0, T) + \int_0^t \sigma e^{-a(T-u)} \, dW_u - \frac{\sigma^2}{4a} \left[ e^{-2a(T-t)} - e^{-2aT} \right].
\]

From the property of the Itô integral,

\[
\mathbb{E} \left[ \int_0^t \sigma e^{-a(T-u)} \, dW_u \right] = 0
\]

and

\[
\mathbb{E} \left[ \left( \int_0^t \sigma e^{-a(T-u)} \, dW_u \right)^2 \right] = \mathbb{E} \left[ \int_0^t \sigma^2 e^{-2a(T-u)} \, du \right] = \frac{\sigma^2}{2a} \left[ e^{-2a(T-t)} - e^{-2aT} \right].
\]

Since \( \int_0^t \sigma e^{-a(T-u)} \, dW_u \) is in the form \( \int_0^t f(u) \, dW_u \), from Problem 3.2.2.4 (page 126) we can easily show that \( \int_0^t \sigma e^{-a(T-u)} \, dW_u \) follows a normal distribution.

In addition, since we can also write the Itô integral as

\[
\int_0^t \sigma e^{-a(T-u)} \, dW_u = \lim_{n \to \infty} \sum_{i=0}^{n-1} \sigma e^{-a(T-t_i)}(W_{t_{i+1}} - W_{t_i})
\]

where \( t_i = \frac{it}{n}, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \) and due to the stationary increment of a standard Wiener process, each term of \( W_{t_{i+1}} - W_{t_i} \sim \mathcal{N} \left( 0, \frac{t}{n} \right) \) is normally distributed multiplied by a deterministic term and we can deduce that

\[
\int_0^t \sigma e^{-a(T-u)} \, dW_u \sim \mathcal{N} \left( 0, \frac{\sigma^2}{2a} \left[ e^{-2a(T-t)} - e^{-2aT} \right] \right).
\]

Hence,

\[
\log F(t, T) \sim \mathcal{N} \left( \log F(0, T) - \frac{\sigma^2}{4a} \left[ e^{-2a(T-t)} - e^{-2aT} \right], \frac{\sigma^2}{2a} \left[ e^{-2a(T-t)} - e^{-2aT} \right] \right)
\]

which implies

\[
\mathbb{E} \left[ F(t, T) \mid F(0, T) \right] = F(0, T)
\]

and

\[
\text{Var} \left[ F(t, T) \mid F(0, T) \right] = F(0, T)^2 \exp \left( \frac{\sigma^2}{2a} \left[ e^{-2a(T-t)} - e^{-2aT} \right] - 1 \right).
\]

\(\square\)
18. Constant Elasticity of Variance Model. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process with respect to the filtration \(\mathcal{F}_t, t \geq 0\). Suppose \(S_t > 0\) follows a constant elasticity of variance (CEV) model of the form

\[
dS_t = rS_t dt + \sigma(S_t, t)S_t dW_t
\]

where \(r\) is a constant and \(\sigma(S_t, t)\) is a local volatility function. By setting \(\sigma(S_t, t) = \alpha S_t^{\beta - 1}\) with \(\alpha > 0\) and \(0 < \beta < 1\), show using Itô’s formula that \(\sigma(S_t, t)\) satisfies

\[
\frac{d\sigma(S_t, t)}{\sigma(S_t, t)} = (\beta - 1) \left\{ \left[ r + \frac{1}{2}(\beta - 2)\sigma(S_t, t)^2 \right] dt + \sigma(S_t, t) dW_t \right\}.
\]

Finally, conditional on \(S_t\) show that for \(t < T\),

\[
S_T = e^{\int_t^T e^{-r(1-\beta)u} S_u^{1-\beta} + \alpha \int_t^T e^{-r(1-\beta)u} dW_u} [1]^{1-\beta}.
\]

**Solution:** From Itô’s formula,

\[
\frac{d\sigma(S_t, t)}{\sigma(S_t, t)} = (\beta - 1) \left\{ \left[ r + \frac{1}{2}(\beta - 2)\sigma(S_t, t)^2 \right] dt + \sigma(S_t, t) dW_t \right\}.
\]

Therefore,

\[
\frac{d\sigma(S_t, t)}{\sigma(S_t, t)} = (\beta - 1) \left\{ \left[ r + \frac{1}{2}(\beta - 2)\sigma(S_t, t)^2 \right] dt + \sigma(S_t, t) dW_t \right\}.
\]

To find the solution of the CEV model, let \(X_t = e^{-r_t S_t}\) and by Itô’s formula

\[
\frac{dX_t}{X_t} = \frac{dX_t}{dt} dt + \frac{dX_t}{dS_t} dS_t + \frac{1}{2} \frac{d^2X_t}{dS_t^2} (dS_t)^2 + \ldots
\]

\[
= -re^{-r_t}S_t dt + e^{-r_t} \left( rS_t dt + \alpha S_t^\beta dW_t \right)
\]

\[
= \alpha e^{-r_t} S_t^\beta dW_t
\]

\[
= \alpha e^{-r_t(1-\beta)} X_t^\beta dW_t.
\]

Taking integrals

\[
\int_t^T \frac{dX_u}{X_u^\beta} = \alpha \int_t^T e^{-r_t(1-\beta)u} dW_u.
\]
we have

\[ X_t^{1-\beta} = X_t + a \int_t^T e^{-r(1-\beta)u} \ dW_u \]

or

\[ X_t = \left[ X_t^{1-\beta} + a \int_t^T e^{-r(1-\beta)u} \ dW_u \right]^{1/(1-\beta)} \]

19. Geometric Average. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W_t : t \geq 0\}\) be a standard Wiener process with respect to the filtration \(\mathcal{F}_t, t \geq 0\). Suppose \(S_t\) satisfies the following geometric Brownian motion model:

\[ dS_t = \mu S_t du + \sigma S_t dW_t, \]

where \(\mu\) and \(\sigma\) are constant parameters. By considering a geometric average of \(S_t\)

\[ G_t = e^{\frac{1}{2} \int_0^t S_u du}, \quad G_0 = S_0 \]

show that \(G_t\) satisfies the following SDE:

\[ \frac{dG_t}{G_t} = \frac{1}{2} \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \frac{\sigma}{\sqrt{3}} d\bar{W}_t \]

where \(\bar{W}_t = \sqrt{\frac{3}{t}} \int_0^t W_u \ du\).

Under what condition is this SDE valid?

Solution: Using the steps described in Problem 3.2.2.7 (page 129) for \(u < t\),

\[ S_t = S_0 e^{\left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t} \]

or

\[ \log S_t = \log S_0 + \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t. \]

Taking the natural logarithm of \(G_t\),

\[ \log G_t = \frac{1}{t} \int_0^t \log S_u \ du \]

\[ = \frac{1}{t} \int_0^t \left[ \log S_0 + \left( \mu - \frac{1}{2} \sigma^2 \right) u + \sigma W_u \right] du \]
3.2.2 One-Dimensional Diffusion Process

\[
= \frac{1}{t} \int_0^t \log S_0 \, du + \frac{1}{t} \int_0^t \left( \mu - \frac{1}{2} \sigma^2 \right) u \, du + \frac{\sigma}{t} \int_0^t W_u \, du
\]

\[
= \log S_0 + \frac{1}{2} \left( \mu - \frac{1}{2} \sigma^2 \right) t + \frac{\sigma}{t} \int_0^t W_u \, du.
\]

From Problem 3.2.1.13 (page 115), using integration by parts we can write

\[
\int_0^t W_u \, du = \int_0^t (t - u) \, dW_u
\]

and we can deduce that \( \int_0^t W_u \, du \sim \mathcal{N} \left( 0, \frac{1}{3} t^3 \right) \) and hence

\[
\frac{\sigma}{t} \int_0^t W_u \, du \sim \mathcal{N} \left( 0, \frac{1}{3} \sigma^2 t \right) \quad \text{or} \quad \frac{\sigma}{\sqrt{3}} \tilde{W}_t \sim \mathcal{N} \left( 0, \frac{1}{3} \sigma^2 t \right)
\]

where \( \tilde{W}_t = \frac{\sqrt{3}}{t} \int_0^t W_u \, du \sim \mathcal{N}(0, t) \). Thus,

\[
\log G_t = \log S_0 + \frac{1}{2} \left( \mu - \frac{1}{2} \sigma^2 \right) t + \frac{\sigma}{\sqrt{3}} \tilde{W}_t
\]

or

\[
\frac{dG_t}{G_t} = \frac{1}{2} \left( \mu - \frac{1}{6} \sigma^2 \right) \, dt + \frac{\sigma}{\sqrt{3}} \, d\tilde{W}_t
\]

with \( G_0 = S_0 \).

However, given that \( \tilde{W}_t \) does not have the stationary increment property (see Problem 3.2.1.13, page 115), this SDE is only valid if the geometric average starts at time \( t = 0 \). 

\[ \Box \]

20. Feynman–Kac Formula for One-Dimensional Diffusion Process. We consider the following PDE problem:

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma(S_t, t)^2 \frac{\partial^2 V}{\partial S_t^2} + \mu(S_t, t) \frac{\partial V}{\partial S_t} - r(t) V(S_t, t) = 0
\]

with boundary condition \( V(S_T, T) = \Psi(S_T) \) where \( \mu, \sigma \) are known functions of \( S_t \) and \( t \), \( r \) and \( \Psi \) are functions of \( t \) and \( S_T \), respectively with \( t < T \). Using Itô’s formula on the process

\[
Z_u = e^{-\int_0^u r(v) \, dv} V(S_u, u)
\]

where \( S_t \) satisfies the generalised SDE

\[
dS_t = \mu(S_t, t) \, dt + \sigma(S_t, t) \, dW_t
\]
such that \( \{ W_t : t \geq 0 \} \) is a standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), show that under the filtration \( \mathcal{F}_t \) the solution of the PDE is given by

\[
V(S_t, t) = \mathbb{E} \left[ e^{-\int_t^T r(s) \, ds} \Psi(S_T) \, \middle| \, \mathcal{F}_t \right].
\]

**Solution:** Let \( g(u) = e^{-\int_u^T r(s) \, ds} \) and hence we can write

\[
Z_u = g(u)V(S_u, u).
\]

By applying Taylor’s expansion and Itô’s formula on \( dZ_u \) we have

\[
dZ_u = \frac{\partial Z_u}{\partial u} du + \frac{\partial Z_u}{\partial S_u} dS_u + \frac{1}{2} \frac{\partial^2 Z_u}{\partial S_u^2} (dS_u)^2 + \ldots
\]

\[
= \left( g(u) \frac{\partial V}{\partial u} + V(S_u, u) \frac{\partial g}{\partial u} \right) du + \left( g(u) \frac{\partial V}{\partial S_u} \right) dS_u + \frac{1}{2} \left( g(u) \frac{\partial^2 V}{\partial S_u^2} \right) (dS_u)^2
\]

\[
= \left( g(u) \frac{\partial V}{\partial u} - r(u)g(u)V(S_u, u) \right) du + \left( g(u) \frac{\partial V}{\partial S_u} \right) \left( \mu(S_u, u) \, du + \sigma(S_u, u) dW_u \right)
\]

\[
\quad + \frac{1}{2} \left( g(u) \frac{\partial^2 V}{\partial S_u^2} \right) (\sigma(S_u, u)^2 du)
\]

\[
= g(u) \left( \frac{\partial V}{\partial u} + \frac{1}{2} \sigma(S_u, u)^2 \frac{\partial^2 V}{\partial S_u^2} + \mu(S_u, u) \frac{\partial V}{\partial S_u} - r(u)V(S_u, u) \right) du
\]

\[
+ g(u) \sigma(S_u, u) \frac{\partial V}{\partial S_u} dW_u
\]

\[
= g(u) \sigma(S_u, u) \frac{\partial V}{\partial S_u} dW_u
\]

since

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma(S_u, u)^2 \frac{\partial^2 V}{\partial S_u^2} + \mu(S_u, u) \frac{\partial V}{\partial S_u} - r(u)V(S_u, u) = 0.
\]

Taking integrals we have

\[
\int_t^T dZ_u = \int_t^T g(u) \sigma(S_u, u) \frac{\partial V}{\partial S_u} dW_u
\]

\[
Z_T - Z_t = \int_t^T e^{-\int_t^t r(s) \, ds} \sigma(S_u, u) \frac{\partial V}{\partial S_u} dW_u.
\]

By taking expectations and using the properties of the Itô integral,

\[
\mathbb{E} \left( Z_T - Z_t \right) = 0 \quad \text{or} \quad \mathbb{E} \left( Z_t \right) = \mathbb{E} \left( Z_T \right)
\]
3.2.2 One-Dimensional Diffusion Process

and hence under the filtration \( \mathcal{F}_t \),

\[
\mathbb{E} \left[ Z_t | \mathcal{F}_t \right] = \mathbb{E} \left[ Z_T | \mathcal{F}_t \right] \\
\mathbb{E} \left[ e^{-\int_t^T r(v) dv} V(S_t, t) | \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^T r(v) dv} V(S_T, T) | \mathcal{F}_t \right] \\
V(S_t, t) = \mathbb{E} \left[ e^{-\int_t^T r(v) dv} \Psi(S_T) | \mathcal{F}_t \right].
\]

\[
\square
\]

21. Backward Kolmogorov Equation for One-Dimensional Diffusion Process. Let \( \{ W_t : t \geq 0 \} \) be a standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). For \( t \in [0, T] \), \( T > 0 \) consider the generalised stochastic differential equation

\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t
\]

where \( \mu(X_t, t) \) and \( \sigma(X_t, t) \) are functions dependent on \( X_t \) and \( t \). By conditioning \( X_t = x \) and \( X_T = y \), let \( p(x, t; y, T) \) be the transition probability density of \( X_T \) at time \( T \) starting at time \( t \) at point \( X_t \). For any function \( \Psi(X_T) \), from the Feynman–Kac formula the function

\[
f(x, t) = \mathbb{E} \left[ e^{-\int_t^T r(u) du} \Psi(X_T) | \mathcal{F}_t \right] \\
= e^{-\int_t^T r(u) du} \int \Psi(y) p(x, t; y, T) \, dy
\]

satisfies the partial differential equation

\[
\frac{\partial f}{\partial t} + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2 f}{\partial x^2} + \mu(x, t) \frac{\partial f}{\partial x} - r(t) f(x, t) = 0
\]

where \( r \) is a time-dependent function.

Show that the transition probability density \( p(t, x; y, T) \) in the \( y \) variable satisfies

\[
\frac{\partial}{\partial t} p(t, x; y, T) + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2}{\partial x^2} p(t, x; y, T) + \mu(x, t) \frac{\partial}{\partial x} p(t, x; y, T) = 0.
\]

Solution: From the Feynman–Kac formula, for any function \( \Psi(X_T) \), the function

\[
f(x, t) = \mathbb{E} \left[ e^{-\int_t^T r(u) du} \Psi(X_T) | \mathcal{F}_t \right] \\
= e^{-\int_t^T r(u) du} \int \Psi(y) p(x, t; y, T) \, dy
\]

satisfies the following PDE:

\[
\frac{\partial f}{\partial t} + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2 f}{\partial x^2} + \mu(x, t) \frac{\partial f}{\partial x} - r(t) f(x, t) = 0.
\]
By differentiation, we have
\[
\frac{\partial f}{\partial t} = r(t)e^{-\int_0^T r(u)du} \int \Psi(y)p(x, t; y, T) \, dy + e^{-\int_0^T r(u)du} \int \Psi(y) \frac{\partial}{\partial t} p(x, t; y, T) \, dy
\]
\[
= r(t)f(x) + e^{-\int_0^T r(u)du} \int \Psi(y) \frac{\partial}{\partial x} p(x, t; y, T) \, dy
\]
\[
\frac{\partial f}{\partial x} = e^{-\int_0^T r(u)du} \int \Psi(y) \frac{\partial}{\partial x} p(x, t; y, T) \, dy \quad \text{and}
\]
\[
\frac{\partial^2 f}{\partial x^2} = e^{-\int_0^T r(u)du} \int \Psi(y) \frac{\partial^2}{\partial x^2} p(x, t; y, T) \, dy.
\]

Substituting the above equations into the PDE, we obtain
\[
e^{-\int_0^T r(u)du} \times \int \Psi(y) \left[ \frac{\partial}{\partial t} p(x, t; y, T) + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2}{\partial x^2} p(x, t; y, T) + \mu(x, t) \frac{\partial}{\partial x} p(x, t; y, T) \right] \, dy = 0.
\]

Finally, irrespective of the choice of \(\Psi(y)\) and \(r(t)\), the transition probability density function \(p(x, t; y, T)\) satisfies
\[
\frac{\partial}{\partial t} p(x, t; y, T) + \frac{1}{2} \sigma(x, t)^2 \frac{\partial^2}{\partial x^2} p(x, t; y, T) + \mu(x, t) \frac{\partial}{\partial x} p(x, t; y, T) = 0.
\]

\[\square\]

22. **Forward Kolmogorov Equation for One-Dimensional Diffusion Process.** Let \(\{W_t : t \geq 0\}\) be a standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). For \(t \in [0, T]\), \(T > 0\) consider the generalised stochastic differential equation
\[
dX_t = \mu(X_t, t) \, dt + \sigma(X_t, t) \, dW_t
\]

where \(\mu(X_t, t)\) and \(\sigma(X_t, t)\) are functions dependent on \(X_t\) and \(t\). Using Itô’s formula on the function \(f(X_t, t)\), show that
\[
f(X_T, T) = f(X_t, t) + \int_t^T \left( \frac{\partial f}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial f}{\partial X_t}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 f}{\partial X_t^2}(X_s, s) \right) ds
\]
\[
+ \int_t^T \sigma(X_s, s) \frac{\partial f}{\partial X_t}(X_s, s) \, dW_s
\]

and taking the expectation conditional on \(X_t = x\), show that in the limit \(T \to t\)
\[
\int_t^T \mathbb{E} \left[ \frac{\partial f}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial f}{\partial X_t}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 f}{\partial X_t^2}(X_s, s) \bigg| X_t = x \right] ds = 0.
3.2.2 One-Dimensional Diffusion Process

Let \( X_T = y \) and define the transition probability density \( p(x, t; y, T) \) of \( X_T \) at time \( T \) starting at time \( t \) at point \( X_t \). By writing the conditional expectation in terms of \( p(x, t; y, T) \) and integrating by parts twice, show that

\[
\frac{\partial}{\partial T} p(x, t; y, T) = \frac{1}{2} \sigma^2 (\sigma (y, T)^2 p(x, t; y, T)) - \frac{\partial}{\partial y} (\mu (y, T) p(x, t; y, T)).
\]

**Solution:** For a suitable function \( f(X_t, t) \) and using Itô’s formula,

\[
\begin{align*}
df(X_t, t) &= \frac{\partial f}{\partial t}(X_t, t)\, dt + \frac{\partial f}{\partial X_t}(X_t, t)\, dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2}(X_t, t)(dX_t)^2 + \ldots \\
&= \frac{\partial f}{\partial t}(X_t, t)\, dt + \frac{\partial f}{\partial X_t}(X_t, t) \left( \mu(X_t, t)\, dt + \sigma(X_t, t)\, dW_t \right) + \frac{1}{2} \sigma(X_t, t)^2 \frac{\partial^2 f}{\partial X_t^2}(X_t, t)\, dt \\
&= \left( \frac{\partial f}{\partial t}(X_t, t) + \mu(X_t, t) \frac{\partial f}{\partial X_t}(X_t, t) + \frac{1}{2} \sigma(X_t, t)^2 \frac{\partial^2 f}{\partial X_t^2}(X_t, t) \right) dt \\
&\quad + \sigma(X_t, t) \frac{\partial f}{\partial X_t}(X_t, t)\, dW_t.
\end{align*}
\]

Taking integrals,

\[
\int_t^T df(X_s, s) = \int_t^T \left( \frac{\partial f}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial f}{\partial X_s}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 f}{\partial X_s^2}(X_s, s) \right) ds \\
+ \int_t^T \sigma(X_s, s) \frac{\partial f}{\partial X_s}(X_s, s)\, dW_s
\]

\( f(X_T, T) = f(X_t, t) \)

\[
\begin{align*}
&+ \int_t^T \left( \frac{\partial f}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial f}{\partial X_s}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 f}{\partial X_s^2}(X_s, s) \right) ds \\
&+ \int_t^T \sigma(X_s, s) \frac{\partial f}{\partial X_s}(X_s, s)\, dW_s.
\end{align*}
\]

By taking conditional expectations given \( X_t = x \) we have

\[
\begin{align*}
\mathbb{E} \left( f(X_T, T) | X_t = x \right) &= \mathbb{E} \left( f(X_t, t) | X_t = x \right) \\
&+ \int_t^T \mathbb{E} \left( \frac{\partial f}{\partial t}(X_s, s) + \mu(X_s, s) \frac{\partial f}{\partial X_s}(X_s, s) + \frac{1}{2} \sigma(X_s, s)^2 \frac{\partial^2 f}{\partial X_s^2}(X_s, s) \bigg| X_t = x \right) ds
\end{align*}
\]

and since in the limit

\[
\lim_{T \to t} \mathbb{E} \left( f(X_T, T) | X_t = x \right) = \mathbb{E} \left( f(X_t, t) | X_t = x \right)
\]
so

\[
\int_0^T \left[ \frac{\partial f}{\partial t}(X_t, s) + \mu(X_t, s) \frac{\partial f}{\partial X_t}(X_t, s) + \frac{1}{2} \sigma(X_t, s)^2 \frac{\partial^2 f}{\partial X_t^2}(X_t, s) \right] ds = 0
\]

or

\[
\int_0^T \int_{-\infty}^{\infty} \left[ \frac{\partial f}{\partial s}(y, s) + \mu(y, s) \frac{\partial f}{\partial y}(y, s) + \frac{1}{2} \sigma(y, s)^2 \frac{\partial^2 f}{\partial y^2}(y, s) \right] p(x, t; y, s) dy ds = 0
\]

where \( p(x, t; y, s) \) is the transition probability density function of \( X_t \) in the \( y \)-variable. Integrating by parts,

\[
\int_0^T \int_{-\infty}^{\infty} \frac{\partial f}{\partial s}(y, s) p(x, t; y, s) dy ds = \int_{-\infty}^{\infty} \int_0^T \frac{\partial f}{\partial s}(y, s) p(x, t; y, s) dy ds dy
\]

\[
= \int_{-\infty}^{\infty} f(y, s) p(x, t; y, s) \bigg|_0^T dy
\]

\[
- \int_{-\infty}^{\infty} \int_t^T f(y, s) \frac{\partial}{\partial s} p(x, t; y, s) ds dy
\]

\[
= - \int_{-\infty}^{\infty} \int_t^T f(y, s) \frac{\partial}{\partial s} p(x, t; y, s) ds dy
\]

\[
= - \int_t^T \int_{-\infty}^{\infty} f(y, s) \frac{\partial}{\partial s} p(x, t; y, s) dy ds
\]

\[
\int_0^T \int_{-\infty}^{\infty} \mu(y, s) \frac{\partial f}{\partial y}(y, s) p(x, t; y, s) dy ds = \int_0^T f(y, s) \mu(y, s) p(x, t; y, s) \bigg|_{-\infty}^{\infty} ds
\]

\[
- \int_0^T \int_{-\infty}^{\infty} f(y, s) \frac{\partial}{\partial y} (\mu(y, s) p(x, t; y, s)) dy ds
\]

\[
= - \int_0^T \int_{-\infty}^{\infty} f(y, s) \frac{\partial}{\partial y} (\mu(y, s) p(x, t; y, s)) dy ds
\]

\[
\int_0^T \int_{-\infty}^{\infty} \frac{1}{2} \sigma(y, s)^2 \frac{\partial^2 f}{\partial y^2}(y, s) p(x, t; y, s) dy ds
\]

\[
= \int_0^T \int_{-\infty}^{\infty} \frac{1}{2} \frac{\partial}{\partial y} f(y, s) \sigma(y, s)^2 p(x, t; y, s) \bigg|_{-\infty}^{\infty} ds
\]

\[
- \int_0^T \int_{-\infty}^{\infty} \frac{1}{2} \frac{\partial}{\partial y} f(y, s) \sigma(y, s)^2 p(x, t; y, s) dy ds
\]

\[
= - \int_0^T \int_{-\infty}^{\infty} \frac{1}{2} \frac{\partial}{\partial y} f(y, s) \sigma(y, s)^2 p(x, t; y, s) dy ds
\]

\[
= - \int_0^T \int_{-\infty}^{\infty} \frac{1}{2} \frac{\partial}{\partial y} (\sigma(y, s)^2 p(x, t; y, s)) \bigg|_{-\infty}^{\infty} ds
\]
3.2.2 One-Dimensional Diffusion Process

\[
+ \int_1^T \int_{-\infty}^\infty \frac{1}{2} f(y, s) \frac{\partial^2}{\partial y^2} (\sigma(y, s)^2 p(x, t; y, s)) \, dy\, ds
\]

\[
= \int_1^T \int_{-\infty}^\infty \frac{1}{2} f(y, s) \frac{\partial^2}{\partial y^2} (\sigma(y, s)^2 p(x, t; y, s)) \, dy\, ds
\]

and by substituting the above relationships into the integro partial differential equation, we have

\[
\int_1^T \int_{-\infty}^\infty f(y, s) \times \left[ \frac{\partial}{\partial s} p(x, t; y, s) + \frac{\partial}{\partial y} (\mu(y, s) p(x, t; y, s)) - \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma(y, s)^2 p(x, t; y, s)) \right] \, dy\, ds = 0.
\]

Finally, by differentiating the above equation with respect to \( T \), we obtain

\[
\int_{-\infty}^\infty f(y, T) \times \left[ \frac{\partial}{\partial T} p(x, t; y, T) + \frac{\partial}{\partial y} (\mu(y, T) p(x, t; y, T)) - \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma(y, T)^2 p(x, t; y, T)) \right] \, dy = 0
\]

and irrespective of the choice of \( f \), the transition probability density function \( p(x, t; y, T) \) satisfies

\[
\frac{\partial}{\partial T} p(x, t; y, T) = \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma(y, T)^2 p(x, t; y, T)) - \frac{\partial}{\partial y} (\mu(y, T) p(x, t; y, T)).
\]

\( \square \)

23. **Backward Kolmogorov Equation for a One-Dimensional Random Walk.** We consider a one-dimensional symmetric random walk where at initial time \( t_0 \), a particle starts at \( x_0 \) and is at position \( x \) at time \( t \). At time \( t + \delta t \), the particle can either move to \( x + \delta x \) or \( x - \delta x \) each with probability \( \frac{1}{2} \). Let \( p(x, t; x_0, t_0) \) denote the probability density of the particle position \( x \) at time \( t \) starting at \( x_0 \) at time \( t_0 \).

By writing the backward equation in a discrete fashion and expanding it using Taylor’s series, show that for \( \delta x = \sqrt{\delta t} \) and in the limit \( \delta t \to 0 \)

\[
\frac{\partial p(x, t; x_0, t_0)}{\partial t} = -\frac{1}{2} \frac{\partial^2 p(x, t; x_0, t_0)}{\partial x^2}.
\]

**Solution:** By denoting \( p(x, t; x_0, t_0) \) as the probability density function of the particle position \( x \) at time \( t \), hence the discrete model of the backward equation is

\[
p(x, t; x_0, t_0) = \frac{1}{2} p(x + \delta x, t + \delta t; x_0, t_0) + \frac{1}{2} p(x - \delta x, t + \delta t; x_0, t_0).
\]

Using Taylor’s series, we have

\[
p(x - \delta x, t + \delta t; x_0, t_0) = p(x, t + \delta t; x_0, t_0) - \frac{\partial p(x, t + \delta t; x_0, t_0)}{\partial x} \delta x
\]

\[
+ \frac{1}{2} \frac{\partial^2 p(x, t + \delta t; x_0, t_0)}{\partial x^2} (-\delta x)^2 + O((\delta x)^3)
\]
24. **Forward Kolmogorov Equation for a One-Dimensional Random Walk.** We consider a one-dimensional symmetric random walk where at initial time $t_0$, a particle starts at $y_0$ and is at position $y$ at terminal time $T > 0$. At time $T - \delta T$, the particle can either move to $y + \delta y$ or $y - \delta y$ each with probability $\frac{1}{2}$. Let $p(y; T; y_0, t_0)$ denote the probability density of the position $y$ at time $T$ starting at $y_0$ at time $t_0$. By writing the forward equation in a discrete fashion and expanding it using Taylor’s series, show that for $\delta y = \sqrt{\delta T}$ and in the limit $\delta T \to 0$

$$\frac{\partial p(y; T; y_0, t_0)}{\partial T} = \frac{1}{2} \frac{\partial^2 p(y; T; y_0, t_0)}{\partial y^2}.$$ 

**Solution:** By denoting $p(y; T; y_0, t_0)$ as the probability density function of the particle position $y$ at time $T$, hence the discrete model of the forward equation is

$$p(y; T; y_0, t_0) = \frac{1}{2} p(y - \delta y, T - \delta T; y_0, t_0) + \frac{1}{2} p(y + \delta y, T - \delta T; y_0, t_0).$$

By expanding $p(y - \delta y, T - \delta T; y_0, t_0)$ and $p(y + \delta y, T - \delta T; y_0, t_0)$ using Taylor’s series, we have

$$p(y - \delta y, T - \delta T; y_0, t_0) = p(y, T - \delta T; y_0, t_0) - \frac{\partial p(y, T - \delta T; y_0, t_0)}{\partial y} \delta y$$

$$+ \frac{1}{2} \frac{\partial^2 p(y, T - \delta T; y_0, t_0)}{\partial y^2} (-\delta y)^2 + O((\delta y)^3)$$

$$p(y + \delta y, T - \delta T; y_0, t_0) = p(y, T - \delta T; y_0, t_0) + \frac{\partial p(y, T - \delta T; y_0, t_0)}{\partial y} \delta y$$

$$+ \frac{1}{2} \frac{\partial^2 p(y, T - \delta T; y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta y)^3).$$
3.2.3 Multi-Dimensional Diffusion Process

Let

\[ p(y + \delta y, T - \delta T; y_0, t_0) = p(y, T - \delta T; y_0, t_0) + \frac{\partial p(y, T - \delta T; y_0, t_0)}{\partial y} \delta y \]

\[ + \frac{1}{2} \frac{\partial^2 p(y, T - \delta T; y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta y)^3). \]

Substituting the above two equations into the discrete forward equation,

\[ p(y, T; y_0, t_0) = p(y, T - \delta T; y_0, t_0) + \frac{1}{2} \frac{\partial^2 p(y, T - \delta T; y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta y)^3). \]

By setting \( \delta y = \sqrt{\delta T} \), dividing the equation by \( \delta T \) and taking limits \( \delta T \to 0 \)

\[ \lim_{\delta T \to 0} \frac{p(y, T; y_0, t_0) - p(y, T - \delta T; y_0, t_0)}{\delta T} = \lim_{\delta T \to 0} \left[ \frac{1}{2} \frac{\partial^2 p(y, T - \delta T; y_0, t_0)}{\partial y^2} + O(\sqrt{\delta T}) \right] \]

we finally have

\[ \frac{\partial p(y, T; y_0, t_0)}{\partial T} = \frac{1}{2} \frac{\partial^2 p(y, T; y_0, t_0)}{\partial y^2}. \]

3.2.3 Multi-Dimensional Diffusion Process

1. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and consider \( n \) assets with prices \( S^{(i)}_t, i = 1, 2, \ldots, n \) satisfying the SDEs

\[ dS^{(i)}_t = \mu^{(i)} S^{(i)}_t dt + \sigma^{(i)} S^{(i)}_t dW^{(i)}_t \]

\[ \left( dW^{(i)}_t \right) \left( dW^{(j)}_t \right) = \rho^{(ij)} dt \]

where \( \{ W^{(i)}_t : t \geq 0 \}, i = 1, 2, \ldots, n \) are standard Wiener processes, \( \rho^{(ij)} \in (-1, 1), i \neq j \)

and \( \rho^{(ii)} = 1. \)

By considering the function \( f(S^{(1)}_t, S^{(2)}_t, \ldots, S^{(n)}_t) \), show using Itô’s formula that

\[ df(S^{(1)}_t, S^{(2)}_t, \ldots, S^{(n)}_t) = \sum_{i=1}^{n} \mu^{(i)} S^{(i)}_t \frac{df}{\partial S^{(i)}_t} dt + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho^{(ij)} \sigma^{(i)} S^{(i)}_t \sigma^{(j)} S^{(j)}_t \frac{\partial^2 f}{\partial S^{(i)}_t \partial S^{(j)}_t} dt \]

\[ + \sum_{i=1}^{n} \sigma^{(i)} S^{(i)}_t \frac{df}{\partial S^{(i)}_t} dW^{(i)}_t. \]

**Solution:** By expanding \( df(S^{(1)}_t, S^{(2)}_t, \ldots, S^{(n)}_t) \) using Taylor’s formula,

\[ df(S^{(1)}_t, S^{(2)}_t, \ldots, S^{(n)}_t) = \sum_{i=1}^{n} \frac{df}{\partial S^{(i)}_t} dS^{(i)}_t + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial S^{(i)}_t \partial S^{(j)}_t} dS^{(i)}_t dS^{(j)}_t + \ldots \]
3.2.3 Multi-Dimensional Diffusion Process

Let

\[ \text{By letting } dU_t = \mu U_t dt + \sigma U_t dV_t \]

where \( \mu = \mu^1 - \mu^2 + (\sigma^1)^2 - \rho \sigma^1 \sigma^2, \sigma = \sqrt{(\sigma^1)^2 + (\sigma^2)^2 - 2 \rho \sigma^1 \sigma^2} \) and

\[ V_t = \frac{\sigma^1 W_t^1 - \sigma^2 W_t^2}{\sqrt{(\sigma^1)^2 + (\sigma^2)^2 - 2 \rho \sigma^1 \sigma^2}}. \]

Finally, show that \( \{ V_t : t \geq 0 \} \) is a standard Wiener process.

**Solution:** By using Taylor’s expansion and applying Itô’s formula,

\[ dU_t = d \left( \frac{S_t^{(1)}}{S_t^{(2)}} \right) \]

\[ = \frac{1}{S_t^{(2)}} dS_t^{(1)} - \frac{S_t^{(1)}}{(S_t^{(2)})^2} dS_t^{(2)} + \frac{S_t^{(1)}}{(S_t^{(2)})^3} (dS_t^{(2)})^2 - \frac{1}{(S_t^{(2)})^2} dS_t^{(1)} dS_t^{(2)} + \ldots \]
3.2.3 Multi-Dimensional Diffusion Process

\[ = \mu(1)U_t dt + \sigma(1)U_t dW_t^{(1)} - \left( \mu(2)U_t dt + \sigma(2)U_t dW_t^{(2)} \right) \]
\[ + (\sigma(2)^2)U_t dt - \rho \sigma(1) \sigma(2) U_t dt \]
\[ = \left( \mu(1) - \mu(2) + (\sigma(2)^2) \right) U_t dt + U_t \left( \sigma(1) dW_t^{(1)} - \sigma(2) dW_t^{(2)} \right) \]
\[ = \left( \mu(1) - \mu(2) + (\sigma(2)^2) \right) U_t dt + \sqrt{(\sigma(1)^2 + (\sigma(2)^2))} U_t dV_t. \]

Therefore,

\[ dU_t = \mu U_t dt + \sigma U_t dV_t \]

where \( \mu = \mu(1) - \mu(2) + (\sigma(2)^2) - \rho \sigma(1) \sigma(2), \sigma = \sqrt{(\sigma(1)^2 + (\sigma(2)^2))} \) and

\[ V_t = \frac{\sigma(1) W_t^{(1)} - \sigma(2) W_t^{(2)}}{\sqrt{(\sigma(1)^2 + (\sigma(2)^2))}}. \]

To show that \( \{V_t : t \geq 0\} \) is a standard Wiener process, we first show that \( V_t \sim \mathcal{N}(0, t) \).

Taking expectations,

\[ E(V_t) = \frac{\sigma(1) E(W_t^{(1)}) - \sigma(2) E(W_t^{(2)})}{\sqrt{(\sigma(1)^2 + (\sigma(2)^2))}} = 0 \]
\[ E(V_t^2) = E \left[ \frac{(\sigma(1)^2 W_t^{(1)^2} + (\sigma(2)^2 W_t^{(2)^2} - 2 \sigma(1) \sigma(2) W_t^{(1)} W_t^{(2)})}{(\sigma(1)^2 + (\sigma(2)^2))} \right] \]
\[ = \frac{(\sigma(1)^2 t + (\sigma(2)^2) t - 2 \rho \sigma(1) \sigma(2) t}{(\sigma(1)^2 + (\sigma(2)^2))} \]
\[ = t. \]

Given that \( W_t^{(1)} \sim \mathcal{N}(0, t), W_t^{(2)} \sim \mathcal{N}(0, t) \) and a linear combination of normal variates is also normal, therefore \( V_t \sim \mathcal{N}(0, t) \).

To show that \( \{V_t : t \geq 0\} \) is a standard Wiener process, we note the following:

(a) \( V_0 = 0 \) and it is clear that \( V_t \) has continuous sample paths for \( t \geq 0 \).
(b) For \( t > 0, s > 0 \) we have

\[ W_{t+s}^{(1)} - W_t^{(1)} \sim \mathcal{N}(0, s), \quad W_{t+s}^{(2)} - W_t^{(2)} \sim \mathcal{N}(0, s) \]
\[ \text{Cov} \left( W_{t+s}^{(1)} - W_t^{(1)}, W_{t+s}^{(2)} - W_t^{(2)} \right) = \rho s \]

and hence

\[ V_{t+s} - V_t = \frac{\sigma(1) (W_{t+s}^{(1)} - W_t^{(1)}) - \sigma(2) (W_{t+s}^{(2)} - W_t^{(2)})}{\sqrt{(\sigma(1)^2 + (\sigma(2)^2))}}. \]
3.2.3 Multi-Dimensional Diffusion Process

Let

From the results of (a)–(c) we have shown that (c) For the SDEs

Since

Therefore,

For \( t > 0, s > 0 \), to show that \( V_{t+s} - V_t \perp V_t \) we note that

Since \( V_t \sim \mathcal{N}(0, t) \), \( V_{t+s} - V_t \sim \mathcal{N}(0, s) \) and the joint distribution of \( V_t \) and \( V_{t+s} - V_t \) is a bivariate normal (see Problem 2.2.1.5, page 58), if \( \text{Cov}(V_{t+s} - V_t, V_t) = 0 \) then \( V_{t+s} - V_t \perp V_t \).

From the results of (a)–(c) we have shown that \( V_t \) is a standard Wiener process.

\[\square\]

3. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. We consider two assets with prices \( S_t^{(1)}, S_t^{(2)} \) satisfying the SDEs

\[dS_t^{(1)} = \mu^{(1)} S_t^{(1)} dt + \sigma^{(1)} S_t^{(1)} dW_t^{(1)}\]

\[dS_t^{(2)} = \mu^{(2)} S_t^{(2)} dt + \sigma^{(2)} S_t^{(2)} dW_t^{(2)}\]

\[dW_t^{(1)} dW_t^{(2)} = \rho dt\]
3.2.3 Multi-Dimensional Diffusion Process

where \( \mu^{(1)}, \mu^{(2)}, \sigma^{(1)}, \sigma^{(2)} \) are constants and \( \{ W_t^{(1)} : t \geq 0 \}, \{ W_t^{(2)} : t \geq 0 \} \) are standard Wiener processes with correlation \( \rho \).

By letting \( U_t = S_t^{(1)} S_t^{(2)} \), show that the SDE satisfied by \( U_t \) is

\[
dU_t = \mu U_t dt + \sigma U_t dV_t
\]

where \( \mu = \mu^{(1)} + \mu^{(2)} + \rho \sigma^{(1)} \sigma^{(2)}, \sigma = \sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + 2 \rho \sigma^{(1)} \sigma^{(2)}} \) and

\[
V_t = \frac{\sigma^{(1)} W_t^{(1)} + \sigma^{(2)} W_t^{(2)}}{\sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + 2 \rho \sigma^{(1)} \sigma^{(2)}}}
\]

Show that \( \{ V_t : t \geq 0 \} \) is a standard Wiener process.

**Solution:** From Itô’s formula,

\[
dU_t = d(S_t^{(1)} S_t^{(2)})
\]

\[
= S_t^{(1)} dS_t^{(2)} + S_t^{(2)} dS_t^{(1)} + (dS_t^{(1)}) (dS_t^{(2)}) + \ldots
\]

\[
= S_t^{(1)} \left( \mu^{(1)} S_t^{(1)} dt + \sigma^{(1)} S_t^{(1)} dW_t^{(1)} \right) + S_t^{(2)} \left( \mu^{(2)} S_t^{(2)} dt + \sigma^{(2)} S_t^{(2)} dW_t^{(2)} \right)
\]

\[
+ \rho \sigma^{(1)} \sigma^{(2)} S_t^{(1)} S_t^{(2)} dt
\]

\[
= (\mu^{(1)} + \mu^{(2)} + \rho \sigma^{(1)} \sigma^{(2)}) U_t dt + U_t \left( \sigma^{(1)} dW_t^{(1)} + \sigma^{(2)} dW_t^{(2)} \right)
\]

\[
= (\mu^{(1)} + \mu^{(2)} + \rho \sigma^{(1)} \sigma^{(2)}) U_t dt + V_t dW_t
\]

Therefore,

\[
dU_t = \mu U_t dt + \sigma U_t dV_t
\]

where \( \mu = \mu^{(1)} + \mu^{(2)} + \rho \sigma^{(1)} \sigma^{(2)}, \sigma = \sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + 2 \rho \sigma^{(1)} \sigma^{(2)}} \) and

\[
V_t = \frac{\sigma^{(1)} W_t^{(1)} + \sigma^{(2)} W_t^{(2)}}{\sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + 2 \rho \sigma^{(1)} \sigma^{(2)}}}
\]

Following the same procedure as described in Problem 3.2.3.2 (page 156), we can show that

\[
V_t = \frac{\sigma^{(1)} W_t^{(1)} + \sigma^{(2)} W_t^{(2)}}{\sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + 2 \rho \sigma^{(1)} \sigma^{(2)}}} \sim N(0,t)
\]

and is also a standard Wiener process.

\[ \square \]

4. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. We consider three assets with prices \( S_t^{(1)}, S_t^{(2)} \) and \( S_t^{(3)} \) satisfying the SDEs

\[
dS_t^{(1)} = \mu^{(1)} S_t^{(1)} dt + \sigma^{(1)} S_t^{(1)} dW_t^{(1)}
\]

\[
dS_t^{(2)} = \mu^{(2)} S_t^{(2)} dt + \sigma^{(2)} S_t^{(2)} dW_t^{(2)}
\]
3.2.3 Multi-Dimensional Diffusion Process

\[ dS_t^{(3)} = \mu^{(3)} S_t^{(3)} dt + \sigma^{(3)} S_t^{(3)} dW_t^{(3)} \]

\[ dW_t^{(i)} dW_t^{(j)} = \rho_{ij} dt, \quad i \neq j \quad \text{and} \quad i, j = 1, 2, 3 \]

where \( \mu^{(1)}, \mu^{(2)}, \mu^{(3)}, \sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)} \) are constants and \( \{ W_t^{(1)} : t \geq 0 \}, \{ W_t^{(2)} : t \geq 0 \}, \{ W_t^{(3)} : t \geq 0 \} \) are standard Wiener processes with correlations \( \rho_{ij} \) and \( \sigma^{(i)} \) and \( (dW_t^{(i)})^2 = dt \) for \( i, j = 1, 2, 3 \).

By letting \( U_t = \frac{S_t^{(1)} S_t^{(2)}}{S_t^{(3)}} \), show that the SDE satisfied by \( U_t \) is

\[ dU_t = \mu U_t dt + \sigma U_t dV_t \]

where \( \mu = \mu^{(1)} + \mu^{(2)} - \mu^{(3)} + \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} - \rho_{23} \sigma^{(2)} \sigma^{(3)} \), \( \sigma^2 = (\sigma^{(1)})^2 + (\sigma^{(2)})^2 + (\sigma^{(3)})^2 + 2 \rho_{12} \sigma^{(1)} \sigma^{(2)} - 2 \rho_{13} \sigma^{(1)} \sigma^{(3)} - 2 \rho_{23} \sigma^{(2)} \sigma^{(3)} \) and

\[ V_t = \frac{\sigma^{(1)} W_t^{(1)} + \sigma^{(2)} W_t^{(2)} - \sigma^{(3)} W_t^{(3)}}{\sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + (\sigma^{(3)})^2 + 2 \rho_{12} \sigma^{(1)} \sigma^{(2)} - 2 \rho_{13} \sigma^{(1)} \sigma^{(3)} - 2 \rho_{23} \sigma^{(2)} \sigma^{(3)}}} \]

Show that \( \{ V_t : t \geq 0 \} \) is a standard Wiener process.

**Solution:** From Itô’s lemma,

\[ dU_t = \frac{\partial U_t}{\partial t} dt + \frac{\partial U_t}{\partial S_t^{(1)}} dS_t^{(1)} + \frac{\partial U_t}{\partial S_t^{(2)}} dS_t^{(2)} + \frac{\partial U_t}{\partial S_t^{(3)}} dS_t^{(3)} + \frac{1}{2} \left( \frac{\partial^2 U_t}{\partial (S_t^{(1)})^2} \right) (dS_t^{(1)})^2 + \frac{1}{2} \left( \frac{\partial^2 U_t}{\partial (S_t^{(2)})^2} \right) (dS_t^{(2)})^2 + \frac{1}{2} \left( \frac{\partial^2 U_t}{\partial (S_t^{(3)})^2} \right) (dS_t^{(3)})^2 \]

\[ + \frac{\partial^2 U_t}{\partial S_t^{(1)} \partial S_t^{(2)}} dS_t^{(1)} ds_t^{(2)} + \frac{\partial^2 U_t}{\partial S_t^{(1)} \partial S_t^{(3)}} dS_t^{(1)} ds_t^{(3)} + \frac{\partial^2 U_t}{\partial S_t^{(2)} \partial S_t^{(3)}} dS_t^{(2)} ds_t^{(3)} + \ldots \]

\[ = \frac{S_t^{(2)}}{S_t^{(3)}} \left( \mu^{(1)} S_t^{(1)} dt + \sigma^{(1)} S_t^{(1)} dW_t^{(1)} \right) + \frac{S_t^{(1)}}{S_t^{(3)}} \left( \mu^{(2)} S_t^{(2)} dt + \sigma^{(2)} S_t^{(2)} dW_t^{(2)} \right) \]

\[ - \frac{S_t^{(1)}}{S_t^{(3)}} \left( \mu^{(3)} S_t^{(3)} dt + \sigma^{(3)} S_t^{(3)} dW_t^{(3)} \right) + \frac{S_t^{(2)}}{(S_t^{(3)})^2} \left( \rho_{12} \sigma^{(1)} \sigma^{(2)} S_t^{(1)} S_t^{(2)} dt \right) \]

\[ + \frac{1}{S_t^{(3)}} \left( \rho_{12} \sigma^{(1)} \sigma^{(2)} S_t^{(1)} S_t^{(2)} dt \right) - \frac{S_t^{(2)}}{(S_t^{(3)})^2} \left( \rho_{23} \sigma^{(2)} \sigma^{(3)} S_t^{(2)} S_t^{(3)} dt \right) \]

\[ = (\mu^{(1)} + \mu^{(2)} - \mu^{(3)} + \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} - \rho_{23} \sigma^{(2)} \sigma^{(3)}) U_t dt \]

\[ + U_t \left( \sigma^{(1)} dW_t^{(1)} + \sigma^{(2)} dW_t^{(2)} - \sigma^{(3)} dW_t^{(3)} \right) \]

\[ = (\mu^{(1)} + \mu^{(2)} - \mu^{(3)} + \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} - \rho_{23} \sigma^{(2)} \sigma^{(3)}) U_t dt \]
3.2.3 Multi-Dimensional Diffusion Process

\[ + \sqrt{\sigma^2 + \sigma^2 + \rho \sigma \sigma \sigma - 2 \rho \sigma \sigma \sigma - 2 \rho \sigma \sigma \sigma} \, dV_t \]

which therefore becomes

\[ dV_t = \mu V_t \, dt + \sigma V_t \, dW_t \]

where \( \mu = \mu^1 + \mu^2 - \mu^3 + \rho \sigma \sigma \sigma - \rho \sigma \sigma \sigma - \rho \sigma \sigma \sigma \), \( \sigma^2 = (\sigma^1)^2 + (\sigma^2)^2 + (\sigma^3)^2 + 2 \rho \sigma \sigma \sigma - 2 \rho \sigma \sigma \sigma - 2 \rho \sigma \sigma \sigma \) and

\[ V_t = \frac{\sigma^1 W_t^1 - \sigma^2 W_t^2 - \sigma^3 W_t^3}{\sqrt{(\sigma^1)^2 + (\sigma^2)^2 + (\sigma^3)^2}} \]

Following the steps described in Problem 3.2.3.2 (page 156), we can easily show that

\[ V_t = \frac{\sigma^1 W_t^1 + \sigma^2 W_t^2 - \sigma^3 W_t^3}{\sqrt{(\sigma^1)^2 + (\sigma^2)^2 + (\sigma^3)^2}} \sim \mathcal{N}(0, 1) \]

and is also a standard Wiener process.

5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. We consider three assets with prices \(S_t^{(1)}, S_t^{(2)}\) and \(S_t^{(3)}\) satisfying the SDEs

\[
\begin{align*}
    dS_t^{(1)} &= \mu^{(1)} S_t^{(1)} \, dt + \sigma^{(1)} S_t^{(1)} \, dW_t^{(1)} \\
    dS_t^{(2)} &= \mu^{(2)} S_t^{(2)} \, dt + \sigma^{(2)} S_t^{(2)} \, dW_t^{(2)} \\
    dS_t^{(3)} &= \mu^{(3)} S_t^{(3)} \, dt + \sigma^{(3)} S_t^{(3)} \, dW_t^{(3)} \\
    dW_t^{(i)} dW_t^{(j)} &= \rho_{ij} dt, \quad i \neq j \quad \text{and} \quad i, j = 1, 2, 3
\end{align*}
\]

where \(\mu^{(1)}, \mu^{(2)}, \mu^{(3)}, \sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}\) are constants and \(\{W_t^{(1)} : t \geq 0\}, \{W_t^{(2)} : t \geq 0\}, \{W_t^{(3)} : t \geq 0\}\) are standard Wiener processes with correlations \(dW_t^{(i)} dW_t^{(j)} = \rho_{ij} dt, i \neq j \) and \((dW_t^{(i)})^2 = dt\) for \(i, j = 1, 2, 3\).

By letting \(U_t = S_t^{(1)} / (S_t^{(2)} S_t^{(3)})\), show that the SDE satisfied by \(U_t\) is

\[ dU_t = \mu U_t \, dt + \sigma U_t \, dV_t \]

where \(\mu = \mu^{(1)} - \mu^{(2)} + \sigma^{(1)^2} - 2\rho \sigma \sigma \sigma - 2\rho \sigma \sigma \sigma + 2\rho \sigma \sigma \sigma - 2\rho \sigma \sigma \sigma + 2\rho \sigma \sigma \sigma\) and

\[ V_t = \frac{\sigma^1 W_t^1 - \sigma^2 W_t^2 - \sigma^3 W_t^3}{\sqrt{(\sigma^1)^2 + (\sigma^2)^2 + (\sigma^3)^2}} \]

Show that \(\{V_t : t \geq 0\}\) is a standard Wiener process.
Solution: Applying Itô’s lemma,

\[
dU_i = \frac{\partial U_i}{\partial S_t^{(1)}} dS_t^{(1)} + \frac{\partial U_i}{\partial S_t^{(2)}} dS_t^{(2)} + \frac{\partial U_i}{\partial S_t^{(3)}} dS_t^{(3)}
\]
\[+ \frac{1}{2} \frac{\partial^2 U_i}{\partial (S_t^{(1)})^2} (dS_t^{(1)})^2 + \frac{1}{2} \frac{\partial^2 U_i}{\partial (S_t^{(2)})^2} (dS_t^{(2)})^2 + \frac{1}{2} \frac{\partial^2 U_i}{\partial (S_t^{(3)})^2} (dS_t^{(3)})^2
\]
\[+ \frac{\partial^2 U_i}{\partial S_t^{(1)} \partial S_t^{(2)}} dS_t^{(1)} dS_t^{(2)} + \frac{\partial^2 U_i}{\partial S_t^{(1)} \partial S_t^{(3)}} dS_t^{(1)} dS_t^{(3)} + \frac{\partial^2 U_i}{\partial S_t^{(2)} \partial S_t^{(3)}} dS_t^{(2)} dS_t^{(3)} + \ldots
\]

\[= \frac{1}{S_t^{(1)} S_t^{(3)}} \left( \mu^{(1)} S_t^{(1)} dt + \sigma^{(1)} S_t^{(1)} dW_t^{(1)} \right) - \frac{S_t^{(1)}}{(S_t^{(2)})^2 S_t^{(3)}} \left( \mu^{(2)} S_t^{(2)} dt + \sigma^{(2)} S_t^{(2)} dW_t^{(2)} \right)
\]
\[+ \frac{S_t^{(1)}}{S_t^{(2)} (S_t^{(3)})^2} \left( (\sigma^{(3)} S_t^{(3)})^2 dt \right) - \frac{1}{(S_t^{(2)})^2 S_t^{(3)}} \left( \rho_{12} \sigma^{(1)} \sigma^{(2)} S_t^{(1)} S_t^{(2)} dt \right)
\]
\[- \frac{1}{S_t^{(2)} (S_t^{(3)})^2} \left( \rho_{13} \sigma^{(1)} \sigma^{(3)} S_t^{(1)} S_t^{(3)} dt \right) + \frac{S_t^{(1)}}{(S_t^{(2)})^3 S_t^{(3)}} \left( \rho_{23} \sigma^{(2)} \sigma^{(3)} S_t^{(2)} S_t^{(3)} dt \right)
\]

\[= (\mu^{(1)} - \mu^{(2)} - \mu^{(3)}) + (\sigma^{(2)})^2 - \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} + \rho_{23} \sigma^{(2)} \sigma^{(3)}) U_i dt
\]
\[+ U_i \left( \sigma^{(1)} dW_t^{(1)} - \sigma^{(2)} dW_t^{(2)} - \sigma^{(3)} dW_t^{(3)} \right)
\]
\[= (\mu^{(1)} - \mu^{(2)} - \mu^{(3)}) + (\sigma^{(2)})^2 - \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} + \rho_{23} \sigma^{(2)} \sigma^{(3)}) U_i dt
\]
\[+ \sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + (\sigma^{(3)})^2} \sim \mathcal{N}(0, t)
\]

We can therefore write

\[dU_i = \mu U_i dt + \sigma U_i dV_i
\]

where \( \mu = \mu^{(1)} - \mu^{(2)} - \mu^{(3)} + (\sigma^{(2)})^2 - \rho_{12} \sigma^{(1)} \sigma^{(2)} - \rho_{13} \sigma^{(1)} \sigma^{(3)} + \rho_{23} \sigma^{(2)} \sigma^{(3)}, \)
\( \sigma^2 = (\sigma^{(1)})^2 + (\sigma^{(2)})^2 + (\sigma^{(3)})^2 - 2 \rho_{12} \sigma^{(1)} \sigma^{(2)} - 2 \rho_{13} \sigma^{(1)} \sigma^{(3)} + 2 \rho_{23} \sigma^{(2)} \sigma^{(3)}, \)
\( V_i = \frac{\sigma^{(1)} W_t^{(1)} - \sigma^{(2)} W_t^{(2)} - \sigma^{(3)} W_t^{(3)}}{\sqrt{(\sigma^{(1)})^2 + (\sigma^{(2)})^2 + (\sigma^{(3)})^2 - 2 \rho_{12} \sigma^{(1)} \sigma^{(2)} - 2 \rho_{13} \sigma^{(1)} \sigma^{(3)} + 2 \rho_{23} \sigma^{(2)} \sigma^{(3)}}}.
\]

As described in Problem 3.2.3.2 (page 156) we can easily show that

\( V_i \sim \mathcal{N}(0, t) \)

and is a standard Wiener process.
3.2.3 Multi-Dimensional Diffusion Process

6. Bessel Process. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})\) be an \(n\)-dimensional standard Wiener process, with \(W_t^{(i)}, i = 1, 2, \ldots, n\) being independent one-dimensional standard Wiener processes. By setting

\[
X_t = \sqrt{(W_t^{(1)})^2 + (W_t^{(2)})^2 + \ldots + (W_t^{(n)})^2}
\]

show that

\[
W_t = \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} + \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} + \ldots + \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)}
\]

is a standard Wiener process.

Show also that \(X_t\) satisfies the \(n\)-dimensional Bessel process with SDE

\[
dX_t = \left(\frac{n-1}{2X_t}\right) dt + dW_t
\]

where \(X_0 = 0\).

Solution: We first need to show that \(W_t \sim \mathcal{N}(0, t)\). Using the properties of the stochastic Itô integral and the independence property of \(W_t^{(i)}, i = 1, 2, \ldots, n\),

\[
\mathbb{E}(W_t) = \mathbb{E} \left[ \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} + \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} + \ldots + \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right]
\]

\[
= \mathbb{E} \left( \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} \right) + \mathbb{E} \left( \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} \right) + \ldots + \mathbb{E} \left( \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right)
\]

\[
= 0
\]

\[
\mathbb{E}(W_t^2) = \mathbb{E} \left[ \left( \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} \right)^2 + \left( \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} \right)^2 + \ldots + \left( \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right)^2 \right]
\]

\[
+ 2 \sum_{i=1}^n \sum_{j=1}^n \left( \int_0^t \frac{W_s^{(i)}}{X_s} \, dW_s^{(i)} \right) \left( \int_0^t \frac{W_s^{(j)}}{X_s} \, dW_s^{(j)} \right)
\]

\[
= \mathbb{E} \left[ \left( \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} \right)^2 \right] + \mathbb{E} \left[ \left( \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} \right)^2 \right] \]

\[
+ \ldots + \mathbb{E} \left[ \left( \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right)^2 \right]
\]
Because $W_t$ is a function of $n$ independent normal variates, so $W_t \sim \mathcal{N}(0, t)$. To show that $W_t$ is also a standard Wiener process, we note the following:

(a) $W_0 = \lim_{t \to 0} \int_0^t \frac{W_s^{(1)}}{X_s} \, ds + \int_0^t \frac{W_s^{(2)}}{X_s} \, ds + \cdots + \int_0^t \frac{W_s^{(n)}}{X_s} \, ds = 0$ and it is clear that $W_t$ has continuous sample paths for $t \geq 0$.

(b) For $t > 0, u > 0$

$$W_{t+u} - W_t = \int_0^{t+u} \frac{W_s^{(1)}}{X_s} \, ds + \int_0^{t+u} \frac{W_s^{(2)}}{X_s} \, ds + \cdots + \int_0^{t+u} \frac{W_s^{(n)}}{X_s} \, ds - \int_0^t \frac{W_s^{(1)}}{X_s} \, ds - \int_0^t \frac{W_s^{(2)}}{X_s} \, ds - \cdots - \int_0^t \frac{W_s^{(n)}}{X_s} \, ds$$

$$= \int_t^{t+u} \frac{W_s^{(1)}}{X_s} \, ds + \int_t^{t+u} \frac{W_s^{(2)}}{X_s} \, ds + \cdots + \int_t^{t+u} \frac{W_s^{(n)}}{X_s} \, ds.$$

Taking expectations,

$$\mathbb{E}(W_{t+u} - W_t) = \mathbb{E}\left( \int_t^{t+u} \frac{W_s^{(1)}}{X_s} \, ds \right) + \mathbb{E}\left( \int_t^{t+u} \frac{W_s^{(2)}}{X_s} \, ds \right) + \cdots + \mathbb{E}\left( \int_t^{t+u} \frac{W_s^{(n)}}{X_s} \, ds \right)$$

$$= 0$$

and due to the independence of $W_t^{(i)}, i = 1, 2, \ldots, n$,

$$\mathbb{E}\left[ (W_{t+u} - W_t)^2 \right]$$

$$= \mathbb{E}\left[ \left( \int_t^{t+u} \frac{W_s^{(1)}}{X_s} \, ds \right)^2 + \left( \int_t^{t+u} \frac{W_s^{(2)}}{X_s} \, ds \right)^2 + \cdots + \left( \int_t^{t+u} \frac{W_s^{(n)}}{X_s} \, ds \right)^2 \right]$$

$$+ 2 \sum_{i=1}^n \sum_{j=1}^n \left( \int_t^t \frac{W_s^{(i)}}{X_s} \, ds \right) \left( \int_t^{t+u} \frac{W_s^{(j)}}{X_s} \, ds \right)$$
3.2.3 Multi-Dimensional Diffusion Process

\[
E \left[ \left( \int_t^{t+u} \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} \right)^2 \right] + E \left[ \left( \int_t^{t+u} \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} \right)^2 \right]
\]

\[
+ \ldots + E \left[ \left( \int_t^{t+u} \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right)^2 \right]
\]

\[
= E \left[ \int_t^{t+u} \left( \frac{W_s^{(1)}}{X_s} \right)^2 \, ds \right] + E \left[ \int_t^{t+u} \left( \frac{W_s^{(2)}}{X_s} \right)^2 \, ds \right] + \ldots + E \left[ \int_t^{t+u} \left( \frac{W_s^{(n)}}{X_s} \right)^2 \, ds \right]
\]

\[
= E \left[ \int_t^{t+u} ds \right]
\]

\[
= t - t = 0.
\]

Thus, \( W_{t+u} - W_t \sim \mathcal{N}(0, u) \).

(c) To show that \( W_{t+u} - W_t \perp W_t \) we note that from the independent increment property of \( W_t^{(i)} \), \( i = 1, 2, \ldots, n, \)

\[
E \left[ (W_{t+u} - W_t) \, W_t \right]
\]

\[
= E \left( W_{t+u} W_t \right) - E \left( W_t^2 \right)
\]

\[
= E \left[ \left( \int_0^{t+u} \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} \right) + \int_0^{t+u} \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} + \ldots + \int_0^{t+u} \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right]
\]

\[
\times \left( \int_0^t \frac{W_s^{(1)}}{X_s} \, dW_s^{(1)} + \int_0^t \frac{W_s^{(2)}}{X_s} \, dW_s^{(2)} + \ldots + \int_0^t \frac{W_s^{(n)}}{X_s} \, dW_s^{(n)} \right) \right] - E\left( W_t^2 \right)
\]

\[
= \sum_{i=1}^n \left[ \left( \int_0^{t+u} \frac{W_s^{(i)}}{X_s} \, dW_s^{(i)} \right)^2 \right] + \sum_{i=1}^n \sum_{j=1}^n \left[ \left( \int_0^{t+u} \frac{W_s^{(i)}}{X_s} \, dW_s^{(i)} \right) \left( \int_0^t \frac{W_s^{(j)}}{X_s} \, dW_s^{(j)} \right) \right] - E\left( W_t^2 \right)
\]

\[
= \sum_{i=1}^n \left[ \int_0^{t} \left( \frac{W_s^{(i)}}{X_s} \right)^2 \, ds \right] - E\left( W_t^2 \right)
\]

\[
= t - t = 0.
\]

Since \( W_t \sim \mathcal{N}(0, t) \), \( W_{t+u} - W_t \sim \mathcal{N}(0, u) \) and the joint distribution of \( W_t \) and \( W_{t+u} - W_t \) is a bivariate normal (see Problem 2.2.1.5, page 58), if \( \text{Cov}(W_{t+u} - W_t, W_t) = 0 \) then \( W_{t+u} - W_t \perp W_t \).
Thus, from the results of (a)–(c) we have shown that $W_t$ is a standard Wiener process. From Itô’s formula and given $dW_t^i dW_t^j = 0$, $i \neq j$,

$$dX_i = \sum_{i=1}^n \frac{\partial X_i}{\partial W_t^i} dW_t^i + \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 X_i}{\partial (W_t^i)^2} (dW_t^i)^2$$

$$= \sum_{i=1}^n \frac{\partial X_i}{\partial W_t^i} dW_t^i + \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 X_i}{\partial (W_t^i)^2} dt$$

$$= \sum_{i=1}^n \frac{W_t^i}{X_t} dW_t^i + \frac{1}{2} \sum_{i=1}^n \frac{X_t - X_t^{-1} (W_t^i)^2}{X_t^3} dt$$

$$= dW_t + \frac{1}{2} n \frac{X_t^2}{X_t^3} - \sum_{i=1}^n \left( \frac{W_t^i)^2}{X_t^3} \right) dt$$

Therefore,

$$X_t = \sqrt{(W_t^{(1)})^2 + (W_t^{(2)})^2 + \ldots + (W_t^{(n)})^2}$$

satisfies the SDE

$$dX_t = \left( \frac{n-1}{2X_t} \right) dt + dW_t.$$

7. **Forward–Spot Price Relationship II.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We consider the forward price $F(t,T)$ of an asset $S_t$ satisfying the SDE

$$\frac{dF(t,T)}{F(t,T)} = \sigma_1(t,T) dW_t^{(1)} + \sigma_2(t,T) dW_t^{(2)}$$

where $\sigma_1(t,T) > 0$, $\sigma_2(t,T) > 0$ are time-dependent volatilities and $\{W_t^{(1)} : t \geq 0\}$, $\{W_t^{(2)} : t \geq 0\}$ are standard Wiener processes with correlation $\rho \in (-1, 1)$. Given the relationship $F(t,T) = S_t e^{r(T-t)}$ where $r$ is the risk-free interest rate, show that

$$\frac{dS_t}{S_t} = \left\{ \frac{\partial \log F(0,t)}{\partial t} - \int_0^t \left[ \sigma_1(u,t) \frac{\partial \sigma_1(u,t)}{\partial t} + \rho \left( \sigma_1(u,t) \frac{\partial \sigma_2(u,t)}{\partial t} + \sigma_2(u,t) \frac{\partial \sigma_1(u,t)}{\partial t} \right) \right] du + \int_0^t \frac{\partial \sigma_2(u,t)}{\partial t} dW_u^{(1)} + \int_0^t \frac{\partial \sigma_2(u,t)}{\partial t} dW_u^{(2)} \right\} dt + \sigma_1(t,T) dW_t^{(1)} + \sigma_2(t,T) dW_t^{(2)}.$$
where
\[ \sigma(t, t) = \sqrt{\sigma_1(t, t)^2 + 2\rho\sigma_1(t, t)\sigma_2(t, t) + \sigma_2(t, t)^2} \]
and
\[ W_t = \frac{\sigma_1(t, t)W_t^{(1)} + \sigma_2(t, t)W_t^{(2)}}{\sqrt{\sigma_1(t, t)^2 + 2\rho\sigma_1(t, t)\sigma_2(t, t) + \sigma_2(t, t)^2}} \]
is a standard Wiener process.

**Solution:** Expanding \( \log F(t, T) \) using Taylor’s theorem and then applying Itô’s lemma, we have
\[
d\log F(t, T) = \frac{1}{F(t, T)}dF(t, T) - \frac{1}{2F(t, T)^2}dF(t, T)^2 + \ldots
\]
\[
= \sigma_1(t, T) dW_t^{(1)} + \sigma_2(t, T) dW_t^{(2)}
\]
\[
- \frac{1}{2} \left[ \sigma_1(t, T)^2 + \sigma_2(t, T)^2 + 2\rho\sigma_1(t, T)\sigma_2(t, T) \right] dt
\]
and taking integrals,
\[
\log \left( \frac{F(t, T)}{F(0, T)} \right) = \int_0^t \sigma_1(u, T) dW_u^{(1)} + \int_0^t \sigma_2(u, T) dW_u^{(2)}
\]
\[
- \frac{1}{2} \int_0^t \left[ \sigma_1(u, T)^2 + \sigma_2(u, T)^2 + 2\rho\sigma_1(u, T)\sigma_2(u, T) \right] du
\]
We finally have
\[
F(t, T) = F(0, T)e^{-\frac{1}{2} \int_0^t \left[ \sigma_1(u, T)^2 + \sigma_2(u, T)^2 + 2\rho\sigma_1(u, T)\sigma_2(u, T) \right] du + \int_0^t \sigma_1(u, T) dW_u^{(1)} + \int_0^t \sigma_2(u, T) dW_u^{(2)}}.
\]
By setting \( T = t \) and taking note that \( S_t = F(t, t) \), the spot price \( S_t \) has the expression
\[
S_t = F(0, t)e^{-\frac{1}{2} \int_0^t \left[ \sigma_1(u, T)^2 + \sigma_2(u, T)^2 + 2\rho\sigma_1(u, T)\sigma_2(u, T) \right] du + \int_0^t \sigma_1(u, T) dW_u^{(1)} + \int_0^t \sigma_2(u, T) dW_u^{(2)}}.
\]
To find the SDE for \( S_t \), we apply Itô’s lemma
\[
dS_t = \frac{\partial S_t}{\partial t}dt + \frac{\partial S_t}{\partial W_t^{(1)}}dW_t^{(1)} + \frac{\partial S_t}{\partial W_t^{(2)}}dW_t^{(2)}
\]
\[
+ \frac{1}{2} \frac{\partial^2 S_t}{\partial (W_t^{(1)})^2}(dW_t^{(1)})^2 + \frac{1}{2} \frac{\partial^2 S_t}{\partial (W_t^{(2)})^2}(dW_t^{(2)})^2 + \frac{\partial^2 S_t}{\partial W_t^{(1)}\partial W_t^{(2)}}dW_t^{(1)}dW_t^{(2)} + \ldots
\]
\[
= \left( \frac{\partial S_t}{\partial t} + \frac{1}{2} \frac{\partial^2 S_t}{\partial (W_t^{(1)})^2} + \frac{1}{2} \frac{\partial^2 S_t}{\partial (W_t^{(2)})^2} + \rho \frac{\partial^2 S_t}{\partial W_t^{(1)}\partial W_t^{(2)}} \right) dt + \frac{\partial S_t}{\partial W_t^{(1)}}dW_t^{(1)}
\]
\[+ \frac{\partial S_t}{\partial W_t^{(2)}}dW_t^{(2)}. \]
3.2.3 Multi-Dimensional Diffusion Process

Taking partial differentiations of $S_t$, we have

$$
\frac{\partial S_t}{\partial t} = \frac{\partial \log F(0, t)}{\partial t} - \frac{1}{2} \left[ \sigma_1(t, t)^2 + \sigma_2(t, t)^2 + 2 \rho \sigma_1(t, t) \sigma_2(t, t) \right] \\
- \int_0^t \left[ \sigma_1(u, t) \frac{\partial \sigma_1(u, t)}{\partial t} + \rho \left( \sigma_1(u, t) \frac{\partial \sigma_2(u, t)}{\partial t} + \sigma_2(u, t) \frac{\partial \sigma_1(u, t)}{\partial t} \right) \\
+ \sigma_2(u, t) \frac{\partial \sigma_2(u, t)}{\partial t} \right] du \\
+ \int_0^t \frac{\partial \sigma_1(u, t)}{\partial t} dW_u^{(1)} + \int_0^t \frac{\partial \sigma_2(u, t)}{\partial t} dW_u^{(2)}, \\
$$

$$
\frac{\partial S_t}{\partial W_i^{(1)}} = \sigma_1(t, t) S_t, \\
\frac{\partial S_t}{\partial W_i^{(2)}} = \sigma_2(t, t) S_t, \\
\frac{\partial^2 S_t}{\partial (W_i^{(1)})^2} = \sigma_1(t, t)^2 S_t, \\
\frac{\partial^2 S_t}{\partial (W_i^{(2)})^2} = \sigma_2(t, t)^2 S_t, \\
\frac{\partial^2 S_t}{\partial W_i^{(1)} \partial W_i^{(2)}} = \sigma_1(t, t) \sigma_2(t, t) S_t
$$

and substituting them into the SDE we eventually have

$$
\frac{dS_t}{S_t} = \left\{ \frac{\partial \log F(0, t)}{\partial t} - \int_0^t \left[ \sigma_1(u, t) \frac{\partial \sigma_1(u, t)}{\partial t} + \rho \left( \sigma_1(u, t) \frac{\partial \sigma_2(u, t)}{\partial t} + \sigma_2(u, t) \frac{\partial \sigma_1(u, t)}{\partial t} \right) \\
+ \sigma_2(u, t) \frac{\partial \sigma_2(u, t)}{\partial t} \right] du \\
+ \int_0^t \frac{\partial \sigma_1(u, t)}{\partial t} dW_u^{(1)} + \int_0^t \frac{\partial \sigma_2(u, t)}{\partial t} dW_u^{(2)} \right\} dt \\
+ \int_0^t \frac{\partial \sigma_1(u, t)}{\partial t} dW_u^{(1)} + \int_0^t \frac{\partial \sigma_2(u, t)}{\partial t} dW_u^{(2)} dt \\
+ \sigma(t, t) dW_t
$$

where

$$
\sigma(t, t) = \sqrt{\sigma_1(t, t)^2 + 2 \rho \sigma_1(t, t) \sigma_2(t, t) + \sigma_2(t, t)^2}
$$

and from the steps discussed in Problem 3.2.3.2 (page 156) we can prove that

$$
W_i = \frac{\sigma_1(t, t) W_i^{(1)} + \sigma_2(t, t) W_i^{(2)}}{\sqrt{\sigma_1(t, t)^2 + 2 \rho \sigma_1(t, t) \sigma_2(t, t) + \sigma_2(t, t)^2}} \sim \mathcal{N}(0, t)
$$

and is also a standard Wiener process.

\[\Box\]
8. **Gabillon 2-Factor Model.** Let \( \{ W^{(1)}_t : t \geq 0 \} \) and \( \{ W^{(2)}_t : t \geq 0 \} \) be standard Wiener processes on the probability space \((\Omega, \mathcal{F}, P)\) with correlation \( \rho \in (-1, 1) \). Suppose the forward curve \( F(t, T) \) follows the process

\[
\frac{dF(t, T)}{F(t, T)} = \sigma_s e^{-\alpha(t-T)} dW^{(s)}_t + \sigma_l (1 - e^{-\alpha(T-t)}) dW^{(l)}_t
\]

where \( t < T, \alpha \) is the mean-reversion parameter and \( \sigma_s \) and \( \sigma_l \) are the short-term and long-term volatilities, respectively.

By setting

\[
dW^{(s)}_t = dW^{(1)}_t \quad \text{and} \quad dW^{(l)}_t = \rho dW^{(1)}_t + \sqrt{1 - \rho^2} dW^{(2)}_t
\]

where \( W^{(1)}_t \) and \( W^{(2)}_t \) are standard Wiener processes, \( W^{(1)}_t \perp W^{(2)}_t \), show that

\[
\frac{dF(t, T)}{F(t, T)} = \sigma(t, T) \, dW_t
\]

where

\[
\sigma(t, T) = \sqrt{\sigma_s^2 + (\sigma_l^2 - 2 \rho \sigma_s \sigma_l + \alpha^2) e^{-2\alpha(t-T)} + 2 \left( \rho \sigma_s \sigma_l - \alpha \right) e^{-\alpha(T-t)}}
\]

and

\[
W_t = \left( \frac{\sigma_s e^{-\alpha(t-T)} + \rho \sigma_l (1 - e^{-\alpha(T-t)})}{\sqrt{\sigma_l^2 + (\sigma_l^2 - 2 \rho \sigma_s \sigma_l + \sigma_s^2) e^{-2\alpha(T-t)} + 2 \left( \rho \sigma_s \sigma_l - \sigma_s \right) e^{-\alpha(T-t)}}} \right) W^{(1)}_t + \frac{\sqrt{1 - \rho^2} \left( 1 - e^{-\alpha(T-t)} \right)}{\sqrt{\sigma_l^2 + (\sigma_l^2 - 2 \rho \sigma_s \sigma_l + \sigma_s^2) e^{-2\alpha(T-t)} + 2 \left( \rho \sigma_s \sigma_l - \sigma_s \right) e^{-\alpha(T-t)}}} W^{(2)}_t
\]

is a standard Wiener process.

Finally, show that conditional on \( F(0, T), F(t, T) \) follows a lognormal distribution with mean

\[
\mathbb{E} \left[ F(t, T) \mid F(0, T) \right] = F(0, T)
\]

and variance

\[
\text{Var} \left[ F(t, T) \mid F(0, T) \right] = F(0, T)^2 \exp \left\{ \sigma_s^2 t + \left( \sigma_l^2 - 2 \rho \sigma_s \sigma_l + \sigma_s^2 \right) \frac{\left( e^{-2\alpha(T-t)} - e^{-2\alpha T} \right)}{2\alpha} + 2 \left( \rho \sigma_s \sigma_l - \sigma_s \right) \frac{\left( e^{-\alpha(T-t)} - e^{-\alpha T} \right)}{\alpha} - 1 \right\}.
\]

**Solution:** By defining \( dW^{(s)}_t = dW^{(1)}_t \) and \( dW^{(l)}_t = \rho dW^{(1)}_t + \sqrt{1 - \rho^2} dW^{(2)}_t \) such that \( W^{(1)}_t \perp W^{(2)}_t \), the SDE of \( F(t, T) \) can be expressed as

\[
\frac{dF(t, T)}{F(t, T)} = \sigma_s e^{-\alpha(t-T)} dW^{(s)}_t + \alpha_l (1 - e^{-\alpha(T-t)}) \left( \rho dW^{(1)}_t + \sqrt{1 - \rho^2} dW^{(2)}_t \right)
\]

\[
= \left( \sigma_s e^{-\alpha(t-T)} + \rho \sigma_l (1 - e^{-\alpha(T-t)}) \right) dW^{(1)}_t + \sqrt{1 - \rho^2} \sigma_l (1 - e^{-\alpha(T-t)}) dW^{(2)}_t
\]

\[
= \sigma(t, T) \, dW_t
\]
where, owing to the fact that \(dW_t^{(1)} \cdot dW_t^{(2)} = 0\), we have

\[
\sigma(t, T)^2 = (\sigma_t e^{-a(T-t)} + \rho \sigma_t (1 - e^{-a(T-t)})^2) + (1 - \rho^2) \sigma_t^2 (1 - e^{-a(T-t)})^2
\]

\[
= \sigma_t^2 e^{-2a(T-t)} + 2\rho \sigma_t \sigma e^{-a(T-t)} (1 - e^{-a(T-t)}) + \sigma_t^2 (1 - e^{-a(T-t)})^2
\]

\[
= \sigma_t^2 + (\sigma_t^2 - 2\rho \sigma_t \sigma + \sigma_t^2) e^{-2a(T-t)} + 2(\rho \sigma_t \sigma - \sigma_t^2) e^{-a(T-t)}
\]

and using the steps described in Problem 3.2.3.2 (page 156), we can easily show that

\[
W_t = \left( \sigma_t e^{-a(T-t)} + \rho \sigma_t (1 - e^{-a(T-t)}) \right) W_t^{(1)} + \sqrt{1 - \rho^2} \sigma_t (1 - e^{-a(T-t)}) W_t^{(2)}
\]

\[
\sim \mathcal{N}(0, t)
\]

is also a standard Wiener process.

By expanding \(d \log F(t, T)\) using Taylor’s theorem and applying Itô’s formula, we have

\[
d \log F(t, T) = \frac{dF(t, T)}{F(t, T)} - \frac{1}{2} \left( \frac{dF(t, T)}{F(t, T)} \right)^2 + \ldots
\]

\[
= \sigma(t, T) dW_t - \frac{1}{2} \sigma(t, T)^2 dt.
\]

Taking integrals from 0 to \(t\),

\[
\int_0^t d \log F(u, T) = \int_0^t \sigma(u, T) dW_u - \frac{1}{2} \int_0^t \sigma(u, T)^2 du
\]

\[
\log F(t, T) = \log F(0, T) + \int_0^t \sigma(u, T) dW_u - \frac{1}{2} \int_0^t \sigma(u, T)^2 du.
\]

From the property of Itô’s integral, we have

\[
\mathbb{E} \left[ \int_0^t \sigma(u, T) dW_u \right] = 0
\]

and

\[
\mathbb{E} \left[ \left( \int_0^t \sigma(u, T) dW_u \right)^2 \right] = \mathbb{E} \left[ \int_0^t \sigma(u, T)^2 du \right] = \int_0^t \sigma(u, T)^2 du.
\]

Since the Itô integral \(\int_0^t \sigma(u, T) dW_u\) is in the form \(\int_0^t f(u) dW_u\), from Problem 3.2.2.4 (page 126) we can easily prove that \(\int_0^t \sigma(u, T) dW_u\) follows a normal distribution,

\[
\int_0^t \sigma(u, T) dW_u \sim \mathcal{N} \left( 0, \int_0^t \sigma(u, T)^2 du \right)
\]

and hence

\[
\log F(t, T) \sim \mathcal{N} \left( \log F(0, T) - \frac{1}{2} \int_0^t \sigma(u, T)^2 du, \int_0^t \sigma(u, T)^2 du \right).
\]
3.2.3 Multi-Dimensional Diffusion Process

Solving the integral,

\[
\int_0^t \sigma(u, T)^2 \, du = \int_0^t \left[ \sigma_t^2 + \left( \sigma_t^2 - 2 \rho \sigma \sigma_l + \sigma_l^2 \right) e^{-2\alpha(T-u)} + 2 \left( \rho \sigma \sigma_l - \sigma_l^2 \right) e^{-\alpha(T-u)} \right] \, du \\
= \sigma_t^2 u + \left( \sigma_t^2 - 2 \rho \sigma \sigma_l + \sigma_l^2 \right) \frac{e^{-2\alpha(T-t)}}{2\alpha} + 2 \left( \rho \sigma \sigma_l - \sigma_l^2 \right) \frac{e^{-\alpha(T-t)}}{\alpha} \bigg|_0^t \\
= \sigma_t^2 t + \left( \sigma_t^2 - 2 \rho \sigma \sigma_l + \sigma_l^2 \right) \left( \frac{e^{-2\alpha(T-t)}}{2\alpha} - \frac{e^{-\alpha(T-t)}}{\alpha} \right) \\
+ 2 \left( \rho \sigma \sigma_l - \sigma_l^2 \right) \left( \frac{e^{-\alpha(T-t)}}{\alpha} - 1 \right).
\]

Since \( F(t, T) \) conditional on \( F(0, T) \) follows a lognormal distribution, we have

\[
\mathbb{E}[F(t, T)|F(0, T)] = F(0, T)
\]

and

\[
\text{Var} \{F(t, T)|F(0, T)\} = F(0, T)^2 \exp \left\{ \sigma_t^2 t + \left( \sigma_t^2 - 2 \rho \sigma \sigma_l + \sigma_l^2 \right) \left( \frac{e^{-2\alpha(T-t)}}{2\alpha} - \frac{e^{-\alpha(T-t)}}{\alpha} \right) \right. \\
+ 2 \left( \rho \sigma \sigma_l - \sigma_l^2 \right) \left( 1 - \frac{e^{-\alpha(T-t)}}{\alpha} \right) \left. \right\}.
\]

9. Integrated Square-Root Process. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W_t : t \geq 0\} \) be a standard Wiener process. Suppose \( X_t \) follows the CIR model with SDE

\[
dX_t = \kappa(\theta - X_t) \, dt + \sigma \sqrt{X_t} \, dW_t, \quad X_0 > 0
\]

where \( \kappa, \theta \) and \( \sigma \) are constants. We consider the integral

\[
Y_t = \int_0^t X_u \, du, \quad X_0 > 0
\]

as the integrated square-root process of \( X_t \) up to time \( t \) from initial time \( t_0, t_0 < t \).

Show that for \( n \geq 1, m \geq 1, n, m \in \mathbb{N} \), the process \( X_t^n Y_t^m \) satisfies the SDE

\[
\frac{d(X_t^n Y_t^m)}{X_t^n Y_t^m} = \left[ \left( n \kappa (\theta - X_t) + \frac{1}{2} n(n-1) \sigma^2 \right) X_t^{-1} + m X_t^{-1} \right] dt + n \sigma X_t^{-1} \, dW_t
\]

and hence show that \( \mathbb{E}(X_t^n Y_t^m | X_{t_0}) \) satisfies the following first-order ordinary differential equation

\[
\frac{d}{dt} \mathbb{E} \left( X_t^n Y_t^m | X_{t_0} \right) = n \kappa \theta \mathbb{E} \left( X_t^{n-1} Y_t^m | X_{t_0} \right) - n \kappa \mathbb{E} \left( X_t^n Y_t^m | X_{t_0} \right) \\
+ \frac{1}{2} n(n-1) \sigma^2 \mathbb{E} \left( X_t^n Y_t^m | X_{t_0} \right) + m \mathbb{E} \left( X_t^{n+1} Y_t^{m-1} | X_{t_0} \right).
\]

Finally, find the first two moments of \( Y_t \), given \( X_{t_0} \).
3.2.3 Multi-Dimensional Diffusion Process

Solution: From Taylor’s theorem and subsequently applying Itô’s formula, we can write

\[
d(X^n_t) = nX^n_t \, dX_t + \frac{1}{2} n(n-1)X^{n-2}_t (dX_t^2) + \ldots
\]

\[
= nX^n_t \, [\kappa (\theta - X_t) \, dt + \sigma \sqrt{X_t} \, dW_t] + \frac{1}{2} n(n-1)X^{n-2}_t (\sigma^2 X_t \, dt)
\]

\[
= \left[ n \kappa (\theta - X_t) X^{n-1}_t + \frac{1}{2} n(n-1)\sigma^2 X^{n-1}_t \right] dt + n\sigma X^{n-\frac{1}{2}}_t dW_t
\]

and

\[
d(Y^m_t) = mY^m_t \, dY_t + \frac{1}{2} m(m-1)Y^{m-2}_t (dY_t)^2 + \ldots
\]

\[
= mX_t Y^{m-1}_t \, dt.
\]

Thus,

\[
d(X^n_t Y^m_t) = Y^m \, d(X^n_t) + X^n_t \, d(Y^m_t) + d(X^n_t) d(Y^m_t)
\]

\[
= \left[ \left( n \kappa (\theta - X_t) + \frac{1}{2} n(n-1)\sigma^2 \right) X^{n-1}_t Y^m_t \right] dt + n\sigma X^{n-\frac{1}{2}}_t Y^m_t dW_t
\]

\[+ mX^{n+1}_t Y^{m-1}_t \, dt
\]

or

\[
d(X^n_t Y^m_t) = \left[ \left( n \kappa (\theta - X_t) + \frac{1}{2} n(n-1)\sigma^2 \right) X^{n-1}_t Y^m_t + mX_t Y^{m-1}_t \right] \, dt + n\sigma X^{n-\frac{1}{2}}_t Y^m_t \, dW_t.
\]

Taking integrals,

\[
\int_{t_0}^t d(X^n_u Y^m_u) = \int_{t_0}^t \left[ n \kappa (\theta - X_u) X^{n-1}_u Y^m_u + \frac{1}{2} n(n-1)\sigma^2 X^{n-1}_u Y^m_u + mX^{n+1}_u Y^{m-1}_u \right] \, du
\]

\[+ \int_{t_0}^t n\sigma X^{n-\frac{1}{2}}_u \, Y^m_u \, dW_u
\]

\[
X^n_t Y^m_t = X^n_{t_0} Y^m_{t_0} + \int_{t_0}^t \left[ n \kappa (\theta - X_u) X^{n-1}_u Y^m_u + \frac{1}{2} n(n-1)\sigma^2 X^{n-1}_u Y^m_u + mX^{n+1}_u Y^{m-1}_u \right] \, du
\]

\[+ \int_{t_0}^t n\sigma X^{n-\frac{1}{2}}_u \, Y^m_u \, dW_u
\]

and taking expectations given \(X_{t_0}\) and because \(Y_{t_0} = \int_{t_0}^t X_u \, du = 0\), we have

\[
\mathbb{E}[X^n_t Y^m_t | X_{t_0}] = \int_{t_0}^t \mathbb{E}[n \kappa (\theta - X_u) X^{n-1}_u Y^m_u | X_{t_0}] \, du + \int_{t_0}^t \mathbb{E} \left[ \frac{1}{2} n(n-1)\sigma^2 X^{n-1}_u Y^m_u \right] \, du
\]

\[+ \int_{t_0}^t n\sigma X^{n-\frac{1}{2}}_u \, Y^m_u \, dW_u
\]
3.2.3 Multi-Dimensional Diffusion Process

\[ + \int_{t_0}^t \mathbb{E} \left[ m X_u^{n+1} Y_{u-1} \big| X_{t_0} \right] du + \int_{t_0}^t \mathbb{E} \left[ n \sigma X_u^{n-1} Y_u \right] X_{t_0} \right] dW_u \]

\[ = \int_{t_0}^t \mathbb{E} \left[ n \kappa (\theta - X_u) X_u^{n-1} Y_u \big| X_{t_0} \right] du + \int_{t_0}^t \mathbb{E} \left[ \frac{1}{2} n(n-1) \sigma^2 X_u^{n-1} Y_u \big| X_{t_0} \right] du \]

\[ + \int_{t_0}^t \mathbb{E} \left[ m X_u^{n+1} Y_{u-1} \big| X_{t_0} \right] du. \]

Differentiating with respect to \( t \) yields

\[ \frac{d}{dt} \mathbb{E} \left[ X_t^n Y_t \big| X_{t_0} \right] = n \kappa \mathbb{E} \left[ (\theta - X_t) X_t^{n-1} Y_t \big| X_{t_0} \right] + \frac{1}{2} n(n-1) \sigma^2 \mathbb{E} \left[ X_t^{n-1} Y_t \big| X_{t_0} \right] + m \mathbb{E} \left[ X_t^{n+1} Y_{t-1} \big| X_{t_0} \right]. \]

To find the mean of \( Y_t \) given \( X_{t_0} \), we let \( n = 0 \) and \( m = 1 \) so that

\[ \frac{d}{dt} \mathbb{E} \left[ Y_t \big| X_{t_0} \right] = \mathbb{E} \left[ X_t \big| X_{t_0} \right]. \]

From Problem 3.2.2.12 (page 135), we have

\[ \mathbb{E} \left[ X_t \big| X_{t_0} \right] = X_{t_0} e^{-\kappa(t-t_0)} + \theta \left( 1 - e^{-\kappa(t-t_0)} \right) \]

and therefore,

\[ \mathbb{E} \left[ Y_t \big| X_{t_0} \right] = \int_{t_0}^t \left[ X_{t_0} e^{-\kappa(u-t_0)} + \theta \left( 1 - e^{-\kappa(u-t_0)} \right) \right] du \]

\[ = \theta(t-t_0) + \frac{1}{\kappa} \left( X_{t_0} - \theta \right) (1 - e^{-\kappa(t-t_0)}). \]

For the second moment of \( Y_t \) given \( X_{t_0} \), we set \( n = 0 \) and \( m = 2 \) so that

\[ \frac{d}{dt} \mathbb{E} \left[ Y_t^2 \big| X_{t_0} \right] = 2 \mathbb{E} \left[ X_t Y_t \big| X_{t_0} \right] \]

and to find \( \mathbb{E}[X Y_t | X_{t_0}] \), we set \( n = 1 \) and \( m = 1 \) so that

\[ \frac{d}{dt} \mathbb{E} \left[ X_t Y_t \big| X_{t_0} \right] = \kappa \theta \mathbb{E} \left[ Y_t \big| X_{t_0} \right] - \kappa \mathbb{E} \left[ X_t Y_t \big| X_{t_0} \right] + \mathbb{E} \left[ X_t^2 \big| X_{t_0} \right] \]

or

\[ \frac{d}{dt} \mathbb{E} \left[ X_t Y_t \big| X_{t_0} \right] + \kappa \mathbb{E} \left[ X_t Y_t \big| X_{t_0} \right] = \kappa \theta \mathbb{E} \left[ Y_t \big| X_{t_0} \right] + \mathbb{E} \left[ X_t^2 \big| X_{t_0} \right]. \]
By setting the integrating factor $I = e^{\xi t}$, the solution of the first-order differential equation is

$$\frac{d}{dt} \left[ e^{\xi t} \mathbb{E} \left[ X_t Y_t | X_{t_0} \right] \right] = \kappa \theta \mathbb{E} \left[ Y_t | X_{t_0} \right] + \mathbb{E} \left[ X_t^2 | X_{t_0} \right]$$

or

$$\mathbb{E} \left[ X_t Y_t | X_{t_0} \right] = e^{-\xi t} \int_{t_0}^{t} \left\{ \kappa \theta \mathbb{E} \left[ Y_u | X_{t_0} \right] + \mathbb{E} \left[ X_u^2 | X_{t_0} \right] \right\} du.$$  

Since

$$\mathbb{E} \left[ Y_t | X_{t_0} \right] = \theta (t - t_0) + \frac{1}{\kappa} \left( X_{t_0} - \theta \right) \left( 1 - e^{-\kappa (t - t_0)} \right)$$

and from Problem 3.2.2.12 (page 135),

$$\mathbb{E} \left[ X_t^2 | X_{t_0} \right] = X_{t_0}^2 e^{-2\kappa (t - t_0)} + \left( \frac{2\kappa \theta + \sigma^2}{\kappa} \right) \left( X_{t_0}^2 - \theta \right) \left( e^{-\kappa (t - t_0)} - e^{-2\kappa (t - t_0)} \right)$$

we can easily show that

$$\mathbb{E} \left[ X_t Y_t | X_{t_0} \right] = \frac{1}{2} \kappa \theta^2 (t - t_0)^2 e^{-\kappa t} + \theta (t - t_0) \left( X_{t_0} + \frac{\sigma^2}{\kappa} \right) e^{-\kappa t}$$

$$+ \left( \frac{\kappa \theta + \sigma^2}{\kappa} \right) \left( X_{t_0} - \theta \right) \left( 1 - e^{-\kappa (t - t_0)} \right) e^{-\kappa t}$$

$$+ \frac{1}{2\kappa} \left[ X_{t_0}^2 - \left( \frac{2\kappa \theta + \sigma^2}{\kappa} \right) \left( X_{t_0} - \frac{1}{2} \theta \right) \right] \left( 1 - e^{-2\kappa (t - t_0)} \right) e^{-\kappa t}.$$  

Finally, by substituting the result of $\mathbb{E}[X_t Y_t | X_{t_0}]$ into

$$\mathbb{E} \left[ Y_t^2 | X_{t_0} \right] = 2 \int_{t_0}^{t} \mathbb{E} \left[ X_u Y_u | X_{t_0} \right]$$

and using integration by parts, we eventually arrive at

$$\mathbb{E} \left[ Y_t^2 | X_{t_0} \right] = \frac{1}{\kappa^2} \left[ X_{t_0}^2 + \left( 4\theta + \frac{\sigma^2}{\kappa} \right) X_{t_0} + \theta \left( \theta - \frac{\sigma^2}{2\kappa} \right) \right] \left( 1 - e^{-\kappa (t - t_0)} \right) e^{-\kappa t_0}$$

$$- \frac{\theta}{\kappa} \left[ 2(X_{t_0} + 1) + \frac{\sigma^2}{\kappa} + t - t_0 \right] (t - t_0) e^{-\kappa t}$$

$$- \left( \frac{\kappa \theta + \sigma^2}{\kappa^3} \right) \left( X_{t_0} - \theta \right) \left( 1 - e^{-2\kappa (t - t_0)} \right) e^{-\kappa t_0}$$

$$- \frac{1}{3\kappa^2} \left[ X_{t_0}^2 - \left( \frac{2\kappa \theta + \sigma^2}{\kappa} \right) \left( X_{t_0} - \frac{1}{2} \theta \right) \right] \left( 1 - e^{-3\kappa (t - t_0)} \right) e^{-\kappa t_0}. \square$$
3.2.3 Multi-Dimensional Diffusion Process

10. **Heston Model.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{W^S_t : t \geq 0\}, \{W^\sigma_t : t \geq 0\}\) be two standard Wiener processes with correlation \(\rho \in (-1, 1)\). Suppose the asset price \(S_t\) takes the form

\[
dS_t = \mu S_t dt + \sigma_t S_t dW^S_t, \quad S_0 > 0
\]

\[
d\sigma^2_t = \kappa(\theta - \sigma^2_t) dt + \alpha \sigma_t dW^\sigma_t, \quad \sigma_0 > 0
\]

where \(\mu, \kappa, \theta\) and \(\alpha\) are constants, and \(\sigma_t\) is the stochastic volatility of \(S_t\). By defining \(\{B_t : t \geq 0\}\) as a standard Wiener process where \(B_t \perp \perp W_t^\sigma\), show that we can write

\[
W_t^S = \rho W_t^\sigma + \sqrt{1 - \rho^2} B_t.
\]

Using the above relation show also that for \(0 \leq t \leq T\) we can write

\[
\log \left( \frac{S_T}{S_t} \right) = \mu(T - t) - \frac{1}{2} \int_t^T \sigma^2_u du + \int_t^T \sigma_u dB_u
\]

where \(\xi_t = e^{\rho \int_0^t \sigma_u dW_u^\sigma - \frac{1}{2} \rho^2 \int_0^t \sigma^2_u du}\).

Prove that the relation of \(\xi_T\) and \(\xi_t\) can be expressed as

\[
\xi_T = \xi_t + \rho \int_t^T \sigma_u \xi_u dW_u^\sigma.
\]

Conditional on \(\mathcal{F}_t\) and \(\{\sigma_u : t \leq u \leq T\}\), show that

\[
\log \left( \frac{S_T}{S_t} \right) \bigg| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \sim \mathcal{N} \left( \mu - \frac{1}{2} \sigma^2_{RMS}(T - t), \sigma^2_{RMS}(T - t) \right)
\]

where \(\sigma_{RMS} = \sqrt{\frac{1 - \rho^2}{T - t} \int_t^T \sigma_u^2 du} \).

**Solution:** From \(W_t^S = \rho W_t^\sigma + \sqrt{1 - \rho^2} B_t\) we have

\[
\mathbb{E}(W_t^S) = \mathbb{E}(\rho W_t^\sigma + \sqrt{1 - \rho^2} B_t) = \rho \mathbb{E}(W_t^\sigma) + \sqrt{1 - \rho^2} \mathbb{E}(B_t) = 0
\]

and

\[
\text{Var} (W_t^S) = \text{Var} (\rho W_t^\sigma + \sqrt{1 - \rho^2} B_t) = \rho^2 \text{Var} (W_t^\sigma) + (1 - \rho^2) \text{Var} (B_t) = t.
\]

Given both \(W_t^\sigma \sim \mathcal{N}(0, t)\) and \(B_t \sim \mathcal{N}(0, t)\), therefore

\[
\rho W_t^\sigma + \sqrt{1 - \rho^2} B_t \sim \mathcal{N}(0, t).
\]
3.2.3 Multi-Dimensional Diffusion Process

In addition, using Itô’s formula and taking note that \( W_t^\sigma \perp B_t \),

\[
dW_t^S \cdot dW_t^\sigma = d(\rho W_t^\sigma + \sqrt{1 - \rho^2} B_t) \cdot dW_t^\sigma
\]

\[
= (\rho dW_t^\sigma + \sqrt{1 - \rho^2} dB_t) \cdot dW_t^\sigma
\]

\[
= \rho (dW_t^\sigma)^2 + \sqrt{1 - \rho^2} dB_t \cdot dW_t^\sigma
\]

\[
= \rho dt.
\]

Thus, we can write

\[
W_t^S = \rho W_t^\sigma + \sqrt{1 - \rho^2} B_t.
\]

Writing the SDEs of \( S_t \) and \( \sigma_t^2 \) in terms of \( W_t^\sigma \) and \( B_t \),

\[
dS_t = \mu S_t dt + \sigma_t S_t (\rho dW_t^\sigma + \sqrt{1 - \rho^2} dB_t)
\]

\[
d\sigma_t^2 = \kappa(\theta - \sigma_t^2) dt + \alpha \sigma_t dW_t^\sigma
\]

and following Itô’s lemma,

\[
d\log S_t = \frac{dS_t}{S_t} - \frac{1}{2} \left( \frac{dS_t}{S_t} \right)^2 + \ldots
\]

\[
= \mu dt + \sigma_t S_t (\rho dW_t^\sigma + \sqrt{1 - \rho^2} dB_t) - \frac{1}{2} \sigma_t^2 (\rho^2 dt + (1 - \rho^2) dt)
\]

\[
= \left( \mu - \frac{1}{2} \sigma_t^2 \right) dt + \rho \sigma_t dW_t^\sigma + \sqrt{1 - \rho^2} \sigma_t dB_t.
\]

and taking integrals,

\[
\int_t^T d \log S_u = \int_t^T \left( \mu - \frac{1}{2} \sigma_u^2 \right) du + \rho \int_t^T \sigma_u dW_u^\sigma + \sqrt{1 - \rho^2} \int_t^T \sigma_u dB_u
\]

\[
\log S_T = \log S_t + \mu(T - t) - \frac{1}{2} (1 - \rho^2) \int_t^T \sigma_u^2 du + \rho \int_t^T \sigma_u dW_u^\sigma + \sqrt{1 - \rho^2} \int_t^T \sigma_u dB_u.
\]

By setting

\[
\xi_t = e^{\int_0^t \sigma_u dW_u^\sigma - \frac{1}{2} \rho^2 \int_0^t \sigma_u^2 du}
\]

and \( \xi_T = e^{\int_0^T \sigma_u dW_u^\sigma - \frac{1}{2} \rho^2 \int_0^T \sigma_u^2 du} \),

we have

\[
\log S_T = \log S_t + \mu(T - t) - \frac{1}{2} (1 - \rho^2) \int_t^T \sigma_u^2 du + \sqrt{1 - \rho^2} \int_t^T \sigma_u dB_u + \log \xi_T - \log \xi_t
\]

or

\[
\log \left( \frac{S_T/\xi_T}{S_t/\xi_t} \right) = \mu(T - t) - \frac{1}{2} (1 - \rho^2) \int_t^T \sigma_u^2 du + \sqrt{1 - \rho^2} \int_t^T \sigma_u dB_u.
\]
To show the relation between $\xi_T$ and $\xi_t$, from Taylor’s theorem and then using Itô’s lemma,

$$d\xi_t = \frac{\partial \xi_t}{\partial t} dt + \frac{\partial \xi_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 \xi_t}{\partial t^2} (dt)^2 + \frac{1}{2} \frac{\partial^2 \xi_t}{\partial (W_t)^2} (dW_t)^2 + \ldots$$

$$= -\frac{1}{2} \rho^2 \sigma_t^2 \xi_t dt + \rho \sigma_t \xi_t dW_t + \frac{1}{2} \rho^2 \sigma_t^2 \xi_t dt$$

$$= \rho \sigma_t \xi_t dW_t.$$

Taking integrals,

$$\int_t^T d\xi_u = \int_t^T \rho \sigma_u \xi_u dW_u$$

or

$$\xi_T = \xi_t + \rho \int_t^T \sigma_u \xi_u dW_u.$$

Finally, conditional on $\mathcal{F}_t$ and $\{\sigma_u : t \leq u \leq T\}$,

$$\mathbb{E} \left[ \log \left( \frac{S_T}{S_t} / \xi_T / \xi_t \right) \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right] = \mu(T-t) - \frac{1}{2} (1 - \rho^2) \int_t^T \sigma_u^2 \ du$$

and using the properties of Itô’s integral,

$$\text{Var} \left[ \log \left( \frac{S_T}{S_t} / \xi_T / \xi_t \right) \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right]$$

$$= (1 - \rho^2) \text{Var} \left[ \int_t^T \sigma_u dW_u \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right]$$

$$= (1 - \rho^2) \mathbb{E} \left[ \left( \int_t^T \sigma_u dW_u \right)^2 \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right]$$

$$- (1 - \rho^2) \mathbb{E} \left[ \int_t^T \sigma_u dW_u \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right]^2$$

$$= (1 - \rho^2) \mathbb{E} \left[ \int_t^T \sigma_u^2 \ du \left| \mathcal{F}_t, \{\sigma_u : t \leq u \leq T\} \right. \right]$$

$$= (1 - \rho^2) \sigma_{\text{RMS}}^2 (T-t)$$
where \( \sigma_{\text{RMS}}^2 = \left( 1 - \frac{\rho^2}{T-t} \right) \int_s^T \sigma_u^2 \, du \).

Since \( \int_s^T \sigma_u \, dB_u \) is normally distributed, therefore

\[
\log \left( \frac{S_T}{S_t} \right) \bigg| \mathcal{F}_t, \{ \sigma_u : t \leq u \leq T \} \sim \mathcal{N} \left[ \left( \mu - \frac{1}{2} \sigma_{\text{RMS}}^2 \right) (T-t), \sigma_{\text{RMS}}^2(T-t) \right].
\]

N.B. Given that the stochastic volatility \( \sigma_t \) follows a CIR process, from the results of Problem 3.2.3.9 (page 171) we can easily find the mean and variance of \( \sigma_{\text{RMS}}^2 \).

\[\square\]

11. **Feynman–Kac Formula for Multi-Dimensional Diffusion Process.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( S_t = (s^{(1)}_t, s^{(2)}_t, \ldots, s^{(n)}_t) \). We consider the following PDE problem:

\[
\frac{\partial V}{\partial t}(S_t, t) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i(S^{(i)}_t, t)\sigma_j(S^{(j)}_t, t) \frac{\partial^2 V}{\partial S^{(i)}_t \partial S^{(j)}_t}(S_t, t) + \sum_{i=1}^n \mu_i(S^{(i)}_t, t) \frac{\partial V}{\partial S^{(i)}_t}(S_t, t) - r(t)V(S_t, t) = 0
\]

with boundary condition \( V(S_T, T) = \Psi(S_T) \) where \( \mu, \sigma \) are known functions of \( S^{(i)}_t \) and \( t \), \( r \) and \( \Psi \) are functions of \( t \) and \( S_T \), respectively where \( t < T \). Using Itô’s formula on the process,

\[
Z_u = e^{-\int_u^T r(v) \, dv} V(S_u, u)
\]

where \( S^{(i)}_t \) satisfies the generalised SDE

\[
dS^{(i)}_t = \mu_i(S^{(i)}_t, t) \, dt + \sigma_i(S^{(i)}_t, t) \, dW^{(i)}_t
\]

such that \( \{ W^{(i)}_t : t \geq 0 \} \) is a standard Wiener process and \( dW^{(i)}_t \cdot dW^{(j)}_t = \rho_{ij} \, dt, \rho_{ij} \in (-1, 1) \) for \( i \neq j \) and \( \rho_{ij} = 1 \) for \( i = j \) where \( i, j = 1, 2, \ldots, n \), show that under the filtration \( \mathcal{F}_t \), the solution of the PDE is given by

\[
V(S_t, t) = \mathbb{E} \left[ e^{-\int_t^T r(v) \, dv} \Psi(S_T) \bigg| \mathcal{F}_t \right].
\]

**Solution:** In analogy with Problem 3.2.2.20 (page 147), we let \( g(u) = e^{-\int_u^T r(v) \, dv} \) and set

\[
Z_u = g(u)V(S_u, u).
\]

By applying Taylor’s expansion and Itô’s formula on \( dZ_u \), we have

\[
dZ_u = \frac{\partial Z_u}{\partial u} \, du + \sum_{i=1}^n \frac{\partial Z_u}{\partial S^{(i)}_u} dS^{(i)}_u + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 Z_u}{\partial S^{(i)}_u \partial S^{(j)}_u} dS^{(i)}_u dS^{(j)}_u + \ldots
\]
3.2.3 Multi-Dimensional Diffusion Process

By integrating both sides of

\[ \frac{dg}{du} + V(S_u, u) \frac{dg}{dS_u} du + \sum_{i=1}^{n} \left( g(u) \frac{\partial V}{\partial S_u^{(i)}} \right) dS_u^{(i)} \]

\[ + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( g(u) \frac{\partial^2 V}{\partial S_u^{(i)} \partial S_u^{(j)}} \right) dS_u^{(i)} dS_u^{(j)} \]

\[ = \left( g(u) \frac{\partial V}{\partial u} - r(u)g(u)V(S_u, u) \right) du \]

\[ + \sum_{i=1}^{n} \left( g(u) \frac{\partial V}{\partial S_u^{(i)}} \right) \left( \mu(S_u^{(i)}, u) du + \sigma(S_u^{(i)}, u) dW_u^{(i)} \right) \]

\[ + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( g(u) \frac{\partial^2 V}{\partial S_u^{(i)} \partial S_u^{(j)}} \right) \left( \rho_{ij}(S_u^{(i)}, u) \sigma(S_u^{(i)}, u) \sigma(S_u^{(j)}, u) dt \right) \]

\[ = g(u) \left( \frac{\partial V}{\partial u}(S_u, u) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}(S_u^{(i)}, u) \sigma(S_u^{(i)}, u) \sigma(S_u^{(j)}, u) \frac{\partial^2 V}{\partial S_u^{(i)} \partial S_u^{(j)}} \right) \]

\[ + \sum_{i=1}^{n} \mu(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}}(S_u, u) - r(u)V(S_u, u) \]

\[ + g(u) \sum_{i=1}^{n} \sigma(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}} dW_u^{(i)} \]

\[ = g(u) \sum_{i=1}^{n} \sigma(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}} dW_u^{(i)} \]

since

\[ \frac{\partial V}{\partial u}(S_u, u) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}(S_u^{(i)}, u) \sigma(S_u^{(i)}, u) \frac{\partial^2 V}{\partial S_u^{(i)} \partial S_u^{(j)}}(S_u, u) \]

\[ + \sum_{i=1}^{n} \mu(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}}(S_u, u) - r(u)V(S_u, u) = 0. \]

By integrating both sides of \( dZ_u \) we have

\[ \int_{t}^{T} dZ_u = \sum_{i=1}^{n} \left\{ \int_{t}^{T} g(u)\sigma(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}} dW_u^{(i)} \right\} \]

\[ Z_T - Z_t = \sum_{i=1}^{n} \left\{ \int_{t}^{T} e^{-\int_{t}^{s} r(v) dv} \sigma(S_u^{(i)}, u) \frac{\partial V}{\partial S_u^{(i)}} dW_u^{(i)} \right\}. \]

Taking expectations and using the property of Itô calculus,

\[ \mathbb{E}(Z_T - Z_t) = 0 \quad \text{or} \quad \mathbb{E}(Z_t) = \mathbb{E}(Z_T). \]
Therefore, under the filtration $\mathcal{F}_t$, \[
\mathbb{E}(Z_t | \mathcal{F}_t) = \mathbb{E}(Z_T | \mathcal{F}_t) \\
\mathbb{E} \left[ e^{-\int_t^T r(s) ds} V(S_t, t) | \mathcal{F}_t \right] = \mathbb{E} \left[ e^{-\int_t^T r(s) ds} V(S_T, T) | \mathcal{F}_t \right] \\
V(S_t, t) = \mathbb{E} \left[ e^{-\int_t^T r(s) ds} \Psi(S_T) | \mathcal{F}_t \right]. \]

\[\square\]

12. **Backward Kolmogorov Equation for a Two-Dimensional Random Walk.** We consider a two-dimensional symmetric random walk where at initial time $t_0$, a particle starts at $(x_0, y_0)$ and is at position $(x, y)$ at time $t$. At time $t + \delta t$, the particle can either move to $(x + \delta x, y)$, $(x - \delta x, y)$, $(x, y + \delta y)$, or $(x, y - \delta y)$ each with probability $\frac{1}{4}$. Let $p(x, y; t, x_0, y_0, t_0)$ denote the probability density of the particle position $(x, y)$ at time $t$ starting at $(x_0, y_0)$ at time $t_0$. By writing the backward equation in a discrete fashion and expanding it using Taylor’s series, show that for $\delta x = \delta y = \sqrt{\delta t}$ and in the limit $\delta t \to 0$, \[
\frac{\partial p(x, y; t, x_0, y_0, t_0)}{\partial t} = -\frac{1}{4} \left( \frac{\partial^2 p(x, y; t, x_0, y_0, t_0)}{\partial x^2} + \frac{\partial^2 p(x, y; t, x_0, y_0, t_0)}{\partial y^2} \right). 
\]

**Solution:** By denoting $p(x, y; t, x_0, y_0, t_0)$ as the probability density function of the particle position $(x, y)$ at time $t$, the discrete model of the backward equation is \[
p(x, y; t, x_0, y_0, t_0) = \frac{1}{4} p(x - \delta x, y, t + \delta t; x_0, y_0, t_0) + \frac{1}{4} p(x + \delta x, y, t + \delta t; x_0, y_0, t_0) \\
+ \frac{1}{4} p(x, y - \delta y, t + \delta t; x_0, y_0, t_0) + \frac{1}{4} p(x, y + \delta y, t + \delta t; x_0, y_0, t_0). \]

Expanding $p(x - \delta x, y, t + \delta t; x_0, y_0, t_0)$, $p(x + \delta x, y, t + \delta t; x_0, y_0, t_0)$, $p(x, y - \delta y, t + \delta t; x_0, y_0, t_0)$ and $p(x, y + \delta y, t + \delta t; x_0, y_0, t_0)$ using Taylor’s series, we have \[
p(x - \delta x, y, t + \delta t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) - \frac{\partial p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial x} \delta x \\
+ \frac{1}{2} \frac{\partial^2 p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial x^2} (-\delta x)^2 + O((\delta x)^3) \\
p(x + \delta x, y, t + \delta t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) + \frac{\partial p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial x} \delta x \\
+ \frac{1}{2} \frac{\partial^2 p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial x^2} (\delta x)^2 + O((\delta x)^3) \\
p(x, y - \delta y, t + \delta t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) - \frac{\partial p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial y} \delta y \\
+ \frac{1}{2} \frac{\partial^2 p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial y^2} (-\delta y)^2 + O((\delta y)^3) \\
p(x, y + \delta y, t + \delta t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) + \frac{\partial p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial y} \delta y \\
+ \frac{1}{2} \frac{\partial^2 p(x, y; t + \delta t; x_0, y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta y)^3). \]
3.2.3 Multi-Dimensional Diffusion Process

and

\[
p(x, y + \delta y, t + \delta t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) + \frac{\partial p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial y} \delta y \\
+ \frac{1}{2} \frac{\partial^2 p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta y)^3).
\]

By substituting the above equations into the discrete backward equation,

\[
p(x, y, t; x_0, y_0, t_0) = p(x, y, t + \delta t; x_0, y_0, t_0) + \frac{1}{4} \frac{\partial^2 p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial x^2} (\delta x)^2 \\
+ \frac{1}{4} \frac{\partial^2 p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial y^2} (\delta y)^2 + O((\delta x)^3) + O((\delta y)^3).
\]

Setting \( \delta x = \delta y = \sqrt{\delta t} \) and dividing the equation by \( \delta t \) and in the limit \( \delta t \to 0 \)

\[
\lim_{\delta t \to 0} \left[ p(x, y, t + \delta t; x_0, y_0, t_0) - p(x, y, t; x_0, y_0, t_0) \right] = - \lim_{\delta t \to 0} \frac{1}{4} \left( \frac{\partial^2 p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial x^2} \right) \\
- \lim_{\delta t \to 0} \frac{1}{4} \left( \frac{\partial^2 p(x, y, t + \delta t; x_0, y_0, t_0)}{\partial y^2} \right) \\
+ \lim_{\delta t \to 0} O(\sqrt{\delta t})
\]

we eventually arrive at

\[
\frac{\partial p(x, y, t; x_0, y_0, t_0)}{\partial t} = - \frac{1}{4} \left( \frac{\partial^2 p(x, y, t; x_0, y_0, t_0)}{\partial x^2} + \frac{\partial^2 p(x, y, t; x_0, y_0, t_0)}{\partial y^2} \right).
\]

13. **Forward Kolmogorov Equation for a Two-Dimensional Random Walk.** We consider a two-dimensional symmetric random walk where at initial time \( t_0 \), a particle starts at \( (X_0, Y_0) \) and is at position \( (X, Y) \) at terminal time \( T > 0 \). At time \( T - \delta T \), the particle can either move to \( (X + \delta X, Y), (X - \delta X, Y), (X, Y + \delta Y) \) or \( (X, Y - \delta Y) \) each with probability \( \frac{1}{4} \). Let \( p(X, Y, T; X_0, Y_0, t_0) \) denote the probability density of the position \( (X, Y) \) at time \( T \) starting at \( (X_0, Y_0) \) at time \( t_0 \).

By writing the forward equation in a discrete fashion and expanding it using Taylor’s series, show that for \( \delta X = \delta Y = \sqrt{\delta T} \) and in the limit \( \delta T \to 0 \)

\[
\frac{\partial p(X, Y, T; X_0, Y_0, t_0)}{\partial T} = \frac{1}{4} \left( \frac{\partial^2 p(X, Y, T; X_0, Y_0, t_0)}{\partial X^2} + \frac{\partial^2 p(X, Y, T; X_0, Y_0, t_0)}{\partial Y^2} \right).
\]
3.2.3 Multi-Dimensional Diffusion Process

Solution: By denoting \( p(X, Y, T; X_0, Y_0, t_0) \) as the probability density function of the particle position \((X, Y)\) at time \(T\) starting at \((X_0, Y_0)\) at initial time \(t_0\), the discrete model of the forward equation is

\[
p(X, Y, T; X_0, Y_0, t_0) = \frac{1}{4} p(X - \delta X, Y, T - \delta T; X_0, Y_0, t_0) + \frac{1}{4} p(X + \delta X, Y, T - \delta T; X_0, Y_0, t_0) + \frac{1}{4} p(X, Y - \delta Y, T - \delta T; X_0, Y_0, t_0) + \frac{1}{4} p(X, Y + \delta Y, T - \delta T; X_0, Y_0, t_0).
\]

Expanding \( p(X - \delta X, Y, T - \delta T; X_0, Y_0, t_0) \), \( p(X + \delta X, Y, T - \delta T; X_0, Y_0, t_0) \), \( p(X, Y - \delta Y, T - \delta T; X_0, Y_0, t_0) \) and \( p(X, Y + \delta Y, T - \delta T; X_0, Y_0, t_0) \) using Taylor’s series, we have

\[
p(X - \delta X, Y, T - \delta T; X_0, Y_0, t_0) = p(X, Y, T - \delta T; X_0, Y_0, t_0) - \frac{\partial p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X} \delta X + \frac{1}{2} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X^2} (\delta X)^2 + O((\delta X)^3)
\]

\[
p(X + \delta X, Y, T - \delta T; X_0, Y_0, t_0) = p(X, Y, T - \delta T; X_0, Y_0, t_0) + \frac{\partial p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X} \delta X + \frac{1}{2} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X^2} (\delta X)^2 + O((\delta X)^3)
\]

\[
p(X, Y - \delta Y, T - \delta T; X_0, Y_0, t_0) = p(X, Y, T - \delta T; X_0, Y_0, t_0) - \frac{\partial p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y} \delta Y + \frac{1}{2} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y^2} (\delta Y)^2 + O((\delta Y)^3)
\]

and

\[
p(X, Y + \delta Y, T - \delta T; X_0, Y_0, t_0) = p(X, Y, T - \delta T; X_0, Y_0, t_0) + \frac{\partial p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y} \delta Y + \frac{1}{2} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y^2} (\delta Y)^2 + O((\delta Y)^3).
\]
3.2.3 Multi-Dimensional Diffusion Process

By substituting the above equations into the discrete forward equation,

\[ p(X, Y, T; X_0, Y_0, t_0) = p(X, Y, T - \delta T; X_0, Y_0, t_0) + \frac{1}{4} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X^2} (\delta X)^2 + \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y^2} (\delta Y)^2 + O((\delta X)^3) + O((\delta Y)^3). \]

Setting \( \delta X = \delta Y = \sqrt{\delta T} \), dividing the equation by \( \delta T \) and in the limit \( \delta T \to 0 \)

\[ \lim_{\delta T \to 0} \left[ \frac{p(X, Y, T; X_0, Y_0, t_0) - p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\delta T} \right] \]

\[ = \lim_{\delta T \to 0} \frac{1}{4} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial X^2} \]

\[ + \lim_{\delta T \to 0} \frac{1}{4} \frac{\partial^2 p(X, Y, T - \delta T; X_0, Y_0, t_0)}{\partial Y^2} \]

\[ + \lim_{\delta T \to 0} O(\sqrt{\delta T}) \]

we eventually arrive at

\[ \frac{\partial p(X, Y, T; X_0, Y_0, t_0)}{\partial T} = \frac{1}{4} \left( \frac{\partial^2 p(X, Y, T; X_0, Y_0, t_0)}{\partial X^2} + \frac{\partial^2 p(X, Y, T; X_0, Y_0, t_0)}{\partial Y^2} \right). \]
Change of Measure

In finance, derivative instruments such as options, swaps or futures can be used for both hedging and speculation purposes. In a hedging scenario, traders can reduce their risk exposure by buying and selling derivatives against fluctuations in the movement of underlying risky asset prices such as stocks and commodities. Conversely, in a speculation scenario, traders can also use derivatives to profit in the future direction of underlying prices. For example, if a trader expects an asset price to rise in the future, then he/she can sell put options (i.e., the purchaser of the put options pays an initial premium to the seller and has the right but not the obligation to sell the shares back to the seller at an agreed price should the share price drop below it at the option expiry date). Given the purchaser of the put option is unlikely to exercise the option, the seller would be most likely to profit from the premium paid by the purchaser. From the point of view of trading such contracts, we would like to price contingent claims (or payoffs of derivative securities such as options) in such a way that there is no arbitrage opportunity (or no risk-free profits). By doing so we will ensure that even though two traders may differ in their estimate of the stock price direction, yet they will still agree on the price of the derivative security. In order to accomplish this we can rely on Girsanov’s theorem, which tells us how a stochastic process can have a drift change (but not volatility) under a change of measure. With the application of this important result to finance we can convert the underlying stock prices under the physical measure (or real-world measure) into the risk-neutral measure (or equivalent martingale measure) where all the current stock prices are equal to their expected future prices discounted at the risk-free rate. This is in contrast to using the physical measure, where the derivative security prices will vary greatly since the underlying assets will differ in degrees of risk from each other.

4.1 INTRODUCTION

From the seminal work of Black, Scholes and Merton in using diffusion processes and martingales to price contingent claims such as options on risky asset prices, the theory of mathematical finance is one of the most successful applications of probability theory. Fundamentally, the Black–Scholes model is concerned with an economy consisting of two assets, a risky asset (stock) whose price $S_t$, $t \geq 0$ is a stochastic process and a risk-free asset (bond or money market account) whose value $B_t$ grows at a continuously compounded interest rate. Here we can assume that $S_t$ and $B_t$ satisfy the following equations

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t, \quad \frac{dB_t}{B_t} = r_t dt$$

where $\mu_t$ is the stock price growth rate, $\sigma_t$ is the stock price volatility, $r_t$ is the risk-free rate and $W_t$ is a standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the financial market we would like to price a contingent claim $\Psi(S_T)$ at time $t \leq T$ in such a way that no risk-free profit or arbitrage opportunity exists. But before we define the notion of arbitrage we need some
financial terminologies. The following first two definitions refer to the concepts of trading strategy and self-financing trading strategy, which form the basis of creating a martingale framework to price a contingent claim.

**Definition 4.1(a) (Trading Strategy)** In a continuous time setting, at time $t \in [0, T]$ we consider an economy which consists of a non-dividend-paying risky asset $S_t$ and a risk-free asset $B_t$. A trading strategy (or portfolio) is a pair $(\phi_t, \psi_t)$ of stochastic processes which are adapted to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$ holding $\phi_t$ shares of $S_t$ and $\psi_t$ units invested in $B_t$. Therefore, at time $t$ the value of the portfolio, $\Pi_t$, is

$$\Pi_t = \phi_t S_t + \psi_t B_t.$$

**Definition 4.1(b) (Self-Financing Trading Strategy)** At time $t \in [0, T]$, the trading strategy $(\phi_t, \psi_t)$ of holding $\phi_t$ shares of non-dividend-paying risky asset $S_t$ and $\psi_t$ units in risk-free asset $B_t$ having a portfolio value

$$\Pi_t = \phi_t S_t + \psi_t B_t,$$

is called self-financing (or self-financing portfolio) if

$$d\Pi_t = \phi_t dS_t + \psi_t dB_t$$

which implies the change in the portfolio value is due to changes in the market conditions and not to either infusion or extraction of funds.

Note that the change in the portfolio value is only attributed to $S_t$ and $B_t$, rather than $\phi_t$ and $\psi_t$. However, given that we cannot deposit or withdraw cash in the portfolio there will be restrictions imposed on the pair $(\phi_t, \psi_t)$. In order to keep the trading strategy self-financing we can see that if $\phi_t$ is increased then $\psi_t$ will be decreased, and vice versa. Thus, when we make a choice of either $\phi_t$ or $\psi_t$, the other can easily be found.

**Definition 4.1(c) (Admissible Trading Strategy)** At time $t \in [0, T]$, the trading strategy $(\phi_t, \psi_t)$ of holding $\phi_t$ shares of non-dividend-paying risky asset $S_t$ and $\psi_t$ units in risk-free asset $B_t$ having a portfolio value

$$\Pi_t = \phi_t S_t + \psi_t B_t,$$

is admissible if it is self-financing and if $\Pi_t \geq -\alpha$ almost surely for some $\alpha > 0$ and $t \in [0, T]$.

From the above definition, the existence of a negative portfolio value which is bounded from below almost surely shows that the investor cannot go too far into debt and thus have a finite credit line. Once we have restricted ourselves to a class of admissible strategies, the absence of arbitrage opportunities (“no free lunch”) can be ensured when pricing contingent claims. But before we discuss its precise terminology we need to define first the concept of attainable contingent claim.

**Definition 4.1(d) (Attainable Contingent Claim)** Consider a trading strategy $(\phi_t, \psi_t)$ at time $t \in [0, T]$ of holding $\phi_t$ shares of non-dividend-paying risky asset $S_t$ and $\psi_t$ units in risk-free asset $B_t$ having a portfolio value

$$\Pi_t = \phi_t S_t + \psi_t B_t.$$
The contingent claim $\Psi(S_T)$ is attainable if there exists an admissible strategy worth $\Pi_T = \Psi(S_T)$ at exercise time $T$.

**Definition 4.1(e) (Arbitrage)** An arbitrage opportunity is an admissible strategy if the following criteria are satisfied:

(i) $\Pi_0 = 0$ (no net investment initially)
(ii) $\mathbb{P}(\Pi_T \geq 0) = 1$ (always win at time $T$)
(iii) $\mathbb{P}(\Pi_T > 0) > 0$ (making a positive return on investment at time $T$).

For risky assets which are traded in financial markets, the pricing of contingent claims based on the underlying assets is determined based on the presence/absence of arbitrageable opportunities as well as whether the market is complete/incomplete. The following definition discusses the concept of a complete market.

**Definition 4.1(f) (Complete Market)** A market is said to be complete if every contingent claim is attainable (i.e., can be replicated by a self-financing trading strategy). Otherwise, it is incomplete.

In general, when we have $N_A$ number of traded assets (excluding risk-free assets) and $N_R$ sources of risk, the financial “rule of thumb” is as follows:

- If $N_A < N_R$ then the market has no arbitrage and is incomplete.
- If $N_A = N_R$ then the market has no arbitrage and is complete.
- If $N_A > N_R$ then the market has arbitrage.

Within the framework of a Black–Scholes model, the following theorem states the existence of a trading strategy pair $(\phi_t, \psi_t)$ in a portfolio.

**Theorem 4.2 (One-Dimensional Martingale Representation Theorem)** Let $\{W_t : 0 \leq t \leq T\}$ be a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_t, 0 \leq t \leq T$ be the filtration generated by $W_t$. If $M_t, 0 \leq t \leq T$ is a martingale with respect to this filtration then there exists an adapted process $\gamma_t, 0 \leq t \leq T$ such that

$$M_t = M_0 + \int_0^t \gamma_u \, dW_u, \quad 0 \leq t \leq T.$$

From the martingale representation theorem it follows that martingales can be represented as Itô integrals. However, the theorem only states that an adapted process $\gamma_t$ exists but does not provide a method to find it explicitly. Owing to the complexity of the proof which involves functional analysis, the details are omitted in this book.

To show the relationship between the martingale representation theorem and the trading strategy pair $(\psi_t, \phi_t)$, if we set $B_0 = 1$ so that $B_t = e^{\int_0^t r_u \, du}$ then the discounted portfolio value can be written as

$$\tilde{\Pi}_t = B_t^{-1} \Pi_t.$$
Therefore, the trading strategy pair \((\phi_t, \psi_t)\) is self-financing if and only if the discounted portfolio value can be written as a stochastic integral

\[
\tilde{\Pi}_t = \tilde{\Pi}_0 + \int_0^t \phi_v \, d(B_t^{-1}S_v).
\]

By assuming that \((\phi_t, \psi_t)\) is a self-financing trading strategy which replicates the contingent claim \(\Psi(S_T)\), it is hoped that the discounted risky asset value \(B_t^{-1}S_t\) will be a martingale (or \(\mathbb{P}\)-martingale), so that by taking expectations under the \(\mathbb{P}\) measure

\[
\mathbb{E}^\mathbb{P}\left[\tilde{\Pi}_T \mid \mathcal{F}_T\right] = \mathbb{E}^\mathbb{P}\left[\tilde{\Pi}_0 \mid \mathcal{F}_T\right] + \mathbb{E}^\mathbb{P}\left[\int_0^T \phi_v \, d(B_v^{-1}S_v) \mid \mathcal{F}_T\right] = \tilde{\Pi}_0 + \int_0^T \phi_v \, d(B_v^{-1}S_v) = \tilde{\Pi}_T
\]

or

\[
\Pi_T = \mathbb{E}^\mathbb{P}\left[e^{-\int_0^T r_u \, du} \, \tilde{\Pi}_T \mid \mathcal{F}_T\right] = \mathbb{E}^\mathbb{P}\left[e^{-\int_0^T r_u \, du} \, \Psi(S_T) \mid \mathcal{F}_T\right]
\]

since \(\Pi_T = \Psi(S_T)\). Although \(B_t^{-1}S_t\) is not a \(\mathbb{P}\)-martingale on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), there exists another equivalent martingale measure \(Q\) (or risk-neutral measure) such that \(B_t^{-1}S_t\) is a \(Q\)-martingale on the probability space \((\Omega, \mathcal{F}, Q)\). Before we state the theorem we first present a few intermediate results which will lead to the main result.

**Definition 4.3** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions. Let \(Q\) be another probability measure on \((\Omega, \mathcal{F}, Q)\). Assume that for every \(A \in \mathcal{F}\) satisfying \(\mathbb{P}(A) = 0\), we also have \(Q(A) = 0\), then we say \(Q\) is absolutely continuous with respect to \(\mathbb{P}\) on \(\mathcal{F}\) and we write it as \(Q \ll \mathbb{P}\).

**Theorem 4.4 (Radon–Nikodým Theorem)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions. Let \(Q\) be another probability measure on \((\Omega, \mathcal{F}, Q)\). Under the assumption that \(Q \ll \mathbb{P}\), there exists a non-negative random variable \(Z\) such that

\[
\frac{dQ}{d\mathbb{P}} = Z
\]

and we call \(Z\) the Radon–Nikodým derivative of \(Q\) with respect to \(\mathbb{P}\).

Take note that in finance we need a stronger statement – that is, \(Q\) to be equivalent to \(\mathbb{P}\), \(Q \sim \mathbb{P}\) which is \(Q \ll \mathbb{P}\) and \(\mathbb{P} \ll Q\). By imposing the condition \(Q\) to be equivalent to \(\mathbb{P}\), if an event cannot occur under the \(\mathbb{P}\) measure then it also cannot occur under the \(Q\) measure and vice versa. By doing so, we can now state the following definition of the equivalent martingale measure for asset prices which pay no dividends.

**Definition 4.5 (Equivalent Martingale Measure)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions and let \(Q\) be another probability measure on \((\Omega, \mathcal{F}, Q)\). The probability measure \(Q\) is said to be an equivalent martingale measure (or risk-neutral measure) if it satisfies:

- \(Q\) is equivalent to \(\mathbb{P}\), \(Q \sim \mathbb{P}\);
4.1 INTRODUCTION

- the discounted price processes \( \{ B_t^{-1} S_t^{(i)} \} \), \( i = 1, 2, \ldots, m \) are martingales under \( \mathbb{Q} \), that is

\[
\mathbb{E}^{\mathbb{Q}} \left[ B_u^{-1} S_t^{(i)} \mid \mathcal{F}_t \right] = B_t^{-1} S_t^{(i)}
\]

for all \( 0 \leq t \leq u \leq T \).

Note that for the case of dividend-paying assets, the discounted values of the asset prices with the dividends reinvested are \( \mathbb{Q} \)-martingales.

Once the equivalent martingale measure is defined, the next question is can we transform a diffusion process into a martingale by changing the probability measure? The answer is yes where the transformation can be established using Girsanov’s theorem, which is instrumental in risk-neutral pricing for derivatives.

**Theorem 4.6 (One-Dimensional Girsanov Theorem)** Let \( \{ W_t : 0 \leq t \leq T \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and let \( \mathcal{F}_t, 0 \leq t \leq T \) be the associated Wiener process filtration. Suppose \( \theta_t \) is an adapted process, \( 0 \leq t \leq T \) and consider

\[
Z_t = e^{-\int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 ds}.
\]

If

\[
\mathbb{E}^{\mathbb{P}} \left( \int_0^T \theta_t^2 dt \right) < \infty
\]

then \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \). By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that

\[
\mathbb{E}^{\mathbb{P}} \left( d\mathbb{Q} \mid \mathcal{F}_t \right) = d\mathbb{Q} \mid \mathcal{F}_t = Z_t,
\]

then

\[
\tilde{W}_t = W_t + \int_0^t \theta_u \ du
\]

is a \( \mathbb{Q} \)-standard Wiener process.

In probability theory, the importance of Girsanov’s theorem cannot be understated as it provides a formal concept of how stochastic processes change under changes in measure. The theorem is especially important in the theory of financial mathematics as it tells us how to convert from the physical measure \( \mathbb{P} \) to the risk-neutral measure \( \mathbb{Q} \). In short, from the application of this theorem we can change from a Wiener process with drift to a standard Wiener process.

In contrast, the converse of Girsanov’s theorem says that every equivalent measure is given by a change in drift. Thus, by changing the measure it is equivalent to changing the drift and hence in the Black–Scholes model there is only one equivalent risk-neutral measure. Otherwise we would have multiple arbitrage-free derivative prices.

**Corollary 4.7** If \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and \( \mathbb{Q} \) is equivalent to \( \mathbb{P} \) then there exists an adapted process \( \theta_t, 0 \leq t \leq T \) such that:
4.1 INTRODUCTION

(i) \( \tilde{W}_t = W_t + \int_0^t \theta_u \, du \) is a \( Q \)-standard Wiener process,

(ii) the Radon–Nikodym derivative of \( Q \) with respect to \( P \) is

\[
\mathbb{E}^P \left( \frac{dQ}{dP} \bigg| \mathcal{F}_t \right) = \frac{dQ}{dP} \bigg|_{\mathcal{F}_t} = Z_t = e^{-\int_0^t \theta_u \, dW_u - \frac{1}{2} \int_0^t \theta_u^2 \, du}
\]

for \( 0 \leq t \leq T \).

By comparing the martingale representation theorem with Girsanov’s theorem we can see that in the former the filtration generated by the standard Wiener process is more restrictive than the assumption given in Girsanov’s theorem. By including this extra restriction on the filtration in Girsanov’s theorem we have the following corollary.

**Corollary 4.8** Let \( \{W_t : 0 \leq t \leq T\} \) be a \( P \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, P)\) and let \( \mathcal{F}_t, 0 \leq t \leq T \) be the filtration generated by \( W_t \). By assuming the one-dimensional Girsanov theorem holds and if \( \tilde{M}_t, 0 \leq t \leq T \) is a \( Q \)-martingale, there exists an adapted process \( \{\tilde{\gamma}_t : 0 \leq t \leq T\} \) such that

\[
\tilde{M}_t = \tilde{M}_0 + \int_0^t \tilde{\gamma}_u \, d\tilde{W}_u, \quad 0 \leq t \leq T
\]

where \( \tilde{W}_t \) is a \( Q \)-standard Wiener process.

By extending the one-dimensional Girsanov theorem to multiple risky assets where each of the assets is a random component driven by an independent standard Wiener process, we state the following multi-dimensional Girsanov theorem.

**Theorem 4.9 (Multi-Dimensional Girsanov Theorem)** Let \( W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})^T \) be an \( n \)-dimensional \( P \)-standard Wiener process, with \( \{W_t^{(i)}\}_{0 \leq t \leq T}, i = 1, 2, \ldots, n \) being an independent one-dimensional \( P \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, P)\) and let \( \mathcal{F}_t, 0 \leq t \leq T \) be the associated Wiener process filtration. Suppose we have an \( n \)-dimensional adapted process \( \theta_t = (\theta_t^{(1)}, \theta_t^{(2)}, \ldots, \theta_t^{(n)})^T, 0 \leq t \leq T \) and we consider

\[
Z_t = \exp \left\{ -\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|_2^2 \, du \right\}.
\]

If

\[
\mathbb{E}^P \left( \exp \left\{ \frac{1}{2} \int_0^T \|\theta_t\|_2^2 \, dt \right\} \right) < \infty
\]

where

\[
\|\theta_t\|_2 = \sqrt{(\theta_t^{(1)})^2 + (\theta_t^{(2)})^2 + \ldots + (\theta_t^{(n)})^2}
\]

then \( Z_t \) is a positive \( P \)-martingale for \( 0 \leq t \leq T \). By changing the measure \( P \) to a measure \( Q \) such that

\[
\mathbb{E}^P \left( \frac{dQ}{dP} \bigg| \mathcal{F}_t \right) = \frac{dQ}{dP} \bigg|_{\mathcal{F}_t} = Z_t
\]
4.1 INTRODUCTION

For all \(0 \leq t \leq T\) then

\[
\tilde{W}_t = W_t + \int_0^t \theta_u \, du
\]

is an \(n\)-dimensional \(\mathbb{Q}\)-standard Wiener process where \(\tilde{W}_t = (\tilde{W}_t^{(1)}, \tilde{W}_t^{(2)}, \ldots, \tilde{W}_t^{(n)})^T\) and \(\tilde{W}_t^{(i)} = W_t^{(i)} + \int_0^t \theta_u^{(i)} \, du, \ i = 1, 2, \ldots, n\) such that the component processes of \(\tilde{W}_t\) are independent under \(\mathbb{Q}\).

By amalgamating the results of the multi-dimensional Girsanov theorem we also state the following multi-dimensional martingale representation theorem.

**Theorem 4.10 (Multi-Dimensional Martingale Representation Theorem)** Let \(W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})^T\) be an \(n\)-dimensional \(\mathbb{P}\)-standard Wiener process, with \(\{W_t^{(i)}\}_{0 \leq t \leq T}\), \(i = 1, 2, \ldots, n\) being an independent one-dimensional \(\mathbb{P}\)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and let \(\mathcal{F}_t, 0 \leq t \leq T\) be the filtration generated by \(W_t\). If \(M_t, 0 \leq t \leq T\) is a martingale with respect to this filtration, then there exists an \(n\)-dimensional adapted process \(\gamma_t = (\gamma_t^{(1)}, \gamma_t^{(2)}, \ldots, \gamma_t^{(n)})^T, 0 \leq t \leq T\) such that

\[
M_t = M_0 + \int_0^t \gamma_u \cdot dW_u, \ \ 0 \leq t \leq T.
\]

By assuming the multi-dimensional Girsanov theorem holds and if \(\tilde{M}_t, 0 \leq t \leq T\) is a \(\mathbb{Q}\)-martingale there exists an \(n\)-dimensional adapted process \(\tilde{\gamma}_t = (\tilde{\gamma}_t^{(1)}, \tilde{\gamma}_t^{(2)}, \ldots, \tilde{\gamma}_t^{(n)})^T, 0 \leq t \leq T\) such that

\[
\tilde{M}_t = \tilde{M}_0 + \int_0^t \tilde{\gamma}_u \cdot d\tilde{W}_u, \ \ 0 \leq t \leq T
\]

where \(\tilde{W}_t = (\tilde{W}_t^{(1)}, \tilde{W}_t^{(2)}, \ldots, \tilde{W}_t^{(n)})^T\) is an \(n\)-dimensional \(\mathbb{Q}\)-standard Wiener process.

In our earlier discussion we used the risk-free asset \(B_t\) to construct our trading strategy where, under the risk-neutral measure \(\mathbb{Q}\) the discounted risky asset price \(B_t^{-1} S_t\) is a \(\mathbb{Q}\)-martingale, that is

\[
\mathbb{E}^\mathbb{Q}[B_T^{-1} S_T | \mathcal{F}_t] = B_t^{-1} S_t.
\]

In finance terminology the risk-free asset \(B_t\) which does the discounting is called the numéraire.

**Definition 4.11** A numéraire is an asset with positive price process which pays no dividends.

Suppose there is another non-dividend-paying asset \(N_t\) with strictly positive price process and under the risk-neutral measure \(\mathbb{Q}\), the discounted price \(B_t^{-1} N_t\) is a \(\mathbb{Q}\)-martingale. From Girsanov’s theorem we can define a new probability measure \(\mathbb{Q}^N\) given by the Radon–Nikodym derivative

\[
\frac{d\mathbb{Q}^N}{d\mathbb{Q}}|_{\mathcal{F}_t} = \frac{N_t}{N_0} \frac{B_t}{B_0}.
\]

By changing the \(\mathbb{Q}\) measure to an equivalent \(\mathbb{Q}^N\) measure, the discounted price \(N_t^{-1} S_t\) is a \(\mathbb{Q}^N\)-martingale such that

\[
\mathbb{E}^{\mathbb{Q}^N}[N_T^{-1} S_T | \mathcal{F}_t] = N_t^{-1} S_t.
\]
Thus, we can deduce that

\[ B_t \mathbb{E}^Q \left[ \frac{S_T}{B_T} \left| \mathcal{F}_t \right. \right] = N_t \mathbb{E}^Q \left[ \frac{S_T}{N_T} \left| \mathcal{F}_t \right. \right] \]

and by returning to the discussion of the self-financing trading strategy, for a contingent claim \( \Psi(S_T) \) we will also eventually obtain

\[ B_t \mathbb{E}^Q \left[ \frac{\Psi(S_T)}{B_T} \left| \mathcal{F}_t \right. \right] = N_t \mathbb{E}^Q \left[ \frac{\Psi(S_T)}{N_T} \left| \mathcal{F}_t \right. \right] . \]

The idea discussed above is known as the change of numéraire technique and is often used to price complicated derivatives such as convertible bonds, foreign currency and interest rate derivatives.

### 4.2 PROBLEMS AND SOLUTIONS

#### 4.2.1 Martingale Representation Theorem

1. Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process and let \( X \) be a real-valued random variable on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) such that \( \mathbb{E}^\mathbb{P}(|X|^2) < \infty \). For the process

\[ M_t = \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t) \]

show that \( \mathbb{E}^\mathbb{P}(M_t^2) < \infty \) and that \( M_t \) is a \( \mathbb{P} \)-martingale with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T \) generated by \( W_t \).

**Solution:** By definition,

\[ \mathbb{E}^\mathbb{P}(M_t^2) = \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t)^2 \right] . \]

From Jensen’s inequality (see Problem 1.2.3.14, page 48) we set \( \varphi(x) = x^2 \), which is a convex function. By substituting \( x = \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t) \) we have

\[ \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t)^2 \leq \mathbb{E}^\mathbb{P}(X^2|\mathcal{F}_t) . \]

Taking the expectation,

\[ \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t)^2 \right] \leq \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X^2|\mathcal{F}_t) \right] \]

and from the tower property (see Problem 1.2.3.11, page 46)

\[ \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X^2|\mathcal{F}_t) \right] = \mathbb{E}^\mathbb{P}(X^2) . \]

Thus,

\[ \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t)^2 \right] \leq \mathbb{E}^\mathbb{P}(X^2) \]

and because \( \mathbb{E}^\mathbb{P}(|X|^2) < \infty \), so \( \mathbb{E}^\mathbb{P}(M_t^2) = \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X|\mathcal{F}_t)^2 \right] < \infty . \)
4.2.1 Martingale Representation Theorem

To show that $M_t$ is a $\mathbb{P}$-martingale we note that:

(a) Under the filtration $\mathcal{F}_s$, $0 \leq s \leq t$ and using the tower property (see Problem 1.2.3.11, page 46),

$$
\mathbb{E}^\mathbb{P}(M_t | \mathcal{F}_s) = \mathbb{E}^\mathbb{P}\left[ \mathbb{E}^\mathbb{P}(X | \mathcal{F}_s) \right] = \mathbb{E}^\mathbb{P}(X | \mathcal{F}_s) = M_s.
$$

(b) Since $\mathbb{E}^\mathbb{P}(|X|^2) < \infty$ we can deduce, using Hölder’s inequality, that

$$
\mathbb{E}^\mathbb{P}(|M_t|) = \mathbb{E}^\mathbb{P}\left[\mathbb{E}^\mathbb{P}(|X| | \mathcal{F}_s)\right] \leq \mathbb{E}^\mathbb{P}\left[\mathbb{E}^\mathbb{P}(|X| | \mathcal{F}_s)\right] \leq \sqrt{\mathbb{E}^\mathbb{P}(|X|^2)} < \infty.
$$

(c) $M_t$ is clearly $\mathcal{F}_t$-adapted for $0 \leq t \leq T$.

From the results of (a)–(c) we have shown that $M_t$ is a $\mathbb{P}$-martingale with respect to $\mathcal{F}_t$, $0 \leq t \leq T$.

\[\square\]

2. Let $\{W_t : t \geq 0\}$ be a $\mathbb{P}$-standard Wiener process and let $X$ be a real-valued random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}^\mathbb{P}(|X|^2) < \infty$. Given that the process $M_t = \mathbb{E}^\mathbb{P}(X | \mathcal{F}_t)$ is a $\mathbb{P}$-martingale with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$ generated by $W_t$, then from the martingale representation theorem there exists an adapted process $\gamma_u, 0 \leq u \leq T$ such that

$$
M_t = M_0 + \int_0^t \gamma_u \, dW_u, \quad 0 \leq t \leq T.
$$

Show that

(a) if $X = W_T$ then $\gamma_t = 1$,

(b) if $X = W_T^2$ then $\gamma_t = 2W_t$,

(c) if $X = W_T^3$ then $\gamma_t = 3(W_T^2 + T - t)$,

(d) if $X = e^{\sigma W_T}$, $\sigma \in \mathbb{R}$ then $\gamma_t = \sigma e^{\sigma W_T + \frac{1}{2} \sigma^2 (T-t)}$, for $0 \leq t \leq T$.

**Solution:** To find the adapted process $\gamma_t$ we note that from Itô’s formula,

$$
dM_t = \frac{\partial M_t}{\partial t} \, dt + \frac{\partial M_t}{\partial W_t} \, dW_t + \frac{1}{2} \frac{\partial^2 M_t}{\partial W_t^2} \, dW_t^2 + \ldots = \gamma_t \, dW_t
$$

for $0 \leq t \leq T$.

(a) For $M_t = \mathbb{E}^\mathbb{P}(W_T | \mathcal{F}_t)$ and since $W_t$ is a $\mathbb{P}$-martingale (see Problem 2.2.3.1, page 71),

$$
M_t = \mathbb{E}^\mathbb{P}(W_T | \mathcal{F}_t) = W_t
$$

and hence, for $0 \leq t \leq T$,

$$
\gamma_t = \frac{\partial M_t}{\partial W_t} = 1.$$
4.2.2 Girsanov’s Theorem

(b) For $M_t = \mathbb{E}^P \left( W_t^2 \middle| \mathcal{F}_t \right)$ and since $W_t^2 - t$ is a $\mathbb{P}$-martingale (see Problem 2.2.3.2, page 72),

$$\mathbb{E}^P \left( W_t^2 - T \middle| \mathcal{F}_t \right) = W_t^2 - t$$

or

$$M_t = \mathbb{E}^P \left( W_t^2 \middle| \mathcal{F}_t \right) = W_t^2 + T - t.$$

Thus,

$$\gamma_t = \frac{\partial M_t}{\partial W_t} = 2W_t$$

for $0 \leq t \leq T$.

(c) If $M_t = \mathbb{E}^P \left( W_T^3 \middle| \mathcal{F}_t \right)$ and since $W_t^3 - 3tW_t$ is a $\mathbb{P}$-martingale (see Problem 2.2.3.4, page 73),

$$\mathbb{E}^P \left( W_T^3 - 3TW_T \middle| \mathcal{F}_t \right) = W_t^3 - 3tW_t$$

or

$$M_t = \mathbb{E}^P \left( W_t^3 \middle| \mathcal{F}_t \right) = W_t^3 + 3(T - t)W_t.$$

Therefore,

$$\gamma_t = \frac{\partial M_t}{\partial W_t} = 3 \left( W_t^2 + T - t \right)$$

for $0 \leq t \leq T$.

(d) If $M_t = \mathbb{E}^P \left( e^{\sigma W_t} \middle| \mathcal{F}_t \right)$ and since $e^{\sigma W_t - \frac{1}{2} \sigma^2 t}$ is a $\mathbb{P}$-martingale (see Problem 2.2.3.3, page 72),

$$\mathbb{E}^P \left( e^{\sigma W_t - \frac{1}{2} \sigma^2 t} \middle| \mathcal{F}_t \right) = e^{\sigma W_t - \frac{1}{2} \sigma^2 t}$$

or

$$M_t = \mathbb{E}^P \left( e^{\sigma W_t} \middle| \mathcal{F}_t \right) = e^{\sigma W_t + \frac{1}{2} \sigma^2 (T-t)}.$$

Thus,

$$\gamma_t = \frac{\partial M_t}{\partial W_t} = \sigma e^{\sigma W_t + \frac{1}{2} \sigma^2 (T-t)}$$

for $0 \leq t \leq T$.

\[ \square \]

4.2.2 Girsanov’s Theorem

1. **Novikov’s Condition I.** Let $\{ W_t : t \geq 0 \}$ be a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\theta_t$ be an adapted process, $0 \leq t \leq T$. By considering

$$Z_t = e^{-\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds}$$
4.2.2 Girsanov’s Theorem

and if

\[ \mathbb{E}^\mathbb{P} \left( e^{\int_0^T \frac{1}{2} \Theta_s^2 \, ds} \right) < \infty \]

then, without using Itô’s formula, show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

**Solution:** Using the properties of stochastic integrals (see Problems 3.2.1.10, page 112; 3.2.1.11, page 113; and 3.2.2.4, page 126), we can deduce

\[ \int_0^t \Theta_s \, dW_s \sim \mathcal{N} \left( 0, \int_0^t \Theta_s^2 \, ds \right) . \]

In the same vein, for any \( u < t \), the random variable

\[ \int_u^t \Theta_s \, dW_s \sim \mathcal{N} \left( 0, \int_u^t \Theta_s^2 \, ds \right) \]

and \( f_u^a \Theta_s \, dW_s \perp \int_u^t \Theta_s \, dW_s \). To show that \( Z_t \) is a \( \mathbb{P} \)-martingale, we note the following.

(a) Under the filtration \( \mathcal{F}_u \),

\[
\mathbb{E}^\mathbb{P} \left( Z_t \mid \mathcal{F}_u \right) = \mathbb{E}^\mathbb{P} \left( e^{\int_0^t \Theta_s \, dW_s - \frac{1}{2} \int_0^t \Theta_s^2 \, ds} \middle| \mathcal{F}_u \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \mathbb{E}^\mathbb{P} \left( e^{-\int_0^t \Theta_s \, dW_s + \frac{1}{2} \int_0^t \Theta_s^2 \, ds} \mid \mathcal{F}_u \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \mathbb{E}^\mathbb{P} \left( e^{-\int_0^u \Theta_s \, dW_s + \int_u^t \Theta_s \, dW_s} \mid \mathcal{F}_u \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \mathbb{E}^\mathbb{P} \left( e^{-\int_0^u \Theta_s \, dW_s} \right) \mathbb{E}^\mathbb{P} \left( e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \cdot e^{-\frac{1}{2} \int_0^u \Theta_s^2 \, ds} \\
= e^{-\int_0^a \Theta_s \, dW_s - \frac{1}{2} \int_0^t \Theta_s^2 \, ds}.
\]

Therefore, \( \mathbb{E}^\mathbb{P} \left( Z_t \mid \mathcal{F}_u \right) = Z_u \).

(b) Taking the expectation of \( |Z_t| \),

\[
\mathbb{E}^\mathbb{P} \left( |Z_t| \right) = \mathbb{E}^\mathbb{P} \left( e^{-\int_0^t \Theta_s \, dW_s + \frac{1}{2} \int_0^t \Theta_s^2 \, ds} \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \mathbb{E}^\mathbb{P} \left( e^{\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \right) \\
= e^{-\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \cdot e^{\frac{1}{2} \int_0^t \Theta_s^2 \, ds} \\
= 1 < \infty
\]

since, for \( X \sim \log-N(\mu, \sigma^2) \), \( \mathbb{E}^\mathbb{P}(X) = e^{\mu + \frac{1}{2} \sigma^2} \).

(c) Because \( Z_t \) is a function of \( W_t \), it is \( \mathcal{F}_t \)-adapted.
From the results of (a)–(c) we have shown that \( \{Z_t : 0 \leq t \leq T\} \) is a \( \mathbb{P} \)-martingale and because \( Z_t > 0 \), \( \{Z_t\}_{0 \leq t \leq T} \) is a positive \( \mathbb{P} \)-martingale.

\[\square\]

2. **Novikov’s Condition II.** Let \( \{W_t : t \geq 0\} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and let \( \theta_t \) be an adapted process, \( 0 \leq t \leq T \). By considering
\[Z_t = e^{-\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds},\]
and if
\[\mathbb{E}^\mathbb{P} \left( e^{\frac{1}{2} \int_0^T \theta_t^2 dt} \right) < \infty\]
then, using Itô’s formula, show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

**Solution:** Let \( Z_t = f(X_t) = e^{X_t} \) where \( X_t = -\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds \) and by applying Taylor’s expansion and subsequently Itô’s formula,
\[dZ_t = \frac{\partial f}{\partial X_t} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial X_t^2} (dX_t)^2 + \ldots = e^{X_t} (-\theta_t dW_t - \frac{1}{2} \theta_t^2 dt) + \frac{1}{2} e^{X_t} \theta_t^2 dt = -\theta_t Z_t dW_t.\]

Integrating both sides of the equation from \( u \) to \( t \), where \( u < t \),
\[\int_u^t dZ_s = -\int_u^t \theta_s Z_s dW_s \]
\[Z_t = Z_u - \int_u^t \theta_s Z_s dW_s.\]

Under the filtration \( \mathcal{F}_u \) and because \( \int_u^t \theta_s Z_s dW_s \) is independent of \( \mathcal{F}_u \), we have
\[\mathbb{E}^\mathbb{P} \left( Z_t | \mathcal{F}_u \right) = Z_u\]
where \( \mathbb{E}^\mathbb{P} \left( \int_u^t \theta_s Z_s dW_s \right) \bigg| \mathcal{F}_u = \mathbb{E}^\mathbb{P} \left( \int_u^t \theta_s Z_s dW_s \right) = 0 \). Using the same steps as described in Problem 4.2.2.1 (page 194), we can also show that \( \mathbb{E}^\mathbb{P} \left( |Z_t| \right) < \infty \). Since \( Z_t \) is \( \mathcal{F}_t \)-adapted and \( Z_t > 0 \), we have shown that \( \{Z_t : 0 \leq t \leq T\} \) is a positive \( \mathbb{P} \)-martingale.

\[\square\]

3. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions. Let \( \mathbb{Q} \) be another probability measure on \((\Omega, \mathcal{F}, \mathbb{Q})\) such that \( \mathbb{Q} \) is absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \). Using the Radon–Nikodým theorem show that if \( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t = Z_t \) for all \( 0 \leq t \leq T \), then \( Z_t \) is a positive \( \mathbb{P} \)-martingale.
4.2.2 Girsanov’s Theorem

Solution: By definition, the probability measures $\mathbb{P}$ and $\mathbb{Q}$ are functions

$$\mathbb{P} : \Omega \mapsto [0,1] \quad \text{and} \quad \mathbb{Q} : \Omega \mapsto [0,1].$$

Because $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}$ then, from the Radon–Nikodým theorem,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} : \Omega \mapsto (0, \infty) \quad \text{and} \quad \mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} > 0 \right) = 1.$$

Furthermore,

$$\mathbb{E}^\mathbb{P} \left[ \frac{d\mathbb{Q}}{d\mathbb{P}} \right] = \int_\Omega \frac{d\mathbb{Q}}{d\mathbb{P}} \cdot d\mathbb{P} = \int_\Omega d\mathbb{Q} = 1.$$

Given $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = Z_t$, we can therefore deduce that $Z_t > 0$ for all $0 \leq t \leq T$. To show that $Z_t$ is a positive martingale we have the following:

(a) Under the filtration $\mathcal{F}_t$, we define the Radon–Nikodým derivative as

$$Z_t = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right).$$

For $0 \leq s \leq t \leq T$, and using the tower property of conditional expectation (see Problem 1.2.3.11, page 46),

$$\mathbb{E}^\mathbb{P} \left( Z_t \big| \mathcal{F}_s \right) = \mathbb{E}^\mathbb{P} \left[ \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) \bigg| \mathcal{F}_s \right] = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_s \right) = Z_s.$$

(b) Since $\frac{d\mathbb{Q}}{d\mathbb{P}} > 0$, so $|Z_t| = \left| \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} \right| = \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t}$. Therefore, for $0 \leq t \leq T$ and using the tower property of conditional expectation (see Problem 1.2.3.11, page 46),

$$\mathbb{E}^\mathbb{P} (|Z_t|) = \mathbb{E}^\mathbb{P} \left[ \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right]$$

$$\mathbb{E}^\mathbb{P} \left[ \mathbb{E}^\mathbb{P} (|Z_t|) \right] = \mathbb{E}^\mathbb{P} \left[ \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) \right]$$

we have

$$\mathbb{E}^\mathbb{P} \left( |Z_t| \right) = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) = 1 < \infty.$$

(c) $Z_t$ is clearly $\mathcal{F}_t$-adapted for $0 \leq t \leq T$.

From the results of (a)–(c) we have shown that $Z_t$ is a positive $\mathbb{P}$-martingale.
4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space satisfying the usual conditions. Let $\mathbb{Q}$ be another probability measure on $(\Omega, \mathcal{F}, \mathbb{Q})$ such that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}$. Let $\frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t$ for all $0 \leq t \leq T$, so that $Z_t$ is a positive $\mathbb{P}$-martingale. By letting $\{X_t : 0 \leq t \leq T\}$ be an $\mathcal{F}_t$ measurable random variable, show using the Radon–Nikodým theorem that

$$\mathbb{E}^\mathbb{Q}(X_t) = \mathbb{E}^\mathbb{P} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right) \right]$$

and hence deduce that for $A \in \mathcal{F}_t$,

$$\int_A X_t \, d\mathbb{Q} = \int_A X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right) \, d\mathbb{P}.$$ 

**Solution:** From the definition of the Radon–Nikodým derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$ where $Z$ is a positive random variable, we can write

$$\int_\Omega X_t \, d\mathbb{Q} = \int_\Omega X_t Z \, d\mathbb{P}$$

$$\mathbb{E}^\mathbb{Q}(X_t) = \mathbb{E}^\mathbb{P}(X_t Z).$$

Using the tower property of conditional expectation (see Problem 1.2.3.11, page 46),

$$\mathbb{E}^\mathbb{Q}(X_t) = \mathbb{E}^\mathbb{P}(X_t Z)$$

$$= \mathbb{E}^\mathbb{P}[\mathbb{E}^\mathbb{P}(X_t Z | \mathcal{F}_t)]$$

$$= \mathbb{E}^\mathbb{P}[X_t \mathbb{E}^\mathbb{P}(Z | \mathcal{F}_t)].$$

Under the filtration $\mathcal{F}_t$, we define the Radon–Nikodým derivative as

$$Z_t = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t}$$

and therefore

$$\mathbb{E}^\mathbb{Q}(X_t) = \mathbb{E}^\mathbb{P}(X_t Z_t) \quad \text{or} \quad \mathbb{E}^\mathbb{Q}(X_t) = \mathbb{E}^\mathbb{P} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right) \right].$$

By defining

$$\mathbb{1}_A = \begin{cases} 1 & \text{if } A \in \mathcal{F}_t \\ 0 & \text{otherwise} \end{cases}$$

and if $X_t$ is $\mathcal{F}_t$ measurable, then for any $A \in \mathcal{F}_t$,

$$\mathbb{E}^\mathbb{Q}(\mathbb{1}_A X_t) = \mathbb{E}^\mathbb{P}(\mathbb{1}_A X_t Z_t).$$
4.2.2 Girsanov’s Theorem

which is equivalent to

\[ \int_A X_t \, d\mathbb{Q} = \int_A X_t Z_t \, d\mathbb{P} \]

or

\[ \int_A X_t \, d\mathbb{Q} = \int_A X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right) \, d\mathbb{P}. \]

\[ \square \]

5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions. Let \(\mathbb{Q}\) be another probability measure on \((\Omega, \mathcal{F}, \mathbb{Q})\) such that \(\mathbb{Q}\) is absolutely continuous with respect to \(\mathbb{P}\) on \(\mathcal{F}\). Let \(\frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t\) for all \(0 \leq t \leq T\), so that \(Z_t\) is a positive \(\mathbb{P}\)-martingale. By letting \(\{X_t : 0 \leq t \leq T\}\) be an \(\mathcal{F}_t\) measurable random variable, show using the Radon–Nikodým theorem that

\[ \mathbb{E}^\mathbb{P} (X_t) = \mathbb{E}^\mathbb{Q} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right)^{-1} \right] \]

and hence deduce that for \(A \in \mathcal{F}_t\),

\[ \int_A X_t \, d\mathbb{P} = \int_A X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right)^{-1} \, d\mathbb{Q}. \]

**Solution:** From the definition of the Radon–Nikodým derivative \(\frac{d\mathbb{Q}}{d\mathbb{P}} = Z\) where \(Z\) is a positive random variable, we can write

\[ \int_\Omega X_t \, d\mathbb{P} = \int_\Omega X_t Z^{-1} \, d\mathbb{Q} \]

\[ \mathbb{E}^\mathbb{P} (X_t) = \mathbb{E}^\mathbb{Q} (X_t Z^{-1}). \]

Using the tower property of conditional expectation (see Problem 1.2.3.11, page 46),

\[ \mathbb{E}^\mathbb{P} (X_t) = \mathbb{E}^\mathbb{Q} (X_t Z^{-1}) \]

\[ = \mathbb{E}^\mathbb{Q} \left[ \mathbb{E}^\mathbb{Q} (X_t Z^{-1} | \mathcal{F}_t) \right] \]

\[ = \mathbb{E}^\mathbb{Q} \left[ X_t \mathbb{E}^\mathbb{Q} (Z^{-1} | \mathcal{F}_t) \right]. \]

Under the filtration \(\mathcal{F}_t\), we define the Radon–Nikodým derivative as

\[ Z_t = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \]

and therefore

\[ \mathbb{E}^\mathbb{P} (X_t) = \mathbb{E}^\mathbb{Q} (X_t Z_t^{-1}) \quad \text{or} \quad \mathbb{E}^\mathbb{P} (X_t) = \mathbb{E}^\mathbb{Q} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right)^{-1} \right]. \]
By defining
\[ 1_A = \begin{cases} 1 & \text{if } A \in \mathcal{F}_t \\ 0 & \text{otherwise} \end{cases} \]
and if \( X_t \) is \( \mathcal{F}_t \) measurable, then for any \( A \in \mathcal{F}_t \),
\[ \mathbb{E}^{\mathbb{P}}(1_A X_t) = \mathbb{E}^{\mathbb{Q}}(1_A X_t Z_t^{-1}) \]
which is equivalent to
\[ \int A X_t \, d\mathbb{P} = \int A X_t Z_t^{-1} \, d\mathbb{Q} \]
or
\[ \int A X_t \, d\mathbb{P} = \left( \int_A \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \bigg|_{\mathcal{F}_t} \right)^{-1} d\mathbb{Q}. \]
\( \square \)

6. Let \( \{W_t : 0 \leq t \leq T\} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and let \( \mathcal{F}_t \), \( 0 \leq t \leq T \) be the associated Wiener process filtration. By defining
\[ \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t = e^{-\int_0^t \theta_u \, dW_u - \frac{1}{2} \int_0^t \theta_u^2 \, du} \]
where for all \( 0 \leq t \leq T \) the adapted process \( \theta_t \) satisfies \( \mathbb{E}^{\mathbb{P}} \left( e^{\frac{1}{2} \int_0^T \theta_t^2 \, dt} \right) < \infty \), show that \( \mathbb{Q} \) is a probability measure.

**Solution:** From Problem 4.2.2.4 (page 198) we can deduce that for \( A \in \mathcal{F}_t \),
\[ \int A \, d\mathbb{Q} = \int A \, d\mathbb{P} \]
or
\[ \mathbb{Q}(A) = \mathbb{E}^{\mathbb{P}}(Z_t 1_A) \]
where
\[ 1_A = \begin{cases} 1 & \text{if } A \in \mathcal{F}_t \\ 0 & \text{otherwise} \end{cases} \]
In addition, from Problem 4.2.2.1 (page 194) we can deduce that the adapted process \( \theta_t \) follows
\[ \int_0^t \theta_s \, dW_s \sim \mathcal{N} \left( 0, \int_0^t \theta_s^2 \, ds \right) \]
so that
\[ \mathbb{E}^{\mathbb{P}}(Z_t) = e^{-\frac{1}{2} \int_0^t \theta_u^2 \, du} \mathbb{E}^{\mathbb{P}} \left( e^{-\frac{1}{2} \int_0^t \theta_u \, dW_u} \right) = e^{-\frac{1}{2} \int_0^t \theta_u^2 \, du} \cdot e^{\frac{1}{2} \int_0^t \theta_u^2 \, du} = 1. \]
Since \( \mathbb{P} \) is a probability measure and based on the above results, \( \mathbb{Q} \) is also a probability measure because
4.2.2 Girsanov’s Theorem

(a) \( \mathbb{Q}(\emptyset) = \mathbb{E}^\mathbb{P}(Z_t \mathbb{1}_\emptyset) = \mathbb{E}^\mathbb{P}(Z_t) \mathbb{P}(\emptyset) = 0 \) since \( \mathbb{P}(\emptyset) = 0 \).

(b) \( \mathbb{Q}(\Omega) = \mathbb{E}^\mathbb{P}(Z_t \mathbb{1}_\Omega) = \mathbb{E}^\mathbb{P}(Z_t) \mathbb{P}(\Omega) = 1 \) since \( \mathbb{P}(\Omega) = 1 \) and \( \mathbb{E}^\mathbb{P}(Z_t) = 1 \).

(c) For \( A_1, A_2, \ldots \in \mathcal{F}_t \), \( A_i \cap A_j = \emptyset \), \( i \neq j \), \( i, j \in \{1, 2, \ldots \} \)

\[
\bigcup_{i=1}^\infty \mathbb{Q}(A_i) = \bigcup_{i=1}^\infty \mathbb{E}^\mathbb{P}(Z_t \mathbb{1}_{A_i}) = \bigcup_{i=1}^\infty \mathbb{E}^\mathbb{P}(Z_t) \mathbb{P}(A_i) = \bigcup_{i=1}^\infty \mathbb{Q}(A_i).
\]

\[\square\]

7. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space satisfying the usual conditions. Let \( \mathbb{Q} \) be another probability measure on \((\Omega, \mathcal{F}, \mathbb{Q})\) such that \( \mathbb{Q} \) is absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \). Let \( \frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_s} = Z_s \) for all \( 0 \leq t \leq T \), so that \( Z_t \) is a positive \( \mathbb{P} \)-martingale. By letting \( \{X_t : 0 \leq t \leq T\} \) be an \( \mathcal{F}_t \) measurable random variable and using the partial averaging property given as

\[
\int_A \mathbb{E}^\mathbb{P}(Y \mid \mathcal{G}) d\mathbb{P} = \int_A Y d\mathbb{P}, \quad \text{for all } A \in \mathcal{G}
\]

where \( Y \) is a random variable on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and \( \mathcal{G} \) is a sub-\( \sigma \)-algebra of \( \mathcal{F} \) show that for \( 0 \leq s \leq t \leq T \),

\[
\mathbb{E}^\mathbb{Q}(X_t \mid \mathcal{F}_s) = \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_s} \right)^{-1} \mathbb{E}^\mathbb{P} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_s} \right) \mid \mathcal{F}_s \right].
\]

**Solution:** First, note that \( \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_s} \right)^{-1} \mathbb{E}^\mathbb{P} \left[ X_t \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \big|_{\mathcal{F}_s} \right) \mid \mathcal{F}_s \right] = Z_s^{-1} \mathbb{E}^\mathbb{P}(X_t Z_s \mid \mathcal{F}_s) \) is \( \mathcal{F}_s \)-measurable. For any \( A \in \mathcal{F}_s \), and using the results of Problem 4.2.2.4 (page 198) as well as partial averaging, we have

\[
\int_A Z_s^{-1} \mathbb{E}^\mathbb{P}(X_t Z_s \mid \mathcal{F}_s) \ d\mathbb{Q} = \int_A Z_s^{-1} \mathbb{E}^\mathbb{P}(X_t Z_s \mid \mathcal{F}_s) \cdot Z_s \ d\mathbb{P}
\]

\[= \int_A X_t Z_s \ d\mathbb{P}
\]

\[= \int_A X_t \ d\mathbb{Q}.
\]

Using the partial averaging once again, we have

\[
\int_A X_t \ d\mathbb{Q} = \int_A \mathbb{E}^\mathbb{Q}(X_t \mid \mathcal{F}_s) \ d\mathbb{Q}.
\]
Therefore,
\[
\int_A Z_t^{-1} \mathbb{E}^P (X_t | \mathcal{F}_s) \, dQ = \int_A \mathbb{E}^Q (X_t | \mathcal{F}_s) \, dQ
\]
or
\[
Z_t^{-1} \mathbb{E}^P (X_t | \mathcal{F}_s) = \mathbb{E}^Q (X_t | \mathcal{F}_s). 
\]
By substituting \( Z_t = \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \) we have
\[
\mathbb{E}^Q (X_t | \mathcal{F}_s) = \left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right)^{-1} \mathbb{E}^P \left[ X_t \left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right) \bigg| \mathcal{F}_s \right].
\]
\( \square \)

8. Let \((\Omega, \mathcal{F}, P)\) be the probability space satisfying the usual conditions. Let \(Q\) be another probability measure on \((\Omega, \mathcal{F}, P)\) such that \(Q\) is absolutely continuous with respect to \(P\) on \(\mathcal{F}\). Let \(\frac{dQ}{dP} \bigg|_{\mathcal{F}_s} = Z_t \) for all \(0 \leq t \leq T\), so that \(Z_t\) is a positive \(P\)-martingale. By letting \(\{X_t : 0 \leq t \leq T\}\) be an \(\mathcal{F}_s\)-measurable random variable and using the partial averaging property given as
\[
\int_A \mathbb{E}^P (Y | \mathcal{G}) \, dP = \int_A Y \, dP, \quad \text{for all} \quad A \in \mathcal{G}
\]
where \(Y\) is a random variable on the probability space \((\Omega, \mathcal{F}, P)\) and \(\mathcal{G}\) is a sub-\(\sigma\)-algebra of \(\mathcal{F}\) show that for \(0 \leq s \leq t \leq T\),
\[
\mathbb{E}^P (X_t | \mathcal{F}_s) = \left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right)^{-1} \mathbb{E}^P \left[ X_t \left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right) \bigg| \mathcal{F}_s \right].
\]
**Solution:** First, note that
\[
\left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right)^{-1} \mathbb{E}^P \left[ X_t \left( \frac{dQ}{dP} \bigg|_{\mathcal{F}_s} \right) \bigg| \mathcal{F}_s \right] = Z_s \mathbb{E}^Q (X_t Z_t^{-1} | \mathcal{F}_s)
\]
is \(\mathcal{F}_s\) measurable. For any \(A \in \mathcal{F}_s\), and using the results of Problem 4.2.2.5 (page 199) as well as partial averaging, we have
\[
\int_A Z_s \mathbb{E}^Q (X_t Z_t^{-1} | \mathcal{F}_s) \, dP = \int_A Z_s \mathbb{E}^Q (X_t Z_t^{-1} | \mathcal{F}_s) \cdot Z_t^{-1} \, dQ
\]
\[
= \int_A Z_t^{-1} \, dQ
\]
\[
= \int_A X_t \, dP.
\]
Using the partial averaging once again, we have
\[
\int_A X_t \, dP = \int_A \mathbb{E}^P (X_t | \mathcal{F}_s) \, dP.
\]
4.2.2 Girsanov’s Theorem

Therefore,
\[ \int_A Z_s \mathbb{E}^Q (X_t | \mathcal{F}_s) \, d\mathbb{P} = \int_A \mathbb{E}^P (X_t | \mathcal{F}_s) \, d\mathbb{P} \]
or
\[ Z_s \mathbb{E}^Q (X_t | \mathcal{F}_s) = \mathbb{E}^P (X_t | \mathcal{F}_s) . \]

By substituting \( Z_t = \frac{dQ}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \), we have
\[ \mathbb{E}^P (X_t | \mathcal{F}_s) = \left( \frac{dQ}{d\mathbb{P}} \bigg|_{\mathcal{F}_s} \right) \mathbb{E}^Q \left[ X_t \left( \frac{dQ}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} \right)^{-1} \bigg| \mathcal{F}_s \right] . \]

\( \square \)

9. Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and suppose \( \theta_t \) is an adapted process, \( 0 \leq t \leq T \). We consider
\[ Z_t = e^{-\int_0^t \theta_s \, dW_s - \frac{1}{2} \int_0^t \theta_s^2 \, ds} \]
and if
\[ \mathbb{E}^P \left( e^{\frac{1}{2} \int_0^T \theta_s^2 \, ds} \right) < \infty \]
show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that \( \frac{dQ}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t \), show that
\[ \tilde{W}_t = W_t + \int_0^t \theta_u \, du \]
is a \( \mathbb{Q} \)-martingale.

**Solution:** The first part of the result is given in Problem 4.2.2.1 (page 194) or Problem 4.2.2.2 (page 196). To show that \( \tilde{W}_t \) is a \( \mathbb{Q} \)-martingale we note the following:

(a) Let \( 0 \leq s \leq t \leq T \), under the filtration \( \mathcal{F}_s \) and using the result of Problem 4.2.2.7 (page 201), we have
\[ \mathbb{E}^Q \left( \tilde{W}_t | \mathcal{F}_s \right) = \frac{1}{Z_s} \mathbb{E}^P \left( \tilde{W}_t Z_t | \mathcal{F}_s \right) . \]

From Itô’s formula and using the results of Problem 4.2.2.2 (page 196),
\[ d \left( \tilde{W}_t Z_t \right) = \tilde{W}_t dZ_t + Z_t d\tilde{W}_t + d\tilde{W}_t dZ_t \]
\[ = \tilde{W}_t (-\theta_t Z_t dW_t) + Z_t (dW_t + \theta_t dt) + (dW_t + \theta_t dt) (-\theta_t Z_t dW_t) \]
\[ = \left( 1 - \theta_t \tilde{W}_t \right) Z_t dW_t . \]
By integrating both sides of the equation from $s$ to $t$, where $s < t$,

$$
\int_s^t d\left(\tilde{W}_u Z_u\right) = \int_s^t \left(1 - \theta_u \tilde{W}_u\right) Z_u \, dW_u
$$

$$
\tilde{W}_t Z_t = \tilde{W}_s Z_s + \int_s^t \left(1 - \theta_u \tilde{W}_u\right) Z_u \, dW_u.
$$

Under the filtration $\mathcal{F}_t$ and because $\int_s^t \left(1 - \theta_u \tilde{W}_u\right) Z_u \, dW_u \perp \mathcal{F}_s$,

$$
\mathbb{E}^\mathbb{P} \left(\tilde{W}_t Z_t \mid \mathcal{F}_s\right) = \tilde{W}_s Z_s
$$

where $\mathbb{E}^\mathbb{P} \left(\int_s^t \left(1 - \theta_u \tilde{W}_u\right) Z_u \, dW_u \mid \mathcal{F}_s\right) = \mathbb{E}^\mathbb{P} \left(\int_s^t \left(1 - \theta_u \tilde{W}_u\right) Z_u \, dW_u\right) = 0$.

Thus,

$$
\mathbb{E}^\mathbb{Q} \left(\tilde{W}_t \mid \mathcal{F}_s\right) = \frac{1}{Z_s} \mathbb{E}^\mathbb{P} \left(\tilde{W}_t Z_t \mid \mathcal{F}_s\right) = \frac{1}{Z_s} \tilde{W}_s = \tilde{W}_s.
$$

(b) For $0 \leq t \leq T$,

$$
|\tilde{W}_t| = |W_t + \int_0^t \theta_u \, du| \leq |W_t| + \left|\int_0^t \theta_u \, du\right|.
$$

Taking expectations under the $\mathbb{Q}$ measure, using Hölder’s inequality (see Problem 1.2.3.2, page 41) and taking note that the positive random variable $Z_t \sim \log-N\left(-\frac{1}{2} \int_0^t \theta^2_s \, ds, \int_0^t \theta^2_s \, ds\right)$ under $\mathbb{P}$, we have

$$
\mathbb{E}^\mathbb{Q} \left(|\tilde{W}_t|\right) \leq \mathbb{E}^\mathbb{Q} \left(|W_t|\right) + \left|\int_0^t \theta_u \, du\right|
$$

$$
= \mathbb{E}^\mathbb{P} \left(|W_t| Z_t\right) + \left|\int_0^t \theta_u \, du\right|
$$

$$
= \mathbb{E}^\mathbb{P} \left(|W_t| Z_t\right) + \left|\int_0^t \theta_u \, du\right|
$$

$$
\leq \sqrt{\mathbb{E}^\mathbb{P} \left(W_t^2\right) \mathbb{E}^\mathbb{P} \left(Z_t^2\right)} + \left|\int_0^t \theta_u \, du\right|
$$

$$
= \sqrt{te^\frac{1}{2} \int_0^T \theta^2_u \, du} + \left|\int_0^t \theta_u \, du\right|
$$

$$
< \infty
$$

since $\mathbb{E}^\mathbb{P} \left(e^{\frac{1}{2} \int_0^T \theta^2_u \, du}\right) < \infty$.  

4.2.2 Girsanov’s Theorem

(c) Because \( \tilde{W}_t \) is a function of \( W_t \), it is \( \mathcal{F}_t \)-adapted.

From the results of (a)–(c), we have shown that \( \tilde{W}_t, 0 \leq t \leq T \) is a \( \mathbb{Q} \)-martingale.

\[ \square \]

10. One-Dimensional Girsanov Theorem. Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and suppose \( \theta_t \) is an adapted process, \( 0 \leq t \leq T \). We consider

\[ Z_t = e^{-\int_0^t \theta_s \, dW_s} - \frac{1}{2} \int_0^t \theta_s^2 \, ds \]

and if

\[ \mathbb{E}^\mathbb{P} \left( e^{\frac{1}{2} \int_0^T \theta_s^2 \, ds} \right) < \infty \]

show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that \( d\mathbb{Q} \bigg|_{\mathcal{F}_t} = Z_t \), show that

\[ \tilde{W}_t = W_t + \int_0^t \theta_u \, du \]

is a \( \mathbb{Q} \)-standard Wiener process.

Solution: The first part of the result is given in Problem 4.2.2.1 (page 194) or Problem 4.2.2.2 (page 196). As for the second part of the result, we can prove it using two methods.

Method 1. We first need to show that under the \( \mathbb{Q} \) measure, \( \tilde{W}_t \sim \mathcal{N}(0, t) \) and to do this we rely on the moment generating function. By definition, for a constant \( \psi \),

\[ M_{\tilde{W}_t}(\psi) = \mathbb{E}^\mathbb{Q} \left( e^{\psi \tilde{W}_t} \right) \]

\[ = \mathbb{E}^\mathbb{P} \left( e^{\psi \tilde{W}_t} Z_t \right) \]

\[ = \mathbb{E}^\mathbb{P} \left( e^{\psi W_t + \psi \int_0^t \theta_s \, dW_s} - \frac{1}{2} \int_0^t \theta_s^2 \, ds \right) \]

\[ = e^{\int_0^t (\psi \theta_s - \frac{1}{2} \theta_s^2) \, ds} \mathbb{E}^\mathbb{P} \left( e^{\psi W_t - \int_0^t \theta_s \, dW_s} \right) \]

\[ = e^{\int_0^t (\psi \theta_s - \frac{1}{2} \theta_s^2) \, ds} \mathbb{E}^\mathbb{P} \left( e^{\int_0^t \psi \, dW_s - \int_0^t \theta_s \, dW_s} \right) \]

\[ = e^{\int_0^t (\psi \theta_s - \frac{1}{2} \theta_s^2) \, ds} \mathbb{E}^\mathbb{P} \left( e^{\int_0^t (\psi - \theta_s) \, dW_s} \right) \]

Using the properties of the stochastic Itô integral under the \( \mathbb{P} \)-measure

\[ \int_0^t (\psi - \theta_s) \, dW_s \sim \mathcal{N} \left( 0, \int_0^t (\psi - \theta_s)^2 \, ds \right) \]

therefore

\[ \mathbb{E}^\mathbb{P} \left( e^{\int_0^t (\psi - \theta_s) \, dW_s} \right) = e^{\frac{1}{2} \int_0^t (\psi - \theta_s)^2 \, ds} \]
and hence
\[ M_{\tilde{W}_t}^{\psi} = e^{\frac{1}{2} \psi^2 t} \]
which is a moment generating function of a normal distribution with mean zero and variance \( t \).

Thus, under the \( \mathbb{Q} \) measure, \( \tilde{W}_t \sim \mathcal{N}(0, t) \).

Finally, to show that \( \tilde{W}_t, 0 \leq t \leq T \) is a \( \mathbb{Q} \)-standard Wiener process we have the following:

(a) \( \tilde{W}_0 = 0 \) and \( \tilde{W}_t \) certainly has continuous sample paths for \( t > 0 \).

(b) For \( t > 0 \) and \( s > 0 \), \( \tilde{W}_{t+s} - \tilde{W}_t \sim \mathcal{N}(0, s) \) since
\[
\mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \right) = \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} \right) - \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_t \right) = 0
\]
and from the results of Problem 2.2.1.4 (page 57),
\[
\text{Var}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \right) = \text{Var}^{\mathbb{Q}} \left( \tilde{W}_{t+s} \right) + \text{Var}^{\mathbb{Q}} \left( \tilde{W}_t \right) - 2 \text{Cov}^{\mathbb{Q}} \left( \tilde{W}_{t+s}, \tilde{W}_t \right)
= t + s + t - 2 \min\{t + s, t\}
= s.
\]

(c) Because \( \tilde{W}_t \) is a \( \mathbb{Q} \)-martingale (see Problem 4.2.2.9, page 203), then for \( t > 0 \), \( s > 0 \) and under the filtration \( \mathcal{F}_t \),
\[
\mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} \mid \mathcal{F}_t \right) = \tilde{W}_t.
\]

Since we can write
\[
\mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} \mid \mathcal{F}_t \right) = \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t + \tilde{W}_t \mid \mathcal{F}_t \right)
= \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \mid \mathcal{F}_t \right) + \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_t \mid \mathcal{F}_t \right)
= \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \mid \mathcal{F}_t \right) + \tilde{W}_t
\]
then
\[
\mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \mid \mathcal{F}_t \right) = \mathbb{E}^{\mathbb{Q}} \left( \tilde{W}_{t+s} - \tilde{W}_t \right) = 0
\]
which implies \( \tilde{W}_{t+s} - \tilde{W}_t \perp \mathcal{F}_t \). Therefore, \( \tilde{W}_{t+s} - \tilde{W}_t \perp \tilde{W}_t \).

From the results of (a)–(c) we have shown \( \tilde{W}_t, 0 \leq t \leq T \) is a \( \mathbb{Q} \)-standard Wiener process.

Method 2. The process \( \tilde{W}_t, 0 \leq t \leq T \) is a \( \mathbb{Q} \)-martingale with the following properties:

(i) \( \tilde{W}_0 = 0 \) and has continuous paths for \( t > 0 \).

(ii) By setting \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = t, n \in \mathbb{N} \) the quadratic variation of \( \tilde{W}_t \) is
\[
\lim_{n \to \infty} \sum_{i=1}^{n-1} \left( \tilde{W}_{t_{i+1}} - \tilde{W}_{t_i} \right)^2 = \lim_{n \to \infty} \sum_{i=1}^{n-1} \left( W_{t_{i+1}} - W_{t_i} + \int_0^{t_{i+1}} \theta_u \, du - \int_0^{t_i} \theta_u \, du \right)^2
= \lim_{n \to \infty} \sum_{i=1}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right)^2
\]
4.2.2 Girsanov’s Theorem

\[ + \lim_{n \to \infty} \sum_{i=1}^{n-1} 2 \left(W_{t_i} - W_{t_i - 1}\right) \left(\int_{t_i}^{t_{i+1}} \theta_u \, du\right) \]

\[ + \lim_{n \to \infty} \sum_{i=1}^{n-1} \left(\int_{t_i}^{t_{i+1}} \theta_u \, du\right)^2 \]

\[ = t \]

since the quadratic variation of \(W_t\) is \(t\) and \(\lim_{n \to \infty} \int_{t_i}^{t_{i+1}} \theta_u \, du = 0\). Informally, we can write \(d\tilde{W}_t = dt\).

Based on properties (i) and (ii), and from the Lévy characterisation theorem (see Problem 3.2.1.15, page 119), we can deduce that \(\tilde{W}_t, 0 \leq t \leq T\) is a \(\mathbb{Q}\)-standard Wiener process.

\[ \square \]

11. Multi-Dimensional Novikov Condition. Let \(W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})^T\) be an \(n\)-dimensional \(\mathbb{P}\)-standard Wiener process, with \(\{W_t^{(i)} : 0 \leq t \leq T\}, \; i = 1, 2, \ldots, n\) an independent one-dimensional \(\mathbb{P}\)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Suppose we have an \(n\)-dimensional adapted process \(\theta_t = (\theta_t^{(1)}, \theta_t^{(2)}, \ldots, \theta_t^{(n)})^T\), \(0 \leq t \leq T\). By considering

\[ Z_t = \exp \left\{ -\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|_2^2 \, du \right\} \]

and if

\[ E^\mathbb{P} \left( \exp \left\{ \frac{1}{2} \int_0^T \|\theta_t\|_2^2 \, dt \right\} \right) < \infty \]

where

\[ \|\theta\|_2 = \sqrt{\left(\theta_t^{(1)}\right)^2 + \left(\theta_t^{(2)}\right)^2 + \ldots + \left(\theta_t^{(n)}\right)^2} \]

show that \(Z_t\) is a positive \(\mathbb{P}\)-martingale.

**Solution:** To show that \(Z_t\) is a positive \(\mathbb{P}\)-martingale, we have the following.

(a) We can write \(Z_t\) as

\[ Z_t = \exp \left\{ -\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|_2^2 \, du \right\} \]

\[ = \exp \left\{ -\int_0^t \sum_{i=1}^n \theta_u^{(i)} dW_u^{(i)} - \frac{1}{2} \int_0^t \sum_{i=1}^n \left(\theta_u^{(i)}\right)^2 \, du \right\} \]

\[ = \exp \left\{ \sum_{i=1}^n \left( -\int_0^t \theta_u^{(i)} dW_u^{(i)} - \frac{1}{2} \int_0^t \left(\theta_u^{(i)}\right)^2 \, du \right) \right\} \]
Multi-Dimensional Girsanov Theorem

4.2.2 Girsanov’s Theorem

\[ = \prod_{i=1}^{n} \exp \left\{ - \int_{0}^{t} \dot{\theta}_u(i) dW_u(i) - \frac{1}{2} \int_{0}^{t} (\dot{\theta}_u(i))^2 du \right\} \]

\[ = \prod_{i=1}^{n} Z_t(i) \]

where \( Z_t(i) = e^{-\int_{0}^{t} \dot{\theta}_u(i) dW_u(i) - \frac{1}{2} \int_{0}^{t} (\dot{\theta}_u(i))^2 du} \). From Problem 4.2.2.1 (page 194), we have shown that \( Z_t(i) \) is a positive \( \mathbb{P} \)-martingale and since \( \{ W_t(i) : 0 \leq t \leq T \}, i = 1, 2, \ldots, n \) is an independent one-dimensional \( \mathbb{P} \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), so for \( 0 \leq s \leq t \),

\[ \mathbb{E}^\mathbb{P}(Z_t | \mathcal{F}_s) = \prod_{i=1}^{n} \mathbb{E}^\mathbb{P}(Z_t(i) | \mathcal{F}_s) = \prod_{i=1}^{n} Z_t(i) = Z_s. \]

(b) Furthermore, because \( \mathbb{E}^\mathbb{P}(|Z_t(i)|) < \infty \), hence

\[ \mathbb{E}^\mathbb{P}(|Z_t|) = \mathbb{E}^\mathbb{P}\left( \prod_{i=1}^{n} |Z_t(i)| \right) \leq \prod_{i=1}^{n} \mathbb{E}^\mathbb{P}(|Z_t(i)|) \leq \prod_{i=1}^{n} \mathbb{E}^\mathbb{P}(|Z_s(i)|) < \infty. \]

(c) \( Z_t \) is clearly \( \mathcal{F}_t \)-adapted for \( 0 \leq t \leq T \).

From the results of (a)–(c) and since \( Z_t > 0 \), we have shown that \( Z_t \) is a positive \( \mathbb{P} \)-martingale.

\( \square \)

12. **Multi-Dimensional Girsanov Theorem.** Let \( W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})^T \) be an \( n \)-dimensional \( \mathbb{P} \)-standard Wiener process, with \( \{ W_t(i) : 0 \leq t \leq T \}, i = 1, 2, \ldots, n \) an independent one-dimensional \( \mathbb{P} \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Suppose we have an \( n \)-dimensional adapted process \( \theta_t = (\theta_t^{(1)}, \theta_t^{(2)}, \ldots, \theta_t^{(n)})^T \), \( 0 \leq t \leq T \). By considering

\[ Z_t = \exp \left\{ - \int_{0}^{t} \dot{\theta}_u \cdot dW_u - \frac{1}{2} \int_{0}^{t} \| \dot{\theta}_u \|_2^2 du \right\} \]

and if

\[ \mathbb{E}^\mathbb{P}\left( \exp \left\{ \frac{1}{2} \int_{0}^{T} \| \dot{\theta}_u \|_2^2 dt \right\} \right) < \infty \]

where

\[ \| \dot{\theta}_t \|_2 = \sqrt{(\dot{\theta}_t^{(1)})^2 + (\dot{\theta}_t^{(2)})^2 + \ldots + (\dot{\theta}_t^{(n)})^2} \]

show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that \( \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = Z_t \) for all \( 0 \leq t \leq T \), show that

\[ \tilde{W}_t = W_t + \int_{0}^{t} \dot{\theta}_u \, du \]
4.2.2 Girsanov’s Theorem

is an $n$-dimensional $\mathbb{Q}$-standard Wiener process where $\tilde{W}_t = (\tilde{W}_t^{(1)}, \tilde{W}_t^{(2)}, \ldots, \tilde{W}_t^{(n)})^T$ and $\tilde{W}_t^{(i)} = W_t^{(i)} + \int_0^t \theta_u^{(i)} du, i = 1, 2, \ldots, n$ such that the component processes of $\tilde{W}_t$ are independent under $\mathbb{Q}$.

**Solution:** The first part of the result is given in Problem 4.2.2.11 (page 207).

As for the second part of the result, we note that given $\tilde{W}_t = (\tilde{W}_t^{(1)}, \tilde{W}_t^{(2)}, \ldots, \tilde{W}_t^{(n)})^T$, and following the steps given in Problem 4.2.2.9 (page 203), we can easily show that for each $i = 1, 2, \ldots, n$,

$$\tilde{W}_t^{(i)} = W_t^{(i)} + \int_0^t \theta_u^{(i)} du$$

is a $\mathbb{Q}$-martingale. Because $\tilde{W}_0^{(i)} = 0$, and has continuous sample paths with quadratic variation equal to $t$ (see Problem 4.2.2.10, page 205), then from the multi-dimensional Lévy characterisation theorem (see Problem 3.2.1.16, page 121) we can show that $\tilde{W}_t^{(i)}$, $i = 1, 2, \ldots, n$ is a $\mathbb{Q}$-standard Wiener process. Finally, since $W_t^{(i)} \perp W_t^{(j)}, i \neq j, i, j = 1, 2, \ldots, n$ we have

$$d\tilde{W}_t^{(i)} \cdot d\tilde{W}_t^{(j)} = (dW_t^{(i)} + \theta_t^{(i)} dt) (dW_t^{(j)} + \theta_t^{(j)} dt)$$

$$= dW_t^{(i)} dW_t^{(j)} + \theta_t^{(i)} dW_t^{(j)} dt + \theta_t^{(j)} dW_t^{(i)} dt + \theta_t^{(i)} \theta_t^{(j)} dt^2$$

$$= 0.$$

Thus, from the multi-dimensional Lévy characterisation theorem (see Problem 3.2.1.16, page 121) we can deduce that $\tilde{W}_t, 0 \leq t \leq T$ is an $n$-dimensional $\mathbb{Q}$-standard Wiener process such that each component process of $\tilde{W}_t$ is independent under $\mathbb{Q}$.

13. **Convergence Issue of Equivalent Measures.** Let $W_t = (W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)})^T$ be an $n$-dimensional $\mathbb{P}$-standard Wiener process, with $\{W_t^{(i)} : 0 \leq t \leq T\}, i = 1, 2, \ldots, n$ an independent one-dimensional $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose we have an $n$-dimensional adapted process $\theta_t = (\theta_t^{(1)}, \theta_t^{(2)}, \ldots, \theta_t^{(n)})^T, 0 \leq t \leq T$.

We consider

$$Z_t = \exp \left\{ -\int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|_2^2 du \right\}$$

and if

$$\mathbb{E}^{\mathbb{P}} \left( \exp \left\{ \frac{1}{2} \int_0^T \|\theta_t\|_2^2 dt \right\} \right) < \infty$$

where

$$\|\theta_t\|_2 = \sqrt{(\theta_t^{(1)})^2 + (\theta_t^{(2)})^2 + \ldots + (\theta_t^{(n)})^2}$$

show that $Z_t$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$. 

By changing the measure $\mathbb{P}$ to a measure $\mathbb{Q}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}}_{\mathcal{F}_t} = Z_t$ for all $0 \leq t \leq T$, show that

$$\tilde{W}_t = W_t + \int_0^t \theta_u \, du$$

is an $n$-dimensional $\mathbb{Q}$-standard Wiener process where $\tilde{W}_i = (\tilde{W}^{(1)}_i, \tilde{W}^{(2)}_i, \ldots, \tilde{W}^{(n)}_i)$ and $\tilde{W}^{(i)}_t = W^{(i)}_t + \int_0^t \theta^{(i)}_u \, du, i = 1, 2, \ldots, n$ such that the component processes of $\tilde{W}_t$ are independent under $\mathbb{Q}$. Let $X^{(i)} = \int_0^t \theta^{(i)}_u \, dW^{(i)}_u, i = 1, 2, \ldots, n$ be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assume $\theta^{(1)}_t = \theta^{(2)}_t = \ldots = \theta^{(n)}_t = \theta_t$ where $\theta_t$ is an adapted process, $0 \leq t \leq T$. For $n \to \infty$ show that under the $\mathbb{P}$ measure

$$\mathbb{P} \left( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X^{(i)} = 0 \right) = 1$$

and under the equivalent $\mathbb{Q}$ measure

$$\mathbb{Q} \left( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X^{(i)} = 0 \right) = 0.$$

Discuss the implications on the equivalent measures $\mathbb{P}$ and $\mathbb{Q}$.

**Solution:** The first two results are given in Problem 4.2.2.12 (page 208). As for the final result, because $\theta^{(1)}_t = \theta^{(2)}_t = \ldots = \theta^{(n)}_t = \theta_t$ then under the $\mathbb{P}$ measure we can deduce that

$$X^{(i)} = \int_0^t \theta^{(i)}_u \, dW^{(i)}_u \sim \mathcal{N} \left( 0, \int_0^t \theta^2_u \, du \right)$$

for $i = 1, 2, \ldots, n$. Because $\{W^{(i)}_t : 0 \leq t \leq T\}, i = 1, 2, \ldots, n$ are independent and identically, distributed,

$$\frac{1}{n} \sum_{i=1}^{n} X^{(i)} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \int_0^t \theta^{(i)}_u \, dW^{(i)}_u \right\} \sim \mathcal{N} \left( 0, \frac{1}{n} \int_0^t \theta^2_u \, du \right)$$

or

$$\frac{1}{n} \sum_{i=1}^{n} X^{(i)} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \int_0^t \theta^{(i)}_u \, dW^{(i)}_u \right\} = \phi \sqrt{\frac{1}{n} \int_0^t \theta^2_u \, du}, \quad \phi \sim \mathcal{N}(0, 1).$$

Taking limits $n \to \infty$, we have

$$\mathbb{P} \left( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X^{(i)} = 0 \right) = 1.$$
4.2.2 Girsanov’s Theorem

On the contrary, under the measure \( \mathbb{Q} \) for \( i = 1, 2, \ldots, n \),

\[
\tilde{W}_t^{(i)} = W_t^{(i)} + \int_0^t \theta_u^{(i)} \, du
\]

is a \( \mathbb{Q} \)-standard Wiener process and \( \tilde{W}_t^{(1)}, \tilde{W}_t^{(2)}, \ldots, \tilde{W}_t^{(n)} \) are independent and identically distributed. Thus,

\[
d\tilde{W}_t^{(i)} = dW_t^{(i)} + \theta_t^{(i)} \, dt
\]

and in turn we can write

\[
\int_0^t \theta_u^{(i)} \, dW_u^{(i)} = \int_0^t \theta_u^{(i)} \, d\tilde{W}_u^{(i)} - \int_0^t (\theta_u^{(i)})^2 \, du.
\]

Because \( \theta_t^{(1)} = \theta_t^{(2)} = \ldots = \theta_t^{(n)} = \theta_t \) and under the \( \mathbb{Q} \) measure,

\[
X^{(i)} = \int_0^t \theta_u^{(i)} \, dW_u^{(i)} \sim \mathcal{N} \left( - \int_0^t \theta_u^2 \, du, \frac{1}{n} \int_0^t \theta_u^2 \, du \right).
\]

Since \( \{ \tilde{W}_t^{(i)} \}_{0 \leq t \leq T} \), \( i = 1, 2, \ldots, n \) are independent and identically distributed,

\[
\frac{1}{n} \sum_{i=1}^n X^{(i)} = \frac{1}{n} \sum_{i=1}^n \left\{ \int_0^t \theta_u^{(i)} \, dW_u^{(i)} \right\} \sim \mathcal{N} \left( - \int_0^t \theta_u^2 \, du, \frac{1}{n} \int_0^t \theta_u^2 \, du \right)
\]

or

\[
\frac{1}{n} \sum_{i=1}^n X^{(i)} = \frac{1}{n} \sum_{i=1}^n \left\{ \int_0^t \theta_u^{(i)} \, dW_u^{(i)} \right\} = - \int_0^t \theta_u^2 \, du + \phi \sqrt{\frac{1}{n} \int_0^t \theta_u^2 \, du}, \quad \phi \sim \mathcal{N}(0, 1).
\]

By setting \( n \to \infty \),

\[
\mathbb{Q} \left( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n X^{(i)} = 0 \right) = 0.
\]

Therefore, in the limit \( n \to \infty \), the measures \( \mathbb{P} \) and \( \mathbb{Q} \) fail to be equivalent.

\[\square\]

14. Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). By considering a Wiener process with drift

\[
\hat{W}_t = \alpha t + W_t
\]

and by using Girsanov’s theorem to find a measure \( \mathbb{Q} \) under which \( \hat{W}_t \) is a standard Wiener process, show that the joint density of

\[
\hat{W}_t \quad \text{and} \quad \hat{M}_t = \max_{0 \leq s \leq T} \hat{W}_s
\]
under \( \mathbb{P} \) is
\[
f^\mathbb{P}_{\hat{M}_t, \hat{W}_t}(x, w) = \begin{cases} \frac{2(2x-w)}{r\sqrt{2\pi}} e^{-\frac{1}{2}w^2 - \frac{1}{2}(2x-w)^2} & x \geq 0, w \leq x \\ 0 & \text{otherwise.} \end{cases}
\]

Finally, deduce that the joint density of \( \hat{W}_t \) and \( \hat{m}_t = \min_{0 \leq s \leq t} \hat{W}_s \)
under \( \mathbb{P} \) is
\[
f^\mathbb{P}_{\hat{m}_t, \hat{W}_t}(y, w) = \begin{cases} \frac{-2(2y-w)}{r\sqrt{2\pi}} e^{-\frac{1}{2}w^2 - \frac{1}{2}(2y-w)^2} & y \leq 0, y \leq w \\ 0 & \text{otherwise.} \end{cases}
\]

**Solution:** By defining \( \hat{W}_t = \alpha t + W_t \) and \( \hat{M}_t = \max_{0 \leq s \leq t} \hat{W}_s \)
using Itô's formula we can write \( d\hat{W}_t = d\hat{W}_t \)
where \( \hat{W}_t = W_t + \int_0^t \alpha \, du \). From Girsanov's theorem there exists an equivalent probability measure \( \mathbb{Q} \) on the filtration \( \mathcal{F}_s \), \( 0 \leq s \leq t \) defined by the Radon–Nikodým derivative
\[
Z_s = e^{-\int_0^s \alpha \, dW_u - \frac{1}{2} \int_0^s \alpha^2 \, du}
= e^{-\alpha W_s - \frac{1}{2} \alpha^2 s}
= e^{-\alpha (\hat{W}_s - \alpha s) - \frac{1}{2} \alpha^2 s}
= e^{-\alpha \hat{W}_s + \frac{1}{2} \alpha^2 s}
\]
so that \( \hat{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process.

From Problem 2.2.5.4 (page 86), under \( \mathbb{Q} \), the joint probability distribution of \( (\hat{M}_t, \hat{W}_t) \)
can be written as
\[
f^\mathbb{Q}_{\hat{M}_t, \hat{W}_t}(x, w) = \begin{cases} \frac{2(2x-w)}{r\sqrt{2\pi}} e^{-\frac{1}{2}w^2 - \frac{1}{2}(2x-w)^2} & x \geq 0, w \leq x \\ 0 & \text{otherwise.} \end{cases}
\]

In order to find the joint cumulative distribution function of \( (\hat{M}_t, \hat{W}_t) \) under \( \mathbb{P} \), we use the result given in Problem 4.2.2.5 (page 199)
\[
\mathbb{P}\left( \hat{M}_t \leq x, \hat{W}_t \leq w \right) = \mathbb{E}^\mathbb{P}\left( \mathbb{1}_{\{\hat{M}_t \leq x, \hat{W}_t \leq w\}} \right)
= \mathbb{E}^\mathbb{Q}\left( Z^{-1}_t \mathbb{1}_{\{\hat{M}_t \leq x, \hat{W}_t \leq w\}} \right)
\]
4.2.2 Girsanov’s Theorem

$$\mathbb{E}^\mathbb{Q} \left( e^{a\hat{W}_t - \frac{1}{2}a^2t I_{\{\hat{M}_t \leq x, \hat{W}_t \leq w\}} } \right)$$

$$= \int_{-\infty}^{w} \int_{-\infty}^{x} e^{a\hat{W}_t - \frac{1}{2}a^2t} f^\mathbb{Q}_{\hat{M}_t, \hat{W}_t}(u, v) dudv.$$ 

Since \(\hat{W}_0 = 0\), \(\hat{M}_t \geq 0\) therefore \(\hat{W}_t \leq \hat{M}_t\). By definition, the joint density of \((\hat{M}_t, \hat{W}_t)\) under \(\mathbb{P}\) is

$$f^\mathbb{P}_{\hat{M}_t, \hat{W}_t}(x, w) = \frac{\partial^2}{\partial x \partial w} \mathbb{P} \left( \hat{M}_t \leq x, \hat{W}_t \leq w \right)$$

which implies

$$f^\mathbb{P}_{\hat{M}_t, \hat{W}_t}(x, w) = \begin{cases} 
\frac{2(2x-w)}{t\sqrt{2\pi t}} e^{aw - \frac{1}{2}a^2t - \frac{1}{2}(2x-w)^2} & y \geq 0, w \leq x \\
0 & \text{otherwise.}
\end{cases}$$

For the case of the joint density of

$$\hat{W}_t = at + W_t \quad \text{and} \quad \hat{M}_t = \min_{0 \leq s \leq t} \hat{W}_s$$

from Problem 2.2.5.5 (page 88), the joint probability distribution of \((\hat{M}_t, \hat{W}_t)\) under \(\mathbb{Q}\) is

$$f^\mathbb{Q}_{\hat{M}_t, \hat{W}_t}(y, w) = \begin{cases} 
\frac{-2(2y-w)}{t\sqrt{2\pi t}} e^{aw - \frac{1}{2}a^2t - \frac{1}{2}(2y-w)^2} & y \leq 0, y \leq w \\
0 & \text{otherwise.}
\end{cases}$$

Using the same techniques as discussed earlier, the joint cumulative distribution of \((\hat{M}_t, \hat{W}_t)\) under \(\mathbb{P}\) is

$$\mathbb{P} \left( \hat{M}_t \leq y, \hat{W}_t \leq w \right) = \mathbb{E}^\mathbb{P} \left( I_{\{\hat{M}_t \leq y, \hat{W}_t \leq w\}} \right)$$

$$= \mathbb{E}^\mathbb{Q} \left( Z_t^{-1} I_{\{\hat{M}_t \leq y, \hat{W}_t \leq w\}} \right)$$

$$= \mathbb{E}^\mathbb{Q} \left( e^{a\hat{W}_t - \frac{1}{2}a^2t I_{\{\hat{M}_t \leq y, \hat{W}_t \leq w\}} } \right)$$

$$= \int_{-\infty}^{w} \int_{-\infty}^{y} e^{a\hat{W}_t - \frac{1}{2}a^2t} f^\mathbb{Q}_{\hat{M}_t, \hat{W}_t}(u, v) dudv.$$ 

Therefore,

$$f^\mathbb{P}_{\hat{M}_t, \hat{W}_t}(y, w) = \frac{\partial^2}{\partial y \partial w} \mathbb{P} \left( \hat{M}_t \leq y, \hat{W}_t \leq w \right)$$

$$= \begin{cases} 
\frac{-2(2y-w)}{t\sqrt{2\pi t}} e^{aw - \frac{1}{2}a^2t - \frac{1}{2}(2y-w)^2} & y \leq 0, y \leq w \\
0 & \text{otherwise.}
\end{cases}$$

\(\square\)
15. **Running Maximum and Minimum of a Wiener Process.** Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) and let 
\[
X_t = \nu + \mu t + \sigma W_t
\]
be a Wiener process starting from \( \nu \in \mathbb{R} \) with drift \( \mu \in \mathbb{R} \) and volatility \( \sigma > 0 \). Let the running maximum of the process \( X_t \) up to time \( t \) be defined as 
\[
M_{Xt}^t = \max_{0 \leq s \leq t} X_s.
\]
Using Girsanov’s theorem to find a measure \( \mathbb{Q} \) under which \( X_t \) is a standard Wiener process, show that the cumulative distribution function of the running maximum is 
\[
\mathbb{P} (M_{Xt}^t \leq x) = \Phi \left( \frac{x - \nu - \mu t}{\sigma \sqrt{t}} \right) - e^{\frac{2\mu(x+)}{\sigma^2}} \Phi \left( \frac{x + \nu - \mu t}{\sigma \sqrt{t}} \right), \quad x \geq \nu
\]
where \( \Phi(\cdot) \) is the standard normal cumulative distribution function. Finally, deduce that the cumulative distribution function of the running minimum 
\[
m_{Xt}^t = \min_{0 \leq s \leq t} X_s
\]
is 
\[
\mathbb{P} (m_{Xt}^t \leq x) = \Phi \left( \frac{x - \nu - \mu t}{\sigma \sqrt{t}} \right) + e^{\frac{2\mu(x+)}{\sigma^2}} \Phi \left( \frac{x + \nu + \mu t}{\sigma \sqrt{t}} \right), \quad x \leq \nu.
\]
Find the probability density functions for \( M_{Xt}^t \) and \( m_{Xt}^t \).

**Solution:** Let \( Y_t = \frac{X_t - \nu}{\sigma} \) such that 
\[
Y_t = \alpha t + W_t
\]
where \( \alpha = \mu / \sigma \). Following this, we can define 
\[
M_{Yt}^t = \max_{0 \leq s \leq t} Y_s
\]
to be the running maximum of \( Y_t = \frac{X_t - \nu}{\sigma} \). From Problem 4.2.2.14 (page 211) we can write the joint density of \( M_{Yt}^t \) and \( Y_t \) as 
\[
f_{M_{Yt}^t, Y_t}^\mathbb{P}(y, w_y) = \begin{cases} 
\frac{2(2y-w_y)}{\sqrt{2\pi}y} e^{w_y - \frac{1}{2} y^2 - \frac{1}{2} \left(2w_y - w\right)^2} & y \geq 0, w_y \leq y \\
0 & \text{otherwise.}
\end{cases}
\]
Given that the pair of random variables \( (M_{Yt}^t, Y_t) \) only take values from the set 
\[
\{(y, w_y) : y \geq 0, y \geq w_y \}
\]
then in order to get the cumulative distribution function of $M^y_t$ we compute over the shaded region given in Figure 4.1

$$
\Pr(M^y_t \leq y) = \int_{-\infty}^{0} \int_{0}^{y} f_{M^y_t, Y_t}^p(u, v) \, du \, dv + \int_{0}^{y} \int_{v}^{y} f_{M^y_t, Y_t}^p(u, v) \, du \, dv
$$
such that

$$
\int_{-\infty}^{0} \int_{0}^{y} f_{M^y_t, Y_t}^p(u, v) \, du \, dv = \int_{-\infty}^{0} \int_{0}^{y} \frac{2(2u - v)}{t\sqrt{2\pi}} e^{-\frac{1}{2}u^2t - \frac{1}{2t}(2u-v)^2} \, du \, dv
$$

$$
= - \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2t - \frac{1}{2t}(2u-v)^2} \bigg|_{u=y} \, dv
$$

$$
= - \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{0} e^{-\frac{1}{2}u^2t - \frac{1}{2t}(2y-v)^2} \, dv
$$

$$
+ \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{0} e^{-\frac{1}{2}u^2t - \frac{1}{2t}v^2} \, dv
$$

and using similar steps we have

$$
\int_{0}^{y} \int_{v}^{y} f_{M^y_t, Y_t}^p(u, v) \, du \, dv = - \frac{1}{\sqrt{2\pi t}} \int_{0}^{y} e^{-\frac{1}{2}a^2t - \frac{1}{2t}(2y-v)^2} \, dv
$$

$$
+ \frac{1}{\sqrt{2\pi t}} \int_{0}^{y} e^{-\frac{1}{2}a^2t - \frac{1}{2t}v^2} \, dv.
$$
Therefore,
\[ P(M_t^Y \leq y) = -\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{\frac{-1}{2} u^2 - \frac{1}{2} (2y-u)^2} du + \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{\frac{-1}{2} u^2 - \frac{1}{2} (2y-u)^2} du \]
and by completing the squares we have
\[ \alpha v - \frac{1}{2} \alpha^2 t - \frac{1}{2} (2y - v)^2 = -\frac{1}{2} (v - 2y - \alpha t)^2 + 2ay \]
\[ \alpha v - \frac{1}{2} \alpha^2 t - \frac{1}{2} v^2 = -\frac{1}{2} (v - \alpha t)^2. \]
Thus, for \( y \geq 0 \),
\begin{align*}
P(M_t^Y \leq y) &= -\frac{e^{2ay}}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{-\frac{1}{2} (v - 2y - \alpha t)^2} dv + \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{y} e^{-\frac{1}{2} (v - \alpha t)^2} dv \\
&= -e^{2ay} \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} dz + \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} dz \\
&= -e^{2ay} \Phi \left( \frac{-y - \alpha t}{\sqrt{t}} \right) + \Phi \left( \frac{y - \alpha t}{\sqrt{t}} \right)
\end{align*}
or
\[ P(M_t^Y \leq y) = \Phi \left( \frac{y - \alpha t}{\sqrt{t}} \right) - e^{2ay} \Phi \left( \frac{-y - \alpha t}{\sqrt{t}} \right). \]

By substituting
\[ M_t^Y = \frac{M_t^X - \nu}{\sigma}, \quad y = \frac{x - \nu}{\sigma}, \quad \text{and} \quad a = \frac{\mu}{\sigma} \]
we eventually have
\begin{align*}
P(M_t^Y \leq x) &= \Phi \left( \frac{x - \nu - \mu t}{\sigma \sqrt{t}} \right) - e^{-\frac{2\mu(x+y)}{\sigma^2}} \Phi \left( \frac{-x + \nu - \mu t}{\sigma \sqrt{t}} \right), \quad x \geq \nu.
\end{align*}

Finally, we focus on the minimum value of \( X_t \). Take note that \( Y_0 = 0 \) and hence by defining
\[ m_t^Y = \min_{0 \leq s \leq t} Y_s \]
we have \( m_t^Y \leq 0 \). Therefore, for \( y \leq 0 \) and following the symmetry of the standard Wiener process,
\[ P \left( m_t^Y \leq y \right) = P \left( \min_{0 \leq s \leq t} Y_s \leq y \right) \]
\[ = P \left( -\max_{0 \leq s \leq t} \{-Y_s\} \leq y \right) \]
\[\begin{align*}
\mathbb{P}
&= \mathbb{P}
\left(-\max_{0 \leq s \leq t} \{-as - W_s\} \leq y\right) \\
&= \mathbb{P}
\left(-\max_{0 \leq s \leq t} \{-as + W_s\} \leq y\right) \\
&= \mathbb{P}
\left(\max_{0 \leq s \leq t} \tilde{Y}_s \geq -y\right)
\end{align*}\]

where \(\tilde{Y}_t = -at + W_t\). Thus,

\[\begin{align*}
\mathbb{P}(mY_t \leq y) &= 1 - \mathbb{P}
\left(\max_{0 \leq s \leq t} \tilde{Y}_s \leq -y\right) \\
&= 1 - \Phi\left(\frac{-y - at}{\sqrt{t}}\right) + e^{2\alpha y} \Phi\left(\frac{y + at}{\sqrt{t}}\right) \\
&= \Phi\left(\frac{y - at}{\sqrt{t}}\right) + e^{2\alpha y} \Phi\left(\frac{y + at}{\sqrt{t}}\right).
\end{align*}\]

Substituting

\[m_t^Y = \frac{m_t^X - \nu}{\sigma}, \quad y = \frac{x - \nu}{\sigma} \quad \text{and} \quad \alpha = \frac{\mu}{\sigma}\]

we have

\[\mathbb{P}(m_t^X \leq x) = \Phi\left(\frac{x - \nu - \mu t}{\sigma \sqrt{t}}\right) + e^{2\alpha (x - \nu)} \Phi\left(\frac{x - \nu + \mu t}{\sigma \sqrt{t}}\right), \quad x \leq \nu.\]

By denoting \(f_{M_t^X}(x)\) and \(f_{m_t^X}(x)\) as the probability density functions for \(M_t^X\) and \(m_t^X\), respectively, therefore

\[f_{M_t^X}(x) = \frac{d}{dx} \mathbb{P}(M_t^X \leq x) \]

\[= \frac{d}{dx}\left[\Phi\left(\frac{x - \nu - \mu t}{\sigma \sqrt{t}}\right) - e^{2\alpha (x - \nu)} \Phi\left(\frac{-x + \nu - \mu t}{\sigma \sqrt{t}}\right)\right] \]

\[= \frac{1}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2}\left(\frac{x - \nu - \mu t}{\sigma \sqrt{t}}\right)^2} - \frac{2\mu}{\sigma^2} e^{2\alpha (x - \nu)} \Phi\left(\frac{-x + \nu - \mu t}{\sigma \sqrt{t}}\right) \]

\[+ \frac{1}{\sigma \sqrt{2\pi t}} e^{\frac{2\alpha (x - \nu)}{\sigma^2}} \cdot \frac{1}{2}\left(\frac{x - \nu + \mu t}{\sigma \sqrt{t}}\right)^2 \]

\[= \frac{2}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2}\left(\frac{x - \nu - \mu t}{\sigma \sqrt{t}}\right)^2} - \frac{2\mu}{\sigma^2} e^{2\alpha (x - \nu)} \Phi\left(\frac{-x + \nu - \mu t}{\sigma \sqrt{t}}\right), \quad x \geq \nu\]
4.2.2 Girsanov’s Theorem

\[ f_{m_t}(x) = \frac{d}{dx} \mathbb{P}(m_t^X \leq x) \]

\[ = \frac{d}{dx} \left[ \Phi \left( \frac{x - \nu - \mu t}{\sigma \sqrt{t}} \right) + e^{\frac{2\mu(x-\nu)}{\sigma^2}} \Phi \left( \frac{x - \nu + \mu t}{\sigma \sqrt{t}} \right) \right] \]

\[ = \frac{1}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2} \left( \frac{x-\nu-\mu t}{\sigma \sqrt{t}} \right)^2} + \frac{2\mu}{\sigma^2} e^{\frac{2\mu(x-\nu)}{\sigma^2}} \Phi \left( \frac{x - \nu + \mu t}{\sigma \sqrt{t}} \right) + \frac{1}{\sigma \sqrt{2\pi t}} e^{\frac{2\mu(x-\nu)}{\sigma^2} - \frac{1}{2} \left( \frac{x-\nu+\mu t}{\sigma \sqrt{t}} \right)^2} \]

\[ = \frac{2}{\sigma \sqrt{2\pi t}} e^{-\frac{1}{2} \left( \frac{x-\nu-\mu t}{\sigma \sqrt{t}} \right)^2} + \frac{2\mu}{\sigma^2} e^{\frac{2\mu(x-\nu)}{\sigma^2}} \Phi \left( \frac{x - \nu + \mu t}{\sigma \sqrt{t}} \right) \]

\[ , \quad x \leq \nu. \]

\[ \square \]

16. First Passage Time Density of a Standard Wiener Process Hitting a Sloping Line. Let \( \{W_t : t \geq 0\} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). By setting \( T_{a+\beta t} \) as a stopping time such that \( T_{a+\beta t} = \inf \{ t \geq 0 : W_t = a + \beta t \} \), \( a, \beta \in \mathbb{R}, a \neq 0 \) show that the probability density function of \( T_{a+\beta t} \) is given as

\[ f_{T_{a+\beta t}}(t) = \frac{|a|}{t \sqrt{2\pi t}} e^{-\frac{1}{2} \beta^2 t}. \]

Solution: By defining \( \tilde{W}_t = W_t - \int_0^t \beta \, dW_u \) such that

\[ T_{a+\beta t} = \inf \{ t \geq 0 : \tilde{W}_t = a \} \]

then, from Girsanov’s theorem, there exists an equivalent probability measure \( \mathbb{Q} \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \) defined by the Radon–Nikodým derivative

\[ Z_s = e^{\int_0^s \beta \, dW_u - \frac{1}{2} \int_0^s \beta^2 \, du} \]

\[ = e^{\beta \mathbb{W}_s - \frac{1}{2} \beta^2 s} \]

\[ = e^{\beta (\tilde{W}_s + \beta s) - \frac{1}{2} \beta^2 s} \]

\[ = e^{\beta \tilde{W}_s + \frac{1}{2} \beta^2 s} \]

so that \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process. From Problem 2.2.5.3 (page 85), the probability density function of \( T_{a+\beta t} = \inf \{ t \geq 0 : \tilde{W}_t = a \} \) under \( \mathbb{Q} \) is therefore

\[ \frac{d}{dt} f_{T_{a+\beta t}}(t) = \frac{|a|}{t \sqrt{2\pi t}} e^{-\frac{1}{2} \beta^2 t}. \]

To find the probability density of \( T_{a+\beta t} \) under \( \mathbb{P} \) we note that

\[ \mathbb{P} \left( T_{a+\beta t} \leq t \right) = \mathbb{E}^{\mathbb{P}} \left( 1_{T_{a+\beta t}} \right) \]
4.2.2 Girsanov’s Theorem

\[
\mathbb{E}^Q \left( Z_t^{-1} \mathbb{I}_{T_{\alpha + \beta \leq t}} \right) \\
= \mathbb{E}^Q \left( e^{-\beta \tilde{W}_t - \frac{1}{2} \beta^2 t} \mathbb{I}_{T_{\alpha + \beta \leq t}} \right) \\
= \mathbb{E}^Q \left( e^{-a \beta - \frac{1}{2} \beta^2 t} \mathbb{I}_{T_{\alpha + \beta \leq t}} \right) \\
= \int_0^t e^{-a \beta - \frac{1}{2} \beta^2 u} f_{T_{\alpha + \beta}}(u) \, du \\
= \int_0^t e^{-a \beta - \frac{1}{2} \beta^2 u} \left[ \frac{|\alpha|}{u \sqrt{2 \pi u}} e^{-\frac{1}{2} (a + \beta u)^2} \right] \, du.
\]

Since

\[
f_{T_{\alpha + \beta}}(t) = \frac{d}{dt} \mathbb{P}(T_{\alpha + \beta} \leq t)
\]

we therefore have

\[
f_{T_{\alpha + \beta}}(t) = \frac{|\alpha|}{t \sqrt{2 \pi t}} e^{-\frac{1}{2} (a + \beta t)^2}.
\]

\[\square\]

17. Let \( \{ W_t : t \geq 0 \} \) be a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), and let \( \mathcal{F}_t \) be the filtration generated by \( W_t \). Suppose \( \theta_t \) is an adapted process, \( 0 \leq t \leq T \) and by considering

\[
Z_t = e^{-\int_0^t \theta_s \, dW_s - \frac{1}{2} \int_0^t \theta_s^2 \, ds}
\]

and if

\[
\mathbb{E}^P \left( e^{\frac{1}{2} \int_0^T \theta_t^2 \, dt} \right) < \infty
\]

then \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \). By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that \( \frac{d\mathbb{Q}}{d\mathbb{P}} |_{\mathcal{F}_t} = Z_t \), from Girsanov’s theorem

\[
\tilde{W}_t = W_t + \int_0^t \theta_u \, du, \quad 0 \leq t \leq T
\]

is a \( \mathbb{Q} \)-standard Wiener process. If \( \tilde{M}_t, 0 \leq t \leq T \) is a \( \mathbb{Q} \)-martingale, then show that \( M_t = Z_t \tilde{M}_t, 0 \leq t \leq T \) is a \( \mathbb{P} \)-martingale.

Using the martingale representation theorem, show that there exists an adapted process \( \{ \tilde{\gamma}_t : 0 \leq t \leq T \} \) such that

\[
\tilde{M}_t = \tilde{M}_0 + \int_0^t \tilde{\gamma}_u \, d\tilde{W}_u, \quad 0 \leq t \leq T.
\]
Solution: To show that $M_t = Z_t \widetilde{M}_t$ is a $\mathbb{P}$-martingale, we note the following:

(a) From Problem 4.2.2.8 (page 202),

$$\mathbb{E}^\mathbb{P}(M_t | \mathcal{F}_s) = Z_s \mathbb{E}^\mathbb{Q}\left( M_t Z_t^{-1} | \mathcal{F}_s \right) = Z_s \widetilde{M}_s = M_s.$$ 

(b) From Problem 4.2.2.5 (page 199),

$$\mathbb{E}^\mathbb{P}(|M_t|) = \mathbb{E}^\mathbb{Q}\left( |M_t| \frac{1}{Z_t} \right) = \mathbb{E}^\mathbb{Q}\left( |M_t Z_t^{-1}| \right) = \mathbb{E}^\mathbb{Q}\left( |\widetilde{M}_t| \right) < \infty$$

since $\widetilde{M}_t$ is a $\mathbb{Q}$-martingale.

(c) $M_t = Z_t \widetilde{M}_t$ is clearly $\mathcal{F}_t$-adapted.

From the results of (a)–(c) we have shown that $M_t = Z_t \widetilde{M}_t$ is a $\mathbb{P}$-martingale.

Using Itô’s formula on $d \widetilde{M}_t = d(M_t Z_t^{-1})$ and since $dZ_t = -\theta_t Z_t dW_t$ (see Problem 4.2.2.2, page 196) and $d \tilde{W}_t = dW_t + \theta_t dt$, we have

$$d \left( M_t Z_t^{-1} \right) = \frac{1}{Z_t} dM_t - \frac{M_t}{Z_t^2} dZ_t + \frac{M_t}{Z_t^3} (dZ_t)^2 - \frac{1}{Z_t^2} dM_t dZ_t$$

$$= \frac{1}{Z_t} dM_t - \frac{M_t}{Z_t^2} (-\theta_t Z_t dW_t) + \frac{M_t}{Z_t^3} (\theta_t^2 Z_t^2 dt) - \frac{1}{Z_t^2} dM_t (-\theta_t Z_t dW_t)$$

$$= \frac{1}{Z_t} dM_t + \frac{\theta_t}{Z_t} dM_t dW_t + \frac{M_t \theta_t}{Z_t} (dW_t + \theta_t dt)$$

$$= \frac{1}{Z_t} dM_t + \frac{\theta_t}{Z_t} dM_t dW_t + \frac{M_t \theta_t}{Z_t} d\tilde{W}_t.$$ 

Since $M_t$ is a $\mathbb{P}$-martingale then, from the martingale representation theorem, there exists an adapted process $\gamma_t$, $0 \leq t \leq T$ such that

$$M_t = M_0 + \int_0^t \gamma_u dW_u, \quad 0 \leq t \leq T$$

or

$$dM_t = \gamma_t dW_t$$

and hence

$$d \widetilde{M}_t = \frac{\gamma_t}{Z_t} dW_t + \frac{\gamma_t \theta_t}{Z_t} dt + \frac{M_t \theta_t}{Z_t} d\tilde{W}_t$$

$$= \frac{\gamma_t}{Z_t} (d\widetilde{W}_t - \theta_t dt) + \frac{\gamma_t \theta_t}{Z_t} dt + \frac{M_t \theta_t}{Z_t} d\tilde{W}_t$$

$$= \left( \frac{\gamma_t}{Z_t} + \frac{M_t \theta_t}{Z_t} \right) d\tilde{W}_t$$

$$= \tilde{\gamma}_t d\tilde{W}_t.$$
4.2.3 Risk-Neutral Measure

where \( \tilde{\gamma}_t = \frac{\gamma_t + M_t \theta_t}{Z_t} \). Integrating on both sides, we therefore have

\[
\tilde{M}_t = \tilde{M}_0 + \int_0^t \tilde{\gamma}_u d\tilde{W}_u, \quad 0 \leq t \leq T.
\]

\( \square \)

4.2.3 Risk-Neutral Measure

1. Geometric Brownian Motion. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes

\[
dB_t = r_t B_t dt, \quad dS_t = \mu_t S_t dt + \sigma_t S_t dW_t
\]

where \( r_t \) is the risk-free rate, \( \mu_t \) is the stock price drift rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

From the following discounted stock price process

\[
X_t = e^{-\int_0^t r_u du} S_t
\]

show, using Girsanov’s theorem, that by changing the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), \( X_t \) is a \( \mathbb{Q} \)-martingale and by applying the martingale representation theorem show also that

\[
X_t = X_0 + \int_0^t \sigma_u X_u d\tilde{W}_u, \quad 0 \leq t \leq T
\]

where \( \tilde{W}_t = W_t + \int_0^t \lambda_u du \) is a \( \mathbb{Q} \)-standard Wiener process with \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \) defined as the market price of risk.

Finally, show that under the \( \mathbb{Q} \) measure the stock price follows

\[
dS_t = r_t S_t dt + \sigma_t S_t d\tilde{W}_t.
\]

Solution: By expanding \( dX_t \) using Taylor’s theorem and then applying Itô’s formula,

\[
dX_t = \frac{\partial X_t}{\partial t} dt + \frac{\partial X_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial S_t^2} dS_t^2 + \frac{1}{2} \frac{\partial^2 X_t}{\partial S_t^2} dS_t^2 + \ldots
\]

\[
= -r_t X_t dt + e^{-\int_0^t r_u du} dS_t
\]

\[
= (\mu_t - r_t) X_t dt + \sigma_t X_t d\tilde{W}_t
\]

\[
= \sigma_t X_t \left[ \frac{\mu_t - r_t}{\sigma_t} \right] dt + d\tilde{W}_t
\]

\[
= \sigma_t X_t d\tilde{W}_t
\]
where \( \tilde{W}_t = W_t + \int_0^t \lambda_u du \) such that \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \). From Girsanov’s theorem there exists an equivalent martingale measure or risk-neutral measure on the filtration \( \mathcal{F}_s \), \( 0 \leq s \leq t \) defined by the Radon–Nikodým derivative
\[
Z_s = e^{-\int_0^s \lambda_u du - \frac{1}{2} \int_0^s \lambda_u^2 du}
\]
so that \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process. Given that under the risk-neutral measure \( \mathbb{Q} \), the discounted stock price diffusion process
\[
dX_t = \sigma_t X_t d\tilde{W}_t
\]
has no \( dt \) term, then \( X_t \) is a \( \mathbb{Q} \)-martingale (see Problem 3.2.2.5, page 127). Thus, from the martingale representation theorem, there exists an adapted process \( \gamma_u \), \( 0 \leq u \leq T \) such that
\[
X_t = X_0 + \int_0^t \gamma_u d\tilde{W}_u, \quad 0 \leq t \leq T
\]
or
\[
X_t = X_0 + \int_0^t \sigma_u X_u d\tilde{W}_u, \quad 0 \leq t \leq T
\]
such that under \( \mathbb{Q} \), the process \( \int_0^t \sigma_u X_u d\tilde{W}_u \) is a \( \mathbb{Q} \)-martingale.

Finally, by substituting \( dW_t = d\tilde{W}_t + \lambda_t dt \) into \( dS_t = \mu_t S_t dt + \sigma_t S_t dW_t \), the stock price diffusion process under the risk-neutral measure becomes
\[
dS_t = r_t S_t dt + \sigma_t S_t d\tilde{W}_t.
\]

2. **Arithmetic Brownian Motion.** Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes
\[
dB_t = r_t B_t dt, \quad dS_t = \mu_t dt + \sigma_t dW_t
\]
where \( r_t \) is the risk-free rate, \( \mu_t \) is the stock price drift rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{W_t : 0 \leq t \leq T\} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

From the following discounted stock price process
\[
X_t = e^{-\int_0^t r_u du} S_t
\]
show, using Girsanov’s theorem, that by changing the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), \( X_t \) is a \( \mathbb{Q} \)-martingale and by applying the martingale representation theorem show also that
\[
X_t = X_0 + \int_0^t \sigma_u B_u^{-1} d\tilde{W}_u, \quad 0 \leq t \leq T
\]
where \( \tilde{W}_t = W_t + \int_0^t \lambda_u \, du \) is a \( \mathbb{Q} \)-standard Wiener process such that

\[
\lambda_t = \frac{\mu_t - r_S t}{\sigma_t}
\]

is defined as the market price of risk.

Finally, show that under the \( \mathbb{Q} \) measure the stock price follows

\[
dS_t = r_S S_t dt + \sigma_t \, d\tilde{W}_t.
\]

**Solution:** By expanding \( dX_t \) using Taylor’s theorem and then applying Itô’s formula,

\[
dX_t = \frac{\partial X_t}{\partial t} dt + \frac{\partial X_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial t^2} dt^2 + \frac{1}{2} \frac{\partial^2 X_t}{\partial S_t^2} dS_t^2 + \ldots
\]

\[
= -r_X X_t dt + e^{-\int_0^t \lambda_u \, du} dS_t
\]

\[
= e^{-\int_0^t \lambda_u \, du} (\mu_t - r_S t) dt + \sigma_t e^{-\int_0^t \lambda_u \, du} dW_t
\]

\[
= \sigma_t e^{-\int_0^t \lambda_u \, du} \left[ \left( \frac{\mu_t - r_S t}{\sigma_t} \right) dt + dW_t \right]
\]

\[
= \sigma_t B_t^{-1} d\tilde{W}_t
\]

where \( B_t = e^{\int_0^t \lambda_u \, du}, \tilde{W}_t = W_t + \int_0^t \lambda_u \, du \) such that \( \lambda_t = \frac{\mu_t - r_S t}{\sigma_t} \). From Girsanov’s theorem there exists an equivalent martingale measure or risk-neutral measure on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \) defined by the Radon–Nikodym derivative

\[
Z_s = e^{-\int_0^s \lambda_u \, du - \frac{1}{2} \int_0^s \lambda_u^2 \, du}
\]

so that \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process. Given that under the risk-neutral measure \( \mathbb{Q} \), the discounted stock price diffusion process

\[
dx_t = \sigma_t B_t^{-1} d\tilde{W}_t
\]

has no \( dt \) term, then \( X_t \) is a \( \mathbb{Q} \)-martingale (see Problem 3.2.2.5, page 127). Thus, from the martingale representation theorem, there exists an adapted process \( \gamma_u, 0 \leq u \leq T \) such that

\[
X_t = X_0 + \int_0^t \gamma_u \, d\tilde{W}_u, \quad 0 \leq t \leq T
\]

or

\[
X_t = X_0 + \int_0^t \sigma_u B_u^{-1} \, d\tilde{W}_u, \quad 0 \leq t \leq T
\]

such that under \( \mathbb{Q} \), the process \( \int_0^t \sigma_u B_u^{-1} \, d\tilde{W}_u \) is a \( \mathbb{Q} \)-martingale.
Finally, by substituting \( dW_t = d\tilde{W}_t + \lambda_t dt \) into \( dS_t = \mu_t dt + \sigma_t dW_t \), the stock price diffusion process under the risk-neutral measure becomes

\[
dS_t = r_S dt + \sigma_t d\tilde{W}_t.
\]

\[\square\]

3. *Discounted Portfolio.* Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes

\[
dB_t = r_t B_t dt, \quad dS_t = \mu_t S_t dt + \sigma_t S_t dW_t \]

where \( r_t \) is the risk-free rate, \( \mu_t \) is the stock price drift rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \), holding \( \phi_t \) shares of stock and \( \psi_t \) units being invested in a risk-free asset. From the following discounted portfolio value

\[
Y_t = e^{-\int_0^t r_u du} \Pi_t
\]

show, using Girsanov’s theorem, that by changing the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), then the discounted portfolio \( Y_t \) is a \( \mathbb{Q} \)-martingale.

**Solution:** At time \( t \), the portfolio \( \Pi_t \) is valued as

\[
\Pi_t = \phi_t S_t + \psi_t B_t
\]

and the differential of the portfolio is

\[
d\Pi_t = \phi_t dS_t + \psi_t r_t B_t dt
\]

\[
= \phi_t (\mu_t S_t dt + \sigma_t S_t dW_t) + \psi_t r_t B_t dt
\]

\[
= r_t \Pi_t dt + \phi_t (\mu_t - r_t) S_t dt + \phi_t \sigma_t S_t dW_t
\]

\[
= r_t \Pi_t dt + \phi_t \sigma_t (\lambda dt + dW_t)
\]

where \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \). By expanding \( dY_t \) using Taylor’s theorem and then applying Itô’s formula,

\[
dY_t = \frac{\partial Y_t}{\partial t} dt + \frac{\partial Y_t}{\partial \Pi_t} d\Pi_t + \frac{1}{2} \frac{\partial^2 Y_t}{\partial \Pi_t^2} (d\Pi_t)^2 + \ldots
\]

\[
= -r_t Y_t dt + e^{-\int_0^t r_u du} d\Pi_t
\]

\[
= -r_t e^{-\int_0^t r_u du} \Pi_t dt + e^{-\int_0^t r_u du} (r_t \Pi_t dt + \phi_t \sigma_t (\lambda dt + dW_t))
\]

\[
= \phi_t d \left( e^{-\int_0^t r_u du} S_t \right)
\]
4.2.3 Risk-Neutral Measure

where \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \) and from Problem 4.2.3.1 (page 221),
\[
\frac{d}{dt} \left( e^{-\int_0^t r_u du} S_t \right) = \sigma_t e^{-\int_0^t r_u du} S_t (\lambda_t dt + dW_t)
\]
\[
= \sigma_t e^{-\int_0^t r_u du} S_t d\tilde{W}_t
\]
where \( \tilde{W}_t = W_t + \int_0^t \lambda_u du \).

From Girsanov’s theorem, there exists an equivalent martingale measure or risk-neutral measure on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \) defined by the Radon–Nikodým derivative
\[
Z_s = e^{-\int_0^s \lambda_u du - \frac{1}{2} \int_0^s \lambda_u^2 du}
\]
so that \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process. Given that under the risk-neutral measure \( \mathbb{Q} \), the discounted stock price diffusion process
\[
dY_t = \phi_t \sigma_t e^{-\int_0^t r_u du} S_t d\tilde{W}_t
\]
does not have a \( dt \) term, then the discounted portfolio \( Y_t \) is a \( \mathbb{Q} \)-martingale (see Problem 3.2.2.5, page 127).

\( \square \)

4. Self-Financing Trading Strategy. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes
\[
\frac{dB_t}{B_t} = r_t dt, \quad \frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t
\]
where \( r_t \) is the risk-free rate, \( \mu_t \) is the stock price drift rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \) holding \( \phi_t \) shares of stock and \( \psi_t \) units being invested in a risk-free asset. Under the risk-neutral measure \( \mathbb{Q} \), the following discounted portfolio value
\[
Y_t = e^{-\int_0^t r_u du} \Pi_t
\]
is a \( \mathbb{Q} \)-martingale. Using the martingale representation theorem, show that the portfolio \((\phi_t, \psi_t)\) trading strategy has the values
\[
\phi_t = \gamma_t \frac{B_t}{S_t}, \quad \psi_t = \frac{\sigma_t \Pi_t - \gamma_t B_t}{\sigma_t S_t}, \quad 0 \leq t \leq T
\]
where \( \gamma_t, 0 \leq t \leq T \) is an adapted process.

**Solution:** Since \( Y_t \) is a \( \mathbb{Q} \)-martingale with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T \) (see Problem 4.2.3.3, page 224) and
\[
dY_t = \phi_t \sigma_t e^{-\int_0^t r_u du} S_t d\tilde{W}_t
\]
where \( \tilde{W}_t = W_t + \int_0^t \lambda_u \, du \), \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \) then, from the martingale representation theorem, there exists an adapted process \( \gamma_t, 0 \leq t \leq T \) such that

\[
Y_t = Y_0 + \int_0^t \gamma_u \, d\tilde{W}_u, \quad 0 \leq t \leq T.
\]

Taking integrals of

\[
dY_t = \phi_t \sigma_t e^{-\int_0^t r_s \, ds} S_t d\tilde{W}_t,
\]

\[
\int_0^t dY_v = \int_0^t \phi_v \sigma_v e^{-\int_0^v r_s \, ds} S_v d\tilde{W}_v
\]

\[
Y_t = Y_0 + \int_0^t \phi_v \sigma_v e^{-\int_0^v r_s \, ds} S_v d\tilde{W}_v.
\]

Therefore, for \( 0 \leq t \leq T \), we can set

\[
\gamma_t = \phi_t \sigma_t e^{-\int_0^t r_s \, ds} S_t
\]

or

\[
\phi_t = \frac{\gamma_t e^{\int_0^t r_s \, ds}}{\sigma_t S_t} = \frac{\gamma_t B_t}{\sigma_t S_t}
\]

where \( B_t = e^{\int_0^t r_s \, ds} \). Since \( \psi_t B_t = \Pi_t - \phi_t S_t \), therefore

\[
\psi_t = \frac{\sigma_t \Pi_t - \gamma_t B_t}{\sigma_t B_t}.
\]

\[\square\]

5. **Self-Financing Portfolio.** Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes

\[
dB_t = r_t B_t \, dt, \quad dS_t = \mu_t S_t \, dt + \sigma_t S_t \, dW_t
\]

where \( r_t \) is the risk-free rate, \( \mu_t \) is the stock price drift rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \) holding \( \phi_t \) shares of stock and \( \psi_t \) units being invested in a risk-free asset. For each of the following choices of \( \phi_t \):

(a) \( \phi_t = \alpha \), where \( \alpha \) is a constant
(b) \( \phi_t = S_t^n, n > 0 \)
(c) \( \phi_t = \int_0^t S_u^n \, du, n > 0 \)

find the corresponding \( \psi_t \) so that the trading strategy at time \( t \), \((\phi_t, \psi_t)\), is self-financing.

**Solution:** By definition, the portfolio at time \( t \), \( \Pi_t = \phi_t S_t + \psi_t B_t \) is self-financing if \( d\Pi_t = \phi_t dS_t + \psi_t dB_t \). Taking note that \( B_t = e^{\int_0^t r_s \, ds} \) satisfies \( dB_t = r_t B_t \, dt \) and from Itô’s lemma, we have \((dS_t)^2 = \sigma_t^2 S_t^2 \, dt\) and \((dS_t)^\nu = 0\) for \( \nu \geq 3 \). Then, for each of the following cases:
4.2.3 Risk-Neutral Measure

(a) If $\phi_t = \alpha$, we have $\Pi_t = \alpha S_t + \psi_t B_t$ and in differential form

$$d\Pi_t = \alpha dS_t + \psi_t dB_t$$

and in order for the portfolio to be self-financing (i.e., $d\Pi_t = \alpha dS_t + \psi_t dB_t$) we have $B_t d\psi_t = 0$ or $\psi_t = \beta$ for some constant $\beta$.

(b) If $\phi_t = S^n_t$, we have $\Pi_t = S^{n+1}_t + \psi_t B_t$ and in differential form

$$d\Pi_t = (n + 1)S^n_t dS_t + \frac{1}{2} n(n + 1) S^{n-1}_t (dS_t)^2 + \psi_t dB_t + B_t d\psi_t$$

and for the portfolio to be self-financing (i.e., $d\Pi_t = S^n_t dS_t + \psi_t dB_t$) we therefore set

$$nS^n_t dS_t + \frac{1}{2} n(n + 1) S^{n-1}_t (dS_t)^2 + B_t d\psi_t = 0$$

or

$$\psi_t = - \int_0^t \frac{S^n_u}{B_u} dS_u - \frac{1}{2} \int_0^t n(n + 1) \sigma^2_u S^{n+1}_u dS_u - \frac{1}{2} \int_0^t \int_0^t e^{-\int_0^v r_u dW_u} dS_v dW_v$$

(c) If $\phi_t = \int_0^t S^n_u du$, we have $\Pi_t = \left( \int_0^t S^n_u du \right) S_t + \psi_t B_t$ and in differential form

$$d\Pi_t = S^{n+1}_t dt + \left( \int_0^t S^n_u du \right) dS_t + \psi_t dB_t + B_t d\psi_t$$

and for the portfolio to be self-financing (i.e., $d\Pi_t = \left( \int_0^t S^n_u du \right) dS_t + \psi_t dB_t$) we require

$$S^{n+1}_t dt + B_t d\psi_t = 0$$

or

$$\psi_t = - \int_0^t \frac{S^{n+1}_u}{B_u} du = - \int_0^t e^{-\int_0^v r_u dW_u} S^{n+1}_u du.$$
At time $t$, we consider a trader who has a portfolio valued at $\Pi_t$ holding $\phi_t$ shares of stock and $\psi_t$ units being invested in a risk-free asset. From the following discounted portfolio value

$$Y_t = e^{-\int_0^t r_u du} \Pi_t$$

show, using Girsanov’s theorem, that by changing the measure $\mathbb{P}$ to an equivalent risk-neutral measure $\mathbb{Q}$, then the discounted portfolio $Y_t$ is a $\mathbb{Q}$-martingale.

Show also that under the $\mathbb{Q}$-measure the stock price follows

$$dS_t = (r_t - D_t)S_t dt + \sigma_t S_t d\tilde{W}_t$$

where $\tilde{W}_t = W_t + \int_0^t \lambda_u du$ is a $\mathbb{Q}$-standard Wiener process such that

$$\lambda_t = \frac{\mu_t - r_t}{\sigma_t}$$

is defined as the market price of risk.

**Solution:** At time $t$, the portfolio $\Pi_t$ is valued as

$$\Pi_t = \phi_t S_t + \psi_t B_t$$

and since the trader will receive $D_t S_t dt$ for every stock held and because the trader holds $\phi_t$ of the stock, then in differential form

$$d\Pi_t = \phi_t dS_t + \phi_t D_t S_t dt + \psi_t dB_t$$

$$= \phi_t [(\mu_t - D_t)S_t dt + \sigma_t S_t dW_t] + \phi_t D_t S_t dt + \psi_t r_t B_t dt$$

$$= r_t \Pi_t dt + \phi_t (\mu_t - r_t) S_t dt + \phi_t \sigma_t S_t dW_t$$

$$= r_t \Pi_t dt + \phi_t \sigma_t S_t (\lambda_t dt + dW_t)$$

where $\lambda_t = \frac{\mu_t - r_t}{\sigma_t}$. Following Problem 4.2.3.3 (page 224),

$$dY_t = d\left( e^{-\int_0^t r_u du} \Pi_t \right) = \phi \sigma_t e^{-\int_0^t r_u du} S_t d\tilde{W}_t$$

such that

$$\tilde{W}_t = W_t + \int_0^t \lambda_u du.$$
4.2.3 Risk-Neutral Measure

7. Stock Price with Continuous Dividend Yield (Arithmetic Brownian Motion). Consider an economy consisting of a risk-free asset and a dividend-paying stock price (risky asset). At time $t$, the risk-free asset $B_t$ and the stock price $S_t$ have the following diffusion processes

$$dB_t = r_t B_t dt, \quad dS_t = (\mu_t - D_t) dt + \sigma_t dW_t$$

such that $r_t$ is the risk-free rate, $\mu_t$ is the stock price drift rate, $D_t$ is the continuous dividend yield, $\sigma_t$ is the stock price volatility (which are all time dependent) and $\{W_t : 0 \leq t \leq T\}$ is a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

At time $t$, we consider a trader who has a portfolio valued at $\Pi_t$ holding $\phi_t$ shares of stock and $\psi_t$ units being invested in a risk-free asset. From the following discounted portfolio value

$$Y_t = e^{-\int_0^t r_u du} \Pi_t$$

show, using Girsanov’s theorem, that by changing the measure $\mathbb{P}$ to an equivalent risk-neutral measure $\mathbb{Q}$, then the discounted portfolio $Y_t$ is a $\mathbb{Q}$-martingale. Show also that under the $\mathbb{Q}$-measure the stock price follows

$$dS_t = (r_t - D_t) S_t dt + \sigma_t d\tilde{W}_t$$

where $\tilde{W}_t = W_t + \int_0^t \lambda_u du$ is a $\mathbb{Q}$-standard Wiener process such that

$$\lambda_t = \frac{\mu_t - r_t S_t}{\sigma_t}$$

is defined as the market price of risk.

**Solution:** At time $t$, the portfolio $\Pi_t$ is valued as

$$\Pi_t = \phi_t S_t + \psi_t B_t$$

and since the trader will receive $D_t dt$ for every stock held, then in differential form

$$d\Pi_t = \phi_t (dS_t + D_t dt) + \psi_t dB_t$$

$$= \phi_t (\mu_t dt + \sigma_t dW_t) + \psi_t r_t B_t dt$$

$$= r_t \Pi_t dt + \phi_t (\mu_t - r_t S_t) dt + \phi_t \sigma_t dW_t$$

$$= r_t \Pi_t dt + \phi_t \sigma_t (\lambda_t dt + dW_t)$$

where $\lambda_t = \frac{\mu_t - r_t S_t}{\sigma_t}$.

Following Problem 4.2.3.3 (page 224), the discounted portfolio becomes

$$dY_t = d \left( e^{-\int_0^t r_u du} \Pi_t \right) = \phi_t \sigma_t e^{-\int_0^t r_u du} \widehat{dW}_t$$

such that

$$\widehat{W}_t = W_t + \int_0^t \lambda_u du.$$
By applying Girsanov’s theorem to change the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), under which \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process, the discounted portfolio \( Y_t \) is a \( \mathbb{Q} \)-martingale.

By substituting \( dW_t = d\tilde{W}_t + \lambda_t dt \) into \( dS_t = (\mu_t - D_t)dt + \sigma_t dW_t \), the diffusion process under the risk-neutral measure becomes

\[
dS_t = (r_t - D_t)S_tdt + \sigma_t d\tilde{W}_t. \tag*{\Box}
\]

8. Commodity Price with Cost of Carry. Consider an economy consisting of a risk-free asset and a commodity price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the commodity price \( S_t \) have the following diffusion processes

\[
\begin{align*}
    dB_t &= r_t B_t dt, \\
    dS_t &= \mu_t S_t dt + C_t dt + \sigma_t S_t dW_t
\end{align*}
\]

such that \( r_t \) is the risk-free rate, \( \mu_t \) is the commodity price drift rate, \( C_t > 0 \) is the cost of carry for storage per unit time, \( \sigma_t \) is the commodity price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \) holding \( \phi_t \) units of commodity and \( \psi_t \) units being invested in a risk-free asset. From the following discounted portfolio value

\[
Y_t = e^{-\int_0^t r_u du} \Pi_t
\]

show, using Girsanov’s theorem, that by changing the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), then the discounted portfolio \( Y_t \) is a \( \mathbb{Q} \)-martingale.

Show also that under the \( \mathbb{Q} \)-measure the commodity price follows

\[
dS_t = r_t S_t dt + C_t dt + \sigma_t S_t d\tilde{W}_t
\]

where \( \tilde{W}_t = W_t + \int_0^t \lambda_u du \) is a \( \mathbb{Q} \)-standard Wiener process such that

\[
\lambda_t = \frac{\mu_t - r_t}{\sigma_t}
\]

is defined as the market price of risk.

**Solution:** At time \( t \), the portfolio \( \Pi_t \) is valued as

\[
\Pi_t = \phi_t S_t + \psi_t B_t
\]

and since the trader will pay \( C_t dt \) for storage, and because the trader holds \( \phi_t \) of the commodity, then in differential form

\[
d\Pi_t = \phi_t dS_t - \phi_t C_t dt + \psi_t r_t B_t dt
\]

\[
= \phi_t \left[ \mu_t S_t dt + C_t dt + \sigma_t S_t d\tilde{W}_t \right] - \phi_t C_t dt + \psi_t r_t B_t dt
\]
4.2.3 Risk-Neutral Measure

\[ r_t \Pi_t dt + \phi_t (\mu_t - r_t) S_t dt + \phi_t \sigma_t S_t dW_t \]

\[ = r_t \Pi_t dt + \phi_t \sigma_t S_t (\lambda_t dt + dW_t) \]

where \( \lambda_t = \frac{\mu_t - r_t}{\sigma_t} \). Following Problem 4.2.3.3 (page 224),

\[ dY_t = d \left( e^{-\int_0^t r_u du} \Pi_t \right) = \phi_t \sigma_t e^{-\int_0^t r_u du} S_t d\tilde{W}_t \]

such that

\[ \tilde{W}_t = W_t + \int_0^t \lambda_u du. \]

By applying Girsanov’s theorem to change the measure \( \mathbb{P} \) to an equivalent risk-neutral measure \( \mathbb{Q} \), under which \( \tilde{W}_t \) is a \( \mathbb{Q} \)-standard Wiener process, the discounted portfolio \( Y_t \) is a \( \mathbb{Q} \)-martingale.

By substituting \( dW_t = d\tilde{W}_t + \lambda_t dt \) into \( dS_t = \mu_t S_t dt + C_t dt + \sigma_t S_t dW_t \), the commodity price diffusion process under the risk-neutral measure becomes

\[ dS_t = r_t S_t dt + C_t dt + \sigma_t S_t d\tilde{W}_t. \]

\[ \Box \]

9. Pricing a Security Derivative. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes

\[ dB_t = r_t B_t dt, \quad dS_t = (\mu_t - D_t) S_t dt + \sigma_t S_t dW_t \]

where \( r_t \) is the risk-free rate, \( D_t \) is the continuous dividend yield, \( \mu_t \) is the stock price growth rate, \( \sigma_t \) is the stock price volatility (which are all time dependent) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \) holding \( \phi_t \) shares of stock and \( \psi_t \) units being invested in a risk-free asset. Let \( \Psi(S_T) \) represents the payoff of a derivative security at time \( T \), \( 0 \leq t \leq T \) such that \( \Psi(S_T) \) is \( \mathcal{F}_T \) measurable. Assuming the trader begins with an initial capital \( \Pi_0 \) and a trading strategy \( (\phi_t, \psi_t) \), \( 0 \leq t \leq T \) such that the portfolio value at time \( T \) is

\[ \Pi_T = \Psi(S_T) \]

almost surely.

show that under the risk-neutral measure \( \mathbb{Q} \),

\[ \Psi(S_t) = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^T r_u du} \Psi(S_T) \bigg| \mathcal{F}_t \right], \quad 0 \leq t \leq T. \]

Solution: At time \( t \), in order to calculate the price of a derivative security \( \Psi(S_t), 0 \leq t \leq T \) with payoff at time \( T \) given as \( \Psi(S_T) \), we note that because \( e^{-\int_0^t r_u du} \Pi_t \) is a \( \mathbb{Q} \)-martingale and \( \Pi_T = \Psi(S_T) \) almost surely,

\[ e^{-\int_0^T r_u du} \Pi_T = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^T r_u du} \Pi_T \bigg| \mathcal{F}_T \right] = \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^T r_u du} \Psi(S_T) \bigg| \mathcal{F}_T \right]. \]
Since the derivative security $\Psi$ and the portfolio $\Pi$ have identical values at $T$, and by assuming the trader begins with an initial capital $\Pi_0$, the value of the portfolio $\Pi_t = \Psi(S_t)$ for $0 \leq t \leq T$. Thus, we can write

$$
\Psi(S_t) = \mathbb{E}^Q \left[ e^{-\int_0^T r_u du} \Psi(S_T) \mid \mathcal{F}_t \right], \quad 0 \leq t \leq T.
$$

\[\Box\]

10. **First Fundamental Theorem of Asset Pricing.** Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time $t$, the risk-free asset $B_t$ and the stock price $S_t$ have the following diffusion processes

$$
dB_t = r_t B_t dt, \quad dS_t = (\mu_t - D_t)S_t dt + \sigma_t S_t dW_t,
$$

where $r_t$ is the risk-free rate, $\mu_t$ is the stock price growth rate, $D_t$ is the continuous dividend yield, $\sigma_t$ is the stock price volatility (which are all time dependent) and $\{W_t : 0 \leq t \leq T\}$ is a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

At time $t$, $0 \leq t \leq T$ we consider a trader who has a self-financing portfolio valued at $\Pi_t$ and we define an arbitrage strategy such that the following criteria are satisfied:

(i) $\Pi_0 = 0$
(ii) $\mathbb{P}(\Pi_T \geq 0) = 1$
(iii) $\mathbb{P}(\Pi_T > 0) > 0$.

Prove that if the trading strategy has a risk-neutral probability measure $\mathbb{Q}$, then it does not admit any arbitrage opportunities.

**Solution:** We prove this result via contradiction.

Suppose the trading strategy has a risk-neutral probability measure $\mathbb{Q}$ and it admits arbitrage opportunities. Since the discounted portfolio $e^{-\int_0^T r_u du} \Pi_t$ is a $\mathbb{Q}$-martingale, by letting $\Pi_0 = 0$, and from the optional stopping theorem,

$$
\mathbb{E}^Q \left( e^{-\int_0^T r_u du} \Pi_T \right) = \Pi_0 = 0.
$$

Since $\mathbb{P}(\Pi_T \geq 0) = 1$ and because $\mathbb{P}$ is equivalent to $\mathbb{Q}$, therefore $\mathbb{Q}(\Pi_T \geq 0) = 1$. In addition, since $\Pi_T \geq 0$ therefore $e^{-\int_0^T r_u du} \Pi_T \geq 0$. By definition,

$$
\mathbb{E}^Q \left( e^{-\int_0^T r_u du} \Pi_T \right) = \int_0^\infty \mathbb{Q} \left( e^{-\int_0^T r_u du} \Pi_T \geq x \right) dx = \int_0^\infty \mathbb{Q} \left( \Pi_T \geq xe^{\int_0^T r_u du} \right) dx = 0
$$

which implies

$$
\mathbb{Q}(\Pi_T \geq 0) = 0.
$$

Because $\mathbb{Q}$ is equivalent to $\mathbb{P}$, therefore

$$
\mathbb{P}(\Pi_T \geq 0) = 0
$$

which is a contradiction to the fact that there is an arbitrage opportunity. \[\Box\]
4.2.3 Risk-Neutral Measure

11. Second Fundamental Theorem of Asset Pricing. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time $t$, the risk-free asset $B_t$ and the stock price $S_t$ have the following diffusion processes

$$
\begin{align*}
    dB_t &= r_t B_t \, dt, \quad dS_t = (\mu_t - D_t) S_t \, dt + \sigma_t S_t \, dW_t 
\end{align*}
$$

where $r_t$ is the risk-free rate, $\mu_t$ is the stock price growth rate, $D_t$ is the continuous dividend yield, $\sigma_t$ is the stock price volatility (which are all time dependent) and $\{W_t : 0 \leq t \leq T\}$ is a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We consider a market model under the risk-neutral measure $\mathbb{Q}$ where at each time $t$, $0 \leq t \leq T$, the portfolio $\Pi_t$ is constructed with $\psi_t$ number of stocks and $\phi_t$ units of non-risky assets. Prove that if the market is complete (i.e., every derivative security can be hedged), then the risk-neutral measure is unique.

Solution: Suppose the market model has two risk-neutral measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$. Let $A \in \mathcal{F}_T$, and we consider the payoff of a derivative security

$$
\Psi(S_T) = \begin{cases} 
    e^\int_0^T r_u \, du & A \in \mathcal{F}_T \\
    0 & A \notin \mathcal{F}_T.
\end{cases}
$$

As the market model is complete, there is a portfolio value process $\Pi_t$, $0 \leq t \leq T$ with the same initial condition $\Pi_0$ that satisfies $\Pi_T = \Psi(S_T)$. Since both $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are risk-neutral measures, the discounted portfolio $e^{-\int_0^T r_u \, du} \Pi_t$ is a martingale under $\mathbb{Q}_1$ and $\mathbb{Q}_2$. From the optional stopping theorem we can write

$$
\begin{align*}
    \mathbb{Q}_1(A) &= \mathbb{E}^{\mathbb{Q}_1} \left[ e^{-\int_0^T r_u \, du} \Psi(S_T) \right] = \mathbb{E}^{\mathbb{Q}_1} \left( e^{-\int_0^T r_u \, du} \Pi_T \right) = \Pi_0 \\
    \mathbb{Q}_2(A) &= \mathbb{E}^{\mathbb{Q}_2} \left[ e^{-\int_0^T r_u \, du} \Psi(S_T) \right] = \mathbb{E}^{\mathbb{Q}_2} \left( e^{-\int_0^T r_u \, du} \Pi_T \right) = \Pi_0.
\end{align*}
$$

Therefore,

$$
\mathbb{Q}_1 = \mathbb{Q}_2
$$

since $A$ is an arbitrary set.

12. Change of Numéraire. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time $t$, the risk-free asset $B_t$ and the stock price $S_t$ have the following diffusion processes

$$
\begin{align*}
    dB_t &= r_t B_t \, dt, \quad dS_t = (\mu_t - D_t) S_t \, dt + \sigma_t S_t \, dW_t 
\end{align*}
$$

where $r_t$ is the risk-free rate, $\mu_t$ is the stock drift rate, $D_t$ is the continuous dividend yield, $\sigma_t$ is the stock price volatility (which are all time dependent) and $\{W_t : 0 \leq t \leq T\}$ is a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. 

Let $\Psi(S_T)$ represent the payoff of a derivative security at time $T$, $0 \leq t \leq T$ such that $\Psi(S_T)$ is $\mathcal{F}_T$ measurable and under the risk-neutral measure $Q$, let the numéraire $N_t^{(i)}$, $i = 1, 2$ be a strictly positive price process for a non-dividend-paying asset with the following diffusion process:

$$dN_t^{(i)} = r_t N_t^{(i)} dt + \nu_t^{(i)} N_t^{(i)} dB_t^{(i)}$$

where $\nu_t^{(i)}$ is the volatility and $B_t^{(i)}$, $0 \leq t \leq T$ is a $Q$-standard Wiener process. Show that for $0 \leq t \leq T$,

$$N_t^{(1)} E^{Q^{(1)}} \left( \frac{\Psi(S_T)}{N_T} \bigg| \mathcal{F}_t \right) = N_t^{(2)} E^{Q^{(2)}} \left( \frac{\Psi(S_T)}{N_T} \bigg| \mathcal{F}_t \right)$$

and

$$dQ^{(1)} \bigg|_{\mathcal{F}_t} = N_t^{(1)} \bigg/ N_t^{(2)}$$

where $N_t^{(i)}$ is a numéraire and $Q^{(i)}$ is the measure under which the stock price discounted by $N_t^{(i)}$ is a martingale.

**Solution:** By solving $dB_t = r_t B_t dt$ we have $B_t = e^{\int_0^t r_u du}$. For $i = 1, 2$ and given $B_t = e^{\int_0^t r_u du}$, from Problem 4.2.3.1 (page 221) we can deduce that

$$d(B_t^{-1} N_t^{(i)}) = B_t^{-1} N_t^{(i)} dB_t^{(i)}$$

is a $Q$-martingale and from Itô’s formula we have

$$d(\log (B_t^{-1} N_t)) = d(\log (B_t^{-1} N_t^{(i)})) + \frac{1}{2} \left( \frac{d \left( B_t^{-1} N_t^{(i)} \right)}{B_t^{-1} N_t^{(i)}} \right)^2 + \ldots$$

$$= \nu_t^{(i)} dW_t^{(i)} - \frac{1}{2} (\nu_t^{(i)})^2 dt.$$

Taking integrals,

$$\log \left( \frac{B_T^{-1} N_T}{B_0^{-1} N_0} \right) = \int_0^T \nu_u^{(i)} dW_u^{(i)} - \int_0^T \frac{1}{2} (\nu_u^{(i)})^2 du$$

or

$$B_T^{-1} N_T = N_0^{(i)} e^{\int_0^T \nu_u^{(i)} dW_u^{(i)} - \frac{1}{2} \int_0^T (\nu_u^{(i)})^2 du}$$

since $B_0 = 1$. Using Girsanov’s theorem to change from the $Q$ measure to an equivalent $Q^{(i)}$ measure, the Radon–Nikodým derivative is

$$\frac{dQ^{(i)}}{dQ} \bigg|_{\mathcal{F}_t} = e^{-\int_0^t (\nu_u) dW_u^{(i)} - \frac{1}{2} \int_0^t (\nu_u)^2 du} = \frac{N_t^{(i)}}{N_0^{(i)} B_t}$$

so that $W_t^{(i)} = W_t^{(i)} - \int_0^t \nu_u^{(i)} du$, $0 \leq t \leq T$ is a $Q^{(i)}$-standard Wiener process.
4.2.3 Risk-Neutral Measure

Under the risk-neutral measure $\mathbb{Q}$, the stock price follows

$$dS_t = (r_t - D_t)S_t dt + \sigma_t S_t d\tilde{W}_t$$

where $\tilde{W}_t = W_t + \int_0^t \left( \frac{\mu_u - r_u}{\sigma_u} \right) du$ is a $\mathbb{Q}$-standard Wiener process. By changing the $\mathbb{Q}$ measure to an equivalent $\mathbb{Q}^{(i)}$ measure, the discounted stock price

$$\{N_t^{(i)}\}^{-1} S_t$$

is a $\mathbb{Q}^{(i)}$-martingale. Thus, for a derivative payoff $\Psi(S_T)$,

$$\mathbb{E}^{\mathbb{Q}} (\Psi(S_T) | \mathcal{F}_t) = \left( \frac{d\mathbb{Q}^{(i)}}{d\mathbb{Q}} \right) \left| \mathcal{F}_t \right. \left. \left[ \Psi(S_T) \left( \frac{d\mathbb{Q}^{(i)}}{d\mathbb{Q}} \big| \mathcal{F}_T \right) \right]^{-1} \right| \mathcal{F}_t$$

$$= \frac{N_t^{(i)}}{N_T^{(i)}} \mathbb{E}^{\mathbb{Q}^{(i)}} \left( \Psi(S_T) \left| \mathcal{F}_t \right. \right)$$

$$= \frac{B_T N_t^{(i)}}{B_t} \mathbb{E}^{\mathbb{Q}^{(i)}} \left( \Psi(S_T) \left| \mathcal{F}_t \right. \right).$$

Thus,

$$N_t^{(1)} \mathbb{E}^{\mathbb{Q}^{(1)}} \left( \frac{\Psi(S_T)}{N_T^{(1)}} \big| \mathcal{F}_t \right) = N_t^{(2)} \mathbb{E}^{\mathbb{Q}^{(2)}} \left( \frac{\Psi(S_T)}{N_T^{(2)}} \big| \mathcal{F}_t \right).$$

To find $\frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}^{(2)}}$ on $\mathcal{F}_t$ we note that

$$\mathbb{E}^{\mathbb{Q}^{(1)}} \left( \frac{N_0^{(1)} \Psi(S_t)}{N_t^{(1)}} \big| \mathcal{F}_0 \right) = \mathbb{E}^{\mathbb{Q}^{(2)}} \left( \frac{N_0^{(2)} \Psi(S_t)}{N_t^{(2)}} \big| \mathcal{F}_0 \right).$$

By letting $X_t$, $0 \leq t \leq T$ be an $\mathcal{F}_t$ measurable random variable, from Problem 4.2.2.4 (page 198)

$$\mathbb{E}^{\mathbb{Q}^{(1)}} (X_t) = \mathbb{E}^{\mathbb{Q}^{(2)}} \left[ X_t \left( \frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}^{(2)}} \big| \mathcal{F}_t \right) \right].$$

We can therefore deduce that

$$\frac{N_0^{(1)} \Psi(S_t)}{N_t^{(1)}} \left( \frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}^{(2)}} \big| \mathcal{F}_t \right) = \frac{N_0^{(2)} \Psi(S_t)}{N_t^{(2)}}$$

or

$$\frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}^{(2)}} \big| \mathcal{F}_t = \frac{N_t^{(1)}}{N_0^{(1)}} / \frac{N_t^{(2)}}{N_0^{(2)}}.$$

N.B. An alternative method is to let $\frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}^{(2)}} = \frac{d\mathbb{Q}^{(1)}}{d\mathbb{Q}} / \frac{d\mathbb{Q}^{(2)}}{d\mathbb{Q}}$ and the result follows. □
13. **Black–Scholes Equation.** Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time $t$, the risk-free asset $B_t$ and the stock price $S_t$ have the following diffusion processes

$$dB_t = rB_t dt, \quad dS_t = \mu S_t dt + \sigma S_t dW_t$$

where $r$ is the risk-free rate, $\mu$ is the stock price drift rate, $\sigma$ is the stock price volatility (which are all constants) and $\{W_t : 0 \leq t \leq T\}$ is a $\mathbb{P}$-standard Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

At time $t$, we consider a trader who has a portfolio valued at $\Pi_t$, given as

$$\Pi_t = \phi_t S_t + \psi_t B_t$$

where he holds $\phi_t$ shares of stock and $\psi_t$ units being invested in a risk-free asset. Show that the portfolio $(\phi_t, \psi_t)$ is self-financing if and only if $\Pi_t$ satisfies the Black–Scholes equation

$$\frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} + r S_t \frac{\partial \Pi_t}{\partial S_t} - r \Pi_t = 0.$$

**Solution:** By applying Taylor’s theorem on $d\Pi_t$,

$$d\Pi_t = \frac{\partial \Pi_t}{\partial t} dt + \frac{\partial \Pi_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 \Pi_t}{\partial S_t^2} dS_t^2 + \ldots$$

and since $dS_t^2 = (\mu S_t dt + \sigma S_t dW_t)^2 = \sigma^2 S_t^2 dt$ such that $(dt)^\nu = 0$, $\nu \geq 2$, we have

$$d\Pi_t = \frac{\partial \Pi_t}{\partial t} dt + \frac{\partial \Pi_t}{\partial S_t} dS_t + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} dt$$

$$= \frac{\partial \Pi_t}{\partial S_t} dS_t + \left( \frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} \right) dt.$$ 

In contrast, the portfolio $(\phi_t, \psi_t)$ is self-financing if and only if

$$d\Pi_t = \phi_t dS_t + \psi_t dB_t = \phi_t dS_t + rB_t \psi_t dt.$$

By equating both of the equations we have

$$rB_t \psi_t = \frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} \quad \text{and} \quad \phi_t = \frac{\partial \Pi_t}{\partial S_t}$$

and substituting the above two equations into

$$\Pi_t = \phi_t S_t + \psi_t B_t$$

we have

$$\frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} + r S_t \frac{\partial \Pi_t}{\partial S_t} - r \Pi_t = 0.$$
4.2.3 Risk-Neutral Measure

N.B. Take note that the Black–Scholes equation does not contain the growth parameter 
\( \mu \), which means that the value of a self-financing portfolio is independent of how rapidly 
or slowly a stock grows. In essence, the only parameter from the stochastic differential 
equation \( dS_t = \mu S_t dt + \sigma S_t dW_t \) that affects the value of the portfolio is the volatility.

14. Black–Scholes Equation with Stock Paying Continuous Dividend Yield. Consider an economy consisting of a risk-free asset and a stock price (risky asset). At time \( t \), the risk-free asset \( B_t \) and the stock price \( S_t \) have the following diffusion processes

\[
\begin{align*}
\text{dB}_t &= rB_t \, dt, \\
\text{dS}_t &= (\mu - D)S_t \, dt + \sigma S_t \, dW_t
\end{align*}
\]

where \( r \) is the risk-free rate, \( \mu \) is the stock price drift rate, \( D \) is the continuous dividend yield, \( \sigma \) is the stock price volatility (which are all constants) and \( \{ W_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

At time \( t \), we consider a trader who has a portfolio valued at \( \Pi_t \), given as

\[
\Pi_t = \phi_t S_t + \psi_t B_t
\]

where he holds \( \phi_t \) shares of stock and \( \psi_t \) units being invested in a risk-free asset. Show that the portfolio \( (\phi_t, \psi_t) \) is self-financing if and only if \( \Pi_t \) satisfies the Black–Scholes equation with continuous dividend yield

\[
\frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} + (r - D)S_t \frac{\partial \Pi_t}{\partial S_t} - r \Pi_t = 0.
\]

Solution: By applying Taylor’s theorem on \( d\Pi_t \),

\[
d\Pi_t = \frac{\partial \Pi_t}{\partial t} dt + \frac{\partial \Pi_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 \Pi_t}{\partial S_t^2} dS_t^2 + \ldots
\]

and since \( dS_t^2 = ((\mu - D)S_t dt + \sigma S_t dW_t)^2 = \sigma^2 S_t^2 dt \) such that \( (dt)^v = 0 \), \( v \geq 2 \), we have

\[
d\Pi_t = \frac{\partial \Pi_t}{\partial t} dt + \frac{\partial \Pi_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 \Pi_t}{\partial S_t^2} dS_t^2
\]

\[
= \frac{\partial \Pi_t}{\partial S_t} dS_t + \left( \frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} \right) dt.
\]

Since the trader will receive \( DS_t dt \) for every stock held, the portfolio \( (\phi_t, \psi_t) \) is self-financing if and only if

\[
d\Pi_t = \phi_t dS_t + \psi_t dB_t + \phi_t DS_t \, dt = \phi_t dS_t + (rB_t \psi_t + \phi_t DS_t) \, dt.
\]

By equating both of the equations we have

\[
rB_t \psi_t + \phi_t DS_t = \frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} \quad \text{and} \quad \phi_t = \frac{\partial \Pi_t}{\partial S_t}.
\]
and substituting the above two equations into

\[ \Pi_t = \phi_t S_t + \psi_t B_t \]

we have

\[
\frac{\partial \Pi_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \Pi_t}{\partial S_t^2} + (r - D) S_t \frac{\partial \Pi_t}{\partial S_t} - r \Pi_t = 0.
\]

\[ \square \]

15. Foreign Exchange Rate under Domestic Risk-Neutral Measure. We consider a foreign exchange (FX) market which at time \( t \) consists of a foreign-to-domestic FX spot rate \( X_t \), a risk-free asset in domestic currency \( B^d_t \) and a risk-free asset in foreign currency \( B^f_t \). Here, \( X_t B^f_t \) denotes the foreign risk-free asset quoted in domestic currency. Assume that the evolution of these values has the following diffusion processes

\[
dX_t = \mu_t X_t dt + \sigma_t X_t dW^x_t, \quad dB^d_t = r^d_t B^d_t dt \quad \text{and} \quad dB^f_t = r^f_t B^f_t dt
\]

where \( \mu_t \) is the drift parameter, \( \sigma_t \) is the volatility parameter, \( r^d_t \) is the domestic risk-free rate and \( r^f_t \) is the foreign risk-free rate (which are all time dependent) and \( \{ W^x_t : 0 \leq t \leq T \} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

From the discounted foreign risk-free asset in domestic currency

\[ \tilde{X}_t = \frac{X_t B^f_t}{B^d_t} \]

show, using Girsanov’s theorem, that by changing the measure \( \mathbb{P} \) to an equivalent domestic risk-neutral measure \( \mathbb{Q}^d \) then \( \tilde{X}_t \) is a \( \mathbb{Q}^d \)-martingale.

Finally, show that under the \( \mathbb{Q}^d \) measure the FX rate follows

\[
dX_t = (r^d_t - r^f_t) X_t + \sigma_t X_t d\tilde{W}^d_t
\]

where \( \tilde{W}^d_t = W^x_t + \int_0^t \lambda_u du \) is a \( \mathbb{Q}^d \)-standard Wiener process with

\[ \lambda_t = \frac{\mu_t + r^f_t - r^d_t}{\sigma_t}. \]

**Solution:** By applying Taylor’s theorem on \( d\tilde{X}_t \),

\[
d\tilde{X}_t = \frac{\partial \tilde{X}_t}{\partial X_t} dX_t + \frac{\partial \tilde{X}_t}{\partial B^d_t} dB^d_t + \frac{\partial \tilde{X}_t}{\partial B^f_t} dB^f_t + \frac{1}{2} \frac{\partial^2 \tilde{X}_t}{\partial X_t^2} (dX_t)^2 + \frac{1}{2} \frac{\partial^2 \tilde{X}_t}{\partial B^d_t^2} (dB^d_t)^2 + \frac{1}{2} \frac{\partial^2 \tilde{X}_t}{\partial B^f_t^2} (dB^f_t)^2 \\
+ \frac{\partial^2 \tilde{X}_t}{\partial X_t \partial B^d_t} (dX_t)(dB^d_t) + \frac{\partial^2 \tilde{X}_t}{\partial X_t \partial B^f_t} (dX_t)(dB^f_t) + \frac{\partial^2 \tilde{X}_t}{\partial B^d_t \partial B^f_t} (dB^d_t)(dB^f_t) + \ldots
\]
and from Itô’s formula,

\[
d\tilde{X}_t = \frac{\partial \tilde{X}_t}{\partial X_t} dX_t + \frac{\partial \tilde{X}_t}{\partial B^d_t} dB^d_t + \frac{\partial \tilde{X}_t}{\partial B^f_t} dB^f_t + \frac{1}{2} \sigma_t^2 \frac{\partial^2 \tilde{X}_t}{\partial X_t^2} dt
\]

\[
d\tilde{X}_t \quad \frac{\partial \tilde{X}_t}{\partial X_t} \frac{\partial \tilde{X}_t}{\partial X_t} dX_t + \frac{\partial \tilde{X}_t}{\partial B^d_t} dB^d_t + \frac{\partial \tilde{X}_t}{\partial B^f_t} dB^f_t + \frac{1}{2} \sigma_t^2 \frac{\partial^2 \tilde{X}_t}{\partial X_t^2} dt
\]

\[
= \left( \frac{B^d_t}{B^f_t} \right) \left( \mu_t X_t dt + \sigma_t X_t dW^f_t \right) + \left( \frac{X_t}{B^d_t} B^f_t \right) \frac{r^f_t}{\sigma_t} dt - \left( \frac{X_t}{B^d_t} \frac{B^f_t}{B^d_t} \right) \frac{r^f_t}{\sigma_t} dB^d_t dt
\]

\[
= \left( \mu_t + \frac{r^f_t - r^d_t}{\sigma_t} \right) \tilde{X}_t dt + \sigma_t \tilde{X}_t dW^f_t
\]

\[
= \sigma_t \tilde{X}_t \left[ \frac{\mu_t + r^f_t - r^d_t}{\sigma_t} \right] dt + dW^f_t
\]

where \( \tilde{W}^d_t = W^d_t + \int_0^t \lambda_u du \) such that \( \lambda_t = \frac{\mu_t + r^f_t - r^d_t}{\sigma_t} \). From Girsanov’s theorem there exists an equivalent domestic risk-neutral measure \( \mathbb{Q}^d \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \) defined by the Radon–Nikodým derivative

\[
Z_s = e^{-\int_0^s \lambda_u du - \frac{1}{2} \int_0^s \sigma_u^2 dW^d_u}
\]

so that \( \tilde{W}^d_t \) is a \( \mathbb{Q}^d \)-standard Wiener process.

Therefore, by substituting

\[
dW^f_t = d\tilde{W}^d_t - \left( \frac{\mu_t + r^f_t - r^d_t}{\sigma_t} \right) dt
\]

into \( dX_t = \mu_t X_t dt + \sigma_t X_t dW^f_t \), under the domestic risk-neutral measure \( \mathbb{Q}^d \), the FX spot rate SDE is

\[
dX_t = \left( r^f_t - r^d_t \right) X_t + \sigma_t X_t d\tilde{W}^d_t.
\]

\[\square\]

16. **Foreign Exchange Rate under Foreign Risk-Neutral Measure.** We consider a foreign exchange (FX) market which at time \( t \) consists of a foreign-to-domestic FX spot rate \( X_t \), a risk-free asset in domestic currency \( B^d_t \) and a risk-free asset in foreign currency \( B^f_t \). Here, \( B^d_t / X_t \) denotes the domestic risk-free asset quoted in foreign currency. Assume that the evolution of these values has the following diffusion processes

\[
dX_t = \mu_t X_t dt + \sigma_t X_t dW^x_t, \quad dB^d_t = r^d_t B^d_t dt \quad \text{and} \quad dB^f_t = r^f_t B^f_t
d\tilde{X}_t = \frac{B^d_t}{X_t B^f_t}
\]

where \( \mu_t \) is the drift parameter, \( \sigma_t \) is the volatility parameter, \( r^d_t \) is the domestic risk-free rate and \( r^f_t \) is the foreign risk-free rate (which are all time dependent) and \( \{W^x_t : 0 \leq t \leq T\} \) is a \( \mathbb{P} \)-standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \).

From the discounted domestic risk-free asset in foreign currency

\[\tilde{X}_t = \frac{B^d_t}{X_t B^f_t}\]
show, using Girsanov’s theorem, that by changing the measure $\mathbb{P}$ to an equivalent foreign risk-neutral measure $\mathbb{Q}'$ then $\tilde{X}_t$ is a $\mathbb{Q}'$-martingale.

Finally, show that under the $\mathbb{Q}'$ measure the FX rate follows

$$dX_t = (r^d_t - r^f_t + \sigma_X^2)X_t + \sigma_X d\tilde{W}^f_t$$

where $\tilde{W}^f_t = W^x_t + \int_0^t \lambda_u \, du$ is a $\mathbb{Q}'$-standard Wiener process with $\lambda_t = \frac{\mu_t + r^f_t - r^d_t - \sigma_t^2}{\sigma_t}$.

**Solution:** By applying Taylor’s theorem on $d\tilde{X}_t$,

$$d\tilde{X}_t = \frac{\partial \tilde{X}_t}{\partial X_t} dX_t + \frac{\partial \tilde{X}_t}{\partial B^d_t} dB^d_t + \frac{\partial \tilde{X}_t}{\partial B^f_t} dB^f_t + \frac{1}{2} \frac{\partial^2 \tilde{X}_t}{\partial X_t^2} (dX_t)^2 + \frac{1}{2} \frac{\partial \tilde{X}_t}{\partial B^d_t} (dB^d_t)^2 + \frac{1}{2} \frac{\partial \tilde{X}_t}{\partial B^f_t} (dB^f_t)^2 + \ldots$$

and from Itô’s formula,

$$d\tilde{X}_t = \frac{\partial \tilde{X}_t}{\partial X_t} dX_t + \frac{\partial \tilde{X}_t}{\partial B^d_t} dB^d_t + \frac{\partial \tilde{X}_t}{\partial B^f_t} dB^f_t + \frac{1}{2} \sigma^2_X \frac{\partial^2 \tilde{X}_t}{\partial X_t^2} dt$$

$$= -\left( \frac{B^d_t}{X_t^2 B^f_t} \right) (\mu_t X_t dt + \sigma_X dW^x_t) - \left( \frac{B^d_t}{X_t (B^f_t)^2} \right) B^f_t r^f_t dt + \left( \frac{1}{X_t B^f_t} \right) r^f_t dB^f_t$$

$$+ \left( \frac{B^d_t}{X^2 B^f_t} \right) \sigma^2_X dt$$

$$= \left( r^d_t - r^f_t + \sigma^2_t - \mu_t \right) \tilde{X}_t dt - \sigma_t \tilde{X}_t dW^x_t$$

$$= -\sigma_t \tilde{X}_t \left[ \left( \mu_t + r^f_t - r^d_t - \sigma^2_t \right) dt + dW^x_t \right]$$

$$= -\sigma_t \tilde{X}_t d\tilde{W}^f_t$$

where $\tilde{W}^f_t = W^x_t + \int_0^t \lambda_u \, du$ such that $\lambda_t = \frac{\mu_t + r^f_t - r^d_t - \sigma^2_t}{\sigma_t}$. From Girsanov’s theorem there exists an equivalent foreign risk-neutral measure $\mathbb{Q}'$ on the filtration $\mathcal{F}_t$, $0 \leq s \leq t$ defined by the Radon–Nikodým derivative

$$Z_s = e^{-\int_0^s \lambda_u \, du - \frac{1}{2} \int_0^s \lambda_u^2 \, du}$$

so that $\tilde{W}^f_t$ is a $\mathbb{Q}'$-standard Wiener process.
4.2.3 Risk-Neutral Measure

Therefore, by substituting
\[ dW_x^t = \tilde{d}W_f^t - \left( \frac{\mu_t + r_f^t - i_t^f - \sigma_t^2}{\sigma_t} \right) dt \]
into \( dX_t = \mu_tX_t dt + \sigma_tX_t dW_x^t \), under the foreign risk-neutral measure \( Q_f \), the FX spot rate SDE is
\[ dX_t = (r_f^t - i_t^f + \sigma_t^2)X_t + \sigma_tX_t d\tilde{W}_x^t. \]
\[ \square \]

17. Foreign-Denominated Stock Price under Domestic Risk-Neutral Measure. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{W^s_t : 0 \leq t \leq T\} \) and \( \{W^x_t : 0 \leq t \leq T\} \) be \( \mathbb{P} \)-standard Wiener processes on the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \). Let \( S_t \) and \( X_t \) denote the stock price quoted in foreign currency and the foreign-to-domestic exchange rate, respectively, each having the following SDEs
\[
\begin{align*}
    dS_t &= (\mu_s - D_s)S_t dt + \sigma_s S_t dW^s_t \\
    dX_t &= \mu_x X_t dt + \sigma_x X_t dW^x_t \\
    dW^s_t \cdot dW^x_t &= \rho dt, \quad \rho \in (-1, 1)
\end{align*}
\]
where \( \mu_s, D_s \) and \( \sigma_s \) are the stock price drift, continuous dividend yield and volatility, respectively, whilst \( \mu_x \) and \( \sigma_x \) are the exchange rate drift and volatility, respectively. Here, we assume \( W^s_t \) and \( W^x_t \) are correlated with correlation coefficient \( \rho \in (-1, 1) \). In addition, let \( B^f_t \) and \( B^d_t \) be the risk-free assets in foreign and domestic currencies, respectively, having the following differential equations
\[
\begin{align*}
    dB^f_t &= r_f B^f_t dt \quad \text{and} \quad dB^d_t = r_d B^d_t dt
\end{align*}
\]
where \( r_f \) and \( r_d \) are the foreign and domestic risk-free rates.

Show that for the stock price denominated in domestic currency, \( X_t S_t \) follows the diffusion process
\[
\frac{d(X_t S_t)}{X_t S_t} = \left( \mu_s + \mu_x + \rho \sigma_s \sigma_x - D_s \right) dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2} dW^{xx}_t
\]
where \( W^{xx}_t = \frac{\sigma_s W^s_t + \sigma_x W^x_t}{\sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}} \) is a \( \mathbb{P} \)-standard Wiener process.

Using Girsanov’s theorem show that under the domestic risk-neutral measure \( Q^d \), the stock price denominated in domestic currency has the diffusion process
\[
\frac{d(X_t S_t)}{X_t S_t} = (r_d - D_s) dt + \sqrt{\sigma_x^2 + 2 \rho \sigma_s \sigma_x + \sigma_s^2} d\tilde{W}^{xx}_t
\]
where \( \tilde{W}^{xx}_t = W^{xx}_t + \left( \frac{\mu_s + \mu_x + \rho \sigma_s \sigma_x - r_d}{\sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}} \right) t \) is a \( Q^d \)-standard Wiener process.
Solution: From Problem 3.2.3.3 (page 158) we can easily show that for $X_tS_t$, its diffusion process is

$$\frac{d(X_tS_t)}{X_tS_t} = (\mu_s + \mu_x + \rho \sigma_s \sigma_x - D_s)dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}dW^{xs}_t$$

where $W^{xs}_t = \frac{\sigma_s W^s_t + \sigma_x W^x_t}{\sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}}$ is a $\mathbb{P}$-standard Wiener process.

At time $t$, we let the portfolio $\Pi_t$ be valued as

$$\Pi_t = \phi_t U_t + \psi_t B^d_t$$

where $\phi_t$ and $\psi_t$ are the units invested in $U_t = X_t S_t$ and the risk-free asset $B^d_t$, respectively. Taking note that the holder will receive $D_s U_t dt$ for every stock held, then

$$d\Pi_t = \phi_t (dU_t + D_s U_t dt) + \psi_t, r_d B^d_t dt$$

$$= \phi_t \left( (\mu_s + \mu_x + \rho \sigma_s \sigma_x)U_t dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}U_t dW^{xs}_t \right) + \psi_t, r_d B^d_t dt$$

$$= r_d \Pi_t dt + \phi_t \left( (\mu_s + \mu_x + \rho \sigma_s \sigma_x - r_d)U_t dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}U_t dW^{xs}_t \right)$$

$$= r_d \Pi_t dt + \phi_t \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2} U_t d\tilde{W}^{xs}_t$$

where $\tilde{W}^{xs}_t = \lambda t + W^{xs}_t$ such that $\lambda = \frac{\mu_s + \mu_x + \rho \sigma_s \sigma_x - r_d}{\sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}}$.

By applying Girsanov’s theorem to change the measure $\mathbb{P}$ to an equivalent risk-neutral measure $\mathbb{Q}^d$, under which $\tilde{W}^{xs}_t$ is a $\mathbb{Q}^d$-standard Wiener process, then the discounted portfolio $e^{-rt} \Pi_t$ is a $\mathbb{Q}^d$-martingale.

Finally, by substituting $dW^{xs}_t = d\tilde{W}^{xs}_t - \lambda dt$ into

$$\frac{d(X_tS_t)}{X_tS_t} = (\mu_s + \mu_x + \rho \sigma_s \sigma_x - D_s)dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}dW^{xs}_t$$

the stock price diffusion process under the risk-neutral measure $\mathbb{Q}^d$ becomes

$$\frac{d(X_tS_t)}{X_tS_t} = (r_d - D_s)dt + \sqrt{\sigma_s^2 + 2 \rho \sigma_s \sigma_x + \sigma_x^2}d\tilde{W}^{xs}_t.$$
5 Poisson Process

In mathematical finance the most important stochastic process is the Wiener process, which is used to model continuous asset price paths. The next important stochastic process is the Poisson process, used to model discontinuous random variables. Although time is continuous, the variable is discontinuous where it can represent a "jump" in an asset price (e.g., electricity prices or a credit risk event, such as describing default and rating migration scenarios). In this chapter we will discuss the Poisson process and some generalisations of it, such as the compound Poisson process and the Cox process (or doubly stochastic Poisson process) that are widely used in credit risk theory as well as in modelling energy prices.

5.1 INTRODUCTION

In this section, before we provide the definition of a Poisson process, we first define what a counting process is.

**Definition 5.1 (Counting Process)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A stochastic process \(\{N_t : t \geq 0\}\) where \(N_t\) denotes the number of events that have occurred in the interval \((0, t]\) is said to be a counting process if it has the following properties:

(a) \(N_t \geq 0\);
(b) \(N_t\) is an integer value;
(c) for \(s < t\), \(N_s \leq N_t\) and \(N_t - N_s\) is the number of events occurring in \((s, t]\).

Once we have defined a counting process we can now give a proper definition of a Poisson process. Basically, there are three definitions of a Poisson process (or homogeneous Poisson process) and they are all equivalent.

**Definition 5.2(a)** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A Poisson process (or homogeneous Poisson process) \(\{N_t : t \geq 0\}\) with intensity \(\lambda > 0\) is a counting process with the following properties:

(a) \(N_0 = 0\);
(b) \(N_t\) has independent and stationary increments;
(c) the sample paths \(N_t\) have jump discontinuities of unit magnitude such that for \(h > 0\)

\[
\mathbb{P}(N_{t+h} = i + j | N_t = i) = \begin{cases} 1 - \lambda h + o(h) & j = 0 \\ \lambda h + o(h) & j = 1 \\ o(h) & j > 1. \end{cases}
\]
5.1 INTRODUCTION

Definition 5.2(b) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A Poisson process (or homogeneous Poisson process) \(\{N_t : t \geq 0\}\) with intensity \(\lambda > 0\) is a counting process with the following properties:

(a) \(N_0 = 0\);
(b) \(N_t\) has independent and stationary increments;
(c) \(\mathbb{P}(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, k = 0, 1, 2, \ldots\)

Definition 5.2(c) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A Poisson process (or homogeneous Poisson process) \(\{N_t : t \geq 0\}\) with intensity \(\lambda > 0\) is a counting process with the following properties:

(a) \(N_0 = 0\);
(b) the inter-arrival times of events \(\tau_1, \tau_2, \ldots\) form a sequence of independent and identically distributed random variables where \(\tau_i \sim \text{Exp}(\lambda), i = 1, 2, \ldots\);
(c) the sample paths \(N_t\) have jump discontinuities of unit magnitude such that for \(h > 0\)

\[
\mathbb{P}(N_{t+h} = i + j | N_t = i) = \begin{cases} 
1 - \lambda h + o(h) & j = 0 \\
\lambda h + o(h) & j = 1 \\
o(h) & j > 1.
\end{cases}
\]

Like a standard Wiener process, the Poisson process has independent and stationary increments and so it is also a Markov process.

Theorem 5.3 (Markov Property) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. The Poisson process \(\{N_t : t \geq 0\}\) is a Markov process such that the conditional distribution of \(N_t\) given the filtration \(\mathcal{F}_s, s < t\) depends only on \(N_s\).

In addition, the Poisson process also has a strong Markov property.

Theorem 5.4 (Strong Markov Property) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. If \(\{N_t : t \geq 0\}\) is a Poisson process and given \(\mathcal{F}_t\) is the filtration up to time \(t\) then for \(s > 0, N_{t+s} - N_t \perp \mathcal{F}_s\).

By modifying the intensity parameter we can construct further definitions of a Poisson process and the following definitions (which are all equivalent) describe a non-homogeneous Poisson process with intensity \(\lambda_t\) being a deterministic function of time.

Definition 5.5(a) Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. A Poisson process (or non-homogeneous Poisson process) \(\{N_t : t \geq 0\}\) with intensity function \(\lambda_t : \mathbb{R}^+ \mapsto \mathbb{R}^+\) is a counting process with the following properties:

(a) \(N_0 = 0\);
(b) \(N_t\) has independent and stationary increments;
5.1 INTRODUCTION

(c) the sample paths $N_t$ have jump discontinuities of unit magnitude such that for $h > 0$

$$\mathbb{P}(N_{t+h} = i + j | N_t = i) = \begin{cases} 
1 - \lambda_i h + o(h) & j = 0 \\
\lambda_i h + o(h) & j = 1 \\
o(h) & j > 1.
\end{cases}$$

**Definition 5.5(b)** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A Poisson process (or non-homogeneous Poisson process) $\{N_t : t \geq 0\}$ with intensity function $\lambda_t : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a counting process with the following properties:

(a) $N_0 = 0$;
(b) $N_t$ has independent and stationary increments;
(c) $\mathbb{P}(N_t = k) = \frac{(\Lambda_t)^k e^{-\Lambda_t}}{k!}$, $k = 0, 1, 2, \ldots$ where $\Lambda_t = \int_0^t \lambda_u du$ is known as the intensity measure (hazard function or cumulative intensity).

It should be noted that if $\lambda_t$ is itself a random process, then $\{N_t : t \geq 0\}$ is called a doubly stochastic Poisson process, conditional Poisson process or Cox process. Given that $\lambda_t$ is a random process, the resulting stochastic process $\{\Lambda_t : t \geq 0\}$ defined as

$$\Lambda_t = \int_0^t \lambda_u du$$

is known as the hazard process. In credit risk modelling, due to the stochastic process of the intensity, the Cox process can be used to model the random occurrence of a default event, or even the number of contingent claims in actuarial models (claims that can be made only if one or more specified events occurs).

Given that the Poisson process $N_t$ with intensity $\lambda$ is an increasing counting process, therefore it is not a martingale. However, in its compensated form, $N_t - \lambda t$ is a martingale.

**Theorem 5.6** Let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$. By defining the compensated Poisson process $\hat{N}_t$ as

$$\hat{N}_t = N_t - \lambda t$$

then $\hat{N}_t$ is a martingale.

In the following we define an important generalisation of the Poisson process known as a compound Poisson process, where the jump size (or amplitude) can be modelled as a random variable.

**Definition 5.7 (Compound Poisson Process)** Let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, and let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables
with common mean $\mathbb{E}(X) = \mu$ and variance $\text{Var}(X) = \sigma^2$. Assume also that $X_1, X_2, \ldots$ are independent of $N_t$. The compound Poisson process $M_t$ is defined as

$$M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0.$$  

From the definition we can deduce that although the jumps in $N_t$ are always of one unit, the jumps in $M_t$ are of random size since $X_t$ is a random variable. The only similarity between the two processes is that the jumps in $N_t$ and $M_t$ occur at the same time. Furthermore, like the Poisson process, the compound Poisson process $M_t$ also has the independent and stationary increments property.

Another important generalisation of the Poisson process is to add a drift and a standard Wiener process term to generate a jump diffusion process of the form

$$dX_t = \mu(X_{t^-}, t) \, dt + \sigma(X_{t^-}, t) \, dW_t + J(X_{t^-}, t) \, dN_t$$

where $J(X_{t^-}, t)$ is a random variable denoting the size of the jump and

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda \, dt \\ 0 & \text{with probability } 1 - \lambda \, dt. \end{cases}$$

In the presence of jumps, $X_t$ is a càdlàg process, where it is continuous from the right

$$\lim_{u \uparrow t} X_u = \lim_{u \to t^-} X_u = X_t$$

which is also inclusive of any jumps at time $t$. To specify the value just before a jump we write $X_{t^-}$, which is the limit from the left:

$$\lim_{u \downarrow t} X_u = \lim_{u \to t^-} X_u = X_{t^-}.$$

Let $X_t$ be a càdlàg process with jump times $\tau_1 < \tau_2 < \ldots$. The integral of a function $\psi_t$ with respect to $X_t$ is defined by the pathwise Lebesgue–Stieltjes integral

$$\int_0^t \psi_s \, dX_s = \sum_{0 < s \leq t} \psi_s \Delta X_s = \sum_{0 < s \leq t} \psi_s (X_s - X_{s^-}) = \sum_{s=1}^{N_t} \psi_s$$

where the size of the jump, $\Delta X_s$ is denoted by

$$\Delta X_s = X_s - X_{s^-}$$

and given there is no jump at time zero we therefore set $\Delta X_0 = 0$.

**Theorem 5.8 (One-Dimensional Ito Formula for Jump Diffusion Process)** Consider a stochastic process $X_t$ satisfying the following SDE:

$$dX_t = \mu(X_{t^-}, t) \, dt + \sigma(X_{t^-}, t) \, dW_t + J(X_{t^-}, t) \, dN_t$$

or in integrated form

$$X_t = X_0 + \int_0^t \mu(X_{s^-}, s) \, ds + \int_0^t \sigma(X_{s^-}, s) \, dW_s + \int_0^t J(X_{s^-}, s) \, dN_s$$
5.1 INTRODUCTION

with \( \int_0^t \{ |\mu(X_r,s)| + \sigma(X_r,s)^2 \} \, ds < \infty \). Then, for a function \( f(X_r,t) \), the stochastic process \( Y_t = f(X_t,t) \) satisfies

\[
dY_t = \frac{\partial f}{\partial t}(X_{t-},t) \, dt + \frac{\partial f}{\partial X_t}(X_{t-},t) \, dX_t + \frac{1}{2!} \int_0^t \frac{\partial^2 f}{\partial X_t^2}(X_{t-},t)(dX_t)^2 + \frac{1}{3!} \int_0^t \frac{\partial^3 f}{\partial X_t^3}(X_{t-},t)(dX_t)^3 + \ldots
\]

\[
= \left[ \frac{\partial f}{\partial t}(X_{t-},t) + \mu(X_{t-},t) \frac{\partial f}{\partial X_t}(X_{t-},t) + \frac{1}{2} \sigma(X_{t-},t)^2 \frac{\partial^2 f}{\partial X_t^2}(X_{t-},t) \right] \, dt
\]

\[+ \sigma(X_{t-},t) \frac{\partial f}{\partial X_t}(X_{t-},t) \, dW_t + [f(X_t + J_{t-},t) - f(X_{t-},t)] \, dN_t
\]

where \( J_{t-} = J(X_{t-},t) \), \( (dX_t)^m, m = 1, 2, \ldots \) are expressed according to the rule

\[
(dt)^2 = (dt)^3 = \ldots = 0
\]

\[
(dW_t)^2 = dt, \quad (dW_t)^3 = (dW_t)^4 = \ldots = 0
\]

\[
(dN_t)^2 = (dN_t)^3 = \ldots = dN_t
\]

\[dW_t \, dt = dN_t \, dt = dN_t \, dW_t = 0.
\]

In integrated form

\[
Y_t = Y_0 + \int_0^t \frac{\partial f}{\partial t}(X_{s-},s) \, ds + \int_0^t \frac{\partial f}{\partial X_s}(X_{s-},s) \, dX_s
\]

\[+ \frac{1}{2!} \int_0^t \frac{\partial^2 f}{\partial X_s^2}(X_{s-},s)(dX_s)^2 + \frac{1}{3!} \int_0^t \frac{\partial^3 f}{\partial X_s^3}(X_{s-},s)(dX_s)^3 + \ldots
\]

\[= Y_0 + \int_0^t \left[ \frac{\partial f}{\partial t}(X_{s-},s) + \mu(X_{s-},s) \frac{\partial f}{\partial X_s}(X_{s-},s) + \frac{1}{2} \sigma(X_{s-},s)^2 \frac{\partial^2 f}{\partial X_s^2}(X_{s-},s) \right] \, ds
\]

\[+ \int_0^t \sigma(X_{s-},s) \frac{\partial f}{\partial X_s}(X_{s-},s) \, dW_s + \int_0^t [f(X_t + J_{s-},s) - f(X_{s-},s)] \, dN_s
\]

where

\[
\int_0^t [f(X_t + J_{s-},s) - f(X_{s-},s)] \, dN_s = \sum_{0 < s \leq t} [f(X_s,s) - f(X_{s-},s)] \, \Delta N_s
\]

\[= \sum_{i=1}^{N_t} [f(X_{s_i},s) - f(X_{s_{i-}},s)].
\]

The following is the multi-dimensional version of Itô’s formula for a jump diffusion process.

**Theorem 5.9 (Multi-Dimensional Itô Formula for Jump Diffusion Process)** Consider the stochastic processes \( X^{(i)}_t \), \( i = 1, 2, \ldots, n \) each satisfying the following SDE:

\[
dX^{(i)}_t = \mu^{(i)}(X^{(i)}_t,t) \, dt + \sigma^{(i)}(X^{(i)}_t,t) \, dW_t + J^{(i)}(X^{(i)}_t,t) \, dN_t
\]
or in integrated form

\[ X^{(i)}_t = X^{(i)}_0 + \int_0^t \mu(X^{(i)}_s, s) \, ds + \int_0^t \sigma(X^{(i)}_s, s) \, dW_t + \int_0^t J(X^{(i)}_s, s) \, dN_t \]

with \( \int_0^t \left\{ |\mu(X^{(i)}_s, s)| + \sigma(X^{(i)}_s, s)^2 \right\} \, ds < \infty \). Then for a function \( f(X^{(1)}_t, X^{(2)}_t, \ldots, X^{(n)}_t, t) \), the stochastic process \( Y_t = f(X^{(1)}_t, X^{(2)}_t, \ldots, X^{(n)}_t, t) \) satisfies

\[
dY_t = \left[ \sum_{i=1}^n \frac{\partial f}{\partial X^{(i)}_t}(X^{(1)}_t, X^{(2)}_t, \ldots, X^{(n)}_t, t) \right] dt + \sum_{i=1}^n \sum_{j=1}^n 1 \cdot \left( \sum_{i=1}^n \frac{\partial^2 f}{\partial X^{(i)}_t \partial X^{(j)}_t}(X^{(1)}_t, X^{(2)}_t, \ldots, X^{(n)}_t, t) \right) dx^{(i)}_t \, dx^{(j)}_t + \ldots
\]

where \( J^{(i)}_t = J(X^{(i)}_t, t) \), \((dX)_m^n, m = 1, 2, \ldots \) are expressed according to the rule

\( (dt)^2 = (dt)^3 = \ldots = 0 \)

\( (dW_t)^2 = dt \), \( (dW_t)^3 = (dW_t)^4 = \ldots = 0 \)

\( (dN_t)^2 = (dN_t)^3 = \ldots = dN_t \)

\( dW_t dt = dN_t dt = dN_t dW_t = 0 \).

In integrated form

\[
Y_t = Y_0 + \int_0^t \left[ \sum_{i=1}^n \frac{\partial f}{\partial X^{(i)}_t}(X^{(1)}_s, X^{(2)}_s, \ldots, X^{(n)}_s, s) \right] ds + \int_0^t \left[ \sum_{i=1}^n \sum_{j=1}^n 1 \cdot \left( \sum_{i=1}^n \frac{\partial^2 f}{\partial X^{(i)}_t \partial X^{(j)}_t}(X^{(1)}_s, X^{(2)}_s, \ldots, X^{(n)}_s, s) \right) dx^{(i)}_s \, dx^{(j)}_s \right] + \ldots
\]

\[
= Y_0 + \int_0^t \left[ \sum_{i=1}^n \frac{\partial f}{\partial X^{(i)}_s}(X^{(1)}_s, X^{(2)}_s, \ldots, X^{(n)}_s, s) + \sum_{i=1}^n \mu(X^{(i)}_s, s) \frac{\partial f}{\partial X^{(i)}_t}(X^{(1)}_s, X^{(2)}_s, \ldots, X^{(n)}_s, s) \right] dt
\]
5.1 INTRODUCTION

\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{2} \sigma(x^{(i)}, s) \sigma(x^{(j)}, s) \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) \right] ds \]

\[ + \int_0^t \left[ \sum_{i=1}^{n} \sigma(x^{(i)}, s) \frac{\partial f}{\partial x^{(i)}}(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) dW_s \right] \]

\[ + \int_0^t \left[ f(x^{(1)} + j^{(1)}, x^{(2)} + j^{(2)}, \ldots, x^{(n)} + j^{(n)}, s) - f(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) \right] dN_s \]

where

\[ \int_0^t \left[ f(x^{(1)} + j^{(1)}, x^{(2)} + j^{(2)}, \ldots, x^{(n)} + j^{(n)}, s) - f(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) \right] dN_s \]

\[ = \sum_{0 \leq j \leq t} \left[ f(x^{(1)} + j^{(1)}, x^{(2)} + j^{(2)}, \ldots, x^{(n)} + j^{(n)}, s) - f(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) \right] \Delta N_s \]

\[ = \sum_{i=1}^{N_t} \left[ f(x^{(1)} + j^{(1)}, x^{(2)} + j^{(2)}, \ldots, x^{(n)} + j^{(n)}, s) - f(x^{(1)}, x^{(2)}, \ldots, x^{(n)}, s) \right]. \]

By recalling that we can use Girsanov’s theorem to change the measure so that a Wiener process (standard Wiener process with drift) becomes a standard Wiener process, we can also change the measure for both Poisson and compound Poisson processes. For a Poisson process, the outcome of the change of measure affects the intensity whilst for a compound Poisson process, the change of measure affects both the intensity and the distribution of the jump amplitudes. In the following we provide general results for the change of measure for Poisson and compound Poisson processes.

**Theorem 5.10 (Girsanov’s Theorem for Poisson Process)** Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \). Let \( \eta > 0 \) and consider the Radon–Nikodym derivative process

\[ Z_t = e^{(\lambda-\eta)t} \left( \frac{1}{\lambda} \right)^N_t. \]

By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that

\[ \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = Z_t \]

then, under the \( \mathbb{Q} \) measure, \( N_t \sim \text{Poisson}(\eta t) \) with intensity \( \eta > 0 \).

**Theorem 5.11 (Girsanov’s Theorem for Compound Poisson Process)** Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \) and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables where each \( X_i \), \( i = 1, 2, \ldots \) has a probability density function \( f^\mathbb{P}(X_i) \) under the \( \mathbb{P} \) measure. Let \( X_1, X_2, \ldots \) also be independent of \( N_t \) and \( W_t \). From the definition of the compound Poisson process

\[ M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \]
we let $\eta > 0$ and consider the Radon–Nikodým derivative process

$$Z_t = \begin{cases} e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \frac{\eta Q(X_i)}{\lambda P(X_i)} & \text{if } X_i \text{ is discrete, } P(X_i) > 0 \\
\exp(\lambda - \eta \int_0^t \theta_u \, du - \frac{1}{2} \int_0^t \theta_u^2 \, du) & \text{if } X_i \text{ is continuous} \end{cases}$$

where $Q(X_i) (f^Q(X_i))$ is the probability mass (density) function of $X_i$, $i = 1, 2, \ldots$ under the $Q$ measure. For the case of continuous random variables $X_i$, $i = 1, 2, \ldots$ we also let $Q$ be absolutely continuous with respect to $P$ on $\mathcal{F}$. By changing the measure $P$ to measure $Q$ such that

$$\mathbb{E}^P \left( \frac{dQ}{dP} \mid \mathcal{F}_t \right) = \frac{dQ}{dP} \bigg|_{\mathcal{F}_t} = Z_t$$

then, under the $Q$ measure, $N_t$ is a Poisson process with intensity $\eta > 0$ and $X_i$, $i = 1, 2, \ldots$ is a sequence of independent and identically distributed random variables with probability mass (density) functions $Q(X_i) (f^Q(X_i))$, $i = 1, 2, \ldots$.

By augmenting the Wiener process to the Poisson and compound Poisson processes we have the following results.

**Theorem 5.12 (Girsanov’s Theorem for Poisson Process and Standard Wiener Process)** Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : 0 \leq t \leq T\}$ be a standard Wiener process defined on the probability space $(\Omega, \mathcal{F}, P)$ with respect to the filtration $\mathcal{F}_t$, $t \geq 0$. Suppose $\theta_t$, $0 \leq t \leq T$ is an adapted process and $\eta > 0$. We consider the Radon–Nikodým derivative process

$$Z_t = Z_t^{(1)} \cdot Z_t^{(2)}$$

such that

$$Z_t^{(1)} = e^{(\lambda - \eta)t} \left( \frac{\eta}{\lambda} \right)^{N_t}$$

and

$$Z_t^{(2)} = e^{-\int_0^t \theta_u \, du} - \frac{1}{2} \int_0^t \theta_u^2 \, du.$$

By changing the measure $P$ to measure $Q$ such that

$$\mathbb{E}^P \left( \frac{dQ}{dP} \mid \mathcal{F}_t \right) = \frac{dQ}{dP} \bigg|_{\mathcal{F}_t} = Z_t$$

then, under the $Q$ measure, $N_t$ is a Poisson process with intensity $\eta > 0$, the process $\tilde{W}_t = W_t + \int_0^t \theta_u \, du$ is a $Q$-standard Wiener process and $N_t \perp \tilde{W}_t$.

**Theorem 5.13 (Girsanov’s Theorem for Compound Poisson Process and Standard Wiener Process)** Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : 0 \leq t \leq T\}$ be a standard Wiener process defined on the probability space $(\Omega, \mathcal{F}, P)$.
5.2.1 Properties of Poisson Process

with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Suppose $\theta_t$, $0 \leq t \leq T$ is an adapted process, $\eta > 0$, $X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables where each $X_i, i = 1, 2, \ldots$ has a probability mass (density) function $\mathbb{P}(X_i) > 0$ ($f^\mathbb{P}(X_i) > 0$) under the $\mathbb{P}$ measure and assume also that the sequence $X_1, X_2, \ldots$ is independent of $N_t$. We consider the Radon–Nikodým derivative process

$$Z_t = Z_t^{(1)} \cdot Z_t^{(2)}$$

such that

$$Z_t^{(1)} = \begin{cases} 
  e^{(i - \eta)t} \frac{N_t \eta \mathbb{Q}(X_i)}{\lambda \mathbb{P}(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\
  e^{(i - \eta)t} \frac{N_t \eta f^\mathbb{Q}(X_i)}{\lambda f^\mathbb{P}(X_i)} & \text{if } X_i \text{ is continuous}
\end{cases}$$

and

$$Z_t^{(2)} = e^{-\int_0^t \theta_u du} \cdot \frac{1}{2} \int_0^t \theta_u^2 du$$

where $\mathbb{Q}(X_i)$ ($f^\mathbb{Q}(X_i)$) is the probability mass (density) function of $X_i, i = 1, 2, \ldots$ under the $\mathbb{Q}$ measure and $\mathbb{E}^\mathbb{P} \left( e^{\frac{1}{2} \int_0^t \theta_u^2 du} \right) < \infty$. For the case of continuous random variables $X_i, i = 1, 2, \ldots$ we also let $\mathbb{Q}$ be absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}$. By changing the measure $\mathbb{P}$ to measure $\mathbb{Q}$ such that

$$\mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t = Z_t$$

then, under the $\mathbb{Q}$ measure, the process $M_t = \sum_{i=1}^N X_i$ is a compound Poisson process with intensity $\eta > 0$ where $X_i, i = 1, 2, \ldots$ is a sequence of independent and identically distributed random variables with probability mass (density) functions $\mathbb{Q}(X_i)$ ($f^\mathbb{Q}(X_i)$), $i = 1, 2, \ldots$, the process $\tilde{W}_t = W_t + \int_0^t \theta_u du$ is a $\mathbb{Q}$ standard Wiener process and $M_t \perp \tilde{W}_t$.

### 5.2 PROBLEMS AND SOLUTIONS

5.2.1 Properties of Poisson Process

1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$. Show that $\text{Cov} (N_s, N_t) = \lambda \min \{s, t\}$ and deduce that the correlation coefficient of $N_s$ and $N_t$ is

$$\rho = \sqrt{\frac{\min \{s, t\}}{\max \{s, t\}}}.$$

**Solution:** Since $N_t \sim \text{Poisson}(\lambda t)$ with $\mathbb{E} (N_t) = \lambda t$, $\mathbb{E} (N_t^2) = \lambda t + \lambda^2 t^2$ and by definition

$$\text{Cov} (N_s, N_t) = \mathbb{E} (N_s N_t) - \mathbb{E} (N_s) \mathbb{E} (N_t).$$
For $s \leq t$, using the stationary and independent properties of a Poisson process,
\[
\mathbb{E} \left( N_s N_t \right) = \mathbb{E} \left( N_s (N_t - N_s) + N^2_s \right) = \mathbb{E} (N_s (N_t - N_s)) + \mathbb{E} (N^2_s),
\]
where $N_s \perp N_t - N_s$.
\[
= \lambda s \cdot \lambda (t - s) + \lambda s + \lambda^2 s^2 = \lambda^2 st + \lambda s
\]
which implies
\[
\text{Cov}(N_s, N_t) = \lambda^2 st + \lambda s - \lambda^2 st = \lambda s.
\]
Similarly, for $t \leq s$,
\[
\mathbb{E} \left( N_s N_t \right) = \mathbb{E} \left( N_t (N_s - N_t) + N^2_t \right) = \mathbb{E} (N_t (N_s - N_t)) + \mathbb{E} (N^2_t),
\]
where $N_t \perp N_s - N_t$.
\[
= \lambda t \cdot \lambda (s - t) + \lambda t + \lambda^2 t^2 = \lambda^2 st + \lambda t
\]
and hence
\[
\text{Cov}(N_s, N_t) = \lambda^2 st + \lambda t - \lambda^2 st = \lambda t.
\]
Thus, $\text{Cov}(N_s, N_t) = \lambda \min\{s, t\}$.

By definition, the correlation coefficient of $N_s$ and $N_t$ is defined as
\[
\rho = \frac{\text{Cov}(N_s, N_t)}{\sqrt{\text{Var}(N_s)} \sqrt{\text{Var}(N_t)}} = \frac{\lambda \min\{s, t\}}{\lambda \sqrt{st}} = \frac{\min\{s, t\}}{\sqrt{st}}.
\]
For $s \leq t$,
\[
\rho = \frac{s}{\sqrt{st}} = \sqrt{\frac{s}{t}}
\]
whilst for $s > t$,
\[
\rho = \frac{t}{\sqrt{st}} = \sqrt{\frac{t}{s}}.
\]
Therefore, $\rho = \sqrt{\frac{\min\{s, t\}}{\max\{s, t\}}}$.
5.2.1 Properties of Poisson Process

2. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{N_t : t \geq 0\}\) be a Poisson process with intensity \(\lambda > 0\) satisfying the following conditions for small \(h > 0\) and \(k \in \mathbb{N}\):
   
   (a) \(N_0 = 0\) and \(N_t\) is a non-decreasing counting process;
   
   \[
   \mathbb{P}(N_t = i) = \begin{cases} 
   1 - \lambda h + o(h) & j = 0 \\
   \lambda h + o(h) & j = 1 \\
   o(h) & j > 1
   \end{cases}
   \]

   (b) \(\mathbb{P}(N_{t+h} = i + j | N_t = i) = \frac{\lambda^j}{j!} e^{-\lambda t} h^j \mathbb{P}(N_t = k)\)

   (c) \(N_t\) has independent stationary increments.

   By setting \(p_k(t) = \mathbb{P}(N_t = k)\), show that

   \[
   p_k'(t) = \begin{cases} 
   -\lambda p_0(t) & k = 0 \\
   \lambda p_{k-1}(t) - \lambda p_k(t) & k \neq 0
   \end{cases}
   \]

   with boundary condition

   \[
   p_k(0) = \begin{cases} 
   1 & k = 0 \\
   0 & k \neq 0
   \end{cases}
   \]

   By solving the differential-difference equation using mathematical induction or the moment generating function, show that

   \[
   \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots
   \]

   **Solution:** By definition, for \(k \neq 0\),

   \[
   \mathbb{P}(N_{t+h} = k) = \sum_j \mathbb{P}(N_{t+h} = k | N_t = j) \mathbb{P}(N_t = j)
   \]

   \[
   = \mathbb{P}(N_{t+h} = k | N_t = k-1) \mathbb{P}(N_t = k-1) + \mathbb{P}(N_{t+h} = k | N_t = k) \mathbb{P}(N_t = k) + o(h)
   \]

   \[
   = (\lambda h + o(h)) \mathbb{P}(N_t = k-1) + (1 - \lambda h + o(h)) \mathbb{P}(N_t = k).
   \]

   By setting \(p_k(t + h) = \mathbb{P}(N_{t+h} = k)\), \(p_{k-1}(t) = \mathbb{P}(N_t = k-1)\) and \(p_k(t) = \mathbb{P}(N_t = k)\), therefore for \(k \neq 0\),

   \[
   p_k'(t + h) = \lambda h p_{k-1}(t) + (1 - \lambda h) p_k(t) + o(h)
   \]

   \[
   \lim_{h \to 0} \frac{p_k(t + h) - p_k(t)}{h} = \lambda p_{k-1}(t) - \lambda p_k(t) + \lim_{h \to 0} \frac{o(h)}{h}
   \]

   and we will have

   \[
   p_k'(t) = \lambda p_{k-1}(t) - \lambda p_k(t), \quad k \neq 0
   \]

   For the case when \(k = 0\),

   \[
   \mathbb{P}(N_{t+h} = 0) = \mathbb{P}(N_{t+h} = 0 | N_t = 0) \mathbb{P}(N_t = 0) + o(h)
   \]

   \[
   = (1 - \lambda h + o(h)) \mathbb{P}(N_t = 0) + o(h)
   \]

   or

   \[
   p_0(t + h) = (1 - \lambda h) p_0(t) + o(h)
   \]
where \( p_0(t + h) = \mathbb{P}(N_{t+h} = 0) \) and \( p_0(t) = \mathbb{P}(N_t = 0) \). By subtracting \( p_0(t) \) from both sides of the equation, dividing by \( h \) and letting \( h \to 0 \), we will have

\[
p'_0(t) = -\lambda p_0(t).
\]

Knowing that \( \mathbb{P}(N_0 = 0) = 1 \), hence

\[
p'_k(t) = \begin{cases} 
-\lambda p_0(t) & k = 0 \\
\lambda p_{k-1}(t) - \lambda p_k(t) & k \neq 0
\end{cases}
\]

with boundary condition

\[
p_k(0) = \begin{cases} 
1 & k = 0 \\
0 & k \neq 0
\end{cases}
\]

The following are two methods to solve the differential-difference equation.

**Method I: Mathematical Induction.** Let \( k = 0 \). To solve

\[ p'_0(t) + \lambda p_0(t) = 0 \]

we let the integrating factor \( I = e^{\lambda t} \). By multiplying the integrating factor on both sides of the equation and taking note that \( p_0(0) = \mathbb{P}(N_0 = 0) = 1 \), we have

\[ \mathbb{P}(N_t = 0) = e^{-\lambda t}. \]

Therefore, the result is true for \( k = 0 \).

Assume the result is true for \( k = j, j = 0, 1, 2, \ldots \) such that

\[ \mathbb{P}(N_t = j) = \frac{(\lambda t)^j e^{-\lambda t}}{j!}. \]

For \( k = j + 1 \) the differential-difference equation is

\[ p'_{j+1}(t) + \lambda p_{j+1}(t) = \frac{\lambda (\lambda t)^j e^{-\lambda t}}{j!}. \]

Setting the integrating factor \( I = e^{\lambda t} \) and multiplying it by both sides of the differential equation,

\[ \frac{d}{dt}(e^{\lambda t}p_{j+1}(t)) = \frac{\lambda^{j+1}p_j}{j!} \]

or

\[ e^{\lambda t}p_{j+1}(t) = \frac{(\lambda t)^{j+1} e^{-\lambda t}}{(j+1)!} + C \]

where \( C \) is a constant. Knowing that \( p_{j+1}(0) = \mathbb{P}(N_0 = j + 1) = 0 \) we have \( C = 0 \) and hence

\[ \mathbb{P}(N_t = j + 1) = \frac{(\lambda t)^{j+1} e^{-\lambda t}}{(j+1)!}. \]

Therefore, the result is also true for \( k = j + 1 \). Thus, using mathematical induction we have shown

\[ \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots \]
5.2.1 Properties of Poisson Process

**Method II: Moment Generating Function.** Let

\[ M_{N_t}(t, u) = \mathbb{E}(e^{uN_t}) = \sum_{k=0}^{\infty} e^{uk} \mathbb{P}(N_t = k) = \sum_{k=0}^{\infty} e^{uk} p_k(t) \]

be defined as the moment generating function of the Poisson process. Taking the first partial derivative with respect to \( t \),

\[
\frac{\partial M_{N_t}(t, u)}{\partial t} = \sum_{k=0}^{\infty} e^{uk} p'_k(t)
\]

\[
= \lambda \sum_{k=1}^{\infty} e^{uk} p_{k-1}(t) - \lambda \sum_{k=0}^{\infty} e^{uk} p_k(t)
\]

\[
= \lambda e^{ut} \sum_{k=1}^{\infty} e^{uk} p_{k-1}(t) - \lambda \sum_{k=0}^{\infty} e^{uk} p_k(t)
\]

or

\[
\frac{\partial M_{N_t}(t, u)}{\partial t} = \lambda(e^u - 1)M_{N_t}(t, u)
\]

with boundary condition \( M_{N_t}(0, u) = 1 \).

By setting the integrating factor \( I = e^{-\lambda(e^u-1)t} \) and multiplying it by the first-order differential equation, we have

\[
\frac{d}{dt} \left( M_{N_t}(t, u)e^{-\lambda(e^u-1)t} \right) = 0 \quad \text{or} \quad M_{N_t}(t, u)e^{-\lambda(e^u-1)t} = C
\]

where \( C \) is a constant. Since \( M_{N_t}(0, u) = 1 \) we have \( C = 1 \) and hence

\[ M_{N_t}(t, u) = e^{\lambda(e^u-1)t} \]

By definition, \( M_{N_t}(t, u) = \sum_{k=0}^{\infty} e^{uk} \mathbb{P}(N_t = k) \) and using Taylor’s series expansion we have

\[ M_{N_t}(t, u) = e^{\lambda(e^u-1)t} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda te^u)^k}{k!} = \sum_{k=0}^{\infty} e^{uk} \left( \frac{(\lambda t)^k}{k!} e^{-\lambda t} \right) \]

and hence

\[ \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots \]

\[ \square \]

3. **Pure Birth Process.** Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space. If \( \{N_t : t \geq 0\} \) is a Poisson process with intensity \( \lambda > 0 \) then for small \( h > 0 \) and \( k \in \mathbb{N} \), show that it satisfies the following property:

\[
\mathbb{P}(N_{t+h} = k + j|N_t = k) = \begin{cases} 
1 - \lambda h + o(h) & j = 0 \\
\lambda h + o(h) & j = 1 \\
o(h) & j > 1.
\end{cases}
\]
Solution: We first consider \( j = 0 \) where
\[
\mathbb{P}(N_{t+h} = k | N_t = k) = \frac{\mathbb{P}(N_{t+h} = k, N_t = k)}{\mathbb{P}(N_t = k)}
\]
\[
= \frac{\mathbb{P}(\text{zero arrival between } (t, t+h), N_t = k)}{\mathbb{P}(N_t = k)}.
\]
Since the event \( \{ N_t = k \} \) relates to arrival during the time interval \([0, t]\) and the event \( \{ \text{zero arrival between } (t, t+h) \} \) relates to arrivals after time \( t \), both of the events are independent and because \( N_t \) has stationary increments,
\[
\mathbb{P}(N_{t+h} = k | N_t = k) = \frac{\mathbb{P}(\text{zero arrival between } (t, t+h)) \mathbb{P}(N_t = k)}{\mathbb{P}(N_t = k)} = \mathbb{P}(N_{t+h} - N_t = 0) = e^{-\lambda h}
\]
\[
= 1 - \lambda h + \frac{(\lambda h)^2}{2} - \frac{(\lambda h)^3}{3!} + \ldots
= 1 - \lambda h + o(h).
\]
Similarly,
\[
\mathbb{P}(N_{t+h} = k+1 | N_t = k) = \mathbb{P}(1 \text{ arrival between } (t, t+h) |
\]
\[
= \mathbb{P}(N_{t+h} - N_t = 1) = \lambda he^{-\lambda h}
\]
\[
= \lambda h \left( 1 - \lambda h + \frac{(\lambda h)^2}{2} - \frac{(\lambda h)^3}{3!} + \ldots \right) = \lambda h + o(h)
\]
and finally
\[
\mathbb{P}(N_{t+h} > k + 1 | N_t = k) = 1 - \mathbb{P}(N_{t+h} = k | N_{t+h} = k) - \mathbb{P}(N_{t+h} = k+1 | N_t = k)
\]
\[
= 1 - (1 - \lambda h + o(h)) - \lambda h + o(h) = o(h).
\]
\[\Box\]

4. Arrival Time Distribution. Let the inter-arrival times of “jump” events \( \tau_1, \tau_2, \ldots \) be a sequence of independent and identically distributed random variables where each \( \tau_i \sim \text{Exp}(\lambda), \lambda > 0, i = 1, 2, \ldots \) has a probability density function \( f_{\tau}(t) = \lambda e^{-\lambda t}, t \geq 0 \). By defining the arrival time of the \( n \)-th jump event as
\[
T_n = \sum_{i=1}^{n} \tau_i
\]
5.2.1 Properties of Poisson Process

where \( \tau_i = T_i - T_{i-1} \), show that the arrival time \( T_n \), \( n \geq 1 \) follows a gamma distribution, \( T_n \sim \text{Gamma}(n, \lambda) \) with probability density function given as

\[
f_{T_n}(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda e^{-\lambda t}, \quad t \geq 0.
\]

Let \( N_t \) be the number of jumps that occur at or before time \( t \). Explain why for \( k \geq 1 \),

\[
P(N_t \geq k) = P(T_k \leq t)
\]

and for \( k = 0 \),

\[
P(N_t = 0) = P(T_1 > t).
\]

Using the above results show that \( N_t \) is a Poisson process with intensity \( \lambda \) having the probability mass function

\[
P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots
\]

**Solution:** The first part of the proof is analogous to Problem 1.2.2.5 (page 16) and we shall omit it. As for the remaining part of the proof we note that for \( k \geq 1 \), \( N_t \geq k \) if and only if there are at least \( k \) jump events by time \( t \), which implies that \( T_k \) (the time of the \( k \)-th jump) is less than or equal to \( t \). Thus,

\[
P(N_t \geq k) = P(T_k \leq t), \quad k \geq 1.
\]

On the contrary, for \( k = 0 \), \( N_t = 0 \) if and only if there are zero jumps by time \( t \), which implies \( T_1 \) (the time of the first jump) occurs after time \( t \). Therefore,

\[
P(N_t = 0) = P(T_1 > t).
\]

For \( k = 0 \),

\[
P(N_t = 0) = P(T_1 > t) = P(\tau_1 > t) = \int_t^\infty \lambda e^{-\lambda t} \, ds = e^{-\lambda t}.
\]

For \( k \geq 1 \),

\[
P(N_t = k) = P(N_t \geq k) - P(N_t \geq k + 1).
\]

By solving

\[
P(N_t \geq k + 1) = P(T_{k+1} \leq t) = \int_0^t \frac{(\lambda s)^k}{k!} \lambda e^{-\lambda s} \, ds
\]

and using integration by parts

\[
\int_a^b u(x) \frac{d}{dx} v(x) \, dx = u(x)v(x) \big|_a^b - \int_a^b v(x) \frac{d}{dx} u(x) \, dx,
\]

we set

\[
u = \frac{(\lambda s)^k}{k!} \quad \Rightarrow \frac{du}{ds} = \frac{\lambda^k s^{k-1}}{(k-1)!}
\]
Thus,
\[ \frac{dv}{ds} = \lambda e^{-\lambda s} \Rightarrow v = -e^{-\lambda s}. \]

Thus, \[ \mathbb{P}(N_t \geq k+1) = - \left[ \frac{(\lambda s)^k}{k!} e^{-\lambda s} \right]_{s=0}^{s=t} + \int_0^t \frac{(\lambda s)^{k-1}}{(k-1)!} \lambda e^{-\lambda s} ds = - \frac{(\lambda t)^k}{k!} e^{-\lambda t} + \mathbb{P}(N_t \geq k). \]

Therefore, for \( k \geq 1 \)
\[ \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots \]

Since \( 0! = 1 \), in general we can express the probability mass function of a Poisson process with intensity \( \lambda > 0 \) as
\[ \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \ldots \]

5. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \). If \( \tau_1, \tau_2, \ldots \) is a sequence of inter-arrival times, show that the random variables are mutually independent and each follows an exponential distribution with parameter \( \lambda \).

**Solution:** We prove this result via mathematical induction.

Since \( \mathbb{P}(\tau_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t} \)
the density of \( \tau_1 \) is
\[ f_{\tau_1}(t) = \frac{d}{dt} \mathbb{P}(\tau_1 < t) = \frac{d}{dt} (1 - e^{-\lambda t}) = \lambda e^{-\lambda t} \]
and hence \( \tau_1 \sim \text{Exp}(\lambda) \).

By conditioning on \( \tau_1 \) and since \( \tau_1 < \tau_2 \),
\[ \mathbb{P}(\tau_2 > t|\tau_1 = t_1) = \mathbb{P}(\text{zero arrival in } (t_1, t_1 + t])|\tau_1 = t_1). \]

Since the event \( \{\tau_1 = t_1\} \) is only related to arrivals during the time interval \([0, t_1]\) and the event \( \{\text{zero arrival in } (t_1, t_1 + t]\) relates to arrivals during the time interval \((t_1, t_1 + t]\), these two events are independent. Thus, we have
\[ \mathbb{P}(\tau_2 > t|\tau_1 = t_1) = \mathbb{P}(\text{zero arrival in } (t_1, t_1 + t]] = \mathbb{P}(\tau_2 > t) = e^{-\lambda t} \]
with \( \tau_2 \perp \tau_1 \) and \( \tau_2 \sim \text{Exp}(\lambda) \).

We assume that for \( n > 1 \), \( \tau_n \perp \tau_i \) and each \( \tau_i \sim \text{Exp}(\lambda) \), \( i = 1, 2, \ldots, n-1 \) such that
\[ \mathbb{P}(\tau_n > t|\tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_{n-1} = t_{n-1}) = \mathbb{P}(\text{zero arrival in } (T_{n-1}, T_{n-1} + t]) \]
\[ = e^{-\lambda t} \]
where \( T_{n-1} = \sum_{i=1}^{n-1} t_i \).
5.2.1 Properties of Poisson Process

Conditional on \( \tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_n = t_n \), such that \( \tau_1 < \tau_2 < \ldots < \tau_n < \tau_{n+1} \) we have

\[
P(\tau_{n+1} > t | \tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_n = t_n) = P(\text{zero arrival in } (T_n, T_n + t]) | \tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_n = t_n)
\]

where \( T_n = \sum_{i=1}^{n} T_i \). Since the events \{ zero arrival in \( (T_n, T_n + t] \) \} and \( \{ \tau_i = t_i \}, \ i = 1, 2, \ldots, n \) are mutually independent, then

\[
P(\tau_{n+1} > t | \tau_1 = t_1, \tau_2 = t_2, \ldots, \tau_n = t_n) = P(\text{zero arrival in } (T_n, T_n + t]) = P(\tau_{n+1} > t) = e^{-\lambda t}
\]

which implies \( \tau_{n+1} \perp \tau_i, i = 1, 2, \ldots, n \) and \( \tau_{n+1} \sim \text{Exp}(\lambda) \).

Thus, from mathematical induction, the claim is true for all \( n \geq 1 \).

\[\square\]

6. Stationary and Independent Increments. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \). Suppose \( 0 = t_0 < t_1 < \ldots < t_n \), using the property of the Poisson process

\[
P(N_{t_i+h} = v | N_{t_i} = w) = \begin{cases} 1 - \lambda h + o(h) & v = w \\ \lambda h + o(h) & v = w + 1 \\ o(h) & v > w + 1 \end{cases}
\]

show that the increments

\[N_i - N_{t_{i-1}}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}\]

are stationary and independent with the distribution of \( N_{t_i} - N_{t_{i-1}} \), the same as the distribution of \( N_{t_{i-1}} \).

**Solution:** To show that \( \{N_t : t \geq 0\} \) is stationary we note that for \( 0 \leq t_{i-1} < t_i, m \geq 0 \)

\[
P(N_{t_i} - N_{t_{i-1}} = m) = \sum_u P(N_{t_i} = u + m, N_{t_{i-1}} = u)
\]

\[
= \sum_u P(N_{t_i} = u + m | N_{t_{i-1}} = u) P(N_{t_{i-1}} = u)
\]

\[
= \sum_u p_{u,u+m}(t_{i-1}, t_i) P(N_{t_{i-1}} = u)
\]

where \( p_{u,u+m}(t_{i-1}, t_i) = P(N_{t_i} = u + m | N_{t_{i-1}} = u) \).

From the property of the Poisson process

\[
P(N_{t_i+h} = v | N_{t_i} = w) = \begin{cases} 1 - \lambda h + o(h) & v = w \\ \lambda h + o(h) & v = w + 1 \\ o(h) & v > w + 1 \end{cases}
\]
we can write
\[
P(N_{t+h} = v|N_{t-1} = u) = P(N_{t+h} = v|N_{t} = v - 1)P(N_t = v - 1|N_{t-1} = u) + \sum_{u<v} P(N_{t+h} = v|N_{t} = u)P(N_t = u|N_{t-1} = u) = (\lambda h + o(h))P(N_t = v - 1|N_{t-1} = u) + (1 - \lambda h + o(h))P(N_t = v|N_{t-1} = u) + o(h)
\]
or
\[
p_{0,0}'(t_{i-1}, t_i) = -\lambda p_{0,0}(t_{i-1}, t_i), \quad p_{u,v}'(t_{i-1}, t_i) = -\lambda p_{u,v}(t_{i-1}, t_i) + \lambda p_{u,v-1}(t_{i-1}, t_i), \quad v = u + 1, u + 2, \ldots
\]
where
\[
p_{u,v}'(t_{i-1}, t_i) = \lim_{h \to 0} \frac{P(N_{t+h} = v|N_{t-1} = u) - P(N_{t} = v|N_{t-1} = u)}{h}
\]
Using the same steps as discussed in Problem 5.2.1.2 (page 253), we can deduce that
\[
p_{u,v}(t_{i-1}, t_i) = \frac{\lambda^{u-v}(t_i - t_{i-1})^{v-u} e^{-\lambda(t_i - t_{i-1})}}{(v-u)!}, \quad v = u + 1, u + 2, \ldots
\]
By substituting the above result into \(P(N_t - N_{t-1} = m)\),
\[
P(N_t - N_{t-1} = m) = \sum_u p_{u,v-1}(t_{i-1}, t_i)P(N_{t-1} = u) = \frac{\lambda^m(t_i - t_{i-1})^m e^{-\lambda(t_i - t_{i-1})}}{m!} \sum_u P(N_{t-1} = u) = \frac{\lambda^m(t_i - t_{i-1})^m e^{-\lambda(t_i - t_{i-1})}}{m!}
\]
which is a probability mass function for the Poisson \((\lambda(t_i - t_{i-1}))\) distribution. Thus, for all \(i = 1, 2, \ldots, n\), \(N_t - N_{t-1}\) is a stationary process and has the same distribution as Poisson \((\lambda(t_i - t_{i-1}))\).
Finally, to show \(\{N_t : t \geq 0\}\) has independent increments we note for \(m_i \geq 0, m_j \geq 0, i < j, i, j = 1, 2, \ldots, n\) and from the stationary property of the Poisson process,
\[
P(N_j - N_{j-1} = m_j|N_t - N_{t-1} = m_i) = \frac{P(N_j - N_{j-1} = m_j, N_t - N_{t-1} = m_i)}{P(N_t - N_{t-1} = m_i)} = \frac{P(N_j - N_{j-1} = m_j, N_{t-1} - N_{t-1} = m_i)}{P(N_t - N_{t-1} = m_i)} = \frac{P(\text{arrivals in } (t_{j-1}, t_j) \cap \text{arrivals in } (t_{i-1}, t_i))}{P(\text{arrivals in } (t_{i-1}, t_i))}.
\]
5.2.1 Properties of Poisson Process

Because \( i < j \) and since there are no overlapping events between the time intervals \((t_{i-1}, t_i]\) and \((t_{j-1}, t_j]\), thus for \( i < j, i, j = 1, 2, \ldots, n \)

\[
P(N_j - N_{j-1} = m_j | N_i - N_{i-1} = m_i) = P(\text{arrivals in } (t_{i-1}, t_i]) = P(N_j - N_{j-1} = m_j).
\]

**N.B.** In general, by denoting \( \mathcal{F}_{k-1} \) as the \( \sigma \)-algebra of information and observing the Poisson process \( N_t \) for \( 0 \leq t \leq t_{k-1} \), we can deduce that \( N_{t_k} - N_{t_{k-1}} \) is independent of \( \mathcal{F}_{k-1} \).

7. **Superposition.** Let \( N^{(1)}_t, N^{(2)}_t, \ldots, N^{(n)}_t \) be \( n \) independent Poisson processes with intensities \( \lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(n)} \), respectively, where the sequence of processes is defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). Show that \( N_i^{(1)} + N_i^{(2)} + \ldots + N_i^{(n)} \) is a Poisson process with intensity \( \lambda^{(1)} + \lambda^{(2)} + \ldots + \lambda^{(n)} \).

**Solution:** For \( N_i^{(i)} \sim \text{Poisson} \left( \lambda^{(i)} t \right) \) for \( i = 1, 2, \ldots, n \) the moment generating function for a Poisson process is

\[
M_{N_i^{(i)}}(t, u) = \mathbb{E} \left( e^{u N_i^{(i)}} \right) = e^{-\lambda^{(i)} t} \sum_{x=0}^{\infty} \frac{e^{u x} (\lambda^{(i)} t)^x}{x!} = e^{-\lambda^{(i)} t} \sum_{x=0}^{\infty} \frac{u x (\lambda^{(i)} t)^x}{x!} = e^{u (\lambda^{(i)} t) - 1}
\]

where \( u \in \mathbb{R} \).

By setting \( N_t = N_t^{(1)} + N_t^{(2)} + \ldots + N_t^{(n)} \) such that \( N_t^{(i)} \perp N_t^{(j)} \) for \( i \neq j, i, j = 1, 2, \ldots, n \) the moment generating function for \( N_t \) is

\[
M_{N_t}(t, u) = \mathbb{E} \left( e^{u N_t} \right) = \mathbb{E} \left( e^{u (N_t^{(1)} + N_t^{(2)} + \ldots + N_t^{(n)})} \right) = e^{-\lambda^{(1)} t} \mathbb{E} \left( e^{u N_t^{(1)}} \right) \mathbb{E} \left( e^{u N_t^{(2)}} \right) \ldots \mathbb{E} \left( e^{u N_t^{(n)}} \right) = e^{(\lambda^{(1)} + \lambda^{(2)} + \ldots + \lambda^{(n)}) t - 1}.
\]

Therefore,

\[
N_t^{(1)} + N_t^{(2)} + \ldots + N_t^{(n)} \sim \text{Poisson} \left( (\lambda^{(1)} + \lambda^{(2)} + \ldots + \lambda^{(n)}) t \right).
\]

8. **Markov Property.** Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). Show that if \( f \) is a continuous function then there exists another continuous function \( g \) such that

\[
\mathbb{E} \left[ f(N_t) | \mathcal{F}_u \right] = g(N_u)
\]

for \( 0 \leq u \leq t \).

**Solution:** For \( 0 \leq u \leq t \) we can write

\[
\mathbb{E} \left[ f(N_t) | \mathcal{F}_u \right] = \mathbb{E} \left[ f(N_t - N_u + N_u) | \mathcal{F}_u \right].
\]
5.2.1 Properties of Poisson Process

Since \( N_t - N_u \perp \mathcal{F}_u \) and \( N_u \) is \( \mathcal{F}_u \) measurable, by setting \( N_u = x \) where \( x \) is a constant value,

\[
\mathbb{E} \left[ f(N_t - N_u + x) \Big| \mathcal{F}_u \right] = \mathbb{E} \left[ f(N_t - N_u + x) \right].
\]

Because \( N_t - N_u \sim \text{Poisson}(\lambda(t - u)) \), we can write \( \mathbb{E} \left[ f(N_t - N_u + x) \right] \) as

\[
\mathbb{E} \left[ f(N_t - N_u + x) \right] = \sum_{k=0}^{\infty} f(k+x) e^{-\lambda(t-u)} (\frac{\lambda(t-u)^k}{k!}).
\]

By setting \( \tau = t - u \) and \( y = k + x \), we can rewrite \( \mathbb{E} \left[ f(N_t - N_u + x) \right] = \mathbb{E} \left[ f(N_t - N_u + N_u) \right] \) as

\[
\mathbb{E} \left[ f(N_t - N_u + x) \right] = \sum_{y=0}^{\infty} f(y+N_u) \frac{e^{-\lambda \tau} (\lambda \tau)^y}{y!}.
\]

Since the only information from the filtration \( \mathcal{F}_u \) is \( N_u \), then

\[ \mathbb{E} \left[ f(N_t) \Big| \mathcal{F}_u \right] = g(N_u) \]

where

\[ g(N_u) = \sum_{y=0}^{\infty} f(y+N_u) \frac{e^{-\lambda \tau} (\lambda \tau)^y}{y!}. \]

\[ \blacksquare \]

9. **Compensated Poisson Process.** Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \). By defining the compensated Poisson process, \( \hat{N}_t \), as

\[ \hat{N}_t = N_t - \lambda t \]

show that \( \hat{N}_t \) is a martingale.

**Solution:** To show that \( \hat{N}_t \) is a martingale, for \( 0 \leq s < t \) we note the following.

(a) Since the increment \( N_t - N_s \) is independent of \( \mathcal{F}_s \) and has expected value \( \lambda(t-s) \), we can write

\[
\mathbb{E} \left( \hat{N}_t \Big| \mathcal{F}_s \right) = \mathbb{E} \left( \hat{N}_t - \hat{N}_s + \hat{N}_s \Big| \mathcal{F}_s \right)
\]

\[
= \mathbb{E} \left( \hat{N}_t - \hat{N}_s \Big| \mathcal{F}_s \right) + \mathbb{E} \left( \hat{N}_s \Big| \mathcal{F}_s \right)
\]

\[
= \mathbb{E} \left( N_t - N_s - \lambda(t-s) \Big| \mathcal{F}_s \right) + \hat{N}_s
\]

\[
= \mathbb{E} \left( N_t - N_s \right) - \lambda(t-s) + \hat{N}_s
\]

\[ = \hat{N}_s. \]
5.2.1 Properties of Poisson Process

10. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). By defining the compensated Poisson process, \( \hat{N}_t \) as

\[
\hat{N}_t = N_t - \lambda t
\]

show that \( \hat{N}_t^2 - \lambda t \) is a martingale.

**Solution:** To show that \( \hat{N}_t^2 - \lambda t \) is a martingale, for \( 0 \leq s < t \) we note the following.

(a) Since the increment \( N_t - N_s \perp \mathcal{F}_s \) and has expected value \( \lambda(t-s) \), and because \( N_t - N_s \perp N_s \), we can write

\[
\mathbb{E} \left( \hat{N}_t^2 - \lambda t \mid \mathcal{F}_s \right) = \mathbb{E} \left[ (\hat{N}_t - \hat{N}_s)^2 \right] + \lambda t
\]

\[
= \mathbb{E} \left[ (\hat{N}_t - \hat{N}_s)^2 \right] + 2 \mathbb{E} \left[ (\hat{N}_t - \hat{N}_s) \hat{N}_s \right] + \mathbb{E} \left( \hat{N}_s^2 \right) - \lambda t
\]

\[
= \mathbb{E} \left[ (N_t - N_s)^2 \right] - \lambda (t-s) + \mathbb{E} \left( \hat{N}_s^2 \right) + \lambda t
\]

\[
= \hat{N}_s^2 - \lambda s.
\]

(b) Since \( \left| \hat{N}_t^2 - \lambda t \right| \leq (N_t - \lambda t)^2 + \lambda t \therefore \)

\[
\mathbb{E} \left( \left| \hat{N}_t^2 - \lambda t \right| \right) \leq \mathbb{E} \left[ (N_t - \lambda t)^2 \right] + \lambda t = 2\lambda t < \infty
\]

as \( \mathbb{E} \left[ (N_t - \lambda t)^2 \right] = \text{Var} \left( N_t \right) = \lambda t. \)

(c) The process \( \hat{N}_t^2 - \lambda = (N_t - \lambda t)^2 - \lambda t \) is clearly \( \mathcal{F}_t \)-adapted.

From (a)-(c) we have shown that \( \hat{N}_t^2 - \lambda t \) is a martingale.

11. *Exponential Martingale Process.* Let \( N_t \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). Show that for \( u \in \mathbb{R} \),

\[
X_t = e^{\lambda N_t - \lambda t(e^u - 1)}
\]

is a martingale.

**Solution:** Given that \( N_t \sim \text{Poisson}(\lambda t) \) then for \( u \in \mathbb{R} \),

\[
\mathbb{E}(e^{uN_t}) = e^{-\lambda t} \sum_{x=0}^{\infty} \frac{e^{ux} (\lambda t)^x}{x!} = e^{-\lambda t} \sum_{x=0}^{\infty} \frac{(\lambda t e^u)^x}{x!} = e^{\lambda t(e^u - 1)}
\]
and hence for $0 \leq s < t$ and because $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$,
\[
\mathbb{E}(e^{a(N_t - N_s)}) = e^{\lambda(t-s)(e^a-1)}.
\]
To show that $X_t = e^{\lambda N_t - \lambda t(e^a-1)}$ is a martingale, for $0 \leq s < t$ we have the following:

(a) Since $N_t - N_s \perp \mathcal{F}_s$ and has stationary increment, therefore $e^{\lambda N_t - \lambda t(e^a-1)}$ is independent of $e^{N_s}$ Thus, we can write
\[
\mathbb{E}(X_t|\mathcal{F}_s) = \mathbb{E}(e^{\lambda N_t - \lambda t(e^a-1)}|\mathcal{F}_s)
= e^{-\lambda t(e^a-1)} \mathbb{E}(e^{\lambda N_s}|\mathcal{F}_s)
= e^{-\lambda t(e^a-1)} \mathbb{E}(e^{\lambda N_s} - \lambda t(e^a-1))
= e^{\lambda N_s - \lambda t(e^a-1)}
= X_s.
\]

(b) Since $|X_t| = \left|e^{\lambda N_t - \lambda t(e^a-1)}\right| = e^{\lambda N_t - \lambda t(e^a-1)}$, we can write
\[
\mathbb{E}(|X_t|) = \mathbb{E}(e^{\lambda N_t - \lambda t(e^a-1)}) = e^{-\lambda t(e^a-1)} \mathbb{E}(e^{\lambda N_t}) = e^{-\lambda t(e^a-1)} \cdot e^{\lambda t(e^a-1)} = 1 < \infty.
\]

(c) The process $X_t = e^{\lambda N_t - \lambda t(e^a-1)}$ is clearly $\mathcal{F}_t$-adapted.

From (a)–(c) we have shown that $X_t = e^{\lambda N_t - \lambda t(e^a-1)}$ is a martingale.

\[\square\]

12. Compound Poisson Process. Let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, and let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with common mean $\mathbb{E}(X_i) = \mathbb{E}(X) = \mu$ and variance $\mathbb{Var}(X_i) = \mathbb{Var}(X) = \sigma^2$. Let $X_1, X_2, \ldots$ be independent of $N_t$. By defining the compound Poisson process $M_t$ as
\[
M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0
\]
show that the moment generating function for $M_t$ is
\[
\varphi_{M_t}(t, u) = e^{\lambda t (\varphi_X(u)-1)}, \quad u \in \mathbb{R}
\]
where $\varphi_X(u) = \mathbb{E}(e^{uX})$.

Further, show that $\mathbb{E}(M_t) = \mu \lambda t$ and $\mathbb{Var}(M_t) = (\mu^2 + \sigma^2) \lambda t$.

Solution: By definition, for $u \in \mathbb{R}$
\[
\varphi_{M_t}(t, u) = \mathbb{E}(e^{uX}) = \mathbb{E}(e^{u \sum_{i=1}^{N_t} X_i})
\]
5.2.1 Properties of Poisson Process

\[ = \mathbb{E} \left[ \mathbb{E} \left( e^{\sum_{i=1}^{N_t} X_i} \mid N_t \right) \right] \]

\[ = \sum_{n=0}^{\infty} \mathbb{E} \left( e^{\sum_{i=1}^{n} X_i} \mid N_t = n \right) \mathbb{P}(N_t = n). \]

Since \( X_1, X_2, \ldots, X_n \) are independent and identically distributed, we can therefore write

\[ \varphi_{M_t}(t, u) = \sum_{n=0}^{\infty} \mathbb{E}(e^{nX})^n \mathbb{P}(N_t = n) = \sum_{n=0}^{\infty} \mathbb{E}(e^{nX})^n \frac{e^{-\lambda t}(\lambda t)^n}{n!} = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t \varphi_X(t))^n}{n!} = e^{\lambda t(\varphi_X(u)-1)}. \]

By taking first and second partial derivatives of \( \varphi_{M_t}(t, u) \) with respect to \( u \), we have

\[ \frac{\partial \varphi_{M_t}(t, u)}{\partial u} = \lambda t \varphi_X'(u) e^{\lambda t(\varphi_X(u)-1)} \]

and

\[ \frac{\partial^2 \varphi_{M_t}(t, u)}{\partial u^2} = \left\{ \lambda t \varphi_X''(u) + (\lambda t)^2 \left( \varphi_X'(u)^2 \right) \right\} e^{\lambda t(\varphi_X(u)-1)}. \]

Since \( \varphi_X(0) = 1 \), therefore

\[ \mathbb{E}(M_t) = \frac{\partial \varphi_{M_t}(t, 0)}{\partial u} = \lambda t \varphi_X'(0) = \lambda t \mathbb{E}(X) = \mu \lambda t \]

and

\[ \mathbb{E} \left( M_t^2 \right) = \frac{\partial^2 \varphi_{M_t}(t, 0)}{\partial u^2} = \lambda t \varphi_X''(0) + (\lambda t)^2 \left( \varphi_X'(0)^2 \right) = \lambda t \mathbb{E}(X^2) + (\lambda t)^2 \mathbb{E}(X)^2. \]

Thus, the variance of \( Y \) is

\[ \text{Var}(M_t) = \mathbb{E} \left( M_t^2 \right) - \mathbb{E}(M_t)^2 = \lambda t \mathbb{E}(X^2) = \lambda t \left[ \text{Var}(X) + \mathbb{E}(X)^2 \right] = (\mu^2 + \sigma^2) \lambda t. \]

\[ \square \]

13. Martingale Properties of Compound Poisson Process. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \), and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables with common mean \( \mathbb{E}(X_i) = \mathbb{E}(X) = \mu \) and variance \( \text{Var}(X_i) = \text{Var}(X) = \sigma^2 \). Let \( X_1, X_2, \ldots \) be independent of \( N_t \). By defining the compound Poisson process \( M_t \) as

\[ M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \]
5.2.1 Properties of Poisson Process

and assuming \( \mathbb{E}(|X|) < \infty \), show that for \( 0 \leq s < t \)

\[
\mathbb{E}(M_t | \mathcal{F}_s) = \begin{cases} 
M_s & \text{if } \mu = 0 \\
\geq M_s & \text{if } \mu > 0 \\
\leq M_s & \text{if } \mu < 0.
\end{cases}
\]

Solution: To show that \( M_t \) can be a martingale, submartingale or supermartingale, for \( 0 \leq s < t \) we have the following.

(a) \( \mathbb{E}(M_t | \mathcal{F}_s) = \mathbb{E}(M_t - M_s + M_s | \mathcal{F}_s) = \mathbb{E}(M_t - M_s | \mathcal{F}_s) + \mathbb{E}(M_s | \mathcal{F}_s) = \mu \lambda (t - s) + M_s \)

since the increment \( M_t - M_s \) is independent of \( \mathcal{F}_s \) and has mean \( \mu \lambda (t - s) \).

(b) \( \mathbb{E}(M_t) = \mathbb{E}\left( \left\lfloor \sum_{i=1}^{N_t} X_i \right\rfloor \right) \leq \mathbb{E}\left( \sum_{i=1}^{N_t} |X_i| \right) = \lambda t \mathbb{E}(|X|) < \infty \) since \( \mathbb{E}(|X|) < \infty \).

(c) The process \( M_t = \sum_{i=1}^{N_t} X_i \) is clearly \( \mathcal{F}_t \)-adapted.

We can therefore deduce that \( M_t \) is a martingale, submartingale or supermartingale by setting \( \mu = 0, \mu > 0 \) or \( \mu < 0 \), respectively.

\[ \square \]

14. Compensated Compound Poisson Process. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables with common mean \( \mathbb{E}(X) = \mathbb{E}(X) = \mu \) and variance \( \text{Var}(X) = \text{Var}(X) = \sigma^2 \). Let \( X_1, X_2, \ldots \) be independent of \( N_t \). By defining the compound Poisson process \( M_t \) as

\[
M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0
\]

and assuming \( \mathbb{E}(|X|) < \infty \), show that the compensated compound Poisson process

\[
\hat{M}_t = M_t - \mu \lambda t
\]

is a martingale.

Solution: To show that \( \hat{M}_t \) is a martingale, for \( 0 \leq s < t \) we have the following.

(a) Given the increment \( M_t - M_s \) is independent of \( \mathcal{F}_s \) and has mean \( \mu \lambda (t - s) \),

\[
\mathbb{E}\left( \left\lfloor \hat{M}_t \right\rfloor | \mathcal{F}_s \right) = \mathbb{E}\left( \hat{M}_t - \hat{M}_s + \hat{M}_s | \mathcal{F}_s \right)
= \mathbb{E}\left( M_t - M_s - \mu \lambda (t - s) + M_s - \mu \lambda s | \mathcal{F}_s \right)
= \mathbb{E}\left( M_t - M_s | \mathcal{F}_s \right) - \mu \lambda (t - s) + M_s - \mu \lambda s
= \mu \lambda (t - s) - \mu \lambda (t - s) + M_s - \mu \lambda s
= M_s - \mu \lambda s.
\]

(b) \( \mathbb{E}\left| \left\lfloor \hat{M}_t \right\rfloor \right| = \mathbb{E}\left( \left| \sum_{i=1}^{N_t} X_i - \mu \lambda t \right| \right) \leq \mathbb{E}\left( \sum_{i=1}^{N_t} |X_i| \right) + \mu \lambda t = \lambda t \mathbb{E}(|X|) + \mu \lambda t < \infty \)

since \( \mathbb{E}(|X|) < \infty \).
5.2.1 Properties of Poisson Process

(c) The process \( \hat{M}_t = \sum_{i=1}^{N_t} X_i - \mu \lambda t \) is clearly \( \mathcal{F}_t \)-adapted.

From the results of (a)–(c) we have shown that \( \hat{M}_t = \sum_{i=1}^{N_t} X_i - \mu \lambda t \) is a martingale.

\[ \Box \]

15. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \). Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables with common mean \( \mathbb{E}(X_i) = \mathbb{E}(X) = \mu \) and variance \( \text{Var}(X_i) = \text{Var}(X) = \sigma^2 \). Let \( X_1, X_2, \ldots \) be independent of \( N_t \). By defining the compound Poisson process \( M_t \) as

\[ M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \]

show that for \( 0 = t_0 < t_1 < t_2 < \ldots < t_n \) the increments

\[ M_{t_1} - M_{t_0}, M_{t_2} - M_{t_1}, \ldots, M_{t_n} - M_{t_{n-1}} \]

are independent, stationary and the distribution of \( M_{t_k} - M_{t_{k-1}} \) is the same as the distribution of \( M_{t_{k-1}} - t_{k-1} \).

**Solution:** We first show that the increment \( M_{t_k} - M_{t_{k-1}}, k = 1, 2, \ldots, n \) is stationary and has the same distribution as \( M_{t_{k-1}} - t_{k-1} \).

Using the technique of moment generating function, for \( u \in \mathbb{R} \)

\[
\phi_{M_{t_k} - M_{t_{k-1}}}(u) = \mathbb{E}\left(e^{u(M_{t_k} - M_{t_{k-1}})}\right) = \mathbb{E}\left(e^{u(\sum_{i=1}^{N_k} X_i - N_{t_{k-1}})}\right) = \mathbb{E}\left(e^{u \sum_{i=1}^{N_k} X_i} \right) = \phi_{M_{t_{k-1}}} (u).
\]

Therefore, \( M_{t_k} - M_{t_{k-1}}, k = 1, 2, \ldots, n \) is stationary and has the same distribution as \( M_{t_{k-1}} - t_{k-1} \).

Finally, to show that \( M_{t_i} - M_{t_{i-1}} \) is independent of \( M_{t_j} - M_{t_{j-1}}, i, j = 1, 2, \ldots, n \) and \( i \neq j \), using the joint moment generating function, for \( u, v \in \mathbb{R} \)

\[
\phi_{(M_{t_i} - M_{t_{i-1}})(M_{t_j} - M_{t_{j-1}})}(u, v) = \mathbb{E}\left(e^{u(M_{t_i} - M_{t_{i-1}})} + v(M_{t_j} - M_{t_{j-1}})\right) = \mathbb{E}\left(e^{u \sum_{i=1}^{N_i} X_i} + v \sum_{i=1}^{N_j} X_i\right) = \mathbb{E}\left(e^{u \sum_{i=1}^{N_i} X_i} \right) \mathbb{E}\left(e^{v \sum_{i=1}^{N_j} X_i}\right) = \phi_{M_{t_i} - t_{i-1}}(u) \phi_{M_{t_j} - t_{j-1}}(v).
\]
268 5.2.1 Properties of Poisson Process

\[ \phi_{M_{t_i - 1}}(u) \phi_{M_{t_j - 1}}(v) \]

since \(X_1, X_2, \ldots\) is a sequence of independent and identically distributed random variables and also the intervals \([N_{t_{i-1}} + 1, N_{t_{i}}]\) and \([N_{t_{j-1}} + 1, N_{t_{j}}]\) have no overlapping events.

Thus, \(M_{t_i} - M_{t_{i-1}} \perp M_{t_j} - M_{t_{j-1}}, \text{ for all } i, j = 1, 2, \ldots, n, \ i \neq j.\)

N.B. Since the compound Poisson process \(M_t\) has increments which are independent and stationary, therefore from Problem 5.2.1.8 (page 261) we can deduce that \(M_t\) is Markov. That is, if \(f\) is a continuous function then there exists another continuous function \(g\) such that

\[ \mathbb{E}[f(M_t) \mid \mathcal{F}_u] = g(M_u) \]

for \(0 \leq u \leq t.\)

\(\square\)

16. Decomposition of a Compound Poisson Process (Finite Jump Size). Let \(N_t\) be a Poisson process with intensity \(\lambda > 0\) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \(\mathcal{F}_t\). Let \(X_1, X_2, \ldots\) be a sequence of independent and identically distributed random variables, and let \(X_1, X_2, \ldots\) be independent of \(N_t\) such that

\[ \mathbb{P}(X = x_k) = p(x_k) \quad \text{and} \quad \sum_{k=1}^{K} \mathbb{P}(X = x_k) = 1 \]

where \(X_i \overset{d}{=} X, i = 1, 2, \ldots\) and \(x_1, x_2, \ldots, x_K\) is a finite set of non-zero numbers. By defining the compound Poisson process as

\[ M_t = \sum_{i=1}^{N_t} X_i \]

show that \(M_t\) and \(N_t\) can be written as

\[ M_t = \sum_{k=1}^{K} x_k N_t^{(k)} \quad \text{and} \quad N_t = \sum_{k=1}^{K} N_t^{(k)} \]

where \(N_t^{(k)} \sim \text{Poisson}(\lambda t p(x_k)), k = 1, 2, \ldots, K\) is a sequence of independent and identical Poisson processes.

Solution: From Problem 5.2.1.12 (page 264), the moment generating function for \(M_t\) is

\[ \phi_{M_t}(t, u) = e^{t(\phi_X(u) - 1)}, \quad u \in \mathbb{R} \]

where \(\phi_X(u) = \mathbb{E}(e^{uX}).\) Given \(\mathbb{P}(X = x_k) = p(x_k), k = 1, 2, \ldots, K\) we have

\[ \phi_X(u) = \mathbb{E}(e^{uX}) \]

\[ = \sum_{k=1}^{K} e^{u x_k} \mathbb{P}(X = x_k) \]

\[ = \sum_{k=1}^{K} e^{u x_k} p(x_k). \]
5.2.1 Properties of Poisson Process

Substituting $\varphi_x(u)$ into the moment generating function for $M_t$, 

$$\varphi_{M_t}(t, u) = e^{\lambda \left( \sum_{k=1}^{K} e^{u k} \mathbb{P}(X_k) - 1 \right)}.$$ 

Because $\sum_{k=1}^{K} \mathbb{P}(x_k) = 1$, 

$$\varphi_{M_t}(t, u) = e^{\lambda \left( \sum_{k=1}^{K} e^{u k} \mathbb{P}(x_k) - \sum_{k=1}^{K} \mathbb{P}(x_k) \right)}$$ 

$$= \prod_{k=1}^{K} e^{\lambda \mathbb{P}(x_k) (e^{u k} - 1)}$$ 

$$= \prod_{k=1}^{K} \mathbb{E} \left( e^{u x_k N_t^{(k)}} \right)$$

which is a product of $K$ independent moment generating functions of the random variable $x_k N_t^{(k)}$ where $N_t^{(k)} \sim \text{Poisson} \left( \lambda t \mathbb{P}(x_k) \right)$. Thus, 

$$\varphi_{M_t}(t, u) = \prod_{k=1}^{K} \mathbb{E} \left( e^{u x_k N_t^{(k)}} \right) = \mathbb{E} \left( e^{u \sum_{k=1}^{K} x_k N_t^{(k)}} \right)$$

which implies $M_t = \sum_{k=1}^{K} x_k N_t^{(k)}$.

Given that $N_t^{(k)} \sim \text{Poisson} \left( \lambda t \mathbb{P}(x_k) \right), k = 1, 2, \ldots , K$ is a sequence of independent and identically distributed Poisson processes, $\sum_{k=1}^{K} N_t^{(k)} \sim \text{Poisson} \left( \sum_{k=1}^{K} \lambda t \mathbb{P}(x_k) \right) \text{ or } \sum_{k=1}^{K} N_t^{(k)} \sim \text{Poisson}(\lambda t)$ and hence $N_t = \sum_{k=1}^{K} N_t^{(k)}$. 

\[\square\]

17. Let $\{X_1, X_2, \ldots \}$ be a series of independent and identically distributed random variables with moment generating function 

$$M_X(\xi) = \mathbb{E} \left( e^{\xi X} \right), \quad \xi \in \mathbb{R}$$

where $X \overset{d}{=} X_i, i = 1, 2, \ldots$ and let $\{t_1, t_2, \ldots \}$ be its corresponding sequence of independent and identically distributed random jump times of a Poisson process $\{N_t : t \geq 0\}$ with intensity parameter $\lambda > 0$ where $t_i \sim \text{Exp}(\lambda), i = 1, 2, \ldots$ Let $X_1, X_2, \ldots$ be independent of $N_t$.

Show that for $n \geq 1, \kappa > 0$ the process $\{Y_t : t \geq 0\}$ with initial condition $Y_0 = 0$ given as 

$$Y_t = \sum_{i=1}^{n} e^{-\kappa(t-i)} X_i$$

has moment generating function $M_Y(\xi, t) = \mathbb{E} \left( e^{\xi Y} \right)$ given by 

$$M_Y(\xi, t) = \prod_{i=1}^{n} \left\{ \int_{0}^{t} M_X(\xi e^{-\kappa s}) \lambda e^{-\lambda(t-s)} ds \right\}$$
and
\[ E(Y_t) = n \mu \left( \frac{\lambda}{\lambda - \kappa} \right) (e^{-\kappa t} - e^{-\lambda t}) (1 - e^{-\lambda t})^{n-1} \]

\[ \text{Var} (Y_t) = n(\mu^2 + \sigma^2) \left( \frac{\lambda}{\lambda - 2\kappa} \right) (e^{-2\kappa t} - e^{-\lambda t}) (1 - e^{-\lambda t})^{n-1} \]

\[ + n(n-1)\mu^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 (e^{-2\kappa t} - e^{-\lambda t})^2 (1 - e^{-\lambda t})^{n-2} \]

\[ - n^2 \mu^2 \left( \frac{\lambda}{\lambda - \kappa} \right)^2 (e^{-\kappa t} - e^{-\lambda t})^2 (1 - e^{-\lambda t})^{2(n-1)} \]

where \( E(X_i) = \mu \) and \( \text{Var} (X_i) = \sigma^2 \) for \( i = 1, 2, \ldots \)

**Solution:** By definition,

\[ M_{Y} (\xi, t) = E (e^X Y_t) = E \left( e^{\xi \sum_{i=1}^{n} e^{-\xi t l_{i-1}} X_i} \right) = \prod_{i=1}^{n} E \left( e^{\xi e^{-\kappa_{l_{i-1}}X_i}} \right) = \prod_{i=1}^{n} M_{X_i} (\xi e^{-\kappa_{l_{i-1}}}). \]

Using the tower property,

\[ M_{X_i} (\xi e^{-\kappa_{l_{i-1}}}) = E \left( e^{\xi e^{-\kappa_{l_{i-1}}}} X_i \right) \]

\[ = E \left[ E \left( e^{\xi e^{-\kappa_{l_{i-1}}} X_i} \mid t_i = s \right) \right] \]

\[ = \int_{0}^{t} E \left( e^{\xi e^{-\kappa_{l_{i-1}}} X_i} \right) \lambda e^{-\lambda s} \, ds \]

\[ = \int_{0}^{t} M_{X_i} (\xi e^{-\kappa_{l_{i-1}}}) \lambda e^{-\lambda s} \, ds \]

\[ = \int_{0}^{t} M_{X_i} (\xi e^{-\kappa_{l_{i-1}}}) \lambda e^{-\lambda (t-s)} \, ds. \]

Therefore,

\[ M_{Y} (\xi, t) = \prod_{i=1}^{n} \left\{ \int_{0}^{t} M_{X_i} (\xi e^{-\kappa_{l_{i-1}}}) \lambda e^{-\lambda (t-s)} \, ds \right\}. \]

By setting \( \tilde{\xi} = \xi e^{-\kappa s} \) and \( \tilde{M}_{X_i} (\tilde{\xi}) = \int_{0}^{t} M_{X_i} (\xi e^{-\kappa s}) \lambda e^{-\lambda (t-s)} \, ds \), and differentiating \( M_{Y} (\xi, t) \) with respect to \( \xi \), we have

\[ \frac{\partial M_{Y} (\xi, t)}{\partial \tilde{\xi}} = \sum_{i=1}^{n} \left\{ \frac{\partial M_{X_i} (\tilde{\xi})}{\partial \tilde{\xi}} \prod_{j=1}^{n} M_{X_j} (\tilde{\xi}) \right\} \]

\[ = \sum_{i=1}^{n} \left\{ \left[ \int_{0}^{t} \frac{\partial M_{X_i} (\tilde{\xi})}{\partial \tilde{\xi}} \cdot \frac{\partial \tilde{\xi}}{\partial \xi} \cdot \lambda e^{-\lambda (t-s)} \, ds \right] \prod_{j=1}^{n} M_{X_j} (\tilde{\xi}) \right\}. \]
5.2.1 Properties of Poisson Process

By substituting \( \xi = 0 \) and taking note that \( M_X(0) = 1 \), \( \frac{\partial M_X(0)}{\partial \xi} = \text{E}(X) = \mu \) and
\( \frac{\partial^2 M_X(0)}{\partial \xi^2} = \text{E}(X^2) = \mu^2 + \sigma^2 \), we therefore have

\[
\text{E}(Y_i) = \frac{\partial M_Y(0, t)}{\partial \xi} = \sum_{i=1}^{n} \left[ \int_0^t \mu e^{-\lambda t} e^{(\lambda - \kappa) s} ds \right] \prod_{j=1 \atop j \neq i}^{n} \left[ \int_0^t \lambda e^{-\lambda (t-s)} ds \right]
\]

\[
= n \mu \left( \frac{\lambda}{\lambda - \kappa} \right) (e^{-\kappa t} - e^{-\lambda t}) (1 - e^{-\lambda t})^{n-1}
\]

and

\[
\text{E}(Y_i^2) = \frac{\partial^2 M_Y(0, t)}{\partial \xi^2} = \sum_{i=1}^{n} \left[ \int_0^t (\mu^2 + \sigma^2) \lambda e^{-\lambda t} e^{(\lambda - 2\kappa) s} ds \right] \prod_{j=1 \atop j \neq i}^{n} \left[ \int_0^t \lambda e^{-\lambda (t-s)} ds \right]
\]
5.2.1 Properties of Poisson Process

Therefore,

\[
\mathbb{E}(Y_t) = \mathbb{E}(Y^2_t) - \mathbb{E}(Y_t)^2
\]

\[
= n(\mu^2 + \sigma^2) \left( \frac{\lambda}{\lambda - 2\kappa} \right) \left( e^{-2\kappa t} - e^{-\lambda t} \right) (1 - e^{-\lambda t})^{n-1}
\]

\[
+ n(n-1)\mu^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \left( e^{-2\kappa t} - e^{-\lambda t} \right)^2 (1 - e^{-\lambda t})^{n-2}
\]

\[
- n^2 \mu^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \left( e^{-\kappa t} - e^{-\lambda t} \right)^2 (1 - e^{-\lambda t})^{2(n-1)}\]

18. Let \( \{X_1, X_2, \ldots, X_n\} \) be a series of independent and identically distributed random variables with moment generating function

\[
M_X(\xi) = \mathbb{E}(e^{\xi X}), \quad \xi \in \mathbb{R}
\]

where \( X \overset{d}{=} X_i, \ i = 1, 2, \ldots \) and the \( \{t_1, t_2, \ldots\} \) is its corresponding sequence of random jump times of a Poisson process \( \{N_t : t \geq 0\} \) with intensity parameter \( \lambda > 0 \) where \( t_i \sim \text{Exp} (\lambda), \ i = 1, 2, \ldots \) In addition, let the sequence \( X_1, X_2, \ldots \) be independent of \( N_t \). Show that for \( \kappa > 0 \) the process \( \{Y_t : t \geq 0\} \) with initial condition \( Y_0 = 0 \) given as

\[
Y_t = \sum_{i=1}^{N_t} e^{-\kappa (t-t_i)} X_i
\]

has moment generating function \( M_Y(\xi, t) = \mathbb{E}(e^{\xi Y_t}) \) given by

\[
M_Y(\xi, t) = \exp \left\{ \lambda \int_0^t (M_X(\xi e^{-\kappa s}) - 1) \ ds \right\}
\]

and

\[
\mathbb{E}(Y_t) = \frac{\lambda \mu}{\kappa} \left( 1 - e^{-\kappa t} \right)
\]

\[
\text{Var}(Y_t) = \frac{\lambda}{2\kappa} \left( \mu^2 + \sigma^2 \right) \left( 1 - e^{-2\kappa t} \right)
\]

where \( \mathbb{E}(X_i) = \mu \) and \( \text{Var}(X_i) = \sigma^2 \) for \( i = 1, 2, \ldots \)
5.2.1 Properties of Poisson Process

Solution: Given the mutual independence of the random variables \( X_i \) and \( X_j, \) \( i \neq j \) the conditional expectation of the process \( Y_t \) given the first jump time, \( t_1 = s \) is

\[
\mathbb{E} \left( e^{\xi Y_{t_1}} \middle| t_1 = s \right) = \mathbb{E} \left( \exp \left\{ \xi \sum_{i=1}^{N_t} e^{-\kappa(t-t_i)} X_i \right\} \middle| t_1 = s \right)
\]

\[
= \mathbb{E} \left( \exp \left\{ \xi e^{-\kappa(t-s)} X_1 \right\} \middle| t_1 = s \right) \mathbb{E} \left( \exp \left\{ \xi \sum_{i=2}^{N_t} e^{-\kappa(t-t_i)} X_i \right\} \middle| t_1 = s \right).
\]

From the properties of the Poisson process, the sum conditioned on \( t_1 = s \) has the same unconditional distribution as the original sum, starting with the first jump \( i = 1 \) until time \( t - s, \) and we therefore have

\[
\mathbb{E} \left( e^{\xi Y_{t_1}} \middle| t_1 = s \right) = \mathbb{E} \left( \exp \left\{ \xi e^{-\kappa(t-s)} X_1 \right\} \right) M_Y(\xi, t - s)
\]

\[
= M_X(\xi e^{-\kappa(t-s)}) M_Y(\xi, t - s).
\]

Since the first jump time is exponentially distributed, \( t_1 \sim \text{Exp}(\lambda) \) and from the tower property

\[
M_Y(\xi, t) = \mathbb{E} \left[ \mathbb{E} \left( e^{\xi Y_{t_1}} \middle| t_1 = s \right) \right]
\]

\[
= \int_0^t \mathbb{E} \left( e^{\xi Y_{t_1}} \middle| t_1 = s \right) \lambda e^{-\lambda s} ds
\]

\[
= \int_0^t M_X(\xi e^{-\kappa(t-s)}) M_Y(\xi, t - s) \lambda e^{-\lambda s} ds
\]

\[
= \int_0^t M_X(\xi e^{-\kappa(t-s)}) M_Y(\xi, s) \lambda e^{-\lambda(t-s)} ds.
\]

Differentiating the integral with respect to \( t, \) we have

\[
\frac{\partial M_Y(\xi, t)}{\partial t} = M_X(\xi e^{-\kappa t}) M_Y(\xi, t) \lambda - \lambda \int_0^t M_X(\xi e^{-\kappa s}) M_Y(\xi, s) \lambda e^{-\lambda(t-s)} ds
\]

\[
= \lambda (M_X(\xi e^{-\kappa t}) - 1) M_Y(\xi, t)
\]

and solving the first-order differential equation,

\[
M_Y(\xi, t) = \exp \left( \lambda \int_0^t \left\{ M_X(\xi e^{-\kappa s}) - 1 \right\} ds \right).
\]

Setting \( \bar{\xi} = \xi e^{-\kappa t} \) and differentiating \( M_Y(\xi, t) \) with respect to \( \xi, \) we have

\[
\frac{\partial M_Y(\xi, t)}{\partial \xi} = \frac{\partial}{\partial \bar{\xi}} \left( \lambda \int_0^t \left\{ M_X(\bar{\xi}) - 1 \right\} ds \right) M_Y(\xi, t)
\]

\[
= \lambda \left\{ \int_0^t e^{-\kappa s} \frac{\partial M_X(\bar{\xi})}{\partial \bar{\xi}} ds \right\} M_Y(\xi, t)
\]
and differentiating the above equation with respect to $\xi$,
\[
\frac{\partial^2 M_Y(\xi, t)}{\partial \xi^2} = \lambda \frac{\partial}{\partial \xi} \left\{ \int_0^t e^{-\xi s} \frac{\partial M_X(\xi)}{\partial \xi} ds \right\} M_Y(\xi, t) + \lambda \left\{ \int_0^t e^{-\xi s} \frac{\partial M_X(\xi)}{\partial \xi} ds \right\} \frac{\partial M_Y(\xi, t)}{\partial \xi}
\]
\[
= \lambda \left\{ \int_0^t e^{-2\xi s} \frac{\partial^2 M_X(\xi)}{\partial \xi^2} ds \right\} M_Y(\xi, t) + \lambda \left\{ \int_0^t e^{-\xi s} \frac{\partial M_X(\xi)}{\partial \xi} ds \right\} \frac{\partial M_Y(\xi, t)}{\partial \xi}.
\]

By substituting $\xi = 0$, and since $M_Y(0, t) = 1$, we therefore have
\[
\mathbb{E}(Y_t) = \frac{\partial M_Y(0, t)}{\partial \xi} = \frac{\lambda \mu}{\lambda} \left( 1 - e^{-\lambda t} \right)
\]
\[
\mathbb{E}(Y_t^2) = \frac{\partial^2 M_Y(0, t)}{\partial \xi^2} = \frac{\lambda}{2\lambda} \left( \mu^2 + \sigma^2 \right) (1 - e^{-2\lambda t}) + \mathbb{E}(Y_t)^2
\]
and hence
\[
\text{Var}(Y_t) = \frac{\lambda}{2\lambda} (\mu^2 + \sigma^2) (1 - e^{-2\lambda t})
\]
where $\mathbb{E}(X_i) = \mu$ and $\text{Var}(X_i) = \sigma^2$ for $i = 1, 2, \ldots$

\[\square\]

19. Let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, and let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables which are also independent of $N_t$. By defining the compound Poisson process $M_t$ as
\[
M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0
\]
show that its quadratic variation is
\[
\langle M_t, M_t \rangle = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( M_{i+1} - M_i \right)^2 = \sum_{i=1}^{N_t} X_i^2
\]
where $t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N}$ and deduce that $dM_t \cdot dM_t = X_t^2 dN_t$ where $X_t$ is the jump size if $N$ jumps at $t$.

**Solution:** From the definition of quadratic variation
\[
\langle M_t, M_t \rangle = \lim_{n \to \infty} \sum_{k=0}^{n-1} \left( M_{k+1} - M_k \right)^2
\]
where $t_k = tk/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N}$, we can set
\[
\langle M_t, M_t \rangle = \lim_{n \to \infty} \sum_{k=0}^{n-1} \left( \sum_{i=k+1}^{N_{t_{k+1}}} X_i - \sum_{i=k}^{N_{t_k}} X_i \right)^2 = \lim_{n \to \infty} \sum_{i=N_{t_0}}^{N_{t_n}} X_i^2 = \sum_{i=1}^{N_{t_n}} X_i^2
\]
since $N_{t_0} = N_0 = 0$. 

5.2.1 Properties of Poisson Process

From the definition
\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \left( M_{t_{i+1}} - M_{t_i} \right)^2 = \int_0^t (dM_s)^2 = \sum_{i=1}^{N_t} X_i^2
\]
and since
\[
\sum_{i=1}^{N_t} X_i^2 = \int_0^t X_u^2 \, dN_u
\]
where \( X_t \) is the jump size before a jump event at time \( t \), then by differentiating both sides with respect to \( t \) we finally have
\[
dM_t \cdot dM_t = X_t^2 \, dN_t
\]
where \( X_t \) is the jump size if \( N_j \) jumps at \( t \).

\[\square\]

20. Let \( \{ N_t : t \geq 0 \} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \) and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \). By defining the compound Poisson process \( M_t \) as

\[
M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0
\]
show that the cross-variation between \( M_t \) and \( N_t \) is

\[
\langle M, N \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( M_{t_{i+1}} - M_{t_i} \right) \left( N_{t_{i+1}} - N_{t_i} \right) = \sum_{i=1}^{N_t} X_i
\]
where \( t_i = it/n, \quad 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, \quad n \in \mathbb{N} \) and deduce that \( dM_t \cdot dN_t = X_t \, dN_t \).

**Solution:** From the definition of cross-variation, for \( t_k = kt/n, \quad 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, \quad n \in \mathbb{N} \) and since we can also write \( N_t = \sum_{i=1}^{N_t} 1 \),

\[
\langle M, N \rangle_t = \lim_{n \to \infty} \sum_{k=0}^{n-1} \left( M_{t_{k+1}} - M_{t_k} \right) \left( N_{t_{k+1}} - N_{t_k} \right)
\]

\[
= \lim_{n \to \infty} \sum_{k=0}^{n-1} \left( \sum_{i=1}^{N_{t_{k+1}}} - \sum_{i=1}^{N_{t_k}} \right) \left( \sum_{i=1}^{N_{t_{k+1}}} - \sum_{i=1}^{N_{t_k}} \right)
\]

\[
= \lim_{n \to \infty} \sum_{i=N_{t_0}}^{N_t} X_i \cdot 1
\]

\[
= \sum_{i=1}^{N_t} X_i
\]
since \( N_{t_0} = N_0 = 0 \).

Given
\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \left( M_{t_{i+1}} - M_{t_i} \right) \left( N_{t_{i+1}} - N_{t_i} \right) = \int_0^t dM_s \cdot dN_s = \sum_{i=1}^{N_t} X_i
\]
and because we can define
\[ \sum_{i=1}^{N_t} X_i = \int_0^t X_s dN_t \]
where \( X_i \) is the jump size before a jump event at time \( t \), then by differentiating the integrals with respect to \( t \) we have \( dM_t \cdot dN_t = X_t dN_t \) where \( X_t \) is the jump size if \( N \) jumps at \( t \).

\[ \square \]

21. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \). Show that its quadratic variation is
\[ \langle N, N \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( N_{i+1} - N_i \right)^2 = N_t \]
where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N} \) and deduce that \( dN_t \cdot dN_t = dN_t \).

**Solution:** Since \( N_t \) is a counting process, we can define the compound Poisson process \( M_t \) as
\[ M_t = \sum_{i=1}^{N_t} X_i \]
where \( X_i \) is a sequence of independent and identically distributed random variables which are also independent of \( N_t \). Using the results in Problem 5.2.1.19 (page 274),
\[ \langle M, M \rangle_t = \sum_{i=1}^{N_t} X_i^2. \]
By setting \( X_i = 1 \) for all \( i = 1, 2, \ldots, N_t \), then \( M_t = N_t \) which implies
\[ \langle N, N \rangle_t = \sum_{i=1}^{N_t} 1 = N_t. \]
By definition,
\[ \langle N, N \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( N_{i+1} - N_i \right)^2 \]
where \( t_j = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N} \) and since
\[ \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( N_{i+1} - N_i \right)^2 = \int_0^t (dN_t)^2 = N_t \]
then, by differentiating both sides with respect to \( t \), we have \( dN_t \cdot dN_t = dN_t \).

\[ \square \]
5.2.1 Properties of Poisson Process

22. Let \( \{W_t : t \geq 0\} \) be a standard Wiener process and \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Show that the cross-variation between \( W_t \) and \( N_t \) is

\[
\langle W, N \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (N_{t_{i+1}} - N_{t_i}) = 0
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N} \) and deduce that \( dW_t \cdot dN_t = 0 \).

**Solution:** Since

\[
\left| \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (N_{t_{i+1}} - N_{t_i}) \right| \leq \max_{0 \leq k \leq n-1} |W_{t_k} - W_{t_{k+1}}| \left| \sum_{i=0}^{n-1} (N_{t_{i+1}} - N_{t_i}) \right|
\]

and because \( W_t \) is continuous, we have

\[
\lim_{n \to \infty} \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| = 0.
\]

Therefore, we conclude that

\[
\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (N_{t_{i+1}} - N_{t_i}) \right| \leq 0
\]

and hence

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (N_{t_{i+1}} - N_{t_i}) = 0.
\]

Finally, because

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i}) (N_{t_{i+1}} - N_{t_i}) = \int_0^t dW_s \cdot dN_s = 0
\]

then, by differentiating both sides with respect to \( t \), we can deduce that \( dW_t \cdot dN_t = 0 \).

\( \square \)

23. Let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t \) and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \). By defining the compound Poisson process \( M_t \) as

\[
M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0
\]
show that the cross-variation between \( W_t \) and \( M_t \) is

\[
\langle W, M \rangle_t = \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) = 0
\]

where \( t_i = it/n, 0 = t_0 < t_1 < t_2 < \ldots < t_n = t, n \in \mathbb{N} \) and deduce that \( dW_t \cdot dM_t = 0 \).

**Solution:** Since we can write

\[
\left| \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) \right| \leq \max_{0 \leq k \leq n-1} \left| \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) \right|
\]

and because \( W_t \) is continuous, we have

\[
\lim_{n \to \infty} \max_{0 \leq k \leq n-1} \left| W_{t_{k+1}} - W_{t_k} \right| = 0.
\]

Thus, we conclude that

\[
\left| \lim_{n \to \infty} \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) \right| \leq 0
\]

or

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) = 0.
\]

Finally, because

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \left( W_{t_{i+1}} - W_{t_i} \right) \left( M_{t_{i+1}} - M_{t_i} \right) = \int_0^t dW_s \cdot dM_s = 0
\]

then, by differentiating both sides with respect to \( t \), we can deduce that \( dW_t \cdot dM_t = 0 \). \( \square \)

24. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) and \( \{W_t : t \geq 0\} \) be a standard Wiener process relative to the same filtration \( \mathcal{F}_t, t \geq 0 \). By considering the process

\[
Z_t = e^{u_1N_t + u_2W_t}
\]

where \( u_1, u_2 \in \mathbb{R} \), use Itō’s formula to find an SDE for \( Z_t \).

By setting \( m_t = \mathbb{E} (Z_t) \) show that solution to the SDE can be expressed as

\[
\frac{dm_t}{dt} - \left( (e^{u_1} - 1) \lambda + \frac{1}{2} u_2^2 \right) m_t = 0.
\]

Solve the differential equation to find \( \mathbb{E} (Z_t) \) and deduce that \( N_t \perp \perp W_t \).
5.2.1 Properties of Poisson Process

**Solution:** By expanding $Z_t$ using Taylor’s theorem and taking note that $(dW_t)^2 = dt$, $(dN_t)^2 = (dN_t)^3 = \cdots = dN_t^r, dW_t dN_t = 0,$

\[
dZ_t = \frac{\partial Z_t}{\partial N_t} dN_t + \frac{\partial Z_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 Z_t}{\partial N_t^2} (dN_t)^2 + \frac{1}{2} \frac{\partial^2 Z_t}{\partial W_t^2} (dW_t)^2 + \cdots
\]

\[+ \frac{\partial^3 Z_t}{\partial N_t \partial W_t} (dN_t dW_t) + \cdots + \frac{1}{3!} \frac{\partial^3 Z_t}{\partial N_t^3} (dN_t)^3 + \cdots
\]

\[= \left( u_1 + \frac{1}{2} u_1^2 + \frac{1}{3!} u_1^3 + \cdots \right) Z_t dN_t + u_2 Z_t dW_t + \frac{1}{2} u_2^2 Z_t dt
\]

\[= (e^{u_1} - 1) Z_t dN_t + u_2 Z_t dW_t + \frac{1}{2} u_2^2 Z_t dt.
\]

Taking integrals, we have

\[
\int_0^t dZ_s = (e^{u_1} - 1) \int_0^t Z_s dN_s + u_2 \int_0^t Z_s dW_s + \frac{1}{2} u_2^2 \int_0^t Z_s ds
\]

\[Z_t - 1 = (e^{u_1} - 1) \int_0^t Z_s d\hat{N}_s + u_2 \int_0^t Z_s dW_s + \left[ \lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 \right] \int_0^t Z_s ds
\]

where $Z_0 = 1$ and $\hat{N}_t = N_t - \lambda t$. Taking expectations,

\[
\mathbb{E}(Z_t) = 1 + \left[ \lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 \right] \int_0^t \mathbb{E}(Z_s) ds
\]

where since both $W_t$ and $\hat{N}_t$ are martingales we therefore have

\[
\mathbb{E} \left( \int_0^t Z_s dW_s \right) = 0 \quad \text{and} \quad \mathbb{E} \left( \int_0^t Z_s d\hat{N}_s \right) = 0.
\]

By differentiating the integral equation,

\[
\frac{d}{dt} \mathbb{E}(Z_t) = \left[ \lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 \right] \mathbb{E}(Z_t)
\]

or

\[
\frac{dm_t}{dt} - \left[ \lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 \right] m_t = 0
\]

where $m_t = \mathbb{E}(Z_t)$.

By setting the integrating factor to be $I = e^{-\int (\lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2) dt} = e^{-\lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 t}$ and multiplying the differential equation with $I$, we have

\[
\frac{d}{dt} \left( m_t e^{-\lambda (e^{u_1} - 1) - \frac{1}{2} u_2^2 t} \right) = 0 \quad \text{or} \quad e^{-\lambda (e^{u_1} - 1) - \frac{1}{2} u_2^2 t} \mathbb{E}(e^{u_1 N_t + u_2 W_t}) = C
\]

where $C$ is a constant. Since $\mathbb{E}(e^{u_1 N_0 + u_2 W_0}) = 1$ therefore $C = 1$, and hence we finally obtain

\[
\mathbb{E}(e^{u_1 N_t + u_2 W_t}) = e^{\lambda (e^{u_1} - 1) + \frac{1}{2} u_2^2 t}.
\]
Since the joint moment generating function of
\[ \mathbb{E} \left( e^{u_1 N_t + u_2 W_t} \right) = e^{\lambda t (e^{u_1} - 1)} \cdot e^{\frac{1}{2} u_2^2 t} \]
can be expressed as a product of the moment generating functions for \( N_t \) and \( W_t \), respectively, we can deduce that \( N_t \) and \( W_t \) are independent.

\( \square \)

25. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) and \( \{W_t : t \geq 0\} \) be a standard Wiener process relative to the same filtration \( \mathcal{F}_t \), \( t \geq 0 \). Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \) and \( W_t \). By defining a compound Poisson process \( M_t \) as
\[ M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \]
and by considering the process
\[ Z_t = e^{u_1 M_t + u_2 W_t}, \]
where \( u_1, u_2 \in \mathbb{R} \), use Itô’s formula to find an SDE for \( Z_t \).

By setting \( m_t = \mathbb{E} (Z_t) \) show that solution to the SDE can be expressed as
\[ \frac{dm_t}{dt} = -\left( (e^{\varphi X(u_1)} - 1) \lambda + \frac{1}{2} u_2^2 \right) m_t = 0 \]
where \( \varphi X(u_1) = \mathbb{E} (e^{u_1 X_i}) \) is the moment generating function of \( X_i \), which is the jump size if \( N \) jumps at time \( t \).

Finally, solve the differential equation to find \( \mathbb{E} (Z_t) \) and deduce that \( M_t \perp W_t \).

**Solution:** By expanding \( Z_t \) using Taylor’s theorem and taking note that \((dW_t)^2 = dt\), \( dM_t dW_t = 0 \), \((dN_t)^k = dN_t \) and \((dM_t)^k = X_i^k dN_t \) for \( k = 1, 2, \ldots \),

\[
\begin{align*}
\frac{dZ_t}{dt} &= \frac{\partial Z_t}{\partial M_t} dM_t + \frac{\partial Z_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 Z_t}{\partial M_t^2} (dM_t)^2 + \frac{1}{2} \frac{\partial^2 Z_t}{\partial W_t^2} (dW_t)^2 \\
&\quad \quad + \frac{\partial^2 Z_t}{\partial M_t \partial W_t} (dM_t dW_t) + \frac{1}{3!} \frac{\partial^3 Z_t}{\partial M_t^3} (dM_t)^3 + \cdots \\
&= (u_1 X_t + \frac{1}{2} u_1^2 X_t^2 + \frac{1}{3!} u_1^3 X_t^3 + \cdots) Z_t dN_t + u_2 Z_t dW_t + \frac{1}{2} u_2^2 Z_t dt \\
&= (e^{u_1 X_t} - 1) Z_t dN_t + u_2 Z_t dW_t + \frac{1}{2} u_2^2 Z_t dt.
\end{align*}
\]

Integrating, we have

\[ \int_0^t dZ_s = \int_0^t (e^{u_1 X_s} - 1) Z_s dN_s + u_2 \int_0^t Z_s dW_s + \frac{1}{2} u_2^2 \int_0^t Z_s ds \]

\[ Z_t - 1 = \int_0^t (e^{u_1 X_s} - 1) Z_s d\hat{N}_s + u_2 \int_0^t Z_s dW_s + \int_0^t \left[ \lambda (e^{u_1 X_s} - 1) + \frac{1}{2} u_2^2 \right] Z_s ds \]
5.2.2 Jump Diffusion Process

where $Z_0 = 1$ and $\hat{N}_t = N_t - \lambda t$. Taking expectations and because the jump size variable $X_t$ is independent of $N_t$ and $W_t$, we have

$$E(Z_t) = 1 + \left[ \lambda \left( E(e^{\mu_1 X_t}) - 1 \right) + \frac{1}{2} \mu_2^2 \right] \int_0^t E(Z_s) \, ds$$

since $E\left( \int_0^t Z_s dW_s \right) = 0$ and $E\left( \int_0^t Z_s d\hat{N}_s \right) = 0$.

By differentiating the integral equation,

$$\frac{d}{dt} E(Z_t) = \left[ \lambda \left( E(e^{\mu_1 X_t}) - 1 \right) + \frac{1}{2} \mu_2^2 \right] E(Z_t)$$

or

$$\frac{dm_t}{dt} - \left[ \lambda \left( \varphi_X(u_1) - 1 \right) + \frac{1}{2} \mu_2^2 \right] m_t = 0$$

where $m_t = E(Z_t)$ and $\varphi_X(u_1) = E(e^{\mu_1 X_t})$.

By setting the integrating factor to be $I = e^{-\int (\lambda \left( \varphi_X(u_1) - 1 \right) + \frac{1}{2} \mu_2^2) \, dt}$ and multiplying the differential equation with $I$, we have

$$\frac{d}{dt} \left( m_t e^{-\lambda \left( \varphi_X(u_1) - 1 \right) - \frac{1}{2} \mu_2^2 t} \right) = 0 \quad \text{or} \quad e^{-\lambda \left( \varphi_X(u_1) - 1 \right) - \frac{1}{2} \mu_2^2 t} E(e^{\mu_1 M_t + \mu_2 W_t}) = C$$

where $C$ is a constant. Since $E(e^{\mu_1 M_0 + \mu_2 W_0}) = 1$ therefore $C = 1$, and hence we finally obtain

$$E\left( e^{\mu_1 M_t + \mu_2 W_t} \right) = e^{\lambda \left( \varphi_X(u_1) - 1 \right) + \frac{1}{2} \mu_2^2 t}.$$

Since the joint moment generating function of

$$E\left( e^{\mu_1 M_t + \mu_2 W_t} \right) = e^{\lambda \left( \varphi_X(u_1) - 1 \right) + \frac{1}{2} \mu_2^2 t}$$

can be expressed as a product of the moment generating functions for $M_t$ and $W_t$, respectively, we can deduce that $M_t$ and $W_t$ are independent.

\[\Box\]

5.2.2 Jump Diffusion Process

1. **Pure Jump Process.** Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ relative to the filtration $\mathcal{F}_t$, $t \geq 0$. Suppose $S_t$ follows a pure jump process

$$\frac{dS_t}{S_{t-}} = (J_t - 1) dN_t$$

where $J_t$ is the jump size variable if $N$ jumps at time $t$, $J_t \perp N_t$ and

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt. \end{cases}$$

Explain why the term in $dN_t$ is $J_t - 1$ and not $J_t$. 

Show that the above differential equation can also be written as
\[
\frac{dS_t}{S_{t^-}} = dM_t
\]
where \( M_t = \sum_{i=1}^{N_t} (J_i - 1) \) is a compound Poisson process such that \( J_i, i = 1, 2, \ldots \) is a sequence of independent and identically distributed random variables which are also independent of \( N_t \).

**Solution:** Let \( S_{t^-} \) be the value of \( S_t \) just before a jump and assume there occurs an instantaneous jump (i.e., \( dN_t = 1 \)) in which \( S_t \) changes from \( S_{t^-} \) to \( J_t S_{t^-} \) where \( J_t \) is the jump size. Thus,
\[
dS_t = J_t S_{t^-} - S_{t^-} = (J_t - 1)S_{t^-}
\]
or
\[
\frac{dS_t}{S_{t^-}} = (J_t - 1).
\]
Therefore, we can write
\[
\frac{dS_t}{S_{t^-}} = (J_t - 1)dN_t
\]
where
\[
dN_t = \begin{cases} 
1 & \text{with probability } \lambda dt \\
0 & \text{with probability } 1 - \lambda dt.
\end{cases}
\]
Given the compound Poisson process
\[
M_t = \sum_{i=1}^{N_t} (J_i - 1)
\]
we let \( M_t^- \) be the value of \( M_t \) just before a jump event. If \( N \) jumps at time \( t \) then
\[
dM_t = M_t - M_t^- = J_t - 1.
\]
Thus, in general, we can write
\[
dM_t = (J_t - 1)dN_t
\]
which implies the pure jump process can also be expressed as \( \frac{dS_t}{S_{t^-}} = dM_t \).

2. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) relative to the filtration \( \mathcal{F}_t, t \geq 0 \). Suppose \( S_t \) follows a pure jump process
\[
\frac{dS_t}{S_{t^-}} = (J_t - 1)dN_t
\]
where $J_t$ is the jump size variable if $N$ jumps at time $t$ and

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda \, dt \\ 0 & \text{with probability } 1 - \lambda \, dt. \end{cases}$$

Assume $J_t$ follows a lognormal distribution such that $\log J_t \sim \mathcal{N}(\mu_J, \sigma_J^2)$ and $J_t$ is also independent of $N_t$. By applying Itô’s formula on $\log S_t$ and taking integrals show that for $t < T$,

$$S_T = S_t \prod_{i=1}^{N_{T-t}} J_i$$

provided $J_i \in (0, 2]$, $i = 1, 2, \ldots$ is a sequence of independent and identically distributed jump size random variables which are independent of $N_t$.

Given $S_t$ and $N_{T-t} = n$, show that $S_T$ follows a lognormal distribution with mean

$$\mathbb{E} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t e^{\mu_J + \frac{1}{2} \sigma_J^2}$$

and variance

$$\text{Var} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t^2 (e^{\sigma_J^2} - 1)e^{\mu_J + \frac{1}{2} \sigma_J^2}.$$ 

Finally, given only $S_t$, show that

$$\mathbb{E} \left( S_T \mid S_t \right) = S_t \exp \left\{ \lambda \left( e^{\mu_J + \frac{1}{2} \sigma_J^2} - 1 \right) (T - t) \right\}$$

and

$$\text{Var} \left( S_T \mid S_t \right) = S_t^2 \left[ \exp \left\{ \lambda (T - t) \left( e^{2(\mu_J + \sigma_J^2)} - 1 \right) \right\} - \exp \left\{ 2\lambda (T - t) \left( e^{\mu_J + \frac{1}{2} \sigma_J^2} - 1 \right) \right\} \right].$$

**Solution:** By letting $S_{t-}$ denote the value of $S_t$ before a jump event, and expanding $d(\log S_t)$ using Taylor’s theorem and taking note that $dN_t \cdot dN_t = dN_t$,

$$d(\log S_t) = \frac{dS_t}{S_{t-}} - \frac{1}{2} \left( \frac{dS_t}{S_{t-}} \right)^2 + \frac{1}{3} \left( \frac{dS_t}{S_{t-}} \right)^3 - \frac{1}{4} \left( \frac{dS_t}{S_{t-}} \right)^4 + \ldots$$

$$= \left\{ (J_t - 1) - \frac{1}{2} (J_t - 1)^2 + \frac{1}{3} (J_t - 1)^3 - \frac{1}{4} (J_t - 1)^4 + \ldots \right\} dN_t$$

$$= \log J_t \, dN_t$$

provided $-1 < J_t - 1 \leq 1$ or $0 < J_t \leq 2$.

By taking integrals, we have

$$\int_t^T d(\log S_u) = \int_t^T \log J_u \, dN_u$$

$$\log \left( \frac{S_T}{S_t} \right) = \sum_{i=1}^{N_{T-t}} \log J_i$$
5.2.2 Jump Diffusion Process

or

\[ S_T = S_t \prod_{i=1}^{N_{T-t}} J_i \]

where \( J_i \in (0, 2] \) is the jump size occurring at time instant \( t_i \) and \( N_{T-t} = N_T - N_t \) is the total number of jumps in the time interval \((t, T]\).

Since \( \log J_i \sim N(\mu_J, \sigma_J^2) \), \( i = 1, 2, \ldots, N_{T-t} \) are independent and identically distributed, and conditional on \( S_t \) and \( N_{T-t} = n \),

\[ \log S_T = \log S_t + \sum_{i=1}^{N_{T-t}} \log J_i \]

follows a normal distribution. Therefore, the mean and variance of \( S_T \) are

\[ \mathbb{E}(S_T | S_t, N_{T-t} = n) = S_t e^{\mu_J + \frac{1}{2} \sigma_J^2} \]

and

\[ \text{Var}(S_T | S_t, N_{T-t} = n) = S_t^2 \left( e^{2\sigma_J^2} - 1 \right) e^{(2\mu_J + \sigma_J^2)} \]

respectively.

Finally, conditional only on \( S_t \), by definition

\[
\mathbb{E}(S_T | S_t) = \mathbb{E} \left( S_t \prod_{i=1}^{N_{T-t}} J_i | S_t \right) \\
= S_t \mathbb{E} \left( \prod_{i=1}^{N_{T-t}} J_i | S_t \right) \\
= S_t \mathbb{E} \left( \sum_{i=1}^{N_{T-t}} \log J_i \right) \\
= S_t \mathbb{E} \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} \right) .
\]

By applying the tower property and from Problem 5.2.1.12 (page 264), we have

\[
\mathbb{E}(S_T | S_t) = S_t \mathbb{E} \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} | N_{T-t} \right) \\
= S_t \exp \left\{ \lambda(T-t) \left[ \mathbb{E}(e^{\log J}) - 1 \right] \right\} \\
= S_t \exp \left\{ \lambda(T-t) \left[ \mathbb{E}(J) - 1 \right] \right\} .
\]

Since \( \mathbb{E}(J) = e^{\mu_J + \frac{1}{2} \sigma_J^2} \) therefore

\[ \mathbb{E}(S_T | S_t) = S_t \exp \left\{ \lambda \left( e^{\mu_J + \frac{1}{2} \sigma_J^2} - 1 \right) (T-t) \right\} . \]
For the case of variance of $S_T$ conditional on $S_t$, by definition

$$\text{Var}(S_T|S_t) = \text{Var} \left( S_t \prod_{i=1}^{N_{T-t}} J_i \bigg| S_t \right)$$

$$= S_t^2 \text{Var} \left( \prod_{i=1}^{N_{T-t}} J_i \right)$$

$$= S_t^2 \text{Var} \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} \right)$$

$$= S_t^2 \left\{ \mathbb{E} \left( e^{2 \sum_{i=1}^{N_{T-t}} \log J_i} \right) - \left( \mathbb{E} \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} \right) \right)^2 \right\}$$

By applying the tower property and from Problem 5.2.1.12 (page 264), we have

$$\text{Var}(S_T|S_t) = S_t^2 \mathbb{E} \left[ \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} \right) \right] - S_t^2 \left\{ \mathbb{E} \left( e^{\sum_{i=1}^{N_{T-t}} \log J_i} \right) \right\}^2$$

$$= S_t^2 \exp \left\{ \lambda (T-t) \left[ \mathbb{E} \left( J_t^2 \right) - 1 \right] \right\} - S_t^2 \exp \left\{ 2 \lambda (T-t) \left[ \mathbb{E} \left( J_t \right) - 1 \right] \right\}.$$

Since $\mathbb{E}(J_t) = e^{\mu J + \frac{1}{2} \sigma^2 J}$ and $\mathbb{E}(J_t^2) = e^{2(\mu J + \sigma^2 J)}$, therefore

$$\text{Var}(S_T|S_t) = S_t^2 \left[ \exp \left\{ \lambda (T-t) \left( e^{2(\mu J + \sigma^2 J)} - 1 \right) \right\} - \exp \left\{ 2 \lambda (T-t) \left( e^{\mu J + \sigma^2 J} - 1 \right) \right\} \right].$$

\[\square\]

3. **Merton’s Model.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : t \geq 0\}$ be a standard Wiener process relative to the same filtration $\mathcal{F}_t$, $t \geq 0$. Suppose $S_t$ follows a jump diffusion process with the following SDE

$$\frac{dS_t}{S_t} = (\mu - D) dt + \sigma dW_t + (J_t - 1)dN_t$$

where

$$dN_t = \begin{cases} 
1 & \text{with probability } \lambda dt \\
0 & \text{with probability } 1 - \lambda dt 
\end{cases}$$

with constants $\mu$, $D$, and $\sigma$ being the drift, continuous dividend yield and volatility, respectively, and $J_t$ the jump variable such that $\log J_t \sim \mathcal{N}(\mu J, \sigma^2 J)$. Assume that $J_t$, $W_t$, and $N_t$ are mutually independent. In addition, we also let $J_i$, $i = 1, 2, \ldots$ be a sequence of independent and identically distributed jump size random variables which are also independent of $N_t$ and $W_t$. 


5.2.2 Jump Diffusion Process

By applying Itô’s formula on $\log S_t$ and taking integrals show that for $t < T$,

$$S_T = S_0 e^{(\mu-D)T + \sigma W_T} \prod_{j=1}^{N_T} J_t$$

provided $J_t \in (0, 2]$ and $W_T \sim N(0, T-t)$.

Given $S_t$ and $N_{T-t} = n$, show that $S_T$ follows a lognormal distribution with mean

$$\mathbb{E} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t e^{(\mu-D)(T-t) + n(\mu_t + \frac{1}{2} \sigma^2_t)}$$

and variance

$$\text{Var} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t^2 \left( e^{\sigma^2 (T-t) + n \sigma^2} - 1 \right) e^{2(\mu-D)(T-t) + n(2\mu_t + \sigma^2_t)}.$$

Finally, conditional only on $S_t$, show that

$$\mathbb{E} \left( S_T \mid S_t \right) = S_t e^{(\mu-D)(T-t)} \exp \left\{ \lambda \left( e^{\mu_t + \frac{1}{2} \sigma^2} - 1 \right) (T-t) \right\}$$

and

$$\text{Var} \left( S_T \mid S_t \right) = S_t^2 e^{2(\mu-D)(T-t)} \left( e^{\sigma^2 (T-t)} - 1 \right) \left[ \exp \left\{ \lambda(T-t) \left( e^{2\mu_t + \sigma^2} - 1 \right) \right\} - \exp \left\{ 2\lambda(T-t) \left( e^{\mu_t + \frac{1}{2} \sigma^2} - 1 \right) \right\} \right].$$

**Solution:** By letting $S_{t-}$ denote the value of $S_t$ before a jump event, and expanding $d(\log S_t)$ using both Taylor’s theorem and Itô’s lemma,

$$d(\log S_t) = \frac{dS_t}{S_{t-}} - \frac{1}{2} \left( \frac{dS_t}{S_{t-}} \right)^2 + \frac{1}{3} \left( \frac{dS_t}{S_{t-}} \right)^3 - \frac{1}{4} \left( \frac{dS_t}{S_{t-}} \right)^4 + \ldots$$

$$= (\mu - D) dt + \sigma dW_t + (J_t - 1)dN_t - \frac{1}{2}(\sigma^2 dt + (J_t - 1)^2 dN_t)$$

$$+ \frac{1}{3}(J_t - 1)^3 dN_t - \frac{1}{4}(J_t - 1)^4 dN_t + \ldots$$

$$= \left( \mu - D - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t$$

$$+ \left\{ (J_t - 1) - \frac{1}{2}(J_t - 1)^2 + \frac{1}{3}(J_t - 1)^3 - \frac{1}{4}(J_t - 1)^4 + \ldots \right\} dN_t$$

$$= \left( \mu - D - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t + \log J_t dN_t$$

provided $-1 < J_t - 1 \leq 1$ or $0 < J_t \leq 2$.

By taking integrals, we have

$$\int_t^T d(\log S_u) = \int_t^T \left( \mu - D - \frac{1}{2} \sigma^2 \right) du + \int_t^T \sigma dW_u + \int_t^T \log J_u dN_u$$
5.2.2 Jump Diffusion Process

\[
\log \left( \frac{S_T}{S_t} \right) = \left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma W_{T-t} + \sum_{i=1}^{N_{T-t}} \log J_i
\]

or

\[
S_T = S_t e^{\left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma W_{T-t}} \prod_{i=1}^{N_{T-t}} J_i
\]

where \( J_i \in (0, 2) \) is the random jump size occurring at time \( t_i \) and \( N_{T-t} = N_T - N_t \) is the total number of jumps in the time interval \((t, T)\).

Since \( W_{T-t} \sim \mathcal{N}(0, T-t), \log J_i \sim \mathcal{N}(\mu_J, \sigma_J^2), i = 1, 2, \ldots \) and by independence,

\[
\log S_T | \{ S_t, N_{T-t} = n \} \sim \mathcal{N} \left[ \log S_t + \left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) + n \mu_J, \sigma^2(T-t) + n \sigma_J^2 \right].
\]

Therefore, conditional on \( S_t \) and \( N_{T-t} = n \), \( S_T \) follows a lognormal distribution with mean

\[
\mathbb{E} \left( S_T | S_t, N_{T-t} = n \right) = S_t e^{(\mu - D)(T-t) + n(\mu_J + \frac{1}{2} \sigma_J^2)}
\]

and variance

\[
\text{Var} \left( S_T | S_t, N_{T-t} = n \right) = S_t^2 \left( e^{\sigma^2(T-t) + n \sigma_J^2} - 1 \right) e^{2(\mu - D)(T-t) + n(2\mu_J + \sigma_J^2)}.
\]

Finally, given only \( S_t \) and from the mutual independence of a Wiener process and a compound Poisson process, we have

\[
\mathbb{E}(S_T | S_t) = \mathbb{E} \left[ S_t e^{\left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma W_{T-t}} \prod_{i=1}^{N_{T-t}} J_i | S_t \right]
\]

\[
= S_t e^{\left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) \mathbb{E}(e^{\sigma W_{T-t}})} \mathbb{E} \left( \prod_{i=1}^{N_{T-t}} J_i \right)
\]

\[
= S_t e^{(\mu - D)(T-t)} \mathbb{E} \left( \prod_{i=1}^{N_{T-t}} J_i \right).
\]

Since \( \mathbb{E}(e^{\sigma W_{T-t}}) = e^{\frac{1}{2} \sigma^2(T-t)} \), from Problem 5.2.2.2 (page 282) we therefore have

\[
\mathbb{E}(S_T | S_t) = S_t e^{(\mu - D)(T-t)} \exp \left\{ \lambda \left( e^{\mu_J + \frac{1}{2} \sigma_J^2} - 1 \right) (T-t) \right\}.
\]

As for the variance of \( S_T \) conditional on \( S_t \), we can write

\[
\text{Var}(S_T | S_t) = \text{Var} \left[ S_t e^{\left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t) + \sigma W_{T-t}} \prod_{i=1}^{N_{T-t}} J_i | S_t \right]
\]

\[
= S_t^2 e^{2\left( \mu - D - \frac{1}{2} \sigma^2 \right) (T-t)} \text{Var}(e^{\sigma W_{T-t}}) \mathbb{E} \left( \prod_{i=1}^{N_{T-t}} J_i \right).
\]
Since $\text{Var}(e^{\lambda W_{t-T}}) = (e^{\lambda^2(T-t)} - 1)e^{\lambda^2(T-t)}$ and using the results from Problem 5.2.2.3 (page 285), we finally have

$$\text{Var}(S_T|S_t) = S_t^2 e^{2(\mu-D)(T-t)}(e^{\lambda^2(T-t)} - 1) \left[ \exp \left\{ \lambda(T-t) \left( e^{2(\mu_D+\sigma^2_T)} - 1 \right) \right\} \right.$$

$$\left. - \exp \left\{ 2\lambda(T-t) \left( e^{\mu_D+\frac{1}{2}\sigma^2_T} - 1 \right) \right\} \right].$$

\[\square\]

4. \textit{Ornstein–Uhlenbeck Process with Jumps.} Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{N_t : t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : t \geq 0\}$ be a standard Wiener process relative to the same filtration $\mathcal{F}_t$, $t \geq 0$. Suppose $S_t$ follows an Ornstein–Uhlenbeck process with jumps of the form

$$dS_t = \kappa(\theta - S_t)\, dt + \sigma dW_t + \log J_t dN_t,$$

where

$$dN_t = \begin{cases} 
1 & \text{with probability } \lambda \, dt \\
0 & \text{with probability } 1 - \lambda \, dt 
\end{cases}$$

with $\kappa$, $\theta$ and $\sigma$ being constant parameters such that $\kappa$ is the mean-reversion rate, $\theta$ is the long-term mean and $\sigma$ is the volatility. The random variable $J_t$ is the jump amplitude such that $\log J_t \sim N(\mu_J, \sigma^2_J)$, and $W_t$, $N_t$, and $J_t$ are mutually independent. Suppose the sequence of jump amplitudes $J_i$, $i = 1, 2, \ldots$ is independent and identically distributed, and also independent of $N_t$ and $W_t$.

Using Ito's lemma on $e^{\kappa S_t}$ show that for $t < T$,

$$S_T = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \int_t^T \sigma e^{-\kappa(\tau-t)} \, dW_\tau + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i,$$

and conditional on $S_t$ and $N_{T-t} = n$,

$$\mathbb{E} \left( S_T | S_t, N_{T-t} = n \right) = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \eta \mu_J \left( \frac{\lambda}{\lambda - \kappa} \right) \left[ e^{-\kappa(T-t)} - e^{-\kappa(T-t)} \right] \left[ 1 - e^{-\lambda(T-T_i)} \right]^{n-1},$$

and

$$\text{Var}(S_T | S_t, N_{T-t} = n) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] + n \left( \mu_J^2 + \sigma_J^2 \right) \left( \frac{\lambda}{\lambda - \kappa} \right) \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{n-1} + n(n-1)\mu_J^2 \left( \frac{\lambda}{\lambda - \kappa} \right)^2 \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right]^2 \left[ 1 - e^{-\lambda(T-t)} \right]^{n-2} - n^2 \mu_J^2 \left( \frac{\lambda}{\lambda - \kappa} \right)^2 \left[ e^{-\kappa(T-t)} - e^{-\lambda(T-t)} \right]^2 \left[ 1 - e^{-\lambda(T-t)} \right]^{2(n-1)}.$$
5.2.2 Jump Diffusion Process

Finally, conditional on \( S_t \), show also that

\[
\mathbb{E} (S_T|S_t) = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\lambda \mu_J}{\kappa} \left[ 1 - e^{-\kappa(T-t)} \right]
\]

and

\[
\text{Var} (S_T|S_t) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] + \frac{\lambda}{2\kappa} \left( \mu_J^2 + \sigma_J^2 \right) \left[ 1 - e^{-\kappa(T-t)} \right].
\]

What is the distribution of \( S_T \) given \( S_t \)?

**Solution:** In order to solve the jump diffusion process we note that

\[
d(e^{\kappa t} S_t) = \kappa e^{\kappa t} S_t \, dt + e^{\kappa t} dS_t + \kappa^2 e^{\kappa t} S_t (dt)^2 + \ldots
\]

and using Itô’s lemma,

\[
d(e^{\kappa t} S_t) = \kappa \theta e^{\kappa t} dt + \sigma e^{\kappa t} dW_t + e^{\kappa t} \log J_t dN_t.
\]

For \( t < T \) the solution is represented by

\[
S_T = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \int_t^T \sigma e^{-\kappa(T-u)} dW_u + \int_t^T e^{-\kappa(T-u)} \log J_u dN_u.
\]

Since

\[
\int_t^T e^{-\kappa(T-u)} \log J_u dN_u = \sum_{i=1}^{N_T} e^{-\kappa(T-t_i)} \log J_i = \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i
\]

where \( J_i \) is the random jump size occurring at time \( t_i \) and \( N_{T-t} = N_T - N_t \) is the total number of jumps in the time interval \((t, T]\), therefore

\[
S_T = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \int_t^T \sigma e^{-\kappa(T-u)} dW_u + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i.
\]

Given \( S_t \), from the properties of Itô’s integral we have

\[
\mathbb{E} \left[ \int_t^T \sigma e^{-\kappa(T-u)} dW_u \right] = 0
\]

and

\[
\mathbb{E} \left[ \left( \int_t^T \sigma e^{-\kappa(T-u)} dW_u \right)^2 \right] = \mathbb{E} \left[ \int_t^T \sigma^2 e^{-2\kappa(T-u)} \, du \right] = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right].
\]

Thus, given \( S_t \) and \( N_{T-t} = n \), from Problem 5.2.1.17 (page 269) we can easily obtain

\[
\mathbb{E} \left[ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i \middle| S_t, N_{T-t} = n \right] = n \mu_J \frac{\lambda}{\lambda - \kappa} \left[ e^{-\kappa(T-t)} - e^{-\kappa(T-t_i)} \right] \left[ 1 - e^{-\kappa(T-t_i)} \right]^{n-1}
\]
and
\[
\text{Var} \left[ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t-i)} \log J_i \bigg| S_t, N_{T-t} = n \right] \\
= n(\mu_j^2 + \sigma_j^2) \left( \frac{\lambda}{\lambda - 2\kappa} \right) \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{n-1} \\
+ n(n-1)\mu_j^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \\
\times \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{n-2} \\
- n^2 \mu_j^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \\
\times \left[ e^{-\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{2(n-1)}.
\]

Therefore,
\[
\mathbb{E} \left( S_T | S_t, N_{T-t} = n \right) = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] \\
+ n \mu_j \left( \frac{\lambda}{\lambda - \kappa} \right) \left[ e^{-\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{n-1}
\]
and
\[
\text{Var} \left( S_T | S_t, N_{T-t} = n \right) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \\
+ n(\mu_j^2 + \sigma_j^2) \left( \frac{\lambda}{\lambda - 2\kappa} \right) \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right] \left[ 1 - e^{-\lambda(T-t)} \right]^{n-1} \\
+ n(n-1)\mu_j^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \left[ e^{-2\kappa(T-t)} - e^{-\lambda(T-t)} \right]^{2} \left[ 1 - e^{-\lambda(T-t)} \right]^{n-2} \\
- n^2 \mu_j^2 \left( \frac{\lambda}{\lambda - 2\kappa} \right)^2 \left[ e^{-\kappa(T-t)} - e^{-\lambda(T-t)} \right]^{2} \left[ 1 - e^{-\lambda(T-t)} \right]^{2(n-1)}.
\]

Finally, conditional on \( S_t \), from Problem 5.2.1.18 (page 272) we have
\[
\mathbb{E} \left[ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t-i)} \log J_i \right] = \frac{\mu_j}{\kappa} \left[ 1 - e^{-\kappa(T-t)} \right]
\]
and
\[
\text{Var} \left[ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t-i)} \log J_i \right] = \frac{\lambda}{2\kappa} \left( \mu_j^2 + \sigma_j^2 \right) \left[ 1 - e^{-2\kappa(T-t)} \right].
\]

Therefore,
\[
\mathbb{E} \left( S_T | S_t \right) = S_t e^{-\kappa(T-t)} + \theta \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\lambda \mu_j}{\kappa} \left[ 1 - e^{-\kappa(T-t)} \right]
\]
and
\[
\text{Var} \left( S_T | S_t \right) = \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] + \frac{\lambda}{2\kappa} \left( \mu_j^2 + \sigma_j^2 \right) \left[ 1 - e^{-\kappa(T-t)} \right].
\]

Without the jump component, \( S_T \) given \( S_t \) is normally distributed (see Problem 3.2.2.10, page 132). With the inclusion of a jump component, the term \( \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t-i)} \log J_i \)
5.2.2 Jump Diffusion Process

consists of summing over the product of random jump times in the exponent and jump amplitudes where we do not have an explicit expression for the distribution. Thus, the distribution of \( S_T \) conditional on \( S_t \) is not known.

\[\Box\]

5. Geometric Mean-Reverting Jump Diffusion Model. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \( \{N_t : t \geq 0\} \) be a Poisson process with intensity \( \lambda > 0 \) and \( \{W_t : t \geq 0\} \) be a standard Wiener process relative to the same filtration \( \mathcal{F}_t, t \geq 0 \). Suppose \( S_t \) follows a geometric mean-reverting jump diffusion process of the form

\[
dS_t = \kappa(\theta - \log S_t) \, dt + \sigma dW_t + (J_t - 1) \, dN_t
\]

where

\[
dN_t = \begin{cases} 1 & \text{with probability } \lambda \, dt \\ 0 & \text{with probability } 1 - \lambda \, dt \end{cases}
\]

with \( \kappa, \theta \) and \( \sigma \) being constant parameters. Here, \( \kappa \) is the mean-reversion rate, \( \theta \) is the long-term mean and \( \sigma \) is the volatility. The random variable \( J_t \) is the jump amplitude such that \( \log J_t \sim \mathcal{N}(\mu_J, \sigma_J^2) \), and \( W_t, N_t \) and \( J_t \) are mutually independent. Suppose the sequence of jump amplitudes \( J_i, i = 1, 2, \ldots \) is independent and identically distributed, and also independent of \( N_t \) and \( W_t \).

By applying Itô’s formula on \( X_t = \log S_t \), show that

\[
dx_t = \left( \kappa(\theta - X_t) - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t + \log J_t dN_t.
\]

Using Itô’s lemma on \( e^{\kappa X_t} \) and taking integrals show that for \( t < T \),

\[
\log S_T = (\log S_t) e^{-\kappa(T-t)} + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right]
+ \int_t^T \sigma e^{-\kappa(s-t)} dW_s + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_t.
\]

Conditional on \( S_t \) and \( N_{T-t} = n \), show that

\[
\mathbb{E} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t e^{-\kappa(T-t)} \exp \left\{ \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\sigma^2}{4\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\}
\times \left[ \lambda \int_t^T e^{\mu_J e^{-\kappa s} + \frac{1}{2} \sigma_J^2 e^{-2\kappa s - \lambda(T-t_s)}} ds \right]^n
\]

and

\[
\text{Var} \left( S_T \mid S_t, N_{T-t} = n \right) = S_t^2 e^{-\kappa(T-t)} \exp \left\{ 2 \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] \right\}
\times \left[ \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\} - 1 \right] \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\}
\]
5.2.2 Jump Diffusion Process

Finally, given only $S_t$, show that

$$\mathbb{E}(S_T|S_t) = S_t \exp \left\{ e^{-\kappa(T-t)} + \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\sigma^2}{4\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\}$$

and

$$\text{Var}(S_T|S_t) = S_t^2 \exp \left\{ 2e^{-\kappa(T-t)} + 2 \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] \right\}$$

$$\times \left[ \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-\kappa(T-t)} \right] \right\} - 1 \right] \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\} e^{-\lambda(T-t)}$$

$$\times \left[ \exp \left\{ \lambda \int_t^T e^{\mu_j e^{-\kappa s} + \frac{1}{2} \sigma_j^2 e^{-2\kappa s}} \, ds - \lambda(T-t) \right\} \right].$$

**Solution:** We first let $S_{t^-}$ denote the value of $S_t$ before a jump event and by expanding $d(\log S_t)$ using Taylor’s theorem and subsequently applying Itô’s lemma, we have

$$d(\log S_t) = \frac{dS_t}{S_{t^-}} - \frac{1}{2} \left( \frac{dS_t}{S_{t^-}} \right)^2 + \frac{1}{3} \left( \frac{dS_t}{S_{t^-}} \right)^3 - \frac{1}{4} \left( \frac{dS_t}{S_{t^-}} \right)^4 + \ldots$$

$$= \kappa(\theta - \log S_{t^-}) \, dt + \sigma dW_t + (J_t - 1) dN_t - \frac{1}{2} \sigma^2 \, dt + (J_t - 1)^2 dN_t$$

$$+ \frac{1}{3} (J_t - 1)^3 dN_t - \frac{1}{4} (J_t - 1)^4 dN_t + \ldots$$

$$= \left[ \kappa(\theta - \log S_{t^-}) - \frac{1}{2} \sigma^2 \right] \, dt + \sigma dW_t + \log J_t dN_t$$

provided the random jump amplitude $-1 < J_t - 1 \leq 1$ or $0 < J_t < 2$.

Setting $X_t = \log S_t$, we can redefine the above SDE into an Ornstein–Uhlenbeck process with jumps

$$dX_t = \kappa(\theta - X_{t^-}) \, dt + \sigma dW_t + \log J_t dN_t$$

and applying Taylor’s theorem and then Itô’s lemma on $d(e^{\xi t} X_t)$, we have

$$d(e^{\xi t} X_t) = \kappa e^{\xi t} X_{t^-} \, dt + e^{\xi t} dX_t + \kappa^2 e^{\xi t} X_{t^-} (dt)^2 + \ldots$$

$$= \kappa \theta e^{\xi t} \, dt + \sigma e^{\xi t} dW_t + e^{\xi t} \log J_t dN_t.$$
5.2.2 Jump Diffusion Process

Taking integrals for \( t < T \),
\[
\int_{t}^{T} d(e^{\xi X_s}) = \int_{t}^{T} \left( \kappa \theta e^{\xi s} - \frac{1}{2} \sigma^2 e^{2 \xi s} \right) ds + \int_{t}^{T} \sigma e^{\xi s} dW_s + \int_{t}^{T} e^{\xi s} \log J_s dN_s
\]
or
\[
X_T = X_t e^{e^{-\kappa(T-t)}} + \left( \theta - \frac{\sigma^2}{2 \kappa} \right) [1 - e^{-\kappa(T-t)}] + \int_{t}^{T} \sigma e^{-\kappa(s)} dW_s + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i.
\]
Substituting \( X_T = \log S_T \) and \( X_t = \log S_t \), we finally have
\[
\log S_T = (\log S_t) e^{-\kappa(T-t)} + \left( \theta - \frac{\sigma^2}{2 \kappa} \right) [1 - e^{-\kappa(T-t)}] + \int_{t}^{T} \sigma e^{-\kappa(s)} dW_s + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i.
\]
and
\[
S_T = S_t e^{-\kappa(T-t)} \exp \left\{ \left( \theta - \frac{\sigma^2}{2 \kappa} \right) [1 - e^{-\kappa(T-t)}] + \int_{t}^{T} \sigma e^{-\kappa(s)} dW_s + \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t_i)} \log J_i \right\}.
\]
To find the expectation and variance of \( S_T \) given \( S_t \), we note that from Itô’s integral
\[
\int_{t}^{T} \sigma e^{-\kappa(s)} dW_s \sim \mathcal{N} \left( 0, \frac{\sigma^2}{2 \kappa} [1 - e^{-2\kappa(T-t)}] \right)
\]
and hence
\[
\mathbb{E} \left[ \exp \left\{ \int_{t}^{T} \sigma e^{-\kappa(s)} dW_s \right\} \right] = \exp \left\{ \frac{\sigma^2}{4 \kappa} [1 - e^{-2\kappa(T-t)}] \right\}
\]
and
\[
\text{Var} \left[ \exp \left\{ \int_{t}^{T} \sigma e^{-\kappa(s)} dW_s \right\} \right] = \left[ \exp \left\{ \frac{\sigma^2}{2 \kappa} [1 - e^{-2\kappa(T-t)}] \right\} - 1 \right] \exp \left\{ \frac{\sigma^2}{2 \kappa} [1 - e^{-2\kappa(T-t)}] \right\}.
\]
Conditional on \( N_{T-t} = n \), from Problem 5.2.1.17 (page 269) and by setting \( \xi = 1 \), \( X = \log J \) and due to the independence of \( \log J_i, i = 1, 2, \ldots, n \), we can write
\[
\mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{n} e^{-\kappa(T-t_i)} \log J_i \right\} \right] = \prod_{i=1}^{n} \left\{ \int_{t}^{T} M_{\log J_i} (e^{-\kappa s}) \lambda e^{-\lambda (T-t-s)} ds \right\}
\]
\[
= \left[ \lambda \int_{t}^{T} e^{\mu J e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s} - \lambda(T-t-s)} ds \right]^{n}
\]
Finally, by treating $N_{T-t}$ as a random variable, from Problem 5.2.1.18 (page 272) and by setting $\xi = 1$ and $X = \log J$, we can express

$$
\mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t)} \log J_i \right\} \right] = \exp \left\{ \lambda \int_t^T M_{\log J_i}(e^{-\kappa s}) - 1 \, ds \right\}
$$

and

$$
\mathbb{E} \left[ \exp \left\{ 2 \sum_{i=1}^{n} e^{-\kappa(T-t)} \log J_i \right\} \right] = \prod_{i=1}^n \left\{ \int_t^T M_{\log J_i}(2e^{-\kappa s}) \lambda e^{-\lambda(T-t)} \, ds \right\}^n = \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^n.
$$

Therefore,

$$
\mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{n} e^{-\kappa(T-t)} \log J_i \right\} \right] = \mathbb{E} \left[ \exp \left\{ 2 \sum_{i=1}^{n} e^{-\kappa(T-t)} \log J_i \right\} \right] - \mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{n} e^{-\kappa(T-t)} \log J_i \right\} \right]^2
$$

$$
= \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^n - \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^{2n}.
$$

Conditional on $S_t$ and $N_{T-t} = n$ and due to the independence of $W_t$ and $N_t$, we can easily show that

$$
\mathbb{E} \left( S_T | S_t, N_{T-t} = n \right) = S_t e^{-\kappa(T-t)} \exp \left\{ \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\sigma^2}{4\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\}
$$

$$
\times \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^n
$$

and

$$
\mathbb{E} \left( S_T | S_t, N_{T-t} = n \right) = S_t e^{-\kappa(T-t)} \exp \left\{ 2 \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] \right\}
$$

$$
\times \left[ \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\} - 1 \right] \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\}
$$

$$
\times \left\{ \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^n - \left[ \lambda \int_t^T e^{2\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds \right]^{2n} \right\}.
$$

Finally, by treating $N_{T-t}$ as a random variable, from Problem 5.2.1.18 (page 272) and by setting $\xi = 1$ and $X = \log J$, we can express

$$
\mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{N_{T-t}} e^{-\kappa(T-t)} \log J_i \right\} \right] = \exp \left\{ \lambda \int_t^T \left[ M_{\log J_i}(e^{-\kappa s}) - 1 \right] \, ds \right\}
$$

$$
= \exp \left\{ \lambda \int_t^T e^{\mu \xi + \frac{\sigma^2}{2} e^{-2\xi s - \lambda(T-t)}} \, ds - \lambda(T-t) \right\}.$$
and

\[
\mathbb{E}\left[ \exp \left\{ 2 \sum_{i=1}^{N_{t-t}} e^{-\kappa(T-t_i)} \log J_i \right\} \right] = \exp \left\{ \lambda \int_t^T [M_{\log J}(2e^{-\kappa s}) - 1] \, ds \right\} = \exp \left\{ \lambda \int_t^T e^{2\mu e^{-\kappa s} + \sigma^2 e^{-2\kappa s}} \, ds - \lambda(T-t) \right\}
\]

and subsequently

\[
\text{Var} \left[ \exp \left\{ \sum_{i=1}^{N_{t-t}} e^{-\kappa(T-t_i)} \log J_i \right\} \right] = \mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{N_{t-t}} e^{-\kappa(T-t_i)} \log J_i \right\} \right] - \mathbb{E} \left[ \exp \left\{ \sum_{i=1}^{N_{t-t}} e^{-\kappa(T-t_i)} \log J_i \right\} \right]^2 = \exp \left\{ \lambda \int_t^T e^{2\mu e^{-\kappa s} + \sigma^2 e^{-2\kappa s}} \, ds - \lambda(T-t) \right\}
- \exp \left\{ 2\lambda \int_t^T e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} \, ds - 2\lambda(T-t) \right\}
= e^{-\lambda(T-t)} \left[ \exp \left\{ \lambda \int_t^T e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} \left( e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} - 2 \right) \, ds \right\} - e^{-\lambda(T-t)} \right].
\]

Conditional only on \(S_t\) and from the independence of \(W_i\) and \(N_i\), we have

\[
\mathbb{E}(S_T|S_t) = S_t e^{-\kappa(T-t)} \exp \left\{ \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] + \frac{\sigma^2}{4\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] 
+ \lambda \int_t^T e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} \, ds - \lambda(T-t) \right\}
\]

and

\[
\text{Var}(S_T|S_t) = S_t^2 e^{-\kappa(T-t)} \exp \left\{ 2 \left( \theta - \frac{\sigma^2}{2\kappa} \right) \left[ 1 - e^{-\kappa(T-t)} \right] \right\}
\times \left[ \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\} - 1 \right] \exp \left\{ \frac{\sigma^2}{2\kappa} \left[ 1 - e^{-2\kappa(T-t)} \right] \right\} e^{-\lambda(T-t)}
\times \left[ \exp \left\{ \lambda \int_t^T e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} \left( e^{\mu e^{-\kappa s} + \frac{1}{2} \sigma^2 e^{-2\kappa s}} - 2 \right) \, ds \right\} - e^{-\lambda(T-t)} \right].
\]

\[\square\]

6. **Kou’s Model.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and let \(\{N_t : t \geq 0\}\) be a Poisson process with intensity \(\lambda > 0\) and \(\{W_t : t \geq 0\}\) be a standard Wiener process relative to the same
5.2.2 Jump Diffusion Process

filtration \( \mathcal{F}_t, t \geq 0 \). Suppose \( S_t \) follows a jump diffusion process with the following SDE

\[
\frac{dS_t}{S_t} = (\mu - D) + \sigma dW_t + (J_t - 1) dN_t,
\]

where

\[
dN_t = \begin{cases} 
1 & \text{with probability } \lambda dt \\
0 & \text{with probability } 1 - \lambda dt
\end{cases}
\]

with constants \( \mu, D \) and \( \sigma > 0 \) being the drift, continuous dividend yield and volatility, respectively. The jump variable is denoted by \( J_t \), where \( X_t = \log J_t \) follows an asymmetric double exponential distribution with density function

\[
f_{X_t}(x) = \begin{cases} 
pe^{-\alpha x} & x \geq 0 \\
(1 - p)\beta e^{\beta x} & x > 0
\end{cases}
\]

where \( 0 \leq p \leq 1, \alpha > 1 \) and \( \beta > 0 \). Assume that \( J_t, W_t \) and \( N_t \) are mutually independent and let \( J_i, i = 1, 2, \ldots \) be a sequence of independent and identically distributed random variables which are independent of \( N_t \) and \( W_t \).

By applying Itô’s formula on \( \log S_t \) and taking integrals show that for \( t < T \),

\[
S_T = S_t e^{(\mu - D - \frac{1}{2} \sigma^2)(T-t) + \sigma W_{T-t} - \lambda \sum_{i=1}^{N_{T-t}} J_i}
\]

provided the random variable \( J_t \in (0, 2] \), \( i = 1, 2, \ldots, N_{T-t} \) and \( W_{T-t} \sim \mathcal{N}(0, T-t) \).

Given \( S_t \) and \( N_{T-t} = n \), show that the mean and variance of \( S_T \) are

\[
\mathbb{E}(S_T | S_t, N_{T-t} = n) = S_t e^{(\mu - D)(T-t)} \left[ \frac{p}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1} \right]^n
\]

and

\[
\text{Var}(S_T | S_t, N_{T-t} = n) = S_t^2 e^{2(\mu - D)(T-t)} \left( e^{\sigma^2(T-t)} - 1 \right) \times \left\{ \frac{p}{\alpha - 2} + (1-p) \frac{\beta}{\beta + 2} - \left[ \frac{p}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1} \right]^2 \right\}^n.
\]

Finally, conditional only on \( S_t \), show that

\[
\mathbb{E}(S_T | S_t) = S_t e^{(\mu - D)(T-t)} \left[ \frac{p}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1} \right] \exp \left\{ \lambda \left( e^{\mu_t + \frac{1}{2} \sigma^2} - 1 \right) (T-t) \right\}
\]

and

\[
\text{Var}(S_T | S_t) = S_t^2 e^{2(\mu - D)(T-t)} \left( e^{\sigma^2(T-t)} - 1 \right) \left[ \frac{p}{\alpha - 2} + (1-p) \frac{\beta}{\beta + 2} \right] \times \left[ \exp \left\{ \lambda (T-t) \left( e^{2(\mu_t + \sigma^2)} - 1 \right) \right\} - \exp \left\{ 2\lambda (T-t) \left( e^{\mu_t + \frac{1}{2} \sigma^2} - 1 \right) \right\} \right].
\]
5.2.2 Jump Diffusion Process

**Solution:** Using the same steps as described in Problem 5.2.2.3 (page 285), for \( t < T \) we can show the solution of the jump diffusion process is

\[
S_T = S_t e^{\left( \mu - \frac{1}{2} \sigma^2 \right)(T-t) + \sigma W_{T-t} \prod_{i=1}^{N_{T-t}} J_i}
\]

provided \( J_i \in (0, 2], \ i = 1, 2, \ldots, N_{T-t} \).

To find the mean and variance of \( J_t \) we note that

\[
\mathbb{E}(J_t) = \mathbb{E}(e^{X_t})
\]

\[
= \int_{-\infty}^{\infty} e^{x} f_{X_t}(x) \, dx
\]

\[
= \int_{-\infty}^{0} (1-p)\beta e^{(1+\beta)x} \, dx + \int_{0}^{\infty} pae^{(1-\alpha)x} \, dx
\]

\[
= p \frac{\alpha}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1}
\]

and

\[
\mathbb{E}(J_t^2) = \mathbb{E}(e^{2X_t})
\]

\[
= \int_{-\infty}^{\infty} e^{2x} f_{X_t}(x) \, dx
\]

\[
= \int_{-\infty}^{0} (1-p)\beta e^{(2+\beta)x} \, dx + \int_{0}^{\infty} pae^{(2-\alpha)x} \, dx
\]

\[
= p \frac{\alpha}{\alpha - 2} + (1-p) \frac{\beta}{\beta + 2}
\]

so that

\[
\text{Var}(J_t) = \mathbb{E}(J_t^2) - \mathbb{E}(J_t)^2
\]

\[
= p \frac{\alpha}{\alpha - 2} + (1-p) \frac{\beta}{\beta + 2} - \left[ p \frac{\alpha}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1} \right]^2.
\]

Since \( W_{T-t} \sim \mathcal{N}(0, T-t) \) so that

\[
\mathbb{E}(e^{\sigma W_{T-t}}) = e^{\frac{1}{2}\sigma^2(T-t)} \quad \text{and} \quad \text{Var}(e^{\sigma W_{T-t}}) = \left( e^{\sigma^2(T-t)} - 1 \right) e^{\sigma^2(T-t)}
\]

and conditional on \( S_t, N_{T-t} = n \) and from independence, we have

\[
\mathbb{E}(S_T | S_t, N_{T-t} = n) = S_t e^{(\mu - D)(T-t) - \frac{1}{2} \sigma^2(T-t)} \mathbb{E}(e^{\sigma W_{T-t}}) \prod_{i=1}^{n} \mathbb{E}(J_i)
\]

\[
= S_t e^{(\mu - D)(T-t)} \left[ p \frac{\alpha}{\alpha - 1} + (1-p) \frac{\beta}{\beta + 1} \right]^n.
\]
5.2.3 Girsanov’s Theorem for Jump Processes

298

5.2.3 Girsanov’s Theorem for Jump Processes

1. Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T\). Let \( \eta > 0 \) and consider the Radon–Nikodým derivative process

\[
Z_t = e^{(\lambda - \eta)N_t},
\]

Show that

\[
\frac{dZ_t}{Z_{t^-}} = (\lambda - \eta) \, dt - \left(\frac{\lambda - \eta}{\lambda}\right) \, dN_t,
\]

for \( 0 \leq t < T \).

**Solution:** From Taylor’s theorem,

\[
dZ_t = \frac{\partial Z_t}{\partial t} \, dt + \frac{\partial Z_t}{\partial N_t} \, dN_t + \frac{\partial^2 Z_t}{\partial t \partial N_t} (dN_t, dt) + \frac{1}{2!} \frac{\partial^2 Z_t}{\partial N_t^2} (dN_t)^2 + \frac{1}{3!} \frac{\partial^3 Z_t}{\partial N_t^3} (dN_t)^3 + \ldots
\]
5.2.3 Girsanov’s Theorem for Jump Processes

Since \(dN_t dt = 0\), \((dN_t)^2 = (dN_t)^3 = \ldots = dN_t\), we have

\[
dZ_t = \frac{\partial Z_t}{\partial t} dt + \left( \frac{\partial Z_t}{\partial N_t} + \frac{1}{2!} \frac{\partial^2 Z_t}{\partial N_t^2} + \frac{1}{3!} \frac{\partial^3 Z_t}{\partial N_t^3} + \ldots \right) dN_t.
\]

From \(Z_t = e^{(\lambda - \eta)u} \left( \frac{\eta}{\lambda} \right)^{N_t}\) we can express

\[
\frac{\partial Z_t}{\partial t} = (\lambda - \eta) e^{(\lambda - \eta)u} \left( \frac{\eta}{\lambda} \right)^{N_t} = (\lambda - \eta)Z_t
\]

\[
\frac{\partial Z_t}{\partial N_t} = e^{(\lambda - \eta)u} \log \left( \frac{\eta}{\lambda} \right) \left( \frac{\eta}{\lambda} \right)^{N_t} = \log \left( \frac{\eta}{\lambda} \right) Z_t
\]

\[
\frac{\partial^2 Z_t}{\partial N_t^2} = \log \left( \frac{\eta}{\lambda} \right) \frac{\partial Z_t}{\partial N_t} = \left[ \log \left( \frac{\eta}{\lambda} \right) \right]^2 Z_t.
\]

In general, we can deduce that

\[
\frac{\partial^m Z_t}{\partial N_t^m} = \left[ \log \left( \frac{\eta}{\lambda} \right) \right]^m Z_t, \quad m = 1, 2, \ldots
\]

Therefore,

\[
dZ_t = (\lambda - \eta)Z_t dt + \left\{ \log \left( \frac{\eta}{\lambda} \right) + \frac{1}{2!} \left[ \log \left( \frac{\eta}{\lambda} \right) \right]^2 + \frac{1}{3!} \left[ \log \left( \frac{\eta}{\lambda} \right) \right]^3 + \ldots \right\} Z_t dN_t
\]

\[
= (\lambda - \eta)Z_t dt + \left\{ e^{\log \left( \frac{\eta}{\lambda} \right)} - 1 \right\} Z_t dN_t
\]

\[
= (\lambda - \eta)Z_t dt + \left( \frac{\eta - \lambda}{\lambda} \right) Z_t dN_t.
\]

By letting \(Z_{t^-}\) denote the value of \(Z_t\) before a jump event, we have

\[
\frac{dZ_t}{Z_{t^-}} = (\lambda - \eta) dt - \left( \frac{\lambda - \eta}{\lambda} \right) dN_t.
\]

\[\square\]

2. Let \(\{N_t : 0 \leq t \leq T\}\) be a Poisson process with intensity \(\lambda > 0\) defined on the probability space \((\Omega, \mathcal{F}, P)\) with respect to the filtration \(\mathcal{F}_t, 0 \leq t \leq T\). Let \(\eta > 0\) and consider the Radon–Nikodým derivative process

\[Z_t = e^{(\lambda - \eta)u} \left( \frac{\eta}{\lambda} \right)^{N_t}.
\]

Show that \(E^P (Z_t) = 1\) and \(Z_t\) is a positive \(P\)-martingale for \(0 \leq t \leq T\).
5.2.3 Girsanov’s Theorem for Jump Processes

**Solution:** From Problem 5.2.3.1 (page 298), \( Z_t \) satisfies

\[
\frac{dZ_t}{Z_t} = (\lambda - \eta) \, dt - \left( \frac{\eta - \lambda}{\lambda} \right) \, dN_t
\]

\[
= \left( \frac{\eta - \lambda}{\lambda} \right) (dN_t - \lambda \, dt)
\]

\[
= \left( \frac{\eta - \lambda}{\lambda} \right) d(N_t - \lambda t)
\]

\[
= \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_t
\]

where \( \hat{N}_t = N_t - \lambda t \). Since \( \hat{N}_t = N_t - \lambda t \) is a \( \mathbb{P} \)-martingale, therefore \( \left( \frac{\eta - \lambda}{\lambda} \right) \hat{N}_t \) is also a \( \mathbb{P} \)-martingale.

Taking integrals,

\[
\int_0^t dZ_u = \int_0^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u
\]

\[
Z_t = Z_0 + \int_0^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u
\]

and taking expectations,

\[
\mathbb{E}^\mathbb{P} (Z_t) = \mathbb{E}^\mathbb{P} (Z_0) + \mathbb{E}^\mathbb{P} \left[ \int_0^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u \right] = 1
\]

since \( Z_0 = 1 \) and due to \( \hat{N}_t \) being a \( \mathbb{P} \)-martingale, \( \mathbb{E}^\mathbb{P} \left[ \int_0^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u \right] = 0 \). Thus, \( \mathbb{E}^\mathbb{P} (Z_t) = 1 \) for all \( 0 \leq t \leq T \).

To show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale, by taking integrals for \( s < t \),

\[
\int_s^t dZ_u = \int_s^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u
\]

\[
Z_t - Z_s = \int_s^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u.
\]

Taking expectations under the filtration \( \mathcal{F}_s \), \( s < t \),

\[
\mathbb{E}^\mathbb{P} (Z_t - Z_s | \mathcal{F}_s) = \mathbb{E}^\mathbb{P} \left[ \int_s^t Z_u \left( \frac{\eta - \lambda}{\lambda} \right) d\hat{N}_u \bigg | \mathcal{F}_s \right]
\]

and because \( \hat{N}_t \) is a \( \mathbb{P} \)-martingale, therefore

\[
\mathbb{E}^\mathbb{P} (Z_t - Z_s | \mathcal{F}_s) = 0 \quad \text{or} \quad \mathbb{E}^\mathbb{P} (Z_t | \mathcal{F}_s) = Z_s.
\]
In addition, since $Z_t > 0$ for all $0 \leq t \leq T$, therefore $|Z_t| = Z_t$ and hence $\mathbb{E}^P (|Z_t|) = \mathbb{E}^P (Z_t) = 1 < \infty$ for all $0 \leq t \leq T$. Finally, because $Z_t$ is $\mathcal{F}_t$ adapted, we can conclude that $Z_t$ is a positive $\mathbb{P}$-martingale for all $0 \leq t \leq T$.

5.2.3 Girsanov’s Theorem for Jump Processes

3. Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Let $\eta > 0$ and consider the Radon–Nikodým derivative process

$$Z_t = e^{(\lambda - \eta)t} \left( \frac{\eta}{\lambda} \right)^{N_t}.$$

By changing the measure $\mathbb{P}$ to a measure $\mathbb{Q}$ such that

$$\mathbb{E}^P \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t = Z_t$$

show that under the $\mathbb{Q}$ measure, $N_t \sim \text{Poisson}(\eta)$ with intensity $\eta > 0$ for $0 \leq t \leq T$.

**Solution:** To solve this problem we need to show

$$\mathbb{E}^Q \left( e^{uN_t} \right) = e^{\lambda t(e^u-1)}$$

which is the moment generating function of a Poisson process, $N_t \sim \text{Poisson}(\lambda t)$, $0 \leq t \leq T$.

For $u \in \mathbb{R}$, using the moment generating function approach,

$$\mathbb{E}^Q \left( e^{uN_t} \right) = \mathbb{E}^P \left( e^{uN_t} \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right)$$

$$= \mathbb{E}^P \left[ e^{uN_t} e^{(\lambda - \eta)t} \left( \frac{\eta}{\lambda} \right)^{N_t} \right]$$

$$= e^{(\lambda - \eta)t} \mathbb{E}^P \left[ e^{(u + \log \left( \frac{\eta}{\lambda} \right))N_t} \right]$$

$$= e^{(\lambda - \eta)t} \exp \left\{ \lambda t \left( e^{u + \log \left( \frac{\eta}{\lambda} \right)} - 1 \right) \right\}$$

since the moment generating function of $N_t \sim \text{Poisson}(\lambda t)$ is

$$\mathbb{E}^P \left( e^{mN_t} \right) = e^{\lambda t(e^m-1)}$$

for all $m \in \mathbb{R}$ and hence

$$\mathbb{E}^Q \left( e^{uN_t} \right) = e^{\eta t(e^u-1)}$$

which is the moment generating function of a Poisson process with intensity $\eta$. Therefore, under the $\mathbb{Q}$ measure, $N_t \sim \text{Poisson}(\eta)$ for $0 \leq t \leq T$. 

\[\square\]
4. Let \( \{ N_t : 0 \leq t \leq T \} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with respect to the filtration \( \mathcal{F}_t, t \geq 0 \). Suppose \( \theta_t \) is an adapted process, \( 0 \leq t \leq T \) and \( \eta > 0 \). We consider the Radon–Nikodým derivative process

\[
Z_t = Z_t^{(1)} \cdot Z_t^{(2)},
\]

such that

\[
Z_t^{(1)} = e^{(\lambda - \eta \theta_t)/\lambda} \quad \text{and} \quad Z_t^{(2)} = e^{-\int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du}.
\]

Show that \( \mathbb{E}^\mathbb{P}(Z_T) = 1 \) and \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( 0 \leq t \leq T \).

**Solution:** From Problem 4.2.2.1 (page 194) and Problem 5.2.3.2 (page 299), we have shown that \( \mathbb{E}^\mathbb{P}(Z_t^{(1)}) = 1 \), \( \mathbb{E}^\mathbb{P}(Z_t^{(2)}) = 1 \) and both \( Z_t^{(1)} \) and \( Z_t^{(2)} \) are \( \mathbb{P} \)-martingales for \( 0 \leq t \leq T \).

From Itô’s lemma,

\[
d(Z_t^{(1)} Z_t^{(2)}) = Z_t^{(1)} dZ_t^{(2)} + Z_t^{(2)} dZ_t^{(1)} + dZ_t^{(1)} dZ_t^{(2)}
\]

where \( Z_t^{(1)} \) is the value of \( Z_t^{(1)} \) before a jump event. Since \( dZ_t^{(1)} dZ_t^{(2)} = 0 \) thus, by taking integrals,

\[
\int_0^t d(Z_u^{(1)} Z_u^{(2)}) = \int_0^t Z_u^{(1)} dZ_u^{(2)} + \int_0^t Z_u^{(2)} dZ_u^{(1)}
\]

\[
Z_t^{(1)} Z_t^{(2)} = Z_0^{(1)} Z_0^{(2)} + \int_0^t Z_u^{(1)} dZ_u^{(2)} + \int_0^t Z_u^{(2)} dZ_u^{(1)}.
\]

Taking expectations under the \( \mathbb{P} \) measure,

\[
\mathbb{E}^\mathbb{P}(Z_t^{(1)} Z_t^{(2)}) = \mathbb{E}^\mathbb{P}(Z_0^{(1)} Z_0^{(2)}) = 1
\]

since \( Z_0^{(1)} = 1 \), \( Z_0^{(2)} = 1 \) and both \( Z_t^{(1)} \) and \( Z_t^{(2)} \) are \( \mathbb{P} \)-martingales.

To show that \( Z_t = Z_t^{(1)} \cdot Z_t^{(2)} \) is a positive \( \mathbb{P} \)-martingale, by taking integrals for \( s < t \) of

\[
d(Z_s^{(1)} Z_s^{(2)}) = Z_s^{(1)} dZ_s^{(2)} + Z_s^{(2)} dZ_s^{(1)}
\]

we have

\[
Z_t^{(1)} Z_t^{(2)} = Z_s^{(1)} Z_s^{(2)} + \int_s^t Z_u^{(1)} dZ_u^{(2)} + \int_s^t Z_u^{(2)} dZ_u^{(1)}
\]

and taking expectations with respect to the filtration \( \mathcal{F}_s, s < t \)

\[
\mathbb{E}^\mathbb{P}(Z_s^{(1)} Z_s^{(2)} | \mathcal{F}_s) = Z_s^{(1)} Z_s^{(2)}
\]

since \( \mathbb{E}^\mathbb{P}(\int_s^t Z_u^{(1)} dZ_u^{(2)} | \mathcal{F}_s) = \mathbb{E}^\mathbb{P}(\int_s^t Z_u^{(2)} dZ_u^{(1)} | \mathcal{F}_s) = 0. \)
5.2.3 Girsanov’s Theorem for Jump Processes

In addition, since $Z_t^{(1)} > 0$ and $Z_t^{(2)} > 0$ for all $0 \leq t \leq T$, therefore $|Z_t^{(1)}Z_t^{(2)}| = Z_t^{(1)}Z_t^{(2)}$ and hence $\mathbb{E}^P\left(|Z_t|\right) = \mathbb{E}^P(Z_t) = 1 < \infty$ for $0 \leq t \leq T$. Finally, because $Z_t$ is also $\mathcal{F}_t$-adapted we can conclude that $Z_t$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$. □

5. Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Suppose $\theta_t$ is an adapted process, $0 \leq t \leq T$ and $\eta > 0$. We consider the Radon–Nikodým derivative process

$$Z_t = Z_t^{(1)} \cdot Z_t^{(2)}$$

such that

$$Z_t^{(1)} = e^{(\lambda - \eta)N_t \lambda}$$

and

$$Z_t^{(2)} = e^{-\int_0^t \theta_udW_u - \frac{1}{2} \int_0^t \theta_u^2 du}.$$

By changing the measure $\mathbb{P}$ to measure $\mathbb{Q}$ such that

$$\mathbb{E}^P\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\Big| \mathcal{F}_t\right) = \frac{d\mathbb{Q}}{d\mathbb{P}} = Z_t$$

show that, under the $\mathbb{Q}$ measure, $N_t$ is a Poisson process with intensity $\eta > 0$, the process $\tilde{W}_t = W_t + \int_0^t \theta_u du$ is a $\mathbb{Q}$-standard Wiener process and $N_t \perp \tilde{W}_t$ for $0 \leq t \leq T$.

**Solution:** The first two results are given in Problem 5.2.3.3 (page 301) and Problem 4.2.2.10 (page 205).

To prove the final result we need to show that, under the $\mathbb{Q}$ measure,

$$\mathbb{E}^Q\left(e^{\alpha N_t + \alpha \tilde{W}_t}\right) = e^{\eta(e^\alpha - 1)} \cdot e^{\frac{1}{2} \alpha^2 t}$$

which is a joint product of independent moment generating functions of $N_t \sim \text{Poisson}(\eta t)$ and $\tilde{W}_t \sim \mathcal{N}(0, t)$.

From the definition

$$\mathbb{E}^Q\left(e^{\alpha N_t + \alpha \tilde{W}_t}\right) = \mathbb{E}^P\left(e^{\alpha N_t + \alpha \tilde{W}_t}Z_t\right) = \mathbb{E}^P\left(e^{\alpha N_t}Z_t^{(1)} \cdot e^{\alpha \tilde{W}_t}Z_t^{(2)}\right)$$

we let

$$\bar{Z}_t = \bar{Z}_t^{(1)} \cdot \bar{Z}_t^{(2)}$$

where

$$\bar{Z}_t^{(1)} = e^{\alpha N_t}Z_t^{(1)} \quad \text{and} \quad \bar{Z}_t^{(2)} = e^{\alpha \tilde{W}_t}Z_t^{(2)}.$$

Using Itô’s lemma, we have

$$d\bar{Z}_t = d\left(\bar{Z}_t^{(1)} \bar{Z}_t^{(2)}\right) = \bar{Z}_t^{(1)} d\bar{Z}_t^{(2)} + \bar{Z}_t^{(2)} d\bar{Z}_t^{(1)} + d\bar{Z}_t^{(1)} d\bar{Z}_t^{(2)}$$

where $\bar{Z}_t^{(1)}$ is the value of $\bar{Z}_t^{(1)}$ before a jump event.
For $d\bar{Z}^{(1)}_t$ we can expand using Taylor’s theorem,

\[
d\bar{Z}^{(1)}_t = \frac{\partial \bar{Z}^{(1)}_t}{\partial N_t} dN_t + \frac{\partial \bar{Z}^{(1)}_t}{\partial Z^{(1)}_t} dZ^{(1)}_t + \frac{1}{2!} \left[ \frac{\partial^2 \bar{Z}^{(1)}_t}{\partial N_t^2} (dN_t)^2 + 2 \frac{\partial^2 \bar{Z}^{(1)}_t}{\partial N_t \partial Z^{(1)}_t} (dN_t)(dZ^{(1)}_t) + \frac{\partial^2 \bar{Z}^{(1)}_t}{\partial Z^{(1)}_t^2} (dZ^{(1)}_t)^2 \right] + \frac{1}{3!} \left[ \frac{\partial^3 \bar{Z}^{(1)}_t}{\partial N_t^3} (dN_t)^3 + 3 \frac{\partial^3 \bar{Z}^{(1)}_t}{\partial N_t^2 \partial Z^{(1)}_t} (dN_t)^2 (dZ^{(1)}_t) + \frac{\partial^3 \bar{Z}^{(1)}_t}{\partial N_t \partial Z^{(1)}_t^2} (dN_t)(dZ^{(1)}_t)^2 + \frac{\partial^3 \bar{Z}^{(1)}_t}{\partial Z^{(1)}_t^3} (dZ^{(1)}_t)^3 \right] + \ldots
\]

Because $dZ^{(1)}_t = \left( \frac{\eta - \lambda}{\lambda} \right) \bar{Z}^{(1)}_t d\hat{N}_t$, $Z^{(1)}_t = e^{\eta N_t} Z^{(1)}_t$ and $d\hat{N}_t = dN_t - \lambda \, dt$, we can write

\[
d\bar{Z}^{(1)}_t = u_t \bar{Z}^{(1)}_t dN_t + e^{\eta N_t} dZ^{(1)}_t + \frac{1}{2!} \left[ u_t^2 \bar{Z}^{(1)}_t dN_t + 2 \left( \frac{\eta - \lambda}{\lambda} \right) u_t \bar{Z}^{(1)}_t dN_t d\hat{N}_t \right] + \frac{1}{3!} \left[ u_t^3 \bar{Z}^{(1)}_t dN_t + 3 \left( \frac{\eta - \lambda}{\lambda} \right) u_t^2 \bar{Z}^{(1)}_t dN_t d\hat{N}_t \right] + \ldots
\]

\[
= \left[ u_t + \frac{1}{2!} u_t^2 + \frac{1}{3!} u_t^3 + \ldots \right] \bar{Z}^{(1)}_t dN_t + \left( \frac{\eta - \lambda}{\lambda} \right) e^{\eta N_t} Z^{(1)}_t d\hat{N}_t + \left( \frac{\eta - \lambda}{\lambda} \right) u_t \bar{Z}^{(1)}_t dN_t d\hat{N}_t + \left( \frac{\eta - \lambda}{\lambda} \right) u_t^2 \bar{Z}^{(1)}_t dN_t d\hat{N}_t + \ldots
\]

\[
= (e^{\eta t} - 1) \bar{Z}^{(1)}_t dN_t + \left( \frac{\eta - \lambda}{\lambda} \right) \bar{Z}^{(1)}_t d\hat{N}_t + \left( \frac{\eta - \lambda}{\lambda} \right) (e^{\eta t} - 1) \bar{Z}^{(1)}_t dN_t d\hat{N}_t
\]

since $(dN_t)^2 = dN_t$ and $dN_t \, dt = 0$.

For the case of $d\bar{Z}^{(2)}_t$, by applying Taylor’s theorem

\[
d\bar{Z}^{(2)}_t = \frac{\partial \bar{Z}^{(2)}_t}{\partial \tilde{W}_t} d\tilde{W}_t + \frac{\partial \bar{Z}^{(2)}_t}{\partial Z^{(2)}_t} dZ^{(2)}_t + \frac{1}{2!} \left[ \frac{\partial^2 \bar{Z}^{(2)}_t}{\partial (\tilde{W}_t)^2} (d\tilde{W}_t)^2 + 2 \frac{\partial^2 \bar{Z}^{(2)}_t}{\partial \tilde{W}_t \partial Z^{(2)}_t} (d\tilde{W}_t)(dZ^{(2)}_t) + \frac{\partial^2 \bar{Z}^{(2)}_t}{\partial Z^{(2)}_t^2} (dZ^{(2)}_t)^2 \right] + \ldots
\]
and since $d\tilde{W}_t = dW_t + \theta_t dt$ and $dZ_t^{(2)} = -\theta_t Z_t^{(2)} dW_t$, (see Problem 4.2.2.2, page 196), we can write

$$d\tilde{Z}_t^{(2)} = \frac{1}{2} u_t^2 \tilde{Z}_t^{(2)} dt + (u_t - \theta_t) \tilde{Z}_t^{(2)} dW_t.$$ 

Since $d\hat{N}_t = dN_t - \lambda dt$, $dN_t dW_t = 0$, $dW_t dt = 0$, $dN_t dt = 0$, $(dW_t)^2 = dt$ and $(dt)^2 = 0$, we have

$$d\tilde{Z}_t^{(1)} d\tilde{Z}_t^{(2)} = \frac{1}{2} u_t^2 \tilde{Z}_t^{(1)} dt + (u_t - \theta_t) \tilde{Z}_t^{(1)} dW_t.$$ 

Thus, by letting $d\hat{N}_t = d\hat{N}_t + \lambda dt$, we have

$$d\tilde{Z}_t = \frac{1}{2} u_t^2 \tilde{Z}_t dt + (u_t - \theta_t) \tilde{Z}_t dW_t + \left( \frac{\eta - \lambda}{\lambda} \right) \tilde{Z}_t d\hat{N}_t.$$ 

Taking integrals, we have

$$\tilde{Z}_t = 1 + \int_0^t \left[ \eta (e^{\eta s} - 1) + \frac{1}{2} u_s^2 \right] \tilde{Z}_s ds + \int_0^t (u_s - \theta_s) \tilde{Z}_s dW_s + \int_0^t \left( \frac{\eta e^{\eta s} - \lambda}{\lambda} \right) \tilde{Z}_s d\hat{N}_s,$$

where $\tilde{Z}_0 = 1$, and because $\hat{N}_t$ and $W_t$ are $\mathbb{P}$-martingales, thus by taking expectations under the $\mathbb{P}$ measure we have

$$\mathbb{E}^\mathbb{P} \left( \tilde{Z}_t \right) = 1 + \int_0^t \left[ \eta (e^{\eta s} - 1) + \frac{1}{2} u_s^2 \right] \mathbb{E}^\mathbb{P} \left( \tilde{Z}_s \right) ds.$$

By differentiating the above equation with respect to $t$,

$$\frac{d}{dt} \mathbb{E}^\mathbb{P} \left( \tilde{Z}_t \right) = \left[ \eta (e^{\eta t} - 1) + \frac{1}{2} u_t^2 \right] \mathbb{E}^\mathbb{P} \left( \tilde{Z}_t \right),$$

or

$$\frac{d}{dt} m_t - \left[ \eta (e^{\eta t} - 1) + \frac{1}{2} u_t^2 \right] m_t = 0,$$

where $m_t = \mathbb{E}^\mathbb{P} \left( \tilde{Z}_t \right)$. 

Thus, by letting $dN_t = d\hat{N}_t + \lambda dt$, we have

$$d\tilde{Z}_t = \frac{1}{2} u_t^2 \tilde{Z}_t dt + (u_t - \theta_t) \tilde{Z}_t dW_t + \left( \frac{\eta - \lambda}{\lambda} \right) \tilde{Z}_t d\hat{N}_t.$$ 

Taking integrals, we have

$$\tilde{Z}_t = 1 + \int_0^t \left[ \eta (e^{\eta s} - 1) + \frac{1}{2} u_s^2 \right] \tilde{Z}_s ds + \int_0^t (u_s - \theta_s) \tilde{Z}_s dW_s + \int_0^t \left( \frac{\eta e^{\eta s} - \lambda}{\lambda} \right) \tilde{Z}_s d\hat{N}_s,$$

where $\tilde{Z}_0 = 1$, and because $\hat{N}_t$ and $W_t$ are $\mathbb{P}$-martingales, thus by taking expectations under the $\mathbb{P}$ measure we have

$$\mathbb{E}^\mathbb{P} \left( \tilde{Z}_t \right) = 1 + \int_0^t \left[ \eta (e^{\eta s} - 1) + \frac{1}{2} u_s^2 \right] \mathbb{E}^\mathbb{P} \left( \tilde{Z}_s \right) ds.$$
By setting the integrating factor to be 
\[ I = e^{-\int (\eta (e^{u_1} - 1) + \frac{1}{2} u_2^2) \, dt} = e^{-\eta (e^{u_1} - 1) + \frac{1}{2} u_2^2} \] 
and multiplying it by the first-order ordinary differential equation, we have
\[ \frac{d}{dt} \left( m_t e^{-\eta_t (e^{u_1} - 1) + \frac{1}{2} u_2^2} \right) = 0 \quad \text{or} \quad e^{-\eta_t (e^{u_1} - 1) + \frac{1}{2} u_2^2} E^P (Z_t) = C \]
where \( C \) is a constant. Since \( E^P (Z_0) = E^P \left( e^{\eta_t (e^{u_1} - 1) - \frac{1}{2} u_2^2} \right) = 1 \), therefore \( C = 1 \).
Thus, we will have
\[ E^Q \left( e^{\mu_t N_t + u_2 \tilde{W}_t} \right) = E^P (Z_t) = e^{\eta_t (e^{u_1} - 1) + \frac{1}{2} u_2^2}. \]
Since the joint moment generating function of \( E^Q \left( e^{\mu_t N_t + u_2 \tilde{W}_t} \right) = e^{\eta (e^{u_1} - 1) - \frac{1}{2} u_2^2} \) can be expressed as a product of the moment generating functions for \( N_t \sim \text{Poisson}(\lambda t) \) and \( \tilde{W}_t \sim \mathcal{N}(0, t) \), respectively, we can deduce that \( N_t \) and \( \tilde{W}_t \) are independent. \( \square \)

6. Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, P) \) with respect to the filtration \( \mathcal{F}_t, t \geq 0 \). Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables with common probability mass function \( P(X = x_k) = p(x_k) > 0, k = 1, 2, \ldots, K \) and \( \sum_{k=1}^{K} P(X = x_k) = 1 \). Let \( X_1, X_2, \ldots \) also be independent of \( N_t \). From the definition of a compound Poisson process
\[ M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0 \]
show that \( M_t \) and \( N_t \) can be expressed as
\[ M_t = \sum_{k=1}^{K} x_k N_t^{(k)}, \quad N_t = \sum_{k=1}^{K} N_t^{(k)} \]
where \( N_t^{(k)} \sim \text{Poisson}(\lambda_t p(x_k)) \) such that \( N_t^{(i)} \perp N_t^{(j)}, i \neq j \).
Let \( \eta_1, \eta_2, \ldots, \eta_K > 0 \) and consider the Radon–Nikodým derivative process
\[ Z_t = \prod_{k=1}^{K} Z_t^{(k)} \]
where \( Z_t^{(k)} = e^{(\lambda_k - \eta_k) t} \left( \frac{\eta_k}{\lambda_k} \right)^{N_t^{(k)}} \) such that \( \lambda_k = \lambda p(x_k), k = 1, 2, \ldots, K \).
Show that \( E^P (Z_t) = 1 \) and \( Z_t \) is a positive \( P \)-martingale for \( 0 \leq t \leq T \).

**Solution:** The first part of the result follows from Problem 5.2.1.16 (page 268).
From Problem 5.2.3.1 (page 298), for each \( k = 1, 2, \ldots, K \), \( Z^{(k)}_t \) satisfies

\[
\frac{dZ_t^{(k)}}{Z_t^{(k)}} = \left( \frac{\eta_k - \lambda_k}{\lambda_k} \right) d\hat{N}_t^{(k)}
\]

where \( \hat{N}_t^{(k)} = N_t^{(k)} - \lambda_k t \), \( \lambda_k = \lambda t p(x_k) \), therefore we can easily show that \( \hat{N}_t^{(k)} = N_t^{(k)} - \lambda_k t \) is a \( \mathbb{P} \)-martingale. Thus, \( \left( \frac{\eta_k - \lambda_k}{\lambda_k} \right) \hat{N}_t^{(k)} \) is also a \( \mathbb{P} \)-martingale and subsequently,

\[
\mathbb{E}^{\mathbb{P}}(Z_t^{(k)}) = 1 \quad \text{and} \quad Z_t^{(k)} \quad \text{is a} \quad \mathbb{P} \text{-martingale for} \quad 0 \leq t \leq T.
\]

For the case

\[
Z_t = \prod_{i=1}^{K} Z_t^{(k)}
\]

and because the Poisson processes \( N_t^{(i)} \) and \( N_t^{(j)} \), \( i \neq j \) have no simultaneous jumps, we can deduce that \( Z_t^{(i)} \perp Z_t^{(j)} \), \( i \neq j \). Therefore,

\[
\mathbb{E}^{\mathbb{P}}(Z_t) = \mathbb{E}^{\mathbb{P}}\left( \prod_{i=1}^{K} Z_t^{(k)} \right) = \prod_{i=1}^{K} \mathbb{E}^{\mathbb{P}}(Z_t^{(k)}) = \prod_{i=1}^{K} 1 = 1
\]

for \( 0 \leq t \leq T \).

Finally, to show that \( Z_t \) is a positive \( \mathbb{P} \)-martingale, \( 0 \leq t \leq T \) we prove this result via mathematical induction. Let \( K = 1 \), then obviously \( Z_t \) is a positive \( \mathbb{P} \)-martingale. Assume that \( Z_t \) is a positive \( \mathbb{P} \)-martingale for \( K = m \), \( m \geq 1 \). For \( K = m + 1 \) and because \( \prod_{k=1}^{m} Z_t^{(k)} \) and \( Z_t^{(m+1)} \) have no simultaneous jumps, by Itô's lemma

\[
d \left( Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} Z_t^{(m+1)} \right) = Z_t^{(m+1)} d \left( Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} \right) + \left( Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} \right) dZ_t^{(m+1)}
\]

where \( Z_t^{(k)} \) is the value of \( Z_t^{(k)} \) before a jump event.

Since \( N_t^{(i)} \perp N_t^{(j)} \), \( i \neq j \) therefore \( d \left( Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} \right) dZ_t^{(m+1)} = 0 \), and by taking integrals for \( s < t \),

\[
Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} Z_t^{(m+1)} = Z_s^{(1)} Z_s^{(2)} \cdots Z_s^{(m)} Z_s^{(m+1)} + \int_{s}^{t} Z_u^{(m+1)} d \left( Z_u^{(1)} Z_u^{(2)} \cdots Z_u^{(m)} \right)
\]

\[
\quad + \int_{s}^{t} \left( Z_u^{(1)} Z_u^{(2)} \cdots Z_u^{(m)} \right) dZ_u^{(m+1)}.
\]

Because \( \prod_{k=1}^{m} Z_t^{(k)} \) and \( Z_t^{(m+1)} \) are positive \( \mathbb{P} \)-martingales and the integrands are left-continuous, taking expectations with respect to the filtration \( \mathcal{F}_s \), \( s < t \)

\[
\mathbb{E}^{\mathbb{P}} \left( Z_t^{(1)} Z_t^{(2)} \cdots Z_t^{(m)} Z_t^{(m+1)} \bigg| \mathcal{F}_s \right) = Z_s^{(1)} Z_s^{(2)} \cdots Z_s^{(m)} Z_s^{(m+1)}.
\]
5.2.3 Girsanov’s Theorem for Jump Processes

In addition, since $Z_i^{(i)} > 0$, $i = 1, 2, \ldots, m + 1$ for all $t \geq 0$, therefore

$$\prod_{k=1}^{m+1} Z_t^{(k)} = \prod_{k=1}^{m+1} Z_t^{(k)}$$

and hence for $t \geq 0$,

$$\mathbb{E}^\mathbb{P} \left( \prod_{k=1}^{m+1} Z_t^{(k)} \right) = \mathbb{E}^\mathbb{P} \left( \prod_{k=1}^{m+1} Z_t^{(k)} \right) = 1 < \infty.$$

Finally, because $\prod_{k=1}^{m+1} Z_t^{(k)}$ is $\mathcal{F}_t$-adapted, the process $\prod_{k=1}^{m+1} Z_t^{(k)}$ is a positive $\mathbb{P}$-martingale.

Thus, from mathematical induction, $Z_t = \prod_{k=1}^K Z_t^{(k)}$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$.

\[ \Box \]

7. Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with common probability mass function $\mathbb{P}(X = x_k) = p(x_k) > 0$, $k = 1, 2, \ldots, K$ and $\sum_{k=1}^K p(X = x_k) = 1$. Let $X_1, X_2, \ldots$ also be independent of $N_t$. From the definition of a compound Poisson process

$$M_t = \sum_{i=1}^{N_t} X_i, \quad t \geq 0$$

we can set $M_t$ and $N_t$ as

$$M_t = \sum_{k=1}^K x_k N_t^{(k)}, \quad N_t = \sum_{k=1}^K N_t^{(k)}$$

where $N_t^{(k)} \sim \text{Poisson} \left( \lambda p(x_k) \right)$ is the number of jumps in $M_t$ of size $x_k$ up to and including time $t$ such that $N_t^{(i)} \perp N_t^{(j)}$, $i \neq j$.

Let $\eta_1, \eta_2, \ldots, \eta_K > 0$ and consider the Radon–Nikodým derivative process

$$Z_t = \prod_{k=1}^K Z_t^{(k)}$$

where $Z_t^{(k)} = e^{(\lambda_k - \eta_k) N_t^{(k)}}$ such that $\lambda_k = \lambda p(x_k)$, $k = 1, 2, \ldots, K$. By changing the measure $\mathbb{P}$ to a measure $\mathbb{Q}$ such that

$$\mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = Z_t$$
5.2.3 Girsanov’s Theorem for Jump Processes 309

show that under the \( \mathbb{Q} \) measure, \( M_t \) is a compound Poisson process with intensity \( \eta = \sum_{k=1}^{K} \eta_k > 0 \) and \( X_1, X_2, \ldots \) is a sequence of independent and identically distributed random variables with common probability mass function

\[
\mathbb{Q}(X = x_k) = q(x_k) = \frac{\eta_k}{\eta_1 + \eta_2 + \ldots + \eta_K}
\]

such that \( \sum_{k=1}^{K} \mathbb{Q}(X = x_k) = 1 \).

Solution: From the independence of \( N_t^{(1)}, N_t^{(2)}, \ldots, N_t^{(K)} \) under \( \mathbb{P} \), for \( u \in \mathbb{R} \) and using the moment generating function approach

\[
\mathbb{E}^{\mathbb{Q}}(e^{uM_t}) = \mathbb{E}^{\mathbb{P}} \left( e^{uM_t} \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right)
\]

\[
= \mathbb{E}^{\mathbb{P}} \left[ e^{u \sum_{k=1}^{K} \lambda_k \eta_k / \lambda_k} \prod_{k=1}^{K} e^{(\lambda_k - \eta_k)(\eta_k / \lambda_k) N_t^{(k)}} \right]
\]

\[
= \prod_{k=1}^{K} e^{(\lambda_k - \eta_k) u} \mathbb{E}^{\mathbb{P}} \left[ e^{(ux_k + \log(\eta_k / \lambda_k)) N_t^{(k)}} \right]
\]

From Problem 5.2.1.14 (page 266), for a constant \( m > 0 \), \( \mathbb{E}^{\mathbb{P}}(e^{mN_t^{(k)}}) = e^{\lambda p(x_k) e^{m-1}} \) therefore

\[
\mathbb{E}^{\mathbb{Q}}(e^{uM_t}) = \prod_{k=1}^{K} e^{(\lambda_k - \eta_k) u} \cdot e^{\lambda p(x_k) e^{m-1}}
\]

\[
= \prod_{k=1}^{K} e^{(\lambda_k - \eta_k) u} \cdot e^{\lambda (\log(\eta_k / \lambda_k) e^{m-1})}
\]

\[
= \prod_{k=1}^{K} e^{(\lambda_k - \eta_k) u} (\eta_k / \lambda_k) e^{\lambda u_k - 1}
\]

\[
= e^{\eta u (\sum_{k=1}^{K} q(x_k) e^{u_k} - 1)}
\]

which is the moment generating function for a compound Poisson process with intensity \( \eta > 0 \) and jump size distribution \( \mathbb{Q}(X_i = x_i) = q(x_i) = \frac{\eta_i}{\eta}, i = 1, 2, \ldots \)

\( \square \)

8. Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \). Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \), and define the compound Poisson process as

\[
M_t = \sum_{i=1}^{N_t} X_i, \quad 0 \leq t \leq T.
\]
By changing the measure \( \mathbb{P} \) to a measure \( \mathbb{Q} \) such that

\[
\mathbb{E}^\mathbb{P} \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t
\]

show that for \( \eta > 0 \) the Radon–Nikodým derivative process \( Z_t \) can be written as

\[
Z_t = \begin{cases} 
  e^{(\lambda - \eta)x} \prod_{i=1}^{N_t} \frac{\eta \mathbb{Q}(X_i)}{\lambda \mathbb{P}(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\
  e^{(\lambda - \eta)x} \prod_{i=1}^{N_t} \frac{\eta f^{\mathbb{Q}}(X_i)}{\lambda f^{\mathbb{P}}(X_i)} & \text{if } X_i \text{ is continuous}
\end{cases}
\]

where \( \mathbb{P}(X) (f^{\mathbb{P}}(X)) \) and \( \mathbb{Q}(X) (f^{\mathbb{Q}}(X)) \) are the probability mass (density) functions of \( \mathbb{P} \) and \( \mathbb{Q} \), respectively. Note that for the case of continuous variables \( X_i, i = 1, 2, \ldots \), we assume \( \mathbb{Q} \) is absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \).

**Solution:** We first consider the discrete case where we let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables with common probability mass function \( \mathbb{P}(X = x_k) = p(x_k) > 0, k = 1, 2, \ldots, K \) and \( \sum_{k=1}^{K} \mathbb{P}(X = x_k) = 1 \).

From Problem 5.2.3.7 (page 308), the compound Poisson process \( M_t \) and the Poisson process \( N_t \) can be expressed as

\[
M_t = \sum_{k=1}^{K} x_k N_t^{(k)}, \quad N_t = \sum_{k=1}^{K} N_t^{(k)}
\]

where \( N_t^{(k)} \sim \text{Poisson} (\lambda p(x_k)) \) such that \( N_t^{(i)} \perp \perp N_t^{(j)}, i \neq j \). By setting \( \eta_1, \eta_2, \ldots, \eta_K > 0 \), we consider the Radon–Nikodým derivative process

\[
Z_t = \prod_{k=1}^{K} \frac{\eta_k}{\lambda_k} N_t^{(k)}
\]

where \( Z_t^{(k)} = e^{(\lambda - \eta_k)x} \left( \frac{\eta_k}{\lambda_k} \right)^{N_t^{(k)}} \) and \( \lambda_k = \lambda p(x_k) \) for \( k = 1, 2, \ldots, K \). Thus, under the \( \mathbb{Q} \) measure, \( M_t \) is a compound Poisson process with intensity \( \eta = \sum_{k=1}^{K} \eta_k > 0 \) and \( X_1, X_2, \ldots \) is a sequence of independent and identically distributed random variables with common probability mass function

\[
\mathbb{Q}(X = x_k) = q(x_k) = \frac{\eta_k}{\eta_1 + \eta_2 + \ldots + \eta_K}.
\]

Therefore,

\[
Z_t = \prod_{k=1}^{K} Z_t^{(k)} = \prod_{k=1}^{K} e^{(\lambda_k - \eta_k)x} \left( \frac{\eta_k}{\lambda_k} \right)^{N_t^{(k)}}
\]
5.2.3 Girsanov’s Theorem for Jump Processes

\[ e^{\sum_{k=1}^{K} (x_k - \eta_k) \frac{\eta_k}{\lambda_k}} \prod_{k=1}^{K} \left( \frac{\eta_k}{\lambda_k} \right)^{N_k} \]

\[ = e^{(\lambda - \eta) t} \prod_{k=1}^{K} \left( \frac{\eta_k}{\lambda_k} \right)^{N_k} \]

\[ = e^{(\lambda - \eta) N \left( X_i \right)} \prod_{i=1}^{N} \frac{\eta Q \left( X_i \right)}{\lambda P \left( X_i \right)}. \]

By analogy with the discrete case, we can deduce that when \( X_i, i = 1, 2, \ldots \) are continuous random variables the Radon–Nikodým derivative process can also be written as

\[ Z_t = e^{(\lambda - \eta) t} \prod_{i=1}^{N_t} \frac{\eta f \mathbb{Q} \left( X_i \right)}{\lambda f \mathbb{P} \left( X_i \right)}, \quad 0 \leq t \leq T \]

where \( f \mathbb{P} (X) \) and \( f \mathbb{Q} (X) \) are the probability density functions of \( \mathbb{P} \) and \( \mathbb{Q} \) measures, respectively, and \( \mathbb{Q} \) is absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \).

\[ \square \]

9. Let \( \{ N_t : 0 \leq t \leq T \} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t, 0 \leq t \leq T \) and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \). Under the \( \mathbb{P} \) measure, each \( X_i, i = 1, 2, \ldots \) has a probability mass (density) function \( \mathbb{P} (X_i) (f \mathbb{P} (X_i)) \). From the definition of the compound Poisson process

\[ M_t = \sum_{i=1}^{N_t} X_i, \quad 0 \leq t \leq T \]

we let \( \eta > 0 \) and consider the Radon–Nikodým derivative process

\[ Z_t = \begin{cases} 
 e^{(\lambda - \eta) t} \prod_{i=1}^{N_t} \frac{\eta Q \left( X_i \right)}{\lambda P \left( X_i \right)} & \text{if } X_i \text{ is discrete, } \mathbb{P} (X_i) > 0 \\
 e^{(\lambda - \eta) t} \prod_{i=1}^{N_t} \frac{\eta Q \left( X_i \right)}{\lambda P \left( X_i \right)} & \text{if } X_i \text{ is continuous}
\end{cases} \]

where \( Q (X_i) (f \mathbb{Q} (X_i)) \) is the probability mass (density) function of \( X_i, i = 1, 2, \ldots \) under the \( \mathbb{Q} \) measure. For the case of continuous random variables \( X_i, i = 1, 2, \ldots \) we also let \( \mathbb{Q} \) be absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \). By changing the measure \( \mathbb{P} \) to measure \( \mathbb{Q} \) such that

\[ \mathbb{E}^{\mathbb{P}} \left[ \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right] = \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t} = Z_t \]

show that under the \( \mathbb{Q} \) measure for \( 0 \leq t \leq T, M_t \) is a compound Poisson process with intensity \( \eta > 0 \) and \( X_i, i = 1, 2, \ldots \) is a sequence of independent and identically
distributed random variables with probability mass (density) functions \( \mathbb{Q}(X_i) (f^\mathbb{Q}(X_i)) \), \( i = 1, 2, \ldots \)

**Solution:** To solve this problem we need to show

\[
\mathbb{E}^\mathbb{Q}(e^{\mu M_t}) = e^{\eta(\varphi^\mathbb{Q}(u) - 1)}, \quad u \in \mathbb{R}
\]

which is the moment generating function of a compound Poisson process with intensity \( \eta \) such that

\[
\varphi^\mathbb{Q}_X(u) = \mathbb{E}^\mathbb{Q}(e^{\mu X}) = \begin{cases} 
\sum_{i=1}^{N_t} e^{\mu X_i} \mathbb{Q}(X_i) & \text{if } X \text{ is discrete} \\
\int_{-\infty}^{\infty} e^{\mu x} f^\mathbb{Q}(x) \, dx & \text{if } X \text{ is continuous}.
\end{cases}
\]

Without loss of generality we consider only the continuous case where, by definition,

\[
\mathbb{E}^\mathbb{Q}(e^{\mu M_t}) = \mathbb{E}^\mathbb{P}\left(e^{\mu M_t} \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \mathbb{E}^\mathbb{P}\left(e^{\mu M_t} Z_t \right)
\]

\[
\mathbb{E}^\mathbb{P}\left(e^{\mu \sum_{i=1}^{N_t} X_i} \prod_{i=1}^{N_t} \eta f^\mathbb{Q}(X_i) \bigg| N_t \right) = \mathbb{E}^\mathbb{P}\left(e^{\mu \sum_{i=1}^{N_t} X_i} \prod_{i=1}^{N_t} \frac{\eta f^\mathbb{Q}(X_i)}{f^\mathbb{P}(X_i)} \bigg| N_t \right)
\]

\[
= e^{(\lambda - \eta) u} \mathbb{P}(N_t = n) \sum_{n=0}^{\infty} e^{n (\lambda - \eta) u} \mathbb{E}^\mathbb{P}\left(e^{\mu \sum_{i=1}^{N_t} X_i} \prod_{i=1}^{N_t} \frac{\eta f^\mathbb{Q}(X_i)}{f^\mathbb{P}(X_i)} \bigg| N_t \right)\mathbb{P}(N_t = n)
\]

Since \( X_1, X_2, \ldots \) are independent and identically distributed random variables and because \( N_t \sim \text{Poisson}(\lambda t) \), we can express

\[
\mathbb{E}^\mathbb{Q}(e^{\mu M_t}) = e^{(\lambda - \eta) u} \sum_{n=0}^{\infty} \mathbb{E}^\mathbb{P}\left(\prod_{i=1}^{N_t} \left( \frac{\eta}{\lambda} e^{\mu X_i} \frac{f^\mathbb{Q}(X_i)}{f^\mathbb{P}(X_i)} \right) \bigg| N_t \right) \mathbb{P}(N_t = n)
\]

\[
= e^{(\lambda - \eta) u} \sum_{n=0}^{\infty} \left[ \mathbb{E}^\mathbb{P}\left( \frac{\eta}{\lambda} e^{\mu X} f^\mathbb{Q}(X) \bigg| f^\mathbb{P}(X) \right) \right]^n e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\]

Because

\[
\mathbb{E}^\mathbb{P}\left( \frac{\eta}{\lambda} e^{\mu X} f^\mathbb{Q}(X) \bigg| f^\mathbb{P}(X) \right) = \int_{-\infty}^{\infty} \frac{\eta}{\lambda} e^{\mu x} f^\mathbb{Q}(x) \cdot f^\mathbb{P}(x) \, dx
\]

\[
= \int_{-\infty}^{\infty} \frac{\eta}{\lambda} e^{\mu x} f^\mathbb{Q}(x) \, dx
\]

\[
= \frac{\eta}{\lambda} \varphi^\mathbb{Q}_X(u)
\]
5.2.3 Girsanov’s Theorem for Jump Processes

where \( \varphi^Q_X(u) = \int_{-\infty}^{\infty} e^{ux} f^Q(x) \, dx \), then

\[
\mathbb{E}^Q(e^{\lambda M_t}) = e^{(\lambda - \eta)t} \sum_{n=0}^{\infty} \left( \eta \varphi^Q_X(u) \right)^n \frac{e^{-\lambda t}(\lambda t)^n}{n!}
\]

\[
= e^{-\eta t} \sum_{n=0}^{\infty} \left( \eta \varphi^Q_X(u) \right)^n \frac{e^{-\lambda t}(\lambda t)^n}{n!}
\]

\[
= e^{\eta(\varphi^Q_X(u) - 1)}.
\]

Based on the moment generating function we can deduce that under the \( \mathbb{Q} \) measure, \( M_t \) is a compound Poisson process with intensity \( \eta > 0 \) and \( X_1, X_2, \ldots \) are also independent and identically distributed random variables with probability density function \( f^Q(X_i) \), \( i = 1, 2, \ldots \)

\[\square\]

10. Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \) and let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables which are also independent of \( N_t \). Under the \( \mathbb{P} \) measure each \( X_i, i = 1, 2, \ldots \) has a probability mass (density) function \( \mathbb{P}(X_i) (f^P(X_i)) \). From the definition of the compound Poisson process

\[ M_t = \sum_{i=1}^{N_t} X_i, \quad 0 \leq t \leq T \]

we let \( \eta > 0 \) and consider the Radon–Nikodym derivative process

\[
Z_t = \begin{cases} 
\frac{e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \eta^Q(X_i) \lambda^P(X_i)}{\lambda^P(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\
\frac{e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \eta^Q(X_i) \lambda^P(X_i)}{\lambda^P(X_i)} & \text{if } X_i \text{ is continuous}
\end{cases}
\]

where \( Q(X_i) (f^Q(X_i)) \) is the probability mass (density) function of \( X_i, i = 1, 2, \ldots \) under the \( \mathbb{Q} \) measure. For the case of continuous random variables \( X_i, i = 1, 2, \ldots \) we also let \( \mathbb{Q} \) be absolutely continuous with respect to \( \mathbb{P} \) on \( \mathcal{F} \).

Show that \( \mathbb{E}^P(Z_t) = 1 \) and \( Z_t \) is a positive \( \mathbb{P} \) martingale for \( 0 \leq t \leq T \).

**Solution:** Without loss of generality we consider only the case of continuous random variables \( X_i, i = 1, 2, \ldots \)

Let \( Z_t = e^{(\lambda - \eta)t} G_t \), where \( G_t = \prod_{i=1}^{N_t} \lambda^P(X_i) \) is defined as a pure jump process, and let \( Z_t \) denote the value of \( Z_t \) before a jump event.
From Itô’s lemma,

\[ dZ_t = \frac{\partial Z_t}{\partial t} dt + \frac{\partial Z_t}{\partial G_t} dG_t + \frac{1}{2} \frac{\partial^2 Z_t}{\partial G_t^2} (dG_t)^2 + \ldots \]

\[ = (\lambda - \eta)e^{(\lambda - \eta)G_t} dt + e^{(\lambda - \eta)G_t} dG_t \]

\[ = (\lambda - \eta)Z_t + e^{(\lambda - \eta)G_t} dG_t. \]

Let \( G_t^- \) be the value \( G_t \) just before a jump event and assume there occurs an instantaneous jump at time \( t \). Therefore,

\[ dG_t = G_t - G_t^- \]

\[ = \frac{\eta f^Q(X_t)}{\lambda f^P(X_t)} G_t^- - G_t^- \]

\[ = \left( \frac{\eta f^Q(X_t)}{\lambda f^P(X_t)} - 1 \right) G_t^- \]

where \( \frac{\eta f^Q(X_t)}{\lambda f^P(X_t)} \) is the size of the jump variable if \( N \) jumps at time \( t \).

By defining a new compound Poisson process

\[ H_t = \sum_{i=1}^{N_t} \frac{\eta f^Q(X_i)}{\lambda f^P(X_i)} \]

we let \( H_t^- \) be the value of \( H_t \) just before a jump. At jump time \( t \) (which is also the jump time of \( M_t \) and \( G_t \)), we have

\[ dH_t = H_t - H_t^- = \frac{\eta f^Q(X_t)}{\lambda f^P(X_t)}. \]

Therefore, at jump times of \( N \) we can write

\[ \frac{dG_t}{G_t^-} = dH_t - 1 \]

and in general we can write

\[ \frac{dG_t}{G_t^-} = dH_t - dN_t \]

where

\[ dN_t = \begin{cases} 1 & \text{with probability } \lambda \ dt \\ 0 & \text{with probability } 1 - \lambda \ dt \end{cases} \]

and

\[ dH_t = \frac{\eta f^Q(X_t)}{\lambda f^P(X_t)} dN_t. \]
Thus,
\[dZ_t = (\lambda - \eta)Z_t \, dt + e^{(\lambda - \eta)t}(dH_t - dN_t)G_t\]
\[= (\lambda - \eta)Z_t \, dt + Z_t \, dH_t - Z_t \, dN_t\]
or
\[dZ_t \over Z_t = d(H_t - \eta t) - d(N_t - \lambda t)\]

Since \(\hat{N}_t = N_t - \lambda t\) is a \(\mathbb{P}\)-martingale we now need to show that \(\hat{H}_t = H_t - \eta t\) is also a \(\mathbb{P}\)-martingale. By definition,
\[\mathbb{E}^\mathbb{P}\left[\eta f_\mathbb{P}(X) \over \lambda f_\mathbb{P}(X)\right] = \frac{\eta}{\lambda} \int_{-\infty}^{\infty} f_\mathbb{P}(x) \cdot f_\mathbb{P}(x) \, dx = \frac{\eta}{\lambda}\]

since \(\int_{-\infty}^{\infty} f_\mathbb{P}(x) \, dx = 1\). From Problem 5.2.1.14 (page 266), we can therefore deduce that
\[\hat{H}_t = H_t - \mathbb{E}^\mathbb{P}\left[\eta f_\mathbb{P}(X) \over \lambda f_\mathbb{P}(X)\right] \lambda t = H_t - \eta t\]
is a \(\mathbb{P}\)-martingale as well.

Substituting \(\hat{N}_t = N_t - \lambda t\) and \(\hat{H}_t = H_t - \eta t\) into the stochastic differential equation, we have
\[dZ_t \over Z_t = d\hat{H}_t - d\hat{N}_t\]

and taking integrals, for \(t \geq 0\)
\[\int_0^t dZ_u = \int_0^t Z_u \, d\hat{H}_u - \int_0^t Z_u \, d\hat{N}_u\]
\[Z_t = Z_0 + \int_0^t Z_u \, d\hat{H}_u - \int_0^t Z_u \, d\hat{N}_u\]

Taking expectations,
\[\mathbb{E}^\mathbb{P}(Z_t) = \mathbb{E}^\mathbb{P}(Z_0) + \mathbb{E}^\mathbb{P}\left[\int_0^t Z_u \, d\hat{H}_u\right] - \mathbb{E}^\mathbb{P}\left[\int_0^t Z_u \, d\hat{N}_u\right] = 1\]
since \(Z_0 = 1\), \(\hat{H}_t\) and \(\hat{N}_t\) are both \(\mathbb{P}\)-martingales.

To show that \(Z_t\) is a positive \(\mathbb{P}\)-martingale, by taking integrals for \(s < t\)
\[\int_s^t dZ_u = \int_s^t Z_u \, d\hat{H}_u - \int_s^t Z_u \, d\hat{N}_u\]
\[Z_t - Z_s = \int_s^t Z_u \, d\hat{H}_u - \int_s^t Z_u \, d\hat{N}_u\]
and taking expectations under the filtration $\mathcal{F}_t$,

$$
\mathbb{E}^\mathbb{P} [Z_t - Z_s | \mathcal{F}_s] = \mathbb{E}^\mathbb{P} \left[ \int_s^t Z_u \, d\widehat{H}_u \bigg| \mathcal{F}_s \right] - \mathbb{E}^\mathbb{P} \left[ \int_s^t Z_u \, d\widehat{N}_u \bigg| \mathcal{F}_s \right].
$$

Because $\widehat{H}_t$ and $\widehat{N}_t$ are $\mathbb{P}$-martingales, therefore

$$
\mathbb{E}^\mathbb{P} [Z_t | \mathcal{F}_s] = Z_s.
$$

In addition, since $Z_t > 0$ for $0 \leq t \leq T$, then $|Z_t| = Z_t$ and hence $\mathbb{E}^\mathbb{P} (|Z_t|) = \mathbb{E}^\mathbb{P} (Z_t) = 1 < \infty$ for $0 \leq t \leq T$. Because $Z_t$ is also $\mathcal{F}_t$-adapted, we can conclude that $Z_t$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$.

\[\square\]

11. Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : 0 \leq t \leq T\}$ be a standard Wiener process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Suppose $\theta_t$ is an adapted process, $0 \leq t \leq T$ and $\eta > 0$. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables where each $X_i$, $i = 1, 2, \ldots$ has a probability mass (density) function $\mathbb{P}(X_i)$ under the $\mathbb{P}$ measure. Let $X_1, X_2, \ldots$ also be independent of $N_t$ and $W_t$.

From the definition of the compound Poisson process

$$
M_t = \sum_{i=1}^{N_t} X_i, \quad 0 \leq t \leq T
$$

we consider the Radon–Nikodým derivative process

$$
Z_t = Z_t^{(1)} \cdot Z_t^{(2)}
$$

such that

$$
Z_t^{(1)} = \begin{cases} 
\exp(\lambda - \eta \lambda) \prod_{i=1}^{N_t} \frac{\eta Q(X_i)}{\lambda \mathbb{P}(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\
\exp(\lambda - \eta \lambda) \prod_{i=1}^{N_t} \frac{\eta f^Q(X_i)}{\lambda f^\mathbb{P}(X_i)} & \text{if } X_i \text{ is continuous}
\end{cases}
$$

and

$$
Z_t^{(2)} = \exp \left( \int_0^t \theta_u \, dW_u - \frac{1}{2} \int_0^t \theta_u^2 \, du \right)
$$

where $Q(X_i) (f^Q(X_i))$ is the probability mass (density) function of $X_i$, $i = 1, 2, \ldots$ under the $Q$ measure and $\mathbb{E}^\mathbb{P} \left( \exp \left( \frac{1}{2} \int_0^T \theta_u^2 \, du \right) \right) < \infty$. For the case of continuous random variables $X_i$, $i = 1, 2, \ldots$ we also let $Q$ be absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}$.

Show that $\mathbb{E}^\mathbb{P} (Z_t) = 1$ and $Z_t$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$. 

5.2.3 Girsanov’s Theorem for Jump Processes

Solution: From Problem 4.2.2.1 (page 194) and Problem 5.2.3.10 (page 313), we have shown that $\mathbb{E}^\mathbb{P} (Z_t^{(1)}) = 1$, $\mathbb{E}^\mathbb{P} (Z_t^{(2)}) = 1$, $Z_t^{(1)}$ and $Z_t^{(2)}$ are both $\mathbb{P}$-martingales. From Itô’s lemma,

$$d(Z_t^{(1)} Z_t^{(2)}) = Z_t^{(1)} dZ_t^{(2)} + Z_t^{(2)} d + dZ_t^{(1)} dZ_t^{(2)}$$

where $Z_t^{(1)}$ is the value of $Z_t^{(1)}$ before a jump event. Since $dZ_t^{(1)} dZ_t^{(2)} = 0$, by taking integrals,

$$\int_0^t d(Z_u^{(1)} Z_u^{(2)}) = \int_0^t Z_u^{(1)} dZ_u^{(2)} + \int_0^t Z_u^{(2)} dZ_u^{(1)}$$

$$Z_t^{(1)} Z_t^{(2)} = Z_0^{(1)} Z_0^{(2)} + \int_0^t Z_u^{(1)} dZ_u^{(2)} + \int_0^t Z_u^{(2)} dZ_u^{(1)}.$$ 

Taking expectations under the $\mathbb{P}$ measure,

$$\mathbb{E}^\mathbb{P} (Z_t^{(1)} Z_t^{(2)}) = \mathbb{E}^\mathbb{P} (Z_0^{(1)} Z_0^{(2)}) = 1$$

since $Z_0^{(1)} = 1$, $Z_0^{(2)} = 1$ and both $Z_t^{(1)}$ and $Z_t^{(2)}$ are $\mathbb{P}$-martingales.

To show that $Z_t = Z_t^{(1)} Z_t^{(2)}$ is a positive $\mathbb{P}$-martingale, by taking integrals for $s < t$ of

$$d(Z_t^{(1)} Z_t^{(2)}) = Z_t^{(1)} dZ_t^{(2)} + Z_t^{(2)} dZ_t^{(1)}$$

we have

$$Z_t^{(1)} Z_t^{(2)} = Z_s^{(1)} Z_s^{(2)} + \int_s^t Z_u^{(1)} dZ_u^{(2)} + \int_s^t Z_u^{(2)} dZ_u^{(1)}$$

and taking expectations with respect to the filtration $\mathcal{F}_s$, $s < t$

$$\mathbb{E}^\mathbb{P} (Z_t^{(1)} Z_t^{(2)} | \mathcal{F}_s) = Z_s^{(1)} Z_s^{(2)}$$

since $\mathbb{E}^\mathbb{P} \left( \int_s^t Z_u^{(1)} dZ_u^{(2)} | \mathcal{F}_s \right) = \mathbb{E}^\mathbb{P} \left( \int_s^t Z_u^{(2)} dZ_u^{(1)} | \mathcal{F}_s \right) = 0$.

In addition, since $Z_t^{(1)} > 0$ and $Z_t^{(2)} > 0$ for all $0 \leq t \leq T$, therefore $|Z_t^{(1)} Z_t^{(2)}| = Z_t^{(1)} Z_t^{(2)}$ and hence $\mathbb{E}^\mathbb{P} (|Z_t|) = \mathbb{E}^\mathbb{P} (Z_t) = 1 < \infty$ for $0 \leq t \leq T$. Finally, because $Z_t$ is also $\mathcal{F}_t$-adapted, we can conclude that $Z_t$ is a positive $\mathbb{P}$-martingale for $0 \leq t \leq T$.

*\[\square\]

12. Let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ and $\{W_t : 0 \leq t \leq T\}$ be a standard Wiener process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathcal{F}_t$, $0 \leq t \leq T$. Suppose $\theta_t$ is an adapted process, $0 \leq t \leq T$ and $\eta > 0$. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables where each $X_i$, $i = 1, 2, \ldots$ has a probability mass (density) function $P(X_i)$ ($f^\mathbb{P}(X_i)$) under the $\mathbb{P}$ measure. Let $X_1, X_2, \ldots$ also be independent of $N_t$ and $W_t$.

From the definition of the compound Poisson process

$$M_t = \sum_{i=1}^{N_t} X_i, \quad 0 \leq t \leq T$$
we consider the Radon–Nikodým derivative process
\[ Z_t = Z_t^{(1)} \cdot Z_t^{(2)} \]
such that
\[
Z_t^{(1)} = \begin{cases} 
  e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \frac{\eta Q(X_i)}{\lambda P(X_i)} & \text{if } X_i \text{ is discrete, } P(X_i) > 0 \\
  e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \frac{\eta f(X_i)}{\lambda f(X_i)} & \text{if } X_i \text{ is continuous}
\end{cases}
\]
and
\[ Z_t^{(2)} = e^{-\int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds} \]
where \( Q(X_i) (f^Q(X_i)) \) is the probability mass (density) function of \( X_i, i = 1, 2, \ldots \) under the \( Q \) measure and \( \mathbb{E}^P \left( e^{\frac{1}{2} \int_0^t \theta_u^2 du} \right) < \infty \). For the case of continuous random variables \( X_i, i = 1, 2, \ldots \) we also let \( Q \) be absolutely continuous with respect to \( P \) on \( \mathcal{F} \).

By changing the measure \( P \) to measure \( Q \) such that
\[ \mathbb{E}^P \left( \frac{dQ}{dP} \right)_{\mathcal{F}_t} = Z_t \]
show that under the \( Q \) measure, the process \( M_t = \sum_{i=1}^{N_t} X_i \) is a compound Poisson process with intensity \( \eta > 0 \) where \( X_i, i = 1, 2, \ldots \) is a sequence of independent and identically distributed random variables with probability mass (density) functions \( Q(X_i) (f^Q(X_i)), i = 1, 2, \ldots \), the process \( \tilde{W}_t = W_t + \int_0^t \theta_u du \) is a \( Q \)-standard Wiener process and \( M_t \perp \tilde{W}_t \).

**Solution:** The first two results are given in Problem 5.2.3.9 (page 311) and Problem 4.2.2.10 (page 205).

To prove the final result we need to show that, under the \( Q \) measure,
\[ \mathbb{E}^Q \left( e^{\mu_1 M_t + \mu_2 \tilde{W}_t} \right) = e^{\eta^Q(\mu_1 - 1)} \cdot e^{\frac{1}{2} \mu_2^2 t} \]
which is a joint product of independent moment generating functions of \( M_t \) and \( \tilde{W}_t \sim \mathcal{N}(0, t) \).

By definition,
\[ \mathbb{E}^Q \left( e^{\mu_1 M_t + \mu_2 \tilde{W}_t} \right) = \mathbb{E}^P \left( e^{\mu_1 M_t + \mu_2 \tilde{W}_t} Z_t \right) = \mathbb{E}^P \left( e^{\mu_1 M_t} Z_t^{(1)} \cdot e^{\mu_2 \tilde{W}_t} Z_t^{(2)} \right) \]
and we let
\[ Z_t = Z_t^{(1)} \cdot Z_t^{(2)} \]
where
\[ Z_t^{(1)} = e^{\mu_1 M_t} Z_t^{(1)} \quad \text{and} \quad Z_t^{(2)} = e^{\mu_2 \tilde{W}_t} Z_t^{(2)}. \]
By setting \( \bar{Z}_i^{(1)} \) to denote the value of \( \bar{Z}_i^{(1)} \) before a jump event and using Itô’s lemma, we have
\[
dz = d(\bar{Z}_i^{(1)} - \bar{Z}_i^{(2)}) = \bar{Z}_i^{(1)} d\bar{Z}_i^{(2)} + \bar{Z}_i^{(2)} d\bar{Z}_i^{(1)} + d\bar{Z}_i^{(1)} d\bar{Z}_i^{(2)}.
\]
For \( d\bar{Z}_i^{(1)} \) we can expand, using Taylor’s theorem,
\[
d\bar{Z}_i^{(1)} = \frac{\partial \bar{Z}_i^{(1)}}{\partial M_i} dM_i + \frac{\partial^2 \bar{Z}_i^{(1)}}{\partial M_i^2} (dM_i)^2 + \frac{\partial^3 \bar{Z}_i^{(1)}}{\partial M_i^3} (dM_i)^3 + \frac{\partial^4 \bar{Z}_i^{(1)}}{\partial M_i^4} (dM_i)^4 + \cdots
\]
Because \( dM_i = X_t dN_i \), \( dz_i^{(1)} = Z_i^{(1)} (d\bar{H}_i - d\bar{N}_i) \) where
\[
d\bar{H}_i = \begin{cases} \eta^Q(X_i) \frac{dN_i}{\lambda P(X_i)} - \eta \ dt & \text{if } X_i \text{ is discrete, } P(X_i) > 0 \\ \eta^f(X_i) \frac{dN_i}{\lambda f^P(X_i)} - \eta \ dt & \text{if } X_i \text{ is continuous} \end{cases}
\]
and \( d\bar{N}_i = N_i - \lambda dt \) (see Problem 5.2.3.10, page 313), we can write
\[
d\bar{Z}_i^{(1)} = (u_1 X_i) \bar{Z}_i^{(1)} dN_i + \bar{Z}_i^{(1)} (d\bar{H}_i - d\bar{N}_i) + \frac{1}{21} \left[ (u_1 X_i)^2 \bar{Z}_i^{(1)} dN_i + 2u_1 X_i \bar{Z}_i^{(1)} dN_i (d\bar{H}_i - d\bar{N}_i) \right] + \cdots
\]
\[
= \left[ (u_1 X_i)^3 \bar{Z}_i^{(1)} dN_i + 3(u_1 X_i)^2 \bar{Z}_i^{(1)} dN_i (d\bar{H}_i - d\bar{N}_i) \right] + \cdots
\]
\[
= \left[ (u_1 X_i)^3 \bar{Z}_i^{(1)} dN_i + 3(u_1 X_i)^2 \bar{Z}_i^{(1)} dN_i (d\bar{H}_i - d\bar{N}_i) \right] + \cdots
\]
\[
= \left[ (u_1 X_i)^3 \bar{Z}_i^{(1)} dN_i + 3(u_1 X_i)^2 \bar{Z}_i^{(1)} dN_i (d\bar{H}_i - d\bar{N}_i) \right] + \cdots
\]
\[
= (e^{\mu X_i} - 1) \bar{Z}_i^{(1)} dN_i + \bar{Z}_i^{(1)} (d\bar{H}_i - d\bar{N}_i) + (e^{\mu X_i} - 1) \bar{Z}_i^{(1)} dN_i (d\bar{H}_i - d\bar{N}_i).
\]
Since
\[
d\bar{H}_i - d\bar{N}_i = \begin{cases} \eta^Q(X_i) \frac{dN_i}{\lambda P(X_i)} - 1 \ dN_i - (\eta - \lambda) dt & \text{if } X_i \text{ is discrete, } P(X_i) > 0 \\ \eta^f(X_i) \frac{dN_i}{\lambda f^P(X_i)} - 1 \ dN_i - (\eta - \lambda) dt & \text{if } X_i \text{ is continuous} \end{cases}
\]
and
\[ dN_t(d\hat{H}_t - d\hat{N}_t) = \begin{cases} \eta Q(X_t) \frac{P(X_t)}{P(X_t)} - 1 \quad &\text{if } X_t \text{ is discrete, } P(X_t) > 0 \\ \eta f'(X_t) \frac{P(X_t)}{P(X_t)} - 1 \quad &\text{if } X_t \text{ is continuous} \end{cases} \]

therefore
\[ d\tilde{Z}_t^{(n)} = \left\{ \begin{array}{l} \eta Q(X_t) \left( e^{\mu_t X_t - 1} \right) \tilde{Z}_t^{(1)} dN_t + \tilde{Z}_t^{(1)} \left( d\hat{H}_t - d\hat{N}_t \right) \quad \text{if } X_t \text{ is discrete, } P(X_t) > 0 \\ \eta f'(X_t) \left( e^{\mu_t X_t - 1} \right) \tilde{Z}_t^{(1)} dN_t + \tilde{Z}_t^{(1)} \left( d\hat{H}_t - d\hat{N}_t \right) \quad \text{if } X_t \text{ is continuous} \end{array} \right. \]

For the case of \( d\tilde{Z}_t^{(2)} \), by applying Taylor’s theorem
\[
d\tilde{Z}_t^{(2)} = \frac{\partial \tilde{Z}_t^{(2)}}{\partial \tilde{W}_t} d\tilde{W}_t + \frac{\partial \tilde{Z}_t^{(2)}}{\partial \tilde{Z}_t^{(2)}} d\tilde{Z}_t^{(2)} + \frac{1}{2!} \left[ \frac{\partial^2 \tilde{Z}_t^{(2)}}{\partial (\tilde{W}_t)^2} (d\tilde{W}_t)^2 + 2 \frac{\partial^2 \tilde{Z}_t^{(2)}}{\partial \tilde{W}_t \partial \tilde{Z}_t^{(2)}} (d\tilde{W}_t)(d\tilde{Z}_t^{(2)}) + \frac{\partial^2 \tilde{Z}_t^{(2)}}{\partial (\tilde{Z}_t^{(2)})^2} (d\tilde{Z}_t^{(2)})^2 \right] + \ldots
\]

and since \( d\tilde{W}_t = dW_t + \theta \ dt \) and \( d\tilde{Z}_t^{(2)} = -\theta \tilde{Z}_t^{(2)} dW_t \) (see Problem 4.2.2.2, page 196), we can write
\[
d\tilde{Z}_t^{(2)} = \frac{1}{2} u_2 \tilde{Z}_t^{(2)} \ dt + (u_2 - \theta) \tilde{Z}_t^{(2)} dW_t.
\]

Since \( dN_t dW_t = 0 \) we have
\[
\tilde{Z}_t^{(1)} d\tilde{Z}_t^{(2)} = \frac{1}{2} u_2 \tilde{Z}_t^{(2)} \ dt + (u_2 - \theta) \tilde{Z}_t^{(2)} dW_t,
\]

\[
\tilde{Z}_t^{(2)} d\tilde{Z}_t^{(1)} = \begin{cases} \eta Q(X_t) \left( e^{\mu_t X_t - 1} \right) \tilde{Z}_t dN_t + \tilde{Z}_t \left( d\hat{H}_t - d\hat{N}_t \right) \quad \text{if } X_t \text{ is discrete, } P(X_t) > 0 \\ \eta f'(X_t) \left( e^{\mu_t X_t - 1} \right) \tilde{Z}_t dN_t + \tilde{Z}_t \left( d\hat{H}_t - d\hat{N}_t \right) \quad \text{if } X_t \text{ is continuous} \end{cases}
\]

and
\[
d\tilde{Z}_t^{(1)} d\tilde{Z}_t^{(2)} = 0.
\]

Without loss of generality, we let \( X_t \) be a continuous random variable and by letting \( dN_t = d\hat{N}_t + \lambda \ dt \) we have
\[
d\tilde{Z}_t = \frac{1}{2} u_2 \tilde{Z}_t \ dt + (u_2 - \theta) \tilde{Z}_t dW_t + \eta f'(X_t) \left( e^{\mu_t X_t - 1} \right) \tilde{Z}_t dN_t + \tilde{Z}_t \left( d\hat{H}_t - d\hat{N}_t \right)
\]
5.2.3 Girsanov’s Theorem for Jump Processes

\[
\begin{align*}
&= \left[ \frac{\eta f^Q(X_t)}{f^P(X_t)} e^{\mu_1 X_t} - 1 \right] \int \frac{1}{2} u^2 \right] Z_t \, dt + (u_2 - \theta_t) Z_t \, dW_t \\
&\quad + \int \left[ \frac{\eta f^Q(X_t)}{f^P(X_t)} e^{\mu_1 X_t} - 1 \right] Z_t \, d\hat{N}_t + Z_t (d\hat{H}_t - d\hat{N}_t).
\end{align*}
\]

Taking integrals,

\[
Z_t = 1 + \int_0^t \left[ \frac{\eta f^Q(X_s)}{f^P(X_s)} e^{\mu_1 X_s} - 1 \right] Z_s \, ds + \int_0^t (u_2 - \theta_s) Z_s \, dW_s \\
\quad + \int_0^t \left[ \frac{\eta f^Q(X_s)}{f^P(X_s)} e^{\mu_1 X_s} - 1 \right] Z_s \, d\hat{N}_s + \int_0^t \left( Z_t (d\hat{H}_s - d\hat{N}_s)
\right)
\]

where \( Z_0 = 1 \) and because the jump size variable \( X_t \) is independent of \( N_t \) and \( W_t \), and \( \hat{N}_t \), \( \hat{H}_t \) and \( W_t \) are \( \mathbb{P} \)-martingales, by taking expectations under the \( \mathbb{P} \) measure we have

\[
\mathbb{E}^\mathbb{P} \left( Z_t \right) = 1 + \int_0^t \left\{ \mathbb{E}^\mathbb{P} \left[ \frac{\eta f^Q(X_s)}{f^P(X_s)} e^{\mu_1 X_s} - 1 \right] + \frac{1}{2} u_s^2 \right\} \mathbb{E}^\mathbb{P} \left( Z_s \right) \, ds.
\]

By differentiating the integrals with respect to \( t \),

\[
\frac{d}{dt} \mathbb{E}^\mathbb{P} \left( Z_t \right) = \left\{ \mathbb{E}^\mathbb{P} \left[ \frac{\eta f^Q(X_t)}{f^P(X_t)} e^{\mu_1 X_t} - 1 \right] + \frac{1}{2} u_t^2 \right\} \mathbb{E}^\mathbb{P} \left( Z_t \right)
\]

or

\[
\frac{dm_t}{dt} - \left[ \eta \left( \phi^Q_X(u_t) - 1 \right) + \frac{1}{2} u_t^2 \right] m_t = 0
\]

where \( m_t = \mathbb{E}^\mathbb{P} \left( Z_t \right) \) and \( \phi^Q_X(u_t) = \mathbb{E}^\mathbb{Q} \left( e^{\mu_1 X_t} \right) \).

By setting the integrating factor to be \( I = e^{-\int (\eta \phi_X(u_t) - 1 + \frac{1}{2} u_t^2) \, dt} = e^{-\int (\eta \phi_X(u_t) - 1 + \frac{1}{2} u_t^2) \, dt} \) and multiplying the differential equation with \( I \), we have

\[
\frac{d}{dt} \left( m_t e^{-\eta (\phi_X(u_t) - 1) - \frac{1}{2} u_t^2} \right) = 0 \quad \text{or} \quad e^{-\eta (\phi_X(u_t) - 1) - \frac{1}{2} u_t^2} \mathbb{E}^\mathbb{P} \left( Z_t \right) = C
\]

where \( C \) is a constant. Since \( \mathbb{E}^\mathbb{P} \left( Z_0 \right) = \mathbb{E}^\mathbb{P} \left( e^{\mu_1 M_0} Z_0^{(1)} \cdot e^{\mu_2 \hat{W}_0} Z_0^{(2)} \right) = 1 \), therefore \( C = 1 \). Thus, we finally obtain

\[
\mathbb{E}^\mathbb{Q} \left( e^{\mu_1 M_t + \mu_2 \hat{W}_t} \right) = \mathbb{E}^\mathbb{P} \left( Z_t \right) = e^{\eta (\phi_X(u_t) - 1) + \frac{1}{2} u_t^2}.
\]

Since the joint moment generating function of

\[
\mathbb{E}^\mathbb{Q} \left( e^{\mu_1 M_t + \mu_2 \hat{W}_t} \right) = e^{\eta (\phi_X(u_t) - 1) + \frac{1}{2} u_t^2}
\]
can be expressed as a product of the moment generating functions for $M_t$ (which is a compound Poisson process with intensity $\eta$) and $\tilde{W}_t \sim \mathcal{N}(0, t)$, respectively, we can deduce that $M_t$ and $\tilde{W}_t$ are independent.

\textit{N.B.} The same conclusion can also be obtained for the case of treating $X_t$ as a discrete random variable.

\begin{align*}
\frac{dS_t}{S_t} &= (J - 1)dN_t
\end{align*}

where $J$ is a constant jump amplitude if $N$ jumps at time $t$ and

\[ dN_t = \begin{cases} 1 & \text{with probability } \lambda \ dt \\ 0 & \text{with probability } 1 - \lambda dt. \end{cases} \]

Let $r$ be the risk-free interest rate.

By considering the Radon–Nikodým derivative process, for $\eta > 0$

\[ Z_t = e^{(\lambda - \eta)\frac{\eta t}{\lambda} N_t}, \quad 0 \leq t \leq T \]

show that by changing the measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$, the above stochastic differential equation is

\begin{align*}
\frac{dS_t}{S_t} &= (J - 1)dN_t
\end{align*}

where $N_t \sim \text{Poisson}(\eta t)$, $\eta = r(J - 1)^{-1} > 0$ provided $J > 1$.

Is the market arbitrage free and complete under the $\mathbb{Q}$ measure?

**Solution:** Let $\eta > 0$ and from the Radon–Nikodým derivative process

\[ Z_t = e^{(\lambda - \eta)\frac{\eta t}{\lambda} N_t}. \]

By changing the measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$ on the filtration $\mathcal{F}_s$, $0 \leq s \leq t$ then under $\mathbb{Q}$, $N_t \sim \text{Poisson}(\eta t)$ and the discounted asset price $e^{-rt}S_t$ is a $\mathbb{Q}$-martingale.

By setting $S_t$ to denote the value of $S_t$ before a jump event and by expanding $d(e^{-rt}S_t)$, we have

\begin{align*}
\frac{d(e^{-rt}S_t)}{e^{-rt}S_t} &= -re^{-rt}S_t \ dt + e^{-rt}dS_t \\
&= -re^{-rt}S_t \ dt + e^{-rt}S_t(J - 1)dN_t \\
&= e^{-rt}(r + \eta(J - 1)) \ dt + e^{-rt}S_t(J - 1)(dN_t - \eta dt).
\end{align*}
5.2.4 Risk-Neutral Measure for Jump Processes

Since the compensated Poisson process $N_t - \eta t$ is a $\mathbb{Q}$-martingale (see Problem 5.2.1.9, page 262), then, in order for $e^{-rt}S_t$ to be a $\mathbb{Q}$-martingale, we can set

$$\eta = r(J - 1)^{-1}.$$

Since $\eta > 0$ we require $J > 1$. Thus, under the risk-neutral measure $\mathbb{Q}$,

$$\frac{dS_t}{S_{r^{-}}}= (J - 1)dN_t,$$

where $N_t \sim \text{Poisson}(\eta t)$.

The market is arbitrage free since we can construct a risk-neutral measure $\mathbb{Q}$ on the filtration $\mathcal{F}_s, 0 \leq s \leq t$. Since $\mathbb{Q}$ is unique, the market is also complete.

2. Pure Jump Process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{N_t : 0 \leq t \leq T\}$ be a Poisson process with intensity $\lambda > 0$ relative to the filtration $\mathcal{F}_t, 0 \leq t \leq T$. Suppose $S_t$ follows a pure jump process

$$\frac{dS_t}{S_{r^{-}}}= (J_t - 1)dN_t,$$

where $J_t$ is the jump variable with mean $\mathbb{E}^\mathbb{P}(J_t) = \bar{J}$ if $N$ jumps at time $t$ and

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda \, dt \\ 0 & \text{with probability } 1 - \lambda \, dt. \end{cases}$$

Let $r$ be the risk-free interest rate. Assume that $J_t$ is independent of $N_t$ and consider the Radon–Nikodym derivative process

$$Z_t = \begin{cases} e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \frac{\eta Q(X_i)}{\lambda P(X_i)} & \text{if } X_i \text{ is discrete, } P(X_i) > 0 \\ e^{(\lambda - \eta)t} \prod_{i=1}^{N_t} \frac{\eta f^Q(X_i)}{\lambda f^P(X_i)} & \text{if } X_i \text{ is continuous} \end{cases}$$

where $\eta > 0, X_i = J_i - 1, P(X) (f^P(X))$ and $Q(X) (f^Q(X))$ are the probability mass (density) functions of $P$ and $Q$, respectively.

Show, by changing the measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$, the above SDE can be written as

$$\frac{dS_t}{S_{r^{-}}}= dM_t,$$

where $M_t = \sum_{i=1}^{N_t} (J_i - 1)$ is a compound Poisson process with $J_i, i = 1, 2, \ldots$ a sequence of independent and identically distributed jump amplitude random variables having intensity

$$\eta = \frac{r}{\mathbb{E}^\mathbb{Q}(J_i) - 1} > 0$$

provided $\mathbb{E}^\mathbb{Q}(J_i) > 1$. 

Is the market arbitrage free and complete under the $\mathbb{Q}$ measure?

By assuming that the jump amplitude remains unchanged with the change of measure, find the corresponding SDE under the equivalent martingale risk-neutral measure, $\mathbb{Q}_M$.

Solution: From Problem 5.2.2.1 (page 281) we can express the pure jump process as

$$\frac{dS_t}{S_t} = dM_t$$

where $M_t = \sum_{i=1}^{N_t} (J_i - 1)$ is a compound Poisson process with intensity $\lambda > 0$.

Under the risk-neutral measure $\mathbb{Q}$ on the filtration $\mathcal{F}_s$, $0 \leq s \leq t$ we let $\eta > 0$ and by considering the Radon–Nikodým derivative process $Z_t = \begin{cases} \exp\left(\lambda - \eta \right) t \prod_{i=1}^{N_t} \frac{\mathbb{Q}(X_i)}{\mathbb{P}(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\ \exp\left(\lambda - \eta \right) t \prod_{i=1}^{N_t} \frac{\mathbb{Q}(X_i)}{\mathbb{P}(X_i)} & \text{if } X_i \text{ is continuous} \end{cases}$

where $X_i = J_i - 1$, $\mathbb{P}(X) (f^\mathbb{P}(X))$ and $\mathbb{Q}(X) (f^\mathbb{Q}(X))$ are the probability mass (density) functions of $\mathbb{P}$ and $\mathbb{Q}$ respectively, then by changing the measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$ we have that, under $\mathbb{Q}$, $M_t = \sum_{i=1}^{N_t} (J_i - 1)$ is a compound Poisson process with intensity $\eta > 0$ and the discounted asset price $e^{-rt}S_t$ is a $\mathbb{Q}$-martingale.

By letting $S_{t^-} = S_t$ to denote the value of $S_t$ before a jump event and expanding $d(e^{-rt}S_t)$, we have

$$d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}dS_t + e^{-rt}dM_t$$

$$= e^{-rt}S_t (-r + \eta \mathbb{E}^\mathbb{Q}(J_t - 1)) dt + e^{-rt}S_t (dM_t - \eta \mathbb{E}^\mathbb{Q}(J_t - 1) dt).$$

Given that the compensated compound Poisson process $M_t - \eta \mathbb{E}^\mathbb{Q}(J_t - 1)t$ is a $\mathbb{Q}$-martingale (see Problem 5.2.1.14, page 266), and in order for $e^{-rt}S_t$ to be a $\mathbb{Q}$-martingale, we can set

$$\eta = \frac{r}{\mathbb{E}^\mathbb{Q}(J_t) - 1}$$

provided $\mathbb{E}^\mathbb{Q}(J_t) > 1$.

Thus, under the risk-neutral measure $\mathbb{Q}$ we have

$$\frac{dS_t}{S_{t^-}} = (J_t - 1)dN_t$$

where $N_t \sim \text{Poisson}(\eta t)$, $\eta = r(\mathbb{E}^\mathbb{Q}(J_t - 1))^{-1} > 0$.

The market is arbitrage free since we can construct a risk-neutral measure $\mathbb{Q}$ on the filtration $\mathcal{F}_s$, $0 \leq s \leq t$. However, the market model is incomplete as $\mathbb{Q}$ is not unique given that we can vary the jump distribution $J_t$. 
5.2.4 Risk-Neutral Measure for Jump Processes

By assuming that the distribution of the jump amplitude \( J_t \) remains unchanged under the change of measure, there exists an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \sim \mathbb{Q} \) such that \( e^{-rS_t} \) is a \( \mathbb{Q}_M \)-martingale. Thus, under the equivalent martingale risk-neutral measure \( \mathbb{Q}_M \),

\[
\mathbb{E}^{\mathbb{Q}_M} (J_t - 1) = \mathbb{E}^{\mathbb{P}} (J_t - 1) = \bar{J} - 1
\]

and the corresponding SDE is

\[
\frac{dS_t}{S_t} = (J_t - 1) dN_t
\]

where \( N_t \sim \text{Poisson} (\eta t) \), \( \eta = r(\bar{J} - 1)^{-1} \) provided \( \bar{J} > 1 \).

The market model is arbitrage free since we can construct an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \) on the filtration \( \mathcal{F}_s \), \( 0 \leq s \leq t \). However, the market is incomplete as \( \mathbb{Q}_M \) is not unique given that we can still vary the jump size distribution \( J_t \) and hence \( J \).

3. Simple Jump Diffusion Process. Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) and \( \{W_t : 0 \leq t \leq T\} \) be a standard Wiener process defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_t \), \( 0 \leq t \leq T \). Suppose the stock price \( S_t \) follows a jump diffusion process

\[
\frac{dS_t}{S_t} = (\mu - D) dt + \sigma dW_t + (J - 1) dN_t
\]

where \( W_t \perp N_t \), \( J \) is a constant jump amplitude if \( N \) jumps at time \( t \) and

\[
dN_t = \begin{cases} 
1 \text{ with probability } \lambda dt \\
0 \text{ with probability } 1 - \lambda dt
\end{cases}
\]

The constant parameters \( \mu, D \) and \( \sigma \) are the drift, continuous dividend yield and volatility, respectively. In addition, let \( B_t \) be the risk-free asset having the following differential equation:

\[
dB_t = rB_t dt
\]

where \( r \) is the risk-free interest rate.

By considering the Radon–Nikodým derivative process

\[
Z_t = Z_t^{(1)} \cdot Z_t^{(2)}
\]

such that

\[
Z_t^{(1)} = e^{(\lambda - \eta)\left( \frac{\eta}{\lambda} \right) N_t}
\]

and

\[
Z_t^{(2)} = e^{-\int_0^t \theta dW_u - \frac{1}{2} \int_0^t \theta^2 du}
\]

where \( \theta \in \mathbb{R}, \eta > 0 \) and \( \mathbb{E}^{\mathbb{P}} \left( e^{\frac{1}{2} \int_0^T \theta^2 du} \right) < \infty \) show that, under the risk-neutral measure \( \mathbb{Q} \), the SDE can be written as

\[
\frac{dS_t}{S_t} = (r - D - \eta(J - 1)) dt + \sigma d\widetilde{W}_t + (J - 1) dN_t
\]
5.2.4 Risk-Neutral Measure for Jump Processes

where \( \tilde{W}_t = W_t + \left( \frac{\mu - r + \eta(J - 1)}{\sigma} \right) t \) is the \( \mathbb{Q} \)-standard Wiener process and \( N_t \sim \text{Poisson}(\eta t) \).

Is the market arbitrage free and complete under the \( \mathbb{Q} \) measure?

By assuming the jump component is uncorrelated with the market (i.e., the jump intensity remains unchanged with the change of measure), find the corresponding SDE under the equivalent martingale risk-neutral measure \( \mathbb{Q}_M \).

Is the market arbitrage free and complete under the \( \mathbb{Q}_M \) measure?

**Solution:** Let \( \theta \in \mathbb{R} \) and \( \eta > 0 \) and consider the Radon–Nikodym derivative process

\[
Z_t = Z_t^{(1)} : Z_t^{(2)}
\]

such that

\[
Z_t^{(1)} = e^{(\xi - \eta)\theta_N} \left( \frac{\eta}{\lambda} \right)^N
\]

and

\[
Z_t^{(2)} = e^{-\int_0^t \theta dW - \frac{1}{2} \int_0^t \theta^2 dt}
\]

where \( \mathbb{E}^\mathbb{P}\left( e^{\frac{1}{2} \int_0^T \theta^2 du} \right) < \infty \). By changing the measure \( \mathbb{P} \) to the risk-neutral measure \( \mathbb{Q} \) on the filtration \( \mathcal{F}_s \), \( 0 \leq s \leq t \) such that

\[
\mathbb{E}^\mathbb{P}\left( \frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t \right) = \frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t = Z_t
\]

then from Problem 5.2.3.5 (page 303), under the \( \mathbb{Q} \) measure, the Poisson process \( N_t \sim \text{Poisson}(\eta t) \) has intensity \( \eta > 0 \), the process \( \tilde{W}_t = W_t + \int_0^t \theta du \) is a \( \mathbb{Q} \)-standard Wiener process and \( N_t \perp \tilde{W}_t \).

In the presence of dividends, the strategy of holding a single stock is no longer self-financing as it pays out dividends at a rate \( DS_i dt \). By letting \( S_i \) denote the value of \( S_i \) before a jump event, at time \( t \) we let the portfolio \( \Pi_t \) be valued as

\[
\Pi_t = \phi_t S_t + \psi_t B_t
\]

where \( \phi_t \) and \( \psi_t \) are the units invested in \( S_t \) and the risk-free asset \( B_t \), respectively. Given that the holder will receive \( DS_i dt \) for every stock held, then

\[
d\Pi_t = \phi_t (dS_t + DS_i dt) + \psi_t rB_t dt
= \phi_t \left[ \mu S_t dt + \sigma S_t dW_i + (J - 1)S_t dN_i \right] + \psi_t rB_t dt
= r\Pi_t dt + \phi_t S_t \left[ (\mu - r) dt + \sigma dW_i + (J - 1) dN_i \right].
\]

By substituting \( W_t = \tilde{W}_t - \int_0^t \theta du \) and taking note that the compensated Poisson process \( N_t - \eta t \) is a \( \mathbb{Q} \)-martingale, we have

\[
d\Pi_t = r\Pi_t dt + \phi_t S_t \left[ (\mu - r - \sigma \theta + \eta (J - 1)) dt + \sigma d\tilde{W}_t + (J - 1) (dN_i - \eta dt) \right].
\]
5.2.4 Risk-Neutral Measure for Jump Processes

Since both \( \tilde{W}_t \) and \( N_t - \eta t \) are \( \mathbb{Q} \)-martingales and in order for the discounted portfolio \( e^{-rt} \Pi_t \) to be a \( \mathbb{Q} \) martingale

\[
\begin{align*}
    d(e^{-rt} \Pi_t) &= -re^{-rt} \Pi_t \, dt + e^{-rt} d\Pi_t \\
    &= e^{-rt} \phi_t S_t \left[ (\mu - r - \sigma \theta + \eta (J - 1)) \, dt + \sigma d\tilde{W}_t + (J - 1)(dN_t - \eta \, dt) \right]
\end{align*}
\]

we set

\[
    \theta = \frac{\mu - r + \eta (J - 1)}{\sigma}.
\]

Therefore, the jump diffusion process under \( \mathbb{Q} \) is

\[
    \frac{dS_t}{S_t} = (\mu - D) \, dt + \sigma \left( d\tilde{W}_t - \theta \, dt \right) + (J - 1) dN_t
\]

where \( \tilde{W}_t = W_t - \left( \frac{\mu - r + \eta (J - 1)}{\sigma} \right) t \) is a \( \mathbb{Q} \)-standard Wiener process and \( N_t \sim \text{Poisson}(\eta t) \).

The market is arbitrage free since we can construct a risk-neutral measure \( \mathbb{Q} \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \). The market model is incomplete as \( \mathbb{Q} \) is not unique, since there are more degrees of freedom in selecting \( J \) and \( \eta \).

Under the assumption that the jump component is uncorrelated with the market, there exists an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \sim \mathbb{Q} \) such that \( e^{-rt} \Pi_t \) is a \( \mathbb{Q}_M \)-martingale.

Therefore, under \( \mathbb{Q}_M \),

\[
    N_t \sim \text{Poisson}(\lambda t)
\]

and the SDE becomes

\[
    \frac{dS_t}{S_t} = (r - D - \lambda (J - 1)) \, dt + \sigma d\tilde{W}_t + (J - 1) dN_t
\]

where \( \tilde{W}_t = W_t - \left( \frac{\mu - r + \lambda (J - 1)}{\sigma} \right) t \) is a \( \mathbb{Q}_M \)-standard Wiener process and \( N_t \sim \text{Poisson}(\lambda t) \).

The market is arbitrage free since we can construct an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \). However, the market is not complete as \( \mathbb{Q}_M \) is not unique, since we can still vary the jump amplitude \( J \).

\[\square\]

4. General Jump Diffusion Process (Merton’s Model). Let \( \{N_t : 0 \leq t \leq T\} \) be a Poisson process with intensity \( \lambda > 0 \) and \( \{W_t : 0 \leq t \leq T\} \) be a standard Wiener process defined on the probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with respect to the filtration \( \mathcal{F}_s, 0 \leq s \leq t \). Suppose the asset price \( S_t \) follows a jump diffusion process

\[
    \frac{dS_t}{S_t} = (\mu - D) \, dt + \sigma dW_t + (J_t - 1) dN_t
\]
where $J_t$ is the jump variable with mean $\mathbb{E}^\mathbb{P}(J_t) = \bar{J}$ if $N$ jumps at time $t$ and

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt. \end{cases}$$

Assume that $W_t$, $N_t$, and $J_t$ are mutually independent. The constant parameters $\mu$, $D$ and $\sigma$ are the drift, continuous dividend yield and volatility, respectively. In addition, let $B_t$ be the risk-free asset having the following differential equation:

$$dB_t = rB_t \, dt$$

where $r$ is the risk-free interest rate.

By considering the Radon–Nikodým derivative process

$$Z_t = Z_t^{(1)} \cdot Z_t^{(2)}$$

such that

$$Z_t^{(1)} = \begin{cases} e^{(d-\eta)\sum_{i=1}^{N_t} \eta Q(X_i) \mathbb{P}(X_i)} & \text{if } X_t \text{ is discrete, } \mathbb{P}(X_t) > 0 \\ e^{(d-\eta)\sum_{i=1}^{N_t} \eta f^Q(X_i) \mathbb{F}^\mathbb{P}(X_i)} & \text{if } X_t \text{ is continuous} \end{cases}$$

and

$$Z_t^{(2)} = e^{-\int_0^t \theta dW_u - \frac{1}{2} \int_0^t \theta^2 \, du}$$

where $\theta \in \mathbb{R}$, $\eta > 0$, $Q(X_i)$ ($f^Q(X_i)$) is the probability mass (density) function of $X_i = J_i - 1$, $i = 1, 2, \ldots$ under the $\mathbb{Q}$ measure and $\mathbb{E}^\mathbb{P} \left( e^{\frac{1}{2} \int_0^T \theta^2 \, du} \right) < \infty$ show that, under risk-neutral measure $\mathbb{Q}$, the above SDE can be written as

$$\frac{dS_t}{S_t} = (r - D - \eta \mathbb{E}^\mathbb{Q}((J_t - 1)) \, dt + \sigma d\tilde{W}_t + (J_t - 1) dN_t$$

where $\tilde{W}_t = W_t - \left( \frac{\mu - r + \eta \mathbb{E}^\mathbb{Q}(J_t - 1)}{\sigma} \right) t$ is the $\mathbb{Q}$-standard Wiener process and $N_t$ ~ Poisson($\eta$), $\eta > 0$.

Is the market model arbitrage free and complete under the $\mathbb{Q}$ measure?

By assuming the jump component is uncorrelated with the market (i.e., the jump intensity and jump size distribution remain unchanged with the change of measure), find the corresponding SDE under the equivalent martingale risk-neutral measure $\mathbb{Q}_M$.

Is the market model arbitrage free and complete under the $\mathbb{Q}_M$ measure?

**Solution:** From Problem 5.2.2.1 (page 281) we can express the jump diffusion process as

$$\frac{dS_t}{S_t} = (\mu - D) \, dt + \sigma dW_t + dM_t.$$
5.2.4 Risk-Neutral Measure for Jump Processes

where $M_t = \sum_{i=1}^{N_t} (J_i - 1)$ is a compound Poisson process with intensity $\lambda > 0$ such that the jump size (or amplitude) $J_i$, $i = 1, 2, \ldots$ is a sequence of independent and identically distributed random variables.

Let $\theta \in \mathbb{R}$, $\eta > 0$ and $X_t = J_t - 1$, $i = 1, 2, \ldots$ be a sequence of independent and identically distributed random variables where each $X_t$, $i = 1, 2, \ldots$ has a probability mass (density) function $\mathbb{P}(X_i) > 0$ under the $\mathbb{P}$ measure. We consider the Radon–Nikodým derivative process

$$Z_t = Z_t^{(1)} \cdot Z_t^{(2)}$$

such that

$$Z_t^{(1)} = \begin{cases} e^{(\lambda - \eta) t} \prod_{i=1}^{N_t} \frac{\eta Q(X_i)}{\lambda P(X_i)} & \text{if } X_i \text{ is discrete, } \mathbb{P}(X_i) > 0 \\ e^{(\lambda - \eta) t} \prod_{i=1}^{N_t} \frac{\eta f Q(X_i)}{\lambda f P(X_i)} & \text{if } X_i \text{ is continuous} \end{cases}$$

and

$$Z_t^{(2)} = e^{-\int_0^t \theta dW_u - \frac{1}{2} \int_0^t \theta^2 du}$$

where $Q(X_i)$ ($f Q(X_i)$) is the probability mass (density) function of $X_i$, $i = 1, 2, \ldots$ under the $Q$ measure and $\mathbb{E}^Q \left( e^{\frac{1}{2} \int_0^T \theta^2 du} \right) < \infty$. By changing the measure $\mathbb{P}$ to the risk-neutral measure $Q$ on the filtration $\mathcal{F}_s$, $0 \leq s \leq t$ such that

$$\mathbb{E}^\mathbb{P} \left( \frac{dQ}{d\mathbb{P}} \bigg| \mathcal{F}_t \right) = \frac{dQ}{d\mathbb{P}} \bigg| \mathcal{F}_t = Z_t$$

then from Problem 5.2.3.12 (page 317), under $Q$, the compound Poisson process $M_t = \sum_{i=1}^{N_t} (J_i - 1)$ has intensity $\eta > 0$, the process $\tilde{W}_t = W_t + \int_0^t \theta du$ is a $Q$-standard Wiener process and $M_t \perp \tilde{W}_t$.

In the presence of dividends, the simple strategy of holding a single risky asset is no longer self-financing as the asset pays out dividends at a rate $DS_t \, dt$. By letting $S_t^-$ denote the value of $S_t$ before a jump event at time $t$, we let the portfolio $\Pi_t$ be valued as

$$\Pi_t = \phi_t S_t^+ + \psi_t B_t$$

where $\phi_t$ and $\psi_t$ are the units invested in $S_t$ and the risk-free asset $B^d_t$, respectively. Given that the holder will receive $DS_t \, dt$ for every risky asset held, then

$$d\Pi_t = \phi_t(dS_t + DS_t \, dt) + \psi_t r B_t \, dt$$

$$= \phi_t \left[ \mu S_t \, dt + \sigma S_t dW_t + dM_t \right] + \psi_t r B_t \, dt$$

$$= r \Pi_t \, dt + \phi_t S_t \left[ (\mu - r) \, dt + \sigma dW_t + dM_t \right].$$
By substituting \( W_t = \tilde{W}_t - \int_0^t \theta \, du \) and taking note that the compensated compound Poisson process \( M_t - \eta \mathbb{E}^\mathbb{Q}(J_t - 1) t \) is a \( \mathbb{Q} \)-martingale, we have
\[
d\Pi_t = r \Pi_t \, dt + \phi_t S_t \left[ \left( \mu - r - \sigma \theta + \eta \mathbb{E}^\mathbb{Q}(J_t - 1) \right) \, dt + \sigma d\tilde{W}_t + dM_t - \eta \mathbb{E}^\mathbb{Q}(J_t - 1) \, dt \right].
\]
Since both \( \tilde{W}_t \) and \( M_t - \eta \mathbb{E}^\mathbb{Q}(J_t - 1) t \) are \( \mathbb{Q} \)-martingales and in order for the discounted portfolio \( e^{-rt} \Pi_t \) to be a \( \mathbb{Q} \)-martingale
\[
d(e^{-rt} \Pi_t) = -re^{-rt} \Pi_t \, dt + e^{-rt} d\Pi_t
\]
we set
\[
\theta = \frac{\mu - r + \eta \mathbb{E}^\mathbb{Q}(J_t - 1)}{\sigma}.
\]
Therefore, the jump diffusion process under \( \mathbb{Q} \) is
\[
\frac{dS_t}{S_{t-}} = (\mu - D) \, dt + \sigma \left( d\tilde{W}_t - \theta \, dt \right) + (J_t - 1) dN_t
\]
\[
= (r - D - \eta \mathbb{E}^\mathbb{Q}(J_t - 1)) \, dt + \sigma d\tilde{W}_t + (J_t - 1) dN_t
\]
where \( \tilde{W}_t = W_t - \left( \frac{\mu - r + \eta \mathbb{E}^\mathbb{Q}(J_t - 1)}{\sigma} \right) t \) and \( N_t \sim \text{Poisson}(\eta t) \).
The market is arbitrage free since we can construct a risk-neutral measure \( \mathbb{Q} \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \). However, the market model is incomplete as \( \mathbb{Q} \) is not unique, given that there are more degrees of freedom in selecting \( \eta \) and \( J_t \).
Under the assumption that the jump component is uncorrelated with the market, there exists an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \sim \mathbb{Q} \) such that \( e^{-rt} \Pi_t \) is a \( \mathbb{Q}_M \)-martingale and hence
\[
\mathbb{E}^{\mathbb{Q}_M}(J_t - 1) = \mathbb{E}^{\mathbb{P}}(J_t - 1) = J - 1 \quad \text{and} \quad N_t \sim \text{Poisson}(\lambda t).
\]
Thus, under \( \mathbb{Q}_M \) the jump diffusion process becomes
\[
\frac{dS_t}{S_{t-}} = (r - D - \lambda(J - 1)) \, dt + \sigma d\tilde{W}_t + (J_t - 1) dN_t
\]
where \( \tilde{W}_t = W_t - \left( \frac{\mu - r + \lambda(J - 1)}{\sigma} \right) t \) is a \( \mathbb{Q}_M \)-standard Wiener process and \( N_t \sim \text{Poisson}(\lambda t) \).
The market is arbitrage free since we can construct an equivalent martingale risk-neutral measure \( \mathbb{Q}_M \) on the filtration \( \mathcal{F}_s, 0 \leq s \leq t \). However, the market is incomplete as \( \mathbb{Q}_M \) is not unique, since we can still vary the jump size distribution \( J_t \).
\( \Box \)
Appendix A
Mathematics Formulae

Indices
\[ x^a x^b = x^{a+b}, \quad \frac{x^a}{x^b} = x^{a-b}, \quad (x^a)^b = (x^b)^a = x^{ab} \]
\[ x^{-a} = \frac{1}{x^a}, \quad \left(\frac{x}{y}\right)^a = \frac{x^a}{y^a}, \quad x^0 = 1. \]

Surds
\[ x^{\frac{1}{a}} = \sqrt[a]{x}, \quad \sqrt{xy} = \sqrt{x} \sqrt{y}, \quad \frac{\sqrt{x}}{\sqrt{y}} = \frac{\sqrt{\sqrt{x}}}{\sqrt{\sqrt{y}}} \]
\[ \left(\sqrt{\sqrt{x}}\right)^a = x, \quad \sqrt{\sqrt[3]{x}} = \sqrt[3]{\sqrt{x}}, \quad \left(\sqrt[3]{x}\right)^b = x^{\frac{b}{3}}. \]

Exponential and Natural Logarithm
\[ e^x e^y = e^{x+y}, \quad (e^x)^y = (e^y)^x = e^{xy}, \quad e^0 = 1 \]
\[ \log (xy) = \log x + \log y, \quad \log \left(\frac{x}{y}\right) = \log x - \log y, \quad \log x^y = y \log x \]
\[ \log e^x = x, \quad e^{\log x} = x, \quad e^{a \log x} = x^a. \]

Quadratic Equation
For constants \(a, b \) and \(c\), the roots of a quadratic equation \(ax^2 + bx + c = 0\) are
\[ x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \]

Binomial Formula
\[ \binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \]
\[ (x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^{n-k} y^k. \]
Series

**Arithmetic:** For initial term $a$ and common difference $d$, the $n$-th term is

$$T_n = a + (n - 1)d$$

and the sum of $n$ terms is

$$S_n = \frac{1}{2}n[2a + (n - 1)d].$$

**Geometric:** For initial term $a$ and common ratio $r$, the $n$-th term is

$$T_n = ar^{n-1}$$

the sum of $n$ terms is

$$S_n = \frac{a(1 - r^n)}{1 - r},$$

and the sum of infinite terms is

$$\lim_{n \to \infty} S_n = \frac{a}{1 - r}, \quad |r| < 1.$$

Summation

For $n \in \mathbb{Z}^+$,

$$\sum_{k=1}^{n} k = \frac{1}{2}n(n + 1), \quad \sum_{k=1}^{n} k^2 = \frac{1}{6}n(n + 1)(2n + 1), \quad \sum_{k=1}^{n} k^3 = \left[\frac{1}{2}n(n + 1)\right]^2.$$

Let $a_1, a_2, \ldots$ be a sequence of numbers:

- If $\sum_{n=1}^{\infty} a_n < \infty \Rightarrow \lim_{n \to \infty} a_n = 0$.
- If $\lim_{n \to \infty} a_n \neq 0 \Rightarrow \sum_{n=1}^{\infty} a_n = \infty$.

Trigonometric Functions

$$\sin(-x) = -\sin x, \quad \cos(-x) = \cos x, \quad \tan x = \frac{\sin x}{\cos x},$$

$$\csc x = \frac{1}{\sin x}, \quad \sec x = \frac{1}{\cos x}, \quad \cot x = \frac{1}{\tan x},$$

$$\cos^2 x + \sin^2 x = 1, \quad \tan^2 x + 1 = \sec^2 x, \quad \cot^2 x + 1 = \csc^2 x,$$

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y,$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y,$$

$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}.$$
Appendix A

Hyperbolic Functions

\[
\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2}, \quad \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}
\]

\[
\begin{align*}
\text{csch} x &= \frac{1}{\sinh x}, & \text{sech} x &= \frac{1}{\cosh x}, & \coth x &= \frac{1}{\tanh x}, \\
\sinh(-x) &= -\sinh x, & \cosh(-x) &= \cosh x, & \tanh(-x) &= -\tanh x
\end{align*}
\]

\[
\begin{align*}
\cosh^2 x - \sinh^2 x &= 1, & \coth^2 x - 1 &= \text{csch}^2 x, & 1 - \tanh^2 x &= \text{sech}^2 x
\end{align*}
\]

\[
\begin{align*}
\sinh(x \pm y) &= \sinh x \cosh y \pm \cosh x \sinh y, \\
\cosh(x \pm y) &= \cosh x \cosh y \pm \sinh x \sinh y, \\
\tanh(x \pm y) &= \frac{\tanh x \pm \tanh y}{1 \pm \tanh x \tanh y}.
\end{align*}
\]

Complex Numbers

Let \( w = u + iv \) and \( z = x + iy \) where \( u, v, x, y \in \mathbb{R}, \) \( i = \sqrt{-1} \) and \( i^2 = -1 \) then

\[
\begin{align*}
\frac{w}{z} &= \left( \frac{ux + vy}{x^2 + y^2} \right) + \left( \frac{vx - uy}{x^2 + y^2} \right)i, \\
\frac{\bar{z}}{z} &= \frac{\overline{w}}{\overline{z}}.
\end{align*}
\]

De Moivre’s Formula: Let \( z = x + iy \) where \( x, y \in \mathbb{R} \) and we can write

\[
z = r(\cos \theta + i \sin \theta), \quad r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left( \frac{y}{x} \right).
\]

For \( n \in \mathbb{Z} \)

\[
[r(\cos \theta + i \sin \theta)]^n = r^n[\cos(n\theta) + i \sin(n\theta)].
\]

Euler’s Formula: For \( \theta \in \mathbb{R} \)

\[
e^{i\theta} = \cos \theta + i \sin \theta.
\]

Derivatives

If \( f(x) \) and \( g(x) \) are differentiable functions of \( x \) and \( a \) and \( b \) are constants

**Sum Rule:**

\[
\frac{d}{dx} (af(x) + bg(x)) = af'(x) + bg'(x)
\]

**Product/Chain Rule:**

\[
\frac{d}{dx} (f(x)g(x)) = f'(x)g(x) + f(x)g'(x)
\]
Quotient Rule: \[
\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \quad g(x) \neq 0
\]

where \( \frac{d}{dx}f(x) = f'(x) \) and \( \frac{d}{dx}g(x) = g'(x) \).

If \( f(z) \) is a differentiable function of \( z \) and \( z = z(x) \) is a differentiable function of \( x \), then
\[
\frac{d}{dx}f(z(x)) = f'(z(x))z'(x).
\]

If \( x = x(s), \ y = y(s) \) and \( F(u, v) = f(x(u, v), y(u, v)) \), then
\[
\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial u}, \quad \frac{\partial F}{\partial v} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial v}.
\]

Standard Differentiations

If \( f(x) \) and \( g(x) \) are differentiable functions of \( x \) and \( a \) and \( b \) are constants
\[
\frac{d}{dx}a = 0, \quad \frac{d}{dx}[f(x)]^n = n[f(x)]^{n-1}f'(x)
\]
\[
\frac{d}{dx}e^{f(x)} = f'(x)e^{f(x)}, \quad \frac{d}{dx}\log(f(x)) = \frac{f'(x)}{f(x)}, \quad \frac{d}{dx}a^{f(x)} = f'(x)a^{f(x)} \log a
\]
\[
\frac{d}{dx}\sin(ax) = a \cos x, \quad \frac{d}{dx}\cos(ax) = -a \sin(ax), \quad \frac{d}{dx}\tan(ax) = a \sec^2 x
\]
\[
\frac{d}{dx}\sinh(ax) = a \cosh(ax), \quad \frac{d}{dx}\cosh(ax) = a \sinh(ax), \quad \frac{d}{dx}\tanh(ax) = a \sech^2 (ax)
\]

where \( f'(x) = \frac{d}{dx}f(x) \).

Taylor Series

If \( f(x) \) is an analytic function of \( x \), then for small \( h \)
\[
f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{1}{2!}f''(x_0)h^2 + \frac{1}{3!}f'''(x_0)h^3 + \ldots
\]
Appendix A

If $f(x, y)$ is an analytic function of $x$ and $y$, then for small $\Delta x, \Delta y$

$$f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x} \Delta x + \frac{\partial f(x_0, y_0)}{\partial y} \Delta y$$
$$+ \frac{1}{2!} \left[ \frac{\partial^2 f(x_0, y_0)}{\partial x^2} (\Delta x)^2 + 2 \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} (\Delta x)(\Delta y) + \frac{\partial^2 f(x_0, y_0)}{\partial y^2} (\Delta y)^2 \right]$$
$$+ \frac{1}{3!} \left[ \frac{\partial^3 f(x_0, y_0)}{\partial x^3} (\Delta x)^3 + 3 \frac{\partial^3 f(x_0, y_0)}{\partial x^2 \partial y} (\Delta x)^2 (\Delta y) + 3 \frac{\partial^3 f(x_0, y_0)}{\partial x \partial y^2} (\Delta x)(\Delta y)^2 + \frac{\partial^3 f(x_0, y_0)}{\partial y^3} (\Delta y)^3 \right] + \ldots$$

Maclaurin Series

Taylor series expansion of a function about $x_0 = 0$

$$\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \ldots, \quad |x| < 1$$
$$\frac{1}{1 - x} = 1 + x + x^2 + x^3 + \ldots, \quad |x| < 1$$
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \quad \text{for all } x$$
$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots, \quad \text{for all } x$$
$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, \quad x \in (-1, 1]$$
$$\log(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} + \frac{x^4}{4} + \ldots, \quad |x| < 1$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots, \quad \text{for all } x$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots, \quad \text{for all } x$$
$$\tan x = x + \frac{x^3}{3} + \frac{2}{15} x^5 + \frac{17}{315} x^7 + \ldots, \quad |x| < \frac{\pi}{2}$$
$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \ldots, \quad \text{for all } x$$
$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \ldots, \quad \text{for all } x$$
$$\tanh x = x - \frac{x^3}{3} + \frac{2}{15} x^5 - \frac{17}{315} x^7 + \ldots, \quad |x| < \frac{\pi}{2}.$$

Landau Symbols and Asymptotics

Let $f(x)$ and $g(x)$ be two functions defined on some subsets of real numbers, then as $x \to x_0$
Appendix A

- \( f(x) = O(g(x)) \) if there exists a constant \( K > 0 \) and \( \delta > 0 \) such that \( |f(x)| \leq K|g(x)| \) for \( |x - x_0| < \delta \).
- \( f(x) = o(g(x)) \) if \( \lim_{x \to x_0} \frac{f(x)}{g(x)} = 0 \).
- \( f(x) \sim g(x) \) if \( \lim_{x \to x_0} \frac{f(x)}{g(x)} = 1 \).

**L’Hospital Rule**

Let \( f \) and \( g \) be differentiable on \( a \in \mathbb{R} \) such that \( g'(x) \neq 0 \) in an interval around \( a \), except possibly at \( a \) itself. Suppose that

\[
\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0
\]

or

\[
\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = \pm \infty
\]

then

\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.
\]

**Indefinite Integrals**

If \( F(x) \) is a differentiable function and \( f(x) \) is its derivative, then

\[
\int f(x) \, dx = F(x) + c
\]

where \( F'(x) = \frac{d}{dx} F(x) = f(x) \) and \( c \) is an arbitrary constant.

If \( f(x) \) is a continuous function then

\[
\frac{d}{dx} \int f(x) \, dx = f(x).
\]

**Standard Indefinite Integrals**

If \( f(x) \) is a differentiable function of \( x \) and \( a \) and \( b \) are constants

\[
\int a \, dx = ax + c, \quad \int (ax + b)^n \, dx = \frac{(ax + b)^{n+1}}{a(n + 1)} + c, \quad n \neq -1
\]

\[
\int \frac{f'(x)}{f(x)} \, dx = \log|f(x)| + c, \quad \int e^{f(x)} \, dx = \frac{1}{f'(x)} e^{f(x)} + c
\]

\[
\int \log(ax) \, dx = x \log(ax) - ax + c, \quad \int a^x \, dx = \frac{a^x}{\log a} + c
\]

\[
\int \sin(ax) \, dx = -\frac{1}{a} \cos(ax) + c, \quad \int \cos(ax) \, dx = \frac{1}{a} \sin(ax) + c
\]
Appendix A

\[ \int \sinh(ax) \, dx = \frac{1}{a} \cosh(ax) + c, \quad \int \cosh(ax) \, dx = \frac{1}{a} \sinh(ax) + c \]

where \( c \) is an arbitrary constant.

**Definite Integrals**

If \( F(x) \) is a differentiable function and \( f(x) \) is its derivative and is continuous on a close interval \([a, b]\), then

\[ \int_a^b f(x) \, dx = F(b) - F(a) \]

where \( F'(x) = \frac{d}{dx} F(x) = f(x) \).

If \( f(x) \) and \( g(x) \) are integrable functions then

\[ \int_a^b f(x) \, dx = 0, \quad \int_a^b f(x) \, dx = -\int_b^a f(x) \, dx \]

\[ \int_a^b [af(x) + \beta g(x)] \, dx = a \int_a^b f(x) \, dx + \beta \int_a^b g(x) \, dx, \quad \alpha, \beta \text{ are constants} \]

\[ \int_a^b f(x) \, dx = \int_c^a f(x) \, dx + \int_c^b f(x) \, dx, \quad c \in [a, b]. \]

**Derivatives of Definite Integrals**

If \( f(t) \) is a continuous function of \( t \) and \( a(x) \) and \( b(x) \) are continuous functions of \( x \)

\[ \frac{d}{dx} \int_{a(x)}^{b(x)} f(t) \, dt = f(b(x)) \frac{d}{dx} b(x) - f(a(x)) \frac{d}{dx} a(x) \]

\[ \frac{d}{dx} \int_{a(x)}^{b(x)} dt = \frac{d}{dx} b(x) - \frac{d}{dx} a(x). \]

If \( g(x, t) \) is a differentiable function of two variables then

\[ \frac{d}{dx} \int_{a(x)}^{b(x)} g(x, t) \, dt = g(x, b(x)) \frac{d}{dx} b(x) - g(x, a(x)) \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{\partial g(x, t)}{\partial x} \, dt. \]

**Integration by Parts**

For definite integrals

\[ \int_a^b u(x)v'(x) \, dx = u(x)v(x) \bigg|_a^b - \int_a^b v(x)u'(x) \, dx. \]

where \( u'(x) = \frac{d}{dx} u(x) \) and \( v'(x) = \frac{d}{dx} v(x) \).
Integration by Substitution

If \( f(x) \) is a continuous function of \( x \) and \( g' \) is continuous on the closed interval \([a, b]\) then
\[
\int_{g(a)}^{g(b)} f(x) \, dx = \int_{a}^{b} f(g(u))g'(u) \, du.
\]

Gamma Function

The gamma function is defined as
\[
\Gamma(z) = \int_{0}^{\infty} t^{z-1} e^{-t} \, dt
\]
such that
\[
\Gamma(z + 1) = z\Gamma(z), \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma(n) = (n-1)! \text{ for } n \in \mathbb{N}.
\]

Beta Function

The beta function is defined as
\[
B(x, y) = \int_{0}^{1} t^{x-1}(1-t)^{y-1} \, dt \quad \text{for } x > 0, y > 0
\]
such that
\[
B(x, y) = B(y, x), \quad B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.
\]
In addition
\[
\int_{0}^{u} t^{x-1}(u-t)^{y-1} \, dt = u^{x+y-1}B(x, y).
\]

Convex Function

A set \( \Omega \) in a vector space over \( \mathbb{R} \) is called a convex set if for \( x, y \in \Omega \), \( x \neq y \) and for any \( \lambda \in (0, 1) \),
\[
\lambda x + (1-\lambda)y \in \Omega.
\]
Let \( \Omega \) be a convex set in a vector space over \( \mathbb{R} \). A function \( f : \Omega \rightarrow \mathbb{R} \) is called a convex function if for \( x, y \in \Omega \), \( x \neq y \) and for any \( \lambda \in (0, 1) \),
\[
f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).
\]
If the inequality is strict then \( f \) is strictly convex.
If \( f \) is convex and differentiable on \( \mathbb{R} \), then
\[
f(x) \geq f(y) + f'(y)(x - y).
\]
If \( f \) is a twice continuously differentiable function on \( \mathbb{R} \), then \( f \) is convex if and only if \( f'' \geq 0 \).
If \( f'' > 0 \) then \( f \) is strictly convex.
Appendix A

$f$ is a (strictly) concave function if $-f$ is a (strictly) convex function.

**Dirac Delta Function**

The Dirac delta function is defined as

$$
\delta(x) = \begin{cases} 
0 & x \neq 0 \\
\infty & x = 0 
\end{cases}
$$

and for a continuous function $f(x)$ and a constant $a$, we have

$$
\int_{-\infty}^{\infty} f(x) \delta(x) \, dx = f(0), \quad \int_{-\infty}^{\infty} f(x) \delta(x - a) \, dx = f(a).
$$

**Heaviside Step Function**

The Heaviside step function, $H(x)$ is defined as the integral of the Dirac delta function given as

$$
H(x) = \int_{-\infty}^{x} \delta(s) \, ds = \begin{cases} 
0 & x < 0 \\
1 & x \geq 0 
\end{cases}
$$

**Fubini’s Theorem**

Suppose $f(x, y)$ is $A \times B$ measurable and if $\int_{A \times B} |f(x, y)| \, d(x, y) < \infty$ then

$$
\int_{A \times B} f(x, y) \, d(x, y) = \int_{A} \left( \int_{B} f(x, y) \, dy \right) \, dx = \int_{B} \left( \int_{A} f(x, y) \, dx \right) \, dy.
$$
Appendix B

Probability Theory Formulae

Probability Concepts

Let \( A \) and \( B \) be events of the sample space \( \Omega \) with probabilities \( \mathbb{P}(A) \in [0, 1] \) and \( \mathbb{P}(B) \in [0, 1] \), then

Complement:

\[
\mathbb{P}(A^c) = 1 - \mathbb{P}(A).
\]

Conditional:

\[
\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.
\]

Independence: The events \( A \) and \( B \) are independent if and only if

\[
\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).
\]

Mutually Exclusive: The events \( A \) and \( B \) are mutually exclusive if and only if

\[
\mathbb{P}(A \cap B) = 0.
\]

Union:

\[
\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).
\]

Intersection:

\[
\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A).
\]

Partition:

\[
\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c).
\]

Bayes’ Rule

Let \( A \) and \( B \) be events of the sample space \( \Omega \) with probabilities \( \mathbb{P}(A) \in [0, 1] \) and \( \mathbb{P}(B) \in [0, 1] \), then

\[
\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.
\]
Indicator Function

The indicator function $1_A$ of an event $A$ of a sample space $\Omega$ is a function $1_A : \Omega \mapsto \mathbb{R}$ defined as

$$1_A(\omega) = \begin{cases} 
1 & \text{if } \omega \in A \\
0 & \text{if } \omega \in A^c.
\end{cases}$$

Properties: For events $A$ and $B$ of the sample space $\Omega$

$$1_{A^c} = 1 - 1_A, \quad 1_{A \cap B} = 1_A 1_B, \quad 1_{A \cup B} = 1_A + 1_B - 1_A 1_B$$

$$\mathbb{E}(1_A) = \mathbb{P}(A), \quad \text{Var}(1_A) = \mathbb{P}(A)\mathbb{P}(A^c), \quad \text{Cov}(1_A, 1_B) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B).$$

Discrete Random Variables

Univariate Case

Let $X$ be a discrete random variable whose possible values are $x = x_1, x_2, \ldots$, and let $\mathbb{P}(X = x)$ be the probability mass function.

Total Probability of All Possible Values:

$$\sum_{k=1}^{\infty} \mathbb{P}(X = x_k) = 1.$$

Cumulative Distribution Function:

$$\mathbb{P}(X \leq x_n) = \sum_{k=1}^{n} \mathbb{P}(X = x_k).$$

Expectation:

$$\mathbb{E}(X) = \mu = \sum_{k=1}^{\infty} x_k \mathbb{P}(X = x_k).$$

Variance:

$$\text{Var}(X) = \sigma^2$$

$$= \mathbb{E}[(X - \mu)^2]$$

$$= \sum_{k=1}^{\infty} (x_k - \mu)^2 \mathbb{P}(X = x_k)$$

$$= \sum_{k=1}^{\infty} x_k^2 \mathbb{P}(X = x_k) - \mu^2$$

$$= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.$$

Moment Generating Function:

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{k=1}^{\infty} e^{tx_k} \mathbb{P}(X = x_k), \quad t \in \mathbb{R}.$$
Appendix B

Characteristic Function:

\[ \varphi_X(t) = \mathbb{E}(e^{itX}) = \sum_{k=1}^{\infty} e^{itx_k} \mathbb{P}(X = x_k), \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}. \]

Bivariate Case

Let \( X \) and \( Y \) be discrete random variables whose possible values are \( x = x_1, x_2, \ldots \) and \( y = y_1, y_2, \ldots \), respectively, and let \( \mathbb{P}(X = x, Y = y) \) be the joint probability mass function.

Total Probability of All Possible Values:

\[
\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \mathbb{P}(X = x_j, Y = y_k) = 1.
\]

Joint Cumulative Distribution Function:

\[
\mathbb{P}(X \leq x_n, Y \leq y_m) = \sum_{j=1}^{n} \sum_{k=1}^{m} \mathbb{P}(X = x_j, Y = y_k).
\]

Marginal Probability Mass Function:

\[
\mathbb{P}(X = x) = \sum_{k=1}^{\infty} \mathbb{P}(X = x, Y = y_k), \quad \mathbb{P}(Y = y) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_j, Y = y).
\]

Conditional Probability Mass Function:

\[
\mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}, \quad \mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)}.
\]

Conditional Expectation:

\[
\mathbb{E}(X|Y) = \mu_{x|y} = \sum_{j=1}^{n} x_j \mathbb{P}(X = x_j|Y = y), \quad \mathbb{E}(Y|X) = \mu_{y|x} = \sum_{k=1}^{m} y_k \mathbb{P}(Y = y_k|X = x).
\]

Conditional Variance:

\[
\text{Var}(X|Y) = \sigma_{x|y}^2 = \mathbb{E}[(X - \mu_{x|y})^2|Y]
\]
\[
= \sum_{j=1}^{\infty} (x_j - \mu_{x|y})^2 \mathbb{P}(X = x_j|Y = y)
\]
\[
= \sum_{j=1}^{\infty} x_j^2 \mathbb{P}(X = x_j|Y = y) - \mu_{x|y}^2
\]
\[ \text{Var}(Y|X) = \sigma_{Y|x}^2 \]
\[ = \mathbb{E}[(Y - \mu_{Y|x})^2|X] \]
\[ = \sum_{k=1}^{\infty} (y_k - \mu_{Y|x})^2 \mathbb{P}(Y = y_k|X = x) \]
\[ = \sum_{k=1}^{\infty} y_k^2 \mathbb{P}(Y = y_k|X = x) - \mu_{Y|x}^2. \]

\text{Covariance:}\ For \ \mathbb{E}(X) = \mu_x \text{ and } \mathbb{E}(Y) = \mu_y
\[ \text{Cov}(X, Y) = \mathbb{E}[(X - \mu_x)(Y - \mu_y)] \]
\[ = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (x_j - \mu_x)(y_k - \mu_y) \mathbb{P}(X = x_j, Y = y_k) \]
\[ = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} x_j y_k \mathbb{P}(X = x_j, Y = y_k) - \mu_x \mu_y \]
\[ = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \]

\text{Joint Moment Generating Function:}\ For \ s, t \in \mathbb{R}
\[ M_{XY}(s, t) = \mathbb{E}(e^{sX+tY}) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} e^{s x_j + t y_k} \mathbb{P}(X = x_j, Y = y_k). \]

\text{Joint Characteristic Function:}\ For \ i = \sqrt{-1} \text{ and } s, t \in \mathbb{R}
\[ \varphi_{XY}(s, t) = \mathbb{E}(e^{i s X + i t Y}) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} e^{i x_j s + i y_k t} \mathbb{P}(X = x_j, Y = y_k). \]

\text{Independence: } X \text{ and } Y \text{ are independent if and only if}
\begin{itemize}
  \item \( \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y). \)
  \item \( M_{XY}(s, t) = \mathbb{E}(e^{sX+tY}) = \mathbb{E}(e^{sX})\mathbb{E}(e^{tY}) = M_X(s)M_Y(t). \)
  \item \( \varphi_{XY}(s, t) = \mathbb{E}(e^{i s X + i t Y}) = \mathbb{E}(e^{i s X})\mathbb{E}(e^{i t Y}) = \varphi_X(s)\varphi_Y(t). \)
\end{itemize}

\textbf{Continuous Random Variables}

\textbf{Univariate Case}

Let \( X \) be a continuous random variable whose values \( x \in \mathbb{R} \) and let \( f_X(x) \) be the probability density function.

\textit{Total Probability of All Possible Values:}
\[ \int_{-\infty}^{\infty} f_X(x) \, dx = 1. \]
Appendix B

Evaluating Probability:
\[ \mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) \, dx. \]

Cumulative Distribution Function:
\[ F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(x) \, dx. \]

Probability Density Function:
\[ f_X(x) = \frac{d}{dx} F_X(x). \]

Expectation:
\[ \mathbb{E}(X) = \mu = \int_{-\infty}^{\infty} x f_X(x) \, dx. \]

Variance:
\[ \text{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) \, dx = \int_{-\infty}^{\infty} x^2 f_X(x) \, dx - \mu^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2. \]

Moment Generating Function:
\[ M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx, \quad t \in \mathbb{R}. \]

Characteristic Function:
\[ \varphi_X(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f_X(x) \, dx, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}. \]

Probability Density Function of a Dependent Variable:

Let the random variable \( Y = g(X) \). If \( g \) is monotonic then the probability density function of \( Y \) is
\[ f_Y(y) = f_X \left( g^{-1}(y) \right) \left| \frac{d}{dy} g^{-1}(y) \right|^{-1} \]
where \( g^{-1} \) denotes the inverse function.

Bivariate Case

Let \( X \) and \( Y \) be two continuous random variables whose values \( x \in \mathbb{R} \) and \( y \in \mathbb{R} \), and let \( f_{XY}(x,y) \) be the joint probability density function.

Total Probability of All Possible Values:
\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y) \, dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y) \, dydx = 1. \]
Joint Cumulative Distribution Function:

$$F_{XY}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x, y) \, dy \, dx = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x, y) \, dx \, dy.$$  

Evaluating Joint Probability:

$$\mathbb{P}(x_a \leq X \leq x_b, y_a \leq Y \leq y_b) = \int_{x_a}^{x_b} \int_{y_a}^{y_b} f_{XY}(x, y) \, dy \, dx = \int_{y_a}^{y_b} \int_{x_a}^{x_b} f_{XY}(x, y) \, dx \, dy = F_{XY}(x_b, y_b) - F_{XY}(x_b, y_a) - F_{XY}(x_a, y_b) + F_{XY}(x_a, y_a).$$  

Joint Probability Density Function:

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) = \frac{\partial^2}{\partial y \partial x} F_{XY}(x, y).$$  

Marginal Probability Density Function:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx.$$  

Conditional Probability Density Function:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}, \quad f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}.$$  

Conditional Expectation:

$$\mathbb{E}(X|Y) = \mu_{x|y} = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) \, dx, \quad \mathbb{E}(Y|X) = \mu_{y|x} = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) \, dy.$$  

Conditional Variance:

$$\text{Var}(X|Y) = \sigma^2_{x|y}$$

$$= \mathbb{E}[(X - \mu_{x|y})^2|Y]$$

$$= \int_{-\infty}^{\infty} (x - \mu_{x|y})^2 f_{X|Y}(x|y) \, dx$$

$$= \int_{-\infty}^{\infty} x^2 f_{X|Y}(x|y) \, dx - \mu^2_{x|y}$$

$$\text{Var}(Y|X) = \sigma^2_{y|x}$$

$$= \mathbb{E}[(Y - \mu_{y|x})^2|X]$$

$$= \int_{-\infty}^{\infty} (y - \mu_{y|x})^2 f_{Y|X}(y|x) \, dy.$$
then the joint probability density function of $g$ is

\[ \int_{-\infty}^{\infty} y^2 f_{Y\mid X}(y|x) \, dy - \mu_{y|x}^2. \]

Covariance: For $\mathbb{E}(X) = \mu_x$ and $\mathbb{E}(Y) = \mu_y$

\[ \text{Cov}(X, Y) = \mathbb{E}[(X - \mu_x)(Y - \mu_y)] \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y)f_{XY}(x,y) \, dx \, dy \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_{XY}(x,y) \, dx \, dy - \mu_x\mu_y \]

\[ = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \]

Joint Moment Generating Function: For $t, s \in \mathbb{R}$

\[ M_{XY}(s,t) = \mathbb{E}(e^{sx+ty}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{sx+ty}f_{XY}(x,y) \, dx \, dy. \]

Joint Characteristic Function: For $i = \sqrt{-1}$ and $t, s \in \mathbb{R}$

\[ \phi_{XY}(s,t) = \mathbb{E}(e^{ixY+iy}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ixy+iy}f_{XY}(x,y) \, dx \, dy. \]

Independence: $X$ and $Y$ are independent if and only if

- $f_{XY}(x,y) = f_X(x)f_Y(y)$.
- $M_{XY}(s,t) = \mathbb{E}(e^{sx+ty}) = \mathbb{E}(e^{sx})\mathbb{E}(e^{ty}) = M_X(s)M_Y(t)$.
- $\phi_{XY}(s,t) = \mathbb{E}(e^{ixY+iy}) = \mathbb{E}(e^{ixY})\mathbb{E}(e^{iy}) = \phi_X(s)\phi_Y(t)$.

Joint Probability Density Function of Dependent Variables: Let the random variables $U = g(X, Y), V = h(X, Y).$ If $u = g(x,y)$ and $v = h(x,y)$ can be uniquely solved for $x$ and $y$ in terms of $u$ and $v$ with solutions given by, say, $x = p(u,v)$ and $y = q(u,v)$ and the functions $g$ and $h$ have continuous partial derivatives at all points $(x,y)$ such that the determinant

\[ J(x,y) = \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \frac{\partial g}{\partial x} \frac{\partial h}{\partial y} - \frac{\partial g}{\partial y} \frac{\partial h}{\partial x} \neq 0 \]

then the joint probability density function of $U$ and $V$ is

\[ f_{UV}(u,v) = f_{XY}(x,y)|J(x,y)|^{-1} \]

where $x = p(u,v)$ and $y = q(u,v)$.

Properties of Expectation and Variance

Let $X$ and $Y$ be two random variables and for constants $a$ and $b$

\[ \mathbb{E}(aX + b) = a\mathbb{E}(X) + b, \quad \text{Var}(aX + b) = a^2\text{Var}(X) \]

\[ \mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y), \quad \text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y). \]
Properties of Moment Generating and Characteristic Functions

If a random variable $X$ has moments up to $k$-th order where $k$ is a non-negative integer, then

$$
\mathbb{E}(X^k) = \frac{d^k}{dt^k} M_X(t) \bigg|_{t=0} = i^{-k} \frac{d^k}{dt^k} \varphi_X(t) \bigg|_{t=0}
$$

where $i = \sqrt{-1}$.

If the bivariate random variables $X$ and $Y$ have moments up to $m + n = k$ where $m$, $n$ and $k$ are non-negative integers, then

$$
\mathbb{E}(X^m Y^n) = \frac{d^k}{ds^m dt^n} M_{XY}(s, t) \bigg|_{s=0, t=0} = i^{-k} \frac{d^k}{ds^m dt^n} \varphi_{XY}(s, t) \bigg|_{s=0, t=0}
$$

where $i = \sqrt{-1}$.

Correlation Coefficient

Let $X$ and $Y$ be two random variables with means $\mu_x$ and $\mu_y$ and variances $\sigma_x^2$ and $\sigma_y^2$. The correlation coefficient $\rho_{xy}$ between $X$ and $Y$ is defined as

$$
\rho_{xy} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} = \frac{\mathbb{E}[(X - \mu_x)(Y - \mu_y)]}{\sigma_x \sigma_y}.
$$

Important information:

- $\rho_{xy}$ measures only the linear dependency between $X$ and $Y$.
- $-1 \leq \rho_{xy} \leq 1$.
- If $\rho_{xy} = 0$ then $X$ and $Y$ are uncorrelated.
- If $X$ and $Y$ are independent then $\rho_{xy} = 0$. However, the converse is not true.
- If $X$ and $Y$ are jointly normally distributed then $X$ and $Y$ are independent if and only if $\rho_{XY} = 0$.

Convolution

If $X$ and $Y$ are independent discrete random variables with probability mass functions $\mathbb{P}(X = x)$ and $\mathbb{P}(Y = y)$, respectively, then the probability mass function for $Z = X + Y$ is

$$
\mathbb{P}(Z = z) = \sum_x \mathbb{P}(X = x) \mathbb{P}(Y = z - x) = \sum_y \mathbb{P}(X = z - y) \mathbb{P}(Y = y).
$$

If $X$ and $Y$ are independent continuous random variables with probability density functions $f_X(x)$ and $f_Y(y)$, respectively, then the probability density function for $Z = X + Y$ is

$$
f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) \, dx = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) \, dy.
$$
Appendix B

Discrete Distributions

**Bernoulli:** A random variable $X$ is said to follow a Bernoulli distribution, $X \sim \text{Bernoulli}(p)$, where $p \in [0, 1]$ is the probability of success and the probability mass function is given as

$$P(X = x) = p^x(1 - p)^{1-x}, \quad x = 0, 1$$

where $\mathbb{E}(X) = p$ and $\text{Var}(X) = p - p^2$. The moment generating function is

$$M_X(t) = 1 - p + pe^t, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = 1 - p + pe^{it}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$ 

**Geometric:** A random variable $X$ is said to follow a geometric distribution, $X \sim \text{Geometric}(p)$, where $p \in [0, 1]$ is the probability of success and the probability mass function is given as

$$\mathbb{P}(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, \ldots$$

where $\mathbb{E}(X) = \frac{1}{p}$ and $\text{Var}(X) = \frac{1 - p}{p^2}$. The moment generating function is

$$M_X(t) = \frac{p}{1 - (1 - p)e^t}, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = \frac{p}{1 - (1 - p)e^{it}}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$ 

**Binomial:** A random variable $X$ is said to follow a binomial distribution, $X \sim \text{Binomial}(n, p)$, $p \in [0, 1]$, where $p \in [0, 1]$ is the probability of success and $n \in \mathbb{N}_0$ is the number of trials and the probability mass function is given as

$$\mathbb{P}(X = x) = \binom{n}{x} p^x(1 - p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

where $\mathbb{E}(X) = np$ and $\text{Var}(X) = np(1 - p)$. The moment generating function is

$$M_X(t) = (1 - p + pe^t)^n, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = (1 - p + pe^{it})^n, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$
Negative Binomial: A random variable $X$ is said to follow a negative binomial distribution, $X \sim \text{NB}(r, p)$ where $p \in [0, 1]$ is the probability of success and $r$ is the number of successes accumulated and the probability mass function is given as

$$P(X = x) = \binom{x - 1}{r - 1} p^r (1 - p)^{x-r}, \quad x = r, r + 1, r + 2, \ldots$$

where $\mathbb{E}(X) = \frac{r}{p}$ and $\text{Var}(X) = \frac{r(1 - p)}{p^2}$. The moment generating function is

$$M_X(t) = \left( \frac{1 - p}{1 - pe^t} \right)^r, \quad t < -\log p$$

and the corresponding characteristic function is

$$\varphi_X(t) = \left( \frac{1 - p}{1 - pe^{it}} \right)^r, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$

Poisson: A random variable $X$ is said to follow a Poisson distribution, $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$ with probability mass function given as

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

where $\mathbb{E}(X) = \lambda$ and $\text{Var}(X) = \lambda$. The moment generating function is

$$M_X(t) = e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = e^{\lambda(e^{it} - 1)}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$

Continuous Distributions

Uniform: A random variable $X$ is said to follow a uniform distribution, $X \sim \mathcal{U}(a, b)$, $a < b$ with probability density function given as

$$f_X(x) = \frac{1}{b - a}, \quad a < x < b$$

where $\mathbb{E}(X) = \frac{a + b}{2}$ and $\text{Var}(X) = \frac{(b-a)^2}{12}$. The moment generating function is

$$M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = \frac{e^{tb} - e^{ta}}{it(b-a)}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$
Appendix B

Normal: A random variable $X$ is said to follow a normal distribution, $X \sim \mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$ with probability density function given as

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}, \quad x \in \mathbb{R}$$

where $\mathbb{E}(X) = \mu$ and $\text{Var}(X) = \sigma^2$. The moment generating function is

$$M_X(t) = e^{\mu t + \frac{1}{2} \sigma^2 t^2}, \quad t \in \mathbb{R}$$

and the corresponding characteristic function is

$$\varphi_X(t) = e^{\mu t - \frac{1}{2} \sigma^2 t^2}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$

Lognormal: A random variable $X$ is said to follow a lognormal distribution, $X \sim \log\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$ with probability density function given as

$$f_X(x) = \frac{1}{x \sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log x - \mu}{\sigma} \right)^2}, \quad x > 0$$

where $\mathbb{E}(X) = e^{\mu + \frac{1}{2} \sigma^2}$ and $\text{Var}(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$. The moment generating function is

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} e^{\mu t + \frac{1}{2} \sigma^2 t^2}, \quad t \leq 0$$

and the corresponding characteristic function is

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} e^{\mu t + \frac{1}{2} \sigma^2 t^2}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$

Exponential: A random variable $X$ is said to follow an exponential distribution, $X \sim \text{Exp}(\lambda)$, $\lambda > 0$ with probability density function

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

where $\mathbb{E}(X) = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$. The moment generating function is

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda$$

and the corresponding characteristic function is

$$\varphi_X(t) = \frac{\lambda}{\lambda - it}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.$$
Gamma: A random variable $X$ is said to follow a gamma distribution, $X \sim \text{Gamma}(\alpha, \lambda)$, $\alpha, \lambda > 0$ with probability density function given as
\[
f_X(x) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, \quad x \geq 0
\]
such that
\[
\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} \, dx
\]
where $\mathbb{E}(X) = \frac{\alpha}{\lambda}$ and $\text{Var}(X) = \frac{\alpha}{\lambda^2}$. The moment generating function is
\[
M_X(t) = \left( \frac{\lambda}{\lambda - t} \right)^\alpha, \quad t < \lambda
\]
and the corresponding characteristic function is
\[
\varphi_X(t) = \left( \frac{\lambda}{\lambda - it} \right)^\alpha, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.
\]
Chi-Square: A random variable $X$ is said to follow a chi-square distribution, $X \sim \chi^2(\nu)$, $\nu \in \mathbb{N}$ with probability density function given as
\[
f_X(x) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2} - 1} e^{-\frac{x}{2}}, \quad x \geq 0
\]
such that
\[
\Gamma\left(\frac{\nu}{2}\right) = \int_0^\infty e^{-x} x^{\frac{\nu}{2} - 1} \, dx
\]
where $\mathbb{E}(X) = \nu$ and $\text{Var}(X) = 2\nu$. The moment generating function is
\[
M_X(t) = (1 - 2t)^{-\frac{\nu}{2}}, \quad -\frac{1}{2} < t < \frac{1}{2}
\]
and the corresponding characteristic function is
\[
M_X(t) = (1 - 2it)^{-\frac{\nu}{2}}, \quad i = \sqrt{-1} \text{ and } t \in \mathbb{R}.
\]
Bivariate Normal: The random variables $X$ and $Y$ with means $\mu_x$, $\mu_y$, variances $\sigma_x^2$, $\sigma_y^2$, and correlation coefficient $\rho_{xy} \in (-1, 1)$ is said to follow a joint normal distribution, $(X, Y) \sim \mathcal{N}_2(\mu, \Sigma)$ where $\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \rho_{xy} \sigma_x \sigma_y \\ \rho_{xy} \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}$ with joint probability density function given as
\[
f_{XY}(x, y) = \frac{1}{2\pi\sigma_x \sigma_y \sqrt{1 - \rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right]}, \quad x, y \in \mathbb{R}.
Appendix B

The moment generating function is
\[ M_{XY}(s, t) = e^{\mu_X s + \mu_Y t + \frac{1}{2} (\sigma_X^2 s^2 + 2 \rho_{XY} \sigma_X \sigma_Y s t + \sigma_Y^2 t^2)}, \quad s, t \in \mathbb{R} \]
and the corresponding characteristic function is
\[ \phi_{XY}(s, t) = e^{i \mu_X s + i \mu_Y t - \frac{1}{2} (\sigma_X^2 s^2 + 2 \rho_{XY} \sigma_X \sigma_Y s t + \sigma_Y^2 t^2)}, \quad i = \sqrt{-1} \text{ and } s, t \in \mathbb{R}. \]

**Multivariate Normal:** The random vector \( X = (X_1, X_2, \ldots, X_n) \) is said to follow a multivariate normal distribution, \( X \sim \mathcal{N}_n(\mu, \Sigma) \) where
\[
\mu = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \ldots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_1, X_2) & \text{Var}(X_2) & \ldots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_1, X_n) & \text{Cov}(X_2, X_n) & \ldots & \text{Var}(X_n) \end{bmatrix}
\]
with probability density function given as
\[ f_X(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} x^T \Sigma^{-1} x}. \]

The moment generating function is
\[ M_X(t_1, t_2, \ldots, t_n) = e^{\mu^T t + \frac{1}{2} t^T \Sigma t}, \quad t = (t_1, t_2, \ldots, t_n)^T \]
and the corresponding characteristic function is
\[ \phi_X(t_1, t_2, \ldots, t_n) = e^{i \mu^T t - \frac{1}{2} t^T \Sigma t}, \quad t = (t_1, t_2, \ldots, t_n)^T. \]

**Integrable and Square Integrable Random Variables**

Let \( X \) be a real-valued random variable

- If \( \mathbb{E}(|X|) < \infty \) then \( X \) is an integrable random variable.
- If \( \mathbb{E}(X^2) < \infty \) then \( X \) is a square integrable random variable.

**Convergence of Random Variables**

Let \( X, X_1, X_2, \ldots, X_n \) be a sequence of random variables. Then

(a) \( X_n \xrightarrow{a.s.} X \) converges almost surely if
\[ \mathbb{P} \left( \lim_{n \to \infty} X_n = X \right) = 1. \]
(b) $X_n \overset{r}{\to} X$ converges in the $r$-th mean, $r \geq 1$, if $\mathbb{E}(|X_n^r|) < \infty$ and
\[
\lim_{n \to \infty} \mathbb{E}(|X_n - X|^r) = 0.
\]
(c) $X_n \overset{P}{\to} X$ converges in probability, if for all $\varepsilon > 0$,
\[
\lim_{n \to \infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.
\]
(d) $X_n \overset{D}{\to} X$ converges in distribution, if for all $x \in \mathbb{R}$,
\[
\lim_{n \to \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x).
\]

**Relationship Between Modes of Convergence**

For any $r \geq 1$
\[
\begin{align*}
\left\{ X_n \overset{a.s.}{\to} X \right\} &\Rightarrow \left\{ X_n \overset{P}{\to} X \right\} \Rightarrow \left\{ X_n \overset{D}{\to} X \right\}.
\end{align*}
\]
If $r > s \geq 1$ then
\[
\left\{ X_n \overset{r}{\to} X \right\} \Rightarrow \left\{ X_n \overset{s}{\to} X \right\}.
\]

**Dominated Convergence Theorem**

If $X_n \overset{a.s.}{\to} X$ and for any $n \in \mathbb{N}$ we have $|X_n| < Y$ for some $Y$ such that $\mathbb{E}(|Y|) < \infty$, then $\mathbb{E}(|X_n|) < \infty$ and
\[
\lim_{n \to \infty} \mathbb{E}(X_n) = \mathbb{E}(X).
\]

**Monotone Convergence Theorem**

If $0 \leq X_n \leq X_{n+1}$ and $X_n \overset{a.s.}{\to} X$ for any $n \in \mathbb{N}$ then
\[
\lim_{n \to \infty} \mathbb{E}(X_n) = \mathbb{E}(X).
\]

**The Weak Law of Large Numbers**

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with common mean $\mu \in \mathbb{R}$. Then for any $\varepsilon > 0$,
\[
\mathbb{P} \left( \lim_{n \to \infty} \left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right| \geq \varepsilon \right) = 0.
\]
Appendix B

The Strong Law of Large Numbers

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with common mean $\mu \in \mathbb{R}$. Then

$$\mathbb{P} \left( \lim_{n \to \infty} \frac{X_1 + X_2 + \ldots + X_n}{n} = \mu \right) = 1.$$  

The Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with common mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ and denote the sample mean as

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

where $\mathbb{E}(\overline{X}) = \mu$ and $\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$. By defining

$$Z_n = \frac{\overline{X} - \mathbb{E}(\overline{X})}{\sqrt{\text{Var}(\overline{X})}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

then for $n \to \infty$,

$$\lim_{n \to \infty} Z_n = \lim_{n \to \infty} \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} \mathcal{N}(0, 1).$$

That is, $Z_n$ follows a standard normal distribution asymptotically.
Appendix C  Differential Equations Formulae

Separable Equations

The form
\[ \frac{dy}{dx} = f(x)g(y) \]
has a solution
\[ \int \frac{1}{g(y)} \, dy = \int f(x) \, dx. \]

If \( g(y) \) is a linear equation and if \( y_1 \) and \( y_2 \) are two solutions, then \( y_3 = ay_1 + by_2 \) is also a solution for constant \( a \) and \( b \).

First-Order Ordinary Differential Equations

**General Linear Equation:** The general form of a first-order ordinary differential equation
\[ \frac{dy}{dx} + f(x)y = g(x) \]
has a solution
\[ y = I(x)^{-1} \int I(u)g(u) \, du + C \]
where \( I(x) = e^{\int f(x) \, dx} \) is the integrating factor and \( C \) is a constant.

**Bernoulli Differential Equation:** For \( n \neq 1 \), the Bernoulli differential equation has the form
\[ \frac{dy}{dx} + P(x)y = Q(x)y^n \]
which, by setting \( w = \frac{1}{y^{n-1}} \), can be transformed to a general linear ordinary differential equation of the form
\[ \frac{dw}{dx} + (1 - n)P(x)w = (1 - n)Q(x) \]
with a particular solution
\[ w = (1 - n)I(x)^{-1} \int I(u)Q(u) \, du \]
where \( I(x) = e^{(1-n)\int p(x)dx} \) is the integrating factor. The solution to the Bernoulli differential equation becomes

\[
y = \left\{ (1-n)I(x)^{-1} \int I(u)Q(u) \, du \right\}^{-\frac{1}{n-1}} + C
\]

where \( C \) is a constant value.

**Second-Order Ordinary Differential Equations**

**General Linear Equation:** For a homogeneous equation,

\[
a \frac{d^2y}{dx^2} + b \frac{dy}{dx} + cy = 0.
\]

By setting \( y = e^{ax} \) the differential equation has a general solution based on the characteristic equation

\[
a u^2 + bu + c = 0
\]

such that \( m_1 \) and \( m_2 \) are the roots of the quadratic equation, and if

- \( m_1, m_2 \in \mathbb{R}, m_1 \neq m_2 \) then \( y = Ae^{m_1x} + Be^{m_2x} \)
- \( m_1, m_2 \in \mathbb{R}, m_1 = m_2 = m \) then \( y = e^{mx}(A + Bx) \)
- \( m_1, m_2 \in \mathbb{C}, m_1 = \alpha + i\beta, m_2 = \alpha - i\beta \) then \( y = e^{\alpha x}[A \cos(\beta x) + B \sin(\beta x)] \)

where \( A, B \) are constants.

**Cauchy–Euler Equation:** For a homogeneous equation,

\[
a x^2 \frac{d^2y}{dx^2} + bx \frac{dy}{dx} + cy = 0.
\]

By setting \( y = x^n \) the Cauchy–Euler equation has a general solution based on the characteristic equation

\[
a u^2 + (b - a)u + c = 0
\]

such that \( m_1 \) and \( m_2 \) are the roots of the quadratic equation, and if

- \( m_1, m_2 \in \mathbb{R}, m_1 \neq m_2 \) then \( y = Ax^{m_1} + Bx^{m_2} \)
- \( m_1, m_2 \in \mathbb{R}, m_1 = m_2 = m \) then \( y = x^n(A + B \log x) \)
- \( m_1, m_2 \in \mathbb{C}, m_1 = \alpha + i\beta, m_2 = \alpha - i\beta \) then \( y = x^n[A \cos(\beta \log x) + B \sin(\beta \log x)] \)

where \( A, B \) are constants.

**Variation of Parameters:** For a general non-homogeneous second-order differential equation,

\[
a(x) \frac{d^2y}{dx^2} + b(x) \frac{dy}{dx} + c(x) = f(x)
\]
Appendix C

has the solution

\[ y = y_c + y_p \]

where \( y_c \), the complementary function, satisfies the homogeneous equation

\[
a(x) \frac{d^2 y_c}{dx^2} + b(x) \frac{d y_c}{dx} + c(x) = 0
\]

and \( y_p \), the particular integral, satisfies

\[
a(x) \frac{d^2 y_p}{dx^2} + b(x) \frac{d y_p}{dx} + c(x) = f(x).
\]

Let \( y_c = C_1 y_c^{(1)}(x) + C_2 y_c^{(2)}(x) \) where \( C_1 \) and \( C_2 \) are constants, then the particular solution to the non-homogeneous second-order differential equation is

\[
y_p = -y_c^{(1)}(x) \int \frac{y_c^{(2)}(x)f(x)}{a(x) W(y_c^{(1)}(x), y_c^{(2)}(x))} \, dx + y_c^{(2)}(x) \int \frac{y_c^{(1)}(x)f(x)}{a(x) W(y_c^{(1)}(x), y_c^{(2)}(x))} \, dx
\]

where \( W(y_c^{(1)}(x), y_c^{(2)}(x)) \) is the Wronskian defined as

\[
W(y_c^{(1)}(x), y_c^{(2)}(x)) = \begin{vmatrix}
  y_c^{(1)}(x) & y_c^{(2)}(x) \\
  \frac{d}{dx} y_c^{(1)}(x) & \frac{d}{dx} y_c^{(2)}(x)
\end{vmatrix} = y_c^{(1)}(x) \frac{d}{dx} y_c^{(2)}(x) - y_c^{(2)}(x) \frac{d}{dx} y_c^{(1)}(x) \neq 0.
\]

**Homogeneous Heat Equations**

*Initial Value Problem on an Infinite Interval:* The diffusion equation of the form

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad \alpha > 0, \quad -\infty < x < \infty, \quad t > 0
\]

with initial condition \( u(x, 0) = f(x) \) has a solution

\[
u(x, t) = \frac{1}{2 \sqrt{\pi \alpha t}} \int_{-\infty}^{\infty} f(z) e^{-\frac{(x-z)^2}{4\alpha t}} \, dz.
\]

*Initial Value Problem on a Semi-Infinite Interval:* The diffusion equation of the form

\[
\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad \alpha > 0, \quad 0 \leq x < \infty, \quad t > 0
\]
with

- initial condition \( u(x, 0) = f(x) \) and boundary condition \( u(0, t) = 0 \) has a solution

\[
    u(x, t) = \frac{1}{2\sqrt{\pi \alpha t}} \int_{0}^{\infty} f(z) \left[ e^{-\frac{(x-z)^2}{4\alpha t}} - e^{-\frac{(x+z)^2}{4\alpha t}} \right] \, dz
\]

- initial condition \( u(x, 0) = f(x) \) and boundary condition \( u_x(0, t) = 0 \) has a solution

\[
    u(x, t) = \frac{1}{2\sqrt{\pi \alpha t}} \int_{0}^{\infty} f(z) \left[ e^{-\frac{(x-z)^2}{4\alpha t}} + e^{-\frac{(x+z)^2}{4\alpha t}} \right] \, dz
\]

- initial condition \( u(x, 0) = 0 \) and boundary condition \( u(0, t) = g(t) \) has a solution

\[
    u(x, t) = \frac{x}{2\sqrt{\pi \alpha}} \int_{0}^{t} \frac{1}{\sqrt{t-w}} g(w) e^{-\frac{x^2}{4\alpha(t-w)}} \, dw.
\]

### Stochastic Differential Equations

Suppose \( X_t, Y_t \) and \( Z_t \) are Itô processes satisfying the following stochastic differential equations:

\[
    dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_x^t
\]
\[
    dY_t = \mu(Y_t, t)dt + \sigma(Y_t, t)dW_y^t
\]
\[
    dZ_t = \mu(Z_t, t)dt + \sigma(Z_t, t)dW_z^t
\]

where \( W_x^t, W_y^t \) and \( W_z^t \) are standard Wiener processes.

**Reciprocal:**

\[
    d \left( \frac{1}{X_t} \right) = -\frac{dX_t}{X_t} + \left( \frac{dX_t}{X_t} \right)^2.
\]

**Product:**

\[
    \frac{d(X_tY_t)}{X_tY_t} = \frac{dX_t}{X_t} + \frac{dY_t}{Y_t} + \frac{dX_t}{X_t} \frac{dY_t}{Y_t}.
\]

**Quotient:**

\[
    \frac{d \left( \frac{X_t}{Y_t} \right)}{X_tY_t} = \frac{dX_t}{X_t} - \frac{dY_t}{Y_t} - \frac{dX_t}{X_t} \frac{dY_t}{Y_t} + \left( \frac{dY_t}{Y_t} \right)^2.
\]

**Product and Quotient I:**

\[
    \frac{d \left( \frac{X_tY_t}{Z_t} \right)}{X_tY_t} = \frac{dX_t}{X_t} + \frac{dY_t}{Y_t} - \frac{dZ_t}{Z_t} + \frac{dX_t}{X_t} \frac{dY_t}{Y_t} - \frac{dX_t}{X_t} \frac{dZ_t}{Z_t} - \frac{dY_t}{Y_t} \frac{dZ_t}{Z_t} + \left( \frac{dZ_t}{Z_t} \right)^2.
\]
Appendix C

**Product and Quotient II:**

\[
\frac{d}{d \left( \frac{X_t}{Y_t Z_t} \right)} = \frac{dX_t}{X_t} - \frac{dY_t}{Y_t} - \frac{dZ_t}{Z_t} - \frac{dX_t dY_t}{X_t Y_t} - \frac{dX_t dZ_t}{X_t Z_t} + \frac{dY_t dZ_t}{Y_t Z_t}.
\]

**Black–Scholes Model**

**Black–Scholes Equation (Continuous Dividend Yield):** At time \( t \), let the asset price \( S_t \) follows a geometric Brownian motion

\[
\frac{dS_t}{S_t} = (\mu - D)dt + \sigma dW_t
\]

where \( \mu \) is the drift parameter, \( D \) is the continuous dividend yield, \( \sigma \) is the volatility parameter and \( W_t \) is a standard Wiener process. For a European-style derivative \( V(S_t, t) \) written on the asset \( S_t \), it satisfies the Black–Scholes equation with continuous dividend yield

\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} + (r - D)S_t \frac{\partial V}{\partial S_t} - rV(S_t, t) = 0
\]

where \( r \) is the risk-free interest rate. The parameters \( \mu, r, D \) and \( \sigma \) can be either constants, deterministic functions or stochastic processes.

**European Options:** For a European option having the payoff

\[
\Psi(S_T) = \max\{\delta(S_T - K), 0\}
\]

where \( \delta \in \{-1, 1\}, K \) is the strike price and \( T \) is the option expiry time, and if \( r, D \) and \( \sigma \) are constants the European option price at time \( t < T \) is

\[
V(S_t, t; K, T) = \delta S_t e^{-D(T-t)} \Phi(\delta d_+) - \delta K e^{-r(T-t)} \Phi(\delta d_-)
\]

where \( d_{\pm} = \frac{\log(S_t/K) + (r-D \pm \frac{1}{2} \sigma^2)(T-t)}{\sigma \sqrt{T-t}} \) and \( \Phi(\cdot) \) is the cumulative distribution function of a standard normal.

**Reflection Principle:** If \( V(S_t, t) \) is a solution of the Black–Scholes equation then for a constant \( B > 0 \), the function

\[
U(S_t, t) = \left( \frac{S_t}{B} \right)^{2a} V \left( \frac{B^2}{S_t}, t \right), \quad a = \frac{1}{2} \left( 1 - \frac{r-D}{\frac{1}{2} \sigma^2} \right)
\]

also satisfies the Black–Scholes equation.
Black Model

Black Equation: At time $t$, let the asset price $S_t$ follows a geometric Brownian motion

$$\frac{dS_t}{S_t} = (\mu - D)dt + \sigma dW_t$$

where $\mu$ is the drift parameter, $D$ is the continuous dividend yield, $\sigma$ is the volatility parameter and $W_t$ is a standard Wiener process. Consider the price of a futures contract maturing at time $T > t$ on the asset $S_t$ as

$$F(t, T) = S_t e^{r(T-t)}$$

where $r$ is the risk-free interest rate. For a European option on futures $V(F(t, T), t)$ written on a futures contract $F(t, T)$, it satisfies the Black equation

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 F(t, T)^2 \frac{\partial^2 V}{\partial F^2} - rV(F(t, T), t) = 0.$$ 

The parameters $\mu$, $r$, $D$ and $\sigma$ can be either constants, deterministic functions or stochastic processes.

European Options on Futures: For a European option on futures having the payoff

$$\Psi(F(T, T)) = \max\{\delta(F(T, T) - K), 0\}$$

where $\delta \in \{-1, 1\}$, $K$ is the strike price and $T$ is the option expiry time, and if $r$ and $\sigma$ are constants the price of a European option on futures at time $t < T$ is

$$V(F(t, T), t; K, T) = \delta e^{-r(T-t)} [F(t, T)\Phi(\delta d_+) - K\Phi(\delta d_-)]$$

where $d_\pm = \frac{\log(F(t, T)/K)}{\sigma \sqrt{T-t}} \pm \frac{1}{2} \sigma^2(T-t)$ and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal.

Reflection Principle: If $V(F(t, T), t)$ is a solution of the Black equation then for a constant $B > 0$, the function

$$U(F(t, T), t) = \frac{F(t, T)}{B} V\left(\frac{B^2}{F(t, T)}, t\right)$$

also satisfies the Black equation.

Garman–Kohlhagen Model

Garman–Kohlhagen Equation: At time $t$, let the foreign-to-domestic exchange rate $X_t$ follows a geometric Brownian motion

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t$$
where $\mu$ is the drift parameter, $\sigma$ is the volatility parameter and $W_t$ is a standard Wiener process. For a European-style derivative $V(X_t, t)$ which depends on $X_t$, it satisfies the Garman–Kohlhagen equation
\[
\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 X_t^2 \frac{\partial^2 V}{\partial X_t^2} + (r_d - r_f)X_t \frac{\partial V}{\partial X_t} - r_d V(X_t, t) = 0
\]
where $r_d$ and $r_f$ are the domestic and foreign currency risk-free interest rates. The parameters $\mu$, $r_d$, $r_f$ and $\sigma$ can be either constants, deterministic functions or stochastic processes.

**European Options:** For a European option having the payoff
\[
\Psi(X_T) = \max\{\delta(X_T - K), 0\}
\]
where $\delta \in \{-1, 1\}$, $K$ is the strike price and $T$ is the option expiry time, and if $r_d$, $r_f$ and $\sigma$ are constants the European option price (domestic currency in one unit of foreign currency) at time $t < T$ is
\[
V(X_t, t, K, T) = \delta X_t e^{-r_f(T-t)} \Phi(\delta d_+) - \delta Ke^{-r_f(T-t)} \Phi(\delta d_-)
\]
where $d_\pm = \log(X_t/K) + (r_d - r_f \pm \frac{1}{2} \sigma^2)(T-t)/\sigma \sqrt{T-t}$ and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal.

**Reflection Principle:** If $V(X_t, t)$ is a solution of the Garman–Kohlhagen equation then for a constant $B > 0$ the function
\[
U(X_t, t) = \left(\frac{X_t}{B}\right)^{2\alpha} V\left(\frac{B^2}{X_t}, t\right), \quad \alpha = \frac{1}{2} \left(1 - \frac{r_d - r_f}{\frac{1}{2} \sigma^2}\right)
\]
also satisfies the Garman–Kohlhagen equation.
Bibliography


SET NOTATION

∈ is an element of
∉ is not an element of
Ω sample space
ℰ universal set
∅ empty set
A subset of Ω
Ac complement of set A
|A| cardinality of A
ℕ set of natural numbers, \{1, 2, 3, \ldots\}
ℕ₀ set of natural numbers including zero, \{0, 1, 2, \ldots\}
ℤ set of integers, \{0, ±1, ±2, ±3, \ldots\}
ℤ⁺ set of positive integers, \{1, 2, 3, \ldots\}
ℝ set of real numbers
ℝ⁺ set of positive real numbers, \{x ∈ ℝ : x > 0\}
ℂ set of complex numbers
A × B cartesian product of sets A and B, A × B = \{(a, b) : a ∈ A, b ∈ B\}
a ∼ b a is equivalent to b
⊂ subset
⊂⊂ proper subset
∩ intersection
∪ union
\setminus difference
Δ symmetric difference
sup supremum or least upper bound
inf infimum or greatest lower bound
[a, b] the closed interval \{x ∈ ℝ : a ≤ x ≤ b\}
[a, b) the interval \{x ∈ ℝ : a ≤ x < b\}
(a, b] the interval \{x ∈ ℝ : a < x ≤ b\}
(a, b) the open interval \{x ∈ ℝ : a < x < b\}
ℱ, ℰ, ℋ σ-algebra (or σ-fields)
MATHEMATICAL NOTATION

\[ x^+ = \max\{x, 0\} \]
\[ x^- = \min\{x, 0\} \]
\[ |x| \text{ largest integer not greater than and equal to } x, \max\{m \in \mathbb{Z} \mid m \leq x\} \]
\[ [x] \text{ smallest integer greater than and equal to } x, \min\{n \in \mathbb{Z} \mid n \geq x\} \]
\[ x \lor y \text{ max}\{x, y\} \]
\[ x \land y \text{ min}\{x, y\} \]
\[ i = \sqrt{-1} \]
\[ \infty \text{ infinity} \]
\[ \exists \text{ there exists} \]
\[ \exists! \text{ there exists a unique} \]
\[ \forall \text{ for all} \]
\[ \approx \text{ approximately equal to} \]
\[ p \implies q \text{ } p \text{ implies } q \]
\[ p \iff q \text{ } p \text{ is implied by } q \]
\[ f : X \mapsto Y \text{ } f \text{ is a function where every element of } X \text{ has an image in } Y \]
\[ f(x) \text{ the value of the function } f \text{ at } x \]
\[ \lim_{x \to a} f(x) \text{ limit of } f(x) \text{ as } x \text{ tends to } a \]
\[ \delta x, \Delta x \text{ increment of } x \]
\[ f^{-1}(x) \text{ the inverse function of } f(x) \]
\[ f'(x), f''(x) \text{ the first and second-order derivative of the function } f(x) \]
\[ \frac{dy}{dx}, \frac{d^2y}{dx^2} \text{ first and second-order derivative of } y \text{ with respect to } x \]
\[ \int y \, dx, \int_a^b y \, dx \text{ the indefinite and definite integral of } y \text{ with respect to } x \]
\[ \frac{\partial f}{\partial x_i}, \frac{\partial^2 f}{\partial x_i^2} \text{ first and second-order partial derivative of } f \text{ with respect to } x_i \]
\[ \frac{\partial^2 f}{\partial x_i \partial x_j} \text{ second-order partial derivative of } f \text{ with respect to } x_i \text{ and } x_j \]
\[ \log_a x \text{ logarithm of } x \text{ to the base } a \]
\[ \log x \text{ natural logarithm of } x \]
\[ \sum_{i=1}^n a_i \text{ } a_1 + a_2 + \ldots + a_n \]
\[ \prod_{i=1}^n a_i \text{ } a_1 \times a_2 \times \ldots \times a_n \]
\[ |a| \text{ modulus of } a \]
\[ (\sqrt[n]{a})^m \]
\[ a \frac{n!}{k!(n-k)!} \text{ for } n, k \in \mathbb{Z}^+ \]
\[ \delta(x) \text{ Dirac delta function} \]
\[ H(x) \text{ Heaviside step function} \]
Notation

\begin{align*}
\Gamma(t) & \quad \text{gamma function} \\
B(x, y) & \quad \text{beta function} \\
a & \quad \text{a vector } a \\
|a| & \quad \text{magnitude of a vector } a \\
a \cdot b & \quad \text{scalar or dot product of vectors } a \text{ and } b \\
a \times b & \quad \text{vector or cross-product of vectors } a \text{ and } b \\
M & \quad \text{a matrix } M \\
M^T & \quad \text{transpose of a matrix } M \\
M^{-1} & \quad \text{inverse of a square matrix } M \\
|M| & \quad \text{determinant of a square matrix } M
\end{align*}

\textbf{PROBABILITY NOTATION}

\begin{align*}
A, B, C & \quad \text{events} \\
\mathbb{I}_A & \quad \text{indicator of the event } A \\
\mathbb{P}, \mathbb{Q} & \quad \text{probability measures} \\
\mathbb{P}(A) & \quad \text{probability of event } A \\
\mathbb{P}(A|B) & \quad \text{probability of event } A \text{ conditional on event } B \\
X, Y, Z & \quad \text{random variables} \\
X, Y, Z & \quad \text{random vectors} \\
\mathbb{P}(X = x) & \quad \text{probability mass function of a discrete random variable } X \\
f_X(x) & \quad \text{probability density function of a continuous random variable } X \\
F_X(x), \mathbb{P}(X \leq x) & \quad \text{cumulative distribution function of a random variable } X \\
M_X(t) & \quad \text{moment generating function of a random variable } X \\
\varphi_X(t) & \quad \text{characteristic function of a random variable } X \\
P(X = x, Y = y) & \quad \text{joint probability mass function of discrete variables } X \text{ and } Y \\
f_{XY}(x, y) & \quad \text{joint probability density function of continuous random variables } X \text{ and } Y \\
F_{XY}(x, y), \mathbb{P}(X \leq x, Y \leq y) & \quad \text{joint cumulative distribution function of random variables } X \text{ and } Y \\
M_{XY}(s, t) & \quad \text{joint moment generating function of random variables } X \text{ and } Y \\
\varphi_{XY}(s, t) & \quad \text{joint characteristic function of random variables } X \text{ and } Y \\
p(x; t; y, T) & \quad \text{transition probability density of } y \text{ at time } T \text{ starting at time } t \text{ at point } x \\
\sim & \quad \text{is distributed as} \\
\not\sim & \quad \text{is not distributed as} \\
\ast & \quad \text{is approximately distributed as} \\
\xrightarrow{a.s} & \quad \text{converges almost surely} \\
\xrightarrow{r} & \quad \text{converges in the } r\text{-th mean} \\
\xrightarrow{p} & \quad \text{converges in probability} \\
\xrightarrow{D} & \quad \text{converges in distribution} \\
X \overset{d}{=} Y & \quad \text{X and } Y \text{ are identically distributed random variables}
\end{align*}
<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X \perp Y$</td>
<td>$X$ and $Y$ are independent random variables</td>
</tr>
<tr>
<td>$X \not\perp Y$</td>
<td>$X$ and $Y$ are not independent random variables</td>
</tr>
<tr>
<td>$\mathbb{E}(X)$</td>
<td>Expectation of random variable $X$</td>
</tr>
<tr>
<td>$\mathbb{E}(X</td>
<td>\mathcal{F})$</td>
</tr>
<tr>
<td>$\text{Var}(X)$</td>
<td>Variance of random variable $X$</td>
</tr>
<tr>
<td>$\text{Var}(X</td>
<td>\mathcal{F})$</td>
</tr>
<tr>
<td>$\text{Cov}(X,Y)$</td>
<td>Covariance of random variables $X$ and $Y$</td>
</tr>
<tr>
<td>$\rho_{xy}$</td>
<td>Correlation between random variables $X$ and $Y$</td>
</tr>
<tr>
<td>Bernoulli($p$)</td>
<td>Bernoulli distribution with mean $p$ and variance $p(1-p)$</td>
</tr>
<tr>
<td>Geometric($p$)</td>
<td>Geometric distribution with mean $p^{-1}$ and variance $(1-p)p^{-2}$</td>
</tr>
<tr>
<td>Binomial($n,p$)</td>
<td>Binomial distribution with mean $np$ and variance $np(1-p)$</td>
</tr>
<tr>
<td>BN($n,r$)</td>
<td>Negative binomial distribution with mean $rp$ and variance $r(1-p)p^{-2}$</td>
</tr>
<tr>
<td>Poisson($\lambda$)</td>
<td>Poisson distribution with mean $\lambda$ and variance $\lambda$</td>
</tr>
<tr>
<td>Exp($\lambda$)</td>
<td>Exponential distribution with mean $\lambda^{-1}$ and variance $\lambda^{-2}$</td>
</tr>
<tr>
<td>Gamma($\alpha,\lambda$)</td>
<td>Gamma distribution with mean $\alpha \lambda^{-1}$ and variance $\alpha \lambda^{-2}$</td>
</tr>
<tr>
<td>$\mathcal{U}(a,b)$</td>
<td>Uniform distribution with mean $\frac{1}{2}(a+b)$ and variance $\frac{1}{12}(b-a)^2$</td>
</tr>
<tr>
<td>$\mathcal{N}(\mu,\sigma^2)$</td>
<td>Normal distribution with mean $\mu$ and variance $\sigma^2$</td>
</tr>
<tr>
<td>log-$\mathcal{N}(\mu,\sigma^2)$</td>
<td>Lognormal distribution with mean $e^{\mu + \frac{1}{2}\sigma^2}$ and variance $(e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$</td>
</tr>
<tr>
<td>$\chi^2(k)$</td>
<td>Chi-square distribution with mean $k$ and variance $2k$</td>
</tr>
<tr>
<td>$\mathcal{N}_n(\mu,\Sigma)$</td>
<td>Multivariate normal distribution with $n$-dimensional mean vector $\mu$ and $n \times n$ covariance matrix $\Sigma$</td>
</tr>
<tr>
<td>$\Phi(\cdot)$, $\Phi(x)$</td>
<td>Cumulative distribution function of a standard normal</td>
</tr>
<tr>
<td>$\Phi(x,y,\rho_{xy})$</td>
<td>Cumulative distribution function of a standard bivariate normal with correlation coefficient $\rho_{xy}$</td>
</tr>
<tr>
<td>$W_t$</td>
<td>Standard Wiener process, $W_t \sim \mathcal{N}(0,t)$</td>
</tr>
<tr>
<td>$N_t$</td>
<td>Poisson process, $N_t \sim \text{Poisson}(\lambda t)$</td>
</tr>
</tbody>
</table>
Index

adapted stochastic processes, 2, 303
admissible trading strategy, 186
American options, 53
appendices, 1, 331–63
arbitrage, 53, 185–7, 232, 322–30
arithmetic Brownian motion, 128–9, 222–3, 229–30
see also Bachelier model
stock price with continuous dividend yield, 229–30
arithmetic series, formulae, 332
arrival time distribution, Poisson process, 256–8
see also stock...
fundamental theories, 232–5
attainable contingent claim, 186–7
Bachelier model
see also arithmetic Brownian motion
definition and formulae, 128–9
backward Kolmogorov equation, 97, 98–102, 149–50, 153–4, 180–1
see also diffusion; parabolic...
definition, 97, 98–9, 149–50, 153–4, 180–1
multi-dimensional diffusion process, 99–102, 155–83
one-dimensional diffusion process, 87–8, 123–55
one-dimensional random walk, 153–4
two-dimensional random walk, 180–1
Bayes’ Formula, 7
Bayes’ rule, 341
Bernoulli differential equation, 357–8
Bernoulli distribution, 11, 12, 13, 14–15, 349
Bessel process, 163–6
beta function, 338
binomial distribution, 12–14, 349–50
bivariate continuous random variables, 345–7
see also continuous...
bivariate discrete random variables, 343–4
see also discrete...
bivariate normal distribution, 27–9, 34–40, 60–2, 158, 165, 352–3
see also normal distribution
covariance, 28–9, 57–9, 158, 165
marginal distributions, 27
Black equation, 362
Black model, 362
Black–Scholes equation, 54, 96, 236–8, 361
see also partial differential equations
reflection principle, 54–5, 361
Black–Scholes model, 129–30, 185, 236–8, 361
see also geometric Brownian motion
Bonferroni’s inequality, 7
Boole’s inequality, 6, 7, 10
Borel–Cantelli lemma, 10–11
Brownian bridge process, 137–8
Brownian motion, 51–93, 95–183, 185–92, 221–4, 227–8, 361–3
see also arithmetic...; diffusion...; geometric...; random walks; Wiener processes
definitions and formulae, 51, 128–32, 138–9, 221–4, 227–30
càdlàg process, 246
Cauchy–Euler equation, 358–9
cdf see cumulative distribution function
central limit theorem, 12, 13, 57, 355
CEV see constant elasticity of variance model
change of measure, 43, 185–242, 249–51, 301–30
see also Girsanov’s theorem
definitions, 185–92, 249–51
Chebyshev’s inequality, 40–1, 75
chi-square distribution, 24–6, 352
CIR see Cox–Ingersoll–Ross model
Clewlow–Strickland 1-factor model, 141–4
compensated Poisson process, 245, 262–3, 266–7, 323–6, 330
see also Poisson...
complement, probability concepts, 1, 4–5, 341
complete market, 187, 233, 322–30
see also Poisson...
decomposition, 268–9, 306–8
Girsanov’s theorem, 249–51
martingales, 265–7, 323–5
concave function, 339
conditional, probability concepts, 341
conditional Jensen’s inequality, properties of conditional expectation, 3, 48–9, 75–6, 192–4
conditional probability density function formulae, 346
mass function formulae, 343
properties of conditional expectation, 3, 43–5
conditional variance, 343, 346
constant elasticity of variance model (CEV), 145–6
contingent claims, 185–92, 245
see also derivatives
continuous distributions, 350–3
convergence of random variables, 10–11, 209–11, 353–4
convex function, 3, 42, 48–9, 75–6, 192–4, 338–9
convolution formulae, 25–6, 348
countable unions, 1–2
counting process, 243–330
see also Poisson...
definition and formulae, 243–4
covariance, 28–9, 57–60, 64–8, 117–18, 157–8, 165–6, 251–2, 344, 347
matrices, 59, 64–8
covariance of bivariate normal distribution, 28–9, 57–60, 158, 165–6
covariance of two standard Wiener processes, 57–60, 64–8
Cox process (doubly stochastic Poisson process), 243, 245
see also Poisson process
Cox–Ingersoll–Ross model (CIR), 135–7, 171–4, 178
cumulative distribution function (cdf), 16–20, 22–4, 30–2, 37–40, 214–18, 342, 343, 345, 361–3
cumulative intensity, 245
see also intensity...
De Moivre’s formula, 333
De Morgan’s law, 4, 7, 10
decomposition of a compound Poisson process, 268–9, 306–8
see also ordinary...; partial...; stochastic... differential-difference equations, 253–5
Dirac delta function, 339
see also Heaviside step function
discounted portfolio value, 187–92, 224–7, 228–32, 242, 324–30
discrete distributions, 349–50
discrete-time martingales, 53
dominated convergence theorem, 354
Donsker theorem, 56–7
doob’s maximal inequality, 76–80
elementary process, 96
equivalent martingale measure, 185, 188–9, 209–11, 222–5, 230–1, 234–5, 238–42, 325–30
see also risk-neutral...
Euler’s formula, 333
events, definition, 1, 243–51
exclusion for probability, definition, 7–10
exponential martingale process, 263–4
see also martingales
Feynman–Kac theorem, 97–102, 147–9, 178–80
see also diffusion
multi-dimensional diffusion process, 178–80
one-dimensional diffusion process, 147–9, 178
fields, definition, 1–2
first fundamental theorem of asset pricing, definition, 232
first passage time density function, 85–9, 218–19
first passage time of a standard Wiener processes, 53, 76–89, 218–19
hitting a sloping line, 218–19
Laplace transform, 83–4
Fokker–Planck equation see forward Kolmogorov equation
folded normal distribution, 22–4, 72
see also normal distribution
foreign exchange (FX), 51, 54, 192, 238–42, 362
foreign-denominated stock price under
domestic risk-neutral measure, 241–2
risk-neutral measure, 238–42
forward curve from an asset price following a
geometric Brownian motion, 138–9
forward curve from an asset price following a
geometric mean-reverting
process, 139–40, 169–71
forward curves, 138–44
forward Kolmogorov equation, 97, 98–102,
150–3, 154–5, 181–3
see also diffusion; parabolic...
multi-dimensional diffusion process, 102
one-dimensional diffusion process, 150–3
one-dimensional random walk, 154–5
two-dimensional random walk, 180–3
Fubini’s theorem, 339
FX see foreign exchange
Gabillon 2-factor model, 169–71
gamma distribution, 16–17, 257, 352
Garman–Kohlhagen equation/model, 362–3
general probability theory, concepts, 1–49,
185–92, 341–55
generalised Brownian motion, 130–2
generalised Itô integral, 118–19
geometric average, 146–7
geometric Brownian motion, 70–1, 129–32,
138–9, 146–7, 221–2, 227–8, 361, 362–3
see also Black–Scholes options pricing
model; Brownian motion
definition and formulae, 129–32, 138–9,
146–7, 221–2, 227–8
forward curve from an asset price, 138–9
Markov property, 70–1
stock price with continuous dividend
yield, 227–8
geometric distribution, 349
geometric mean-reverting process, 134–5,
139–40, 169–71, 291–5
definition and formulae, 134–5, 139–40,
169–71, 291–5
forward curve from an asset price, 139–40,
169–71
jump-diffusion process, 291–5
definition and formulae, 133–2
Girsanov’s theorem, 185, 189–92, 194–242,
249–51, 298–322
see also real-world measure; risk-neutral
measure
corollaries, 189–90
definitions, 185, 189–92, 194–225, 249–51, 298–322
formulae, 189–92, 194–225, 249–51,
298–322
jump processes, 298–322
Poisson process, 249–51, 298–322
running maximum and minimum of a Wiener
process, 214–18
hazard function, 245
see also intensity...
hazard process, 245
heat equations, 359–60
Heaviside step function, 339
see also Dirac delta function
Heston stochastic volatility model, 175–8
Hölder’s inequality, 41–2, 72–4, 79, 193,
204–5
homogeneous heat equations, 359–60
homogeneous Poisson process see Poisson
process
hyperbolic functions, 74, 333
inclusion and exclusion for probability,
definition, 7–10
incomplete markets, 187, 324–30
independence, properties, 3, 11, 47–8, 163–6,
210, 259–61
independent events, probability
concepts, 259–61, 341, 344, 347
integrable random variable, 3, 44, 45–9, 353
see also random variables
integral calculus, definition, 7–10
integrated square-root process, 171–4
integration by parts, 19–20, 115–18, 147, 151,
174, 257–8, 337
Itô integral, 102–7, 108–116, 118–20, 144,
148–9, 163–6, 177–8, 187–92, 195–6,
205–7
Itô isometry, 113–14
Itô processes, 54, 95–123, 360–1
Itô’s theorem see Itô’s lemma
Itô’s lemma, 96–7, 99–100, 102–27, 129–33,
134–7, 139–46, 147–53, 155–62, 166–8,
170–80, 196, 203–5, 212–13, 220–7,
234–41, 246–51, 278–80, 283–98,
307–8, 314–22
see also stochastic differential equations;
Taylor series
multi-dimensional Itô formulae for
jump-diffusion process, 247–9
one-dimensional Itô formulae for
jump-diffusion process, 246–7
Jensen’s inequality, properties of conditional
expectation, 3, 48–9, 75–6, 192–4
joint characteristic function, 344, 347
joint cumulative distribution function, 16–17, 32, 212–13, 343, 345
joint distribution of standard Wiener processes, 58–60, 64, 158, 165–6
joint moment generating function, 122–3, 267, 280–1, 306, 318–19, 321–2, 344, 347
joint probability density function, 16–17, 27–34, 86–9, 211–18, 345–6, 352
joint probability mass function, definition, 343
see also diffusion; Poisson...; Wiener...
concepts, 246–51, 281–98, 325–30
general mean-reverting process, 291–5
Merton’s model, 285–8, 327–30
multi-dimensional Itô formulae, 247–9
one-dimensional Itô formulae, 246–7
pure jump process, 281–5, 323–5, 327–30
simple jump-diffusion process, 325–7
jumps, 95, 243–330
see also Poisson process
Girsanov’s theorem, 298–322
risk-neutral measure, 322–30
Kou’s model, 295–8
Laplace transform of first passage time, 83–4
laws of large numbers, formulae, 354–5
Lebesgue–Stieltjes integral, 246
Lévy processes, 55, 119–23, 207, 209
L’Hospital rule, formulae, 336
linearity, properties of conditional expectation, 3, 44, 45
Maclaurin series, 335
marginal distributions of bivariate normal distribution, 27
marginal probability density function, 27–9, 346
marginal probability mass function, 343
Markov property of a geometric Brownian motion, 70–1
Markov property of Poisson process, 244–5, 261–2, 268–9
Markov property of Wiener processes, 52, 68–71, 97
definition and formulae, 51–2, 68–71
Markov’s inequality, definition and formulae, 40
martingale representation theorem, 187–8, 190, 192–4, 219–26
see also equivalent...; stochastic processes;
Wiener processes
compound Poisson process, 265–8, 322–4
continuous processes property, 53
discrete processes property, 53
theorems, 53–5, 187–92
maximum of two correlated normal distributions, definition, 29–32
mean value theorem, 105–6
mean-reversion, 134–5, 139–44, 169–71, 288–95
see also geometric mean-reverting process
measurability, properties of conditional expectation, 3, 43–4, 45, 199–200
measurable space, definition, 2
measure theory, definition, 2
measures, 2, 185–242, 249–51, 301–30
see also Girsanov’s theorem; real-world...;
risk-neutral...
change of measure, 185–242, 249–51, 301–30
Merton’s model, definition and formulae, 285–8, 327–30
minimum and maximum of two correlated normal distributions, 29–32
Minkowski’s inequality, 42
monotone convergence theorem, 354
monotonicity, properties of conditional expectation, 3, 44–45
multi-dimensional diffusion process
see also diffusion
backward Kolmogorov equation, 101–2
definition and formulae, 99–102, 155–83
Feynman–Kac theorem, 178–80
forward Kolmogorov equation, 101
problems and solutions, 155–83
multi-dimensional Girsanov theorem, 190–1, 208–9, 249–51
multi-dimensional Itô formulae, 99–100
multi-dimensional Itô formulae jump-diffusion process, 247–9
multi-dimensional Lévy characterisation theorem, 121–3, 209
multi-dimensional martingale representation theorem, 191–2
multi-dimensional Novikov condition, 207–8
multi-dimensional Wiener processes, 54, 64–8, 99–102, 163–6
multiplication, probability concepts, 341
multivariate normal distribution, 353
see also normal distribution
mutually exclusive events, probability concepts, 12, 341
negative binomial distribution, 350
see also bivariate...; folded...; log...; multivariate...
minimum and maximum of two correlated normal distributions, 29–32
Novikov’s conditions, 194–6, 207–8
numéraire, definition, 191–2
one-dimensional diffusion process, 97–9, 123–55, 178
see also diffusion
backward Kolmogorov equation, 149–50 definition and formulæ, 97–9, 147–53 Feynman–Kac formulæ, 147–9, 178 forward Kolmogorov equation, 150–53 one-dimensional Girsanov theorem, 189–90, 205–7 one-dimensional Itô formulæ for jump-diffusion process, 246–7 one-dimensional Lévy characterisation theorem, 119–21, 207 one-dimensional martingale representation theorem, 187–8, 190 one-dimensional random walk see also random walks
reflection principle (continued)
Garman–Kohlhagen equation, 362–3
Wiener processes, 53–5, 60, 84–9
risk-neutral measure, 53, 185, 188–242, 322–30
see also equivalent martingale...; Girsanov’s theorem
definition, 185, 188–9, 221–42, 322–30
FX, 238–42
jump processes, 322–30
problems and solutions, 221–42, 322–30
running maximum and minimum of a Wiener process, Girsanov’s theorem, 214–18

sample space, definition, 1, 4, 341–2
scaled symmetric random walk, 51–5
see also random walks
SDEs see stochastic differential equations
second fundamental theorem of asset pricing, definition, 233

second-order ordinary differential equations
formulae, 358–9
variation of parameters, 358–9
self-financing trading strategy, 186–8, 192, 225–7, 236–8, 326–30
sets, 1, 2–11
definition, 1
σ-sigma-algebra, 1–5, 43–9, 201–5, 261
simple jump process, 322–3
simple jump-diffusion process, 325–7
simple process see elementary process
skew, 54
speculation uses of derivatives, 185
square integrable random variable, 95–6, 353
covariance of two standard Wiener processes, 57–60, 64–8, 117–18
definition, 52, 54, 189–90, 250–1, 325–6
joint distribution of standard Wiener processes, 58–63, 64, 158–9, 165–6
stationary and independent increments, Poisson process, 259–81
stochastic differential equations
(SDEs), 95–183, 237–42, 246–51, 315–30, 360–1
definition, 95–102, 246–9
integral calculus contrasts, 95–6
stochastic processes, 2, 51, 52, 185–242, 243–330
see also martingales; Poisson...; Wiener...
definitions, 2, 51, 52
stochastic volatility, 54, 95, 175–8
stopping times, Wiener processes, 53–5, 80–9, 218–19, 232
Stratonovich integral, 103–6
the strong law of large numbers, formulae, 355
strong Markov property, 52, 84–9
supermartingales, 53, 76–9
symmetric random walk, 51–68, 88–9, 153–5, 180–3
see also random walks
time inversion, Wiener processes, 62–3
time reversal, Wiener processes, 63–4
time shifting, Wiener processes, 61
total probability of all possible values, formulae, 342–5
tower property, conditional expectation, 3, 46, 49, 192–4, 197–9, 270–4, 284–5
trading strategy, 186–8, 191–2, 225–7, 236–8, 326–30
transition probability density function, 98–102, 149–53
see also backward Kolmogorov...; forward Kolmogorov...
two-dimensional random walk
see also random walks
backward Kolmogorov equation, 180–1
forward Kolmogorov equation, 181–3
uniform distribution, 350
union, probability concepts, 341
univariate continuous random variables, 344–7
see also continuous...
univariate discrete random variables, 342–4
see also discrete...
see also covariance...
constant elasticity of variance model, 145–6
see also local...; stochastic...
the weak law of large numbers, formulae, 354
see also Brownian motion; diffusion...;
martingales; random walks
covariance of two standard Wiener processes, 57–60, 64–8, 115–18
Index 379

first passage time, 53, 76–89, 218–19
joint distribution of standard Wiener processes, 58–60, 64, 158, 165–6
Markov property, 52, 68–71, 97–102
multi-dimensional Wiener processes, 54, 64–8, 99–102, 163–6

reflection principle, 54–5, 60, 84–9
running maximum and minimum of a Wiener process, 214–18
WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.