THEORY OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH JUMPS AND APPLICATIONS

MATHEMATICAL AND ANALYTICAL TECHNIQUES WITH APPLICATIONS TO ENGINEERING
MATHEMATICAL AND ANALYTICAL TECHNIQUES WITH APPLICATIONS TO ENGINEERING

Alan Jeffrey, Consulting Editor

Published:

Inverse Problems
A. G. Ramm

Singular Perturbation Theory
R. S. Johnson

Methods for Constructing Exact Solutions of Partial Differential Equations with Applications
S. V. Meleshko

Stochastic Differential Equations with Applications
R. Situ

Forthcoming:

The Fast Solution of Boundary Integral Equations
S. Rjasanow and O. Steinbach
## Contents

Preface xi

Acknowledgement xvii

Abbreviations and Some Explanations xix

I Stochastic Differential Equations with Jumps in $\mathbb{R}^d$ 1

1 Martingale Theory and the Stochastic Integral for Point Processes 3

1.1 Concept of a Martingale . . . . . . . . . . . . . . . . . . . . . . 3
1.2 Stopping Times. Predictable Process . . . . . . . . . . . . . . . 5
1.3 Martingales with Discrete Time . . . . . . . . . . . . . . . . . . 8
1.4 Uniform Integrability and Martingales . . . . . . . . . . . . . . 12
1.5 Martingales with Continuous Time . . . . . . . . . . . . . . . . . 17
1.6 Doob-Meyer Decomposition Theorem . . . . . . . . . . . . . . . 19
1.7 Poisson Random Measure and Its Existence . . . . . . . . . . . . 26
1.8 Poisson Point Process and Its Existence . . . . . . . . . . . . . 28
1.9 Stochastic Integral for Point Process. Square Integrable Mar-
ingales . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

2 Brownian Motion, Stochastic Integral and Ito’s Formula 39
2.1 Brownian Motion and Its Nowhere Differentiability
2.2 Spaces \( L^0 \) and \( L^2 \)
2.3 Ito's Integrals on \( L^2 \)
2.4 Ito's Integrals on \( L^{2,\text{loc}} \)
2.5 Stochastic Integrals with respect to Martingales
2.6 Ito's Formula for Continuous Semi-Martingales
2.7 Ito's Formula for Semi-Martingales with Jumps
2.8 Ito's Formula for \( d \)-dimensional Semi-Martingales. Integration by Parts
2.9 Independence of BM and Poisson Point Processes
2.10 Some Examples
2.11 Strong Markov Property of BM and Poisson Point Processes
2.12 Martingale Representation Theorem

3 Stochastic Differential Equations
3.1 Strong Solutions to SDE with Jumps
3.1.1 Notation
3.1.2 A Priori Estimate and Uniqueness of Solutions
3.1.3 Existence of Solutions for the Lipschitzian Case
3.2 Exponential Solutions to Linear SDE with Jumps
3.3 Girsanov Transformation and Weak Solutions of SDE with Jumps
3.4 Examples of Weak Solutions

4 Some Useful Tools in Stochastic Differential Equations
4.1 Yamada-Watanabe Type Theorem
4.2 Tanaka Type Formula and Some Applications
4.2.1 Localization Technique
4.2.2 Tanaka Type Formula in \( d \)-Dimensional Space
4.2.3 Applications to Pathwise Uniqueness and Convergence of Solutions
4.2.4 Tanaka Type Formual in 1-Dimensional Space
4.2.5 Tanaka Type Formula in The Component Form
4.2.6 Pathwise Uniqueness of solutions
4.3 Local Time and Occupation Density Formula
4.4 Krylov Estimation
4.4.1 The case for 1-dimensional space
4.4.2 The Case for \( d \)-dimensional space
4.4.3 Applications to Convergence of Solutions to SDE with Jumps

5 Stochastic Differential Equations with Non-Lipschitzian Coefficients
5.1 Strong Solutions. Continuous Coefficients with \( \rho \)- Conditions
5.2 The Skorohod Weak Convergence Technique
Contents vii

5.3 Weak Solutions. Continuous Coefficients .......................... 147
5.4 Existence of Strong Solutions and Applications to ODE .... 153
5.5 Weak Solutions. Measurable Coefficient Case ................. 153

II Applications 161

6 How to Use the Stochastic Calculus to Solve SDE 163
6.1 The Foundation of Applications: Ito's Formula and Girsanov's
    Theorem ........................................... 163
6.2 More Useful Examples .................................... 167

7 Linear and Non-linear Filtering 169
7.1 Solutions of SDE with Functional Coefficients and Girsanov
    Theorems ........................................... 169
7.2 Martingale Representation Theorems (Functional Coefficient
    Case) ............................................. 177
7.3 Non-linear Filtering Equation ................................. 180
7.4 Optimal Linear Filtering .................................... 191
7.5 Continuous Linear Filtering. Kalman-Bucy Equation ........... 194
7.6 Kalman-Bucy Equation in Multi-Dimensional Case ............. 196
7.7 More General Continuous Linear Filtering ...................... 197
7.8 Zakai Equation ......................................... 201
7.9 Examples on Linear Filtering ................................. 203

8 Option Pricing in a Financial Market and BSDE 205
8.1 Introduction ............................................ 205
8.2 A More Detailed Derivation of the BSDE for Option Pricing.. 208
8.3 Existence of Solutions with Bounded Stopping Times ....... 209
  8.3.1 The General Model and its Explanation ................ 209
  8.3.2 A Priori Estimate and Uniqueness of a Solution ....... 213
  8.3.3 Existence of Solutions for the Lipschitzian Case ....... 215
8.4 Explanation of the Solution of BSDE to Option Pricing .... 219
  8.4.1 Continuous Case .................................. 219
  8.4.2 Discontinuous Case ................................ 220
8.5 Black-Scholes Formula for Option Pricing. Two Approaches 223
8.6 Black-Scholes Formula for Markets with Jumps ................. 229
8.7 More General Wealth Processes and BSDEs ...................... 234
8.8 Existence of Solutions for Non-Lipschitzian Case ............ 236
8.9 Convergence of Solutions ................................ 239
8.10 Explanation of Solutions of BSDEs to Financial Markets ... 241
8.11 Comparison Theorem for BSDE with Jumps .................... 243
8.12 Explanation of Comparison Theorem. Arbitrage-Free Market 250
8.13 Solutions for Unbounded (Terminal) Stopping Times ....... 254
8.14 Minimal Solution for BSDE with Discontinuous Drift ....... 258
8.15 Existence of Non-Lipschitzian Optimal Control. BSDE Case 262
8.16 Existence of Discontinuous Optimal Control. BSDEs in $\mathbb{R}^1$. 267
8.17 Application to PDE. Feynman-Kac Formula 271

9 Optimal Consumption by H-J-B Equation and Lagrange Method 277
9.1 Optimal Consumption 277
9.2 Optimization for a Financial Market with Jumps by the Lagrange Method 279
  9.2.1 Introduction 280
  9.2.2 Models 280
  9.2.3 Main Theorem and Proof 282
  9.2.4 Applications 286
  9.2.5 Concluding Remarks 290

10 Comparison Theorem and Stochastic Pathwise Control 291
10.1 Comparison for Solutions of Stochastic Differential Equations 292
  10.1.1 1-Dimensional Space Case 292
  10.1.2 Component Comparison in $d$-Dimensional Space 293
  10.1.3 Applications to Existence of Strong Solutions. Weaker Conditions. 294
10.2 Weak and Pathwise Uniqueness for 1-Dimensional SDE with Jumps 298
10.3 Strong Solutions for 1-Dimensional SDE with Jumps 300
  10.3.1 Non-Degenerate Case 300
  10.3.2 Degenerate and Partially-Degenerate Case 303
10.4 Stochastic Pathwise Bang-Bang Control for a Non-linear System 312
  10.4.1 Non-Degenerate Case 312
  10.4.2 Partially-Degenerate Case 316
10.5 Bang-Bang Control for $d$-Dimensional Non-linear Systems 319
  10.5.1 Non-Degenerate Case 319
  10.5.2 Partially-Degenerate Case 322

11 Stochastic Population Control and Reflecting SDE 329
11.1 Introduction 330
11.2 Notation 332
11.3 Skorohod's Problem and its Solutions 335
11.4 Moment Estimates and Uniqueness of Solutions to RSDE 342
11.5 Solutions for RSDE with Jumps and with Continuous Coefficients 345
11.6 Solutions for RSDE with Jumps and with Discontinuous Coefficients 349
11.7 Solutions to Population SDE and Their Properties 352
11.8 Comparison of Solutions and Stochastic Population Control 363
11.9 Caculation of Solutions to Population RSDE .............. 372

12 Maximum Principle for Stochastic Systems with Jumps 377
  12.1 Introduction ............................................. 377
  12.2 Basic Assumption and Notation ................................. 378
  12.3 Maximum Principle and Adjoint Equation as BSDE with Jumps .............. 379
  12.4 A Simple Example ....................................... 380
  12.5 Intuitive thinking on the Maximum Principle .............. 381
  12.6 Some Lemmas ........................................... 383
  12.7 Proof of Theorem 354 .................................. 386

A A Short Review on Basic Probability Theory 389
  A.1 Probability Space, Random Variable and Mathematical Expectation .............. 389
  A.2 Gaussian Vectors and Poisson Random Variables .............. 392
  A.3 Conditional Mathematical Expectation and its Properties ....... 395
  A.4 Random Processes and the Kolmogorov Theorem .............. 397

B Space D and Skorohod's Metric 401

C Monotone Class Theorems. Convergence of Random Processes 407
  C.1 Monotone Class Theorems ................................. 407
  C.2 Convergence of Random Variables ............................... 409
  C.3 Convergence of Random Processes and Stochastic Integrals 411

References 415

Index 431
Preface

Stochastic differential equations (SDEs) were first initiated and developed by K. Ito (1942). Today they have become a very powerful tool applied to Mathematics, Physics, Chemistry, Biology, Medical science, and almost all sciences. Let us explain why we need SDEs, and how the contents in this book have been arranged.

In nature, physics, society, engineering and so on we always meet two kinds of functions with respect to time: one is deterministic, and another is random. For example, in a financial market we deposit money $\pi_t$ in a bank. This can be seen as our having bought some units $\eta_t^0$ of a bond, where the bond's price $P_t^0$ satisfies the following ordinary differential equation

$$dP_t^0 = P_t^0 r_t dt, \quad P_0^0 = 1, \quad t \in [0, T],$$

where $r_t$ is the rate of the bond, and the money that we deposit in the bank is $\pi_t = \eta_t^0 P_t^0 = \eta_t^0 \exp[\int_0^t r_s ds]$. Obviously, usually, $P_t^0 = \exp[\int_0^t r_s ds]$ is non-random, since the rate $r_t$ is usually deterministic. However, if we want to buy some stocks from the market, each stock's price is random. For simplicity let us assume that in the financial market there is only one stock, and its price is $P_t^1$. Obviously, it will satisfy a differential equation as follows:

$$dP_t^1 = P_t^1 (b_t dt + d(\text{a stochastic perturbation})), \quad P_0^1 = P_0^1, \quad t \in [0, T],$$

where all of the above processes are 1-dimensional. Here the stochastic perturbation is very important, because it influences the price of the stock, which will cause us to earn or lose money if we buy the stock. One important problem arises naturally. How can we model this stochastic perturbation? Can we make calculations to get the solution of the stock's price $P_t^1$, as we do in the case of the bond's price $P_t^0$? The answer is positive, usually a
continuous stochastic perturbation will be modeled by a stochastic integral 
\[ \int_0^t \sigma_s dw_s, \] where \( w_t, t \geq 0 \) is the so-called Brownian Motion process (BM), or the Wiener process. The 1—dimensional BM \( w_t, t \geq 0 \) has the following nice properties: 1) (Independent increment property). It has an independent increment property, that is, for any \( 0 < t_1 < \cdots < t_n \) the system \( \{w_0, w_{t_1} - w_0, w_{t_2} - w_{t_1}, \cdots, w_{t_n} - w_{t_{n-1}} \} \) is an independent system. Or say, the increments, which happen in disjoint time intervals, occurred independently. 2) (Normal distribution property). Each increment is Normally distributed. That is, for any \( 0 < s < t \) the increment \( W_t - W_s \) on this time interval is a normal random variable with mean \( m \), and variance \( \sigma^2(t - s) \). We write this as \( W_t - W_s \sim N(m, \sigma^2(t - s)) \). 3) (Stationary distribution property). The probability distribution of each increment only depends on the length of the time interval, and it does not depend on the starting point of the time interval. That is, the \( m \) and \( \sigma^2 \) appearing in property 3) are constants. 4) (Continuous trajectory property). Its trajectory is continuous. That is BM \( w_t, t \geq 0 \) is continuous in \( t \).

Since the simplest or say, the most basic continuous stochastic perturbation, intuitively will have the above four properties, the modeling of the general continuous stochastic perturbation by a stochastic integral with respect to this basic BM \( W_t, t \geq 0 \) is quite natural. However, the 1—dimensional BM also has some strange property: Even though it is continuous in \( t \), it is nowhere differentiable in \( t \). So we cannot define the stochastic integral \( \int_0^t \sigma_s(\omega)dw_\omega(\omega) \) for each given \( \omega \). That is why K. Ito (1942) invented a completely new way to define this stochastic integral.

Our first task in this book is to introduce the Ito stochastic integral and discuss its properties for later applications.

After we have understood the stochastic integral \( \int_0^t \sigma_s(\omega)dw_\omega(\omega) \) we can study the following general stochastic differential equation (SDE):

\[ x_t = x_0 + \int_0^t \beta(s,x_s)ds + \int_0^t \sigma(s,x_s)dw_s, \quad t \geq 0, \]

or equivalently, we write

\[ dx_t = \beta(t,x_t)dt + \sigma(t,x_t)dw_t, \quad x_0 = x_0, \quad t \geq 0. \quad (1) \]

Returning to the stock’s price equation, we naturally consider it as the following SDE:

\[ dP_t^1 = P_t^1(b_t dt + \sigma_t dw_t), \quad P_0^1 = P_0^1, \quad t \in [0,T]. \quad (2) \]

Comparing this to the solution of \( P_0^0 \), one naturally asks could the solution of this SDE be \( P_t^1 = P_0^1 \exp[\int_0^t b_s ds + \int_0^t \sigma_s dw_s] \)? To check this guess, obviously if we can have a differential rule to perform differentiation on \( P_t^1 \exp x_t \), where \( x_t = \int_0^t b_s ds + \int_0^t \sigma_s dw_s \), then we can make the check. Or more generally, if we have an \( f(x) \in C^2(\mathbb{R}) \) and \( dx_t = b_t dt + \sigma_t dw_t \), can we have \( df'(x_t) = f'(x_t)dx_t = f'(x_t)(b_t dt + \sigma_t dw_t) \)?
If as in the deterministic case, this differential rule holds true, then we immediately see that \( P^1_t = P^1_0 \exp[\int_0^t b_s ds + \int_0^t \sigma_s dw_s] \) satisfies (2). Unfortunately, such a differential rule is not true. K. Ito (1942) has found that it should obey another differential rule - the so-called Ito's formula:

\[
\frac{df'(x_t)}{dt} = f'(x_t) dx_t + \frac{1}{2} f''(x_t) |\sigma_t|^2 dt.
\]

By this rule one easily checks that

\[
P_t = P^1_0 \exp[\int_0^t b_s ds + \int_0^t \sigma_s dw_s]
\]

is a solution of (2), and \( \tilde{P}^1_t = P^1_0 \exp[\int_0^t b_s ds + \int_0^t \sigma_s dw_s] \) actually satisfies another SDE:

\[
d\tilde{P}^1_t = \tilde{P}^1_t \left[ \int_0^t b_s ds + \int_0^t \sigma_s dw_s + \frac{1}{2} \int_0^t |\sigma_s|^2 ds \right],
\]

and \( \tilde{P}^1_t = P^1_t, \forall t \in [0, T] \).

Our second task in this book is to establish the Ito formula and discuss its applications: solving SDE and solving other problems.

However, even if we have a powerful Ito formula (or say, Ito's differential rule) in our hand, we still need to discuss how to solve the general SDE, because usually, the form of the solution of SDE is not easy to guess. Moreover, for solutions of SDE, we actually meet a more complicated and hence also a more interesting case. Consider some physical quantity \( x_t \) determined by dynamics. If this dynamics is deterministic, that is, it is not disturbed by any random noises, say such that

\[
dx_t = b(t, x_t) dt, x_0 = x_0, t \geq 0;
\]

then solving this ODE we immediately get this quantity \( x_t \). However, if the dynamics are disturbed by some continuous random noise, say such that

\[
dx_t = b(t, x_t) dt + \sigma(t, x_t) dW_t, x_0 = x_0, t \geq 0;
\]

then for the amount \( x_t \), or say, the solution of this SDE, two situations can arise. The first one is, if we think that the random noise - BM \( W_s, s \leq t \), is an input, then after disturbing the dynamics we get an output \( x_t \). This means that the solution \( x_t \) is a functional of the given noise - BM \( W_s, s \leq t \) for each \( t \). We will call such a solution a strong solution. Another situation is that for a given noise we cannot immediately find the solution. However, we can find a random process \( \tilde{x}_t, t \geq 0 \), and maybe another random noise that is also a BM \( \tilde{W}_t, t \geq 0 \), such that \( (\tilde{x}_t, \tilde{W}_t), t \geq 0 \) satisfy the SDE

\[
d\tilde{x}_t = \tilde{b}(t, \tilde{x}_t) dt + \tilde{\sigma}(t, \tilde{x}_t) d\tilde{W}_t, \tilde{x}_0 = x_0, t \geq 0.
\]

In this case we will call \( (\tilde{x}_t, \tilde{W}_t), t \geq 0 \) or simply, \( \tilde{x}_t, t \geq 0 \), a weak solution of the original SDE. Obviously, from the engineering point of view the strong solutions is more realistic and useful. However, since, if the strong solution \( x_t, t \geq 0 \) exists, then all finite dimensional probability distributions of \( (x_t, W_t)_{t \geq 0} \) are the same as that of \( (\tilde{x}_t, \tilde{W}_t)_{t \geq 0} \). So, if we do not know the existence of a strong solution, but we do know the existence of a weak solution \( (\tilde{x}_t, \tilde{W}_t), t \geq 0 \), then from the probability point of view the weak solution can still help us in some sense. Therefore, for solutions of SDEs there are two kinds that need to be considered: strong solutions and weak solutions.

Our third task in this book is to introduce the concepts of solutions and to discuss their existence and uniqueness and the related important theory. (For example, Girsanov's theorem and the martingale representation theorem, the first of which can help us find weak solutions, while the second
one is necessary for finding the solutions of backward SDE and the filtering problem considered later).

Since, actually, in the realistic world we will always meet some jump type stochastic perturbation, in this book we also consider stochastic integrals with respect to a Poisson counting measure (which is generated by a Poisson point process), the Ito formula and SDE for this case. (To find the reason why we consider the Poisson point process and its related integral as a jump type stochastic perturbation see the subsection "The General Model and its Explanation" in Chapter 8 - "Option Pricing in a Financial Market and BSDE").

The first three Chapters are intended to solve the above three tasks. They are the basic foundation of the SDE theory and its applications.

However, interesting and important things for SDE do not only come from the above mentioned three chapters, where they exhibit the following facts: the definition of Ito's stochastic integrals and Ito's differential rule are completely different from the deterministic case, etc. The interesting and important things also come from the following facts:

1) For an ordinary differential equation (ODE) \(dx_t = \tilde{b}(t,x_t)dt, x_0 = x_0, t \geq 0\) if \(\tilde{b}(t,x)\) is only bounded and jointly continuous, then even though the solution exists, is not necessary unique. However, for the SDE (1) in one-dimensional case if \(\tilde{b}(t,x)\) and \(\tilde{\sigma}(t,x)\) are only bounded and jointly Borel-measurable, and \(|\tilde{\sigma}(t,x)|^{-1}\) is also bounded and \(\tilde{\sigma}(t,x)\) is Lipschitz continuous in \(x\), then (1) will have a unique strong solution. (Here "strong" means that \(x_t\) is \(\mathcal{F}_t\)-measurable). This means that adding a non-degenerate stochastic perturbation term into the differential equation, can even improve the nice property of the solution.

2) The stochastic perturbation term has an important practical meaning in some cases and it cannot be discarded. For example, in the investment problem and the option pricing problem from a Financial Market, the investment portfolio actually appears as the coefficient of the stochastic integral in an SDE, where the stochastic integral acts like a stochastic perturbation term.

3) The solutions of SDEs and backward SDEs can help us to explain the solutions of some deterministic partial differential equations (PDEs) with integral terms (the Feynman-Kac formula) and even to guess and find the solution of a PDE, for example, the solution of the PDE for the price of an option can be solved by a solution of a BSDE - the Black-Scholes formula.

4) More and more.

So we have many reasons to study the SDE theory and its applications more deeply and carefully. That is why we have a Chapter that discusses useful tools for SDE, and a Chapter for the solutions of an SDE with non-Lipschitzian coefficients. These are Chapter 4 and 5.

The above concerns the first part of our book, which represents the theory and general background of the SDE.
The second part of our book is about the Applications.

We first provide a short Chapter to help the reader to take a quick look at how to use Stochastic Analysis (the theory in the first part), to solve an SDE.

Then we discuss the estimation problem for a signal process: the so-called filtering problem, where the general linear and non-linear filtering problem for continuous SDE systems and SDE systems with jumps, the Kalman-Bucy filtering equation for continuous systems, and the Zakai equation for non-linear filtering, etc. are also considered.

Since, now, research on mathematical finance, and in particular on the option pricing problem for the financial market has become very popular, we also provide a Chapter that discusses the option pricing problem and backward SDE, where the famous Black-Scholes formulas for a market with or without jumps are derived using a probability and a PDE approach, respectively; and the arbitrage-free market is also discussed. The interesting thing here is that we deal with the mathematical financial problem by using the backward stochastic differential equation (BSDE) technique, which now becomes very powerful when treating many problems in the financial market, in mathematics and in other sciences.

Since deterministic population control has proved to be important and efficient, and the stochastic population control is more realistic, we also provide a Chapter that develops the stochastic population control problem by using the reflecting SDE approach, where the existence, the comparison and the calculation of the population solution and the optimal stochastic population control are established.

Besides, the stochastic Lagrange method for the stochastic optimal control, the non-linear pathwise stochastic optimal control, and the Maximum Principle (that is, the necessary conditions for a stochastic optimal control) are also formulated and developed in specific Chapters, respectively.

For the convenience of the readers three Appendixes are also provided: giving a short review on basic probability theory, space $D$ and Skorohod's metric, and monotone class theorems and the convergence of random processes.

We suggest that the reader studies the book as follows:

For readers who are mainly interested in Applications, the following approach may be considered: Appendix A → Chapter 1 → Chapter 2 → Chapter 3 → Chapter 6 → Any Chapter in The second part "Applications" except Chapter 10, and at any time return to read the related sections in Chapters 4 and 5, or Appendixes B and C, when necessary. However, to read Chapter 10, knowledge of Chapters 4 and 5 and Appendixes B and C are necessary.
Acknowledgement

The author would like to express his sincere thanks to Professor Alan Jeffrey for kindly recommending publication of this book, and for his interest in the book from the very beginning to the very end, and for offering many valuable and important suggestions. Without his help this book would not have been possible.
Abbreviations and Some Explanations

All important statements and results, like, Definitions, Lemmas, Propositions, Theorems, Corollaries, Remarks and Examples are numbered in sequential order throughout the whole book. So, it is easy to find where they are located. For example, Lemma 22 follows Definition 21, and Theorem 394 is just after Remark 393, etc. However, the numbers of equations are arranged independently in each Chapter and each Appendix. For example, (3.25) means equation 25 in Chapter 3, and (C.4) means the equation 4 in Appendix C.

The following abbreviations are frequently used in this book.

- **a.e.** almost everywhere.
- **a.s.** almost sure.
- **BM** Brownian Motion.
- **BSDE** backward stochastic differential equation.
- **FSDE** forward stochastic differential equation.
- **H-J-B** equation Hamilton-Jacobi-Bellman equation.
- **IDE** integral-differential equation.
- **ODE** ordinary differential equation.
- **PDE** partial differential equation.
- **RCLL** right continuous with left limit.
- **SDE** stochastic differential equation.
- **$P-a.s$** almost sure in probability $P$.
- **$a^+$** $\max\{a, 0\}$.
- **$a^-$** $\max\{-a, 0\}$.
- **$a \lor b$** $\max\{a, b\}$.
- **$a \land b$** $\min\{a, b\}$. 
\(\mu << \nu\) measure \(\mu\) is absolutely continuous with respect to \(\nu\); that is, for any measurable set \(A\), \(\nu(A) = 0\) implies that \(\mu(A) = 0\).

\(\xi_n \to \xi\), a.s. \(\xi_n\) converges to \(\xi\), almost surely; that is, \(\xi_n(\omega) \to \xi(\omega)\) for all \(\omega\) except at the points \(\omega \in A\), where \(P(A) = 0\).

\(\xi_n \to \xi\), in \(P\) \(\xi_n\) converges to \(\xi\) in probability; that is, \(\forall \varepsilon > 0\),

\[
\lim_{n \to \infty} P(\omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon) = 0.
\]

\# \{\cdot\} the numbers of \(\cdot\) counted in the set \{\cdot\}.

\(\sigma(x, s \leq t)\) the smallest \(\sigma\)-field, which makes all \(x_s, s \leq t\) measurable.

\(E[\xi|\eta]\) \(E[\xi|\sigma(\eta)]\). It means the conditional expectation of \(\xi\) given \(\sigma(\eta)\).

The following notations can be found on the corresponding pages. For example, \(\mathfrak{F}, 4,387\) means that notation \(\mathfrak{F}\) can be found in page 4 and page 387.

\(\Omega, 4,387\)
\(\mathfrak{F}, 4,387\)
\(\mathfrak{F}_t, 4\)
\(\mathcal{F}_p, 34\)
\(\mathcal{F}^1_p, 34\)
\(\mathcal{F}^2_p, 34\)
\(\mathcal{F}^{2,\text{loc}}_p, 34\)
\(\mathcal{E}^2_p (R^{d\otimes d_2}), 212\)
\(\mathcal{L}_0^2, 43\)
\(\mathcal{L}^2, 43\)
\(\mathcal{L}^2, 43\)
\(\mathcal{L}^2_{\text{loc}}, 48\)
\(\mathcal{L}^2_{M}, 50\)
\(\mathcal{L}^4_{\text{loc}}, 51\)
\(L^2_\Omega (R^{d\otimes d_1}), 212\)
\(L^2_\mathcal{F} (R^{d\otimes d}), 212\)
\(L^1(\Omega, \mathfrak{F}, P). 12\)
\(\mathcal{M}^2, 35\)
\(\mathcal{M}^{2,\text{c}}, 37\)
\(\mathcal{M}^{2,\text{loc}}, 35\)
\(\mathcal{M}^2,\text{rdlt}, 37\)
\(\mathcal{M}^2_{\text{T}}, 35\)
\(\mathcal{M}^{2,\text{c}}_{\text{T}}, 37\)
\(\mathcal{M}^2_{\text{T}}\text{loc}, 35\)
\(\mathcal{M}^2_{\text{T}},\text{rdlt}, 37\)
\(\mathcal{O}, 7\)
\(\mathcal{P}, 7\)
\(P, 4,387\)
\(S^2_\mathfrak{F} (R^d), 212\)
Part I

Stochastic Differential Equations with Jumps in $\mathbb{R}^d$
1

Martingale Theory and the Stochastic Integral for Point Processes

A stochastic integral is a kind of integral quite different from the usual deterministic integral. However, its theory has broad and important applications in Science, Mathematics itself, Economic, Finance, and elsewhere. A stochastic integral can be completely characterized by martingale theory. In this chapter we will discuss the elementary martingale theory, which forms the foundation of stochastic analysis and stochastic integral. As a first step we also introduce the stochastic integral with respect to a Point process.

1.1 Concept of a Martingale.

In some sense the martingale conception can be explained by a fair game. Let us interprete it as follows:

In a game suppose that a person at the present time $s$ has wealth $x_s$ for the game, and at the future time $t$ he will have the wealth $x_t$. The expected money for this person at the future time $t$ is naturally expressed as $E[x_t|\mathcal{F}_s]$, where $E[\cdot]$ means the expectation value of ·, $\mathcal{F}_s$ means the information up to time $s$, which is known by the gambler, and $E[\cdot|\mathcal{F}_s]$ is the conditional expectation value of · under given $\mathcal{F}_s$. Obviously, if the game is fair, then it should be

$$E[x_t|\mathcal{F}_s] = x_s, \forall t \geq s.$$

This is exactly the definition of a martingale for a random process $x_t, t \geq 0$. Let us make it more explicit for later development.
Let \((\Omega, \mathcal{F}, P)\) be a probability space, \(\{\mathcal{F}_t\}_{t \geq 0}\) be an information family (in Mathematics, we call it a \(\sigma\)-algebra family or a \(\sigma\)-field family, see Appendix A), which satisfies the so-called “Usual Conditions”:

1. \(\mathcal{F}_s \subset \mathcal{F}_t\), as \(0 \leq s \leq t\);  
2. \(\mathcal{F}_{t^+} := \bigcap_{h > 0} \mathcal{F}_{t+h}\).

Here condition (i) means that the information increases with time, and condition (ii) that the information is right continuous, or say, \(\mathcal{F}_{t+h} \downarrow \mathcal{F}_t\), as \(h \downarrow 0\). In this case we call \(\{\mathcal{F}_t\}_{t \geq 0}\) a \(\sigma\)-field filtration.

**Definition 1** A real random process \(\{x_t\}_{t \geq 0}\) is called a martingale (supermartingale, submartingale) with respect to \(\{\mathcal{F}_t\}_{t \geq 0}\), or \(\{x_t, \mathcal{F}_t\}_{t \geq 0}\) is a martingale (supermartingale, submartingale), if

1. \(x_t\) is integrable for each \(t \geq 0\); that is, \(E|x_t| < \infty, \forall t \geq 0\);
2. \(x_t\) is \(\mathcal{F}_t\)-adapted; that is, for each \(t \geq 0\), \(x_t\) is \(\mathcal{F}_t\)-measurable;  
3. \(E(x_t|\mathcal{F}_s) = x_s\), (respectively, \(\leq, \geq\)), a.s. \(\forall 0 \leq s \leq t\).

For the random process \(\{x_t\}_{t \in [0, T]}\) and the random process \(\{x_n\}_{n=1}^{\infty}\) with discrete time similar definitions can be given.

**Example 2** If \(\{x_t\}_{t \geq 0}\) is a random process with independent increments; that is, \(\forall 0 < t_1 < t_2 < \cdots < t_n\), the family of random variables

\[\{x_0, x_1 - x_0, x_2 - x_1, \ldots, x_n - x_{n-1}\}\]

is independent, and the increment \(x_t - x_s, \forall t > s\), is integrable and with non-negative expectation, moreover, \(x_0\) is also integrable, then \(\{x_t\}_{t \geq 0}\) is a submartingale with respect to \(\{\mathcal{F}_t^x\}_{t \geq 0}\), where \(\mathcal{F}_t^x = \sigma(x_s, s \leq t)\), which is a \(\sigma\)-field generated by \(\{x_s, s \leq t\}\) (that is, the smallest \(\sigma\)-field which makes all \(x_s, s \leq t\) measurable) and makes a completion.

In fact, by independent and non-negative increments,

\[0 \leq E(x_t - x_s) = E[(x_t - x_s)|\mathcal{F}_s], \forall t \geq s.\]

Hence the conclusion is reached.

**Example 3** If \(\{x_t\}_{t \geq 0}\) is a submartingale, let \(y_t := x_t \vee 0 = \max(x_t, 0)\), then \(\{y_t\}_{t \geq 0}\) is still a submartingale.

In fact, since \(f(x) = x \vee 0\) is a convex function, hence by Jensen’s inequality for the conditional expectation

\[E[x_t \vee 0|\mathcal{F}_s] \geq E[x_t|\mathcal{F}_s] \vee E[0|\mathcal{F}_s] \geq x_s \vee 0, \forall t \geq s.\]

So the conclusion is true.

**Example 4** If \(\{x_t\}_{t \geq 0}\) is a martingale, then \(\{|x_t|\}_{t \geq 0}\) is a submartingale.

In fact, by Jensen’s inequality

\[E[|x_t| |\mathcal{F}_s] \geq |E[x_t|\mathcal{F}_s]| = |x_s|, \forall t \geq s.\]

Thus the conclusion is deduced.

Martingales, submartingales and supermartingales have many important and useful properties, which make them become powerful tools in dealing with many theoretical and practical problems in Science, Finance and elsewhere. Among them the martingale inequalities, the limit theorems, and the
1.2 Stopping Times. Predictable Process

**Definition 5** A random variable \( \tau(\omega) \in [0, \infty] \) is called a \( \mathcal{F}_t \)-stopping time, or simply, a stopping time, if for any \( (\infty > t \geq 0) \), \( \{ \tau(\omega) \leq t \} \in \mathcal{F}_t \).

The intuitive interpretation of a stopping time is as follows: If a gambler has a right to stop his gamble at any time \( \tau(\omega) \), he would of course like to choose the best time to stop. Suppose he stops his game before time \( t \), i.e. he likes to make \( \tau(\omega) < t \), then the maximum information he can get about his decision is only the information up to \( t \), i.e. \( \{ \tau(\omega) \leq t \} \in \mathcal{F}_t \). The trivial example for a stopping time is \( \tau(\omega) = t \), \( \forall \omega \in \Omega \). That is to say, any constant time \( t \) actually is a stopping time.

For a discrete random variable \( \tau(\omega) \in \{0, 1, 2, \cdots, \infty\} \) the definition can be reduced to that \( \tau(\omega) \) is a stopping time, if for any \( n \in \{0, 1, 2, \cdots\} \), \( \{\tau(\omega) = n\} \in \mathcal{F}_n \), since \( \{\tau(\omega) = n\} = \{\tau(\omega) \leq n\} - \{\tau(\omega) \leq n - 1\} \), and \( \{\tau(\omega) \leq n\} = \bigcup_{k=1}^n \{\tau(\omega) = k\} \). The following examples of stopping time are useful later.

**Example 6** Let \( B \) be a Borel set in \( \mathbb{R}^1 \) and \( \{x_n\}_{n=1}^{\infty} \) be a sequence of real \( \mathcal{F}_t \)-adapted random variables. Define the first hitting time \( \tau_B(\omega) \) to the set \( B \) (i.e. the first time that \( \{x_n\}_{n=1}^{\infty} \) hits \( B \)) by
\[
\tau_B(\omega) = \inf \{n : x_n(\omega) \in B\}.
\]
Then \( \tau_B(\omega) \) is a discrete stopping time.

In fact,
\[
\{\tau_B(\omega) = n\} = \bigcap_{k=1}^{n-1} \{x_k \in B^c\} \cap \{x_n \in B\} \in \mathcal{F}_n.
\]

For a general random process with continuous time parameter \( t \) we have the following similar example.

**Example 7** Let \( x_t \) be a \( d \)-dimensional right continuous \( \mathcal{F}_t \)-adapted process and let \( A \) be an open set in \( \mathbb{R}^d \). Denote the first hitting time \( \sigma_A(\omega) \) to \( A \) by
\[
\sigma_A(\omega) = \inf \{t > 0 : x_t(\omega) \in A\}.
\]
Then \( \sigma_A(\omega) \) is a stopping time.

In fact, by the open set property and the right continuity of \( x_t \) one has that
\[
\{\sigma_A(\omega) \leq t\} = \bigcap_{n=1}^{\infty} \{\sigma_A(\omega) < t + \frac{1}{n}\}
= \bigcap_{n=1}^{\infty} \bigcup_{r \in Q, r < t+1/n} \{x_r(\omega) \in A\} \in \mathcal{F}_{t+0} = \mathcal{F}_t,
\]
where \( Q \) is the set of all rational numbers.

The following properties of general stopping times will be useful later.

**Lemma 8** \( \tau(\omega) \) is a stopping time, if and only if \( \{ \tau(\omega) < t \} \in \mathcal{F}_t, \forall t. \)

**Proof.** \( \Rightarrow \): \( \{ \tau(\omega) < t \} = \bigcup_{n=1}^{\infty} \{ \tau(\omega) \leq t - \frac{1}{n} \} \in \mathcal{F}_t. \)

\( \Leftarrow \): \( \{ \tau(\omega) \leq t \} = \bigcap_{n=1}^{\infty} \{ \tau(\omega) < t + \frac{1}{n} \} \in \mathcal{F}_{t+0} = \mathcal{F}_t. \)

**Lemma 9** Let \( \sigma, \tau, \sigma_n, n = 1, 2, \cdots \) be stopping times. Then

(i) \( \sigma \wedge \tau, \sigma \vee \tau, \)

(ii) \( \sigma = \lim_{n \to \infty} \sigma_n, \) when \( \sigma_n \uparrow \) or \( \sigma_n \downarrow, \)

are all stopping times.

**Proof.** (i): \( \{ \sigma \wedge \tau \leq t \} = \{ \sigma \leq t \} \cup \{ \tau \leq t \} \in \mathcal{F}_t, \)

\( \{ \sigma \vee \tau \leq t \} = \{ \sigma \leq t \} \cap \{ \tau \leq t \} \in \mathcal{F}_t. \)

(ii): If \( \sigma_n \uparrow \sigma, \) then

\( \{ \sigma \leq t \} = \bigcap_{n=1}^{\infty} \{ \sigma_n \leq t \} \in \mathcal{F}_t. \)

If \( \sigma_n \downarrow \sigma, \) then

\( \{ \sigma < t \} = \bigcup_{n=1}^{\infty} \{ \sigma_n < t \} \in \mathcal{F}_t. \)

By Lemma 8 \( \sigma \) is a stopping time. \( \blacksquare \)

Now let us introduce a \( \sigma^- \) field which describes the information obtained up to stopping time \( \tau. \) Set

\( \mathcal{F}_\tau = \{ A \in \mathcal{F}_\infty : \forall t \in [0, \infty), A \cap \{ \tau(\omega) \leq t \} \in \mathcal{F}_t \}, \)

where we naturally define that \( \mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t, \) i.e. the smallest \( \sigma^- \)field including all \( \mathcal{F}_t, t \in [0, \infty). \) Obviously, \( \mathcal{F}_\tau \) is a \( \sigma^- \) algebra, and if \( \tau(\omega) \equiv t, \) then \( \mathcal{F}_\tau = \mathcal{F}_t. \)

**Proposition 10** Let \( \sigma, \tau, \sigma_n, n = 1, 2, \cdots \) be stopping times.

(1) If \( \sigma(\omega) \leq \tau(\omega), \forall \omega, \) then \( \mathcal{F}_\sigma \subseteq \mathcal{F}_\tau. \)

(2) If \( \sigma_n(\omega) \downarrow \sigma(\omega), \forall \omega, \) then \( \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n} = \mathcal{F}_{\sigma}. \)

(3) \( \sigma \in \mathcal{F}_{\sigma}. \) (We use \( f \in \mathcal{F}_{\sigma} \) to mean that \( f \) is \( \mathcal{F}_{\sigma} \)-measurable).

**Proof.** (1): \( A \cap \{ \tau \leq t \} = (A \cap \{ \sigma \leq t \}) \cap \{ \tau \leq t \} \in \mathcal{F}_t. \)

(2): By (1) \( \mathcal{F}_\sigma \subseteq \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n}. \) Conversely, if \( A \in \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n}, \) then

\( A \cap \{ \sigma_n < t \} = \bigcup_{k=1}^{\infty} (A \cap \{ \sigma_n \leq t - \frac{1}{k} \}) \in \mathcal{F}_t, \forall t \geq 0, \forall n. \)

Hence \( A \cap \{ \sigma < t \} = \bigcup_{k=1}^{\infty} (A \cap \{ \sigma_n < t \}) \in \mathcal{F}_t, \)

\( A \cap \{ \sigma \leq t \} = \bigcap_{k=1}^{\infty} (A \cap \{ \sigma < t + \frac{1}{k} \}) \in \mathcal{F}_{t+0} = \mathcal{F}_t, \) i.e. \( A \in \mathcal{F}_\sigma. \)

(3): For any constant \( 0 \leq c < 1 \) one has that \( \{ \sigma \leq c \} \cap \{ \sigma \leq t \} \in \mathcal{F}_{t \wedge c} \subset \mathcal{F}_t, \) so \( \{ \sigma \leq c \} \in \mathcal{F}_{\sigma}. \)

It is natural to ask that if \( \{ x_t \}_{t \geq 0} \) is \( \mathcal{F}_t \)-adapted, and \( \sigma \) is a stopping time, is it true that \( x_\sigma \in \mathcal{F}_{\sigma}? \) Generally speaking, it is not true. However, if \( \{ x_t \}_{t \geq 0} \) is a progressive measurable process, then it is correct. Let us introduce such a related concept.

**Definition 11** An \( R^d \)-valued process \( \{ x_t \}_{t \geq 0} \) is called measurable (respectively, progressive measurable), if the mapping

\( (t, \omega) \in [0, \infty) \times \Omega \to x_t(\omega) \in R^d \)
(respectively, for each \( t \geq 0 \), \((s, \omega) \in [0, t] \times \Omega \rightarrow x_t(\omega) \in R^d\)) is \( \mathcal{B}([0, \infty)) \times \mathcal{F}(R^d) - \)measurable
(respectively, \( \mathcal{B}([0, t]) \times \mathcal{F}_t / \mathcal{B}(R^d) - \)measurable);
that is, \( \{(t, \omega) : x_t(\omega) \in B \in \mathcal{B}([0, \infty)) \times \mathcal{F}, \forall B \in \mathcal{B}(R^d)\) (respectively, \( \{(s, \omega) : s \in [0, t], x_s(\omega) \in B \} \in \mathcal{B}([0, t]) \times \mathcal{F}_t, \forall B \in \mathcal{B}(R^d)\).

Let us introduce two useful \( \sigma \)-algebras as follows: Denote by \( \mathcal{P} \) (respectively, \( \mathcal{O} \)) as the smallest \( \sigma \)-algebra on \([0, \infty) \times \Omega \) such that all left-continuous (respectively, right-continuous) \( \mathcal{F}_t \)-adapted processes
\[ y_t(\omega) : [0, \infty) \times \Omega \rightarrow x_t(\omega) \in R^d \]
are measurable. \( \mathcal{P} \) (respectively, \( \mathcal{O} \)) is called the predictable (respectively, optional) \( \sigma \)-algebra. Thus, the following definition is natural.

**Definition 12** A process \( \{x_t \}_{t \geq 0} \) is called predictable (optional), if the mapping
\[ (t, \omega) \in [0, \infty) \times \Omega \rightarrow x_t(\omega) \in R^d \]
is \( \mathcal{P} / \mathcal{B}(R^d) - \)measurable (respectively \( \mathcal{O} / \mathcal{B}(R^d) - \)measurable).

Let us use the notation \( f \in \mathcal{P} \) to mean that \( f \) is \( \mathcal{P} \)-measurable; etc. It is easily seen that the following relations hold:
\[ f \in \mathcal{P} \Rightarrow f \in \mathcal{O} \Rightarrow f \text{ is progressive measurable} \Rightarrow f \text{ is measurable and } \mathcal{F}_t \text{-adapted.} \]

We only need to show the first two implications. The last one is obvious.

Assume that \( \{x_t \}_{t \geq 0} \) is left-continuous, let \( x_k^n = x_{\frac{k}{2^n}}, k = 0, 1, \ldots; n = 1, 2, \ldots \). Then obviously, \( x_k^n \) is right-continuous, and by the left-continuity of \( x_t \), \( x_k^n(\omega) \rightarrow x_t(\omega), \text{as } n \rightarrow \infty, \forall t, \forall \omega \). So \( \{x_t \}_{t \geq 0} \in \mathcal{O} \).

From this one sees that \( \mathcal{P} \subset \mathcal{O} \). Let us show that \( \{x_t \}_{t \geq 0} \in \mathcal{O} \) implies that \( \{x_t \}_{t \geq 0} \) is progressive measurable. For this for each given \( t \geq 0 \) we show that \( \{x_s \}_{s \geq t} \) restricted on \((s, \omega) \in [0, t] \times \Omega \) is \( \mathcal{B}([0, t]) \times \mathcal{F}_t - \)measurable.

In fact, without loss of generality we may assume that \( \{x_t \}_{t \geq 0} \) is right-continuous. Now for each given \( t \geq 0 \), let \( x_k^n = x_{\frac{k}{2^n}}, k = 0, 1, \ldots, 2^n - 1; n = 1, 2, \ldots \). Then obviously, \( \{x_k^n \}_s \in \mathcal{B}([0, t]) \times \mathcal{F}_t - \)measurable, so is \( \{x_s \}_s \in [0, t], \text{ since by the right continuity of } x_t \) we have that as \( n \rightarrow \infty, x_k^n(\omega) \rightarrow x_s(\omega), \forall s \in [0, t], \forall \omega \).

Let us show the following

**Theorem 13** If \( \{x_t \}_{t \geq 0} \) is a \( R^d \)-valued progressive measurable process, then for each stopping time \( \sigma \), \( Z_{\sigma I_{\sigma < \infty}} \) is \( \mathcal{F}_\sigma - \)measurable.

We will use the composition of measurable maps to show this theorem. For this we need the following lemma.

**Lemma 14** If \( f_i \) is a measurable mapping from \((\Omega, \mathcal{F})\) to \((\Omega'_i, \mathcal{F}'_i), i = 1, 2, \ldots \); then
\[ f(\omega) = (f_1(\omega), f_2(\omega), \ldots) \]
is a measurable mapping from \((\Omega, \mathcal{F})\) to \((\Omega'_1 \times \Omega'_2 \times \cdots, \mathcal{F}'_1 \times \mathcal{F}'_2 \times \cdots)\).
In fact, for any $B_i \in \mathcal{F}'_i, i = 1, 2, \ldots, f^{-1}(B_1 \times B_2 \times \cdots) = \cap_{i=1}^{\infty} f_i^{-1}(B_i) \in \mathcal{F}$. So $f^{-1}(\mathcal{F}'_1 \times \mathcal{F}'_2 \times \cdots) \subset \mathcal{F}$.

Now let us prove Theorem 13.

**Proof.** Let $\Omega_\sigma = \{ \sigma < \infty \}$. We need to show that $x_\sigma$ is a measurable mapping from $(\Omega_\sigma, \mathcal{F}_\sigma)$ to $(R^d, \mathcal{B}(R^d))$. For any given $t \geq 0$ by Proposition 10 $\Omega_\sigma \in \mathcal{F}_\sigma$. So by the definition of $\mathcal{F}_\sigma$, $\sigma$ is a measurable mapping from $\{ \sigma < t \}$ to $(R^d, \mathcal{B}(R^d))$. Hence $x_\sigma(\omega) = g_2 \circ g_1(\omega)$ is a measurable mapping from $\{ \sigma < t \} \times \Omega, \mathcal{B}(R^d)$ to $(R^d, \mathcal{B}(R^d))$. This shows that for any $B \in \mathcal{B}(R^d)$, $\{ x_\sigma I_{\sigma < \infty} \in B \} \in \mathcal{F}_\sigma$. Since $t \geq 0$ is arbitrary by definition $\{ x_\sigma I_{\sigma < \infty} \in B \} \in \mathcal{F}_\sigma$. ■

### 1.3 Martingales with Discrete Time

First we will show the Doob’s stopping theorem (or called Doob’s optional sampling theorem) for bounded stopping times.

**Theorem 15** Let $\{ x_n \}_{n=0,1,2,\ldots}$ be a martingale (supermartingale, submartingale), $\sigma \leq \tau$ be two bounded stopping times. Then $\{ x_n \}_{n=0,1,2,\ldots}$ is a strong martingale (respectively, strong supermartingale, strong submartingale), i.e.

$$E[x_\tau | \mathcal{F}_\sigma] = x_\sigma \text{ (respectively,} \leq, \geq, \text{) a.s.}$$

**Proof.** We only prove the conclusion for the case of submartingale. By assumption there exists a natural number $0 \leq n_0$ such that $\tau \leq n_0$. So $|x_{\tau}| \leq \max\{|x_n|, n = 0, 1, 2, \ldots, n_0\} \leq \sum_{n=0}^{n_0} |x_n|$. So $E|x_{\tau}| < \infty$. By the same manner $E| x_\sigma | < \infty$. Note that by the definition of a stopping time and $\mathcal{F}_\sigma$ for $A \in \mathcal{F}_\sigma$ and $0 \leq n \leq n_0$

$$A \cap \{ \sigma = n \} \cap \{ \tau > n \} \in \mathcal{F}_n.$$

Now suppose $\tau - \sigma \leq 1$ in addition. Then by the definition of a submartingale

$$\int_A (x_\sigma - x_{\tau}) dP = \sum_{n=0}^{n_0} \int_{A \cap (\sigma = n) \cap (\tau > n)} (x_n - x_{n+1}) dP \leq 0.$$  

In the general case set $T_n = \tau \wedge (\sigma + n), n = 1, 2, \ldots, n_0$. Then all $T_n$ are stopping times, and

$$\sigma \leq T_1 \leq T_2 \leq \cdots \leq T_{n_0} = \tau, \quad T_1 - \sigma \leq 1, \quad T_{n+1} - T_n \leq 1, \quad n = 1, 2, \ldots, n_0 - 1.$$  

Let $A \in \mathcal{F}_\sigma \subset \mathcal{F}_{T_{n_0}}$. Then by the above conclusion

$$\int_A x_\sigma dP \leq \int_A x_{T_1} dP \leq \cdots \leq \int_A x_\tau dP.$$  

The proof is complete. ■

Now we have the following martingale inequality:

**Theorem 16** Let $\{ x_n \}_{n=0,1,2,\ldots}$ be a submartingale. Then for every $\lambda > 0$ and natural number $N$
1.3 Martingales with Discrete Time

\[ \lambda P(\max_{0 \leq n \leq N} x_n \geq \lambda) \leq E(x_N I_{\max_{0 \leq n \leq N} x_n \geq \lambda}) \leq E(x_N^+) \leq E|x_N|, \]
and
\[ \lambda P(\min_{0 \leq n \leq N} x_n \leq -\lambda) \leq -E x_0 + E(x_N I_{\min_{0 \leq n \leq N} x_n > -\lambda}) \]
\[ \leq E x_0^+ + E(x_N^+) \leq E|x_0| + E|x_N|. \]

**Proof.** Let us use the first hitting time technique and strong submartingale property to show this theorem. Set
\[ \sigma = \min\{n \leq N : x_n \geq \lambda\}; \sigma = N, \text{ if } \{\cdot\} = \emptyset. \]
Then \( \sigma \) is a bounded stopping time. By Theorem 15
\[ E x_N \geq E x_{\sigma} = E x_{\sigma} I_{\max_{0 \leq n \leq N} x_n \geq \lambda} + E x_N I_{\max_{0 \leq n \leq N} x_n < \lambda} \]
\[ \geq \lambda P(\max_{0 \leq n \leq N} x_n \geq \lambda) + E x_N I_{\max_{0 \leq n \leq N} x_n < \lambda}. \]

Transferring the last term to the left hand side, we obtain the first inequality. Now set
\[ \tau = \min\{n \leq N : x_n \leq -\lambda\}; \tau = N, \text{ if } \{\cdot\} = \emptyset. \]
Then
\[ E x_0 \leq E x_{\tau} = E x_{\tau} I_{\min_{0 \leq n \leq N} x_n \leq -\lambda} + E x_N I_{\min_{0 \leq n \leq N} x_n > -\lambda} \]
\[ \leq -\lambda P(\min_{0 \leq n \leq N} x_n \leq -\lambda) + E(x_N I_{\min_{0 \leq n \leq N} x_n > -\lambda}). \]
Thus the second inequality is derived. \( \blacksquare \)

**Corollary 17** 1) Assume that \( \{x_n\}_{n=0,1,...} \) is a real submartingale such that \( E((x_n^+)^p) < \infty, n = 0, 1, \ldots \), for some \( p \geq 1 \). Then for every \( N \), and \( \lambda > 0 \),
\[ P(\max_{0 \leq n \leq N} x_n^+ \geq \lambda) \leq E((x_N^+)^p)/\lambda^p, \]
and if \( p > 1 \),
\[ E(\max_{0 \leq n \leq N} (x_n^+)^p) \leq \left(\frac{p}{p-1}\right)^p E((x_N^+)^p). \]
2) If \( \{x_n\}_{n=0,1,...} \) is a real martingale such that \( E(|x_n|^p) < \infty, n = 0, 1, \ldots \), then the conclusions in 1) hold true for \( x_n^+ \) and \( x_n^- \) replaced by \( |x_n| \) and \( |x_N| \), respectively.

**Proof.** 1): By Example 3 \( \{x_n^+\}_{n=0,1,...} \) is a non-negative submartingale. Using Jensen’s inequality again one has that \( \{(x_n^+)\}_{n=0,1,...} \) is still a non-negative submartingale. Hence the first inequality is obtained from Theorem 16. Now if \( p > 1 \), set \( \xi = \max_{0 \leq n \leq N} (x_n^+) \). then by Theorem 16 again one has that
\[ \lambda P(\xi \geq \lambda) \leq E x_N^+ I_{\xi \geq \lambda}. \]
Hence using Fubini’s theorem and Hölder’s inequality one derives that
\[ E(\xi^p) = E \int_0^\xi p\lambda^{p-1}d\lambda = E \int_0^\xi p\lambda^{p-1}I_{\lambda \leq \xi}d\lambda = p \int_0^\infty \lambda^{p-1}P(\xi \geq \lambda)d\lambda \]
\[ \leq p \int_0^\infty \lambda^{p-2}E(x_N^+ I_{\xi \geq \lambda})d\lambda = p \frac{1}{p-1}E(\xi^p)^{\frac{1}{p}}. \]
Now if \( E(\xi^p) = 0 \), then the second inequality is trivial. If \( E(\xi^p) > 0 \), dividing both sides by \( [E(\xi^p)]^{(p-1)/p} \), the second inequality is also obtained.
2): If \( \{x_n\}_{n=0,1,...} \) is a real martingale, then by Jensen’s inequality we have that \( \{|x_n|\}_{n=0,1,...} \) is a submartingale, and \( |x_n|^+ = |x_n| \). So by 1) the conclusions are derived in this case. \( \blacksquare \)
1. Martingale Theory and the Stochastic Integral for Point Processes

In the following we will show the upcrossing inequality for a submartingale, which is the basis for proving the important limit property of a submartingale. First we introduce some notations.

For a real \( \mathcal{F}_t \)-adapted process \( \{x_n\}_{n=0,1,...} \) and an interval \([a, b]\), where \( b > a \), let
\[
\tau_1 = \min\{n \geq 0 : x_n \leq a\},
\tau_2 = \min\{n \geq \tau_1 : x_n \geq b\},
\ldots,
\tau_{2n+1} = \min\{n \geq \tau_{2n} : x_n \leq a\},
\tau_{2n+2} = \min\{n \geq \tau_{2n+1} : x_n \geq b\},
\ldots;
\]
where we recall that \( \min \phi = +\infty \). Then \( \{\tau_n\} \) is an increasing sequence of stopping times. In fact, \( \forall k \geq 0, \)
\[
\{\tau_1 = k\} = \{x_0 > a, x_1 > a, \ldots, x_{k-1} > a, x_k \leq a\} \in \mathcal{F}_k;
\{\tau_{2k} = k\} = \bigcup_{j=0}^{k-1} \{\tau_1 = j, \tau_2 = k\}
= \bigcup_{j=0}^{k-1} \{\tau_1 = j, x_j \leq a, x_{j+1} < b, \ldots, x_{k-1} < b, x_k \geq b\} \in \mathcal{F}_k;
\{\tau_{2k+1} = k\} = \bigcup_{j=1}^{k-1} \{\tau_2 = j, \tau_3 = k\}
= \bigcup_{j=1}^{k-1} \{\tau_2 = j, x_j \geq b, x_{j+1} > a, \ldots, x_{k-1} > a, x_k \leq a\} \in \mathcal{F}_k.
\]
Hence \( \tau_1, \tau_2, \) and \( \tau_3 \) are stopping times. The proofs for the rest are similar.

Now set
\[
U^b_a[x(.), N](\omega) = \max\{k \geq 1 : \tau_{2k}(\omega) \leq N\},
D^b_a[x(.), N](\omega) = \max\{k \geq 1 : \tau_{2k-1}(\omega) \leq N\}.
\]
Obviously the first one is the number connected to the upcrossing of \( \{x_n\}_{n=0}^N \) for the interval \([a, b]\), and the second one is the number connected to the downcrossing of \( \{x_n\}_{n=0}^N \) for the interval \([a, b]\).

**Theorem 18** If \( \{x_n\}_{n=0}^\infty \) is a submartingale, then for each \( N \geq 1, n \geq 0 \) and \( a < b \)
\[
EU^b_a[x(.), N] \leq \frac{1}{b-a} \left( E(\{x_N - a\}^+ - \{x_0 - a\}^+) \right),
\]
\[
P(U^b_a(x(.), N) \geq n) \leq \frac{1}{b-a} E(\{x_N - a\}^+ U^b_a[x(.), N]=n],
\]
\[
ED^b_a(x(.), N] \leq \frac{1}{b-a} E(x_N - b)^+,
\]
\[
P(D^b_a[x(.), N] \geq n + 1) \leq \frac{1}{b-a} E(x_N - b)^+ D^b_a[x(.), N]=n].
\]

**Proof.** For a submartingale \( \{x_n\}_{n=0}^\infty \) by Example 3 one sees that \( \{y_n\}_{n=0}^\infty = \{(x_n - a)^{+}\}_{n=0}^\infty \) is a non-negative submartingale. Clearly \( U^{b-a}_0[y(.), N](\omega) = U^b_a[x(.), N](\omega) \). Again define \( \tau_1, \tau_2, \ldots \) as above, but with \( x, a, \) and \( b \) replaced by \( y, 0, \) and \( b - a \) respectively. Then if \( 2k > N \)
\[
E(y_N - y_0) = E \sum_{n=0}^{2k}(y_{r_n \wedge N} - y_{r_{n-1} \wedge N}) = E \sum_{n=1}^{k}(y_{r_{2n} \wedge N} - y_{r_{2n-1} \wedge N})
+ \sum_{n=0}^{k-1} E(y_{r_{2n} \wedge N} - y_{r_{2n+1} \wedge N}) \geq (b - a) E U^{b-a}_0[y(.), N],
\]
where we have used the fact that \( \{y_n\}_{n=0}^\infty \) is a submartingale, and hence a strong submartingale (Theorem 15), so \( E(y_{r_{2n+1} \wedge N} - y_{r_{2n} \wedge N}) \geq 0; \) and \( y_0 \geq 0, \) \( \forall n. \) The first inequality is proved. Now observe that
\[
0 \geq E(y_{r_{2n} \wedge N} - y_{r_{2n+1} \wedge N})
= E[(y_{r_{2n} \wedge N} - y_{r_{2n+1} \wedge N})(I_{r_{2n} \leq N < r_{2n+1} + I_{r_{2n+1} \leq N}})]
\]
1.3 Martingales with Discrete Time

\[ E[(b - a - yN)I_{\tau_{2n} \leq N < \tau_{2n+1}} + (b - a)I_{\tau_{2n+1} \leq N}] = E(b - a)I_{\tau_{2n} \leq N} - E(yNI_{\tau_{2n} \leq N < \tau_{2n+1}}). \]

Since \( \{U_0^{b-a}[y(.)], N \geq n\} = \{N \geq \tau_{2n}\} \) and

\[ \{\tau_{2n} \leq N < \tau_{2n} + 1\} \subset \{\tau_{2n} \leq N < \tau_{2n+2}\} = \{U_0^{b-a}[y(.)], N = n\}. \]

Hence we find that \( E(yNI_{U_0^{b-a}[x(.)],N}=n} \geq (b - a)P(U_0^{b-a}[y(.)], N) \geq n\).

For the downcrossing inequality we have to discuss \( \{x_n\}_{n=0}^{\infty} \) itself directly, since \( \{x_n \land 0\}_{n=0}^{\infty} \) is not a submartingale. Let us set \( y_n = x_n - b \). Then \( \{y_n\}_{n=0}^{\infty} \) is still a submartingale, and

\[ D_{(b-a)}^0[y(.)], N](\omega) = D_a^b[x(.)], N](\omega). \]

Again define \( \tau_1, \tau_2, \ldots \) as above but with \( x, a, \) and \( b \) replaced by \( y, -(b-a), \) and \( 0 \) respectively. We will now use another method to show the last two inequalities. First, for the fourth inequality we have that as \( n \geq 1 \)

\[ 0 \geq E(y_{\tau_{2n} \land N} - y_{\tau_{2n+1} \land N}) \]

\[ = E[(0 - (x_N - b))I_{\tau_{2n} \leq N < \tau_{2n+1} + (b - a)I_{\tau_{2n+1} \leq N}]. \]

Since

\[ \{D_a^b[x(.)], N\} \geq n + 1\} = \{D_{(b-a)}^0[y(.)], N\} \geq n + 1\} \]

\[ = \{N \geq \tau_{2n+2}\} \subset \{N \geq \tau_{2n+1}\} \]

and \( \{\tau_{2n} \leq N < \tau_{2n+1}\} \subset \{\tau_{2n} \leq N < \tau_{2n+2}\} = \{D_a^b[x(.)], N\} = n\}. \]

Hence it follows that

\[ E(x_N - b)^+I_D[\tau,x(.)], N=0} \geq (b - a)P(D_a^b[x(.)], N) \geq n + 1\}. \]

The fourth inequality holds. Now taking the summation for \( n \geq 0 \) it yields

\[ E(x_N - b)^+ \geq (b - a)\sum_{n=0}^{\infty} P(D_a^b[x(.)], N) \geq n + 1\}

\[ = (b - a)\sum_{n=0}^{\infty} nP(D_a^b[x(.)], N) = n) = (b - a)ED_a^b[x(.)], N]. \]

The third inequality is also established. ■

**Corollary 19** If \( \{x_n\}_{n=0}^{\infty} \) is a supermartingale, then for each \( N \geq 1, n \geq 0 \) and \( a < b \)

\[ E[u_a^b[x(.)], N] \leq \frac{1}{b-a}E[(x_N - a)^-], \]

\[ P(U_a^b[x(.)], N) \geq n + 1\} \leq \frac{1}{b-a}E[(x_N - a)^-I_{U_a^b[x(.)], N} = n\}]. \]

\[ ED_a^b[x(.)], N \leq \frac{1}{b-a}E[(x_N - b)^- - (x_0 - b)^-], \]

\[ P(D_a^b[x(.)], N) \geq n\} \leq \frac{1}{(b-a)}E(x_N - b)^-I_{D_a^b[x(.)], N} = n\}]. \]

**Proof.** Let \( y_n = -x_n \). Then \( \{y_n\}_{n=0}^{\infty} \) is a submartingale. Hence

\[ U_a^b[x(.)], N] = D_{-a}^b[y(.)], N], \]

and

\[ D_a^b[x(.)], N] = U_{-a}^b[y(.)], N]. \]

Applying Theorem 18 we arrive at the results. ■

Theorem 18 and Corollary 19 are the classical crossing theorems on martingale. We can derive some other useful crossing results which are very useful in the mathematical finance. Here we apply some of them to derive the important limit theorem on martingales.
Theorem 20 If \( \{x_n\}_{n=0}^{\infty} \) is a submartingale such that there exists a subsequence of \( \{n\} \), denote it by \( \{n_k\} \), such that
\[
\sup_k E x_{n_k}^+ < \infty,
\]
then \( x_\infty = \lim_{n \to \infty} x_n \) exists a.s., and \( x_\infty \) is integrable. In particular, if \( x_n \leq 0, \forall n \), then condition (1.1) is obviously satisfied, and in this case \( \forall n \ E[x_\infty | \mathcal{F}_n] \geq x_n, \ a.s. \).

Proof. First, clearly
\[
\text{condition (1.1) } \iff \sup_n E x_n^+ < \infty \iff \sup_n E |x_n| < \infty.
\]
In fact, by the properties of submartingales one has that
\[
\mathbb{E}[x_n] \leq \mathbb{E}[x_{n_k}], \ \forall k.
\]
Hence the equivalent relations hold. Now let \( U^b_a(x(.)) = \lim_{N \to \infty} U^b_a(x(.), N) \).

Then by Theorem 18
\[
\mathbb{E}U^b_a(x(.)) \leq \frac{1}{b-a} \sup_{k} E(x_N - a)^+ < \infty.
\]
Hence \( U_a^b(x(.)) < \infty, \ a.s. \) Let
\[
W = \bigcup_{a,b \in \mathcal{Q}, a < b} W_{a,b} = \bigcup_{a,b \in \mathcal{Q}, a < b} \{\lim_n x_n < a < b < \lim_n x_n\}.
\]
Then
\[
P(W) \leq \sum_{a,b \in \mathcal{Q}, a < b} P(W_{a,b}) \leq \sum_{a,b \in \mathcal{Q}, a < b} P(U^b_a(x(.)) = \infty) = 0.
\]
Now we can let \( x_\infty(\omega) = \lim_{n \to \infty} x_n(\omega) \), as \( \omega \notin W \), and \( x_\infty(\omega) = 0 \), as \( \omega \in W \). By Fatou’s lemma
\[
\mathbb{E}|x_\infty| \leq \sup_n \mathbb{E}|x_n| < \infty.
\]
Hence \( x_\infty \) is integrable. In the case \( x_n \leq 0, \forall n \), by the definition of a submartingale
\[
0 \geq \mathbb{E}[x_m | \mathcal{F}_n] \geq x_n, \ a.s. \ \forall m.
\]
Again by Fatou’s lemma letting \( m \to \infty \) one reaches the final conclusion.

1.4 Uniform Integrability and Martingales

It is well known in the theory of real analysis that if a sequence of measurable functions is dominated by an integrable function, then one can take the limit under the integral sign for the function sequence. That is the famous Lebesgue’s dominated convergence theorem. However, sometimes it is difficult to find such a dominated function. In this case the uniform integrability of that function sequence can be a great help. Actually, in many cases it is a powerful tool.

Definition 21 A family of functions \( A \subset L^1(\Omega, \mathcal{F}, P) \) is called uniform integrable, if \( \lim_{\lambda \to \infty} \sup_{f \in A} \mathbb{E}(|f| I_{|f| > \lambda}) = 0 \), where \( L^1(\Omega, \mathcal{F}, P) \) is the totality of random variables \( \xi \), (that is, all \( \xi \) are \( \mathcal{F} \)-measurable) such that \( \mathbb{E} |\xi| < \infty \).
Lemma 22 Suppose that \( \{x_n\}_{n=1}^{\infty} \subset L^1(\Omega, \mathcal{F}, P) \) is uniformly integrable, and as \( n \to \infty \),
\[ x_n \to x, \text{ in probability,} \]
i.e. \( \forall \varepsilon > 0, \lim_{n \to \infty} P(|x_n - x| > \varepsilon) = 0 \), then
\[ \lim_{n \to \infty} E|x_n - x| = 0. \] (i.e. \( x_n \to x, \text{ in } L^1(\Omega, \mathcal{F}, P) \)).
In particular, \( \lim_{n \to \infty} E x_n = Ex \)

Proof. In fact, \( \forall \varepsilon > 0 \),
\[ E|x_n - x| \leq E(|x_n - x| I_{|x_n - x| > \lambda}) + E(|x_n - x| I_{|x_n - x| \leq \lambda}) = I_1^{n, \lambda} + I_2^{n, \lambda}. \]
Hence one can take a \( \lambda \) large enough such that \( I_1^{n, \lambda} < \varepsilon/2 \), since clearly \( \{x_n - x\}_{n=1}^{\infty} \) is uniformly integrable. Then for this fixed \( \lambda \) by using Lebesgue’s dominated convergence theorem one can have a sufficiently large \( N \) such that as \( n \geq N \), \( I_2^{n, \lambda} < \varepsilon/2. \)

For the sufficient conditions of uniform integrability of a family \( A \) we have

Lemma 23 Suppose that \( A \subset L^1(\Omega, \mathcal{F}, P) \). Any one of the following conditions makes \( A \) uniformly integrable:
1) There exists an integrable \( g \in L^1(\Omega, \mathcal{F}, P) \) such that
\[ |x| \leq g, \forall x \in A. \]
2) There exists a \( p > 1 \) such that \( \sup_{x \in A} E|x(\omega)|^p < \infty. \)

Proof. 1): Since as \( \lambda \to \infty \)
\[ \sup_{x \in A} P(|x(\omega)| > \lambda) \leq \frac{1}{\lambda} \sup_{x \in A} E|x(\omega)| \leq \frac{1}{\lambda} E|g(\omega)| \to 0. \]
So by the integrability of \( g \) one has that as \( \lambda \to \infty \)
\[ E|x(\omega)| I_{|x(\omega)| > \lambda} \leq E|g(\omega)| I_{|x(\omega)| > \lambda} \to 0, \text{ uniformly w.r.t. } x \in A. \]
2): Since \( \sup_{x \in A} P(|x(\omega)| > \lambda) \leq \frac{1}{\lambda} \sup_{x \in A} E|x(\omega)| \to 0, \text{ as } \lambda \to \infty. \) So as \( \lambda \to \infty \)
\[ \sup_{x \in A} E|x(\omega)| I_{|x(\omega)| > \lambda} \leq \sup_{x \in A} (E|x(\omega)|^p)^{1/p} \sup_{x \in A} [P(|x(\omega)| > \lambda)]^{(p-1)/p} \to 0. \]
Now we know that the uniform integrability condition is weaker than the domination condition. Actually, it is also the necessary condition for the \( L^1 \)-convergence of the sequence of integrable random variables or, say, integrable functions.

Theorem 24 Suppose that \( \{x_n\}_{n=1}^{\infty} \subset L^1(\Omega, \mathcal{F}, P) \). Then the following two statements are equivalent:
1) \( \{x_n\}_{n=1}^{\infty} \) is uniformly integrable.
2) \( \sup_{n=1,2,\ldots} E|x_n| < \infty; \text{ and } \forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall B \in \mathcal{F}, \text{ as } P(B) < \delta \)
\[ \sup_{n=1,2,\ldots} E|x_n| I_B < \varepsilon. \]
Furthermore, if there exists an \( x \in L^1(\Omega, \mathcal{F}, P) \) such that as \( n \to \infty \),
\[ x_n \to x, \text{ in probability; then the following statement is also equivalent to } 1): \]
3) \( x_n \to x, \text{ in } L^1(\Omega, \mathcal{F}, P). \)
1. Martingale Theory and the Stochastic Integral for Point Processes

Proof. Since $1) \implies 3)$ is already proved in Lemma 22, we will show that $3) \implies 2) \implies 1) \implies 2$.

$1) \implies 2)$: Take a $\lambda_0$ large enough such that $\sup_{n=1,2,\ldots} E |x_n| I_{|x_n|>\lambda_0} < 1$. Then

$$\sup_{n=1,2,\ldots} E |x_n| \leq \lambda_0 + 1.$$}

On the other hand, for any $B \in \mathcal{F}$ since

$$E |x_n| I_B = E |x_n| I_{|x_n|>\lambda} \cap B + E |x_n| I_{|x_n|\leq\lambda} \cap B \leq \sup_{n=1,2,\ldots} E |x_n| I_{|x_n|>\lambda} + \lambda P(B) = I_1^{\lambda} + I_2^{\lambda} B.$$

Hence $\forall \varepsilon > 0$ one can take a $\lambda > 0$ large enough such that $I_1^{\lambda} < \frac{\varepsilon}{2}$, then let $\delta_\varepsilon = \frac{\varepsilon}{2\lambda}$. For this $\delta_\varepsilon > 0$ one has that $\forall B \in \mathcal{F}, P(B) < \delta_\varepsilon \implies \sup_{n=1,2,\ldots} E |x_n| I_B < \varepsilon$.

$2) \implies 1)$: $\forall \varepsilon > 0$ Take $\delta > 0$ such that $2)$ holds. Since

$$P(|x_n| > \lambda) \leq \frac{1}{\lambda} \sup_{n=1,2,\ldots} E |x_n|,$$

Hence one can take an $N$ large enough such that as $\lambda > N,$

$$P(|x_n| > \lambda) < \delta, \forall n = 1, 2, \ldots.$$}

Thus by $2)$ as $\lambda > N,$

$$E |x_n| I_{|x_n|>\lambda} < \varepsilon, \forall n = 1, 2, \ldots.$$}

$3) \implies 2)$: Take an $N_0$ large enough such that as $n > N_0,$

$$E |x_n| < 1.$$}

Thus

$$\sup_{n=1,2,\ldots} E |x_n| \leq \max\{1 + E |x|, E |x_1|, \ldots, E |x_{N_0}|\} < \infty.$$}

On the other hand, observe that

$$E |x_n| I_B \leq E |x_n - x| + E |x| I_B.$$}

Hence $\forall \varepsilon > 0$, one can take an $N_\varepsilon$ large enough such that as $n > N_\varepsilon,$

$$E |x_n| < \frac{\varepsilon}{2}.$$}

Then take a $\delta > 0$ small enough such that $\forall B \in \mathcal{F}$, as $P(B) < \delta,$

$$\max_{n=1,2,\ldots, N_\varepsilon} \{E |x_n| I_B\} < \varepsilon, \text{ and } E |x| I_B < \varepsilon / 2.$$}

Thus as $P(B) < \delta$, $E |x_n| I_B < \varepsilon, \forall n = 1, 2, \ldots$. $\blacksquare$

Now let us use uniform integrability as a tool to study the martingales.

Theorem 25 If $\{x_n\}_{n=0}^\infty$ is a submartingale such that $\{x_n^+\}_{n=0}^\infty$ is uniformly integrable, then $x_\infty = \lim_{n \to \infty} x_n$ exists, a.s., and

$$E[x_\infty | \mathcal{F}_n] \geq x_n, \forall n,$$

i.e. $\{x_n\}_{n=0,1,2,\ldots, \infty}$ is also a submartingale, and we call it a right-closed submartingale.

This theorem actually tells us that a uniformly integrable submartingale is a right-closed submartingale.

Proof. By uniform integrability one has that

$$\sup_{n=0,1,2,\ldots} E x_n^+ < \infty.$$}

Hence applying Theorem 20 one has that $x_\infty = \lim_{n \to \infty} x_n$ exists, a.s. Now by the submartingale property $\{x_n^+\}_{n=0}^\infty$ is also a submartingale (Example 3). Hence for any $a > 0$, and $B \in \mathcal{F}_n$, as $m \geq n,$

$$\int_B [(-a) \vee x_n] dP \leq \int_B [(-a) \vee x_m] dP.$$}

Letting $m \to \infty$ by the uniform integrability of $\{x_n^+\}_{n=0}^\infty$ one has that