Introduction to Financial Mathematics

Hyun-Gon Ryu

Abstract. last draft : September 12, 2007

This documents are lecture note on intensive course for financial mathematics.
Nowadays, there are various textbooks for financial mathematics by many authors as Øksendal, Shereve, Kratzas, Cont, Krylov, etc.
To reach research level in financial mathematics, one must know about various topics as calculus, linear algebra, ODE, measure theory, probability theory, mathematical statistics, stochastic process, stochastic calculus, nemerical analysis, PDE, functional analysis. they also need to obtain the skill for SAS, programming with C, fortran and MATLAB, GAUSS, excel VBA.
I will summurize what to know in this field, and write down the knowhow to shorten the path.

Contents

CHAPTER 1

Mathematical Preliminaries

1. probability and measure theory

To study financial mathematics, some preliminaries are needed. first concept is measure and probability. I recommend to read Durret, Allen Gut, Gnedenko, Breiman and Olav Kallenberg's books for graduate level probability and check details in the textbooks about measure theory by for Frank Jones, Zigmund and Folland after read this note.

Definition 1.1. Probability Space $(\Omega, \mathcal{F}, P)$
A set Ω is a sample space
A collection $\mathcal{F}$ of subsets of Ω is a *σ-field* if

$\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$

for any $A \in \mathcal{F}, A^c := \Omega \backslash A \in \mathcal{F}$

for any countable collection $\{A\}_{n=1}^{\infty}$ of sets in $\mathcal{F}, \cup_{n=1}^{\infty} A_n \in \mathcal{F}$

The third property is called *countable additivity* or *σ-additivity*

A $[0, \infty]$-valued set function $\mu : \mathcal{F} \to [0, \infty]$ is a *measure* if

$$\mu(\emptyset) = 0$$

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n) A_n \in \mathcal{F} \text{ with } A_n \cap A_m = \emptyset, n \neq m$$

If $\mu(\Omega) = 1$, μ is a *probability measure* and we usually note P instead of μ. Specially, we call $(\Omega, \mathcal{F}, P)$ *Probability Space.*

Definition 1.2. Borel σ-field
Let $\mathbb{R}$ be the sample space and let $\mathcal{B} = \mathcal{B}(\mathbb{R})$ be the *σ-field* generated by all the open set in $\mathbb{R}$. $\mathcal{B}$ is called the *Borel σ-field* on $\mathbb{R}$
In other words, the Borel sigma algebra is equal to the intersection of all sigma algebras $\mathcal{A}$ of X having the property that every open set of X is an element of $\mathcal{A}$. An element of $\mathcal{B}$ is called a Borel subset of X, or a Borel set.

Definition 1.3. Measurable
A function $f : X \to \mathbb{R}$ is *measurable* if, for every real number a, the set $\{x \in X : f(x) > a\}$ is *measurable.*
When $X = \mathbb{R}$ with Lebesque measure μ, or more generally any Borel measure, then all continuous functions are measurable. In fact, practically any function that can be described is measurable. The measurable functions form one of the most general classes of real functions.

Definition 1.4. Random Variables
Let Probability Space $(\Omega, \mathcal{F}, P)$ and $(\mathbb{R}, \mathcal{B})$ be given. Measurable function $X : \Omega \to \mathbb{R}$ is a

random variable on Ω if

$$X^{-1}(A) \in \mathcal{F} \text{ for any } A \in \mathcal{B}$$

In this case

$$P_X(A) := P(X^{-1}(A))$$

is a probability measure on $(\mathbb{R}, \mathcal{B})$. P_X is called the *distribution* or *law* of X

DEFINITION 1.5. Lebesgue Integral
The *integral* of a measurable function $f\colon X \to \mathbb{R} \cup \{\pm\infty\}$ on a measure space $(X, \mathcal{B}, \mu)$ is written

$$\int_X f \; d\mu \text{ or just } \int f.$$

- for any $A \in \mathcal{F}$,

$$\int_X 1_A \; d\mu := \mu(A).$$

- If f is a simple function as $f = \sum_{i=1}^n a_k 1_{A_i}, a_i \in \mathbb{R}$

  for any $A_i \in \mathcal{F}$ and $a_i, \in \mathbb{R}$

$$\int_X f \; d\mu := \sum_{i=1}^n a_i \int_X 1_{A_i} \; d\mu = \sum_{i=1}^n a_i \mu(A_i).$$

- If f is a nonnegative measurable function

$$\int_X f \; d\mu := \sup \left\{ \int_X h \; d\mu : h \text{ is simple and } h(x) \le f(x) \text{ for all } \; x \in X \right\}.$$

- For f any measurable function write $f = f^+ - f^-$ where $f^+ := \max(f, 0)$ and $f^- := \max(-f, 0)$,
  so that $|f| = f^+ + f^-$, and define the integral of f as

$$\int_X f \; d\mu := \int_X f^+ \; d\mu - \int_X f^- \; d\mu,$$

  provided that $\int_X f^+ \; d\mu$ and $\int_X f^- \; d\mu$ are not both ∞.

- In the case of probability space $(\Omega, \mathcal{F}, P)$, we usually use the notation EX instead of $\int_\Omega X dP$;

$$EX = \int_\Omega X dP$$

$$E(X; A) = \int_A X dP = \int_\Omega X 1_A dP$$

DEFINITION 1.6. Product measure
Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two σ-finite measure spaces, i.e.,

$$\Omega_1 = \cup_{n=1}^\infty A_n, \mu_1(A_n) < \infty \text{ for all } n$$

$$\Omega_2 = \cup_{n=1}^\infty B_n, \mu_2(B_n) < \infty \text{ for all } n$$

We define $(\Omega, \mathcal{F})$ by

$$\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) | \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

$$\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2 = \sigma(A \times B | A \in \mathcal{F}_1, B \in \mathcal{F}_2)$$

Then, there exists a unique measure $\mu := \mu_1 \times \mu_2$ on $(\Omega, \mathcal{F})$ such for all $A \in \mathcal{F}_1, B \in \mathcal{F}_2$, $\mu(A \times B) = \mu_1(A) \times \mu_2(B)$ is called called product measure and $(\Omega, \mathcal{F}, \mu) := (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ is called the product measure space.

THEOREM 1.7. *Tonelli and Fubini*

$$f \geq 0 \Rightarrow \int_\Omega f d\mu = \int_{\Omega_2} \left(\int_{\Omega_1} f d\mu_1 \right) d\mu_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f d\mu_2 \right) d\mu_1$$

$$\int_\Omega |f| d\mu < \infty \Rightarrow \int_\Omega f d\mu = \int_{\Omega_2} \left(\int_{\Omega_1} f d\mu_1 \right) d\mu_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f d\mu_2 \right) d\mu_1$$

REMARK 1.8. some useful examples on probability

- $X \geq 0 \Rightarrow EX = \int_0^\infty P(X > t) dt$
- $p > 0, X \geq 0 \Rightarrow EX^p = \int_0^\infty pt^{p-1} P(X > t) dt$
- $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$
- $\mu(\{x \in X;\ |f(x)| \geq t\}) \leq \frac{1}{t^2} \int_X f^2 d\mu$ or $\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$

THEOREM 1.9. *Monotone convergence theorem, MCT*

$$f_n \geq 0, f_n \uparrow f \Rightarrow \int_\Omega f_n d\mu \to \int_\Omega f d\mu$$

THEOREM 1.10. *Fatou's lemma*

$$f_n \leq 0 \Rightarrow \int_\Omega \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu$$

THEOREM 1.11. *Dominated convergence theorem, DCT*

$$f_n \to f, |f_n| \leq g, \int_\Omega g d\mu \leq \infty \Rightarrow \int_\Omega f_n d\mu \to \int_\Omega f d\mu$$

DEFINITION 1.12. Convergence

- X_n converges to X almost surely if $P\left(\omega; \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1$
  denote this by $X_n \xrightarrow{a.s.} X$
- X_n converges to X in rth mean for $r > 0$ if $E|X_n - X|^e \to 0$
  denote this by $X_n \xrightarrow{r} X$
- X_n converges to X in probability if $P(|X_n - X| > \varepsilon) \to 0 \quad$ a $n \to \infty$
  denote this by $X_n \xrightarrow{p} X$
- X_n converges to X in distribution if $\rho_n(x) \to \rho(x)$ at every point x
  denote this by $X_n \xrightarrow{D} X$

Definition 1.13. weakly converges
Let μ_n, μ be finite nonnegative measure on $(\Omega, \mathcal{B}(\Omega))$ we say μ_n *converges weakly* to μ and write $\mu_n \xrightarrow{w} \mu$ if for every bounded continuous function f

$$\int_\Omega f \mu_n(dx) \longrightarrow \int_\Omega f \mu(dx)$$

Definition 1.14. Conditional Expectation Let $(\Omega, \mathcal{F}, P)$ be a proabbility measure space.
Let $\mathcal{G}$ be a sub σ-field of $\mathcal{F}$, that is, $\mathcal{G} \subset \mathcal{F}$
Let $X : \Omega \to \mathbb{R}$ be a $\mathcal{F}$-measurable random variable with $E|X| < \infty$.
then, there exists a random variable $Y \in \mathcal{G}$ such that

$$E(X; A) = E(Y; A) \text{ for any } A \in \mathcal{G}$$

Y is called the *conditional expectation* of X wrt $\mathcal{G}$ and it is denoted by $E(X|\mathcal{G})$.

Theorem 1.15. *Basic property of conditional expectation*

(1) *linearity* $E(\alpha X + \beta Y|\mathcal{G}) = \alpha E(X|\mathcal{G}) + \beta E(Y|\mathcal{G})$
(2) *Monotonicity* $E(X|\mathcal{G}) \geq E(Y|\mathcal{G})$ *if* $X \geq Y$
(3) *Tower* $E(E(X|\mathcal{G})|\mathcal{H}) = E(E(X|\mathcal{H})|\mathcal{G}) = E(X|\mathcal{H})$ *if* $\mathcal{H} \subset \mathcal{G}$
(4) $E(X|\mathcal{G}) = X$ *if* $X \in \mathcal{G}$
(5) $E(X|\mathcal{G}) = EX$ *if* X *and* Y *are independent.*
(6) $E(E(X|\mathcal{G})) = EX$
(7) $E(XY|\mathcal{G}) = XE(Y|\mathcal{G})$ *if* $X \in \mathcal{G}$
(8) *Jensen's ineq.* $E(\varphi(X)|\mathcal{G}) \geq \varphi(E(X|\mathcal{G}))$ *for convex function* φ
(9) *MCT* $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ *if* $X_n \geq 0$ *and* $X_n \uparrow X$

Definition 1.16. Characteristic function
Let X be a random variable defined on $(\Omega, \mathcal{F}, P)$ and taking values in $\mathbb{R}$ with probability law p_X. Its *characteristic function* $\phi_X : \mathbb{R} \to \mathbb{C}$ is defined by

$$\begin{aligned} \phi_X(u) :=& \mathbb{E}(e^{iuX}) = \int_\Omega e^{iuX(\omega)} P(d\omega) \\ &= \int_\mathbb{R} e^{iuy} P_X(dy) \end{aligned}$$

for each $u \in \mathbb{R}$.

Definition 1.17. Convolution
Let $\mathcal{M}_1(\mathbb{R})$ denote the borel probability measure on $\mathbb{R}$, We define the *convolution* of two probability measures as follows:

$$(\mu_1 * \mu_2)(A) := \int_\mathbb{R} \mu_1(A - x)\mu_2(dx)$$

for each $\mu_i \in \mathcal{M}_1(\mathbb{R}), i = 1, 2$ and each $A \in \mathcal{B}(\mathbb{R})$ where we denote $A - x = \{y - x | y \in A\}$

PROPOSITION 1.18. *$f \in \mathbf{B}_b(\mathbb{R})$, then for all $\mu_i \in \mathcal{M}_1(\mathbb{R}), i = 1, 2$*

$$\int_{\mathbb{R}} f(y)(\mu_1 * \mu_2)(dy) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x + y)\mu_1(dy)\mu_2(dx)$$

COROLLARY 1.19. *For each $f \in \mathbf{B}_b(\mathbb{R})$,*

$$\mathbb{E}(f(X_1 + X_2)) = \int_{\mathbb{R}} f(z)(\mu_1 * \mu_2)(dz)$$

therefore,

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_x(x - y) f_Y(y) dy$$

DEFINITION 1.20. infinitely divisible Let X be a random variable taking values in $\mathbb{R}$ with law μ_X. We say that X is *infinitely divisible* if, for all $n \in \mathbb{N}$, there exist i.i.d. random variables $Y_1^{(n)}, \ldots, Y_n^{(n)}$ such that

$$X \stackrel{\mathrm{d}}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}$$

Let $\phi_X(u) = \mathbb{E}(e^{iuX})$ denote the characteristic function of X, where $uin\mathbb{R}$.

PROPOSITION 1.21. *The following are equivalent :*

(1) *X is infinitely divisible*
(2) *μ_X has a convolution nth root that is itself the lwa of a random bariable, for each $n \in \mathbb{N}$,*
(3) *ϕ_X has an nth root that is itself the characteristic function of a random variable, for each $n \in \mathbb{N}$*

REMARK 1.22. $\mu \in \mathcal{M}_1(\mathbb{R})$ is infinitely divisible iff $\mu^{1/n} \in \mathcal{M}_1(\mathbb{R})$ for each $\phi_\mu(x) = [\phi_{\mu^{1/n}}(x)]^n$

2. mathematical statistics

2.1. linear algebra.

2.2. estimation.

2.3. test.

2.4. regression model.

2.5. time-series analysis.

3. Stochastic Process

3.1. Basic of Stochastic process.

Definition 3.1. Filtration
A *filtration* is an increasing sequence of σ-algebras on a measurable space. That is, given a measurable space $(\Omega, \mathcal{F})$, a filtration is a sequence of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$with $\mathcal{F}_t \subseteq \mathcal{F}$ for each t and $t_1 \leq t_2 \Rightarrow \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$. In finance, we consider it as information up to time t.

Definition 3.2. Stochastic Process
Given a probability space $(\Omega, \mathcal{F}, P)$, a *stochastic process* with state space X is a collection of X-valued random variables index by a set T. That is, a Stochastic process X is a collection

$$\{X_t : t \in T\}$$

where each F_t is an X-valued random variable.

Definition 3.3. $\mathcal{F}_t$-measurable, adapted and predictable
A random variable X is said to be *$\mathcal{F}_t$-measurable*(or measurable with respect to $\mathcal{F}_t$) if $\{x_1 < X \leq x_2\} \in \mathcal{F}$ for any $x_1 \leq x_2$. A stochastic process $\{X_t; t = 0, 1, \ldots, T\}$ is said to be *adapted* to the filtration $\{\mathcal{F}_t\}$, if each X_t is measurable with respect to $\mathcal{F}_t$. The process $\{X_t\}$ is called *predictable* if X_t is measurable with respect to $\mathcal{F}_{t-1}$ for all $t = 1, 2, \ldots, T$

Remark 3.4. if X_t is adapted to its own filtration $\mathcal{F}^X = \sigma\{X_s | 0 \leq s \leq t\}$ is usually called the *natural filtration* and then, $\mathbb{E}[X_s|\mathcal{F}_s] = X_s$ a.s.

Definition 3.5. stopping time
A *stopping time* is a random variable $\tau : \Omega \to [0, \infty]$ for which the event $(\tau \leq t) \in \mathcal{F}_t$ for each $t \geq 0$.

first hitting time τ_A is defined by

$$\tau_A = \inf\{t \geq 0; X_t \in A\}$$

Definition 3.6. Martingale
Let X be an adapted process defined on a filtered probability space that also satisfies the integrability requirement $\mathbb{E}[|X_s||\mathcal{F}_s] \leq \infty$ for all $t \geq 0$. We say that is a *martingale* if ,for all $0 \leq s < t < \infty$

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \quad \text{a.s.}$$

Remark 3.7. $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$ then submartingale, $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$ then supermartingale

Definition 3.8. semimartingale
A process X_t is called a *semimartingale*, if X_t can be decoposed as follows : $X_t = M_t + A_t$ where M_t is a local martingale, and A_t is an adapted process with finite variation.

Definition 3.9. Markov
Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a filtration $\mathcal{F}_t, t \geq 0$ Let $\{X_t\}_{t\geq 0}$ be an adapted process. We say that X_t is a *Markov process* if, for all $f \in \mathbf{B}_b(\mathbb{R})$ $0 \leq s < t < \infty$

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s] \quad \text{a.s.}$$

Definition 3.10. semigroup operator
With Markov process X_t, let define the *Operator* $T_{s,t}, 0 \leq s \leq t < \infty$ from $\mathbf{B}_b(\mathbb{R})$ to the Banach space of all bounded functions on $\mathbb{R}$ by

$$(T_{s,t}f)(x) = \mathbb{E}[f(X_t)|X_s = x]$$

then, The Operator is called *semigroup Operator* if

$$(T_{s+t}f)(x) = (T_t f)(x)(T_s f)(x)$$

REMARK 3.11. If X_t is an arbitrary Markov process, Operator $T_{s,t}, 0 \leq s \leq t < \infty$ on X_t is semigroup operator

DEFINITION 3.12. Transition probability
we call the mappings $p_{s,t}$ *transition probabilities*, as they give the probabilities of transition of the process from the point x at time s to the set A at time t.

$$p_{s,t}(x, A) = P(X_t \in A | X_s = x)$$

REMARK 3.13. If X_t is an arbitrary Markov process, $(T_{s,t}f)(x) = \int_{\mathbb{R}} f(y) p_{s,t}(x, dy)$
If X be lévy process and let q_t be law of X_t then $p_{s,t}(x, A) = q_{t-s}(A - x)$

stationary, gaussian, markov process, point process
levy process cadlag poisson process brown motion martingale

3.2. Poisson Process.

DEFINITION 3.14. Point Process
A process X_t is called *gaussian* iff for every finite set of indice $t_1, \cdots, t_k$ in T,

$$\tilde{\mathbf{X}}_{t_1,\cdots,t_k} = (\tilde{\mathbf{X}}_{t_1}, \cdots, \tilde{\mathbf{X}}_{t_k})$$

DEFINITION 3.15. Poisson Process
A process X_t is called *poisson* if

(1) it has independent increment

(2)the increment is stationary in time : $N_{t+s} - N_t depends only on s$

(3)the process admits at most single jump : $P(N_{t+h} - N_t) \geq 2 | N_t) = \wr(h)$

3.3. Brownian Motion.

DEFINITION 3.16. Gaussian Process
A process X_t is called *gaussian* iff for every finite set of indice $t_1, \cdots, t_k$ in T,

$$\tilde{\mathbf{X}}_{t_1,\cdots,t_k} = (\tilde{\mathbf{X}}_{t_1}, \cdots, \tilde{\mathbf{X}}_{t_k})$$

is a vector-valued Gaussian random variable.
OU process is stationary gaussian process.

DEFINITION 3.17. Stationary Process
a *stationary process* is a stochastic process whose probability distribution at a fixed time or position is the same for all times or positions. As a result, parameters such as the mean and variance, if they exist, also do not change over time or position.

DEFINITION 3.18. Stable Process
a *stationary process* is a stochastic process whose distribution is *stable* if

$$\frac{X_1 + X_2 + \cdots + X_n - d_n}{c_n} \Rightarrow X$$

DEFINITION 3.19. α - Stable Process
a *stationary process* is a stochastic process whose distribution is emphstable if

$$\frac{X_1 + X_2 + \cdots + X_n - d_n}{c_n} \Rightarrow X$$

Definition 3.20. α - similar process
a process X_t is α - *similar process* if

$$\sqrt[\alpha]{c} X_t \sim X_{t/c}$$

Specially, if $\alpha = 1/2$, we call it *self-similar process.* Brownian motion and gaussian process is self-similar process.

stable, stationary, similar properties are related.

Definition 3.21. Brownian Motion
A stochastic process, that is a collection of random variables $\{B_t; 0let < \infty\}$ defined on the same probability space$(\Omega, \mathcal{F}, P)$ is a *Brownian Motion* if a random process X_t is defined by

$$\text{Increament of } B_t \text{ , } B_t - B_s \text{is independent to } \mathcal{F}_s := \sigma\{B_t; 0 \leq u \leq s\}$$
$$\text{Increament of } B_t \text{ , } B_t - B_s \sim N(0, t-s) \text{ for } t > s$$
$$t \mapsto B_t \text{ is continuous with probability 1.}$$

where $\{t_0, \cdots, t_{N+1}\}$ satisfying

Remark 3.22. BM is not stationary, but it has stationary increments

Theorem 3.23. *Property of Brown Motion*

(1) $\mathbb{E}[B_t] = 0$

(2) $\mathbb{E}[B_s B_t] = s \wedge t$

(3) $\mathbb{E}\left[e^{\lambda B_t}\right] = e^{\lambda^2 t/s}$

(4) $\mathbb{E}\left[e^{i\theta B_t}\right] = e^{-\theta^2 t/s}$

(5) B_t *is a martingale*

3.4. levy process.

Definition 3.24. stochastically continuous
A stochastic process X_t is called *stochastically continuous* if

$$\lim_{t' \to t} P(|X_{t'} - X_t| > \varepsilon) = 0$$
$$X_{t'} \xrightarrow{P} X_t \quad \text{as } t' \to t$$

Definition 3.25. càdlàg
A stochastic process is called *càdlàg* if it has a.s. right continuous paths and the limits from the left exist. and the stochastic process is continuous if its paths are a.s. continuous.
If f is *càdlàg function* we will denote the left limit at each point $t \in (a, b]$ as

$$f(t-) = \lim_{s \to t} f(s)$$
$$f(t-) = f(t) \text{ iff } f \text{ is continuous at } t$$
$$\text{jump at } t \text{ by } \Delta f(t) = f(t) - f(t-)$$

Clearly, a càdlàg function can only have jump discontinuous. and every a càdlàg function is Borel measurable.

DEFINITION 3.26. Levy process A stochastic process X_t is called *levy process* if process is stochastic continuous
increment of process is independent and stationary.

CHAPTER 2

Stochastic Calculus

DEFINITION 0.27. π, λ- system Let Ω be the sample sample.
We call $\mathcal{A}$,the subset of Ω be π-*system*, if

$$A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$$

We call $\mathcal{C}$,the subset of Ω be λ-*system*, if

$$\Omega \in \mathcal{C}$$
$$A, B \in \mathcal{C} \text{ and } B \subset A \Rightarrow A \setminus B \in \mathcal{A}$$
$$A_n \in \mathcal{C} \text{ and } A_n \subset A_{n+1} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{C}$$

THEOREM 0.28. *Dynkin's π, λ- system*

$$\mathcal{A} \subset \mathcal{C} \Rightarrow \sigma(\mathcal{A}) \subset \mathcal{C}$$

1. Construction of Itô integral

DEFINITION 1.1. p^{th}- variation process
In general, the p^{th}- *variation process* for a random process X_t is defined by

$$\langle X, X \rangle_T^p(\omega) := \lim_{\Delta t_k \to 0} \sum_{k=0} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega))|^p$$

where $\{t_0, \cdots, t_{N+1}\}$ satisfying $0 = t_0 < t_1 < \cdots < t_N < T_{N+1} = T$ is a partition of $[0, T]$

DEFINITION 1.2. Quadratic Variation
We specially call *Quadratic Variation* which is 2^{th}- variation as below

$$\begin{aligned} Q(X) &:= \lim_{\Delta t \to 0} Q_n(X) \\ &= \lim_{\Delta t \to 0} \sum_{k=0}^{n-1} (X(t_{k+1} - X(t_k))^2 \end{aligned}$$

REMARK 1.3. Quadratic Variation of Brown Motion. if X is a differentiable function of t, Quadratic variation is zero, but for Brown motion, Quadratic variation is T a.s. and $\mathbb{E}[Q_n] = T$ and $Var[Q_n] = 3T^2/n$ for any n

DEFINITION 1.4. Total Variation
Let $\gamma : [a, b] \to X$ be a function mapping an interval $[a, b]$ to a metric space (X, d). We say that γ is of *bounded variation* if there is a constant M such that, for each partition $P = \{a = t_0 < t_1 < \cdots < t_n = b\}$ of $[a, b]$,

$$v(\gamma, P) = \sum_{k=1}^{n} d(\gamma(t_k), \gamma(t_{k-1})) \le M.$$

The *total variation* V_γ of γ is defined by

$$V_\gamma = \sup\{v(\gamma, P) : P \text{ is a partition of } [a, b]\}.$$

It can be shown that, if X is either $\mathbb{R}$ or $\mathbb{C}$, every smooth (or piecewise smooth) function $\gamma : [a, b] \to X$ is of bounded variation, and

$$V_\gamma = \int_a^b |\gamma'(t)| dt.$$

Also, if γ is of bounded variation and $f : [a, b] \to X$ is continuous, then the Riemann-Stieltjes integral $\int_a^b f d\gamma$ is finite.

REMARK 1.5. Total Variation of Brown Motion. if X_t is Brown motion, Total Variation is unbounded almost surely.

1.1. Riemann-Stieltjes Integral.

DEFINITION 1.6. Riemann-Stieltjes Integral
Let f ϕ be two functions which are defined and finite on a finite interval $[a, b]$. Let $\Gamma = \{a = x_0 < x_2 < \cdots < x_m = b\}$ is a *partition* of $[a, b]$, we select intermediate points $\{\xi_i\}_{i=1}^m$ satisfying $x_{i-1} \le \xi_i \le x_i$ and write

$$R_\Gamma = \sum_{i=1}^m f(\xi_i)[\phi(x_i) - \phi(x_{i-1})].$$

R_Γ is called a *Riemann-Stieltjes sum* for Γ, is depend on ξ_i, f,ϕ and interval$[a, b]$. If

$$I = \lim_{|\Gamma|\to 0} R_\Gamma$$

exist and is finite, for $\varepsilon > 0, \delta > 0$ $|I - R_\Gamma| < \varepsilon$ for any Γ satisfying $|\Gamma| < \delta$, then I is called the *Riemann-Stieltjes integral* of f with respect to ϕ on $[a, b]$, and denoted

$$I = \int_a^b f(x) d\phi(x) = \int_a^b f d\phi$$

if $\phi(x) = x$ then *Riemann-Stieltjes integral* is *Riemann integral*

1.2. Motivation. Let consider stochastic process $dX_t = \mu(t)dt + \sigma(t)dBt$ If it is integrable, its integral is

$$X_t = \int_0^t \mu(s)ds \quad + \int_0^t \sigma(s)dBt$$

first term of right side is Riemann integral, second term of right side is similiar with Riemann steiltjes integral with $f(t) = \sigma(t)$, $\phi(t) = B_t$
In the property of *Riemann-Stieltjes integral* ϕ must be finite variation. but B_t is not Bounded Variation. and 2-variation is finite. therefore, we can construct its integral with 2-variational sense.

For define $\int_0^t BsdBt$
Let check the properties.

$$\begin{aligned} S_n(t^*) &= \sum B(t^*)\Delta B_i \quad t^* = (t_1^*, \cdots, t_n^*) \\ &= \sum B(t_{i-1})(B_i - B_{i-1}) \quad \text{ choose } t_i^* = t_{i-1} \\ &= \frac{1}{2}\sum (B_i^2 - B_{i-1}^2)(B_{t_i} - B_{t_{i-1}})^2 \quad a(b-a) = \frac{b^2 - a^2 - (b-a)^2}{2} \\ &= \frac{1}{2}[(B_t^2 - B_0^2) + \sum (B_{t_i} - B_{t_{i-1}})^2] \quad \text{ expansion} \\ &= \frac{1}{2}B_t^2 + \frac{1}{2}\sum (B_{t_i} - B_{t_{i-1}})^2 \end{aligned}$$

Let define $Q_n = \sum(B_{t_i} - B_{t_{i-1}})^2 = E[\Delta B_t^2]$ Consider $E[Q_n]$, $Var[Q_n]$

$$\begin{aligned}
E[Q_n] &= E[\sum(B_{t_i} - B_{t_{i-1}})^2] \\
&= \sum E[(B_{t_i} - B_{t_{i-1}})^2] \\
&= \sum t_i - t_{i-1} \quad \text{property of BM} \\
&= t - 0 \quad \text{expansion} \\
E[\Delta B_t^2] &= t
\end{aligned}$$

$$\begin{aligned}
Var[Q_n] &= E[Q_n^2] - E[Q_n]^2 \\
&= E[(\sum(B_{t_i} - B_{t_{i-1}})^2)^2] - E[\sum(B_{t_i} - B_{t_{i-1}})^2]^2 \\
&= E[\sum(B_{t_i} - B_{t_{i-1}})^4 + \sum_{i \neq j}(B_{t_i} - B_{i-1})(B_{t_j} - B_{t_{i-1}})] \\
&\quad - [\sum_i E[(B_{t_i} - B_{t_{i-1}})^2]^2 + \sum_{i \neq j} E[(B_{t_i} - B_{t_{i-1}})]E[(B_{t_j} - B_{t_{j-1}})]] \\
&= E[\sum(B_{t_i} - B_{t_{i-1}})^4] - \sum_i E[(B_{t_i} - B_{t_{i-1}})^2]^2 \\
&= A + B \quad B = \sum_i \Delta t^2 \\
E[\sum_i(B_{t_i} - B_{t_{i-1}})^4] &= \sum E(B_{t_i} - B_{t_{i-1}})^4 \\
&= \sum_i E(B_{t_i - t_{i-1}} - B_0)^4 = \sum_i E(B_{t_i - t_{i-1}})^4 \\
&= \sum_i (E\sqrt{t_i - t_{i-1}} B_1)^4 \\
&= \sum_i \sqrt{t_i - t_{i-1}}^4 E(B_1)^4 \\
&= \sum_i \Delta t^2 E(B_1)^4 = \sum_i \Delta t^2 3 \\
Var[Q_n] &= \sum_i \Delta t^2 3 - \sum_i \Delta t^2 \\
&= (3-1)\sum_i \Delta t^2 \leq 2|\Delta t| \to 0 \quad \text{as } t \to 0
\end{aligned}$$

In conclusion,

$$E[(Q_n - t)^2] = Var[Q_n] \to 0$$

therefore, we consider norm $\| \cdot \| = E[(\cdot)^2]$ and its equality is mean square sense's equation

$$Q_n \xrightarrow{m.s.} t \quad , \quad S_n \xrightarrow{m.s.} \frac{1}{2}B_t^2 - \frac{1}{2}t^2$$

Let define

$$\int_0^t f dB_t := m.s. \lim \int_0^t \phi_n^f dB_t$$

$$E[(f - \phi_n^f)^2] \to 0$$

2. Itô's lemma

THEOREM 2.1. *Itô's lemma Let X_t be an Itô process.*

$$d(X_t) = \mu(t, X)dt + \sigma(t, X)dW_t$$

where W_t is Wiener process, and let $f(t, X)$ be a function with with continuous second derivatives. then. $f(t, X)$ is also an Itô process and

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}dX + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}dX^2 \\ &= [\frac{\partial f}{\partial t} + \mu(t, X)\frac{\partial f}{\partial X} + \sigma(t, X)^2\frac{1}{2}\frac{\partial^2 f}{\partial X^2}]dt + \sigma(t, X)\frac{\partial f}{\partial X}dW_t \end{aligned}$$

REMARK 2.2. Example of Geometric Brownian Motion
Let consider Geometric Brownian Motion X_t

$$dX_t = \mu X_t dt + \sigma X_t dB_t$$

let $f(t, x) = ln(x)$ with $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = \frac{1}{x}$, $\frac{\partial^2 f}{\partial x^2} = -\frac{1}{x^2}$
by ito lemma,

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}dX + \frac{1}{2}\frac{\partial f}{\partial X}dX^2 \\ &= 0dt + \frac{1}{X_t}[\mu X_t dt + \sigma X_t dB_t] - \frac{1}{2}\frac{1}{X_t^2}[\mu X_t dt + \sigma X_t dB_t]^2 \\ &= [\mu + \frac{1}{2}\sigma^2]dt + \sigma dW_t \\ d(ln(X_t)) &= [\mu - \frac{1}{2}\sigma^2]dt + \sigma dW_t \\ ln(X_t) &= ln(X_0) + \int_0^t [\mu - \frac{1}{2}\sigma^2]ds + \int_0^t \sigma dW_s \\ ln(X_t) &= ln(X_0) + [\mu - \frac{1}{2}\sigma^2]t + \sigma W_t \\ X_t &= X_0 e^{[\mu - \frac{1}{2}\sigma^2]t + \sigma W_t} \end{aligned}$$

REMARK 2.3. Example of OU process
Let consider Ornstein-Uhlenbeck process(or Vasicket process) X_t

$$dX_t = \alpha(m - X_t)dt + \beta dB_t$$

let $f(t, x) = xe^{\alpha t}$ with $\frac{\partial f}{\partial t} = \alpha x e^{\alpha t}$, $\frac{\partial f}{\partial x} = e^{\alpha t}$, $\frac{\partial^2 f}{\partial x^2} = 0$
by ito lemma,

$$
\begin{aligned}
df(t,X_t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X}dX + \frac{1}{2}\frac{\partial f}{\partial X}dX^2 \\
&= (\alpha X_t e^{\alpha t})dt + e^{\alpha t}[\alpha(m - X_t)dt + \beta dB_t] + 0[\alpha(m - X_t)dt + \beta dB_t]^2 \\
&= (\alpha X_t e^{\alpha t} + e^{\alpha t}\alpha(m - X_t))dt + \beta e^{\alpha t} dB_t \\
&= (\alpha m e^{\alpha t})dt + \beta e^{\alpha t} dB_t \\
d(X_t e^{\alpha t}) &= (\alpha m e^{\alpha t})dt + \beta e^{\alpha t} dB_t \\
X_t e^{\alpha t} &= X_0 e^{\alpha 0} + \int_0^t (\alpha m e^{\alpha s})ds + \int_0^t \beta e^{\alpha s} dB_s \\
X_t &= X_0 e^{-\alpha t} + m(1 - e^{-\alpha t}) + \int_0^t \beta e^{\alpha(s-t)} dB_s
\end{aligned}
$$

thus, $\mathbb{E}[X_t] = X_0 e^{-\alpha t} + m(1 - e^{-\alpha t})$
C.F. The Ornstein-Uhlenbeck process is a stationary Gaussian process.

Theorem 2.4. *product formula*

$$d(X_tY_t) = (dX_t)Y_t + X_t(dY_t) + (dX_t)(dY_t)$$

Theorem 2.5. *Itô isometry*

$$\mathbb{E}\left[(\int_0^T f(t)dB_t)^2\right] = \mathbb{E}\left[\int_0^T f(t)^2 dt\right]$$

Remark 2.6. Example of Itô isometry
Let X_t OU process, Covariance

$$
\begin{aligned}
Cov(X_s, X_t) &= \mathbb{E}[(X_s - \mathbb{E}[X_s])(X_t - \mathbb{E}[X_t])] \\
&= \mathbb{E}[(X_s - (X_0 e^{-\alpha s} + m(1 - e^{-\alpha s})))(X_t - (X_0 e^{-\alpha t} + m(1 - e^{-\alpha t})))] \\
&= \mathbb{E}[((X_0 e^{-\alpha s} + m(1 - e^{-\alpha s}) + \int_0^t \beta e^{\alpha(u-s)} dB_u) - (X_0 e^{-\alpha s} + m(1 - e^{-\alpha s}))) \\
&\qquad ((X_0 e^{-\alpha t} + m(1 - e^{-\alpha t}) + \int_0^t \beta e^{\alpha(v-t)} dB_v) - (X_0 e^{-\alpha t} + m(1 - e^{-\alpha t})))] \\
&= \mathbb{E}[(\int_0^t \beta e^{\alpha(u-s)} dB_u)(\int_0^t \beta e^{\alpha(v-t)} dB_v)] \\
&= \beta^2 e^{-\alpha(s+t)} \mathbb{E}[\int_0^t e^{\alpha u} dB_u \int_0^t e^{\alpha v} dB_v] \\
&= \beta^2 e^{-\alpha(s+t)} \mathbb{E}[(\int_0^{t\wedge s} e^{\alpha u} dB_u)^2] \quad \text{by ito isometry} \\
&= \beta^2 e^{-\alpha(s+t)} \mathbb{E}[\int_0^{t\wedge s} e^{2\alpha u} du] \\
&= \frac{\beta^2}{e}^{-\alpha(s+t)} 2\alpha \mathbb{E}[(e^{2\alpha(t\wedge s)} - 1)] \\
&= \frac{\beta^2}{e}^{-\alpha(s+t)} 2\alpha (e^{2\alpha(t\wedge s)} - 1)
\end{aligned}
$$

3. Relationship between SDE and PDE

DEFINITION 3.1. infinitesimal generator
Let X_t be an Ito diffusion in $\mathbb{R}$ The *infinitesimal generator* $\mathcal{A}$ of X_t is defined by

$$\mathcal{A}f(x) := \lim_{t\downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} \quad x \in \mathbb{R}$$

THEOREM 3.2. *generator of ito diffusion*
Let X_t be stochastic process as

$$X_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t$$

then infinitesimal generator of X_t is

$$\begin{aligned}\mathcal{A}f(X_t) &= \lim_{t\downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} \quad x \in \mathbb{R} \\ &= \mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\end{aligned}$$

REMARK 3.3. generator of brown motion is half laplacian that satisfies $u_t = \frac{1}{2}u_{xx}$

THEOREM 3.4. *Dynkins' formula*
Let τ be a stopping time with $\mathbb{E}^x[\tau] < \infty$ and let $f \in C_0^2$ then, Dynkins' formula hold

$$\mathbb{E}^x[f(X_\tau)] = f(x) + \mathbb{E}^x[\int_0^\tau \mathcal{A}f(X_s)ds]$$

Let X_t be an Ito diffusion in $\mathbb{R}$ with generator $\mathcal{A}$. If we choose $f \in C_0^2(\mathbb{R})$ $\tau = t$ in Dynkin's formula, we see that

$$u(t, x) = \mathbb{E}^x[f(X_t)]$$

is differentiable with respect to t and

$$\frac{\partial u}{\partial t} = \mathbb{E}^x[\mathcal{A}f(X_t)]$$

THEOREM 3.5. *Komogorov's Backward Equation*
Let $f \in C_0^2(\mathbb{R})$ Define

$$u(t, x) = \mathbb{E}^x[f(X_t)]$$

then $u(t, \cdot) \in \mathcal{D}_\mathcal{A}$ for each t and

$$\begin{aligned}\frac{\partial u}{\partial t} &= \mathcal{A}u \text{ for } t > o, x \in \mathbb{R} \\ u(0, x) &= f(x) \text{ for } x \in \mathbb{R}\end{aligned}$$

THEOREM 3.6. *Komogorov's Forward Equation*
Let X_t be an Ito diffusion in $\mathbb{R}$ with generator $\mathcal{A}$ and assume that the transition measure of X_t has a density $p_t(x, y)$, i.e.;

$$\mathbb{E}^x[f(X_t)] \int_\mathbb{R} f(y)p_t(x, y)dy; \qquad f \in C_0^2$$

Assume that $y \to p_t(x,y)$ is smooth for each t, x. than

$$\frac{d}{dt}p_t(x,y) = \mathcal{A}^* p_t(x,y)$$
$$= -\tfrac{\partial}{\partial y}[\mu p_t(x,y)] + \tfrac{1}{2}\tfrac{\partial^2}{\partial y^2}[\sigma^2 p_t(x,y)]$$

where $\mathcal{A}^* f$ is the adjoint operator of $\mathcal{A}$

THEOREM 3.7. *Original The Feynmann-Kac formula*
Let $f \in C_0^2(\mathbb{R})$ and $q \in C_0^2(\mathbb{R})$. Assume that q is lower bounded. Put

$$v(t,x) = \mathbb{E}^x[e^{-\int_0^t q(X_s)ds} f(X_t)]$$

then

$$\frac{\partial v}{\partial t} = \mathcal{A}v - qv \text{ for } t > o, x \in \mathbb{R}$$
$$v(0,x) = f(x) \text{ for } x \in \mathbb{R}$$

THEOREM 3.8. *simple Feynmann-Kac formula*
Let $f \in C_0^2(\mathbb{R})$ and $q \in C_0^2(\mathbb{R})$. Assume that q is lower bounded. Put

$$f(t,x) = \mathbb{E}^x[\mathcal{H}(T, S_T)]$$

then

$$\frac{\partial f}{\partial t} + \mathcal{A}f = 0$$
$$f(T,x) = \mathcal{H}(T,S_T) x \in \mathbb{R}$$

THEOREM 3.9. *discounted The Feynmann-Kac formula*
Let $f \in C_0^2(\mathbb{R})$. Assume that q is lower bounded. Put

$$f(t,x) = \mathbb{E}^x[e^{-(T-t)r}\mathcal{H}(T,S_T)]$$

then

$$\frac{\partial f}{\partial t} + \mathcal{A}f = rf$$
$$f(T,x) = \mathcal{H}(T,S_T) \quad \text{for } x \in \mathbb{R}$$

REMARK 3.10. Feynmann-Kac formula for option pricing
Let S_t is GBM, Let β_t is riskless asset, Call Option price is

$$c(t,x) = \mathbb{E}^x[e^{-(T-t)r}\mathcal{H}(T,S_T,K)]$$

then by Feynmann-Kac formula

$$\frac{\partial f}{\partial t} + rx\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} = rf$$
$$f(T,x) = \mathcal{H}(T,S_T)$$

$$f_t + rxf_x + \frac{1}{2}\sigma^2 x^2 f_{xx} - rf = 0$$
$$f(T,x) = \mathcal{H}(T,S_T)$$

this is black sholes PDE

CHAPTER 3

Partial Differential Equations

0.1. solving heat equation.

0.2. solving by green function.

0.3. HJB equation.

0.4. Finite Difference Method.

0.5. sobolev spaces and functional analysis.

0.6. Finite Element Method.

0.7. Sparse Grid Method.

CHAPTER 4

Girsanov theorem

0.8. Girsanov theorem.

0.9. Numeraire.

0.10. Fundamental Theorem of Asset Pricing.

CHAPTER 5

The Classical Black-Sholes Option Pricing Model

The famous Option Pricing Model is Merton, Black....

1. Model

European Options
Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

American Options
Payoff Function is $H(T, S_T) = \sup_{t \le T}[S_t - K, 0]$

Asian Option
Payoff Function is $H(T, S_T) = \max[\frac{1}{T}\int_0^T S_t dt - K, 0]$

and there are various exotic options as barrier, double barrier, ladder, etc.

1.1. Black-Scholes Option Pricing. Let consider asset S_t with riskless asset β_t

$dS_t = \mu S_t dt + \sigma S_t dB_t$ GBM

$d\beta_t = rdt$

European Call Options
Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

In this model, they assume some no-arbitrage argument, constant interest rate, no tax and no transaction costs, short selling is possible.
The details of these assumptions are well explained by Hull(1990), Shreve(2001).
by risk neutral measure with market price of risk $\theta = \frac{r-\mu}{\sigma}$

$dS_t^\star = rS_t^\star dt + \sigma S_t^\star dB_t^\star$ GBM

Call Option price is

$$C_t = e^{-(T-t)r}\mathbb{E}^\star[(S_T^\star - K, 0)^+]$$

with $S_t^\star = S_0^\star e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t^\star}$ and its law is lognormal.

2. Pricing Formula

The Pricing formula is given

$$C_t = x\mathbf{N}(d_+(T-t, x)) - Ke^{-r(T-t)}\mathbf{N}(d_-(T-t, x))$$

where

$$d_\pm(\tau, x) = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\frac{x}{K} + (r \pm \frac{1}{2}\sigma^2)\tau\right]$$

$$\mathbf{N}(z) = \frac{1}{\sqrt{2\phi}}\int_{-\infty}^{z} e^{-y^2/2}dy$$

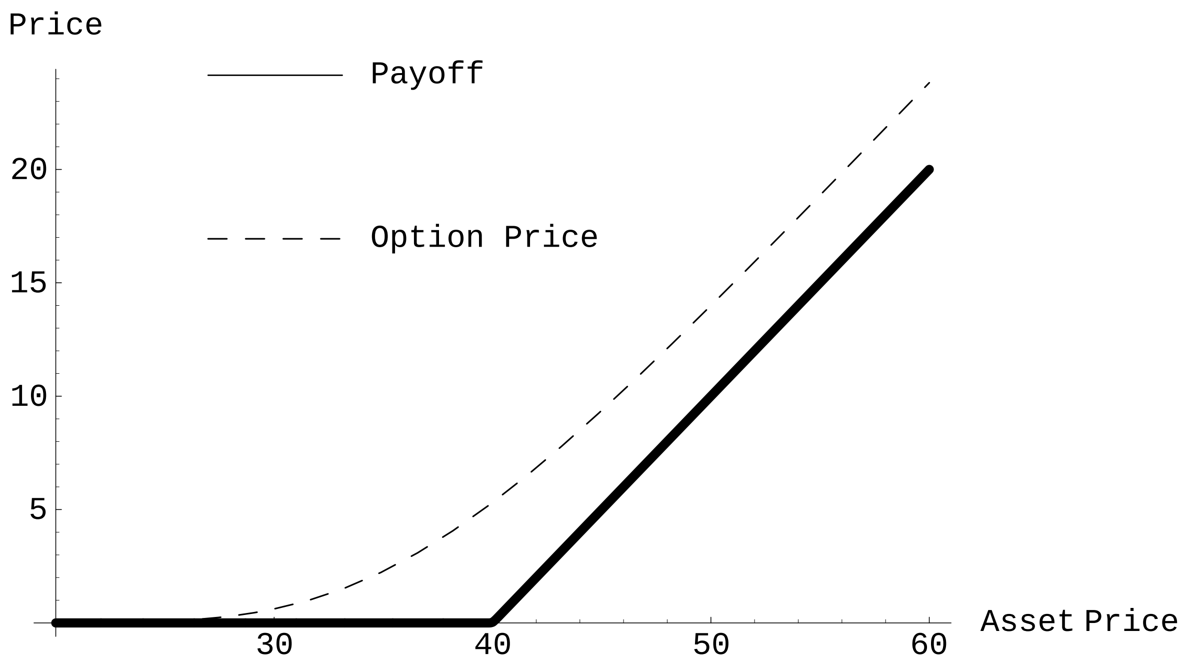

FIGURE 1. The Price of European Call Option on the stock price

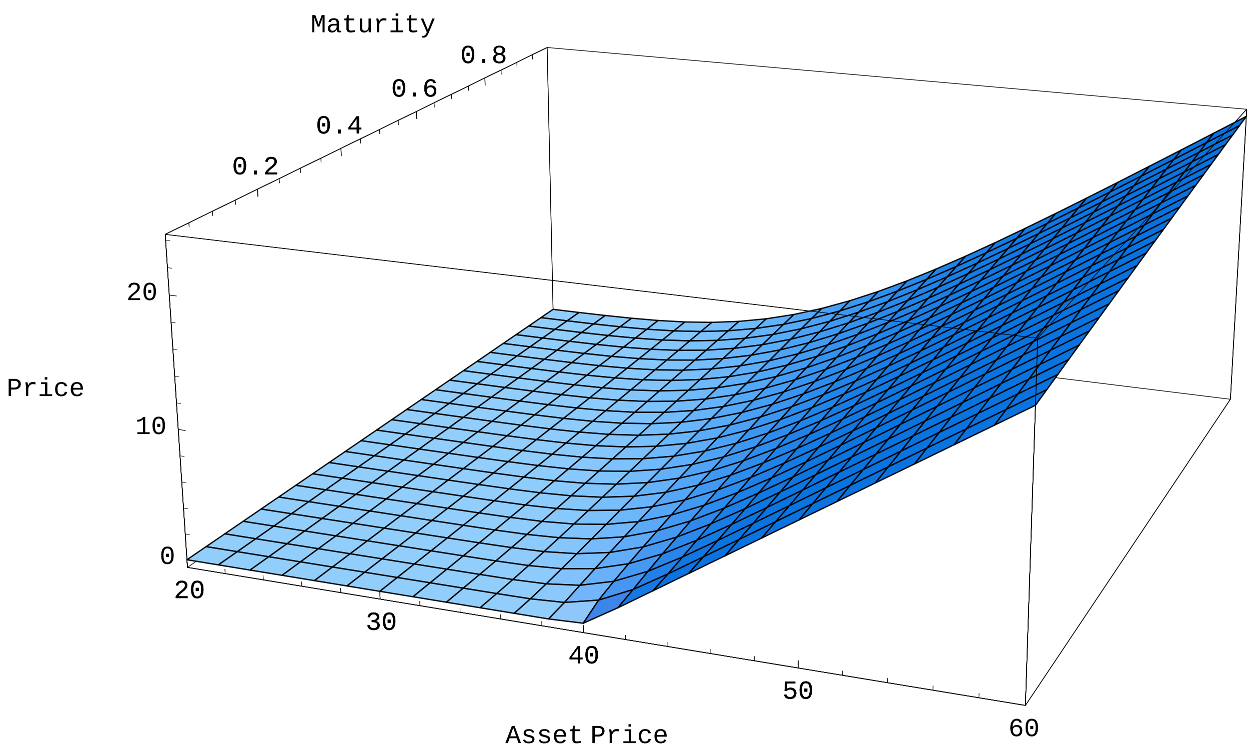

FIGURE 2. Numerical solution for the Black-Scholes European option pricing function with $\sigma = 0.2$

3. Greeks

3.0.1. *Delta* Δ.

3.0.2. *Gamma* Γ.

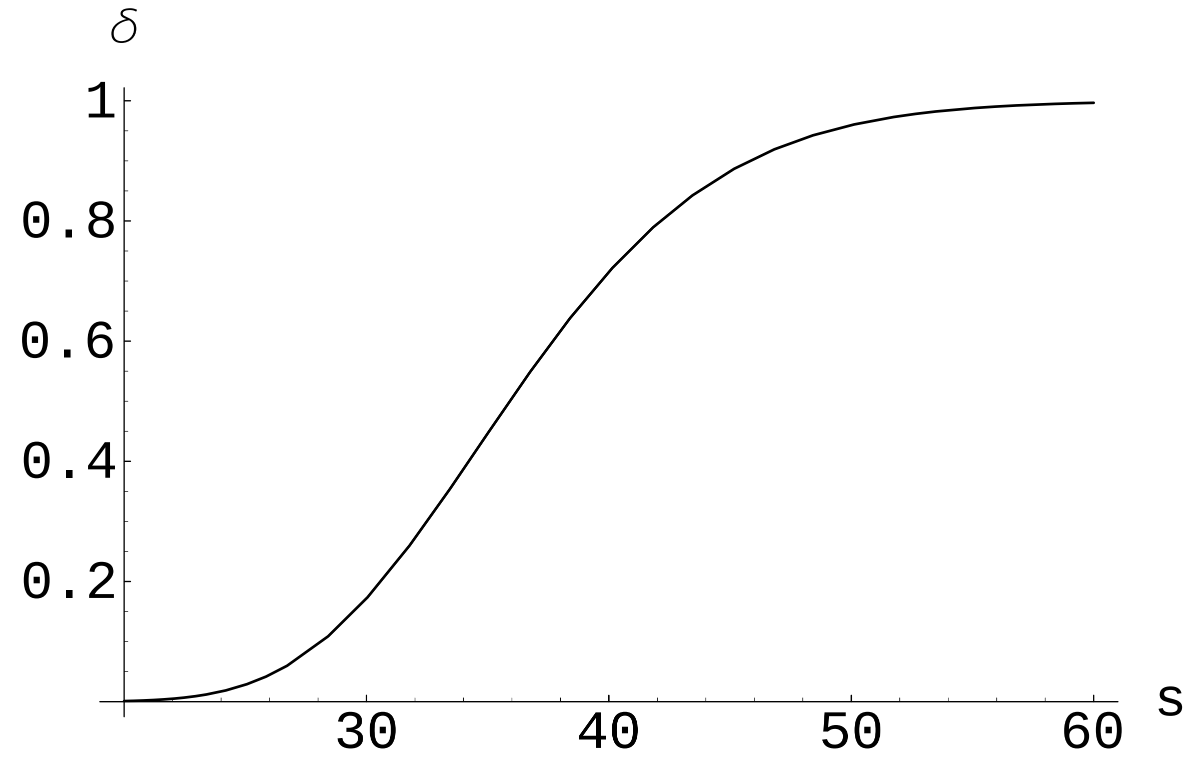

FIGURE 3. Δ on asset price

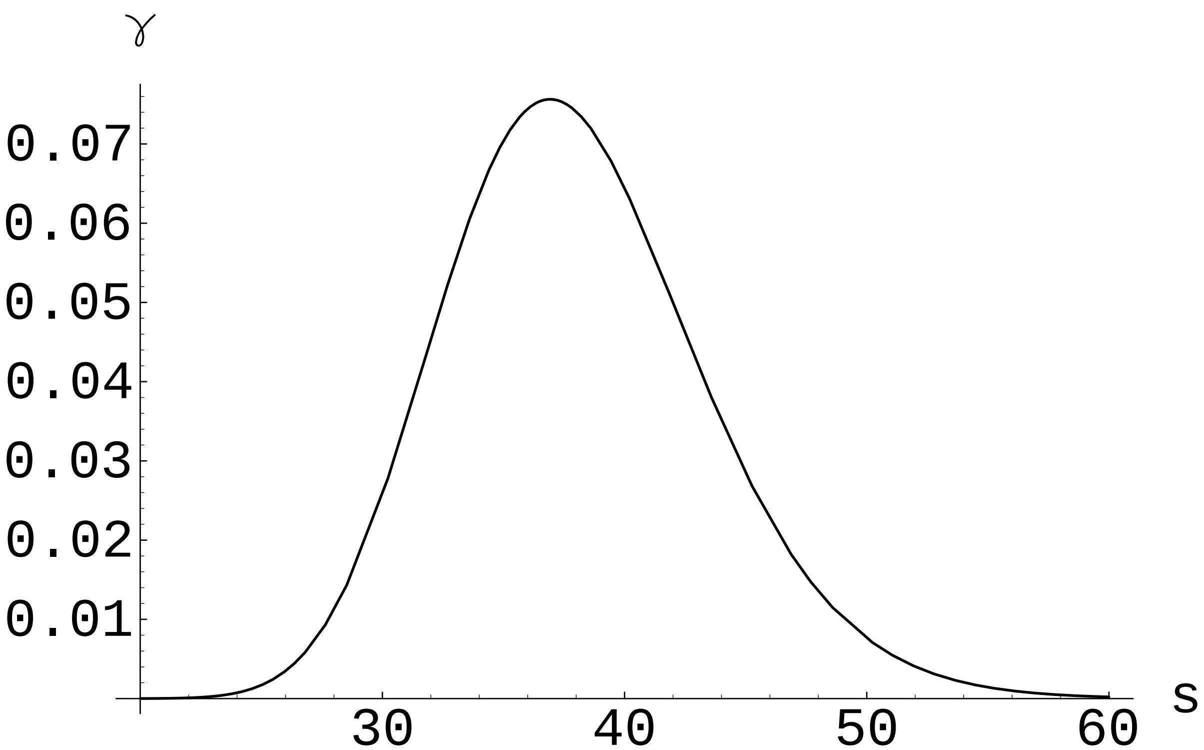

FIGURE 4. Γ on asset price

3.0.3. *Vega* ν.

3.0.4. *Rho* ρ.

3.0.5. *Theta* θ.

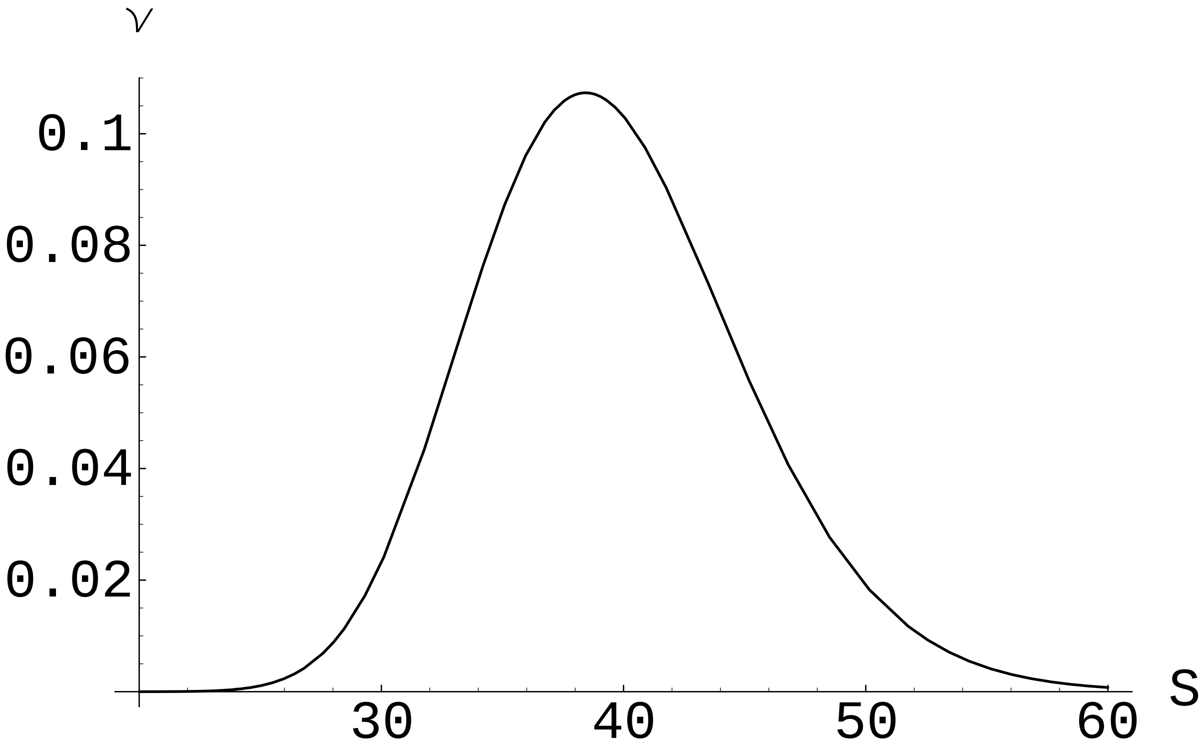

FIGURE 5. ν on asset price

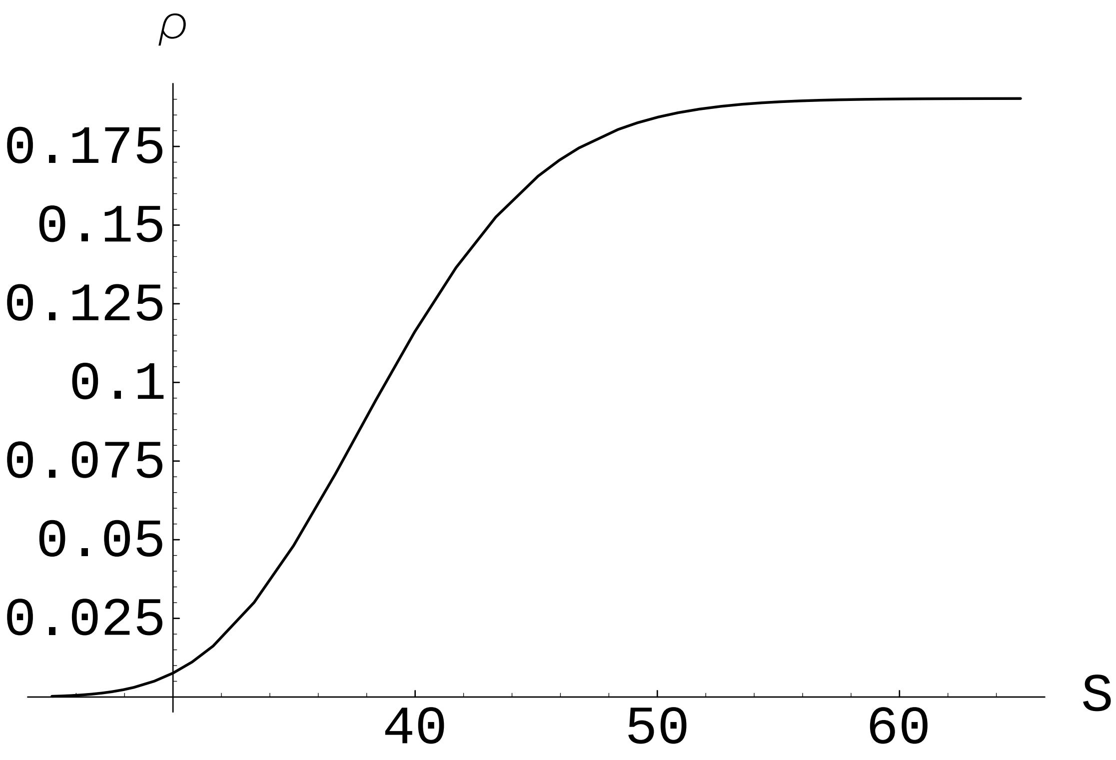

FIGURE 6. ρ on asset price

4. Exotic Option pricing : Powered call option

Let consider asset γ powered process S_t with riskless asset β_t

$$dS_t^\gamma = \mu S_t^\gamma dt + \sigma S_t^\gamma dB_t$$

$$d\beta_t = rdt$$

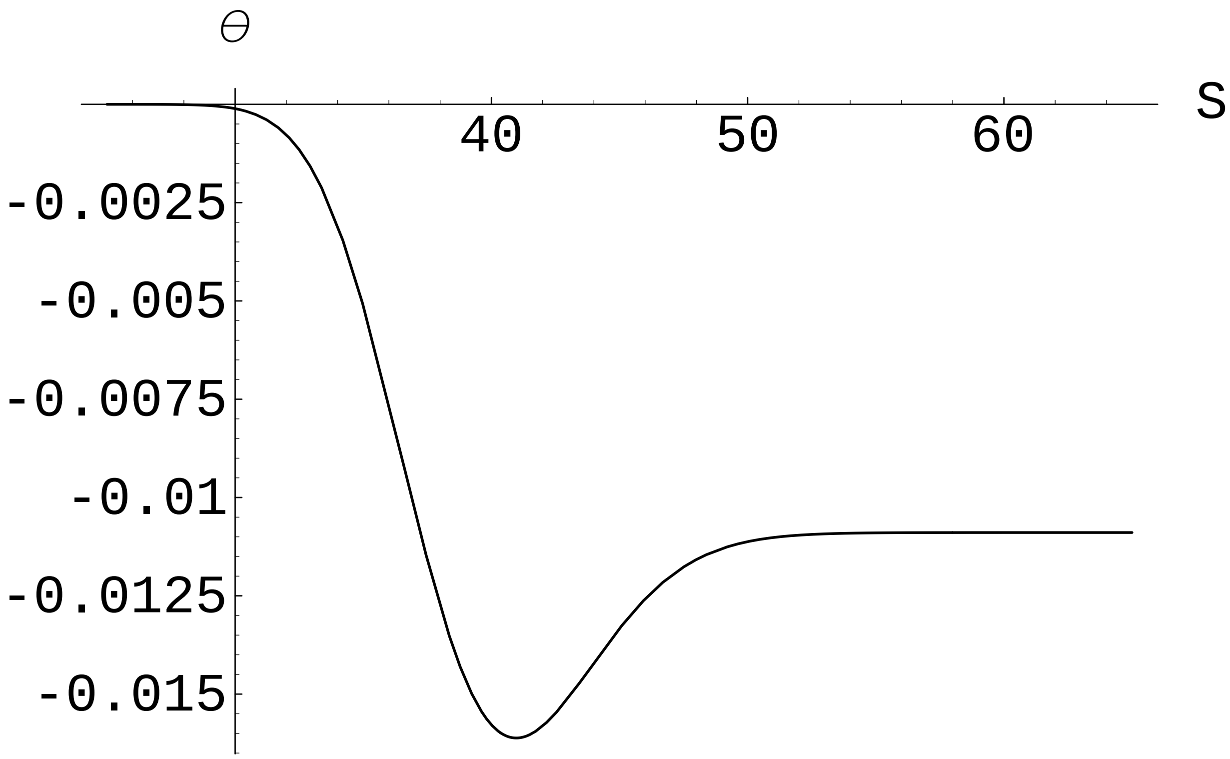

FIGURE 7. θ on asset price

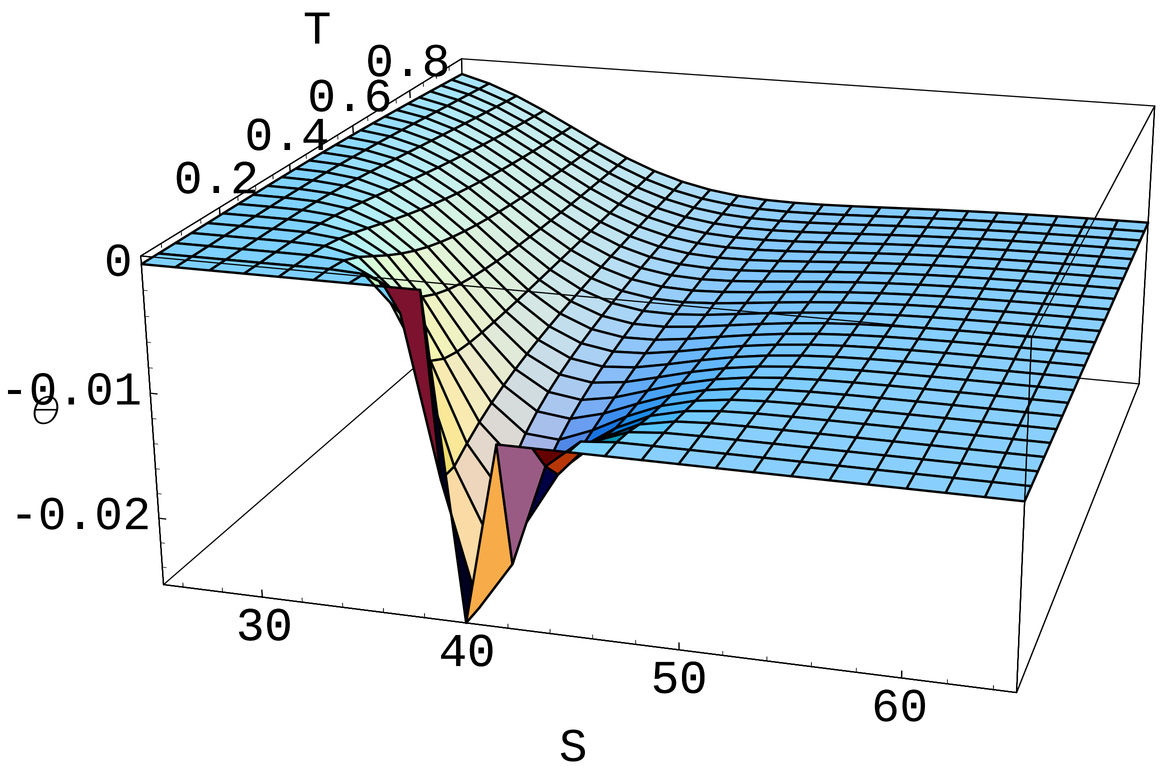

FIGURE 8. θ on asset price and maturity

European Powered Call Options
Payoff Function is $H(T, S_T) = \max[S_T^{\gamma} - K, 0]$

CHAPTER 6

Monte Carlo simulation

0.1. Introduction to Monte Carlo Simulation.

Definition 0.1. simple Monte Carlo Simulation We can get the integration by many trials $X_1, \cdots, X_N$

$$\mathbb{E}[X] = \frac{1}{N}\sum_{i=1}^{N} X_i$$

In BS option pricing model, for $i = 1, \cdots, N$

$$\begin{aligned} C_t &= e^{-(T-t)r}\mathbb{E}^\star[(S_T^\star - K, 0)^+] \\ &= e^{-(T-t)r}\frac{1}{N}\sum_{i=1}^{N}(S_{T,i}^\star - K, 0)^+ \end{aligned}$$

with $S_t^\star = S_0^\star e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t^\star}$ in this model law of $B_t^\star$ is normal distribution, therefore we can easily simulate $S_t^\star$ by random number generating.

0.2. Quasi Monte Carlo Simulation.

0.3. Variance Reduction Method.

0.4. Control Variate Method.

CHAPTER 7

Various Approach for fitting to Empirical Data

After 1983 of Stock Market Crisis, introduction of stochastic volatility

1. Empirical Evidence for Option price did not follow BS formula

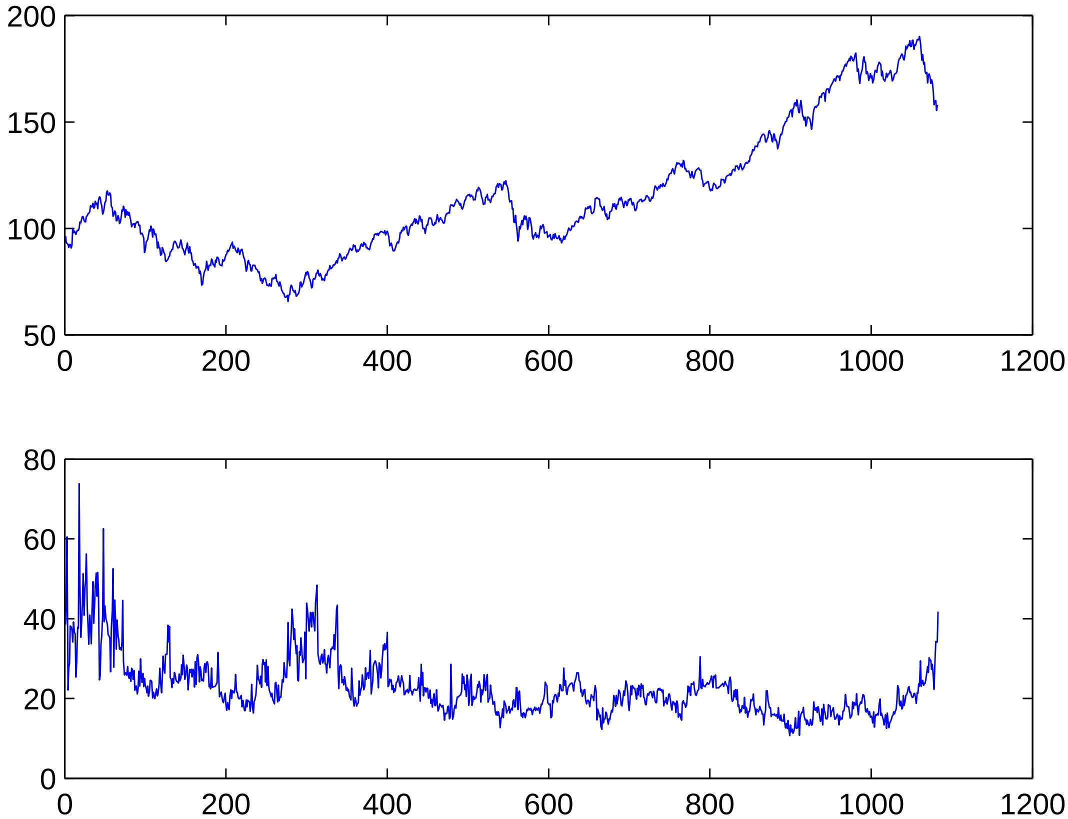

FIGURE 1. historical chart of KOSPI 200 data from Jan. 20 2000 to Aug. 17 2007 and its volatilities

1.1. Fat-tail of distribution. In Black-Scholes Model, they assume the underlying asset's dynamics follows geometric brownian motion, that means its log return is normal distribution. But, empirical evidence revealed that the fat-tail.

1.2. Skewness of Volatility Curves.

2. Constant Elasticity Volatility model

Let consider Constant elasticity of variance process as asset price S_t

$$dS_t = \mu S_t dt + \sigma S_t^{\alpha\backslash 2} dB_t$$

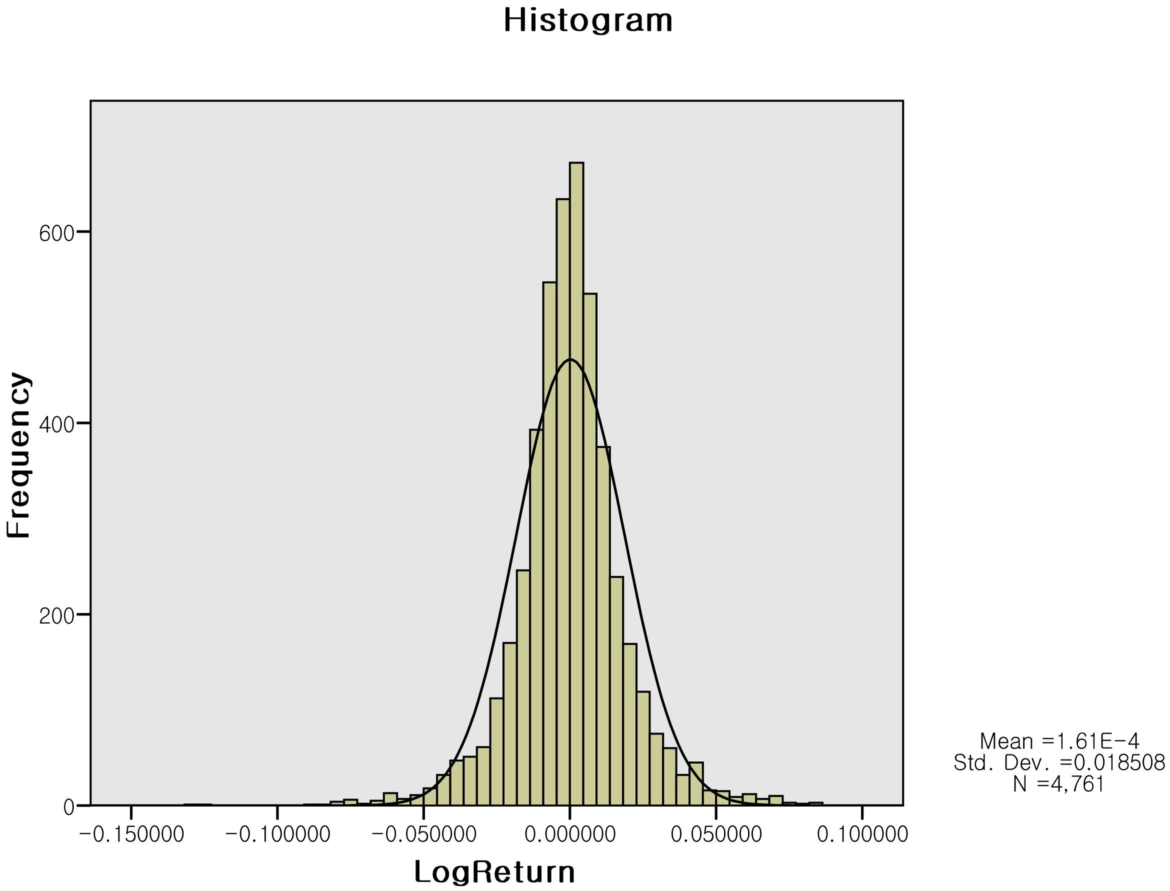

FIGURE 2. historical chart of KOSPI 200 data from Jan. 20 2000 to Aug. 17 2007 and its volatilities

and Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

for solving SDE, let $Y = S_t^{2-\alpha}$. then by ito lemma with $\frac{\partial Y}{\partial t} = 0$, $\frac{\partial Y}{\partial S} = (2-\alpha)S_t^{1-\alpha}$, $\frac{\partial^2 Y}{\partial S^2} = (2-\alpha)(1-\alpha)S_t^{-\alpha}$ we have

$$\begin{aligned}
dY_t &= Y_t dt + Y_s dS_t + \frac{1}{2}Y_{ss}(dS_t)^2 \\
&= 0dt + \left((2-\alpha)S_t^{1-\alpha}\right) dS_t + \frac{1}{2}\left((2-\alpha)(1-\alpha)S_t^{-\alpha}\right)(dS_t)^2 \\
&= \left((2-\alpha)S_t^{1-\alpha}\right)\left(\mu S_t dt + \sigma S_t^{\alpha\backslash 2} dB_t\right) + \frac{1}{2}\left((2-\alpha)(1-\alpha)S_t^{-\alpha}\right)(\mu S_t dt + \sigma S_t^{\alpha\backslash 2} dB_t)^2 \\
&= \left[\mu(2-\alpha)S_t^{2-\alpha} + \frac{1}{2}\sigma^2(\alpha-1)(\alpha-2)\right] dt + \left[\sigma(2-\alpha)S_t^{\frac{2-\alpha}{2}}\right] dB_t \\
dY_t &= \left(\mu(2-\alpha)Y_t + \frac{1}{2}\sigma^2(\alpha-1)(\alpha-2)\right) dt + \sigma(2-\alpha)\sqrt{Y_t}dB_t
\end{aligned}$$

This process is CIR process with $Y_t = S_t^{2-\alpha}$ Therefore, we can derive density function of Y_t by The Kolmogorov Forward Equation

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial Y}[\mu(2-\alpha)Y_t + \frac{1}{2}\sigma^2(\alpha-1)(\alpha-2)]f + \frac{1}{2}\frac{\partial^2}{\partial Y^2}[\sigma^2(2-\alpha)^2 Y_t]f$$

by solving PDE, we get the law of S_t then, we can get option price with it.

3. Stochastic Volatility Model with One-Factor model

We donote the price of the underlying by X_t and model it as the solution of the stochastic differential equation :

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t \tag{3.1}$$

$$\sigma_t = f(Y_t).$$

$$dY_t = \alpha(m - Y_t)dt + \beta d\hat{Z}_t$$

where is a standard Brownian motion, and its covariation with

$$d\langle W, \hat{Z}\rangle_t = \rho dt \tag{3.2}$$

therefore we use below for inconvince

$$\hat{Z}_t = \rho W_t + \sqrt{1-\rho^2} Z_t \tag{3.3}$$

W_t Z_t are independent Brownian motions so that

$$\begin{pmatrix} W_t \\ \hat{Z}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} W_t \\ Z_t \end{pmatrix} \tag{3.4}$$

in terms of the small parameter ε the rate of mean reversion α or its inverse, the typical correlation time of (Y_t) which we will consider as a small quantity denoted by

$$\varepsilon = \frac{1}{\alpha} \tag{3.5}$$

In the OU case the variance ν^2 is given by in terms of ε

$$\beta = \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}} \tag{3.6}$$

The model can be rewritten under the small parameter ε :

$$dX_t^\varepsilon = \mu X_t^\varepsilon dt + f(Y_t^\varepsilon) X_t^\varepsilon dW_t \tag{3.7}$$

$$dY_t^\varepsilon = \frac{1}{\varepsilon}(m - Y_t^\varepsilon)dt + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}} d\hat{Z}_t \tag{3.8}$$

3.1. Risk Neutral Measure. The market chooses one of these for pricing and we write next the stochastic differential equations that model this choice in terms of the following two-dimensional standard Brownian motion under the risk-neutral measure:

$$W_t^\star = W_t + \int_0^t \frac{\mu - r}{f(Y_s)} ds$$

$$Z_t^\star = Z_t + \int_0^t \gamma(Y_s) ds$$

where we assume that $\gamma(y)$ is smooth bounded functions of y only. We introduce the combined market prices of volatility risk Λ defiend by

$$\Lambda(y) = \frac{\rho(\mu - r)}{f(y)} + \gamma(y)\sqrt{1-\rho^2} \tag{3.9}$$

The model can be rewritten under risk-neutral probability measure as

$$dX_t^\varepsilon = rX_t^\varepsilon dt + f(Y_t^\varepsilon)X_t^\varepsilon dW_t^\star \tag{3.10}$$

$$dY_t^\varepsilon = \big[\frac{1}{\varepsilon}(m - Y_t^\varepsilon) + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}\Lambda(Y_t^\varepsilon)\big]dt + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}d\hat{Z}_t^\star \tag{3.11}$$

we also write

$$\hat{Z}_t^\star = \rho W_t^\star + \sqrt{1-\rho^2}Z_t^\star \tag{3.12}$$

with $|\rho| < 1$, where $W_t^\star$ $Z_t^\star$ are two independent standard Brownian motions under $I\!P^{\star(\gamma)}$

3.2. Assymptotics. We denote this price by$P^\varepsilon(t,x,y)$, and we know that by risk-neutral pricing,

$$P^\varepsilon(t,x,y) = I\!E^{\star(\gamma)}\big\{e^{-r(T-t)}h(X_T^\varepsilon)|X_t^\varepsilon = x, Y_t^\varepsilon = y\big\} \tag{3.13}$$

by Feynman-Kac, partial differential equation for (15) is below

$$\begin{aligned}\frac{\partial P^\varepsilon}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2 P^\varepsilon}{\partial x^2} + \frac{\rho\nu\sqrt{2}}{\sqrt{\varepsilon}}xf(y)\frac{\partial^2 P^\varepsilon}{\partial x\partial y} + \frac{\nu^2}{\varepsilon}\frac{\partial^2 P^\varepsilon}{\partial y^2}\\ + r\big(x\frac{\partial P^\varepsilon}{\partial x} - P^\varepsilon\big) + \big[\frac{1}{\varepsilon}(m - Y_t^\varepsilon) + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}\Lambda(Y_t^\varepsilon)\big]\frac{\partial P^\varepsilon}{\partial y} = 0\end{aligned} \tag{3.14}$$

which has to be solved for $t < T$ with the terminal condition

$$P^\varepsilon(T,x,y) = h(x) \tag{3.15}$$

The partial differential equation involves terms of order $\mathcal{O}(1/\varepsilon), \mathcal{O}(1/\sqrt{\varepsilon})$, and $\mathcal{O}(1)$. In order to account for these three different orders, we introduce the following convenient notation :

$$\mathcal{L}^\varepsilon := \frac{1}{\varepsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}_1 + \mathcal{L}_2 \text{ and} \tag{3.16}$$

$$\mathcal{L}_0 := (m-y)\frac{\partial}{\partial y} + \nu^2\frac{\partial^2}{\partial y^2} \tag{3.17}$$

$$\mathcal{L}_1 := \sqrt{2}\nu\rho x f(y)\frac{\partial^2}{\partial x\partial y} - \sqrt{2}\nu\Lambda(y)\frac{\partial}{\partial y} \tag{3.18}$$

$$\mathcal{L}_2 := \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2}{\partial x^2} + r\big(x\frac{\partial}{\partial x} - \cdot\big) \tag{3.19}$$

$$\tag{3.20}$$

With this notatoin, the pricing partial differential equation becomes

$$\left(\frac{1}{\varepsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}_1 + \mathcal{L}_2\right)P^\varepsilon = 0 \tag{3.21}$$

with terminal conditoin

$$P^\varepsilon(T,x,y) = h(x) \tag{3.22}$$

The method is to expand the solution P^ε in power of $\sqrt{\varepsilon}$,

$$P^\varepsilon = P_0 + \sqrt{\varepsilon}P_1 + \varepsilon P_2 + \varepsilon\sqrt{\varepsilon}P_3 + \varepsilon^2 P_4 + \cdots, \tag{3.23}$$

where $P_0, P_1, ...$ are functin of (t,x,y) to be determined such that $P_0(T,x,y) = h(x)$. We are primarily interested in ther first two terms $P_0 + \sqrt{\varepsilon}P_1$. The terminal condition for the second term is $P_1(T,x,y) = 0$. Substituting (28) into (23) lead to

$$
\begin{aligned}
\frac{1}{\varepsilon}\mathcal{L}_0P_0 + \frac{1}{\sqrt{\varepsilon}}&(\mathcal{L}_0P_1 + \mathcal{L}_1P_0)\\
&+ (\mathcal{L}_0P_2 + \mathcal{L}_1P_1 + \mathcal{L}_2P_0)\\
&+ \sqrt{\varepsilon}(\mathcal{L}_0P_3 + \mathcal{L}_1P_2 + \mathcal{L}_2P_1)\\
&+ \cdots\\
&= 0
\end{aligned} \tag{3.24}
$$

3.2.1. *Deriving terms.* equation terms of $\mathcal{O}(1/\varepsilon)$ we must have

$$\mathcal{L}_0P_0 = 0 \tag{3.25}$$

The operator $\mathcal{L}_0$ given by (20) is the generator of an ergodic Markov process and acts only on the y variable; hence P_0 must be a constant with respect to that variable, which implies that

$$P_0 = P_0(t, x) \tag{3.26}$$

a function of (t, x) only.

Similarly, in order to eliminate the terms in $\mathcal{O}(1/\sqrt{\varepsilon})$, we must have

$$\mathcal{L}_0P_1 + \mathcal{L}_1P_0 = 0 \tag{3.27}$$

The operator $\mathcal{L}_1$ given by (21) takes derivatives with respect to y, and we therefore deduce from (31) that $\mathcal{L}_1P_0 = 0$ and consequently that $\mathcal{L}_0P_1 = 0$. Using the same argument as for (30), it is clear that

$$P_1 = P_1(t, x) \tag{3.28}$$

a function of (t, x) only. This implies in particular that the combination of the first two terms $P_0 + \sqrt{\varepsilon}P_1$ will not depend on the present volatility

The equation terms of $\mathcal{O}(1)$ give:

$$\mathcal{L}_0P_2 + \mathcal{L}_1P_1 + \mathcal{L}_2P_0 = 0 \tag{3.29}$$

3.2.2. *Poisson equation.* From (33) we know that $\mathcal{L}_1P_1 = 0$, so this equation reduced to

$$\mathcal{L}_0P_2 + \mathcal{L}_2P_0 = 0 \tag{3.30}$$

The variabvle x being fixed, $\mathcal{L}_2P_0$ is a function of y since $\mathcal{L}_2$ involves $f(y)$. Focusing on the y dependence only, equation (35) is of the form

$$\mathcal{L}_0\chi + g = 0 \tag{3.31}$$

which is known as a Poisson equation for $\chi(y)$ with respect to the operator $\mathcal{L}_0$ in the variable y. This equation does not have a solution unless the function $g(y)$ is centered with respect to the invariant distribution of the Markov process Y whose infinitesimal generator is $\mathcal{L}_0$.

We donote the invariant distribution of Y by

$$\Phi(y) = \frac{1}{\sqrt{2\pi\nu}}e^{-(y-m)^2/2\nu^2} \tag{3.32}$$

The centeringf condition

$$\langle g\rangle = \int g(y)\Phi(y)dy = 0 \tag{3.33}$$

is necessary for poisson equation to admit a solution, as can be seen from the following calculation using integral by part and use the definition of the adjoint operator $\mathcal{L}_0^\star$ and its property $\mathcal{L}_0^\star\Phi(y) = 0$ Because the OU process has invariant measure

In our situation, the centering condition in equation gives

$$\langle \mathcal{L}_0 P_0 \rangle = 0 \tag{3.34}$$

since P_0 does not depend on y, this is $\langle \mathcal{L}_0 \rangle P_0 = 0$ and, from the definition of $\langle \mathcal{L}_2 \rangle = \mathcal{L}_{BS}(\bar{\sigma})$ where the effective volatility $\bar{\sigma}$ is defined by $\bar{\sigma} = \langle f^2 \rangle$

3.2.3. *zero-order term.* Therefore, the zero-order term $P_0(t,x)$is the solution of the Black-Scholes equation

$$\mathcal{L}_{BS}(\bar{\sigma}) P_0 = 0 \tag{3.35}$$

with the terminal condition $P_0(T,x) = h(x)$.

As the centering condition is satisfied we can write

$$\begin{aligned} \mathcal{L}_2 P_0 =& \mathcal{L}_2 P_0 - \langle \mathcal{L}_2 P_0 \rangle \\ =& \frac{1}{2}(f(y)^2 - \bar{\sigma}^2) x^2 \frac{\partial^2 P_0}{\partial x^2} \end{aligned} \tag{3.36}$$

The second-order correction P_2, solution of the Poisson equation $\mathcal{L}_0 P_2 + \mathcal{L}_2 P_0 = 0$, is then given by

$$\begin{aligned} \mathcal{L}_0 P_2 &= -\mathcal{L}_2 P_0 \\ &= -(\mathcal{L}_2 P_0 - \langle \mathcal{L}_2 P_0 \rangle) \\ &= -\frac{1}{2}(f(y)^2 - \bar{\sigma}^2) x^2 \frac{\partial^2 P_0}{\partial x^2} \end{aligned} \tag{3.37}$$

therefore

$$\begin{aligned} P_2(t,x,y) =& -\frac{1}{2} \mathcal{L}_0^{-1} (f(y)^2 - \bar{\sigma}^2) x^2 \frac{\partial^2 P_0}{\partial x^2} \\ =& -\frac{1}{2}(\phi(y) + c(t,x)) x^2 \frac{\partial^2 P_0}{\partial x^2} \end{aligned} \tag{3.38}$$

where $\phi(y)$ is a solution of the poisson equation

$$\mathcal{L}_0 \phi(y) = f(y)^2 - \langle f(y)^2 \rangle \tag{3.39}$$

and $c(t,x)$ is a constant in y that may depend on (t,x).

3.3. first correction. The equation terms of $\mathcal{O}(\sqrt{\varepsilon})$ give:

$$\mathcal{L}_0 P_3 + \mathcal{L}_1 P_2 + \mathcal{L}_2 P_1 = 0 \tag{3.40}$$

This is again a Poisson equation for P_3 with respect to $\mathcal{L}_0$, which requires the centering or solvability condition

$$\langle \mathcal{L}_1 P_2 + \mathcal{L}_2 P_1 \rangle = 0 \tag{3.41}$$

Using the computation of P_2, the fact that P_1 does not depend on y and $\langle \mathcal{L}_2 \rangle = \mathcal{L}_{BS}(\bar{\sigma})$ where the effective volatility $\bar{\sigma}$ is defined by $\bar{\sigma} = \langle f^2 \rangle$, we deduce that

$$\mathcal{L}_{BS}(\bar{\sigma}) P_1 = \frac{1}{2} \langle \mathcal{L}_1 \phi(y) \rangle x^2 \frac{\partial^2 P_0}{\partial x^2} \tag{3.42}$$

we then check that it is given by

$$\begin{aligned} \langle \mathcal{L}_2 \rangle P1 =& \mathcal{L}_{BS}(\bar{\sigma}) P_1 \\ =& \frac{1}{2} \langle \mathcal{L}_1 \phi(y) \rangle x^2 \frac{\partial^2 P_0}{\partial x^2} \end{aligned} \tag{3.43}$$

Notice that $\mathcal{L}_1 c = 0$, since $\mathcal{L}_1$ take derivatives with respect to y and $c(t,x)$ is independent of y. Using definition of $\mathcal{L}_1$, one can computed the operator

$$\begin{aligned}\langle \mathcal{L}_1 \phi(y) \cdot \rangle =& \langle \mathcal{L}_1 \phi(y) \cdot \rangle \\ =& (\sqrt{2}\nu\rho \langle f(y)\phi(y)' \rangle x \frac{\partial}{\partial x} - \sqrt{2}\nu \langle \Lambda(y)\phi(y)' \rangle \cdot\end{aligned} \tag{3.44}$$

the sixth equality is given by $P_1(t,x), P_0(t,x)$ does not depend on y. and finally derived the equation for $P_1(t,x)$:

$$\mathcal{L}_{BS}(\bar{\sigma})P_1 = \frac{\sqrt{2}}{2}\nu\rho\langle f\phi' \rangle x^3 \frac{\partial^3 P_0}{\partial x^3} + \left(\sqrt{2}\nu\rho\langle f\phi' \rangle - \frac{\sqrt{2}}{2}\nu\langle \Lambda\phi' \rangle \right) \frac{\partial^2 P_0}{\partial x^2} \tag{3.45}$$

with the terminal condition $P_1(T,x) = 0$.

At this stage it is convenient to introduce notation for the first small correction, for approximation of $P^\varepsilon = P_0 + \sqrt{\varepsilon}P_1 + \varepsilon P_2 + \varepsilon\sqrt{\varepsilon}P_3 + \varepsilon^2 P_4 + \cdots$ are below:

$$\tilde{P}_1(t,x) := \sqrt{\varepsilon}P_1(t,x) \tag{3.46}$$

which is the solution of

$$\mathcal{L}_{BS}(\bar{\sigma})\tilde{P}_1 = \mathcal{H}(t,x)P_0 \tag{3.47}$$

where the first correction operator $\mathcal{H}(t,x)$ defined by

$$\mathcal{H}(t,x) := V_2 x^2 \frac{\partial^2}{\partial x^2} + V_3 x^3 \frac{\partial^3}{\partial x^3} \tag{3.48}$$

V_2 V_3 are two small coefficients, given in terms of $\alpha = 1/\varepsilon$ by

$$V_2 = \frac{\nu}{\sqrt{2\alpha}}(2\rho\langle f\phi' \rangle - \langle \Lambda\phi' \rangle) \tag{3.49}$$

$$V_3 = \frac{\rho\nu}{\sqrt{2\alpha}}\langle f\phi' \rangle \tag{3.50}$$

The first correction satisfies the inhomogenous Black-Scholes PDE equation with a zero terminal condition and a small source term computed from derivatives of the leading term $P_0(t,x)$. the solution of $\tilde{P}_1$ is explicitly given by

$$\begin{aligned}\tilde{P}_1 =& -(T-t)\left(V_2 \frac{\partial^2}{\partial x^2} + V_3 x^3 + V_3 \frac{\partial^3}{\partial x^3} \right) P_0 \\ =& -(T-t)\mathcal{H}(t,x)P_0\end{aligned} \tag{3.51}$$

$$\mathcal{L}_{BS}(\bar{\sigma})\left(x^n \frac{\partial^n}{\partial x^n}(t,x)P_0 \right) = x^n \frac{\partial^n}{\partial x^n}\mathcal{L}_{BS}(\bar{\sigma})P_0 \tag{3.52}$$

for any positive integer n

Lastly, the corrected price is given explicitly by

$$\begin{aligned}P^\varepsilon &\approx \tilde{P}_1 \\ &= -(T-t)\left(V_2 \frac{\partial^2}{\partial x^2} + V_3 x^3 + V_3 \frac{\partial^3}{\partial x^3} \right) P_0 \\ &= P_0 - (T-t)\mathcal{H}(t,x)P_0\end{aligned} \tag{3.53}$$

where $P_0 = P_{BS}(\bar{\sigma})$,the Black-Scholes price with constant volatility $\bar{\sigma} = \langle f(y) \rangle$

4. Jump Diffusion Model

Let consider asset S_t with jump Y_t

$$dS_t = \mu S_t dt + \sigma S_t dB_t^{(1)} + dY_t$$
$$Y_t := \quad \text{jump process}$$

where Y_t as poisson process, compound process, cox process etc.
and Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

5. Stochastic Volatility Model with Two-Factor model

5.1. Fast scale volatility factor.

5.2. Slow scale volatility factor.

5.3. Risk Neutral Measure.

5.4. Asymptotics.

5.4.1. *Deriving terms.*
5.4.2. *Expansion in the fast scale.*
5.4.3. *Expansion of P_1^{ε}.*

5.5. The first Correction.

5.6. Jump Diffusion Model. Let consider asset S_t with jump Y_t

$$dS_t = \mu S_t dt + \sigma S_t dB_t^{(1)} + dY_t$$
$$Y_t := \quad \text{jump process}$$

where Y_t as poisson process, compound process, cox process etc.
and Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

we can but we must find generator of Y_t to get option price with jump process.

5.7. lèvy process Model. Let consider asset S_t as lèvy process.
CGMY process , Variance Gamma process , α-stable process is famous lèvy process.
and Payoff Function is $H(T, S_T) = \max[S_T - K, 0]$

CHAPTER 8

Financial Application : Term Structure Model

1. Classical Model

Forward rate
Vasicek Model : Check on Calculation of OU process

1.1. HJM framework.

1.2. LIBOR market model.

CHAPTER 9

Financial Application : Credit Risk Modeling

0.3. default.

0.4. hazard rate.

0.5. cox process.

0.6. CDS.

0.7. copula.

CHAPTER 10

stochastic optimal control

0.8. kalman filter.

0.9. optimal stopping.

0.10. optimal control.

0.11. portfolio optimization.