1 Fundamental contents

1. Understand the abstract concepts: random experiment, sample space, σ-algebra, probability measure, probability space, filtration, random variables, stochastic processes, Markov process, martingale process, Random walk, Brownian motion, Poisson process, Geometric Brownian motion, Ito lemma, Ito integral, solving stochastic differential equation

2. Chapter 1: basic about probability space: random experiment, sample space, σ-algebra, probability measure, probability space, filtration, random variables, probability distribution, joint distribution, conditional distribution, expectation, variance, covariance. Focusing on Bernoulli, uniform, Binomial, Negative binomial, geometric, normal, Poisson, exponential distribution.

3. Chapter 2: Introduction to stochastic processes: discrete and continuous random process, filtration, filtered probability space, adapted process, Markov process and Poisson process.

4. Chapter 3: Brownian motion. Random walk, Brownian motion: maximum of Brownian, zeros of Brownian motion, first passage time distribution, Brownian with drift, simulating, Brownian bridge... Martingale process

5. Chapter 4: Ito's integral definition, Ito's formula, derive solutions for Arithmetic Brownian motion, geometric Brownian motion, Ornstein-Ulenbeck Process

2 Fundamental skills

1. Compute the probability of an event, the expectation variance and covariance of common random variables: binomial, normal, Poisson, exponential random variables

Distribution	Probability mass function
Bernoulli	$P(X=k)=\begin{cases} p & \text{if } k=1 \\ 1-p & \text{if } k=0 \end{cases}$
Binomial	$P(X=k)=C_n^k p^k(1-p)^{n-k}, k=0,...,n$
Poisson	$P(X=k)=e^{-\lambda}\dfrac{\lambda^k}{k!}, k\in\mathbb{N}$
Uniform	$P(X=k)=\dfrac{1}{n}, k=1,2,...,n$

Distribution	Probability density function
Uniform	$f(x)=\dfrac{1}{b-a}, \forall\ a\le x\le b$
Exponential	$f(x)=\lambda e^{-\lambda x}, x>0$

Table 1: Some common discrete and continuous distributions

Let X be a random variable defined on a probability space $(\Omega,\mathcal{F},\mathcal{P})$. The unconditional expectation of X is

$$E(X)=\begin{cases} \displaystyle\sum_{k=1}^{\infty} x_k P(X=x_k), \text{if } X \text{ is discrete with values } \{x_1,x_2,...\} \\ \displaystyle\int_{-\infty}^{\infty} x f(x)dx, \text{if } X \text{ is continuous with density function } f \end{cases}$$

Let g is a real-valued function. $E(g(X)) = \begin{cases} \sum_{k=1}^{\infty} g(x_k)P(X = x_k), \\ \int_{-\infty}^{\infty} g(x)f(x)dx \end{cases}$ The unconditional variance of X as

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

Let $Z = g(X, Y)$ be a function of two random variables.

- $E[g(X,Y)] = \begin{cases} \sum_{x \in \mathbf{S}} \sum_{y \in \mathbf{T}} g(x,y)\mathbf{P}(X = x, Y = y) \\ \int_{x \in \mathbf{S}} \int_{y \in \mathbf{T}} f(x,y)g(x,y)dydx. \end{cases}$
- Let X and Y be independent random variables. Then for any functions g and h, $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$.

Distribution	Probability mass function	E	Var
Bernoulli	$P(X = k) = \begin{cases} p & \text{if } k = 1 \\ 1 - p & \text{if } k = 0 \end{cases}$	p	$p(1-p)$
Binomial	$P(X = k) = C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k \in \mathbb{N}$	λ	λ
Uniform	$P(X = k) = \frac{1}{n}, k = 1, 2, ..., n$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$

Distribution	Probability density function	E	Var
Uniform	$f(x) = \frac{1}{b-a}, \forall\ a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$f(x) = \lambda e^{-\lambda x}, x > 0$	$1/\lambda$	$1/\lambda^2$

Table 2: Some common continuous distributions

- $Cov(X, Y) = E[XY] - E[X]E[Y] = E[(X - E[X])(Y - E[Y])]$.
- $Cov(X, Y) = 0$ if X and Y are independent.
- $Corr(X, Y) = \frac{Cov(X, Y)}{SD[X]SD[Y]}$.
- For constants $a < b$ and $c < d$,

$$P(a \leq X \leq b, c \leq Y \leq d) = \begin{cases} \sum_{x=a}^{b} \sum_{y=c}^{d} \mathbf{P}(X = x, Y = y), \\ \int_a^b \int_c^d f(x,y)dydx \end{cases}$$

- if X and Y are discrete then the probability mass function P_X, P_Y are determined as: $P_X(X = x) = \sum_{y=-\infty}^{\infty} \mathbf{P}(X = x, Y = y)$ and $P_Y(Y = y) = \sum_{x=-\infty}^{\infty} \mathbf{P}(X = x, Y = y)$.

If X and Y are continuous then the probability density f_X, f_Y are determined by: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$.

Law of total probability

Let $B_1, ..., B_k$ be a sequence of events that partition the sample space. That is, the B_i are mutually exclusive (disjoint) and their union is equal to Ω. Then, for any event A, we have

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

2. Compute the conditional probability $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

   - If X and Y are jointly distributed discrete random variables, then the conditional probability mass function of Y given $X = x$ is $P(Y = y|X = x) = \dfrac{P(X = x, Y = y)}{P(X = x)}$ defined when $P(X = x) > 0$.
   - For continuous random variable X and Y, the conditional density function of Y given $X = x$ is

     $$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)},$$

     where f_X is the marginal density function of X, which is computed by the formula: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$.
   - $E(Y|X = x) = \begin{cases} \sum_y yP(Y = y|X = x), \text{if X is discrete} \\ \int_{-\infty}^{\infty} yf_{Y|X}(y|x)dy, \text{if X is continuous} \end{cases}$
   - $E(Y|X = x)$ is a function of x, i.e., the result depends on the value of x

   Law of total expectation

   Let $A_1, ..., A_k$ be a sequence of events that partition the sample space. We have $E(Y) = \sum_{i=1}^{k} E(Y|A_i)P(A_i)$.

3. Compute the probability and expectation of some events and find the long-term behavior of Markov process.
   Let S be a discrete set. A Markov chain is a sequence of random variables $X_0, X_1, ...$ taking values in S with the property that

   $$P(X_{n+1} = j|X_0 = x_0, ..., X_{n-1} = x_{n-1}, X_n = i) = P(X_{n+1} = j|X_n = i),$$

   for all $x_0, x_1, ..., x_{n-1}, i, j \in S$ and $n \geq 0$. The set S is the state space of the Markov chain.

   - If $X_n = i$, we say that the chain visits state i, or hits i, at time n.
   - A Markov chain is time-homogeneous if $P(X_{n+1} = j|X_n = i) = P(X_1 = j|X_0 = i) = P_{ij}$, for all $n \geq 0$.
   - Matrix P, with ij-th elements P_{ij}: one-step transition matrix.
   - For states i and j, $n \geq 1$, $P(X_n = j|X_0 = i)$ is the probability that the chain started at stage i and hits state j in n steps.
   - The n-step transition matrix of the Markov chain includes ij-th entries $P(X_n = j|X_0 = i)$.

- Let $X_0, X_1, ...$ be a Markov chain with transition matrix P. The matrix P^n is the n-step transition matrix of the chain, i.e., $P_{ij}^n = P(X_n = j|X_0 = i), \quad \forall\ n \geq 0, i, j$.
- Do not confuse P_{ij}^n, the ijth entry of the matrix P^n, with $(P_{ij})^n$, the number P_{ij} raised to nth power.
- $P(X_n = j) = \sum_i P(X_0 = i)P(X_n = j|X_0 = i) = \sum_i \alpha_i P_{ij}^n$.
- αP^n is the distribution of X_n and $P(X_n = j) = (\alpha P^n)_j$, the j-th element of $\alpha P^n, \ \forall j$.

Long-term behavior of a discrete Markov chain

- $P(X_n = j|X_0 = i) = \mathbf{P}_{ij}^n$
- $P(X_n = j) = (\alpha \mathbf{P}^n)_j$
- What happens if n increases to ∞? Does the distribution of X_n still depend on initial distribution α.
- If $\lim\limits_{n\to\infty} (\alpha \mathbf{P}^n)_j = \lambda_j, \forall\ \alpha$, then $\lambda = (\lambda_1, ..., \lambda_n, ...)$ is called the limiting distribution of the Markov chain.
- λ_j as the long-term probability that the chain hits state j
- If $\lim\limits_{n\to\infty} \mathbf{P}^n = A$ and all rows of A are the same with vector λ then λ is the limiting distribution.

Compute limiting distribution

(a) Check whether $\mathbf{P}$ is a regular matrix.

(b) Solve the linear equation system $\pi \mathbf{P} = \pi$.

(c) Conclude the solution π is the limiting distribution.

4. Poisson process

A Poisson process with parameter λ is a counting process $(N_t)_{t\geq 0}$ with the following properties:

(a) $N_0 = 0$.

(b) For all $t > 0$, N_t has a Poisson distribution with parameter λt.

(c) (Stationary increments) For all $s, t > 0, N_{t+s} - N_s$ has the same distribution as N_t. That is,

$$P(N_{t+s} - N_s = k) = P(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!},$$

for $k = 0, 1,$

(d) (Independent increments) For $0 \leq q < r \leq s < t, N_t - N_s$ and $N_r - N_q$ are independent random variables.

- The distribution of the number of arrivals in an interval depends only on the length of the interval
- The number of arrivals on disjoint intervals are independent random variables
- $E(N_t) = \lambda t$

5. How to prove a process be a Martingale process

Consider a continuous-time stochastic process $(X_t)_{t\geq 0}$ adapted to the filtration $\{\mathcal{F}_t\}$ and satisfies the condition $E(|X_t|) < \infty, \quad \forall t \geq 0$. If $E(X_t|\mathcal{F}_s) = X_s, \ \forall 0 \leq s < t$: martingale process (no tendency to fall or rise)

6. Brownian motion.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathcal{P})$ be a filtered probability space and let $(B_t)_{0\le t<\infty}$ be an adapted process of this space. The process $(B_t)_{0\le t<\infty}$ is called a standard Brownian motion if it satisfies the following properties:

(a) $B_0 = 0$

(b) Independent increments: $B_t - B_s$ is independent of $\mathcal{F}_s$, $0 \le s < t$. That means

$$P(B_t - B_s \le k|\mathcal{F}_s) = P(B_t - B_s \le k).$$

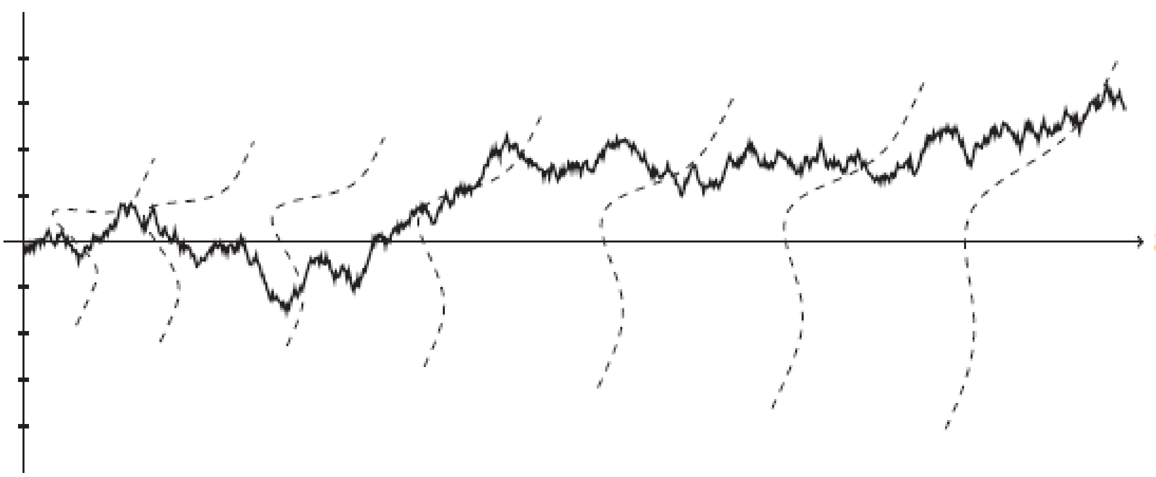
(c) Stationary increments:

$$B_t - B_s \stackrel{d}{=} B_{t-s} \sim N(0, \sqrt{t-s}),\ \forall\ 0 < s < t.$$

(d) Continuous paths: all sample paths of process $(B_t)_{t\ge 0}$ are almost surely continuous, i.e.

$$P(\omega \in \Omega | B_t(\omega) \text{ is a continuous sample path}) = 1.$$

Brownian motion can be thought of as the motion of a particle that diffuses randomly along a line. At each point t, the particle's position B_t is normally distributed about the line with variance t, i.e., $B_t \sim N(0, t)$. As t increases, the particle's position is more diffuse.



First passage time distribution

- The key reflection equality

$$P(\tau_m \le t, B_t \le w) = P(B_t \ge 2m - w), \quad \forall w \le m, m > 0.$$

$$P(\tau_m \le t, B_t \ge w) = P(B_t \le 2m - w), \quad \forall w \ge m, m < 0$$

- For all $m \ne 0$, the random variable τ_m has cumulative distribution function

$$P(\tau_m \le t) = \frac{2}{\sqrt{2\pi}} \int_{m/\sqrt{t}}^{\infty} e^{-y^2/2} dy$$

and the density function:

$$f_{\tau_m}(t) = \frac{d}{dt} P(\tau_m \le t) = \frac{|m|}{t\sqrt{2\pi t}} e^{-m^2/(2t)}, t \ge 0.$$

Brownian motion with drift

- For real value μ and $\sigma > 0$, the process defined by $W_t = \mu t + \sigma B_t$, for $t \ge 0$, is called Brownian motion with drift parameter μ and variance parameter σ^2.
- For $s, t > 0$, $W_{t+s} - W_t \sim N(\mu s, \sigma^2 s)$.
- $P(W_t - W_s \le k|\mathcal{F}_s) = P(\mu(t-s) + \sigma(B_t - B_s) \le k|\mathcal{F}_s) = P(B_t - B_s \le \dfrac{k - \mu(t-s)}{\sigma}|\mathcal{F}_s) = P(B_t - B_s \le \dfrac{k - \mu(t-s)}{\sigma}) = P(\mu(t-s) + \sigma(B_t - B_s) \le k) = P(W_t - W_s \le k)$.
- $W_t - W_s = \mu(t-s) + \sigma(B_t - B_s) \stackrel{d}{=} \mu(t-s) + \sigma B_{t-s} \stackrel{d}{=} W_{t-s},\ \forall\ 0 < s < t.$

7. Computing stochastic integral by definition

Let $\{(W_t)_{t\ge 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and let $\mathcal{F}_t$ be the associated filtration. We consider a function $f(W_t, t)$, which is $\mathcal{F}_t$-measurable (i.e., $f(W_t, t)$ is a stochastic process adapted to the filtration of a Brownian motion W_t) and is square-integrable $E[\int_0^t f(W_s, s)^2 ds] < \infty$, for all $t > 0$. The Ito integral of $f(W_t, t)$, with respect to the Brownian motion W_t is defined as:

$$I_t = \int_0^t f(W_s, s) dW_s = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i})$$

where $t_i = it/n, i = 0...n$.

The Stratonovich integral of $f(W_t, t)$, with respect to the Brownian motion W_t is defined as:

$$I_t = \int_0^t f(W_s, s) dW_s = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(\frac{W_{t_i} + W_{t_{i+1}}}{2}, t_i)(W_{t_{i+1}} - W_{t_i})$$

where $t_i = it/n, i = 0...n$.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let $\{W_t, t \ge 0\}$ be a standard Wiener process. Show that it has finite quadratic variation such that $< W, W >_t = \lim_{n\to\infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t$, where $t_i = it/n, i = 0...n$. Finally, deduce that $dW_t^2 = dt$.

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let $\{W_t, t \ge 0\}$ be a standard Wiener process. Show that the following cross-variation between W_t and t, and the quadratic variation of t are:
$\lim_{n\to\infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})(t_{i+1} - t_i) = 0$ and $\lim_{n\to\infty} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = 0$,
where $t_i = it/n$, $i = 0, ..., n$. Finally, deduce that $dW_t.dt = 0$ and $dt^2 = 0$.

The Ito integral has the following properties:
i) If $I_t = \int_0^t f(W_s, s)dW_s$ and $J_t = \int_0^t g(W_s, s)dW_s$ then

$$I_t \pm J_t = \int_0^t [f(W_s, s) \pm g(W_s, s)]dW_s$$

and for a constant c, $cI_t = \int_0^t cf(W_s, s)dW_s$.
ii) I_t is a martingale.
iii) $E[I_t^2] = E(\int_0^t f(W_s, s)^2 ds)$.

Let $\{(W_t)_{t\geq 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and let $\mathcal{F}_t$ be the associated filtration. We consider a function $f(W_t, t)$, which is $\mathcal{F}_t$-measurable (i.e., $f(W_t, t)$ is a stochastic process adapted to the filtration of a Brownian motion W_t) and is square-integrable $E[\int_0^t f(W_s, s)^2 ds] < \infty$, for all $t > 0$. The Ito integral of $f(W_t, t)$, with respect to the Brownian motion W_t is defined as:

$$I_t = \int_0^t f(W_s, s)dW_s = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i})$$

where $t_i = it/n, i = 0...n$.

8. Computing Ito-Doeblin formula for Ito process

For one-variable functions, their second-order Taylor series expansions have the form:

$$df(x) = f'(x)dx + \frac{1}{2}f''(x)dx^2.$$

Substituting $x = X_t$ in the above formula, we obtain:

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)dX_t^2 = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(\Delta_t dW_t + \Phi_t dt)^2 = f'(X_t)dX_t + \frac{1}{2}f''(X_t)\Delta_t^2 dt.$$

Thus we obtain $df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)\Delta_t^2 dt$. By integrating both sides, we obtain the integral form:

$$\int_0^t df(X_s) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)\Delta_s^2 ds$$

or

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)\Delta_s^2 ds$$

For two-variable functions, their second-order Taylor series expansions have the form:

$$df(x,t) = f'_x(x,t)dx + f'_t(x,t)dt + \frac{1}{2}f''_{x^2}(x,t)dx^2 + f''_{xt}(x,t)dxdt + \frac{1}{2}f''_{t^2}(x,t)dt^2.$$

Substituting $x = X_t, t = Y_t$ in the above formula, we obtain:

$$\begin{aligned} df(X_t, Y_t) &= f'_x(X_t, Y_t)dX_t + f'_t(X_t, Y_t)dY_t + \frac{1}{2}f''_{x^2}(X_t, Y_t)dX_t^2 + f''_{xt}(X_t, Y_t)dX_t dY_t + \frac{1}{2}f''_{t^2}(X_t, Y_t)dY_t^2 \\ &= f'_x(X_t, Y_t)dX_t + f'_t(X_t, Y_t)dY_t + \frac{1}{2}f''_{x^2}(X_t, Y_t)a_t^2 dt + f''_{xt}(X_t, Y_t)a_t c_t dt + \frac{1}{2}f''_{t^2}(X_t, Y_t)c_t^2 dt \end{aligned}$$

We now consider some special cases. Let $f(x,t) = xt$ then:

$$\begin{aligned} dX_tY_t &= Y_t dX_t + X_t dY_t + dX_t dY_t \\ &= Y_t dX_t + X_t dY_t + a_t c_t dt \end{aligned}$$

Substituting $x = W_t, y = t$ in the general formula and ignore all the terms associated with dt^m, with $m > 1$, we obtain:

$$df(W_t,t) = f'_x(W_t,t)dW_t + f'_t(W_t,t)dt + \frac{1}{2}f''_{x^2}(W_t,t)dW_t^2 = f'_x(W_t,t)dW_t + [f'_t(W_t,t) + \frac{1}{2}f''_{x^2}(W_t,t)]dt.$$

By integrating both sides, we obtain the integral form:

$$f(W_t,t) - f(W_0,0) = \int_0^t f'_x(W_s,s)dW_s + \int_0^t f'_t(W_s,s)ds + \int_0^t \frac{1}{2}f''_{x^2}(W_s,s)ds$$

9. Solving stochastic differential equation
A diffusion-type stochastic differential equation (SDE) can be described as

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

where W_t is a standard Brownian motion. We call $\mu(X_t,t)$ the drift and $\sigma(X_t,t)$ be the volatitility of the process. The integral form of the SDE is

$$X_t = X_0 + \int_0^t \mu(X_s,s)ds + \int_0^t \sigma(X_s,s)dW_s,$$

with $X_0 = x_0$ be the initial state of the solution process.

Our objective to solve a SDE:

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

is to find all stochastic processes satisfying the SDE. Such processes are called Ito processes.

Let $\{(W_t)_{t\geq 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Suppose that the process $(X_t)_{t\geq 0}$ is governed by the SDE: $dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t$, where μ and σ are functions of X_t and t. Show that X_t is a martingale if $\mu(X_t,t) = 0$.

(Bachelier Model: Arithmetic Brownian motion). Let $\{(W_t)_{t\geq 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Suppose that the process $(X_t)_{t\geq 0}$ is governed by the SDE:

$$dX_t = \mu dt + \sigma dW_t$$

where μ and σ are constant. Show that the random variable $(X_T|X_t = x)$ follows a normal distribution with mean $x + \mu(T-t)$ and variance $\sigma^2(T-t)$.

(Black-Scholes model: Geometric Brownian motion). Let $\{(W_t)_{t\geq 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Suppose that the process $(X_t)_{t\geq 0}$ is governed by the geometric Brownian motion:

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

where μ and σ are constant.

i) By applying Ito formula to $Y_t = \log X_t$, show that $X_T = X_t e^{(\mu - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T - W_t)}$.

ii) Show that the random variable $(X_T | X_t = x)$ follows a lognormal distribution with mean $xe^{\mu(T-t)}$ and variance $x^2(e^{\sigma^2(T-t)} - 1)e^{2\mu(T-t)}$.

(Ornstein-Ulenbeck Process). Let $\{(W_t)_{t\geq 0}\}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Suppose that the process $(X_t)_{t\geq 0}$ is governed by the Ornstein-Uhlenbeck process:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t$$

where κ, θ and σ are constant.

i) By applying Ito formula to $Y_t = e^{\kappa t} X_t$, show that $X_T = X_t e^{(-\kappa)(T-t)} + \theta[1 - e^{-\kappa(T-t)}] + \int_t^T \sigma e^{-\kappa(T-s)} dW_s$.

ii) Show that the random variable $(X_T | X_t = x)$ follows a normal distribution with mean

$$xe^{-\kappa(T-t)} + \theta[1 - e^{-\kappa(T-t)}]$$

and variance

$$\frac{\sigma^2}{2\kappa}[1 - e^{-2\kappa(T-t)}].$$

References