# Ch 1. Wiener Process (Brownian Motion)

**I. Introduction of Wiener Process**

**II. Itô's Lemma**

**III. Stochastic Integral**

**IV. Solve Stochastic Differential Equations with Stochastic Integral**

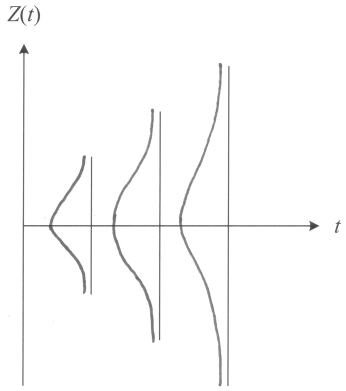
- This chapter introduces the stochastic process (especially the Wiener process), Itô's Lemma, and the stochastic intergral. The knowledge of the stochastic process is the foundation of derivative pricing and thus indispensable in the field of financial engineering.

- This course, however, is not a mathematic course. The only goal of this chapter is to build students up with enough knowledge about the stochastic process and thus to be able to understand academic papers associated with derivative pricing.

## I. Introduction of Wiener Process

- Wiener process, also called Brownian motion, is a kind of Markov stochastic process.
  - ⊙ Stochastic process: whose value changes over time in an uncertain way, and thus we only know the distribution of the possible values of the process at any time point. (In contrast to the stochastic process, a deterministic process is with an exact value at any time point.)
  - ⊙ Markov process: the likelihood of the state at any future time point depends only on its present state but not on any past states.
  - ⊙ In a word, the Markov stochastic process is a particular type of stochastic process where only the current value of a variable is relevant for predicting the future movement.
  - ⊙ The Wiener process $Z(t)$ is in essence a series of normally distributed random variables, and for later time points, the variances of these normally distributed random variables increase to reflect that it is more uncertain (thus more difficult) to predict the value of the process after a longer period of time. See Figure 1-1 for illustration.

**Figure 1-1**




- Instead of assuming $Z(t) \sim ND(0,t)$, which cannot support algebraic calculations, the Wiener process $dZ$ is introduced.

- $\triangle Z \equiv \varepsilon\sqrt{\triangle t}$ (change in a time interval $\Delta t$)

  $\varepsilon \sim ND(0{,}1) \Rightarrow \Delta Z$ follows a normal distribution

  $$\Rightarrow \begin{cases} E[\Delta Z] &= 0 \\ \mathrm{var}(\Delta Z) &= \Delta t \Rightarrow \mathrm{std}(\Delta Z) = \sqrt{\Delta t} \end{cases}$$

- $Z(T) - Z(0) = \sum\limits_{i=1}^{n} \varepsilon_i \sqrt{\triangle t} = \sum\limits_{i=1}^{n} \triangle Z_i$, where $n = \frac{T}{\triangle t}$

  $\Rightarrow Z(T) - Z(0)$ also follows a normal distribution

  $$\Rightarrow \begin{cases} E[Z(T) - Z(0)] &= 0 \\ \mathrm{var}(Z(T) - Z(0)) &\underset{\uparrow}{=} n \cdot \Delta t = T \Rightarrow \mathrm{std}(Z(T) - Z(0)) = \mathrm{std}(Z(T)) = \sqrt{T} \end{cases}$$

  Variances are additive because any pair of $\Delta Z_i$ and $\Delta Z_j$ $(i \neq j)$ are assumed to be independent. $Z(0) = 0$ if there is no further assumption.

- As $n \to \infty$, $\triangle t$ converges to 0 and is denoted as $dt$, which means an infinitesimal time interval. Correspondingly, $\Delta Z$ is redenoted as $dZ$.

- In conclusion, $dZ$ is noting more than a notation. It is invented to simplify the representation of the Wiener process, which is in essence a series of normal distributions with variances in proportional to time.

- Main properties of Wiener process $\{Z(t)\}$ for $t \geq 0$:

  (i) (Normal increments) $Z(t) - Z(s) \sim ND(0, t-s)$.

  (ii) (Independence of increments) $Z(t) - Z(s)$ and $Z(u)$ are independent, for $u \leq s < t$.

  (iii) (Continuity of the path) $Z(t)$ is a continuous function of $t$.

> Other properties:
>
> ⊙ Jagged path: not monotone in any interval, no matter how small a interval is.
>
> ⊙ None-differentiable everywhere: $Z(t)$ is continuous but with infinitely many edges.
>
> ⊙ Infinite variation on any interval: $V_Z([a,b]) = \infty$.
>
> > variation of a real-valued function $g$ on $[a,b]$:
> >
> > $$V_g([a,b]) = \sup_{\mathbf{P}} \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})|, \; a = t_1 < t_2 < \cdots < t_n = b,$$
> >
> > where $\mathbf{P}$ is the set of all possible partitions with mesh size going to zero as $n$ goes to infinity.
>
> ⊙ Quadratic variation on $[0,t]$ is $t$.
>
> $$[Z,Z](t) = [Z,Z]([0,t]) = \sup_{\mathbf{P}} \sum_{i=1}^{n} |Z(t_i) - Z(t_{i-1})|^2 = t \text{ (will be proved later)}$$
>
> ⊙ $\text{cov}(Z(t), Z(s)) = E[Z(t)Z(s)] - E[Z(t)]E[Z(s)] = E[Z(t)Z(s)]$.
>
> (If $s < t, Z(t) = Z(s) + Z(t) - Z(s)$.)
>
> $= E[Z^2(s)] + E[Z(s)(Z(t) - Z(s))] = E[Z^2(s)] = \text{var}(Z(s)) = s = \min(t,s)$.
>
> (The covariance is the length of the overlapping time period (or the sharing path) between $Z(t)$ and $Z(s)$.)

- Generalized Wiener process

  $dX = adt + bdZ$

$$\Rightarrow \begin{cases} E[dX] &= adt \\ \text{var}(dX) &= b^2 dt \Rightarrow \text{std}(dX) = b\sqrt{dt} \end{cases}$$

$\Rightarrow dX \sim ND(adt, b^2 dt)$

$\Rightarrow X(T) - X(0) = \sum_{i=1}^{n} \Delta X_i \sim ND(aT, b^2 T)$

- Itô process (also called diffusion process) (Kiyoshi Itô, a Japanese mathematician, deceased in 2008 at the age of 93.)

  $dX{=}a(X,t)dt{+}b(X,t)dZ$

  $\downarrow \qquad \downarrow$

  drift and volatility (Both are not constants, so it is no more simple to derive $E[dX]$ and $\text{var}(dX)$)

  (Both generalized Wiener processes and Itô process are called stochastic differential equation (SDE).)

- For the stock price, it is commonly assumed to follow an Itô process

  $dS = \mu S dt + \sigma S dZ$

  $\Rightarrow \frac{dS}{S} = \mu dt + \sigma dZ$ (also known as the geometric Brownian motion, GBM)

  $\Rightarrow \frac{dS}{S} \sim ND(\mu dt, \sigma^2 dt)$

  > $\frac{d\ln S}{dS} = \frac{1}{S} \Rightarrow d\ln S = \frac{dS}{S}$ (WRONG!) (Note that this differential result is true only when $S$ is a real-number variable. This type of differentiation CANNOT be applied to stochastic processes. The stochastic calculus is not exactly the same as the calculus for real-number variables, since any change of a stochastic process must come through the passage of a time period.)
  >
  > In fact, stock prices follow the lognormal distribution based on the above assumption of the geometric Brownian motion, but it does not mean $d\ln S \sim ND(\mu dt, \sigma^2 dt)$.

- (Advanced content) Stochastic volatility (SV) process for the stock price (Heston (1993)):

$$dS = \mu S dt + \sqrt{V} S dZ_S,$$
$$dV = \kappa(\theta - V)dt + \sigma_V \sqrt{V} dZ_V,$$

  where $V$ denotes the stochastic variance and $\text{corr}(dZ_S, dZ_V) = \rho_{SV}$.

- (Advanced content) Jump-diffusion process for the stock price (Merton (1976)):

$$dS = (\mu - \lambda E[Y_S - 1])S dt + \sigma S dZ + (Y_S - 1) S dq,$$

where $dq$ is a Poisson (counting) process with the jump intensity $\lambda$, i.e., the probability of an event occuring during a time interval of length $\Delta t$ is

$$\begin{cases} \text{Prob}\ \{\text{the event does not occur in } (t, t+\Delta t], \text{ i.e., } dq = 0\} & = & 1 - \lambda\Delta t - \lambda^2 (\Delta t)^2 - ... \\ \text{Prob}\ \{\text{the event occurs once in } (t, t+\Delta t], \text{ i.e., } dq = 1\} & = & \lambda\Delta t \\ \text{Prob}\ \{\text{the events occur twice in } (t, t+\Delta t], \text{ i.e., } dq = 2\} & = & \lambda^2 (\Delta t)^2 \to 0 \\ \vdots & & \end{cases},$$

and the random variable $(Y_S - 1)$ is the random percentage change in the stock price if the Poisson events occur. Merton (1976) considers $\ln Y_S \sim ND(\mu_J, \sigma_J^2)$. Note that $dZ$, $Y_S$, and $dq$ are assumed to be mutually independent. The introduction of the term $(\lambda E[Y_S - 1])$ in the drift is to maintain the growth rate of $S$ to be $\mu$. This is because $E[(Y_S - 1)dq] = E[Y_S - 1]E[dq] = E[Y_S - 1]\lambda dt$.

> If $Y_S$ follows the lognormal distribution, $E[Y_S - 1] = E[Y_S] - 1 = e^{E[\ln Y_S] + \frac{1}{2}\text{var}(\ln Y_S)}$ $-1 = e^{\mu_J + \frac{1}{2}\sigma_J^2} - 1$.

## II. Itô's Lemma

- Itô's Lemma is essentially based on the Taylor series:

$$\begin{aligned} f(x,y) = {} & f(x_0,y_0) + \tfrac{\partial f}{\partial x}(x - x_0) + \tfrac{\partial f}{\partial y}(y - y_0) \\ & + \tfrac{1}{2!}\left[\tfrac{\partial^2 f}{\partial x^2}(x - x_0)^2 + 2\tfrac{\partial^2 f}{\partial x \partial y}(x - x_0)(y - y_0) + \tfrac{\partial^2 f}{\partial y^2}(y - y_0)^2\right] + \cdots \end{aligned}$$

Using the Itô's Lemma to derive a stochastic differential equation:

Given $dX = a(X,t)dt + b(X,t)dZ$, and $f(X,t)$ as a function of $X$ and $t$, the stochastic differential equation for $f$ can be derived as follows.

$$df = \left(\tfrac{\partial f}{\partial t} + \tfrac{\partial f}{\partial X}a + \tfrac{1}{2}\tfrac{\partial^2 f}{\partial X^2}b^2\right)dt + \left(\tfrac{\partial f}{\partial X}b\right)dZ,$$

where $a$ and $b$ are the abbreviations of $a(X,t)$ and $b(X,t)$.

The Itô's Lemma holds under the following approximations:
(i)
$(dt)^1 \quad \to dt$
$(dt)^{1.5} \to 0$
$(dt)^2 \quad \to 0$
$\vdots$

(ii)
$dZ \cdot dZ = ?$
By definition, $dZ \cdot dZ = \varepsilon^2 \cdot dt$.
$\because \varepsilon \sim ND(0,1)$
$\therefore \text{var}(\varepsilon) = 1 \Rightarrow E[\varepsilon^2] - (E[\varepsilon])^2 = 1 \Rightarrow E[\varepsilon^2] = 1 \Rightarrow E[(dZ)^2] = dt$
In addition, $\text{var}((dZ)^2) = \text{var}(\varepsilon^2 dt) = (dt)^2 \text{var}(\varepsilon^2) \to 0$ (because $(dt)^2 \to 0$)
$\Rightarrow dZ \cdot dZ \overset{a.s.}{=} dt$ ("a.s." means "almost surely": an event happens almost surely if it happens with probability 1.)

Itô's Lemma vs. differentiation of a deterministic function of time.

* For a deterministic function of time $f(t)$, if $\frac{df}{dt} = g(t)$, we can interpret that with an infinitesimal change of $dt$, the change in $f$ is $g(t)dt$, which is deterministic.
* The interpretation of the Itô's Lemma: with a infinitesimal change of $dt$, the change in $f$ is $(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial X}a + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}b^2)dt + (\frac{\partial f}{\partial X}b)dZ$. Note that the first term plays a similar role as $g(t)dt$, but the second term tells us that the change in $f$ is random.
* To apply the Itô's Lemma is similar to taking the differentiation for stochastic processes.

- Based on the result of $dZ \cdot dZ = (dZ)^2 = dt$, it is straightforward to infer that the quadratic variation of the Wiener process over $[0,t]$, i.e., $[Z,Z](t) = [Z,Z]([0,t]) = \sup\limits_{\mathbf{P}} \sum\limits_{i=1}^{n} |Z(t_i) - Z(t_{i-1})|^2$, equals $t$.

Similar to the derivation of the Itô's Lemma that $E[(dZ)^2] = dt$ and $\text{var}((dZ)^2) \to 0$ when $n \to \infty$ $(dt \to 0)$, $(Z(t_i) - Z(t_{i-1}))^2$ converges to $t_i - t_{i-1}$ almost surely if $(t_i - t_{i-1})$ is very small. This is because
$E[(Z(t_i) - Z(t_{i-1}))^2] = E[\varepsilon^2(t_i - t_{t-1})] = t_i - t_{t-1}$, and
$\text{var}((Z(t_i) - Z(t_{i-1}))^2) = \text{var}(\varepsilon^2(t_i - t_{t-1})) = (t_i - t_{t-1})^2 \text{var}(\varepsilon^2) \to 0$.
Thus, we can conclude that when $n \to \infty$ $(t_i - t_{t-1} \to 0)$, $\sup\limits_{\mathbf{P}} \sum\limits_{i=1}^{n} (Z(t_i) - Z(t_{i-1}))^2 = t$.

- Example 1: Apply the Itô's Lemma to $f = \ln S$, given $dS = \mu S dt + \sigma S dZ$.

$$\Rightarrow d\ln S = (0 + \tfrac{1}{S} \cdot \mu S - \tfrac{1}{2}\tfrac{1}{S^2} \cdot \sigma^2 S^2)dt + \tfrac{1}{S}\sigma S dZ$$
$$= (\mu - \tfrac{\sigma^2}{2})dt + \sigma dZ.$$

⊙ $\Delta \ln S = (\mu - \frac{\sigma^2}{2})\Delta t + \sigma \Delta Z$

$\Rightarrow \ln S_{t+\Delta t} - \ln S_t = (\mu - \frac{\sigma^2}{2})\Delta t + \sigma\Delta Z \sim ND((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2 \Delta t)$

$\Rightarrow \ln S_{t+\Delta t} \sim ND(\ln S_t + (\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2\Delta t)$.

⊙ Consider $\frac{T-t}{n} = \Delta t$,

$$\begin{cases} \ln S_{t+\Delta t} - \ln S_t & \sim \quad ND((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2\Delta t) \\ \ln S_{t+2\Delta t} - \ln S_{t+\Delta t} & \sim \quad ND((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2\Delta t) \\ & \vdots \\ \ln S_T - \ln S_{T-\Delta t} & \sim \quad ND((\mu - \frac{\sigma^2}{2})\Delta t, \sigma^2\Delta t) \end{cases}$$

$\Rightarrow \ln S_T - \ln S_t \sim ND((\mu - \frac{\sigma^2}{2})n\Delta t, \sigma^2 n\Delta t)$

$\Rightarrow \ln S_T - \ln S_t \sim ND((\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t))$

$\Rightarrow \ln S_T \sim ND(\ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t))$

$\Rightarrow$ The stock price at any future time point is lognormally distributed.

> Another derivation: apply the stochastic integral on the both sides of the equation
> $\Rightarrow \int_t^T d\ln S_\tau = \underbrace{\int_t^T (\mu - \frac{\sigma^2}{2})d\tau}_{\downarrow} + \int_t^T \sigma dZ(\tau)$
>
> Since the integrand is a constant and the variable $\tau$ is a real-number variable, it is simply an integral for a real-number variable.
>
> $\Rightarrow \ln S_\tau|_t^T = (\mu - \frac{\sigma^2}{2})(T-t) + \sigma(\underline{Z(\tau)|_t^T})$
>
> $\downarrow$
>
> $Z(T) - Z(t) \equiv \triangle Z(T-t) \sim ND(0, T-t)$
>
> $\Rightarrow \ln S_T - \ln S_t \sim ND((\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t))$

- Example 2: Given $dS = (\mu - q)Sdt + \sigma S dZ$, for $f = Se^{-q(T-t)} - Ke^{-r(T-t)}$, where $f$ is the value of a forward or futures, $q$ is the dividend yield, $K$ is the delivery price, and $T$ is the delivery date, one can derive that $df = (\mu S e^{-q(T-t)} - rKe^{-r(T-t)})dt + \sigma S e^{-q(T-t)} dZ$.

- Example 3: Given $dS = (\mu - q)Sdt + \sigma S dZ$, for $F = Se^{(r-q)(T-t)}$, where $F$ is the forward or futures price of a stock and $q$ and $T$ denote the divided yield and delivery date, respectively, it can be obtained via the Itô's Lemma that $dF = (\mu - r)Fdt + \sigma F dZ$.

- Itô's Lemma for multiple variates

  $\frac{dS}{S} = \mu_S dt + \sigma_S dZ_S$ (foreign stock price),

  $\frac{dX}{X} = \mu_X dt + \sigma_X dZ_X$ (exchange rate: 1 foreign dollar $=$ $X$ domestic dollars).

  Define $f = S \cdot X$ (the value of a foreign stock share in units of domestic dollars)

  $$df = [\frac{\partial f}{\partial t} + \frac{\partial f}{\partial S} \cdot \mu_S S + \frac{\partial f}{\partial X} \cdot \mu_X X + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial S^2} \cdot \sigma_S^2 S^2 + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial X^2} \cdot \sigma_X^2 X^2 +$$
  $$\frac{\partial^2 f}{\partial S \partial X} \cdot \rho_{XS} \cdot \sigma_S \cdot \sigma_X \cdot S \cdot X]dt + \frac{\partial f}{\partial S}\sigma_S S dZ_S + \frac{\partial f}{\partial X}\sigma_X X dZ_X$$

  $\Rightarrow df = [\mu_S XS + \mu_X XS + \rho_{XS}\sigma_S\sigma_X SX]dt + \sigma_S XS dZ_S + \sigma_X XS dZ_X$

  $\Rightarrow \frac{df}{f} = (\mu_S + \mu_X + \rho_{XS}\sigma_S\sigma_X)dt + \sigma_S dZ_S + \sigma_X dZ_X$ (because $f = SX$).

  $dZ_S \cdot dZ_X = \varepsilon_S\sqrt{dt} \cdot \varepsilon_X\sqrt{dt} = \varepsilon_S\varepsilon_X dt$
  $\Rightarrow E[dZ_S \cdot dZ_X] = E[\varepsilon_S\varepsilon_X]dt = \rho_{XS}dt$
  $\quad \mathrm{var}(dZ_S \cdot dZ_X) = (dt)^2 \mathrm{var}(\varepsilon_S\varepsilon_X) \to 0$
  $\Rightarrow dZ_S \cdot dZ_X \overset{a.s.}{=} \rho_{XS}dt.$

- (Advanced content) Given $dS = (\mu - \lambda K_Y)Sdt + \sigma SdZ + (Y_S - 1)Sdq$, where $K_Y = E[Y_S - 1]$ and $f(S,t)$ as a function of $S$ and $t$, the Itô's Lemma implies

  $$df = \{\frac{\partial f}{\partial t} + \frac{\partial f}{\partial S}(\mu - \lambda K_Y)S + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 + \lambda E[f(SY_S,t) - f(S,t)]\}dt$$
  $$+ \frac{\partial f}{\partial S}\sigma S dZ + (Y_f - 1)f dq,$$

  where $\lambda E[f(SY_S,t) - f(S,t)]dt$ is the expected jump effect on $f$, and $(Y_f - 1)dq$ is introduced to capture the unexpected (zero-mean) jump effect on $f$, where $(Y_f - 1)$ is the random percentage change in $f$ if the Poisson event occurs. Note that $\lambda E[f(SY_S,t) - f(S,t)]dt + (Y_f - 1)f dq$ represents the total effect on $f$ if the Poisson event occurs.

  ⊙ Suppose $f = \ln S$, the Itô's Lemma implies

  $$d\ln S = (\mu - \lambda K_Y - \tfrac{1}{2}\sigma^2)dt + \sigma dZ + J_{\ln S},$$

  where $J_{\ln S}$ represents the total effect on $\ln S$ due to the random jump in $S$.

  If the jump occurs in $S$ at $t$, we can obtain
  $$\frac{S(t^+) - S(t)}{S(t)} = (Y_S - 1),$$
  since $(Y_S - 1)$ is the precentage change if the jump occurs (when $dq = 1$).
  Rewriting the above equation leads to
  $S(t^+) - S(t) = (Y_S - 1)S(t) = Y_S S(t) - S(t) \Rightarrow S(t^+) = Y_S S(t).$

  The random jump in $\ln S$ at $t$, if the Poisson event occurs (when $dq = 1$), is
  $\ln S(t^+) - \ln S(t) = \ln Y_S + \ln S(t) - \ln S(t) = \ln Y_S.$

According to the above inference, we can express the total jump effect by

$$J_{\ln S} = \ln Y_S dq,$$

and thus

$$d \ln S = (\mu - \tfrac{1}{2}\sigma^2 - \lambda K_Y)dt + \sigma dZ + \ln Y_S dq.$$

$*$ Note that the expected jump effect in $\ln S$ is $E[J_{\ln S}] = E[\ln Y_S dq] = \lambda E[\ln Y_S]dt$, and the unexpected jump effect $(Y_f - 1)f dq$ is set to be equal to $\ln Y_S dq - \lambda E[\ln Y_S]dt$.

### III. Stochastic Integral

- Stochastic integral (or called Itô integral or Itô calculus): allows one to integrate one stochastic process (the integrand) over another stochastic process (the integrator). Usually, the integrator is a Wiener process.

- Integral over a stochastic process: $\int_a^b X(\tau)dZ(\tau)$, where $X(\tau)$ can be a deterministic function or a stochastic process, and $dZ(\tau)$ is a Wiener process. (vs. integral over a real-number variable: $\int_a^b f(y)dy$, where $f(y)$ is a deterministic function of the real-number variable $y$)

- Three cases of $X(\tau)$ are discussed: simple deterministic processes, simple predictable processes, and general predictable processes (or say Itô processes).

- Stochastic integral for "simple deterministic" processes

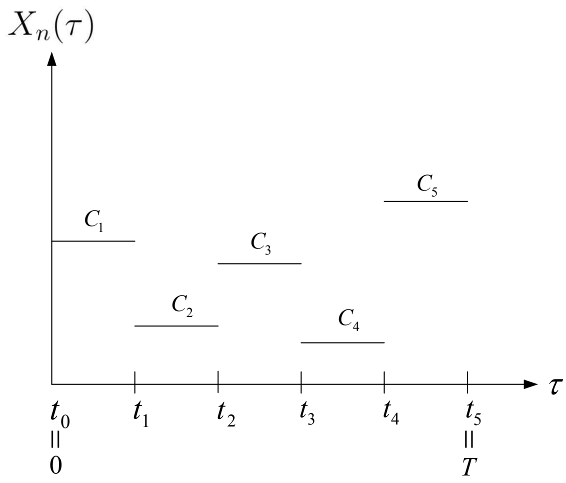
  If $X(\tau)$ is a deterministic process, given any value of $t$, the value of $X(\tau)$ can be known exactly. Therefore, in an infinitesimal time interval, $(t_{i-1}, t_i]$, the value of $X(\tau)$ can be approximated by a constant $C_i$. The term "simple" means to approximate the process by a step function. Denote the step function as $X_n(\tau)$, where $n$ is the number of partitions in $[0, T]$. (In contrast, if $X(\tau)$ is a stochastic process, given any value of $\tau$, we only know the distribution of possible values for $X(\tau)$.)

- For simple deterministic processes, we can define the stochastic integral as follows. (This definition is similar to the rectangle method to define the integral over a real-number variable.)

$$\int_0^T X_n(\tau)dZ(\tau) \equiv \sum_{i=1}^{n} C_i(Z(t_i) - Z(t_{i-1})) \sim ND(0, \sum_{i=1}^{n} C_i^2(t_i - t_{i-1})),$$

  where $X_n(\tau)$ is the discrete-time, step-function approximation of $X(\tau)$ given that $[0, T]$ is partitioned into $n$ (equal-size) intervals.

**Figure 1-2**




$*$ Finally, $\int_0^T X(\tau)dZ(\tau)$ is defined to be $\lim_{n\to\infty} \int_0^T X_n(\tau)dZ(\tau)$.

$*$ Note also the variance $\lim_{n\to\infty} \sum_{i=1}^{n} C_i^2(t_i - t_{i-1})$ converges to $\int_0^T X^2(\tau)d\tau$ by definition.

$*$ Consequently, the stochastic integral for a deterministic process yields a normal distribution with a zero mean and a variance equal to $\int_0^T X^2(\tau)d\tau$. (consistent with Properties (iii) and (iv) of Itô integral on p. 1-13)

(i) According to the above definition, if $X(t) = 1$, the result of the stochastic integral is consistent with the definition of the Wiener process.

$$\int_0^T X(\tau)dZ(\tau) = \int_0^T dZ(\tau) = Z(\tau)|_0^T = Z(T) - Z(0) \sim ND(0,T)$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} (Z(t_i) - Z(t_{i-1})) \sim ND(0, \lim_{n\to\infty} \sum_{i=1}^{n} (t_i - t_{i-1})) = ND(0,T)$$

(ii) Alternative way to calculate the resulting variance of the stochastic integral:

$$\text{var}(\int XdZ) = E[(\int XdZ)^2] - (E[\int XdZ])^2 = E[(\int XdZ)^2] = E[(\lim_{n\to\infty} \sum_{i=1}^{n} C_i(Z(t_i) - Z(t_{i-1})))^2]$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} C_i C_j E[(Z(t_i) - Z(t_{i-1}))(Z(t_j) - Z(t_{j-1}))]$$

$\uparrow$

calculate the squared term in the expectation, and then apply the distributive property of the expectation over the addition and scaler multiplication

$$= \lim_{n\to\infty} \sum_{i=1}^{n} C_i^2(t_i - t_{i-1}).$$

$\uparrow$

because $\text{cov}(Z(t_i) - Z(t_{i-1}), Z(t_j) - Z(t_{j-1})) = 0$, and $\text{var}(Z(t_i) - Z(t_{i-1})) = t_i - t_{i-1}$

- "Simple predictable" process: in the time interval $(t_{i-1}, t_i]$, the constant $C_i$ is replaced by a "random variable" $\xi_i$, which depends on the values of $Z(t)$ for $t \leq t_{i-1}$, but not on values of $Z(t)$ for $t > t_{i-1}$. Therefore, $X_n(t)$ is defined as follows.

$$X_n(t) = \xi I_{\{t|t=0\}} + \sum_{i=1}^{n} \xi_i I_{\{t|t_{i-1}<t\leq t_i\}},$$

where $I$ is a indicator function and $\xi$ is a constant. The corresponding stochastic integral is defined as follows.

$$\int_0^T X_n(\tau)dZ(\tau) \equiv \sum_{i=1}^{n} \xi_i (Z(t_i) - Z(t_{i-1})).$$

- The reason for the name "predictable":

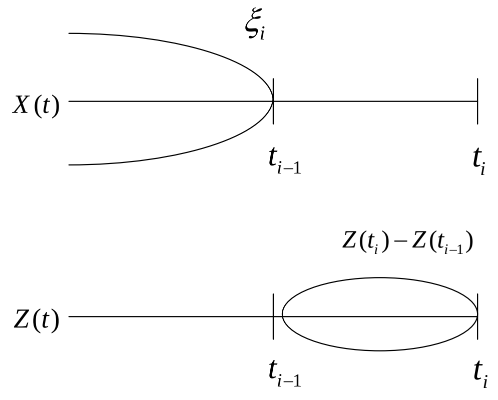
1. The value of $X(t)$ for $(t_{i-1}, t_i]$, $\xi_i$, is determined based on the information set formed by $\{Z(t)\}$ until $t_{i-1}$, denoted by $\mathcal{F}_{t_{i-1}}$. It is also called that $\xi_i$ is $\mathcal{F}_{t_{i-1}}$-measurable. (See Figure 1-3.)

2. In contrast, the value of $Z(t_i) - Z(t_{i-1})$ will not realize until the time point $t_i$, i.e., this value will be known based on the information set $\mathcal{F}_{t_i}$. In other words, $Z(t_i)$ is $\mathcal{F}_{t_i}$-measurable. (See Figure 1-3.)

3. Therefore, we say that $X(t)$ is "predictable" since we know its realized value just before the time point at which $Z(t)$ is realized.

4. In the continuous-time model, $Z(t)$ is $\mathcal{F}_t$-measurable (the realized value is known at $t$). For any process that we can know its realized value just before $t$, we call this process to be $\mathcal{F}_{t-}$-measurable and thus "predictable".

**Figure 1-3**



- Stochastic integral of “general predictable” processes

  Let $X_n(t)$ be a sequence of simple predictable processes (which can be approximated by a step function with a series of predictable random variables) convergent in probability to the process $X(t)$, which is “general predictable” (i.e., $X(t)$ is predictable and $\int_0^T X^2(\tau)d\tau < \infty$). The sequence of their integrals $\int_0^T X_n(\tau)dZ(\tau)$ also converges to $\int_0^T X(\tau)dZ(\tau)$ in probability, i.e.,

$$\lim_{n\to\infty} \int_0^T X_n(\tau)dZ(\tau) = \int_0^T X(\tau)dZ(\tau).$$

  (Convergence in probability: the probability of an unusual outcome becomes smaller and smaller as the sequence progresses.)

  (In some text books, the general predictable process is also known as the predictable process for short.)

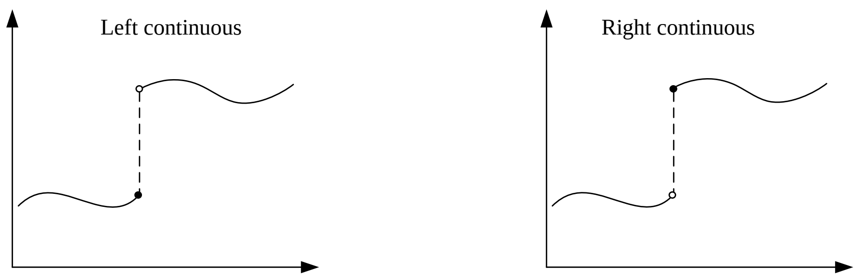
- Any “adapted” and “left continuous” process is a “predictable” process.

  A process is an adapted process iff it is $\mathcal{F}_t$ measurable. For example, the Wiener process $Z(t)$ is an adapted process.

  A left-continuous function is a function which is continuous at all points when approached from the left. In addition, a function is continuous if and only if it is both right-continuous and left-continuous. Since $Z(t)$ is a continuous function of $t$, it must be left-continuous.

  Thus, we can conclude that Wiener process $Z(t)$ is a predictable process, so $Z(t)$ itself (or even all Itô processes) can be the integrand in a stochastic intrgral. This is also the reason for the name of the Itô integral.

**Figure 1-4**



- Solve $\int_0^T Z(\tau)dZ(\tau)$, given $Z(0) = 0$.

  Define $X_n(t) = \sum_{i=1}^{n} Z(t_{i-1})I_{\{t|t_{i-1}<t\leq t_i\}}$ ($\lim_{n\to\infty} X_n(t)$ converging to $Z(t)$ in probability)

  $$\begin{aligned}\int_0^T X_n(\tau)dZ(\tau) &= \sum_{i=1}^{n} Z(t_{i-1})(Z(t_i) - Z(t_{i-1})) \\ &= \tfrac{1}{2}\sum_{i=1}^{n}[(Z(t_i))^2 - (Z(t_{i-1}))^2 - (Z(t_i) - Z(t_{i-1}))^2] \\ &= \tfrac{1}{2}(Z(T))^2 - \tfrac{1}{2}(Z(0))^2 - \tfrac{1}{2}\sum_{i=1}^{n}(Z(t_i) - Z(t_{i-1}))^2\end{aligned}$$

  $\Rightarrow \int_0^T Z(\tau)dZ(\tau) = \lim_{n\to\infty} \int_0^T X_n(\tau)dZ(\tau) = \frac{1}{2}(Z(T))^2 - \frac{1}{2}T.$

  > When $n \to \infty$, $(t_i - t_{i-1}) \to 0$. Therefore, $E[(Z(t_i) - Z(t_{i-1}))^2] = t_i - t_{i-1}$ and $\mathrm{var}((Z(t_i) - Z(t_{i-1}))^2) = \mathrm{var}(\varepsilon^2(t_i - t_{i-1})) = (t_i - t_{i-1})^2\mathrm{var}(\varepsilon^2) \to 0$. Note that the same argument is employed to derive the quadratic variation of the Wiener process on p. 1-6.

- Properties of Itô Integral:

  (i) $\int_0^T(\alpha X(\tau) + \beta Y(\tau))dZ(\tau) = \alpha\int_0^T X(\tau)dZ(\tau) + \beta\int_0^T Y(\tau)dZ(\tau)$ (distributive property).

  (ii) $\int_0^T I_{\{\tau|a<\tau\leq b\}}(\tau)dZ(\tau) = Z(b) - Z(a)$, $0 < a < b < T$.

  (iii) $E[\int_0^T X(\tau)dZ(\tau)] = 0$.

  (iv) $\mathrm{var}(\int_0^T X(\tau)dZ(\tau)) = E[(\int_0^T X(\tau)dZ(\tau))^2] = \int_0^T E[X^2(\tau)]d\tau$ (Itô Isometry).

  $*$ Note that the above properties are applicable for any Itô processes of $X(t)$ and $Y(t)$. For other stochastic processes, not all of the above properties hold.

  $*$ Intuition for Property (iii): By considering $X_n(t) = \sum_{i=1}^{n} X(t_{i-1})I_{\{t|t_{i-1}<t\leq t_i\}}$, one can obtain $E[\int_0^T X(\tau)dZ(\tau)] = E[\lim_{n\to\infty}\int_0^T X_n(\tau)dZ(\tau)] = E[\lim_{n\to\infty}\sum_{i=1}^{n} X(t_{i-1})(Z(t_i)-Z(t_{i-1}))]$. Since $X(t_{i-1})$ can be correlated with $Z(t)$ for $t \leq t_{i-1}$ but is independent of $Z(t_i) - Z(t_{i-1})$, $E[X(t_{i-1})(Z(t_i)-Z(t_{i-1}))] = \mathrm{cov}(X(t_{i-1}), Z(t_i)-Z(t_{i-1}))+E[X(t_{i-1})]E[Z(t_i)-Z(t_{i-1})] = 0$.

- What are $E[\int_0^T Z(\tau)dZ(\tau)]$ and $\mathrm{var}(\int_0^T Z(\tau)dZ(\tau))$?

  (i) $\because E[(Z(T))^2] = \mathrm{var}(Z(T)) + E[Z(T)]^2 = T$

  $\therefore E[\int_0^T Z(\tau)dZ(\tau)] = E[\frac{1}{2}(Z(T))^2 - \frac{1}{2}T] = 0.$

  (Property (iii) can be applied to obtaining the identical result directly.)

(ii) $\text{var}(\int_0^T Z(\tau)dZ(\tau)) = \frac{1}{4}\text{var}((Z(T))^2)$

$$= \frac{1}{4}\{E[(Z(T))^4] - E[(Z(T))^2]^2\} = \frac{1}{4}\{3T^2 - T^2\} = \frac{T^2}{2}.$$

> If $x \sim ND(\mu, \sigma^2)$, then $E[x^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$.
> Since $Z(T) \sim ND(0, T)$, we can derive $E[(Z(T))^4] = 3T^2$.

Apply Property (iv) to finding $\text{var}(\int_0^T Z(\tau)dZ(\tau))$ as follows:
$\text{var}(\int_0^T Z(\tau)dZ(\tau)) = \int_0^T E[(Z(\tau))^2]d\tau = \int_0^T \tau d\tau = \frac{1}{2}\tau^2|_0^T = \frac{T^2}{2}.$

## IV. Solve Stochastic Differential Equations with Stochastic Integral

- How to solve $X(t)$ systematically with the stochastic integral is one of the major applications of the stochastic integral.

- Given the Ornstein-Uhlenbeck process $dX(t) = -\alpha X(t)dt + \sigma dZ(t)$ and $X(0)$, solve $X(t)$.

> An important property of the Ornstein-Uhlenbeck process is that its drift-term setting renders the dynamicof $X(t)$ to be mean-reverting around 0.

According to the stochastic integral, $X(t)$ should satisfy

$$X(t) = X(0) + \int_0^t \mu(X,\tau)d\tau + \int_0^t \sigma(X,\tau)dZ(\tau).$$

> Since $\mu(X,t) = -\alpha X(t)$ is a function of $X(t)$, $\mu(X,t)$ is a stochastic process as well. Moreover, since the value of $\mu(X,t)$ is unknown due to the unsolved $X(t)$. Thus, we cannot derive $X(t)$ by applying the stochastic integral directly.

Define $Y(t) = X(t)e^{\alpha t} \Rightarrow dY(t) = e^{\alpha t}dX(t) + \alpha e^{\alpha t}X(t)dt$

(through the stochastic product rule introduced later)

$$= e^{\alpha t}[-\alpha X(t)dt + \sigma dZ(t)] + \alpha e^{\alpha t}X(t)dt$$
$$= \sigma e^{\alpha t}dZ(t)$$

> By applying the Itô's Lemma to $Y(t) = X(t)e^{\alpha t}$, since $\frac{\partial Y}{\partial t} = X\alpha e^{\alpha t}$, $\frac{\partial Y}{\partial X} = e^{\alpha t}$, and $\frac{\partial^2 Y}{\partial X^2} = 0$, $dY(t) = (\frac{\partial Y}{\partial t} + \frac{\partial Y}{\partial X}(-\alpha X) + \frac{1}{2}\frac{\partial^2 Y}{\partial X^2}\sigma^2)dt + \frac{\partial Y}{\partial X}\sigma dZ(t) = (X\alpha e^{\alpha t} + e^{\alpha t}(-\alpha X) + 0)dt$ $+e^{\alpha t}\sigma dZ(t) = \sigma e^{\alpha t}dZ(t)$.

$$\Rightarrow Y(t) = Y(0) + \int_0^t \sigma e^{\alpha\tau} dZ(\tau)$$

$$\downarrow$$

a simple deterministic process

$$\Rightarrow X(t) = e^{-\alpha t}(Y(0) + \int_0^t \sigma e^{\alpha\tau} dZ(\tau)), \text{ where } Y(0) = X(0).$$

* The idea behind the transformation $Y(t) = X(t)e^{\alpha t}$ is to find a function $Y(X,t)$ such that the drift term of $dY(t)$ is zero, and $X(t)$ or $Y(t)$ does not appear in the volatility term of $dY(t)$. Specifically, derive $Y(X,t)$ by simultaneously solving $\frac{\partial Y}{\partial t} + \frac{\partial Y}{\partial X}(-\alpha X) + \frac{1}{2}\frac{\partial^2 Y}{\partial X^2}\sigma^2 = 0$ and ensuring neither $X(t)$ nor $Y(t)$ appearing in $\frac{\partial Y}{\partial X}\sigma$.

* A more general (but still problem dependent) version of the above method is to introduce a transformation $Y(X,t)$ such that the drift and volatility terms of $dY(t)$ can be 0, constant, deterministic functions of $t$, or functions of a known stochastic process $U(t)$ (but can be neither functions of $X(t)$ or $Y(t)$).

* Later a systematical method is introduced to apply the stochastic integral to solving *linear* stochastic differential equations, which are commonly assumed (higher than 80% of probability) in the field of financial engineering.

- Solution of a linear stochastic differential equation (where $\alpha(t)$, $\beta(t)$, $\gamma(t)$, and $\delta(t)$ are deterministic functions of $t$):

  Given $dX(t)=(\alpha(t) + \beta(t)X(t))dt + (\gamma(t) + \delta(t)X(t))dZ(t)$, where the drift and volatility terms are linear functions of $X(t)$ and $\alpha(t)$, $\beta(t)$, $\gamma(t)$, and $\delta(t)$ are deterministic functions of $t$. Solve $X(t)$ provided that $X(0)$ is known.

  (i) We start to solve the SDE given $\alpha(t) = \gamma(t) = 0$.

  $$dU(t) = \beta(t)U(t)dt + \delta(t)U(t)dZ(t)$$

  $$\Rightarrow \frac{dU(t)}{U(t)} = \beta(t)dt + \delta(t)dZ(t)$$

  (Note that $U(t)$ is similar to $S(t)$, so we can apply the result on p. 1-6 to solve $U(t)$.)

  $$\Rightarrow U(t) = \underbrace{U(0)\exp(\int_0^t (\beta(\tau) - \frac{1}{2}\delta^2(\tau))d\tau + \int_0^t \delta(\tau)dZ(\tau))}_{(1)}$$

  (Note that $U(t)$ follows a lognormal distribution except in the case of $\delta(t) = 0$.)

  (ii) Consider $X(t) = U(t) \cdot V(t)$, and $U(0) = 1$ and $V(0) = X(0)$,

  where $dU(t) = \beta(t)U(t)dt + \delta(t)U(t)dZ(t)$,

  $dV(t) = a(t)dt + b(t)dZ(t)$. (Note $a(t)$ and $b(t)$ could be stochastic processes.)

$\odot$ Stochastic product rule:

$$dX(t) = d(U(t)V(t)) = dU(t)V(t) + U(t)dV(t) + d[U,V](t),$$

where $d[U,V](t) = dU(t)dV(t) = \sigma_U \sigma_V dt$, where $\sigma_U$ and $\sigma_V$ represent the volatility terms of $dU(t)$ and $dV(t)$, respectively.

(It can be proved by performing the Itô's Lemma and the derivation is similar to that of illustrating the multivariate Itô's Lemma on p. 1-8.)

(Note the stochastic product rule is exclusively applied to $X(t)$, which is the product of two (stochastic) processes sharing a common $dZ$.)

$\Rightarrow dX(t) = [\beta(t)U(t)dt+\delta(t)U(t)dZ(t)]V(t)+U(t)[a(t)dt+b(t)dZ(t)]+\delta(t)U(t)b(t)dt.$

$\odot$ Extension to the integration by parts for stochastic processes:

$$U(t)V(t) - U(0)V(0) = \int_0^t V(\tau)dU(\tau) + \int_0^t U(\tau)dV(\tau) + [U,V](t),$$

where $[U,V](t) = \lim_{n\to\infty} \sum_{i=1}^{n} (U(t_i) - U(t_{i-1}))(V(t_i) - V(t_{i-1}))$ (quadratic covariation).

$\odot$ Solve $a(t)$ and $b(t)$ by comparing with the originally $dX(t)$ process:

$\Rightarrow b(t) \cdot U(t) = \gamma(t),\ a(t) \cdot U(t) = \alpha(t) - \delta(t) \cdot \gamma(t)$

$\Rightarrow b(t) = \frac{\gamma(t)}{U(t)}, \quad a(t) = \frac{\alpha(t)-\delta(t)\gamma(t)}{U(t)}$

$$\Rightarrow V(t) = \underbrace{V(0) + \int_0^t \frac{\alpha(\tau) - \delta(\tau)\gamma(\tau)}{U(\tau)} d\tau + \int_0^t \frac{\gamma(\tau)}{U(\tau)} dZ(\tau)}_{(2) \quad \text{where } V(0) = X(0)},$$

$\Rightarrow X(t) = U(t) \cdot V(t) = (1) \times (2)$

- For nonlinear stochastic differential equation, one needs to employ the method introduced on p. 1-15 to solve it.

- Brownian bridge (pinned Brownian motion):

  $dX(t) = \frac{b-X(t)}{T-t}dt + dZ(t),\ 0 \le t \le T,\ X(0) = a$

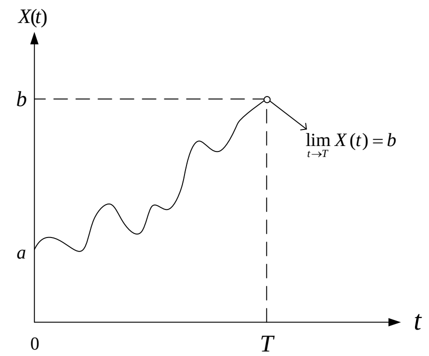
  $\Rightarrow \alpha(t) = \frac{b}{T-t},\ \beta(t) = \frac{-1}{T-t},\ \gamma(t) = 1,\ \delta(t) = 0$

$$U(t) = U(0)\exp\left(\int_0^t (\beta(\tau) - \tfrac{1}{2}\delta^2(\tau))d\tau + \int_0^t \delta(\tau)dZ(\tau)\right)$$
$$= \exp\left(\int_0^t \tfrac{-1}{T-\tau}d\tau\right) = \exp(\ln(T-\tau)|_0^t)$$
$$= \exp\left(\ln \tfrac{T-t}{T}\right) = \tfrac{T-t}{T}$$
$$b(t) = \tfrac{T}{T-t}$$
$$a(t) = \frac{\frac{b}{T-t} - 0\cdot 1}{\frac{T-t}{T}} = \frac{bT}{(T-t)^2}$$
$$V(t) = \underbrace{V(0)}_{X(0)=a} + \underbrace{\int_0^t \frac{bT}{(T-\tau)^2}d\tau}_{\frac{bT}{T-t}-b} + \int_0^t \tfrac{T}{T-\tau}dZ(\tau)$$
$$X(t) = U(t)\cdot V(t) = \tfrac{T-t}{T}\left[a + \tfrac{bT}{T-t} - b + T\int_0^t \tfrac{1}{T-\tau}dZ(\tau)\right]$$

$$X(t) = a(1-\tfrac{t}{T}) + b\tfrac{t}{T} + (T-t)\int_0^t \tfrac{1}{T-\tau}dZ(\tau),\ 0 \le t < T$$

$X(0) = a$, and $\lim_{t\to T} X(t) = b$ (the reason for the given name)
$X(t)$ follows a normal distribution for any $t \in (0, T)$
$E[X(t)] = a(1-\frac{t}{T}) + b\frac{t}{T}$
$\mathrm{var}(X(t)) = t - \frac{t^2}{T} = \frac{Tt-t^2}{T} = \frac{t(T-t)}{T}$
$\mathrm{cov}(X(t), X(s)) = \min(s,t) - st/T$

**Figure 1-5**




- The Brownian bridge is suited to formulate the process of the zero-coupon bond price because the bond price today is known and the bond value is equal to its face value on the maturity date. The disadvantage of formulating the bond price to follow the Brownian bridge is that the price of a zero-coupon bond could be negative due to the normal distribution of $dZ(t)$ in $dX(t)$.

⊙ Evaluate $E[\max(P_t - K, 0)]$, where $P_t \sim ND(\mu, \sigma^2)$.

$$E[\max(P_t - K, 0)] = \int_K^\infty (P_t - K)\tfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\tfrac{(P_t-\mu)^2}{2\sigma^2}\right)dP_t$$

Consider $x = \frac{P_t - \mu}{\sigma} \Rightarrow dP_t = \sigma dx$

$= \int_{\frac{K-\mu}{\sigma}}^{\infty} (x\sigma + \mu - K)\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx$

$= \sigma \int_{\frac{K-\mu}{\sigma}}^{\infty} x\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx + (\mu - K) \int_{\frac{K-\mu}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx$

$= \sigma \frac{-1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})|_{\frac{K-\mu}{\sigma}}^{\infty} + (\mu - K)[1 - N(\frac{K-\mu}{\sigma})]$

$= \sigma \frac{1}{\sqrt{2\pi}} \exp(-\frac{(\frac{K-\mu}{\sigma})^2}{2}) + (\mu - K)N(\frac{\mu-K}{\sigma})$

$= \sigma n(\frac{\mu-K}{\sigma}) + (\mu - K)N(\frac{\mu-K}{\sigma}),$

where $n(d) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{d^2}{2})$, and $N(c) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx$

- (Supplement) Given $X(t) = a(1 - \frac{t}{T}) + b\frac{t}{T} + (T - t)\int_0^t \frac{1}{T-\tau}dZ(\tau)$,

  prove (i) $\text{var}(X(t)) = \frac{t(T-t)}{T}$.

  (ii) $\text{cov}(X(t), X(s)) = s - \frac{st}{T}$ (if $t > s$).

  (i) According to the fourth property of Itô integral, that is,

  $\text{var}(\int_0^T X(\tau)dZ(\tau)) = \int_0^T E[X^2(\tau)]d\tau$, we can derive

$$\begin{aligned}\text{var}(X(t)) &= (T-t)^2 \int_0^t (\tfrac{1}{T-\tau})^2 d\tau = (T-t)^2\ ((T-\tau)^{-1}|_0^t) \\ &= (T-t)^2(\tfrac{1}{T-t} - \tfrac{1}{T}) = \tfrac{t(T-t)}{T}\end{aligned}$$

  (Note that $a(1 - \frac{t}{T}) + b\frac{t}{T}$ in $X(t)$ contributes nothing to $\text{var}(X(t))$.)

  (ii) $\text{cov}(X(t), X(s))$

  $= \text{cov}(X(s) + X(t) - X(s), X(s))$ (assume $s < t$)

  $= \text{var}(X(s)) + \text{cov}(X(t) - X(s), X(s))$

  $= \frac{s(T-s)}{T} + \text{cov}((T-t)\int_0^t \frac{1}{T-\tau}dZ(\tau) - (T-s)\int_0^s \frac{1}{T-\tau}dZ(\tau), (T-s)\int_0^s \frac{1}{T-\tau}dZ(\tau))$

  $(T-t)\int_0^t \frac{1}{T-\tau}dZ(\tau) - (T-s)\int_0^s \frac{1}{T-\tau}dZ(\tau)$

  $\overset{T-s=T-t+t-s}{\Longrightarrow} (T-t)\int_s^t \frac{1}{T-\tau}dZ(\tau) - (t-s)\int_0^s \frac{1}{T-\tau}dZ(\tau)$

  $= \frac{s(T-s)}{T} - (t-s)(T-s)\ \text{var}(\int_0^s \frac{1}{T-\tau}dZ(\tau))$

  $= \frac{s(T-s)}{T} - (t-s)(T-s)(\int_0^s (\frac{1}{T-\tau})^2 d\tau)$

  $= \frac{s(T-s)}{T} - (t-s)(T-s)(\frac{1}{(T-s)} - \frac{1}{T})$

  $= \frac{sT-s^2}{T} - (t-s)(T-s)(\frac{s}{T(T-s)})$

  $= \frac{sT-s^2-st+s^2}{T} = s - \frac{st}{T}$