Stochastic Process

H.C.

Latest Update: March 24, 2020

Contents

Chapter 1

Preliminaries

1.1 General stochastic process

$n \geq k$ is used as an alternative for the statement $n \in \mathbb{Z}_{\geq k} = \mathbb{Z} \cap [k, \infty)$. $t \geq s$ is used as an alternative for the statement $t \in \mathbb{R}_{\geq s} = \mathbb{R} \cap [s, \infty)$.

Definition 1.1.1 (stochastic process) For a given probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and a measurable space $(S, \mathcal{S})$, a *stochastic process* is a collection of S-valued random variables on $(\Omega, \mathcal{F}, \mathrm{P})$ indexed by some set T, which can be written as $X = \{X_t : X_t \text{ is a random variable on } (\Omega, \mathcal{F}, \mathrm{P}), t \in T\}$ or $X = (X_t)_{t\in T}$ or $X : \Omega \times T \to S$. This mathematical space S is called its state space.

Note the identification (up to bijections) among the collection of mappings $\{X_t \in S^\Omega : \sigma(X_t) \in \mathcal{F}, t \in T\}$, the mapping $X_\cdot(-) : T \to S^\Omega, t \mapsto (\omega \mapsto X_t(\omega))$, the mapping $X_\cdot(\cdot) : \Omega \times T \to S, (\omega, t) \mapsto X_t(\omega)$ and the mapping $X_-(\cdot) : \Omega \to S^T, \omega \mapsto (t \mapsto X_t(\omega))$, each of which can be denoted by X.

The following proposition actually gives an equivalent definition of stochastic process.

Proposition 1.1.1 (measurability of $X : \Omega \to S^T$) Define the evaluation mappings

$$\begin{aligned} \pi_t : S^T &\longrightarrow S \\ f &\longmapsto f(t) \end{aligned}$$

and the σ-algebra on S^T

$$\mathcal{S}^T := \sigma\left(\prod_{t\in T} \pi_t\right).$$

A function $X : \Omega \to S^T$ is $\mathcal{F}/\mathcal{S}^T$-measurable iff $X_t : \Omega \to S$ is $\mathcal{F}/\mathcal{S}$-measurable for every $t \in T$.

Definition 1.1.2 (measurable stochastic process) A stochastic process X defined on $(\Omega, \mathcal{F}, \mathrm{P})$ is measurable if, for all $A \in \mathcal{B}\left(\mathbf{R}^N\right)$

$$\{(t, \omega) : X_t(\omega) \in A\} \in \mathcal{B}([0, \infty)) \otimes \mathcal{F}$$

Definition 1.1.3 (the distribution of a process) The distribution of a process $X : \Omega \to S^T$ is the pushforward measure $P \circ X^{-1}$.

We use the notation $X \overset{d}{=} Y$ to represent that X and Y have the same distribution.

Definition 1.1.4 (family of finite dimensional distributions) The family

$$\mathfrak{D}_X := \left\{\mu_{(t_1,t_2,\cdots,t_k)} := \mathrm{P} \circ (X_{t_1}, \cdots, X_{t_k})^{-1} : (t_1, t_2, \cdots, t_k) \in T^k, k \geq 1\right\}$$

of probability distributions is called the *family of finite dimensional distributions (fdds) associated with the stochastic process* $(X_t)_{t\in T}$.

Proposition 1.1.2 Let X, Y be processes on $(\Omega, \mathcal{F}, \mathrm{P})$ with paths in S^T. Then $X \overset{d}{=} Y$ iff

$$(X_{t_1}, \cdots, X_{t_n}) \overset{d}{=} (Y_{t_1}, \cdots, Y_{t_n}), \quad \forall t_1, \cdots, t_n \in T, \quad n \geq 1.$$

Definition 1.1.5 Let X, Y be stochastic processes from $\Omega \times T$ to S. X is a *modification* of Y iff

$$\forall t \in T,\ \mathrm{P}(X_t = Y_t) = 1$$

and X is *indistinguishable* from Y iff

$$\mathrm{P}(X = Y) = \mathrm{P}(\forall t \in T,\ X_t = Y_t) = 1.$$

If X and Y are indistinguishable, they are modifications of each other.

Proposition 1.1.3 (transfer of regularity) Let S be a a separable metric space and X, Y be processes on $(\Omega, \mathcal{F}, \mathrm{P})$ with paths in $U \subset S^T$ such that $X \overset{d}{=} Y$. Assume that Y has paths in some set $U \subset S^T$ that is Borel for the σ-algebra $\mathcal{U} = (\mathcal{B}(S))^T \cap U$. Then even X has a modification with paths in $U \subset S^T$.

Definition 1.1.6 (Independence of stochastic processes) N stochastic processes $X^{(1)}, X^{(2)}, \cdots, X^{(N)}$ defined on the same probability space $(\Omega, \mathcal{F}, P)$ are said to be independent if for all $n \geq 1$ and for all $t_1, \cdots, t_n \in T$, the N random vectors $\left(X_{t_1}^{(1)}, \cdots, X_{t_n}^{(1)}\right), \left(X_{t_1}^{(2)}, \cdots, X_{t_n}^{(2)}\right), \cdots, \left(X_{t_1}^{(N)}, \cdots, X_{t_n}^{(N)}\right)$ are independent, i.e. if

$$F_{X_{t_1}^{(1)}, \cdots, X_{t_n}^{(1)}, \cdots, X_{t_1}^{(N)}, \cdots, X_{t_n}^{(N)}}\left(x_1^{(1)}, \cdots, x_n^{(1)}, \cdots, x_1^{(N)}, \cdots, x_n^{(N)}\right) = \prod_{i=1}^{N} F_{X_{t_1}^{(i)}, \cdots, X_{t_n}^{(i)}}\left(x_1^{(i)}, \cdots, x_n^{(i)}\right).$$

1.2 Stochastic process valued in $\mathbb{R}$

For simplicity, we always assume that $T = \mathbb{R}_{\geq 0}$ or $T = \mathbb{Z}_{\geq 0}$ and that $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proposition 1.2.1 (consistency conditions) Given a family of finite dimensional distributions,

$$\mathfrak{D} = \left\{\mu_{(\alpha_1, \alpha_2, \cdots, \alpha_k)} : (\alpha_1, \alpha_2, \cdots, \alpha_k) \in T^k, k \geq 1\right\},$$

it satisfies the following consistency conditions: for any $k \geq 2$, $(t_1, t_2, \cdots, t_k) \in T^k$, and any $B_1, B_2, \cdots, B_k$ in $\mathcal{B}(\mathbb{R})$,

(C1) $\mu_{(t_1, t_2, \cdots, t_k)}(B_1 \times \cdots \times B_{k-1} \times \mathbb{R}) = \mu_{(t_1, t_2, \cdots, t_k)}(B_1 \times \cdots \times B_{k-1})$

(C2) For any permutation $(i_1, i_2, \cdots, i_k)$ of $(1, 2, \cdots, k)$,

$$\mu_{\left(t_{i_1}, t_{i_2}, \cdots, t_{i_k}\right)}(B_{i_1} \times B_{i_2} \times \cdots \times B_{i_k}) = \mu_{(t_1, t_2, \cdots, t_k)}(B_1 \times B_2 \times \cdots \times B_k)$$

Theorem 1.2.1 (Kolmogorov's consistency theorem) Let T be a nonempty set. Let

$$\mathfrak{D}_T = \left\{\nu_{(\alpha_1, \alpha_2, \cdots, \alpha_k)} : (\alpha_1, \alpha_2, \cdots, \alpha_k) \in T^k, k \geq 1\right\}$$

be a family of probability distributions such that for each $(t_1, t_2, \cdots, t_k) \in T^k, k \geq 1$

(i) $\nu_{(\alpha_1, \alpha_2, \cdots, \alpha_k)}$ is a probability distribution on $\left(\mathbb{R}^k, \mathcal{B}\left(\mathbb{R}^k\right)\right)$

(ii) consistency conditions C1 and C2 hold

Then, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stochastic process $(X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, P)$ such that D_T is the family of finite dimensional distributions associated with $(X_t)_{t\in T}$.

Definition 1.2.1 (strictly stationary process) Let $(X_t)_{t\in T}$ be a stochastic process and let $F_X(x_{t_1+\tau}, \cdots, x_{t_k+\tau})$ represent the distribution function of the joint distribution of $(X_t)_{t\in T}$ at times $t_1 + \tau, \cdots, t_k + \tau$. Then, $(X_t)_{t\in T}$ is said to be strictly stationary if, for all k, for all τ, and for all $t_1, \cdots, t_k$,

$$F_X(x_{t_1+\tau}, \cdots, x_{t_k+\tau}) = F_X(x_{t_1}, \cdots, x_{t_k}).$$

Definition 1.2.2 (independent increments) A stochastic process $(X_t)_{t\in T}$ has *independent increments* if for every $n \in \mathbb{N}_+$ and any $t_1 \le t_2 \le \cdots \le t_n$, the increments $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.

Definition 1.2.3 (stationary increments) A stochastic process $(X_t)_{t\in T}$ has *stationary increments* if for all $s < t$, the probability distribution of the increments $X_t - X_s$ depends only on $t - s$.

Chapter 2

Poisson Process

2.1 Poisson Process

Definition 2.1.1 (Poisson process (I)) A stochastic process $(N_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is said to be a *Poisson process* with rate $\lambda > 0$ if

(i) $N_0 = 0$;

(ii) $(N_t)_{t\geq 0}$ has independent increments: for any $n \in \mathbb{N}_+$ and any $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, the increment $N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \cdots, N_{t_n} - N_{t_{n-1}}$ are independent;

(iii) for any $0 \leq s < t$, $N_t - N_s \sim Pois(\lambda(t-s))$, that is

$$\mathrm{P}(N_t - N_s = k) = e^{-\lambda(t-s)} \frac{\lambda(t-s)^k}{k!} \quad (k = 0, 1, 2, \cdots).$$

Definition 2.1.2 (counting process) A *counting process* is a stochastic process $(N_t)_{t\geq 0}$ with values that are non-negative, integer, and non-decreasing:

(i) $N_t \geq 0$;

(ii) N_t is an integer;

(iii) If $0 \leq s \leq t$, then $N_s \leq N_t$.

For any $0 \leq s < t$, the counting process $N_t - N_s$ represents the number of events that occurred on $(s, t]$.

Definition 2.1.3 (Poisson process (II)) A counting process $(N_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is said to be a *Poisson process* with rate $\lambda > 0$ if

(i) $N_0 = 0$;

(ii) $(N_t)_{t\geq 0}$ has independent increments;

(iii) For all $t \geq 0$, $\mathrm{P}(N_{t+h} - N_t = 1) = \lambda h + o(h)$ when $h \to 0$;

(iv) For all $t \geq 0$, $\mathrm{P}(N_{t+h} - N_t \geq 2) = o(h)$ when $h \to 0$;

Definition 2.1.4 (Poisson process (III)) A stochastic process $(N_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is said to be a *Poisson process* with rate $\lambda > 0$ if

$$N_t = \sum_{n=1}^{\infty} n I_{[T_n, T_{n+1})}(t),$$

where $T_n = X_1 + X_2 + \cdots + X_n$ and X_i i.i.d $\sim Exp(\lambda)$ (Here the pdf of $Exp(\lambda)$ is taken as $\lambda e^{-\lambda x} I_{(0,+\infty)}(x)$).

Proposition 2.1.1 Definition 2.1, 2.3, 2.4 are equivalent definitions of Poisson process.

Proof.

- Definition 2.1 $\implies$ Definition 2.3

  Here we are only to show the implication of Definition 2.3(iii) and Definition 2.3(iv). Since $N_{t+h} - N_t \sim Pois(\lambda h)$, when $h \to 0$ we have

$$\begin{aligned} \mathrm{P}(N_{t+h} - N_t = 1) &= e^{-\lambda h}\lambda h = (1 - \lambda h + o(h))\lambda h = \lambda h + o(h), \\ \mathrm{P}(N_{t+h} - N_t \geq 2) &= 1 - \mathrm{P}(N_{t+h} - N_t = 0) - \mathrm{P}(N_{t+h} - N_t = 1) \\ &= 1 - e^{-\lambda h} - e^{-\lambda h}\lambda h \\ &= 1 - (1 - \lambda h + o(h)) - (\lambda h + o(h)) \\ &= o(h). \end{aligned}$$

- Definition 2.3 $\implies$ Definition 2.1

  Only Definition 2.1(iii) needs to be derived. Given the Laplace transform of the nonnegative random variables N_t and N_{t+h}

$$L_{N_t}(u) = \mathrm{E}[e^{-uN_t}], \quad L_{N_{t+h}}(u) = \mathrm{E}[e^{-uN_{t+h}}], \quad u \geq 0,$$

  according to Definition 2.3(ii) we can obtain

$$\begin{aligned} L_{N_{t+h}}(u) &= \mathrm{E}[e^{-uN_{t+h}}] \\ &= \mathrm{E}[e^{-uN_t} e^{-u(N_{t+h} - N_t)}] \\ &= \mathrm{E}[e^{-uN_t}]\mathrm{E}[e^{-u(N_{t+h} - N_t)}] \\ &= L_{N_t}(u)\mathrm{E}[e^{-u(N_{t+h} - N_t)}]. \end{aligned}$$

  Note that

$$\begin{aligned} &\mathrm{E}[e^{-u(N_{t+h} - N_t)}] \\ &= e^0 \mathrm{P}(N_{t+h} - N_t = 0) + e^{-u}\mathrm{P}(N_{t+h} - N_t = 1) + \sum_{j=2}^{\infty} e^{-un}\mathrm{P}(N_{t+h} - N_t = j) \\ &= 1 - \lambda h + o(h) + e^{-u}(\lambda h + o(h)) + o(h) \\ &= 1 - \lambda h + e^{-u}\lambda h + o(h) \quad (h \to 0). \end{aligned}$$

  Denote $g(t+h) = L_{N_{t+h}}(u)$ and $g(t) = L_{N_t}(u)$ for some fixed u and then we get

$$\frac{g(t+h) - g(t)}{h} = \frac{g(t)(1 - \lambda h + e^{-u}\lambda h + o(h)) - g(t)}{h} = g(t)\lambda(e^{-u} - 1) + \frac{o(h)}{h}.$$

  Letting $h \to 0$ yields the differential equation

$$g'(t) = g(t)\lambda(e^{-u} - 1).$$

  The initial condition $g(0) = \mathrm{E}[e^{-uN_0}] = 1$ determines a special solution of the equation

$$g(t) = L_{N_t}(u) = e^{\lambda t(e^{-u} - 1)},$$

which coincides with the Laplace transform of Poisson distribution $Pois(\lambda t)$. Since Laplace transform uniquely determines the distribution, we can thus conclude $N_t \sim Pois(\lambda t)$. Given any $r \geq 0$, define a stochastic process $N'_t = N_{r+t} - N_r$ and we can check that $(N'_t)_{t\geq 0}$ is also a counting process satisfying all the contitions in Definition 2.3. Hence by repeating the proof above we can show $N'_t \sim Pois(\lambda t)$, which is equivalent to Definition 2.1(iii).

- Definition 2.1 $\implies$ Definition 2.4
Let $T_n = \inf\{t \geq 0 : N_t = n\}$ for $n \in \mathbb{N}_+$. Note that given any $t \geq 0$, $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$. Thus we have

$$N_t = \sum_{n=1}^{\infty} nI_{N_t=n} = \sum_{n=1}^{\infty} nI_{[T_n,T_{n+1})}(t).$$

Let $X_1 = T_1, X_n = T_n - T_{n-1} (n \geq 2)$. Since $\mathrm{P}(X_1 > t) = \mathrm{P}(N_t = 0) = e^{-\lambda t}$, we see $X_1 \sim Exp(\lambda)$. Since

$$\mathrm{P}(X_2 > t|X_1 = t_1) = \mathrm{P}(X_2 > t|X_1 = t_1)$$

When $n \geq 2$, since

$$\begin{aligned}
&\mathrm{P}(X_n > t|X_{n-1} = t_{n-1}, \cdots, X_1 = t_1)\\
=&\mathrm{P}(T_n - T_{n-1} > t|T_{n-1} - T_{n-2} = t_{n-1}, \cdots, T_1 = t_1) \qquad (\text{let } s_n = t_n + \cdots + t_1)\\
=&\mathrm{P}(T_n > s_{n-1} + t|T_{n-1} = s_{n-1}, \cdots, T_1 = s_1)\\
=&\mathrm{P}(N_{s_{n-1}+t} = n-1|N_{s_{n-1}} = n-1) \qquad (\text{memoryless property of } (N_t))\\
=&\mathrm{P}(N_{s_{n-1}+t} - N_{s_{n-1}} = 0|N_s = n-1)\\
=&\mathrm{P}(N_{s_{n-1}+t} - N_{s_{n-1}} = 0)\\
=&e^{-\lambda t},
\end{aligned}$$

it is plain to show that $\{X_i\}$ is sequence of independent random variable. Furthermore, we have

$$\mathrm{P}(X_n > t) = \mathrm{E}[\mathrm{P}(X_n > t|X_{n-1}, \cdots, X_1)] = e^{-\lambda t},$$

which implies X_i i.i.d $\sim Exp(\lambda)$,

- Definition 2.4 $\implies$ Definition 2.1
Clearly $N_0 = 0$ holds. Since $T_n = X_1 + X_2 + \cdots + X_n$ and X_i i.i.d $\sim Exp(\lambda)$, we can deduce the jointly probability density function of $(T_1, T_2, \cdots, T_m)$

$$\begin{aligned}
f_S(y_1, y_2, \cdots, y_m) &= f_X(y_1, y_2 - y_1, \cdots, y_m - y_{m-1}) \left|\frac{\partial(x_1, \cdots, x_m)}{\partial(y_1, \cdots, y_m)}\right|\\
&= \lambda^m e^{-\lambda y_m} I_{\{0\leq y_1<\cdots<y_m\}}.
\end{aligned}$$

Thus for any $1 \leq j_1 < j_2 < \cdots < j_n$, the jointly probability density function of $(T_{j_1}, T_{j_2}, \cdots, T_{j_n})$ is

$$\frac{y_1^{j_1-1}}{(j_1-1)!} \frac{(y_2 - y_1)^{j_2-j_1-1}}{(j_2 - j_1 - 1)!} \cdots \frac{(y_n - y_{n-1})^{j_n-j_{n-1}-1}}{(j_n - j_{n-1} - 1)!} \lambda^{j_n} e^{-\lambda y_n} I_{\{0\leq y_1<\cdots<y_n\}}$$

Note

$$N_t = \sum_{n=1}^{\infty} nI_{[T_n,T_{n+1})}(t)$$

implies $\{N_t = n\} = \{T_n \le t < T_{n+1}\}$. For any $n \in \mathbb{N}_+$ and any $0 \le t_1 < t_2 < \cdots < t_n$, we have

$$
\begin{aligned}
&\mathrm{P}(N_{t_1} = j_1, N_{t_1} - N_{t_2} = j_2, \cdots, N_{t_n} - N_{t_{n-1}} = j_n) \\
=& \mathrm{P}(N_{t_1} = j_1, N_{t_1} = j_1 + j_2, \cdots, N_{t_n} = j_1 + \cdots + j_n) \quad (\text{let } k_n = j_1 + \cdots j_n) \\
=& \mathrm{P}(T_{k_1} \le t_1, T_{k_1+1} > t_1, T_{k_2} \le t_2, T_{k_2+1} > t_2, \cdots, T_{k_n} \le t_n, T_{k_n+1} > t_n) \\
=& \int_{y_{k_1} \le t_1 < y_{k_1+1} < \cdots < y_{k_n} \le t_n < y_{k_n+1}} \frac{y_{k_1}^{k_1-1}}{(k_1-1)!} \cdots \frac{(y_{k_n} - y_{k_{n-1}+1})^{k_n - k_{n-1} - 2}}{(k_n - k_{n-1} - 2)!} \lambda^{k_n+1} e^{-\lambda y_{k_n+1}} dy_1 \cdots dy_{k_n+1} \\
=& \int_{y_{k_1} \le t_1 < y_{k_1+1} < \cdots < y_{k_n} \le t_n} \frac{y_{k_1}^{k_1-1}}{(k_1-1)!} \cdots \frac{(y_{k_n} - y_{k_{n-1}+1})^{k_n - k_{n-1} - 2}}{(k_n - k_{n-1} - 2)!} \lambda^{k_n} dy_{k_1} \cdots dy_{k_n} \int_{t_n}^{\infty} -de^{-\lambda y_{k_n+1}} \\
=& \int_{y_{k_1} \le t_1 < y_{k_1+1} < \cdots < t_{n-1} < y_{k_{n-1}+1} < t_n} \\
& \frac{y_{k_1}^{k_1-1}}{(k_1-1)!} \cdots \frac{(y_{k_{n-1}} - y_{k_{n-2}+1})^{k_{n-1} - k_{n-2} - 2}}{(k_n - k_{n-1} - 2)!} \lambda^{k_n} e^{-\lambda t_n} dy_1 \cdots dy_{k_{n-1}+1} \int_{y_{k_{n-1}+1}}^{t_n} d\frac{(y_{k_n} - y_{k_{n-1}+1})^{k_n - k_{n-1} - 1}}{(k_n - k_{n-1} - 1)!} \\
=& \int_{y_{k_1} \le t_1 < y_{k_1+1} < \cdots < t_{n-1} < y_{k_{n-1}+1} < t_n} \\
& \frac{y_{k_1}^{k_1-1}}{(k_1-1)!} \cdots \frac{(y_{k_{n-1}} - y_{k_{n-2}+1})^{k_{n-1} - k_{n-2} - 2}}{(k_n - k_{n-1} - 2)!} \lambda^{k_n} e^{-\lambda t_n} dy_1 \cdots dy_{k_{n-1}} \int_{t_{n-1}}^{t_n} -d\frac{(t_n - y_{k_{n-1}})^{k_n - k_{n-1}}}{(k_n - k_{n-1})!} \\
=& \int_{y_{k_1} \le t_1 < y_{k_1+1} < \cdots < t_{n-1} < y_{k_{n-1}+1} < t_n} \\
& \frac{y_{k_1}^{k_1-1}}{(k_1-1)!} \cdots \frac{(y_{k_{n-1}} - y_{k_{n-2}+1})^{k_{n-1} - k_{n-2} - 2}}{(k_n - k_{n-1} - 2)!} \lambda^{k_n} e^{-\lambda t_n} dy_1 \cdots dy_{k_{n-1}} \frac{(t_n - t_{n-1})^{k_n - k_{n-1}}}{(k_n - k_{n-1})!} \\
=& \cdots \\
=& \lambda^{k_n} e^{-\lambda t_n} \frac{t_1^{k_1}}{k_1!} \frac{(t_2 - t_1)^{k_2 - k_1}}{(k_2 - k_1)!} \cdots \frac{(t_n - t_{n-1})^{k_n - k_{n-1}}}{(k_n - k_{n-1})!} \\
=& e^{-\lambda t_1} \frac{(\lambda t_1)^{j_1}}{j_1} e^{-\lambda(t_2 - t_1)} \frac{(\lambda(t_2 - t_1))^{j_2}}{j_2} \cdots e^{-\lambda(t_n - t_{n-1})} \frac{(\lambda(t_n - t_{n-1}))^{j_n}}{j_n!}
\end{aligned}
$$

Therefore, we conclude $N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \cdots, N_{t_n} - N_{t_{n-1}}$ are independent and for any $0 \le s < t$, $N_t - N_s \sim \mathit{Pois}(\lambda(t-s))$.

□

Proposition 2.1.2 Let $(N_t)_{t \ge 0}$ be a Poisson process.

1. $N_t \sim \mathit{Pois}(\lambda t)$, $\mathrm{E}[N_t] = \mathrm{Var}(N_t) = \lambda t$.

2. For $0 \le s \le t$, $\mathrm{E}[N_t N_s] = \lambda^2 ts + \lambda s$, $\mathrm{Cov}(E_t, E_s) = \lambda s$.

3. For $0 \le s \le t$, $\mathrm{E}[N_t | N_s] = N_s + \lambda(t-s)$. So Poisson process is a submartingale.

4. Poisson process is a Markov process. For $0 \le t_1 < t_2 < \cdots < t_n$ and $0 \le k_1 \le k_2 \le \cdots \le k_n$,

$$
\begin{aligned}
&\mathrm{P}(N_{t_n} = k_n | N_{t_{n-1}} = k_{n-1}, \cdots, N_{t_1} = k_1) \\
=& \mathrm{P}(N_{t_n} = k_n | N_{t_{n-1}} = k_{n-1}) \\
=& \mathrm{P}(N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}) \\
=& e^{-\lambda(t_n - t_{n-1})} \frac{(\lambda(t_n - t_{n-1}))^{k_n - k_{n-1}}}{(k_n - k_{n-1})!}.
\end{aligned}
$$

Proof. Apply Definition 2.1(ii) and it is straightforward to show the properties.

□

2.2 Compound Poisson Process

Definition 2.2.1 (compound Poisson distribution) Suppose that $N \sim Pois(\lambda)$ and that $Z_1, Z_2, Z_3, \cdots$ are i.i.d. random variables independent of N with a probability measure $v(dy)$ on $\mathbb{R}$. Then the probability distribution of the sum of N i.i.d. random variables

$$Y = \sum_{n=1}^{N} Z_n$$

is a *compound Poisson distribution.*

Definition 2.2.2 (compound Poisson process) *A compound Poisson process*, parameterised by a rate $\lambda > 0$ and jump size distribution $v(dy)$, is a process $(Y_t)_{t\geq 0}$ given by

$$Y_t = \sum_{n=1}^{N_t} Z_n,$$

where $(N_t)_{t\geq 0}$ is a Poisson process with rate λ, and $(Z_n)_{n\in\mathbb{N}_+}$ are independent and identically distributed random variables with distribution $v(dy)$, which are also independent of $(N_t)_{t\geq 0}$.

Proposition 2.2.1 Let $(Y_t)_{t\geq 0}$ be a compound Poisson process with a rate λ and jump $(Z_n)_{n\in\mathbb{N}_+}$. For convenience, assume $Z_n \stackrel{d}{=} Z$ and $\mathrm{E}[Z^2] < +\infty$.

1. $\mathrm{E}[Y_t] = \lambda t \mathrm{E}[Z]$.
2. $\mathrm{Var}(Y_t) = \lambda t \mathrm{E}[Z^2]$.
3. The moment generating function $M_{Y_t}(a) = \mathrm{E}[e^{aY_t}] = e^{\lambda t(\mathrm{E}[e^{aZ}]-1)} = e^{\lambda t(M_Z(a)-1)}$

Proof.

1. Since Z_n is independent of N_t, we have

$$\mathrm{E}[Y_t] = \mathrm{E}[\mathrm{E}[Y_t|N_t]] = \mathrm{E}\left[\mathrm{E}\left[\left.\sum_{n=1}^{N_t} Z_n \right| N_t\right]\right] = \mathrm{E}\left[\sum_{n=1}^{N_t} \mathrm{E}\left[\left. Z_n \right| N_t\right]\right] = \mathrm{E}[N_t Z] = \mathrm{E}[N_t]\mathrm{E}[Z] = \lambda t \mathrm{E}[Z].$$

2. Since Z_n is independent of N_t, by the law of total variance $\mathrm{Var}(Y_t)$ can be calculated as

$$\begin{aligned}
\mathrm{Var}(Y_t) &= \mathrm{E}[\mathrm{Var}(Y_t|N_t)] + \mathrm{Var}(\mathrm{E}[Y_t|N_t]) \\
&= \mathrm{E}[N_t \mathrm{Var}(Z)] + \mathrm{Var}(N_t \mathrm{E}[Z]) \\
&= \mathrm{Var}(Z)\mathrm{E}[N_t] + \mathrm{E}[Z]^2 \mathrm{Var}(N_t) \\
&= \lambda t \mathrm{Var}(Z) + \lambda t \mathrm{E}[Z]^2 \\
&= \lambda t \mathrm{E}[Z^2].
\end{aligned}$$

3. Make similar use of the dependence of $(Z_n)_{n\in\mathbb{N}_+}$ and N_t to get

$$\begin{aligned}
\mathrm{E}\left[e^{aY_t}\right] &= \mathrm{E}\left[e^{a\left(Z_1+Z_2+\cdots+Z_{N_t}\right)}\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[e^{a\left(Z_1+Z_2+\cdots+Z_{N_t}\right)}\middle| N_t\right]\right] \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left[e^{a\left(Z_1+Z_2+\cdots+Z_{N_t}\right)}\middle| N_t = n\right] \mathrm{P}\left(N_t = n\right) \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left[e^{a(Z_1+Z_2+\cdots+Z_n)}\middle| N_t = n\right] e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left[e^{aZ_1}e^{aZ_2}\cdots e^{aZ_n}\right] e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left[e^{aZ}\right]^n e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} \frac{\left(\lambda t \mathrm{E}\left[e^{aZ}\right]\right)^n}{n!} \\
&= e^{\lambda t\left(\mathrm{E}\left[e^{aZ}\right]-1\right)}.
\end{aligned}$$

□

Let $Y = (Y_t)_{t\geq 0}$ be a compound Poisson process with a rate λ and jump $(Z_n)_{n\in\mathbb{N}_+}$. Let $T_n = \inf\{t \geq 0 : N_t = n\}$ be the time when the nth event happens. Then the Itô integral of a stochastic process K with respect to Y is

$$\int_0^t K dY = \sum_{n=0}^{N_t} K_{T_n}(Y_{T_n} - Y_{T_n-}) = \sum_{n=0}^{N_t} K_{T_n} Z_n.$$

Chapter 3

Markov Chain

3.1 Discrete-time Markov Chain

Definition 3.1.1 (discrete-time Markov chain) A *discrete-time Markov chain* on a countable state space S is a sequence of random variables $X_0, X_1, X_2, \cdots$ with the Markov property, namely that $\forall n \geq 0,\ \forall j, i_0, i_1, \cdots, i_n \in S$,

$$\mathrm{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = \mathrm{P}(X_{n+1} = j | X_n = i_n),$$

if both conditional probabilities are well defined, i.e., if $\mathrm{P}(X_0 = i_0, \cdots, X_n = i_n) > 0$.

Markov property may be interpreted as stating that the probability of moving to the next state depends only on the present state and not on the previous states.
A discrete-time Markov chain $(X_n)_{n \geq 0}$ is *time-homogeneous* if

$$\mathrm{P}(X_{n+2} = j | X_{n+1} = i) = \mathrm{P}(X_{n+1} = j | X_n = i)$$

for all $n \geq 0$ and all $i, j \in S$. We only focus on the discrete-time time-homogeneous Markov chain on a countable space if nothing is specified.
One can consider the (possibly infinite-dimensional) *transition matrix* $P = (p_{ij})_{i,j \in S}$, where

$$p_{ij} = \mathrm{P}(X_{n+1} = j | X_n = i)$$

is called *one-step transition probability.* By the law of total probability it is clear to see

$$\sum_{j \in S} p_{ij} = 1, \quad \forall i \in S.$$

Similarly we can define *n-step transition matrix* $P^{(n)} = \left(p_{ij}^{(n)}\right)_{i,j \in S}$, where *n-step transition probabilities* $p_{ij}^{(n)}$ is the probability that a process in state i will be in state j after n additional transitions. That is,

$$p_{ij}^{(n)} = \mathrm{P}(X_{k+n} = j | X_k = i).$$

As is shown below, the Chapman–Kolmogorov equation enables us to calculate n-step transition matrix readily.

Proposition 3.1.1 (Chapman–Kolmogorov equation) Let $(X_n)_{n \geq 0}$ be a discrete-time Markov chain on a countable state space S. The *Chapman–Kolmogorov equation* states that

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}, \quad \forall i, j \in S,$$

or alternatively in matrix form

$$P^{(n+m)} = P^{(n)} P^{(m)}.$$

Proof.

$$\begin{aligned}
p_{ij}^{(n+m)} &= \mathrm{P}(X_{n+m} = j | X_0 = i) \\
&= \sum_{k \in S} \mathrm{P}(X_{n+m} = j, X_n = k | X_0 = i) \\
&= \sum_{k \in S} \mathrm{P}(X_n = k | X_0 = i) \mathrm{P}(X_{n+m} = j | X_0 = i, X_n = k) \\
&= \sum_{k \in S} \mathrm{P}(X_n = k | X_0 = i) \mathrm{P}(X_{n+m} = j | X_n = k) \\
&= \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}
\end{aligned}$$

□

Of course $P^{(1)} = P$. Thus by iteration we show $P^{(n)}$ coincides with P^n.
Let $\pi(n) = \left(p_i^{(n)}\right)_{i \in S}$ denote the probability distribution of X_n, where $p_i^{(n)} = \mathrm{P}(X_n = i)$. Then we have

$$\pi(n) = \pi(0) P^n.$$

3.2 Continuous-time Markov Chain

Definition 3.2.1 (continuous-time Markov chain) A *continuous-time Markov chain* on a countable state space S is a stochastic $(X_t)_{t \geq 0}$ with the Markov property: for all $n \geq 0$, all $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$, and all $j, i_0, \cdots, i_n \in S$,

$$\mathrm{P}(X_{t_{n+1}} = j | X_{t_0} = i_1, X_{t_1} = i_2, \cdots, X_{t_n} = i_n) = \mathrm{P}(X_{t_{n+1}} = j | X_{t_n} = i_n),$$

if both conditional probabilities are well defined, i.e., if $\mathrm{P}(X_1 = i_1, \cdots, X_n = i_n) > 0$.

A continuous-time Markov chain $(X_t)_{t \geq 0}$ is *time-homogeneous* if

$$\mathrm{P}(X_{s+t} = j | X_s = i) = \mathrm{P}(X_t = j | X_0 = i) = p_{ij}(t)$$

for all $s, t \geq 0$ and all $i, j \in S$.

Example 3.2.1 Poisson progress $(N_t)_{t \geq 0}$ is continuous-time time-homogeneous Markov chain.

We only focus on the continuous-time time-homogeneous Markov chain on a countable space if nothing is specified. In this case, transition probability $p_{ij}(t)$ depends on the interarrival t from state i to state j. One can define the *transition matrix* $P(t) = (p_{ij}(t))_{i,j \in S}$ and we have the similar properties of $p_{ij}(t)$ like the concrete-time case.

Proposition 3.2.1 Let $(X_t)_{t \geq 0}$ be a continuous-time Markov chain on a countable state space S.

1. For all $t \geq 0$ and all $j \in S$,

$$\sum_{j \in S} p_{ij}(t) = 1.$$

2. The *Chapman–Kolmogorov equation* states that

$$p_{ij}(t+s) = \sum_{k \in S} p_{ik}(t) p_{kj}(s), \quad \forall i, j \in S,$$

   or alternatively in matrix form

$$P(t+s) = P(t) P(s).$$

Proof.

1. It follows by the law of total probability.
2. Imitate the proof in Proposition 4.1 and the result is straightforward.

$\square$

Definition 3.2.2 A continuous-time Markov chain is *regular* if it satisfy the following condition

$$\lim_{t\to 0} p_{ij}(t) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Since $p_{ij}(0) = \mathrm{P}(X_t = j|X_0 = i) = \delta_{ij}$, regularity implies $p_{ij}(t)$ is continuous at $t = 0$. Furthermore, we have the following stronger results. We always assume continuous-time Markov chain is regular afterwards.

Lemma 3.2.1 If a continuous-time Markov chain is regular, for any fixed $i, j \in S$, $p_{ij}(t)$ is uniformly continuous with respect to t.

Proof. Since when $h > 0$

$$\begin{aligned} p_{ij}(t+h) - p_{ij}(t) &= \sum_{k\in S} p_{ik}(h)p_{kj}(t) - p_{ij}(t) \\ &= p_{ii}(h)p_{ij}(t) - p_{ij}(t) + \sum_{k\in S\setminus\{i\}} p_{ik}(h)p_{kj}(t) \\ &= -(1-p_{ii}(h))p_{ij}(t) + \sum_{k\in S\setminus\{i\}} p_{ik}(h)p_{kj}(t), \end{aligned}$$

we have

$$\begin{aligned} &p_{ij}(t+h) - p_{ij}(t) \geq -(1-p_{ii}(h))p_{ij}(t) \geq -(1-p_{ii}(h)), \\ &p_{ij}(t+h) - p_{ij}(t) \leq \sum_{k\in S\setminus\{i\}} p_{ik}(h)p_{kj}(t) \leq \sum_{k\in S\setminus\{i\}} p_{ik}(h) = 1 - p_{ii}(h), \end{aligned}$$

which implies

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h).$$

When $h < 0$ in a similar way we can get

$$|p_{ij}(t) - p_{ij}(t+h)| \leq 1 - p_{ii}(-h).$$

In general

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(|h|).$$

According to regular condition we conclude for any $t \geq 0$,

$$\lim_{h\to 0} |p_{ij}(t+h) - p_{ij}(t)| = 0,$$

that is, $p_{ij}(t)$ is uniformly continuous with respect to t on $[0, \infty)$.

$\square$

If $p_{ij}(t)$ is differentiable, define the *transition rate*

$$q_{ij} = \left.\frac{dp_{ij}(t)}{dt}\right|_{t=0} = \lim_{h\to 0} \frac{p_{ij}(h) - p_{ij}(0)}{h}.$$

The q_{ij} can be seen as measuring how quickly the transition from i to j happens. Then define the *transition rate matrix* $Q = (q_{ij})_{i,j \in S}$ with dimensions equal to that of the state space.
Since $P(0) = I$, it can be shown that

$$\mathrm{P}(X_{t+h} = j | X_t = i) = p_{ij}(h) = \delta_{ij} + q_{ij}h + o(h).$$

As an intermediate consequence, if the state space S is finite, we have

$$\sum_{j \in S} q_{ij} = 0,$$

which follows by

$$1 = \sum_{j \in S} p_{ij}(h) = \sum_{j \in S} [\delta_{ij} + q_{ij}h + o(h)] = 1 + \sum_{j \in S} q_{ij}h + o(h).$$

Theorem 3.2.1 Let $(X_t)_{t \geq 0}$ be a continuous-time Markov chain on a countable state space S.

1. Kolmogorov forward equation:
$$p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) q_{kj}$$
or alternatively
$$P'(t) = P(t)Q(t).$$

2. Kolmogorov backward equation:
$$p'_{ij}(t) = \sum_{k \in S} q_{ik} p_{kj}(t)$$
or alternatively
$$P'(t) = Q(t)P(t).$$

Chapter 4

Brownian Motion

4.1 1-dimensional Brownian Motion

Definition 4.1.1 (Brownian motion) A stochastic process $(B_t)_{t\geq 0}$ is called a *Brownian motion* if

1. $B_0 = 0$ a.s.
2. $(B_t)_{t\geq 0}$ has continuous path, that is $t \mapsto B_t$ is almost surely continuous.
3. $(B_t)_{t\geq 0}$ has independent and stationary increments.
4. For $t > 0$, $B_t \sim N(0, t)$.

Definition 4.1.2 (Gaussian process) A stochastic process $(X_t)_{t\in T}$ is a *Gaussian process* if and only if for every finite set of indices $t_1, \cdots, t_n$ in the index set T, $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ follows multivariate normal distribution $N(\mu, \Sigma)$.

Theorem 4.1.1 $B = (B_t)_{t\geq 0}$ is a Brownian motion if and only if B is a Gaussian process satisfying

1. $B_0 = 0$,
2. B has continuous paths,
3. For all $t \geq 0$, $\mathrm{E}[B_t] = 0$,
4. For all $s, t \geq 0$, $\mathrm{E}[B_s B_t] = s \wedge t$.

Proposition 4.1.1 Let $B = (B_t)_{t\geq 0}$ be a Brownian motion.

1. For $k \geq 1$, $\mathrm{E}[B_t^{2k-1}] = 0$, $\mathrm{E}[B_t^{2k}] = t^k(2k-1)!!$.
2. B is a Markov process.
3. B is a martingale.

Theorem 4.1.2 The quadratic variation of a Brownian motion B exists, and is given by $\langle B\rangle_t = t$.

Proof. Given a partition P of the interval $[0, t]$, we have

$$\begin{aligned}
\mathrm{E}\left[\sum_{k=1}^{n}(B_{t_k} - B_{t_{k-1}})^2\right] &= \sum_{k=1}^{n}\mathrm{E}\left[(B_{t_k} - B_{t_{k-1}})^2\right] \\
&= \sum_{k=1}^{n}(t_k - t_{k-1}) \\
&= t
\end{aligned}$$

and

$$
\begin{aligned}
\operatorname{Var}\left(\sum_{k=1}^{n}(B_{t_k}-B_{t_{k-1}})^2\right) &= \sum_{k=1}^{n}\operatorname{Var}\left((B_{t_k}-B_{t_{k-1}})^2\right) \\
&= \sum_{k=1}^{n}\mathrm{E}\left[(B_{t_k}-B_{t_{k-1}})^4\right] - \sum_{k=1}^{n}\left(\mathrm{E}\left[(B_{t_k}-B_{t_{k-1}})^2\right]\right)^2 \\
&= \sum_{k=1}^{n}3(t_k-t_{k-1})^2 - \sum_{k=1}^{n}(t_k-t_{k-1})^2 \\
&= 2\sum_{k=1}^{n}(t_k-t_{k-1})^2 \\
&\le 2\|P\|\sum_{k=1}^{n}(t_k-t_{k-1}) \\
&= 2\|P\|t.
\end{aligned}
$$

Since

$$
\lim_{\|P\|\to 0}\mathrm{E}\left[\left(\sum_{k=1}^{n}(B_{t_k}-B_{t_{k-1}})^2 - t\right)^2\right] = \lim_{\|P\|\to 0}\operatorname{Var}\left(\sum_{k=1}^{n}(B_{t_k}-B_{t_{k-1}})^2\right) \le \lim_{\|P\|\to 0} 2\|P\|t = 0,
$$

we conclude

$$
[B]_t = \lim_{\|P\|\to 0}\sum_{k=1}^{n}(B_{t_k}-B_{t_{k-1}})^2 = t \quad \text{a.s.}
$$

□

4.2 N-dimensional Brownian Motion

Definition 4.2.1 (N-dimensional Brownian Motion) The N-dimemsional stochastic process $B = \left(B^{(1)}, B^{(2)}, \cdots, B^{(N)}\right)$ is *a* (standard) N-dimensional Brownian motion if the N-components $B^{(i)}$ are independent one-dimensional standard Brownian motions.

Theorem 4.2.1 (Lévy characterisation) Let $M = \left(M^{(1)}, M^{(2)}, \cdots, M^{(N)}\right)$ be a N-dimensional stochastic process where $M^{(i)} \in \mathscr{M}_0^{\mathrm{loc}}(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathrm{P})$. Then M is an N-dimensional Brownian Motion if and only if

$$
\left\langle M^{(i)}, M^{(j)} \right\rangle_t = \delta_{ij} t, \qquad \forall t \ge 0.
$$

Chapter 5

Martingale

5.1 Basic Notion

Definition 5.1.1 (filtration) Suppose $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and let T be a linearly ordered index set such as $\mathbb{N}$ or $\mathbb{R}_{\geq 0}$. For every $t \in T$ let $\mathcal{F}_t$ be a sub-σ-algebra of $\mathcal{F}$. Then

$$\mathbb{F} = (\mathcal{F}_t)_{t \in T}$$

is called a *filtration* on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ if $\mathcal{F}_k \subset \mathcal{F}_\ell$ for all $k \leq \ell$.

Definition 5.1.2 (filtered probability space) If $(\Omega, \mathcal{F}, \mathrm{P})$ is a probability space and $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$ is a filtration on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, then $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathrm{P})$ is called a *filtered probability space*.

Definition 5.1.3 (adapted process) A stochastic process $X = (X_t)_{t \in T}$ is called *adapted* (to the filtration $(\mathcal{F}_t)_{t \in T}$) if for any $t \in T, X_t$ is $\mathcal{F}_t$ -measurable, that is, $\sigma(X_t) \subset \mathcal{F}_t$ for all $t \in T$.

A stochastic process $(X_t)_{t \in T}$ is said to be defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathrm{P})$ if $(X_n)_{t \in T}$ is defined on $(\Omega, \mathcal{F}, \mathrm{P})$ and adapted to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$.

The definition of $\mathcal{F}_\infty$ is usually specified as

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{t \in T} \mathcal{F}_t\right)$$

Definition 5.1.4 (right-continuous filtration) Given a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$, define

$$\mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s \quad , \forall t \in T.$$

Then $\mathbb{F}^+ := (\mathcal{F}_{t+})_{t \in T}$ is a filtration. The filtration $\mathbb{F}$ is called *right-continuous* if and only if $\mathbb{F}^+ = \mathbb{F}$.

Definition 5.1.5 (complete filtration) Let

$$\mathcal{N}_P := \{A \subset 2^\Omega | A \subset B \text{ for a } B \text{ with } \mathrm{P}(B) = 0\}$$

be the set of all sets that are contained within a P-null set. A filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$ is called a *complete filtration*, if every $\mathcal{F}_t$ contains $\mathcal{N}_P$. This is equivalent to $(\Omega, \mathcal{F}_t, \mathrm{P})$ being a complete measure space for every $t \in T$.

Definition 5.1.6 (natural filtration) Let $X = (X_t)_{t\geq 0}$ be a stochastic process on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Then

$$\mathcal{F}_t^X = \sigma(\{X_s : 0 \leq s \leq t\})$$

is a σ-algebra and $(\mathcal{F}_t^X)_{t\geq 0}$ is a filtration that X is adapted to. We call $(\mathcal{F}_t^X)_{t\geq 0}$ the *natural filtration* induced by the stochastic process X. $(\mathcal{F}_t^X)_{t\geq 0}$ is the minimum filtration which X is adapted to.

Definition 5.1.7 (stopping time) Let τ be a random variable defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathrm{P})$ with values in T. Then τ is called a *stopping time* (with respect to the filtration $(\mathcal{F}_t)_{t\in T}$), if the following condition holds:

$$\forall t \in T,\ \{\tau \leq t\} \in \mathcal{F}_t$$

or equivalently

$$X_t := 1_{\tau\leq t} = \begin{cases} 1 & \text{if } \tau \leq t \\ 0 & \text{if } \tau > t \end{cases}$$

is adapted to $(\mathcal{F}_t)_{t\in T}$.

Definition 5.1.8 (stopped process) Let $X = (X_t)_{t\in T}$ be a stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathrm{P})$ and τ be a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\in T}$. The stopping process X^τ is defined as $(X_{\tau\wedge t})_{t\in T}$, where

$$\begin{aligned} X_{\tau\wedge t} : \Omega &\longrightarrow S \\ \omega &\longmapsto X_{\tau(\omega)\wedge t}(\omega). \end{aligned}$$

It is useful to observe that, if μ is another stopping time, then

$$(X^\tau)^\mu = (X^\mu)^\tau = X^{\mu\wedge\tau}.$$

Proposition 5.1.1 Let $(X_t)_{t\geq 0}$ be an adapted process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathrm{P})$ with values in a metric space (E, d).

1. Assume that the sample paths of X are right-continuous, and let O be an open subset of E. Then

   $$\tau_O = \inf\{t \geq 0 : X_t \in O\}$$

   is a stopping time of the filtration $\mathbb{F}^+$.

2. Assume that the sample paths of X are continuous, and let F be a closed subset of E. Then

   $$\tau_F = \inf\{t \geq 0 : X_t \in F\}$$

   is a stopping time of the filtration $\mathbb{F}$.

5.2 Discrete-time Martingale

Definition 5.2.1 (discrete-time martingale) A discrete-time stochastic process $M = (M_n)_{n\geq 0}$ defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ is a *martingale* if it satisfies

1. For $n \geq 0$, $\mathrm{E}[|M_n|] < +\infty$;

2. For $n \geq 0$, $\mathrm{E}[M_{n+1}|\mathcal{F}_n] = M_n$.

Definition 5.2.2 (discrete-time submartingale) A discrete-time *submartingale* is a stochastic process $M = (M_n)_{n\geq 0}$ consisting of integrable random variables satisfying for $n \geq 0$

$$\mathrm{E}[M_{n+1}|\mathcal{F}_n] \geq M_n.$$

Definition 5.2.3 (discrete-time supermartingale) A discrete-time *supermartingale* is a stochastic process $M = (M_n)_{n\geq 0}$ consisting of integrable random variables satisfying for $n \geq 0$

$$\mathrm{E}[M_{n+1}|\mathcal{F}_n] \leq M_n.$$

Example 5.2.1 Suppose $(M_n)_{n\geq 0}$ is a martingale defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and that $\phi : \mathbb{R} \to \mathbb{R}$ is convex. If $\phi(M_n)$ is integrable for $n \geq 0$, then $(\phi(M_n))_{n\geq 0}$ is a submartingale.

Definition 5.2.4 (stopping time in discrete-time case) Let τ be a random variable defined on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with values in $\mathbb{N}\cup\{+\infty\}$. Then τ is called a stopping time (with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$), if the following condition holds:

$$\forall n \in \mathbb{N},\ \{\tau \leq n\} \in \mathcal{F}_n$$

or equivalently

$$\forall n \in \mathbb{N},\ \{\tau = n\} \in \mathcal{F}_n.$$

Since $\{\tau = \infty\}^c = \bigcup\limits_{n\geq 0} \{\tau = n\} \in \mathcal{F}_\infty$, we can deduce that $\{\tau = \infty\} \in \mathcal{F}_\infty$.

Example 5.2.2 Given a discrete-time stochastic process $(X_n)_{n\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and a Borel set B,

$$\tau = \inf\{n \geq 0 : X_n \in B\}$$

is a stopping time of the filtration $(\mathcal{F}_n)_{n\geq 0}$, called the *first hitting time*. ($\inf \varnothing = \infty$)

Definition 5.2.5 (martingale transform) The process $\widetilde{M} = (\widetilde{M}_n)_{n\geq 0}$ defined by setting $\widetilde{M}_0 = M_0$ and by setting

$$\widetilde{M}_n = M_0 + A_1 (M_1 - M_0) + A_2 (M_2 - M_1) + \cdots + A_n (M_n - M_{n-1})$$

for $n \geq 1$ is called the martingale transform of M by A.

Theorem 5.2.1 (martingale transform theorem) If $M = (M_n)_{n\geq 0}$ is a martingale defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and if $A = (A_n)_{n\geq 0}$ is predicted process with respect to $(\mathcal{F}_n)_{n\geq 0}$, then the martingale transform $\widetilde{M}$ of M by A is itself a martingale with respect to $(\mathcal{F}_n)_{n\geq 0}$.

Theorem 5.2.2 (stopping time theorem) If $M = (M_n)_{n\geq 0}$ is a martingale defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and τ is a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$, then the stopped process $M^\tau = (M_{\tau\wedge n})_{n\geq 0}$ is also a martingale defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and $\mathrm{E}[M_{\tau\wedge n}] = \mathrm{E}[M_0]$ for $n \geq 0$.

Theorem 5.2.3 (Doob's optional sampling theorem) Let $M = (M_n)_{n\geq 0}$ be a martingale defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ and τ be a stopping time with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$. Suppose $\mathrm{P}(\tau < \infty) = 1$ and M^τ is L^1-bounded, then $\mathrm{E}[M_\tau] = \mathrm{E}[M_0]$.

Proof. Since $\mathrm{P}(\tau < \infty) = 1$, $X_{\tau\wedge n} \xrightarrow{a.s.} X_\tau$ and $|X_\tau| \leq K < \infty$ and hence $\mathrm{E}[|X_\tau|] < \infty$. Thus, $\mathrm{E}[|X_\tau - X_{\tau\wedge n}|] \leq 2K\mathrm{P}(\tau > n) \to 0$.

□

5.3 Continuous-time Martingale

Definition 5.3.1 (continuous-time martingale) A continuous-time stochastic process $M = (M_t)_{t\geq 0}$ defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ is a *martingale* if it satisfies

1. For $t \geq 0$, $\mathrm{E}[|M_t|] < +\infty$, that is, M_t is L^1-bounded;
2. For $0 \leq t \leq s < +\infty$, $\mathrm{E}[\, M_s | \mathcal{F}_t \,] = M_t$.

Definition 5.3.2 (continuous martingale) A continuous-time martingale $M = (M_t)_{t\geq 0}$ defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ is continuous if the paths of M are almost surely continuous. That is, there exists an $\Omega_0 \subset \Omega$ with $P(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ the function

$$\begin{aligned} \gamma_\omega : [0, \infty) &\longrightarrow \mathbb{R} \\ t &\longmapsto X_t(\omega) \end{aligned}$$

is continuous.

Definition 5.3.3 (L^p-bounded martingale) A martingale $M = (M_t)_{t\geq 0}$ is said to be L^p-*bounded* if

$$\sup_{t\geq 0} \mathrm{E}\left[|M_t|^p\right] < \infty.$$

Definition 5.3.4 (uniform integrability) A class $\mathcal{C}$ of random variables is called *uniformly integrable* if given $\varepsilon > 0$, there exists $K \in [0, \infty)$ such that

$$\mathrm{E}\left(|X| 1_{|X|\geq K}\right) \leq \varepsilon \text{ for all } X \in \mathcal{C}.$$

Theorem 5.3.1 (Doob's maximal inequalities in continuous time) If $M = (M_t)_{t\geq 0}$ is a continuous nonnegative submartingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ and $\lambda > 0$, then for all $p \geq 1$ we have

$$\lambda^p \mathrm{P}\left(\sup_{0\leq t\leq T} M_t > \lambda\right) \leq \mathrm{E}\left[M_T^p\right]$$

and, if $M_T \in L^p(\Omega, \mathcal{F}, \mathrm{P})$ for some $p > 1$, then we also have

$$\left\|\sup_{0\leq t\leq T} M_t\right\|_p \leq \frac{p}{p-1} \left\|M_T\right\|_p.$$

Theorem 5.3.2 (martingale convergence theorems in continuous time) Let $M = (M_t)_{t\geq 0}$ be a continuous martingale on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$. If M satisfies $\mathrm{E}\left[|M_t|^p\right] \leq B < \infty$ for some $p > 1$ and all $t \geq 0$, then there exists a random variable $M_\infty \in L^p(\Omega, \mathcal{F}, \mathrm{P})$ with $\mathrm{E}\left[|M_\infty|^p\right] \leq B$ such that

$$\mathrm{P}\left(\lim_{t\to\infty} M_t = M_\infty\right) = 1 \text{ and } \lim_{t\to\infty} \left\|M_t - M_\infty\right\|_p = 0.$$

Also, if M satisfies $\mathrm{E}\left[|M_t|\right] \leq B < \infty$ for all $t \geq 0$, then there exists a random variable $M_\infty \in L^1(\Omega, \mathcal{F}, \mathrm{P})$ with $\mathrm{E}\left[|M_\infty|\right] \leq B$ such that

$$\mathrm{P}\left(\lim_{t\to\infty} M_t = M_\infty\right) = 1.$$

According to the theorem 5.3.2, M_∞ is well defined for any L^p-bounded martingale M.

Proposition 5.3.1 (Hilbert spaces $\mathscr{M}_0^2$ and $\mathscr{M}_{0,c}^2$) Let $\mathscr{M}_0$ denote the collection of all the martingales M defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ with initial value $M_0 = 0$ a.s.. All the L^2-bounded martingales $M \in \mathscr{M}_0$ constitute a Hilbert space, which is denoted by $\mathscr{M}_0^2$, with the inner product defined as

$$(M, N)_{\mathscr{M}_0^2} := (M_\infty, N_\infty)_{L^2} = \mathrm{E}\left[M_\infty N_\infty\right].$$

All the L^2-bounded continuous martingales $M \in \mathscr{M}_0$ constitute a Hilbert space $\mathscr{M}_{0,c}^2$, which is a closed subspace of $\mathscr{M}_0^2$. It follows that $\mathscr{M}_{0,c}^2 \subset \mathscr{M}_0^2 \subset \mathscr{M}_0$.

Definition 5.3.5 (quadratic variation) Suppose that $X = (X_t)_{t\geq 0}$ is a real-valued stochastic process defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$. The *quadratic variation* of X (if exists) is defined as the stochastic process $\langle X\rangle = (\langle X\rangle_t)_{t\geq 0}$ satisfying that for all $t \geq 0$, for all $\varepsilon > 0$,

$$\lim_{\left\|P_{[0,t]}\right\|\to 0} \mathrm{P}\left(\left|\sum_{k=1}^{n}(X_{t_k} - X_{t_{k-1}})^2 - \langle X\rangle_t\right| > \varepsilon\right) = 0$$

where $P_{[0,t]} = \{(t_0, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = t\}$ ranges over partitions of the interval $[0, t]$ and the norm of the partition $P_{[0,t]}$ is the length of the longest of these subintervals, namely

$$\left\|P_{[0,t]}\right\| = \max_{1\leq k\leq n} |t_k - t_{k-1}|.$$

A càdlàg (French: "continue à droite, limite à gauche"), RCLL ("right continuous with left limits"), or corlol ("continuous on (the) right, limit on (the) left") function is a function defined on the real numbers (or a subset of them) that is everywhere right-continuous and has left limits everywhere.

Theorem 5.3.3 (Doob–Meyer decomposition theorem) Let Z be a càdlàg supermartingale satisfying

1. $Z_0 = 0$ a.s;
2. The collection $\{Z_\tau | \tau \text{ is a finite-valued stopping time}\}$ is uniformly integrable.

Then there exists a unique, increasing, predictable process A with $A_0 = 0$ such that $M = Z - A$ is a uniformly integrable martingale.

Definition 5.3.6 (finite variation) A process X is said to have *finite variation* if it has bounded variation over every finite time interval with probability 1.

The quadratic variation exists for all continuous finite variation processes, and is zero.

Proposition 5.3.2 If $M \in \mathscr{M}_{0,c}^2$, then its quadratic variation $\langle M\rangle$ exists and has finite variation. The almost sure limit of $\langle M\rangle_t$ as $t \to \infty$ exists and is denoted by

$$\langle M\rangle_\infty = \lim_{t\to\infty} \langle M\rangle_t \quad \text{a.s.}$$

Moreover, $\langle M\rangle_\infty$ is integrable, and satisfies

$$\mathrm{E}\left[\langle M\rangle_\infty\right] = \mathrm{E}\left[M_\infty^2\right].$$

Definition 5.3.7 (bracket process) The bracket process of two processes X and Y is

$$\langle X, Y\rangle := \frac{1}{4}\left(\langle X + Y\rangle - \langle X - Y\rangle\right)$$

if both $\langle X + Y\rangle$ and $\langle X - Y\rangle$ exist.

Proposition 5.3.3 If $M, N \in \mathscr{M}_{0,c}^2$, then $\langle M, N\rangle$ exists and $MN - \langle M, N\rangle$ is a uniformly integrable martingale. Consequently, the almost sure limit of $\langle M, N\rangle_t$ as $t \to \infty$ exists and is denoted by

$$\langle M, N\rangle_\infty = \lim_{t\to\infty} \langle M, N\rangle_t \quad \text{a.s.}$$

Moreover, $\langle M, N\rangle_\infty$ is integrable, and satisfies

$$\mathrm{E}\left[\langle M, N\rangle_\infty\right] = \mathrm{E}\left[M_\infty N_\infty\right].$$

Proposition 5.3.4 For all $\alpha, \beta \in \mathbb{R}$, $M, M', N \in \mathscr{M}_{0,c}^2$,

1. $\langle \alpha M + \beta M', N\rangle = \alpha\langle M, N\rangle + \beta \langle M', N\rangle$
2. $\langle M, N\rangle = \langle N, M\rangle$
3. $\langle M, M\rangle = \langle M\rangle \geq 0$ and $\langle M\rangle = 0 \iff M = 0$

Proposition 5.3.5

$$\int_0^t |X_s|\,|Y_s|\,d\langle M, N\rangle_s \leq \left(\int_0^t X_s^2 d\langle M\rangle_s\right)^{\frac{1}{2}} \left(\int_0^t Y_s^2 d\langle N\rangle_s\right)^{\frac{1}{2}}.$$

Proposition 5.3.6

$$\langle M^\tau, N^\tau\rangle = \langle M^\tau, N\rangle = \langle M, N\rangle^\tau.$$

5.4 Continuous Local Martingale

Definition 5.4.1 (continuous local martingale) An adapted process $M = (M_t)_{t\geq 0}$ with continuous sample paths is called a *continuous local martingale* if there exists a nondecreasing sequence $(\tau_n)_{n\geq 0}$ of stopping times such that $\tau_n \uparrow \infty$ and, for every n, the stopped process M^{τ_n} is a martingale.

The sequence of stopping times $(\tau_n)_{n\geq 0}$ is called the *localizing sequence* for (or is said to reduce) M if $\tau_n \uparrow \infty$ and, for every n, the stopped process M^{τ_n} is a martingale.

Proposition 5.4.1 (linear space $\mathscr{M}_{\mathbf{0,c}}^{\mathbf{loc}}$) Let $\mathscr{M}_0^{\mathrm{loc}}(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ denote the collection of all the local martingales M defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ with initial value $M_0 = 0$ a.s.. For simplicity we will just denote $\mathscr{M}_{0,c}^{\mathrm{loc}}$ when the underlying filtered probability space is clear. All the continuous local martingales $M \in \mathscr{M}_0^{\mathrm{loc}}$ constitute a vector space, which is denoted by $\mathscr{M}_{0,c}^{\mathrm{loc}}$.

Proposition 5.4.2

1. $\mathscr{M}_{0,c} \subset \mathscr{M}_{0,c}^{\mathrm{loc}}$, and for any $M \in \mathscr{M}_{0,c}$ the sequence $\tau_n = n$ $(n \geq 0)$ reduces M.
2. A nonnegative continuous local martingale M such that $M_0 \in L^1(\Omega, \mathcal{F}_0, \mathrm{P})$ is a supermartingale.
3. A continuous local martingale M such that there exists a random variable $Z \in L^1(\Omega, \mathcal{F}, \mathrm{P})$ with $|M_t| \leq Z$ for every $t \geq 0$ (in particular a bounded continuous local martingale) is a uniformly integrable martingale.
4. For $M \in \mathscr{M}_{0,c}^{\mathrm{loc}}$ and a stopping time τ, we have $M^\tau \in \mathscr{M}_{0,c}^{\mathrm{loc}}$.
5. For $M \in \mathscr{M}_{0,c}^{\mathrm{loc}}$, the sequence $\tau_n = \inf\{t \geq 0 : |M_t| \geq n\}$ $(n \geq 0)$ reduces M.
6. If $(\tau_n)_{n\geq 0}$ reduces M and $(\upsilon_n)_{n\geq 0}$ is a sequence of stopping times such that $\upsilon_n \uparrow \infty$, then the sequence $(\tau_n \wedge \upsilon_n)_{n\geq 0}$ also reduces M.

Proposition 5.4.3 (existence of quadratic variation) Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale. There exists an increasing finite variation process $Q = (Q)_{t\geq 0}$, which is unique up to indistinguishability, such that $(M_t^2 - Q_t)_{t\geq 0}$ is a continuous local martingale. Furthermore, Q is exactly the quadratic variation of M.

Proposition 5.4.4 If $M, N \in \mathscr{M}_{0,c}^{\text{loc}}$, the bracket process of M and N is well defined as

$$\langle M, N\rangle_t := \frac{1}{4}\left(\langle M+N\rangle_t - \langle M-N\rangle_t\right).$$

Furthermore, for all $t \geq 0$,

$$\langle M, N\rangle_t = \operatorname*{plim}_{\left\|P_{[0,t]}\right\|\to 0} \sum_{k=1}^{n} \left(M_{t_k} - M_{t_{k-1}}\right)\left(N_{t_k} - N_{t_{k-1}}\right),$$

where $P_{[0,t]} = \{(t_0, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = t\}$ ranges over partitions of the interval $[0,t]$ and

$$\left\|P_{[0,t]}\right\| = \max_{1\leq k\leq n} |t_k - t_{k-1}|.$$

5.5 Continuous Semimartingales

Definition 5.5.1 (continuous semimartingale) A process $X = (X_t)_{t\geq 0}$ is a continuous semimartingale if it can be written in the form

$$X_t = M_t + A_t$$

where M is a continuous local martingale and A is a continuous finite variation process.

The decomposition $X = M + A$ is unique up to indistinguishability.

Definition 5.5.2 (bracket process) Let $X = M + A$ and $Y = M' + A'$ be the canonical decompositions of two continuous semimartingales X and Y. The bracket $\langle X, Y\rangle$ is the finite variation process defined by

$$\langle X, Y\rangle_t = \langle M, M'\rangle_t$$

In particular, we have $\langle X\rangle_t = \langle M\rangle_t$.

Proposition 5.5.1 Assume X and Y are two continuous semimartingales. For all $t \geq 0$,

$$\sum_{k=1}^{n} \left(M_{t_k} - M_{t_{k-1}}\right)\left(N_{t_k} - N_{t_{k-1}}\right) \xrightarrow{p} \langle M, N\rangle_t, \quad \left\|P_{[0,t]}\right\| \longrightarrow 0,$$

where $P_{[0,t]} = \{(t_0, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = t\}$ ranges over partitions of the interval $[0,t]$ and

$$\left\|P_{[0,t]}\right\| = \max_{1\leq k\leq n} |t_k - t_{k-1}|.$$

Chapter 6

Stochastic Integration

If not specified explicitly, the stochastic processes and random variables are always assumed to be defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$.

6.1 Stochastic Integrals for L^2-Bounded Martingales

Definition 6.1.1 (progressively measurable) Let $\Phi = (\Phi_t)_{t\geq 0}$ be a stochastic process defined on filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$. If for all $T \geq 0$, the mapping

$$\begin{aligned} \Phi^{(T)} : \Omega \times [0,T] &\longrightarrow S \\ (\omega, t) &\longmapsto \Phi_t(\omega) \end{aligned}$$

is $\mathcal{F}_T \otimes \mathcal{B}([0,T])$ -measurable, we say Φ is *progressively measurable*.

Proposition 6.1.1 (Hilbert space $\mathscr{L}^2(M)$) Suppose $M \in \mathscr{M}^2_{0,c}$. Define

$$\begin{aligned} \mathrm{P}_M : \mathcal{F} \otimes \mathcal{B}([0,\infty)) &\longrightarrow S \\ A &\longmapsto \mathrm{E}\left[\int_0^\infty \mathbf{1}_A(\omega, s) d\langle M\rangle_s\right] = \int_\Omega \left[\int_0^\infty \mathbf{1}_A(\omega, s) d\langle M\rangle_s(\omega)\right] d\mathrm{P} \end{aligned}$$

Then $(\Omega \times [0,\infty), \mathcal{F} \otimes \mathcal{B}([0,\infty)), \mathrm{P}_M)$ is a measure space. Let

$$\mathscr{L}^2(M) = \{\Phi \in L^2\left(\Omega \times [0,\infty), \mathcal{F} \otimes \mathcal{B}([0,\infty)), \mathrm{P}_M\right) : \Phi \text{ is progressively measurable}\}.$$

$\mathscr{L}^2(M)$ is a closed subspace of $L^2\left(\Omega \times [0,\infty), \mathcal{F} \otimes \mathcal{B}([0,\infty)), \mathrm{P}_M\right)$ and also a Hilbert space, with the inner product written as

$$(\Phi, \Psi)_{\mathscr{L}^2(M)} = \mathrm{E}\left[\int_0^\infty \Phi_s \Psi_s d\langle M\rangle_s\right].$$

The associated norm is

$$\|\Phi\|_{\mathscr{L}^2(M)} = \left(\mathrm{E}\left[\int_0^\infty \Phi_s^2 d\langle M\rangle_s\right]\right)^{\frac{1}{2}}.$$

Therefore, $\mathscr{L}^2(M)$ consists of all the progressive processes Φ such that

$$\mathrm{E}\left[\int_0^\infty \Phi_s^2 d\langle M\rangle_s\right] < \infty$$

with the identifications for all processes that only differ on P_M-null sets.

Definition 6.1.2 (elementary process) An elementary process is a progressive process of the form

$$\Phi_t(\omega) = \sum_{i=0}^{p-1} \Phi_{(i)}(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

where $0 = t_0 < t_1 < t_2 < \cdots < t_p$ and for every $i \in \{0, 1, \cdots, p-1\}$, $\Phi_{(i)}$ is a bounded $\mathscr{F}_{t_i}$ -measurable random variable.

The set $\mathscr{E}$ of all elementary processes forms a linear subspace of $L^2(M)$. To be precise, we should here say "equivalence classes of elementary processes" "(recall that Φ and Φ' are identified in $\mathscr{L}^2(M)$ if $\|\Phi - \Phi'\|_{\mathscr{L}^2(M)} = 0$).

Proposition 6.1.2 For every $M \in \mathscr{M}_{0,c}^2, \mathscr{E}$ is dense in $\mathscr{L}^2(M)$.

Theorem 6.1.1 Let $M \in \mathscr{M}_{0,c}^2$. For every $\Phi \in \mathscr{E}$ of the form

$$\Phi_t(\omega) = \sum_{i=0}^{p-1} \Phi_{(i)}(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

the formula

$$(\Phi \cdot M)_t = \int_0^t \Phi_s dM_s := \sum_{i=0}^{p-1} \Phi_{(i)} \left(M_{t_{i+1} \wedge t} - M_{t_i \wedge t} \right)$$

defines a process $\Phi \cdot M \in \mathscr{M}_{0,c}^2$. The mapping $I_M^* : \mathscr{E} \to \mathscr{M}_{0,c}^2,\ \Phi \mapsto \Phi \cdot M$ can extend to a linear isometry

$$\begin{aligned} I_M : \mathscr{L}^2(M) &\longrightarrow \mathscr{M}_{0,c}^2 \\ \Phi &\longmapsto \Phi \cdot M, \end{aligned}$$

which means

$$\|\Phi \cdot M\|_{\mathscr{M}_c^2} = \left(\mathrm{E}\left[\left(\int_0^\infty \Phi_s dM_s \right)^2 \right] \right)^{\frac{1}{2}} = \|\Phi\|_{\mathscr{L}^2(M)} = \left(\mathrm{E}\left[\int_0^\infty \Phi_s^2 d\langle M \rangle_s \right] \right)^{\frac{1}{2}}.$$

Furthermore, $\Phi \cdot M$ is the unique martingale in $\mathscr{M}_{0,c}^2$ that satisfies the property

$$\begin{gathered} \langle \Phi \cdot M, N \rangle = \Phi \cdot \langle M, N \rangle, \quad \forall N \in \mathscr{M}_{0,c}^2, \\ \left\langle \int_0^\cdot \Phi_s dM_s, N \right\rangle_t = \int_0^t \Phi_s d \left\langle M, N \right\rangle_s, \quad \forall N \in \mathscr{M}_{0,c}^2,\ t \in [0, \infty). \end{gathered}$$

We call $\Phi \cdot M$ the stochastic integral of Φ with respect to M.

Proposition 6.1.3 Assume that $M, N \in \mathscr{M}_{0,c}^2$, $\Phi \in \mathscr{L}^2(M)$, $\Psi \in \mathscr{L}^2(N)$. Then

$$\left\langle \int_0^\cdot \Phi_s \mathrm{d}M_s, \int_0^\cdot \Psi_s \mathrm{d}N_s \right\rangle_t = \int_0^t \Phi_s \Psi_s \mathrm{d}\langle M, N \rangle_s, \quad \forall t \in [0, \infty].$$

Proposition 6.1.4 If τ is a stopping time with respect to $(\mathcal{F}_t)_{t \geq 0}$, we have

$$\begin{gathered} \left(\mathbf{1}_{[0,\tau]} \Phi \right) \cdot M = (\Phi \cdot M)^\tau = \Phi \cdot M^\tau, \\ \int_0^t \mathbf{1}_{[0,\tau]}(s) \Phi_s dM_s = \int_0^{\tau \wedge t} \Phi_s dM_s = \int_0^t \Phi_s dM_s^\tau, \quad \forall t \in [0, \infty]. \end{gathered}$$

6.2 Stochastic Integrals for Continuous Local Martingales

We will now use extend the definition of $\Phi \cdot M$ to an arbitrary continuous local martingale. If $M \in \mathscr{M}_{0,c}^{\mathrm{loc}}$, we write $\mathscr{L}_{\mathrm{loc}}^2(M)$ for the set of all progressive processes Φ such that for all $t \geq 0$,

$$\int_0^t \Phi_s^2 \mathrm{d}\langle M\rangle_s < \infty.$$

For future reference, we note that $\mathscr{L}_{\mathrm{loc}}^2(M)$ can again be viewed as an "ordinary" L^2-space and thus has a Hilbert space structure. Clearly we see $\mathscr{L}^2(M) \subset \mathscr{L}_{\mathrm{loc}}^2(M)$ for $M \in \mathscr{M}_{0,c}^2$.

Theorem 6.2.1 (stochastic integrals for continuous local martingales) Let $M \in \mathscr{M}_{0,c}^{\mathrm{loc}}$. For every $\Phi \in \mathscr{L}_{\mathrm{loc}}^2(M)$ there exists a unique continuous local martingale in $\mathscr{M}_{0,c}^{\mathrm{loc}}$, which is denoted by $\Phi \cdot M$ or $\int_0^{\cdot} \Phi_s dM_s$, such that for every $N \in \mathscr{M}_{0,c}^{\mathrm{loc}}$,

$$\langle \Phi \cdot M, N\rangle = \Phi \cdot \langle M, N\rangle.$$

If $M' \in \mathscr{M}_c^{\mathrm{loc}}$, then $M' - M_0 \in \mathscr{L}_{\mathrm{loc}}^2(M)$ and we can define

$$\int_0^{\cdot} \Phi_s dM_s' := \Phi \cdot (M' - M_0).$$

If $\Phi \in \mathscr{L}_{\mathrm{loc}}^2\,(M)$ and Ψ is a progressive process, we have $\Psi \in \mathscr{L}_{\mathrm{loc}}^2\,(\Phi \cdot M)$ if and only if $\Phi\Psi \in \mathscr{L}_{\mathrm{loc}}^2(M)$, and then

$$\Phi \cdot (\Psi \cdot M) = (\Phi\Psi) \cdot M.$$

Finally, if $M \in \mathscr{M}_{0,c}^2$, and $\Phi \in \mathscr{L}^2(M)$, the definition of $\Phi \cdot M$ is consistent with that of Theorem 5.4.

Proposition 6.2.1 If τ is a stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$, we have

$$\left(\mathbf{1}_{[0,\tau]}\Phi\right) \cdot M = (\Phi \cdot M)^\tau = \Phi \cdot M^\tau.$$

6.3 Stochastic Integrals for Continuous Semimartingales

We finally extend the definition of stochastic integrals to continuous semimartingales.

Definition 6.3.1 (locally bounded) We say that a progressive process Φ is locally bounded if

$$\forall t \geq 0, \quad \sup_{s \leq t} |\Phi_s| < \infty \quad \text{ a.s.}$$

or equivalently there exist a sequence of stopping times $(\tau_n)_{n\geq 0}$ and a sequence of constants $(C_n)_{n\geq 0}$ such that

$$\forall n \geq 0, \quad \forall t \geq 0, \quad |\Phi_t^{\tau_n}| \leq C_n \quad \text{ a.s.}$$

In particular, any adapted process with continuous sample paths is a locally bounded progressive process. If Φ is (progressive and) locally bounded, then for every finite variation process V, we have

$$\forall t \geq 0, \quad \int_0^t |\Phi_s|\,|\mathrm{d}V_s| < \infty, \quad \text{ a.s.}$$

and similarly $\Phi \in \mathscr{L}_{1\mathrm{oc}}^2(M)$ for every continuous local martingale M.

Definition 6.3.2 (stochastic integrals for continuous semimartingales) Let X be a continuous semimartingale and let $X = M + V$ be its canonical decomposition. If Φ is a locally bounded progressive process, the stochastic integral $\Phi \cdot X$ is the continuous semimartingale with canonical decomposition

$$\Phi \cdot X = \Phi \cdot M + \Phi \cdot V$$

and we write

$$(\Phi \cdot X)_t = \int_0^t \Phi_s \mathrm{d}X_s$$

Proposition 6.3.1

1. The mapping $(\Phi, X) \mapsto \Phi \cdot X$ is bilinear.
2. $\Phi \cdot (\Psi \cdot X) = (\Phi\Psi) \cdot X$, if Φ and Ψ are progressive and locally bounded.
3. For every stopping time τ, $(\Phi \cdot X)^\tau = \Phi \mathbf{1}_{[0,\tau]} \cdot X = \Phi \cdot X^\tau$.
4. If X is a continuous local martingale, resp. if X is a finite variation process, then the same holds for $\Phi \cdot X$.
5. If H is of the form
$$\Phi_s(\omega) = \sum_{i=0}^{p-1} \Phi_{(i)}(\omega) \mathbf{1}_{(t_i, t_i+1)}(s),$$
where $0 = t_0 < t_1 < \cdots < t_p$, and, for every $i \in \{0, 1, \cdots, p-1\}$, $H_{(i)}$ is $\mathcal{F}_{t_i}$-measurable, then
$$\int_0^t \Phi_s \mathrm{d}X_s = \sum_{i=0}^{p-1} \Phi_{(i)} \left(X_{t_{i+1} \wedge t} - X_{t_i \wedge t} \right).$$

Proposition 6.3.2 Assume that Φ is a process with continuous sample paths and that X is continuous semimartingales. Then, for every $t > 0$,

$$\sum_{k=1}^{n} \Phi_{t_{k-1}} \left(X_{t_k} - X_{t_{k-1}} \right) \xrightarrow{p} \int_0^t \Phi_s \mathrm{d}X_s, \quad \left\| P_{[0,t]} \right\| \longrightarrow 0,$$

where $P_{[0,t]} = \{(t_0, \cdots, t_n) : 0 = t_0 < t_1 < \cdots < t_n = t\}$ ranges over partitions of the interval $[0,t]$ and

$$\left\| P_{[0,t]} \right\| = \max_{1 \le k \le n} |t_k - t_{k-1}|.$$

6.4 Itô's Formula

Theorem 6.4.1 (Itô's formula) Let $X^1, \cdots, X^p$ be p continuous semimartingales, and let $F \in C^2(\mathbb{R}^p)$ be a twice continuously differentiable real function. Then, for every $t \ge 0$

$$\begin{aligned} F\left(X_t^1, \cdots, X_t^p\right) = F\left(X_0^1, \cdots, X_0^p\right) &+ \sum_{i=1}^{p} \int_0^t \frac{\partial F}{\partial x^i} \left(X_s^1, \cdots, X_s^p\right) dX_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^{p} \int_0^t \frac{\partial^2 F}{\partial x^i \partial x^j} \left(X_s^1, \cdots, X_s^p\right) d\left\langle X^i, X^j \right\rangle_s. \end{aligned}$$

Corollary 6.4.1 (formula of integration by parts) Take $p = 2$ and $F(x,y) = xy$ in the theorem. If X and Y are two continuous semimartingales, we have

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s \mathrm{d}Y_s + \int_0^t Y_s \mathrm{d}X_s + \langle X, Y \rangle_t.$$

In particular, if $Y = X$

$$X_t^2 = X_0^2 + 2 \int_0^t X_s \mathrm{d}X_s + \langle X \rangle_t.$$

Corollary 6.4.2 1. Let $F \in C^2(\mathbb{R})$.

$$F\left(B_t\right) = F\left(B_0\right) + \int_0^t F'\left(B_s\right) dB_s + \frac{1}{2}\int_0^t F''\left(B_s\right) ds.$$

2. Let $F(t,x) \in C^2(\mathbb{R}^2)$.

$$F\left(t, B_t\right) = F\left(0, B_0\right) + \int_0^t \frac{\partial F}{\partial x}\left(s, B_s\right) dB_s + \int_0^t \left(\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\right)\left(s, B_s\right) ds$$

Appendix

1.Properties of Common Distributions

Distribution	pmf $\mathrm{P}(X=k)$	Support	Mean	Variance
Bernoulli $B(1,p)$	$p^k(1-p)^{1-k}$	$\{0,1\}$	p	$p(1-p)$
Binomial $B(n,p)$	$\binom{n}{k}p^k(1-p)^{n-k}$	$\{0,\cdots,n\}$	np	$np(1-p)$
Negative Binomial $NB(r,p)$	$\binom{k+r-1}{k}(1-p)^r p^k$	$\mathbb{N}$	$\dfrac{pr}{1-p}$	$\dfrac{pr}{(1-p)^2}$
Poisson $Pois(\lambda)$	$\dfrac{\lambda^k e^{-\lambda}}{k!}$	$\mathbb{N}$	λ	λ
Geometric $Geo(p)$	$(1-p)^{k-1}p$	$\mathbb{N}_+$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Hypergeometric $H(N,K,n)$	$\dfrac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$	$\{0,\cdots,\min(n,K)\}$	$n\dfrac{K}{N}$	$n\dfrac{K}{N}\dfrac{N-K}{N}\dfrac{N-n}{N-1}$
Uniform (discrete) $DU(a,b)$	$\dfrac{1}{n}$	$\{a,a+1,\ldots,b\}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a+1)^2-1}{12}$

Distribution	pdf	Mean	Variance
Degenerate δ_a	$I_{\{a\}}(x)$	a	0
Uniform (continuous) $U(a,b)$	$\frac{1}{b-a} I_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential $Exp(\lambda) = \Gamma(1,\lambda)$	$\lambda e^{-\lambda x} I_{[0,+\infty)}(x)$	λ^{-1}	λ^{-2}
Normal $N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Log-normal $LogN(\mu,\sigma^2)$	$\frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$	$e^{\gamma+\frac{\sigma^2}{2}}$	$e^{2\left(\gamma+\sigma^2\right)} - e^{2\gamma+\sigma^2}$
Gamma $\Gamma(\alpha,\beta)$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} I_{(0,+\infty)}(x)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Beta $B(\alpha,\beta)$	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)} I_{(0,1)}(x)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Chi-squared $\chi^2_k = \Gamma(\frac{k}{2},\frac{1}{2})$	$\frac{1}{2^{k/2}\Gamma(k/2)}\ x^{k/2-1} e^{-x/2} I_{(0,+\infty)}(x)$	k	$2k$
Student's t t_ν	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	0	$\frac{\nu}{\nu-2}$ for $\nu > 2$

2.Generating Function & Characteristic Function

Distribution	Moment-generating function	Characteristic function
Degenerate δ_a	e^{ta}	e^{ita}
Bernoulli $B(1,p)$	$1-p+pe^t$	$1-p+pe^{it}$
Geometric $Geo(p)$	$\frac{pe^t}{1-(1-p)e^t}$, $t<-\ln(1-p)$	$\frac{pe^{it}}{1-(1-p)e^{it}}$
Binomial $B(n,p)$	$(1-p+pe^t)^n$	$(1-p+pe^{it})^n$
Negative Binomial $NB(r,p)$	$\frac{(1-p)^r}{(1-pe^t)^r}$	$\frac{(1-p)^r}{(1-pe^{it})^r}$
Poisson $Pois(\lambda)$	$e^{\lambda(e^t-1)}$	$e^{\lambda(e^{it}-1)}$
Uniform (continuous) $U(a,b)$	$\begin{cases} \frac{\mathrm{e}^{tb}-\mathrm{e}^{ta}}{t(b-a)} & \text{for } t\neq 0 \\ 1 & \text{for } t=0 \end{cases}$	$\begin{cases} \frac{\mathrm{e}^{itb}-\mathrm{e}^{ita}}{it(b-a)} & \text{for } t\neq 0 \\ 1 & \text{for } t=0 \end{cases}$
Uniform (discrete) $DU(a,b)$	$\frac{e^{at}-e^{(b+1)t}}{(b-a+1)\,(1-e^t)}$	$\frac{e^{it\mu}}{(b-a+1)\,(1-e^{it})}$
Laplace $L(\mu,b)$	$\frac{e^{t\mu}}{1-b^2t^2}$, $\lvert t\rvert<1/b$	$\frac{e^{it\mu}}{1+b^2t^2}$
Normal $N(\mu,\sigma^2)$	$e^{t\mu+\frac{1}{2}\sigma^2t^2}$	$e^{it\mu-\frac{1}{2}\sigma^2t^2}$
Chi-squared χ^2_k	$(1-2t)^{-\frac{k}{2}}$	$(1-2it)^{-\frac{k}{2}}$
Noncentral chi-squared $\chi^2_k(\lambda)$	$e^{\lambda t/(1-2t)}(1-2t)^{-\frac{k}{2}}$	$e^{i\lambda t/(1-2it)}(1-2it)^{-\frac{k}{2}}$
Gamma $\Gamma(\alpha,\beta)$	$\left(1-\frac{t}{\beta}\right)^{-\alpha}$, $t<\beta$	$\left(1-\frac{it}{\beta}\right)^{-\alpha}$
Beta $B(\alpha,\beta)$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$	${}_1F_1(\alpha;\alpha+\beta;it)$
Exponential $Exp(\lambda)$	$\frac{\lambda}{\lambda-t}, t<\lambda$	$\frac{\lambda}{\lambda-it}$
Multivariate normal $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$	$e^{\mathbf{t}^T\left(\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{\Sigma}\mathbf{t}\right)}$	$e^{\mathbf{t}^T\left(i\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{\Sigma}\mathbf{t}\right)}$
Cauchy $Cauchy(\mu,\theta)$	Does not exist	$e^{it\mu-\theta\lvert t\rvert}$
Multivariate Cauchy $MultiCauchy(\boldsymbol{\mu},\boldsymbol{\Sigma})$	Does not exist	$e^{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}-\sqrt{\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}}}$

Bibliography

[1] Gall J F L. Brownian Motion, Martingales, and Stochastic Calculus[J]. 2018.

[2] Morimoto H. Stochastic control and mathematical modeling[M]. Cambridge University Press, 2010.

[3] Athreya K B, Lahiri S N. Measure theory and probability theory[M]. Springer Science & Business Media, 2006.

[4] Pivato M. Stoshastic processes and stochastic integration[J]. 1999.