# Stochastic Process

## 1 Preliminaries

**Definition 1.1 (stochastic process)** For a given probability space $(\Omega, \mathcal{F}, P)$ and a measurable space $(S, \mathcal{E})$, a *stochastic process* is a collection of $S$-valued random variables on $(\Omega, \mathcal{F}, P)$ indexed by some set $T$, which can be written as $X = \{X(t) : t \in T\}$ or $X = (X_t)_{t \in T}$ or $X : \Omega \times T \to S$. This mathematical space $S$ is called its state space.

For convenience, we always assume $T$ is a totally ordered set and denote the collection of all finite subsets of $T$ by $\mathcal{I}_T$, namely

$$\mathcal{I}_T = \{\{t_1, t_2, \cdots, t_n\} : t_1, \cdots, t_n \in T,\ n \geq 1\}.$$

**Definition 1.2 (independent increments)** A stochastic process $(X_t)_{t \in T}$ has *independent increments* if for every $n \in \mathbb{N}_+$ and any $t_1 \leq t_2 \leq \cdots \leq t_n$, 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 1.3 (independent increments)** A stochastic process $(X_t)_{t \in T}$ has *independent increments* if for every $n \in \mathbb{N}_+$ and any $t_1 \leq t_2 \leq \cdots \leq t_n$, 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 1.4 (strictly stationary process)** Let $(X_t)_{t \in T}$ be a stochastic process and let $F_X(x_{t_1+\tau}, \ldots, x_{t_k+\tau})$ represent the distribution function of the joint distribution of $(X_t)_{t \in T}$ at times $t_1 + \tau, \ldots, t_k + \tau$ . Then, $(X_t)_{t \in T}$ is said to be strictly stationary if, for all $k$, for all $\tau$, and for all $t_1, \ldots, t_k$,

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

## 2 Poisson Process

**Definition 2.1 (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 2.2 (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 $0 \le s < t$, the counting process $N_t - N_s$ represents the number of events that occurred on $(s, t]$.

**Definition 2.3 (Poisson process (II))** A counting process $(N_t)_{t\ge 0}$ 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) $(N_t)_{t\ge 0}$ has independent increments;

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

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

**Definition 2.4 (Poisson process (III))** A stochastic process $(N_t)_{t\ge 0}$ 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** Definition 2.1, 2.3, 2.4 are equivalent definitions of Poisson process.

*Proof.*

- <u>Definition 2.1 $\Longrightarrow$ Definition 2.3</u>

  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

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

- <u>Definition 2.3 $\Longrightarrow$ Definition 2.1</u>

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

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

  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).

- <u>Definition 2.1 $\implies$ Definition 2.4</u>
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 $\sim Exp(\lambda)$,

- <u>Definition 2.4 $\Longrightarrow$ Definition 2.1</u>

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 $(T_1, T_2, \cdots, T_m)$

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

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

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

Note

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

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 $0 \le s < t$, $N_t - N_s \sim Pois(\lambda(t-s))$.

$\square$

**Proposition 2.2** Let $(N_t)_{t\geq 0}$ be a Poisson process.

1. $N_t \sim Pois(\lambda t)$, $\mathrm{E}[N_t] = \mathrm{Var}(N_t) = \lambda t$.
2. For $0 \leq s \leq t$, $\mathrm{E}[N_t N_s] = \lambda^2 ts + \lambda s$, $\mathrm{Cov}(E_t, E_s) = \lambda s$.
3. For $0 \leq s \leq t$, $\mathrm{E}[N_t|N_s] = N_s + \lambda(t-s)$. So Poisson process is a submartingale.
4. Poisson process is a Markov process. For $0 \leq t_1 < t_2 < \cdots < t_n$ and $0 \leq k_1 \leq k_2 \leq \cdots \leq k_n$,

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

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

# 3 Compound Poisson Process

**Definition 3.1 (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 3.2 (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 $\lambda$, 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 3.1** Let $(Y_t)_{t\geq 0}$ be a compound Poisson process with a rate $\lambda$ and jump $(Z_n)_{n\in\mathbb{N}_+}$. For convenience, assume $Z_n \overset{d}{=} Z$ and $\mathrm{E}[Z^2] < +\infty$.

1. $\mathrm{E}[Y_t] = \lambda t E[Z]$.
2. $\mathrm{Var}(Y_t) = \lambda t E[Z^2]$.
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

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

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

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

# 4 Markov Chain

## 4.1 Discrete-time Markov Chain

**Definition 4.1 (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 $\forall n\geq 0,\ \forall j, i_0, i_1, \cdots, i_n\in S$,

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

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

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

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

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

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,

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

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

**Proposition 4.1 (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.*

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

□

Of course $P^{(1)} = P$. Thus by iteration we show $P^{(n)}$ coincides with $P^n$.

## 4.2 Continuous-time Markov Chain

**Definition 4.2 (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 $j, i_0, \cdots, i_n \in S$,

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

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

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

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

**Example 4.1** 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 4.2** 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.

□

**Definition 4.3** 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 $p_{ij}(0) = \mathrm{P}(X_t = j \mid X_0 = i) = \delta_{ij}$, 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 4.1** 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

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

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 $[0, \infty)$.

□

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

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

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 4.1** 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

$$P'(t) = Q(t)P(t).$$

# 5 Martingale

**Definition 5.1 (conditional expectation)** Let $X$ be a $\mathcal{F}$-measurable random variable on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ such that $\mathrm{E}[|X|] < \infty$. Given a $\sigma$-algebra $\mathcal{G} \subset \mathcal{F}$, a random variable $Z$ that is $\mathcal{G}$-measurable and satisfies

$$\mathrm{E}[XI_A] = \mathrm{E}[ZI_A] \quad \text{for all } A \in \mathcal{G}$$

is called the *conditional expectation* of $Y$ given $\mathcal{G}$ and is written as $\mathrm{E}(X \mid \mathcal{G})$.

In probability theory, we show that conditional expectation exists and is unique up to absolutely surely equality. If not pointed out explicitly, all equalities and inequalities involving conditional expectation are considered to hold absolutely surely.

**Proposition 5.1** Let $X$, $Y$, $X_n$ be integrable $\mathcal{F}$-measurable random variables on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and $\mathcal{G} \subset \mathcal{F}$ be a $\sigma$-algebra.

1. If $a$, $b$ are constants, then $\mathrm{E}[aX + bY \mid \mathcal{F}] = a\mathrm{E}[X \mid \mathcal{F}] + b\mathrm{E}[Y \mid \mathcal{F}]$.
2. If $X$ equals a constant $a$, then $\mathrm{E}[X \mid \mathcal{F}] = a$.
3. If $X \geq Y$, then $\mathrm{E}[X \mid \mathcal{F}] \geq \mathrm{E}[Y \mid \mathcal{F}]$.
4. $|\mathrm{E}[X \mid \mathcal{F}]| \leq \mathrm{E}[|X| \mid \mathcal{F}]$.
5. If $\phi$ is a convex function on $\mathbb{R}$ and $\phi(X)$ is integrable, then $\phi(\mathrm{E}[X \mid \mathcal{F}]) \leq \mathrm{E}[\phi(X) \mid \mathcal{F}]$.
6. If $\lim\limits_{n\to\infty} X_n = X$ and $|X_n| \leq X$, then $\lim\limits_{n\to\infty} \mathrm{E}[X_n \mid \mathcal{F}] = \mathrm{E}[X \mid \mathcal{F}]$.
7. $\mathrm{E}[\mathrm{E}[X \mid \mathcal{F}]] = \mathrm{E}[X]$.
8. $\mathrm{E}[\mathrm{E}[X \mid \mathcal{G}] \mid \mathcal{F}] = \mathrm{E}[\mathrm{E}[X \mid \mathcal{F}] \mid \mathcal{G}] = \mathrm{E}[X \mid \mathcal{G}]$
9. If $X$ and $\mathcal{F}$ are independent, that is, whenever $A \in \sigma(X)$ and $B \in \mathcal{F}$, $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$, then $\mathrm{E}[X \mid \mathcal{F}] = X$.
10. If $Z$ is $\mathcal{F}$-measurable and $ZX$ is integrable, then $\mathrm{E}[ZX \mid \mathcal{F}] = Z\mathrm{E}[X \mid \mathcal{F}]$

**Definition 5.2 (filtration)** Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and let $I$ be a linearly ordered index set such as $\mathbb{N}$ or $\mathbb{R}_{\geq 0}$. For every $i \in I$ let $\mathcal{F}_i$ be a sub-$\sigma$-algebra of $\mathcal{F}$. Then

$$\mathbb{F} = (\mathcal{F}_i)_{i \in I}$$

is called a *filtration* if $\mathcal{F}_k \subset \mathcal{F}_\ell$ for all $k \leq \ell$.

If $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$ is a filtration, then $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in T}, \mathrm{P})$ is called a *filtered probability space*. 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 $\mathbb{F} = (\mathcal{F}_t)_{t \in T}$, that is, $\sigma(X_t) \subset \mathcal{F}_t$ for all $t \in T$.

**Example 5.1** 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 $\sigma$-algebra and $(\mathcal{F}_t^X)_{t \geq 0}$ is a filtration that $X$ is adapted to. And $(\mathcal{F}_t^X)_{t \geq 0}$ is called the filtration induced by the stochastic process $X$.

**Definition 5.3 (stopping time)** Let $\tau$ 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 $\tau$ is called a *stopping time* (with respect to the filtration $(\mathcal{F}_t)_{t \in T}$), if the following condition holds:

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

**Definition 5.4 (stopping time in discrete-time case)** Let $\tau$ be a random variable defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathrm{P})$ with values in $\mathbb{N} \cup \{+\infty\}$. Then $\tau$ is called a stopping time (with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$), if the following condition holds:

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

**Definition 5.5 (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 $n \geq 0$, $\mathrm{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$.

**Definition 5.6 (submartingale)** A discrete-time *submartingale* is a stochastic process $M = (M_n)_{n\geq 0}$ consisting of integrable random variables satisfying for $n \geq 0$

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

**Definition 5.7 (supermartingale)** A discrete-time *supermartingale* is a stochastic process $M = (M_n)_{n\geq 0}$ consisting of integrable random variables satisfying for $n \geq 0$

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

# 6 Brownian Motion

**Definition 6.1 (Brownian motion)** A stochastic process $(B_t)_{t\geq 0}$ is called a *Brownian motion* if

1. $B_0 = 0$.
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 6.2 (Gaussian process)** A stochastic process $(X_t)_{t\in T}$ is a *Gaussian process* if and only if for every finite set of indices $t_1, \dots, t_n$ in the index set $T$, $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ follows multivariate normal distribution $N(\mu, \Sigma)$.

**Theorem 6.1** $B = (B_t)_{t\geq 0}$ 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 6.1** 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.

**Definition 6.3 (quadratic variation)** Suppose that $(X_t)_{t\geq 0}$ is a real-valued stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. The quadratic variation of $(X_t)_{t\geq 0}$ is the stochastic process, written as $[X]_t$, defined as

$$[X]_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2 \quad \text{a.e.},$$

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

$$\|P\| = \max_{1\leq k\leq n} |t_k - t_{k-1}|.$$

**Theorem 6.2** The quadratic variation of a Brownian motion $B$ exists, and is given by $[B]_t = t$.

*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

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

Since

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

we conclude

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

$\square$

# 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 $Pois(\lambda)$ | $\dfrac{\lambda^k e^{-\lambda}}{k!}$ | $\mathbb{N}$ | $\lambda$ | $\lambda$ |
| 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 $\delta_a$ | $I_{\{a\}}(x)$ | $a$ | $0$ |
| Uniform (continuous) $U(a,b)$ | $\dfrac{1}{b-a}I_{[a,b]}(x)$ | $\dfrac{a+b}{2}$ | $\dfrac{(b-a)^2}{12}$ |
| Exponential $Exp(\lambda)=\Gamma(1,\lambda)$ | $\lambda e^{-\lambda x}I_{[0,+\infty)}(x)$ | $\lambda^{-1}$ | $\lambda^{-2}$ |
| Normal $N(\mu,\sigma^2)$ | $\dfrac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | $\mu$ | $\sigma^2$ |
| Gamma $\Gamma(\alpha,\beta)$ | $\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}I_{(0,+\infty)}(x)$ | $\dfrac{\alpha}{\beta}$ | $\dfrac{\alpha}{\beta^2}$ |
| Chi-squared $\chi^2_k=\Gamma(\frac{k}{2},\frac{1}{2})$ | $\dfrac{1}{2^{k/2}\Gamma(k/2)}\ x^{k/2-1}e^{-x/2}I_{(0,+\infty)}(x)$ | $k$ | $2k$ |
| Student's t $t_\nu$ | $\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\dfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$ | $0$ | $\dfrac{\nu}{\nu-2}$ for $\nu>2$ |

## 2.Generating Function & Characteristic Function

| Distribution | Moment-generating function | Characteristic function |
|---|---|---|
| Degenerate $\delta_a$ | $e^{ta}$ | $e^{ita}$ |
| 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 $Pois(\lambda)$ | $e^{\lambda\left(e^t-1\right)}$ | $e^{\lambda\left(e^{it}-1\right)}$ |
| 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) $DU(a,b)$ | $\frac{e^{at}-e^{(b+1)t}}{(b-a+1)\left(1-e^t\right)}$ | $\frac{e^{it\mu}}{(b-a+1)\left(1-e^{it}\right)}$ |
| 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 $\chi_k^2$ | $(1-2t)^{-\frac{k}{2}}$ | $(1-2it)^{-\frac{k}{2}}$ |
| Noncentral chi-squared $\chi_k^2(\lambda)$ | $e^{\lambda t/(1-2t)}(1-2t)^{-\frac{k}{2}}$ | $e^{i\lambda t/(1-2it)}(1-2it)^{-\frac{k}{2}}$ |
| Gamma $\Gamma(\alpha,\beta)$ | $\left(1-\frac{t}{\beta}\right)^{-\alpha}$, $t<\beta$ | $\left(1-\frac{it}{\beta}\right)^{-\alpha}$ |
| 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 | $e^{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}-\sqrt{\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}}}$ |