Definition 1.1 (Stochastic process)

Family of random variables $\{Y(t): \ t \in \mathcal{T}\}$ is called a stochastic process. It is called discrete if the set $\mathcal{T}$ is finite or countable, and it is called continuous otherwise. Also, the set $\mathcal{S}$ such that $Y(t) \in \mathcal{S}$ is called the state-space of the stochastic process $\{Y(t): \ t \in \mathcal{T}\}$. Note that $Y(t): \Omega \times \mathcal{T} \to \mathcal{S} \subseteq [0, \ \infty)$.

Definition 1.2 (Non-homegenious Markov chain)

Stochastic process $\{Y(t): \ t \in \mathcal{T}\}$ is a Markov chain if (a) it is discrete, (b) has finite number of states, and (c)

$$\begin{aligned} & \mathbb{P}(Y(t+1) = s | \ Y(t) = s_t, Y(t-1) = s_{t-1}, \ldots, Y(0) = s_0) \\ = \ & \mathbb{P}(Y(t+1) = s | \ Y(t) = s_t) \end{aligned}$$

for all $t = 1, 2, 3, \ldots$ and all $s_0, s_1, \ldots, s_t, s \in \mathcal{S}$.

We will use the notation $p_x^{i,j} = \mathbb{P}(Y(x+1) = j|\ Y(x) = i)$ for all $x = 1, 2, 3, \ldots$ and all $i,\ j \in \mathcal{S}$.

Definition 1.3 (Transition probability matrix)

Let $m = \#(\mathcal{S})$, then matrix $P_n \in Mat_{m\times m}([0,\ 1])$ that has $p_{x+n}^{i,j}$ as its i-th row and j-th column entry is called transition probability matrix, and it gathers probabilities that (x) 'jumps' from the i-th state at time n to the j-th state at time $n+1$.

Note

We must have

$$\sum_{j=1}^{m} p_{x+n}^{i,j} = \sum_{j=1}^{m} \mathbb{P}(Y(x+n+1) = j|\ Y(x+n) = i) = 1.$$

This does not have to be so for the sum of columns of the matrix P_n.

Example 1.1 (Simplest alive-dead situation)

Set $\mathcal{S} = \{1,\ 2\}$, where $\{1\}$ denotes "alive" and $\{2\}$ denotes "dead". Then

$$P_n = \begin{pmatrix} p_{x+n} & q_{x+n} \\ 0 & 1 \end{pmatrix},$$

where obviously $(P_n)_{1,1} + (P_n)_{1,2} = p_{x+n} + q_{x+n} = 1$ and also $(P_n)_{2,1} + (P_n)_{2,2} = 0 + 1 = 1$.

Example 1.2 (Multiple-decrement situation)

Set $\mathcal{S} = \{0,\ 1,\ 2, \ldots, m\}$, where $\{0\}$ denotes "alive" and $\{1,\ 2, \ldots, m\}$ denote various reasons of decrement. Then the transition matrix P_n is $(m+1) \times (m+1)$ as following

$$P_n = \begin{pmatrix} p_{x+n}^{(\tau)} & q_{x+n}^{(1)} & q_{x+n}^{(2)} & \cdots & q_{x+n}^{(m)} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

where obviously the sum of row elements is equal to one.

Example 1.3 (Multiple-life situation)

Set $\mathcal{S} = \{0,\ 1,\ 2, 3\}$, where $\{0\}$ denotes "both are alive", $\{1\}$ denotes "(x) is alive and (y) is dead", $\{2\}$ denotes "(x) is dead and (y) is alive", $\{3\}$ denotes "both are dead". Then the transition matrix P_n is as following under the assumption of independence of the future life-time random variables

$$P_n = \begin{pmatrix} p_{x+n}p_{y+n} & p_{x+n}q_{y+n} & q_{x+n}p_{y+n} & q_{x+n}q_{y+n} \\ 0 & p_{x+n} & 0 & q_{x+n} \\ 0 & 0 & p_{y+n} & q_{y+n} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where obviously the sum of row elements is equal to one.

Example 1.4 (Disability situation)

Set $\mathcal{S} = \{0,\ 1,\ 2, 3\}$, where $\{0\}$ denotes "(x) is active", $\{1\}$ denotes "(x) is temporarily disabled", $\{2\}$ denotes "(x) is permanently disabled", $\{3\}$ denotes "(x) is dead". Then the transition matrix P_n is such that $(P_n)_{2,0} = (P_n)_{2,1} = 0$. $(P_n)_{3,3} = 1$ and so $(P_n)_{3,0} = (P_n)_{3,1} = (P_n)_{3,2} = 0$, whereas other probabilities are chosen so that they reflect observations.

ther popular examples are driving ratings and continuing care retirement communities.

Proposition 1.1

Let $Y(x)$ be a non-homogeneous Markov Chain with $\mathcal{S} = \{1, \ldots, m\}$, then

$$ {}_k(P_n)_{i,j} := \mathbb{P}(Y(x+n+k) = j |\, Y(x+n) = i) = (P_n P_{n+1} \cdots P_{n+k-1})_{i,j} $$

for non-negative and integer x, k, n and $i, j \in \mathcal{S}$.

Proof.

We will prove by induction. To start off, let $x = n = 0$ for simplicity of the exposition and without loss of generality. So what is the probability $\mathbb{P}(Y(2) = j \| Y(0) = i)$? □

Proof.

$$
\begin{aligned}
{}_2(P_0)_{i,j} &= \mathbb{P}(Y(2)=j|\ Y(0)=i) \\
&= \sum_{l=1}^{m} \mathbb{P}(Y(2)=j|\ Y(1)=l, Y(0)=i)\mathbb{P}(Y(1)=l|\ Y(0)=i) \\
&= \sum_{l=1}^{m} \mathbb{P}(Y(2)=j|\ Y(1)=l)\mathbb{P}(Y(1)=l|\ Y(0)=i) \\
&= \sum_{l=1}^{m} (P_1)_{l,j}(P_0)_{i,l} = \sum_{l=1}^{m} (P_0)_{i,l}(P_1)_{l,j} \\
&= (P_0 P_1)_{i,j} \stackrel{(?)}{=} (P^2)_{i,j}
\end{aligned}
$$

for all $i, j \in \mathcal{S}$ and the last equality holds if the Markov Chain is homogeneous. □

Proof.

Further assume that
$\mathbb{P}(Y(k-1)=j|\ Y(0)=i)=(P_0P_1\cdots P_{k-2})_{i,j}$ for all $i,j\in\mathcal{S}$, then

$$
\begin{aligned}
{}_k(P_0)_{i,j} &= \mathbb{P}(Y(k)=j|\ Y(0)=i) \\
&= \sum_{l=1}^{m}\mathbb{P}(Y(k)=j|\ Y(k-1)=l,Y(0)=i)\mathbb{P}(Y(k-1)=l|\ Y(0)=i \\
&= \sum_{l=1}^{m}\mathbb{P}(Y(k)=j|\ Y(k-1)=l)(P_0P_1\cdots P_{k-2})_{i,l} \\
&= \sum_{l=1}^{m}(P_{k-1})_{l,j}(P_0P_1\cdots P_{k-2})_{i,l}=\sum_{l=1}^{m}(P_0P_1\cdots P_{k-2})_{i,l}(P_{k-1})_{l,j} \\
&= (P_0P_1\cdots P_{k-1})_{i,j}\overset{(?)}{=}(P^k)_{i,j}
\end{aligned}
$$

for all $i,j\in\mathcal{S}$ and the last equality holds if the Markov Chain is

Corollary 1.1

Let $Y(0),\ Y(1), \ldots$ be a non-homogeneous Markov Chain with $\mathcal{S} = \{1, \ldots, m\}$ and transition probability matrices $P_0,\ P_1, \ldots$, then

$$_k(P_n) = P_n \times P_{n+1} \times \cdots \times P_{n+k-1}$$

for non-negative and integer k, n. Also, if the Markov Chain is homogeneous, then

$$_k(P_n) = (P_n^k)$$

for non-negative and integer k, n.

Denote by π the distribution of the random variable $Y(x)$, and let $\pi_k = \mathbb{P}(Y(x) = k),\ k \in \mathcal{S}$.

Corollary 1.2

We have that the distribution of the random variable $Y(x+k)$ is $\pi'(P_0 P_1 \times \cdots \times P_{k-1})$.

Proof.

We have

$$\mathbb{P}(Y(x+k) = j) = \sum_{i=1}^{m} \mathbb{P}(Y(x+k) = j, Y(x) = i)$$

$$= \sum_{i=1}^{m} \mathbb{P}(Y(x+k) = j|\ Y(x) = i)\mathbb{P}(Y(x) = i)$$

$$= \sum_{i=1}^{m} (P_0 P_1 \times \cdots P_{k-1})_{i,j}\pi_i.$$

Proof.

Last line yields

$$\mathbb{P}(Y(x+k)=j)=\sum_{i=1}^{m}\pi_i(P_0P_1\times\cdots P_{k-1})_{i,j},$$

which establishes the assertion. □

Proposition 1.2

Consider again a non-homogeneous Markov Chain $\{Y(x),\ x\in\mathcal{T}\}$*, then*

$$\begin{aligned}
& \mathbb{P}(Y(x+n+1)=Y(x+n+2)=\cdots=Y(x+n+k)=i|\ Y(x+ \\
= & (P_n)_{i,i}(P_{n+1})_{i,i}\times\cdots\times(P_{n+k-1})_{i,i},
\end{aligned}$$

for non-negative and integer n,k *and* $i\in\mathbb{S}$*.*

Proof.

As we have, by conditioning and evoking the Markovian property,

$$
\begin{aligned}
& \mathbb{P}(Y(x+n+k)=i,\ Y(x+n+k-1)=i,\ \ldots,\ Y(x+n+ \\
= \quad & \mathbb{P}(Y(x+n+k)=i|\ Y(x+n+k-1)=i,\ \ldots,\ Y(x+n+ \\
\times \quad & \mathbb{P}(Y(x+n+k-1)=i|\ Y(x+n+k-2)=i,\ \ldots,\ Y(x+ \\
\times\cdots\times \quad & \mathbb{P}(Y(x+n)=i) \\
= \quad & \mathbb{P}(Y(x+n+k)=i|\ Y(x+n+k-1)=i) \\
\times \quad & \mathbb{P}(Y(x+n+k-1)=i|\ Y(x+n+k-2)=i) \\
\times\cdots\times \quad & \mathbb{P}(Y(x+n+1)=i|\ Y(x+n)=i)\times\mathbb{P}(Y(x+n)=i),
\end{aligned}
$$

then the conditional probability in the assertion follows. □

From now and on consider a continuous stochastic process $\{Y_x(t)\}_{t\geq 0}$ with the state space $\mathcal{S} = \{0,\ 1,\ \dots, m\},\ m \in \mathbb{N}$ and $\mathcal{T} = [0,\ \infty)$. We assume that:

(a) For any $s \geq 0$ and $i, j \in \mathcal{S}$, the conditional probability

$$\mathbb{P}(Y_x(t+s) = j|\ Y_x(t) = i)$$

is independent on the history of the process for all times before $t \in [0,\ \infty)$.

(b) For any time length $h > 0$,

$$\mathbb{P}(\text{2 or more transitions occure within h}) = o(h),$$

where we say that the function $f(h)$ is $o(h)$ if $lim_{h\to 0} f(h)/h = 0$.

We will use the notation

$$ {}_t p_x^{i,j} = \mathbb{P}(Y_x(t) = j |\ Y_x = i),\ i, j \in \mathcal{S},\ t \geq 0 $$

and

$$ {}_t p_x^{\overline{i,i}} = \mathbb{P}(Y_x(s) = j \forall s \in [0,\ t] |\ Y_x = i),\ i \in \mathcal{S},\ t \geq 0 $$

(c) The function $t \mapsto {}_t p_x^{i,j}$ is differentiable for all $t \in (0,\ \infty)$.

Note that now we can define the force of transition as following

$$ \mu_x^{i,j} = \lim_{h \downarrow 0} \frac{{}_h p_x^{i,j}}{h},\ \text{ for } i \neq j \in \mathcal{S}. $$

Note that we can say equivalently that

$$ {}_hp_x^{i,j} = h \times \mu_x^{i,j} + o(h), \;\; \text{for } i \neq j \in \mathcal{S} $$

or in other words

$$ {}_hp_x^{i,j} \approx h \times \mu_x^{i,j}, \;\; \text{for } i \neq j \in \mathcal{S}. $$

The latter expression should remind you the simple alive-dead framework.

Proposition 1.3

For a general multiple-state model, we have for $h > 0,\ x \geq 0$ and $i \in \mathcal{S}$,

$$ {}_hp_x^{i,i} = {}_hp_x^{\overline{i,i}} + o(h) $$

Proof.

The right hand side is obtained by the law of total probability:

$$
\begin{aligned}
{}_h p_x^{i,i} &= \mathbb{P}\left(Y_x(h) = i \mid Y_x = i\right) \\
&= \mathbb{P}\left(Y_x(h) = i \mid Y_x = i, \exists t \in [0,\ h) : Y_x(t) \neq i\right) \\
&\times \mathbb{P}\left(\exists t \in [0,\ h) : Y_x(t) \neq i \mid Y_x = i\right) \\
&+ \mathbb{P}\left(Y_x(h) = i \mid Y_x = i, \forall t \in [0,\ h) : Y_x(t) = i\right) \\
&\times \mathbb{P}\left(\forall t \in [0,\ h) : Y_x(t) = i \mid Y_x = i\right) \\
&= \mathbb{P}\left(Y_x(h) = i, Y_x = i, \exists t \in [0,\ h) : Y_x(t) \neq i\right) / \mathbb{P}(Y_x = i) \\
&+ \mathbb{P}\left(Y_x(h) = i, Y_x = i, \forall t \in [0,\ h) : Y_x(t) = i\right) / \mathbb{P}(Y_x = i) \\
&= \mathbb{P}\left(Y_x(h) = i, \exists t \in [0,\ h) : Y_x(t) \neq i \mid Y_x = i\right) \\
&+ \mathbb{P}\left(Y_x(h) = i, \forall t \in [0,\ h) : Y_x(t) = i \mid Y_x = i\right) \\
&= o(h) + {}_h p_x^{\overline{i,i}},
\end{aligned}
$$

which completes the proof. □

Proposition 1.4

For any multiple-state model and for $h > 0$, $x \geq 0$ and $i, j \in \mathcal{S}$, we have

$$ {}_hp_x^{\overline{i,i}} = 1 - h \sum_{j=0, j\neq i}^{m} \mu_x^{i,j} + o(h) $$

Proof.

We have

$$ 1 = {}_hp_x^{\overline{i,i}} + \sum_{j=0, j\neq i}^{m} {}_hp_x^{i,j} + o(h), $$

or

$$ 1 = {}_hp_x^{\overline{i,i}} + h \sum_{j=0, j\neq i}^{m} \mu_x^{i,j} + m \times o(h), $$

which completes the proof since $m \times o(h) = o(h)$. □

Proposition 1.5

For any multiple-state model, we have for $h \geq 0$ and $i \in \mathcal{S}$,

$$ {}_h p_x^{\overline{i,i}} = \exp\left\{ -\int_0^h \sum_{j=0, j\neq i}^{m} \mu_x^{i,j}(s)ds \right\} $$

Proof.

Start with the observation

$$ {}_{h+\Delta h} p_x^{\overline{i,i}} = {}_h p_x^{\overline{i,i}} \times {}_{\Delta h} p_{x+h}^{\overline{i,i}}, $$

which is true for all $h \geq 0$ because $\{Y_x\}_{h\geq 0}$ is a Markov process. □

Proof.

Further evoking Proposition 1.4, we obtain

$$ {}_{h+\Delta h}p_x^{\overline{i,i}} = {}_h p_x^{\overline{i,i}} \times \left(1 - \Delta h \sum_{j=0, j\neq i}^{m} \mu_x^{i,j}(h) + o(\Delta h) \right) $$

or

$$ {}_{h+\Delta h}p_x^{\overline{i,i}} - {}_h p_x^{\overline{i,i}} = -{}_h p_x^{\overline{i,i}} \times \Delta h \sum_{j=0, j\neq i}^{m} \mu_x^{i,j}(h) + o(\Delta h) $$

or for $\Delta h > 0$

$$ \frac{{}_{h+\Delta h}p_x^{\overline{i,i}} - {}_h p_x^{\overline{i,i}}}{\Delta h} = -{}_h p_x^{\overline{i,i}} \sum_{j=0, j\neq i}^{m} \mu_x^{i,j}(h) + \frac{o(\Delta h)}{\Delta h}. $$

Then take $\Delta h \downarrow 0$, and obtain □

Proof.

$$\frac{d}{dh}{}_hp_x^{\overline{i,i}} = -{}_hp_x^{\overline{i,i}} \sum_{j=0, j\neq i}^{m} \mu_x^{i,j}(h).$$

This is an ODE we have already seen, and its solution is exactly the assertion of this proposition.. □

Think of the ODE

$$f'(h) = g(f(h), h),$$

that has the initial condition $f(0) = c > 0$; set $f(h) = {}_hp_x$,$g(f(h), h) = -{}_hp_x \times \mu_x(h)$ such that ${}_0p_x = 1$. The solution is

$${}_hp_x = \exp\left\{-\int_0^h \mu_x(s)ds\right\}.$$

Proposition 1.6

For any multiple-state model with $\mathcal{S} = \{0,\ 1,\ \ldots, m \in \mathbb{N}\}$, $i, j \in \mathcal{S}$, *we have*

$$\frac{d}{dh}\,{}_h p_x^{i,j} = \sum_{k \neq j} {}_h p_x^{i,k} \mu_x^{k,j}(h) - {}_h p_x^{i,j} \sum_{k \neq j} \mu_x^{j,k}(h).$$

Proof.

Start with the expression

$$\begin{aligned} {}_{h+\Delta h} p_x^{i,j} &= \sum_{k \in \mathcal{S}} {}_h p_x^{i,k} \times {}_{\Delta h} p_{x+h}^{k,j} \\ &= \sum_{k \neq j} {}_h p_x^{i,k} \times {}_{\Delta h} p_{x+h}^{k,j} + {}_h p_x^{i,j} \times {}_{\Delta h} p_{x+h}^{j,j}, \end{aligned}$$

which hold by conditioning. □

Proof.

Further we have

$$
\begin{aligned}
& {}_{h+\Delta h}p_x^{i,j} \\
= & \sum_{k\neq j} {}_hp_x^{i,k} \times (\Delta h \mu_x^{k,j}(h) + o(\Delta h)) + {}_hp_x^{i,j} \times \left(1 - \sum_{k\neq j} {}_{\Delta h}p_{x+h}^{j,k}\right),
\end{aligned}
$$

which yields

$$
\begin{aligned}
& {}_{h+\Delta h}p_x^{i,j} - {}_hp_x^{i,j} \\
= & \sum_{k\neq j} {}_hp_x^{i,k} \times (\Delta h \mu_x^{k,j}(h) + o(\Delta h)) \\
- & {}_hp_x^{i,j} \times \sum_{k\neq j} \left(\Delta h \mu_x^{j,k}(h) + o(\Delta h)\right).
\end{aligned}
$$

Now divide by $\Delta h > 0$ throughout and get □

Proof.

$$\begin{aligned} & \frac{{}_{h+\Delta h}p_x^{i,j} - {}_h p_x^{i,j}}{\Delta h} \\ = \; & \sum_{k \neq j} {}_h p_x^{i,k} \times \mu_x^{k,j}(h) \\ - \; & {}_h p_x^{i,j} \times \sum_{k \neq j} \mu_x^{j,k}(h) + o(\Delta h). \end{aligned}$$

Finally, take the limit $\Delta h \downarrow 0$, and the assertion follows. □