# STOCHASTIC PROCESSES

STAT 150


Instructor: Benson Au

Kelvin Lee

UC Berkeley

# Contents

# 1 Probability Review

## 1.1 Basic Definitions

**Definition 1.1.1** (Probability Space). A *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ is a triple consisting of a set $\Omega$ called the *sample space*, a set $\mathcal{F} \subseteq \Omega$ satisfying certain closure properties, and a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ that assigns probabilities to events in a coherent way.

**Requirements for $\mathcal{F}$:**

(i) $\Omega \in \mathcal{F}$.

(ii) If $E \in \mathcal{F}$, then $E^c \in \mathcal{F}$.

(iii) If $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$, then

$$\bigcap_{i=1}^{\infty} E_i \in \mathcal{F}.$$

**Requirements for $\mathbb{P}$:**

(i) $\mathbb{P}(\Omega) = 1$.

(ii) If $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ are pairwise disjoint (meaning $E_i \cap E_j = \varnothing$ for $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}\left(E_i\right)$$

**Definition 1.1.2** (Random Variable). A *random variable* is a function $X : \Omega \to \mathbb{R}$ such that $X^{-1}(B) \in \mathcal{F}$ whenever $B$ is a "nice" subset of $\mathbb{R}$.

**Example 1.1.3.** $\Omega = \{H, T\}, \quad \mathcal{F} = 2^{\Omega}, \quad \mathbb{P}(\{H\}) = \frac{1}{2}$. $X(H) = 1, \quad X(T) = 0.$

$$\mathbb{P}(X = 1) = \mathbb{P}(\{H\}) = \frac{1}{2}, \quad \mathbb{P}(X = 0) = \mathbb{P}(\{T\}) = \frac{1}{2}.$$

## 1.2 Overview

**Definition 1.2.1** (Stochastic Process). A *stochastic process* is a collection $\{X_t : t \in T\}$ of random variables $X_t : \Omega \to S \subseteq \mathbb{R}$ all defined on the some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here $T$ is some index set (typically representing time) and $S$ is the *state space*. One write this as

$X : \Omega \times T \to S$, $(w, t) \mapsto X_t(\omega)$. For a given outcome $\omega \in \Omega$, we get a sample path trajectory $X(\omega) : T \to S, t \mapsto X_t(\omega)$. A stochastic process can then be thought of as a random function.

The theme of this course is what can we say about the distribution of trajectories?

**Example 1.2.2** (Branching Process (DTDS)). $X_0 = 1$, one individual in the 0th generation individuals produce a random number of offspring, i.i.d. $(\xi_i^{(n)})_{i\in\mathbb{N},n\in\mathbb{N}_0}$.

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n)}.$$

One interesting question would be what is $\mathbb{P}(X_n = 0 \text{ eventually})$, the probability of dying out?

**Example 1.2.3** (Poisson Process (CTDS)). Recall that the Poisson distribution is used to model the number of occurrences of a rare event in some fixed period of time. The Poisson process $(N_t)_{t\geq 0}$ models the number of occurrences throughout time. $N_t = \#$ of occurrences by time $t$.

## 1.3 Useful Properties

(i) *(DeMorgan)*

$$(E \cup F)^c = E^c \cap F^c, \quad (E \cap F)^c = E^c \cup F^c.$$

(ii) *(Complementation)*

$$\mathbb{P}(E) = 1 - \mathbb{P}\left(E^c\right).$$

(iii) *(Inclusion-exclusion)*

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$$

$$\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{j=1}^{n}(-1)^{j-1} \sum_{S\in[n]:|S|=j} \mathbb{P}\left(\bigcap_{i\in S} E_i\right).$$

(iv) *(Partitioning)* If $\bigsqcup_{i=1}^{\infty} E_i = \Omega$, then

$$\mathbb{P}(F) = \mathbb{P}\left(\bigsqcup_{i=1}^{\infty}(F \cap E_i)\right) = \sum_{i=1}^{\infty} \mathbb{P}(F \cap E_i)$$

## 1.4 Conditional Probability

**Conditioning:** For $\mathbb{P}(F) > 0$,

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

$\mathbb{P}(\cdot \mid F)$ defines a new probability measure on $(\Omega, \mathcal{F})$.

**Multiplication rule:**

$$\mathbb{P}(E \cap F) = \mathbb{P}(F)\mathbb{P}(E \mid F).$$

If $\bigsqcup_{i=1}^{\infty} F_i = \Omega$, then

$$\mathbb{P}(E) = \sum_{i=1}^{\infty} \mathbb{P}(E \cap F_i) = \sum_{i=1}^{\infty} \mathbb{P}(F_i)\mathbb{P}(E \mid F_i).$$

**Bayes' rule:**

$$\mathbb{P}(F_j \mid E) = \frac{\mathbb{P}(F_j)\mathbb{P}(E \mid F_j)}{\sum_{i=1}^{\infty} \mathbb{P}(F_i)\mathbb{P}(E \mid F_i)}$$

## 1.5 Random Variables

### 1.5.1 Discrete Random Variables

If $X : \Omega \to S \subseteq \mathbb{R}$ is discrete,

$$\mathbb{P}(X \in E) = \sum_{x \in E} \mathbb{P}(X = x) = \sum_{x \in E \cap S} \mathbb{P}(X = x).$$

### 1.5.2 Indicator Random Variable

$$X(\omega) = \mathbb{I}_E(\omega) = \begin{cases} 1 & \omega \in E \\ 0 & \omega \notin E. \end{cases}$$

#### 1.5.2.1 Binomial Random Variable

$$X = \sum_{i=1}^{n} \mathbb{I}_{E_i}, \quad \mathbb{P}(E_i) = p.$$

$$p_X(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

### 1.5.3 Continuous Random Variables

If $X$ continuous,

$$\mathbb{P}(X \in E) = \int_E f_X(x)dx.$$

$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)dx.$$

#### 1.5.3.1 Exponential Random Variable

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{I}_{x \geq 0}.$$

#### 1.5.3.2 Gaussian Random Variable

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}.$$

### 1.5.4 Cumulative Distribution Function (CDF)

$F_X : \mathbb{R} \to [0, 1]$,

$$F_X(r) = \mathbb{P}(X \leq r) = \mathbb{P}(X \in (-\infty, r]).$$

If $X$ is discrete,

$$F_X(r) = \sum_{x_i \leq r} p_X(x_i).$$

If $X$ is continuous,

$$F_X(r) = \int_{-\infty}^{r} f_X(x) dx.$$

$$\frac{d}{dr} F_X(r) = f_X(r).$$

### 1.5.5 Expectation

#### 1.5.5.1 Discrete case

$$\mathbb{E}[X] = \sum_{x_i \in S} x_i \mathbb{P}(X = x_i).$$

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n).$$

$$\mathbb{E}[g(X)] = \sum_{x_i \in S} g(x_i) \mathbb{P}(X = x_i).$$

#### 1.5.5.2 Continuous case

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx.$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \mathbb{P}(X \geq x) dx.$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

### 1.5.6 Variance

$$\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

### 1.5.7 Moments

$$\mathbb{E}[X^m] = \int_0^\infty m x^{m-1} \mathbb{P}(X \geq x) dx.$$

### 1.5.8 Joint Distribution

#### 1.5.8.1 Discrete

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

#### 1.5.8.2 Continuous

$$\mathbb{P}((X, Y) \in E) = \int\int_E f_{X,Y}(x, y) dx dy$$

#### 1.5.8.3 Marginal Distribution

$$p_X(x) = \sum_{y \in S_Y} p_{X,Y}(x, y).$$

$$f_X(x) = \int_{y \in S_y} f_{X,Y}(x, y) dy$$

### 1.5.9 Independence

$$p_{X,Y}(x, y) = p_X(x) p_Y(y).$$

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

$$\mathbb{P}(X \leq x, Y \leq y) = F_X(x) F_Y(y).$$

### 1.5.10 Linearity of Expectation

$$\mathbb{E}\left[\sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n c_i \mathbb{E}[X_i]$$

If $(X_i)_{i=1}^n$ independent,

$$(g(X_i))_{i=1}^n$$

independent.

$$\mathbb{E}\left[\prod_{i=1}^n g(X_i)\right] = \prod_{i=1}^n \mathbb{E}[g(x_i)]$$

$$\operatorname{Var}\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \operatorname{Var}(x_i)$$

In general,

$$\operatorname{Var}\left(\sum_{i=1}^{n} x_i\right) = \sum_{i,j=1}^{n} \operatorname{Cov}(x_i, x_j)$$

### 1.5.11 Convolution

**Discrete case:** $X, Y$ discrete $X \perp\!\!\!\perp Y$

$$\begin{aligned}
\mathbb{P}(X+Y=z) &= \sum_{Y} \mathbb{P}(X+Y=z, Y=y) \\
&= \sum_{y} \mathbb{P}(X=z-y, Y=y) \\
&= \sum_{y} \mathbb{P}(X=z-y)\mathbb{P}(Y=y) \quad \left(= \sum_{x} \mathbb{P}(X=x)\mathbb{P}(Y=z-x)\right).
\end{aligned}$$

If $X, Y$ are $\mathbb{Z}$-valued, this becomes

$$\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k=-\infty}^{\infty} \mathbb{P}(X=n-k)\mathbb{P}(Y=y) \\
&= \sum_{h=-\infty}^{\infty} \mathbb{P}(X=k)\mathbb{P}(Y=n-k) \\
&= (\mathbb{P}_X * \mathbb{P}_Y)(n)
\end{aligned}$$

**Example 1.5.1** (Poisson). $X \sim$ Poisson $(\lambda)$, $Y \sim$ Poisson $(\mu)$, $X+Y \sim$ Poisson $(\lambda+\mu)$

$$\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{h=0}^{n} \mathbb{P}(X=k)P(Y=n-k) \\
&= \sum_{k=0}^{n} e^{-\lambda}\frac{\lambda^k}{\mu!} e^{-\mu}\frac{\mu^{n-\mu}}{n-k!} \\
&= e^{-(\lambda+\mu)} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\lambda^k \mu^{n-k} \\
&= \frac{e^{-(\lambda+\mu)}}{n!}(\lambda+\mu)^n \\
&= \mathbb{P}(Z=n)
\end{aligned}$$

where $Z \sim \text{Poisson}(\lambda+\mu)$.

**Continuous case:** $X, Y$ continuous

$$\begin{aligned}
\mathbb{P}(X+Y \leqslant z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_X(x) f_Y(y)\,dy\,dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{z} f_X(x) f_Y(y-x)\,dy\,dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_X(x) f_Y(y-x)\,dx\,dy
\end{aligned}$$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx = \int_{-\infty}^{\infty} f_Y(y) f_X(z-y)\,dy = f_X * f_Y.$$

**Example 1.5.2** (Convolution in uniform distributions). $X, Y \sim U[0,1]$, $X \perp\!\!\!\perp Y$.

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$
$$f_X(x) = \mathbb{I}_{[0,1]}(x) \qquad f_Y(y) = \mathbb{I}_{[0,1]}(y)$$

so

$$\begin{aligned} f_{X+Y}(z) &= \int_{x\in[0,1], z-x\in[0,1]} 1dx \\ &= \int_{x\in[0,1], x\in[-1+z,z]} 1dx \\ &= \int_{\max(0,-1+z)}^{\min(1,z)} 1dx \\ &= \min(1,z) - \max(0,-1+z). \end{aligned}$$

### 1.5.12 Gamma Distribution

**Definition 1.5.3** (Gamma function). Let $\alpha > 0$. The *gamma function* $\Gamma : (0,\infty) \to (0,\infty)$ is defined by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx = \mathbb{E}[X^{\alpha-1}]$$

where $X \sim \text{Exp}(1)$ Let $\alpha, \lambda > 0$. The $\text{Gamma}(\alpha, \lambda)$ distribution is defined by

$$f(x) = \frac{\lambda^{\alpha} x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} \mathbb{I}_{x\geq 0}.$$

**Exercise 1.5.4.** $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$. (Hint: use induction)

### 1.5.13 Moment Generating Function

**Definition 1.5.5** (MGF). For a random variable $X$, the *moment generating function* (MGF) is the function $M_X : \mathbb{R} \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$,

$$M_X(t) = \mathbb{E}[e^{tX}].$$

If $M_X(t) < +\infty$ for $t \in (-\epsilon, \epsilon)$, then

$$M_x(t) = \sum_{k=0}^{\infty} \frac{t^k \mathbb{E}[x^k]}{k!} \text{ for } |t| < \epsilon$$

For independent RVs $(X_i)_{i=1}^n$,

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t)$$

**Exercise 1.5.6.** If $X \sim \text{Exp}(\lambda)$, then $M_x(t) = \frac{\lambda}{\lambda - t}$ if $t < \lambda, +\infty$ otherwise.
If $X \sim \text{Gamma}(n, \lambda)$, then

$$M_x(t) = \left(\frac{\lambda}{\lambda - t}\right)^n.$$

If $X \sim \text{Gamma}(\alpha, \lambda)$, then

$$M_x(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}.$$

## 1.6 Conditional Probability (Cont'd)

**Exercise 1.6.1** (Generalization). $(X_i)_{i=1}^n, (Y_j)_{j=1}^m$

$$p_{X_1,\ldots,X_n|Y_1,\ldots,Y_m}(x_1,\ldots,x_n \mid y_1,\ldots,y_m) = ?$$

**Example 1.6.2.** Let $M \in \mathbb{N}$ and $p, q \in (0,1)$. Consider $N \sim \text{Bin}(M,q)$ and $X \sim \text{Bin}(N,p)$. What is the distribution of $X$?

$$\begin{aligned}
\mathbb{P}(X=k) &= \sum_{n=0}^{M} \mathbb{P}(N=n)\mathbb{P}(X=k \mid N=n) \\
&= \sum_{n=0}^{M} \binom{M}{n} q^n (1-n)^{M-n} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{p^k}{k!} \sum_{n=k}^{M} \frac{M!}{(M-n)!(n-k)!} q^n (1-q)^{M-n} (1-p)^{n-k} \\
&= \frac{p^k}{k!(M-k)!} \sum_{n=k}^{M} \frac{M!(M-k)!}{(M-n)!(n-k)!} q^n (1-q)^{M-n} (1-p)^{n-k} \\
&= \binom{M}{k} p^k q^k \sum_{n=k}^{M} \binom{M-k}{n-k} q^{n-k} (1-q)^{M-j} (1-p)^{n-k} \\
&= \binom{M}{k} p^k q^k \sum_{t=0}^{M-k} \binom{M-k}{t} (q(1-p))^t (1-q)^{M-k-t} \\
&= \binom{M}{k} (pq)^k (q(1-p) + (1-q))^{M-k} \\
&= \binom{M}{k} (pq)^k (1-pq)^{M-k}.
\end{aligned}$$

Thus, $X \sim \text{Bin}(M, pq)$.

**Remark.** What if $k > n$ in $\mathbb{P}(X=k \mid N=n)$ above in the first line? The probability is simply 0.

**Question.** Why does this answer make sense?

**Answer.** Think about retesting whenever we succeeded for the first $M$ trials. Then $X$ is simply the number of trials with double successes, thus we have the $pq$ parameter.

**Exercise 1.6.3.** Consider $N \sim \text{Poisson}(\lambda), X \sim \text{Bin}(N,p)$. What is the distribution of $X$?

**Answer.** $X \sim \text{Poisson}(\lambda p)$.

**Question.** How can we interpret this?

**Answer.** We can interpret $X$ as the number of customers visiting a store who purchase something.

### 1.6.1 Conditional Expectation

For $X, Y$ discrete, $g : \mathbb{R} \to \mathbb{R}$. Assume $\mathbb{E}[|g(X)|] = \sum_x |g(x)p_X(x) < \infty$.

**Definition 1.6.4** (Conditional expectation). The *conditional expectation* is defined as

$$\mathbb{E}[g(X) = y] = \sum_x g(x) p_{X|Y}(x|y)$$

if $p_Y(y) \neq 0$.

**Remark.** Note that $\mathbb{E}[g(X) \mid Y = y]$ is a real number, whereas $\mathbb{E}[g(X) \mid Y]$ is a random variable.

#### 1.6.1.1 Tower Property

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[g(X) \mid Y]] &= \mathbb{E}\left[\sum_y \mathbb{E}[g(X) \mid Y = y]\right] \\
&= \sum_y \mathbb{E}[g(X) \mid Y = y] p_Y(y) \\
&= \sum_y \sum_x g(x) p_{X|Y}(x|y) p_Y(y) \\
&= \sum_x g(x) \sum_y p_{X|Y}(x|y) p_Y(y) \\
&= \sum_x g(x) p_X(x) \\
&= \mathbb{E}[g(X)].
\end{aligned}$$

**Remark.** One intuitive example would be considering the averages of heights of students from a classroom. We divide it into several groups and let $Y$ denote the whichever group we select and let $\mathbb{E}[g(X) \mid Y]$ be the average of those from group $Y$. Then the average height of the entire classroom $\mathbb{E}[g(X)]$ is equivalent to the average of the average of heights of each group, which is $\mathbb{E}[\mathbb{E}[g(X) \mid Y]]$.

**Properties of conditional expectations:**

1. $\mathbb{E}\left[c_1 g\left(x_1\right) + c_2 h\left(x_2\right) \mid Y = y\right] = c_1 E\left[g\left(X_1\right) \mid Y = y\right] + c_2 E\left[h\left(X_2\right) \mid Y = y\right]$
2. If $g \geq 0$, then $E[g(x) \mid Y = y] \geq 0$.
3. $\mathbb{E}[f(X, Y) \mid Y = y] = \mathbb{E}[f(X, y) \mid Y = y]$.
4. If $X \perp\!\!\!\perp Y, \mathbb{E}[g(X) \mid Y = y] = \mathbb{E}[g(X)]$
5. $\mathbb{E}[g(x)h(y) \mid Y = y] = h(y)\mathbb{E}[g(x) \mid Y = y]$
6. $\mathbb{E}[g(x)h(y)] = \sum_y h(y) E[g(x) \mid Y = y] p_y(y) = \mathbb{E}[h(Y)\mathbb{E}[g(X) \mid Y]]$

*Proof of 3.*

$$\begin{aligned}
\mathbb{E}[f(X,Y) \mid Y=y] &= \sum_{x,z} f(x,z) p_{X,Y|Y}(x,z|y) \\
&= \sum_{x,z} f(x,z) \frac{p_{X,Y,Y}(x,z,y)}{p_Y(y)} \\
&= \sum_{x} f(x,y) \frac{p_{X,Y}(x,y)}{p_Y(y)} \\
&= \mathbb{E}[f(X,y) \mid Y=y].
\end{aligned}$$

□

**Remark.** $\mathbb{E}[f(X,y)] \neq \mathbb{E}[f(X,y) \mid Y=y]$.

# 2 Random Sums

**Definition 2.0.1.** Let $(\xi_i)_{i=1}^{\infty}$ be i.i.d random variables, $N$ be a $\mathbb{N}_0$-valued random variable, $N \perp\!\!\!\perp (\xi_i)_{i=1}^{\infty}$. The *random sum* is defined as

$$X = \sum_{i=1}^{N} \xi_i = \sum_{n=0}^{\infty} \left( \sum_{i=1}^{n} \xi_i \right) \mathbf{1}_{N=n} = \begin{cases} \sum_{i=1}^{n} \xi_i & \text{if } N = n \geq 1 \\ 0 & \text{if } N = 0. \end{cases}$$

**Question.** What is the distribution of $X$?

Let $X, N$ be random variables. $N$ is $\mathbb{N}_0$-valued. The condition CDF is

$$F_{X|N}(x|n) = \mathbb{P}(X \leq x \mid N = n)$$

if $\mathbb{P}(N = n) \neq 0$. The is an actual CDF, but for the random variable $X \mid N = n$.

Suppose that $X$ is continuous and $F_{X|N}(x|n)$ is a differentiable function of $x$ for each $n$ such that $p_N(n) > 0$. The conditional PDF is

$$f_{X|N}(x|n) = \frac{d}{dx} F_{X|N}(x|n).$$

$$\begin{aligned} \int_a^b f_{X|N}(x|n) dx &= F_{X|N}(b|n) - F_{X|N}(a|n) \\ &= \mathbb{P}(X \in [a, b] \mid N = n). \end{aligned}$$

**Answer.**

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{n=0}^{\infty} \mathbb{P}(N = n) \mathbb{P}(X \leq x \mid N = n).$$

$$f_X(x) = \sum_{n=0}^{\infty} \mathbb{P}(N = n) f_{X|N}(x|n).$$

## 2.1 Mean and Variance of Random Sums

Assume $\mathbb{E}[N] = \nu$ and $\mathbb{E}[\xi_i] = \mu$. Then

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X \mid N]] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{N} \xi_i \mid N\right]\right] \\
&= \mathbb{E}[N\mathbb{E}[\xi_1]] \\
&= \mathbb{E}[N\mu] \\
&= \mu\nu.
\end{aligned}$$

# 3 Markov Chains

## 3.1 Discrete-time Markov Chains

**Definition 3.1.1** (Markov process)**.** A is a stochastic process $(X_t)_{t\in T}$ such that the future, given the present, is independent of the past.

**Definition 3.1.2** (Markov property)**.** The *Markov property* for a DTDS stochastic process is

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}\left(X_{n+1} = j \mid X_n = i\right)$$

**Example 3.1.3** (Gambler's ruin)**.** $(X_n)_{n=0}^{\infty}, X_n =$ your wealth after $n$ turns. Stop if $X_n = 0$ or 5. Each play, you win \$1 with probability $p$ and lose \$1 with probability $1-p$ independently of all previous plays. This process satisfies the markov property.

**Example 3.1.4** (Ehrenfest model)**.** Box of $N$ particles. $X_n =$ number of particles on the left side at time $n$. $N - X_n$ be the number of particles on the other side.

$$\mathbb{P}(X_{n+1} = i+1 \mid X_n = i) = \frac{N-i}{N}$$
$$\mathbb{P}(X_{n+1} = i-1 \mid X_n = i) = \frac{i}{N}.$$

**Theorem 3.1.5.**
Joint PMF of the Markov Chain is determined by initial distribution and $P = (p_{i,j})_{i,j\in S}$.

*Proof.*

$$\begin{aligned}\mathbb{P}(X_n = i_n, \ldots, X_0 = i_0) &= \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)\mathbb{P}(X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ &= p_{i_{n-1},i_n} p_{i_{n-2},i_{n-1}} \cdots p_{i_0,i_1} \mathbb{P}(X_0 = i_0).\end{aligned}$$

□

### 3.1.1 $n$-step transition probabilities

$$p_{i,j} = \mathbb{P}(X_{n+1} = j \mid X_n = i).$$

**Theorem 3.1.6.**

$$p_{i,j}^{(m)} = \mathbb{P}(X_{n+m} = j \mid X_n = i) = (P^m)_{i,j}.$$

*Proof.*

$$\begin{aligned}
\mathbb{P}(X_{n+m+1} = j \mid X_n = i) &= \sum_{k \in S} \mathbb{P}(X_{n+m+1} = j, X_{n+m} = k \mid X_n = i) \\
&= \sum_{k \in S} \mathbb{P}(X_{n+m+1} = j \mid X_{n+m} = k)\mathbb{P}(X_{n+m} = k \mid X_n = i).
\end{aligned}$$

□

**Example 3.1.7.**

$$\begin{aligned}
p_{i,j}^{(2)} &= \mathbb{P}(X_2 = j \mid X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(X_2 = j, X_1 = k \mid X_0 = i) \\
&= \sum_{k \in S} \mathbb{P}(X_2 = j \mid X_1 = k, X_0 = i)\mathbb{P}(X_1 = k \mid X_0 = i) \\
&= \sum_{k \in S} P_{i,k} P_{k,j} \\
&= (P^2)_{i,j}
\end{aligned}$$

**Example 3.1.8** (Inventory model)**.** $X_n$ = inventory that you have of this product after the $n$th business day. If $X_n \leq s$, place an order that brings inventory back to $S$ by next morning. $\xi_n$ = demand on $n$th day and $(\xi_n)$ are i.i.d..

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \begin{cases} \mathbb{P}(\xi_{n+1} = S - j) & \text{if } i \leq s \\ \mathbb{P}(\xi_{n+1} = i - j) & \text{if } i > s. \end{cases}$$

$\lim_{n\to\infty} \mathbb{P}(X_n < 0)$ = chance of excess demand.

## 3.2 First Step Analysis

Consider $(X_n)_{n\geq 0}$ Markov chain on $\{1, \ldots, r\} \cup \{r+1, \ldots, N\}$ where $\{1, \ldots, r\}$ are the *transient states* and $\{r+1, \ldots, N\}$ are the *absorbing states* such that

$$\begin{aligned}
\lim_{n\to\infty} p_{i,j}^{(n)} &= 0 \qquad \forall i, j \in \{1, \ldots, r\} \\
\lim_{n\to\infty} p_{i,i}^{(n)} &= 1 \qquad \forall i \in \{r+1, \ldots, N\}
\end{aligned}$$

Then we can express the transition matrix $P$ as

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix},$$

where $Q$ and $R$ is some transition matrices for the corresponding partitioned states and 0 is the zero matrix and $I$ is the identity matrix.

Let $T = \min\{n \geq 0 : X_n \geq r+1\}$ be the time of absorption and $X_T$ be the state we get absorbed into. Define $u_{i,k} = \mathbb{P}(X_T = k \mid X_0 = i)$. Then we have

$$\begin{aligned}
u_{i,k} &= \sum_{j=1}^{N} \mathbb{P}(X_T = k, X_1 = j \mid X_0 = i) \\
&= \sum_{j=1}^{N} \mathbb{P}(X_T = k \mid X_1 = j, X_0 = i)\mathbb{P}(X_1 = j \mid X_0 = i) \\
&= \sum_{j=1}^{N} p_{i,j}\mathbb{P}(X_T = k \mid X_1 = j) \\
&= \sum_{j=1}^{N} \mathbb{P}(X_T = k \mid X_0 = j) \\
&= \sum_{j=1}^{r} p_{i,j}u_{j,k} + \sum_{j=r+1, j\neq k}^{N} p_{i,j}u_{j,k} + p_{i,k}u_{k,k}.
\end{aligned}$$

Thus,

$$u_{i,k} = \sum_{j=1}^{r} P_{i,j}u_{j,k} + p_{i,k}$$

Hence, we have

$$U = QU + R \implies (I-Q)U = R \implies U = (I-Q)^{-1}R,$$

where $U$ contains all the $(u_{i,k})_{i\in\{1,\ldots,r\},k\in\{r+1,\ldots,N\}}$.

### 3.2.1 The General Absorbing Markov Chain

Let's suppose that associated with each transient state $i$ is a rate $g(i)$ and that we wish to determine the mean total rate that is accumulated up to absorption. Let $v_i$ be this mean total amount, where the subscript $i$ denotes the starting position $X_0 = i$, i.e.,

$$v_i = \mathbb{E}\left[\sum_{n=0}^{T-1} g(X_n) \mid X_0 = i\right]$$

The choice $g = 1$ will give $v_i = \mathbb{E}[T \mid X_0 = i]$. We can also write for $i \in \{1, \ldots, r\}$ that

$$\begin{aligned}
v_i &= g(i) + \mathbb{E}\left[\sum_{n=1}^{T-1} g(X_n) \mid X_0 = i\right] \\
&= g(i) + \sum_{j=1}^{N} p_{i,j}v_j (= \sum_{j=1}^{N} p_{i,j}(g(i) + v_i)).
\end{aligned}$$

Then we can condense this into the following form

$$v = g + Qv \implies v = (I-Q)^{-1}g.$$

where $v = (v_i)_{i\in\{1,\ldots,r\}}$ and $g = (g(i))_{i\in\{1,\ldots,r\}}$.