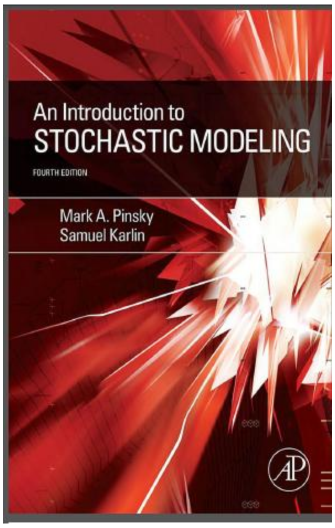
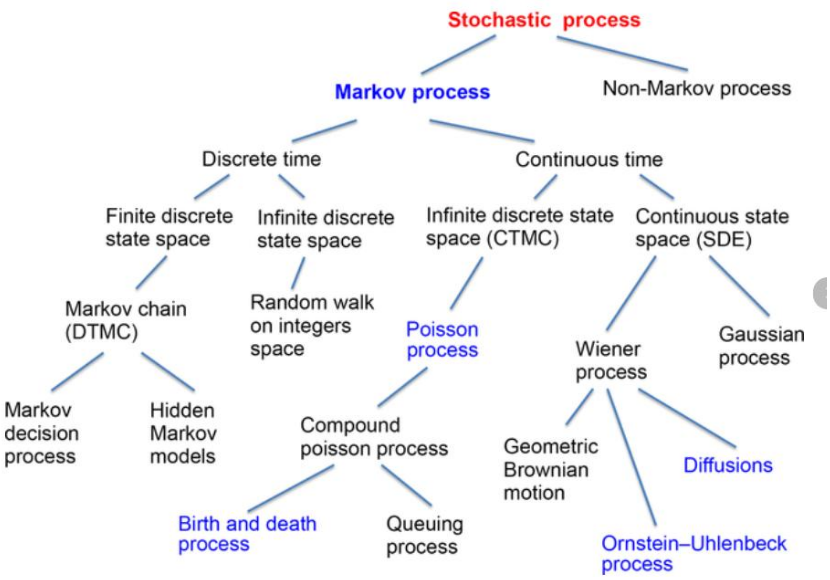
King Saud University
College of sciences
Department of Mathematics

Math-380
Terms & Formulas



Prepared by:
Lecturer: Fawaz bin Saud Alotaibi



1 Introduction

Term or Formula	Explanation
stochastic process	family of random variables X(t)
index set T	T = {0, 1, 2, . . .} or T = [0,∞)
state space	Range of possible values for X_t
random variable X	variable that takes on its values by chance
sample space Ω	a set Ω whose elements ω correspond to the possible outcomes of an experiment
Events	the family of events, a collection $\mathcal{F}$ of subsets A of Ω: we say that the event A *occurs* if the outcome ω of the experiment is an element of A; and the *probability measure*, a function P defined on $\mathcal{F}$ and satisfying (a) $0 = P[\varnothing] \le P[A] \le P[\Omega] = 1 \quad \text{for } A \in \mathcal{F}$ ($\varnothing$ = the empty set) (b) $P\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{\infty} P[A_n]$, (2.19) if the events $A_1, A_2, \ldots$ are disjoint; i.e., if $A_i \cap A_j = \varnothing$ when $i \neq j$.
probability space	The triple $(\Omega, \mathcal{F}, P)$
the law of total probability,	$\Pr\{A\} = \sum_{i=1}^{\infty} \Pr\{A \mid B_i\} \Pr\{B_i\}.$ $\Pr\{A \cap B_i\} = \Pr\{A \mid B_i\} \Pr\{B_i\}$ into $\Pr\{A\} = \Sigma_{i=1}^{\infty} \Pr\{A \cap B_i\}$, where $\Omega = B_1 \cup B_2 \cup \cdots$ and $B_i \cap B_j = \varnothing$ if $i \neq j$ (cf. Section 2.1), to yield
distribution function of X	$F(x) = \Pr\{X \le x\}, \qquad -\infty < x < +\infty,$
Discrete X	set of distinct values $x_1, x_2, \ldots$ such that $a_i = \Pr\{X = x_i\} > 0$ for $i = 1, 2, \ldots$ and $\Sigma_i a_i = 1$. The function
probability mass function for X	$p(x_i) = p_X(x_i) = a_i \qquad \text{for } i = 1, 2, \ldots$
Continuous X	If $\Pr\{X = x\} = 0$ for every value of x,

1 Introduction

probability density function for X	function of x. If there is a nonnegative function $f(x) = f_X(x)$ defined for $-\infty < x < \infty$ such that $$\Pr\{a < X \le b\} = \int_a^b f(x)\,dx \quad \text{for } -\infty < a < b < \infty, \qquad (2.2)$$
distribution function of X	$$F(x) = \int_{-\infty}^{x} f(\xi)\,d\xi, \qquad -\infty < x < \infty.$$ $$f(x) = \frac{d}{dx}F(x) = F'(x), \qquad -\infty < x < \infty.$$
the expectation of g(X)	$$E[g(X)] = \sum_{i=1}^{\infty} g(x_i)\Pr\{X = x_i\},$$ If X is a discrete random variable
the expectation of g(X)	$$E[g(X)] = \int g(x)f_X(x)\,dx.$$ If X is continuous
mth moment of X	$$E[X^m] = \sum_i x_i^m \Pr\{X = x_i\},$$ If X is a discrete random variable
	$$E[X^m] = \int_{-\infty}^{+\infty} x^m f(x)\,dx,$$ If X is continuous
mean or expected value of X (m_X or μ_X.)	m=1 in mth moment of X
variance of X (σ_X^2 or $\mathrm{Var}[X]$.)	$$\mathrm{Var}[X] = E[(X-\mu)^2] = E[X^2] - \mu^2.$$

1 Introduction

Bernoulli Distribution (with parameter p)	$E[X] = p$ and $\mathrm{Var}[X] = p(1-p)$, Its mass prob. fn is $pr(X=x) = p^{x}(1-p)^{1-x}, \quad x \in \{0,1\}$
Binomial Distribution (with parameters n and p)	$Y = \mathbf{1}(A_1) + \cdots + \mathbf{1}(A_n)$ $E[Y] = E[\mathbf{1}(A_1)] + \cdots + E[\mathbf{1}(A_n)] = np,$ $\mathrm{Var}[Y] = \mathrm{Var}[\mathbf{1}(A_1)] + \cdots + \mathrm{Var}[\mathbf{1}(A_n)] = np(1-p).$ $p_Y(k) = \Pr\{Y = k\}$ $= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad \text{for } k = 0, 1, \ldots, n. \qquad (3.3)$
Poisson Distribution (with parameter k > 0)	$E[X] = \lambda$, $Var(X) = \lambda$ $p(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{for } k = 0, 1, \ldots.$
Geometric Distribution (with parameter p)	$E[Z] = \frac{1-p}{p}; \quad \mathrm{Var}[Z] = \frac{1-p}{p^2}.$ $p_Z(k) = p(1-p)^k \quad \text{for } k = 0, 1, \ldots,$
Uniform Distribution	$E[U] = \frac{1}{2}(a+b)$ and $\mathrm{Var}[U] = \frac{(b-a)^2}{12}.$ $f_U(u) = \begin{cases} \frac{1}{b-a} & \text{for } a \le u \le b, \\ 0 & \text{elsewhere.} \end{cases}$, $F_U(x) = \begin{cases} 0 & \text{for } u \le a, \\ \frac{x-a}{b-a} & \text{for } a < x \le b, \\ 1 & \text{for } x > b, \end{cases}$
Exponential Distribution (with parameter $\lambda > 0$)	$E[T] = \frac{1}{\lambda}$ and $\mathrm{Var}[T] = \frac{1}{\lambda^2}.$ $f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{for } t \ge 0, \\ 0 & \text{for } t > 0. \end{cases}$ $F_T(t) = \begin{cases} 1 - e^{-\lambda t} & \text{for } t \ge 0, \\ 0 & \text{for } t > 0, \end{cases}$,
memoryless property	There is no memory in the sense that $\Pr\{T - t > x \mid T > t\} = e^{-\lambda x} = \Pr\{T > x\}$, and an item that has survived for t units of time has a remaining lifetime that is statistically the same as that for a new item.

1 Introduction

hazard rate or failure rate r(s)	The exponential distribution is uniquely the continuous distribution with the *constant* failure rate $r(t) \equiv \lambda$.
Gamma Distribution (with parameters $\alpha > 0, \lambda > 0$) $\Gamma(n) = (n-1)!$	$$E[X_\alpha] = \frac{\alpha}{\lambda} \quad \text{and} \quad \operatorname{Var}[X_\alpha] = \frac{\alpha}{\lambda^2},$$ $$f(x) = \frac{\lambda}{\Gamma(\alpha)} (\lambda x)^{\alpha-1} e^{-\lambda x} \qquad \text{for } x > 0.$$
joint distribution function	$$F_{XY}(x, y) = F(x, y) = \Pr\{X \leq x \text{ and } Y \leq y\}.$$ $$F(x,y) = pr(X \leq x, Y \leq y) = \sum_{X \leq x} \sum_{Y \leq y} p(x,y)$$
(joint) probability density	$f_{X,Y}(x,y)$ $$F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(\xi, \eta)\, d\eta\, d\xi \qquad \text{for all } x, y.$$ $$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$
(joint) probability mass	$p_{X,Y}(x,y) = pr(X = x, Y = y)$ دالة الكتلة الاحتمالية المشتركة $pr(a \leq X \leq b;\ c \leq Y \leq d) = \sum_{X=a}^{b} \sum_{Y=c}^{d} p(x,y)$
marginal density function of X	$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$$
marginal distribution function of X	$$F_X(x) = \lim_{y \to \infty} F(x, y)$$
marginal density function of Y	$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\, dx.$$
marginal distribution function of Y	$$F_Y(y) = \lim_{y \to \infty} F(x, y)$$

1 Introduction

marginal mass function of X	$p_X(x) = \sum_y p_{X,Y}(x,y)$
marginal mass function of Y	$p_Y(y) = \sum_x p_{X,Y}(x,y)$
covariance of X and Y	Cov[X, Y], is the product moment $\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X\mu_Y$
X and Y are said to be uncorrelated	$\sigma_{XY} = 0.$
correlation coefficient	$\rho = \sigma_{XY}/\sigma_X\sigma_Y$ for which $-1 \leq \rho \leq +1.$
Normal Distribution (with parameters $\mu > 0, \sigma^2 > 0$)	probability density function $\phi(x;\, \mu,\, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty.$
standard normal distribution $\mu = 0, \sigma^2 = 1$	$\phi(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}, \qquad -\infty < \xi < \infty,$ **Cdf:** $\Phi(x) = \int_{-\infty}^{x} \phi(\xi)\, d\xi, \qquad -\infty < x < \infty.$
The joint normal (or bivariate normal) distribution for random variables X, Y	$\phi_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2} Q(x, y)\right\}, \quad -\infty < x, y < \infty.$
	$Q(x, y) = \frac{1}{1-\rho^2}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}.$
Moments $-1 < \rho < 1$	$E[X] = \mu_X, \qquad E[Y] = \mu_Y,$ $\mathrm{Var}[X] = \sigma_X^2, \qquad \mathrm{Var}[Y] = \sigma_Y^2.$ $\mathrm{Cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = \rho\sigma_X\sigma_Y.$

1 Introduction

correlation coefficient	ρ

2 Conditional Probability and Conditional Expectation

Term or Formula	Explanation
conditional probability mass function $p_{XY}(x\|y)$ of X given Y =y	$p_{XY}(x\|y) = \frac{\Pr\{X = x \text{ and } Y = y\}}{\Pr\{Y = y\}}$ $p_{X\|Y}(x\|y) = \frac{p_{XY}(x,y)}{p_Y(y)}$ if $p_Y(y) > 0; \quad x, y = 0, 1, \ldots.$
law of total probability " The marginal distribution of X"	$\Pr\{X = x\} = \sum_{y=0}^{\infty} p_{XY}(x\|y)p_Y(y).$
the conditional expected value of g(X) given Y = y	$E[g(X)\|Y = y] = \sum_x g(x)p_{XY}(x\|y) \qquad \text{if } p_Y(y) > 0,$
The law of total probability for conditional expectation	$E[g(X)] = \sum_y E[g(X)\|Y = y]p_Y(y).$ $E[g(X)] = E\{E[g(X)\|Y]\}.$
conditional probability density function $f_{XY}(x\|y)$ for the random variable X given that Y = y	$f_{XY}(x\|y) = \frac{f_{XY}(x, y)}{f_Y(y)} \qquad \text{if } f_Y(y) > 0,$
The conditional distribution function for X given Y = y	$F_{XY}(x\|y) = \int_{-\infty}^{x} f_{XY}(\xi\|y)\, d\xi \qquad \text{if } f_Y(y) > 0.$
the conditional expectation of g(X) given that Y = y	$E[g(X)\|Y = y] = \int g(x)f_{XY}(x\|y)\, dx \qquad \text{if } f_Y(y) > 0.$

2 Conditional Probability and Conditional Expectation

The law of total probability for conditional expectation	$$E[g(X)] = E\{E[g(X)\|Y]\} = \int E[g(X)\|Y = y] f_Y(y)\, dy.$$
conditional distribution function	random sum $X = \xi_1 + \cdots + \xi_N$ when $\xi_1, \xi_2, \ldots$ are continuous random variables, *conditional distribution function* $F_{X\|N}(x\|n)$ of the random variable X given that $N = n$ to be $$F_{X\|N}(x\|n) = \frac{\Pr\{X \le x \text{ and } N = n\}}{\Pr\{N = n\}} \quad \text{if } \Pr\{N = n\} > 0, \qquad (3.2)$$
conditional probability density function	*conditional probability density function* $f_{X\|N}(x\|n)$ for the random variable X given that $N = n$ by setting $$f_{X\|N}(x\|n) = \frac{d}{dx} F_{X\|N}(x\|n) \quad \text{if } \Pr\{N = n\} > 0. \qquad (3.3)$$ $$\Pr\{a \le X < b, N = n\} = \int_a^b f_{X\|N}(x\|n) p_N(n)\, dx$$ for $a < b$ and where $p_N(n) = \Pr\{N = n\}$. $$f_X(x) = \sum_{n=0}^{\infty} f_{X\|N}(x\|n) p_N(n).$$
The conditional expectation of g(X)	$$E[g(X)\|N = n] = \int g(x) f_{X\|N}(x\|n)\, dx.$$
the law of total probability	$$E[g(X)] = \sum_{n=0}^{\infty} E[g(X)\|N = n]\, p_N(n) = E\{E[g(X)\|N]\}.$$
The Moments of a Random Sum	Let us assume that ξ_k and N have the finite moments $$E[\xi_k] = \mu, \quad \text{Var}[\xi_k] = \sigma^2, \qquad E[N] = \nu, \quad \text{Var}[N] = \tau^2, \qquad (3.8)$$ and determine the mean and variance for $X = \xi_1 + \cdots + \xi_N$ as defined in (3.1). The derivation provides practice in manipulating conditional expectations, and the results, $$E[X] = \mu\nu, \qquad \text{Var}[X] = \nu\sigma^2 + \mu^2\tau^2, \qquad (3.9)$$
the probability density function for the fixed sum $\xi_1 + \xi_2 + \cdots + \xi_n$	the n-fold convolution of the density $f(z)$, denoted by $f^{(n)}(z)$ and recursively defined by $$f^{(1)}(z) = f(z)$$ and $$f^{(n)}(z) = \int f^{(n-1)}(z-u) f(u)\mathrm{d}u \quad \text{for } n > 1.$$

2 Conditional Probability and Conditional Expectation

marginal density function of X= $\xi_1 + \xi_2 + \cdots + \xi_n$	$$f_X(x) = \sum_{n=1}^{\infty} f^{(n)}(x) p_N(n).$$ given that N = n ≥ 1.
Law of Total Variance	$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y\|X)] + \mathrm{Var}(E[Y\|X])$$
stochastic process {X,,; n = 0, 1.... } is a Martingale	For n=0,1,..., (a) $E[\|X_n\|] < \infty$, and (b) $E[X_{n+1}\|X_0, \ldots, X_n] = X_n$.
	Extension: $$E[X_m\|X_0, \ldots, X_n] = X_n \quad \text{for } m \geq n.$$

3 Markov Chains: Introduction

<table>
<tr><th>Term or Formula</th><th>Explanation</th></tr>
<tr><td rowspan="2">A Markov process $\{X_t\}$ (Discrete)</td><td>$$\Pr\{X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i\} = \Pr\{X_{n+1} = j \mid X_n = i\}$$ for all time points n and all states $i_0, \ldots, i_{n-1}, i, j$.</td></tr>
<tr><td>$$\begin{aligned}&\Pr\{X_{n+1} = j_1, \ldots, X_{n+m} = j_m \mid X_0 = i_0, \ldots, X_n = i_n\}\\ &= \Pr\{X_{n+1} = j_1, \ldots, X_{n+m} = j_m \mid X_n = i_n\}\end{aligned} \qquad (3.9)$$ for all time points n, m and all states $i_0, \ldots, i_n, j_1, \ldots, j_m$. In other words, once (3.9) is established for the value $m = 1$, it holds for all $m \geq 1$ as well.</td></tr>
<tr><td>one-step transition probability</td><td>$$P_{ij}^{n,n+1} = \Pr\{X_{n+1} = j \mid X_n = i\}.$$</td></tr>
<tr><td>stationary transition probabilities</td><td>When the one-step transition probabilities are independent of the time variable n
Then $P_{ij}^{n,n+1} = P_{ij}$ is independent of n, and P_{ij} is the conditional probability that the state value undergoes a transition from i to j in one trial.</td></tr>
<tr><td>Markov matrix or transition probability matrix of the process</td><td>$$\mathbf{P} = \begin{Vmatrix} P_{00} & P_{01} & P_{02} & P_{03} & \cdots \\ P_{10} & P_{11} & P_{12} & P_{13} & \cdots \\ P_{20} & P_{21} & P_{22} & P_{23} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & P_{i3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}$$ $$P_{ij} \geq 0 \quad \text{for } i, j = 0, 1, 2, \ldots,$$ $$\sum_{j=0}^{\infty} P_{ij} = 1 \quad \text{for } i = 0, 1, 2, \ldots.$$</td></tr>
<tr><td></td><td>$$\begin{aligned}&\Pr\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}\\ &= p_{i_0} P_{i_0, i_1} \cdots P_{i_{n-2}, i_{n-1}} P_{i_{n-1}, i_n}.\end{aligned} \qquad (1.8)$$ This shows that all finite-dimensional probabilities are specified once the transition probabilities and initial distribution are given, and in this sense the process is defined by these quantities.
Let $\Pr\{X_0 = i\} = p_i$.</td></tr>
<tr><td>The n-step transition probabilities of a Markov chain</td><td>$$P_{ij}^{(n)} = \Pr\{X_{m+n} = j \mid X_m = i\}.$$ $$P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)},$$</td></tr>
</table>

3 Markov Chains: Introduction

	$P_{ij}^{(0)} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$
	$\mathbf{P}^{(n)} = \underbrace{\mathbf{P} \times \mathbf{P} \times \cdots \times \mathbf{P}}_{n \text{ factors}} = \mathbf{P}^n;$ (2.3) in other words, the n-step transition probabilities $P_{ij}^{(n)}$ are the entries in the matrix $\mathbf{P}^n$, the nth power of $\mathbf{P}$.
i is absorbing state	$P_{ii} = 1$
Absorbing Markov chain	A Markov chain is an absorbing chain if (1) there's at least one absorbing state (2) there's at least one transition from non-absorbing state to absorbing state
Time of absorption, (or hitting time, or reaching time)	$\boxed{T = \min\{n \geq 0,\ X_n = 0,\ \text{or } X_n = 2\}}$
The probability that the Markov chain ends in state 0, starting from state 1	$\boxed{u = \mathbb{P}\{X_T = 0 \mid X_0 = 1\}}.$ $\boxed{u = P_{10} + P_{11} u}$
The expected (mean) time for absorption, or mean hitting time, starting from state 1	$\boxed{v = \mathbb{E}[T \mid X_0 = 1]}.$ $\boxed{v = 1 + P_{11} v}$
The case of four state Markov chains	$u_i = \mathbb{P}\{X_T = 0 \mid X_0 = i\},\ \text{for } i = 1,\ 2,$ $u_1 = P_{10} + P_{11} u_1 + P_{12} u_2$ $u_2 = P_{20} + P_{21} u_1 + P_{22} u_2$

3 Markov Chains: Introduction

<table>
<tr><td></td><td>

$$v_i = \mathbb{E}[T|X_0 = i], \text{ for } i = 1,\ 2.$$

$$v_1 = 1 + p_{11} v_1 + p_{12} v_2$$

$$v_2 = 1 + p_{21} v_1 + p_{22} v_2$$

</td></tr>
<tr><td></td><td>

$$v_i = \mathbb{E}[T|X_0 = i],$$

I=0,1,2

$$v_0 = 1 + p_{00} v_0 + p_{01} v_1 + p_{02} v_2$$

$$v_1 = 1 + \cancel{p_{10} v_0} + p_{11} v_1 + p_{12} v_2$$

$$v_2 = 1 + \cancel{p_{20} v_0} + \cancel{p_{21}} v_1 + p_{22} v_2$$

</td></tr>
<tr><td>

A one-dimensional random walk

</td><td>

is a Markov chain whose state space is a finite or infinite subset $a, a + 1, \ldots, b$ of the integers, in which the particle, if it is in state i, can in a single transition either stay in i or move to one of the neighboring states $i - 1, i + 1$.

</td></tr>
<tr><td>

the transition matrix of a random walk

</td><td>

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & 2 & & i-1 & i & i+1 & & \end{array} \\ \left\| \begin{array}{ccccccccc} r_0 & p_0 & 0 & \cdots & 0 & \cdots & & & \\ q_1 & r_1 & p_1 & \cdots & 0 & \cdots & & & \\ 0 & q_2 & r_2 & \cdots & 0 & \cdots & & & \\ & & \ddots & & & & & & \\ & & 0 & & q_i & r_i & p_i & 0 & \\ & & & \ddots & & & & & \ddots \end{array} \right\| \end{array}, \tag{5.9}$$

where $p_i > 0$, $q_i > 0$, $r_i \geq 0$, and $q_i + r_i + p_i = 1$, $i = 1, 2, \ldots$ $(i \geq 1)$, $p_0 \geq 0$, $r_0 \geq 0$, $r_0 + p_0 = 1$. Specifically, if $X_n = i$, then for $i \geq 1$,

$$\Pr\{X_{n+1} = i + 1 | X_n = i\} = p_i,$$

$$\Pr\{X_{n+1} = i - 1 | X_n = i\} = q_i,$$

$$\Pr\{X_{n+1} = i | X_n = i\} = r_i,$$

with the obvious modifications holding for $i = 0$.

The designation "random walk" seems apt, since a realization of the process describes the path of a person (suitably intoxicated) moving randomly one step forward or backward.

</td></tr>
</table>

3 Markov Chains: Introduction

Gambler Ruin	player A Starting with fortune \$ i ثروة / مبلغ / الرصيد player B \$ $N-i$, N is the total amount $u_i = \Pr\{X_n \text{ reaches state } 0 \text{ before state } N \mid X_0 = i\}$ $= \begin{cases} \dfrac{N-i}{N} & \text{when } p = q = \frac{1}{2}, \\ \dfrac{(q/p)^i - (q/p)^N}{1-(q/p)^N} & \text{when } p \neq q. \end{cases}$ **where X,, is player A's fortune at stage n**

4 The Long Run Behavior of Markov Chains

<table>
<tr><th>Term or Formula</th><th>Explanation</th></tr>
<tr><td>Markov chain, is called regular</td><td>when raised to some
power k, the matrix Pk has all of its elements strictly positive</td></tr>
<tr><td>limiting probability distribution</td><td>$\pi = (\pi_0, \pi_1, \ldots, \pi_N)$,
where $\pi_j > 0$ for $j = 0, 1, \ldots, N$ and $\Sigma_j \pi_j = 1$, and this distribution is independent of the initial state.
or, in terms of the Markov chain $\{X_n\}$,
$$\lim_{n\to\infty} \Pr\{X_n = j \mid X_0 = i\} = \pi_j > 0 \qquad \text{for } j = 0, 1, \ldots, N.$$
Theorem 1.1 Let P be a regular transition probability matrix on the states 0, 1, ..., N. Then the limiting distribution $\pi = (\pi_0, \pi_1, \ldots \pi_N)$ is the unique nonnegative solution of the equations
$$\pi_j = \sum_{k=0}^{N} \pi_k P_{kj}, \qquad j = 0, 1, \ldots, N, \tag{1.2}$$
$$\sum_{k=0}^{N} \pi_k = 1. \tag{1.3}$$</td></tr>
<tr><td></td><td>$$\mathbf{P} = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-a & a \\ 1 & b & 1-b \end{array}, \qquad \text{where } 0 < a, b < 1,$$
$$\mathbf{P}^n = \frac{1}{a+b}\begin{Vmatrix} b & a \\ b & a \end{Vmatrix} + \frac{(1-a-b)^n}{a+b}\begin{Vmatrix} a & -a \\ -b & b \end{Vmatrix}.$$
$$\lim_{n\to\infty} \mathbf{P}^n = \begin{Vmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{Vmatrix}.$$</td></tr>
<tr><td>Long run mean cost per unit time</td><td>$$= \sum_{j=0}^{N} \pi_j c_j.$$
π_j also gives the long run mean fraction of time that the process $\{X_n\}$ is in state j. Thus if each visit to state j incurs a “cost” of c_j, then the long run mean cost per unit time associated with this Markov chain is</td></tr>
</table>

5 Poisson Processes

Term or Formula	Explanation
The Poisson distribution	with parameter $\mu > 0$ is given by $$p_k = \frac{e^{-\mu}\mu^k}{k!} \quad \text{for } k = 0, 1, \ldots.$$
	E(X)=μ , VAR(X)=$\sigma_X{}^2 = \mu$
Poisson process	**Definition** *A Poisson process of intensity, or rate, $\lambda > 0$ is an integer-valued stochastic process $\{X(t); t \geq 0\}$ for which* (i) *for any time points $t_0 = 0 < t_1 < t_2 < \cdots < t_n$, the process increments* $$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$ *are independent random variables;* (ii) *for $s \geq 0$ and $t > 0$, the random variable $X(s + t) - X(s)$ has the Poisson distribution* $$\Pr\{X(s+t) - X(s) = k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad \textit{for } k = 0, 1, \ldots;$$ (iii) $X(0) = 0$.
	In particular, observe that if $X(t)$ is a Poisson process of rate $\lambda > 0$, then the moments are $$E[X(t)] = \lambda t \quad \text{and} \quad \mathrm{Var}[X(t)] = \sigma^2_{X(t)} = \lambda t.$$
nonhomogeneous or nonstationary Poisson process	$$\lambda = \lambda(t)$$ If $X(t)$ is a nonhomogeneous Poisson process with rate $\lambda(t)$, then an increment $X(t) - X(s)$, giving the number of events in an interval $(s, t]$, has a Poisson distribution with parameter $\int_s^t \lambda(u)\,du$, and increments over disjoint intervals are independent random variables.

5 Poisson Processes

<table>
<tr><td rowspan="2">A compound Poisson process</td><td>Given a Poisson process $X(t)$ of rate $\lambda > 0$, suppose that each event has associated with it a random variable, possibly representing a value or a cost. Examples will appear shortly. The successive values $Y_1, Y_2, \ldots$ are assumed to be independent, independent of the Poisson process, and random variables sharing the common distribution function
$$G(y) = \Pr\{Y_k \le y\}.$$
A *compound Poisson process* is the cumulative value process defined by
$$Z(t) = \sum_{k=1}^{X(t)} Y_k \qquad \text{for } t \ge 0. \tag{6.1}$$</td></tr>
<tr><td>Consider the compound Poisson process $Z(t) = \sum_{k=1}^{X(t)} Y_k$. If $\lambda > 0$ is the rate for the process $X(t)$ and $\mu = E[Y_1]$ and $\nu^2 = \mathrm{Var}[Y_1]$ are the common mean and variance for $Y_1, Y_2, \ldots$, then the moments of $Z(t)$ can be determined from the random sums formulas of II, Section 3.2 and are
$$E[Z(t)] = \lambda\mu t; \qquad \mathrm{Var}[Z(t)] = \lambda(\nu^2 + \mu^2)t. \tag{6.2}$$</td></tr>
<tr><td>A marked Poisson process</td><td>A marked Poisson process is the sequence of pairs $(W_1, Y_1), (W_2, Y_2), \ldots$, where $W_1, W_2, \ldots$ are the waiting times or event times in the Poisson process $X(t)$.
Both compound Poisson and marked Poisson processes appear often as models of physical phenomena.</td></tr>
</table>

6 Continuous Time Markov Chains

Term or Formula	Explanation
continuous time, discrete state, Markov processes	$\{X(t); 0 \le t < \infty\}$ where the possible values of $X(t)$ are the nonnegative integers. the transition probability function for $t > 0$, $$P_{ij}(t) = \Pr\{X(t+u) = j \mid X(u) = i\}, \qquad i, j = 0, 1, 2, \ldots,$$ is independent of $u \ge 0$.
pure birth process	Consider a sequence of positive numbers, $\{\lambda_k\}$. We define a pure birth process as a Markov process satisfying the following postulates: (1) $\Pr\{X(t+h) - X(t) = 1 \mid X(t) = k\} = \lambda_k h + o_{1,k}(h)$ $(h \to 0+)$. (2) $\Pr\{X(t+h) - X(t) = 0 \mid X(t) = k\} = 1 - \lambda_k h + o_{2,k}(h)$. (1.1) (3) $\Pr\{X(t+h) - X(t) < 0 \mid X(t) = k\} = 0$ $(k \ge 0)$. As a matter of convenience we often add the postulate (4) $X(0) = 0$.
	With this postulate $X(t)$ does not denote the population size but, rather, the number of births in the time interval $(0, t]$.
$P_n(t)$	We define $P_n(t) = \Pr\{X(t) = n\}$, assuming $X(0) = 0$.
	$$P_0(t) = e^{-\lambda_0 t},$$ $$P_1(t) = \lambda_0\left(\frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} + \frac{1}{\lambda_0 - \lambda_1} e^{-\lambda_1 t}\right), \qquad (1.7)$$
	$$P_n(t) = \Pr\{X(t) = n \mid X(0) = 0\} = \lambda_0 \cdots \lambda_{n-1}[B_{0,n}e^{-\lambda_0 t} + \cdots + B_{n,n}e^{-\lambda_n t}] \quad \text{for } n > 1, \qquad (1.8)$$ where $$B_{0,n} = \frac{1}{(\lambda_1 - \lambda_0)\cdots(\lambda_n - \lambda_0)},$$ $$B_{k,n} = \frac{1}{(\lambda_0 - \lambda_k)\cdots(\lambda_{k-1} - \lambda_k)(\lambda_{k+1} - \lambda_k)\cdots(\lambda_n - \lambda_k)} \qquad (1.9)$$ for $0 < k < n$, and $$B_{n,n} = \frac{1}{(\lambda_0 - \lambda_n)\cdots(\lambda_{n-1} - \lambda_n)}.$$
pure death proces	Complementing the increasing pure birth process is the decreasing pure death process. It moves successively through states $N, N-1, \ldots, 2, 1$ and ultimately is absorbed in state 0 (extinction).

6 Continuous Time Markov Chains

pure death proces	Alternatively, we have the infinitesimal description of a pure death process as a Markov process $X(t)$ whose state space is 0, 1, ..., N and for which (i) $\Pr\{X(t+h) = k-1 \mid X(t) = k\} = \mu_k h + o(h),\ k = 1, \ldots, N;$ (ii) $\Pr\{X(t+h) = k \mid X(t) = k\} = 1 - \mu_k h + o(h),\ k = 1, \ldots, N;$ (2.1) (iii) $\Pr\{X(t+h) > k \mid X(t) = k\} = 0,\ k = 0, 1, \ldots, N.$ The parameter μ_k is the "death rate" operating or in effect while the process sojourns in state k. It is a common and useful convention to assign $\mu_0 = 0$.
$P_n(t)$	When the death parameters $\mu_1, \mu_2, \ldots, \mu_N$ are distinct, that is, $\mu_j \neq \mu_k$ if $j \neq k$, then we have the explicit transition probabilities $$P_N(t) = e^{-\mu_N t};$$ and for $n < N$, $$\begin{aligned} P_n(t) &= \Pr\{X(t) = n \mid X(0) = N\} \\ &= \mu_{n+1}\mu_{n+2}\cdots\mu_N[A_{n,n}e^{-\mu_n t} + \cdots + A_{N,n}e^{-\mu_N t}], \end{aligned} \tag{2.2}$$ where $$A_{k,n} = \frac{1}{(\mu_N - \mu_k)\cdots(\mu_{k+1} - \mu_k)(\mu_{k-1} - \mu_k)\cdots(\mu_n - \mu_k)}.$$

8 Brownian Motion and Related Processes

Term or Formula	Explanation
Brownian motion process	**is an example of a continuous-time, continuous-state-space Markov process**
	Let $B(t)$ be the y component (as a function of time) of a particle in Brownian motion. Let x be the position of the particle at time t_0; i.e., $B(t_0) = x$. Let $p(y, t\|x)$ be the probability density function, in y, of $B(t_0 + t)$, given that $B(t_0) = x$.
	Since $p(y, t\|x)$ is a probability density function in y, we have the properties $$p(y, t\|x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} p(y, t\|x)\, dy = 1. \qquad (1.1)$$
	$$\lim_{t \to 0} p(y, t\|x) = 0 \quad \text{for} \quad y \neq x.$$
	$p(y, t\|x)$ must satisfy the partial differential equation $$\frac{\partial p}{\partial t} = \tfrac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2}. \qquad (1.3)$$ This is called the *diffusion equation*, and σ^2 is the *diffusion coefficient*,
Brownian motion with diffusion coefficient σ^2	**Definition** Brownian motion with diffusion coefficient σ^2 is a stochastic process $\{B(t);\, t \geq 0\}$ with the properties: (a) Every increment $B(s + t) - B(s)$ is normally distributed with mean zero and variance $\sigma^2 t$; $\sigma^2 > 0$ is a fixed parameter.

8 Brownian Motion and Related Processes

	(b) For every pair of disjoint time intervals $(t_1, t_2]$, $(t_3, t_4]$, with $0 \le t_1 < t_2 \le t_3 < t_4$, the increments $B(t_4) - B(t_3)$ and $B(t_2) - B(t_1)$ are independent random variables, and similarly for n disjoint time intervals, where n is an arbitrary positive integer. (c) $B(0) = 0$, and $B(t)$ is continuous as a function of t.
standard Brownian motion	for a standard Brownian motion ($\sigma^2 = 1$), we have $\Pr\{B(s+t) \le y \mid B(s) = x\} = \Pr\{B(s+t) - B(s) \le y - x\}$ $= \Phi\left(\frac{y-x}{\sqrt{t}}\right).$
covariance of the Brownian motion	$\operatorname{Cov}[B(s), B(t)] = \sigma^2 \min\{s, t\}, \qquad \text{for } s, t \ge 0.$