Stochastic Modelling Cheatsheet

Probability

Countability

A is called **countable** if A is finite or it's possible to enumerate A by N.

Countable: N, Z, Q

Uncountable: R, Q^c

Expectation

Linear: $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$

Monotone: if $X \leq Y$ then $E[X] \leq E[Y]$

Tail Sum Formula

If X is nonnegative, Continuous r.v.:

$$E[X] = \int_0^\infty [1 - F(x)]dx = \int_0^\infty P(X > x)dx$$

Discrete r.v.:

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

Moment Generating Function

$m_x(t) = E[e^{tX}] \quad t \in R$

Distribution of X is determined by mgf provided expectation is finite in a neighborhood of the origin.

$$\frac{d^n}{dt^n} m_x(0) = E[X^n]$$

X, Y independent, $Z = X + Y$ is determined by:

$m_Z(t) = E[e^{tZ}] = E[e^{tX}] \times E[e^{tY}] = m_X(t) \times m_Y(t)$

Probability Generating Function

$P_X(t) = E[t^X] \quad |t| \leq 1$

$P'_X(1) = E[X]$, $P''_X(1) = E[X(X-1)]$

Conditional Probability

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

$$P_{X|Y} = P(X = x|Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)} = \frac{p_{XY}(x, y)}{p_Y(y)}$$

$E\{g(X)\} = E[E\{g(X)|Y\}]$

Mixed Case (X continuous, N discrete):

$$F_{X|N}(x|n) = \frac{P(X \leq x \cap N = n)}{P(N = n)}$$

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

Random Sums

$X = \sum_{n=1}^{N} \xi_n$, ξ_n's are i.i.d and $\xi_n \perp N$

If $E[\xi_n] = \mu$, $Var[\xi_n] = \sigma^2$, $E[N] = \nu$, $Var[n] = \tau^2$,

$E[X] = \mu\nu$, $Var[X] = \mu^2\tau^2 + \nu\sigma^2$

$m_x(t) = E[m_\xi^N(t)] = P_N(m_\xi(t))$

if $a < 0 < b$:

$$P(a < X \leq b) = \int_a^b \{f^{(n)}(x) p_N(n)\}dx + p_N(0)$$

if $a < b < 0$ or $0 < a < b$:

$$P(a < X \leq b) = \int_a^b \{f^{(n)}(x) p_N(n)\}dx$$

Markov Chain

Definition

A **Markov Chain** $X\{t\}$ is a stochastic process such that simultaneously

1. Countable state space S
2. Discrete time $T = N_0 = 0, 1, 2, \ldots$
3. Satisfies Markov property
4. Satisfies stationary transition property

Markov Property

$P(X_{t+1}|X_0 = x_0, \ldots, X_t = x_t) = P(X_{t+1}|X_t = x_t)$

Stationary Transition Property

$P(X_{t+1}|X_t = x) = P(X_1|X_0 = x)$

Absorbing State

State a is absorbing whenever

$P(a, a) = 1$ and $P(a, y) = 0$ for $y \neq a$

Recurrence and Transience

Hitting Time

$T_A = inf\{\ t > 0 : X_t \in A\}$

First Step Analysis

$$\begin{aligned} U_{ik} &:= P(\text{Absorbed in state k}) \\ &= P_{ik} + \sum_{j=0}^{r-1} U_{jk} P_{ij} \end{aligned}$$

$$\begin{aligned} w_i &:= \sum_{n=0}^{T_A - 1} g(X_n) \\ &= g(i) + \sum_{j=0}^{r-1} P_{ij} w_j \end{aligned}$$

$$\begin{aligned} v_i &:= E_i[T_A] \\ &= 1 + \sum_{j=0}^{r-1} P_{ij} v_j \end{aligned}$$

Hitting Probability

$\rho_{xy} := P_x(T_y < \infty) = P(T_y < \infty | X_0 = x)$

$$\rho_{xC} = \sum_{z \in C} P(x, z) + \sum_{z \in S_T} P(x, z)\rho_{zC}$$

A state y is called

- Recurrent whenever $\rho_{yy} = 1$
- Transient whenever $\rho_{yy} < 1$

Recurrent states need not to be absorbing. Absorbing states are recurrent.

Markovian Martingale Property

$$\sum_{y \in S} yP(x, y) = x \text{ for all x} \in S$$

For a Markovian Martingale with state Space $S = 0, 1, \ldots, d$, and only state 0 and d are absorbing. We have $\rho_{xd} = x/d$

Branching Process

Offspring $\xi, \xi_k, \xi_{k,t+1}$ are i.i.d r.v.'s

$X_{t+1} = \sum_{k=1}^{X_t} \xi_{k,t+1}$

$P(x, y) = P\{\sum_{k=1}^{x} \xi_k = y\}$

Probability generating function: $P_\xi(u) = E[u^\xi]$

Possibility of extinction:

if $P(\xi = 1) = 1$ then $\rho_{01} = 0$

if $P(\xi = 1) < 1$ and $E[\xi] = P'_\xi(1) \leq 1$ then $\rho_{01} = 1$

if $E[\xi] > 1$ then $\rho_{01} < 1$,

ρ_{10} is the smallest nonnegative solution to $u = P_\xi(u)$

Stationary Distribution

1. π is a valid pmf
2. $\pi P = \pi$

Existence of an unique π i.f.f. an **irreducible positive recurrent** chain

Steady State Distribution

If exists a steady state distribution π :

1. $\pi = lim_{t \to \infty} \pi_0 P^t$ for all initial state π_0
2. Existence: π is a stationary distribution
3. Uniqueness: π is the only stationary distribution

Not the other way around: Unique stationary distribution does not guarantee a steady state distribution.

Birth & Death Chain

$$P(x,y) = \begin{cases} q_x, & if\ y = x-1 \\ r_x, & if\ y = x \\ p_x, & if\ y = x+1 \\ 0, & otherwise \end{cases}$$

$irreducible \Leftrightarrow q_x, p_x > 0\ for\ all\ x > 0\ and\ p_0 > 0$

$\gamma_0 = 1,\ \gamma_y = \frac{\prod_{i=1}^{y} q_i}{\prod_{i=1}^{y} p_i} = \prod_{i=1}^{y} \frac{q_i}{p_i}$

$P_x(T_a < T_b) = \frac{\sum_{y=x}^{b-1} \gamma_y}{\sum_{z=a}^{b-1} \gamma_z}$

$\rho_{10} = \lim_{b \to \infty} P_1(T_0 < T_b) = 1 - 1/\sum_{z=0}^{\infty} \gamma_z$

$Recurrent \Leftrightarrow \sum_{z=0}^{\infty} \gamma_z = \infty$

$\delta_0 := 1$ and $\delta_y := \frac{\rho_0 \times ... \times \rho_{y-1}}{q_1 \times ... \times q_y}$

Stationary distribution exists: $\sum_{y=0}^{\infty} \delta_y < \infty$ $(p < 1/2)$

$\pi(x) = \frac{\delta_x}{\sum_{y=0}^{\infty} \delta_y}$

Stationary distribution = S.S.D. only when $r \neq 0$ (aperiodic)

Positive/Null Recurrent

$\left\{ \begin{matrix} Positive \\ Null \end{matrix} \right\}$ recurrent if $m_y = E_y[T_y] = \left\{ \begin{matrix} < \infty \\ = \infty \end{matrix} \right\}$

If S is finite:

- S has at least one positive recurrent state
- S has no null recurrent state

The existence and uniqueness of stationary distribution

An **irreducible and positive recurrent** chain has an unique stationary distribution π with $\pi(x) = 1/m_y$

Null positive recurrent chain doesn't have a stationary distribution.

The proportion of time spent in state y is $\frac{\rho_{xy}}{m_y}$

Period of a State

$x \in S$ $I_x = \{n \geq 1 :\ P^n(x,x) > 0\}$

$d_x :=$ g.c.d of I_x

d_x is the period of state x

if $x \leftrightarrow y$ then $d_x = d_y$

Aperiodic: An irreducible Markov chain with $d_x = 1$ for all $x \in S$

Regular Chain

For some $n \in N$, P^n has strongly positive entires:

$P^n(x,y) > 0$ for all $x, y \in S$

$d = d_x = 1$

Existence of Steady State Distribution

When a Markov chain is:

1. irreducible
2. positive recurrent
3. aperiodic

Pure Jump Process

1. State space S is countable
2. Sample paths are **right-continuous**

Time of explosion: $\tau_\infty := \lim_{n \to \infty} \tau_n$

Pure Jump Markov Process

$P_{x_0}\{X(s_1) = x_1, ..., X(s_n) = x_n)\} = \prod_{i=0}^{n-1} P_{x_i, x_{i+1}}(s_{i+1} - s_i)$

Embedded Markov Chain

Let $\tau_0 = 0 < \tau_1 \leq \tau_2 \leq \tau_n < \infty$ be the holding times

Let $X_0 \neq X_1, X_1 \neq X_2, ..., X_{n-1}$ be the associated states

$q_x = 1/E_x(\tau_1)$, $q_{xx} = -q_x$

$Q_{xy} = P_x\{X(\tau_1) = y\}$ (X non-absorbing)

$q_{xy} = q_x Q_{xy}$

$\{X_n\}$ is a Markov chain with initial distribution π_0 and trasition matrix Q

As long as $\tau_n < \infty$,

$$X(t) = \sum_{k-0}^{n} X_k 1_{[\tau_k, \tau_{k+1})}(t)$$

Chapman-Kolmogorov Equation

$P_{xy} = \sum_{z \in S} P_{xz}(t) P_{zy}(s)$

Forward and Backward Equations

$$P'_{xy}(t) = \sum_{z \in S} q_{xz} P_{zy}(t)$$

$$P'_{xt}(t) = \sum_{z \in S} P_{xz}(t) q_{zy}$$

Birth & Death Process

The Condition of Explosion

$$\tau_\infty = lim_{n \to \infty} \tau_n < \infty \Leftrightarrow \sum_{x=0}^{\infty} \frac{1}{\lambda_x} + \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} \frac{\mu_x ... \mu_{x-y+1}}{\lambda_x ... \lambda_{x-y}}$$

$\Rightarrow$ no explosion for sublinear birth rates

Stationary distribution:

$$\delta_0 = 0, \delta_y = \frac{\lambda_0 ... \lambda_y - 1}{\mu_1 ... \mu_y}$$

$$\pi(x) = \frac{\delta_x}{\sum_{y=0}^{\infty} \delta_y}$$

$X(t)$ transient $\Leftrightarrow X_n$ transient $\Leftrightarrow \sum_{y=1}^{\infty} \frac{\mu_1 ... \mu_y}{\lambda_1 ... \lambda_y} < \infty$

Linear ODE

The solution to $f'(t) = -cf(t) + g(t), t \geq 0$ is

$f(t) = f(0)e^{(-ct)} + \int_0^t e^{-c(t-s))} g(s) ds$

Poisson Process

Homogeneous

$X(0) = 0$

$P_{xy}(t) = e^{-\lambda t}(\lambda t)^{y-x}/(y-x)!$, if x $\leq$ y

$X(t) - X(s) \sim Poisson\{\lambda(t-s)\}$

$\tau_k - \tau_{k-1} \sim Exponential(\lambda)$

mgf $m_X(s) = E[e^{sX}] = exp(\lambda_X(e^t - 1))$

Superposition

If $X \sim Poisson(\lambda_X)$ and $Y \sim Poisson(\lambda_Y)$, X, Y independent, then $X + Y \sim Poisson(\lambda_X + \lambda_Y)$

Thinning

If $\xi_1, \xi_2, ... \overset{i.i.d}{\sim} Bernoulli(p)$ then $\sum_{l=1}^{X} \xi_k \sim Poisson(p\lambda_X)$

Inhomogeneous

Homogeneous: $\lambda_X = \lambda > 0$

Inhomogeneous: $\lambda_X(t) \geq 0$,

$\Lambda(t) = E[X(t)] = \int_0^t \lambda_X(u) du$

$X(t) - X(s) \sim Poisson\{\Lambda_X(t) - \Lambda_X(s)\}$

Satisfies Markov property, but not stationary transition property

Compound Poisson Process

Let $X(t) \sim Poisson(\lambda)$,

Y_n are i.i.d r.v.'s with $G(y) = P(Y_n \leq y)$

$$Z(t) = \sum_{k=1}^{X(t)} Y_k$$

Let $E[Y_n] = \mu$ and $Var(Y_n) = \sigma^2$ Then

$E[Z(t)] = \lambda\mu t$ and $Var[Z(t)] = \lambda(\mu^2 + \sigma^2)t$

$$F_{Z(t)}(t) = e^{(-\lambda t)} 1_{[0,\infty)}(x) + \sum_{n=1}^{\infty} \frac{e^{(-\lambda t)}(\lambda t)^n}{n!} G^{(n)}(x)$$

Failure Time (critical value a):

$\{F > t\} = \{Z(t) < a\}$, $P(F > t) = F_{Z(t)}(a)$

$E[F] = \int_0^\infty P(F > t) dt = \lambda^{-1} \sum_{n=0}^{\infty} G^{(n)}(a)$

If $Y_n \sim Exponential(\mu)$, then

$E[F] = \frac{1 + \mu a}{\lambda}$

Long Run Behaviour

Solve $\pi q = 0$.

$$\pi(x) = \begin{cases} \frac{1}{q_x m_x} & x \in C \\ 0 & x \notin C \end{cases}$$

C is the union of positive recurrent communication classes.

Irreducible positive recurrent process $\Rightarrow$ steady state distribution exists.

Infinite Server Queue

$\lambda_x = \lambda$, $\mu_x = \mu x$
$\pi(x) = \frac{(\lambda/\mu)^x}{x!} e^{-\lambda/\mu}$

Finite(N-server) Server Queue

$$\lambda_x = \lambda, \mu_x = \begin{cases} x\mu & a \le x < N \\ N_\mu & x \ge N \end{cases}$$

Gaussian Process

Univariate Gaussian Distribution

$X \sim Normal(\mu, \sigma^2)$
mgf $m_X(T) = exp(\mu t + \sigma^2 t^2/2)$

Bivariate Gaussian Distribution

$X = (X_1, X_2)' \sim N(\mu, \Sigma)$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$$

mgf $m_X(T) = exp(\mu' t + t'\Sigma' t/2)$

Multivariate Gaussian Distribution

pdf exist $\Leftrightarrow det\Sigma_x \neq 0$ Then

$$f(x) = \frac{1}{(2\pi)^{d/2}\sqrt{det(\Sigma)}} exp(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu))$$

Conditional Distribution: Bivariate

$$E[X_1|X_2] = E[x_1] + \frac{Cov(X_1,X_2)}{Var(X_2)}(X_2 - E[X_2])$$

$$Var(X_1|X_2) = Var(X_1) - \frac{Cov(X_1,X_2)^2}{Var(X_2)}$$

$X_1|X_2 = x_2 \ N(E[X_1|X_2 = x_2], Var(X_1|X_2 = x2))$

Conditional Distribution: Multivariate

$$\mu_X = (\mu_X', \mu_Y')', \Sigma_Z = \begin{pmatrix} \Sigma_X & C_{XY}' \\ C_{XY} & \Sigma_Y \end{pmatrix}$$

Find b s.t. $\Sigma_Y b' = C_{XY}$
$\mu_{X|Y=y} = E[X] + b(y - E[Y])$
$\Sigma_{X|Y=y} = \Sigma_X - bC_{XY}$
$X|Y = y \sim N(\mu_{X|Y=y}, \Sigma_{X|Y=y})$

Gaussian Process

$\{X(t)\}_{t\in[0,\infty]}$, for all d, $t_1, ...t_d \in [0,\infty)$,
$(X(t_1), ..., X(t_d))$ is a d-dimensional Gaussian vector.

Brownian Motion

$\{W(t) : t \in [0,\infty)\}, W(0) = 0$
For $t \ge s \ge 0, W(t) - W(s) \sim N(0, \sigma^2(t-s))$
Increments of non-overlapping time intervals are independent
$r_W(s,t) = Cov(W(s), W(t)) = \sigma^2 min\{s,t\}$

Brownian Bridge

$B = \{B(t) : 0 \le t \le 1\}$
$\mu_B(t) = 0,\ r_B(s,t) = \sigma^2\{min\{s,t\} - st\}$

Geometric Brownian Motion

$S(t) := s_0 e^{\alpha t + \sigma W(t)}$
If $s_0 = 1$, $\alpha = 0$, and $\sigma = 1$, then $\mu_S(t) = e^{t/2}$ and
$r_S(s,t) = e^{(s+t)/2}(e^{min\{s,t\}} - 1)$
Log-normal distribution of r.v. X:
$E[X] = exp(\mu + \sigma^2/2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} exp(-\frac{(\ln x - \mu)^2}{2\sigma^2})$$

Reflection Principle

$P(W(t) > 0) = P(W(t) < 0) = 1/2$
$P(W(t) > a) = P(T_a \le t)/2$
$F_{T_1}(t) = P(|W(t)| > a) = 2(1 - \Phi(a/\sqrt{\sigma^2 t}))$
Let $M(t) := max_{0\le s\le t} W(s)$, $M_t \overset{D}{=} |W(t)|$

Brownian Martingale

One of NASC: $\mu_X(t) \equiv const$

Donsker's Theorem

Let $\xi, \xi_1, \xi_2, ...$ be i.i.d. r.v.'s with $E[\xi] = 0$ and $Var(\xi) = \sigma^2$
Let $S_n = \sum_{k=1}^n \xi_k$. For $\triangle_t, \triangle_x > 0$ define:
$X_\triangle(t) := X_{\triangle_t,\triangle_x}(t) := \triangle_x S_{[t/\triangle_t]}$
$\mu_{X_\triangle}(t) = \triangle_x E(S_{[t/\triangle_t]}) \equiv 0$
$r_{X_\triangle}(t) = \triangle_x^2 E(S_{[t/\triangle_t]} S_{[s/\triangle_t]}) = \triangle_x^2 \sigma_2 min\{[t/\triangle_t], [s/\triangle_t]\}$

take $\triangle_x := \sqrt{\triangle_t}$

Theorem: $X_\triangle \overset{D}{\to} W \qquad \triangle_t = \triangle_x^2 \downarrow 0$

CLT: $\triangle_t = 1/n$

$$P(\frac{S_n}{\sqrt{n}} \le x) = P(X_\triangle(1) \le x) \approx P(W(1) \le x) = \Phi(\frac{x}{\sigma})$$

Heat Equation

$$\frac{\partial}{\partial t} f(t,x) = \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2} f(t,x)$$

Boundary Crossing with Drift

$$P(T_b^\mu < T_a^\mu) = \frac{exp(-2\mu a/\sigma^2) - 1}{exp(-2\mu a/\sigma^2) - exp(-2\mu b/\sigma^2)}$$

Integrated Brownian Motion

$$Y(s) = \int_0^s W(u)du$$

Gaussian(the limit of a linear combination of normal r.v.)
$\mu_Y = E[Y(s)] = E[\int_0^s W(u)du] = \int_0^s \mu_W(u)du = 0$
$r_Y(s,t) = Cov\{Y(s), Y(t)\} = \frac{s^2(3t-s)}{6}, s \le t$
Not Markovian $\Rightarrow Cov\{Y(t) - Y(s), Y(s) = s^2(t-s)/2 \neq 0\}$
$Cov\{Y(s), W(t)\} = smin\{s,t\} - (min\{s,t\}^2/2)$

White Noise

$$Y = \int_a^b g(t)W'(t)dt = \int_a^b g(t)dW(t)$$

$$= g(b)W(b) - g(a)W(a) - \int_a^b g'(t)W(t)dt$$

White noise: $W'(t)$ or $dW(t)$
Gaussian
$E[y] = 0$
$r_Y(s,t) = \sigma^2 \int_0^{min\{s,t\}} g^2(u)du$
$r_Y(t,t) = \sigma^2 \int_0^t g^2(u)du$
if $g(t) = 1$, then $Y(t) = W(t)$