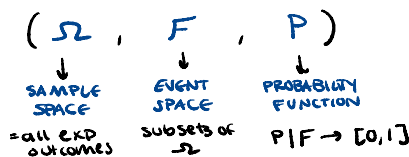


PROBABILITY SPACES



AXIOMS

1. NONNEGATIVITY: $P(E) \geq 0 \quad \forall E \in \mathcal{F}$
2. UNIT MEASURE: $P(\Omega) = 1$
3. COUNTABLE ADDITIVITY: $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$
 $\uparrow$ For disjoint events E_i

CONDITIONALITY

BAYES RULE = $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}$
 $\uparrow$ POSTERIOR $\uparrow$ EVIDENCE $\uparrow$ PRIOR

TOTAL PROBABILITY RULE $P(A) = \sum_i P(A \cap B_i)$
 $= \sum_i P(A|B_i)P(B_i)$

INDEPENDENCE

* Knowing one event occurred does NOT change probability of another event

$P(A|B) = P(A)$ $P(A \cap B) = P(A)P(B)$

INCLUSION / EXCLUSION

$|A \cup B| = |A| + |B| - |A \cap B|$
 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 $P(\bigcup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k+1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right)$

UNION BOUND $P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$

RANDOM VARIABLES

function from sample space (Ω) to subset $\mathbb{R}$
 $X: \Omega \rightarrow \mathbb{R}$

DISCRETE random variables have PROBABILITY MASS FUNCTIONS

$P_X(x) = P(X=x) = \sum_{\omega \in \Omega} \mathbb{1}_{\{X(\omega)=x\}} P(\omega)$
 ① Find all outcomes that give rise to $X=x$
 ② Add permutations to get $P_X(x)$

EXPECTATION = weighted avg of a RV
 $E[X] = \sum_x x P_X(x)$ **MEAN**

① LINEAR: $E[aX+b] = aE[X] + b$
 $E[X+Y] = E[X] + E[Y]$
 ② LOTUS: $E[g(X)] = \sum_x g(x) P_X(x)$

③ CONDITIONAL EXPECTATION: $E[X|Y=y] = \sum_x x p_{X|Y}(x|y)$

④ TOTAL EXPECTATION: $E[X] = \sum_y P_Y(y) E[X|Y=y]$

IMPORTANT DISCRETE RVs

NAME	PMF	EXPECTATION	VARIANCE	WHEN TO USE
Unif(a,b)	$P(X=i) = \frac{1}{b-a+1}$	$E[X] = \frac{a+b}{2}$	$Var(X) = \frac{(b-a+1)^2 - 1}{12}$	DICE ROLL
Bernoulli(p)	$P(X=0) = 1-p$ $P(X=1) = p$	$E[X] = p$	$Var(X) = p(1-p)$	SINGLE COIN FLIP
Binomial(n,p)	$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$	$E[X] = np$	$Var(X) = np(1-p)$	# of heads in n flips
Geometric(p)	$P(X=i) = (1-p)^{i-1} p$	$E[X] = \frac{1}{p}$	$Var(X) = \frac{1-p}{p^2}$	# of tosses UNTIL head
Poisson(λ)	$P(X=i) = \frac{\lambda^i}{i!} e^{-\lambda}$	$E[X] = \lambda$	$Var(X) = \lambda$	RARE EVENT in some interval

IMPORTANT DENSITIES

NAME	PDF	CDF	$E[X]$	$Var(X)$
Unif(a,b)	$f_X(x) = \frac{1}{b-a}$	$F_X = \frac{x-a}{b-a}$	$E[X] = \frac{a+b}{2}$	$Var(X) = \frac{(b-a)^2}{12}$
Exp(λ)	$f_X(x) = \lambda e^{-\lambda x}$	$F_X = 1 - e^{-\lambda x}$	$E[X] = \frac{1}{\lambda}$	$Var(X) = \frac{1}{\lambda^2}$
$N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$B(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$	$E[X] = \mu$	$Var(X) = \sigma^2$

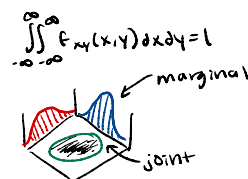
CONTINUOUS RANDOM VARS

CDF = cumulative density function $\rightarrow F_X(x) = P(X \leq x)$ * RV is less than or equal to x
 PDF = probability density function $\rightarrow f_X(x) = \frac{d}{dx} F_X(x)$ * RV is equal to specific value
*** Integrate PDF to get CDF**

JOINT DENSITIES

probability that x & y take on specific values simultaneously

$P((X,Y) \in A) = \iint_A f_{X,Y}(x,y) dx dy$
 JOINT PDF
 MARGINAL: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
 $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$



EXPECTATION: $E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$

$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ * relating JOINT to CONDITIONAL

LAW OF TOTAL EXPECTATION

$E[X] = E[E[X|Y]]$ * cannot just treat RV as a constant

$E[E[X|Y]] = \begin{cases} \sum_y E[X|Y=y] P_Y(y) & \text{DISCRETE} \\ \int_{-\infty}^{\infty} E[X|Y=y] f_Y(y) dy & \text{CONT.} \end{cases}$

LAW OF TOTAL VARIANCE

$Var(X) = E[Var(X|Y)] + Var(E[X|Y])$

MEMORYLESS PROPERTY

$P(X > a+b | X > a) = P(X > b)$

VARIANCE = measure of spread

$Var(X) = E[(X - E[X])^2]$

$Var(X) = E[X^2] - E[X]^2$

$X \sim \text{EXP}(\lambda)$
 * FOR EXPONENTIAL
 * P of event happening in next ϵ min is SAME regardless of how long you're waited

PROOF: $P(X > a+b, X > a) = \frac{P(X > a+b)}{P(X > a)}$
 $= \frac{P(X > a+b)}{P(X > a)} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} = e^{-\lambda b} = P(X > b)$

$Var(aX+b) = a^2 Var(X)$

CONVOLUTION

Let $Z = X + Y$ for independent X & Y
DISCRETE: $P(Z=z) = \sum_{k=-\infty}^{\infty} P(X=k)P(Y=z-k)$

CONTINUOUS: $f_z(z) = (f_x * f_y)(z) = \int_{-\infty}^{\infty} f_x(x)f_y(z-x)dx$

COVARIANCE = when 1 RV changes, does the other change the same

$COV(X,Y) = E[(X-E[X])(Y-E[Y])]$
 $= E[XY] - E[X]E[Y]$

PROPERTIES

- 1) $COV(X,Y) = COV(Y,X)$
- 2) $COV(aX+bY, cZ) = ac COV(X,Z) + bc COV(Y,Z)$
- 3) X & Y indep $\rightarrow COV(X,Y) = \emptyset$
* If $COV(X,Y) = \emptyset \rightarrow$ not necessarily indep
ADD PROOF HERE??

4) $COV(X,X) = Var(X)$
 $\rho(X,Y) = \frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}$

$Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) + \sum_{\{i,j\} \neq \{i,i\}} 2COV(X_i,X_j)$

ORDER STATISTICS

$X_1 \dots X_n$ be iid RVs w/ $f_X(x)$ and $F_X(x)$
 k^{th} order statistic
 $F_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F_X(x)]^{k-1} [1-F_X(x)]^{n-k} f_X(x)$

$k=1$ smaller than x
 $k=n$ larger than x
Exactly 1 item = x

MIN/MAX TRICKS

$\max(X,Y) \rightarrow$ both must be below
 $P(\max(X,Y) \leq t) = P(X \leq t \text{ and } Y \leq t)$
 $\min(X,Y) \rightarrow$ both have NOT happened
 $P(\min(X,Y) > t) = P(X > t \text{ and } Y > t)$
 $\rightarrow P(\min > t) = [1 - F_X(t)] \cdot [1 - F_Y(t)]$

SURVIVAL FUNCTION

$S(t) = P(X > t) = 1 - F(X \leq t)$

TAIL SUM = $E[X] = \int_0^{\infty} P(X > t) dt$

TOTAL EXPECTATION EXAMPLE

Ex) $Z \sim N(\mu, \sigma^2)$ $A \sim \text{UNIF}(0, 2^3)$ $E[Z^2 + \text{var}(Z) \cdot E[A^2]]$
 $E[A] = E[E[A|Z]] = E[\frac{0+Z^3}{2}] = \frac{1}{2} E[Z^3]$
 $= \frac{1}{2} (\sigma^3 + \mu^3)$

HELPFUL CALCULUS

$\frac{d}{dx} \int_{a(x)}^{b(x)} f(t) dt = f(b(x))b'(x) - f(a(x))a'(x)$

Taylor series: $e^u = \sum_{x=0}^{\infty} \frac{u^x}{x!}$

$\int_a^b f(t) dt = F(b) - F(a)$

chain rule: $\frac{d}{dx} f(g(x)) = f'(g(x))g'(x)$

product rule: $\frac{d}{dx} (f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

TO complete square for $x^2 + bx \rightarrow$
add + subtract $(\frac{b}{2})^2$
to get $(x - \frac{b}{2})^2$

MOMENT GENERATING FUNCTION

$M_X(t) = E[e^{tx}] \rightarrow \begin{cases} M(s) = \sum_{x=-\infty}^{\infty} e^{sx} P_X(x) & \text{DISCRETE} \\ M(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx & \text{CONTINUOUS} \end{cases}$

JOINT MGF: $M_{\vec{X}}(\vec{t}) = E[e^{\vec{t}^T \vec{X}}]$ where $\vec{t}^T \vec{X} = t_1x_1 + t_2x_2 + \dots + t_nx_n$

* TAKE DERIVATIVE OF TRANSFORM TO EXTRACT MOMENTS

$\left. \frac{d}{ds} M(s) \right|_{s=0} = \int_{-\infty}^{\infty} x f_X(x) dx = E[X]$ first moment

$\left. \frac{d^n}{ds^n} M(s) \right|_{s=0} = \int_{-\infty}^{\infty} x^n f_X(x) dx = E[X^n]$ n^{th} moment formula

* SUM OF VARIABLES = PRODUCT OF MGF

$M_{X+Y}(s) = E[e^{s(X+Y)}] = E[e^{sX}] E[e^{sY}] = M_X(s) M_Y(s)$

* AFFINE TRANSFORMS

$M_{aX+b}(s) = e^{bs} M_X(as)$

* 2 INDEP VARIABLES

$M_{X,Y}(s,t) = M_X(s) M_Y(t)$

IMPORTANT TRANSFORMS

BERNOULLI	$M_X(s) = 1 - p + pe^s$	NORMAL	$M_X(s) = e^{\frac{\sigma^2 s^2}{2} + \mu s}$
BINOMIAL	$M_X(s) = (1 - p + pe^s)^n$	EXPONENTIAL	$M_X(s) = \frac{\lambda}{\lambda - s}$
GEOMETRIC	$M_X(s) = \frac{pe^s}{1 - (1-p)e^s}$	POISSON	$M_X(s) = e^{\lambda(e^s - 1)}$

BINOMIAL MGF TRANSFORM

$Z = X_1 + \dots + X_n$ where X_i binomial
 $M_Z(s) = M_{X_1}(s) \dots M_{X_n}(s)$
 $= (1 - p + pe^s)^n \leftarrow$ ALSO BINOMIAL

POISSON TRANSFORM

$M_X(s) = e^{\lambda(e^s - 1)}$
 $Z = X + Y$
 $M_Y(s) = e^{\lambda(e^s - 1)}$
 $M_Z(s) = M_X(s) M_Y(s)$
 $= e^{\lambda(e^s - 1)} \cdot e^{\lambda(e^s - 1)}$
 $= e^{2\lambda(e^s - 1)}$

NORMAL TRANSFORM

$X_1 \sim N(\mu_1, \sigma_1^2)$ $Y = X_1 + X_2$
 $X_2 \sim N(\mu_2, \sigma_2^2)$
 $M_Y(s) = M_{X_1}(s) M_{X_2}(s)$
 $= e^{\frac{\mu_1 s + \sigma_1^2 s^2}{2}} \cdot e^{\frac{\mu_2 s + \sigma_2^2 s^2}{2}}$
 $= e^{\frac{(\mu_1 + \mu_2)s + (\sigma_1^2 + \sigma_2^2)s^2}{2}}$

INDICATOR TRICKS

* When looking for count of something
 $\rightarrow$ decompose into sums of indicator (Bernoulli) variables

$X = \# \text{ of H in } n \text{ coin flips}$
 $1_i = \begin{cases} 1 & \text{if flip } i \text{ is heads} \\ 0 & \text{o/w} \end{cases}$ probability of event
 $X = \sum_{i=1}^n 1_i$ * $E[1_i] = p$
* $Var(1_i) = p(1-p)$
 $E[X] = E[\sum_{i=1}^n 1_i] = \sum_{i=1}^n E[1_i]$
 $= \sum_{i=1}^n \frac{1}{2} \rightarrow p(n)$

CONCENTRATION BOUNDS

MARKOV INEQUALITY: (if X nonnegative values)
 $P(X \geq a) \leq \frac{E[X]}{a}$
 for all $a > 0$

Chebyshev Inequality: (mean μ , var σ^2)
 $P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$
 for all $c > 0$

WLLN: sample mean of large # of iid RV is very close to true mean
 Sample mean: $M_n = \frac{X_1 + \dots + X_n}{n}$

Strong LLN: $P(|M_n - \mu| \geq \epsilon) = P(|\frac{X_1 + \dots + X_n}{n} - \mu| \geq \epsilon) \rightarrow 0$ as $n \rightarrow \infty$

* Bulk of dist near μ

CONVERGENCE

almost sure $\rightarrow$ in probability $\rightarrow$ in distribution

almost surely $P(\lim_{n \rightarrow \infty} X_n = X) = 1$ $X_n \xrightarrow{a.s.} X$

$\rightarrow$ SLLN: if $(X_n)_{n=1}^{\infty}$ iid w/ finite mean then sample mean $\bar{X}_n$ converges a.s. to true mean

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{a.s.} E(X_1)$$

$\rightarrow$ BOREL CANTELLI LEMMA

let $(A_n)_{n=1}^{\infty}$ be collection events

A_n i.o. = $\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$

- ① If $\sum_{n=1}^{\infty} P(A_n) < \infty \rightarrow P(A_n \text{ i.o.}) = 0$ (ie $X_n \xrightarrow{a.s.} X$)
- ② If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and $(A_n)_{n=1}^{\infty}$ are indep then $P(A_n \text{ i.o.}) = 1$ ie $X_n \not\xrightarrow{a.s.} X$

Convergence in probability: $\lim_{n \rightarrow \infty} P(|Y_n - a| \geq \epsilon) = 0$ $X_n \xrightarrow{P} X$
 Y_n converges to a in probability
 $P(|Y_n - a| \geq \epsilon) \leq \delta$
 ϵ = accuracy level
 δ = confidence level

Convergence in distribution: $\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x)$

* does not imply convergence in probability
 Standard score of sample mean converges in distribution to normal distribution

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \xrightarrow{d} N(0, 1)$$

- ① standardize RV by subtracting mean & dividing variance
- ② Apply CLT

$\rightarrow$ POISSON LIMIT THEOREM $X_n \sim \text{Binomial}(n, p_n)$
 $\lim_{n \rightarrow \infty} np_n = \lambda > 0$
 $X \sim \text{Poisson}(\lambda) \rightarrow X_n \xrightarrow{d} X$

$\rightarrow$ LIMIT OF GEOMETRIC DISTRIBUTION $X_n \sim \text{Geom}(p_n)$
 $\lim_{n \rightarrow \infty} \frac{p_n}{n} = \lambda > 0$

$X \sim \text{Exponential}(\lambda) \rightarrow X_n \xrightarrow{d} X$

ENTROPY = expected surprise = expected uncertainty

$$H(X) = \sum_{x \in X} P_X(x) \log \frac{1}{P_X(x)} = E\left[\log \frac{1}{P_X(X)}\right]$$

* BASE 2 * concave
 * nonnegative since $P_X(x) \in [0, 1] \rightarrow \frac{1}{P_X(x)} \geq 1 \rightarrow \log \frac{1}{P_X(x)} \geq 0$

$$H(X, Y) = \sum_{x, y \in X \times Y} P_{X,Y}(x, y) \log \frac{1}{P_{X,Y}(x, y)}$$

CONDITIONAL ENTROPY

$$H(Y|X) = \sum_{x \in X} P_X(x) \sum_{y \in Y} P_{Y|X}(y|x) \log \frac{1}{P_{Y|X}(y|x)}$$

$$= \sum_{x \in X} P_X(x) H(Y|X=x)$$

* info gained from given we have x

CHAIN RULE

$$H(X, Y) = H(X) + H(Y|X)$$

$$* H(X, Y) = H(Y, X) * H(X, Y) \leq H(X)$$

$$* H(X, Y) \leq H(X) + H(Y) \text{ equality iff } X, Y \text{ independent}$$

SOURCE CODING

mapping from symbols x in source alphabet X to bit strings
 $l(x)$ = length of binary string description of x

source coding theorem

$$\lim_{n \rightarrow \infty} E\left[\frac{1}{n} l(X_1 \dots X_n)\right] \leq H(X) + \epsilon$$

typical set for sequence of random vars: set of sequences w/ probability $\approx 2^{-nH(X)} = 2^{E \log P_X(X)^n}$

ϵ -typical set: $2^{-n(H(X)+\epsilon)} \leq P_{X^n}(X_1 \dots X_n) \leq 2^{-n(H(X)-\epsilon)}$

$\rightarrow$ size $\approx 2^{nH(X)}$ # of typical sequences $\approx 2^{nH(X)}$

AEP = asymptotic equipartition property

if $(X_n)_{n=1}^{\infty}$ iid over finite alphabet: $-\frac{1}{n} \log_2 P_X(X_1 \dots X_n) \xrightarrow{a.s.} H(X)$

$$P(2^{-n(H(X)+\epsilon)} < P_{X^n}(X^n) < 2^{-n(H(X)-\epsilon)}) \rightarrow 1$$

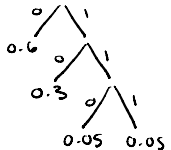
* we only need to care abt encoding sequences in typical set

of bits required to encode set of size S : $\lceil \log_2 S \rceil$

of bits to assign code to each sequence in typical set $\leq n(H(X) + \epsilon)$

HUFFMAN CODING

= optimal way to code random variable X in prefix-free way



$m=4$ values
 $P = (0.6, 0.3, 0.05, 0.05)$
 codewords $w = (0, 10, 110, 111)$
 lengths $l = (1, 2, 3, 3)$

example prefix free code

Huffman coding: frequent symbols get short codes
 rare symbols get long codes

* achieves MIN possible expected length L

$$* E[l] = \sum p_i l_i \quad H(X) \leq L \leq H(X) + 1$$

$\rightarrow$ determined by tree depth

* If min expected length for single RV

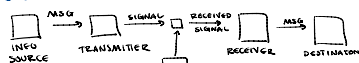
$\rightarrow$ Use Huffman

* If min expected length for sequences of RV $\rightarrow$ AEP/Entropy

- ① list all probabilities
- ② pick 2 smallest
- ③ Merge them
Parent = $P_{min_1} + P_{min_2}$
- ④ Treat parent as new symbol & repeat process
- ⑤ Assign bits (0 $\rightarrow$ left, 1 $\rightarrow$ right)

SHANNON CODING

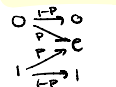
- ① src msg compressed
- ② redundancy added to deal w/ channel noise
- ③ coded msg sent through channel



BINARY SYMMETRIC CHANNEL



BINARY ERASURE CHANNEL



if using BEC channel n times w/ erasure prob p :

- ① info-theoretically max length $m = Cn$ bits
- ② max rate w/ Shannon scheme
 $\text{Error} \leq 2^{-n\epsilon}$
 $m = (1-p-\epsilon)n$ bits

CAPACITY: $C = 1 - H(p)$
 achieved when $P(X) = \text{uniform}$

$$\text{CAPACITY: } C = 1 - p$$

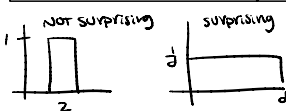
GENERAL FORMULA:
 $C = \max_{P(X)} I(X; Y)$

MUTUAL INFORMATION

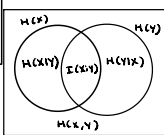
$$I(X; Y) = H(X) - H(X|Y)$$

$$= H(X) + H(Y) - H(X, Y)$$

$$= H(Y) - H(Y|X)$$



* $H(X)$ maximized by uniform distribution



KL DIVERGENCE

* How diff are 2 distributions

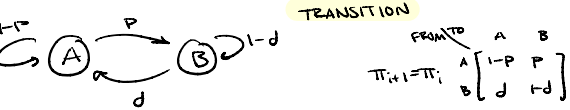
$$D(P||Q) = \sum P(x) \log \left(\frac{P(x)}{Q(x)} \right)$$

$I(X; Y) = D_{KL}(P_{X,Y} || P_X P_Y)$

cross entropy

$$- \sum P(x) \log(Q(x)) = H(P) + D_{KL}(P||Q)$$

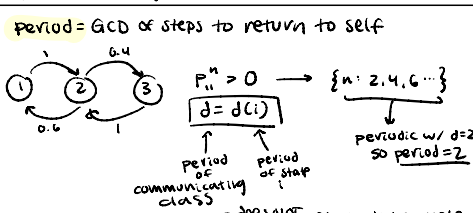
DISCRETE TIME MARKOV CHAINS



STATIONARY DISTRIBUTION = steady state
 $\pi^* = P\pi^*$
Solve: $\pi^* P = \pi^*$
 $\sum \pi_i = 1$ (Normalization)
Balance equations: $\sum \pi_i P_{ij} = \pi_j$ * FLOW IN = FLOW OUT
Use if chain has cycles

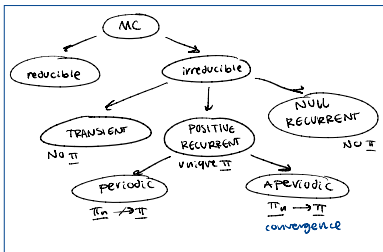
DETAILED BALANCE: $\pi_i P_{ij} = \pi_j P_{ji}$ * $i \rightarrow j$ flow is same as $j \rightarrow i$ flow
Use for birth-death chains

MARKOV PROPERTY = $P(X_{n+1}=j | X_n=i, X_{n-1}, \dots, X_0) = P(X_{n+1}=j | X_n=i) = P_{ij}$
* future only depends on present state



$d=1 \rightarrow$ APERIODIC
 $d>1 \rightarrow$ PERIODIC
CLASS PROPERTY

GENERAL FSE for DTMC $X = \{1, 2, \dots, N\}$
 $T_A = \min \{n \geq 0 | X(n) \in A\}$ $A \subset X$
Mean hitting time: $f_A(i) = E[T_A | X_0=i]$

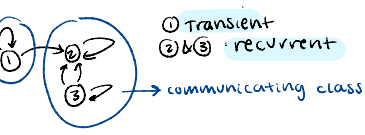


RANDOM WALK ON GRAPH
States = nodes of the graph
 $P_{ij} = \begin{cases} \frac{1}{d_i} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$ d_i = degree of node i
Undirected connected graph stationary dist
* ALWAYS REVERSIBLE
 $\pi_i = \frac{d_i}{2|E|}$
Handshaking lemma: sum of all degrees is 2 * edges

CLASSIFICATION

irreducible: access any state from any state
reducible: CANNOT access any state from any state (multiple communicating classes)

TRANSIENT: if given we start at state i there's nonzero probability we never return to that state
opposite is recurrent: no chance you will not return always chance of return



if MC is irreducible AND recurrent then
positive recurrent: $E[T_i | X_0=i] < \infty$ aka expected time to return is FINITE
means a unique stationary dist exists
 $\pi_i = \frac{1}{E[T_i]}$

null recurrent: $E[T_i | X_0=i] = \infty$ aka expected time to return is INFINITE
does NOT have a stationary dist
ex) start @ 0 and move left/right w/ p=0.5

HITTING TIME = how many time steps to reach certain state
starting at state A: how many time steps to reach E
 T_E is hitting time to hit E from A
USE FIRST STEP EQUATIONS

Mean hitting time: $f_E(A) = E[T_E | X_0=A]$
 $f(i) = E[T_E | X_0=i]$
 $f(A) = 1 + \frac{1}{2} f(B) + \frac{1}{2} f(C)$
 $f(B) = 1 + f(C)$
 $f(C) = 1 + f(A)$
 $f(C) = 1 + \frac{1}{2} f(A) + \frac{1}{2} f(B) + \frac{1}{2} f(C)$
 $f(C) = 0$
probability of flow to other state

ACCUMULATING REWARDS
collect reward until you hit hitting time
 $g(i) = E[\sum_{n=0}^{T-1} g(X_n) | X_0=i]$ $\forall i \in X$
 $g(i) = \begin{cases} g(i) & \text{if } i \in A \\ g(i) + \sum_{j \in A} P_{ij} g(j) & \text{if } i \notin A \end{cases}$
reward function value at that state

condition for reversibility
detailed balance: $\pi(i) P_{ij} = \pi(j) P_{ji}$
Equations

$\pi(x) P(x,y) = \pi(y) P(y,x)$ for all $x,y \in X$ is REVERSIBLE
forward transition: $P_{ij} = P(X_{n+1}=j | X_n=i)$
backward transition: $Q_{ji} = P(X_n=i | X_{n+1}=j)$
* reversible if backward transition = forward transition
 $Q_{ji} = \frac{\pi_i P_{ij}}{\pi_j}$

METROPOLIS HASTINGS
generate samples from target distribution
 $\pi(x) \rightarrow$ construct MC so stationary dist is exactly $\pi(x)$
current state is $X_t = i$
1. Pick candidate next state j from proposal dist $q(i,j)$
2. calculate acceptance probability: $\alpha(i,j) = \min(1, \frac{\pi(j) q(j,i)}{\pi(i) q(i,j)})$
3. Accept/reject by flipping biased coin w/ P(heads) = $\alpha(i,j)$
heads $\rightarrow X_{t+1} = j \rightarrow$ move to new state
tails $\rightarrow X_{t+1} = i \rightarrow$ stay where you are
 $P(i,j) = g(i,j) \alpha(i,j) = g(i,j) \min\{\frac{\pi(j) q(j,i)}{\pi(i) q(i,j)}, 1\}$
SATISFIES DETAILED BALANCE
 $\pi P = \pi P$
 $\pi P_{ij} = \frac{d_j}{2|E|} \frac{1}{d_i} = \frac{1}{2|E|}$
 $\pi_j P_{ji} = \frac{d_i}{2|E|} \frac{1}{d_j} = \frac{1}{2|E|}$
DBE
 $\pi_i P_{ij} = \pi_j P_{ji}$

BIG THEOREM: if MC is irreducible, aperiodic, and positive recurrent:
THEN:
1. Unique stationary π exists
2. π is unique soln to $\pi P = \pi$ (and $\sum \pi_i = 1$)
3. $\lim_{n \rightarrow \infty} P(X_n=j) = \pi_j$
4. $E[T_i] = \frac{1}{\pi_i} \rightarrow$ expected time to return

POISSON PROCESS

$P(k, \tau)$ = prob of k arrivals in interval of duration τ

N_τ : # of arrivals w/ rate λ over interval τ

PMF: $P_{N_\tau}(k) = P(k, \tau) = e^{-\lambda\tau} \frac{(\lambda\tau)^k}{k!}$ $k=0,1,\dots$

$E[N_\tau] = \lambda\tau$
 $\text{Var}(N_\tau) = \lambda\tau$ * POISSON DIST

TIME T until first arrival

PDF: $\lambda e^{-\lambda t}$ $t \geq 0$
 ECT: $\frac{1}{\lambda}$
 $\text{Var}(T) = \frac{1}{\lambda^2}$ * EXPONENTIAL DIST

k^{th} arrival time: $Y_k = T_1 + T_2 + \dots + T_k$

k^{th} interarrival time: $T_k = Y_k - Y_{k-1}$ where $T_1 = Y_1$

$E[Y_k] = ECT_1 + \dots + ECT_k = \frac{k}{\lambda}$
 $\text{Var}(Y_k) = \text{Var}(T_1) + \dots + \text{Var}(T_k) = \frac{k}{\lambda^2}$

PDF Y_k : $f_{Y_k}(y) = \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!}$ $y \geq 0$
 Erlang Dist

MEMORYLESSNESS for interarrival times

$P(T > s + t | T > s) = P(T > t) = e^{-\lambda t}$

SPLITTING: each arrival kept w/ prob p discarded w/ prob $1-p$

$\hookrightarrow$ **PPCAP**

MERGING: 2 indep PP $\rightarrow$ PP(λ_1) & PP(λ_2) & merging whenever arrival occurs in either process

$\hookrightarrow$ **PP($\lambda_1 + \lambda_2$)**

$P(\text{arrival from first process}) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ $P(\text{arrival from second process}) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$

RANDOM INCIDENCE PARADOX: if you arrive at bus stop at random time: avg length of interval you fall into is DOUBLE the avg interarrival time

$E[\text{backward}] = \frac{1}{\lambda}$ $E[\text{forward}] = \frac{1}{\lambda}$

$E[L] = E[LB] + E[FB] = \frac{2}{\lambda}$

CTMC

= transition rates instead of probabilities

q_{ij} = rate for transition $i \rightarrow j$

Rate Matrix: $Q \in \mathbb{R}^{(N+1) \times (N+1)}$

$Q = \begin{bmatrix} -\sum_{j=1}^N \lambda_{0j} & \lambda_{01} & \dots & \lambda_{0N} \\ \lambda_{10} & -\sum_{j=0,2}^N \lambda_{1j} & \dots & \lambda_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N0} & \lambda_{N1} & \dots & -\sum_{j=0}^{N-1} \lambda_{Nj} \end{bmatrix}$

$\begin{cases} q_{ij} \geq 0 \quad \forall i \neq j \text{ (OFF DIAGONAL)} \\ \sum_j q_{ij} = 0 \quad \leftarrow \text{row sums to 0} \end{cases}$

$q_i = \sum_{j=1, j \neq i}^N \lambda_{i \rightarrow j} = -Q(i,i)$

HOLDING TIME: time in state i before moving

$T_i \sim \text{Exp}(q_i) \rightarrow E[T_i] = \frac{1}{q_i}$ where $q_i = -Q(i,i)$

STATIONARY DISTRIBUTION solve ① $\pi Q = 0$
 ② $\sum_i \pi_i = 1$

DETAILED BALANCE EQN $\pi_i q_{ij} = \pi_j q_{ji}$

TRANSITION PROBABILITIES: $P(t) = e^{tQ}$

- KOLMOGOROV FORWARD EQN: $\frac{dP}{dt} = P(t)Q$

BACKWARD EQN: $\frac{dP}{dt} = QP(t)$

EMBEDDED JUMP CHAIN = modelling CTMC as DTMC

$P_{ij} = \frac{q_{ij}}{q_i}$ * does NOT necessarily have same stationary dist
 $P_{ii} = 0$ $q_i = -Q(i,i)$

HITTING TIME EQN: reach state A from state i :

$M_i = \frac{1}{q_i} + \sum_{j \notin A} P_{ij} M_j$

UNIFORMIZATION: constructing DTMC w/ same stationary as its CTMC

$P = I + \frac{1}{q} Q$ add self-loop where necessary $\rightarrow P_{ii} = 1 - \frac{q_i}{q} \leftarrow \max_i q_i$ (max of diagonals)

RANDOM GRAPHS

$G(n, p)$

$\uparrow$
 n vertices

$\leftarrow$ $p \in [0, 1]$
 $\binom{n}{2}$ edges present w/ probability p

$E[M] = \# \text{ of edges} = \binom{n}{2} p$

D degree = $D \sim \text{Binomial}(n-1, p)$
 $P(D=d) = \binom{n-1}{d} p^d (1-p)^{n-1-d}$

$E[D] = (n-1)p$

$P(\text{any vertex isolated}) = (1-p)^{n-1}$

Expected number of triangles $\rightarrow \binom{n}{3}$ possible triangles

$E[X_{\text{triangles}}] = \binom{n}{3} p^3$

$E[X_{\text{isolated vertices}}] = n(1-p)^{n-1}$

MLE

= value of param θ that MAX $P(x_1, \dots, x_n | \theta)$ over all θ

DISCRETE $\hat{\theta}_n = \arg \max_{\theta} P_n(x_1, \dots, x_n | \theta)$

CONTINUOUS $\hat{\theta}_n = \arg \max_{\theta} f_n(x_1, \dots, x_n | \theta)$

$\rightarrow$ observations x_i independent:
 likelihood function: $P_n(x_1, \dots, x_n | \theta) = \prod_{i=1}^n P(x_i | \theta)$

* if θ appears in bounds of distribution
 $x \sim \text{Unif}(0, \theta) \rightarrow$ use $1_{\{x_i \leq \theta\}} \rightarrow \hat{\theta}_{MLE} = \max(x_1, \dots, x_n)$

$x \sim \text{Unif}(0, b) \rightarrow \hat{\theta}_{MLE} = \min(x_1, \dots, x_n)$

MAP

= select value $\hat{\theta}$ that maximizes θ over the posterior dist

GOAL: $\hat{\theta}_{MAP} = \arg \max_{\theta} P(\theta | x) = \arg \max_{\theta} \frac{P(x | \theta) P(\theta)}{P(x)}$

= $\arg \max_{\theta} P(x | \theta) P(\theta)$

* log, take derivative wrt θ , solve

$\hat{\theta}_{MAP}(\theta) = \hat{\theta}_{MLE}(\theta) + \ln \pi(\theta)$

$\lim_{n \rightarrow \infty} \hat{\theta}_{MAP} = \hat{\theta}_{MLE}$ $\hookrightarrow$ MLE + penalty term

HYPOTHESIS TESTING

decision that maximizes PCD subject to PFA

H_0 = null hypothesis = no fire = $f(x | H_0)$
 H_1 = alternative hypothesis = fire = $f(x | H_1)$

Likelihood Ratio: $L(x) = \frac{f_{H_1}(x)}{f_{H_0}(x)}$

Rejection region: data that leads you to REJECT null
 $R: \{x: L(x) > k\}$

PFA = Type I error = $\alpha = P(\text{Reject } H_0 | H_0 \text{ true}) = P(L(x) > k | H_0)$

PCD = Power of the test = $P(\text{Reject } H_0 | H_1 \text{ true}) = P(L(x) > k | H_1)$

① Find likelihood ratio

② Look at behavior of $L(x)$

- strictly increasing $\rightarrow$ isolate $x \rightarrow \Delta(x) > k \rightarrow x > t(k)$ (right tail)
- symmetric $\rightarrow \Delta(x) > k \leftrightarrow |x| > c$ (both tails)
- strictly decreasing $\rightarrow$ isolate x & FLIP sign $\rightarrow \Delta(x) > k \leftrightarrow x < c$ (left tail)

③ Find distribution ASSUMING H_0 is true

④ Solve for exact critical value using α

$\hat{x} = \begin{cases} 1 & \text{if } L(x) > \lambda \\ 0 & \text{if } L(x) \leq \lambda \end{cases} \quad (\text{w/p } \gamma)$

$\rightarrow$ randomization

G_0 on n vertices w/ m edges

$P(G_0 = G) = p^m (1-p)^{\binom{n}{2} - m}$

Phase transitions

$p < \frac{1}{n}$: graph is forest of small components $\rightarrow$ degree $\rightarrow 0$
 $p = \frac{1}{n}$: critical threshold $\rightarrow$ $n \rightarrow \infty$
 $p > \frac{1}{n}$: giant component of size $O(n)$ emerges $\rightarrow$ $n \rightarrow \infty$
 $p \gg \frac{1}{n}$: completely connected w/ high probability $\rightarrow$ $n \rightarrow \infty$

(if $p(n) = \frac{\lambda \ln n}{n}$ $\lambda < 1$ then $P(G(n, p(n)) \text{ is connected}) \rightarrow 0$
 $\lambda > 1$ then $P(G(n, p(n)) \text{ is connected}) \rightarrow 1$)

① Write likelihood: $\mathcal{L}(\theta | x) = P(x | \theta) = \prod_{i=1}^n f(x_i | \theta)$

② take log: $\log(\mathcal{L}(\theta | x)) = \log(P(x | \theta)) = \sum_{i=1}^n \log f(x_i | \theta)$

③ Maximize: $\hat{\theta}_{MLE} = \arg \max_{\theta} \log(P(x | \theta))$
 $\hookrightarrow$ take deriv wrt $\theta \rightarrow$ set to 0

USE = linear least squares estimator

* Goal: minimize $E\|X - bY\|^2$

$$E\|X - bY\|^2 = E\|X\|^2 + \frac{\text{cov}(X,Y)}{\text{var}(Y)}(Y - E[Y])$$

$$E\|Y - bX\|^2 = E\|Y\|^2 + \frac{\text{cov}(X,Y)}{\text{var}(X)}(X - E[X])$$

$$\text{cov}(X,Y) = E\|XY\| - E[X]E[Y]$$

* In projection sense: $X - E\|X - bY\|$ orthogonal

to every linear function of Y

MMSE = minimum mean squares estimator

Goal: finding $g(Y)$ that minimizes $E\|X - g(Y)\|^2$

$$g(Y) = E\|X\|$$

Error is $E\|\text{var}(X|Y)\|$

$$\textcircled{1} \text{ Find marginal PDF: } f_Y(y) = \int f_{X,Y}(x,y) dx$$

$$\textcircled{2} \text{ find cond PDF: } f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$\textcircled{3} \hat{x}_{\text{mmse}} = E\|X|Y=y\| = \int x f_{X|Y}(x|y) dx$$

MMSE = USE WHEN JOINTLY GAUSSIAN

HILBERT SPACE = treating RV in vector space

$$\textcircled{1} \text{ Inner Product: } \langle X,Y \rangle = E\|XY\|$$

$$\textcircled{2} \text{ Norm (length): } \|X\| = \sqrt{\langle X,X \rangle} = \sqrt{E\|X^2\|}$$

$$\|X - Y\| = \sqrt{E\|(X - Y)^2\|}$$

$$\textcircled{3} \text{ Orthogonality: } \langle X,Y \rangle = 0 \iff E\|XY\| = 0$$

* Shortest distance from a point to a subspace is straight line ORTHOGONAL to that subspace

$$\langle X - \hat{X}, Z \rangle = 0$$

↑ $\hat{X}$ every vector inside space H

$$\textcircled{4} \text{ Zero mean shortcuts } E\|X\| = E\|Y\| = 0$$

$$\text{inner product: } \langle X,Y \rangle = \text{cov}(X,Y)$$

$$\text{norm squared: } \|X\|^2 = \text{var}(X)$$

$$\text{orthogonality: } \langle X,Y \rangle = 0 \iff \text{cov}(X,Y) = 0$$

USE = projecting X onto straight line/plane

$$\hat{X} = \text{Proj}_Y(X) = \frac{\langle X,Y \rangle}{\|Y\|^2} Y = \frac{\text{cov}(X,Y)}{\text{var}(Y)} Y$$

MMSE = projecting onto curved surface

$$\hat{X} = E\|X|Y\|$$

JOINTLY GAUSSIAN

For jointly Gaussian RV $X = [X_1, X_2, \dots, X_n]^T$

$$\text{Joint PDF: } f(X) = \frac{1}{(2\pi)^n \det(\Sigma)} \exp\left(-\frac{1}{2}(X - \mu)^T \Sigma^{-1}(X - \mu)\right)$$

$$\Sigma_{ij} = \text{cov}(X_i, X_j) \quad \Sigma = AA^T$$

RULES

$$\textcircled{1} \text{ INDEPENDENT } \iff \text{UNCORRELATED}$$

$$X_i, X_j \text{ indep } \iff \text{cov}(X_i, X_j) = 0$$

$$\textcircled{2} \text{ LINEAR TRANSFORMATIONS ALSO GAUSSIAN}$$

$$Y = AX + b \rightarrow Y \sim N(A\mu + b, A\Sigma A^T)$$

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n \rightarrow N(\mu_Y, \sigma_Y^2)$$

$$\textcircled{3} \text{ If set of variables is JOINTLY GAUSSIAN}$$

↳ any subset is also JOINTLY GAUSSIAN

* If X is Gaussian & Y is Gaussian $\rightarrow$ does NOT guarantee joint Gaussian system

$$X \sim N(0,1)$$

$$Y \sim BX \text{ where } B \sim \text{Bernoulli}(\frac{1}{2})$$

$$\textcircled{4} \text{ If jointly Gaussian, conditional dist ALSO Gaussian}$$

$$X|Y = y \sim N(\mu_{X|Y}, \Sigma_{X|Y})$$

CONCENTRATION INEQUALITIES

Markov: If X can take only nonnegative: $P(X \geq a) \leq \frac{E\|X\|}{a}$

Chebyshev: mean μ and variance σ^2 : $P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$

what distance from the mean am I willing to tolerate?

↳ is within 2% of true mean
prob that μ is so if $40 \times 5 \times 60$

$$\text{Chernoff: } P(X \geq a) = \frac{E\|e^{tX}\|}{e^{ta}} = \frac{M_X(t)}{e^{ta}}$$

$$\text{Chebyshev: } P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

* n = sample size

E = tolerance (within 0.05, deviates by no more than 10%)

δ = failure $\rightarrow$ 95% confidence = $\delta = 0.05$ (usually RHS)

EXAMPLES

min order statistic for n i.i.d. $[0,1]$ RV = $\frac{m}{n+1}$

↳ $E\|Y_1\|$ is orthogonal proj so $Z = E\|Y_1\| \perp Y$

↳ Probability node is NOT connected to S^* = $\prod_{i=1}^n (1-p_i)^{n-k}$

↳ P(node is isolated and node j is isolated) = P(node is isolated | node j is isolated) P(node is isolated)

↳ min of $2 \exp(-\lambda_1 \lambda_2) = \frac{1}{\lambda_1 \lambda_2} \rightarrow$ Min N exponentials = $\exp(-N)$

GLOBAL BALANCE eqn = total flow in = total flow out

↳ for birth-death setting: $(n_1, n_2) \rightarrow (n_1+1, n_2)$

$(n_1, n_2) \rightarrow (n_1, n_2+1)$ GLOBAL BALANCE

$(n_1, n_2) \rightarrow (n_1-1, n_2)$ GLOBAL BALANCE

$(n_1, n_2) \rightarrow (n_1, n_2-1)$

CONVERGING IN PROBABILITY

$$\lim_{n \rightarrow \infty} P\|X_n \in C\| = 1 \iff \lim_{n \rightarrow \infty} P\|X_n \notin C\| = 0$$

$$Y_n = \min(X_1, \dots, X_n) \text{ where } X_i \sim \text{Unif}(0,1)$$

$$\text{set up prob } P(Y_n \in C) = P(\min(X_i) \in C) = \prod_{i=1}^n P(X_i \in C)$$

$$\text{set up limit } \lim_{n \rightarrow \infty} P(Y_n \in C) = \lim_{n \rightarrow \infty} \prod_{i=1}^n P(X_i \in C)$$

$$\text{take log } \lim_{n \rightarrow \infty} \sum_{i=1}^n \ln(P(X_i \in C)) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right)$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

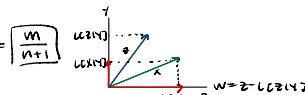
$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \ln\left(1 - \frac{1}{n}\right) = 0$$



GLOBAL BALANCE
 $(n_1, n_2) \rightarrow (n_1+1, n_2)$ GLOBAL BALANCE
 $(n_1, n_2) \rightarrow (n_1, n_2+1)$ GLOBAL BALANCE
 $(n_1, n_2) \rightarrow (n_1-1, n_2)$ GLOBAL BALANCE
 $(n_1, n_2) \rightarrow (n_1, n_2-1)$ GLOBAL BALANCE

BEC expected packets

$$E = \frac{1}{1-p} \rightarrow$$

if waiting for 2 out of 3 packets $\rightarrow$ model as geometric($\frac{2}{3}$)

if equal ($\frac{1}{3}$) probability

$$\text{Expected packets} \rightarrow \frac{1}{1-p} \left(1 - \frac{1}{n}\right)$$

Can also set up $X_i \sim \text{Geom}(1-p)$ for packets after first packet

NEWMAN TESTABLE HYPOTHESIS TEST

$$\Delta(x) = \frac{1}{2} \left(\frac{x}{n} \right)^2$$

$$\textcircled{1} \text{ SET UP } P(\Delta(X) > k | H_0) \leq 0.05 \leftarrow \text{PFA}$$

$$\textcircled{2} \text{ SIMILARITY: } \Delta(x) > k \iff x > \frac{\ln(2k)}{\ln(n)} + 1$$

$$\textcircled{3} \text{ PUT THRESHOLD INTO OG PROB STATEMENT}$$

$$P(X > \frac{\ln(2k)}{\ln(n)} + 1 | H_0) = 0.5$$

$$0.5^{P(n)} = 0.05 \leftarrow \text{set to PFA}$$

$$\textcircled{4} \text{ CHECK THRESHOLD & RANDOMIZE}$$

$$P_{H_0}(X > 5) = 0.5^5 = 0.03125 < 0.05$$

$$P_{H_0}(X = 5) = 0.5(0.5)^4 = 0.03125$$

$$\rightarrow \text{randomize: } 0.03125 + 0.03125 = 0.0625$$

$$\text{reject } H_0 \text{ if } x > 5$$

$$\text{reject } H_0 \text{ w/ prob } 6 \text{ if } x = 5$$

$$\text{accept } H_0 \text{ if } x < 5$$

$$\text{optimal decision rule}$$

$$\text{Bounded MLE}$$

$$\text{MAE of } E\|X\| = \left(1 - \frac{1}{n}\right)^{-1} \rightarrow$$

$$P(X \leq 1) \leq \left(1 - \frac{1}{n}\right)^{-1} e^{-1} \rightarrow$$

find value of t between 0 and 1 which minimizes this expr

$$\text{Serial \# UNIFORM } (1, N)$$

$$P(X_1 = x_1, X_2 = x_2 | N=n) = \frac{1}{n(n-1)}$$

$$\rightarrow \text{argmin } \frac{n(n-1)}{2} \rightarrow$$

increasing so MLE at LEFT endpoint

so max X_1, X_2

$$f_{X,Y}(x,y) = \frac{f_{X|Y}(x|y) f_Y(y)}{f_X(x)}$$

$$E\|X|Y=y\| = \int x f_{X|Y}(x|y) dx$$

$$E\|X\| = \int \int x f_{X,Y}(x,y) dx dy$$

$$\text{HYPOTHESIS EXAMPLE: } P(Y=1|X=0) = 1 - (1-p)^n \leq \beta$$

$$\rightarrow b = \left\lceil \frac{\log(1-\beta)}{\log(1-p)} \right\rceil$$

round down bc decreasing

stopping can wait until you're needs