Probability Review

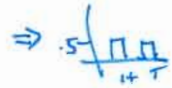
①

X = Random Variable

Value of X depends on prob dist

Discrete $\quad X = \{H, T\} \quad P(X=H) = \frac{1}{2}, \; P(X=T) = \frac{1}{2}$] Dice $X = \{1, \ldots, 6\}$

- biased coin $\rightarrow$ .4, .6

probability mass function : $P(X=x) = \begin{cases} 1/2 & x = H \\ 1/2 & x = T \end{cases} \Rightarrow$ 

Continuous $\quad x \in [0,1]$ uniform, $\quad x \in \mathbb{R}$ gaussian

probability density function $\rightarrow$ derivative of function cumulative distribution

uniform: $F_X(x) = P(X < x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \qquad f_x(x) = F_x'(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \le x \le 1 \\ 0 & x > 1 \end{cases}$

gaussian: $\frac{1}{2}\left[1 + Q\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

Expected Value: not most-likely (mode)

long-term average of samples from X

discrete $\quad E[X] = \sum_x x \, P(X=x)$

continuous $\quad E[X] = \int_{-\infty}^{\infty} x f_x(x)\, dx$

Variance: measure of statistical dispersion

square of the standard deviation

$$\left.\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= E\left[(X - E(X))^2\right] \end{aligned}\right] \quad \begin{aligned} \mu &= E(X) \\ \sigma^2 &= Var(X) \end{aligned}$$

discrete: $\quad Var(X) = \sum_x (x-\mu)^2 P(X=x)$

continuous: $\quad Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_x(x)\, dx$

~~Random Processes~~

~~RP = RV changes over time~~

~~Examples~~

Joint distributions: X, Y two random variables

independent: no statistical dependency between values / Ex: two die, two tosses of coin, roulette

dependent: values related

knowing value of one changes distribution of other

Ex: X, σ

counting cards

Joint Distributions

$p(x, y)$ = multidimensional function

$= p(x)\,p(y)$ if independent

conditional distributions $\quad P(x|y) \rightarrow$ know value of y, how affect x?

$$P(x|y) = \frac{P(x, y)}{P(y)} = P(x) \text{ if indep}$$

Random Sequences

$x[n]$ = random sequence

each value from prob distn

Ex: coin tosses – independent

Ex: $x[n] = nY$ – dependent

– know ~~x[0]~~ $x[1]$, know all

vocab ~~start~~ $x[n]$ stationary if it has the same distribution $\forall n$

$x[1] \sim \mathcal{N}(0,1) \Rightarrow x[100] \sim \mathcal{N}(0,1)$

i.i.d. $\Rightarrow$ independent, identically distributed

white $\Rightarrow$ $x[a]$ ind $x[b]$ $\forall a \neq b$

i.i.d. $\Leftrightarrow$ white, stationary

Random Process

continuous version of random sequence $\Rightarrow x(t)$

Ex: white noise $\quad x(t) \sim \mathcal{N}(0, \sigma^2) \quad$ i.i.d.

$x(t) = Y\sin(t) \Rightarrow$ not i.i.d.

Networks: # of packets waiting to be transmitted in a queue at time t

Markov chains

Random sequence $X[n]$

Key property: $X[n]$ depends on $X[n-1]$
but not $X[n-2], \dots, X[0]$

$$\Rightarrow P[X[n], X[n-1], \dots, X[1], X[0]] = P[X[n], X[n-1]]$$

$$\Rightarrow P[X[n] \mid X[n-1], \dots] = P[X[n] \mid X[n-1]]$$

Ex: state machine
$X \in \{A, B, C\}$

Transition Probability matrix
Σ of each row $= 1$

$$Q = \begin{bmatrix} A\to A & A\to C \\ C\to A & C\to C \end{bmatrix}$$

$$q_{ij} = P[X[n] = j \mid X[n-1] = i]$$

Let $V[n] = \left[P[X[n] = A] \quad P[X[n] = B] \quad \cdots \right]$

if $X[n-1] = i$, $P[X[n] = j] = q_{ij}$

$$V[n] = V[n-1]\, Q$$

$$= [P_A, P_B] \begin{bmatrix} P_{A\to A} & P_{A\to B} \\ P_{B\to A} & P_{B\to B} \end{bmatrix}$$

$$= \left[P_A P_{A\to A} + P_B P_{B\to A}, \quad P_A P_{A\to B} + P_B P_{B\to B} \right]$$

induction:
$$\begin{aligned} V[n] &= V[n-1]\, Q \\ &= (V[n-2]\, Q)\, Q \\ &= V[n-2]\, Q^2 \\ &= V[n-i]\, Q^i \\ &= V[0]\, Q^n \end{aligned}$$

[Chapman-Kolmogorov Equations]

$V[0]$ = initial state distn
$\Downarrow$
always start A
$\Rightarrow [1, 0 \dots 0]$

Markov chains

steady-state distribution

long-term, what is the distribution of $X[n]$?

$\bar{X}$ = dist of $\lim_{n \to \infty} X[n]$

$P(\bar{X}) = v[0]\, Q^{\infty}$

Can solve for as $P(\bar{x}) = P(\bar{x})\, Q$

n equations, n unknowns

⇒ multiple solutions, need $v[0]$ to know for sure

Continuous-Time Markov chains

random processes instead of random sequences

stay at current state until jump to new state

- X_1 and X_2 represent random variables indicating transition times
- if $X_1 < X_2$, will jump to B after X_1

  $X_2 < X_1$, will jump to C after X_2

$P(X_1 = X_2) = 0$ ⇒ continuous

CTMC typically assume exponential distribution

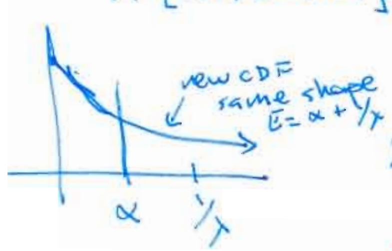
reason: memoryless.

$X \sim \exp(\lambda) \Rightarrow P_X(x) = \lambda e^{-x\lambda}$

$E[X] = 1/\lambda$

$1/\lambda$ = expected time ⇒ λ = expected rate

why memoryless?
- α units of time have passed, timer not expired
- expected time left $1/\lambda - \alpha$? NO
- expected time left still $1/\lambda$



$P[X > s+t \mid X > s] = P[X > t]$

Special Type of Markov Chain

Linear:

$$\begin{tikzcd} 0 \arrow[r, bend left, "\lambda_1"] & 1 \arrow[l, bend left, "\mu_1"] \arrow[r, bend left, "\lambda_2"] & 2 \arrow[l, bend left, "\mu_2"] \arrow[r, bend left, "\lambda_3"] & 3 \arrow[l, bend left, "\mu_3"] \arrow[r, bend left, "\lambda_4"] & {} \arrow[l, bend left, "\mu_4"] \end{tikzcd}$$

known as "Birth-Death" process

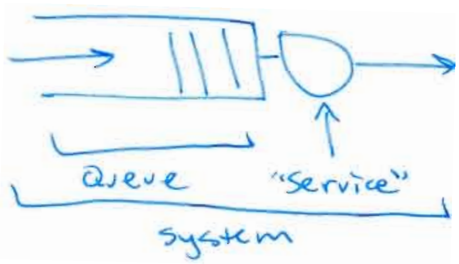
Represents population

$\lambda_i = \lambda \Rightarrow$ λ = Birth rate
μ = Death rate

$\lambda > \mu \Rightarrow$ infinite, long term
$\lambda < \mu \Rightarrow$ small population
$\lambda = \mu \Rightarrow$ undefined, finite

Queues

FIFO = First In, First Out
FILO = First In, last out → stack



- Number of queue elements is a birth-death process
- birth = new job
- death = service complete

Arrival Process: interarrival times are exponentially dist
⇒ Poisson Process

$$P[X(t+\tau) - X(t) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}$$

Service Process: completion time exponential
Poisson if queue nonempty
- can't finish job if none exist

If arrival and service processes are exponential / Poisson then queue is called "M/M/1"

$\underbrace{M/M}_{\text{Memoryless in/out}}/\underbrace{1}_{\text{one server}}$

other types: M = Memoryless
D = ~~[illegible]~~ Deterministic
G = General (~~[illegible]~~ not memoryless)

multiple servers: 1, n, ∞

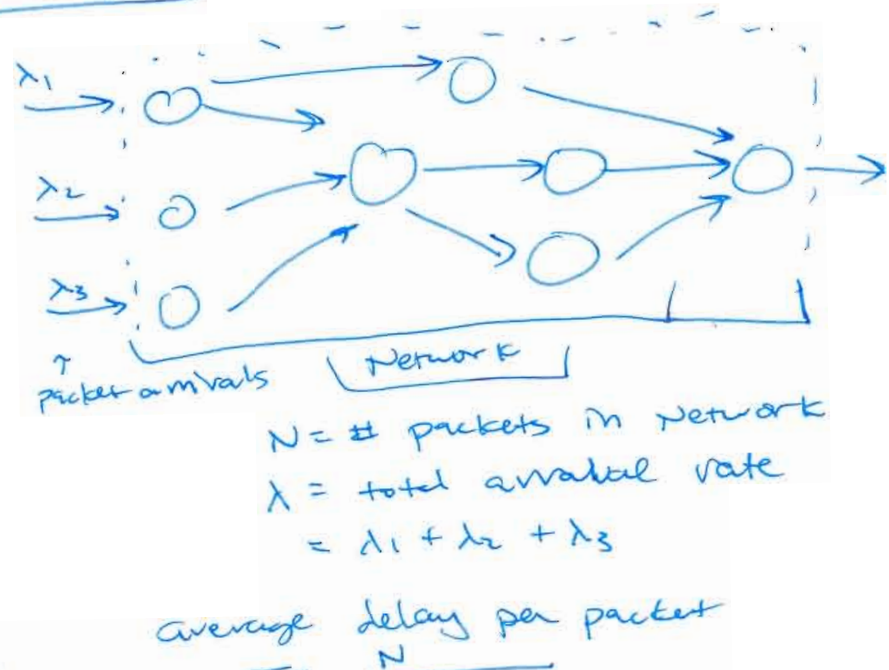
Queues

Things of interest: Instant / Avg Queue length
~~Average~~ Time in Queue
Time in system (Q + service)
Number of people in queue

N = # in ~~queue~~ system N_Q = # in queue
W = avg wait time queue
T = " " system
λ = arrival rate
μ = service rate

Little's Law: $N = \lambda T$ $N_Q = \lambda W$
$N_Q = N - 1$

Examples



N = # packets in network
λ = total arrival rate
$= \lambda_1 + \lambda_2 + \lambda_3$

average delay per packet
$$T = \frac{N}{\lambda_1 + \lambda_2 + \lambda_3}$$

~~M/M/1 Queues~~

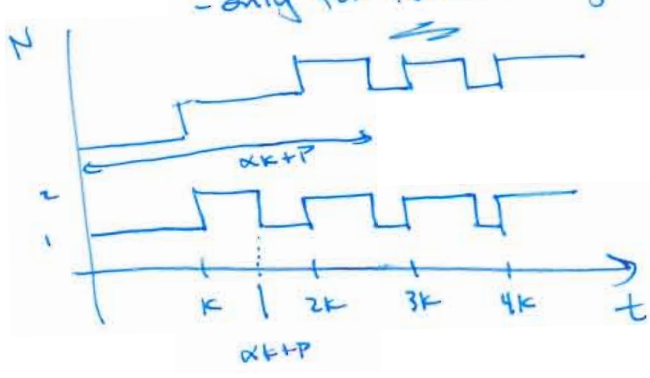
Ex 2 k seconds → new packet
each requires αk time to transmit
processing time = P
interarrival times const
$\alpha < 1$

$T = \alpha k + P$

$\lambda = \frac{1}{k}$

$N = \lambda T = \alpha + \frac{P}{k}$

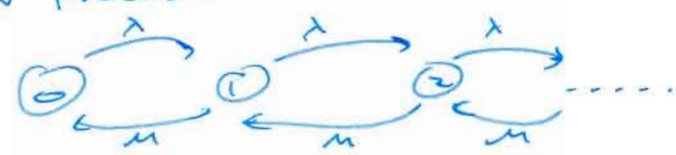
- only for time averages



D/D/1 queue

M/M/1 Queue

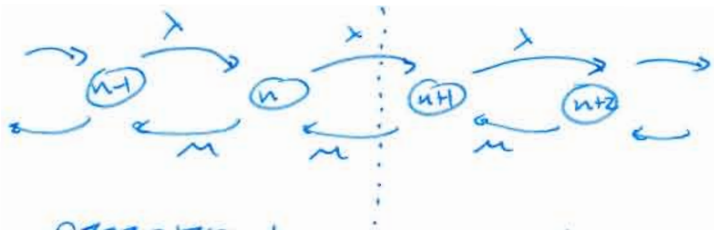
Markov Process



state = # in system = $X(t)$

want to know steady-state distribution

$$p_j = \lim_{t\to\infty} P[X(t)=j]$$



transitions $n \to n+1$ are within 1 of transitions $n+1 \to n$

$$P[n \to n+1] = p_n \lambda$$

$$P[n+1 \to n] = p_{n+1} \mu$$

$$\delta = \{0,1\}$$

over time $t \Rightarrow p_n \lambda t = p_{n+1} \mu t + \delta$

$$\Rightarrow p_n \lambda = p_{n+1} \mu + \frac{\delta}{t}$$

$$\lim_{t\to\infty} \frac{\delta}{t} = 0 \Rightarrow \text{asymptotically } p_n \lambda = p_{n+1} \mu$$

let $\rho = \lambda/\mu \Rightarrow$

$$\begin{aligned} p_{n+1} &= \rho p_n \\ &= \rho^2 p_{n-1} \\ &= \rho^3 p_{n-2} \\ &\vdots \\ p_n &= \rho^n p_0 \end{aligned}$$

prob of all states must sum to 1:

$$1 = \sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \rho^n p_0 = p_0 \sum_{n=0}^{\infty} \rho^n = \frac{p_0}{1-\rho}$$

for $\rho < 1$

therefore $1-\rho = p_0$

$$p_n = \rho^n (1-\rho)$$

M/M/1 Queue

Steady-state queue length:

$$N = \lim_{t\to\infty} E[X(t)]$$

$$= \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n \rho^n (1-\rho)$$

$$= \rho(1-\rho) \sum_{n=0}^{\infty} \underbrace{n\rho^{n-1}}_{\text{derivative of } \rho^n}$$

$$= \rho(1-\rho) \frac{d}{d\rho}\left(\sum_{n=0}^{\infty} \rho^n\right)$$

for $\rho < 1$

$$= \rho(1-\rho) \frac{d}{d\rho}\left(\frac{1}{1-\rho}\right)$$

$$= \rho(1-\rho) \frac{1}{(1-\rho)^2} = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

System wait time per Little's law:

$$T = \frac{N}{\lambda} = \frac{\rho}{\lambda(1-\rho)} = \frac{1}{\mu-\lambda}$$

Q wait time

$$W = \frac{1}{\mu-\lambda} - \frac{1}{\mu} = \frac{\rho}{\mu-\lambda}$$

in Q

$$N_Q = \lambda W = \frac{\rho^2}{1-\rho}$$

<u>overall</u>: $\lambda = \lambda \quad N = \frac{\lambda}{\mu-\lambda} \quad T = \frac{1}{\mu-\lambda} \quad W = \frac{\rho}{\mu-\lambda} \quad N_Q = \frac{\rho^2}{1-\rho}$

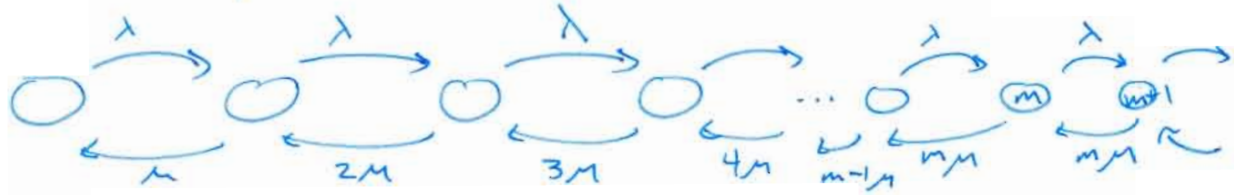
where $\rho = \lambda/\mu$

othere types of queues

after MM/m

$M/M/m \rightarrow m$ servers

- ~~telecom~~ nets: m channels to transmit
- phone: m circuits available



$\rho = \frac{\lambda}{m\mu}$ $\quad P_Q$ = prob of queuing $= P[x(t) \geq m]$

Erlang C $\rightarrow \; = \dfrac{p_0 (m\rho)^m}{m!\,(1-\rho)}$ $\qquad p_0 = \left[\sum_{k=0}^{m-1} \dfrac{(m\rho)^k}{k!} + \dfrac{(m\rho)^m}{m!\,(1-\rho)}\right]^{-1}$

$$p_n = \begin{cases} p_0 \dfrac{(m\rho)^n}{n!} & n \leq m \\ p_0 \dfrac{m^m \rho^n}{m!} & n > m \end{cases}$$

$W = \dfrac{\rho P_Q}{\lambda(1-\rho)}$ $\qquad N_Q = \dfrac{\rho P_Q}{1-\rho}$

$T = \dfrac{1}{\mu} + W$ $\qquad N = m\rho \; \dfrac{\rho P_Q}{1-\rho}$

$M/M/m/m \rightarrow$ no queuing, $(m+1)$th call is dropped

$$P_m = \frac{(\lambda/\mu)^m / m!}{\sum_{n=0}^{m} (\lambda/\mu)^n / n!} \qquad \text{Erlang - B}$$

$M/M/1/m \Rightarrow$ Finite storage

$p_n = p_0 \rho^n$

states Σ to 1 $\quad 1 = \sum_{n=0}^{m} p_n = p_0 \sum_{n=0}^{m} \rho^n = p_0 \left(\dfrac{\rho(1-\rho^m)}{1-\rho} \right)$

$$p_0 = \frac{1-\rho}{1-\rho^{m+1}}$$

$$p_n = \frac{\rho^n (1-\rho)}{1-\rho^{m+1}} \qquad 0 \leq n \leq m$$

$$N = \sum_{n=0}^{m} n p_n = \sum_{n=0}^{m} n \frac{\rho^n (1-\rho)}{1-\rho^{m+1}} = \frac{\rho(1-\rho)}{1-\rho^{m+1}} \sum_{n=0}^{m} n \rho^{n-1}$$

$$= \frac{\rho(1-\rho)}{1-\rho^{m+1}} \frac{d}{d\rho}\left(\frac{\rho(1-\rho^m)}{1-\rho} \right)$$

$$= \frac{\rho^{m+1}(m\rho - m - 1) - \rho^3 + \rho^2 + \rho}{(1-\rho)(1-\rho^{m+1})}$$