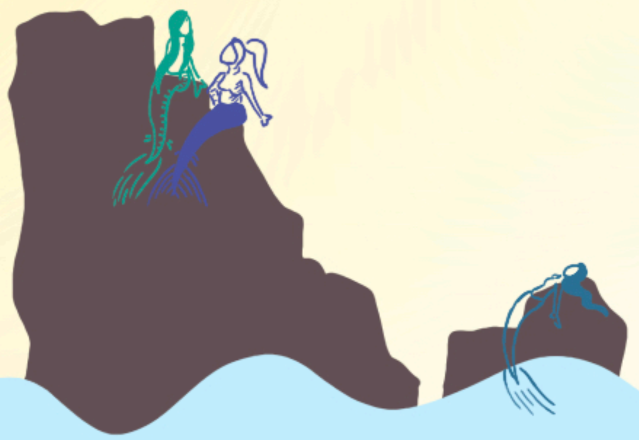
Good Luck

Stochastic

Final Study Guide



Basics

- What is a Stochastic Process?

A probability model that describes the evolution of a system evolving randomly in time.

- We describe a Stochastic Process by describing it's ... Time and State Space

Time
- Discrete Time $\{X_n, n \geq 0\}$ a countable amount of time
- Continuous Time $\{X(t), t \geq 0\}$ an ∞ and uncountable amount of time

State Space
- Discrete State Space: a countable number of outcomes
- Continuous State Space: an ∞ and uncountable number of outcomes

RV Review

- We'll need some basic RVs for when we start exploring CTMC's...

Exponential Distribution

pdf → $f(x) = \lambda e^{-\lambda x}, x \geq 0$

cdf → $F(x) = P(X \leq x) = 1 - e^{-\lambda x}, x \geq 0$

expected value → $E[X] = \frac{1}{\lambda}$

variance → $Var(X) = \frac{1}{\lambda^2}$

the exponential distribution has the memoryless property which means the distribution does not depend on what has happened previously in the distribution.

Erlang Distribution

pdf → $f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{k-1}}{(k-1)!}$

cdf → $F(x) = 1 - \sum_{r=0}^{k-1} e^{-\lambda x} \frac{(\lambda x)^r}{r!}$

expected value → $E[X] = \frac{k}{\lambda}$

variance → $Var(X) = \frac{k}{\lambda^2}$

Poisson Distribution

pmf → $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$

expected value → $E[X] = \lambda$

variance → $Var(X) = \lambda$

properties

Sum of exponentials: X_i are iid $Exp(\lambda)$ rv's and $Z_n = X_1 + X_2 + \cdots + X_n$, then... $Z_n \sim Erl(n, \lambda)$

Minimum of exponentials: X_i are ind $Exp(\lambda_i)$ rv's and $X = \min\{X_1, X_2, \ldots, X_n\}$; $X \sim Exp(\lambda)$, $\lambda = \sum_{i=1}^{n} \lambda_i$

Distribution of (Z, X): The Z and X above are ind rvs w/ dist... $P(Z=i)P(X>x) = \frac{\lambda_i}{\lambda} e^{-\lambda x}$

Sum of Poisson: X_i are ind $Po(\lambda_i)$ rv's and $Z_n = X_1 + X_2 + \cdots + X_n$; $Z_n \sim Po(\lambda)$, $\lambda = \sum_{i=1}^{n} \lambda_i$

DTMC

Discrete Time Markov Chain

$\{X_n, n \geq 0\}$

X_n = the state of a system at time n for $n = 0, 1, 2, \ldots$ (n is discrete)

what do we need to describe X_n... time homo. DTMC
- $S = \{1, 2, \ldots, N\}$
- $P = [p_{i,j}]$ = transition prob. matrix
- $\vec{a} = [a_1, \ldots, a_N]$ = initial dist.
- $\vec{c} = [c(1), \ldots, c(N)]$ = cost incurred at state n.

If we could predict X_{n+1} by just knowing X_n (instead of needing $X_n, X_{n-1}, X_{n-2}, \ldots$) our lives would be a lot easier. A system with the property has the Markov Property.

DTMC: X_n is a discrete time markov chain if for $\forall\, i, j \in S$: $P(X_{n+1}=j \mid X_n=i, X_{n-1}, \ldots, X_0) = P(X_{n+1}=j \mid X_n=i)$

X_n is also time homogenous if... $P(X_{n+1}=j \mid X_n=i) = P(X_1=j \mid X_0=i)$

let's define these... this is a... one step transition probability $= p_{i,j}$

there exists a matrix containing all $p_{i,j}$'s: $P = [p_{i,j}]$

jumping off of a 1 step transition prob. we can make an n-step transition probability: $p^{(n)}_{i,j} = P(X_n=j \mid X_0=i)$, $P^{(n)} = [p^{(n)}_{i,j}] = P^n$

Chapman-Kolmogorov Eq. $P^{(n+m)} = P^{(n)} \cdot P^{(m)} = P^n P^m$

we could also make a transient dist. a.k.a $P(X_n=j) = a_j^{(n)}$; $\vec{a}$ = initial dist; $\vec{a}^{(n)} = \vec{a}P^n$

Applications

Occupancy Time $= m_{i,j}(n)$ = expected # of times system visits state j over $\{0, \ldots, n\}$ if it starts in state i

$M(n) = \begin{bmatrix} m_{1,1}(n) & \cdots & m_{1,N}(n) \\ \vdots & & \vdots \\ m_{N,1}(n) & \cdots & m_{N,N}(n) \end{bmatrix} = \sum_{r=0}^{n} P^r$

expected total cost over finite n if you start in state i $= g(i,n)$; $\vec{g}(n) = M(n) \cdot \vec{c}$

long run expected cost per unit time $= g = g(i) = \sum_{j=1}^{N} \pi_j c(j) = \vec{\pi} \cdot \vec{c}$

First Passage Time = time when a system first passes into/enters a given state or subset of states

A = subset of states; $m_i(A)$ = expected first passage time into subset A $= 1 + \sum_{j \notin A} p_{i,j} m_j(A)$, for $i \notin A$

Limiting Behavior

Limiting Dist. $\vec{\pi} = [\pi_1, \pi_2, \ldots, \pi_N]$, $\pi_j = \lim_{n\to\infty} P(X_n=j)$; the limiting dist is how the system acts as time → ∞; it's a steady state. we find it by solving $\pi_j = \sum_{i=1}^{N} \pi_i p_{i,j}$, $j \in S$; $1 = \sum_{j=1}^{N} \pi_j$

Stationary Dist. $\vec{\pi}^* = [\pi^*_1, \ldots, \pi^*_N]$; if you choose $\vec{\pi}^*$ as your initial dist the dist. will remain the same for all n; if $\vec{\pi}$ exists then $\vec{\pi}^* = \vec{\pi}$

Occupancy Dist. $\hat{\pi} = [\hat{\pi}_1, \ldots, \hat{\pi}_N]$; the fraction of time a system spends in state j; if $\vec{\pi}$ exists then $\hat{\pi} = \vec{\pi}$

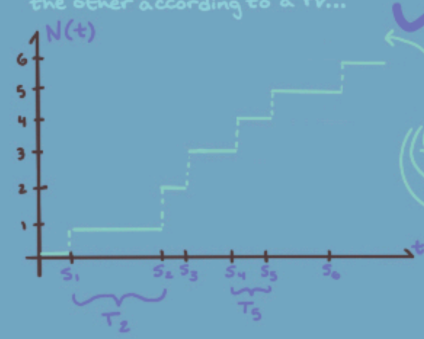
unique dist.
> A DTMC has a unique occupancy and stationary dist. if it is irreducible ↳ (you can go from any state to any state)
> A DTMC has a unique limiting dist. if it's irreducible and aperiodic ↳ (a dtmc w/ period d returns to starting state only at times d, 2d, 3d, ...)

CTMC

Continuous Time Markov Chain

Poisson Process

this system is tracking occurences that occur 1 after the other according to a rv...



- $N(t)$ = total # of events over $(0, t]$
- S_n = occurence time of n^{th} event
- $T_n = S_n - S_{n-1}$ = time between n^{th} and $(n-1)^{th}$ event

Poisson Process
> the stochastic process $\{N(t), t \geq 0\}$ is a Poisson Process w/ rate λ if $\{T_n, n \geq 1\}$ is a sequence of iid $exp(\lambda)$ rv's.
> if $\{N(t), t \geq 0\} \sim PP(\lambda)$, then... $N(t) \sim Po(\lambda t)$
> if some events in a PP are special w/ prob. p we can thin the PP to just those special events... $\{R(t), t \geq 0\} \sim PP(\lambda p)$

$\{X(t), t \geq 0\}$

a stochastic process $\{X(t), t \geq 0\}$ is a CTMC if...

$P(X(s+t)=j \mid X(s)=i, X(u), 0 \leq u \leq s) = P(X(s+t)=j \mid X(s)=i)$

and is time homogenous if... $P(X(s+t)=j \mid X(s)=i) = P(X(t)=j \mid X(0)=i)$

this is very analogus to our DTMC defs, and like DTMC's we can define...

$p_{i,j}(t) = P(X(t)=j \mid X(0)=i)$ $\quad P(t) = [p_{i,j}(t)]$

$P(t) = \sum_{k=0}^{\infty} e^{-rt} \frac{(rt)^k}{k!} \hat{P}^k$

$r \geq \max\{r_1, \ldots, r_N\}$

$\hat{P} = [\hat{p}_{i,j}]$

$\hat{p}_{i,j} = \begin{cases} 1 - \frac{r_i}{r} &, \text{if } i=j \\ \frac{r_{i,j}}{r} &, \text{if } i \neq j \end{cases}$

$X(t)$ = state of a system at time t

starts in state i → stays in state i for $exp(r_i)$ amount of time

then transitions to state $j \neq i$ w/ prob $p_{i,j}$ → $p_{i,i} = 0$

stays in state j for $exp(r_j)$ amount of time...

$p_{i,j} \neq p_{i,j}(t)$

$p_{i,j}$ = prob system moves to state j after leaving i

what we will use most of the time to describe X(t) is...
- r_i = parameter describing the amount of time the system stays in i.
- $p_{i,j}$ = Prob of system moving to state j after state i
- $r_{i,j} = r_i \cdot p_{i,j}$ → $r_i = \sum_{j=1}^{N} r_{i,j}$ → $p_{i,j} = \frac{r_{i,j}}{r_i}$

$R = [r_{i,j}]$ = the rate matrix

The process as define above is a CTMC, because the time between occurences are exp. and therefore memoryless.

NOTE: We can use a PP to describe arrivals to a CTMC system (since a PP is made up of exp rv)

Applications

many of these are very similar to their DTMC definitions

occupancy time of state j until time T starting from state i. $= m_{i,j}(T) = \int_0^T p_{i,j}(t)\,dt$

expected total cost up to finite time T starting from state i $= \vec{g}(T) = M(T)\vec{c}$ → cost vector

long run cost per unit time $= g = \sum_{j=1}^{N} p_j c(j)$

First passage times = time when system first passes into a set of states A; solve $r_i m_i = 1 + \sum_{j \notin A} r_{i,j} m_j$ for $i \notin A$

Limiting behavior

a CTMC is irreducible if it's corresponding DTMC is irreducible

Limiting Dist. an irreducible CTMC has a unique limiting distribution $\vec{p}$ given by... $p_j r_j = \sum_{i=1}^{N} p_i r_{i,j}$, $1 = \sum_{i=1}^{N} p_i$, $1 \leq j \leq N$

Stationary Dist. if $\vec{p}$ the lim. dist. exists for a CTMC, then it's also the stationary dist. If the initial dist for the CTMC is $\vec{p}$, then the dist will stay at $\vec{p}$ for $\forall t$.

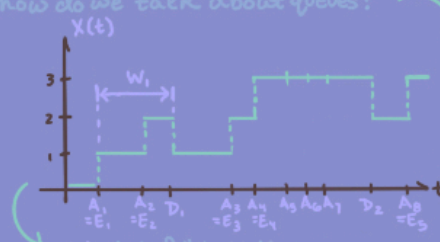
Occupancy Dist. if $\vec{p}$ the lim dist exists for a CTMC, then it's also the occupancy dist aka the fraction of long run time a system spends in a given state.

QUEUES

one common type of CTMC is a queuing model...

how do we talk about queues?

4 attributes... __/__/__/__
- Arrival Process: interarrival times can be... M → exponential, G → general, D → deterministic, E_K → Erlang
- Service Times: service times can be... M → exponential, G → general, D → deterministic, E_K → Erlang
- # of Servers: denoted s
- Capacity: denoted K; if blank assume ∞



- $X(t)$ = total # in system
- A_n = time of n^{th} arrival
- E_n = time of n^{th} entry (can't enter if system is full)
- D_n = time of n^{th} departure
- W_n = time spent by n^{th} cust. in system

more defs...

- L = expected # cust in system in steady state
- λ = arrival rate
- W = expected time spent in system by cust. in steady state.

Little's Law $L = \lambda W$

general definitions...
- $N(t)$ = # of cust arrived by time t
- $D(t)$ = # of cust departed by time t
- $X_n = X(D_n^+)$ = # of cust left after n^{th} departure
- $X_n^* = X(E_n^-)$ = # of cust before n^{th} entry.
- $\hat{X}_n = X(A_n^-)$ = # of cust in system before n^{th} arrival.
- $p_j = \lim_{t\to\infty} P(X(t)=j)$ = long run fraction of time the system has j cust.
- $\pi_j = \lim_{n\to\infty} P(X_n=j)$ = long run fraction of departs that leave behind j cust.
- $\pi_j^* = \lim_{n\to\infty} P(X_n^*=j)$ = long run fraction of entering cust. that see j people ahead.
- $\hat{\pi}_j = \lim_{n\to\infty} P(\hat{X}_n=j)$ = long run fraction of arrivals that see j people ahead.

FACT! if one of the limits exists... $\pi_j = \pi_j^*$ and if arrival is Po... $p_j = \hat{\pi}_j$ → PASTA

M/M/_/K

exponential arrival $= \lambda$, service $= \mu$; finite

M/M/1/K

traffic intensity $= \rho = \frac{\lambda}{\mu}$

limiting prob. $= p_i(K) = \begin{cases} \frac{1-\rho}{1-\rho^{K+1}} \rho^i &, \text{if } \rho \neq 1 \\ \frac{1}{K+1} &, \text{if } \rho = 1 \end{cases}$

server idleness $= p_0(K)$

system is full = long run frac. of cust lost $= p_K(K)$

prob. an entering cust sees i cust ahead of them $= \pi_i^* = \frac{p_i(K)}{1-p_K(K)}$

expected # in system $= L = \sum_{i=0}^{K} i\, p_i(K)$

expected waiting time (arrivals) $= W = L/\lambda$

expected waiting time (entries) $= W = \frac{L}{\lambda(1-p_K(K))}$

↳ also interpret as... expected rate of revenue = expected pay by entering cust $\cdot \lambda(1-p_K)$

use these to find more quantities for M/M/s/K

M/M/s/K

traffic intensity $= \rho = \frac{\lambda}{s\mu}$

limiting prob. $= p_i(K) = p_0(K)\rho_i$

$\rho_i = \begin{cases} \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i &, \text{if } 0 \leq i \leq s-1 \\ \frac{s^s}{s!}\rho^i &, \text{if } s \leq i \leq K \end{cases}$

$p_0(K) = \left(\sum_{i=0}^{s-1} \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i + \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \cdot \frac{1-\rho^{K-s+1}}{1-\rho} \right)^{-1}$

M/M/K/K

limiting prob. $= p_i(K) = \frac{\rho_i}{\sum_{j=0}^{K}\rho_j}$

$\rho_i = \frac{(\lambda/\mu)^i}{i!}$

M/M/_

∞ capacity

M/M/1

traffic intensity $= \rho = \frac{\lambda}{\mu}$

queue is stable if and only if $\rho < 1$

limiting prob. $= p_i = (1-\rho)\rho^i$

$\hat{\pi}_i = \pi_i^* = \pi_i = p_i = (1-\rho)\rho^i$

prob. server is idle $= p_0 = 1-\rho$

prob. server is busy = exp. # busy servers $= 1 - p_0 = \rho$

exp. # in system $= L = \frac{\rho}{1-\rho}$ $\quad$ exp. waiting time $= W = L/\lambda = \frac{1}{\mu-\lambda}$

M/M/∞ → always stable

Limiting prob. $= p_i = e^{-\lambda/\mu} \frac{(\lambda/\mu)^i}{i!}$

exp. # in system $= L = \frac{\lambda}{\mu}$

M/M/s

stable if... traffic intensity $= \rho = \frac{\lambda}{s\mu} < 1$

Limiting prob. $= p_i = p_0 \rho_i$

$\rho_i = \begin{cases} \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i &, \text{if } 0 \leq i \leq s-1 \\ \frac{s^s}{s!}\rho^i &, \text{if } i \geq s \end{cases}$

$p_0 = \left(\sum_{i=0}^{s-1} \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i + \frac{s^s}{s!} \cdot \frac{\rho^s}{1-\rho} \right)^{-1}$

$\hat{\pi}_i = \pi_i^* = \pi_i = p_i = p_0\rho_i$

prob. all servers are busy $= \frac{p_s}{1-\rho}$

exp. # of busy servers $= B = \frac{\lambda}{\mu}$

exp. # in system $= L = \frac{\lambda}{\mu} + p_s \frac{\rho}{(1-\rho)^2}$

exp. waiting time $= W = \frac{L}{\lambda} = \frac{1}{\mu} + \frac{1}{s\mu}\frac{p_s}{(1-\rho)^2}$

M/G/1

Not a CTMC! exponential arrivals, general service times, single server

arrival times $= \lambda$

service times mean $= \tau$

variance $= \sigma^2$

2nd mom. $= s^2 = \sigma^2 + \tau^2$

traffic int. $= \rho = \lambda\tau < 1$

prob. server is busy $= \rho$

prob server is idle $= 1-\rho$

exp. time in queue $= W_q = \frac{\lambda s^2}{2(1-\rho)}$ → exp. time in system $= W = \tau + W_q$

exp. # in queue $= L_q = \frac{\lambda^2 s^2}{2(1-\rho)}$ → exp. # in system $= L = \rho + L_q$

G/M/1

Not a CTMC! gen arrival times, exponential service times, single server

service times mean $= \tau = 1/\mu$

arrival times mean $= 1/\lambda$

traffic int. $= \rho = \lambda\tau = \frac{\lambda}{\mu} < 1$

Lim. prob. $= p_j = \rho\pi^*_{j-1}$

$\pi_j^* = (1-\alpha)\alpha^j$ ↳ α is the sol. to the func. eq.

prob. server is busy $= \rho$

prob. server is idle $= p_0 = 1-\rho$

exp. # in system $= L = \frac{\rho}{1-\alpha}$

exp. wait time $= W = L/\lambda = \frac{1}{\mu}\frac{1}{1-\alpha}$

Networks

multiple queues!

Jackson Networks
- N single station queues
- i^{th} station has s_i servers
- unlimited capacity at each station
- arrive at station i from outside network according to $PP(\lambda_i)$ indp of each other
- service times at station i are iid $exp(\mu_i)$ rv
- after leaving station i the cust goes to station j w/ prob. $p_{i,j}$ or leave w/ prob. r_i

$P = [p_{i,j}]$ $\vec{r} = [r_i]$; $\sum_{j=1}^{N} p_{i,j} + r_i = 1, \forall i$

for station j...
- λ_j = external arrivals to j
- b_j = internal arrivals to j
- $a_j = \lambda_j + b_j$ = total arrivals

$b_j = \sum_{i=1}^{N} a_i p_{i,j}$

$a_j = \lambda_j + \sum_{i=1}^{N} a_i p_{i,j}$ → $\vec{a} = \vec{\lambda} + \vec{a}P$

stability if... $a_i < s_i\mu_i, \forall i$

$X_i(t)$ = # of cust. in i^{th} service station of a JN @ time t

$p(n_1, n_2, \ldots, n_N) = p_1(n_1)p_2(n_2)\cdots p_N(n_N)$ in steady state

$p_i(n) = p_i(0)\rho_i(n)$ (station i, n cust.) node i behaves like M/M/s queue w/ a_i, μ_i

expected # cust. @ station i $= L_i = \frac{\rho_i}{1-\rho_i}$, $\rho_i = \frac{a_i}{\mu_i}$ if $s = 1$

expected # cust in total $= L = \sum_{i=1}^{N} L_i$