STOCHASTIC midterm 1

PT 1 Introduction

What is a Stochastic process?

- A probability model that describes a system that evolves randomly over time...
- A Stochastic model can be on...

Discrete Time

- a countable measurement of time (ie 1,2,3...)

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

X_n = random state of system at time n

n takes values $n = 0,1,2,\dots$

$X_n \in S$ → values X can take the State Space

Continous Time

- An uncountable/infinite measure of time

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

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

t takes values $t \geq 0$, $[0, \infty)$

$X(t) \in S$ → values X can take, the State Space

Discrete State Space

- A countable list of things that can happen

Continous State Space

- An uncountable/infinite list of things that can happen.

how can we work with DTDS stoch. models...

↳ discrete time discrete state space

PT 2 Basics of Discrete Time

- We have a discrete time stoch. model X_n and now we want to use it to predict what the state will be at a certain time. of course we would assume predicting X_{n+1} would require info from $X_0, X_1, X_2, \dots, X_n$, but if $X_0, X_1, \dots, X_{n-1}$ was redundent and only X_n mattered in predicting X_{n+1} that would make things a lot easier...

Markov Chains

- A stochastic process $\{X_n, n \geq 0\}$ on state space S is a discrete time Markov Chain if for all i and j in S...

$$P(X_{n+1}=j \mid X_n=i, X_{n-1}, \dots, X_0) = P(X_{n+1}=j \mid X_n=i)$$

X_{n+1} only relies on X_n

- A DTMC is Time Homogenous if for all $n = 0,1,\dots$

$$\underbrace{P(X_{n+1}=j \mid X_n=i)}_{\text{one step transition probability}} = P(X_1=j \mid X_0=i)$$

→ the time doesn't matter

↳ $p_{i,j}$ Transition Probability Matrix →

$$P = \begin{bmatrix} & & \\ & p_{i,j} & \\ & & \end{bmatrix} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{N,1} & \cdots & \cdots & p_{N,N} \end{bmatrix}$$

→ each row adds to 1
→ row = starting state
so entry $p_{i,j}$ = prob of starting at i and ending at j
column = ending state

- we can also draw as graph...

$$\begin{tikzcd} \textcircled{1} \arrow[r, bend left, "p_{1,2}"] \arrow[loop left, "p_{1,1}"] & \textcircled{2} \arrow[l, bend left, "p_{2,1}"] \arrow[r, bend left, "p_{2,3}"] & \textcircled{3} \arrow[l, bend left, "p_{3,2}"] \arrow[loop right, "p_{3,3}"] \end{tikzcd}$$

→ prob. of starting at 2 and going to 3
A node for each State Space

- for now, we will only look at time homogenous dtmcs...

Transient Distributions

- we have $\{X_n, n \geq 0\}$ → time homo dtmc → $\in S = \{1,2,3,\dots,N\}$ → transition Prob. matrix P

NEW! initial dist. $\vec{a}$

$a_i = P(X_0 = i)$, $1 \leq i \leq N$

$\vec{a} = [a_1, \dots, a_N]$ → adds to 1

prob a matrix will start in state 1...N

- Transient Distribution → the prob. of the process being in state j.

$$P(X_n=j) = \sum_{i=1}^{N} a_i \underbrace{P(X_n=j \mid X_0=i)}_{\text{n-step transition probability}} = a_j^{(n)} = [\vec{a}P^n]_j$$

much like a 1 step transition probability but a n sized step

n-step transition probability → $p_{i,j}^{(n)}$

n-step transition probability matrix →

$$P^{(n)} = \begin{bmatrix} p_{i,j}^{(n)} \end{bmatrix} = \begin{bmatrix} p_{1,1}^{(n)} & \cdots & p_{1,N}^{(n)} \\ \vdots & & \vdots \\ p_{N,1}^{(n)} & \cdots & p_{N,N}^{(n)} \end{bmatrix}$$

- Some important theorems about n step transition Probability matrix

$P^{(n)} = P^n$ → nth power of matrix P

$a^{(n)} = \vec{a}P^n$

$P^{(n+m)} = P^{(n)} \cdot P^{(m)}$

- Transient PMF vector →
- the probability X_n will be in State space j at time n is...

$\vec{a}^{(n)} = [a_1^{(n)}, a_2^{(n)}, \dots a_j^{(n)}, \dots, a_N^{(n)}]$

$a_j^{(n)}$ → $P(X_n = j)$

PT 3 Discrete Time Applications

- Lets see what we can do with DTMCs...

Occupancy Times

- We have our $\{X_n, n \geq 0\}$, lets say we want to study the amount of time a DTMC will spend in a given state during a given time...

$N_j(n)$ = # of times the DTMC visits state j over $\{0,1,2,\dots,n\}$

$m_{i,j}(n) = E[N_j(n) \mid X_0 = i]$ = Occupancy Time

$$M(n) = \begin{bmatrix} m_{1,1}(n) & \cdots & m_{1,N}(n) \\ \vdots & & \vdots \\ m_{N,1}(n) & \cdots & m_{N,N}(n) \end{bmatrix} = \text{Occupancy Time Matrix}$$

how long a system spends in time j if $X_0 = i$ up to n

↳ $M(n) = \sum_{r=0}^{n} P^r$

Limiting Behavior

- What if we want to study what happens to X_n as $n \to \infty$. To look at this question we first need two defs...
- Irreducibility → a dtmc where for every i and j in S there is a $k > 0$ such that... $P(X_k = j \mid X_0 = i)$. AKA you can go from any state i to any state j in some # of steps.
- Periodicity → if we have an irreducible dtmc and d is the largest integer such that... $P(X_n = i \mid X_0 = i) > 0$ and n is an integer multiple of d.
  - Periodic → if $d > 1$
  - aperiodic → if $d = 1$

  AKA basically just wondering if the system is in a cycle or oscillating.

- Now lets talk about... Limiting Distributions aka steady state distr.
  - as the system carries on to ∞ our $P(X_n = j)$ aka $a_j^{(n)}$ might approach a steady state...
  - if a limiting distr. exists it satisfies...

$\vec{\pi} = [\pi_1, \pi_2, \dots, \pi_N]$ → for states 1,2,...,N

$\pi_j = \lim_{n \to \infty} P(X_n = j)$

$\underbrace{\pi_j = \sum_{i=1}^{N} \pi_i p_{i,j}}_{\text{balance eq}}$ $\quad \underbrace{1 = \sum_{j=1}^{N} \pi_j}_{\text{normalizing eq}}$ → matrix form $\vec{\pi} = \vec{\pi}P$

- But even if X_n oscilates and does not approach a limit we can still solve the balance and normalizing equation. The solution then is the...

Stationary Distribution → $\vec{\pi}^* = [\pi_1^*, \pi_2^*, \dots, \pi_N^*]$
  - same #s as Limiting Distr.
  - if we choose $\vec{\pi}^*$ as our initial dist. & then the pmf of X_n will remain the same for all n!

↳ when a limiting dist. exists it is also a stationary dist.

- One last interpretation... Occupancy Distribution
  - same #s as other two, again found from b + n equations.
  - the % of time X_n spends in each state j.

$$\hat{\pi}_j = \lim_{n \to \infty} \frac{E[N_j(n) \mid X_0 = i]}{n+1}$$

$$\vec{\hat{\pi}} = [\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_N]$$

- A unique dist. exists when it does not depend on initial dist.
- How do we know when a unique limiting, stationary, or occupancy dist exists?

Unique Stationary Dist.
- ☐ irreducible
- ☐ = Unique Occupancy

Unique Occupancy Dist.
- ☐ irreducible
- ☐ = Unique Stationary

Unique Limiting Dist.
- ☐ irreducible
- ☐ aperiodic
- ☐ = other 2

Cost Models

- We can add costs to our models ("cost" can mean price, reward, profit, etc.) lets def. some new varsiables...

$C(i)$ = random cost per visit to i → $C(X_r)$ = cost at time r

$c(i)$ = expected cost per visit to i ↳ $c(X_r)$ = expected cost at time r

- Expected Total Cost over Finite time →

$g(i,n) = E\left[\sum_{r=0}^{n} c(X_r) \mid X_0 = i\right]$ → up to time n, if we start in state i

$\vec{g}(n) = M(n) \cdot \vec{c}$ → $\vec{c}$ = vector of cost at each state → $M(n)$ = occupancy time matrix

- $g(i,n) \to \infty$ as $n \to \infty$ (duh), so if we wanted to look at long run in a smarter way we would look at Long Run Cost Rate → (X_n has to be irreducible)

expected cost per unit time in long run.

$g(i) = \lim_{n \to \infty} \frac{g(i,n)}{n+1}$ → $g = g(i) = \sum_{j=1}^{N} \hat{\pi}_j c(j)$ → matrix form $g = \vec{\hat{\pi}}\vec{c}$

First Passage Times

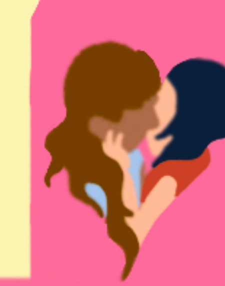
- The random time at which a stochastic process "first passes into" a given subset.

$T = \min\{n \geq 0 : X_n = N\}$ or $T = \min\{n \geq 0 : X_n \in A\}$

$m_i = E[T \mid X_0 = i]$ $\quad$ $m_i = E[T \mid X_0 = i]$

- Expected First Passage Times → $m_i = 1 + \sum_{j=1}^{N-1} p_{i,j} m_j$ $\quad$ $m_i(A) = 1 + \sum_{j \notin A} p_{i,j} m_j(A)$

↳ matrix form $\vec{m}(A) = e + P(A)\vec{m}(A)$



PT 4 Next Steps

- Now we want to move on to continuous time stochastic models. But before we can do that we need to review some random variables...

Exponential RVs

- The Exponential Random Variable is a RV w/ the pdf and the cdf and other properties

$f(x) = \lambda e^{-\lambda x}$, $x \geq 0$ $\quad$ $F(x) = P(X \leq x) = 1 - e^{-\lambda x}$ $\quad$ $E[X] = 1/\lambda$, $Var(X) = 1/\lambda^2$

- Write as $X \sim Exp(\lambda)$
- An interesting property of the Exponential RV is the Memoryless Property which means...

$$P(X > t+s \mid X > s) = P(X > t), \text{ for } s,t \geq 0$$

↳ a common example is a lightbulb that has prob. q of lasting 1 month and a prob. q of lasting a 2nd. It's like it doesn't remember how long its been working for

- only the Exponential RV has the memoryless property.
- A distribution related to the Exponential is the Erlang Rv with pdf and cdf and properties

$f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{k-1}}{(k-1)!}$ $\quad$ $F(x) = P(X \leq x) = 1 - \sum_{r=0}^{k-1} e^{-\lambda x} \frac{(\lambda x)^r}{r!}$ $\quad$ $E[X] = \frac{k}{\lambda}$, $Var(X) = \frac{k}{\lambda^2}$

- We write it as $X \sim Erl(k, \lambda)$
- we can see the exponential and the erlangs similarities but the real reason we put them together is because if we take the Sum of iid $Exp(\lambda)$ →

$$Z_n = X_1 + X_2 + \dots + X_n$$

then $Z_n \sim Erl(n, \lambda)$

- We can also look at a rv X that is the Minimum of ind $Exp(\lambda_i)$ →

$$X = \min\{X_1, X_2, \dots, X_k\}$$

then $X \sim Exp(\lambda)$ with $\lambda = \sum_{i=1}^{k} \lambda_i$

- An interesting property of the minimum above is that the rv X is independent of which X_i is the minimum.
  - $Z = i$ if $X_i = X$, then $P(Z=i, X>x) = P(Z=i)P(X>x)$

↳ in the context of an ex. the min time it takes 1 of k lightbulbs to fail is indp. of which of the k lightbulbs fails first.

$P(Z=i) = \frac{\lambda_i}{\lambda}$ → prob. that lightbulb i fails first.

Poisson RVs

- poisson rv can be dif. to conceptualize, one way to think about them is by thinking about Binomial RVs with pmf and if we study it as $n \to \infty$ and $p \to 0$...

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \lim_{n \to \infty, p \to 0} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

and this is...

- Poisson RV → a rv with pmf $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$ and other properties

$$E[X] = Var(X) = \lambda$$

- We write is as $X \sim P_o(\lambda)$
- If we take the Sum of ind. $P_o(\lambda)$ →

$$Z_n = X_1 + X_2 + \dots X_n$$

then $Z_n \sim P_o(\lambda)$ with $\lambda = \sum_{i=1}^{n} \lambda_i$

XOXO

good luck