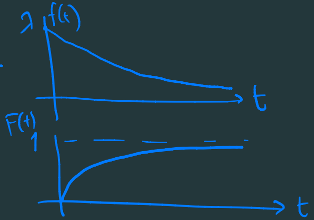
exponential distribution cheat sheet $\lambda \equiv 1/\text{sec} \Rightarrow \lambda \equiv$ 'rate'

can extend DTMCs to continuous time using special properties of the exponential distribution/poisson process

suppose $T \sim$exponential(λ), then: $\mathbb{E}[T] = 1/\lambda$



- pdf: $f_T(t) =$ $\lambda e^{-\lambda t} \quad \forall\, t \geq 0$
- cdf: $F_T(t) = \mathbb{P}[T \leq t] =$ $1 - e^{-\lambda t}$
- (memorylessness): cdf of T knowing it is bigger than t?

$$\underbrace{\mathbb{P}[T \leq t + x | T > t]}_{} = \frac{\mathbb{P}[t < T \leq t+x]}{\mathbb{P}[T > t]} = \frac{F(t+x) - F(t)}{1 - F(t)} \quad \text{(any r.v.)}$$

$$= \frac{e^{-\lambda t} - e^{-(t+x)\lambda}}{e^{-t\lambda}} = 1 - e^{-x\lambda} = F(x)$$

$= \mathbb{P}[T \leq x]$

iff $T \sim \mathcal{E}xp(\lambda)$

(or for discrete, $T \sim \text{Geom}(p)$)

the exponential distribution: properties

suppose $T_1, T_2, \ldots, T_n$ are all independently exponentially distributed, with $T_i \sim Exp(\lambda_i)$.

- (minimum of exponentials): let $T_{\min} = \min\{T_i | i \in \{1, 2, \ldots, n\}\}$
  distribution of $T_{\min}$?

  $$T_{\min} \sim \mathbb{P}\left[\min(T_1, T_2, \ldots, T_n) \geq t\right] = \mathbb{P}[T_1 \geq t, T_2 \geq t, \ldots T_n \geq t] \overset{\text{independence}}{=} \prod_{i=1}^{n} \mathbb{P}[T_i \geq t]$$

  $$= \prod_{i=1}^{n} e^{-\lambda_i t}$$

  $$= e^{-(\sum \lambda_i) t} \sim Exp\left(\sum_{i=1}^{n} \lambda_i\right)$$

- (first arrival): let $I_{\min} = \arg\min\{T_i | i \in \{1, 2, \ldots, n\}\}$
  distribution of $I_{\min}$?

  $$I_{\min} \sim \mathbb{P}[I_{\min} = i] = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

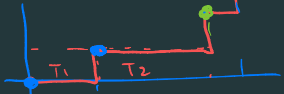
  (multinomial dist^n)

Poisson process cheat sheet

given Poisson processes $X(t) \sim PP(\lambda)$ - Counting process with iid Exp interarrival times

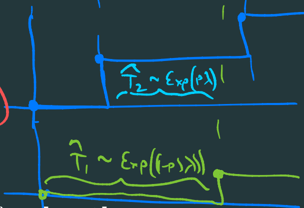
- (inter-arrival times): let $\{A_1, A_2, A_3 \dots\}$ be the arrival times of the agents; then

  $$T_i = A_i - A_{i-1} \sim \text{Exp}(\lambda)$$



- (splitting): suppose we probabilistically split arrivals from $X(t)$ to $Y(t)$ with probability p, else to $Z(t) = X(t) - Y(t)$

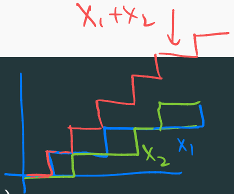
  $$Y(t) \sim PP(p\lambda) \quad \text{and} \quad Y(t) \perp\!\!\!\perp Z(t)$$
  $$Z(t) \sim PP((1-p)\lambda)$$



- (time-varying rate): a time-varying rate of $\lambda(t) \in [0, \lambda^\star]$ is equivalent to a $PP(\lambda^\star)$ for which arrivals at time t are thinned with probability $p(t) = \lambda(t)/\lambda^\star$

later in course

Poisson process cheat sheet (contnd)



given independent Poisson processes
$X_1(t) \sim PP(\lambda_1), X_2(t) \sim PP(\lambda_2), , X_3(t) \sim PP(\lambda_3)$

- (superposition): suppose $S(t) = X_1(t) + X_2(t) + X_3(t)$

  $S(t) \sim$ $PP(\lambda_1 + \lambda_2 + \lambda_3)$

- (first arrival): let $I_{\min}$ = the identity (i.e., $\{1, 2, 3\}$) of the first arrival among the three processes; distribution of $I_{\min}$?

  $I_{\min} \sim$ i wp $\lambda_i / \sum_{i=1}^{3} \lambda_i$

- # of arrivals in $[t, t+s] = N(t+s) - N(t) \sim \mathrm{Poi}(\lambda s)$

  $\frac{e^{-\lambda s}(\lambda s)^k}{k!}$

- Arrivals in disjoint intervals are independent

class poll: spreading a rumor

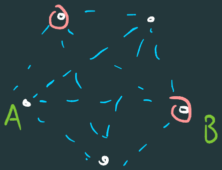
we model rumor spreading among n people using a Markovian model:

– each pair of people (i, j) independently meet after Exponential$(1/\tau)$ time

– when a person in the know meets someone who is unaware, then the rumor spreads

suppose at time t, there are $N(t)$ people who know the rumor

what is the distribution of the time T after which the number of people in the know increases to $N(t) + 1$?



(a) Exponential $\left(N(t)/\tau\right)$

(b) Exponential $\left(N(t)\,\tau\right)$

(c) Poisson $\left(N(t)(n-N(t))/\tau\right)$

(d) Exponential $\left(N(t)(n-N(t))/\tau\right)$

(e) Exponential $\left(N(t)(N(t)-1)/2\tau\right)$

Rumor process - $N(t)$ = # of people (out of n) who know rumor at time t

$N(t) = 2$ $\qquad$ $n - N(t) = 3$

Bad idea for Sim

- Generate ordered list of who meets whom

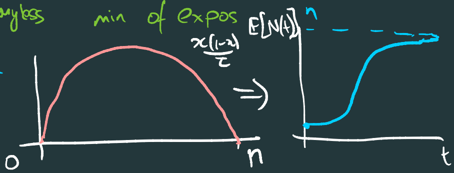
$(1,3)\ (2,3)\ (1,5)\ \ldots$

- **Better Idea** - ONLY simulate relevant events

i.e., meetings between those in the know $(N(t))$ and those who do not know $(n - N(t))$ $\Rightarrow$ $N(t)(n-N(t))$ pairs

$$\Rightarrow \text{Earliest relevant event} = \underbrace{t}_{\text{memoryless}} + \underbrace{\mathrm{Exp}\left(N(t)(n-N(t))/\tau\right)}_{\text{min of expos}}$$

- **Note** - What is the rate of spread as a fn of number in the know



ORIE 4580/5580: Simulation Modeling and Analysis

Unit 10: intro to Markov chains $\begin{pmatrix} \text{discrete time MC} \\ \text{cont time MC} \end{pmatrix}$

Sid Banerjee

School of ORIE, Cornell University

rest of course: the 'simulation hierarchy'

aim: simulate some complex real-world system to measure/optimize/control

– can do so in 3 ways of decreasing complexity (ie, faster sim)

discrete-event simulations (closest to reality)

- most general framework
- allows detailed modeling, general distributions
- complex; takes time to code/execute

Great for convincing 'non-experts'

- Not good for testing / experimenting

simplify

'Markovian models'

- need inter-event times to be memoryless (i.e., Exponentially distributed)
- easier to simulate (no event-list needed), good approx for large systems
- can give spurious insights, hide critical issues

closed-form solutions (ORIE 3510)

- e.g. queueing models - PK formula
- have formulas for steady-state performance measures
- restrictive assumptions

Main use in Sim

- Good test cases

today: intro to Markov chains + Markovian simulation models

random process

random process (rvs with an index)

indexed collection of rvs $X_t \in \mathcal{S}$, one for each $t \in T$ ← $\{0,1,2,\dots\}$ - discrete time

– $\mathcal{S}$: state space, $\mathcal{T}$: index set

$\mathbb{R}^2$ - spatial proc. $[0,\infty)$ - cont time

Markov property

random process X_t has the Markov property if the probability of moving to a future state only depends on the present state and not on past states

Eg - $T = \{0,1,2,\dots\}$ X_t is Markovian iff

$$\mathbb{P}[X_t \le x \mid \underbrace{X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2} \dots X_0 = x_0}_{\text{history}}] = \mathbb{P}[X_t \le x \mid X_{t-1} = x_{t-1}]$$

more generally

$$\mathbb{P}[X_t \le x \mid X_{t-2} = a, X_{t-4} = b, X_0 = c] = \mathbb{P}[X_t \le x \mid X_{t-2} = a]$$

ie - most recent state is all you need

Note - Every process is Markovian (by expanding the state)

random process

random process

indexed collection of rvs $X_t \in \mathcal{S}$, one for each $t \in T$

– $\mathcal{S}$: state space, $\mathcal{T}$: index set

Markov property

random process X_t has the Markov property if the probability of moving to a future state only depends on the present state and not on past states

four types

chains

- $\mathcal{S}$ discrete, $\mathcal{T}$ discrete: discrete-time Markov chain (DTMC)
  – random walk on integers
- $\mathcal{S}$ discrete, $\mathcal{T}$ continuous: continuous-time Markov chain (CTMC)
  – Poisson process

things we care about in this course

process

- $\mathcal{S}$ continuous, $\mathcal{T}$ discrete:
  – random walk on the reals

training a learning algorithm

- $\mathcal{S}$ continuous, $\mathcal{T}$ continuous: Markov process
  – Brownian motion

physics sim

financial models

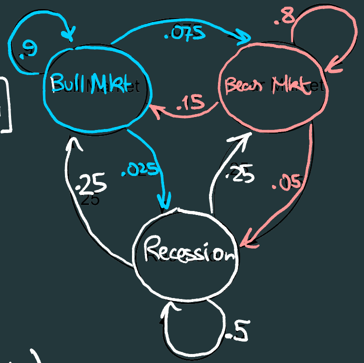
(5582)

Markov chains: basic definition (discrete time - DTMC)

Collection of rv $X_1, X_2, \ldots$

s.t

$$\mathbb{P}[X_t = x \mid X_{t-1}, X_{t-2}, \ldots X_0] = \mathbb{P}[X_t = x \mid X_{t-1}]$$



$\Rightarrow$ 2 good ways of representing

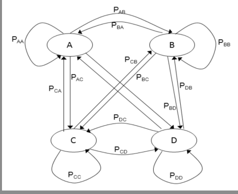
1) $\boxed{\text{Transition diagram/graph}}$

- Imagine a particle moving on a graph (weighted, directed)

- Edge $i \to j$ has probability $P_{ij} = \mathbb{P}[X_{t+1} = j \mid X_t = i]$

Property - Sum of wts of outgoing edges = 1

Markov chains: transition-diagram and transition matrix



$$\mathbf{P}=\begin{bmatrix} p_{AA} & p_{AB} & p_{AC} & p_{AD} \\ p_{BA} & p_{BB} & p_{BC} & p_{BD} \\ p_{CA} & p_{CB} & p_{CC} & p_{CD} \\ p_{DA} & p_{DB} & p_{DC} & p_{DD} \end{bmatrix}$$

Eg - Consider transition diag on left

$\Rightarrow$ matrix $P = (P_{ij})_{i,j \in S}$

For stock mkt =

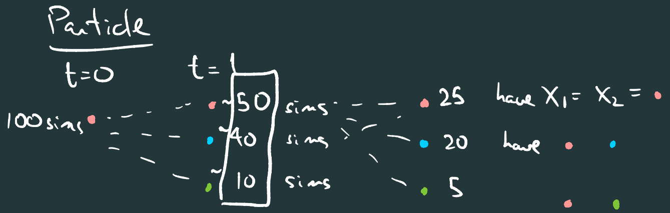
$$\begin{array}{c|ccc} & \text{Bull} & \text{Bear} & \text{Rec} \\ \hline \text{Bull} & 0.9 & 0.075 & 0.025 \\ \text{Bear} & \cdot & \cdot & \\ \text{Rec} & \vdots & & \cdot \end{array}$$

$\leftarrow$ outward edges from 'Bull Mkt' node

- P_{ij} = Entry in row i, column j = $\mathbb{P}[X_{t+1}=j \mid X_t = i]$
- Note - ALL ROWS SUM to 1 ('stochastic matrix')

  (and ALL $P_{ij} \in [0,1]$)

Natural view associated with matrix representation



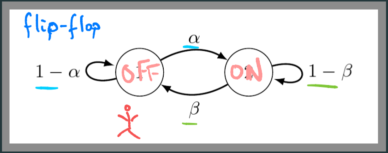
Natural Summary $\equiv$ $\pi_t \equiv$ 'distrib of sim state of particle at time t'

$$(\pi_t(\cdot)\ \pi_t(\cdot)\ \pi_t(\cdot))^T \in \mathbb{R}^n$$

Property - $\pi_t^T = \pi_{t-1}^T P$

Markov chains: two viewpoints

(Revising lecture 19: discrete-time Markov Chains)



1) Random particle view

2) Flow view | Distribution view

RP view.

- $\alpha, \beta \in [0,1]$

$X_0 = 1$

wp α $X_1 = 2$ $X_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$1-\beta$ β $X_2 = 2$ $X_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

wp $1-\alpha$ $X_1 = 1$ $X_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

α $1-\alpha$ $X_2 = 1$ $X_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

alt representation $X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ State of particle

different

$\mathbb{P}[X_0 = i] = \underline{\Pi_0} = \begin{pmatrix} \mathbb{P}[X_0 = 1] \\ \mathbb{P}[X_0 = 2] \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ prob distn over states

$\Pi_1 = \begin{pmatrix} 1-\alpha \\ \alpha \end{pmatrix}$

$$\boxed{\Pi_t^T = \Pi_{t-1}^T P}$$

α $1-\alpha$ 1 unit of water

$1-\alpha$ α β $1-\beta$

$(1-\alpha)^2 + \alpha\beta$ $\alpha(1-\beta) + \alpha(1-\alpha)$

$$\begin{bmatrix} \Pi_2(1) & \Pi_2(2) \end{bmatrix} = (1-\alpha \ \ \alpha) \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Markov chains: long-term behavior

3 possible behaviors in the <u>long term</u> (ie, What is X_t or π_t as $t \to \infty$)

1) Absorption - Particle gets 'stuck' in some state

ie, $X_{\tau+1} = X_{\tau+2} = X_{\tau+3} = \dots$ after some time τ

$$\left(\text{ie, } \lim_{t \to \infty} X_t = X \right)$$

2) Transience - Particle goes off to ∞

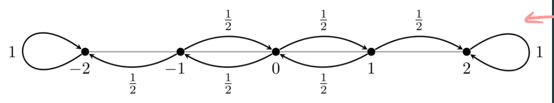
3) Recurrence - State of particle remains random (but distrib^n π_t 'settles down')

- Null recurrence ~ 'longer oscillations'
- Positive Recurrence - $\lim_{t \to \infty} \pi_t = \pi$

'starts looking like an iid r.v.' $\longrightarrow$

Markov chains: absorbing chain (Gambler's Ruin)

2 gamblers A, B, with \$2 each, play repeated games



X_t = Amt of money A has won till round t

- each game ≡ coin toss wp ½, H ⇒ B gives \$1 to A, T ≡ A gives \$1 to B
- continues till someone goes bankrupt.

Q: What happens to X_t as $t \to \infty$?

A: Gets ABSORBED in $\{-2, 2\}$

A ruined

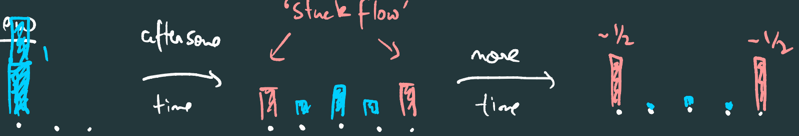
B ruined

particle view

In particular $X_\infty = \lim_{t \to \infty} X_t = \begin{cases} -2 & \text{wp } \tfrac{1}{2} \\ 2 & \text{wp } \tfrac{1}{2} \end{cases}$

Flow view

π_0

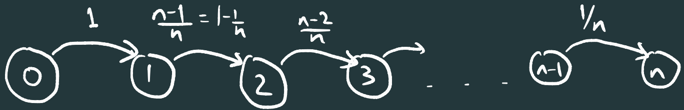


example: the coupon collector

a brand of cereal always distributes a baseball card in every cereal box, chosen randomly from a set of n distinct cards

Markov chain model for number of cards owned by a collector:

X_t = # of UNIQUE cards in first t, $X_0 = 0$, $X_t \in \{0, 1, \dots, n\}$



- Absorbing - $\lim_{t \to \infty} X_t = n$, $\lim_{t \to \infty} \pi_t = (0\,0\,0 \dots 0\,1)^T$

- Q: $\mathbb{E}[\text{Time to collect all } n \text{ coupons}] = \sum_{i=1}^{n} \underbrace{\mathbb{E}[\text{time to go from } i \text{ to } i+1 \text{ unique coupons}]}_{\text{Geom}\left(\frac{n-i}{n}\right)}$

$$= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \dots + n = n\left(\frac{1}{n} + \frac{1}{n-1} + \dots + 1\right)$$

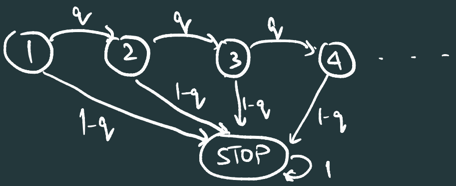
example: coin tosses and geometric rv. (Pure Absorbing)

recall the Geometric rv $p(k) = q^{k-1}(1-q) \, \forall \, k \in \{1, 2, \ldots\}$

we can view this as a Markov chain as follows: X_t = # of coin tosses till we see H (if $\mathbb{P}[\text{coin} = H] = 1-q$)

$X_0 = 1$

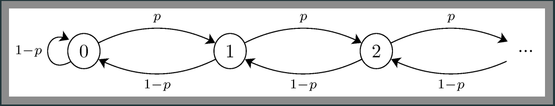
$X_\infty \triangleq$ 'state from which we jump to STOP'



$$\mathbb{P}[X_\infty = i] = q^{i-1}(1-q)$$

$$P = \begin{array}{c c} & \begin{array}{ccccc} \text{STOP} & 1 & 2 & & \end{array} \\ \begin{array}{c} \text{Stop} \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \\ 1-q & 0 & q & 0 & 0 & \cdots \\ 1-q & 0 & 0 & q & 0 & \\ 1-q & 0 & 0 & 0 & q & \\ & \vdots & & & & \end{pmatrix} \end{array}$$

Markov chains: transient/recurrent chain (Birth-Death process)



$X_0 = 0$

$X_t \in \{0,1,2,\dots\}$

If $X_t = i \Rightarrow X_{t+1} = \begin{cases} i-1 & \text{wp } 1-p \\ i+1 & \text{wp } p \end{cases}$

(transient)

- $\lim_{t\to\infty} X_t$ can either go to ∞ (if $p > 1-p$)

  get absorbed at 0 (if $p = 0$)

  get 'stabilized', ie, positive recurrent (if $p < 1-p \Rightarrow \lim_{t\to\infty} \pi_t = \pi$)

  (if $p = 1/2 = 1-p$, then null recurrent)

undesirable in sims / real systems

similar in behavior

next few slides - Give 'intuition' about $\lim_{t\to\infty} \pi_t$

Markov chains: steady-state behavior (simpler example of recurrence)

- Q: Suppose $\lim_{t\to\infty} \pi_t = \underbrace{\pi}_{}$, what does π look like?

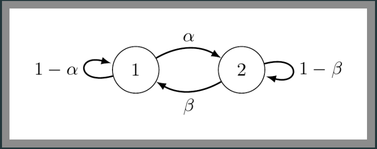
'steady-state probability'

(Such a π exists, and is unique under some condns) → Flow View

(Global Balance)

$$\boxed{\pi^T = \pi^T P}$$

example: birth-death 'with reflection at 2'



ie - π is a distribution on $\{1,2\}$ such that $\pi(1)+\pi(2)=1$

AND $(\pi(1), \pi(2)) = \begin{pmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{pmatrix} (\pi(1), \pi(2))$

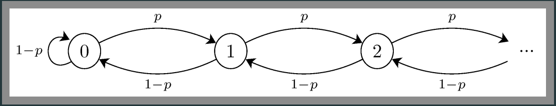
(Local Balance)

Particle View - as $t \to \infty$, both X_t and X_{t+1} have 'same' distribⁿ π

$\Rightarrow$ Avg # of times $X_t = 1, X_{t+1} = 2$ = Avg # of time $X_t = 2, X_{t+1} = 1$

$\Rightarrow \pi(1) \cdot \alpha = (1-\pi(1)) \beta \Rightarrow \pi = \begin{pmatrix} \beta/\alpha+\beta \\ \alpha/\alpha+\beta \end{pmatrix}$

Markov chains: steady-state for infinite chains



- Solving $\pi^T = \pi^T P$ is tricky (∞ system of eqns...)
- Particle view -

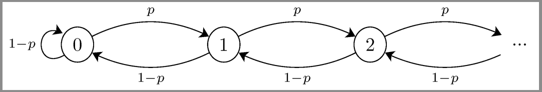
FLOW $\pi(i) \cdot p$

$\pi(i+1) \cdot (1-p)$

$$\Rightarrow \pi(i) = \left(\frac{p}{1-p}\right)\pi(i-1) = \left(\frac{p}{1-p}\right)^2 \pi(i-2) \ldots = \left(\frac{p}{1-p}\right)^i \pi(0)$$

$$\text{AND} \quad \sum_{i=0}^{\infty} \pi(i) = 1 \Rightarrow \pi(0) \underbrace{\sum_{i=0}^{\infty} \left(\frac{p}{1-p}\right)^i}_{\text{Needs to be finite} \Rightarrow p < 1-p} = 1 \quad \leftarrow \frac{1}{1 - p/1-p} = \frac{(1-p)}{(1-2p)}$$

Markov chains: the ergodic theorem



assuming $p \in (0, \frac{1}{2})$.

$$\pi(i) = \left(\frac{1-2p}{1-p}\right)\left(\frac{p}{1-p}\right)^i, \; i \geq 0$$

Why do we care about the steady-state?

Because as $t \to \infty$, X_t 'appears' like an iid sample from π

How can we formalize appears?

Consider any function $g(\cdot)$ s.t. $E_{Y\sim\pi}[g(Y)] < \infty$. Then

$$\boxed{\text{Time Avg}} \quad \underbrace{\lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=0}^{T-1} g(X_t)\right]}_{\text{Can measure via simulation}} = \underbrace{E_{Y\sim\pi}\left[g(Y)\right]}_{\text{- Can compute in closed form if } \pi \text{ is known}} \quad \boxed{\text{Space Avg}}$$

sample path of MC

sample from π