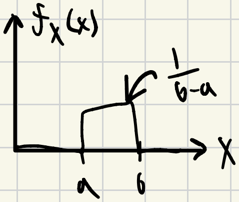
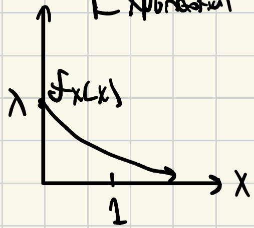
Lec 7

Continuous RVs

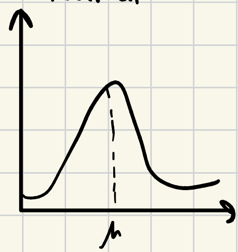
Uniform



Exponential

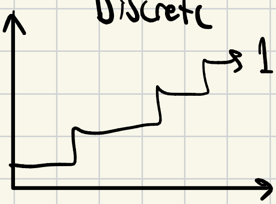


Normal

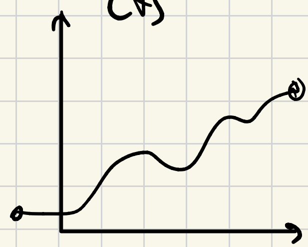


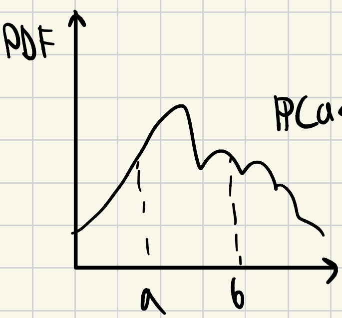
Cumulative Distribution Func

Discrete

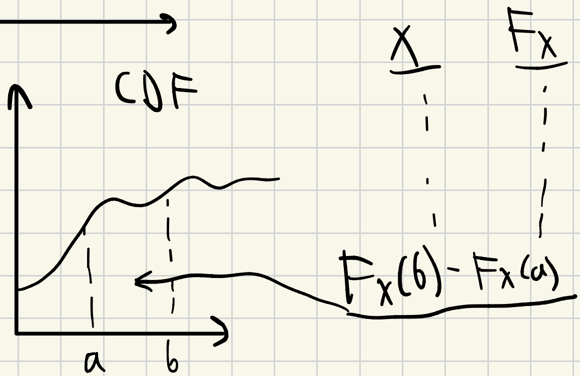


Cts





$\mathbb{P}(a \leq X \leq b)$



Continuous Bayes Rule

Discrete:

B event

$A_1, \dots, A_n$ disjoint events partition Ω

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)}{\mathbb{P}(B)}$$

$$= \frac{\mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)}{\mathbb{P}(B \mid A_1) \cdot \mathbb{P}(A_1) \cdots \mathbb{P}(B \mid A_n) \cdot \mathbb{P}(A_n)}$$

ex: test:

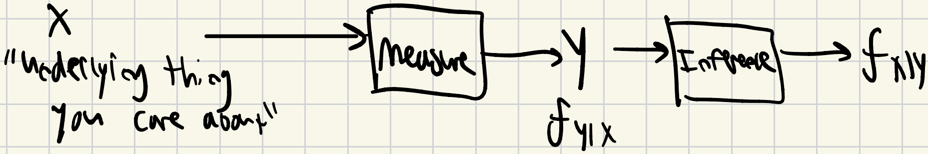
B: pos result

A_i: event of underlying condition

$\mathbb{P}(B \mid A_i) =$ accuracy

$\mathbb{P}(A_i) =$ prior

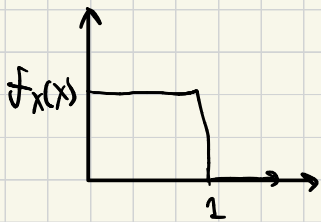
Continuous Case

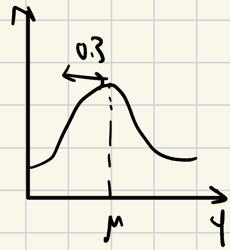


X: proportion of cells w/ Her2 receptor, Uniform $[0,1]$

Y: measurement of this proportion using measurement tool

$$P(Y|X=x) = N(\mu = x, \sigma^2 = 0.3^2)$$





$$f_{X|y} = \frac{f_{Y|X} \cdot f_X \leftarrow \text{Prior}}{f_Y \leftarrow \text{overall prob of measurements } y}$$

$$f_Y(y) = \int_x \underbrace{f_{X,Y}(x,y)\,dx}_{\text{marginal}} = \int_0^1 \underbrace{f_{Y|X}}_{\text{Conditional}} \cdot f_X\,dx = \int_0^1 N(x, 0.3^2) \cdot 1$$

$1/(b-a) = 1$

$$f_{X|y} = \frac{N(x, 0.3^2)\cdot 1}{\int_0^1 N(x, 0.3^2)\,dx} \Bigg\} \; f_{X|y}(x \in A \mid Y = 0.8), \quad A \to [0.0, 0.1]$$

Now let X: discrete RV
- 0: Her2 −
- 1: Her2 +

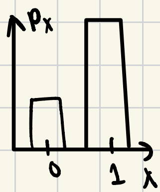
$P(X|Y=y)$

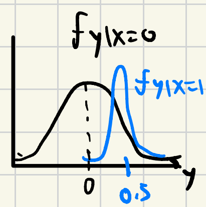
$f_{Y|X} = N(x, 0.3^2)$

$P(Y|X=0) = N(0, 0.3^2)$

but, $P(Y|X=1) = N(.5, 0.1^2)$

$P(X=0) = .1$, $P(X=1) = .9$





$$\text{Want } P(X=0|Y=y) = \frac{P(Y|X=0)\cdot P(X=0) \leftarrow \text{Prior}}{P(Y)}$$

$$P(Y) = P(Y|X=0)\cdot P(X=0) + P(Y|X=1)\cdot P(X=1)$$

$$= \frac{N(0, .3^2)\cdot 0.1}{N(0,.3^2)\cdot 0.1 + N(.5, .1^2)\cdot 0.9}, \quad \text{if } y = 0 \approx .99$$

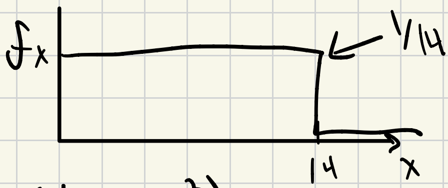
$y = .25 \approx .37$

$y = .5 \approx 0.1$

COVID Tracing

Y: measurement of viral load

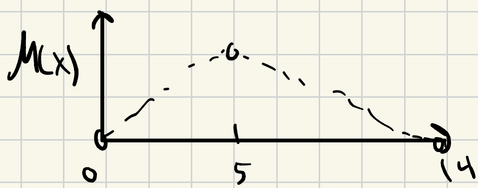
X: number of days since Covid → want to infer



$$f_{Y|X} \sim N(M(x), \sigma^2)$$

$$M(x) = \begin{cases} x & \text{if } 0 \le x \le t_{peak} \\ t_{peak} - 0.5\cdot(x - t_{peak}) & \text{if } t_{peak} \le x \le 14 \end{cases}$$

$t_{peak} = 5$



Observe $Y = 4$

$$f_{X|Y}(x \mid Y=4) = \frac{f_{Y|X}(y|x)\, f(x)}{f(y)}$$

$$= \frac{\dots \cdot 1/14}{}$$

Don't care about non dependent parts of X → $f_{Y|X} = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-M(x))^2/2\sigma^2}$

$X = 0$, $M(0) = 0$

$X = 4$, $M(4) = 1$

$M(0) \to 3.3\cdot 10^{-9}$ in the exponential

$M(13) \to .011$ in the exponential

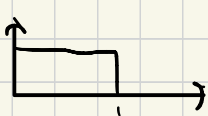
Ch 4 Derived Dists

X, Y $\quad Y = g(X)$ PDF of X: f_X What f_Y

$f_X(x) \to f_Y = f_X(g^{-1}(y))$ if g invertible?

Example

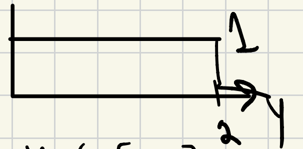
$Y = 2X$, $X \sim$ uniform $[0,1]$



$f_X = 1$ for $[0,1]$

$Y/2 = X$

$f_Y \overset{?}{=} \underbrace{f_X(Y/2)}$



$f_Y = 1$ for $\forall y \in [0,2]$, this pdf isn't normalized

but if $Y = X + c$ or $Y = -X$

Derived Dists

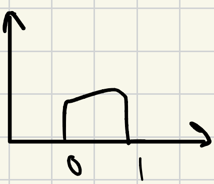
pdf of X

1) Find F_Y using $F_Y(y) = \mathbb{P}(g(X) \le y) = \int_{\{x \mid g(x) \le y\}} f_X(x)\,dx$

2) $F_Y(y) \longrightarrow$ differentiate to get $f_Y(y)$, $f_Y = \dfrac{dF_Y(y)}{dy}$

Example: $X \sim$ Uniform $[0,1]$, $Y = \sqrt{X}$

1) CDF of y: $F_Y = \mathbb{P}(Y \le y) = \mathbb{P}(\sqrt{X} \le y) = \mathbb{P}(X \le y^2) = y^2$



$F_X(x) = x$

2) $\dfrac{dF_Y}{dy} = \dfrac{d}{dy}(y^2) = 2y = f_Y$

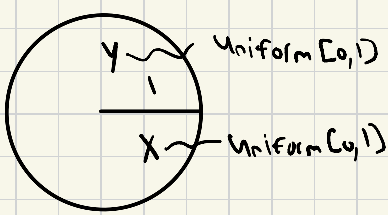
ex 2: $Y = aX + b$ Show $f_Y = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

$F_Y = \mathbb{P}(Y \le y) = \mathbb{P}(aX + b \le y) = \mathbb{P}\left(X \le \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$

$\dfrac{dF_Y}{dy} = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$

EX: 2 archers shooting

PDF: distance of center-losing shot



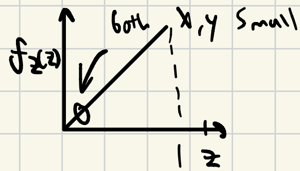
1) $X \to$ Shooter 1 $\quad Z = \max(X,Y)$

$Y \to$ Shooter 2

$$\forall z \in [0,1], \ \mathbb{P}(X \le z) = \mathbb{P}(Y \le z) = z$$

$$F_Z(z) = \mathbb{P}(\max(X,Y) \le z)$$
$$= \mathbb{P}(X \le z, Y \le z) = \mathbb{P}(X \le z) \cdot \mathbb{P}(Y \le z)$$
$$= F_X(z) \cdot F_Y(z) = z^2$$

2) $\frac{dF_Z}{dz} = 2z \quad 0 \le z \le 1$



if $Z = X+Y$

$$p_Z(z) = \mathbb{P}(X+Y=z)$$

4.2 - 4.3

- $\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$

When $\mathrm{Cov}(X,Y) = 0$, X and Y uncorrelated

- $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(X, aY+b) = a \cdot \mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X, Y+Z) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$
- $\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$
- if $X_1, \dots, X_n$ finite variance

$$\mathrm{Var}(X_1+X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\,\mathrm{Cov}(X_1, X_2)$$

4.3

$$\mathbb{E}[\mathbb{E}[X|Y]] = \begin{cases} \sum_y \mathbb{E}[X|Y=y]\, p_Y(y) & \text{discrete } Y \\ \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y]\, f_Y(y)\, dy & \text{continuous } Y \end{cases}$$

$\Rightarrow$ if X has finite expectation
$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

Ex: Stick length ℓ, break it uniformly, keep left, repeat. $\mathbb{E}$(length after 2 breaks?)
Y = length after first break
X = length after 2nd

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[\tfrac{Y}{2}] = \tfrac{1}{2}\mathbb{E}[Y] = \tfrac{1}{4}\ell$$

Conditional Expectation as estimator

if Y gives info about X the estimator of X given Y
$$\hat{X} = \mathbb{E}[X|Y]$$

estimation error
$$\tilde{X} = \hat{X} - X$$

- $\mathrm{Cov}(\hat{X}, \tilde{X}) = 0 \Rightarrow \mathrm{Var}(X) = \mathrm{Var}(\tilde{X}) + \mathrm{Var}(\hat{X})$
- $\mathrm{Var}(X|Y) = \mathbb{E}[(X - \mathbb{E}[X|Y])^2 | Y] = \mathbb{E}[\tilde{X}^2 | Y]$, note $\mathbb{E}(\tilde{X}) = 0$
- $\mathrm{Var}(\tilde{X}) = \mathbb{E}(\tilde{X}^2) = \mathbb{E}[\mathbb{E}[\tilde{X}^2|Y]] = \mathbb{E}[\mathrm{Var}(X|Y)] \Rightarrow$

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathbb{E}[X|Y])$$

Lecture 8

if $Y = g(X)$, know f_X

derived dists: ① $F_Y(y)$ ② $f_Y(y) = \frac{dF_Y}{dy}$

$Z = X+Y$, p_Z, $F_Z = \mathbb{P}(Z \le z)$, X, Y indep

$$\mathbb{P}(Z \le z \mid X = x) = \mathbb{P}(X+Y \le z \mid X = x)$$
$$= \mathbb{P}(Y \le z - x)$$

Approach:
$\mathbb{P}(Z \le z \mid X = x)$
mult by $F_X \to F_{Z,X}$
marginals $\to F_Z$

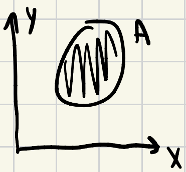
$$f_{Z|X} = f_Y(z-x)$$

$$f_{X,Z} = f_{Z|X} \cdot f_X = f_Y(z-x) \cdot f_X$$

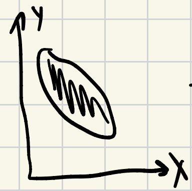
$$f_Z = \int_{-\infty}^{\infty} f_{X,Z}\, dx = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx \leftarrow \text{convolution}$$

Cov

$\text{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, can be pos, neg



X, Y uniform dist in A
+ cov



− cov

1) $\text{Cov}(X, X) = \text{Var}(X)$
2) $\text{Cov}(X, aY + b) = a\text{Cov}(X, Y)$
3) $\text{Cov}(X, Y+Z) = \text{Cov}(X,Y) + \text{Cov}(X,Z)$

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \to [-1, 1] \text{ corr. coeff.}$$

n indep coin tosses $X + Y = n$
$X = \#H$
$Y = \#T$

want $\rho(X,Y) = \frac{\mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]}{\cdots}$

$$= \frac{-\mathbb{E}[(X - \mathbb{E}(X))^2]}{\cdots} = \frac{-\text{Var}(X)}{\cdots}$$

$$= \frac{-\text{Var}(X)}{\sqrt{\text{Var}(X)}^2} = -1$$

if $Y = cX + b \to \rho = \pm 1$ exercise

Var of Sum of R.V

$X = X_1 + \dots + X_n$

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2)$$

$Cov(X_1,X_2) = 0$ if X_1, X_2 indp

Example n People throw their hat into a box

$X = \#$ of ppl that get their own hat back

$\mathbb{E}(X)$, $Var(X)$

X_i = Person i gets their hat $\to \mathbb{1}_i$

→ Bernoulli param $p = 1/n$

$\mathbb{E}(X) = 1$

$$Var(X) = \sum var(X_i) + \sum cov(X_i, X_j)$$

$$\sum \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 \qquad \mathbb{E}(X_i, X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$$

$= \sum \mathbb{E}(X_i) - \mathbb{E}(X_i)^2$

$= \sum \frac{1}{n} - \frac{1}{n^2}$

$= 1 - \frac{1}{n} + 2\left(\sum \mathbb{E}(X_i, X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)\right)$

$\mathbb{P}(X_i = 1, X_j = 1) = \mathbb{P}(X_i = 1 \mid X_j = 1)\cdot \mathbb{P}(X_j = 1) = \frac{1}{n-1}\cdot\frac{1}{n}$

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

Conditioning

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}(X \mid Y = y)\cdot \mathbb{P}(Y = y)$$

Lec 9

$\mathbb{E}(X|Y) \to RV$: $\mathbb{E}[\mathbb{E}(X|Y)] = \mathbb{E}[X]$

$X = \underbrace{\mathbb{E}[X|Y]}_{\hat{X}} + \underbrace{\text{error}(\mathbb{E}[X|Y] - X)}_{\tilde{X}}$ = estimator + error

$\hat{X} = \mathbb{E}[X|Y]$

$\tilde{X} = \hat{X} - X$

Observe : $\mathbb{E}[\tilde{X}|Y] = \mathbb{E}[\hat{X} - X|Y] = \mathbb{E}[\hat{X}|Y] - \mathbb{E}[X|Y]$

$= \hat{X} - \hat{X} = 0$

$\hat{X}$ uncorr w/ $\tilde{X}$ $\to$ $\frac{\mathrm{Cov}(\hat{X},\tilde{X})}{\sqrt{\mathrm{Var}(\hat{X})\mathrm{Var}(\tilde{X})}} \to 0 - 0 = 0$

$\mathrm{Var}(X) = \mathrm{Var}(\hat{X}) + \mathrm{Var}(\tilde{X}) + 0$

Conditional Variance: $\mathrm{Var}(X|Y) = \mathbb{E}[(X - \mathbb{E}(X|Y))^2|Y]$

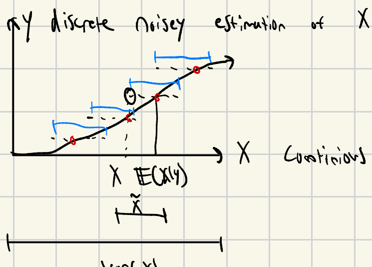
$X - \mathbb{E}(X|Y) = x - \hat{x} = \tilde{x}$

So $\mathrm{Var}(X|Y) = \mathbb{E}(\tilde{X}^2|Y)$

$\mathrm{Var}(\tilde{X}) = \mathbb{E}(\tilde{X}^2) - \mathbb{E}(\tilde{X})^2 = \mathbb{E}(\tilde{X}^2) = \mathbb{E}[\mathbb{E}(\tilde{X}^2|Y)]$

$\mathrm{Var}(\hat{X}) = \mathrm{Var}(\mathbb{E}(X|Y))$

Law of Total Variance: $\mathrm{Var}(X) = \underbrace{\mathrm{Var}(\mathbb{E}(X|Y))}_{\text{across group variance}} + \underbrace{\mathbb{E}(\mathrm{Var}(X|Y))}_{\text{within group variance}}$



Example : Discrete Y: School of Student

X: math test score of student

$\mathrm{Var}(\mathbb{E}(X|Y))$

Variation of mean score of students given schools

Coin flips : X: # H in n flips $\to$ Binomial RV $\mathbb{E}(X) = np$ $\mathrm{Var}(X) = np(1-p)$

Y: Prob(head) uniform $[0,1]$

$\mathbb{E}(X|Y=y) = ny$

$\mathrm{Var}(X|Y) = ny(1-y)$

$\mathrm{Var}(X) = \mathbb{E}(\mathrm{Var}(X|Y)) + \mathrm{Var}(\mathbb{E}(X|Y))$

$= \mathbb{E}(nY(1-Y)) + \mathrm{Var}(nY)$

$= n\mathbb{E}(Y) - n\mathbb{E}(Y^2) + n^2\mathrm{Var}(Y)$

$= n/2 - n/3 = \boxed{n/6 + n^2/12}$

Transforms (Moment generating)

example: 2 pizza places
orders per hr Poisson: Shop 1: λ, Shop 2: μ

Poisson PMF $\mathbb{P}(X=k)\ e^{-\lambda}\dfrac{\lambda^k}{k!}$

$$\mathbb{P}(X+Y=k) = \sum_{i=0}^{k} \mathbb{P}(X=i)\cdot\mathbb{P}(Y=k-i) \quad \text{discrete convolution}$$

$$= \sum_{i=0}^{k} e^{-\lambda}\frac{\lambda^k}{k!} e^{-\mu}\frac{\mu^{k-i}}{(k-i)!} \quad \text{horrible to solve}$$

moment generating transform for X

$$M_X(s) = \mathbb{E}[e^{sX}] \rightarrow \sum_x e^{sx} p_X(x) \quad \text{discrete}$$

$$\rightarrow \int_{-\infty}^{\infty} e^{sx} f_X(x)\,dx \quad \text{continuous}$$

Example

$$p_X(x) = \begin{cases} 1/2 & x=2 \\ 1/6 & x=3 \\ 1/3 & x=5 \end{cases} \qquad M_X(s) = e^{2s}/2 + e^{3s}/6 + e^{5s}/3$$

Example Poisson RV

$$M_X(s) = \sum_{x=0}^{\infty} e^{sx}\cdot\frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^{s}\lambda)^x}{x!} = e^{-\lambda}\cdot e^{e^{s}\lambda} = \boxed{e^{\lambda(e^s-1)}}$$

transform $\rightarrow$ moments

cts case $M_X(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x)\,dx$

1) $\dfrac{d}{ds}M_X(s) = \int_{-\infty}^{\infty} x e^{sx} f_X(x)\,dx$, if $\dfrac{d^2}{ds^2}M_X(0)$ then $\mathbb{E}(X^2)$

2) at $s=0$ $\int_{-\infty}^{\infty} x f_X(x)\,dx = \mathbb{E}(X)$

for Poisson: $M_X(s) = e^{\lambda(e^s-1)}$, $\dfrac{d}{ds} = e^{-\lambda}\lambda e^{s} e^{\lambda e^s}$, at $s=0 \rightarrow \lambda$

if $Z = X+Y$ indep X,Y $\quad M_Z(s) = \mathbb{E}[e^{sZ}] = \mathbb{E}[e^{sX+sY}] = \mathbb{E}[e^{sX}]\cdot\mathbb{E}[e^{sY}] = M_X(s)M_Y(s)$

if $X_1,\dots,X_n$ Bernoulli $\rightarrow X = \sum X_i$ is binomial, $\mathbb{E}[e^{sX}] = \mathbb{E}[e^{s\sum X_i}] = \prod_{i=0}^{n}\mathbb{E}[e^{sX_i}]$

$M_{X_i}(s) = \mathbb{E}[e^{sX_i}] = e^{s0}(1-p) + e^{s}p = 1-p+pe^s$

$= (1-p+pe^s)^n$

$$\frac{d}{ds}(1-p+pe^{s})^{n} = n(1-p+pe^{s})^{n-1}\cdot pe^{s} \quad \text{at} \quad s=0 \quad \rightarrow np$$

---

a) Show $\mathbb{P}(X \geq a) \leq e^{-sa} M(s)$

$$\mathbb{P}(e^{sX} \geq e^{sa}) \leq \frac{\mathbb{E}(e^{sX})}{e^{sa}} \quad \text{True by Markovs}$$

b) $\mathbb{P}(X \leq a) \leq \dfrac{\mathbb{E}[e^{sX}]}{e^{sa}} \leftarrow a \in \mathbb{N}$

$s > 0$

$$\underbrace{1-\mathbb{P}(e^{sX} \geq e^{sa})} \leq \frac{\mathbb{E}[e^{sX}]}{e^{sa}}$$

$$\frac{1}{2} \leq \frac{\mathbb{E}[e^{sX}]}{e^{sa}}$$

$$e^{sa} \leq \mathbb{E}[2e^{sX}]$$

Bernoulli Review

- Discrete time: Success prob p
- # of arrivals in n time slots $\sim$ Binom
- Interarrival $\sim$ Geo
- Time to k arrivals $\sim$ Pascal
- Memorylessness

$Y_k =$ Time until kth arrival $= T_1 + \dots + T_k$, $T_i =$ indep geo(p)

$$\mathbb{P}(Y_k = t) = \binom{t-1}{k-1} p^k (1-p)^{t-k}, \quad t = k, k+1, \dots$$

↑ pascal PMF

Consider we start watching @ t

$$B_1, \dots, B_t, \underbrace{B_{t+1}, B_{t+2}}$$
$$ \; X_1 = B_{t+1}, \; X_2 = B_{t+2}$$

$B_i \sim$ Bernoulli(p)

$X_i \sim$ Bernoulli(p)

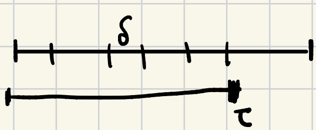
# of arrivals in disjoint time intervals are indep

for small δ prob of k arrivals in it where $\lambda =$ "arrival rate"

$$\mathbb{P}(k, \delta) = \begin{cases} 1 - \lambda\delta & \text{if } k = 0 \\ \lambda\delta & k = 1 \\ 0 & k > 1 \end{cases} + O(\delta^2)$$

$$\lim_{\delta \to 0} \frac{\mathbb{P}(1, \delta)}{\delta} = \lambda$$

$$\mathbb{E}[\#\text{ of arrivals } [0, \delta]] = \lambda\delta$$



$n = \frac{\tau}{\delta}$ $\quad p = \lambda\delta = \frac{\lambda\tau}{n}$

Binary arrivals since slots so small

$\mathbb{P}(k \text{ arrivals}) = \binom{m}{k}\left(\frac{\lambda\tau}{m}\right)^k\left(1-\frac{\lambda\tau}{m}\right)^{m-k}$

$\delta \to 0$

$m \to \infty$

$\lambda\tau = mp$ ← $\mathbb{E}$ success in bernoulli process

↑

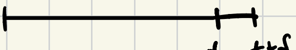
Arrival rate · time

$\mathbb{E}$ arrivals

Example: $\lambda = 5$ emails per hr, you check every $1/2$ hr

$\lambda t = 2.5$

$\mathbb{P}(0, 1/2) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = \frac{1}{e^{2.5}} = \mathbb{P}$ 0 messages seen



$f_{Y_k}(t)\delta = \mathbb{P}(t \le Y_k \le t+\delta)$

Need $k-1$ arrivals in $[0,t]$

So $\mathbb{P}(k-1 \text{ in } [0,t]) \cdot \lambda\delta$

$= \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} \cdot \lambda\delta$

Y_k = time of kth arrival ~ Erlang

$$f_{Y_k}(y) = \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!}$$

time of first arrival $(k=1)$: exponential $\lambda e^{-\lambda y}$ $y \ge 0$

the time to the next arrival is indep of the past.

$Y_2 = T_1 + T_2$ $\quad T_i \sim \exp(\lambda)$ indep