1st Qaqish Review

(1) - 3 problems, 2 hrs per problem
- biggest source of mistakes → not reading the question properly

→ spend 30 min reading the problem
↳ you <u>need</u> to understand the problem
→ "I wish we could give negative grades but we cannot". If you don't understand, you won't get any points

(2) Do not jump to quick conclusions
example: there might be a way to get expectation w/o doing integral
example: if you have conditional, you can find expectation, variances, co-variances by double expectation
↳ one of the most useful for all of statistics not just tests

(3) Don't make assumption that if you can't do part a, you can't do part b

(4) Is this estimator biased or unbiased?
→ knee jerk → calculate expected value
→ Q. doesn't say calculate expected value

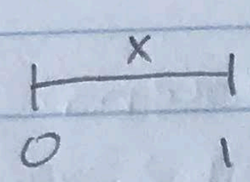
Example: $\bar{X}^2$ estimator of μ^2
$\mu = E(X)$

Jensen's inequality: $g(X)$ convex, $E[g(X)] > g(E[X])$
$g(X) = X^2$
$E(X^2) > (EX)^2 = \mu^2$

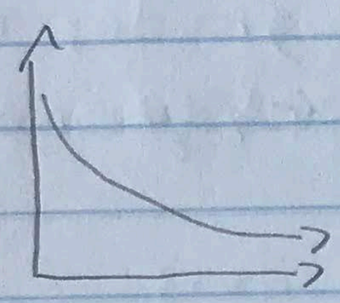
- Ask yourself, if it is 0.1, what can I do? What about 0.2? Try to find a pattern for general expression

- Try to not use your cheat sheet

$\boxed{2017}$

1) 

$X \perp Y$



$f_Y(y) = e^{-y}, \; y > 0$

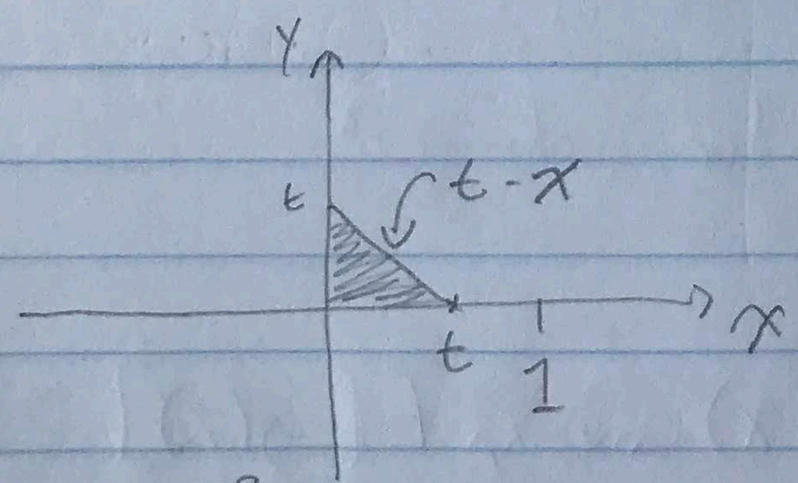
a) and b) are trying to get you to find the expression of T, b/c it is both parts

- "If you don't remember it off the top of your head, don't use it. You probably don't understand it"

---

a) $X \in (0,1)$ $\quad T = X + Y$
$Y \in (0, \infty)$

Given $t \in (0,1)$, $P(X+Y \le t)$



What is the region that corresponds to this event?

$\{X+Y \le t\}$
$\{Y \le t - X\}$

Don't think of probabilities, think of events

$\boxed{2017}$

1a)

You have the joint probability is a volume (independent, multiply together) and integrate over this region

Integrate:

x from 0 to t,
for each x, y from 0 to $t-x$

$$\underbrace{\int_{x=0}^{x=t} \int_{y=0}^{y=t-x}}_{\text{most errors are in the ranges!!}} f_{X,Y}(x,y)\,dy\,dx$$

$$\int_0^{t-x} e^{-y}\,dy$$
$$= 1 - e^{-(t-x)}$$ ← (do integral or know it is exponential CDF

$$\int_0^t \{1 - e^{-(t-x)}\}\,dx$$

comes out, constant

$$t - e^{-t}\int_0^t e^x\,dx = t - e^{-t}(e^t - 1) = t - 1 + e^{-t}$$

$0 < t < 1$

When calculating propabilities, check: that it is always positive and less than 1

$t-1$ is negative, since $t \in (0,1)$

$e^{-t} = 1 - t + \frac{t^2}{2} + \dots$ ← taylor series, which is $> t-1$, so this probability is positive

$\boxed{2017}$

1b) Given $t \in (1, \infty)$
Different

c) $T = X + Y$
$ET = EX + EY = \frac{1}{2} + 1$
$Var(T) = Var(X) + Var(Y) + 0$
b/c $X \perp Y$ so covariance is 0

$Corr(X, T)$

first find covariance:
$cov(X, T) = cov(X, X+Y)$
$= cov(X, X) + cov(X, Y)$
$= var(X) + 0$
$= \frac{1}{12}$ ← b/c X is uniform

$$Corr(X, T) = \frac{1/12}{\sqrt{\frac{1}{12} \cdot \frac{13}{12}}} = \frac{1}{\sqrt{13}}$$

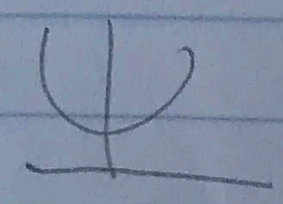
sq root of product of the variances

linear regression of X on T, on the population

e) $E[a + bT - X] = 0$
$= a + b(1.5) - 0.5 = 0$
$b = \frac{0.5 - a}{1.5}$

plug in here to get a

$Var(a + bT - X) = 0 + var(bT) + var(X) - 2cov(bT, X)$
$= \frac{13}{12}b^2 + \frac{1}{12} - 2b\left(\frac{1}{12}\right)$
$= 13b^2 - 2b + 1 \longrightarrow$



minimize ↑, take derivative:
$26b - 2 = 0$
$b = 13$

☆ Good notation is critical on the exam!! ☆

$\boxed{2017}$

1f) Problem: do something for general n
Start w/ Z_1 then try Z_2

R	B
3	3
Z_n	$6-Z_n$

$P(\text{Red Ball} \mid Z_n \text{ currently})$
$= Z_n/6$

$P(\text{Blue} \mid Z_n) = 1 - \frac{Z_n}{6}$

	Red:	Blue
Draw Red $\Rightarrow$ replace w/ blue	$Z_n - 1$	$6 - Z_n + 1$
Draw Blue $\Rightarrow$ replace w/ red	$Z_n + 1$	$6 - Z_n - 1$

It cannot stay at Z_n, it has to go up by 1, or down by 1

$P(Z_{n+1} = Z_n - 1 \mid Z_n) = \frac{Z_n}{6}$

$P(Z_{n+1} = Z_n + 1 \mid Z_n) = 1 - \frac{Z_n}{6}$

If you see something on right hand side of probability, you know it must be conditional. It is a R.V

$$E[Z_{n+1} \mid Z_n] = (Z_n - 1)\frac{Z_n}{6} + (Z_n + 1)\left(1 - \frac{Z_n}{6}\right)$$
$$= -\frac{Z_n}{6} + Z_n + 1 - \frac{Z_n}{6} = 1 + \frac{2Z_n}{3} = E[Z_{n+1} \mid Z_n]$$

$Z_0 = 3$, $E[Z_1 \mid Z_0 = 3] = 3$
$E[Z_2 \mid Z_1] = 1 + \frac{2Z_1}{3}$ } pattern, then take double expectation

Stochastic Process

$\boxed{2017}$

1f)

$$E[Z_{n+1}] = 1 + \frac{2}{3}E(Z_n)$$

$$a = 1 + \frac{2}{3}a$$

$$a = 3$$

$$E[Y] = E\Big[\underbrace{E[Y \mid X]}_{\text{function of } X}\Big]$$

$$E[Z_{n+1}] = E\Big[E[Z_{n+1} \mid Z_n]\Big]$$

2018 #3 Review 6/12/19

3) $\sim Bern(p)$

a) $L(p|y) = \prod_{i=1}^{n} \{ p^{y_i} (1-p)^{1-y_i} \}$

$$= p^{\sum_{i=1}^{n} y_i} (1-p)^{n - \sum_{i=1}^{n} y_i}$$

$$\ell(p|y) = \left(\sum_{i=1}^{n} y_i\right) \log(p) + \left(n - \sum y_i\right) \log(1-p)$$

$$\frac{\partial \ell}{\partial p} = \frac{\sum_1^n y_i}{p} - \frac{n - \sum_1^n y_i}{1-p} = 0$$

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}$$

$$\frac{\partial^2 \ell(p|y)}{\partial p^2} = \frac{-n\bar{y}}{p^2} - \frac{n(1-\bar{y})}{(1-p)^2}$$

local max, since continuous, also a global max

(plug in) at $p = \bar{y} \Rightarrow \frac{-n}{\bar{y}} - \frac{n}{1-\bar{y}} = \frac{n\bar{y} - n\bar{y} - n}{\bar{y}(1-\bar{y})} = \left\{ \frac{-n}{\bar{y}(1-\bar{y})} \right\}$

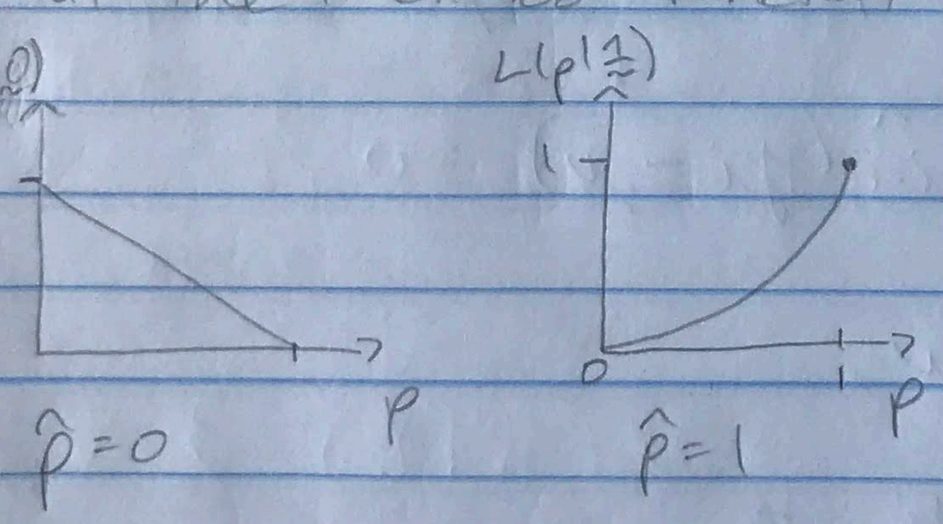
So, you need to look at the likelihood function

some probability that $\bar{y} = 0$ or 1

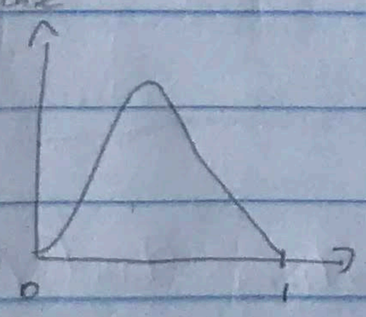
$P(\bar{y} = 1) = p^n > 0$

$P(\bar{y} = 0) = (1-p)^n > 0$

first consider these two cases



(still fits general case b/c $\hat{p} = \bar{y}$)



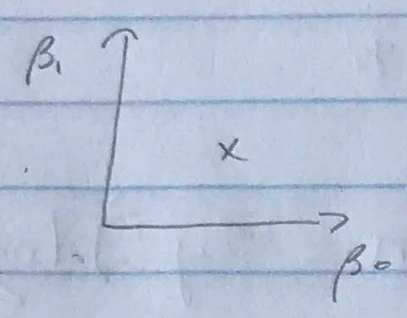
only applicable in this case

#3, 2018

3c) $p_i = E[Y_i] = \dfrac{1}{1+e^{\beta_0 - \beta_1 x_i}}$ $\quad x_1, \dots x_n$ given

always need to think about

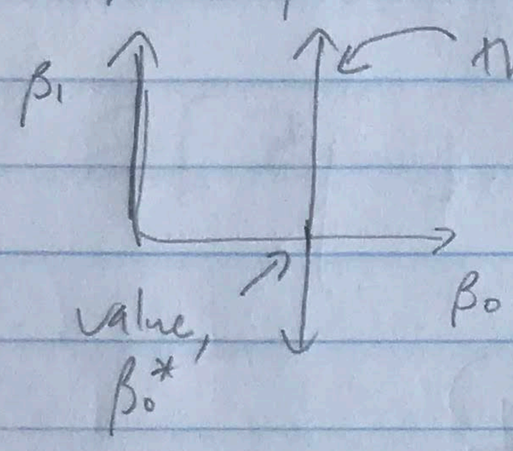
if β_0 and β_1 unknown, then the parameter space:



$\mathbb{R}^2$

x ← any point is a posibility, parameter space is $\mathbb{R}^2$

But, if β_0 is fixed/known/given, then parameter space



this line is the parameter space, $\mathbb{R}^1$

Show: $\sum_{i=1}^{n} x_i(y_i - p_i) = 0$

Problem should say that x_i's are given.

Other example: $Y_1, \dots, Y_n$ iid, $E[Y_i] = \mu$

$\hat{\mu}$ satisfies: $\sum_{i=1}^{n} (y_i - \mu) = 0$

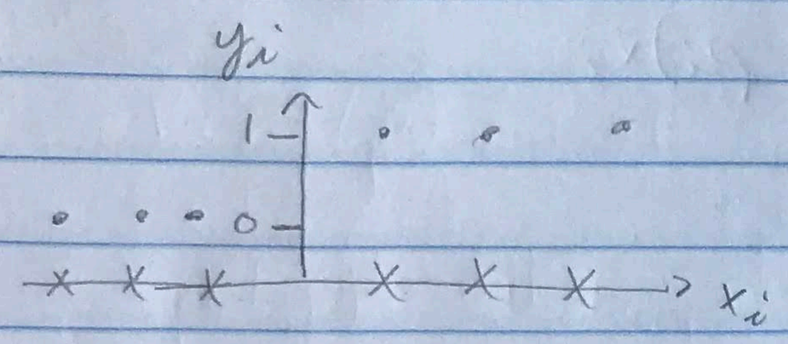
$\sum_{i=1}^{n} y_i = n\mu$

$\mu = \dfrac{\sum_1^n y_i}{n} = \bar{y}$

This problem is making the assumption that MLE of β_1 exists. It might not

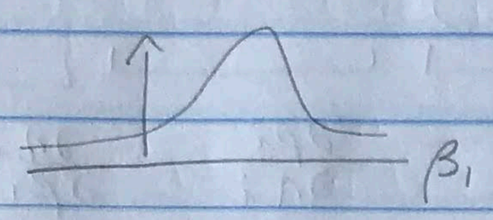
2018, #3

c)



Here, MLE is ∞ or doesn't exist

In this case, if you increase β_1, the MLE will increase?



← make this assumption about what the MLE looks like.

$$L(\beta_1 \mid y) = \prod_{i=1}^{n} \{ p_i^{y_i} (1-p_i)^{1-y_i} \}$$

independant

$$= \prod_{i=1}^{n} \left\{ \left(\frac{p_i}{1-p_i}\right)^{y_i} (1-p_i) \right\}$$

$$\ell(\beta_1 \mid y) = \sum_{i=1}^{n} \{ y_i \operatorname{logit} p_i + \log(1-p_i) \}$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \operatorname{logit} p_i = \eta_i$$

$$\eta_i = -\beta_0 + \beta_1 x_i$$

[Chain rule: $\frac{\partial}{\partial \beta_1} = \frac{\partial}{\partial \eta_i} \frac{\partial \eta_i}{\beta_1}$

$$p_i = \frac{1}{1+e^{-\eta_i}}$$

$$\frac{\partial p_i}{\partial \beta_1} = \underbrace{\frac{\partial p_i}{\partial \eta_i}}_{p_i(1-p_i)} \underbrace{\frac{\partial \eta_i}{\partial \beta_1}}_{x_i}$$

reciprocal

$$\frac{\partial \eta_i}{\partial p_i} = \frac{\partial}{\partial p_i} \log \frac{p_i}{1-p_i} = \frac{1}{p_i(1-p_i)}$$

2018 #3

c) $$\frac{\partial \ell}{\partial \beta_1} = \sum_{i=1}^{n} y_i \cancel{p_i(1-p_i)} x_i$$

$$\frac{\partial}{\partial p_i} \quad \frac{\partial p_i}{\partial \eta_i} \quad \frac{\partial \eta_i}{\partial \beta_1}$$

$$\frac{\partial \ell}{\partial \beta_1} = \sum_{i=1}^{n} \underbrace{\left\{ \frac{y_i}{p_i(1-p_i)} - \frac{1}{1-p_i} \right\}}_{\frac{\partial}{\partial p_i}} \underbrace{p_i(1-p_i)}_{\frac{\partial p_i}{\partial \eta_i}} \underbrace{x_i}_{\frac{\partial \eta_i}{\partial \beta_1}}$$

$$= \sum_{i=1}^{n} \frac{y_i - p_i}{\cancel{p_i(1-p_i)}} \cancel{p_i(1-p_i)} x_i$$

cancellations ok if $p_i \neq 0, 1$

$$= \sum_{i=1}^{n} x_i (y_i - p_i)$$

These are all exponential families

d) $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$

Didn't ask for UMP. Said, find asymptotic test, large sample. It said find <u>an</u> asymptotic → it doesn't matter: LRT, score, Wald

↳ based on estimates

1 parameter problem. $\frac{\partial \ell(\beta_1 | \underline{Y})}{\partial \beta_1} = \sum_{i=1}^{n} x_i (Y_i - p_i)$ ↑ R.V.

2nd path

Information function is variance of score function?

$$I(P_i) = \sum_{i=1}^{n} x_i^2 p_i(1-p_i)$$

2018 #3

d) Wald:

$$Z = \frac{\hat{\beta}_1 - 0}{\sqrt{\frac{1}{I(\hat{\beta}_1)}}} \sim N(0,1)$$

p_i depends on β_1, so plug in $\hat{\beta}_1$ for β_1

LRT: $T = 2\{\ell(\hat{\beta}_1) - \ell(0)\} \sim \chi^2_{(1)}$ b/c there is one constraint

$\ell(0) < \ell(\hat{\beta}_1)$, so don't need -2

Reject H_0 if $T > \chi^2_{1,1-\alpha}$

$1-\alpha$ quantile of the χ^2_1 distribution

Reject H_0 if $Z > |\Phi^{-1}(1-\frac{\alpha}{2})|$

e) $\theta = \frac{K}{\beta_1}$, $\hat{\theta} = \frac{K}{\hat{\beta}_1}$

Some students try to find the standard error, need Delta method

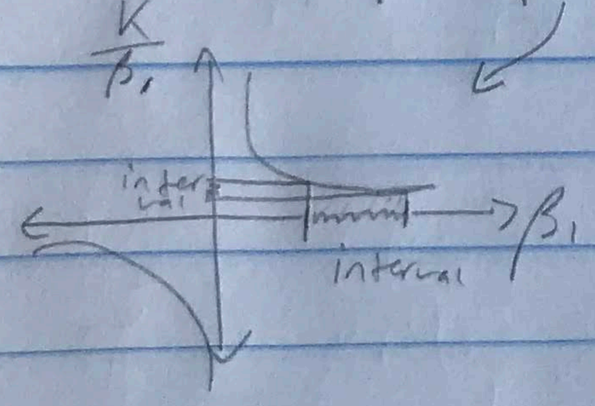
$\frac{K}{\beta_1}$ is monotone function of β_1

So, make CI for β_1 itself

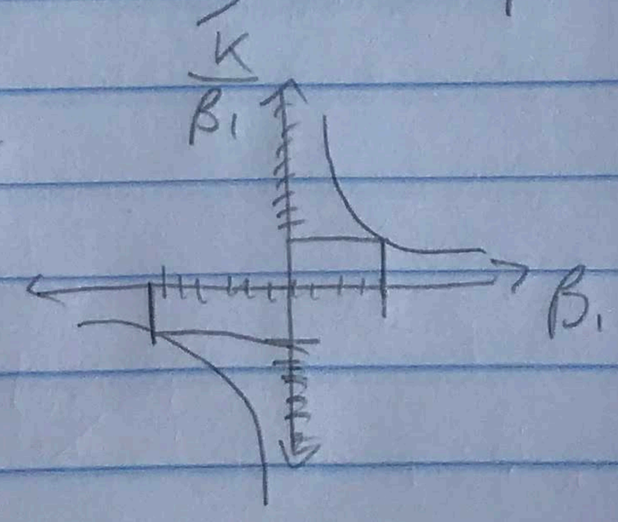
$$\hat{\beta}_1 \pm Z_{1-\alpha/2}\sqrt{\frac{1}{I(\hat{\beta}_1)}}$$

Then, apply $\frac{K}{\beta_1}$ transformation to upper and lower limits

assuming K is positive



or



If β_1 interval contains 0, then the K/β_1 interval is union of two infinite intervals

MLE

$U(0, \theta)$

MLE

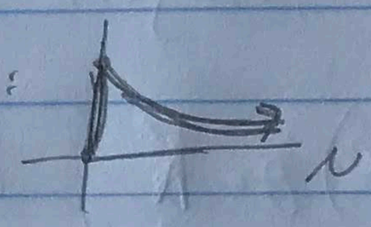
MLE is product of two things, Indicator and the function

UMVUE problems → have to show completeness → read about it → students don't do it right

can discuss it

Think about the boundary, about the shape of the likelihood

Poi: $\frac{e^{-\lambda} \lambda^{y}}{y!}$ if all y are 0: $e^{-\lambda}$:



Qaqish Review 6/19/19

$\boxed{2017}$ 2c)

- write down pdf of Y_1 to Y_n
  Gamma, $\alpha = 1 \longrightarrow$ exponential:

$$f_Y(y \mid \alpha = 1, \beta) = \frac{e^{-y/\beta}}{\beta} \qquad y > 0, \quad \alpha = 1, \beta > 0$$

b/c exp fam + can'd ratio

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y} \mid \alpha = 1, \beta) = \beta^{-n} e^{-\underbrace{\sum_{i=1}^{n} y_i}_{\text{DATA } y}/\beta}$$

min suff. stat.

- all you need for likelihood is minimal sufficient statistic (up to a proportionality constant)
- same for log likelihood plus a constant

exponential family: function of data times function of parameter

$$L(\beta \mid \underset{\sim}{y}) = \exp\{T(\underset{\sim}{y})\, W(\beta) - K(\beta)\}$$

Domain: $\beta > 0$ — important to write b/c of open set stuff

$$= \exp\left\{-\frac{U}{\beta} - n \log \beta\right\} \qquad U = \sum_{i=1}^{n} Y_i$$

The density of $\underset{\sim}{y}$ (a vector) their joint distribution belongs to a 1 parameter exponential family.

U is complete because $\{W(\beta), \beta > 0\} = (-\infty, 0)$ in $\mathbb{R}^1$

$W(\beta) = -1/\beta \to (-\infty, 0)$

Because the range of $W(\beta)$ contains an open set in $\mathbb{R}^1$, e.g. $(-2, -1)$

Can't just say that since its exponential family it is CSS. Need the extra condition that it contains an open set → show it.

If you just say "b/c it's exponential family" this shows lack of understanding of completeness. Completeness is a mathematical concept, not statistical. Example:

$$\left.\begin{array}{l} X \sim \text{Poisson}(\theta) \\ Y \sim \text{Poisson}(\theta^2) \end{array}\right\} \text{independent} \qquad \theta > 0$$

what is min. sufficient stat?

$$f_{X,Y}(x,y|\theta) = \frac{e^{-\theta}\theta^x}{x!}\,\frac{e^{-\theta^2}\theta^{2y}}{y!}$$

$$= \frac{e^{-\theta-\theta^2}\theta^{x+2y}}{x!\,y!}$$

$$= \exp\{\overbrace{(x+2y)}\overbrace{\log\theta}^{w(\theta)} - \overbrace{(\theta+\theta^2)}^{K(\theta)} + c(x,y)\}$$

minimal suff stat is $(x+2y)$, the range of θ is $(-\infty, \infty)$, so also complete

(x, y) is a sufficient statistic. It is two dimensional, but not minimal sufficient. Is it complete?

Can we find a function of x and y that has expectation zero?

$x \to E(x) = \theta \qquad x^2 \to E(x^2) = \theta + \theta^2$

$y \to E(y) = \theta^2$

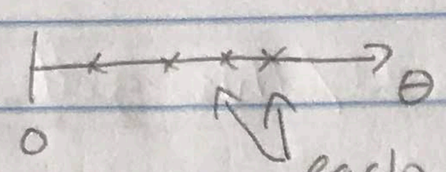
$\forall \theta > 0$

$E[x^2 - y - x] = 0$ So, not complete.

Completeness is a property for each statistic.
This is a shorthand:

Completeness is a property of a family (set) of distributions.

$\{\text{Poisson}(\theta),\ \theta > 0\}$ ← a set of distributions



each point represents one theta on real number line. Each is a different distribution

Think about the family of distributions, not the statistic

Only minimal sufficient statistic can be complete.

$Y \sim \text{Bin}(2, \theta)$
$Y \in \{0, 1, 2\}$

$\theta = \{\frac{1}{4}, \frac{1}{2}\}$ ← only two points

y	$P(Y=y)$
0	$(1-\theta)^2$
1	$2\theta(1-\theta)$
2	θ^2
	1

y	$P(Y=y)$	$\theta = \frac{1}{4}$	$\theta = \frac{1}{2}$
0	$(1-\theta)^2$	$9/16$	$1/4 = 4/16$
1	$2\theta(1-\theta)$	$6/16$	$2/4 = 8/16$
2	θ^2	$1/16$	$1/4 = 4/16$
		1	1

Y is not complete

Y only takes on three values

proportional to

$\theta = \frac{1}{4}$: 9 6 1

$\theta = \frac{1}{2}$: 4 8 4

$$\rightarrow \begin{bmatrix} 9 & 6 & 1 \\ 4 & 8 & 4 \end{bmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This parameter space is not complete $\rightarrow$ $\Theta = \{\frac{1}{4}, \frac{1}{2}\}$

There are infinite solutions for f_1, f_2, f_3

If you set $f_1 = 1$, then:

$$\begin{bmatrix} 6 & 1 \\ 8 & 4 \end{bmatrix} \begin{pmatrix} f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} -9 \\ -4 \end{pmatrix}$$

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \leftarrow \text{w/ this get } E(\quad) = 0$$

functions that are not zero, where no matter what the true parameter is, the expectation is zero.

Change problem to: $\Theta = \{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}\}$, will not find f_1, f_2, f_3 where $E(\ )$ is zero

Or, if $\Theta = \{\frac{1}{4}, 3/4, \frac{1}{2}\}$

This parameter space is complete

y	$\theta = 3/4$
0	1/16
1	6/16
2	9/16

$P(Y=0) = P_0$
$P(Y=1) = P_1$
$P(Y=2) = P_2$

Expectation: $5 \cdot P_0 + 3 \cdot P_1 + 7 \cdot P_2 = 0$

you cannot solve this system $\rightarrow$

$$\begin{bmatrix} 9 & 6 & 1 \\ 4 & 8 & 4 \\ 1 & 6 & 9 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

So complete Binomial has property that any three different values of Θ will

- This parameter space doesn't contain an open set, but it is still complete

$E[I(Y_1 > t)]$

obvious estimator is the indicator of the event

$E[\underbrace{I(A)}_{\text{R.V.,}}] = P(A)$

unbiased estimator of probability of event A

parameter

- completeness doesn't impact ability to estimate

What you are really talking about is the family of distributions that that statistic belongs to

UMVUE's doesn't always exist
↳ had one in previous exam

If you can't apply Lemon-Scheffe, UMVUE might still exist.

To find UMVUE:

usually a lot of work

- have statistic, $T(Y)$, and $E[T(Y)] = \theta$ unbiased
  then, compute $var(T(Y))$
  compute CRLB, it is UMVUE if the same
-

(X, Y), $Y_1 \dots Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$ known

$\Phi\left(\frac{-\mu}{\sigma}\right) = \theta = P\left(Z \le \frac{-\mu}{\sigma}\right) = P(\underbrace{Z\sigma + \mu}_{\sim N(\mu, \sigma^2) \to \text{distribution of } Y_1 \text{ to } Y_n} \le 0) = P(Y_1 \le 0) = \theta$

CDF of standard normal

$V_1 = I(Y_1 \le 0)$, so V_1 is unbiased estimator of θ.

$\bar{Y}$ is MSS (min suff stat), $-\infty < \mu < \infty$

what is the bridge between Y_1 and $\bar{Y}$ → Rao Blackwell Theory

$E[V_1 \mid \bar{Y}] = g(\bar{Y})$, $E(g(\bar{Y})) = \theta$

$\boxed{2016}$

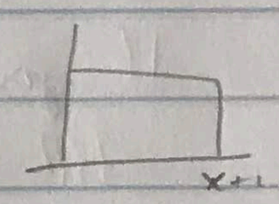
1) $\underbrace{0000}\,1\,\underbrace{000}\,1$ ← 2nd success

# of Zeros $= X$ $\qquad \binom{x+1}{1}$

a) $P(X=x) = (x+1)p^2q^x \quad , \; x = 0, 1, 2 \ldots$

$Y \mid X = x \sim U(0, x+1)$

Y is continuous: $U(0,1)$, etc.



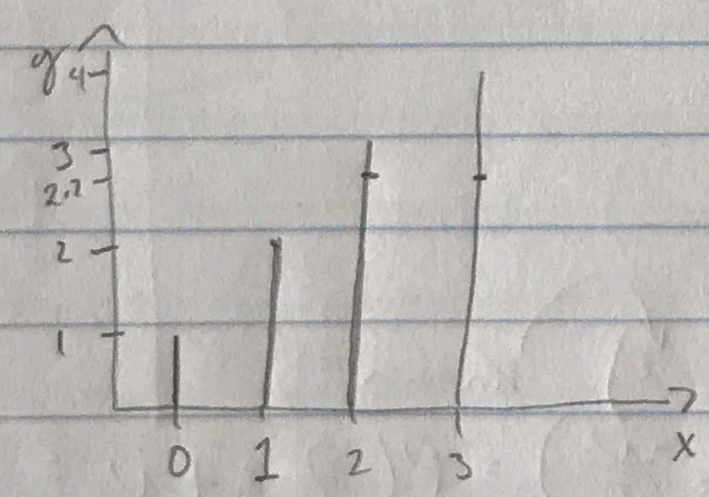
$f_{X,Y}(x,y) = f_X(x) \cdot f_{Y|X}(y|x)$

$= (x+1)p^2q^x \cdot \frac{I(0<y<x+1)}{x+1}$

not done until you write down the range ↳ bivariate

$\{(x,y) : x \in \{0, 1, \ldots\}, \; 0 < y < x+1\}$

Draw the support:



lines in $\mathbb{R}^2$

if $y = 2.7$, then x can be $2, 3, \ldots$

b) $f_Y(y) = \sum_{x \in A(y)} f_{X,Y}(x,y) \qquad A(y) = \{\lfloor y \rfloor, \lfloor y \rfloor + 1, \ldots\}$

$y^* = \lfloor y \rfloor$

$= \sum_{x=y^*}^{x=\infty} \cancel{(x+1)} p^2 q^x \frac{I(0<y<x+1)}{\cancel{x+1}}$

$= p^2 \sum_{y^*}^{\infty} q^x$ ← geometric summation

$= \frac{p^2 q^{y^*}}{1-q} = pq^{y^*}, \quad 0 < y < \infty$ ★

↳ $= pq^{\lfloor y \rfloor}$

1c) $E[Y]$ → Don't get integral of $f_Y(y)$ → will get the right answer but too much work

$$E[Y|X] = \frac{X+1}{2}$$

$$E[Y] = E\left[\frac{X+1}{2}\right] = \frac{\frac{2q}{p}+1}{2}$$

d) $Cov(X, Y) = E[Cov(X, Y|X)] + Cov(E[Y|X], E[X|X])$

$= E[0] + Cov\left(\frac{X+1}{2}, X\right)$

$= \frac{1}{2} Var(X)$

e) Guess → Cov is zero

f) Cov is zero, but need to check.
Easiest to check ranges

$T = 2Y - X$

range: $(0, X+1)$

range $2Y$: $(0, 2X+2)$, now $-X$

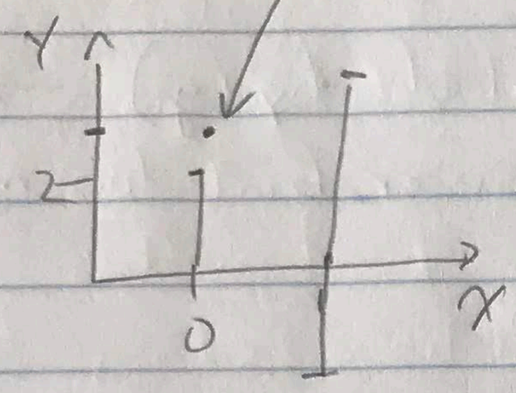
X	Y	$2Y$	T
0	(0, 1)	(0, 2)	(0, 2)
1		(0, 4)	(-1, 3)

→ They are not independent

The range of T depends on X

The conditional range of T given $X = x$ depends on X

1f)



at the point, the value of the joint is not equal to the product of the marginals

Qaqish Review 6/26/19

Completeness

$Y_1, \dots, Y_n$ iid $N(\mu, \mu^2)$, $\mu > 0$

$$\prod_{i=1}^{n} \frac{e^{-\frac{1}{2\mu^2}(y_i-\mu)^2}}{\sqrt{2\pi\mu^2}}$$

$$= \exp\left\{ \sum_{i=1}^{n} \left[-\frac{1}{2}\log\mu^2 - \frac{y_i^2 - 2y_i\mu + \mu^2}{2\mu^2} \right] \right\} c$$

$$= \exp\left\{ \left(\sum_{i=1}^{n} y_i^2\right)\left(-\frac{1}{2\mu^2}\right) + \left(\sum_{i=1}^{n} y_i\right)\left(\frac{1}{\mu}\right) - \frac{n}{2} - \frac{n}{2}\log\mu^2 \right\} c$$

$$T_1(y) \cdot w_1(\mu) + T_2(y) \cdot w_2(\mu) + K(\mu) \} \cdot c$$

MSS is $\left(\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} Y_i^2\right)$

Complete? Have to look at parameter space

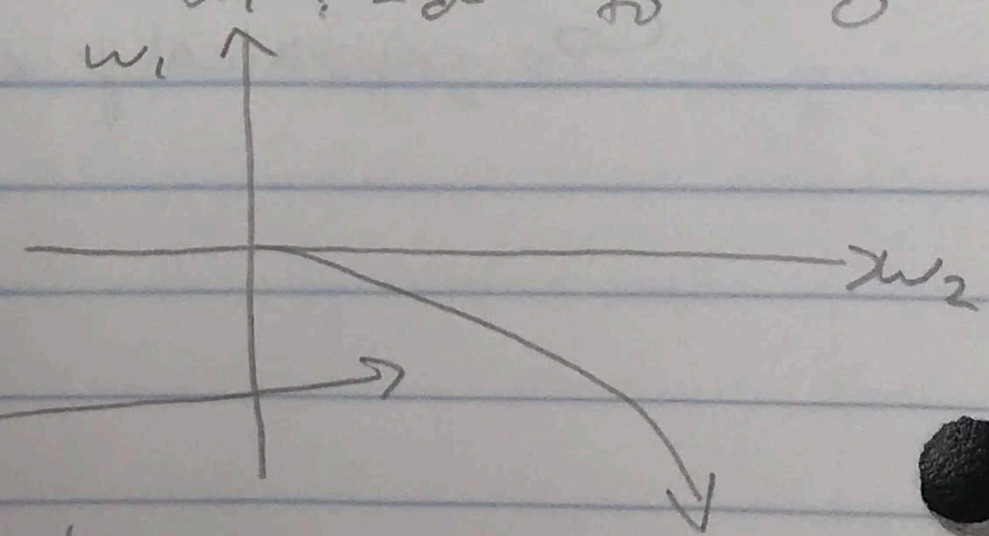
Need to look at $w_1(\mu)$, $w_2(\mu)$ and parameter space $\mu > 0$:

$(w_1(\mu), w_2(\mu))$ $\mu > 0$

$\left(-\frac{1}{2\mu^2}, \frac{1}{\mu}\right)$

$\left(-\frac{1}{2}\{w_2(\mu)\}^2, w_2(\mu)\right)$

μ : 0 to ∞
w_2 : ∞ to 0
w_1 : $-\infty$ to 0



w_1 and w_2 only take values on this line

★ It doesn't contain an open set in R^2 (which would be a disk or a square). So, we don't know if its complete or not. ★

	E	var
$\sum_{i=1}^{n} Y_i$	$n\mu$	$n\mu^2$

$\sum_{i=1}^{n} Y_i^2 \quad 2n\mu^2$

apply linearity of expectation

$$E[Y_i^2] = (EY_i)^2 + var(Y_i)$$
$$= \mu^2 + \mu^2 = 2\mu^2$$

	E
$\left(\sum_{i=1}^{n} Y_i\right)^2$	$n^2\mu^2 + n\mu^2 = (n^2+n)\mu^2$

$$E\left(\frac{\sum_{i=1}^{n} Y_i^2}{2n}\right) = \mu^2 \qquad E\left(\frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n^2+n}\right) = \mu^2$$

$n > 1$

$$E\left[\frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n^2+n} - \frac{\sum_{i=1}^{n} Y_i^2}{2n}\right] = 0 \quad \forall \mu > 0$$

So, this shows that this MSS is not complete.

The only things that can be complete are MSS. Things that are not minimal can't be complete.

Suppose I have 1 observation, $N(\mu, \mu^2)$
$\mu > 0$

Find an exact 95% CI for μ

$$\frac{Y-\mu}{\mu} \sim N(0,1)$$

Answer: use a pivot

?

$$P\left(-1.96 < \frac{Y-\mu}{\mu} < 1.96\right) = 0.95$$

get μ by itself?

Always keep pivots in mind.
Pivots are not statistics, b/c it involves an unknown parameter

$$Y \sim Exp(\mu)$$

pivot $\rightarrow$ $\frac{Y}{\mu} \sim Exp(1)$

$\boxed{2017}$ 3

3a) CLT

b) $F_n(x) \to$ avg of Bernoulli's, use CLT

c) a particular choice of value $F(x)$

x given, so $p = F(x)$

$H_0: p = \frac{1}{2}$ vs $H_1: p \neq \frac{1}{2}$

$Y_1, \ldots, Y_n$ iid $Ber(p)$

$$L(p|y) = p^{\sum y_i}(1-p)^{n-\sum y_i}$$

$$= p^{n\bar{y}}(1-p)^{n(1-\bar{y})}$$

$$= \{p^{\bar{y}}(1-p)^{1-\bar{y}}\}^n$$

$$\hat{p} = \bar{y}$$

$$L(\hat{p}|y) = \{\bar{y}^{\bar{y}}(1-\bar{y})^{1-\bar{y}}\}^n$$ ← largest possible value

Under H_0: $L(\frac{1}{2}|y) = \{\frac{1}{2}^{\bar{y}}(\frac{1}{2})^{1-\bar{y}}\}^n = \{\frac{1}{2}\}^n$

switched from how we did it

$$\frac{L(\hat{p}|y)}{L(\frac{1}{2}|y)} = \{2\bar{y}^{\bar{y}}(1-\bar{y})^{1-\bar{y}}\}^n$$

$$\log() = n\{\log 2 + \bar{y}\log\bar{y} + (1-\bar{y})\log(1-\bar{y})\}$$

remember it is χ^2 and the df is the number of restrictions

χ^2_1 (b/c one constraint)

function of $\bar{y}$, a transform of $\bar{y}$: $g(\bar{y})$

$g(\bar{y})$

CLT: $\sqrt{n}(\bar{y} - p) \xrightarrow{d} N(0, p(1-p))$, then use delta method

as $n \to \infty$

3c) but problem:

$g'(\frac{1}{2}) = 0$

so go to next derivative

$g''(\frac{1}{2}) \neq 0$

$n(\bar{y} - p)^2 \longrightarrow (?)\ \chi^2_1$

Different number of contraints:

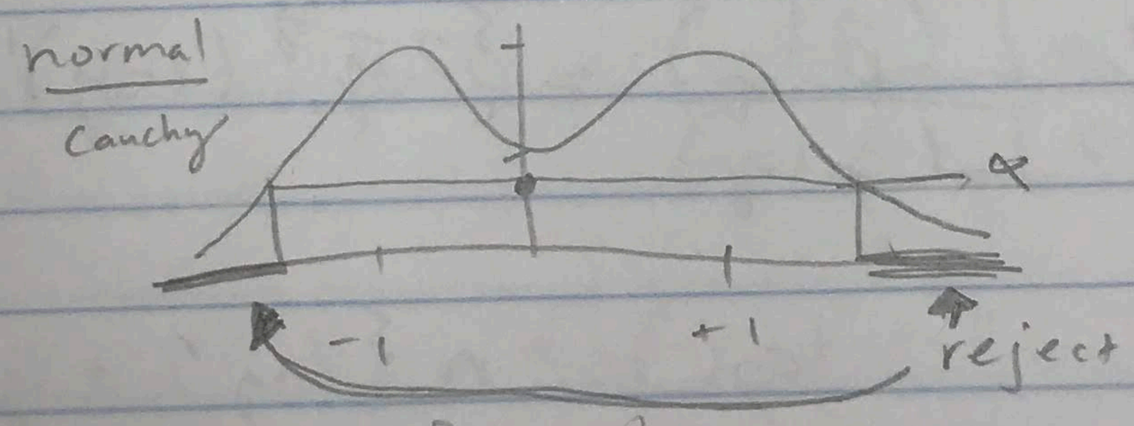
$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} \longrightarrow \theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$$

$$\begin{pmatrix} \theta_1 - \theta_2 \\ \theta_1 - \theta_3 \\ \theta_1 - \theta_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

← 3 restrictions, so LRT is χ^2_3

d) $G(x) \sim N$, $H(x) \sim$ Cauchy

likelihood ratio: $\frac{G(x)}{H(x)}$



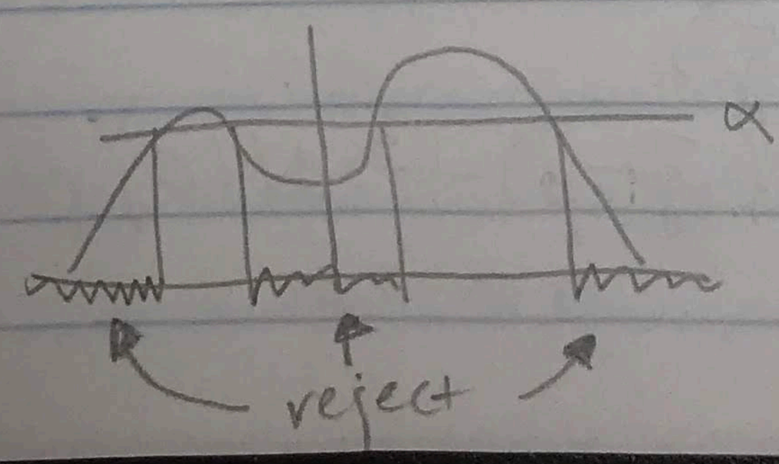
What's the critical region? is the question

H_0 : normal

H_1 : Cauchy

So reject H_0 if ratio $\frac{H_0}{H_1} < k$

for small α, rejects if value is too large or too small, for large α reject if x too large, too small or close to zero



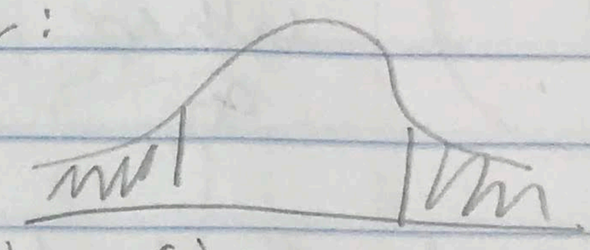
as α gets bigger

d) can define critical region in terms of x or the ratio

e) power: prob of rejecting H_0 given H_1 is true

? $P(|X| > \Phi^{-1}(1-\frac{\alpha}{2}) \mid X \sim \text{Cauchy})$

Cauchy:



Need Cauchy CDF

↳ do integral, has tan()

Same:

$$\underset{\sim F(1,1)}{X^2} > \{\Phi^{-1}(\alpha)\}^2$$

$1 -$ CDF of $F(1,1)$

$\boxed{2016}$

3) a) X is a_1 or a_2 or a_3 or a_4

if $X = a_1$, MLE of θ is θ_2

if $X = a_2$, MLE is θ_1

$= a_3$ θ_3

$= a_4$ θ_2

b) LRT is function of X value you'll observe

	a_1	a_2	a_3	a_4	
MLE	θ_2	θ_1	θ_3	θ_2	
H_1: $L(\text{MLE})$	0.4	0.4	0.5	0.3	← max value of likelihood
H_0: $L(\theta_1)$	0.3	0.4	0.1	0.2	
ratio $\frac{H_1}{H_0}$:	$4/3 = 1\frac{1}{3}$	1	5	$3/2 = 1\frac{1}{2}$	reject null if $\frac{H_1}{H_0} \geq k$

reject if ratio $\frac{H_1}{H_0} \geq 5$

i.e if $x = a_3$, otherwise fail to reject.

3b) if $\alpha = 0.3$

				0.1
Values:	5	$1\frac{1}{2}$	$1\frac{1}{3}$	1

?

where you draw your line is based on α

look at value of likelihood ratio statistic not x.

You can achieve other alphas but that test is not going to be UMP.

looking w/in a row: pmf
looking w/in column: likelihood

c) What we did (?)
d) Point vs Point

	a_1	a_2	a_3	a_4
	0.3	0.4	0.1	0.2
	0.4	0.1	0.2	0.3
$\frac{L(\theta_1)}{L(\theta_2)}$ ratio:	3/4	4	1/2	2/3

reject

Sort them:	4	3/4	2/3	1/2
Probability under H_0:	0.4	.3	0.2	$\boxed{0.1}$

reject if too small
want type 1 error 0.1
reject if this ratio if $\frac{1}{2}$ or less
UMP b/c point vs. point, relies on N-P Lemma

e) θ_1 against θ_2 UMP is what we got
the question is

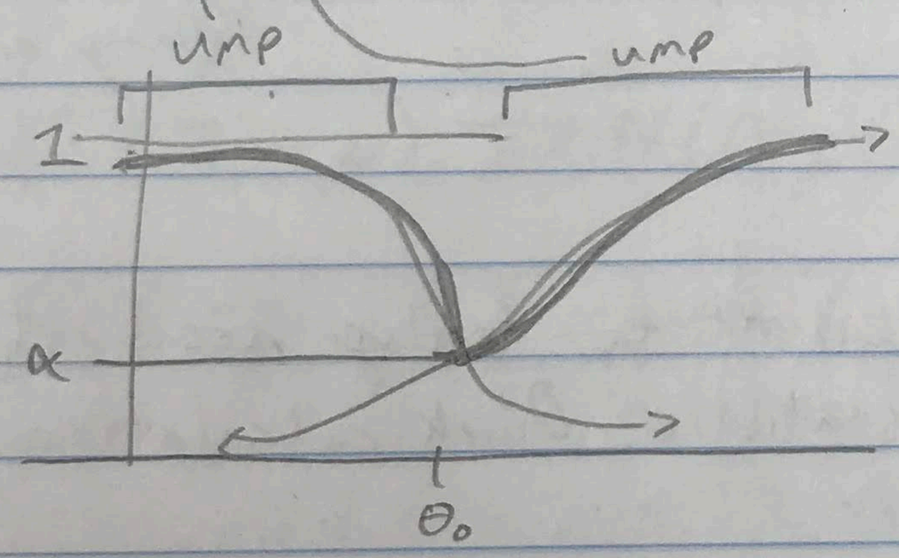
θ_1 against θ_3 :

	a_1	a_2	a_3	a_4
$\frac{L(\theta_1)}{L(\theta_3)}$ =	$\frac{3}{2}$	4	$\frac{1}{5}$	1

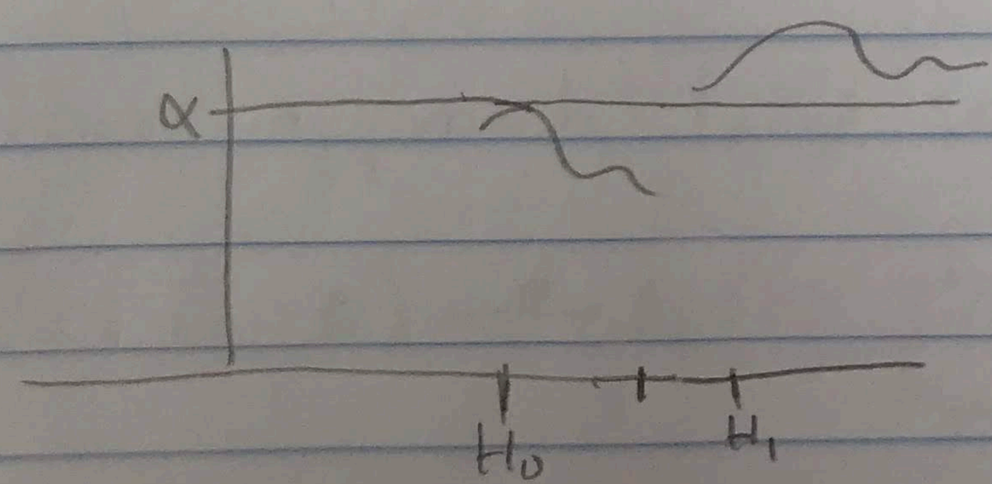
Probability under H_0 : 0.1

Against θ_2 and against θ_3, this is UMP, if $x = a_3$

$\theta = \theta_0$ vs $\theta > \theta_0$
$\theta = \theta_0$ vs $\theta < \theta_0$
$\theta = \theta_0$ vs $\theta \neq \theta_0$



Unbiased test: any test as a size probability of rejecting H_0 when H_0 true, which is α. Unbiased test: power curve never dips below α. (when H_1 is true)



prefer this to be larger than the second. This is true of unbiased tests.

$\rightarrow P(\text{reject } H_0 \mid H_0 \text{ False})$
$P(\text{reject } H_0 \mid H_0 \text{ True})$

$\boxed{2015}$

3e) For LRT just need likelihood function

Likelihood:

$L_1(\theta | \underset{\sim}{x}) \cdot L_2(\beta | y)$ ← maximize them separately under alternative

$H_1: \hat{\theta} = \frac{1}{\bar{x}}, \hat{\beta} = \frac{1}{\bar{y}}$ ← MLE separately

$\theta = \beta$ → H_0: X's and Y's are all iid, so can look at all of them together:

One sample w/size $2n$

(1/mean)

$$\hat{\theta} = \hat{\beta} = \frac{2n}{n\bar{X} + n\bar{Y}}$$ } reciprocal of the mean of the combined sample b/c H_0 is they have the same distribution

$X_1, \ldots, X_n \sim Exp(\theta)$ iid

$Y_1, \ldots, Y_n \sim Exp(\beta)$ iid

$exp(\theta)$ or $exp(\beta)$ (they are same under H_0)

Score: take derivative, only need H_0. Calculate variance of score statistic. Quick calculation

7/3/19 Review

$\boxed{2015}$ 3) a) $\hat{\theta} = \frac{1}{\bar{x}}$

b) $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, ___)$

know it's Normal b/c

$(T - \mu)$

CLT $\rightarrow$ $\sqrt{n}(\bar{x} - \frac{1}{\theta}) \xrightarrow{d} N(0, \frac{1}{\theta^2})$

$\mu = EX = \frac{1}{\theta}$

$\sigma^2 = Var\, X = \frac{1}{\theta^2} = \mu^2$

$\sqrt{n}(\frac{1}{\bar{x}} - \theta) \xrightarrow{d} N(0, \frac{1}{\theta^2} \cdot ___)$ ← Delta Method

$g(\bar{x}) = \frac{1}{\bar{x}}$, $g(t) = \frac{1}{t}$

$g'(t) = -1/t^2$ substitute in μ for t

$T = \bar{X}$

$\sqrt{n}(T - \mu) \xrightarrow{N} N(0, \mu^2)$ ← this is our starting point

$\sqrt{n}(\frac{1}{\bar{x}} - \frac{1}{\mu}) \xrightarrow{N} N(0, \mu^2 \cdot (-\frac{1}{\mu^2})^2)$

$\equiv N(0, \frac{1}{\mu^2})$

$\equiv N(0, \theta^2)$

Can only use b/c $\hat{\theta}$ is function of $\bar{X}$, and have CLT w/ $\bar{x}$

Other approach: large sample properties of MLE

$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \frac{1}{i(\theta)})$ as $n \rightarrow \infty$ always

information for 1 (or avg of all)

$\boxed{2015}$ 3 b)

$: \theta e^{-\theta x_i}$ (when iid, just do 1 observation)

$\ell(\theta) = \log\theta - \theta x_i$

$\frac{\partial \ell}{\partial \theta} = \frac{1}{\theta} - x_i \longrightarrow$ get var $\rightarrow i(\theta) = \frac{1}{\theta^2}$

2nd method is easier + more general

CRLB: just $\frac{1}{i(\theta)}$, so <u>asymptotically</u>

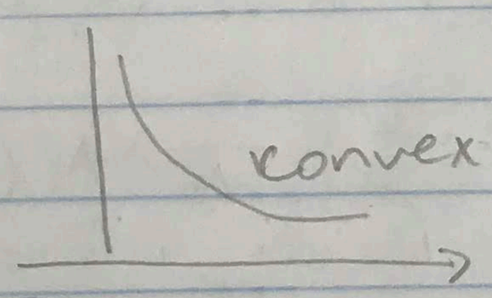
it does achieve the CRLB.

$\rightarrow$ Is $\hat{\theta}$ biased or unbiased?

$E(\bar{X}) = \frac{1}{\theta}$

$E[\frac{1}{\bar{X}}] \overset{?}{=} \theta$ $\leftarrow$ not linear function, so not equal

$\bar{X}$ is positive



<u>Jensen's</u>

if g convex, $E[g(X)] > g(E[X])$

so

$E[\hat{\theta}] = E[\frac{1}{\bar{X}}] > \frac{1}{E\bar{X}} = \frac{1}{1/\theta} = \theta$

On exam, don't try to figure out exact expected value. Waste of time.

Side: $Y \sim$ Gamma, $E[\frac{1}{Y^r}]$ $r = 1, 2, 3, \ldots \rightarrow$ rearrange

$\boxed{2015}$ 1) d-f

d) Use induction for these kinds of problems

Assuming relationship is true for n, show it is true for $n+1$

and show it's true for 1 $(\in \mathbb{Z}_1)$

e) hard to find prob distribution marginally, but can condition on X_n

Step	R	B	Total (fixed)
0	a	b	$a+b$
n	X_n	$a+b-X_n$	

event R

$P(\text{red} \mid X_n) = \frac{X_n}{a+b}$ ← linear function of X_n

double expectation:

prob is $E[I(R) \mid X_n]$

$$EE[I(R) \mid X_n] = E[I(R)] = P(R)$$

$$\downarrow$$

$$E[X_n / a+b]$$

f) $P(Y = a+b) = 1$ → means Y is a constant equal to $a+b$ (which is total)

Given $\epsilon > 0$

$P(|X_n - (a+b)| > \epsilon) \to 0$ as $n \to \infty$

(X_n is only integer, $a+b$ is integer, ϵ can be 0, 1

→ $P(X_n = a+b) \to 1$

$\boxed{2015}$ 1f)

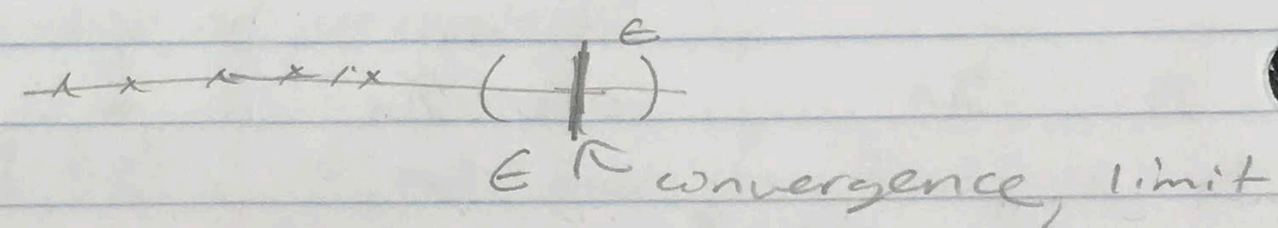
prob. $\overbrace{P(R_n)}^{\text{sequence}} \longrightarrow 1$ as $n \to \infty$

Given $\epsilon > 0$ $\exists$ an integer m such that
(there is)

$$|P(R_n) - 1| < \epsilon \quad \forall n > m$$

only values $P(R_n)$ takes $\longrightarrow P(R_n) \in \left\{ \frac{0}{a+b}, \frac{1}{a+b}, \dots, \frac{a+b-1}{a+b}, 1 \right\}$

1) mathematical convergence from e)



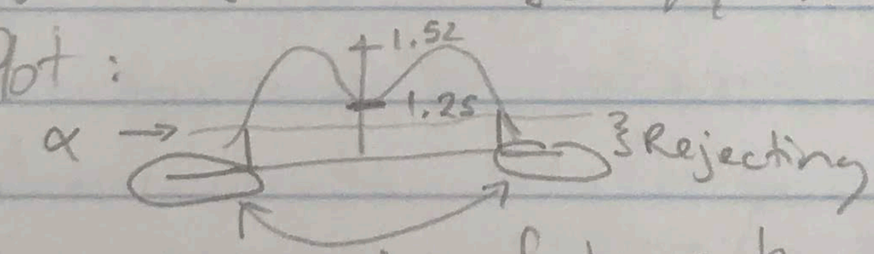
after m, every single point will be in the ϵ interval

$\rightarrow$ in this problem, the numbers are fixed and if you're within $1/2$ of $\frac{1}{a+b}$, then you have to be at 1.

$\boxed{2017}$ 3d, e)

d) Give answer: use N-P Lemma

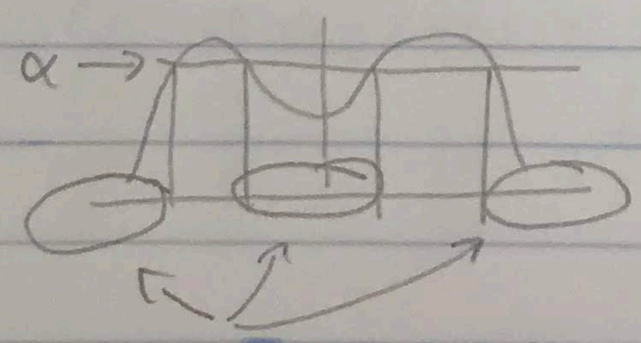
Plot:



$R: \{x_1 < -c\} \cup \{x_1 > c\}$
$P(x_1 \in R \mid H_0) = \alpha$

union of these two areas is your rejection region

α is always under the null, $N(0,1)$ so find critical region here



Rejection Region

α: 0.63 or 0.32

e) Just use the definition of power

$\boxed{2017}$ 3 c)

Y_i's $\rightarrow$ Berroulli's

$Y_1, \ldots, Y_n$ iid Bern(θ)

$H_0: \theta = \frac{1}{2}$ $\quad$ $H_1: \theta \neq \frac{1}{2}$

$\hat{\theta} = \bar{Y}$

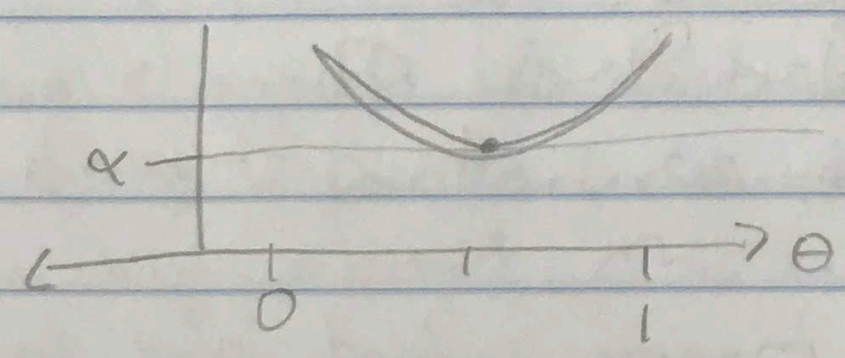
$$\text{LRT: } \frac{H_0}{H_1} = \left(\frac{\left(\frac{1}{2}\right)}{\left\{ \bar{Y}^{\bar{Y}} (1-\bar{Y})^{1-\bar{Y}} \right\}} \right)^n$$

take log, -2

asymptotic is easier

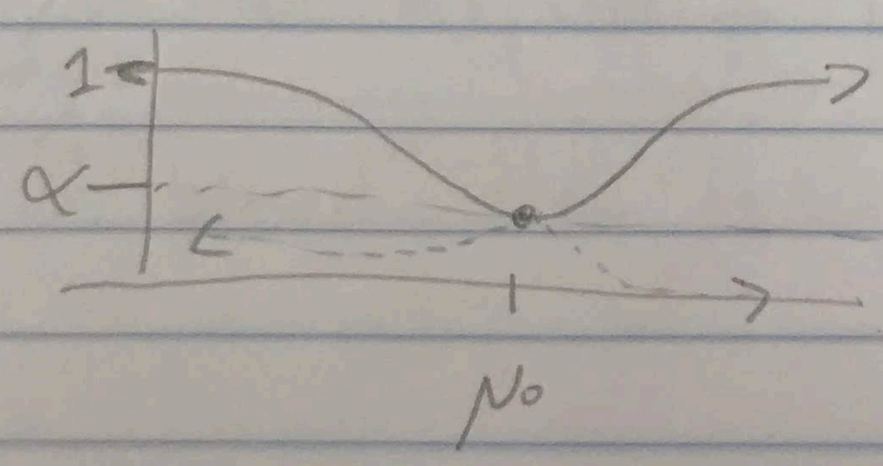
3e) Can't draw power function b/c simple vs simple

Sketching Power Function



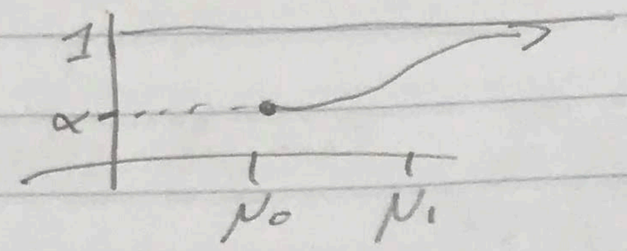
$N(\mu, 1)$

$H_0: \mu = \mu_0$ $\quad$ $H_1: \mu \neq \mu_0$ $\quad$ Z statistic



Want to guarantee α is no larger than a certain place

One-sided Test:



$\boxed{2015}$ 4a)

this transformation preserves the order

$$Y_i = \frac{X_i - \theta}{a} \longrightarrow X_i = aY_i + \theta$$

b) Derive pdf of X_n, calculate expected value, it is a linear

$$Y_{(n)} = \frac{X_{(n)} - \theta}{a} \longrightarrow X_{(n)} = aY_{(n)} - \theta$$

$\sim$ Beta, but can also find CDF, derivative

$E(Y_{(n)})$

CRLB $\longrightarrow$ b/c log likelihood is not differentiable anywhere $\theta = X_i$?
CRLB requires derivatives. There is an indicator function, so not differentiable

For MLE's always imagine what plot looks like.

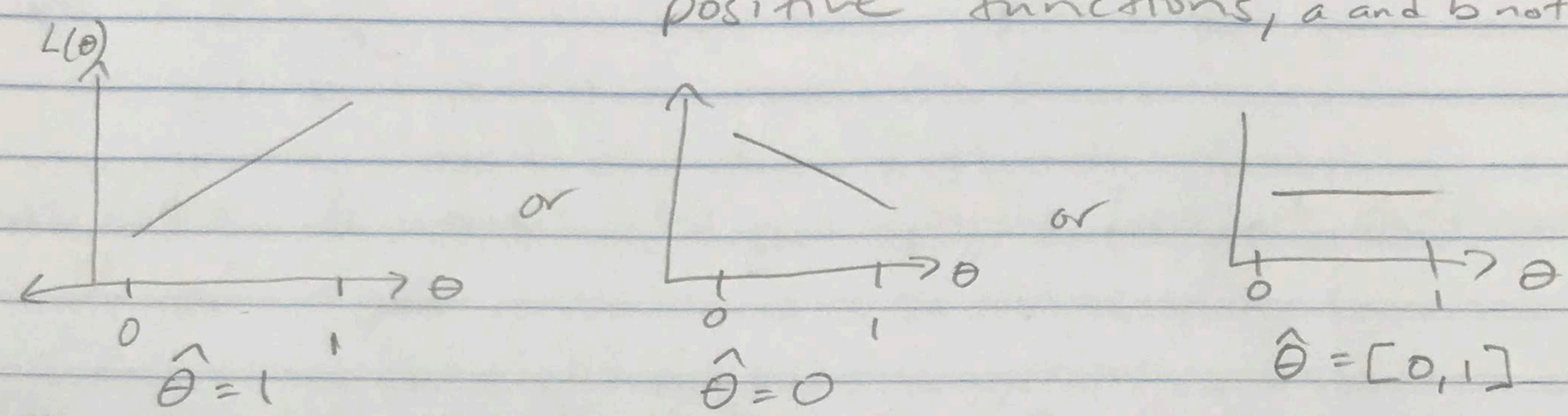
There was a problem where the density looked like:

$$f_X(x|\theta) = \theta \cdot a(x) + (1-\theta)\, b(x) = \theta \cdot a(x) + b(x) - \theta b(x)$$

$$\theta \in [0, 1]$$

a is given, only one observation

Find mle: ($a(x)$ is pdf, so is $b(x)$, both positive functions, a and b not same)



$\hat{\theta} = 1$ $\quad$ $\hat{\theta} = 0$ $\quad$ $\hat{\theta} = [0, 1]$

$$f_X(x|\theta) = \theta\{a(x) - b(x)\} + b(x)$$

↳ linear function of θ

slope is $\{a(x) - b(x)\}$

↳ if positive, then $\hat{\theta} = 1$

if negative, then $\hat{\theta} = 0$

→ if 0, then $\hat{\theta} = [0, 1]$

doesn't happen b/c X is continuous so probability of it being a particular number is zero, so don't need to worry about this case

Don't be mechanical. Open up your mind and understand the problem the way it is.

$\boxed{\text{Try to find this problem}}$

$\hat{\theta}$ biased? How to fix bias?

1 observation, $Y \sim Bern(\theta)$

MLE of $\theta = \bar{Y} = Y \longrightarrow$ could be 0 or 1

MLE Binomial: $\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n}{n}$

Jagish Review

$\boxed{2014}$ 1e

1e) $\begin{pmatrix} X_i \\ Y_i \end{pmatrix}$, $E\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \underset{\sim}{\mu}$, iid

If X, Y are bivariate normal, $\bar{X}, \bar{Y}$ are bivariate normal

$$\text{cov}\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \underset{2\times 2}{\Sigma} = \begin{bmatrix} \text{var}(X_i) & \text{cov}(X_i, Y_i) \\ & \text{var}(Y_i) \end{bmatrix}$$

$$E\begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \underset{\sim}{\mu}$$

$$\text{cov}\begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} = \frac{1}{n}\Sigma = \begin{bmatrix} \sigma_{11}/n & \sigma_{12}/n \\ & \sigma_{22}/n \end{bmatrix}$$

$$\frac{\sigma_{12}/n}{\sqrt{\frac{\sigma_{11}}{n} \cdot \frac{\sigma_{22}}{n}}} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{12}}} = \text{corr}(X_i, Y_i)$$ ← not necessary

New Notation:

$$\begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix} \sim MVN(\underset{K\times 1}{\underset{\sim}{\mu}}, \underset{K\times K}{\Sigma})$$

multivariate normal

Single: $$\frac{\exp\{-(x_i - \mu_i)^2 / 2\sigma_{ii}\}}{\sqrt{2\pi\sigma_{ii}}}$$

if iid multiple: (product) $$\frac{\exp\{-\sum_{i=1}^{K}(x_i - \mu_i)^2 / 2\sigma_{ii}\}}{(2\pi)^{K/2}\left(\prod_{i=1}^{K}\sigma_{ii}\right)^{1/2}}$$

notation: $|\Sigma|^{K/2}$

Determinant of Σ ^ $K/2$

General sigma, not iid (Σ is not diagonal)

$(\underset{\sim}{x}-\underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{x}-\underset{\sim}{\mu})$ in exponent

$$\frac{\exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})^T \Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\right.}{(2\pi)^{k/2}|\Sigma|^{k/2}}$$

$\boxed{2014}$

2a)

$Y|X \sim \text{Poisson}(X)$

$E[\text{common colds} \mid X] = X$, $X \sim U(0,2)$

Y

- if any RV shows up on right side for expectation + probability, it must be conditional on that RV

☆ $Y \sim \text{Poisson}(X)$ ← total nonsense!! It must be $\boxed{Y|X}$ (unless little x is known constant)

g) $E[X \mid \underbrace{Y=0}_{\text{no common colds}}]$ Y, discrete, X continuous

d) $E[Y-X] \overset{?}{=} 0$ ← if unbiased

$= E[E[Y-X|X]] = E[X-X] = 0$

Prediction mean squared error: $E(Y-X)^2 =$

$E(Y-X)^2 = \text{Bias}^2 + \text{var}(Y) = (E(Y-X))^2 + \text{var}(Y)$

$= 0 + 4/3$

$\text{Var}(Y) = 1 + \frac{1}{3} = \frac{4}{3}$ ← conditional var

$\boxed{2014}$

2e) $E(X-a-bY)=0$, $Var(X-a-bY) \leftarrow$ min.

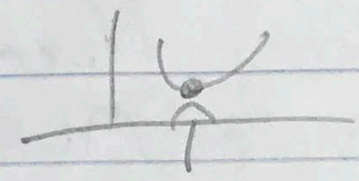
$\hookrightarrow = E(X) - a - bE[Y]$

$= 1 - a - b(1) = 0 = \boxed{a = 1-b}$

$E[Y] = EE(Y|X) = E(X) = 1$

$Var(X-a-bY) = Var(X-bY)$ (constant doesn't matter)

$= \underbrace{Var(X)}_{\text{out of our control, so minimize this}} + \underbrace{b^2 Var(Y) - 2b \cdot cov(X,Y)}$

$b^2 Var(Y) - 2b\,cov(X,Y) \longrightarrow$ 

Derivative: $\frac{\partial}{\partial b} = 2b\,Var(Y) - 2cov(X,Y) = 0$

$\Rightarrow \boxed{b = \frac{cov(X,Y)}{Var(Y)} \quad a = 1-b}$

$\boxed{b = \frac{1/3}{4/3} = \frac{1}{4} \quad , \quad a = 3/4}$

$\boxed{\text{look up}}$

$cov(X,Y) =$ apply double expectation: $1/3$
(it's the same as the variance conditional)

$T = \frac{3}{4} + \frac{1}{4}Y \leftarrow$ new predictor

$E[T] = \frac{3}{4} + \frac{1}{4} = 1 = EX$

smaller than original

$Var(T-X) = Var(\frac{1}{4}Y - X) = \frac{1}{16} \cdot \underset{\uparrow \text{ Var Y}}{\frac{4}{3}} + \underset{\uparrow \text{ Var X}}{\frac{1}{3}} - \frac{1}{2}\underset{\uparrow \text{ covariance}}{\left(\frac{1}{3}\right)} = \frac{1}{4}$

This is linear regression (apparently)

Predict X from Y

from 663: $E[X] + \frac{cov(X,Y)}{var(Y)}(Y - EY)$
(notation is flipped)

$$= 1 + \frac{1/3}{4/3}(Y-1)$$

$$= 1/4\, Y + 3/4$$

best linear unbiased predictor (linear in Y)

Normal, σ^2, S^2 is unbiased
MLE is $\frac{n-1}{n}S^2$, there is an $a+bS^2$
that has smaller MSE but is biased

$\boxed{2014}$ 2a) $Y|X \sim Poisson(X)$, $X \sim U(0,2)$
Find: $E[X|Y=0]$
no obvious way to do this. Find pdf of X for $Y=0$

$$f_{X|Y}(x|y=0) = \frac{f_{X,Y}(x,0)}{f_Y(0)}$$

$$= \frac{\frac{1}{2}\cdot e^{-x}}{\int_0^2 \frac{1}{2}\cdot e^{-x}\,dx} \qquad \boxed{0<x<2}$$

density of X

marginal of y: joint density integrated over x

density of y given X, at $y=0$

$$= \frac{e^{-x}/2}{\frac{1}{2}[-e^{-x}]_0^2} = \frac{\frac{1}{2}e^{-x}}{\frac{1}{2}[e^{-x}]_2^0} = \frac{e^{-x}}{1-e^{-2}} = f_{X|Y}(x,y), \quad x \in (0,2)$$

$$E[X|Y=0] = \int_0^2 \frac{xe^{-x}}{1-e^{-2}}\,dx = \frac{[-xe^{-x}(x+1)]_0^2}{1-e^{-2}}$$

2a) $\int x^n e^{-x}\,dx$ ← occurs a lot in statistics, so he just knew the formula

$$f_X(x)\cdot f_{Y|X}(0|x) = \frac{1}{2}\cdot e^{-x}$$

$\boxed{2014}$ 4) $X_1, \dots X_n \sim iid \sim N(0, \sigma^2)$

$H_0: \sigma = \sigma_0$ vs $H_1: \sigma \neq \sigma_0$

$$\delta(\underset{\sim}{x}) = 1 \text{ if } \sum_{i=1}^{n} x_i^2 < c_1 \text{ or } > c_2$$

d) N-P Lemma to get UMP

sum of squares will be larger than

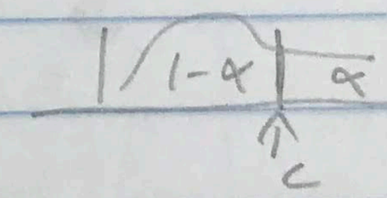
N-P lemma: H_0 $\sigma = \sigma_0$ vs H_1: $\sigma = \sigma_1 > \sigma_0$

↳ gives you the most powerful test for H_0 vs H_1 for its size α. (when H_1 is one point)

Rejection Rule: Reject H_0 if $\sum_{i=1}^{n} x_i^2 > c$

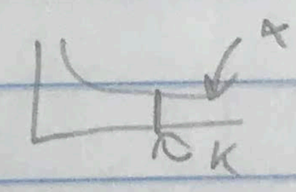
Chose c on alpha and the distribution of the quantity under H_0

here: $\sigma_0^2\ \chi^2_n$:



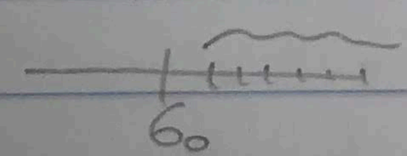
Another test: reject H_0 if $x_1^2 > K$

$\sigma_0^2\ \chi_1^2$

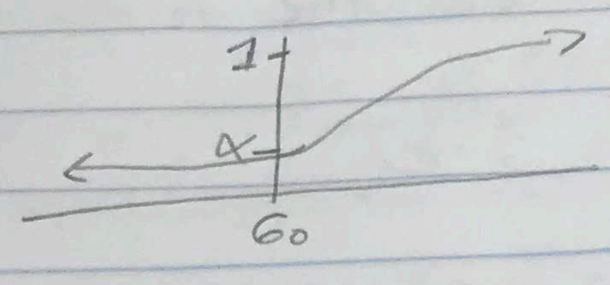


but not most powerful

The next step to uniformly most powerful. B/c the rejection region doesn't depend on σ_1 at all, only on σ_0, and so the test won't change as long as the σ_1 is $> \sigma_0$



$\boxed{2014}$ 4e) ~~When $H_0: \sigma = \sigma_0$ H_1 $\sigma \neq \sigma_1$~~



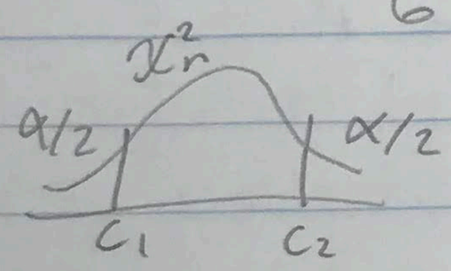
$$\beta(\sigma^2) = P(\delta = 1 \mid \sigma^2)$$
$$= P\left(c_1 < \sum_{i=1}^{n} x_i^2 < c_2 \mid \sigma^2\right)$$

get c_1 and c_2 from χ^2_n

Beginning: $H_0: \sigma = \sigma_0$ H_1 $\sigma \neq \sigma_0$

$$\frac{\sum x_i^2}{\sigma^2} \sim \chi^2_n$$

Chi-squared, n



Inverse CDF

$$c_1 = \sigma_0^2 F^{-1}_{\chi^2_n}(\alpha/2), \quad c_2 = \sigma_0^2 F^{-1}_{\chi^2_n}(1-\alpha/2)$$

$$= P\left(\frac{c_1}{\sigma^2} < \underbrace{\frac{\sum x_i^2}{\sigma^2}}_{\chi^2_n} < \frac{c_2}{\sigma^2}\right)$$

σ^2 is true value

$$= 1 - P\left(\frac{c_1}{\sigma^2} < \chi^2_n < \frac{c_2}{\sigma^2}\right)$$

c_1, c_2 are fixed quantities

Prob of rejection →

$$= 1 - P\left(\frac{\sigma_0^2}{\sigma^2} F^{-1}_{\chi^2_n}(\alpha/2) < \chi^2_n < \frac{c_2}{\sigma^2}\right)$$

Do over

4e) $H_0: \sigma = \sigma_0$ vs $H_1: \sigma > \sigma_0$

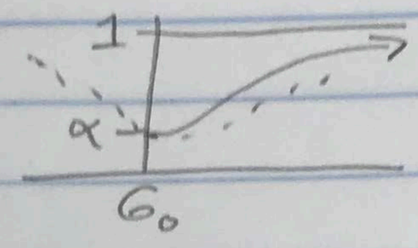
$\delta = 1 - I(c_1 < \sum x_i^2 < c_2)$

UMP: Reject if $\sum x_i^2 > k$

$P(\sum x_i^2 > k \mid H_0) = \alpha$

If hypothesis is $\sigma = \sigma_1$

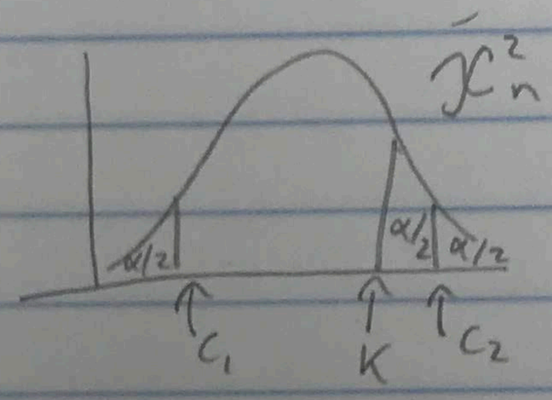
- Part d) —
Part a) - - - -



- Power function for δ:
$P(\sum X_i^2 < c_1 \text{ or } \sum X_i^2 > c_2 \mid \sigma^2)$

- Power for UMP:
$P(\sum X_i^2 > k)$

Under H_0, both will give 5%



$\leftarrow \dfrac{\sum X_i^2}{\sigma_0^2}$ under H_0

$\dfrac{\sum X_i^2}{\sigma_0^2}$ under alternative is not χ^2_n, need: $\underbrace{\dfrac{\sigma_1^2}{\sigma_0^2}}$ $\boxed{\dfrac{\sum X_i^2}{\sigma_1^2}}$

$\uparrow \chi^2_n$

this is > 1, so it's shifting to the right (stretching the distribution)

Review Session MS Theory - 2018 Exam #1 & 2

$\boxed{Y|X \sim N(X, X) \qquad X \sim U(0,1)}$ → Given

#1a)

KNOW YOUR ITERATIVE/DOUBLE EXPECTATION/VARIANCE/COVARIANCE FORMULAS

b)

$E(Y) = E[E(Y|X)] \qquad Y|X \sim N(X, X)$

$= E(X) = \frac{1+0}{2} = 1/2$

$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)]$

$= E(X) + Var(X) = \frac{1+0}{2} + \frac{(1-0)^2}{12}$

$= 7/12$

$Cov(Y, X)$

Method #1: $= \underbrace{E(XY)} - E(Y)E(X) = 1/3 - 1/2 \cdot 1/2$

$= 1/12$

Apply double expectation

$E(XY) = E[E(XY|X)] = E[X\,E(Y|X)] = E(X^2)$

$= Var(X) + [E(X)]^2$

$Var(X) = E(X^2) - [E(X)]^2 \qquad = 1/12 + 1/4 = 1/3$

Method #2:

$Cov(Y, X) = E[Cov(X, Y|X)] + Cov(E(X|X), E(Y|X))$

Can condition on any rv in the same space, as long as you use the same one

$$\mathrm{cov}(X, Y \mid X) = E[XY \mid X] - E(Y \mid X)\,E(X \mid X)$$
$$= XE(Y \mid X) - XE(Y \mid X) = 0$$

$$\mathrm{Cov}(Y, X) = 0 + \mathrm{Cov}(X, X) = 0 + \mathrm{Var}(X) = 1/12$$

b) $$\mathrm{Cov}(Y-X, X) = E[X(Y-X)] - E(Y-X)E(X)$$
$$= [E(XY) - E(X^2)] - [E(Y) - E(X)]\,E(X)]$$

Applying linearity of expectation

$$= 0 - 0 = 0$$

In part 1a, we showed $E(XY) = E(X^2)$ and $E(Y) = E(X)$

OR

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

This will be 0 because

Covariance of X & something conditioned on X

$$E(Y-X \mid X)$$
$$= E(Y \mid X) - E(X \mid X)$$
$$= X - X = 0$$

$$\mathrm{Cov}(k, X) = 0$$

constant

So, $\mathrm{Cov}(Y-X, X) = 0$

Now, are $Y-X$, X independent? Let $T = Y - X$.

Showing Independence:

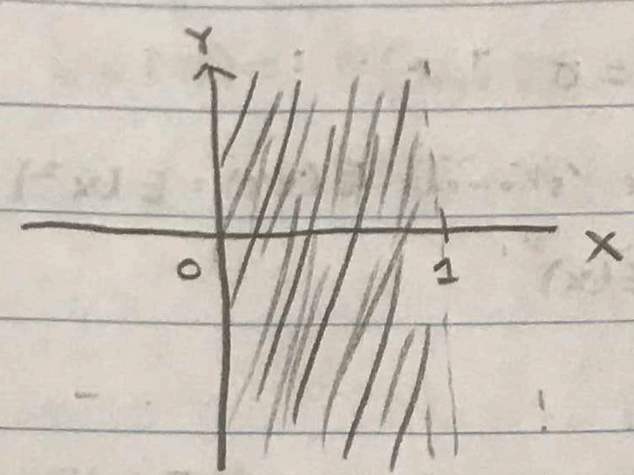
1) $f_{T,X}(t,x) = f_T(t) \cdot f_X(x)$ for all (t,x)

2) $F_{T,X}(t,x) = F_T(t) \cdot F_X(x)$ for all (t,x)

These two methods are about finding the joint pdf/cdf, then factoring into distinct/separate functions of T and X for independence. DO NOT DO THIS HERE.

First, you'd need to find $f_Y(y) = \int_{x \in X} f_X(x) \cdot f_{Y|X}(y|x)\,dx$

here is 0 to 1 b/c $X \sim U(0,1)$



3) Show that the distribution $T|X$ does not depend on X for all (t,x)

$Y|X \sim N(X, X) \Rightarrow$ Given

Thus,

$Y - X|X \sim N(0, X)$ We let $T = Y - X$

$\Downarrow$

$f_{T|X}(t|x) = \frac{1}{\sqrt{2\pi x}} e^{-t^2/2x}$ Since $T|X$ does depend on X for some (t,x), T and X not independent

The key here is recognizing that b/c $Y|X$ is normal, so is $Y-X|X$.

Again, avoid integrating where you can

c) $\boxed{\text{Double expectation of probabilities} \quad P(A) = E[P(A|X)]}$ ☆

$$P(X > Y) = 1 - \underbrace{P(Y<X)}_{} = 1 - 1/2 = 1/2$$

$$\underbrace{P(Y<X)}_{A} = E\Big[\underbrace{P(Y<X|X)}_{}\Big] = E(1/2) = 1/2$$

Here, we are looking at P where $Y|X < X$. Since $Y|X \sim N(X, X)$, by symmetry, the P that $Y|X$ is less than its mean is $1/2$.

d) $$P(Z_{n+1} < Z_n \mid \underbrace{Z_n = \max(Z_1, \dots, Z_n)}_{\text{Call } A}) = \underbrace{1 - P(Z_{n+1} > Z_n \mid A)}_{}$$

Easier to work like this b/c we want/now have $Z_{n+1} >$ the largest of $Z_1, \dots, Z_n$

Apply Bayes' Rules

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(Z_{n+1} > Z_n \mid A) = \frac{P(A, Z_{n+1} > Z_n)}{P(A)}$$

B/c $Z_1, \dots, Z_{n+1}$ is a sequence of iid rvs, this becomes a counting problem

$$P(A) = \frac{(n-1)!}{n!}$$

$\to Z_n$ is fixed at the top spot, so $(n-1)!$ free spots

$\to n!$ ways to order $Z_1, \dots, Z_n$

$$= 1/n$$

Z_n, Z_{n+1} fixed at top and; $n-1$ free

$$P(A, Z_{n+1} > Z_n) = \frac{(n-1)!}{(n+1)!}$$

$\rightarrow (n+1)!$ orderings of $Z_1, \dots, Z_{n+1}$

$$P(Z_{n+1} < Z_n \mid A) = 1 - \frac{\frac{(n-1)!}{(n+1)!}}{\frac{(n-1)!}{n!}} = 1 - \frac{n!}{(n+1)!} = 1 - \frac{1}{n+1}$$

$$= \boxed{\frac{n}{n+1}}$$

e) Let $W = Z_1 - Z_2$

B/c Z_1, Z_2 are iid $N(0,1)$, $W \sim N(0,2)$

We could then say $T = \frac{Z_1 - Z_2}{\sqrt{2}} \sim N(0,1)$

$$E[|T|] = \int_{-\infty}^{\infty} |t| \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt$$

$\rightarrow$ or recognize both pieces are symmetrical about 0

$$= \int_{-\infty}^{0} -t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt + \int_{0}^{\infty} t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt$$

2a)

$$= 2\int_{0}^{\infty} t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt = \frac{\sqrt{2}}{\sqrt{\pi}} \int_{0}^{\infty} t e^{-t^2/2} \, dt$$

Let $u = -t^2/2$

$du = -t \, dt$

$$= -\frac{\sqrt{2}}{\sqrt{\pi}} \int_{0}^{\infty} e^{u} \, du$$

$$= -\sqrt{\frac{2}{\pi}} \left[e^{-t^2/2} \Big|_{0}^{\infty} \right]$$

$$= -\sqrt{\frac{2}{\pi}} \left[0 - 1 \right] = \sqrt{\frac{2}{\pi}}$$

We want $E[|Z_1 - Z_2|] = E[|W|] = \sqrt{2}\, E[|T|] = \frac{2}{\sqrt{\pi}}$

Qaqish also did this.

$$2\int_0^\infty t \frac{e^{-t^2/2}}{\sqrt{2\pi}} = 2\int_0^\infty t\, \phi(t)\, dt$$

$$\phi(t) = \frac{e^{-t^2/2}}{\sqrt{2\pi}}$$

$$\frac{d\,\phi(t)}{dt} = \frac{e^{-t^2/2}}{\sqrt{2\pi}} \cdot -t = -t\,\phi(t)$$

Thus, $2\int_0^\infty t\,\phi(t)\,dt = 2\left[\phi(t)\right]\Big|_\infty^0$

as you're integrating $-\frac{d\,\phi(t)}{dt}$ w/ respect to t

2a) First, it should be evident that $E(\bar{Y}) = \mu$ by WLLN

or $E\left(\frac{Y_1 + Y_2}{2}\right) = \frac{1}{2}\left[E(Y_1) + E(Y_2)\right] = \frac{1}{2}[\mu + \mu] = \mu$

$$f_Y(y|\mu) = \prod_{i=1}^{2} \frac{1}{\mu} e^{-y_i/\mu} = \left(\frac{1}{\mu}\right)^2 e^{-\sum_{i=1}^{2} y_i/\mu}$$

This can be written in exp. family form

$$h(y) = 1 \quad c(\theta) = \left(\frac{1}{\mu}\right)^2 \quad w(\theta) = -\frac{1}{\mu} \quad t(y) = \sum_{i=1}^{2} y_i$$

I'm not sure what $\{(w_1(\theta), \dots, w_k(\theta)) : \theta \in \Theta\}$ contains an open set in R^k means, but since $w(\theta) = -\frac{1}{\mu}$ and $\mu > 0$, I'm assuming it's okay.

So, $T(Y) = \sum_{i=1}^{2} Y_i$ is a CSS for μ.

Since $E(\bar{Y}) = \mu$, so $\bar{Y}$ is an unbiased estimator for μ and $\bar{Y}$ is a function only of $\sum_{i=1}^{2} Y_i$ (a CSS for μ), by Lehmann-Scheffe Theorem, $\bar{Y}$ is the UMVUE for μ. c)

b) $X_1 = Y_{(1)} = \min(Y_1, Y_2)$

$$F_{Y_1}(y) = P(Y_{(1)} \le y)$$
$$= 1 - P(Y_{(1)} \ge y)$$
$$= 1 - [P(Y_1 \ge y)]^2$$
$$= 1 - [1 - P(Y_1 \le y)]^2 = 1 - [e^{-y/\mu}]^2$$

$$P(Y_1 \le y) = \int_0^y \frac{1}{\mu} e^{-y/\mu}\, dy = -\int_0^y e^t\, dt$$

$t = \frac{-y}{\mu}$

$dt = -\frac{1}{\mu}\, dy$

$$= -\left[e^{-y/\mu} \Big|_0^y \right]$$
$$= -\left[e^{-y/\mu} - 1 \right]$$
$$= 1 - e^{-y/\mu}$$

$$f_{Y_{(1)}}(y) = 0 - 2 \cdot [e^{-y/\mu}] \cdot [e^{-y/\mu}] \cdot \left(-\frac{1}{\mu}\right)$$
$$= \frac{2}{\mu} e^{-2y/\mu} \qquad 0 < y < \infty, \quad \mu > 0$$

Thus, $X_1 = \min(Y_1, Y_2)$ is distributed as exponential w/ mean $\mu/2$

Think of finding the expected value of a die roll

Basically, we can "break up" or partition the E(R) into the two possible distinct events multiplied by the probability of that event occurring

c) $E[R] = \underbrace{E[R \mid Y_1 > Y_2]}_{\text{Event 1}} \underbrace{P(Y_1 > Y_2)}_{P(\text{Event 1})} + \underbrace{E[R \mid Y_2 > Y_1]}_{\text{Event 2}} \underbrace{P(Y_2 > Y_1)}_{P(\text{Event 2})}$

Since R is $|Y_1 - Y_2|$, ~~there~~ possibilities are $Y_1 > Y_2$ or $Y_2 > Y_1$.

$P(Y_1 > Y_2) = P(Y_2 > Y_1) = 1/2$ b/c we have two iid continuous random variables

$E[R \mid Y_1 > Y_2] = E[Y_1 - Y_2 \mid Y_1 > Y_2]$

$E[R \mid Y_2 > Y_1] = E[Y_2 - Y_1 \mid Y_2 > Y_1]$

Written this way b/c

$R = \max(Y_1, Y_2) - \min(Y_1, Y_2)$

$E[R \mid Y_1 > Y_2] = E[E[Y_1 - Y_2 \mid Y_1 > Y_2, \underline{Y_2}]]$

We use iterative expectation & add a second condition so Y_2 is treated as a constant

Apply the memoryless property & apply to expectations

$\Rightarrow$ Essentially $E[Y_1 - Y_2 \mid Y_1 > Y_2, Y_2] = E[Y_1 + k \mid Y_1 > k]$

$= E[Y_1] = \mu$

By the same logic, $E[R \mid Y_2 > Y_1] = E[E[Y_2 - Y_1 \mid Y_2 > Y_1, Y_1]]$

$= E[Y_2] = \mu$

So, $E[R] = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$

Using mgfs to find the distribution of R

$M_{Y_1-Y_2}(t) = M_{Y_1}(t) \cdot M_{Y_2}(-t)$

As Y_1, Y_2 are iid, the difference of the two can be expressed as the product of their mgfs

$$= \frac{1}{1-\mu t} \cdot \frac{1}{1+\mu t} = \frac{1}{1^2-\mu^2 t^2}$$

This is a double exp. / Laplace distribution

w/ $\mu = 0$ and $\sigma = \mu > 0$ (in our dist. table)

$$f_T(t) = \frac{1}{2\mu} e^{-|t|/\mu} \qquad -\infty < t < \infty$$

$R = |T|$ => monotone from $(-\infty, 0)$ & monotone over $(0, \infty)$

$(-\infty, 0)$ -> Abs. value removed; $t<0$, so two negatives cancel, then plug in $-r$ for t

$R = |T|$; $T = -R$, $t < 0$

$$f_R(r) = f_T(-r) \left| \frac{d}{dr} - r \right| = \frac{1}{2\mu} e^{-r/\mu}$$

$(0, \infty)$ -> $t > 0$, so we have $e^{-t/\mu}$; plug in r for t

$R = |T|$; $T = R$, $t > 0$

$$f_R(r) = f_T(r) \left| \frac{d}{dr} r \right| = \frac{1}{2\mu} e^{-r/\mu}$$

Summing the two pieces, you get

$$f_R(r) = \frac{2}{2\mu} e^{-r/\mu} = \frac{1}{\mu} e^{-r/\mu} \qquad \mu > 0, \quad r > 0$$

Thus $R \sim Exp(\mu)$, so $Var(R) = \mu^2$

d) Recognize that $Y_1 + Y_2 = X_1 + X_2$

$$E(Y_1) + E(Y_2) = E(X_1) + E(X_2)$$

$$E(X_2) = E(Y_1) + E(Y_2) - E(X_1)$$

$$= \mu + \mu - \mu/2 = \frac{3\mu}{2}$$

$$Var(\underbrace{Y_1 + Y_2}_{Y_1,\, Y_2 \text{ iid}}) = Var(\underbrace{X_1 + X_2}_{X_1,\, X_2 \text{ as order stats are not iid}})$$

$$Var(Y_1) + Var(Y_2) = Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2)$$

$$Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$$

$$= E(Y_1 Y_2) - E(X_1)E(X_2)$$

$$= \underbrace{E(Y_1)E(Y_2)}_{\text{b/c } Y_1, Y_2 \text{ iid}} - E(X_1)E(X_2)$$

$$= \mu^2 - \mu/2 \cdot \frac{3\mu}{2} = \mu^2 - \frac{3\mu^2}{4} = \frac{\mu^2}{4}$$

$$Var(X_2) = \mu^2 + \mu^2 - \frac{\mu^2}{4} - \frac{2\mu^2}{4} = \frac{8\mu^2}{4} - \frac{3\mu^2}{4} = \frac{5\mu^2}{4}$$

e) The key is first recognizing that $E[R|\bar{Y}] = \overbrace{g(\bar{Y})}^{\text{Some function of } \bar{Y}}$.

Then, from part A, recognize $\bar{Y}$ is the UMVUE & CSS for μ

$$E[E[R|\bar{Y}]] = E[g(\bar{Y})] = E(R) = \mu$$

1) Using def of CSS:

$$E[g(\bar{Y})] = \mu \quad \& \quad E(\bar{Y}) = \mu$$

Let $h(\bar{Y}) = g(\bar{Y}) - \bar{Y}$

$$E[\underbrace{g(\bar{Y}) - \bar{Y}}_{h(\bar{Y}) \text{ is a function of a CSS of } \mu}] = 0$$

Thus, $h = 0$ is the only function that has $E(h(\bar{Y})) = 0$ for all $\mu > 0$. Since $h(\bar{Y})$ is 0 function, $g(\bar{Y}) = \bar{Y}$, so $E[R|\bar{Y}] = \bar{Y}$.

2) Uniqueness of UMVUE

$E[g(\bar{Y})] = \mu$ & $g(\bar{Y})$ is function of only CSS of $\mu(\bar{Y})$

By Lehmann-Scheffe, $g(\bar{Y})$ is UMVUE of μ. But $\bar{Y}$ is the UMVUE of μ. By uniqueness of UMVUE, $g(\bar{Y}) = E[R|\bar{Y}] = \bar{Y}$