1^{st} Qaqish Review

- 3 problems, 2 hrs per problem

(1) • biggest sourse 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
$\mu = E(X)$

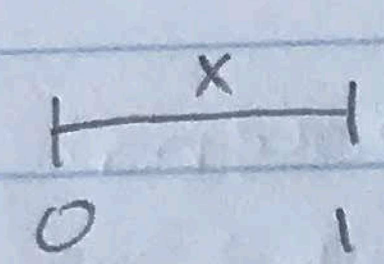
Jensen's inequality: $g(x)$ convex, $E[g(x)] > g(E[X])$
$g(x) = x^2$
$E(x^2) > (Ex)^2 = \mu^2$

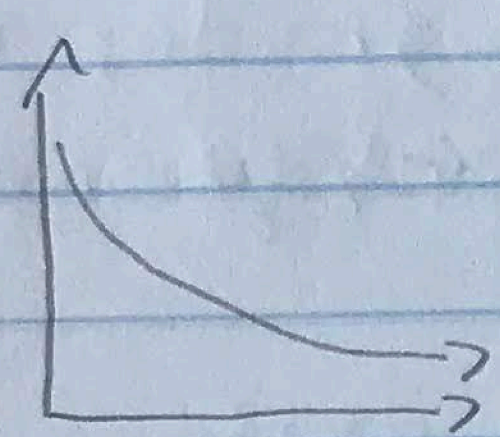
Ψ

- 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)  $X \perp Y$



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

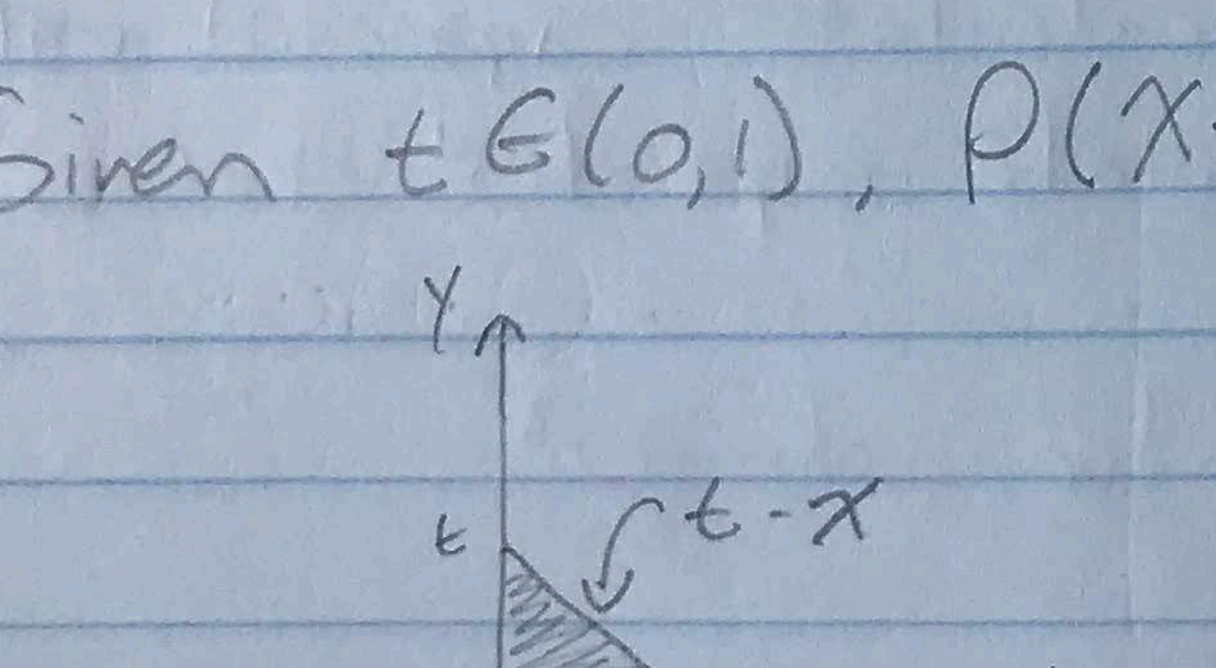
a) and b) are trying to get you to find the expression of T, b/c it is both parts

- "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 $t \in (0,1)$, $P(X+Y \leq t)$



What is the region that corresponds to this event?

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

Don't think of probabilities, think of events

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

$$\underbrace{\int_{x=0}^{x=t} \int_{y=0}^{y=t-x}}_{} f_{X,Y}(x,y)\,dy\,dx$$

most errors are in the ranges!!

$$\int_0^{t-x} e^{-y}\,dy$$
$$= 1 - e^{-(t-x)}$$ ← (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 $t \in (0,1)$

$e^{-t} = 1 - t + \frac{t^2}{2} + \ldots$ ← 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

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

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 ⟹ replace w/ blue	$Z_n - 1$	$6 - Z_n + 1$
Draw Blue ⟹ 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) $L(p|y) = \prod_{i=1}^{n} \{ p^{y_i} (1-p)^{1-y_i} \}$

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

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

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

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

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

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

local max, since continuous, also a global max

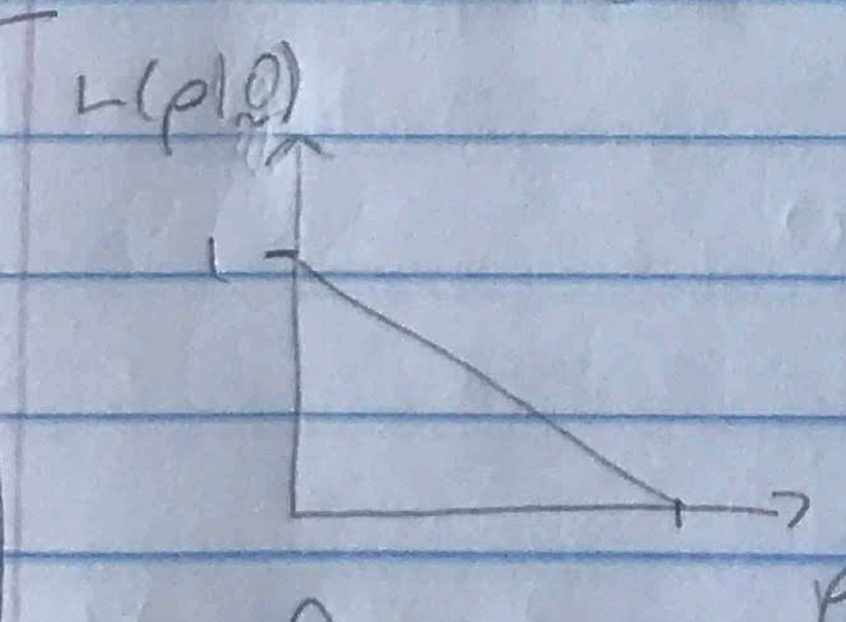
some probability that $\bar{y} = 0$ or 1

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

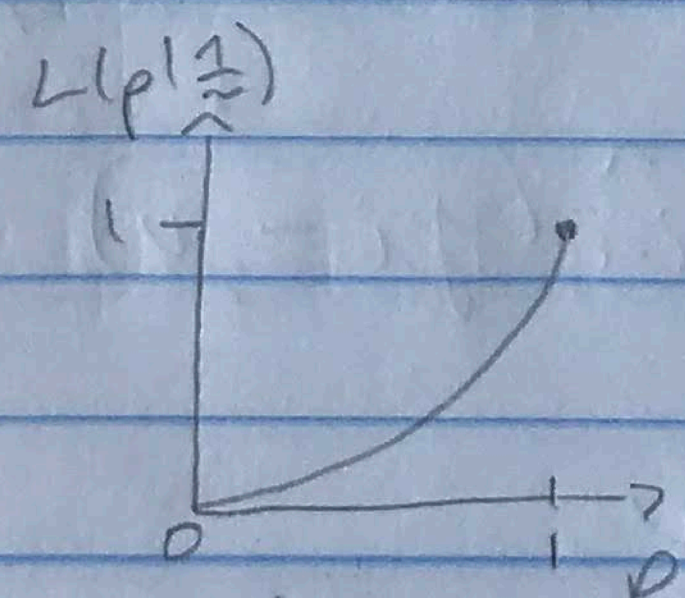
$P(\bar{y} = 0) = (1-p)^n > 0$

So, you need to look at the likelihood function

first consider these two cases

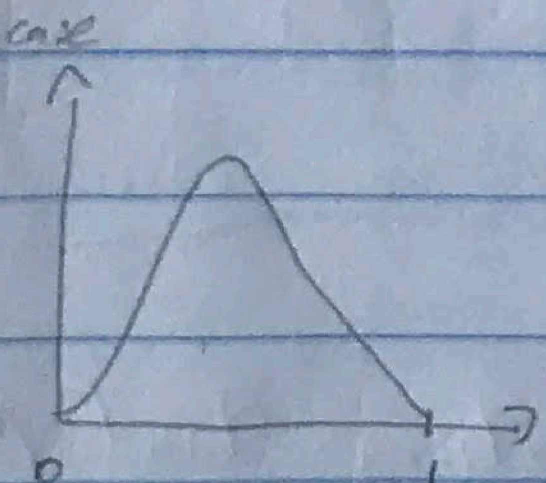


$\hat{p} = 0$



$\hat{p} = 1$

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

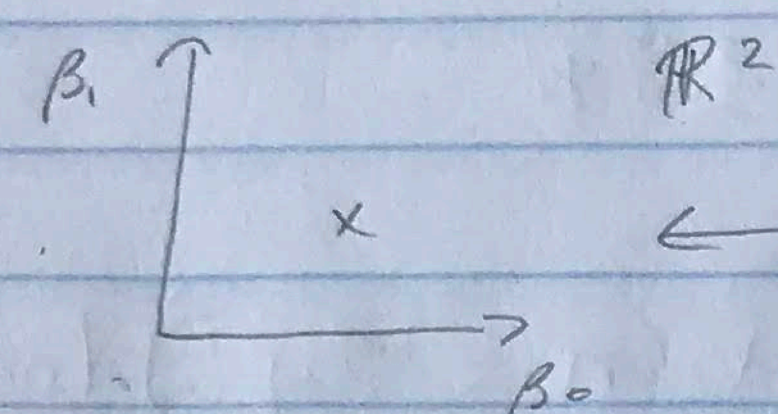


only applicable in this case

#3, 2018

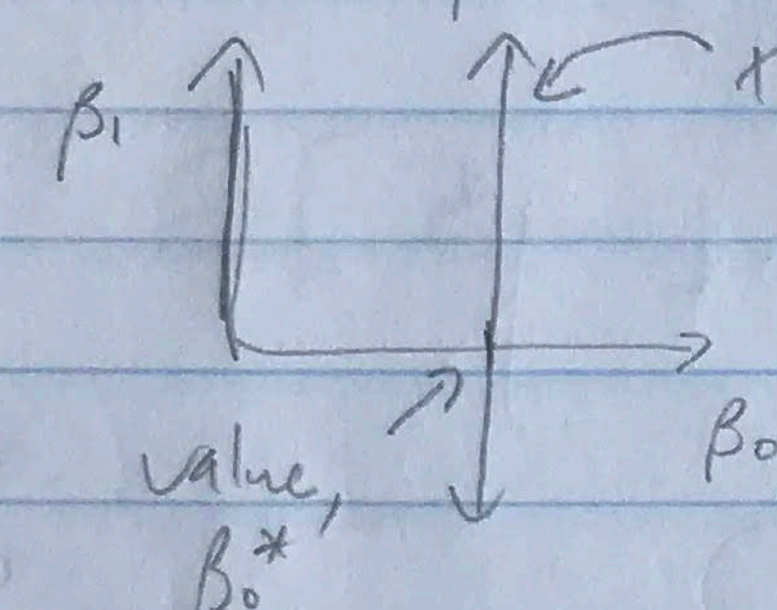
3c) $P_i = E[Y_i] = \dfrac{1}{1+e^{\beta_0 - \beta_1 x_i}}$ $x_1, \dots x_n$ given

if β_0 and β_1 unknown, then the $\underbrace{\text{parameter space}}$: (always need to think about)



← 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 parameterspace, $\mathbb{R}^1$

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

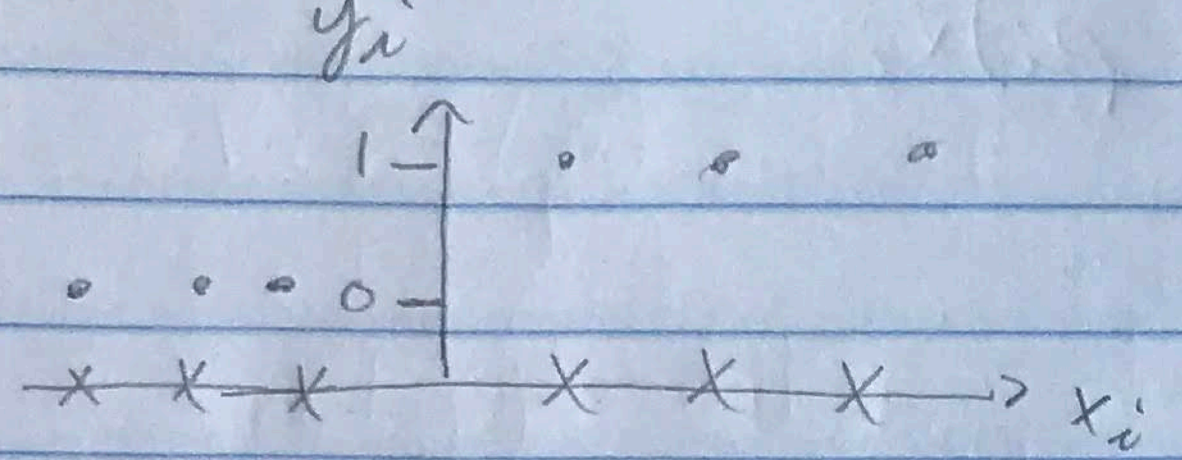
$\sum_{i=1}^{n} y_i = n\mu$

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

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

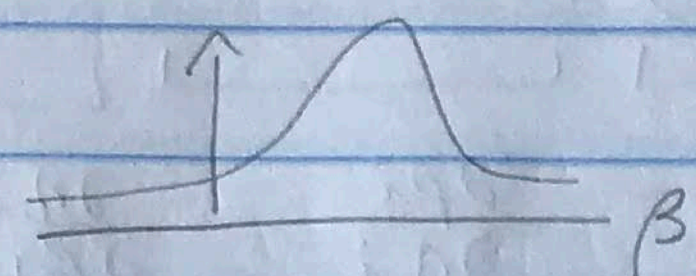
2018, #3

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.

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

independent

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

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

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

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

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

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

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

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

cancellation is 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. $\frac{\partial \ell(\beta_1 \mid \underline{Y})}{\partial \beta_1} = \sum_{i=1}^{n} x_i (Y_i - p_i)$ ↑ 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

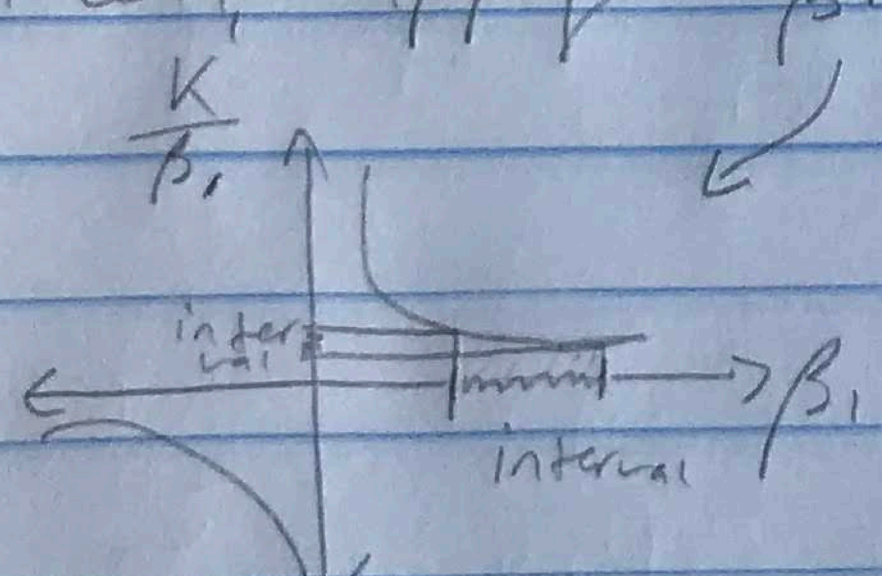
$\frac{K}{\beta_1}$ is monotone function of β_1

So, make CI for β_1 itself

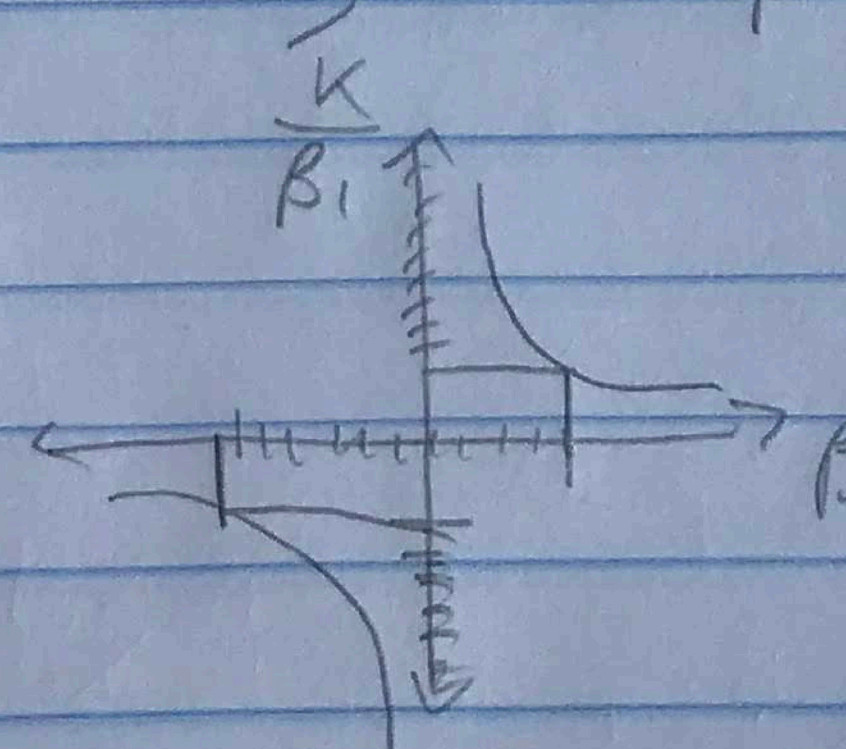
$$\beta_1 \pm Z_{1-\alpha/2}\sqrt{\frac{1}{I(\hat{\beta}_1)}}$$

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

assuming k is positive



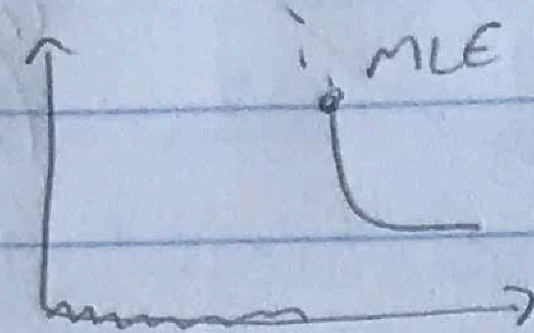
or



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

MLE

$U(0, \theta)$



MLE 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: $\frac{e^{-\nu} \nu^{y}}{y!}$ if all y are 0: $e^{-\nu}$: