Statistics Comprehensive Viva Questions

B. Stat 22–25

April 2025

***Please list the main questions, and suggestions if any***

# 28/04/2025

## Aishik Raysharma (BS2204)

Questions:

KD: What do you understand by skewness? Define. How does the measure defined by you explain the intuition? … AKG gave a hint …

SKD: Suppose you have n iid observations from U(1-theta,1+theta). Suggest a pivotal quantity for theta….did….is it a function of the minimal sufficient statistic? …Couldn't answer…what is the minimal sufficient statistic in this case?... Calculated….

KD: Discuss the paper presented by Tanishq. SKD: What is the exact expression of the FDR that the procedure controls? AKG: Compare with Bonferroni.

Comments/Suggestions:

## Akurathi Syam Prakash (BS2205)

Questions:

X ~ NBin (r, p) (single observation)

Test wheather p= 1/2 or not, construct estimate for p

Is it consistent or not biased or not (r/X)

explain kendalls Tau

give an example where correlation fails but kendalls Tau gives better results

Does Corr =1 => Tau =1 ? Yes

Does Tau =1 => Corr =1 ? No

If corr =1 can introducing one point make it negative (Yes)

If Tau =1 can introducing one point make it negative (No)

Asked to explain swagato’s paper, I told haven't read few papers. Then asked to explain sourav’s paper

Comments/Suggestions:

## Ayush Mandal (BS2212)

Questions:

D. Paul: Under heteroscedastic residuals, what can we do meaning estimate of beta, then why weights are 1/sigma2(Xi) better than any other weight

Skd: X1, X2,..., Xn~ Exp(lamda) find Pivot, CI, then given ci = (0.5,5.5) and Ho: lamda=1 vs H1: lamda ≠ 1, why will you accept it

KD: Explain Urjit paper intuition , aim

Comments/Suggestions:

## Drohan Chakraborty (BS2215)

KD,AKG,SKD

Questions:

KD: Which topics would you prefer? Although there should not be any preferences ideally.

Me: Stat 2, Parametric

KD: Ei sobai same keno bolcho….er baire thekei proshno kori?

Me: Sir ektu oshubidha hobe korle

SKD: Achha Bootstrap Jackknife jano to? Okhan theke kori tahole?

Me: Ha sir.

KD: Suppose the data obtained from a certain distribution is (1,2,...,7). A bootstrap sample of size 7 (srswr) is drawn from the original sample. What is the probability that the median of the bootstrapped sample is 3?

(Answered)

KD: (to SKD) Achha tumi tahole parametric theke jigesh kore nao kichu.

SKD: Suppose X\_1,...,X\_n are independent with X\_i being sampled from Bernoulli(1/(1+exp{i\theta})). How will you test the hypothesis that \theta=0 against the alternate that \theta≠0? Try to name a good test with some desirable properties.

(Answered, fumbled a lot)

KD: Explain Tanishq’s paper

(Explained)

SKD: What is the definition of power in FWER?

(Couldn't give a proper answer)

Comments/Suggestions:

## Koustav Chowdhury (BS2217)

Questions:

1.

Comments/Suggestions:

## Koustav Goswami (BS2218)

Questions:

1.

Comments/Suggestions:

## Rishika Gupta (BS2221)

Questions:

1.What is one factor Anova ? Define the model. Why are constraints necessary in the model? Is there any particular constraint that is good to use , which and why?

2.⁠Define sufficient statistics. X\_1,X\_2,X\_3 ~ iid Ber(theta) . Write down some sufficient statistics for it. Write a 2 dimensional sufficient statistics. How many sets are there in the partition created by the statistic X bar? Create a statistic that is not sufficient and has a partition with sets greater than the one by X bar , and another statistic with less sets in the partition. Is X\_1+2X\_2+ 4X\_3 sufficient statistic?

3.⁠Explain Saptashwa ‘s paper

Comments/Suggestions: meow

## Ritwick Pal (BS2222)

Questions:

1.

Comments/Suggestions:

## Rudrayan Kundu (BS2224)

Questions:

1.

Comments/Suggestions:

## Saptashwa Baisya (BS2226)

Questions:

1.

Comments/Suggestions:

## Shambo Saha (BS2227)

Questions:

1. (SKD) X has density f\_θ. If I say this has MLR property, what does it mean? Define MLR property.

Me: Gave the definition

Does the increasing property have to hold for all real x?

Me: No, only for x in the union of support of f\_θ1 and f\_θ2.

What is uniformly least powerful test?

Me: Defined it just like UMP

For H0: θ=θ0 vs H1: θ<θ0, find ULP. Justify without using Karlin-Rubin.

Me: If the family has MLR in T(X), ULP rejects when T(X)>c, c found from size criterion. The test is equivalent to a likelihood ratio criterion, and proof can can be completed in the same way as for NP lemma.

2. (KD) Do you know partial correction? Define it.

Me: Defines it

Is it possible that X,Y are positively correlated but partial correlation of X,Y given some Z is negative? Give intuition/example.

Me: Suppose Z is a covariate, and B is an independent Bernoulli. Say X=Z+B, Y=Z-B, then it holds.

Now give a real life example.

Me: Suppose two companies are trading with each other, so if one incurs profit, other incurs loss. Then if overall the economy is growing, then both will have an upward trend of profits, then the overall correlation is positive. But conditioning on profit incurred from other trades, their profits will be anti-correlated.

3. (AKG) Do you know V-statistic?

Me: I only know for bivariate, and gave the expression.

Give an estimate for μ².

Me: The nonparametric MLE gives Xbar².

Is this a V-statistic?

Me: Yes, since I can write it as a V-statistic type sum.

Is this unbiased?

Me: No, Xi² terms will introduce bias.

What do you get if I do the jackknife bias correction? Is this a U-statistic?

Me: Wrote down the associated U-statistic, yes it is a U-statistic. Said the general result was proved in class.

If the bias of an estimate is of the order 1/√n, will jackknife bias correction reduce the order? Prove it.

Me: Simple computation of expectation, then Taylor series.

4. Explain Shreetama's paper.

Comments/Suggestions: SKD didn't explicitly say to not use KR theorem but I realised from his manner that he wanted a more fundamental justification. AKG seemed a little dissatisfied when I didn't compute the bias corrected jackknife estimator for μ² and just used the result.

## Shreetama Bhuniya (BS2229)

Questions:

1.

Comments/Suggestions:

## Somarddha Das (BS2232)

Kd, Akg, SD, PC

Questions:

1. PC: What have I taught?

Me: estimators, comparing estimators,..

PC: If variance of an estimator is less than another, then is it better?.(.no.. we compare mse..) Give 2 estimators T\_1, T\_2 such that var(T\_1) less than var(T\_2) but mse is large?

Me: N(mu,1)... Xbar and 1

PC: no.. mse is zero if my is 1

Me: poi(lambda).. (-1)^X for e^-2lambda (not true)

PC: leave it

2. Kd: how will you test if a coin is biased?

Me: Binomial pmf… said asymptotic normal test when asked for another

3. Akg: How will you test for 2 coins p1=p2?

Me: 2 sample test .. asymptotic normal

For exact test.. gave the conditional test

Asked saptashwa's paper.

Comments/Suggestions:

## Soumil Amul Sanghrajka (BS2233)

Questions:

1. PC: what all have I taught you

Response: UI, LR, ...

PC: in Stat 3

Response: fisher schoring, estimating g(theta)....

PC: Asked continuos mapping theorem.

Follow up questions by PC:

Is theta(hat) unbiased enough or something else is needed?

What kind of function is g ( I said cont. And differentiable as we use taylor series, but only cont. Is enough)

PC: write down: root(n) (theta hat - theta) follows normal (0, sigma^2(theta) ) asymp.

Now what is the asymp distribution of g(theta\_hat).

PC: What kind of a function is g. (This time cont. And differentiable😇)

PC: Why is variance stabilizing transform used, how to get it.

Made me write down: g'(theta)\_square\* (sigma\_sq) = const.

PC: Now, it can happen that using this, g might not be invertible. That's a problem right?

Answer: we get the condition that g' = root(const/sigma\_sq) which is +ve and g is cont. Hence g is monotonic. It has to be invertible for variance stabilizing transform.

DePaul: X1, ......Xn iid N(mu, 1)

Find unbiased estimate of mu^3 and get UMVUE.

Response: (glad I didn't say X1\*X2\*X3 as the answer, cause I had to find UMVUE next. With the method I did, I got unbiased estimate of u^3 as a function of Xbar which is a CSS so the estimate was UMVUE) I said let me check expectation of Xbar^3 and see if I can adjust it .

Wrote, distribution of Xbar, now Expectation of (Xbar - mu)^3 = 0, we know E(Xbar) and E(Xbar\_sq), so we get unbiased estimate of mu^3.

KD: Explain Supratim's paper.

Note: I made a rookie mistake in expanding (Xbar - mu)^3.

PC said go to high school. 🪦🪦

Comments/Suggestions:

## Srijan Saha (BS2238)

KD, AKG, SKD

**Prelude:** I arrived with a walker (my left ankle was plastered due to ligament injury :\_) ).

KD: “Parbe uthte?” (Can you walk up the stage?). I demonstrated (as if I had another choice).

SKD was generous enough to provide me with chalk and duster ( :) ).

Despite already being explained to AKG during the group presentation about the ankle and **firsthand seeing that** earlier, he asked as if he totally forgot…

AKG: “What happened to your leg?”

KD: “Football!”

AKG: “Oh! In the ISI playground?”

Me: “Yes”

AKG: “When will the fracture be repaired?”

Me: “I didn’t fracture, the ligament is injured”.

AKG (with utter confusion): “Ligament kibhabe fracture hote pare” (How can the ligament be fractured?)

KD (Diplomatically): “Na, chhirte to pare” (No, it can tear).

AKG: “Oh! Kintu paa er patay kon ligament thake je okhane plaster korte holo?” (Oh! But which ligament is in the sole of your feet so that you had to plaster like that?)

Me: “There are (in fact) ligaments in the ankle.”

AKG (Moment of Realization): “Oh! I see.”

AKG’s doubt (as well as his viva questions) were over.

KD asked about preference. I said Stat 2, Stat 3 and Parametric Inference.

Now the **Questions**:

1. (By KD) X1,...,Xn and Y1,...,Yn are two samples. One of them is from Cauchy, another from Normal, but you don't know which is from which distribution. You can take as many samples as you want. How will you identify the distribution of the sample?

KD started by saying, “**Suppose**, there are two distributions - say Normal and Cauchy”. I mentioned the Likelihood Ratio. I didn’t initially understand the focus on the particular example of Cauchy. I fumbled a bit, but once KD clarified that if I knew “something” about Cauchy, I got the answer. So basically plot the r-th running mean against r (=1,2,...). By Strong Law of Large Numbers, it should converge for Normal, so the graph will hit an asymptote. For Cauchy, the mean doesn’t exist, so the graph will keep oscillating.

[A better solution that I thought later of, was to take the running mean of absolute values, instead of just the values itself. Then for Normal, it converges **finitely**. For Cauchy, it converges (or diverges) to **Infinity** by the version of SLLN for infinite mean. However, that’s way above the paygrade, I acknowledge, considering the question was based on “Simulation” from Stat 2]

AKG supported my earlier answer of Likelihood and said, “You use that for a finite sample. It works well then.” I asked, to their surprise, “I thought so. But, why does it work? Not clear to me”. It caught them off guard.

AKG (Looking around at other professors): “Keno kaaj kore? Pore ek din kotha hobe sei niye” (Why does it work? Well, let’s talk about that another day, not today).

2. (By SKD) Define CSS (Complete Sufficient Statistic). Prove |X| (Norm) and X/|X| (Unit Vector) are independent, if X ~ N(0,c^2\*I\_p) (p-variate normal) using CSS.

Easy (but lengthy, I realized once I started explaining). Defined Sufficient Statistic. Defined Complete Statistic. Stated the result for CSS for Full Rank Exponential Family. Hence showed using the pdf of Normal, that |X|^2 was CSS. Said that X/|X| is ancillary for the parameter. SKD asked why. I said it is a scale family. He said, “So what?” Then I proved explicitly why the distribution is parameter-free. By Basu’s Theorem, since CSS and Ancillary statistic are independent, so is |X|^2 and X/|X|. Then |X| is (one-one) function of |X|^2. So |X| and X/|X| are independent.

After all these explanations, he said that I am missing a minute detail and hinted at “Natural Parameter” in the definition of Full Rank Exponential Family. I couldn’t understand well then, can you?

3. (Again KD) Explain Souvik's Paper (one of the difficult ones imo). I explained the core idea of the paper to KD. He specifically asked some questions to check if I knew the setup of the problem, such as the scheme of the sampling (Cluster), weighting method (IPW) etc. He was pretty satisfied (I wasn’t, I felt like I fumbled and could not remember the formulae).

**Comments/Suggestions:** Be prepared to explain the core ideas of the difficult papers and be honest about how much you understood and you didn’t. They may be impressed at your efforts only. You don’t have to be precise.

Same for other questions (especially if they are creative). Think aloud, share the ideas and do the computations to support your ideas. Don’t discourage your own ideas. Explain why you thought of it in the first place (In my case, for the first question, I clarified I didn’t understand the emphasis was on Cauchy distribution).

Also for easy but lengthy questions, don’t rush to answer. Form the answer fully in your mind, then proceed to explain.

And maybe, when they ask for topic preference, hint at the one you presented during the topic presentation (if you are confident with it). Prepare that one preferred topic well beforehand.

And last but not the least, “Don’t let them know your next move!” (referring to counter-questioning AKG). Jokes, please don’t take it seriously (: .

## Suvankar Saha (BS2244)

PC, AKG, KD, D Paul present

Questions:

KD : What is your preference?

Me: Any topic

KD: Probal Da apni qn korun.

PC: Ami ki ki porie6ilm?

Me: Estimation, Testing, regression etc. Sir Stat IV ta ektu weak.

PC: Egulo tomar strong !

Me : Thik thak.

Here are the questions:

PC: Consider N(mue, sigma²). List some estimates of sigma².

Wrote U.E. and MLE of sigma².

PC: Which one has a smaller MSE?

PC: Among the class c times sum(X\_i - Xbar)²?

KD: Write pdf of Laplace.

KD: How do you generate from Laplace?

Answered the one using coinflip (or rademacher).

KD: You have two machines, each of them generated 10 samples. One machine generates from Standard Normal another from standard Laplace. So you have two sets of data each containing 10 data points. How do you guess the distributions? … Without hypothesis testing…what do you guess just by seeing the data points.

PC: Suppose now you have two sets of data points each containing a single observation instead of 10. How do you guess now?

PC: Find the probability that your guess is wrong.

AKG: In your first question you generated from Laplace in two steps. How do you generate in just one step.

KD: Explain Urjit’s paper.

Comments/Suggestions:

## Swastika Dey (BS2246)

PC, Dp, Kd, Akg present

I entered with a pale face. (I wasn't so much afraid but PC was there, so to try to avoid any mishap I acted like that🙃) They told me to sit in the chair first.

Kd asked, “which are your preferred area? “

Me: I haven't read lsm properly(Told that because I knew I was not able to handle if PC asks something on that).

Kd: Baki sob porecho? Konta preferrable?

Me: Sob i porechi motamoti, konotai temon preferrable noy.

Then Kd asked PC: “ Probal da, will you ask first? “ ( My Bp raised slightly)

PC: “No. You ask first “.

Questions:

Kd: “how to measure kurtosis?”

Me: “it measures the heaviness of the tail and peakedness. “ Also told the formula, m\_4/m\_2^2 and when it is lepto, meso, platykurtic.

Kd: “Take a chalk and explain by drawing graphs “

Me: Drew the graphs and explained which one is lepto, meso or platykurtic. (He looked satisfied)

Kd: “OK, now tell why we use m\_4 and why > or < or = 3 work? “

Me: (Thinking… Thinking …nothing was coming in my mind.)

Kd: “Acha, Charo, bodh hoy jano na, Debashis da, tumi jigges koro”

Dp: “X(single observation ) poi(θ), find ue of θ^2.

Me: (Take some time to compute) I answered “sir, it is coming x(x-1)”.

Dp: “OK, right. Now tell same for n iid poi(θ) random variables. “

Me: (without writing anything) Sir, x\_1x\_2.

Dp: (He thought I told him to repeat the question) “No, I said no iid random variables, not only x\_1, x\_2. “

Akg: “Na, o answer ta x\_1x\_2 bolche”.

Dp: “Ooh yes! That's correct, but give me another ue.”

PC interrupted him…

PC: “First tell me what is the umvue of θ^2 for single observation from poi(θ). “

Me: (At first I thought he wanted me to show using the power series convergence that x(x-1) is the only possible ue and hence must be umvue) “Sir, here x(x-1) is the only ue and hence umvue, because T(X) to be ue E(T(x)) = θ^2 and hence sum(T(x) θ^x / x!) = sum(θ^2 θ^x/x!) = sum(x(x-1) θ^x/x!) and by power series comparison…)

PC stopped me and told: “ OK, but I am wanting something else, what can you tell about x here? Poisson and some other distribution has a very special property, what is it? “

Me: (then I realised he is wanting me to show umvue through css) “sir, this belong to exponential family and hence x is css so being an unbiased func of css x(x-1)is umvue “

PC: “Yes, correct, now find a umvue for n iid poi(θ) “

Me: (Now I got the idea that he is wanting unbiased estimate which is a func of css, i.e, X\_ber. Started to look for ue of θ^2 in terms of X\_ber. First I did a computation mistake.) “sir, here X\_ber(X\_ber - 1) (error) is a ue of θ^2 and being a func of css it is umvue. “

PC : “ You have done some computation mistake, find where and correct that “

Me: (realised in a while, I have a 1/n in var(X\_ber) computation) “ yes sir, sorry, it is X\_ber (X\_ber - 1/n). “

PC : (looked satisfied) “ KIRANMOY, amar hoe geche, tomader kichu thakle dhoro”

Kd (to Akg) : “kichu dhorbe Anil da? “

Akg : “na, na, taratari koro, paper ta dhore nao, bari jabo”(He was seeming exhausted and impatient and was like he is wishing as soon as possible all vivas be over!!!)

Kd: ” tell about Sourav's paper “

Me: started to tell, messed up, can't explain properly. Also Akg repeatedly told, “taratari nao”, then he told, “darao, sudhu estimation method ta bolo(when I was telling why we need to take a reparametrization of beta here).“ I wrote the log likelihood and told that from score funcs equal to 0, we can't solve so do fisher scoring…and the initial estimate of β is… “

Akg: (impatiently) (actually it was at about 6:40 pm and two more students were left) “thikache chole jao, osob bolte hobe na, board ta muche chole jao“. (Also I don’t think he was satisfied by my explanation.)

Comments/Suggestions: Professors help a lot. 🙂

## Tanishq Kumar Prasad (BS2247)

Questions:

*(enters)*

AKG: Esho esho, tomake to shokal thekei dekhchi dariye acho.

KD: O kalkeo ashbe dekhbe. Probal da kicchu jigesh korben?

PC: Na amar kichu korar nei aage tumi koro.

KD: Ki preference tomar? [Me: Off stat na korlei holo.]

KD: Na na oshob korbo na. Bolo what is skewness? [Me: (writes definition)]

KD: Definition ta baad dao, tumi intuition ta dao. [Me: Skewness is basically expectation of (X-EX/SD(X))^3; it indicates which side the tail of the data lies, since (X-EX)/SD(X) tells about the deviation from the centre. We take the exponent as three as it will maintain the sign of the extremity.]

KD: Tahole three er jaygaye jekono odd number cholbe to? [Me: Yes]

KD: Do you know any other measure of skewness? [Me: Nah]

KD: Not even Bauli’s Measure of Skewness? [Me: Never heard of that *(All were surprised, sem 1 stuff)]*

PC: Amar mone ache ami porai ni eta, kintu Kiranmoy er note a dekhechilam [*(KD laughs in agreement )* Me: Nonetheless amar mone nei]

PC: Ok tahole derive korabo. Consider a distribution without any moments. How would you define a measure of skewness here? [Me: I will replace mean by median and standard deviation by interquartile range.]

PC: What will the definition be then? [Me: (wrote the same definition as original skewness, but with r in the power, r being any odd number, replacing mean with median and sd with iqr)]

PC: This is a mean based definition since you are averaging, do you think it will be good in this case? [Me: No moments do not exist (replaces mean with median)]

PC: See for r=1, this is zero (infact it is always 0, no clue why I wrote that). This is not something very good. Tell me how the quartiles look for left, right and non skewed. [Me: (drew them)]

PC: What do you observe? [ Me: q3 is closer to q2 in negatively skewed, q1 is closer to q2 in positively skewed. So we can take the difference of their distances in numerator]

PC: Fine.

AKG: Ebar ektu inference dhori, suppose you have an iid sample from U(0,\theta), will UMP for H0: \theta = \theta\_0 vs H1: \theta > \theta\_0 exist? [Me: Yes (argued using test free of alternate \theta)]

PC: Is any special property being used here? [Me: Yes, MLR]

AKG: What about \theta \neq \theta\_0? [Me: exists, we can create a test which matches with the power function for both \theta > \theta\_0 and \theta < \theta\_0]

PC: Alright, so same argument should hold with normal; UMP exists there right? [Me: No, for normal using uniqueness of UMP if a test matches with the power function of \theta>\theta\_0, it cannot match with the other]

PC: Then this argument should also hold in uniform case. Why does it not work there? [Me: (confused a bit, thinking)

PC: (hint) How did you conclude in normal it cannot happen? [Me: Uniqueness of UMP, so if it were unique here then same argument would hold, that means it is not unique here]

PC: Yes, there are several UMPs here [Me: Yes, so we can choose UMPs in a way that the power function matches]

AKG: Why is UMP not unique here? [Me: Support depends on parameter]

KD: Explain Soumil’s paper. [Me: (proceeds to start with Shreetama’s paper (wtf?) since they are similar, realise what I’m doing, then turn around and say sorry this was Shreetama’s)]

Comments/Suggestions: Speak slowly and carefully. Do not panic or overload your mind before the viva. For the papers, being thorough with the core ideas should be enough; not much stress was laid on data analytic results/simulations.

## Shreyansh Mukhopadhyay (BS2147)

Questions:

1.

Comments/Suggestions:

# 29/04/2025

## Aaryan Kumar (BS2201)

Questions:

1. KD -

Why we can't simultaneously minimize type 1 and type 2 error ? Is there any case where we can minimize simultaneously?(This can happen in sequential tests.)

2.AKG-

You have a distribution in the support [0,10] prove or disprove that standard deviation<=5.

3.Explain rupsa paper .

Comments/Suggestions:

## Abhroneel Ghosh (BS2202)

Questions:

1. AKG

You have a distribution in the support [0,10] prove or disprove that standard deviation<=5.

You must have learnt some relation between standard deviation and range in school. (sd^2 <= R^2/4)

2. De Paul

X ~ Bern(p). Can you find an unbiased estimate of p^2. (No, only for linear functions of p).

3. KD

Is the MLE always unbiased? Example where it is not.

Is the MLE always consistent? Example where it is not.

AKG

Xi ~ Unif(theta). MLE? Is it consistent? Does it converge a.s? What is Borel Cantelli?

4. KD

Explain Sourav’s paper

Comments/Suggestions:

## Aman Das (BS2206)

Questions:

1. 200 students, to estimate number of girls

* 20 size sample, 2 girls detected

2. Estimate number of girls Total.

X1, X2 ~ uniform(θ, θ+1)

Test θ=1 vs θ=1.2

3. Soumil paper

Comments/Suggestions:

## Anant Consul (BS2207)

Questions:

1.

Comments/Suggestions:

## Avirup Chakraborty (BS2211)

Questions:

1.

Comments/Suggestions:

## Bibhabasu Mandal (BS2213)

Questions:

1.

Comments/Suggestions:

## Biswarup Chakraborty (BS2214)

Questions:

1.

Comments/Suggestions:

## Raj Pratap Singh (BS2219)

Questions:

1.

Comments/Suggestions:

## Rajsuvro Adhikari (BS2220)

Questions:

1. Present KD AKG PC De Paul(very unlucky PC arrived just before my presentation😭)

Kd : Come sit . Probal da tumi ee jiggesh koro naki (Probal da you start)

PC: Suppose you have two multivariate distributions how to check if their means are equal.

Me: looking at PC with a blank face

Pc: ok let me make the problem easier. Write on board

X1,.....Xn follows MVN(mu1, Sigma1)

Y1,.....Yn follows MVN(mu2, Sigma2)

What assumptions would you take to test mu1=mu2? What is the test?

Me: Assumption : Both Sigmas are equal. I will do LR.

PC:Why do you need Sigmas equal?

Me: Answered

PC: Ok now tell me the LR test statistic.

Me: I can derive but I don't remember.

PC: No don't derive . Ok tell me some known test to check.

Me : Hotelling T^2

Pc: Write it down.

Me :Wrote.

PC : Now tell the LR test statistic

Me: I don't remember again.

PC : He seems not so prepared with stat 3.

Ask him something else

Me: {man hi man me--- WTF}

Kd: Debashis ask something

De paul: You have y=alpha+beta\*x + epsilon .

Also -1<beta<1. What would you report your Beta\_hat\_ls and alpha\_hat\_ls.

Me: Now after pc thaap I went full on bengali with de paul . (Said sir alpha r beta axis gulo hole function ta ekta chong(🎉) er moto dekhte hbe ). Drew picture of a convex function in and told the answer.

Debashis Paul: Ok you have done it graphically how would you do it analytically.

Me: beta\_hat\_ls is mle for beta and it must belong to parameter space . For a restricted parameter space I will report 1 if beta\_hat\_ls>1 , beta\_hat when in between and -1 when beta\_hat\_ls<-1 .

De paul: ek ee holo . Give some other method .

Me: Again standing with a blank face . ( Later got to know the ans is change it to a ridge ) .

De paul: Ok you did it correct but not what I wanted to hear .

KD(saviour comes) : Explain saptaswa's paper.

Me:Explained it completely.( Only paper I was very confident ) No one listened except KD . No qns from paper.

Pc: Arekjon ke ki kore nebe .

KD: Na asole bole diechi oder 2 to ee aste abar . Rajsuvro rub the board and you can go.

Viva over

Me : Ebar ektu shanti te ghumabo .

Comments/Suggestions: PC wants the results to be at your fingertips (whatever he taught) . Think out loud . That helps a lot.

## Rupsa Ray (BS2225)

Questions:

1. Regularity condition of crlb …what is crlb…why full support needed…how to get umvue for uniform (0,theta)

2. How to get Conf Ellipsoid for BVN ….min CE why centered around mean?

Explain Tanishq paper main idea

Comments/Suggestions: Ans slowly… ask for time rather saying a wrong ans

## SK Zaheen (BS2230)

Questions:

1. Is MLE always consistent?

2. X follows Poi(lambda). Give an unbiased estimate of e^(-lambda).

3. X\_1, … ,X\_n iid Poi(lambda). Give an unbiased estimate of e^(-lambda). Give its UMVUE.

4. Explain Rupsa’s paper.

Comments/Suggestions:

## Soham Purkait (BS2231)

Questions:

1.

Comments/Suggestions:

## Soupayan Dasgupta (BS2235)

Questions:

what is Kendall tau why is it used ? In 2 data set if Pearson correlation is 1 for both can the pooled correlation be negative? Is it true for kendall tau?if Pearson is 1 or -1 is Kendall tau is also 1 or -1

Is umvue unique, consistent

Explain souvik paper

Comments/Suggestions:

## Sourav De (BS2236)

Questions:

1. What is Cluster Sampling, and 2-stage Sampling?

[Ans: 👍]

Explain using an example.

[Ans: Consider the situation when our survey objective is to measure the effectiveness of a government scheme in West Bengal. Divide West Bengal into districts. In case of Cluster Sampling, randomly choose districts and sample every block in the selected districts. In case of 2-stage Sampling, randomly choose districts and randomly choose blocks within each selected district and sample.]

Which is better in this scenario ?

[Ans: 2-stage Sampling is cheaper to conduct.]

Can we perform Stratified Sampling in this scenario? (Mention advantages and disadvantages)

[Ans: Stratified Sampling is more expensive to conduct, since the construction of homogeneous stratas requires prior survey.]

2. Is MLE always consistent?

[Ans: No]

Give an example of an inconsistent MLE.

[Ans: 🥲]

Is UMVUE always consistent?

[Ans: YES (did not ask for the proof)]

3. Consider X1, X2, … , Xn ~ Unif(theta - ½, theta + ½). Find the MLE.

[Ans: theta\_mle is any estimate in the interval (X\_(n) - ½, X\_(1) + ½)]

Explain.

[Ans: 👍]

Is any estimator in that interval consistent?

[Ans: YES. Any estimator of the form lambda(X\_(n) - ½) + (1 - lambda)(X\_(1) + ½) is consistent.]

Explain.

[Ans: 👍]

If we put lambda\_n instead of lambda, will the estimate be consistent?

[Ans: YES. (asked for proof)]

What will be the UMVUE of theta?

[Ans: Sir, I know very few methods of constructing UMVUE:

I can construct UMVUE using LS Theorem: Here (X\_(1), X\_(n)) is minimal sufficient, but not complete. Therefore, I cannot use this method.

Since not all the conditions for CRLB inequity are met, I cannot check if the variance of any unbiased estimate attains the corresponding lower bound.

Sir, I cannot answer this question. 🥲]

4. Explain Souvik's paper.

[Ans: Sir, I have prepared all the papers, except the papers presented by Souvik and Swagoto.]

Why have you left out these papers?

[Ans: 🤫]

Explain Urjit’s paper.

[Ans: 👍]

Comments/Suggestions:

The sample survey questions were asked by KD and the other questions were asked by PC.

## Souvik Roy (BS2237)

Questions:

1.

Comments/Suggestions:

## Sunirban Sarkar (BS2242)

Questions:

1. Samples from 2 different poisson distribution

Test for lamda1=lamda2

2 . How to simulate from normal poisson and bin distribution

3 Say something about multinomial distribution

4 Soumil’s paper

Comments/Suggestions:

## Supratim Das (BS2243)

Questions:

Professors present: KD, AKG, PC

1. KD: How would you sample from bivariate normal (0,1, 1,3,½)? (Box Muller transform, then linear transforms). Follow up from PC: Is this transform unique? Well no.. Gave a cheeky trivial transform, realized there were non trivial transforms available

2. PC: How to test equality of dispersions in multivariate normal case? Give LRT statistic without derivation.

3. Rupsa’s paper

Comments/Suggestions: KD really helps. PC is cooperative but expects prompt and accurate answers. As for AKG,... All the best to those who’re reading this.

## Swagato Das (BS2245)

Questions:

1.

Comments/Suggestions:

## Urjit Paul Chowdhury (BS2248)

Questions:

1.

Comments/Suggestions:

## Yuvraj Dutta (BS2249)

Questions:

Profs Present: KD, AKG and PC

(Entry)

KD: Sit, sit. Have you prepared for everything or just some selected topics? Can you answer questions from any topic?

(I’m fine with everything. As for whether I can answer everything or not—that depends on what you ask :) )

PC: What have I taught you in Stat 3 and Stat 4?

(Mentioned whatever topics I liked and remembered while skipping the others.)

PC: Write down X₁, X₂, …, Xₙ ~ iid N(θ, 1). Now give me a consistent estimate of θ.

(Immediately mentioned the sample mean.)

PC: Explain why.

(Explained.)

PC: Now give me another consistent estimate for θ, different from the sample mean.

(Mentioned the sample median.)

PC: Is it really consistent? Are you sure? Realllyyyyy???!?

(Like yeah man, of course—but the man did not stop trying to mess with me till the end.

I tried to explain using the asymptotic distribution of the sample median, but he refused to accept that argument and asked me to suppose he doesn’t know about the asymptotic distribution of the sample median.)

PC: So explain why the sample median is consistent for θ without using the asymptotic normality of the median.

(I mentioned the Strong Law of Large Numbers for the median, but he said that there is NO SLLN for medians. I was like, tf you mean bro, of course something like that exists. It might not be called that exactly, but I can prove it for you right now.)

PC: Prove it then.

(Went ahead and proved it. But before I used the Chebyshev’s inequality, while equating two probabilities after transforming the r.v.’s, I made a mistake and wrote an “=” instead of “≤” and the guy just tried to mess me up over that. I was like, it doesn’t really matter whether it’s “=” or “≤” since it’ll go to 0 anyways, but he said, “Write it properly or else just leave the viva.” So I gave up, apologized, and just went ahead and finished the proof.)

PC: I’m done with this question. I’m good now. You all ask. (Started scrolling through his phone again.)

KD: Suppose X₁, X₂, …, Xₙ ~ iid f\_θ, where θ is the mean parameter of f, and assume f is a “good” density.

(I was like, what do you even mean by a good density? He said that I can take whatever assumptions I want on f.)

Now you get an estimate θ̂ of θ using Monte Carlo/SLLN-type stuff. Later on we find out that X₁, X₂, …, Xₙ ~ iid g\_δ rather than f\_θ, where δ is the mean parameter of g. How will you transform your variable θ̂ to get an appropriate estimate for δ?

(The guy wanted to ask a question on importance sampling but messed up the question. I was like, the SLLN says the sample mean converges to the mean of the true distribution of the Xᵢ’s, not the assumed one. So if g is the true distribution, we just use the same estimate.)

KD: Oh, right. In that case, what if you don’t know whether they come from g or f? How will you decide? Give me an intuitive way that uses some statistics but doesn’t need to be statistically the best way.

(I just mentioned that since f and g are known, we can do one-sample KS tests for both and choose the distribution with the higher p-value.)

AKG: KS-tests?!? What is KS?

(Kolmogorov-Smirnov Test.)

AKG wasn’t convinced but KD was.

(So I went on to mention the “statistically correct” way of doing a Neyman-Pearson test here. Also mentioned that we did a similar thing for a single sample where f ~ N(0,1) and g ~ Cauchy(0), and we simplified it to get a rejection region based on |X|.)

KD: Then assume f ~ N(0,1) and g ~ Cauchy(0), it won’t matter. Now answer.

(I was like, if n is very large (~10,000), then we look at the sample mean. If it converges to something, it comes from Normal. If not—if it's erratic—it comes from Cauchy due to SLLN.)

KD: Correct. So what if n is not that large?

(Then we can proceed with the Neyman-Pearson-type test for X₁, …, Xₙ. But I can’t simplify the statistic as I did in the single-observation case. It might be possible, maybe, but I for sure can’t do it.)

KD: Why is that? Try to simplify it.

(I was like... no. Just stared at him for 5–6 seconds until AKG signaled to KD that it won’t be possible, so let’s move on.)

KD: Sure, so suppose you can’t simplify it. Then what do you do? How do you get the cutoffs and everything?

(I said that I’ll bootstrap the distribution of the statistic under the null and explained why it would be invariant, etc.)

AKG: What is bootstrap?

(Explained.)

AKG: Right, so now what if you have two samples, X and Y, such that one of them is Normal and the other one is Cauchy. How will you statistically choose which one is which?

(Said that we can look at the test where the Null is (X,Y) ~ (N(0,1), Cauchy(0)) and the Alternative is (X,Y) ~ (Cauchy(0), N(0,1)). I’ll calculate the two joint probabilities/likelihoods and do something similar to Neyman-Pearson again.)

Both KD and AKG were convinced and said this is what they wanted to hear.

KD: Let’s move on to the papers then. Explain Sourav’s paper.

(Did so.)

KD: That’s it, you can leave now.

Comments/Suggestions: