Main Concepts of Chapter 9 (§9.1 – 9.4)

- Goal: find estimators of unknown paramaters, using various methods and tools.
  - can determine whether an estimator is better than another estimator.

§9.2: Methods of estimation

1) producing a method of moments estimator, or MME

I.e.,
- equate population moments $E(X^r)$ to sample moments $\frac{\sum_{i=1}^{n} X_i^r}{n}$.
- number of moments needed (values of r) corresponds to # of unknown params. Two unknowns? Need $E(X) = \bar{X}$ and $E(X^2) = \frac{\sum_{i=1}^{n} X_i^2}{n}$, solve each.
- The MME $\hat{\theta}$ is the value such that $E(X^r) = \frac{\sum_{i=1}^{n} X_i^r}{n}$. put in terms of statistics
- when $r=1$, $\hat{\theta}$ is the value s.t. $E(X) = \bar{X}$ (assuming only one unknown param θ).
- If don't know special distribution, then ~~might~~ need to derive $E(X)$, $E(X^2)$, etc. manually using pdf.
- If know ~~the~~ an estimator, then given specific values for each R.V. of some given sample size n, can obtain an MM estimate.

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

- When deriving pop. moments, put them in terms of the parameters.
  e.g., for r.s. from $N(\mu, \sigma^2)$, $E(X) = \mu$ and $E(X^2) = Var(X) + [E(X)]^2 = \sigma^2 + \mu^2$

- Invariance Property: If want MME of a function of θ, say $\tau(\theta)$, then $\hat{\tau}(\theta) = \tau(\hat{\theta})$ can be used.

2) producing a maximum likelihood estimator (MLE)

i.e., find $\hat{\theta}$ that maximizes $L(\theta)$

- use $L(\theta) = \underbrace{\prod_{i=1}^{n} f(x_i; \theta)}_{\text{technically a joint pdf, originally}}$ (one parameter, θ), maybe need $\ln L(\theta)$, then maximize. (take deriv. w.r.t. θ, set $=0$, etc.)

- might need one or more indicator functions depending on whether the params appear in the support of the pdfs.
  - if so, then will be unable to maximize via derivatives. Look @ behavior of L ($\uparrow$'ing, $\downarrow$'ing, constant).

- In general, $E(\bar{X}) = E(X)$, and $Var(\bar{X}) = \frac{Var(X)}{n}$, assuming have indep. R.V.s.
  - Invariance property holds.

§ 9.3 – Criteria for Evaluating Estimators

- sps you have a parameter $\tau(\theta)$ (might be a fcn of the parameter); then an estimator T of $\tau(\theta)$ is unbiased for $\tau(\theta)$ if $E(T) = \tau(\theta)$. ($\forall \theta \in \Omega$)
- unbiasedness is good...

pdf of $X_{n:n}$: $f_{X_{n:n}}(x_{n:n}) = n\left[F_X(x_{n:n})\right]^{n-1} \cdot f_X(x_{n:n})$ (and support)

pdf of $X_{1:n}$: $f_{X_{1:n}}(x_{1:n}) = n\left[1 - F_X(x_{1:n})\right]^{n-1} \cdot f_X(x_{1:n})$ (and support)

- the distribution/shape of an unbiased estimator will be centered at $\tau(\theta)$, symmetric.
  - between two unbiased estimators of $\tau(\theta)$, the estimator with smaller variance is better (however, this can change as n changes).
- To properly compare an unbiased estimator with ~~another~~ a biased estimator (the latter of which may prove to be better!), we use mean square error (MSE)

T an estimator of $\tau(\theta)$ →
  - $MSE(T) = Var(T) + [b(T)]^2$,

    where $b(T)$ is the bias of T, with $b(T) = E(T) - \tau(\theta)$.
  - (Also, $MSE(T) = E([T - \tau(\theta)]^2)$, but above is the computational formula
- Can get a (lower bd) min. on the variance of unbiased estimators... UMVUE stuff...

See Notes / Handout + Thm on back. (CRLB)

If an estimator $\hat{\theta}$ is UMVUE of θ, then $E(\hat{\theta}) = \theta$ and $Var(\hat{\theta}) = CRLB$.

§ 9.4 – Asymptotic Properties of the MLE

- MLEs possess nice properties for large sample sizes

Next Chapter: Ch. 9 - Point Estimation

- observe sample, make inferences about population (opposite of Ch. 7, Ch. 8, where we assumed we knew the pop. params, made inferences about a random sample of that pop.)

- Scenario: We will take a r. sample from a distribution with pdf $f(x;\theta)$, where the form of f is known but the value of θ is unknown.

parameter (describes population)

(capital omega)

  - Let Ω be the parameter space, that is, the set of all possible values that θ can take.
    - θ can be a vector, in which case we denote it as $\underset{\sim}{\theta}$.

Ex: 1) $f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x>0$, $\theta>0$.

Here, $\Omega = (0,\infty)$ (excluding zero cuz in denom)

2) $f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$.

Here, $\underset{\sim}{\theta} = (\mu,\sigma^2)$, ← multiple params, so have a vector,

so $\Omega = (-\infty,\infty)\times(0,\infty)$,

Goal: To find an estimator (a statistic) for θ.

an "estimate" is a particular values obtained using the "estimator", which is like a general form.

§9.2 - Methods of Estimation

I. Method of Moments

- produce the (particular) method of moments estimator (MMEs)

Idea: Equate the population moment with the sample moment.

- Recall: population moments are $E(X^r)$ (r = 1 is first moment, E(X), or the mean) (r = 2 is 2nd moment, $E(X^2)$)
  in terms of parameters

  sample moments are $\frac{\sum_{i=1}^{n} X_i^r}{n}$.

  r=1 gives you $\overline{X}$ $\left(= \frac{\sum_{i=1}^{n} X_i}{n}\right)$

  average of [redacted] rth population moments, up to n of them.

- The MME, $\hat{\theta}$, is the value such that

$$E(X^r) = \frac{\sum_{i=1}^{n} X_i^r}{n}.$$

- Note: Always start with r=1 and use as many moments as the number of parameters to be estimated.
  - need a moment for each parameter. For Ex 1) (on previous page), need one moment (cuz only had one parameter, θ.). For Ex 2) (on previous page), need two moments (cuz two params). (μ and σ^2)

Ex 1) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $f(x;\theta) = \theta x^{\theta - 1}$, $0 < x < 1$, $\theta > 0$. Derive the MME of θ.

- one parameter (θ), so need one moment. ~~[illegible]~~

  1st population moment: (r=1) $E(X)$,

$$E(X) = \int_0^1 x\theta x^{\theta-1}\,dx = \int_0^1 \theta x^{\theta}\,dx = \theta \frac{x^{\theta+1}}{\theta+1}\Big|_0^1 = \frac{\theta}{\theta+1}$$

(θ is param)

- Find $\hat{\theta}$ such that $\frac{\hat{\theta}}{\hat{\theta}+1} = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$ ← 1st sample moment. Equate to 1st pop. moment.

1st sample moment is always the sample mean.

$$\Rightarrow \hat{\theta} = \overline{X}\hat{\theta} + \overline{X}$$

$$\Rightarrow \hat{\theta}(1-\overline{X}) = \overline{X} \Rightarrow \boxed{\hat{\theta} = \frac{\overline{X}}{1-\overline{X}}}$$

always > 0, always > 0 ✓

estimator — is in terms of statistics ($\overline{X}$) ✓

statistic

Suppose we take a r.s. of size 5 and observe

lower case → $x_1 = .2$

cuz → $x_2 = .8$

are specific values for the R.V. $x_3 = .6$

$x_4 = .1$

$x_5 = .3$

$$\frac{x_1+x_2+x_3+x_4+x_5}{5} = \frac{0.2+0.8+0.6+0.1+0.3}{5} = \frac{2}{5} = \bar{x}$$

MM estimate $\hat{\theta} = \dfrac{\bar{x}}{1-\bar{x}}$, where $\bar{x} = \frac{2}{5}$

a value of our statistic (as opposed to MM estimator)

$$= \frac{\left(\frac{2}{5}\right)}{\frac{3}{5}} = \boxed{\frac{2}{3}}$$

Ex 2) $X_1, X_2, \dots, X_n$ a r.s. from $N(\mu, \sigma^2)$.

Derive MMEs of $\underset{\sim}{\theta} = (\mu, \sigma^2)$.

· pop moments: $E(X) = \mu$

r=1 →

r=2 → $E(X^2) = Var(X) + [E(X)]^2 = \sigma^2 + \mu^2$

express in terms of the params

· sample moments: 1st: $\bar{X}$

will not change

2nd: $\dfrac{\sum_{i=1}^n X_i^2}{n}$

· Find $\hat{\mu}$ and $\hat{\sigma}^2$ such that $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 + \hat{\mu}^2 = \sum_{i=1}^n \frac{X_i^2}{n}$:

$\boxed{\hat{\mu} = \bar{X}}$ ← already done. Can plug in,

$$\hat{\sigma}^2 + \hat{\mu}^2 = \frac{\sum_{i=1}^n X_i^2}{n}$$

want an n-1 in denom to look like S^2

memorize proof

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2}{n} - (\bar{X})^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n} = \boxed{\frac{n-1}{n} S^2}$$

· Note, if we want the MME of a function of θ, say $\tau(\theta)$, then use

(capital tau, with a hat) $\hat{\tau}(\theta) = \hat{\tau}(\hat{\theta})$. So the MME of σ is $\hat{\sigma} = \sqrt{\frac{n-1}{n} S^2} = \sqrt{\frac{n-1}{n}}\, S$

the estimator of the function is the function evaluated @ the estimator

Ex 3) Let $X_1, X_2, \ldots, X_n$ be a r.s. from

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!}, \quad x = 0,1,2,\ldots$$

· Want: The MME of $\tau(\theta) = P(X=0) = e^{-\theta}$.

· Note: $X \sim POI(\theta)$ ← helpful to recognize special dist cuz then can get 1st pop moment immediately from cheat sheet

· Need: $\hat{\theta}$

(mean for $POI(\theta)$)

pop moment: $E(X) = \theta$ ← in terms of param

sample moment: $\bar{X}$ ← equate pop moment to sample moment

So the MME of θ is $\hat{\theta} = \bar{X}$.

put a hat on the pop moment if its simple

Thus,

$$\hat{\tau}(\theta) = e^{-\hat{\theta}} = e^{-\bar{X}}.$$

Ex 4) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} EV(\theta, \eta)$. ← (extreme value distribution) ← see cheat sheet (or table in beginning of the textbook)

· Find MMEs of θ and η. (I.e., find $\underset{\sim}{\theta} = (\theta, \eta)$.)

1st pop moment: $E(X) = \eta - \gamma\theta$, where γ = Euler's constant
$r=1$

2nd pop moment: $E(X^2) = Var(X) + [E(X)]^2$
$r=2$

$$= \frac{\pi^2\theta^2}{6} + (\eta - \gamma\theta)^2$$

The more samples, the more info you have about your population. ~~Depends~~ Usefulness depends on how the sample was taken and the sample size.

· Find $\hat{\theta}$ and $\hat{\eta}$ such that

1st pop moment → $\hat{\eta} - \gamma\hat{\theta} = \bar{X}$ ← 1st sample moment ← ①

and 2nd pop moment → $\frac{\pi^2\hat{\theta}^2}{6} + (\hat{\eta} - \gamma\hat{\theta})^2 = \frac{\sum_{i=1}^n X_i^2}{n}$ ← 2nd sample moment ← ②

Use ① in ② to get

$$\frac{\pi^2\hat{\theta}^2}{6} + \bar{X}^2 = \frac{\sum_{i=1}^n X_i^2}{n} \Rightarrow \frac{\pi^2\hat{\theta}^2}{6} = \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 = \frac{n-1}{n}S^2$$

1 param left still, solve for it

$$\Rightarrow \hat{\theta}^2 = \frac{6}{\pi^2} \cdot \frac{n-1}{n} S^2$$

want $\hat{\theta}$ not $\hat{\theta}^2$

$$\Rightarrow \boxed{\hat{\theta} = \frac{\sqrt{6}}{\pi}\sqrt{\frac{n-1}{n}} \cdot S}$$

cont'd →

Go back to (1):

$$\hat{\eta} = \bar{X} + \gamma \hat{\theta}$$

$$\Rightarrow \boxed{\hat{\eta} = \bar{X} + \gamma \frac{\sqrt{6}}{\pi} \sqrt{\frac{n-1}{n}} \cdot S}$$

II. Method of Maximum Likelihood

- produces the maximum likelihood estimators (MLEs), which possess desirable large-sample properties.

Def: The likelihood function, $L(\theta)$, is the joint pdf of $X_1, X_2, \dots, X_n$ evaluated at $x_1, x_2, \dots, x_n$. I.e.,

$$L(\theta) = f(x_1, x_2, \dots, x_n ; \theta)$$

the only unknown, so denoted as just a fcn of θ.

specific values (known), so the only unknown value is θ

If $X_1, X_2, \dots, X_n$ form a r.s. from $f(x;\theta)$, then

$$L(\theta) = f(x_1, x_2, \dots, x_n ; \theta)$$

$$= f(x_1;\theta) \cdot f(x_2;\theta) \cdots f(x_n;\theta) \quad \text{cuz independent}$$

$$= \prod_{i=1}^{n} f(x_i ; \theta)$$

gives us the chance that we will observe the data

· Goal: find $\hat{\theta}$ that maximizes $L(\theta)$.

(take derivatives, set equal to 0, etc.)

- alternatively, we can take the log-likelihood function $\ln(L(\theta))$ and maximize it.

· helpful depending on the form of the pdf we obtain.

# of derivatives depends on # of params we want to estimate.

Sometimes easier to take log version first.

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $GEO(p)$.

Derive the MLE of p.

Recall: $f(x;p) = pq^{x-1}, \quad x = 1, 2, \ldots$

need write in terms of p $\rightarrow$ $= p(1-p)^{x-1}$

$\ln(a \cdot b) = \ln(a) + \ln(b)$

$\ln L(p) = \ln\left(p^n \cdot (1-p)^{\sum_{i=1}^n x_i - n}\right)$

$= \ln(p^n) + \ln\left((1-p)^{\sum_{i=1}^n x_i - n}\right)$

$= n \ln(p) + \left(\sum_{i=1}^n x_i - n\right) \ln(1-p)$

So $L(p) = \prod_{i=1}^n p(1-p)^{x_i - 1}$

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

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

like a constant cuz do $\frac{d}{dp}$ w.r.t. p

painless way is to use log-likelihood version

$\Rightarrow \ln L(p) = n \ln(p) + \left(\sum_{i=1}^n x_i - n\right) \ln(1-p)$ $\longleftarrow$ want to maximize this, so take derivative now

So $\frac{d}{dp}\left[\ln L(p)\right] = \frac{n}{p} + \frac{\left(\sum_{i=1}^n x_i - n\right)}{1-p} \cdot (-1)$ $\leftarrow$ chain rule

now set $=$ to 0 (at this point, start using hats) (find $\hat{p}$ so that deriv set $=$ to 0 gives maximum)

$$\Rightarrow \frac{n}{\hat{p}} - \frac{\sum_{i=1}^n x_i - n}{1 - \hat{p}} = 0$$

$\vdots$

$$\Rightarrow \hat{p} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}} = \text{maximum-likelihood estimate of } p.$$

little $\bar{x}$ $\leftarrow$ specific value

Notice that in this particular example, MME = MLE.

$\frac{1}{p} = \bar{X} \Rightarrow \hat{p} = \frac{1}{\bar{x}}$

1st pop mom. $\nearrow$ 1st samp mom.

~~Ask~~ example w/ more than 1 param involved? later notes cont'd

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $f(x;\theta) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$

derive the MLE of θ.

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$= \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n \cdot \prod_{i=1}^{n} x_i^{\theta-1}$$

can't do sum of exponents b/c the bases are all different! (Clue to use ln likelihood)

so $\ln L(\theta) = n\ln(\theta) + \sum_{i=1}^{n}\left[(\theta-1)\ln(x_i)\right]$

$$= n\ln(\theta) + (\theta-1)\sum_{i=1}^{n}\ln(x_i)$$

Now $\frac{d}{d\theta}\ln L(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n}\ln(x_i)$; set = to 0:

$$\Rightarrow \frac{n}{\hat{\theta}} + \sum_{i=1}^{n}\ln(x_i) = 0 \Rightarrow \frac{n}{\hat{\theta}} = -\sum_{i=1}^{n}\ln x_i$$

$$\Rightarrow \hat{\theta} = \frac{-n}{\sum_{i=1}^{n}\ln(x_i)}$$

negative makes sense b/c $0<x<1$ so the $\ln(x_i)$ will be negative. $\theta > 0$ given, so the $-n$ is necessary.

MME: $\hat{\theta} = \frac{\bar{X}}{1-\bar{X}}$

for comparison

MLE

(cannot write in terms of $\bar{X}$)

MLestimate

(don't need convert to MLestimator unless explicitly asked)

notes cont'd ⟶

[Invariance Property, then another MLE example but where we need the Indicator Function]

Ask more than one param MLE?

Thm: Invariance Property

If $\hat{\theta}$ is the MLE of θ and if $u(\theta)$ is some function of θ, then the MLE of $u(\theta)$ is $u(\hat{\theta})$.

Example: Let $X_1, X_2, \ldots, X_n$ be a r. sample from GEO(p). Want the MLE of $P(X_1 > k)$.

- Recall: $P(X_1 > k) = q^k$ ← property of GEO distribution
$= (1-p)^k = u(p)$.

- We found that the MLE of p is $\hat{p} = \frac{1}{\bar{X}}$, ~~ask~~ general result...? Yes, for a GEO(p) for param. p.

so MLE of $u(p) = (1-p)^k$ is $\hat{u}(p) = u(\hat{p}) = \widehat{P(X_1 > k)}$
$= (1-\hat{p})^k$
$= \left(1 - \frac{1}{\bar{X}}\right)^k$ (estimator)

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from UNIF$(0, \theta)$. Derive MLE of θ.

Recall: $f(x_i; \theta) = \frac{1}{\theta}, \; 0 < x_i < \theta$ ← parameter is part of the support ☆

- So, use the indicator function: ☆

$$I\{A\} = \begin{cases} 1 & \text{if event } A \text{ occurred} \\ 0 & \text{o.w.} \end{cases}$$

e.g., if x_i negative, then get 0 (takes care of the "o.w." scenario)

- So $f(x_i; \theta) = \frac{1}{\theta} I\{0 < x_i < \theta\}$. Now,

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I\{0 < x_i < \theta\}$$

for this product to be nonzero as a whole, it must be true that none of these are zero.

$$= \frac{1}{\theta^n} \underbrace{\prod_{i=1}^{n} I\{0 < x_i < \theta\}}$$

$= 1$ only if every x_i is $0 < x_i < \theta$. Can replace with the ind. fcn of: $I\{0 < x_{1:n} < x_{n:n} < \theta\}$ (= 1 so long as the min is smaller than the max and the min is > 0, max $< \theta$) (if allowed the max to be $> \theta$, then would allow for zero case)

so $L(\theta) = \frac{1}{\theta^n} I\{0 < x_{1:n} < x_{n:n} < \theta\}$

$= \frac{1}{\theta^n}(1)$ provided this is true

Now, could try maximize like normal:

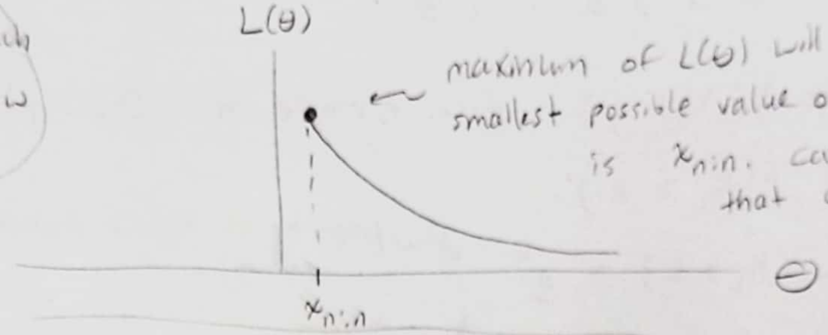
$\Rightarrow \frac{d}{d\theta} L(\theta) = -n\,\theta^{-n-1}$

$= \cancel{\frac{-n}{\theta^{n+1}} \overset{set}{=} 0}$

But this method doesn't work cuz can never $= 0$ (can't solve for θ)

~~Cont~~

Use diff. approach to figure out how to maximize:



maximum of $L(\theta)$ will be the smallest possible value of θ, which is $x_{n:n}$. cuz we established that $0 < x_{1:n} < x_{n:n} < \theta$ must be true.

For $L(\theta) = \frac{1}{\theta^n}$, as $\theta \to \infty$, $L(\theta) = \frac{1}{\theta^n} \to 0$ (decreasing) so graph looks like this

$\boxed{\text{So the MLE of } \theta \text{ is } \hat{\theta} = x_{n:n}}$

[start of notes from 2/23/22] handout, page 3, page 4, ~~page 1~~, page 2, ~~page 3, then handout, then~~

[handout] normal derivation w/ MLE

STAT 480B

Examples for MLE

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x;\theta) = 1,\ \theta - 1/2 < x < \theta + 1/2$. Derive the MLE of θ.

2. Suppose a random sample of size 10 from the distribution given above resulted in the following values:

1.23 1.14 1.35 1.02 1.15 0.61 0.89 0.66 0.90 1.48

Determine an MLE of θ. $+0.5 = 1.11$ $-0.5 = 0.98$

$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} I\{\theta - \frac{1}{2} < x_i < \theta + \frac{1}{2}\}$

(Indicator fn = 0 or 1. So max is when = 1)

$\Rightarrow L(\theta) = I\{\theta - \frac{1}{2} < x_{1:n} < x_{n:n} < \theta + \frac{1}{2}\}$. Now maximize this.

- $L(\theta) = 1$ when $\hat{\theta} - \frac{1}{2} < x_{1:n}$ and $x_{n:n} < \hat{\theta} + \frac{1}{2}$. [Solve for $\hat{\theta}$.]

~~ask~~ that here? no

$\Rightarrow \hat{\theta} < x_{1:n} + \frac{1}{2}$, $\Rightarrow x_{n:n} - \frac{1}{2} < \hat{\theta}$

$\Rightarrow \boxed{x_{n:n} - \frac{1}{2} < \hat{\theta} < x_{1:n} + \frac{1}{2}}$ ← (can stop here, but can go stricter)

[any value of $\hat{\theta}$ in these bds will maximize $L(\theta)$, i.e., will be an MLE of θ]

$\Rightarrow$ MLE is not unique, so choose a convenient value that you know will be in these bds.

Think of this like $a < \hat{\theta} < b$. Then two convenient values of $\hat{\theta}$ between a and b would be the midpt ($\hat{\theta} = \frac{b-a}{2}$) and the average ($\hat{\theta} = \frac{a+b}{2}$).

So one convenient value of $\hat{\theta}$ is the average: $\hat{\theta} = \frac{x_{n:n} - \frac{1}{2} + x_{1:n} + \frac{1}{2}}{2} = \boxed{\frac{x_{n:n} + x_{1:n}}{2}}$

2) Want MLestimate, cuz specific values given.

- Well, any $\hat{\theta}$ in the interval $x_{n:n} - \frac{1}{2} < \hat{\theta} < x_{1:n} + \frac{1}{2}$ will be an MLestimate.

Out of the given values, $x_{n:n} = 1.48$. so $\hat{\theta} > 1.48 - \frac{1}{2} = 0.98$.

Also, $x_{1:n} = 0.61$, so $\hat{\theta} < 0.61 + \frac{1}{2} = 1.11$. So any $\hat{\theta}$ in the interval $0.98 < \hat{\theta} < 1.11$

- A convenient value of such an MLestimate would be $\frac{x_{n:n} + x_{1:n}}{2} = \frac{1.48 + 0.61}{2} = \boxed{1.045}$

§ 9.2 Notes (MLE stuff) cont'd: (still from Wed. 2/23/22 lecture)

· example with more than one parameter needing to be estimated: (where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$):

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $N(\mu, \sigma^2)$.
Derive the MLE of $\underset{\sim}{\theta} = (\mu, \sigma^2)$.

$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$

$$\cdot L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2)$$

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

our param is σ^2, not σ. Should get everything strictly in terms of how the params appear first. So rewrite the $\frac{1}{\sigma}$ as $\frac{1}{\sqrt{\sigma^2}}$, etc.

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\cdot\sqrt{\sigma^2}} \cdot e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}} = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2}(x_i-\mu)^2/\sigma^2}$$

$$= \left[(2\pi\sigma^2)^{-1/2}\right]^n \cdot e^{\sum_{i=1}^{n}\left[-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}\right]}$$

$$= (2\pi\sigma^2)^{-n/2} \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\Rightarrow \ln L(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \cancel{\ln(e)}^{(1)}$$

· Now, create a system of eqns by taking the partial deriv. of this w.r.t. each param.

$$\frac{d}{d\mu}\ln L(\mu, \sigma^2) = \frac{d}{d\mu}\left(\cancel{-\frac{n}{2}\ln(2\pi\sigma^2)}^{0}\right) - \frac{d}{d\mu}\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= -\frac{1}{2\sigma^2}\frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i-\mu)^2\right) = \frac{-1}{2\sigma^2}\sum_{i=1}^{n}\frac{d}{d\mu}(x_i-\mu)^2$$

$$= \frac{-1}{2\sigma^2}\cdot\sum_{i=1}^{n} 2(x_i-\mu)^{(1)}(-1) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu).$$

$$\frac{d}{d\sigma^2}\ln L(\mu, \sigma^2) = \frac{d}{d\sigma^2}\left(-\frac{n}{2}\ln(2\pi\sigma^2)\right) - \frac{1}{2}\frac{d}{d\sigma^2}\left((\sigma^2)^{-1}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

do w.r.t. σ^2, not σ.

$$= -\frac{n}{2}\cdot\frac{1}{2\pi\sigma^2}(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\left[-(\sigma^2)^{-2}\right]$$

$$= \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

now set each of these equal to zero and solve the system of eqns.

So we have $\frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\hat{\mu}) = 0$ and $\frac{-n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = 0$.

$$\Rightarrow \sum_{i=1}^{n}(x_i-\hat{\mu}) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i - n\hat{\mu} = 0$$

$$\Rightarrow \sum x_i = n\hat{\mu}$$

$$\Rightarrow \hat{\mu} = \frac{\sum x_i}{n} \Rightarrow \boxed{\hat{\mu} = \bar{x}}$$

$$\Rightarrow \frac{1}{2\hat{\sigma}^2}\left[-n + \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2\right] = 0$$

$$\Rightarrow -n + \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Rightarrow \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = n$$

$$\Rightarrow \boxed{\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n}} \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

$$\Rightarrow \boxed{\hat{\sigma}^2 = \frac{(n-1)}{n}S^2}$$

example where more than one parameter, and one of them requires indicator function.

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $f(x;\mu,\sigma) = \frac{1}{\sigma} e^{\frac{-(x-\mu)}{\sigma}}$,

where $x_i > \mu$ and $\sigma > 0$. Derive the MLE of $\underset{\sim}{\theta} = (\mu, \sigma)$.

$$L(\mu,\sigma) = \prod_{i=1}^{n} f(x_i;\mu,\sigma) \cdot I\{x_i > \mu\}$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma} e^{\frac{-(x_i-\mu)}{\sigma}} \cdot I\{x_i > \mu\}$$

$$= \frac{1}{\sigma^n} e^{\sum_{i=1}^{n}\left(-\frac{(x_i-\mu)}{\sigma}\right)} \cdot \prod_{i=1}^{n} I\{x_i > \mu\}$$

all $x_i > \mu$, so $x_{1:n} > \mu$

$$= \left(\frac{1}{\sigma^n} \cdot e^{\sum_{i=1}^{n}\left(\frac{-(x_i-\mu)}{\sigma}\right)}\right) \cdot I\{x_{1:n} > \mu\}$$

ask

ask what do if both params rely on x_i? multiply by another $I\{A\}$?

ask | Yes, or try putting as one $I\{A\}$.

can't have the params dep. on one another cuz r.s., right?

Yes, but that's something different.

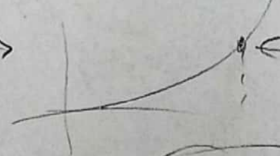
- To get MLE of μ, analyze the behavior of $L(\mu,\sigma)$ as a function of μ. Because μ, the param, is in the indicator fn. Can't take deriv. right away)

$$e^{\sum_{i=1}^{n}\left(\frac{-(x_i-\mu)}{\sigma}\right)} = e^{-\frac{1}{\sigma}\sum_{i=1}^{n}(x_i-\mu)} = e^{-\frac{1}{\sigma}\left(\sum_{i=1}^{n} x_i - n\mu\right)} = e^{\frac{-\sum_{i=1}^{n} x_i}{\sigma}} \cdot e^{\frac{n\mu}{\sigma}},$$

$$\text{so } L(\mu,\sigma) = \underbrace{\frac{1}{\sigma^n} \cdot e^{\frac{-\sum_{i=1}^{n} x_i}{\sigma}}}_{\text{like a positive constant}} \cdot e^{\frac{n\mu}{\sigma}} \cdot I\{x_{1:n} \geq \mu\}$$

enforces...

increasing fcn



stops @ $x_{1:n}$ cuz $x_{1:n} > \mu$

pdf cont., so L cont., so equality irrelevant.

$$\Rightarrow \hat{\mu} = x_{1:n}$$

try writing $I\{A\}$ as one if possible. Otherwise, do product of two of them.

now go back and do for σ; don't need indicator fcn. Use deriv., etc.

finish notes

§9.3 – Criteria for Evaluating Estimators

out of diff. estimators, decide which is the better one to use. — depends on criteria we are interested in. ask e.g., unbiasedness. :)

- book: "Several properties of estimators would appear to be desirable, including unbiasedness."

· **Def:** An estimator T is unbiased for $\tau(\theta)$ if $E(T) = \tau(\theta)$ for all $\theta \in \Omega$. (O.w., T is unbiased for $\tau(\theta)$.)

- book: "If an unbiased estimator is used to assign a value of $\tau(\theta)$, then the correct value of $\tau(\theta)$ may not be achieved by any given estimate, t, but the "average" value of T will be $\tau(\theta)$."

· Ex) Let $X_1, X_2, \dots, X_n$ be a r. sample from $N(\mu, \sigma^2)$. Are the MLEs of μ and σ^2 unbiased?

· Recall: → $MLE(\mu) = \hat{\mu} = \bar{X}$. (by a previous result for a r.s from $N(\mu,\sigma^2)$) (see §9.2)

· $E(\bar{X}) = \mu$ ✓

→ $\underset{\text{estimate}}{MLE(\sigma^2)} = \hat{\sigma}^2 = \dfrac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n}$.

ask What is our $\tau(\theta)$?

· We also found that $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \dfrac{(n-1)S^2}{n}$, which shows us that $\sum_{i=1}^n (x_i - \hat{\mu})^2 = (n-1)S^2$.

MLestimator: $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$

· Realize that $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. Therefore, [NTS $E(\hat{\sigma}^2) = \sigma^2$]

so $\bar{X}$ is unbiased for μ, but $\hat{\sigma}^2$ is not unbiased for σ^2. S^2 is unbiased for σ^2, tho.

$$E(\hat{\sigma}^2) = E\left(\frac{\sum_{i=1}^n (X_i - \hat{\mu})^2}{n}\right) = E\left(\frac{(n-1)S^2}{n}\right) = E\left(\frac{\sigma^2}{\sigma^2}\frac{(n-1)S^2}{n}\right) = E\Big(\frac{\sigma^2}{n} \cdot \underbrace{\frac{(n-1)S^2}{\sigma^2}}_{=Y \sim \chi^2(n-1)}\Big)$$

$$= E\left(\frac{\sigma^2}{n} \cdot Y\right) = \frac{\sigma^2}{n}E(Y) = \frac{\sigma^2}{n}(n-1) = \frac{(n-1)\sigma^2}{n} \neq \sigma^2.$$

~~ask~~ when do hats go away dep. on derivation process. Need to go away somehow.

So $\hat{\sigma}^2$ is biased for σ^2. Though $\frac{n-1}{n}$ is just a constant, multiply by reciprocal...

But, $\frac{n}{n-1}\hat{\sigma}^2$ is unbiased for σ^2.

$$\Rightarrow \frac{n}{n-1}\hat{\sigma}^2 = \frac{\cancel{n}}{n-1} \cdot \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\cancel{n}} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = S^2, \text{ so } \boxed{S^2 \text{ is unbiased for } \sigma^2.}$$

Ex) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $f(x;\theta) = \frac{1}{\theta}$, $0 < x < \theta$.

I.e., $X \sim UNIF(0, \theta)$.

support dep. on param, so

Recall that we found that the MLE for θ is $\hat{\theta} = X_{n:n}$.

Q: Is $\hat{\theta} = X_{n:n}$ unbiased for θ?

Want: $E(X_{n:n})$. First need pdf of $X_{n:n}$.

ask?

$$f(x_{n:n}) = n\left[\underbrace{F(x_{n:n})}_{=P(X \le x_{n:n})}\right]^{n-1} f(x_{n:n})$$ (recall this from 480A)

$$= n\left[\frac{x_{n:n}}{\theta}\right]^{n-1}\frac{1}{\theta}$$

$$= \frac{n(x_{n:n})^{n-1}}{\theta^n}, \quad 0 < x_{n:n} < \theta$$

x_1	x_2	x_3	x_4	x_5	x_6	$\cdots$	x_n
1	2	3	4	5	6		n

have n of these, but only 1 can be the max. $\binom{n}{1}$

the $n-1$ remaining must be smaller than

So,

$$E(X_{n:n}) = \int_0^\theta x_{n:n} f(x_{n:n})\,dx_{n:n}$$

$$= \int_0^\theta x_{n:n} \cdot \frac{n(x_{n:n})^{n-1}}{\theta^n}\,dx_{n:n}$$

?

$$= \frac{n}{\theta^n}\int_0^\theta (x_{n:n})^n\,dx_{n:n} = \frac{n}{\theta^n}\left[\frac{(x_{n:n})^{n+1}}{n+1}\right]_0^\theta$$

$$= \frac{n}{\theta^n(n+1)}\left[\theta^{n+1} - 0\right],$$ so $E(X_{n:n}) = \frac{n}{n+1}\theta$,

So $\hat{\theta} = X_{n:n}$ is biased for θ

(But, $\frac{n+1}{n}X_{n:n}$ is unbiased for θ.) $\tau(\theta)$

unbiasedness a good quality for an estimator we not over- or under-estimating true value.

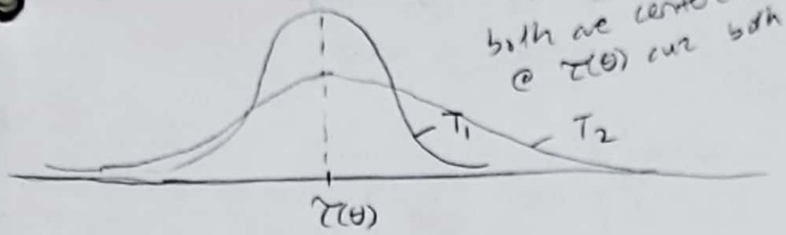
- Sps we use the MME instead:

$E(X) = \frac{\theta}{2} = \bar{X} \Rightarrow \hat{\theta} = 2\bar{X} = MME$. Is the MME unbiased for θ?

$E(2\bar{X}) = 2E(\bar{X}) = 2 \cdot \frac{\theta}{2} = \theta$. So yes, $2\bar{X}$ is unbiased for θ.

- Found two unbiased estimators. $\frac{n+1}{n}X_{n:n}$ and $2\bar{X}$. Which of them is better?

· Consider 2 estimators T_1 and T_2 whose distributions are shown below



both are centered @ $\tau(\theta)$ cuz both are unbiased

Each of T_1 and T_2 could work, but T_2 is way more spread out relative to the mean. This describes the variance.

So now we have to compare their variances.

T_1 is better than T_2 since $Var(T_1) < Var(T_2)$

· Let's go back to the UNIF$(0, \theta)$ case.

$$Var(\underbrace{2\bar{X}}_{T_1}) = 4Var(\bar{X}) = 4 \cdot \frac{Var(X)}{n} = \frac{4}{n} \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

· Now let's get $Var\left(\frac{n+1}{n} X_{n:n}\right)$.

$$E(X_{n:n}) = \frac{n}{n+1}\theta$$

$$Var\left(\frac{n+1}{n} X_{n:n}\right) = \left(\frac{n+1}{n}\right)^2 Var(X_{n:n})$$

$$= \left(\frac{n+1}{n}\right)^2 \left[E(X_{n:n}^2) - [E(X_{n:n})]^2\right],$$

where $E(X_{n:n}^2) = \int_0^\theta x_{n:n}^2 \cdot f(x_{n:n})\,dx_{n:n}$

$$= \int_0^\theta x_{n:n}^2 \cdot n \cdot \frac{x_{n:n}^{n-1}}{\theta^n}\,dx_{n:n} = \frac{n}{\theta^n}\int_0^\theta x_{n:n}^{n+1}\,dx_{n:n}$$

$$= \frac{n}{\theta^n}\left[\frac{x_{n:n}^{n+2}}{n+2}\right]_0^\theta = \frac{n}{\theta^n(n+2)}\left[\theta^{n+2} - 0\right]$$

$$= \frac{n}{n+2}\theta^2$$

thus $Var(\underbrace{\tfrac{n+1}{n} X_{n:n}}_{T_2}) = \frac{(n+1)^2}{n^2}\left[\frac{n\theta^2}{n+2} - \left(\frac{n}{n+1}\theta\right)^2\right]$

skip a couple steps...

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

$$= \frac{\theta^2}{n(n+2)} \cdot \left[\cancel{n^2} + \cancel{2n} + 1 - \cancel{n^2} - \cancel{2n}\right] = \frac{\theta^2}{n(n+2)}$$

The better estimator is the one w/ the lower variance.

Which is the better estimator?

· If $n=1$: $Var(T_1) = \frac{\theta^2}{3} = Var(T_2)$ (plug n=1 into each)

$n>1$: $3 < n+2$, so $\frac{1}{3} > \frac{1}{n+2}$, so $\frac{\theta^2}{3n} > \frac{\theta^2}{(n+2)n}$.

I.e., $Var(T_2) < Var(T_1)$

So $\hat{\theta} = \frac{n+1}{n} X_{n:n}$ is a better estimator than $2\bar{X}$.

[determine based on sample size desired]

· It is possible to find a lower bound for the variance of unbiased estimators.

- Want: uniformly minimum variance unbiased estimator (UMVUE)

Def: Let $X_1, X_2, \ldots, X_n$ be a r.s. from $f(x;\theta)$. An estimator T^* of $\tau(\theta)$ is UMVUE for $\tau(\theta)$ if

① T^* is unbiased for $\tau(\theta)$. I.e., $E(T^*) = \tau(\theta)$.

② $Var(T^*) \leq Var(T)$ where T is any other unbiased estimator of $\tau(\theta)$ for all $\theta \in \Omega$.

$Var(T^*) \leq Var(T)$, unbiased $\to Var(T) \geq CRLB$, so if $Var(T^*_{candidate}) = CRLB$, then candidate is T^*.

· [See handout]

· Ex) Let $X_1, \ldots, X_n$ be a r.s. from $f(x;\mu) = \frac{\mu^x e^{-\mu}}{x!}$, $x = 0, 1, 2, \ldots$

I.e., $X \sim POI(\mu)$. Find a UMVUE for μ.

θ in τ: $\to \theta = \mu$, $\tau(\theta) = \tau(\mu) = \mu$

· First need to find an estimator using one of our methods. Sps we choose to use MME. Well, $\hat{\mu} = \bar{X}$.

· Second, see if $\hat{\mu}$ is unbiased. Well, $E(\bar{X}) = \mu$. ✓

so $\bar{X}$ is a candidate for T^* by 1st pt.

· So $\bar{X}$ is unbiased for μ. · Third, get $Var(\bar{X})$. Well, $Var(\bar{X}) = \frac{Var(X)}{n} = \frac{\mu}{n}$.

· Fourth, get the CRLB. Need τ'. Well, $\tau(\mu) = \mu$, so $\tau'(\mu) = \frac{d}{d\mu}(\mu) = 1$.

Now, $\ln f(X;\mu) = X\ln\mu - \mu - \ln(X!) \Rightarrow \frac{d}{d\mu}\ln f(X;\mu) = \frac{X}{\mu} - 1$ →

Only applicable when support does not depend on θ.
(Ie, not applicable to ones requiring an Indicator fcn)

Kruh-mare Raow

STAT 480b

Cramer-Rao Lower Bound (CRLB) for Unbiased Estimators

gives lower bd. for variances

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $f(x;\theta)$ and let T be an unbiased estimator of $\tau(\theta)$. Then the Cramer-Rao lower bound for unbiased estimators of $\tau(\theta)$ is given by

$$\mathbf{Var}(T) \geq \text{CRLB} = \frac{[\tau'(\theta)]^2}{n\,\mathbf{E}\left(\left[\frac{\partial}{\partial\theta}\, ln\ f(X;\theta)\right]^2\right)} = \frac{[\tau'(\theta)]^2}{\underbrace{-n\,\mathbf{E}\left[\frac{\partial^2}{\partial\theta^2}\, ln\ f(X;\theta)\right]}}.$$

we used this one cuz recognized the variance....

else if easier to get the expected value, no square outside

Q8 repeat

NOTES:

Can do both tho.

(i) An unbiased estimator T is UMVUE for $\tau(\theta)$ if $\mathbf{Var}(T) = \text{CRLB}$.

(ii) If $\mathbf{Var}(T) > \text{CRLB}$, it does not necessarily mean that T is not a UMVUE for $\tau(\theta)$.

(not an iff statement!)
(may not retain the lower bd)

Derivation on Pg 305, 306.

She will tell us which method to use to make things easier.

~~Q8~~ Thm on back

★ asfu

Theorem: If an unbiased estimator of $\tau(\theta)$ exists, the variance of which achieves the CRLB, then only a **linear** function of $\tau(\theta)$ will admit an unbiased estimator, the variance of which achieves the corresponding CRLB. (new CRLB, but guaranteed to find Var(new estimator) = new CRLB

Ex) Let $X_1, X_2, \ldots, X_n \sim POI(\mu)$

Recall: $\bar{X}$ is UMVUE of μ. ($\tau(\theta)$) — linear function of the one we orig. estimated is this

$\Rightarrow E(\bar{X}) = \mu$ and $Var(\bar{X}) = CRLB$.

- Sps we want to estimate $\tau(\mu) = 1 + 2\mu$. (old $\tau(\theta)$)
  - What would be our estimator T?
    - Using Invariance property, $T = 1 + 2\bar{X}$.
  - Q: Is T unbiased?

$E(T) = E(1 + 2\bar{X}) = 1 + 2E(\bar{X})$

$= 1 + 2\mu$

Since $E(T) = \tau(\mu)$, T is unbiased for $\tau(\mu)$.

$\therefore$ only linear fcns of $\tau(\theta)$, like $\tau(\mu) = 1 + 2\mu$ ($\tau(\theta)$), should give an unbiased estimator... not guaranteed. Must check still.

They WILL.

And we can expect, w/o checking, that its CRLB will = its variance.

NOT the CRLB from before. (different one)

- Now, let's consider $Var(T)$.

$Var(T) = Var(1 + 2\bar{X}) = Var(2\bar{X}) = 4\,Var(\bar{X}) = 4\,\frac{Var(X)}{n} = 4\frac{\mu}{n}$

- $CRLB = \dfrac{[\tau'(\mu)]^2}{nE\left(\left[\frac{d}{d\mu}\ln f(x;\mu)\right]^2\right)}$ $\quad \tau'(\mu) = 2$

$= \frac{n}{\mu}$ from prev. example

$= \dfrac{(2)^2}{\frac{n}{\mu}} = \dfrac{4\mu}{n}$

So $Var(T) = CRLB$. ✓

$\therefore T = 1 + 2\bar{X}$ is UMVUE for $\tau(\mu) = 1 + 2\mu$.

If τ is linear, then can expect to be able to obtain an unbiased estimator of τ and that its variance = its CRLB.

If τ nonlinear, can make no conclusions from theorem. That is, can still be unbiased and can still be UMVUE, but its variance will NOT equal its CRLB. (b/c not linear)

$$E\left(\frac{d}{d\mu}\ln f(X;\mu)\right)^2 = E\left(\frac{X}{\mu}-1\right)^2 = E\left(\frac{X-\mu}{\mu}\right)^2 = \frac{1}{\mu^2}\underbrace{E(X-\mu)^2}_{=Var(X)=\mu} = \frac{\mu}{\mu^2} = \frac{1}{\mu}$$

· Putting it all together,

$$CRLB = \frac{(1)^2}{n\left(\frac{1}{\mu}\right)} = \frac{\mu}{n}.$$ Now compare w/ $Var(\bar{X})$ from 3rd step.

Well, $Var(\bar{X}) = \frac{\mu}{n} = CRLB$. Var of estimator = Var of lower bd., so Var of any other estimator will also ~~be~~

$\boxed{\therefore \bar{X} \text{ is UMVUE for } \mu.}$

~~find an unbiased estimator,~~

cont. of Pg 2, Handout

· Now, sps we want to estimate $P(X=0) = e^{-\mu}$,

- NOT a linear fn of μ, so $e^{-\mu}$ will not have an unbiased estimator whose variance is equal to the appropriate CRLB.

- cannot conclude that it ~~doesn't~~ isn't UMVUE, tho.

Ex #23, Pg 331) Let $X_1, \dots, X_n$ be a r.s. $\sim N(\overset{E(X_i)}{0}, \overset{Var(X_i)}{\theta})$.

(a) Is the MLE $\hat{\theta}$ an unbiased estimator of θ?

- from #4d, MLE of θ is $\hat{\theta} = \dfrac{\sum_{i=1}^{n} X_i^2}{n}$.

First, $E(\hat{\theta}) = E\left(\sum_{i=1}^{n} X_i^2 / n\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i^2\right) \overset{\text{linear operator}}{=} \frac{1}{n} \sum_{i=1}^{n} E(X_i^2)$

$= \frac{1}{n} \sum_{i=1}^{n} \left(Var(X_i) + (E(X_i))^2\right)$

$= \frac{1}{n} \sum_{i=1}^{n} (\theta + \cancel{0^2}) = \frac{1}{n} \sum_{i=1}^{n} \theta = \frac{1}{\cancel{n}} \cancel{n}\theta = \theta$,

So $E(\hat{\theta}) = \theta$, and $\hat{\theta}$ is unbiased for θ.

(b) Is $\hat{\theta}$ UMVUE for θ?

Check: is $Var(\hat{\theta}) = CRLB$?

The CRLB is: $\dfrac{[\tau'(\theta)]^2}{n E\left(\left[\frac{d}{d\theta} \ln f(X; 0, \theta)\right]^2\right)} = \dfrac{[\tau'(\theta)]^2}{-n E\left[\frac{d^2}{d\theta^2} \ln f(X; 0, \theta)\right]}$

- $\tau(\theta) = \theta$ here, so $\tau'(\theta) = 1$
- $\ln f(x; 0, \theta) = \ln\left(\frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2}\frac{x^2}{\theta}}\right) = -\frac{1}{2}\ln(2\pi\theta) - \frac{1}{2}\frac{x^2}{\theta}$

$\Rightarrow \frac{d}{d\theta} \ln f(x, 0, \theta) = -\frac{1}{2} \cdot \frac{1}{\cancel{2\pi}\theta}(\cancel{2\pi}) + \frac{1}{2}\frac{x^2}{\theta^2} = -\frac{1}{2\theta} + \frac{1}{2}\frac{x^2}{\theta^2}$

- $E\left(\left[-\frac{1}{2\theta} + \frac{1}{2}\frac{x^2}{\theta^2}\right]\right)^2$ will be a lot of work. Would need $E(X^4)$, we don't know this for a normal.
- So do 2nd deriv. w.r.t. θ: $\frac{d^2}{d\theta^2} \ln f(x; 0, \theta) = \frac{d}{d\theta}\left(-\frac{1}{2\theta} + \frac{1}{2}\frac{x^2}{\theta^2}\right)$

$= \frac{1}{2\theta^2} - \frac{x^2}{\theta^3}$ — looks more doable; won't have to square the whole thing. we can handle $E(X^2)$, not $E(X^4)$.

So,

- $E\left[\frac{d^2}{d\theta^2} \ln f(x; 0, \theta)\right] = E\left[\frac{1}{2\theta^2} - \frac{x^2}{\theta^3}\right] = \frac{1}{2\theta^2} - \frac{1}{\theta^3} \underbrace{E(X^2)}_{(=\theta)}$

$= \frac{1}{2\theta^2} - \frac{1}{\theta^3}\left(Var(X) + (E(X))^2\right) = \dots = \frac{1}{2\theta^2} - \frac{1}{\theta^2} = -\frac{1}{2\theta^2}$

Thus..., $\longrightarrow$

Thus,

$$CRLB = \frac{(1)^2}{-n\left(-\frac{1}{2\theta^2}\right)} = \frac{2\theta^2}{n}$$

- Now, need $Var(\hat{\theta})$.

X_i's indep., so X_i^2 indep., so sum can be broken up. Ie., can swap Var w/ Σ.

$$Var(\hat{\theta}) = Var\left(\frac{\sum_{i=1}^{n} X_i^2}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i^2\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i^2) \text{ due to indep.}$$

Note: $X_i \sim N(0,\theta)$, so $Z_i = \frac{X_i - (0)}{\sqrt{\theta}} = \frac{X_i}{\sqrt{\theta}} \sim N(0,1)$,

so let $U_i = Z_i^2 = \frac{X_i^2}{\theta} \sim \chi^2(1)$.

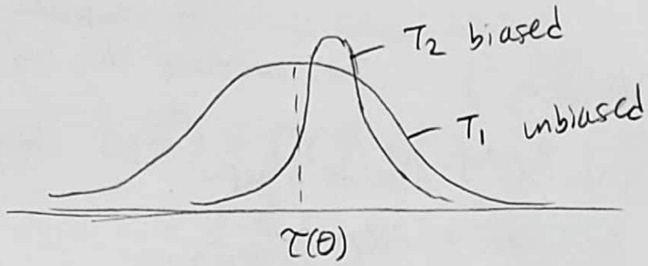
$\Rightarrow X_i^2 = U_i\theta$,

so $Var(\hat{\theta}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(\theta U_i) = \frac{1}{n^2}\sum_{i=1}^{n} \theta^2 \underbrace{Var(U_i)}_{=2(1)=2}$ $\chi^2(1)$

$$= \frac{1}{n^2}\sum_{i=1}^{n} 2\theta^2 = \frac{1}{n^2} 2n\theta^2 = \frac{2\theta^2}{n}.$$

So, since $Var(\hat{\theta}) = CRLB$, we can conclude $\hat{\theta} = \frac{\sum_{i=1}^{n} X_i^2}{n}$ is UMVUE for θ.

Now, let's consider estimators T_1 and T_2 for $\tau(\theta)$, whose distributions are below;



— can see T_1 is unbiased, but T_2 ~~is better~~ targets $\tau(\theta)$ better, (less variance). Lastly we compare their means to figure out whether to use T_1 or T_2:

To compare estimators, we can use the mean square error (MSE);

· Def; If T is an estimator of $\tau(\theta)$, then the bias of T is given by

$$b(T) = E(T) - \tau(\theta)$$

and the mean square error (MSE) is given by

$$MSE(T) = E(T - \tau(\theta))^2$$

becomes $E(T)$ when unbiased so we get Variance.

· Thm; If T is an estimator of $\tau(\theta)$,

computational formula → $$\boxed{MSE(T) = Var(T) + [b(T)]^2.}$$

want smaller MSE.

Pf: $MSE(T) = E([T - \tau(\theta)]^2)$ +- 0

$= E([\underbrace{T - E(T)}_{a} + \underbrace{E(T) - \tau(\theta)}_{b}]^2)$ ← $E([a+b]^2)$

$= E[(T - E(T))^2 + 2(T - E(T))(\underbrace{E(T) - \tau(\theta)}_{\text{a constant}}) + (\underbrace{E(T) - \tau(\theta)}_{\text{a constant}})^2]$

use linearity → $= E[(T - E(T))^2] + 2(E(T) - \tau(\theta))E[T - E(T)] + [E(T) - \tau(\theta)]^2$

$= Var(T)^2 + \cancel{2[E(T) - \tau(\theta)][E(T) - E(T)]} + \underbrace{[E(T) - \tau(\theta)]^2}_{= [b(T)]^2}$

$= \boxed{Var(T) + [b(T)]^2.}$ ∎

3/14/22 Notes over §9.3 cont'd

asl "uniformly"?

· Recall: T is an estimator of τ. Then

$$MSE(T) = Var(T) + [b(T)]^2,$$

where $b(T) = E(T) - \tau$.

It's not possible to find an estimator that has the min MSE of all estimators, cuz of the bias.

- Note: MSE is used for comparing estimators.

Ex) Let $X_1, X_2, \dots, X_n$ be a r.s. from $EXP(1, \eta)$. (scale: 1, shape: η)

(a) Derive the MME of η and call it $\hat{\eta}_1$.

(b) Derive the MLE of η and call it $\hat{\eta}_2$.

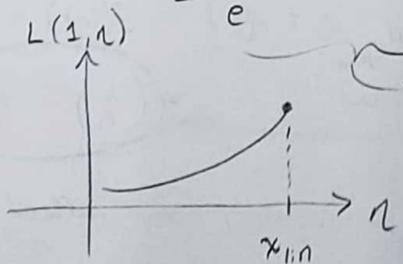
(c) Compare $MSE(\hat{\eta}_1)$ and $MSE(\hat{\eta}_2)$.

(a) $E(X) = \eta + 1$, so

$$\hat{\eta}_1 + 1 = \bar{X} \Rightarrow \hat{\eta}_1 = \bar{X} - 1.$$

(b) $L(1, \eta) = \prod_{i=1}^{n} e^{-(x_i - \eta)} \cdot I\{x_i > \eta\}$ cuz support is $x > \eta$ — every x_i must be $> \eta$.

$$= e^{-\sum_{i=1}^{n}(x_i - \eta)} \cdot I\{x_{1:n} > \eta\}$$



$\Rightarrow$ L is an ↑'ing fcn of η, so L is maximized at biggest possible value of η, which is $x_{1:n}$.

Thus, $\hat{\eta}_2 = x_{1:n}$.

asl. True always if indep.

$E(\bar{X}) = E(X)$? For r.s., always true.

(c) $b(\hat{\eta}_1) = E(\hat{\eta}_1) - \eta = E(\bar{X} - 1) - \eta = E(\bar{X}) - 1 - \eta = (\eta + 1) - 1 - \eta = 0$, (where $E(\bar{X}) = E(X)$)

so $\hat{\eta}_1$ is unbiased for η.

$Var(\hat{\eta}_1) = Var(\bar{X} - 1) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{(1)^2}{n} = \frac{1}{n}$,

so $MSE(\hat{\eta}_1) = \frac{1}{n} + (0)^2 = \frac{1}{n}$.

$= X_{1:n}$

$b(\hat{\eta}_2) = E(\hat{\eta}_2) - \eta$

$f_X(x) = e^{-(x-\eta)}, \quad x > \eta$

need: pdf of $Y = X_{1:n}$

$f_Y(y) = n[1 - F_X(y)]^{n-1} f_X(y)$,

where $F_X(y) = \int_\eta^y e^{-(x-\eta)}\,dx$ ← let $u = x-\eta$, then $du = dx$

$= \int_0^{y-\eta} e^{-u}\,du = -e^{-u}\Big|_0^{y-\eta} = 1 - e^{-(y-\eta)}$

Thus, $f_Y(y) = n\left[1 - (1 - e^{-(y-\eta)})\right]^{n-1} e^{-(y-\eta)}$

$= n\left[e^{-(y-\eta)}\right]^{n-1} e^{-(y-\eta)} = n \cdot e^{-(y-\eta)\cdot n}, \quad y > \eta$

$X_{1:n} > \eta$, $X_{1:n} = Y$

can make look nicer

use transformation method

Let $w = y - \eta$, then $y = w + \eta$, so $\frac{dy}{dw} = 1$, $\Rightarrow dy = dw$.

So, $f_W(w) = ne^{-(w)n} \cdot |1|$, $w > 0$ ← Jacobian

$= ne^{-nw}, \quad w > 0$

now won't have to integrate all over again later

$\therefore W = X_{1:n} - \eta \sim EXP(\frac{1}{n})$

So, $E(\hat{\eta}_2) = E(X_{1:n}) = E(\underbrace{X_{1:n} - \eta}_{=W} + \eta) = E(W) + \eta$

$= \left(\frac{1}{n}\right) + \eta$

So, $b(\hat{\eta}_2) = \left(\frac{1}{n} + \cancel{\eta}\right) - \cancel{\eta} = \frac{1}{n}$.

constant, doesn't affect

Now, $Var(\hat{\eta}_2) = Var(\underbrace{X_{1:n} - \eta}_{=W} + \eta) = Var(W) = \frac{1}{n^2}$.

$\therefore MSE(\hat{\eta}_2) = \frac{1}{n^2} + \left(\frac{1}{n}\right)^2 = \frac{2}{n^2}$.

~~MSE~~

x	$MSE(\hat{\tau}_1)$ $= \frac{1}{n}$	$MSE(\hat{\tau}_2)$ $= \frac{2}{n^2}$	
1	1	2	$\hat{\tau}_1$ better here cuz $MSE(\hat{\tau}_1) < MSE(\hat{\tau}_2)$
2	$\frac{1}{2}$	$\frac{1}{2}$	$MSE(\hat{\tau}_1) = MSE(\hat{\tau}_2)$
3	$\frac{1}{3}$	$\frac{2}{9}$	$>$
			⋮

Brute-force way of comparing

For $x \geq 3$, will find $MSE(\hat{\tau}_1) > MSE(\hat{\tau}_2)$.

An alternative way to compare:

Take $MSE(\hat{\tau}_1) - MSE(\hat{\tau}_2) = \frac{1}{n} - \frac{2}{n^2}$

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

$= \underbrace{\frac{1}{n} \cdot \frac{n-2}{n}}_{>0}$

works cuz can factor out a $\frac{1}{n}$.

when $n < 2$, get negative #, so $MSE(\hat{\tau}_1) < MSE(\hat{\tau}_2)$.

When $n = 2$, get 0, so both ok then.

When $n > 2$, get positive #, so $MSE(\hat{\tau}_1) > MSE(\hat{\tau}_2)$.

Want the smaller MSE always cuz want variance to be small and bias to be small.

So better one is dependent on sample size n.

e.g., if $n = 5$, want to use $\hat{\tau}_2$.

(So no Indicator fns allowed)

§9.4 Notes

STAT 480B

Asymptotic Properties of the MLE

~~Use~~ Should be able to get MLE using derivatives only, support does not dep. on param.

Thm Under certain regularity conditions, the solutions $\hat{\theta}_n$ of maximum likelihood equations have the following properties:

fixed n

# of observations

1. The MLE $\hat{\theta}_n$ exists and is unique.

(for fixed n)

2. The MLE $\hat{\theta}_n$ is a consistent estimator of θ, that is, for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}[|\hat{\theta}_n - \theta| < \epsilon] = 1.$$

conv. in probability. Ie, the limiting dist of $\hat{\Theta}_n$ is degenerate at Θ.

for every $\theta \in \Omega$. Ie., as $n \to \infty$, $\hat{\Theta}_n \xrightarrow{P} \Theta$.

(as $n\to\infty$, estimator will hit the param we are interested in)

3. $\hat{\theta}_n$ is asymptotically normal with asymptotic mean θ and and asymptotic variance

$$\frac{1}{n\,\mathbf{E}[\frac{\partial}{\partial\theta}\ln f(X;\theta)]^2}.$$

I.e, $E(\hat{\Theta}_n) = \Theta$ when $n \to \infty$

⇒ unbiased as $n\to\infty$ ~~cste~~

4. $\hat{\theta}_n$ is asymptotically efficient.

Ie, Variance = CRLB

~~sometimes needed if to try to calculate the above Expected Value; end up with $E(X-\mu)^2$, then this is the Variance of X, which is~~

where does the n come from?

why $\hat{\Theta}_n$? ~~Considering all possible~~

sequence is formed by talking about increasing the # of observed events $(X_1, X_2, X_3, \ldots)$ (which is n) invar'prop.

The Delta Method: If $\tau(\theta)$ is a function with a non-zero derivative, then $\hat{\tau}_n = \tau(\hat{\theta})$ is asymptotically normal with asymptotic mean $\tau(\theta)$ and asymptotic variance $[\tau'(\theta)]^2$CRLB, that is,

square this

$$\hat{\tau}_n = \tau(\hat{\theta}) \sim AN\left[\tau(\theta), \frac{[\tau'(\theta)]^2}{n\,\mathbf{E}[\frac{\partial}{\partial\theta}\ln f(X;\theta)]^2}\right].$$

sps want estimate K^2. Then $K^2 = \tau(K)$, which has nonzero deriv. $2K$.

Exam: §9.1 - 9.4.

For large n, $\hat{\theta} \sim AN\left(\theta, \frac{1}{nE[\frac{\partial}{\partial\theta}\ln f(X;\theta)]^2}\right)$. I.e., $\hat{\theta}_n \xrightarrow{P} \hat{\theta}$.

Now, for a fn of $\hat{\theta}_n$, we also have

$\hat{\tau}(\theta) = \tau(\hat{\theta}) \sim AN($

same $\hat{\theta}$

I.e, $\hat{\tau}_n(\theta) = \tau(\hat{\theta}_n) \longrightarrow \tau(\hat{\theta})$. If all conditions met, then this is same as $\tau(\hat{\theta}_n) \xrightarrow{P} \tau(\hat{\theta})$

Ex) Consider a r.s. of size n from $f(x; K) = \frac{K}{(1+x)^{K+1}}$, $x > 0$, $K > 0$

Derive the asymptotic distribution of the MLE, $\hat{K}$.

· Find $\hat{K}$.

$$L(K) = \prod_{i=1}^{n} \frac{K}{(1+x_i)^{K+1}} = K^n \cdot \prod_{i=1}^{n} (1+x_i)^{-(K+1)}$$

$$\Rightarrow \ln L(K) = n \ln(K) - \sum_{i=1}^{n} (K+1) \ln(1+x_i)$$

$$= n \ln(K) - (K+1) \sum_{i=1}^{n} \ln(1+x_i)$$

$$\Rightarrow \frac{d}{dK} \ln L(K) = \left(\frac{n}{K} - \sum_{i=1}^{n} \ln(1+x_i) \right) \quad \text{set} = 0, \text{ hats}$$

$$\Rightarrow \frac{n}{\hat{K}} = \sum_{i=1}^{n} \ln(1+x_i) \Rightarrow \hat{K} = \frac{n}{\sum_{i=1}^{n} \ln(1+x_i)}$$

$\hat{K} \sim AN$ with

mean $= K$ and variance $= \frac{1}{nE\left[\frac{d}{dK} \ln f(X;K)\right]^2} = \frac{1}{-nE\left[\frac{d^2}{dK^2} \ln f(X;K)\right]}$

Now,

$f(X;K) = \frac{K}{(1+X)^{K+1}}$, so $\ln f(X;K) = \ln(K) - (K+1)\ln(1+X)$,

so $\frac{d}{dK} \ln f(X;K) = \frac{1}{K} - \ln(1+X)$

$\Rightarrow \frac{d^2}{dK^2} \ln f(X;K) = -\frac{1}{K^2}$. So variance $= \frac{1}{-nE\left(-\frac{1}{K^2}\right)} = \frac{K^2}{n}$

$\therefore \hat{K} \sim AN\left(K, \frac{K^2}{n}\right)$

so as $n \to \infty$, will approach a Normal dist w/ these params. And as $n \to \infty$, variance $= \frac{K^2}{n} \to 0$. (which is good)

Chapter 9 Suggested Exercises

$|e^{ax}dx$
$= \frac{1}{a}e^{ax}$

$\int_1^\infty a x^a dx$
$= a\int_1^\infty x^a dx$
$= a\left[\frac{x^{a+1}}{a+1}\right.$

2/16: Problems 1bc, 2a-d, 3, 4.

1bc) Find method of moments estimators (MMEs) of θ based on a random sample $X_1, \ldots, X_n$ from each of the following pdfs:

(b) $f(x;\theta) = (\theta+1)x^{-\theta-2}$; $1 < x$, zero otherwise; $0 < \theta$.

- one parameter (θ), so need just one moment (i.e., $r=1$).
  - 1st population moment ($E(X^r)$ where $r=1$): $E(X)$, where using definition

$$E(X) = \int_1^\infty x\cdot(\theta+1)x^{-\theta-2}dx = \int_1^\infty (\theta+1)x^{-\theta-1}dx = (\theta+1)\int_1^\infty x^{-\theta-1}dx$$

$$= (\theta+1)\left[\frac{x^{-\theta-\cancel{1}+\cancel{1}}}{-\theta-\cancel{1}+\cancel{1}}\right]_1^\infty = \frac{\theta+1}{-\theta}\left[\frac{1}{x^\theta}\right]_1^\infty = -\frac{(\theta+1)}{\theta}\left[\cancel{0} - \frac{1}{1^\theta}\right] \quad = -1$$

$$= \frac{\theta+1}{\theta} = 1 + \frac{1}{\theta}$$

- 1st sample moment: $r=1$ for $\frac{\sum_{i=1}^n X_i^r}{n}$, so $\frac{\sum_{i=1}^n X_i}{n} = \overline{X}$ ← capital $\overline{X}$
- Now we find the MME $\hat{\theta}$ such that $1 + \frac{1}{\hat{\theta}} = \overline{X}$.

$$\Rightarrow \frac{1}{\hat\theta} = \overline{X} - 1 \quad \Rightarrow \boxed{\hat\theta = \frac{1}{\overline{X}-1}} \quad ✓$$

Cheat sheets?

(c) $f(x;\theta) = \theta^2 x e^{-\theta x}$; $0 < x$, zero otherwise; $0 < \theta$.

- one parameter, so only need moments for $r=1$:
  - 1st pop. moment: $E(X^{(1)}) = E(X)$, where

$$E(X) = \int_0^\infty x\cdot\theta^2 x e^{-\theta x}dx = \theta^2\int_0^\infty x^2 e^{-\theta x}dx$$

let $u = x^2$, $\int dv = \int e^{-\theta x}dx$
$\Rightarrow \frac{du}{dx} = 2x \Rightarrow du = 2x\,dx$ ↳ $\Rightarrow v = -\frac{1}{\theta}e^{-\theta x}$

$$= \theta^2\left[(x^2)\left(-\tfrac{1}{\theta}e^{-\theta x}\right)\Big|_0^\infty - \int_0^\infty\left(-\tfrac{1}{\theta}e^{-\theta x}\cdot 2x\,dx\right)\right]$$

$$= \theta^2\left[-\frac{x^2}{\theta}e^{-\theta x}\Big|_0^\infty\right.$$

yeah this one just sucks, but would be helpful to review for the sake of practice for an exam. (tedious)

2a-d) Find the MMEs based on a random sample of size n from each of the following distributions (see Appendix B):

(a) $X_i \sim NB(3, p)$.

$\underbrace{3}_{r\text{th success}}$, $\underbrace{p}_{\text{prob of success}}$

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

· two parameters:

don't need use definition; know we have an NB dist., whose mean is given by $\frac{r}{p}$, variance given by $\frac{rq}{p^2}$

1st pop. moment: $E(X) = \frac{3}{p}$

~~2nd pop. moment: $E(X^2) = Var(X) + [E(X)]^2$~~

~~$= \frac{3(1-p)}{p^2} + \left(\frac{3}{p}\right)^2 = \frac{3-3p}{p^2} + \frac{9}{p^2} = \frac{12-3p}{p^2}$~~

1st sample moment: $\bar{X}$

~~2nd sample moment: $\frac{\sum_{i=1}^{n} X_i^2}{n}$~~

· Now, ~~find $\hat{r}$ and~~ $\hat{p}$ such that $\frac{3}{\hat{p}} = \bar{X}$ ~~and $\frac{12-3\hat{p}}{\hat{p}^2} = \frac{\sum_{i=1}^{n} X_i^2}{n}$.~~

~~$\boxed{\hat{r} = 3}$ (given)~~

Ask sufficient?

$\Rightarrow \boxed{\hat{p} = \frac{3}{\bar{X}}}$

~~$\Rightarrow n(12-3\hat{p}) = \hat{p}^2 \sum_{i=1}^{n} X_i^2$~~

how solve?

~~Ask~~

~~ask~~

do I need to match the "index" of the parameter to the index of the moments used? no

or do I only care about # of unknown parameters? Yes

all that needed to be found, looks like expected value sort

$\frac{3}{\hat{p}} = \bar{X}$

(trick only works for MME cuz uses $E(X)$.

Don't need deal with 2nd moments. # of moments = # of unknown params. Only p is unknown, so just look at the 1st pop and 1st samp. moments and solve for $\hat{p}$.

$\boxed{\hat{p} = \frac{3}{\bar{X}}}$

~~(b) $X_i \sim GAM(2, \kappa)$~~

(b) $X_i \sim GAM(2, \kappa)$ (θ κ). (c) $X_i \sim WEI(\theta, 1/2)$ (θ β), (d) $X_i \sim DE(\theta, \eta)$ (both unknown)

(b) only one unknown param, κ, so just need $r=1$ moments

1st pop moment ($r=1$) → $E(X) = E(X)_{GAM(\theta,\kappa)} = \kappa\theta = 2\kappa$

Want MME $\hat{\kappa}$ such that $2\hat{\kappa} = \bar{X}$ (1st sample moment). So $\boxed{\kappa = \frac{1}{2}\bar{X}}$

(c) again, only one unknown param, θ, so

$$E(X) = E(X)_{WEI(\theta,\beta)} = \theta\Gamma(1+\tfrac{1}{\beta}) = \theta\cdot\Gamma(1+(\tfrac{1}{1/2})) = \theta\cdot\Gamma(1+2)$$

$$= \theta\cdot\Gamma(3) = \theta(3-1)! = \theta\cdot 2! = 2\theta$$

So, want MME ~~s.t.~~ $\hat{\theta}$ s.t. $2\hat{\theta} = \bar{X}$. So $\boxed{\hat{\theta} = \frac{1}{2}\bar{X}.}$

(d) Two unknowns this time.

1st pop moment: $E(X) = E(X)_{DE(\theta,\eta)} = \eta$

1st sample moment: $\bar{X}$. So want $\hat{\eta} = \bar{X}$. ← $\hat{\eta}$ already in terms of just its own stuff. I.e., don't have the other parameter involved, so we're done solving for $\hat{\eta}$ already. all in terms of statistics ✓

$Var(X) = E(X^2) - [E(X)]^2$
$\Rightarrow E(X^2) = Var(X) + \dots$ Right idea? ask Yes

2nd pop moment: $E(X^2) = Var(X) + [E(X)]^2$

$$= Var(X)_{DE(\theta,\eta)} + [E(X)_{DE(\theta,\eta)}]^2$$

$$= 2\theta^2 + [\eta]^2$$ ← all in terms of the params already ✓

2nd sample moment: $\frac{\sum_{i=1}^n X_i^2}{n}$

So we want $2\hat{\theta}^2 + \hat{\eta}^2 = \frac{\sum_{i=1}^n X_i^2}{n}$. Here, we know already that $\boxed{\hat{\eta} = \bar{X},}$ ask Yes

so plug it in and solve for $\hat{\theta}$.

$$\Rightarrow 2\hat{\theta}^2 + \bar{X}^2 = \frac{\sum_{i=1}^n X_i^2}{n} \Rightarrow 2\hat{\theta}^2 = \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n}\cdot\frac{(n-1)}{(n-1)}$$

$$\Rightarrow 2\hat{\theta}^2 = \frac{1}{n}\cdot\frac{(n-1)}{(n-1)}\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) = \frac{n-1}{n}\cdot\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1} = \frac{n-1}{n}S^2$$

$$\Rightarrow 2\hat{\theta}^2 = \frac{n-1}{n}S^2 \Rightarrow \hat{\theta}^2 = \frac{1}{2}\frac{(n-1)}{n}S^2 \Rightarrow \hat{\theta} = \sqrt{\frac{S^2(n-1)}{2n}} \Rightarrow \boxed{\hat{\theta} = S\sqrt{\frac{n-1}{2n}}}$$

Can I leave in terms of $\bar{X}$? ask Yes.

Her answer: $\hat{\theta} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} = \sqrt{\frac{\sum_{i=1}^n X_i^2}{n} - (\bar{X})^2}$ ask

$\hat{\eta} = \bar{X} - \theta$?

$E(S^2) = \sigma^2$ just for Normal. $E(\bar{X}) = E(X)$ works for all.

Ask $Var(S^2)$ for anything other than a Normal dist?

2d) (Rewritten w/ corrections)

Find the MMEs based on a r.s. of size n, ~~from~~ $X_i \sim DE(\theta, \eta)$.

got it correct

3) Find maximum likelihood estimators (MLEs) for θ based on a random sample of size n for each of the pdfs in Exercise 1.

(1b) $f(x;\theta) = (\theta+1)x^{-\theta-2}$; $1<x$, zero o.w ; $0<\theta$. (from R.sample $X_1,\dots,X_n$)

· Goal: find $\hat{\theta}$ that maximizes $L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$

· $L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} (\theta+1)x_i^{-\theta-2} = (\theta+1)^n \cdot \underbrace{x_i^{\sum_{i=1}^{n}(-\theta-2)}}$

wrong, need same base to do this.

$= (\theta+1)^n x_i^{-\sum_{i=1}^{n}(\theta+2)} = (\theta+1)^n x_i^{-n(\theta+2)}$

$\Rightarrow \ln L(\theta) = \ln\left((\theta+1)^n \cdot x_i^{-n(\theta+2)}\right) = \ln(\theta+1)^n + \ln(x_i)^{-n(\theta+2)}$

$= n\ln(\theta+1) - n(\theta+2)\ln(x_i)$ $\quad n[\ln(\theta+1) - (\theta+2)\ln(x_i)]$

$\Rightarrow \frac{d}{d\theta}\ln L(\theta) = \frac{d}{d\theta}(n\ln(\theta+1)) - \frac{d}{d\theta}(n(\theta+2)\ln(x_i))$

$= \frac{n}{\theta+1}(1) - \frac{d}{d\theta}(n\ln(x_i)\,\theta) - \frac{d}{d\theta}(n\ln(x_i)\,2) \to 0$

$= \frac{n}{\theta+1} - n\ln(x_i)$, set equal to 0; solve:

ask

help

$\frac{n}{\hat{\theta}+1} - n\ln(x_i) = 0 \Rightarrow \frac{1}{\hat{\theta}+1} = \ln(x_i) \Rightarrow \hat{\theta}+1 = \frac{1}{\ln(x_i)}$

$\Rightarrow \hat{\theta} = \underbrace{\frac{1}{\ln(x_i)}} - 1$

this term is wrong

ask

what does answer need be in terms of?

doesn't matter, solve for param.

$\frac{d}{dx}2^x = \ln(2)\cdot 2^x$

$L(\theta) = (\theta+1)^n \cdot x_i^{-n(\theta+2)}$

$\frac{d}{d\theta}L(\theta) = \frac{d}{d\theta}\left[(\theta+1)^n \cdot x_i^{-n(\theta+2)}\right]$ use product rule

$= (\theta+1)^n \cdot \frac{d}{d\theta}\left(x_i^{-n(\theta+2)}\right) + \frac{d}{d\theta}\left[(\theta+1)^n\right]\cdot x_i^{-n(\theta+2)}$

$= (\theta+1)^n \cdot \ln(x_i)\cdot x_i^{-n(\theta+2)}\cdot(-n) + (n)(\theta+1)^{n-1}(1)\cdot x_i^{-n(\theta+2)}$

$= n(\theta+1)^{n-1} x_i^{-n(\theta+2)} - n(\theta+1)^n \ln(x_i)\cdot x_i^{-n(\theta+2)}$, set = to 0, solve for θ...

yeah don't do this

let $u = \theta+2$,

$\frac{d}{d\theta} x_i^{-n(\theta+2)} = \frac{d}{du} x_i^{-nu} \cdot \frac{d}{d\theta}(u+2)$

$= \ln(x_i)\cdot x_i^{-nu}\cdot(-n)\cdot(1)$

see back

(1b) $f(x;\theta) = (\theta+1)x^{-\theta-2}$; $1 < x$, zero o.w. ; $0 < \theta$.

$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} (\theta+1) x_i^{-(\theta+2)} = (\theta+1)^n \cdot \prod_{i=1}^{n} x_i^{-(\theta+2)}$

$\Rightarrow \ln L(\theta) = n\ln(\theta+1) + \ln x_1^{-(\theta+2)} + \ln x_2^{-(\theta+2)} + \dots + \ln x_n^{-(\theta+2)}$

$= n\ln(\theta+1) - (\theta+2)\sum_{i=1}^{n} \ln(x_i)$

$\Rightarrow \frac{d}{d\theta}\ln L(\theta) = \frac{n}{\theta+1} - \sum_{i=1}^{n}\ln(x_i) \overset{set}{=} 0$

let $u = \theta+2$

Then do $\frac{d}{du}\left[-u \underbrace{\sum_{i=1}^{n}\ln(x_i)}_{constant}\right]\cdot\frac{d}{d\theta}(\theta+2)$

$= -\sum_{i=1}^{n}\ln(x_i)\cdot(1)$

$\Rightarrow \frac{n}{\hat\theta+1} = \sum_{i=1}^{n}\ln(x_i) \Rightarrow \frac{\hat\theta+1}{n} = \frac{1}{\sum_{i=1}^{n}\ln(x_i)} \Rightarrow \hat\theta+1 = \frac{n}{\sum_{i=1}^{n}\ln(x_i)}$

$\Rightarrow \boxed{\hat\theta = \frac{n}{\sum_{i=1}^{n}\ln(x_i)} - 1}$ ✓

ask Valid? Yes. And then once $\hat\theta < 0$, we don't include those values, implicitly. Need satisfy conditions of pdf.

$x_i > 1$, so $\ln(x_i) > 0$, so $\sum_{i=1}^{n}\ln(x_i) > 0$.

But $\hat\theta > 0$ only so long as $\sum_{i=1}^{n}\ln(x_i) < 1$ ... ~~ask~~

How do we know $\theta > 0$ will be maintained? Isn't it possible that $\hat\theta \le 0$ eventually? (Yes)

(1c) $f(x;\theta) = \theta^2 x e^{-\theta x}$; $0 < x$, zero o.w. ; $0 < \theta$.

$L(\theta) = \prod_{i=1}^{n} f(x;\theta) = \prod_{i=1}^{n}\theta^2 x_i e^{-\theta x_i}$

$= \theta^2 x_1 e^{-\theta x_1}\cdot\theta^2 x_2 e^{-\theta x_2}\cdots\theta^2 x_n e^{-\theta x_n}$

$= (\theta^2)^n\cdot\prod_{i=1}^{n} x_i \cdot e^{\sum_{i=1}^{n}(-\theta x_i)}$

$= \theta^{2n}\cdot\prod_{i=1}^{n} x_i\cdot e^{n(-\theta)\sum_{i=1}^{n} x_i}$

$= \ln(x_1\cdot x_2\cdots x_n)$

$= \ln x_1 + \ln x_2 + \ln x_3 \dots + \ln x_n$

$\Rightarrow \ln L(\theta) = \ln\left[\theta^{2n}\cdot\prod_{i=1}^{n}x_i\cdot e^{-n\theta\sum_{i=1}^{n}x_i}\right] = \ln\theta^{2n} + \ln e^{-n\theta\sum_{i=1}^{n}x_i} + \ln\left(\prod_{i=1}^{n}x_i\right)$

$= 2n\ln\theta - n\theta\sum_{i=1}^{n}x_i\,(1) + \ln\left(\prod_{i=1}^{n}x_i\right)$

$\Rightarrow \frac{d}{d\theta}\ln L(\theta) = \frac{d}{d\theta}\left(2n\ln(\theta)\right) - \frac{d}{d\theta}\left(n\theta\sum_{i=1}^{n}x_i\right) + \cancel{\frac{d}{d\theta}\left(\ln\left(\prod_{i=1}^{n}x_i\right)\right)} \to 0$

$= \frac{2n}{\theta} - n\cdot\sum_{i=1}^{n}x_i$ ← $\bar{x} = \frac{\sum_{i=1}^{n}x_i}{n}$, so $\sum_{i=1}^{n}x_i = n\bar{x}$

$= \frac{2n}{\theta} - n\bar{x}$ ← set = to 0, solve for θ

$\Rightarrow \frac{2n}{\hat\theta} - n\bar{x} = 0 \Rightarrow \frac{2n}{\hat\theta} = n\bar{x} \Rightarrow \frac{2}{\hat\theta} = \bar{x} \Rightarrow \frac{\hat\theta}{2} = \frac{1}{\bar{x}} \Rightarrow \boxed{\hat\theta = \frac{2}{\bar{x}}}$

ask why? How identify? No way to, would be told which method.

note: same as what get from MME, coincidentally

$$L(p) = \frac{p^{3n}}{2^n}(1-p)^{\sum_{i=1}^{n}(x_i-3)} \cdot \prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]$$

$$\Rightarrow \ln L(p) = 3n\ln(p) - n\ln(2) + \sum_{i=1}^{n}(x_i-3)\ln(1-p) + \ln\left(\prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]\right)$$

$$\ln\left(\prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]\right) = \ln\left((x_1-1)(x_1-2)\cdot(x_2-1)(x_2-2)\cdots(x_n-1)(x_n-2)\right)$$

$$= \ln\left[(x_1-1)(x_1-2)\right] + \ln\left[(x_2-1)(x_2-2)\right] + \cdots + \ln\left[(x_n-1)(x_n-2)\right]$$

$$= \ln(x_1-1) + \ln(x_1-2) + \ln(x_2-1) + \ln(x_2-2) + \cdots + \ln(x_n-1) + \ln(x_n-2)$$

$$= \ln(x_1-1) + \ln(x_2-1) + \cdots + \ln(x_n-1) + \ln(x_1-2) + \ln(x_2-2) + \cdots + \ln(x_n-2)$$

$$= \sum_{i=1}^{n}\ln(x_i-1) + \sum_{i=1}^{n}\ln(x_i-2)$$

so $\ln\left(\prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]\right) = \sum_{i=1}^{n}\ln(x_i-1) + \sum_{i=1}^{n}\ln(x_i-2)$.

other way to see: treat $\prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]$ as $\left[\prod_{i=1}^{n}(x_i-1)\right]\left[\prod_{i=1}^{n}(x_i-2)\right]$.

Then $\ln\left(\left[\prod_{i=1}^{n}(x_i-1)\right]\cdot\left[\prod_{i=1}^{n}(x_i-2)\right]\right) = \ln\left[\prod_{i=1}^{n}(x_i-1)\right] + \ln\left[\prod_{i=1}^{n}(x_i-2)\right]$

I.e., $\ln\left[\prod_{i=1}^{n}\left[(x_i-1)(x_i-2)\right]\right] = \sum_{i=1}^{n}\ln\left[(x_i-1)(x_i-2)\right]$ $= \sum_{i=1}^{n}\ln(x_i-1) + \sum_{i=1}^{n}\ln(x_i-2)$

$= \sum_{i=1}^{n}\left[\ln(x_i-1) + \ln(x_i-2)\right] =$

So anyway,

$$\ln L(p) = 3n\ln(p) - n\ln(2) + \sum_{i=1}^{n}(x_i-3)\ln(1-p) + \sum_{i=1}^{n}\ln(x_i-1) + \sum_{i=1}^{n}\ln(x_i-2)$$

$$\Rightarrow \frac{d}{dp}\ln L(p) = \frac{3n}{p} - 0 + \sum_{i=1}^{n}(x_i-3)\cdot\frac{1}{1-p}(-1) + 0 + 0$$

$$= \frac{3n}{p} - \frac{\sum_{i=1}^{n}(x_i-3)}{1-p} \overset{\text{set}}{=} 0$$

$$\frac{3n}{\hat{p}} = \frac{\sum_{i=1}^{n}(x_i-3)}{1-\hat{p}} \Rightarrow 3n - 3n\hat{p} = \hat{p}\sum_{i=1}^{n}(x_i-3) \Rightarrow -3n\hat{p} - \hat{p}\sum_{i=1}^{n}(x_i-3) = -3n$$

$$\Rightarrow \hat{p}\left(3n + \sum_{i=1}^{n}(x_i-3)\right) = 3n \Rightarrow \hat{p} = \frac{3n}{3n + \sum_{i=1}^{n}(x_i-3)} = \frac{3n}{\cancel{3n} + \sum_{i=1}^{n}(x_i) - \cancel{3n}} = 3\cdot\frac{n}{\sum_{i=1}^{n}x_i} = 3\cdot\frac{1}{\bar{x}} = \boxed{\frac{3}{\bar{x}}}$$

4c) $X_i \sim NB(3, p)$, find MLE

$\binom{n}{r} = \frac{n!}{\;}$

- first, what is $f(x_i; 3, p)$?

$$f(x; 3, p) = \binom{x-1}{3-1} p^3 q^{x-3}$$

$$= \binom{x-1}{2} p^3 (1-p)^{x-3}$$

$$= \frac{(x-1)!}{2!\,(x-3)!} p^3 (1-p)^{x-3}$$

$$\frac{(x-1)!}{2(x-3)!} = \frac{(x-1)(x-2)\cancel{(x-3)!}}{2\cancel{(x-3)!}} = \frac{(x-1)(x-2)}{2}$$

$$= \frac{(x-1)(x-2)}{2} p^3 (1-p)^{x-3}$$

just p; it's the only unknown param

$$\Rightarrow L(p) = \prod_{i=1}^{n} f(x_i; 3, p)$$

$$= \prod_{i=1}^{n} \frac{(x_i-1)(x_i-2)}{2} p^3 (1-p)^{x_i-3}$$

$$= \frac{p^{3n}}{2^n}\left[\prod_{i=1}^{n}(x_i-1)(x_i-2)\right](1-p)^{\sum_{i=1}^{n}(x_i-3)}$$

cuz same base

$$\ln\left(\prod_{i=1}^{n}(x_i-1)(x_i-2)\right) = \sum_{i=1}^{n}\ln\left[(x_i-1)(x_i-2)\right] = \sum_{i=1}^{n}\left[\ln(x_i-1)+\ln(x_i-2)\right]$$

$$\Rightarrow \ln L(p) = 3n\ln(p) - n\ln(2) + \sum_{i=1}^{n}\left[\ln(x_i-1)+\ln(x_i-2)\right] + \left[\sum_{i=1}^{n}(x_i-3)\ln(1-p)\right]$$

so $$\frac{d}{dp}\ln L(p) = \frac{3n}{p} - 0 + 0 + \frac{\sum_{i=1}^{n}(x_i-3)}{1-p}(-1)$$

$$= \frac{3n}{p} - \frac{\sum_{i=1}^{n}(x_i-3)}{1-p} \xRightarrow{\text{set } = 0} \frac{3n}{\hat{p}} = \frac{\sum_{i=1}^{n}(x_i-3)}{1-\hat{p}}$$

$$\Rightarrow 3n - 3n\hat{p} = \hat{p}\sum_{i=1}^{n}(x_i-3) \Rightarrow 3n = \hat{p}\left[\underbrace{\sum_{i=1}^{n}(x_i-3) + 3n}_{= \sum_{i=1}^{n} x_i - \cancel{3n}}\right]$$

$$\Rightarrow 3n = \hat{p}\sum_{i=1}^{n} x_i$$

$$\Rightarrow \hat{p} = \frac{3n}{\sum_{i=1}^{n} x_i} = \frac{3}{\bar{x}}$$ ← estimate w/ little x's

so estimator would be $\boxed{\hat{p} = \frac{3}{\bar{X}}}$

5) Find the MLE for θ based on a r.s. of size n from a dist. w/ pdf

$$f(x;\theta) = \begin{cases} 2\theta^2 x^{-3}, & \theta \le x \\ 0, & x < \theta \end{cases} ; \quad 0 < \theta.$$

ask What if the other part was something other than 0? Or a different param?

Only care about nonzero pdf? We are trying to maximize $L(\theta)$, so if we use $f(x) = 0$, then this is not useful. By definition, a pdf is never negative, so $f(x) = 0$ should not be what yields a maximum for L.

$$L(\theta) = \prod_{i=1}^{n} 2\theta^2 x_i^{-3} \cdot I\{\theta \le x_i\}$$

$$= 2^n(\theta^2)^n \cdot \prod_{i=1}^{n} x_i^{-3} \cdot \underbrace{\prod_{i=1}^{n} I\{\theta \le x_i\}}$$

for our $I\{A\}$ fcn to be nonzero, we need every x_i to be $\geq \theta$ (i.e., so that our pdf is never possibly zero) cuz support

$$= 2^n \cdot \theta^{2n} \cdot \prod_{i=1}^{n} x_i^{-3} \cdot I\{\theta \le x_{1:n}\}$$

$\therefore$ because an indicator fcn is involved, to maximize L, we must analyze L's behavior as a fcn of θ.

$$L(\theta) = \underbrace{2^n \cdot \prod_{i=1}^{n} x_i^{-3}}_{\text{like constants, } > 0} \cdot \theta^{2n}$$

here, as $\theta \uparrow$'s, $\theta^{2n} \uparrow$'s, so $L \uparrow$'s.

$\therefore$ L is an increasing fcn, which attains a max value when θ is largest. Well, $\theta \le x_{1:n}$, so L is maximized when $\boxed{\hat{\theta} = x_{1:n}}$ ✓

$$f(x) = \begin{cases} g(x), & x > \theta \\ h(x), & x < 2 \text{ or something} \end{cases}$$

Typically, when support dep. on θ, will not have a piecewise.

Should never encounter a piecewise w/ a nonzero $h(x)$ cuz when do $L = \prod f(x_i;\theta)$, we don't know where the x_i's lie.

ask? memorize? MLE/MME results for $N(\mu, \sigma^2)$

Meh.

8) $X_1, X_2, \ldots, X_n$ r.s. from $N(\mu, \sigma^2)$. Want MLE of

a) $P(X > c)$.

Recall: MLEs of μ and σ^2 are $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} X_i^2}{n} - (\bar{X})^2$

$= \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$

$P(X > c) = 1 - P(X < c)$

$= 1 - P\left(Z < \frac{c-\mu}{\sigma}\right)$ want to use table

fcn of our parameters $\tau(\mu, \sigma^2)$

$= 1 - \Phi\left(\frac{c-\mu}{\sigma}\right)$ ← cuz for an arbitrary c prob. is know expressed in terms of our params.

so $\widehat{P(X>c)} = \boxed{1 - \Phi\left(\frac{c-\hat{\mu}}{\hat{\sigma}}\right), \text{ where } \hat{\mu} = \bar{X} \text{ and } \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}}$

csh

if leave as $1 - P\left(Z < \frac{c-\hat{\mu}}{\hat{\sigma}}\right)$, then need say we must then use table.

Φ notation used to say we're using the stan. normal dist.

---

as $n \to \infty$, the MLE will always be AN with a mean = param (so unbiased), and its asymptotic variance will be the CRLB.

(I.e., finite sample variance) not the same as its normal variance

csh

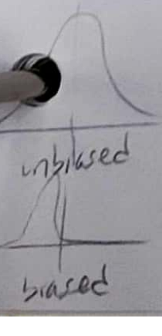
$\therefore$ For large n, have unbiasedness and (asymptotic) variance = CRLB, so UMVUE.

For large n, in other words, an MLE becomes UMVUE. I.e., what it converges to is UMVUE. It converges to a statistic.

If get to end and CRLB ≠ Variance, then can conclude nothing.

– so then how do we then say whether UMVUE?

We don't. We say inconclusive and we're done.



estimators can be
- unbiased,
- UMVUE,
- biased but low variance, etc.

a biased-but-small variance T_2 might be better than a UMVUE T_1.

#13) $L(\mu, \sigma_1^2, \sigma_2^2) = \underbrace{f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)}_{\text{joint pdf}}$ ← each should have the $; \sigma_1^2, \sigma_2^2$ etc.

μ fcn of x_i and of y_i

c of 2 indep. $= \prod_{i=1}^{n} f_1(x_i) \cdot f_2(y_i)$

#15) $X \sim BIN(n, p)$, $\hat{p} = \frac{X}{n}$ ← given the estimator

(a) Find c s.t. $E[c\hat{p}(1-\hat{p})] = p(1-p)$. (I.e, s.t. $c\hat{p}(1-\hat{p})$ is unbiased for $p(1-p)$.)

$E[c\,\hat{p}(1-\hat{p})] = c\,E\left[\frac{X}{n}\left(1-\frac{X}{n}\right)\right]$

$= c\,E\left[\frac{X}{n} - \frac{X^2}{n^2}\right]$

$= c\left[E\left(\frac{X}{n}\right) - E\left(\frac{X^2}{n^2}\right)\right] = c\left[\frac{1}{n}\overbrace{E(X)}^{=np} - \frac{1}{n^2}E(X^2)\right]$ ← $= \underbrace{Var(X)}_{np q} + \underbrace{[E(X)]^2}_{(np)^2}$

$= c\left[\frac{1}{n}(np) - \frac{1}{n^2}\left(np(1-p) + (np)^2\right)\right]$

$= c\left[p - \frac{1}{n^2}(np - np^2 + n^2p^2)\right]$

$= c\left[p - \frac{p}{n} + \frac{p^2}{n} - p^2\right]$

$= cp\left[1 - \frac{1}{n} + \frac{p}{n} - p\right]$ ~~$= cp\left[1 - p + \frac{p}{n} - \frac{1}{n}\right]$~~

$= cp\left[\left(1-\frac{1}{n}\right) - p\left(1-\frac{1}{n}\right)\right]$ ~~$= cp\,[(1)(1-p) - \frac{1}{n}(p+$~~

$= cp(1-p)\left(1-\frac{1}{n}\right)$

Thus $c\left(1-\frac{1}{n}\right)p\cancel{(1-p)} \overset{\text{must}}{=} p\cancel{(1-p)}$ ~~$\Rightarrow c - \frac{c}{n} = 0 \Rightarrow c\left(1-\frac{1}{n}\right)$~~

$\Rightarrow c\left(1-\frac{1}{n}\right) = 1 \Rightarrow c = \frac{1}{1-\frac{1}{n}} = \frac{1}{\frac{n-1}{n}} = \boxed{\frac{n}{n-1}}$

(so $\frac{n}{n-1}\hat{p}(1-\hat{p})$ is unbiased for $p(1-p)$)

biasedness can be componentwise

Sps didn't do pt. (a) of #15, what if try do Part (b)?

I.e., $X \sim BIN(n,p)$, $Var(X) = np(1-p)$
given estimate

$$E(\hat{p}) = \frac{n}{n+1}p \,, \quad \hat{p}_1 = \frac{n+1}{n}\hat{p}$$

$X_1, X_2, \ldots, X_n$ r.s. $UNIF(\theta-1, \theta+1)$. Let $Y = X_{1:n}$. Want $f_Y(y)$.

17b) $E(X_{1:n}) = \theta - \frac{n-1}{n+1}$, $Y = X_{1:n}$

$$f_Y(y) = n\left[1 - F_X(y)\right]^{n-1} f_X(y), \text{ where } F_X(y) = \int_{\theta-1}^{y} \frac{1}{(\theta+1)-(\theta-1)}\,dx$$

$y < \theta+1$

$$= \frac{1}{2}\cdot x \Big|_{\theta-1}^{y} = \frac{1}{2}(y-\theta+1)$$

$$= n\left[1 - \frac{1}{2}(y-\theta+1)\right]^{n-1}\cdot\left(\frac{1}{2}\right)$$

$$= n\left[\frac{2-y+\theta-1}{2}\right]^{n-1}\cdot\frac{1}{2} = \frac{n}{2^n}\cdot(1-y+\theta)^{n-1}, \quad \theta-1 < y < \theta+1$$

Now, want $E(X_{1:n}) = E(Y)$.

$$E(Y) = \int_{\theta-1}^{\theta+1} y\cdot\frac{n}{2^n}(1-y+\theta)^{n-1}\,dy$$

Let $u = 1-y+\theta$, then $du = -dy$.

$u_1 = 1 - \cancel{(\theta+1)} + \cancel{\theta} = 0$

$u_2 = 1 - (\theta+1) + \theta = 2$

$\Rightarrow y = 1+\theta-u$

$$= \frac{n}{2^n}\int_2^0 (1+\theta-u)u^{n-1}(-du)$$

$$= \frac{n}{2^n}\int_0^2 \left[(1+\theta)u^{n-1} - u^n\right]du$$

$$= \frac{n}{2^n}\left[(1+\theta)\frac{u^n}{n} - \frac{u^{n+1}}{n+1}\Big|_0^2\right] = \frac{n}{2^n}\left[(1+\theta)\frac{2^n}{n} - \frac{2^{n+1}}{n+1}\right]$$

So $$E(Y) = \frac{n}{\cancel{2^n}}\cdot\cancel{2^n}\left(\frac{1+\theta}{n} - \frac{2}{n+1}\right) = \cancel{n}\left(\frac{(1+\theta)(n+1) - 2(n)}{\cancel{n}(n+1)}\right) = \frac{n+1+\theta n+\theta-2n}{n+1}$$

$$= \frac{-n+1+\theta(n+1)}{n+1} = \frac{\theta\cancel{(n+1)}}{\cancel{n+1}} - \frac{(n-1)}{n+1} = \boxed{\theta - \frac{n-1}{n+1}}$$

same idea for $X_{n:n}$.

21) Consider a r.s. of size n from a Bernoulli distribution, $X_i \sim BIN(1, p)$.

one trial is performed. We'll call this k, I guess. So $k=1$.

(a) Find the CRLB for the variances of unbiased estimators of p.

- first, need an ~~unbiased~~ estimator $\hat{p}$ of p. Use either MLE or MME.

1st pop moment: $E(X)$, where $E(X) = kp$ ~ mean for BIN dist.

1st sample moment: $\frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}$ We have $k=1$, so $E(X) = p$.

- So we have $\hat{p} = \bar{X}$.

- Check if unbiased:

~~else~~ Can I skip straight to $= E(X) = p$? Yes

$$E(\hat{p}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}(p) = \frac{1}{n}\cdot np = p$$

~~So $\hat{p}$ is~~

- Have that $\hat{p} = \bar{X}$ is an [unbiased] estimator for p. ~~So, can begin to explore/find the CLRB.~~ Here, we are saying that $\hat{p} = \bar{X}$ is an [unbiased] estimator of $\tau(p) = p$.

not needed to get a CRLB.

$$CRLB = \frac{[\tau'(p)]^2}{n \cdot E\left(\left[\frac{d}{dp}\ln f(X;p)\right]^2\right)} = \frac{[\tau'(p)]^2}{-n\, E\left(\frac{d^2}{dp^2}\ln f(X;p)\right)}$$

- $\tau(p) = p$, so $\tau'(p) = 1$.

$f(x;p) = p^x q^{1-x} = p^x (1-p)^{1-x}$, $x = 0, 1$,

so $\ln f(x;p) = \ln(p^x) + \ln(1-p)^{1-x} = x\ln(p) + (1-x)\ln(1-p)$

$\Rightarrow \frac{d}{dp}\ln f(x;p) = \frac{d}{dp}\left[x\ln(p) + (1-x)\ln(1-p)\right]$

$$= \frac{x}{p} + (1-x)\cdot\frac{-1}{1-p} = \frac{x(1-p) - p(1-x)}{p(1-p)} = \frac{x - \cancel{xp} - p + \cancel{xp}}{p(1-p)}$$

$$= \frac{x-p}{p(1-p)}$$

- so $E\left(\left[\frac{d}{dp}\ln f(X;p)\right]^2\right) = E\left(\left[\frac{X-p}{p(1-p)}\right]^2\right) = \frac{1}{p^2(1-p)^2}\cdot E\left([X-p]^2\right)$ $= Var(X) = p(1-p)$

$$= \frac{1}{p^2(1-p)^2}\cdot p(1-p) = \frac{1}{p(1-p)}$$

- Altogether, $CRLB = \frac{(1)^2}{n\cdot\left(\frac{1}{p(1-p)}\right)} = \frac{1}{\frac{n}{p(1-p)}} = \frac{p(1-p)}{n}$

$Var(\hat{p}) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{p(1-p)}{n}$

$Var(\hat{p}) = CRLB$, $\hat{p}$ unbiased, so $\hat{p} = \bar{X}$ is UMVUE for p.

$Var(\bar{X}) = Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{1}{n^2}\cdot np(1-p) = \frac{p(1-p)}{n} = \frac{Var(X)}{n}$

All of this can be found w/o knowing whether the estimator ~~is~~ is unbiased already. Yes

~~else~~

(b) Find the CRLB for the variances of unbiased estimators of $p(1-p)$.

not linear, so theorem not applicable.

· Here, $\tau(p) = p(1-p)$.

$$CRLB = \frac{[\tau'(p)]^2}{nE\left(\left[\frac{d}{dp}\ln f(X;p)\right]^2\right)} = \frac{[\tau'(p)]^2}{-nE\left(\frac{d^2}{dp^2}\ln f(X;p)\right)}$$

Ask

Shouldn't we first verify/find an unbiased estimator of $p(1-p)$, though?

No, unless want UMVUE.

· From (a), we already know the denom. Only $\tau(p)$ has changed.

$\tau(p) = p(1-p) = p - p^2$, so $\tau'(p) = 1-2p$.

$\therefore$,

$$CRLB = \frac{(1-2p)^2}{\frac{n}{p(1-p)}} = \frac{(1-4p+4p^2)(p(1-p))}{n} = \frac{(1-4p+4p^2)(p-p^2)}{n}$$

$$= \frac{p-p^2-4p^2+4p^3+4p^3-4p^4}{n} = \frac{-4p^4+8p^3-5p^2+p}{n}$$

$$= \frac{(1-2p)^2 p(1-p)}{n}$$ ✓

(c) Find a UMVUE of p.

Though unneeded to complete part (a), I showed that

$Var(\hat{p}) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{p(1-p)}{n} = CRLB$. Thus, since $Var(\hat{p}) = CRLB$ and $\hat{p}$ is an unbiased estimator for p, it follows that $\hat{p} = \boxed{\bar{X}}$ is a UMVUE of p. ✓

If didn't know (a),

would need $E[\hat{\tau}(p)] = \tau(p)$ and $Var(\hat{\tau}(p)) = Var(\tau(\hat{p})) = CRLB$.

(in order to complete (b), that is)

21(b), seeing if have an unbiased estimator of $p(1-p)$...

Need unbiased estimator [redacted] of $p(1-p)$. ← still asymptotically normal

cuz it's an MLE

- we know the MME of p is $\hat{p} = \bar{X}$
- We want the MME of a function of p, namely the MME of $\tau(p) = p(1-p)$.

By the Invariance Property, the MME of $\tau(p) = p(1-p)$ is given by $\hat{\tau}(p) = \tau(\hat{p}) = \hat{p}(1-\hat{p}) = \bar{X}(1-\bar{X})$.

ask $= E(\bar{X}\cdot\bar{X})$
$= E(\bar{X})\cdot E(\bar{X})$?
No, b/c they both depend on the same X's. Change one, change other.

- We want this to be unbiased, so check:

$$E(\hat{\tau}(p)) = E(\bar{X}(1-\bar{X})) = E(\bar{X} - \bar{X}^2) = E(\bar{X}) - E(\bar{X}^2)$$

$$= E(X) - \left[Var(\bar{X}) + [E(\bar{X})]^2\right] = p - \left[\frac{Var(X)}{n} + [E(X)]^2\right]$$

allowed? ok

Yes

$$= p - \left[\frac{p(1-p)}{n} + (p)^2\right] = p - \frac{(p-p^2)}{n} - p^2 = \frac{np}{n} - \frac{(p-p^2)}{n} - \frac{np^2}{n}$$

$$= \frac{np - p + p^2 - np^2}{n} = \frac{p(n-1) - p^2(n-1)}{n} = \frac{(p-p^2)(n-1)}{n}$$

$$= \frac{n-1}{n}\cdot p(1-p) \neq p(1-p)$$, so is not unbiased.

should we be able to anticipate this result, since

~~Thm says only lin. fcns will have a variance~~

ask

No – Thm says nothing about unbiasedness conclusions, only CRLB equality...

we can anticipate the variance of unbiased estimators of $p(1-p)$ will not equal its corresponding CRLB cuz nonlinear and what is known about (a), but cannot conclude anything about UMVUE. By the Thm.

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

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

$$\Rightarrow E(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2$$

$$= \frac{Var(X)}{n} + [E(X)]^2$$

$$= \frac{p(1-p)}{n} + (p)^2$$

28a-d) Let $X_1, \ldots, X_n$ be a random sample from $EXP(\theta)$, and define $\hat{\theta}_1 = \bar{X}$ and $\hat{\theta}_2 = n\bar{X}/(n+1)$.

(a) Find the variances of $\hat{\theta}_1$ and $\hat{\theta}_2$.

$$Var(\hat{\theta}_1) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{\theta^2}{n}.$$

$$Var(\hat{\theta}_2) = Var\left(\frac{n\bar{X}}{n+1}\right) = \left(\frac{n}{n+1}\right)^2 Var(\bar{X}) = \frac{n^2}{(n+1)^2}\left(\frac{\theta^2}{n}\right) = \frac{n\theta^2}{(n+1)^2}.$$

(b) Find the MSEs of $\hat{\theta}_1$ and $\hat{\theta}_2$.

$MSE(T) = Var(T) + [b(T)]^2$,

$b(T) = E(T) - \underbrace{\tau(\theta)}_{\text{the original parameter being estimated}}$

$$MSE(\hat{\theta}_1) = Var(\hat{\theta}_1) + [b(\hat{\theta}_1)]^2$$

$$= \frac{\theta^2}{n} + \left[E(\hat{\theta}_1) - (\theta)\right]^2$$

26 & 34

$$= \frac{\theta^2}{n} + \left[E(\bar{X}) - \theta\right]^2 = \frac{\theta^2}{n} + \left[E(X) - \theta\right]^2 = \frac{\theta^2}{n} + \left[\cancel{(\theta) - \theta}\right]^2$$

means $\hat{\theta}_1 = \bar{X}$ is unbiased

$$MSE(\hat{\theta}_2) = Var(\hat{\theta}_2) + [b(\hat{\theta}_2)]^2$$

$$= \frac{n\theta^2}{(n+1)^2} + \left[E(\hat{\theta}_2) - (\theta)\right]^2 = \frac{n\theta^2}{(n+1)^2} + \left[E\left(\frac{n}{n+1}\cdot\bar{X}\right) - \theta\right]^2$$

$$= \frac{n\theta^2}{(n+1)^2} + \left[\frac{n}{n+1}E(X) - \theta\right]^2 = \frac{n\theta^2}{(n+1)^2} + \left[\frac{n\theta}{n+1} - \theta\right]^2$$

$$= \frac{n\theta^2}{(n+1)^2} + \left[\frac{n\theta - \theta(n+1)}{n+1}\right]^2 = \frac{n\theta^2}{(n+1)^2} + \frac{(\cancel{n\theta} - \cancel{n\theta} - \theta)^2}{(n+1)^2}$$

$$= \frac{n\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} = \frac{n\theta^2 + \theta^2}{(n+1)^2} = \frac{(n+1)\theta^2}{(n+1)^2} = \frac{\theta^2}{n+1}.$$

ask you have $\cancel{\frac{\theta^2}{n+1}}$?

(c) Compare the variances of $\hat{\theta}_1$ and $\hat{\theta}_2$ for $n=2$.

$Var(\hat{\theta}_1) = \frac{\theta^2}{n}$, so for $n=2$, $Var(\hat{\theta}_1) = \frac{\theta^2}{2} = \frac{9\theta^2}{18}$

$Var(\hat{\theta}_2) = \frac{n\theta^2}{(n+1)^2}$, so for $n=2$, $Var(\hat{\theta}_2) = \frac{2\theta^2}{9} = \frac{4\theta^2}{18}$

I.e., $Var(\hat{\theta}_2) < Var(\hat{\theta}_1)$ when $n=2$.

If two unbiased estimators, then compare their variances. This is still the MSE.

ask So...? Can we make any conclusions? —depends. Are you interested in smaller-variance estimators?

We should only make conclusion based off of the MSEs, right?

(d) Compare the MSEs of $\hat{\Theta}_1$ and $\hat{\Theta}_2$ for $n < 2$.

$MSE(\hat{\Theta}_1) = \frac{\Theta^2}{n}$, so for $n=2$ $MSE(\hat{\Theta}_1) = \frac{\Theta^2}{2}$. $= \frac{3\Theta^2}{6}$

$MSE(\hat{\Theta}_2) = \frac{\Theta^2}{n+1}$, so for $n=2$ $MSE(\hat{\Theta}_2) = \frac{\Theta^2}{3}$. $= \frac{2\Theta^2}{6}$

$\boxed{\text{So for } n=2,\ MSE(\hat{\Theta}_2) < MSE(\hat{\Theta}_1).}$

~~Ask~~ Conclusion?

34 a-d, f) Consider a r.s. of size n from a dist. with discrete pdf

$f(x;p) = p(1-p)^x$; $x = 0, 1, \dots$, 0 otherwise.

compare with pdf of GEO(p). $f(x;p) = pq^{x-1}$, $x = 1, 2, \dots$

Ask ?

(a) Find the MLE of p.

$$L(p) = \prod_{i=1}^{n} f(x_i;p) = \prod_{i=1}^{n} p(1-p)^{x_i} = p^n \cdot (1-p)^{\sum_{i=1}^n x_i}$$

$$\Rightarrow \ln(L(p)) = n\ln(p) + \left(\sum_{i=1}^n x_i\right)\ln(1-p)$$

$$\Rightarrow \frac{d}{dp}\ln L(p) = \frac{n}{p} + \sum_{i=1}^n x_i \cdot \frac{1}{1-p}(-1) = \frac{n}{p} - \sum_{i=1}^n x_i \cdot \frac{1}{1-p}$$

HATS

$$\overset{set}{=} 0 \Rightarrow \frac{n}{\hat{p}} = \frac{\sum_{i=1}^n x_i}{1-\hat{p}} \Rightarrow n - n\hat{p} = \hat{p}\sum_{i=1}^n x_i \Rightarrow -n\hat{p} - \hat{p}\sum_{i=1}^n x_i = -n$$

$$\Rightarrow n\hat{p} + \hat{p}\sum_{i=1}^n x_i = n \Rightarrow \hat{p}\left(n + \sum_{i=1}^n x_i\right) = n \Rightarrow \boxed{\hat{p} = \frac{n}{n + \sum_{i=1}^n x_i}}$$ ✓

$= \frac{1}{\bar{x}+1}$

$0 < \hat{p} \le 1$

(b) Find the MLE of $\theta = \frac{1-p}{p}$.

$\hat{p}$ is the MLE of p, and $\theta = \frac{1-p}{p} = u(p)$ is a fcn of p, so the MLE of $\theta = u(p)$ is $\hat{\theta} = u(\hat{p}) = \frac{1-\hat{p}}{\hat{p}}$

$$= \frac{1 - \frac{n}{n+\sum_{i=1}^n x_i}}{\frac{n}{n+\sum_{i=1}^n x_i}} = \frac{1}{\frac{n}{n+\sum_{i=1}^n x_i}} - 1 = \frac{n + \sum_{i=1}^n x_i}{n} - 1$$

MLE $\hat{\theta}$ of θ.

$$= \cancel{1} + \frac{\sum_{i=1}^n x_i}{n} - \cancel{1} = \boxed{\bar{x}}$$ ✓

(c) Find the CRLB for variances of unbiased estimators of θ.

Have $\tau(\theta) = \theta$, where $\hat{\theta} = \bar{X}$ is an estimator of $\tau(\theta)$.

· Is it unbiased? Check.

cuz X_i's indep. $(x = 0, 1, 2, \dots)$

$$E(\hat{\theta}) = E(\bar{X}) = E(X) = \sum_{\text{all } x} x \cdot f(x;p) = \sum_{x=0}^{\infty} x \cdot f(x;p)$$

must be absolutely conv., o.w. E(X) DNE.

$$= \sum_{x=0}^{\infty} x p(1-p)^x = p(1-p)\sum_{x=0}^{\infty} \underbrace{x(1-p)^{x-1}}_{= -\frac{d}{dp}(1-p)^x} = -p(1-p)\frac{d}{dp}\underbrace{\sum_{x=0}^{\infty}(1-p)^x}_{= \frac{1}{1-(1-p)} = \left(\frac{1}{p}\right)}$$ geom.

$$= -p(1-p)\frac{d}{dp}\left(\frac{1}{p}\right) = \dots = \frac{1-p}{p}$$

another way to check

Don't need to check for unbiasedness. Only care about it when looking @ whether UMVUE.

(ask) Example where need use Thm?

~~Let's assume it's unbiased. Then,~~

$$\cancel{CRLB = \frac{[\tau(\theta)]^2}{n E\left(\left[\frac{\partial}{\partial\theta} \ln f(X;\theta\right.\right.}}$$

~~don't have this pdf... seems like theorem is needed, but~~

pt(c) $\rightarrow$ $\theta = \frac{1-p}{p} = \tau(p)$ here. We do the CRLB all in terms of p stuff.

$$CRLB = \frac{[\tau'(p)]^2}{n E\left(\left[\frac{\partial}{\partial p} \ln f(X;p)\right]^2\right)} = \frac{[\tau'(p)]^2}{-n E\left(\frac{\partial^2}{\partial p^2} \ln f(X;p)\right)}$$

- $\tau(p) = \frac{1-p}{p} = \frac{1}{p} - 1 = p^{-1} - 1$, so $\tau'(p) = -\frac{1}{p^2}$.
- $f(X;p) = p(1-p)^X$, $X = 0, 1, 2, \ldots$

$\Rightarrow \ln f(X;p) = \ln(p) + X \ln(1-p)$

$\Rightarrow \frac{\partial}{\partial p} \ln f(X;p) = \frac{1}{p} - \frac{X}{1-p} \leftarrow = \frac{(1-p) - X(p)}{p(1-p)} = \frac{1-p-Xp}{p(1-p)} = \frac{-p(1+X)+1}{p(1-p)}$

$\Rightarrow E\left(\left[\frac{\partial}{\partial p} \ln f(X;p)\right]^2\right) = E\left[\left(\frac{1}{p} - \frac{X}{1-p}\right)^2\right] =$ not good, once FOIL will get $E(X^2)$ at some pt.

CRLB can be found w/o needing to first locate an unbiased estimator.

do instead:

$$\frac{\partial^2}{\partial p^2} \ln f(X;p) = -\frac{1}{p^2} - \frac{X}{(1-p)^2}$$

$$\Rightarrow E\left[\frac{\partial^2}{\partial p^2} \ln f(X;p)\right] = E\left[\frac{1}{-p^2} - \frac{X}{(1-p)^2}\right] = \frac{-1}{p^2} - \frac{1}{(1-p)^2} E(X)$$

Let $X = Y - 1$ where $Y \sim f_Y(y) = p(1-p)^{y-1}$, $y = 1, 2, \ldots$ (Y is Geometric)

Then $E(X) = E(Y-1) = E(Y) - 1 = \frac{1}{p} - 1 = \frac{1-p}{p}$, so

$$E\left[\frac{\partial^2}{\partial p^2} \ln f(X;p)\right] = -\frac{1}{p^2} - \frac{\cancel{(1-p)}}{p(1-p)^{\cancel{2}}} = -\frac{1}{p^2} - \frac{1}{p(1-p)} = \frac{-(1-p) - (p)}{p^2(1-p)}$$

$$= \frac{-1}{p^2(1-p)}.$$ Finally,

$$CRLB = \frac{\left[-\frac{1}{p^2}\right]^2}{-n\left[\frac{-1}{p^2(1-p)}\right]} = \frac{1}{p^4} \cdot \frac{p^2(1-p)}{n} = \boxed{\frac{(1-p)}{np^2}}$$

~~34, part (f)) Find the asymptotic distribution of the MLE of θ.~~

~~· The MLE of θ we found is $\hat{\theta} = \bar{X}$~~

~~· The idea, then, is that $\hat{\theta} \sim AN$ with (asymptotic) mean θ and with asymptotic variance $\frac{1}{nE([\frac{\partial}{\partial}}$~~

~~Help?~~

d) Got $\hat{\theta} = \bar{X}$ from (b).

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

so $\hat{\theta}$ unbiased for θ.

$Var(\bar{X}) = \frac{Var(X)}{n} = \frac{Var(Y-1)}{n} = \frac{Var(Y)}{n} = \frac{1}{n} \cdot \left(\frac{1-p}{p^2}\right) = CRLB.$ so yes, UMVUE

34f) Find the Asymptotic distribution of the MLE of θ.

· Earlier, we obtained the MLE $\hat{\theta} = \bar{X}$ of the parameter θ, and we were able to do so w/o needing to analyze the behavior of $L(\theta)$ or deal w/ any Indicator fns. I.e., only derivatives were needed. Thus, the large sample properties apply to this MLE.

So, $\hat{\theta} = \bar{X} \sim AN$ with mean θ and asymptotic variance $=$ CRLB.

In part (c), we got that the CRLB is given by $\frac{1-p}{np^2}$.

$$\boxed{\text{Hence } \hat{\theta} = \bar{X} \sim AN\left(\theta, \frac{1-p}{np^2}\right).}$$

Which is to say,

$$\hat{\theta}_n = \bar{X} \xrightarrow{d} \hat{\theta} \sim N\left(\theta, \frac{1-p}{np^2}\right).$$

still a statistic

~~ask~~

Matthew Miller - Problem set 4, due 3/2/22 $\frac{55}{55}$ excellent!

1) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x;\theta) = \frac{1}{\theta} x^{\frac{1-\theta}{\theta}}$, $0<x<1$, $\theta>0$.

(a) Derive the MME of θ.

· One unknown parameter, so need only the first moments (i.e., for $r=1$).

1st population moment: $E(X)$, where

$$E(X) = \int_0^1 x \frac{1}{\theta} x^{\frac{1}{\theta}-1}\,dx = \frac{1}{\theta}\int_0^1 x^{\frac{1}{\theta}}\,dx = \frac{1}{\theta}\left[\frac{x^{\frac{1}{\theta}+1}}{\frac{1}{\theta}+1}\right]_0^1$$

$$= \frac{1}{\cancel{\theta}}\cdot\frac{\cancel{\theta}}{(1+\theta)}\left[(1)^{\frac{1}{\theta}+1} - \cancel{(0)^{\frac{1}{\theta}+1}}\right] = \frac{1}{1+\theta}$$

$$\frac{1}{\frac{1}{\theta}+1} = \frac{1}{\frac{1+\theta}{\theta}} = \frac{\theta}{1+\theta}$$

5/5

1st sample moment: $\overline{X}$

· Setting them equal to one another, we want $\hat{\theta}$ such that $\frac{1}{1+\hat{\theta}} = \overline{X}$.

$$\frac{1}{1+\hat{\theta}} = \overline{X} \Rightarrow 1+\hat{\theta} = \frac{1}{\overline{X}} \Rightarrow \boxed{\hat{\theta} = \frac{1}{\overline{X}} - 1.}$$

(b) Is the MME you obtained in (a) unbiased for θ? Justify.

— skip this part —

(c) Derive the MLE of θ.

· $f(x_i;\theta) = \frac{1}{\theta} x_i^{\frac{1}{\theta}-1}$, $0<x_i<1$, $\theta>0$.

$$\Rightarrow L(\theta) = \prod_{i=1}^n f(x_i;\theta) = \prod_{i=1}^n \frac{1}{\theta} x_i^{\frac{1}{\theta}-1} = \left(\frac{1}{\theta}\right)^n \cdot \prod_{i=1}^n x_i^{\frac{1}{\theta}-1}$$

8/8

$$\Rightarrow \ln L(\theta) = n\ln\left(\frac{1}{\theta}\right) + \ln\left[x_1^{\frac{1}{\theta}-1}\cdot x_2^{\frac{1}{\theta}-1}\cdots x_n^{\frac{1}{\theta}-1}\right]$$

$$= n[\overbrace{\ln(1)}^{0} - \ln(\theta)] + \ln x_1^{\frac{1}{\theta}-1} + \ln x_2^{\frac{1}{\theta}-1} + \ldots + \ln x_n^{\frac{1}{\theta}-1}$$

$$= -n\ln(\theta) + \left(\tfrac{1}{\theta}-1\right)\ln x_1 + \left(\tfrac{1}{\theta}-1\right)\ln x_2 + \ldots + \left(\tfrac{1}{\theta}-1\right)\ln x_n$$

$$= -n\ln(\theta) + \left(\tfrac{1}{\theta}-1\right)\left[\ln x_1 + \ln x_2 + \ldots + \ln x_n\right]$$

$$= -n\ln(\theta) + \left(\tfrac{1}{\theta}-1\right)\cdot\sum_{i=1}^n \ln x_i$$

$$\Rightarrow \frac{d}{d\theta}\ln L(\theta) = \frac{-n}{\theta} + \sum_{i=1}^n \ln x_i \cdot \frac{d}{d\theta}\left(\theta^{-1}-1\right) = \frac{-n}{\theta} + \left(\sum_{i=1}^n \ln(x_i)\right)\left(-\theta^{-2}\right)$$

$$= -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^n \ln(x_i)$$

→ cont'd

· Setting this equal to zero and solving for $\hat{\theta}$,

$$-\frac{n}{\hat{\theta}} - \frac{1}{\hat{\theta}^2}\sum_{i=1}^{n} \ln(x_i) = 0 \Rightarrow -\frac{n}{\hat{\theta}} = \frac{\sum_{i=1}^{n}\ln(x_i)}{\hat{\theta}^2}$$

$$\Rightarrow \hat{\theta}\sum_{i=1}^{n}\ln(x_i) = -n\hat{\theta}^2 \Rightarrow n\hat{\theta}^2 + \hat{\theta}\sum_{i=1}^{n}\ln(x_i) = 0$$

$$\Rightarrow \hat{\theta}\left(n\hat{\theta} + \sum_{i=1}^{n}\ln(x_i)\right) = 0$$

ask $\Rightarrow \cancel{\hat{\theta} = 0}$ or $n\hat{\theta} + \sum_{i=1}^{n}\ln(x_i) = 0$

(we given that $\theta > 0$, so not a solution.) $\Rightarrow n\hat{\theta} = -\sum_{i=1}^{n}\ln(x_i) \Rightarrow \boxed{\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)}$

[note that $0 < x_i < 1$ means $\sum_{i=1}^{n}\ln(x_i)$ will be negative, but the $-\frac{1}{n}$ makes $-\frac{1}{n}\sum_{i=1}^{n}\ln(x_i) > 0$, so $\hat{\theta} > 0$ as is required. Can also use 2nd derivative test to verify this solution is, indeed, a local maximum.].

(d) Is the MLE that you obtained in (c) unbiased for θ? Justify.

To show unbiasedness, we need to show that $E\left(-\frac{1}{n}\sum_{i=1}^{n}\ln(X_i)\right) = \theta$.

· $E\left(-\frac{1}{n}\sum_{i=1}^{n}\ln(X_i)\right) = -\frac{1}{n}E\left(\sum_{i=1}^{n}\ln(X_i)\right) \overset{\text{by linearity of expected value operator.}}{=} -\frac{1}{n}\sum_{i=1}^{n}[E(\ln(X_i))]$

· Need pdf of the r. variable $\ln(X)$; use transformation technique.

For efficiency, let $Y = -\ln(X)$. Then $-Y = \ln X \Rightarrow e^{-Y} = X = w(Y)$.

Thus, $f_Y(y) = f_X(w(y))\cdot\left|\frac{d}{dy}w(y)\right| = \frac{1}{\theta}(e^{-y})^{\frac{1}{\theta}-1}\cdot\left|\frac{d}{dy}e^{-y}\right|$

$= \frac{1}{\theta}e^{y-\frac{y}{\theta}}\cdot|-e^{-y}| = \frac{1}{\theta}e^{y-\frac{y}{\theta}}\cdot e^{-y} = \frac{1}{\theta}\cdot e^{-y/\theta}$,

where $0 < x < 1 \Rightarrow 0 < e^{-y} < 1 \Rightarrow -y < \ln(1) \Rightarrow -y < 0 \Rightarrow y > 0$.

So $Y \sim EXP(\theta)$. That is, for each $-\ln(X_i)$, we have $Y_i = -\ln(X_i) \sim EXP(\theta)$.

Hence,

$E(\hat{\theta}) = E\left(-\frac{1}{n}\sum_{i=1}^{n}(\ln X_i)\right) = \frac{1}{n}E\left(\sum_{i=1}^{n}(-\ln X_i)\right) \overset{\text{due to linearity of expected value}}{=} \frac{1}{n}\cdot\sum_{i=1}^{n}(E(-\ln X_i))$

$= \frac{1}{n}\sum_{i=1}^{n}(E(Y_i)) = \frac{1}{n}\sum_{i=1}^{n}(\theta) = \frac{1}{n}\cdot n\theta = \theta$.

$\boxed{E(\hat{\theta}) = \theta\text{, so the MLE } \hat{\theta} \text{ is unbiased for } \theta.}$

10/10

(e) Obtain the MLE of $P(X < c)$.

- $P(X < c) = F_X(c)$, where

$$F_X(x) = \int_0^x \frac{1}{\theta} t^{\frac{1}{\theta}-1} dt = \frac{1}{\theta}\int_0^x t^{\frac{1}{\theta}-1} dt = \frac{1}{\theta}\left(\theta t^{\frac{1}{\theta}}\right)\Big|_0^x = x^{1/\theta}, \quad 0 < x < 1.$$

That is,

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ x^{1/\theta}, & 0 < x < 1 \\ 1, & x \geq 1. \end{cases}$$

The constant c is some fixed value; though it is not specified exactly where c exists in the support, it only really makes sense to discuss the scenario in which $0 < c < 1$, as otherwise $F_X(c) = 0$ or $F_X(c) = 1$, which is trivially true. That is, this information does not help us, at least not in the context of discussing an MLE of a function of θ. (Since θ isn't involved at all.)

- Basically, we need only consider when $0 < c < 1$, in which case we have

$$P(X < c) = F_X(c) = c^{1/\theta}.$$

- We know the MLE of θ is $\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n} \ln(x_i)$, and $P(X < c) = F_X(c)$ is a function of θ. Hence, the MLE of $P(X < c) = u(\theta)$ is

$$\hat{u}(\theta) = u(\hat{\theta}) = c^{\frac{1}{\hat{\theta}}} = \boxed{c^{\frac{1}{-\frac{1}{n}\sum_{i=1}^{n} \ln(x_i)}}}$$

$$= c^{-\frac{N}{\sum_{i=1}^{N} \ln x_i}}$$

Uniform dist
$f(x) = \frac{1}{b-a}, \ a < x < b$

2) Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f(x; \theta_1, \theta_2) = \frac{1}{\theta_2 + 3 - 2\theta_1}, \quad 2\theta_1 < x < \theta_2 + 3$$

both parameters in the support $\Rightarrow$ need one indicator fcn.

(a) Derive the MLE of $\underset{\sim}{\theta} = (\theta_1, \theta_2)$.

$$L(\theta_1, \theta_2) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2) \cdot I\{2\theta_1 < x_i < \theta_2 + 3\}$$

$$= \prod_{i=1}^{n} \left(\frac{1}{\theta_2 + 3 - 2\theta_1}\right) \cdot I\{2\theta_1 < x_i < \theta_2 + 3\}$$

$$= \left(\frac{1}{\theta_2 + 3 - 2\theta_1}\right)^n \cdot \prod_{i=1}^{n} I\{2\theta_1 < x_i < \theta_2 + 3\}$$

all of this must be true simultaneously, for each x_i, for $I\{A\} \neq 0$.

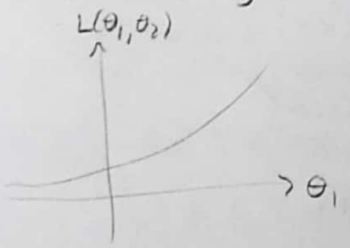
$$= \frac{1}{(\theta_2 + 3 - 2\theta_1)^n} \cdot I\{2\theta_1 < x_{1:n} < x_{n:n} < \theta_2 + 3\}.$$

To maximize L, we analyze the behavior of $L(\theta_1, \theta_2)$ as a function of each parameter, one at a time.

7/7

$L(\theta_1, \theta_2)$ as a function of θ_1: $L(\theta_1, \theta_2) = \frac{1}{(\underbrace{\theta_2 + 3}_{\text{constant}} - 2\theta_1)^n}$ as $\theta_1 \uparrow$'s, the denominator $\downarrow$'s, so $L \uparrow$'s.

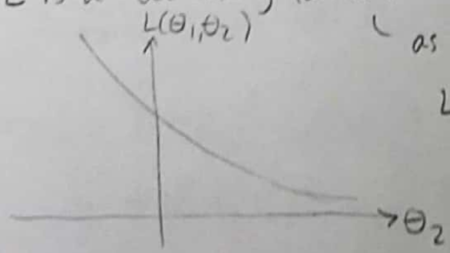
- L is an increasing function



as θ_1 increases, $L(\theta_1, \theta_2)$ increases. Well, the largest that θ_1 is allowed to be is dictated by $2\theta_1 < x_{1:n}$, hence $\theta_1 < \frac{x_{1:n}}{2}$. Thus $\boxed{\hat{\theta}_1 = \frac{x_{1:n}}{2}}$.

$L(\theta_1, \theta_2)$ as a function of θ_2: $L(\theta_1, \theta_2) = \frac{1}{(\theta_2 + \underbrace{3 - 2\theta_1}_{\text{constant}})^n}$ As $\theta_2 \uparrow$'s, the denominator $\uparrow$'s, so $L \downarrow$'s.

- L is a decreasing function



as θ_2 increases, $L(\theta_1, \theta_2)$ decreases. The largest value of L, in this case, occurs when θ_2 is smallest. Well, $x_{n:n} < \theta_2 + 3$, which means $x_{n:n} - 3 < \theta_2$. So θ_2 is smallest when at $x_{n:n} - 3$, and thus $\boxed{\hat{\theta}_2 = x_{n:n} - 3}$.

$\boxed{\text{I.e., the MLE of } \underset{\sim}{\theta} = (\theta_1, \theta_2) \text{ is } \hat{\underset{\sim}{\theta}} = \left(\frac{x_{1:n}}{2}, x_{n:n} - 3\right)}$.

(b) Is the MLE you obtained in (a) unbiased for $\underset{\sim}{\theta}$? Justify.

- Need to check whether $E(\hat{\theta}_1) = E\left(\frac{X_{1:n}}{2}\right) = \theta_1$ and whether $E(\hat{\theta}_2) = E(X_{n:n} - 3) = \theta_2$.
- Need pdfs of $X_{1:n}$ and $X_{n:n}$.
- First need CDF of $F_{X_i}(x)$.

$$F_{X_i}(x) = \int_{2\theta_1}^{x} \frac{1}{\theta_2+3-2\theta_1}\,dt = \frac{1}{\theta_2+3-2\theta_1}(t)\Big|_{2\theta_1}^{x} = \frac{x}{\theta_2+3-2\theta_1} - \frac{2\theta_1}{\theta_2+3-2\theta_1},$$

$2\theta_1 < x < \theta_2+3$

I.e.,

$$F_{X_i}(x) = \begin{cases} 0, & x \le 2\theta_1, \\ \dfrac{x-2\theta_1}{\theta_2+3-2\theta_1}, & 2\theta_1 < x < \theta_2+3 \\ 1, & x \ge \theta_2+3. \end{cases}$$

20/20

- $F_{X_{n:n}}(x_{n:n}) = P(X_{n:n} \le x_{n:n}) = P(\text{all } X_i\text{'s} \le x_{n:n}) = [F_{X_i}(x_{n:n})]^n$

$$= \left(\frac{x_{n:n}-2\theta_1}{\theta_2+3-2\theta_1}\right)^n, \quad 2\theta_1 < x_{n:n} < \theta_2+3,$$

so $f_{X_{n:n}}(x_{n:n}) = \dfrac{d}{dx_{n:n}}\left(\dfrac{x_{n:n}-2\theta_1}{\theta_2+3-2\theta_1}\right)^n$

$$= n\left(\frac{x_{n:n}-2\theta_1}{\theta_2+3-2\theta_1}\right)^{n-1}\left(\frac{1}{\theta_2+3-2\theta_1}\right), \quad 2\theta_1 < x_{n:n} < \theta_2+3$$

$\longrightarrow$ Thus $E(\hat{\theta}_2) = E(X_{n:n}-3) = E(X_{n:n}) - 3$

$$= \int_{2\theta_1}^{\theta_2+3} x_{n:n}\cdot f_{X_{n:n}}(x_{n:n})\,dx_{n:n} - 3$$

$$= \int_{2\theta_1}^{\theta_2+3} x_{n:n}\cdot n\left(\frac{x_{n:n}-2\theta_1}{\theta_2+3-2\theta_1}\right)^{n-1}\left(\frac{1}{\theta_2+3-2\theta_1}\right)dx_{n:n} - 3$$

$$= \frac{n}{\theta_2+3-2\theta_1}\cdot\frac{1}{(\theta_2+3-2\theta_1)^{n-1}}\int_{2\theta_1}^{\theta_2+3} x_{n:n}(x_{n:n}-2\theta_1)^{n-1}\,dx_{n:n} - 3$$

$$= \frac{n}{(\theta_2+3-2\theta_1)^n}\int_{2\theta_1}^{\theta_2+3} x_{n:n}(x_{n:n}-2\theta_1)^{n-1}\,dx_{n:n} - 3 \ldots\ldots$$

cont'd $\rightarrow$

So far, $E(\hat{\theta}_2) = \frac{n}{(\theta_2+3-2\theta_1)^n} \int_{2\theta_1}^{\theta_2+3} x_{n:n}(x_{n:n}-2\theta_1)^{n-1}\,dx_{n:n} - 3.$

let $u = x_{n:n}$, $\int dv = \int (x_{n:n}-2\theta_1)^{n-1}$

$\Rightarrow \frac{du}{dx_{n:n}} = 1 \qquad \Rightarrow v = \frac{(x_{n:n}-2\theta_1)^{n-\cancel{1}+\cancel{1}}}{n-\cancel{1}+\cancel{1}} = \frac{1}{n}(x_{n:n}-2\theta_1)^n$

$\Rightarrow du = dx_{n:n}$

or use u-substitution with $u = x_{n:n} - 2\theta_1$

So...

$$= \frac{n}{(\theta_2+3-2\theta_1)^n}\left[(x_{n:n})\cdot\frac{1}{n}(x_{n:n}-2\theta_1)^n\Big|_{2\theta_1}^{\theta_2+3} - \int_{2\theta_1}^{\theta_2+3}\frac{1}{n}(x_{n:n}-2\theta_1)^n\,dx_{n:n}\right] - 3$$

ask

easier way...?

$$= \frac{\cancel{n}}{(\theta_2+3-2\theta_1)^n}\left[\frac{1}{\cancel{n}}\left((\theta_2+3)(\theta_2+3-2\theta_1)^n - \cancel{(2\theta_1)\cdot(2\theta_1-2\theta_1)^n}^{\,0}\right) - \frac{1}{\cancel{n}}\left(\frac{(x_{n:n}-2\theta_1)^{n+1}}{n+1}\Big|_{2\theta_1}^{\theta_2+3}\right)\right] - 3$$

$$= \frac{1}{(\theta_2+3-2\theta_1)^n}\left[(\theta_2+3)(\theta_2+3-2\theta_1)^n - \frac{1}{n+1}\left((\theta_2+3-2\theta_1)^{n+1} - \cancel{(2\theta_1-2\theta_1)^{n+1}}^{\,0}\right)\right] - 3$$

$$= \frac{1}{(\theta_2+3-2\theta_1)^n}\left[(\theta_2+3)(\theta_2+3-2\theta_1)^n - \frac{1}{(n+1)}(\theta_2+3-2\theta_1)^{n+1}\right] - 3$$

$$= \frac{1}{\cancel{(\theta_2+3-2\theta_1)^n}}\cancel{(\theta_2+3-2\theta_1)^n}\left[\theta_2+\cancel{3} - \frac{(\theta_2+3-2\theta_1)}{n+1}\right] - \cancel{3}$$

ask

$$= \theta_2 - \frac{\theta_2+3-2\theta_1}{n+1} \neq \theta_2.$$

So $\hat{\theta}_2$ is biased for θ_2.

As for $E(\hat{\theta}_1)$, now we need $F_{X_{1:n}}(x_{1:n})$ so that we can get $f_{X_{1:n}}(x_{1:n})$.

$\cdot\, F_{X_{1:n}}(x_{1:n}) = P(X_{1:n} \le x_{1:n}) = 1 - P(X_{1:n} > x_{1:n}) = 1 - P(\text{all } X_i\text{'s} > x_{1:n})$

$$= 1 - [P(X_i > x_{1:n})]^n = 1 - [1 - P(X_i \le x_{1:n})]^n = 1 - [1 - F_{X_i}(x_{1:n})]^n$$

$$= 1 - \left[1 - \left(\frac{x_{1:n}-2\theta_1}{\theta_2+3-2\theta_1}\right)\right]^n = 1 - \left[\frac{\theta_2+3-\cancel{2\theta_1}-x_{1:n}+\cancel{2\theta_1}}{\theta_2+3-2\theta_1}\right]^n$$

$$= 1 - \left[\frac{\theta_2+3-x_{1:n}}{\theta_2+3-2\theta_1}\right]^n, \quad 2\theta_1 < x_{1:n} < \theta_2+3,$$

so $f_{X_{1:n}}(x_{1:n}) = \cancel{0} - n\left(\frac{\theta_2+3-x_{1:n}}{\theta_2+3-2\theta_1}\right)^{n-1}\cdot\left(\frac{-1}{\theta_2+3-2\theta_1}\right)$

$$= \frac{n}{\theta_2+3-2\theta_1}\cdot\left(\frac{\theta_2+3-x_{1:n}}{\theta_2+3-2\theta_1}\right)^{n-1}, \quad 2\theta_1 < x_{1:n} < \theta_2+3$$

cont'd →

So,

$$\longrightarrow E(\hat{\theta}_1) = E\left(\frac{X_{1:n}}{2}\right) = \frac{1}{2}E(X_{1:n}) = \frac{1}{2}\cdot\int_{2\theta_1}^{\theta_2+3} x_{1:n} f_{X_{1:n}}(x_{1:n})\,dx_{1:n}$$

$$= \frac{1}{2}\int_{2\theta_1}^{\theta_2+3} x_{1:n}\frac{n}{(\theta_2+3-2\theta_1)}\cdot\left(\frac{\theta_2+3-x_{1:n}}{\theta_2+3-2\theta_1}\right)^{n-1} dx_{1:n}$$

$$= \frac{n}{2(\theta_2+3-2\theta_1)^n}\int_{2\theta_1}^{\theta_2+3} x_{1:n}(\theta_2+3-x_{1:n})^{n-1}\,dx_{1:n}.$$

$$\left[\begin{array}{ll} \text{Let } u = x_{1:n} & , \text{let } \int dv = \int(\theta_2+3-x_{1:n})^{n-1}dx_{1:n} \\ \Rightarrow du = dx_{1:n} & \Rightarrow v = -\frac{1}{n}(\theta_2+3-x_{1:n})^n \end{array}\right]$$

or use u-substitution with $u = \theta_2+3-X_{1:n}$

$$= \frac{n}{2(\theta_2+3-2\theta_1)^n}\left[(x_{1:n})\left(-\frac{1}{n}(\theta_2+3-x_{1:n})^n\right)\Big|_{2\theta_1}^{\theta_2+3} - \int_{2\theta_1}^{\theta_2+3}\left(-\frac{1}{n}(\theta_2+3-x_{1:n})^n\right)dx_{1:n}\right]$$

$$= \frac{n}{2(\theta_2+3-2\theta_1)^n}\left[-\frac{1}{n}\left((\theta_2+3)\cancelto{0}{(\theta_2+3-(\theta_2+3))^n} - (2\theta_1)(\theta_2+3-2\theta_1)^n\right) + \frac{1}{n}\int_{2\theta_1}^{\theta_2+3}(\theta_2+3-x_{1:n})^n\,dx_{1:n}\right]$$

$$= \frac{\cancel{n}}{2(\theta_2+3-2\theta_1)^n}\left[\frac{1}{\cancel{n}}\left(2\theta_1(\theta_2+3-2\theta_1)^n\right) + \frac{1}{\cancel{n}}\left(-\frac{(\theta_2+3-x_{1:n})^{n+1}}{n+1}\Big|_{2\theta_1}^{\theta_2+3}\right)\right]$$

$$= \frac{1}{2(\theta_2+3-2\theta_1)^n}\left[2\theta_1(\theta_2+3-2\theta_1)^n - \frac{1}{(n+1)}\left(\cancelto{0}{(\theta_2+3-(\theta_2+3))^{n+1}} - (\theta_2+3-(2\theta_1))^{n+1}\right)\right]$$

$$= \frac{1}{2(\theta_2+3-2\theta_1)^n}\left[2\theta_1(\theta_2+3-2\theta_1)^n + \frac{(\theta_2+3-2\theta_1)^{n+1}}{n+1}\right]$$

$$= \frac{1}{2\cancel{(\theta_2+3-2\theta_1)^n}}\cancel{(\theta_2+3-2\theta_1)^n}\left[2\theta_1 + \frac{\theta_2+3-2\theta_1}{n+1}\right]$$

$$= \theta_1 + \frac{\theta_2+3-2\theta_1}{2(n+1)} \neq \theta_2$$

$\boxed{\text{So } \hat{\theta}_1 \text{ is biased for } \theta_1.}$

$$\longrightarrow \boxed{\text{In total, } \hat{\underset{\sim}{\theta}} = (\hat{\theta}_1, \hat{\theta}_2) = \left(\frac{x_{1:n}}{2}, x_{n:n}-3\right) \text{ is biased for } \underset{\sim}{\theta} = (\theta_1, \theta_2)}.$$

89.5%

100 excellent!
/100 (!!)

STAT 480b
Exam 2
March 21, 2022

Name: Matthew

Using a separate sheet of paper for each number and only one side of each sheet, solve the following problems completely and neatly. Use the appropriate notation and if applicable, encircle your final answer. **No solution, no credit.** Good Luck!

1. Suppose $X_1, X_2, \ldots, X_n$ form a random sample from a distribution with pdf
$$f(x;\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta, \quad \theta > 0.$$
   (a) (7 points) Derive the method of moments estimator of θ.
   (b) (7 points) Derive the maximum likelihood estimator of θ.
   (c) (16 points) **(Undergraduate students only:)** Derive the mean and variance of the MME that you obtained in (a). Is the MME an unbiased estimator of θ? If not, make adjustments to make it an unbiased estimator of θ.
   (d) (16 points) **(Graduate students only:)** Is the MLE that you obtained in (b) an unbiased estimator of θ? If not, make adjustments to make it an unbiased estimator of θ.

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, 25)$.
   (a) (10 points) Derive the maximum likelihood estimator of μ.
   (b) (14 points) Is the MLE that you obtained in (a) a UMVUE of μ? Justify your answer.
   (c) (10 points) Derive the maximum likelihood estimator of the 90^{th} percentile of X.
   (d) (10 points) Is the MLE that you obtained in (c) a UMVUE of the 90^{th} percentile of X? Justify your answer.
   (e) (6 points) Determine the asymptotic distribution of the MLE of μ.

3. (20 points) Suppose $Y_1, Y_2, \ldots, Y_n$ form a random sample from a distribution with pdf
$$f(y;\mu) = \frac{1}{\mu}\exp\left(-\frac{y}{\mu}\right), \quad y > 0, \quad \mu > 0.$$
Consider two estimators of μ given by $T_1 = \overline{Y}$ and $T_2 = \frac{n}{n+1}\overline{Y}$. Using the mean squared error (MSE) as a criterion, which estimator will you prefer? Justify your choice.

Matthew Miller Exam # 2

1) Sps $X_1, X_2, \dots, X_n$ form a r.s. from a dist. with pdf

$$f(x;\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta, \quad \theta > 0.$$

(a) To derive the MME of θ, we need only the population and sample moments for $r=1$, as there is only one unknown parameter.

1st pop. moment: $E(X)$, where

$$E(X) = \int_0^\theta x \cdot \frac{2x}{\theta^2}\,dx = \frac{2}{\theta^2}\int_0^\theta x^2 dx = \frac{2}{\theta^2}\cdot\left(\frac{1}{3}x^3\Big|_{x=0}^{x=\theta}\right)$$

$$= \frac{2}{\theta^2}\cdot\frac{1}{3}(\theta^3 - \cancel{0^3}) = \frac{2}{3}\cdot\frac{\theta^3}{\theta^2} = \frac{2}{3}\theta.$$

1st sample moment: $\overline{X}$

So we want the MME $\hat{\theta}$ of θ such that $\frac{2}{3}\hat{\theta} = \overline{X}$.

$$\frac{2}{3}\hat{\theta} = \overline{X} \Rightarrow \boxed{\hat{\theta} = \frac{3}{2}\overline{X}.}$$

7/7

(b) To derive the MLE of θ, we will need an indicator function since θ is in the support.

$$L(\theta) = \prod_{i=1}^n f(x_i;\theta)\cdot I\{0 < x_i < \theta\}$$

$$= \prod_{i=1}^n \frac{2x_i}{\theta^2}\cdot I\{0 < x_i < \theta\}$$

$$= \left(\frac{2}{\theta^2}\right)^n \cdot \prod_{i=1}^n x_i \cdot \prod_{i=1}^n I\{0 < x_i < \theta\}$$

$$= \frac{2^n}{\theta^{2n}}\cdot\prod_{i=1}^n x_i \cdot I\{0 < x_{1:n} < x_{n:n} < \theta\}.$$

Due to the neccessity of an indicator function, we maximize $L(\theta)$ by analyzing the behavior of L as a function of θ.

$$L(\theta) = \frac{2^n}{\theta^{2n}}\cdot\prod_{i=1}^n x_i \quad \leftarrow \text{here, as } \theta\uparrow\text{'s, } \frac{2^n\cdot\prod_{i=1}^n x_i}{\theta^{2n}} \downarrow\text{'s,}$$

meaning L is a decreasing function. The largest value of L is therefore achieved when θ is smallest. We have that $x_{n:n} < \theta$, hence it must be true that $\boxed{\hat{\theta} = X_{n:n}}$ is the estimator that maximizes L.

7/7

1)(d) From part (b), an MLE $\hat{\theta} = X_{n:n}$ was obtained for θ.
To check whether it is an unbiased estimator, we will need to check whether $E(\hat{\theta}) = E(X_{n:n}) = \theta$.

- First, we will need the pdf of $X_{n:n}$. Let $Y = X_{n:n}$.

Then,
$$f_Y(y) = n \cdot [F_X(y)]^{n-1} \cdot f_X(y), \quad 0 < y < \theta,$$

where
$$F_X(y) = \int_0^y \frac{2x}{\theta^2} dx = \frac{2}{\theta^2} \int_0^y x\,dx = \frac{\cancel{2}}{\theta^2}\left(\frac{1}{2}x^2 \Big|_{x=0}^{x=y}\right)$$
$$= \frac{1}{\theta^2} \cdot (y^2 - 0^2) = \frac{1}{\theta^2} y^2$$

16/16

So,
$$f_Y(y) = n\left[\frac{1}{\theta^2}y^2\right]^{n-1} \cdot \frac{2y}{\theta^2} = \frac{2n}{\theta^2} \cdot \frac{1}{(\theta^2)^{(n-1)}} \cdot y^1 \cdot (y^2)^{(n-1)}$$
$$= \frac{2n}{\theta^{2n-2} \cdot \theta^2} \cdot y^{2n-2} \cdot y^1 = \frac{2n}{\theta^{2n-\cancel{2}+\cancel{2}}} \cdot y^{2n-2+1} = \frac{2n}{\theta^{2n}} y^{2n-1},$$
$$0 < y < \theta$$

Thus, we see that
$$E(\hat{\theta}) = E(X_{n:n}) = E(Y) = \int_0^\theta y \cdot \frac{2n}{\theta^{2n}} \cdot y^{2n-1}\, dy$$
$$= \int_0^\theta \frac{2n}{\theta^{2n}} y^{2n}\, dy = \frac{2n}{\theta^{2n}} \cdot \int_0^\theta y^{2n}\, dy = \frac{2n}{\theta^{2n}}\left[\frac{y^{2n+1}}{2n+1}\right]_0^\theta$$
$$= \frac{2n}{\theta^{2n}(2n+1)} \cdot \theta^{2n+1} = \frac{\theta^{\cancel{2n}+1}}{\theta^{\cancel{2n}}} \cdot \frac{2n}{2n+1} = \theta \cdot \frac{2n}{2n+1} \neq \theta.$$

So the MLE obtained in (b) is $\boxed{\text{biased}}$.

→ To adjust this into an unbiased estimator, consider the estimator $\boxed{\frac{2n+1}{2n}\hat{\theta}}$. Then $E\left(\frac{2n+1}{2n}\hat{\theta}\right) = \frac{2n+1}{2n} E(\hat{\theta}) = \frac{\cancel{2n+1}}{\cancel{2n}} \cdot \left(\theta \cdot \frac{\cancel{2n}}{\cancel{2n+1}}\right) = \theta$

$$= \frac{2n+1}{2n} X_{n:n}$$

2) Let $X_1, X_2, \ldots, X_n$ be a r.s. from $N(\mu, \overset{\sigma^2}{25})$.

(a) Since only one parameter is unknown, namely μ, we need ~~&~~ L to be only a function of μ.

· pdf of normal R.V: $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $x \in \mathbb{R}$.

So for our r.s.,

$$f(x;\mu) = \frac{1}{\sqrt{2\pi}(5)} e^{-\frac{1}{2}\frac{(x-\mu)^2}{25}}, \quad x \in \mathbb{R}. \quad \text{Now,}$$

$$L(\mu) = \prod_{i=1}^{n} f(x_i;\mu) = \prod_{i=1}^{n} \left(\frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{50}(x_i-\mu)^2} \right)$$

$$= \frac{1}{(5\sqrt{2\pi})^n} \cdot e^{\sum_{i=1}^{n}\left[-\frac{1}{50}(x_i-\mu)^2\right]} = \frac{1}{(5\sqrt{2\pi})^n} \cdot e^{-\frac{1}{50}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\Rightarrow \ln L(\mu) = \overset{\to 0}{\ln(1)} - n\ln(5\sqrt{2\pi}) - \frac{1}{50}\sum_{i=1}^{n}(x_i-\mu)^2$$

10/10

$$\Rightarrow \frac{d}{d\mu}\ln L(\mu) = -(0) - \frac{1}{50}\frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= -\frac{1}{50}\left[\sum_{i=1}^{n}\frac{d}{d\mu}(x_i-\mu)^2\right] = -\frac{1}{50}\left[\sum_{i=1}^{n}(-2(x_i-\mu))\right]$$

$$= \frac{1}{25}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{25}\left[\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}(\mu)\right] = \frac{1}{25}\sum_{i=1}^{n}x_i - \frac{n\mu}{25}$$

$\Rightarrow$ set $= 0$,

$$\frac{1}{25}\sum_{i=1}^{n}x_i = \frac{n\hat{\mu}}{25} \Rightarrow \sum_{i=1}^{n}x_i = n\hat{\mu} \Rightarrow \hat{\mu} = \frac{\sum_{i=1}^{n}x_i}{n} = \bar{x},$$

Thus the ML estimator would be $\boxed{\hat{\mu} = \bar{X}.}$

(b) We need that $E(\hat{\mu}) = \mu$ and that $Var(\hat{\mu}) = CRLB$.

· First, $E(\hat{\mu}) = E(\bar{X}) = E(X) = \mu$, so $\hat{\mu} = \bar{X}$ is unbiased. ✓

· Second, we need the $Var(\hat{\mu})$ and the CRLB.

· $Var(\hat{\mu}) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{25}{n}$.

cont'd ⟶

2 cont'd (b)

- We have $\hat{\mu} = \bar{X}$ as the estimator for $\tau(\mu) = \mu$.
- So $\tau'(\mu) = \frac{d}{d\mu}(\mu) = 1$.
- $f(X;\mu) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{50}(x-\mu)^2}$, so

$$\ln f(X;\mu) = \cancel{\ln(1)}^{0} - \ln(5\sqrt{2\pi}) - \frac{1}{50}(x-\mu)^2$$
$$= -\ln(5\sqrt{2\pi}) - \frac{1}{50}(x-\mu)^2$$

$$\Rightarrow \frac{d}{d\mu}\ln f(X;\mu) = \cancel{\frac{d}{d\mu}(-\ln(5\sqrt{2\pi}))}^{0} - \frac{1}{50}\frac{d}{d\mu}(x-\mu)^2$$
$$= \frac{1}{25}(x-\mu)$$

$$\Rightarrow \left[\frac{d}{d\mu}\ln f(X;\mu)\right]^2 = \left(\frac{1}{25}(x-\mu)\right)^2 = \frac{1}{625}(x-\mu)^2$$

$$\Rightarrow E\left(\left[\frac{d}{d\mu}\ln f(X;\mu)\right]^2\right) = E\left(\frac{1}{625}(x-\mu)^2\right) = \frac{1}{625}E(x-\mu)^2$$
$$= \frac{1}{625}Var(X) = \frac{1}{625}(25) = \frac{1}{25}.$$

14/14

Altogether,

$$CRLB = \frac{[\tau'(\mu)]^2}{n\,E\left(\left[\frac{d}{d\mu}\ln f(X;\mu)\right]^2\right)} = \frac{(1)^2}{n\left(\frac{1}{25}\right)} = \frac{1}{\frac{n}{25}} = \frac{25}{n}.$$

We have that $Var(\hat{\mu}) = \frac{25}{n} = CRLB$, and that $\hat{\mu}$ is unbiased,

$\boxed{\text{hence } \hat{\mu} = \bar{X} \text{ is a UMVUE of } \mu.}$

(c) The 90^{th} percentile of the R.V. X is the number c such that $P(X<c) = 0.90$. Here, $P(X<c) = P\left(\frac{X-\mu}{5} < \frac{c-\mu}{5}\right) = P\left(Z < \frac{c-\mu}{5}\right)$, so we have $P\left(Z < \frac{c-\mu}{5}\right) = 0.90$, or $\Phi\left(\frac{c-\mu}{5}\right) = 0.90$.

10/10

Using Table 3, this means that $\frac{c-\mu}{5} = 1.282$, so $c-\mu = 5(1.282)$

$\Rightarrow c = 6.41 + \mu$

cont'd →

cont'd (c)

So we have that $c = \tau(\mu) = 6.41 + \mu$, which is a linear function of μ. From (b) we know that there exists an unbiased estimator of $\tau(\mu)$ whose variance will equal its CRLB, since $\hat{\mu} = \bar{X}$ is UMVUE of μ.

So, $\boxed{\hat{\tau}(\mu) = \tau(\hat{\mu}) = 6.41 + \bar{X}}$ is the resulting ML estimator for $c = \tau(\mu) = 6.41 + \mu$.

For (d), didn't need to solve for it, could have just cited that we can obtain UMVUE for sure.

(d) First, we check that $E(\hat{\tau}(\mu)) = \tau(\mu)$:

$$E(\hat{\tau}(\mu)) = E(\tau(\hat{\mu})) = E(6.41 + \bar{X}) = 6.41 + E(\bar{X})$$
$$= 6.41 + E(X) = 6.41 + \mu = \tau(\mu) = c,$$

so $\hat{\tau}(\mu) = 6.41 + \bar{X}$ is unbiased for $\tau(\mu) = 6.41 + \mu$.

10/10

Now,

$$Var(\hat{\tau}(\mu)) = Var(6.41 + \bar{X}) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{25}{n};$$

$$CRLB = \frac{[\tau'(\mu)]^2}{nE([\frac{d}{d\mu}\ln f(X;\mu)]^2)},$$ where from part (b) we already know that $nE(\frac{d}{d\mu}\ln f(X;\mu))^2 = \frac{1}{25}$. Now, $\tau'(\mu) = \frac{d}{d\mu}(6.41+\mu) = 1$,

so we find that the $CRLB = \frac{(1)^2}{n/25} = \frac{25}{n} = Var(\hat{\tau}(\mu))$.

$\therefore$ the MLE $\hat{\tau}(\mu) = 6.41 + \bar{X}$ is a UMVUE of $c = \tau(\mu) = 6.41 + \mu$,

(since $\hat{\tau}(\mu)$ is unbiased for $\tau(\mu)$ and since $Var(\hat{\tau}(\mu)) = CRLB$)

(e) We have that the MLE of μ is $\hat{\mu} = \bar{X}$, which was found under satisfactory conditions (i.e., no indicator functions involved). Thus,

6/6

$$\hat{\mu}_n \xrightarrow{d} \hat{\mu} \sim N(\mu, CRLB), \text{ where } CRLB = \frac{25}{n}.$$

$$\boxed{\text{That is, } \hat{\mu} \sim AN(\mu, \tfrac{25}{n}).}$$

5) Sps $Y_1, Y_2, \cdots, Y_n$ form a r.s. from a dist. w/ pdf

$$f(y;\mu) = \frac{1}{\mu} e^{(-y/\mu)}, \quad y > 0, \mu > 0.$$

Let $\underline{T_1 = \bar{Y}}$ and $\underline{T_2 = \frac{n}{n+1}\bar{Y}}$ each be estimators of μ.

· $MSE(T_1) = Var(T_1) + [b(T_1)]^2$, where $b(T_1) = E(T_1) - \mu$.

- note that the given pdf suggests $Y_1, Y_2, \ldots, Y_n \sim EXP(\mu)$,

so $E(Y) = \mu$ and $Var(Y) = \mu^2$.

· Continuing,

$$MSE(T_1) = Var(T_1) + [b(T_1)]^2 = Var(\bar{Y}) + [E(T_1) - \mu]^2$$

$$= \frac{Var(Y)}{n} + [E(\bar{Y}) - \mu]^2 = \frac{\mu^2}{n} + [E(Y) - \mu]^2$$

$$= \frac{\mu^2}{n} + [\cancel{\mu - \mu}]^{2 \to 0} = \frac{\mu^2}{n}$$

($\bar{Y}$ unbiased)

20/20

$$MSE(T_2) = Var(T_2) + [b(T_2)]^2 = Var\left(\frac{n}{n+1}\bar{Y}\right) + [E(T_2) - \mu]^2$$

$$= \frac{n^2}{(n+1)^2} \frac{Var(Y)}{n} + \left[E\left(\frac{n}{n+1}\bar{Y}\right) - \mu\right]^2$$

$$= \frac{n^2}{(n+1)^2} \frac{\mu^2}{n} + \left[\frac{n}{n+1} E(Y) - \mu\right]^2 = \frac{n^2}{(n+1)^2} \frac{\mu^2}{n} + \left[\frac{n}{n+1}\mu - \mu\right]^2$$

$$= \frac{n^2\mu^2}{(n+1)^2} \cdot \frac{1}{n} + \left[\frac{\cancel{n\mu} - \cancel{n\mu} - \mu}{n+1}\right]^2 = \frac{n\mu^2}{(n+1)^2} + \frac{\mu^2}{(n+1)^2}$$

$$= \frac{\mu^2(n+1)}{(n+1)^2} = \frac{\mu^2}{n+1}$$

	$MSE(T_1)$		$MSE(T_2)$
$n=1$	μ^2	$>$	$\mu^2/2$
$n=2$	$\mu^2/2$	$>$	$\mu^2/3$
$n=3$	$\mu^2/3$	$>$	$\mu^2/4$
			⋮

Based on MSE alone, T_2 would be the better estimator; for all $n \in \mathbb{N}$, $MSE(T_2) < MSE(T_1)$, making T_2 preferrable (even though T_1 happens to be unbiased).

$f(x;\theta) = \frac{2x}{\theta^2}$, $0 < x < \theta$, $(\theta > 0)$.

(Scratch Work)

$0 < x < \theta$, i.e. $0 < x_{n:n} < \theta$

(all x_i's are between 0 and θ, so the largest x_i, $x_{n:n}$, is also between 0 and θ. $Y = X_{n:n}$, so sprt is $0 < y < \theta$.

$$f_Y(y) = n\left[\frac{1}{\theta^2}y^2\right]^{n-1} \cdot \frac{2y}{\theta^2}$$

$$= n \cdot \left(\frac{1}{\theta^2}\right)^{n-1} \cdot (y^2)^{n-1} \cdot \frac{2}{\theta^2} \cdot y^1$$

$$= \frac{n \cdot 2}{(\theta^2)^{n-1} \cdot (\theta^2)^1} \cdot y^{2(n-1)} \cdot y^1 = \frac{2n}{\theta^{2n-2} \cdot \theta^2} \cdot y^{2n-2+1}$$

$$= \frac{2n}{\theta^{2n-\cancel{2}+\cancel{2}}} \cdot y^{2n-1} = \frac{2n}{\theta^{2n}} \cdot y^{2n-1}$$

$$\frac{2n}{\theta^{2n}} \int_0^\theta y^{2n}\, dy$$

$$= \frac{2n}{\theta^{2n}}\left[\frac{y^{2n+1}}{2n+1}\right]$$

$$\ln L(\mu) = -n\ln(5\sqrt{2\pi}) - \frac{1}{50}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{d}{d\mu}\ln L(\mu) = (0) - \frac{1}{50}\frac{d}{d\mu}\left(\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

let $w = x_i - \mu$.
then do

$$\frac{d}{dw}(w)^2 \cdot \frac{d}{d\mu}(x_i - \mu)$$

$$= 2w \cdot (-1)$$

$$= 2(x_i - \mu)(-1) = -2(x_i - \mu)$$

$$= -\frac{1}{50}\left[\sum_{i=1}^{n}\frac{d}{d\mu}(x_i-\mu)^2\right]$$

$$= -\frac{1}{50}\left[\sum_{i=1}^{n}(-2)(x_i-\mu)\right]$$

$$\cancel{= \frac{1}{50}} = \frac{1}{25}\sum_{i=1}^{n}(x_i - \mu)$$

$$= \frac{1}{25}\left[x_1 - \mu + x_2 - \mu + \dots + x_n - \mu\right]$$

$$= \frac{1}{25}\left[x_1 + x_2 + \dots + x_n + (-\mu) + (-\mu) + \dots + (-\mu)\right]$$

$$= \frac{1}{25}\left[\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}(-\mu)\right]$$

$$= \frac{1}{25}\left[\sum_{i=1}^{n}x_i - n\mu\right]$$

$$\frac{\mu^2(n^2+1)}{(n+1)^2} = \frac{\mu^2 n^2 + \mu^2}{n^2 + 2n + 1}$$

$$\frac{\mu^2(2)}{4} = \frac{\mu^2}{2} \qquad \frac{\mu^2(5)}{9}$$

$$\frac{\mu^2}{2} = \frac{9\mu^2}{18} \qquad \frac{5\mu^2}{9} = \frac{10\mu^2}{18}$$

$$\frac{\mu^2(10)}{\cdot} = \frac{30}{4\pi}\mu^2 \text{ vs } \frac{16\mu^2}{\cdot}$$

CamScanner

90^{th} percentile of R.V. X is the number c such that $P(X<c) = 0.90$.

$X \sim N(\mu, 25)$, so to get c we'd need first convert to stan. normal:

$$Z = \frac{X-\mu}{\sigma} = \frac{X-\mu}{5}$$

so $$P(X<c) = P\left(\frac{X-\mu}{5} < \frac{c-\mu}{5}\right) = 0.90$$

$$= P\left(Z < \frac{c-\mu}{5}\right) = 0.90$$

$$= \underbrace{\Phi\left(\frac{c-\mu}{5}\right)}_{\text{fcn of } \mu} = 0.90$$

$\hat{\mu} = \bar{X}$ is unbiased MLE of μ, also its CRLB = its variance.

If have $\tau(\mu)$ a lin. fcn of μ, then by Theorem, it follows that

Scratch paper