SC482: FINAL

Name: HUAN Q. BUI

May 15, 2020

Question 1 $\quad X_i \sim Exp(\theta)$

$$f(x|\theta) = \frac{1}{\theta} e^{-x/\theta} \quad ; \quad x > 0.$$

(A) Find mle...

$$\mathcal{L}(\theta) = \frac{1}{\theta^n} e^{-\Sigma x_i/\theta}$$

$$l(\theta) = \ln \mathcal{L}(\theta) = -n \ln \theta - \frac{1}{\theta} \Sigma x_i$$

$$\partial_\theta l(\theta) = 0 \Rightarrow \frac{-n}{\cancel{\theta}} + \frac{1}{\theta^{\cancel{2}}} \Sigma x_i = 0$$

$$\Rightarrow \hat{\theta} = \Sigma x_i / n \Rightarrow \boxed{\hat{\theta} = \bar{X}}$$

(B) Is $\bar{X}$ efficient?

- $E[\hat{\theta}] = E[\bar{X}] = \frac{1}{n} E[\Sigma x_i] = \frac{1}{n} \Sigma E[x_i]$

$$= \frac{1}{n} \cdot n \cdot \theta = \theta \Rightarrow \hat{\theta} \text{ unbiased}.$$

- $Var(\hat{\theta}) = \frac{1}{n^2} Var(\Sigma x_i) = \frac{1}{n^2} \Sigma Var(x_i) = \frac{n\theta^2}{n^2} = \frac{\theta^2}{n}.$

- $I(\theta) = n I(\theta) = Var(\partial_\theta l(\theta)) = Var\left[\frac{-n}{\theta} + \frac{1}{\theta^2} \Sigma x_i\right]$

$$= Var\left[\frac{1}{\theta^2} \Sigma x_i\right] = \frac{1}{\theta^4} \Sigma Var(x_i) = \frac{1}{\theta^4} \cdot n\theta^2 = \frac{n}{\theta^2}$$

$\Rightarrow \underbrace{Var(\hat{\theta}) = \frac{1}{I(\theta)}}_{\text{(CRLB)}}, \ \hat{\theta} \text{ unbiased} \Rightarrow \boxed{\hat{\theta} \text{ is efficient}}$

(c) Find MVUE for θ

- Hence $\frac{\sum X_i}{n}$ as ~~suffic~~ unbiased estimator for θ.

- Now, by factorization theorem --

$$\mathcal{L}(x|\theta) = \frac{1}{\theta^n} \exp\left[-\frac{1}{\theta} \sum x_i\right]$$

$\sum x_i$ is a sufficient statistic for θ.

- Because $\bar{X}$ is a ~~function~~ function of the sufficient statistic $\sum X_i$ and $\bar{X}$ is an unbiased estimator for θ.

$\Rightarrow$ Rao - Blackwell says $\bar{X}$ is the MVUE for θ.

Alternatively, since $\sum X_i$ is both sufficient & complete (pdf is a member of the scalar exp class) and $\frac{\sum X_i}{n} = \bar{X}$ is an unbiased estimator for θ.

$\Rightarrow$ $\bar{X}$ is the MVUE for θ.

Question 2 $X_i \sim \text{Ray}(\theta)$

$$f(x) = \frac{2x}{\theta} e^{-x^2/\theta} \quad ; \quad x > 0$$

(A) <u>Find sufficient statistic for θ</u>

$$\mathcal{L}(\theta) = \left(\frac{2}{\theta}\right)^n \left(\prod^n x_i\right) \exp\left\{\frac{-1}{\theta} \sum x_i^2\right\}$$

$$\begin{cases} K_1(Y_1, \theta) = \left(\frac{2}{\theta}\right)^n \exp\left\{\frac{-1}{\theta} \sum x_i^2\right\} \\ K_2(x_i \ldots) = \prod^n x_i \end{cases}$$

by factorization theorem $\Rightarrow$ $\boxed{Y_1 = \sum^n x_i^2}$ is one sufficient statistic.

(B) <u>MVUE for θ</u>

$$E[Y_1] = E\left[\sum^n x_i^2\right] = \sum E[x_i^2] = n E[x_i^2]$$

$$E[x_i^2] = \int_0^\infty \frac{2x^3}{\theta} e^{-x^2/\theta} dx \overset{x^2 = u}{=\!=} \int_0^\infty \frac{1}{\theta} u e^{-u/\theta} du = \theta$$

$$= \theta \qquad (= \text{Exp}[u] \; ; \; u \sim \text{Exp}(\theta))$$

$$\Rightarrow E[Y_1] = n\theta \Rightarrow \boxed{\hat{\theta} = \frac{1}{n} \sum x_i^2} \Rightarrow E[\hat{\theta}] = \theta \Rightarrow \boxed{\frac{1}{n} \sum x_i^2} \text{ is the MVUE}$$

~~unbias~~ unbiased est. for θ

$\hat{\theta}$ unbiased, and a function of sufficient Y_1

(c) <u>To show that the MVUE is unique...</u>

We know that

$Y_1 = \sum x_i^2$ is a sufficient statistic for θ.

and $\hat{\theta} = \frac{1}{n} Y_1 = \frac{1}{n} \sum x_i^2$ is a function of Y_1 and

and unbiased estimator for θ.

$\hat{\theta}$ is the unique MVUE for θ <u>if</u> $\{ f_{Y_1}(y_1; \theta) : \theta \in \Omega \}$

is <u>complete</u>.

(Lehmann - Scheffé)

$\boxed{\text{i.e.}}$ we want to show that ~~it~~ with $\{ f_{Y_1}(y_1; \theta) \}$

if $E(u(Y_1)) = 0 \quad \forall \, \theta \in \Omega$

then $u(y_1) \equiv 0$ except on a set of points that

has probability zero for each $f_{Y_1}(y; \theta)$ in the

family.

$\square$

$\boxed{\text{Question 3}}$ Y is a single observation...

$$f(y|\theta) = \theta y^{\theta-1} \quad ; \quad 0 \leq y \leq 1.$$

4 Find most powerful test of $H_0: \theta = 1$
$H_A: \theta = 2$

Find form of rejection region ($\alpha = 0.05$)

Find specific values of test statistic for which H_0 is rejected

Well...

$$\frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_A)} = \frac{1}{2y} < K \Rightarrow y > K'$$

with $\theta_0 \to 1$, $\theta_A \to 2$

rejection region has the form: $\boxed{C = \{y \in [0,1] ; y \geq c\}}$

for some constant c.

~~with~~ Integrating under the null... $\theta_0 = 1 \Rightarrow f(y|\theta) = 1$

$$\int_{K'}^{1} 1 \, dy = 0.05 \Rightarrow K' = 1 - 0.05 = \boxed{0.95}$$

$\Rightarrow$ We reject if $\boxed{y > 0.95 \Rightarrow C_{\alpha=0.05} = \{y \in [0,1] ; y > 0.95\}}$

This is the most powerful test for $H_0: \theta = 1$ vs. $H_A: \theta = 2$

(B) Is the test UMP for $H_a: \theta > 1$?

~~Obviously, the form of the rejection region DEPENDS on the specific value of the alternative parameter θ_A~~

~~When~~

As long as $\theta_A > 1$, the form of the rejection region will not depend on the specific value of θ_A because

$$\frac{L(\theta_0)}{L(\theta_A)} = \frac{1}{\theta_A y^{\theta_A - 1}} \sim \frac{1}{y^T} < k \text{ where } T > 0$$

$\Rightarrow$ still reject if $y > k'$.

$\Rightarrow$ $\boxed{\text{YES, the test is still UMP}}$

(c) Find power for $H_0: \theta_0 = 1$, $H_A: \theta_A = 2$

$$\text{Power} = P(y > 0.95 \mid \theta = 2) = \int_{0.95}^{1} 2y\,dy = y^2 \Big|_{0.95}^{1} =$$

$$= 1 - 0.95^2 = \boxed{\cancel{}}$$

$$= 0.0975$$

D **Likelihood ratio test** $H_0: \theta_0 = 1$; $H_A: \theta_A > 1$.

$$f(y|\theta) = \theta y^{\theta - 1}, \quad 0 \le y \le 1$$

Null space : $\{\theta : \theta = 1\}$

Alternative space: $\{\theta : \theta > 1\}$

Find the mle for θ:

$$\mathcal{L}(\theta) = f(y|\theta) = \theta y^{\theta-1}$$

$\rightarrow \ln(\mathcal{L}(\theta)) = \ell(\theta) = \ln\theta + (\theta - 1)\ln y$

$\rightarrow \partial_\theta \ell(\theta) = \frac{1}{\theta} + \ln y = 0 \Leftrightarrow \boxed{\hat{\theta} = \frac{-1}{\ln(y)}}$

$$\rightarrow \boxed{\Lambda = \frac{\mathcal{L}(\hat{\theta}_0)}{\mathcal{L}(\hat{\theta}_A)} = \frac{\theta_0 y^{\theta_0 - 1}}{\hat{\theta}_A y^{\hat{\theta}_A - 1}} = \frac{1}{\frac{-1}{\ln(y)} \cdot y^{-1/\ln y - 1}}}$$

Asymptotically ... $-2\ln\Lambda \sim \chi^2(1)$

$\rightarrow$ reject if $\boxed{-2\ln\Lambda > \chi^2_{1,\,0.05} = 3,84.}$

~~table~~ $\Leftrightarrow$ if $\Lambda < 0.14667$. Let $z = \ln y$

$\rightarrow \Lambda = \frac{-z}{(e^z)^{-1/z - 1}} = \frac{-z}{e^{-1-z}} < 0.14667 \Leftrightarrow$ this is a transcendent al-equation

... I can't find the values for y with which we reject H_0 ~~[illegible]~~ b/c this is transcendental

7 bis

... we reject if $-2\ln\Lambda > \chi^2_{1,\,0.05} = 3.84$

this means if $\Lambda < 0.14667$.

now...

$$\Lambda = \frac{1}{\frac{-1}{\ln y}\, y^{-1/\ln y - 1}} = \frac{-\ln y}{1} \cdot y^{\left(1/\ln y + 1\right)}$$

we want to find y such that $\Lambda < 0.14667$.

$(0 \leq y \leq 1)$

to do this, we can ask mathematica:

to find the intersection of the graphs $\Lambda(y)$ and the line $y = 0.14667$.

From there, we can find where $\Lambda(y) < 0.14667$.

...

$\Rightarrow$ This is a transcendental equation, so we can't solve this by hand.

$\boxed{\text{Question 4}}$ $X \sim Poi(\lambda)$

$$p(x|\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad ; \quad x = 0, 1, 2, \dots$$

(A) Exp class ...

$$p(x|\lambda) = \exp\left\{ -\lambda + x \ln \lambda - \ln(X!) \right\} \; ; \; \begin{matrix} x \in \mathbb{N}_+ \\ \lambda \in \mathbb{R} \cup \infty \end{matrix}$$

$\Downarrow$

$p \in$ Exponential class of dist.

$$\begin{cases} p(\theta) = p(\lambda) = + \ln(\lambda) \\ k(x) = X \\ H(x) = -\ln(X!) \\ q(\theta) = q(\lambda) = -\lambda \end{cases}$$

(B) Show that the complete sufficient statistic for this dist also belongs to the exp family...

- Note that $Poi(\lambda) \in$ regular exponential class of pmf

$\Rightarrow$ the statistic $Y_1 = \sum K(x_i) = X$ is a complete sufficient statistic for λ. (Theorem)

- If we have a sample of iid $X_i \sim Poi(\lambda)$ then $Y_1 = \sum^{n} X_i$ is a complete sufficient statistic for λ. (Theorem).

$\rightarrow$ We know that $\boxed{\sum X_i \sim Poi(\sum \lambda) = Poi(n\lambda)}$. So Y_1 is also a member of the exponential family, by a similar argument (to (A)). $\square$

Explicitly ...

$$p(Y = \cancel{x_i} \mid \lambda) = \frac{e^{-n\lambda}(n\lambda)^{y}}{y!}$$

$$= \exp\left\{-n\lambda + y\ln(n\lambda) - \ln(y!)\right\}$$

$$y \in \mathbb{N}\ ; \quad n\lambda \in \mathbb{R}^{+} \cup \infty$$

$$\begin{cases} p(n\lambda) = \ln(n\lambda) \\ k(y) = y \\ H(y) = -\ln(y!) \\ q(n\lambda) = -n\lambda \end{cases}$$

$\rightarrow$ so, see that $p_Y(y \mid n\lambda)$ is also a member of the exponential class ...

Question 5 $X_1, \dots X_n$; $X_i \sim Uni(0, \theta)$

(A) MLE for $E[X_1]$ & $Var(X_1)$

$$E[X_1] = \frac{\theta}{2} \quad ; \quad Var[X_1] = \frac{\theta^2}{12}.$$

→ to find mle for θ ⇒ need mle for θ!

we know that $\hat{\theta} = \max_i (X_i) = Y_n$

so, by invariance property of mle...

$$\widehat{E[X_1]} = \frac{\hat{\theta}}{2} = \frac{\max(X_i)}{2} = \boxed{\frac{Y_n}{2}}$$

$$\widehat{Var[X_i]} = \frac{\hat{\theta}^2}{12} = \frac{\max^2(X_i)}{12} = \boxed{\frac{Y_n^2}{12}}$$

(B) Find minimal sufficient statistic for θ

$$f(x_1 \dots x_n) = \frac{1}{\theta^n} I(X > 0) I(x < \theta)$$

By factorization...
$$\begin{cases} K_1(Y_n, \theta) = \frac{1}{\theta^n} I(Y_n < \theta) \\ K_2[X_1 \dots X_n] = I[X_{(1)} > 0] \end{cases}$$

We see that $Y_{(n)}$ (the max) is sufficient for θ.

→ since we can no longer reduce... → Yn is minimal sufficient for θ

(c) <u>Show... if $\theta > 0$ then Y_n is complete...</u>

We have $g_{(n)}(y) = n y^{n-1} \theta^{-n}$; $0 \le y \le \theta$

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

Suppose $u(y)$ is a function s.t. $E(u(Y)) = 0$.

then

$$E[u(y)] = \int_0^\theta u(y)\, n y^{n-1} \theta^{-n} \, dt = 0$$

taking ∂_θ ... we get

$$0 = \underbrace{u(\theta)}\, \underbrace{n \cdot \theta^{n-1} \cdot \theta^{-n}}_{\neq 0} = 0$$ (fundamental theorem of calc)

$\Rightarrow u(\theta) = 0$ identically $\forall\, \theta > 0 \Rightarrow \forall\, y$ too

So, the family of $f_{Y_{(n)}}(y)$ is complete.

$\Rightarrow$ $\boxed{Y_{(n)} \text{ is a complete (minimal) sufficient statistic for } \theta.}$

D Show: if $\theta > 1$ then Y_n is not complete

if $\theta > 1$, then by a similar argument we will get

$$0 = u(\theta)\, n\theta^{n-1}\theta^{-n} = (u(\theta)) \cdot \frac{n}{\theta} = 0 \;\forall \theta.$$

Now, $\theta > 1 \rightarrow u(\theta) = 0 \;\forall\; \theta > 1$ only.

$\Rightarrow$ The condition $E(u(y)) = 0$ hence only requires

$u(y) = 0 \;\forall\; y > 1$ only, not $\forall y \in \text{Supp}\{f\}$.

$\Rightarrow$ $\boxed{Y_n \text{ is not (necessarily) complete.}}$

Question 6

A True

B True ... (mle's are asymptotically efficient)

C False

D True

E True

F False (e.g. $U(0, \theta)$ ... mle is good but regular conditions are not satisfied).

G True

Name: ______________________________

Statistics 482 Spring 2020
Final

15 May 2020

- This is an open-book, open-note, closed-internet exam. I have also provided you with properties of common distributions.
- Do not use Wolfram Alpha to obtain integrals. Enough work must be shown for me to tell that the answer was not just copied.
- All work must be your own. You may not give or receive any kind of aid, either verbally, visually, or otherwise, during this exam. No other sources may be consulted, except as specified above.
- **The exam has 100 possible points. There are 6 questions and 11 pages, including this cover page. You have 3 hours to complete the exam so plan your time accordingly. I have included the possible points next to each problem.**
- Some questions are more difficult than others, and the questions may not be in order of difficulty. Don't spend too much time on any one question; if you get stuck, go on and try another part.
- Whenever possible, show your work and explain your reasoning. In case you make a mistake, I can more easily give you partial credit if you explain your steps.
- Some parts of a question may require the answer to an earlier part of the question. If you can't solve the earlier part, you can still receive partial credit for the latter parts: make up a reasonable answer for the earlier part and use that in solving for the latter parts.
- **Please upload your answers to Moodle in a single document with filename LASTNAME_482_Final.pdf.**

Question 1 (18 points total)

Suppose that X_1, X_2,…,X_n are independent and identically distributed from a distribution with density function,

$$f(x\mid\theta)=\begin{cases}\left(\dfrac{1}{\theta}\right)e^{-x/\theta}; & x>0\\ 0; & \text{elsewhere}\end{cases}$$

where θ is the unknown parameter.

A. (4 points) Find the maximum likelihood estimator of the parameter, θ.

B. (8 points) Is the maximum likelihood estimator in (A) efficient? Show or explain.

Question 1 continued…

C. (6 points) Now find the MVUE for θ.

Question 2 (10 points total)
Consider a random sample, X_1, X_2,…, X_n, from a Rayleigh distribution with density function,

$$f(x) = \begin{cases} \left(\dfrac{2x}{\theta}\right)e^{-x^2/\theta}; & x > 0 \\ 0; & \text{elsewhere} \end{cases}$$

A. (3 points) Find a sufficient statistic for θ.

B. (5 points) Using the sufficient statistic in (A), find the MVUE for θ. Hint: Start by taking the expectation of the sufficient statistic.

C. (2 points) What would you need to show in order to prove that the MVUE you found in (B) is unique?

Question 3 (26 points total)
Suppose *Y* is a random sample of size 1 (a single random observation) from a population with distribution,

$$f(y|\theta) = \begin{cases} \theta y^{\theta-1}; & 0 \leq y \leq 1 \\ 0; & elsewhere \end{cases}$$

A. (8 points) Find a most powerful test of $H_0: \theta = 1\ vs. H_A: \theta = 2$. Find the form of the rejection region, we well as making sure to show your work. Also find the specific values of your test statistic for which the null hypothesis would be rejected if $\alpha = 0.05$.

B. (2 points) Is the test in (A) uniformly most powerful for alternatives of the form, $\theta > 1$? Show or explain.

Question 3 continued…

C. (6 points) Find the power for the test in (A).

D. (10 points) Find the likelihood ratio test for $H_0\colon \theta = 1\ vs.\ H_A\colon \theta > 1$. Make sure to find the values of the test statistic for which you would reject if $\alpha = 0.05$.

Question 4 (12 points total)

Consider a random variable, X, with a Poisson(λ) distribution with probability function,

$p(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$; x = 0, 1, 2,…

A. (4 points) Show that the random variable *X* belongs to the exponential class of distributions. Specify the functions: $p(\theta), k(x), H(x), \text{and } q(\theta)$.

B. (8 points) Now show that the complete sufficient statistic for this distribution also belongs to the exponential family.

Question 5 (20 points total)

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables from a Uniform(0, θ) distribution.

A. **(5 points)** Find the maximum likelihood estimator for both $E[X_1]$ and $Var[X_1]$.

B. **(3 points)** Find a minimal sufficient statistic, *Y*, for θ.

C. **(5 points)** Show that, if $\theta > 0$, then the sufficient statistic you found in (B) is complete. Note that the density function for the max is given by,

$$\boldsymbol{g_{(n)}(y) = \left(\frac{n}{\theta^n}\right) y^{n-1}; \quad 0 \le y \le \theta.}$$

Question 5 continued…

D. (5 points) Show that if $\theta > 1$, then the sufficient statistic you found in (B) is *not* complete.

Question 6 (14 points total)
For each of the following statements, say whether the statement is true or false.

A. A minimax loss function helps you to find the decision that provides you with the best "worst-case" scenario.

True **False**

B. The maximum likelihood estimator will have the smallest variance of all estimators if the sample size is large.

True **False**

C. A minimal statistic is always unique.

True **False**

D. An ancillary statistic, by itself, provides us with no information about the unknown parameter.

True **False**

E. A Wald test evaluates the distance from the null value, θ_0, to the value of the estimate of θ that maximizes the likelihood function.

True **False**

F. Maximum likelihood estimation provides you with a valid estimate only if regularity conditions are satisfied.

True **False**

G. Likelihood, score, and Wald tests are asymptotically equivalent.

True **False**