Probability Theory and Statistics 3, Assignment 4

Jiyeon Kim [*], Younjoo Mo [†]

November 20, 2020

Exercise 1

a.

$$f(x;\theta) = \frac{\theta}{x^2} I_{(\theta,\infty)}(x) \quad (0 < \theta < \infty) \tag{1}$$

$$f(x_1,\ldots,x_n;\theta) \;=\; \frac{\theta^n}{\prod_n^{i=1} x_1^2} \prod_n^{i=1} I_{(\theta,\infty)}(x_i) \tag{2}$$

$$= \frac{\theta^n}{\prod_n^{i=1} x_1^2} I_{(\theta,\infty)}(x_{1:n}) \tag{3}$$

$\prod_n^{i=1} I_{(\theta,\infty)}$ assumes value 1 if and only if in certain domain.

$\therefore$ we could write $\prod_n^{i=1} I_{(\theta,\infty)}(x_i) = I_{(\theta,\infty)}(x_{1:n})$

Next, let's divide $f(x;\theta)$ into two parts(in order to use factorisation criterion)

$$f(x_1,\ldots,x_n;\theta) = g(x_{1:n};\theta)h(x_1,\ldots,x_n) \tag{4}$$

$$\Longrightarrow g(x_{1:n};\theta) \;=\; \theta^n I_{(\theta,\infty)}(x_{1:n}) \tag{5}$$

$$= h(\vec{x}) = \frac{1}{\prod_n^{i=1} x_i^2} \tag{6}$$

* Student number: 12470236
† Studentnumber: 12475440

Random variable $X_1, \ldots, X_n$ has joint pdf $f(x_1, \ldots, x_n; \theta)$

In (5),(6)th line, $g(x_{1:n}; \theta)$ doesn't depend on $x_1, \ldots, x_n$ except through s ,while $h(\vec{x})$ doesn't depend on θ

Therefore we can conclude that by theorem 10.2.1(Factorisation criterion), $x_{1:n}$ is sufficient for θ

Next, to check whether $x_{1:n}$ is complete , we have to calculate the pdf first

$$F_X(x) = \int_\theta^x f(z;\theta)dz \tag{7}$$
$$= \int_\theta^x \frac{\theta}{z^2}dz \tag{8}$$
$$= \theta(1/\theta - 1/x), \quad (x > 0) \tag{9}$$
$$= 1 - \frac{\theta}{x}, \quad (x > 0) \tag{10}$$

So that the CDF of S is as follows.

$$\Longrightarrow F_S(s) \; = \; 1 - (\frac{\theta}{s})^n, \quad (s > 0) \tag{11}$$
$$f_S(s) \; = \; \frac{n\theta^n}{s^{n+1}}, \quad (0 < s < \infty) \tag{12}$$

We use definition 10.4.1

$$
\begin{align}
& E[u(s)] = 0 \quad (\forall s > 0) \tag{13} \\
& \Longrightarrow \int_{\theta}^{\infty} u(s) n \theta^{n} s^{-n-1} ds = 0, \quad (\forall s > 0) \tag{14} \\
& \Longrightarrow \int_{\theta}^{\infty} u(s) s^{-n-1} ds = 0, \quad (\forall s > 0) \tag{15} \\
& \Longrightarrow \frac{d}{d\theta} \int_{\theta}^{\infty} u(s) s^{-n-1} ds = \frac{d}{d\theta} 0 \quad , \quad (\forall s > 0) \tag{16} \\
& \Longrightarrow u(\theta) \theta^{-n-1} = 0, \quad (\forall \theta > 0) \tag{17} \\
& \Longrightarrow u(\theta) = 0, \quad (\forall \theta > 0) \tag{18} \\
& \Longrightarrow u(s) = 0, \quad (\forall \theta > 0) \tag{19}
\end{align}
$$

we proved that sufficient statistic we calculated $x_{1:n}$ is also complete.

b.

From sub-question (b) we calculated the following equation

$$f_S(s) \;=\; n\frac{\theta^n}{s^{n+1}} \quad (\theta < s < \infty) \tag{20}$$

To find UMVUE we are going to calculate expectation of S

$$
\begin{aligned}
E[S] \;&=\; \int_0^\infty s\frac{n\theta^n}{s^{n+1}}ds && (21)\\
&= n\theta^n \int_0^\infty s^{-n}ds && (22)\\
&= \frac{n\theta^n}{-n+1}\left[s^{-n+1}\right]_\theta^\infty && (23)\\
&= \frac{n\theta^n}{-n+1}(-\theta)^{-n+1} && (24)\\
&= \frac{n}{n-1}\theta && (25)
\end{aligned}
$$

Therefore, $T^* = \frac{n-1}{n}s$ is an unbiased estimator of θ as $E[T^*] = \theta$.

We already proved in previous sub-question (a) that S is a sufficient and complete statistic for θ and from the above result that T^* is a function of S.

Since it satisfies two conditions of theorem 10.4.1(Lehamann-Scheffe'), we can conclude that T^* is the UMVUE of θ

c. We've got this equation from sub-question (b)

$$f_S(s) = n\frac{\theta^n}{s^{n+1}} \quad (\theta < s < \infty) \tag{26}$$

To find UMVUE we are going to calculate expectation of $1/S\frac{n+1}{n}$ and see if it is unbiased.

$$E[(\frac{n+1}{n})(\frac{1}{s})] = \frac{n+1}{n}\int_0^\infty \frac{1}{s}\frac{n\theta^n}{s^{n+1}}ds \tag{27}$$

$$= \frac{n+1}{n}n\theta^n\int_0^\infty s^{-n-2}ds \tag{28}$$

$$= -\frac{n+1}{n}\frac{n\theta^n}{n+1}\left[s^{-n-1}\right]_\theta^\infty \tag{29}$$

$$= -\theta^n(-\theta)^{-n-1} \tag{30}$$

$$= \frac{1}{\theta} \tag{31}$$

Therefore,we can conclude that $T^{**} = \frac{n+1}{n}s^{-1}$ is an unbiased estimator of $\frac{1}{\theta}$.

We already proved in previous sub-question (a) that S is a sufficient and complete statistic for $\frac{1}{\theta}$ and from the above result that T^{**} is a function of S.

Since the findings satisfies two conditions of theorem 10.4.1(Lehamann-Scheffe') , we can conclude that T^{**} is the UMVUE of $\frac{1}{\theta}$

d.

In order to show that $\frac{X_{2:n}}{X_{1:n}}$ and $X_{1:n}$ are stochastically independent,

We have to first find the pdf of Y using transformation method

Let's define $Y = \frac{X}{\theta}$

$$f_Y(y) = f_X(x^{-1}(y))|J| \; = \; \left[\frac{\theta}{(\theta y)^2} I_{\theta:\infty}(z)\right]\theta \; = \; \frac{1}{y^2} I_{1:\infty}(z), \quad (y > 1) \tag{32}$$

Here, we substituted (θy) in x place .

$$Z = \frac{X_{2:n}}{X_{1:n}} \; = \; \frac{\frac{X_{2:n}}{\theta}}{\frac{X_{1:n}}{\theta}} \; = \; \frac{Y_{2:n}}{Y_{1:n}} \tag{33}$$

$(Y_1, \ldots, Y_n) = Z$ is free of θ since pdf of y and joint pdf y is free of θ.

In sub-question (a) we proved that $S = x_{1:n}$ has complete and sufficient statistic for θ.

$\therefore$ by Theorem 10.4.7(Basu) $Z = \frac{X_{2:n}}{X_{1:n}}$ and S are stochastically independent

Exercise 3

a.

i) Finding S which is a sufficient statistic for μ

$$f(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n}\left[e^{-\mu}\cdot\frac{\mu^{x_i}}{x_i!}\right] \tag{34}$$

$$= e^{-n\mu}\cdot\mu^{\Sigma x_i}\cdot\frac{1}{\prod x_i} \tag{35}$$

By the factorization criterion, (36)

$$f(x_1,\ldots,x_n;\theta) = g(s;\mu)h(x_1,\ldots,x_n) \tag{37}$$

$$\Rightarrow \begin{cases} g(s;\mu) = e^{-n\mu}\cdot\mu^{\Sigma x_i} \\ h(x_1,\ldots,x_n) = \frac{1}{\prod x_i} \end{cases} \tag{38}$$

$$\therefore S = \sum_{i=1}^{n} X_i \tag{39}$$

ii) Showing that S is also a complete statistic

$$S = X_1+\ldots+X_n \sim POI(n\mu) \tag{40}$$

Let $n\mu = \theta$ (41)

$$\Rightarrow f_S(s) = e^{-\theta}\cdot\frac{\theta^s}{s!}\ , \quad (\theta > 0,\ s = 0,1,\ldots) \tag{42}$$

Let $u(S)$ is a function of S (43)

$$\Rightarrow E[u(S)] = \sum_{s=0}^{\infty}\left[u(s)\cdot e^{-\theta}\cdot\frac{\theta^s}{s!}\right] \tag{44}$$

$$\text{Assume that } E[u(S)] = 0 \tag{45}$$

$$e^{-\theta} \neq 0 \text{ and } \theta^s \neq 0 \tag{46}$$

$$\Rightarrow \text{all } \frac{u(s)}{s!} = 0 \tag{47}$$

$$\Rightarrow u(S) = 0 \tag{48}$$

$$\Rightarrow E[u(S)] = 0 \text{ implies } u(S) = 0 \tag{49}$$

$$\Rightarrow \text{By Definition 10.4.1, } \{f_S(s;\theta) \mid \theta \in \Omega\} \text{ is complete.} \tag{50}$$

$$\Rightarrow S \text{ is a complete statistic for } \theta = n\mu. \tag{51}$$

$$\therefore S = \sum_{i=1}^{n} X_i \text{ is a C\&S for } \mu. \tag{52}$$

b.

$$T \sim BIN(1, p) \tag{53}$$

$$f_T(t) = p^t (1-p)^{1-t}, \quad (t = 0, 1) \tag{54}$$

$$= (1-p)\exp\left[t \cdot \ln\left(\frac{p}{1-p}\right)\right] \tag{55}$$

$$\Rightarrow \text{By Definition 10.4.2,} \tag{56}$$

$$f_T(t) \text{ is a member of } REC(q_1) \text{ with } q_1(p) = \ln\left(\frac{p}{1-p}\right) \text{ and } r_1(t) = t. \tag{57}$$

$$\Rightarrow \text{By Theorem 10.4.2, } S_1 = \sum_{i=1}^{n} r_1(T_i) = \sum_{i=1}^{n} T_i \text{ is a C\&S for } p. \tag{58}$$

$$\bar{T}_n \text{ is an unbiased estimator for } p. \quad \because E[\bar{T}_n] = E[T] = p \tag{59}$$

$$\bar{T}_n \text{ is a function of } S_1. \quad \because \bar{T}_n = \frac{S_1}{n} \tag{60}$$

$$\therefore \text{By Theorem 10.4.1, } \bar{T}_n \text{ is a UMVUE of } p. \tag{61}$$

c.

In this subquestion, T means the T of the subquestion 3b.

$$Y_n = T_1 + T_2 + \ldots + T_n \tag{62}$$

$$\Rightarrow \hat{p} = \frac{Y_n}{n} = \bar{T}_n \tag{63}$$

We need two things to use Definition 9.4.4 to solve this subquestion.

1. $\{J_n\}$, an asymptotically unbiased sequence of estimator for p

2. CRLB for $Var[J_n]$

$\Rightarrow$ Objective : $ae(J_n)$, the asymptotic efficiency of $\{J_n\}$

i) Finding $\{J_n\}$

$$\text{By Definition 9.4.3, } \hat{p} = \bar{T}_n \text{ is asymptotically unbiased for } p \tag{64}$$

$$\because \lim_{n\to\infty} E[\bar{T}_n] = E[E(T)] = p \tag{65}$$

$$\therefore \{J_n\} = \{\bar{T}_n\} \tag{66}$$

ii) Compute CRLB for $Var[J_n]$

$$Var[J_n] \ \geq \ CRLB \tag{67}$$

$$\Leftrightarrow Var[\bar{T}_n] \geq \frac{[\tau'(\mu)]^2}{-nE\left[\frac{\partial^2}{\partial\mu^2}\ln f(x;\mu)\right]} \ , \quad (\tau(\mu) = p = e^{-\mu}) \tag{68}$$

$$\begin{cases} [\tau'(\mu)]^2 = e^{-2\mu} \\ E\left[\frac{\partial^2}{\partial\mu^2}\ln f(x;\mu)\right] = E\left[\frac{\partial^2}{\partial\mu^2}\big(-\mu + x\ln\mu - \ln(x!)\big)\right] = E\left[-\frac{X}{\mu^2}\right] = -\frac{1}{\mu} \end{cases} \tag{69}$$

$$\text{Substitute (69) to (68)} \tag{70}$$

$$Var[\bar{T}_n] \ \geq \ \frac{\mu e^{-2\mu}}{n} = CRLB \tag{71}$$

If we check if there is a linear relationship between $\bar{T}_n$ and $\tau(\mu) = e^{-\mu}$,

$$\bar{T}_n \stackrel{?}{=} \ a\left(\frac{\partial}{\partial\mu}\ln f\left(x_1,\ldots,x_n;\mu\right)\right) + e^{-\mu} \tag{72}$$

$$\text{LHS } = \exp\left(-\bar{X}_n\right) \tag{73}$$

$$\text{RHS } = a\left(\frac{X_1+\ldots+X_n-n\mu}{\mu}\right) + e^{-\mu} \tag{74}$$

We can see that LHS $\neq$ RHS. However, if we take the limit of both sides and let $n \to \infty$, LHS becomes equal to RHS. Because $\bar{X}_n \to \mu$ and $\sum X_i - n\mu \to 0$ as

$n \to \infty$. Therefore, $Var[\bar{T}_n] = CRLB$ as n approaches to ∞.

$$\text{By Definition 9.4.4,} \tag{75}$$

$$\text{Since } \{J_n\} = \{\bar{T}_n\}, \tag{76}$$

$$ae(\bar{T}_n) \tag{77}$$

$$= are(\bar{T}_n, \bar{T}_n^*) = \lim_{n\to\infty} \frac{Var[\bar{T}_n^*]}{Var[\bar{T}_n]} = \lim_{n\to\infty} \frac{\text{CRLB}}{nVar[T]} \tag{78}$$

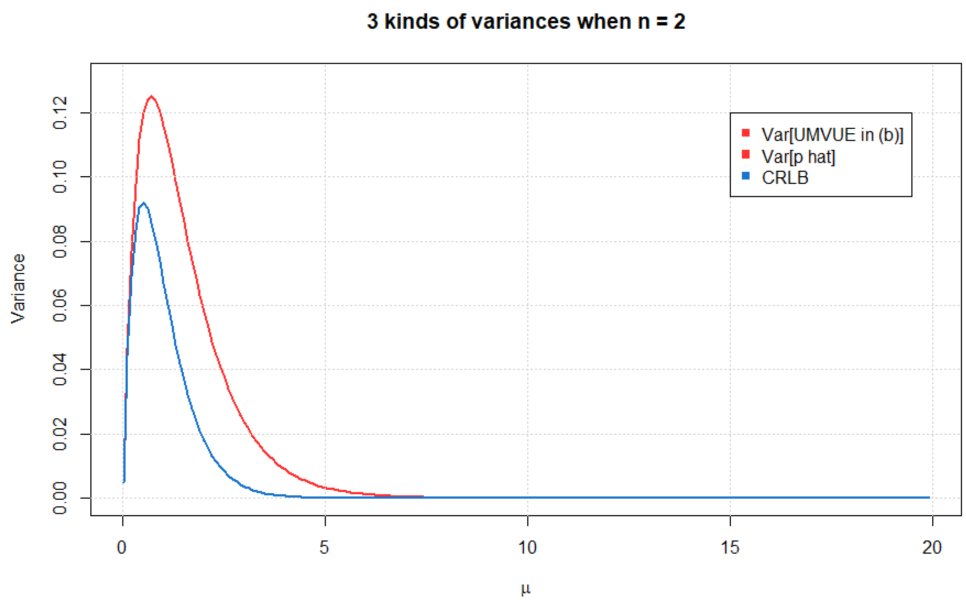
$$= \lim_{n\to\infty} \frac{\left(\frac{\mu e^{-2\mu}}{n}\right)}{ne^{-\mu}(1-e^{-\mu})} = 1 \tag{79}$$

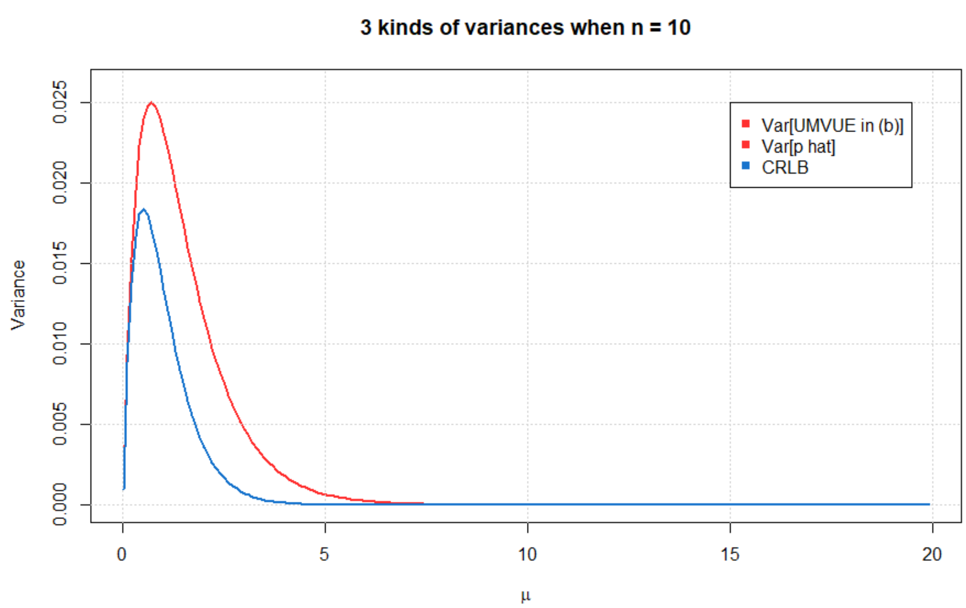
$$\therefore \text{The asymptotic efficiency of } \{\bar{T}_n\} = 1 \tag{80}$$

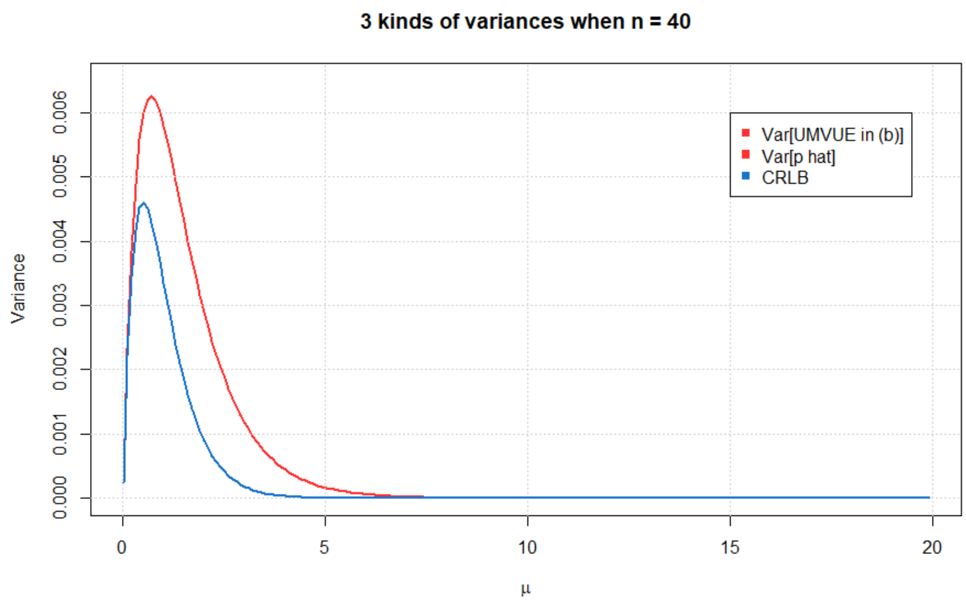
d.

Among three types of variances required to graph, $Var[$ UMVUE in (b)$] = Var[\hat{p}]$. UMVUE in (b) is $\bar{T}_n = \frac{T_1+...+T_n}{n}$, and its numerator means how many $x_i = 0$ among n samples, which is same meaning as Y_n. Therefore, $\bar{T}_n = \frac{Y_n}{n} = \hat{p} \Rightarrow Var[\bar{T}_n] = Var[\hat{p}] = e^{-\mu}(1-e^{-\mu})$. This is illustrated as the red line in the graphs.

On the other hand, It has been shown that $\bar{T}_n$ is an unbiased estimator of p in (59). Thus, CRLB of $Var[\bar{T}_n]$ can be 'the CRLB of unbiased estimators of p'. The CRLB of $Var[\bar{T}_n]$ has been computed as $\frac{\mu e^{-2\mu}}{n}$ in (71). This is illustrated as the blue line in the graphs.


3 kinds of variances when n = 2
Variance
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
5
10
15
20
μ
Var[UMVUE in (b)]
Var[p hat]
CRLB


3 kinds of variances when n = 10
Variance
0.000
0.005
0.010
0.015
0.020
0.025
0
5
10
15
20
μ
Var[UMVUE in (b)]
Var[p hat]
CRLB


3 kinds of variances when n = 40
Variance
0.000
0.001
0.002
0.003
0.004
0.005
0.006
Var[UMVUE in (b)]
Var[p hat]
CRLB
0
5
10
15
20
μ