Stat511(Section001): Homework #7

Due on Mar. 16, 2022 at 5:00pm

Instructor:Xuming He

Tiejin Chen **tiejin@umich.edu**

Problem 1

(a)

First we compute the MLE of p. We have the pmf of X_i is:

$$f(x) = p^x(1-p)^{1-x}$$

Thus we have the los-likelihood function of $X_1, .., X_n$ is:

$$logL(X_1, .., X_n) = \sum_{i=1}^{n}(x_i log(p) + (1-x_i)log(1-p)$$

Then we have:

$$\frac{\partial logL(X_1, .., X_n)}{\partial p} = 0 \rightarrow \hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Now let us consider the variance of $\hat{p}$, we know:

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

Now consider the Cramer-Rao inequality. We have:

$$E_p(\frac{1}{n}\sum_{i=1}^{n} x_i) = p, \frac{dp}{dp} = 1$$

$$-E[\frac{\partial^2}{\partial p^2} log f(x|p)] = E(\frac{x}{p^2} + \frac{1-x}{(1-p)^2}) = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}$$

Thus the Cramer-Rao lower bound is:

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

Thus variance of the maximum likelihood estimator of p attains Carmer-Rao lower bound.

(b)

Since for all x_i, they are iid, then we have:

$$E(x_1x_2x_3x_4) = E(x_1)E(x_2)E(x_3)E(x_4) = p^4$$

Thus it is a unbiasd estimator of p^4. And we know one complete sufficient statistic for p is $\sum_{i=1}^{n}$, then we have:

$$\phi(T) = E(x_1x_2x_3x_4|\sum_{i=1}^{n} x_i = t) = P(x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 1|T) = \frac{p(x_1, x_2 = 1, x_3 = 1, x_4 = 1)p(\sum_{i=5}^{n} x_i = t-4)}{p(\sum_{i=1}^{n} x_i = t)}$$

By Bayes' rule. Then we have:

$$\phi(T) = \frac{(n-4Ct-4)p^t(1-p)^{n-t}}{(nCt)p^t(1-p)^{n-t}} = \frac{n-4Ct-4}{nCt} = \frac{t(t-1)(t-2)(t-3)}{n(n-1)(n-2)(n-3)}$$

is the best unbiased estimator of p where $t = \sum_{i=1}^{n} x_i$ and $t \geq 3$. And we also know p must be greater than 0. Thus we can get the best unbiased estimator of p is $\max(0, \frac{t(t-1)(t-2)(t-3)}{n(n-1)(n-2)(n-3)})$.

Problem 2

(a)
We have:

$$logL(x_i, ..., x_n) \propto -\frac{n}{2}log(\sigma^2) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial logL(x_i, ..., x_n)}{\partial \sigma^p} = \frac{\partial logL(x_i, ..., x_n)}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \sigma^p}$$

And we have:

$$\frac{\partial \sigma^2}{\partial \sigma^p} = \frac{\partial t^{\frac{2}{p}}}{\partial t} = \frac{2}{p}t^{\frac{2}{p}-1} = \frac{2}{p}\sigma^{2-p}$$

when $t = \sigma^p$. Thus:

$$\frac{\partial logL(x_i, ..., x_n)}{\partial \sigma^p} = \frac{2}{p}\sigma^{2-p}(-\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4})$$

$$E[(\frac{\partial logL(x_i, ..., x_n)}{\partial \sigma^p})^2] = E\{[\frac{2}{p}\sigma^{2-p}(\frac{\sum_{i=1}^{n}(x_i-\mu)^2 - n\sigma^2}{2\sigma^4})]^2\} = \frac{4}{p^2}\sigma^{4-2p}\frac{Var(\sum_{i=1}^{n}(x_i-\mu)^2)}{4\sigma^8}$$

Thus we can have:

$$I_n(\sigma^p) = \frac{2n\sigma^{-2p}}{p^2}$$

(b)
Consider $\frac{d}{d\theta}E_\theta W(x)$, for any unbiased estimator $W(x)$, we have:

$$E_\theta W(x) = \theta$$

Thus:

$$\frac{d}{d\theta}E_\theta W(x) = 1$$

for every θ.Thus, the Cramer-Rao lower bound is just $\frac{1}{I_n(\sigma^p)}$. Thus it is:

$$\frac{p^2}{2n\sigma^{-2p}}$$

Problem 3

(a)
We have the likelihood function of uniform distribution is:

$$L(x_1, ..., x_n)\frac{1}{\theta^n}\prod_{i=1}^{n}I(0 \le x_i \le \theta) = \frac{1}{\theta^n}I(x_{(n)} \le \theta)I(x_{(1)>0})$$

And we know $\frac{1}{\theta^n}$ is a decreasing function of θ. So to make likelihood as large as possible, θ need to as small as possible. And the smallest θ meet the requirement is $x_{(n)}$. Thus the MLE of θ is $x_{(n)}$. Now we consider $var(x_{(n)})$, we have:

$$E(x_{(n)}) = \int_{x=0}^{\theta} nx\frac{x^{n-1}}{\theta^n}dx = \frac{n}{n+1}\theta$$

And

$$E(x_{(n)}^2) = \int_{x=0}^{\theta} nx^2\frac{x^{n-1}}{\theta^n}dx = \frac{n}{n+2}\theta^2$$

Thus we have:

$$Var(x_{(n)}) = E(x_{(n)}^2) - E(x_{(n)})^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n}{(n+2)(n+1)^2}\theta^2$$

Problem 3 continued on next page. . .

We have the variance decrease at the rate of $\frac{1}{n^2}$ instead of $\frac{1}{n}$.

(b)

It does not hold. Consider the Cramer-Rao lower bound for MLE of θ, we have:

$$I(\theta) = nE[(\frac{\partial}{\partial\theta}log(1/\theta))^2] = \frac{n}{\theta^2}$$

$$(\frac{d}{d\theta}E(x_{(n)}))^2 = \frac{n^2}{(n+1)^2}$$

Thus the Cramer-Rao lower bound is $\frac{n}{(n+1)^2}\theta^2$. And $Var(x_{(n)}) = \frac{n}{(n+2)(n+1)^2}\theta^2 < \frac{n}{(n+1)^2}\theta^2$. And thus it does not hold. For the reason, We have:

$$\frac{d}{d\theta}E(x_{(n)}) = \frac{n}{n+1}$$

$$\int_{x=0}^{\theta}\frac{\partial}{\partial\theta}max(x)\frac{1}{\theta}dx = -\int_{x=0}^{\theta}\frac{x}{-\theta^2}dx = -\frac{1}{2}$$

They are not equal and the requirement of Cramer-Rao Inequality does not hold. Therefore, Cramer-Rao lower bound does not hold.

Problem 4

(a)

No, We cannot. In this problem, we do not konw the distribution of Y_i and X_i,we know that $Y|X \sim N(\beta X, \sigma^2)$. However, we cannot assume X_i is given here. Thus we cannot use $Y_i|X_i$. We can only have:

$$Y_i - \beta X_i \sim N(0, \sigma^2), f(Y_i - \beta X_i = z_i) = \frac{1}{\sqrt{2\pi}\sigma}exp(-\frac{z_i^2}{2\sigma^2})$$

We cannot further simplify this function because it is a function of z. And if we plug $z_i = Y_i - \beta X_i$, we cannot use exponitional family of distribution of (X, Y), And thus we cannot get a complete sufficient statistic for β. Hence we cannot use both when we know σ^2 or not.