Statistical Inference Assignment 6

Junhao Yuan (20307130129)

November 3, 2022

Problem 1.

Suppose $X_1, \cdots, X_n \overset{iid}{\sim} B(p)$.

(a) Show that the variance of the MLE of p attains the Cramer-Rao lower bound.

(b) For $n \geq 4$, show that the product $X_1X_2X_3X_4$ is an unbiased estimator of p^4, and use this fact to find the best unbiased estimator of p^4.

Solution.

(a) It's easy to compute that the MLE of p is $\hat{p} = \sum_{i=1}^n X_i/n$ which is unbiased, and

$$\mathbb{V}(\hat{p}) = \frac{p(1-p)}{n}.$$

The fisher information of the Bernoulli distribution is

$$\begin{aligned} I(p) &= -\mathbb{E}\left(\frac{\partial^2}{\partial p^2} log(p^x(1-p)^{1-x})\right) \\ &= \mathbb{E}\left(\frac{x}{p^2} + \frac{1-x}{(1-p)^2}\right) \\ &= \frac{1}{p(1-p)}. \end{aligned}$$

Therefore, the C-R lower bound is

$$\frac{1}{nI(p)} = \frac{p(1-p)}{n},$$

which is equal to the variance of $\hat{p}$.

(b) Since X_1, X_2, X_3 and X_4 are independent and identically distributed,

$$\mathbb{E}(X_1X_2X_3X_4) = \prod_{i=1}^4 \mathbb{E}(X_i) = p^4.$$

Thus, $X_1X_2X_3X_4$ is an unbiased estimator of p^4.

We have already known that $T = \sum_{i=1}^n X_i \sim B(n,p)$ is a complete sufficient statistics. Therefore,

$$\phi(T) = \mathbb{E}\left(X_1X_2X_3X_4 | T\right)$$

is the best unbiased estimator of p^4. Here,

$$\begin{aligned}
\phi(t) &= \mathbb{E}\left(X_1X_2X_3X_4|T=t\right) \\
&= P(X_1X_2X_3X_4|T=t) \\
&= \frac{P(X_1=X_2=X_3=X_4=1, \sum_{i=5}^n X_i = t-4)}{P(T=t)} \\
&= \frac{p^4 \cdot \binom{n-4}{t-4} p^{t-4}(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}} \\
&= \frac{\binom{n-4}{t-4}}{\binom{n}{t}} \qquad (1)
\end{aligned}$$

when $n \geq 5$. When $n = 4$, $\mathbb{E}\left(X_1X_2X_3X_4|T\right) = 1$ and (1) also equals to zero. Hence, the best unbiased estimator of p^4 is

$$\phi(T) = \frac{\binom{n-4}{T-4}}{\binom{n}{T}},$$

where $T = \sum_{i=1}^n X_i$.

Problem 2.

Let $X_1, \cdots, X_n \overset{iid}{\sim} P(\lambda)$, and let $\bar{X}$ and S^2 denote the sample mean and variance respectively.

(a) Prove that $\bar{X}$ is the best unbiased estimator of λ without using the Cramer-Rao theorem.

(b) Prove that $\mathbb{E}(S^2|\bar{X}) = \bar{X}$ and use it to show that $\mathbb{V}(S^2) > \mathbb{V}(\bar{X})$.

Solution.

(a) We can write the joint pmf of the samples as

$$\begin{aligned}
f(X|\lambda) &= \prod_{i=1}^n e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} \\
&= \frac{e^{-n\lambda}}{\prod_{i=1}^n x_i!} e^{log(\lambda)\sum X_i} \\
&= \frac{e^{-ne^{\mu}}}{\prod_{i=1}^n X_i} e^{\mu \sum X_i},
\end{aligned}$$

where $\mu = log(\lambda) \in \mathbb{R}$. Since there exists an open set in the natural parameter space of μ, $T = \sum_{i=1}^n X_i$ is a complete sufficient statistics, which means any estimators based on T is the best unbiased estimator of its expected value. Since $\bar{X} = T/n$ and $\mathbb{E}(\bar{X}) = \lambda$, we can conclude that $\bar{X}$ is the unique best unbiased estimator of λ.

(b) It's easy to verify that $\bar{X}$ is also a complete sufficient statistics. Since S^2 is an unbiased estimator of λ, $\phi(\bar{X}) = \mathbb{E}(S^2|\bar{X})$ is also the best unbiased estimator of λ. However, the best unbiased estimator of λ is unique, which means $\mathbb{E}(S^2|\bar{X}) = \bar{X}$.

Since

$$\begin{aligned}
\mathbb{V}(S^2) &= \mathbb{V}(\mathbb{E}(S^2|\bar{X})) + \mathbb{E}(\mathbb{V}(S^2|\bar{X})) \\
&= \mathbb{V}(\bar{X}) + \mathbb{E}(\mathbb{V}(S^2|\bar{X})),
\end{aligned}$$

if $\mathbb{E}(\mathbb{V}(S^2|\bar{X})) > 0$, we can prove that $\mathbb{V}(S^2) > \mathbb{V}(\bar{X})$. Now let's prove it.

If $\mathbb{E}(\mathbb{V}(S^2|\bar{X})) = 0$, $\mathbb{V}(S^2|\bar{X})$ is an unbiased estimator of 0. But $\mathbb{V}(S^2|\bar{X})$ is a function based on $\bar{X}$, a complete sufficient statistics, which means $\mathbb{V}(S^2|\bar{X}) = 0$ with probability 1. This implies that given $\bar{X}$, S^2 is constant, which is not possible. Thus, we can conclude that $\mathbb{E}(\mathbb{V}(S^2|\bar{X})) > 0$.

PROBLEM 3.

Suppose $X_1, \cdots, X_n \stackrel{iid}{\sim} B(p)$. Find the UMVUE of $p(1-p)$. Make sure to prove that the estimator is indeed a UMVUE of $p(1-p)$.

SOLUTION.

Define $\phi(X) = \mathbb{I}(X_1 = 1, X_2 = 0)$, then we have $\mathbb{E}(\phi(X)) = p(1-p)$, which means $\phi(X)$ is an unbiased estimator of $p(1-p)$. Note that $T = \sum_{i=1}^n X_i$ is a complete sufficient statistics. Therefore, $g(T) = \mathbb{E}(\phi(X)|T)$ is also an unbiased estimator of $p(1-p)$, and by Lehmann-Scheff theorem, is the unique best unbiased estimator. And

$$\begin{aligned}
g(t) &= P(X_1 = 1, X_2 = 0|T = t) \\
&= \frac{P(X_1 = 1, X_2 = 0, \sum_{i=3}^n X_i = t-1)}{P(\sum_{i=1}^n X_i = t)} \\
&= \frac{p(1-p) \cdot \binom{n-2}{t-1} p^{t-1}(1-p)^{n-t-1}}{\binom{n}{t} p^t (1-p)^{n-t}} \\
&= \frac{t(n-t)}{(n-1)n}.
\end{aligned}$$

PROBLEM 4.

Prove the following statement:

Let T be a complete sufficient statistics for θ and let $\phi(T)$ be any estimator based on T. Then $\phi(T)$ is the unique unbiased estimator of its expected value.

SOLUTION.

By Rao-BlackWell, if we want to find the best unbiased estimator, we need only consider unbiased estimators based on T. Since T is complete, for any estimators of 0 based on T, which satisfies

$$\mathbb{E}(g(T)) = 0,$$

we have $g(T) = 0$ with probability 1. Hence, there is no unbiased estimator of 0 except 0 itself, which means $\phi(T)$ is uncorrelated with all unbiased estimators of 0. Therefore, $\phi(T)$ is the best unbiased estimator of its expected value, and furthermore, it is unique.