STAT 530 Homework 6

2022-03-18

(1) Problem 7.55 (a) and (b) Casella & Berger: For each of the following pdfs, let $X_1, \dots, X_n$ be a sample from that distribution. In each case, find the best unbiased estimator of θ^r.

(a) $f(x \mid \theta) = \frac{1}{\theta}, \quad 0 < x < \theta, r < n$

It is known that $T = X_{(n)}$ is a minimal sufficient and complete statistic (example 6.2.23). Using theorem 7.2.23 from Casella and Berger, we can find a UMVUE $g(X_{(n)})$. We first calculate the density of T:

$$
\begin{aligned}
f_T(t) &= \frac{\partial}{\partial t} P(X_{(n)} \le t) = \frac{\partial}{\partial t}[F_X(t)]^n \\
&= n[F_X(t)]^{n-1} f_X(t) \\
&= \frac{nt^{n-1}}{\theta^n} I(0 < t < \theta)
\end{aligned}
$$

Then the UMVUE $g(T)$ must satisfy

$$\theta^r = \int_0^\theta g(t) \frac{nt^{n-1}}{\theta^n} dt$$

Taking the derivative of both sides with respect to θ, we have

$$
\begin{aligned}
r\theta^{r-1} &= -n\theta^{-(n+1)} \int_0^\theta g(t) n t^{n-1} dt + \theta^n g(\theta) n \theta^{n-1} \\
&= -\frac{n}{\theta} \int_0^\theta g(t) \frac{nt^{n-1}}{\theta^n} dt + \frac{ng(\theta)}{\theta} \\
&= -\frac{n\theta^r}{\theta} + \frac{ng(\theta)}{\theta}.
\end{aligned}
$$

Now we can solve for $g(\theta)$:

$$
\begin{aligned}
r\theta^{r-1} &= -\frac{n\theta^r}{\theta} + \frac{ng(\theta)}{\theta} \\
r\theta^r &= -n\theta^r + ng(\theta) \\
\theta^r(r+n) &= ng(\theta) \\
\frac{\theta^r(r+n)}{n} &= g(\theta)
\end{aligned}
$$

Replacing θ is the complete sufficient statistic $X_{(n)}$, we have that the UMVUE for θ^r is

$$g(X_{(n)}) = \frac{X_{(n)}^r (n+r)}{n}$$

(b) $f(x \mid \theta) = e^{-(x-\theta)}, \quad x > \theta$

Here, it can be shown that $T = X_{(1)}$ is a complete and sufficient statistic. By the Neyman-Fisher factorization criterion, T is sufficient, and we can show that T is complete as follows:

First we find the density function $f_T(t)$:

$$\begin{aligned} f_T(t) &= \frac{\partial}{\partial t} 1 - [1 - F_X(t)]^n = n[1 - F_X(t)]^{n-1} f_X(t) \\ &= ne^{-n(t-\theta)} I(t > \theta). \end{aligned}$$

Now, suppose there exists a function $g(T)$ such that $Eg(T) = 0$ for all θ. Then we have

$$0 = \int_\theta^\infty g(t) ne^{n\theta} e^{-nt} dt = -ne^{n\theta} \int_\infty^\theta g(t) e^{-nt} dt$$

Taking the partial derivative with respect to θ on both sides, we get

$$0 = n \int_\theta^\infty ng(t) e^{-n(t-\theta)} dt - ng(\theta).$$

The first term, $nEg(T)$, is equal to zero. But then we have that $g(t) = 0$ for all θ. So by the definition of completeness, T is a complete statistic.

Now that we have established that T is complete and sufficient, we know that an unbiased estimator $\varphi(T)$ of θ^r which is a function of only T will be the UMVUE for θ^r. So we have

$$\begin{aligned} \theta^r &= E\varphi(T) \\ &= \int_\theta^\infty \phi(t) ne^{-n(t-\theta)} dt \end{aligned}$$

Taking the derivative with respect to θ on both sides, we get

$$\begin{aligned} r\theta^{r-1} &= n \int_\theta^\infty \varphi ne^{-n(t-\theta)} dt - n\varphi(\theta) \\ &= \theta^r - n\varphi(\theta). \end{aligned}$$

We then have that φ must satisfy $\varphi(\theta) = \frac{\theta^{r-1}(n\theta - r)}{n}$, so the UMVUE for θ^r must be

$$\varphi(X_{(1)}) = \frac{X_{(1)}^{r-1}(nX_{(1)} - r)}{n}.$$

(2) Suppose $X_1, \ldots, X_n \overset{iid}{\sim}$ Poisson(λ), where

$$P(X = x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, \ldots; \quad 0 \leq \lambda < \infty.$$

Find the UMVUE of $\tau(\lambda) = \lambda^r$ for some positive integer r.

It is known that $T = \bar{X}$ is a complete, sufficient statistic. If we can find a function $\varphi(T)$ that is unbiased for λ^r, then $\phi(T)$ will be the UMVUE for λ^r. So suppose there does exist a function $\phi(T)$ that is an unbiased estimator of λ^r. Then because $T \sim$ Poisson$(n\lambda)$, we have

$$\lambda^r = \sum_{t=0}^{\infty} \varphi(T) \frac{e^{-n\lambda}(n\lambda)^t}{t!}.$$

Rewriting this equality, we obtain

$$\begin{aligned} \lambda^r e^{n\lambda} &= \sum_{t=0}^{\infty} \varphi(t) \frac{(n\lambda)^t}{t!} \\ \sum_{i=0}^{\infty} \frac{n^i \lambda^{r+i}}{i!} &= \sum_{t=0}^{\infty} \varphi(t) \frac{(n\lambda)^t}{t!} \end{aligned}$$

Since both sides of the equality are polynomials in λ, each term on the left hand side must have the same coefficient as the term on the right side with the corresponding degree. Because the smallest power of λ on the left hand side is r, $\varphi(t) = 0$ for $0 \leq t < r$. For $t \geq r$, we see that

$$\frac{n^{t-r}\lambda^t}{(t-r)!} = \varphi(t) \frac{(n\lambda)^t}{t!}$$

Solving for $\varphi(t)$, we find that the UMVUE for λ^r is

$$\varphi(t) = \begin{cases} 0 & \text{for } 0 \leq t < r \\ \frac{n^{-r} t!}{(t-r)!} & \text{for } t \geq r \end{cases}$$

(3) Prove the following claims:

(a) Suppose $\hat{\theta}$ is the unique Bayes estimator, then $\hat{\theta}$ is admissible.
(b) Suppose θ^* is the unique minimax estimator, then θ^* is admissible.

(4) For this question, we will study the *breakdown value* in greater depth. The textbook definition of a breakdown value is given below:

Definition 10.2.2 Let $X_{(1)} < \cdots < X_{(n)}$ be an ordered sample of size n, and let T_n be a statistic based on this sample. T_n has *breakdown value* $b, 0 \leq b \leq 1$, if, for every $\epsilon > 0$,

$$\lim_{X_{(\{(1-b)n\})} \to \infty} T_n < \infty \quad \text{and} \quad \lim_{X_{(\{(1-(b+\epsilon))n\})} \to \infty} T_n = \infty$$

(Recall Definition 5.4.2 on percentile notation)

Where $\{b\}$ is the number b rounded to the nearest integer. That is, if i is an integer and $i - 0.5 \leq b < i + 0.5$, then $\{b\} = i$. The textbook also claims that the sample median M_n has a breakdown value of 50%. These do not make sense. For example, consider $n = 10$, $b = 50\%$, and $\epsilon = 0.01$, then $\{(1-b)n\} = \{(1-(b+\epsilon))n\} = 5$. Obviously we cannot have

$$\lim_{X_{(\{(1-b)n\})} \to \infty} M_n = \lim_{X_{(5)} \to \infty} M_n < \infty \quad \text{and} \quad \lim_{X_{(\{(1-(b+\epsilon))n\})} \to \infty} T_n = \lim_{X_{(5)} \to \infty} M_n = \infty$$

at the same time.

Now consider replacing the equations in Definition 10.2.2 by

$$\lim_{X_{(\lfloor(1-b)n\rfloor)}\to\infty} T_n < \infty \quad \text{and} \quad \lim_{X_{(\lfloor(1-(b+\epsilon))n\rfloor)}\to\infty} T_n = \infty,$$

where $\lfloor b \rfloor$ is the greatest integer less than or equal to b, that is, the floor function of b. Show that, under the new definition, the sample median M_n has a breakdown value of $\frac{\lfloor\frac{n-3}{2}\rfloor}{n}$ (assume $n \geq 3$). Obviously, this converges to 50% as $n \to \infty$.