Matthew Wilder
MATH 326 - Spring 2022
Homework 04
Due: Saturday 02/19/22 at 03:00

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform distribution on the interval $(\theta,\ \theta+1)$. Let

$$\hat{\theta}_1 = \bar{X} - \frac{1}{2}, \quad \text{and} \quad \hat{\theta}_2 = X_{(n)} - \frac{n}{n+1}.$$

   (a) Show that both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ.

   Solution: We need to show that $E[\,\hat{\theta}_1\,] = \theta$. By our given defintion of $\hat{\theta}_1$,

$$E[\,\hat{\theta}\,] = E\left(\bar{X} - \frac{1}{2}\right) = E(\bar{X}) - \frac{1}{2}$$

   By the definition of a uniform distribution,

$$\mu = \frac{b+a}{2} \iff \mu = \frac{\theta + (1+\theta)}{2} \iff \mu = \theta + \frac{1}{2}$$

$$E[\,\bar{X}\,] = E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{n\mu}{n} = \mu = \theta + \frac{1}{2}$$

   We now have all the parts needed to compute $E[\,\hat{\theta}_1\,]$. Thus,

$$[\,\hat{\theta}_1\,] = E(\bar{X}) - \frac{1}{2} = \theta - \frac{1}{2} + \frac{1}{2} = \theta.$$

   Therefore $\hat{\theta}_1$ is unbiased. Next we will show for $\hat{\theta}_2$.

$$\hat{\theta}_2 = X_{(n)} - \frac{n}{n+1}$$

$$g_{(n)}(y) = n\Big[F(y)\Big]^{n-1} f(y) = n\,[y-\theta]^{n-1}$$

   Again, by the deinition of a uniform distribution,

$$f(y) = \frac{1}{b-a} = \frac{1}{(\theta+1)-\theta} = 1 \quad \text{and} \quad F(y) = \int_\theta^y 1\,dt = y - \theta.$$

   From old notes,

$$E[\,X_{(n)}\,] = n\int_\theta^{\theta+1} y[y-\theta]^{n-1}\,dy$$

Recall IBP formula: $\int fg' = fg - \int f'g$.

Let $f = y$, $\quad f' = dy$, $\quad g' = (y-\theta)^{n-1}dy$, and $g = \frac{(y-\theta)^n}{n}$. Then,

$$\begin{aligned}
n\int_\theta^{\theta+1} y[y-\theta]^{n-1}\,dy &= n\int_\theta^{\theta+1} fg' \\
&= \frac{\cancel{n}y(y-\theta)^n}{\cancel{n}}\Bigg|_{y=\theta}^{y=\theta+1} - \int_\theta^{\theta+1} \frac{\cancel{n}(y-\theta)^n}{\cancel{n}}\,dy \\
&= \theta + 1 - \left[\frac{(y-\theta)^{n+1}}{n+1}\right]_{y=\theta}^{y=\theta+1} \\
&= \theta + 1 - \left[\frac{(\theta+1-\theta)^{n+1}}{n+1} - \frac{(\theta-\theta)^{n+1}}{n+1}\right] \\
&= \frac{(\theta+1)(n+1)}{n+1} - \frac{1}{n+1} \\
&= \frac{\theta n + \theta + n + 1 - 1}{n+1} \\
&= \frac{\theta(n+1) + n}{n+1} \\
&= \theta + \frac{n}{n+1}
\end{aligned}$$

Substituting this value back into $E[\,\hat{\theta}_2\,]$, we get that

$$E[\,\hat{\theta}_2\,] = E\left[X_{(n)} - \frac{n}{n+1}\right] = \theta + \frac{n}{n+1} - \frac{n}{n+1} = \theta.$$

Therefore, $\hat{\theta}_2$ is unbiased.

(b) Show that both estimators are consistent estimators.

Solution: Using the variances computed in part (c),

$$\lim_{n\to\infty} V(\hat{\theta}_1) = \lim_{n\to\infty} \frac{1}{12n} = 0$$

Therefore $\hat{\theta}_1$ is a consistent estimator.

$$\begin{aligned}
\lim_{n\to\infty} V(\hat{\theta}_2) &= \lim_{n\to\infty}\left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right) \\
&= \lim_{n\to\infty}\frac{n}{n+2} - \lim_{n\to\infty}\frac{n^2}{n^2+2n+1} \\
&= 1 - 1 \\
&= 0
\end{aligned}$$

Therefore $\hat{\theta}_2$ is a consistent estimator.

(c) Find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$.

Solution: We need to compute the variances of $\hat{\theta}_1$ and $\hat{\theta}_2$. For $\hat{\theta}_1$,

$$V(\hat{\theta}_1) = V\left(\bar{X} - \frac{1}{2}\right) = V(\bar{X}) - 0 = V(\bar{X}).$$

By the definition of uniform distributions,

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{((\theta+1)-\theta)^2}{12} = \frac{1}{12} = V(X_i)$$

Then,

$$V(\bar{X}) = V\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n^2}\left[V(X_1) + \cdots V(X_n)\right] = \frac{1}{n^2} \cdot \frac{n}{12} = \frac{1}{12n}.$$

Thus, $V(\hat{\theta}_1) = \frac{1}{12n}$. As for $\hat{\theta}_2$,

$$V(\hat{\theta}_2) = V\left(X_{(n)} - \frac{n}{n+1}\right) = V(X_{(n)}) - 0$$

$$V(X_{(n)}) = E(X_{(n)}^2) - \left[E(X_{(n)})\right]^2$$

$$\begin{aligned}
E(X_{(n)}^2) &= n\int_{\theta}^{\theta+1} x^2(x-\theta)^{n-1}\,dx \qquad \text{let } g(x) = u = (x-\theta) \\
&= n\int_{g(\theta)}^{g(\theta+1)} (u+\theta)^2(u)^{n-1}\,du \\
&= n\int_0^1 (u^2 + 2\theta u + \theta^2)(u)^{n-1}\,du \\
&= n\int_0^1 u^{n+1} + 2\theta u^n + \theta^2 u^{n-1}\,du \\
&= n\left[\frac{u^{n+2}}{n+2} + \frac{2\theta u^{n+1}}{n+1} + \frac{\theta^2 u^n}{n}\right]_{u=0}^{u=1} \\
&= n\left[\frac{1}{n+2} + \frac{2\theta}{n+1} + \frac{\theta^2}{n}\right] \\
&= \frac{n}{n+2} + \frac{2\theta n}{n+1} + \theta^2
\end{aligned}$$

From part (a) we know $E(X_{(n)})$, so

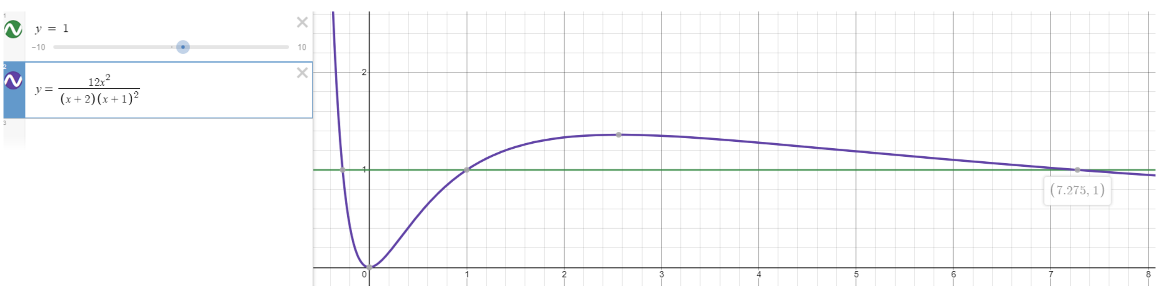
$$\begin{aligned}
V(X_{(n)}) &= E(X_{(n)}^2) - \left[E(X_{(n)})\right]^2 \\
&= \frac{n}{n+2} + \frac{2\theta n}{n+1} + \theta^2 - \left(\theta + \frac{n}{n+1}\right)^2 \\
&= \frac{n}{n+2} + \frac{2\theta n}{n+1} + \theta^2 - \left(\theta^2 + \frac{2\theta n}{n+1} + \frac{n^2}{(n+1)^2}\right) \\
&= \frac{n}{n+2} - \frac{n^2}{(n+1)^2} \\
&= V(\hat{\theta}_2)
\end{aligned}$$

Then,

$$\text{eff}(\hat{\theta}_1,\, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)} = \frac{\frac{n}{n+2} - \frac{n^2}{(n+1)^2}}{\frac{1}{12n}} = 12n\left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right) = \frac{12n^2}{(n+2)(n+1)^2}$$

(d) Which is the better estimator and why?

Solution: $\hat{\theta}_1$ is a better estimator when $n \in [1, 7]$, since the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is greater than 1 in that interval.
"Look at this photo**graph**" (Chad Kroeger):



The intersection of the graph is at 7.275 and it can be shown that the efficiency is monotonically decreasing for all $n \geq 3$.

$$\lim_{n\to\infty} \text{eff}(\hat{\theta}_1,\, \hat{\theta}_2) = \lim_{n\to\infty} \frac{12n^2}{(n+2)(n+1)^2} = 0$$

Therefore, for $n \in [8, \infty)$, $\hat{\theta}_2$ is a better estimator. Since this covers a wider range of possibilities, $\hat{\theta}_2$ is generally better.

2. Suppose the population has a gamma distribution and we know β but α is unknown. Let $X_1, X_2, \ldots, X_n$ denote a random sample from the distribution. Determine the likelihood function, compute the factorization, and using the Factorization Theorem, show that

$$T = \sum_{i=1}^{n} \ln(X_i)$$

is a sufficient statistic for α.

Solution: The gamma distribution has the probability distribution function

$$f(x) = \frac{\lambda^{\alpha} e^{-x\lambda}}{\Gamma\alpha} x^{\alpha-1} \qquad \lambda = \frac{1}{\beta}$$

Then,

$$\begin{aligned} L(X_1, \ldots, X_n \mid \alpha) &= \prod_{i=1}^{n} f(X_i \mid \alpha) \\ &= \prod_{i=1}^{n} \frac{\lambda^{\alpha} e^{-X\lambda}}{\Gamma\alpha} X^{\alpha-1} \\ &= \left(\frac{\lambda^{\alpha}}{\Gamma\alpha}\right)^{n} \prod_{i=1}^{n} e^{-X\lambda} X^{\alpha-1} \end{aligned}$$

$$g(X_i \mid \alpha) = \left(\frac{\lambda^{\alpha}}{\Gamma\alpha}\right) \prod_{i=1}^{n} X_i^{\alpha-1} \underline{=} (\alpha-1) \sum_{i=1}^{n} \ln X_i$$

And

$$h(X_i) = e^{-\lambda \sum_{i=1}^{n} X_i}$$

Then by the factorization theorem, since $gh = L$,

$$T = \sum_{i=1}^{n} \ln(X_i)$$

3. Let $X_1, X_2, \ldots, X_n$ be iid from a Bernoulli distribution with probability p. We are going to construct the MVUE for variance pq. Recall that in class we showed that $S = \sum X_i$ is a sufficient statistic for p.

   (a) Define the statistic

$$T = \tau(x_1, \ldots, x_n) = \begin{cases} 1 & X_1 = 1 \text{ and } X_2 = 0 \\ 0 & \text{otherwise} \end{cases}$$

   Show that τ is an unbiased estimator for pq.

   Solution: We need to show that $E[\,T\,] = pq$

$$\begin{aligned} E[\,T\,] &= 1 \cdot P(X = 1 \text{ and } X = 0) \\ &= P(X = 1) \cdot P(X = 0) \\ &= p^1 q^0 \cdot p^0 q^1 \\ &= pq \end{aligned}$$

   Therefore T is an unbiased estimator.

   (b) Show that

$$P(T = 1 \mid S = s) = \frac{s(n-s)}{n(n-1)}$$

   Solution: By the definition of a conditional probability,

$$\begin{aligned} P(T = 1 \mid S = s) &= \frac{P(T = 1 \cap S = s)}{P(S = s)} \\ &= P(X_1 = 1,\ X_2 = 0,\ S = \sum_{i=3}^{n} X_i - 1) \\ &= \frac{pq \cdot \binom{n-2}{s-1} p^{s-1} q^{n-s-1}}{\binom{n}{s} p^s q^{n-s}} \\ &= \frac{\binom{n-2}{s-1}}{\binom{n}{s}} \\ &= \frac{(n-2)!}{(s-1)!\,(n-2-(s-1))!} \cdot \frac{s!}{n!\,(n-s)!} \\ &= \frac{s(n-s)}{n(n-1)} \end{aligned}$$

(c) Using the Rao-Blackwell Theorem states that to find an MVUE of pq, we define a new statistic $\phi(s) = E[T \mid S = s]$. Show that

$$\phi(s) = \frac{n}{n-1}[\bar{X}(1-\bar{X})]$$

is the minimum variance unbiased estimator of pq.

Solution:

$$\begin{aligned}
\phi(s) &= E[T \mid S = s] \\
&= 1 \cdot P(T = 1 \mid S = s) + 0 \cdot P(T = 0 \mid S = s) \\
&= \frac{s(n-s)}{n(n-1)} && \text{via part B} \\
&= \frac{\sum X_i(n - \sum X_i)}{n(n-1)} && \text{Substituting } S = \sum X_i \\
&= \frac{\bar{X}(n - \sum X_i)}{n-1} && \text{Substituting } \bar{X} = \frac{1}{n}\sum X_i \\
&= \frac{n \cdot \frac{1}{n}\bar{X}(n - \sum X_i)}{n-1} && \text{multiply by fancy 1: } \frac{n}{n} \\
&= \frac{n\bar{X}(\frac{n}{n} - \frac{1}{n}\sum X_i)}{n-1} && \text{distribution of } \frac{1}{n} \\
&= \frac{n\bar{X}(1-\bar{X})}{n-1} && \text{Substituting } \bar{X} = \frac{1}{n}\sum X_i \\
&= \frac{n}{n-1}[\bar{X}(1-\bar{X})]
\end{aligned}$$

Since S is a sufficient stat for p, $\phi(s)$ is the MVUE by the Rao Blackwell Theorem.

Thanks. #1 was a delightful, albeit time consuming, work of calculus art.