1. A random sample $X_1, \ldots, X_n$ is drawn from a population with pdf

$$f(x|\theta) = \frac{1}{2}(1 + \theta x)\, I_{(-1,1)}(x)$$

for $-1 < \theta < 1$. Find a consistent estimator of θ and show that it is consistent.

Solution:

Note that the mean of the population is given by

$$\begin{aligned} E(X) &= \int_{-1}^{1} \frac{x}{2} + \frac{\theta x^2}{2}\, dx \\ &= \frac{x^2}{4} + \frac{\theta x^3}{6}\bigg|_{-1}^{1} \\ &= \frac{\theta}{3}. \end{aligned}$$

The method of moments estimator of θ is $\widetilde{\Theta} = 3\overline{X}$. Note that this estimator is unbiased by linearity of E. To show that this estimator is consistent, note that the mean squared error is given by

$$Var_\theta(3\overline{X}) - Bias(\theta)^2 = 9\, Var_\theta(\overline{X}) = \frac{9\sigma_\theta^2}{n},$$

where σ_θ^2 is given by the finite integral $\int_{-1}^{1} x^2 f(x|\theta)$. Hence,

$$MSE_{\widetilde{\Theta}}(\theta) = \frac{9\sigma_\theta^2}{n} \to 0 \quad \text{as } n \to \infty.$$

$\square$

2. Let $X_1, \ldots, X_n$ be a random sample from $N(\theta, 1)$. Consider two estimators of $\tau(\theta) = P_\theta(X > 0)$ given by

$$T_1 = \frac{1}{n}\sum_{i=1}^{n} I_{(0,\infty)}(X_i) \quad \text{and} \quad T_2 = \Phi(\widehat{\Theta})$$

where $\widehat{\Theta}$ is MLE of θ.

(a) Show that both estimators are weakly consistent

Solution:

Note that T_1 is unbiased for $\tau(\theta) = P_\theta(X > 0)$ since

$$\begin{aligned}
E(T_1) &= E\left[\frac{1}{n}\sum_{i=1}^{n} I_{(0,\infty)}(X_i)\right] \\
&= \frac{1}{\cancel{n}}\sum_{i=1}^{\cancel{n}} E\left[I_{(0,\infty)}(X)\right] \\
&= P_\theta(X > 0).
\end{aligned}$$

Note also

$$\begin{aligned}
\tau(\theta) &= E_\theta[I_{(0,\infty)}(X)] = P_\theta(X > 0) \\
&= 1 - P_\theta(X - \theta \leq -\theta) \\
&= 1 - \Phi(-\theta) \\
&= \Phi(\theta).
\end{aligned}$$

So,

$$\begin{aligned}
MSE_{T_1}(\theta) &= Var_\theta(T_1) \\
&= E_\theta\left[\frac{1}{n}\sum_{i=1}^{n} I_{(0,\infty)}(X_i)\right]^2 - \left(E_\theta \frac{1}{n}\sum_{i=1}^{n} I_{(0,\infty)}(X_i)\right)^2 \\
&= \frac{1}{n^2}\sum_{i=j} E_\theta\left[I_{(0,\infty)}(X_i)^2\right] + \frac{1}{n^2}\sum\sum_{i\neq j} E_\theta\left[I_{(0,\infty)}(X_i) I_{(0,\infty)}(X_j)\right] - \left[\frac{1}{n}\sum_{i=1}^{n} E_\theta(X_i)\right]^2 \\
&= \frac{1}{n}\Phi(\theta) + \frac{n(n-1)}{n^2}[\Phi(\theta)]^2 - [\Phi(\theta)]^2 \quad \text{by independence in the second term} \\
&= \frac{1}{n}\left[\Phi(\theta) - (\Phi(\theta))^2\right] \to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

So T_1 is strongly consistent, and thus, weakly consistent.

For T_2, recall that the MLE for a normal distribution is given by $\widehat{\Theta} = \overline{X}$. By the weak law of large numbers we have that $\overline{X} \xrightarrow{p} \theta$, and since Φ is continuous $\Phi(\overline{X}) \xrightarrow{p} \Phi(\theta)$. Hence $\widehat{\Theta}$ is weakly consistent.

$\square$

(b) Are both estimators asymptotically efficient? Justify your answers.

Solution:

Let us first calculate the quantity

$$\begin{aligned} v(\theta) &= \frac{(\tau'(\theta))^2}{I(\theta)} \\ &= \frac{\phi(\theta)}{E_\theta(X-\theta)^2} \\ &= \frac{\phi(\theta)}{Var_\theta(X)} \\ &= \phi(\theta). \end{aligned}$$

Now, let us calculate the asymptotic distributions of both T_1 and T_2. For T_1, note that $E_\theta(I_{(0,1)}(X_i)) = \Phi(\theta)$ and $Var_\theta(I_{(0,\infty)}(X)) = \Phi(\theta) - (\Phi(\theta))^2$, so by the CLT

$$\sqrt{n}(T_1 - \Phi(\theta)) \mathrel{\dot{\underset{\cdot}{\sim}}} N(0, \Phi(\theta) - (\Phi(\theta))^2).$$

Note that $\Phi(0) - \Phi(0)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$ and $\phi(0) = \frac{1}{\sqrt{2\pi}}$, so these distributions do not coincide.

For $T_2 = \overline{X}$, note that

$$\sqrt{n}(\overline{X} - \theta) \mathrel{\dot{\underset{\cdot}{\sim}}} N(0, 1),$$

so by the Δ-method

$$\sqrt{n}(\Phi(\overline{X}) - \Phi(\theta)) \mathrel{\dot{\underset{\cdot}{\sim}}} N(0, \phi(\theta)).$$

Hence $\widehat{\Theta}$ is asymptotically efficient.

$\square$

3. Let $X_1, \ldots, X_n$ be a random sample from the density

$$f(x|\theta) = \theta(1+x)^{-(1+\theta)} I_{(0,\infty)}(x) \quad \text{for} \quad \theta > 0.$$

The MLE for θ is

$$\widehat{\Theta} = \frac{n}{\sum_{i=1}^n \log(1+X_i)}.$$

Show that this estimator is consistent and asymptotically efficient.

Solution:

Let $Y_i = \log(1+X_i)$. Then $Y_i \sim \frac{dx}{dy} f(x(y)|\theta)$ where

$$\frac{dx}{dy} f(x(y)|\theta) = \frac{d}{dy}\left[e^y - 1\right] \theta e^{-y(1+\theta)} I_{(0,\infty)}(e^y - 1) = \theta e^{-\theta} I_{(0,\infty)}(y).$$

This is an exponential density so $\widehat{\Theta} = 1/\overline{Y}$. Hence

$$\begin{aligned} E_\theta(\widehat{\Theta}) &= \int_0^\infty \frac{1}{y} \cdot \frac{(n\theta)^n}{\Gamma(n)} y^{n-1} e^{-n\theta y} dy \\ &= \frac{\Gamma(n-1)(n\theta)^n}{\Gamma(n)(n\theta)^{n-1}} \cdot 1 \\ &= \frac{n\theta}{n-1}, \end{aligned}$$

and similarly

$$\begin{aligned} Var_\theta(\widehat{\Theta}) &= \frac{\Gamma(n-2)(n\theta)^n}{\Gamma(n)(n\theta)^{n-2}} - \left(\frac{n\theta}{n-1}\right)^2 \\ &= \frac{(n\theta)^2}{(n-1)(n-2)} - \left(\frac{n\theta}{n-1}\right)^2 \\ &= \frac{(n\theta)^2}{(n-1)^2(n-2)}. \end{aligned}$$

So the MSE is given by

$$\begin{aligned} MSE_{\widehat{\Theta}}(\theta) &= Var_\theta(\widehat{\Theta}) - Bias_\theta(\widehat{\Theta})^2 \\ &= \frac{(n\theta)^2}{(n-1)^2(n-2)} - \theta^2\left(\frac{n}{n-1} - 1\right)^2. \end{aligned}$$

Note that each term goes to 0 as $n \to \infty$, hence the statistic is strongly consistent.

For asymptotic efficiency, recall for the exponential distribution $E_\theta(Y_i) = \frac{1}{\theta}$ and $Var_\theta(Y_i) = \frac{1}{\theta^2}$. We invoke the Δ-method with $g(\overline{Y}) = 1/\overline{Y}$ and $\frac{1}{\theta} g'(1/\theta) = -\theta$ to obtain

$$\sqrt{n}\left(1/\overline{Y} - \theta\right) \dot{\underset{\cdot}{\sim}} N(0, \theta^2).$$

The Fisher-information is

$$\begin{aligned} I(\theta) &= -E_\theta \left[\frac{\partial^2}{\partial \theta^2} f(X|\theta) \right] \\ &= -E_\theta \left[\frac{-1}{\theta^2} \right] \\ &= \frac{1}{\theta^2}. \end{aligned}$$

So, the asymptotic variance implied by the Cramér-Rao is

$$v(\theta) = \frac{(\tau'(\theta))^2}{I(\theta)} = \theta$$

This matches the asymptotic variance of $\widehat{\Theta}$, hence the statistic is asymptotically efficient.

$\square$

4. Let $Z_1, \ldots, Z_n$ be a random sample from $N(0, \theta^2), \theta >)$. Define $X_i = |Z_i|$. Consider two estimators of θ^2 given by

$$T_1 = \frac{\sum X_i^2}{n} \quad \text{and} \quad T_2 = \frac{\sum X_i^2}{n+2}.$$

(a) Show that both estimators are consistent.

Solution:

Recall from homework 5 that the family of estimators

$$T_c = c \sum_{i=1}^{n} X_i^2$$

has

$$E_\theta(T_c) = cn\theta^2 \quad \text{and} \quad Var_\theta(T_c) = 2nc^2\theta^4$$

by recognizing that $\sum \frac{X_i}{\theta} \sim \chi_n^2$. Hence

$$\begin{aligned} MSE_{T_1}(\theta) &= 2n \left(\frac{1}{n^2} \right) \theta^4 - \cancel{(\theta^2 - \theta^2)^2} \\ &= \frac{2\theta^4}{n} \to 0 \quad \text{as } n \to \infty, \end{aligned}$$

and

$$\begin{aligned} MSE_{T_2}(\theta) &= 2n \left(\frac{1}{n+2} \right)^2 \theta^4 - \left(\frac{n\theta^2}{n+2} - \theta^2 \right)^2 \\ &= \theta^4 \frac{2n}{(n+2)^2} - \theta^4 \left(\frac{n}{n+2} - 1 \right)^2 \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

The second convergence follows from the fact that the quadratic denominator of the first term dominates the linear numerator, and in the second term $n/(n+2) \to 1$ as $n \to \infty$.

□

(b) Find the asymptotic distribution of T_1.

Solution:

Note that $\frac{X_1^2}{\theta^2} \sim \chi_1^2$, so by CLT

$$\sqrt{n}(\overline{X} - \theta^2) \dot{\underset{\cdot}{\sim}} N(0, 2\theta^4).$$

□

5. MGB VIII: 1[b]. Let X be a single observation from the density

$$f(x|\theta) = \theta x^{\theta-1} I_{(0,1)}(x),$$

where $\theta > 0$. Show that $(Y/2, Y)$ is a confidence interval for θ. Find its confidence coefficient. Also, find a better confidence interval for θ. Define $Y = -1/\log X$.

Solution:

Note that

$$\frac{Y}{2} < \theta < Y \iff 1 < \frac{Y}{\theta} < 2.$$

$(T_1, T_2) = (Y/2, Y)$ is a confidence interval provided the following integral is free from θ,

$$\begin{aligned} P(\theta < Y < 2\theta) &= \int_\theta^{2\theta} f_Y(y|\theta)dy \\ &= \int_\theta^{2\theta} \frac{dx}{dy} \cdot f_X(x(y)|\theta)dy \\ &= \int_\theta^{2\theta} \frac{e^{-1/y}}{y^2} \theta e^{(1-\theta)/y} \\ &= \int_\theta^{2\theta} \theta e^{-\theta/y} dy \\ &= e^{-1/2} - e^{-1} \end{aligned}$$

Hence Y/θ is pivotal, and $(Y/2, Y)$ is a $(e^{-1/2} - e^{-1}) \cdot 100\%$ confidence interval. Since Y/θ is pivotal, another confidence interval is given by any q_1, q_2 satisfying

$$\int_{q_1\theta}^{q_2\theta} \theta e^{-\theta/y} dy = e^{-q_1} - e^{-q_2} \stackrel{\dagger}{=} e^{-1/2} - e^{-1}.$$

The expected width of the interval is given by $E(Y)(q_1 - q_2)$ which we can minimize subject to †. I.e.

$$q_2 = -\log\left(e^{-q_1} + e^{-1} - e^{-1/2}\right)$$

and

$$\begin{aligned} \frac{d}{dq_1}(q_1 - q_2) &= 1 - \frac{dq_2}{dq_1} \\ &= 1 + \frac{-e^{-q_1}}{e^{-q_1} + e^{-1} - e^{-1/2}} \\ &= \frac{e^{-1} - e^{-1/2}}{e^{-q_1} + e^{-1} - e^{-1/2}}. \end{aligned}$$

This derivative is always negative, hence the width is decreasing with respect to q_1. Thus, $q_1 = 0$ and $q_2 = -\log\left(1 + e^{-1} - e^{-1/2}\right)$ which is also the minimum expected width for the pivotal statistic $1/Y$. Note $q_2 \approx 0.2726637 < 1/2$.

□

6. MGB VIII 4. Let $X_1, \ldots, X_n$ be a random sample from $f(x|\theta) = I_{(\theta-1/2,\theta+1/2)}(x)$. Let $Y_1 < \cdots < Y_n$ be the corresponding ordered sample. Show that (Y_1, Y_n) is a confidence interval for θ. Find its confidence coefficient.

Solution:

We calculate

$$\begin{aligned}
P(Y_1 < \theta < Y_n) &= P(\theta < Y_n) - P(\theta < Y_n \text{ and } \theta \leq Y_1) && \text{since } A \cap B = A \setminus (A \cap B^c) \\
&= P(\theta < Y_n) - (\theta \leq Y_1) && \text{since } \theta \leq Y_1 \implies \theta \leq Y_n \\
&= P(Y_1 \leq \theta) - P(Y_n \leq \theta) \\
&= 1 - \big[1 - F(\theta)\big]^n - \big[F(\theta)\big]^n \\
&= 1 - \big[1 - (\theta - \theta + 1/2)\big]^n - \big[\theta - \theta + 1/2\big]^n && \text{since } F(x) = \theta - x + 1/2 \\
&= 1 - \left[\frac{1}{2}\right]^{n-1}.
\end{aligned}$$

Thus (Y_1, Y_n) is a confidence interval with confidence $1 - (1/2)^{n-1}$. Note that the confidence approaches 1 as $n \to \infty$.

□