Sufficient, Complete, and UMVU Estimators

This problem comes from the Fall 2009 pre-qual.

Let $\underline{X} = (X_1, \ldots, X_n)^T$ from Uniform$(\theta_1 - \theta_2, \theta_1 + \theta_2)$ where $\theta_1 \in \mathcal{R}$ and $\theta_2 > 0$ are unknown. Consider estimating θ_1, θ_2.

(a) Show $T(\underline{X}) = (X_{(1)}, X_{(n)})$ is sufficient and complete.

Sufficiency: Using the Fisher-Neyman Factorization,

$$\begin{aligned}
f(\underline{X}) = \prod_{i=1}^{n} f(x_i) &== \left(\frac{1}{(\theta_1+\theta_2)-(\theta_1-\theta_2)}\right)^n \cdot \prod \mathbf{1}(x_i > \theta_1 - \theta_2) \cdot \mathbf{1}(x_i < \theta_1 + \theta_2)) \\
&= \left(\frac{1}{2\theta_2}\right)^n \mathbf{1}(X_{(1)} > \theta_1 - \theta_2) \cdot \mathbf{1}(X_{(n)} < \theta_1 + \theta_2)
\end{aligned}$$

Noting $h(X) = 1$ and $g(\theta, T(\underline{X})$ is everything else, we see we can split the joint pdf into a piece that depends only on data, and a piece where the random variable only interacts with the parameters through the sufficient statistics, $T(\underline{X}) = (X_{(1)}, X_{(n)})$.

Completeness: Recall that a statistic is complete iff

$$E_\theta(g(T(\underline{X}))) = 0, \ \ \forall \theta \ \Rightarrow \ g(T(\underline{X}) = 0 \text{ a.s.}$$

For ease of notation, let $R = X_{(n)} - X_{(1)}$. The joint pdf of the min and max is given by

$$\frac{n!}{(n-2)!}\frac{1}{(2\theta_2)^2}\left(\frac{R}{2\theta_2}\right)^{n-2}$$

$$\begin{aligned}
E_\theta(g(T(\underline{X}))) &= \int_{\underline{\theta}} \frac{n\cdot(n-1)}{(2\theta_2)^n} \cdot R^{n-2} g(X_{(1)}, X_{(n)}) dT(\underline{X}) = 0 \\
&\Rightarrow \int_{\underline{\theta}} R^{n-2} g(X_{(1)}, X_{(n)}) dT(\underline{X}) = 0 \\
&\Rightarrow \int_{\underline{\theta}} \left(R^{n-2} g(X_{(1)}, X_{(n)})\right)^+ dT(\underline{X}) = \int_{\underline{\theta}} \left(R^{n-2} g(X_{(1)}, X_{(n)})\right)^- dT(\underline{X})
\end{aligned}$$

Since both integrands are positive, this suggests that

$$\left(R^{n-2} g(X_{(1)}, X_{(n)})\right)^+ dT(\underline{X}) = \left(R^{n-2} g(X_{(1)}, X_{(n)})\right)^-,$$

which will occur only when $R^{n-2}g(X_{(1)}, X_{(n)}) = 0$ almost surely. Since the range is a positive value, this implies that $g(T(\underline{X}) = 0$ almost surely. This proves that the min and max are complete.

Show that

$$E(X_{(1)}) = \theta_1 - \theta_2 + \frac{2\theta_2}{n+1}$$

and

$$E(X_{(n)}) = \theta_1 + \theta_2 - \frac{2\theta_2}{n+1}$$

It can be shown that

$$f(X_{(1)}) = \frac{n}{2\theta_2}\left(1 - \frac{X_{(1)} - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n-1}$$

$$f(X_{(n)}) = \frac{n}{(2\theta_2)^n}\left(X_{(n)} - (\theta_1 - \theta_2)\right)^{n-1}$$

$$\begin{aligned}
E(X_{(1)}) &= \int_{\theta_1-\theta_2}^{\theta_1+\theta_2} \frac{nx}{2\theta_2}\left(1 - \frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n-1} dx \\
&= \left[-x\left(1 - \frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n}\right]_{\theta_1-\theta_2}^{\theta_1+\theta_2} + \int_{\theta_1-\theta_2}^{\theta_1+\theta_2} \left(1 - \frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n} dx \\
&= \theta_1 - \theta_2 - \frac{2\theta_2}{n+1}\left[\left(1 - \frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n+1}\right]_{\theta_1-\theta_2}^{\theta_1+\theta_2} \\
&= \theta_1 - \theta_2 + \frac{2\theta_2}{n+1} \\
E(X_{(n)}) &= \int_{\theta_1-\theta_2}^{\theta_1+\theta_2} \frac{nx}{2\theta_2}\left(\frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n-1} dx \\
&= \left[x\left(\frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n}\right]_{\theta_1-\theta_2}^{\theta_1+\theta_2} - \int_{\theta_1-\theta_2}^{\theta_1+\theta_2} \left(\frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n} dx \\
&= \theta_1 + \theta_2 - \frac{2\theta_2}{n+1}\left[\left(\frac{x - (\theta_1 - \theta_2)}{2\theta_2}\right)^{n+1}\right]_{\theta_1-\theta_2}^{\theta_1+\theta_2} \\
&= \theta_1 + \theta_2 - \frac{2\theta_2}{n+1}
\end{aligned}$$

Find the UMVUE for θ_1 and θ_2.

Find, we find unbiased estimators for θ_1 and θ_2 using the sample min and max.

$$E(X_{(1)} + X_{(n)}) = \left(\theta_1 - \theta_2 + \frac{2\theta_2}{n+1}\right) + \left(\theta_1 + \theta_2 - \frac{2\theta_2}{n+1}\right) = 2\theta_1$$

$$\Rightarrow E\left[\frac{(X_{(1)}+X_{(n)})}{2}\right]=\theta_1$$

$$E(X_{(n)}-X_{(1)})=\left(\theta_1+\theta_2-\frac{2\theta_2}{n+1}\right)-\left(\theta_1-\theta_2+\frac{2\theta_2}{n+1}\right)=2\theta_2-\frac{4\theta_2}{n+1}$$

$$\Rightarrow E\left[\frac{(n+1)\left(X_{(n)}-X_{(1)}\right)}{2n-2}\right]=\theta_2$$

Since they depend only on sufficient and complete statistics, the Lehmann-Scheffé Theorem tells us that the two estimators are UMVUE for θ_1, θ_2.