Remarks · An UMVUE does not always exist, even when unbiased estimators do.

A useful first step is to find a lower bound for the variance of unbiased estimators of $\tau(\theta)$.

Suppose that $X_1, \dots, X_n$ has joint pdf/pmf $f(\underline{x}|\theta)$ where $X_1, \dots, X_n$ are not necessarily independent. Consider the following regularity conditions.

1. Θ is an open interval on the real line
2. $\frac{\partial}{\partial\theta} \log f(\underline{x}|\theta)$ exists and is finite $\forall\, \theta \in \Theta$ and all $\underline{x} \in \underline{\mathcal{X}}$ (joint support)
3. $\frac{\partial}{\partial\theta} \int f(\underline{x}|\theta)\, d\underline{x} = \int \frac{\partial}{\partial\theta} f(\underline{x}|\theta)\, d\underline{x}$ (for the continuous case, with analogous condition for discrete case.)
4. $\frac{\partial}{\partial\theta} \int W(\underline{x}) f(\underline{x}|\theta)\, d\underline{x} = \int W(\underline{x}) \frac{\partial}{\partial\theta} f(\underline{x}|\theta)\, d\underline{x}$ for $W(\underline{X})$ an unbiased estimator of $\tau(\theta)$.
5. $0 < E_\theta\left[\left(\frac{\partial}{\partial\theta} \log f(\underline{X}|\theta)\right)^2\right] < \infty \quad \forall\, \theta \in \Theta$.

Theorem 7.3.9 (Cramér-Rao Inequality)

under conditions 1.-5.,

$$Var_\theta(W(\underline{X})) \geq \frac{(\tau'(\theta))^2}{E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\underline{X}|\theta)\right)^2\right]}$$

Cramér-Rao Lower Bound (CRLB)

when $W(\underline{X})$ is an unbiased estimator of $\tau(\theta)$.

proof: Omitted.

Corollary 7.3.10 (iid case)

If $X_1, \dots, X_n \overset{iid}{\sim} f(x|\theta)$, then under the regularity conditions,

$$Var_\theta(W(\underline{X})) \geq \frac{(\tau'(\theta))^2}{n\, E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right]}.$$

Def : $I(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\underline{X}|\theta)\right)^2\right]$, assuming $0 < I(\theta) < \infty$, is called Fisher's Information in $\underline{X}$ concerning θ.

It follows from Theorem 7.3.9 that

$$Var_\theta(W(\underline{X})) \geq \frac{(\tau'(\theta))^2}{I(\theta)}.$$

Lemma 7.3.11 If $f(x|\theta)$ is twice differentiable wrt θ and

Result holds for random variable X or random vector $\underline{X}$.

$$\frac{d}{d\theta} E_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) = \int \frac{\partial}{\partial\theta}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\cdot f(x|\theta)\right]dx,$$

then $E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right] = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right)$.

Remarks: 2. If the range of $\underline{X}$ depends on θ, (as is the case for Uniform $[\theta-\frac{1}{2},\theta+\frac{1}{2}]$,) then $f(X|\theta)$ does not satisfy the regularity conditions.

3. If $f(x|\theta)$ is a pdf/pmf in the exponential family, then the regularity conditions are satisfied.

4. Lemma 7.3.11 holds for any $f(x|\theta)$ in the exponential family.

Ex: $X_1,\dots,X_n \overset{iid}{\sim}$ Bernoulli(θ), $0\le\theta\le 1$.

Estimate θ. Can we find an UMVUE of θ?

Recall that $\tilde{\theta} = \hat{\theta} = \bar{X}$ and $E_\theta(\bar{X}) = \theta$, $Var_\theta(\bar{X}) = \frac{\theta(1-\theta)}{n}$.

We have $f(\underline{x}|\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\prod_{i=1}^n I_{\{0,1\}}(x_i)$ so

$$\frac{\partial}{\partial\theta}\log f(\underline{X}|\theta) = \frac{\partial}{\partial\theta}\left(\sum_{}^{n} x_i\log\theta + (n-\sum x_i)\log(1-\theta) + \log\prod_{i=1}^n I_{\{0,1\}}(x_i)\right)$$

$$= \frac{\sum^n x_i}{\theta} - \frac{n-\sum x_i}{1-\theta} = \frac{n(\bar{x}-\theta)}{\theta(1-\theta)}.$$

Then $E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\underline{X}|\theta)\right)^2\right] = E_\theta\left[\frac{n^2(\bar{X}-\theta)^2}{\theta^2(1-\theta)^2}\right] = \frac{n^2}{\theta^2(1-\theta)^2}\frac{\theta(1-\theta)}{n} = \frac{n}{\theta(1-\theta)}$,

Also, $\tau(\theta)=\theta$ so $\tau'(\theta)=1$ and for any unbiased estimator $W(\underline{X})$ of θ,

$$Var_\theta(W(\underline{X})) \ge \frac{1}{\frac{n}{\theta(1-\theta)}} = \frac{\theta(1-\theta)}{n}.$$

The CRLB coincides with $Var_\theta(\bar{X})$.

i.e., $\bar{X}$ is an UMVUE of θ. (Is it unique?)

If $\tau(\theta)=\theta(1-\theta)$, then $\tau'(\theta)=1-2\theta$ and for any unbiased estimator $W(\underline{X})$ of $\tau(\theta)$, $Var_\theta(W(\underline{X})) \ge \frac{(1-2\theta)^2\,\theta(1-\theta)}{n}$. Is there an UMVUE of $\theta(1-\theta)$?

Sufficiency and Unbiasedness (Section 7.3.3)

Theorem 7.3.17 (Rao-Blackwell Theorem)
If W is an unbiased estimator of $\tau(\theta)$ and T is a sufficient statistic for θ, then $\phi(T) = E(W|T)$ is an unbiased estimator of $\tau(\theta)$ and $Var_\theta\, \phi(T) \le Var_\theta(W)$ $\forall\, \theta$. i.e., $\phi(T)$ is a uniformly better estimator of $\tau(\theta)$ than W.

proof: Since T is sufficient and $E(W|T)$ does not depend on θ, it follows that $\phi(T) = E(W|T)$ is an estimator. Now

$$E_\theta(\phi(T)) = E_\theta(E(W|T)) = E_\theta(W) = \tau(\theta).$$

$$\begin{aligned} Var_\theta(W) &= Var_\theta(E(W|T)) + E_\theta(Var(W|T)) \\ &= Var_\theta(\phi(T)) + E_\theta(Var(W|T)) \\ &\ge Var_\theta(\phi(T)). \end{aligned}$$

Remark: The Rao-Blackwell Theorem gives us an improved unbiased estimator of $\tau(\theta)$, but does it give us the UMVUE?

Ex: $X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$, $0 \le \theta \le 1$.
Recall that a sufficient statistic for θ is $T(\underset{\sim}{X}) = \sum_{i=1}^n X_i$.
Note that $E_\theta(X_1(1-X_2)) = E_\theta(X_1)\, E_\theta(1-X_2) = \theta(1-\theta)$ so $X_1(1-X_2)$ is an unbiased estimator of $\theta(1-\theta)$. Applying the Rao-Blackwell Theorem

$$E(X_1(1-X_2) \mid T=t) = \frac{P_\theta(X_1=1, X_2=0, \sum_1^n X_i = t)}{P_\theta(\sum_1^n X_i = t)}$$

$$= \frac{P_\theta(X_1=1, X_2=0, X_3+\dots+X_n = t-1)}{P_\theta(\sum_1^n X_i = t)}$$

$$= \frac{\theta(1-\theta)\binom{n-2}{t-1}\theta^{t-1}(1-\theta)^{(n-2)-(t-1)}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} = \frac{\binom{n-2}{t-1}}{\binom{n}{t}} = \frac{t(n-t)}{n(n-1)}.$$

That is, $E(X_1(1-X_2)|T) = \dfrac{T(n-T)}{n(n-1)} = \dfrac{\sum_1^n X_i(n - \sum_1^n X_i)}{n(n-1)} = \dfrac{n}{n-1}\bar{X}(1-\bar{X})$.

Note that $\dfrac{n}{n-1}\bar{X}(1-\bar{X}) = S^2$ and one can show that

$E_\theta(S^2) = \sigma^2 = \theta(1-\theta)$.

$Var_\theta(S^2) = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right) = \frac{\theta(1-\theta)}{n}\left[1 - \theta(1-\theta)\frac{4n-6}{n-1}\right] \ge \text{CRLB}$

Question: Can we do better?

Theorem 7.3.19 : If an UMVUE of $\tau(\theta)$ exists, then the UMVUE is unique.

Theorem 7.3.23 (Lehmann - Scheffé Theorem).

If T is a complete sufficient statistic for θ and if $\phi(T)$, a function of T, is an unbiased estimator of $\tau(\theta)$, then $\phi(T)$ is the UMVUE of $\tau(\theta)$.

proof: Let W be any other unbiased estimator of $\tau(\theta) = E_\theta(\phi(T))$. It suffices to show that $Var_\theta(\phi(T)) \le Var_\theta(W)\ \forall\ \theta$. Applying the Rao Blackwell Theorem to W results in an unbiased estimator $W^* = E(W|T)$ and $Var_\theta(W^*) \le Var_\theta(W)$ $\forall\ \theta$. Also $0 = E_\theta(\phi(T) - W^*)$ and $\phi(T) - W^*$ is a function of T so by the completeness of T, $P_\theta(\phi(T) = W^*) = 1$, so $Var_\theta(\phi(T)) = Var_\theta(W^*)$.

Ex: $X_1, \ldots, X_n \overset{iid}{\sim}$ Bernoulli (θ), $0 \le \theta \le 1$.

Recall that $T = \sum_{i=1}^n X_i = n\bar{X}$ is a complete sufficient statistic for θ.

Since $E_\theta(S^2) = E_\theta\left(\frac{n}{n-1} \bar{X}(1-\bar{X})\right) = \theta(1-\theta)$, the sample variance S^2 is the UMVUE of the population variance $\theta(1-\theta)$.

Note that the variance of the UMVUE does not attain the CRLB.

Ex: $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, where $\underset{\sim}{\theta} = (\mu, \sigma^2)$ and $\tau(\theta) = \mu$.

Recall that $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ is complete and sufficient for for $\underset{\sim}{\theta}$ by the Theorem 6.2.25 and the Factorization Theorem. Furthermore, $\bar{X}$ is a function of of $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ and $E_{\underset{\sim}{\theta}}(\bar{X}) = \mu$ so $\bar{X}$ is UMVUE by the Lehmann- Scheffé Theorem.

Ex: $X_1, \ldots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$, $\lambda > 0$.

$$f(\underline{x}|\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} I_{\{0,1,\ldots\}}(x_i) = \prod_{i=1}^{n}\left(\frac{1}{x_i!} I_{\{0,1,\ldots\}}(x_i)\right) e^{-\lambda n} \cdot e^{(\log\lambda)\sum_{i=1}^{n} x_i}$$

$$= h(\underline{x})c(\lambda)\, e^{(\log\lambda)\sum_{i=1}^{n} x_i}$$ So by Theorem 6.2.25 and Factorization Theorem it follows that $\sum_{i=1}^{n} X_i$ is complete sufficient for λ. To find the UMVUE of λ itself, it suffices to find a function of $\sum_{i=1}^{n} X_i$ whose expectation is λ. It obvious that $E_\lambda(\bar{X}) = \lambda$ so by the Lehmann-Scheffe' Theorem, $\bar{X}$ is UMVUE for λ.

Now considering estimating $\tau(\lambda) = e^{-\lambda} = P(X_i = 0)$.

Let's derive the UMVUE of $e^{-\lambda}$ by calculating the conditional expectation of some unbiased estimator given the complete sufficient statistic $\sum_{i=1}^{n} X_i$.

Note that $I_{\{0\}}(X_1)$ is unbiased for $e^{-\lambda}$.

$$I_{\{0\}}(X_1) = \begin{cases} 1, & \text{if } X_1 = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E_\lambda\left(I_{\{0\}}(X_1)\right) = 1 \cdot P(X_1 = 0) = e^{-\lambda}$$

$$E\left(I_{\{0\}}(X_1) \,\middle|\, \sum_{i=1}^{n} X_i = t\right) = 1 \cdot P\left(X_1 = 0 \,\middle|\, \sum_{i=1}^{n} X_i = t\right)$$

$$= \frac{P\left(X_1 = 0, \sum_{i=2}^{n} X_i = t\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)} = \frac{P(X_1 = 0)\, P\left(\sum_{i=2}^{n} X_i = t\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)}$$

Recall that

$\sum_{i=2}^{n} X_i \sim \text{Poisson}((n-1)\lambda)$

$\sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$

$$= \frac{e^{-\lambda} \cdot e^{-(n-1)\lambda}[(n-1)\lambda]^t/t!}{e^{-n\lambda}(n\lambda)^t/t!} = \left(\frac{n-1}{n}\right)^t \text{ for } n > 1.$$

It follows that the UMVUE of $e^{-\lambda}$ is $\left(\frac{n-1}{n}\right)^{\sum_{i=1}^{n} X_i}$ for $n > 1$ (and simply $I_{\{0\}}(X_1)$ for $n=1$.)

Remark: The Cramér-Rao Inequality, Rao-Blackwell Theorem, and Lehmann-Scheffe' Theorem for $\tau(\theta)$ in the unidimensional case can be extended to the higher dimensional case.