Chapter 6

UMVUEs and the FCRLB

Warning: UMVUE theory is rarely used in practice unless the UMVUE U_n of θ satisfies $U_n = a_n \hat{\theta}_{MLE}$ where a_n is a constant that could depend on the sample size n. UMVUE theory tends to be somewhat useful if the data is iid from a 1P–REF.

6.1 MSE and Bias

Definition 6.1. Let the sample $\boldsymbol{Y} = (Y_1, ..., Y_n)$ where $\boldsymbol{Y}$ has a pdf or pmf $f(\boldsymbol{y}|\boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \Theta$. Assume all relevant expectations exist. Let $\tau(\boldsymbol{\theta})$ be a real valued function of $\boldsymbol{\theta}$, and let $T \equiv T(Y_1, ..., Y_n)$ be an estimator of $\tau(\boldsymbol{\theta})$. The **bias** of the estimator T for $\tau(\boldsymbol{\theta})$ is

$$B(T) \equiv B_{\tau(\boldsymbol{\theta})}(T) \equiv \text{Bias(T)} \equiv \text{Bias}_{\tau(\boldsymbol{\theta})}(\text{T}) = \text{E}_{\boldsymbol{\theta}}(\text{T}) - \tau(\boldsymbol{\theta}). \tag{6.1}$$

The *mean squared error* (**MSE**) of an estimator T for $\tau(\boldsymbol{\theta})$ is

$$\text{MSE(T)} \equiv \text{MSE}_{\tau(\boldsymbol{\theta})}(\text{T}) = \text{E}_{\boldsymbol{\theta}}[(\text{T} - \tau(\boldsymbol{\theta}))^2]$$

$$= Var_{\boldsymbol{\theta}}(T) + [\text{Bias}_{\tau(\boldsymbol{\theta})}(\text{T})]^2. \tag{6.2}$$

T is an *unbiased estimator* of $\tau(\boldsymbol{\theta})$ if

$$E_{\boldsymbol{\theta}}(T) = \tau(\boldsymbol{\theta}) \tag{6.3}$$

for all $\boldsymbol{\theta} \in \Theta$. Notice that $\text{Bias}_{\tau(\boldsymbol{\theta})}(\text{T}) = 0$ for all $\boldsymbol{\theta} \in \Theta$ if T is an unbiased estimator of $\tau(\boldsymbol{\theta})$.

Notice that the bias and MSE are functions of $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \Theta$. If $MSE_{\tau(\boldsymbol{\theta})}(T_1) < MSE_{\tau(\boldsymbol{\theta})}(T_2)$ for all $\boldsymbol{\theta} \in \Theta$, then T_1 is "a better estimator" of $\tau(\boldsymbol{\theta})$ than T_2. So estimators with small MSE are judged to be better than ones with large MSE. Often T_1 has smaller MSE than T_2 for some $\boldsymbol{\theta}$ but larger MSE for other values of $\boldsymbol{\theta}$.

Often θ is real valued. A common problem considers a class of estimators $T_k(\boldsymbol{Y})$ of $\tau(\theta)$ where $k \in \Lambda$. Find the MSE as a function of k and then find the value $k_o \in \Lambda$ that is the global minimizer of $MSE(k) \equiv MSE(T_k)$. This type of problem is a lot like the MLE problem except you need to find the global min rather than the global max.

This type of problem can often be done if $T_k = kW_1(\boldsymbol{X}) + (1-k)W_2(\boldsymbol{X})$ where both W_1 and W_2 are unbiased estimators of $\tau(\theta)$ and $0 \le k \le 1$.

Example 6.1. If $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$ then $k_o = n+1$ will minimize the MSE for estimators of σ^2 of the form

$$S^2(k) = \frac{1}{k}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

where $k > 0$. See Problem 6.2.

Example 6.2. Find the bias and MSE (as a function of n and c) of an estimator $T = c\sum_{i=1}^n Y_i$ or ($T = b\overline{Y}$) of θ when $Y_1, ..., Y_n$ are iid with $E(Y_1) = \mu = h(\theta)$ and $V(Y_i) = \sigma^2$.
Solution: $E(T) = c\sum_{i=1}^n E(Y_i) = nc\mu$, $V(T) = c^2\sum_{i=1}^n V(Y_i) = nc^2\sigma^2$, $B(T) = E(T) - \theta$ and $MSE(T) = V(T) + [B(T)]^2$. (For $T = b\overline{Y}$, use $c = b/n$.)

Example 6.3. Suppose that $Y_1, ..., Y_n$ are independent binomial(m_i, ρ) where the $m_i \ge 1$ are known constants. Let

$$T_1 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n m_i} \quad \text{and} \quad \mathrm{T_2} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\frac{\mathrm{Y_i}}{\mathrm{m_i}}$$

be estimators of ρ.

a) Find MSE(T_1).

b) Find MSE(T_2).

c) Which estimator is better?

Hint: by the arithmetic–geometric–harmonic mean inequality,

$$\frac{1}{n}\sum_{i=1}^{n} m_i \geq \frac{n}{\sum_{i=1}^{n}\frac{1}{m_i}}.$$

Solution: a)

$$E(T_1) = \frac{\sum_{i=1}^{n} E(Y_i)}{\sum_{i=1}^{n} m_i} = \frac{\sum_{i=1}^{n} m_i \rho}{\sum_{i=1}^{n} m_i} = \rho,$$

so $\text{MSE}(T_1) = V(T_1) =$

$$\frac{1}{(\sum_{i=1}^{n} m_i)^2} V(\sum_{i=1}^{n} Y_i) = \frac{1}{(\sum_{i=1}^{n} m_i)^2}\sum_{i=1}^{n} V(Y_i) = \frac{1}{(\sum_{i=1}^{n} m_i)^2}\sum_{i=1}^{n} m_i\rho(1-\rho)$$

$$= \frac{\rho(1-\rho)}{\sum_{i=1}^{n} m_i}.$$

b)

$$E(T_2) = \frac{1}{n}\sum_{i=1}^{n}\frac{E(Y_i)}{m_i} = \frac{1}{n}\sum_{i=1}^{n}\frac{m_i\rho}{m_i} = \frac{1}{n}\sum_{i=1}^{n}\rho = \rho,$$

so $\text{MSE}(T_2) = V(T_2) =$

$$\frac{1}{n^2} V(\sum_{i=1}^{n}\frac{Y_i}{m_i}) = \frac{1}{n^2}\sum_{i=1}^{n} V(\frac{Y_i}{m_i}) = \frac{1}{n^2}\sum_{i=1}^{n}\frac{V(Y_i)}{(m_i)^2} = \frac{1}{n^2}\sum_{i=1}^{n}\frac{m_i\rho(1-\rho)}{(m_i)^2}$$

$$= \frac{\rho(1-\rho)}{n^2}\sum_{i=1}^{n}\frac{1}{m_i}.$$

c) The hint

$$\frac{1}{n}\sum_{i=1}^{n} m_i \geq \frac{n}{\sum_{i=1}^{n}\frac{1}{m_i}}$$

implies that

$$\frac{n}{\sum_{i=1}^{n} m_i} \leq \frac{\sum_{i=1}^{n}\frac{1}{m_i}}{n} \quad \text{and} \quad \frac{1}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{m_i}} \leq \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}\frac{1}{\mathrm{m_i}}}{\mathrm{n}^2}.$$

Hence $\text{MSE}(T_1) \leq \text{MSE}(T_2)$, and T_1 is better.

6.2 Exponential Families, UMVUEs and the FCRLB.

Definition 6.2. Let the sample $\boldsymbol{Y} = (Y_1, ..., Y_n)$ where $\boldsymbol{Y}$ has a pdf or pmf $f(\boldsymbol{y}|\theta)$ for $\theta \in \Theta$. Assume all relevant expectations exist. Let $\tau(\theta)$ be a real valued function of θ, and let $U \equiv U(Y_1, ..., Y_n)$ be an estimator of $\tau(\theta)$. Then U is the *uniformly minimum variance unbiased estimator* (**UMVUE**) of $\tau(\theta)$ if U is an unbiased estimator of $\tau(\theta)$ and if $\text{Var}_\theta(U) \le \text{Var}_\theta(\text{W})$ for all $\theta \in \Theta$ where W is any other unbiased estimator of $\tau(\theta)$.

The following theorem is the most useful method for finding UMVUEs since if $Y_1, ..., Y_n$ are iid from a 1P–REF $f(y|\theta) = h(y)c(\theta)\exp[w(\theta)t(y)]$ where $\eta = w(\theta) \in \Omega = (a, b)$ and $a < b$ are not necessarily finite, then $T(\boldsymbol{Y}) = \sum_{i=1}^n t(Y_i)$ is a complete sufficient statistic. It will turn out that $E_\theta[W(\boldsymbol{Y})|T(\boldsymbol{Y})] \equiv E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ does not depend on θ. Hence $U = E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ is a statistic.

Theorem 6.1, Lehmann-Scheffé UMVUE (LSU) Theorem: If $T(\boldsymbol{Y})$ is a complete sufficient statistic for θ, then

$$U = g(T(\boldsymbol{Y})) \tag{6.4}$$

is the UMVUE of its expectation $E_\theta(U) = E_\theta[g(T(\boldsymbol{Y}))]$. In particular, if $W(\boldsymbol{Y})$ is *any unbiased estimator* of $\tau(\theta)$, then

$$U \equiv g(T(\boldsymbol{Y})) = E[W(\boldsymbol{Y})|T(\boldsymbol{Y})] \tag{6.5}$$

is the UMVUE of $\tau(\theta)$. If $V_\theta(U) < \infty$ for all $\theta \in \Theta$, then U is the unique UMVUE of $\tau(\theta) = E_\theta[g(T(\boldsymbol{Y}))]$.

The process (6.5) is called Rao-Blackwellization because of the following theorem. The theorem is also called the Rao-Blackwell-Lehmann-Scheffé theorem.

Theorem 6.2, Rao-Blackwell Theorem. Let $W \equiv W(\boldsymbol{Y})$ be an unbiased estimator of $\tau(\theta)$ and let $T \equiv T(\boldsymbol{Y})$ be a sufficient statistic for θ. Then $\phi(T) = E[W|T]$ is an unbiased estimator of $\tau(\theta)$ and $\text{VAR}_\theta[\phi(T)] \le \text{VAR}_\theta(W)$ for all $\theta \in \Theta$.

Proof. Notice that $\phi(T)$ does not depend on θ by the definition of a sufficient statistic, and that $\phi(T)$ is an unbiased estimator for $\tau(\theta)$ since

$\tau(\theta) = E_\theta(W) = E_\theta(E(W|T)) = E_\theta(\phi(T))$ by iterated expectations (Theorem 2.10). By Steiner's formula (Theorem 2.11), $\text{VAR}_\theta(W) =$

$$E_\theta[\text{VAR}(W|T)] + \text{VAR}_\theta[E(W|T)] \geq \text{VAR}_\theta[E(W|T)] = \text{VAR}_\theta[\phi(T)]. \quad \text{QED}$$

Tips for finding the UMVUE:

i) From the LSU Theorem, if $T(\boldsymbol{Y})$ is complete sufficient statistic and $g(T(\boldsymbol{Y}))$ is a real valued function, then $U = g(T(\boldsymbol{Y}))$ is **the UMVUE of its expectation** $E_\theta[g(T(\boldsymbol{Y}))]$.

ii) Given a complete sufficient statistic $T(\boldsymbol{Y})$ (eg $T(\boldsymbol{Y}) = \sum_{i=1}^n t(Y_i)$ if the data are iid from a 1P–REF), the first method for finding the UMVUE of $\tau(\theta)$ is to guess g and show that $E_\theta[g(T(\boldsymbol{Y}))] = \tau(\theta)$ for all θ.

iii) If $T(\boldsymbol{Y})$ is complete, the second method is to find **any unbiased estimator** $W(\boldsymbol{Y})$ of $\tau(\theta)$. Then $U(\boldsymbol{Y}) = E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ is the UMVUE of $\tau(\boldsymbol{\theta})$.

This problem is often very hard because guessing g or finding an unbiased estimator W and computing $E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ tend to be difficult. Write down the two methods for finding the UMVUE and simplify $E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ as far as you can for partial credit. If you are asked to find the UMVUE of $\tau(\theta)$, see if an unbiased estimator $W(\boldsymbol{Y})$ is given in the problem. Also check whether you are asked to compute $E[W(\boldsymbol{Y})|T(\boldsymbol{Y}) = t]$ anywhere.

iv) The following facts can be useful for computing the conditional expectation via Rao-Blackwellization (see problems 6.7, 6.10 and 6.12). Suppose $Y_1, ..., Y_n$ are iid with finite expectation.
a) Then $E[Y_1| \sum_{i=1}^n Y_i = x] = x/n$.
b) If the Y_i are iid Poisson(θ), then $(Y_1| \sum_{i=1}^n Y_i = x) \sim bin(x, 1/n)$.
c) If the Y_i are iid Bernoulli Ber(ρ), then $(Y_1| \sum_{i=1}^n Y_i = x) \sim Ber(x/n)$.
d) If the Y_i are iid N(μ, σ^2), then $(Y_1| \sum_{i=1}^n Y_i = x) \sim N[x/n, \sigma^2(1 - 1/n)]$.

Often students will be asked to compute a lower bound on the variance of unbiased estimators of $\eta = \tau(\theta)$ when θ is a scalar.

Definition 6.3. Let $\boldsymbol{Y} = (Y_1, ..., Y_n)$ have a pdf or pmf $f(\boldsymbol{y}|\theta)$. Then the **information number** or **Fisher Information** is

$$I_{\boldsymbol{Y}}(\theta) \equiv I_n(\theta) = E_\theta\left(\left[\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))\right]^2\right). \tag{6.6}$$

Let $\eta = \tau(\theta)$ where $\tau'(\theta) \neq 0$. Then

$$I_n(\eta) \equiv I_n(\tau(\theta)) = \frac{I_n(\theta)}{[\tau'(\theta)]^2.} \tag{6.7}$$

Theorem 6.3. a) Equations (6.6) and (6.7) agree if $\tau'(\theta)$ is continuous, $\tau'(\theta) \neq 0$, and $\tau(\theta)$ is one to one and onto so that an inverse function exists such that $\theta = \tau^{-1}(\eta)$

b) If the $Y_1 \equiv Y$ is from a 1P–REF, then the Fisher information in a sample of size one is

$$I_1(\theta) = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log(f(Y|\theta))\right]. \tag{6.8}$$

c) If the $Y_1, ..., Y_n$ are iid from a 1P–REF, then

$$I_n(\theta) = nI_1(\theta). \tag{6.9}$$

Hence if $\tau'(\theta)$ exists and is continuous and if $\tau'(\theta) \neq 0$, then

$$I_n(\tau(\theta)) = \frac{nI_1(\theta)}{[\tau'(\theta)]^2}. \tag{6.10}$$

Proof. a) See Lehmann (1999, p. 467–468).

b) The proof will be for a pdf. For a pmf replace the integrals by sums. By Remark 3.2, the integral and differentiation operators of all orders can be interchanged. Note that

$$0 = E\left[\frac{\partial}{\partial\theta}\log(f(Y|\theta))\right] \tag{6.11}$$

since

$$\frac{\partial}{\partial\theta}1 = 0 = \frac{\partial}{\partial\theta}\int f(y|\theta)dy = \int \frac{\partial}{\partial\theta}f(y|\theta)dy = \int \frac{\frac{\partial}{\partial\theta}f(y|\theta)}{f(y|\theta)}f(y|\theta)dy$$

or

$$0 = \frac{\partial}{\partial\theta}\int f(y|\theta)dy = \int\left[\frac{\partial}{\partial\theta}\log(f(y|\theta))\right]f(y|\theta)dy$$

which is (6.11). Taking 2nd derivatives of the above expression gives

$$0 = \frac{\partial^2}{\partial\theta^2} \int f(y|\theta) dy = \frac{\partial}{\partial\theta} \int \left[\frac{\partial}{\partial\theta} \log(f(y|\theta))\right] f(y|\theta) dy =$$

$$\int \frac{\partial}{\partial\theta} \left(\left[\frac{\partial}{\partial\theta} \log(f(y|\theta))\right] f(y|\theta) \right) dy =$$

$$\int \left[\frac{\partial^2}{\partial\theta^2} \log(f(y|\theta))\right] f(y|\theta) dy + \int \left[\frac{\partial}{\partial\theta} \log(f(y|\theta))\right] \left[\frac{\partial}{\partial\theta} f(y|\theta)\right] \frac{f(y|\theta)}{f(y|\theta)} dy$$

$$= \int \left[\frac{\partial^2}{\partial\theta^2} \log(f(y|\theta))\right] f(y|\theta) dy + \int \left[\frac{\partial}{\partial\theta} \log(f(y|\theta))\right]^2 f(y|\theta) dy$$

or

$$I_1(\theta) = E_\theta[(\frac{\partial}{\partial\theta} \log f(Y|\theta))^2] = -E_\theta \left[\frac{\partial^2}{\partial\theta^2} \log(f(Y|\theta))\right].$$

c) By independence,

$$I_n(\theta) = E_\theta \left[\left(\frac{\partial}{\partial\theta} \log(\prod_{i=1}^n f(Y_i|\theta)) \right)^2 \right] = E_\theta \left[\left(\frac{\partial}{\partial\theta} \sum_{i=1}^n \log(f(Y_i|\theta)) \right)^2 \right] =$$

$$E_\theta \left[\left(\frac{\partial}{\partial\theta} \sum_{i=1}^n \log(f(Y_i|\theta)) \right) \left(\frac{\partial}{\partial\theta} \sum_{j=1}^n \log(f(Y_j|\theta)) \right) \right] =$$

$$E_\theta \left[\left(\sum_{i=1}^n \frac{\partial}{\partial\theta} \log(f(Y_i|\theta)) \right) \left(\sum_{j=1}^n \frac{\partial}{\partial\theta} \log(f(Y_j|\theta)) \right) \right] =$$

$$\sum_{i=1}^n E_\theta \left[\left(\frac{\partial}{\partial\theta} \log(f(Y_i|\theta)) \right)^2 \right] +$$

$$\sum_{i} \sum_{\neq j} E_\theta \left[\left(\frac{\partial}{\partial\theta} \log(f(Y_i|\theta)) \right) \left(\frac{\partial}{\partial\theta} \log(f(Y_j|\theta)) \right) \right].$$

Hence

$$I_n(\theta) = nI_1(\theta) + \sum_{i} \sum_{\neq j} E_\theta \left[\left(\frac{\partial}{\partial\theta} \log(f(Y_i|\theta)) \right) \right] E_\theta \left[\left(\frac{\partial}{\partial\theta} \log(f(Y_j|\theta)) \right) \right]$$

by independence. Hence

$$I_n(\theta) = nI_1(\theta) + n(n-1)\left[E_\theta\left(\frac{\partial}{\partial\theta}\log(f(Y_j|\theta))\right)\right]^2$$

since the Y_i are iid. Thus $I_n(\theta) = nI_1(\theta)$ by Equation (6.11) which holds since the Y_i are iid from a 1P–REF. QED

Definition 6.4. Let $\boldsymbol{Y} = (Y_1, ..., Y_n)$ be the data, and consider $\tau(\theta)$ where $\tau'(\theta) \neq 0$. The quantity

$$\mathrm{FCRLB_n}(\tau(\theta)) = \frac{[\tau'(\theta)]^2}{\mathrm{I_n}(\theta)}$$

is called the **Fréchet Cramér Rao lower bound** (FCRLB) for the variance of unbiased estimators of $\tau(\theta)$. In particular, if $\tau(\theta) = \theta$, then $\mathrm{FCRLB}_n(\theta) = \frac{1}{I_n(\theta)}$. The FCRLB is often called the Cramér Rao lower bound (CRLB).

Theorem 6.4, Fréchet Cramér Rao Lower Bound or Information Inequality. Let $Y_1, ..., Y_n$ be iid from a 1P–REF with pdf or pmf $f(y|\theta)$. Let $W(Y_1, ..., Y_n) = W(\boldsymbol{Y})$ be any unbiased estimator of $\tau(\theta) \equiv E_\theta W(\boldsymbol{Y})$. Then

$$\mathrm{VAR}_\theta(W(\boldsymbol{Y})) \geq FCRLB_n(\tau(\theta)) = \frac{[\tau'(\theta)]^2}{I_n(\theta)} = \frac{[\tau'(\theta)]^2}{nI_1(\theta)}.$$

Proof. By Definition 6.4 and Theorem 6.3c,

$$FCRLB_n(\tau(\theta)) = \frac{[\tau'(\theta)]^2}{I_n(\theta)} = \frac{[\tau'(\theta)]^2}{nI_1(\theta)}.$$

Since the Y_i are iid from a 1P–REF, by Remark 3.2 the derivative and integral or sum operators can be interchanged when finding the derivative of $E_\theta h(\boldsymbol{Y})$ if $E_\theta|h(\boldsymbol{Y})| < \infty$. The following argument will be for pdfs. For pmfs, replace the integrals by appropriate sums. Following Casella and Berger (2002, p. 335-8), the Cauchy Schwarz Inequality is

$$[\mathrm{Cov(X,Y)}]^2 \leq \mathrm{V(X)V(Y)}, \quad \text{or} \quad \mathrm{V(X)} \geq \frac{[\mathrm{Cov(X,Y)}]^2}{\mathrm{V(Y)}}.$$

Hence

$$V_\theta(W(\boldsymbol{Y})) \geq \frac{(\mathrm{Cov}_\theta[\mathrm{W}(\boldsymbol{Y}), \frac{\partial}{\partial\theta}\log(\mathrm{f}(\boldsymbol{Y}|\theta))])^2}{V_\theta[\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))]}. \tag{6.12}$$

Now

$$E_\theta[\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))] = E_\theta\left[\frac{\frac{\partial}{\partial\theta}f(\boldsymbol{Y}|\theta)}{f(\boldsymbol{Y}|\theta)}\right]$$

since the derivative of $\log(h(t))$ is $h'(t)/h(t)$. By the definition of expectation,

$$E_\theta[\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))] = \int\cdots\int_{\mathcal{Y}}\frac{\frac{\partial}{\partial\theta}f(\boldsymbol{y}|\theta)}{f(\boldsymbol{y}|\theta)}f(\boldsymbol{y}|\theta)d\boldsymbol{y}$$

$$= \int\cdots\int_{\mathcal{Y}}\frac{\partial}{\partial\theta}f(\boldsymbol{y}|\theta)d\boldsymbol{y} = \frac{d}{d\theta}\int\cdots\int_{\mathcal{Y}}f(\boldsymbol{y}|\theta)d\boldsymbol{y} = \frac{d}{d\theta}1 = 0.$$

Notice that $f(\boldsymbol{y}|\theta) > 0$ on the support $\mathcal{Y}$, that the $f(\boldsymbol{y}|\theta)$ cancelled in the 2nd term, that the derivative was moved outside of the integral by Remark 3.2, and that the integral of $f(\boldsymbol{y}|\theta)$ on the support $\mathcal{Y}$ is equal to 1.

This result implies that

$$\text{Cov}_\theta[\text{W}(\boldsymbol{Y}), \frac{\partial}{\partial\theta}\log(\text{f}(\boldsymbol{Y}|\theta))] = \text{E}_\theta[\text{W}(\boldsymbol{Y}) \quad \frac{\partial}{\partial\theta}\log(\text{f}(\boldsymbol{Y}|\theta))]$$

$$= E_\theta\left[\frac{W(\boldsymbol{Y}) \quad (\frac{\partial}{\partial\theta}f(\boldsymbol{Y}|\theta))}{f(\boldsymbol{Y}|\theta)}\right]$$

since the derivative of $\log(h(t))$ is $h'(t)/h(t)$. By the definition of expectation, the right hand side is equal to

$$\int\cdots\int_{\mathcal{Y}}\frac{W(\boldsymbol{y})\frac{\partial}{\partial\theta}f(\boldsymbol{y}|\theta)}{f(\boldsymbol{y}|\theta)}f(\boldsymbol{y}|\theta)d\boldsymbol{y} = \frac{d}{d\theta}\int\cdots\int_{\mathcal{Y}}W(\boldsymbol{y})f(\boldsymbol{y}|\theta)d\boldsymbol{y}$$

$$= \frac{d}{d\theta}E_\theta W(\boldsymbol{Y}) = \tau'(\theta) = \text{Cov}_\theta[\text{W}(\boldsymbol{Y}), \frac{\partial}{\partial\theta}\log(\text{f}(\boldsymbol{Y}|\theta))]. \quad (6.13)$$

Since

$$E_\theta[\frac{\partial}{\partial\theta}\log \ f(\boldsymbol{Y}|\theta)] = 0,$$

$$V_\theta[\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))] = E_\theta([\frac{\partial}{\partial\theta}\log(f(\boldsymbol{Y}|\theta))]^2) = I_n(\theta) \quad (6.14)$$

by Definition 6.3. Plugging (6.13) and (6.14) into (6.12) gives the result. QED

Theorem 6.4 is not very useful in applications. If the data are iid from a 1P–REF then $FCRLB_n(\tau(\theta)) = [\tau'(\theta)]^2/[nI_1(\theta)]$ by Theorem 6.4. Notice

that $W(\boldsymbol{Y})$ is an unbiased estimator of $\tau(\theta)$ since $E_\theta W(\boldsymbol{Y}) = \tau(\theta)$. Hence if the data are iid from a 1P–REF and if $\text{VAR}_\theta(W(\boldsymbol{Y})) = FCRLB_n(\tau(\theta))$ for all $\theta \in \Theta$ then $W(\boldsymbol{Y})$ is the UMVUE of $\tau(\theta)$; however, this technique for finding a UMVUE rarely works since typically equality holds only if
1) the data come from a 1P–REF with complete sufficient statistic T, and
2) $W = a + bT$ is a linear function of T.
The FCRLB inequality will typically be strict for nonlinear functions of T if the data is iid from a 1P–REF. If T is complete, $g(T)$ is the UMVUE of its expectation, and determining that T is the complete sufficient statistic from a 1P–REF is simpler than computing $\text{VAR}_\theta(W)$ and $\text{FCRLB}_n(\tau(\theta))$. If the family is not an exponential family, the FCRLB may **not be a lower bound** on the variance of unbiased estimators of $\tau(\theta)$.

Example 6.4. Let $Y_1, ..., Y_n$ be iid random variables with pdf

$$f(y) = \frac{2}{\sqrt{2\pi}\lambda}\frac{1}{y} I_{[0,1]}(y) \exp\left[\frac{-(\log(y))^2}{2\lambda^2}\right]$$

where $\lambda > 0$. Then $[\log(Y_i)]^2 \sim G(1/2, 2\lambda^2) \sim \lambda^2 \chi^2_1$.

a) Find the uniformly minimum variance estimator (UMVUE) of λ^2.

b) Find the information number $I_1(\lambda)$.

c) Find the Fréchet Cramér Rao lower bound (FCRLB) for estimating $\tau(\lambda) = \lambda^2$.

Solution. a) This is a one parameter exponential family with complete sufficient statistic $T_n = \sum_{i=1}^n [\log(Y_i)]^2$. Now $E(T_n) = nE([\log(Y_i)]^2) = n\lambda^2$. Hence $E(T_n/n) = \lambda^2$ and T_n/n is the UMVUE of λ^2 by the LSU Theorem.

b) Now

$$\log(f(y|\lambda)) = \log(2/\sqrt{2\pi}) - \log(\lambda) - \log(y) - \frac{[\log(y)]^2}{2\lambda^2}.$$

Hence

$$\frac{d}{d\lambda}\log(f(y|\lambda)) = \frac{-1}{\lambda} + \frac{[\log(y)]^2}{\lambda^3},$$

and

$$\frac{d^2}{d\lambda^2}\log(f(y|\lambda)) = \frac{1}{\lambda^2} - \frac{3[\log(y)]^2}{\lambda^4}.$$

Thus

$$I_1(\lambda) = -E\left[\frac{1}{\lambda^2} - \frac{3[\log(Y)]^2}{\lambda^4}\right] = \frac{-1}{\lambda^2} + \frac{3\lambda^2}{\lambda^4} = \frac{2}{\lambda^2}.$$

c)

$$FCRLB_n(\tau(\lambda)) = \frac{[\tau'(\lambda)]^2}{nI_1(\lambda)}.$$

Now $\tau(\lambda) = \lambda^2$ and $\tau'(\lambda) = 2\lambda$. So

$$FCRLB_n(\tau(\lambda)) = \frac{4\lambda^2}{n2/\lambda^2} = \frac{2\lambda^4}{n}.$$

Example 6.5. Suppose that $X_1, ..., X_n$ are iid Bernoulli(p) where $n \geq 2$ and $0 < p < 1$ is the unknown parameter.

a) Derive the UMVUE of $\tau(p)$, where $\tau(p) = e^2(p(1-p))$.

b) Find the FCRLB for estimating $\tau(p) = e^2(p(1-p))$.

Solution: a) Consider the statistic $W = X_1(1-X_2)$ which is an unbiased estimator of $\tau(p) = p(1-p)$. The statistic $T = \sum_{i=1}^n X_i$ is both complete and sufficient. The possible values of W are 0 or 1. Then $U = \phi(T)$ where

$$\begin{aligned}
\phi(t) &= E[X_1(1-X_2)|T=t] \\
&= 0P[X_1(1-X_2)=0|T=t] + 1P[X_1(1-X_2)=1|T=t] \\
&= P[X_1(1-X_2)=1|T=t] \\
&= \frac{P[X_1=1, X_2=0 \text{ and } \sum_{\rm i=1}^{\rm n} \rm X_i = t]}{P[\sum_{i=1}^n X_i = t]} \\
&= \frac{P[X_1=1]P[X_2=0]P[\sum_{i=3}^n X_i = t-1]}{P[\sum_{i=1}^n X_i = t]}.
\end{aligned}$$

Now $\sum_{i=3}^n X_i$ is $Bin(n-2, p)$ and $\sum_{i=1}^n X_i$ is $Bin(n, p)$. Thus

$$\phi(t) = \frac{p(1-p)[\binom{n-2}{t-1}p^{t-1}(1-p)^{n-t-1}]}{\binom{n}{t}p^t(1-p)^{n-t}}$$

$$= \frac{\binom{n-2}{t-1}}{\binom{n}{t}} = \frac{(n-2)!}{(t-1)!(n-2-t+1)!} \; \frac{t(t-1)!(n-t)(n-t-1)!}{n(n-1)(n-2)!} = \frac{t(n-t)}{n(n-1)}$$

$$= \frac{\frac{t}{n}(n - n\frac{t}{n})}{n-1} = \frac{\frac{t}{n}n(1-\frac{t}{n})}{n-1} = \frac{n}{n-1}\overline{x}(1-\overline{x}).$$

Thus $\frac{n}{n-1}\overline{X}(1-\overline{X})$ is the UMVUE of $p(1-p)$ and $U = e^2\frac{n}{n-1}\overline{X}(1-\overline{X})$ is the UMVUE of $\tau(p) = e^2p(1-p)$.

Alternatively, $\overline{X}$ is a complete sufficient statistic, so try an estimator of the form $U = a(\overline{X})^2 + b\overline{X} + c$. Then U is the UMVUE if $E_p(U) = e^2 p(1-p) = e^2(p - p^2)$. Now $E(\overline{X}) = E(X_1) = p$ and $V(\overline{X}) = V(X_1)/n = p(1-p)/n$ since $\sum X_i \sim Bin(n,p)$. So $E[(\overline{X})^2] = V(\overline{X}) + [E(\overline{X})]^2 = p(1-p)/n + p^2$. So $E_p(U) = a[p(1-p)/n] + ap^2 + bp + c$

$$= \frac{ap}{n} - \frac{ap^2}{n} + ap^2 + bp + c = (\frac{a}{n} + b)p + (a - \frac{a}{n})p^2 + c.$$

So $c = 0$ and $a - \frac{a}{n} = a\frac{n-1}{n} = -e^2$ or

$$a = \frac{-n}{n-1}e^2.$$

Hence $\frac{a}{n} + b = e^2$ or

$$b = e^2 - \frac{a}{n} = e^2 + \frac{n}{n(n-1)}e^2 = \frac{n}{n-1}e^2.$$

So

$$U = \frac{-n}{n-1}e^2(\overline{X})^2 + \frac{n}{n-1}e^2\overline{X} = \frac{n}{n-1}e^2\overline{X}(1-\overline{X}).$$

b) The FCRLB for $\tau(p)$ is $[\tau'(p)]^2/nI_1(p)$. Now $f(x) = p^x(1-p)^{1-x}$, so $\log f(x) = x\log(p) + (1-x)\log(1-p)$. Hence

$$\frac{\partial \log f}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p}$$

and

$$\frac{\partial^2 \log f}{\partial p^2} = \frac{-x}{p^2} - \frac{1-x}{(1-p)^2}.$$

So

$$I_1(p) = -E(\frac{\partial^2 \log f}{\partial p^2}) = -(\frac{-p}{p^2} - \frac{1-p}{(1-p)^2}) = \frac{1}{p(1-p)}.$$

So

$$FCRLB_n = \frac{[e^2(1-2p)]^2}{\frac{n}{p(1-p)}} = \frac{e^4(1-2p)^2 p(1-p)}{n}.$$

Example 6.6. Let $X_1, ..., X_n$ be iid random variables with pdf

$$f(x) = \frac{1}{\lambda}\phi x^{\phi-1}\frac{1}{1+x^\phi}\exp\left[-\frac{1}{\lambda}\log(1+x^\phi)\right]$$

where x, ϕ, and λ are all positive. If ϕ is known, find the uniformly minimum unbiased estimator of λ using the fact that $\log(1+X_i^{\phi}) \sim$ Gamma $(\nu = 1, \lambda)$.

Solution: This is a regular one parameter exponential family with complete sufficient statistic $T_n = \sum_{i=1}^n \log(1+X_i^{\phi}) \sim G(n, \lambda)$. Hence $E(T_n) = n\lambda$ and T_n/n is the UMVUE of λ.

6.3 Summary

1) The **bias** of the estimator T for $\tau(\boldsymbol{\theta})$ is

$$B(T) \equiv B_{\tau(\boldsymbol{\theta})}(T) \equiv \text{Bias}_{\tau(\boldsymbol{\theta})}(\text{T}) = \text{E}_{\boldsymbol{\theta}}\text{T} - \tau(\boldsymbol{\theta})$$

and the MSE is

$$\text{MSE}_{\tau(\boldsymbol{\theta})}(\text{T}) = \text{E}_{\boldsymbol{\theta}}[(\text{T} - \tau(\boldsymbol{\theta}))^2] = \text{V}_{\boldsymbol{\theta}}(\text{T}) + [\text{Bias}_{\tau(\boldsymbol{\theta})}(\text{T})]^2.$$

2) T is an *unbiased estimator* of $\tau(\boldsymbol{\theta})$ if $E_{\boldsymbol{\theta}}[T] = \tau(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Theta$.

3) Let $U \equiv U(Y_1, ..., Y_n)$ be an estimator of $\tau(\theta)$. Then U is the **UMVUE** of $\tau(\theta)$ if U is an unbiased estimator of $\tau(\theta)$ and if $\text{VAR}_\theta(U) \leq \text{VAR}_\theta(W)$ for all $\theta \in \Theta$ where W is any other unbiased estimator of $\tau(\theta)$.

4) If $Y_1, ..., Y_n$ are iid from a 1P–REF $f(y|\theta) = h(y)c(\theta)\exp[w(\theta)t(y)]$ where $\eta = w(\theta) \in \Omega = (a, b)$, and if $T \equiv T(\boldsymbol{Y}) = \sum_{i=1}^n t(Y_i)$, then by the LSU Theorem, $g(T)$ is the UMVUE of its expectation $\tau(\theta) = E_\theta(g(T))$.

5) Given a complete sufficient statistic $T(\boldsymbol{Y})$ and any unbiased estimator $W(\boldsymbol{Y})$ of $\tau(\theta)$, then $U(\boldsymbol{Y}) = E[W(\boldsymbol{Y})|T(\boldsymbol{Y})]$ is the UMVUE of $\tau(\theta)$.

7) $I_n(\theta) = E_\theta[(\frac{\partial}{\partial\theta} \log f(\boldsymbol{Y}|\theta))^2]$.

8) $\text{FCRLB}_n(\tau(\theta)) = \dfrac{[\tau'(\theta)]^2}{I_n(\theta)}$.

9) If $Y_1, ..., Y_n$ are iid from a 1P–REF $f(y|\theta) = h(y)c(\theta)\exp[w(\theta)t(y)]$, then a)

$$I_1(\theta) = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\log(f(Y|\theta))\right].$$

b)

$$I_n(\tau(\theta)) = \frac{nI_1(\theta)}{[\tau'(\theta)]^2}.$$

c)

$$FCRLB_n(\tau(\theta)) = \frac{[\tau'(\theta)]^2}{nI_1(\theta)}.$$

d) Information inequality: Let $Y_1, ..., Y_n$ be iid from a 1P–REF and let $W(\boldsymbol{Y})$ be any unbiased estimator of $\tau(\theta) \equiv E_\theta[W(\boldsymbol{Y})]$. Then

$$\text{VAR}_\theta(W(\boldsymbol{Y})) \geq FCRLB_n(\tau(\theta)) = \frac{[\tau'(\theta)]^2}{nI_1(\theta)}.$$

e) Rule of thumb for a 1P–REF: Let $T(\boldsymbol{Y}) = \sum_{i=1}^n t(Y_i)$ and $\tau(\theta) = E_\theta(g(T(\boldsymbol{Y}))$. Then $g(T(\boldsymbol{Y}))$ is the UMVUE of $\tau(\theta)$ by LSU, but the information inequality is strict for nonlinear functions $g(T(\boldsymbol{Y}))$. Expect the equality

$$\text{VAR}_\theta(g(T(\boldsymbol{Y})) = \frac{[\tau'(\theta)]^2}{nI_1(\theta)}$$

only if g is a linear function, ie, $g(T) = a + bT$ for some fixed constants a and b.

10) If the family is not an exponential family, the FCRLB may **not be a lower bound** on the variance of unbiased estimators of $\tau(\theta)$.

6.4 Complements

For a more precise statement of when the FCRLB is achieved and for some counterexamples, see Wijsman (1973) and Joshi (1976). Although the FCRLB is not very useful for finding UMVUEs, similar ideas are useful for finding the asymptotic variances of UMVUEs and MLEs. See Chapter 8 and Portnoy (1977).

Karakostas (1985) has useful references for UMVUEs. Also see Guenther (1978).

6.5 Problems

PROBLEMS WITH AN ASTERISK * ARE ESPECIALLY USEFUL.

Refer to Chapter 10 for the pdf or pmf of the distributions in the problems below.

6.1*. Let W be an estimator of $\tau(\theta)$. Show that

$$MSE_{\tau(\theta)}(W) = Var_\theta(W) + [Bias_{\tau(\theta)}(W)]^2.$$

6.2. (Aug. 2002 QUAL): Let $X_1, ..., X_n$ be independent identically distributed random variable from a $N(\mu, \sigma^2)$ distribution. Hence $E(X_1) = \mu$ and $VAR(X_1) = \sigma^2$. Consider estimators of σ^2 of the form

$$S^2(k) = \frac{1}{k}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

where $k > 0$ is a constant to be chosen. Determine the value of k which gives the smallest mean square error. (Hint: Find the MSE as a function of k, then take derivatives with respect to k. Also, use Theorem 4.1c and Remark 5.1 VII.)

6.3. Let $X_1, ..., X_n$ be iid N$(\mu, 1)$ random variables. Find $\tau(\mu)$ such that $T(X_1, ..., X_n) = (\sum_{i=1}^n X_i)^2$ is the UMVUE of $\tau(\mu)$.

6.4. Let $X \sim N(\mu, \sigma^2)$ where σ^2 is known. Find the Fisher information $I_1(\mu)$.

6.5. Let $X \sim N(\mu, \sigma^2)$ where μ is known. Find the Fisher information $I_1(\sigma^2)$.

6.6. Let $X_1, ..., X_n$ be iid $N(\mu, \sigma^2)$ random variables where μ is **known** and $\sigma^2 > 0$. Then $W = \sum_{i=1}^n (X_i - \mu)^2$ is a complete sufficient statistic and $W \sim \sigma^2 \chi^2_n$. From Chapter 10,

$$EY^k = \frac{2^k \Gamma(k + n/2)}{\Gamma(n/2)}$$

if $Y \sim \chi^2_n$. Hence

$$T_k(X_1, ..., X_n) \equiv \frac{\Gamma(n/2) W^k}{2^k \Gamma(k + n/2)}$$

is the UMVUE of $\tau_k(\sigma^2) = \sigma^{2k}$ for $k > 0$. Note that $\tau_k(\theta) = (\theta)^k$ and $\theta = \sigma^2$.

a) Show that

$$Var_\theta T_k(X_1, ..., X_n) = \sigma^{4k}\left[\frac{\Gamma(n/2)\Gamma(2k + n/2)}{\Gamma(k + n/2)\Gamma(k + n/2)} - 1\right] \equiv c_k \sigma^{4k}.$$

b) Let $k = 2$ and show that $\mathrm{Var}_\theta[T_2] - FCRLB(\tau_2(\theta)) > 0$ where FCRLB$(\tau_2(\theta))$ is for estimating $\tau_2(\sigma^2) = \sigma^4$ and $\theta = \sigma^2$.

6.7. (Jan. 2001 Qual): Let $X_1, ..., X_n$ be independent, identically distributed $N(\mu, 1)$ random variables where μ is unknown and $n \geq 2$. Let t be a fixed real number. Then the expectation

$$E_\mu(I_{(-\infty,t]}(X_1)) = P_\mu(X_1 \leq t) = \Phi(t - \mu)$$

for all μ where $\Phi(x)$ is the cumulative distribution function of a $N(0, 1)$ random variable.

a) Show that the sample mean $\overline{X}$ is a sufficient statistic for μ.

b) Explain why (or show that) $\overline{X}$ is a complete sufficient statistic for μ.

c) Using the fact that the conditional distribution of X_1 given $\overline{X} = \overline{x}$ is the $N(\overline{x}, 1 - 1/n)$ distribution where the second parameter $1 - 1/n$ is the variance of conditional distribution, find

$$E_\mu(I_{(-\infty,t]}(X_1)|\overline{X} = \overline{x}) = E_\mu[I_{(-\infty,t]}(W)]$$

where $W \sim N(\overline{x}, 1 - 1/n)$. (Hint: your answer should be $\Phi(g(\overline{x}))$ for some function g.)

d) What is the uniformly minimum variance unbiased estimator for $\Phi(t - \mu)$?

Problems from old quizzes and exams.

6.8. Suppose that X is Poisson with pmf

$$f(x|\lambda) = P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

where $x = 0, 1, ...$ and $\lambda > 0$. Find the Fisher information $I_1(\lambda)$.

6.9. Let $X_1, ..., X_n$ be iid Exponential(β) random variables and $Y_1, ..., Y_m$ iid Exponential($\beta/2$) random variables. Assume that the Y_i's and X_j's are independent.

a) Find the joint pdf $f(x_1, ..., x_n, y_1, ..., y_m)$ and show that this pdf is a regular exponential family with complete sufficient statistic $T = \sum_{i=1}^n X_i + 2\sum_{i=1}^m Y_i$.

b) Find the function $\tau(\beta)$ such that T is the UMVUE of $\tau(\beta)$. (Hint: find $E_\beta[T]$. The theorems of this chapter apply since $X_1, ..., X_n, 2Y_1, ..., 2Y_m$ are iid.)

6.10. Let $X_1, ..., X_n$ be independent, identically distributed $N(\mu, 1)$ random variables where μ is unknown.

a) Find $E_\mu[X_1^2]$.

b) Using the fact that the conditional distribution of X_1 given $\overline{X} = \overline{x}$ is the $N(\overline{x}, 1 - 1/n)$ distribution where the second parameter $1 - 1/n$ is the variance of conditional distribution, find

$$E_\mu(X_1^2 | \overline{X} = \overline{x}).$$

[Hint: this expected value is equal to $E(W^2)$ where $W \sim N(\overline{x}, 1 - 1/n)$.]

c) What is the MLE for $\mu^2 + 1$? (Hint: you may use the fact that the MLE for μ is $\overline{X}$.)

d) What is the uniformly minimum variance unbiased estimator for μ^2+1? Explain.

6.11. Let $X_1, ..., X_n$ be a random sample from a Poisson(λ) population.

a) Find the Fréchet Cramér Rao lower bound $FCRLB_n(\lambda^2)$ for the variance of an unbiased estimator of $\tau(\lambda) = \lambda^2$.

b) The UMVUE for λ^2 is $T(X_1, ..., X_n) = (\overline{X})^2 - \overline{X}/n$. Will $Var_\lambda[T] = FCRLB_n(\lambda^2)$ or will $Var_\lambda[T] > FCRLB_n(\lambda^2)$? Explain. (Hint: use the rule of thumb 9e from Section 6.3.)

6.12. Let $X_1, ..., X_n$ be independent, identically distributed Poisson(λ) random variables where $\lambda > 0$ is unknown.

a) Find $E_\lambda[X_1^2]$.

b) Using the fact that the conditional distribution of X_1 given $\sum_{i=1}^n X_i = y$ is the Binomial($y, 1/n$) distribution, find

$$E_\lambda(X_1^2 | \sum_{i=1}^n X_i = y).$$

c) Find $\tau(\lambda)$ such that $E_\lambda(X_1^2 | \sum_{i=1}^n X_i)$ is the uniformly minimum variance unbiased estimator for $\tau(\lambda)$.

6.13. Let $X_1, ..., X_n$ be iid Bernoulli(ρ) random variables.

a) Find the Fisher information $I_1(\rho)$.

b) Find the Fréchet Cramér Rao lower bound for unbiased estimators of $\tau(\rho) = \rho$.

c) The MLE for ρ is $\overline{X}$. Find Var($\overline{X}$).

d) Does the MLE achieve the FCRLB? Is this surprising? Explain.

6.14. (Jan. 2003 Qual): Let $X_1, ..., X_n$ be independent, identically distributed exponential(θ) random variables where $\theta > 0$ is unknown. Consider the class of estimators of θ

$$\{T_n(c) = c \sum_{i=1}^n X_i \;\; | \;\; c > 0\}.$$

Determine the value of c that minimizes the mean square error MSE. Show work and prove that your value of c is indeed the global minimizer.

6.15. Let $X_1, ..., X_n$ be iid from a distribution with pdf

$$f(x|\theta) = \theta x^{\theta-1} I(0 < x < 1), \quad \theta > 0.$$

a) Find the MLE of θ.

b) What is the MLE of $1/\theta^2$? Explain.

c) Find the Fisher information $I_1(\theta)$. You may use the fact that $-\log(X) \sim$ exponential($1/\theta$).

d) Find the Fréchet Cramér Rao lower bound for unbiased estimators of $\tau(\theta) = 1/\theta^2$.

6.16. Let $X_1, ..., X_n$ be iid random variables with $E(X) = \mu$ and $Var(X) = 1$. Suppose that $T = \sum_{i=1}^n X_i$ is a complete sufficient statistic. Find the UMVUE of μ^2.

6.17. Let $X_1, ..., X_n$ be iid exponential(λ) random variables.

a) Find $I_1(\lambda)$.

b) Find the FCRLB for estimating $\tau(\lambda) = \lambda^2$.

c) If $T = \sum_{i=1}^n X_i$, it can be shown that the UMVUE of λ^2 is

$$W = \frac{\Gamma(n)}{\Gamma(2+n)} T^2.$$

Do you think that $Var_\lambda(W)$ is equal to the FCRLB in part b)? Explain briefly.

6.18. Let $X_1, ..., X_n$ be iid $N(\mu, \sigma^2)$ where μ is known and $n > 1$. Suppose interest is in estimating $\theta = \sigma^2$. You should have memorized the fact that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

a) Find the MSE of S^2 for estimating σ^2.

b) Find the MSE of T for estimating σ^2 where

$$T = \frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2.$$

6.19. (Aug. 2000 SIU, 1995 Univ. Minn. Qual): Let $X_1, ..., X_n$ be independent identically distributed random variable from a $N(\mu, \sigma^2)$ distribution. Hence $E(X_1) = \mu$ and $VAR(X_1) = \sigma^2$. Suppose that μ is known and consider estimates of σ^2 of the form

$$S^2(k) = \frac{1}{k}\sum_{i=1}^n (X_i - \mu)^2$$

where k is a constant to be chosen. Note: $E(\chi^2_m) = m$ and $VAR(\chi^2_m) = 2m$. Determine the value of k which gives the smallest mean square error. (Hint: Find the MSE as a function of k, then take derivatives with respect to k.)

6.20. (Aug. 2001 Qual): Let $X_1, ..., X_n$ be independent identically distributed random variables with pdf

$$f(x|\theta) = \frac{2x}{\theta} e^{-x^2/\theta}, \quad x > 0$$

and $f(x|\theta) = 0$ for $x \le 0$.

a) Show that X_1^2 is an unbiased estimator of θ. (Hint: use the substitution $W = X^2$ and find the pdf of W or use u-substitution with $u = x^2/\theta$.)

b) Find the Cramer-Rao lower bound for the variance of an unbiased estimator of θ.

c) Find the uniformly minimum variance unbiased estimator (UMVUE) of θ.

6.21. (Aug. 2001 Qual): See Mukhopadhyay (2000, p. 377). Let $X_1, ..., X_n$ be iid $N(\theta, \theta^2)$ normal random variables with mean θ and variance θ^2. Let

$$T_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

and let

$$T_2 = c_n S = c_n \sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}}$$

where the constant c_n is such that $E_\theta[c_n S] = \theta$. You do not need to find the constant c_n. Consider estimators $W(\alpha)$ of θ of the form.

$$W(\alpha) = \alpha T_1 + (1-\alpha) T_2$$

where $0 \le \alpha \le 1$.

a) Find the variance

$$Var_\theta[W(\alpha)] = Var_\theta(\alpha T_1 + (1-\alpha)T_2).$$

b) Find the mean square error of $W(\alpha)$ in terms of $Var_\theta(T_1), Var_\theta(T_2)$ and α.

c) Assume that

$$Var_\theta(T_2) \approx \frac{\theta^2}{2n}.$$

Determine the value of α that gives the smallest mean square error. (Hint: Find the MSE as a function of α, then take the derivative with respect to α. Set the derivative equal to zero and use the above approximation for $Var_\theta(T_2)$. Show that your value of α is indeed the global minimizer.)

6.22. (Aug. 2003 Qual): Suppose that $X_1, ..., X_n$ are iid normal distribution with mean 0 and variance σ^2. Consider the following estimators: $T_1 = \frac{1}{2}|X_1 - X_2|$ and $T_2 = \sqrt{\frac{1}{n}\sum_{i=1}^n X_i^2}$.

a) Is T_1 unbiased for σ? Evaluate the mean square error (MSE) of T_1.

b) Is T_2 unbiased for σ? If not, find a suitable multiple of T_2 which is unbiased for σ.

6.23. (Aug. 2003 Qual): Let $X_1, ..., X_n$ be independent identically distributed random variables with pdf (probability density function)

$$f(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right)$$

where x and λ are both positive. Find the uniformly minimum variance unbiased estimator (UMVUE) of λ^2.

6.24. (Jan. 2004 Qual): Let $X_1, ..., X_n$ be independent identically distributed random variables with pdf (probability density function)

$$f(x) = \sqrt{\frac{\sigma}{2\pi x^3}} \exp\left(-\frac{\sigma}{2x}\right)$$

where x and σ are both positive. Then $X_i = \dfrac{\sigma}{W_i}$ where $W_i \sim \chi^2_1$. Find the uniformly minimum variance unbiased estimator (UMVUE) of $\dfrac{1}{\sigma}$.

6.25. (Jan. 2004 Qual): Let $X_1, ..., X_n$ be a random sample from the distribution with density

$$f(x) = \begin{cases} \frac{2x}{\theta^2}, & 0 < x < \theta \\ 0 & \text{elsewhere} \end{cases}$$

Let $T = max(X_1, ..., X_n)$. To estimate θ consider estimators of the form CT. Determine the value of C which gives the smallest mean square error.

6.26. (Aug. 2004 Qual): Let $X_1, ..., X_n$ be a random sample from a distribution with pdf

$$f(x) = \frac{2x}{\theta^2}, \ 0 < x < \theta.$$

Let $T = c\overline{X}$ be an estimator of θ where c is a constant.

a) Find the mean square error (MSE) of T as a function of c (and of θ and n).

b) Find the value c that minimizes the MSE. Prove that your value is the minimizer.

6.27. (Aug. 2004 Qual): Suppose that $X_1, ..., X_n$ are iid Bernoulli(p) where $n \geq 2$ and $0 < p < 1$ is the unknown parameter.

a) Derive the UMVUE of $\nu(p)$, where $\nu(p) = e^2(p(1-p))$.

b) Find the Cramér Rao lower bound for estimating $\nu(p) = e^2(p(1-p))$.

6.28. Let $X_1, ..., X_n$ be independent identically distributed Poisson(λ) random variables. Find the UMVUE of

$$\frac{\lambda}{n} + \lambda^2.$$

6.29. Let $Y_1, ..., Y_n$ be iid Poisson(θ) random variables.

a) Find the UMVUE for θ.

b) Find the Fisher information $I_1(\theta)$.

c) Find the FCRLB for unbiased estimators of $\tau(\theta) = \theta$.

d) The MLE for θ is $\overline{Y}$. Find Var($\overline{Y}$).

e) Does the MLE achieve the FCRLB? Is this surprising? Explain.

6.30. (Jan. 2009 Qual): Suppose that $Y_1, ..., Y_n$ are independent binomial(m_i, ρ) where the $m_i \geq 1$ are known constants. Let

$$T_1 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n m_i} \quad \text{and} \quad \mathrm{T}_2 = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\mathrm{Y_i}}{\mathrm{m_i}}$$

be estimators of ρ.

a) Find MSE(T_1).

b) Find MSE(T_2).

c) Which estimator is better?

Hint: by the arithmetic–geometric–harmonic mean inequality,

$$\frac{1}{n} \sum_{i=1}^n m_i \geq \frac{1}{n} \sum_{i=1}^n \frac{1}{m_i}.$$

6.31. (Sept. 2010 Qual): Let $Y_1, ..., Y_n$ be iid gamma($\alpha = 10, \beta$) random variables. Let $T = c\overline{Y}$ be an estimator of β where c is a constant.

a) Find the mean square error (MSE) of T as a function of c (and of β and n).

b) Find the value c that minimizes the MSE. Prove that your value is the minimizer.

6.32. (Jan. 2011 Qual): Let $Y_1, ..., Y_n$ be independent identically distributed random variables with pdf (probability density function)

$$f(y) = (2 - 2y)I_{(0,1)}(y)\ \nu\ \exp[(1 - \nu)(-\log(2y - y^2))]$$

where $\nu > 0$ and $n > 1$. The indicator $I_{(0,1)}(y) = 1$ if $0 < y < 1$ and $I_{(0,1)}(y) = 0$, otherwise.

a) Find a complete sufficient statistic.

b) Find the Fisher information $I_1(\nu)$ if $n = 1$.

c) Find the Cramer Rao lower bound (CRLB) for estimating $1/\nu$.

d) Find the uniformly minimum unbiased estimator (UMVUE) of ν.

Hint: You may use the fact that $T_n = -\sum_{i=1}^n \log(2Y_i - Y_i^2) \sim G(n, 1/\nu)$, and

$$E(T_n^r) = \frac{1}{\nu^r}\ \frac{\Gamma(r+n)}{\Gamma(n)}$$

for $r > -n$. Also $\Gamma(1 + x) = x\Gamma(x)$ for $x > 0$.

6.33. (Sept. 2011 Qual): Let $Y_1, ..., Y_n$ be iid random variables from a distribution with pdf

$$f(y) = \frac{\theta}{2(1 + |y|)^{\theta+1}}$$

where $\theta > 0$ and y is real. Then $W = \log(1 + |Y|)$ has pdf $f(w) = \theta e^{-w\theta}$ for $w > 0$.

a) Find a complete sufficient statistic.

b) Find the (Fisher) information number $I_1(\theta)$.

c) Find the uniformly minimum variance unbiased estimator (UMVUE) for θ.