Stat 513
HW 1 Solutions
January 30, 2004

- 7.3.8
Let $f_1(x)$ and $f_2(x)$ be p.d.f's with known means μ_1 and μ_2 respectively. Consider the following contaminated p.d.f. for $0 \le \theta \le 1$

$$f(x|\theta) = \theta f_1(x) + (1-\theta) f_2(x)$$

Let $X_1, ..., X_n$ be a sample from $f(x|\theta)$, find the MME of θ.

$$\begin{aligned} \overline{X} = E(f(x|\theta)) &= \theta E(f_1(x)) + (1-\theta)E(f_2(x)) = \theta\mu_1 + (1-\theta)\mu_2 = \theta(\mu_1 - \mu_2) + \mu_2 \\ &\Rightarrow \frac{\overline{X} - \mu_2}{\mu_1 - \mu_2} = \tilde{\theta} \end{aligned}$$

- 7.3.11
Let $X_1, ..., X_n$ be sample from p.d.f. $f(x|\theta) = \left(\frac{\theta}{x^2}\right) 1_{[\theta,\infty)}(x)$ for some $\theta > 0$. Find the MME for θ.
First we need to know what the expectation for this p.d.f. is.

$$E(x|\theta) = \int_\theta^\infty x \cdot \frac{\theta}{x^2} dx = \theta ln(x)|_\theta^\infty \Rightarrow \text{ undefined}$$

Note: All k moments for $k \ge 1$ do not exist (convience yourself of this, just write out $E(X^2)$ and $E(X^3)$). Thus a method of moments estimator does not exist. Some people tried a "half-moment" estimator, but method of moments estimators were defined for k being an integer greater than or equal to 1.

- 7.5.1
Let $X_1, ..., X_n$ *iid* $B(1, p)$, and let $T_n = X_1 + ... + X_n$ and $\hat{p}_n = \frac{T_n}{n}$.

  (a) Show $\hat{p}_n$ is a "BRUE" for p.
  There are *two* parts to this, unbiasedness, and minimum variance.
  Unbiased:

$$E(\hat{p}_n) = \frac{1}{n}\sum_{i=1}^n E(X_i) = \frac{np}{n} = p$$

Minimum Variance:

$$\begin{aligned}
Var(\hat{p}_n) &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n} \\
L_i(p) &= p^x(1-p)^{1-x} \\
\mathcal{L}_i(p) &= xln(p) + (1-x)ln(1-p) \\
\mathcal{U}_i(p) &= \frac{x}{p} - \frac{1-x}{1-p} \\
I_i(p) &= -E\left(-\frac{x}{p^2} - \frac{1-x}{(1-p)^2}\right) = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)} \\
\text{Since we have iid } X_i\text{'s }\ I_n(p) &= \frac{n}{p(1-p)} \\
\Rightarrow Var(\hat{p}_n) &\geq \frac{[1]^2}{\frac{n}{p(1-p)}} = \frac{p(1-p)}{n} \Rightarrow \hat{p}_n \text{ is the BRUE}
\end{aligned}$$

(b) Show $T_n(n-T_n)(n-1-T_n)/d_n$ is unbiased for pq^2 and find the CR-LB for this where $d_n = n(n-1)(n-2)$.

Unbiasedness:
This is basically just algebra:

$$E\left(\frac{T_n(n-T_n)(n-1-T_n)}{n(n-1)(n-2)}\right) = \frac{1}{d_n}[n(n-1)E(T_n) + (1-2n)E(T_n^2) + E(T_n^3)]$$

From previous homework #3 in Stat 512:

$$\begin{aligned}
E(T_n) &= np \\
E(T_n^2) &= (np)^2 + np(1-p) \\
E(T_n^3) &= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np
\end{aligned}$$

$$\begin{aligned}
\text{Substitute in with algebra} &\Rightarrow \frac{n(n-1)(n-2)(p^3-2p^2+p)}{d_n} \\
&= p(1-p)^2 = pq^2
\end{aligned}$$

CR-LB:
This is the new part.
Note: pq^2 is a function of p. Let $g(p) = p(1-p)^2 = p - 2p^2 + p^3$. Then $g'(p) = 1 - 4p + 3p^2$
Now using previous results from part (a) and letting T_n^b denote the estimator of pq^2 from

above:

$$Var(T_n^b) \geq \frac{[g'(p)]^2}{I_n(p)} = \frac{[1-4p+3p^2]^2}{\frac{n}{p(1-p)}} = \frac{[1-4p+3p^2]^2 p(1-p)}{n}$$

(c) Show that $T_n(n-T_n)/(n(n-1))$ is unbiased for pq and has LARGER variance than the MLE $\hat{p}\hat{q}$.
Unbiasedness:

$$\begin{aligned} E\left[\frac{T_n(n-T_n)}{n(n-1)}\right] &= \frac{1}{n(n-1)}E[nT_n - T_n^2] \\ &= \frac{1}{n(n-1)}\left(nE[T_n] - Var[T_n] - (E[T_n])^2\right) \\ &= \frac{(n^2p - np(1-p) - n^2p^2)}{n(n-1)} = \frac{np - np^2 - p + p^2}{(n-1)} \\ &= \frac{(n-1)(p-p^2)}{(n-1)} = p(1-p) = pq \end{aligned}$$

Variance:

$$\begin{aligned} \hat{p}\hat{q} &= \frac{T_n}{n}(1-\frac{T_n}{n}) \text{ by invariance of MLE's} \\ Var(\hat{p}\hat{q}) = Var[\frac{1}{n^2}T_n(n-T_n)] &= \frac{1}{n^4}Var[T_n(n-T_n)] \\ Var[\frac{1}{n(n-1)}T_n(n-T_n)] &= \frac{1}{n^2(n-1)^2}Var[T_n(n-T_n)] \end{aligned}$$

Now since $\frac{1}{n^4} < \frac{1}{n^2(n-1)^2}$, the MLE has a smaller variance.

- 7.5.2

(a) Let $X_1, ..., X_n$ *iid* $N(\mu, 1)$, and T_{1n} be the MLE of μ^2 and T_{2n} be an unbiased estimator of μ^2. Calculate the CR-LB for each and compare using consistency and unbiasedness.
Recall that the mle of μ is $\overline{X}$, so by invariance of MLE, the MLE of μ^2 is $\overline{X}^2$. Also, many people specified a specific unbiased estimator for T_{2n}, but the CR-LB holds for every unbiased estimator.

$$L(\mu) = \prod_{i=1}^{n}(2\pi)^{-n/2}e^{-\frac{1}{2}(x_i-\mu)^2}$$

$$
\begin{aligned}
\mathcal{L}_i(\mu) &= -\frac{1}{2}(x_i-\mu)^2 - \frac{n}{2}\log(2\pi) \\
\mathcal{U}_i(\mu) &= (x_i-\mu) \\
I_i(\mu) &= 1 \\
E(T_{1n}) &= E(\overline{X}^2) = Var(\overline{X}) + E(\overline{X})^2 \\
&= \frac{1}{n} + \mu^2 \\
bias(T_{1n}) &= E(T_{1n}) - \mu^2 = \frac{1}{n} \\
b'(T_{1n}) &= 0 \\
bias(T_{2n}) &= 0 \text{ by assumption} \\
g(\mu) &= \mu^2 \\
g'(\mu) &= 2\mu \text{ for both T}
\end{aligned}
$$

So the CR-LB for both T_{1n} and T_{2n} is $\frac{4\mu^2}{n}$. The decision whether to use T_{1n} or T_{2n} depends entirely upon the variance of T_{2n}. The no-brainer is to subtract off the bias from T_{1n} and use that as your estimator.

(b) Find CR-LB for the unbiased estimator of $P(X > 2\mu)$.

$$
\begin{aligned}
P(X > 2\mu) &= P(X - \mu > \mu) \\
&= P(Z > \mu)\,,\; Z \sim N(0,1) \\
&= \Phi(\mu) \\
g(\mu) &= \Phi(\mu) \\
g'(\mu) &= \phi(\mu) \\
&= (2\pi)^{-1/2} e^{-\frac{\mu^2}{2}} \\
Var(T) &\geq \frac{[g'(\mu)]^2}{n} \\
&= \frac{e^{-\mu^2}}{2\pi n}
\end{aligned}
$$

- 7.5.3
  Let $X_1, ..., X_n$ *iid* $\Gamma(\alpha,\lambda)$, T_1 be an unbiased estimator of α with λ known, and T_2 be an unbiased estimator of λ with α known. Find the CR-LB for $Var(T_1)$ and $Var(T_2)$. Is either estimator 'BRUE'?

$$
L_i(\alpha,\lambda) = \frac{1}{\lambda^{\alpha+1}\Gamma(\alpha+1)} x_i^{\alpha} e^{-x_i/\lambda}
$$

$$
\begin{aligned}
\mathcal{L}_i(\alpha,\lambda) &= -(\alpha+1)\log(\lambda) - \log(\Gamma(\alpha+1)) + \alpha\log(x_i) - \frac{x_i}{\lambda} \\
\mathcal{U}_i(\alpha|\lambda) &= -\log\lambda - \Psi(\alpha+1) + \log(x_i) \\
& \quad \text{Where } \Psi(x) = \tfrac{\Gamma'(x)}{\Gamma(x)} \\
I(\alpha|\lambda) &= E(-\Psi'(\alpha+1)) = -\Psi'(\alpha+1) \\
\mathcal{U}_i(\lambda|\alpha) &= -\frac{\alpha+1}{\lambda} + \frac{x_i}{\lambda^2} \\
&= \frac{\alpha+1}{\lambda^2}(\frac{x_i}{\alpha+1} - \lambda) \\
I(\lambda|\alpha) &= E(-\frac{\alpha+1}{\lambda^2} + \frac{2x_i}{\lambda^3}) \\
&= \frac{2\alpha}{\lambda^2} - \frac{\alpha+1}{\lambda^2} \\
&= \frac{\alpha-1}{\lambda^2}
\end{aligned}
$$

Notice that we could put the score function in the form $A(\lambda)(T_2(X)-\lambda)$ for $T_2 = \frac{\overline{X}}{\alpha+1}$, so it is BRUE. We cannot change the digamma into that form, so no T_1 is BRUE. The CR-LB for T_1 is $\frac{-1}{\Psi'(\alpha+1)}$, and the CR-LB for T_2 is $\frac{\lambda^2}{\alpha-1}$.

- 7.5.7
  Let $X_1, ..., X_n$ *iid* $Pois(\lambda)$ and $T_n = (\overline{X})^2 - \overline{X}$. Find the lower bound for $E(T_n - \lambda^2)^2$, then find a function $g(T_n)$ that is unbiased for λ^2 and find the lower bound for this $g(T_n)$.

$$
\begin{aligned}
E(T_n) &= E(\overline{X}^2) - E(\overline{X}) \\
&= Var(\overline{X}) + E(\overline{X})^2 - E(\overline{X}) \\
&= \frac{\lambda}{n} + \lambda^2 - \lambda \\
&= \lambda^2 - \frac{n-1}{n}\lambda \\
E(T_n - \lambda^2)^2 &= MSE(T_n) \\
&= Var(T_n) + b^2(T_n) \\
bias(T_n) &= -\frac{n-1}{n}\lambda \\
b'(\lambda) &= -\frac{n-1}{n} \\
g(\lambda) &= \lambda^2 \\
g'(\lambda) &= 2\lambda \\
I(\lambda) &= \frac{1}{\lambda} \text{ by Ex 7.5.6} \\
Var(T_n) &\geq \frac{\lambda(2\lambda - \frac{n-1}{n})^2}{n}
\end{aligned}
$$

$$MSE(T_n) \geq \frac{\lambda(2\lambda - \frac{n-1}{n})^2}{n} + \frac{\lambda^2(n-1)^2}{n^2}$$

- 7.5.10
  Let $X_1, ..., X_n$ *iid* $N(\mu, \sigma^2)$ with both μ and σ^2 known. Three possible esimators for σ^2 are:

$$t_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)} = s^2, \ \ t_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \hat{\sigma}^2, \ \ t_3 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n+1)}$$

  Compare with respect to bias, variance, and mean squared error.
  Bias:

$$\begin{aligned} E(t_1) &= \sigma^2 \text{ (done in previous homework) } \Rightarrow b(t_1) = 0 \\ E(t_2) &= \frac{n-1}{n} \cdot E(t_1) = \frac{n-1}{n}\sigma^2 \Rightarrow b(t_2) = -\frac{1}{n}\sigma^2 \\ E(t_3) &= \frac{n-1}{n+1} \cdot E(t_1) = \frac{n-1}{n+1}\sigma^2 \Rightarrow b(t_3) = -\frac{2}{n+1}\sigma^2 \end{aligned}$$

  Variance:
  Note: Since the X_i's are normal, $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1} = \Gamma(\frac{n-1}{2}, 2)$.
  Thus,

$$\begin{aligned} \left(\frac{n-1}{\sigma^2}\right)^2 Var(s^2) &= 2(n-1) \text{ (since } Var(\Gamma(a,b)) = ab^2) \\ \Rightarrow Var(s^2) &= 2(n-1)\frac{\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{(n-1)} \end{aligned}$$

$$\begin{aligned} Var(t_1) &= Var(s^2) = \frac{1}{(n-1)}2\sigma^4 \\ Var(t_2) &= \left(\frac{n-1}{n}\right)^2 Var(s^2) = \frac{(n-1)}{n^2}2\sigma^4 \\ Var(t_3) &= \left(\frac{n-1}{n+1}\right)^2 Var(s^2) = \frac{(n-1)}{(n+1)^2}2\sigma^4 \end{aligned}$$

  Mean Squared Error:

$$MSE(t_1) = \frac{1}{(n-1)}2\sigma^4 + 0 = \frac{2}{(n-1)}\sigma^4$$

$$
\begin{aligned}
MSE(t_2) &= \frac{(n-1)}{n^2}2\sigma^4 + \left(-\frac{1}{n}\sigma^2\right)^2 = \frac{2n-1}{n^2}\sigma^4 \\
MSE(t_3) &= \frac{(n-1)}{(n+1)^2}2\sigma^4 + \left(-\frac{2}{n+1}\sigma^2\right)^2 = \frac{2}{n+1}\sigma^4
\end{aligned}
$$

Moral of the story: There are instances where you can beat an unbiased estimator in MSE with a bias one. Thus, while it seems like a good idea to look at unbiased estimators, they aren't neccessarily the only estimators that are reasonable for a given situation. Also, MLE's aren't always the best idea for an estimator either (t_2 was the MLE and wasn't the best in anything).

- Supplemental Problem
  Let $Y_1, ..., Y_n$ *iid* $B(1, p_i)$ with

$$
logit\ (p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_i
$$

for known covariate x_i. For simplicity, let $x_i\beta = \beta_0 + \beta_1 x_i$
Note:

$$
\frac{p_i}{1-p_i} = e^{x_i\beta} \Longrightarrow p_i = \frac{e^{x_i\beta}}{1+e^{x_i\beta}}
$$

(i) Likelihood:

$$
\begin{aligned}
L(\beta) &= \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i} = \prod_{i=1}^{n}\left(\frac{p_i}{1-p_i}\right)^{y_i}(1-p_i) \\
\text{substitute in for } p_i &\Rightarrow \prod_{i=1}^{n}\frac{e^{(x_i\beta)y_i}}{1+e^{x_i\beta}}
\end{aligned}
$$

(ii) Log-likelihood:

$$
\mathcal{L}(\beta) = \sum_{i=1}^{n}\left[(x_i\beta)y_i - ln(1+e^{x_i\beta})\right]
$$

(iii) Score:

$$
\begin{aligned}
\beta_0 &\Rightarrow \sum_{i=1}^{n}\left[y_i - \frac{1}{(1+e^{x_i\beta})}e^{x_i\beta}\right] \\
\beta 1 &\Rightarrow \sum_{i=1}^{n}\left[x_i y_i - \frac{x_i}{(1+e^{x_i\beta})}e^{x_i\beta}\right]
\end{aligned}
$$

(iv) Fisher's information:
Four parts here for a 2x2 information matrix.

$$
\begin{aligned}
-E\left(\frac{d\ score(\beta_0)}{d\ \beta_0}\right) &= \sum_{i=1}^{n} \frac{e^{-x_i\beta}}{(1+e^{-x_i\beta})^2} \\
-E\left(\frac{d\ score(\beta_0)}{d\ \beta_1}\right) = -E\left(\frac{d\ score(\beta_1)}{d\ \beta_0}\right) &= \sum_{i=1}^{n} \frac{x_i e^{-x_i\beta}}{(1+e^{-x_i\beta})^2} \\
-E\left(\frac{d\ score(\beta_1)}{d\ \beta_1}\right) &= \sum_{i=1}^{n} \frac{x_i^2 e^{-x_i\beta}}{(1+e^{-x_i\beta})^2}
\end{aligned}
$$