Stat Inf HW3 solutions 4/10/08

7.9 This is a uniform$(0,\theta)$ model. So $\mathrm{E}X = (0+\theta)/2 = \theta/2$. The method of moments estimator is the solution to the equation $\tilde{\theta}/2 = \bar{X}$, that is, $\tilde{\theta} = 2\bar{X}$. Because $\tilde{\theta}$ is a simple function of the sample mean, its mean and variance are easy to calculate. We have

$$\mathrm{E}\tilde{\theta} = 2\mathrm{E}\bar{X} = 2\mathrm{E}X = 2\frac{\theta}{2} = \theta, \quad \text{and} \quad \mathrm{Var}\,\tilde{\theta} = 4\mathrm{Var}\,\bar{X} = 4\frac{\theta^2/12}{n} = \frac{\theta^2}{3n}.$$

The likelihood function is

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\theta} I_{[0,\theta]}(x_i) = \frac{1}{\theta^n} I_{[0,\theta]}(x_{(n)}) I_{[0,\infty)}(x_{(1)}),$$

where $x_{(1)}$ and $x_{(n)}$ are the smallest and largest order statistics. For $\theta \geq x_{(n)}$, $L = 1/\theta^n$, a decreasing function. So for $\theta \geq x_{(n)}$, L is maximized at $\hat{\theta} = x_{(n)}$. $L = 0$ for $\theta < x_{(n)}$. So the overall maximum, the MLE, is $\hat{\theta} = X_{(n)}$. The pdf of $\hat{\theta} = X_{(n)}$ is nx^{n-1}/θ^n, $0 \leq x \leq \theta$. This can be used to calculate

$$\mathrm{E}\hat{\theta} = \frac{n}{n+1}\theta, \quad \mathrm{E}\hat{\theta}^2 = \frac{n}{n+2}\theta^2 \quad \text{and} \quad \mathrm{Var}\,\hat{\theta} = \frac{n\theta^2}{(n+2)(n+1)^2}.$$

$\tilde{\theta}$ is an unbiased estimator of θ; $\hat{\theta}$ is a biased estimator. If n is large, the bias is not large because $n/(n+1)$ is close to one. But if n is small, the bias is quite large. On the other hand, $\mathrm{Var}\,\hat{\theta} < \mathrm{Var}\,\tilde{\theta}$ for all θ. So, if n is large, $\hat{\theta}$ is probably preferable to $\tilde{\theta}$.

7.10 a. $f(\mathbf{x}|\theta) = \prod_i \frac{\alpha}{\beta^\alpha} x_i^{\alpha-1} I_{[0,\beta]}(x_i) = \left(\frac{\alpha}{\beta^\alpha}\right)^n \left(\prod_i x_i\right)^{\alpha-1} I_{(-\infty,\beta]}(x_{(n)}) I_{[0,\infty)}(x_{(1)}) = L(\alpha,\beta|\mathbf{x})$. By the Factorization Theorem, $(\prod_i X_i, X_{(n)})$ are sufficient.

b. For any fixed α, $L(\alpha,\beta|\mathbf{x}) = 0$ if $\beta < x_{(n)}$, and $L(\alpha,\beta|\mathbf{x})$ a decreasing function of β if $\beta \geq x_{(n)}$. Thus, $X_{(n)}$ is the MLE of β. For the MLE of α calculate

$$\frac{\partial}{\partial\alpha}\log L = \frac{\partial}{\partial\alpha}\left[n\log\alpha - n\alpha\log\beta + (\alpha-1)\log\prod_i x_i\right] = \frac{n}{\alpha} - n\log\beta + \log\prod_i x_i.$$

Set the derivative equal to zero and use $\hat{\beta} = X_{(n)}$ to obtain

$$\hat{\alpha} = \frac{n}{n\log X_{(n)} - \log\prod_i X_i} = \left[\frac{1}{n}\sum_i (\log X_{(n)} - \log X_i)\right]^{-1}.$$

The second derivative is $-n/\alpha^2 < 0$, so this is the MLE.

c. $X_{(n)} = 25.0$, $\log\prod_i X_i = \sum_i \log X_i = 43.95 \Rightarrow \hat{\beta} = 25.0$, $\hat{\alpha} = 12.59$.

7.12 $X_i \sim$ iid Bernoulli(θ), $0 \le \theta \le 1/2$.

a. method of moments:

$$\mathrm{E}\,X = \theta = \frac{1}{n}\sum_i X_i = \bar{X} \quad \Rightarrow \quad \tilde{\theta} = \bar{X}.$$

MLE: In Example 7.2.7, we showed that $L(\theta|\mathbf{x})$ is increasing for $\theta \le \bar{x}$ and is decreasing for $\theta \ge \bar{x}$. Remember that $0 \le \theta \le 1/2$ in this exercise. Therefore, when $\bar{X} \le 1/2$, $\bar{X}$ is the MLE of θ, because $\bar{X}$ is the overall maximum of $L(\theta|\mathbf{x})$. When $\bar{X} > 1/2$, $L(\theta|\mathbf{x})$ is an increasing function of θ on $[0, 1/2]$ and obtains its maximum at the upper bound of θ which is $1/2$. So the MLE is $\hat{\theta} = \min\{\bar{X}, 1/2\}$.

b. The MSE of $\tilde{\theta}$ is $\mathrm{MSE}(\tilde{\theta}) = \mathrm{Var}\,\tilde{\theta} + \mathrm{bias}(\tilde{\theta})^2 = (\theta(1-\theta)/n) + 0^2 = \theta(1-\theta)/n$. There is no simple formula for $\mathrm{MSE}(\hat{\theta})$, but an expression is

$$\begin{aligned}
\mathrm{MSE}(\hat{\theta}) &= \mathrm{E}(\hat{\theta}-\theta)^2 = \sum_{y=0}^{n}(\hat{\theta}-\theta)^2\binom{n}{y}\theta^y(1-\theta)^{n-y} \\
&= \sum_{y=0}^{[n/2]}\left(\frac{y}{n}-\theta\right)^2\binom{n}{y}\theta^y(1-\theta)^{n-y} + \sum_{y=[n/2]+1}^{n}\left(\frac{1}{2}-\theta\right)^2\binom{n}{y}\theta^y(1-\theta)^{n-y},
\end{aligned}$$

where $Y = \sum_i X_i \sim$ binomial(n,θ) and $[n/2] = n/2$, if n is even, and $[n/2] = (n-1)/2$, if n is odd.

c. Using the notation used in (b), we have

$$\mathrm{MSE}(\tilde{\theta}) = \mathrm{E}(\bar{X}-\theta)^2 = \sum_{y=0}^{n}\left(\frac{y}{n}-\theta\right)^2\binom{n}{y}\theta^y(1-\theta)^{n-y}.$$

Therefore,

$$\begin{aligned}
\mathrm{MSE}(\tilde{\theta}) - \mathrm{MSE}(\hat{\theta}) &= \sum_{y=[n/2]+1}^{n}\left[\left(\frac{y}{n}-\theta\right)^2 - \left(\frac{1}{2}-\theta\right)^2\right]\binom{n}{y}\theta^y(1-\theta)^{n-y} \\
&= \sum_{y=[n/2]+1}^{n}\left(\frac{y}{n}+\frac{1}{2}-2\theta\right)\left(\frac{y}{n}-\frac{1}{2}\right)\binom{n}{y}\theta^y(1-\theta)^{n-y}.
\end{aligned}$$

The facts that $y/n > 1/2$ in the sum and $\theta \le 1/2$ imply that every term in the sum is positive. Therefore $\mathrm{MSE}(\hat{\theta}) < \mathrm{MSE}(\tilde{\theta})$ for every θ in $0 < \theta \le 1/2$. (Note: $\mathrm{MSE}(\hat{\theta}) = \mathrm{MSE}(\tilde{\theta}) = 0$ at $\theta = 0$.)

7.24 For n observations, $Y = \sum_i X_i \sim \text{Poisson}(n\lambda)$.

a. The marginal pmf of Y is

$$\begin{aligned} m(y) &= \int_0^\infty \frac{(n\lambda)^y e^{-n\lambda}}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta}\, d\lambda \\ &= \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \int_0^\infty \lambda^{(y+\alpha)-1} e^{-\frac{\lambda}{\beta/(n\beta+1)}}\, d\lambda \;=\; \frac{n^y}{y!\Gamma(\alpha)\beta^\alpha} \Gamma(y+\alpha) \left(\frac{\beta}{n\beta+1}\right)^{y+\alpha}. \end{aligned}$$

Thus,

$$\pi(\lambda|y) = \frac{f(y|\lambda)\pi(\lambda)}{m(y)} = \frac{\lambda^{(y+\alpha)-1} e^{-\frac{\lambda}{\beta/(n\beta+1)}}}{\Gamma(y+\alpha)\left(\frac{\beta}{n\beta+1}\right)^{y+\alpha}} \sim \text{gamma}\left(y+\alpha, \frac{\beta}{n\beta+1}\right).$$

b.

$$\begin{aligned} \mathrm{E}(\lambda|y) &= (y+\alpha)\frac{\beta}{n\beta+1} \;=\; \frac{\beta}{n\beta+1} y + \frac{1}{n\beta+1}(\alpha\beta). \\ \mathrm{Var}(\lambda|y) &= (y+\alpha)\frac{\beta^2}{(n\beta+1)^2}. \end{aligned}$$

7.37 To find a best unbiased estimator of θ, first find a complete sufficient statistic. The joint pdf is

$$f(\mathrm{x}|\theta) = \left(\frac{1}{2\theta}\right)^n \prod_i I_{(-\theta,\theta)}(x_i) = \left(\frac{1}{2\theta}\right)^n I_{[0,\theta)}(\max_i |x_i|).$$

By the Factorization Theorem, $\max_i |X_i|$ is a sufficient statistic. To check that it is a complete sufficient statistic, let $Y = \max_i |X_i|$. Note that the pdf of Y is $f_Y(y) = ny^{n-1}/\theta^n$, $0 < y < \theta$. Suppose $g(y)$ is a function such that

$$\mathrm{E}\, g(Y) = \int_0^\theta \frac{ny^{n-1}}{\theta^n} g(y)\, dy = 0, \text{ for all } \theta.$$

Taking derivatives shows that $\theta^{n-1} g(\theta) = 0$, for all θ. So $g(\theta) = 0$, for all θ, and $Y = \max_i |X_i|$ is a complete sufficient statistic. Now

$$\mathrm{E}\, Y = \int_0^\theta y \frac{ny^{n-1}}{\theta^n}\, dy = \frac{n}{n+1}\theta \quad \Rightarrow \quad \mathrm{E}\left(\frac{n+1}{n} Y\right) = \theta.$$

Therefore $\frac{n+1}{n}\max_i |X_i|$ is a best unbiased estimator for θ because it is a function of a complete sufficient statistic. (Note that $\left(X_{(1)}, X_{(n)}\right)$ is not a minimal sufficient statistic (recall Exercise 5.36). It is for $\theta < X_i < 2\theta$, $-2\theta < X_i < \theta$, $4\theta < X_i < 6\theta$, etc., but not when the range is symmetric about zero. Then $\max_i |X_i|$ is minimal sufficient.)

7.38 Use Corollary 7.3.15.

a.

$$\begin{aligned} \frac{\partial}{\partial\theta} \log L(\theta|\mathrm{x}) &= \frac{\partial}{\partial\theta} \log \prod_i \theta x_i^{\theta-1} \;=\; \frac{\partial}{\partial\theta} \sum_i \left[\log\theta + (\theta-1)\log x_i\right] \\ &= \sum_i \left[\frac{1}{\theta} + \log x_i\right] \;=\; -n\left[-\sum_i \frac{\log x_i}{n} - \frac{1}{\theta}\right]. \end{aligned}$$

Thus, $-\sum_i \log X_i / n$ is the UMVUE of $1/\theta$ and attains the Cramér-Rao bound.

b.

$$\begin{aligned}\frac{\partial}{\partial\theta}\log L(\theta|\mathrm{x}) &= \frac{\partial}{\partial\theta}\log\prod_i\frac{\log\theta}{\theta-1}\theta^{x_i} = \frac{\partial}{\partial\theta}\sum_i[\log\log\theta-\log(\theta-1)+x_i\log\theta]\\ &= \sum_i\left(\frac{1}{\theta\log\theta}-\frac{1}{\theta-1}\right)+\frac{1}{\theta}\sum_i x_i = \frac{n}{\theta\log\theta}-\frac{n}{\theta-1}+\frac{n\bar{x}}{\theta}\\ &= \frac{n}{\theta}\left[\bar{x}-\left(\frac{\theta}{\theta-1}-\frac{1}{\log\theta}\right)\right].\end{aligned}$$

Thus, $\bar{X}$ is the UMVUE of $\frac{\theta}{\theta-1}-\frac{1}{\log\theta}$ and attains the Cramér-Rao lower bound.

Note: We claim that if $\frac{\partial}{\partial\theta}\log L(\theta|\mathbf{X}) = a(\theta)[W(\mathbf{X})-\tau(\theta)]$, then $\mathrm{E}W(\mathbf{X}) = \tau(\theta)$, because under the condition of the Cramér-Rao Theorem, $\mathrm{E}\frac{\partial}{\partial\theta}\log L(\theta|\mathrm{x}) = 0$. To be rigorous, we need to check the "interchange differentiation and integration" condition. Both (a) and (b) are exponential families, and this condition is satisfied for all exponential families.

7.40

$$\begin{aligned}\frac{\partial}{\partial\theta}\log L(\theta|\mathrm{x}) &= \frac{\partial}{\partial p}\log\prod_i p^{x_i}(1-p)^{1-x_i} = \frac{\partial}{\partial p}\sum_i x_i\log p+(1-x_i)\log(1-p)\\ &= \sum_i\left[\frac{x_i}{p}-\frac{(1-x_i)}{1-p}\right] = \frac{n\bar{x}}{p}-\frac{n-n\bar{x}}{1-p} = \frac{n}{p(1-p)}[\bar{x}-p].\end{aligned}$$

By Corollary 7.3.15, $\bar{X}$ is the UMVUE of p and attains the Cramér-Rao lower bound. Alternatively, we could calculate

$$\begin{aligned}&-n\mathrm{E}_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right)\\ &= -n\mathrm{E}\left(\frac{\partial^2}{\partial p^2}\log\left[p^X(1-p)^{1-X}\right]\right) = -n\mathrm{E}\left(\frac{\partial^2}{\partial p^2}[X\log p+(1-X)\log(1-p)]\right)\\ &= -n\mathrm{E}\left(\frac{\partial}{\partial p}\left[\frac{X}{p}-\frac{(1-X)}{1-p}\right]\right) = -n\mathrm{E}\left(\frac{-X}{p^2}-\frac{1-X}{(1-p)^2}\right)\\ &= -n\left(-\frac{1}{p}-\frac{1}{1-p}\right) = \frac{n}{p(1-p)}.\end{aligned}$$

Then using $\tau(\theta) = p$ and $\tau'(\theta) = 1$,

$$\frac{\tau'(\theta)}{-n\mathrm{E}_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right)} = \frac{1}{n/p(1-p)} = \frac{p(1-p)}{n} = \mathrm{Var}\bar{X}.$$

We know that $\mathrm{E}\bar{X} = p$. Thus, $\bar{X}$ attains the Cramér-Rao bound.

7.41 a. $\mathrm{E}\left(\sum_i a_iX_i\right) = \sum_i a_i\mathrm{E}\,X_i = \sum_i a_i\mu = \mu\sum_i a_i = \mu$. Hence the estimator is unbiased.

b. $\mathrm{Var}\left(\sum_i a_iX_i\right) = \sum_i a_i^2\mathrm{Var}\,X_i = \sum_i a_i^2\sigma^2 = \sigma^2\sum_i a_i^2$. Therefore, we need to minimize $\sum_i a_i^2$, subject to the constraint $\sum_i a_i = 1$. Add and subtract the mean of the a_i, $1/n$, to get

$$\sum_i a_i^2 = \sum_i\left[\left(a_i-\frac{1}{n}\right)+\frac{1}{n}\right]^2 = \sum_i\left(a_i-\frac{1}{n}\right)^2+\frac{1}{n},$$

because the cross-term is zero. Hence, $\sum_i a_i^2$ is minimized by choosing $a_i = 1/n$ for all i. Thus, $\sum_i(1/n)X_i = \bar{X}$ has the minimum variance among all linear unbiased estimators.

7.48 a. The Cramér-Rao Lower Bound for unbiased estimates of p is

$$\frac{\left[\frac{d}{dp}p\right]^2}{-n\mathrm{E}\frac{d^2}{dp^2}\log L(p|X)} = \frac{1}{-n\mathrm{E}\left\{\frac{d^2}{dp^2}\log[p^X(1-p)^{1-X}]\right\}} = \frac{1}{-n\mathrm{E}\left\{-\frac{X}{p^2}-\frac{(1-X)}{(1-p)^2}\right\}} = \frac{p(1-p)}{n},$$

because $\mathrm{E}\,X = p$. The MLE of p is $\hat{p} = \sum_i X_i/n$, with $\mathrm{E}\,\hat{p} = p$ and $\mathrm{Var}\,\hat{p} = p(1-p)/n$. Thus $\hat{p}$ attains the CRLB and is the best unbiased estimator of p.

b. By independence, $\mathrm{E}(X_1X_2X_3X_4) = \prod_i \mathrm{E}\,X_i = p^4$, so the estimator is unbiased. Because $\sum_i X_i$ is a complete sufficient statistic, Theorems 7.3.17 and 7.3.23 imply that $\mathrm{E}(X_1X_2X_3X_4|\sum_i X_i)$ is the best unbiased estimator of p^4. Evaluating this yields

$$\begin{aligned}\mathrm{E}\left(X_1X_2X_3X_4\middle|\sum_i X_i = t\right) &= \frac{P(X_1 = X_2 = X_3 = X_4 = 1, \sum_{i=5}^n X_i = t-4)}{P(\sum_i X_i = t)} \\ &= \frac{p^4\binom{n-4}{t-4}p^{t-4}(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}} = \binom{n-4}{t-4}\Big/\binom{n}{t},\end{aligned}$$

for $t \geq 4$. For $t < 4$ one of the X_is must be zero, so the estimator is $\mathrm{E}(X_1X_2X_3X_4|\sum_i X_i = t) = 0$.

7.50 a. $\mathrm{E}(a\bar{X} + (1-a)cS) = a\mathrm{E}\,\bar{X} + (1-a)\mathrm{E}(cS) = a\theta + (1-a)\theta = \theta$. So $a\bar{X} + (1-a)cS$ is an unbiased estimator of θ.

b. Because $\bar{X}$ and S^2 are independent for this normal model, $\mathrm{Var}(a\bar{X}+(1-a)cS) = a^2V_1 + (1-a)^2V_2$, where $V_1 = \mathrm{Var}\bar{X} = \theta^2/n$ and $V_2 = \mathrm{Var}(cS) = c^2\mathrm{E}\,S^2 - \theta^2 = c^2\theta^2 - \theta^2 = (c^2-1)\theta^2$. Use calculus to show that this quadratic function of a is minimized at

$$a = \frac{V_2}{V_1+V_2} = \frac{(c^2-1)\theta^2}{((1/n)+c^2-1)\theta^2} = \frac{(c^2-1)}{((1/n)+c^2-1)}.$$

c. Use the factorization in Example 6.2.9, with the special values $\mu = \theta$ and $\sigma^2 = \theta^2$, to show that $(\bar{X}, S^2)$ is sufficient. $\mathrm{E}(\bar{X} - cS) = \theta - \theta = 0$, for all θ. So $\bar{X} - cS$ is a nonzero function of $(\bar{X}, S^2)$ whose expected value is always zero. Thus $(\bar{X}, S^2)$ is not complete.