Mathematical Statistics I
Assignment 6 - Solutions

1. (a) T is a Bernoulli random variable. Hence,

$$\mathrm{E}_p(T) = P_p(T=1) = P_p\left(\sum_{i=1}^{n} X_i > X_{n+1}\right) = h(p).$$

(b) $\sum_{i=1}^{n+1} X_i$ is a complete sufficient statistic for θ, so $\mathrm{E}\left(T \middle| \sum_{i=1}^{n+1} X_i\right)$ is the best unbiased estimator of $h(p)$. We have

$$\begin{aligned}\mathrm{E}\left(T \middle| \sum_{i=1}^{n+1} X_i = y\right) &= P\left(\sum_{i=1}^{n} X_i > X_{n+1} \middle| \sum_{i=1}^{n+1} X_i = y\right) \\ &= P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = y\right) \Big/ P\left(\sum_{i=1}^{n+1} X_i = y\right)\end{aligned}$$

The denominator equals $\binom{n+1}{y} p^y (1-p)^{n+1-y}$. If $y=0$ the numerator is

$$P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = 0\right) = 0.$$

If $y > 0$ the numerator is

$$P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = y, X_{n+1} = 0\right) + P\left(\sum_{i=1}^{n} X_i > X_{n+1}, \sum_{i=1}^{n+1} X_i = y, X_{n+1} = 1\right)$$

which equals

$$P\left(\sum_{i=1}^{n} X_i > 0, \sum_{i=1}^{n} X_i = y\right) P\left(X_{n+1} = 0\right) + P\left(\sum_{i=1}^{n} X_i > 1, \sum_{i=1}^{n} X_i = y-1\right) P\left(X_{n+1} = 1\right).$$

For all $y > 0$,

$$P\left(\sum_{i=1}^{n} X_i > 0, \sum_{i=1}^{n} X_i = y\right) = P\left(\sum_{i=1}^{n} X_i = y\right) = \binom{n}{y} p^y (1-p)^{n-y}.$$

If $y = 1$ or 2, then

$$P\left(\sum_{i=1}^{n} X_i > 1, \sum_{i=1}^{n} X_i = y-1\right) = 0.$$

And if $y > 2$, then

$$P\left(\sum_{i=1}^{n} X_i > 1, \sum_{i=1}^{n} X_i = y-1\right) = P\left(\sum_{i=1}^{n} X_i = y-1\right) = \binom{n}{y-1} p^{y-1}(1-p)^{n-y+1}.$$

Therefore, the UMVUE is

$$\mathrm{E}\left(T \middle| \sum_{i=1}^{n+1} X_i = y\right) = \begin{cases} 0 & \text{if } y = 0 \\ \frac{\binom{n}{y} p^y (1-p)^{n-y}(1-p)}{\binom{n+1}{y} p^y (1-p)^{n-y+1}} = \frac{\binom{n}{y}}{\binom{n+1}{y}} = \frac{1}{(n+1)(n+1-y)} & \text{if } y = 1 \text{ or } 2 \\ \frac{\left(\binom{n}{y} + \binom{n}{y-1}\right) p^y (1-p)^{n-y+1}}{\binom{n+1}{y} p^y (1-p)^{n-y+1}} = \frac{\binom{n}{y} + \binom{n}{y-1}}{\binom{n+1}{y}} = 1 & \text{if } y > 2. \end{cases}$$

2. (a) First, we have to show that $Y = X_{(n)}$ is a sufficient statistic. The pdf of Y is

$$\begin{aligned} f_Y(y;\theta) &= P(Y \le y) - P(Y \le y-1) \\ &= [y/\theta]^n - [(y-1)/\theta]^n, \quad y = 1, 2, \ldots, \theta. \end{aligned}$$

Since the quotient,

$$\frac{f(x_1, x_2, \ldots, x_n; \theta)}{f_Y(y;\theta)} = \frac{(1/\theta)^n}{(1/\theta)^n \left[y^n - (y-1)^n\right]}.$$

does not depend upon θ, by definition, Y is a sufficient statistic for θ. Now we have to prove the completeness. Let $u(Y)$ be any function of the sufficient statistic Y. Assume $E_\theta(u(Y)) = 0$, we have,

$$\sum_{y=1}^{\theta} u(y) f_Y(y;\theta) = (1/\theta)^n \sum_{y=1}^{\theta} u(y)\left[y^{n+1} - (y-1)^{n+1}\right] = 0,$$

note that for $y = 1, \ldots, \theta$, $(1/\theta)^n \left[y^{n+1} - (y-1)^{n+1}\right] > 0$ hence $E_\theta(u(Y)) = 0$, for $\theta \in \mathbb{N}$ implies that $u(x) = 0$ for $x \in \mathbb{N}$. Therefore u is zero with probability 1 and Y is indeed a complete sufficient statistic.

(b) Let $v(Y) = \left[Y^{n+1} - (Y-1)^{n+1}\right] / \left[Y^n - (Y-1)^n\right]$, by Lehmann and Scheffé's Theorem it suffice to show that $v(Y)$ is an unbiased estimator of θ. The expected value of $v(Y)$ is,

$$\sum_{y=1}^{\theta} v(y) f_Y(y;\theta) = (1/\theta^n) \sum_{y=1}^{\theta} \left[y^{n+1} - (y-1)^{n+1}\right].$$

Clearly, by substituting $y = 1, 2, \ldots, \theta$, the summation equals θ^{n+1}, hence

$$E[v(Y)] = \left(\frac{1}{\theta^n}\right) \theta^{n+1} = \theta.$$

3. Notice that $X_1, \ldots, X_n$ are iid observations from an exponential family with pdf of the form

$$f(x;\theta) = h(x)c(\theta)\exp\left(w(\theta)\,t(x)\right),$$

where $h(x) = 1$, $c(\theta) = w(\theta) = 1/\theta$, $t(x) = x$. Hence the statistic

$$T(\boldsymbol{X}) = \sum_{i=1}^{n} t(X_i) = \sum_{i=1}^{n} X_i$$

is complete. Now note that $E(nX_{(1)}|\sum_i X_i)$ is a function of T, hence by Lehmann and Scheffé's Theorem, $E(nX_{(1)}|\sum_i X_i)$ is a unique MVUE for it's expected value, which is $E(nX_{(1)})$. We know that $nX_{(1)} \sim E(\theta)$, hence $E(nX_{(1)}) = \theta$, so our initial problem of finding $E(nX_{(1)}|\sum_i X_i)$ turns into achieving the UMVUE for θ which is easily seen to be $\overline{X}$ for that $E(\overline{X}) = \theta$ and $\overline{X}$ is a function of the complete statistic T.

4. Let $X_{(j)}$ be the jth order statistic. Then $\left(X_{(1)}, X_{(n)}\right)$ is complete and sufficient for (θ_1, θ_2). Hence, it suffices to find a function of $\left(X_{(1)}, X_{(n)}\right)$ that is unbiased for the parameter of interest. Let $Y_i = \left[X_i - (\theta_1 - \theta_2)\right] / (2\theta_2)$, $i = 1, \ldots, n$. Then Y_i's are independent and identically distributed as the uniform distribution on the interval $(0,1)$. Let $Y_{(j)}$ be the jth order statistic of Y_i's. Then,

$$\begin{aligned} E\left(X_{(n)}\right) &= 2\theta_2 E\left(Y_{(n)}\right) + \theta_1 - \theta_2 \\ &= 2\theta_2 n \int_0^1 y^n dy + \theta_1 - \theta_2 \\ &= \frac{2\theta_2 n}{n+1} + \theta_1 - \theta_2 \end{aligned}$$

and

$$\begin{aligned} E\left(X_{(1)}\right) &= 2\theta_2 E\left(Y_{(1)}\right) + \theta_1 - \theta_2 \\ &= 2\theta_2 n \int_0^1 y(1-y)^{n-1} dy + \theta_1 - \theta_2 \\ &= -\frac{2\theta_2 n}{n+1} + \theta_1 + \theta_2 \end{aligned}$$

Hence, $E\left(X_{(n)} + X_{(1)}\right)/2 = \theta_1$ and $E\left(X_{(n)} - X_{(1)}\right) = 2\theta_2(n-1)/(n+1)$. Therefore, the UMVUE's of θ_1 and θ_2 are, respectively, $\left(X_{(n)} + X_{(1)}\right)/2$ and $(n+1)\left(X_{(n)} + X_{(1)}\right)/[2(n-1)]$.

5. All parts can be solved by this general method. Suppose $X \sim f(x;\theta) = c(\theta)m(x)$, $\theta < x < b$. Then $1/c(\theta) = \int_\theta^b m(x)dx$, and the cdf of X is $F(x) = c(\theta)/c(x)$, $\theta < x < b$. Let $Y = X_{(1)}$ be the smallest order statistic. Note that Y is a complete sufficient statistic. Thus, by Lehmann and Scheffé's Theorem, any function $T(Y)$ that is an unbiased estimator of $h(\theta)$ is the best unbiased estimator of $h(\theta)$. The pdf of Y is $g(y;\theta) = nc(\theta)^n m(y)/c(y)^{n-1}$, $\theta < y < b$. Now consider the equations

$$\int_\theta^b f(x;\theta)dx = 1 \quad \text{and} \quad \int_\theta^b T(y)g(y;\theta)dy = h(\theta)$$

which are equivalent to

$$\int_\theta^b m(x)dx = \frac{1}{c(\theta)} \quad \text{and} \quad n\int_\theta^b T(y)m(y)/c(y)^{n-1}dy = h(\theta)/c(\theta)^n$$

Differentiating both sides of these two equations with respect to θ and using the Fundamental Theorem of Calculus yields

$$m(\theta) = \frac{c'(\theta)}{c(\theta)^2} \quad \text{and} \quad -\frac{nT(\theta)m(\theta)}{c(\theta)^{n-1}} = \frac{c(\theta)^n h'(\theta) - h(\theta)nc(\theta)^{n-1}c'(\theta)}{c(\theta)^{2n}}.$$

Change θs to ys and solve these two equations for $T(y)$ to get the best unbiased estimator of $h(\theta)$ is

$$T(y) = h(y) - \frac{h'(y)}{nm(y)c(y)}.$$

Note that for $h(\theta) = \theta^r$, $h'(\theta) = r\theta^{r-1}$.
(a) For this pdf, $m(x) = e^{-x}$ and $c(\theta) = e^\theta$. Hence

$$T(y) = y^r - \frac{ry^{r-1}}{ne^{-y}e^y} = y^r - \frac{ry^{r-1}}{n}.$$

(b) For this pdf, $m(x) = e^{-x}$ and $c(\theta) = 1/\left(e^{-\theta} - e^{-b}\right)$. Hence

$$T(y) = y^r - \frac{ry^{r-1}}{ne^{-y}}\left(e^{-y} - e^{-b}\right) = y^r - \frac{ry^{r-1}\left(1 - e^{-(b-y)}\right)}{n}.$$