STAT 611 Homework 4 Solutions

1. (a) The statistic T is a Bernoulli random variable so

$$\mathbb{E}_p[T] = \mathbb{P}_p(T=1) = \mathbb{P}_p\left(\sum_{i=1}^{n} X_i > X_{n+1}\right) = h(p)$$

(b) Recall that $S(\mathbf{X}) = \sum_{i=1}^{n+1} X_i$ is a complete sufficient statistic for θ. Conditioning T on S produces the best unbiased estimator of $h(p)$:

$$\begin{aligned} g(T) &= \mathbb{E}\left(T \,\middle|\, \sum_{i=1}^{n+1} X_i = \xi\right) = \mathbb{P}\left(\sum_{i=1}^{n} X_i > X_{n+1} \,\middle|\, \sum_{i=1}^{n+1} X_i = \xi\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{n} X_i > X_{n+1},\ \sum_{i=1}^{n+1} X_i = \xi\right)\left[\mathbb{P}\left(\sum_{i=1}^{n+1} X_i = \xi\right)\right]^{-1} \end{aligned}$$

First, we have that

$$\left(\sum_{i=1}^{n+1} X_i = \xi\right) = \binom{n+1}{\xi} p^{\xi}(1-p)^{n+1-\xi}$$

Case 1: When $\xi = 0$, the numerator is zero since the sum of no collection of Bernoulli random variables can strictly exceed zero, that is,

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

Case 2: When $\xi > 0$, the numerator becomes

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} X_i > X_{n+1},\ \sum_{i=1}^{n+1} X_i = \xi\right) &= \mathbb{P}\left(\sum_{i=1}^{n} X_i > X_{n+1},\ \sum_{i=1}^{n} X_i = \xi, X_{n+1} = 0\right) \\ &\quad + \mathbb{P}\left(\sum_{i=1}^{n} X_i > X_{n+1},\ \sum_{i=1}^{n} X_i = \xi, X_{n+1} = 1\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{n} X_i > 0,\ \sum_{i=1}^{n} X_i = \xi\right)\mathbb{P}(X_{n+1} = 0) \\ &\quad + \mathbb{P}\left(\sum_{i=1}^{n} X_i > 1,\ \sum_{i=1}^{n} X_i = \xi - 1\right)\mathbb{P}(X_{n+1} = 1) \end{aligned}$$

Now, for all $\xi > 0$, we have that

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

But if $\xi = 1$ or $\xi = 2$, notice that

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

and when $\xi > 2$,

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

Given all this, we can write down the full form of the UMVUE of $h(p)$:

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

2. (a) Since we are only given one observation, the likelihood function is

$$L(\theta|X) = f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|} \tag{1}$$

and so the score function is

$$l(\theta|X) := \log L(\theta|X) = |x| \log\left(\frac{\theta}{2}\right) + (1 - |x|)\log(1-\theta)$$

which has first partial derivative w.r.t. θ of

$$\frac{\partial}{\partial\theta} l(\theta|X) = \frac{|x|}{\theta} - \frac{1-|x|}{1-\theta}$$

Equating this to zero, we find that the MLE of θ is $\hat{\theta} = |x|$. Since the next higher partial derivative w.r.t. θ is

$$\frac{\partial^2}{\partial\theta^2} l(\theta|X) = -\frac{\theta^2 - 2|x|\theta + |x|}{(\theta-1)^2\theta^2}$$

which is negative for all $\theta \in [0,1]$, the MLE above is the unique solution.

(b) The expectation of T is

$$\mathbb{E}[T] = 2\mathbb{P}(X = 1) = 2\left(\frac{\theta}{2}\right) = \theta$$

Hence, T is unbiased.

(c) Notice that the density can be written as

$$\begin{aligned} f(x|\theta) &= \left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|} \\ &= (1-\theta)\exp\left(|x|\log\left(\frac{\theta}{2(1-\theta)}\right)\right) \end{aligned}$$

Thus, f is a one-parameter exponential family meaning that $S(X) = |X|$ is a sufficient statistic for θ. So by the Rao-Blackwell Theorem, since T is unbiased, the estimator

$$T^*(X) = \mathbb{E}[T(X) \mid S(X)] = \mathbb{E}[T(X) \mid |X|]$$

is a better estimator than T. In fact,

$$\mathbb{E}[T(X) \mid |X|] = 2 \times \frac{1}{2}|X| = |X|$$

(d) Now assume that $X_1, \ldots, X_n$ are iid according to the density in (1). Then the statistic $S(X) = \sum_{i=1}^n |X_i|$ is a complete sufficient statistic for θ meaning that

$$\mathbb{E}[T(X) \mid S(X)] = \mathbb{E}\left[T(X) \;\middle|\; \sum_{i=1}^n |X_i|\right]$$

is the best unbiased estimator. But since

$$\mathbb{P}\left[X_1 = 1 \;\middle|\; \sum_{i=1}^n |X_i|\right] = \frac{k}{n},$$

this simplifies to

$$\mathbb{E}\left[T(X) \;\middle|\; \sum_{i=1}^n |X_i|\right] = \frac{1}{n}\sum_{i=1}^n |X_i|$$

3. (a) We are interested in estimating $\tau(\theta) = 1/\theta$. Consider the estimator

$$W(X,Y) = \frac{X+Y}{2}$$

which has expectation $1/\theta$ meaning that it is unbiased. The joint pdf of X and Y (which satisfies the assumptions of the Cramér-Rao theorem) is

$$f(x,y|\theta) = \theta^2 e^{-\theta(x+y)}$$

which has score function

$$\frac{\partial}{\partial\theta}\log f(x,y|\theta) = \frac{2}{\theta} - (x+y) = \underbrace{-2}_{a(\theta)}\left(\underbrace{\frac{x+y}{2}}_{W(X)} - \underbrace{\frac{1}{\theta}}_{\tau(\theta)}\right)$$

By the attainment theorem, $W(X)$ is the UMVUE for $\tau(\theta)$.

(b) The MSE of W (remember that W is unbiased) is

$$MSE(W(X,Y)) = \text{Var}\left(\frac{X+Y}{2}\right) = \frac{1}{2\theta^2}$$

The MSE of the geometric mean $\sqrt{XY}$ is

$$\begin{aligned}
MSE(\sqrt{XY}) &= \mathbb{E}\left[\left(\sqrt{XY} - \frac{1}{\theta}\right)\right] \\
&= \mathbb{E}\left[XY - \frac{2}{\theta}\sqrt{XY} + \frac{1}{\theta^2}\right] \\
&= \mathbb{E}[XY] - \frac{2}{\theta}\mathbb{E}[\sqrt{XY}] + \frac{1}{\theta^2} \\
&= \frac{2}{\theta^2} - \frac{2}{\theta}\mathbb{E}[\sqrt{X}]^2 \qquad \text{since } X \text{ and } Y \text{ are iid}
\end{aligned}$$

Hence, $MSE(W(X,Y)) \leq MSE(\sqrt{XY})$ if, and only if,

$$\mathbb{E}[\sqrt{X}] > \frac{1}{2}\sqrt{\frac{3}{\theta}}$$

But,

$$\mathbb{E}[\sqrt{X}] = \int \sqrt{x}\theta e^{\theta x} = \frac{1}{2}\sqrt{\frac{\pi}{\theta}} > \frac{1}{2}\sqrt{\frac{3}{\theta}}$$

for all $\theta \in [0, \infty)$. Thus, $MSE(\sqrt{XY}) < MSE(W(X,Y))$.