# HW4

Sam Olson

## Problem 1

**Problem 6.2, Casella and Berger (2nd Edition)**

**6.2** Let $X_1, \dots, X_n$ be independent random variables with densities

$$f_{X_i}(x|\theta) = \begin{cases} e^{\theta - x} & x \geq i\theta \\ 0 & x < i\theta. \end{cases}$$

Prove that $T = \min_i(X_i/i)$ is a sufficient statistic for $\theta$.

### Answer

Start by noting the Factorization Thm.: a statistic $T(X)$ is sufficient for $\theta$ if the joint pdf can be expressed in the form:

$$f(x_1, \dots, x_n|\theta) = g(T(X), \theta)h(x_1, \dots, x_n),$$

where $g(T(X), \theta)$ is a function depending on $\theta$ and the data only through $T(X)$, and $h(x_1, \dots, x_n)$ is a function that does not depend on $\theta$.

We are given that $X_1, \dots, X_n$ are iid.

$$f_{X_i}(x|\theta) = \begin{cases} e^{\theta - x} & x \geq i\theta, \\ 0 & x < i\theta \end{cases}$$

Making the joint pdf of $X_1, \dots, X_n$:

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} f_{X_i}(x_i|\theta) = \prod_{i=1}^{n} e^{\theta - x_i} \cdot I_{[i\theta, +\infty)}(x_i)$$

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} e^{\theta - x_i} \cdot I_{[i\theta, +\infty)}(x_i)$$

So we have two products to consider. The first:

$$\prod_{i=1}^{n} e^{\theta - x_i} = e^{n\theta - \sum_{i=1}^{n} x_i}$$

And for the second:

$$\prod_{i=1}^{n} I_{[i\theta,+\infty)}(x_i) = I_{[\theta,+\infty)}\left(\min_i (x_i/i)\right)$$

Noting that the condition $x_i \geq i\theta$ for all $i$ is equivalent to $\min_i(x_i/i) \geq \theta$.

Taken together, the joint pdf is:

$$f(x_1, \ldots, x_n|\theta) = e^{n\theta} \cdot I_{[\theta,+\infty)}\left(\min_i (x_i/i)\right) \cdot e^{-\sum_{i=1}^{n} x_i}$$

Let $T(X) = \min_i(X_i/i)$, such that we have:

$$f(x_1, \ldots, x_n|\theta) = e^{n\theta} \cdot I_{[\theta,+\infty)}(T(X)) \cdot e^{-\sum_{i=1}^{n} x_i}$$

Where:

$$g(T(X), \theta) = e^{n\theta} \cdot I_{[\theta,+\infty)}(T(X))$$

And

$$h(x_1, \ldots, x_n) = e^{-\sum_{i=1}^{n} x_i}$$

So we've effectively met our condition required by the Factorization Thm., i.e. one factor $g(T(X), \theta)$ depends on $\theta$ only through $T(X)$, and $h(x_1, \ldots, x_n)$ is independent of $\theta$, so $T(X) = \min_i(X_i/i)$ is a sufficient statistic for $\theta$.

# Problem 2

**Example of Rao-Blackwell theorem, which is largely a STAT 5420 problem in computation.**

Let $X_1$ and $X_2$ be iid Bernoulli($p$), $0 < p < 1$.

## a)

Show $S = X_1 + X_2$ is Sufficient for p

### Answer

By the Factorization Theorem, a statistic $S$ is sufficient for $p$ if the joint pmf can be written in the form:

$$f(x_1, x_2|p) = g(S, p) \cdot h(x_1, x_2)$$

, i.e. as the product of two functions, one of which is not dependent upon the parameters of interest, p.

The joint pmf of $X_1, X_2$, noting the two random variables are iid Bernoulli($p$), is:

$$f(x_1, x_2|p) = p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} = p^{x_1+x_2}(1-p)^{2-(x_1+x_2)}$$

Let $S = X_1 + X_2$, and rewrite the above:

$$f(x_1, x_2|p) = p^S(1-p)^{2-S}$$

Since this is of the form $g(S, p) \cdot h(x_1, x_2)$ with $h(x_1, x_2) = 1$, it follows that $S$ is sufficient for $p$ by the Factorization Thm.

## b)

Identify the conditional probability $P(X_1 = x|S = s)$; you should know which values of $x, s$ to consider.

### Answer

We compute:

$$P(X_1 = x|S = s) = \frac{P(X_1 = x, S = s)}{P(S = s)}$$

Generally speaking, we know the range of possible values of S, that is $S \in [0, 2]$.

Thus, for possible values of $S$, consider the cases:

(0): If $S = 0$, then $X_1 = 0$ and $X_2 = 0$, so:

$$P(X_1 = 0|S = 0) = 1$$

(1): If $S = 2$, then $X_1 = 1$ and $X_2 = 1$, so:

$$P(X_1 = 1|S = 2) = 1$$

(2): If $S = 1$, then either:

$X_1 = 0, X_2 = 1$, or $X_1 = 1, X_2 = 0$, both events being equally likeliy (equal probability),

$$P(X_1 = 1|S = 1) = P(X_1 = 0|S = 1) = \frac{1}{2}$$

Taking points (0) through (2) together gives us:

$$P(X_1 = x|S = s) = \begin{cases} 1 & \text{if } s = 0 \text{ and } x = 0, \text{ or if } s = 2 \text{ and } x = 1, \\ \frac{1}{2} & \text{if } s = 1 \text{ and } x \in \{0, 1\}, \\ 0 & \text{otherwise} \end{cases}$$

## c)

Find the conditional expectation $T \equiv E(X_1|S)$, i.e., as a function of the possibilities of $S$. Note that $T$ is a statistic.

**Answer**

Using the values from part (b):

$$T = E(X_1|S) = \begin{cases} 0 & S = 0, \\ \frac{1}{2} & S = 1, \\ 1 & S = 2 \end{cases}$$

$T$ is a statistic, noted.

## d)

Show $X_1$ and $T$ are both unbiased for $p$.

**Answer**

For $X_1$:

$$E_p(X_1) = p$$

Noting the distributional properties of $X_1 \sim$ Bernoulli(p).

For $T$, noting properties of expectation:

$$E_p(T) = \sum_{s=0}^{2} E(X_1|S = s)P(S = s)$$

Substituting:

$$E_p(T) = 0 \cdot (1 - p)^2 + \frac{1}{2} \cdot 2p(1 - p) + 1 \cdot p^2 = p(1 - p) + p^2 = p$$

Thus, both $X_1$ and $T$ are unbiased estimators of $p$.

## e)

Show $\text{Var}_p(T) \leq \text{Var}_p(X_1)$, for any $p$.

**Answer**

By invoking the Rao-Blackwell Thm., we know:

$$\text{Var}_p(T) \leq \text{Var}_p(X_1)$$

Alternatively, consider that since $X_1 \sim \text{Bernoulli}(p)$, we know its variance is given by:

$$\text{Var}_p(X_1) = p(1-p)$$

For $T$:

$$\text{Var}_p(T) = E_p(T^2) - (E_p(T))^2$$

We may then solve for $E_p(T^2)$:

$$E_p(T^2) = 0^2 \cdot (1-p)^2 + \left(\frac{1}{2}\right)^2 \cdot 2p(1-p) + 1^2 \cdot p^2 = \frac{p(1-p)}{2} + p^2$$

Thus,

$$\text{Var}_p(T) = \left(\frac{p(1-p)}{2} + p^2\right) - p^2 = \frac{p(1-p)}{2}$$

Since

$$\frac{p(1-p)}{2} \leq p(1-p)$$

it follows that:

$$\text{Var}_p(T) \leq \text{Var}_p(X_1)$$

as expected from Rao-Blackwell.

# Problem 3

**Problem 6.21 a)-b), Casella and Berger (2nd Edition)**

**6.21** Let $X$ be one observation from the pdf

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|}, \quad x = -1, 0, 1, \quad 0 \le \theta \le 1.$$

## a)

Is $X$ a complete sufficient statistic?

### Answer

Since $X$ is the only observation, it is sufficient for $\theta$ as it is the entirety of the data (all the information).

To determine whether $X$ is complete, we then need to check whether the only function $g(X)$ satisfying $E[g(X)] = 0$ for all $\theta$ is the zero function.

To that end note:

$$E[g(X)] = \sum_{x \in \{-1,0,1\}} g(x) f(x|\theta)$$

Using the given density:

$$E[g(X)] = g(-1)P(X=-1) + g(0)P(X=0) + g(1)P(X=1) == \frac{\theta}{2}g(-1) + (1-\theta)g(0) + \frac{\theta}{2}g(1)$$

Noting the Law of Total Probability.

Since this must be zero for all $\theta \in [0,1]$, we then have:

$$\theta\left(\frac{g(-1)+g(1)}{2} - g(0)\right) + g(0) = 0$$

However, for this to be true for all $\theta$, both coefficients must be zero:

$$\frac{g(-1)+g(1)}{2} - g(0) = 0 \rightarrow g(0) = 0$$

Using $g(0) = 0$, the first equation gives us:

$$\frac{g(-1)+g(1)}{2} = 0 \rightarrow g(-1) + g(1) = 0$$

So $X$ is not complete, as we have identified a function that is not the zero function such that $g(-1) = 1, g(1) = -1, g(0) = 0$.

## b)

Is $|X|$ a complete sufficient statistic?

**Answer**

Note again the Factorization Thm. for determining sufficiency.

To that end, the pdf is:

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

As defined, the above pdf depends on $X$ only through $|X|$, so the conditional distribution of $X$ given $|X|$ does not depend on $\theta$. So $|X|$ is sufficient via the Factorization Thm. Another way to say this is that we have two functions, one which entirely depends on $\theta$ and one that does not (in this case, the 1 function), i.e. $f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|} \cdot 1$.

Next, we check completeness, using the same criteria used in part a).

Again, note the conditional pdf of $|X|$ given above, and that $|X|$ is always positive by construction. Taken together, for the purposes of determining the underlying pmf, we have:

$$P(|X| = 0) = 1 - \theta, \text{ and } P(|X| = 1) = \theta$$

This is the pmf of a Bernoulli with $p = \theta$! Given this result, we may then note that the Bernoulli family is complete, meaning we cannot find a function that is not the zero function satisfying $E[g(X)] = 0$ for some function g. And as $|X|$ is Bernoulli distributed, it is a complete sufficient statistic.

Note: Part of the completeness argument is based on assumption that we know the Binomial family is a complete family of distributions, and Bernoulli being a Binomial distribution with n=1 (a specific instance of a Binomial).

# Problem 4

**Problem 6.24, Casella and Berger (2nd Edition)**

**6.24** Consider the following family of distributions:

$$\mathcal{P} = \{P_\lambda(X = x) : P_\lambda(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}; x = 0, 1, 2, \ldots; \lambda = 0 \text{ or } 1\}$$

This is a Poisson family with $\lambda$ restricted to be 0 or 1. Show that the family $\mathcal{P}$ is not complete, demonstrating that completeness can be dependent on the range of the parameter. (See Exercises 6.15 and 6.18.)

## Answer

To show that $\mathcal{P}$ is not complete, we must find a nonzero function $h(X)$ such that:

$$E_\lambda[h(X)] = 0, \quad \text{for all } \lambda \in \{0, 1\}$$

By definition, a family of distributions is complete if the only function satisfying this expectation condition is the zero function.

As given, we only consider values for which $\lambda = 0, 1$.

For $\lambda = 0$, the Poisson distribution degenerates to:

$$P_{\lambda=0}(X = x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \neq 0 \end{cases}$$

So it's expectation ios:

$$E_{\lambda=0}[h(X)] = h(0) \text{ so, for } E_{\lambda=0}[h(X)] = 0 \to h(0) = 0$$

Then, $\lambda = 1$, $X \sim \text{Poisson}(1)$, giving expectation:

$$E_{\lambda=1}[h(X)] = \sum_{x=0}^{\infty} h(x) \cdot \frac{1^x e^{-1}}{x!} = e^{-1} \sum_{x=0}^{\infty} \frac{h(x)}{x!}$$

As noted previously, for $h(0) = 0$, this simplifies to:

$$\sum_{x=1}^{\infty} \frac{h(x)}{x!} = 0$$

Taken together, we must have a function $h(X)$ that satisfies:

$$\sum_{x=1}^{\infty} \frac{h(x)}{x!} = 0, \quad h(0) = 0$$

A simple choice is:

$$h(0) = 0, \quad h(1) = 1, \quad h(2) = -2, \quad h(x) = 0 \text{ for } x \geq 3$$

Computing the sum:

$$\sum_{x=1}^{\infty} \frac{h(x)}{x!} = \frac{h(1)}{1!} + \frac{h(2)}{2!} + \sum_{x=3}^{\infty} \frac{h(x)}{x!} = \frac{1}{1} + \frac{-2}{2} + 0 = 1 - 1 = 0$$

Thus, $E_\lambda[h(X)] = 0$ for both $\lambda = 0$ and $\lambda = 1$, yet $h(X)$ is not the zero function! This is proof that the family $\mathcal{P}$ as defined is not complete (illustrating that completeness can be dependent on the range of the parameter).

# Problem 5

**Problem 7.57, Casella and Berger (2nd Edition)** You may assume $n \geq 3$.

One has to Rao-Blackwellize on the complete/sufficient statistic here

$$\sum_{i=1}^{n+1} X_i.$$

**7.57** Let $X_1, \ldots, X_{n+1}$ be iid Bernoulli$(p)$, and define the function $h(p)$ by

$$h(p) = P\left(\sum_{i=1}^{n} X_i > X_{n+1} \middle| p\right),$$

the probability that the first $n$ observations exceed the $(n+1)$st.

## a)

Show that

$$T(X_1, \ldots, X_{n+1}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} X_i > X_{n+1}, \\ 0 & \text{otherwise,} \end{cases}$$

is an unbiased estimator of $h(p)$.

### Answer

For $T(X_1, \ldots, X_{n+1})$, as given, we must check unbiasedness by showing it's expectation is equal to $h(p)$.

With $T$ as an indicator function of the event $\sum_{i=1}^{n} X_i > X_{n+1}$, and $h(p) = P\left(\sum_{i=1}^{n} X_i > X_{n+1} | p\right)$, we have:

$$E_p[T] = P_p(T = 1) = P_p\left(\sum_{i=1}^{n} X_i > X_{n+1}\right) = h(p)$$

Thus, $T(X)$ is an unbiased estimator of $h(p)$.

## b)

Find the best unbiased estimator of $h(p)$.

### Answer

Since $\sum_{i=1}^{n+1} X_i$ is a complete sufficient statistic for $p$, we can use Rao-Blackwell (More Lehmann–Scheffé given the complete sufficient statistic), specifically by finding the conditional expectation of $T(X)$ (estimator of $h(p)$) from part a) conditioned on a complete and sufficient statistic to find the UMVUE. So that's the "plan".

The idea here is our best unbiased estimator of $h(p)$ is of the form:

$$T^*(X) = E[T(X)|S = \sum_{i=1}^{n+1} X_i]$$

With the goal of calculating $T^*(X)$.

To that end, as given from part a), $T(X)$ is defined as:

$$E\left[T \mid \sum_{i=1}^{n+1} X_i = y\right] = P\left(\sum_{i=1}^{n} X_i > X_{n+1} \middle| \sum_{i=1}^{n+1} X_i = y\right)$$

As $X_{n+1}$ is binary, there are two cases to check for to then invoke the Law of Total Probability. These are:

**(1)** $\quad X_{n+1} = 0$

If $X_{n+1} = 0$, then $\sum_{i=1}^{n} X_i = y - 0 = y$. Since $y \geq 1$, the event $\sum_{i=1}^{n} X_i > X_{n+1}$ always holds:

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

**(2)** $\quad X_{n+1} = 1$

If $X_{n+1} = 1$, then $\sum_{i=1}^{n} X_i = y - 1$, so $\sum_{i=1}^{n} X_i > X_{n+1}$ only holds when $y - 1 \geq 1$, i.e., when $y \geq 2$:

$$P\left(\sum_{i=1}^{n} X_i > X_{n+1} \mid \sum_{i=1}^{n+1} X_i = y, X_{n+1} = 1\right) = I_{y\geq 2}.$$

To combine cases (1) and (2), we note that $X_{n+1} \sim \text{Bernoulli}(p)$, such that the probability of both cases is:

$$P(X_{n+1} = 1 \mid \sum_{i=1}^{n+1} X_i = y) = \frac{y}{n+1}$$

And

$$P(X_{n+1} = 0 \mid \sum_{i=1}^{n+1} X_i = y) = \frac{n+1-y}{n+1}$$

Then, invoking the Law of Total Probability:

$$P\left(\sum_{i=1}^{n} X_i > X_{n+1} \mid \sum_{i=1}^{n+1} X_i = y\right) = \left(1 \cdot \frac{n+1-y}{n+1}\right) + \left(I_{y\geq 2} \cdot \frac{y}{n+1}\right)$$

Using the above formula, we take expectation:

$$E\left[T \mid \sum_{i=1}^{n+1} X_i = y\right] = \begin{cases} 0 & y = 0, \\ \frac{\binom{n}{y}}{\binom{n+1}{y}} & y = 1 \text{ or } 2, \\ 1 & y > 2 \end{cases}$$

Simplifying:

$$T^*(X) = \begin{cases} 0 & y = 0, \\ \frac{\binom{n}{y}}{\binom{n+1}{y}} & y = 1 \text{ or } 2, \\ 1 & y > 2 \end{cases} = \begin{cases} 0 & y = 0, \\ \frac{n}{n+1} & y = 1, \\ \frac{n-1}{n+1} & y = 2, \\ 1 & y > 2 \end{cases}$$

**Some Additional Algebra For Justifying the Above Cases**

**$\mathbf{y = 0}$**

For y $= 0$, $X_i = 0 \quad \forall i$, so $\sum_{i=1}^n X_i = 0$, and $\sum_{i=1}^n X_i > X_{n+1}$ has probability zero (does not occur). So we have:

$$E\left[T \mid \sum_{i=1}^{n+1} X_i = 0\right] = 0$$

**$\mathbf{y = 1}$**

For y $= 1$, $X_{n+1} = 0$, so we have:

$$P(\sum_{i=1}^{n} X_i = 1 \mid \sum_{i=1}^{n+1} X_i = 1) = \frac{\binom{n}{1} p(1-p)^{n-1}(1-p)}{\binom{n+1}{1} p(1-p)^n} = \frac{\binom{n}{1}}{\binom{n+1}{1}} = \frac{n}{n+1}$$

**$\mathbf{y = 2}$**

For y $= 2$:

$$P(\sum_{i=1}^{n} X_i = 2 \mid \sum_{i=1}^{n+1} X_i = 2) = \frac{\binom{n}{2} p^2(1-p)^{n-2}(1-p)}{\binom{n+1}{2} p^2(1-p)^{n-1}} = \frac{\binom{n}{2}}{\binom{n+1}{2}} = \frac{n-1}{n+1}$$

**$\mathbf{y > 2}$**

For $y > 2$:

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