# Stat 346/446: Theoretical Statistics II: Practice Exercises 3 Solutions

## Textbook Exercises

**6.3** (346 & 446)Let $X_1, \dots, X_n$ be a random sample from the pdf

$$f(x|\mu, \sigma) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, \quad \mu < x < \infty, \quad 0 < \sigma < \infty.$$

Find a two-dimensional sufficient statistic for $(\mu, \sigma)$.
Let $x_{(1)} = \min_i x_i$. Then the joint pdf is

$$f(x_1, \dots, x_n|\mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma} e^{-(x_i-\mu)/\sigma} I_{(\mu,\infty)}(x_i)$$

$$= \left(\frac{e^{\mu/\sigma}}{\sigma}\right)^n e^{-\sum x_i/\sigma} I_{(\mu,\infty)}(x_{(1)}) \cdot 1.$$

We identify

$$g(x_{(1)}, \sum_i x_i|\mu, \sigma) = \left(\frac{e^{\mu/\sigma}}{\sigma}\right)^n e^{-\sum_i x_i/\sigma} I_{(\mu,\infty)}(x_{(1)}).$$

$$h(x) = 1$$

Thus, by the Factorization Theorem, $(X_{(1)}, \sum_i X_i)$ is a sufficient statistic for $(\mu, \sigma)$.

**6.17** (346 & 446) Let $X_1, \dots, X_n$ be iid with geometric distribution

$$P_\theta(X = x) = \theta(1-\theta)^{x-1}, \quad x = 1, 2, \dots, \quad 0 < \theta < 1.$$

Show that $\sum X_i$ is sufficient for $\theta$, and find the family of distributions of $\sum X_i$. Is the family complete?
The population pmf is given by:

$$f(x|\theta) = \theta(1-\theta)^{x-1} = \frac{\theta}{1-\theta} e^{\log(1-\theta)x},$$

which is an exponential family with $t(x) = x$. By Theorem 6.2.10, a pmf of the form

$$f(x|\theta) = h(x)c(\theta) \exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right)$$

belongs to the exponential family, where $\theta = (\theta_1, \theta_2, \dots, \theta_d)$ with $d \le k$. Then, the statistic

$$T(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \dots, \sum_{j=1}^{n} t_k(X_j)\right)$$

is a sufficient statistic for $\theta$. In this case, $t(x) = x$, so the sufficient statistic is:

$$T(\mathbf{X}) = \sum_{i=1}^{n} X_i.$$

By Theorem 6.2.25, in an exponential family of the form:

$$f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{j=1}^{k} w(\theta_j)t_j(x)\right),$$

the statistic

$$T(\mathbf{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \sum_{i=1}^{n} t_2(X_i), \dots, \sum_{i=1}^{n} t_k(X_i)\right)$$

is complete as long as the parameter space $\Theta$ contains an open set in $\mathbb{R}^k$. Since our pmf belongs to the exponential family and the parameter space satisfies the necessary conditions, $\sum_{i=1}^{n} X_i$ is complete. Since $X_i \sim \text{Geometric}(\theta)$, we have:

$$\sum_{i=1}^{n} X_i - n \sim \text{Negative Binomial}(n, \theta).$$

**6.20b** (346 & 446)For each of the following pdfs let $X_1, \dots, X_n$ be iid observations. Find a complete sufficient statistic, or show that one does not exist.

(b) $f(x|\theta) = \frac{\theta}{(1+x)^{1+\theta}}, \quad 0 < x < \infty, \quad \theta > 0$

We are given that $X_1, \dots, X_n$ are iid observations from the pdf:

$$f(x|\theta) = \frac{\theta}{(1+x)^{1+\theta}}, \quad 0 < x < \infty, \quad \theta > 0.$$

The pdf can be rewritten as:

$$f(x|\theta) = \exp\left(\log\theta - (1+\theta)\log(1+x)\right).$$

Comparing with the general exponential family form:

$$f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{j=1}^{k} w(\theta_j)t_j(x)\right),$$

we identify:

- $t(x) = \log(1+x)$,
- $w(\theta) = -(1+\theta)$,
- $c(\theta) = e^{\log\theta} = \theta$,
- $h(x) = 1$.

Thus, the given pdf belongs to the exponential family. By **Theorem 6.2.10**, the statistic:

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

is a sufficient statistic for $\theta$. By **Theorem 6.2.25**, since $\sum_{i=1}^{n} t(X_i)$ is the natural sufficient statistic in an exponential family and the parameter space $\Theta = (0, \infty)$ contains an open set in $\mathbb{R}$, it is also complete. Thus, the complete and sufficient statistic for $\theta$ is:

$$T(\mathbf{X}) = \sum_{i=1}^{n} \log(1+X_i).$$

**7.59** (446) Let $X_1, \ldots, X_n$ be iid $n(\mu, \sigma^2)$. Find the best unbiased estimator of $\sigma^p$, where $p$ is a known positive constant, not necessarily an integer.
We know $T = (n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Then

$$E[T^{p/2}] = \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} \int_0^\infty t^{\frac{p+n-1}{2}-1} e^{-t/2}\, dt$$

$$= \frac{2^{p/2}\Gamma\left(\frac{p+n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = C_{p,n}.$$

Thus,

$$E\left(\frac{(n-1)S^2}{\sigma^2}\right)^{p/2} = C_{p,n},$$

so

$$\frac{(n-1)^{p/2}S^p}{C_{p,n}}$$

is an unbiased estimator of $\sigma^p$. From Theorem 6.2.25, $(\bar{X}, S^2)$ is a complete, sufficient statistic. The unbiased estimator

$$\frac{(n-1)^{p/2}S^p}{C_{p,n}}$$

is a function of $(\bar{X}, S^2)$. Hence, it is the best unbiased estimator.

**7.60** (446) Let $X_1, \ldots, X_n$ be iid gamma$(\alpha, \beta)$ with $\alpha$ known. Find the best unbiased estimator of $1/\beta$.
Let $X_1, \ldots, X_n$ be iid Gamma$(\alpha, \beta)$ with $\alpha$ known. For a gamma-distributed random variable $X_i \sim$ Gamma$(\alpha, \beta)$, the mean and variance are:

$$E[X_i] = \frac{\alpha}{\beta}, \quad \mathrm{Var}(X_i) = \frac{\alpha}{\beta^2}.$$

Since the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is an unbiased estimator of $E[X_i]$, we get:

$$E[\bar{X}] = \frac{\alpha}{\beta}.$$

Rearranging,

$$E\left[\frac{1}{\alpha}\bar{X}\right] = \frac{1}{\beta}.$$

Thus, the unbiased estimator of $1/\beta$ is:

$$\hat{\theta} = \frac{\bar{X}}{\alpha} = \frac{1}{n\alpha}\sum_{i=1}^n X_i.$$

- The sufficient statistic for $\beta$ is $\sum X_i$ (since the sum of gamma variables follows another gamma distribution).
- By Theorem **6.2.25** (Exponential Family Completeness), $\sum X_i$ is also **complete**.
- Since our estimator $\hat{\theta}$ is a function of this complete sufficient statistic, it is the **best unbiased estimator (UMVUE)** by the Lehmann-Scheffé theorem.

$$\hat{\theta} = \frac{\bar{X}}{\alpha} = \frac{1}{n\alpha}\sum_{i=1}^n X_i$$

is the best unbiased estimator of $1/\beta$.

**10.1** (346 & 446) A random sample $X_1, \dots, X_n$ is drawn from a population with pdf

$$f(x|\theta) = \frac{1}{2}(1 + \theta x), \quad -1 < x < 1, \quad -1 < \theta < 1.$$

Find a consistent estimator of $\theta$ and show that it is consistent.
First, note that

$$E(X) = \int_{-1}^{1} x \frac{1}{2}(1 + \theta x)\, dx.$$

$$E(X) = \frac{1}{2}\left[\frac{x^2}{2} + \frac{\theta x^3}{3}\right]_{-1}^{1}.$$

$$= \frac{1}{2}\left(\frac{1}{2} + \frac{\theta}{3} - \frac{1}{2} - \frac{\theta}{3}\right) = \frac{\theta}{3}.$$

By WLLN, we have

$$\bar{X}_n \xrightarrow{P} \frac{\theta}{3} \quad \text{as } n \to \infty.$$

Since $h(x) = 3x$ is a continuous function, applying the continuous mapping theorem,

$$3\bar{X}_n \xrightarrow{P} \theta \quad \text{as } n \to \infty.$$

Thus, $3\bar{X}_n$ is a consistent estimator of $\theta$.

**OR:**
Alternatively,

$$E(\bar{X}_n) = \frac{\theta}{3} \quad \Rightarrow \quad E(3\bar{X}_n) = \theta.$$

Thus, the bias of $3\bar{X}_n$ is zero, meaning it is an unbiased estimator. We compute $E(X^2)$:

$$E(X^2) = \int_{-1}^{1} x^2 \frac{1}{2}(1 + \theta x)\, dx.$$

Evaluating,

$$E(X^2) = \frac{1}{2}\left[\frac{x^3}{3} + \frac{\theta x^4}{4}\right]_{-1}^{1}.$$

$$= \frac{1}{2}\left(\frac{1}{3} + \frac{\theta}{4} - \frac{1}{3} - \frac{\theta}{4}\right) = \frac{1}{3}.$$

Thus, the variance of $X$ is

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{3} - \frac{\theta^2}{9} = \frac{1}{3} - \frac{\theta^2}{9}.$$

For the sample mean,

$$\mathrm{Var}(3\bar{X}_n) = 9 \cdot \frac{\mathrm{Var}(X)}{n} = 9\left(\frac{1}{3} - \frac{\theta^2}{9}\right)\frac{1}{n}.$$

$$= \frac{9}{n}\left(\frac{1}{3} - \frac{\theta^2}{9}\right) \to 0 \quad \text{as } n \to \infty.$$

Thus, $3\bar{X}_n$ is a consistent estimator of $\theta$.