Introduction to Statistics - Homework #5

- Use Appendix C: Distribution tables (e.g. Z-table) from page 408 to 417 in the textbook if necessary.

Exercise 1

Suppose we obtain a random sample $X_1, X_2, \ldots, X_n$ from a population with probability density function

$$f(x) = \begin{cases} \frac{1}{3\theta}, & -\theta < x < 2\theta,\ (\theta > 0), \\ 0, & \text{otherwise.} \end{cases}$$

Answer the following questions.

(1) Find the constant k such that the estimator $k\bar{X}$ is an unbiased estimator for θ, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Solution:

$$E[X] = \int_{-\theta}^{2\theta} x \cdot \frac{1}{3\theta}\, dx = \frac{\theta}{2}, \qquad E[k\bar{X}] = \frac{k\theta}{2}.$$

Thus, for unbiasedness we require $k = 2$.

(2) For the value of k obtained in (1), show that $k\bar{X}$ is a consistent estimator for θ.

Solution:

$$E[X^2] = \int_{-\theta}^{2\theta} x^2 \cdot \frac{1}{3\theta}\, dx = \theta^2, \quad Var(X) = E[X^2] - (E[X])^2 = \theta^2 - \left(\frac{\theta}{2}\right)^2 = \frac{3}{4}\theta^2.$$

Since $X \sim U(-\theta, 2\theta)$,

$$Var(X) = \frac{(2\theta - (-\theta))^2}{12} = \frac{9\theta^2}{12} = \frac{3}{4}\theta^2,$$

consistent with the above result. Therefore,

$$Var(2\bar{X}) = \frac{4}{n}Var(X) = \frac{4}{n} \cdot \frac{3}{4}\theta^2 = \frac{3\theta^2}{n}.$$

As $n \to \infty$,

$$\lim_{n\to\infty} E[2\bar{X}] = \theta, \qquad \lim_{n\to\infty} Var(2\bar{X}) = 0.$$

Hence, $2\bar{X}$ is a consistent estimator for θ.

Exercise 2

Suppose $X_1, X_2, \ldots, X_n$ are obtained from a Bernoulli distribution $B(1, p)$. Answer the following questions.

(1) Consider the estimators

$$\hat{p}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{p}_2 = \frac{1}{2}(X_1 + X_2).$$

Show that $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators of p.

Solution: Since $X_i \sim B(1, p)$, we have $E[X_i] = p$ and $Var(X_i) = p(1-p)$ for each i. For $\hat{p}_1$:

$$E[\hat{p}_1] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} \cdot np = p.$$

Note that $\hat{p}_2 = \frac{1}{2}\sum_{i=1}^{2} X_i$ is exactly $\hat{p}_1$ with $n = 2$, so by the same formula,

$$E[\hat{p}_2] = p.$$

Hence both $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators.

(2) Between $\hat{p}_1$ and $\hat{p}_2$, determine which estimator is more efficient (assume $n > 2$).

Solution: From (1), both are unbiased. Since $X_1, \ldots, X_n$ are independent,

$$Var(\hat{p}_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} \cdot n\,p(1-p) = \frac{p(1-p)}{n}.$$

Note that $\hat{p}_2 = \frac{1}{2}\sum_{i=1}^{2} X_i$ is exactly $\hat{p}_1$ with $n = 2$, so by the same formula,

$$Var(\hat{p}_2) = \frac{p(1-p)}{2}.$$

If $n > 2$, then $Var(\hat{p}_1) < Var(\hat{p}_2)$. Therefore, $\hat{p}_1$ is more efficient than $\hat{p}_2$.

(3) Show that the estimator $\hat{p}_1$ is a consistent estimator for p.

Solution: From (2), $E[\hat{p}_1] = p$ and $Var(\hat{p}_1) = \frac{p(1-p)}{n}$. Therefore,

$$\lim_{n\to\infty} E[\hat{p}_1] = p, \qquad \lim_{n\to\infty} Var(\hat{p}_1) = \lim_{n\to\infty} \frac{p(1-p)}{n} = 0.$$

Therefore, $\hat{p}_1$ is a consistent estimator for p.