Problem Set 6
"Statistics: Samples and Estimators"

(6.1) [Sample versus population] You are trying to explain the concept of estimation to a friend who has never taken statistics.

a. "If an estimator is just a random number," your friend says, "then it does not necessarily equal the true population value. How can it tell us anything about the population?" How do you respond?
b. "If there are lots of different possible estimators for any particular population value," your friend says, "and they are all random, why does it even matter which estimator we use?" How do you respond?

(6.2) [Sample versus population] Which of the following is a random variable in the classical sense? Justify each answer.

a. The average income of all families in Des Moines, Iowa this year.
b. The average income of a sample of families chosen by taking every hundredth listing in the telephone book, beginning at a point chosen by throwing a dart at the first page.

(6.3) [Sample versus population] Which of the following is a random variable in the classical sense? Justify each answer.

a. The average age of persons living in Iowa whose Social Security number ends in the same two digits as the Dow Jones Industrial Average at tomorrow's stock market closing.
b. The average age of all persons living in Iowa.

(6.4) [Properties of estimators] Suppose we wish to estimate the mean μ of a certain population using just 5 random observations X_1 through X_5. Two estimators for μ are proposed, as follows.

$$\hat{\mu}_1 = \frac{1}{6}\sum_{i=1}^{5} X_i \quad \text{and} \quad \hat{\mu}_2 = \frac{1}{5}\sum_{i=1}^{5} X_i$$

Analyze the behavior of these estimators assuming that the true population mean is **12** and the true population variance is **30**.

a. Use the properties of expectation to compute the means of the estimators, that is, $E(\hat{\mu}_1)$ and $E(\hat{\mu}_2)$.
b. Compute the biases of the estimators, that is, $Bias(\hat{\mu}_1)$ and $Bias(\hat{\mu}_2)$. Is either estimator unbiased?
c. Use the properties of variance to compute the variances of the estimators, that is, $Var(\hat{\mu}_1)$ and $Var(\hat{\mu}_2)$. Which estimator has lower variance?
d. Compute the mean squared errors of the estimators, that is, $MSE(\hat{\mu}_1)$ and $MSE(\hat{\mu}_2)$. Which estimator has lower mean squared error?

(6.5) [Properties of estimators] Suppose we wish to estimate the mean income of persons in a population. Denote this true population mean μ and the true population variance σ_2. Since it is impractical to collect this information for the entire population, we have collected this information for n=10 persons selected at random. Denote their observed incomes X_1 through X_{10}. To estimate the mean income in the entire population, you propose the following estimator:

$$\hat{\mu}_1 = \frac{1}{10}\sum_{i=1}^{10} X_i$$

But your lazy research assistant proposes the following estimator:

$$\hat{\mu}_2 = \frac{1}{5}\left(X_2 + X_4 + X_6 + X_8 + X_{10}\right)$$

which is based only on every other observation. Analyze the behavior of these estimators assuming that the true population mean is **40** and the true population variance is **60**.

a. Use the properties of expectation to compute the means of the estimators, that is, $E(\hat{\mu}_1)$ and $E(\hat{\mu}_2)$.
b. Compute the biases of the estimators, that is, $Bias(\hat{\mu}_1)$ and $Bias(\hat{\mu}_2)$. Is either estimator unbiased?
c. Use the properties of variance to compute the variances of the estimators, that is, $Var(\hat{\mu}_1)$ and $Var(\hat{\mu}_2)$. Which estimator has lower variance?
d. Compute the mean squared errors of the estimators, that is, $MSE(\hat{\mu}_1)$ and $MSE(\hat{\mu}_2)$. Which estimator has lower mean squared error?

(6.6) [Properties of estimators] Suppose we have a random sample $X_1, X_2, \ldots, X_n$ from a population with unknown true mean μ and true variance σ^2. Consider the estimator $\hat{\mu}$ given by the following formula.

$$\hat{\mu} = \left(\frac{1}{n+3} \right) \sum_{i=1}^{n} X_i$$

a. Prove that $E(\hat{\mu}) = \left(\frac{n}{n+3} \right) \mu$.
b. Find a formula for $Bias(\hat{\mu})$.
c. Is $\hat{\mu}$ unbiased? Justify your answer.
d. Is $\hat{\mu}$ asymptotically unbiased? Justify your answer.
e. Prove that $Var(\hat{\mu}) = \frac{n\,\sigma^2}{(n+3)^2}$.
f. Find a formula for $MSE(\hat{\mu})$.
g. Is $\hat{\mu}$ consistent? Justify your answer.

(6.7) [Properties of estimators] Suppose we have a random sample of four observations X_1, X_2, X_3, X_4 from a population with unknown true mean μ and true variance σ^2. Consider the two estimators for μ given by the following formulas.

$$\hat{\mu}_1 = 0.25 \sum_{i=1}^{4} X_i \qquad \text{and} \qquad \hat{\mu}_2 = 0.1X_1 + 0.1X_2 + 0.4X_3 + 0.4X_4$$

a. Use the properties of expectation to compute the means of the estimators, that is, $E(\hat{\mu}_1)$ and $E(\hat{\mu}_2)$.
b. Compute the biases of the estimators, that is, $Bias(\hat{\mu}_1)$ and $Bias(\hat{\mu}_2)$. Is either estimator unbiased?
c. Use the properties of variance to compute the variances of the estimators, that is, $Var(\hat{\mu}_1)$ and $Var(\hat{\mu}_2)$. Which estimator has lower variance?
d. Compute the mean squared errors of the estimators, that is, $MSE(\hat{\mu}_1)$ and $MSE(\hat{\mu}_2)$. Which estimator has lower mean squared error?
e. Which estimator is better? Why?

[end of problem set]