Point Estimation

Definition

A point estimate of some population parameter θ is a single numerical value $\hat{\theta}$ of a statistic $\hat{\Theta}$. The statistic $\hat{\Theta}$ is called the point estimator.

Estimation problems occur frequently in engineering. We often need to estimate

- The mean μ of a single population
- The variance σ^2 (or standard deviation σ) of a single population
- The proportion p of items in a population that belong to a class of interest
- The difference in means of two populations, $\mu_1 - \mu_2$
- The difference in two population proportions, $p_1 - p_2$

Reasonable point estimates of these parameters are as follows:

- For μ, the estimate is $\hat{\mu} = \bar{x}$, the sample mean.
- For σ^2, the estimate is $\hat{\sigma}^2 = s^2$, the sample variance.
- For p, the estimate is $\hat{p} = x/n$, the sample proportion, where x is the number of items in a random sample of size n that belong to the class of interest.
- For $\mu_1 - \mu_2$, the estimate is $\hat{\mu}_1 - \hat{\mu}_2 = \bar{x}_1 - \bar{x}_2$, the difference between the sample means of two independent random samples.
- For $p_1 - p_2$, the estimate is $\hat{p}_1 - \hat{p}_2$, the difference between two sample proportions computed from two independent random samples.

A number of important properties are required of good point estimators. Two of the most important of these properties are the following:

1. The point estimator should be **unbiased.** That is, the expected value of the point estimator should be the parameter being estimated.
2. The point estimator should have **minimum variance.** Any point estimator is a random variable. Thus, a minimum variance point estimator should have a variance that is smaller than the variance of any other point estimator of that parameter.

The *sample* mean and variance $\bar{x}$ and s^2 are unbiased estimators of the *population* mean and variance μ and σ^2, respectively. That is,

$$E(\bar{x}) = \mu \quad \text{and} \quad E\left(s^2\right) = \sigma^2$$

where the operator E is simply the expected value operator, a shorthand way of writing the process of finding the mean of a random variable. (See the supplemental material for this chapter for more information about mathematical expectation.)

Unbiased Estimators

Definition

The point estimator $\hat{\Theta}$ is an unbiased estimator for the parameter θ if

$$E(\hat{\Theta}) = \theta \qquad (7\text{-}5)$$

If the estimator is not unbiased, then the difference

$$E(\hat{\Theta}) - \theta \qquad (7\text{-}6)$$

is called the bias of the estimator $\hat{\Theta}$.

Example

Suppose that X is a random variable with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the population represented by X. Show that the sample mean $\overline{X}$ and sample variance S^2 are unbiased estimators of μ and σ^2, respectively.

First consider the sample mean. In Equation 5.40a in Chapter 5, we showed that $E(\overline{X}) = \mu$. Therefore, the sample mean $\overline{X}$ is an unbiased estimator of the population mean μ.

Now consider the sample variance. We have

$$E(S^2) = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}\right] = \frac{1}{n-1} E \sum_{i=1}^{n}(X_i - \overline{X})^2$$

$$= \frac{1}{n-1} E \sum_{i=1}^{n}(X_i^2 + \overline{X}^2 - 2\overline{X}X_i) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\overline{X}^2)\right]$$

Example (continued)

The last equality follows from Equation 5-37 in Chapter 5. However, since $E(X_i^2) = \mu^2 + \sigma^2$ and $E(\overline{X}^2) = \mu^2 + \sigma^2/n$, we have

$$\begin{aligned} E(S^2) &= \frac{1}{n-1}\left[\sum_{i=1}^{n} (\mu^2 + \sigma^2) - n(\mu^2 + \sigma^2/n)\right] \\ &= \frac{1}{n-1}(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2) \\ &= \sigma^2 \end{aligned}$$

Therefore, the sample variance S^2 is an unbiased estimator of the population variance σ^2.

The sample standard deviation s is *not* an unbiased estimator of the population standard deviation σ. It can be shown that

$$E(s) = \left(\frac{2}{n-1}\right)^{1/2} \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}\sigma$$
$$= c_4\sigma \qquad (3\text{-}17)$$

Appendix Table VI gives values of c_4 for sample sizes $2 \le n \le 25$. We can obtain an unbiased estimate of the standard deviation from

$$\hat{\sigma} = \frac{s}{c_4} \qquad (3\text{-}18)$$

•As n gets large the bias goes to zero