

distribution for A is smaller than the standard deviation of the distribution for B , indicating that, over a large number of samples, the values of A cluster more closely around the unknown population parameter than do the values of B . Stated differently, the probability that A is close to the parameter value is higher than the probability that B is close to the parameter value.

In sum, to make an inference about a population parameter, we use the sample statistic with a sampling distribution that is unbiased and has a small standard deviation (usually smaller than the standard deviation of other unbiased sample statistics). The derivation of this sample statistic will not concern us, because the "best" statistic for estimating specific parameters is a matter of record. We will simply present an unbiased estimator with its standard deviation for each population parameter we consider. [Note: The standard deviation of the sampling distribution of a statistic is also called the **standard error of the statistic**.]

EXAMPLE 6.4

BIASED AND UNBIASED ESTIMATORS

Problem In Example 6.1, we found the sampling distributions of the sample mean \bar{x} and the sample median M for random samples of $n = 3$ measurements from a population defined by the following probability distribution:

x	0	3	12
$p(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

The sampling distributions of \bar{x} and M were found to be as follows:

\bar{x}	0	1	2	3	4	5	6	8	9	12
$p(\bar{x})$	$\frac{8}{64}$	$\frac{12}{64}$	$\frac{6}{64}$	$\frac{1}{64}$	$\frac{12}{64}$	$\frac{12}{64}$	$\frac{3}{64}$	$\frac{6}{64}$	$\frac{3}{64}$	$\frac{1}{64}$
M	0	3	12							
$p(M)$	$\frac{32}{64}$	$\frac{22}{64}$	$\frac{10}{64}$							

- Show that \bar{x} is an unbiased estimator of μ in this situation.
- Show that M is a biased estimator of μ in this situation.

Solution

- The expected value of a discrete random variable x (see Section 4.3) is defined as $E(x) = \sum xp(x)$, where the summation is over all values of x . Then

$$E(x) = \mu = \sum xp(x) = (0)\left(\frac{1}{2}\right) + (3)\left(\frac{1}{4}\right) + (12)\left(\frac{1}{4}\right) = 3.75$$

The expected value of the discrete random variable \bar{x} is

$$E(\bar{x}) = \sum (\bar{x})p(\bar{x})$$

summed over all values of \bar{x} , or

$$E(\bar{x}) = (0)\left(\frac{8}{64}\right) + (1)\left(\frac{12}{64}\right) + 2\left(\frac{6}{64}\right) + \cdots + (12)\left(\frac{1}{64}\right) = 3.75$$

Since $E(\bar{x}) = \mu$, \bar{x} is an unbiased estimator of μ .

- The expected value of the sample median M is

$$E(M) = \sum Mp(M) = (0)\left(\frac{32}{64}\right) + (3)\left(\frac{22}{64}\right) + (12)\left(\frac{10}{64}\right) = 4.56 \approx 2.91$$

Since the expected value of M is not equal to μ ($\mu = 3.75$), the sample median M is a biased estimator of μ .

SECTION 6.2 Properties of Sampling Distributions: Unbiasedness and Minimum Variance

287

EXAMPLE 6.5

VARIANCE OF ESTIMATORS

Problem Refer to Example 6.4 and find the standard deviations of the sampling distributions of \bar{x} and M . Which statistic would appear to be a better estimator of μ ?

Solution The variance of the sampling distribution of \bar{x} (we denote it by the symbol $\sigma_{\bar{x}}^2$) is found to be

$$\sigma_{\bar{x}}^2 = E\{[\bar{x} - E(\bar{x})]^2\} = \sum(\bar{x} - \mu)^2 p(\bar{x})$$

where, from Example 6.4,

$$E(\bar{x}) = \mu = 3.75$$

Then

$$\sigma_{\bar{x}}^2 = (0-3.75)^2(8/64) + (1-3.75)^2(12/64) + \dots + (12-3.75)^2(1/64)$$

$$= \cancel{(0-3.75)^2(8/64)} + \cancel{(1-3.75)^2(12/64)} + \dots + \cancel{(12-3.75)^2(1/64)}$$

$$= 8.6667 = 8.0625$$

and

$$\sigma_{\bar{x}} = \sqrt{\frac{8.6667}{8.0625}} = 2.94 \quad 2.84$$

Similarly, the variance of the sampling distribution of M (we denote it by σ_M^2) is

$$\sigma_M^2 = E\{[M - E(M)]^2\}$$

where, from Example 6.4, the expected value of M is $E(M) = 2.91$. Then

$$\sigma_M^2 = E\{[M - E(M)]^2\} = \sum[M - E(M)]^2 p(M)$$

$$= \cancel{(0-2.91)^2(7/64)} + \cancel{(2-2.91)^2(13/64)} + \cancel{(12-2.91)^2(7/64)} - 20.9136$$

$$= (6-2.91)^2(32/64) + (3-2.91)^2(22/64) + (12-2.91)^2(10/64)$$

and

$$\sigma_M = \sqrt{\frac{20.9136}{17.1475}} = 4.59 \quad 4.14$$

Which statistic appears to be the better estimator for the population mean μ , the sample mean \bar{x} or the median M ? To answer this question, we compare the sampling distributions of the two statistics. The sampling distribution of the sample median M is biased (i.e., it is located to the left of the mean μ), and its standard deviation $\sigma_M = 4.59$ is much larger than the standard deviation of the sampling distribution of \bar{x} , $\sigma_{\bar{x}} = 2.84$. Consequently, for the population in question, the sample mean \bar{x} would be a better estimator of the population mean μ than the sample median M would be.

Look Back Ideally, we desire an estimator that is unbiased and has the smallest variance among all unbiased estimators. We call this statistic the **minimum-variance unbiased estimator (MVUE)**.

Exercises 6.10–6.20

Understanding the Principles

- 6.10 What is a point estimator of a population parameter?
 6.11 What is the difference between a biased and unbiased estimator?
 6.12 What is the MVUE for a parameter?
 6.13 What are the properties of an ideal estimator?

Learning the Mechanics

- 6.14 Consider the following probability distribution:

x	0	1	4
$p(x)$	$1/3$	$1/3$	$1/3$