TABLE 10.1 A Comparison of Observed Values of $s_{\bar{X}}$ (Based on 10,000 $\bar{X}$'s) and Theoretical Values of $\sigma_{\bar{X}}$ for Various Sample Sizes When the Three Parent Populations (Normal, Rectangular and Skewed) Sampled Have Equal Means and Standard Deviations (μ = 100 and σ = 15)

	$s_{\bar{X}}$					
Parent Population	*n = 1*	*n = 2*	*n = 5*	*n = 10*	*n = 25*	*n = 100*
Normal	14.90	10.61	6.74	4.81	2.96	1.498
Rectangular	15.03	10.69	6.66	4.70	2.97	1.487
Skewed	14.98	10.49	6.63	4.70	2.98	1.479
$\sigma_{\bar{X}} = \sigma/\sqrt{n}$	15	10.61	6.71	4.74	3.00	1.500

(where μ = 100 and σ = 15) is 2.66% for the normal distribution (see left-most figure in Panel A), 1.89% for the rectangular distribution (see middle figure in Panel A), and 2.58% for the skewed distribution (see right-most figure in Panel A) shown in Figure 10.3.

Note that Panels B, C, and D of Figure 10.3 are empirical sampling distributions in which n = 1, 2, and 5, respectively. For example, for Panel D, a sample of five observations was selected randomly from the normal parent population, the mean of these five observations was computed, and this process was repeated 10,000 times. The left-most figure in Panel D is the frequency distribution of these 10,000 means[12]—that is, this figure is an empirical sampling distribution of the mean when μ = 100, σ = 15, and n = 5. If the process had been repeated, not 10,000 but 1,000,000 times, the empirical sampling distribution would have become almost perfectly symmetrical and normal. The small amount of irregularity evident in the sampling distributions from the normal population would virtually disappear and the empirical sampling distribution would coincide with the theoretical sampling distribution.

Observe that the mean of all of the sampling distributions, the mean of the $\bar{X}$'s, is approximately 100 = μ in each figure. Indeed, the expression $E(\bar{X}) = \mu$ is another way of saying that the mean of the sampling distribution of an infinite number of samples (not just 10,000 as in Figure 10.3) is the parameter μ. In Panels B to E of Figure 10.3, the sample sizes are small (1, 2, 5, and 10); some degree of non-normality in the parent population continues to be evident in the sampling distributions, but progressively less so as n increases. Panel F gives the three corresponding empirical sampling distributions when n was increased to 25. Notice that the sampling distributions for the normal, rectangular, and skewed populations are very similar, yet n is only 25. To the untrained eye, the distributions in Panel F may not appear to be normal, but this is only because the vertical (Percent) axis has been scaled uniformly in Panels A to F so that the decrease in the variability of the sampling distribution would be evident. The sampling distributions in Figure 10.3 demonstrate that even in non-normal distributions the standard deviation of the $\bar{X}$'s—that is, the standard error of the mean—equals the standard deviation of the parent population divided by the square root of the sample size: $\sigma_{\bar{X}} = \sigma/\sqrt{n}$.

In Table 10.1, the standard errors, $s_{\bar{X}}$ (each based on 10,000 means for the various sampling distributions), are reported along with the theoretical value, $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. For

[12]The authors are indebted to George Kretke for this computer simulation demonstration. It is estimated that this demonstration done by hand using a table of random numbers and a hand calculator would have required approximately 2,500 hours—approximately one full working year!

example, when samples of n = 25 were drawn from a skewed parent population, the resulting 10,000 sample means had a standard deviation of 2.98, which agrees almost perfectly with the theoretical standard error of the mean, $\sigma/\sqrt{n} = 15/\sqrt{25} = 3$. In other words, *even when the parent population is not normal, the formula $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ accurately depicts the degree of variability in the sampling distribution.*

10.16 THE USE OF SAMPLING DISTRIBUTIONS

The notion of a sampling distribution is used by the mathematical statistician to derive the techniques of inferential statistics. Researchers do not create their own sampling distribution by repeatedly drawing samples from a population. That would be not be feasible and is unnecessary. In practice, only one sample of n cases is drawn; then the theory underlying the sampling distribution is used to establish a confidence interval. For example, an investigator might draw a sample of n = 200 cases and establish a single confidence interval, say, the .95 confidence interval, around $\bar{X}$. Many samples are not drawn to construct an actual sampling distribution of $\bar{X}$. Instead, one has a single interval, extending perhaps from 46.5 to 51.5. Is μ in this interval? It is not possible to know for certain. Is it rational to act as though μ is in this interval? Indeed it is, since in the long run μ would be missing from only 5% of the .95*CI*'s. The technique of interval estimation is based on the theoretical concept of the sampling distribution with its notion of infinitely many samples drawn and their means distributed in some known fashion.

10.17 PROOF[13] THAT $\sigma_{\bar{X}}^2 = \sigma^2/n$

The proof is straightforward that whatever the shape of the parent population, the sampling distribution of $\bar{X}$ has a mean of μ and variance of σ^2/n, where μ and σ^2 are the mean and variance of the population sampled and n is the sample size. X is the variable being measured on the population; its mean is μ and its variance is σ^2. A random sample of size n has a first observation X_1, a second observation X_2, ..., and an nth observation X_n; X_1 is merely the *first* score chosen in each sample, not the smallest score. Therefore, the collection of all possible X_1's, that is, all first scores chosen in all possible random samples from the population, forms a population with mean μ and variance σ. Thus, $X_1, X_2, \ldots, X_n$ are each random variables from a population with a mean of μ and a variance of σ^2.

The sample mean equals $(X_1 + X_2 + \ldots + X_n)/n$. The mean of the sampling distribution of means equals the *expected value* of $\bar{X}$:

$$E(\bar{X}) = E\left[\frac{(X_1 + X_2 + \ldots + X_n)}{n}\right]$$

$$= \left(\frac{1}{n}\right)E(X_1 + X_2 + \ldots + X_n)$$

$$= \left(\frac{1}{n}\right)[E(X_1) + E(X_2) + \ldots + E(X_n)]$$

[13]This section is a more mathematical presentation of the concepts presented intuitively in Sections 10.11 and 10.12. Although not essential for a conceptual understanding, it is provided for students who desire a closer look at the related mathematics.

Now X_1 has the same distribution over samples as does X_2 or any other X_i. Its population mean and variance are μ and σ; hence, the last term in the equation above equals:

$$E(\overline{X}) = \left(\tfrac{1}{n}\right)(\mu + \mu + \ldots + \mu) = \left(\tfrac{1}{n}\right)(n\mu) = \mu$$

Stated in words, regardless of the shape of the parent population, the expected value of $\overline{X}$ is the mean μ of the sampling distribution of the sample means, which is also the mean of the population being sampled. How much will $\overline{X}$ vary from sample to sample? If each sample contains more than one observation, the variance of the sample means will be smaller than the variance of the parent population. Notice that if $n = 1$, the sampling distribution of the mean would be the same as the frequency distribution of the parent population, as shown in Panel B of Figure 10.3. If samples with $n = 1$ were repeatedly drawn and a sampling distribution of these "sample means" were constructed, the variance of the original population, σ^2, and the variance of the sampling distribution of the "means" of samples of size 1 would be the same; $\sigma^2/n = \sigma^2/1 = \sigma^2$:

$$E(s_{\overline{X}}^2) = \sigma_{\overline{X}}^2$$

What is the variance of the means of samples of size $n = 2$ from a population? Let the population variance be σ^2. For each sample, $\overline{X} = (X_1 + X_2)/2$ is calculated. X_1 and X_2 are arbitrary designations for the first and second observations randomly drawn and are not related to the size of the scores. Consequently, over all random samples, X_1 has variance σ^2 and so does X_2. Because the samples are randomly drawn, there is no relationship ($\rho = 0$) between the values of the first and second observations in any sample. Since X_1 and X_2 are uncorrelated, the correlation, and the covariance, is zero between the first and second observations in a sample over infinitely many random samples from a population.

Now the variance over random samples of $\overline{X} = (X_1 + X_2)/2$ is denoted as follows:

$$\sigma_{\overline{X}}^2 = \sigma_{(X_1+X_2)/2}^2$$

When a variable is multiplied by a constant (in this instance, ½), the variance of the resulting variable is the original variance multiplied by the square of the constant (Section 5.11). Therefore,

$$\sigma_{(X_1+X_2)(1/2)}^2 = \left(\tfrac{1}{2}\right)^2 \sigma_{(X_1+X_2)}^2$$

If two variables are uncorrelated, then the variance of the sum of the two variables is the sum of their variances (Equation 7.8). We argued that X_1 and X_2 are uncorrelated:

$$\left(\tfrac{1}{2}\right)^2 \sigma_{X_1+X_2}^2 = \tfrac{1}{4}(\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2}) = \tfrac{1}{4}(\sigma_{X_1}^2 + \sigma_{X_2}^2)$$

The variance of X_1 over repeated random samples is σ^2, and so is the variance of $X_2 : \sigma_{X_1}^2 = \sigma_{X_2}^2 = \sigma^2$. Therefore, the equation can be written as follows:

$$\left(\tfrac{1}{4}\right)(\sigma_{X_1}^2 + \sigma_{X_2}^2) = \left(\tfrac{1}{4}\right)(2\sigma^2) = \sigma^2/2$$

The equation expresses the conclusion of the argument: The variance of the mean of samples of size 2 from a population with variance σ^2 is equal to $\sigma^2/2$. In this instance, $n = 2$ and $\sigma_{\overline{X}}^2 = \sigma^2/2$. This is no coincidence. It is true in general that for random samples of size n, $\sigma_{\overline{X}}^2 = \sigma^2/n$.

If random samples of size n are taken from a population with variance σ^2, then the variance of the mean, $\overline{X} = (X_1 + X_2 + \ldots + X_n)/n$, over n samples is given by:

$$\sigma_{\overline{X}}^2 = \sigma_{(X_1+X_2+\ldots+X_n)/n}^2$$

The right-hand side of the previous equation shows $\sigma_{\overline{X}}^2$ to be the variance of $(1/n)$ times the variance of the sum of the n *uncorrelated* variables $X_1, X_2, \ldots X_n$. Therefore:

$$\sigma_{\overline{X}}^2 = \left(\tfrac{1}{n}\right)^2 \sigma_{(X_1+X_2+\ldots+X_n)}^2$$

Each variable $X_i (i = 1, 2, \ldots, n)$ has a variance of σ^2 and is uncorrelated with the other $n - 1$ variables. Therefore, the variance of the sum of the n uncorrelated variables is the sum of the variances of the variables, because each of the $n(n - 1)/2$ covariance is 0. Thus:

$$\left(\tfrac{1}{n}\right)^2 \sigma_{(X_1+X_2+\ldots+X_n)}^2 = \left(\tfrac{1}{n}\right)^2 (\sigma_{X_1}^2 + \sigma_{X_2}^2 + \ldots + \sigma_{X_n}^2)$$

Because each variable has the same variance σ^2, the previous equation can be written as:

$$\left(\tfrac{1}{n}\right)^2 (\sigma^2 + \sigma^2 + \ldots + \sigma^2) = \left(\tfrac{1}{n}\right)^2 (n\sigma^2) = \frac{\sigma^2}{n} \quad \textbf{(10.7)}$$

A fundamental relationship is expressed in Equation 10.7: *The variance of the means of random samples of size* n *from a population with variance* σ^2 *is equal to* $\sigma^2/n = \sigma_{\overline{X}}^2$. The expression $\sigma_{\overline{X}}^2$ has traditionally been called the *variance error of the mean*. Its positive square root (Equation 10.8), is the *standard error of the mean:*

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \quad \textbf{(10.8)}$$

The standard error of the mean, $\sigma_{\overline{X}}$, is the standard deviation of the sampling distribution of the means of an infinite number of samples, each of size n, from a population with variance σ^2. Notice that in Table 10.1 and Figure 10.3 that the population from which samples were drawn had a standard deviation of 15, and the standard deviation of the sampling distribution of means of random samples of various sizes is in agreement with Equation 10.8. For example, when $n = 10$: $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 15/\sqrt{10} = 4.74$; whereas in Table 10.1, the values of $s_{\bar{X}}$ were 4.81, 4.70, and 4.70 for the normal, rectangular, and skewed distributions, respectively.

10.18 PROPERTIES OF ESTIMATORS

An estimate is a value of a sample statistic that equals the value of a population parameter plus some amount of error. For example, the sample mean $\bar{X}$ is an estimator of the population mean μ. There is a close analogy between the way in which a sample mean is calculated and the way in which one might calculate a population mean. It is logical to think of $\bar{X}$ as estimating μ. However, there are other ways of treating sample data to arrive at a value that estimates μ. Why not use the sample median or the sample mode as an estimate of μ? It is certainly possible to do this; however, by the criteria used in assessing the properties of an estimator, $\bar{X}$ turns out to be a better estimator of μ than either the sample median or the sample mode (Section 4.16).

In the following three sections, the properties of estimators of parameters will be examined. What are the different ways in which parameters can be estimated? Is one estimator to be preferred over all others for estimating a certain parameter, and why? The properties of unbiasedness, consistency, and efficiency will be considered.

10.19 UNBIASEDNESS

As discussed in Section 5.13, an estimator, $\hat{\theta}$, is said to be unbiased for estimating a parameter, θ, if the mean of the sampling distribution of the sample estimates equals the value of the parameter being estimated. Equivalently, an estimator, $\hat{\theta}$, is unbiased if its expected value, $E(\hat{\theta})$, is equal to the parameter, θ, being estimated.

Whatever the nature of the population being sampled, the sample mean, $\bar{X}$, is an unbiased estimator of the population mean, μ. Notice in Figure 10.3 that the value of the population mean, μ, is 100 and that the mean of any of the empirical sampling distributions of $\bar{X}$ is approximately 100. This illustrates the unbiasedness of $\bar{X}$ as an estimator of μ. If samples are drawn randomly from a normal distribution or some other symmetric distribution, then the sample median is also an unbiased estimator of the population mean, μ. In other words, the average of the medians on an infinite number of random samples from a normal distribution equals μ, the mean of the normal distribution (which is, of course, also its median).

There are many examples of biased estimators. Suppose one wishes to estimate ρ, the correlation between two variables in the population. Imagine that for a particular population $\rho = .75$. The mean of the sampling distribution of the sample correlation coefficient, r, will be slightly less than .75 for any finite sample size. Thus, r is a negatively biased estimator of ρ.[14] When the expected value of a statistic is less than the parameter being estimated, it is said to be *negatively biased*. Conversely, if $E(\hat{\theta}) > \theta$, $\hat{\theta}$ is said to be *positively biased*. In Equation 5.5, the variance in a sample is defined as $s^2 = \Sigma(X_i - \bar{X})^2/(n-1)$. It might have been more natural to measure variability by simply taking the average of the n squared

[14] The extent of the bias is exceedingly small—less than 1% if $n > 25$ (Olkin, 1967, p. 111).

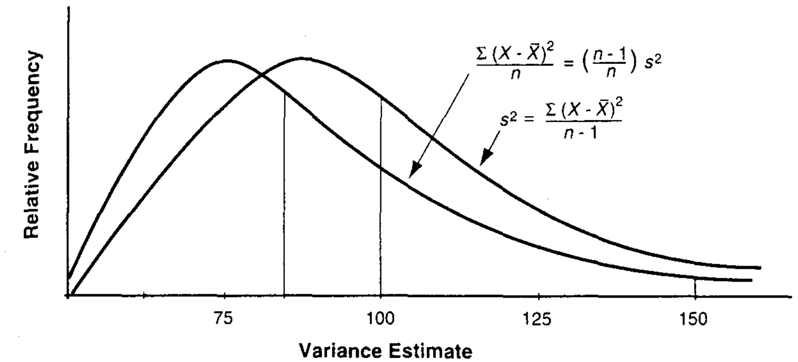

FIGURE 10.4 Sampling Distribution of s^2 and $\Sigma(X_i - \bar{X})^2/6$ for Random Samples of Size 6 from a Normal Distribution with Variance $\sigma^2 = 100$

deviations around the sample mean. Instead, it was decided to place $(n-1)$ in the denominator of s^2 because the quantity s^2 is an unbiased estimator of the population variance σ^2; whereas, $\Sigma(X_i - \bar{X})^2/n$ is negatively biased as an estimator of σ^2. That is:

$$E\left[\frac{\sum_i (X_i - \bar{X})^2}{n}\right] < \sigma^2$$

Suppose that we took *many* random samples from any population with variance σ^2 and calculated s^2 each time. The average of these sample variances would equal σ^2. Hence, s^2 is an unbiased estimator of σ^2. If instead, $\Sigma(X_i - \bar{X})^2/n$ had been calculated on each sample, the average of these quantities would have been smaller than σ^2, namely $[(n-1)/n]\sigma^2$. Of course, if n were quite large—100 or more for example—the difference between s^2 and $\Sigma(X_i - \bar{X})^2/n$ would be very small, because the value of $(n-1)/n$ (the ratio of the degrees of freedom to sample size) would approach 1, and the estimator would contain only a small bias as an estimator of σ^2.

Suppose that one has a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 100$. If an infinite number of random samples of size $n = 6$ were drawn from the population and both s^2 and $\Sigma(X_i - \bar{X})^2/6$ were calculated for each sample, the two sampling distributions in Figure 10.4 would be obtained.

Notice that the mean of the sampling distribution of s^2 is 100 the value of σ^2. This illustrates the unbiasedness of s^2 in this instance, that is, $E(s^2) = \sigma^2$. The mean of the sampling distribution of $\Sigma(X_i - \bar{X})^2/6$ is equal to 83.33. In this instance, the bias introduced into the estimation of σ^2 by using n in place of $n-1$ in the denominator of the sample variance is sizable, that is, $(n-1)/n = 5/6$ here.

How was it decided that the denominator of the sample variance should be $n-1$ (i.e., the degrees of freedom, ν) in order for $E(s^2) = \sigma^2$? It was not determined empirically or by

Statistical Methods in Education and Psychology

Third Edition

Gene V Glass
Arizona State University, Tempe

and
Kenneth D. Hopkins
University of Colorado, Boulder

Allyn and Bacon
Boston • London • Toronto • Sydney • Tokyo • Singapore