9

SAMPLING DISTRIBUTIONS

RESEARCH EXAMPLE NINE

In a study of a sample of 59 deaf children, Hirshoren, Hurley, and Kavale (1979) found that the average IQ score on the Wechsler Intelligence Scale for Children-Revised (WISC-R) was 88.07. The population mean for hearing children on the WISC-R is $\mu = 100$. They wanted to provide standardization information on the revised WISC for future research with children with impaired hearing. As part of this research they wanted to ask, "Is the population average IQ for deaf children equal to 100?"

They could see that 88.07 is not equal to 100, but 88.07 is the value of a *sample* mean $\bar{X}$, and the question is about the *population* mean of IQ for deaf children. Obviously the sample mean is not 100, but what about the population mean? Could they have obtained a sample with $\bar{X} = 88.07$ from a population with mean equal to 100? Yes, it is *possible* to get an $\bar{X}$ this low if $\mu = 100$, but what about the *probability* of getting such an $\bar{X}$? Since the important question involves the probability of such an occurrence, we should restate the question more precisely. What is the probability of getting 88.07 or lower as a sample mean ($\bar{X}$) if the population mean is $\mu = 100$? The answer to this question will give us information as to how unusual this value of 88.07 is *if* $\mu = 100$. If 88.07 is an unusual value from a distribution with $\mu = 100$, then the probability will be low and we would doubt that $\mu = 100$ for the population of IQ scores for deaf children.

Another way of thinking about this problem is to ask, "If we repeatedly sampled 59 deaf children and computed $\bar{X}$ on each sample, what would be the mean of the distribution of those $\bar{X}$'s? Would it be 100? Now it is impractical to actually repeatedly sample in this fashion, so we must answer these questions with our single $\bar{X} = 88.07$. We are led back to the question we asked in the preceding paragraph: *If* $\mu = 100$, what is the probability of getting $\bar{X} = 88.07$ or lower? This question about the probability of $\bar{X}$ if $\mu = 100$ is going to require us to gain information about the distribution of $\bar{X}$'s. Knowledge of the key facts about the distribution of $\bar{X}$ will be crucial to our decision making about μ.

To answer any question about the probability of $\bar{X}$, we have to learn about the distribution of $\bar{X}$, called the *sampling distribution* of $\bar{X}$. With knowledge of the mean, variance, and shape of the sampling distribution of the statistic $\bar{X}$, we can find probabilities associated with $\bar{X}$. We can find the probability of $\bar{X}$ being less than or equal to 88.07 if $\mu = 100$, and ultimately we can make a decision about the parameter μ. We will be able to decide whether $\mu = 100$ is a plausible value once we learn about the sampling distribution of $\bar{X}$, which is the focus of this chapter.

Just to satisfy your curiosity, the probability of getting a sample which has $\bar{X} = 88.07$ from a population with $\mu = 100$ is very small. Thus, we would decide to reject the idea that $\mu = 100$ for the population of IQ scores for deaf children. This research is referred to several times in this and future chapters. In these chapters we see how to find the probability of $\bar{X}$ and how to make decisions about parameters.

9.1 INTRODUCTION AND REVIEW

When the concept of a sampling distribution was introduced in Chapter One, it was included in a list of seven topics common to all inferential statistical methods: descriptive statistics, probability, estimation, variability of statistics, distributions of statistics, theoretical reference distributions, and hypothesis testing. Since Chapter One, we have spent most of our time in descriptive statistics, with coverage of one theoretical reference distribution (normal distributions, Chapter Five) and probability (Chapter Eight). Now we are ready to cover three of these topics in this chapter: estimation, variability of statistics, and distributions of statistics.[1]

Let's use the final examination scores for the students taught with a new (experimental) method as an example (originally presented as part of a two-group study, Table 1.4). The final examination scores and the values of $\bar{X}$ and $\overset{*}{s}^2$ are given in Table 9.1.

A Brief Review

In addition to those seven topics in "Overview of Statistics," the whole process of drawing samples from populations is important to this chapter. Recall the

TABLE 9.1
Final Examination Scores for New Teaching Method

92	86	83	88
85	87	78	81
91	92	75	87
99	94	90	92

$$\bar{X} = \frac{\Sigma X}{N} = \frac{1400}{16} = 87.5$$

$$\overset{*}{s}^2 = \frac{N\Sigma X^2 - (\Sigma X)^2}{N^2}$$

$$= \frac{(16)(123{,}072) - 1400^2}{16^2} = \frac{9152}{256}$$

$$= 35.75$$

[1]Periodically you should review these seven topics to maintain your perspective on statistics. If you have forgotten what was discussed earlier about these topics, now is a good time for such a review. Go to Section 1.8, read through the "Preview of Inferential Statistics," then return here.

definitions of the following three concepts: population, sample, and random sampling. A *population* is the target group for our inferences, some large group of subjects, or an entire set of subjects, objects, measurements, or events all of which share some common characteristic. The population for the experimental teaching method would be any students who might take this introductory course in the near future if the new method were implemented. A *sample* is some subgroup, subset, or part of the population and is the small group of subjects used in the actual research. The sample here is obviously the 16 students from the experimental group or their scores on the final examination. *Random sampling* is a procedure in which selection of any one observation from a population is independent of the selection of any other observation from the same population. In this teaching method example, random sampling would not have been used. Students were accepted into the research as they enrolled in the course, and the researcher "judged" that the sample did not differ from that expected from random sampling. Remember that random sampling and random assignment of subjects to groups are different processes.

Sampling: More Information

Sampling is the process of selecting the subgroup of the population called the *sample*. Random sampling reduces the chances of systematic biases in the sample if the correct population is sampled. This means that there could be bias in any one sample, but the average bias over a long run of samples is zero. However, it is possible to draw a random sample and reach an incorrect conclusion. This can occur when the sample is randomly selected from the wrong population.

An example is the random sample taken by the *Literary Digest* in the 1936 Presidential election. From the results of a preelection poll, Landon was predicted to win over Roosevelt. The *Literary Digest* had accurately predicted the 1932 election, using the same process as used in 1936, but in 1936 the prediction was wrong. What happened? We might be tempted to answer "sampling variability or chance," but there was actually a severe bias in the population sampled. The sample was randomly selected from subscribers to *Literary Digest* who had telephones. First, their subscribers may not have represented the voting population. Second, people who had telephones in their home may not have represented the voting population. In contrast with today, in 1936 only a small percentage of the population had telephones since telephones were concentrated in urban areas and could be afforded only by the wealthy. In this case, random sampling did not guarantee freedom from bias because the wrong population was sampled. As can be seen from this example, statistics used in the inferential process are no better than the sampling. If the sampling is inferior, then the statistics and the inferential process are inferior.

In most behavioral science research, the difficulty of actually accomplishing random sampling promotes use of volunteers or some other form of judgment sampling. Use of volunteers may introduce biases into the process, as might any judgment sampling procedure. The researcher makes a judgment that the sample is as though it were randomly selected from a given population, and this judgment and its quality determine the quality of the inferences. However they are taken, most samples are treated as if they are random. Because of this assumption and because of the mathematical simplicity which random sampling brings to statistical procedures, random sampling is assumed throughout the remainder of this text.

Another topic in the area of sampling is the issue of sampling with or without replacement. *Replacement* refers to replacing a subject or score in the population after it has been sampled and used in the research, where it is available to be sampled again. Sampling without replacement is an accurate model for what behavioral scientists actually do, but sampling with replacement is the model upon which most statistical procedures are based. Resolution of this problem comes through the fact that for large populations, the probabilities associated with the two methods of sampling are nearly equal. Thus, there are no practical differences between sampling large populations with or without replacement.

9.2 THREE DIFFERENT TYPES OF DISTRIBUTIONS

We sampled the population of students taught with the new method and obtained the above sample which gave $\overline{X} = 87.5$. What would happen if we took another sample of $N = 16$ scores? Would we get the same scores as in the first sample? It is not likely. Would we get the same $\overline{X}$ as in the first sample? Not likely. What if we took a third sample? And a fourth? What if we took infinitely many samples from our population, computed $\overline{X}$ for each sample, and put these $\overline{X}$'s in a distribution? We would get a distribution of $\overline{X}$, called a *sampling distribution,* like that illustrated in Figure 9.1. And $\overline{X} = 87.5$ would be one value of $\overline{X}$ in this distribution.

Now, we never actually do this repeated sampling to obtain sampling distributions; rather, we get needed information about sampling distributions from mathematical statistics. This information comes to us as results from mathematical calculations or from computer simulations called Monte Carlo (from gambling simulations) studies. Still, it is helpful to conceive of sampling distributions as if they were formed by repeatedly drawing samples of a given size from a population, even though we don't get our information in this way. As you see in Figure 9.1, the population is a distribution of X's, the sample is a distribution of X's, but the sampling distribution of $\overline{X}$ is a distribution of $\overline{X}$'s. In the coverage of populations, samples, and sampling distributions which follows, keep Figure 9.1 in mind and refer to it often, looking for differences and similarities in these three types of distributions.

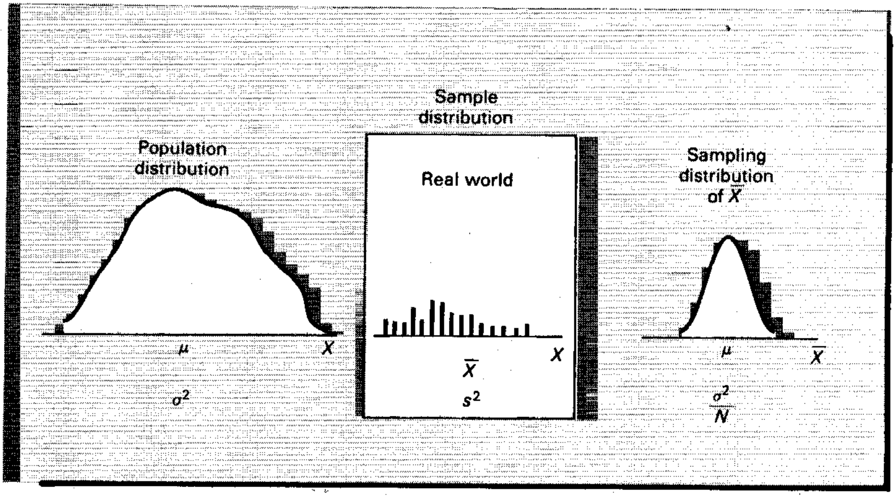

FIGURE 9.1 Three different types of distributions.

9.3 POPULATION DISTRIBUTION

Populations Are Usually Large

The characteristics of populations are that they are usually large, unobtainable, and hypothetical. Populations are large because behavioral science researchers do want to generalize to some broadly defined group even though their sampling process is restricted to some local, narrowly defined group. Remember, it is the target of the inferences which defines the population. In our example, the researcher wants to generalize to all students who might take the course taught by the new method. This would include present and future students at the university where the research was done and any other institution which might adopt the new teaching method. Although populations are usually large, most statistical procedures idealize this largeness to assume that the population is infinite. If there are infinitely many scores in the population, then infinitely many samples of size N could be selected.

Populations Are Usually Unobtainable and Hypothetical

By the very fact that populations are usually large, they are usually unobtainable. This does not mean that the population is unknown: The population is known in the sense that we define the characteristics for inclusion in the population. However, usually not all subjects in the population can be identified, assigned a unique number for the purpose of random sampling, or measured on the dependent variable. Additionally, not all subjects can be given the conditions of the experiment, and so the population is hypothetical. These three characteristics (large, unobtainable, and hypothetical) emphasize that the main thrust of statistics is generalization from the known to the unknown. Sample statistics are known, and population parameters are generally unknown. For final examination scores for students taught by the new method, the population mean is unknown.

Populations Described

Population distribution Distribution of raw scores X in the population

Because population parameters are unknown, I have not given any formulas for computing them. Rather, I merely defined symbols for parameters, such as μ and σ^2. I usually draw any picture of the population as an ambiguously shaped, continuous curve such as in Figure 9.1, so you won't get any false impression about the population. Note that the **population distribution** is one of raw scores X and has parameters as summary characteristics.

Populations and Inference

Any interest in the population distribution is in terms of the final generalization: The researcher wants to be able to make some decisions about parameters of this distribution or about the distribution itself. In the teaching method example, we want to be able to generalize to all students who might take this course taught by the new method. We are interested in current students and future students, and we want to generalize the results from our sample to these students and from our known value of $\overline{X}$ to the unknown value of μ for the population of these students. The first step in the decision process is to draw a random sample of size N from the population.

9.4 SAMPLE DISTRIBUTION

Samples Described

Sample distribution Distribution of raw scores X in the sample

We now have N values of X (N raw scores); we have a sample. Use Figure 9.1 to compare the sample distribution with the population distribution. Note that like the population, the **sample distribution** is a distribution of X. Also, the sample has its summary characteristics, but these are not parameters; they are the statistics $\overline{X}$ and $\overset{*}{s}{}^2$. In further contrast to the population, the sample distribution is discrete and is shown as a histogram. Also the sample differs from the population in that the sample is known, obtainable, and real. We can describe the shape of the sample distribution because we have it. For these reasons, the distribution of the sample is in the region of Figure 9.1 labeled "Real World." In our example, we have the $N = 16$ final examination scores with $\overline{X} = 87.5$ and $\overset{*}{s}{}^2 = 35.75$.

Samples and Inference

The distribution of these N scores is generally not of too much interest itself since our focus is on the statistics. We are interested in the value of $\overline{X}$ or $\overset{*}{s}^2$ or some other statistic computed from the N scores and our ability to infer from the statistic to its corresponding parameter. In the example we are interested in the value of $\overline{X} = 87.5$ and in using it to make decisions about the population mean μ. We use the sampling distribution as the basis for our inference from a statistic to a parameter.

9.5 SAMPLING DISTRIBUTION

Statistics Have Variability

One very important fact is that statistics have variability; that is, sample statistics may differ from sample to sample. Although the researcher actually computes only one statistic, such as $\overline{X}$ for the final examination scores above, you must realize that *if* the researcher were to draw another random sample with $N = 16$ from the same population, the second $\overline{X}$ would most likely not equal the $\overline{X}$ from the first sample. Another way of conceptualizing this important fact is to say that the statistic is a random variable which can take on many potentially different values before the sample is actually drawn and the statistic is computed. Thus, $\overline{X} = 87.5$ is just one of many values which the researcher could have obtained. The sample which gave $\overline{X} = 87.5$ just happened to be selected.

Statistics Have Distributions

The conceptualization of any statistic as a random variable also helps us to see that every statistic has a distribution. That is, any statistic that can be computed from sample data could take on many potentially different values which theoretically have a distribution. To students who first hear of this idea, the notion of a statistic having a distribution seems a bit ridiculous. With great doubt in their voices they say, "You mean that my single value of $\overline{X} = 87.5$ has a distribution? Of course, when there is only a single value of a statistic in hand, there *is* only one value and it is difficult to conceive of the distribution from which the statistic has come. But the fact remains that there exists a distribution of the potential values of the statistic. Before the sample is actually drawn and the statistic computed, any number of values for the statistic potentially exist. Hence, the actual value of $\overline{X}$ which is computed from the data is only one value from potentially many values of $\overline{X}$ which could have been obtained in this sampling procedure. The distribution of any statistic is called the *sampling distribution* of that statistic. For example, the distribution of $\overline{X}$ is called the sampling distribution of $\overline{X}$. Similarly, every statistic which can be computed from the data has a distribution of all possible values of the statistic, which is called the sampling distribution of that statistic.

Sampling Distributions Defined

Here is a more complete definition of a sampling distribution:

Sampling distribution
Distribution of all possible values of a statistic

The **sampling distribution** of a statistic is a distribution which is as if it had been formed by drawing infinitely many samples of a given size *N* from some population, computing the statistic on the scores for each sample, and arranging these infinitely many statistics in a distribution.

Several important points should be noted:

1. The size of each sample is *N*, which is the number of scores in the sample distribution. Do not confuse *N* with the number of statistics in the sampling distribution.
2. It is as if all possible values of the statistic have been computed, so there are infinitely many statistics in the sampling distribution.
3. All statistics in a sampling distribution are as if they have been computed from samples of a common size *N*.
4. The sampling distribution of any statistic is theoretical and is never actually obtained by repeated sampling. This point is obvious since it is impossible to draw infinitely many samples.
5. For each different statistic computed from every different size sample, there is a distinct sampling distribution.
6. The sampling distribution of any statistic is a probability distribution (a theoretical relative-frequency distribution) from which we can compute probabilities and make decisions.

Sampling Distributions Described

Examine Figure 9.1 where the sampling distribution of $\overline{X}$ has been drawn along with the population and sample distributions. Use this sampling distribution as an example, and compare and contrast it to the other distributions. Like the population distribution, the sampling distribution is theoretical, thus not in the real world, and is drawn as a smooth, continuous curve; also, it has parameters as summary characteristics. However, unlike the other two distributions, this sampling distribution is a distribution of statistics, $\overline{X}$'s, not of raw scores. Another way of saying this is that the sampling distribution has $\overline{X}$ as its random variable, rather than X.

Recall the sampling distribution of $\overline{X}$ which we built in Chapter Eight. The population had only three scores (1, 2, and 3) with a large equal frequency

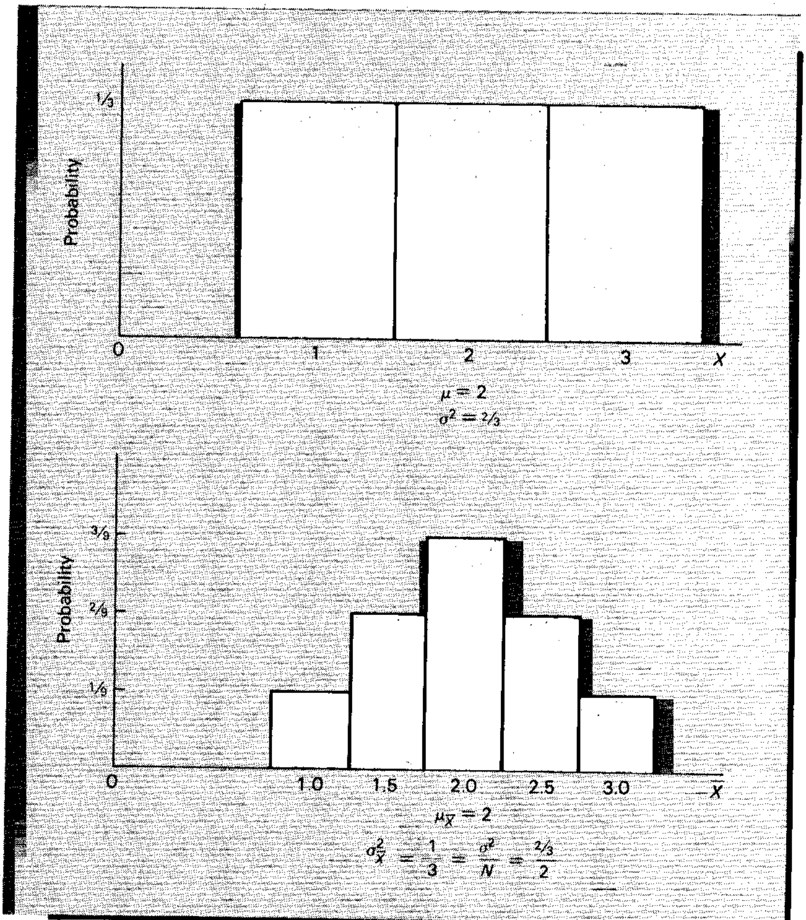

FIGURE 9.2
Population with three values and sampling distribution of $\overline{X}$, $N = 2$.

of each. We took all possible samples of size $N = 2$ and computed $\overline{X}$ on each. Then we computed the probabilities for the distribution of $\overline{X}$'s. This sampling distribution of $\overline{X}$ is reproduced in Figure 9.2. This sampling distribution is atypical of the sampling distribution of $\overline{X}$ for several reasons. First, the population has a very small number of possible values. Second, the sample size $N = 2$ is small. This combination gives a sampling distribution of $\overline{X}$ which is deceptively simple. It has a small number of $\overline{X}$ values and is discrete. But it does show several key features of the sampling distribution of $\overline{X}$ which would be much more difficult to demonstrate with a larger population or sample size. To start, the mean of the sampling distribution of $\overline{X}$ is the same as the mean of the population $\mu = 2$. Then the variance of the sampling distribution of $\overline{X}$ is $\sigma^2_{\overline{X}} = 1/3$, which is the same as $\sigma^2/N = (2/3)/2$. Finally, even with the limitations mentioned, a normal distribution would give a good approximation to the shape of the sampling distribution of $\overline{X}$.

Sampling Distributions and Inference

Since we cannot actually form sampling distributions by repeated sampling, we must turn to mathematical statistics to provide us with all the necessary information about the distribution. We rely on the results of several theorems to give us facts needed to completely specify a sampling distribution apart from actual repeated sampling. The information usually desired for a sampling distribution of a statistic includes the theoretical distribution which it exactly follows, or at least closely approximates, and the parameters of the sampling distribution. Once we know these important characteristics, there is no need for actually obtaining the sampling distribution by repeated sampling because we can use a tabled theoretical distribution such as the normal distribution to make probability statements about the statistics.

For the sampling distribution of $\overline{X}$, the facts provided are that if we have a random sample of N independent X's and

X is distributed normally with mean μ and variance σ^2

then

$\overline{X}$ is distributed normally with mean μ and variance σ^2/N

Shape of the sampling distribution of $\overline{X}$
If X is distributed normally, then $\overline{X}$ is also distributed normally

Parameters of the sampling distribution of $\overline{X}$
Sampling distribution of $\overline{X}$ has mean equal to μ and variance equal to σ^2/N

Thus, if X is normally distributed, then $\overline{X}$ is also normally distributed with the same mean μ but with variance equal to the population variance divided by the sample size, σ^2/N. We know the sampling distribution of $\overline{X}$ is fit by the theoretical normal distribution and has parameters μ and σ^2/N (mean and variance). Note that as N increases, the variance of the sampling distribution of $\overline{X}$, σ^2/N, decreases.

For a summary of these three types of distributions, examine Table 9.2. Use a piece of paper to cover up the answers in the body of the table to quiz yourself about this information.

If there were three or more of the entries in the table which you did not know, you should review the material in this section before proceeding to the next topic.

Need for Sampling Distributions

We need sampling distributions to obtain probabilities about statistics for decision making about parameters. In Chapter One, we used probability in decision making about the fairness of a coin. In that problem, the use of *fair* in connection with a coin was equivalent with saying that the probability of a head was .5, or $p = .5$. We can consider p to be a parameter, and our decision making about the fairness of the coin was decision making about p. We counted the number of heads to get some indication of the fairness of the coin, and we could consider the number of heads as the statistic. So we asked,

TABLE 9.2
Summary of Population, Sample, and Sampling Distributions

	Population Distribution	Sample Distribution	Sampling Distribution of $\overline{X}$
Random variable (what is in distribution)	X (raw score)	X (raw score)	$\overline{X}$ (statistic)
Summary characteristics	μ σ^2	$\overline{X}$ $\hat{s}^2$	μ σ^2/N
Size of distribution	Large	Small	Large
Obtainable?	No	Yes	No
Hypothetical or real?	Hypothetical	Real	Hypothetical
Continuous or discrete?	Continuous	Discrete	Continuous
Shape	Generally unknown	Can describe	Normal if population is normal

What is the probability of getting 10 heads on 10 tosses of a coin if it is fair? And we decided that the answer $1/1024 = .009766$ made us doubt the premise of the fair coin. We had to be able to get the probability of the obtained statistic (number of heads) in order to make a decision about the parameter ($p = .5$ or "coin is fair"). Reread the last sentence because this is the heart of the issue: *Without sampling distributions of statistics we could not find probabilities of obtained statistics and could not make decisions about parameters.*

Now shift your thinking to the sampling distribution of $\overline{X}$. In using $\overline{X}$ to make decisions about μ, we need a probability from the sampling distribution of $\overline{X}$. Without the sampling distribution of $\overline{X}$, we could not obtain probabilities associated with the obtained $\overline{X}$ and could not make decisions about the parameter μ. We need probabilities of statistics which we can get only from sampling distributions of statistics. The entire process of inferential statistics is thus dependent on sampling distributions.

The logic of the need for sampling distributions can be summarized as follows:

1. The population and its parameters are unknown, yet we want to make decisions about them.
2. The sample and its statistics are known, and the statistics are estimates of parameters. However, we cannot simply use a statistic as equal to a parameter and make our decision directly because statistics have variability.
3. We use the sampling distribution of a statistic to quantify the information about the variability of the statistic into probability. Thus, we make our decision indirectly, from the sample statistic through the sampling distribution to the population parameter. For example, we calculate $\overline{X}$ from the

sample, refer it to the sampling distribution of $\overline{X}$ for a probability statement, and then use the probability statement to make a decision about the population mean μ.

Sampling distributions serve as the bridge between the known and the unknown—the statistic and the parameter. We use the sampling distribution of the statistic to obtain a probability to make the inference from the obtained statistic to the unobtainable parameter.

9.6 SAMPLING DISTRIBUTION OF $\overline{X}$ AS AN EXAMPLE

Any statistic which can be computed from the sample has a sampling distribution. Not only is there a sampling distribution of $\overline{X}$, but also there is a sampling distribution of $\hat{s}^2$, a sampling distribution of X_{50}, etc. Several other sampling distributions are covered later, but now we want to use the sampling distribution of $\overline{X}$ as an example.

Information about the Sampling Distribution of $\overline{X}$

We already have considerable information about the sampling distribution of $\overline{X}$, but we want to review what we know and add to this knowledge. Given that the definition of a sampling distribution includes the concept of infinitely many (all possible) samples, we know that

1. The mean of the sampling distribution of $\overline{X}$ is μ.
2. The variance of the sampling distribution of $\overline{X}$ is $\sigma^2_{\overline{X}} = \sigma^2/N$.
3. The shape of the sampling distribution of $\overline{X}$ is normal if the population is normal in shape.

Note that the mean of the sampling distribution of $\overline{X}$ is *exactly* equal to the mean of the population. Many students miss this point, so it bears repeating. Let's use as an example the sampling distribution of the mean of IQ scores introduced in Research Example Nine. If the mean of the population is μ = 100, then the mean of the sampling distribution of $\overline{X}$ is μ = 100. Whatever the value of μ, even though it is usually unknown, the mean of the sampling distribution of $\overline{X}$ has exactly the same value. The sampling distribution of $\overline{X}$'s calculated from IQ scores of infinitely many repeated samples of 59 deaf children has the same mean as the population of IQ scores of all deaf children, whatever its value.

The standard deviation of the sampling distribution of $\overline{X}$ is

$$\sigma_{\overline{X}} = \sqrt{\frac{\sigma^2}{N}} = \frac{\sigma}{\sqrt{N}}$$

Standard error of the mean
Standard deviation of $\bar{X}$ is $\sigma/\sqrt{N}$

and it is given a special name, the **standard error of the mean.** Note that $\sigma_{\bar{X}}$ is a standard deviation and is conceptually the same as other standard deviations. For the IQ scores in Research Example Nine, the value of σ is the square root of 225, or $\sigma = 15$. Since $N = 59$, the value of the standard error of the mean of these IQ scores is $\sigma_{\bar{X}} = 15/\sqrt{59} = 1.95$.

It bears repeating that the information we have on the sampling distribution of $\bar{X}$ depends on the concept of infinitely many samples. Anything less than infinitely many samples, such as a computer sampling experiment would have, leads to an *approximation* of the sampling distribution and only approximations of μ, σ^2/N, and normality.

Let's return to the teaching method example. Suppose we know somehow that $\mu = 84.75$ and $\sigma^2 = 40$ for the population. You have just obtained the sample in Table 9.1 of $N = 16$ final examination scores with $\bar{X} = 87.5$. Assume that final examination scores are normally distributed in the population. Now list everything you know about the sampling distribution of $\bar{X}$ *for this situation.* Go back through this section, and be as specific as you can about the information concerning the sampling distribution of $\bar{X}$. Are you finished? The mean of the sampling distribution of $\bar{X}$ would be $\mu = 84.75$; the variance of the sampling distribution of $\bar{X}$ would be $\sigma^2/N = 40/16 = 2.5$; the shape of the sampling distribution of $\bar{X}$ would be normal; and the standard error of the mean would be $\sigma_{\bar{X}} = \sigma/\sqrt{N} = \sqrt{2.5} = 1.58$.

Defining $z_{\bar{X}}$

When we need to compute probabilities about $\bar{X}$'s, we use the above information about the sampling distribution of $\bar{X}$. Looking at the sampling distribution of $\bar{X}$, we realize that if the population is normal in shape, then the sampling distribution of $\bar{X}$ is normal in shape. We can use Table C to look up probabilities if we can change the sampling distribution of $\bar{X}$ from a normal distribution with mean μ and standard deviation $\sigma/\sqrt{N}$ to the standard normal distribution with mean of 0 and standard deviation of 1.

z score for $\bar{X}$
$z_{\bar{X}}$ is formed by subtracting μ from $\bar{X}$ and dividing by the standard error of $\bar{X}$

To compute probabilities in the sampling distribution of $\bar{X}$, we need to form a **z score for $\bar{X}$** which is

$$z_{\bar{X}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \tag{9.1}$$

Remember that z doesn't change the shape of the distribution, so we can use a normal distribution for $z_{\bar{X}}$ as we could for $\bar{X}$. Notice that formula 9.1 takes on the same form as z for raw scores given in formula 5.2: something minus its mean, divided by its standard deviation. Here, we wish to get a z score for $\bar{X}$, so we subtract its mean μ and divide by its standard deviation $\sigma/\sqrt{N}$. Even though this formula is slightly more complicated than that for a z score given in formula 5.2, the general form is the same. Using formula 9.1, we can convert $\bar{X}$ to $z_{\bar{X}}$ and then use Table C to obtain the desired probability.

For Research Example Nine, $\mu = 100$, $\sigma = 15$, $N = 59$, and $\overline{X} = 88.07$. The value of $z_{\overline{X}}$ would be

$$z_{\overline{X}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}} = \frac{88.07 - 100}{15/\sqrt{59}} = -6.11$$

From what we know about z scores, this value of -6.11 tells us that $\overline{X} = 88.07$ is more than 6 standard deviations of $\overline{X}$ (standard errors) below 100. This value of $\overline{X}$ would be very unusual if the population had a mean of 100.

Is $z_{\overline{X}}$ Practical?

From examining the formula for $z_{\overline{X}}$, we can see that we need to have numerical values for μ and σ^2 before we can obtain a numerical value for $z_{\overline{X}}$. If we are dealing with a random variable that is a score on a standardized test, such as IQ scores, GRE scores, etc., then we can easily use the formula for $z_{\overline{X}}$, since μ and σ^2 are known in the form of norms for norming populations for standardized tests. Sometimes we want to test whether the population we sample is the same as that on which norms were calculated (the norming population). So we want to test whether our unknown μ is the same as the norm value. For example, is the mean (μ) IQ of firefighters the same as $\mu = 100$ for the entire population? Also, if previous research or some specific theory gives us the values of μ and σ^2 for a specific population, we can proceed to compute $z_{\overline{X}}$. However, for most research in the behavioral sciences, the values of μ and σ^2 are never known since the dependent variables are not standardized tests. Because of this, use of $z_{\overline{X}}$ is restricted to hypothetical examples used in the teaching of statistics or to one of the situations given here. In spite of its restricted use, $z_{\overline{X}}$ and probabilities computed from $z_{\overline{X}}$ illustrate all the desired principles of decision making.

Computing Probabilities

Suppose we have IQ scores ($\mu = 100$, $\sigma = 15$) for nine firefighters. If $\overline{X} = 105$ for the $N = 9$ firefighters, what is $p(\overline{X} \geq 105)$? Since

$$z_{\overline{X}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}} = \frac{105 - 100}{15/\sqrt{9}} = \frac{5}{5} = 1.00$$

and

$$p(z_{\overline{X}} \geq 1.00) = .1587 \qquad \text{(from Table C)}$$

then

$$p(\overline{X} \geq 105) = .1587$$

TABLE 9.3
Computation of $z_{\overline{X}}$ and $p(\overline{X} \geq 87.5)$

1. Obtain all values necessary to use formula 9.1: $\overline{X} = 87.5$, $\mu = 84.75$, $\sigma = \sqrt{40} = 6.32$, and $N = 16$.
2. Put these values in formula 9.1 and compute $z_{\overline{X}}$:

$$z_{\overline{X}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}} = \frac{87.5 - 84.75}{6.32/\sqrt{16}}$$
$$= \frac{2.75}{6.32/4} = \frac{2.75}{1.58}$$
$$= 1.74$$

3. Refer $z_{\overline{X}}$ to Table C, restating the probability question about $\overline{X}$ into a probability question about $z_{\overline{X}}$:

$$p(\overline{X} \geq 87.5) = p(z_{\overline{X}} \geq 1.74) = .0409 \qquad \text{from Table C}$$

As a second example, use the teaching method example in Table 9.1. Suppose that previous research told us that $\mu = 84.75$, $\sigma^2 = 40$, and the population is normal. Given these values of μ and σ^2, we can compute $z_{\overline{X}}$ for $\overline{X} = 87.5$ and find the probability of getting an $\overline{X}$ this large or larger. Table 9.3 shows these computations. Figure 9.3 shows the area in a normal distribution for $\overline{X}$'s larger than 87.5 (and for $z_{\overline{X}}$ larger than 1.74).

These computations are based on the knowledge of μ and σ^2 and on the assumption that the raw scores X are normally distributed in the population.

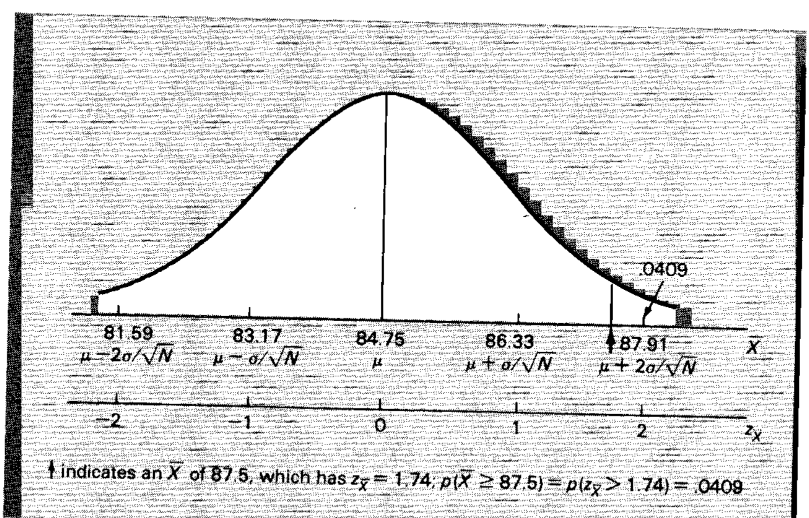

FIGURE 9.3
Sampling distribution of $\overline{X}$ and $z_{\overline{X}}$ ($\mu = 84.75$, $\sigma = 6.32$): $N = 16$, $\sigma/\sqrt{N} = 1.58$, $\overline{X} = 87.5$, $z_{\overline{X}} = 1.74$.

For practice, find $p(\overline{X} \geq 89)$, using the values for μ and σ^2 given in the second example. You should get

$$p(\overline{X} \geq 89) = p(z_{\overline{X}} \geq 2.69) = .0036$$

As N Increases, σ^2/N Decreases

Let's return to the fact that σ^2/N, the variance of the sampling distribution of $\overline{X}$, decreases as N increases. We can now illustrate this fact by computing the variance for $\overline{X}$ for various sized samples. Table 9.4 contains such computations for sample sizes of 3, 5, and 15 and compares the values of σ^2/N for these samples to the variance of the population.

As N increased from 3 to 15, σ^2/N decreased from 75 to 15, when the variance of the population was $\sigma^2 = 225$. As sample size increases, it becomes more unlikely that we will obtain a sample which consists of nothing but extreme values; so the values of $\overline{X}$ tend to cluster more closely around μ, which gives less variability. Figure 9.4 shows the distribution for the population and the sampling distributions of $\overline{X}$ for $N = 3, 5$, and 15. For most statistics, the variability of the statistic decreases as N increases, as we have seen here for $\overline{X}$.

9.7 MEANS OF SAMPLING DISTRIBUTIONS: UNBIASED ESTIMATES

Estimation
Calculation of an approximate value of a parameter

Point estimate
Using a statistic as a single value (point) to estimate a parameter

Estimation is related to sampling distributions. Remember that **estimation** is the calculation of an approximate value of a parameter. If we want some idea of the value of the population mean, we can calculate the sample mean as an estimate. When we use the sample mean as an estimate of the population mean, $\overline{X}$ is called a **point estimate**. We realize that the sample mean is not equal to the population mean, but it is an available indicator of the value of μ, so we

TABLE 9.4
Computation of $\sigma^2_{\overline{X}} = \sigma^2/N$ for Various Sized Samples from a Normal Population with $\mu = 100$ and $\sigma^2 = 225$

	Variance
Single score	$\sigma^2 = 225$
$N = 3$	$\sigma^2/N = 225/3 = 75$
$N = 5$	$\sigma^2/N = 225/5 = 45$
$N = 15$	$\sigma^2/N = 225/15 = 15$

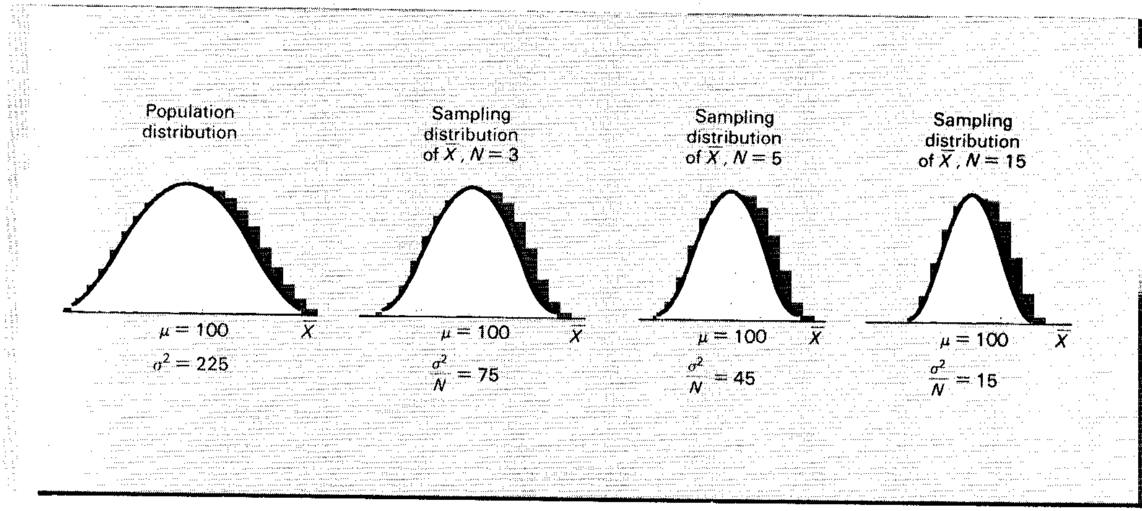

FIGURE 9.4
Distribution for X and $\bar{X}$.

use $\bar{X}$ as our estimate. When we do not have a population parameter, we can estimate it with the corresponding sample statistic.[2]

Parameters and Estimates

Table 9.5 lists population parameters that *could* be estimated by statistics which we have covered so far in the text. Your might be surprised by some of the combinations, such as the sample median as an estimate of the population mean. Since any statistic *can* be used to estimate any population parameter, you need to know about the quality of the estimate, which statistics are good estimates. For example, is X_{50} a good estimate of μ? The ultimate question concerns the quality of the statistic as an estimate of a given parameter: Is the statistic a good estimate?[3]

---

[2]Another type of estimation is called *interval estimation.* The goal of interval estimation is to obtain an interval of potential values for a parameter. Rather than giving just $\bar{X} = 87.5$ as our best estimate of μ, we could calculate an interval of potential values of μ. Interval estimation is covered in Section 12.5.

[3]Pay attention to the last two entries in the right pair of columns of Table 9.5. These entries give the variance and standard error of $\bar{X}$ and their estimates. Since N is a constant, only σ^2 or σ needs to be estimated by a corresponding statistic, $\overset{*}{s}{}^2$ or $\overset{*}{s}$, and combined with N or $\sqrt{N}$ to give the estimates. Any of the statistics using $\overset{*}{s}{}^2$ or $\overset{*}{s}$ have a problem in that they are not good estimates, as we will see shortly.

TABLE 9.5
Parameters and Statistics as Point Estimates

Parameter	Statistic	Parameter	Statistic
Population mode	Mode	Population range	Range
Population median	X_{50}	σ^2	$\overset{*}{s}{}^2$
μ	$\overline{X}$	σ	$\overset{*}{s}$
μ	X_{50}	$\sigma^2_{Y \cdot X}$	$s^2_{Y \cdot X}$
μ	Mode	$\sigma^2_{\overline{X}} = \sigma^2/N$	$\overset{*}{s}{}^2/N$
ρ	r	$\sigma_{\overline{X}} = \sigma/\sqrt{N}$	$\overset{*}{s}/\sqrt{N}$

Unbiased

Unfortunately, there are many definitions of *good* statistics as estimates. Let's concentrate on two definitions, one in this section and one in the next. The first definition of a *good estimate* is one which is unbiased:

A statistic is an *unbiased* estimate of a parameter if the mean of the sampling distribution of the statistic is equal to the parameter.

Unbiased
A statistic is an unbiased estimate of a parameter if the mean of its sampling distribution is equal to the parameter

Although we can make no statement about equality of the parameter and the statistic itself, **unbiased** tells us that the parameter and the mean of the sampling distribution are equal. Another way of saying this is that *on the average* the statistic will equal the parameter. This *does not* say that the statistic equals the parameter; rather, the mean of the statistic's sampling distribution equals the parameter.

$\overline{X}$ Is Unbiased

The sample mean $\overline{X}$ is an unbiased estimate of the population mean μ. Check this last statement against the definition of *unbiased* given above: Is the mean of the sampling distribution of the statistic $\overline{X}$ equal to the parameter μ? Yes, and this means that even though we cannot be guaranteed that our single sampled value of $\overline{X}$ equals μ, we know that $\overline{X}$ is a statistic whose average value equals μ. If we were to repeatedly draw samples of size N from the same population, some of the $\overline{X}$'s would be too high and others would be too low, but the average value of the $\overline{X}$'s would be right on target. *Unbiased* means that on the average, $\overline{X}$ is free from any systematic tendencies to be larger or smaller than μ. One important additional comment is that $\overline{X}$ is an unbiased estimate of μ regardless of the shape of the population of raw scores. No restrictions are put on the unbiasedness of $\overline{X}$ as an estimate of μ.

The property of unbiasedness helps us to correctly choose $\overline{X}$ as the best estimate of μ among those available. The sample median is unbiased as an estimate of μ only if the population is symmetric. If the population is not symmetric, then the mean of the sampling distribution of X_{50} is not equal to μ. Because of this, if you had reason to believe that your population was not symmetric and wanted your statistic to be an unbiased estimate of μ, you would choose $\overline{X}$ over X_{50}. Little is known about the mode as an estimate, but speculation would lead us to believe that the mode also would not be unbiased if the population were not symmetric. Consideration of the property of unbiasedness would lead us to choose $\overline{X}$ to estimate μ, which fits with a commonsense notion that the sample mean should be the best estimate of the population mean.

Biased and Unbiased Sample Variance

Unfortunately, common sense is not always correct, and we realize this when we look at the sample variance $\overset{*}{s}{}^2$ as an estimate of σ^2. We have used the sample variance $\overset{*}{s}{}^2$ for descriptive purposes as a measure of the variability of the sample. But $\overset{*}{s}{}^2$ has a serious drawback as an estimate of σ^2: $\overset{*}{s}{}^2$ is not unbiased. Remember, for a statistic to be an unbiased estimate of σ^2, the mean of the sampling distribution of the statistic must equal σ^2. The mean of the sampling distribution of $\overset{*}{s}{}^2$ is

$$\text{Mean of sampling distribution of } \overset{*}{s}{}^2 = \frac{\sigma^2(N - 1)}{N} \tag{9.2}$$

which is not exactly σ^2. On the average, $\overset{*}{s}{}^2$ is too small as an estimate of σ^2. This is not to say that every value of $\overset{*}{s}{}^2$ is smaller than σ^2, but that the average of all possible values of $\overset{*}{s}{}^2$ is smaller than σ^2. For example, if $N = 10$, then the mean of the sampling distribution of $\overset{*}{s}{}^2$ is $0.9\sigma^2$, so the bias in $\overset{*}{s}{}^2$ is that $\overset{*}{s}{}^2$ is too small on the average.

Reconsider the sampling experiment in Chapter Eight. We sampled from a population which had only three scores (1, 2, and 3), each with a large equal frequency. We took all possible samples of $N = 2$. Now we want to compute the sample variance $\overset{*}{s}{}^2$ for each sample and arrange the $\overset{*}{s}{}^2$ in a distribution. Figure 9.5 contains the possible samples, the sample variances $\overset{*}{s}{}^2$, and the sampling distribution of $\overset{*}{s}{}^2$. The mean of the sampling distribution of $\overset{*}{s}{}^2$ for $N = 2$ from this small population is .33, which is $\sigma^2(N - 1)/N = (0.67)(1)/2$. Thus $\overset{*}{s}{}^2$ is too small on the average. As a sidelight, note that the sampling distribution of $\overset{*}{s}{}^2$ is positively skewed. See Box 9.1 for an explanation of *why* $\overset{*}{s}{}^2$ is too small on the average.

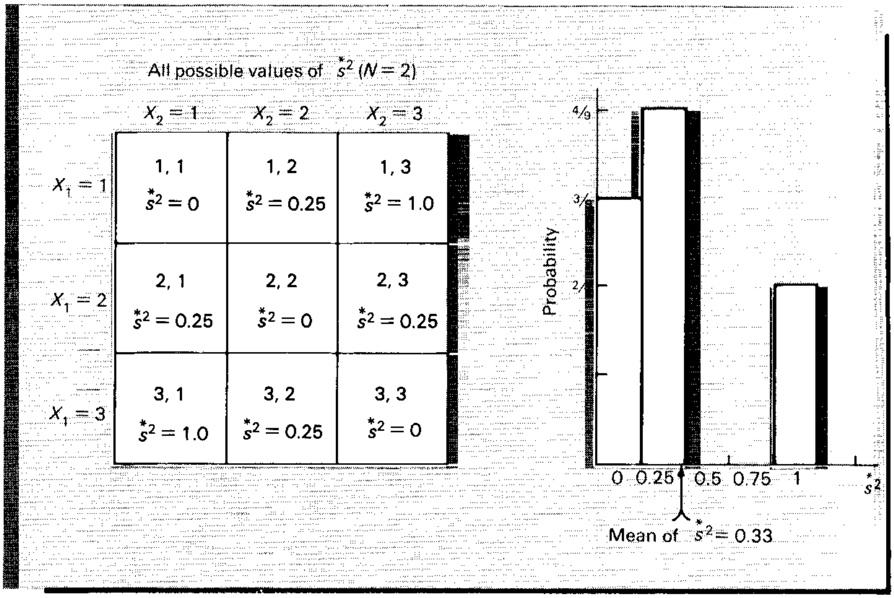

FIGURE 9.5 Sampling distribution of $\overset{*}{s}^2$ from population in Figure 9.2, $\sigma^2 = 2/3 = .67$.

We correct the tendency to underestimate σ^2 by dividing $\Sigma(X - \overline{X})^2$ by $N - 1$ instead of N. This gives a new measure of sample variability,[4] the **unbiased sample variance** s^2.

Unbiased sample variance
s^2 is a measure of variability in the sample which is formed by dividing the sum of squares by $N - 1$; the mean of the sampling distribution of s^2 is σ^2

$$s^2 = \frac{\Sigma(X - \overline{X})^2}{N - 1} \tag{9.3}$$

We take formula 9.3 as the definitional formula for s^2 and use the following formula for computational purposes:

$$s^2 = \frac{N\Sigma X^2 - (\Sigma X)^2}{N(N - 1)} \tag{9.4}$$

We now have s^2 as an unbiased estimate of σ^2. On the average, use of s^2 as an estimate of σ^2 will be correct. However, since s^2 is a statistic and has variability, we cannot be certain how close any one sample statistic s^2 will be to σ^2, but we can be confident that the mean of the sampling distribution of s^2 is exactly equal to σ^2. Figure 9.6 shows the sampling distribution of all possible values of s^2 for the sampling experiment from the sample population

[4]One of my students says that the way he remembers s^2 as unbiased and $\overset{*}{s}^2$ as biased is that s^2 is "unstarred and unbiased."

BOX 9.1 THE BIAS IN $\overset{*}{s}^2$

As an estimate of the population variance σ^2, the sample variance $\overset{*}{s}^2$ is too small on the average. Why? To see that the bias in $\overset{*}{s}^2$ is toward $\overset{*}{s}^2$ being too small, consider the "ideal" estimate of population variance given as

$$\frac{\Sigma(X - \mu)^2}{N}$$

This estimate of σ^2 has no bias; that is, the average of the sampling distribution of $\Sigma(X - \mu)^2/N$ is σ^2. Unfortunately, we never know the value of μ, so we cannot compute $\Sigma(X - \mu)^2/N$. Using $\overline{X}$ as an estimate of μ gives $\overset{*}{s}^2 = \Sigma(X - \overline{X})^2/N$. Remember that $\overline{X}$ minimizes the sum of squared deviations. That is, $\Sigma(X - \overline{X})^2$ is as small as it can be and smaller than $\Sigma(X - \mu)^2$. By this least-squares property of $\overline{X}$, when we substitute $\overline{X}$ for μ in $\Sigma(X - \mu)^2/N$, we make it too small. The exact amount by which it is too small is shown when we consider the mean of the numerator of the ideal estimate. The mean of the sampling distribution of $\Sigma(X - \mu)^2$ can be shown to be

$$\text{Mean of sampling distribution of } \Sigma(X - \mu)^2 = N\sigma^2$$

When we substitute $\overline{X}$ for μ, we get

$$\text{Mean of sampling distribution of } \Sigma(X - \overline{X})^2 = N\sigma^2 - \sigma^2$$

That is, the numerator is too small by σ^2. Rewriting $N\sigma^2 - \sigma^2$ as $\sigma^2(N - 1)$ gives the solution to the problem:

$$\text{Mean of sampling distribution of } \Sigma(X - \overline{X})^2 = \sigma^2(N - 1)$$

Since $\Sigma(X - \overline{X})^2$ has an average value of $\sigma^2(N - 1)$, all we have to do is divide this numerator by $N - 1$ instead of N, and we have an unbiased estimate of σ^2, or

$$\text{Mean of sampling distribution of } \frac{\Sigma(X - \overline{X})^2}{N - 1} = \sigma^2$$

We call $s^2 = \Sigma(X - \overline{X})^2/(N - 1)$ the *unbiased sample variance.*

Another way of looking at the problem is to correct the bias in $\overset{*}{s}^2$. All we need to do is to multiply $\overset{*}{s}^2$ by $N/(N - 1)$, giving

$$s^2 = (\overset{*}{s}^2)\frac{N}{N - 1} = \frac{\Sigma(X - \overline{X})^2}{N}\frac{N}{N - 1} = \frac{\Sigma(X - \overline{X})^2}{N - 1}$$

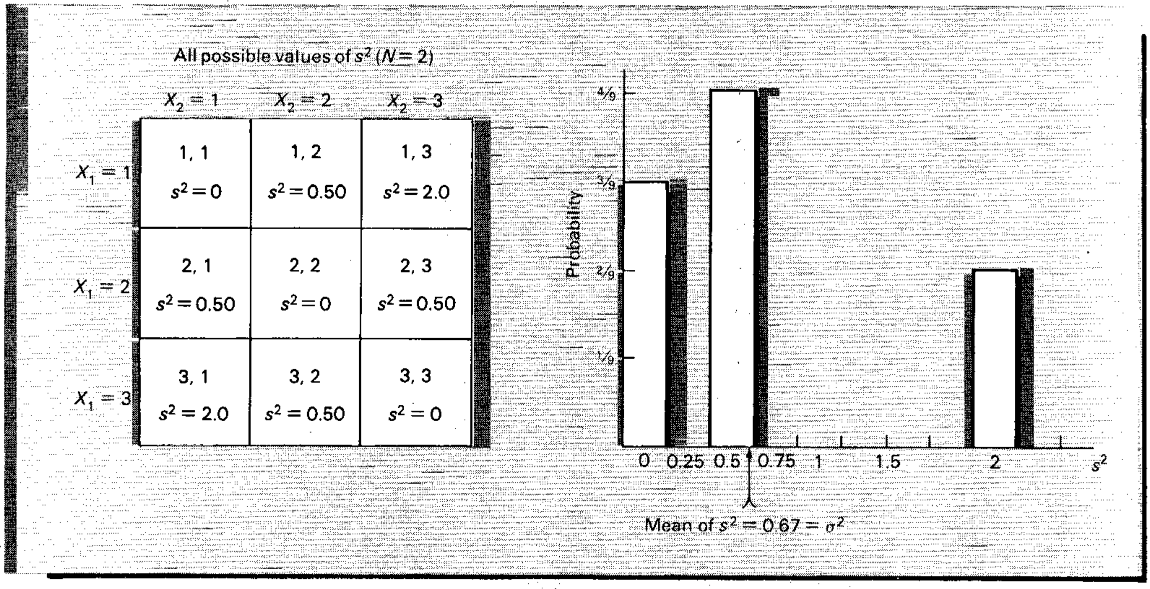

FIGURE 9.6
Sampling distribution of s^2 from population in Figure 9.2, $\sigma^2 = 2/3 = .67$.

with scores of 1, 2, and 3. The mean of the sampling distribution of s^2 is exactly $\sigma^2 = .67$ because s^2 is unbiased.

Computation of s^2 is very similar to computation of $\overset{*}{s}^2$, except for the denominator. Table 9.6 shows the computation of s^2 for the data of Table 9.1. Use of formula 9.4 gives $s^2 = 38.13$, which corrects for $\overset{*}{s}^2 = 35.75$ being too small. The following is important:

The biased sample variance $\overset{*}{s}^2$ is not used further in this text, but the unbiased sample variance s^2 is used whenever an estimate of σ^2 is needed.

TABLE 9.6
Computation of s^2 for Data of Table 9.1

$$
\begin{aligned}
s^2 &= \frac{N\Sigma X^2 - (\Sigma X)^2}{N(N-1)} \\
&= \frac{(16)(123{,}072) - 1400^2}{(16)(16-1)} = \frac{9152}{(16)(15)} = \frac{9152}{240} \\
&= 38.13333 = 38.13 \\
\overset{*}{s}^2 &= 35.75
\end{aligned}
$$

TABLE 9.7
Unbiased Estimates of Parameters

Parameter	Statistic	Parameter	Statistic
μ	$\overline{X}$	σ^2	s^2
ρ	r (if $\rho = 0$)	σ^2/N	s^2/N
$\sigma^2_{Y \cdot X}$	$s^2_{Y \cdot X}$		

To form an unbiased estimate of σ^2/N, we substitute s^2 for σ^2, yielding s^2/N. Whenever an estimate of σ^2/N is needed, s^2/N is used. Note that the general principle in finding an unbiased estimate of a parameter is to find a statistic whose sampling distribution has its mean equal to the parameter. Table 9.7 shows parameters and their *unbiased* estimates.

Biased Standard Deviation

Biased standard deviation
The square root of s^2 gives the standard deviation s, which is biased

Neither $\overset{*}{s}$ nor s is an unbiased estimate of σ. The reason is that the property of unbiasedness is not retained for a nonlinear transformation such as square root. So $s = \sqrt{s^2}$ is not an unbiased estimate of σ. However, the bias is appreciable only for small samples, and as N increases, the bias decreases. For $N = 2$, the mean of the sampling distribution of s is 0.798 σ, and for $N = 10$ it is 0.973 σ.[5] For subsequent statistics which use s as an estimate of σ, the degree of bias is taken into account in the statistic and in its sampling distribution.

The property of unbiasedness is a good place to emphasize Polya's first point in *How to Solve It,*[6] understanding the problem. There, he asks, "What is the unknown?" In understanding the concept of unbiased estimators, the key is the mean of the sampling distribution of a statistic. The mean of the sampling distribution is the unknown. If we find that the mean of a statistic's sampling distribution is the parameter we desire to estimate, then the statistic is unbiased. For example, in estimating μ with $\overline{X}$, the mean of $\overline{X}$'s sampling distribution is μ, so $\overline{X}$ is unbiased. As we conclude this section on the property of unbiasedness, check that you know the unknown for each of these statistics in Table 9.8. For each parameter-statistic pair, give the mean of the statistic's sampling distribution and answer yes or no to the question, Is the statistic unbiased? Cover up the answers in the right of the table so you can use them to check your answers (all of which are available in this section). If you did not get all the answers correct, review the information in this section. Any

---

[5]See Marascuilo and McSweeney, 1977, p. 81.

[6]See Section 1.7.

TABLE 9.8
Quiz on Unbiased Estimates

Desired Parameter	Statistic	Mean of Statistic's Sampling Distribution	Unbiased?
μ	$\overline{X}$	μ	Yes
μ	s^2	σ^2	No
σ^2	$\overset{*}{s}{}^2$	$\frac{\sigma^2(N-1)}{N}$	No
σ^2	s^2	σ^2	Yes
σ	$s\ (N = 10)$	$0.973\ \sigma$	No

statistic is an unbiased estimate of a desired parameter if the sampling distribution of the statistic has its mean equal to the parameter.

This section has shown us the importance of sampling distributions in connection with a part of inferential statistics, estimation. The mean of the sampling distribution of a statistic tells us whether a statistic is unbiased as an estimate of a parameter. Sampling distributions continue to play a crucial role in inferential statistics as we examine the variance of sampling distributions for another property of statistics as estimates. The next section on variance of sampling distributions is optional, but Section 9.9 on the shape of sampling distributions is *not* optional. Be sure you don't miss it.

9.8 VARIABILITY OF SAMPLING DISTRIBUTIONS: EFFICIENCY*

Efficiency Defined

As we have mentioned, the mean and variance of sampling distributions are important to inferential statistics. We have seen how the variance of the sampling distribution of $\overline{X}$ plays an important part in forming $z_{\overline{X}}$, as shown in formula 9.1. Then $z_{\overline{X}}$ is used in computing probabilities for $\overline{X}$ to make decisions about μ. For most descriptive statistics, the variance of the sampling distribution does not play such a direct role in the formation of the statistic used in the decision making. However, variability of sampling distributions is important when we consider these statistics as estimates of parameters. Generally speaking, the less variable the statistic, the better it is as an estimate. For example, which $\overline{X}$ do you think is a better estimate of μ, $\overline{X}$ from a sample of size $N = 2$ or $\overline{X}$ from a sample of size $N = 100$? Both $\overline{X}$'s are unbiased, but the $\overline{X}$ from the larger sample is less variable ($\sigma^2/N = \sigma^2/100$) than the $\overline{X}$ from the smaller sample ($\sigma^2/N = \sigma^2/2$). This property is labeled *efficiency*, and the $\overline{X}$ from $N = 100$ is more efficient than $\overline{X}$ from $N = 2$.

Relative Efficiency and Examples

Efficient
A statistic is efficient if the variance of its sampling distribution is small

Efficiency as a property of estimates is a direct function of the variance of the sampling distribution of the statistics. If the variance of the sampling distribution of a statistic is small, the statistic is **efficient.** In practice, efficiency is a relative property since we compare the efficiency of two different estimates. So if we have two statistics which could be chosen as an estimate of a parameter, we want to compare the variance of their sampling distributions and choose the statistic with the smallest variance.

For example, for normal populations, both $\overline{X}$ and X_{50} are unbiased estimates of μ. Since we cannot distinguish between these two statistics for this situation on the basis of unbiasedness, we turn to the property of efficiency. The variance of the sampling distribution of the sample mean is less than that of the sample median, so $\overline{X}$ is more efficient than X_{50} as an estimate of μ. For normal populations we would choose $\overline{X}$ as the best estimate of μ because it is both unbiased and efficient.

When considering measures of variability, we would not use the range because it is more variable than s^2 or s. Other words used to describe efficiency are *stability* and *reliability*, so the sample range is said to be inefficient, unstable, and unreliable.

Having discussed the mean and variance of sampling distributions, we now turn to the shape of the sampling distribution of $\overline{X}$ and the Central Limit Theorem.

9.9 SHAPE OF SAMPLING DISTRIBUTIONS: CENTRAL LIMIT THEOREM

In Chapter One we discussed theoretical reference distributions used to approximate sampling distributions of statistics. We mentioned that we rely on certain theorems and the results of research that tell us the shape of any given sampling distribution, which can often be closely approximated by a known theoretical distribution. This section covers the shape of the sampling distribution of $\overline{X}$. Shapes of other sampling distributions are discussed throughout the remainder of the text.

Central Limit Theorem Defined

Under the assumption that the population is normally distributed, we know that the sampling distribution of $\overline{X}$ is also normally distributed. However, most researchers do not know the actual shape of the sampled population; or if they do know the shape, it is not normal. So most researchers would not be able to assert that the population sampled was normal. This position represents a serious dilemma since what we know so far indicates that $\overline{X}$ is normally distributed only if X is normally distributed, and we need to be able to compute probabilities for $\overline{X}$. Remember that we want to use $\overline{X}$ to make decisions about

Central Limit Theorem
The theorem which states that the sampling distribution of $\bar{X}$ approaches a normal shape as N approaches infinity

μ, and we need probabilities from the sampling distribution of $\overline{X}$ to make these decisions. The researcher who is in this position is rescued from the dilemma by one of the most amazing theorems in all mathematics, the **Central Limit Theorem:**

If the researcher is sampling independently from a population which has mean μ and variance σ^2, then as sample size N approaches infinity, the sampling distribution of the sample mean $\bar{X}$ approaches normality without regard to the shape of the sampled population.

Central Limit Theorem Applied: Only $\bar{X}$

The Central Limit Theorem says that as N approaches infinity, the distribution of $\overline{X}$ will approach a normal distribution with mean μ and variance σ^2/N. Another way to state this theorem is that the larger the sample size, the more closely the sampling distribution of $\overline{X}$ will be approximated by the normal distribution. At this point, many students say, "The theorem sounds great, but we don't have infinitely many rats for our experiment. How does the theorem help for realistic sample sizes?" As long as the population is not very unusual in shape, the approximation is quite good for even moderate sample sizes such as those used by behavioral scientists. For samples of even 25 or 30 from most populations, the sampling distribution of $\overline{X}$ approaches a normal shape. Figure 9.7 shows several populations and the sampling distribution of $\overline{X}$ for $N = 2$ and $N = 30$.

The sampling distribution of $\overline{X}$ approaches normality fairly rapidly as N increases. So for many applied situations, we can proceed to use the normal distribution to compute probabilities for the sampling distribution of $\overline{X}$, even though we cannot say that the population is normally distributed. Please note that the Central Limit Theorem applies *only* to $\overline{X}$. Other statistics have other theorems which tell us the shape of their sampling distributions, but the Central Limit Theorem does not apply to them.

9.10 SUMMARY AND COMPUTATION

Summary

The important concept of sampling distributions includes three general types of distributions. We randomly sample scores from a population distribution to get our sample distribution. The sampling distribution of any statistic is as if it had been formed by repeatedly drawing such samples of a given size, computing the statistic, and arranging the statistics in a distribution. Decision making about parameters involves the use of the probability of the statistic from the sampling distribution.

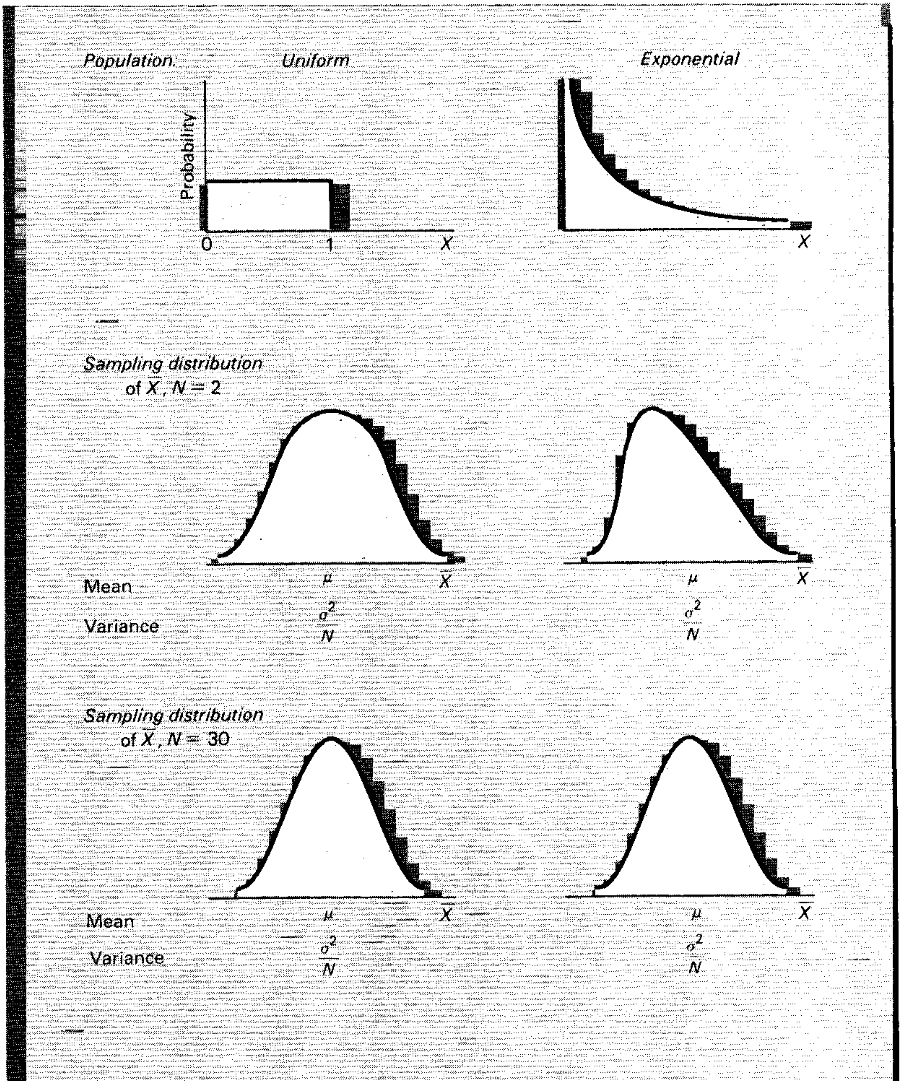

FIGURE 9.7
Central Limit Theorem effect for small samples: sampling distribution of $\bar{X}$.

If the X's are independent and X is distributed normally with mean of μ and variance of σ^2, then $\bar{X}$ is distributed normally with mean of μ and variance of σ^2/N. Probabilities for $\bar{X}$ are obtained from the standard normal distribution by transforming $\bar{X}$ to $z_{\bar{X}}$.

An unbiased estimate has a sampling distribution with its mean equal to the parameter estimated. Examples of unbiased estimators are $\bar{X}$ and s^2. An efficient estimate has a sampling distribution with small variance, and for a normal population $\bar{X}$ and s^2 are efficient relative to other estimates of μ and

σ^2. The Central Limit Theorem lets us use a normal distribution to approximate the distribution of $\overline{X}$ even when the population shape is unknown.

Key terms introduced in this chapter are

Population distribution	Standard error of the mean
Sample distribution	Unbiased
Sampling distribution	Relatively efficient*
σ^2/N	Central Limit Theorem

Computation

Since much of this chapter is conceptually based, there is very little computation, none of which is new. We computed $z_{\overline{X}}$ which used the basic operations learned in Chapter Five, so we have nothing new to add. None of the exercises which follow need to be done on the computer, so use them to help understand the important concepts covered above.

EXERCISES

Section(s)

9.1

1. A university researcher randomly samples 10 subjects from a list of 175 volunteers from introductory psychology. Is the population actually sampled "all students in introductory psychology this semester?" Why or why not?
2. How could judgment sampling justify the population stated in exercise 1? Could judgment sampling stretch the population to all college students in the United States enrolled in a similar course? Why or why not?

9.2 to 9.5

3. Without referring to Figure 9.1 or Table 6.2, compare the population distribution, the sample distribution, and the sampling distribution of $\overline{X}$ on the following:
   a. Random variable in the distribution
   b. Shape of the distribution
   c. Mean of the distribution
   d. Variance of the distribution
   e. Reality of obtaining the distribution
4. Describe each of the following populations in terms of large (or small), unobtainable (or obtainable), and hypothetical (or real):
   a. Whooping cranes
   b. All college students taking any statistics course in the next 3 yr
   c. All rats deprived of food for 24 h
   d. All professional basketball players

9.6 5. Suppose you computed the average IQ of the students in your statistics class and found it to be $\overline{X} = 115$. If IQ is normally distributed with mean of 100 and variance of 225, and there are 25 students in your class, what is the probability of getting 115 or larger for the value of $\overline{X}$?

6. If $N = 100$, how does the probability in exercise 5 change? For $N = 9$? If this population is not normal, how would the Central Limit Theorem influence our ability to do these computations?

7. Suppose you work in a car manufacturing plant where the average number of defects per car is 16 with a standard deviation of 3. If your quality control program samples four cars which have $\overline{X} = 18$ defects, what is the probability of getting an $\overline{X}$ greater than or equal to 18? State the assumptions you made in order to calculate this probability. Would you act to make changes in the assembly line?

8. If your sample in exercise 7 used nine cars which had $\overline{X} = 18$ defects, what is the probability of getting an $\overline{X}$ greater than or equal to 18? Would you act to make changes in the assembly line? Explain the difference between your answer to this question and exercise 7, given that $\overline{X} = 18$ in both cases.

9. Suppose a psychological clinic claims that its patients treated for anxiety are "normal" on this mood state after 4 weeks of therapy. The norms for a standard anxiety test are 25 for the mean and 5 for the standard deviation. If you randomly sample $N = 16$ patients and obtain an anxiety score for each, yielding $\overline{X} = 27.8$, what is the probability of getting an $\overline{X}$ this large or larger?

9.7 10. Is *unbiased* the only way to define a good estimate? Are all statistics unbiased? Do all statistics have a sampling distribution? Does the Central Limit Theorem apply to all statistics?

11. *True or False?* An unbiased statistic is equal to the parameter estimated. Defend your answer.