Example 4.22 illustrates for a very small population that we could in fact enumerate every possible sample of size 2 selected from the population and then compute all possible values of the sample mean. The next example will illustrate the properties of the sample mean, $\bar{y}$, when sampling from a larger population. This example will illustrate that the behavior of $\bar{y}$ as an estimator of $\mu$ depends on the sample size, $n$. Later in this chapter, we will illustrate the effect of the shape of the population distribution on the sampling distribution of $\bar{y}$.

**Example 4.23**

In this example, the population values are known and, hence, we can compute the exact values of the population mean, $\mu$, and population standard deviation, $\sigma$. We will then examine the behavior of $\bar{y}$ based on samples of size $n = 5, 10,$ and $25$ selected from the population. The population consists of $500$ pennies from which we compute the age of each penny: Age = 2008 – Date on penny. The histogram of the 500 ages is displayed in Figure 4.20(a). The shape is skewed to the right with a very long right tail. The mean and standard deviation are computed to be $\mu = 13.468$ years and $\sigma = 11.164$ years. In order to generate the sampling distribution of $\bar{y}$ for $n = 5$, we would need to generate all possible samples of size $n = 5$ and then compute the $\bar{y}$ from each of these samples. This would be an enormous task since there are $255,244,687,600$ possible samples of size 5 that could be selected from a population of 500 elements. The number of possible samples of size 10 or 25 is so large it makes even the national debt look small. Thus, we will use a computer program to select 25,000 samples of size 5 from the population of 500 pennies. For example, the first sample consists of pennies with ages 4, 12, 26, 16, and 9. The sample mean $\bar{y} = (4 + 12 + 26 + 16 + 9)/5 = 13.4$. We repeat 25,000 times the process of selecting 5 pennies, recording their ages, $y_1, y_2, y_3, y_4, y_5$, and then computing $\bar{y} = (y_1 + y_2 + y_3 + y_4 + y_5)/5$. The 25,000 values for $\bar{y}$ are then plotted in a frequency histogram, called the sampling distribution of $\bar{y}$ for $n = 5$. A similar procedure is followed for samples of size $n = 10$ and $n = 25$. The sampling distributions obtained are displayed in Figures 4.20(b)–(d).

Note that all three sampling distributions have nearly the same central value, approximately 13.5. (See Table 4.11.) The mean values of $\bar{y}$ for the three samples are nearly the same as the population mean, $\mu = 13.468$. In fact, if we had generated all possible samples for all three values of $n$, the mean of the possible values of $\bar{y}$ would agree exactly with $\mu$. 

![Figure 4.19](image-url)
Chapter 4  Probability and Probability Distributions

FIGURE 4.20

(a) Histogram of ages for 500 pennies

(b) Sampling distribution of $\bar{y}$ for $n = 5$

(c) Sampling distribution of $\bar{y}$ for $n = 10$

(d) Sampling distribution of $\bar{y}$ for $n = 25$
The next characteristic to notice about the three histograms is their shape. All three are somewhat symmetric in shape, achieving a nearly normal distribution shape when \( n = 25 \). However, the histogram for \( n = 5 \) is more spread out than the histogram based on \( n = 10 \), which, in turn, is more spread out than the histogram based on \( n = 25 \). When \( n \) is small, we are much more likely to obtain a value of \( y \) far from \( \mu \) than when \( n \) is larger. What causes this increased dispersion in the values of \( y \)? A single extreme \( y \), either large or small relative to \( \mu \), in the sample has a greater influence on the size of \( \bar{y} \) when \( n \) is small than when \( n \) is large. Thus, sample means based on small \( n \) are less accurate in their estimation of \( \mu \) than their large-sample counterparts.

Table 4.11 contains summary statistics for the sampling distribution of \( y \). The sampling distribution of \( \bar{y} \) has mean \( \mu_\bar{y} \) and standard deviation \( \sigma_\bar{y} \), which are related to the population mean, \( \mu \), and standard deviation, \( \sigma \), by the following relationship:

\[
\mu_\bar{y} = \mu \quad \sigma_\bar{y} = \frac{\sigma}{\sqrt{n}}
\]

From Table 4.11, we note that the three sampling deviations have means that are approximately equal to the population mean. Also, the three sampling deviations have standard deviations that are approximately equal to \( \sigma/\sqrt{n} \). If we had generated all possible values of \( y \), then the standard deviation of \( \bar{y} \) would equal \( \sigma/\sqrt{n} \) exactly. This quantity, \( \sigma_\bar{y} = \sigma/\sqrt{n} \), is called the **standard error of \( \bar{y} \)**.

The next characteristic to notice about the three histograms is their shape. All three are somewhat symmetric in shape, achieving a nearly normal distribution shape when \( n = 25 \). However, the histogram for \( n = 5 \) is more spread out than the histogram based on \( n = 10 \), which, in turn, is more spread out than the histogram based on \( n = 25 \). When \( n \) is small, we are much more likely to obtain a value of \( y \) far from \( \mu \) than when \( n \) is larger. What causes this increased dispersion in the values of \( y \)? A single extreme \( y \), either large or small relative to \( \mu \), in the sample has a greater influence on the size of \( \bar{y} \) when \( n \) is small than when \( n \) is large. Thus, sample means based on small \( n \) are less accurate in their estimation of \( \mu \) than their large-sample counterparts.

Table 4.11 contains summary statistics for the sampling distribution of \( y \). The sampling distribution of \( \bar{y} \) has mean \( \mu_\bar{y} \) and standard deviation \( \sigma_\bar{y} \), which are related to the population mean, \( \mu \), and standard deviation, \( \sigma \), by the following relationship:

\[
\mu_\bar{y} = \mu \quad \sigma_\bar{y} = \frac{\sigma}{\sqrt{n}}
\]

From Table 4.11, we note that the three sampling deviations have means that are approximately equal to the population mean. Also, the three sampling deviations have standard deviations that are approximately equal to \( \sigma/\sqrt{n} \). If we had generated all possible values of \( y \), then the standard deviation of \( \bar{y} \) would equal \( \sigma/\sqrt{n} \) exactly. This quantity, \( \sigma_\bar{y} = \sigma/\sqrt{n} \), is called the **standard error of \( \bar{y} \)**.

Quite a few of the more common sample statistics, such as the sample median and the sample standard deviation, have sampling distributions that are nearly normal for moderately sized values of \( n \). We can observe this behavior by computing the sample median and sample standard deviation from each of the three sets of 25,000 sample \( (n = 5, 10, 25) \) selected from the population of 500 pennies. The resulting sampling distributions are displayed in Figures 4.21(a)–(d), for the sample median, and Figures 4.22(a)–(d), for the sample standard deviation. The sampling distribution of both the median and the standard deviation are more highly skewed in comparison to the sampling distribution of the sample mean. In fact, the value of \( n \) at which the sampling distributions of the sample median and standard deviation have a nearly normal shape is much larger than the value required for the sample mean. A series of theorems in mathematical statistics called the **Central Limit Theorems** provide theoretical justification for our approximating the true sampling distribution of many sample statistics with the normal distribution. We will discuss one such theorem for the sample mean. Similar theorems exist for the sample median, sample standard deviation, and the sample proportion.
**FIGURE 4.21**

(a) Histogram of ages for 500 pennies

(b) Sampling distribution of median for $n = 5$

(c) Sampling distribution of median for $n = 10$

(d) Sampling distribution of median for $n = 25$
4.12 Sampling Distributions

**FIGURE 4.22**

(a) Histogram of ages for 500 pennies

(b) Sampling distribution of standard deviation for $n = 5$

(c) Sampling distribution of standard deviation for $n = 10$

(d) Sampling distribution of standard deviation for $n = 25$
Figure 4.20 illustrates the Central Limit Theorem. Figure 4.20(a) displays the distribution of the measurements \( y \) in the population from which the samples are to be drawn. No specific shape was required for these measurements for the Central Limit Theorem to be validated. Figures 4.20(b)–(d) illustrate the sampling distribution for the sample mean when \( n \) is 5, 10, and 25, respectively. We note that even for a very small sample size, \( n = 10 \), the shape of the sampling distribution of \( \bar{y} \) is very similar to that of a normal distribution. This is not true in general. If the population distribution had many extreme values or several modes, the sampling distribution of \( \bar{y} \) would require \( n \) to be considerably larger in order to achieve a symmetric bell shape.

We have seen that the sample size \( n \) has an effect on the shape of the sampling distribution of \( \bar{y} \). The shape of the distribution of the population measurements also will affect the shape of the sampling distribution of \( \bar{y} \). Figures 4.23 and 4.24 illustrate the effect of the population shape on the shape of the sampling distribution of \( \bar{y} \). In Figure 4.23, the population measurements have a normal distribution. The sampling distribution of \( \bar{y} \) is exactly a normal distribution for all values of \( n \), as is illustrated for \( n = 5, 10, \) and 25 in Figure 4.23. When the population distribution is nonnormal, as depicted in Figure 4.24, the sampling distribution of \( \bar{y} \) will not have a normal shape for small \( n \) (see Figure 4.24 with \( n = 5 \)). However, for \( n = 10 \) and 25, the sampling distributions are nearly normal in shape, as can be seen in Figure 4.24.