CHAPTER 3

NUMERICAL DESCRIPTIVE MEASURES

Prem Mann, *Introductory Statistics*, *7/E*

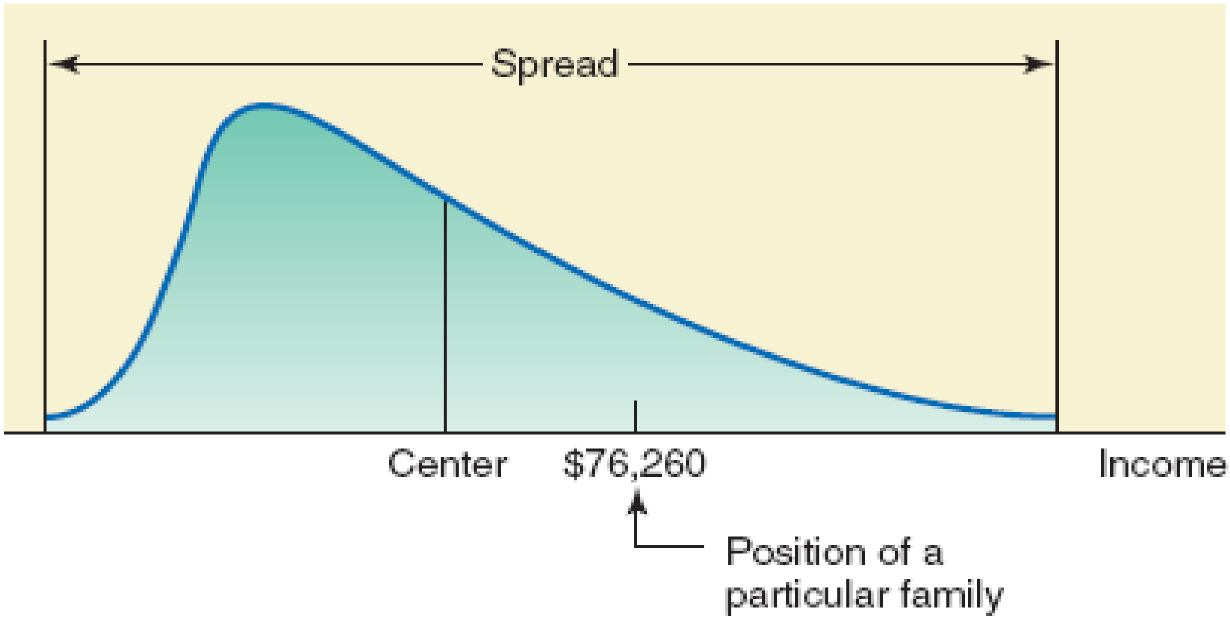
Opening Example

During the 2008 season, among all baseball teams, the New York Yankees drew the highest average of spectators to the games. (See Case Study 3-1). Despite baseball being "America's National Pastime," attendance at Major League Baseball games varies from team to team. What may attract fans to baseball fields? Is it the number of championships won, market size, fan loyalty, or just love for the team or game?

3.1 Measures of Central Tendency for Ungrouped Data

- Mean
- Median
- Mode
- Relationships among the Mean, Median, and Mode

Figure 3.1



Prem Mann, *Introductory Statistics*, *7/E*

Mean

The ***mean for ungrouped data*** is obtained by dividing the sum of all values by the number of values in the data set. Thus,

Mean for population data: $$\mu = \frac{\sum x}{N}$$

Mean for sample data: $$\bar{x} = \frac{\sum x}{n}$$

Example 3-1

Table 3.1 lists the total sales (rounded to billions of dollars) of six U.S. companies for 2008.

Table 3.1 2008 Sales of Six U.S. Companies

Company	Total Sales (billions of dollars)
General Motors	149
Wal-Mart Stores	406
General Electric	183
Citigroup	107
Exxon Mobil	426
Verizon Communication	97

Find the 2008 mean sales for these six companies.

Example 3-1: Solution

$$\sum x = x_1 + x_2 + x_3 + x_4 + x_5 + x_6$$

$$= 149 + 406 + 183 + 107 + 426 + 97 = 1368$$

$$\bar{x} = \frac{\sum x}{n} = \frac{1368}{6} = 228 = \$228 \text{ Billion}$$

Thus, the mean 2008 sales of these six companies was 228, or $228 billion.

Example 3-2

The following are the ages (in years) of all eight employees of a small company:

53 32 61 27 39 44 49 57

Find the mean age of these employees.

Example 3-2: Solution

$$\mu = \frac{\sum x}{N} = \frac{362}{8} = 45.25 \text{ years}$$

Thus, the mean age of all eight employees of this company is 45.25 years, or 45 years and 3 months.

Example 3-3

Table 3.2 lists the total philanthropic givings (in million dollars) by six companies during 2007.

Corporation	Money Given in 2007 (millions of dollars)
CVS	22.4
Best Buy	31.8
Staples	19.8
Walgreen	9.0
Lowe's	27.5
Wal-Mart	337.9

Example 3-3

Notice that the charitable contributions made by Wal-Mart are very large compared to those of other companies. Hence, it is an outlier. Show how the inclusion of this outlier affects the value of the mean.

Example 3-3: Solution

If we do not include the charitable givings of Wal-Mart (the outlier), the mean of the charitable contributions of the fiver companies is

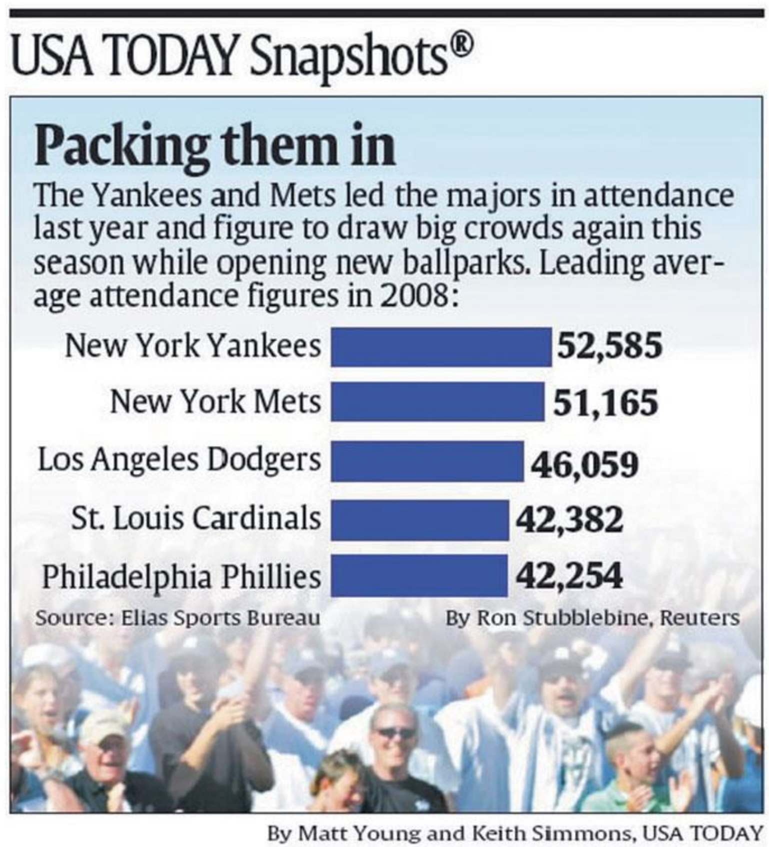
$$\textbf{Mean} = \frac{22.4 + 31.8 + 19.8 + 9.0 + 27.5}{5} = \textbf{\$22.1 million}$$

Example 3-3: Solution

Now, to see the impact of the outlier on the value of the mean, we include the contributions of Wal-Mart and find the mean contributions of the six companies. This mean is

$$\textbf{Mean} = \frac{22.4 + 31.8 + 19.8 + 9.0 + 27.5 + 337.9}{6} = \$74.73 \textbf{ million}$$

Case Study 3-1 Average Attendance at Baseball Games



Median

Definition

The ***median*** is the value of the middle term in a data set that has been ranked in increasing order.

The calculation of the median consists of the following two steps:

1. Rank the data set in increasing order.
2. Find the middle term. The value of this term is the median.

Example 3-4

The following data give the prices (in thousands of dollars) of seven houses selected from all houses sold last month in a city.

312 257 421 289 526 374 497

Find the median.

Example 3-4: Solution

First, we rank the given data in increasing order as follows:

257   289   312   374   421   497   526

Since there are seven homes in this data set and the middle term is the fourth term,

257   289   312 $\boxed{374}$ 421   497   526

↑

Median

Thus, the median price of a house is 374.

Example 3-5

Table 3.3 gives the 2008 profits (rounded to billions of dollars) of 12 companies selected from all over the world.

Table 3.3 Profits of 12 Companies for 2008

Find the median of these data.

Company	2008 Profits (billions of dollars)
Merck & Co	8
IBM	12
Unilever	7
Microsoft	17
Petrobras	14
Exxon Mobil	45
Lukoil	10
AT&T	13
Nestlé	17
Vodafone	13
Deutsche Bank	9
China Mobile	11

Example 3-5: Solution

First we rank the given profits as follows:

7 8 9 10 11 12 13 13 14 17 17 45

There are 12 values in this data set. Because there is an even number of values in the data set, the median is given by the average of the two middle values.

Example 3-5: Solution

The two middle values are the sixth and seventh in the foregoing list of data, and these two values are 12 and 13.

$$\textbf{Median} = \frac{\mathbf{12+13}}{\mathbf{2}} = \frac{\mathbf{25}}{\mathbf{2}} = \mathbf{12.5} = \mathbf{\$12.5\ billion}$$

Thus, the median profit of these 12 companies is $12.5 billion.

Median

The median gives the center of a histogram, with half the data values to the left of the median and half to the right of the median. The advantage of using the median as a measure of central tendency is that it is not influenced by outliers. Consequently, the median is preferred over the mean as a measure of central tendency for data sets that contain outliers.

Case Study 3-2 The Gender Pay Gap

Mode

Definition

The ***mode*** is the value that occurs with the highest frequency in a data set.

Example 3-6

The following data give the speeds (in miles per hour) of eight cars that were stopped on I-95 for speeding violations.

77   82   74   81   79   84   74   78

Find the mode.

Example 3-6: Solution

In this data set, 74 occurs twice and each of the remaining values occurs only once. Because 74 occurs with the highest frequency, it is the mode. Therefore,

Mode = **74 miles per hour**

Mode

- A major shortcoming of the mode is that a data set may have none or may have more than one mode, whereas it will have only one mean and only one median.
  - Unimodal: A data set with only one mode.
  - Bimodal: A data set with two modes.
  - Multimodal: A data set with more than two modes.

Example 3-7 (Data set with no mode)

Last year's incomes of five randomly selected families were \$76,150. \$95,750, \$124,985, \$87,490, and \$53,740. Find the mode.

Example 3-7: Solution

Because each value in this data set occurs only once, this data set contains **no mode**.

Example 3-8 (Data set with two modes)

Refer to the data on 2008 profits of 12 companies given in Table 3.3 of Example 3-5. Find the mode for these data.

Example 3-8: Solution

In the data given in Example 3-5, each of two values 13 and 17 occurs twice, and each of the remaining values occurs only once. Therefore, that data set has two modes: **$13 billion** and **$ 17 billion**.

Example 3-9 (Data set with three modes)

The ages of 10 randomly selected students from a class are 21, 19, 27, 22, 29, 19, 25, 21, 22 and 30 years, respectively. Find the mode.

Example 3-9: Solution

This data set has three modes: **19**, **21** and **22**. Each of these three values occurs with a (highest) frequency of 2.

Mode

One advantage of the mode is that it can be calculated for both kinds of data - quantitative and qualitative - whereas the mean and median can be calculated for only quantitative data.

Example 3-10

The status of five students who are members of the student senate at a college are senior, sophomore, senior, junior, senior. Find the mode.

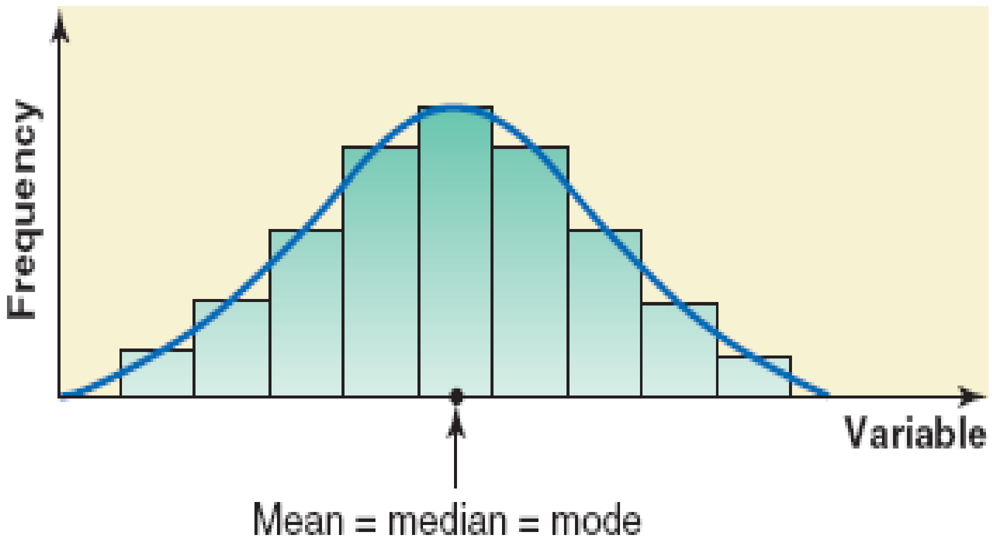
Example 3-10: Solution

Because **senior** occurs more frequently than the other categories, it is the mode for this data set. We cannot calculate the mean and median for this data set.

Relationships among the Mean, Median, and Mode

1. For a symmetric histogram and frequency curve with one peak (Figure 3.2), the values of the mean, median, and mode are identical, and they lie at the center of the distribution.

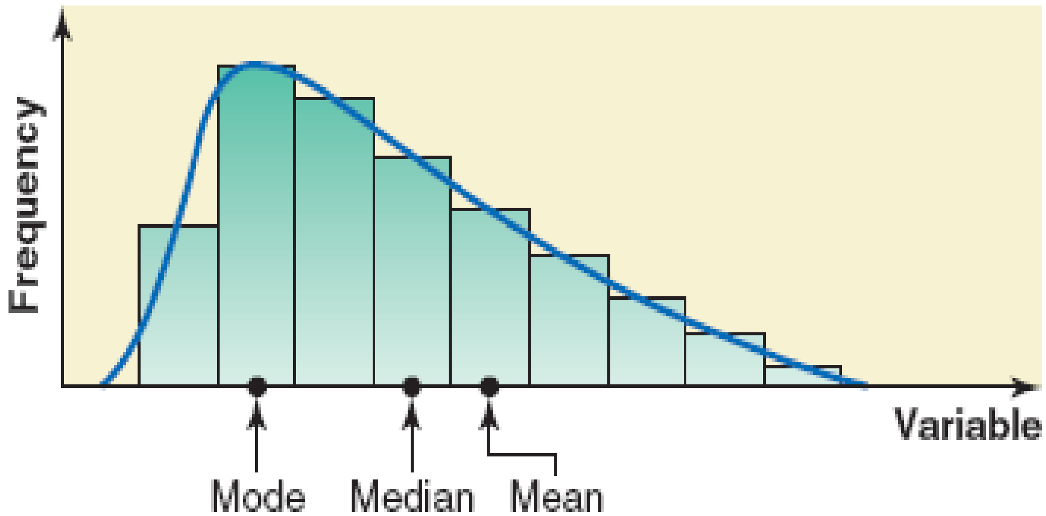
Figure 3.2 Mean, median, and mode for a symmetric histogram and frequency curve.



Relationships among the Mean, Median, and Mode

2. For a histogram and a frequency curve skewed to the right (Figure 3.3), the value of the mean is the largest, that of the mode is the smallest, and the value of the median lies between these two. (Notice that the mode always occurs at the peak point.) The value of the mean is the largest in this case because it is sensitive to outliers that occur in the right tail. These outliers pull the mean to the right.

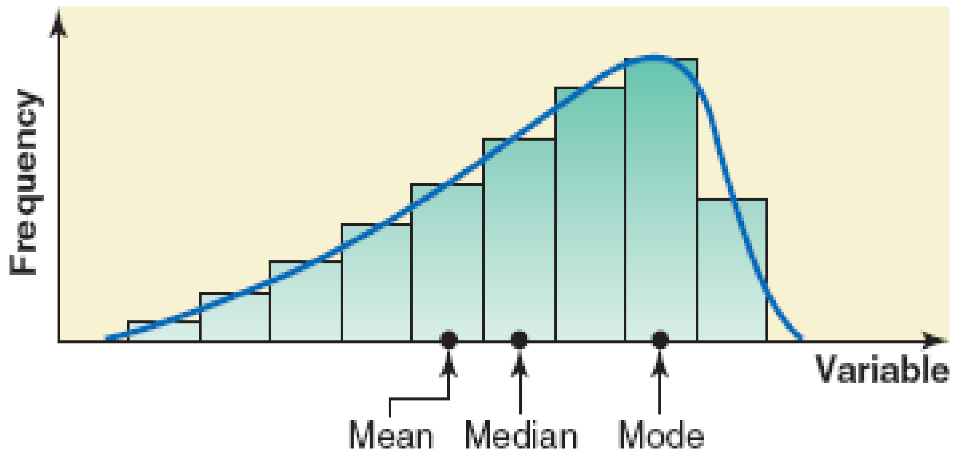
Figure 3.3 Mean, median, and mode for a histogram and frequency curve skewed to the right.



Relationships among the Mean, Median, and Mode

3. If a histogram and a distribution curve are skewed to the left (Figure 3.4), the value of the mean is the smallest and that of the mode is the largest, with the value of the median lying between these two. In this case, the outliers in the left tail pull the mean to the left.

Figure 3.4 Mean, median, and mode for a histogram and frequency curve skewed to the right.



6.3 THE CENTRAL LIMIT THEOREM

In this section, we will consider one of the most remarkable results in probability — namely, the *central limit theorem*. Loosely speaking, this theorem asserts that the sum of a large number of independent random variables has a distribution that is approximately normal. Hence, it not only provides a simple method for computing approximate probabilities for sums of independent random variables, but it also helps explain the remarkable fact that the empirical frequencies of so many natural populations exhibit a bell-shaped (that is, a normal) curve.

In its simplest form, the central limit theorem is as follows:

Theorem 6.3.1 The Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables each having mean μ and variance σ^2. Then for n large, the distribution of

$$X_1 + \cdots + X_n$$

is approximately normal with mean $n\mu$ and variance $n\sigma^2$.

It follows from the central limit theorem that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

is approximately a standard normal random variable; thus, for n large,

$$P\left\{\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < x\right\} \approx P\{Z < x\}$$

where Z is a standard normal random variable.

EXAMPLE 6.3a An insurance company has 25,000 automobile policy holders. If the yearly claim of a policy holder is a random variable with mean 320 and standard deviation 540, approximate the probability that the total yearly claim exceeds 8.3 million.

SOLUTION Let X denote the total yearly claim. Number the policy holders, and let X_i denote the yearly claim of policy holder i. With $n = 25{,}000$, we have from the central limit theorem that $X = \sum_{i=1}^{n} X_i$ will have approximately a normal distribution with mean $320 \times 25{,}000 = 8 \times 10^6$ and standard deviation $540\sqrt{25{,}000} = 8.5381 \times 10^4$. Therefore,

$$\begin{aligned}P\{X > 8.3 \times 10^6\} &= P\left\{\frac{X - 8 \times 10^6}{8.5381 \times 10^4} > \frac{8.3 \times 10^6 - 8 \times 10^6}{8.5381 \times 10^4}\right\} \\ &= P\left\{\frac{X - 8 \times 10^6}{8.5381 \times 10^4} > \frac{.3 \times 10^6}{8.5381 \times 10^4}\right\}\end{aligned}$$

$$\approx P\{Z > 3.51\} \qquad \text{where } Z \text{ is a standard normal}$$
$$\approx .00023$$

Thus, there are only 2.3 chances out of 10,000 that the total yearly claim will exceed 8.3 million. ■

EXAMPLE 6.3b Civil engineers believe that W, the amount of weight (in units of 1,000 pounds) that a certain span of a bridge can withstand without structural damage resulting, is normally distributed with mean 400 and standard deviation 40. Suppose that the weight (again, in units of 1,000 pounds) of a car is a random variable with mean 3 and standard deviation .3. How many cars would have to be on the bridge span for the probability of structural damage to exceed .1?

SOLUTION Let P_n denote the probability of structural damage when there are n cars on the bridge. That is,

$$\begin{aligned} P_n &= P\{X_1 + \cdots + X_n \geq W\} \\ &= P\{X_1 + \cdots + X_n - W \geq 0\} \end{aligned}$$

where X_i is the weight of the ith car, $i = 1, \ldots, n$. Now it follows from the central limit theorem that $\sum_{i=1}^{n} X_i$ is approximately normal with mean $3n$ and variance $.09n$. Hence, since W is independent of the X_i, $i = 1, \ldots, n$, and is also normal, it follows that $\sum_{i=1}^{n} X_i - W$ is approximately normal, with mean and variance given by

$$E\left[\sum_{1}^{n} X_i - W\right] = 3n - 400$$

$$\text{Var}\left(\sum_{1}^{n} X_i - W\right) = \text{Var}\left(\sum_{1}^{n} X_i\right) + \text{Var}(W) = .09n + 1{,}600$$

Therefore, if we let

$$Z = \frac{\sum_{i=1}^{n} X_i - W - (3n - 400)}{\sqrt{.09n + 1{,}600}}$$

then

$$P_n = P\left\{Z \geq \frac{-(3n - 400)}{\sqrt{.09n + 1{,}600}}\right\}$$

where Z is approximately a standard normal random variable. Now $P\{Z \geq 1.28\} \approx .1$, and so if the number of cars n is such that

$$\frac{400 - 3n}{\sqrt{.09n + 1{,}600}} \leq 1.28$$

or

$$n \geq 117$$

then there is at least 1 chance in 10 that structural damage will occur. ■

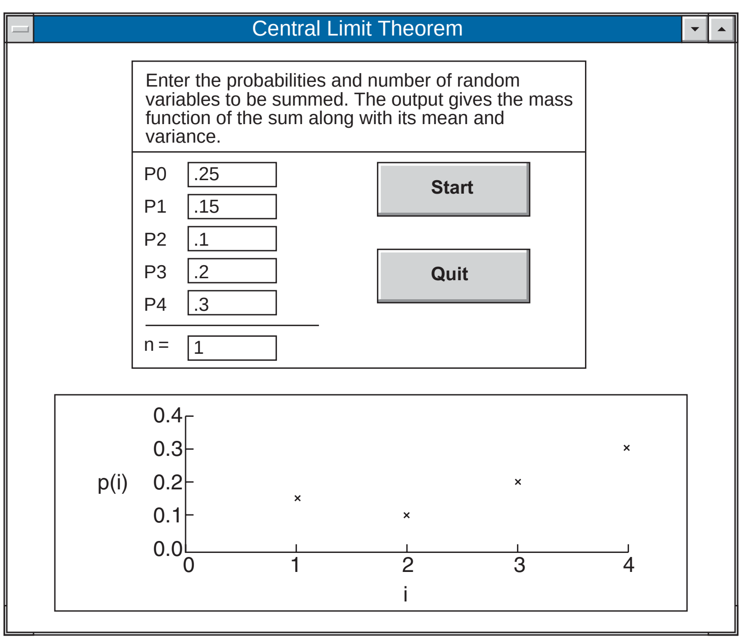
The central limit theorem is illustrated by Program 6.1 on the text disk. This program plots the probability mass function of the sum of n independent and identically distributed random variables that each take on one of the values 0, 1, 2, 3, 4. When using it, one enters the probabilities of these five values, and the desired value of n. Figures 6.2(a)–(f) give the resulting plot for a specified set of probabilities when $n = 1, 3, 5, 10, 25, 100$.

One of the most important applications of the central limit theorem is in regard to binomial random variables. Since such a random variable X having parameters (n, p) represents the number of successes in n independent trials when each trial is a success with probability p, we can express it as

$$X = X_1 + \cdots + X_n$$

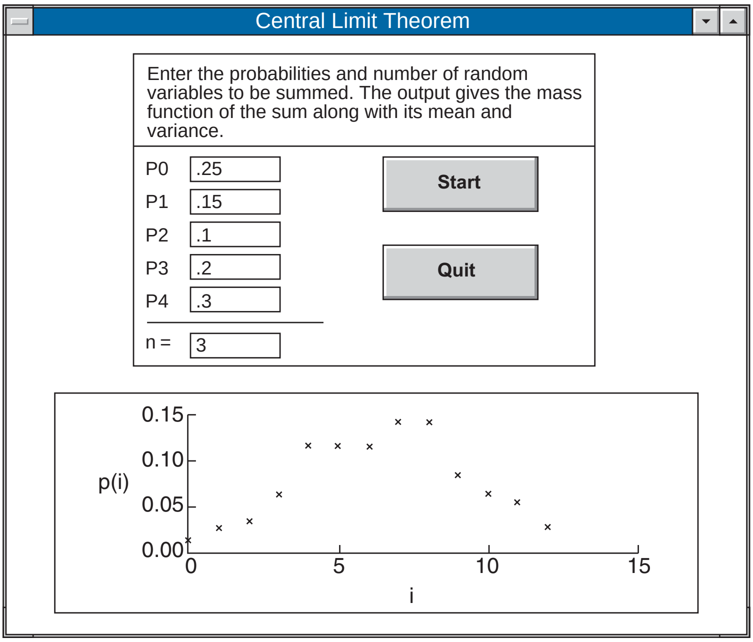
where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th trial is a success} \\ 0 & \text{otherwise} \end{cases}$$

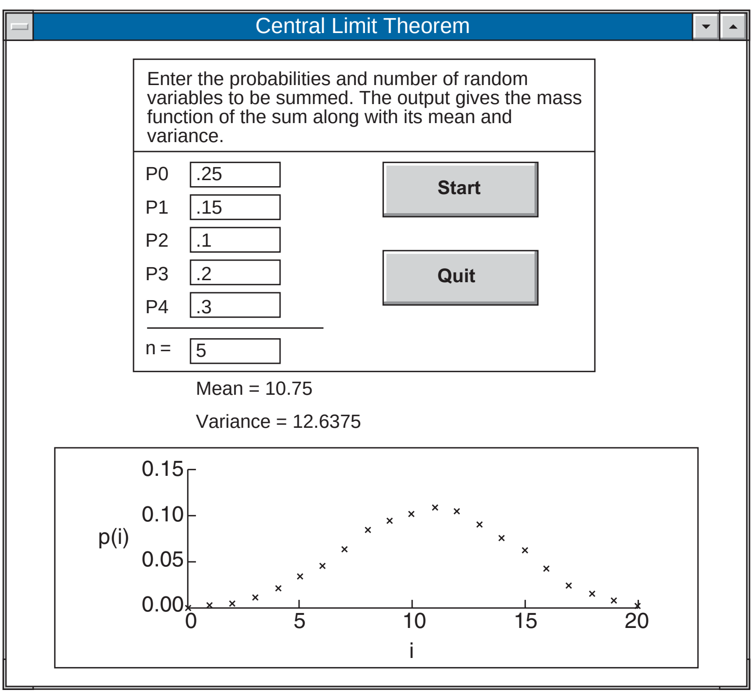


(a)

FIGURE 6.2 *(a)* $n = 1$, *(b)* $n = 3$, *(c)* $n = 5$, *(d)* $n = 10$, *(e)* $n = 25$, *(f)* $n = 100$.

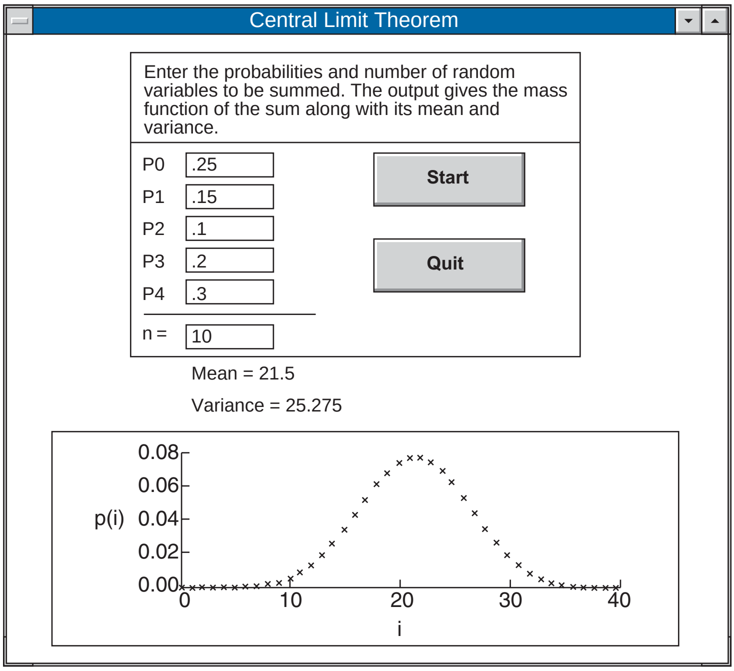


(b)

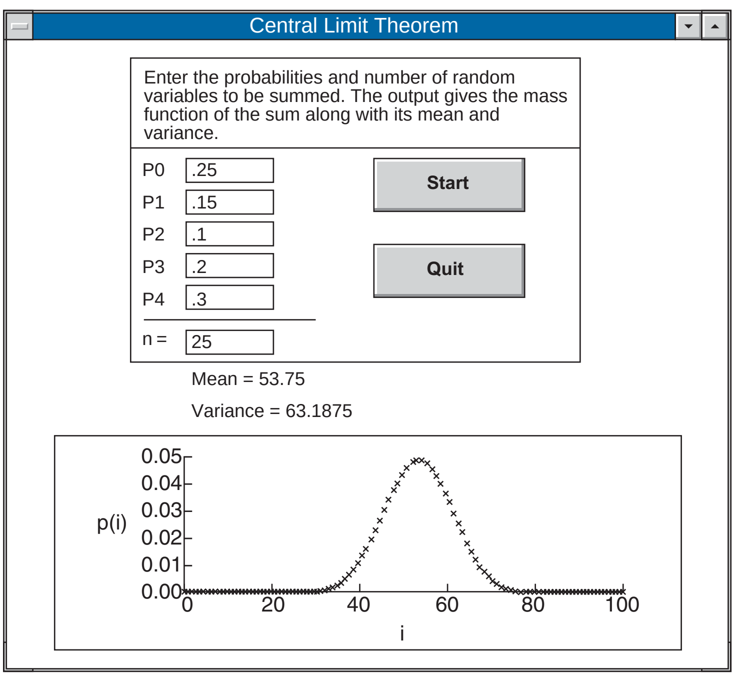


(c)

FIGURE 6.2 *(continued)*

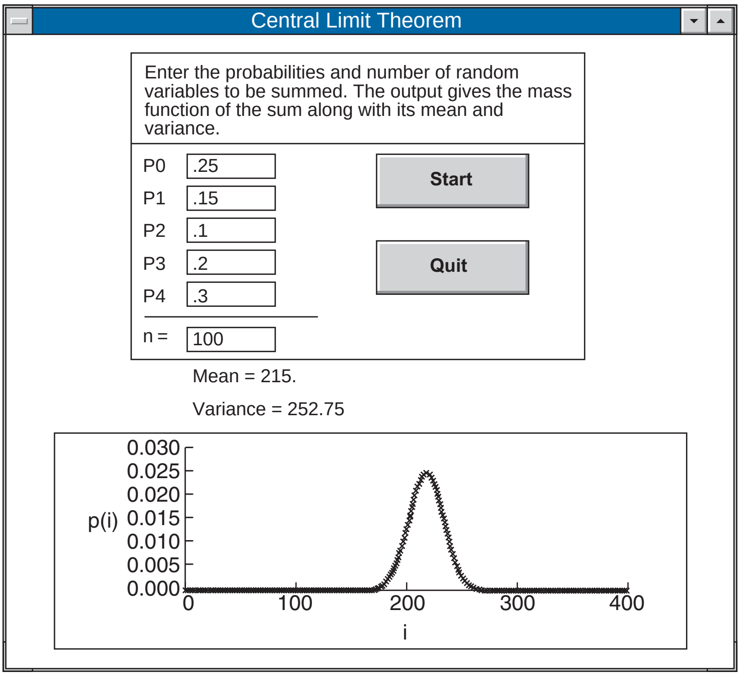


(d)



(e)

FIGURE 6.2 *(continued)*



(f)

FIGURE 6.2 *(continued)*

Because

$$E[X_i] = p, \qquad \mathrm{Var}(X_i) = p(1-p)$$

it follows from the central limit theorem that for n large

$$\frac{X - np}{\sqrt{np(1-p)}}$$

will approximately be a standard normal random variable [see Figure 6.3, which graphically illustrates how the probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger].

EXAMPLE 6.3c The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

SOLUTION Let X denote the number of students that attend; then assuming that each accepted applicant will independently attend, it follows that X is a binomial random

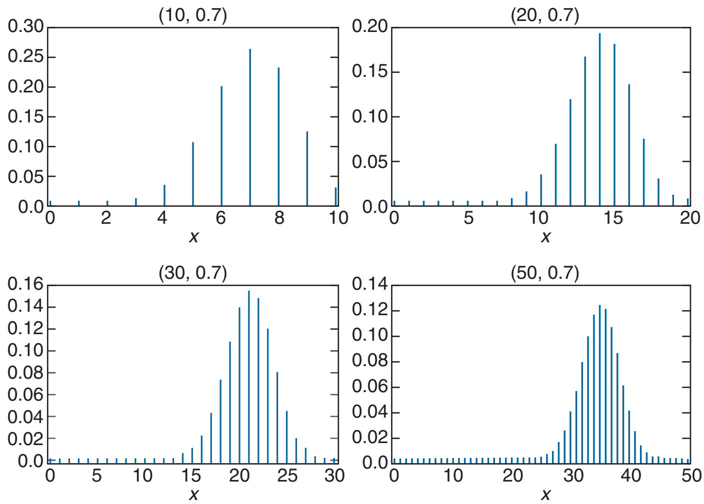


FIGURE 6.3 *Binomial probability mass functions converging to the normal density.*

variable with parameters $n = 450$ and $p = .3$. Since the binomial is a discrete and the normal a continuous distribution, it is best to compute $P\{X = i\}$ as $P\{i - .5 < X < i + .5\}$ when applying the normal approximation (this is called the continuity correction). This yields the approximation

$$P\{X > 150.5\} = P\left\{\frac{X - (450)(.3)}{\sqrt{450(.3)(.7)}} \geq \frac{150.5 - (450)(.3)}{\sqrt{450(.3)(.7)}}\right\}$$
$$\approx P\{Z > 1.59\} = .06$$

Hence, only 6 percent of the time do more than 150 of the first 450 accepted actually attend. ■

It should be noted that we now have two possible approximations to binomial probabilities: The Poisson approximation, which yields a good approximation when n is large and p small, and the normal approximation, which can be shown to be quite good when $np(1 - p)$ is large. [The normal approximation will, in general, be quite good for values of n satisfying $np(1 - p) \geq 10$.]

6.3.1 Approximate Distribution of the Sample Mean

Let $X_1, \ldots, X_n$ be a sample from a population having mean μ and variance σ^2. The central limit theorem can be used to approximate the distribution of the sample mean

$$\overline{X} = \sum_{i=1}^{n} X_i/n$$

Since a constant multiple of a normal random variable is also normal, it follows from the central limit theorem that $\overline{X}$ will be approximately normal when the sample size n is large. Since the sample mean has expected value μ and standard deviation $\sigma/\sqrt{n}$, it then follows that

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

has approximately a standard normal distribution.

EXAMPLE 6.3d The weights of a population of workers have mean 167 and standard deviation 27.

(a) If a sample of 36 workers is chosen, approximate the probability that the sample mean of their weights lies between 163 and 170.

(b) Repeat part (a) when the sample is of size 144.

SOLUTION Let Z be a standard normal random variable.

(a) It follows from the central limit theorem that $\overline{X}$ is approximately normal with mean 167 and standard deviation $27/\sqrt{36} = 4.5$. Therefore,

$$\begin{aligned} P\{163 < \overline{X} < 170\} &= P\left\{\frac{163 - 167}{4.5} < \frac{\overline{X} - 167}{4.5} < \frac{170 - 167}{4.5}\right\} \\ &= P\left\{-.8889 < \frac{\overline{X} - 167}{4.5} < .8889\right\} \\ &\approx 2P\{Z < .8889\} - 1 \\ &\approx .6259 \end{aligned}$$

(b) For a sample of size 144, the sample mean will be approximately normal with mean 167 and standard deviation $27/\sqrt{144} = 2.25$. Therefore,

$$\begin{aligned} P\{163 < \overline{X} < 170\} &= P\left\{\frac{163 - 167}{2.25} < \frac{\overline{X} - 167}{2.25} < \frac{170 - 167}{2.25}\right\} \\ &= P\left\{-1.7778 < \frac{\overline{X} - 167}{4.5} < 1.7778\right\} \\ &\approx 2P\{Z < 1.7778\} - 1 \\ &\approx .9246 \end{aligned}$$

Thus increasing the sample size from 36 to 144 increases the probability from .6259 to .9246. ■

EXAMPLE 6.3e An astronomer wants to measure the distance from her observatory to a distant star. However, due to atmospheric disturbances, any measurement will not yield the exact distance d. As a result, the astronomer has decided to make a series of measurements and then use their average value as an estimate of the actual distance. If the astronomer believes that the values of the successive measurements are independent random variables with a mean of d light years and a standard deviation of 2 light years, how many measurements need she make to be at least 95 percent certain that her estimate is accurate to within $\pm .5$ light years?

SOLUTION If the astronomer makes n measurements, then $\overline{X}$, the sample mean of these measurements, will be approximately a normal random variable with mean d and standard deviation $2/\sqrt{n}$. Thus, the probability that it will lie between $d \pm .5$ is obtained as follows:

$$\begin{aligned}
P\{-.5 < \overline{X} - d < .5\} &= P\left\{\frac{-.5}{2/\sqrt{n}} < \frac{\overline{X} - d}{2/\sqrt{n}} < \frac{.5}{2/\sqrt{n}}\right\} \\
&\approx P\{-\sqrt{n}/4 < Z < \sqrt{n}/4\} \\
&= 2P\{Z < \sqrt{n}/4\} - 1
\end{aligned}$$

where Z is a standard normal random variable.

Thus, the astronomer should make n measurements, where n is such that

$$2P\{Z < \sqrt{n}/4\} - 1 \geq .95$$

or, equivalently,

$$P\{Z < \sqrt{n}/4\} \geq .975$$

Since $P\{Z < 1.96\} = .975$, it follows that n should be chosen so that

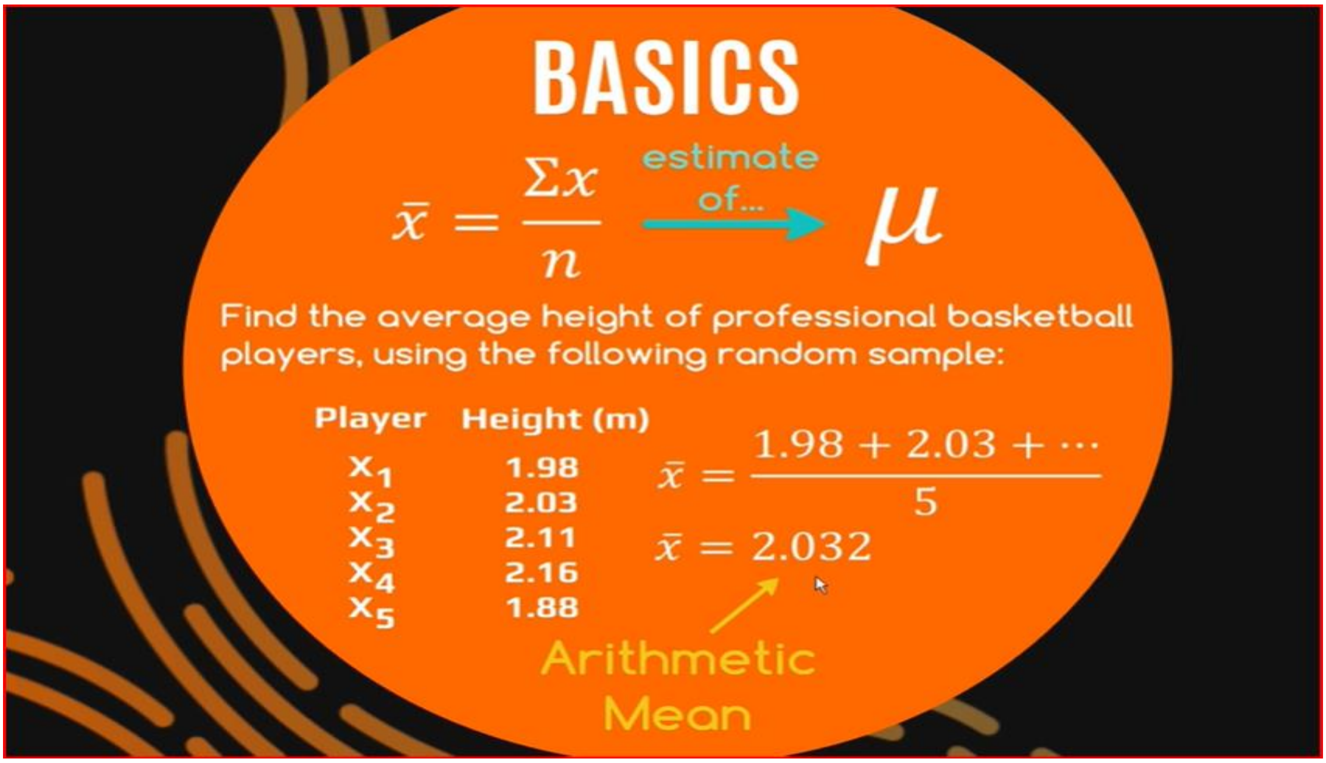
$$\sqrt{n}/4 \geq 1.96$$

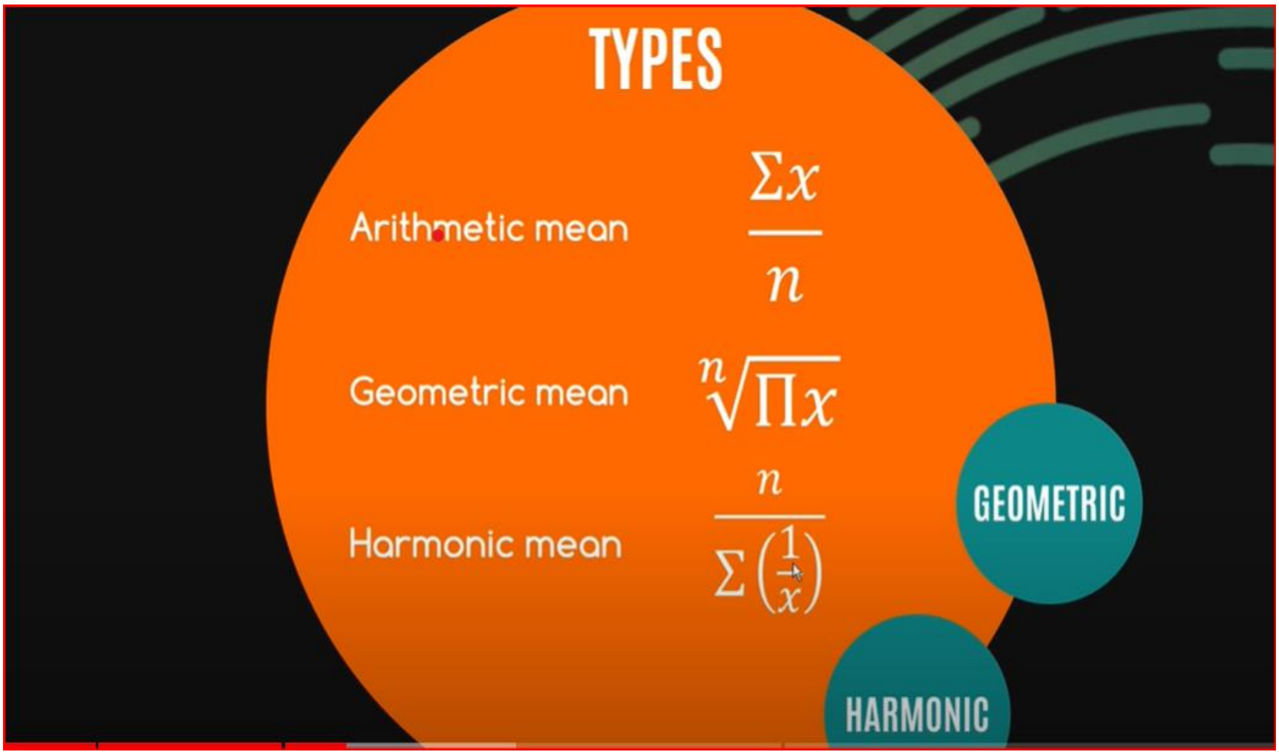
That is, at least 62 observations are necessary. ■

6.3.2 How Large a Sample Is Needed?

The central limit theorem leaves open the question of how large the sample size n needs to be for the normal approximation to be valid, and indeed the answer depends on the population distribution of the sample data. For instance, if the underlying population distribution is normal, then the sample mean $\overline{X}$ will also be normal regardless of the sample size. A general rule of thumb is that one can be confident of the normal approximation whenever the sample size n is at least 30. That is, practically speaking, no matter how nonnormal the underlying population distribution is, the sample mean of a sample of size at least 30 will be approximately normal. In most cases, the normal approximation is valid for much

Basis idea about Arithmetic mean, Geometric mean, Harmonic mean





The arithmetic mean, often simply called the average, is the sum of a set of numbers divided by the count of those numbers. It is the most commonly used measure of central tendency.

Mathematical formula:

$$\text{Arithmetic Mean} = \frac{x_1+x_2+...+x_n}{n}$$

Where:

- $x_1, x_2, ..., x_n$ are the individual numbers in the set.
- n is the count of numbers in the set.

Example:

Consider the set of numbers: 2, 5, 8, 11, 14. The arithmetic mean is calculated as follows:

$$\text{Arithmetic Mean} = \frac{2+5+8+11+14}{5} = \frac{40}{5} = 8$$

Geometric Mean:

The geometric mean is the nth root of the product of n numbers. It is used when dealing with quantities that are proportional to each other.

Mathematical formula:

$$\text{Geometric Mean} = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}$$

Where:

- $x_1, x_2, ..., x_n$ are the individual numbers in the set.
- n is the count of numbers in the set.

Example:

Consider the set of numbers: 2, 4, 8, 16. The geometric mean is calculated as follows:

$$\text{Geometric Mean} = \sqrt[4]{2 \cdot 4 \cdot 8 \cdot 16} = \sqrt[4]{1024} = 8$$

Harmonic Mean:

The harmonic mean is the reciprocal of the arithmetic mean of the reciprocals of a set of numbers. It is often used in situations where rates or ratios are involved.

Mathematical formula:

$$\text{Harmonic Mean} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

Where:

- $x_1, x_2, \ldots, x_n$ are the individual numbers in the set.
- n is the count of numbers in the set.

Example:

Consider the set of numbers: 2, 4, 8, 16. The harmonic mean is calculated as follows:

$$\text{Harmonic Mean} = \frac{4}{\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16}} = \frac{4}{\frac{8+4+2+1}{16}} = \frac{4}{\frac{15}{16}} = \frac{64}{15}$$

In summary AM$\geq$ $GM \geq HM$

In summary:

Arithmetic mean **is suitable for situations where the data is evenly distributed.**

Geometric mean **is useful for situations involving growth rates, ratios, or proportions.**

Harmonic mean **is appropriate when dealing with rates, such as speed or efficiency.**

Finding Mode and Median for Grouped Data

Mode:

The mode is the value or values that occur most frequently. For grouped data, you can find the mode using the formula:

$$\text{Mode} = L + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times c$$

Where:

- L is the lower class boundary of the modal class,
- f_1 is the frequency of the modal class,
- f_0 is the frequency of the class before the modal class,
- f_2 is the frequency of the class after the modal class,
- c is the width of the class interval.

Median:

For grouped data, the median can be found using the formula:

$$\text{Median} = L + \left(\frac{\frac{N}{2} - F}{f}\right) \times c$$

Where:

- L is the lower class boundary of the median class,
- N is the total number of observations,
- F is the cumulative frequency of the class before the median class,
- f is the frequency of the median class,
- c is the width of the class interval.

Let's go through an example to illustrate these concepts.

Consider the following grouped data:

Class Interval	Frequency
$10-20$	5
$20-30$	8
$30-40$	12
$40-50$	6
$50-60$	9

1. **Mode:**
   - Modal class: 30-40
   - $L = 30$ (lower boundary of the modal class)
   - $f_1 = 12$ (frequency of the modal class)
   - $f_0 = 8$ (frequency of the class before the modal class)
   - $f_2 = 6$ (frequency of the class after the modal class)
   - $c = 10$ (width of the class interval)

   Substituting these values into the mode formula:

   $\text{Mode} = 30 + \left(\frac{12-8}{2\times12-8-6}\right) \times 10$

2. **Median:**
   - Total number of observations (N) = Sum of frequencies = $5 + 8 + 12 + 6 + 9 = 40$
   - Median class: 30-40
   - $L = 30$ (lower boundary of the median class)
   - $F = 5 + 8 = 13$ (cumulative frequency of the class before the median class)
   - $f = 12$ (frequency of the median class)
   - $c = 10$ (width of the class interval)

   Substituting these values into the median formula:

   $\text{Median} = 30 + \left(\frac{\frac{40}{2}-13}{12}\right) \times 10$

Now, you can calculate these values to find the mode and median for the given data.