Lecture 9: Measures of Central Tendency and Sampling Distributions

Assist. Prof. Dr. Emel YAVUZ DUMAN

Introduction to Probability and Statistics
İstanbul Kültür University
Faculty of Engineering
Outline

1. Measures of Central Tendency

2. Sampling Distributions
   - Population and Sample, Statistical Inference
   - Sampling With and Without Replacement
   - Random Samples

3. The Sampling Distribution of the Mean

4. The Sampling Distribution of the Mean: Finite Population
Outline

1. Measures of Central Tendency

2. Sampling Distributions
   - Population and Sample. Statistical Inference
   - Sampling With and Without Replacement
   - Random Samples

3. The Sampling Distribution of the Mean

4. The Sampling Distribution of the Mean: Finite Population
Introduction

A measure of *central tendency* is a single value that attempts to describe a set of data by identifying the central position within that set of data.
A measure of *central tendency* is a single value that attempts to describe a set of data by identifying the central position within that set of data. As such, measures of central tendency are sometimes called measures of *central location*.
A measure of *central tendency* is a single value that attempts to describe a set of data by identifying the central position within that set of data. As such, measures of central tendency are sometimes called measures of *central location*. They are also classed as summary statistics.
A measure of *central tendency* is a single value that attempts to describe a set of data by identifying the central position within that set of data. As such, measures of central tendency are sometimes called measures of *central location*. They are also classed as summary statistics. The mean (often called the average) is most likely the measure of central tendency that you are most familiar with, but there are others, such as the median and the mode. The *mean*, *median* and *mode* are all valid measures of central tendency, but under different conditions, some measures of central tendency become more appropriate to use than others.
Introduction

A measure of *central tendency* is a single value that attempts to describe a set of data by identifying the central position within that set of data. As such, measures of central tendency are sometimes called measures of *central location*. They are also classed as summary statistics. The mean (often called the average) is most likely the measure of central tendency that you are most familiar with, but there are others, such as the median and the mode. The *mean*, *median* and *mode* are all valid measures of central tendency, but under different conditions, some measures of central tendency become more appropriate to use than others. In the following, we will look at the mean, mode and median, and learn how to calculate them and under what conditions they are most appropriate to be used.
The mean (or average) is the most popular and well known measure of central tendency.
The *mean* (or *average*) is the most popular and well known measure of central tendency. It can be used with both discrete and continuous data, although its use is most often with continuous data.
The *mean* (or *average*) is the most popular and well known measure of central tendency. It can be used with both discrete and continuous data, although its use is most often with continuous data.

**Definition 1**

The mean is equal to the sum of all the values in the data set divided by the number of values in the data set when we are dealing with discrete random variables.
The mean (or average) is the most popular and well known measure of central tendency. It can be used with both discrete and continuous data, although its use is most often with continuous data.

**Definition 1**

The mean is equal to the sum of all the values in the data set divided by the number of values in the data set when we are dealing with discrete random variables.

So, if we have $n$ values in a data set and they have values $x_1, x_2, \cdots, x_n$, the sample mean, usually denoted by $\overline{x}$ is:

$$
\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \sum_{k=1}^{n} \frac{x_k}{n}.
$$
Example 2

A football team keep records of the number of goals it scores per match during a season:

<table>
<thead>
<tr>
<th>No. of goals</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Find the mean number of goals per match.
Solution. The table above can be used, with a third column added.

<table>
<thead>
<tr>
<th>No. of goals</th>
<th>Frequency</th>
<th>No. of goals × Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>0 × 8 = 0</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>1 × 10 = 10</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>2 × 12 = 24</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3 × 3 = 9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4 × 5 = 20</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>5 × 2 = 10</td>
</tr>
<tr>
<td>Totals</td>
<td>40 (total matches)</td>
<td>73 (total goals)</td>
</tr>
</tbody>
</table>

Mean = \( \bar{x} = \frac{73}{40} = 1.825 \).
Definition 3

The *median* is the middle score for a set of data that has been arranged in order of magnitude.
Median

**Definition 3**

The *median* is the middle score for a set of data that has been arranged in order of magnitude.

In order to calculate the median, suppose we have the data below:

65 55 89 56 35 14 56 55 87 45 92
Median

**Definition 3**

The *median* is the middle score for a set of data that has been arranged in order of magnitude.

In order to calculate the median, suppose we have the data below:

65 55 89 56 35 14 56 55 87 45 92

We first need to rearrange that data into order of magnitude (smallest first):

14 35 45 55 55 56 56 65 87 89 92
**Median**

**Definition 3**

The *median* is the middle score for a set of data that has been arranged in order of magnitude.

In order to calculate the median, suppose we have the data below:

65 55 89 56 35 14 56 55 87 45 92

We first need to rearrange that data into order of magnitude (smallest first):

14 35 45 55 55 56 56 65 87 89 92

Our median mark is the middle mark - in this case, 56. It is the middle mark because there are 5 scores before it and 5 scores after it.
Median

This works fine when you have an odd number of scores, but what happens when you have an even number of scores?
Median

This works fine when you have an odd number of scores, but what happens when you have an even number of scores? What if you had only 10 scores?
Median

This works fine when you have an odd number of scores, but what happens when you have an even number of scores? What if you had only 10 scores? Well, you simply have to take the middle two scores and average the result.
Median

This works fine when you have an odd number of scores, but what happens when you have an even number of scores? What if you had only 10 scores? Well, you simply have to take the middle two scores and average the result. So, if we look at the example below:

65 55 89 56 35 14 56 55 87 45
This works fine when you have an odd number of scores, but what happens when you have an even number of scores? What if you had only 10 scores? Well, you simply have to take the middle two scores and average the result. So, if we look at the example below:

65 55 89 56 35 14 56 55 87 45

We again rearrange that data into order of magnitude (smallest first):

14 35 45 55 56 56 65 87 89
This works fine when you have an odd number of scores, but what happens when you have an even number of scores? What if you had only 10 scores? Well, you simply have to take the middle two scores and average the result. So, if we look at the example below:

\[
65 \ 55 \ 89 \ 56 \ 35 \ 14 \ 56 \ 55 \ 87 \ 45
\]

We again rearrange that data into order of magnitude (smallest first):

\[
14 \ 35 \ 45 \ 55 \ 55 \ 56 \ 56 \ 65 \ 87 \ 89
\]

Only now we have to take the 5th and 6th score in our data set and average them to get a median of 55.5.
Example 4

Seven basketball players shoot 30 free throws during a practice session. The numbers of baskets they make are listed below. What is the median number of baskets made?

22 23 11 18 22 20 15
Example 4

Seven basketball players shoot 30 free throws during a practice session. The numbers of baskets they make are listed below. What is the median number of baskets made?

22 23 11 18 22 20 15

Solution. Here are the scores in ascending order.
Example 4

Seven basketball players shoot 30 free throws during a practice session. The numbers of baskets they make are listed below. What is the median number of baskets made?

22 23 11 18 22 20 15

Solution. Here are the scores in ascending order.

11 15 18 20 22 22 23

The median number of baskets is 20 because there are three scores above 20 and three scores below 20.
Example 5

Twelve members of a gym class, some in good physical condition and some in not-so-good physical condition, see how many sit-ups they can complete in a minute. Here are their scores.

2 3 6 10 12 12 14 15 15 15 24 25

What is the median number of sit-ups?
Example 5

Twelve members of a gym class, some in good physical condition and some in not-so-good physical condition, see how many sit-ups they can complete in a minute. Here are their scores.

\[
2 \ 3 \ 6 \ 10 \ 12 \ 12 \ 14 \ 15 \ 15 \ 15 \ 24 \ 25
\]

What is the median number of sit-ups?

Solution. The median is 13, because there are six scores below 13 and six scores above 13.
Example 5

Twelve members of a gym class, some in good physical condition and some in not-so-good physical condition, see how many sit-ups they can complete in a minute. Here are their scores.

```
2 3 6 10 12 12 14 15 15 15 24 25
```

What is the median number of sit-ups?

**Solution.** The median is 13, because there are six scores below 13 and six scores above 13. Note that the median does not necessarily have to be an existing score.
Example 5

Twelve members of a gym class, some in good physical condition and some in not-so-good physical condition, see how many sit-ups they can complete in a minute. Here are their scores.

2 3 6 10 12 12 14 15 15 15 24 25

What is the median number of sit-ups?

Solution. The median is 13, because there are six scores below 13 and six scores above 13. Note that the median does not necessarily have to be an existing score. In this case, no one completed exactly 13 sit-ups.
Mode

**Definition 6**

The *mode* is the most frequently occurring value in a set of values.
Definition 6

The *mode* is the most frequently occurring value in a set of values.

The mode is the most frequent score in our data set.
Mode

**Definition 6**

The *mode* is the most frequently occurring value in a set of values.

The mode is the most frequent score in our data set. On a histogram it represents the highest bar in a bar chart or histogram.
Mode

Definition 6

The *mode* is the most frequently occurring value in a set of values.

The mode is the most frequent score in our data set. On a histogram it represents the highest bar in a bar chart or histogram. You can, therefore, sometimes consider the mode as being the most popular option.
Definition 6

The *mode* is the most frequently occurring value in a set of values.

The mode is the most frequent score in our data set. On a histogram it represents the highest bar in a bar chart or histogram. You can, therefore, sometimes consider the mode as being the most popular option. An example of a mode is presented below:
Mode

Normally, the mode is used for categorical data where we wish to know which is the most common category, as illustrated below:

We can see above that the most common form of transport, in this particular data set, is the bus.
Mode

However, one of the problems with the mode is that it is not unique, so it leaves us with problems when we have two or more values that share the highest frequency, such as below:
Mode

Example 7

Here we have the number of items found by 11 children in a scavenger hunt. What was the modal number of items found?

14 6 11 8 7 20 11 3 7 5 7
Mode

Solution. If there are not too many numbers, a simple list of scores will do.
Mode

**Solution.** If there are not too many numbers, a simple list of scores will do. However, if there are many scores, you will need to put the scores in order and then create a frequency table.
Mode

**Solution.** If there are not too many numbers, a simple list of scores will do. However, if there are many scores, you will need to put the scores in order and then create a frequency table. Here are the previous scores in a descending order frequency table:

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
Mode

**Solution.** If there are not too many numbers, a simple list of scores will do. However, if there are many scores, you will need to put the scores in order and then create a frequency table. Here are the previous scores in a descending order frequency table:

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

The mode is 7, because there are more 7s than any other number.
**Mode**

**Solution.** If there are not too many numbers, a simple list of scores will do. However, if there are many scores, you will need to put the scores in order and then create a frequency table. Here are the previous scores in a descending order frequency table:

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

The mode is 7, because there are more 7s than any other number. Note that the number of scores on either side of the mode does not have to be equal.
Example 8

To find the mode of the number of days in each month:

<table>
<thead>
<tr>
<th>Month</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>31</td>
</tr>
<tr>
<td>February</td>
<td>28</td>
</tr>
<tr>
<td>March</td>
<td>31</td>
</tr>
<tr>
<td>April</td>
<td>30</td>
</tr>
<tr>
<td>May</td>
<td>31</td>
</tr>
<tr>
<td>June</td>
<td>30</td>
</tr>
<tr>
<td>July</td>
<td>31</td>
</tr>
<tr>
<td>August</td>
<td>31</td>
</tr>
<tr>
<td>September</td>
<td>30</td>
</tr>
<tr>
<td>October</td>
<td>31</td>
</tr>
<tr>
<td>November</td>
<td>30</td>
</tr>
<tr>
<td>December</td>
<td>31</td>
</tr>
</tbody>
</table>
Measures of Central Tendency

Sampling Distributions

The Sampling Distribution of the Mean

The Sampling Distribution of the Median

Mode

Example 8

To find the mode of the number of days in each month:

<table>
<thead>
<tr>
<th>Month</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>31</td>
</tr>
<tr>
<td>February</td>
<td>28</td>
</tr>
<tr>
<td>March</td>
<td>31</td>
</tr>
<tr>
<td>April</td>
<td>30</td>
</tr>
<tr>
<td>May</td>
<td>31</td>
</tr>
<tr>
<td>June</td>
<td>30</td>
</tr>
<tr>
<td>July</td>
<td>31</td>
</tr>
<tr>
<td>August</td>
<td>31</td>
</tr>
<tr>
<td>September</td>
<td>30</td>
</tr>
<tr>
<td>October</td>
<td>31</td>
</tr>
<tr>
<td>November</td>
<td>30</td>
</tr>
<tr>
<td>December</td>
<td>31</td>
</tr>
</tbody>
</table>

Solution. 7 months have a 31 days, 4 months have a total of 30 days and only 1 month has a total of 28 days (29 in a leap year).
**Example 8**

To find the mode of the number of days in each month:

<table>
<thead>
<tr>
<th>Month</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>31</td>
</tr>
<tr>
<td>February</td>
<td>28</td>
</tr>
<tr>
<td>March</td>
<td>31</td>
</tr>
<tr>
<td>April</td>
<td>30</td>
</tr>
<tr>
<td>May</td>
<td>31</td>
</tr>
<tr>
<td>June</td>
<td>30</td>
</tr>
<tr>
<td>July</td>
<td>31</td>
</tr>
<tr>
<td>August</td>
<td>31</td>
</tr>
<tr>
<td>September</td>
<td>30</td>
</tr>
<tr>
<td>October</td>
<td>31</td>
</tr>
<tr>
<td>November</td>
<td>30</td>
</tr>
<tr>
<td>December</td>
<td>31</td>
</tr>
</tbody>
</table>

**Solution.** 7 months have a 31 days, 4 months have a total of 30 days and only 1 month has a total of 28 days (29 in a leap year). The mode is therefore, 31.
Mode

Some data sets may have more than one mode:
Mode

Some data sets may have more than one mode: 1, 3, 3, 4, 4, 5 for example, has two most frequently occurring numbers (3 and 4) this is known as a *bimodal set*.
Some data sets may have more than one mode: 1, 3, 3, 4, 4, 5 for example, has two most frequently occurring numbers (3 and 4) this is known as a *bimodal set*. Data sets with more than two modes are referred to as *multimodal data sets*. 
Mode

Some data sets may have more than one mode: 1, 3, 3, 4, 4, 5 for example, has two most frequently occurring numbers (3 and 4) this is known as a *bimodal set*. Data sets with more than two modes are referred to as *multimodal data sets*. If a data set contains only unique numbers then calculating the mode is more problematic.
Some data sets may have more than one mode: 1, 3, 3, 4, 4, 5 for example, has two most frequently occurring numbers (3 and 4) this is known as a *bimodal set*. Data sets with more than two modes are referred to as *multimodal data sets*. If a data set contains only unique numbers then calculating the mode is more problematic.

It is usually perfectly acceptable to say there is no mode, but if a mode has to be found then the usual way is to create number ranges and then count the one with the most points in it.
Mode

For example from a set of data showing the speed of passing cars we see that out of 9 cars the recorded speeds are:

34 42 39 41 50 48 49 33 47
Mode

For example from a set of data showing the speed of passing cars we see that out of 9 cars the recorded speeds are:

34 42 39 41 50 48 49 33 47

These numbers are all unique (each only occurs once), there is no mode.
Mode

For example from a set of data showing the speed of passing cars we see that out of 9 cars the recorded speeds are:

34 42 39 41 50 48 49 33 47

These numbers are all unique (each only occurs once), there is no mode. In order to find a mode we build categories on an even scale:

Mode

For example from a set of data showing the speed of passing cars we see that out of 9 cars the recorded speeds are:

34 42 39 41 50 48 49 33 47

These numbers are all unique (each only occurs once), there is no mode. In order to find a mode we build categories on an even scale:


Then work out how many of the values fall into each category, how many times a number between 30 and 32 occurs, etc.
Mode

For example from a set of data showing the speed of passing cars we see that out of 9 cars the recorded speeds are:

34 42 39 41 50 48 49 33 47

These numbers are all unique (each only occurs once), there is no mode. In order to find a mode we build categories on an even scale:


Then work out how many of the values fall into each category, how many times a number between 30 and 32 occurs, etc.

30–32 = 0
33–35 = 2
36–38 = 0
39–41 = 2
42–44 = 1
45–47 = 1
48–50 = 3
The category with the most values is 48-50 with 3 values.
The category with the most values is 48-50 with 3 values. We can take the mid value of the category to estimate the mode at 49.
Mode

The category with the most values is 48-50 with 3 values. We can take the mid value of the category to estimate the mode at 49. This method of calculating the mode is not ideal as, depending on the categories you define, the mode may be different.
Example 9

The table below gives data on the heights, in cm, of 51 children.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$140 \leq h &lt; 150$</td>
<td>6</td>
</tr>
<tr>
<td>$150 \leq h &lt; 160$</td>
<td>16</td>
</tr>
<tr>
<td>$160 \leq h &lt; 170$</td>
<td>21</td>
</tr>
<tr>
<td>$170 \leq h &lt; 180$</td>
<td>8</td>
</tr>
</tbody>
</table>

(a) Estimate the mean, (b) estimate the median and (c) find the modal class.
**Solution. (a)** To estimate the mean, the mid-point of each interval should be used

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Mid-point</th>
<th>Frequency</th>
<th>Mid-point $\times$ Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$140 \leq h &lt; 150$</td>
<td>145</td>
<td>6</td>
<td>$145 \times 6 = 870$</td>
</tr>
<tr>
<td>$150 \leq h &lt; 160$</td>
<td>155</td>
<td>16</td>
<td>$155 \times 16 = 2480$</td>
</tr>
<tr>
<td>$160 \leq h &lt; 170$</td>
<td>165</td>
<td>21</td>
<td>$165 \times 21 = 3465$</td>
</tr>
<tr>
<td>$170 \leq h &lt; 180$</td>
<td>175</td>
<td>8</td>
<td>$175 \times 8 = 1400$</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>51</strong></td>
<td></td>
<td><strong>8215</strong></td>
</tr>
</tbody>
</table>
**Solution. (a)** To estimate the mean, the mid-point of each interval should be used.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Mid-point</th>
<th>Frequency</th>
<th>Mid-point × Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>140 ≤ ( h ) &lt; 150</td>
<td>145</td>
<td>6</td>
<td>145 × 6 = 870</td>
</tr>
<tr>
<td>150 ≤ ( h ) &lt; 160</td>
<td>155</td>
<td>16</td>
<td>155 × 16 = 2480</td>
</tr>
<tr>
<td>160 ≤ ( h ) &lt; 170</td>
<td>165</td>
<td>21</td>
<td>165 × 21 = 3465</td>
</tr>
<tr>
<td>170 ≤ ( h ) &lt; 180</td>
<td>175</td>
<td>8</td>
<td>175 × 8 = 1400</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>51</strong></td>
<td></td>
<td><strong>8215</strong></td>
</tr>
</tbody>
</table>

\[
\text{Mean } = \overline{X} = \frac{8215}{51} = 161 \text{ (to the nearest cm)}
\]
(b) the median is the 26th value. In this case it lies in the $160 \leq h < 170$ class interval. The 4th value in the interval is needed. It is estimated as

$$160 + \frac{4}{21} \times 10 = 162 \text{ (to the nearest cm).}$$
(b) the median is the 26th value. In this case it lies in the $160 \leq h < 170$ class interval. The 4th value in the interval is needed. It is estimated as

$$160 + \frac{4}{21} \times 10 = 162 \text{ (to the nearest cm).}$$

(c) The modal class is $160 \leq h < 170$ as it contains the most values.
Note.

Example 9 uses what are called *continuous data*, since height can be of any value (other examples of continuous data are weight, temperature, area, volume and time).
Example 9 uses what are called *continuous data*, since height can be of any value (other examples of continuous data are weight, temperature, area, volume and time).

The next example uses *discrete data*, that is, data which can take only a particular value, such as integers 1, 2, 3, ··· in this case.
Example 9 uses what are called *continuous data*, since height can be of any value (other examples of continuous data are weight, temperature, area, volume and time).

The next example uses *discrete data*, that is, data which can take only a particular value, such as integers 1, 2, 3, ⋯ in this case.

The calculations for mean and mode are not effected but estimation of the median requires replacing the *discrete* grouped data with an approximate *continuous* interval, like continuity correction.
Example 10

The number of days that children were missing from school due to sickness in one year was recorded.

<table>
<thead>
<tr>
<th>Number of days off sick</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 5</td>
<td>12</td>
</tr>
<tr>
<td>6 – 10</td>
<td>11</td>
</tr>
<tr>
<td>11 – 15</td>
<td>10</td>
</tr>
<tr>
<td>16 – 20</td>
<td>4</td>
</tr>
<tr>
<td>21 – 25</td>
<td>3</td>
</tr>
</tbody>
</table>

(a) Estimate the mean, (b) estimate the median and (c) find the modal class.
Solution. (a) The estimate is made by assuming that all the values in a class interval are equal to the midpoint of the class interval.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Mid-point</th>
<th>Frequency</th>
<th>Mid-point × Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 5</td>
<td>3</td>
<td>12</td>
<td>3 × 12 = 36</td>
</tr>
<tr>
<td>6 – 10</td>
<td>8</td>
<td>11</td>
<td>8 × 11 = 88</td>
</tr>
<tr>
<td>11 – 15</td>
<td>13</td>
<td>10</td>
<td>13 × 10 = 130</td>
</tr>
<tr>
<td>16 – 20</td>
<td>18</td>
<td>4</td>
<td>18 × 4 = 72</td>
</tr>
<tr>
<td>21 – 25</td>
<td>23</td>
<td>3</td>
<td>23 × 3 = 69</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>40</td>
<td>395</td>
</tr>
</tbody>
</table>
Solution. (a) The estimate is made by assuming that all the values in a class interval are equal to the midpoint of the class interval

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>Mid-point</th>
<th>Frequency</th>
<th>Mid-point \times Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 − 5</td>
<td>3</td>
<td>12</td>
<td>3 \times 12 = 36</td>
</tr>
<tr>
<td>6 − 10</td>
<td>8</td>
<td>11</td>
<td>8 \times 11 = 88</td>
</tr>
<tr>
<td>11 − 15</td>
<td>13</td>
<td>10</td>
<td>13 \times 10 = 130</td>
</tr>
<tr>
<td>16 − 20</td>
<td>18</td>
<td>4</td>
<td>18 \times 4 = 72</td>
</tr>
<tr>
<td>21 − 25</td>
<td>23</td>
<td>3</td>
<td>23 \times 3 = 69</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>40</strong></td>
<td><strong>395</strong></td>
<td></td>
</tr>
</tbody>
</table>

Mean = \overline{x} = \frac{395}{40} = 9.925 days.
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values.
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the $6 - 10$ class interval, which is really the $5.5 - 10.5$ class interval, so this interval contains the median.
(b) As there are 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the 6 – 10 class interval, which is really the 5.5 – 10.5 class interval, so this interval contains the median.

As there are 12 values in the first class interval, the median is found by considering 8th and 9th values of the second interval.
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the 6 – 10 class interval, which is really the 5.5 – 10.5 class interval, so this interval contains the median.

As there are 12 values in the first class interval, the median is found by considering 8th and 9th values of the second interval. As there are 11 values in the second interval, the median is estimated as being

\[
\frac{8.5}{11}
\]

of the way along the second interval.
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the 6 – 10 class interval, which is really the 5.5 – 10.5 class interval, so this interval contains the median.

As there are 12 values in the first class interval, the median is found by considering 8th and 9th values of the second interval. As there are 11 values in the second interval, the median is estimated as being

\[
\frac{8.5}{11}
\]
/of the way along the second interval. But the length of the second interval is 10.5 – 5.5 = 5, so the median is estimated by

\[
\frac{8.5}{11} \times 5 = 3.86
\]

from the start of this interval.
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the 6 – 10 class interval, which is really the 5.5 – 10.5 class interval, so this interval contains the median.

As there are 12 values in the first class interval, the median is found by considering 8th and 9th values of the second interval. As there are 11 values in the second interval, the median is estimated as being

\[
\frac{8.5}{11}
\]

of the way along the second interval. But the length of the second interval is 10.5 – 5.5 = 5, so the median is estimated by

\[
\frac{8.5}{11} \times 5 = 3.86
\]

from the start of this interval. Therefore the median is estimated as

\[
5.5 + 3.86 = 9.36.
\]
(b) As there 40 pupils, we need to consider the mean of 20th and 21st values. These both lie in the 6 – 10 class interval, which is really the 5.5 – 10.5 class interval, so this interval contains the median.

As there are 12 values in the first class interval, the median is found by considering 8th and 9th values of the second interval. As there are 11 values in the second interval, the median is estimated as being \[ \frac{8.5}{11} \]

of the way along the second interval. But the length of the second interval is 10.5 – 5.5 = 5, so the median is estimated by

\[ \frac{8.5}{11} \times 5 = 3.86 \]

from the start of this interval. Therefore the median is estimated as

\[ 5.5 + 3.86 = 9.36. \]

(c) The modal class is 1 – 5, as this class contains the most
We see that if the mean is lower than the mode, the distribution is negatively skewed. Conversely, if the mean is higher than the mode, the distribution is positively skewed. Similarly, one can tell from the shape of the distribution where the mean, median, and mode will fall. If a distribution is negatively skewed, the mean must be lower than the mode. Conversely, if a distribution is positively skewed, the mean must be higher than the mode.
Outline

1. Measures of Central Tendency

2. Sampling Distributions
   - Population and Sample. Statistical Inference
   - Sampling With and Without Replacement
   - Random Samples

3. The Sampling Distribution of the Mean

4. The Sampling Distribution of the Mean: Finite Population
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects.
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group,
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group, called the *population*,
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group, called the *population*, which may be difficult or impossible to do,
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group, called the *population*, which may be difficult or impossible to do, we may examine only a small part of this population, which is called a *sample*. 
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group, called the *population*, which may be difficult or impossible to do, we may examine only a small part of this population, which is called a *sample*. We do this with the aim of inferring certain facts about the population from results found in the sample, a process known as *statistical inference*. 
Often in practice we are interested in drawing valid conclusions about a large group of individuals or objects. Instead of examining the entire group, called the *population*, which may be difficult or impossible to do, we may examine only a small part of this population, which is called a *sample*. We do this with the aim of inferring certain facts about the population from results found in the sample, a process known as *statistical inference*. The process of obtaining samples is called *sampling*.
Example 11

We may wish to draw conclusions about the heights (or weights) of 12,000 adult students (the population) by examining only 100 students (a sample) selected from this population.
Example 11

We may wish to draw conclusions about the heights (or weights) of 12,000 adult students (the population) by examining only 100 students (a sample) selected from this population.

Example 12

We may wish to draw conclusions about the percentage of defective bolts produced in a factory during a given 6-day week by examining 20 bolts each day produced at various times during the day. In this case all bolts produced during the week comprise the population, while the 120 selected bolts constitute a sample.
Example 13

We may wish to draw conclusions about the fairness of a particular coin by tossing it repeatedly. The population consists of all possible tosses of the coin. A sample could be obtained by examining, say, the first 60 tosses of the coin and noting the percentages of heads and tails.
Example 13

We may wish to draw conclusions about the fairness of a particular coin by tossing it repeatedly. The population consists of all possible tosses of the coin. A sample could be obtained by examining, say, the first 60 tosses of the coin and noting the percentages of heads and tails.

Example 14

We may wish to draw conclusions about the colors of 200 marbles (the population) in an urn by selecting a sample of 20 marbles from the urn, where each marble selected is returned after its color is observed.
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778.567.”
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778.567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects.
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778.567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects. In Example 11 we can speak of the population of 12,000 heights (or weights) while in Example 14 we can speak of the population of all 200 colors in the urn (some of which may be the same).
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778.567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects. In Example 11 we can speak of the population of 12,000 heights (or weights) while in Example 14 we can speak of the population of all 200 colors in the urn (some of which may be the same). Third, the population can be finite or infinite, the number being called the population size, usually denoted by $N$. 
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778,567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects. In Example 11 we can speak of the population of 12,000 heights (or weights) while in Example 14 we can speak of the population of all 200 colors in the urn (some of which may be the same). Third, the population can be finite or infinite, the number being called the population size, usually denoted by $N$. Similarly the number in the sample is called the sample size, denoted by $n$, and is generally finite.
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778,567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects. In Example 11 we can speak of the population of 12,000 heights (or weights) while in Example 14 we can speak of the population of all 200 colors in the urn (some of which may be the same). Third, the population can be finite or infinite, the number being called the population size, usually denoted by $N$. Similarly the number in the sample is called the sample size, denoted by $n$, and is generally finite. In Example 11, $N = 12,000$, $n = 100$, while in Example 13, $N$ is infinite, $n = 60$. 
Several things should be noted. First, the word population does not necessarily have the same meaning as in everyday language, such as “the population of Abuja is 778,567.” Second, the word population is often used to denote the observations or measurements rather than the individuals or objects. In Example 11 we can speak of the population of 12,000 heights (or weights) while in Example 14 we can speak of the population of all 200 colors in the urn (some of which may be the same). Third, the population can be finite or infinite, the number being called the population size, usually denoted by $N$. Similarly the number in the sample is called the sample size, denoted by $n$, and is generally finite. In Example 11, $N = 12,000$, $n = 100$, while in Example 13, $N$ is infinite, $n = 60$.

**Definition 15 (Population)**

A set of numbers from which a sample is drawn is referred to as a *population*. The distribution of the numbers constituting a population is called *population distribution*. 
If we draw an object from an urn, we have the choice of replacing or not replacing the object into the urn before we draw again.
Sampling With and Without Replacement

If we draw an object from an urn, we have the choice of replacing or not replacing the object into the urn before we draw again. In the first case a particular object can come up again and again, whereas in the second it can come up only once.
If we draw an object from an urn, we have the choice of replacing or not replacing the object into the urn before we draw again. In the first case a particular object can come up again and again, whereas in the second it can come up only once. Sampling where each member of a population may be chosen more than once is called sampling with replacement, while sampling where each member cannot be chosen more than once is called sampling without replacement.
Sampling With and Without Replacement

If we draw an object from an urn, we have the choice of replacing or not replacing the object into the urn before we draw again. In the first case a particular object can come up again and again, whereas in the second it can come up only once. Sampling where each member of a population may be chosen more than once is called sampling with replacement, while sampling where each member cannot be chosen more than once is called sampling without replacement.

A finite population that is sampled with replacement can theoretically be considered infinite since samples of any size can be drawn without exhausting the population.
Sampling With and Without Replacement

If we draw an object from an urn, we have the choice of replacing or not replacing the object into the urn before we draw again. In the first case a particular object can come up again and again, whereas in the second it can come up only once. Sampling where each member of a population may be chosen more than once is called sampling with replacement, while sampling where each member cannot be chosen more than once is called sampling without replacement.

A finite population that is sampled with replacement can theoretically be considered infinite since samples of any size can be drawn without exhausting the population. For most practical purposes, sampling from a finite population that is very large can be considered as sampling from an infinite population.
Random Samples

Clearly, the reliability of conclusions drawn concerning a population depends on whether the sample is properly chosen so as to represent the population sufficiently well, and one of the important problems of statistical inference is just how to choose a sample.
Random Samples

Clearly, the reliability of conclusions drawn concerning a population depends on whether the sample is properly chosen so as to represent the population sufficiently well, and one of the important problems of statistical inference is just how to choose a sample. One way to do this for finite populations is to make sure that each member of the population has the same chance of being in the sample, which is then often called a *random sample*.
Random Samples

Clearly, the reliability of conclusions drawn concerning a population depends on whether the sample is properly chosen so as to represent the population sufficiently well, and one of the important problems of statistical inference is just how to choose a sample. One way to do this for finite populations is to make sure that each member of the population has the same chance of being in the sample, which is then often called a random sample.

**Definition 16 (Random Sample)**

If $X_1, X_2, \cdots, X_n$ are independent and identically distributed random variables, we say that they constitute a random sample from the infinite population given by their common distribution.
If \( f(x_1, x_2, \cdots, x_n) \) is the value of the joint distribution of such set of random variables at \((x_1, x_2, \cdots, x_n)\), by virtue of independence we can write
\[
    f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i)
\]
where \( f(x_i) \) is the value of the population distribution at \( x_i \).
Statistical inferences are usually based on *statistics*, that is, on random variables that are functions of a set of random variables $X_1, X_2, \cdots, X_n$ constituting a random sample.
Statistical inferences are usually based on statistics, that is, on random variables that are functions of a set of random variables $X_1, X_2, \cdots, X_n$ constituting a random sample. Typical of what we mean by “statistic” are the sample mean and sample variance.
Statistical inferences are usually based on statistics, that is, on random variables that are functions of a set of random variables $X_1, X_2, \cdots, X_n$ constituting a random sample. Typical of what we mean by “statistic” are the sample mean and sample variance.

**Definition 17 (Sample Mean and Sample Variance)**

If $X_1, X_2, \cdots, X_n$ are constitute a random sample, then the *sample mean* is given by

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

and the *sample variance* is given by

$$
S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n - 1}.
$$
It is common practice also to apply the terms “random sample”, “statistics”, “sample mean” and “sample variance” to the values of the random variables instead of the random variables themselves.
It is common practice also to apply the terms “random sample”, “statistics”, “sample mean” and “sample variance” to the values of the random variables instead of the random variables themselves. Intuitively, this makes more sense and it conforms with colloquial usage.
It is common practice also to apply the terms “random sample”, “statistics”, “sample mean” and “sample variance” to the values of the random variables instead of the random variables themselves. Intuitively, this makes more sense and it conforms with colloquial usage. Thus we may calculate

\[\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}\]  
and  
\[s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}\]

for observed sample data and refer to these statistics as the sample mean and the sample variance.
It is common practice also to apply the terms “random sample”, “statistics”, “sample mean” and “sample variance” to the values of the random variables instead of the random variables themselves. Intuitively, this makes more sense and it conforms with colloquial usage. Thus we may calculate

\[ \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n - 1} \]

for observed sample data and refer to these statistics as the sample mean and the sample variance. Here, \( x_i, \overline{x}, \) and \( s^2 \) are values of the corresponding random variables \( X_i, \overline{X}, \) and \( S^2. \)
It is common practice also to apply the terms “random sample”, “statistics”, “sample mean” and “sample variance” to the values of the random variables instead of the random variables themselves. Intuitively, this makes more sense and it conforms with colloquial usage. Thus we may calculate

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}
\]

for observed sample data and refer to these statistics as the sample mean and the sample variance. Here, \(x_i, \bar{x},\) and \(s^2\) are values of the corresponding random variables \(X_i, \overline{X},\) and \(S^2.\) Indeed, the formula for \(\bar{x}\) and \(s^2\) are used even when we deal with any kind of data, not necessarily sample data, in which case we refer to \(\bar{x}\) and \(s^2\) simply as the mean and the variance.
Example 18

If a sample of size 5 results in the sample values 7, 9, 1, 6, 2, then the sample mean is
Example 18

If a sample of size 5 results in the sample values 7, 9, 1, 6, 2, then the sample mean is

\[
\bar{x} = \frac{7 + 9 + 1 + 6 + 2}{5} = 5.
\]
Outline

1. Measures of Central Tendency

2. Sampling Distributions
   - Population and Sample, Statistical Inference
   - Sampling With and Without Replacement
   - Random Samples

3. The Sampling Distribution of the Mean

4. The Sampling Distribution of the Mean: Finite Population
Let $f(x)$ be the probability distribution of some given population from which we draw a sample of size $n$. 
Let $f(x)$ be the probability distribution of some given population from which we draw a sample of size $n$. Then it is natural to look for the probability distribution of the sample statistic $\bar{X}$, which is called the sampling distribution for the sample mean, or the sampling distribution of means.
Let $f(x)$ be the probability distribution of some given population from which we draw a sample of size $n$. Then it is natural to look for the probability distribution of the sample statistic $\bar{X}$, which is called the sampling distribution for the sample mean, or the sampling distribution of means. The following theorems are important in this connection.
Theorem 19

If $X_1, X_2, \cdots, X_n$ are constitute a random sample from an infinite population with mean $\mu$ and the variance $\sigma^2$, then

$$E(\overline{X}) = \mu \text{ and } \text{var}(\overline{X}) = \frac{\sigma^2}{n}.$$
Theorem 19

If $X_1, X_2, \cdots, X_n$ are constitute a random sample from an infinite population with mean $\mu$ and the variance $\sigma^2$, then

$$E(\bar{X}) = \mu \text{ and } \text{var}(\bar{X}) = \frac{\sigma^2}{n}.$$  

Proof. $X_1, X_2, \cdots, X_n$ are random variables having the same distribution as the population, which has mean $\mu$, we have

$$E(X_k) = \mu, \; k = 1, 2, \cdots n.$$
Theorem 19

If \( X_1, X_2, \cdots, X_n \) are constitute a random sample from an infinite population with mean \( \mu \) and the variance \( \sigma^2 \), then

\[
E(\bar{X}) = \mu \text{ and } \text{var}(\bar{X}) = \frac{\sigma^2}{n}.
\]

Proof. \( X_1, X_2, \cdots, X_n \) are random variables having the same distribution as the population, which has mean \( \mu \), we have

\[
E(X_k) = \mu, \quad k = 1, 2, \cdots n.
\]

Then since the sample mean is defined as

\[
\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}
\]

we have as required
**Theorem 19**

If \(X_1, X_2, \cdots, X_n\) are constitute a random sample from an infinite population with mean \(\mu\) and the variance \(\sigma^2\), then

\[
E(\overline{X}) = \mu \text{ and } \text{var}(\overline{X}) = \frac{\sigma^2}{n}.
\]

**Proof.** \(X_1, X_2, \cdots, X_n\) are random variables having the same distribution as the population, which has mean \(\mu\), we have

\[
E(X_k) = \mu, \quad k = 1, 2, \cdots, n.
\]

Then since the sample mean is defined as

\[
\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}
\]

we have as required

\[
E(\overline{X}) = \frac{1}{n}[E(X_1) + E(X_2) + \cdots + E(X_n)] = \frac{1}{n}(n\mu) = \mu.
\]
On the other hand, since $X_1, X_2, \cdots, X_n$ are independent and

$$
\overline{X} = \frac{X_1}{n} + \frac{X_2}{n} + \cdots + \frac{X_n}{n}
$$

we have that
On the other hand, since $X_1, X_2, \cdots, X_n$ are independent and

$$
\bar{X} = \frac{X_1}{n} + \frac{X_2}{n} + \cdots + \frac{X_n}{n}
$$

we have that

$$
\text{var}(\bar{X}) = \frac{1}{n^2} \text{var}(X_1) + \frac{1}{n^2} \text{var}(X_2) + \cdots + \frac{1}{n^2} \text{var}(X_n) = n \left( \frac{1}{n^2} \sigma^2 \right) = \frac{\sigma^2}{n}.
$$
Example 20

A population consists of three housing units, where the value of $X$, the number of rooms for rent in each unit, is shown in the illustration.

Consider drawing a random sample of size 2 with replacement. Denote by $X_1$ and $X_2$ the observation of $X$ obtained in the first and second drawing, respectively. (a) Find the sampling distribution of $\overline{X} = (X_1 + X_2)/2$. (b) Calculate the mean and standard deviation for the population distribution and for the distribution of $\overline{X}$. Verify the relation $E(\overline{X}) = \mu$ and $\sigma_{\overline{X}} = \sigma/\sqrt{n}$. 
Solution.

The population distribution of $X$ given in the following table, which formalizes the fact that each of the $X$ values 2, 3 and 4 occurs in $1/3$ of the population of housing units.

<table>
<thead>
<tr>
<th>$x$</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(x)$</td>
<td>$1/3$</td>
<td>$1/3$</td>
<td>$1/3$</td>
</tr>
</tbody>
</table>

The Population Distribution

Because each unit is equally likely to be selected, the observation $X_1$ from the first drawing has the same distribution as given in the following table. Since the sampling is with replacement, the second observation $X_2$ also has this same distribution.
Solution.

The possible samples \((x_1, x_2)\) of size 2 and the corresponding values of \(\bar{X}\) are

\[
\begin{array}{c|cccccccc}
(x_1, x_2) & (2, 2) & (2, 3) & (2, 4) & (3, 2) & (3, 3) & (3, 4) & (4, 2) & (4, 3) & (4, 4) \\
\bar{X} = \frac{x_1 + x_2}{2} & 2 & 2.5 & 3 & 2.5 & 3 & 3.5 & 3 & 3.5 & 4 \\
\end{array}
\]

The nine possible samples are equally likely so, for instance \(P(\bar{X} = 2.5) = 2/9\). Continuing in this manner, we obtain the distribution of \(\bar{X}\) is

\[
\begin{array}{c|cccccc}
\text{Value of } \bar{X} & 2 & 2.5 & 3 & 3.5 & 4 \\
\text{Probability} & 1/9 & 2/9 & 3/9 & 2/9 & 1/9 \\
\end{array}
\]

The Probability Distribution of \(\bar{X}\)
Solution.

Population Distribution.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$f(x)$</th>
<th>$xf(x)$</th>
<th>$x^2f(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/3</td>
<td>2/3</td>
<td>4/3</td>
</tr>
<tr>
<td>3</td>
<td>1/3</td>
<td>3/3</td>
<td>9/3</td>
</tr>
<tr>
<td>4</td>
<td>1/3</td>
<td>4/3</td>
<td>16/3</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>3</td>
<td>29/3</td>
</tr>
</tbody>
</table>

$\mu = 3, \quad \sigma^2 = \frac{29}{3} - 3^2 = \frac{2}{3}$

Distribution of $\bar{X}$.

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$f(\bar{x})$</th>
<th>$\bar{x}f(\bar{x})$</th>
<th>$\bar{x}^2f(\bar{x})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/9</td>
<td>2/9</td>
<td>4/9</td>
</tr>
<tr>
<td>2.5</td>
<td>2/9</td>
<td>5/9</td>
<td>12.5/9</td>
</tr>
<tr>
<td>3</td>
<td>3/9</td>
<td>9/9</td>
<td>27/9</td>
</tr>
<tr>
<td>3.5</td>
<td>2/9</td>
<td>7/9</td>
<td>24.5/9</td>
</tr>
<tr>
<td>4</td>
<td>1/9</td>
<td>4/9</td>
<td>16/9</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>3</td>
<td>84/9</td>
</tr>
</tbody>
</table>

$E(\bar{X}) = 3 = \mu, \quad var(\bar{X}) = \frac{84}{9} - 3^2 = \frac{1}{3}$. 
It is customary to write \( E(\overline{X}) \) as \( \mu_{\overline{X}} \) and \( var(\overline{X}) \) as \( \sigma^2_{\overline{X}} \) and \( \sigma_{\overline{X}} \) as the \textit{standard error of the mean}. 
It is customary to write $E(\overline{X})$ as $\mu_{\overline{X}}$ and $\text{var}(\overline{X})$ as $\sigma^2_{\overline{X}}$ and $\sigma_{\overline{X}}$ as the standard error of the mean. The formula for the standard error of the mean, $\sigma_{\overline{X}} = \sigma / \sqrt{n}$, shows that the standard deviation of the distribution of $\overline{X}$ decreases when $n$, the sample size, is increases.
It is customary to write $E(\bar{X})$ as $\mu_{\bar{X}}$ and $\text{var}(\bar{X})$ as $\sigma^2_{\bar{X}}$ and $\sigma_{\bar{X}}$ as the standard error of the mean. The formula for the standard error of the mean, $\sigma_{\bar{X}} = \sigma / \sqrt{n}$, shows that the standard deviation of the distribution of $\bar{X}$ decreases when $n$, the sample size, is increases. This means that when $n$ becomes larger and we actually have more information (the values of more random variables), we can expect values of $\bar{X}$ to be closer to $\mu$, the quantity that they are intended to estimate.
It is customary to write $E(\overline{X})$ as $\mu_\overline{X}$ and $\text{var}(\overline{X})$ as $\sigma^2_\overline{X}$ and $\sigma_\overline{X}$ as the standard error of the mean. The formula for the standard error of the mean, $\sigma_\overline{X} = \sigma/\sqrt{n}$, shows that the standard deviation of the distribution of $\overline{X}$ decreases when $n$, the sample size, increases. This means that when $n$ becomes larger and we actually have more information (the values of more random variables), we can expect values of $\overline{X}$ to be closer to $\mu$, the quantity that they are intended to estimate. If we use Chebyshev’s theorem, we can express this formally in the following way:
It is customary to write $E(\bar{X})$ as $\mu_{\bar{X}}$ and $\text{var}(\bar{X})$ as $\sigma^2_{\bar{X}}$ and $\sigma_{\bar{X}}$ as the *standard error of the mean*. The formula for the standard error of the mean, $\sigma_{\bar{X}} = \sigma / \sqrt{n}$, shows that the standard deviation of the distribution of $\bar{X}$ decreases when $n$, the *sample size*, is increases. This means that when $n$ becomes larger and we actually have more information (the values of more random variables), we can expect values of $\bar{X}$ to be closer to $\mu$, the quantity that they are intended to estimate. If we use Chebyshev’s theorem, we can express this formally in the following way:

**Theorem 21 (Law of Large Numbers)**

*For any positive constant $c$, the probability that $\bar{X}$ will take on a value between $\mu - c$ and $\mu + c$ is at least*

$$1 - \frac{\sigma^2}{nc^2}.$$

*When $n \to \infty$, this probability approaches 1.*
Theorem 22 (Central Limit Theorem)

If $X_1, X_2, \cdots, X_n$ are constitute a random sample from an infinite population with mean $\mu$, the variance $\sigma^2$, and the moment-generating function $M_X(t)$, then the limiting distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

as $n \to \infty$ is the standard normal distribution.
Theorem 22 (Central Limit Theorem)

If $X_1, X_2, \cdots, X_n$ are constitute a random sample from an infinite population with mean $\mu$, the variance $\sigma^2$, and the moment-generating function $M_X(t)$, then the limiting distribution of

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

as $n \to \infty$ is the standard normal distribution.

Sometimes, the central limit theorem is interpreted incorrectly as implying that the distribution of $\overline{X}$ approaches a normal distribution when $n \to \infty$. 
Theorem 22 (Central Limit Theorem)

If \(X_1, X_2, \ldots, X_n\) are constitute a random sample from an infinite population with mean \(\mu\), the variance \(\sigma^2\), and the moment-generating function \(M_X(t)\), then the limiting distribution of

\[
Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}
\]

as \(n \to \infty\) is the standard normal distribution.

Sometimes, the central limit theorem is interpreted incorrectly as implying that the distribution of \(\overline{X}\) approaches a normal distribution when \(n \to \infty\). This is incorrect because \(\text{var}(\overline{X}) \to 0\) when \(n \to \infty\); on the other hand, the central limit theorem does justify approximating the distribution of \(\overline{X}\) with a normal distribution having the mean \(\mu\) and the variance \(\sigma^2/n\) when \(n\) is large.
Theorem 22 (Central Limit Theorem)

If $X_1, X_2, \cdots, X_n$ are constitute a random sample from an infinite population with mean $\mu$, the variance $\sigma^2$, and the moment-generating function $M_X(t)$, then the limiting distribution of

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

as $n \to \infty$ is the standard normal distribution.

Sometimes, the central limit theorem is interpreted incorrectly as implying that the distribution of $\overline{X}$ approaches a normal distribution when $n \to \infty$. This is incorrect because $\text{var}(\overline{X}) \to 0$ when $n \to \infty$; on the other hand, the central limit theorem does justify approximating the distribution of $\overline{X}$ with a normal distribution having the mean $\mu$ and the variance $\sigma^2/n$ when $n$ is large. In practice, this approximation is used when $n \geq 30$ regardless of the actual shape of the population sampled.
Example 23

A soft drink vending machine is set so that the amount of drink dispensed is a random variable with mean of 200 milliliters and a standard deviation of 15 milliliters. What is the probability that the average (mean) amount dispensed in a random sample of size 36 is at least 204 milliliters?
Example 23

A soft drink vending machine is set so that the amount of drink dispensed is a random variable with mean of 200 milliliters and a standard deviation of 15 milliliters. What is the probability that the average (mean) amount dispensed in a random sample of size 36 is at least 204 milliliters?

Solution. According to Theorem 19, the distribution of $\bar{X}$ has the mean $\mu_{\bar{X}} = 200$ and the standard deviation $\sigma_{\bar{X}} = \frac{15}{\sqrt{36}} = 2.5$, and according to the central limit theorem, this distribution is approximately normal. Since $z = \frac{204 - 200}{2.5} = 1.6$, we see that

$$P(\bar{X} \geq 204) \approx P(Z \geq 1.6) = 0.5 - 0.4452 = 0.0548.$$
It is of interest to note that when the population we are sampling is normal, the distribution of $\bar{X}$ is a normal distribution regardless of the size of $n$. 
It is of interest to note that when the population we are sampling is normal, the distribution of $\bar{X}$ is a normal distribution regardless of the size of $n$.

**Theorem 24**

*If $\bar{X}$ is the mean of a random sample of size $n$ from a normal population with mean $\mu$ and the variance $\sigma^2$, its sampling distribution is a normal distribution with mean $\mu$ and the variance $\sigma^2/n$.***
Outline

1. Measures of Central Tendency
2. Sampling Distributions
   - Population and Sample. Statistical Inference
   - Sampling With and Without Replacement
   - Random Samples
3. The Sampling Distribution of the Mean
4. The Sampling Distribution of the Mean: Finite Population
If an experiment consists of selecting one or more values from a finite set of numbers \( \{ c_1, c_2, \cdots, c_N \} \), this set is referred to as a finite population of size \( N \).
If an experiment consists of selecting one or more values from a finite set of numbers \( \{c_1, c_2, \cdots, c_N\} \), this set is referred to as a finite population of size \( N \). In the definition that follows, it will be assumed that we are sampling without replacement from a finite population of size \( N \).
The Sampling Distribution of the Mean: Finite Population

If an experiment consists of selecting one or more values from a finite set of numbers \( \{c_1, c_2, \cdots, c_N\} \), this set is referred to as a finite population of size \( N \). In the definition that follows, it will be assumed that we are sampling without replacement from a finite population of size \( N \).

**Definition 25 (Random Sample-Finite Population)**

If \( X_1 \) is the first value drawn from a finite population of size \( N \), \( X_2 \) is the second value drawn, \( \ldots \), \( X_n \) is the \( n \)th value drawn, and the joint probability distribution of these \( n \) random variables is given by

\[
f(x_1, x_2, \cdots, x_n) = \frac{1}{N(N - 1) \cdots (N - n + 1)}
\]

for each ordered \( n \)-tuple of values of these random variables, then \( X_1, X_2, \cdots, X_n \) are said to constitute a random sample from the given finite population.
From the joint probability distribution of Definition 25, it follows that the probability for each subset \( n \) of the \( N \) elements of the finite population (regardless of the order in which the values are drawn) is

\[
\frac{n!}{N(N - 1)\cdots(N - n + 1)} = \frac{1}{\binom{N}{n}}.
\]
From the joint probability distribution of Definition 25, it follows that the probability for each subset $n$ of the $N$ elements of the finite population (regardless of the order in which the values are drawn) is

$$\frac{n!}{N(N - 1) \cdots (N - n + 1)} = \frac{1}{\binom{N}{n}}.$$

This is often given as an alternative definition or as a criterion for the selection of a random sample of size $n$ from a finite population size $N$:
From the joint probability distribution of Definition 25, it follows that the probability for each subset $n$ of the $N$ elements of the finite population (regardless of the order in which the values are drawn) is

$$\frac{n!}{N(N - 1)\cdots(N - n + 1)} = \frac{1}{\binom{N}{n}}.$$

This is often given as an alternative definition or as a criterion for the selection of a random sample of size $n$ from a finite population size $N$: Each of the $\binom{N}{n}$ possible samples must have the same probability.
From the joint probability distribution of Definition 25, it follows that the probability for each subset $n$ of the $N$ elements of the finite population (regardless of the order in which the values are drawn) is

$$\frac{n!}{N(N-1)\cdots(N-n+1)} = \frac{1}{\binom{N}{n}}.$$ 

This is often given as an alternative definition or as a criterion for the selection of a random sample of size $n$ from a finite population size $N$: Each of the $\binom{N}{n}$ possible samples must have the same probability.

It also follows from the joint probability distribution of 25 that the marginal distribution of $X_r$ is given by

$$f(x_r) = \frac{1}{N} \text{ for } x_r = c_1, c_2, \cdots, c_N$$

for $r = 1, 2, \cdots, n$, and we refer to the mean and the variance of this discrete uniform distribution as the mean and the variance of the finite population.
Definition 26 (Sample Mean and Variance-Finite Population)

The *sample mean* and the *sample variance* of the finite population \( \{c_1, c_2, \ldots, c_N\} \) are

\[
\mu = \sum_{i=1}^{N} c_i \frac{1}{N} \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{N} (c_i - \mu)^2 \frac{1}{N}.
\]
Definition 26 (Sample Mean and Variance-Finite Population)

The sample mean and the sample variance of the finite population \{c_1, c_2, \ldots, c_N\} are

\[
\mu = \frac{1}{N} \sum_{i=1}^{N} c_i \quad \text{and} \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (c_i - \mu)^2.
\]

Finally, it follows from the joint probability distribution of 25 that the joint marginal distribution of any two of the random variables \(X_1, X_2, \ldots, X_n\) is given by

\[
g(x_r, x_s) = \frac{1}{N(N - 1)}
\]

for each ordered pair of elements of the finite population.
Theorem 27

If $X_r$ and $X_s$ are the $r$th and $s$th random variables of a random sample of size $n$ drawn from the finite population $\{c_1, c_2, \cdots, c_N\}$, then

$$\text{cov}(X_r, X_s) = -\frac{\sigma^2}{N - 1}.$$
Theorem 27

If $X_r$ and $X_s$ are the $r$th and $s$th random variables of a random sample of size $n$ drawn from the finite population $\{c_1, c_2, \cdots, c_N\}$, then

$$\text{cov}(X_r, X_s) = -\frac{\sigma^2}{N - 1}.$$

Theorem 28

If $\bar{X}$ is the mean of a random sample of size $n$ taken without replacement from a finite population of size $N$ with mean $\mu$ and the variance $\sigma^2$, then

$$E(\bar{X}) = \mu \text{ and } \text{var}(\bar{X}) = \frac{\sigma^2}{n} \frac{N - n}{N - 1}.$$
It is of interest to note that the formulas we obtained for $\text{var}(\bar{X})$ in Theorems Thm9.8 and 28 differ only by the \textit{finite population correction factor} $\frac{N-n}{N-1}$. 
It is of interest to note that the formulas we obtained for $\text{var}(\bar{X})$ in Theorems Thm9.8 and 28 differ only by the finite population correction factor $\frac{N-n}{N-1}$. Indeed, when $N$ is large compared to $n$, the difference between the two formulas for $\text{var}(\bar{X})$ is usually negligible, and the formula $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ is often used as an approximation when we are sampling from a large finite population.
It is of interest to note that the formulas we obtained for \( \text{var}(\bar{X}) \) in Theorems Thm9.8 and 28 differ only by the finite population correction factor \( \frac{N-n}{N-1} \). Indeed, when \( N \) is large compared to \( n \), the difference between the two formulas for \( \text{var}(\bar{X}) \) is usually negligible, and the formula \( \sigma_{\bar{X}} = \sigma / \sqrt{n} \) is often used as an approximation when we are sampling from a large finite population. A general rule of thumb is to use this approximation when the sampling does not constitute more than 5 percent of the population.
Example 29

A population consists of the five numbers 2, 3, 6, 8, 11. Consider all possible samples of size two which can be drawn with replacement from this population. Find (a) the mean of the population, (b) the standard deviation of the population, (c) the mean of the sampling distribution of means, (d) the standard deviation of the sampling distribution of means, i.e., the standard error of means.
Example 29

A population consists of the five numbers 2, 3, 6, 8, 11. Consider all possible samples of size two which can be drawn with replacement from this population. Find (a) the mean of the population, (b) the standard deviation of the population, (c) the mean of the sampling distribution of means, (d) the standard deviation of the sampling distribution of means, i.e., the standard error of means.

Example 30

Solve Example 29 in case sampling is without replacement.
Example 31

Assume that the heights of 3000 male students at a university are normally distributed with mean 68.0 inches and standard deviation 3.0 inches. If 80 samples consisting of 25 students each are obtained, what would be the mean and standard deviation of the resulting sample of means if sampling were done (a) with replacement, (b) without replacement?
Example 31

Assume that the heights of 3000 male students at a university are normally distributed with mean 68.0 inches and standard deviation 3.0 inches. If 80 samples consisting of 25 students each are obtained, what would be the mean and standard deviation of the resulting sample of means if sampling were done (a) with replacement, (b) without replacement?

Example 32

In how many samples of Example 31 would you expect to find the mean (a) between 66.8 and 68.3 inches, (b) less than 66.4 inches?
Example 33

Five hundred ball bearings have a mean weight of 5.02 oz and a standard deviation of 0.30 oz. Find the probability that a random sample of 100 ball bearings chosen from this group will have a combined weight, (a) between 496 and 500 oz, (b) more than 510 oz.
Thank You!!!