Chapter 6: SAMPLING DISTRIBUTIONS

Read Section 1.5

Graphical methods may not always be sufficient for describing data. **Numerical measures** can be created for both *populations* and *samples*.

**Definition** A numerical descriptive measure calculated for a *population* is called a ________

A numerical descriptive measure calculated for a *sample* is called a ________.

<table>
<thead>
<tr>
<th>Population Parameter</th>
<th>Sample Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
</tr>
<tr>
<td>Binomial Proportion</td>
<td></td>
</tr>
</tbody>
</table>

Statistics vary from sample to sample and hence are random variables. The probability distributions for statistics are called ____________

**Definition** The **sampling distribution of a statistic** is the probability distribution for the possible values of the statistic that results when random samples of size $n$ are repeatedly drawn from the population.
Example A population consists of \( N = 5 \) numbers: 3, 6, 9, 12, 15. Draw samples of size \( n = 3 \) without replacement.

1. For each possible sample calculate the mean and the median.

<table>
<thead>
<tr>
<th></th>
<th>( \bar{x} )</th>
<th>median</th>
<th></th>
<th>( \bar{x} )</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td>3, 6, 9</td>
<td>8</td>
<td>6</td>
<td>3, 12, 15</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>3, 6, 12</td>
<td></td>
<td></td>
<td>6, 9, 12</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>3, 6, 15</td>
<td>8</td>
<td>6</td>
<td>6, 9, 15</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>3, 9, 12</td>
<td>8</td>
<td>9</td>
<td>6, 12, 15</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>3, 9, 15</td>
<td>9</td>
<td>9</td>
<td>9, 12, 15</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

2. Find the sampling distributions for the sample mean and the sample median
Consider all possible samples of size $n$ that can be drawn from a given population. For each sample we can compute the sample mean. In this way we obtain the **sampling distribution of means** or the **sampling distribution of the mean**.

**Example** A population consists of $N = 5$ numbers: 3, 6, 9, 12, 15. Draw samples of size $n = 3$ without replacement. Then, let $\bar{x}$ the sample mean. In the previous page we had:

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$p(\bar{x})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>$\frac{1}{10}$</td>
</tr>
<tr>
<td>7</td>
<td>$\frac{1}{10}$</td>
</tr>
<tr>
<td>8</td>
<td>$\frac{2}{10}$</td>
</tr>
<tr>
<td>9</td>
<td>$\frac{2}{10}$</td>
</tr>
<tr>
<td>10</td>
<td>$\frac{2}{10}$</td>
</tr>
<tr>
<td>11</td>
<td>$\frac{1}{10}$</td>
</tr>
<tr>
<td>12</td>
<td>$\frac{1}{10}$</td>
</tr>
</tbody>
</table>

Find the mean of the population and the mean of the random variable $\bar{x}$. What do you notice?
Theorem: We denote the mean and the standard deviation of the sampling distribution of means by \( \mu_x \) and \( \sigma_x \) and the population mean and standard deviation by \( \mu \) and \( \sigma \), respectively. If the population is infinite or sampling is with replacement then

\[
\mu_x = \quad \sigma_x =
\]

The Central Limit Theorem

If random samples of \( n \) observations are drawn from a nonnormal population with finite \( \mu \) and standard deviation \( \sigma \), then, when \( n \) is large, the sampling distribution of the sample mean \( \bar{x} \) is approximately normally distributed, with mean \( \mu \) and standard deviation \( \frac{\sigma}{\sqrt{n}} \). The approximation becomes more accurate as \( n \) becomes large.

NOTE:

✓ The Central Limit Theorem also implies that the sum of \( n \) measurements is approximately normal with mean \( n\mu \) and standard deviation \( \sigma\sqrt{n} \)

✓ Many statistics that are used for statistical inference are \textbf{sums} or \textbf{averages} of sample measurements.

✓ When \( n \) is large, these statistics will have approximately normal distributions.

How Large is Large?

• If the sample is \textit{normal}, then the sampling distribution of \( \bar{x} \) will also be normal, no matter what the sample size.

• When the sample population is approximately \textit{symmetric}, the distribution becomes approximately normal for relatively small values of \( n \).

• When the sample population is \textit{skewed}, the sample size must be at least 30 before the sampling distribution of \( \bar{x} \) becomes approximately normal.
The Sampling Distribution of the Sample Mean

A random sample of size \( n \) is selected from a population with mean \( \mu \) and standard deviation \( \sigma \).

✓ The sampling distribution of the sample mean \( \bar{x} \) will have:

Mean: _____ and standard deviation: _____

✓ If the original population is normal, the sampling distribution will be normal for any sample size.

✓ If the original population is nonnormal, the sampling distribution will be normal when \( n \) is large.

✓ If the sampling distribution of \( \bar{x} \) is normal or approximately normal, standardize or rescale the interval of interest in terms of

\[
z =
\]

✓ Find the appropriate area using Table 3.

Example A random sample of size \( n = 16 \) from a normal distribution with \( \mu = 10 \) and \( \sigma = 8 \). Find \( P(\bar{x} > 12) \)
Example  A soda filling machine is supposed to fill cans of soda with 12 fluid ounces. Suppose that
the fills are actually normally distributed with a mean of 12.1 oz and a standard deviation of .2 oz.
What is the probability that the average fill for a 6-pack of soda is less than 12 oz?

Example  The weights of trout at a trout farm are normally distributed with mean 1 Kg and standard
deviation 0.25 Kg.

1. Find the probability that a trout chosen at random will weigh more than 1.25 Kg?

2. Find the probability that the mean weight of a sample of 10 trout chosen at random will be
less than 0.9 Kg?
Example The scores on a general test follow a normal distribution with mean 450 and standard deviation 50. It is highly desirable to score over 480 on this exam. In one location 25 people sign up to take the exam. The average score of these 25 people exceeds 490. Is this odd? Should the test center investigate?

Example Consider the height of an adult male. It’s approximately normally distributed with mean 70 inches and standard deviation 4. Sixteen adult males are in a pit which is 98 feet deep. They decide to stand on one another (feet to head), hoping that the person on top can grip the top of the pit and get out, and, hence go for help. What’s the probability that their plan succeeds?
There are many practical examples of the binomial random variable $x$. One common application involves consumer preference or opinion polls, in which we use a random sample of $n$ people to estimate the proportion $p$ of people in the population who have a specific characteristic. If $x$ of the sampled people have this characteristic, then the sample proportion

$$\hat{p} = \frac{x}{n}$$

can be used to estimate the population proportion $p$.

Then, the sampling distribution of the sample proportion $\hat{p}$ will have:

Mean: $E(\hat{p}) = \mu_{\hat{p}}$

Standard Deviation: $\sigma_{\hat{p}}$

NOTE: When the sample size $n$ is large, the sampling distribution of $\hat{p}$ can be approximated by a normal distribution. The approximation will be adequate if $np \geq 15$ and $nq \geq 15$.

Example Out of 300 students in the school, 225 passed an exam. What would be the mean of the sampling distribution of the proportion of students who passed the exam in the school?

You take a sample of 10 of these students. What is the standard deviation of $\hat{p}$?
Example: According to a recent poll, 66% of adults who use the Internet have paid to download music. In a random sample of $n = 1,000$ adults who use the Internet, let $\hat{p}$ the proportion who have paid to download music. Answer the following:

1. What are the mean and standard deviation of $\hat{p}$?

2. Describe the shape of the sampling distribution of $\hat{p}$.

3. What is the probability that the sample proportion lies between 60% and 70%?

4. If someone tells you that more than 85% of adults paid to download music from the internet, do you believe it?
Example A random sample of $n = 258$ measurements is drawn from a binomial population with probability of success $0.76$.

1. Find the mean and standard deviation of $\hat{p}$

2. Describe the shape of the sampling distribution of $\hat{p}$.

3. Find $P(\hat{p} < .9)$
Example: The proportion of individuals with an Rh–positive blood type is 85%. You have a random sample of 500 individuals.

1. What are the mean and standard deviation of \( \hat{p} \), the sample proportion with Rh–positive blood type?

2. Is the distribution of \( \hat{p} \) approximately normal? Justify your answer.

3. What is the probability that the sample proportion lies between 83% and 88%?

4. 95% of the time, the sample proportion would lie between what limits?
Example  The contents of bags of oats are normally distributed with mean 3.05 lbs and standard deviation 0.08 lbs. Suppose that 36 bags are selected at random.

1. Describe the shape of the sampling distribution of $\bar{x}$?

2. Find $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$.

3. What is the probability that the sample has a mean weight of contents between 3 lbs and 3.15 lbs?
Example A city currently does not have a National Football League team. 54% of all the city's residents are in favor of attracting an NFL team. A random sample of 1000 of the city's residents is selected, and asked if they would want an NFL team.

1. Describe the shape of the sampling distribution of \( \hat{p} \)

2. Find \( E(\hat{p}) \) and \( \sigma_{\hat{p}} \).

3. What is the probability the percentage of those residents polled who are in favor of attracting an NFL team is less than 50%?

HOMEWORK

Chapter 6 # 21, 27, 31, 37, 38(a,b,c), 40, 41, 49, 51, 55, 57, 59, 71, 73, 80.