Chapter 9

Inferring Population Means

Sample Means of Random Samples

- In Chapter 7, we learned to estimate population parameters by collecting a random sample from that population.
  - We used the collected data to calculate a statistic, and this statistic was used to estimate the parameter.
  - We used the statistic $\hat{p}$ to estimate the parameter p.
- In Chapter 9, we're going to learn to use the statistic $\overline{x}$ to estimate the parameter μ.
  - In general, we want to know how close our estimate is to the truth, i.e. we need to know how far away that statistic is, typically, from the parameter.

Sample Means of Random Samples

- Just as we did in Chapter 7 with $\hat{p}$, we now examine three characteristics of the behavior of the sample mean: its accuracy, its precision, and its probability distribution.
- By understanding these characteristics, we'll be able to measure how well our estimate performs and thus make better decisions.

Statistic (based on data)		Parameter (typically unknown)	
Sample mean	$\bar{x}$	Population mean	μ (mu)
Sample standard deviation	s	Population standard deviation	σ (sigma)
Sample variance	s^2	Population variance	σ^2
Sample proportion	$\hat{p}$ (p-hat)	Population proportion	p

Accuracy and Precision of a Sample Mean

- The reason why the sample mean is a useful estimator for the population mean is that the sample mean is accurate and, with a sufficiently large sample size, very precise.
- The accuracy of an estimator is measured by the bias, and the precision is measured by the standard error.
- The sample mean is unbiased (bias is 0) when estimating the population mean, i.e. on average, the sample mean is the same as the population mean.
- The precision of the sample mean depends on the variability in the population, but the more observations we collect, the more precise the sample mean becomes.

Example

Clicker!

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- Is the mean value of 243 seconds an example of a parameter or statistic?

A. Parameter

B. Statistic

Example

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- Is the mean value of 243 seconds an example of a parameter or statistic?

It is a parameter. Parameter is a numerical value that characterizes some aspect of the population.

Example

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- What should the student expect the average song length to be for his playlist?

Example

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- What should the student expect the average song length to be for his playlist?

The sample mean length can vary, but it is typically the same as the population mean of 243 seconds.

Example

Clicker!

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- What is the standard error for the mean song length of 25 randomly selected songs?

$$SE = \frac{\sigma}{\sqrt{n}}$$

A. $\frac{243}{\sqrt{25}}$

B. $\frac{93}{\sqrt{25}}$

C. $\frac{25}{\sqrt{25}}$

Example

A student's iTunes music library has a very large number of songs. The mean length of the songs is 243 seconds, and the standard deviation is 93 seconds. Using his digital music player, this student will create a playlist that consists of 25 randomly selected songs.

- What is the standard error for the mean song length of 25 randomly selected songs?

$$\frac{\sigma}{\sqrt{n}} = \frac{93}{\sqrt{25}} = 18.6 \text{ seconds}$$

Central Limit Theorem - Means

- If we draw repeated random samples of the same size, n, from some population and measure the mean, $\bar{x}$, we see in each sample, then the collection of means will pile up around the underlying population mean, μ.
- A histogram of the sample means can be modeled well by a Normal model.
- The Central Limit Theorem (CLT) states that the distribution of $\bar{x}$ is approximately Normal with mean equal to the population, μ, and the standard deviation (standard error, since its a sampling distribution) is equal to

$$\bar{x} \sim N\left(\text{mean}\left(\bar{x}\right) = \mu,\ \text{SD}\left(\bar{x}\right) = \frac{\sigma}{\sqrt{n}}\right)$$

Check the Conditions for CLT

- Random and Independent: The sample is collected randomly and the trials are independent of each other.
- Large Sample:
    - If the population distribution is normal, then the sampling will be normal as well, regardless of the sample size.
    - OR if the population distribution is not normal, then we need a large enough sample ($n \geq 25$) to ensure that the sampling distribution will be normal.
- Big Population: If the sample is collected without replacement, then the population size is at least 10 times the sample size.

The t-distribution

- The hypothesis tests and confidence intervals that we will use for estimating and testing the mean are based on a statistic called t-statistic:

$$t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}$$

- Typically, we almost never know the value of the population standard deviation, σ.
- Instead, we replace it with an estimate: the sample standard deviation, s.

The t-distribution

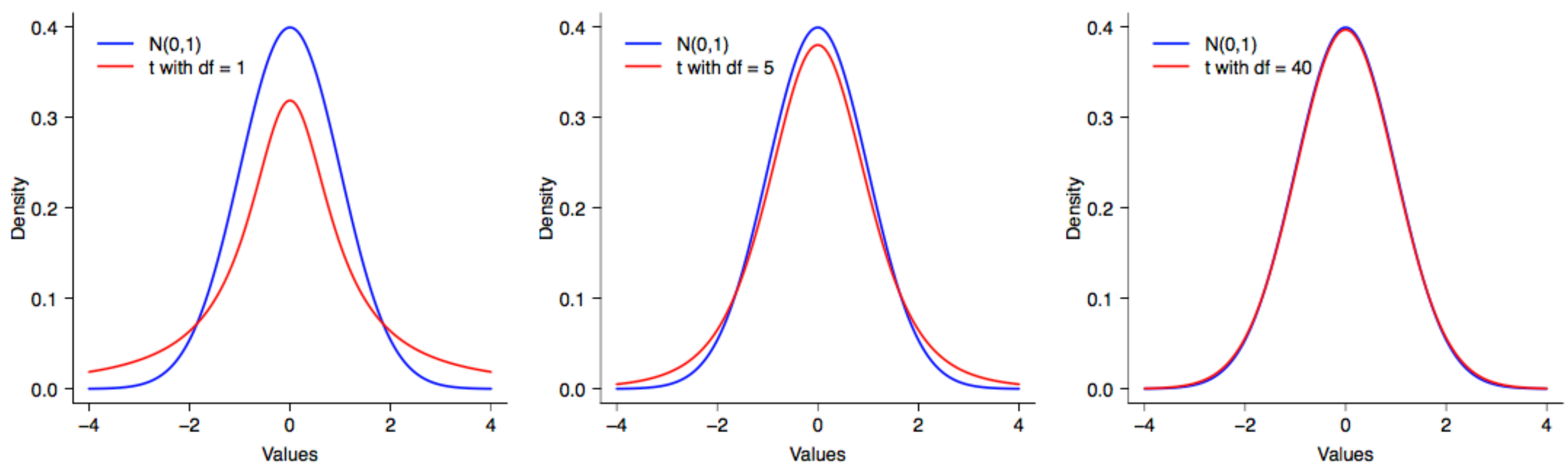
- The t-statistic does not follow the Normal distribution.
- The t-statistic is more variable that the z-statistic. One reason is that the denominator changes with each sample.
- Instead, if the three conditions for using the CLT hold, the t-statistic follows a distribution called the t-distribution.
- The t-distribution is a better model than the Normal for the sampling distribution of $\bar{x}$ when σ is unknown.

The t-distribution

- The t-distribution shares many characteristics with the N(0,1) distribution.
    - Both are symmetric, unimodal and might be described as “bell-shaped”.
    - The difference is that the t-distribution has thicker tails. This means that it is more likely that we will see extreme values (values far from 0).

The t-distribution

- The t-distribution's shape depends on the degrees of freedom (df).
    - The number of degrees of freedom is usually an integer: 1, 2, 3, etc.
    - If df is small, then the t-distribution has thick tails. The larger the df, the thinner the tails. When df is infinitely large, the t-distribution is exactly the same as N(0,1) distribution.



CI and Hypothesis Tests for Means

- Two techniques are used for answering questions about the population mean.
  - Confidence intervals
  - Hypothesis tests
- Confidence intervals are used for estimating parameter values.
- Hypothesis tests are used for deciding whether a parameter's value is one thing or another.

Confidence Intervals for Means

- We estimate the unknown σ (population standard deviation) by the known s (sample standard deviation) and calculate the standard error of $\bar{x}$.

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}$$

- The confidence interval for a mean is:

$$\bar{x} \pm t^{*}_{df} \frac{s}{\sqrt{n}}$$

where df = n - 1 and t* is found using the t-distribution.

Example

A researcher is interested in estimating how much college students spend on airfare in a year. A random sample of 20 college students yielded a mean of $555 with a standard deviation of $210. Assume airfare is distributed nearly normal. Estimate the true population mean of the amount of money college students spend on airfare in a year with a 95% confidence interval.

First, we must check the conditions to use the CLT.

Example

A researcher is interested in estimating how much college students spend on airfare in a year. A random sample of 20 college students yielded a mean of $555 with a standard deviation of $210. Assume airfare is distributed nearly normal. Estimate the true population mean of the amount of money college students spend on airfare in a year with a 95% confidence interval.

Random and independent: We have a random sample of 20 college students and we assume the students are independent of each other due to the random sample.

Large sample: If the population distribution is normal, then the sampling will be normal as well, regardless of the sample size.

Big population: If the sample is collected without replacement, then the population size is at least 10 times the sample size. 10 x 20 = 200.

Example

A researcher is interested in estimating how much college students spend on airfare in a year. A random sample of 20 college students yielded a mean of $555 with a standard deviation of $210. Assume airfare is distributed nearly normal. Estimate the true population mean of the amount of money college students spend on airfare in a year with a 95% confidence interval.

Next, find the 95% confidence interval.

$$\bar{x} = 555,\ s=210,\ n=20$$

Example

Next, find the 95% confidence interval.

$\bar{x}$ = 555, s=210, n=20

$$\bar{x} \pm t^*_{df} \frac{s}{\sqrt{n}}$$

$$df = n - 1 = 20 - 1 = 19$$

$$555 \pm 2.093 \frac{210}{\sqrt{20}}$$

$$(457, 653)$$

	Confidence Level			
	80%	90%	95%	98%
	Right-Tail Probability			
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528

Example

A researcher is interested in estimating how much college students spend on airfare in a year. A random sample of 20 college students yielded a mean of $555 with a standard deviation of $210. Assume airfare is distributed nearly normal. Estimate the true population mean of the amount of money college students spend on airfare in a year with a 95% confidence interval.

Finally, interpret the confidence interval.

We are 95% confident that the true population mean of the amount of money all college students spend on airfare in a year is between $457 and $653.

Confidence Intervals

- The confidence level is a measure of how well the method used to produce the confidence interval performs.
- We can interpret the confidence level to mean that if we were to take many random samples of the same size from the same population, and for each random sample calculate a confidence interval, then the confidence level is the proportion of intervals that contain the population parameter.
- We prefer confidence intervals that have 90% or higher confidence levels because then we know that the process that produced these levels is a good process.
- A confidence level, such as 90%, is not a probability. Saying we are 90% confident the mean is between 21.1 and 21.3 minutes doesn't mean that there is a 90% chance that the mean is between these two values. It either is or is not. There is no probability about it.

Hypothesis Testing for One Mean

- We use the same steps that we used in Chapter 8 for hypothesis testing (with some adjustments for the mean).
- Steps for hypothesis testing for one mean:
    1. Write the null and alternative hypotheses.
    2. Distinguish the type of test and check the CLT conditions.
    3. Find the test statistic.
    4. Find the p-value.
    5. Interpret the results.

Step 1: Hypothesize

- First we find the null hypothesis and alternative hypothesis.
- Hypotheses are always statements about the population parameter.
- In this chapter, we focus on the mean of the population.
- The exact form of the alternative hypothesis depends on the research question.

Two-Tailed	One-Tailed (Left)	One-Tailed (Right)
$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu > \mu_0$

Step 2: Type of test and CLT conditions

- Next, we want to distinguish which type of test we will use.
  - Since we have one sample, we will be using the one-sample t-test.
  - Since we typically don't have σ we will be using the t-test.
- We continue to check the CLT condition (same as the ones from last lecture):
  - Random and Independent: The sample is collected randomly and the trials are independent of each other.
  - Large Sample: If the population distribution is normal, then the sampling will be normal as well, regardless of the sample size. If not, the sample size needs to be larger than or equal to 25.
  - Big Population: If the sample is collected without replacement, then the population size is at least 10 times the sample size.

Step 3: The Test Statistic

- Now, we find the test statistic.
  - We know we will need to use the one-sample t-test hence we will need to find the t-statistic.
  - The idea is to compare the observed value of the sample mean to the value claimed by the null hypothesis.

$$t = \frac{\bar{x} - \mu}{SE}$$

where $SE = \frac{s}{\sqrt{n}}$

Step 3: The Test Statistic

- The test statistic works because it compares the value of the parameter that the null hypothesis says is true to the estimate of that value that we actually observed in our data.
- If the estimate is close to the null hypothesis value, then the t-statistic is close to 0.
- If the estimate is far from the null hypothesis value, then the t-statistic is far from 0.
- The farther the t-statistic is from 0, the worse things look for the null hypothesis.

Step 4: The p-value

- How unusual is a value, according to the null hypothesis?
- The p-value tells us exactly that! It is the probability of getting a t-statistic as extreme as or more extreme than what we observed.
- The p-value is found using the test statistic and the degrees of freedom which is df=n-1.
- The p-value can be computed with technology or by using the t-table in the appendix of your textbook.

Step 5: Interpret the results

- We compare the p-value to a given α:
  - If p-value < α, we reject the H_0.
    - There is sufficient evidence to suggest that H_a is plausible.
  - If p-value > α, we fail to reject the H_0.
    - There is not sufficient evidence to suggest that H_a is plausible.
- Lastly, write the results in context.

Hypothesis Testing for One Mean

A research conducted a few years ago showed that the mean amount of money spent on airfare by college students was \$555. A survey given to 20 randomly selected college students yielded a mean of \$570 with a standard deviation of \$40.

- Is there significant evidence to suggest that the mean amount of money spent on airfare by college students has increased over the years? Assume that the distribution of amount of money spent on airfare by college students is nearly normal. Use significance level 0.01.

Hypothesis Testing for One Mean

A research conducted a few years ago showed that the mean amount of money spent on airfare by college students was $555. A survey given to 20 randomly selected college students yielded a mean of $570 with a standard deviation of $40.

Clicker!

- What is the population mean?

  A. $555

  B. $570

  C. $40

Hypothesis Testing for One Mean

A research conducted a few years ago showed that the mean amount of money spent on airfare by college students was $555. A survey given to 20 randomly selected college students yielded a mean of $570 with a standard deviation of $40.

Clicker!

- What is the sample mean?

  A. $555

  B. $570

  C. $40

Hypothesis Testing for One Mean

A research conducted a few years ago showed that the mean amount of money spent on airfare by college students was $555. A survey given to 20 randomly selected college students yielded a mean of $570 with a standard deviation of $40.

Clicker!

- What is the sample standard deviation?

  A. $555

  B. $570

  C. $40

Hypothesis Testing for One Mean

A research conducted a few years ago showed that the mean amount of money spent on airfare by college students was $555. A survey given to 20 randomly selected college students yielded a mean of $570 with a standard deviation of $40.

- Are we given the population standard deviation?

  A. Yes

  B. No

Clicker!

Hypothesis Testing for One Mean

- Population mean was \$555.
- New sample mean is \$570.
- We are testing the claim that the population mean is greater than \$555.
- The new data indicate that it may have changed since \$570 is clearly greater than \$555.
- But is this difference statistically significant, i.e. are the data inconsistent enough? We do a formal hypothesis test to answer this question.

Hypothesis Testing for One Mean

1. Hypotheses: state the null and alternative hypotheses.

$$H_0: \mu=555$$

$$H_a: \mu>555$$

2. Type of test: one-sample t-test

Check the conditions:

- Random and independent: random sample of 20 students
- Large sample: the population distribution is nearly normal
- Big population: 10 x 20 = 200. There are more than 200 college students.

Hypothesis Testing for One Mean

3. Test Statistic: We now find the t-statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{570 - 555}{\frac{40}{\sqrt{20}}} = 1.677$$

$$df = n - 1 = 20 - 1 = 19$$

Hypothesis Testing for One Mean

4. P-value:

For a t-statistic, we can only find a range, not an exact value.

t=1.677

df = 19

$$0.05 < p\text{-value} < 0.10$$

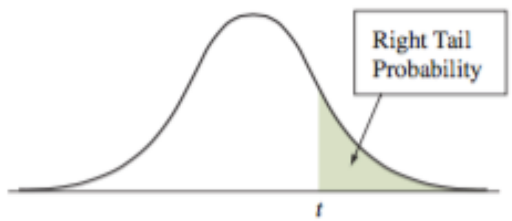


Table 4: *t* Distribution Critical Values

	Confidence Level					
	80%	90%	95%	98%	99%	99.8%
	Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
1	3.078	6.314	12.706	31.821	63.656	318.289
2	1.886	2.920	4.303	6.965	9.925	22.328
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.894
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.611
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552

Hypothesis Testing for One Mean

5. Interpret:

α=0.01 and $0.05 < \text{p-value} < 0.10$

p-value > α, so we fail to reject the null hypothesis.

The data do not provide convincing evidence to suggest that the mean amount of money spent by college students on airfare yearly has increased over the years.

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Determine whether the mean weight for 20-year old vegetarian women is significantly less than 128, using a significance level of 0.05.

Hypothesis Testing for One Mean

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- What is the population mean?

Clicker!

A. 128

B. 122

C. 40

D. 15

Hypothesis Testing for One Mean

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- What is the sample mean?

Clicker!

A. 128

B. 122

C. 40

D. 15

Hypothesis Testing for One Mean

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- What is the sample standard deviation?

Clicker!

A. 128

B. 122

C. 40

D. 15

Hypothesis Testing for One Mean

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- What is the sample size?

Clicker!

A. 128

B. 122

C. 40

D. 15

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Determine whether the mean weight for 20-year old vegetarian women is significantly less than 128, using a significance level of 0.05.

$$\mu = 128,\ \bar{x} = 122,\ s = 15,\ n = 40$$

Example

1. Hypotheses: state the null and alternative hypotheses.

$$H_0: \mu=128$$

$$H_a: \mu<128$$

2. Type of test: one-sample t-test

Check the conditions:

- Random and independent: random sample and independence indicated in problem
- Large sample: $n = 40>25$
- Big population: $10 \times 40 = 400$

Example

3. Test Statistic: We now find the t-statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{122 - 128}{\frac{15}{\sqrt{40}}} = -2.53$$

$$df = n - 1 = 40 - 1 = 39$$

Example

4. P-value:

For a t-statistic, we can only find a range, not an exact value.

t = -2.53

df = 39

Area is less than

-2.53.

$$0.005 < p\text{-value} < 0.010$$

	Confidence Level					
	80%	90%	95%	98%	99%	99.8%
	Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
1	3.078	6.314	12.706	31.821	63.656	318.289
2	1.886	2.920	4.303	6.965	9.925	22.328
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.894
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261

Example

5. Interpret:

$\alpha=0.05$ and $0.005 < \text{p-value} < 0.010$

p-value < α, so we reject the null hypothesis.

The data provide convincing evidence to suggest the mean for vegetarian women is less than 128.

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Now suppose the sample consists of 100 vegetarian women who are 20 years old, and repeat the test.
- Determine whether the mean weight for 20-year old vegetarian women is significantly less than 128, using a significance level of 0.05.

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Now suppose the sample consists of 100 vegetarian women who are 20 years old, and repeat the test.
- Determine whether the mean weight for 20-year old vegetarian women is significantly less than 128, using a significance level of 0.05.

$$\mu = 128,\ \bar{x} = 122,\ s = 15,\ n = 100$$

Example

1. Hypotheses: state the null and alternative hypotheses.

$$H_0: \mu=128$$

$$H_a: \mu<128$$

2. Type of test: one-sample t-test

Check the conditions:

- Random and independent: random sample and independence indicated in problem
- Large sample: n = 100>25
- Big population: 10 x 100 = 1000

Example

3. Test Statistic: We now find the t-statistic.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{122 - 128}{\frac{15}{\sqrt{100}}} = -4$$

$$df = n - 1 = 100 - 1 = 99$$

Example

4. P-value:

For a t-statistic, we can only find a range, not an exact value.

t = -4

df = 99

Area is less than -4.

$$p\text{-value} < 0.001$$

	Confidence Level					
	80%	90%	95%	98%	99%	99.8%
	Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
1	3.078	6.314	12.706	31.821	63.656	318.289
2	1.886	2.920	4.303	6.965	9.925	22.328
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.894
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.091

Example

5. Interpret:

$\alpha=0.05$ and p-value <0.001

p-value $< \alpha$, so we reject the null hypothesis.

The data provide convincing evidence to suggest the mean for vegetarian women is less than 128.

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Explain what causes the difference between the p-values for parts a and b.

Example

The mean weight of all 20-year old women is 128 pounds. A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation of 15 pounds. The women's measurements were independent of each other.

- Explain what causes the difference between the p-values for parts a and b.

The larger the sample size, n, the smaller the standard error (narrower and taller sampling distribution). The shaded area in the tails is smaller hence the smaller p-value.

Comparing Two Population Means

- Do people comprehend better when they read on paper rather than on a computer screen?
- Do men spend less time doing laundry than women? If so, how much less?
- These questions can be answered, in part, by comparing the means of two populations.
- In Ch 7.5, we learned that when we are looking at the difference between two samples, we subtract one sample statistic from the other.
- From subtraction we learn that,
    - if the result is positive, the first mean is greater than the second.
    - if the result is negative, the first mean is less than the second.
    - if the result is 0, the means are equal.

Comparing Two Population Means

- When comparing two populations, it is important to pay attention to whether the data sampled from the population are two independent samples or dependent samples (paired samples).
- Usually dependence occurs when the objects in your sample are all measured twice (such as “before” and “after” comparisons) or when the objects are related somehow (for example, siblings or spouses) or when the experimenters have deliberately matched subjects in the groups to have similar characteristics.

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- A researcher wants to know whether pulse rates of people go down after brief medication. She collects the pulse rates of a random sample of people before medication and then collects their pulse rates after medication.

  A. Paired

  B. Independent

Clicker!

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- A researcher wants to know whether pulse rates of people go down after brief medication. She collects the pulse rates of a random sample of people before medication and then collects their pulse rates after medication.

  A. Paired

  B. Independent

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- A researcher wants to compare the hand-eye coordination of men and women. She finds a random sample of 50 men and 50 women, asks them to thread a needle, and determines how long they take to do the job.

  A. Paired

  B. Independent

Clicker!

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- A researcher wants to compare the hand-eye coordination of men and women. She finds a random sample of 50 men and 50 women, asks them to thread a needle, and determines how long they take to do the job.

  A. Paired

  B. Independent

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- The number of hours of sleep husbands and wives get are compared to see whether the means are different.

  A. Paired

  B. Independent

Clicker!

Independent or Dependent Samples?

- Does the situation involve two independent samples or paired samples?
- The number of hours of sleep husbands and wives get are compared to see whether the means are different.

  A. Paired

  B. Independent

Confidence Intervals for 2 Sample Means (Independent Samples)

- We need to check the CLT conditions:
  - Random Samples and Independence: Both samples are randomly taken from their populations, or subjects are randomly assigned to one of the two groups, and each observation is independent of any other.
  - Independent Samples: The two samples are independent of each other (not paired or dependent).
  - Large Samples: The populations are approximately Normal, or the sample size in each sample is 25 or more.

Confidence Intervals for 2 Sample Means (Independent Samples)

- The formula for the confidence interval for two sample means (independent samples) is

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t^{*}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- The multiplier t* is based on an approximate t-distribution. The degrees of freedom (df) is equal to the smaller of $n_1 - 1$ and $n_2 - 1$.

Interpreting Confidence Intervals for 2 Sample Means (Independent Samples)

- The most important thing to look for is whether or not the interval includes 0.
- If it doesn't include 0, then we have evidence that the two means are different from each other.
    - If the interval contains all positive values, then we are confident that the first mean is greater than the second mean.
    - If the interval contains all negative values, then we are confident that the first mean is less than the second mean.
- If it does include 0, then it is plausible that the two means are equal to each other.
- You should also pay attention to how great or how small the difference between the two means could be, especially if the interval includes 0.

Example

Lower back pain is a common complaint that is hard to treat effectively. Does regular physical therapy help? A randomized experiment assigned patients visiting their doctor for lower back pain to two groups: 102 received an examination and advice from a physical therapist; another 97 received regular physical therapy for up to five weeks. After a year, the change in their level of disability (0% to 100%) was assessed by a doctor who did not know which treatment the patients had received. Most changes were negative, that is, the level of disability went down.

- Construct a 95% confidence interval for the difference between mean therapy and mean advice.

	n	Mean change	s
Therapy	97	-3.83	1.15
Advice	102	-3.20	1.59

Example

- The 95% confidence interval for the difference in decrease in back pain as a results of therapy therapy vs advice.

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (-3.83 - (-3.20)) \pm 1.984 \sqrt{\frac{1.15^2}{97} + \frac{1.59^2}{102}} = (-1.02, -0.24)$$

$$t^* = 1.984$$

	80%	90%	95%	98%	99%	99.8%
			Right-Tail Probability			
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174

Since the interval (-1.02, -0.24) does not contain 0, we have evidence that the therapy mean and advice mean are different from each other. The interval contains negative mean values, hence we are confident that the therapy mean is less than the advice mean.

Hypothesis Testing for 2 Sample Means

- Step 1: Hypothesize

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 > \mu_2 \quad \text{or} \quad \mu_1 < \mu_2 \quad \text{or} \quad \mu_1 \neq \mu_2$$

- Step 2: Two-sample t-test and CLT Conditions:
    - Random Samples and Independent Observations
    - Independent Samples
    - Large Sample
- Step 3: Test statistic $t = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
- Step 4: P-value
- Step 5: Interpret the results

Example

Lower back pain is a common complaint that is hard to treat effectively. Does regular physical therapy help? A randomized experiment assigned patients visiting their doctor for lower back pain to two groups: 102 received an examination and advice from a physical therapist; another 97 received regular physical therapy for up to five weeks. After a year, the change in their level of disability (0% to 100%) was assessed by a doctor who did not know which treatment the patients had received. Most changes were negative, that is, the level of disability went down.

- Is there evidence to suggest that the results of regular therapy are different from just receiving advice?

	n	Mean change	s
Therapy	97	-3.83	1.15
Advice	102	-3.20	1.59

Example

1. Hypotheses: state the null and alternative hypotheses.

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

2. Type of test: two-sample t-test

Check the conditions:

- Random sample and independent observations: randomized experiment is noted.
- Independent samples: we can assume the samples are independent of each other.
- Large samples: the therapy sample size is 97 and the advice sample size is 102, both are larger than 25.

Example

3. Test Statistic: We now find the t-statistic.

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{-3.83 - (-3.20)}{\sqrt{\dfrac{1.15^2}{97} + \dfrac{1.59^2}{102}}} = \frac{-0.63}{0.196} = -3.2143$$

$$df = \min(n_1 - 1, n_2 - 1) = \min(97 - 1, 102 - 1) = \min(96, 101) = 96$$

Example

4. P-value:

For a t-statistic, we can only find a range, not an exact value.

If negative, can look at the positive table due to symmetry.

t= -3.2143

df = 96

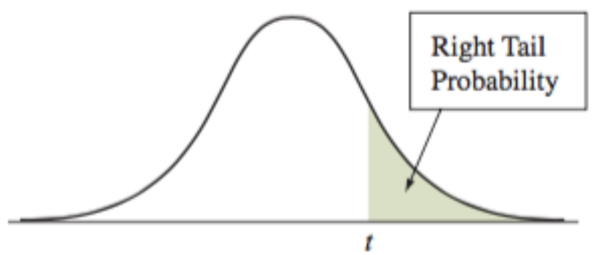


Table 4: t Distribution Critical Values

	Confidence Level					
	80%	90%	95%	98%	99%	99.8%
	Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174

Since we have a two-sided test, we multiply the p-value by 2. We find p-value < 0.001 hence p-value < 2x.0001

$$p\text{-value} < 0.002$$

Example

5. Interpret:

$\alpha=0.05$ and p-value <0.002.

p-value $< \alpha$, so we reject the null hypothesis.

The data provide convincing evidence to suggest that the results of regular therapy are different from just receiving advice.

Confidence Intervals for the Mean of a Difference (Dependent or Paired Samples)

- With paired samples, we turn two samples into one, by finding the difference in each pair.
- For example, researchers wanted to know whether our sensitivity to smells is different when we were sitting up compared to when we are lying down. They came up with a measure for a person's ability of smell and they measured a sample of people twice: once when they were lying down and once when they were sitting upright.
- The data for both groups come from the same people. Hence these are dependent (or paired) samples.
- When we are dealing with paired samples, we need to transform the original data from two variables into a single variable that contains the difference between the scores in group 1 and group 2. It doesn't matter which is group 1 or 2 but we need to remember our choice.
- Once this is done, we can continue to compute a one-sample confidence interval and hypothesis test.

Check the Conditions for CLT

- NOTE: Remember you are viewing this data as a one-sample so you will check the conditions for a one-sample (Ch 9.2).
- Random and Independent: The sample is collected randomly and the trials are independent of each other.
- Large Sample:
  - If the population distribution is normal, then the sampling will be normal as well, regardless of the sample size.
  - OR if the population distribution is not normal, then we need a large enough sample ($n \geq 25$) to ensure that the sampling distribution will be normal.
- Big Population: If the sample is collected without replacement, then the population size is at least 10 times the sample size.

Example

- We return to the example of smelling ability (sitting upright or lying down).
- The first few lines of the original data are shown below.

Subject Number	Gender	Sitting Upright	Lying Down
1	Female	13.5	13.25
2	Female	13.5	13
3	Female	12.75	11.5
4	Male	12.5	12.5

- We create a new variable called “difference” and define it to be the difference between smelling ability sitting upright and lying down. The new variable is shown below.

Subject	Gender	Sitting Upright	Lying Down	Difference
1	Female	13.5	13.25	0.25
2	Female	13.5	13	0.50
3	Female	12.75	11.5	1.25
4	Male	12.5	12.5	0

Example

- Here are the summary statistics for Sitting Upright, Lying Down and Difference.

Variable	n	Sample Mean	Sample Standard Deviation
Sitting Upright	36	11.47	3.26
Lying Down	36	10.60	3.06
Difference	36	0.87	2.39

- After verifying that the CLT conditions hold (they do hold), we can find a 95% confidence interval for the mean difference using the one-sample formula from Ch 9.3.

$$\overline{\mathrm{x}} \pm t_{df}^{*}\frac{s}{\sqrt{n}} = 0.87 \pm 2.03\frac{2.39}{\sqrt{36}} = (0.06,\ 1.68)$$

$$df = n - 1 = 36 - 1 = 35$$

Example

- We find the 95% confidence interval to be (0.06, 1.68).
- Since the interval contains only positive values, we conclude that the mean for the first group (sitting upright) is higher than the mean for the second group (lying down).
- For this reason, we are confident that smelling sensitivity is greater when sitting upright than when lying down. This difference could be fairly small, 0.06 units, or as large as 1.68 units.

Hypothesis Testing for the Mean of a Difference (Dependent or Paired Samples)

- As with confidence intervals, we convert the two variables into a difference variable, and our hypotheses are now not about the individual groups, but about the difference of the two groups.
- Let's consider an example. The diet program known as The Zone promises that you'll lose weight, burn fat, and not feel hungry. The diet requires that you eat 30% protein, 30% fat, and 40% carbs, and it also imposes restrictions on the times at which you eat your meals and snacks. Can people lose weight on The Zone?
- The subjects were randomly assigned to The Zone diet. For each subject, weight was measured at 0 months and at 2 months. Because the subjects appear in both groups, the data are paired.

Example

- Instead of considering weight at 0 months and weight at 2 months, as separate variables, we will calculate the difference in weight and name this variable “difference”.

Difference = (weight at 2 months) - (weight at 0 months)

- The summary statistics: $\bar{x} = -3.795$ kg, $s = 3.5903$ kg, $n = 40$
- Our hypotheses are now about just one mean, the mean of difference:

$$H_0: \mu_{difference} = 0 \quad (\text{or } \mu_{2months} = \mu_{0months})$$

$$H_a: \mu_{difference} < 0 \quad (\text{or } \mu_{2months} < \mu_{0months})$$

Example

- Next we find our t-statistic.

$$t = \frac{\bar{x}_{difference} - 0}{\frac{s_{difference}}{\sqrt{n}}} = \frac{-3.795}{\frac{3.5903}{\sqrt{40}}} = -6.685$$

- To find the p-value use the t-distribution with df = n-1 = 40-1=39. We want the area to the left of the -6.685 and we find the p-value to be extremely small (nearly 0).

- Since p-value≈0 < α=0.05, we reject the null hypothesis and conclude that the typical subject in the study really did lose weight on The Zone diet.