8.1 REVIEW AND PREVIEW

DEFINITION

In statistics, a **hypothesis** is a claim or statement about a property of the population.

A **hypothesis test (aka test of significance)** is a procedure for testing a claim about a property of a population.

8.2 BASICS OF HYPOTHESIS TESTING

PART 1: BASICS CONCEPTS OF HYPOTHESIS TESTING

The methods presented in this chapter are based on the rare event rule for inferential statistics.

RARE EVENT RULE FOR INFERENTIAL STATISTICS

If, under a given assumption, the probability of a particular observed event is extremely small, we conclude that the assumption is probably not correct.

WORKING WITH THE STATED CLAIM: NULL AND ALTERNATIVE HYPOTHESES

The **null hypothesis** denoted by H_0 is a statement that the value of a population parameter is equal to some claimed value. The term null is used to indicate no change or no effect or no difference.

The **alternative hypothesis** denoted by H_A or H_1 or H_a is the statement that the parameter has a value that somehow differs from the null hypothesis.

For the methods of this chapter, the symbolic form of the alternative hypothesis must use one of these symbols: $<$, $>$, $\neq$.

IDENTIFYING H_0 **AND** H_a

START

1

- Identify the specific claim or hypothesis to be tested
- Express it in symbolic form

2

- Give the symbolic form that must be true when the original claim is false

3

- Using the two symbolic expressions obtained so far, identify the null hypothesis H_0 and the alternative hypothesis H_a
- H_a is the symbolic expression that does not contain equality
- H_0 is the symbolic expression that the parameter equals the fixed value being considered

Example 1: Examine the given statement, then express the null hypothesis and the alternative hypothesis in symbolic form.

a. The majority of college students have credit cards.

$H_0: p \le \frac{1}{2}$

$H_a: p > \frac{1}{2}$

b. The mean weight of plastic discarded by households in one week is less than 1 kg.

$H_0: \mu \ge 1$

$H_a: \mu < 1$

CONVERTING SAMPLE DATA TO A TEST STATISTIC

Test statistic for proportion:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

p is the proportion we assume to be true under the null.
$\hat{p}$ is statistic from sample.

Test statistic for mean:

sample mean

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{OR} \quad t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

μ is the value of the mean we assume to be true under the null hypothesis.

Example 2: Find the value of the test statistic. The claim is that less than ½ of adults in the United States have carbon monoxide detectors. A KRC Research survey of 1005 adults resulted in 462 who have carbon monoxide detectors.

1

$H_0: p \ge \frac{1}{2}$

$H_a: p < \frac{1}{2}$

2 Test Stat.

$$\frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.46 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1005}}} \approx \boxed{-2.54}$$

$x = 462$

$n = 1005$

$p = 0.5$

$q = 1 - p$

$q = 1 - 0.5$

$q = 0.5$

$\hat{p} = \frac{x}{n} = \frac{462}{1005}$

$\hat{p} = 0.46$

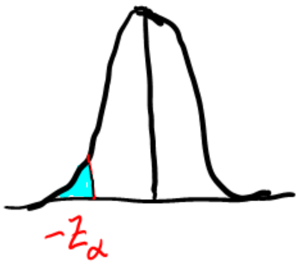
TOOLS FOR ASSESSING THE TEST STATISTIC: CRITICAL REGION, SIGNIFICANCE LEVEL, CRITICAL VALUE, AND *P*-VALUE

The test statistic alone usually does not give us enough information to make a decision about the claim being tested. The following tools can be used to understand and interpret the test statistic.

- The **critical region (aka rejection region)** is the set of all values of the test statistic that cause us to reject the null hypothesis H_0
- The **significance level (denoted by α)** is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true. If the test statistic falls in the critical region, we reject the null hypothesis, so α is the probability of making the mistake of rejection the null hypothesis when it is true.
- A **critical value** is any value that separates the critical region from the values of the test statistic that do not lead to rejection of the null hypothesis. The critical values depend on the nature of the null

hypothesis, the sampling distribution that applies, and the significance level α. The procedure can be summarized as follows:

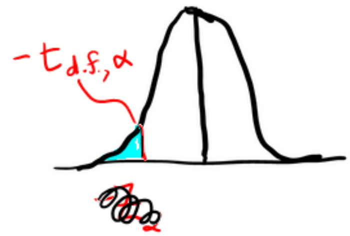
Critical region in the left tail:



$H_0: p \geq p_0$ or $H_0: \mu \geq \mu_0$

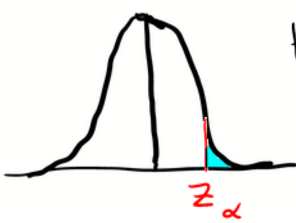
$H_a: p < p_0$ $\quad$ $H_a: \mu < \mu_0$

if σ is known



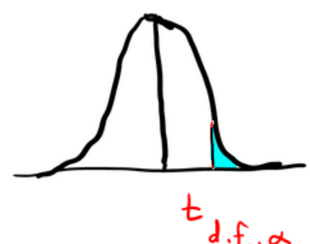
$H_0: \mu \geq \mu_0$
$H_a: \mu < \mu_0$ $\Big\}$ σ is unknown

Critical region in the right tail:



$H_0: p \leq p_0$ or $H_0: \mu \leq \mu_0$

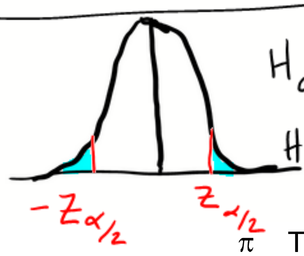
$H_a: p > p_0$ $\quad$ $H_a: \mu > \mu_0$



$H_0: \mu \leq \mu_0$

$H_a: \mu > \mu_0$

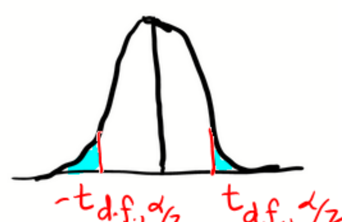
Critical region in two tails:



$H_0: p = p_0$ or $H_0: \mu = \mu_0$

$H_a: p \neq p_0$ $\quad$ $H_a: \mu \neq \mu_0$

σ is known



$H_0: \mu = \mu_0$

$H_a: \mu \neq \mu_0$

σ is unknown

π The ***P*-value (aka *p*-value or probability value)** is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true. *P*-values can be found after finding the area beyond the test statistic.

$<$

$P(Z < \text{test statistic})$

or

$P(t < \text{test statistic})$

$>$

$P(Z > \text{test statistic})$

or

$P(t > \text{test statistis}$

$\neq$

(test stat negative)

$2P(Z \neq \text{test statistic})$

or

$2P(t \neq \text{test statistic})$

$\neq$

(test stat positive)

$2P(Z \neq \text{test statistic})$

$2P(t \neq \text{test statistic})$

DECISIONS AND CONCLUSIONS

***P*-value method:** Using the significance level α:

If P-value $\leq \alpha$, reject H_0

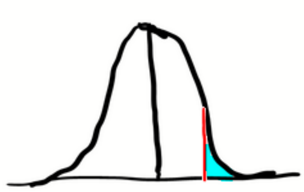
If P-value $> \alpha$, fail to reject H_0

Traditional method: If the test statistic falls within the critical (rejection) region, reject H_0. If the test statistic does not fall within the critical region, fail to reject H_0.

Confidence intervals: A confidence interval estimate of a population parameter contains the likely values of that parameter. If a confidence interval does not include a claimed value of a population parameter, reject that claim.

Example 3: Use the given information to find P-value.

a. The test statistic in a right-tailed test is $z = 2.50$



test stat is $z = 2.50$

$P(z > 2.50) = 1 - 0.9938$

$= \boxed{0.0062}$ ← p-value

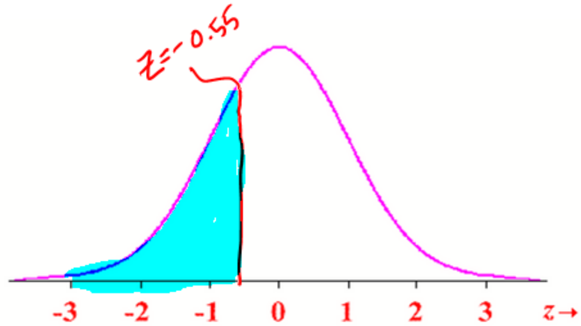
or

$P(z > 2.50) = P(z < -2.50)$

$= \boxed{0.0062}$

b. The test statistic in a two-tailed test is z = -0.55

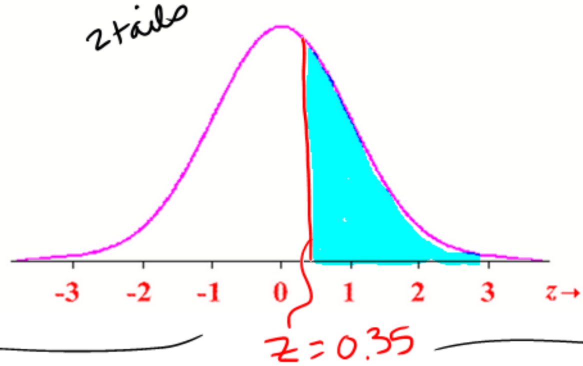
test stat.



$$\text{P-value} = 2P(z < \underbrace{-0.55}) = 2(0.2912) = \boxed{0.5824}$$

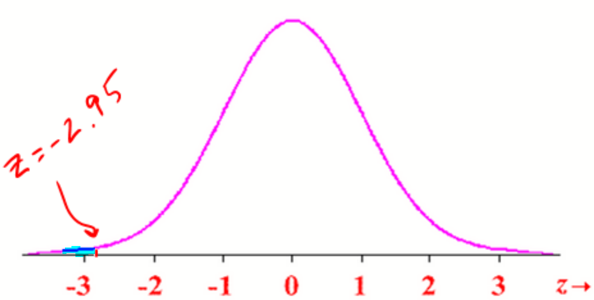
c. With $H_1: p \neq \frac{3}{4}$, the test statistic is z = 0.35

2 tails



$$\text{p-value} = 2P(z > \underline{0.35}) = 2[1 - P(z < 0.35)] = 2(1 - 0.6368) = 2(0.3632) = \boxed{0.7264}$$

d. With $H_1: p < 0.777$, the test statistic is z = -2.95



$$\text{p-value} = P(z < \underline{-2.95}) = \boxed{0.0016}$$

Example 4: State the final conclusion in simple non-technical terms. Be sure to address the original claim. Original claim: The percentage of on-time U.S. airline flights is less than 75%. Initial conclusion: Reject the null hypothesis.

$H_a: p < 0.75 \rightarrow H_0$ was rejected

There is significant evidence to support the claim that the percentage of on-time U.S. airline flights is less than 75%.

ERRORS IN HYPOTHESIS TESTS

		TRUE STATE OF NATURE	
		THE NULL HYPOTHESIS IS TRUE	THE NULL HYPOTHESIS IS FALSE
DECISION	We decide to reject H_0	TYPE I ERROR	CORRECT DECISION
	We fail to reject H_0	CORRECT DECISION	TYPE II ERROR

Example 5: Identify the type I error and the type II error that correspond to the given hypothesis. The percentage of Americans who believe that life exists only on earth is equal to 20%.

$H_0: p = 0.20$

$H_a: p \neq 0.20$

Type I error: Rejecting the claim that 20% of Americans believe life exists only on Earth when that proportion is true.

Type II error: Failing to reject the claim that 20% of Americans believe life exists only on Earth when the proportion is really different than 20%.

COMPREHENSIVE HYPOTHESIS TEST

CONFIDENCE INTERVAL METHOD

For 2-tailed hypothesis tests construct a confidence interval with a confidence level of $1-\alpha$; but for a 1-tailed hypothesis test with significance level α, construct a confidence interval of $1-2\alpha$.

A confidence interval estimate of a population proportion contains the likely values of that parameter. We should therefore reject a claim that the population parameter has a value that is not included in the confidence interval.

8.3 TESTING A CLAIM ABOUT A PROPORTION

PART 1: BASIC METHODS OF TESTING CLAIMS ABOUT A POPULATION PROPORTION p

OBJECTIVE

Test a claim about a pop. proportion

NOTATION

$n =$ Sample size or # of trials

$p =$ pop. proportion assumed true under H_0

$\hat{p} = \frac{x}{n}$ → sample proportion

$q =$ $1-p$

REQUIREMENTS

1. The sample observations are a simple random sample.
2. The conditions for a binomial distribution are satisfied.

   Fixed # of independent trials having constant probabilities and each trial has 2 outcome categories

3. The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied so the binomial dist. of sample proportions can be approximated by a normal dist. with $\mu = np$ and $\sigma = \sqrt{npq}$. Note that p is the assumed proportion used in the null.

TEST STATISTIC FOR TESTING A CLAIM ABOUT A PROPORTION

$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$

$P-$values:

Critical values:

Table A-2

FINDING THE NUMBER OF SUCCESSES x

Computer software and calculator designed for hypothesis tests of proportions usually require input consisting of the sample size n and the number of successes x, but the sample proportion is often given instead of x.

$\hat{p} = \dfrac{x}{n} \longleftrightarrow x = n\hat{p}$

Example 1: Identify the indicated values. Use the normal distribution as an approximation to the binomial distribution. In a survey, 1864 out of 2246 randomly selected adults in the United States said that texting while driving should be illegal (based on data from Zogby International). Consider a hypothesis test that uses a 0.05 significance level to test the claim that more than 80% of adults believe that texting while driving should be illegal.

a. What is the test statistic?

$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$

$z = \dfrac{0.8299 - 0.8}{\sqrt{\dfrac{(0.8)(0.2)}{2246}}}$

$\boxed{z = 3.5426}$

$x = 1864$

$n = 2246$

$p = 0.80$

$q = 1 - 0.8 = 0.2$

$\hat{p} = \dfrac{1864}{2246}$

$\hat{p} = 0.8299$

Right 1-tailed test

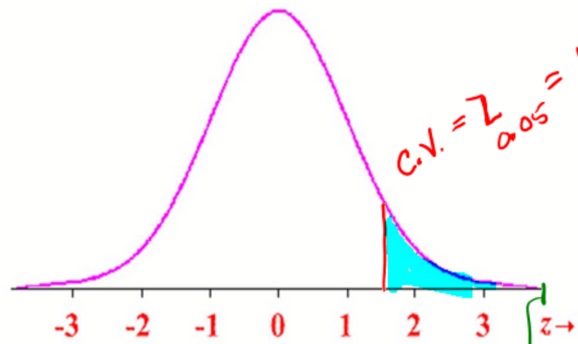
$np \geq 5$ ✓

$nq \geq 5$ ✓122

b. What is the critical value?

1-tailed test (right tail), $\alpha = 0.05$

$C.V. = z_{0.05} = \boxed{1.645}$

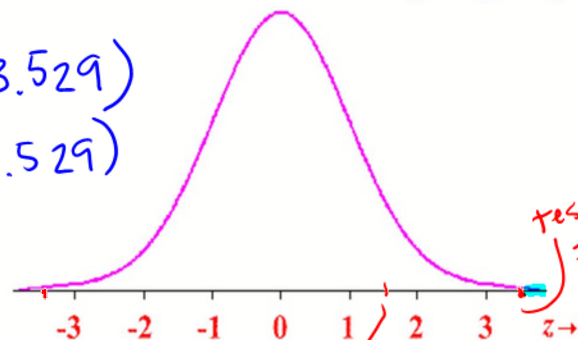


$C.V. = z_{0.05} = 1.645$

$z = 3.529$ which falls in rejection region.

c. What is the *P*-value?

$P\text{-value} = P(z > 3.529)$
$= 1 - P(z < 3.529)$
$= 1 - 0.9999$
$= \boxed{0.0001}$



test stat $z = 3.529$

$z_{0.05} = 1.645$

$P(z > 3.529) = 0.0001 < 0.05$ so reject H_0

d. What is the conclusion?

There is significant evidence at the 5% level to reject the null hypothesis and support the claim that more than 80% of adults believe that texting while driving should be illegal.

Example 2: The company Drug Test Success provides a "1-Panel-THC" test for marijuana usage. Among 300 tested subjects, results from 27 subjects were wrong (either a false positive or a false negative). Use a 0.05 significance level to test the claim that less than 10% of the test results are wrong. Does the test appear to be good for most purposes?

a. Identify the null hypothesis

$H_0: p \geq 0.1$

b. Identify the alternative hypothesis

$H_a: p < 0.1$

$x = 27$
$n = 300$
$\hat{p} = \frac{27}{300} = 0.09$
$p = 0.1$
$q = 0.9$
$np \geq 5 \rightarrow$ yes
$nq \geq 5 \rightarrow$ yes

left 1-tail test
$\alpha = 0.05$

c. Identify the test statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.09 - 0.1}{\sqrt{\frac{(0.1)(0.9)}{300}}}$$

$z \approx \boxed{-0.5774}$

Stat | test | 1-propZtest . . . calculate

1-PropZTest
prop<0.1
z=-0.5773502692
p=0.2818513864
$\hat{p}$=0.09
n=300

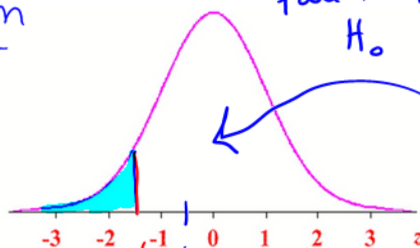
d. Identify the *P*-value or critical value(s)

Crit. Region

$z_{\alpha} = z_{0.05}$

$= -1.645$

Fail to reject H_0

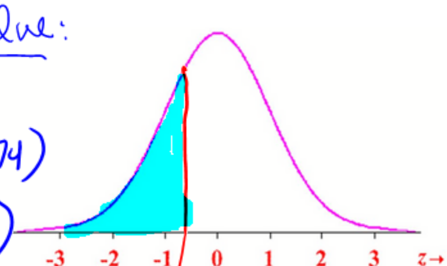


↳ test stat: $z \doteq -0.5774$

P-value:

P-value
$= P(z < -0.5774)$
$\approx P(z < -0.58)$
$= 0.2810$

$0.2810 > 0.05$
so we fail to reject H_0.



test stat $z = -0.5774$

e. What is your final conclusion?

There is not significant evidence at the 5% level to support the claim that less than 10% of the tests are wrong. It would depend on the philosophy of a company.

Example 3: In recent years, the town of Newport experienced an arrest rate of 25% for robberies (based on FBI data). The new sheriff compiles records showing that among 30 recent robberies, the arrest rate is 30%, so she claims that her arrest rate is greater than the 25% rate in the past. Is there sufficient evidence to support her claim that the arrest rate is greater than 25%?

a. Identify the null hypothesis

$H_0: p \leq 0.25$ or $p = 0.25$

how it's given in homework

b. Identify the alternative hypothesis

$H_a: p > 0.25$

$n = 30$
$p = 0.25$
$q = 0.75$
$np = 30(.25) > 5$ ✓
$nq = 30(.75) > 5$ ✓
$\hat{p} = 0.30$
$\hat{p} = \frac{x}{n} \Rightarrow 0.3 = \frac{x}{30}$
need for calculator $\rightarrow x = 9$
$\alpha = 0.05$ (common α level)
1-tailed test (right)

c. Identify the test statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.3 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{30}}}$$

$z \approx 0.63$

d. Identify the *P*-value or critical value(s)

$C.V. = z_{\alpha} = z_{0.05} = 1.645$

$z = 0.63$ doesn't lie in the crit. region so fail to reject H_0

$p\text{-value} = P(z > \text{test stat})$

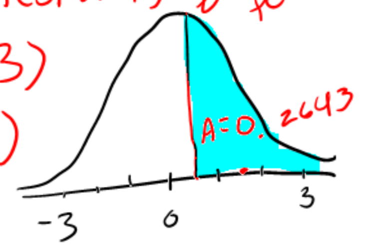
$= P(z > 0.63)$

$= 1 - P(z < 0.63)$

$= 1 - 0.7357$

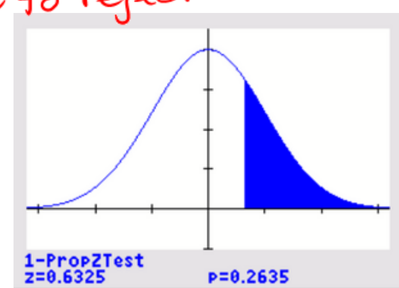
$= 0.2643 \rightarrow 0.2643 > 0.05 = \alpha$

so we fail to reject



1-PropZTest
prop>0.25
z=0.632455532
p=0.2635445651
p̂=0.3
n=30

calculate



draw

stat test 1propztest

e. What is your final conclusion?

There's not significant evidence at the 5% level to support the new sheriff's claim that her arrest rate for robberies is greater than 25%.

8.4 TESTING A CLAIM ABOUT A MEAN: SIGMA KNOWN

TESTING CLAIMS ABOUT A POPULATION MEAN (WITH σ KNOWN)
OBJECTIVE
NOTATION

$n =$ sample siz

$\mu_{\bar{x}} =$ pop. mean of all sample means (assumed true under H_0)

$\bar{x} =$ sample mean

$\sigma =$ pop. standard deviation

REQUIREMENTS

1. The sample is a simple random Sample (SRS).
2. The value of the pop. standard deviation σ is know. (unlikely in real world)
3. The population is normally distributed and/or $n > 30$.

TEST STATISTIC FOR TESTING A CLAIM ABOUT A MEAN (WITH σ KNOWN)

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

P – values:

Critical values:

Table A-2 similar technique to 8.3

$\mu_{\bar{x}} = 3.5$, $\bar{x} = 2.9375$, $n = 40$, $\sigma = 1.7078$, $\alpha = 0.05$, 2-tailed test

Example 1: When a fair die is rolled many times, the outcomes of 1, 2, 3, 4, 5, and 6 are equally likely, so the mean of the outcomes should be 3.5. The author drilled holes into a die and loaded it by inserting lead weights, then rolled it 40 times to obtain a mean of 2.9375. Assume that the standard deviation of the outcomes is 1.7078, which is the standard deviation for a fair die. Use a 0.05 significance level to test the claim that outcomes from the loaded die have a mean different from the value of 3.5 expected with a fair die.

a. Identify the null hypothesis

$H_0: \mu_{\bar{x}} = 3.5$

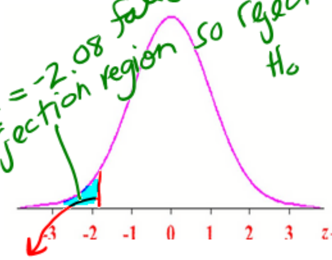
b. Identify the alternative hypothesis

$H_a: \mu_{\bar{x}} \neq 3.5$

c. Identify the test statistic

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} \rightarrow z = \frac{2.9375 - 3.5}{\frac{1.7078}{\sqrt{40}}} \rightarrow z = -2.08$$

$z = -2.08$ falls in rejection region so reject H_0



$-z_{\alpha/2} = -z_{.025} = -1.96$

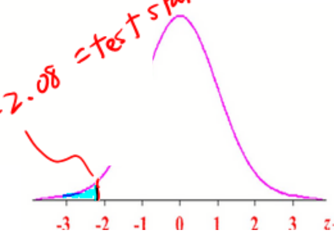
Stat → test → Z-test ... Calculate

Z-Test
μ≠3.5
z=-2.083125874
p=0.0372396217
x̄=2.9375
n=40

d. Identify the *P*-value or critical value(s)

$$\text{p-value} = 2P(z < \text{test stat})$$
$$= 2P(z < -2.08)$$
$$= 2(0.0188)$$
$$= 0.0376$$

$z = -2.08$ = test stat



Since $\cancel{0.0188}$ 0.0376 < $\cancel{0.025}$ 0.05 (α) we reject H_0

e. What is your final conclusion?

There's significant evidence at the 5% to support the claim that the mean of the loaded die is different than the mean of the fair die.

Example 2: Listed below are recorded speeds (in mi/h) of randomly selected cars traveling on a section of Highway 405 in Los Angeles (based on data from Sigalert). That part of the highway has a posted speed limit of 65 mi/h. Assume that the standard deviation of speeds is 5.7 mi/h and use a 0.01 significance level to test the claim that the sample data is from a population with a mean greater than 65 mi/h.

68 68 72 73 65 74 73 72 68 65 65 73 66 71 68 74 66 71 65 73
59 75 70 56 66 75 68 75 62 72 60 73 61 75 58 74 60 73 58 75

a. Identify the null hypothesis

b. Identify the alternative hypothesis

c. Identify the test statistic

d. Identify the *P*-value or critical value(s)

e. What is your final conclusion?

8.5 TESTING A CLAIM ABOUT A MEAN: SIGMA NOT KNOWN

TESTING CLAIMS ABOUT A POPULATION MEAN (WITH σ NOT KNOWN)

OBJECTIVE

* We only know s, not σ

NOTATION

$n =$ sample size

$\mu_{\bar{x}} =$ pop mean of the sample means assumed true under H_0

$\bar{x} =$ sample mean

$s =$ sample standard deviation

REQUIREMENTS

1. The sample is a simple random sample (SRS).
2. The value of the population standard deviation σ is not known.
3. **The** population **is** normally distributed **and/or** n>30.

TEST STATISTIC FOR TESTING A CLAIM ABOUT A MEAN (WITH σ KNOWN)

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

$P -$ values: Table A-3 or calculator

Critical values:

CHOOSING THE CORRECT METHOD

When testing a claim about a population mean, first be sure that the sample data have been collected with an appropriate Sampling method. If we have a simple random Sample, a hypothesis test of a claim about μ might use the Student t - dist., the normal distribution, or it might require nonparametric methods or bootstrapping resampling techniques.

To test a Claim **about a** pop mean**, use the** Student t - dist. **when the sample is a** SRS, σ **is** not Known**, and** one **or** both **of these conditions is satisfied:**

The population **is** normally **distributed or** $n > 30$**.**

Example 1: Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, Student t distribution, or neither.

a. Claim about FICO credit scores of adults: $\mu = 678,\ n = 12,\ \bar{x} = 719,\ s = 92.$ The sample data appear to come from a population with a distribution that is not normal and σ is not known.

NEITHER

b. Claim about daily rainfall amounts in Boston: $\mu < 0.20 \text{ in.},\ n = 52,\ \bar{x} = 0.10 \text{ in.},\ s = 0.26 \text{ in.}$ The sample data appear to come from a population with a distribution that is very far from normal, and σ is known.

NORMAL DIST.

FINDING *P*-VALUES WITH THE STUDENT *t* DISTRIBUTION

1. Use software or a graphing calculator.
2. If technology is not available, use Table A-3 to identify a range of values as follows: Use the number of degrees of freedom to locate the relevant row of Table A-3, then determine where the test statistic lies relative to the t-values in that row. Based on a comparison of the t test statistic and the t values in the row of Table A-3, identify a range of values by referring to the area values given at the top of Table A-3.

Example 2: Either use technology to find the *P*-value or use Table A-3 to find a range of values for the *P*-value.

a. Movie Viewer Ratings: Two-tailed test with $n = 15$, and test statistic $t = 1.495$.

$\boxed{0.1 < \text{p-value} < 0.2}$

d.f. = 14

b. Body Temperatures: Test a claim about the mean body temperature of healthy adults. Left-tailed test with $n = 11$ and test statistic $t = -3.518$.

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

d.f. = 10

* The positive t-values in the table yield the same areas as the negatives.

Example 3: Assume that a SRS has been selected from a normally distributed population and test the given claim. A SRS of 40 recorded speeds (in mi/h) is observed from cars traveling on a section of Highway 405 in Los Angeles. The sample has a mean of 68.4 mi/h and a standard deviation of 5.7 mi/h (based on data from Sigalert). Use a 0.05 significance level to test the claim that the mean speed of all cars is greater that the posted speed limit of 65 mi/h.

a. Identify the null hypothesis

$H_0: \mu_{\bar{x}} \le 65$ or $\mu_{\bar{x}} = 65$

b. Identify the alternative hypothesis

$H_a: \mu_{\bar{x}} > 65$

n = 40
$\bar{x} = 68.4$
$\boxed{s = 5.7}$
$\mu_{\bar{x}} = 65$
$\alpha = 0.05$
1-tailed test (right)
t-test

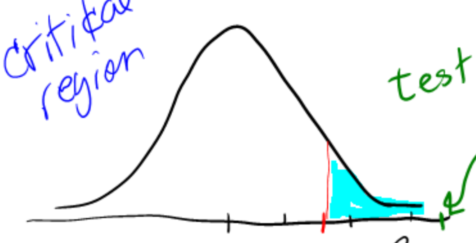
c. Identify the test statistic

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} \rightarrow t = \frac{68.4 - 65}{\frac{5.7}{\sqrt{40}}} \rightarrow t \approx 3.77$$

d. Identify the *P*-value or critical value(s)

d.f. = 39

critical region



test stat = t = 3.77

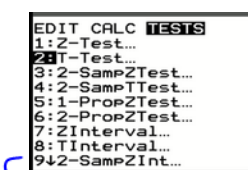
falls in rejection region so reject H_0

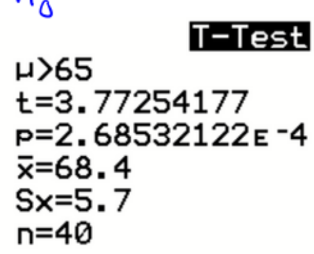
CV = $t_{(t_{39,\,0.05})}$ = 1.685

p-value: $P(t > 3.77)$

p-value < 0.005

p-value < 0.005 < 0.05 so we reject H_0





e. What is your final conclusion?

There's significant evidence to support the claim that cars are traveling faster than 65mph on that stretch of the 405.

d.f. = 14 , $t = 1.495$

Since $1.761 < 1.495 < 1.345$

$0.10 < \text{p-value} < 0.20$

TABLE A-3 *t* Distribution

Degrees of Freedom	.005 (one tail) .01 (two tails)	.01 (one tail) .02 (two tails)	.025 (one tail) .05 (two tails)	.05 (one tail) .10 (two tails)	.10 (one tail) .20 (two tails)	.25 (one tail) .50 (two tails)
1	63.657	31.821	12.706	6.314	3.078	1.000
2	9.925	6.965	4.303	2.920	1.886	.816
3	5.841	4.541	3.182	2.353	1.638	.765
4	4.604	3.747	2.776	2.132	1.533	.741
5	4.032	3.365	2.571	2.015	1.476	.727
6	3.707	3.143	2.447	1.943	1.440	.718
7	3.500	2.998	2.365	1.895	1.415	.711
8	3.355	2.896	2.306	1.860	1.397	.706
9	3.250	2.821	2.262	1.833	1.383	.703
10	3.169	2.764	2.228	1.812	1.372	.700
11	3.106	2.718	2.201	1.796	1.363	.697
12	3.054	2.681	2.179	1.782	1.356	.696
13	3.012	2.650	2.160	1.771	1.350	.694
14	2.977	2.625	2.145	1.761	1.345	.692

1.495

Since $-3.518 < -3.169$

$\text{p-value} < 0.005$

d.f. = 10, $t = -3.518$

TABLE A-3 *t* Distribution

Degrees of Freedom	.005 (one tail) .01 (two tails)	.01 (one tail) .02 (two tails)	.025 (one tail) .05 (two tails)	.05 (one tail) .10 (two tails)	.10 (one tail) .20 (two tails)	.25 (one tail) .50 (two tails)
1	63.657	31.821	12.706	6.314	3.078	1.000
2	9.925	6.965	4.303	2.920	1.886	.816
3	5.841	4.541	3.182	2.353	1.638	.765
4	4.604	3.747	2.776	2.132	1.533	.741
5	4.032	3.365	2.571	2.015	1.476	.727
6	3.707	3.143	2.447	1.943	1.440	.718
7	3.500	2.998	2.365	1.895	1.415	.711
8	3.355	2.896	2.306	1.860	1.397	.706
9	3.250	2.821	2.262	1.833	1.383	.703
10	−3.169	−2.764	−2.228	−1.812	−1.372	−.700

−3.518

Example 2: Assume that a SRS has been selected from a normally distributed population and test the given claim. The trend of thinner Miss America winners has generated charges that the contest encourages unhealthy diet habits among young women. Listed below are body mass indexes (BMI) of recent Miss America winners. Use a 0.01 significance level to test the claim that recent Miss America winners are from a population with a mean BMI less than 20.16, which was the BMI for winners from the 1920s and 1930s.

19.5 20.3 19.6 20.2 17.8 17.9 19.1 18.8 17.6 16.8

a. Identify the null hypothesis

$H_0: \mu_{\bar{x}} \geq 20.16$ or $\mu_{\bar{x}} = 20.16$

b. Identify the alternative hypothesis

$H_a: \mu_{\bar{x}} < 20.16$

$n = 10$

$\bar{x} = 18.76$

$s = 1.19$

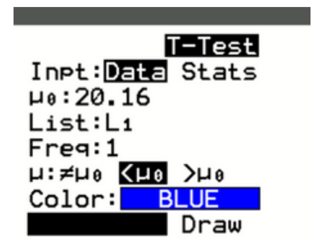
→ t-dist

1-tail, left

$\alpha = 0.01$

c. Identify the test statistic

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}} \rightarrow t = \frac{18.76 - 20.16}{\frac{1.19}{\sqrt{10}}} \rightarrow t \approx -3.72$$



d. Identify the *P*-value or critical value(s)

C.V: $t_{9, 0.01} = -2.821$

Since $-3.72 < -2.821$, our test stat falls in rejection region so we reject H_0

table A-3

P-value:

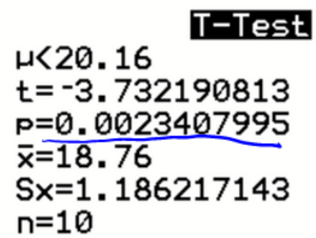
↓

P-value < 0.005

since 0.005 < 0.01

P-value < α

we reject H_0.

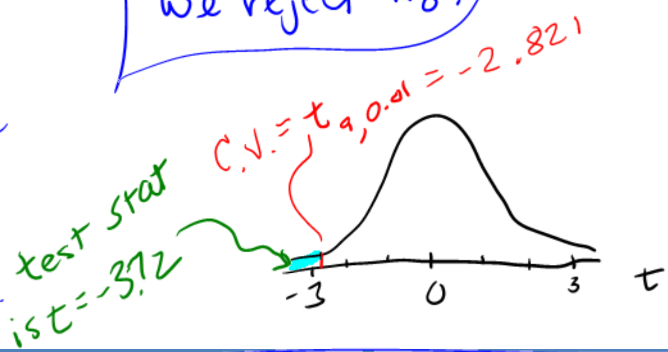


→ test stat

→ P-value

e. What is your final conclusion?

There's significant evidence to support the claim that the B.M.I. of recent Miss America winners is less than winners from the 20's and 30's.



9.2 INFERENCES ABOUT TWO PROPORTIONS

OBJECTIVES

Test a claim about 2 pop. proportions or construct a CI of the difference between 2 pop. proportions

NOTATION FOR TWO PROPORTIONS

$p_1 =$ pop. proportion for pop 1

$\hat{p}_1 = \frac{x_1}{n_1}$

$n_1 =$ sample size for pop. 1

$\hat{q}_1 = 1 - \hat{p}_1$

$x_1 =$ # of successes in the sample from pop. 1

The corresponding notations $p_2, n_2, x_2, \hat{p}_2$, and $\hat{q}_2$ apply to population 2.

POOLED SAMPLE PROPORTION

The __pooled__ __sample__ __proportion__ is denoted by _$\bar{p}$_ and is given by:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}, \quad \bar{q} = 1 - \bar{p}$$

REQUIREMENTS

1. The __sample__ __prop.__ are from __2__ __simple random__ samples that are __independent__.

2. For each of the __2__ samples, the number of __successes__ is __at__ __least__ __5__ and the number of __failures__ is at __least__ __5__. That is, __$np \geq 5$__ and __$nq \geq 5$__ for each of the two samples.

TEST STATISTIC FOR TWO PROPORTIONS (WITH $H_0: p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}_1\bar{q}_1}{n_1} + \frac{\bar{p}_2\bar{q}_2}{n_2}}}$$

$p_1 - p_2 = 0$ under the null

$\hat{p}_1 = \frac{x_1}{n_1}$ and $\hat{p}_2 = \frac{x_2}{n_2}$

P – value:

Critical values: Table A-2 or calculator

CONFIDENCE INTERVAL ESTIMATE OF $p_1 = p_2$

The confidence interval estimate of the difference $p_1 - p_2$ is:

$$\hat{p}_1 - \hat{p}_2 - E < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + E$$

where the margin of error E is given by

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Rounding: Round the confidence interval limits to 3 significant digits.

CAUTION!!! When testing a claim about 2 population proportions, the p-value method and the critical method are equivalent, but they are not equivalent to the confidence interval method!!! If you want to test a claim about 2 pop proportions,

use the C.V. **method or the** p-value **method; if you want to** estimate **the** difference **between** 2 pop proportions**, use a** confidence interval.

Example 1: In a 1993 survey of 560 college students, 171 said they used illegal drugs during the previous year. In a recent survey of 720 college students, 263 said that they used illegal drugs during the previous year (based on data from the National Center for Addiction and Substance Abuse at Colombia University). Use a 0.05 significance level to test the claim that the proportion of college students using illegal drugs in 1993 was less than it is now.

$H_0: p_1 = p_2$ $\qquad$ $H_a: p_1 < p_2$

$p_1 - p_2 = 0$

$n_1 = 560$

$x_1 = 171$

$\hat{p}_1 = \frac{171}{560} \approx 0.305$

$\hat{q}_1 = 0.695$

$n_2 = 720$

$x_2 = 263$

$\hat{p}_2 = \frac{263}{720} \approx 0.365$

$\hat{q}_2 = 0.635$

$\alpha = 0.05$

$p_1 < p_2$

$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$

$\bar{p} = \frac{171 + 263}{560 + 720}$

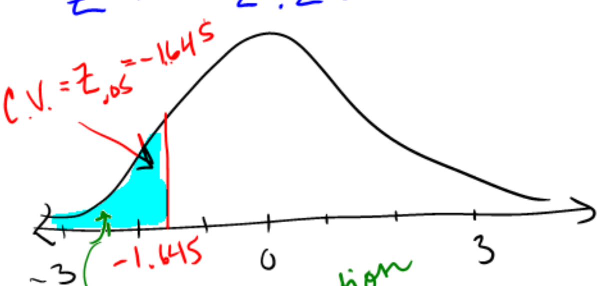
$\bar{p} \approx 0.339$

$\bar{q} \approx 0.661$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

$$z = \frac{(0.305 - 0.365) - (0)}{\sqrt{\frac{(0.339)(0.661)}{560} + \frac{(0.339)(0.661)}{720}}}$$

$z \approx -2.25$



lies in the rejection region so we reject H_0.

Conclusion:

There's significant evidence at the 5% level to support the claim that the proportion of college students using illegal drugs in 1993 was less than it is now.

Example 2: Among 2739 female atom bomb survivors, 1397 developed thyroid diseases. Among 1352 male atom bomb survivors, 436 developed thyroid diseases (based on data from "Radiation Dose-Response Relationships for Thyroid Nodules and Autoimmune Thyroid Diseases in Hiroshima and Nagasaki Atomic Bomb Survivors 55-58 Years After Radiation Exposure," by Imaizumi, et al., *Journal of the American Medical Association,* Vol. 295, No. 9).

a. Use a 0.01 significance level to test the claim that the female survivors and male survivors have different rates of thyroid diseases.

$p_1 = p_2$

$p_1 - p_2 = 0$

$\alpha = 0.01$

2-tailed test

$n_1 = 2739$

$x_1 = 1397$

$\hat{p}_1 = \frac{1397}{2739}$

$\hat{p}_1 \approx 0.510$

$\hat{q}_1 = 0.490$

$n_2 = 1352$

$x_2 = 436$

$\hat{p}_2 \approx 0.322$

$\hat{q}_2 = 0.678$

$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$

$\bar{p} = \frac{1397+436}{2739+1352}$

$\bar{p} \approx 0.448$

$\bar{q} = 0.552$

$H_0: p_1 = p_2$ vs $H_a: p_1 \neq p_2$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

$$z = \frac{(0.510 - 0.322) - (0)}{\sqrt{\frac{(0.448)(0.552)}{2739} + \frac{(0.448)(0.552)}{1352}}}$$

$\rightarrow z \simeq 11.37$

p-value $= 2P(z > \text{test stat})$

$= 2P(z > 11.37)$

$= 2[1 - P(z < 11.37)]$

$= 2[1 - 0.9999]$

$= 2[0.0001] < \alpha = 0.05$

reject H_0

0.0002 p-value

There's sig. evidence to support the claim the female survivors & male survivors have different rates of thyroid diseases.

b. Construct the confidence interval corresponding to the hypothesis test conducted with a 0.01 significance level.

$$\hat{p}_1 - \hat{p}_2 - E < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + E$$

$$0.510 - 0.322 - E < p_1 - p_2 < 0.510 - 0.322 + E$$

$$0.188 - 0.0409 < p_1 - p_2 < 0.188 + 0.0409$$

$$\boxed{0.1471 < p_1 - p_2 < 0.2289}$$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

2.575

$$E = z_{0.005}\sqrt{\frac{(.510)(.490)}{2739} + \frac{(.322)(.678)}{1352}}$$

$E \approx$

c. What conclusion does the confidence interval suggest?

We are 99% confident the the difference between the rates of thyroid disease between female survivors and male survivors lies between 0.1471 and 0.2289. Since 0 is not a likely difference, the CI suggest the these rates are different.

9.3 INFERENCES ABOUT TWO MEANS: INDEPENDENT SAMPLES

INDEPENDENT SAMPLES WITH σ_1 AND σ_2 UNKNOWN AND NOT ASSUMED EQUAL

DEFINITION

Two samples are **independent** if the sample values from one population are not related to or somehow naturally paired or matched with the sample values from the other population.

Two samples are **dependent** if the sample values are paired.

Inferences about Means of Two Independent Populations, With σ_1 and σ_2 Unknown and Not Assumed to be Equal

NOTATION

Population 1:

$\mu_1 =$ mean of pop. 1 assumed true under the null

$s_1 =$ sample standard deviation from pop. 1

σ_1 is unknown

$\bar{x}_1 =$ sample mean from pop. 1

$n_1 =$ sample size of pop. 1

The corresponding notations for μ_2, s_2, $\bar{x}_2$, σ_2, and n_2 apply to population 2.

REQUIREMENTS

1. σ_1 and σ_2 are unknown and it is not assumed that σ_1 and σ_2 are equal.

2. The 2 samples are independent.

3. Both samples are simple random samples.

4. Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

HYPOTHESIS TEST STATISTIC FOR TWO MEANS: INDEPENDENT SAMPLES

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Degrees of Freedom: When finding critical values or p-value, use the following for determining the number of degrees of freedom.

1. In this book we use the conservative estimate: df = smaller of $n_1 - 1$ and $n_2 - 1$.

2. Statistical software packages typically use the more accurate but more difficult estimate given below:

$$\text{df} = \frac{(A+B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}}, \quad A = \frac{s_1^{\,2}}{n_1}, \quad B = \frac{s_2^{\,2}}{n_2}$$

P-**values and critical values:** Use Table A-3.

CONFIDENCE INTERVAL ESTIMATE OF $\mu_1 - \mu_2$: INDEPENDENT SAMPLES

The confidence interval estimate of the difference $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 - E < \mu_1 - \mu_2 < \bar{x}_1 - \bar{x}_2 + E$$

$$E = t_{d.f.,\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

and the number of degrees of freedom df is as described above for hypothesis tests.

EQUIVALENCE OF METHODS

Example 1: Determine whether the samples are independent or dependent.

a. To test the effectiveness of Lipitor, cholesterol levels are measured in 250 subjects before and after Lipitor treatments.

Dependent

b. On each of 40 different days, the author measured the voltage supplied to his home and he also measured the voltage produced by his gasoline powered generator.

Independent

Example 2: Assume that the two samples are independent simple random samples selected from normally distributed populations. Do not assume that the population standard deviations are equal. A simple random sample of 13 four-cylinder cars is obtained, and the braking distances are measured. The mean braking distance is 137.5 feet and the standard deviation is 5.8 feet. A SRS of 12 six-cylinder cars is obtained and the braking distances have a mean of 136.3 feet with a standard deviation of 9.7 feet (based on Data Set 16 in Appendix B).

a. Construct a 90% CI estimate of the difference between the mean braking distance of four-cylinder cars and six-cylinder cars.

$n_1 = 13$
$\bar{x}_1 = 137.5$
$s_1 = 5.8$
smaller $n_2 = 12$
$\bar{x}_2 = 136.3$
$s_2 = 9.7$
$\alpha = 0.10$
$\alpha/2 = 0.05$
$d.f. = 12 - 1 = 11$

$$E = t_{d.f., \alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$E = t_{11, 0.05}\sqrt{\frac{(5.8)^2}{13} + \frac{(9.7)^2}{12}}$$

$$E = 1.796(3.2293) \approx 5.800$$

$$\bar{x}_1 - \bar{x}_2 - E < \mu_1 - \mu_2 < \bar{x}_1 - \bar{x}_2 + E$$

$$137 - 136.3 - 5.800 < \mu_1 - \mu_2 < 137 - 136.3 + 5.800$$

$$\boxed{-5.1 < \mu_1 - \mu_2 < 6.5}$$

b. Does there appear to be a difference between the two means?

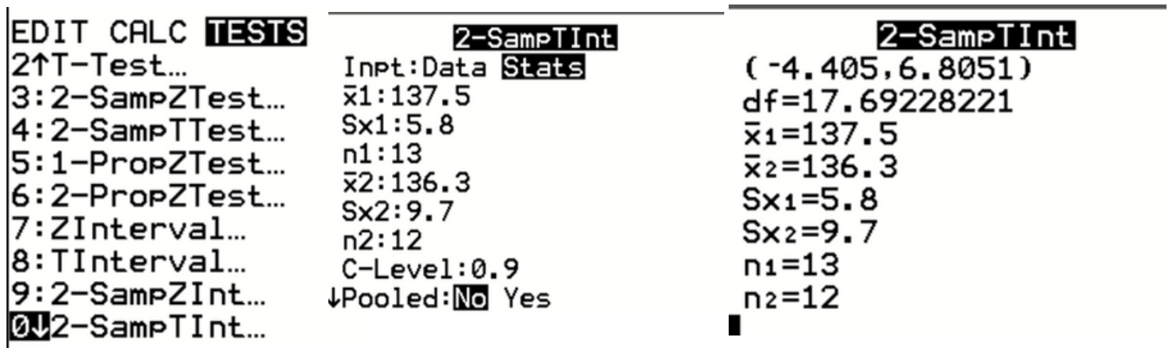
No since 0 is a likely value of the difference between the population means.

c. Use a 0.05 significance level to test the claim that the mean braking distance of four-cylinder cars is greater than the mean braking distance of six-cylinder cars.

A 1-tailed test at $\alpha = 0.05$ will yield the same conclusion as the 90% CI for the diff of the pop. means.

There is not significant evidence at the 5% level to support the claim that the mean braking distance of four-cylinder cars is greater than the mean braking distance of six-cylinder cars.

TI-83/84 PLUS



9.4 INFERENCES FROM DEPENDENT SAMPLES

Key Concept…

In this section we present methods for testing hypotheses and constructing confidence intervals involving the __mean__ of the __differences__ of the __value__ of two __dependent samples__. With __dependent__ samples, there is some __relationship__ whereby each value in one sample is __paired__ with a __corresponding__ value in the other sample. Here are two typical examples of dependent samples:

- Each pair of sample values consists of two measurements from the __same__ subject
- Each pair of sample values consists of a __matched pair__.

Because the hypothesis test and CI use the same __distribution__ and __standard deviation__, they are __equivalent__ in the sense that they result in the __same conclusions__. Consequently, the __null__ hypothesis that the __mean difference equals 0__ can be tested by determining whether the __confidence interval__ includes __0__. There are no exact procedures for dealing with __dependent__ samples, but the __t distribution__ serves as a reasonably good approximation, so the following methods are commonly used.

Inferences about Means of Two Dependent Populations

NOTATION

d = individual difference between the 2 values in the matched pair

s_d = sample standard dev. of the paired differences

μ_d = mean value of diff. d for pop. of all pairs of data

$\bar{d}$ = mean value of the paired differences for the paired sample data

n = # of pairs of data

REQUIREMENTS

1. The sample data are paired/dependent.
2. The samples are S R S.
3. Either or both of these conditions are satisfied: The number of pairs of data is large ($n>30$) or the pairs of values have distributions that are from a population that is approximately normal.

HYPOTHESIS TEST FOR DEPENDENT SAMPLES

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

Degrees of Freedom: $n-1$

***P*-values and critical values:** Use Table A-3.

CONFIDENCE INTERVALS FOR DEPENDENT SAMPLES

$$\bar{d} - E < \mu_d < \bar{d} + E$$

where

$$E = t_{d.f.,\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right)$$

and

$$d.f. = n-1$$

Example 1: Assume that the paired sample data are SRSs and that the differences have a distribution that is approximately normal.

a. Listed below are BMIs of college students.

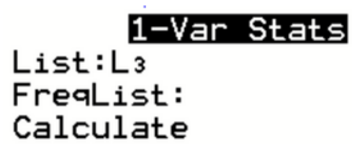
April BMI	20.15	19.24	20.77	23.85	21.32
September BMI	20.68	19.48	19.59	24.57	20.96

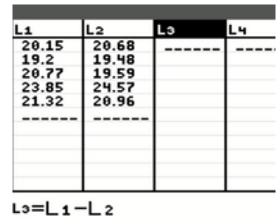
i. Use a 0.05 significance level to test the claim that the mean change in BMI for all students is equal to 0.

$H_0: \mu_d = 0$

$H_a: \mu_d \neq 0$

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$





$$t = \frac{0.002 - 0}{\frac{0.7745}{\sqrt{5}}}$$

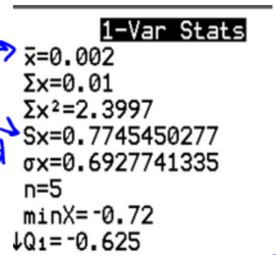
$\rightarrow t \approx 0.00577$

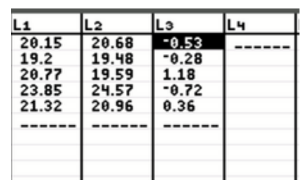
p-value > 0.20

and 0.20 > 0.05

we fail to reject Ho.

$\bar{d} \rightarrow$ $s_d \rightarrow$

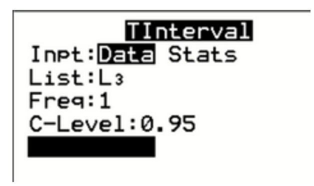


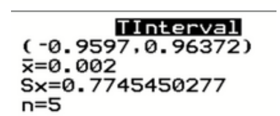


Not sig. evid. @ 5% level to claim that the BMI is different.

ii. Construct a 95% CI estimate of the change in BMI during freshman year.

$$\boxed{-0.9597 < \mu_d < 0.96372}$$





iii. Does the CI include zero, and what does that suggest about BMI during freshman year?

The CI includes zero which suggests that there's no significant difference in BMI from September to April at the 5% level.