Packet 9: One-Sample Hypothesis Test for the Population Mean Textbook pages: 586 – 589; 597 – 601

After completing this material, you should be able to:

- conduct a test of hypothesis for μ using the appropriate format.
- state when it is valid to use this procedure.
- define Type I and Type II errors in terms of the problem.
- discuss the consequences of these errors.

Recall: What are the steps in a hypothesis test?

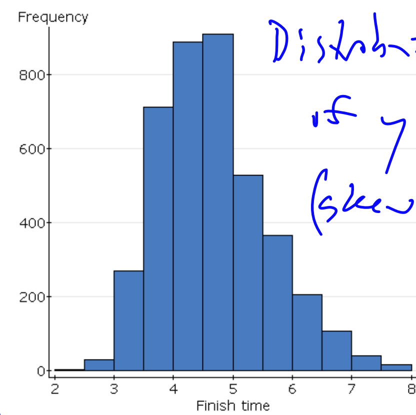
1. State the null & the alternative hypothesis H_0 & H_a;
2. Determine the significance level α & the decision rule;
3. Calculate the test statistic & corresponding probability;
4. Reach our conclusion; reject (or not) the null H_0.

Example: The Flying Pig Marathon is held in Cincinnati every May. The StatCrunch output below summarizes the time it took for all runners in 2012 to finish the 26.2 mile race. An organizer of the marathon plans to sample 35 finishers from 2016 to see if the time it takes to complete the race has changed.

$n = 35$

Column	Mean	Std. dev.	Min	Q1	Median	Q3	Max
Finish time	4.6738087	0.89921376	2.3677778	4.0030556	4.5675	5.2158333	7.7769444

μ σ



Distribution of y (skewed)

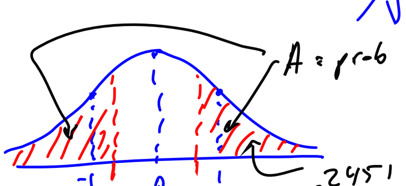
Describe the sampling distribution of the sample mean finish time when samples of 35 runners is taken.

1. shape: normal ($n \geq 30$)
2. center: $\mu_{\bar{y}} = \mu_y = \mu = 4.6738$
3. spread: $\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{.8992}{\sqrt{35}} = .1520$

The sample of 35 runners is taken and the sample mean is found to be 4.569 hours. Use the sample to test the claim made by the organizer at a significance level of 0.05.

$\alpha = .05$ $\bar{y} = 4.569$

1. $H_0: \mu = 4.6738$; $H_a: \mu \neq 4.6738$
2. $\alpha = .05$; Reject the null in favor of H_a if prob. < α.
3. TS: $z = \frac{\bar{y} - \mu}{\sigma/\sqrt{n}} = \frac{4.569 - 4.6738}{.152} = -.69$



$A = 2 * .2451 = .4902 > \alpha$

Two-sided

4. With α of .05 we fail to reject the null: there is not significant evidence that the mean has changed.

Example: A city has instituted an educational program that they hope will reduce the number of violent crimes committed in the city. The program is quite expensive, so the mayor wants to be sure that it is worthwhile. In the past, the city experienced 5.2 violent crimes per day, on average. After implementing the program, a random sample of 36 days yielded an average crime rate of 4.7 violent crimes with a standard deviation of 4.16. Test the appropriate hypotheses to see if the program is successful at reducing violent crimes using a significance level of 0.10.

$\alpha = .10$

What variable was recorded? What parameter will be tested?

Violent crimes per day (quantitative)

What hypotheses should be tested?

$H_0: \mu = 5.2$ $H_a: \mu < 5.2$ violent crimes per day.

Assign the appropriate notation to the values given in the problem.

$\mu = 5.2$ $\bar{y} = 4.7$ (ⁿ̈ - no σ)

$n = 36$ $s = 4.16$

Problem!!	We'd like to use $z = \frac{\bar{y} - \mu}{\sigma/\sqrt{n}}$ - but we don't know σ. Idea: let's just use s in place of σ. - so we create a new test statistic:

What is the effect of replacing σ with s? We no longer have a normal sampling distribution for the sample mean. We have a t-distribution:

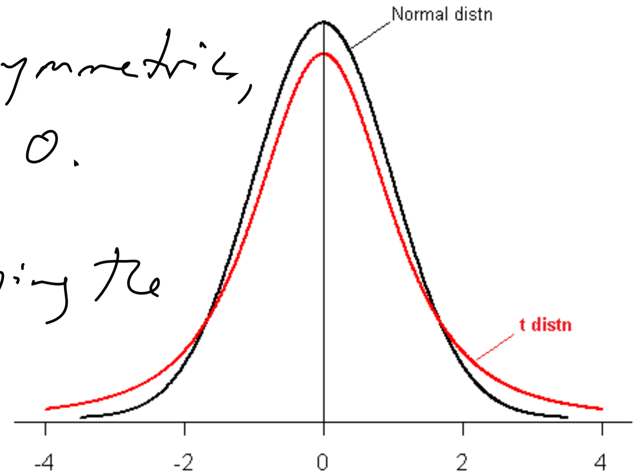
$$t = \frac{\bar{y} - \mu}{s/\sqrt{n}}$$

t is not normal - but it's really close. $\bar{y}$ follows a t-distribution w/ n-1 degrees of freedom.

How is the t-distribution different than the normal distribution?

Similar: bell-shaped, symmetric, unimodal, centered at 0.

Different: t requires knowing the degrees of freedom.
t is a little squatter than a normal - wider & shorter.



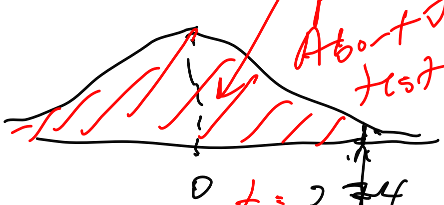
Technically we would need to use a *t-table* to find probabilities associated with the t-distribution. Because the t-distribution is indexed by its degrees of freedom, it is impossible to have a single table of these probabilities. Instead of using a table to find probabilities, we'll rely on StatCrunch to calculate these probabilities.

Let's look at a couple of examples of p-values to refresh how the probabilities are found:

Suppose we are testing the following null hypothesis, H_0: $\mu = 5$, and the test statistic calculated is $t = 2.74$. For each of the three alternatives below, draw a picture representing the p-value for the test:

H_a: $\mu < 5$

huge prob. Abort the test

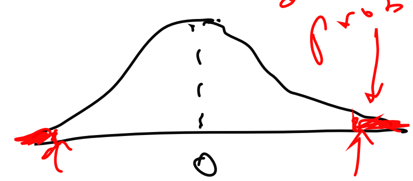


H_a: $\mu > 5$

looks like we'll reject the null

H_a: $\mu \neq 5$

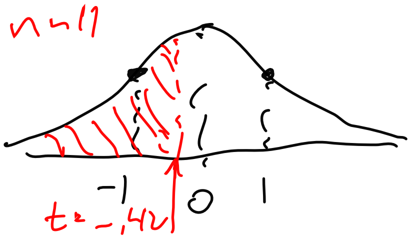
double this probability



Suppose we are testing the following null hypothesis, H_0: $\mu = 12000$, and the test statistic calculated is $t = -0.42$. For each of the three alternatives below, draw a picture representing the p-value for the test:

H_a: $\mu < 12000$

We can't reject the null



H_a: $\mu > 12000$

We can't reject the null

H_a: $\mu \neq 12000$

Huge probability!

Can't reject the null.

What patterns do you notice in the calculation of the p-values and the alternative hypotheses?

(1) The < & > probabilities sum to 1: A & 1-A

(2) The ≠ probability is twice one of the other two probabilities (the smaller of the two).

Back to the example: A city has instituted an educational program that they hope will reduce the number of violent crimes committed in the city. The program is quite expensive, so the mayor wants to be sure that it is worthwhile. In the past, the city experienced 5.2 violent crimes per day, on average. After implementing the program, a random sample of 36 days yielded an average crime rate of 4.7 violent crimes with a standard deviation of 4.16. Test the appropriate hypotheses to see if the program is successful at reducing violent crimes using a significance level of 0.10.

$\alpha = .10$

① $H_0: \mu = 5.2$ versus $H_a: \mu < 5.2$

② $\alpha = .10$ Reject H_0 if prob $< \alpha = .10$.

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
~~μ ≠~~	~~4.7~~	~~0.69333333~~	~~35~~	~~-0.72115385~~	~~0.4756~~

$\neq$ .2378

Hypothesis test results:

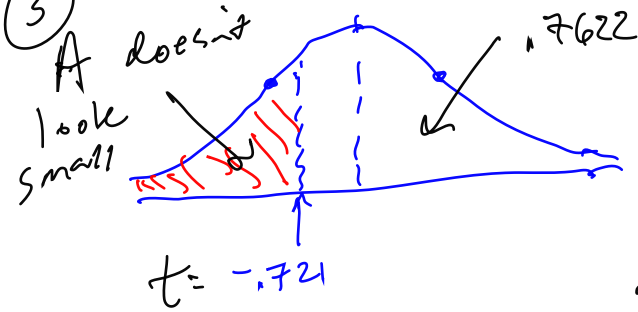
Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ <	4.7	0.69333333	35	-0.72115385	0.2378

Hypothesis test results:

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
~~μ >~~	~~4.7~~	~~0.69333333~~	~~35~~	~~-0.72115385~~	~~0.7622~~

sum to 1

③



A ≈ .2378

prob of a t that extreme or more so is $.2378 > \alpha = .10$

④ Because the probability of a value of t of $-.721$ or less (more extreme to the left) is $> \alpha$, we fail to reject the null of no change in the number of violent crimes.

Bad news - we spent all that money but didn't have a significant impact.

What type of error is possible in the hypothesis test which was conducted?

We failed to reject the null

Type II - fail to reject a false null.

Maybe our program was successful - but the sample data just doesn't show it.

Example: Acetaminophen is an active ingredient found in hundreds of over-the-counter and prescription medicines, such as pain relievers, cough suppressants, and cold medications. It is safe and effective when used correctly, but there are consequences when too much or too little is taken. A researcher believes that the mean amount of acetaminophen per tablet in a particular brand of cold medicine differs from the claimed dosage of 600 mg. To investigate the claim, a random sample of 100 tablets is taken, and the average acetaminophen content was found to be 603.7 mg with a standard deviation of 9.782 mg.

Ha

$\mu = 600$ $\bar{Y} = 603.7$

$n = 100$ $s = 9.782$

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ ≠	603.7	0.9782	99		0.0003

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	603.7	0.9782	99		0.9999

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	603.7	0.9782	99		0.0001

Those add to 1 for < & > probabilities

1 $H_0: \mu = 600$

$H_a: \mu \neq 600$

2 $\alpha = .01$ Reject H_0 if prob $< .01 = \alpha$.

3 T.S. $t = \dfrac{\bar{y} - \mu}{s/\sqrt{n}}$

$= \dfrac{603.7 - 600}{9.782/\sqrt{100}} = \dfrac{3.7}{.9782} = \boxed{3.78}$

Whoa! That's crazy big.

prob = .0003 < α

4 Reject the null of $\mu = 600$ in favor of the alternative $\mu \neq 600$; in fact it appears that this product contains more of the drug than it should.

Example: A local business is considering a new location. A market analyst has determined that the location will have the greatest likelihood of success if the number of people who pass the storefront during the lunch rush averages 100 or more people. The storefront is observed for 25 weekdays.

$n = 25$ Randomly?

What variable will be recorded?

# of people passing at lunch hour. (each day)

What hypotheses should be tested in this study?

1 $H_0: \mu = 100$ $H_a: \mu > 100$

In the context of the problem, describe a Type I and Type II error. What could potential consequences be if these errors were to occur?

Type I: reject a true null. $\mu = 100$ but we claim $\mu > 100$; so we open the business, but fail for lack of customers (100 is on the edge).

Type II: fail to reject a false null. In fact $\mu > 100$, & we should be profitable. We fear failure, so we don't open (even though we could have succeeded).

Remember that the significance level controls the probability of a Type I error occurring – based on your description, what significance level would you recommend setting?

(2) $\alpha = .01$; Reject H_0: probability $< \alpha = .01$

I fear a Type I error, & fear it a lot (α small).

A random sample of 25 days was taken, and the number of people passing by the storefront was recorded. The mean number of people who passed the store was 105.678 with a standard deviation of 18.932. Use these to test the hypotheses above at the significance level that was determined.

$s = 18.932$

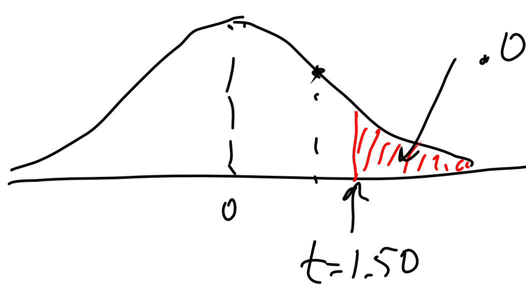
$\bar{y} = 105.678$

(3) TS.

$$t = \frac{105.678 - 100}{18.932/\sqrt{25}} = 1.50$$

Possible p-values (only one is appropriate):
$\neq$ p-value = 0.1468 ← Twice .0734
$>$ p-value = 0.0734
$<$ p-value = 0.9266

} sum to 1



.0734 = prob. of a result this extreme or more

so: $.0734 > .01$

(4) We fail to reject the null of $\mu = 100$ in favor of opening a store! We fear that we won't have enough customers.*

* Both of the assumptions were false!

When the t-distribution is used for an inference, what assumptions must be satisfied?

(1) A random sample of quantitative data

(2) Data is sampled from a normal distribution or n is very large ($n > 100$).