CHAPTER 9
Testing a Claim

9.3
Tests About a Population Mean

The Practice of Statistics, 5th Edition
Starnes, Tabor, Yates, Moore

Bedford Freeman Worth Publishers

Tests About a Population Mean

Learning Objectives

After this section, you should be able to:

- STATE and CHECK the Random, 10%, and Normal/Large Sample conditions for performing a significance test about a population mean.
- PERFORM a significance test about a population mean.
- USE a confidence interval to draw a conclusion for a two-sided test about a population parameter.
- PERFORM a significance test about a mean difference using paired data.

Introduction

Inference about a population mean uses a distribution with degrees of freedom, except in the rare case when the population standard deviation σ is known.

Carrying Out a Significance Test for μ

In an earlier example, a company claimed to have developed a new AAA battery that lasts longer than its regular AAA batteries. Based on years of experience, the company knows that its regular AAA batteries last for 30 hours of continuous use, on average.

An SRS of 15 new batteries lasted an average of 33.9 hours with a standard deviation of 9.8 hours.

Do these data give convincing evidence that the new batteries last longer on average?

To find out, we must perform a significance test of

$H_0: \mu = 30$ hours
$H_a: \mu > 30$ hours

where μ = the true mean lifetime of the new deluxe AAA batteries.

Check conditions (the same conditions as for a confidence interval of a population mean).

What are the 3 conditions?

Carrying Out a Significance Test for

The test statistic measures in standardized units how far the sample result diverges from the parameter value specified by H_o.

$$\text{test statistic} = \frac{\text{statistic - parameter}}{\text{standard deviation of statistic}}$$

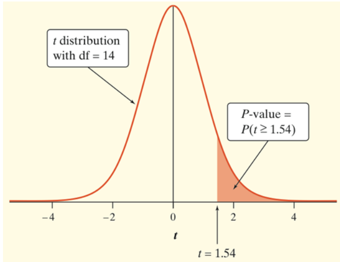
For the population mean, because is not known, we use rather than , and we use instead of .

$$t = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}}$$

In our example, t =

Carrying Out a Significance Test for μ

	Upper-tail probability *p*		
df	.10	.05	.025
13	1.350	1.771	2.160
14	1.345	1.761	2.145
15	1.341	1.753	3.131
	80%	90%	95%
	Confidence level *C*		



Using the table, the p-value for this test is between 0.05 and 0.10.

Using technology, use "tcdf(lower,upper, df) = 0.0729. This is a more accurate value.

Note: The t-table (Table B) reads the opposite of the z-table.

- The z-table has the z-values on the outside and the probabilities on the inside.
- The t-table has the degrees of freedom, the confidence levels and the probabilities on the outside and the t-values on the inside.

(compare the 2 tables)

Also, Table B only shows values for positive values of t and it shows the area to the RIGHT of the t-value.

Table B has limitations of only showing df of 1 - 30, then skips to 40, 50, etc. It also has limited probabilities.

Given these limitations, use the calculator.

Just in Case: Using Table B

- If the df value you need isn't provided in Table B, use the next **lower** df that is available.
- Table B shows probabilities only for positive values of *t*. To find a *P*-value for a negative value of *t*, we use the symmetry of the *t* distributions.

On the calculator:

[2nd] [VARS] tcdf(lower, upper, df)

tcdf(lower:1.54, upper:10,000, df:14) = .0729268628

For a 2-sided test, multiply by 2

The Relationship between Two-Sided Tests and Confidence Intervals

This relationship is even stronger for means than for proportions because the standard error is used in both calculations:

Test statistic: $t = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}}$ Confidence interval: $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$

Example: a 2-sided tests at significance level 0.05 and a 95% confidence interval give similar information about the parameter.

When the 2-sided test rejects H_o, the confidence interval will not contain the hypothesized value μ_o.

When the 2-sided tests fails to reject H_o, the confidence interval will contain μ_o.

Inference for Means: Paired Data

Comparing two observations on the same individuals or one observation on each of 2 similar individuals result in paired data.

Compare by finding the differences in each pair (subtracting).

Check the same conditions, and then use one-sample procedures to perform inference about the mean difference, μ_d.

Example: Paired data and one-sample t procedures

Researchers designed an experiment to study the effects of caffeine withdrawal. They recruited 11 volunteers who were diagnosed as being caffeine dependent to serve as subjects. Each subject was barred from coffee, colas, and other substances with caffeine for the duration of the experiment.

During one two-day period, subjects took capsules containing their normal caffeine intake. During another two-day period, they took placebo capsules. The order in which subjects took caffeine and the placebo was randomized. At the end of each two-day period, a test for depression was given to all 11 subjects.

Researchers wanted to know if being deprived of caffeine would lead to an increase in depression.

Results of a caffeine-deprivation study

Subject	Depression (caffeine)	Depression (placebo)	Difference (placebo – caffeine)
1	5	16	11
2	5	23	18
3	4	5	1
4	3	7	4
5	8	14	6
6	5	24	19
7	0	6	6
8	0	3	3
9	2	15	13
10	11	12	1
11	1	0	-1

Example: Paired data and one-sample t procedures

(a)Why did researchers randomly assign the order in which subjects received placebo and caffeine?

Example: Paired data and one-sample t procedures

(b) Carry out a test to investigate the researchers' question.

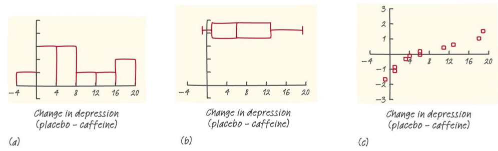
State

Example: Paired data and one-sample t procedures

Plan: If the conditions are met, we should do a paired t test for μ_d

Example: Paired data and one-sample t procedures

• Normal/Large Sample: We don't know whether the actual distribution of difference in depression scores (placebo -caffeine) is Normal. With such a small sample size (n = 11), we need to examine the data to see if it's safe to use t procedures.

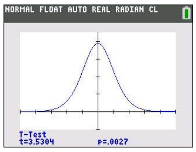


The histogram has an irregular shape with so few values; the boxplot shows some right-skewness but not outliers; and the Normal probability plot is slightly curved, indicating some slight skewness. With no outliers or strong skewness, the t procedures should be pretty accurate.

Example: Paired data and one-sample t procedures

Do: We entered the differences in list1 and then used the calculator's t-test with "Draw" command,

• Test statistic t = 3.53



P-value 0.0027, which is the area to the right of t = 3.53 on the t distribution curve with df = 11 -1 = 10.

Conclude: (a) Because the P-value is 0.0027, we reject H_0: $\mu_d = 0$.

We have convincing evidence that the true mean difference (placebo - caffeine) in depression score is positive for subjects like these.

Using Tests Wisely

Carrying out a significance test is often quite simple, especially if you use a calculator or computer. Using tests wisely is not so simple. Here are some points to keep in mind when using or interpreting significance tests.

How Large a Sample Do I Need?

- A smaller significance level requires stronger evidence to reject the null hypothesis.
- Higher power gives a better chance of detecting a difference when it really exists.
- At any significance level and desired power, detecting a small difference between the null and alternative parameter values requires a larger sample than detecting a large difference.

Using Tests Wisely

Statistical Significance and Practical Importance

When a null hypothesis ("no effect" or "no difference") can be rejected at the usual levels ($\alpha = 0.05$ or $\alpha = 0.01$), there is good evidence of a difference. But that difference may be very small. When large samples are available, even tiny deviations from the null hypothesis will be significant.

Beware of Multiple Analyses

Statistical significance ought to mean that you have found a difference that you were looking for. The reasoning behind statistical significance works well if you decide what difference you are seeking, design a study to search for it, and use a significance test to weigh the evidence you get. In other settings, significance may have little meaning.

Tests About a Population Mean

Section Summary

In this section, we learned how to…

- STATE and CHECK the Random, 10%, and Normal/Large Sample conditions for performing a significance test about a population mean.
- PERFORM a significance test about a population mean.
- USE a confidence interval to draw a conclusion for a two-sided test about a population parameter.
- PERFORM a significance test about a mean difference using paired data.