

<http://www.Rcoding.org>

Lecture Notes -- Nicholas Karlson

Math 1001: Introduction to Statistics – recall that lecture notes are available on the web at www.Rcoding.org as well as screencasts (these include audio). Posting ANONYMOUS questions, comments, and suggestions can be done at www.Rcoding.org.

Next, one sample t -test lecture continued...

Note: there is a lot of terminology/jargon associated with a one sample t -test (as you can see in your textbook). Working examples is a great way to make sense of this jargon and learn how to apply statistical theory in practice to real world

problems. So we will work a lot of examples!

Last time we covered the basic concepts for a one sample *t*-test. Recall:

SAMPLE MEAN: \bar{x} (pronounce it x-bar) of n numbers is found by summing all n numbers and then dividing this sum by n .

Using mathematical notation, the definition is:

$$\bar{x} = (\sum_{i=1}^n x_i)/n = (x_1 + x_2 + x_3 + \dots + x_n)/n$$

Example: for the two numbers 6 and 4 the sample mean \bar{x} is $(6 + 4)/2$ which equals 5.

Also:

SAMPLE VARIANCE: S^2 (pronounce it S-squared) of n numbers is defined as:

$$S^2 = \{\sum_{i=1}^n (x_i - \bar{x})^2\}/(n - 1)$$

And the square root of the sample variance S^2 equals S is called the **SAMPLE STANDARD DEVIATION**.

Example: for the two numbers 6 and 4 the sample variance is:

$$\{(6 - 5)^2 + (4 - 5)^2\}/(2 - 1)$$

which equals 2. Thus $S^2 = 2$, and $S = \sqrt{2}$, i.e., the sample standard deviation of 6 and 4 is the square root of 2.

More key concepts previously covered:

The t probability distribution has a parameter called degrees of freedom (d.f.). This parameter determines the shape of the t -distribution. **When a t distribution's degrees of freedom parameter is 30 or more, then often in practice it is considered "equivalent" to a normal distribution with mean 0 and**

variance 1 (i.e., the standard normal distribution). With smaller d.f. values, the t -distribution is still symmetrical and bell shaped like a standard normal, but with “fatter” tails than the standard normal.

Statistical Hypothesis Testing: choosing between two well defined hypotheses using a test statistic calculated from observed data.

Null Hypothesis (H_0): a state of the world that might be tentatively accepted over a competing hypothesis (H_A) as a result of a statistical test calculated from observed data.

Example: H_0 : the average age in Europe is 26.2 years.

Alternative Hypothesis (H_A): a state of the world that is competing against the null hypothesis.

Example: H_A : the average age in Europe is less than 26.2 years.

Test Statistic (TS): A statistic calculated from observed data. If the observed value of the TS falls into a test's critical region, the null hypothesis is rejected in favour of the alternative hypothesis.

The Distribution of a Test Statistic (TS): The probability distribution of the test statistic.

Example: For a one sample t -test with $n = 10$ data points (observations), the test statistic has a t distribution with parameter (degrees of freedom) $n - 1 = 10 - 1 = 9$.

P-value: Assuming the null hypothesis is true, a P-value is the probability of obtaining a test statistic value as extreme (or more extreme)

than the test statistic that was calculated from the actual data.

Significance Level of a Test: Often denoted by the Greek letter α (alpha), is the probability of rejecting the null hypothesis when it is in fact true. Rejecting the null hypothesis when it should be accepted is also referred to as a Type I error. Standard levels of significance are 5% and 1%. Alpha is used to determine the critical value(s) and region(s) for a test and it is also the probability of a Type I error.

One Sample *t*-Test: A small-sample statistical test used to investigate the location of the mean of some population. The population data is assumed to be normal or close to

normally distributed. The data used for the test is selected randomly and independently.

Important result from statistical theory: If the population sampled is normally (or approximately) distributed then the statistic:

$$t = (\bar{x} - \mu) \div (S/\sqrt{n})$$

has a *t*-distribution with $n - 1$ degrees of freedom. Recall that \bar{x} is the sample mean, μ (μ) is the population's true mean (unknown in practice), S is the sample standard deviation, and n is the number of data points (observations).

The test statistic for a one-sample *t*-test is calculated by replacing μ in the above

equation by μ_0 (where μ_0 is the mean of the population under the null-hypothesis).

For a one-tailed t -test the critical value c is the point on the x -axis of the plotted t -distribution (with $n - 1$ d.f.) such that the area under the t -curve to the x -axis (starting at the critical value and heading away from 0) equals α (alpha).

The value of c is found from statistical tables or software that tabulates the value of c for a given area under a t -curve. So the smaller the level of significance (α), the more extreme a test statistic must be to pass into the rejection region of a test, and the smaller the probability of making a Type I error (α).

Accept or Reject the Null Hypothesis at the alpha-level of significance: If the calculated t -

test statistic is more extreme than the critical value, reject the null hypothesis in favour of the alternative hypothesis, otherwise accept the null hypothesis at significance level alpha.

NOTE: some researchers prefer simply to report the resulting P-value of a test. If a P-value is smaller than a given level of significance (e.g., α_1), then the test will reject the null hypothesis at significance level α_1 .

END OF RECAP

NEXT AN EXAMPLE

Points Highlighted in the upcoming example:

- 1. The setup of the null hypothesis and alternative hypothesis.**
- 2. The determination of the P-value.**

3. The concept and the use of the significance level (the example will use 5%).

Problem Summary: Is average SUV fuel efficiency 9 litres per 100km driven as stated by the manufacturer?

This is the null hypothesis or the claim we are testing. We are interested in an “average” which corresponds to the **true mean fuel-rating** for all such vehicles. **This is the parameter of interest and it is denoted by the Greek letter mu: μ .** The null hypothesis can be stated as:

$$H_0: \mu = 9$$

Is the manufacturer overstating the fuel rating? This is the alternative hypothesis we

wish to investigate. In other words is the fuel rating worse than advertised? Is it less than 9, i.e., is μ less than 9? **The alternative hypothesis can be stated as:**

$$H_A: \mu < 9$$

This is **a one-tailed or one-sided test** as we are interested in values for the true mean μ that are less than 9. This also implies that the critical value and **rejection region associated with this test will be located in the left tail of the t -distribution.**

In sum, the test can be written as:

$$H_0: \mu = 9$$

Versus (vs.)

$$H_A: \mu < 9$$

We also decide to test H_0 vs. H_A at a significance level of 5%.

Now we will **assume** that the population of Hybrid SUV's from which we are sampling is **normal or approximately normally distributed**.

Next assume that a random sample of 6 SUV's is taken and the fuel rating for each car is measured. Because the sample size is small ($n = 6$) a t -test will be used rather than a z -test (z -tests are based on the standard normal distribution and were covered previously).

The resulting 6 data values (fuel ratings) are 8.8, 10.2, 11.0, 9.2, 9.7, and 8.6.

In a nutshell, we need to use this data to **1)** calculate a test statistic, **2)** use a t -distribution with $n - 1 = 5$ degrees of freedom (d.f.) along

with the test statistic to calculate the test's P-value, and **3)** find a critical region associated with a 5% level of significance to see if the test statistic falls into it (rejecting the null hypothesis only if this happens).

SO:

1) Calculate a test statistic:

The test statistic as given in your textbook is:

$$t_0 = (\bar{x} - \mu_0) \div (S/\sqrt{n})$$

To calculate t_0 we need the values of \bar{x} (sample mean), μ_0 (mean under the null hypothesis), S (the sample standard deviation), and n (sample size).

From the problem statement we have:

$$\mu_0 = 9 \text{ and } n = 6$$

And using the formulas:

$$\bar{x} = (\sum_{i=1}^n x_i)/n = (x_1 + x_2 + x_3 + \dots + x_n)/n$$

$$s^2 = \{\sum_{i=1}^n (x_i - \bar{x})^2\}/(n - 1)$$

We get:

$$\bar{x} = \frac{8.8 + 10.2 + 11.0 + 9.2 + 9.7 + 8.6}{6} = 9.583$$

And

$$s =$$

$$\left\{ \frac{(8.8-9.6)^2 + (10.2-9.6)^2 + (11.0-9.6)^2 + (9.2-9.6)^2 + (9.7-9.6)^2 + (8.6-9.6)^2}{5} \right\}^{1/2}$$

which equals 0.909 (or 0.91 rounded).

So:

$$t_0 = (9.583 - 9) \div (0.909/\sqrt{6})$$

This equals 1.571 (or 1.57 rounded).

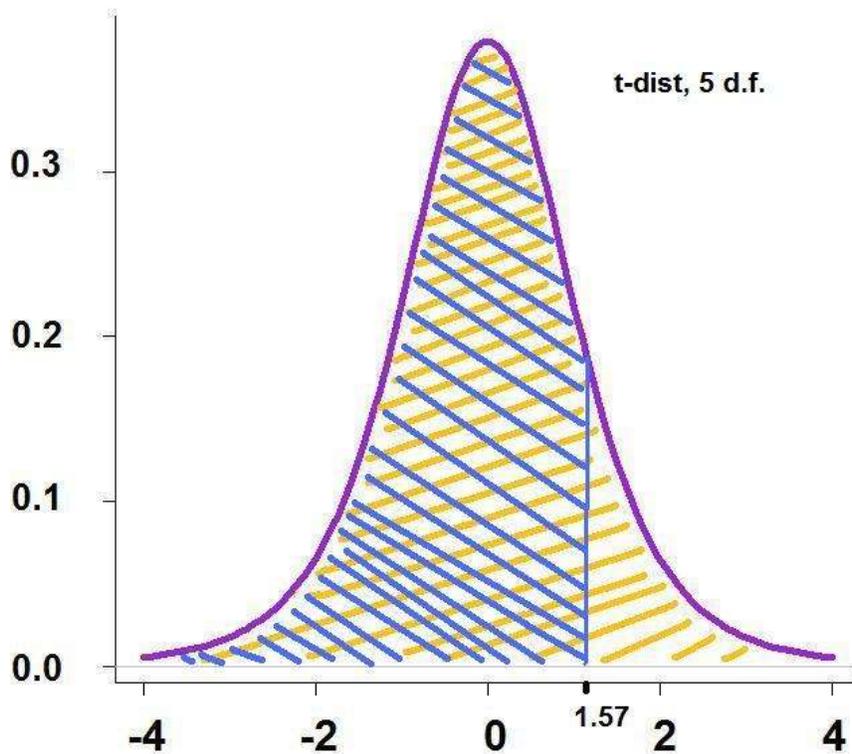
NEXT:

2) use a t-distribution with $n - 1 = 5$ degrees of freedom (d.f.) along with the test statistic to calculate the test's P-value.

Recall: Assuming the null hypothesis is true, a P-value is the probability of obtaining a test statistic value as extreme (or more extreme) than the test statistic that was calculated from the actual data.

For our SUV test, the P-value is 0.91. This can be seen in the following graph (the area of the blue and gold hash is the probability of observing more extreme values of the test statistic assuming the null hypothesis is true):

The calculated test statistic = 1.57, the P-value is the area to the left of 1.57 (shaded blue) which = 0.91



Note finding this area (probability) is usually done with a computer. In R (see Rcoding.org), the command you type in is:

```
pt(c(1.57), df=5, lower.tail=TRUE)
```

The above R command will return 0.91.

Also the P-value is given when using Minitab to do the test (rather than doing it by hand):

Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help

Session

Test of $\mu = 9$ vs < 9

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
C1	6	9.583	0.909	0.371	10.331	1.57	0.912

One-Sample T: Mileage

Test of $\mu = 9$ vs < 9

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
Mileage	6	9.583	0.909	0.371	10.331	1.57	0.912

Worksheet1 ***

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mileage												
1	8.8												
2	10.2												
3	11.0												
4	9.2												
5	9.7												
6	8.6												
7													

Current Worksheet: Worksheet1 Editable

In R:

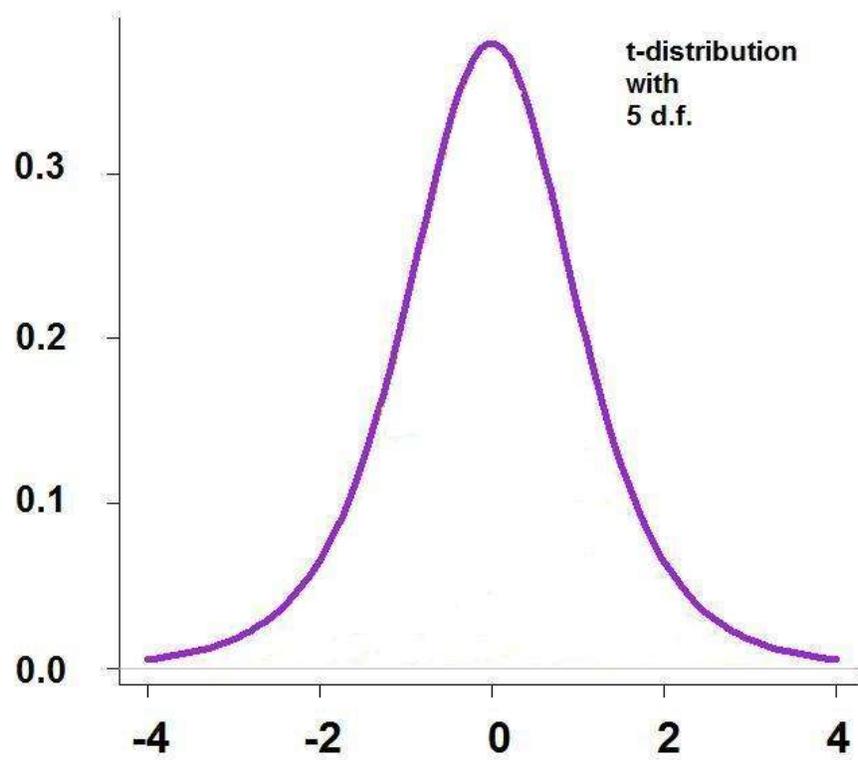
not be rejected since we have P-value = 0.91 > 0.05 = level of significance.

For completeness let's find the critical region and show that the test statistic 1.57 does not fall in this region.

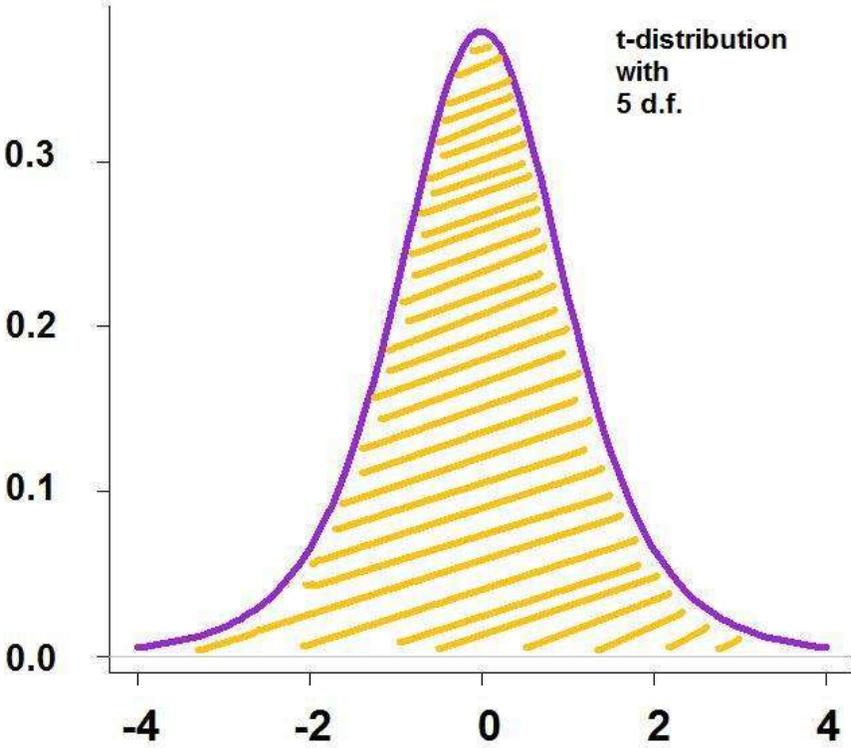
NEXT

3) Find a critical region associated with a 5% level of significance to see if the test statistic falls into it (rejecting the null hypothesis only if this happens).

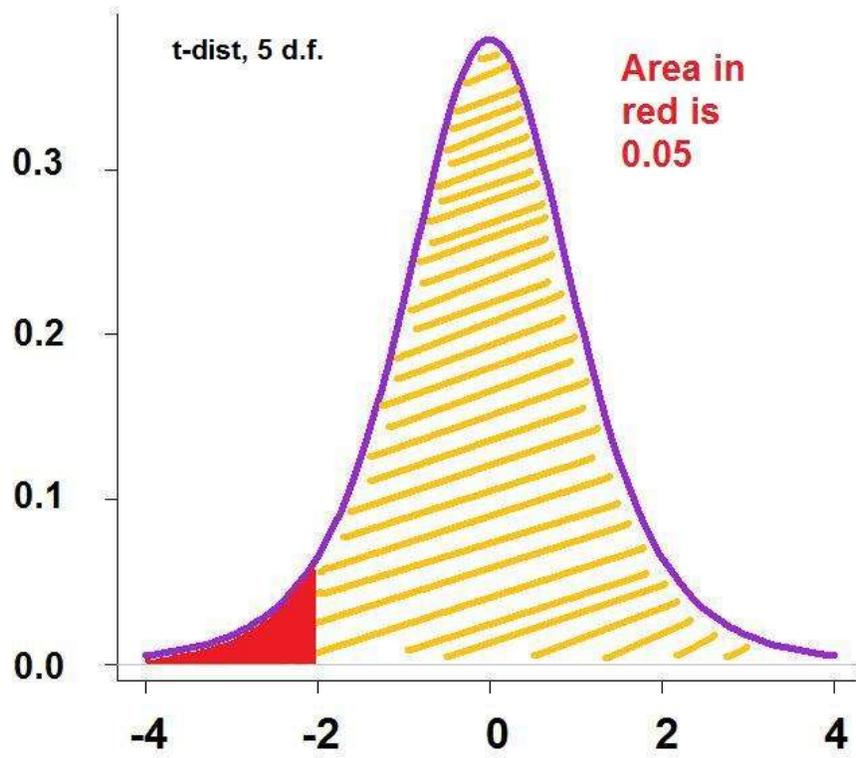
First let's try a graphical solution...

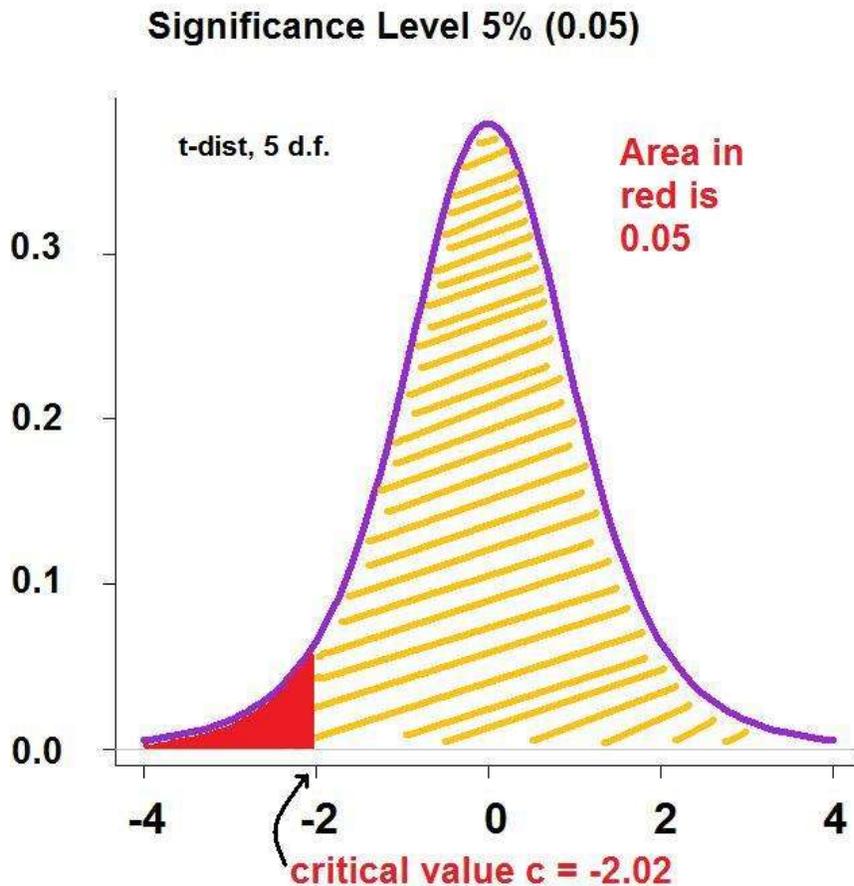


Total Area under t-curve equals 1



Significance Level 5% (0.05)





Note: the probability of making a **Type I** error (0.05 from the level of significance of the test) comes from the fact that **even if the null hypothesis is true, there is a 5% chance or 0.05 probability that the calculated test statistic will fall into the red area (the rejection region).**

Note finding c is usually done with a computer or tables in a book (as we have covered previously). In R (see Rcoding.org), the command to find c is:

```
qt (c (.05) ,df=5 ,lower.tail=TRUE)
```

The above R command will return -2.02.

TEST RESULT

The calculated test statistic is 1.57, it does not fall into the critical region, so do not reject the null hypothesis in favour of the alternative hypothesis at the 5% level of significance ($\alpha = 0.05$) .

Recall from previous lecture notes that α is also the probability of rejecting the null hypothesis when it should be accepted (i.e., the probability of making a Type I error).

The P-value is quite large as well, so **the data do not provide evidence against the manufacturer's claim of 9 litres for every 100 kilometres at any standard levels of significance (e.g. 1%, 5%, or even 10%).**

IN SUMMARY

For a one-sample t -test:

- 1) Determine what is to be tested**
- 2) Determine the null and alternative hypotheses**
- 3) Set a level of significance for the test**
- 4) Collect the data (a random sample from a population that is normal or approximately normally distributed)**
- 5) Calculate a critical region (based on the level of significance) for the test**
- 6) Calculate the test statistic and P-value**

7) State the result of the t-test, level of significance, and P-value

8) Report results

ESPECIALLY NOTE:

1. How we setup the null and alternative hypotheses.

2. How we determined the P-value for the test.

3. The concept and the use of the significance level (the example used 5%) as related to a test's rejection region and probability of a Type I error.

Questions?

Remember you can ask questions or send comments anonymously (visit www.Rcoding.org for more details).