

LESSON QUESTIONS: Hypothesis testing

FOCUS QUESTION: How can I tell whether the test group is different from the control group?

Contents

- [EXAMPLE 1: Load the consolidated sleep diary data](#)
- [EXAMPLE 2: Does subject 1 sleep 8 hours on average?](#)
- [EXAMPLE 3: Is it possible that the students in section 2 sleep 8 hours on average?](#)
- [EXAMPLE 4: Do the students in sections 2 and 3 sleep a different amount?](#)
- [EXAMPLE 5: Do section 2 and 3 students sleep differently at the 0.01 significance level?](#)
- [EXAMPLE 6: Do section 2 students sleep more than section 3 students?](#)
- [EXAMPLE 7: Do section 2 students sleep more than section 3 students \(fewer assumptions\)?](#)

EXAMPLE 1: Load the consolidated sleep diary data

```
load diaries.mat;

sleepHours = (wakeTimes - bedTimes)*24;
```

EXAMPLE 2: Does subject 1 sleep 8 hours on average?

```
[h1, p1, ci] = ttest(sleepHours(:, 1), 8);

fprintf(['Does subject 1 sleep 8 hours on average?\n\t' ...

        'h = %g, p = %g, ci = [%g, %g]\n'], h1, p1, ci);
```

Does subject 1 sleep 8 hours on average?

```
h = 0, p = 0.113092, ci = [7.21123, 8.09036]
```

Questions	Answers
What is the purpose of the <code>ttest</code> function?	The <code>ttest</code> uses knowledge about a sample to determine whether a population mean is likely to be different from a specified value. The sample should be drawn at random from a normally distributed population.
What should I conclude if this <code>ttest</code> returns <code>h1 = 0</code> ?	Based on the evidence provided by this sample, it is possible that the true mean average sleep time of subject 1 is 8 hours. However, you do not have evidence to make a definitive conclusion. (Presume innocent until proven guilty!)

Questions	Answers
Does <code>ttest</code> make any assumptions about the distribution of the population?	Yes, <code>ttest</code> assumes that the population is normally distributed.
When should I use <code>ttest</code> ?	Use <code>ttest</code> when you want to know whether a sample comes from population that has a different mean than a specified value.
Should I apply the <code>ttest</code> to the sample mean?	No, this a point of confusion. The <code>ttest</code> function requires the entire sample not just the sample mean.
What is the <code>p1</code> ?	The <code>p1</code> is called a p-value. It is the probability that your sample's test statistic would have been observed if the populations were actually the same. A smaller p-value lends more support to the population means being different.

EXAMPLE 3: Is it possible that the students in section 2 sleep 8 hours on average?

```
sleepHoursSec2 = sleepHours(:, section == 2);

[h2, p2, c2] = ttest(sleepHoursSec2(:), 8);

fprintf(['Do section 2 students sleep 8 hours on average?\n\t' ...

        'h = %g, p = %g, ci = [%g, %g]\n'], h2, p2, c2);
```

Do section 2 students sleep 8 hours on average?

```
h = 1, p = 1.91292e-06, ci = [8.30087, 8.71824]
```

Questions	Answers
What is the null hypothesis for this <code>ttest</code> ?	The null hypothesis is that students in section 2 sleep 8 hours on average.
What is the alternative hypothesis for this <code>ttest</code> ?	The alternative hypothesis is that section 2 does not get 8 hours sleep on average.
Why bother with the strange null hypothesis terminology?	Many researchers report the results of standard statistical tests in terms of the null hypothesis, p-values, and confidence intervals. In order to understand these results, you must learn to work with the terminology.
What does a <code>ttest</code> return value of <code>h2 = 1</code> mean in terms of the null hypothesis?	A return value of 1 indicates that you should reject the null hypothesis in favor of the alternative hypothesis. In other words, you should conclude that the true mean sleep time for section 2 is not likely to be 8 hours.

Questions	Answers
How unlikely is unlikely?	The default level of significance for <code>ttest</code> is 5%, meaning that less than 5% of the samples randomly selected from a normal distribution with mean 8 would have this statistical characteristic. Roughly, you can think of this as being "95% certain" that the true mean average sleep time for section 2 is not 8 hours.
What does the p-value <code>p2</code> tell me?	The p-value gives the probability that a sample drawn at random from the true population could have produced the observed test statistic if the null hypothesis were actually true. In other words, the p-value is a measure of how probable the observed differences could be due to a bad draw of the random sample. The p-value in this example is miniscule, so it is very unlikely that such a sample could have been drawn by chance from a distribution whose mean was actually 8.
How is the <code>ttest</code> return value <code>h2</code> related to the p-value <code>p2</code> ?	The <code>ttest</code> has a cut-off value called the level of significance or <code>alpha</code> . If the p-value, <code>p2</code> , is less than <code>alpha</code> , then the <code>ttest</code> returns <code>h2 = 1</code> .
How is the value of <code>h2</code> related to the level of significance?	The level of significance is a threshold set by you. By default the <code>ttest</code> uses a 5% significance level corresponding to an <code>alpha</code> of 0.05. A certain function (test statistic) called the t-statistic can be computed for your sample. The distribution of these statistic values computed for samples drawn from a normal distribution is known to follow a specific distribution (the t-distribution with $n-1$ degrees of freedom). If fewer than 5% of the samples from the t-distribution have a test statistic more extreme than that of your sample, the <code>ttest</code> returns a value of 1. Otherwise, the <code>ttest</code> returns a value of 0. This probably seems complicated --- its a good reason to take a statistics course!
Is it necessary to state null and alternative hypotheses when applying the <code>ttest</code> ?	No, you can call the <code>ttest</code> function without explicitly stating these hypotheses.
How should I interpret <code>c2</code> ?	The <code>c2</code> holds the confidence interval for this test. 95% of the samples will have a confidence interval that contains the actual population mean. Since <code>c2</code> is <code>[8.30087, 8.71824]</code> , you can be 95% "sure" that this confidence interval holds the true average sleep time of section 2. In other words, you can be 95% sure that the true mean is in <code>[8.30087, 8.71824]</code> .

EXAMPLE 4: Do the students in sections 2 and 3 sleep a different amount?

Create a new cell in which you type and execute:

```
sleepHoursSec3 = sleepHours(:, section == 3);

[h3, p3, c3] = ttest2(sleepHoursSec2(:), sleepHoursSec3(:));

fprintf(['Do students in sections 2 and 3 get different amounts of sleep on
average?\n\t' ...

        'h = %g, p = %g, ci = [%g, %g]\n'], h3, p3, c3);
```

Do students in sections 2 and 3 get different amounts of sleep on average?

```
h = 1, p = 0.0035873, ci = [0.127686, 0.652328]
```

Questions	Answers
What is the purpose of the <code>ttest2</code> function?	The <code>ttest2</code> determines whether the means of two distinct populations are likely to be different based on a random sample from each population.
When would I use <code>ttest</code> rather than <code>ttest2</code> ?	Use the <code>ttest</code> function when you have a particular mean value in mind and want to determine whether the mean of a single population is likely to be different from that value. Use the <code>ttest2</code> when you want to determine whether the means of two populations are likely to be different.
Should I apply the <code>ttest2</code> to the sample means?	No, this a point of confusion. The <code>ttest2</code> function requires the sample values, not just their respective means.
How should I interpret a <code>ttest2</code> return value of <code>h2 = 1</code> ?	A <code>ttest2</code> return value of <code>h2 = 1</code> indicates that the true average hours of sleep for sections 2 and 3 are likely to be different.
How likely is likely?	The default level of significance for <code>ttest2</code> is 5% (alpha is 0.05), meaning that there is less than a 5% probability that you would observe the test statistic if the average amounts of sleep for the two sections were the same. Roughly speaking, you can be "95% sure" that the two sections have different average amounts of sleep.
What if the <code>ttest2</code> returned <code>h2 = 0</code> ?	If <code>ttest2</code> returns an <code>h2</code> values of 0, the test has not provided evidence that the means are different. You cannot then conclude that the means are the same. (Innocent until proven guilty again!)
Does <code>ttest2</code> make any assumptions about the distribution of the population?	Yes, <code>ttest2</code> assumes that the samples were drawn at random from normally distributed populations.
Can I get a more statistically significant answer?	Yes, the <code>ttest2</code> function has a third argument, <code>alpha</code> , that specifies the significance level. When omitted, this argument is assumed to be 0.05, meaning the 5% significance level.

EXAMPLE 5: Do section 2 and 3 students sleep differently at the 0.01 significance level?

Create a new cell in which you type and execute:

```
[h4, p4, c4] = ttest2(sleepHoursSec2(:), sleepHoursSec3(:), 0.01);  
  
fprintf(['Do students in sections 2 and 3 sleep differently at the 0.01  
significance level?\n\t' ...  
  
      'h = %g, p = %g, ci = [%g, %g]\n'], h4, p4, c4);
```

Do students in sections 2 and 3 sleep differently at the 0.01 significance level?

```
h = 1, p = 0.0035873, ci = [0.0451417, 0.734872]
```

Questions	Answers
What does the 0.01 signify?	The third argument of <code>ttest2</code> is <code>alpha</code> , which corresponds to a 1% significance level. When omitted, this argument is assumed to have a value of 0.05, meaning the 5% significance level. An <code>alpha</code> of 0.01 puts a stronger requirement on rejecting the null hypothesis. If the call to <code>ttest2</code> returned a value of <code>h3 = 1</code> at a significance level of 1% significance level, you could conclude there is less than a 1% probability that true average sleep times of sections 2 and 3 are the same. (The actual interpretation is that fewer than 1% of the samples would have the observed statistic if the true means were actually equal.)
What is the difference between the value of <code>alpha</code> and the significance level?	The value of <code>alpha</code> is always a fraction between 0 and 1. The significance level is the corresponding percentage. For example an <code>alpha</code> of 0.05 corresponds to a 5% significance level, while an <code>alpha</code> of 0.01 corresponds to a 1% significance level. We use the two interchangeably, noting that MATLAB requires a fraction rather than a percentage when specifying significance.

EXAMPLE 6: Do section 2 students sleep more than section 3 students?

```
[h5, p5, c5] = ttest2(sleepHoursSec2(:), sleepHoursSec3(:), 0.05, 'right');

fprintf(['Do students in sections 2 and 3 get more sleep on average?\n\t' ...

        'h = %g, p = %g, ci = [%g, %g]\n'], h5, p5, c5);
```

Questions	Answers
Can I use the <code>ttest2</code> to determine whether the average amount of sleep for section 2 is greater than that of section 3?	Yes, this example shows an example testing whether the first argument has an average that is greater than the second. The alternative hypothesis for this example is: "The mean of the population represented by sample <code>sleepHoursSec2(:)</code> is greater than the mean of the population represented by the sample <code>sleepHoursSec3(:)</code> ".

Do students in sections 2 and 3 get more sleep on average?

```
h = 1, p = 0.00179365, ci = [0.169891, Inf]
```

EXAMPLE 7: Do section 2 students sleep more than section 3 students (fewer assumptions)?

Create a new cell in which you type and execute:

```
[h6, p6, c6] = ttest2(sleepHoursSec2(:), sleepHoursSec3(:), 0.05, 'right',
'unequal');

fprintf(['Do students in sections 2 and 3 get more sleep on average?\n\t' ...
```

```
'h = %g, p = %g, ci = [%g, %g]\n', h6, p6, c6);
```

Do students in sections 2 and 3 get more sleep on average?

```
h = 1, p = 0.00198515, ci = [0.167467, Inf]
```

Questions	Answers
What does the 'unequal' argument mean?	The plain vanilla t-test assumes that the two populations from which the samples were drawn (e.g., sleep times from sections 2 and 3) have the same variances. The 'unequal' parameter specifies that this assumption may not apply and to use a variation of the test that doesn't assume equal variances.

This lesson was written by Kay A. Robbins of the University of Texas at San Antonio and last modified on 18-Oct-2012. Please contact krobbins@cs.utsa.edu with comments or suggestions. The photo is of Sir Ronald Fisher, founder of modern statistics and namesake of the Fisher Iris dataset. (See http://en.wikipedia.org/wiki/File:R._A._Fischer.jpg.)