# Lab 16

# T-testing Body Temperature Data

In this lab we will look at basic t-test hypothesis tests as well as various useful conditional statements in *Stata*.

## About the Data

This article takes data from a paper in the *Journal of the American Medical Association* that examined whether the true mean body temperature is 98.6 degrees Fahrenheit. These data are derived from a data set presented in P.A. Mackowiak, S.S. Wasserman, and M.M. Levine, "A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," *Journal of the American Medical Association*, 268:1578-1580, 1992. Data were constructed to match as closely as possible the histograms and summary statistics presented in that article. There are 130 observations in the data set, with gender, body temp, and resting heart rate recorded for each subject in the study.

# Today's Lab

This week's lab will give you the opportunity to determine what the average body temperature is for yourself. Begin by loading the data into *Stata*.

```
. use http://www.stat.ucla.edu/labs/datasets/bodytemp.dta
```

In your variables window you should have a list of 3 variables. To get a better idea of what information is contained in the data set and what all of these variables mean, we use the describe command, this should always be your first step after loading data in to *Stata*.

```
. describe
```

Make sure that your results window reports 130 observations with 3 variables. Read over the labels for each of the variables.

*Question 1: What does a value of 1 for the "gender" variable represent?*

To get an idea as to how body temperature is distributed, we can create a histogram.

```
. histogram temp, norm
```

The norm option added to the end of the command overlays a normal curve for comparison. This is particularly useful when working with t-tests since one of the assumptions of a t-test is the sample distribution is approximately normally.

*Question 2: Is it reasonable to say that body temps are approximately normally distributed?*

Plot side-by-side histograms for males and females.

```
. histogram temp, norm by(gender)
```

*Question 3: Do female and male body temps seem to be distributed similarly?*

To insert a reference line on a histogram, we add the **xline(#)** option, where # is the location we want the vertical line to be placed at on the $x$-axis. Try the first graph again with a reference line at 98.6 degrees.

```
. histogram temp, norm xline(98.6)
```

*Question 4: What is interesting about this graph? Where is our 98.6 line located in the histogram in relation to the center of the overlaid normal curve?*

Use the summarize command to find the average body temp of the subjects in this study.

```
. summarize temp
```

*Question 5: What is the average body temp in this sample of 130 people? Is this what you expected? Why or why not?*

As you've learned in lecture, we can use a t-test to determine if the average body temp is, in fact, significantly lower than 98.6 degrees.

In *Stata*, a t-test is run using the **ttest** command. There are several options within this command. To get a good idea of what situations the ttest command can be used in, see the help menu.

```
. help ttest
```

Since we saw in our histograms that 98.6 was slightly higher than the center we might want to check if the average body temperature is actually less than 98.6 degrees. We can by test this using the command:

```
. ttest temp=98.6
```

*Question 6: Is this an example of a one or two tailed test? Is this a one or two sample test? What are our null and alternative hypotheses? What is the test statistic generated by Stata for our test? What is the corresponding p-value? State the conclusion of the test.*

When looking at the histograms of body temperature for men and women separately, we observed some slight differences. Let's continue this examination and try to determine if the differences we saw were due to chance or are they statistically significant.

First, re-plot the side-by-side histograms and look at the summary statistics.

```
. histogram temp, by(gender) norm
. sort gender
. by gender:  summarize temp
```

*Question 7: Do men and women seem to have approximately the same body temperature? Are they similarly distributed?*

Just as there was a formal test to check and see if the true mean body temp could be equal to 98.6, there is a t-test to check if men and women have the same mean body temp. Again, we use the **ttest** command, but the format is slightly altered.

*Question 8: What assumptions are we making by running a t-test? Are those assumptions reasonable with this data? Explain using the results from earlier commands.*

```
. ttest temp, by(gender)
```

*Question 9: State the null and alternative hypotheses of interest. Is this a one-tailed or two-tailed test? Is this a one sample or two sample test? What test statistic is generated by Stata for our test? What is the p-value? What can we conclude about how the mean body temperature for men compares to the mean body temperature for women?*

Notice that in this test the assumption of equal variances is reasonable since the larger variance is not more than twice the smaller one. If this was not

the case, then the *Stata* command would change slightly. The format of the **ttest** command if the variances are unequal is:

```
. ttest var1, by(var2) unequal
```

where "var1" is the variable whose mean we are testing, and "var2" is the variable defining two groups. In this particular case, we could use the command:

```
. ttest temp, by(gender) unequal
```

You might notice though, that the resulting output is almost identical to our original t-test output. (In fact, the only difference is the degrees of freedom.) This is because the variances are so close, we can consider them equal for testing purposes.

## Assignment

Using methods similar to those already discussed in this lab, explore resting heart rates. What is the average resting heart rate among women? Among men? In the sample as a whole? Do men and women have significantly different heart rates?

## Extra

Recalling what you've learned in correlation and regression, does there seem to be a relationship between a person's resting heart rate and their body temperature? Explain briefly, but thoroughly.