Revised July 2018

Testing for a Difference between Two Group Means

This set of notes shows how to use Stata to examine differences between two group means of a quantitative variable. It assumes that you have set Stata up on your computer (see the "Getting Started with Stata" handout), and that you have read in the set of data that you want to analyze (see the "Reading in Stata Format (.dta) Data Files" handout).

In Stata, most tasks can be performed either by issuing commands within the "Stata command" window, **or** by using the menus. These notes illustrate both approaches, using the data file "GSS2016.DTA" (this data file is posted here: https://canvas.harvard.edu/courses/53958).

To test the hypothesis that two groups have the same population mean on a quantitative variable using the "Stata Command" window, issue the following command:

```
ttest <varname1>, by(varname2)
```

where you fill in the variable name of the quantitative variable in place of "varname1" and designate your two-group (dichotomous) variable by "varname2".

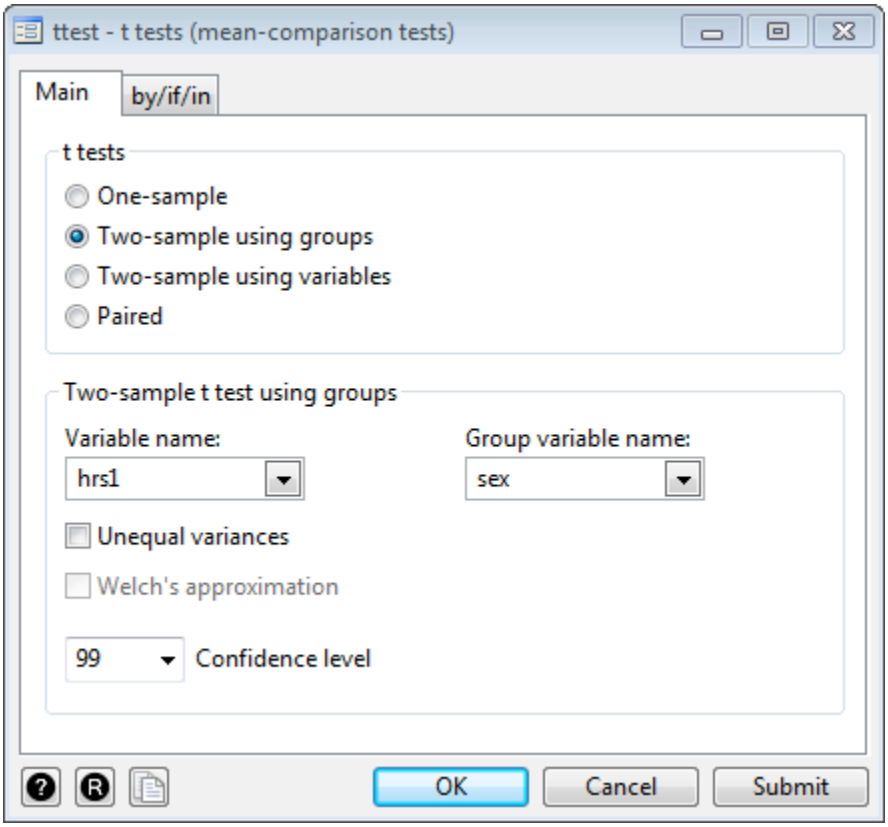
For example, the following command tests the hypothesis that the mean number of hours worked ("hrs1") is the same for men and women:

```
ttest hrs1,by(sex)
```

Using the Stata menus, you can conduct a t-test for a difference of two means as follows:

click on "Statistics"
click on "Summaries, tables, & tests"
click on "Classical tests of hypotheses"
click on "t-test (mean-comparison test)"

A window like the one shown on the next page will open up:



Fill in the name of your quantitative variable in the "Variable name:" box, and the name of your dichotomous group variable in the "Group variable name" box. If you want, change the "Confidence level" (for the "bonus" confidence interval–see below) from its default level of 95. Leave the "Unequal variances" box un-checked.

Then click "Submit" or "OK."

Either way, the following output appears in the "Stata Results" window:

```
. ttest hrs1, by(sex) level(99)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [99% Conf. Interval]
---------+--------------------------------------------------------------------
    male |     792    44.16035    .5228357     14.7139    42.81036    45.51035
  female |     854    37.90398    .4598889    13.43946    36.71673    39.09123
---------+--------------------------------------------------------------------
combined |   1,646    40.91434    .3550878    14.40624    39.99863    41.83005
---------+--------------------------------------------------------------------
    diff |            6.256372    .6939482                4.466803    8.045942
------------------------------------------------------------------------------
    diff = mean(male) - mean(female)                              t =   9.0156
Ho: diff = 0                                     degrees of freedom =     1644

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

This shows that the mean hours worked per week overall is 40.9. Among men it is 44.2, while among women it is 37.9; the difference in means between men and women is estimated as 6.26 hours per week. The estimated standard error of the difference (assuming that the variances of hours worked [in the population] are the same for men and women) is 0.693.

The usual null hypothesis is that the means are the same for men and women. A t-statistic for assessing that hypothesis is (6.26-0)/0.693=9.02. The p-value for the two-sided alternative (middle panel beneath the table) is very small, beneath 0.0001. Since it would be so very unusual to draw a sample in which the t-statistic was above 9 if the null hypothesis were true, we would *reject* the hypothesis of equal means in favor of the two-tailed alternative (middle panel) that men and women work different numbers of hours on average. Because the p value is very small, we are very confident in rejecting the assertion that the group means are equal, at any plausible significance level that we might select.

As a bonus, the output for the t-test includes a confidence interval for the *difference in group means* (bottom line). While we estimate that difference is 6.25 from the sample means themselves, the confidence interval indicates that we are 99% confident that the interval (4.47, 8.05) includes difference of group means in the population.