Revised July 2018

Hypothesis Testing About a Mean or Proportion

This set of notes shows how to use Stata to conduct a hypothesis test about the population mean of a quantitative variable or the population proportion for a dichotomous 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).

A hypothesis test asks if it is plausible that a parameter (here the mean or proportion in the "population") could be equal to a particular value (as specified by the "null hypothesis"). The "p level" associated with the test corresponds to the chance of making a so-called "false rejection" or "Type I" error in inference: concluding that the null hypothesis is not true, when in fact it is. *The smaller the p level, the stronger the evidence against the null hypothesis*.

To test a hypothesis about the population mean using the "Stata Command" window, issue the following command:

```
ttest <varname>==##.#
```

where you fill in the variable name of interest to you in place of "varname" and designate the hypothesized value of the population mean under the null hypothesis in place of "##.#". Note that there are <u>two</u> equal signs between "varname" and "##.#"; this is <u>not</u> a typo.

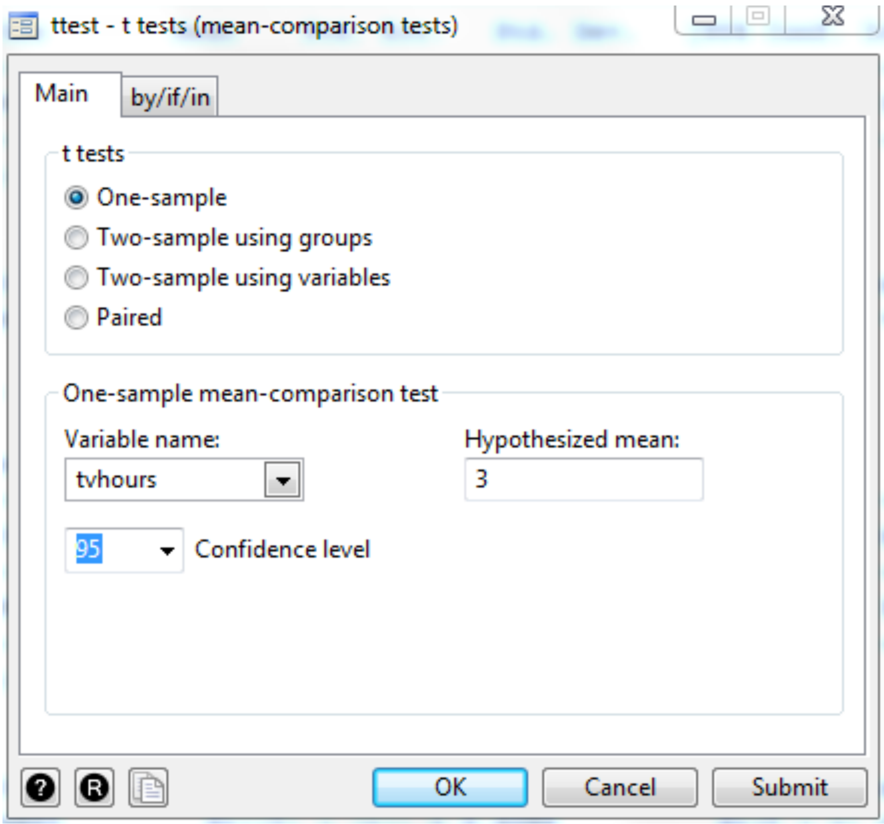
For example, the following command tests the hypothesis that the mean number of hours of television watched per day by an American adult equals 3:

```
ttest tvhours==3
```

Using the Stata menus, you can request a t-test about a population mean as follows:

click on "Statistics"
click on "Summaries, tables, & tests"
click on "Classical tests of hypotheses"
click on "One-sample mean comparison test"

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



Fill in the variable name of interest to you in the “Variable name:” box, and the hypothesized value of the population mean under your null hypothesis in the “Hypothesized mean:” box. If you want, change the Confidence level (for the “bonus” confidence interval–see below) from its default level of 95.

Then click “OK.”

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

```
One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 tvhours |   1,883    3.031333    .0647707    2.810633    2.904303    3.158363
------------------------------------------------------------------------------
    mean = mean(tvhours)                                          t =   0.4838
Ho: mean = 3                                     degrees of freedom =     1882

    Ha: mean < 3                 Ha: mean != 3                 Ha: mean > 3
 Pr(T < t) = 0.6857         Pr(|T| > |t|) = 0.6286          Pr(T > t) = 0.3143
```

The t-statistic for the test is 0.48. The middle panel at the bottom assumes the typical “2-tailed alternative”–prior to conducting the test, we specify that the alternative hypothesis to a mean of 3 is a mean that is *not* 3 (Stata symbolizes “not” by “!=” in the above). The large p-level (about

0.629) tells us that it is quite plausible that we would have gotten a sample with a statistic as large or larger in magnitude (in absolute value—that is, larger than 0.48 or smaller than -0.48) as this, if the mean hours of TV watched in the population really were 3. This is a fairly likely result, then, that provides very weak evidence against the null hypothesis. Hence, we do not have evidence to contest the claim that the mean is 3, though we cannot assert that it actually is 3 on the basis of this test.

The right panel has more of a point of view–its alternative hypothesis asserts that the mean is *greater* than 3. With this "1-tailed alternative", the p level tells us that the chance of observing a sample with a t-statistic this large or larger is also fairly likely (if the mean is really 3), about 0.314. Again, this offers little evidence against the claim that the mean level of television viewing among adults is 3 hours per day.

The p value for the left panel–where the 1-tailed alternative is now that the population mean is *less* than 3–is also large (0.686), and also seems to offer next to no evidence against the null hypothesis that mean hours of TV viewing among adults is 3 (here, if the alternative to a mean of 3 hours is a mean that is less than that). None of these high p levels indicate that the mean television viewing level in the population really *is* 3 hours, however. They just don't indicate that it *isn't* 3 hours.

Indeed, "not rejecting" the null hypothesis–as we would do for all three alternative hypotheses in this example–never amounts to "accepting" the null hypothesis. It only means that we lack evidence against it. That may seem like splitting semantic hairs, but it's an important point to keep in mind.

Note, as a bonus, that the output for the t-test includes a 95% confidence interval for the mean, in this case from 2.90 to 3.16 hours.

Closely related, but not identical, is the drill for testing a hypothesis about a population proportion.

To do this, you must work with a dichotomous variable that is coded 1 if someone has the outcome of interest, and 0 otherwise. You may have to recode your variables in order to do this (see separate handout on how to recode). For example, to code a categorical variable that measures if the number of immigrants in America should change (variable "letin1," 1=increased a lot…5=reduced a lot) as a proportion, the recoding and variable labeling commands might be

```
recode letin1 (1/3=0 "Same or more")(4/5=1 "Fewer"),gen(lessimm)

label variable lessimm "Favors reduced immigration"
```

To test a hypothesis about a population proportion using the "Stata Command" window, issue the following command:

```
prtest <varname>==#.##
```

where you fill in the variable name of interest to you in place of "varname" and designate the hypothesized value of the population proportion under the null hypothesis in place of "##.#". Note that there are two equal signs between "varname" and "##.#"; again, this is not a typo.

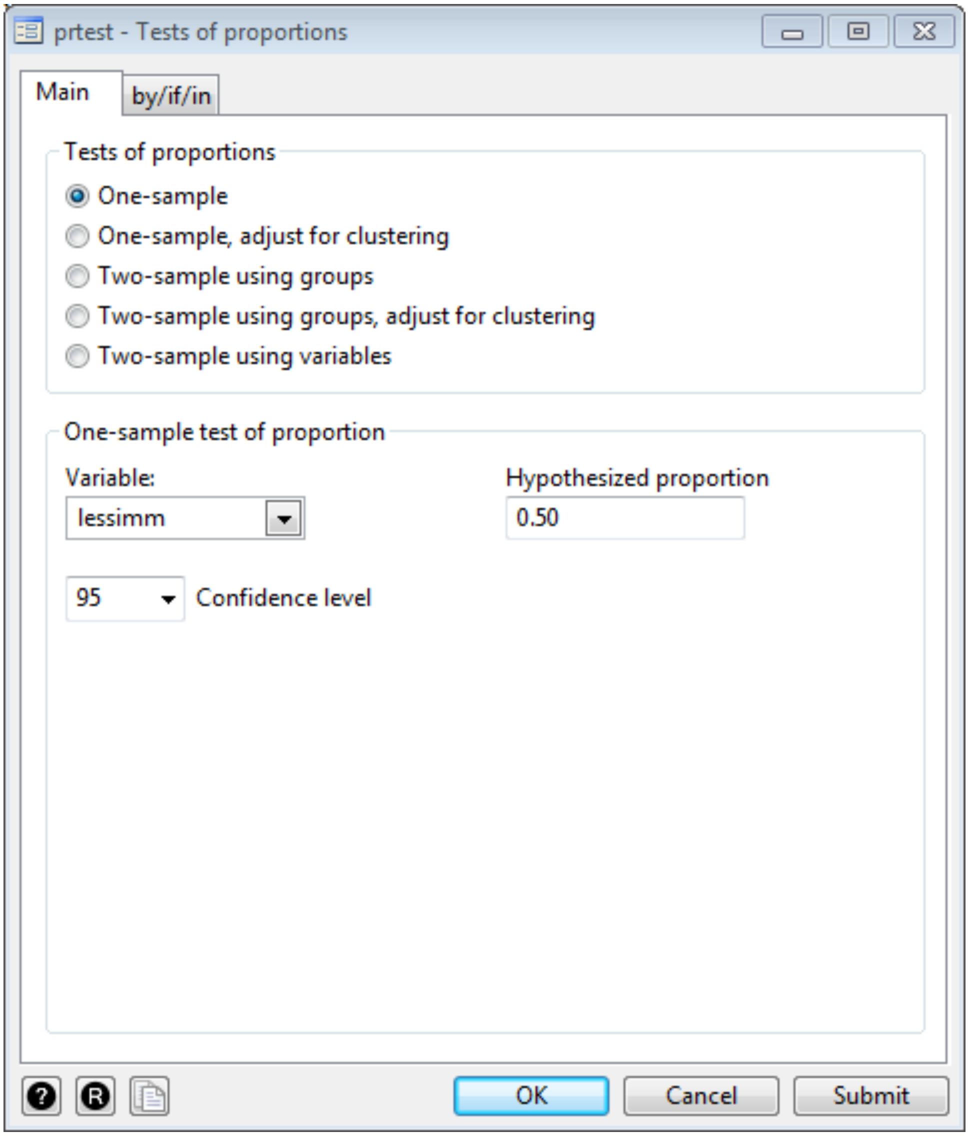
For example, the following command tests the hypothesis that the proportion of adult Americans who favor reduced immigration is 0.5, i.e. that the population is evenly divided between those who want lower immigration and those who don't:

```
prtest lessimm==0.5
```

To accomplish this using the Stata menus, you can request a z-test about a population proportion as follows:

click on "Statistics"
click on "Summaries, tables, & tests"
click on "Classical tests of hypotheses"
click on "Proportion test"

A window like this will open up:



Fill in the variable name of interest to you in the "Variable name:" box, and the hypothetical value of the population proportion under your null hypothesis in the "Hypothesized proportion" box. If you want, change the Confidence level (for the "bonus" confidence interval–see below) from its default level of 95.

Then click "OK."

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

```
One-sample test of proportion                Number of obs     =     1845

------------------------------------------------------------------------------
   Variable |      Mean   Std. Err.                [95% Conf. Interval]
-------------+----------------------------------------------------------------
    lessimm |   .4211382   .0114948                 .3986088    .4436676
------------------------------------------------------------------------------
   p = proportion(lessimm)                                    z =  -6.7748
Ho: p = 0.5

    Ha: p < 0.5              Ha: p != 0.5              Ha: p > 0.5
 Pr(Z < z) = 0.0000       Pr(|Z| > |z|) = 0.0000       Pr(Z > z) = 1.0000
```

The z-statistic for the test is -6.77. The middle panel assumes the typical "2-tailed alternative"– prior to conducting the test, we specify that the alternative to a population proportion of 0.5 is a proportion that is *not* 0.5 (again, Stata symbolizes "not" by "!=" in the above). The small p-level (<0.0001) tells us that it is not at all plausible that we would have gotten a statistic as large or larger in magnitude (in absolute value—that is, larger than 6.77 <u>or</u> smaller than -0.6.77) as this, if the proportion favoring lower immigration in the population really was really 0.5. This is strong evidence against the null hypothesis of an equally divided population. Hence, we have evidence to contest the claim that the population proportion is 0.5. There is a very small chance (<0.0001) that this is an erroneous decision (i.e. that the population really is evenly divided and that our sample is an extremely unusual one). But that possibility is indeed *very* small in this instance.

The left and right panels have more of a point of view–their alternative hypotheses assert that the proportion is *less* than 0.5 (left) or *greater* than 0.5 (right), if they are not evenly divided. With the "1-tailed alternative" at the left, the low p level again tells us that the chance of observing a z-statistic this small (if the population proportion really is 0.5) is very small, less than 0.0001. This offers strong evidence against the claim that population is evenly divided as regards lowering immigration levels, and in favor of the claim that a majority of the population does <u>not</u> favor reducing immigration.

The p value for the right panel–where the 1-tailed alternative is now that the population proportion is *greater* than 0.5–is instead quite large (1.0). This indicates that we would not be at all surprised to get a sample in which the proportion favoring lower immigration is 0.421 (the sample proportion) or bigger, if the alternative to an equally divided population is that a *majority* of Americans favor lower immigration. This high p value indicates that we would prefer the claim of an equal division (the null hypothesis) to a claim that a majority favors a reduction in immigration on the basis of this evidence.

The 95% confidence interval for the population proportion here is (0.399, 0.444).