Immediate commands for confidence intervals and tests of significance

As an alternative to the commands `ttest` and `ci`, we can use the immediate commands for confidence intervals and tests of significance. An immediate command is a command that obtains data not from the data stored in memory but from numbers typed as arguments. Immediate commands, in effect, turn STATA into a glorified hand-calculator.

There are instances where you may not have the data, but you know something about the data and what you do know is adequate to perform the statistical test.

Confidence Intervals

For example suppose we want a 90% confidence interval for $\mu$ and we do not have access to the data but we know that

$$ n = 100 \quad \bar{y} = 50 \quad and \quad s = 8 $$

To construct the confidence interval use the command:

```
. cii 100 50 8, level(90)
```

This gives the following output:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[90% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>50</td>
<td>.8</td>
<td>48.67169 51.32831</td>
</tr>
</tbody>
</table>

A 90% confidence interval for $\mu$ is (48.67, 51.33)
Tests of proportions

To test $H_0 : p = p_0$ using a one-sample test of proportion, use the command:

```
prtesti N X p0, count
```

where $N$ is the sample size, $X$ is the number of “successes” in the observed sample and $p_0$ is the hypothesized population proportion.

**Ex.** A factory claims that less than 10% of the components they produce are defective. A consumer group is skeptical of the claim and checks a SRS of 300 components and finds that 39 are defective. Is there significant evidence to reject the companies claim?

In this example $N=300$, $X=39$ and $p_0=0.10$. To solve this problem, using STATA we can use the command

```
prtesti 300 39 .10, count
```

This gives the following output:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>0.13</td>
<td>0.0194165</td>
<td>0.0919444 0.1680556</td>
</tr>
</tbody>
</table>

Ho: proportion(x) = .1

<table>
<thead>
<tr>
<th>Ho: x &lt; .1</th>
<th>Ha: x != .1</th>
<th>Ha: x &gt; .1</th>
</tr>
</thead>
<tbody>
<tr>
<td>z = 1.732</td>
<td>z = 1.732</td>
<td>z = 1.732</td>
</tr>
<tr>
<td>P &lt; z = 0.9584</td>
<td>P &gt;</td>
<td>z</td>
</tr>
</tbody>
</table>

Since the alternative hypothesis is $H_a$: $p>0.10$, the p-value of the test is 0.0416.
To test $H_0: p_1 = p_2$ using a two-sample test of proportion, use the command:

```
prtesti N1 X1 N2 X2, count
```

where N1 and N2 are the sample sizes and X1 and X2 are the number of “successes” in each of the two samples.

**Ex.** In two separate polls conducted 6 months apart, constituents were asked if they approved of the job their congressman was doing. The results of the polls are listed in the table below:

<table>
<thead>
<tr>
<th>Population</th>
<th>N</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second Survey</td>
<td>1100</td>
<td>700</td>
</tr>
<tr>
<td>First Survey</td>
<td>900</td>
<td>500</td>
</tr>
</tbody>
</table>

Is there significant evidence that his approval rating has increased?

In this example N1=1100, X1=700, N2=900 and X2=500. To test whether there is a significant difference in approval rating between the polls use the command

```
prtesti 1100 700 900 500, count
```

This command gives the following output:

```
Two-sample test of proportion                       x: Number of obs =     1100
                                                 y: Number of obs =      900
-----------------------------------------------------------------------------
Variable |       Mean   Std. Err.       z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 x |   .6363636   .0145041                      .6079362    .6647911
 y |   .5555556   .0165635                      .5230918    .5880194
-------------+----------------------------------------------------------------
  diff |   .0808081   .0220163                       .037657    .1239592
 |  under Ho:   .0220193     3.67   0.000
-----------------------------------------------------------------------------
Ho: proportion(x) - proportion(y) = diff = 0
Ha: diff < 0         Ha: diff ! = 0         Ha: diff > 0
 z =  3.670          z =  3.670          z =  3.670
P < z =  0.9999       P > |z| = 0.0002       P > z = 0.0001
```

From the output we see that the p-value is 0.0001.
Tests of means

To test $H_0 : \mu = \mu_0$ using a one-sample t-test, use the command:

\[ \text{ttesti n ybar s mu0} \]

where \( n \) is the sample size, \( ybar \) is the sample mean, \( s \) is the sample standard deviation and \( \mu_0 \) is the hypothesized sample mean \( \mu_0 \).

Ex. Estimate the mean height of all Columbia students. The population of students has mean \( \mu \) and standard deviation \( \sigma \), both unknown.

We take a sample of 12 students and obtain \( \bar{y} = 66.30 \) and \( s = 4.35 \).

\[ H_0 : \mu = 68 \quad H_a : \mu \neq 68 \]

Can we reject \( H_0 \) at the 5% significance level?

To solve this problem, using STATA we can use the command

\[ \text{ttesti 12 66.3 4.35 68} \]

This gives the following output:

|     Obs   |     Mean   |   Std. Err.   |     Std. Dev.   | [95% Conf. Interval] |
|---------+-----------+-------------+-----------------+---------------------|
|     12  |    66.3    |    1.255737  |       4.35      | 63.53614  69.06386  |

Degrees of freedom: 11

Ho: mean(x) = 68

Ha: mean < 68 \quad Ha: mean \neq 68 \quad Ha: mean > 68

\[ t = -1.3538 \quad t = -1.3538 \quad t = -1.3538 \]

P < t = 0.1015 \quad P > |t| = 0.2030 \quad P > t = 0.8985

Since the alternative hypothesis is two-sided, the p-value is 0.2030.
To test $H_0 : \mu_1 = \mu_2$ using a two-sample t-test we can use the command:

\[ \texttt{ttesti n1 ybar1 s1 n2 ybar2 s2, unequal} \]

where n1 and n2 are the sample sizes, ybar1 and ybar2 are the sample means and s1 and s2 are the sample standard deviations of each of the two samples.

**Ex.** Testing the effect of a new medication on pulse rate - 60 subjects are randomly divided into two groups of 30. One group is given the new medicine and the other a placebo.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample size</th>
<th>Sample mean</th>
<th>Sample standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Medicine</td>
<td>30</td>
<td>65.2</td>
<td>7.8</td>
</tr>
<tr>
<td>2 – Placebo</td>
<td>30</td>
<td>70.3</td>
<td>8.4</td>
</tr>
</tbody>
</table>

Does the medicine reduce pulse rate?

$H_0 : \mu_1 - \mu_2 = 0$ and $H_a : \mu_1 - \mu_2 < 0$

To test this claim, use the command:

\[ \texttt{ttesti 30 65.2 7.8 30 70.3 8.4 , unequal} \]

which gives the following output:

According to the output, the p-value is 0.0090.
Exercise 1:
Do problem 22.22 from the textbook. The equivalent problem in Edition 1 is 22.20.

Solve the problem using STATA and the `prtesti` command. Make sure to hand in your log file and answers to any questions in the text.

Exercise 2:

Solve the problem using STATA and the `ttesti` command. Make sure to hand in your log file and answers to any questions in the text.