Note 7.1 – Finding percentiles of a t-distribution

The Stata command `invttail(n, p)` returns percentiles from the t-distribution with n degrees of freedom where p is the area under the right-tail of the t-distribution’s probability density function. For example, the 95th percentile of a t-distribution with 59 degrees of freedom can be found with the following command:

Stata command:
```
invttail(59, 0.05)
```

Stata output:
```
1.6711
```

Note 7.2 – Binomial confidence intervals

The `ci` command in Stata returns confidence intervals for means, proportions, and counts. The `level( )` option allows you to set the confidence level for the confidence interval. To find a 90 percent confidence interval for twenty observations with four successes, we can use the Stata command shown below which provides (0.0714, 0.4010).

Stata command:
```
cii 20 4, level(90)
```

Stata output:
```
Variable |        Obs        Mean    Std. Err.       [90% Conf. Interval]
-----------------+-----------------------------------------------
         |         20        .2000        .0894427        .0713539    .4010281
```

Note 7.3 – Confidence intervals using the t-distribution

Although hypothesis testing is not introduced until Chapter 8, the `ttest` command in Stata calculates confidence intervals by default. In Example 7.7, we are comparing the mean ages between AML and ALL patients using the Stata command below:
Stata command:

ttest age, by(dx_type) level(99)

From the output shown below, a 99 percent confidence interval for the difference in the mean ages is (1.41, 25.02).

Stata output:

Two-sample t test with equal variances

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[99% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>51</td>
<td>49.86275</td>
<td>2.31195</td>
<td>16.51063</td>
<td>16.51063 - 43.67182</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>36.65</td>
<td>3.99096</td>
<td>17.84812</td>
<td>25.23212 - 48.06788</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>combined</td>
<td>71</td>
<td>46.14085</td>
<td>2.113033</td>
<td>17.80473</td>
<td>40.54574 - 51.73595</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>13.21275</td>
<td>4.456005</td>
<td>1.408892</td>
<td>25.0166</td>
</tr>
</tbody>
</table>

diff = mean(0) - mean(1)                                   t = 2.9652
Ho: diff = 0                                     degrees of freedom = 69
Ha: diff < 0                                Ha: diff != 0                           Ha: diff > 0
Pr(T < t) = 0.9979                              Pr(|T| > |t|) = 0.0042                          Pr(T > t) = 0.0021

Note 7.4 – Confidence intervals for the Pearson correlation coefficient

By using the display command, Stata can be used as a calculator. As an example, we can calculate the lower and upper limits for the correlation coefficient. The lower limit can be found using the Stata command:

Stata commands:

display (exp(2*(0.7882)) - 1)/(exp(2*(0.7882))+1)
* which gives 0.657, and the upper limit can be found using

display (exp(2*(2.0948)) - 1)/(exp(2*(2.0948))+1)
* which gives 0.970