Tools for Confidence Intervals and Hypothesis Tests in Excel

Fall 2001 Professor Paul Glasserman
B6014: Managerial Statistics 403 Uris Hall

1. Confidence Intervals

Although it is a relatively simple matter to compute confidence intervals using just a calculator and a normal table, it is sometimes useful to know how to automate these calculations in a spreadsheet. To illustrate, suppose we have a range of data named DataRange. The following formulas give the lower and upper 95% confidence limits for the mean based on this data:

```
AVERAGE(DatRange)-NORMSINV(0.975)*STDEV(DataRange)/SQRT(COUNT(DataRange))

AVERAGE(DataRange)+NORMSINV(0.975)*STDEV(DataRange)/SQRT(COUNT(DataRange))
```

You probably recognize `AVERAGE` as the mean, `STDEV` as the sample standard deviation, and `COUNT(DataRange)` as the sample size. The function `NORMSINV` is the inverse of the standard normal cumulative distribution; using `NORMSINV` is equivalent to looking up a probability value in the body of the normal table and finding the corresponding z value. In our usual notation, NORMSINV$(1-\alpha) = z_\alpha$, where $P(Z > z_\alpha) = \alpha$; thus, `NORMSINV(0.975)` evaluates to 1.96. A slightly more general implementation of this formula specifies a cell address for the α value and replaces `NORMSINV(0.975)` with NORMSINV$(1-\alpha/2)$. This allows the user to specify the α level in a separate cell.

Had we wanted to compute t-based confidence intervals instead of z-based intervals, we could have used `TINV` instead of `NORMSINV`. This command requires that we specify the degrees of freedom: `TINV(0.975,COUNT(DataRange)-1)`. (Recall that the df is one less than the sample size.)

2. Hypothesis Test on a Mean

Here is a simple way to run a large sample (z-based) hypothesis test on a mean. Suppose we want to test

$$H_0 : \mu \leq \mu_0$$
$$H_1 : \mu > \mu_0$$

based on data in DataRange. The Excel function ZTEST(DataRange,μ_0) gives the p-value for this test based on the data. (In other words, it automatically calculates the Z-statistic and finds how much area there is to the right of Z under the normal curve.) Of course, in practice you would use a specific number rather than μ_0.

If you somehow knew the population standard deviation σ, you could use it by specifying ZTEST(DataRange,μ_0,σ). The optional third argument of ZTEST is the standard deviation. If it is omitted, Excel estimates it from the data.

3. Hypothesis Tests on the Difference of Two Means

Excel contains tools to automate hypothesis tests. To illustrate, suppose we wanted to test the difference in mean returns between Experts and Darts. Pull down the menu `Tools...Data Analysis`. and then select `t-Test: Paired Two Sample for Means`. This will open a dialog box. Specify the following information:

- `Variable 1 Range:` The range of cells containing Expert returns.
- `Variable 2 Rage:` The range of cells containing Dart (or Index) returns.
- `Output Range:` Anywhere convenient.
- You can leave `Hypothesized Mean Difference` blank; the default is 0, which indicates we are testing for the presence of any difference at all.

If you have labels for your data in the cells just above the data itself, then check the `Labels` box; this tells Excel that the first row of data contains labels. This will produce output that is easier to read. Now click on `OK` to execute the hypothesis test. The output contains a lot of information. The most important items are the t-statistic and the p-values. Both 1-sided and 2-sided p-values are reported; it's up to the user to decide which number is relevant.