Coverage Needed for a Given Relative Standard Error (RSE)

This example application is for a weekly sample survey of denatured ethanol production, using a monthly census survey as regressor data. This is at the national level.

Use equation 4 from page 15 of "Projected Variance for the Model-Based Classical Ratio Estimator" at http://interstat.statjournals.net/YEAR/2012/abstracts/1209001.php?Name=209001:

$$a^* = \left\{1 + \left[\left(\frac{RSE(T^*)}{\sigma^*_{e_0}}\right)^2 \Big/ T_x\right]\right\}^{-1}$$

Below are results from a spreadsheet where only $\sigma^*_{e_0}$ needs to be correct if we have the other variables in the equation above, to be able to check on the coverage needed for given estimates of RSE.

Here we use

$$y_i = bx_i + e_{0_i}x_i^{\gamma}\ ; \quad e_i = e_{0_i}x_i^{\gamma}; \quad \gamma = ½\ ; \quad e_{0_i} = (y_i - bx_i)x_i^{-½}\ ;$$

and

$$\sigma^{*2}_{e_0} = \sum_{i=1}^{n} e^2_{0_i}/(n-1)$$

Then, from Excel, this was found:

For an experiment using monthly regressor data three-months previous to the weekly data, with the full cutoff sample that was collected, n=136 out of 180 plus two add-ons, $\sigma^*_{e_0}$ was approximately 0.766. For a reduced cutoff sample, n=67, $\sigma^*_{e_0}$ was approximately 0.919.

For the value of $\sigma^*_{e_0}$, which may be in the thousands, or thousandths, or anything, depending on the units of the data, the relative value is important. Here 1.0 would be relatively large, and may provide a good 'upper limit' on the coverage needed for a given estimated relative standard error (RSE).

Using $\sigma^*_{e_0} = 1.0$, RSE*= 0.025 = 2.5%, T*= 5600, and $T_x = 25962$, we can get an estimate of monthly regressor data coverage (the part for which we need a sample) that will likely provide an RSE estimate of 2.5% or less:

$$a^* = 1/(1 + ((0.025)(5600)/1.0)^2/25962) = 57\%$$

The smallest sample to obtain this coverage would be a cutoff sample. Here n = 67 plus 2 add-ons yields a coverage of 60% with an estimated RSE of 2.1%.