

Calculating Sample Autocorrelations in Excel

A sample autocorrelation is defined as $\hat{r}_k \equiv \frac{\hat{g}_k}{\hat{g}_0} = \frac{\widehat{\text{cov}}(R_{it}, R_{i,t-k})}{\widehat{\text{var}}(R_{it})}$.

In Excel, the tricky part in calculating sample autocorrelations is calculating the sample covariance term.

Suppose you have data as follows:

	A	B
1	R_it	R_jt
2	0.03	0.02
3	0.02	0.05
4	0.05	-0.01
5	-0.01	-0.01
6	0.01	0.02
7	0.03	0.5
8	0.05	-0.2
9	0.04	0.2
10	-0.02	0.1

Calculating the variance for series R_it is straightforward:

=var(A2:A10)

Calculating the k-th lag covariance for series R_it is similar. Consider the k=1 case.

=covar(A3:A10,A2:A9)

Now, consider the k=2 case

=covar(A4:A10,A2:A8)

The pattern should be clear.

A sample cross correlation is defined as $\hat{r}_{ij,k} \equiv \frac{\hat{g}_{ij,k}}{\sqrt{\hat{g}_{i,0}\hat{g}_{j,0}}} = \frac{\widehat{\text{cov}}(R_{it}, R_{j,t-k})}{\sqrt{\widehat{\text{var}}(R_{it}) \cdot \widehat{\text{var}}(R_{jt})}}$.

Calculating the sample variances is straightforward. Calculating the sample covariances is done as follows.

For k=1,

=covar(A3:A10,B2:B9)

For k=2,

=covar(A4:A10,B2:B8)

The pattern should again be clear.

Note, however, that $\hat{\mathbf{r}}_{ij,k} \neq \hat{\mathbf{r}}_{ji,k}$. For $\hat{\mathbf{g}}_{ji,k}$, the sample covariances are calculated as follows.

For k=1,

=covar(B3:B10,A2:A9)

For k=2,

=covar(B4:B10,A2:A8)

Useful Q+A:

Q. I'm confused on question #1. I understand that I need to do a chow test on the excess return model but do I set up two separate regressions (one before and one after the break point) or do I use delta? Also if I use delta, how do I get excel to compute a regression with two variables? Additionally, which data do I use on excel to calculate the RSS (sum of residuals squared)?

A. You can either set up two separate regressions or use a dummy variable. The F-statistic will be exactly the same in both cases. If you use the dummy variable approach, you will have three right hand side variables (r_mt-rf, D_t, and D_t*(r_mt-rf)). To compute a regression with multiple right hand side variables in Excel, simply highlight all of the data in the three columns corresponding to r_mt-rf, D_t, and D_t*(r_mt-rf), rather than one column (r_mt-rf) as in the previous homework, when filling in cell references for the "X" variable. To calculate the RSS, you can get Excel to calculate the residuals. It will put the residual series below the regression estimates. Take the squares of the residuals and sum across time. You can also calculate the residuals manually as follows:

$$e_hat = r_it - rf - alpha_hat - beta_hat*(r_mt-rf) - delta0_hat*D_t - delta1_hat*D_t*(r_mt-rf)$$

Q. For #1, to calculate the F statistic, can I just use the value that Excel generates in the regression output (I used the dummy variable method)? Also when you discussed the

RSS on the discussion page of the website (about taking the squares of the residuals and summing them), that is the restricted RSS, right? Then how do I find the unrestricted, is it given in the Excel regression output?

A. The F-statistic reported by Excel is for a different hypothesis. The null in that case is that all of the parameters for the X variables are equal to zero. When testing the stability of the parameters over time, the hypothesis of interest is whether the deltas are equal to zero only. That is, consider the general regression equation using dummy variables:

$$R_{it-rf} = \alpha + \beta(R_{mt-rf}) + \delta_1 D_t + \delta_2 D_t(R_{mt-rf}) + e_{it}$$

The null hypothesis of stable parameters (alpha and beta) is $H_0: \delta_1 = \delta_2 = 0$. To compute the F-statistic for this hypothesis, use the RSS reported in Excel for the dummy variable regression. This will be your RSS_UR (unrestricted). Then, compare it to the RSS reported in Excel for the basic CAPM regression

$$R_{it-rf} = \alpha + \beta(R_{mt-rf}) + e_{it}$$

This will be your RSS_R (restricted) for calculating the F-statistic. RSS is the sum of the squared residuals. You can calculate it manually given the residuals, but Excel also reports the RSS for a given regression.

Q. I had a few questions on the homework that I wondered if you could help me with? When trying to do a White test in #2, I get an error that the "Y" needs to be a "contiguous reference." What I have done is squared the residuals on the regression statistics from #1. That is in the "Y" box and then for the "x" box I have "rmt-rf" and "rmt-rf squared." Basically it won't run the regression for me with Mobile or with Citicorp.

A. Try lining up the three series next to each other: squared residuals, rmt-rf, (rmt-rf)². Make sure you have the same number of observations for each series. Then, highlight the data in the column with the squared residuals for Y and highlight the data in the two columns with rmt-rf and (rmt-rf)² for X. This should work.