**AW9h Significance of autocorrelation coefficients and evaluation**

Now we have autocorrelations and partial autocorrelations calculated, and the confidence intervals defined, what else might be useful to know about these two data sets, in particular of autocorrelations? Well, in some textbooks and in numerous software printouts you will invariably get some further statistics associated with autocorrelations, so let’s recreate them and provide explanations.

**Example**

In Figure 1 we provided such a typical printout (excluding the confidence interval, just to make the printout less cluttered). We are using the same autocorrelation coefficients as in previous example.



Figure 1

In column D we have calculated the so-called Q-statistic. This statistic is called Ljung-Box statistic and is also explained in the textbook. In this chapter we’ll refer just to the brief method on how to calculate it. The equation for Q-statistic is:

 (1)

Where *n* is the number of observations, *j* is the number of lags and rj are the autocorrelation coefficients.

The Q-Statistic follows the chi-square distribution.

To simplify the “translation” of the equation (1) into Excel syntax, we used a few shortcuts, such as =SUMPRODUCT() function to multiply the squared values of the autocorrelations $r\_{j}^{2}$ with 1/(n-j). In addition to this, column E gives us another set of values, the so-called probability values. These values are related to the Q-Statistic. They define the cumulative probability, up to every lag, that the calculated Q-stat values are not random. We already said that the Q-stat values are distributed in accordance with the chi-squared distribution, where the lag corresponds with the degrees of freedom. To calculate every value, we just use the =CHIDIST() function. Just one brief digression, in cell E6, for example, we see the p-value of 1.47179E-10. This means that we must put 10 zeros before 147179, so in fact the number in cell E6 is 0.000000000147179. The same applies to all the numbers with the E suffix in notation. Figures 2 and 3 show all the Excel formulae used to calculate these values.



Figure 2



Figure 3

The way to understand the probability for Q-stat (column E) is to remind yourself of the hypothesis testing procedure. The H0, i.e. the null hypothesis for Ljung-Box statistic, is that all the data (in this case the autocorrelations) are random. The H1, i.e. the alternative hypothesis, is that the data are not random. We can arbitrarily decide that the level of significance is 0.05. In other words, we want to be 95% certain that we can reject the null hypothesis. The general principle that always applies is: if the calculated probability value is SMALLER than the level of significance (in this case 0.05), we REJECT the null hypothesis. In some textbooks, and software packages, we can also find another method of checking if the autocorrelations (or the partial autocorrelations) are significantly different from zero. This is achieved by using a simple t-test (column F). The t-test statistic is calculated as:

$t\_{r\_{k}}=\frac{r\_{k}}{SE\_{r\_{k}}}$ (2)

Once we calculated the t-value for every autocorrelation, we need to test its significance, i.e. decide if the autocorrelation is significant (non-zero) or insignificant (virtually zero). Conventionally we would have to look in the table for the t-values, but in Excel we can just use the =TDIST() function. Figures 1 and 2 show how to calculate the t-values and the associated probabilities for every autocorrelation coefficient (columns F and G). In columns H we included a brief formula/descriptor to show us if the specific autocorrelation value is significant or not. In summary, we just want to say that the autocorrelation coefficients together with the partial autocorrelations provide essential assistance to characterize the time series and select the correct model for forecasting. The chapters in the textbook that follow will show us the practical value of these two functions and how to use them. This chapter is focused on simple mechanics of how to produce these two functions.