**S9Wj Error and Residual inspection**

At the beginning of the textbook chapter, we mentioned that forecasting errors should be treated as the residual element. In other words, something that is moving completely randomly when observed visually. If the model we used truly fits the actual data, then, everything that is not included in the model, i.e. the errors, will become the residuals that have to behave randomly. If they were not behaving randomly, then our model does not fit well the data set and we need to find a more suitable one.

The lesson here is that regardless of the forecasting method, errors should always be calculated and inspected. Should these errors follow any kind of pattern, then the forecasting method must be treated as suspect. In fact, forecasting errors are often required to adhere to some formal assumptions, such as: **linearity**, **independence**, **normality** and **homoscedasticity**. These are the concepts that we covered in the chapter on linear regression. The procedure is identical, regardless of whether we are working with regression analysis or any time series method. For this reason, we suggest you revisit the relevant section in Chapter 8 of the textbook.

One of the validation methods for confirming that the forecasting method has produced credible results is the autocorrelation analysis of errors. We did not cover this approach in chapter on linear regression, so we’ll do it here.

**Example**

We’ll use Brent oil spot prices in USD per barrel from January 2016 until September 2017. We’ll attempt to fit a trend to this time series, extrapolate it until June 2018 and conclude, by examining the forecasting errors, if this method is appropriate for this time series. Figure 1 shows the whole example.



Figure 1

**Excel solution**

Date Cell A4:A24 Values

Period Cell B4:B24 Values

Time series Cell C4:C24 Values

c= Cell I1 Values

b= Cell I2 Values

Trend Cell D4 Formula: =$I$1\*LN(B4)+$I$2

 Copy down D5:D24

Forecast Cell E25 Formula: =$I$1\*LN(B25)+$I$2

 Copy down E26:E33

Errors Cell F4 Formula: =C4-D4

 Copy down F5:F24

Deviation from

the mean Cell G4 Formula: =F4-AVERAGE($F$4:$F$24)

 Copy down G5:G24

ACF Cell H4 Formula:

=SUMPRODUCT($G$4:INDEX($G$4:$G$24,ROWS(G5:G$24)),$G5:G$24)/DEVSQ($G$4:$G$24)

Copy down H5:H14

SE Cell I4 Formula: =SQRT(1/COUNT($G$4:$G$24))

 Copy down I5:I14

+CI Cell J4 Formula: =1.96\*I4

 Copy down J5:J14

-CI Cell K4 Formula: =-1.96\*I4

 Copy down K5:K14

The trend we decided to use here was a logarithmic trend. The format for the logarithmic trend is given by the equation (1)

y = c ℓn x + b (1)

which implies that it has two parameters, c and b. The values of these parameters were provided directly by Format Trendline Excel function and we put them in the cells I1 and I2. The graph below in Figure 2 shows the plot provided by Excel.



Figure 2

Let’s now focus on errors, which is our aim in this section. Column F in Figure W8.1 contains forecasting errors and the graph in Figure 3 shows how they look as a line graph.



Figure 3

They look reasonably random, but as we said in Chapter 8, they should be subject to further testing of linearity, independence, normality and homoscedasticity. We will not repeat these tests here, but we will take a look at the autocorrelation function of these errors. If they are genuinely random, they should be insignificant and appear to be below the standard error corridor. Figure 4 shows in a graph the results from columns H:K from Figure 1.



Figure 4

As we can see all the error autocorrelation coefficients are non-zero, meaning that they are inside the 95% limits defined by the ±1.96 standard errors. Not a single coefficient is “sticking” above the confidence limits, indicating that there is no correlation among the errors. We concluded that the errors are random and that, therefore, our logarithmic trend approximates the oil prices well, at least for the given time window.