ORIE 4820: Spreadsheet-Based Modeling and Data Analysis
Trend Estimation
Spring 2013

The purpose of this lab exercise is to give your project team an opportunity to ***design a set of demand projection methods*** that will be of use to a Liquair-Pro salesperson in determining a price quote for an ***existing customer***.

The template file for this exercise is ***trend-estimation.xlsm***. Download a copy of the file from the course Blackboard site and *save it on your computer*.

Topics/Tools used to construct the template file:

- ***Parsing text*** and ***converting text to numbers***: **right**, **value**
- ***Using array functions*** to select values from a restricted set: **if**
- ***Transforming data*** for use in regression models: **ln**
- ***Performing dynamic linear regression***: **linest**
- ***Evaluating the goodness of fit*** of simple regression models: **rsq, steyx**
- ***Using regression model output to make projections***

Background:

The workbook ***trend-estimation.xlsm*** consists of a single worksheet that contains customer information and up to seven years' worth of historical quarterly demand data for five existing Liquair-Pro customers (G – K), as well as quarterly data for a potential new customer (L) that is thinking of switching to Liquair-Pro from its current supplier. Already completed on the worksheet is a simple regression tool that determines the ***best fit linear trend line*** and the ***best fit exponential trend line*** for the selected customer's demand data at the annual level. These trend lines are projected forward to ***provide point estimate demand projections for the next five years***.

The linear and exponential regression models were constructed using the steps in Sections 1-3 below. Walk through these steps with your team to ensure you understand how the projections are being made. Then tackle the questions at the bottom of page 3.

Section 1: Summarizing and Transforming Data

Data validation has been used on cell L5 so that the user may select any of the customers from a dropdown list. In the *Selected Customer Demand* table:

(1) Column K has already been populated with the corresponding column of quarterly demand data using the **vlookup** and **match** functions. Note that a zero (0) will result if the source data element is blank.

(2) ***Populate the Year column*** (column L) by *extracting the rightmost four characters* from the corresponding cell in column J using the **right** function and then *converting the result to a numeric value* using the **value** function. For instance, cell L16 should contain "=VALUE(RIGHT(J16,4))".

(3) ***Populate the First and Last Year cells*** (L8 and L9), making sure to *only consider those years where demand is greater than 0* (recall that you can do this by using **if** as an array function). Once this is achieved, populate cell L7 accordingly.

In the *Best Linear Fit* table:

(4) Note that column P has been populated to *display only those years having nonzero demand*.

(5) Summarize the annual demands in column Q using the **sumif** function.

In the *Best Exponential Fit* table:

(6) Note that column X has been populated to *display only those years having nonzero demand*.

(7) Summarize the annual demands in column Y using the **sumif** function.

(8) Take the natural logarithm of these values in column Z using **ln**. We can now use ***simple linear regression*** to get the (transformed) coefficients of the best fit exponential curve.

Section 2: Dynamic Simple Linear Regression

Using the existing infrastructure on the worksheet, we will build a dynamic regression tool that is faster and more convenient than using Excel's static regression tool.

(1) **Regression Coefficients and Standard Error of Coefficients**: The array function **linest** uses the "least squares" method to calculate a straight line that best fits the designated *Y* and *X* data (i.e., the response and explanatory values that you give it), and it ***returns an entire matrix of information***, including the regression line coefficients (i.e., the regression line slope $\hat{b}$ and intercept $\hat{a}$) and their corresponding standard errors. **Linest** can be used for either simple or multiple linear regression. In the *Best Linear Fit* and *Best Exponential Fit* tables, you will use the **linest** function to populate the designated cells with the regression line coefficients and their corresponding standard errors by entering it as an ***array formula***. For instance, in the *Best Linear Fit* table:

   (a) Highlight the four cells R25:S26.

   (b) Type "=LINEST(OFFSET(Q16,0,0,L7,1),OFFSET(O16,0,0,L7,1),TRUE,TRUE)" and press ***Ctrl-Shift-Enter***. It is ***essential*** that you specify the third and fourth arguments of the **linest** function as **TRUE**.

   <u>Note:</u> The **linest** function cannot handle blanks or zeroes. We are using the **offset** function here to ensure that only actual data are passed into the **linest** function. You can also find the regression line coefficients $\hat{b}$ and $\hat{a}$, respectively, using the functions **slope** and **intercept**. However, you need to use the **linest** function to get the standard errors of these coefficients, which are needed if you want to construct confidence intervals for these parameters.

   (c) Do the same for the *Best Exponential Fit* table, using the log demand vector as the first input vector to the **linest** function.

(2) ***R* Square and Standard Error of Estimate for the linear regression line**: Cells R28 and R29 should contain, respectively, the *R* Square value and the standard error of estimate associated with the linear regression line. You can use the Excel functions **rsq** and **steyx**, respectively, to do this**.** Be sure to enter the demand vector as the first vector, and time (t) as the second vector. Do the same for the *Best Exponential Fit* table.

(3) **Fitted Demand and Residual columns**: Cells R16:R22 and AA16:AA22 should contain the ***fitted*** annual demand values corresponding to each year in the dataset (i.e., the predicted values according to the regression coefficients). Cells S16:S22 and AB16:AB22 should contain the associated residual errors. For example, in the *Best Exponential Fit* table, given the t and Y_t values in columns W and Y, columns AA and AB should contain, respectively:

$$\hat{Y}_t = e^{\hat{a}} \cdot e^{\hat{b}t}$$

$$e_t = Y_t - \hat{Y}_t$$

(4) **Annual Demand Prediction**: Cells R34:R38 and Z34:Z38 should contain, respectively, the predicted annual demand for 2012-2016 based on the regression coefficients for the best linear fit and the best exponential fit.

(5) **Standard Error of Estimate for the original annual demand data (s_e)**: In cells S34 and AA34, to get the standard error of estimate associated with the annual demand data, you can use the fact that this value is (essentially) the ***standard deviation of the residual errors***, except that we use $n - 2$ in the denominator (instead of n - 1), since the residuals are based on the estimation of 2 parameters (a and b). That is:

$$s_e = \sqrt{\frac{\sum e_i^{\,2}}{n-2}}$$

(6) **Annual Growth Rate and its Standard Error**: Cell Z30 should contain the ***annual growth rate r*** of the exponential curve corresponding to the regression line. That is:

$$r = e^{\hat{b}} - 1$$

Finally, we will approximate the ***standard error of the growth rate*** in cell Z31 by taking the same transformation on the standard error of $\hat{b}$ (in cell Z26).

<u>Get together with your project team and consider the following questions:</u>

(A) ***For which customers do the linear and/or exponential trend tools appear to be helpful in making demand projections? Why or why not? What are the risks and limitations of using these projections?***

(B) ***What <u>other</u> demand projection methods do you feel may be useful for Liquair-Pro salespeople to have at their disposal?***

(C) ***Assume you want to provide the capability for a Liquair-Pro salesperson to <u>easily</u> incorporate a set of demand projections for an existing customer into a pricing tool. How will you design and implement this process?***