Exploring Logistic Regression with E-STAT Data

Joel Yan, Learning Resources Team, Statistics Canada, yanjoel@statcan.ca, 1-800-465-1222

Downloading Statistics Canada E-STAT Data to TI InterActive! and the TI-83

A. Make Connection to TI-83 (Plus)

1. Connect TI-Graph Link cable to communication port and TI-83
2. In TI InterActive!, select in sequence: **Edit**>**Preferences**>**Communications**

   Set: a) calculator type
   b) com port
   c) TI-Graph Link cable type

B. Access Statistics Canada Website (from a computer connected to the Internet)

1. Select TI InterActive! web browser by clicking the globe icon in the top toolbar near the right
2. Go to: www.statcan.ca
3. Click in sequence: **English**> **Learning Resources**>**3-STAT**>**Accept**

C. Download a Data Set (e.g. Cable TV revenues)

1. Click in sequence: **Communications** (found under **Economy**) > **Data**> **Broadcasting**> **Table 353-001**

   (This sequence selects the table which contains data on the Operating and financial summary of the cable television industry, annual since 1961)

2. Click in sequence:

   Canada under **Geography**
   Operating revenue, total under **Operating and financial detail**
   Scroll down to the year and click on the down arrow to change the **From** year to **1976**
   At the bottom of the page Click **Retrieve as Time series** >

3. Set the Output format selection to: **HTML Table, time as rows**
4. Click: **Go**
5. Highlight the two columns of data in the table.
6. In TI InterActive! browser click **Extract**

D. Name Lists and Copy to TI InterActive! Document

1. In Data Editor window, double click on the list names (Annual & V81330)
2. Edit the list names to year and cablesales
3. Click on the upper left save icon to copy the list table into the TI InterActive! document.

E. Download from TI InterActive! to TI-83

1. Double click on the list table
2. In Data Editor, highlight lists to download by:

   Click first list name
   Hold Shift & click second list name

3. Select in sequence: **File> Export> to TI Calculator**

F. Graphic Manipulation

1. Graph the cable TV revenue in TI InterActive! or your graphing calculator.
2. Calculate a logistic regression curve for the cable TV revenue. Include this logistic regression curve on your graph of cable sales.

G. Graphic Manipulation Instructions to use in TI InterActive!

1. To graph the male life expectancy, double click on the Data editor.

   Control B click on the columns to be graphed, (year and cablesales), then click on the **graph** icon (the icon with the scatterplot). This generates a scatter graph of sales over time.

   Customize the colours and type of graph. You can also add titles, and other labelling by clicking the **Format** button
   When done, save the graph back to your document.
2. To calculate the logistic regression line for males, double click on the table to enter the Data Editor.
3. Control-click on the columns for the logistic regression. This selects the appropriate data for the regression. Then select the **Stat Calculation tool**. Name the regression equation MaleEQ(x) so that you can graph it easily in the function editor.
   a) Calculation type: - logistic regression
   b) Xlist B Year
   c) Ylist B cablesales
   d) Frequency B must be blank
   e) Regression equation - regEQ(x)

      Click the Calculate button at the bottom.
      Then enter a title and Save the results to your document
      Save the table back to your document.
4. Including the line of best fit on the graph.

   Double click on the Graph to bring up the Graph editor

   Click the **Functions** button and then the **f(x)** tab. Now type regEQ(x) as the function to be graphed.

   This plots the regression line on the existing graph.

   Note: you must use the exact name you used for storing the regression equation.
5. Once this is done, answer the questions below.

Questions / Investigations

1. How accurately can we approximate cable TV revenues using a logistic regression?
2. What is it about cable TV revenues that lets it be modelled so well by this exponential / logistic curve?
3. Based on the size of these revenues, what are some reasons for the cable TV industry to be regulated or not regulated by government?
4. Try modelling the data using different types of regression, including exponential regression. Do any of these closely approximate the data observed, for at least part of the time period?