Lab 5 (Week 9 - 10)

STAT 2601 - Business Statistics (2024 Fall)
School of Mathematics and Statistics, Carleton University

Schedule: November 18 - 22, 2024

Q: Income and Credit Card Balance (Correlation and Regression)

A bank issues MasterCard credit cards to its customers. A main factor in determining whether a credit card will be profitable to the bank is the average monthly balance that the customer will maintain on the card that will be subject to finance charges. Bank analysts wish to determine whether there is a relationship between the average monthly credit card balance and the income stated on the original credit card application form. The sample data collected from existing 15 credit card customers is located in a data file named **Lab 5_Data.xlsx**.

(a) Use EXCEL to construct a scatter plot of the Credit Balance (Y) versus Income (X) for the sample data.

(b) Use EXCEL to calculate sample correlation coefficient, r between income and credit balance.

(c) Use EXCEL to develop an estimated simple regression model that could be used to predict the credit balance given the income.

(d) At the 0.05 level of significance, determine whether the credit balance and income are linearly related (Test of slope parameter) using EXCEL output.

EXCEL Instructions

- Open the data set **Lab 5_Data.xlsx**.
- **Scatter Plot:** Select the data (**A2:B16**) > Insert > Insert Scatter Chart (Select the first category). Select the chart and click on "+" that appears at the right corner of the chart > Select Axis Title > Now customize X axis as "Income", Y axis as "Credit Balance", and chart title as "Scatter Plot of Income vs Credit Balance".
- **Correlation Coefficient:** Data > Data Analysis > Correlation > Select data (including variable name) ($A$1:$B$16) in the box of Input Range > Select Labels in Firt Row > OK. Alternatively, you can type in any cell **=CORREL(A2:A16,B2:B16)** to get the sample correlation coefficient.
- **Estimated Simple Regression Model:** Data > Data Analysis > Regression > Input Y Range ($B$2:$B$16) > Input X Range ($A$2:$A$16) > OK.
- **Critical Value and p-value Approach**
  (a) Critical Value Approach: Find critical value (type in any cell **=TINV(.05,13) and hit Enter**) and compare with t test statistic produced in EXCEL output. Please note that the degrees of freedom for t test in testing slope is $(n - k - 1)$ where n is the sample size and k is the number of independent variables.
  (b) p-value Approach: Compare p-value produced in EXCEL output with level of significance $\alpha = 0.05$.