Econ 325 Section 003
Assignment 3

The due date is Friday, October 14, before the class starts.

1. In this Stata exercise, we examine the relationship between CEO salary (salary) and return on equity (roe) for the company. The data also includes a dummy variable for financial firms (finance = 1 if a firm is a financial firm; finance=0 if a firm is not a financial firm).

   Download the Stata data file on CEO salary, return on equity, and dummy variable for financial firms from

   http://faculty.arts.ubc.ca/hkasahara/Econ325/ceosal.dta

   An example of Stata code is available at

   http://faculty.arts.ubc.ca/hkasahara/Econ325/ceosal.do

   (a) Use the Stata command "summarize" together with "if" command with an option "detail" to answer the following questions.
      - How many observations are in the data set? Among them, how many are financial firms and non-financial firms?
      - Is it true that the average and the median CEO salary is higher for financial firms than for non-financial firms?

   (b) What is the median value of return on equity? Define "high return firm" as a firm of which return on equity is higher than the sample median and define "low return firm" as a firm of which return on equity is lower than the sample median. Is it true that, on average, the CEO's in high return firms receive higher salary than those in low return firms in this sample?

   (c) Find the average salary of CEO's of industrial firms with a return on equity in the top 25% of industrial firms or utility firms with a return on equity above the median return of utility firms.

   (d) Generate a couple of scatter plots for CEO salary and return on equity: one for financial firms and one for non-financial firms. Also, compute the sample correlation coefficient between CEO salary and return on equity using the command "correlation" for financial firms and non-financial firms. Are there any positive or negative relation between CEO salary and return on equity for financial firms and for non-financial firms?

2. Please read "Notes on mathematical expectation, variance, and covariance" posted on the course website. Let X and Y be two discrete random variables. The set of possible values for X is $\{x_1, \ldots, x_n\}$; and the set of possible values for Y is $\{y_1, \ldots, y_m\}$. The joint PMF (Probability Mass Function) is given by

$$p_{ij}^{X,Y} = P\left(X = x_i, Y = y_j\right), \qquad i = 1, \ldots n; j = 1, \ldots, m.$$

The marginal PMF of X is

$$p_i^X = P(X = x_i) = \sum_{j=1}^{m} p_{ij}^{X,Y}, \qquad i = 1, \dots n,$$

and the marginal PMF of Y is

$$p_j^Y = P(Y = y_j) = \sum_{i=1}^{n} p_{ij}^{X,Y}, \qquad i = 1, \dots n,$$

Prove the following results for general n and m. For (a)-(g), please make sure that you use the summation operator in your answer. For (h)-(j), you don't necessarily use the summation operator for brevity but use the expectation to derive the result.

(a) $E[X+Y] = E[X] + E[Y]$.

(b) $Cov(X,Y) = E[XY] - E[X]E[Y]$.

(c) If c is a constant, then $Cov(X, c) = 0$.

(d) $Cov(a_1 + b_1 X, a_2 + b_2 Y) = b_1 b_2 Cov(X, Y)$, where a_1, a_2, b_1, and b_2 are some constants.

(e) If X and Y are independent then $Cov(X, Y) = 0$.

(f) Prove that $E_Y[Y] = E_X[E_Y[Y|X]]$.

(g) Let $g_1(x)$ and $g_2(x)$ be some functions of x. Prove that $Var(g_1(X) + g_2(X)) = Var(g_1(X)) + Var(g_2(X)) + 2Cov(g_1(X), g_2(X))$.

(h) Let b be a constant. Show that $E[(X-b)^2] = E(X^2) - 2bE(X) + b^2$. What is the constant value of b that gives the minimum value of $E[(X-b)^2]$?

(i) Define $W = (X - E(X))/\sqrt{Var(X)}$. Prove that $\text{Corr}(X, W) = 1$.

(j) Which of the followings is **True**:

A) $E[Y|X = x_i] = \sum_{j=1}^{m} y_j \frac{p_{ij}^{X,Y}}{p_i^X}$ B) $E[Y|X = x_i] = \sum_{j=1}^{m} y_j p_{ij}^{X,Y}$

C) $E[Y|X = x_i] = \sum_{j=1}^{m} y_j p_j^Y p_i^X$ D) $E[Y|X = x_i] = \sum_{j=1}^{m} y_j \frac{p_{ij}^{X,Y}}{p_j^Y}$.

3. Suppose that both X and Y are random variables and are not constants. Which of the followings is generally **False**?

A) $E[Y|X] = E[Y]$ if X and Y are independent. B) $E[Y|X]$ is always a constant.
C) $E[Y|X]$ is not always a constant but generally a random variable that depends on a random variable X.

4. Please read "Notes on Bernoulli and Binomial Distribution" posted on the course website. Let X_1 and X_2 are two Bernoulli random variables with the probability of success p, where X_1 and X_2 are independent, and $X_i = 0$ with probability $1-p$ and $X_i = 1$ with probability p for $i = 1, 2$. Define a random variable $Y = X_1 - X_2$.

(a) Find the mean and the variance of Y.

(b) What is $E[Y|X_1 = 1]$?