Economics 325–003 **Term I 2014**

Introduction to Empirical Economics **Hiro Kasahara**

Midterm Exam

1. (8 points) Given the data set that contains variables named "salary" and "roe," what Stata command computes the correlation coefficient as in the following output?

```
(obs=209)

             |   salary      roe
-------------+------------------
      salary |   1.0000
         roe |   0.1148   1.0000
```

   Answer: "correlate salary roe" (or "pwcorr salary roe")

2. Multiple Choice Questions (No Explanation Necessary):

   (a) (8 points) 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.

   Answer: B. $E[Y|X]$ is a random variable because it depends on the random variable X.

   (b) (8 points) Suppose that both X and Y are random variables and define $W = (X - E(X))/\sqrt{Var(X)}$ and $Z = (Y - E(Y))/\sqrt{Var(Y)}$. Which of the followings is generally **False**?

   A) $Cov(W, Z) = Corr(X, Y)$ B) $Var(W) = 1$ C) $E[WZ] = Corr(X, Y)$ D) $Var(WZ) = Var(XY)$.

   Answer: D. For A, the proof is available in the note. For B, $Var(W) = Var(X - E(X))/Var(X) = Var(X)/Var(X) = 1$**. For C,** $E(WZ) = E[(X-E(X))(Y-E(Y))]/(\sqrt{Var(X)}\sqrt{Var(Y)}) = Cov(X,Y)/(\sqrt{Var(X)}\sqrt{Var(Y)}) = Corr(X,Y)$**.**

   (c) (8 points) If A and B are independent events with $P(A) = 0.30$ and $P(B) = 0.40$, then the probability that A occurs or B occurs or both occur is:

A) 0.12 B) 0.58 C) 0.70 D) 0.82.

Answer: B. When A and B are independent $P(A \cap B) = P(A)P(B)$**. Therefore,** $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = 0.58$**.**

3. (10 points) In one year, the average stock price of Google Inc. was \$560 with the standard deviation equal to \$30. Using the empirical rule, it can be estimated that approximately 95 % of the stock price of Google Inc. will be in what interval?

Answer: $560 \pm 2 \times 30$ **gives** $[500, 620]$**.**

4. (10 points) Consider the joint probability distribution of (X, Y):

		X	
		0	1
Y	0	0.2	0.4
	1	0.3	0.1

Compute the covariance between X and Y.

Answer: $Cov(X, Y) = E[XY] - E[X]E[Y] = 0.1 - (0.5)(0.4) = -0.1$

5. (10 points) A review of the personnel records of a small corporation has revealed the following information about the number of sick days taken per year and the corresponding probabilities.

Number of Sick Days	0	1	2	3	4	5
Probability	0.05	0.22	0.31	0.27	0.13	0.02

Let A be the event that an employee takes more than 2 sick days (i.e., Number of Sick Days ≥ 3). Compute the probability of event A.

Answer: 0.27 + 0.13 + 0.02 = 0.42

6. (10 points) Given the following table, what is the probability that a randomly selected person is female conditional on the selected person does not support US policy in Iraq?

	US Policy in Iraq		
	Support	Doesn't Support	Row Total
Female	0.2438	0.2862	0.53
Male	0.3762	0.0938	0.47
Column Total	0.62	0.38	1.00

Answer: 0.2862/0.38=0.753

7. 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 function of X and Y is given by $p_{ij}^{X,Y} = P(X = x_i, Y = y_j)$ for $i = 1, \ldots n; j = 1, \ldots, m$. The marginal probability function of X is $p_i^X = P(X = x_i) = \sum_{j=1}^m p_{ij}^{X,Y}$ for $i = 1, \ldots n$, and the marginal probability function of Y is $p_j^Y = P(Y = y_j) = \sum_{i=1}^n p_{ij}^{X,Y}$ for $j = 1, \ldots m$.

Prove the following results for general n and m.

(a) (10 points) Prove that $Cov((X - E(X)), Y) = Cov(X, Y)$.

Answer: $Cov((X-E(X)), Y) = E[\{(X-E(X))-E(X-E(X))\}(Y-E(Y))] = E[\{(X-E(X))-(E(X)-E(X))\}(Y-E(Y))] = E[\{(X-E(X))-0\}(Y-E(Y))] = E[(X-E(X))(Y-E(Y))] = Cov(X, Y)$**.**

(b) (10 points) Prove that $E(XY) = E(X)E(Y)$ if X and Y are independent using the summation operator together with notations $p_{ij}^{X,Y}$, p_i^X, and p_j^Y defined above.

Answer: First, note that $p_{ij}^{X,Y} = p_i^X \times p_j^Y$ if X and Y are independent. Therefore, $E(XY) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_{ij}^{X,Y} = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_i^X \times p_j^Y = \sum_{i=1}^n x_i p_i^X \sum_{j=1}^m y_j p_j^Y = \sum_{i=1}^n x_i p_i^X E(Y) = E(Y) \sum_{i=1}^n x_i p_i^X = E(Y)E(X)$, where the first equality is the definition of expectation, the second equality uses $p_{ij}^{X,Y} = p_i^X \times p_j^Y$, the third equality follows because $x_i p_i^X$ is a common factor across m terms in $\sum_{j=1}^m x_i y_j p_i^X \times p_j^Y = (x_i p_i^X y_1 p_1^Y + \ldots + x_i p_i^X y_m p_m^Y)$, the fourth equality follows from the definition of $E(Y)$, the fifth equality

follows because $E(Y)$ is a common factor in $\sum_{i=1}^{n} x_i p_i^X E(Y) = (x_1 p_1^X E(Y) + ... + x_n p_n^X E(Y))$ and the last equality follows from the definition of $E(X)$.

(c) (8 points) Multiple Choice Questions (No Explanation Necessary). 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) $E[Y|X = x_i] = \sum_{j=1}^{m} y_j \frac{p_{ij}^{X,Y}}{p_j^Y}$.

Answer: A