Economics 325–003 **Term I 2013**

Introduction to Empirical Economics **Hiro Kasahara**

Midterm Exam

1. (10 points) Given the data set that contains variables named "salary" and "roe," what Stata command produces the following output?

```
                 1990 salary, thousands $
-------------------------------------------------------------
      Percentiles      Smallest
 1%          333            223
 5%          448            256
10%          525            333       Obs                 209
25%          736            360       Sum of Wgt.         209

50%         1039                      Mean            1281.12
                        Largest       Std. Dev.      1372.345
75%         1407           4143
90%         1900           6640       Variance        1883332
95%         2327          11233       Skewness       6.854923
99%         6640          14822       Kurtosis       60.54128

                return on equity, 88-90 avg
-------------------------------------------------------------
      Percentiles      Smallest
 1%          2.1             .5
 5%          6.8            1.9
10%          8.9            2.1       Obs                 209
25%         12.4            2.9       Sum of Wgt.         209

50%         15.5                      Mean           17.18421
                        Largest       Std. Dev.      8.518509
75%           20           44.4
90%         26.8           44.5       Variance       72.56499
95%         35.1           48.1       Skewness        1.56082
99%         44.5           56.3       Kurtosis       6.678555
```

Answer: summarize salary roe, detail

2. Multiple Choice Questions (No Explanation Necessary):

   (a) (10 points) Let $a = -E(X)/\sqrt{Var(X)}$ and $b = 1/\sqrt{Var(X)}$, where X is a random variable. Define the random variable $Z = a + bX$. Which of the following statements is **not true**?

A) $E(Z) = 0$, B) $Var(Z) = 1$, C) $Cov(X, Z) = \sqrt{Var(X)}$, D) $E[XZ] = \sqrt{Var(X)}$, E) $Corr(X, Z) = Var(X)$.

Answer: E because $Corr(X, Z) = 1$

(b) (10 points) If A and B are independent events with $P(A) = 0.60$ and $P(B) = 0.70$, then the probability that A occurs or B occurs or both occur is:
A) 1.30, B) 0.88, C) 0.42, D) 0.10.

Answer: B because $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)P(B) = 0.88$**, where the last equality uses the independence of** A **and** B**.**

3. A general contractor has submitted two bids for two projects; A and B. The probability of getting project A is 0.60. The probability of getting project B is 0.75. The probability of getting at least one of the projects is 0.85.

(a) (10 points) What is the probability that the contractor will get both projects?

Answer: $P(A) = 0.60$, $P(B) = 0.75$, **and** $P(A \cup B) = 0.85$. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ **implies that** $0.85 = 0.60 + 0.75 - P(A \cap B)$. **Hence,** $P(A \cap B) = 0.50$.

(b) (10 points) If the contractor gets project A, what is the probability that he will get project B?
Answer: $0.50/0.60 = 0.833$

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 variance of $W = X - Y$.

Answer: Note that $Var(W) = Var(X) + Var(Y) - 2Cov(X,Y) = (E(X^2) - E(X)^2) + (E(Y^2) - E(Y)^2) - 2(E(XY) - E(X)E(Y))$**. It is easy to compute** $E(X^2) = E(X) = 0.5$**,** $E(Y^2) = E(Y) = 0.4$**, and** $E(XY) = 0.1$**. Therefore,** $Var(W) = 0.25 + 0.24 - 2 \times (0.1 - 0.2) = 0.69$**.**

5. (10 points) Let Z_1 and Z_2 are two Bernoulli random variables with the probability of success p, where Z_1 and Z_2 are independent, and $Z_i = 0$ with probability $1 - p$ and $Z_i = 1$ with probability p for $i = 1, 2$. Define a random variable $X = Z_1 + Z_2$. Prove that $Var(X) = 2p(1 - p)$.

   Answer: $E(Z_i) = p$ **and** $Var(Z_i) = (1-p)^2 p + (0-p)^2(1-p) = p(1-p)$ **for** $i = 1, 2$**.** $Var(X) = Var(Z_1) + Var(Z_2) + 2Cov(Z_1, Z_2) = Var(Z_1) + Var(Z_2)$**, where the last equality follows because** Z_1 **and** Z_2 **are independent. Therefore,** $Var(X) = p(1-p) + p(1-p) = 2p(1-p)$**.**

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

   Prove the following results for general n and m. (*Please use the summation operator in the proof*):

   (a) (10 points) Let $P(X|Y = y_j)$ be the conditional probability function of the random variable X given $Y = y_j$. Prove that $\sum_{i=1}^{n} P(X = x_i | Y = y_j) = 1$ for any $j = 1, ..., m$.

   Answer: Note that

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{P_{ij}^{X,Y}}{\sum_{k=1}^{n} P_{kj}^{X,Y}}.$$

   Therefore,

$$\sum_{i=1}^{n} P(X = x_i | Y = y_j) = \sum_{i=1}^{n} \frac{P_{ij}^{X,Y}}{\sum_{k=1}^{n} P_{kj}^{X,Y}} = \frac{1}{\sum_{k=1}^{n} P_{kj}^{X,Y}} \sum_{i=1}^{n} P_{ij}^{X,Y} = 1.$$

   (b) (10 points) Define $W = (X - E(X))/\sqrt{Var(X)}$ and $Z = (Y - E(Y))/\sqrt{Var(Y)}$. Prove that $Cov(W, Z) = Corr(X, Y)$.

Answer: First, we have

$$E[W] = \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{x_i - E(X)}{\sqrt{Var(X)}} P_{ij}^{X,Y} = \frac{1}{\sqrt{Var(X)}} \left(\sum_{i=1}^{n} x_i (\sum_{j=1}^{m} P_{ij}^{X,Y}) - E(X) \right)$$
$$= \frac{1}{\sqrt{Var(X)}} \left(\sum_{i=1}^{n} x_i p_i^X - E(X) \right) = \frac{1}{\sqrt{Var(X)}} (E(X) - E(X)) = 0.$$

Similarly, we may prove that $E(Z) = 0$. Then,

$$Cov(W, Z) = E\left(\frac{X - E(X)}{\sqrt{Var(X)}} \frac{Y - E(Y)}{\sqrt{Var(Y)}} \right) = \frac{1}{\sqrt{Var(X)}\sqrt{Var(Y)}} E\left((X - E(X))(Y - E(Y))\right)$$
$$= \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = Corr(X, Y).$$

7. (10 points) True or False Question.

The joint probability distribution of random variables X and Y is shown in the table below.

	X		
Y	1	2	3
1	0.30	0.18	0.12
2	0.15	0.09	0.06
3	0.05	0.03	0.02

Is it true or false that $Var(X + Y) = Var(X) + Var(Y)$? Please answer "True" or "False." No explanation necessary.

Answer: True. Because $Var(X+Y) = Var(X)+Var(Y)+2Cov(X,Y)$ in general, we need to find out $Cov(X,Y) = 0$ or not. For this, it is useful to use the relationship $Cov(X,Y) = E(XY) - E(X)E(Y)$. Note that $E(XY) = 1*0.3+2*0.33+3*0.17+4*0.09+6*0.09+9*0.02 = 2.55$, $E(X) = 1*0.5+2*0.3+3*0.2 = 1.7$, and $E(Y) = 1*0.6+2*0.3+3*0.1 = 1.5$. Therefore, $Cov(X,Y) = 2.55 - 1.7*1.5 = 0$.