Economics 325–001 **Term I 2019**

Introduction to Empirical Economics **Hiro Kasahara**

Midterm Exam[1]

1. (5 points for each) True or False **Write your answer (i.e., "True" or "False") to your booklet.** No explanation necessary.

   (a) $E[(X - E(X))(Y - E(Y))] = E[X(Y - E(Y))]$. Answer: True because $E[(X - E(X))(Y - E(Y))] = E[X(Y - E(Y)) - E(X)(Y - E(Y))] = E[X(Y - E(Y))] - E[E(X)(Y - E(Y))] = E[X(Y - E(Y))] - E(X)E[(Y - E(Y))] = E[X(Y - E(Y))] - E(X) \times 0 = E[X(Y - E(Y))]$.

   (b) $E\left[\frac{Y}{X}\right] = \frac{E[Y]}{E[X]}$. Answer: False.

   (c) Define $Z = (X - E(X))/\sqrt{Var(X)}$. Then, $\text{Corr}(X, Z) = 1$. Answer: True.

   (d) Let $X_1, X_2, ..., X_n$ are n random variables that are stochastically independent to each other with the same mean $\mu = E[X_i]$ for $i = 1, 2, ..., n$. Then, $\mu = \frac{1}{n}\sum_{i=1}^{n} X_i$ always holds. Answer: False because $\frac{1}{n}\sum_{i=1}^{n} X_i$ is a random variable but μ is constant.

2. (8 points) In one year, the average stock price of Apple Inc. was \$200 with the standard deviation equal to \$10. Using the empirical rule, it can be estimated that approximately 95 % of the stock price of Apple Inc. will be in what interval?

   Answer: $200 \pm 2 \times 10 = [180, 220]$.

3. (8 points) Suppose a personnel officer has 8 candidates to fill 2 positions. Among 8 candidates, 5 candidates are men and 3 candidates are women. If every candidate is equally likely to be chosen, what is the probability that no women will be hired?

   Answer: There are $C^8_2 = \frac{8!}{(8-2)!2!} = 28$ possible basic outcomes and, among them, $C^5_2 = \frac{5!}{(5-2)!2!} = 10$ cases for which no women will be hired. Therefore, the probability that no women will be hired is $\frac{C^5_2}{C^8_2} = \frac{10}{28} = \frac{5}{14}$.

4. There is a new diagnostic test for a disease that occurs in 10 percent of the population. The test will detect a person with the disease 80 percent of the time when a person has the disease. It will, however, say falsely that a person without the disease has the disease about 10 percent of the time. Suppose that the test conducted on a randomly selected person indicates that this person has the disease.

What is the probability that this person has the disease. Define $A = \{\text{positive diagnostic test}\}$ and $B = \{\text{having the disease}\}$.

(a) (7 points) What is the probability of this person having the disease given that his or her diagnostic test is positive?

Answer: Note that $P(A|B) = 0.8$, $P(A|\bar{B}) = 0.1$, $P(B) = 0.1$, and $P(\bar{B}) = 0.9$. Therefore, using the Bayes' theorem,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.1 \times 0.9} = \frac{0.08}{0.08 + 0.09} = 8/17 \approx 0..4706.$$

(b) (7 points) Complete the following table for the joint probability distribution of A and B and **write your answer to your booklet**.

Table I: Joint distribution of A and B

	$\bar{B}$	B	Marginal Prob.
$\bar{A}$	0.81	0.02	0.83
A	0.09	0.08	0.17
Marginal Prob	0.9	0.1	1.00

Answer: $P(A \cap B) = P(A|B)P(B) = 0.8 \times 0.1 = 0.08$, $P(A \cap \bar{B}) = P(A|\bar{B})P(\bar{B}) = 0.1 \times 0.9 = 0.09$, $P(\bar{A} \cap \bar{B}) = P(\bar{A}|\bar{B})P(\bar{B}) = (1 - P(A|\bar{B}))P(\bar{B}) = (1 - 0.1) \times 0.9 = 0.81$, and $P(\bar{A}|B) = (1 - P(A|B))P(B) = (1 - 0.8) \times 0.1 = 0.08$. In sum, we have the following table for the joint distribution of A and B.

5. Let X_1 and X_2 are two Bernoulli random variables with the probability of success p, where X_1 and X_2 are stochastically independent, and $X_i = 0$ with probability $1-p$ and $X_i = 1$ with probability p for $i = 1, 2$. Define a random variable $Y = (X_1 + X_2)/2$.

   (a) (6 points) What is the probability mass function of Y?

   Answer: Note that $Y = X_1 + X_2$ takes 0, 1/2, and 1 when $(X_1, X_2) = (0,0)$, $(X_1, X_2) \in \{(1,0),(0,1)\}$, and $(X_1, X_2) = (1,1)$, which happens with the probabilities of $(1-p)^2$, $2p(1-p)$, and p^2, respectively. Therefore, thel probability mass function of Y is

$$f_Y(y) = \begin{cases} (1-p)^2 & \text{if } Y = 0, \\ 2p(1-p) & \text{if } Y = 1/2, \\ p^2 & \text{if } Y = 1. \end{cases}$$

   (b) (6 points) What are the mean and the variance of Y?

   Answer: Note that $E[X_1] = E[X_2] = p$ and $\text{Var}(X_1) = \text{Var}(X_2) = p(1-p)$. Therefore, $E[(X_1 + X_2)/2] = (1/2)(E[X_1] + E[X_2]) = (1/2)(p+p) = p$ and $Var((X_1+X_2)/2) = (1/2)^2(Var(X_1) + Var(X_2) + Cov(X_1, X_2)) = (1/4)(p(1-p) + p(1-p) + 0) = p(1-p)/2$, where $\text{Cov}(X_1, X_2) = 0$ holds because X_1 and X_2 are stochastically independent.

   (c) (6 points) What is the joint probability mass function of X_1 and Y?

   To derive the joint probability mass function, we need to compute $f(y, x_1) := P(Y = y, X_1 = x_1)$ for all possible pairs of values of (y, x_1). Because $(Y, X_1) = ((X_1+X_2)/2, X_1)$, the possible values of (Y, X_1) are determined by the possible values of (X_1, X_2), where

$$(Y, X_1) = ((X_1 + X_2)/2, X_1) = \begin{cases} (0,0) & \text{if } (X_1, X_2) = (0,0), \\ (1/2, 1) & \text{if } (X_1, X_2) = (1,0), \\ (1/2, 0) & \text{if } (X_1, X_2) = (0,1), \\ (1,1) & \text{if } (X_1, X_2) = (1,1). \end{cases}$$

   Therefore, the joint probability mass function is given by

$$f(y, x_1) := P(Y = y, X_1 = x_1) = \begin{cases} P((X_1, X_2) = (0,0)) = (1-p)^2 & \text{if } (y, x_1) = (0,0), \\ P((X_1, X_2) = (1,0)) = p(1-p) & \text{if } (y, x_1) = (1/2, 1), \\ P((X_1, X_2) = (0,1)) = p(1-p) & \text{if } (y, x_1) = (1/2, 0), \\ P((X_1, X_2) = (1,1)) = p^2 & \text{if } (y, x_1) = (1,1), \\ 0 & \text{otherwise.} \end{cases}$$

	Y = 0	Y = 1/2	Y = 1	Marginal Dist. of X_1
$X_1 = 0$	$(1-p)^2$	$p(1-p)$	0	$1-p$
$X_1 = 1$	0	$p(1-p)$	p^2	p
Marginal Dist. of Y	$(1-p)^2$	$2p(1-p)$	p^2	1.00

The above table summarizes the joint probability mass function of (Y, X_1).

(d) (6 points) Are X_1 and Y stochastically independent? **Prove your claim.** [Full 6 points are given to the proof.]

Answer: X_1 and Y are not stochastically independent. For proof, we need to show $\Pr((Y, X_1) = (y, x_1)) \neq \Pr(Y = y)\Pr(X_1 = x_1)$ for some value of (y, x_1). For example, $\Pr((Y, X_1) = (0, 0)) = (1-p)^2$, $\Pr(Y = 0) = (1-p)^2$, and $\Pr(X_1 = 0) = 1-p$. Therefore, $(1-p)^2 = \Pr((Y, X_1) = (0, 0)) \neq \Pr(Y = 0)\Pr(X_1 = 0) = (1-p)^3$ for any $p \neq 0$. Thus, X_1 and Y are not stochastically independent.

(e) (6 points) What is $E[Y|X_1 = 1] - E[Y|X_1 = 0]$?

Answer: $E[Y|X_1 = 1] - E[Y|X_1 = 0] = E[(X_1 + X_2)/2|X_1 = 1] - E[(X_1 + X_2)/2|X_1 = 0] = E[(1 + X_2)/2|X_1 = 1] - E[(0 + X_2)/2|X_1 = 0] = 1/2 + E[X_2/2|X_1 = 1] - E[X_2/2|X_1 = 0] = 1/2 + E[X_2/2] - E[X_2/2] = 1/2$, where $E[X_2/2|X_1 = 1] - E[X_2/2|X_1 = 0] = E[X_2/2] - E[X_2/2]$ holds because X_1 and X_2 are independent.

6. Let X and Y be two discrete random variables. The set of possible values for X is $\{x_1, \dots, x_n\}$; and the set of possible values for Y is $\{y_1, \dots, y_m\}$. The joint function of X and Y is given by $p_{ij}^{X,Y} = P(X = x_i, Y = y_j)$ for $i = 1, \dots n; j = 1, \dots, m$. The marginal probability function of X is $p_i^X = P(X = x_i) = \sum_{j=1}^{m} p_{ij}^{X,Y}$ for $i = 1, \dots n$, and the marginal probability function of Y is $p_j^Y = P(Y = y_j) = \sum_{i=1}^{n} p_{ij}^{X,Y}$ for $j = 1, \dots m$. The conditional probability of $X = x_i$ given $Y = y_j$ is $P(X = x_i|Y = y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)} = \frac{p_{ij}^{X,Y}}{p_j^Y}$ for $i = 1, \dots n$ and $j = 1, \dots m$.

(a) (10 points) Prove that $E[X|Y = y_j] = E[X]$ holds if X and Y are stochastically independent. **Please use the summation operator in the proof for this question**.

Answer:

$$\begin{aligned}
E[Y|X] &= \sum_{j=1}^{m} y_j \frac{p_{ij}^{X,Y}}{p_i^X} \quad \text{(by definition of conditional expectation)} \\
&= \sum_{j=1}^{m} y_j \frac{p_i^X p_j^Y}{p_i^X} \quad (p_{ij}^{X,Y} = p_i^X p_j^Y \text{ because } X \text{ and } Y \text{ are independent}) \\
&= \sum_{j=1}^{m} y_j p_j^Y \\
&= E[Y].
\end{aligned}$$

(b) (10 points) Define $Z = (X - E(X))/\sqrt{Var(X)}$. Prove that $E[Z] = 0$ and $Var[Z] = 1$. (You don't necessarily need to use the summation operator).

Answer:

$$\begin{aligned}
E[Z] &= E\left((X - E(X))/\sqrt{Var(X)}\right) \\
&= \frac{E[X - E[X]]}{\sqrt{Var(X)}} \\
&= \frac{E[X] - E[X]}{\sqrt{Var(X)}} \\
&= \frac{0}{\sqrt{Var(X)}} = 0.
\end{aligned}$$

$$\begin{aligned}
Var(Z) &= Var((X - E(X))/\sqrt{Var(X)}) \\
&= \frac{Var(X - E(X))}{Var(X)} \quad \text{[Because } Var(bX) = b^2 Var(X) \text{ where } b = 1/\sqrt{Var(X)}.] \\
&= \frac{Var(X)}{Var(X)} \quad \text{[Because } Var(X - E(X)) = E[\{X - E(X) - 0\}^2] = Var(X).] \\
&= 1.
\end{aligned}$$