Economics 325–003/004 **Term I 2012**

Probability and Statistics for Economists **Hiro Kasahara**

Midterm Exam

1. Answer the following questions regarding Stata command. The data set contains variables "lexp" and "region." The label for "lexp" is given by "Life expectancy at birth." The variable "region" contains the information on which region each country belongs to; there are three regions: "Eur & C.Asia," "N.A.," and "S.A.".

   (a) (8 points) What is a Stata command that will produce the following Stata output? (Please write a whole line of command that produces the output.)

```
                  Life expectancy at birth
-------------------------------------------------------------
      Percentiles      Smallest
 1%           54             54
 5%           65             62
10%           67             64       Obs                  68
25%           69             65       Sum of Wgt.          68

50%           73                      Mean           72.27941
                        Largest       Std. Dev.      4.715315
75%           76             79
90%           78             79       Variance        22.2342
95%           79             79       Skewness      -.9080208
99%           79             79       Kurtosis       4.783111
```

   (b) (8 points) What is a Stata command that will produce the following Stata output?

```
-------------------------------------------------------------------------------
-> region = Eur & C.Asia

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        lexp |        44    73.06818    4.150639         65         79

-------------------------------------------------------------------------------
-> region = N.A.

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        lexp |        14    71.21429    6.411271         54         79

-------------------------------------------------------------------------------
-> region = S.A.
```

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        lexp |        10        70.3    3.831159         62         75
```

2. (8 points) In one year, the average stock price of Apple Inc. was \$650 with the standard deviation equal to \$40. Using the empirical rule, it can be estimated that approximately 95 % of the stock price of Apple Inc. will be in what interval?

3. (8 points) Suppose that $P(A) = 0.30$, $P(B|A) = 0.40$, and $P(\bar{B}|\bar{A}) = 0.60$, where $\bar{A}$ and $\bar{B}$ are complement of A and B, respectively. What is the probability of $P(\bar{A}|B)$?

4. Multiple Choice Questions:

   (a) In a recent survey of consumer confidence, 160 respondents were classified by their level of education. The results of the survey are presented below.

		Education		
		High School	Some College	College
Confidence	Medium	13	17	15
	Fair	27	22	13
	High	32	14	7

   i. (8 points) What proportion of respondents had at least some college education and had high confidence?

   A) 0.131 B) 0.242 C) 0.558 D) 0.175

   ii. (8 points) Are the events "had a college education" and "had high confidence" statistically independent?

   A) Yes. B) No. C) There is not sufficient information to determine.

   (b) (8 point) If A and B are independent events with $P(A) = 0.20$ and $P(B) = 0.60$, then $P(A|B)$ is:

   A) 0.20 B) 0.60 C) 0.40 D) 0.80 E) None of the above.

   (c) (8 point) If X and Y are independent random variables, which of the following identities is generally false?

   A) $Cov(X, X+Y) = Var(X)+E[XY]$. B) $Cov(X, Y) = 0$. C) $E(X+Y) = E(X)+E(Y)$.
   D) $Var(X + Y) = Var(X) + Var(Y)$.

5. (8 points) Consider the joint probability distribution of (X, Y):

		X	
		0	1
Y	0	0.1	0.3
	1	0.4	0.2

Compute the variance of $W = X + 2Y$.

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 $p_{ij}^{X,Y} = P(X = x_i, Y = y_j)$ 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 (Please use the summation operator in the proof):

   (a) (10 points) $Cov(a_1 + b_1 X, a_2 + b_2 Y) = b_1 b_2 E[XY] - b_1 b_2 E[X]E[Y]$, where a_1, a_2, b_1, and b_2 are some constant.

   (b) (10 points) $Var(aX - bY) = a^2 Var(X) + b^2 Var(Y) - 2abCov(X, Y)$, where a and b are some constant.

7. (8 points) True or False Question. Please answer "True" or "False." No explanation necessary. Consider two sequence of numbers: $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$. Is it true that the following equation holds in general?

$$\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}.$$