

## Midterm Exam

1. Answer the following questions regarding Stata command.
- (a) (8 points) The data set contains variables “salary” and “roe”. What is a Stata command that will reproduce the following Stata output? (Please write a whole line of command that produces the output.)

Variable	Obs	Mean	Std. Dev.	Min	Max
salary	209	1281.12	1372.345	223	14822
roe	209	17.18421	8.518509	.5	56.3

- (b) (8 points) What is a Stata command that can be used for computing correlation coefficient between “salary” and “roe”?
2. (8 points) In one year, earnings growth of the 500 largest U.S. corporations average 9.2 %; the standard deviation was 3.5%. Using the empirical rule, it can be estimated that approximately 95 % of these earnings growth figures will be in what interval?
3. (8 points) If  $X$  and  $Y$  are random variables with  $\text{Var}(X) = 7.5$ ,  $\text{Var}(Y) = 6$  and  $\text{Cov}(X,Y) = 4$ , what is the value of  $\text{Var}(2X+3Y)$ ?
4. (8 points) Suppose that  $P(A) = 0.40$ ,  $P(B|A) = 0.60$ , and  $P(B|\bar{A}) = 0.70$ , where  $\bar{A}$  and  $\bar{B}$  are complement of  $A$  and  $B$ , respectively. What is the probability of  $P(\bar{A}|\bar{B})$ ?
5. Multiple Choice Questions: A recent marketing survey related consumers’ awareness of a new marketing campaign with their rating of the product. Consumers rated their awareness as low, medium, or high, and rated the product as poor, fair, or good. The results are presented below.

		Awareness		
		Low	Medium	Hight
Rating	Poor	0.10	0.15	0.07
	Fair	0.06	0.11	0.06
	Good	0.07	0.11	0.27

- (a) (6 points) What is the probability that a consumer had low awareness?  
 A) 0.10    B) 0.14    C) 0.23    D) 0.07
- (b) (6 points) What is the probability that a consumer who ranked the product as fair had a high awareness of the ad campaign?  
 A) 0.06    B) 0.26    C) 0.23    D) 0.40
- (c) (6 points) What is the probability that a consumer who had high awareness of the ad campaign ranked the product as good?  
 A) 0.675    B) 0.385    C) 0.775    D) 0.325

6. In a recent survey of 300 teenagers, 62% of the teenagers indicated that they had been to a movie within the past month. 75% of those teenagers who had seen a movie also had gone out to dinner in the past month, while only 64% of the teenagers who had not seen a movie had been out to dinner in the past month. Define the random variables as follows:

$X = 1$  if teenager had been to movie;  $X = 0$  otherwise

$Y = 1$  if teenager had been out to dinner;  $Y = 0$  otherwise

- (a) (14 points) Find the joint probability function of  $X$  and  $Y$ , the marginal distribution of  $X$ , and the marginal distribution of  $Y$ . [Hint: Use the following table and complete empty cells.]

	$X = 0$	$X = 1$	Marginal Dist. of $Y$
$Y = 0$			
$Y = 1$			
Marginal Dist. of $X$		0.62	1.00

- (b) (6 points) Find the conditional probability function of  $X$ , given  $Y = 1$ .
- (c) (6 points) Find the covariance between  $X$  and  $Y$ .

7. 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 PMF 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 PMF of  $Y$  is  $p_j^Y = P(Y = y_j) = \sum_{i=1}^n p_{ij}^{X,Y}$  for  $i = 1, \dots, n$ . Prove the following results for genera  $n$  and  $m$ :

- (a) (8 points)  $Cov(X, Y) = E[XY] - E[X]E[Y]$ .
- (b) (8 points)  $Var(X - Y) = Var(X) + Var(Y) - 2Cov(X, Y)$ .