Midterm Examination II

Econ 103, Statistics for Economists

March 25, 2013

You will have 70 minutes to complete this exam. Graphing calculators, notes, and textbooks are not permitted.

I pledge that, in taking and preparing for this exam, I have abided by the University of Pennsylvania's Code of Academic Integrity. I am aware that any violations of the code will result in a failing grade for this course.

Name: ______________________________

Student ID #: ______________________________

Signature: ______________________________

Question:	1	2	3	4	5	6	7	8	Total
Points:	20	10	10	15	15	25	25	20	140
Score:									

Instructions: Answer all questions in the space provided, continuing on the back of the page if you run out of space. Show your work for full credit but be aware that writing down irrelevant information will not gain you points. Be sure to sign the academic integrity statement above and to write your name and student ID number on *each page* in the space provided. Make sure that you have all pages of the exam before starting.

Warning: If you continue writing after we call time, even if this is only to fill in your name, fifteen points will be deducted from your final score. In addition, one point will be deducted for each page on which you do not write your name and student ID.

1. Suppose that X is a random variable with support $\{1, 2\}$ and Y is a random variable with support $\{0, 1\}$ where X and Y have the following joint distribution:

$$\begin{aligned} p_{XY}(1,0) &= 0.25 \\ p_{XY}(1,1) &= 0.25 \\ p_{XY}(2,0) &= 0.25 \\ p_{XY}(2,1) &= 0.25 \end{aligned}$$

(a) (4 points) Express the joint pmf of X and Y in a 2×2 table.

(b) (4 points) Using the table, calculate the marginal pmfs of X and Y.

(c) (6 points) Calculate the conditional pmf of $Y|X = 1$.

(d) (6 points) What is the covariance between X and Y?

2. Suppose that $X_1 \sim \text{Bernoulli}(p_1)$, $X_2 \sim \text{Bernoulli}(p_2)$ and $X_3 \sim \text{Bernoulli}(p_3)$ where X_1, X_2 and X_3 are mutually independent. Let Y be a random variable that takes on the value one if and only if a *majority* of the random variables X_1, X_2, X_3 take on the value one. In any other situation Y takes on the value zero.

   (a) (2 points) Are X_1, X_2, X_3 iid? Why or why not?

   (b) (2 points) What kind of random variable is Y? You don't need to give the values of any associated parameters: just name the random variable and explain briefly.

   (c) (2 points) Suppose that $p_1 = p_2 = p_3 = p$. In this case what kind of random variable is $X_1 + X_2 + X_3$?

   (d) (4 points) Using your answer to (c), calculate $P(Y = 1)$ for the case in which $p_1 = p_2 = p_3 = p$.

3. This question concerns the Beta(a, b) random variable, a continuous random variable that we did not study in class. The Beta random variable has support $[0, 1]$ and is defined by two parameters a and b, both of which are greater than zero. Its pdf is given by the following expression:
$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}$$
where $B(a, b)$ is the so-called "Beta Function" evaluated at (a, b).

   (a) (6 points) The Beta$(2, 2)$ random variable has pdf $f(x) = 6x(1 - x)$. Calculate its variance using the shortcut rule.

   (b) (2 points) In order for $f(x)$ to be a valid pdf, what must the formula for the Beta Function, $B(a, b)$, be? You do not need to evaluate or simplify any integrals in your answer: simply give an expression for $B(a, b)$.

   (c) (2 points) What is the pdf of the Beta$(1, 1)$ random variable?

4. Let $X \sim N(\mu_X, \sigma_X^2)$ independent of $\epsilon \sim N(0, \sigma_\epsilon^2)$ and define $Y = a + bX + \epsilon$ where a, b are constants.

   (a) (3 points) Calculate the mean of Y.

   (b) (3 points) Calculate the variance of Y.

   (c) (3 points) What kind of random variable is Y? Be sure to give the values of any parameters needed to define its distribution.

   (d) (6 points) Calculate $Cov(X, Y)$.

5. Let $X_1, X_2 \sim$ iid with cumulative distribution function F and define $X^* = \max\{X_1, X_2\}$. Let c be a constant.

   (a) (2 points) In terms of F, what is $P(X_1 \leq c)$? What is $P(X_2 \leq c)$?

   (b) (3 points) In terms of F what is $P(X_1 \leq c \cap X_2 \leq c)$?

(c) (3 points) Explain why $P(X^* \leq c) = P(X_1 \leq c \cap X_2 \leq c)$.

(d) (2 points) Using your answer to (c), what is the CDF of X^*?

(e) (5 points) Now suppose that $X_1, X_2 \sim$ iid Uniform$(0,1)$. Using your answer to part (d), calculate the *probability density function* of X^*.

6. Suppose that $X_1, \ldots, X_n \sim$ iid with mean μ and variance σ^2. This question considers two estimators of μ: the sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $\widehat{\mu} = 0$. To be clear: the estimator $\widehat{\mu}$ is a rule that *ignores* the sample data and simply assumes the population mean is zero. You can think of $\widehat{\mu}$ as a *degenerate random variable*: its support is $\{0\}$ and $P(\widehat{\mu} = 1) = 1$.

   (a) (4 points) Calculate the MSE of $\bar{X}_n$.

   (b) (5 points) Calculate the MSE of $\widehat{\mu}$.

(c) (5 points) Now define the estimator $\tilde{\mu} = \omega\widehat{\mu} + (1-\omega)\bar{X}_n$ where ω is a constant between zero and one. If $\omega = 1$, then $\tilde{\mu} = \widehat{\mu}$; if $\omega = 0$, then $\tilde{\mu} = \bar{X}_n$. When $0 < \omega < 1$, $\tilde{\mu}$ can be thought of as a "compromise" between $\bar{X}_n$ and $\widehat{\mu}$. Calculate the MSE of $\tilde{\mu}$.

(d) (5 points) Using your answer to part (c) calculate ω^*, the value of ω that minimizes $MSE(\tilde{\mu})$. Your answer should be given in terms of σ^2, n and μ^2. For full credit, check the second order condition to make sure you have found a minimum.

(e) (6 points) Explain how ω^* changes with n, $|\mu|$ and σ^2. Note that you can answer this without taking any derivatives by rearranging your answer to part (d).

Name: ____________________ Student ID #: ____________________

7. Suppose that $X_1, \ldots, X_n \sim$ iid $N(\mu, \sigma^2)$. You do not have to explain your answers.

   (a) (5 points) What is the distribution of $\sum_{i=1}^n X_i$?

   (b) (5 points) What is the distribution of $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$?

   (c) (5 points) Let $Y = \sum_{i=1}^n (X_i - \mu)^2$. What is the distribution of Y/σ^2?

   (d) (5 points) Let $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$. What is the distribution of $(n-1)S^2/\sigma^2$?

   (e) (5 points) Let

   $$Z = \frac{\bar{X}_n - \mu}{S/\sqrt{n}}$$

   where $S = \sqrt{S^2}$ and S^2 is as defined in (d). What is the distribution of Z?

8. Using R, I generated a random sample of size 16 from a normal population with mean μ and variance σ^2. The sample mean was approximately -0.5 and the sample variance was approximately 6.

   (a) (5 points) Construct a 95% confidence interval for μ. Your answer may include relevant R commands.

(b) (5 points) If I told you that I used a population variance of 4 to generate my random sample, how would your answer to (a) change? Your final answer should *not* include any R commands.

(c) (5 points) Based on your answer to (b), would you be surprised to learn that the population mean was -1? Why or why not?

(d) (5 points) Construct a 90% confidence interval for the population variance. Your answer may include relevant R commands.