Midterm Examination II

Econ 103, Statistics for Economists

November 11th, 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	Total
Points:	20	15	20	15	30	20	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, twenty-five points will be deducted from your final score. In addition, two points will be deducted for each page on which you do not write your name and student ID.

When asked to identify a random variable on this exan be sure to give any and all parameters of its distribution for full credit.

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 pmf:

$$p_{XY}(1,0) = 0.4 \quad p_{XY}(1,1) = 0.3$$
$$p_{XY}(2,0) = 0.3 \quad p_{XY}(2,1) = 0$$

   (a) (2 points) Express the joint probability mass function (pmf) in a 2×2 table.

   (b) (3 points) Using the table, calculate the marginal pmfs of X and Y.

   (c) (5 points) Calculate the conditional pmfs of $Y|X = 1$ and $Y|X = 2$.

   (d) (3 points) Calculate $E[Y|X = 1]$ and $E[Y|X = 2]$.

Name: ______________________ Student ID #: ______________________

(e) (4 points) Calculate the covariance between X and Y.

(f) (3 points) Are X and Y independent? Explain briefly.

2. The random variables X_1 and X_2 correspond to the annual returns of Stock 1 and Stock 2. Suppose that $E[X_1] = 0.1$, $E[X_2] = 0.3$, $Var(X_1) = Var(X_2) = 1$, and $\rho = Corr(X_1, X_2)$. A portfolio $\Pi(\omega)$ is defined by the proportion ω of Stock 1 that it contains. That is, $\Pi(\omega) = \omega X_1 + (1 - \omega)X_2$ where $0 \leq \omega \leq 1$.

   (a) (3 points) What value of ω gives a portfolio with expected return 0.2?

   (b) (6 points) Suppose that $\omega = 1/4$. In terms of ρ, what is the portfolio variance?

(c) (3 points) Again, suppose that $\omega = 1/4$. What are the maximum and minimum values of the portfolio variance? What are the corresponding values of ρ?

(d) (3 points) If we assume that variance is a reasonable measure of risk, what does your answer to part (c) suggest about the benefits of constructing a portfolio rather than holding only one stock? Explain briefly.

3. Let Y be a continuous random variable with support $[0, 1]$ and pdf $f(y) = Cy^3(1 - y)$.

(a) (5 points) Calculate the value of the constant C in the pdf of Y.

(b) (5 points) Calculate the CDF $F(y_0)$ of Y.

(c) (5 points) Calculate the expected value of Y.

(d) (5 points) Calculate the variance of Y using the shortcut formula.

4. Let $X_1, X_2, \ldots, X_k \sim$ iid $N(\mu_X, \sigma^2)$ independent of $Y_1, Y_2, \ldots, Y_m \sim$ iid $N(\mu_Y, \sigma^2)$ and define $\bar{X}_k = (\sum_{i=1}^{k} X_i)/k$, $\bar{Y}_m = (\sum_{i=1}^{m} Y_i)/m$, $\widehat{\mu} = (\bar{X}_k + \bar{Y}_m)/2$.

   (a) (2 points) What is the sampling distribution of $\bar{X}_k$?

   (b) (2 points) What is the sampling distribution of $\bar{Y}_m$?

   (c) (5 points) Suppose you wanted to estimate $\mu = (\mu_X + \mu_Y)/2$. This is the *midpoint* of the two means μ_X and μ_Y. Show that $\widehat{\mu}$ is an unbiased estimator of μ.

(d) (6 points) What is the sampling distribution of $\widehat{\mu}$?

5. Sara is carrying out a poll to estimate the proportion of Penn Undergraduates who favor legalizing marijuana. Let $p \in [0, 1]$ denote the true, unknown proportion. Sara polls a random sample of n Penn students and counts the total number T who favor legalizing marijuana. To estimate p, she uses $\widehat{p} = (T + 2)/(n + 4)$.

   (a) (3 points) Under random sampling T is a random variable. What kind?

   (b) (3 points) Write down $E[T]$.

   (c) (3 points) Write down $Var(T)$.

   (d) (6 points) Calculate the bias of $\widehat{p}$ and briefly explain the intuition for your result.

(e) (5 points) Calculate $Var(\widehat{p})$.

(f) (5 points) Is $\widehat{p}$ a consistent estimator of p? Explain your answer.

(g) (5 points) Kevin thinks that $\widehat{p}$ is a bad estimator. He tells Sara that she should use $\tilde{p} = T/n$ instead. Briefly argue in favor of Kevin's proposal using what you know about the sampling distributions of $\tilde{p}$ and $\widehat{p}$.

6. This question asks you write R code to make random draws from two distributions related to the normal. You may use any commands you like *except* `rchisq` and `rt`.

   (a) (10 points) Write a function called `my.rchisq` that uses `rnorm` to make a single random draw from a $\chi^2(\nu)$ distribution, where ν is the degrees of freedom. Your function should take a single argument, the degrees of freedom `df`, and return the random draw.

(b) (10 points) Write a function called `my.rt` that uses `rnorm` and `my.rchisq` to make a single random draw from a $t(\nu)$ distribution, where ν is the degrees of freedom. Your function should take a single argument, the degrees of freedom `df`, and return the random draw.

7. Let $X_1, X_2, \ldots, X_n \sim$ iid $N(\mu_X, \sigma_X^2)$ independent of $Y_1, Y_2, \ldots, Y_m \sim$ iid $N(\mu_Y, \sigma_Y^2)$ and define S_X^2 to be the sample variance of the X-observations and S_Y^2 to be the sample variance of the Y-observations.

   (a) (3 points) What is the sampling distribution of $(n-1)S_X^2/\sigma_X^2$? You do not need to explain your answer.

   (b) (5 points) Using your answer to the previous part, derive a $100\times(1-\alpha)\%$ confidence interval for σ_X^2. Express the interval in terms of the appropriate R commands.

Name: ______________________ Student ID #: ______________________

(c) (6 points) What is the sampling distribution of $(S_X^2/\sigma_X^2)/(S_Y^2/\sigma_Y^2)$? Explain.

(d) (6 points) Use your answer to the previous part to propose a procedure for constructing a $(1-\alpha)\times 100\%$ confidence interval for the *ratio of population variances* σ_Y^2/σ_X^2. Express the interval in terms of the appropriate R commands and briefly suggest how we might use it in practice.

Name: ____________________ Student ID #: ____________________