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.

Solution:

		X	
		1	2
Y	0	0.25	0.25
	1	0.25	0.25

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

Solution:

$$\begin{aligned} p_X(1) &= p_{XY}(1,0) + p_{XY}(1,1) = 0.25 + 0.25 = 0.5 \\ p_X(2) &= p_{XY}(2,0) + p_{XY}(2,1) = 0.25 + 0.25 = 0.5 \\ p_Y(0) &= p_{XY}(1,0) + p_{XY}(2,0) = 0.25 + 0.25 = 0.5 \\ p_Y(1) &= p_{XY}(1,1) + p_{XY}(2,1) = 0.25 + 0.25 = 0.5 \end{aligned}$$

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

Solution: The distribution of $Y|X = 1$ is

$$P(Y = 0|X = 1) = \frac{p_{XY}(1,0)}{p_X(1)} = 0.25/0.5 = 0.5$$

$$P(Y = 1|X = 1) = \frac{p_{XY}(1,1)}{p_X(1)} = 0.25/0.5 = 0.5$$

Alternatively, you could simply notice from the table that X and Y are independent so the conditional pmfs are the same as the marginal pmfs.

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

Name: _____ Student ID #: _____

Solution: First, from the marginal distributions, $E[X] = 1.5$ and $E[Y] = 0.5$. Hence $E[X]E[Y] = 0.75$. Second,

$$\begin{aligned} E[XY] &= (0 \cdot 1) \cdot 0.25 + (0 \cdot 2) \cdot 0.25 + (1 \cdot 1) \cdot 0.25 + (1 \cdot 2)0.25 \\ &= 0.25 + 0.5 = 0.75 \end{aligned}$$

Finally $Cov(X, Y) = E[XY] - E[X]E[Y] = 0.75 - 0.75 = 0$. Alternatively, use the fact that independence implies zero covariance to solve this without doing any calculations.

2. Suppose that $X_1 \sim$ Bernoulli(p_1), $X_2 \sim$ Bernoulli(p_2) and $X_3 \sim$ 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?

   Solution: Not necessarily: while the three random variables are independent, they are only identically distributed if $p_1 = p_2 = p_3$.

   (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.

   Solution: Since it can only take on only the values zero and one, Y is a Bernoulli RV.

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

   Solution: In this case $X_1, X_2, X_3 \sim$ iid Bernoulli(p) so their sum is, by definition, a Binomial$(3, p)$ random variable.

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

   Solution: From (c), $X_1 + X_2 + X_3$ is a Binomial$(3, p)$ random variable. Hence its pmf is

$$p(x) = \binom{3}{x} p^x (1-p)^{3-x}$$

By definition, $Y = 1$ when a majority of X_1, X_2, X_3 equal one. Thus we need to calculate the probability that the sum $X_1 + X_2 + X_3$ equals either two or three. Since these are mutually exclusive outcomes, we calculate

$$\begin{aligned} P(Y=1) &= P(X_1+X_2+X_3 \geq 2) = p(2)+p(3) \\ &= \binom{3}{2}p^2(1-p) + \binom{3}{3}p^3 \\ &= 3p^2(1-p)+p^3 \\ &= 3p^2 - 3p^3 + p^3 \\ &= 3p^2 - 2p^3 \end{aligned}$$

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.

Solution: Using the shortcut rule:

$$\begin{aligned} E[X] &= 6\int_0^1 (x^2 - x^3)\,dx = 6\left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Big|_0^1 = 6(1/3 - 1/4) = 1/2 \\ E[X^2] &= 6\int_0^1 (x^3 - x^4)\,dx = 6\left(\frac{x^4}{4} - \frac{x^5}{5}\right)\Big|_0^1 = 6(1/4 - 1/5) = 3/10 \\ Var(X) &= E[X^2] - (E[X])^2 = 3/10 - 1/4 = 1/20 \end{aligned}$$

(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 $B(a, b)$.

Solution: The integral of pdf must equal one over the support. Thus:

$$\int_0^1 \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\,dx = 1$$

Multiplying through by $B(a,b)$, which does not involve x,

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

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

Solution: Substituting $a = b = 1$,

$$f(x) = \frac{x^0(1-x)^0}{B(1,1)} = \frac{1}{B(1,1)}$$

This expression is simply a constant: it does not depend on x. Thus, the only way for it to integrate to one over $[0,1]$ is for $B(1,1)$ to equal one. Alternatively, we could evaluate the integral from part (a):

$$B(1,1) = \int_0^1 1\,dx = 1$$

We see that the Beta$(1,1)$ random variable is the same thing as the Uniform$(0,1)$ 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.

Solution: By the Linearity of Expectation

$$E[Y] = a + bE[X] = a + b\mu_X$$

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

Solution: Since X and ϵ are independent,

$$Var(Y) = b^2 Var(X) + Var(\epsilon) = b^2\sigma_X^2 + \sigma_\epsilon^2$$

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

Solution: Linear combinations of independent normal random variables are themselves normally distributed. Thus, using parts (a) and (b),

$$Y \sim N\left(a + b\mu_X, b^2\sigma_x^2 + \sigma_\epsilon^2\right)$$

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

Solution:

$$\begin{aligned} E[XY] &= E[X(a + bX + \epsilon)] = E[aX] + bE[X^2] + E[X\epsilon] \\ &= a\mu_X + b\left\{Var(X) + (E[X])^2\right\} + \{Cov(X, \epsilon) + E[X]E[\epsilon]\} \\ &= a\mu_X + b(\sigma_X^2 + \mu_X^2) \\ \\ E[X]E[Y] &= \mu_X(a + b\mu_X) = a\mu_X + b\mu_X^2 \\ \\ Cov(X, Y) &= E[XY] - E[X]E[Y] \\ &= a\mu_X + b(\sigma_X^2 + \mu_X^2) - (a\mu_X + b\mu_X^2) = b\sigma_X^2 \end{aligned}$$

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 $P(X_2 \leq c)$?

Solution: By the definition of a CDF and the fact that X_1 and X_2 are identically distributed, $F(c) = P(X_1 \leq c) = P(X_2 \leq c)$.

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

Solution: By independence,

$$P(X_1 \leq c \cap X_2 \leq c) = P(X_1 \leq c)P(X_2 \leq c) = F(c)F(c) = F^2(c)$$

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

Solution: Since X^* is the maximum of X_1 and X_2, the events $X^* \leq c$ and $(X_1 \leq c) \cap (X_2 \leq c)$ are *logically equivalent*. If both X_1 and X_2 are no greater than c then their maximum is no greater than c. Further, if the maximum of X_1 and X_2 is no greater c then it must be the case that both X_1 and X_2 are no greater than c. Since the events are equivalent, their probabilities must be equal: this is the logical equivalence rule from class.

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

Solution: $F^*(x) = F^2(x)$

(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^*.

Solution: For a Uniform(0,1) random variable $F(x) = x$ for $x \in [0, 1]$ while $F(x) = 1$ for $x > 1$ and $F(x) = 0$ for $x < 1$. Hence,

$$F^*(x) = F^2(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

Differentiating, $f^*(x) = 2x$ for $x \in [0, 1]$, zero otherwise.

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 $P(\widehat{\mu} = 0) = 1$.

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

Solution:

$$\begin{aligned} MSE(\bar{X}_n) &= \text{Bias}(\bar{X}_n)^2 + Var(\bar{X}_n) \\ &= \left(E[\bar{X}_n] - \mu\right)^2 + Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\ &= \sigma^2/n \end{aligned}$$

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

Solution: Estimators are random variables under random sampling because the *data* are random. Since this estimator completely ignores the data, it is a constant, hence its variance is zero. Its bias is:

$$E[\widehat{\mu}] - \mu = 0 - \mu = -\mu$$

Thus, since MSE is squared bias plus variance:

$$MSE(\widehat{\mu}) = \mu^2$$

(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 $\tilde{\mu}$.

Solution:

$$\begin{aligned} \text{Bias}(\tilde{\mu}) &= E[\tilde{\mu}] - \mu = E[\omega\widehat{\mu} + (1-\omega)\bar{X}_n] - \mu \\ &= (1-\omega)\mu - \mu = -\omega\mu \end{aligned}$$

$$\begin{aligned} Var(\tilde{\mu}) &= Var\left[\omega\widehat{\mu} + (1-\omega)\bar{X}_n\right] = Var\left[\omega \times 0 + (1-\omega)\bar{X}_n\right] \\ &= (1-\omega)^2 Var(\bar{X}_n) = (1-\omega)^2\sigma^2/n \end{aligned}$$

Therefore,

$$MSE(\tilde{\mu}) = \omega^2\mu^2 + (1-\omega)^2\sigma^2/n$$

(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.

Solution: The first order condition is

$$\begin{aligned} 2\omega\mu^2 - 2(1-\omega)\sigma^2/n &= 0 \\ \omega\mu^2 - \sigma^2/n + \omega\sigma^2/n &= 0 \\ \omega(\mu^2 + \sigma^2/n) &= \sigma^2/n \end{aligned}$$

Hence,

$$\omega^* = \frac{\sigma^2/n}{\mu^2 + \sigma^2/n}$$

The second order condition is $2\mu^2 + 2\sigma^2/n > 0$ which holds because $\sigma^2 > 0$.

(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).

Solution: First, rewrite ω^* as follows:

$$\omega^* = \frac{\sigma^2/n}{\mu^2 + \sigma^2/n} = \frac{\sigma^2}{n\mu^2 + \sigma^2} = \frac{1}{\frac{n\mu^2}{\sigma^2} + 1}$$

As n increases, ω^* falls: because the variance of $\bar{X}_n$ decreases with n we should give $\bar{X}_n$ more weight when n is larger. As σ^2 increases, ω^* increases: because the variance of $\bar{X}_n$ increases with σ^2 we should give $\bar{X}_n$ less weight when σ^2 is larger. As $|\mu|$ increases, ω^* falls: because the bias of $\widehat{\mu}$ increases with $|\mu|$, we should give $\bar{X}_n$ more weight when $|\mu|$ is larger.

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 $\sum_{i=1}^n X_i$?

Solution: $N(n\mu, n\sigma^2)$

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

Solution: $N(\mu, \sigma^2/n)$

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

Solution: $\chi^2(n)$

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

Solution: $\chi^2(n-1)$

(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?

Name: _____ Student ID #: _____

Solution: $t(n-1)$

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.

   Solution: The interval should take the form:

   $$\bar{X}_n \pm \texttt{qt}(1-\alpha/2, \texttt{df} = n-1) \times \frac{S}{\sqrt{n}}$$

   which, in this case, gives $-0.5 \pm \texttt{qt}(0.975, \texttt{df} = 15) \times \sqrt{6}/4$

   (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.

   Solution: The interval takes the form $\bar{X}_n \pm \texttt{qnorm}(1-\alpha/2) \times \sigma/\sqrt{n}$. We know, however, that $\texttt{qnorm}(0.975) \approx 2$. Thus, the interval is -0.5 ± 1, i.e. $(-1.5, 0.5)$.

   (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?

   Solution: This would not be surprising: -1 is well within the 95% CI for μ.

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

   Solution: The interval should take the form

   $$\left[\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right]$$

   $$\begin{aligned} a &= \texttt{qchisq}(\alpha/2, \texttt{df = n - 1}) \\ b &= \texttt{qchisq}(1-\alpha/2, \texttt{df = n - 1}) \end{aligned}$$

   In this case $(n-1)S^2 = 15 * 6 = 90$.

Name: _____ Student ID #: _____