

MIDTERM EXAMINATION #2  
ECON 103, STATISTICS FOR ECONOMISTS

NOVEMBER 3RD, 2014

**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:	24	16	20	20	20	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.

1. Mark each statement as True or False. If you mark a statement as False, provide a brief explanation. If you mark a statement as True, no explanation is needed.

- (a) (4 points) If  $Cov(X, Y) = 0$  then  $E[XY] = E[X]E[Y]$

**Solution:** TRUE

- (b) (4 points) Unlike a probability mass function, a probability density function *can* take on values greater than one.

**Solution:** TRUE

- (c) (4 points) If  $X$  and  $Y$  are uncorrelated then  $Var(X - Y) = Var(X) - Var(Y)$

**Solution:** FALSE: the variance of the difference equals  $Var(X) + Var(Y)$  since the minus one is squared when brought in front of the variance operator.

- (d) (4 points) The pdf  $f(x)$  of a continuous random variable  $X$  gives  $P(X = x)$ .

**Solution:** FALSE: the probability that a continuous random variable takes on *any particular value* is zero. Only regions with positive area have positive probability.

- (e) (4 points) For any random variable  $X$  and any function  $g$ ,  $E[g(X)] = g(E[X])$ .

**Solution:** FALSE: this does not hold in general although it is true for linear functions.

- (f) (4 points) Even if  $\hat{\theta}$  is a biased estimator of  $\theta_0$ , it can *still* be consistent for  $\theta_0$ .

**Solution:** TRUE

2. Unless otherwise specified, answer each part with a *single line* of R code.

- (a) (4 points) Calculate the median of an  $F$  random variable with numerator degrees of freedom 3 and denominator degrees of freedom 8.

**Solution:** `qf(0.5, df1 = 3, df2 = 8)`

- (b) (4 points) Calculate the probability that a Binomial  $n = 20$ ,  $p = 2/3$  random variable takes on a value strictly greater than 15.

```
Solution: 1 - pbinom(15, size = 20, prob = 2/3)
```

- (c) (4 points) Make 100 iid draws from a  $t$  distribution with 5 degrees of freedom.

```
Solution: rt(100, df = 5)
```

- (d) (4 points) Write code to plot the pdf of a standard normal random variable between -3 and 3. You may use more than one line of code in your answer to this part.

```
Solution: Many possible solutions, but the general idea is as follows
```

```
x <- seq(-3, 3, 0.01)
y <- dnorm(x)
plot(x,y, type = 'l')
```

3. (20 points) Write an R function called `CI.normal.mean` that constructs a 90% confidence interval for the mean of a normal population with *known* population variance. Your function should take two arguments: `x` is a vector of sample data, assumed to be a sequence of iid draws from a normal population, and `s` is the population standard deviation, assumed known. Your function should return a vector with two elements: the first is the lower confidence limit and the second is the upper confidence limit.

```
Solution: Various possibilities, but the basic idea is as follows:
```

```
CI.normal.mean <- function(x, s){
n <- length(x)
ME <- qnorm(0.95) * s / sqrt(n)
LCL <- mean(x) - ME
UCL <- mean(x) + ME
return(c(LCL, UCL))
}
```

4. Rodrigo has a bowl containing 10 balls: five of them are red and the rest are blue. Rossa wants to know the fraction of red balls in the bowl so he draws four balls at random with replacement from the bowl.
- (a) (5 points) Let  $X$  be the number of red balls that Rossa draws. What kind of random variable is  $X$ ? Write down its support set, its pmf, and the values of its parameters.

**Solution:** Binomial(4, 1/2) with support set  $\{0, 1, 2, 3, 4\}$  and pmf

$$p(x) = \binom{4}{x} (1/2)^4$$

- (b) (5 points) To estimate the fraction of red balls in the bowl, Rossa divides  $X$ , from the preceding part, by four. Is this estimator unbiased? What is its variance?

**Solution:** We use the fact that the expected value of a Binomial RV is  $np$  while the variance is  $np(1-p)$ . In this case,  $n = 4$  and  $p = 1/2$  so the expected value of  $X$  is 2 while the variance is 1. Hence,

$$E[X/4] = 2/4 = 1/2$$

so this estimator is unbiased and its variance is

$$Var(X/4) = (1/16) \times 1 = 1/16$$

- (c) (5 points) What is the probability that Rossa's estimator from the preceding part will *exactly equal* the true fraction of red balls in the bowl?

**Solution:** We simply need to evaluate  $P(X = 2)$  since  $X = 2$  if and only if  $X/4 = 0.5$ . Using the pmf from above,  $p(2) = \binom{4}{2} (1/2)^4 = 6/16 = 0.375$ .

- (d) (5 points) Now suppose that Rossa decides to make his random draws *without replacement*. Will your answer to the preceding part change? If so, how?

**Solution:** Again, we need to calculate the probability that  $X = 2$ . Because the draws are made with replacement, however, we cannot use the Binomial pdf for the calculations since the draws are no longer independent, nor identically distributed. Since Rossa is still choosing at random, every possible *group* of four balls is equally likely to be drawn. The total number of groups of 4 objects drawn from a set of 10 is  $\binom{10}{4} = 210$ . This is the denominator of our desired probability. For the numerator, we need to count the total number of ways to choose exactly two red balls and two blue balls. There are  $\binom{5}{2} = 10$  ways to choose *which* two red balls are in the group and the same number of ways to choose which two blue balls are in the group. Thus,

$$P(X = 2) = \frac{\binom{5}{2} \binom{5}{2}}{\binom{10}{4}} = 100/210 \approx 0.48$$

So the probability has increased substantially.

5. Let  $X$  and  $Y$  be independent, normally distributed random variables representing the returns of two stocks such that  $\mu_x = \mu_y = 1$  and  $\sigma_x^2 = \sigma_y^2 = 2$ . A portfolio  $\Pi(\omega)$  is a linear combination of the form  $\omega X + (1 - \omega)Y$  where  $\omega \in [0, 1]$  is the fraction of your total funds that are invested in  $X$ .
- (a) (10 points) If you wish to construct the portfolio with the lowest possible variance, what value of  $\omega$  should you choose? Prove your answer.

**Solution:** Since the asset returns are uncorrelated, the portfolio variance is simply  $\omega^2\sigma_x^2 + (1 - \omega)^2\sigma_y^2 = 2\omega^2 + 2(1 - \omega)^2 = 4\omega^2 - 4\omega + 2$ . This is a globally convex function so it has a unique minimum. The first order condition is  $8\omega - 4 = 0$  so the minimizer is  $\omega^* = 1/2$ . The minimum variance portfolio is 50% asset  $X$  and 50% asset  $Y$ .

- (b) (5 points) Calculate the expected value and the variance of the portfolio from the preceding part.

**Solution:** The expected value is  $1/2 \times \mu_x + 1/2 \times \mu_y = 1$  and the variance is  $(1/2)^2 \times \sigma_x^2 + (1/2)^2 \sigma_y^2 = 1$ .

- (c) (5 points) Approximately what is the probability that the minimum variance portfolio will make a *negative* return?

**Solution:** Since the individual assets are normally distributed and the portfolio is a linear combination of them, it too is normally distributed. We already worked out its mean and variance in the preceding part: both are one. The probability that a normal random variable takes on a value within one standard deviation of its mean is approximately 0.68. By symmetry, the probability that it takes on a value more than one standard deviation *below* its mean is approximately 0.16. Therefore, the minimum variance portfolio has approximately a 16% chance of earning a negative return.

6. Let  $X$  be a continuous RV with CDF  $F(x_0) = (x_0^3 + 1)/2$  and support set  $[-1, 1]$ .

- (a) (5 points) Calculate the probability that  $X$  takes on a value *outside*  $(-0.5, 0.5)$ .

**Solution:** First calculate the probability that  $X$  takes on a value *inside* the range, namely:

$$F(1/2) - F(-1/2) = (1/8 + 1)/2 - (-1/8 + 1)/2 = 9/16 - 7/16 = 1/8$$

Hence, by the complement rule, the desired probability is  $7/8$ .

- (b) (5 points) Calculate the pdf of  $X$ .

**Solution:**

$$F'(x) = 3x^2/2$$

- (c) (5 points) Calculate the expected value of  $X$ .

**Solution:** By symmetry, the expected value is zero. By brute-force integration:

$$E[X] = \frac{3}{2} \int_{-1}^1 x^3 dx = \frac{3}{2} \left( \frac{x^4}{4} \right) \Big|_{-1}^1 = 0$$

- (d) (5 points) Calculate the quantile function of  $X$ .

**Solution:** This is simply the inverse of the CDF:

$$Q(p) = F^{-1}(p) = (2p - 1)^{1/3}$$

7. Let  $Y_1, \dots, Y_9 \sim \text{iid } N(\mu = 1, \sigma^2 = 4)$ ,  $X = \frac{1}{9} \sum_{i=1}^9 Y_i$  and  $Z = \frac{1}{8} \sum_{i=1}^9 (Y_i - X)^2$ . Specify the *precise* distribution of each of the random variables listed below, including the values of any and all relevant parameters.

- (a) (5 points)  $X$

**Solution:** Since  $X$  is the sample mean of the  $Y_i$ , it follows a normal distribution with mean  $\mu = 1$  variance  $\sigma^2/n = 4/9$ .

- (b) (5 points)  $\frac{3}{2}(X - 1)$

**Solution:** Since  $X$  is normally distributed with mean one and standard deviation  $2/3$  this random variable, which simply subtracts the mean and divides by

the standard deviation, is  $N(0, 1)$ .

(c) (5 points)  $2Z$

**Solution:**  $Z$  is simply the sample variance of the  $Y_i$ . Hence  $\frac{n-1}{\sigma^2}Z \sim \chi^2(n-1)$ . Here,  $n-1 = 8$  and  $\sigma^2 = 4$  so  $2Z \sim \chi^2(8)$ .

(d) (5 points)  $\frac{3(X-1)}{\sqrt{Z}}$

**Solution:** This is simply  $\sqrt{n}(\bar{Y} - \mu)/S$  which we know from class follows a  $t(n-1)$  distribution. Here  $n-1 = 8$  so it is a  $t$  distribution with 8 degrees of freedom.