Economics 250 — Final Exam   **7 December 2013**

Instructions:

- The exam is 3 hours in length. There are 183 points in total. Allocate your time accordingly.

- Put your name and student number on each answer booklet used.

- You may use a hand calculator: either the standard Casio fx-991MS, or a university-approved calculator with a gold or blue sticker. No red-sticker calculators or other aids are allowed.

- Formulas and tables are printed at the end of the question papers.

- This midterm consists of 13 pages in all: this cover sheet, 5 question pages, 4 formula pages, and 3 statistic table pages. Please ensure you have all questions/sheets!

- This exam is divided into two sections:

  – Section A (page 2, worth 33 marks) consists of 11 very short questions requiring only a small calculation, value lookup, or a couple of words.

  – Section B (pages 3–6, worth 150 marks) consists of 7 longer questions with multiple parts. Show your work: part marks cannot be awarded for wrong answers without calculations.

- Answer all questions. The value of each question is shown in the exam.

- Proctors are unable to respond to queries about the interpretation of exam questions. Do your best to answer the exam questions as they are written.

- This material is copyrighted and is for the sole use of students registered in Economics 250 and writing this exam. This material shall not be distributed or disseminated. Failure to abide by these conditions is a breach of copyright and may also constitute a breach of academic integrity under the University Senate's Academic Integrity Policy Statement.

- Answers will be posted to the course website on the morning after the exam.

- Good luck!

Section A: Very short questions [33 points]

The following questions require only a couple of words, a lookup of a numerical value, or a small calculation. Each question is worth 3 marks. No part marks will be awarded for this section.

1. For each of the two following figures, indicate whether the depicted distribution is left-skewed, right-skewed, or symmetric.

Figure 1: Histogram of income

Figure 2: Histogram of log(income)
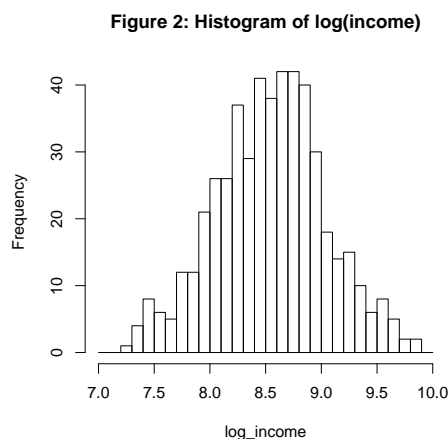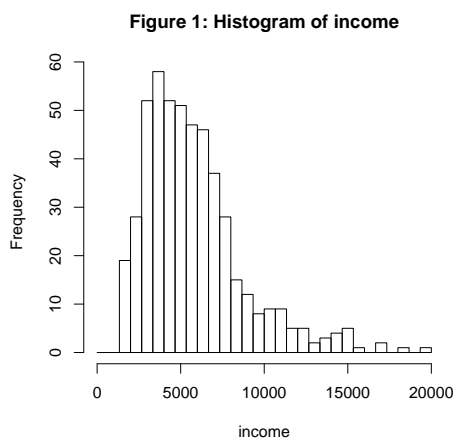
2. What is the critical z^* value you need to use when calculating a 95% confidence interval, assuming that you have calculated $\overline{x}$ and know σ^2?

3. If you obtained $\overline{x}$ from $n = 16$ observations and calculated s^2 (but don't know σ^2), what is the critical t^* value needed to calculate a 99% confidence interval?

4. Suppose you have two independent random samples from the same population with a known standard deviation. The first sample is of size $n = 100$, the second sample is 4 times as large ($n = 400$). How much larger will the standard deviation of the first sample mean be as a multiple of the standard deviation of the second sample mean?

5. If A and B are independent events, $P(A) = 0.4$, and $P(B) = 0.5$, what is $P(A \cap B)$?

6. If A and B are disjoint events, $P(A \cup B) = 0.7$, and $P(A) = 0.5$, what is $P(B)$?

7. If 62% of Queen's students are female, 15% of students own a car, and 6% of students are female *and* own a car, what is the probability that a randomly selected Queen's undergraduate student is female *or* owns a car?

8. If 57% of Acadia University students are female, and 15% of female Acadia students own cars, what is the probability of a randomly selected Acadia student being both female and a car owner?

9. What is the probability that a weighted coin that produces heads 55% of the time will produce exactly 2 heads in 4 flips?

10. If X is distributed $N(2, 2)$, what is $P(X < 0)$?

11. If X is distributed $U(2, 10)$, what is $P(X > 4)$?

Section B: Longer questions [150 points]

1. **[20]** The following is a sample of IQ scores of 8 fifth-grade students:

124	106	124	115	133	116	102	124

   Find the following values:

   a) The sample mean.

   b) The sample variance.

   c) The sample standard deviation.

   d) The sample median.

   e) The mode of the sample.

2. **[20]** The following are the midterm grades of a random sample of 5 students in this semester's Economics 250 course. X denotes the grade on midterm 1, Y denotes the grade on midterm 2. Both values are out of 75.

X	37	63	59	62	46
Y	26	55	58	72	53

   Using the skills you demonstrated in the previous question, you calculate the following:

   $$\overline{x} = 53.4, \quad s_x^2 = 130.3, \quad s_x = 11.41, \quad \overline{y} = 52.8, \quad s_y^2 = 279.7, \quad s_y = 16.72$$

   a) Calculate s_{xy}, the sample covariance of X and Y.

   b) Calculate r_{xy}, the sample correlation of X and Y. Does this value imply a strong or a weak relationship between midterm scores?

   Now suppose that you learn that the true population parameters of X and Y are

   $$\mu_x = 53 \qquad \sigma_x^2 = 163$$
   $$\mu_y = 51 \qquad \sigma_y^2 = 261$$
   $$\sigma_{xy} = 134$$

   c) Find the population correlation, ρ_{xy}.

   Suppose that the random variable $Z = 0.7X + 0.3Y$ is constructed (i.e. Z is a weighted mean that puts 70% of the weight on X and 30% of the weight on Y).

   d) Find μ_z, the mean of Z.

   e) Find σ_z^2, the variance of Z.

3. **[25]** The U.S. Department of Agriculture collects the following data on wheat prices (in dollars per bushel) in two months:

Month	n	$\overline{x}$	s
September	45	$7.43	$0.24
July	90	$7.16	$0.27

The department wishes to use the data to determine whether average wheat prices are different between the two months. Let μ_1 denote the mean price in September, and μ_2 denote the mean price in July.

a) Write down the null and alternative hypotheses for this test.

b) Calculate the t statistic for this test *without* assuming that the two samples have the same standard deviation.

c) Calculate the range in which the p value of this test lies. Your answer should be a range that is as precise as possible, given the available statistical table. You may use an appropriate approximation for the degrees of freedom.

d) Calculate a 95% confidence interval for $\mu_1 - \mu_2$, the price difference between the two months.

4. **[25]** Suppose that there is a lottery which costs $2 to play. If you play, you win a prize with probability 0.3.

a) Suppose you and seven of your friends buy lottery tickets (so there are 8 tickets purchased in all). What is the probability that 2 or fewer of you win a prize?

b) Use your answer to part *a)* to calculate the probability of 3 or more of the group winning a prize.

c) Suppose that you convince 80 people to buy lottery tickets. What is the (approximate) probability that at least 30 of the 80 people win a prize?

The lottery actually pays three possible prizes: with probability 0.01, you win the BIG PRIZE of $50, with probability 0.04 you win the Medium Prize of $10, and with probability 0.25 you win the small prize of $3. (The rest of the time you are told "Better luck next time!": i.e. you lose). Thus the possible lottery outcomes are a discrete random variable, X, with distribution:

Prize:	$50	$10	$3	$0
Probability:	0.01	0.04	0.25	0.7

d) Calculate μ_x, the expected value of lottery winnings.

e) Calculate σ_x, the standard deviation of the lottery.

5. **[20]** Suppose that Pizza Pizza decides to test out a new strategy of including a free apple pie dessert with every delivery order. They decide to test it out in London, Ontario (Canada's most average city). Before the plan is introduced, Pizza Pizza believes that they provide 21% of pizza delivery orders in London.

   Two months after introducing the free dessert, Pizza Pizza hires a researcher to sample pizza delivery choices. The research estimates, from a sample of $n_a = 351$, that 24% ($\widehat{p}_a = 0.24$) of pizza delivery orders are from Pizza Pizza.

   a) Write down the hypotheses that will allow you to test the null that Pizza Pizza's market share has not changed from the pre-apple pie level against the alternative that it has changed.

   b) Calculate a z test statistic associated with the hypothesis from part a), and use it to find the p-value of the test.

   c) Can you reject the null hypothesis at the at the $\alpha = 0.1$ significance level?

   Pizza Pizza then tries a different dessert: they now include a free chocolate brownie with orders instead of apple pie. After another two months, they conduct another survey, which uses a sample of $n_c = 191$ to estimate that Pizza Pizza's new market share is 34% (so $\widehat{p}_c = 0.34$).

   d) Calculate a 95% confidence interval for $p_c - p_a$ (the increase in market share due to the change from apple to chocolate desserts).

6. **[20]** Suppose that, among Queen's students, 89% of female students and 78% of male students like chocolate. (In other words, the probability of liking chocolate conditional on being female is 89%, and 78% conditional on being male). Also suppose that 62% of students are female.

   Let F denote the event that a person is female (and so F^c denotes the event that a person is male), and let C denote the event that a person likes chocolate.

   Note that the formula sheet uses the notation $\overline{A}$ instead of A^c to denote a complement. You may use whichever notation you prefer.

   a) Write down the three probabilities given above as probability statements. For example, $P(A) = 0.1$, $P(A \cup B) = 0.5$, and $P(A|B) = 0.9$ are probability statements for events A and B.

   b) Calculate the probability that a student likes chocolate and is female.

   c) Show that the probability that a student like chocolate and is male equals 0.2964.

   d) Calculate the probability that a student will like chocolate. *Formula sheet hint: F and F^c are "mutually exclusive and exhaustive events."*

   e) Calculate the probability that a student who likes chocolate is female.

   f) Are liking chocolate and being female independent events? Support your answer by showing a specific condition that is or is not satisfied.

7. [20] A researcher wishes to examine housing prices using the following linear regression model:

$$price = \beta_1 + \beta_2 sqrft + \beta_3 bdrms + u$$

where $price$ is the house price in thousands of dollars, $sqrft$ is the size of the house measured in square feet, and $bdrms$ is the number of bedrooms in the house.

Using a data set of $n = 88$ observations, the researcher uses a linear regression program which outputs the following values:

$$\widehat{\beta_1} = -19.31 \quad SE(\widehat{\beta_1}) = 31.047$$
$$\widehat{\beta_2} = 0.1284 \quad SE(\widehat{\beta_2}) = 0.01382$$
$$\widehat{\beta_3} = 15.20 \quad SE(\widehat{\beta_3}) = 9.484$$
$$n = 88$$
$$R^2 = 0.6319$$

where the $SE(\ldots)$ values are the standard errors of the estimates.

a) The first house in the sample has $sqrft = 2438$ and $bdrms = 4$. Find $\widehat{price}$, the price that the estimated regression model predicts for this house.

b) The actual sale price of the house in part *a)* is \$300,000 (so $price = 300.0$). Find the residual for this observation.

c) Give an economic interpretation of the $\widehat{\beta_2} = 0.1284$ coefficient.

d) What percentage of the sample variation in $price$ does the variation in $sqrft$ and $bdrms$ explain?

e) Calculate a 95% confidence interval for β_2. Based on this confidence interval, would you reject the null hypothesis that $\beta_2 = 0$ against the alternative that $\beta_2 \neq 0$?
Hint: Remember that the standard error takes the place of $s/\sqrt{n}$ in the usual confidence interval formula.