Economics 250 — Final Exam **11 April 2015**

Instructions:

- The exam is 3 hours in length. There are 180 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 30 marks) consists of 10 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 9 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.
-
- Good luck!

Section A: Very short questions [30 points]

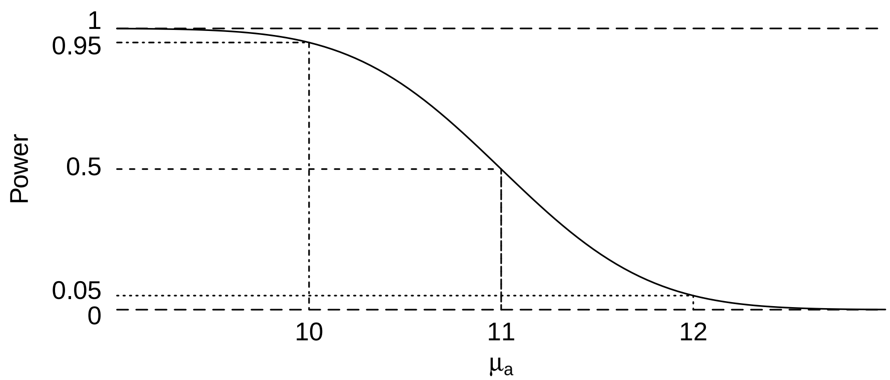
The following questions require only a couple of words, a lookup of a numerical value, or a small calculation. Each question is worth 3 marks.

1. Suppose X is a continuous, right-skewed distribution, and that $Y = \log(X)$ appears (approximately) normally distributed with a mean approximately equal to the median. Would you expect the mean of X to be *above*, *below*, or *approximately equal to* the median of X?

2. You have a sample of $n = 27$ observations and wish to test the hypothesis $\mu = 12$ against the alternative $\mu \neq 12$. The population's standard deviation is unknown. What will be the distribution of your test statistic?

3. What is the critical value of the appropriate distribution needed to calculate a 95% confidence interval for μ from a sample of $n = 27$ observations, without knowing σ?

4. The following graph shows the power of a test conducted at the $\alpha = 0.05$ level. Write down the null and alternative hypotheses of the test.



5. Suppose A and B are disjoint events, and that $P(A) = 0.4$. What is the range of values that $P(B)$ could take?

6. Suppose A and B are independent events, and that $P(A) = 0.4$. What is the range of values that $P(B)$ could take?

7. If 62% of Queen's students are female and 12% of female students own cars, and 14% of students own cars, what is the probability that a randomly selected student car belongs to a female student?

8. What is the probability that a weighted coin that produces heads 75% of the time will produce exactly 9 heads in 12 flips?

9. If X is distributed $\mathcal{U}(-10, 5)$, what is $P(X < 0 \cup X \geq 3)$?

10. Suppose that IQ scores in the general population follow the distribution $\mathcal{N}(100, 15)$. What is the distribution of the mean IQ score of a randomly selected group of 10 people?

Section B: Longer questions [150 points]

1. [**15**] A random variable X takes one of the four values: 0, 5, 6, or 7. The value 0 occurs with probability 0.4; the remaining values all have equal probabilities of occurring.

a) Find the mean and standard deviation of X.

Another random variable Y follows the $\mathcal{U}(1, 3)$ distribution, which has mean 2 and standard deviation 0.577. X and Y are *not* independent: the correlation coefficient between them is -0.5. A third random variable, W, is defined by $W = 2X - 3Y$.

b) Find the mean of W.

c) Find the standard deviation of W.

2. [**15**] The Canadian lottery "Lotto 6/49" is played by choosing any 6 different numbers from 1–49. The jackpot prize (of at least $5 million, and often much larger) is won if all 6 numbers on a ticket match the 6 numbers drawn by the Interprovincial Lottery Corporation. For example, a ticket could contain the numbers 4, 7, 19, 34, 45, 49, and would win the jackpot if the same 6 numbers are drawn (in any order).

You may leave your answers to this question as a multiplication of fractions; it is not necessary to obtain a decimal value.

a) What is the probability that the *first* number drawn by the Lottery Corporation matches any of the 6 numbers on a Lotto 6/49 ticket?

b) If the first draw matches one of the ticket numbers, what is the probability that the second draw matches one of the remaining ticket numbers?

c) Find the probability of winning the jackpot.

3. [**15**] A pollster conducts a poll of eligible, decided voters obtains information on the voting intentions of Canadians among the top three major national parties: Conservative, Liberal, and NDP. Among the 108 surveyed men, 53 intend to vote Conservative and 19 intend to vote NDP. Among surveyed women, 47 intend to vote Conversative and 38 intend to vote NDP. There are 93 Liberal supporters in the sample.

a) Create a two-way table using the above data. Use counts in the table rather than proportions.

b) What is the marginal distribution of support for political parties in the sample?

c) Find the probability that a female voter responded with either Liberal or NDP support.

d) Find the probability that a randomly selected voter in this sample is male and voted for either the Conservative or Liberal parties.

4. [10] Suppose that each student who takes Economics 250 has a 70% chance of passing the course.

a) What is the probability that at least 8 of a class of 10 students pass the course?

b) What is the probability that at least 80 of a class of 100 students pass the course?

5. [20] A study involving 36 tenth grade students measures the students' IQ scores and measures a sample mean of 94.7. IQ scores are known to have a standard deviation of 15.

a) Test the hypothesis at the 95% confidence level that IQ scores equal 100 against the alternative the IQ scores are not equal to 100.

b) What is the *p*-value of your test? Give an interpretation of this *p*-value.

c) Calculate the power of this test when the population mean actually equals 98.

d) Interpret your power value (in words).

6. [20] You collect the following values of waiting times (in days) for ACL repair surgery in a major Canadian city.

107	106	117	162	81	127	156	107

a) Use this data to calculate a test statistic that will allow you to test the hypothesis that the mean waiting time equals 140 against the alternative that the mean waiting time is less than 140 days. What is the distribution of this test statistic?

b) Test the hypothesis of part *a*) at the $\alpha = 0.05$ level.

c) Construct a 95% confidence interval for the waiting time.

7. **[20]** A survey is conducted using 47 six-year-old boys in a sub-Saharan African country. Participants are randomly assigned to one of two groups: the first group (of 21 boys) receives daily nutritional supplements over a four-year period, while the second group (26 boys) receives no treatment.

At the end of the four-year experiment, the height of the participants is measured. The first sample (the treatment group) has a mean height of 138cm with standard deviation of 15.7cm; the second sample has a mean height of 120cm and standard deviation of 13.6cm.

a) Test the hypothesis at the $\alpha = 0.05$ level that the nutritional supplement had no effect against the alternative that it increased heights.

b) Construct a confidence interval at the 98% confidence level for the difference in mean heights.

c) It is discovered that because of an incorrect formula in Excel, the data for the non-treatment group was actually divided by 1.1: the actual heights in the non-treatment sample need to be multiplied by 1.1.

Correct for this error and calculate the p-value for the test.

8. **[20]** Sodexo, the company providing exclusive food service at Queen's, is considering opening a new Tim Hortons franchise at a newly-available location on campus. To help make the decision, they conduct a simple random sample of the campus population and ask the respondents whether people would visit the coffee shop at the proposed location at least once per week. 15 out of 109 responders indicate that they would use the new location regularly.

Since the proportions in this question are relatively close to 0, you should use Wilson's estimate (adding an appropriate number of fake positive and negative responses) throughout this question.

a) Sodexo management will only consider opening the new Tim Hortons if they are confident that more than 10% of the campus population will visit at least once per week. State and perform an appropriate test at the 95% confidence level to determine whether the location will attract the required number of customers.

Sodexo performs a second simple random sample survey to determine whether another location is more suitable. 42 out of 170 participants in this survey indicate that they would visit the second location regularly.

b) Test the hypothesis at the $\alpha = 0.05$ level that the two locations will attract the same number of customers against the alternative that the second location will attract more customers.

c) Construct a 97% confidence interval for $p_2 - p_1$, the difference in customer proportions between the two locations.

9. **[15]** A researcher is studying the performance of high school students at different schools using the following regression model:

$$math10 = \beta_1 + \beta_2 totcomp + \beta_3 enroll + u$$

where:

- $math10$ is the percentage (from 0 to 100) of a school's students who pass a standardized grade 10 math test
- $totcomp$ is the average compension (salary plus benefits) of the school's teachers, measured in dollars
- $enroll$ is the number of students enrolled at the school

Using a data set of $n = 408$ randomly selected schools, the researcher uses a linear regression program which outputs the following values:

	Coefficient	Std. Error
const	8.320	3.487
totcomp	0.0004244	0.00009630
enroll	−0.0001658	0.0002137

R^2	0.050671	Adjusted R^2	0.045983
$F(2, 405)$	10.80850	P-value(F)	0.000027

a) Provide an economic interpretation of the $\widehat{\beta}_2$ coefficient.

b) Perform a hypothesis test to determine whether the data provides evidence at the $\alpha = 0.1$ level that larger schools has a negative effect on the grade 10 math test performance of students. What is the distribution and p-value of your test statistic?

c) What information does the $R^2 = 0.0507$ value provide?

d) One of the schools has $math10 = 51.1$, $totcomp = 39992$, and $enroll = 1116$. Calculate the residual for this observation. Does the fitted model predict a value for this observation that is too high or too low?