QUEEN'S UNIVERSITY
FACULTY OF ARTS AND SCIENCE

APRIL 2011

ECONOMICS 250
Introduction to Statistics

Instructor: Gregor Smith

Instructions:

The exam is three hours in length.

Do all ten (10) questions.

Be sure to show your calculations and intermediate steps.

Put your student number on each answer booklet.

Formulas and tables are printed at the end of this question paper.

You may use a hand calculator. Allowed calculators include those with blue or gold stickers, the Casio 991, the Sharp EL376S, or other non-programmable calculators. No red-sticker calculators or other aids are allowed.

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.

1. A discrete random variable x can take on the values 1, 2, or 3 in its population, each with equal probability. A second random variable y is given by $y = 4 - x$.

(a) Find the mean and variance of x.

(b) Find the correlation and covariance between x and y.

(c) If a third variable, z, is given by $z = 0.5x + 0.5y$ then what is the variance of z?

2. Suppose that 30% of independent bookstores are profitable. There is a 70% probability that, if it is profitable, an independent bookstore will be taken over by a larger chain. Among non-profitable, independent bookstores only 10% are taken over.

(a) What is the probability that an independent bookstore is not profitable and not taken over?

(b) If you observe a takeover, what is the probability that the store in question was profitable?

3. Suppose that $p = 0.25$ is the probability that a European bank is insolvent.

(a) In a random sample of 10 banks, what is the mode of the distribution of the number of insolvent banks?

(b) In a random sample of 10 banks, what is the probability that 3 or more banks are insolvent?

(c) In a random sample of 100 banks, what is the probability that 30 or more banks are insolvent?

4. Suppose that inflation rate, π, in different countries is uniformly distributed with mean 6 and variance 12.

(a) Find the range of possible values for the inflation rate.

(b) Suppose that the unemployment rate, u, across countries is negatively related to the inflation rate like this:

$$u = 10 - 0.5\pi.$$

Find the mean and variance of the unemployment rate.

5. Suppose that we record whether or not a recession is occurring using a discrete, random variable x, with $x = 1$ denoting a recession and $x = 0$ no recession. Now imagine a joint probability density function for the values of x this year and next year, denoted x_1 and x_2. Suppose $\text{Prob}(x_2 = 1|x_1 = 1) = 0.7$ and $\text{Prob}(x_2 = 1|x_1 = 0) = 0.2$.

(a) Find $E(x_2|x_1 = 1)$.

(b) Find the conditional variance of x_2 given that $x_1 = 1$.

6. Suppose that the population of investment returns is normally distributed with mean μ and variance σ^2. An economics student collects a sample of 6 returns with the following values: 3, 2, 4, 2, 7, 6.

(a) Find the sample mean and sample standard deviation.

(b) Find a 99% confidence interval for the population mean.

7. Suppose that a survey of 100 people in the labour force shows that the 8 of them are unemployed. A commentator claims the corresponding proportion in the population is 10%.

(a) Conduct a test of $H_0 : p = 0.10$ vs $H_1 : p < 0.10$ at the 5% significance level.

(b) What is the one-sided p-value?

8. Suppose that a sample of 64 people has an average life expectancy of 78 years with a sample variance 36 years. However, a researcher claims that the average life expectancy in the population is $\mu = 80$ years.

(a) Conduct a test of $H_0 : \mu = 80$ vs $H_1 : \mu \neq 80$ with $\alpha = 0.05$.

(b) Suppose that the researcher's claim is incorrect, and that the true value is $\mu = 78$. Find the power of the test.

9. A researcher interviews a random sample of 1000 workers in Alberta and finds that 200 report switching jobs during the past year. She also interviews a random sample of 800 workers in Saskatchewan and finds 100 report switching jobs during the past year.

Find a 90% confidence interval for the difference between the two population proportions.

10. A company offers a course that claims to improve students' scores on the LSAT (an aptitude test). Suppose that a group of 16 students write the test and receive an average score of 140. They then all take the course and repeat the test with an average score of 145. The sample variance of the change in their scores is 100.

(a) Find a 95% confidence interval for the mean change in the scores.

(b) The company claims the population value of the improvement in LSAT score is 10. Test this hypothesis with $\alpha = 0.05$ vs. the alternative that it is less than 10.

Economics 250
Winter Term 2010
Final Exam Answer Guide

1. (a) $\mu = 2$; $\sigma^2 = \frac{2}{3}$

(b) $\rho = -1$; $\sigma_{xy} = -\frac{2}{3}$ (Note: Finding a correlation greater than 1 in absolute value tells you you have made a calculation error.)

(c) $\sigma_z^2 = 0$.

2. (a) 0.63 or 63%

(b) $\frac{0.21}{0.28} = 0.75$

3. (a) 2

(b) 0.474 or 47.4% (from table 2 or 3)

(c) $x \sim N(25, 18.75)$ so $z = \frac{30-25}{4.330} = 1.154$ so 12.51% (from table 1 using 1.15)
(Or using the midpoint of 1.15 and 1.16 gives 12.40%.)

4. (a) The range is $[0, 12]$.

(b) The mean is 7 and the variance is 3.

5. (a) 0.7

(b) $0.7(1 - 0.7)^2 + 0.3(0 - 0.7)^2 = 0.063 + 0.147 = 0.21$

6. (a) $\overline{x} = 4$ and $s^2 = 4.4$ so $s = 2.0976$.

(b) Using t with 0.05% in the right tail and df 5 gives 4.032,
so the CI is $4 \pm 4.032(0.856) = (0.5472, 7.4528)$

7. Under the null hypothesis $\hat{p} \sim N(0.10, 0.0009)$ so the standard deviation is 0.03. The one-tailed, 5% critical value for z is -1.645 so:

$$-1.645 = \frac{\hat{p}_c - 0.10}{0.03}$$

so that $\hat{p}_c = 0.05065$. The actual sample value is 0.08 which lies above the critical value, so we do not reject H_0.

Alternately, simply construct a test statistic:

$$z = \frac{0.08 - 0.10}{0.03} = -0.66,$$

which lies above -1.645.

Notice that you need to use the value of p under the null (0.10) and *not* the sample value (0.08) to construct the standard error.

(b) Use

$$z = \frac{0.08 - 0.10}{0.03} = -0.666,$$

so that the p-value is 0.2514 (or 0.2546 if rounding to -0.66 rather than -0.67).

8. With $n = 64$ the critical values come from the z-distribution: ± 1.96. Thus the non-rejection region is:

$$80 \pm 1.96\frac{6}{8} = 80 \pm 1.47 = (78.53, 80.47).$$

The sample value, 78, is in the rejection region so we reject H_0 at $\alpha = 0.05$. Alternately, the test statistic is: $z = -2.67$ so we reject.

Notice that the non-rejection region for a two-sided test in general is *not* the same thing as a confidence interval. The former is centred at the value under the null hypothesis; the latter is centred at the sample value. So the two intervals coincide only if the null hypothesis is that the sample statistic is the population parameter.

(b) If the true value is $\mu = 78$ we want to know the probability under this alternative hypothesis that lies below the lower critical value for the original test: 78.53. Under H_1, $\overline{x} \sim N(0.78, 0.5625)$. Standardize:

$$z = \frac{78.53 - 78}{0.75} = 0.7066.$$

The power is the probability we would reject the null when it is false, *i.e.* we would find $\overline{x}$ below 78.53 when $\mu = 78$. That probability is 0.7611 from table 1.

9. In these samples $\hat{p}_x = 0.20$ and $\hat{p}_y = 0.125$. The standard error of the difference is:

$$\sqrt{0.00016 + 0.0001367} = 0.01722.$$

Thus the 90% confidence interval is:

$$0.075 \pm 1.645(0.01722) = 0.075 \pm 0.0283 = (0.0467, 0.1033).$$

10. (a) The confidence interval is

$$5 \pm 2.131\frac{10}{4} = (-0.3275, 10.3275).$$

(b) The test statistic is

$$t = \frac{5 - 10}{2.5} = -2$$

which is less than the critical value of -1.753, so we reject the null hypothesis at this level of significance.