QUEEN'S UNIVERSITY
FACULTY OF ARTS AND SCIENCE

APRIL 2013

ECONOMICS 250
Introduction to Statistics

Instructor: Gregor Smith

Instructions:

The exam is three hours in length.

Do all nine (9) 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 survey of households measures their weekly spending in dollars, recording it within ranges. Suppose 10 households report spending between 0 and 200, 10 households report spending between 200 and 400, and 20 households report spending between 400 and 600.

(a) What is the sample mean household spending?

(b) What is the sample standard deviation?

2. An investment return, labelled r_1, has mean 1 and variance 1. A second return, labelled r_2, has mean 2 and variance 4. A portfolio manager holds an equally-weighted combination (*i.e.* a 50:50 mixture) of the two investments, which has return $r_p = 0.5r_1 + 0.5r_2$.

(a) Suppose that the two returns are independent. Find the mean and variance of the portfolio return, r_p.

(b) Suppose that the covariance between r_1 and r_2 were actually -1. How would that fact change your findings in part (a)?

3. Waiting times at a visa office are continuously and uniformly distributed between 20 minutes and 120 minutes. Suppose that the opportunity cost of waiting is the product of the waiting time and the wage of the person waiting, which is \$20 per hour.

(a) What is the average opportunity cost?

(b) What is the probability of an opportunity cost greater than \$20?

(c) If 10 people make independent visits to the visa office then what is the probability that at least 5 of them wait less than an hour?

4. Suppose that a medical test can report patients as sick or well. Also suppose that, of people who are reported as sick according to the test, 10% are actually well. Suppose that 3% of all people are actually sick. Suppose that 3% of all people are reported as sick according to the test.

(a) What is the probability that someone is sick given that they are reported as sick?

(b) If someone is well, then what is the probability that the test reports that they are sick?

5. Suppose that GDP growth rates in Canada and the US can take on only two values, -1 and 1. The joint distribution of these discrete random variables is:

		US	
		-1	1
Canada	-1	0.1	0.4
	1	0.4	0.1

(a) What is the mean of each growth rate?

(b) What is the correlation between the two growth rates?

(c) Find the conditional mean and the conditional variance of the growth rate in Canada, given that the US growth rate is 1.

6. A researcher measures stock-market returns over 30 days and finds an average return of 2 with a sample standard deviation of 4.

(a) Find a 95% confidence interval for the population average return.

(b) Test the null hypothesis that the population average return is zero against the alternative hypothesis that it is greater than 0, at a significance level of $\alpha = 0.10$.

7. Suppose that a survey of 100 students show that 30 support a specific political party.

(a) Test the null hypothesis that the population proportion that supports this party is 35%, against the alternative hypothesis that the population proportion is less than that, at the significance level $\alpha = 10\%$. Find the value of the sample proportion that divides sample outcomes into a reject region and a do-not-reject region.

(b) Would your decision at this significance level be different if you used the 'plus 4' rule and calculated $\tilde{p}$?

(c) Suppose that an alternative hypothesis, that the proportion is 0.32, is actually correct. Find the probability of type II error using the decision rule you found in part (a).

8. A sample of 16 countries had an average GDP per capita of 31 before the 2008–2010 financial crisis and an average GDP per capita of 29.247 after the financial crisis. The sample standard deviation of the *change* in their GDP per capita was 4.

(a) Find a 99% confidence interval for the average effect of the financial crisis on GDP per capita.

(b) Test the null hypothesis that the average effect is zero against the alternative that it is not zero. Find the P-value for the test.

9. Suppose 8 out of 64 people in Vancouver report having experienced a property crime, while 9 out of 90 people in Toronto do so.

(a) Test the hypothesis that the population crime rates are equal in the two cities against the alternative hypothesis that the crime rate is higher in Vancouver, at the significance level $\alpha = 5\%$.

(b) What is the P-value associated with your test statistic?

Economics 250
Winter Term 2013
Final Exam Answer Guide

1. (a) We use the midpoint in each range for grouped data (see the formula sheet if you have forgotten this). Thus the mean is

$$0.25(100) + 0.25(300) + 0.5(500) = 350.$$

(b) The sample variance is:

$$s^2 = \frac{1}{39}[10(100-350)^2 + 10(300-350)^2 + 20(500-350)^2)] = 28,205.128$$

so the sample standard deviation is $s = 167.94$ dollars.

2. (a) The mean is 1.5 and the variance is 1.25. (Remember that independence implies a covariance of zero.)

(b) The mean would be unchanged. The variance would be:

$$\sigma_p^2 = 0.25(1) + 0.25(4) + 2(0.5)(0.5)(-1) = 1.25 - .5 = 0.75$$

3. (a) The average waiting time is 70 minutes or 1.1666 hours so the average opportunity cost is 23.33 dollars.

(b) The probability of a cost greater than \$20 is the probability of waiting longer than an hour which is 60%. (This can be read off from a picture of the uniform density.)

(c) The probability of 1 person waiting less than an hour is 0.4. Then from Table C the probability of at least 5 successes on 10 trials is 0.367 or 36.7%.

4. (a)

$$P(S|RS) = P(S \cap RS)/P(RS) = 0.027/0.03 = 0.9$$

or 90%

(b)

$$P(RS|W) = P(RS \cap W)/P(W) = 0.003/0.97 = 0.003092$$

5. (a) The mean is zero for each growth rate.

(b) The covariance is -0.6 and the variance of either growth rate is 1 so the correlation is -0.6 also. (Remember that the covariance obviously is negative from the pattern given and also that the correlation must be less than 1 in absolute value.)

(c) Given that the US growth rate is 1 the conditional mean for Canada is 0.8(-1)+0.2(1) = -0.6. The conditional variance is 0.64 (which is less than the unconditional variance).

6. (a) The 95% CI is:

$$2 \pm 2.045 \frac{4}{\sqrt{30}} = 2 \pm 1.4935 = (0.5065, 3.4935)$$

(b) Our test statistic is:

$$t = \frac{2-0}{0.73} = 2.739.$$

From table D, the critical value with df=29 is 1.311 so we reject the null in favour of the alternative.

7. (a) The test statistic is:

$$z = \frac{0.30 - 0.35}{\sqrt{0.35(0.65)/100}} = \frac{-0.05}{.0476} = -1.048$$

which uses the value p_0 to find the standard deviation. The critical value at $\alpha = 0.10$ is -1.28 so we do not reject the null. The critical value for $\hat{p}$ is 0.2889.

(b) If we used $\tilde{p} = 32/104 = 0.3076$ that would give $z = -0.89$ so again we would not reject at this level of significance.

(c) If the true value is 0.32 then our z-statistic for the critical value is is:

$$z = \frac{0.2981 - 0.32}{0.0466} = -0.66$$

(being careful to construct the standard deviation using the value 0.32). From Table A that leaves 0.2514 below and 0.7486 above. The probability of type II error is 74.86%.

8. (a) The 99% confidence interval is

$$-1.753 \pm 2.947\frac{4}{4} = (1.194, -4.7).$$

Many students gave this with the wrong signs.

(b) For this test our test statistic is:

$$t = \frac{-1.753 - 0}{1} = -1.753$$

with 15 df. (Again many students gave this with the wrong sign.) So from Table D the P-value is 2×0.05 (because this is a two-tailed test) or 0.10.

9. (a) The rate in Vancouver is 0.125 and that in Toronto is 0.10, for a difference of 0.025. The pooled estimate of the common proportion is 17/154 or 0.11. The standard deviation then is 0.0511. Thus the test statistic is:

$$z = \frac{0.025 - 0}{0.0511} = 0.489$$

This is much less than 1.645 (for a one-tailed, 5% test) so we do not reject the null that the two population rates are equal, at this level of significance.

(b) Also the P-value (from a statistic of 0.49) is $1 - 0.6879 = 0.3121$.