Economics 250 — Midterm 2 — Answer guide **17 November 2016**

1. [**9**] Suppose that the value of 1 US dollar in Canadian dollars is normally distributed with mean 1.34 and standard deviation 0.02.

The value of 1 Euro in Canadian dollars is also normally distributed with mean 1.44 and standard deviation 0.03.

The correlation between these two exchange rates equals -0.3.

a) Suppose that you have 30 US dollars, but you owe a friend 27 Euros. What is the mean value of your net assets (that is: the Canadian dollar value of what you have minus the Canadian dollar value of what you owe)?

b) What is the distribution of the Canadian dollar value of your net assets?

c) Find the probability that the Canadian dollar value of your net assets is positive.

Answers:

a) Let X denote the value of 1 US dollar and let Y denote the value of 1 Euro. Let W equal net assets, so:

$W = 30X - 27Y$

The mean of W is then:

$\mu_W = 30\mu_X - 27\mu_Y = 30(1.34) - 27(1.44) = 1.32$

b) Since these are both normally distributed, if we add them together it will still be normally distributed. The mean is 1.32, from the previous question. To calculate the standard deviation, we first need the variance:

$\sigma_w^2 = 30^2\sigma_x^2 + 27^2\sigma_y^2 - 2(30)(27)\sigma_{xy}$

To get the last term, we need to use the correlation formula:

$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$, and so $-0.3 = \frac{\sigma_{xy}}{(.02)(.03)}$, which means $\sigma_{xy} = -0.00018$.

That means: $\sigma_w^2 = (900)(.0004) + (729)(.0009) - 1620(-.00018) = 1.3077$, and so $\sigma_w = \sqrt{1.3077} = 1.144$.

And so: $W \sim \mathrm{N}(1.32, 1.144)$.

c) $P(W > 0) = P(Z > \frac{0-1.32}{1.144}) = P(Z > -1.15) = 1 - 0.1251 = 0.8749$

2. **[13]** Suppose that a researcher interesting in studying household income in a developing country samples 20 of the country's households, calculating their income (converted to US dollars). The mean of her sample equals 3000.

Household income in the country is known to have a standard deviation of 1500.

The researcher is interesting in looking for statistically significant evidence in support of an increase in the population from its value ten years ago of 2500.

a) Write down the hypotheses that the researcher wishes to test. Is this a one-tailed or two-tailed test?

b) Find the p-value for this statistical test. Based on this value, can you reject, at the $\alpha = 0.1$ level, that there has been no increase in mean household income?

c) What is the power of this test (with $\alpha = 0.1$) if the population mean income has increased by 300 (that is, if the population mean actually equals 2800)?

d) Without changing α, how could the researcher have decreased the probability of making a Type II error (that is, failing to reject when rejecting is correct) when the true population mean equals 2800?

Answers:

a) The null and alternative hypotheses are:

$$H_0 : \mu = 2500$$
$$H_a : \mu > 2500$$

b) Since $\sigma = \sqrt{2250000} = 1500$ is known, we use a z test:

$$z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{3000 - 2500}{1500/\sqrt{20}}$$
$$= 1.49$$
$$P(Z > 1.49) = 1 - 0.9319 = 0.0681$$

where the final value comes from the table of standard normal probabilities.

Since the p-value of 0.0681 is less than $alpha = 0.1$ we reject the null: there *is* statistically significant evidence that the mean has increased.

c) First we need to figure out the value of $\overline{x}$ where we would start rejecting. For a one-sided, $\alpha = 0.1$ test that happens at $z^* = 1.28$. We find the associated $\overline{x}^*$ from:

$z^* = \frac{\overline{x}^* - 2500}{1500/\sqrt{20}}$

Plugging in $z^* = 1.28$ and rearranging for $\overline{x}^*$ we get: $\overline{x}^* = 2500 + (1.28)(1500/\sqrt{20}) =$ 2929.

The power is then the probability that we get a sample mean bigger than 2929 when sampling from a population with mean 2800:

$$\begin{aligned} Power &= P(\overline{x} > 2929 \mid \mu = 2800) \\ &= P\left(Z > \frac{2929 - 2800}{1500/\sqrt{20}}\right) \\ &= P(Z > 0.38) \\ &= .3520 \end{aligned}$$

This is the power of the test—i.e. the probability of rejecting the null hypothesis—when the population mean equals 2800.

d) The first thing to note: Power is simply the complement of the probability of making a type II error: that is, $Power = 1 - P(\text{Type II})$. So this question is really asking about how the researcher could increase the test's power.

To increase the power of the test at a given alternative and α, the researcher would need to increase n, the sample size.

It isn't necessary to show this, but as an illustrating example, if you increased the sample size from 20 to, say, 80, you'd get $\overline{x}^* = 2500 + (1.28)(1500/\sqrt{80}) = 2715$. The power would then be: $P(\overline{x} > 2715) = P(Z > -0.51) = 0.6950$, and so the probability of making a Type II error would have decreased from $1 - .3520 = 0.6480$ to $1 - .6950 = 0.3050$.

3. [**8**] Suppose that in a very boring class (not ECON 250, of course) held in the late afternoon, any student falls asleep with probability 0.3.

a) If the class has 9 students, what is the probability that 2 or fewer students fall asleep?

b) If the class has 9 students, what is the probability that 3 or more students fall asleep?

c) If the class has 90 students, what is the probability that at least a third of the class will fall asleep?

a) This is simply asking you to use the binomial formula:

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(x = 2) \\ &= \binom{9}{0} 0.3^0 0.7^9 + \binom{9}{1} 0.3^1 0.7^8 + \binom{9}{2} 0.3^2 0.7^7 \\ &= 1(.7)^9 + 9(.3)(.7)^8 + 36(.3)^2(.7)^7 \\ &= 0.463 \end{aligned}$$

b) You *could* do the full calculation:

$$\begin{aligned} P(X \geq 3) &= P(X = 3) + P(X = 4) + \ldots + P(x = 9) \\ &= 0.266827932 + 0.171532242 + 0.073513818 + 0.021003948 \\ &\quad + 0.003857868 + 0.000413343 + 0.000019683 \\ &= 0.537 \end{aligned}$$

but it is far easier to recognize that $X \geq 3$ is the complement of $X \leq 2$ that we answered in the first question, and so the answer must be:

$1 - 0.463 = 0.537$

c) For this we want to use the normal approximation: $X \sim \mathrm{N}(np, \sqrt{np(1-p)})$. Plugging in n and p that is: $X \sim \mathrm{N}(27, 4.347)$.

Thus, using the approximation, the answer is:

$P(X \geq 30) = P(Z \geq \frac{30-27}{4.347}) = P(Z \geq 0.69) = 0.2451$

You can, alternatively, use the correction mentioned in the textbook of using the point halfway between the first included and first excluded values (in this case, 29.5 because it's halfway between 29 and 30). If you did this, you would get:

$P(X \geq 29.5) = P(Z \geq \frac{29.5-27}{4.347}) = P(Z \geq 0.56) = .2877$

which is a little closer to the actual value (0.279). We did not talk about this approximation in class, however, and so using it is not required.

4. [10] A potato farmer is trying out a new variety of potatoes. He is particularly interested in knowing the size of this variety, and so he selects a random sample of 30 potatoes, measures their diameters, and sends you the data.

You plot a histogram of the data, and notice that the sizes do not appear be normally distributed, but instead look quite close to a uniform distribution.

a) Suppose the population distribution of potato diameters was actually $\mathrm{U}[7, 15]$ (which has mean 11 and standard deviation 2.31). What would be the (approximate) distribution of the mean of a sample of size 30 from this population?

Note: do not *continue to use this population assumption for the following questions.*

b) From the farmer's sample, you calculate a mean of 10.2cm and standard deviation of 2.1cm.

Construct a 95% confidence interval for the mean diameter of the farmer's potatoes.

c) The farmer's ideal potato size is 11.1cm. The farmer wants you to test whether or not this is the case for the new variety.

Can you reject that the hypothesis that the mean size equals 11.1cm in favour of the alternative that the mean does not equal 11.1cm at the 95% confidence level?

a) Even though each individual observation is from a uniform distribution, the *mean* of the sample will be very close to $N(11, 2.31/\sqrt{30}) = N(11, 0.422)$ (by the central limit theorem).

b) Since we don't know σ but have to use s, we will have a t distribution. For a sample of $n = 30$ we therefore have $df = 29$ and our critical t^* value for a 95% confidence interval equals (from Table D) 2.045.

Using this, we can calculate the confidence interval:

$$\begin{aligned} &[\overline{x} - t^* \frac{s}{\sqrt{n}}, \overline{x} + t^* \frac{s}{\sqrt{n}}] \\ = &[10.2 - 2.045(2.1/\sqrt{30}), 10.2 + 2.045(2.1/\sqrt{30})] \\ = &[9.416, 10.984] \end{aligned}$$

c) Though this question is worded like a hypothesis testing question, there is a huge shortcut here: because we just worked out the 95% confidence interval in the previous part, *and* because this is a two-sided test, we can immediately see that 11.1cm is not in the confidence interval, and therefore we can reject at the $\alpha = 0.05$ level without needing to calculate a t statistic at all.

If you were to calculate it anyway, the t statistic is:

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}} = \frac{10.2 - 11.1}{2.1/\sqrt{30}} = -2.347$$

For $df = 29$ the one-sided p-value would be between 0.01 and 0.02, and so the two-sided p-value we are interested in is between 0.02 and 0.04, which means it is less than $\alpha = .05$: we reject.