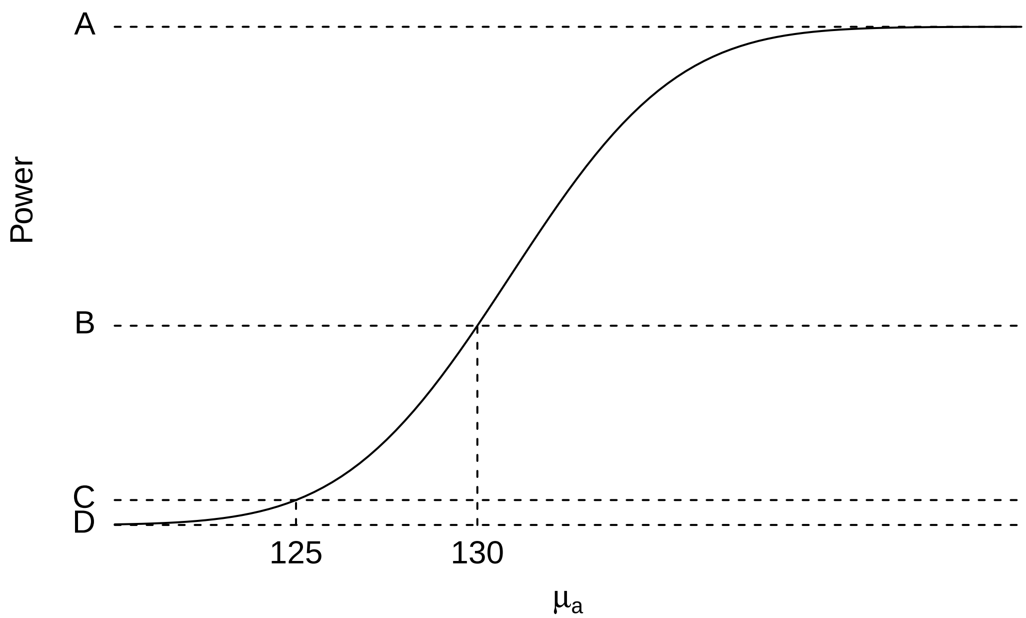
Economics 250 — Midterm 2 (answers) **14 November 2013**

Instructions:

- The exam is 80 minutes in length.
- You may use a hand calculator.
- Hand in your answers. Do not hand in the question and formula sheets.
- Show your work: incorrect answers without any work shown cannot be given partial marks.
- Answer **three** of the **four** questions in the answer booklet provided. If you attempt all four questions, indicate clearly which questions you wish to be graded. If there is no clear indication, only the first three answered questions will be graded.
- Each of the four questions is worth the same amount. The allocation of points within each question is as shown.
- This midterm has 4 pages, including this cover sheet. An additional 6 pages (3 formula sheet pages and 3 statistical tables) are also provided.

1. [**25**] You intend to collect a set of data to estimate μ, the population mean of X, a measure of happiness. You believe, before looking at the data, that the population mean is 125. $\sigma_X = 17.97$ is known.

Before performing the test, you decide on a significance level $\alpha = 0.05$ and sample size $n = 25$. You wish to examine the power of the z test statistic at an economically significant alternative that the mean has changed to 130, and calculate a test power of 0.4. Concerned that your test lacks power, you decide to calculate and graph the power (the probability of rejecting H_0 when H_a is true) of the test statistic for different specific values of μ_a:



a) [**4**] Write down the null (H_0) and alternative (H_a) hypotheses that agree with the graph above. Is this a one-tailed or two-tailed test?

Answer: Since the power goes to 1 far on the right, and goes to 0 far to the left, this is a one-tailed test that rejects for larger values. Thus:

$$H_0 : \mu = 125$$
$$H_a : \mu > 125$$

b) [**4**] Write down the values of A, B, C, and D.

Answer: $A = 1$, $B = 0.4$, $C = \alpha = 0.05$, $D = 0$.

c) [**5**] What is the probability of making a Type I error (rejecting H_0 when H_0 is true)? What is the probability of making a Type II error (failing to reject H_0 when H_a is true) at the specific alternative $\mu_a = 130$?

Answer: $P(\text{Type I}) = \alpha = 0.05$ (by definition of α).

$P(\text{Type II}) = 1 - \text{Power} = 1 - 0.4 = 0.6$

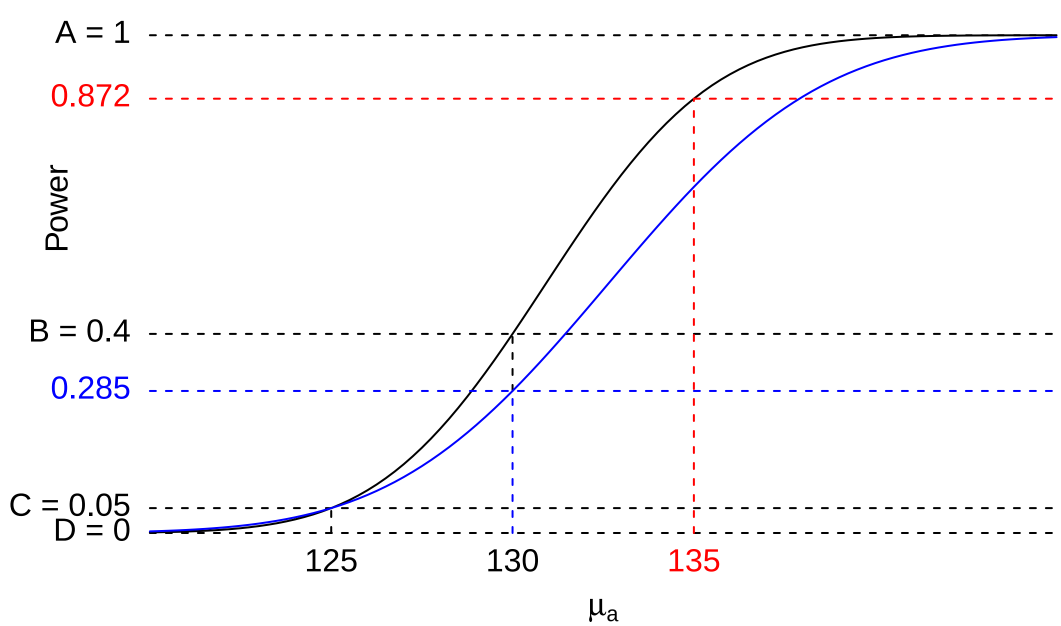
d) [**6**] Would the test power increase, decrease, or stay the same if you calculated the power at $\mu_a = 135$, leaving the other parameters unchanged? Would the *graph* of test power change? If the graph changes: will it be above, equal to, or below B at $\mu_a = 130$; and will it be above, equal to, or below C at $\mu_a = 125$?

Answer: The test power will increase because we are moving to a point to the right of 130 on the graph. The power curve itself won't change (it depends only on n, α, μ_0, and σ): we're just moving along the curve to different point. See the red values and lines in the diagram below for a graphical depiction.

e) [**6**] Would the test power at $\mu_a = 130$ increase, decrease, or stay the same if you reduce the size of the sample from $n = 25$ to $n = 15$, leaving the other parameters unchanged? Would the *graph* of test power change? If the graph changes: will it be above, equal to, or below B at $\mu_a = 130$; and will it be above, equal to, or below C at $\mu_a = 125$?

Answer: A smaller sample has less power to make statistical inference, and so the graph changes. In particular, it will lie below the old graph everywhere to the right of 125, and above the old graph everywhere to the left of 125. Thus it passes below B at 130. Since the power graph always passes through the point $(\mu_0, \alpha) = (125, 0.05)$, changing n affects neither of those values: thus the power will still be equal to $C = \alpha = 0.05$ at $\mu_a = \mu_0 = 125$.

The following diagram depicts the new graph and power levels in blue.



2. **[25]** Suppose that passing Economics 250 increases a person's future earning potential with probability 0.75.

a) **[4]** If you select ten random students who passed Economics 250, what is the probability that exactly seven of them will have higher future incomes?

Answer:

$$P(X=7)=\binom{10}{7}0.75^7(1-0.75)^3=(120)0.75^7 0.25^3=0.250$$

b) **[6]** What is the probability that, of the ten students selected in part *a*), at least eight of them will have higher future incomes?

Answer:

$$\begin{aligned}P(X\geq 8)&=P(X=8)+P(X=9)+P(X=10)\\&=\binom{10}{8}0.75^8 0.25^2+\binom{10}{9}0.75^9 0.25+\binom{10}{10}0.75^{10}\\&=0.282+0.188+0.056\\P(X\geq 8)&=0.526\end{aligned}$$

c) **[5]** What is the probability that, of the ten students selected in part *a*), six or fewer will have higher future income? *Hint: using your answers from parts* a*) and* b*) may help.*

Answer: You can calculate this the long way (see below), but it's much easier to use the two values calculated in parts *a*) and *b*):

$$\begin{aligned}P(X\leq 6)&=1-P(X>6)=1-P(X=7)-P(X\geq 8)\\&=1-0.250-0.526=0.224\end{aligned}$$

Doing this the long way is much more work, but it does work:

$$\begin{aligned}P(X\leq 6)&=P(X=0)+P(X=1)+P(X=2)+\ldots+P(X=6)\\&=0.0000+0.0000+0.0004+0.0031+0.0162+0.0584+0.1460\\&=0.2241\end{aligned}$$

d) **[5]** If you select eighty (80) students who passed Economics 250, what is the probability that at least 55 of them will have higher future incomes? Your answer should use an appropriate approximation.

Answer: Since the sample is large, a normal approximation works reasonably well here:

$$\begin{aligned}P(X\geq 55)&=P\left(Z\geq\frac{55-np}{\sqrt{np(1-p)}}\right)=P\left(Z\geq\frac{55-80(0.75)}{\sqrt{80(0.75)(0.25)}}\right)\\&=P(Z\geq -1.29)=1-P(Z\leq -1.29)=1-0.0985\\P(X\geq 55)&=0.9015\end{aligned}$$

e) **[5]** Of the eighty students selected in part *d*), what is the probability that at least 60 will have higher future incomes? Your answer should use an appropriate approximation.

Answer: Since the approximation says that the binomial distribution is distributed approximately $N(np, \sqrt{np(1-p)}) = N(60, \sqrt{15})$, $P(X \geq 60)$ is just the probability of a normal distribution producing a value above its mean, which is one-half. So: $P(X \geq 60) = 0.5$.

You can plug this into the z formula as well: you'll get $z = 0$, which gives you $P(Z \geq 0) = 0.5$

3. **[25]** A researcher is analyzing the returns of investing in a stock of a hot new company and collects daily data on the stock's performance, which he believes to be normally distributed. From a sample of the daily price over 15 days, he calculates the mean daily percentage change in the stock return to be 0.5.

a) **[6]** The researcher assumes that the variance of the measured variable equals $\sigma^2 = 10$ (and so standard deviation is $\sigma = 3.162$). Using this assumption, find a 97% confidence interval for the mean daily percentage stock return.

Answer: We first need to find z^* that gives a tail probability of 0.015 (so that the area of both tails equals 0.03). This is $z^* = 2.17$, and so the confidence interval is:

$$\left[0.5 - 2.17\frac{\sqrt{10}}{\sqrt{15}}, 0.5 + 2.17\frac{10}{\sqrt{15}}\right]$$
$$= [-1.272, 1.272]$$

b) **[6]** A reporter doing a news story on the company for a local paper interviews the researcher. Upon learning the confidence interval, the reporter responds that the decimal places in the researcher's confidence will confuse his readers (and his editor). The reporter insists on using the narrower confidence interval $[-1, 2]$ instead of what you calculated above.

What confidence level must the reporter use to make the $[-1, 2]$ confidence interval statistically valid? *Hint: start by finding the critical value that would give the desired confidence interval endpoints.*

Answer: To get a left endpoint of -1 we need:

$$0.5 - z^*\frac{\sqrt{10}}{\sqrt{15}} = -1$$

Which can be easily solved for $z^* = 1.837$. (You could also solve using the upper endpoint, and would get exactly the same z^*).

Then what we need is the two times the p-value for this z^*, which comes from the table (using 1.84): $2 \times 0.0329 = 0.0658$. The confidence level is 1 minus this times 100%, or 93.4%.

c) **[7]** Another researcher believes that the first researcher just made up the assumption of $\sigma^2 = 10$. She uses the same 15 observations to calculate $s^2 = 5.92$ (and so $s = 2.433$). Find a 95% confidence interval for the mean daily stock return without relying on the first researcher's assumption.

Answer: First we need the critical value from the t_{14} distribution: $t^* = 2.145$. Now plug this into the confidence interval formula:

$$\begin{aligned}
&\left[\overline{x} - t^* \frac{s}{\sqrt{n}}, \overline{x} + t^* \frac{s}{\sqrt{n}}\right] \\
=&\left[0.5 - 2.145 \frac{2.433}{\sqrt{15}}, 0.5 + 2.145 \frac{2.433}{\sqrt{15}}\right] \\
=&[-0.847, 1.847]
\end{aligned}$$

d) **[6]** If the reporter still wished to use the $[-1, 2]$ confidence interval with the second researcher's analysis, would the confidence level be greater or less than 95%?

Answer: It will be above 95% because -1 and 2 are *outside* the $[-0.847, 1.847]$ confidence interval. Since a higher confidence level value means we are less willing to accept Type I errors, a larger confidence interval is needed for higher confidence in the interval. Thus the confidence level for the $[-1, 2]$ interval will be higher than $[-0.847, 1.847]$.

4. **[25]** A researcher collects data on a country's mean monthly household income level in the month following a major increase in the country's minimum wage rate. She obtains a random sample of 12 of the country's households, collecting household income from each.

Thinking back to her Economics 250 course, the researcher remembers that income is usually strongly right-skewed, but that the logarithm of income is reasonably close to normal. She calculates $x_i = \ln(income_i)$ for each of the 12 households then uses these values to calculate the sample mean, $\overline{x} = 8.700$, and sample variance, $s^2 = 0.0963$ (and thus sample standard deviation $s = 0.310$).

She wishes to use the data to test the null hypothesis that mean household income has not changed from the pre-increase mean of $\mu = \ln(5166) = 8.55$ against the alternative that the mean household income has increased.

a) **[4]** State the null and alternative hypotheses the researcher wishes to test. State the hypotheses in terms of the log-income values (as given above) rather than direct income values.

Answer:

$$H_0 : \mu = 8.55$$
$$H_a : \mu > 8.55$$

where 8.55 is the given natural logarithm of 5166.

It's worth noting in passing that using the direct income values in the hypothesis would represent exactly the same test, but would be slightly confusing because the μ_0 we need to use to calculate a z or t value needs to be the natural logarithm value (since $\overline{x}$ is a mean of natural logaritms).

We could proceed that way, plugging in $\ln(\mu_0)$ instead of μ_0, and get the same result: but it's easier to use the transformed value in the null and alternative hypotheses to avoid this complication.

b) [**5**] Find the test statistic associated with the test, and give as much information about the p-value as is possible given the available statistical tables.

Answer:

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{8.7 - 8.55}{0.310/\sqrt{12}} = 1.676$$

From the $df = n - 1 = 11$ line of Table D, we see that 1.676 is between 1.363 and 1.796, the critical values for $p = .1$ and $p = .05$, thus we can say that the p-value is somewhere between 0.05 and 0.1.

c) [**4**] Can the researcher conclude that income has increased at the $\alpha = 0.05$ significance level? Can the researcher conclude that income has increased at the 95% confidence level?

Answer: No: our p-value is somewhere in $0.05 < p < 0.1$, which means we can't reject at the $\alpha = 0.05$ level since our p-value is bigger than 0.05.

The 95% confidence level question is asking exactly the same thing, so the answer is exactly the same.

d) [**4**] Suppose that the researcher had actually collected the data from $n = 24$ instead of $n = 12$ households and obtained the same $\overline{x}$, s^2, and s values given above. Would the test statistic and p-value change from what you calculated in part *b*)? If so, would they become larger or smaller?

Answer: If n increases, the denominator of the t-statistic $(s/\sqrt{n})$ gets smaller, and so the t-statistic gets bigger. A bigger t-statistic means a smaller p-value. So t goes up and p goes down.

It isn't necessary, but you could also answer the question by finding t and p. You'd get $t = 2.370$, and the corresponding p-value (this time with $df = 23$) would be between 0.01 and 0.02.

e) [**3**] In order to calculate a (two-tailed) confidence interval for the mean household income, you would need to find a critical value, t^*. Write down the critical value you would use to calculate a 98% confidence interval for a data set of $n = 18$ households.

Answer: This is simply a matter of reading the value off the table for $p = 0.01$ and $df = n - 1 = 17$: $t^* = 2.567$.

f) [**5**] If someone gave you a 99% confidence interval for μ of $[8.583, 8.817]$, would you reject $H_0 : \mu = 8.55$ against the two-sided alternative $H_a : \mu \neq 8.55$ at the $\alpha = 0.01$ significance level? At the $\alpha = 0.05$ significance level?

Answer: Yes: since 8.55 is outside the 99% confidence interval, we can reject it at the 99% confidence level, which is the same as saying reject it at the $\alpha = 0.01$ significance level.

At the $\alpha = 0.05$ level, we can reject it as well since the 95% confidence interval will be inside the 99% confidence interval: 8.55 will be even further outside the 95% than it is outside the 99% confidence interval.