Homework 13: ECO220Y – SOLUTIONS

Required Problems:

(1) The hypotheses in formal notation are:

$$H_0: p = 0.2$$
$$H_1: p < 0.2$$

Given $\hat{P} = \frac{40}{400} = 0.1$ and $n = 400$:

$$\text{P-value} = P\left(\hat{P} < 0.1 \mid p = 0.2, n = 400\right) = P\left(Z < \frac{0.1 - 0.2}{\sqrt{\frac{0.2*0.8}{400}}}\right) = P(Z < -5) \cong 0$$

Hence, we have overwhelming support for the conclusion that fewer than 20% of Canadians are in favor of legalizing heroin. The tiny P-value means we can very confidently reject the null and infer the research hypothesis is true.

(2) The hypotheses in formal notation are:

$$H_0: p = 0.15$$
$$H_1: p > 0.15$$

Given $\hat{P} = \frac{16}{100} = 0.16$ and $n = 100$:

$$\text{P-value} = P\left(\hat{P} > 0.16 \mid p = 0.15, n = 100\right) = P\left(Z > \frac{0.16 - 0.15}{\sqrt{\frac{0.15*0.85}{100}}}\right) = P(Z > 0.28) = 0.390$$

Hence, we have insufficient evidence to prove that more than 15% of e-mails are junk. The large P-value means that we fail to reject the null: there is a good chance of observing 16% or more junk mail in a random sample of 100 e-mails even if only 15% of all e-mails were junk. This does NOT mean that we have proven that 15% are junk nor that we accept that 15% are junk. We just do not have *enough* evidence to prove that more than 15% are junk.

(3) The hypotheses in formal notation are:

$$H_0: p = 0.02$$
$$H_1: p \neq 0.02$$

Given $\hat{P} = \frac{30}{1{,}000} = 0.03$ and $n = 1{,}000$:

$$\text{P-value} = P\left(\hat{P} > 0.03\right) + P(\hat{P} < 0.01) = 2 * P\left(Z > \frac{0.03 - 0.02}{\sqrt{\frac{0.02*0.98}{1000}}}\right) = 2 * P(Z > 2.26) = 0.024$$

Hence, we have substantial evidence to prove that the fraction of Ontario Grade 3 students scoring above the 98th percentile on the standardized intelligence test *differs* from 2% (remember that this is a two-tailed test). The small P-value means that it is quite unlikely that we would observe a sample proportion being off by as much as one percentage point because of sampling error. The small P-value means we can be quite confident in our decision to reject the null and infer the research hypothesis is true.

(4) H_0: $\varepsilon = 1$; H_1: $\varepsilon < 1$

(5) H_0: $\varepsilon = -1$; H_1: $\varepsilon > -1$

(6) (a) $H_0: p = 0.512$ versus $H_1: p > 0.512$

$\hat{P} = \frac{7,546}{14,789} = 0.5102441$

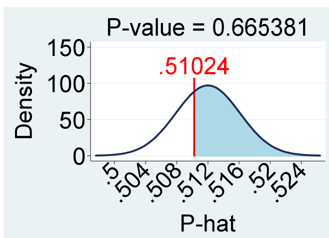
$$z = \frac{0.5102441 - 0.512}{\sqrt{\frac{0.512 * (1 - 0.512)}{14,789}}} = -0.427$$

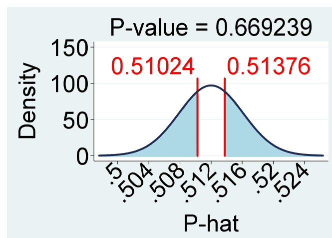
$P(Z > -0.43) = 0.5 + 0.1664 = 0.666$ (Note: We had to round off to -0.43 to use the Normal table, so we get an approximate P-value, which is still very close to the exact answer from software of 0.665.)

(b) For the third child born we have the strongest evidence of interference with the natural chances of having a male. One reason is that families that place a high value on a son will become increasingly desperate if the first two are girls. Remember that these results are pooled for all third born children and not just those of families with two girls already. If we looked at that subgroup alone we would likely find an extremely distorted proportion of boys born.

(c) The correct answer is **(D)**. If you got the wrong answer, you probably skipped drawing the two graphs. Draw a graph illustrating the P-value for the original right-tailed test and draw a graph for the proposed two-tailed test.

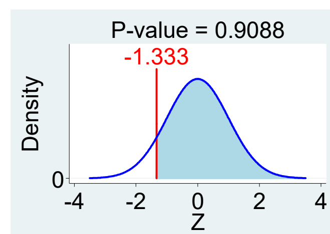
If you are still stuck, see my graphs below.





(7)

$H_0: p = 0.10$
$H_1: p > 0.10$



$$z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

$$-1.33 = \frac{\hat{P} - 0.10}{\sqrt{\frac{0.10(1 - 0.10)}{100}}}$$

$\hat{P} = 0.06$, 6 out of 100 in the sample were defective. We have *no evidence* to support the assertion that more than 10% of the parts are defective: in our sample only 6% were defective, which is the *opposite* of what the research hypothesis says. This yields a very large P-value: the bigger the P-value the weaker the support for the research hypothesis.

(8) In this case, the null hypothesis is $H_0: p = 0.5$ and the research hypothesis is $H_1: p \neq 0.5$. Even though the evidence happens to line up perfectly with the null – $\hat{P} = \frac{50}{100} = 0.5$ – we still CANNOT conclude that the null hypothesis is true. Notice the word *never* in the quote from R.A. Fisher. This is an inconclusive result: fail to reject the null.

(9) The quote is arguing in favor of CI estimation. It points out that CI estimation helps overcome pervasive human misconceptions about sampling error. Interestingly, a 50% CI estimate used to be popular (now 95% is popular).