Homework 15: ECO220Y – SOLUTIONS

Problems:

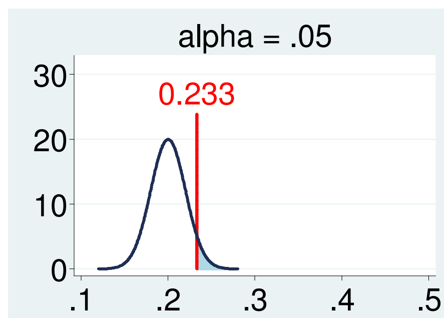
(1) FALSE. REALLY FALSE. The idea that there is a trade-off between the P-value and Type II error is a common misconception based on seriously flawed logic. There is a trade-off between Type I and Type II error but that refers to the significance level (α, the predetermined threshold for the maximum Type I error we are willing to tolerate) and not the P-value (measure of the strength of our actual evidence). It is true that the P-value is the probability of making a Type I error. However, it does NOT follow that a smaller P-value means a larger Type II error. First, the trade-off refers to how big the burden of proof is: the choice of a significance level (α). If you choose a large significance level, such as $\alpha = 0.10$, then you would raise the chance of sending an innocent person to jail (Type I error) and you would decrease the chance of letting a guilty person go free (Type II error). If you choose a smaller significance level, such as $\alpha = 0.01$ (closer to the "beyond a reasonable doubt" standard), then you would lower the chance of sending an innocent person to jail (Type I error) and you would increase the chance of letting a guilty person go free (Type II error). BUT, while you choose the significance level, you do NOT choose the P-value. Hence there is a difference between the significance level, which is the burden of proof, and the P-value, which is the proof/evidence you actually have. We would never say that because we have a huge amount of evidence of a defendant's guilt (very small P-value) that we think there is an increased chance of letting a guilty person go free. That doesn't make sense. We could say that if we REQUIRE a huge amount of evidence to convict (a very small significance level) then we will increase the chance of letting a guilty person go free. Another way to think about this is that there are many factors that affect BOTH the P-value and the chance of making a Type II error: sample size, the standard deviation of the sample proportion, whether research hypothesis is one or two directional, and the value specified in the null hypothesis. Changing any of these underlying factors will change both the P-value and the probability of making a Type II error. Hence we cannot say there is a causal relationship between the P-value and the probability of making a Type II error, which means that we cannot say that changing one would result in a change in the other.

(2) (a) Find the critical value of the hypothesis test.

$P(Z > 1.645) = 0.05$

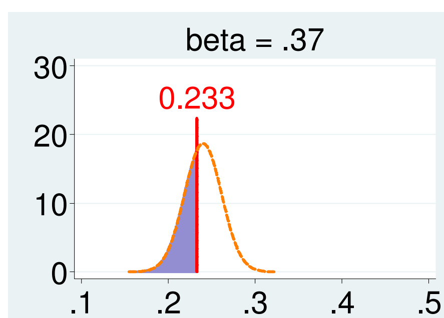
$$P\left(\hat{P} > 1.645 * \sqrt{\frac{0.20(1-0.20)}{400}} + 0.20 \mid H_0 \ is \ true\right) = 0.05$$

$P(\hat{P} > 0.233 \mid H_0 \ is \ true) = 0.05$



Hence the critical value (unstandardized) is $p^* = 0.233$. The rejection region is (0.233, ∞): if our test statistic, $\hat{P}$, lies in this region then we (correctly) reject the false null hypothesis.

If our test statistic, $\hat{P}$, lies outside this region then we (incorrectly) fail to reject the false null hypothesis: make a Type II error. Find the probability of a Type II error.



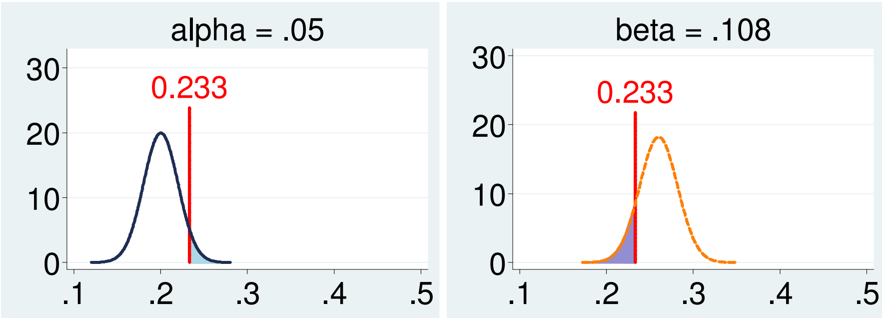
$\beta = P(\hat{P} < 0.233 \mid p = 0.24, \ n = 400)$

$$\beta = P\left(\frac{\hat{P} - 0.24}{\sqrt{\frac{0.24(1-0.24)}{400}}} < \frac{0.233 - 0.24}{\sqrt{\frac{0.24(1-0.24)}{400}}}\right)$$

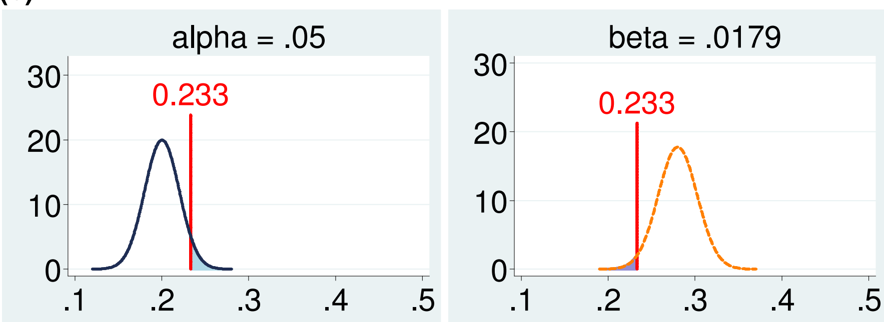
$$\beta = P\left(Z < \frac{0.233 - 0.24}{\sqrt{\frac{0.24(1-0.24)}{400}}}\right)$$

$\beta = P(Z < -0.33) = 0.37$

(b)



(c)

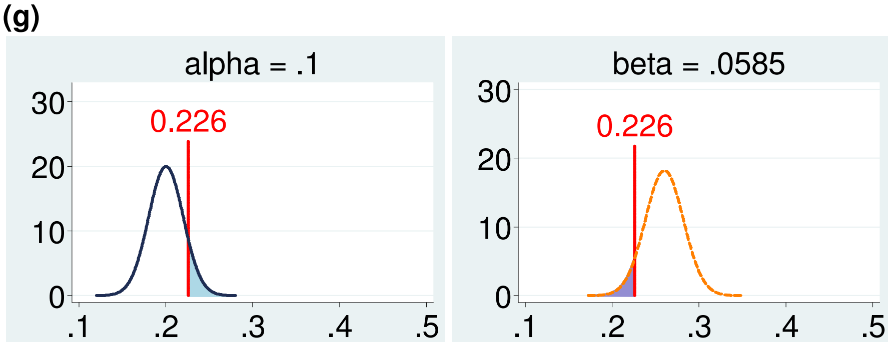


(d) The power increases from (a) to (c) because as the effectiveness of the ad improves – as a higher and higher fraction of the population recall the product – the chance that our random sample will contain a high fraction recalling the ad improves. As the sample proportion increases, our ability to reject the false null hypothesis (which says the fraction is small) improves.

(e) If half of the population recalls the ad and we sample 400 people we will almost surely obtain a high sample proportion that will fall deep into the rejection region and allow us to reject the false null hypothesis that the proportion is only 0.2. With the table you'd approximate the power as 1: the probability of a Type II error in this case is virtually zero (out past the 25^{th} decimal place!).

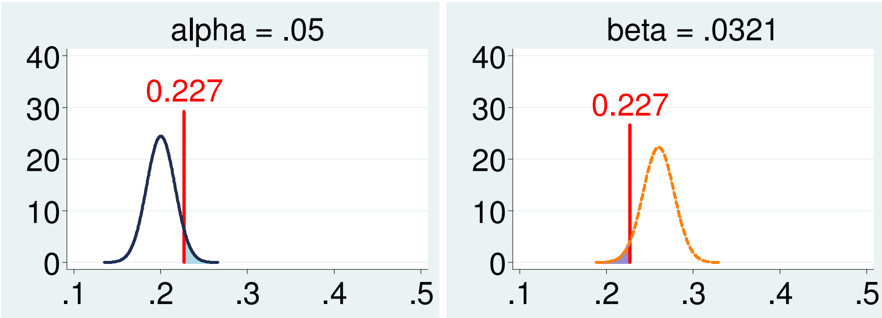
(f) If the proportion of people recalling the ad in the population is only slightly better than the null hypothesis – 20.5% versus 20% -- it is very likely that our sample statistic will be fairly small and provide insufficient proof of our research hypothesis. Power is only 0.083: there is a low chance that we will obtain a sample that will allow us to infer the research hypothesis is true (even though it is in fact true!).

(g)



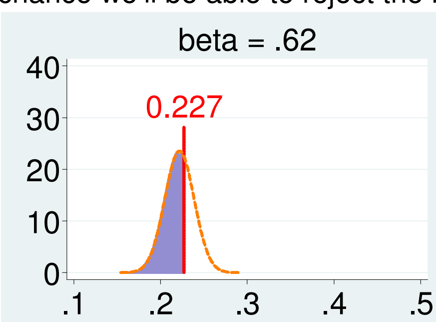
(h) See exercise 44.

(i)



(j) See exercise 44.

(k) There *is* evidence to suggest that the radio station has in fact *exceeded* expectations. The sample proportion, $\hat{P}$, is 22.2% (0.222=133/600), which of course is greater than 20% (expectations). Unfortunately, there is not enough evidence at a 5% significance level to conclude that our sample statistic of 22.2% may not reflect sampling error: that is, getting lucky with 22.2% when in fact the population proportion is only 20%. The problem is that if the population proportion is only a bit higher than expectations (i.e. 22.2% versus 20%) our hypothesis test will have very low power: it is unlikely that we will be able to reject the null hypothesis even though it is wrong and our research hypothesis is correct. To quantify our argument we calculate the power of the test if *we assume* that p = 0.222 and we see the power would be only 0.38 in that case: there is only a 38% chance we'll be able to reject the false null.



Hence the power of the statistical test is 0.63 (=1 – 0.37). This means that if we randomly sample 400 people and ask if they recall the product there is a 63% chance that the statistical test will lead us to conclude correctly that at least 20% recall the product if in fact 24% of the population do. On the dark side, there is a 37% chance that we will fail to find sufficient evidence to support our research hypothesis even though it is in fact true.

(3) **(a)** Choose a significance level. It was not specified so go with the conventional significance level: $\alpha =$ 0.05. Find the rejection region for this one-sided hypothesis test.

Standardized rejection region:

$$P(Z > 1.645) = 0.05$$

Un-standardized region:

$$P(\hat{p} > 1.645 * \sqrt{\frac{0.6(1-0.6)}{100}} + 0.60) = 0.05$$

$$P(\hat{p} > 0.6806) = 0.05$$

Find probability of Type II error (β) if p = 0.58:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.58) = P\left(Z < \frac{0.6806 - 0.58}{\sqrt{\frac{0.58(1-0.58)}{100}}} \right) = P(Z < 2.0383) = 0.9792$$

Find probability of Type II error (β) if p = 0.59:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.59) = P\left(Z < \frac{0.6806 - 0.59}{\sqrt{\frac{0.59(1-0.59)}{100}}} \right) = P(Z < 1.8421) = 0.9673$$

Find probability of Type II error (β) if p = 0.60:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.60) = P\left(Z < \frac{0.6806 - 0.60}{\sqrt{\frac{0.60(1-0.60)}{100}}} \right) = P(Z < 1.645) = 0.95$$

*Note: This is a bit strange because in this case the null hypothesis is NOT false. To understand it think about a value very close to 0.60 (such as 0.6001). In this case the probability of making a Type II error will be 1 - α.

Find probability of Type II error (β) if p = 0.61:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.61) = P\left(Z < \frac{0.6806 - 0.61}{\sqrt{\frac{0.61(1-0.61)}{100}}} \right) = P(Z < 1.4475) = 0.9261$$

Find probability of Type II error (β) if p = 0.62:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.62) = P\left(Z < \frac{0.6806 - 0.62}{\sqrt{\frac{0.62(1-0.62)}{100}}} \right) = P(Z < 1.2485) = 0.8941$$

Find probability of Type II error (β) if p = 0.64:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.64) = P\left(Z < \frac{0.6806 - 0.64}{\sqrt{\frac{0.64(1-0.64)}{100}}} \right) = P(Z < 0.8458) = 0.8012$$

Find probability of Type II error (β) if p = 0.66:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.66) = P\left(Z < \frac{0.6806 - 0.66}{\sqrt{\frac{0.66(1-0.66)}{100}}} \right) = P(Z < 0.4349) = 0.6682$$

Find probability of Type II error (β) if p = 0.68:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.68) = P\left(Z < \frac{0.6806 - 0.68}{\sqrt{\frac{0.68(1-0.68)}{100}}} \right) = P(Z < 0.0129) = 0.5051$$

Find probability of Type II error (β) if p = 0.70:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.70) = P\left(Z < \frac{0.6806 - 0.70}{\sqrt{\frac{0.70(1-0.70)}{100}}} \right) = P(Z < -0.4233) = 0.3360$$

Find probability of Type II error (β) if p = 0.72:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.72) = P\left(Z < \frac{0.6806 - 0.72}{\sqrt{\frac{0.72(1-0.72)}{100}}} \right) = P(Z < -0.8775) = 0.1901$$

Find probability of Type II error (β) if p = 0.74:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.74) = P\left(Z < \frac{0.6806 - 0.74}{\sqrt{\frac{0.74(1-0.74)}{100}}} \right) = P(Z < -1.3542) = 0.0878$$

Find probability of Type II error (β) if p = 0.76:

$$\beta = P(\hat{p} < 0.6806 \mid p = 0.76) = P\left(Z < \frac{0.6806 - 0.76}{\sqrt{\frac{0.76(1-0.76)}{100}}} \right) = P(Z < -1.8591) = 0.0315$$