Lab 5: Power and Sample Size
KEY

Javier E. Flores

March 8, 2019

Total Possible Points: 30

Interval Estimation

Q1) [2 pts] Using the information above, express n as a function of z_{crit}, MOE, and σ.

A1) Using the expression for the MOE and solving for n, we obtain the desired formula:

$$n = \left(\frac{z_{crit}\sigma}{MOE}\right)^2$$

Q2) [2 pts] Would the formula you derived in the previous question be valid if you swapped z_{crit} for t_{crit}? In other words, could the same general formula be used for t-distribution based confidence intervals AND z-distribution based confidence intervals? Why or why not?

A2) We can not swap in t_{crit}. The t-distribution (and hence t_{crit}) is dependent on the degrees of freedom, which is dependent on the sample size n. Therefore, the expression

$$n = \left(\frac{t_{crit}\sigma}{MOE}\right)^2$$

would not correspond to one in which we've solved for n. Finding an expression for n would be much more difficult.

Q3) [2 pts] Preliminary data on the regularly occurring, non-stop flight from Boston to San Francisco, United 433, is provided one the following page. The data describe the airtime in minutes for each of United 433's flights.

Airtime

353	351	377	348	402
358	380	370	351	388
374	346	372	359	381
360	369	356	369	
346	374	407	384	
373	385	356	368	
363	377	360	398	

In 2016, United 433 was scheduled to depart at 6AM EST and arrive before 10AM PST. Due to the time zone difference, the expected flight time is 420 minutes. Using the preliminary data above, determine how large a sample needs to be collected to achieve a 95% confidence interval estimate for the mean airborne time that has a margin of error of 2 minutes. Assume that the data are normally distributed.

A3) From the data, we may obtain an estimate of σ. Then, we plug in the appropriate values for z_{crit} and MOE in the formula derived in the first question:

$$n = \left(\frac{z_{crit}\sigma}{MOE}\right)^2 = \left(\frac{1.96 * 16.16}{2}\right)^2 = 250.8 \approx 251$$

Q4) [2 pts] Derive an expression for n as a function of $\hat{p}$, MOE, and z_{crit}.

A4) In this case, we have that

$$MOE = z_{crit}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Solving for n, we obtain the solution:

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{crit}}{MOE}\right)^2$$

Q5) [2 pts] Suppose you are interested in conducting a poll prior to the 2016 election in order to determine which presidential candidate, Hillary Clinton or Donald Trump, is preferred among Grinnellians. Our interest is in determining the proportion of Grinnellians that support Hillary Clinton. How large of a sample would be needed to estimate this proportion within 3% at the 95% confidence level? Assume that we have no preliminary information to inform what this proportion might be.

A5) Without any preliminary information, we assume each candidate is equally preferred (i.e. $\hat{p} = 0.50$. Using the previously derived formula:

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{crit}}{MOE}\right)^2 = 0.5(1-0.5)\frac{1.96^2}{0.03^2} \approx 1068$$

Q6) [2 pts] Assume that you found an earlier poll estimating the proportion of Hillary supporters to be 42%. How large a sample would be needed to achieve a 3% margin of error in this scenario?

A6) Here, we have preliminary information (i.e. $\hat{p} = 0.42$). Therefore,

$$n = \hat{p}(1-\hat{p})\left(\frac{z_{crit}}{MOE}\right)^2 = 0.42(1-0.42)\frac{1.96^2}{0.03^2} \approx 1040$$

Q7) [2 pts] Consider your answers to questions 5 and 6. Is there some other preliminary estimate that would lead to a larger sample size than what was found in question 5? If so, what is it? If not, why not?

A7) No other estimate for $\hat{p}$ would lead to a larger sample size than what we found in question 5. Intuitively, this can be explained by the fact that 0.5 corresponds to the greatest degree of uncertainty (i.e. we have no preliminary information) which should lead to the largest sample size. Mathematically, consider $\hat{p}(1-\hat{p})$. Plotting this curve for various values of $\hat{p}$, you'll notice that this expression is maximized when $\hat{p} = 0.50$.

Hypothesis Testing

Q8) [3 pts] Opening the app, what does the blue distribution represent? What does the green distribution represent? What do the dashed vertical lines represent?

A8) The blue distribution represents the distribution under the null hypothesis. The green distribution is the distribution assuming a specific effect different from the null value. The dashed vertical lines represent the critical values under the null distribution.

Q9) [2 pts] Why is the light blue region the probability of making a type I error? What is the dark green region the probability of making a type II error?

A9) Recall that the type I error is the probability of rejecting the null hypothesis when it is true. The light blue region is located outside the vertical dashed lines, which corresponds to all instances in which we would reject the null hypothesis.

Recall that the type II error is the probability of failing to reject the null hypothesis when it is false. The dark green region is between the vertical dashed lines, which corresponds to all instances in which we would fail to reject the null hypothesis.

Q10) [3 pts] Using the app, specify $\alpha = 0.05$, $\mu_1 = 0$, and $\sigma_1 = 1$. Create a table which records the power for each of the following effect sizes: 2.1, 2.5, 3, 3.5, 4, 4.5.

A10) The corresponding powers for each of the effect sizes listed are 0.555, 0.705, 0.851, 0.938, 0.979, and 0.994.

Q11) [3 pts] Using the app, specify $\alpha = 0.05$, $\mu_1 = 0$, $\sigma_1 = 1$, $\mu_2 = 3$, and $\sigma_2 = 1$. Create a table which records the power for each of the following significance levels: 0.01, 0.03, 0.05, 0.10, 0.15, 0.20.

A11) The corresponding powers for each of the significance levels listed are 0.664, 0.797, 0.851, 0.912, 0.941, and 0.957

Q12) [2 pts] With an increase in sample size, do you expect the power to increase or decrease? Why? (Hint: Think about how the standard error is related to the sample size!)

A12) With an increase in sample size, we should expect an increase in power. Remember that increasing sample size decreases the variability. Therefore, the distributions shown on the app will become more tightly centered on their means. With a reduced overlap between the green and blue distributions, the type II error reduced and the power is increased in turn.

Q13) [3 pts] In no more than three sentences, summarize the effect that individually varying the effect size, sample size, and significance level has on the power of a test. In other words, for each factor, describe how the power change by varying the factor and keeping all others constant.

A13) Increase any of effect size, significance level, or sample size will lead to an increase in power.