DA 6823
Kilger
Module 2: Part #2 (55 points)

Standard Error of the Estimate + Confidence Intervals + the Logic of Hypothesis Testing + Type 1 and Type II errors

General Instructions: In your own words, answer each of the following questions - don't copy (e.g. cut and paste) some definition out of a book word for word. This is not a group project – you are expected to complete this module on your own. You may refer to text books, online or other sources but not your fellow classmates. If you don't understand the question, feel free to ask the instructor in class, in office hours or in an email.

1. Explain in your own words what a type I error is (4 points)
   When H_o is correct but rejected by the person who is doing the study.

2. Explain in your own words what a type II error is (4 points)
   Opposite of Type I Error, when H_o is accepted or is not rejected when its false.
   you cannot accept a Hnull -1
3. Imagine that you are a cancer researcher who has developed a new test for cancer. Think about what a type I and type II error means for this kind of test. Argue for what you think is the most egregious error – a type I or type II error in this case. (Hint – you can logically argue for either case, just explain why). (6 points)
   H_o = The tumor test indicates that the tumor is a sign of cancer. H_a= The tumor test indicates the tumor is benign.
   Type I Error: The tumor is indicating that the subject has cancer, but the researcher declares the tumor as a benign tumor.
   Type II Error: The tumor test indicates that the tumor is benign, but the researcher declares the tumor to be an indication of cancer.

Type I Error is an egregious error as the test is indicating that the test subject is showing signs of cancer, but the researchers are not looking into it and possibly not gathering more data about cancer.

4. Explain in your own words what the power of a statistical test means (4 points)
   The possibility of correctly rejecting H_o when its false.

5. Name two things that can affect the power of a statistical test (4 points)
   Sample Size.
   How big is the effect.

6. Here are the six steps of hypothesis testing:
    i. State the null (H_{null}) and alternative (H_{alt}) hypotheses
    ii. State the assumptions of the test
    iii. Determine the critical value for the test statistic
    iv. Calculate the value of the test statistic from the data
    v. Compare the calculated and critical values for the test statistic
    vi. Apply the decision rule and interpret the result of the test

We will use a simple chi-square test as our example in this module. Here is the data that examines if there is a relationship between gender and format of book read:

type_book print or ebook read yesterday * sex SEX. Respondent's sex Crosstabulation

			sex SEX. Respondent's sex		
			1 Male	2 Female	Total
type_book print or ebook read yesterday	1.00 read print book	Count	419	635	1054
		Expected Count	407.1	646.9	1054.0
	2.00 read digital book	Count	19	61	80
		Expected Count	30.9	49.1	80.0
Total		Count	438	696	1134
		Expected Count	438.0	696.0	1134.0

The questions on the next page take you through each step applying the chi-square test to this data. Use a stats book and/or the Internet to help you with this but write your answers in your own words, not copy and paste.

a. State the null and alternative hypotheses for this test. (4 points)

H_{null} = Being Males or Females does not affect whether people read digital and printed books.

H_{alt} = Being Males or Females affects the amount of digital and printed books that are being read.

b. State at least one assumption for this test. (3 points)

Gender does affect the type of book.

no 1. the variables are both categorical or nominal in nature.2.

c. Determine the critical value of chi-square that your data will have to exceed in order to reject the null hypothesis. This involves calculating the degrees of freedom for our data as well as looking up the critical value in a chi-square table. Show your work for calculation degrees of freedom. (6 points)

Alpha level = 0.05

Degrees of freedom = (2-1)*(2-1) = 1

Chi-Square Critical Value = 1.05 = 3.841

d. Write out the chi-square formula and then using the data in the table provided above, calculate the chi-square value from the data. Show your work. (10 points)

$X^2 = \sum (O-E)^2/E$

Male Printed: $(419-407.1)^2/407.1 = 0.34785$

Female Printed: $(635-646.9)^2/646.9 = 0.21891$

Male Digital: $(19-30.9)^2/30.9 = 4.58285$

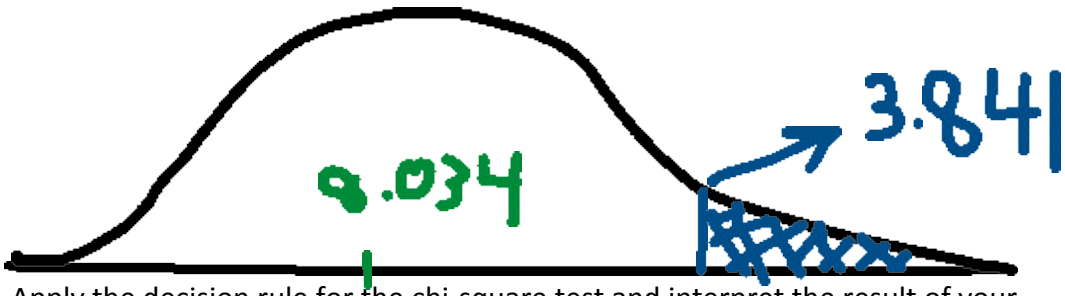
Female Digital: $(61-49.1)^2/49.1 = 2.88411$

0.34785 + 0.21891 + 4.58285 + 2.88411 = 8.03372

e. Compare the chi-square critical value and the chi-square value calculated from the data and draw a rough sketch of a chi square curve and place those two values on the curve. (5 points)

chi-square critical value = 3.841

Chi-square calculated value = 8.034



f. Apply the decision rule for the chi-square test and interpret the result of your analysis. (5 points)

8.034 ≥ 3.841

There is a relationship between gender and book type.

no - 8.034 line should be to the right of 3.841 line -2