Dan Schumacher, hdd249

DA 6823

Kilger

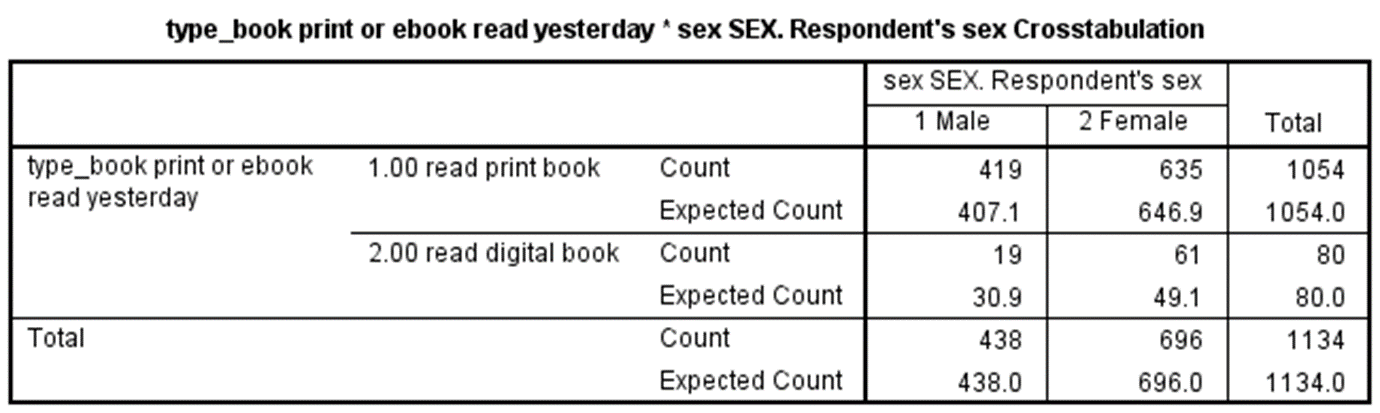
Module 2: Part #2 (55 points)

**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)
   1. **A False Positive.**
   2. **Hey, there is something here! (but there isn’t)**
   3. **Reject a H0 that is actually true**
2. Explain in your own words what a type II error is (4 points)
   1. **A False Negative**
   2. **Hey, there is nothing to see here! (but there totally is)**
   3. **Fail to reject a H0 that is actually false**
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)
   1. **It would be worse to have a type 2 error (false negative)**
   2. **I would rather be told that I have cancer and then, after further testing, find out that I don’t. The alternative is that I assume my health is fine, only to have the cancer grow. Reaction time when dealing with cancer is crucial! The younger you find it the better your odds of fighting it are.**
4. Explain in your own words what the power of a statistical test means (4 points)
   1. **The power of a statistical test is the percentage chance that you don’t have a type 2 error. Or the probability to correctly reject the Null.**
5. Name two things that can affect the power of a statistical test (4 points)
   1. **Alpha value**
   2. **Sample size**
6. Here are the six steps of hypothesis testing:
   * 1. State the null (Hnull) and alternative (Halt) hypotheses
     2. State the assumptions of the test
     3. Determine the critical value for the test statistic
     4. Calculate the value of the test statistic from the data
     5. Compare the calculated and critical values for the test statistic
     6. 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:



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.

1. State the null and alternative hypotheses for this test. (4 points)
   1. **H0: Gender and format of book are unrelated.**
   2. **Halt: Gender and format of book are related**
2. State at least one assumption for this test. (3 points)
   1. **There are 2 nominal variables.**
   2. **Each variable can be counted “each variable can be summarized as a count within one of the cells of the table” - MSDA bootcamp.**
   3. **Data must be substantial enough “80% or more of the cells in the table have to have an expected count of 5 or more and no cell has an expected count of less than 1” -MSDA bootcamp.**
3. 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)
   1. **Alpha = .05**
   2. **Df = (#row-1) \* (#cols-1)**
      1. **(2-1) \* (2-1) = 1 degree of freedom**
   3. **Whats the distribution**
   4. **critical value = 3.841** **A yellow rectangular object with black numbers

      Description automatically generated**
4. 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)
   1. A black background with white text

      Description automatically generated
   2. **Chi^2 = summation of [(observed – expected val)^2 all divided by expected value]**
   3. **Using this quadrant of the chart**

A white rectangular box with black numbers

Description automatically generated

* 1. **(((419-407.1)^2)/407.1) +**

**(((635-646.9)^2)/646.9) +**

**(((19-30.9)^2)/30.9) +**

**(((61-49.1)^2)/49.1) = 8.03371814988**

1. 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)
   1. 8.03 > 3.84
   2. A graph of a function

      Description automatically generated
2. Apply the decision rule for the chi-square test and interpret the result of your analysis. (5 points)
   1. We reject H0 and accept the Halt.
   2. Gender and format of book are related.