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)**

A type I error can be viewed as a false positive and occurs when a tester incorrectly rejects a true null hypothesis. This may lead to reporting that findings are significant when they have actually occurred by chance. Risk of committing a type I error can be reduced by using a lower p-value.

1. **Explain in your own words what a type II error is (4 points)**

A type II error can be viewed as a false negative and occurs when a tester fails to reject a null hypothesis which is really false. In this situation, the tester concludes that there is not a significant effect when in reality there may be. Risk of committing a type II error can be reduced by ensuring the test being performed had enough power (large enough sample size).

1. **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)**

**Most Egregious Error:** Type II Error

**Rationale:** A Type II error would mean that a false negative cancer test result has occurred. This means that the cancer researcher performed a test, failed to reject a null hypothesis, and concluded that there was not a significant effect. In reality, there may have been a significant effect which would have resulted in a positive cancer test.

1. **Explain in your own words what the power of a statistical test means (4 points)**

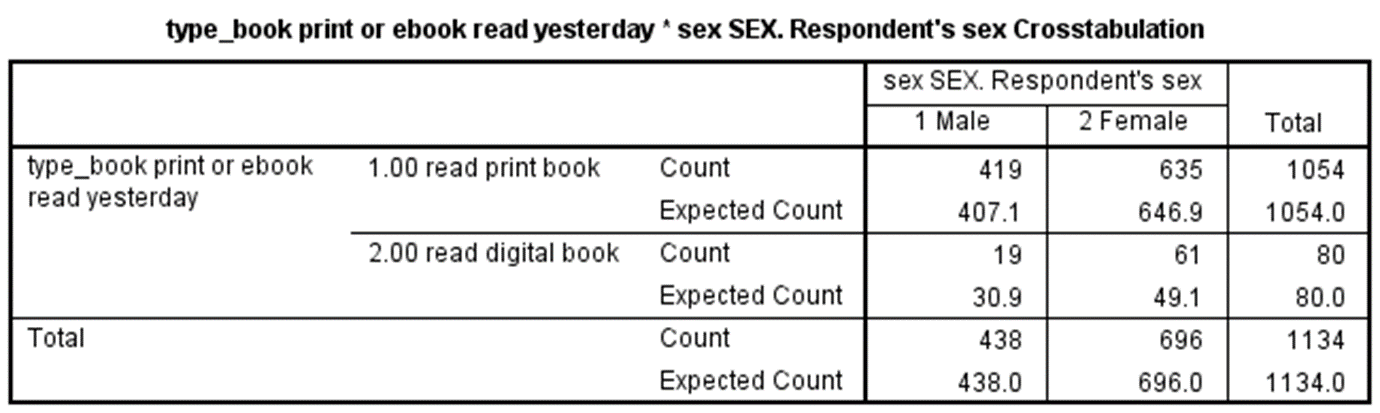
The power of a statistical test is the probability that the test correctly rejects the null hypothesis (the probability of a true positive result).

1. **Name two things that can affect the power of a statistical test (4 points)**

* Significance level
* Sample size

1. **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)**

**Null:** There is no relationship (no significant difference) between gender and format of book read (the variables are independent)

**Alternative:** There is a relationship (significant difference) between gender and format of book read (the variables are dependent)

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

* The data in the cells should be frequencies (counts) of cases rather than percentages or some other transformation of data.
* The categories of the variables are mutually exclusive (meaning that a particular person fits into one and only one category of each of the variables).

1. **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)**

Critical Value = 3.841

* **Calculate the number of degrees of freedom**
  + There are two book formats (print and digital), so 2 -1 = **1 degree of freedom**
* **Find the probability that the phenomenon would occur by chance**
  + P-value of 0.05 or a 5% probability
* **Look up the degrees of freedom and probability in the chi-square table**
  + One degree of freedom and 5% probability is the 3.841 in the chi squared table.



1. **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)**

Chi Square Value = X^2 = 8.03

Text, letter

Description automatically generatedA picture containing text

Description automatically generated

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)**

* Critical value = 3.841
* Chi-square value = 8.03

Diagram

Description automatically generated

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

Chi-Square Value (8.03) > Critical Value (3.841). Therefore, we reject the null. There is a relationship (significant difference) between gender and format of book read (the variables are dependent).