Project 1
Test a Perceptual Phenomenon - The Stroop Effect

Thomas Lindstrom-Vautrin
09/05/16

1. *What is our independent variable? What is our dependent variable?*

Independent Variable: Category (Congruent vs Incongruent)
Dependent Variable: Response Time (Seconds)

2. *What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.*

Null Hypothesis (H_0):

On average, there is no difference between response time for congruent and incongruent sets.

μ_I = mean response time for incongruent sets
μ_C = mean response time for congruent sets
μ_D = mean difference between response times for incongruent and congruent sets (we will also refer to this as the mean **delay** in incongruent sets vs congruent sets)

$\mu_I = \mu_C$
$\mu_I - \mu_C = \mu_D = 0$

Alternate Hypothesis (H_A):

On average, there is a positive difference between response time for congruent and incongruent sets. In other words, the response time for incongruent sets is longer than the response time for congruent sets.

$\mu_I > \mu_C$
$\mu_I - \mu_C = \mu_D > 0$

Statistical Test:

A Student's t-test appears to be the most appropriate test for this experiment as we don't know the population standard deviation and must make inferences based on the sample we have. We will perform the t-test with respect to the distribution of the mean differences (μ_D).

3. *Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.*

Measure of Central Tendency:

The **mean** of the sample delay in incongruent sets vs congruent sets ($\bar{x}_D$) is:

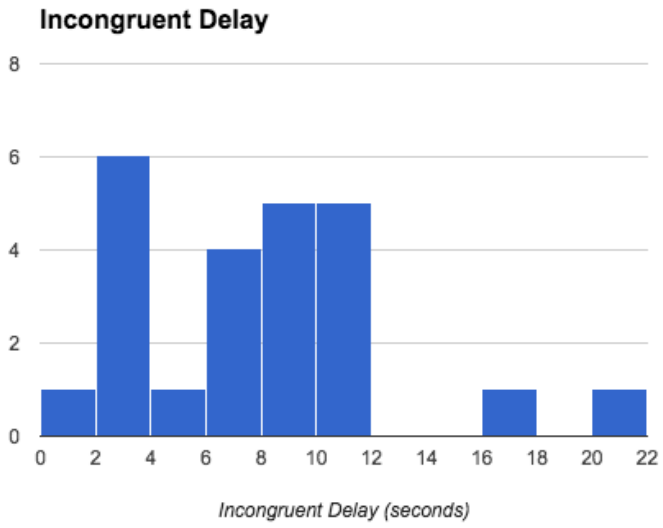
$\bar{x}_D$ = **7.96**

Measure of Variability:

The **standard deviation** of the sample delay in incongruent sets vs congruent sets (σ_D) is:

σ_D = **4.76**

4. *Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.*



This plot skews right with slower outliers between 16 and 18 as well as 20 and 22 seconds. It is somewhat bimodal with a big jump between 2 and 4 seconds and a second wider jump between 6 and 12 seconds. There is no delay below zero.

5. *Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?*

First we must calculate the t-score:

$\bar{x}_D = 7.96$
$\mu_D = 0$
$n = 24$
$S = \sqrt{\frac{\Sigma(x_i - \bar{x}_D)^2}{n-1}} = 4.86$
$t = \frac{\bar{x}_D - \mu_D}{S/\sqrt{n}} = 8.02$

Next we must compare this t-score to our t-critical values for a one-tailed t distribution using the t-table. We see that with n-1 = 23 degrees of freedom (*df*) a tail probability (*p*) of .001 occurs at t-critical value 3.485. Since our t (8.02) significantly exceeds this t-score, we can conclude with more than 99.9% certainty that our sample's greater-than-zero delay for incongruent sets is ***not*** the result of randomly selecting this sample from differences drawn from a large pool of individual response times to tests where we didn't identify the difference between incongruent and congruent sets (meaning some people may have taken two congruent sets, some people may have taken two incongruent sets and the difference may be calculated congruent-incongruent as opposed to incongruent-congruent). Thus **we have rejected the null hypothesis**. This means that the difference between performance on incongruent and congruent tests is significant and we are over 99.9% certain that people in general perform slower on an incongruent test than on a congruent test. These results matched what was expected. It is more difficult to say out loud the colors of words which spell a different color.