P1: Test a Perceptual Phenomenon

Author: Brendan Schell

Date: October 3, 2015

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

The independent variable is the words condition (congruent or incongruent set of words). The dependent variable is the time it takes in seconds to name all of the words' ink colours.

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

An appropriate set of Hypotheses for this task would be as follows:

$$H_0: \mu_{congruent-incongruen} = 0$$
$$H_1: \mu_{congruent-incongruen} < 0$$

Where $\mu_{congruent-incongruent}$ is the mean difference between the time taken to perform the test for congruent words and incongruent words for the population. The null hypothesis is that for the population, the mean amount of time it takes to name the colours of the congruent words is equal to that of the incongruent words. When the mean difference is calculated as per the definition above, this means that the mean difference for the population is equal to 0. The alternative hypothesis is that the mean difference for the population is less than 0. In the alternative hypothesis, it takes less time for the population to name the colours of the congruent words than it does for the incongruent words.

I expect to perform a one-tailed dependent t-test using the mean difference of the data samples to determine whether there is a difference of statistical significance in the amount of time taken to read the congruent and incongruent data for the population.

These would be an appropriate set of hypotheses since it is not known definitively that it takes more time to correctly name incongruent words, but it is the result I expect from this experiment. The t-test would be a suitable test to determine the significance of the results since it can determine significance without known population parameters while the population mean can be approximated as the sample mean.

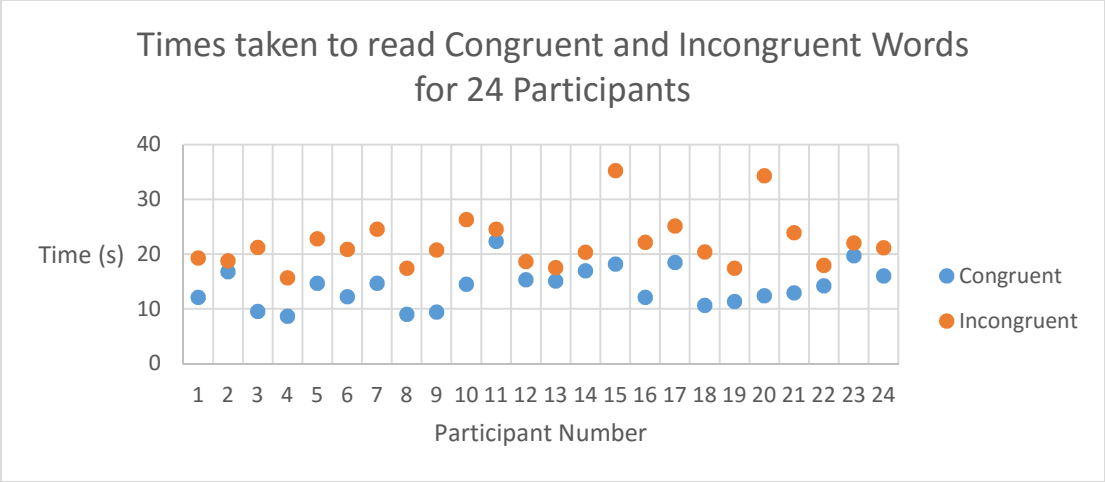
Since the data is taken from the same participants, the data is dependent and will use the parameters for a dependent t-test. A confidence level of 95% will be used.

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

For this dataset, the means and standard errors for the differences respectively are as follows:

$$\bar{x}_{congruent-incongruent} = -7.96$$
$$s_{congruent-incongruent} = 4.86$$

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.**



In the above plot demonstrating the times taken in each sample to read both congruent and incongruent data, it can be seen that it took more time in every instance for a user to read the incongruent words compared to the congruent words. The difference also appears to be multiple seconds in each case, indicating that there could be a significant difference in the results for incongruent vs congruent words.

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?**

Using a confidence level of 95%, a critical t-value of -1.714 was determined for a one-tail test. Since $t_{congruent-incongruent}$ was calculated to be equal to -8.02, this is significantly lower than the criterion value of -1.714. As a result, the null hypothesis is rejected. It can be concluded that there is significant evidence to demonstrate that it takes less time to read the congruent words than the incongruent words for the Stroop test. This result matches up exactly with my expectations and is what would be assumed based on the above plot.

6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions.

I believe that the human brain creates internal shortcuts to access information more quickly and reduce the time to perform certain processes (like how a python dictionary reduces retrieval time). It likely associates the word with the colour so that it can be retrieved more quickly. As a result, the brain sees the visual cue from the word's letters and does not need to process the colour if it is the same. When the association does not actually match, the brain needs to disassociate the connection and process the colour separately from the word (a more time-consuming process since it needs to do two different procedures). A similar process (using a different type of shortcut) would likely result from testing time to name each letter in the alphabet forwards versus backwards. Since the brain has developed a forwards-only association for

retrieving this information, it takes much less time to do this forwards than backwards, even though the same amount of information is being retrieved.