Project 1- Statistics: The Science of Decisions

Questions for Investigation - Stroop task

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

Our independent variable is the condition of the word (e.g. congruent words or incongruent words) and the dependent variable is the time it takes by the individuals to name the ink colors in a list. This dependent variable is one variable that was defined to measure the performance of each individual for naming the word's inks (construct).

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 are:

H0 (hypothesis null): The condition of the word does not affect the time it takes for the individuals to name the ink colors (The time of the first and second column are the same).

H0: Time for reading incongruent words - Time for reading congruent Words = 0

H1(alternative hypothesis): The incongruent conditions reduce the reading performance of individuals comparing to congruent conditions.(A list with incongruent words require a longer time for reading comparing to a list with only congruent words.)

H1: Time for reading incongruent words - Time for reading congruent Words > 0

The statistical test we should perform is Pre-test, Post-test, dependent-samples t-test, one-tailed test, testing two different conditions. In this experiment, the same sample of people has made the same test with two different circumstances (measuring the time to read a list with only congruent words and other with only incongruent words). Moreover, we do not know any population parameters, what indicates the necessity of using t-tests. Finally, the option for one-tailed test occurs due to the alternative hypotheses voiced, stating that the incongruent condition might INCREASE the time for reading the list.

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

Descriptive statistics	Incongruent List	Congruent List	Difference
Mean	22,016	14,051	7,964
Median	21,018	14,357	7,667
Standard Deviation	4,797	3,560	4,865
Variance	12,699	23,0118	23,667
Cohen's			1,637
r^2			0,259

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.

The two graphics below indicates, respectively, the time taken by each individual in the sample to read the incongruent list and the congruent list. Either graphics have the same scale in both axes and the same size in order to enable a better comparison between them. It is striking that the ILRT's histogram is moved to the right when compared to the CLRT's graphic. This might indicate a longer time demanded to read. This assumption is going to be tested on the fifth question, a little bit further on this document.
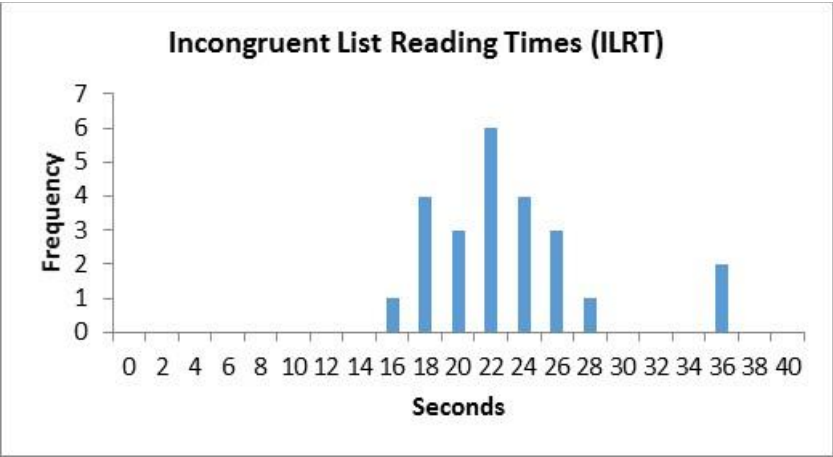

Graphic 1 - Incongruent List Reading Times Histogram
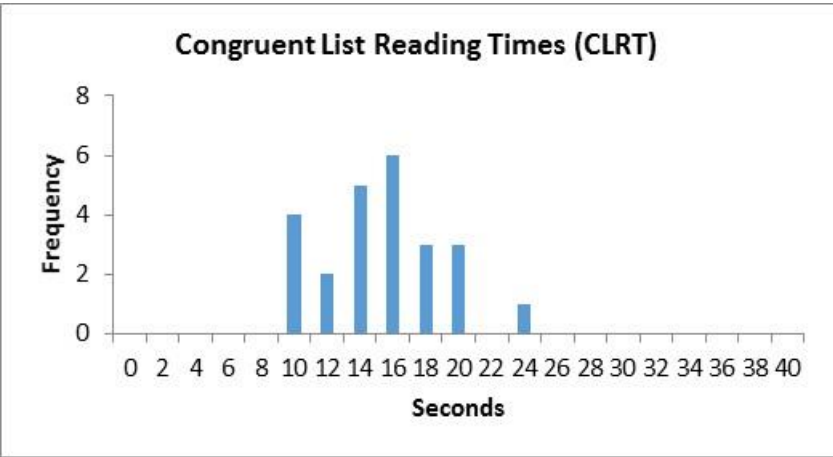
Graphic 2 - Congruent List Reading Times Histogram

The graphic below show us the difference of the time needed for individuals to read the incongruent list and the congruent list. It is important to highlight that there are not any negative values on the sample, what indicates that all the users had a better performance reading the incongruent list comparing to the congruent one.
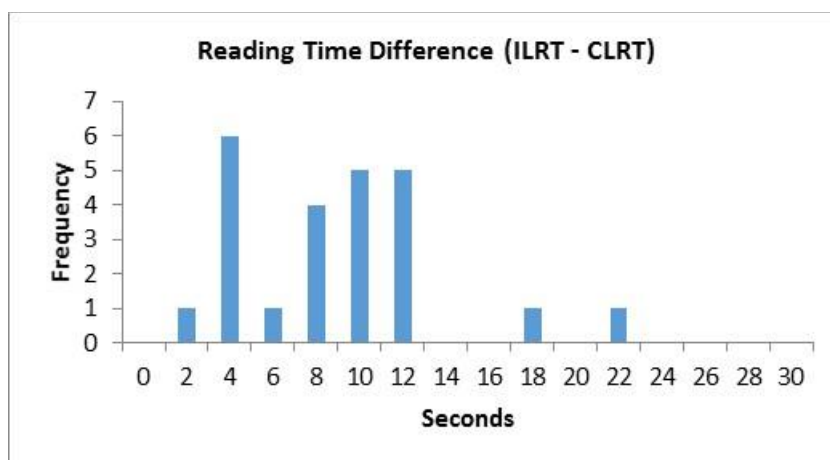
Graphic 3 - Reading Time Difference Histogram

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?

For a confidence level of alfa = 0.0005 and degrees of freedom = 23:

Considering a one-tailed test, according to what it was stated in the first answer:

T Critique	3.768
T Value	8.021

Once the T Value is bigger than the T Critique, we should refuse the null hypothesis with a confidence level of alfa = 0.0005. Indeed the presence of incongruent words hurts the performance of the individuals that read the ink used in which word, as it is expected on beforehand (Since the first time I have tried to do it, I had this impression). This performance of individuals is a construct that is being measured using the needed time to finish naming each of the inks used for writing a list of words.

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 main reason for this significant difference is the faster human processing speed to read a text comparing to identifying a color. Therefore, the individuals end up naming the color according to what the brain did before (reading the word), taking more time to figure out the name of the ink.

A similar task that might present the same behavior is to create two equally-sized words lists, that the first one has geometrical shape's names (square, circle, rectangle, pentagon, hexagon, etc.) inside the same geometrical shaped boxes (congruent list) and the second one has shape's names inside the different geometrical shaped boxes (incongruent list). If you ask someone to name out loud each of the boxes' shapes, probably this person might take more time to name all of them in a incongruent list than in a congruent list.

Biography: https://en.wikipedia.org/wiki/Stroop_effect