Project 1- Statistics

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the *color of the ink* in which the word is printed. The task has two conditions: a congruent words condition, and an incongruent words condition. In the *congruent words* condition, the words being displayed are color words whose names match the colors in which they are printed: for example RED, BLUE. In the *incongruent words* condition, the words displayed are color words whose names do not match the colors in which they are printed: for example PURPLE, ORANGE. In each case, we measure the time it takes to name the ink colors in equally-sized lists. Each participant will go through and record a time from each condition.

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

Independent Variable->the condition that each participant will go through

Dependent Variable-> time it takes to name the ink colors.

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

Here we are using samples from two different populations, i.e., the time it takes to name the ink colors in equally-sized lists under congruent words condition and incongruent words condition.

The null hypothesis is that the two population means will be the same, meaning the times taken by participants to name the ink colors will be similar irrespective of the condition used.

H_0: $\mu_{congruent} = \mu_{incongruent}$

Whereas the alternative hypothesis states that the condition used will influence the time measures. Specifically, the time taken to name the ink colors is more under incongruent word condition.

H_A: $\mu_{congruent} < \mu_{incongruent}$

Based on the above hypotheses, I will conduct a one- tailed t-statistical test in the positive direction. This is because I expect $\mu_{incongruent}$ to be greater than $\mu_{congruent}$. I am using a t-statistical test because I have no details about the population parameters and I am dealing with samples here.

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

Mean of congruent sample, $\mu_{congruent} = 14.05$

Standard deviation of congruent sample, $SD_{congruent} = 3.56$

Mean of incongruent sample, $\mu_{incongruent} = 22.02$

Standard deviation of incongruent sample, $SD_{incongruent} = 4.79$

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.

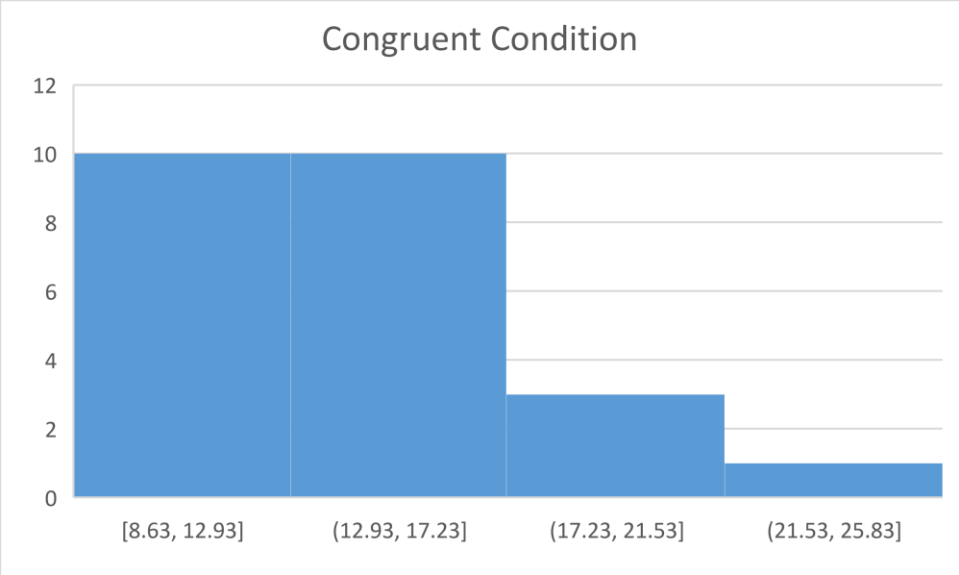

Fig1: Distribution of Congruent condition times

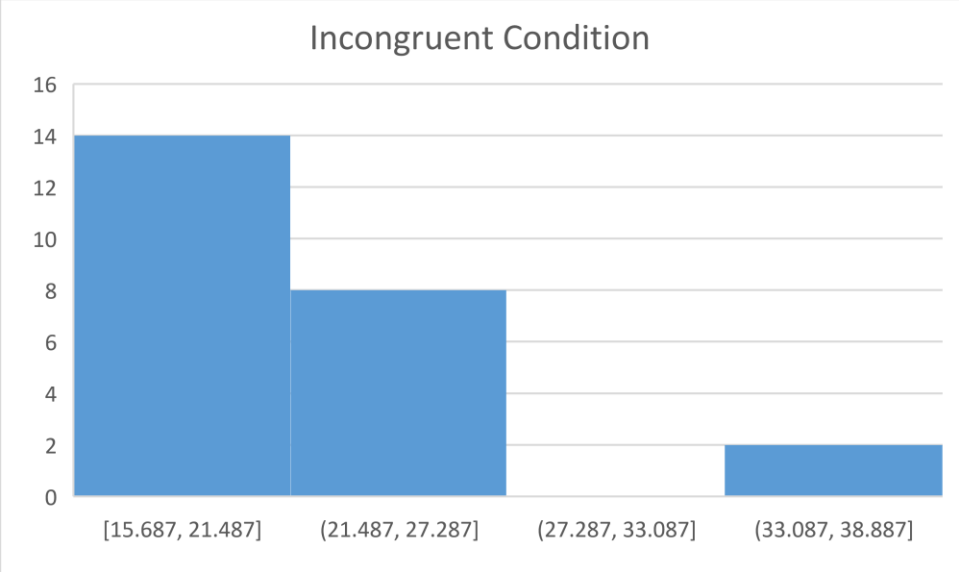

Fig2: Distribution of Incongruent condition times

Both the plots show a positively skewed distribution of sample data.

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?

Point Estimate = $\mu_{incongruent} - \mu_{congruent}$ = 22.02 - 14.05 = 7.97

Standard deviation of differences, S = 4.86

t- statistic = $\mu_{incongruent} - \mu_{congruent}$ / (S/ sqrt(n))

= 7.97/ 4.86(sqrt(24))

= 8.05

Degrees of freedom, df = n-1 = 24 -1 =23

t- critical value for α = 0.05 is 1.714

Since the t-statistic value is way above the t-critical value, I **reject** the null hypothesis. In other words, time taken to name the colors in incongruent condition is more than the time for congruent condition.

Also, p-value, p = 0.0001. This is statistically significant since p< α.

Cohen's d = mean difference/ SD = 7.97/ 4.86 = 1.64

Confidence Interval: (7.97-1.174(0.99), 7.97+1.17(0.99))

= (6.81,9.13)

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!

Our brain takes more time to process a color than a word, i.e., it identifies the word first then the color of the word. Under congruent condition, the brain only needs to process the word since it matches with the color and hence takes less time. But under incongruent condition, the color of the letters needs to be identified which is done only after identifying the word. This leads to confusion and requires more time.

Many variations of stroop test are in existence which display similar results as the original stroop test. One of them is warped words stroop test which follows the same conditions as the original test in addition to the words being printed in such a way that it's difficult to read.

Note: GraphPad software was used to calculate the p-value.

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