The Stroop Effect

Background Information

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?

Ans:

Independent Variable ➔ The condition of the words (Congruent and Incongruent)
Dependent Variable ➔ The time the user takes to say the correct words out loud.

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

Ans:

Ho - **Null Hypothesis**: There is not much difference in the populations' average response time of both congruent and incongruent words. According to the null **hypothesis**, any observed difference in samples is due to chance or sampling error.
H0: μi ≤ μc (μi - population mean of incongruent values, μc - population mean of congruent values)

Ha - **Alternative Hypothesis**: There is a significant difference, positive or negative, in the populations' average response times for congruent and incongruent words.
H1: μi > μc (μi - population mean of incongruent values, μc - population mean of congruent values)

The **Dependent Samples t-**Test is the appropriate statistical test as all the people who take the test are tested under two different conditions,

I will be using a t-test instead of a z-test because 1) the population standard deviation is unknown and 2) the sample set is less than 30. The t-test will be a one tailed t-test i.e. my directional alternative hypothesis is that participant's incongruent sample mean will be larger than the participant's congruent sample mean

A paired t-test (or dependent sample test), will be used because the data set is of one group of participants tested twice under different conditions (word/colour congruency). This will also facilitate either rejecting or accepting the null hypothesis.

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

Ans:

The no. of subjects in each case(population):

- Congruent→24
- Incongruent→24

The mean of the population:

- Congruent→14.0511
- Incongruent→22.0159

The median of the population:

- Congruent→14.3565
- Incongruent→21.0175

The Variance for each case:

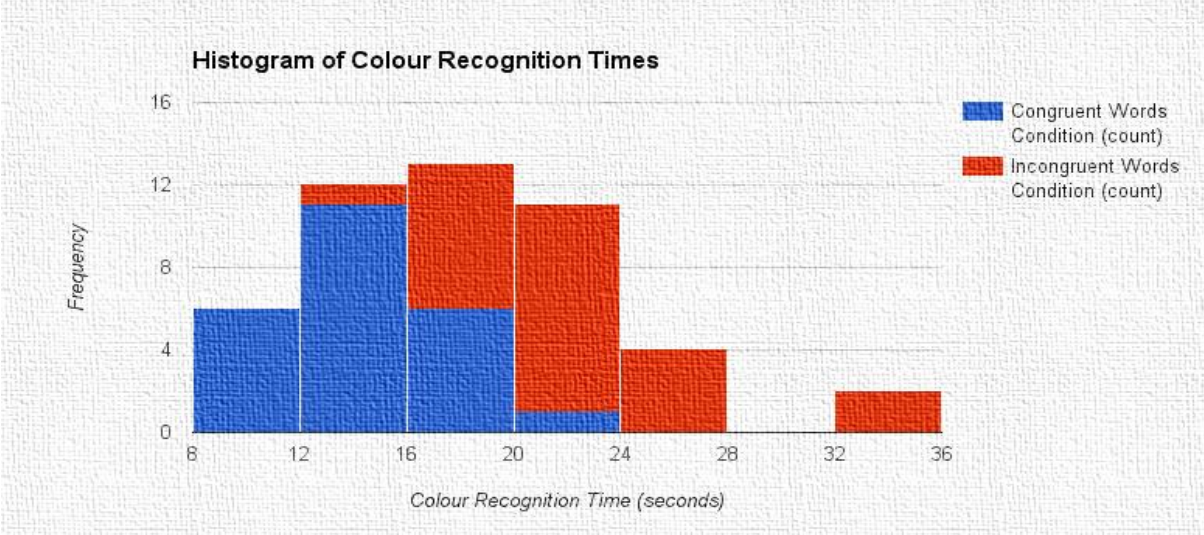
- Congruent→12.6690
- Incongruent→23.0117

The Standard Deviation for each case:

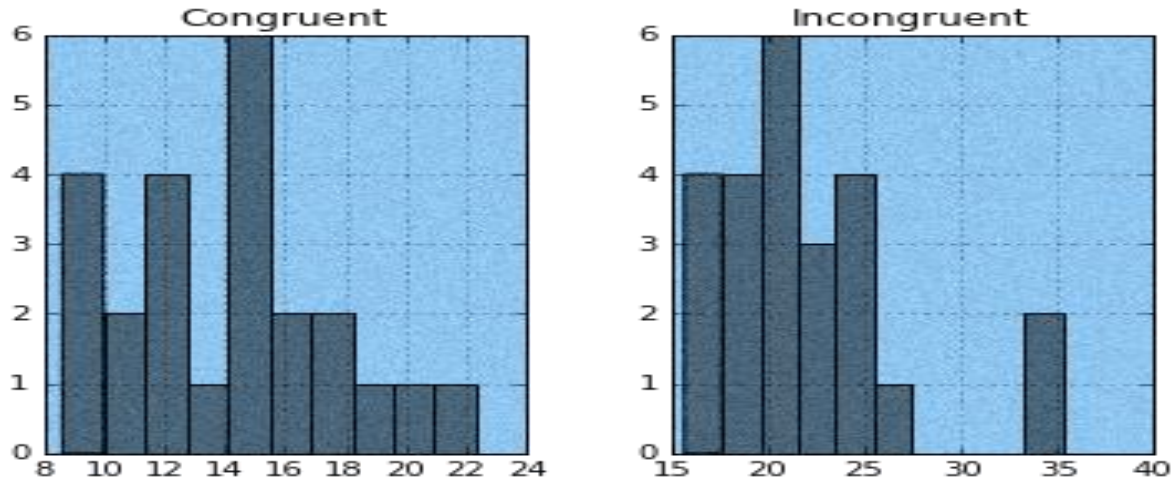
- Congruent→3.5593
- Incongruent→4.7970

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.

Ans:



The bucket range of the congruent words condition data is 8 seconds to 24 seconds. The bucket range of the incongruent words condition data is 12 seconds to 36 seconds, i.e., noticeably to the right of the congruent words condition data.



Above is a histogram which show the distribution of data from both congruent and incongruent conditions.

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?

Ans:

Mean Difference $x = \sum (iw - cw)/n = 7.964$

iw is incongruent, cw is the congruent values, n is the sample set.

Standard Deviation sd = 4.86

Standard Error of Difference $SE = sd/\sqrt{n} = 4.86/\sqrt{24} = 0.99$

t-statistic $T = x/SE = 7.964/0.99 = 8.04$

t-distribution with n-1 degrees of freedom (df = 23). Using the t-distribution table to find p-value…

The value of **p** is < 0.0001. The result is significant at p < 0.05%

Hypothesis:

I reject the null hypothesis, the incongruent words cause a greater response time.

Conclusion:

The results match my expectations.