Data Analyst Nano Degree Program

Project P1

Test a Perceptual Phenomenon

In this project we analyze the experimental psychology (viz.) the Stroop Effect. In psychology, the **Stroop effect** [1] is a demonstration of interference in the reaction time of a task. When the name of a color (e.g., "blue", "green", or "red") is printed in a color that is not denoted by the name (e.g., the word "red" printed in blue ink instead of red ink), naming the color of the word takes longer and is more prone to errors than when the color of the ink matches the name of the color. The effect is named after John Ridley Stroop, who first published the effect in English in 1935. The effect had previously been published in Germany in 1929. The original paper has been one of the most cited papers in the history of experimental psychology, leading to more than 701 replications. The effect has been used to create a psychological test (**Stroop test**) that is widely used in clinical practice and investigation.

In the following analysis, Congruent test means printing the name of the color in the same color as the name and finding out the Reaction time. On the other hand, Incongruent test means, printing the name of the color in a color NOT denoting the color to find out the Reaction time.

The dataset "stroopdata" contains the reaction times for a sample population of 24. The reaction time when the color of the ink matches the name of the color is given in column A (Congruent). And the reaction time when the name of the color is printed in a color that does not denote the name is given in column B (Incongruent).

The dataset can be analyzed as Dependent samples t-test for paired samples; this will be a Within-subject design test with Two conditions under Repeated measures design test (viz.) Reaction time for two types of display of the name of a color (Congruent and Incongruent).

Question 1: Identify variables in the experiment:

Answer:

Reaction time is the dependent variable and the types of displays of the name of the color (Congruent and Incongruent) are the independent variables.

Question 2a: Establish hypotheses

Answer:

The Stroop effect is the finding that naming the colors of color words (e.g. the words 'green', 'red', 'blue', etc.) is easier and quicker if the actual observed colors of the words match the colors that the words denote (e.g. the colors green, red, blue, etc., respectively) than if they do not match.

Reaction time is the time the participant takes to name the color. The printing of the name of a color in the same color is called the Congruent test and the printing of the name of a color in a color that does not denote the color is called the Incongruent test.

What could be the null hypothesis?[2]

A Null hypothesis is a hypothesis that says that there is no statistical significance between the two variables in the hypothesis. It is the hypothesis that the researcher is trying to disprove.

Null Hypothesis:

There is no significant difference between the population reaction times between the Congruent and Incongruent tests, if the tests were extended to everybody.

Note: [3] [3a]Writing Equations and special symbols

If **μc** is the population Mean of reaction time for the Congruent test and **μi** is the population Mean of reaction time for the Incongruent test, then, the null hypothesis can be represented as below:

H_0: (μc – μi) = 0

What could be the alternative hypothesis?

The alternative is the opposite of the null hypothesis. (i.e.) there is a statistical significance between two variables in the hypothesis.

Alternative Hypothesis:

There is significant difference between the population reaction times for the Congruent and Incongruent tests, if the tests were extended to everybody, not just the samples.

It can be represented as below:

H_A: (μc – μi) ≠ 0

The idea of hypothesis testing and statistical inference is that we have limited data, **samples** (generally denoted by X), and from that limited data, we are trying to infer something about the **population** (generally denoted by U) we don't know about.

Question 2b: Establish a statistical test:

Answer:

We will do a t-test for the following reasons:

i) We have less than 30 samples.
ii) The population parameters are not available. Based on this fact and based on the above hypotheses, we will conduct a two-tailed t-test.
iii) We assume that the distributions are Gaussian. ("a typical assumption is that the continuous values associated with each class are distributed according to a Gaussian distribution.") [9]

Question 3: Report Descriptive Statistics

Answer:

Since our point estimates are based on our samples, our point estimate for the difference (**D**) will be:

D = 14.05 – 22.02

= -7.97

Then we calculate the square of the difference D and calculate the Variance by summing up the Squares and dividing by (n-1)

SUM (D2:D25)/23 = 23.67

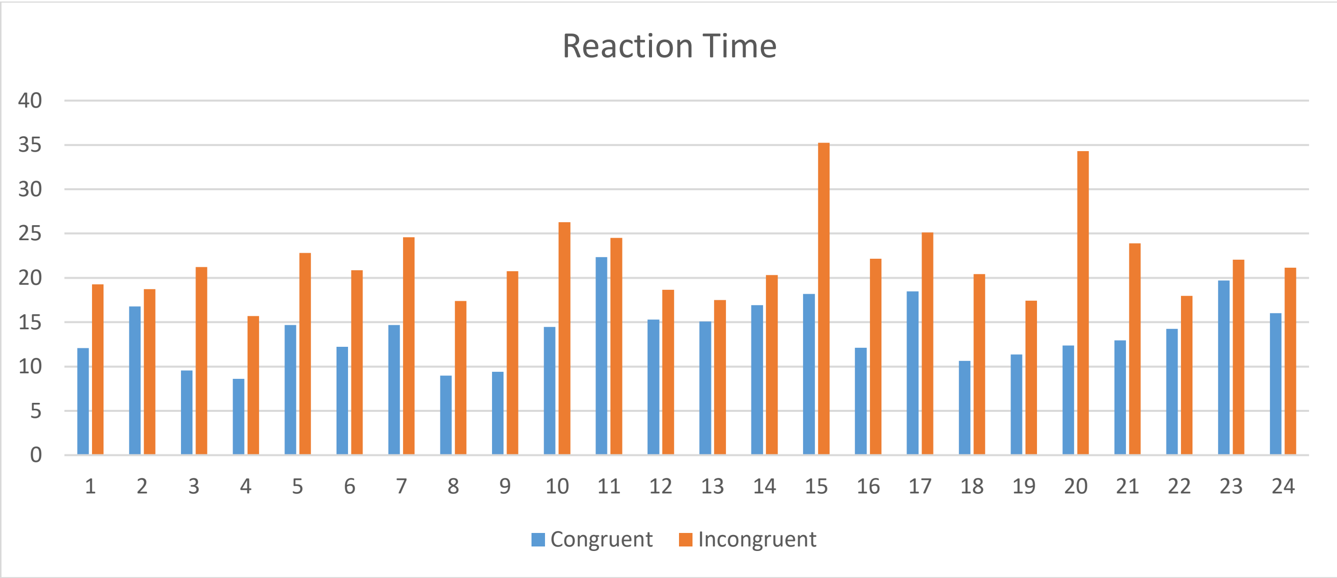
The Square Root of this variance is the Standard Error of the differences denoted by S

S = 4.86

Question 4: Plot the data

Answer:

Plot1: Bar Chart of Reaction Time of Congruent and Incongruent tests



Depicts that the Reaction Time for Incongruent test is always higher than that for Congruent test throughout.

Plot 2: Reaction Time as a Box Plot:

This box plot depicts the relative positions of the values of the two tests. The first quartile, Mean and third quartile values of the Incongruent data lie above the third quartile value of the Congruent data. There are two outliers in the Incongruent data (34.288 and 35.255).

Question 5: Perform the statistical test and interpret your results.

Answer:

The statistical test was performed and the results below.

Results Section [4]:

Inferential Statistics:

Hypothesis test (α = 0.05)

n denotes the sample size = **24**

df (degree of freedom) = **n -1** = **23**

S denotes the Standard Error of the differences OR Standard Deviation using (n-1) as denominator.

S = 4.86

Now let us calculate the t-statistic:

$$t = (\mu_c - \mu_i)/(S/\sqrt{n})$$

$$= (14.05112500 - 22.01591667) / (4.86/\sqrt{24})$$

= - 7.96479167/ (0.992043346)

= - 8.02867

= - 8.03 (rounding off to 2 decimals)

Let us read t-critical for α =.05 and df = 23 (read from the t-table) [5]

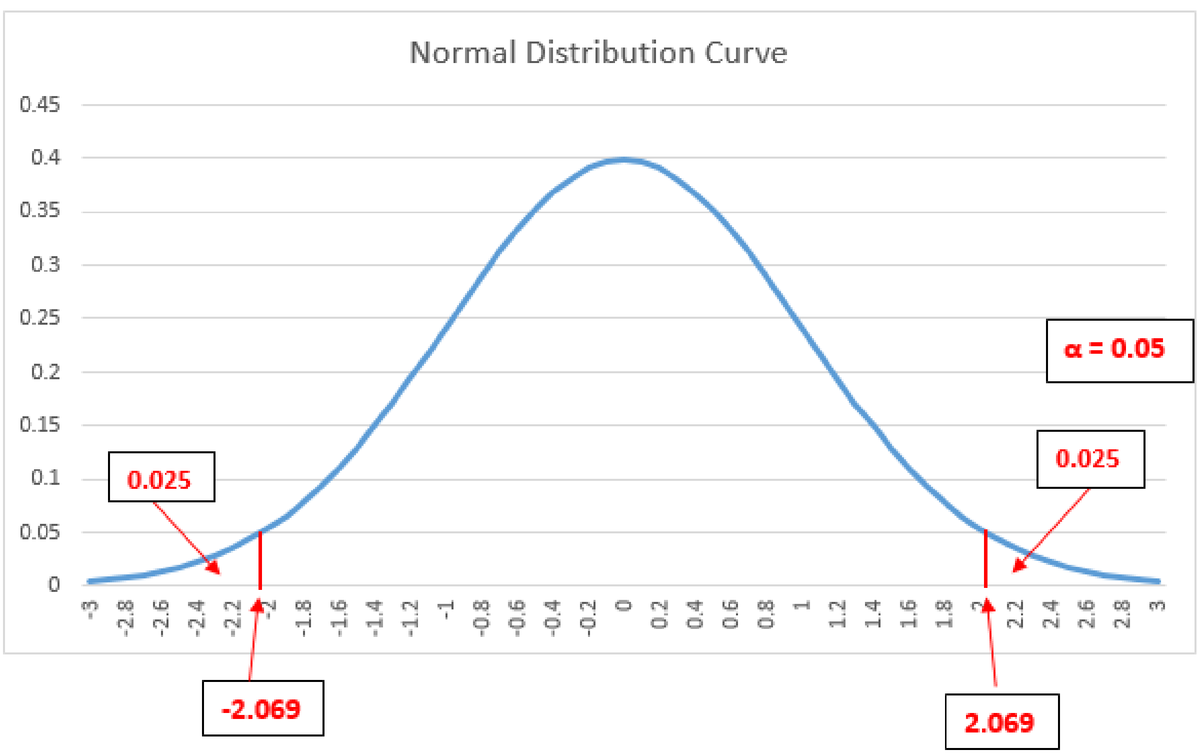
We should remember that the α = .025 for each tail.

t-critical = ± 2.069

t (23) = ± 2.069, P < 0.05, two-tailed

Plot 3:

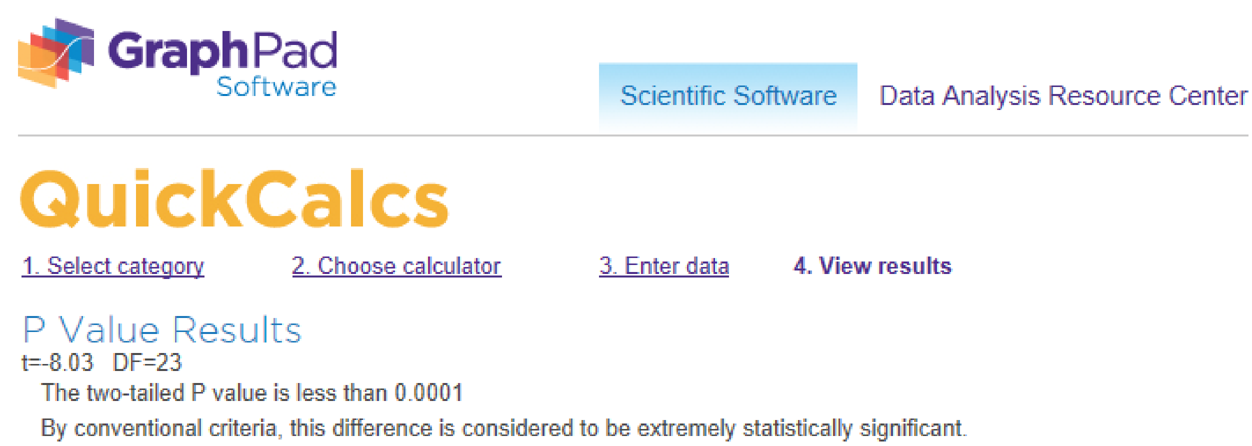
A Normal Distribution curve with the α =0.05 and the t-critical values on either side. The areas beyond the t values ± 2.069 are in the **Critical regions**.



Calculation of P value [6]:

The two-tailed P value for t-statistic = -8.05 and df=23 is less than .0001

The screen shot below shows that the difference is considered to be extremely statistically significant.



Since t-statistic is past the t-critical value and falls in the critical region, we reject the null hypothesis.

This means the reaction limit is significantly lower in the case of Congruent test that the reaction time in the Incongruent test.

Since this is an experimental design, we can make a causal statement that the type of test has a causal effect on the reaction time.

CI (Confidence Level)

$$\textbf{CI} = \mathbf{M_D \pm t_{critical}\,(S_D/\sqrt{n})}$$

$$= -7.97 \pm 2.069\ (4.86/\sqrt{24})$$

$$= -7.97 \pm 2.069\ (.99)$$

$$= -7.97 - 2.05,\ -7.97 + 2.05$$

$$= -10.02,\ -5.92$$

Confidence Level on the mean difference; 95% CI = (-10.02, -5.92)

That means on an average, the participants' reaction time for Congruent test will be less by -10.02 to -5.92 as compared to the reaction time for the Incongruent test.

Now let us calculate the effect size measure: Cohen's d

Cohens' d = M_D/S_D

$$= -7.97/4.86$$

$$= -1.64$$

$$r^2 = \frac{t^2}{t^2+df}$$

$$= \frac{(-8.05)^2}{(-8.05)+23}$$

$$= \frac{64.08}{87.08}$$

$$= .736 \text{ (or) } 73.6\%$$

That means 73.6% of the variability in reaction time is due to the different type test (i.e.) Congruent test.

Question 6: Extending the investigation

Answer:

Meaningfulness of results:

i) What was measured? The reaction time for two types of tests in printing the names of the colors and recognizing them. [7a]

This has a practical importance as these types of tests can be practically applied to test the effectiveness of imparting new tests to students.

ii) Why report Effect Size?

The effect size is the main finding of a quantitative study. While a *P* value can inform the reader whether an effect exists, the *P* value will not reveal the size of the effect. In reporting and interpreting studies, both the substantive significance (effect size) and statistical significance (*P* value) are essential results to be reported.

iii) Why reporting P value is not enough?

Statistical significance is the probability that the observed difference between two groups is due to chance. If the *P* value is larger than the alpha level chosen (eg, .05), any observed difference is assumed to be explained by sampling variability. With a sufficiently large sample, a statistical test will almost always demonstrate a significant difference, unless there is no effect whatsoever, that is, when the effect size is exactly zero; yet very small differences, even if significant, are often meaningless. Thus, reporting only the significant *P* value for an analysis is not adequate for readers to fully understand the results.

Other Investigation:

iv) An investigation of how Stroop interference is affected by the use of color names in a learned language in comparison to a native language [8]

Experiment 1 was conducted in English and a similar experiment was conducted for the reaction time in a second that the participants had acquired outside the sensitive or critical period for language acquisition (I.e. a conscious effort had been made to learn the language). Although it appeared that there could be some Stroop interference on participants' performance in experiment 2, it was not as significant as in experiment 1. It could have been the case that for some participants, the presence of the second language became a bigger distraction than the contradiction between the name and ink color.

References:

For the preparation of this project

[1] https://en.wikipedia.org/wiki/Stroop_effect

[2] https://en.wikipedia.org/wiki/Null_hypothesis

[3] https://www.youtube.com/watch?v=SRGaW3maK38

[3a] https://www.youtube.com/watch?v=eU0Ls3HXGSM

[4] Refer to the Excel Sheet stroopdata.xlsx (that forms part of the

[5] https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg

[6] http://www.graphpad.com/quickcalcs/

[7] https://www.youtube.com/watch?v=_PqnDYMO3lw (Drawing Normal Distribution Curve)

[7a] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3444174/

[8] http://www.stephenpersefoundationlearning.com/wp-content/uploads/2014/06/Stroop-Research-Project1.pdf

[9] http://stats.stackexchange.com/questions/22387/why-it-is-often-assumed-gaussian-distribution