Non-parametric Inference

This analysis aims to compute descriptive statistics and to perform statistical tests (both Parametric and Non-Parametric) on a dataset based on a psychological phenomenon: The Stroop Effect.

This report is structured into four sections: (1) *Introduction*, where the investigation problem is set; (2) *Methodology*, where the analysis is briefed; (3) *Results*, results of various statistical tests applied are reported; and (4) *Conclusion*

Introduction :

Stroop Effect :

The Stroop effect **[1]** is a phenomenon that occurs when you must say the color of a word but not the name of the word. For example, blue might be printed in red and you must say the color rather than the word. It is a demonstration of interference in the reaction time of a task.

These situations are defined as ***congruent***, when the name of color is equivalent to the name, or

incongruent, when the name denotes another color, as shown in the following example sets.

Congruent set of words : **Green Red Blue Orange**
Incongruent set of words : **Green Red Blue Orange**

Dataset :

The dataset used in this analysis is available at **[2]** and refers to the time of reading the ink color either for a list of congruent words and for a list of incongruent words. Each line of the dataset corresponds to a different person, in a total of 24 people.

Independent variable : condition of the set of words, i.e. whether the words are congruent or not.
Dependent variable : varies accordingly to each sample. In this problem, it refers to the time to read and say out loud the color in which a set of words is printed.

Hypothesis to be explored :

Does the time for reading incongruent words differ from that reading congruent ones?

Methodology:

1. Test of Randomness - Runs test using Median
2. Exploratory Data Analysis & GoF Tests - Empirical Distribution Function, QQ Plots, Lilliefors Test of Normality
3. Testing for Association between Congruent and Incongruent times
4. Hypothesis Testing using

a) Non Parametric Tests - Sign Test, Wilcoxon Sign Rank Test
b) Parametric Tests - T Test

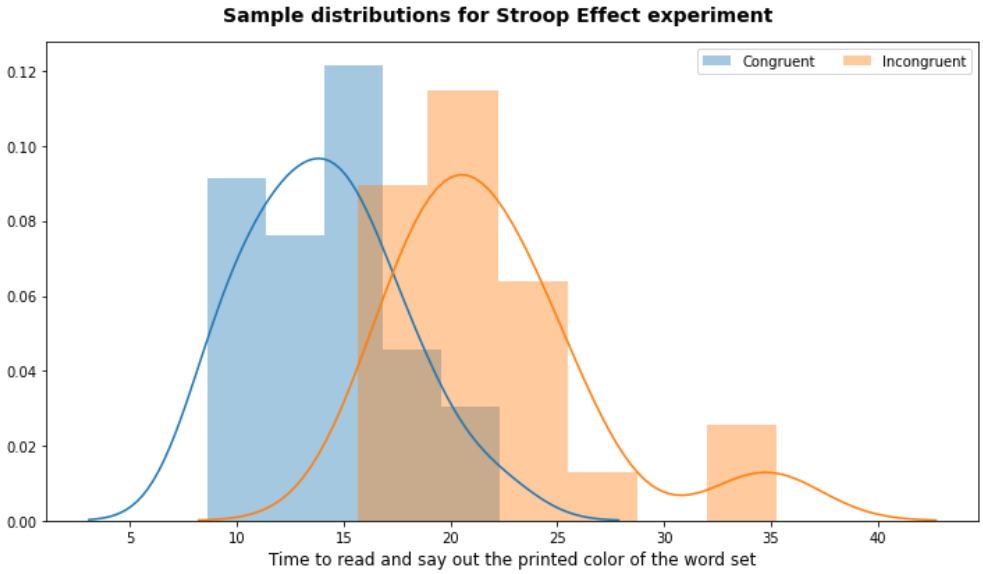
Results:

1. Randomness:

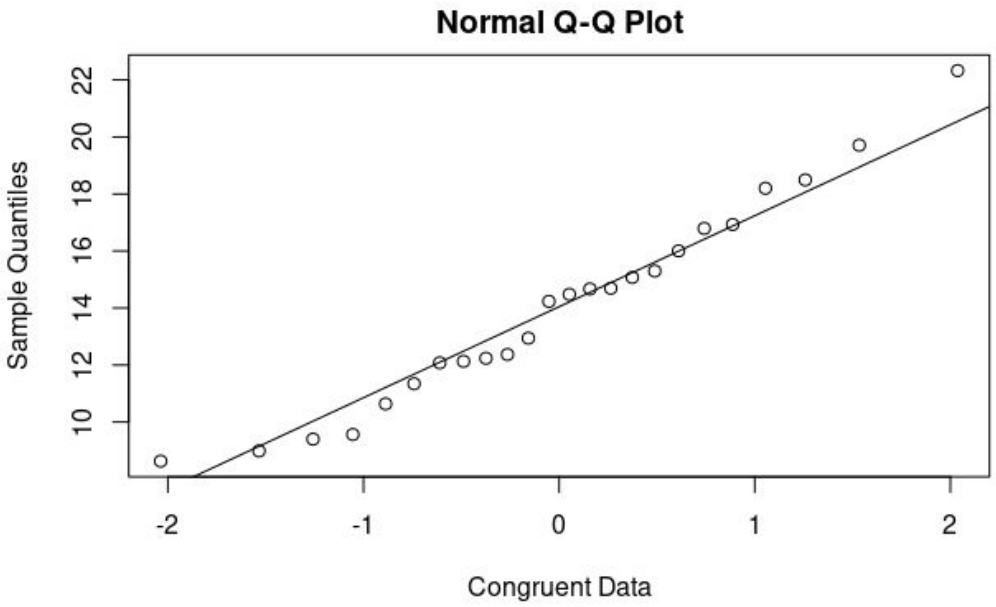
Using the Runs Test ,at a significance level of 0.05, we fail to reject the hypothesis that the samples are randomly generated.

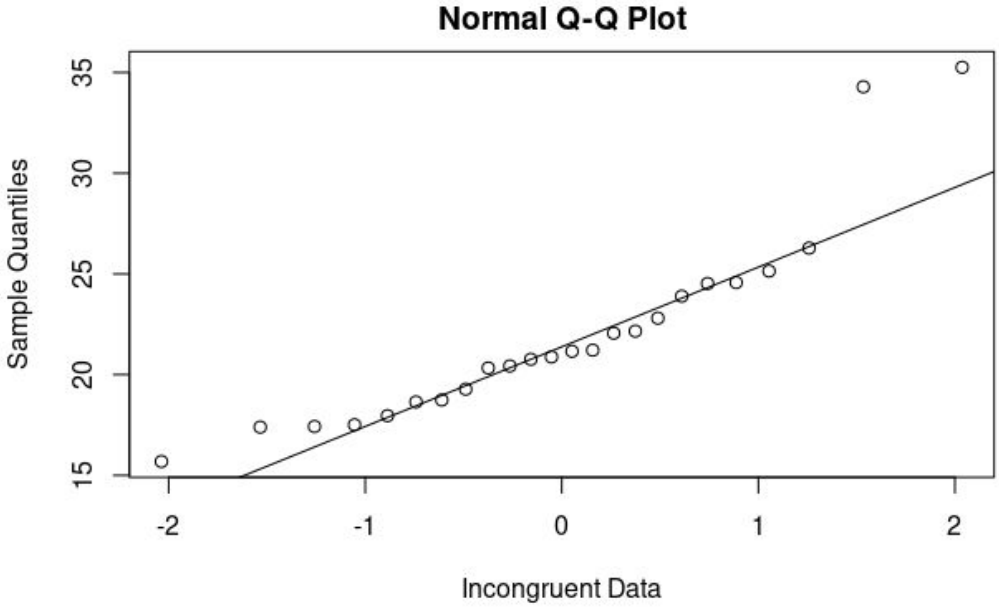
2. Data Analysis and GoF:

Congruent	Incongruent
Min. : 8.63	Min. :15.69
1st Qu.:11.90	1st Qu.:18.72
Median :14.36	Median :21.02
Mean :14.05	Mean :22.02
3rd Qu.:16.20	3rd Qu.:24.05
Max. :22.33	Max. :35.26



*generated from python script





Sample median and sample mean are approximately equal indicating that the data can follow Normal Distribution which is also indicated by QQ Plots.

On running Lilliefors Test of Normality, we fail to reject the null hypothesis that the data follows Normal distribution.

3. Measures of Association:

Using Kendall Tau and Spearman statistics,we have statistical evidence that both the r.v.s are not independent , but indeed associate positively.*(values are reported in R notebook shared)*

4. Hypothesis Testing

We are dealing with a dependent sample - the subject who reads the congruent set of words is the same that reads the incongruent one. (This is also evident from the tests performed above) and we are interested in testing for an average difference among them. Thus we consider paired tests rather than 2 sample tests.

- **Null hypothesis:** the population difference is centered in zero, i.e., the median time/mean time is the same for reading either congruent and incongruent set of words. This definition can be mathematically defined as:
  $H_o : M_D = 0$ (or)
  $H_0 : \mu_d = 0$

- **Alternative hypothesis:** the population difference is nonzero.
  Clearly, time taken to read incongruent words is more than the time taken for congruent words. This is also evident if we observe the previous few measurements we notice the time is shorter for reading the congruent set of words, thereby we can hypothesize that reading congruent words is faster than reading the incongruent ones. Mathematically it is the same as obtaining a negative difference. Thus our alternative hypothesis can be narrowed to the following condition:
  $H_1 : M_D < 0$ (or)
  $H_1 : \mu_d < 0$

a) Non Parametric Tests

$H_o : M_D = 0$
$H_1 : M_D < 0$

Both Sign Test and Wilcoxon Signed Rank Test produce the same p-value resulting in rejection of Null hypothesis at a significance level of 0.05.
P value : 5.96e-08

b) Parametric Tests

$H_0 : \mu_d = 0$
$H_1 : \mu_d < 0$

Left tailed Paired T-Test is used since there is an evidence that the sample follows Normal distribution.

At a significance level of 0.05, a p value of 2.052e-08 is obtained, thus rejecting the null hypothesis.

Hence, by using both parametric and non parametric tests we conclude that there is significant evidence that the reaction times differ.

Observations and Conclusions:

- Both Sign test and Wilcoxon Sign Rank Test produce same results i.e same p value because the sample obtained after differencing has the same sign i.e all values are < 0
- Although the exact sizes of Sign Test and T Test differ, sign test which has less size has more power. Thus, the Sign test is more powerful than the T Test in this case.

  **values are reported in R notebook shared*
- In this analysis we presented the Stroop Effect and applied various tests to collect statistical evidence to support this perceptive phenomenon that **reaction times do differ.** In doing so, we have gone through the steps necessary to test a hypothesis such as testing the randomness and performing exploratory data analysis.

References :

1. **https://www.verywellmind.com/what-is-the-stroop-effect-2795832#:~:text=The%20Stroop%20effect%20is%20a,color%20rather%20than%20the%20word.**
2. *Stroop Effect Dataset*. Available at: **https://www.google.com/url?q=https://drive.google.com/file/d/0B9Yf01UalbUgQXpYb2NhZ29yX1U/view?usp%3Dsharing&sa=D&ust=1525216530059000.**