P1: Test a Perceptual Phenomenon

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.

Analysis:

1. Here the same participants test to conditions; the independent variable is **words condition** and the dependent variable is **the time it takes to name the ink colours**.

2. This is a **within subject design** where participants test two conditions. Therefore, we have paired data; time it takes to name ink colours for congruent words condition and incongruent words condition.
The null hypothesis could be: time taken to name ink colours does is not affected by the words condition (congruent or incongruent)
The alternative hypothesis could be: time taken to name ink colours is affected by words condition.
Let, μ_c be the **population mean** for the time taken to name ink colours for Congruent words condition
   μ_I be the **population mean** for time taken to name ink colours for Incongruent words condition

We can now write our hypothesis as:
 H_O: $\mu_c = \mu_I$
 H_A: $\mu_c \neq \mu_I$
Where H_O is the null hypothesis and H_A is the alternative hypothesis.
We assume that the differences are **Normally distributed.**
 Since, the sample size N=24 is less than 30 and population standard deviation for this task is unknown- a **t-test** is more appropriate. *Also, we wish to determine whether the mean of differences between two paired samples differs from 0 and the effect of words condition on the average time taken to read ink names is unknown, in other words, we do not know if the incongruent words conditions reduces or increases the time taken by participants*. Hence, **a non-directional two tailed paired t-test** should be used.
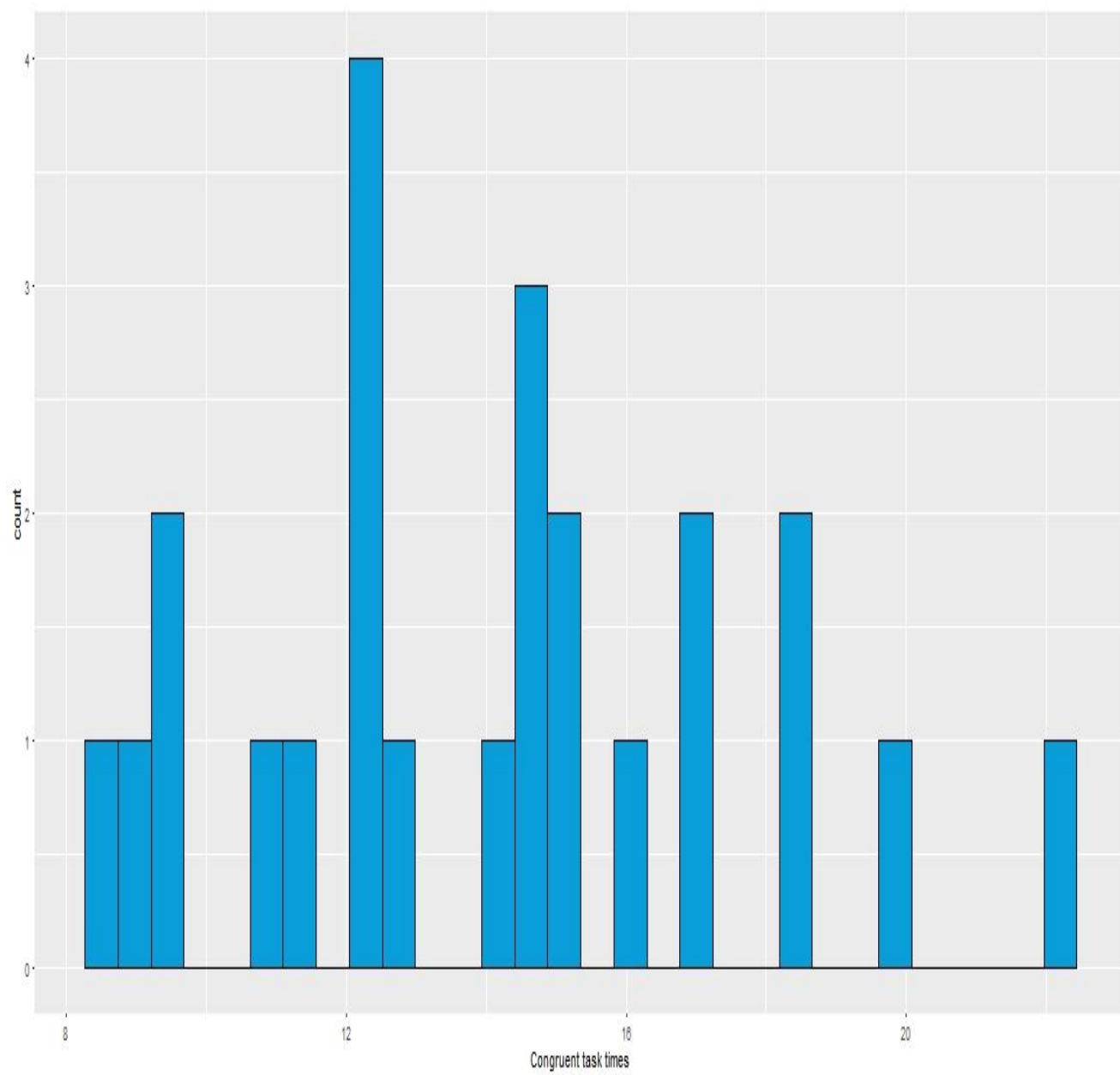
3-4. Descriptive statistics

The following statistics and plots are generated using R and Rstudio:
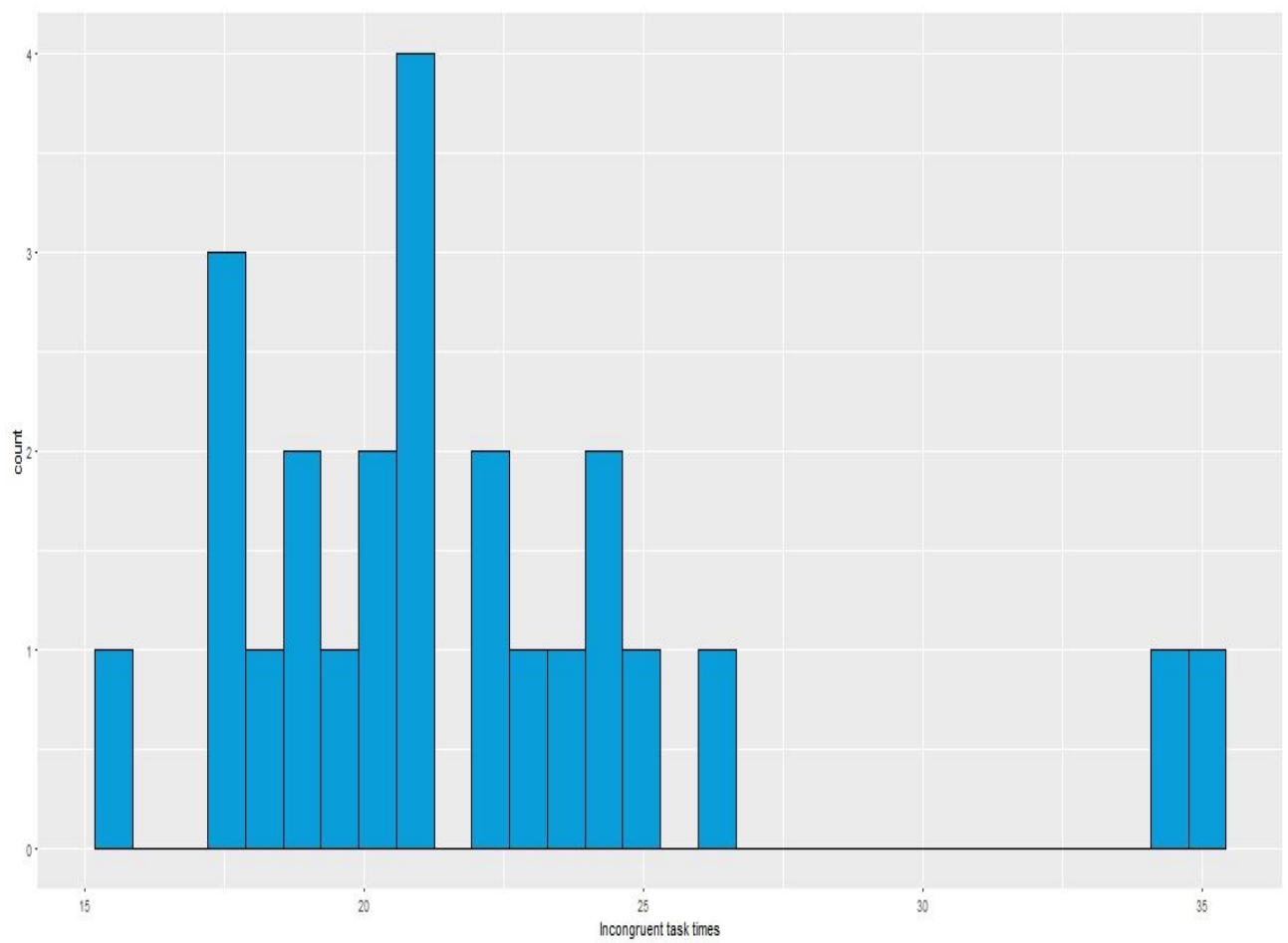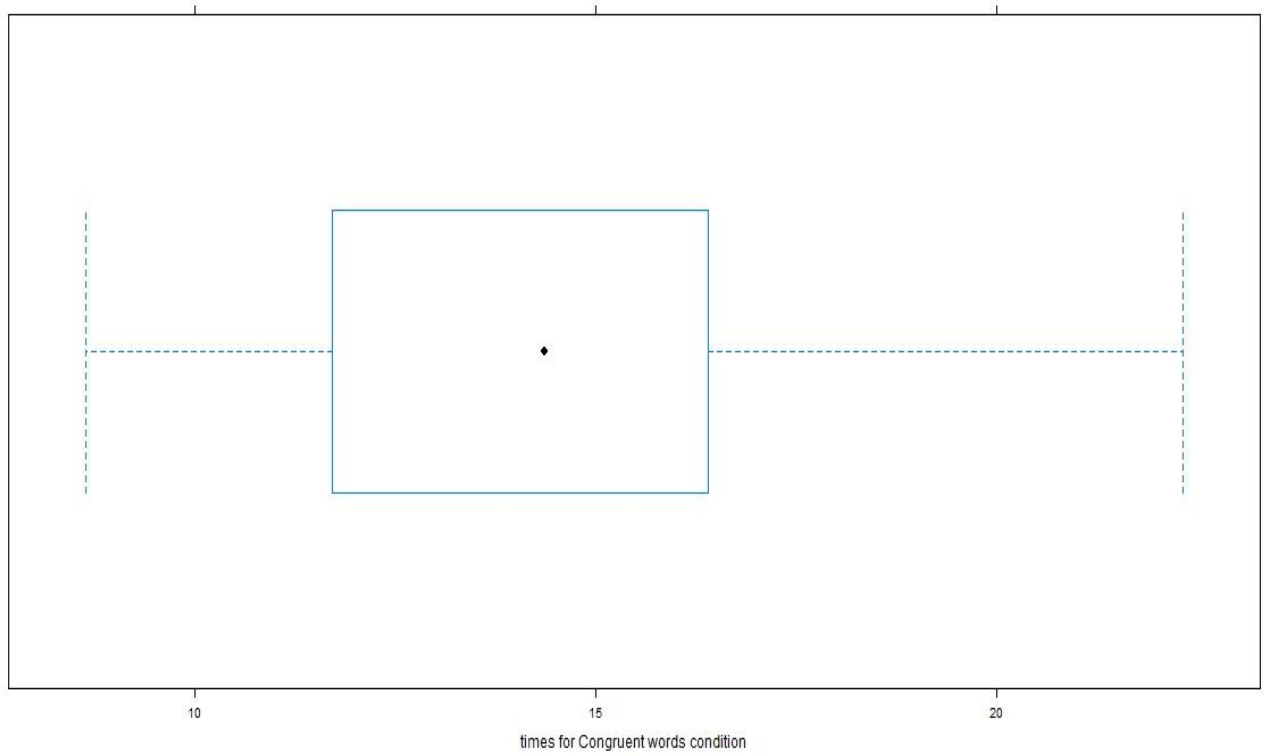
```
Congruent        Incongruent
 Min.   : 8.63    Min.   :15.69
 1st Qu.:11.90    1st Qu.:18.72
 Median: 14.36    Median :21.02
 Mean   :14.05    Mean   :22.02
 3rd Qu.:16.20    3rd Qu.:24.05
 Max.   :22.33    Max.   :35.26
```
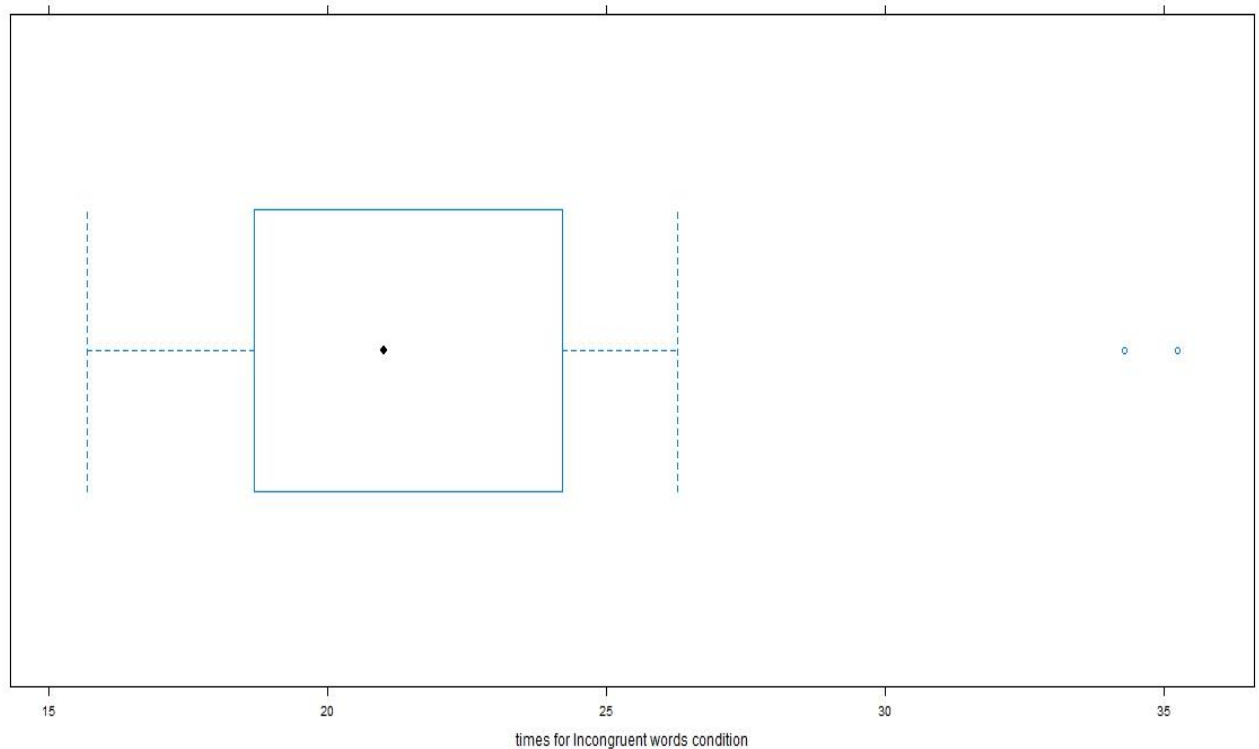
Therefore, the values for mean and standard deviation for the two tasks come out to be:

$\mu_c = 14.05113$ $\mu_I = 22.01592$

$\sigma_c = 3.5593$ $\sigma_I = 4.79706$

times for Congruent words condition

Incongruent task times

times for Incongruent words condition

From the Histograms above, we can see the distribution for Congruent words condition is **Unimodal** whereas for Incongruent words condition it is **Bimodal**. Also, the boxplot shows outliers in case of Incongruent words condition.

Interestingly, the minimum for Incongruent words condition is close to the median for Congruent words condition. Therefore, we can get an idea about the increased average time in case of Incongruent words condition.

5. Now,

μ_c = 14.05113 μ_I= 22.01592   N=24

σ_c= 3.5593 σ_I= 4.79706

Let α=0.05

The degrees of freedom df= 23

Point estimate M_D= μ_I- μ_c= 7.965

Standard deviation of the differences, S_D=4.865

Standard error of the mean difference= $S_D/\sqrt{N}$= 0.993

t- critical values for α=0.05 is $\pm$2.069

t-statistic = M_D/ SEM= 8.0211

Since, our t-statistic falls in the critical region we reject the Null hypothesis.

The 95% Confidence interval for the mean difference will be:

(M_D-(t-critical*SEM), M_D+(t-critical*SEM)) (Assuming t-critical=2.069)

Therefore,

CI: (5.910, 10.02)

$r^2 = t^2/(t^2+df)$

 =0.7367

So, there 73.67% variation in the means for the two conditions.

Using Graphpad, the p-value for two tailed t test is less than 0.0001. Therefore the difference in means is extremely significant. The words condition had a causal effect on the performance of the participants. There is strong evidence that participants took less time in case of congruent words condition compared to incongruent words condition.

This was also expected from the initial plots.