

# Stroop Effect Analysis

*Maitrayee Ranade Bordikar*

*February 28, 2017*

## Overview:

Stroop effect is a psychological phenomenon which demonstrates how interference affects reaction time to a task [1]. It is named after J. Ridley Stroop, who discovered this phenomenon [2]. In Stroop test, colors are written in words with different font colors. When the word and color of the font match each other it's a congruent word and when it does not match it's an incongruent word. Test subjects are asked to name the color of the font and instead of reading the word. The response time to name the color of the font is recorded and grouped by congruent and incongruent words. A statistical analysis of one such recorded data set is presented in this project.

## Independent and dependent variable:

The independent variable in the experiment is if the color of the font matches the word itself or not. Dependent Variable is the response time to name the color of the font.

## Hypothesis:

After establishing the independent and dependent variables it looks that the response time depends on the color of the font and the word. Hence a possible hypothesis could be that the reaction time increases when the color of the font does not match the word i.e. response time for incongruent words is more than the response time for congruent words. Therefore the null hypothesis is that the reaction time is same for both congruent and incongruent words.

Hence,  $H_0 : \mu_C = \mu_I$  at  $\alpha = 0.05$  and  $H_a : \mu_C \neq \mu_I$  at  $\alpha = 0.05$

where,  $H_0$  is the null hypothesis,  $H_a$  is the alternate hypothesis,  $\mu_C$  is the population mean of congruent word's response time,  $\mu_I$  is the population mean on incongruent word's response time and  $\alpha$  is the type 1 error for 95% confidence interval.

## Central tendency and Variance:

The code below calculates the sample mean, median and variance of congruent word's response time and incongruent word's response time.

```
library(ggplot2) ##importing library needed to make the plots
library(reshape2) ##importing library to reshape the data frame

df = read.csv ("stroopdata.csv") ## loading data from file

## combining data into single column for box plot
dfCombined <- melt(df, measure.vars = c("Congruent", "Incongruent"))

meanCong <- mean(df$Congruent) ## calculating congruent mean
meanIncong <- mean(df$Incongruent) ## calculating incongruent mean
varCong <- var(df$Congruent) ## calculating congruent variance
```

```

varIncong <- var(df$Incongruent) ## calculating incongruent variance
medCong <- median(df$Congruent) ## calculating congruent median
medIncong <- median(df$Incongruent) ## calculating incongruent median

## writing the sample mean, median and variance in a tabular format

centralTendancy <- matrix(c(meanCong,medCong, varCong,
                             meanIncong, medIncong, varIncong),ncol=3,byrow=TRUE)
colnames(centralTendancy) <- c("Mean","Median","Variance")
rownames(centralTendancy) <- c("Congruent Words","Incongruent Words")
centralTendancy <- as.table(centralTendancy)
centralTendancy

```

```

##              Mean   Median Variance
## Congruent Words 14.05113 14.35650 12.66903
## Incongruent Words 22.01592 21.01750 23.01176

```

The table shows sample mean, median and variance of congruent word's response time and incongruent word's response time. Mean, median and variance of congruent words response times is 14.05s, 14.36s, and 12.66s respectively. While mean, median and variance of incongruent words response times is 22.06s, 21.017s, and 23.01s respectively. The median and mean of response times of incongruent words differ by a second because of couple of outlier points above 34s. While mean and median of response times of congruent words is approximately same.

## Distribution of data:

The distribution of response times of both the set of words and box plot is plotted using the code below. ggplot package of R is used to make the figures.

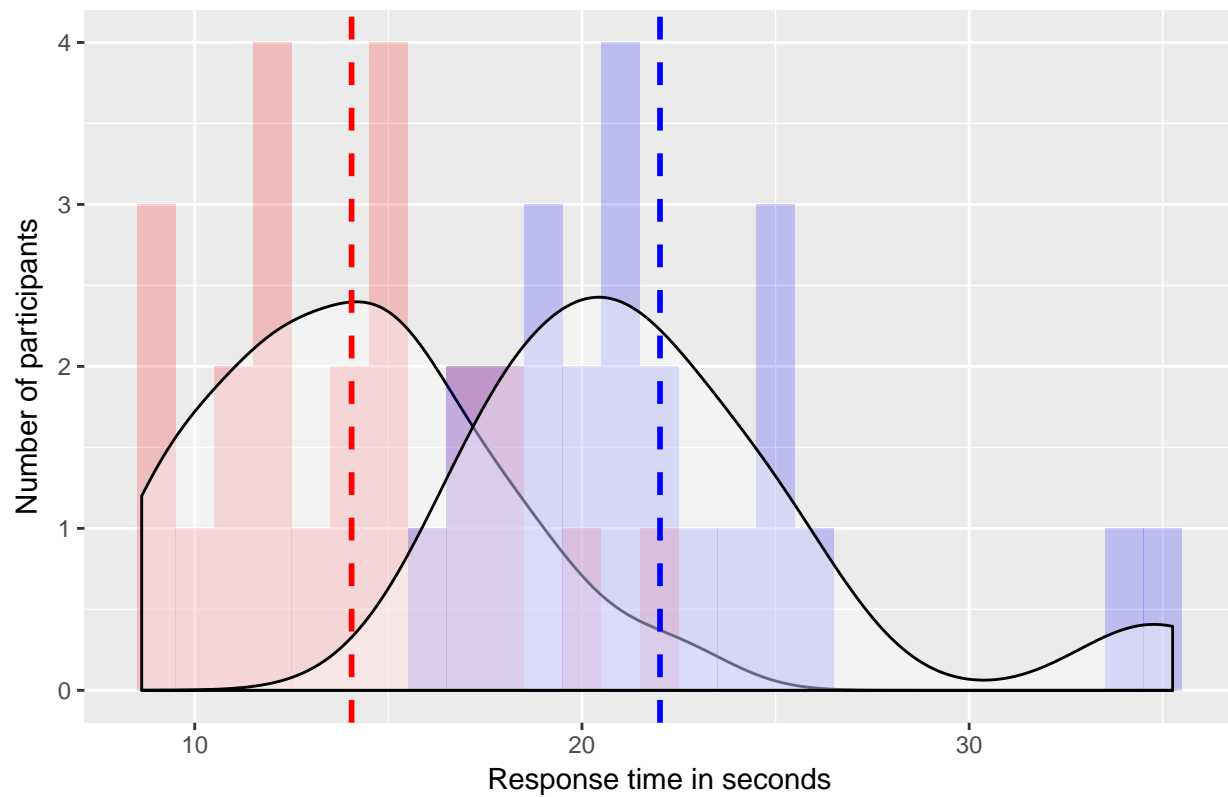
```

## plotting histogram response time by count and overlaying density distribution
ggplot(data=df ) +   geom_histogram (aes(x = df$Congruent,y = ..count..),binwidth=1,
                                alpha=0.2, fill="red", position="identity") +
  geom_density(aes(x = df$Congruent,y = ..count..), alpha=.4, fill="white") +

  geom_histogram (aes(x = df$Incongruent,y = ..count..),binwidth=1,
                                alpha=0.2, fill="blue", position="identity") +
  geom_density(aes(x = df$Incongruent,y = ..count..), alpha=.4, fill="white") +
  theme(legend.position="bottom")+
  geom_vline(aes(xintercept=mean(df$Congruent)),
            color="red", linetype="dashed", size=1)+ ## congruent mean
  geom_vline(aes(xintercept=mean(df$Incongruent)),
            color="blue", linetype="dashed", size=1)+ ## incongruent word mean
  xlab("Response time in seconds") + ylab("Number of participants") +
  ggtitle("Fig. 1: Response times distribution")+
  theme(plot.title = element_text(hjust = 0.5))

```

Fig. 1: Response times distribution



```
## box plot grouped by words
ggplot(dfCombined) +
  geom_boxplot(aes(x=variable, y=value, fill=variable),
               alpha=.5, width=1, position = position_dodge(width = 1),
               outlier.colour = "black", outlier.size = 1)+
  xlab("Words") + ylab("Response") +
  ggtitle("Fig. 2: Response times grouped by words") +
  theme(plot.title = element_text(hjust = 0.5))
```

Fig. 2: Response times grouped by words

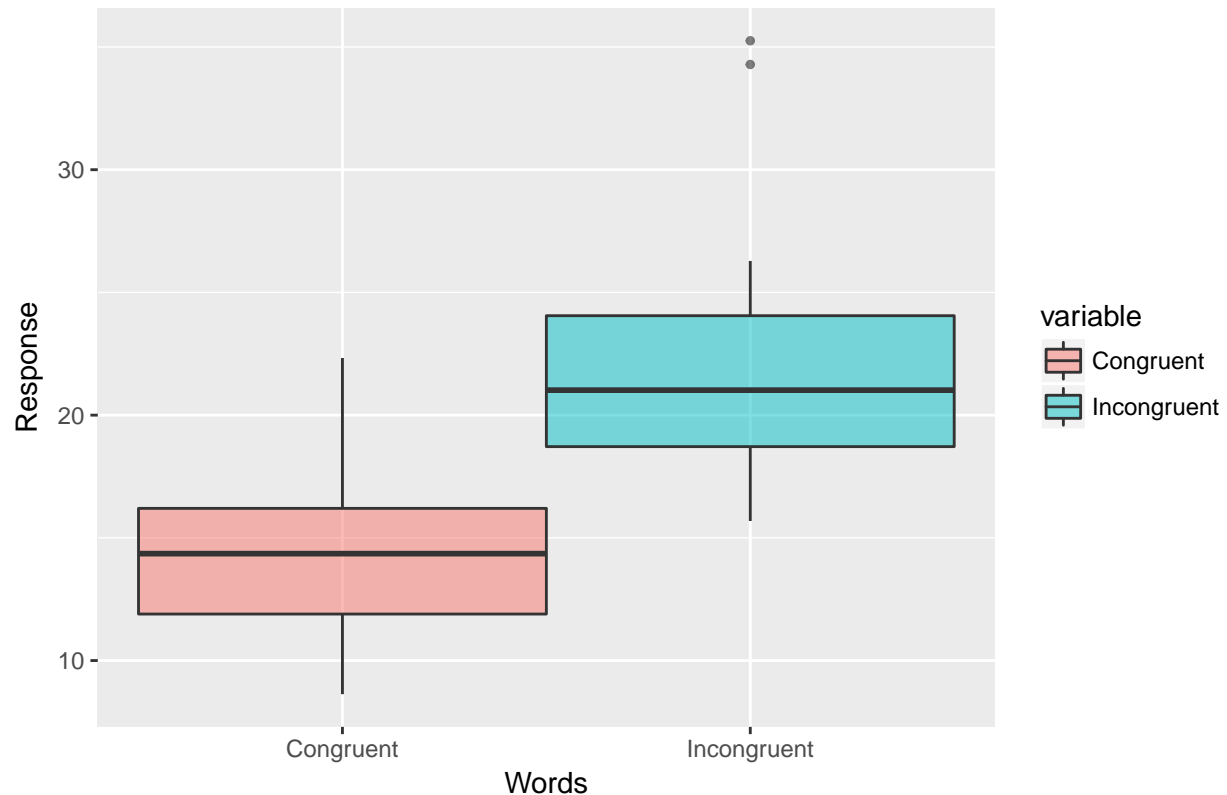


Fig. 1 shows distribution of response time for congruent words and incongruent words overlayed on the same time scale. The red histogram with the overlaying density distribution in white, shows the response time for congruent words while the blue histogram with overlaying density distribution in white again, shows the response time for incongruent words. The red and blue dotted line show the mean of each distribution.

Both set of response time distrubution is approximately normal but the density distribution of incongruent word's response time is shifted to the right on the time scale with its mean larger than the mean of congruent word's response time. The mean of congruent words response time is 14.s and that of incongruent word's response time is 22.s.

Fig. 2 shows the box plot grouped by congruent and incongruent words. The outlier points in the incongruent words response time is shown in black dots.

## Confidence interval and hypothesis testing:

The sample data size is less than 30, and we do not have the knowledge of population standard deviation,  $\sigma$ . Also, sample distribution is normal. Since the alternate hypothesis states that the congruent and incongruent mean is not equal, it is best to use the two tailed t test for the statistical test. The same sample of participants are tested twice and for two different kinds of words, therefore the observations are considered paired. `t.test` function in R is used to calculate the 95% confidece interval and critical value below.

```
##t_words contains results of t.test
tWords <- t.test(df$Congruent, df$Incongruent, paired = TRUE)
tWords_results <- data.frame("Critical_Value" = tWords$p.val,
                             "Conf_Low" = tWords$conf[1], "Conf_High" = tWords$conf[2])
tWords_results ## results put in tabular format
```

```
## Critical_Value Conf_Low Conf_High
## 1 4.103001e-08 -10.01903 -5.910555
```

The 95% confidence interval limits are [-10.02, -5.91]. It does not contain 0. Also p-value or critical value of  $4.1e-08$  is much less than 0.05. The true difference in mean is not equal to zero. Thus we can reject the null hypothesis and conclude that response time increases for incongruent words when the color of the font does not match the word. Thus the alternative hypothesis is true as expected

## Observations:

I took the Stroop test and my response times for congruent and incongruent words were 18.11s and 23.32s respectively. The incongruent words response time is in fact more than the congruent words response time as expected. This phenomenon occurs because of interference with the reaction time. It is much easier to read the words than name the color. Words have a stronger influence than the color itself [2]. Hence when the word differs from the color it takes longer for brain to not read the word and name the color. Perhaps a person who cannot read would have comparable response times of congruent and incongruent words since reading task won't interfere with the color naming task. A color-object interference test similar to Stroop test could be used on subjects that can't read, such as children [3]. In the color-object test line drawings of familiar objects in a color that is congruent (e.g., an orange carrot), incongruent (e.g., a green carrot), or neutral (for objects having no canonical color [e.g., a red book]), and abstract shapes are each drawn in one of six colors and shown to subjects and their response times are recorded. This test also shows results similar to Stroop test.

## References:

1. [https://en.wikipedia.org/wiki/Stroop\\_effect](https://en.wikipedia.org/wiki/Stroop_effect)
2. <https://faculty.washington.edu/chudler/words.html>
3. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2134842/>