STROOP'S EFFECT

Sai Saketh Boyanapalli

October 16, 2017

Stroop's effect:

Famous experiment performed by J. Ridley Stroop who discovered that when a person is said to say the color of ink on a word. It is found that time taken to say color of Congruent words i.e., the words that have same color printed as the word. E.g. Red, Green, Blue. Is less than that of time taken to say the color of Incongruent word i.e., words that have different ink and word. E.g. Red, Green, Blue.

Here I investigate this effect by taking a sample data of participants who have performed this experiment.

```
library(ggplot2)

stroopdata <- read.csv("stroopdata.csv") # read data into data frame
```

Test: We have sample data of 24 participants who have performed stroop's task.

This test uses a within subject design i.e., same user performs both task

Independent variables:

Type of words displayed on screen: Congruent words & Incongruent words.

Dependent variable:

Time, it takes to complete the task.

Here the population parameters are not available.

we can perform a dependent t - test for paired samples.

Research question:

Does it take more time to complete the task with Incongruent words than it takes to complete task with Congruent words?

Hypothesis

$\boldsymbol{H_0}$: The Population mean of time taken to complete the task using Congruent words is less than or equal to that of using Incongruent words, in other words the population mean difference of completing task is less than or equal to zero.

$\boldsymbol{H_0}: \mu_d \leq \mathbf{0}$

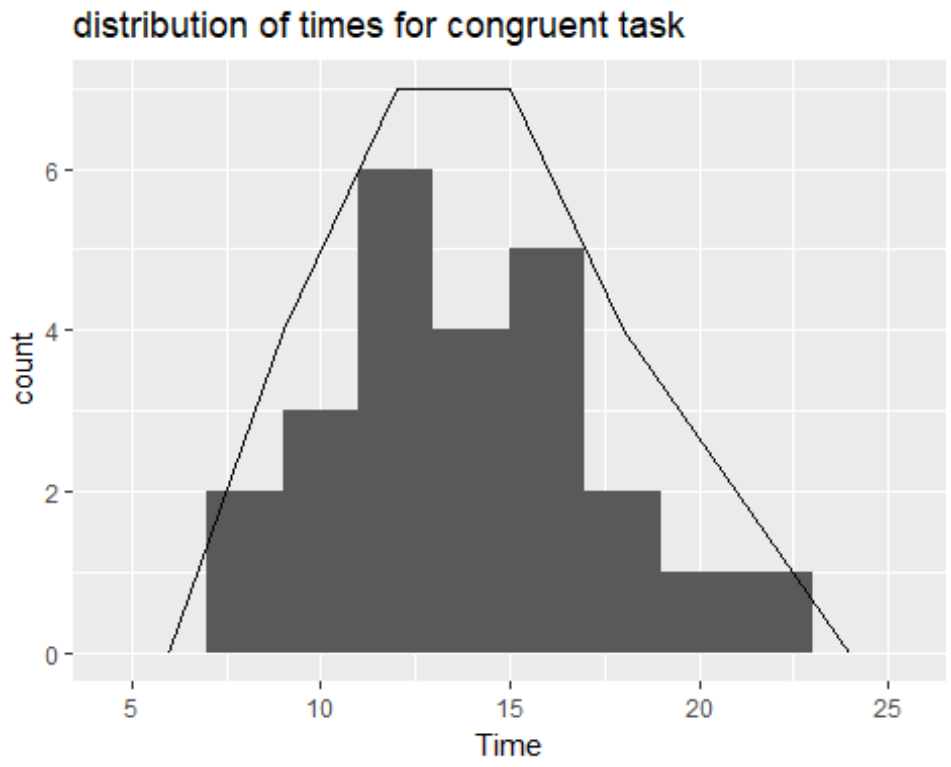
$\boldsymbol{H_1}$: The population mean of time taken to complete the task using Incongruent words is greater than that of using Congruent words, the population mean difference in time taken to complete the tasks is greater than zero.

$\boldsymbol{H_1}: \mu_d > \mathbf{0}$

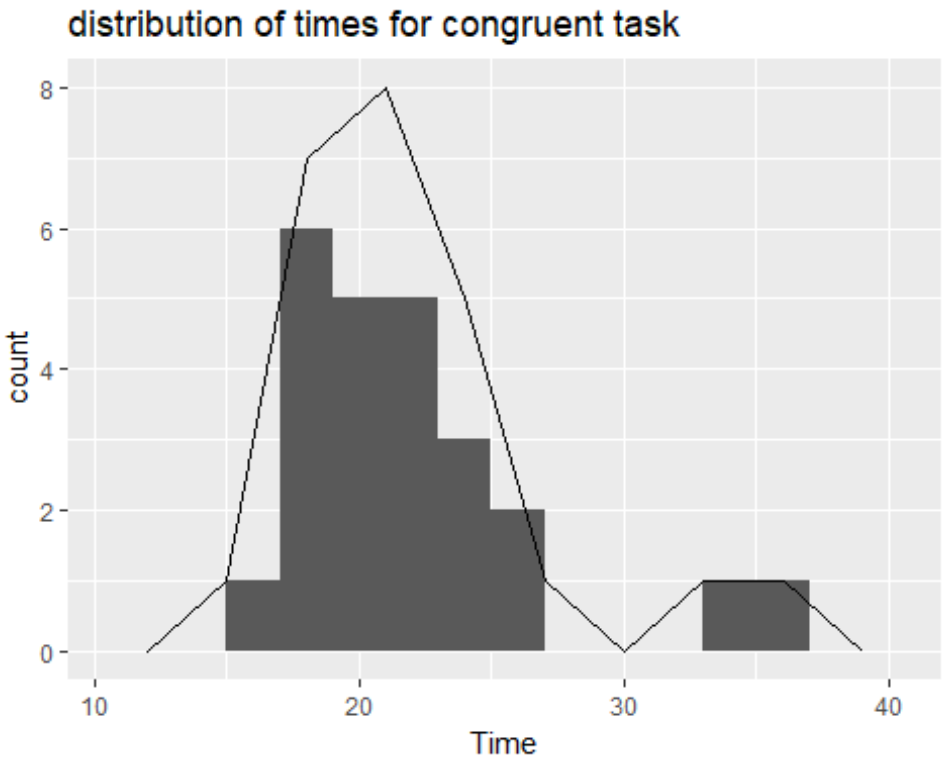
Assumptions

1. dependent variable is continuous. (our dependent variable is time which is consistent)
2. t - test assumes that the data is normally distributed.
3. There are no outlier's in the data.
4. observations are independent of one another.

Let's look at our sample data distributions

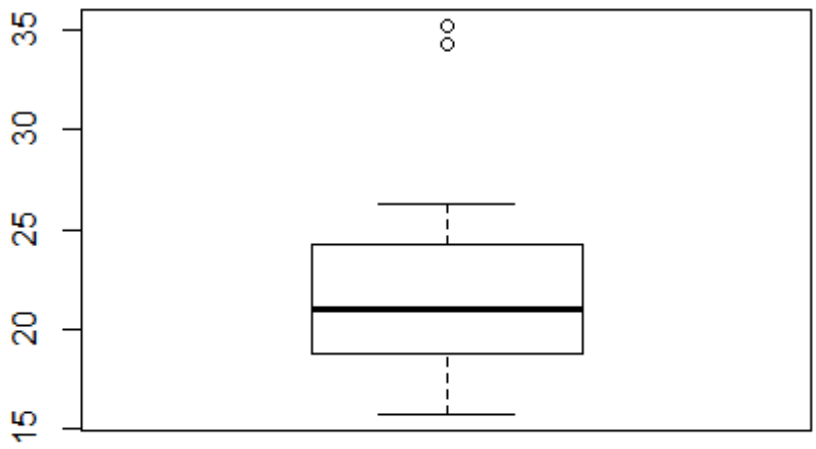


We can see that the data is approximately normally distributed.



Here we see that the data is not normally distributed.

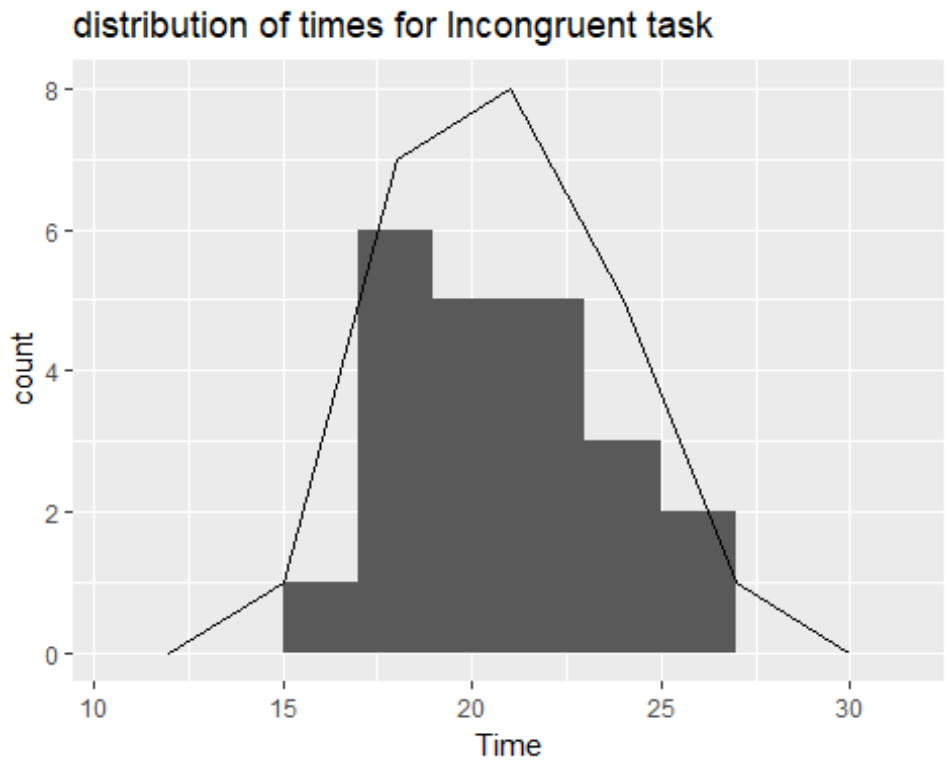
Let's take a look at boxplot of Incongruent task to find outliers in the data.



Box plot shows that they are 2 outliers in the data. Creating a new data frame by removing those outliers in the data.

```
newstroopsData = subset(stroopdata, Incongruent < 30) # subsetting data with incongruent time observations less than 30.
```

let's again look at histogram.



This time it approximates to normal distribution.

Dependent paired t - test.

Sample parameters

```
nrow(newstroopsData) # No of observations in our sample data.

## [1] 22

summary(newstroopsData) # calculates summary statistics

##    Congruent      Incongruent
##  Min.   : 8.63   Min.   :15.69
##  1st Qu.:11.53   1st Qu.:18.67
##  Median :14.36   Median :20.82
##  Mean   :13.94   Mean   :20.86
```

```
##  3rd Qu.:15.83   3rd Qu.:22.64
##  Max.   :22.33   Max.   :26.28
```

Mean, Median of Congruent task are 13.96, 14.36.

similarly Mean, Median of Incongruent task are 20.86, 20.82.

Degree of freedom = n - 1

```
dof <- nrow(newstroopsData) - 1 #degrees of freedom
dof
```

```
## [1] 21
```

```
# sample standard deviation of congruent task
sqrt(sum((newstroopsData$Congruent - mean(newstroopsData$Congruent))^2)/(dof)
)
```

```
## [1] 3.592773
```

```
# sample standard deviation of Incongruent task
sqrt(sum((newstroopsData$Incongruent - mean(newstroopsData$Incongruent))^2)/(
dof))
```

```
## [1] 2.876923
```

```
# Difference of paired observations
newstroopsData$di <- newstroopsData$Incongruent - newstroopsData$Congruent
```

```
# Mean of the differences pair
mean(newstroopsData$di)
```

```
## [1] 6.917318
```

```
# standard deviation of differences
sd(newstroopsData$di)
```

```
## [1] 3.400562
```

```
# standard error of sampling ditribution d bar
sd(newstroopsData$di)/sqrt(nrow(newstroopsData))
```

```
## [1] 0.7250022
```

Performing an Upper tailed t- test on sample data.

Taking true differences in means of populations = 0.

at alpha = 0.05

```
t.test(newstroopsData$Incongruent, newstroopsData$Congruent, paired = T, mu =
 0, alternative = "greater", conf.level = 0.95)
```

```
##
##  Paired t-test
##
## data:  newstroopsData$Incongruent and newstroopsData$Congruent
## t = 9.5411, df = 21, p-value = 2.191e-09
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  5.669776      Inf
## sample estimates:
## mean of the differences
##                6.917318
```

Finding's

I reject the null hypothesis in favor of alternate hypothesis at alpha 0.05.

And conclude that with statistical significance the difference in means is greater than 0.

t - statistic is 9.5411, p - value = 0

95% CONFIDENCE INTERVAL 5.669 – INF

Conclusion

This indeed **strengthens** the theory of Stroops's effect that it takes more time to correctly predict the ink of a word when the ink's color and word don't match with each other.

This is because our brains process words faster than color. This indicates selective attention. That is, it is easy to ignore some features than others. This test is used to measure cognitive functioning of our brain.

Further research can be done by comparing performance of children with adults. Do we tend to do better or worse as we age?

Testing this effect on other languages which are not known to test subject will help us understand the claim that we tend more towards reading the word rather than perceiving color.

References:

http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/

http://www.udacity.com/