

P1: Test a Perceptual Phenomenon :

Vivek Pandey

1st December 2016

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.

Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

1. What is our independent variable? What is our dependent variable?

Answer:

Independent Variable : Type of word set / congruency conditions.

Dependent Variable : time takes to name the ink color of a given word sets.

2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

Answer:

One hypothesis we can use is: there is a difference between the time used to recognize colors under congruent words condition and incongruent words condition, namely, the Stroop Effect is in existence.

The hypothesis test is as follows:

H0: $\mu_{congruent} = \mu_{incongruent}$ On average the time took for population to solve both congruent task and incongruent task is the same.

HA: $\mu_{congruent} \neq \mu_{incongruent}$ On average the time took for population to solve both congruent task and incongruent task is different.

I would prefer to do two-sided paired student T-test (dependent paired T-test). This is because we don't know the mean population times for identification of the colors in incongruent and congruent conditions. Also, it is a paired t-test as the same participants are asked to identify the color of the ink in congruent and incongruent conditions (Dependent sample t test because the same participants are involved in two different conditions).

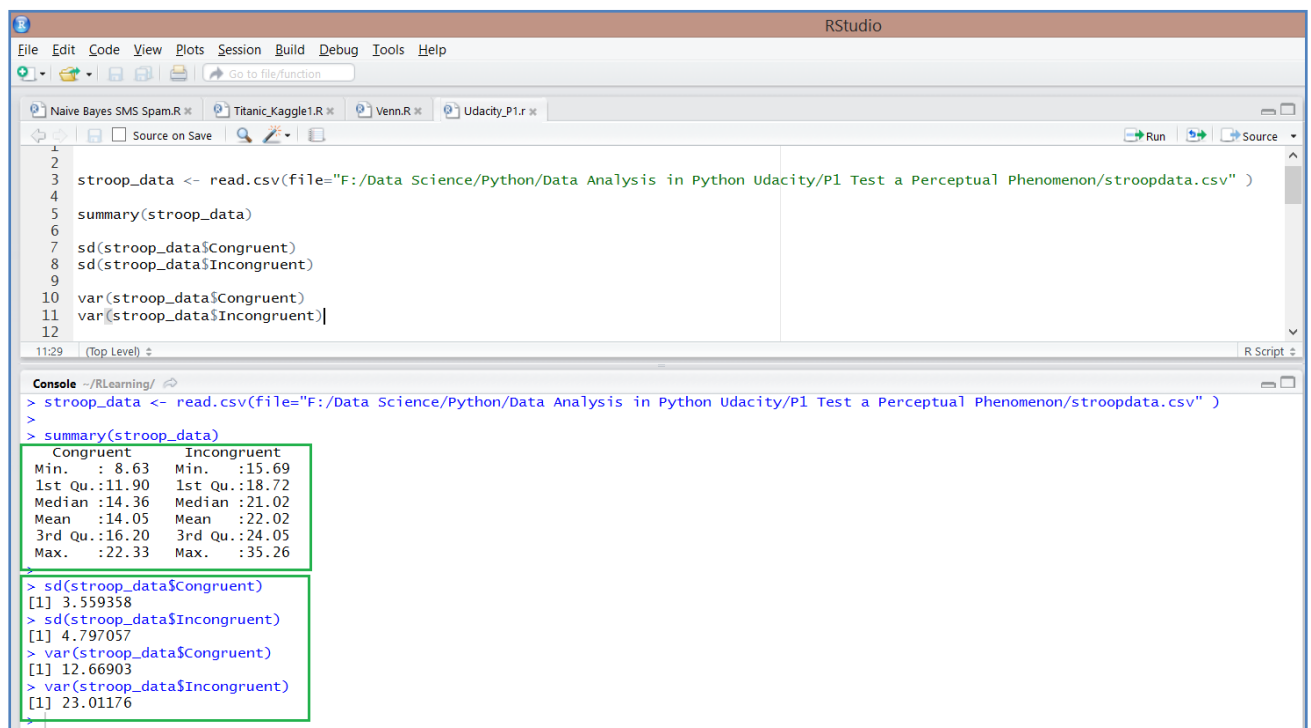
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

Answer:

The measure of tendency that will be used in this situation is mean, and measure of variability is standard deviation (SD) and variance (Var).

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
SD	: 3.56	SD	: 4.797
Var	:12.67	Var	:23.012

Note: Used R studio for these results.



The screenshot shows the RStudio interface. The script editor contains the following code:

```
1  
2  
3 stroop_data <- read.csv(file="F:/Data Science/Python/Data Analysis in Python Udacity/P1 Test a Perceptual Phenomenon/stroopdata.csv" )  
4  
5 summary(stroop_data)  
6  
7 sd(stroop_data$Congruent)  
8 sd(stroop_data$Incongruent)  
9  
10 var(stroop_data$Congruent)  
11 var(stroop_data$Incongruent)|  
12
```

The console output shows the results of these commands:

```
> stroop_data <- read.csv(file="F:/Data Science/Python/Data Analysis in Python Udacity/P1 Test a Perceptual Phenomenon/stroopdata.csv" )  
> summary(stroop_data)  
      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  
> sd(stroop_data$Congruent)  
[1] 3.559358  
> sd(stroop_data$Incongruent)  
[1] 4.797057  
> var(stroop_data$Congruent)  
[1] 12.66903  
> var(stroop_data$Incongruent)  
[1] 23.01176
```

Difference (Incongruent - Congruent)

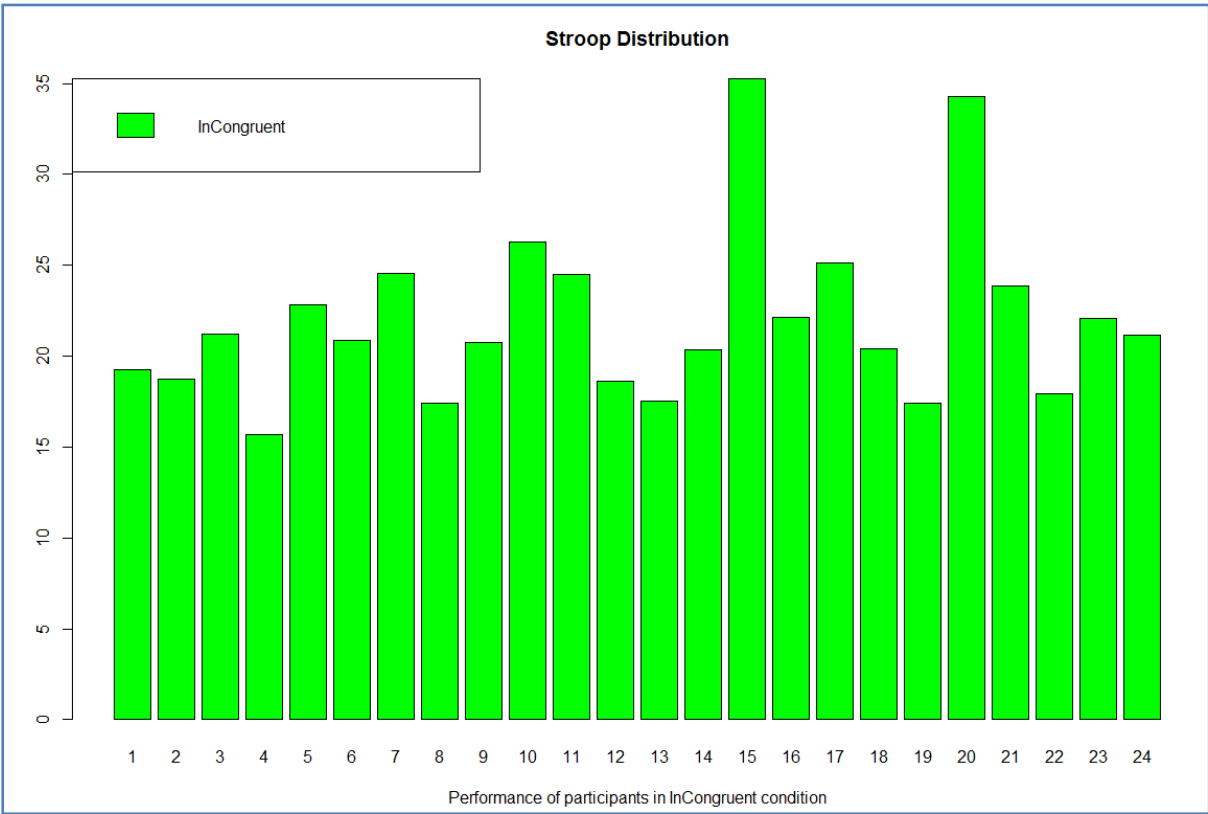
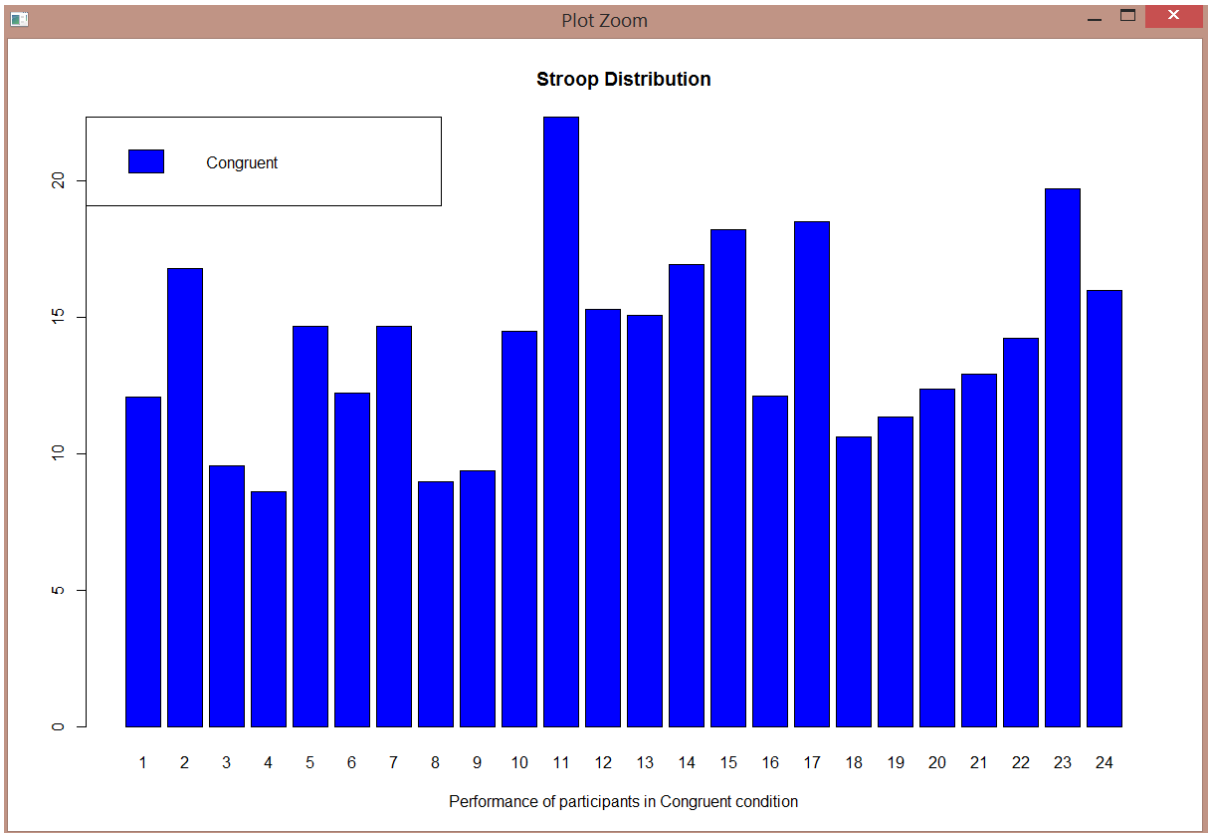
```
> summary ( stroop_data$Incongruent - stroop_data$Congruent )  
      Min. 1st Qu.  Median    Mean  3rd Qu.    Max.     
 1.950   3.646   7.666   7.965  10.260   21.920
```

```
> sd(stroop_data$Incongruent - stroop_data$Congruent )  
[1] 4.864827
```

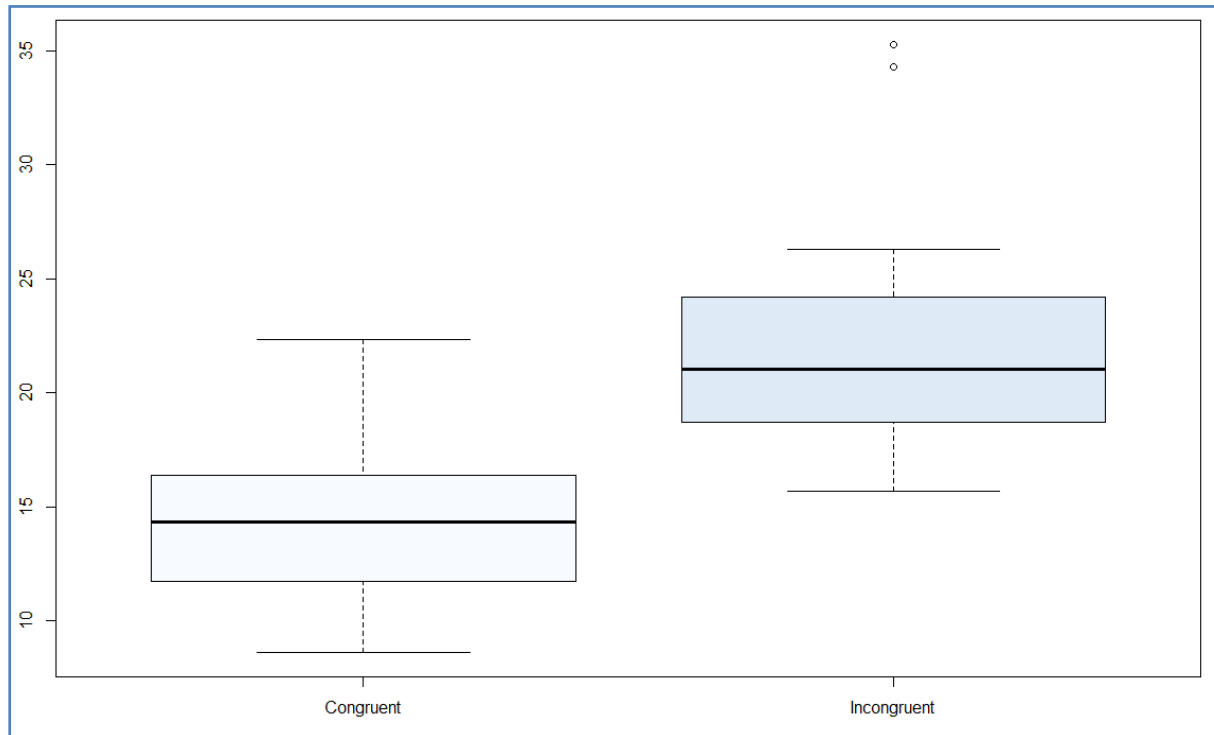
```
> var(stroop_data$Incongruent - stroop_data$Congruent )  
[1] 23.66654
```

4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Answer:



The above chart shows that the time required to name a color in incongruent condition is always more than that in congruent condition.



The boxplot indicates that the two groups have significant difference in median, and the two groups also have different ranges - with the Incongruent words group presenting much longer times.

5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

Answer:

We are performing t-test for hypothesis testing.

Alpha: 0.05

T-critical value : 2.069 for 23 degree of freedom.

n=24

Standard deviation for difference SD (Incongruent - Congruent) = 4.86

T- statistics = mean difference (Mean Incongruent - Mean Congruent) / (SD / sqrt(n))
= 7.96/ (4.86/sqrt(24))
= 8.02

Since t- statistics > t-critical hence we reject the null hypothesis.

As a result, on average the time took for population to solve both congruent task and incongruent task is different.

Confidence Interval : CI

Margin of error ME = t-critical value * Standard error

Standard error= SD / sqrt(n) = 4.86/ sqrt(24)
SE = 0.992043

ME = 2.069 * 0.992043 = 2.052537

mean difference (Mean Incongruent - Mean Congruent) = μ = 7.96

CI ($\mu - ME$, $\mu + ME$)

CI (7.96 - 2.0525 , 7.96 + 2.0525)

CI (5.9075 , 10.0125)

Hence we are 95% confident that participants on average took incongruent task 5.91 to 10.01 seconds longer than congruent task.

REFERENCES:

<http://stattrek.com/probability-distributions/t-distribution.aspx>

https://en.wikipedia.org/wiki/Stroop_effect