

# Test a Perceptual Phenomenon

December 21, 2018

## 0.0.1 Analyzing the Stroop Effect

Perform the analysis in the space below. Remember to follow [the instructions](#) and review the [project rubric](#) before submitting. Once you've completed the analysis and write-up, download this file as a PDF or HTML file, upload that PDF/HTML into the workspace here (click on the orange Jupyter icon in the upper left then Upload), then use the Submit Project button at the bottom of this page. This will create a zip file containing both this .ipynb doc and the PDF/HTML doc that will be submitted for your project.

- (1) What is the independent variable? What is the dependent variable?

Independent Variable: Congruency

Dependent Variable: Time

- (2) What is an appropriate set of hypotheses for this task? Specify your null and alternative hypotheses, and clearly define any notation used. Justify your choices.

Null Hypothesis,  $H_0$  - No change in time between two reading tasks (congruent or incongruent)  
Alternate Hypothesis,  $H_1$  - incongruent task take more time than congruent.

$H_0: \mu_i = \mu_c$  ( $\mu_i$  - population mean of incongruent values,  $\mu_c$  - population mean of congruent values)

$H_1: \mu_i > \mu_c$  ( $\mu_i$  - population mean of incongruent values,  $\mu_c$  - population mean of congruent values)

statistical test dependent t-test (two tailed)

We need to compare the means of two related groups to determine the statistically significant difference between two means. We are assuming distributions are nearly normal and we are comparing 2 dependent samples of data our sample size less than 30 and we don't know the population standard deviations

- (3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability. The name of the data file is 'stroop-data.csv'.

```
In [9]: #import used libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
%matplotlib inline
```

```
In [2]: #read dataset and show first 5 rows
df = pd.read_csv('stroopdata.csv')
df.head()
```

```
Out[2]:
```

	Congruent	Incongruent
0	12.079	19.278
1	16.791	18.741
2	9.564	21.214
3	8.630	15.687
4	14.669	22.803

```
In [12]: #measure a central tendency of each independent variable
congMean = df['Congruent'].mean()
incongMean = df['Incongruent'].mean()
congMean, incongMean
```

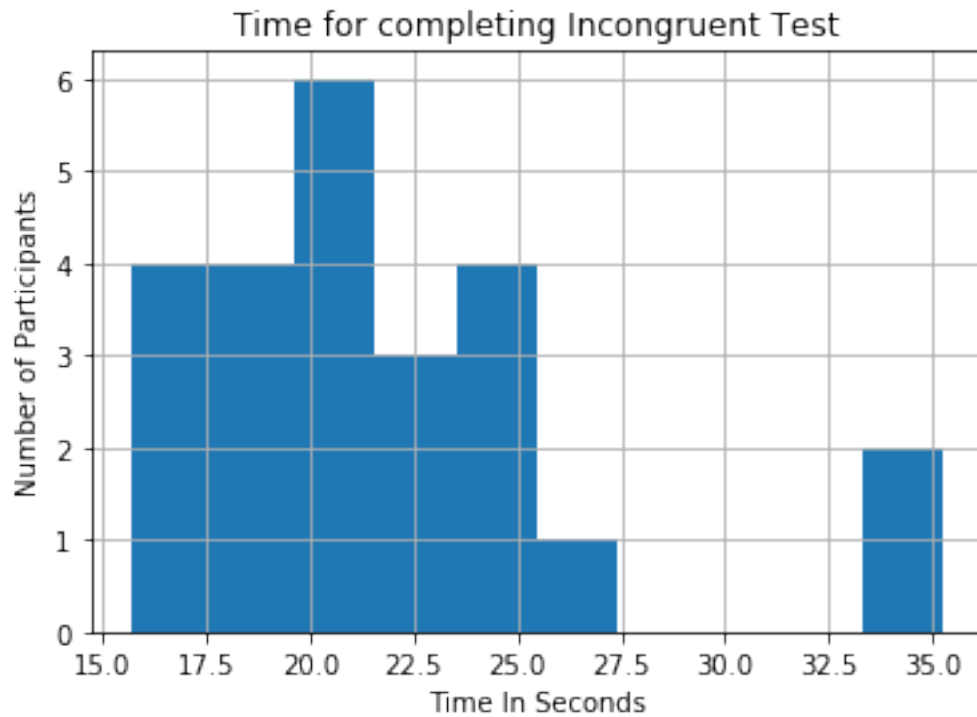
```
Out[12]: (14.051124999999999, 22.015916666666666)
```

```
In [13]: #calculate standard deviation for both variables
congStd = df['Congruent'].std()
incongStd = df['Incongruent'].std()
congStd, incongStd
```

```
Out[13]: (3.5593579576451955, 4.7970571224691376)
```

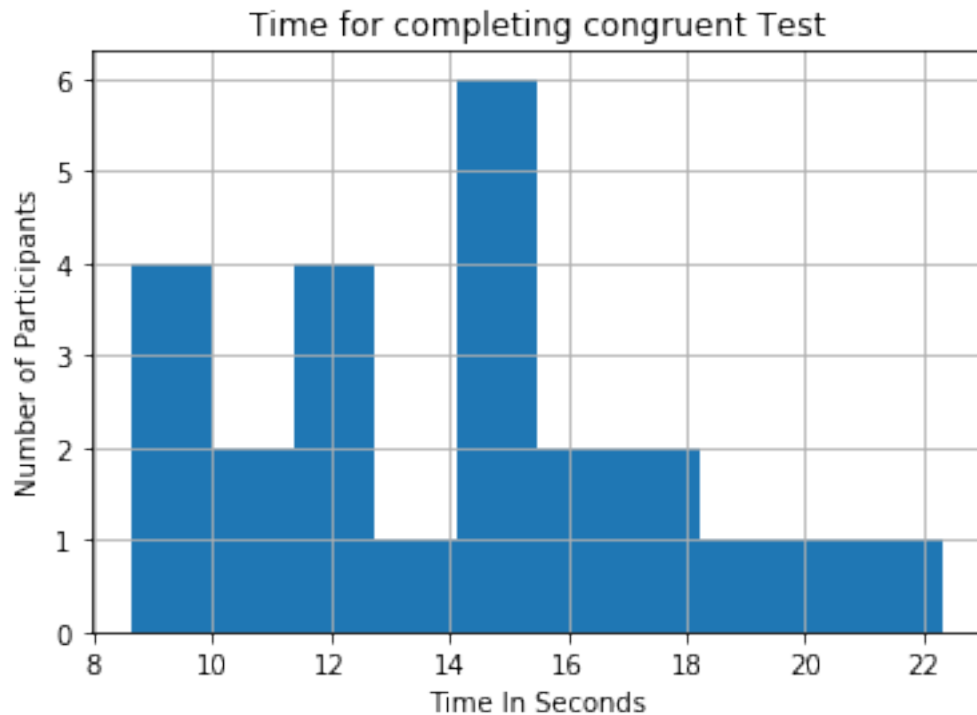
- (4) Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

```
In [7]: #histogram for incongrunt values
df['Incongruent'].hist()
plt.xlabel('Time In Seconds')
plt.ylabel('Number of Participants')
plt.title('Time for completing Incongruent Test');
```



The largest number of participants have solved their test in around 20 seconds, where 2 students are considered as outliers

```
In [8]: #histogram for congruent values
df['Congruent'].hist()
plt.xlabel('Time In Seconds')
plt.ylabel('Number of Participants')
plt.title('Time for completing congruent Test');
```



The largest number of participants have solved their test in around 14~16 seconds.

- (5) Now, perform the statistical test and report your results. What is your confidence level or Type I error associated with your test? What is your conclusion regarding the hypotheses you set up? Did the results match up with your expectations? **Hint:** Think about what is being measured on each individual, and what statistic best captures how an individual reacts in each environment.

```
In [11]: # T_Test
         tStat, pValue = stats.ttest_rel(a=df['Incongruent'],
                                         b=df['Congruent'])
         tStat, pValue
```

```
Out[11]: (8.020706944109957, 4.1030005857111781e-08)
```

```
In [14]: # Confidence
         rejectionConfidence = 1 - pValue
         rejectionConfidence
```

```
Out[14]: 0.99999995896999416
```

With 99% confidence and p-value of 4.1, we reject the null hypothesis. Which means that there is difference in completion time between congruent and incongruent tasks