

Test a Perceptual Phenomenon

September 19, 2017

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 and submit in the next section.

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

The independent variable is the words condition (congruent words or incongruent words).
The dependent variable is the time it takes to name the ink colors in equally sized lists.

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

HO - Null Hypothesis ($i-c = 0$): Time taken to name the ink colors for congruent and incongruent conditions are the same. In other words, there is no difference in the average identification time for both the conditions.

HA - Alternative Hypothesis ($i-c \neq 0$): Time taken to name the ink colors for congruent and incongruent conditions are not the same. In other words, there is a noticeable difference in the average identification time for both the conditions.

A Two-tailed, dependent samples t-Test comparing the difference in average response time for both the conditions should be performed. I believe that Two-tailed test is suitable because, it can be expected that time taken to identify the ink colors for incongruent condition will be different in any direction than that of the congruent condition if at all there is a difference.

t-Test is suitable for this problem because:

We have a sample size less than 30,

Population's standard deviation is unknown,

We can assume that the distributions are Gaussian.

Hence, this is a two directional problem and a two-tailed t-Test is suitable.

- (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 [4]: import os
import pandas as pd
df = pd.read_csv('stroopdata.csv')
df.describe()
```

```
Out[4]:    Congruent  Incongruent
count    24.000000   24.000000
mean     14.051125  22.015917
std      3.559358  4.797057
min      8.630000  15.687000
25%     11.895250  18.716750
50%     14.356500  21.017500
75%     16.200750  24.051500
max     22.328000  35.255000
```

From the above summary:

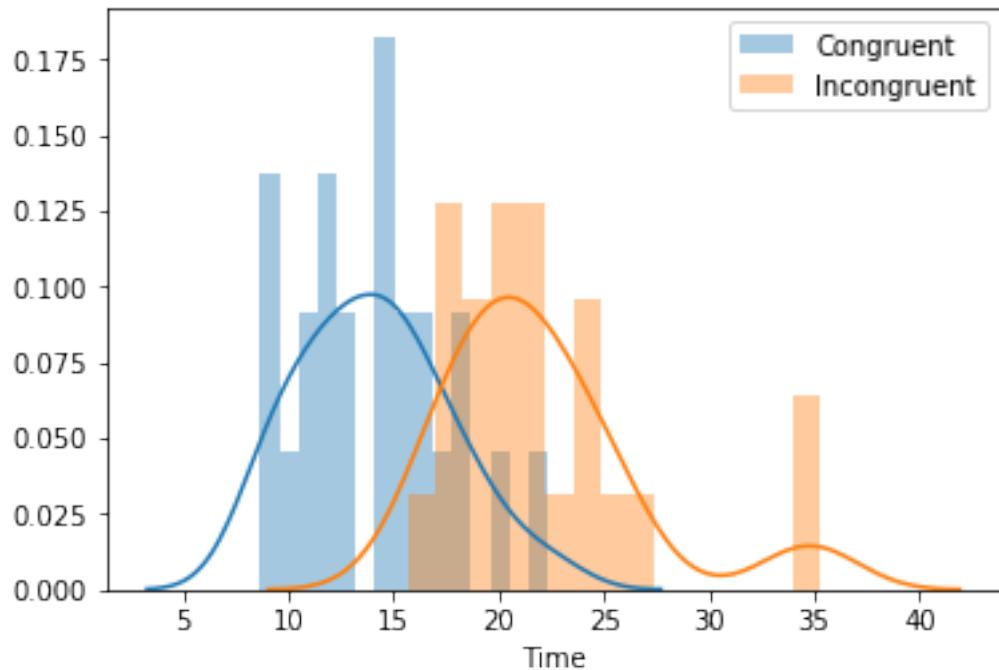
For Congruent Condition: Mean = 14.05, Standard Deviation = 3.56

For Incongruent Condition: Mean = 22.02, Standard Deviation = 4.80

- (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 [18]: import matplotlib.pyplot as plt
        import seaborn as sns
        %matplotlib inline
p1 = sns.distplot(df.Congruent, bins=15, label='Congruent')
p2 = sns.distplot(df.Incongruent, bins=15, label='Incongruent')
plt.xlabel('Time')
plt.legend()
plt.plot()
```

```
Out[18]: []
```



The above plot shows the comparison between both the conditions.

From the plot in blue, i.e the Congruent condition, it is very clear that people are taking less time to name the ink colors for this condition.

On the other hand, from the plot in orange, i.e the Incongruent condition, it is clear that people are taking more time to name the ink colors for this condition when compared to the other condition.

- (5) Now, perform the statistical test and report the results. What is the 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?

```
In [1]: #tD: -7.9648
#S: 4.86482691
#degree of freedom df: 23
#t-stat: -8.020706944
#at 0.05, t-critical: -2.06865761; 2.06865761
#P: 4.103E-08
#95% CI: (-10.019028 -5.910555)
```

The p-value is less than the alpha value 0.05. Hence, I rejected the null hypothesis.

It is clear that the difference between congruence and incongruence condition time difference is statistically significant.

This can conclude that the stroop effect is present. This is what I expected initially.