

Test a Perceptual Phenomenon

October 23, 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?

The independent variables are the Incongruent and Congruent tasks that define the tests independently from the rest of the tests.

The dependent variables are the time it takes to complete said tasks as those are the variables that are dependent on the how the test is performed.

- (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.

The Null hypothesis is that there will be not be a change or a lesser difference between the Incongruent and Congruent mean population test times.

The angles signify a mean, "In" representing the Incongruent Values, Con representing the congruent values.

$$H_0 = \langle In \rangle \leq \langle Con \rangle$$

The Alternative Hypothesis is that there will be a significant increase in time relating from the Incongruent and Congruent mean population test times.

$$H_1 = \langle In \rangle > \langle Con \rangle$$

For this experiment a dependent two tailed. This test should work with the participants being 24 and the participant taking the test twice giving a uniform data structure in the columns, to compare the mean of the variables. With the population we can assume a mostly normal distribution.

- (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 [1]: # Perform the analysis here
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [2]: 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 [3]: # Calculate the mean of both types of tests
```

```
df['Congruent'].mean(), df['Incongruent'].mean()
```

```
Out[3]: (14.051124999999999, 22.015916666666666)
```

```
In [4]: #Calculate the Standard Deviations of both tests
```

```
constd = np.std(df['Congruent'])
```

```
inconstd = np.std(df['Incongruent'])
```

```
print("Congruent STD: ", constd, ', Incongruent STD :', inconstd)
```

```
Congruent STD:  3.48441571277 , Incongruent STD : 4.69605513451
```

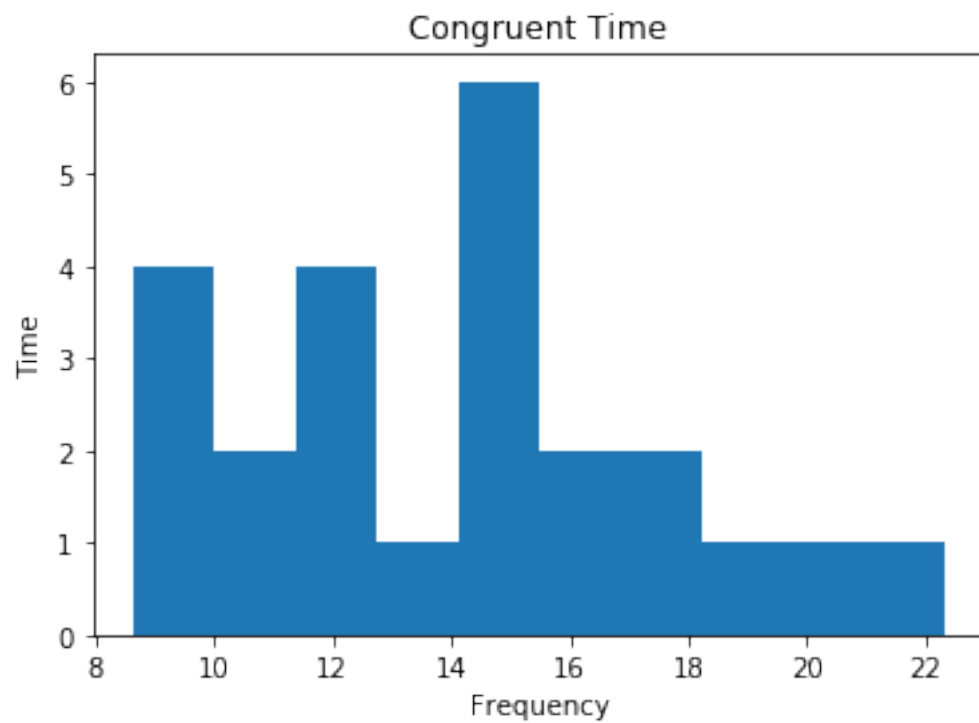
The average time for the Congruent time was 14.05 with a standard deviation of 3.48. The Incongruent tests took an average of 22.02 with a standard deviation of 4.70. These times were rounded up from the second decimal place.

- (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 [5]: # Build the visualizations here
```

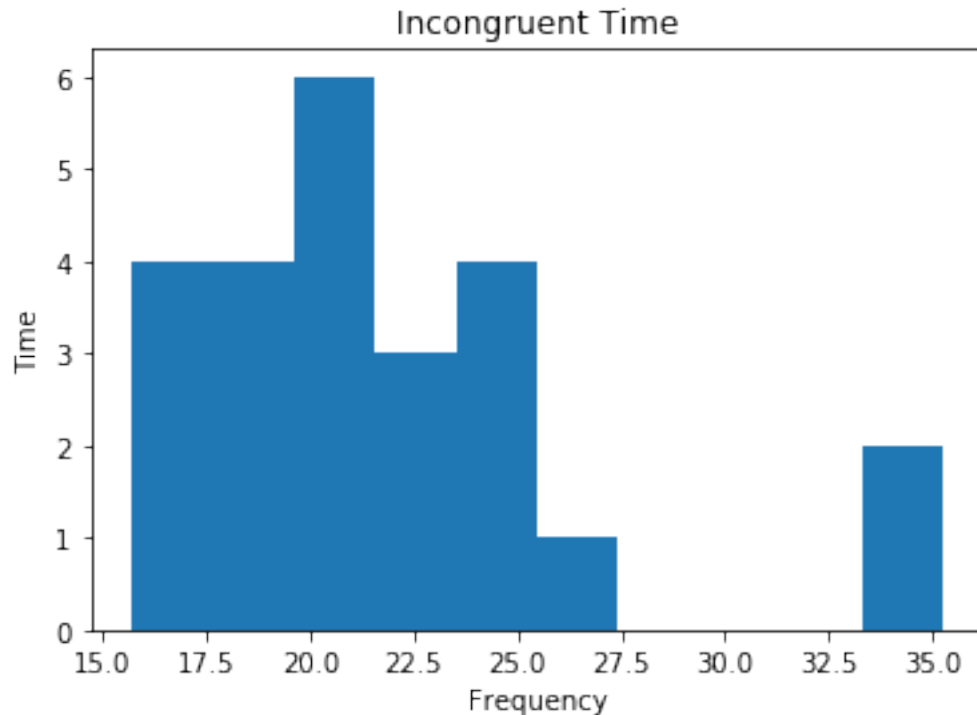
```
plt.hist(df['Congruent']);
plt.title('Congruent Time')
plt.xlabel('Frequency')
plt.ylabel('Time')
```

Out[5]: Text(0,0.5,'Time')



```
In [6]: plt.hist(df['Incongruent']);  
plt.title('Incongruent Time')  
plt.xlabel('Frequency')  
plt.ylabel('Time')
```

Out[6]: Text(0,0.5,'Time')



Both Histograms have a slightly right skewed tail indicating a mean high than the median.

- (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 [21]: # Loading statsmodel

```
import statsmodels.stats.api as sms
from scipy import stats

cong = df['Congruent']
incong = df['Incongruent']

stattest = sms.CompareMeans(sms.DescrStatsW(incong), sms.DescrStatsW(cong))

confres = stattest.tconfint_diff(usevar='unequal')
testres = stats.ttest_rel(df['Congruent'], df['Incongruent'])
critval = stats.t.ppf(0.95, 23)
critconfirm = stats.t.cdf(1.71387, 23)

print(testres, confres, 'Critical T-Value:', critval)
```

```
Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08) (5.504885145169359,
```

With the results of the T-statistic(8.02) being greater than the Critical T Value of (1.714) and the confidence comparison being so far away from zero we can reject the null hypothesis of the test. This means that the congruent and the incongruent produced different results on average

The confidence level is set to 5% in the `tconfint_diff` function giving a 95% confidence in the result of the analysis.

A type one error would be to reject that the congruent and incongruent times are not different when that hypothesis would be true.

- (6) Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!

--write answer here--