Test a Perceptual Phenomenon

February 4, 2019

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?

An independent variable is the variable that is changed, measured, or controlled in a scientific experiment to test the effects on the dependent variable and in this scenario are type of test congruent and incogruent (thoughtco, 2018).

A dependent variable is the variable being tested and measured in a scientific experiment but, it depends on the independent variable in our case is the response time in second(thoughtco, 2018).--

(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.

H0: I C
H1: I > C

The Null hypothesis states that the population mean of incongruent is less than or equal to the population mean of congruent. However, the alternative hypothesis states that the population mean of incongruent is greater than the population mean of congruent.

We assume That the distribution of the differences in the dependent variable between the two related groups should be normally distributed. Normal distributions are symmetric around the center. Also The observations are independent of one another.

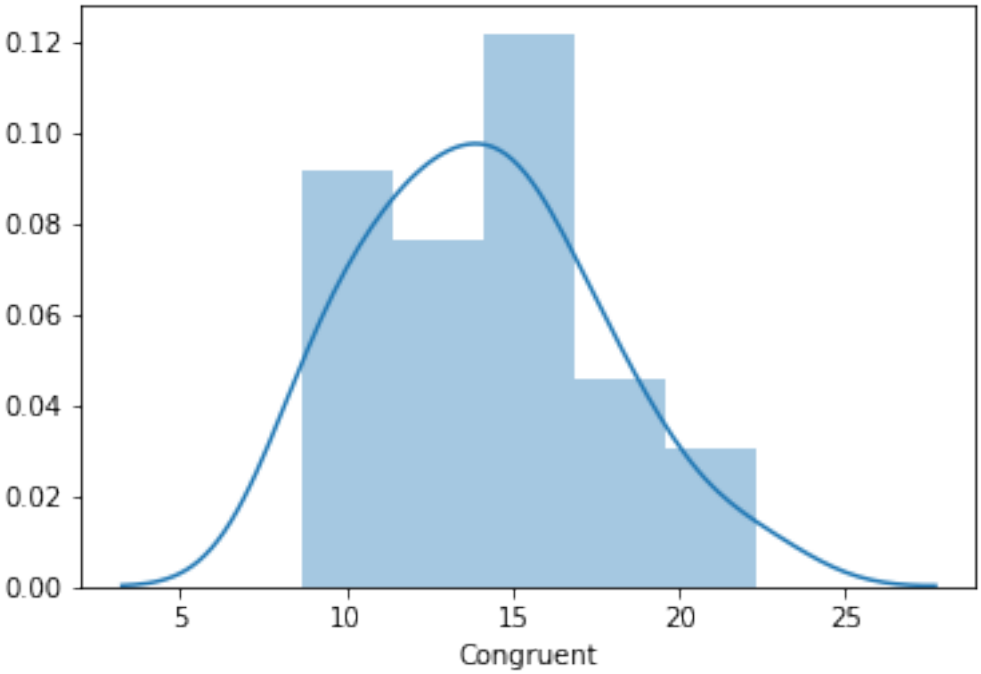
(3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability. The name of the data file is 'stroop-data.csv'.

```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        from scipy.stats import t
        %matplotlib inline
```

```
In [2]: test=pd.read_csv('stroopdata.csv')
        test.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]: test.describe()

Out[3]:        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

In [4]: print(test.mean(axis=0))

        print(test.median(axis=0))
        print("standard deviation for congruent",(np.std(test['Congruent'].values)))
        print("standard deviation for Incongruent",(np.std(test['Incongruent'].values)))

Congruent      14.051125
Incongruent    22.015917
dtype: float64
Congruent      14.3565
Incongruent    21.0175
dtype: float64
standard deviation for congruent 3.48441571277
standard deviation for Incongruent 4.69605513451
```

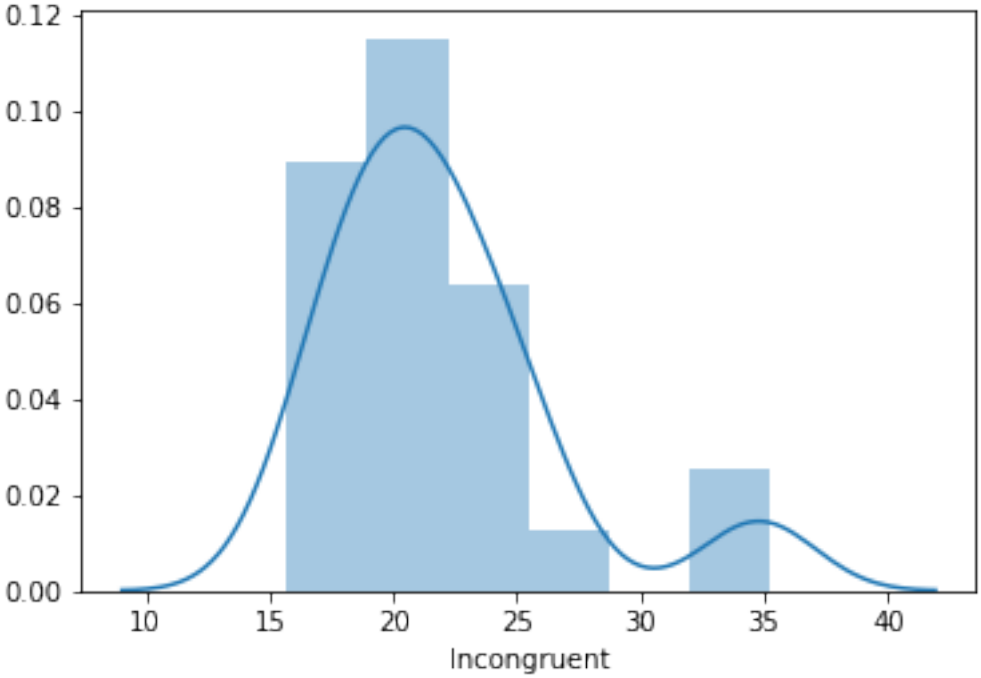
(4) Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

```
In [5]: sns.distplot(test['Congruent']);
```



The data looks normally distributed since the mean is close to the peak

```
In [6]: sns.distplot(test['Incongruent']);
```



The histogram shows that the mean is close to the peak which means it's normally distributed

(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.

The Test is two tailed t- test so find the critical value of t at 99% confidance level, 23 degrees of freedom , and the sample size 24. The dependent t-test compares the mean of two paired groups to see if there are differences between their means.

```
In [7]: print('The sample size is' ,test['Congruent'].size)
        print('The sample size is',test['Incongruent'].size)
```

```
The sample size is 24
The sample size is 24
```

```
In [8]: t.ppf(.99, 23)
```

```
Out[8]: 2.4998667394943976
```

For a confidence level of 99% and 23 degrees of freedom, the t-critical value is 2.4998, and the difference of the means is: 22.02 - 14.05 = 7.97

```
In [9]: test['Difference'] = test['Congruent'] - test['Incongruent']
        print("standard deviation for congruent",(test['Difference'].std(axis=0)))
```

```
standard deviation for congruent 4.86482691036
```

```
In [10]: import math
         7.97/(4.8648 / math.sqrt(24))
```

```
Out[10]: 8.025996238275749
```

The t-statistic (8.026) is greater than our critical value (2.4998),So we can reject the null hypothesis.

The ruslt shows that it takes less time to do the congruent task than it does to do the incongruent task

Reference https://www.thoughtco.com/independent-and-dependent-variables-differences-606115

```
In [ ]:
```