Test a Perceptual Phenomenon

February 23, 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?

Independent variables are the variables with congruent words condition and incongruent word conditions. Whereas, dependent variables here are the time required by the user to finish the task. Here, Independent variables are :- Word conditions Dependent variables :-Response in seconds

(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 Hypothsis - H0 - No change in time between two reading tasks (congruent or incongruent) Alternate Hypothesis - H1 - incongruent task take more time than congruent. H0: i c (i - population mean of incongruent values, c - population mean of congruent values)

H1: i > c (i - population mean of incongruent values, c - population mean of congruent values)

The appropriate statistical test type for this dependent samples test should be t-Test. This is because same subjects are assigned to two different conditions. Since we don't have any population parameters, a z-test would not be appropriate here.

(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 [2]: # Perform the analysis here
        import math
        import numpy as np
        import pandas as pd
        from scipy.stats import t
        import seaborn as sns

        %matplotlib inline
```

```
In [3]: df = pd.read_csv('./stroopdata.csv')

In [6]: df.head()

Out[6]:     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 [7]: df.mean()

Out[7]: Congruent      14.051125
        Incongruent    22.015917
        dtype: float64

In [8]: df.std()

Out[8]: Congruent      3.559358
        Incongruent    4.797057
        dtype: float64

In [11]: print("Congruent standard deviation {0:.3f}".format(np.std(df['Congruent'].values)))
         print("Incongruent standard deviation {0:.3f}".format(np.std(df['Incongruent'].values))

Congruent standard deviation 3.484
Incongruent standard deviation 4.696
```
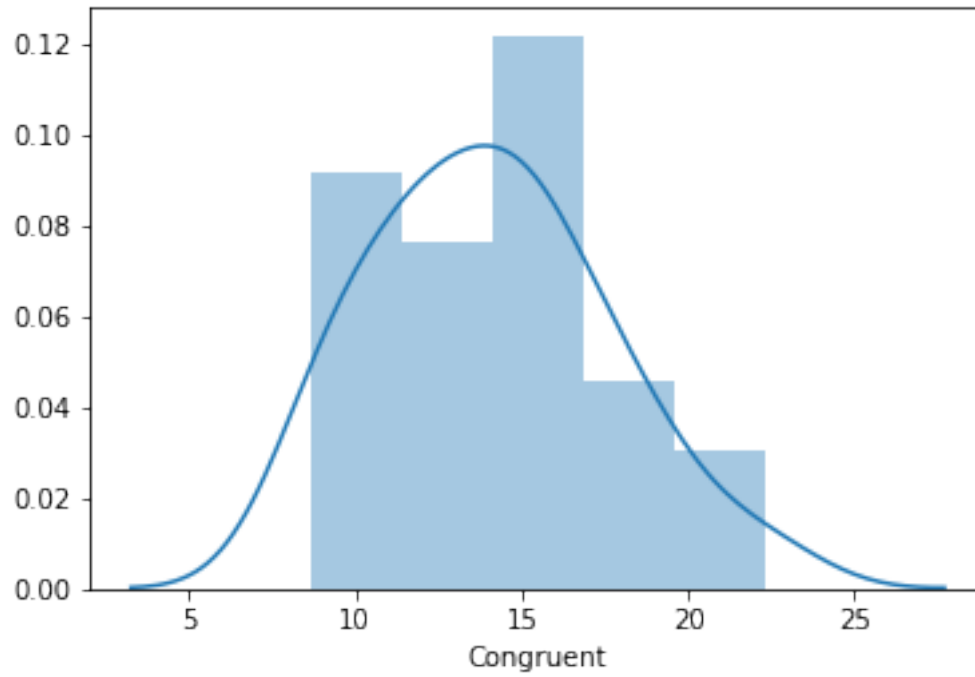
(4) Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.
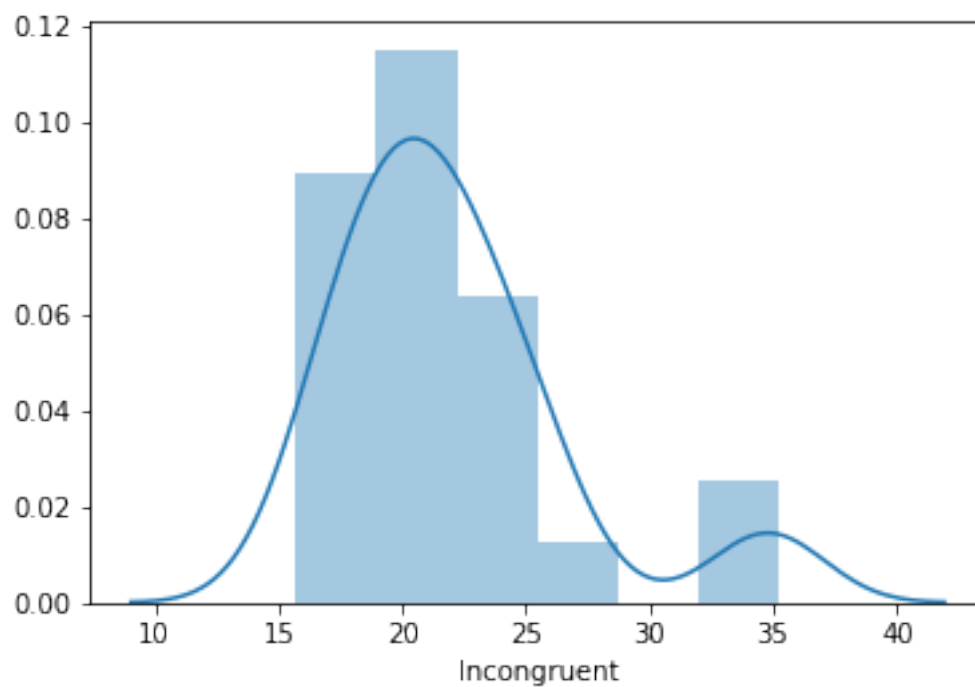
```
In [13]: # Build the visualizations here
         sns.distplot(df['Congruent'])

Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x7f20b8233978>
```

In [14]: sns.distplot(df['Incongruent'])

Out[14]: <matplotlib.axes._subplots.AxesSubplot at 0x7f20b68314a8>

Observation :

From the histogram plots, we can observe that depsite graphs being positively skewed, the mean is closer to the respective peaks hence indicating it's a normal distribution.

(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 [17]: # Perform the statistical test here
         df['Congruent'].size
```

```
Out[17]: 24
```

```
In [18]: df['Incongruent'].size
```

```
Out[18]: 24
```

Now, to calculate T-critical value for 95% confidence level and 23 d.f

```
In [19]: t.ppf(.95, 23)
```

```
Out[19]: 1.7138715277470473
```

For standard deviation,

```
In [22]: df['Difference'] = df['Congruent'] - df['Incongruent']
         print("Congruent standard deviation {0:.4f}".format(df['Difference'].std(axis=0)))
```

Congruent standard deviation 4.8648

For T-static,
Point estimate for the difference of the means is: 22.02 - 14.05 = 7.97

```
In [23]: 7.97/(4.8648 / math.sqrt(24))
```

```
Out[23]: 8.025996238275749
```

CONCLUSION
Out T-static value is greater than our critical value. So, we can reject this null hypothesis (H0).