Test a Perceptual Phenomenon

December 25, 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?

1-independent variable: **Congruent,Incongruent**
2-dependent variable: **The response time to pronounce the color**

(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.

1- H0 (Null Hypothesis)
There is no difference in response time between incongruent and congruent.
H0: (C - I)=0
2- H1 (Alternate Hypothesis)
The response time of the incongruent will not be equal the the response time of the congruent.
H0: (C - I)!=0

(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 matplotlib.pyplot as plt
        import pandas as pd
        import numpy as np
        import seaborn as sns
        import scipy.stats as st

        df = pd.read_csv("stroopdata.csv")

In [2]: df.head(60)
```

```
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
        5       12.238       20.878
        6       14.692       24.572
        7        8.987       17.394
        8        9.401       20.762
        9       14.480       26.282
        10      22.328       24.524
        11      15.298       18.644
        12      15.073       17.510
        13      16.929       20.330
        14      18.200       35.255
        15      12.130       22.158
        16      18.495       25.139
        17      10.639       20.429
        18      11.344       17.425
        19      12.369       34.288
        20      12.944       23.894
        21      14.233       17.960
        22      19.710       22.058
        23      16.004       21.157
```

In [3]: df.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
```

After checking the data and the observations count above, we decided to use the (t test) WHY !!

- because we want to compare two samples with the same size and figure out if the dependant variable which is the response time will be affectedor not.

- The size of sample is small, 30 observations only, so without any dought this is the perfect senario for us to use (t test).

If the size of the samples was big then we would've used the (z test)

```
In [4]: congruent_mean = df['Congruent'].mean()
        incongruent_mean = df['Incongruent'].mean()
        print("the mean of the Congruent is: ")
        print(congruent_mean)
        print("the mean of the Incongruent is: ")
        print(incongruent_mean)

the mean of the Congruent is:
14.051125
the mean of the Incongruent is:
22.0159166667
```

```
In [5]: congruent_std = df['Congruent'].std()
        incongruent_std = df['Incongruent'].std()
        print("the standard deviation of the Congruent is:")
        print(congruent_std)
        print("the standard deviation of the Incongruent is:")
        print(incongruent_std)

the standard deviation of the Congruent is:
3.55935795765
the standard deviation of the Incongruent is:
4.79705712247
```

```
In [6]: print("the standard deviation for Congruent: {0:.5f}".format(np.std(df['Congruent'].valu
        print("the standard deviation for Incongruent: {0:.5f}".format(np.std(df['Incongruent'].

the standard deviation for Congruent: 3.48442
the standard deviation for Incongruent: 4.69606
```

--write answer here--

(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]: a=sns.distplot( df["Congruent"],color='r')
        plt.xlabel('test time',color='r')
        plt.ylabel('# of observations',color='r')
        plt.title('Time VS observations for Congruent',color='r');
```

Time VS observations for Congruent
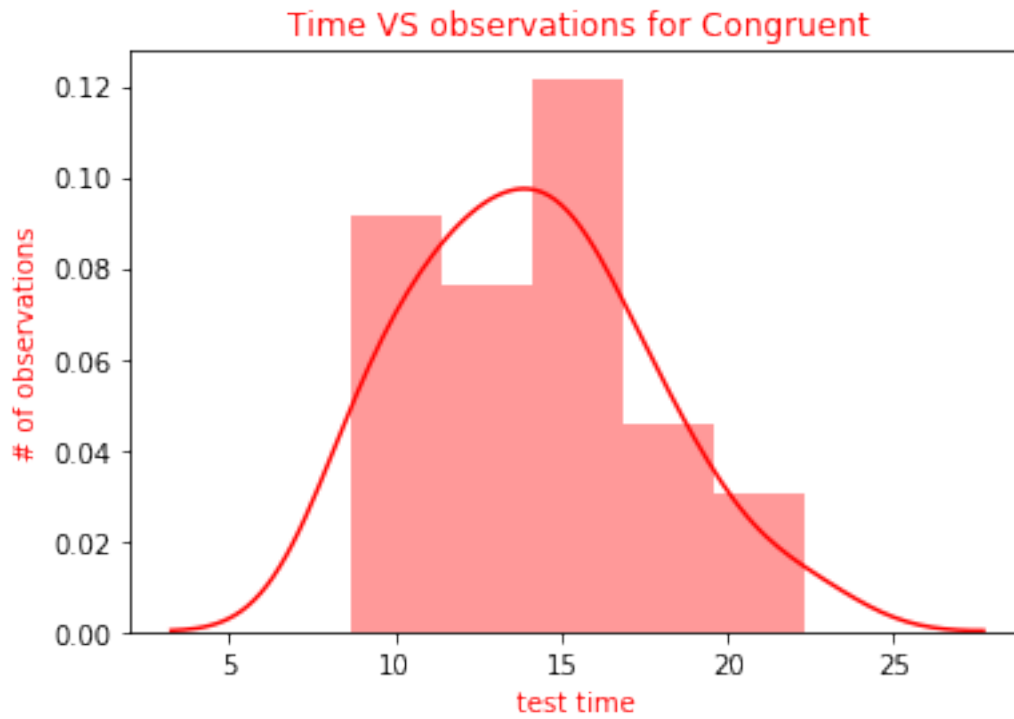
In [8]: a=sns.distplot(df["Incongruent"],color='b')
        plt.xlabel('test time',color='b')
        plt.ylabel('# of observations',color='b')
        plt.title('Time VS observations for Incongruent',color='b');

4

Time VS observations for Incongruent
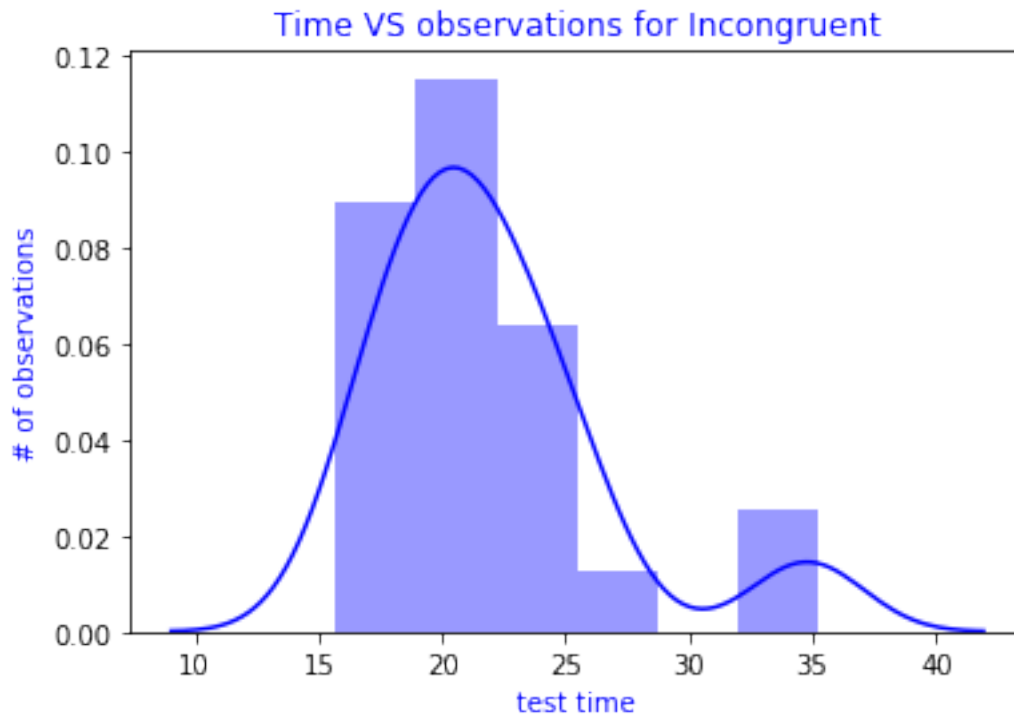
In [9]: sns.distplot(df["Congruent"],color='r')

```
plt.xlabel('test time',color='r')
plt.ylabel('# of observations',color='r')
plt.title('Time VS observations for Congruent',color='r');


sns.distplot( df["Incongruent"],color='b')

plt.xlabel('test time',color='r')
plt.ylabel('# of observations',color='r')
plt.title('Time VS observations for Incongruent',color='r');
```

Time VS observations for Incongruent
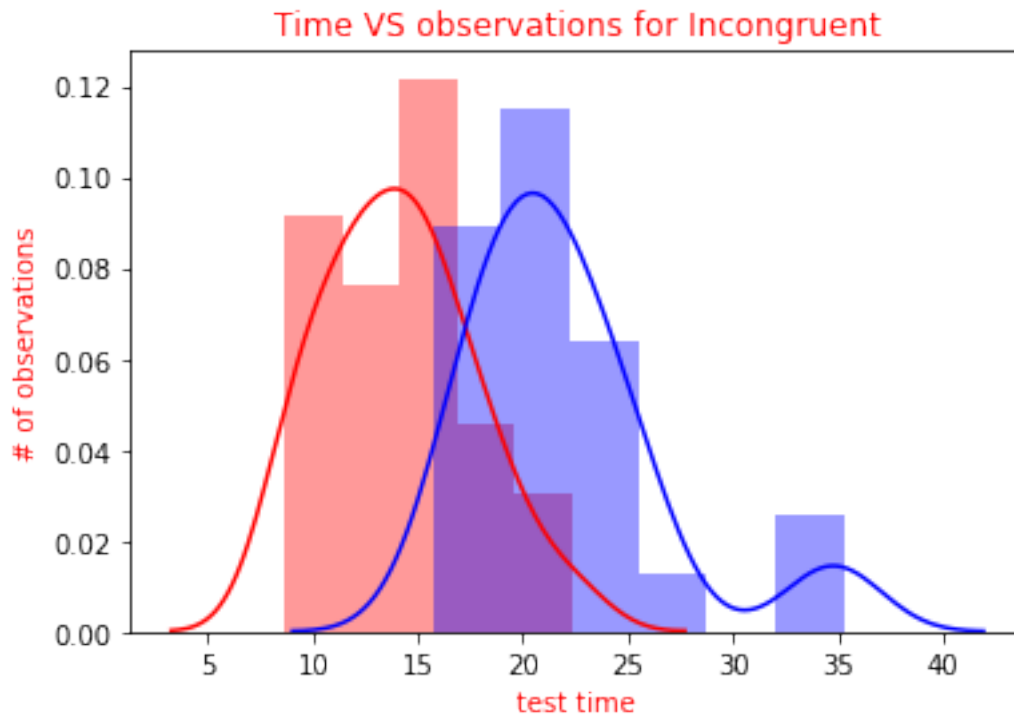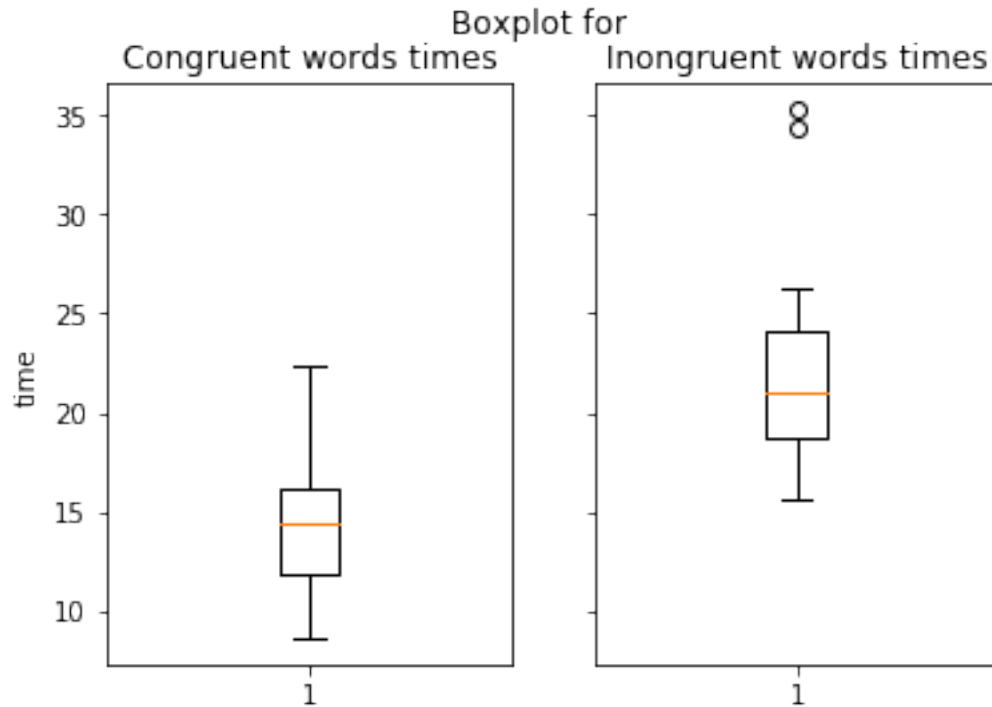
```
In [10]: x = df['Congruent']
         y = df['Incongruent']


         f, (Congruent, Inongruent) = plt.subplots(1, 2, sharey=True)
         f.suptitle('Boxplot for')

         Congruent.boxplot(x)
         Congruent.set_title('Congruent words times')
         Congruent.set_ylabel('time')

         Inongruent.boxplot(y)
         Inongruent.set_title('Inongruent words times')

         plt.show()
```

Boxplot for
Congruent words times   Inongruent words times

the two distributions are similar to each other but the Incongruent has an outliers, we did use the boxplot for this,to make this outliers data more visiable and clear.

(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,p = st.ttest_rel(df['Incongruent'],df['Congruent'],axis=0)

         print(t)

         print(p)

         alpha = 0.005

8.02070694411
4.10300058571e-08
```

Confidence Interval = 95%
= .05
degrees of freedom = 23
one-tailed critical statistic value = 1.714

7

calculated t-value = 8.02
p-value = 0.00000004

A paired t-test was run on a sample of 24 participants to determine whether there was a statistically significant mean difference between the reaction time between the congruent vs. incongruent condition. Typically it takes participants longer to say the ink colors (22.02 ś 4.8 sec) than it does to read the words (14.05 ś 3.56 sec). At the 95% confidence level (= .05) and 23 degrees of freedom, the critical statistic value for a one-tailed test is 1.714. The calculated t-value for the difference between the congruent and incongruent conditions is 8.02. The calculated t-value is greater than the critical t-value (t-table value). Additionally, the p-value is less than the alpha level: $p < .05$; therefore, the null hypothesis is rejected

(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!

Refrences

- scipy-lectures.

- users.sussex.ac.uk.

- programiz.

- towardsdatascience.

- knowledgetack.