Test a Perceptual Phenomenon

January 17, 2019

0.0.1 Analyzing the Stroop Effect

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

the independant variables : Names of colors appeared in black ink(congruent), Names of colors in a different ink than the color named (incongruent) the dependant variables: the reaction time of the individuals

(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: the reaction time for congruent colors and words = the reaction time for incongruent colors and words u(mu: mean of reaction time congruent)= u(mu: mean of reaction time incongruent) H1: the reaction time for congruent colors and words > the reaction time for incongruent colors and words u(mu: mean of reaction time congruent)> u(mu: mean of reaction time incongruent)

(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 'stroopdata.csv'.

```
In [2]: # Perform the analysis here
        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import scipy.stats

        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 [4]: df.shape
```

```
Out[4]: (24, 2)
```

The test to be used in this case is a two sample paired T-Test and this because: we have 2 samples and we are trying to prouve that the difference between the two means (of each sample) is not random but there is a real difference in the reaction time between the 2 tests. In addition, we are using the paired T-test because the same individuals are used for experiments with congruent and incongruent words , we call it self pairing in this test, I assume that the distribution of the mean difference is normal, and also that the pairs of observations are indpendant and random.

```
In [2]: # measures of central tendency used : mean, median=second quartile
        # measures of variability : standard deviation, min, max, first and third quartile

        df.describe()
```

```
Out[2]:         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 [3]: meanC= df['Congruent'].mean()
        meanI= df['Incongruent'].mean()
        varC = df['Congruent'].var()
        varI = df['Incongruent'].var()
        meanC, meanI, varC, varI
```

```
Out[3]: (14.051124999999999,
          22.015916666666666,
          12.669029070652176,
          23.011757036231884)
```

(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 [4]: # Build the visualizations here
        df.plot()
        plt.title('Congruent Vs Incongruent')
        plt.ylabel('Reaction Times in seconds ');
```

Congruent Vs Incongruent
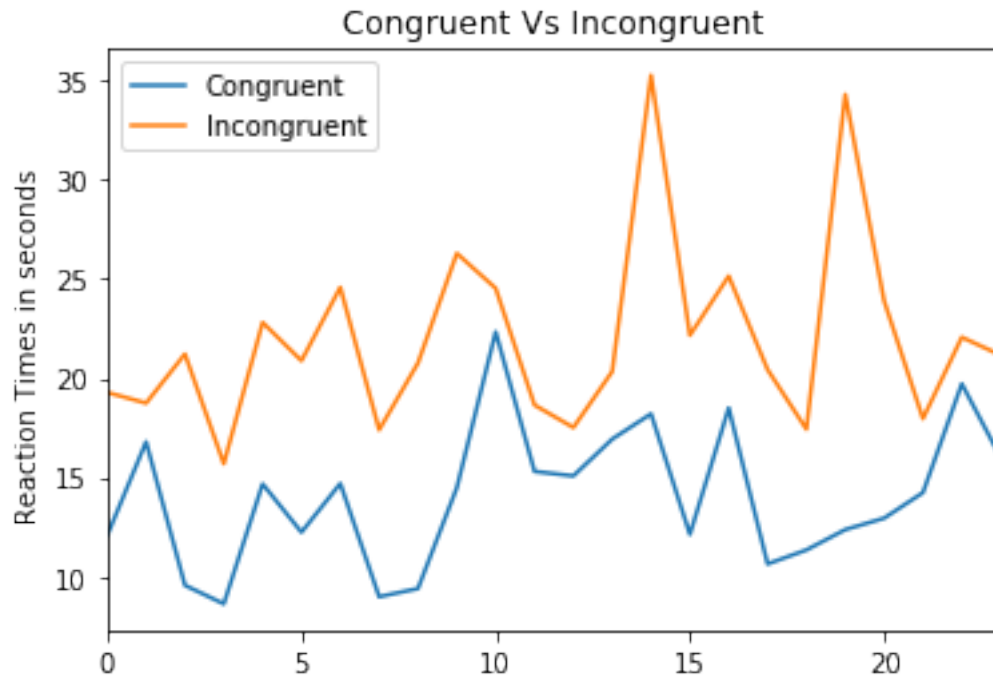
In the graph I noticed that for all individuals it took longuer to react to the incongruent than it took them to react to the congruent words to the congruent.

```
In [5]: x = df['Congruent']
        y = df['Incongruent']
        plt.scatter(x, y)
        plt.ylabel('Incongruent')
        plt.xlabel('Congruent')
        plt.title('Congruent Vs Incongruent reaction times ');

        z = np.polyfit(x, y, 1)
        p = np.poly1d(z)
        plt.plot(x,p(x),"r--");
```

Congruent Vs Incongruent reaction times
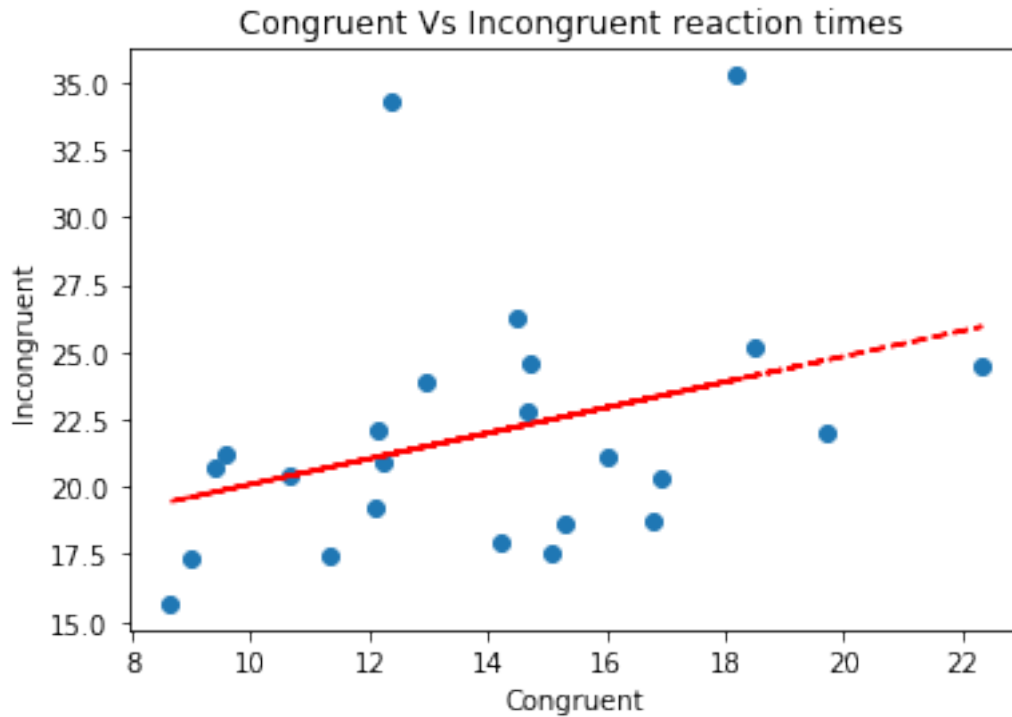
I noticed that there is a positive correlation between congruent and incongruent reaction times.

(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 [7]: # Perform the statistical test here
        # Alpha = 0.05, so if P-value <= alpha we reject the H0, if not than we fail to reject H
        # the confidence level is 95%
        # the variances are inequal
        scipy.stats.ttest_rel (x,y)
```

```
Out[7]: Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08)
```

after running the T-test, the P value = 4.1030e-08, which is less than Alpha= 0.05. This means that I reject the null hypothesis. the results obtained show that no matter how many times I repeat the experiment I will get the same results for more than 99% of the time the results matched my expectations. When I did the experiement I obtained a different reaction time for the congruent and incongruent and the incongruent was bigger

4