Test a Perceptual Phenomenon

January 25, 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 variable:** List of congruent and incongruent words.
- **Dependent variable:** Duration of the time that taken by participants to complete the test.

(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.

- H_0: $_C = _I$ >Null hypothesis: Which means the population of response time for congruent words equals the population response for incongruent words.
- H_1: $_C != _I$ > Alternative hypothesis: Which means the population of response time for congruent words does not equal the population response for incongruent words.

```
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        % matplotlib inline
        df = pd.read_csv('stroopdata.csv')
        len(df)

Out[1]: 24
```

We can notice that the data is a sample of 24 pairs, so I will implement a paired sample t-test to compare the population of response time for congruent and incongruent words.

(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
        print(df)
```

```
        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
```

We can notice that the mean and the std of incongruent are larger than congruent.

(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
        plt.hist(df.Congruent, alpha=0.6, label='Congruent Words')
        plt.hist(df.Incongruent, alpha=0.6, label='Incongruent Words')
        plt.xlabel('Response time (seconds)')
        plt.ylabel('Frequency')

```
plt.legend(loc='upper right')
plt.title('Congruent vs Incongruent');
```
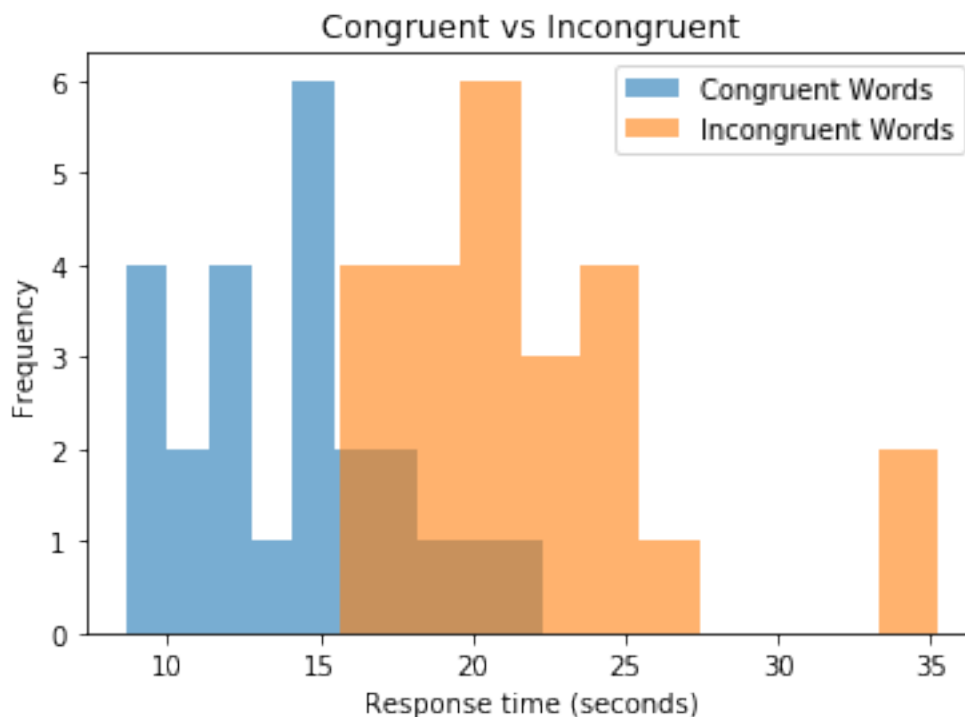

From this histogram we can notice that the response time of the incongruent words is longer than congruent words.

(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 [5]: # Perform the statistical test here
        from scipy.stats import ttest_rel
        ttest_rel(df['Congruent'], df['Incongruent'])
```

```
Out[5]: Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08)
```

The results show that t-statistic equals to -8.02 and the p-value equals 4.1.

Critical value calculation: based on t-table (link in references) - alpha value equals 0.05 - 0.25 for each tail - so t-critical equals +2.069 and -2.069

Since t-statistic (-8.02) is smaller than t-critical (-2.07), we will reject the null hypothesis. So there is a significant difference in response time between congruent and incongruent words. We conclude that reconizing the congruent is faster than incongruent words, which meets our results in Q4.

0.0.2 References

- http://www.ttable.org