Test a Perceptual Phenomenon

January 26, 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 and submit in the next section.

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

The independent variable is the test given (congruent or incongruent). The dependent variable is the participant's reaction time.

(2) What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

An appropriate hypothesis would be that the Stroop effect is indeed present, which would be indicated by the congruent and incongruent datasets being very different from one another. This could be shown by examining the descriptive statistics of each set (for instance, the mean would show this), as well as using a t-Test to determine if the differences between the congruent and incongruent sets have statistical significance.

(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]: # Perform the analysis here

        import pandas as pd

        stroop = pd.read_csv('stroopdata.csv')

        print(stroop)

        stroop.describe()
```

```
   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


Out[1]:          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
```

For the congruent set: The mean is 14.05 and the median is 14.35. The standard deviation is 3.559. For the incongruent set: the mean is 22.02 and the median is 21.02. The standard deviation is 4.797.

In both, the standard deviation is relatively smaller than the mean.

Since the mean and median in both are very close, we know that the distribution will be symmetrical.

Finally, we can see that there is a distinct difference in the timing for each of these datasets- as we initially believed, the Stroop effect could be reasonably occurring as the congruent data set's average is 7 seconds faster than the average of the incongruent data set.

(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
        import numpy as np
        import matplotlib.pyplot as plt
```
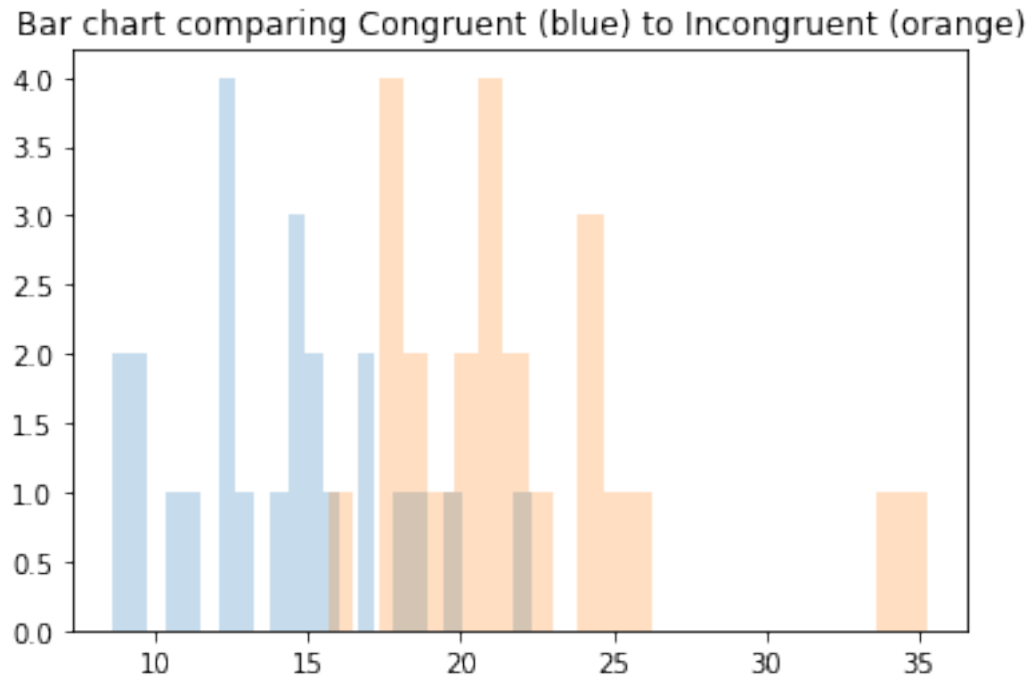
```
x = stroop['Congruent']
y= stroop['Incongruent']

n, bins, patches = plt.hist(x, 24, alpha=0.25)
n, bins, patches = plt.hist(y, 24, alpha=0.25)
plt.title('Bar chart comparing Congruent (blue) to Incongruent (orange)')
plt.show
```

Out[4]: <function matplotlib.pyplot.show>

Bar chart comparing Congruent (blue) to Incongruent (orange)
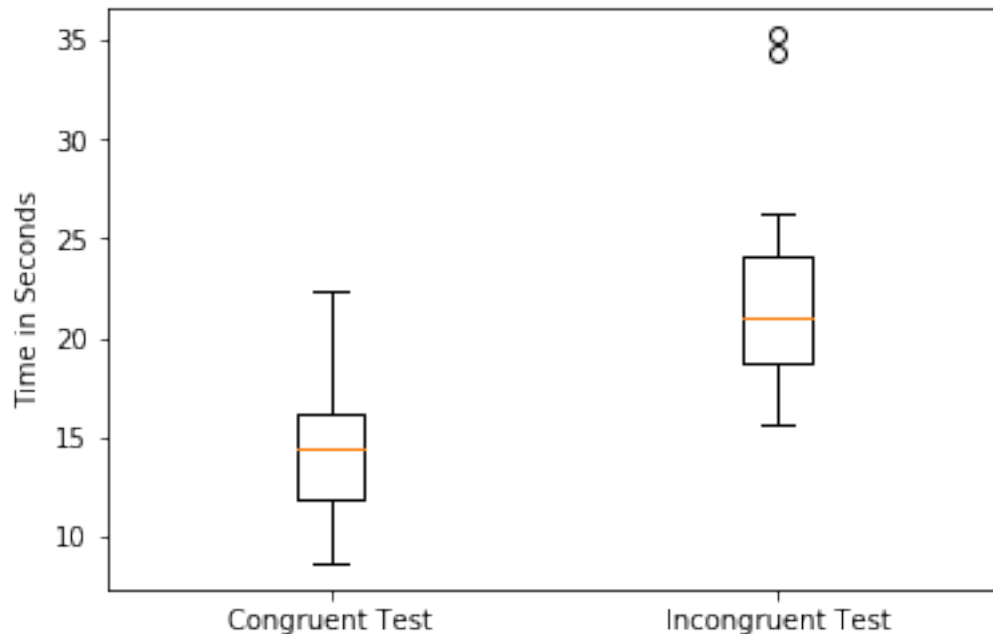
```
In [7]: fig, ax = plt.subplots()
        ax.boxplot(stroop.T)

        plt.ylabel('Time in Seconds')
        plt.xticks([1,2], ['Congruent Test', 'Incongruent Test'])
        plt.show()
```

The first thing I noticed is that there's more overlap that I initially expected between the congruent and incongruent sets (between 15-20 seconds in the first bar chart).

For this reason, I created a box plot, to try to see how much overlap there really was- in fact, the box plot shows they don't overlap that much- more than 75% doesnt overlap with other 75%.

(5) Now, perform the statistical test and report the results. What is the confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

```
In [8]: #check if variance equal or not
        import scipy.stats as stats

        stats.bartlett(x, y)

Out[8]: BartlettResult(statistic=1.9756393730727815, pvalue=0.15985054445707891)

In [9]: #same variance
        #p-val is over 0.05

In [10]: # Perform the statistical test here

         stats.ttest_ind(stroop['Congruent'], stroop['Incongruent'], equal_var=True)

Out[10]: Ttest_indResult(statistic=-6.5322505539032285, pvalue=4.5949489622951854e-08)
```

The p-value for this test is very low (4.59e-08). The statistic is -6.532.
We perform a t-test to see if the results are statistically different from one another.
Since the p-value is well below 0.05, our null hypothesis is wrong- the stroop effect is occurring.

4

6. Optional Question: What is responsible for the effects observed? What is a similar task that would have the same effect?

These effects are caused by the Stroop Effect. It's taking particiapants a longer time to correctly identify the colors in the incongruent test.

A similar task could be trying to identify anything incorrectly labelled- perhaps an apple labeled as 'pear' (though, we'd have to account for other indicators present, as the smell and taste could also affect results).