Test a Perceptual Phenomenon

March 24, 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 : This is the part of the experiment that is changed, it can either be congruent word condition(word & colour match) or incongruent word condition(word & colour do not match) Dependent variable : The dependent variable measures reaction time & it is the part of stroop effect that is measured. In this scenario it is the time it takes to state the name of the ink colors in an equally-sized text list.

(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.

For my test, I used the two tailed t-test.With this, I am able to compare any significant statistical difference between the mean of both paired groups.

I decided to do this because we have no details of the population but just 24 pairs of sample data set.

The same subjects were tested for congruent and incongruent words. By using the same subject to test two different conditions, we eliminate the individual differences that occur between subjects.

Null Hypotesis : No signicant difference exists between the average time it takes to name the Ink colour in both Incongruent & congruent conditions thus both Incongruent & congruent conditions are equal. H0(Null Hypothesis): Time(Incongruent)=Time(congruent) or i=c

Alternative Hypothesis: There is a signicant difference between the average time it takes to name the Ink colour in both Incongruent & congruent conditions thus both Incongruent & congruent conditions are not equal. H1(Null Hypothesis): Time(Incongruent)Time(congruent) or ic

(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 [5]: import pandas as pd

        #Read CSV file & save to ted
        ted = pd.read_csv('stroopdata.csv')
        # print csv data
        print(ted)
        # With describe i show a quick statistical summary of the data
        ted.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
```

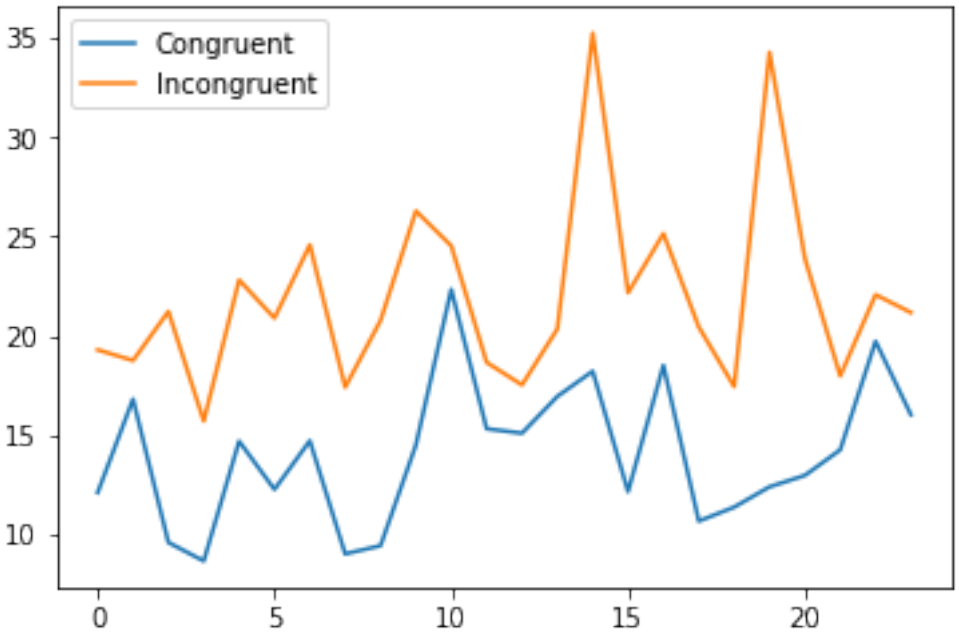
```
Out[5]:        Congruent  Incongruent
        count  24.000000    24.000000
        mean   14.051125    22.015917
        std     3.559358     4.797057
        min     8.630000    15.687000
        25%    11.895250    18.716750
        50%    14.356500    21.017500
        75%    16.200750    24.051500
        max    22.328000    35.255000
```

** central tendency ** * congruent mean = 14.051125 * incongruent mean = 22.015917 ** measure of variability ** * congruent std = 3.559358 * incongruent std = 4.797057

(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.

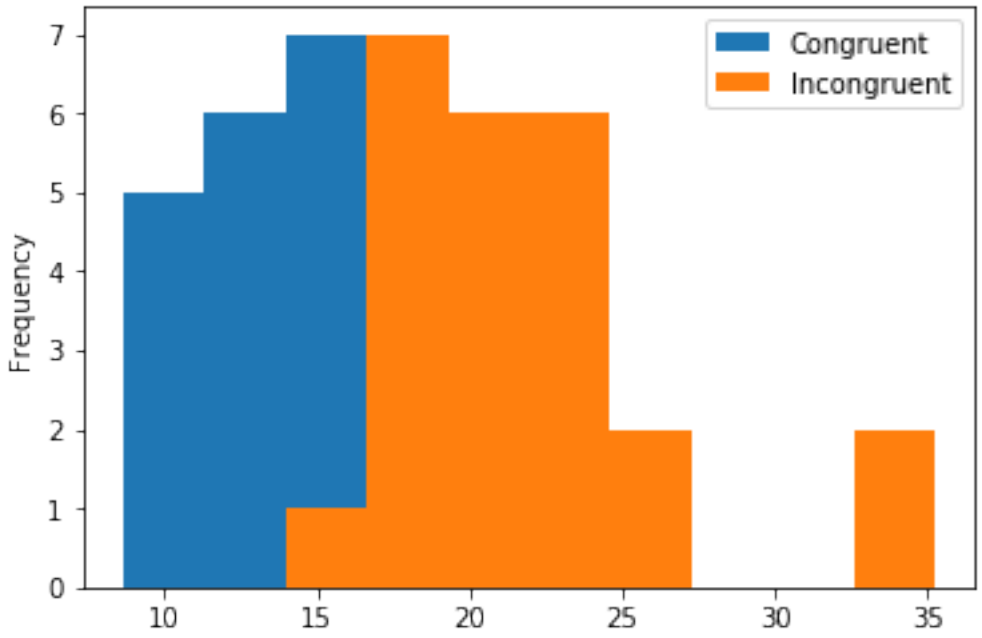
```
In [13]: import pandas as pd
         import matplotlib.pyplot as plt
         %matplotlib inline
         2
         ted = pd.read_csv('stroopdata.csv')
         plt.figure(); ted.plot();
```

```
<matplotlib.figure.Figure at 0x7f8631efd908>
```



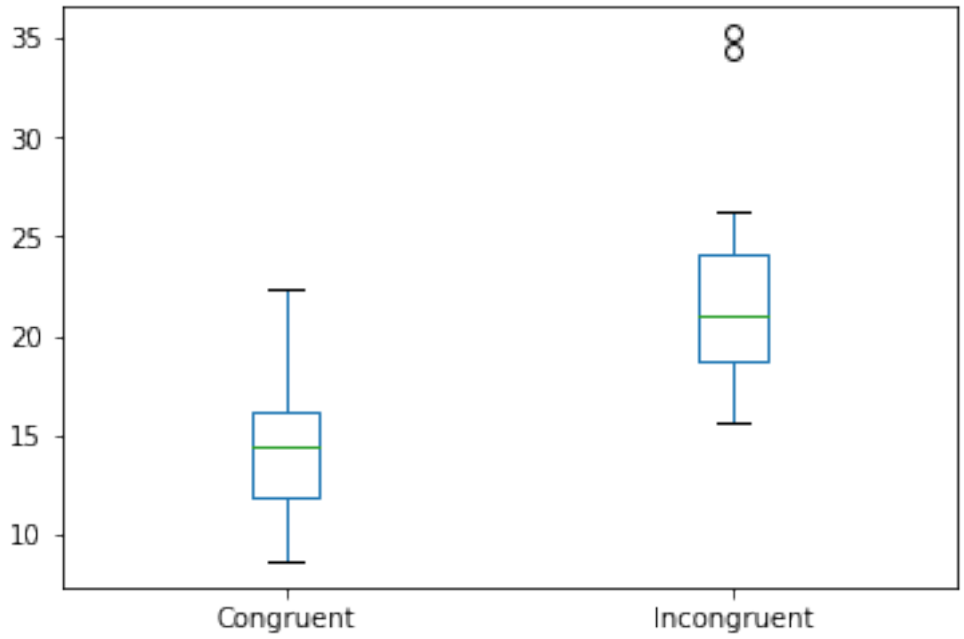

```
In [8]: ted.plot(kind='hist')
```

```
Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0x7f8634424d68>
```




```
In [9]: ted.plot(kind='box')
```

```
Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0x7f8631fa2be0>
```



The histogram shows that the average time/frequency of naming colours within the Incongruent group is higher. Within the box plot we can see that the incongruent group has two outliers around 35. In all, the box plot shows that time taken to name the colour for congruent words is between 8 to 23 and the distribution of time taken to name the colour for incongruent words are between 16 to 36.

(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.

Since the test is a two tailed t- test, its best to find the critical value of t at 95% confidance level. Sample Size(n) = 24 Degree of freedom(ted) = n - 1 = 23

```
In [12]: from scipy import stats
         degreesOfFreedom = 23
         stats.t.ppf(0.975, degreesOfFreedom)

         # Calculates the T-test on two related samples of scores, a and b.
         stats.ttest_rel(ted['Congruent'], ted['Incongruent'])
```

```
Out[12]: Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08)
```

At p value under 0.05, the time taken to name the colours is significatly different between the congruent & incongruent tasks thus it is safe to reject our null hypothesis. It is pretty difficult to name colours at the same speed when the word's meaning and colours match vs when they do not match. The result matches my expectations.

(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!

The brain uses images to associate the shape of a word and its colour. When there is a mismatch, additional time is necessary for the prefrontal cortex to process the information and decide on its meaning.

Explaining this further, two theories can be leveraged

Selective attention theory: This theory stipulates that naming the actual colour of words requires much more attention than simply reading the text that forms the words.

Speed of processing theory: This theory states that people can read words much faster than they can name colours. The speed at which we read makes it much more difficult to name the colour of the word after we have read the word.

Reference

https://www.verywellmind.com/what-is-the-stroop-effect-2795832
https://www.khanacademy.org/math/statistics-probability/significance-tests-one-sample/more-significance-testing-videos/v/one-tailed-and-two-tailed-tests
https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-the-differences-between-one-tailed-and-two-tailed-tests/ https://en.wikipedia.org/wiki/Stroop_effect
https://en.wikipedia.org/wiki/One-_and_two-tailed_tests