

Test a Perceptual Phenomenon

January 10, 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 is the congruency of the colour and text. For the dependent variable, Duration (time) to recognise the colours or the texts

- (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.

We want to prove the Stroop effect, ie, reject null hypothesis. The null hypothesis is that there is no difference in time for people when reading congruent words vs incongruent words. The alternate hypothesis is there is a difference between in time for people when reading congruent words vs incongruent words.

U_c = Population time for congruent words U_i = Population time for incongruent words

$H_0: U_c = U_i$ $H_a: U_c \neq U_i$

For this, we need to conduct a Paired t test. we do a paired test because we need to see the difference of each individual person congruent vs incongruent times. The T test is used to determine if there's a significant difference between the means of two groups of data.

- (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 [2]: import pandas as pd
import numpy as np
import scipy.stats as stats
```

```
In [3]: data = pd.read_csv('stroopdata.csv')
```

```
In [4]: data.describe()
```

```
Out [4]:
```

	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 [5]: data.median()
```

```
Out [5]: Congruent      14.3565
Incongruent    21.0175
dtype: float64
```

```
In [6]: Con = data['Congruent'].values.tolist()
In      = data['Incongruent'].values.tolist()
Con2 = np.round(Con)
In2 = np.round(In)
```

```
In [7]: print("Congruent:")
stats.mode(Con2)
```

Congruent:

```
Out [7]: ModeResult(mode=array([ 12.]), count=array([4]))
```

```
In [8]: print("Incongruent:")
stats.mode(In2)
```

Incongruent:

```
Out [8]: ModeResult(mode=array([ 21.]), count=array([4]))
```

--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.

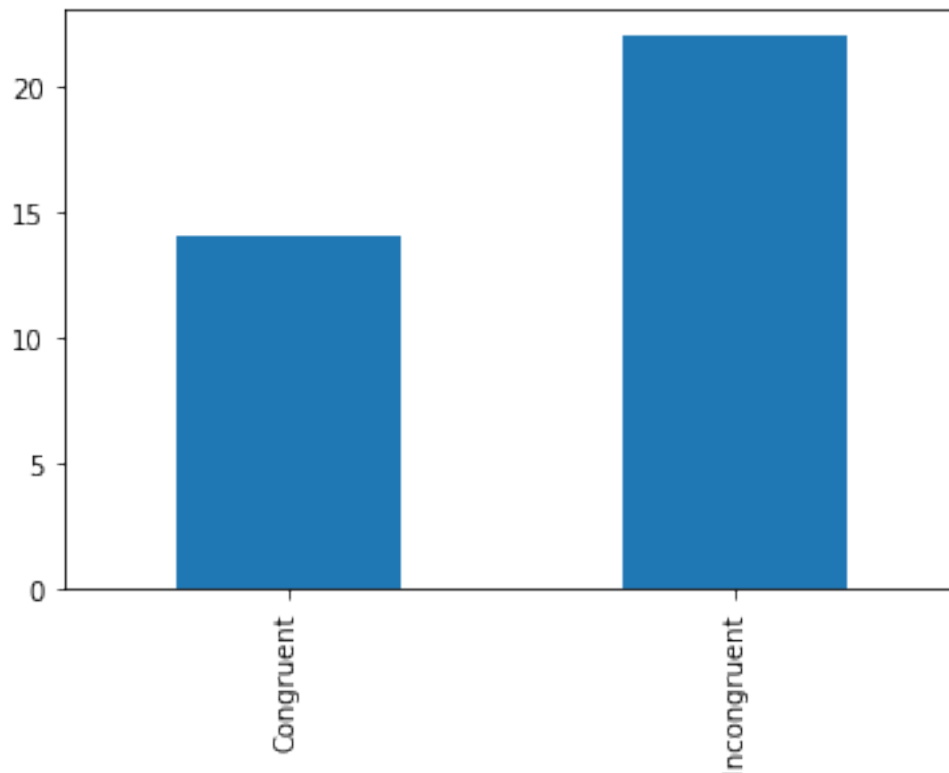
```
In [9]: data.head()
```

```
Out [9]:
```

	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 [11]: data.mean().plot('bar')
```

```
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x7ff32c889208>
```



Just based on the mean, it is clear that the time it took for people to say color words that are Incongruent takes longer/.

- (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 [10]: stats.ttest_rel(Con,In)
```

```
Out[10]: Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08)
```

According to the result, the p value is very small so even with a confidence level of %1 we would reject the null hypothesis. Also the T value is not close to zero which means there will difference when congruent color words vs not congruent words are read. These results are what I expected because of the stroop effect

- (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!