

Test a Perceptual Phenomenon

October 14, 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, 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?

An independent variable is the variable that is changed or controlled in an experiment to test the effects on the dependent variable. A dependent variable is the variable being tested and measured in an experiment. the 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.

H0: $\mu_1 = \mu_2$: Population mean of congruent and Incongruent are same. H1 : $\mu_1 \neq \mu_2$: The population mean of Congruent and Incongruent data differs where

H0 = the null hypothesis, H1 = The alternative hypothesis. μ_1 = the mean of population Incongruent, and μ_2 = the mean of population Congruent.

the statistical test that we can work on for this project is paired t-test, because here we are Testing the difference before and after some treatment(which is the Incongruent) on the same individual which is relevant to the data and our hypothesis.

This test assumes -

1-The differences are of measurement variables.

2-Sampling (or allocation) is random and pairs of observations are independent

3-The distribution of the mean difference is normal.

Here in this task we didn't use z-test as it's used with testing a population proportion , and we didn't also use one sample t-test as it's used with testing a population mean . so , the paired t-test is the most suitable test here that we check or test the change for the same individual after some change in the experiment

- (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 matplotlib.pyplot as plt
```

```
% matplotlib inline
```

```
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 [3]: print('The Congruent mean time is {}'.format(df['Congruent'].mean()))
print('The Incongruent mean time is {}'.format(df['Incongruent'].mean()))
obs_diff = df['Incongruent'].mean() - df['Congruent'].mean()
print('The observatory time difference is {}'.format(obs_diff))
```

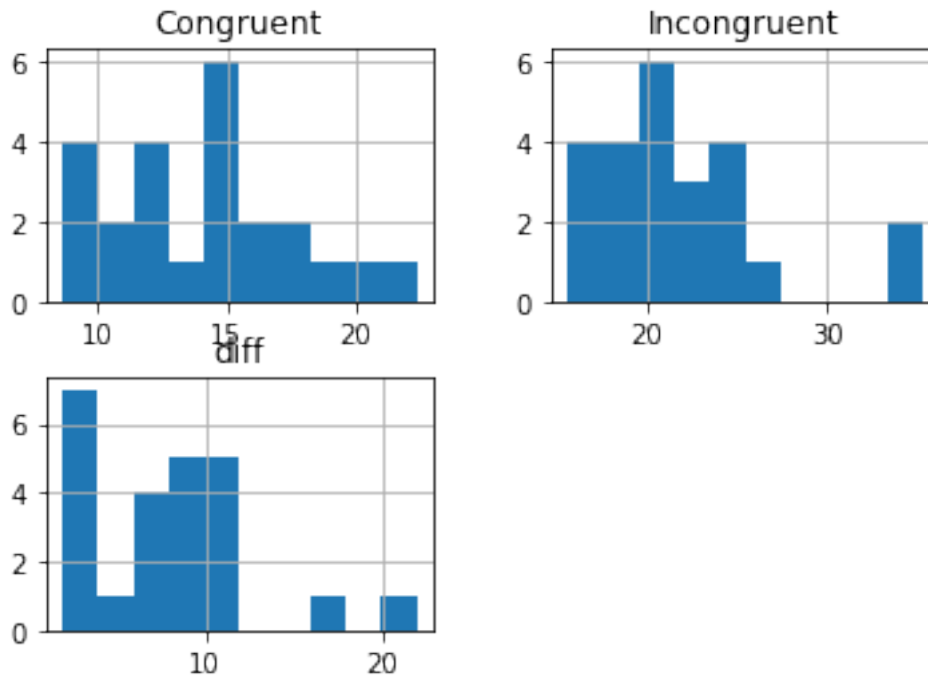
```
The Congruent mean time is 14.051124999999999
The Incongruent mean time is 22.015916666666666
The observatory time difference is 7.964791666666667
```

```
In [4]: print('The Congruent variance is {}'.format(df['Congruent'].var()))
print('The Incongruent variance is {}'.format(df['Incongruent'].var()))
```

```
The Congruent variance is 12.669029070652176
The Incongruent variance is 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 [5]: df['diff'] = df['Incongruent'] - df['Congruent']
df.hist();
```



The Congruent mean is approximately around 14 , while the Incongruent mean is approximately around 22 . The Congruent time is obvious less than that of Incongruent and the Incongruent has Outliers in the distribution. so, it has a large variance.

- (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 [6]: # Here we work on confidence level of 95 %
import math

mean_diff = df['diff'].mean()
std_diff = df['diff'].std()
std_error = std_diff / math.sqrt(24) # number of individuals is 24
T = mean_diff / std_error
T
```

```
Out[6]: 8.020706944109957
```

Looking in the table for $T = 8.020706944109957$ and $df = 23$ -----> $p\text{-value} < 0.0005$ As the p -value is smaller than the type 1 error threshold which is 0.05 . so, we reject the null hypothesis which indicates that The mean Congruent time is less than that of Incongruent

- (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!