Test a Perceptual Phenomenon

Stroop Effect Experiment

Stroop effect is the delay in the response time of a task. In the experiment the participants are given two list of words.

1. In the first list the words are in harmony with the colors it suggests, so this test is called the *congruent test*. The response time in reading the list of words is recorded
2. In the second list the words are not in harmony with the name of the color. The participants next are asked to read the second list which is called the *incongruent test*. The response is again recorded for this test

Reference: https://en.wikipedia.org/wiki/Stroop_effect

Stroop Effect Experiment data

In [35]:

```python
#importing the data

import numpy as np
import pandas as pd

#importing the data as a Dataframe "df_stroop" from the csv file
df_stroop = pd.read_csv('stroopdata.csv')
df_stroop
#https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view
```

Out[35]:

	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

	Congruent	Incongruent
21	14.233	17.960
22	19.710	22.058
23	16.004	21.157

1. Dependent and Independent variables in the Stroop Effect Experiement

The independent variable in the Stroop Effect Experiment:

1. Word and color match - in case of the Congruent the color and the text name are the same
2. Word and color do not match - In case of the Incongruent the color and the text name are different

The dependent variables are:

1. The congruent test is the pre-test which is performed on an individual and incongruent is the post test performed on the same individual, hence these variables are dependent because they are performed by the same individual. The data of pre and post tests are recorded for different individuals. Hence these are dependent samples and are paired.

2. Hypothesis and Statistical test for testing the Hypothesis

The set of hypotheses for this tasks

1. Null hypothesis

$$H_O: \mu_C = \mu_I$$

   Null hypothesis: If the congruent test and incongruent test was extended to everybody not just to the samples, then the mean of the congruent test μ_C for the population will be the same as the mean the incongruent test μ_I for everybody. hence $\mu_C - \mu_I = 0$
2. Alternative hypothesis

$$H_a: \mu_C \neq \mu_I$$

   The alternative hypothesis suggests that the means of the both responses are different i.e if the congruent and incongruent test was extended to everbody then the mean of both the congruent responses μ_C and incongruent responses μ_I will not be the same. i.e belonging to different population set

The Congruent and incongruent samples are paired. Statistical test to reject the H_O will be done using the **T- test** on two dependent samples. The population mean and standard deviation is not known and the **sample size is small (n = 24)** so t-distribution will be used on these dependent samples to test the null hypothesis H$_o$.

references: http://data-blog.udacity.com/posts/2016/10/latex-primer/
https://jupyter.brynmawr.edu/services/public/dblank/Jupyter%20Notebook%20Users%20Manual.ipynb#4.9-LaTeX-Math

3. Descriptive statistics regarding this dataset df_stroop

In [36]:

```python
import math
#df_stroop is a DataFrame which stores the data in columns 'congruent' and 'Incongruent'
#barxc mean of Congruent responses
#barxi mean of Incongruent responses
#sample size n = 24
n = 24
barxc = df_stroop['Congruent'].mean()
barxi = df_stroop['Incongruent'].mean()
print("Congruent_mean,incongruent_mean =",barxc,barxi)

#point_e is a variable which stores point estimate value
point_e =  barxc - barxi
print("point estimate:",point_e)

#Adding a new column 'difference' which stores the value of the difference of Congruent and respec
tive Incongruent values
df_stroop['difference'] = df_stroop['Congruent'] - df_stroop['Incongruent']
```

```
#MUd stores the mean of the difference from the column 'difference'
MUd = df_stroop['difference'].mean()
print("Md:",MUd)

#'difference_sq' stores the squares of the difference of (difference of congruent and incongruent
values) - (mean of the differences)
df_stroop['difference_sq'] = (df_stroop['difference'] - (-7.964791666666667))**2

#SE stores the standard error which is found using Bassel's correction
SE = math.sqrt(df_stroop['difference_sq'].sum()/(n-1))
print("SD:",SE)
```

```
Congruent_mean,incongruent_mean = 14.051125000000004 22.01591666666667
point estimate: -7.964791666666665
Md: -7.964791666666667
SD: 4.864826910359056
```

Means
Calculating mean of the congruent sample and the mean of incongruent sample(rounding to 2 decimal places)

$$\bar{x}_C = 14.05, \bar{x}_I = 22.016$$

$\bar{x}_C$ is the mean of the sample of congruent responses and $\bar{x}_I$ is the mean of the sample of Incongruent responses

Point Estimate

$$PointEstimate = \mu_C - \mu_I = 14.05 - 22.015 = -7.965$$

Since we have only one sample of size 24 from the population of congruent $\mu_C = \bar{x}_C$ Since we have only one sample of size 24 from the population of Ingongruent $\mu_I = \bar{x}_I$ we will find the standard error of the difference so that we know how it compares with other differences.

Standard deviationn of difference
Since we have a small sample size of 24 the standard deviation of the difference of the congruent and respective Incongruent values will be Standard diaviation

$$S_D = \sqrt{\frac{\sum(x_D - \bar{x})^2}{(n-1)}}$$

x_D = differences of congruent and incongruent values
$\bar{x}$ = mean of the differences of congruent and incongruent values
n = sample size = 24
since we have just a sample set we can find the S using Bassel's correction to find the standard deviation of the population

$$S_D = 4.864$$

4. Visualizations that show the distribution of the sample data.

t critical value

A significance level of 0.05 will be used to check for the H_o. Since we are not trying not find if $\mu_C > \mu_I$ or $\mu_C < \mu_I$ i.e we are trying for testing the relationship of the hypothesis in both the direction. Two tail test will be done for n = 24 and degrees of freedom = 23. The critical regions on both the sides will be $\alpha = 0.025$ on each tail.

There are **23 degrees of freedom**. For $\alpha = 0.025$ and 23 degrees of freedom the critical values are -2.069,2.069

critical values

5. Statistical test and the results.

```
#calculating t distribution for paired
#point_e stores the point estimate i.e the difference between µc - µi
#SE is the standard deviation error using the Bessel's correction
#n is the sample size
#the variables are declared in third section
t = point_e/(SE/math.sqrt(n))
print(t)
```

```
-8.020706944109957
```

t distribution for paired samples

$$t = \dfrac{\mu_C - \mu_I}{\dfrac{S}{\sqrt{n}}}$$

$$t = -8.02$$

The t statistic is the part of the left critical region. The H_o **is rejected** stating the alternative hypotheis is true

$$H_a : \mu_C \neq \mu_I$$

which states that the participants who took both the congruent test and the incongruent test did much better in response to congruent test.

So we can say that the participants takes longer time in the incongurent test BLUE RED than the congruent test BLUE RED .

The confidence interval for mean population interval

$$CI = M_D \pm t_critical \times \dfrac{S_D}{\sqrt{n}}$$

CI = confidence interval for $\mu_C - \mu_I$. That is confidence interval of the difference mean values of congruent and incongruent. M_D = $\mu_C - \mu_I$
S_D = standard deviation of the differences = 4.864
$t_critical$ = 2.069 (as foiund from the t table when there are 23 degrees of freedom)
n = 24

```
CI_p = barxc - barxi + 2.069*(SE/math.sqrt(n))
CI_n = barxc - barxi - 2.069*(SE/math.sqrt(n))
CI_p,CI_n
```

```
(-5.91021542131028, -10.019367912023053)
```

If we round it off to integers The critical intervals = (-10,-6) Which tells that the participant is taking 10 seconds to 6 seconds less time in completing the congruent test than the incongruent test on an average.

references: https://en.wikipedia.org/wiki/Stroop_effect http://data-blog.udacity.com/posts/2016/10/latex-primer/
https://jupyter.brynmawr.edu/services/public/dblank/Jupyter%20Notebook%20Users%20Manual.ipynb#4.9-LaTeX-Math