## Test a Perceptual Phenomenon

February 18, 2018

## 0.0.1 Analyzing the Stroop Effect

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

The independent variable is the different participants who and the dependent variable is the time they used in the congruent and incongruent situation.

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

The null hypothesis and the alternative hypothesis should be:

$$H_0: \mu_c = \mu_i,$$
  
 $H_1: \mu_c \neq \mu_i,$ 

Or equally:

$$H_0: \mu_c - \mu_i = 0,$$
  
 $H_1: \mu_c - \mu_i \neq 0,$ 

where  $\mu_0$  is the mean population time in congruent situation and  $\mu_1$  is the mean population time of the incongruent situation.

I plan to perform the paired t-test on the data.

First of all, the sample size is relatively small so we cannot use normal distribution to perform the test, so we have to choose t test.

Second of all, becauese the dependent variable is related in the two groups, we have to use the "matched-pair" t-test rather than the usual two sample t-test.

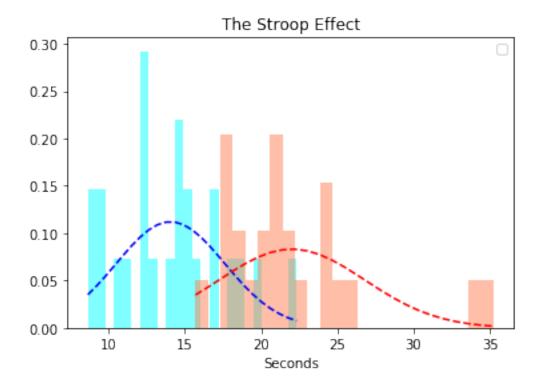
(3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability. The name of the data file is 'stroop-data.csv'.

```
In [42]: # import pandas
    import pandas as pd
    # read data and show the first 10 rows
    data="stroopdata.csv"
    stroop=pd.read_csv(data)
    print(stroop.head(10))
    # print the descriptive statistics
    print("mean is: ",stroop.mean(),"median is: ",stroop.median())
    print("standard deviation is: ",stroop.std())
```

```
Congruent Incongruent
0
      12.079
                   19.278
      16.791
                   18.741
1
2
       9.564
                   21.214
3
                   15.687
       8.630
4
      14.669
                   22.803
5
      12.238
                   20.878
6
      14.692
                   24.572
7
      8.987
                   17.394
                   20.762
8
       9.401
9
      14.480
                   26.282
mean is: Congruent
                          14.051125
               22.015917
Incongruent
dtype: float64 median is: Congruent
                                            14.3565
               21.0175
Incongruent
dtype: float64
standard deviation is:
                         Congruent
                                        3.559358
Incongruent
               4.797057
dtype: float64
```

(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 [31]: # import matplotlib
         import matplotlib.mlab as mlab
         import matplotlib.pyplot as plt
         # set all the parameters
         num_bins = len(before)
         mu1=before.mean()
         sigma1=before.std()
         mu2=after.mean()
         sigma2=after.std()
         # the histogram of the before data
         n, bins, patches = plt.hist(before, num_bins, normed=1, facecolor='aqua', alpha=0.5)
         # add a curve
         y = mlab.normpdf(bins, mu1, sigma1)
         plt.plot(bins, y, 'b--')
         # the histogram of the before data
         n, bins, patches = plt.hist(after, num_bins, normed=1, facecolor='coral', alpha=0.5)
         # add a curve
         y = mlab.normpdf(bins, mu2, sigma2)
         plt.plot(bins, y, 'r--')
         plt.xlabel('Seconds')
         plt.title("The Stroop Effect")
         plt.show()
```



From the plot it is obvious that the incongruent condition does have an effect on the participants since it took them more time to finish the experiment.

(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?

We reject the null hypothesis if the t statistics is less than -2.807 or greater than 2.807.

According to the result of the paired t-test, the t statistics is -8.02 and the p-value is much smaller than 0.01, which means we can reject the null hypothesis with a significance of 0.01.

Therefore, it does match my expectations.