Test a Perceptual Phenomenon

October 24, 2017

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 and submit in the next section.

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

The independent variable in the experiment in the variable is the conditions of words, which has two levels: a congruent words condition, and an incongruent words condition. The dependent variable in the experiment is the time it takes to name the ink colors in equally-sized lists.

(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 is that there is no significant difference in the population average amount of time to name the colors of words in congruent and incongruent conditions.

H0: u1=u2

The alternative hypothesis is that there is significant difference in the population average of the amount of time to name the colors of words in congruent and incongruent conditions.

Ha: u1!=u2

u1 is the population mean of the time it takes in congruent words condition, u2 is the population mean the time it takes in congruent words condition.

Test z and test t are applicable to comparing means. For this test, we have less than 30 samples and we don't know the population's standard deviation. We assume the population is normally distributed. The two samples are dependent since they are performed by the same people. Based on these assumptions, I choose to perform dependent T-test for paired two-samples. The alternative hypothesis is two sided so I choose a two-tailed 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 'stroopdata.csv'.

```
In [10]: from scipy import stats
         import numpy as np

         data=np.genfromtxt('stroopdata.csv',delimiter=',',skip_header=1)
         data_1=data[:,0]
```

```
data_2=data[:,1]
data_1_mean=np.mean(data_1)
data_1_sd=np.sqrt(np.var(data_1,ddof=1))
data_2_mean=np.mean(data_2)
data_2_sd=np.sqrt(np.var(data_2,ddof=1))
str1='In congruent condition,mean= %6.4f standard deviation=%6.4f'
str2='In incongruent condition,mean= %6.4f standard deviation=%6.4f'
print(str1%(data_1_mean,data_1_sd))
print(str2%(data_2_mean,data_2_sd))
```

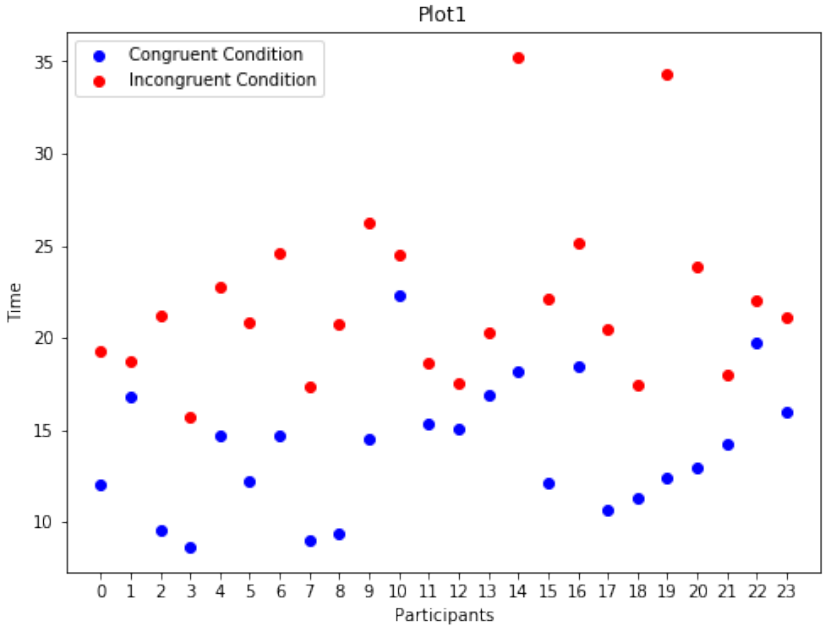
```
In congruent condition,mean= 14.0511 standard deviation=3.5594
In incongruent condition,mean= 22.0159 standard deviation=4.7971
```

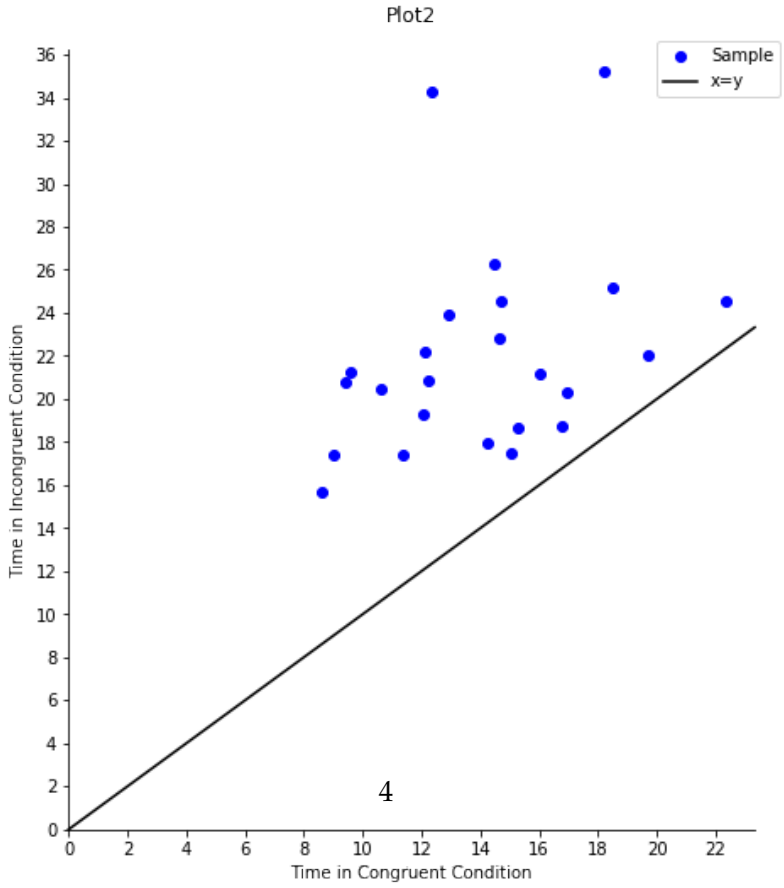
The measure of central tendency is mean and the measure of variability is standard deviation. In congruent condition,mean= 14.0511 standard deviation=3.5594 In incongruent condition,mean= 22.0159 standard deviation=4.7971

(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 [12]: from scipy import stats
         import numpy as np
         import matplotlib.pyplot as plt

         data=np.genfromtxt('stroopdata.csv',delimiter=',',skip_header=1)
         data_1=data[:,0]
         data_2=data[:,1]

         fig=plt.figure(figsize=(8,20))
         sub1=fig.add_subplot(311)
         sub1.plot(range(len(data_1)),data_1,'ob',label='Congruent Condition')
         sub1.plot(range(len(data_2)),data_2,'or',label='Incongruent Condition')
         sub1.legend()
         sub1.set_title('Plot1')
         sub1.set_xticks(np.arange(len(data_1)))
         sub1.set_xlabel('Participants')
         sub1.set_ylabel('Time')

         a= np.arange(0,max(data_1)+1,2)
         b=np.arange(0,max(data_2)+1,2)
         sub2=fig.add_subplot(212)
         sub2.plot(data_1,data_2,'ob',label='Sample')
         sub2.plot((0,max(data_1)+1),(0,max(data_1)+1),'k-',label='x=y')
         sub2.set_title('Plot2')
         sub2.legend(loc='best')
         sub2.set_xlabel('Time in Congruent Condition')
         sub2.set_ylabel('Time in Incongruent Condition')
         sub2.set_xticks(a)
```

```
sub2.set_yticks(b)
sub2.spines['left'].set_position(('data',0))
sub2.spines['bottom'].set_position(('data',0))
sub2.spines['bottom'].set_bounds(0, max(data_1)+1)
sub2.spines['left'].set_bounds(0, max(data_2)+1)
sub2.spines['right'].set_color('none')
sub2.spines['top'].set_color('none')

plt.show()
```


Plot1
Congruent Condition
Incongruent Condition
Time
Participants


Plot2
Sample
x=y
Time in Incongruent Condition
Time in Congruent Condition

In plot 1, we can see that the time taken in incongruent words condition is generally more than the time taken in congruent words condition. In plot 2, we can see that all the observations are above the line(x=y), which means that for each oberservation, the time taken in incongruent words condition is generally more than the time taken in congruent words condition.

(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?

```
In [13]: # Perform the statistical test here
         from scipy import stats
         import numpy as np

         data=np.genfromtxt('stroopdata.csv',delimiter=',',skip_header=1)
         data_1=data[:,0]
         data_2=data[:,1]
         t,p=stats.ttest_rel(data_1,data_2)
         print(t,p)
```

```
-8.02070694411 4.10300058571e-08
```

The t statistic value is -8.0207, p value is 4.10300058571e-08. Confidencen level is 95%, significance level is 0.05. P value is smaller than significance level. So I reject the null hypothesis. The conclusion is that the time it takes to name the ink colors in equally-sized lists in incongruent words condition is different(longer) than the time it takes in congruent condition. The results match up with my expections because I expect that the incongruent words condition will disturb the participants and slow down their reaction.