Test a Perceptual Phenomenon

November 2, 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?

Independent variable : congruent word condition, incongruent word condition
Dependent variable : 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.

Null Hypothesis ($H_0 : \mu_1 = \mu_2$**)** * The average population of time it takes to name the ink colors in equally-sized lists for two groups(congruent and incongruent) are equal.

Alternative Hypothesis ($H_0 : \mu_1 \neq \mu_2$**)** * H_1 : The average population of time it takes to name the ink colors in equally-sized lists for two groups(congruent and incongruent) are not equal.

Statistical Test : dependent t-test (two tailed) Justification : * We need to compare the means of two related groups to determine the statistically significant difference between two means. * We are assuming normal distribution * There are less than 30 samples * it has to be two tailed because our hypothesis is testing equality of two means.

(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 [1]: # code adapted from http://pandas.pydata.org/pandas-docs/stable
        import pandas as pd

        # read csv file and save to df
        df = pd.read_csv('stroopdata.csv')

        # print csv data
        print(df)

        # Describe shows a quick statistic summary of your data
        df.describe()
```

```
    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
21     14.233       17.960
22     19.710       22.058
23     16.004       21.157
```

Out[1]:

	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

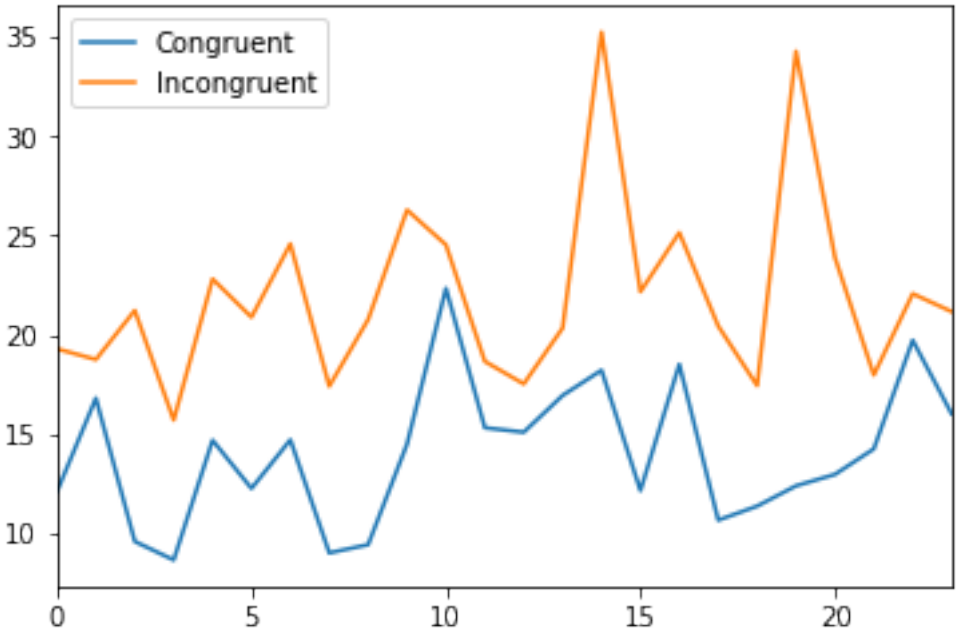
** central tendency ** * congruent mean = 14.0511 * incongruent mean = 22.0159
** measure of variability ** * congruent std = 3.5594 * incongruent std = 4.7971

(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 [2]: # code adapted from http://pandas.pydata.org/pandas-docs/stable/visualization.html#visua
        import pandas as pd
        import matplotlib.pyplot as plt
        %matplotlib inline
```

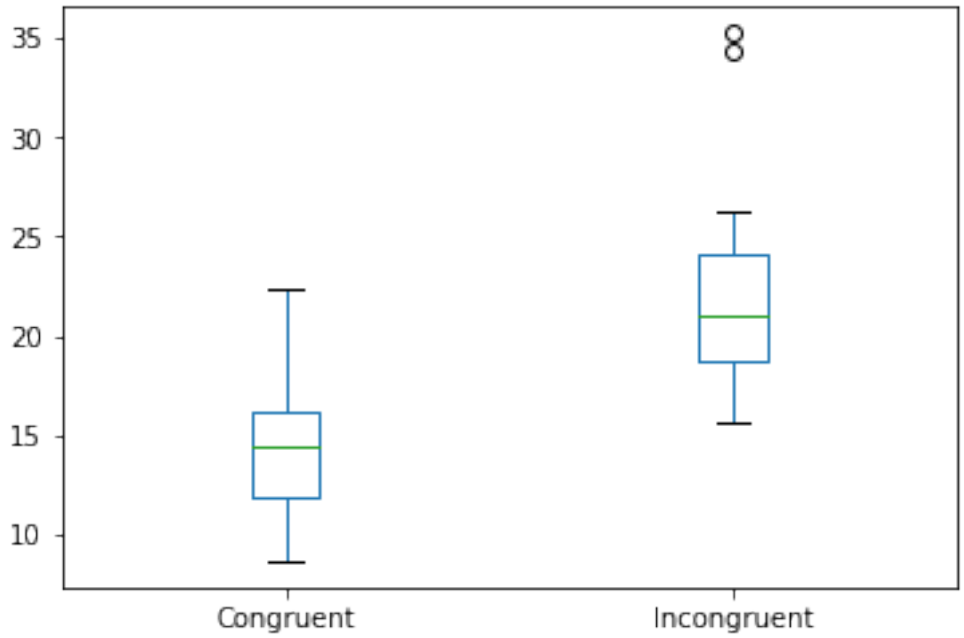
```
        df = pd.read_csv('stroopdata.csv')
        plt.figure(); df.plot();
```

```
<matplotlib.figure.Figure at 0x7f1724ae6be0>
```



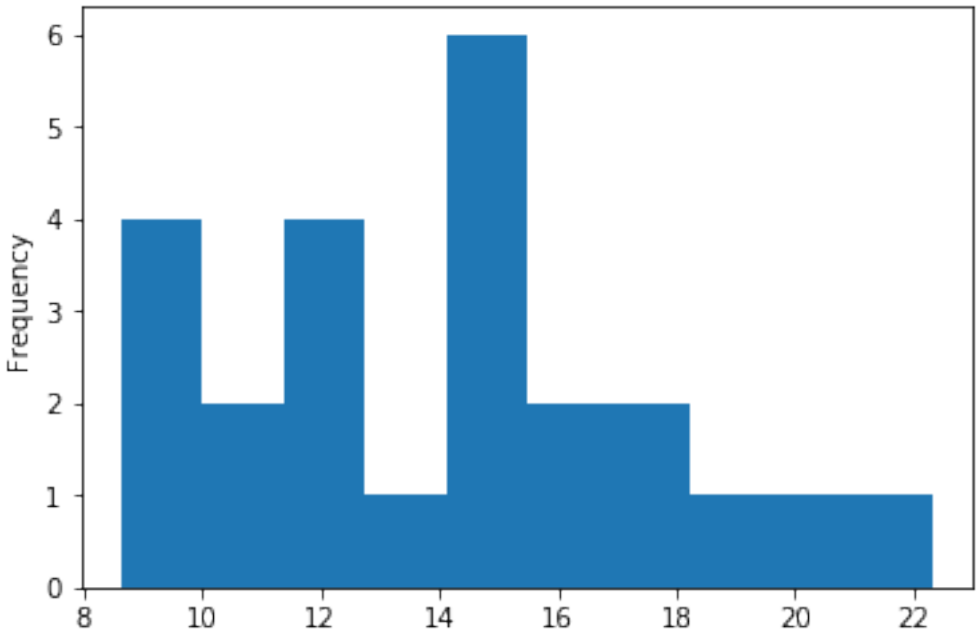

```
In [4]: df.plot(kind='box')
```

```
Out[4]: <matplotlib.axes._subplots.AxesSubplot at 0x7ff965b0fdd8>
```



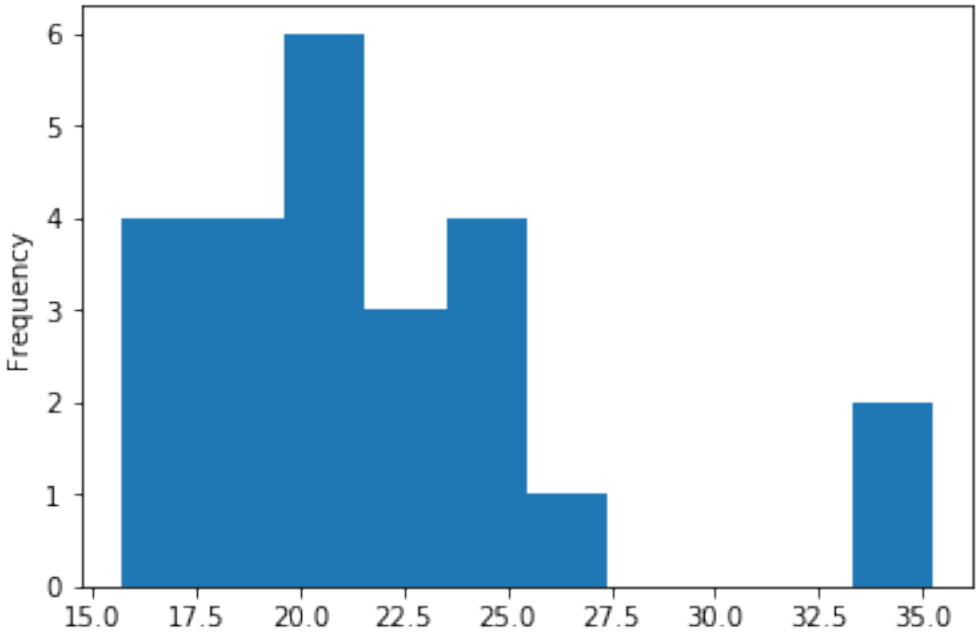

```
In [11]: df['Congruent'].plot(kind="hist")
```

```
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x7f17013cf208>
```



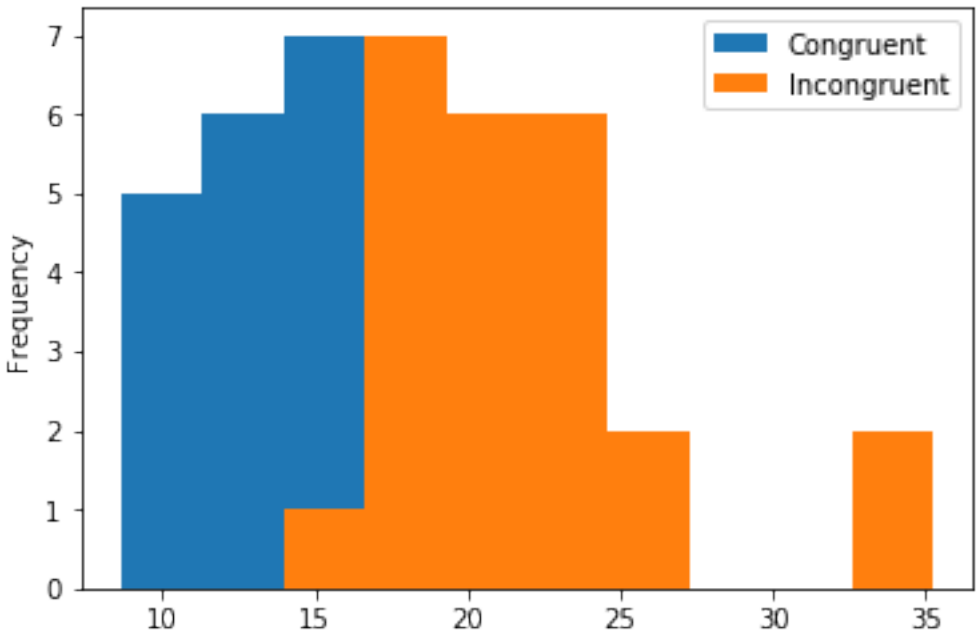
```
In [12]: df['Incongruent'].plot(kind="hist")
```

```
Out[12]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1700fb3780>
```




```
In [21]: df.plot(kind="hist")
```

```
Out[21]: <matplotlib.axes._subplots.AxesSubplot at 0x7f16f2f59630>
```



By looking at the boxplot, we can see that the average completion time of Incongruent group is higher. The box plot also shows that the incongruent group has two outliers. The distribution for both congruent and incongruent group looks normal.

(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 [31]: # Perform the statistical test here
         # https://docs.scipy.org/doc/scipy/reference/tutorial/stats.html
         from scipy import stats
         degreesOfFreedom = 23
         stats.t.ppf(0.975, degreesOfFreedom)

         # Calculates the T-test on TWO RELATED samples of scores, a and b.
         stats.ttest_rel(df['Congruent'], df['Incongruent'])

Out[31]: Ttest_relResult(statistic=-8.020706944109957, pvalue=4.1030005857111781e-08)
```

p value is under 0.05 therefore, we reject our null hypothesis. Thus, we can conclude that incongruent group and congruent group has different average population time to complete tasks.

0.1 Reference

- Latex Symbols
- Dependent T-Test
- Determining Sample Size
- About the null and alternative hypothesis
- Which chart or graph is right for you?
- pandas visualization
- scipy stats