Stroop Effect

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the color of the ink in which the word is printed. The task has two conditions: a congruent words condition, and an incongruent words condition. In the congruent words condition, the words being displayed are color words whose names match the colors in which they are printed: for example RED, BLUE. In the incongruent words condition, the words displayed are color words whose names do not match the colors in which they are printed: for example PURPLE, ORANGE. In each case, we measure the time it takes to name the ink colors in equally -sized lists. Each participant will go through and record a time from each condition.

My score for stroop experiment:

Congruent	Incongruent
11.963	21.44

Questions For Investigation

1. What is our independent variable? What is our dependent variable?

The name of the color with color(congruent or incongruent word) is our independent variable. The time(in seconds) taken by participants to recognise the color of the word is dependent variable.

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

For this task my intuition is that the time taken to say the color in incongruent word, will be greater than that for congruent word. With that theory the null hypothesis H_o will be that incongruent word score population mean μ_i is less than or equal to congruent work score population mean μ_c and alternative hypothesis H_a will be that incongruent word score population mean μ_i is greater than congruent work score population mean μ_c.

$$H_o : \mu_i \leq \mu_c$$

and the alternative hypothesis will be:

$$H_a : \mu_i > \mu_c$$

Since same subject performs the test on both type of words, it is a two conditions dependent samples test. We can not perform a z-test here as we do not have the population mean and standard deviation. The dependent t-test for paired samples will be a better choice of statistical test here.

3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

The congruent sample has a mean of 14.05 sec and standard deviation of 3.56 sec while, the mean of incongruent sample is 22.02 sec and the standard deviation is 4.80 sec. The fastest answer for congruent word is 8.63 sec and that for incogruent word is 15.70 sec. Whereas, The slowest answer for congruent word is 22.33 sec and that for incongruent word is 35.26 sec. The last values are farther mare than 2 std away from the means hence can be outliers. In the presence of outliers median is a better representative of central tendency. The median of congruent words is 14.36 and that for incongruent words is 21.02.

```
In [122]:  import numpy as np
           import pandas as pd
           import matplotlib.pyplot as plt


           stroop = pd.read_csv('stroopdata.csv')
           stroop.describe()
           np.median(stroop['Incongruent'])
           np.median(stroop['Congruent'])
```

```
Out[122]:  14.3565
```

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 [127]:  stroop.hist()
           plt.show()
```
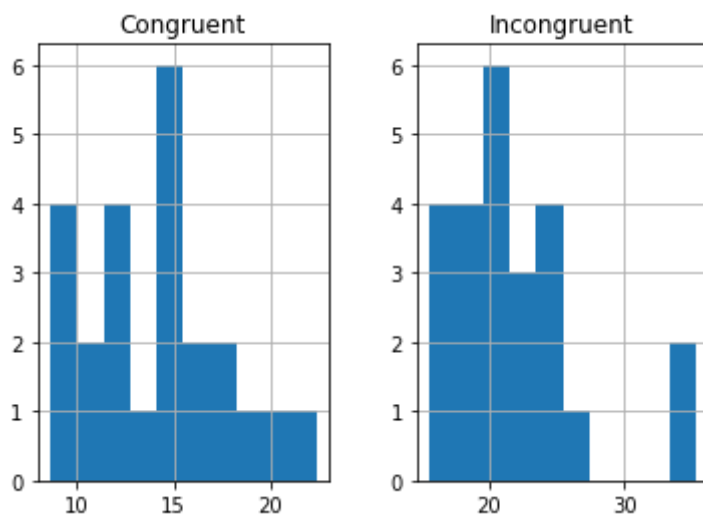

We observe some outliers in the histogram to detect those we will plot boxplot. If we ignore few far lying outliers the histogram shows a normal distribution. The mode for congruent words is at 13 sec and mode for incongruent words is at 21 sec.

```
In [126]: p1 = stroop.boxplot(vert=True,widths = 0.4,patch_artist=True)
          plt.title('Boxplot for congruent and incongruent time data')
          plt.show()
```

Boxplot for congruent and incongruent time data
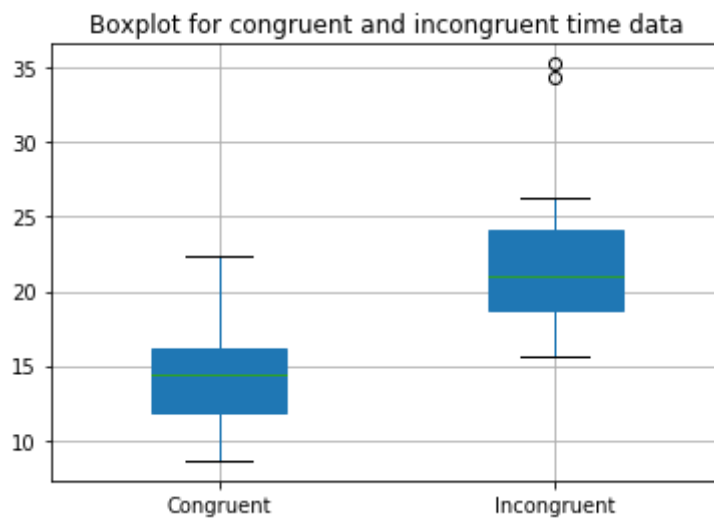
The reason for outliers may be that these participants take longer than usual time to recognise colors. The other reason may be they got interrupted while taking the test.

5. Now, perform the statistical test and report your results. What is your 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?

Size of the paired samples

```
n= 24
```

The degrees of freedom

```
df = n-1 = 23
```

Mean of the two samples:

```
$\mu_{i}$=22.015917 and,
$\mu_{c}$=14.051125
```

point estimate

```
$/mu$ = $\mu_{i}$ - $\mu_{c}$
      = 7.96
```

The sample standard deviation of differences S

```
S= 4.86
```

t$_{critical}$ value

```
We calculate the t$_{critical}$ value for confidence level of 95% i.e. $\alp
ha$ = 0.05. Therefore, for positive
direction one tailed t-test for $\alpha$ = 0.05 with df =23 :
$t_{crit}$= 1.714
```

t-statistic

```
t-statistic can be calculated as-
```

$$t = \frac{\mu}{S/\sqrt{n}}$$

```
t = 8.04
```

p-value

```
    p value corresponding to t = 8.04 and df = 23 is less than 0.0005.
```

```
    Since p-value is less than 0.05 we reject the null hypothesis with 99.9 % co
    nfidence.
```

Cohen's d

```
    Cohen's d = $\frac{(\mu - 0)} {S}$
             = 1.64
    This tells us that mean of congruent and incongruent words are 1.64 standard
     deviation apart.
```

correlation measures

```
    $r^2$=73.75%
```

$$r^2 = \frac{t^2}{(t^2 + df)}$$

$r^2 = 0.7375$

That implies incongruency affects the sample mean by 73.75%.

Conclusion

In this t-test we made an assumption that if a point estimate lies with in critical range of t i.e. with probibility less than 0.05, it is statistically significant. We reject the null hypothesis with a confidence level of 99.9%. In other words, with this statistical test it can be concluded that any participant would take more time to recognize the color of an incongruent word than that of congruent word. This result is in agreement with my expectations.

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!

I think that the reason for this effect lies how our brain works. For most of the people the ability to read a world is faster than the ability to recognise color. On literature survey, I found following theory explananing for Stroop effect[1]:

The words themselves have a strong influence over your ability to say the color. The interference between the different information (what the words say and the color of the words) your brain receives causes a problem. There are two theories that may explain the Stroop effect:

Speed of Processing Theory: the interference occurs because words are read faster than colors are named. Selective Attention Theory: the interference occurs because naming colors requires more attention than reading words. I think that this puzzle would be easier for a very young child than for older children or adults. Try this out on some small kids who know their colors, but cannot yet read! I would imagine that the children would not get confused by this puzzle because the words would not have any meaning to them.

To find the similar experiment we have to think about two or more of human mental skills conflicting with each other. One such example could be-

The participants are shown the name of the shape inside a shape and they have to say the shape. In first group, the name of the shape matches with the shape it is enclosed in and in other group it doesn't.

 Reference

[1] https://faculty.washington.edu/chudler/words.html (https://faculty.washington.edu/chudler/words.html)

In [ ]: