Project: Test a Perceptual Phenomenon

Over View from the project discretion:

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 *colour of the ink i*n which the word is printed. The task has two conditions: a congruent words condition, and an in-congruent words condition. In the *congruent words* condition, the words being displayed are colour words whose names match the colours in which they are printed: for example RED, BLUE. In the *in-congruent words* condition, the words displayed are colour words whose names do not match the colours in which they are printed: for example PURPLE, ORANGE. In each case, we measure the time it takes to name the ink colours in equally-sized lists. Each participant will go through and record a time from each condition.

Below is a dataset which contains results from a number of participants in the task

Congruent	In-congruent
8.63	15.687
8.987	17.394
9.401	20.762
9.564	21.214
10.639	20.429
11.344	17.425
12.079	19.278
12.13	22.158
12.238	20.878
12.369	34.288
12.944	23.894
14.233	17.96
14.48	26.282
14.669	22.803
14.692	24.572
15.073	17.51
15.298	18.644
16.004	21.157
16.791	18.741
16.929	20.33
18.2	35.255
18.495	25.139
19.71	22.058
22.328	24.524

Source :https://docs.google.com/document/d/1-OkpZLjG_kX9J6LIQ5IltsqMzVWjh36QpnP2RYpVdPU/pub?embedded=True

Here we would analyse if the time taken to name the colour of the ink in which the colour name is printed has any effect when the colour name printed matches with ink of colour and when the colour name printed doe not match with ink of colour

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

Independent variable: These are variables that are changed to study the effect on dependent variable. In this case independent variable is **Word Condition** (congruent words condition, and an in-congruent words condition)

Dependent variable: These are variables on which we study the effect when an independent variable is changed. Here dependent variable is the **time taken to tell the colour in which the word is written.**

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

Hypothesis testing and statistical test is done in this experiment to find out if the change in the time taken for the Congruent condition test(Colour name matches with the colour of the ink) and In-Congruent condition test(Colour name does not match with the colour of the ink) is by chance.

In this test we will using a sample of data set(scores of Congruent and In-Congruent test) of randomly participants. Using this sample data, we will infer the behaviour of population.

- Hypotheses test are
  - According to NULL Hypotheses, there is no significant difference between the time taken in Congruent condition(Colour name matches with the colour of the ink) and In-Congruent condition(Colour name does not match with the colour of the ink). That is the difference between the mean of time taken in Congruent condition and In-Congruent condition will be equal to zero

    H0: $\mu_{congruent} - \mu_{incongruent} = 0$

  - According to Alternative Hypotheses states that, there is a significant difference between the time taken in Congruent condition(Colour name matches with the colour of the ink) and In-Congruent condition(Colour name does not match with the colour of the ink).That is the difference between the mean of time taken in Congruent condition and In-Congruent condition will not be equal to zero.

    Ha: $\mu_{congruent} - \mu_{incongruent} != 0$

    - $\mu_{congruent}$ **-->** Mean of time taken in Congruent condition(Colour name matches with the colour of the ink) test by sample population.
    - $\mu_{incongruent}$ **-->** Mean of time taken in the In-Congruent condition(Colour name does not match with the colour of the ink) test by sample population.
    - **H0 -->** Null hypotheses
    - **Ha -->** Alternate hypotheses

- We can perform a **Dependent T test**. The reason for choosing the paired test is because the test are taken by same person in different condition. As the data are dependent, we rule out the use independent T-test. Also if we consider tests taken by 2 different individuals, that can add variance created by other factors.

- We are not considering Z-test as we are not trying to estimate, if the sample is different from the population mean(we don't have data for the population). Here our Sample size 24 and we are trying to calculate population standard deviation from the sample
- (s =-4.86). Moreover, as the data is dependent we go for paired t-test over Z-test.

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

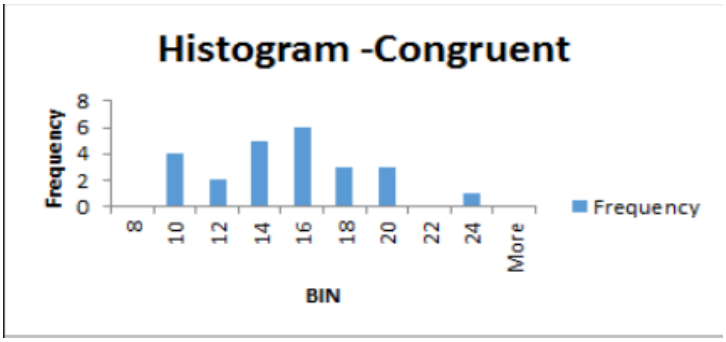
Below are the some of the descriptive statistics

	Congruent	Incongruent	Difference (Congruent – incongruent)
Mean	14.051125	22.015916667	-7.964791666667
Median	14.3565	21.0175	-7.6665
Range	13.698	8.837	19.969
Standard deviation of sample	3.5593579576	4.7970571225	4.8648269103591

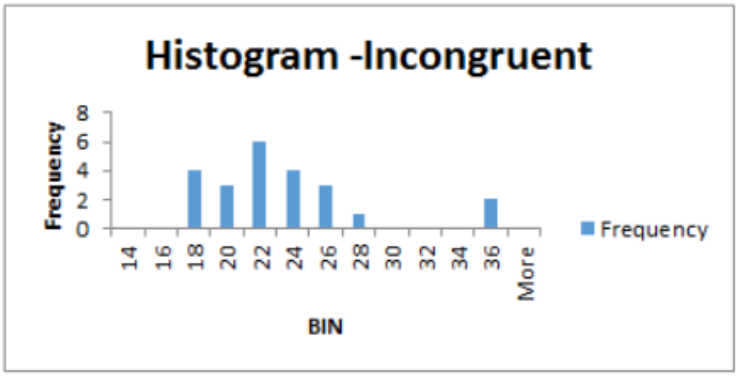
4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Histogram for Time taken by Congruent and In-congruent tests :

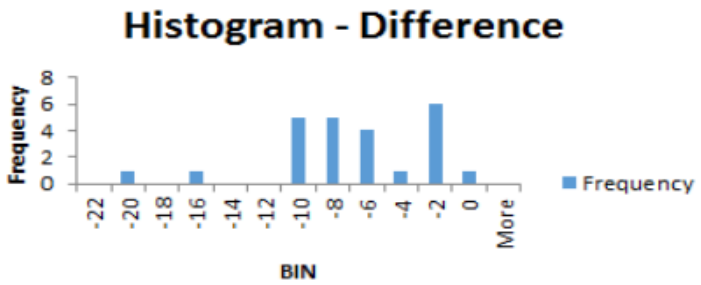
Congruent sample has a normal distribution with mode as 16.



In-congruent sample has a right skewed normal distribution. The skewing has occurred mainly because of 2 values between 34 and 36.



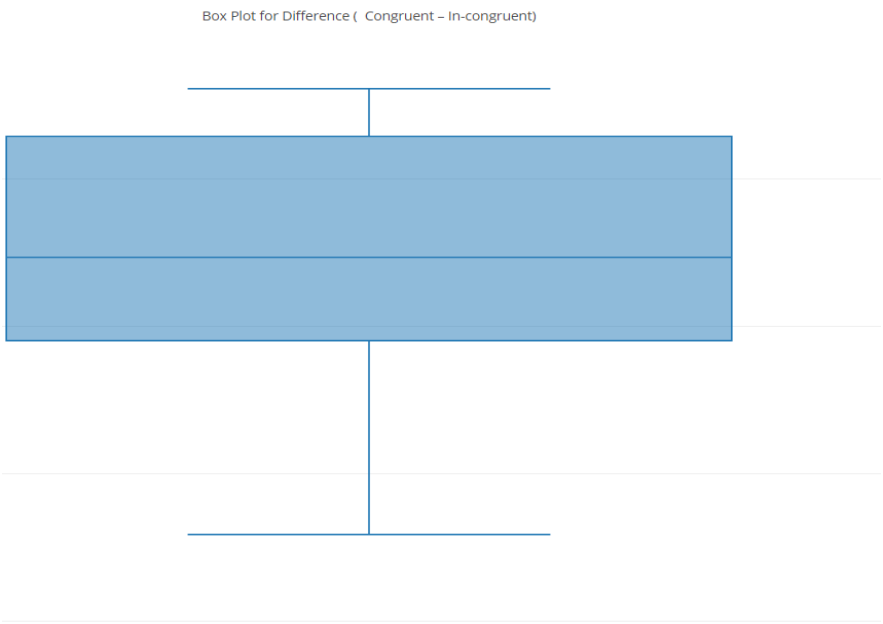
Histogram for Difference of time taken by both the tests (Congruent – In-congruent) is negatively skewed and has outliers.



Box plot for Difference (Congruent – In-congruent)
(Generated from ": https://plot.ly/create/histogram/)

There is one outlier in of samples, which will affect the mean and standard deviation.

Congruent	In-Congruent	Difference
12.369	34.288	**-21.919**



New mean and standard deviation removing outlier

	Congruent	In-Congruent	Difference
Mean	14.12426087	21.482347826087	-7.358086957
Std Dev	3.6208681809	4.11272073857657	3.9378375805

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?

Hypotheses test here would be as below with 95%($\alpha = 0.05$) confidence

H0: $\mu_{congruent} - \mu_{incongruent} = 0$

Ha: $\mu_{congruent} - \mu_{ncongruent}$ **!= 0**

- $\mu_{congruent}$ **-->** Mean of time taken in Congruent condition(Colour name matches with the colour of the ink) test by sample population.
- $\mu_{incongruent}$ **-->** Mean of time taken in the In-Congruent condition(Colour name does not match with the colour of the ink) test by sample population.
- **H0 -->** Null hypotheses
- **Ha -->** Alternate hypotheses

$\mu_D = -7.96$ (Difference between mean values of Congruent condition and In-Congruent condition)

S = 4.86 (Sample deviation)
n = 24 (Sample size)
df = 23 (degrees of freedom. That is one less than sample size)

$t_{critical}$ (2 tail at 95% confidence) = -2.069 /2.069 (T-critical value from t-table for 2 tail test at 95% confidence and df =23. Two tail test is done as the direction is unknown)

Standard Error of mean(SE) = .99 (This is standard deviation divided by square root of sample size)

$t_{statistic} = -8.02$ ($t_{statistic}$ is calculated by dividing Mean difference by standard error)

t(22) = -8.02, p <.00001, Two Tailed

$t_{statistic} < t_{critical}$ (-8.02< -2.069)
$p_{value} < 0.05$

The t-value is much less than the t-critical so we **reject NULL hypothesis**. There is a difference in the time taken to name the colour of the ink in case of Congruent and In-Congruent is not by chance.

6. 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 question

According to Wikipedia Brain needs to separate relevant information from irrelevant information in the task; thus, the focus is placed on the ink colour and not the word. This extra process will take more time. Hence we have higher value value for In-Congruent condition than Congruent condition.