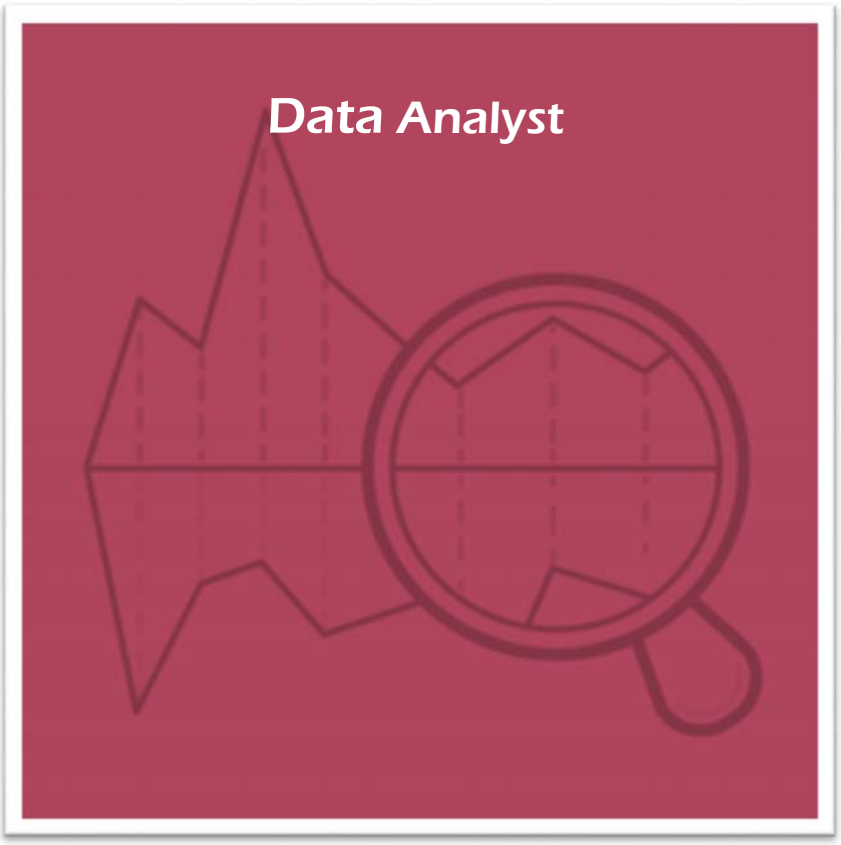


Data Analyst Nano Degree

P1 – TEST A PERCEPTUAL PHENOMENA

Harshit Trivedi | Data Analyst Nano Degree | January 3, 2016

Background Information

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.

Questions For Investigation

What is our independent variable? What is our dependent variable?

Here, our independent variable is **the congruent and incongruent list of colors**. More specifically, it could also be the no. of words in each list, if we want to determine the time taken to go through each list or each color in the list.

Our dependent variable, would therefore be **the time taken to go through the congruent and incongruent color list**, or each word in the list separately.

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

The first and most obvious assignment one can make is that the Stroop Effect exists, that is, there is a significant difference in the time taken to read through the congruent and incongruent list of colors.

More specifically, we can say that

Null Hypothesis (H_0): There is no significant difference between the time taken to read through the congruent and incongruent list of colors, i.e. the mean time required by the population under the congruent condition would be almost equal to the mean time required under the incongruent condition.

$$T_{IL} \sim T_{CL} \approx 0$$

Alternate Hypothesis (H_A): There is a significant difference between the time taken to read through the congruent and incongruent list of colors, i.e., the mean time required by the population under the congruent condition would differ from the mean time required under the incongruent condition.

$$T_{IL} \sim T_{CL} \neq 0$$

To test this, we can employ the "**Dependent t-test**" because we don't know the standard deviation of the population beforehand, and are assuming that the distribution is Gaussian. Also, the population has less than 30 samples, so T-test seems an appropriate choice.

Furthermore, we have to use a **Two-tailed test** because the hypothesis doesn't state a direction and we only have to show that the condition has an impact on the time required, regardless of it being slower or faster.

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

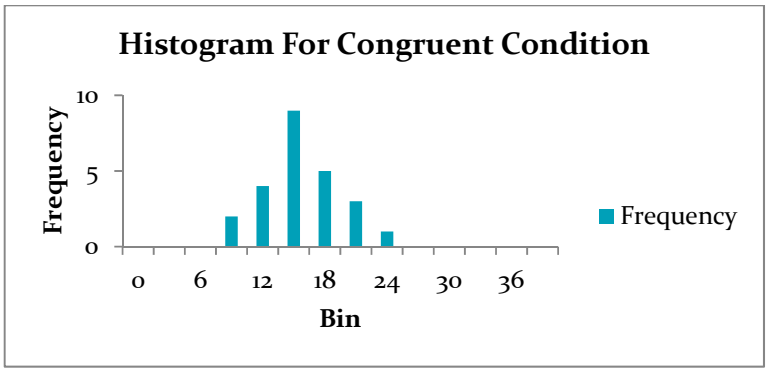
Here are some descriptive statistics regarding this dataset

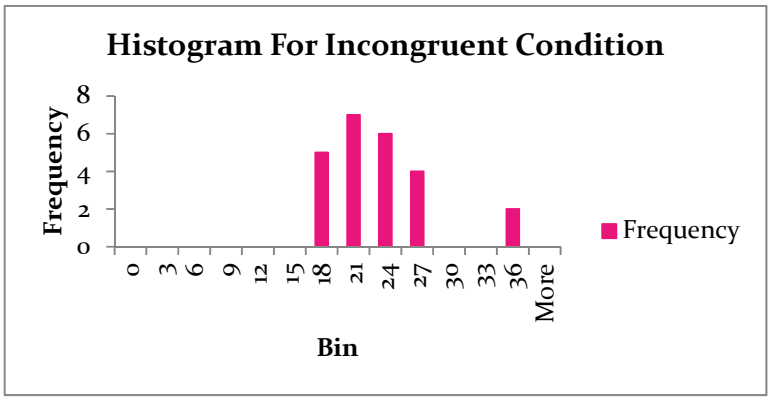
* Values rounded to 3rd decimal place

	Mean (Time)	Median (Time)	Variance (Time)	Std. Deviation (Time)
Congruent	14.051	14.357	12.669	3.559
Incongruent	22.016	21.018	23.012	4.797

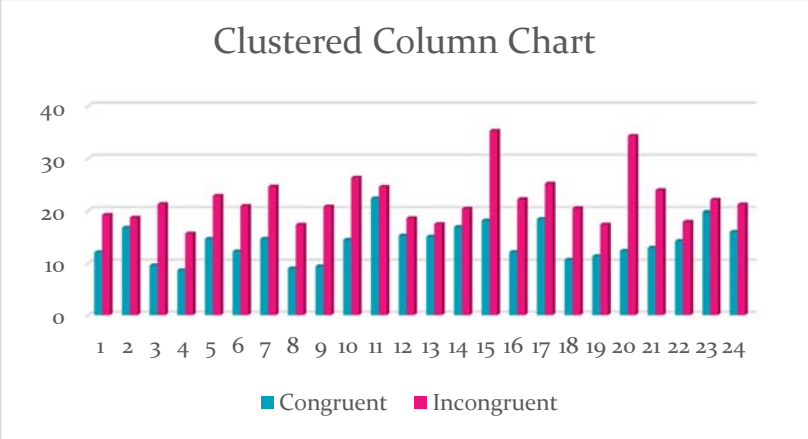
Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Histograms





Clustered Column Chart



Histograms as well as Clustered Column Chart are two fine ways to visualize the overall trend in our data and to note the evident difference between the Congruent v/s Incongruent Reading Conditions.

Both the charts show that the overall time required in the case of Incongruent Reading (in red) is greater than that of Congruent Reading (in blue), though by how much depends from reader to reader.

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?

No. of records in the sample: 24

Mean of Congruent Times: 14.051 s.

Mean of Incongruent Times: 22.016 s.

Mean of $T_{IL} \sim T_{CL}$: 7.964 s.

Std. Deviation of $T_{IL} \sim T_{CL}$: 4.86 s.

T-statistic: (7.964 / (4.86/√24)) = 8.020

Alpha: 0.05 [0.025 for each tail with 23° freedom]

T-critical: ± 2.069 (from t-table)

Confidence Interval: 95%

P-value: 4.103×10^{-8}

Since the P-value is less than 0.05, we can reject the Null Hypothesis (H_o) and conclusively say that there is a significant difference between the time taken to read through the congruent and incongruent list of colors.

This also reinstates my line of expectations as I too experienced a similar result when I took the test:

Congruent Time: 15.513 s. **Incongruent Time: 28.381 s.**

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?

I think the increase in the timings observed for the 2nd case is due to the distraction to the reader while reading the words resulting in a delayed response time.

Over time there have been many variations and modifications to the original Stroop Test like the Emotional Stroop Effect, Numerical Stroop Effect and the Reverse Stroop Effect. I am also aware of other mind benders such as, when you rotate one hand in a circle and move the other vertically for a minute, and then switch motions to the other hand, it is difficult to attain the same speed.