Statistics: The Science of Decisions Project Instructions

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 *in-congruent 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.

NOTES:
- Participant's Task: Say out loud the color of the ink in which the word is printed.
- 2 conditions- Congruent(same word, color) & In-congruent(diff word, color)
- Your Task: Measure the time taken in equally-sized lists.

Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

1. **What is our independent variable? What is our dependent variable?**
   a) **Independent Variable:** The list of words along with their color. Manipulation is done on the basis on this variable.

   b) **Dependent Variable:** The time it takes to name the ink colors in equally sized lists. Time taken varies depending on what independent variable is presented. For eg: RED, BLUE may take less time to read correctly than PURPLE, ORANGE.

2. **What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.**
   a) **Null Hypothesis H_o**: There is no significant time difference(in terms of mean) in the time taken to say the word out loud between the congruent(u_c) and in-congruent(u_i) words condition.
      i. $u_c \sim= u_i$ (time for congruent list similar to in-congruent list) or $u_c - u_i = 0$

   b) **Alternate Hypothesis H_A**: There is a significant difference in the time taken to say the word out loud between the congruent and in-congruent words condition
      i. $u_c != u_i$
      ii. $u_c > u_i$

iii. $u_c < u_i$

Sample Observation:

- Sample Size n = 24
- Population variance is unknown

Statistical Test which can be performed are:

- Paired sample/Dependent t-test: We have 2 sets of observations:
  - Mean time taken to read congruent list
  - Mean time taken to read in-congruent list
  - The above mentioned test is a statistical procedure which is used to determine whether the mean difference between two sets of observations is zero.[1]
  - The Stroop Efect test is a case controlled study where a participant is supplied with a congruent and an in-congruent list of words.
  - The observations are independent.

Now it's your chance to try out the Stroop task for yourself. Go to this link, which has a Java-based applet for performing the Stroop task. Record the times that you received on the task (you do not need to submit your times to the site.) Now, download this dataset which contains results from a number of participants in the task. Each row of the dataset contains the performance for one participant, with the first number their results on the congruent task and the second number their performance on the incongruent task.

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

Calculation done using Excel:

Statistic	Congruent	In-Congruent
N	24	24
Mean	14.05113	22.01592
Median	14.3565	21.0175
Standard Deviation	3.484416	4.696055
Variance	12.14115	22.05293
Standard Error	0.711253	0.958578

Calculation done using matplotlib and pandas:

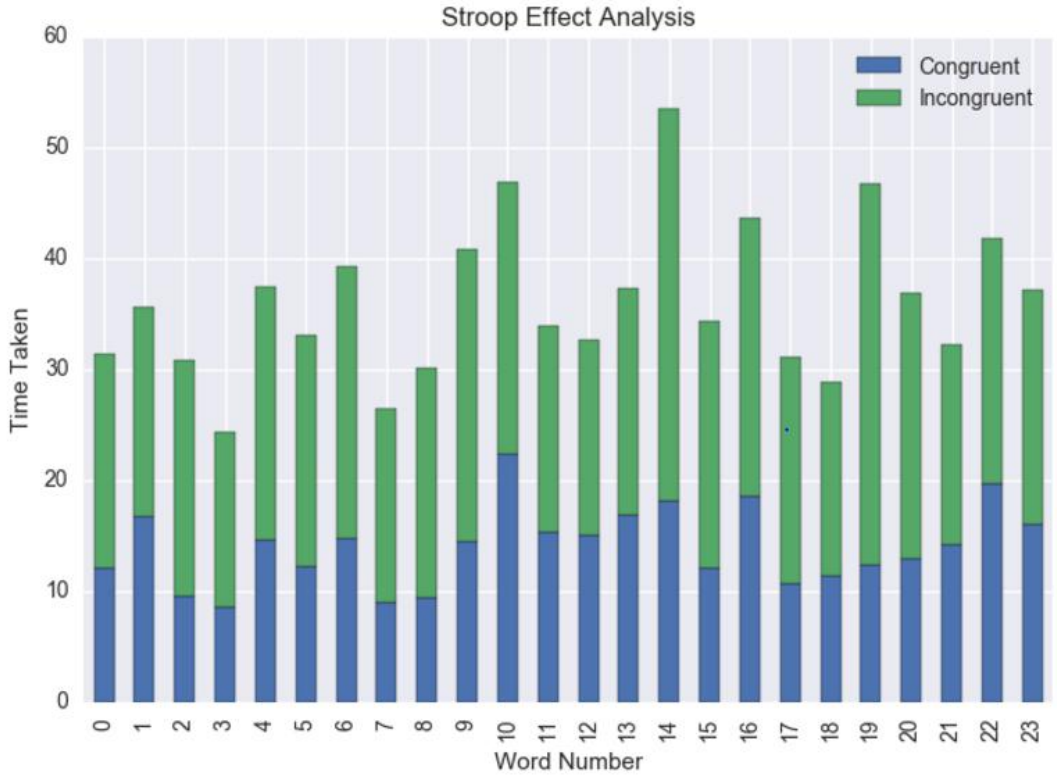
	Congruent	Incongruent
count	24.000000	24.000000
mean	14.051125	22.015917
std	3.559358	4.797057
min	8.630000	15.687000
25%	11.895250	18.716750
50%	14.356500	21.017500
75%	16.200750	24.051500
max	22.328000	35.255000

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.[2]

a) Stacked Histogram :

i. We can see that the time taken to read a word from the in-congruent list is always more via-a-vis congruent list.

ii. The time taken to read in-congruent word is more than double in most of the cases.

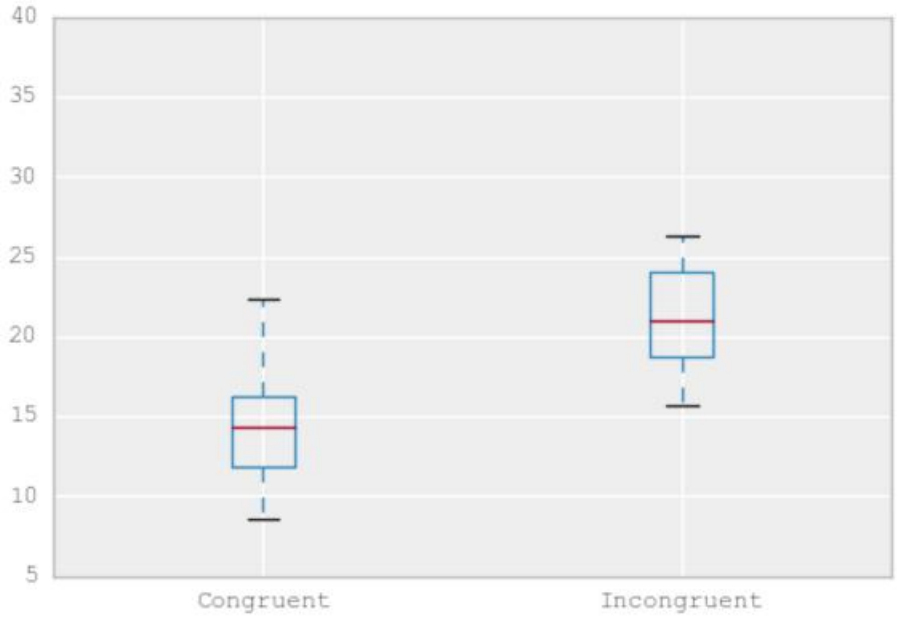


b) Box Plot:

i. The IQR for In-congruent set is more than congruent set.

ii. Using the 'Overall visible spread' to make a comparison when the boxes overlap[3]

1. Difference between Medians (BDM) = (21 - 14) = 7
2. Overall Visible Spread (OVS) = (26 - 23) = 3
3. The Distance Between Medians(DBM) as a percentage of OVS:
   a) DBM = BDM/OVS * 100 = (3/7)*100 = 42.86



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?

a) Alpha= .01, df=23, $t_{critical=}$-2.50, t=-8.02, p-value < 0.001

For the above mentioned confidence level and degrees of freedom, the t critical value for a one-tailed test lies in the negative direction.

The t-statistic lies in the critical region, therefore the null hypothesis, H_o is rejected.

The mean difference between the time to read congruent words and in-congruent words is about 7.96 seconds. In my opinion, this time difference is significant.

The results of the test have matched up 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!

a) I believe our brain is triggered to read the information presented first and then observe the color it is written in. The mind seems to be hard wired in doing so. In order to observe the color before the word, the brain would need to be trained in that manner first.

In psychology, the emotional Stroop task is used as an information-processing approach to assessing emotions. Related to the standard Stroop effect, theemotional Stroop test works by examining the response time of the participant to name colors of negative emotional words.[4]

References:

1. http://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/
2. http://pbpython.com/simple-graphing-pandas.html
3. http://maths.nayland.school.nz/Year_11/AS1.10_Multivar_data/11_Comparing_Boxplots.htm
4. https://en.wikipedia.org/wiki/Emotional_Stroop_test