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 *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

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?

Answer:
Independent variable: The two different conditions: congruent words condition and incongruent words 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.
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.

Answer: My performances on the test: Congruent condition: 9.474s; Incongruent condition: 20.737s.

Null hypotheses: There is no differences in terms of population mean of time taking to name the ink colors between the Incongruent words condition and congruent words condition, or it even takes more population mean of time for congruent words condition than incongruent words condition.
That is: $\mu_{incon} - \mu_{con} \leq 0$, where μ_{incon} refer to the population mean of time consumed by Incongruent words condition, and μ_{con} refer to the population mean of time consumed by congruent condition respectively.

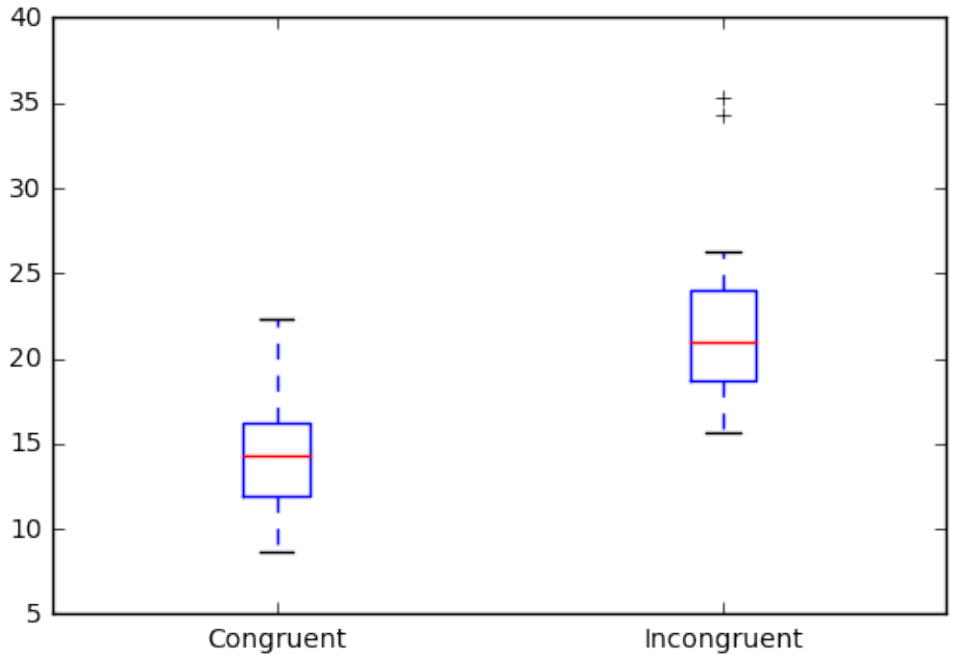
Alternative hypotheses: Incongruent words condition will lead more population mean of time taking to name the ink colors, comparing to the population mean of time consuming by congruent words conditions.
That is: $\mu_{incon} - \mu_{con} > 0$, where μ_{incon} refer to the population mean of time consumed by Incongruent words condition, and μ_{con} refer to the population mean of time consumed by congruent condition respectively.

I will take the provided dataset as two dependent samples, and take a 'pre-test and post-test' kind, one-tailed t-test, with αlevel equals 0.05.

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

Sample Means: $\overline{x}_{incon}$=22.016, $\overline{x}_{con}$=14.051
Sample standard deviation: S_{incon}=3.559, S_{con}=4.797
Standard deviation of sample means'differences: $S_{incon-con}$=4.865

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.



From the Boxplot, we can simply see that almost the whole box, first quartile to the third quartile of the consuming of time of incongruent words condition is higher than congruent words condition's. Therefore, I expect that there is signification longer consuming time for the incongruent words condition then the congruent word condition's.
(For the python code, please refer to another file.)

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?

The confidence level is 95%, corresponding αlevel equals 0.05.
The degrees of freedom: df=n-1=24-1=23
The critical T value: $t_{critical}$=1.714
The t statistic: $t_{statistic}=\frac{\overline{x}incon-\overline{x}con}{S/\sqrt{n}}$=8.021
Since the $t_{statistic}$ > $t_{critical}$, we reject the null hypothesis, that's to say we accept the alternative hypothesis.
Conclusion: Incongruent words condition will lead more time taking to name the ink colors, comparing to the time consuming by congruent words conditions.
That is: $\mu_{incon} > \mu_{con}$

This result does match up with my expectations, especially with the 'Distribution of sample data' plot above.

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!