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?
   Independent variable: type of test
   Dependent variable: count time

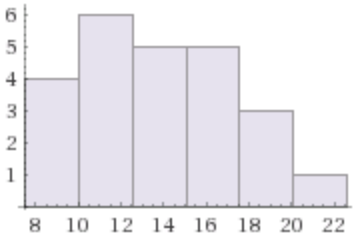
2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.
   Null hypothesis: there's no significant difference between two tests
   H0: μ1 = μ2
   Alternative hypothesis: There is significant difference between tests
   Ha: μ1 != μ2
   μ1: the average result time from congruent test
   μ2: the average result time from incongruent test

Type of test: Dependent t-test for paired two condition samples.
Since we have limited sample data, therefore the population parameter is unknown. In this situation, it's best to conduct t-test. Also two samples for paired with different conditions. It is best to choose Dependent t-test for paired two condition samples for this problem.

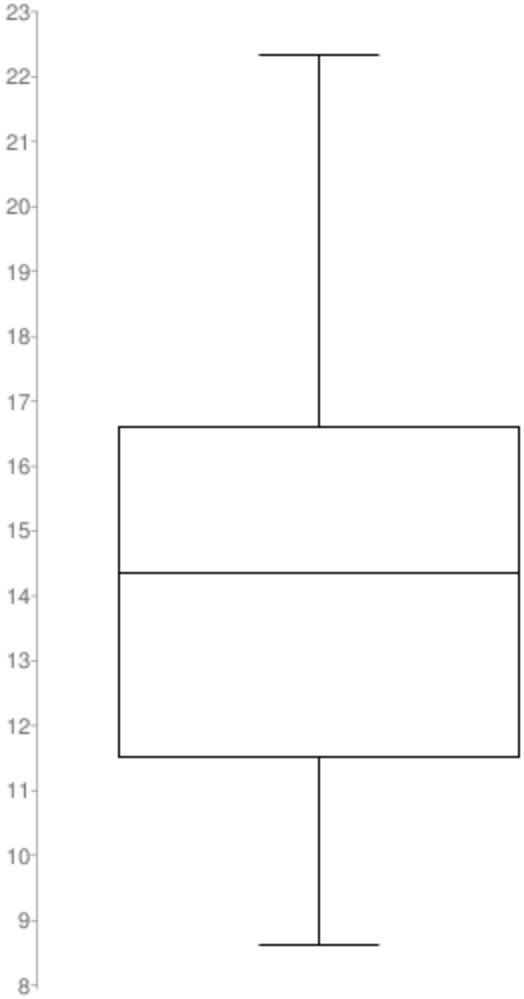
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

mean difference	-7.96479
n	24
df	23
std dev of (a-b)	4.864827
standard error	0.993029

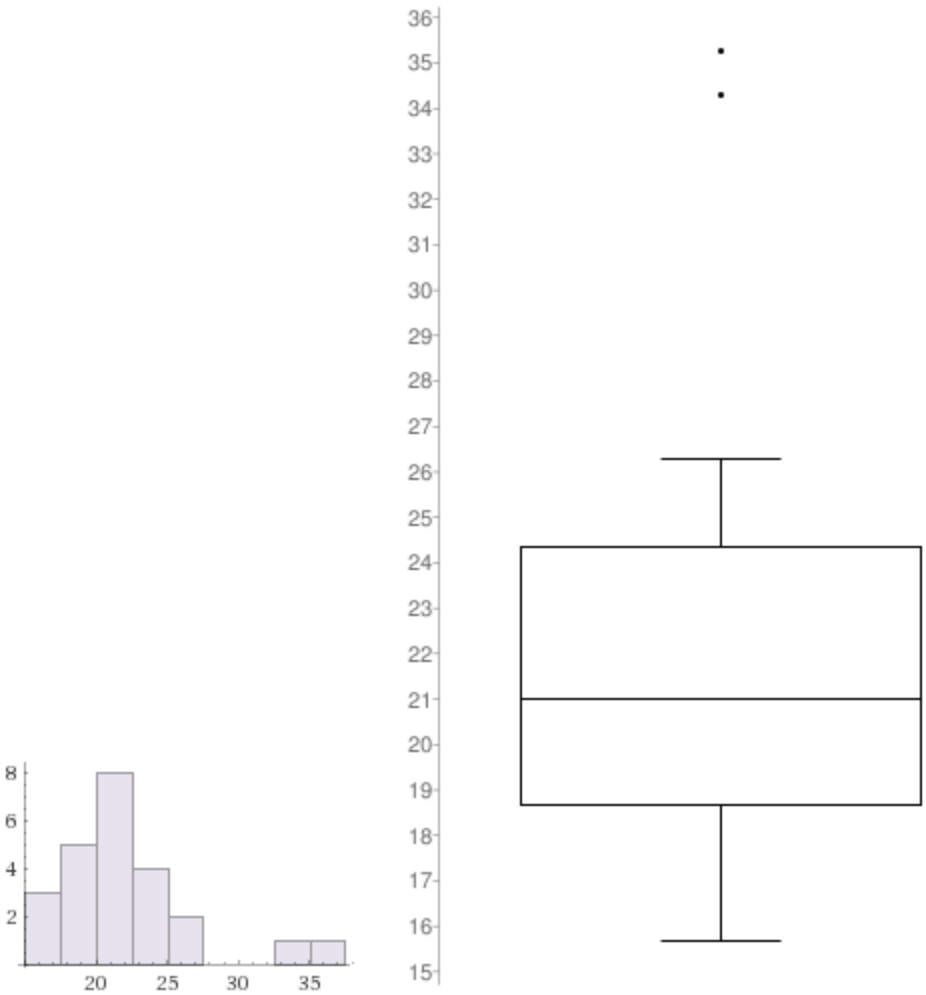
4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.



Histogram for Congruent



Boxplot for Congruent



Histogram for Incongruent Boxplot for Incongruent

Congruent is more in line with normal distribution, and Incongruent is more like positive skewed distributions, with Incongruent has wider spread and more number toward the right tail.

Based on the boxplot, Incongruent set has some outliers.

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?

   Two tailed test, with alpha = 0.05, df= 23, t-critical is about $\pm$2.069. Based on my calculation,

mean difference	-7.96479167
n	24
df	23
std dev of (a-b)	4.86482691
standard error	0.993028635
t-statistic	-8.020706947
t-critical (alpha = 0.05)	-2.069

The t-statistic is less than t-critical on the left side, so it is in the critical region. Therefore it shows that it is significant different. Based on this calculation, I reject the null hypothesis.

This result is expected, since my personal test result showed a significant difference in time from two tests.

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!

Our brain behaves in a way that the brain reacts to the word first, then the color of the word. So there is a delay in brain when it tries to read the secondary parameter. Our brain is trained automatically linked the word, but now it has to interpret the color factor, and break the automatic chain. It has to take more process time to reach the conclusion. Therefore, the response time is much longer.

Similar task:

Recite A to Z alphabet in both ascending order and descending order, and time the result. Then based on the two results to see if there is significant difference between the two.