

Statistics: The Science of Decisions Project Instructions

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

Independent Variable: Color of ink and name of color

Dependent Variable: Time taken to name ink colors from a fix sized list of color names.

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

Let us say population averages for congruent and incongruent tasks are represented as,

Mean time for congruent tasks = μ_c

Mean time for incongruent tasks = μ_i

H_0 : Average time to name colors is the same for congruent and incongruent tasks i.e $\mu_c = \mu_i$

H_a : Average time to name colors is the same for congruent and incongruent tasks is not equal i.e $\mu_c \neq \mu_i$

As each participant goes through congruent and incongruent task, and then times are recorded, so we know samples are dependent.

From histograms we also know that both samples are roughly normally distributed around mean. It is evident that sample size is very small as compared to the overall population. T-Test is a good indicator of whether or not two groups' averages most likely reflects a real difference in the population from which the groups were sampled. So a two-tailed dependent t – test will be a perfect statistical test in this situation.

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

\bar{x}_c = 14.05 seconds (Mean of times for Congruent)

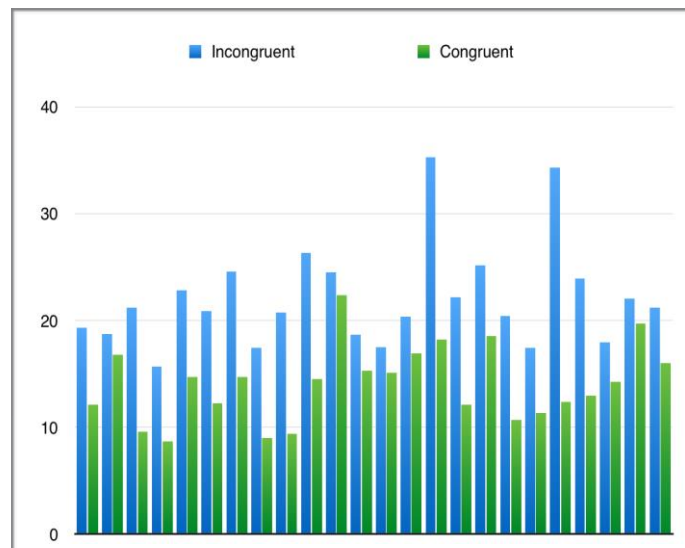
S_c = 3.56 (Standard Deviation for Congruent)

\bar{x}_i = 22.02 seconds (Mean of times for Incongruent)

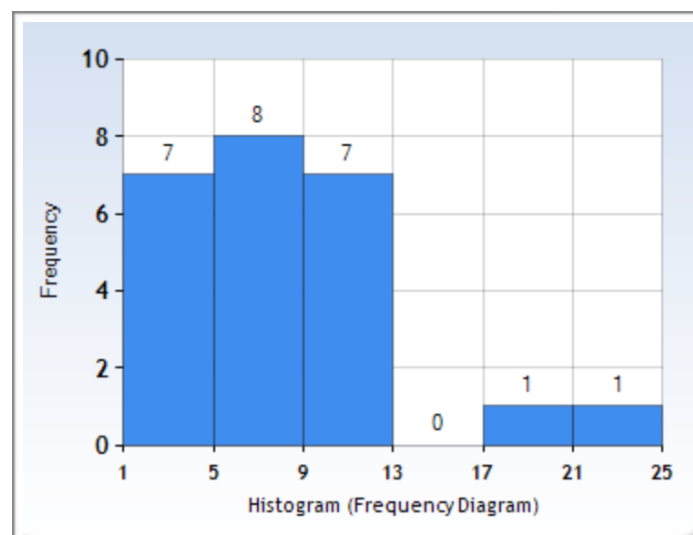
S_i = 4.80 (Standard Deviation for Incongruent)

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.

A. Side by Side Graph for each participant



B. Histogram of differences



From Graph A we observe that some participant finished incongruent test with less difference and some participants finished with more difference. But we also can conclude that all of participants in the sample took more time in incongruent test than congruent test. That gives us a clue that we might have to reject the null hypothesis.

Graph B is histogram of differences. It tells us that the maximum difference is at 8. Majority of people took 8 seconds more than congruent test. There few people who

even took more than 17 seconds than the congruent task. That also gives us a clue that we might have to reject the null hypothesis.

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?

First of all we want to mention that our confidence level is 95%. That means that we will be using $\alpha = 0.05$

Formula for calculating t-statistic is

$$t\text{-value} = (\bar{x}_c - \bar{x}_i) / SE$$

$$\text{Degree of freedom, } df = (n-1) = 23$$

To calculate t-value of sample, we will need the Sample standard deviation of mean differences S_m .

$$S_m = 4.86$$

$$\text{Standard Error, } SE = 0.99$$

Formula for calculating t-statistic is

$$t\text{-value} = -8.03$$

From t-table, for $\alpha = 0.05$, the $t\text{-critical} = -2.069$.

As **$t\text{-value} \ll t\text{-critical}$** , so we reject the Null Hypothesis H_0 because the t-value is in critical region. It takes less time to recognize the color of words with the congruent condition vis-a-vis the incongruent condition.

This is also what we expected from the “Graph A” and “Graph B”