

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?

The Independent Variable : The wording condition, which can be Congruent or Incongruent.

The Dependant Variable: The time it takes to name the ink colors.

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

Hypotheses:

Given the sample we would inference some information on population data, if the test is applied to population the hypotheses will be as follows:

μ_i : The mean of population of Congruent word condition.

μ_c : The mean of population of Incongruent word condition.

Null Hypotheses: The word condition doesn't have an effect on Time it takes to name colors, meaning the response time will be equal on both conditions.

$$H_0: \mu_1 = \mu_2$$

Alternative Hypotheses: The word condition (Incongruent) will have an effect (Increase) on Time it takes to name colors, meaning the response time of congruent condition will be smaller than incongruent word condition.

$$H_1: \mu_1 < \mu_2$$

Statistical Test:

Will perform One-Tailed Dependent-Sample T-test, at Alpha level = 0.05

This is conducted due to many factors:

- 1- The T-test is used to compare the sample means for each condition.
- 2- T-test is more applicable because we don't have the population data.
- 3- The test will be Dependent-Sample, because the participant will be exposed to both condition, meaning the data is of the same sample.
- 4- Directional One-Tailed test is chosen because we want to study the effect of changing the wording condition and how it differs.

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.

Given the dataset below:

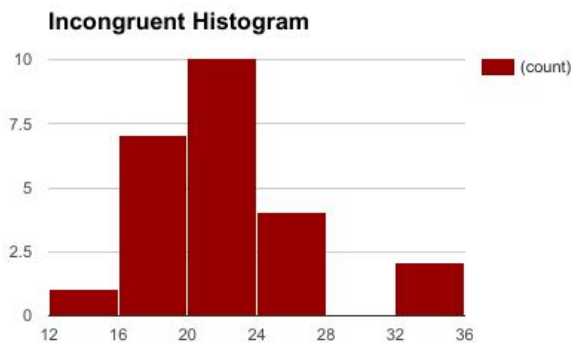
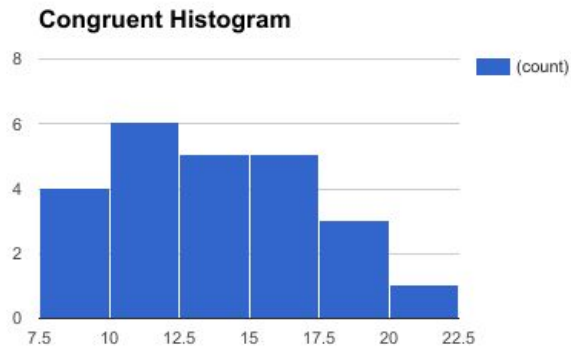
- The Mean of both conditions will be calculated as measure of central tendency, will be denoted:
 \bar{x}_c : Mean of Congruent condition sample.
 \bar{x}_i : Mean of Incongruent condition sample.
 \bar{x}_d : Mean of difference between two samples ($\bar{x}_c - \bar{x}_i$).
- The Standard-Deviation will be calculated as measure of variability, will be denoted:
 S_c : Standard-Deviation of Congruent condition.
 S_i : Standard-Deviation of Incongruent condition.
 S_d : Standard-Deviation of difference between two samples ($\bar{x}_c - \bar{x}_i$).

$\bar{X}_c = 14.051125$	$S_c = 3.559357958$
$\bar{X}_i = 22.01591667$	$S_i = 4.797057122$
$\bar{X}_d = -7.964791667$	$S_d = 4.86482691$

Congruent	Incongruent	$\bar{X}_c - \bar{X}_i$
12.079	19.278	-7.199
16.791	18.741	-1.95
9.564	21.214	-11.65
8.63	15.687	-7.057
14.669	22.803	-8.134
12.238	20.878	-8.64
14.692	24.572	-9.88
8.987	17.394	-8.407
9.401	20.762	-11.361
14.48	26.282	-11.802
22.328	24.524	-2.196
15.298	18.644	-3.346
15.073	17.51	-2.437
16.929	20.33	-3.401
18.2	35.255	-17.055
12.13	22.158	-10.028
18.495	25.139	-6.644
10.639	20.429	-9.79
11.344	17.425	-6.081
12.369	34.288	-21.919
12.944	23.894	-10.95
14.233	17.96	-3.727
19.71	22.058	-2.348
16.004	21.157	-5.153

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 histogram of the two reported sample conditions, it can noticeable that there are differences in the mean of each one, which may be resulted from the wording conditions, but that must be further tested.



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?

Statistical test:

Dependent-Sample T-test, will be tested against positive one-tailed of Alpha level = 0.05

Hypotheses:

$H_0: \mu_c = \mu_i$

$H_1: \mu_c < \mu_i$

Sample Size:

$N = 24$

Degree of Confidence (Df):

$$N-1 = 23$$

$$\alpha = 0.05$$

$$t_{0.05} = 1.714$$

T-test Steps:

1- Calculating the mean of difference between two samples ($\bar{X}_c - \bar{X}_i$):

$$\bar{X}_d = -7.964791667$$

2- Calculating the Standard-Deviation of difference between two samples ($\bar{X}_c - \bar{X}_i$):

$$S_d = 4.86482691$$

3- Calculating the Standard Error

$$SE(d) = 0.9930286348$$

4- Calculating the t-statistic:

$$T = -8.020706944$$

5- Finding the P-Value with $t(23)$:

$$P \approx 0.0001$$

6- Decision:

The P-Value is statistically significant, therefore the null hypotheses is rejected.

$$t(23) = -8.020706944, p < 0.0001$$

This means the wording condition have an effect on the response time. The decrease in concurrent condition is not due to chance.

Note: all calculation can be found in the excel sheet.

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!

Would like to think about it and submit the project again later!