

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

Independent Variable:

Here the independent variable is the test that each participant will be going through i.e. Congruent or Incongruent.

Dependent Variable:

Dependent variable is the reaction time i.e. the time taken to identify the correct answer relate to the test type.

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

Hypothesis:

Here we are considering a set of individuals (who we are assuming that are chosen at random) as our sample as it is not feasible to apply the test for all the people in the world i.e. the whole population. The behavior of the sample will be used to infer regarding the behavior of the whole population. Hence, we'll replace the population mean time of congruent test ( $\mu_c$ ) with " $\alpha$ " and the population mean time of incongruent test ( $\mu_i$ ) with " $\gamma$ ".

**Null Hypothesis ( $H_0$ ):** the mean time taken by the population to perform a congruent test of Stroop effect will be equal to the mean time taken by the population to perform the incongruent test.

$$\mu_c = \mu_i$$

**Alternative Hypothesis ( $H_a$ ):** The mean time taken by the population to perform a congruent test of Stroop effect will not be equal to the mean time taken by the population to perform the incongruent test.

$$\mu_c \neq \mu_i$$

Note:  $\alpha$  = Congruent Test Mean Time &  $\gamma$  = Incongruent Test Mean Time

Statistical Test:

Here, we are going to perform a "**two tailed dependent sample t-test**". The reasons for the chosen statistical test is mentioned below:

- We do not have the population parameters Population Mean & Population SD.
- We are having a sample size which is comparatively small (less than 30)
- We assume that the distribution is Gaussian
- We assume that the sample is chosen at random where probability of selection each sample was equally-likely & independent.
- We are doing measuring the same dependent variable for 2 different test condition on the same sample, where all the participants undergo the same treatment.
- A two tailed dependent t-test is chosen because we do not know the outcome of the direction on the mean after the second treatment
- Both the test is performed by the same set of individuals under the same environmental conditions, i.e. no other external factor which can affect the results of the reaction time is not present.

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

Descriptive Statistics:

We are going to measure the below statistics which will give us an idea about the data we are going to deal with:

- Mean
- Variance
- Standard Deviation

Mean:

Mean Time for Congruent Test ( $\alpha$ ) = 14.05

$$\bar{x} = \frac{\sum_{i=1}^n x}{n} = 14.05$$

Mean Time for Incongruent Test ( $\gamma$ ) = 22.02

$$\bar{y} = \frac{\sum_{i=1}^n y}{n} = 22.02$$

Variance:

Variance of Congruent Test ( $S_x^2$ ): 12.67

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)} = 12.67$$

Variance of Incongruent Test ( $S_y^2$ ): 23.01

$$S_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{(n-1)} = 23.01$$

Standard Deviation:

Standard Deviation of Congruent Test ( $S_x$ ): 3.56

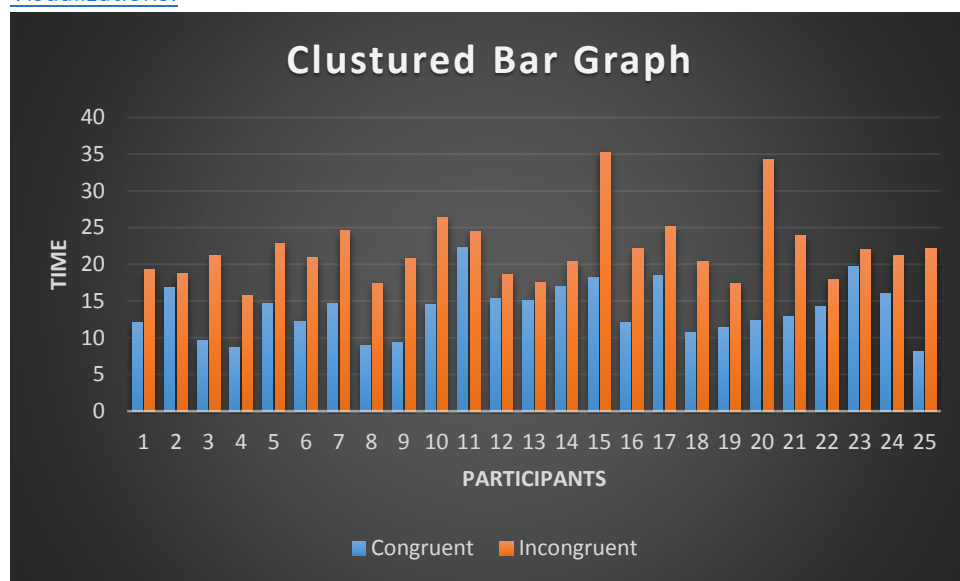
$$S_x = +\sqrt{S_x^2} = 3.56$$

Standard Deviation of Incongruent Test ( $S_y$ ): 4.79

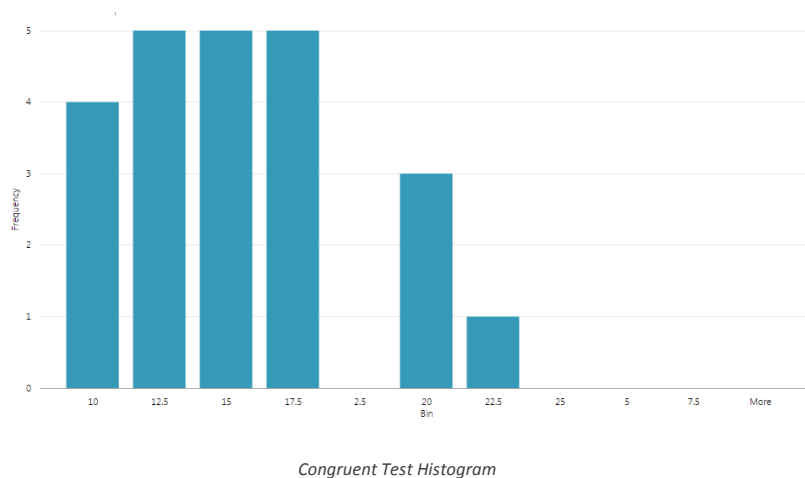
$$S_y = +\sqrt{S_y^2} = 4.79$$

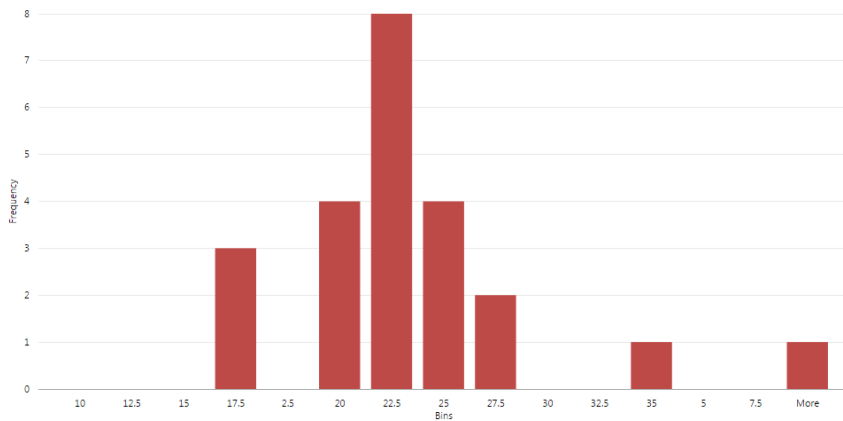
4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Visualizations:



This is a clustered bar graph. In here we can clearly see the time taken for each test by each individual. From the graph we get an idea that the time taken for incongruent test is much more.





*Incongruent Test Histogram*

From the above two Histograms we can clearly see that most of the people completed the Congruent Test within 6-17.5 seconds, whereas most of the people for congruent test completed the test within 16-27.5 seconds.

A significant increase in the time for the completion of test when the data is incongruent is noticed.

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:

Considering confidence level of 95%, i.e.  $\alpha = 0.05$  and  $t_{critical}$  for  $df = 23$  is  $\pm 2.069$  for two tailed distribution.

Congruent Mean ( $\bar{x}$ ) = 14.05

Incongruent Mean ( $\bar{y}$ ) = 22.02

Calculating Difference of Mean:

$$\bar{x}_D = (\bar{x} - \bar{y}) = 14.05 - 22.02 = -7.96$$

$$S_D = \sqrt{\frac{(x_{di} - \bar{x}_D)^2}{(n-1)}} = \sqrt{\frac{544.33044}{23}} = 4.86$$

#### Calculating Standard Error:

$$S.E. = \frac{S_D}{\sqrt{n}} = \frac{4.86}{\sqrt{24}} = 0.99$$

Calculation  $t_{statistic}$ :

$$t_{statistic} = \frac{\bar{x}_D}{S.E.} = \frac{-7.97}{0.99} = -8.051$$

$t_{critical} = \pm 2.069$ , where  $\alpha = 0.05$  &  $df = 23$ , for two-tailed distribution

$P_{value} = 0.0001$ , extremely statistically significant result.

#### Decision:

$t_{statistic}$  (-8.051) is much less than the  $t_{critical}$  (-2.069 in this case) value. Hence, we reject the Null Hypothesis. This leads to our conclusion that, the time taken by a person during the incongruent test will be much higher.

Our Expectation is met.

#### Conclusion:

It takes more time to name the colors of color words if the actual observed colors of the words don't match the colors that the words denote than if they do match, i.e. if the word "GREEN" is shown in 'green' color, it'll take less time to name it than if it is in some other color and we have to name the color.

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!

Responsible:

There is a lag in the brain's ability to recognize the color of the word since the brain reads words faster than it recognizes colors. Color recognition as opposed to reading a word, requires more attention, the brain needs to use more attention to recognize a

color than to word encoding, so it takes a little longer. Since recognizing colors is not an “automatic process” there is hesitancy to respond; whereas, the brain automatically understands the meaning of words as a result of habitual reading.

#### Similar Task:

Similar tasks, can be done with geometric shapes (square, triangle, rectangle etc.), words & colors.

In the first test a person will have to tell the color a geometric shape is of and in the second test the person will have to tell the name of the geometric shape written inside the geometric shape.

#### Tools Used & Websites:

- MS Word 2013
- MS Excel 2013
- MS Power BI
- Wikipedia
- Statistics Leard