Project 1: Test a Perceptual Phenomenon - Stroop Task

Submitted by: Bharteesh Kulkarni

Course: Udacity Data Analyst Nanodegree

Date: 09/04/2015

Introduction

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.

Q1: What is our independent variable? What is our dependent variable?

A: Independent variable in this experiment is the Congruency/Incongruency of color words and their printed colors (whether word name and color of font was same or different). Dependent variable is the average time taken to name the ink color.

Q2: What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

A: For this task, **Paired Sample t-test** would be appropriate to compare two population means, as the two samples are correlated.

A **paired sample t-test** is used to determine whether there is a significant difference between the mean of the same measurement (e.g. time taken) calculated under two different conditions. Both measurements are made on each unit in a sample, and the test is based on the paired differences between these two values.

Following set of hypotheses would be needed:

- Null Hypothesis (H_0: Mean of two samples are equal (or) Difference in mean values is zero.)
- Alternate Hypothesis (H_a: Mean of two samples are not equal (or) Difference in mean values is > 0 or < 0)

Q3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

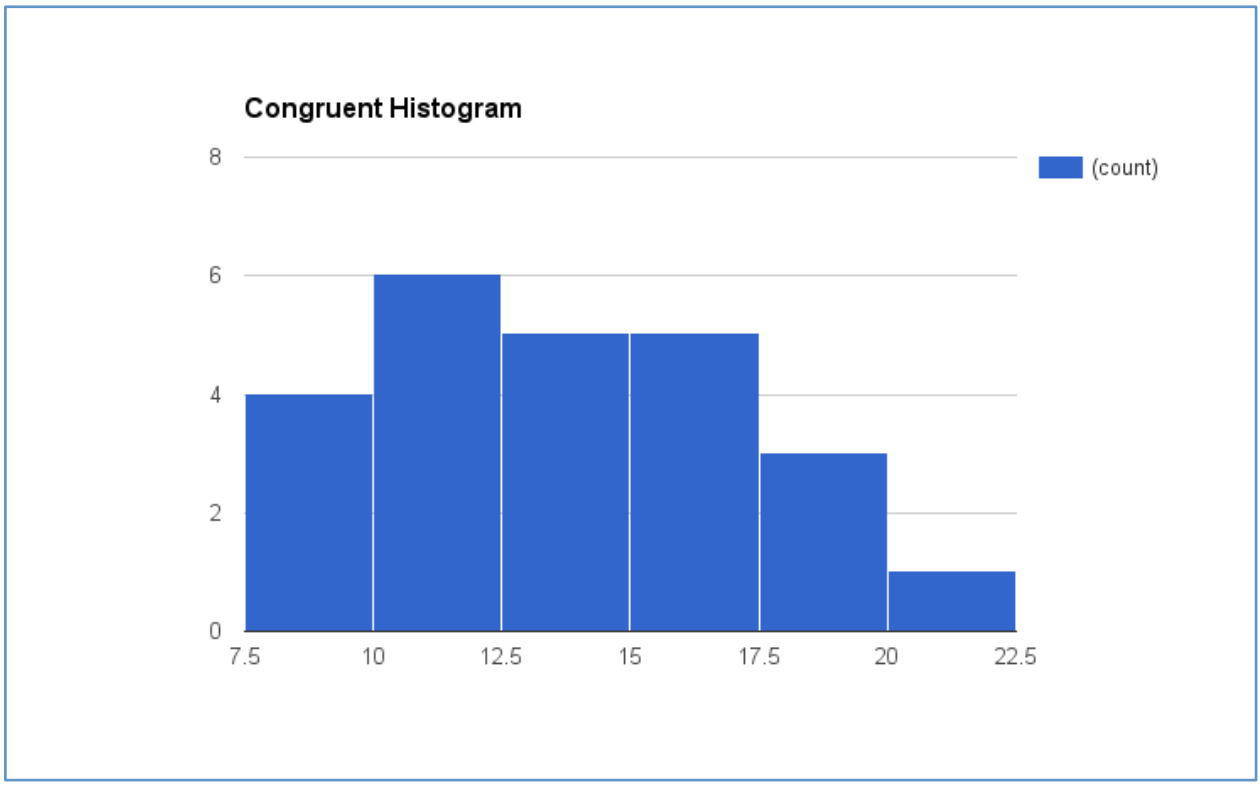
A: Measure of central tendency: Mean values for congruent and incongruent performances are shown below:

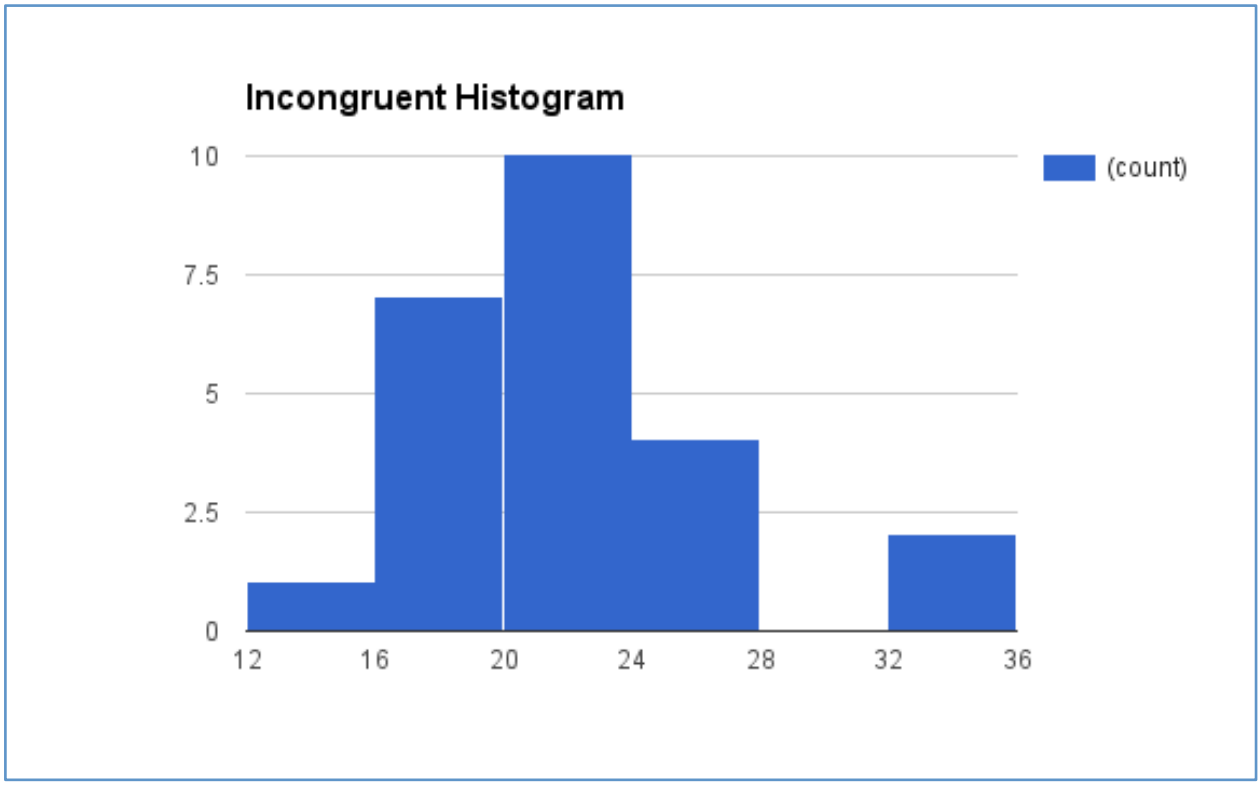
$\mu_{congruent}$ (μ_a) = 14.051, $\mu_{incongruent}$ (μ_b)= 22.015

Measure of variability: Sample Standard deviation for congruent and incongruent performances are shown below:

$\sigma_{congruent}$ (σ_a)= 3.56 , $\sigma_{incongruent}$ (σ_b) = 4.80

Q4: 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.





Observation about Histograms:

Most of the participants took less than 20 seconds to read congruent words. However, to read incongruent words, most of the participants took over 20 seconds.

Q5. 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, mean difference (d) is calculated:

$d = \mu_a - \mu_b = -7.964$

Standard deviation of the differences:

$S_d = 4.865$

Standard error of the mean difference:

$SE(d) = S_d / \sqrt{n} = 0.993$

t-statistic:

$T = d / SE(d) = -8.02$

Degrees of freedom (df) = n – 1 = 23

Confidence level (α)= 0.05

t-critical value from t-table with 23 df and $\alpha\ value\ of\ 0.05$ = 1.714

p-value < 0.0001 (calculated from www.graphpad.com with 23 df)

As the p-value is significantly less than the t-critical value, we reject the Null Hypothesis

Conclusion: Since the difference of Means is a negative number, it can be concluded from this experiment that the average time taken by participants to read Congruent words is less than the average time taken to read Incongruent words.

References

1. https://en.wikipedia.org/wiki/Stroop_effect
2. http://stattrek.com/hypothesis-test/hypothesis-testing.aspx
3. http://www.ats.ucla.edu/stat/mult_pkg/faq/general/tail_tests.htm
4. http://afni.nimh.nih.gov/pub/dist/HOWTO/howto/ht05_group/html/background_ANOVA.html