P1: Test a Perceptual Phenomenon

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.

Raw Data in seconds for each test:

Congruent	Incongruent
12.079	19.278
16.791	18.741
9.564	21.214
8.63	15.687
14.669	22.803
12.238	20.878
14.692	24.572
8.987	17.394
9.401	20.762
14.48	26.282
22.328	24.524
15.298	18.644
15.073	17.51
16.929	20.33
18.2	35.255
12.13	22.158
18.495	25.139
10.639	20.429
11.344	17.425
12.369	34.288
12.944	23.894
14.233	17.96
19.71	22.058
16.004	21.157

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

   The word condition (congruent or incongruent) is the independent variable and the dependent variable is the time (in seconds).

   So the condition of the words (congruent or incongruent) will make the time vary.

   It is the independent variable that causes a change in the dependent variable.

   Independent variable: Word condition
   Dependent variable: Time (in seconds

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

   The null hypothesis is that there is no difference in time while reading the congruent list of words (μ_C) than when reading the incongruent list of words (μ_I)

   $$H_0: \mu_C - \mu_I = 0$$

   The alternative hypothesis is that there is a significant difference in time while reading the congruent list of words versus reading the incongruent list of words taking less time to read the congruent list of words.

   $$H_A: \mu_C - \mu_I < 0$$

   I'm going to perform a one tail in the negative direction statistical test because I want to prove that it takes less time to read the congruent list of words than the incongruent one.

   Will use a **dependent t-test for paired** samples for the following reasons:

   a. We don't know the population standard deviation
   b. The same subject takes the test twice (2 conditions)
   c. $n < 30$

3. *Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.*
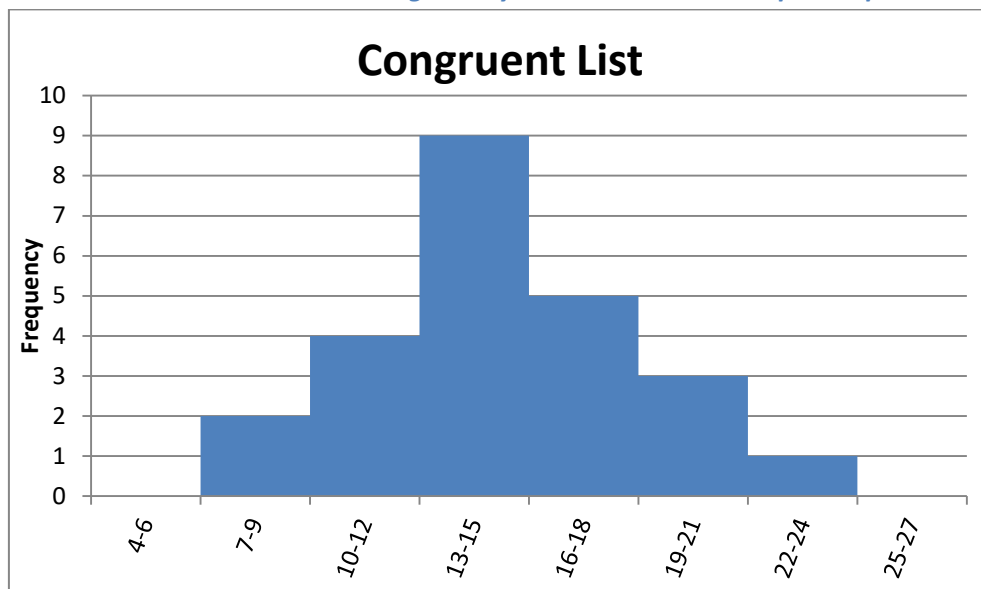
$\bar{x}_C = 14.05$	$S_C = 3.56$	$n_C = 24$	$df_C = 23$
$\bar{x}_I = 22.02$	$S_I = 4.80$	$n_I = 24$	$df_I = 23$
$\bar{x}_{DIF} = -7.96$	$S_{DIF} = 4.86$	$n_{DIF} = 24$	$df_{DIF} = 23$
$SE = 0.64$			

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.*

Congruent List

This is the histogram for the congruent list with a bin size of 3 in a normal distribution we can easily see that the mean is in the 13-15 bar (the mean is actually 14.05), we can also see that the times range from 7 to 24 seconds.
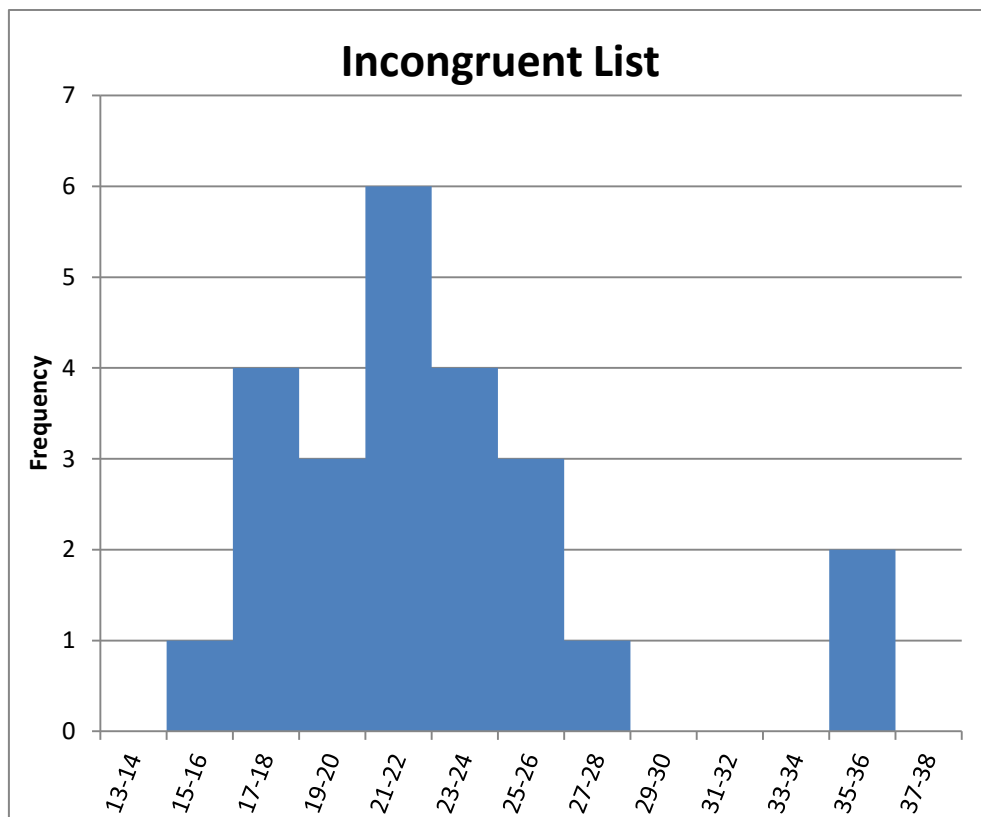
Incongruent List

This histogram is from the incongruent list, with a bin size of 2, it is easy to see the mean here as wee on the 21-22 bar (the mean is 22.02) it is also easy to see that the time range is between 15 up to 36 seconds.

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?*

We will use a 95% Confidence level hence α = 0.05 with a critical value of -1. 714 (since we are testing on the negative side.

α = 0.05 $t_{Critical}$ = - 1.714

$$t = \frac{\overline{x}_C - \overline{x}_I}{S_{DIF}/\sqrt{n}}$$

t = -8.03

We can see that $t_{Critical}$ = - 1.714 and that the actual t-statistic for the values in this project is -8.03 (p value will be less than 0.0001) which is a lot further than our critical point so we "**Reject the null hypothesis**". ~~$H_0: \mu_C - \mu_I = 0$~~

So the experiment result is that there is a significant difference in time while reading the congruent list of words versus reading the incongruent list of words taking less time to read the congruent list of words.

The results do match my expectations.

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!*

I believe it has something to do with the way the brain processes the information, since it is getting 2 different and contradictory values at the same time that relate to the same type of things there is a delay in response time.

Resources:

Udacity Forums
Udacity nanodegree videos
http://www.graphpad.com/quickcalcs/
http://www.ats.ucla.edu/stat/mult_pkg/faq/general/tail_tests.htm
http://www.statstutor.ac.uk/resources/uploaded/paired-t-test.pdf
https://www.ltcconline.net/greenl/courses/201/descstat/mean.htm