Andrew Maner
Udacity Data Analyst Nanodegree
Project 1: Test a Perceptual Phenomenon

Questions for Investigation

Question 1, part 1: What is our independent variable?

The independent variable in a Stroop task experiment is the condition. This is a categorical variable with two levels – *congruent* and *incongruent*. The congruent level corresponds to the list of words whose colorings match their text. (For example, the word "RED" would be colored red.) The incongruent level corresponds to the list of words whose colorings do not match their text. (For example, the word "RED" could be colored green, or any other color in the list, *except for* red.)

Question 1, part 2: What is our dependent variable?

The dependent variable in a Stroop task experiment is the time it takes for a person to name the ink colors in a list of words.

Question 2, part 1: What is an appropriate set of hypotheses for this task?

The default null hypothesis in a test like this would be, "there is no difference in the average times for the two types of lists." In other words, the null hypothesis is saying that the average person would take the same amount of time to read out the colors from congruent and incongruent lists of the same size. We are interested in testing for a difference in the congruent and incongruent population means, so our null hypothesis is written as:

$$H_0: \mu_{I-c} = 0$$

The alternate hypothesis would depend on whether or not the researcher is testing for a difference between the two or believes it will take longer for a person to read the colors from an incongruent list than from a congruent list. If the researcher suspects it takes longer to read from an incongruent list than from a congruent list, then the alternate would be written as:

$$H_A: \mu_{I-C} > 0$$

(If, however, the research simply wants to test for a difference, then the alternate would be written as: $H_A: \mu_{I-C} \neq 0$. We will proceed with the former and test to see if it takes longer to read from an incongruent list.)

Question 2, part 2: What kind of statistical test do you expect to perform?

Since these are paired (dependent) samples and we have fewer than thirty (30) sample pairs and do not know the population variances, I would expect to perform a paired t-test. In order to perform this test, we must assume the congruent and incongruent populations to be normally distributed. The histograms shown in response to question four below do not explicitly preclude this, so we should be able to proceed with the t-test.

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

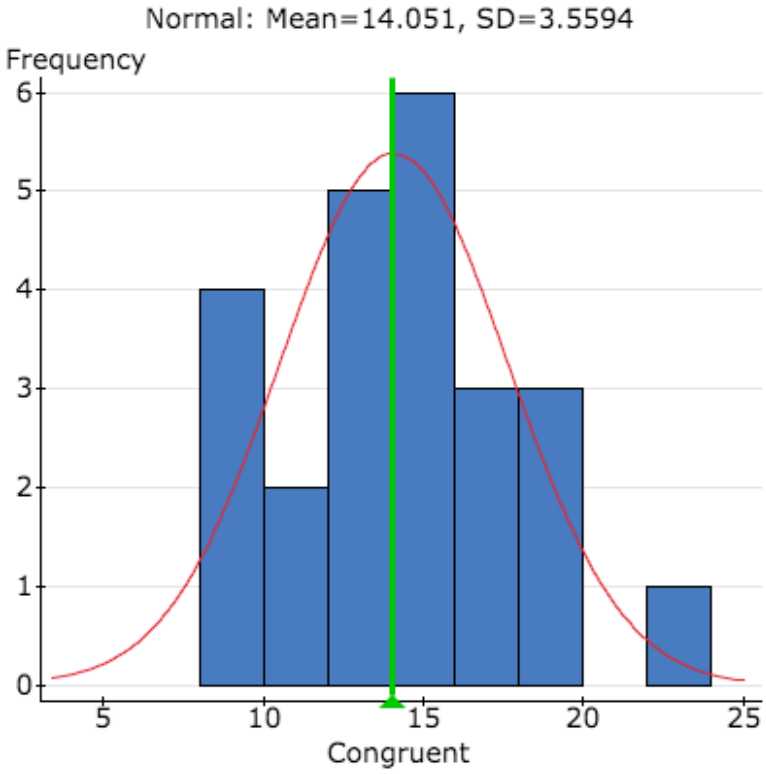
Using StatCrunch [1], I computed the following:

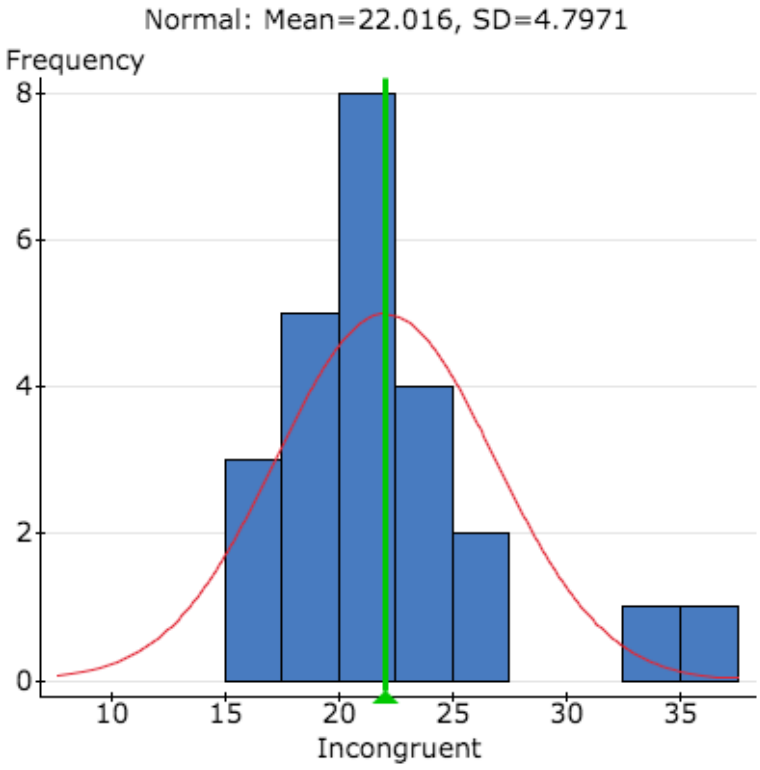
Column	Mean	Variance	Std. dev.	Median	Range	Min	Max	IQR	Mode
Congruent	14.051125	12.669029	3.559358	14.3565	13.698	8.63	22.328	4.686	No unique
Incongruent	22.015917	23.011757	4.7970571	21.0175	19.568	15.687	35.255	5.5165	No unique

The mean of the incongruent sample is 56.7% ((22.016-14.051)/14.051) larger than the mean of the congruent sample, while the median of the incongruent sample is 46.8% larger than that of the congruent sample. The differences seem significant, but a t-test should still be run. It is worth noting that the variability of the incongruent sample is greater – by all measures shown here (variance, standard deviation, range, IQR) – than that of the congruent sample.

Question 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.

(Note: I used StatCrunch to generate these graphs.)


Normal: Mean=14.051, SD=3.5594
Frequency
6
5
4
3
2
1
0
5
10
15
20
25
Congruent


Normal: Mean=22.016, SD=4.7971
Frequency
8
6
4
2
0
10
15
20
25
30
35
Incongruent

Both histograms are shown here with their means and corresponding normal distributions. Both are slightly skewed – the Incongruent sample more so – but neither is so skewed as to prevent us from performing a t-test.

Question 5: Now, perform the statistical test and report your result. 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 result match up to your expectations?

(Note: I used StatCrunch to perform this test.)

Difference	Sample Diff.	Std. Err.	DF	T-Stat	P-value
Incongruent - Congruent	7.9647917	0.99302863	23	8.0207069	<0.0001

I chose a confidence level of 95% (corresponding to a level of significance of 5%), and at this level of significance the critical value is 1.714 (also from StatCrunch). Since the t-stat of 8.021 is *much* greater than the critical value (and the p-value is far lower than the level of significance), we would strongly reject the null hypothesis. In other words, we have evidence to believe that it takes the average subject longer to read colors from an incongruent list than from a congruent list. This result matched up with my expectations, because it matches up with my personal experience with the Stroop task (via the Fit Brains app on iOS) and with my own expectations about processing time. (I imagined it might take longer for people to process the colors from an incongruent list than from a congruent list, simply because you could read out the words form an congruent list without any thought.)

Question 6 (optional), part 1: What do you think is responsible for this effect?

Before I looked into the specifics, I suspected the effect might be down to different processing times. It seemed to me that it might be easier for people to simply *read* than to *interpret*. One of the four theories used to explain the Stroop effect – parallel distributed processing – suggests that our brains develop specific "pathways … for different tasks," and that "some pathways, such as reading, are stronger than others…" [2][3][4] This corresponds roughly to what I was thinking when I saw the results in this dataset.

Question 6 (optional), part 2: Can you think of an alternative or similar task that would result in a similar effect?

A similar task that could result in a similar effect could involve presenting subjects with two different sets of animals (or cars, every day objects, et cetera). In the first set, the images would have the actual names of the animals above them, but in the second set the images would have the names of other animals above them. For example, an image in the first set could be of a dog, and its label would read "DOG." The same image in the second set could have a label that read "HORSE," or

something like that. The subject would be asked to read out the name of the animal displayed. It might be easier to recognize an image and read out its description than to differentiate between the image and its false description.

References

1. www.statcrunch.com
2. https://en.wikipedia.org/wiki/Stroop_effect
3. http://psycnet.apa.org/?&fa=main.doiLanding&doi=10.1037/0033-295X.97.3.332
4. http://www.wisegeek.org/what-is-the-stroop-effect.htm