Nick Shaw
Oct 22, 2015
Udacity Nano Degree P1

1. **What is our independent variable? What is our dependent variable?**
   The independent variable is font color that the word for a color is written.
   The dependent variable is the amount of time it takes to identify a color from a word whose font can be in different 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.**

   Population: Based on the experiments description of the sample, the population in this experiment is people that can read, are not color blind and that have access to the internet.

   μ_c: The mean time for the population to read a congruent word.

   μ_I: The mean time for the population to read an incongruent word.

   Hypotheses:
   H_0: The mean time for the population to read a congruent word is equal to the mean time to read an incongruent word. ($\mu_c = \mu_I$)

   H_A: The average time for the population to read a congruent word is less than the average time to read an incongruent word. ($\mu_c \neq \mu_I$)

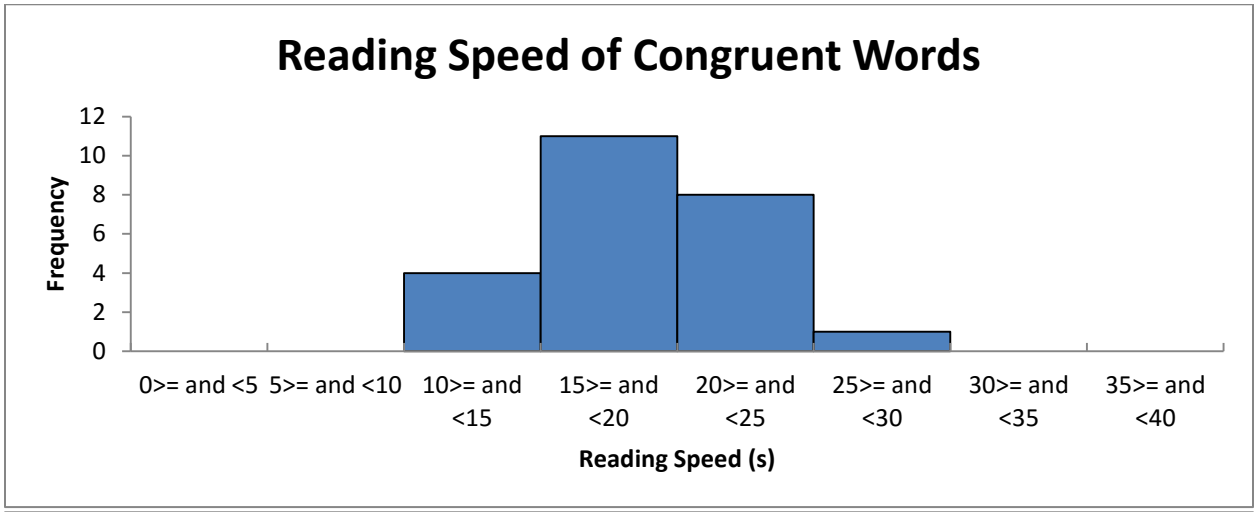
   I expect to perform a dependent t-test (with 2 tails) for paired samples.

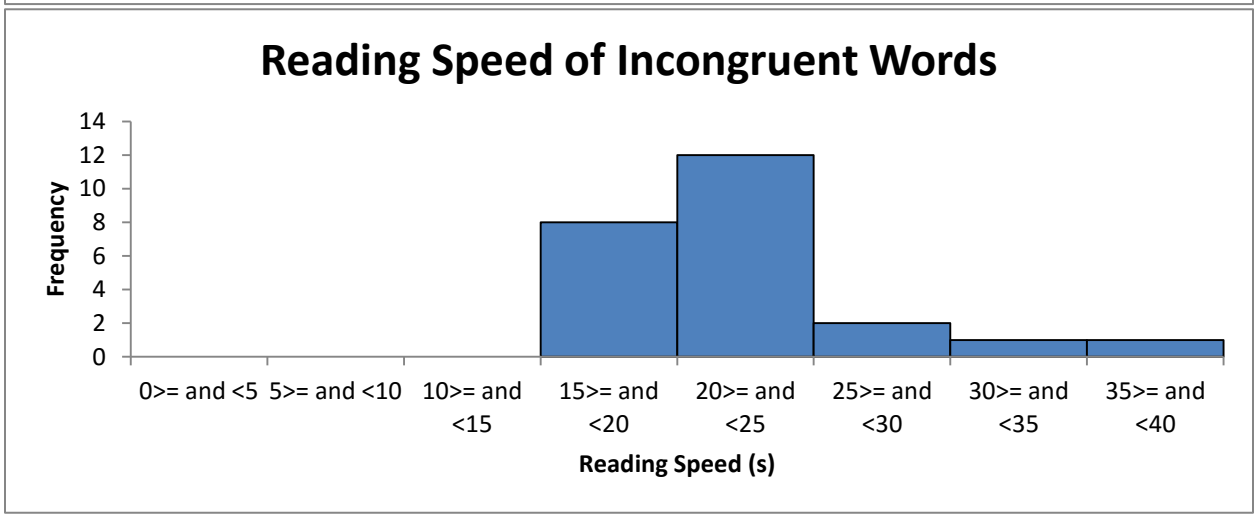
   There are 2 related groups (the 2 groups have the same subjects) so a t-test is valid. The t-test tells if the means of the 2 groups shows any statistical difference which will either prove or disprove the null hypothesis. 2 tails because I am not sure whether people are faster or slower when reading incongruent words.

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

	Congruent	**Incongruent**
mean	14.051125	22.01591667
median	14.3565	21.0175
min	8.63	15.687
max	22.328	35.255
variance	12.66902907	23.01175704
standard deviation	3.559357958	4.797057122

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.**





Most people are between 2 and 15 seconds slower reading the incongruent words. No one is faster reading incongruent words then congruent words and most people are only 15s slower.

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?**

I picked 95% for a confidence level.

The following was generated in MS Excel.

t-Test: Paired Two Sample for Means		
	Congruent	*Incongruent*
Mean	14.051125	22.01591667
Variance	12.66902907	23.01175704
Observations	24	24
Pearson Correlation	0.351819527	
Hypothesized Mean Difference	0	
df	23	
t Stat	-8.020706944	
P(T<=t) one-tail	2.0515E-08	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	4.103E-08	
t Critical two-tail	2.06865761	

The difference between the means of the congruent and incongruent reading times is ~8.02 which is greater than the critical statistic (for 2 tails) of ~2.07. Therefore, the null hypothesis is rejected.

My conclusion based on the experiment is that, reading incongruent words takes more time then reading congruent words.

This matched my expectations. I was thinking that the difference (color and word) caused some conflicts in your brain and slowed the response.

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!**

   From the Wikipedia article (https://en.wikipedia.org/wiki/Stroop_effect) there are a few theories. I like the automaticity theory the best, which basically says we are faster with words because we have trained are brains to quickly recognize and read words.

   Something that might give the same result is printing the names of simple objects on pictures of objects and ask subjects to say the name of the objects. Then (similar to changing the color), try again with the names printed on the object that are different then the object and ask them to say the name of the object (not the word printed).