Udacity Data Analyst Nanodegree

P1: Test a Perceptual Phenomenon

Author: David Venturi

Date: April 22, 2016

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.

Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

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

    Independent variable: the words condition (congruent words or incongruent words)

    Dependent variable: the time it takes to name the ink colors in equally-sized lists

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 should be that the mean time for colour recognition for congruent words is equal to or greater than the mean time for incongruent words, therefore implying a one-tailed test. The alternative hypothesis should be that the congruent words mean is less than the incongruent words mean.

    H_0: $\mu_C \geq \mu_I$

    H_A: $\mu_C < \mu_I$

    where μ is a population mean, the subscript "C" represents the congruent words condition, and the subscript "I" represents the incongruent words condition.

    A one-tailed, dependent samples t-test comparing the difference in means (the time to name the ink colors for congruent words and incongruent words) should be performed. With this test, we seek to determine whether there is enough evidence in the provided sample of data to infer that the

congruent words mean colour recognition time is less than the incongruent words mean colour recognition time for the entire population and not just the sample data.

A t-test is appropriate because the population variance is unknown and the sample size is less than 30. When the sample size is less than 30, the sample data no longer approximate a normal distribution, which makes the use of a Z-value inappropriate.[1] (http://stattrek.com/probability-distributions/t-distribution.aspx) The following assumptions are required for t-tests for dependent means:[2] (http://www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/assump.htm)

- Interval or ratio scale of measurement (approximately interval)
- Random sampling from a defined population
- Samples or sets of data used to produce the difference scores are linked in the population through repeated measurement, natural association, or matching
- Scores are normally distributed in the population; difference scores are normally distributed

A one-tailed test is appropriate under the assumption that incongruent word conditions will not improve recognition times, which is intuitive. The one-tailed test allows for a more scrutinous examination of the negative impact of incongruent word conditions on recognition times.

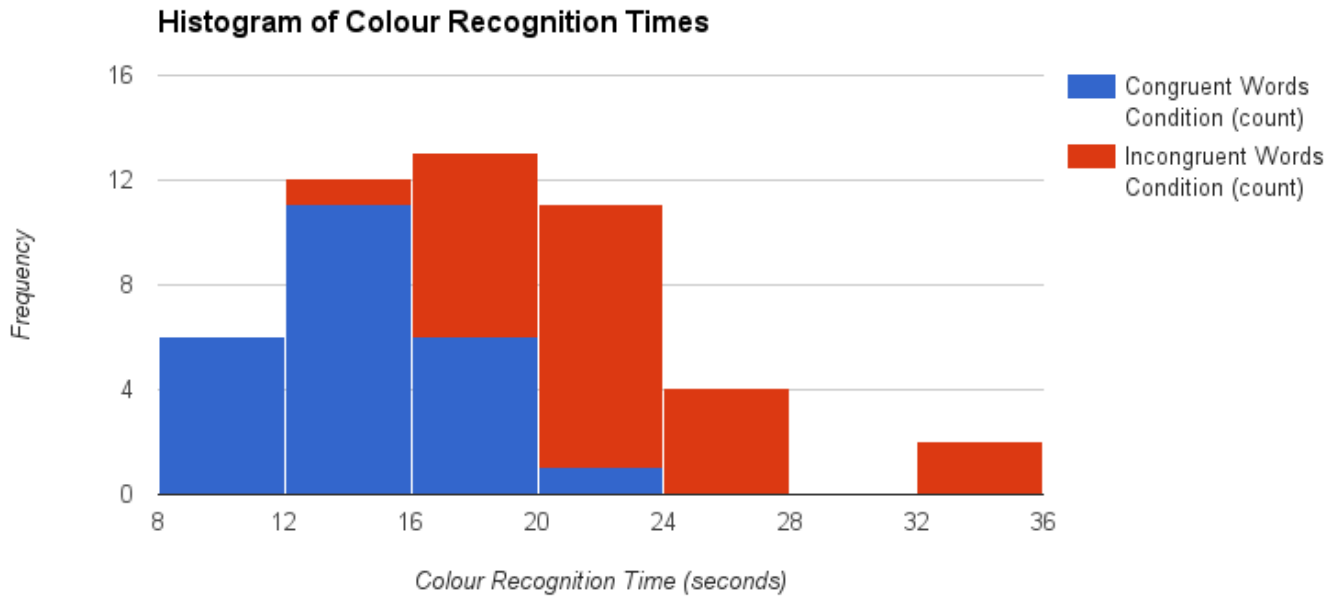
The t-test should be of the dependent samples variety because the same subject is exposed to two conditions and tested for each, which are the defining criteria for "within-subjects" or "repeated-measures" statistical tests.[3] (https://statistics.laerd.com/statistical-guides/dependent-t-test-statistical-guide.php)

Now it's your chance to try out the Stroop task for yourself. Go to this link (https://faculty.washington.edu/chudler/java/ready.html), which has a Java-based applet for performing the Stroop task. Record the times that you received on the task (you do not need to submit your times to the site.) Now, download this dataset (https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view) which contains results from a number of participants in the task. Each row of the dataset contains the performance for one participant, with the first number their results on the congruent task and the second number their performance on the incongruent task.

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

Statistic	**Congruent**	**Incongruent**	**Difference (C-I)**
n	24	24	24
$\bar{x}$	14.05	22.02	-7.96
Median	14.36	21.02	-7.67
s^2	12.67	23.01	23.67
s	3.56	4.80	4.86
SE	0.73	0.98	0.99

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.



The bucket range of the congruent words condition data is 8 seconds to 24 seconds. The bucket range of the incongruent words condition data is 12 seconds to 36 seconds, i.e., noticeably to the right of the congruent words condition data.

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?

$\alpha = .01$

$df = 23$

$t_{crit} = -2.50$

$t = -8.02$

p-value = < .0001

At the 99% confidence level ($\alpha = .01$) and 23 degrees of freedom, the critical statistic value for a one-tailed test in the negative direction is -2.5. The calculated t-statistic for the difference in colour recognition time means of the congruent and incongruent word data is -8.02. Since the t-statistic is in the critical region, the null hypothesis is rejected. With the data presented, it is very unlikely that the 7.96 second difference in mean time for colour recognition for the congruent data vs. the incongruent data is obtained if the two means are actually the same (or if $\mu_C > \mu_I$). By conventional criteria, this difference is considered to be extremely statistically significant.

There is sufficient evidence at the $\alpha = .01$ level of significance to support the claim that it takes less time to recognize the colour of words with the congruent condition compared to words with the incongruent condition.

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!

My hypothesis for the effects observed is that the brain dominantly focuses on reading the word rather than recognizing a colour when the eyes are presented with a coloured word. To recognize a colour, one has to override the brain's natural tendency of reading the word. This override takes time and is likely not always successful, which means re-analyzing a word after the error is recognized, which costs more time.

Numerical/Physical size Stroop tasks, where numerical values and physical size are the factors that contribute to congruency/incongruency, results in a similar effect. It takes longer to recognize the number and physical size (two separate tasks) of small numbers that have a large physical size and large numbers that have a small physical size.[4] (https://en.wikipedia.org/wiki/Numerical_Stroop_effect)

Sources

1. Stat Trek: Student's t Distribution (http://stattrek.com/probability-distributions/t-distribution.aspx)
2. Emory Psychology: Test Assumptions (http://www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/assump.htm)
3. Laerd Statistics: Dependent T-Test for Paired Samples (https://statistics.laerd.com/statistical-guides/dependent-t-test-statistical-guide.php)
4. Wikipedia: Numerical Stroop effect (https://en.wikipedia.org/wiki/Numerical_Stroop_effect)