

Statistics: The Science of Decisions

Project Instructions

Andri Rizhakov
8/17/2015

1. Question 1

The independent variable is the word condition (congruent words condition, or incongruent words condition). The dependent variable is time it takes to name the ink colors. This is because the time to name the color is subject to some preconditioning, in this case the word condition. In other words, in terms of experiment setup or from a mathematical function perspective, the word condition is the input, and would measure the cause of some phenomena. On the other hand, the time to name the color is the output of the experiment (or function), and would measure the effect of the phenomena. Because the output depends on the input, the time it takes to name the ink colors depends on word condition.

2. Question 2

The experiment involves 2 conditions (congruent words condition, or incongruent words condition). The objective is to understand if there is a statistically significant difference between the conditions, i.e., will the time it takes to name the ink colors be statistically different between the congruent words condition versus the incongruent words condition. Therefore, the null hypothesis is that there is no statistical difference between the population mean of the times, and the alternative hypothesis that there is a statistical difference between the population mean of the times. More succinctly:

Appropriate set of hypotheses:

- $H_0: \mu_0 - \mu_A = 0$
- $H_A: \mu_0 - \mu_A \neq 0$

where: H_0 is null hypothesis,

H_A is alternative hypothesis,

μ_0 is population mean of the time in congruent words condition, and

μ_A is population mean of the time in incongruent words condition

The expected statistical test to perform is the two-sided dependent Student's t-test.

- The justification for use of Student's t-test is because the population mean and population standard deviation are not likely to be known. This is likely to be true because in the Problem Statement, there are descriptions of participant tasks, indicating that this

is some sort of controlled study. Furthermore, sample size of each condition is likely to be below 30, a further justification for use of Student's t-test. Even if the sample size is larger than 30, the wider distribution of the t-statistic compared to the z-statistic would give a more conservative answer when rejecting H_0 or failing to reject H_0 . The conservatism comes from the wider distribution: to get the same p-value, one has to have a much higher critical value for the statistic.

- The justification for use of the dependent test is because the samples are paired. Specifically, each sample time for the congruent words condition is a match for a sample time for the incongruent words condition because both values correspond to a study participant that took part in both portions of the experiment sequentially, starting first with the congruent words condition and then proceeding to the incongruent words condition. This comparison of congruent versus incongruent words conditions allows each participant to be their own control [paraphrasing Ref. 8 below] because all other variables for the participant are the same (assuming the time elapsed to do the first congruent words condition and any other effects associated with participating in this sequentially ordered study are minor). Since the only variable that is changing is the words condition, this increases the statistical power of the study and decreases the chance for Type 1 or Type 2 errors.

3. Question 3

The "stroopdata.csv" file was opened with Google Sheets, and analyzed for descriptive statistics. The file use in the analysis is included in the submission folder. Note that because hypothesis from Question 2 response above are formulated as the difference of means from congruent and incongruent words conditions, the statistics and parameters that are reported below will be the result of a the difference distribution of the two samples; in other words, congruent times are subtracted from the incongruent terms to obtain the difference distribution, which is used in the analysis rather than the raw congruent and incongruent condition times.

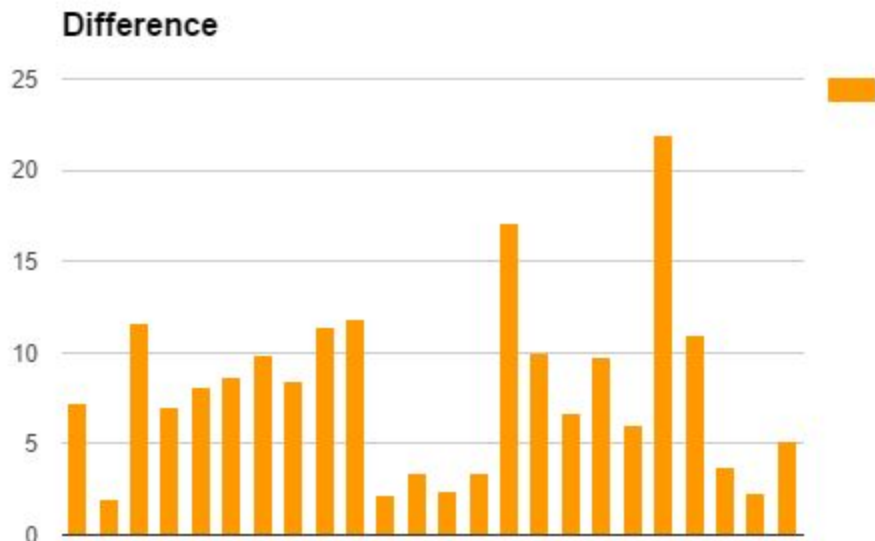
Some measures of central tendency are mean ($M_{\text{difference}} = 7.96479$) and median ($median_{\text{difference}} = 7.66650$). Note that mode is not very helpful because all time values in each sample are unique and thus do not occur more than once.

Some measures of variability are biased sample standard deviation ($s_{\text{biased,difference}} = 4.76240$) and unbiased sample standard deviation ($s_{\text{unbiased,difference}} = 4.86483$). The unbiased sample standard deviation was obtained by applying Bessel's correction.

Note that 6 significant digits were used in the reporting of the descriptive statistics because 5 significant digits were used for each value in the data set. Therefore, to eliminate rounding error, an extra significant digit was carried through in all calculated quantities.

4. Question 4

A histogram was chosen to visualize the data and obtain an intuitive understanding of the data distribution. Congruent and Incongruent data sets were reduced to a “Difference” set by subtracting the Congruent set from the Incongruent one, as discussed earlier in Question 3 response..



The main observations about the data sets include:

- 1) All participants had an increase in time when presented with the Incongruent condition, because when Congruent times were subtracted from Incongruent times the result value was positive.
- 2) The majority of the times appear to be somewhere between 5 and 12, with a handful in the below 5 range, and a few outliers (in the colloquial sense, not statistical; needs to be determined what the true statistical outliers are) at 17 and 22.

5. Question 5

A two-sided Student's t-test was performed to understand if there is a statistical difference between the population means of the 2 conditions of the words.

At a $p = 0.05$ and $df = 23$, the critical t -value for the two-tailed test is 2.069 (from t-table, see References). The t -statistic is calculated on the differenced values of the two conditions:

$$t = (\mu_0 - \mu_A)/(s/\sqrt{n}) = 7.96479/(4.86483/\sqrt{24}) = 8.02071$$

The 95% confidence level for the difference in population means of the time in congruent words condition and the time in incongruent words condition is:

$$\begin{aligned} & (M_{\text{difference}} - t_{0.05}^* s/\sqrt{n}, M_{\text{difference}} + t_{0.05}^* s/\sqrt{n}) \\ &= (7.96479 - 2.069 \cdot 4.86483/\sqrt{24}, 7.96479 + 2.069 \cdot 4.86483/\sqrt{24}) \\ &= (5.91022, 10.0194) \end{aligned}$$

Using the GraphPad Software (see References), the two-tailed P value is less than 0.0001.

In conclusion, the null hypothesis, H_0 , can be rejected because:

- 1) the t -statistic is greater than the critical t -value, the observed event is statistically unlikely to occur from pure randomness in population alone.
[$t(df = 23) = 8.02071, p = 0.05, \text{two-tailed}$].
- 2) actual $p < 0.0001$, again, suggesting that the observed event is statistically unlikely to occur from pure randomness in population alone.
- 3) 95% confidence level for the difference in population means of the condition times is (5.91022, 10.0194), suggesting that there is 95% probability that the true value of the population mean of the differences falls within this interval. Because both confidence interval bounds are greater than 0, this suggests that the difference between the times is non-zero and, furthermore, a positive value.

An interpretation of the results of the analysis can be that there appears to be an increase in time it takes to name the ink colors in the incongruent words condition when followed by congruent words condition. These results match up with my expectations because taking the test first hand made me aware of the increased difficulty in naming the ink color, and thus increasing the time.

6. Question 6

From preliminary online research, there appears a few theories regarding the mechanisms for the effects observed. Neuroanatomical and other experimental psychology tests indicate that additional conflicting data is introduced and the brain struggles with the processing. The struggle comes from processing speed, attention issues, and other neurological circuit issues make the task more complicated, and thus longer to get correct.

There exist similar variations on the Stroop test. For example, instead of the color change that introduces complexity, the words can be warped spatially to study a different aspect of complexity. Additionally, the emotional Stroop test investigates emotional processing in diagnosing depressed individuals.

7. References

References used in answering above questions:

1. Notes from Udacity's Statistics course.
2. Statistics course, problem statement.
3. https://en.wikipedia.org/wiki/Dependent_and_independent_variables.
4. <https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg>.
5. <http://www.graphpad.com/quickcalcs/>.
6. https://en.wikipedia.org/wiki/Confidence_interval.
7. https://en.wikipedia.org/wiki/Stroop_effect.
8. https://en.wikipedia.org/wiki/Student%27s_t-test#Unpaired_and_paired_two-sample_t-tests.

8. Calculations

Calculations were performed in the following files. These are attached to the report submission folder.

1. StroopEffect_data (Google sheet used in analysis). StroopEffect_data.pdf for convenience also included.
2. stroopdata.csv (input)

9. Appendix A. Reviewer Comments and Disposition

Reviewer Comment	Disposition in latest version
<p>Question 2: Establish a hypothesis and statistical test</p> <p>SPECIFICATION</p> <p>An appropriate hypothesis test has been stated along with an appropriate statistical test to apply to collected data, with appropriate justification.</p> <p>DOES NOT MEET SPECIFICATION</p>	<p>Clarification is added in Question 2 section, in the last paragraphs, discussing the expected statistical test.</p>

Reviewer Comments

Almost there! You have correctly declared the set of hypotheses in formal terms pointing out that they are referencing the population means - this is the most common reason to get this section marked as not *meeting specifications* - well done. Also, from the results of Question 5, it is clear you have conducted the correct T-test (dependent). The area that requires more explanation is in the justification of the particular T-test the project is performing. Currently, the justification can be used for both types of T-tests (dependent and independent), but there is a very critical reason why the Dependent T-test is the correct one to choose and the Independent T-test is not. I have attached a link below that should help. So far, this is an amazing first attempt at the most difficult area of this project.

https://en.wikipedia.org/wiki/Student%27s_t-test#Unpaired_and_paired_two-sample_t-tests