Project 1: Statistics - Stroop Effect

Luc Frachon

27 septembre 2016

Abstract

This study provides statistical evidence for the Stroop effect. We find that within 95% confidence, the time taken to name the colours is 5.88 to 10.05 seconds greater in the incongruent condition than in the congruent one. The difference is highly statistically significant and its effect size is over 1.6 standard deviation.

Introduction

This study is performed as part of Project 1 in the Udacity Data Analyst Nanodegree. We are tring to replicate the Stroop effect by subjecting participants to a test and infering the results from a statistical point of view.

This effect was named after John Ridley Stroop, who researched it in his 1935 paper[1] (athough it seems that the effect had previously been discovered and researched earlier in Germany[2]).

Experimental design

A sample of 24 participants (supposed to have been randomly selected from the population of interest) is asked to name a set of colours in two conditions:

- *Congruent words:* Colour names are written in the colour that they represent - e.g. *RED*, *BLUE* etc.
- *Incongruent words:* Colour names are printed in colours that do not match them - e.g. *YELLOW*, *WHITE* etc.

The time taken by each participant to name each set of colours is recorded.

Description of statistical test

Variables

- **Independent variable:** Congruent or incongruent condition
- **Dependent variable:** Time required to name each set of colours (completion time)

Assumptions on variables

This statistical test makes certain assumptions:

- The variable of interest (completion time) is a ratio variable: There is a meaningful zero value and a ratio of 2 between two measurements means twice as much time to complete the assignment.

- As stated above, we assume that the sampling was performed randomly from the population of interest. This is an essential condition to be able to infer conclusions on the population from the statistical tests performed on the sample.
- The difference scores (introduced in section Statistical Test - New dependent variable) are linked through repeated measurement.
- We assume that scores and difference scores are normally distributed in the population.

Test type

Since we only have one sample that goes through two different experiments, we are running a **dependent t-test with paired samples (two conditions)**.

Convention

Let us define the following symbols:

- $\mu^{(C)}$: *population mean, congruent condition*
- $\mu^{(I)}$: *population mean, incongruent condition*

More generally, the (C) superscript refers to the congruent condition while (I) refers to incongruent. Thus, we also define:

- $X_i^{(C)}$: i^{th} observation in the congruent condition
- $X_i^{(I)}$: i^{th} observation in the incongruent condition

Hypotheses and alpha level

The Stroop effect is a well documented phenomenon; we therefore expect the average completion time in the incongruent situation to be longer than in the congruent condition. We therefore run a directional test:

$$H_0 : \mu^{(I)} \leq \mu^{(C)}$$
$$H_A : \mu^{(I)} > \mu^{(C)}$$

Given the fairly small sample sizes, we do not want to be too agressive when setting the alpha level for significance. We choose to set $\alpha = .05$.

Descriptive statistics

The full dataset for both conditions (and the difference in task completion times for each individual) are provided in the appendix. Both experiments use the same sample, of size $n = 24$.

- **Sample means:**
  - $\bar{X}^{(C)} = 14.05$
  - $\bar{X}^{(I)} = 22.02$
- **Sample standard deviations:**
  - $s^{(C)} = 3.56$

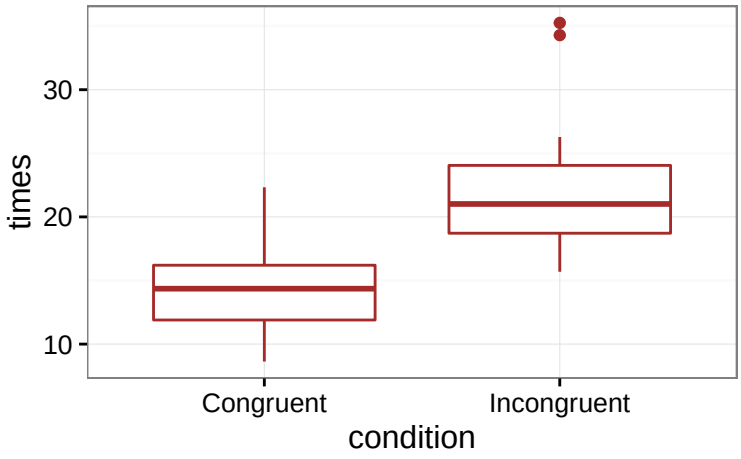

Figure 1: Boxplot of sample distributions for Congruent and Incongruent conditions

- $s^{(I)} = 4.8$

- **Sample distributions:**

From the distributions in figure 1, it seems that for this sample, completion times in the incongruent situation are indeed longer. The effect size seems considerable, although the two distribution overlap. The statistical test should help us assess whether this might be due to chance. There also seems to be more variability in the incongruent times.

Statistical t-test

New dependent variable

We first define a new dependent variable Δ as the difference in completion times between the congruent and incongruent conditions:

$$\Delta_i = X_i^{(I)} - X_i^{(C)}$$

for each $i \in \{1, ..., n\}$

We get the distribution shown in figure 2. Its statistics are:

- **Sample mean:**
  - $\bar{\Delta} = 7.96$
- **Sample standard deviation:**
  - $s_\Delta = 4.86$

New test hypotheses:

As a result of introducing this new variable, our test hypotheses can be re-written as follows:

$$H_0 : \mu_\Delta \leq 0$$
$$H_A : \mu_\Delta > 0$$

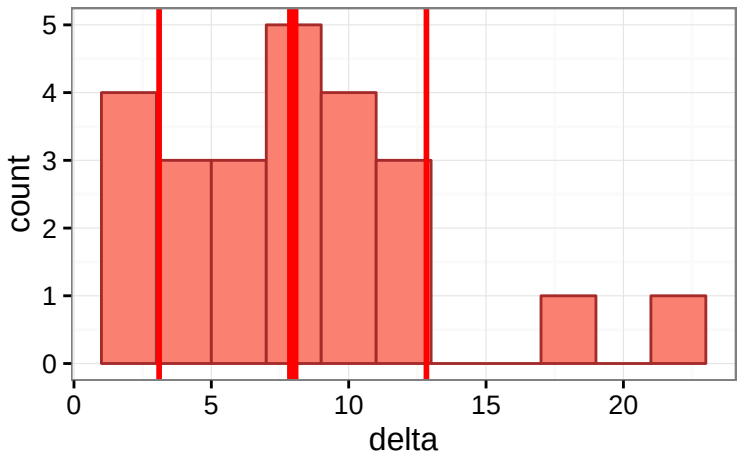

Figure 2: Distribution of the difference in completion times between congruent and incongruent conditions (bin size = 2s). The red vertical lines indicate the sample mean and sample mean +/- standard deviation.

Test results:

Estimation of significance

Our test has one sample of size $n = 24$, therefore it has $df = 23$ degrees of freedom. From the t-table, we get the following critical value for t at $\alpha = .05$:

$$t_{.05} = 1.714$$

The standard error of the mean is given by:

$$S_{\bar{\Delta}} = \frac{s_{\Delta}}{\sqrt{n}}$$

And we find that $S_{\bar{\Delta}} \simeq 0.99$.

We can now calculate the t-statistic with $df = 23$:

$$t = \frac{\bar{\Delta}}{S_{\bar{\Delta}}}$$

Thus $t(23) \simeq 8.02$.

$t > t_{.05}$ and $p < .0001$ which is extremely significant, so **we reject** H_0.

The confidence interval for the difference in completion time is given by the following equation:

$$CI = \bar{\Delta} \pm t_{.025}.S_{\bar{\Delta}}$$

Thus: $CI =$ (5.88, 10.05)

In other words, we have 95% confidence that in the incongruent condition, the population of interest will take 5.88 to 10.05 seconds more to complete the task.

Effect size

We calculate Cohen's d to assess the magnitude of the Stroop effect on this sample:

$$d = \frac{\bar{\Delta}}{s_\Delta}$$

Thus $d \simeq 1.64$. This can be considered as a very large effect size, over 1.6 times the sample standard deviation.

Discussion: Possible causes of the Stroop effect

Several theories have been used to explain the Stroop effect. The most widely accepted cause is given by the "Automaticity model" theory[3].
According to this theory, the adult brain is "wired" to recognise words automatically without much effort. This process cannot be turned off. Conversely, recognising colours requires more effort. As the brain automatically reads words, the participant needs to produce a conscious effort not to say the first word that comes to mind but to look for the name of the hue they recognised. This process increases the time required to complete the task.

This theory seems to fit with our intuitive perception of what is occuring when completing the colour recognition task. It relies on theories about brain plasticity and its ability to create "pathways" (synaptical connections) that "automate" some tasks, ie. allows them to be processed with little or no conscious effort. However if the theory is correct, then given enough practice, the brain should also be able to create connections to automate colour recognition even in the presence incongruent signals. Since this wiring pattern is new, one would expect larger improvement in response times in the incongruent condition than in the congruent one (where the pathways are already in place).

This is precisely what Stroop's paper[1] finds, with an overall improvement of 33.9% for incongruent signals v. 13.9% in the congruent condition. This tends to lend this model credibility. Of course, competing models would need to be evaluated as well against different criteria.

Conclusion

This test establishes with a very high level of significance that the Stroop effect observed in this sample is very unlikely to be due to chance. The effect size is also very large. We do not know the details of the experimental conditions or if the order in which participant took either test was set at random. However assuming that this is the case and that all other factors were held constant, the statistical results allow us to conclude that the incongruent condition most likely causes a large increase in completion times.

This test provides no information that could help us discriminate between different possible explanations for the Stroop effect, however we argue that the most commonly accepted cause - the Automaticity Model – is consistent with the findings that John Ridley Stroop mentionned in his original publication.

References

[1]: http://psychclassics.yorku.ca/Stroop/ "John Ridley Stroop (1935); "Studies of interference in serial verbal reactions"; *Journal of Experimental Psychology*, 18, 643-662."
[2]: https://en.wikipedia.org/wiki/Stroop_effect#cite_ref-Jaensch_2-0 "Wikipedia page on the Stroop Effect."
[3]: http://www.rit.edu/cla/gssp400/sbackground.html "Background on the Stroop effect."

Appendix

Data set from experiments:

	Congruent	Incongruent	delta
1	12.1	19	7.2
2	16.8	19	1.9
3	9.6	21	11.6
4	8.6	16	7.1
5	14.7	23	8.1
6	12.2	21	8.6
7	14.7	25	9.9
8	9.0	17	8.4
9	9.4	21	11.4
10	14.5	26	11.8
11	22.3	25	2.2
12	15.3	19	3.3
13	15.1	18	2.4
14	16.9	20	3.4
15	18.2	35	17.1
16	12.1	22	10.0
17	18.5	25	6.6
18	10.6	20	9.8
19	11.3	17	6.1
20	12.4	34	21.9
21	12.9	24	10.9
22	14.2	18	3.7
23	19.7	22	2.3
24	16.0	21	5.2