Test a Perceptual Phenomenon

October 17th, 2017
By
Beltino Goncalves

Background

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.

Analyzing the Stroop Effect

(1) What is the independent variable? What is the dependent variable?

The independent variable is the color of the ink stating the list of words. This can either be congruent where the color of ink matches the word and incongruent, where the color does not mach the word. The dependent variable is the measured time that it takes for the participant 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.

Null Hypothesis: on average the population mean of the time that it takes to name the ink color in the congruent list should be greater than or equal to the population mean of the time that it take the name the ink color of the incongruent lists.

Alternative Hypothesis: on average the population mean of the time that it takes to name the ink color in the congruent list should be less than the population mean of the time that it takes to name the ink color of the incongruent list.

Mathematical Notation:

$$H0 : \mu \geq \mu0$$
$$HA : \mu < \mu0$$

Denotes,

μ = the population mean of the time that it takes to name the ink color in the congruent list.
$\mu0$ =the population mean of the time that it takes to name the ink color in the incongruent list.

Assumptions,

- **(1)** The population mean is Normality Distributed
- **(2)** The sample data can be estimate population variances.
- **(3)** The variances in samples are different

A significant test of $\alpha = .05$ will be tested on a one-tailed test with test statistic of **t** which will allows us identify the p-value to determine if we reject or accept the null hypothesis.

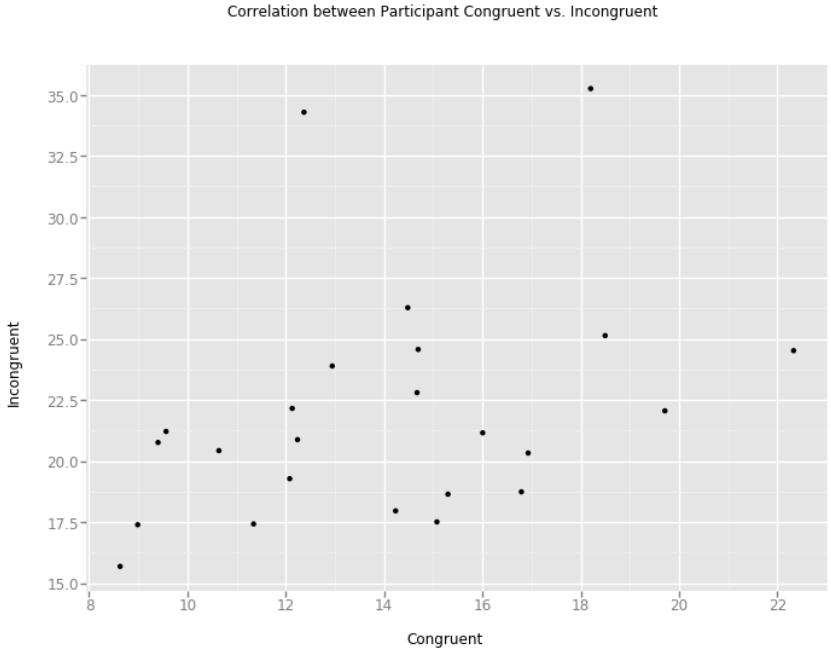
One-tailed test : based on the mathematical notation of $\mu \geq \mu_0$, one sided tailed test will be appropriate.

Test Statistic of t : since the sample size n ≥ 30 or smaller, t test would be appropriate to identify rejects the null hypothesis when $\mathbf{t} < -t_{\alpha,n-1}$ and accepts the null hypothesis when $\mathbf{t} \geq -t_{\alpha,n-1}$ [1].

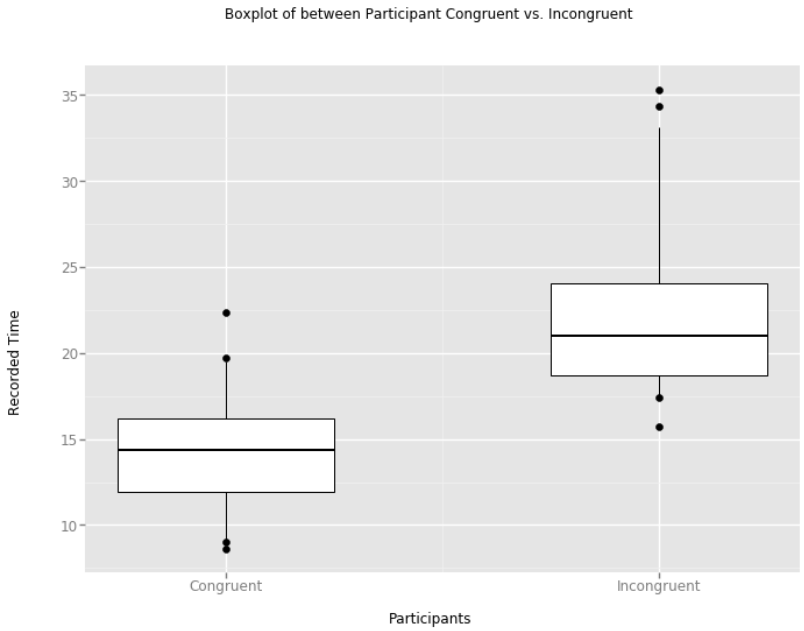
(3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability. The name of the data file is 'stroopdata.csv'.

	Congruent	Incongruent
count	24.000000	24.000000
mean	14.051125	22.015917
std	3.559358	4.797057
min	8.630000	15.687000
25%	11.895250	18.716750
50%	14.356500	21.017500
75%	16.200750	24.051500
max	22.328000	35.255000

(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.



This scatter plot relationship shows us the correlation between congruent and incongruent participant with an positive correlation.



This box plot shows the time differences, denoting that incongruent measurement was greater than congruent.

(5) Now, perform the statistical test and report the results. What is the 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?

Statistical t-Test (Computeded in R)

```
        Paired t-test

        data:  df$Incongruent and df$Congruent
        t = 8.0207, df = 23, p-value = 4.103e-08
        alternative hypothesis: true difference in means is not equal t
        o 0
        95 percent confidence interval:
        5.910555 10.019028
        sample estimates:
        mean of the differences
        7.964792
```

Statistical t-Test (Computeded in Python)

```
# Perform the statistical test here

from scipy import stats

one_paired=stats.ttest_rel(stroop['Incongruent'],stroop['Congruent'])
one_paired
```

```
Ttest_relResult(statistic=8.020706944109957, pvalue=4.1030005857111781e-08)
```

Ttest_relResult(statistic=8.020706944109957,
pvalue=4.1030005857111781e-08)

Used R & Python, to cross check each other.

We reject the null hypothesis based on the p-value: 4.103e -08 < then our denoted α = .05. Our one tailed t-critical value is 1.741[1] at α = .05. and df = 23. Therefore our t-statistic is (8.0207) > t-critical (1.741), stating that we should reject the null hypothesis.

As expected, there is statistically significance that the population mean of the time that it takes to name the ink color in the congruent list is less than the population mean of the time that it takes the name the ink color of the incongruent lists.

(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!

This information was gather from situation [2]

There are several relevant and irrelevant information that must be taken in to consideration:

1. **Processing speed**: This theory suggests there is a lag in the brain's ability to recognize the color of the word since the brain reads words faster than it recognizes colors. This is based on the idea that word processing is significantly faster than color processing.
2. **Selective attention**: Theory that color recognition as opposed to reading a word, requires more attention, the brain needs to use more attention to recognize a color than to word encoding, so it takes a little longer.
3. **Automacity**: It suggests that since recognizing colors is not an "automatic process" there is hesitancy to respond; whereas, the brain automatically understands the meaning of words as a result of habitual reading. This idea is based on the premise that automatic reading does not need controlled attention, but still uses enough attentional resources to reduce the amount of attention accessible for color information processing
4. **Parallel distributed processing**: This theory suggests that as the brain analyzes information, different and specific pathways are developed for different tasks. Some pathways, such as reading, are stronger than others, therefore, it is the strength of the pathway and not the speed of the pathway that is important. In addition, automaticity is a function of the strength of each pathway, hence, when two pathways are activated simultaneously in the Stroop effect, interference occurs between the stronger (word reading) path and the weaker (color naming) path, more specifically when the pathway that leads to the response is the weaker pathway.
5. **Not knowing the meaning of the word**: This could have a tremendous effect on an individual whom English is a second language, if they do not know how to read the English word but knows the color of the ink being displayed. I kid you not, this would be a great effects with children as they are learning the color of ink not how to read the word. So there speed would remain the same based on only knowns colors.

References

[1] Udacity, “ST095 Course Index” accessed October 2017 https://www.udacity.com/wiki/st095/courseindex

[2] Wikipedia, “Stroop effect” accessed October 2017.’https://en.wikipedia.org/wiki/Stroop_effect