DATA ANALYST

Test a Perceptual Phenomenon

Author:
Vayne Lover

School:
Zhejiang University

March 11, 2017

1 Introduction

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participants 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.

2 Variables

In my opinion,if the word is conggruent or in congruent is important for the time to finish the task,so i hold this point of view:

Independent variable : the word is congruent or in congruent.

Dependent variable : the time to finish the recognization task.

3 Hypotheses

$$\mathbf{H_0} : \mathbf{u_1} = \mathbf{u_2}$$
$$\mathbf{H_1} : \mathbf{u_1} \neq \mathbf{u_2}$$

In the equations $\mathbf{H_0}$ is null hypothesis which means there is no significant difference between the word is congruent or incongruent,and $\mathbf{H_1}$ means alternative hypothesis which means there is significant difference between the word is congruent or incongruent.Besides,$\mathbf{u_1}$ means u(congruent),$\mathbf{u_2}$ means u(incongruent).

I will choose samples t-test because :

–i– the scale of the data is small.

–ii– we dont know the population parameters.

–iii– In this case we test out and measure samples from two different populations.

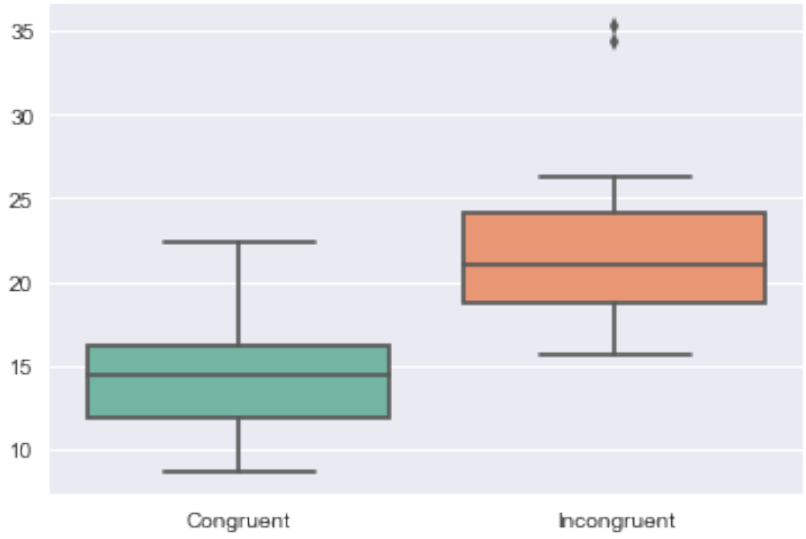
4 Data Explore

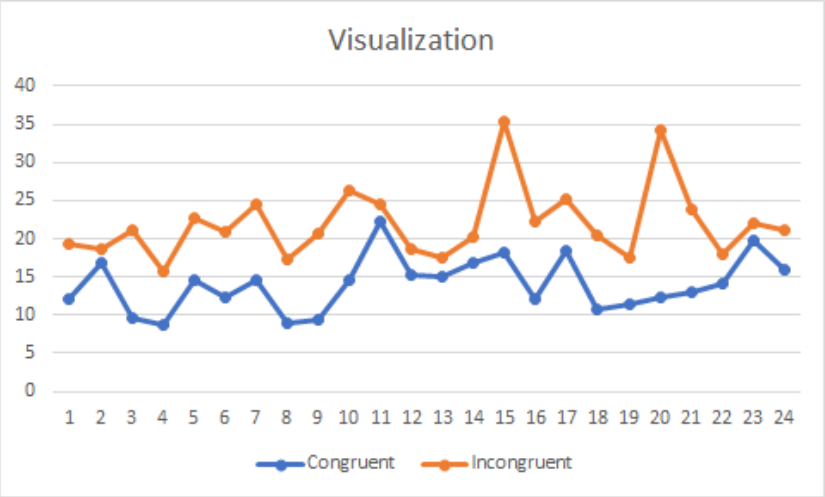
Statistics	Congruent	Incongruent
Min	8.63	15.69
Max	22.33	35.26
Mean	14.05	22.02
Variance	12.67	23.01
Standard Deviation	3.56	4.80

I write some python code to explore the data.The result can be seen above.And we can learn from the table that the minimum and maximum of the Congruent is 8.63 and 22.33 seconds,while they are 15.69 and 35.26 seconds for Incongruent.Besides,the mean of the Congruent is 14.05 s comparing to 22.01 s for Incongruent.What's more,the variance and standard deviation of both are 12.67 s^2,3.56 s and 23.01 s^2,4.80 s.

And wen can know that every statistics of the Incongruent is great than Congruent.

5 Data Visualization





We can learn from figure 1 that Incongruent task have a larger span than Congruent task.Moreover,we know from figure 2 that people usually need more time to finish Incongruent task than Congruent task.

6 Hypothesis Testing

Cause it's samples t-test i need to calculate the $\boldsymbol{d = x - y}$ first.The data can be seen below:

Congruent	Incongruent	d
12.079	19.278	7.199
16.791	18.741	1.950
⋮	⋮	⋮
19.710	22.058	2.348
16.004	21.157	5.153

$$|t| = \frac{\bar{d}}{s_D/\sqrt{n}} = 8.02 > t_{0.025}23 = 2.069$$

We can learn from the T-table that $t_{0.025}23 = 2.069$ which means that t-critical equals to 2.069.And we can use the formula $|t| = \frac{\bar{d}}{s_D/\sqrt{n}}$ to calculate the t-statistic which equals to 8.02.So we can reject the null hypothesis.We can know that there is significant difference between the Congruent task and Incongruent task.

7 Reference

1. https://docs.python.org/3/library/csv.html
2. https://en.wikipedia.org/wiki/Stroop_effect
3. https://en.wikibooks.org/wiki/LaTeX/Mathematics
4. http://seaborn.pydata.org/generated/seaborn.boxplot.html#seaborn.boxplot
5. https://en.wikipedia.org/wiki/Student%27_t-distribution#Table_of_selected_values