UDACITY P1 PROJ

Tuesday, February 21, 2017 4:17 PM

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.

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

Independent variable is: the congruence of word and color.

Dependent variable is: the time which participant taken to finish each experiment.

2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform?

2.1 Hypothesis:

Usually people are familiar with speaking out the word we see, it is a normal habit so our brain processes it very quick so the there is no challenge for people to take the congruent words treatment.

For the incongruent words condition, since brain still works with the usual habit so people might have a little bit difficult to do correctly, it could take more time but not significant if people focus on the challenge.

So my hypothesis is: the time to take the incongruent words condition will be NOT SIGNIFICANTLY DIFFERENT from the other.

μ_c : the mean of time people finish the congruent words condition.

μ_i : the mean of time people finish the incongruent words condition.

H_o: $\mu_c = \mu_i$
H_a: μ_c != μ_i

2.2 Kind of statistical test:

As information provided:

- we don't know the population parameter like mean or standard deviation.
- we are provided the result dataset so we could calculate the sample mean and standard deviation.
---> we will perform T-TEST (two tailed test).

- from the result dataset we could group all participants as a group, and all of them performed the same experiment which has 2 treatments (one is the congruent words, another one is the incongruent words), we collected the result of each person for the comparison.
--> this is DEPENDENT SAMPLE for PAIRED SAMPLES with TWO CONDITIONS.

So I will perform the T-TEST (two tailed test) with the TWO CONDITIONS of DEPENDENT SAMPLES for PAIRED SAMPLES.

3. Some descriptive statistics regarding this dataset.

3.1 Central Tendency: Mean, Median, Mode.

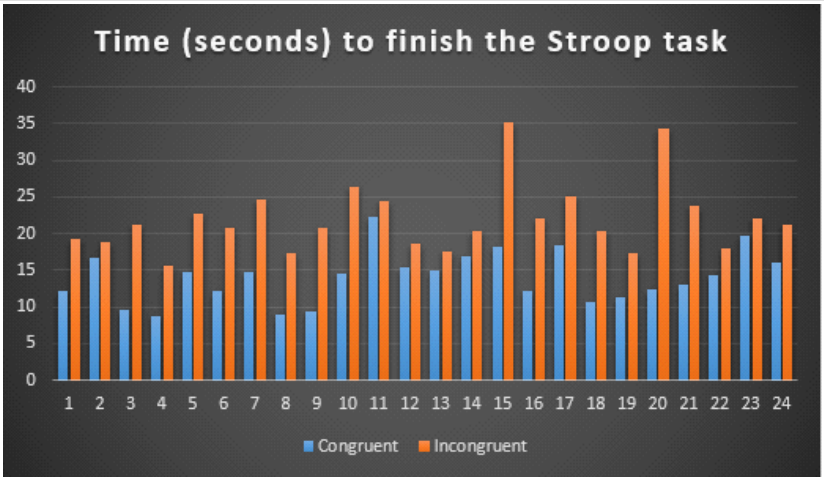
	Congruent	Incongruent
MEAN	14.05	22.02
MEDIAN	14.3565	21.0175
MODE	NA	NA

- Since the values are decimal number --> not found the MODE value.

3.2 Variability: Standard Deviation, Quartiles, IQR, MIN, MAX, Outliers (if available).

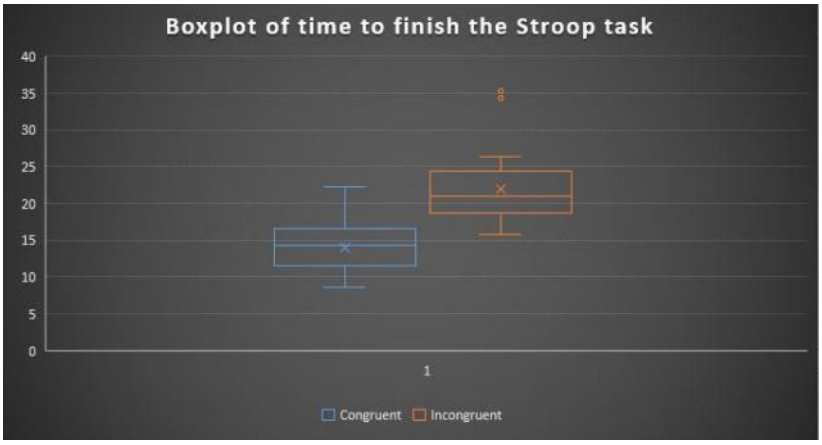
	Congruent	Incongruent
STD	3.56	4.8
1st QUARTILE	11.52775	18.66825
2nd QUARTILE	14.3565	21.0175
3rd QUARTILE	16.59425	24.3665
IQR	5.0665	5.69825
LOWER OUTLIER	3.928	10.120875
UPPER OUTLIER	24.194	32.913875
MIN	8.63	15.687
MAX	22.328	26.282
OUTLIERS	NA	34.288, 35.255

4. Visualizations show the distribution of the sample data.



As the BAR CHART shows, the time taken to finish the Incongruent treatment always greater than the Congruent for each individual person.

Especially there are 2 people took quite long time to finish the Incongruent treatment, they are 15th and 20th, actually they are also the outlier point as I found in the Question 3. They also have the large gap between 2 treatments.



As Boxplot chart show, the boxplot of Incongruent is clearly above the Congruent, it tells that most of responsible values of Incongruent are greater than the Congruent.

Boxplot of Congruent tells that most of the values fall in the interval (11, 17), there is no outliers, data is not much spread out.

Boxplot of Incongruent tells that most of the value fall in the interval (19, 24), there are 2 outliers (shown in Question 3), data is more spread out than the Congruent.

From both charts, seem the time taken to finish Incongruent is significantly greater than the Congruent . I am not fully confident about my hypothesis, but I need to perform the statistical test to ensure that.

5. Performing the statistical test.

Hypothesis: The time taken to finish the INCONGRUENT treatment is NOT SIGNIFICANTLY longer than CONGRUENT.

μ_c : the mean of time people finish the congruent words condition.

μ_i : the mean of time people finish the incongruent words condition.

H_o: $\mu_c = \mu_i$
H_a: μ_c != μ_i

5.1 Calculating t-statistic value:

Using Excel for calculating we will have: μ_c =14.05, μ_i = 22.02, n = 24, S_D =4.86.

then t-statistic = (μ_c - μ_i)/(S_D / sqrt(n)) = (14.05 - 22.02)/(4.86/sqrt(24)) = -8.03.

5.2 Calculating t-critical values at α = 0.01:

Due to the two tailed test --> α/2 = 0.005 and DF = n-1 = 24-1 = 23.

Refer from t-table we have t-critical = +- 2.807.

5.3 Decision: Since 8.03 > 2.807 so we REJECT THE NULL HYPOTHESIS.

5.4 Conclusion: the INCONGRUENT treatment will need significantly longer time to finish than the CONGRUENT treatment. When see a word, the first action of our brain seems to be trying to pronounce it, so it need a significant delayed time if we want to change that habit. This conclusion is different from what I expected.

6.What is responsible for the effects observed? Any an alternative or similar task that would result in a similar effect?

6.1 I think the way human brain working is responsible for the effects. When was a child, we are taught to read. When growing up, every day we see and read words everywhere at any time. When see a word, event our mouth don't speak out but actually our brain pronounce it in silence. It is a habit of human brain, when we want to change the habit, our brain need time to practice and get familiar with new habit.

6.2 Another alternative task could be: when see a word, participants task is saying out the opposite word.