Statistics: The Science of Decisions Project Instructions

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.

Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

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

Dependent variable: Time that participants take to finish saying the color of the words.

~~Independent variable: Number of words, number of colors, font size of the words.~~

Independent variable: The congruent and incongruent words

Reference:
https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4601188734/m-4699459406

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

The set of hypotheses are null hypothesis: $\mu_1=\mu_2$, alternative hypothesis: $\mu_1\neq\mu_2$. Dependent t-test for paired samples and two-tailed test should be performed. If there is no significant difference between two population means, $\mu_1=\mu_2$, this is a null hypotheses. Also, two ~~samples~~ population of the means could be $\mu_1<\mu_2$, or $\mu_1>\mu_2$, the alternative hypothesis is $\mu_1\neq\mu_2$. ~~There are two samples that come from same population by repeating the tests with different conditions and the same person.~~ This task is a repeated measures test, which means it compares the same participant's test results in two separate tests, the results between two tests are dependent, and paired.

Reference:
https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/m-147019329
https://explorable.com/independent-two-sample-t-test?gid=1586
https://explorable.com/dependent-t-test-for-paired-samples?gid=1586
http://stattrek.com/hypothesis-test/difference-in-means.aspx?Tutorial=AP

Now it's your chance to try out the Stroop task for yourself. Go to this link, which has a Java-based applet for performing the Stroop task. Record the times that you received on the task (you do not need to submit your times to the site.) Now, download this dataset which contains results from a number of participants in the task. Each row of the dataset contains the performance for one participant, with the first number their results on the congruent task and the second number their performance on the incongruent task.

3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

Descriptive Statistics		
	Congruent	Incongruent
Mean	14.05	22.02
Mode	12, 15	21
Median	14.36	21.02
Range	13.698	19.568
Interquartile Range (IQR)	4.69	5.52
Variance	12.67	23.01
Standard Deviation	3.56	4.8

Excel sheet:

File: p1_statistics.xlsx

Sheet name: "P1(Q3)"

Reference:

https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4615238546/m-77345075

https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4612348621/e-83934430/m-83664104

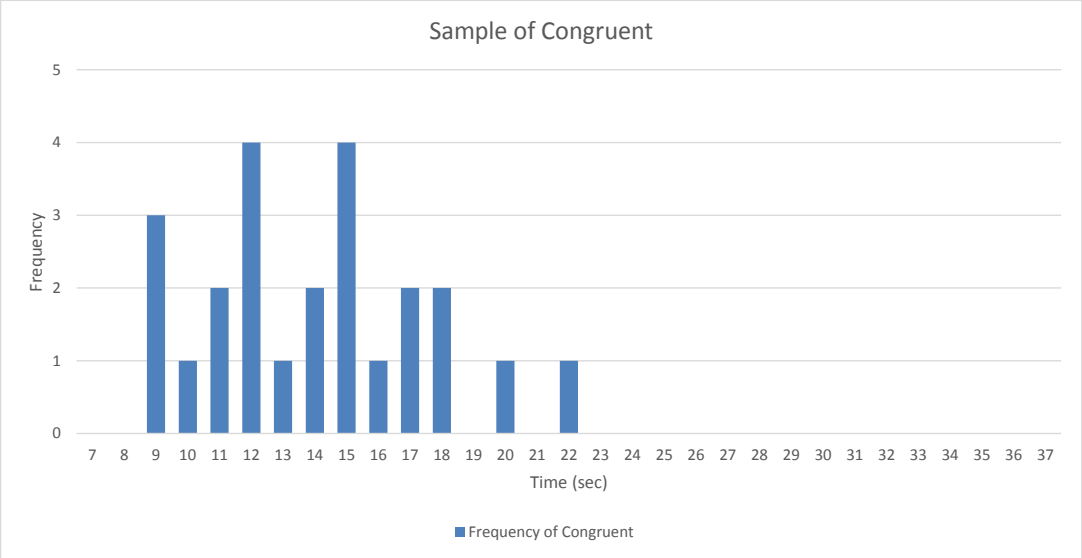
http://webspace.ship.edu/cgboer/descstats.html

http://stattrek.com/descriptive-statistics/variability.aspx?Tutorial=AP

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.

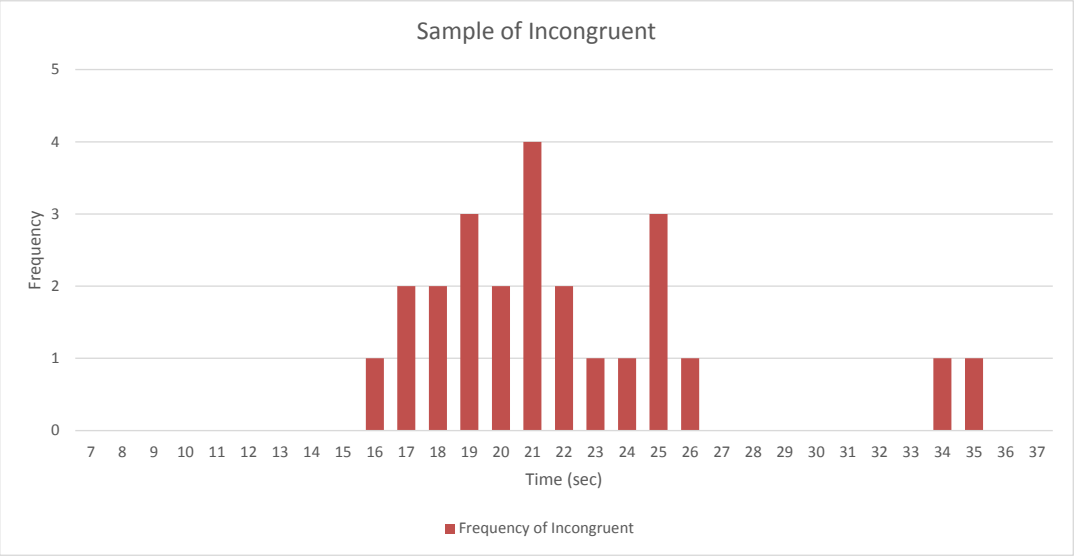
(1) Time distribution of Congruent dataset

The shape of distribution is unimodal and skewed right because I can see the group of three peaks at 9, 12, and 15 seconds.



(2) Time distribution of Incongruent dataset

The shape of distribution is bell shaped because I can see the biggest peak at 21 seconds. The 34 and 35 second records are considered outliers.



Excel sheet:

File: p1_statistics.xlsx

Sheet name: “P1(Q4)”

Reference:

http://mathbitsnotebook.com/Algebra1/StatisticsData/STShapes.html

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

My confidence level is 95% and critical statistic value (α) is 0.05. As the result, the samples rejected the null hypothesis which means two samples of the mean have a significant difference.

The result is t(23)=-8.02, p<.05, two-tailed. The t-critical values are -2.069 and 2.069. The t-statistic value is not in the critical region, and is also located on the negative side. This means the average time of the congruent word condition is significantly faster than the incongruent. Finally, yes, this result matched up with my expectations.

Discritive Statistics

	Congruent	Incongruent
Mean	14.05	22.02

Difference of Means	-7.97	
n	24	
Degrees of freedom	23	
Variance	23.67	
Standard Deviation	4.86	
Standard Error	0.99	
t-statistics	-8.02	
t-critical(p=.05)	-2.069	2.069
Margin of Error	2.05	
Confidence interval(CI)	-10.02	-5.91

Excel sheet:

File: p1_statistics.xlsx

Sheet name: "P1(Q5)"

Reference:

https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4578095863/e-147019320/m-147019322

https://www.udacity.com/course/viewer#!/c-ud134-nd/l-4621269407/e-230229168/m-230229171

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!

In my opinion, the reason for reading incongruent words slower than congruent words, is because reading the words is a much stronger behavior than recognizing the color. When we see any words, we read the words in our mind, and we do this action millions times in our life. On the other hand, we rarely say the color of the font. I think our brain wants to process reading the words first, as it's not used to doing unaccustomed things.

A similar task I've thought of, is putting blue colored cold water and red colored hot water on the tester's hand. In first treatment, an organizer put red and hot/blue and cold on tester's hand. Then the tester reports whether the water is hot or cold. In second treatment, the organizer put red and cold/blue and hot water on tester's hand. The tester has to report hot or cold as well. In this case the dependent variable is the time, until a tester says hot or cold. I think the color will affect the recognition of the heat on our body.

Reference:

https://moodle.unitec.ac.nz/mod/resource/view.php?id=58377&redirect=1