PROJECT: Test a Perceptual Phenomenon

Background Information :

In the Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant has to say out loud the color of the ink in which the word is printed. This task has two conditions: 1) congruent words condition, and 2) 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 YELLOW, PINK. In each case, we measure the time a candidate 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 :

Q1) What is our independent variable? What is our depenedent variable?

A1) The color words whose names match the colors in which they are printed is the independent variable. (i.e. whether the names of the colors match the ink colors). Reaction time – the time each participant takes to name the ink color is the dependent variable.

Q2) What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

A2) One hypotheses here is that there is a difference in reaction time between congruent words condition and incongruent words condition which means Stroop effect exists.It is not possible to perform the test with all population in the world.

Therefore, we use the sample to make inference about the population means. In our case, we will use the sample difference to estimate the population difference. Here is the hypotheses testing and explanation of the symbols I used:

H0 (null hypotheses)

HA (alternative hypotheses)

μcon (population mean of congruent words group)

μin (population mean of incongruent words group)

H0 : μcon - μin = 0

The null hypotheses is that there is no difference in reaction time on average between congruent words condition and incongruent words condition (i.e. the average time required to say the colors of the words between the two conditions are equal).

HA : μcon - μin ≠ 0

The alternative hypotheses is that there is a significant difference in reaction time between these two conditions (i.e. the average time required to say the colors of the words between the two conditions are not the same).

I expect to **perform two-tailed dependent t-test or one tailed dependent test** for paired samples. The reason for two-tailed test is that we want to test whether two means were significantly different from one another.

The reasons for using t-test are that the population standard deviation is unknown and the sample size is less than 30 (i.e. our n=24).

The reasons for using dependent t-test for paired samples:

- We are comparing the means of two related groups to determine whether there is a significant difference between these means. Our test is an example of "repeated-measures" statistical test.

- This means the same participants are tested more than once. Thus, each participant has been measured on two occasions (i.e. congruent words condition and incongruent words condition) for the same dependent variable (i.e. reaction time).

In addition, we need to consider the test assumptions for t-test for dependent means:

➢ The dependent variable is measured on an interval or ratio scale (i.e. reaction time is measure by seconds).

➢ Random sampling from a defined population.

➢ Samples used to produce the different scores are linked in the population through repeated measurement (i.e. the differences in reaction time are linked in the population through congruent words measurement and incongruent words measurement).

➢ Scores are normally distributed in the population; difference scores are normally distributed. However, the t-test for dependent means is considered typically "robust" for violations of normal distribution. This means although our sample time difference distribution is not normal, in most circumstance we can still use t-test for dependent means especially for a two-tailed t-test.

Q3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

A3) **Central Tendency:**

Congruent: Mean=14.051125

Median=14.233

Incongruent: Mean=22.01591667

Median=20.878

Mean and Median are almost the same for both congruent and incongruent but they are not equal. This tells us that the distribution is not perfectly normal.

Considering all data on the congruent and incongruent test groups have unique time, no same time exists as the data is continuous and hence , there is no mode available in both test groups.

Measures of Variability:

Congruent: Variance=12.14115286

Standard Deviation=3.484415713

Incongruent: Variance=22.05293383

Standard Deviation=4.69605513

Congruent	Incongruent	Deviation(C)	Deviation(I)	Squared deviation(C)	Squared deviation(I)	Difference		
8.63	15.687	-5.421125	-6.328916667	29.38859627	40.05518617	-7.057		
8.987	17.394	-5.064125	-4.621916667	25.64536202	21.36211367	-8.407	Significance level=0.005	
9.401	17.425	-4.650125	-4.590916667	21.62366252	21.07651584	-8.024	Standard deviation=4.865	
9.564	17.51	-4.487125	-4.505916667	20.13429077	20.30328501	-7.946	Degree of Freedom=23	
10.639	17.96	-3.412125	-4.055916667	11.64259702	16.45046001	-7.321	Confidence level=2.8475	
11.344	18.644	-2.707125	-3.371916667	7.328525766	11.36982201	-7.3		
12.079	18.741	-1.972125	-3.274916667	3.889277016	10.72507917	-6.662		
12.13	19.278	-1.921125	-2.737916667	3.690721266	7.496187674	-7.148		
12.238	20.33	-1.813125	-1.685916667	3.287422266	2.842315007	-8.092		
12.369	20.429	-1.682125	-1.586916667	2.829544516	2.518304507	-8.06		
12.944	20.762	-1.107125	-1.253916667	1.225725766	1.572307007	-7.818		
14.233	20.878	0.181875	-1.137916667	0.03307851563	1.29485434	-6.645		
14.48	21.157	0.428875	-0.8589166667	0.1839337656	0.7377378403	-6.677		
14.669	21.214	0.617875	-0.8019166667	0.3817695156	0.6430703403	-6.545		
14.692	22.058	0.640875	0.04208333333	0.4107207656	0.001771006944	-7.366		
15.073	22.158	1.021875	0.1420833333	1.044228516	0.02018767361	-7.085		
15.298	22.803	1.246875	0.7870833333	1.554697266	0.6195001736	-7.505		
16.004	23.894	1.952875	1.878083333	3.813720766	3.527197007	-7.89		
16.791	24.524	2.739875	2.508083333	7.506915016	6.290482007	-7.733		
16.929	24.572	2.877875	2.556083333	8.282164516	6.533562007	-7.643		
18.2	25.139	4.148875	3.123083333	17.21316377	9.753649507	-6.939		
18.495	26.282	4.443875	4.266083333	19.74802502	18.19946701	-7.787		
19.71	34.288	5.658875	12.27208333	32.02286627	150.6040293	-14.578		
22.328	35.255	8.276875	13.23908333	68.50665977	175.2733275	-12.927		
Sum:	Sum:			291.3876686	529.2704118			
337.227	528.382			VARIANCE:	VARIANCE:			
Mean:	Mean:			12.14115286	22.05293383			
14.051125	22.01591667			SD:	SD:			
				3.484415713	4.696055135			
	Difference in mean =		-7.964791667					

https://docs.google.com/spreadsheets/d/123Wov_KYwJHZC4FteWVjEbRwTAJO65rNcesH5VZQFX0/edit?usp=sharing

Q4) 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.
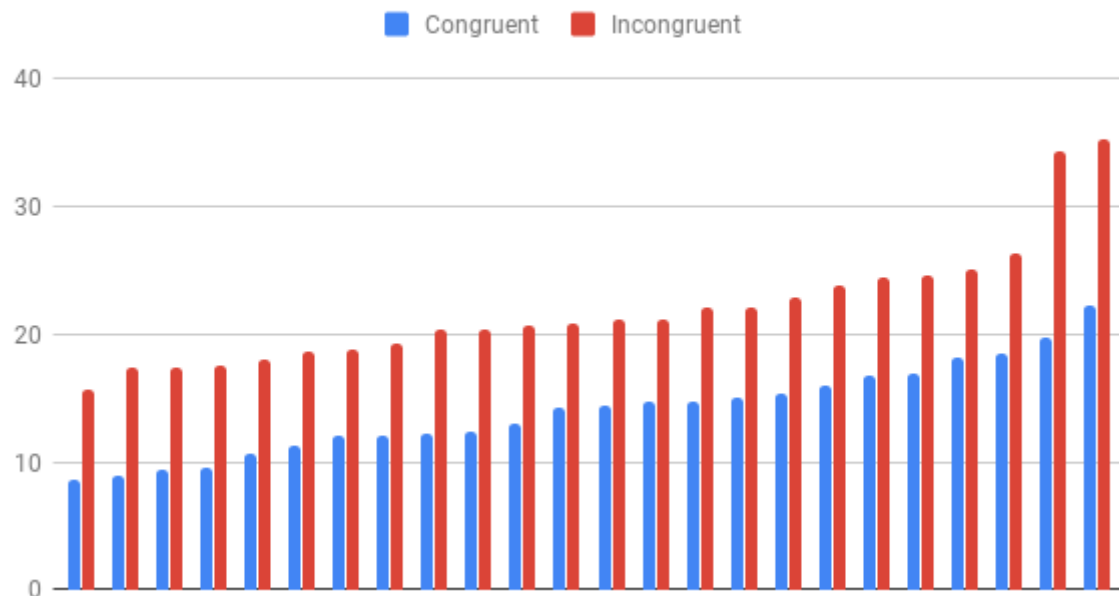
A4) Visulaizations:

1) HISTOGRAM:

Congruent and Incongruent
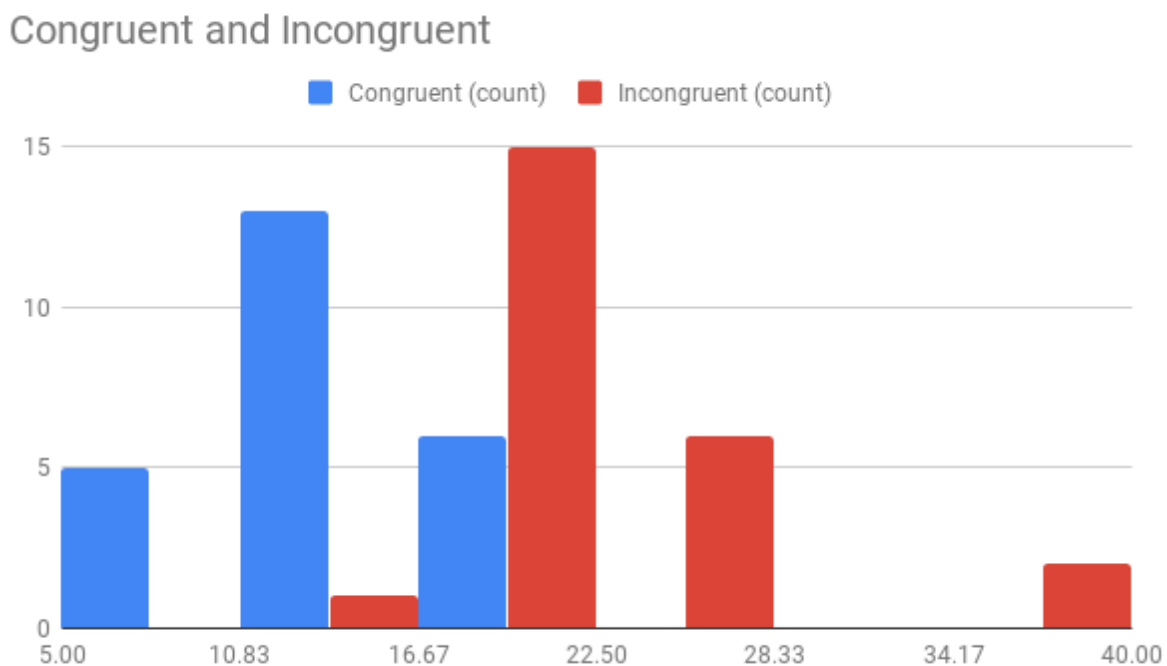
The above histogram shows that the two groups have appreciable difference in their mean, median and mode. Here, the incongruent group needs more time , whatever the measurement maybe.
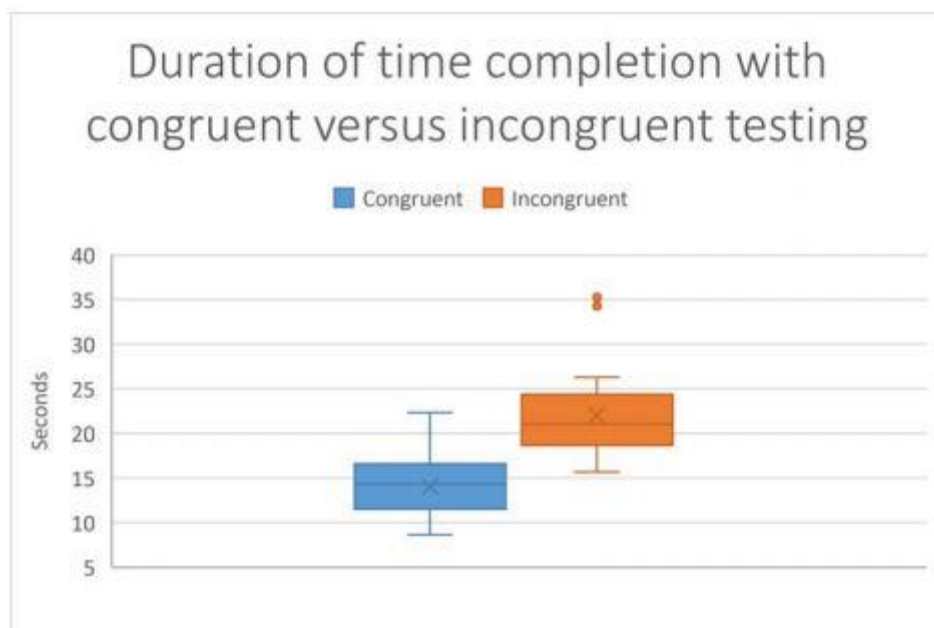
2) BAR CHART:

Congruent and Incongruent

In this chart we can see that the time taken by a candidate to say the colour of the ink is always more in the case of incongruent group as compared to the congruent group.

3) BOX PLOT:

The above box plot shows us a graphical representation of the range of each set at a glance, summarizing and allowing us to compare the two test's data and the distribution of that data. The box plot is simply a quick top down view of the data from the congruent and incongruent tests . In the data we see that there is a significant increase in duration of time taken to finish the test from the congruent test versus the incongruent test as 75% of the data in the incongruent test is above the IQR of the congruent test. That being said about 25% of the data in the congruent test looks like it falls in about 50% to 75% of the incongruent test so there is quite a bit of overlap when considering all data but there is some tendency to believe that incongruent test on average from this plot that summarizes the data will take longer to finish than the congruent test.

Q5) 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?

A5)

Congruent Mean- Incongruent Mean= 14.05-22.02 = -7.97

Variance of difference = 23.667

Standard Deviation of Differences(s) = 4.865

No. of participants in the test(n)=24

T-Statistic=Mean Difference/(s/√n)

$\qquad$ =-7.97/(44.865/√24 = -8.026

Since we have the T-Statistic we can compare it to the T-Critical value to determine if the null hypothesis should be accepted or rejected, using what we

know (24 participants, which means n=24 we can determine the Degrees of Freedom)

Knowing n=24, the degrees of freedom for this t-test = n-1 = 23. Since we are aiming to find with 99%

Confidence level for a one tail our $\alpha = 0.005$

Given df = 23 and $\alpha = 0.005$

@ 99% Confidence Level: need to be between +- 2.807

Given that the T-Statistic found was - 8.026 it leads us to **reject the null hypothesis** due to T-Statistic being in the critical area of tested at 99% Confidence Level.

The t-test confirms what was expected that incongruent tests would normally take longer than congruent tests as congruent test linked both visual colors with the correct printed name of the color while incongruent required to try to separate the visual color and the name of the color which required a bit more time to do.

REFERENCES:

https://faculty.washington.edu/chudler/words.html

https://statistics.laerd.com/statistical-guides/dependent-t-test-statistical-guide.php

https://www.youtube.com/watch?v=_x6j-rH33Oq

https://corporatefinanceinstitute.com/resources/excel/functions/t-confidence-interval-test-excel/

https://stattrek.com/hypothesis-test/hypothesis-testing.aspx