PROJECT-2

TEST A PERCEPTUAL PHENOMENON

1) What is the independent variable? What is our Dependent Variable?

Answer:

Independent Variable

The independent variable is words that are given to the paricipants or their colour.The word was congruent or incongruent was an independent variable

Dependent Variable

The dependent variable is the time taken by all the participants because their time was dependent on the independent variable of the test.

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

Answer:

In this scenario, our hypothesis would be, that giving incongruent words is statistically siginificant or not.There are no information given about the samples other than their congruent and incongruent scores. So we would assume that these samples are randomly choosen which means that the participants are of different age, both male and female and of different education level. In this context :

• μ_1 is the mean of time spent for the participant group of congruent test.

• μ2 is the mean of time spent for the participant group of incongruent test.

NULL Hypothesis :There is no significant difference in the time between congruent and incongruent scores i.e $\mu_1 - \mu_2 \geq 0$.

ALTERNATIVE hypothesis :There would be a significant difference between congruent and incongruent scores i.e $\mu_1 - \mu_2 < 0$.

There are multiple hypothesis test that can be performed like Z-test, t-test, F-test, but here we are taking two different types of test on the same sample and we don't have population parameters available,so here we perform **dependent t-test.** We go with the one tailed or two tailed t-tests to evaluate how significant the incongruent words are on the time taken. We would also give a 95% confidence interval.

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

Answer:

CENTRAL TENDENCY:

	Mean (μ)	Median
CONGRUENT	14.051	14.233
INCONGRUENT	22.016	20.878

As we can see above, the mean and the median is almost same for both the congruent and the incongruent scores but they are not equal. This tells that , the distribution is not perfectly normal otherwise these values would have been same for both.

MEASURE OF VARIABILITY:

Var(congruent) = 12.141

 Var(incongruent) = 22.053

Given we now know the Variance for each of the data sets ,we can use that knowledge to get the sample standard deviation for each of the test groups (congruent and incongruent).

SD= $\sqrt{\text{Variance}}$

 SD (congruent) = 3.559

 SD (incongruent) = 4.797

Congruent	Sum	Mean	Deviation	Squared Deviation	Incongruent	Sum	Mean	Deviation	Squared Deviation	Difference	MeanDifference		
8.63	337.227	14.051125	-5.421125	29.38859627	15.687	528.382	22.01591667	-6.328916667	40.05518617	-7.057	-7.964791667		
8.987			-5.064125	25.64536202	17.394			-4.621916667	21.36211367	-8.407		Significance Level=	0.005
9.401			-4.650125	21.62366252	17.425			-4.590916667	21.07651584	-8.024		Standard Deviation=	4.865
9.564			-4.487125	20.13429077	17.51			-4.505916667	20.30328501	-7.946		Degree of freedom=	23
10.639			-3.412125	11.64259702	17.96			-4.055916667	16.45046001	-7.321		Confidence value=	2.847518564
11.344			-2.707125	7.328525766	18.644			-3.371916667	11.36982201	-7.3			
12.079			-1.972125	3.889277016	18.741			-3.274916667	10.72507917	-6.662			
12.13			-1.921125	3.690721266	19.278			-2.737916667	7.496187674	-7.148			
12.238			-1.813125	3.287422266	20.33			-1.685916667	2.842315007	-8.092			
12.369			-1.682125	2.829544516	20.429			-1.586916667	2.518304507	-8.06			
12.944			-1.107125	1.225725766	20.762			-1.253916667	1.572307007	-7.818			
14.233			0.181875	0.03307851562	20.878			-1.137916667	1.29485434	-6.645			
14.48			0.428875	0.1839337656	21.157			-0.8589166667	0.7377378403	-6.677			
14.669			0.617875	0.3817695156	21.214			-0.8019166667	0.6430703403	-6.545			
14.692			0.640875	0.4107207656	22.058			0.04206333333	0.001771006944	-7.366			
15.073			1.021875	1.044228516	22.158			0.1420833333	0.02018767361	-7.085			
15.298			1.246875	1.554697266	22.803			0.7870833333	0.6195001736	-7.505			
16.004			1.952875	3.813720766	23.894			1.878083333	3.527197007	-7.89			
16.791			2.739875	7.506915016	24.524			2.508083333	6.290482007	-7.733			
16.929			2.877875	8.282164516	24.572			2.556083333	6.533562007	-7.643			
18.2			4.148875	17.21316377	25.139			3.123083333	9.753649507	-6.939			
18.495			4.443875	19.74802502	26.282			4.266083333	18.19946701	-7.787			
19.71			5.658875	32.02286627	34.288			12.27208333	150.6040293	-14.578			
22.328			8.276875	68.50665977	35.255			13.23908333	175.2733275	-12.927			
			0	291.3876686					529.2704118				
				12.66902907 <=variance					23.01175704 <=variance				
				3.559357958 <=sd					4.797057122 <=sd				

https://docs.google.com/spreadsheets/d/1nItMJ_hTkizN6wY1PcAQLkG_5vrDK2d BMlq_S7jfvSI/edit?usp=sharing

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.

Answer:

BOX PLOT:

Duration of time completion with congruent versus incongruent testing
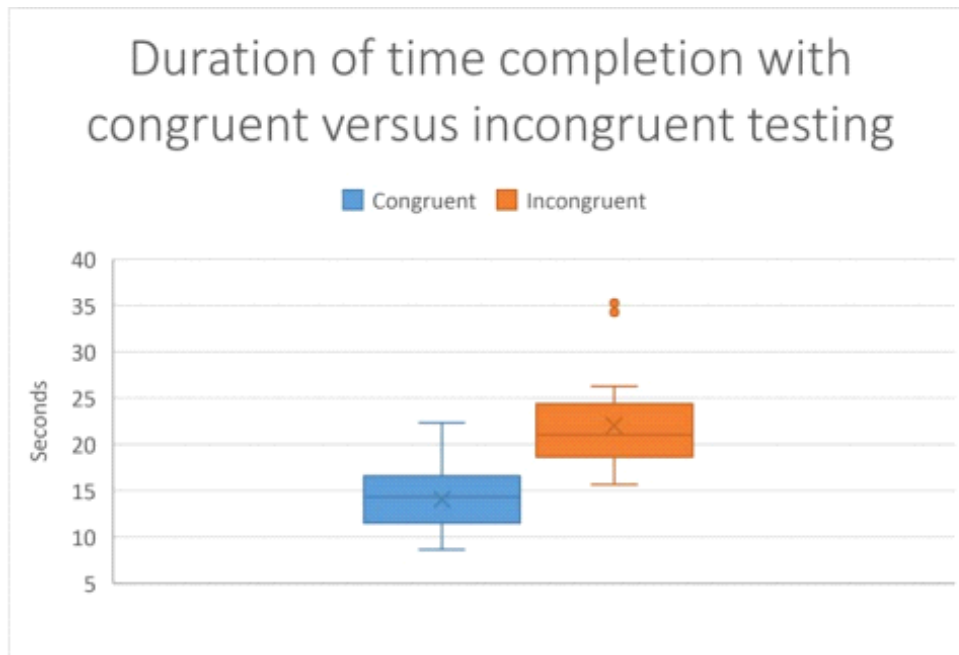
IQR(Congruent) - 4.68   IQR(Incongruent) - 5.51

The first visualization we can use is a box plot that 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 and any indications of symmetry within that data.

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. There is some tendency to believe that incongruent test on average will take longer to finish than the congruent test.

BAR CHART:

Congruent and Incongruent
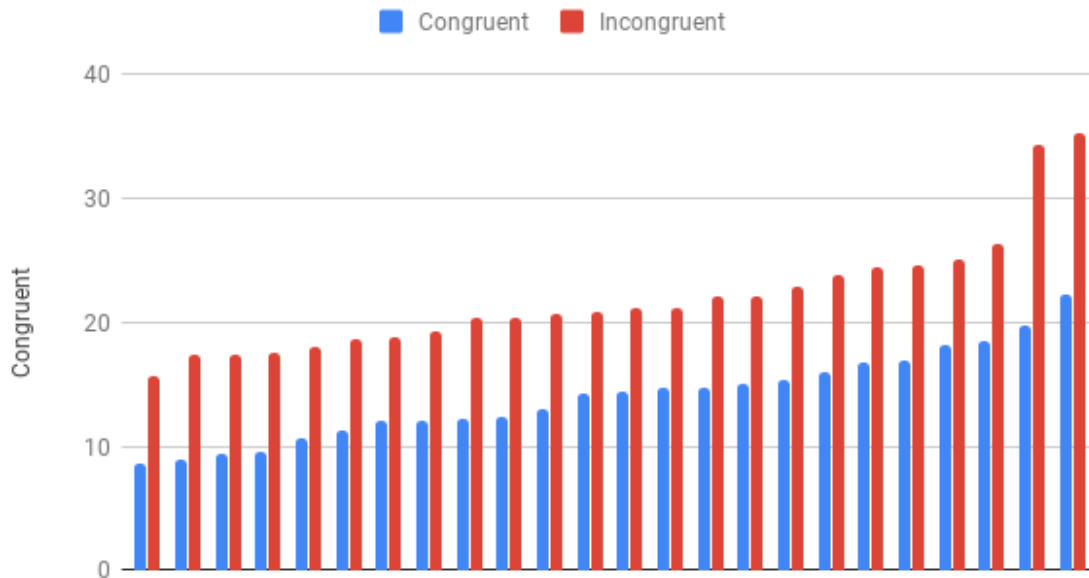
This is the comparison of the congruent and incongruent scores for each indivisuals. It can be observed that the value of the congruent scores in BLUE is lesser than all the incongruent scores colored in RED . This is the obvious situation because people would naturally take less time for pronouncing the congruent words as compared to incongruent ones.

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?

Answer:

- Difference of Congruent And 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/($s/\sqrt{n}$) = -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.

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 +- 2.847

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:

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

2.

https://www.khanacademy.org/math/statistics-probability/confidence-intervals-one-sample/old-confidence-interval-videos/v/t-statistic-confidence-interval

3.

https://en.wikipedia.org/wiki/Stroop_effect

4.

https://support.office.com/en-us/article/create-a-box-plot-10204530-8cdf-40fe-a711-2eb9785e510f