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

What is our independent variable? What is our dependent variable?

The Swoop Test uses a repeated measures design where the same subject receives both tests. This is an example of a within-subject design. Other examples include longitudinal studies where the same subject is recorded at different points in time or testing 2 different interventions. Independent variables are the intervention taken to evaluate the effect on the dependent variable. In the Swoop Test experiment the independent variable is the different tests and the test score is the dependent variable. This is because we change the test to see if there is an effect on the score, so we end up with paired scores.

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

The null hypothesis for this experiment is that the test change has no impact on test score. In other words, the score on the congruent test should be the same as the score on the incongruent test. There are several alternate hypotheses that could be applied to this set. That the score would be worse, better, or just different on the incongruent test in comparison to the initial, congruent test.

For purposes of this evaluation, the alternate hypothesis evaluated will be that the score on the second, incongruous test will be different than the score on the initial test.

The Null and Alternate Hypothesis:

$H_n: \mu_c = \mu_i$ or $\mu_c - \mu_i = 0$

$H_a: \mu_c \neq \mu_i$

There are many statistical tests that can be used to determine significance and correlation. In this analysis, we have paired samples. That is, we are comparing the results from 2 interventions (the tests) applied to the same subject. Also, our null hypothesis is that the difference between the means of each sample will be zero. Finally, our sample distributions both resemble normal distributions with no outliers. Based on this, the most appropriate statistic to use for our analysis is the 2 tailed paired t-test.

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

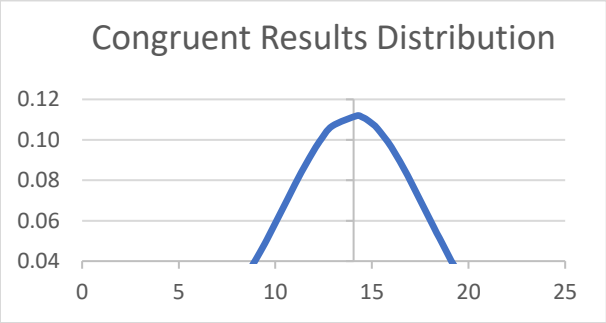
Two samples were provided for review as part of this analysis (Udacity, 2021). The two samples were test scores for the congruent test and the scores for the incongruent test. The data tables reviewed are provided at the end of this analysis for reference.

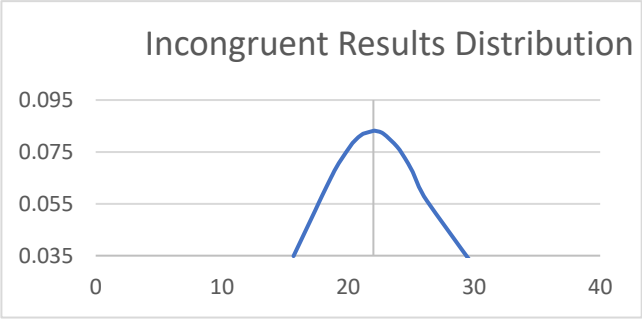
First we can review the measures of central tendency. These descriptive statistics look at the middle point of data and give us some idea of possible distribution of the data. Along with those measures, we can look at the standard deviation for the samples. Standard deviation is a measure of variability that looks at the difference between each value and the mean of the sample. If the differences are small, then the values are close to the mean and there is little variability in your sample. If the differences are larger or not clustered, there is more variability in your sample and thus a larger margin of error compared to the population in question.

The Congruent sample has a mean or average value of 14.05 and a standard deviation of 3.56. This means that assuming a standard distribution, about 95% of the values for the congruent test score fall between about 6.93 and 21.17 (values rounded to the nearest hundredth). In comparison, the Incongruent sample has a mean of 22.05 with a standard deviation of 4.8. These are not wildly different standard deviations but the incongruent sample definitely has more variability

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.

Below is a visualization of the normalized distribution of both samples. We can see that the congruent results have a slight positive skew while the incongruent results are closer to a normal distribution.





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?

When we perform the 2 tailed t-test for paired data on this set, we get a score of -8.02. The formula for this is:

$$t = \frac{M_D - \mu_D}{s_{M_D}}$$

To summarize, it is the difference between the sample means (congruent – incongruent) divided by the standard error of the difference. The standard error of the difference is found by dividing the Standard Deviation of the differences by the square root of the sample. Additional details for the calculations are in the attached excel document

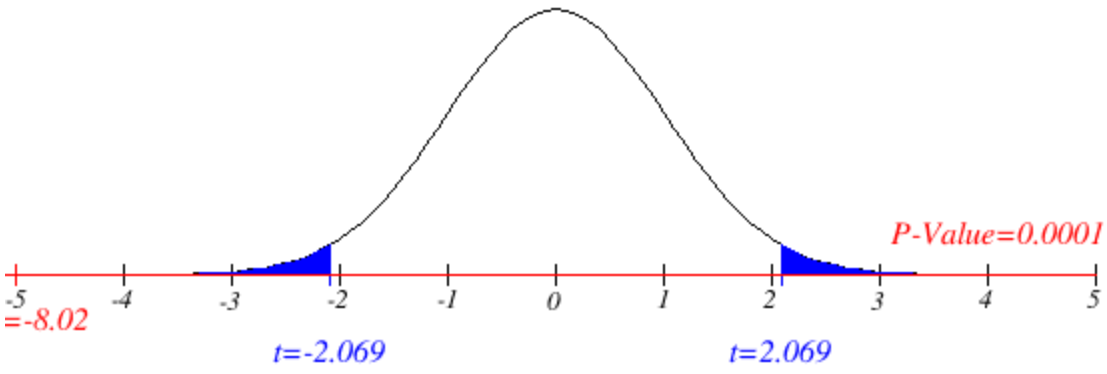
$$Mean_d = Mean_c - Mean_i \quad : \quad Mean_d = 14.36 - 22.02 \quad : Mean_d = -7.96$$

$$SE_d = SD_d \div \sqrt{n} \quad : \quad SE_d = 4.86 \div \sqrt{24} \quad : \quad SE_d = 4.86 \div 4.9 \quad : \quad SE_d = 0.99$$

$$t = -7.96 \div 0.99 = -8.02$$

How do we identify if this is significant? We need to see where this score falls against the critical t-score for the same degrees of freedom. To identify the t-critical value for this assessment, we identify the Degrees of Freedom which is the sample size – 1, so 23. We also determine a level of probability or confidence interval. We do want to be certain of the impact of the intervention so we will use a $p<0.05$ or a 95% confidence interval. In looking at the t-table (https://www.stat.tamu.edu/~lzhou/stat302/T-Table.pdf) , the t-score for p=.05 is 1.714 but because this is a 2 tailed test, the .05 is divided between two tails so we use the value in the 0.025 column which is 2.069.

Based on these, our t-critical value is t(23) = +/- 2.069,$p \approx 0.05$. Our t-score of -8.02 is outside the bounds for the critical value which indicates it is statistically significant. I've included a graph of our value compared to the t-critical value below. Based on this, we can say that it took longer for subjects to complete the incongruent test (M = 220.2, SD = 4.8) than it did for subjects to complete the congruent test, t(50) = -1.676, $p \approx 0.05$ and thus we reject the null hypothesis.



We can see how much of the change can be attributed to the test difference by looking at Cohen's D which is the mean difference divided by the standard deviation of the difference. In this instance we get d=-12.83 which indicates a large effect from the test.

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, this test shows the interesting way in which our brain processes stimuli. The word and the color are somehow linked within our brain – likely for ease of use- and having to process the incongruence takes a bit more processing time. I read about a similar type of test where researchers dyed foods different colors than normal – green mashed potatoes for example. When people ate the food blindfolded, they could easily identify the food but without the blindfold, it was more challenging.

Data Table

Congruent	Incongruent
12.079	19.278
16.791	18.741
9.564	21.214
8.63	15.687
14.669	22.803
12.238	20.878
14.692	24.572
8.987	17.394
9.401	20.762
14.48	26.282
22.328	24.524
15.298	18.644
15.073	17.51
16.929	20.33
18.2	35.255
12.13	22.158
18.495	25.139
10.639	20.429
11.344	17.425
12.369	34.288
12.944	23.894
14.233	17.96

19.71	22.058
16.004	21.157

References:

Udacity Data Set (2021, January 30).
https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view

T-curve graphing:

http://www.imathas.com/stattools/norm.html

T Table

https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg

Other Sites Referenced

https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php

https://towardsdatascience.com/inferential-statistics-series-t-test-using-numpy-2718f8f9bf2f