Jostin Flake

Test a Perceptual Phenomenon

Statistics: The Science of Decisions Project

Premise: The following results section provides a comprehensive analysis of statistics retrieved from the “Stroop” dataset. This set of data was collected by “Udacity,” and as such represents the findings of a two condition experiment, where a random sample of twenty-five individuals, referred to as subjects, were asked to read from two separate lists of words. The first list was made up of a specified number of *congruent* words (e.g., **BLUE**, **ORANGE**, and **BLACK**) while the second list contained the same amount words, except the words in this list, would be *incongruent* (e.g., **RED, WHITE,** and **YELLOW**). Each list represents a distinct and unique condition. For the experiment, the subjects were timed on the number of seconds that it took to read aloud the color and not the text for the complete set of words for each list. This time was then recorded in the data set under the respective condition.

From the premise of this experiment, we can derive both our dependent and independent variables. The **Dependent Variable** is the number of seconds recorded to read every color in each list completely. The **Independent Variable** in this experiment is the type or condition of the list, which is either *congruent* or *incongruent*.

Based on the experimental background information provided in the premise, it is clear to see that the research was conducted to verify the following questions; Will the amount of time it takes to read each color from the incongruent list change significantly from that of the time from the congruent list? Will the time it takes to read aloud the color of the words in the list increase or decrease?

The questions above provide a basis for both the null hypothesis and the alternative hypothesis.

The **Null Hypothesis** will state that the difference between the average number of seconds it takes to read through the *congruent* list (represented by M_C) and the average number of seconds it takes to read through the *incongruent* list (represented by M_I) will be equal to zero. The null hypothesis will be symbolized by the following:

$$H_0: M_C - M_I = 0;$$

or

$$M_C = M_I;$$

Justification to retain the null hypothesis is achieved through the logic that the mean from the *incongruent* sample set (M_I) will be a number equal to the mean from the *congruent* sample (M_C). If the null hypothesis proves to be false, it will be due to the mean from the *incongruent* sample (M_I) being small or large enough to make the difference not equal zero; in this scenario, we would reject the null hypothesis.

Conversely, the **Alternative Hypothesis** states that the difference between the average number of seconds its takes to read through the *congruent* list (represented by M_C), and the average number of seconds it takes to read through the *incongruent* list (represented by M_I) will be not equal zero. The alternative hypothesis will be symbolized by the following:

$$\mathbf{H_A: M_C - M_I \neq 0;}$$

or

$$\mathbf{M_C \neq M_I;}$$

Justification for rejecting the null hypothesis and accept the alternative hypothesis is achieved when it has been verified that the difference from the *incongruent* list mean ($\mathbf{M_I}$) will be a number less than or greater than the mean from the *congruent* sample ($\mathbf{M_C}$). If the alternative hypothesis proves to be false, it will be due to the mean from the *incongruent* sample ($\mathbf{M_I}$) being any number that makes the difference equal to zero.

The data provided does not represent the entire population of data, only a sample set of data for the experiment. Due to this fact, we cannot determine the true population mean(μ) or the true population standard deviation(σ), causing that the hypotheses be determined false or true by performing a statistical t-Test on the sample data. For this experiment, the specific kind of statistical test will be a Dependent paired-samples t-test. The direction of the test will be two-tailed because the hypotheses are verifying simply whether the time is significantly different, for the better or worse. The reason for a Dependent sample statistical t-Test, rather than an Independent Sample, is because the experiment is performed on the same subject twice. The experiment focuses on testing the two conditions on the same subject, which is one of the advantages of the Dependent Sample t-Test. The statistical test will be using a conventional critical alpha level of five percent (**α = 0.05**). This will leave two-point-five percent (.025) in each critical region of our two-tailed test.

Now that the null and alternative hypotheses are formed and the kind of statistical test has been chosen, the first step is to compute descriptive statistical data from the 'Stroop" dataset. Figure 1-1: below displays the sample data for both our *congruent* and *incongruent* samples. A new column has been created to the right of each of the lists of data to hold the values of our statistical data.

Congruent	$X_C = 14.05$	Incongruent	$X_I = 22.02$
12.079	$S_C = 3.56$	19.278	$S_I = 4.80$
16.791	$V_C = 12.67$	18.741	$V_I = 23.01$
9.564	$SE_C = 0.73$	21.214	$SE_I = 0.98$
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	

Figure 1-1: Stroop data and descriptive statistical data

The **sample mean** for the *congruent* samples is represented by (**X_C**) while the sample mean for the *incongruent* samples is represented by (**X_I**). Both were calculated by adding together each data point then dividing the total by the number of samples taken. Final totals were always round to the nearest two decimal places.

Here is the formula symbolized:

$$\boldsymbol{ROUND\left(\frac{\Sigma(X1, X2, \dots\ Xn)}{n}, 2\right)}$$

In the experiment, the sample mean was used as the measurement of centrality. After calculating the sample mean for each list, it was possible to compute our point estimate or the **sample mean difference** (**X_D**), which can be calculated in two ways. The first way is to take our two sample means and find the difference from the two sample means. The second way would be to find the difference from each point in the data and then find the average difference. Both formulas provide the same answer; however, the first way was chosen, and it is symbolized here:

$$\mathbf{X_C - X_I = X_D}$$

The sample mean difference equated to negative seven-point nine seven (**X_D = -7.97**). Figure 1-1: also shows some measurements of variability, such as the **standard deviation**, **variance**, and **standard error** for each sample list. When dealing with paired-samples, it is important to be able to compute the t-statistic. Another descriptive statistic that is required in order to calculate the t-statistic is also a measure of variability. The **standard deviation of differences** (**S_D**) is calculated by taking each point difference and from that finding the standard deviation of those differences.

The formula to find the standard deviation of difference (**S_D**) is symbolized here:

$$C_1 - I_1 = D_1$$

$$C_2 - I_2 = D_2$$

.

.

$$C_n - I_n = D_n$$

The letter "C" represents *congruent*, and the letter "I" represents *incongruent*. Letter "D" represents the *difference*. Therefore, C1 represents the first data point for the *congruent* data, and I1 represents the first data point for the *incongruent* data. Continue calculating the difference between each data point in the list for "n" number of data points. In this dataset, the **sample size** is twenty-five (**n = 24**). Once the difference has been found twenty-five times, it is time to find the **sum of squared differences** (**SS_D**). In order to compute the sum of squared differences (**SS_D**), take the *difference* calculated in the previous step and subtract the sample mean difference (**X_D**) from that number, then squaring the answer. After you have squared each data point, you will sum each point total and divide the sum by our sample size(n) minus one. The reason we divide by our sample size(n) minus one is due to Bessel's correction. This will account for the fact that we are dealing with a sample set of data and not a population set of data. The following formulas symbolize this process here:

$$SS_D = \frac{\Sigma((D1 - XD)^\wedge 2),((D2 - XD)^\wedge 2),\ldots.,((Dn - XD)^\wedge 2)}{n-1}$$

The final step in the calculation of the standard deviation of differences (**S_D**) is simply to find the square root of the sum of squared differences (**SS_D**). Here is the formula:

$$\mathbf{S_D = \sqrt{SS_D}}$$

From the data provided, the sum of squared differences came out to be twenty-two point six-eight (**$SS_D = 22.68$**). The standard deviation of differences for this set of sample data equates to four-point seven six (**$S_D = 4.76$**).

Now that we have acquired some descriptive statistics, this is a good time to provide a histogram chart on the sample data. Figure 2-2: below displays a histogram for both dependent variables for the experiment, that being the *congruent* and *incongruent* lists.

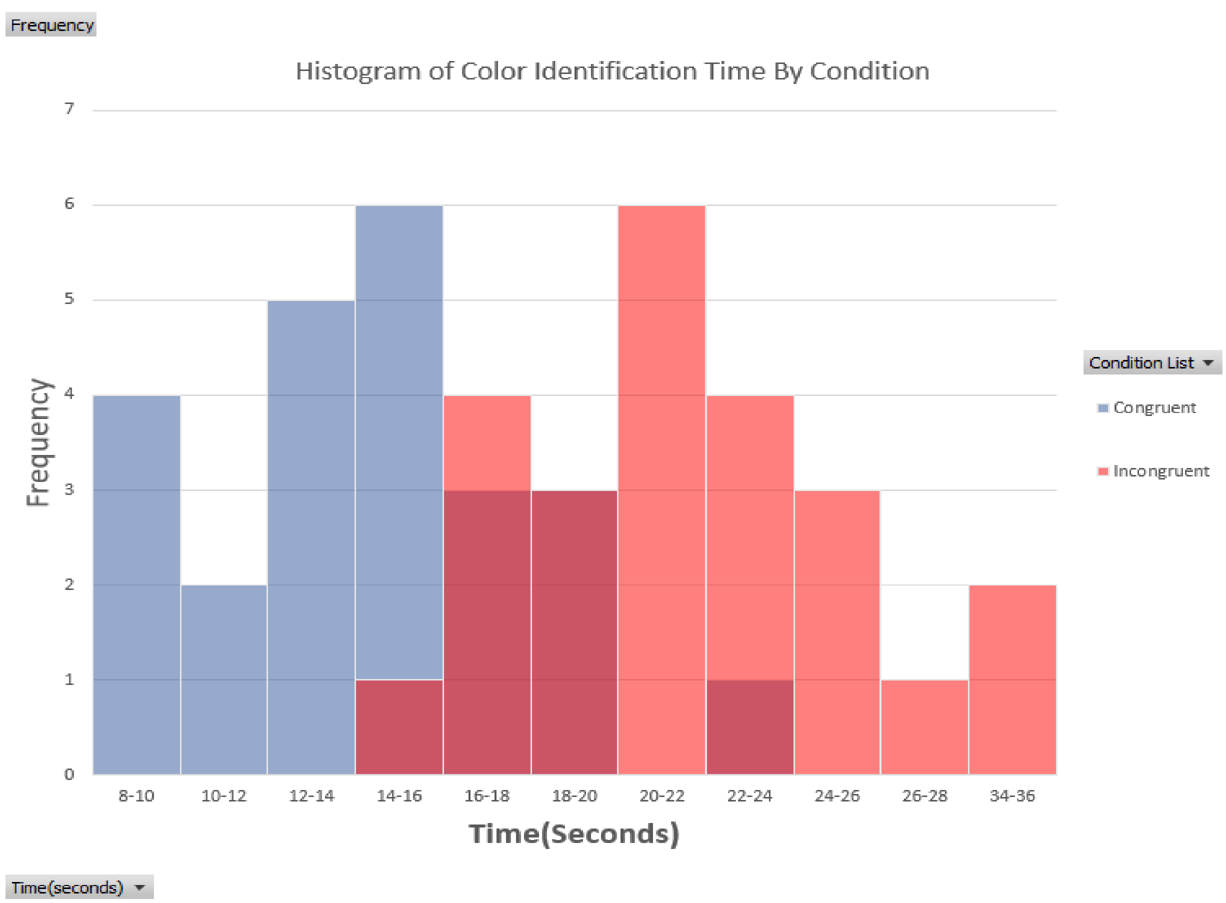

Figure 2-2: Histogram of Color Identification Time by Condition

From observing the graph shown above, notice that the bin size for the x-axis (**Time**) is two. The range of data for the *congruent* sample is from eight to twenty-four, and the range of data for the *incongruent* sample is from fourteen to thirty-six. Also, from the graph we can see that the majority of the *incongruent* sample data falls to the right of the *congruent* sample data, with few to little overlap.

After calculating the required descriptive statistics, it is time to perform the statistical test. To begin, determine the degrees of freedom(**df**), the t-critical value(**t-cv**), and the t-statistic(**t**). The degrees of freedom for a Dependent Sample t-test is calculated by subtracting one from your sample size. The sample size for this experiment was twenty-four, making the degrees of freedom twenty-three (**df = 23).** Next, calculate the t-critical value(**t-cv**), which marks the threshold passing into the two critical regions in the distribution for a two-tailed test. By using a t-table, the values of the degrees of freedom(**df**), and the alpha level(**a**), the t-critical value is determined. Figure 3-3: displays the row and column, which are circled in red, used based on those values, and the t-critical value is highlighted in yellow. Note that for a two-tailed test, the alpha level (0.05) is divided by two to account for two critical regions as opposed to just one. This is the reason the column marked for tail probability *p* (.025) is used when determining the t-critical value. See Figure 3-3: below:

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

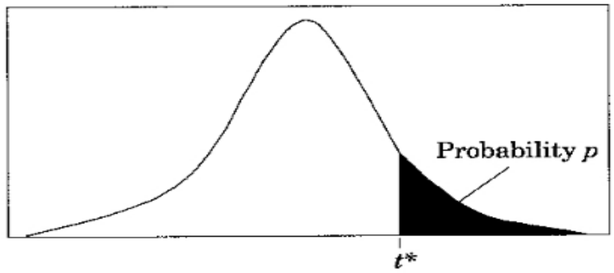

Table B t distribution critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Figure 3-3: t-table

From the t-table, the t-critical value equates to positive and negative two-point zero six-nine (**t-cv = ± 2.069**). It is important to understand the location of the t-critical values(**t-cv**), and the plot point lies within the distribution. Figure 4-4: below shows the critical region shaded in black, and the t-critical values(**t-cv**) plotted on the normal distribution. Notice this represents a confidence level of ninety-five percent.

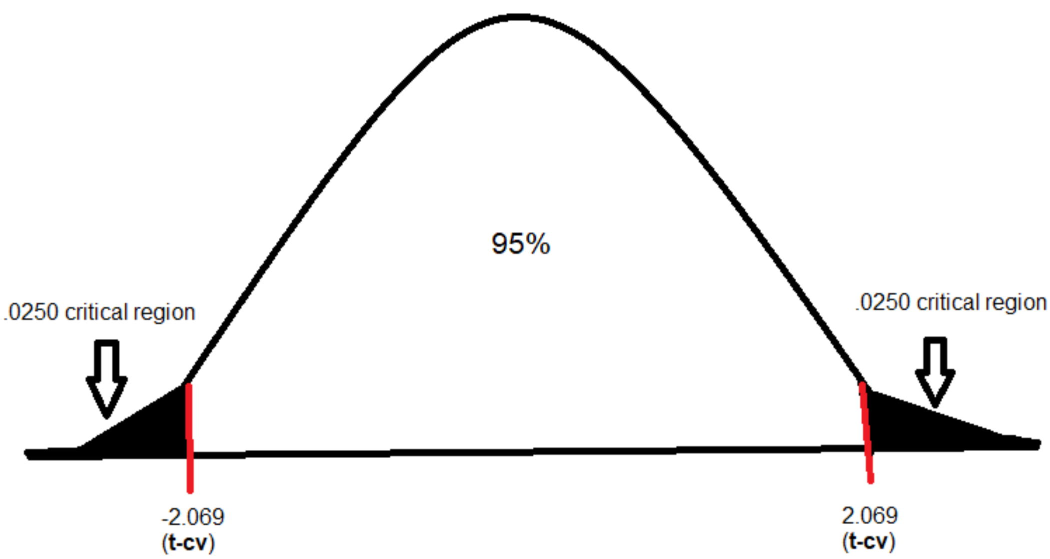

Figure 4-4: t-critical values on a 95% confidence level distribution

The next step is to determine the t-statistic(**t**), which will help us determine whether to reject or retain the null hypothesis(**H_0**). To do this, we took our sample of mean differences (**X_D**) and divided it by the **standard error** (**SEM**). It is important to calculate the standard error (**SEM**) by using the standard deviation of difference divided by the square root of the sample size. Here is the formula symbolized to calculate the standard error: $\mathbf{SEM = (S_D / \sqrt{n})}$

The standard error calculated to be zero-point nine-five (**SEM = 0.97**). Now plug the standard error into the t-statistic formula shown symbolized here: $\mathbf{t = X_D / SEM}$

The t-statistic equated to negative eight-point two-two (**t = -8.22**). After determining the t-statistic, it should be plotted on the distribution to verify wherein the confidence interval the score lies. Figure 5-5: displays this information using the same distribution as Figure 4-4: but with additional information to plot the t-statistic(**t**).

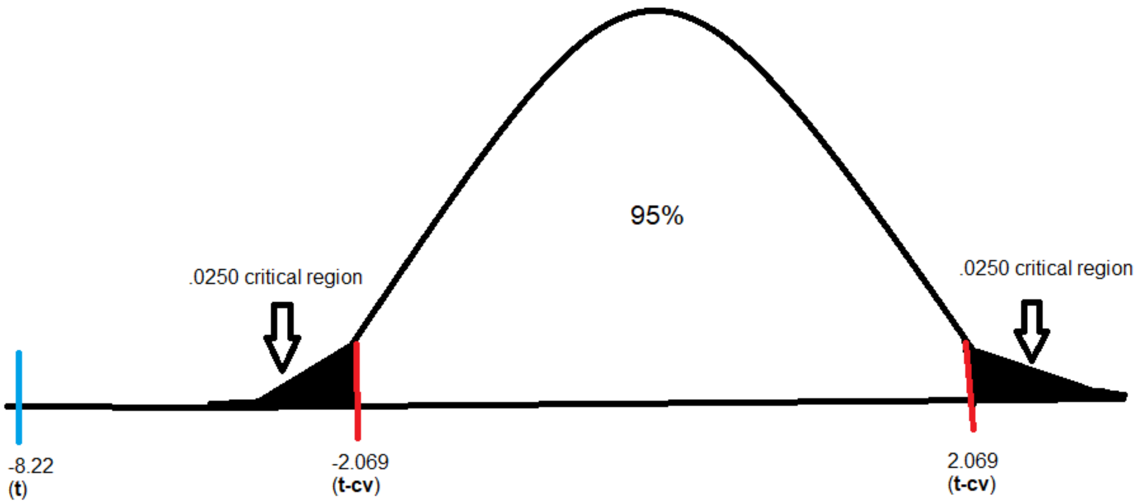

Figure 5-5: t-statistic lies far beyond the negative critical region

From the observations of the distributions of the sample data, it is clear to see that the t-statistic(**t**) lies far beyond the negative critical region. The probability value of obtaining the reflected results is less than point-zero zero zero-one (**p = .0001**)

According to this data, the null hypothesis (**H_0**) is **rejected** as it is very unlikely to obtain a time while reading from the *incongruent* list equal to a time while reading from the *congruent* list. The alternative hypothesis (**H_A**) is accepted as it is safe to say that the mean of times from the *congruent* list is not equal to the mean of times from the *incongruent* list.

$$M_C \neq M_I;$$

This aligns the outcome or results with the expectations set for this experiment.