Marcus A. Streips – Udacity – Project 1 version 2

Statistics: The Science of Decisions Project Instructions

Background Information

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

INDEPENDENT – Word Condition (Incongruent and Congruent)

DEPENDENT – Time (seconds)

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

Ho – Null Hypothesis: $\mu i - \mu c = 0$, the true difference in the means of the incongruent and congruent populations is zero.

Ha – Alternate Hypothesis: $\mu i - \mu c \neq 0$, the true difference in the means of the incongruent and congruent populations is NOT zero.

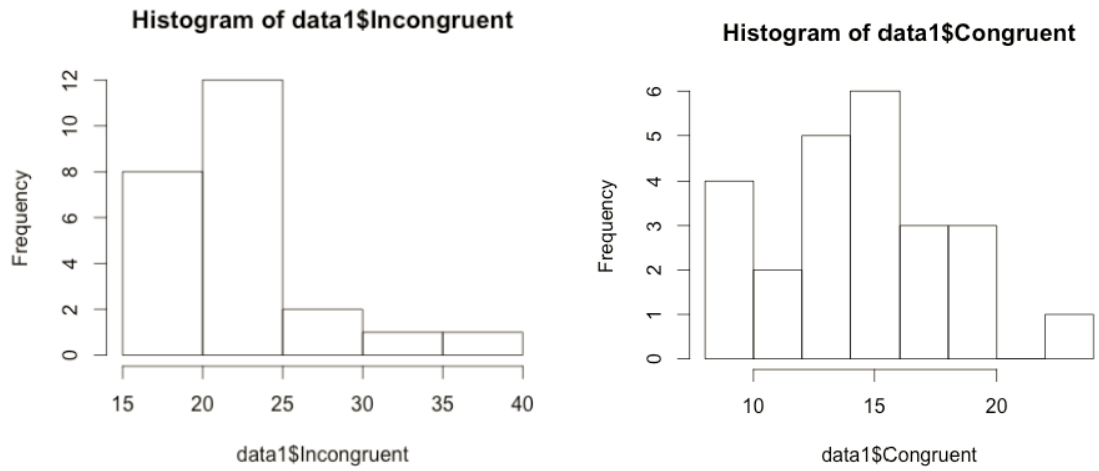
I expect to perform a two-sided, paired (dependent sample) t-test because the of the small sample size (N<30) and since I do not know the population parameters (making a z-test inappropriate). I am assuming a two-sided test, because I am trying to determine whether there is a difference, not whether one is greater or less than the other. Finally, I am choosing to do a paired t-test because the same participant is asked to perform both tests, therefore they are in both populations (not independent).

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

```
> describe(data1)
            vars  n  mean   sd median trimmed  mad   min   max range skew
Congruent      1 24 14.05 3.56  14.36   13.88 3.49  8.63 22.33 13.70 0.37
Incongruent    2 24 22.02 4.80  21.02   21.29 3.89 15.69 35.26 19.57 1.36
            kurtosis   se
Congruent      -0.62 0.73
Incongruent     1.52 0.98
```

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.

Using R – Hist function.



With the congruent data, the range of distributions is smaller, with more samples concentrated at the median. With the incongruent data, the range is larger, with samples concentrated on the left side (a positive or right skewed distribution).

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

From R

```
> t.test(data1$Incongruent, data1$Congruent,paired=TRUE)

        Paired t-test

data:  data1$Incongruent and data1$Congruent
t = 8.0207, df = 23, p-value = 4.103e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  5.910555 10.019028
sample estimates:
mean of the differences
                7.964792
```

From Excel

t-Test: Paired Two Sample for Means

	Variable 1	*Variable 2*
Mean	22.01591667	14.05113
Variance	23.01175704	12.66903
Observations	24	24
Pearson Correlation	0.351819527	
Hypothesized Mean Difference	0	

df	23
t Stat	8.020706944
P(T<=t) one-tail	2.0515E-08
t Critical one-tail	1.713871528
P(T<=t) two-tail	4.103E-08
t Critical two-tail	2.06865761

What is your confidence level and your critical statistic value?

```
95 percent confidence interval:
  5.910555 10.019028
```

`t = 8.0207`

`t Critical two-tail:` 2.06865761

Do you reject the null hypothesis or fail to reject it?

`p-value = 4.103e-08`

p-value < 0.05, REJECT NULL HYPOTHESIS

also, t-value > t-crit.

Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

I conclude that we cannot reject the hypothesis that the mean of the dependent populations are different based on the samples we have drawn. I am 95% that the true difference in the mean of the two populations is between 5.9 and 10.0 seconds. These conclusions are subject to both Type I (false positive) and Type II (false negative) errors.

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!

The brain processes words significantly faster than colors. Plus, there is the fact that the brain automatically understands the meaning of words through habitual reading. A similar effect can be observed in the emotional stroop test where a person that is more depressed is likely to have a more difficult time suppressing irrelevant neutral words than subjects that are not depressed.

(Source: https://en.wikipedia.org/wiki/Stroop_effect)