Statistics: The Science of Decisions

Questions For Investigation

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

Independent Variable: Time (seconds).
Dependent: Frequency (Number of Test Subjects)

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

Null Hypothesis – that the mean completion time for congruent words is equal to the the mean completion time for incongruent words.

Alternate Hypothesis – that the mean completion time for congruent words is NOT equal to the mean completion time for incongruent words.

I expect to perform a two-sided, two-sample t-test assuming unequal variances because the of the small sample size (N<30), because I do not know if the distribution of the Incongruent words is normal, and because the distribution of the two samples looks different, so I am assuming unequal variances. 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.

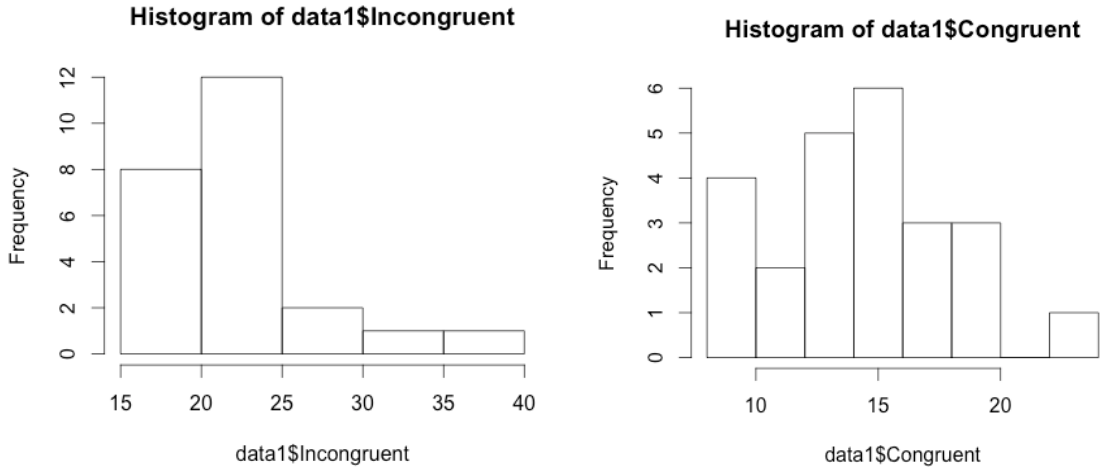
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

Using R – Psych package and describe(data1).

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

Using Excel:

t-Test: Two-Sample Assuming Unequal Variances

	Variable 1	*Variable 2*
Mean	14.051125	22.01591667

Variance	12.66902907	23.01175704
Observations	24	24
Hypothesized Mean Difference	0	
df	42	
t Stat	-6.532250554	
P(T<=t) one-tail	3.40518E-08	
t Critical one-tail	1.681952357	
P(T<=t) two-tail	6.81036E-08	
t Critical two-tail	2.018081703	

Using R:

```
Welch Two Sample t-test

data:  data1$Congruent and data1$Incongruent
t = -6.5323, df = 42.434, p-value = 6.51e-08
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -10.424698  -5.504885
sample estimates:
mean of x mean of y
 14.05113  22.01592
```

What is your confidence level and your critical statistic value?

```
95 percent confidence interval:
 -10.424698  -5.504885

t = -6.5323
```

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

```
p-value = 6.51e-08  REJECT NULL HYPOTHESIS
```

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

Conclude that means of the two sampled populations are not equal. This matches my expectations that the mean time for incongruent words is greater than that of congruent words.

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)