Project 1: Test a Perceptual Phenomenon

Questions for Investigation

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

Our independent variable is the congruency condition - whether the name of the color matches the ink color.
Our dependent variable is the time it takes to each participant to name the ink colors in equally-sized lists.

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

An appropriate set of hypotheses for this task could be as follows:
Null hypothesis: There is no significant difference between the time it takes participants in the Stroop test to complete the congruent task and the time it takes them to complete the incongruent task.
Alternative hypothesis: There is a significant difference between the time it takes participants in the Stroop test to complete the congruent task and the time it takes them to complete the incongruent task.

If we denote the population mean for the congruent task time performances as μ_c and the population mean for the incongruent task time performances as μ_i then the null hypothesis can be written as:
$H_o: \mu_c = \mu_i$
On the other hand the alternative hypothesis can be written as:
$H_a: \mu_c \neq \mu_i$

We will conduct a dependent t-test for paired samples. We have two dependent samples because our subject (people taking part on the test) took the Stroop test the first time with the congruent words condition and the second time with the incongruent words condition. Furthermore, we cannot conduct a z-test because we don't know the population parameters μ and σ for the congruent time results and the incongruent time results.

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

The sample mean for the congruent time results can be denoted as $\bar{x}_c$ and it is equal to:

$\bar{x}_c = 14.05$

Respectively, the sample mean for the incongruent time results can be denoted as $\bar{x}_i$ and it is equal to:

$\bar{x}_i = 22.02$

The sample variance for the congruent time results can be denoted as S^2_c and it is equal to:

$S^2_c = 12.67$

Taking the square root of the variance we can calculate the sample standard deviation for the congruent time results which is equal to:

$S_c = 3.56$

The sample variance for the incongruent time results can be denoted as S^2_i and it is equal to:

$S^2_i = 23.01$

Taking the square root of the variance we can calculate the sample standard deviation for the incongruent time results which is equal to:

$S_c = 4.80$

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.**
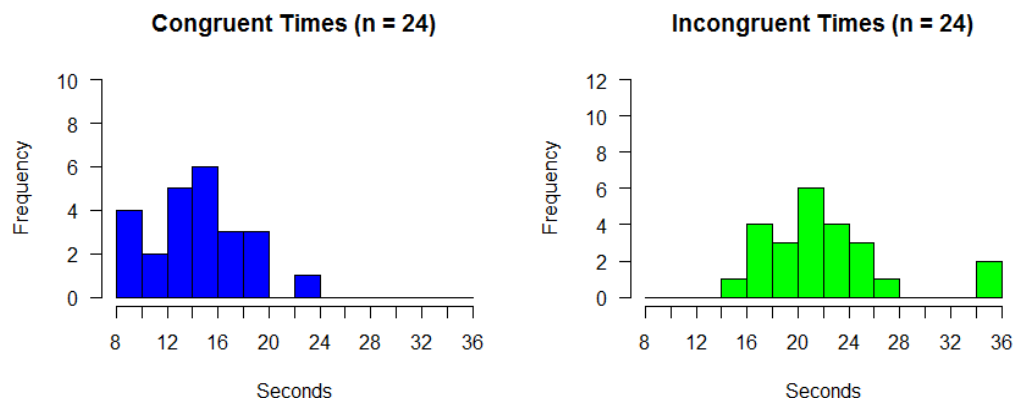
We can notice from the two histogram above that the incongruent times data set is more spread out, its range is larger and the majority of its observations are above 20 (for the congruent data set the majority of its observations lies below 20).

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?**

We will conduct a two tailed paired t-test. We will use a 95% confidence level. Our degrees of freedom for our statistical test is n-1=23. Our t-critical values are: ± 2.069

In order to conduct our t-test we need to compute the t-statistic for our sample data.
To compute the t statistic, we firstly calculate the mean difference (MD) between the two samples:

$$MD = \bar{x}_c - \bar{x}_i = -7.96$$

Then we need to calculate the sample standard deviation of differences (S_D) of the two samples:

$$S_D = 4.86$$

So our t-statistic is equal to:

$$t = \frac{MD}{S_D / \sqrt{n}} = -8.02$$

t(23)=-8.02, p= 0.00000004103
95 percent confidence interval:
 -10.019028, -5.910555

Since the p-value is less than 5%, we reject the null hypothesis that there is no significant difference between the time it takes participants in the Stroop

test to complete the congruent task and the time it takes them to complete the incongruent task in the 95% confidence level. With such a small p-value we can reject the null also in the 99% confidence level.

We can conclude that it takes significantly more time to participants on average to complete the incongruent task than the congruent task in the Stroop test.

The results match up with my expectations because from observing the histograms in question 4 we could expect that there is a significant difference between the time it takes to complete the two tasks.

References

1. Stroop effect- Wikipedia: https://en.wikipedia.org/wiki/Stroop_effect
2. R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.