p1: Statistics: The Science of Decisions Project

Tim Burke

27th February 2017

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

The independent variable is the type of test being conducted, whether the words displayed and their colour are conguent or incongruent. The dependent variable is the time taken for each particpent to complete the test.

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

For this task, an appropriate set of hypotheses that could be considered are:

- H_0 (the null hypothesis): there is no difference in population mean time taken between the two tests. ($\mu_c = \mu_i$)

- H_1 (the alternative hypothesis): there is a difference in the population mean time taken between the two tests. ($\mu_c \neq \mu_i$)

Where μ_c is the population mean time for the congruent test and μ_i is the population mean time taken for the incongruent test.

To test these hypotheses a paired two tailed t-test may be conducted.

The use of a paired t-test is appropriate as the data provided fulfills the following criteria and assumptions:

- the tests were both taken by the same participants.
- there are fewer than 30 samples.
- the population's standard deviation is unknown.
- the distributions are approximately normal (see Q4)

That there are fewer than 30 samples and an unknown population standard deviation would make a z-test inappropriate for this sample.

A two tailed test allows us to test for any difference in the mean times, not merely if one is greater than or less than the other without presuming that congruence of text colour and the word displayed has any effect either way.

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

One measure of central tendency is the mean. Where the word displayed was congruent with it's colour the mean time taken was **14.1** seconds, when the word and colour displayed were incongruent the mean time was greater, **22** seconds.

For measuring variability, we may consider the standard deviation of the times taken for each test. For the congruent test a standard deviation of **3.56** seconds was observed and in the incongruent test a greater variance of **4.8** seconds.

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.

The plot below shows that participants in the group took longer to complete the test when the words were displayed with an incongruent colour than when the word and colour of the text were the same.

It can also be seen that in the results for the incongruent test there are a couple of outlying values at the higher end of the time scale. This perhaps suggests these participants found the test particularly difficult when the words and the colour used to display them differed.
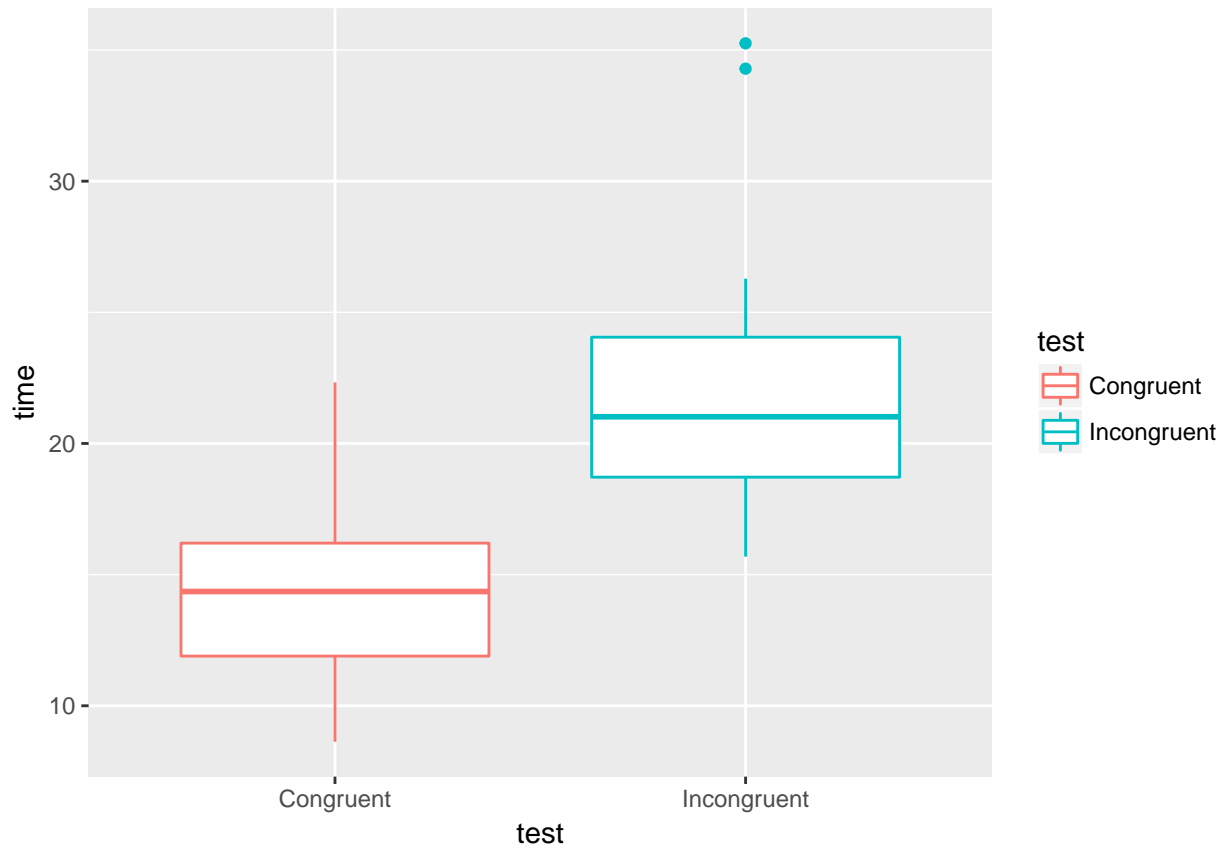
The histograms below also show that the time increased for the test when the words and text colour were incongruent rather than congruent but also that for both tests the times taken roughly form a normal distribution.
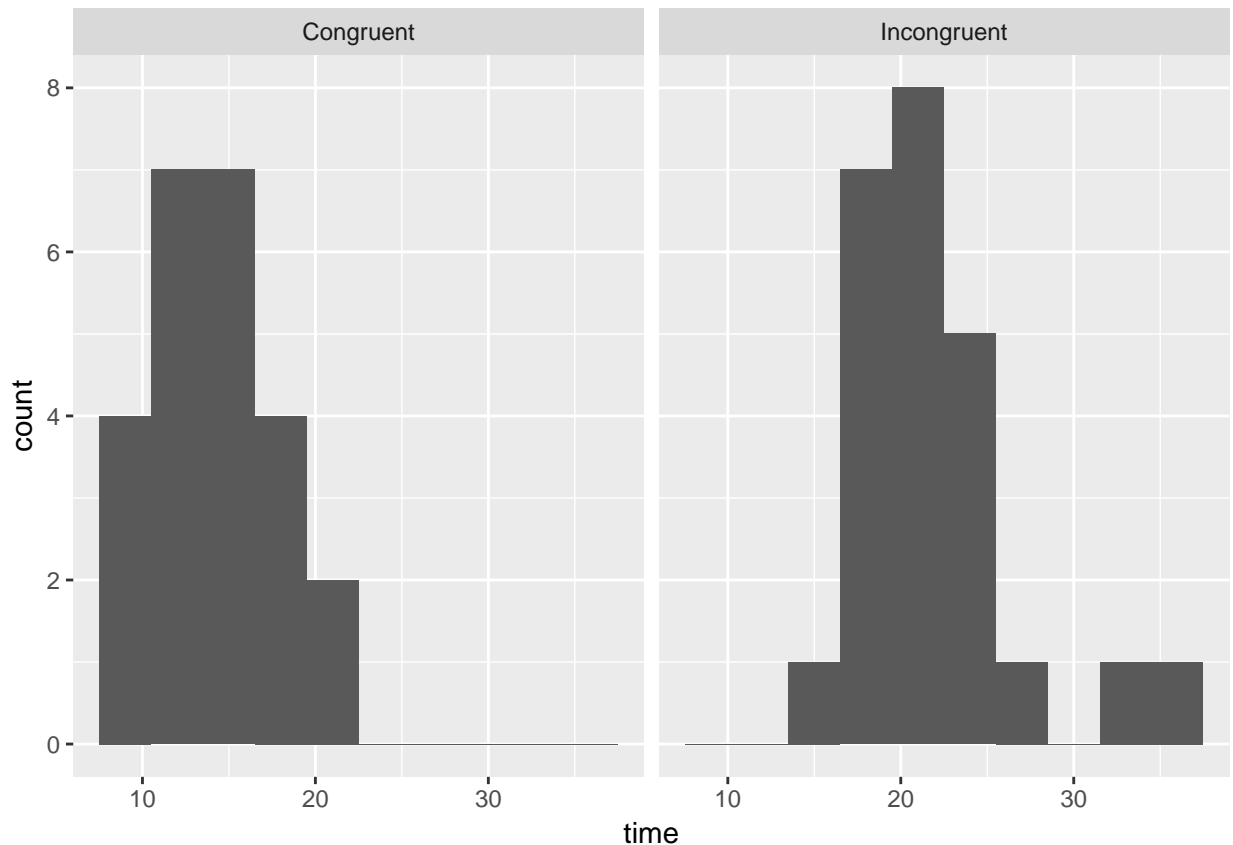
5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

```
stroop <- stroop %>%
  mutate(d_i = Congruent - Incongruent)

mean_time_difference <- stroop$d_i %>%
  mean()

sd_time_difference <- stroop$d_i %>%
  sd()

SE_mean_time_difference <- sd_time_difference / sqrt(24)

t_statistic <- mean_time_difference / SE_mean_time_difference
```
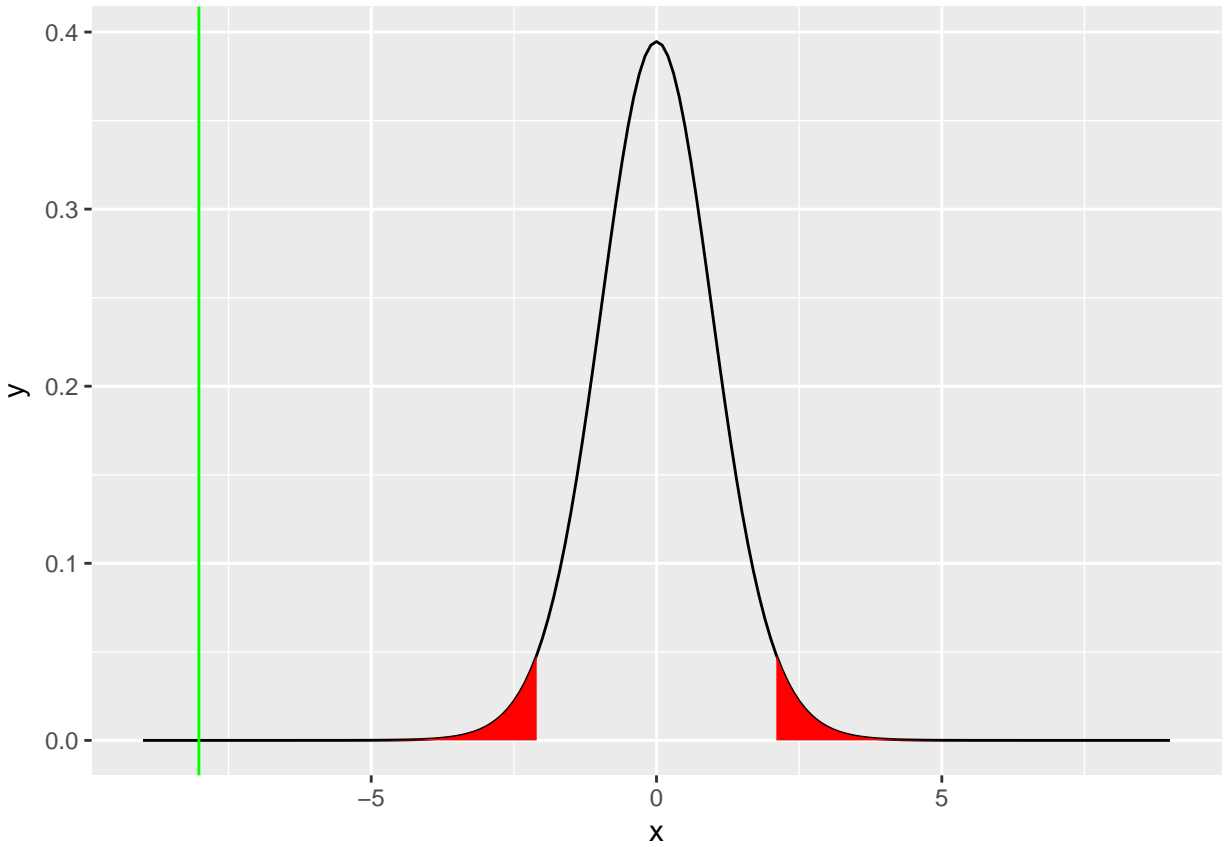
A paired two tailed t-test was performed (as suggested in reponse to part 2) to determine whether there was any difference in the mean time taken to complete the test if the colour of the text matched the word displayed or not. A *confidence level* of **95%** was used to obtain a *critical value* of $\pm$ **2.069** from a t-distribution table.

We can see the t-statistic (highlighted in green), T = -8.0207, lies inside the critical region (highlighted in red) and so the null hypothesis should be rejected; there is evidence that mean time taken to complete the test differs depending on whether the words used are congruent with the colour they are displayed in.