# Statistics: The Science of Decisions Project Instructions

1) What is are the variables?
   a. Independent
      i. 2 list of words where one is congruent whilst the other is incongruent
   b. Dependent
      i. Time taken to recite the word on each of the word list
2) What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.
   a. Hypothesis
      i. Null Hypothesis
         1. To determine whether the difference between population mean of time taken to read the incongruent list, $\mu_1$, and population mean of time taken to read congruent list, $\mu_2$, is equal to 0
         2. $\mu_1 - \mu_2 = 0$
         3. Time taken to read the incongruent list is NOT significantly different from that taken to read the congruent list
      ii. Alternative Hypothesis
         1. The difference between population mean of time taken to read the incongruent list, $\mu_1$, and population mean of time taken to read congruent list, $\mu_2$, not equal to 0
         2. $\mu_1 - \mu_2 \neq 0$
         3. Time taken to read the incongruent list is significantly different from that taken to read the congruent list
   b. I will expect to perform a t-test to see if the null hypothesis is true
      i. This is because
         1. The sample size is relatively small at 24
         2. Population standard deviation is unknown
         3. Furthermore, the 2 sets of readings taken from the same individuals hence they are controlled for the "language ability" of the participant
      ii. Assumptions
         1. $\mu_1 - \mu_2$ has a Gaussian-like distribution
      iii. This is because the 2 sets of readings though taken from the same individuals hence they are dependent and consistent across the "language ability" of the participant
3) Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.
   a. Mean of difference between congruent set and incongruent set is 7.964
   b. Standard derivation of difference between congruent set and incongruent set is 4.864
4) The scatter plot of the congruent vs incongruent numbers shows that the longer it takes for someone to read the congruent list, the more likely it will take that individual longer to read the incongruent list
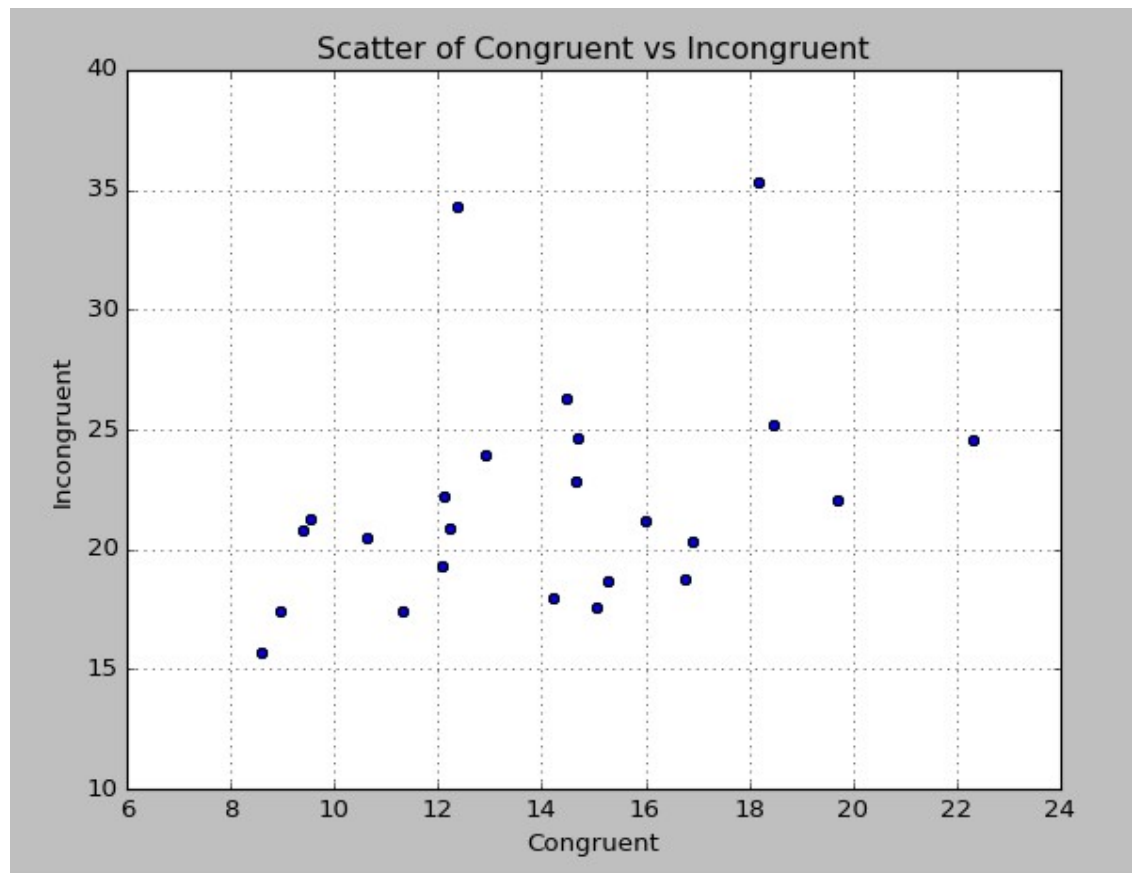
The difference between the time taken to complete the congruent list and incongruent
list is almost bimodal
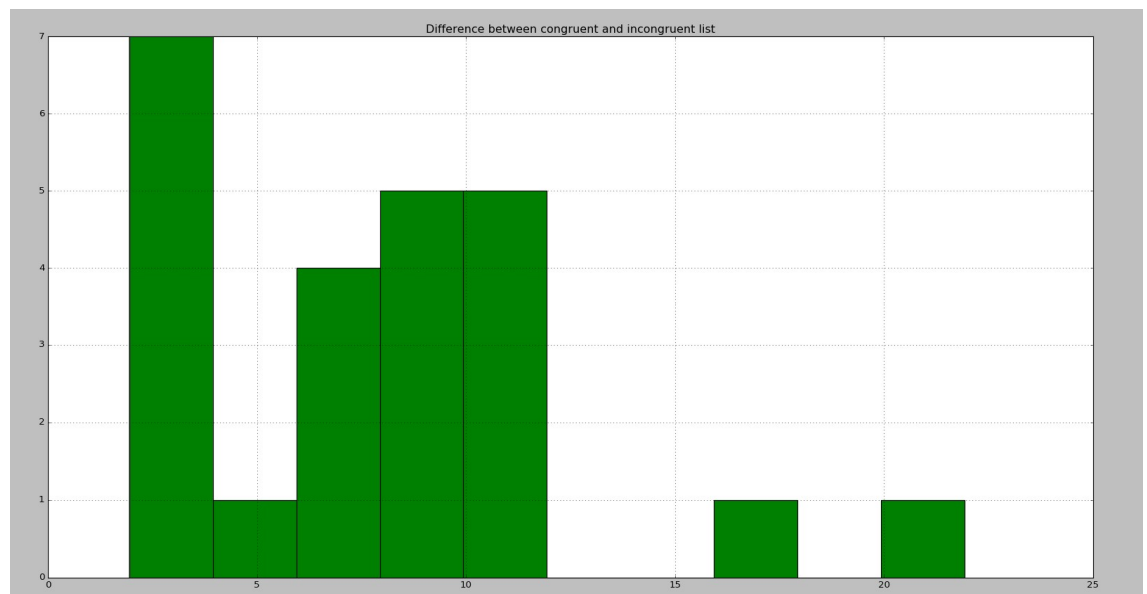


The time taken to finish a congruent and incongruent list is positively skewed where the
bulk of the participants are to the left of the graph with a long tail to the right.
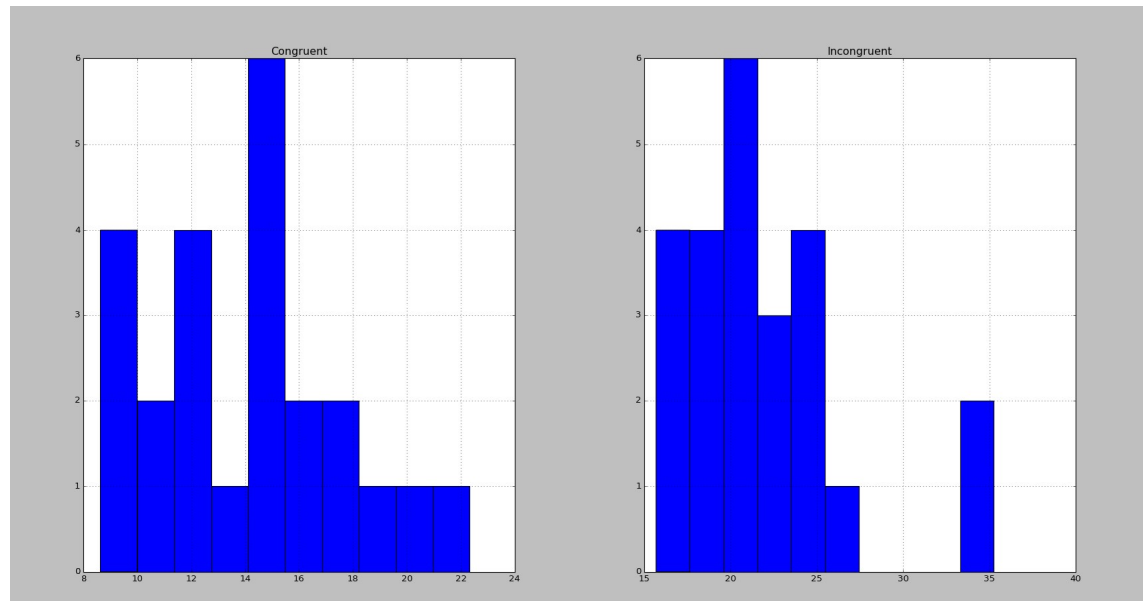
5) Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?
   a. t-statistic = 8.0207
   b. Degree of freedom = 23
   c. p value = less than 0.0001
   d. Critical t value = 2.8073 (99% confidence, 2 tail)
   e. We reject the null hypothesis and the time taken to read incongruent list and congruent list are significantly different.
   f. Yes the results did match my expectations
        i. It is more difficult to read the colour when it is incongruent between the font colour and the word
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!
   a. The difference between reading a word where the font colour and type is consistent comes more natural to people.
   b. An alternative test that can result in a similar effect is a test where instead of reading the word that has a incongruent font colour, participants will be asked to say the font colour with an incongruent word. (e.g. say "Blue" when they see the following phrase RED and "orange" when they see GREEN)