Project I

December 10, 2015

Xiaopeng Hu

Responses to Project Questions

1. a) The independent variable is words condition. This variable has two values here, *e.g.*, 'congruent' and 'incongruent'.
   b) The dependent variable is the completion time.

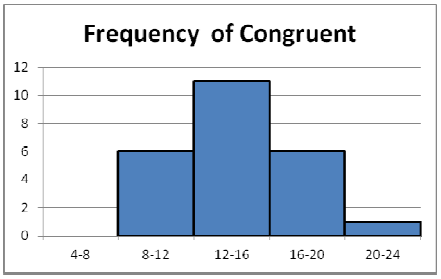
2. a) An appropriate set of hypotheses should be:
   Null hypothesis H_0:
   $\mu C = \mu I$,
   Completion times between incongruent and congruent conditions have no difference.
   Alternative hypothesis H_1:
   $\mu I > \mu C$,
   Completion time of incongruent condition is significantly longer than that of congruent condition.
   μC: completion time of congruent condition.
   μI: completion time of incongruent condition.

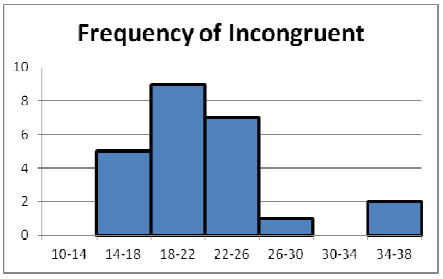
   b) A one-tailed, dependent t-test is suitable to perform a statistics analysis of this dataset.

   c) In this experiment, a sample dataset was collected with a group of 24 people under two conditions. Because the sample size is below 30 and the population standard deviation is unknown [1], a t sample test will be performed. Since the same group of people was tested repeatedly, this is a dependent t sample test. Furthermore, we also want to know if incongruent condition has significantly longer completion time than congruent condition, *e.g.,* we have a directional alternative hypothesis, so a one-tailed dependent t sample test will be employed.

3. a) For dataset of 'congruent', min is 8.630, max is 22.328, mean is 14.051, median is 14.357, Q1 is 11.895, Q3 is 16.20075, IQR is 4.306., sample variance is 12.669, sample standard deviation is 3.559
   b) For dataset of 'incongruent', min is 15.687, max is 35.255, mean is 22.016, median is 21.018, Q1 is 18.717, Q3 is 24.052, IQR is 5.335, sample variance is 23.012, sample standard deviation is 4.797.

4. Visualizations of the sample data. From these two plots, we can say that the distributions are probably normal distribution, but the second one could be a right skewed one. We also can estimate that the mode for dataset of 'congruent' is around 14, while for another dataset is around 20.





5. a) For a one-tailed dependent t-test, a *P* value of 0.05 is used;
   b) It is a dependent t-test and we have 24 testers, so *df* = 23;
   c) According t-score table, for one-tailed t-test, at *df* of 23 and *P* of 0.05, the $t_{\#}$ is 1.714.
   d) To calculate the t value:

con	incon	D(incon-con)
12.079	19.278	7.199
16.791	18.741	1.95
9.564	21.214	11.65
8.63	15.687	7.057
14.669	22.803	8.134
12.238	20.878	8.64
14.692	24.572	9.88
8.987	17.394	8.407
9.401	20.762	11.361
14.48	26.282	11.802
22.328	24.524	2.196
15.298	18.644	3.346
15.073	17.51	2.437
16.929	20.33	3.401
18.2	35.255	17.055
12.13	22.158	10.028
18.495	25.139	6.644
10.639	20.429	9.79
11.344	17.425	6.081
12.369	34.288	21.919
12.944	23.894	10.95
14.233	17.96	3.727
19.71	22.058	2.348

16.004	21.157	5.153

Mean-D = 7.965
Stdev of D = 4.865
SE=(Stdev of D)/sqrt(n) = 4.865/sqrt(24)= 0.993
t = (Mean-D)/SE= 8.021
Therefore we have:
t(23)= 8.021, *P*<.05, one-tailed.
Because this t value locates in the critical region, so we can reject the null hypothesis. This statistic result means that under incongruent word condition, people tend to spend more time to name the ink colors in word-list than under incongruent word condition. This result matches up with my expectations.

6. a) Apparently, the mismatch of the name of color words and the color in which the words are printed will cause confusion on people who read the word list, they need more time to distinguish what they see and what they think, then make a right decision to read the name of the words.
   b) 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!

Reference

1. http://www.statisticshowto.com/when-to-use-a-t-score-vs-z-score/