TEST A PERCEPTUAL PHENOMENON

GRADING PAPER FOR THE COURSE "STATISTICS: THE SCIENCE OF DECISIONS", PART OF UDACITY'S NANODEGREE

Author: Nikita Barsukov, nikita@barsukov.net

INTRODUCTION

This is a write-up for the first course in Udacity's data analyst nanodegree. Here I was required to calculate basic descriptive statistics and perform hypothesis testing on a small dataset representing result from Stroop experiment, a psychological test.

QUESTIONS FOR INVESTIGATION

WHAT IS DEPENDENT AND INDEPENDENT VARIABLE?

There are two measurements in the dataset: "Congruent" and "Incongruent", representing times to complete tasks for words that match the font color, and words that do not match color. Independent variable is whether colors matched words, or did not (congruency of words). Dependent variable here is the time completed by a person to name colors of all the words.

WHAT ARE APPROPRIATE SET OF HYPOTHESIS?

Dataset contains time that it took to name all the colors of words for a situation when colors were the same as words ("Congruent" variable), and when color was different that words ("Incongruent" variable). What we're trying to infer from that dataset is the true difference between these two variables for the **whole population**, not just between these two samples.

Our **null hypothesis** can be "True difference in means of our variables for the entire population *equals 0* when words matched word color, and when words were not the same as their color". **Alternative hypothesis** is "True difference in means of our variables for the entire population is *less than 0* when words matched word color, and when words were not the same as their color".

Let's determine which statistical test we will be using in this paper. Since the dataset is rather small, 24 observations in each dataset, we should be cautious in strictly determining if the distribution of our variables is normal or not. Creating histogram obviously would be meaningless with such a small sample. Let's construct Q-Q plot to see if our distribution is close to normal or not. Dots on both plots in on Figure 1 below are rather close to the straight line. Thus we could accept that the distribution of both our variables is close to normal.

With normally distributed variables either z-test or t-tests are used. Since we do not know standard deviation of our variables for the whole population, and our sample sizes are less than 30, we should use t-test for our hypothesis testing. In addition to that, our alternative hypothesis is about one variable being less than another. Thus we should use one-tail t-test.

Additionally since same people first named colors for congruent words, and then for incongruent words, we're having a case for **paired one-tail t-test.**
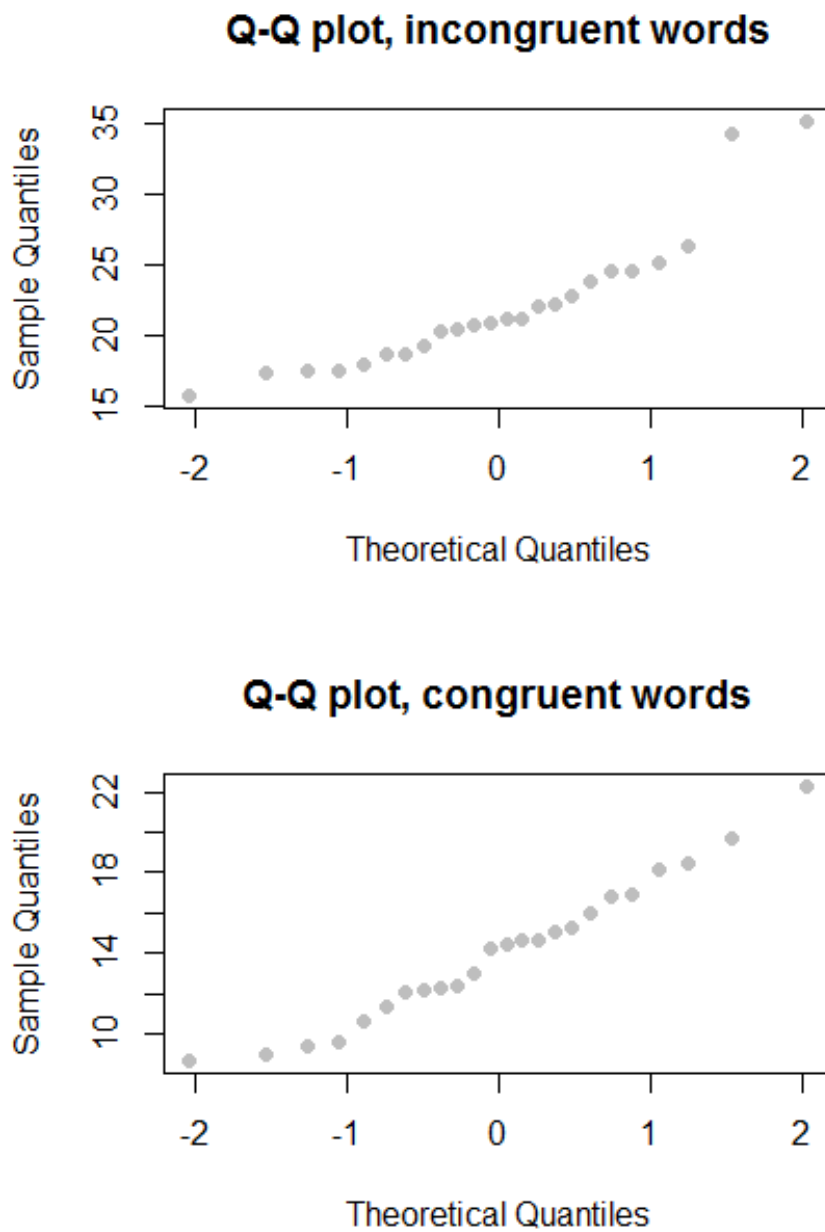
Q-Q plot, incongruent words

Q-Q plot, congruent words

FIGURE 1

DESCRIPTIVE STATISTICS

Mean and standard deviation of two variables are displayed below:

Variable	Mean	Median	Standard Deviation
Congruent	14.05113	14.3565	3.559358
Incongruent	22.01592	21.0175	4.797057

Indeed, mean and median for both variables are rather close, which is an additional indicator that distribution of these two variables is close to normal.

DATA VISUALIZATIONS

I used R programming language to generate boxplot and scatterplot of two variables. As we can see from boxplots in Figure 2 below, distributions differ visibly between each other. On the other hand number of observation is rather small, 24 for both variables. This may lead to inconclusive results from out statistical tests which I aim to perform in next chapter.

Scatterplot on Figure 3 below strongly suggest that these two variables have week correlation with each other. However correlation analysis of these two variables is beyond the scope of this paper.
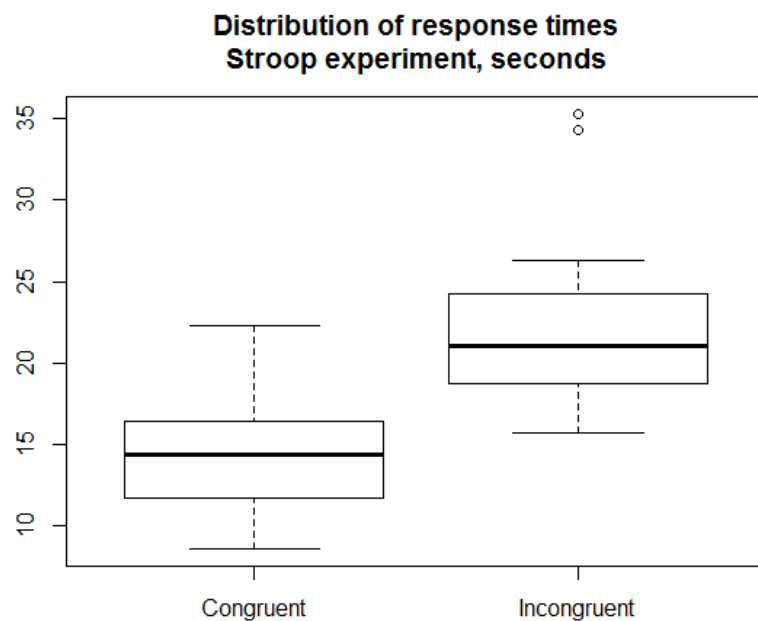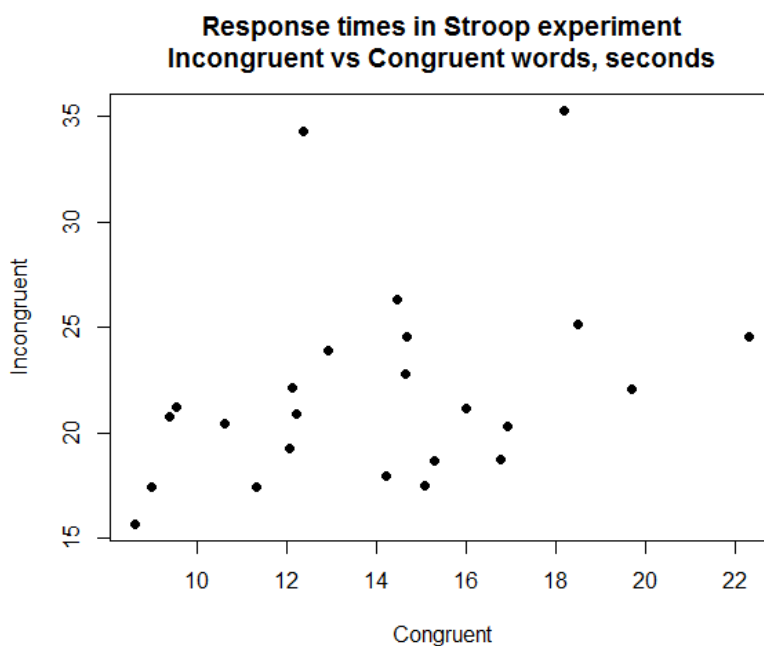
HYPOTHESIS TESTING

One-sided paired t-test with confidence level 99% gave us following results:

Parameter	Value
t-score	-8.0207
degrees of freedom	23
critical t-score at 99%	-2.499867
p-value	~ 0.0000020515 %
99% confidence interval	$-\infty - -5.482352$

We can clearly see that with these results we can **reject null hypothesis** at 99% confidence level. This means that it is almost certain that it takes longer time to correctly name all the colors of words, when words do not match their colors. This corroborates with the original description of this psychological experiment.

SOURCES

1. R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
2. https://en.wikipedia.org/wiki/Normality_test
3. https://en.wikipedia.org/wiki/Q%E2%80%93Q_plot
4. http://stackoverflow.com/questions/11526041/critical-t-values-in-r
5. http://www.statmethods.net/advgraphs/layout.html