UDACITY NANODEGREE PROGRAM

Data Analyst

Project 1: Test A Perceptual Phenomenon

Author: Artem Oppermann

Oldenburg, July 24, 2017

Contents

1 Introduction

In this project I analyze two dependent samples results from the Stroop Effect experiment. The participants in the experiment have to name the color of the ink of a list of words shown on the screen under two different conditions. For the first sample the participants see the words that are congruent with the color of the ink. For the second sample the words are incongruent with their color of ink. The time it takes to finish each test is recorded. The samples contain the measured time for 24 participants.

2 Variables

The dependent variable in this project is the time it takes to name the ink colors. The both task conditions (congruent words and incongruent words condition) are the independent variables.

3 Hypothesis

Let‘s assume μ_0 is the mean-time it takes to finish the tast with congruent words condition and μ_A the mean-time it takes to finish the task with incongruent words condition.

- The null-hypothesis: There is no significant statistical difference between μ_0 and μ_A, $\mathrm{H}_0 : \mu_0 = \mu_A$
- The alternative-hypothesis: The task under the more difficult condition (incongruent words condition) significantly increase the time-mean it takes to finish the task, $\mathrm{H}_A : \mu_A > \mu_0$

Based on the assumption that μ_A is expected to be higher than μ_0 and the fact that the dataset set contains two dependent samples I will perform a positive one-tail t-test later in this project.

4 Descriptive statistics

In the following task I perform some descriptive statistics of the dataset in regards of both conditions. Each sample contains $\mathrm{N} = 24$ time-measurements x_i.

4.1 Congruent Words Condition

Sample mean $\mu = \frac{1}{N}\sum_{n=1}^{N} x_n = 14.051$ sec.
Median = 14.357 sec.
Corrected sample standard deviation $\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \mu)^2} = 3.558$ sec.

4.2 Incongruent Words Condition

Sample mean $\mu = \frac{1}{N}\sum_{n=1}^{N} x_n = 22.016$ sec.
Median = 21.016 sec.
Corrected sample standard deviation $\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \mu)^2} = 4.796$ sec.

5 Data Visualisation

In figure 5.1 we see a histogram of measured time under congruent test condition. The peak of frequencys can be observed around the calculated mean-time of $\mu = 14.051$ sec.

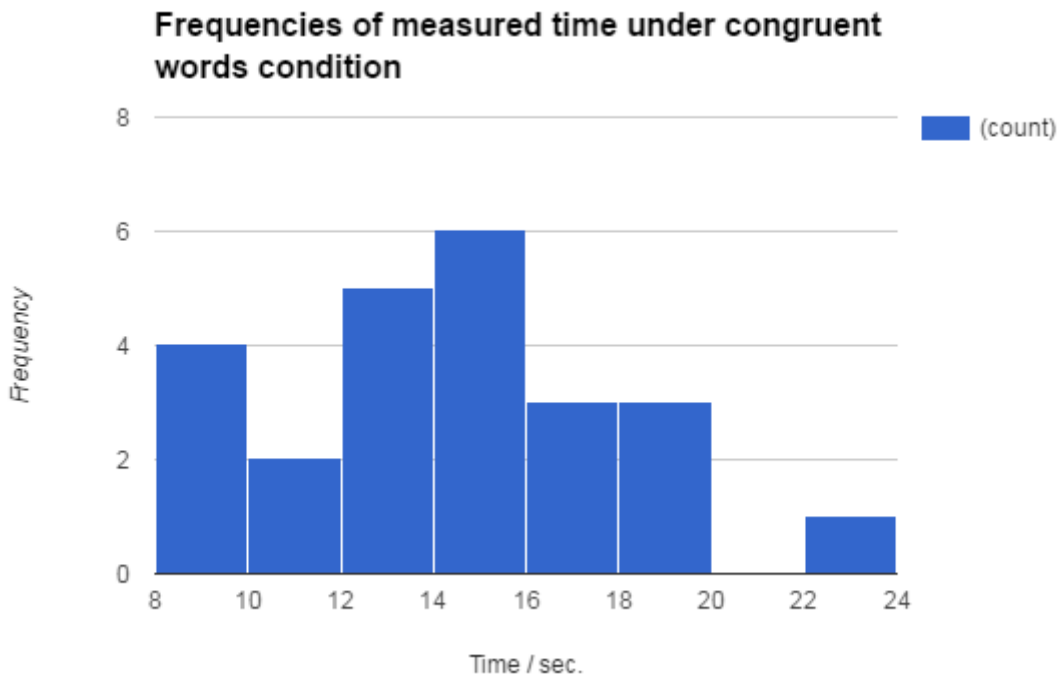

Figure 5.1: Measured time frequencys under congruent test condition, bin-size=2 sec. .

In figure 5.2 we see a histogram of measured time under incongruent test condition. The distribution can roughly be approximated as a normal distri-

bution with the mean around 22.016 sec, according to the previous section. After 34 sec. some outliers can be observed.

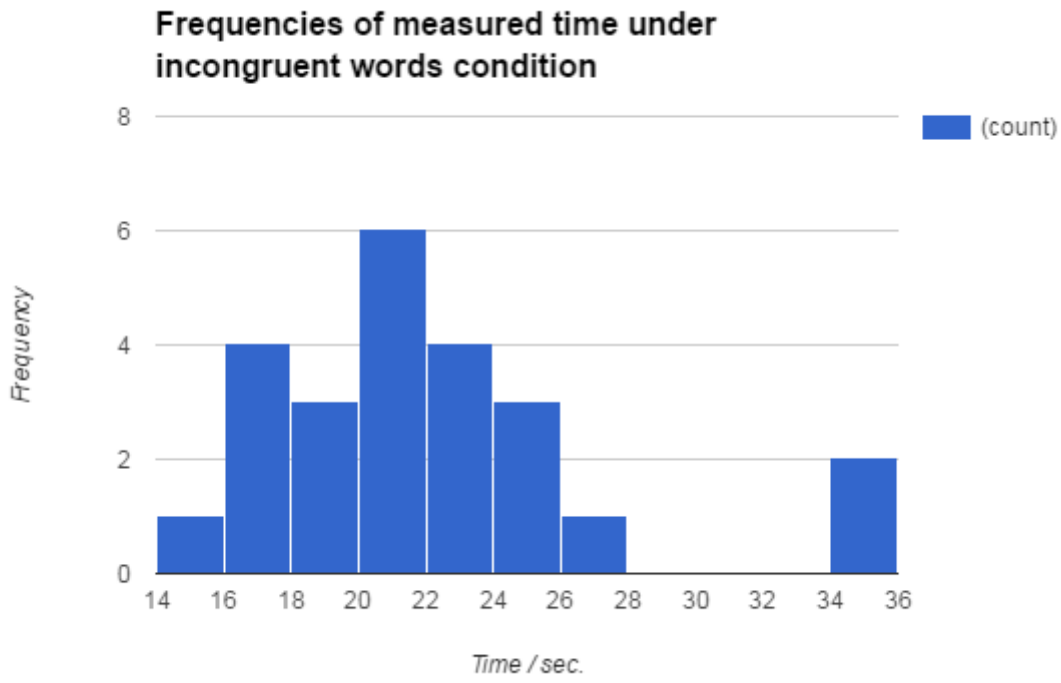

Figure 5.2: Measured time frequencys under incongruent test condition, binsize=2 sec. .

6 t-test

For the following t-test I choose an α-value of α=0.05. In the one-tailed t-test it corresponds to a convidence level of 90%. Each of the two samples consist of n = 24 participants that corresponds to df = n – 1 = 23 degrees of freedom. Based on these values I get the critical t-value of $t_{critical} = 1.714$ from [1].

I calculate the t-statistic according to $t = \frac{\mu_a - \mu_0}{\frac{s}{\sqrt{N}}}$ with s = 4.864 sec. as the standard deviation of the time differences. I get the value of t = 8.029.
Because of $t > t_{critical}$ I reject the null-hypothesis. In other words the probability of getting a sample mean of $\mu_A = 22.016$ sec. from a population with the mean of $\mu_0 = 14.051$ sec. is very small. According to [1] the p-value must be less than 0.0005.

References

[1] http://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf