P6Project: Test a Perceptual Phenomenon

Sara Uenoyama

Questions: https://docs.google.com/document/d/1-0kpZLjG_kX9J6LIQ5IltsqMzVWjh36QpnP2RYpVdPU/pub?embedded=True

Dataset: https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view

1. Our Independent variable and dependent variable
   In this case, our independent variable is the two conditions of words: congruent or incongruent. Our dependent variable is the time it takes for participants to name the ink colors of words on a list.

2. Appropriate hypotheses and statistical test of this task
   Appropriate null hypothesis is that there is no difference in population mean time to name ink colors between congruent words list, μ_1 and incongruent words list, μ_2. On the other hand, alternative hypothesis is that there is a statistical significant difference between the two. We can write down as follows in mathematical way.
    - Null hypotheses: $\mu_1 = \mu_2$
    - Alternative hypotheses: $\mu_1 \neq \mu_2$

   Appropriate test is two-sided t-test for following reasons.
    - We expect that there is a difference in time to read congruent words list vs incongruent words list.
    - We only have sample data, not knowing μ and σ, which is used for z-test.

   Assumptions made for the t-test are
    - The scale of measurement applied to this data collection is continuous scale, "time".
    - The data is collected from randomly selected people from total population.
    - As showed in my answer to question 4, the sample data is approximately normally distributed.
    - The sample size, 24, is merely or not sufficient to draw bell-shaped distribution curve.

3. Descriptive statistics
   Calculating from the dataset, sample mean and sample standard distribution from each group are
    - $\bar{x}_A$ = 14.05
    - s_A = 3.56
    - $\bar{y}_B$ = 22.02
    - s_B = 4.8
   A stands for sample data of congruent words and *B* stands for sample data of incongruent words.

4. Visualization of the distribution of the sample data
   The sampling distribution of congruent words data is normal distribution as showed below.
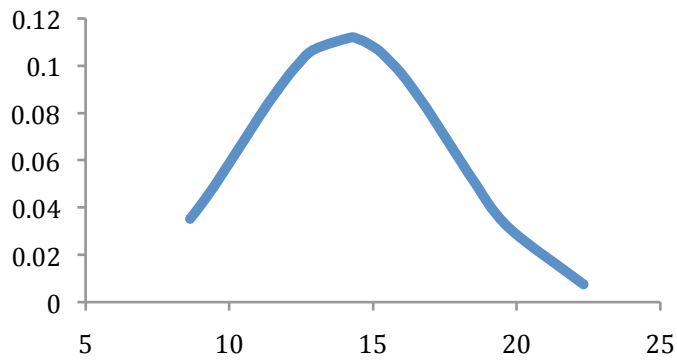
The sample data of incongruent words is normally distributed as well.
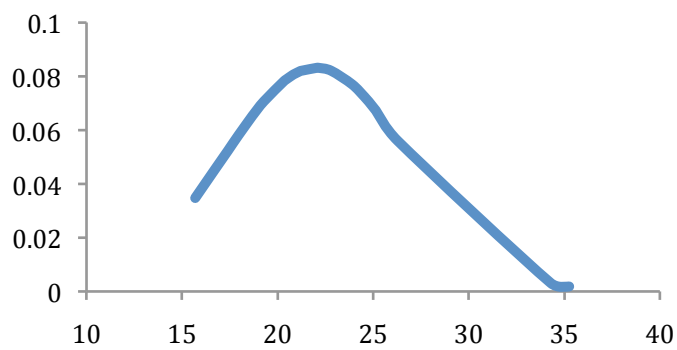
5. Two-tailed t-test
   Calculate $d_i = y_i - x_i$ for each sample data.

	Congruent	Incongruent	di
1	12.079	19.278	7.199
2	16.791	18.741	1.95
3	9.564	21.214	11.65
4	8.63	15.687	7.057
5	14.669	22.803	8.134
6	12.238	20.878	8.64
7	14.692	24.572	9.88
8	8.987	17.394	8.407
9	9.401	20.762	11.361
10	14.48	26.282	11.802
11	22.328	24.524	2.196
12	15.298	18.644	3.346
13	15.073	17.51	2.437
14	16.929	20.33	3.401
15	18.2	35.255	17.055
16	12.13	22.158	10.028
17	18.495	25.139	6.644
18	10.639	20.429	9.79
19	11.344	17.425	6.081
20	12.369	34.288	21.919
21	12.944	23.894	10.95
22	14.233	17.96	3.727
23	19.71	22.058	2.348
24	16.004	21.157	5.153
mean	14.05	22.02	7.96
sd	3.56	4.8	4.86

Mean difference đ is

$$đ = \bar{y}_B - \bar{x}_A = 7.96$$

Standard deviation of the difference is

$$s_d = 4.86$$

Since the sample size is n = 24, degree of freedom is

$$Df = 24 - 1 = 23$$

Standard error of mean difference can be calculated from the number of samples and the sample standard deviations.

$$SE(đ) = s_d / \sqrt{n} = 0.99$$

Using the SEM value, we can calculate t-statics.

$$T\text{-statistics} = đ / SE(đ) = 8.03$$

When α level is 0.05, t-critical value is 2.069.

$$t(23) = \pm 2.069, p = 0.025, \text{two-tailed}$$

Since t-statistics > t (23), we reject the null hypothesis. $p < 0.025$.

In conclusion, in a Stroop effect experiment, the time it takes to name ink color of congruent word and one of incongruent word are different. As we expected, it takes more time to name ink color of words on incongruent words list.

Reference
https://www.investopedia.com/ask/answers/073115/what-assumptions-are-made-when-conducting-ttest.asp
http://www.statstutor.ac.uk/resources/uploaded/paired-t-test.pdf
https://www.rapidtables.com/math/symbols/Statistical_Symbols.html