Uda DANDb P1 - Stroop Effect

Author by Simplato; Date 2017/6/6

1. Variables

- In this Stroop Test, our independent variable is "the two conditions", that is to say **a congruent words condition** and **an incongruent words condition**.
- And our dependent variable is **the time it takes to name the ink colors**.

2. Hypotheses & Type of test

- This test can be considered to be a two-tailed one, due to unpredictability of a direction of the treatment effects. Accordingly, the best null hypothesis is **"A congruent words condition and an incongruent words condition have the same effect on the react time of participants."** To represent the population mean, we use the symbol μ . That way, μ_C is the population mean of time under a congruent words condition, and μ_I is the population mean of time under an incongruent words condition. Then here is the null hypothesis(H_0):

  $$H_0: \mu_C = \mu_I$$

  And the alternative hypothesis (H_A) is **"A congruent words condition and an incongruent words condition have different effects on the react time of participants."** Also, we have a formula like this:

  $$H_A: \mu_C \neq \mu_I$$

- This is a **dependent t-test for paired samples**, which can be inferred from the fact that the participants are measured on the same dependent variable (time) under two different conditions (a congruent and an incongruent words condition).
  A t-test should be adopted rather than a z-test, because of the two factors:
  a) **Unknown population standard deviation;**
  b) **Small sample size.** (In our test, the sample size n = 24 < 30, then a t-test is more suitable.)
  Actually, we must know the standard deviation of the population and our sample size should be above 30 in order for us to be able to use a z-test. Otherwise, we use the t-test.

3. Report the descriptive statistics

- **Sample mean** (μ_C is the sample mean of time under a congruent words condition , μ_I is the sample mean of time under an incongruent words condition.) & **Population mean**:

  According to the Central Limit Theorem, the mean of all samples from the sample population will be approximately equal to the mean of the population, this mean that

$$\mu_C = \bar{X}_C = 14.05 \; ; \; \mu_I = \bar{X}_I = 22.02$$

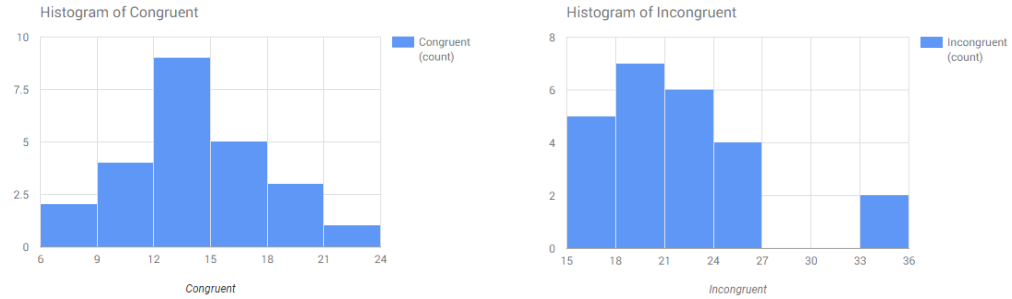
- **Point estimate for $\mu_C - \mu_I$** :

$$M_D = \mu_C - \mu_I = \bar{X}_C - \bar{X}_I = 14.05 - 22.02 = -7.96$$

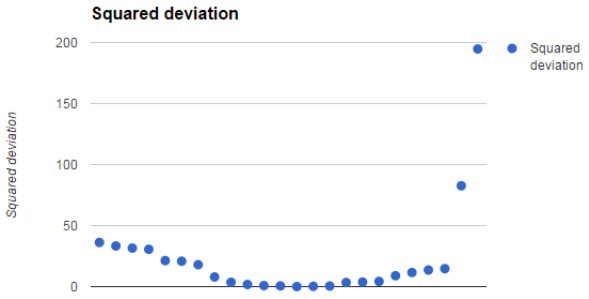
- **Standard deviation of differences (S_D) & Standard error of differences (SE)** :

$$S_D = \sqrt{\frac{\sum(X_C - X_I)^2}{n-1}} = 4.86; \; SE = \frac{S_D}{\sqrt{n}} = .99$$

4. Visualization & Observation



Based on the given sample data of the two original columns, the two histograms above show the distributions of them. It seems that both of them are **positively skewed distributions.** The first distribution, of course, approaches to be normal distribution. Some blanks appear in the second distribution, which may imply there are some **outliers** that strongly bias the mean.



The scatter plot of the squared deviation of difference shows the last two pieces of data should be the very outliers.

5. t-Critical, Reject the null & Conclusion

- **t-Statistic (t):**

$$\text{t-statistic} = \frac{M_D}{SE} = -8.02$$

- **Confidence Level** : 5%
- **Degrees of Freedom (df)**: df = n – 1 = 23
- **t-Critical** :

$$\text{t-critical} = t(23) = \pm 2.069,\ p < .05,\ \text{two-tailed}$$

- **When to reject the null & Conclusion**:

$$\text{t-statistic} < -2.069;$$

So we should **reject the null**, that is to say **"A congruent words condition and an incongruent words condition have different effects on the react time of participants."** The results, I think, perfectly match up my expectations, since the incongruent words would definitely keep the participants from naming the ink colors directly.

- **P-Value** :

The two-tailed P value is less than 0.0001. By conventional criteria, this difference is considered to be extremely statistically significant.

- **Confidence interval (CI) on the mean difference** :

Margin of error = t-critical * SE = -2.05

Lower bound = M_D + Margin of error = -10.02

Upper bound = M_D + Margin of error = -5.91

95% CI : $M_D \pm$ Margin of error = (-10.02, -5.91)

- **Cohen's d** :

$$\text{Cohen's d} = \frac{M_D}{S_D} = -1.64$$

- **R-squared** :

$$r^2 = \frac{t^2}{t^2 + df} = .74$$

Link to Simplato's Stroop Data