Design a stroop effect test

SongTang

9/16/2017

Backgroung Information

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the color of the ink in which the word is printed. The task has two conditions: a congruent words condition, and an incongruent words condition.

- In the congruent words condition, the words being displayed are color words whose names match the colors in which they are printed: for example RED, BLUE.
- In the incongruent words condition, the words displayed are color words whose names do not match the colors in which they are printed: for example PURPLE, ORANGE.

In each case, we measure the time it takes to name the ink colors in equally-sized lists. Each participant will go through and record a time from each condition.

1. What is our independent variable? What is our dependent variable?

Answer:

The independent variable is the congruency condition and incongruency condition - whether the name of the color matches with the ink color or not. The dependent variable is the time it takes to name the ink colors in equally-sized lists.

2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

Answer:

The hypothesis test that can use is: the difference between the time used to identify the colors under congruent words condition and incongruent words condition, namely, the Stroop Effect is in existence.

To be more specifically, here I will refer to the population means of congruence words group and incongruence words group - average times for the respective groups to recognize the colors. By comparing the mean time of two groups, we can find that there the difference exists between the two groups' recognization color time. But the fact is we are not able to collect all the data about time reaction of stroop effect, or do the experiment with all potential subjects in the world. Base on this condition, we need to work with the sample we have to make inferential statistic about the population means, for example, to use the observation means, standard deviation and other statistics to infer about the population means. In this case, the observation is the difference between the two groups' reaction times. With this observation, we can construct new statistics such as means and standard errors.

To achieve this, We will use T-test to verify. T-test assesses whether the means of two groups are statistically different from each other. First of all, we need to address the uncertainty in sample standard error resulted from the unknown population standard deviation; two, we are comparing the means of two groups that are dependent; three, the same subject is involved under both conditions.

Below is the hypothesis to test:

H0: mu_diff = 0 (The real difference between group population means is zero)

HA: mu_diff != 0 (The real difference between group population means is not zero)

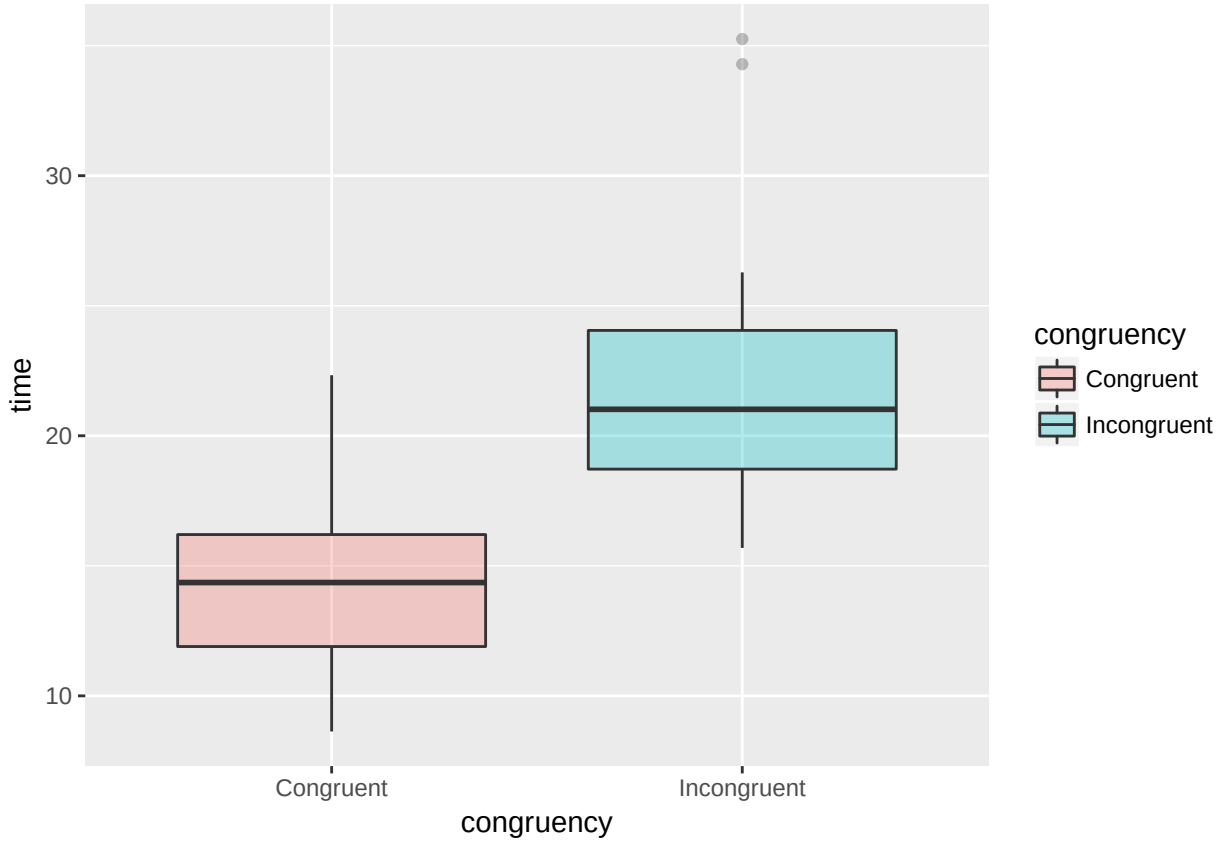
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

```r
# Read in the data
library(readr)
dat <- read.csv("stroopdata.csv")
# Tidy up the data for later analysis
library(tidyr); suppressMessages(library(dplyr))
# Add a column identifying subjects
dat.subject <- mutate(dat, subject = 1:nrow(dat))
# Tidy up data by keeping one variable in one column
tidy.dat <- gather(dat.subject, congruency, time, -subject)
# Calculate the average time for both groups
tidy.dat %>%
    group_by(congruency) %>%
    summarise(round(mean(time), 2), round(median(time), 2), round(sd(time), 2), round(var(time), 2))
```

```
## # A tibble: 2 x 5
##    congruency `round(mean(time), 2)` `round(median(time), 2)`
##         <chr>                  <dbl>                    <dbl>
## 1   Congruent                  14.05                    14.36
## 2 Incongruent                  22.02                    21.02
## # ... with 2 more variables: `round(sd(time), 2)` <dbl>, `round(var(time),
## #   2)` <dbl>
```

4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

```r
library(ggplot2)
b <- ggplot(tidy.dat, aes(y = time, x = congruency, fill = congruency))
b + geom_boxplot(alpha=0.3)
```



Answer

We can conclue from the above plot