Untitled

Kinjal

2 October 2017

Title : Analysis of Stroops file

About Stroop's effect ?

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. Identify the independent and dependent variables in the experiment.

Understanding types of variables : Independent variables are experimental or predictor variables whereas dependent variables are final outcomes.

Based upon above explanation

Independent variables are : Word Condition of congruency/incrogruency.

Dependent Variables is : Time in seconds taken to read out the words correctly.

2. What is the Null and Alternate hypothesis for the experiment.

Null hypothesis : Time in seconds taken to read out incongruent words will be less or same than that of congruent words.Hence there will be no effect of incongruency. H0: mui <= muc (mui - population mean of incongruent values, muc - population mean of congruent values)

Alternate hypotesis : Time in seconds taken to read out incongruent words will be more of than that of congruent words.Hence there will be effect of incongruency. HA: mui > muc (mui - population mean of incongruent values, muc - population mean of congruent values)

3. Statistical test used.

I have used a t-test over z test because population parameters are not know in our dataset and sample size is also small. I have decided to choose a one tailed test is right direction, because my alternate hypothesis will have incongruent reading time more than congruent reading time.Hence I have I want to analyse my direction of test. The test will be repeated order dependent sample test because the test is within subject design , where each subject is assinged two conditions in random order. I have choosen 95% of confidence interval than using 98% confidence interval because looking at the size of the dataset , we are satisfied with lower chance of creating confidence , whereas a 98% confidence interval would have required greater sample size.

Reading the csv file

```
stroop <- read.csv("C:\\Users\\KACHI\\Desktop\\stroopdata.csv")
```

4. Descriptive statistics, including at least one measure of centrality and one measure of variability, have been computed for the dataset's groups.

Below the mean , median , standard deviation(sd) and variance has been created.

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

```
## # A tibble: 2 x 5
##    congruency `mean(time)` `median(time)` `sd(time)` `var(time)`
##         <chr>        <dbl>          <dbl>      <dbl>       <dbl>
## 1   Congruent     14.05113        14.3565   3.559358    12.66903
## 2 Incongruent     22.01592        21.0175   4.797057    23.01176
```

5. One or two visualizations have been created that show off the data, including comments on what can be observed in the plot or plots.

```
stroop$diff <- stroop$Incongruent - stroop$Congruent
```

Diff is a variable that has difference values of congruent and incongruent columns.

Below is the ggplot of difference along with mean , median amd 10 % to 90 % interval.

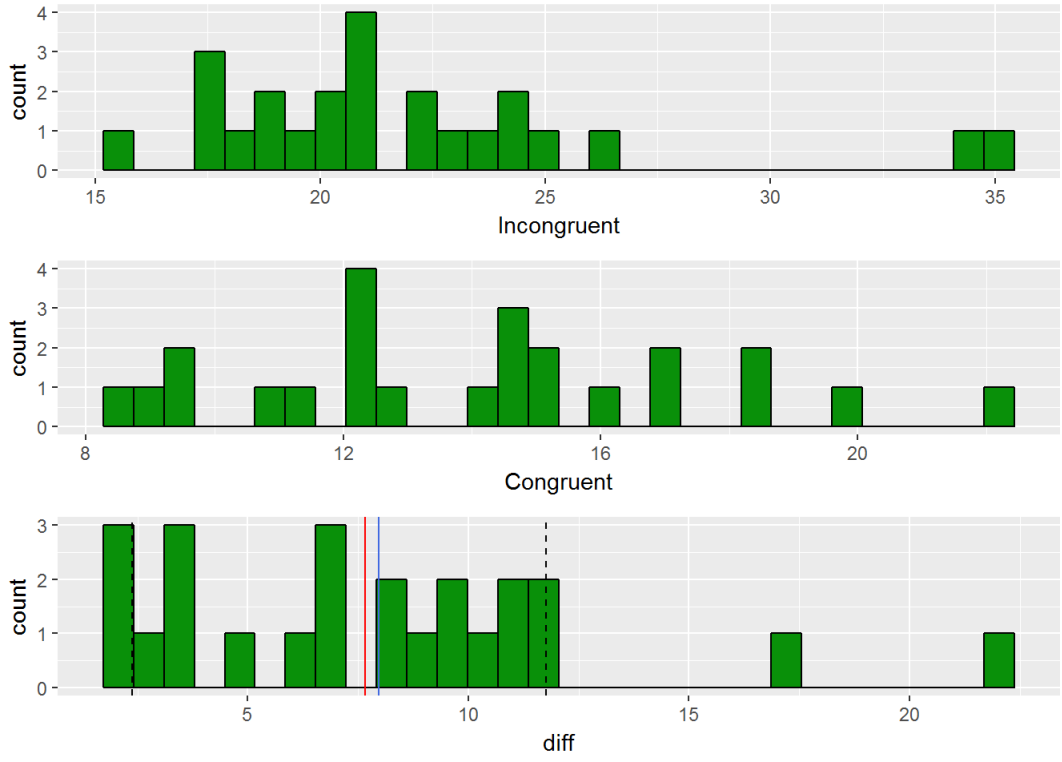
```
library(ggplot2)
library(gridExtra)

q1 <- ggplot(aes(x=diff),
       data =  stroop)+
   geom_histogram(color =I('black'),
                  fill = I('#099009'))+
   geom_vline(xintercept = mean(stroop$diff),
              col = "royalblue")+
   geom_vline(xintercept = median(stroop$diff),
              col = "red")+
   geom_vline(xintercept = quantile(probs = 0.9 ,
                                    stroop$diff),
              linetype = 2)+
   geom_vline(xintercept = quantile(probs = 0.1 ,
                                    stroop$diff),
              linetype = 2)

q2 <- ggplot(aes(x= Congruent),
       data =  stroop)+
   geom_histogram(color =I('black'),
                  fill = I('#099009'))

q3 <- ggplot(aes(x= Incongruent),
       data =  stroop)+
   geom_histogram(color =I('black'),
                  fill = I('#099009'))

grid.arrange(q3, q2,q1, ncol =1)
```



There are three plots of Incongruence , Congruence and the difference between them. As it can be seen the the mean and median of the differences is around 7.5 , there is a drop between 12.5 to 16.5 and between 17.5 to 18.5.

6. A statistical test has been correctly performed and reported, including test statistic, p-value, and test result. The test results are interpreted in terms of the experimental task performed.

Below I have calculated t statistics using a simple in built function .

References for following t test : http://www.instantr.com/2012/12/29/performing-a-one-sample-t-test-in-r/

```
# Verify using the t.test() function
t.test(x=stroop$Incongruent, y=stroop$Congruent,
       alternative = "greater",
       paired = TRUE,
       conf.level = 0.95)
```

```
##
##  Paired t-test
##
## data:  stroop$Incongruent and stroop$Congruent
## t = 8.0207, df = 23, p-value = 2.052e-08
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.262868      Inf
## sample estimates:
## mean of the differences
##                7.964792
```

From the following information : n = 24 DOF : n-1 = 23 alpha = 0.05 t critical value obtained from t table is 2.05 , hence result is statistically important so we have to reject the null.