Examining the Stroop Effect (Udacity)

Stanislav Sajin

February 29, 2016

General Summary

The aim of this project is to examine the Stroop effect. More details about the task can be found here

Question 1: What is our independent variable? What is our dependent variable?

The independent variable in the Stroop task is Congruency, and it has two levels (Congruent and Incongruent). In the congruent condition, the words' color names matches the color of the font they are printed in (e.g., the word RED in a red color ink). In the incongruent condition, the words' color names mismatches the font color (i.e., the word RED in a green color ink).

The dependent variable is reaction time for participants to name the color of the words in the display. Reaction time is measured in seconds, from the onset of the display with words.

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

In this experiment, we hypothesize that the mean performance time for the incongruent condition (??I) will be higher relative to the mean performance time for the congruent condition(??C). In other words, it will take longer for participants to name the color of the words in the incongruent condition relative to the congruent condition. The reason why reaction times will be higher for the incongruent condition is because the automatic process of reading the color label will lead to interference in correctly naming the correct ink color. Ho: ??I<=??C H1: ??I>??C

??I is the population mean time for the incongruent condition ??C is the population mean time for the congruent condition

Because subjects participate in both conditions, this would be considered a within-subjects design, so a parired t-test would be appropriate. Moreover, we select a t.test because our sample size is less than 30 subjects, and we don't know the population standard deviation. We also assume that the distribution the difference in performance times is Gaussian.

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

Below you will find the descriptive statistics (Minimum, 1st Quartile, Median, Mean, 3rd Quartile, and Maximum) for the Congruent and Incongruent Condition

```
##         Congruent Incongruent
## Min.         8.63       15.69
## 1st Qu.     11.90       18.72
## Median      14.36       21.02
## Mean        14.05       22.02
## 3rd Qu.     16.20       24.05
## Max.        22.33       35.26
```

```
##   Congruent Incongruent
##    3.559358    4.797057
```

The mean and median for the incogngruent condition (Mean=22.02; Median=21.02) is higher than for the congruent condition (Mean=14.05; Median=14.36), indicating that participants found the task in the incongruent condition a bit more difficult to perform.

The incongruent condition also shows higher variability (SD=4.797) when compared to the congruent condition (3.55)

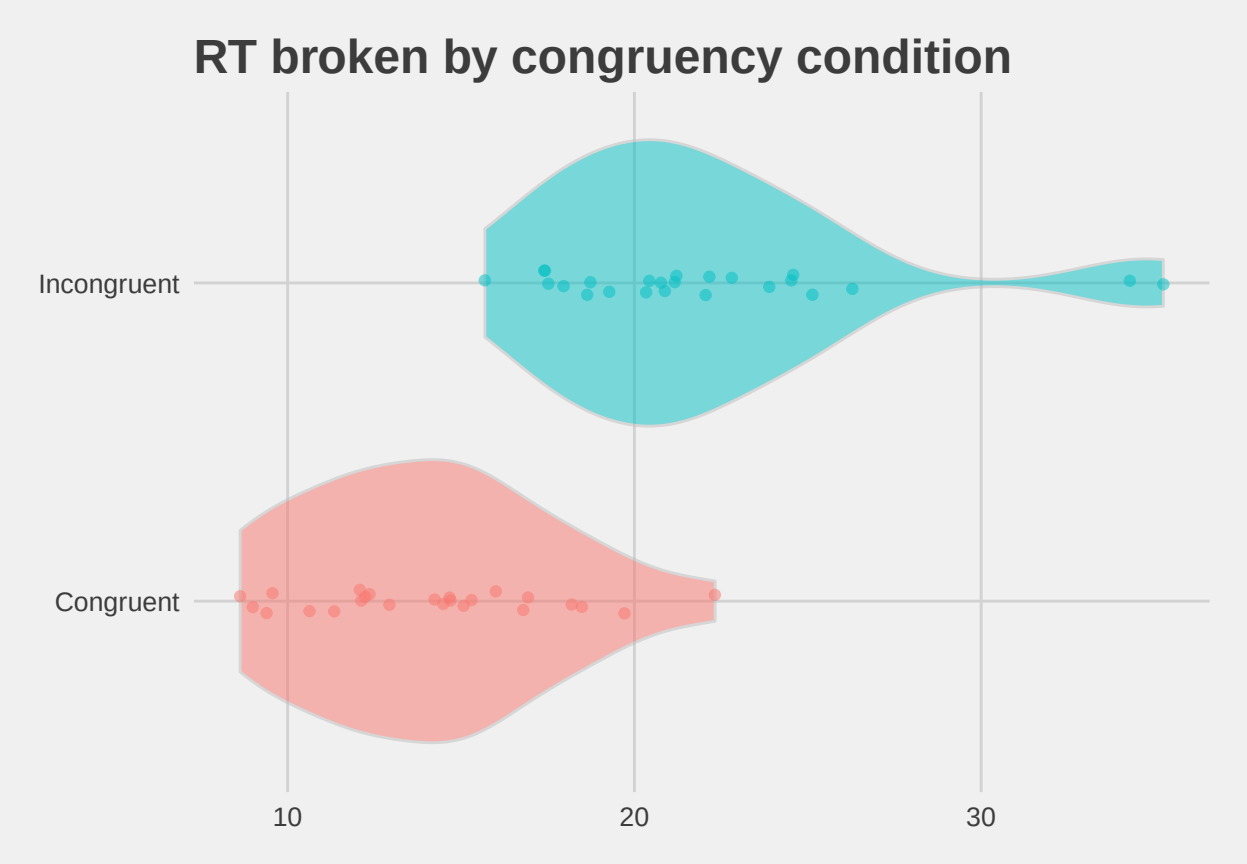
Question 4: Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Before I perform any visualizations, I will transform the data in a tidy data.

```
#load these two packages
library(ggplot2)
library(ggthemes)
library(tidyr)
tidyStroop<-stroop %>% gather("CongruencyCondition","ReactionTime",1:2)
```

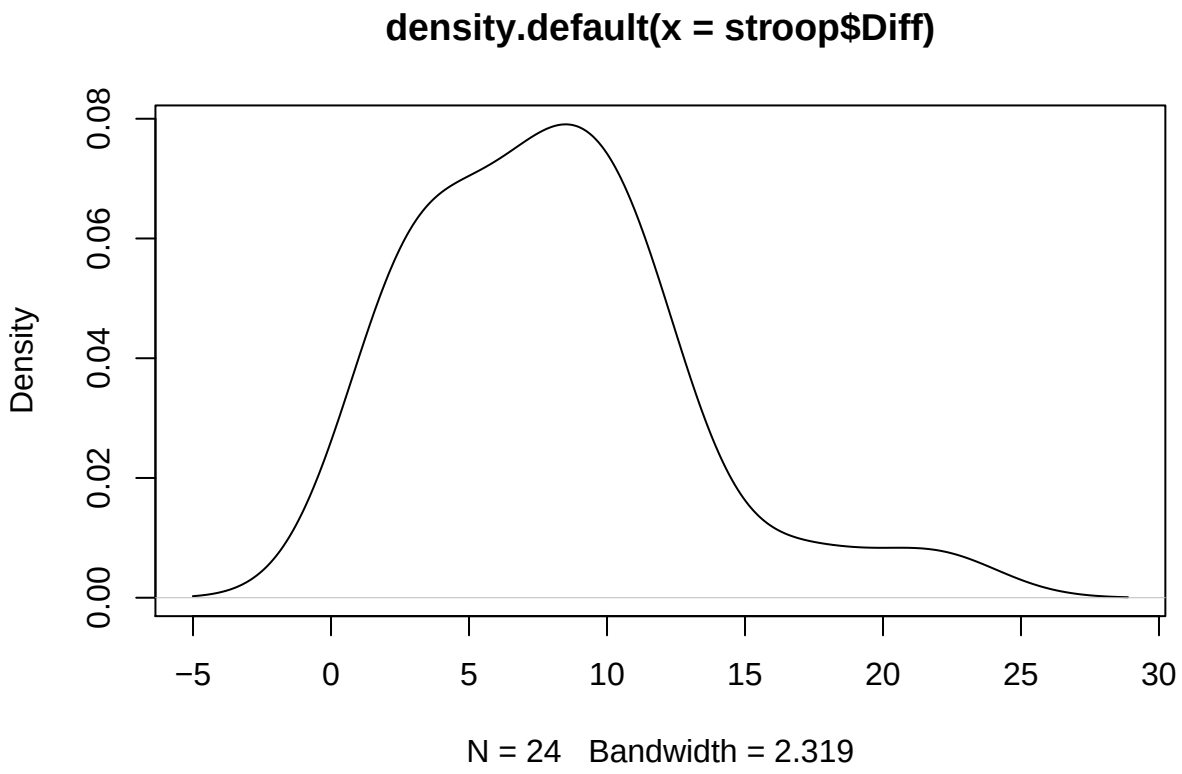
The following plot will show the density distribution of responses for the Congruent and Incongruent Condition.

```
ggplot(tidyStroop, aes(x=CongruencyCondition, y=ReactionTime, fill=CongruencyCondition))+
    geom_violin(alpha=0.5, color="gray")+
    geom_jitter(alpha=0.5, aes(color=CongruencyCondition),
              position = position_jitter(width = 0.1))+coord_flip()+
    labs(title = "RT broken by congruency condition",
           x = " ", y = "Reaction Time (s)")+theme_fivethirtyeight()+theme(legend.position="none")
```



There is one notable about these violin distributions. The plot indicates that two participants provided very extreme responses for the incongruent condition. Since a within-subject test examines the distribution of difference scores, in the next plot I will provide the distribution of the difference in times between the incongruent and congruent condition.

```
stroop$Diff<-stroop$Incongruent-stroop$Congruent
plot(density(stroop$Diff))
```



The plot indicates that there is a positive skew in our distribution. A t-test might not be appropriate if the assumption of normality is not met.

Question 5: Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

Below you'll find the summary of a within subjects t-test (two-tailed). Although the prediction would normally involve a one-tailed t.test, I assume that the 95% CI that the question is asking is for a two-sided CI (which requires a two-tailed t.test)

```
t.test(x=stroop$Congruent,y=stroop$Incongruent, paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  stroop$Congruent and stroop$Incongruent
## t = -8.0207, df = 23, p-value = 4.103e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -10.019028  -5.910555
## sample estimates:
## mean of the differences
##               -7.964792
```

The test shows a very significant difference in performance between the incongruent condition and the congruent condition. Participants were faster at naming the correct color in the congruent condition (t(23)=-8.02, p<.0001; 95% CI [-10.01,-5.91]). The t critical value for alpha of .05 is 2.069 and -2.069. Note that the alpha reprsents the probability for both tails of the t-distribution. Based on this inferential test, we reject the null hypothesis and conclude that the incongruent condition leads to slower response times than the congruent condition. The results did match with our expectations.

A few notes

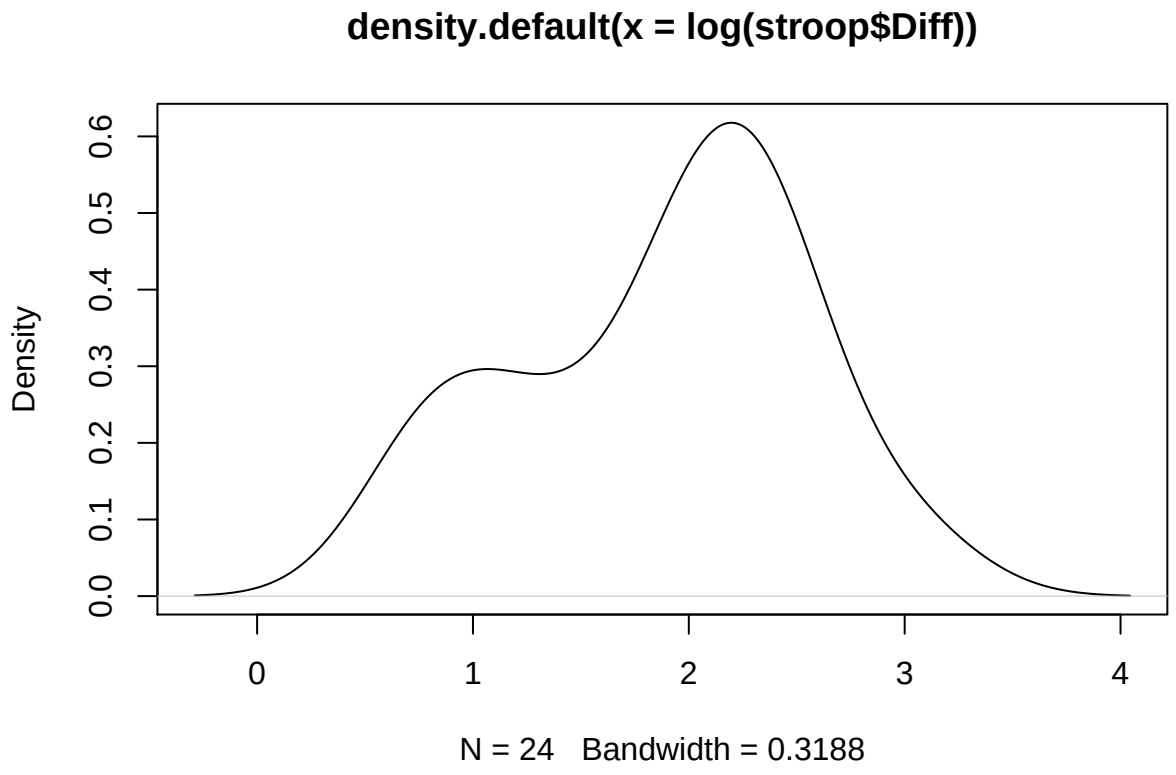
As pointed out in Question 4, the distribution of the difference scores has a positive skew. This suspicion is confirmed through Shapiro-Wilk normality test

```
shapiro.test(stroop$Diff)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  stroop$Diff
## W = 0.91042, p-value = 0.03602
```

Non-normality is a problem, because the two data points at the end of the incongruence condition distribution might be driving our effect. One fix to this would be to work with log-transformed data. The log of the difference is plotted below

```
plot(density(log(stroop$Diff)))
```



It seems like the fix didn't work so well, because now we have a platykurtic distribution, with a negative kurtosis value of -1.0.

It is recommended that the regular t.test be replaced with a permutation t.test, in order to get a better estimate of the inferential values of the Stroop effect.

Question 6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!

The most common theory for the stroop effect is that our brain automatically processes text due to our extensive experience with reading. IN other words, beacuse reading is automatic it leads to interference when we try to name the color of the words. One alternative to this task that should lead to a similar effect would be to ask participants to focus selectively among two sources of information. For instance, a participant might be given a dichotic listening task, where one of the voices is familiar and one is unfamiliar. If participants are asked to attend to the unfamiliar voice, will the familiar voice interfere show any interference? What will happen if participants are asked to attend to the familiar voice, will the unfamiliar voice show interference? Automaticity theory relies on the participant having extensive familiarity with one source of information and would suggest that a strongly familiar voice will lead to greater interference from the familiar voice than from the unfamiliar voice.