Statistics: The Science of Decisions Project Instructions

Scott Tse

6/13/16

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

Our **independent variable** is test words condition which has one of two values, congruent or incongruent. Per the background information, in the congruent condition, the words displayed match the colors in which they are printed and in the incongruent words condition, they do not match.

Our **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.

Since the measurements were performed on the same participant in each observation within the sample, we can consider the data paired since we cannot be sure that the congruent vs incongruent measurements for each student are independent. In looking at paired data, we take the difference between the congruent and incongruent measurement to calculate a point estimate from the sample. For support on statistics, I used the OpenIntro Statistics Text, located at https://www.openintro.org/stat/textbook.php?stat_book=os

We are trying to find out if the Stroop Effect is actually a valid phenomenon affecting our population. If the Stroop Effect is occurring then we should expect our sample to show that there is a difference the mean time it takes to name the colors from the congruent condition vs the incongruent condition. I used a two sided t test of paired observations to evaluate the null hypothesis. I assumed that the difference in pairs were from random and independent samples (assumed that the sample size was less than 10% of the population).

Thus we can set up our hypotheses this way:

> The null hypothesis would imply that there is no Stroop Effect. That is, on average, there is no difference in the time it takes to name the colors from the congruent condition to the incongruent condition in the population. We use the convention INCONGRUENT – CONGRUENT to calculate the difference in each pair and then the mean difference in the population we then call will call u_{diff} . Stated mathematically, we can write the null hypothesis as (H_0: $u_{diff} = 0$).
>
> Conversely, we can then state the null hypothesis in a way that implies that we can observe the Stroop effect in the population. The mean time it takes to name the colors from the congruent condition to the incongruent condition in the population is not zero. (H_A: $u_{diff} \neq 0$).

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

Table 1. Summary statistics for Stroop dataset (time values in seconds)

Mean $\overline{x}_{diff}$	Median	Max	Min	Std Dev S_{diff}	N
7.965	7.667	21.919	1.950	4.865	24

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.

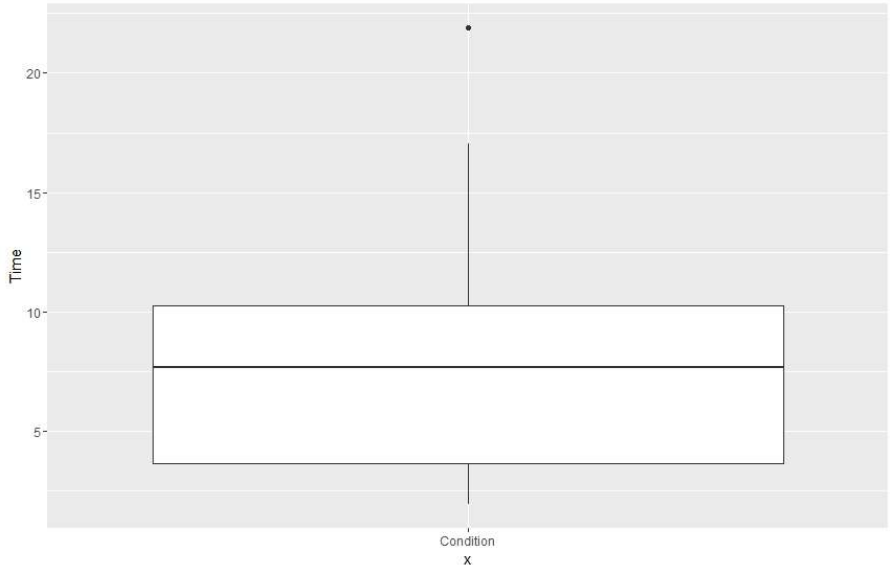

Figure 1. Boxplot of values calculated from X_{diff} ($u_{incongruent}$ - $u_{congruent}$)

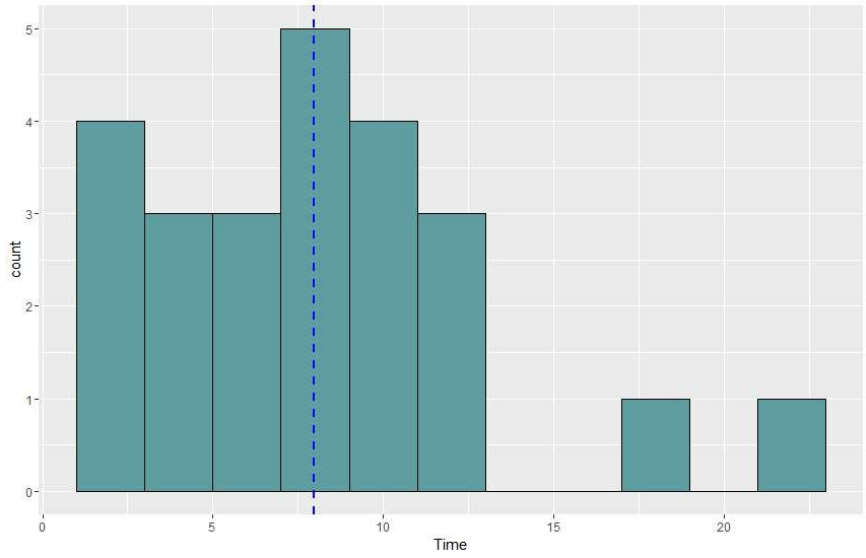

Figure 2. Histogram of values calculated from X_{diff} ($u_{incongruent}$ - $u_{congruent}$)

I used R Studio to generate these visualizations. The boxplot and histogram indicate moderate right skew to the distribution. In addition, there appears to be one outlier in the distribution as well. Although there is an outlier in that data, and moderate skew, our sample size is almost n = 24 so we will assume the normal model. After we generate the P-Value from the t-test we can look at the results to re-evaluate how to handle the outlier.

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?

I used the t test for paired observations to evaluate the null hypothesis. For this hypothesis test, I chose a typical significance value of .05, corresponding to a 95% confidence interval. I used the t test in this case since:

1) the population parameter standard deviation is not known
2) the sample size is small n < 30
3) the sample distribution is slightly skewed and with a single outlier but since n >20 I assume the normal model (sampling distribution is nearly normal). After we generate the P-Value from the t-test we can look at the results to re-evaluate how to handle the outlier if necessary.

For the paired data, as stated above, we have the following null and alternative hypotheses, respectively:

$H_0 : \mu_{diff} = 0$

H_A: $\mu_{diff} \neq 0$

We can evaluate our t statistic as follows based on the point estimates:

$T = \overline{x}_{diff} - 0 / SE_{diff}$

$SE_{diff} = sdiff / \sqrt{n} = 4.865/\sqrt{24} = .993$

$T = (7.965 - 0) / .993$

$T = 8.0211$

The degrees of freedom, df, are 23, so **P value = 0.00000004** which is much less than the significance value **α = .05**

Thus, we can safely reject the null hypothesis, even with an outlier which would make the standard error lower (and hence lower the P value as well) if it were removed from the data set. We indeed have convincing evidence that H_A: $\mu_{diff} \neq 0$, that the difference between incongrurent and congruent condition times is not zero on average in the population. This result did match up with my expectations since the Stroop effect is apparently well known and it is not surprise that is was observed in this sample data as well.

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!

From the Wikipedia entry for Stroop Effect (https://en.wikipedia.org/wiki/Stroop_effect), there appear to be four main theories to explain the cause: Processing Speed, Selective Attention, Automacity, and Parallel Distributed Processing. All fall under what are referred to as "race models" which use the concept that the brain processes different types of information, both relevant and irrelevant, in parallel , but are funelled and "race" into a more central processing function for reporting and decision making. I think these thoeries give plausable explanations for the Stroop effect in general. For instance, in the Processing Speed theory, the Stroop Effect results from the fact that processing words is faster than colors so there is a delay while the brain recognizes color. This does seem to make intuitive sense at first glance.

Since Stroop Effect is defined in general as a demonstration of interference in the reaction time of a task, one could come up with similar tasks. Some examples could be font style on reading speed, or perhaps time to identify which way a person is looking in a photograph depending on where they are located in a photo with regards to a central reference point.

<u>List of Resources Used for Report</u>

- OpenIntro Statistics Text (https://www.openintro.org/stat/textbook.php?stat_book=os)
- Wikipedia entry for Stroop Effect (https://en.wikipedia.org/wiki/Stroop_effect

- OpenIntro Statistics Labs for R: https://www.datacamp.com/courses/statistical-inference-and-data-analysis
- https://www.khanacademy.org/math/probability/statistical-studies/statistical-questions/v/statistics-sample-vs-population-mean
- http://blog.minitab.com/blog/adventures-in-statistics/understanding-hypothesis-tests:-significance-levels-alpha-and-p-values-in-statistics
- P value calculator: http://www.danielsoper.com/statcalc/calculator.aspx?id=8