Statistics: The Science of Decisions Project Instructions

Questions For Investigation

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

The independent variable is whether or not the word condition is congruent.
The dependent variable is the time it takes to name the ink colors.

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

The null hypothesis for this task is that the time taken to name the colors is not significantly different for congruent and incongruent word conditions. The alternative hypothesis is that the time taken is significantly different for congruent and incongruent word conditions.
$H_0: \mu_c = \mu_{ic}$
$H_A: \mu_c \neq \mu_{ic}$
Where μ_c is the mean for the time taken with the congruent conditions and μ_{ic} is the mean for the time taken with the incongruent conditions.

The 2-tailed paired t-test could be used here since it can be used to determine whether the difference between the two dependent populations is significant or not. This test assumes an approximately normal distribution.

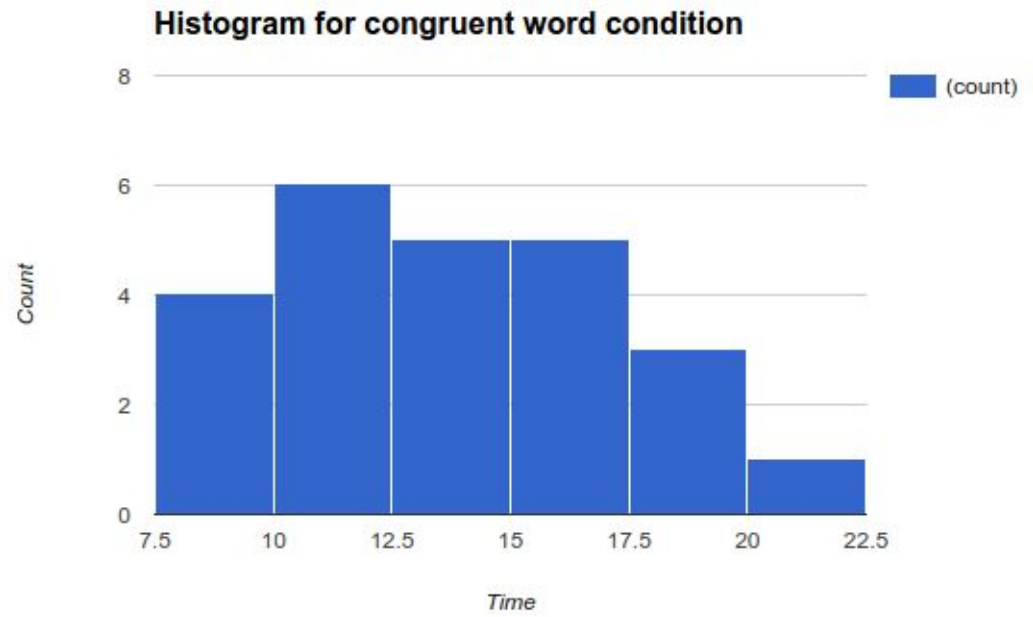
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

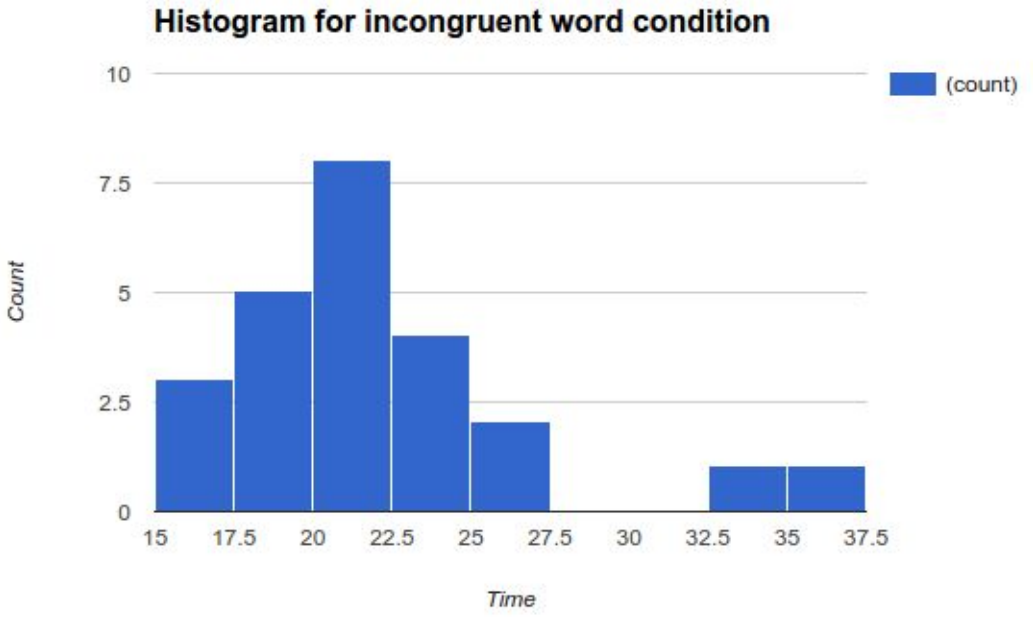
The mean time taken for congruent condition: $\mu_c = 14.051$
The mean time taken for incongruent condition: $\mu_{ic} = 22.016$
Mean of difference between the time taken with congruent and incongruent conditions: $\mu_{c-ic} = -7.965$

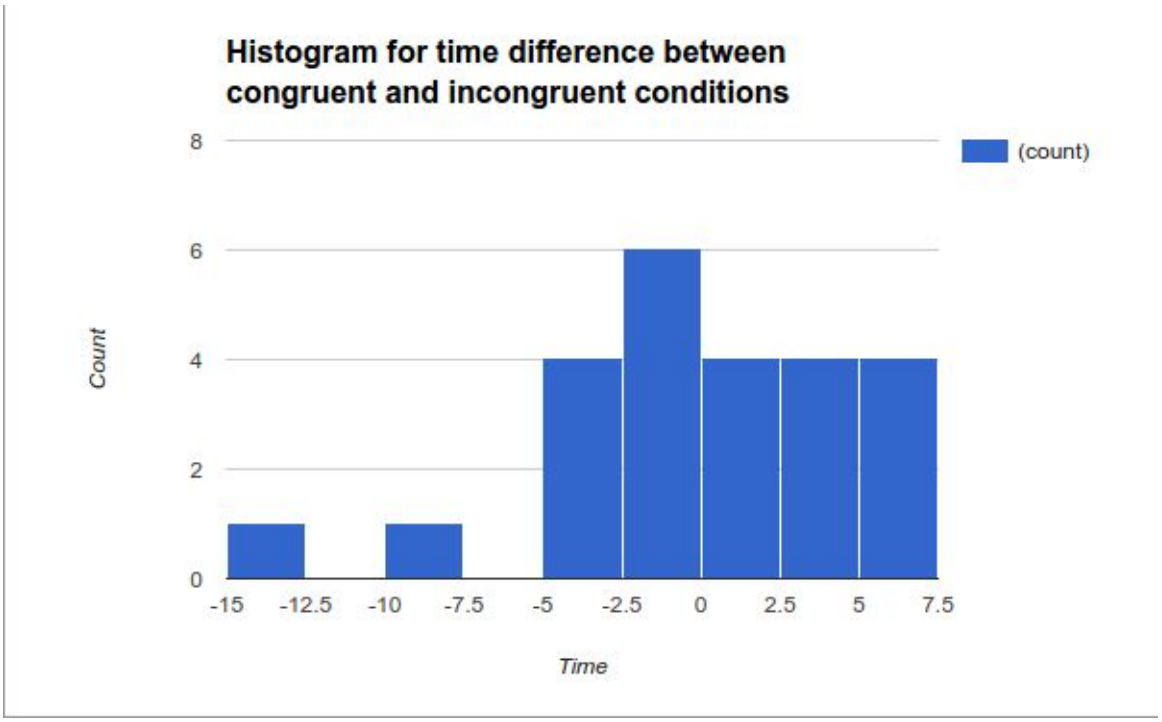
Variance for the time taken for congruent condition: $\sigma^2_c = 12.669$
Variance for the time taken for incongruent condition: $\sigma^2_{ic} = 23.011$
Variance for the difference between the time taken for congruent and incongruent conditions: $\sigma^2_{c-ic} = 23.666$

Standard deviation for the time taken for congruent condition: $\sigma_c = 3.559$
Standard deviation for the time taken for incongruent condition: $\sigma_{ic} = 4.797$
Standard deviation for the difference between the time taken for congruent and incongruent conditions: $\sigma_{c-ic} = 4.865$
Bessel's correction has been applied for variance and standard deviation.

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.







The histograms given above have a bucket size of 2.5 seconds. Though we can see that the distributions are not perfectly normal distributions we can still use the t-test to analyze the data since it is relatively robust to the assumption of normality[ref. #5].

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?

We can take the significance level $\alpha = 0.05$
This will give us a confidence interval of 95%.
The critical t-statistic is: $t_c = \pm 2.069$
From the given data the confidence interval on the mean of difference between the time taken for the congruent and incongruent conditions is:
95%CI = (-10.019 , -5.911)
The degree of freedom for the data-set is 23. The t-statistic and p-value for the set are:
(given by: $t = (\mu_{c-}\mu_{ic})-0 / (\sigma_{c-ic}/\sqrt{n})$)

t(23) = -8.021, p = 0.00000004, two-tailed

This result is statistically significant and thus we reject the null hypothesis. The time taken to perform the task when the colors of the words match the color names is significantly less than the time taken to perform it when the colors of the words don't match the color names. These results did match my expectation.

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!

There are several proposed explanations for the observed effect but the general idea is that the brain finds it easier to read than to identify a color and thus there is confusion in the decision making process due to conflicting information arriving at different times. Another similar task could be to identify the direction in which an arrow is moving while its pointing either in the direction of motion (congruent) or a direction different from the direction of motion (incongruent).

References:
1.https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php
2.https://www.stat.tamu.edu/~lzhou/stat302/T-Table.pdf
3.https://en.wikipedia.org/wiki/Stroop_effect
4.http://www.danielsoper.com/statcalc/calculator.aspx?id=8
5.http://blog.minitab.com/blog/understanding-statistics-and-its-application/what-should-i-do-if-my-data-is-not-normal-v2
All graphs were made in Google sheets