UDACITY

‹ Return to Classroom

Test a Perceptual Phenomenon

REVIEW

HISTORY

Meets Specifications

Next, Large Data Analysis and Our First Prediction Algorithm

Congratulations on completing the course. Next we will be getting deeper into the Python Programming language and doing another sort of statistical test. Remember the level of clarity you have given in this project for that one. The reviews will be held to a similar level of scrutiny. But there is an exciting addition to the statistical analysis. We will be getting into our first Machine Learning and Prediction algorithm! These techniques are at the core of self-driving cars and facial recognition. It is really cool and I hope you like it too. Good luck!

Responses to Project Questions

Q1: Question response correctly identifies the independent and dependent variables in the experiment.

You have correctly identified the Independent and Dependent Variables. Well done!

Q2a: Null and alternative hypotheses are clearly stated in words and mathematically. Symbols in the mathematical statement are defined.

Optimal Way to State the Hypotheses

Null:
The null hypothesis is that the congruent and incongruent samples come from the same general population – meaning even though we are observing a difference in the sample means there is actually no difference in response times between the conditions of the experiment for the population and what we are witnessing is by chance. *One-tail* will go even further and state that the Incongruent population has a lower mean time than the Congruent population.
Two-Tail

$$H_0 : \mu_{congruent} - \mu_{incongruent} = 0$$

One Tail (Preferred)

$$H_0 : \mu_{congruent} - \mu_{incongruent} \geq 0$$

Or
Two-Tail

$$H_0 : \mu_{congruent} = \mu_{incongruent}$$

One-Tail (Preferred)

$$H_1 : \mu_{congruent} \geq \mu_{incongruent}$$

Alternative:
One-Tailed (Preferred)
The alternative hypothesis is that the population_mean(incongruent) is greater than the population_mean(congruent) – meaning that they come from different populations and that the incongruent condition actually does increase response times - the difference in sample means we are witnessing is representative of the general population.

$$H_1 : \mu_{congruent} - \mu_{incongruent} < 0$$

Or

$$H_1 : \mu_{congruent} < \mu_{incongruent}$$

Two-Tailed (similar to the one-tailed but not assuming which is larger (or smaller))
The alternative hypothesis is that there is a difference between the population_means of the congruent and incongruent, however, we are not assuming which is larger or smaller - the difference in sample means we are witnessing is representative of the general population in that they are different, but we are not making the assumption of directionality.

$$H_1 : \mu_{congruent} - \mu_{incongruent} \neq 0$$

Or

$$H_1 : \mu_{congruent} \neq \mu_{incongruent}$$

- **Note** A one-tail assumption is preferred and justified since the question is “does Incongruency increase response times” and all Incongruent times are larger in our sample dataset.
- Notice how the statements of the hypotheses are making sure to convey they reference an assumption about the population **not** the samples since we do not need to make a guess about the samples - we have access to them

have access to them.

Q2b: A statistical test is proposed which will distinguish the proposed hypotheses. Any assumptions made by the statistical test are addressed.

Optimal Answers

Statistical Test
Since **(1)** we do not know the population standard deviation, **(2)** our sample size < 30, **(3)** distributions are normal, and **(4)** the samples from one trial is used in the second trial (same participant is used to test the effect of the conditions - repeated measure), we will use the **Dependent T-test** for paired samples.

- Notice how the justification of the Dependent T-test includes not only the T-test is applicable (#1, #2, #3) but also why the specific T-test (Dependent) is specifically applicable (#4).

Wording the statements like this accurately conveys to the reader you have a firm understanding of the concepts without any uncertainties. Overall, I believe you do know these concepts and I also know it can be hard to articulate them on paper. Being able to explain these concepts to someone else is really good practice and can sometimes mean the difference between memorizing the material and understanding the material.

Q3: Descriptive statistics, including at least one measure of centrality and one measure of variability, have been computed for the dataset's groups.

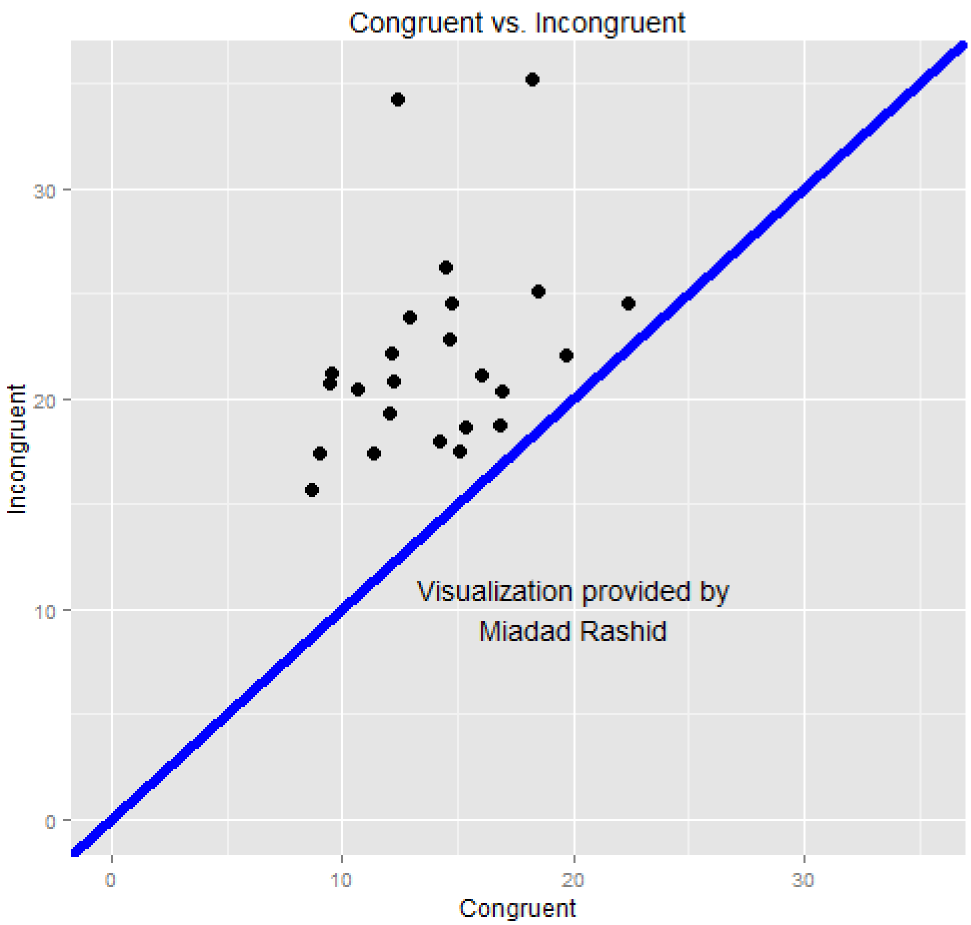
✔ Measurements of centrality are correct
✔ Measurements of Variability are correct

Q4: One or two visualizations have been created that show off the data, including comments on what can be observed in the plot or plots.

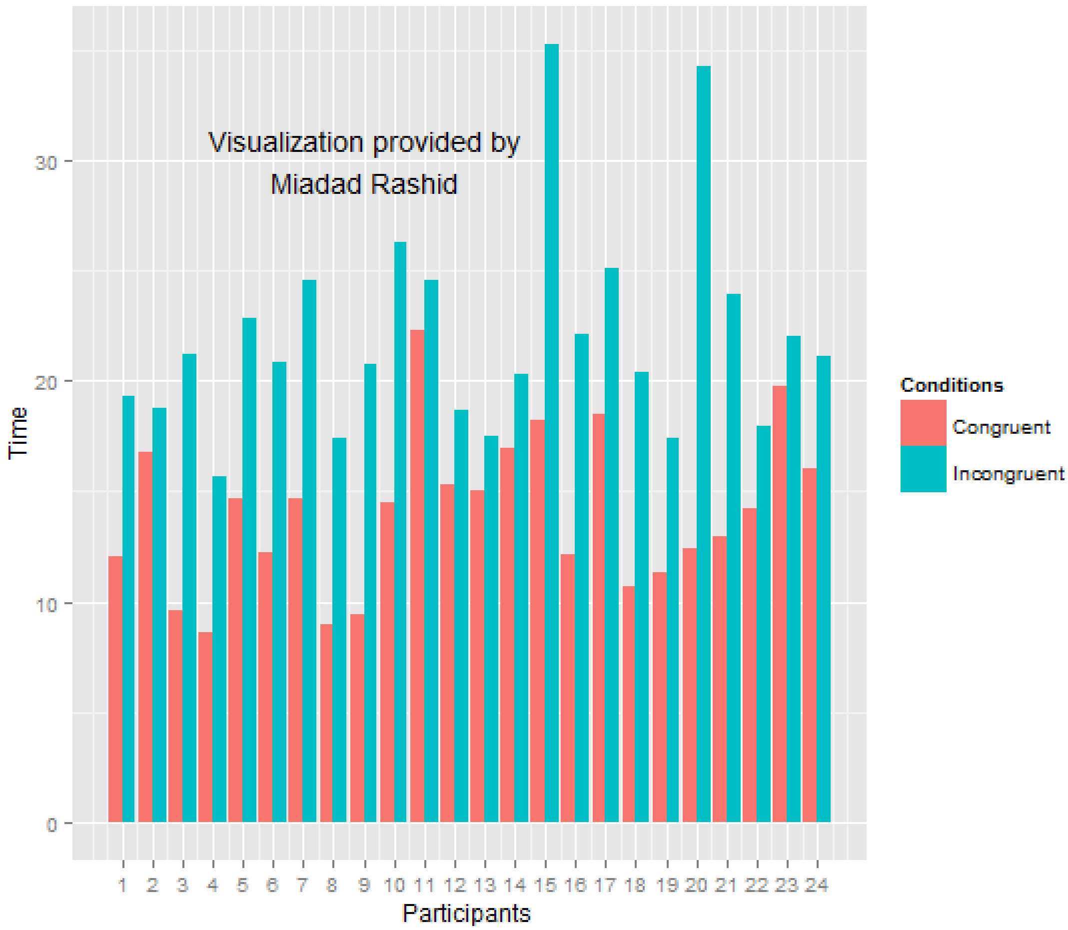
Great work on the visualizations. I have also found the following helpful in my investigation.

Visualizations I found Useful

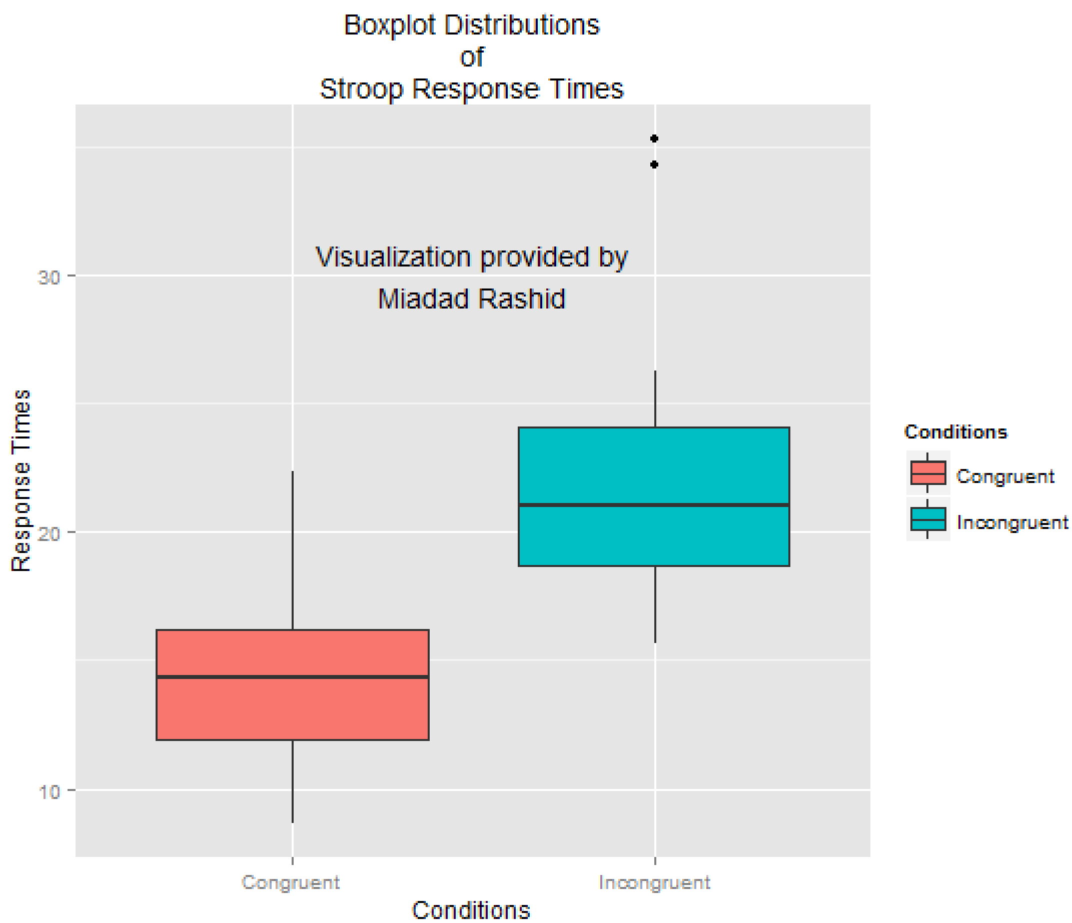
Scatter Plot Comparison If you were to add a line where x = y to a graph of Congruent vs Incongruent, this line would represent if Incongruent was equal to Congruent. You will notice that **all points are above the line.**



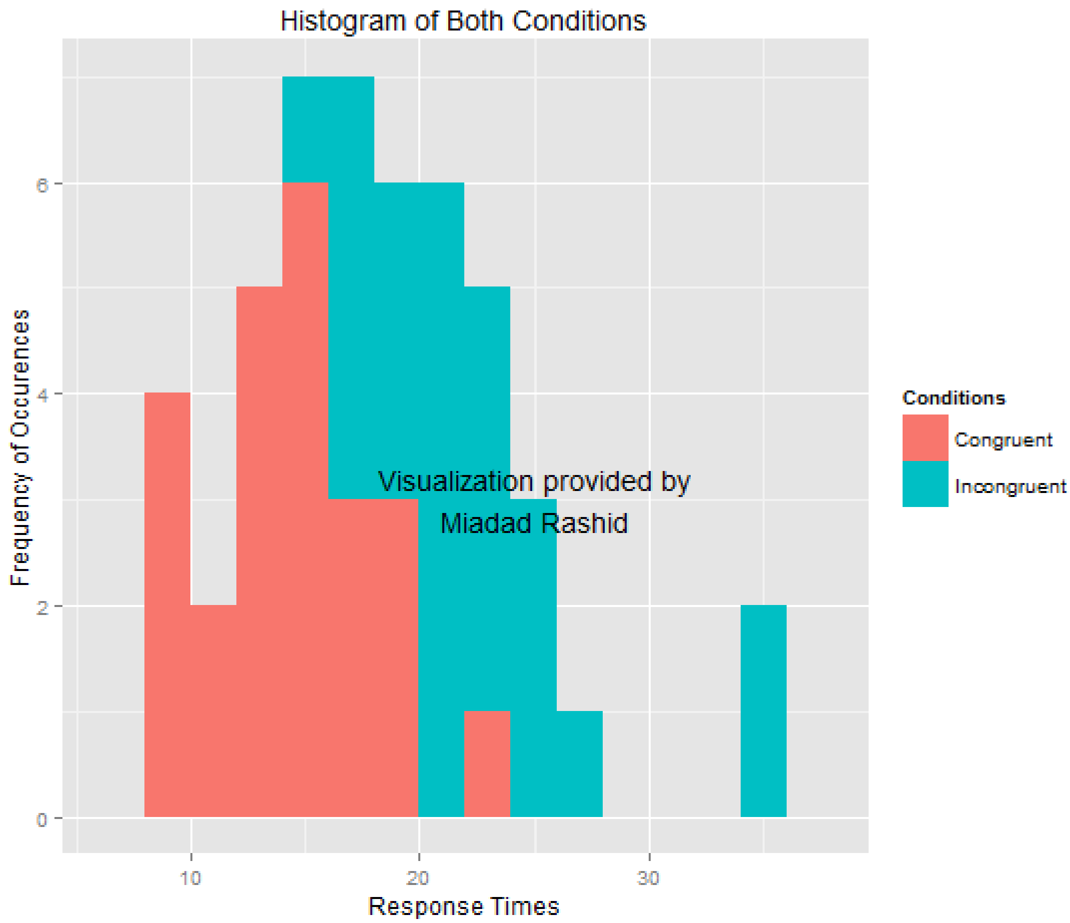
Bar-plot of both Trials Another great tool is a barplot of the 2 conditions side by side. If a bar plot of Congruent and Incongruent samples were overlaid with a bin size = 1 and each bar next to each corresponding bin, it would represent each individual's response time differences. This can offer great insight simultaneously showing times of each individual, and their delays. Here is an example.



Box-Plot Box plots are a great tool for discerning distributions. The upper and lower limits of the boxes show the 75th and 25th quantiles, respectively. The line in the box is the median. Any point outside the region represents some clear outliers.



Histogram of both Trials Histograms are a great way to look at if the distributions follow a probability curve since a lot of Statistical tests have assumptions of probability distributions.

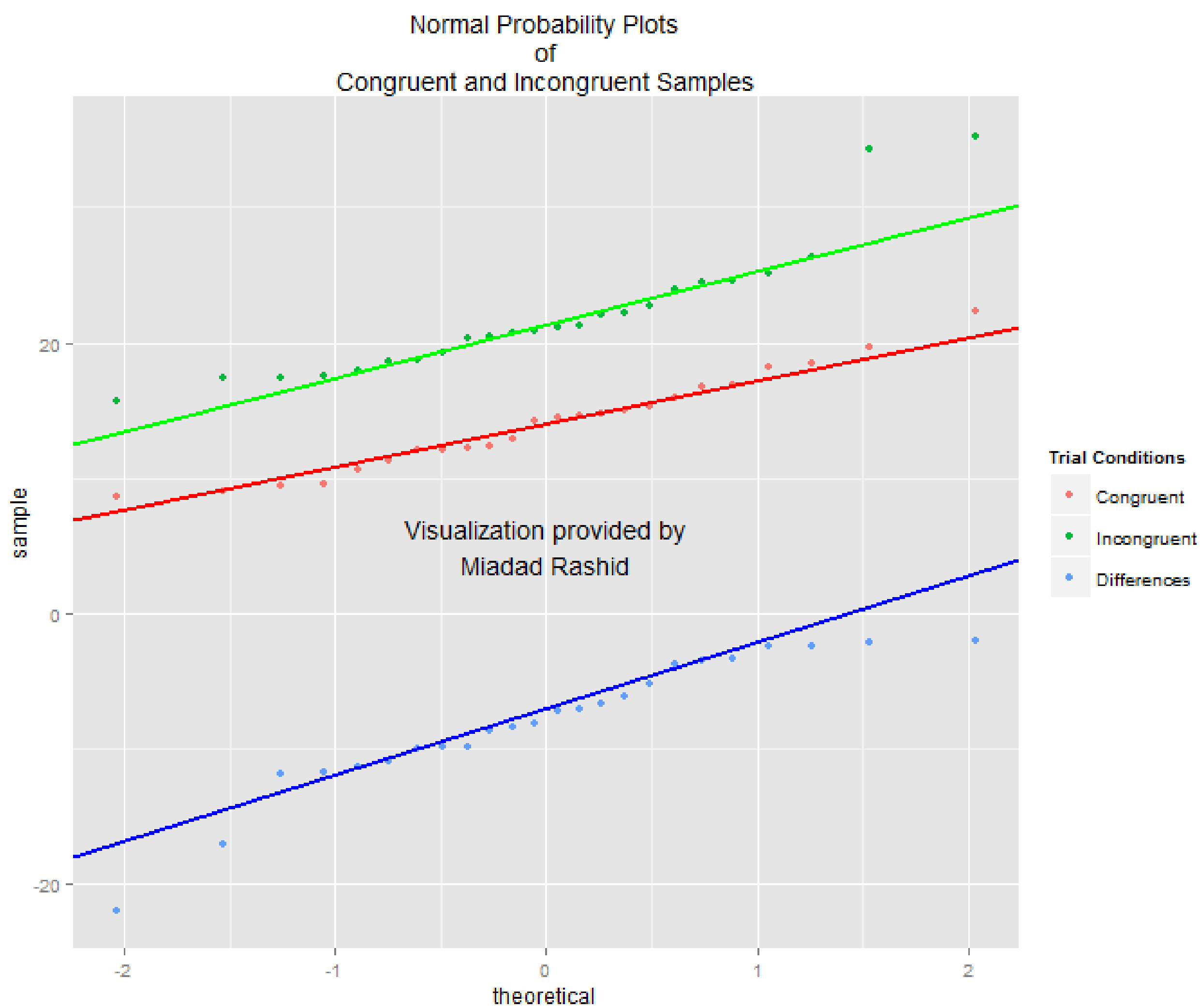


Q5: A statistical test has been correctly performed and reported, including test statistic, p-value, and test result. The test results are interpreted in terms of the experimental task performed. Alternatively, students may use a bootstrapping approach to simulate the results of a traditional hypothesis test.

To Exceed Specifications

For this project, it is ok to *assume* that our distributions are normal but what if they were not? One of the main things before we use a statistical test that assumes normal distribution is to test for normality. There are a couple of ways to test normality of sample distributions besides looking at the histograms. I wanted to talk about some of them as they are a more accurate way to judge the normality of distributions and will be a valuable tool later on in your career. For the level of this class, we let students assume the distributions are normal and therefore can use a T-test which assumes normality. Upon further investigation, we find this is not quite true.

Normal Probability Plot



- The line you see for each conditions represents the line they would follow if they were normally distributed.
- As you can see, the Congruent samples seem to be barely on the line and the Incongruent Samples have some definite divergence from the line. As for the differences between the two, the assessment of normality seems inconclusive, but I would suspect that it is not..

We must be sure so we apply a quantifiable test to determine normality.

Shapiro-Wilk Normality Test

1. A Shapiro-Wilk Test is a statistical test that returns a Wilk Score and p-value. The Null is that the Distribution is Normal. So let's set up an α = 0.05.
   - **Congruent Samples**

```
        Shapiro-Wilk normality test

data:  df$Congruent
```

```
data:  df$Congruent
W = 0.97092, p-value = 0.6898
```

- Fails to reject the Null - it is normally distributed. What do you think the Incongruent Samples will do?

- **Incongruent Samples**

```
        Shapiro-Wilk normality test

data:  df$Incongruent
W = 0.85395, p-value = 0.00259
```

- It **Rejects the Null** - the samples are not normally distributed, as we suspected from the Probability Plots.
- **Differences**

```
        Shapiro-Wilk normality test

data:  diff
W = 0.91042, p-value = 0.03602
```

- It **Rejects the Null** - the samples are not normally distributed.

Although it is out of the scope of the class, we have grounds here to not use a T-test since a T-test assumes normality of the samples in question. We would require a Statistical test that is more robust and resistant to the outliers we are seeing in the plot. This would be a non-parametric test. We will go into this much deeper in the next class so I do not want to ruin it. If you would like to get a head start, start researching non-parametric tests. Which would you use for this dataset? *HINT* remember the samples are paired.

Q6: Hypotheses regarding the reasons for the effect observed are presented. An extension or related experiment to the performed Stroop task is provided, that may produce similar effects.

Thank you for extending and thinking deeper with the investigation. Do you think that the effects of incongruency can be observed in other senses like hearing and touch?

Chemosensory Experiment

http://chemse.oxfordjournals.org/content/32/4/337.full

DOWNLOAD PROJECT

RETURN TO PATH

Rate this review