Statistics: Test a Perceptual Phenomenon

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

The independent variable is the congruency condition - whether the name of the color matches with the ink color, time taken is the dependent variable.

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

One hypothesis we can use is: there is a difference between the time used to recognize colors under congruent words condition and incongruent words condition, namely, the Stroop Effect is in existence.

Specifically, here we referring to the population means of congruence words group and incongruence words group - average times for the respective groups to recognize the colors. By comparing these means directly, we'll be able to tell whether there is a difference between the the two groups' color recognition times. However, it wouldn't be possible to do the experiment with all potential subjects in the world, so we need to work with the sample we have on hand to make inference about the population means, i.e., to use the observation means, sd and other statistics to infer about the population means. In this case, the observation is the difference between the two groups' times. With this new data, we can construct new statistics such as means and standard errors.

To achieve this, we can use a two-sided paired student T-test to verify. This is because: one, we need to address the uncertainty in sample standard error resulted from the unknown population standard deviation; two, we are comparing the means of two groups that are dependent; three, the same subject is involved under both conditions.

Below is the hypothesis to test:

H0: mu_diff = 0 (The real difference between group population means is zero)

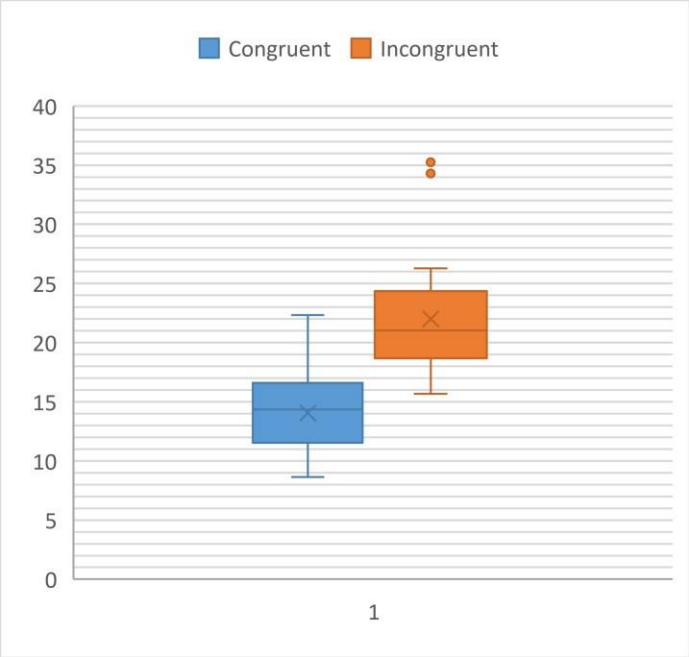
HA: mu_diff != 0 (The real difference between group population means is not zero)

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

Statistic	**Congruent**	**Incongruent**	**Difference (C-I)**
n	24	24	24
$\bar{x}$	14.05	22.02	-7.96
Median	14.36	21.02	-7.67
s^2	12.67	23.01	23.67

s	3.56	4.80	4.86
SE	0.73	0.98	0.99

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.**



Below I have plotted the distributions for Congruent and Incongruent data in the form of box plots. It's very interesting to see the mean of incongruent series be significantly higher than that of the congruent series. I am expecting the results of the t-test to be significant.

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?**

Basically they are trying to prove that when incongruent happen people would spend more time so that if we look to differences between congruent and incongruent date we will have one table:

Congruent	Incongruent	differences
8.63	15.687	7.057

8.987	17.394	8.407
9.401	20.762	11.361
9.564	21.214	11.65
10.639	20.429	9.79
11.344	17.425	6.081
12.079	19.278	7.199
12.13	22.158	10.028
12.238	20.878	8.64
12.369	34.288	21.919
12.944	23.894	10.95
14.233	17.96	3.727
14.48	26.282	11.802
14.669	22.803	8.134
14.692	24.572	9.88
15.073	17.51	2.437
15.298	18.644	3.346
16.004	21.157	5.153
16.791	18.741	1.95
16.929	20.33	3.401
18.2	35.255	17.055
18.495	25.139	6.644
19.71	22.058	2.348
22.328	24.524	2.196

α = .01

df = 23

t_{crit} = -2.50 t = -

8.02 p-value = <

.0001

At the 99% confidence level ($\alpha = .01$) and 23 degrees of freedom, the critical statistic value for a one-tailed test in the negative direction is -2.5. The calculated t-statistic for the difference in colour recognition time means of the congruent and incongruent word data is 8.02. Since the t-statistic is in the critical region, the null hypothesis is rejected. With the data presented, it is very unlikely that the 7.96 second difference in mean time for colour recognition for the congruent data vs. the incongruent data is obtained if the two means are actually the same (or if $\mu_C > \mu_I$). By conventional criteria, this difference is considered to be extremely statistically significant.

There is sufficient evidence at the $\alpha = .01$ level of significance to support the claim that it takes less time to recognize the colour of words with the congruent condition compared to words with the incongruent condition.