

# Hypothesis Testing

Here I am interested to find the relation between the battery in the car and the mileage. Now, by replacing the battery I want to know the effect on car's mileage.

## Testable Hypothesis:

If I replace the battery in my car, **then** I will get better mileage.

## Data Collection:

The data needed here is the car mileage before and after battery replacement. The mileage is calculated by dividing the number of kilometres driven as per the trip meter by the quantity of fuel used by the car. These mileage readings are taken over a period of time before and after battery replacement. And a random sample1 is collected from the mileage data before battery replacement and random sample2 is collected after replacement.

Then, mean value of mileages are calculated for both random samples. Let them be  $\mu_1$  and  $\mu_2$ . Now, the difference between the sample means  $\mu_1$  and  $\mu_2$  is calculated, and we need to determine whether the observed difference between them is statistically significant.

## Classical Hypothesis test:

Is the observed mean mileage after battery replacement is significantly more than before replacement mean?

From here we will be generating,

Null Hypothesis **H0:  $\mu_1 - \mu_2 = 0$**

Alternative Hypothesis **H1:  $\mu_1 - \mu_2 > 0$**

Let the difference(D) between the means is M. Here I am considering my number of sample observations is greater than 30 ( $n > 30$ ), hence I am performing a **Z-test**.

$$P(D \geq M) = P(Z \geq \frac{M - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1 + \sigma_2^2/n_2)}})$$

Here  $\sigma_1^2/n_1, \sigma_2^2/n_2$  = Variation of sample means before and after battery replacement  
 $n_1, n_2$  = Number of sample observations before and after battery replacement

The p-value indicates how likely the data would have occurred by random chance. The level of statistical significance is often expressed as a p-value. The smaller the p-value, the stronger the evidence that you should reject the null hypothesis.

From above equation, p-value (probability value) is calculated, if

p-value  $\leq 0.05$ , H0 is rejected and H1 is accepted

p-value  $> 0.05$ , H0 is accepted and H1 is rejected.

**In order to support my testable hypothesis, the mileage after the battery replacement should be significantly greater than before replacement. i.e.  $p \leq 0.05$ , H1:  $\mu_1 - \mu_2 > 0$  should be accepted.**