

Mini Project 2

Par Inc



January 18, 2019

Saurabh MUDGal

**Table of Contents**

1 Business Objective .............................................................................. 2

2 Proposed Solution ............................................................................... 4

2.1 Hypothesis ........................................................................ 4

2.2 P Valaue ........................................................................... 5

2.3 Descriptive analysis........................................................................ 6

2.4 Question d: ..............................................................

# Business Objective

To increase the market share for Par Inc. , the proposal is to introduce the cut resistant and longer lasting golf ball.

The research team of Par Inc. has come up with the new coating design which resist cut and provide more durable ball .

The test results so far are good but there is concern about effect of new coating on driving distances as it should be comparable with the current model of golf ball. To compare driving distance, 40 balls of both the models are subjected to distance test .

The result of test are contained in attached Golf.csv .

Refer Question and Answer section to deduce the new golf ball distance test analysis :



## Questions

1. Formulate and present the rationale for a hypothesis test that par could use to compare the driving distances of the current and new golf balls

2. Analyze the data to provide the hypothesis testing conclusion. What is the p-value for your test? What is your recommendation for Par Inc.?

3. Provide descriptive statistical summaries of the data for each model

4. What is the 95% confidence interval for the population mean of each model, and what is the 95% confidence interval for the difference between the means of the two population?

5. Do you see a need for larger sample sizes and more testing with the golf balls? Discuss

# Proposed Solution

2.1 To compare the driving distance of both the models NULL hypothesis will be used .Here are the details :

* Sample Size ,N = 40, number of samples =2
* Since there are two set of samples driving distances we have one for older balls and other for new coated one and standard deviation for population is unknows hence we can use **Two Sample T test** .
* Let’s assume that level of significance (alpha) = .05
* Degree of freedom is N-1: 39.(as N=40)

Hypothesis formulation :

Using T Test for Two Sample .

Ho: Mu1 = Mu2

Ha : Mu1 not equal to Mu2

Where Mu1 is the mean for current ball driving distance

And Mu2 is the mean for New ball driving distance .

## 2.2 Pvalue can be calculated with the help of t.test () in R.

t.test(ï..Current,New)

where i..Current is current ball driving distance vector

New is New ball driving distance vector ,

Welch Two Sample t-test

data: ï..Current and New

t = 1.3284, df = 76.852, p-value = 0.188

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.384937 6.934937

sample estimates:

mean of x mean of y

270.275 267.500

alpha is .05 and pvalue is .188 , since pvalue > alpha, hence NULL hypothesis is not rejected .

SO we can say that new ball driving distance is almost similar to current ball driving distance .

Hence with improved cut we have similar driving distance .

We can’t say this while simply calculating the pvalue (pvalue > 5% - u can’t accept the null hypothesis ). we need to have some more test to accept the NULL hypothesis . Here come Power of test .

Do 1 sample test – it will tell where mean is lying with 95% confidence

t.test(current )

t.test(new)

then we can say that there is slight difference in performance of new and old ball

## 2.3. Descriptive statistical summaries of the data for each model

Golfdata is data.frame which hold both the sample

Five point summary for both the samples

summary(golfdata)

ï..Current New

Min. :255.0 Min. :250.0

1st Qu.:263.0 1st Qu.:262.0

Median :270.0 Median :265.0

Mean :270.3 Mean :267.5

3rd Qu.:275.2 3rd Qu.:274.5

Max. :289.0 Max. :289.0

> sd(ï..Current)

[1] 8.752985

> sd(New)

[1] 9.896904

> var(ï..Current)

[1] 76.61474

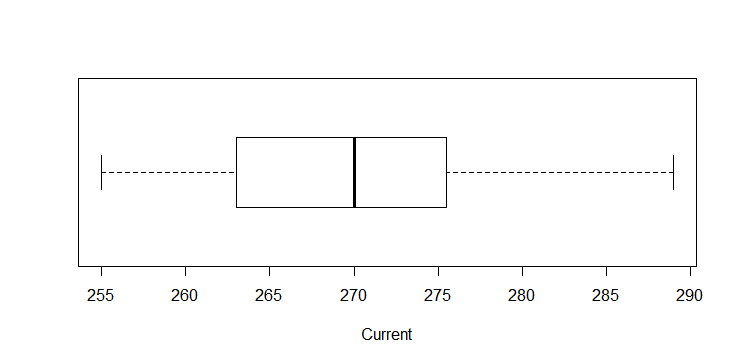
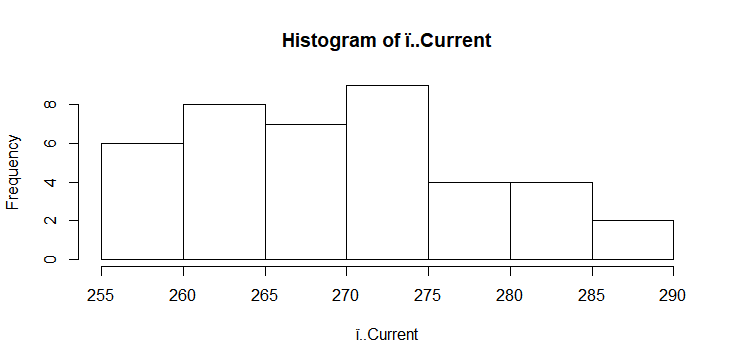
> var(New)

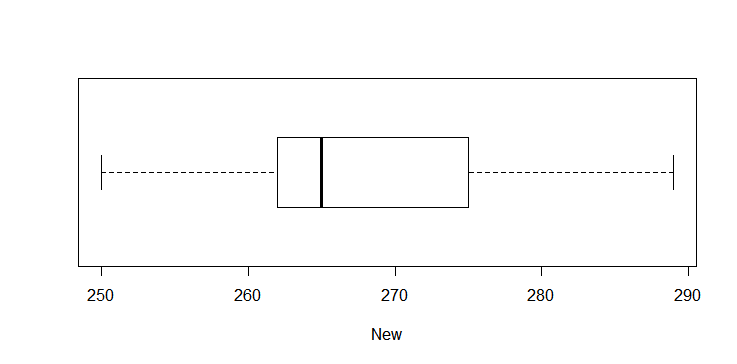
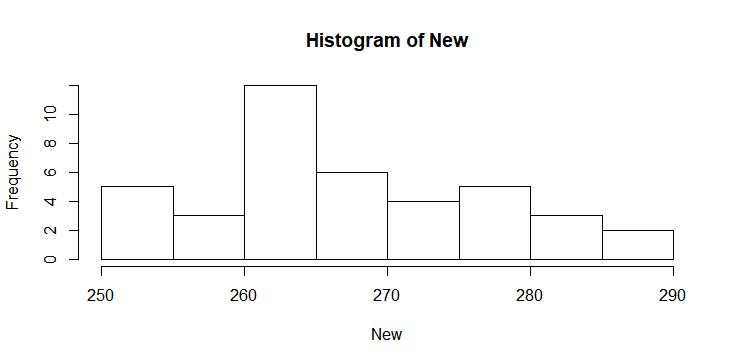
[1] 97.94872

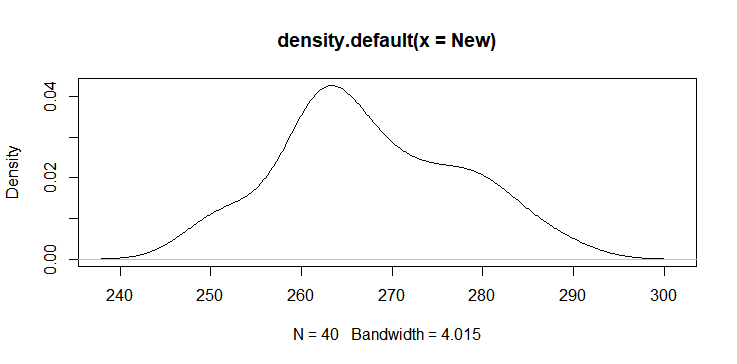
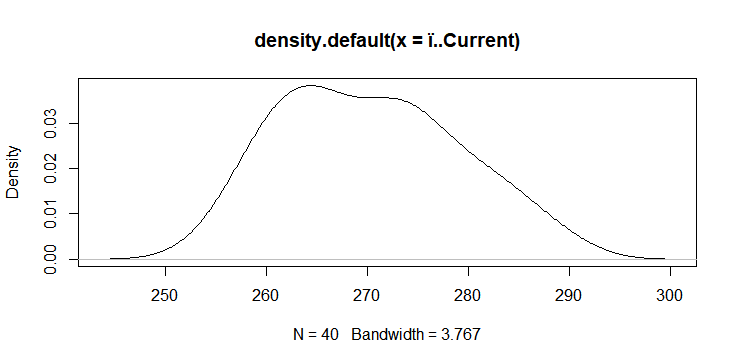
> sd(ï..Current-New)

> sd(ï..Current-New)

[1] 13.74397







* Both the samples seems to be normally distributed
* Mean and median are same for Current sample but mean is slightly higher than median for New sample .This mean that Current sample data is more normally distributed but new sample is skewed toward right .
* Variance for the new ball is 28% more then current ball .
  1. **95% confidence interval for each model and**

t.test also give the confidence level at 95%, we can make out that 95% of time the difference of driving distance is between -1.34 and 6.93 on the lower and higher side respectively .

t confidence interval:

-1.384937 6.934937

**95% confidence interval for difference between the mean.**

**This is calculated as**

**(Mu1-Mu2) +- T \* sqrt(SDa^2/Na+SDb^2/Nb)**

power= pt(.95,39) # 95% confidence interval.

power = .826 `

= Mu1 – Mu2 +- .826(.643)

= 2.8+- .5316

=(2.27,3.3)

**2.5** Since it T-test for the 2 sample , current size of sample 40 is good enough to deduce the insights from the sample data .

MuMdfdffdsdsd4343434

2.5 2DD