Golf Balls

Souradeep Mondal

02/08/2020

## Case Study: Par Inc.

Par Inc., is a major manufacturer of golf equipment. Management believes that Par’s market share could be increasedwith the introduction of a cut-resistant, longer-lasting golf ball. Therefore, the research group at Par has been investigating a new golf ball coating designed to resist cuts and provide a more durable ball. The tests with the coating have been promising. One of the researchers voiced concern about the effect of the new coating on driving distances. Par would like the new cut-resistant ball to offer driving distances comparable to those of the current-model golf ball. To compare the driving distances for the two balls, 40 balls of both the new and current models were subjected to distance tests. The testing was performed with a mechanical hitting machine so that any difference between the mean distances for the two models could be attributed to a difference in the design. The results of the tests, with distances measured to the nearest yard, are contained in the data set “Golf”. Prepare a Managerial Report

Golf<-read.csv("Golf.csv", header = TRUE)  
Golf

## Current New  
## 1 264 277  
## 2 261 269  
## 3 267 263  
## 4 272 266  
## 5 258 262  
## 6 283 251  
## 7 258 262  
## 8 266 289  
## 9 259 286  
## 10 270 264  
## 11 263 274  
## 12 264 266  
## 13 284 262  
## 14 263 271  
## 15 260 260  
## 16 283 281  
## 17 255 250  
## 18 272 263  
## 19 266 278  
## 20 268 264  
## 21 270 272  
## 22 287 259  
## 23 289 264  
## 24 280 280  
## 25 272 274  
## 26 275 281  
## 27 265 276  
## 28 260 269  
## 29 278 268  
## 30 275 262  
## 31 281 283  
## 32 274 250  
## 33 273 253  
## 34 263 260  
## 35 275 270  
## 36 267 263  
## 37 279 261  
## 38 274 255  
## 39 276 263  
## 40 262 279

## Let us take a look at the dataset closely

attach(Golf)  
dim(Golf)

## [1] 40 2

summary(Golf)

## 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

## Preliminary Analysis

Sample Size: 40, Number of Samples = 2 Measurement repeated on Current Golf Balls and New Cut-Resistant Golf Balls Observations:

> Both the samples seems to be normally distributed  
 > Mean and Median values are not that much different  
 > The Current Golf balls data are seems to be left skewed whereas sample of the New Golf balls are normally distributed

## Lets perform few statistical measures

## Standard Deviation  
sd(Current)

## [1] 8.752985

sd(New)

## [1] 9.896904

## Variance  
var(Current)

## [1] 76.61474

var(New)

## [1] 97.94872

## Inter Quartile Range  
IQR(Current)

## [1] 12.25

IQR(New)

## [1] 12.5

## Standard Error  
n = 40  
CurrentSE = (sd(Current)/sqrt(n))  
CurrentSE

## [1] 1.383968

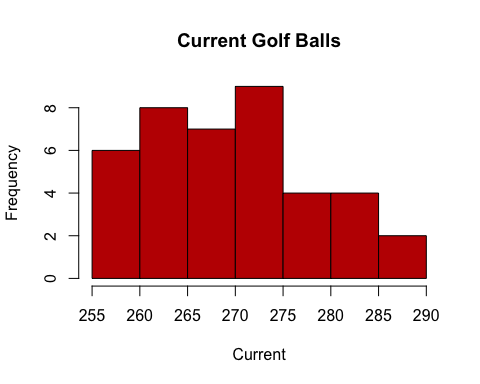
NewSE = (sd(New)/sqrt(n))  
NewSE

## [1] 1.564838

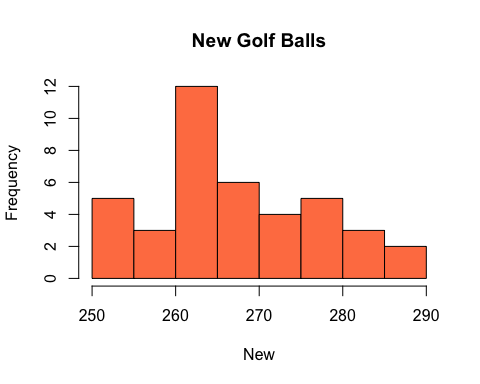
## Data Visualization

## Histograms

hist(Current, col ="#c00000", main = "Current Golf Balls")

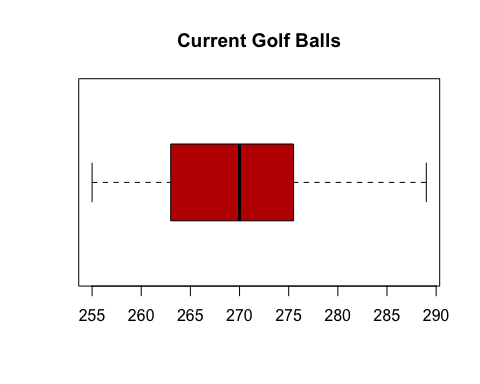


hist(New, col ="coral", main = "New Golf Balls")

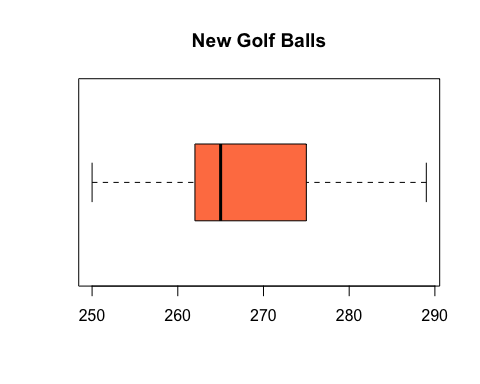


## Boxplots

boxplot(Current,col="#c00000",horizontal =TRUE, main="Current Golf Balls")



boxplot(New,col="coral",horizontal =TRUE, main="New Golf Balls")

 ## Hypothesis Testing Number of samples = 2 Both of the samples are Independent samples Sample Size, N =40 Level of Significance (Alpha) = 0.05 Since Both the Population Standard Deviation (Sigma) is unknown, but the populations are normally distributed we have to use a 2 Sample t-test Mean distance of current-model balls:𝜇current. Mean distance of new cut-resistant balls:𝜇new

## Hypothesis Formulation

NULL HYPOTHESIS  
H0:𝜇current =𝜇new(Mean distance of current balls equals mean distance of new balls)

ALTERNATIVE HYPOTHESIS H1:𝜇current ≠ 𝜇new(Mean distance of current balls is not equal mean distance of new balls)

The ≠ sign in the alternative hypothesis indicates that the test is two-tailed

To further test our hypothesis, we have set α at 0.05 and our rejection criteria is to Reject the Null Hypothesis and accept the Alternative Hypothesis i.e. to prove that there is actually a change in the mean distances for the two models could be attributed to a difference in the design

CI=t.test(Current,New,conf.level =0.95)  
CI

##   
## Welch Two Sample t-test  
##   
## data: 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

Degrees of freedom = 76.852 t = 1.3284 p-value = 0.188

Looking at the descriptive statistics for each samples, we can say that Current Golf balls has a longer range of distance based on the 40 samples with a mean of 270.3 compared to 267.5 for the New Golf balls. The standard deviation of the Current Golf balls is 8.75 and as for the New Golf balls it is 9.89. The mean differences between distances in Current and New golf balls is 2.775(270.3 - 267.5), not showing a significant difference to prove that there is a relevant change between the Current and the New Golf balls. Inter Quartile Ranges(IQR) of both the samples are hardly any different from each other to suggest that there is any significant change.

P-value=0.188>0.05=α. This suggests the following: Failed to reject the null hypothesis i.e. H0 Mean distance of the New cut-resistant Golf balls equalsmean distance of Current model Golf balls The new cut-resistant balls have no significant difference in distance compared to the Current model Golf balls

Thus, it can be recommended for Par Inc. to launch the New cut-resistant Golf balls as it is not impacting the distance aspect as suggested.

## For Current Golf ballat 95% confidence interval  
CICurrent=t.test(Current)  
CICurrent

##   
## One Sample t-test  
##   
## data: Current  
## t = 195.29, df = 39, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 267.4757 273.0743  
## sample estimates:  
## mean of x   
## 270.275

## For New cut-resistantGolf ballat 95% confidence interval  
CINew=t.test(New)  
CINew

##   
## One Sample t-test  
##   
## data: New  
## t = 170.94, df = 39, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 264.3348 270.6652  
## sample estimates:  
## mean of x   
## 267.5

Just by understanding from the confidence intervals of the 2 samples, it can be said that there is no significant difference between the driving distances of the Current and the New Golf balls.

## Is there a need of larger sample sizes and more testing with the golf balls?

P-valuefor this two-tailed test is 0.188, which is greater than level of significance i.e. α=0.05. Hence, H0 will not be rejected which shows that Par Inc. should take a new ball in production as the P-value indicates that there is no significant difference between estimated population mean of current golf ball as well as new golf ball.The 95% confidence interval for the population mean of the current golf ball is 267.4757to 273.0743 and of the model is 264.3348 to 270.6652.It means that the estimated population mean for Par, Inc. should lie within this range for consistent result. However, the 95% confidence interval for the difference between the means of the two populations is 2.775.

#Power of test for sample size calculation  
diff1=Current-New  
diff1

## [1] -13 -8 4 6 -4 32 -4 -23 -27 6 -11 -2 22 -8 0 2 5 9 -12  
## [20] 4 -2 28 25 0 -2 -6 -11 -9 10 13 -2 24 20 3 5 4 18 19  
## [39] 13 -17

detect=mean(diff1)  
detect

## [1] 2.775

sddiff=sd(diff1)  
sddiff

## [1] 13.74397

cohen.d=detect/sddiff  
cohen.d

## [1] 0.2019067

powertest=power.t.test(n=40,d=cohen.d,sig.level =0.05,power =NULL,alternative ="two.sided")  
powertest

##   
## Two-sample t test power calculation   
##   
## n = 40  
## delta = 0.2019067  
## sd = 1  
## sig.level = 0.05  
## power = 0.14274  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

powertestA=power.t.test(n=NULL,d=cohen.d,sig.level =0.05,power=0.95,alternative ="two.sided")  
powertestA

##   
## Two-sample t test power calculation   
##   
## n = 638.484  
## delta = 0.2019067  
## sd = 1  
## sig.level = 0.05  
## power = 0.95  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

powertestB=power.t.test(n=NULL,d=cohen.d,sig.level =0.05,power=0.90,alternative ="two.sided")  
powertestB

##   
## Two-sample t test power calculation   
##   
## n = 516.4577  
## delta = 0.2019067  
## sd = 1  
## sig.level = 0.05  
## power = 0.9  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

Power of test for the sample is 0.14274, which is significantly low for the sample size (n) of 40.Ideally, the Power should be around 0.90 to 0.95 for the sample size.The low power value suggests that the sample size of 40 is too small for this hypothesis testing. Thus, a bigger number sample size is required.

By, the above two data we can say that the present sample size not at all appropriate to prove the hypothesis.

#Conclusion From the given data, it may be concluded that, statistically there is no significant increase/decrease in distance because of the new cut-resistant golf ball. However, our recommendation is that the test be carried out with a larger sample size to improve the accuracy of test result. Other than testing only the driving distance, Par Inc. should also consider other factors like costing, availability of suppliers and the like before finalizing and launching the product if they are looking to test a larger sample size.