Mini Project – Golf's Cut resistant balls

Managerial Model Report

Author: Amit R. Kulkarni

Table of Contents

1. Information of the Case:

Par Inc. want to introduce a new cut resistant golf ball to increase the market share. The ball has been added an additional coating to prevent it from cutting and wear and tear.

Sample of 40 balls have been taken for both new and current ball to test if the newly added coating has been an impact on the distance of the ball travelled when compared to the current ball. Fair testing has been done using a machine so that the impact of the distance is only due to make of the ball.

Below is the data

Current	New
264	277
261	269
267	263
272	266
258	262
283	251
258	262
266	289
259	286
270	264
263	274
264	266
284	262
263	271
260	260
283	281
255	250
272	263
266	278
268	264
270	272
287	259
289	264
280	280
272	274
275	281
265	276
260	269
278	268
275	262
281	283
274	250
273	253
263	260
275	270

267	263
279	261
274	255
276	263
262	279

2. Objective:

The company wants to check if there is any difference in the distance of the ball travelled between the current and the new coated ball.

3. Data Set:

40 observations for both type of balls have been tested and the values are given in the above table.

4. Environment Set Up:

1.1 Installed required packages and invoke the libraries
1.2 Set Up the working directories
1.3 Import and read the dataset. Importing .csv dataset.
1.4 Microsoft Excel

5. Univariate Analysis:

The distance travelled by the balls is discrete variable.

6. Solution:

6.1 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

The balls before the tests and the balls used after the improvement with additional coating to make it cut resistant are independent of each other thus the paired mean test cannot be performed.

The observations are independent from each other and thus "*Testing between Means*" has to be done. A comparison between the current and the improvised balls mean has to be done and check if there has been any impact on the distance to be travelled after the coating.

Basis this we can build our hypothesis.

H_0: μ1 - μ2 = 0

H_1: μ1 - μ2 ≠ 0

Since the population distribution is known and the sample size is 40 applying the *Central Limit Theorem* we can conclude data is normal. ZSTAT test can be used.

Following information is calculated

```
> mean(Current)
[1] 270.275
> mean(New)
[1] 267.5
> sd(Current)
[1] 8.752985
> sd(New)
[1] 9.896904
```

σ_1 (standard Deviation of Distance travelled by the Current Ball) = 8.75

σ_2 (standard Deviation of Distance travelled by the New Ball) = 9.90

n_1 (number of observations for current balls) = 40

n_2 (number of observations for New balls) = 40

$\bar{X}1$ = 270.26

$\bar{X}2$ = 267.50

Standard error can be calculated as

$$\sigma_{M_1-M_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Solving the above equation we get

$\sqrt{}$ (8.75) ^2/40+ (9.90) ^2/40

= 76.56/40 + 98.01/40

= 1.91 + 2.45

= √4.36

Standard Error = 2.09

Calculating the Z value = ($\bar{X}1$ -$\bar{X}2$)-(**μ1 - μ2)**/Standard Error

= (270.28-267.50)/2.09

= 1.33

Since it is a 2 tailed test at 95% confidence interval the probability lies on both side of the distribution with 0.475 area on either side. Referring to the probability distribution value for 0.475 in Z table we get ±1.96. Since our Z value is 1.33 which is within the range of ±1.96 than our z value. Thus the **Ho** will not be rejected.

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

Since it is a 2 tailed test at 95% confidence interval the probability lies on both side of the distribution with 0.475 area on either side. Referring to the probability distribution value for 0.475 in Z table we get ±1.96. Since our Z value is 1.33 which is within the range of ±1.96 than our z value. Thus the **H_o** will not be rejected.

Calculating the p-value

√ (8.75) ^2/40+ (9.90) ^2/40

= 76.56/40 + 98.01/40

= 1.91 + 2.45

= √4.36

Standard Error = 2.09

Calculating the Z value = $(\bar{X}1 - \bar{X}2) - (\mu 1 - \mu 2)$/Standard Error

= (270.28-267.50)/2.09

= 1.33

Using the standard normal distribution table the area to the left of the upper tail distribution at value 1.33 = 1 – 0.9082 = 0.0918

Since it is a 2 tailed test p-value = 2 x 0.0918

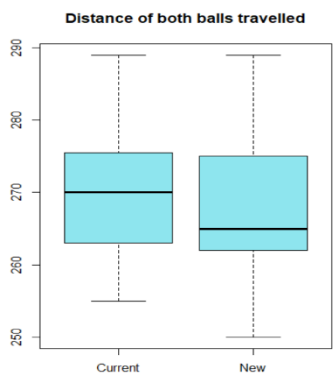
p-value = 0.1836

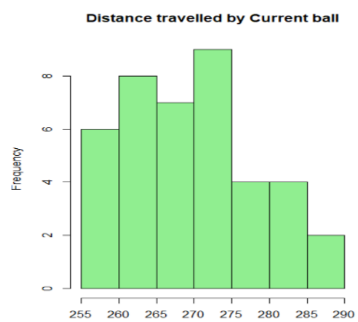
Since the p-value is greater than 0.05 H_o will not be rejected.

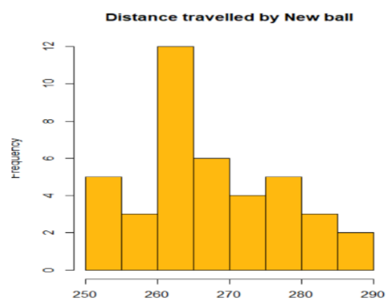
6.3 Provide descriptive statistical summaries of the data for each model

```
[1] "C:/Users/Amit Kulkarni/Documents"
> data<-read.csv("file:///C:/Users/Amit Kulkarni/Documents/R Programming/golf.csv")
> summary(data)
    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  
> |
```

Boxplot and Histograms:







Referring to the graphs above we can see that there is no much difference between the medians of both the population and also the data seems to be nearly normally distributed.

6.4 6.4 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?

For population Current balls below is the information provided

σ_1 (standard Deviation of Distance travelled by the Current Ball) = 8.75

n_1 (number of observations for current balls) = 40

$\bar{X}_1$ = 270.26

At 95% confidence interval the z value is ±1.96

We get the below

270.26 ± 8.75/√40

= 270.26 ±1.38

The mean value will be between 268.88 and 271.64 at 95% confidence

For population of New balls below information is provided

σ_2 (standard Deviation of Distance travelled by the New Ball) = 9.90

n_2 (number of observations for New balls) = 40

$\bar{X}_2$ = 267.50

At 95% confidence interval the z value is

We get the below equation

267.50 ± 9.90/√40

267.50 ± 1.57

The mean value will be between 265.93 and 269.07 at 95% confidence

What is the 95% confidence interval for the difference between the means of the two population?

$\sqrt{}$ (8.75) ^2/40+ (9.90) ^2/40

= 76.56/40 + 98.01/40

= 1.91 + 2.45

= √4.36

Standard Error = 2.09

At 95% confidence level we get the following

For Current Balls the mean distance travelled will be 268.19 to 272.37with 95% confidence

For New Balls the mean distance travelled with be 265.41 to 269.59 with 95% confidence

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

Since the data for 40 have been studied, central limit theorem can be used and as large as 30 samples are sufficient for testing purpose. In our case 40 cases each was used making it a normally distributed data. To do tests on more samples there is always costs involved. One should conduct a cost benefit analysis for testing more samples since it will come with a cost.

However to see what is the optimum level of samples required. It is calculated as below

For Current Balls distance travelled

*At 95% = (1.96*8.75) ^ 2 = ~294 cases needs to tested to be 95% confident*

For New Balls distance travelled

*At 95% = (1.96*9.90) ^ 2 = ~333 cases needs to tested to be 95% confident*

THE END