

|  |
| --- |
| Par Inc., Report  2019 |
|  |
| May 24  Par Inc.,  Authored by: Deepti Lobo Contact: +91 8722901118 lobo.deepti@gmail.com |



**CONTENTS**

[Description 3](#_Toc9572619)

[Assumptions 3](#_Toc9572620)

[Known Facts 3](#_Toc9572621)

[Hypothesis Statements 5](#_Toc9572622)

[Null Hypothesis 5](#_Toc9572623)

[Alternative Hypothesis 5](#_Toc9572624)

[Data Visualization 5](#_Toc9572625)

[Hypothesis Test 6](#_Toc9572626)

[P-Value 6](#_Toc9572627)

[Descriptive Summary 7](#_Toc9572628)

[Confidence Interval 9](#_Toc9572629)

[Reservations About the Results 10](#_Toc9572630)

[Type I & Type II Errors 10](#_Toc9572631)

[Power of The Test and Sample Size 10](#_Toc9572632)

[Sample Size to make Probabilities of Type I and Type II Error Equal 11](#_Toc9572633)

[Recommendation 12](#_Toc9572634)

# Description

Par Inc., is a major manufacturer of golf equipment. Management believes that Par’s market share could be increased with the introduction of a cut-resistant, longer-lasting golf ball.

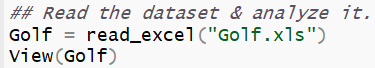
1. Formulate and present the rationale for a hypothesis test that par could use to compare the driving distances of the current and new golf balls.
2. Analyse the data to provide the hypothesis testing conclusion. What is the p-value for your test? What is your recommendation for Par Inc.?
3. Provide 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.

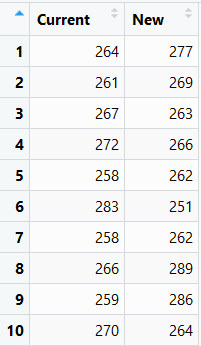
# Assumptions

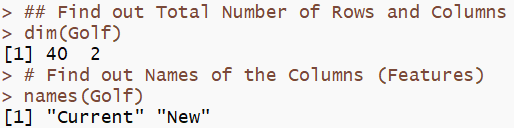
* Required Confidence Level is assumed to be 95%. (α = 0.05, C=0.95).
* We assume that the acceptance level of change in driving distance of 5 yards is acceptable to calculate the Type I and type II errors and required sample size.
* Same test is performed on both the population samples (new & current balls) to calculate the driving distances.

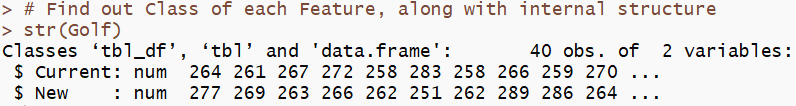
# Known Facts

The decision-making criteria which Par Inc., was considering was based on the comparative study of the driving distance covered by both balls. They have selected 40 balls from each population and subjected to same test and manipulated the driving distance datasets attached below. The 40 balls for each of the two sample sets are selected randomly from two different populations which are completely independent of each other.







The dataset has 40 rows and 2 columns. Both the columns are numeric in nature.

* Currents: containing driving distance of golf balls without coating.
* New: containing driving distances of gold balls with new coating.

For a confidence level of 95%, α = 0.05, from which we also know Z = 1.96.

|  |  |
| --- | --- |
|  |  |

Q1: 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.

# Hypothesis Statements

Par Inc., major objective is to increase their market share by introducing the new cut-resistant, longer-lasting golf ball, but doing so they are also concerned about the impact in the driving distance. Hence, they need to make sure that there is no difference in the mean driving distance of their new golf balls and the current existing ones. So, the hypothesis statements for this analysis would be as follows.,

## **Null Hypothesis**

The mean distance of current balls is equal to the mean distance of the newly redesigned balls.

H0 : µ1 = µ2

## **Alternative Hypothesis**

The mean distance of current balls is not equal to the mean driving distance of the newly redesigned balls. It can be increased or decreased, it doesn’t matter. So this is a two-sided Hypothesis test, as we just needed to check the impact.

Ha : µ1 <> µ2

**Note:**

µ1 = Mean driving distance of current balls

µ2 = Mean driving distance of new cut resistance balls

# Data Visualization

Performing a comparative statistical analysis of the driving distance data provided by Par Inc.,

|  |  |
| --- | --- |
| **CURRENT BALLS** | **NEW BALLS** |
|  |  |
|  | s |
|  |  |

From the distribution plots & histogram, it's evident that both the current and new balls driving distance data are of normally distributed and the data is not skewed much.

From the box plot, it’s evident that there are no significant outliers in the datasets.

# Hypothesis Test

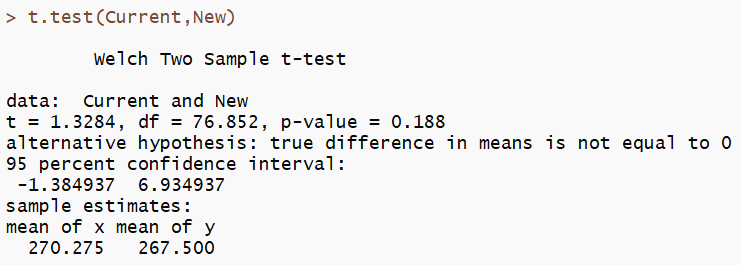
Q2. Analyse the data to provide the hypothesis testing conclusion. What is

the p-value for your test? What is your recommendation for Par Inc.?

## **P-Value**

Currently we know that the sample distribution of both the current and new balls are normal (Parametric datasets) and that they are completely independent of each other.

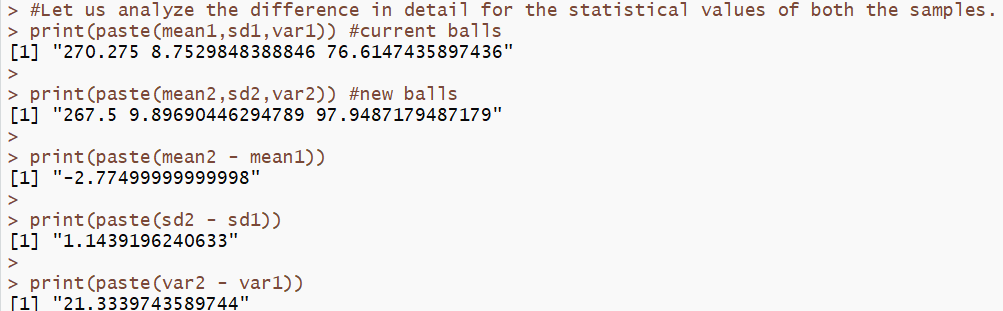
So, for this scenario, we can make use of the **Two Tailed Independent Two Sample** **T Test** to calculate the P-Value



The P-Value (0.188[two-tailed]) > α (0.05), so the **H0 cannot be rejected**. Based on our null hypothesis, now we can say that the mean driving distance of both the new model balls and the current models are equal.

# Descriptive Summary

Q3. Provide descriptive statistical summaries of the data for each model.



The difference in values between the mean and variance of both the sample datasets are minimal.

Considering the sample distribution is subjected to two-sided test, we can calculate the upper & lower limit population means of each model seprately as shown below.

|  |  |
| --- | --- |
| CURRENT BALLS | NEW BALLS |
|  |  |
|  |  |
|  |  |

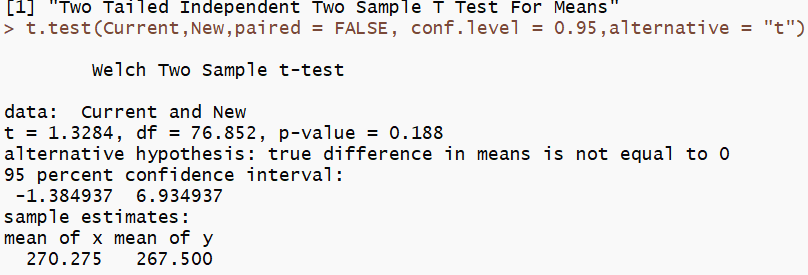
|  |  |  |
| --- | --- | --- |
|  | CURRENT BALLS | NEW BALLS |
| Sample Count | 40 | 40 |
| Mean | 270.275 | 267.5 |
| Sample Variance | 76.614 | 97.948 |
| Sample SD | 8.75 | 9.89 |
| @95% Confidence - α | 0.05 (0.025 + 0.025) | 0.05 (0.025 + 0.025) |
| @95% Confidence - Z | 1.96 | 1.96 |
| Standard Error σ/√n | 1.383 | 1.564 |
| Lower-Limit µ | 267.207 | 264.432 |
| Upper- Limit µ | 272.987 | 270.212 |
|  |  |  |

# Confidence Interval

**Q4. What is the 95% confidence interval for the population mean of each**

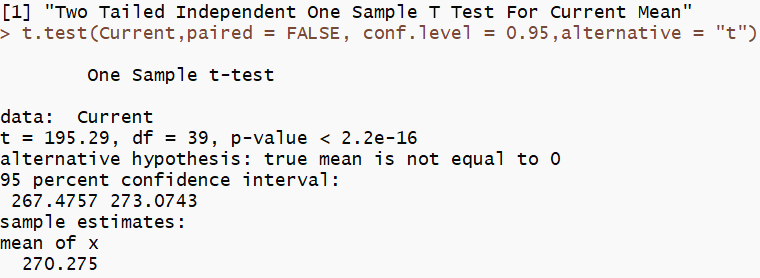
**model, and what is the 95% confidence interval for the difference between**

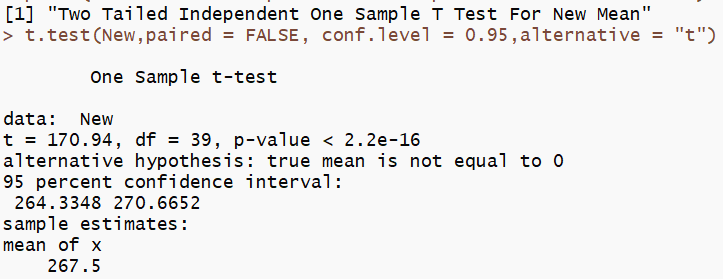
**the means of the two population?**



A two tailed Two Sample T Test was conducted for the difference in means.

* 95% confidence interval for difference in mean is [-1.384937 To 6.934937]





A two tailed One Sample T Test was conducted for Current and New Balls respectively.

* 95% confidence interval for Current balls driving distance mean is [267.4757 To 273.0743]
* 95% confidence interval for New balls driving distance mean is [264.3348 To 270.6652]

# Reservations About the Results

Q5. Do you see a need for larger sample sizes and more testing with the golf balls? Discuss.

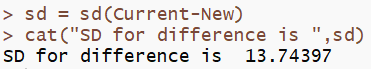
* If we compare the means of the two sample distributions, we see that even though visually it seems as if New coating has effect on the driving distances, statistically it does not.
* The difference in mean in the case of new balls can also be attributed to the higher variance compared to Current balls.
* The variance of New balls driving distances is 97.948 is 28% more than the variance of the driving distances of Current balls 76.61.
* Statistically there is no effect of new coating on driving distances. Though it is suggested to check the effect on the weights and other characteristics like size and shape of the new balls.
* Also, the given sample is from only one golf course, it is advisable that test should perform on different kind of golf courses to take care of the differences in grounds.

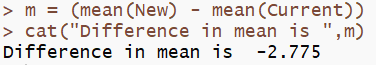
## **Type I & Type II Errors**

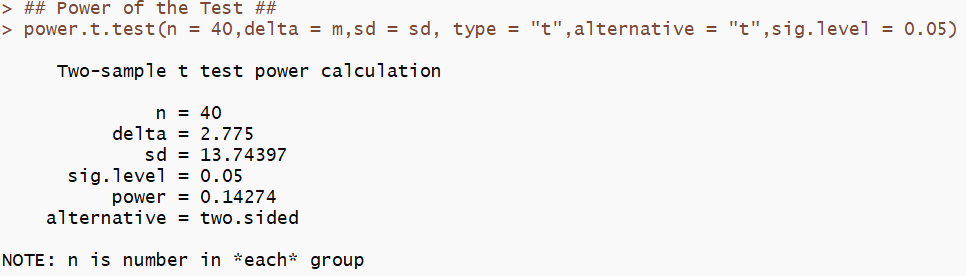
* **Type I Error**α**:** Probability of rejecting null hypothesis when it is true, the probability of a Type I error in hypothesis testing is predetermined by the significance level.
* **Type II error**β**:** Probability of falling to reject the null when it is false. Type II error calculation **depends on the population mean which is unknown**.

## **Power of The Test and Sample Size**

* If alternative hypothesis µ2 − µ1 = µd = 5 yard as per our assumption.
* Null Hypothesis µ2 − µ1 = µd = 0
* First, we need to calculate the probability of Type I error which is predetermined by significance level. If the significance level is 0.05,
* Then Type I error is 0.05 i.e. 5% probability we make Type I error – rejecting null hypothesis when it is true.
* Type II error calculation depends on the value of µ. In this case let’s assume difference between population µ is 5 yards. Let’s also assume that the significance level for the test is 0.05. Then the calculation is as below:
* This is a two tailed test.







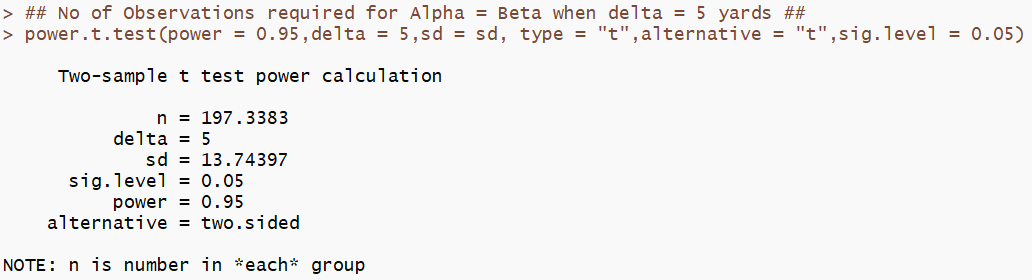
Basically, the power of the test is the probability that we make the right decision when the null is not correct (i.e. we correctly reject it).

## **Sample Size to make Probabilities of Type I and Type II Error Equal**

Let us assume that, we need Type I error and Type II error equal to 0.05

Assuming sample standard deviation is equal to population standard deviation, we can calculate sample size needed as below:

* Null hypothesis’ mean difference µ0 is 0.
* Alternative hypothesis’ mean difference µ1 is 5.
* Sample Standard Deviation is 13.74397.
* α value is 0.05.
* β value is 0.05 i.e. power of the test is 0.95 = 95 %.



Hence, to retain the power, we need to round the value to next whole number. Therefore, we may conclude that we need a sample size of 198 to get the Type I and Type II Errors equal.

# Recommendation

|  |
| --- |
| From the given data, we may statistically conclude that, there is no significant change in driving distance, due to the new coating on the golf balls.  Par Inc., can launch the new model marketing that the model upgrade is done only to improve the ball lifetime and the driving distance will not be increased which the customers usually expect by default when a new golf ball model is launched.  However, our recommendation is that, the test should be carried out with a larger sample size covering number of golf courses (at least five different courses) to improve the accuracy of the test results and negating any effect from one type of ground.  Also, the results need to be interpreted and future actions to be planned with an understanding of the other characteristics of the ball like size, weight, etc. |