# MAT 243 Project Two

Ryan Hatch  
[ryan.hatch@snhu](mailto:ryan.hatch@snhu)  
Southern New Hampshire University

## Introduction: Problem Statement

In this project, I am trying to statistically validate the claims made on the performance of the Los Angeles Lakers basketball team between the periods of 2013-2015. The dataset used contains point scoring records, relative skill ratings of the teams, or the Elo rating, and game outcomes for the teams and seasons, which basically is a historical record of NBA games. The data frame for the Bulls (assigned team,) is “assigned\_team\_df” and the data frame for the Lakers (my chosen team) is “your\_team\_df.” I will be using four variables in the data set to study, the points scored by the team in a game, the measure of relative skill level of the team in the league, the year when the team played the games, and the name of the NBA team.

In order to parse the data that I needed to solve my problem, I applied a number of statistical methods in the analysis, including the testing of the hypothesis. In the analysis, I used a one-sample t-tests to review whether the average relative level of the skills of the two years and the average points scored by the Lakers in the chosen years of 2013 – 2015, and how they differ from hypothesized values. I carried out a proportion z-test to test if the proportion of games won when the score was over a certain number of points is actually a match as it claimed. I then went to conduct a two-sample t-test to establish if the relative skill level mean difference of the Lakers and that of the Chicago Bulls was significant during the years that were used. The results of my analysis would have either confirmed the claims for the performance of the teams or it would have rejected them and help guide with making more informed and tailored decisions in order to improve the strategies used by the teams.

## Introduction: Your Team and the Assigned Team

For my analysis I chose the Los Angeles Lakers for the years between 2013 – 2015. The assigned team I was given for comparative study was the Chicago Bulls, which covered the years between 1996 – 1998.

*Table 1. Information on the Teams*

|  | **Name of Team** | **Years Picked** |
| --- | --- | --- |
| 1. Yours | Lakers | 2013 - 2015 |
| 2. Assigned | Bulls | 1996- 1998 |

## Hypothesis Test for the Population Mean (I)

In hypothesis testing, I used python libraries in order to compute statistical methods to double verify claims about a population mean. This is considered a null and alternative hypothesis test, which ends up forming a claim that needs to be verified. It sets up the null hypothesis, which is representing the claim that one would want to test against an alternative hypothesis which would represent the outcome he or she would expect or would like to prove. The level of significance is then chosen to determine the threshold for rejecting the null hypothesis.

**Null Hypothesis (H0):** The Lakers' average relative skill level for the years 2013-2015 is 1340 and below.

**(H0: μ ≤ 1340)**

**Alternative Hypothesis (H1):** The average number of relative skills levels of Lakers players across the span of the years 2013-2015 is not less than 1340.

**(μ > 1340)**

The degree of significance set for this test was at 5%, which meant that there was a 5% chance of wrongfully rejecting the null hypothesis.

*Table 2: Hypothesis Test for the Population Mean (I)*

| Statistic | Value |
| --- | --- |
| Test Statistic | 2.94 |
| P-value | 0.0034 |

The p-value, 0.0034, lies in the rejection region as it is less than the level of significance, 0.05; therefor, rejecting the null hypothesis. That is, for this specified level of significance, I would conclude that there is enough statistical evidence to say that the average relative skill of the Lakers team during 2013-2015 is more than 1340.  
  
The practical significance of this result is that it demonstrates the Lakers' performance in terms of relative skill level was above a critical low threshold during the specified years. This information can be valuable for the team's management when making decisions in terms of team improvement and strategic planning.

## Hypothesis Test for the Population Mean (II)

The null hypothesis tested the claim that the means of the average points of the Lakers from the year 2013 to 2015 had not been less than 106 points per game. I conducted this test at the 1% level of significance, assuming an unknown population standard deviation for the scores registered.

**Null Hypothesis (H0):** The Lakers have averaged 106 or more points for the seasons 2013-2015.

**(H0: μ ≥ 106)**

**Alternative hypothesis (H1):** The average of Lakers' points during 2013-2015 is less than 106.

**(H1: μ < 106)**

With the level of significance at 1% (or, α = 0.01) and since the P-value (or, 0.0000) is way less than the level of significance (0.01), I rejected the null hypothesis. This was because there was enough evidence to prove the claim of the coach, which was that the average number of points scored by the Lakers for 2013-2015 is less than 106. The practical implication of this is that it points out the poor scoring performance by the Lakers during the specified years, which is below the threshold associated with strong performance in the regular season. This was valuable information for the management team and coaching staff in order to properly identify and tailor what may need to be done in terms of offensive strategy and player development in order to get better results.

*Table 3: Hypothesis Test for the Population Mean (II)*

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -6.91 |
| P-value | 0.0000 |

## Hypothesis Test for the Population Proportion

The claim that was made while analyzing the performance of the Los Angeles Lakers from 2013 to 2015 was the efficiency of this team in the games where they scored more than 102 points.

The management claimed that the Lakers would have an 0.90 probability of winning the game if they scored above 102 points per game. In order to check their hypothesis, I tested the population proportion was at the 5% level of significance.

**Null Hypothesis (H0):** The proportion of the Lakers won games, having scored more than 102 points, is equal to 0.90

**(P=0.90).**

**Alternative Hypothesis (H1):** The proportion of games won by the Lakers when they score over 102 points should not be 0.90.

**(P ≠ 0.90)**

This analysis was done at a level of significance at 5%, which meant that there was a 5% chance that if the null hypothesis was actually true, it would be rejected. And based on the test statistic and the P-value, the null hypothesis would be rejected. The P-value of 0.0000 is very small when comparing it to the level of significance being at 0.05. With this, it strongly suggests a rejection of the null hypothesis. This hypothesis test brought me to the conclusion that the claim made, which was that the Lakers won 90% of the games whenever they had scored over 102, is not possible based off of the data. This has practical significance because of the fact that the differences between these scores, both the perceived and true performance, become quite large in high scoring games.

*Table 4: Hypothesis Test for the Population Proportion*

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -2235.66 |
| P-value | 0.0000 |

## Hypothesis Test for the Difference Between Two Population Means

This would help in testing claims regarding the difference of two population means in a hypothesis testing setting. This calls for setting a null hypothesis, which presumes no difference between the means; and an alternative hypothesis, that two means do show a significant difference. The level of significance is chosen to determine the threshold for rejecting the null hypothesis.

**Null Hypothesis (H0):** There is no difference in the relative skill level between the Lakers in 2013-2015 and the Bulls in 1996-1998.   
**(H0: μLakers = μBulls)**

**The alternative hypothesis (H1):** The mean relative skill level of the Lakers for the years 2013-2015 is not equal to the mean relative skill level of the Bulls for the years 1996-1998.   
**(H1: μLakers ≠ μBulls)**   
  
The significance level adopted for this test was 1%. That is, with the expectation that the null hypothesis is actually not true but is being rejected at a 1% chance.

Since the P-value (0.0000) is less than the level of significance (0.01), we reject the null hypothesis. There is, therefore, sufficient evidence to show that the relative skill level of the Lakers between 2013 and 2015 mean relative to that of the Bulls between 1996 and 1998 differs. The practical implication of the finding is that it brings out a very big difference in the skill levels between the two teams during the corresponding periods. This information can be useful for team management to understand the position of the team as compared to a historically successful team and then devise plans with respect to improving the team or the strategy.

*Table 5: Hypothesis Test for the Difference Between Two Population Means*

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 49.51 |
| P-value | 0.0000 |

## Conclusion

The analysis that was conducted in this project helps to bring more awareness of the Los Angeles Lakers' performance from the years 2013 to 2015. The hypothesis tests conducted helped me to be able to confirm or refute various assertions that were made about the team's performance.

The first hypothesis test helped prove to me that the Lakers' average relative skill level was significantly higher than a critically low threshold, which suggested that the team's skill level during this time period was higher than I initially expected. The second hypothesis test indicated that the team's average points scored per game were significantly lower than 106, which implied that the team's offensive performance was not as strong as desired. The third hypothesis test showed to me that the proportion of games that were won by the Lakers when scoring more than 102 points was significantly lower than 90%. This conclusion was able to help me verify the claim about the team's efficiency in high scoring games.

Also, when comparing the data between the Lakers and the historically successful Chicago Bulls, it showed that there indeed was a significant difference in their relative skill levels. This helped to highlight and bring awareness to the performance gap between the two teams during the selected periods.

The practical significance of the analyses that were conducted is the ability to help bring some awareness to the management and to the coaches on how they can specifically tailor any changes or make any decisions to improve the team’s performance based on the statistics from the parsed data. None the less, the results of these types of analysis easily offer a statistical foundation for evaluating and assessing the team's performance before identifying any potential areas for improvement. Overall, in utilizing statistical data in order to conduct tailored analysis helps to better understand where the strengths and weaknesses are, along with helping the coaches and staff to make more informed and tailored change in order to get the most optimal return out of the players development, strategy, and even overall to improve the team as a whole unit, rather than individually as players. The comparison made between the two teams also provide a benchmark for the Lakers to strive for. Using the data to bring more confidence into the players motivation, can help to contribute to the team's growth and success.