# MAT 243 Project Two Summary Report

Benjamin Leanna

benjamin.leanna@snhu.edu

Southern New Hampshire University

## Introduction: Problem Statement

In this second project I am adding onto our first analysis of the two original teams I was assigned and picked, the Chicago Bulls and the Miami Heat. Again, I am assuming the data analyst role a basketball team and looking to analyze performance patterns between the two. I will be using Python to perform statistical analyses and report my findings for the team’s management.

I will be using different variables including points scored by the team in a game and measurements of relative skill level of the teams in the league throughout specific years when the teams played their games.

The different statistical methods I’ll be using to compare these two teams are ‘ztests’ that determine whether two population means are different when the variances are known and the sample size is large and ‘ttests’ that compares the means of the two samples.

Below is a table of my picked team and assigned team along with their corresponding years played int the league.

Table 1. Information on the Teams

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

## Hypothesis Test for the Population Mean (I)

In the first hypothesis test I am supposing that the relative skill level of 1340 represents a critically low skill level in the league. The management of my team has hypothesized that the average relative skill level of my team is greater than 1340. I am testing the claim using a 5% level of significance. I am also assuming that the population standard deviation for relative skill level is unknown.

Hypothesis testing is used to evaluate two mutually exclusive statements about a population to determine which statement is best supported and then interpret or draw conclusinos about the population using same data.

In this test we are using a null hypothesis, proposing that there is no difference between certain characteristics of a population process, where the average relative skill level of my team is greater than 1340 and an alternative hypothesis, proposing that there is a difference.

Below is a table that measures two points in this test, the test statistic that describes how far my observed data is from the null hypothesis and the P-value which serves as a rejection point to the significance level.

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

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 43.55 |
| P-value | 0.0 |

Looking at this conclusion with the P-value at 0.0 we can conclude on this information that the team’s relative skill level is in fact higher than the 1340 the league set as a standard for critically low.

## Hypothesis Test for the Population Mean (II)

In this next test my team’s coach has now hypothesized that the average number of points scored by my team in the team’s years is less than 106 points. This is what I’ll use as the null hypothesis and the alternate being my team scored higher for those years. Again, I’m assuming that the population standard deviation for points scored is unknown. I am now testing the claim using only a 1% level of significance.

Below here is another table showing the test statistic and P-value of these tests. We can again conclude that the null hypothesis can be rejected and that the teams average score is in indeed lower than the 106 points.

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

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -8.7 |
| P-value | 0.0 |

## Hypothesis Test for the Population Proportion

In this next hypothesis the management claims that the proportion of games that your team wins when scoring 102 or more points is 90%. I’ll be testing this claim with a 5% level of significance to determine the results.

For the calculation of this I’ll be counting the number of game results that ended in a win with the number of total games up against the 102 points. When calculated, it comes out that the percentage of wins by my team when scoring 102 points in a game is 89.2%. I’ve included a table to see the test statistic and P-value as well (which is 0.0) showing that in fact my team does not win over 90% of the time when scoring 120 points in a game.

Table 4: Hypothesis Test for the Population Proportion

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -3404.8 |
| P-value | 0.0 |

## Hypothesis Test for the Difference Between Two Population Means

I will now take the comparison of my team’s skill level of its years (2013-2015) alongside the assigned team’s skill level from their assigned time frame (1996-1998). Here I’ve tested the claim that the skill level of my team is the same as the skill level of the assigned team using a 1% level of significance.

This test is a bit different that the others. In the other examples we had a one tail argument that the results would either be higher of lower. Here we are simply comparing the two teams to see if their skill is the same.

As you can see in the table below our test statistic is very low and our P-value is yet again, zero. Based off this information, since the P-value is lower than the 1% so we can reject the null hypothesis that the team’s average skill level is the same.

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 17.07 |
| P-value | 0.0 |

## Conclusion

In this conclusion, throughout the analysis of the two teams, one can see that the null hypothesis each time was rejected. This showed that the Miami Heat in the years of 2013 to 2015 have scored higher than the relative skill level set by the league, have scored less than the 106 points on average during the regular season, does not win more than 90% of the time when scoring over 102 points in a game, and does not have the same skill level that the Chicago Bulls had in their seasons from 1996 to 1998.

This information is important because if we take the assigned team of the Chicago Bulls, whom were the best team in the league during the years assigned and wanted to make a team that had a higher skill level, then statistically we can compare different parts to see where improvement is needed and adjust. Team’s can also use this information to compare the best teams nowadays to the best teams throughout history and see how the skill level has either increased or decreased. Using this as well, a single team can compare every year it’s been in the league and see the progression (or regression) to further improve.

## Citations

*What Is a Z-Test?* (2021, November 8). Investopedia. https://www.investopedia.com/terms/z/z-test.asp

Bevans, R. (2022, July 9). *An introduction to t-tests*. Scribbr. https://www.scribbr.com/statistics/t-test/

Scribbr. (2022, May 6). *What is a test statistic?* https://www.scribbr.com/frequently-asked-questions/what-is-a-test-statistic/

Mahesh (2021, December 11). *Everything You Need To Know about Hypothesis Testing — Part I*. Medium. https://towardsdatascience.com/everything-you-need-to-know-about-hypothesis-testing-part-i-4de9abebbc8a