NBA: A Look into the Three-Pointer (Part 2)

Quang Nguyen

Question:

Looking at the early years of the introduction of the 3-point shot to the now prolific 3-point era led by guards Steph Curry and Klay Thompson who consistently break 3-point scoring records, have players shot-making abilities significantly changed in recent years to make more 3-point shots than ever before or is it solely due to a greater amount of attempts?

Data Description:

This dataset is an extensive collection of individual statistics for basketball players in the NBA from seasons ending in 1950 to 2017. These individual statistics were compiled from gamelogs, box-scores, and with access to the NBA developer tools [1]. Each row within the raw dataset represents an individual player within the respective season. Each column within the raw dataset for the most part represents a unique statistic tracked by basketball analysts to help generate performance metrics of how well individual basketball players play. Although the raw data that is read in is considered a population and contains statistics from all NBA players from 1950 (the year teams from the Basketball Association of America and the National Basketball League consolidated into the National Basketball Association) to 2017, the data that I will be testing is a small subset of the data [3].

The subsets that I chose to analyze are the 3 point percentages which is calculated by dividing total 3 point shots made by 3 point shots attempted across the league from two 10 season samples - 1985-1995 and 2005-2015 [2]. The reason these two sample of seasons are chosen is it provides variability within the data as the earlier sample demonstrates years that came recently after the introduction of the 3 point shot in 1980 season (which began in 1979) while the latter years coming in the boom of the 3 point shot in media and coverage. This data will help solve the question posed as to whether their shot-making abilities have changed as the data uses percentages rather than just total counts which would not be comparable across the years as that would show alot of variation (i.e. 2011 season that was shortened due to a player lockout versus a normal 81 game season) [4]. By determining if the difference in 3-point shot percentages is statistically different, conclusions can be made as to whether the

increase in 3-point shots made is actually due to more accurate shooting or if it was simply caused by an increase in the number of shots taken.

Test Rationale:

The test I will conduct on my data is an unpaired two-sample t-test which compares two different subjects/samples. Although the data is from the same league, the NBA, every season is different and the ten seasons from the 1985 to 1995 are highly distinct from the seasons between 2005 and 2015 except for a select few players who have had long careers such as Kevin Willis and Kevin Garnett [5]. For this scenario, the t-test is chosen as it will useful to determine whether there is a significant difference between the two sample groups and is commonly used for smaller sample sizes. Additionally, the data of the percentages when plotted is representative for the most part of a normal bell-shaped distribution curve and the variances between the two samples that I will be comparing are not significantly different. The outcome of this test will be able to be generalized to the applicable of the population (after 1980) as it is testing for the significance of the difference between two subset ranges that cover the earlier years of the 3 pointer to the most recent years as well as a fairly large number of years compared to the number of years available for analysis. Thus, by using the outcome of the test which determines whether there is a significant difference between the shooting percentages of the samples can be used to generalize the positive trend of three points made that has been seen since 1980.

Test: Unpaired two-sample t-test

```
setwd('/Users/quang/Desktop/STAT3080/project')
nba.raw = read.csv("nbastats.csv")[-c(1,10:31,41)]
year.list = split(nba.raw, nba.raw$Year)
threepointattempts = sapply(year.list, function(x) sum(x$X3PA))
threepointmade = sapply(year.list, function(x) sum(x$X3P))
threepointpercentages = threepointmade / threepointattempts
df.tpp = data.frame(threepointpercentages)
tppsamp8595 = df.tpp[c(36:45),]
tppsamp0515 = df.tpp[c(56:65),]
res <- t.test(tppsamp0515, tppsamp8595,  mu=0, alternative="two.sided")
res
```

```
##
##  Welch Two Sample t-test
##
## data:  tppsamp0515 and tppsamp8595
## t = 6.5019, df = 10.081, p-value = 6.628e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.02782407 0.05678845
## sample estimates:
## mean of x mean of y
## 0.3574727 0.3151664
```

Conclusion

From the test that I ran on the two groupings of 10 NBA seasons from 1985 to 1995 and 2005 to 2015, I confirmed that there is a noteworthy difference between the 3-point percentages from the earlier years which had lower 3-point percentages as compared to the latter years which had higher 3-point percentages.Using that confirmation, I came to the following conclusions to answer the question that was posed earlier in the project. The consistent increase in 3-points made since the introduction of the shot is not coincidental. This result can be generalized to the seasons that were not included in the tests since 1980 so by looking at the increasing number of 3 points made, I can reasonably deduce that this (so far) gradual trend has also been a result of better shot-making skills by the players themselves. Using my conclusions from the test, I also can infer that over time with more data to analyze and a greater understanding of the 3-point shot, coaches and players alike have seen that the benefits of taking a 3-pointer versus a 2-pointer is worth the risk. This has most likely brought more opportunities for plays to be made that center around the shot so players have been able to get more in-game opportunities and practice with the 3-point shot which in turn betters many players shot-making abilities. This test does not conclude that shot-making abilities and percentages will continue to increase, rather this simply generalizes that the 3-point has seen a rise in importance since its introduction into the NBA. Further research into this topic could delve into a multitude of include evaluating whether or not the 3-point shot has contributed to more points per minute throughout the league, comparing the effectiveness of the 3-point shot versus a normal field goal, and even whether or not the 3-point shot has led to a faster pace of game and more possessions on average. The possibilities are wide and far with the NBA data.

References [#]

1. Dataset retrieved from Kaggle and was mined from Basketball-Reference [1]:
   https://www.kaggle.com/drgilermo/nba-players-stats/home
   https://www.basketball-reference.com/
2. Information about basketball specific metrics [2]:
   https://www.basketball-reference.com/about/glossary.html
3. Information about NBA rule changes throughout history [3]:
   http://www.nba.com/analysis/rules_history.html
4. Article from CNN archives about lockouts and strikes in sports [4]:
   https://www.cnn.com/2013/09/03/us/pro-sports-lockouts-and-strikes-fast-facts/index.html
5. Article about oldest and longest playing NBA players [5]: http://www.oldest.org/sports/nba-players/