Project Part 3

Sam Switz (sds9bt)

Question:

Last summer (2019), the NBA underwent what was one of the biggest free agency periods in its history. Roughly 40 percent of players were, for the most part, free to sign with any other team. Not every free agent changed teams but a number of elite players did end up moving, with Kyrie Irving, Kevin Durant and Kawhi Leonard among others changing teams. For the past 5-10 years, the consensus has been that the Western Conference has more quality teams than the Eastern Conference and therefore is more competitive. Given that roughly 40% of players had expiring contracts last season (2019), I wanted to see if the balance of power between the two conferences had changed at all.

The Data

I used data from FiveThirtyEight and landofbasketball.com to create the data frame for my analysis. I used landofbasketball.com to create columns for whether or not a team made the playoffs, and if they did what seed they were for each year. FiveThirtyEight combines traditional box-score statistics along with more advanced player-tracking data from the NBA and calculates advanced statistics for individual players, RAPTOR and Wins Above Replacement (WAR). RAPTOR is a plus-minus statistic that is supposed to reflect how the modern NBA values players and it is used to calculate WAR. I chose WAR because it's focused on wins, which ultimately determines how good a team is. It gives me a way to measure the overall quality of a team by taking the average of all players while maintaining a player-centric view. I wanted to approach my analysis with a focus on players because my question is about players with expiring contracts. I selected data from the 2014 season to the most recent season because 2014 was the first time player-tracking data was made available, which RAPTOR (and therefore WAR) is heavily reliant upon. Additionally, I will only be using playoff teams because that's where an imbalance of power would have the biggest impact on the league, as each conference has 8 teams qualify for the playoffs.

Generalization

The outcome of the test will determine if the percentage of NBA players with expiring contracts is large enough for a hypothetical free agent class to alter the power balance between the two conferences in the NBA for the next few seasons (as in after a while the results of this test would not necessarily be applicable to the league).

The Test

I will be performing a two-sample randomization test on paired data to try to answer my question. I chose this over a two-sample t-test because my sample size is only 16 and after looking at a histogram, I determined that the data are not normally distributed. I will be pairing Eastern and Western Conference playoff teams based on their seed and then subtracting the Eastern Conference team's average WAR from its Western Conference counterpart and recording the difference. Pairing the data makes sense because if the 6th best Western Conference team was significantly better than the 6th best Eastern Conference team, it would support the claim that one conference is better than the other. However, if the best Western Conference team was significantly better than the worst Eastern Conference team, that wouldn't be helpful in showing an imbalance of power. The assumptions of the test are met as well: the sample (2020 playoff teams) is independent of the population (playoff teams from 2014-2020), the sample is representative of the population as it has the same number of teams from each conference and the population distribution does not have heavy tails. The last assumption I was able to verify by looking at a histogram.

- Hypothesis: the mean of the paired differences for 2020 is different than the mean of the paired differences for the seasons 2014-2019
- Null: the means are the same
- Alpha = 0.05

```
playoffs2 <- playoffs[order(seed),]
p.2020 <- playoffs2[which(playoffs2$year_id == 2020),]
p.rest <- playoffs2[which(playoffs2$year_id < 2020),]

p.2020.w <- p.2020[which(p.2020$conference == "West"),]
p.2020.e <- p.2020[which(p.2020$conference == "East"),]
diffs.2020 <- p.2020.w$avg.war - p.2020.e$avg.war
```

```
p.rest.w <- p.rest[which(p.rest$conference == "West"),]
p.rest.e <- p.rest[which(p.rest$conference == "East"),]
diffs.rest <- p.rest.w$avg.war - p.rest.e$avg.war

samp.diff <- mean(diffs.2020) - mean(diffs.rest)

rand.test <- function(x){
  rand_comb <- sample( c(diffs.2020, diffs.rest) )
  bmean1 <- mean(rand_comb[1:x])
  bmean2 <- mean(rand_comb[(x+1):(length(diffs.2020) + length(diffs.rest))])
  bmean1 - bmean2
}

boot.diffs.null <- replicate(10000, rand.test(length(diffs.2020)))

mean(boot.diffs.null <= samp.diff |
        boot.diffs.null >= 2*mean(boot.diffs.null) - samp.diff)

## [1] 0.1612
```

Test Conclusion

We fail to reject the null hypothesis, as the p-value is larger than the cutoff of 0.05. We do not have enough evidence to say that the means of the paired differences are different for 2020 and the 6 seasons prior. In other words, there is not sufficient evidence to say that the balance of power between the two conferences shifted after the 2019 offseason. To generalize, you would likely need more free agents (a higher percentage of players with expiring contracts) to see a significant change in the balance of power across the two conferences. This is under the assumption that the measurement of a team's quality or value is the average of its players WAR calculated by FiveThirtyEight's RAPTOR metric.

References

1. https://github.com/fivethirtyeight/nba-player-advanced-metrics Source for FiveThirtyEight data

2. https://fivethirtyeight.com/features/how-our-raptor-metric-works/   FiveThirtyEight article explaining what RAPTOR is, how it works and what it's designed to do.

3. https://www.landofbasketball.com/ Source for playoff teams and their seeds