Abstract:

We aimed to examine the "Hot Hand Effect" in basketball shooting which is the belief that players who have made consecutive shots are more likely to continue making shots, and vice versa. To explore this concept we have decided to take a computational approach that will allow us to analyze a larger sample size of data. Our analysis will be statistical in nature and outlined thoroughly in the following methods section. The goal of this project is to evaluate if the effect is measurable and if so to what extent does it impact a players shooting performance.

Methods:

First we wrote a function that will generate a non-streaky distribution of made/missed shots to serve as our baseline. This represents a "non-streaky" shooter.

Next we wrote a program that extracts shot data from every shot taken over the course of the entire 2014 NBA season, from this data we determined every single player's shooting streaks, the length of each streak, and the frequency of each streak length.

We also found the average length and standard deviation of all shooting streaks that occurred that season. This allowed us to calculate the Z-score of each player which represents the streakiness of that player's shooting in the 2014 season compared to the average streakiness of all players. Next, we visualized the differences between the players actual shooting streaks that season and the shooting streaks of a simulated shooter with the real players exact same number of shots taken and shooting average but the simulated shooters shots are distributed in a non-streaky manner which allows us to see if there is a noticeable difference in "regular" shooters vs shooters with a supposed "hot hand".

next, we quantified the differences in the real and simulated streaks by using the Komolgorov-Smirnov test to determine the divergence between the data sets, and we also found the P-value from the tests so that we could either rigorously confirm or deny that there was in fact a hot hand effect for that player in the season.

Finally, we visualized the The Empirical Cumulative Distribution Function (ECDF) of each player to see in what ways the data sets differed for each of our shooters to determine if the player was more or less "streaky" than the simulated dataset.

Conclusion:

Out of the 5 streakiest shooting players in the NBA that season the top two (DJ, AH) had distributions that were not different enough to be statistically significant, the third (BW) showed significant divergence from his simulated counterpart, and the last two (RG, AB) showed a complete divergence from their simulated counterparts. In the case of the last 3 (BW, RG, AB) the difference was great enough to reject the null hypothesis (that both sets of shots came from the same type of distribution). Though based on their ECDF's the only one of the three that diverged in a way that suggested a "hot hand" being present was (BW). So, to conclude even in the streakiest shooters in the NBA there is little evidence to support the "hot hand" but it is observable in some rare cases and more research would need to be done that considers more parameters to reach any definitive conclusion.