The "Hot Hand" Phenomenon in the NBA: A Bayesian Hierarchical Analysis

Nick Hellmer
Bayesian Statistics Final Project

April 21, 2025

Abstract

The "hot hand" phenomenon in basketball is the belief that a player is more likely to make a shot after making previous shots. This project aims to determine if this phenomenon truly exists or if it occurs only due to pure chance (e.g., law of large numbers). Shot level data from the 2015-16 NBA season was analyzed and player-specific probabilities of making a shot when "hot" (defined as making the last n shots) versus "not hot" (making less than the last n shots). The models in this project were created using PyMC, and the posterior distributions were assessed to understand the credibility of the hot-hand effect.

1 Introduction

The presumed "hot hand" effect is commonly understood to be a result of a player's increased confidence or momentum following a sequence of made shots. Although previous research has questioned its legitimacy, this analysis leverages Bayesian hierarchical modeling to evaluate whether players truly receive a field goal percentage boost after consecutive successful attempts.

A binary "hot" label was applied to the data for each shot, based on whether a player made their previous n attempts with no more than t seconds in-between shots. Per-player shot success probabilities were compared between hot and not-hot conditions using a beta-binomial model with partial pooling.

2 Data

The data used for this analysis was the 2015–2016 NBA Shot Logs dataset from Kaggle [Shot Logs]. This data set contains detailed shot level information for each game, including the following attributes:

- Player name
- Shot result (made/missed)
- Game context (time, location, distance, etc.)

2.1 Processing Steps: Hot indicator

1. Sorted the data to ensure consecutive shots for each player are ordered correctly.

2. Added a "time remaining" column, which was used to calculate time between shots for a given player. (see function `compute_time_remaining`)

3. A "hot" indicator was defined and assigned to each shot: 1 if the previous n shots were made, 0 otherwise. Each of the n shots also must occur within t seconds of each other. Lastly, shots must occur in the same game and quarter. (see function `label_hot_shots_quarter_aware`)

4. Removed players with fewer than 50 total shots to ensure sufficient sample size (this was accounted for within each model rather than applied to the overall dataset).

2.2 Outlier Detection

Outliers were removed using a combination of domain knowledge and the modified z-score method. When dealing with skewed data, modified Z-score is a more robust way to identify extreme values compared to the standard Z-score. The modified Z-score for each observation x_i is calculated as:

$$M_i = \frac{0.6745 \cdot (x_i - \text{median}(x))}{\text{MAD}}$$

Given a threshold x, observations with $|M_i| > x$ were considered potential outliers. Outliers were removed after hot shot labeling. (see function `filter_modified_zscore`)

2.2.1 Closest Defender Distance

Wide open shots could create bias, as they often differ from regular in-game shooting conditions. To address this:

- Shots with a closest defender distance modified z-score greater than 2 were removed.

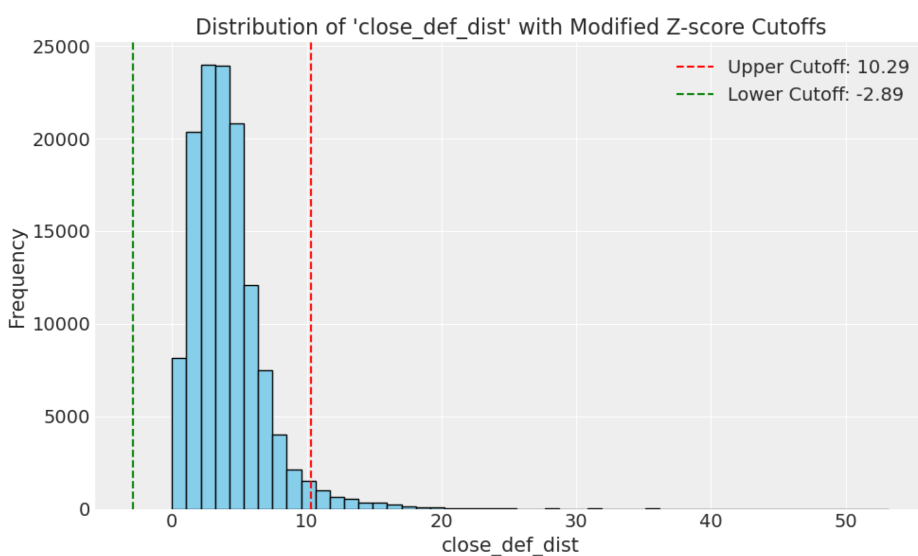

Figure 1: Closest Defender Distance

2.2.2 Shot Distance

For shot distance, a similar logic applies as for closest defender distance. This analysis wants REAL shots, not full court buzzer beaters. Similarly, it does not care for layups or shots from the post. These types of shots usually follow different distributions and models than jump shots as they generally have a much higher field goal percentage.

- Remove all shots less than 15 feet. This threshold was chosen as this is the distance of a free throw. Thus, it is justifiable to claim shots closer than this may not be jump shots (dunks, layups, floaters, hook shots, etc.) and shots further than this are likely jump shots.
- Remove long distance shots with a modified z-score greater than 1.5.

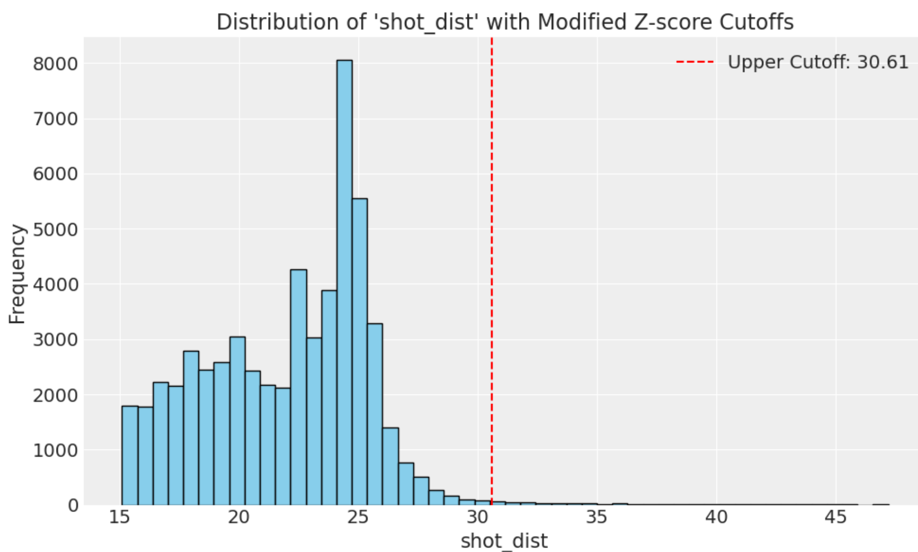

Figure 2: Shot Distance

$z = 1.5$ is just over 30 feet, but this is still well within half court. This is an approriate threshold as players who are labeled ”hot” often attempt a ”heat check”, which is when they shoot from abnormally far due to their increased confidence.

2.2.3 Shot Clock

Forced shots provide bias as they typically have lower field goal percentages. A common example of a forced shot is an attempt occurring in the final seconds of the shot clock. If a team hasn't gotten a good look for the duration of the shot clock, whoever happens to have the ball in the last few seconds just has to give it their best, even if its not good. As for the game clock, shots taken in these final seconds are typically more planned, so these are not being removed. For example, there are usually designed plays leveraged for the final seconds of the game such that a shot occurring in the final seconds is intended rather than forced. *NOTE: shot clock time is blank/NaN when the game clock is less than max shot clock (24 seconds), as the shot clock is turned off when this happens. To ensure these shots are not unintentionally removed, they were excluded from the outlier removal process and then added back at the end.*

- Remove shots on the lower tail with a modified z-score greater than 1.2 (or 1.51 seconds).

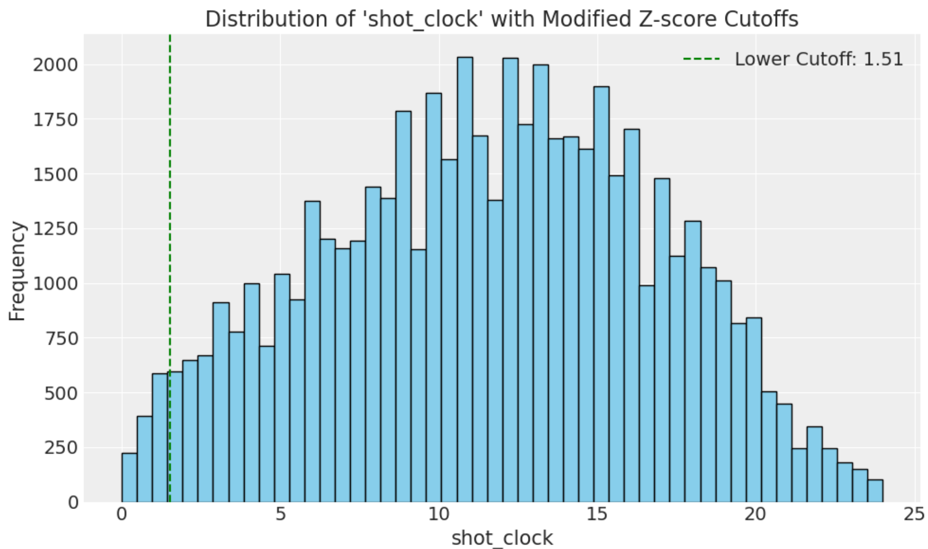

Figure 3: Time Remaining (Shot Clock)

3 Model

3.1 Single-Player Model: Stephen Curry

To build intuition, a simple model was first created for only Stephen Curry. Two independent Beta priors were used for Curry's probability of making a shot when hot and when not hot:

$$\theta_{\text{hot}} \sim \text{Beta}(1, 1)$$
$$\theta_{\text{not}} \sim \text{Beta}(1, 1)$$

The observed data was modeled as:

$$y_{\text{hot}} \sim \text{Binomial}(n_{\text{hot}}, \theta_{\text{hot}})$$
$$y_{\text{not}} \sim \text{Binomial}(n_{\text{not}}, \theta_{\text{not}})$$

3.2 Hierarchical Model Across All Players

Next, a hierarchical model was used where each player had their own probability of making a shot when hot (θ_{h}) and when not hot (θ_{nh}). These player-specific probabilities were drawn from shared priors, enabling the model to account for both individual performance and league variation.

Let $y_{i,h}$ be the number of shots made for player i while labeled hot and $n_{i,h}$ the total number of shots while labeled hot. Similarly, $y_{i,nh}$ and $n_{i,nh}$ referred to the not-hot label.

$$y_{i,h} \sim \text{Binomial}(n_{i,h}, \theta_{i,h}) \quad \text{(Total number of shots made while labeled hot)}$$
$$y_{i,nh} \sim \text{Binomial}(n_{i,nh}, \theta_{i,nh}) \quad \text{(Total number of shots while labeled not-hot)}$$
$$\theta_{i,h} \sim \text{Beta}(\alpha_h, \beta_h) \quad \text{(Player's hot shooting percentage)}$$
$$\theta_{i,nh} \sim \text{Beta}(\alpha_{nh}, \beta_{nh}) \quad \text{(Player's not-hot shooting percentage)}$$

Hyperparameters were assigned the following priors:

$$\alpha_h, \beta_h, \alpha_{nh}, \beta_{nh} \sim \text{HalfNormal}(10)$$

A Binomial likelihood was used to model shot outcomes. This was a clear choice, as each shot represents a single binary trial (make or miss). The Beta distribution was chosen for modeling player shot probabilities because it supports values between 0 and 1 (same domain as shooting percentage) and is the conjugate prior for the Binomial likelihood. Half-Normal priors were used as they are weakly informative and ensure positive values for α and β. Model inference was performed using PyMC.

Two models were created using the above specifications:

- Initial parameters used for models 1 and 2: $n = 3$ and $t = 45$
- Secondary parameters used for model 3 (for sensitivity analysis): $n = 2$ and $t = 60$

4 Results

4.1 Model 1: Stephen Curry

Steph's $P(\theta_{\text{hot}} > \theta_{\text{not}})$ was 0.650 for the single player model (meaning his θ_{hot} was greater than his θ_{not} in 65 percent of samples). The significance of this probability (calculated by comparing posterior draws for each state) becomes clearer when considering Figure 4 below. While his average shooting percentage in hot and not hot states only differed by 4% (0.45 versus 0.41), the model was not designed to analyze the `magnitude` of the hot hand effect, but rather the `likelihood` that the hot hand effect exists, no matter how small or large. In Steph's case, a 65% posterior probability suggests modest support for the presence of a hot-hand effect—even if the size of the difference is small.

NOTE: Since this model was created only for the data associated with Steph Curry, the results differ slightly from his estimates in models 2 and 3. This is due to the differences in filtering for outliers. Although the same filtering bounds were applied, the modified z-scores removed different sets of records when applied globally versus per player. Based on this single player model alone, Steph appears to conform to the presumed hot hand phenomenon.

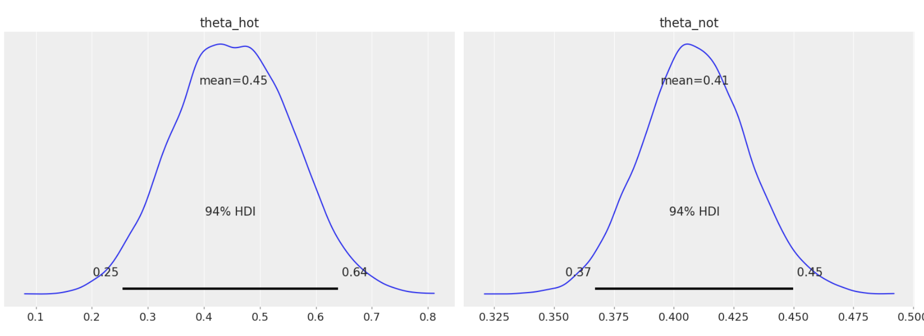

Figure 4: Stephen Curry's Hot and Not-Hot Shooting Percentages

Table 1: Posterior Summary for Steph Curry (Single-Player Model)

Parameter	Mean	SD	3% HDI	97% HDI	ESS (bulk)	$\hat{R}$
θ_{hot}	0.453	0.104	0.255	0.640	20,012	1.000
θ_{not}	0.409	0.022	0.367	0.450	20,204	1.000

4.2 Model 2: Player-Level Comparison

Posterior samples for each player's shooting probabilities were extracted from the model output. The differences between their hot and not-hot state probabilities (i.e., $\theta_{\text{hot},i} - \theta_{\text{not},i}$) were computed for each player to evaluate potential hot-hand effects.

To ensure reliability in the posterior summaries, only players with well-converged parameter estimates (i.e., $\widehat{R} \leq 1.01$ for both θ_{hot} and θ_{not}) were analyzed further. This step helps reduce the risk of interpreting noisy or non-converged posterior estimates.

For each well-converged player, the posterior probability that $\theta_{\text{hot}} > \theta_{\text{not}}$ was estimated by computing the proportion of posterior draws where this condition held true. This provided a Bayesian estimate of the probability that a player experiences the hot hand effect after making successive shots. These values were then mapped to player names and hot shot counts to create a summary table.

The players were ranked by their posterior probability of a hot-hand effect, and the top 10 most likely to "get hot" are shown below.

Table 2: Top 10 Players with Highest Posterior Probability that $\theta_{\text{hot}} > \theta_{\text{not}}$

Player	Hot Shots	$\hat{R}_{\text{hot}}$	$\hat{R}_{\text{not}}$	$P(\theta_{\text{hot}} > \theta_{\text{not}})$
Michael Carter-Williams	4	1.0002	1.0010	0.868
Bojan Bogdanovic	6	1.0001	1.0003	0.864
Trevor Booker	3	1.0001	1.0004	0.861
Donald Sloan	4	1.0002	1.0007	0.804
James Harden	17	0.9999	1.0001	0.804
Giannis Antetokounmpo	2	1.0005	1.0000	0.773
Thaddeus Young	3	1.0000	1.0000	0.746
Rudy Gay	12	1.0002	1.0001	0.742
Mario Chalmers	3	1.0005	1.0006	0.724
Kevin Love	12	1.0000	1.0003	0.718

Some of these top 10 players have few hot shot attempts, meaning they have little data to inform their θ_{hot}. The posterior for these players is more dominated by the prior, so their difference between θs contains more noise. In other words, their estimates are more reflective of prior assumptions rather than actual player performance. To combat this, a threshold was set to remove players with fewer hot shot attempts. The revised top 10 and bottom 10 are shown below.

Table 3: Top 10 Players with ≥ 5 Hot Shots and Highest Posterior Probability that $\theta_{\text{hot}} > \theta_{\text{not}}$

Player	Hot Shots	$\hat{R}_{\text{hot}}$	$\hat{R}_{\text{not}}$	$P(\theta_{\text{hot}} > \theta_{\text{not}})$
Bojan Bogdanovic	6	1.0001	1.0003	0.864
James Harden	17	0.9999	1.0001	0.804
Rudy Gay	12	1.0002	1.0001	0.742
Kevin Love	12	1.0000	1.0003	0.718
Monta Ellis	37	1.0003	1.0006	0.716
Trey Burke	6	1.0010	1.0000	0.690
Nikola Vucevic	13	1.0001	1.0003	0.660
Jordan Hill	10	1.0011	1.0001	0.658
Kobe Bryant	18	1.0005	1.0004	0.657
Andrew Wiggins	8	1.0000	1.0002	0.650

Table 4: Bottom 10 Players with ≥ 5 Hot Shots and Lowest Posterior Probability that $\theta_{\text{hot}} > \theta_{\text{not}}$

Player	Hot Shots	$\hat{R}_{\text{hot}}$	$\hat{R}_{\text{not}}$	$P(\theta_{\text{hot}} > \theta_{\text{not}})$
JJ Redick	14	1.0000	1.0003	0.052
Russell Westbrook	18	1.0002	1.0004	0.063
Chris Paul	22	1.0005	1.0001	0.064
David West	9	1.0003	0.9999	0.068
Dirk Nowitzski	6	1.0001	1.0002	0.097
Courtney Lee	6	1.0001	1.0001	0.101
Brandon Knight	14	1.0002	1.0003	0.107
Marcus Morris	6	1.0002	1.0003	0.150
Joe Johnson	11	1.0002	1.0003	0.153
Darren Collison	5	1.0001	1.0001	0.154

4.3 Model 2: League Wide Inference

Looking across the whole league after filtering for both convergence and a minimum volume of hot shots, 90 players remained in the final analysis. Of those, 34 players ($\approx 37.8\%$) had posterior probabilities $P(\theta_{\text{hot}} > \theta_{\text{not}}) > 0.5$, suggesting a slight lean toward positive hot-hand effects among the filtered subset.

The variance of a Beta distribution is:

$$\text{Var}(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

So the larger the sum of α and β, the smaller the variance. The posterior distributions of α and β from figures 5 and 6 below show that hot-shot probabilities vary more than not-hot probabilities. The hot distribution ($\alpha \approx 10$, $\beta \approx 18$) is broader, while the not-hot distribution ($\alpha \approx 32$, $\beta \approx 54$) is more concentrated, reflecting greater consistency when players are not hot.

4.4 Model 3: Comparing Differences to Model 2

Even when relaxing the hot-streak definition to include more data, the model still finds increased variation (and similar average success) in hot vs. not-hot shooting — suggesting the perceived hot

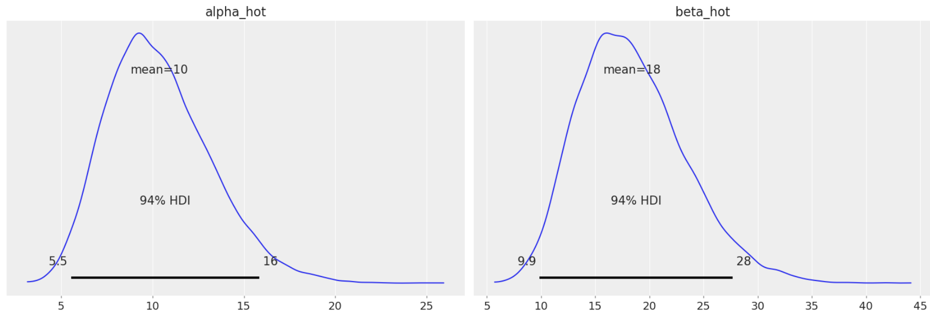

Figure 5: League Wide Hot Hyperparameters

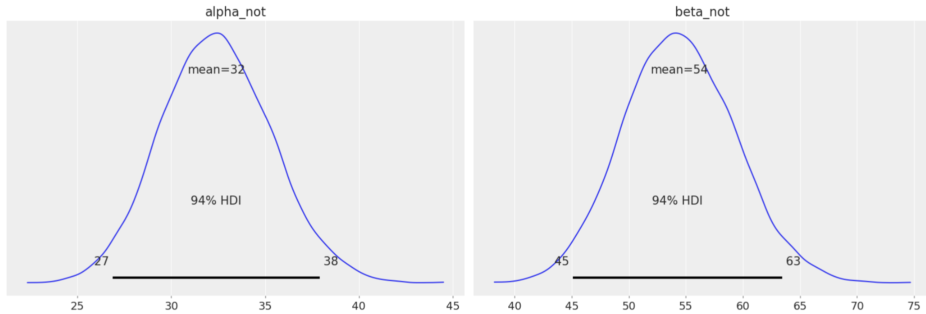

Figure 6: League Wide Not-Hot Hyperparameters

hand effect may be more about variability than a consistent skill boost. See table 5 for model 2 and 3 result comparisons.

Table 5: Comparison of Model 2 and Model 3

n	t	θ_{hot} (mean)	θ_{not} (mean)	θ_{hot} (SD)	θ_{not} (SD)	Total Hot Shots	Total Not-Hot Shots
3	45	0.36	0.37	0.09	0.03	1,212	53,469
2	60	0.37	0.37	0.06	0.03	4,717	49,964

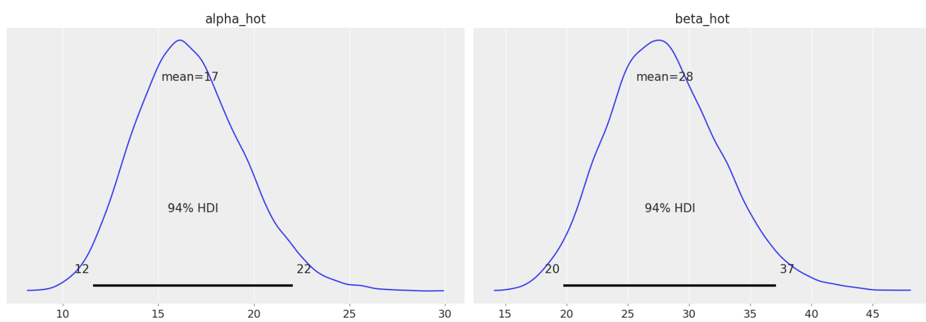

Figure 7: League Wide Hot Hyperparameters

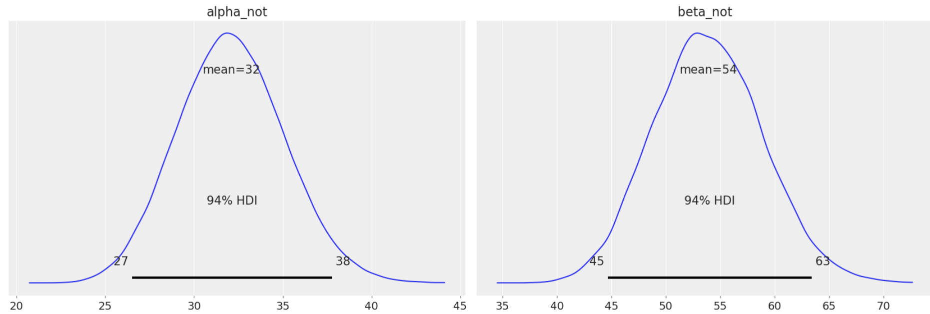

Figure 8: League Wide Not-Hot Hyperparameters

5 Discussion

Key insights:

1. While the hot vs not-hot difference was small or uncertain for most players, the effect was far from uniform.

2. Several players (such as James Harden) showed strong posterior support for a hot-hand effect, while others showed support for the opposite effect (such as Russel Westbrook).

3. These results indicate that the hot hand phenomenon does exist, but it is player-specific rather than a consistent league-wide behavior.

League specific insights:

The means of α and β for the hot condition result in a much wider Beta distribution, whereas the not-hot condition yields a much narrower distribution.

1. When players are hot, there is more variation in shot-making ability.

2. When players are not hot, their shot probabilities are more consistent and predictable.

3. The increased variance in hot-shot probabilities may partly reflect the smaller sample size. A shot is labeled as "hot" only if the previous 3 (or n) shots were made — which is a pretty strict condition. As a result of this strict condition, hot-shot data is extremely limited compared to not-hot shots when $n = 3$ (table 5: 1,212 vs. 53,469), so increased posterior variation in θ_{hot} may reflect uncertainty from low sample sizes rather than meaningful player-level differences.

The mean success rate when hot is nearly identical than when not hot across the league.

1. This suggests there is no strong evidence of a universal league-wide hot-hand effect.

2. However, the increased variance in hot-shot probabilities indicate greater uncertainty and variation among players in the hot state.

3. Combined with the player level results, this supports the idea that while the hot-hand effect may not apply consistently to all players, it does exist meaningfully for some.

Sensitivity Analysis: $n = 2$ **vs.** $n = 3$**, and** $t = 60$ **vs.** $t = 60$

Even when relaxing the hot-streak definition to include more data by changing the conditions from $n = 3$ and $t = 45$ to $n = 2$ and $t = 60$, the model still finds increased variation (and similar average success) in hot vs. not-hot shooting — suggesting the perceived hot hand may be more about variability than a consistent skill boost.

6 Conclusion

The Bayesian hierarchical models applied to NBA shot data provided nuanced results of league wide shot success probabilities under hot and not-hot conditions.

While league-wide averages showed little difference between the two states, the models uncovered meaningful variation at the player level. These player-specific patterns were only visible through a framework that combines individual-level modeling with league-wide priors. Ultimately, the analysis supports the conclusion that the hot-hand phenomenon does exist—but only for some players rather than across the board.

7 Future Enhancements and Next Steps

1. Build a Bayesian logistic regression to model shot success as a function of the number of consecutive prior makes.
2. Explore threshold effects (e.g., is there a distinct boost after 4+ or 5+ consecutive makes?) A larger starting dataset would be required for this, as the number of hot shots would decrease due to the stricter labeling conditions.
3. Consider other player-level attributes not included in this dataset (e.g., position, average minutes played) to see if certain types of players are more prone to hot-hand effects.
4. Extend the model to assess the hot hand effect while considering additional game context, such as game quarter (e.g., does the hot hand effect differ between early and late game periods?) or home vs. away (e.g., does the hot hand effect become more likely for home team players than away team players?).

References

[Shot Logs] NBA Shot Logs Dataset, Kaggle.
`https://www.kaggle.com/code/benhamner/quick-exploration-of-the-nba-shot-logs-data/input`