Advanced Topics in Data Science Project - Group #20

Date: 12/11/25

GitHub: https://github.com/DS301-Project-Group20/Advanced-DS-Project-Hot-Hand

The Hot Hand Phenomenon

Leon Wang | Peter Hu | Yuan Li

Executive summary

1. What is Hot hand: a player has higher than average chance to make the next shot after making a streak of shots. Important in sports and psychology.
   Is hot hand a real phenomenon or a fallacy?
2. Statistical tests – it is about if there are correlations between the outcomes of shots
   controlled experiments (free throw, 3 point contest)
   no hot hand = each shot is i.i.d Bernoulli RV;
   real sports games
   no hot hand = each shot is independent Bernoulli RV (not identical dist.).
3. Traditional Tests for Hot Hand Hypothesis are based on Conditional Probability Test
   P[Make | *k* Makes]–P[Make | *k* Misses] = 0, then no hot hand.
   Conditional probability tests suffers problem of **streak selection bias**
4. Real game situation requires prediction model for shot-make probability.
   Advanced data science models + Large data ⇒ shot-make prediction model.
5. We develop a new shot-make prediction model + a new simplified statistical test for HH
6. ChatGPT correctly interprets statistical meaning of hot hand and can perform some statistical analysis and identify players with hot hand.

Motivation / Related works

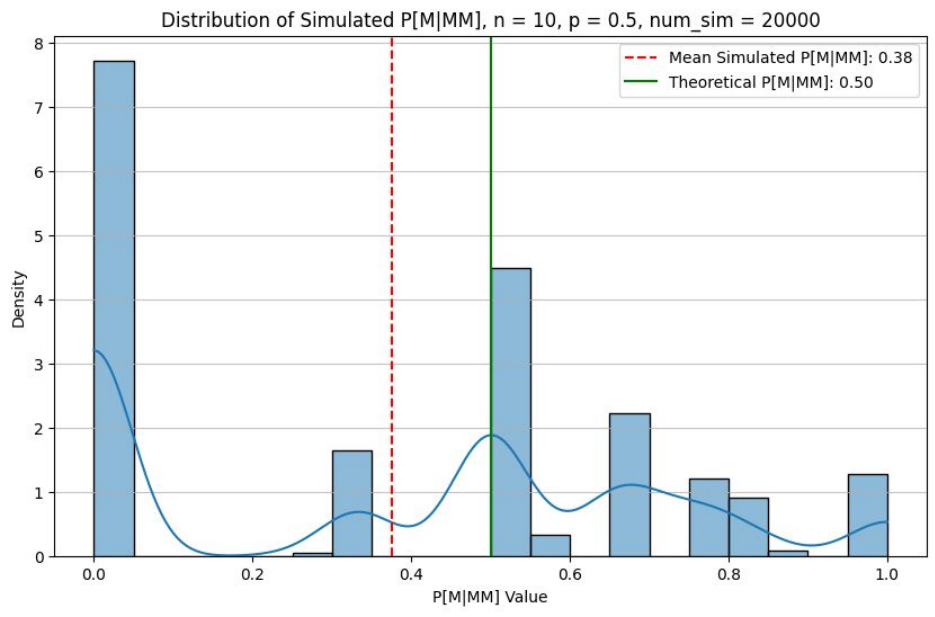
1. GVT (Gilovich, T., Vallone R., and A. Tversky) “The hot hand in basketball: On the misperception of random sequences.” (1985)
   Influential paper. Use controlled experiments and conditional probability test. Find no evidence of hot hand in basketball shooting data.

2. MS (J. B. Miller and A. Sanjurjo.) “Surprised by the hot hand fallacy? A truth in the law of small numbers.” (2018)
   Point out streak selection bias. Use permutation tests to compensate the bias. Reverse GVT’s conclusion and support existence of hot hand.

3. PW (K. Pelechrinis and W. Winston.) “The hot hand in the wild” (2022)
   Use real NBA game data. Develops deep learning shot-make prediction model. Apply conditional probability test with bias adjustment, find evidence of hot hand.

Streak Selection Bias

Where does streak selection bias come from? Because the sequence of shots is finite.
Example: A shooting record with 10 shots [XMMMXXMMXX]. (M = Make, X = Miss)
As a convention, when counting sub-sequence, overlapping is allowed.
Observed P[M|MM] = #(MMM)/{#(MMM) + #(MMX)} = 1/(1+2)=0.333
For a short sequence and P[M] is not too much above 0.5, #(MMM)<#(MMX) on average
So even when the sequence is i.i.d., Observed P[M|MM] < P[M]. Here , 0.38 vs 0.5



Implication of Streak Selection Bias
Even when the shots are iid,
Observed P[M|MM] < P[M].

So when Observed P[M|MM] ≈ P[M], there is positive correlation among shots.

Method/Approach

Our Works:

1. We develop a new shot-make prediction model using LightGBM classifier and it performs better than PW's deep learning model.

2. We develop a new statistical test for hot hand hypothesis based on streak number. It is more intuitive and does not suffer the streak selection bias.

3. We test ChatGPT's responses to hot hand related prompts and compare them with our streak test results and PW's test results.

Implementation/Experimentation

1. We train and test LightGBM prediction model on the same two-season (2013-14, 2014-15) data as used in PW.

2. We apply streak number test on the same data used by PW. We rank the players based on our hot hand index and compare the ranks with PW's hot hand rank.

3. We ran prompts on ChatGPT 5.1 Think version to test its response to hot hand questions. We upload several pairs of players shooting records into ChatGPT and let it compare who has a hot hand.

NBA Facts:
# of teams: 30
# of players: ~ 500
Season: Oct - Apr
# of games (regular): 1230
Total shot attempts per season: ~200,000
Avg starter shot attempts per game: 10~18
Avg shot-make (field goal) percentage: ~47%

New Prediction Model

Results: Shot-Make Prediction Model

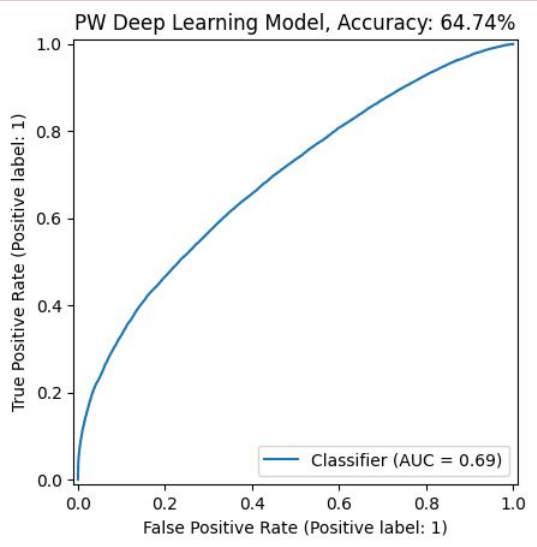
Shot-Make Prediction Models:

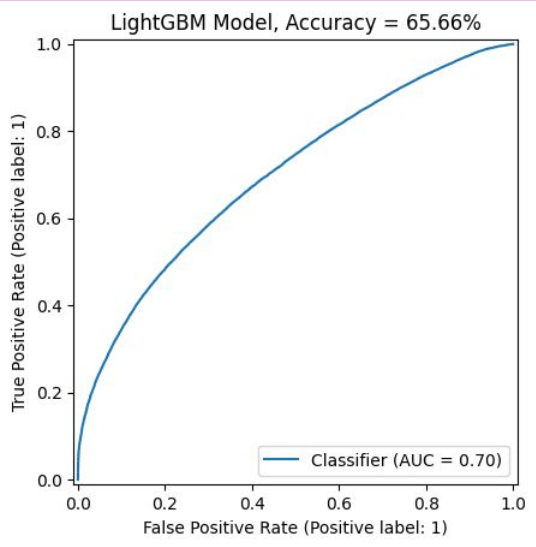
Supervised Training Data: 379873 by 1212

PW Model: Sequential Neural Network,
4 Hidden Layers, 250 Nodes each

LightGBM:
Max_depth: 40
Learning Rate: 0.025
n_estimators: 200

Model: To predict probabilities of a shot is a success. Explanatory variables include: Distance to the basket, Distance of the closest defender, Number of dribbles before shooting, ...





New Hot Hand Test

Results: Streak Number Test

Conditional Probability Tests:
GVT and MS
$P[M|k\ Ms]-P[M|k\ Xs]=0$

PW
$P[M|k\ Ms]_{obs} - P[M|k\ Ms]_{sim}=0$

Issues:
Need to deal with Streak Selection Bias

More noises with higher k

Streak Number Test: Distribution of number of k-streaks for independent Bernoulli RVs can be estimated by simulation. We can calculate the empirical p-value of each player's number of k-streaks.

Player	No. games	Obs 4-streaks	Exp 4-streaks	p-value
Jordan Hill	106	61	20.6	0.12%
Ersan Ilyasova	88	39	13.4	0.47%
Kemba Walker	59	32	11.1	0.79%
Tristan Thompson	140	53	22.5	0.83%
Jose Juan Barea	124	20	6.5	1.06%
Ben McLemore	124	31	10.3	1.14%

No hot hand hypothesis can be rejected if the observed number of k-streaks is significantly higher than expected (p-value < α)
Our tests strongly suggest that hot hand exists in NBA games.

ChatGPT 5.1 Think version

Results: How ChatGPT View Hot Hand

ChatGPT

(1) can correctly identify the statistical meaning of hot hand.

(2) can perform simple statistical estimation on real data

Prompt: An NBA player can have a hot hand during a game. What does having a hot hand mean?

ChatGPT: …a player has a hot hand if the probability of making the next shot is higher after recent makes than after recent misses (or higher than their usual average)

Prompt: Based on the following records, which player do you think has a hot hand?

```
Player 1: Marreese Speights
GAME_ID,MATCHUP,SHOT_NUMBER
21300368 , DEC 17 2013 - GSW vs. NOP , 10011001000
21301123 , APR 02 2014 - GSW @ SAS , 1011001110101011
…

Player 2: Kyle Korver

GAME_ID,MATCHUP,SHOT_NUMBER

21301004 , MAR 18 2014 - ATL vs. TOR , 00001111001000

21301120 , APR 02 2014 - ATL vs. CHI , 110100001
```

ChatGPT: …Speights shows a small “hot-hand-like” effect (better after makes than after misses), Korver does not: his make probability barely changes based on what happened on the previous shot.

Hot Hand Ranks

1. PW's hot hand index: based on how much the observed conditional probability exceeds the simulated conditional probability which assumes independent shots.
2. Our hot hand index: based on p-value of players' number of k-streaks.
3. Hot hand player rankings: based on average indices (weighted on # samples)
4. Two ranks are different but mostly consistent, Spearman's rank correlation 0.81

Player	PW Rank	Our Rank
Kemba Walker	1	19
Jordan Hill	2	2
Ben McLemore	3	1
Boris Diaw	4	10
CJ Miles	5	9
Monta Ellis	6	21
Tristan Thompson	7	4
Chris Bosh	8	20
Kyle Korver	9	53
Rudy Gay	10	16

Player	Our Rank	PW Rank
Ben McLemore	1	3
Jordan Hill	2	2
Marreese Speights	3	59
Tristan Thompson	4	7
Jeremy Lin	5	38
Aaron Brooks	6	36
Markieff Morris	7	15
DeMarre Carroll	8	22
CJ Miles	9	5
Boris Diaw	10	4

Who has hot hand?

We pick several pairs of players, upload their shooting records and ask ChatGPT to evaluate who has hot hand.

Ben McLemore (our rank = 1, PW rank = 3) vs Kemba Walker (our rank = 19, PW rank = 1)
Our #1 vs PW #1. ChatGPT: Ben McLemore has hot hand

Marreese Speights (our rank = 3, PW rank = 59) vs Kyle Korver (our rank = 53, PW rank = 9)
Two inconsistently ranked players. ChatGPT: Marreese Speights has hot hand

CJ Miles (our rank = 9, PW rank = 5) vs Ersan Ilyasova (our rank = 24, PW rank = 26)
Two consistently ranked players. ChatGPT: CJ Miles has hot hand

Jeremy Lin (our rank = 5, PW rank = 38) vs Chris Bosh (our rank = 20, PW rank = 8)
Two moderately inconsistently ranked players. ChatGPT: Jeremy Lin has hot hand

In general, ChatGPT is more aligned to our rank. Higher in our rank, hotter hand by Chat GPT

Conclusion

1. We develop a shot-make prediction model that achieves a better accuracy than PW's deep learning model

2. We develop a new statistical test - streak number test - and find strong evidence that hot hand exists in NBA games.

3. We test ChatGPT's responses to hot hand questions and find it can correctly identify the statistical definition of hot hand and perform some statistical estimation on real data.