Sequential Dependency in Sports Games

David Oluigbo

Introduction

Probability models, such as flipping a fair coin, assume independent event outcomes. However, real-life events are often interconnected rather than independent. Many involve sequential dependencies, requiring humans to recognize and model these patterns to predict outcomes accurately. This phenomenon is especially evident in sports. Momentum, which is the belief that prior successes or failures influence future performances, significantly shapes perceptions among athletes, fans, and bettors. For example, the "hot hand" and "streaky shooter" are popular applications of momentum in basketball, where a player is deemed to have a significantly better chance to make a three-pointer or free throw than normally expected based on their season and career averages (Gilovich, Vallone, & Tversky, 1985).

Related Works

Winning streaks are often perceived as a manifestation of momentum at the team level. These "hot" streaks, defined as three or more consecutive victories uninterrupted by losses or ties, can strongly influence spectators' perceptions, leading them to favor one team over another in sports contests across various leagues. This belief in momentum transcends anecdotal evidence; Vergin found that athletes and fans in Major League Baseball (MLB) and the National Basketball Association (NBA) frequently attributed game outcomes to the presence of momentum (Vergin, 2000). Similarly, in the National Football League (NFL), research revealed that betting volumes increased significantly for teams on winning streaks, despite this approach not reliably resulting in excess returns (Paul, Weinbach, & Humphreys, 2011).

These findings highlight a broader human tendency to perceive patterns and dependencies in sequences of events that are often closer to random than they appear. In sports, momentum is frequently seen as a causal factor, influencing not only the players and teams involved but also the decisions and predictions of fans and bettors. Building on this foundation, my aim is to investigate how individuals' perceptions of momentum, specifically in the form of winning streaks, affect how they predict and model future game outcomes.

Methods

The goal of this study was to compare how humans and probabilistic models predict future game outcomes. By examining participants' probabilistic reasoning and contrasting it with the outputs of mathematical models, this study aimed to identify the ways in which humans process sequential dependencies and integrate contextual information, such as streak lengths and scoring trends, into their predictions.

I recruited a total of 10 human participants for my study, all of whom were college-age students (19–22 years old) actively enrolled at universities in the Boston and Cambridge area.

Participants completed a questionnaire consisting of 47 hypothetical scenarios about sports contests, divided into two distinct sets. The first set consisted of 23 scenarios, and the second set consisted of 24 scenarios. **Participants were informed that, in the absence of other information, the basketball team normally wins 50% of its games.** This base win probability reflects a neutral assumption about team performance, aligning with the focus of this study on understanding how additional contextual factors—such as streak lengths and margins of victory—affect individuals' probability estimates.

First Set of Scenarios. The first set of scenarios was designed to explore participants' predictions based on binary game outcomes (wins or losses) over the team's last 8 games and the length of the team's current winning streak. There were 23 scenarios in this set. For each scenario, only the games prior to the winning streak (excluding the loss immediately preceding the streak) were modeled as Bernoulli trials, where the probability of success (a win) was set at 0.5 to reflect the team's base win probability.

For example, for a team with a 4-game winning streak, the last 5 games were fixed to meet the scenario requirements (4 consecutive wins followed by one preceding loss), while the outcomes of the first three games in the sequence were sampled as Bernoulli trials. This method ensured variability in the sequences presented to participants while adhering to the constraints imposed by the streak length.

The number of unique sequences varied by streak length. For streaks of 1 to 5 games, three different sequences of outcomes were generated for each streak length. For streaks of 6 games, only two unique sequences were possible, as all but one game in the sequence were constrained. For streaks of 7 and 8 games, the outcomes of all 8 games were fixed to meet the scenario requirements, resulting in a single sequence for each streak length.

Participants were asked to estimate the probability that the team would win its next game based on these outcomes. Responses were provided as a probability value between 0 and 100, rounded to the nearest integer. Standardized wording was used across all scenarios, with only the outcomes of the last 8 games and the length of the winning streak varying.

An example scenario from the first set was:

> ”A basketball team is currently on a four-game winning streak. Normally, this team wins 50% of its games. Over the past 8 games, the team has the following results:
>
> W, W, L, L, W, W, W, W (W = win, L = loss).
>
> Based on this information, what is the probability that the team will win its next game?”

Second Set of Scenarios. The second set of scenarios consisted of 24 scenarios. For each of the eight winning streak conditions, one of the game outcome sequences from the first set of scenarios was selected. Then, three new scenarios were generated by adding additional information about the margin of victory or loss for each of the team's last 8 games. For each streak condition, the three scenarios were designed to reflect different trends in the score margins: one with a decreasing trend in the margin of victory, another with an increasing trend in the margin of victory, and one with a more general pattern. This design was intended to test how the margin of victory might affect participants' perceptions of winning streaks as a causal factor in the team's chances of winning the next game.

Instead of presenting exact final scores, participants were shown only the computed margins of victory (positive values) or loss (negative values) for each game. This simplified the presentation while emphasizing trends in performance over absolute scores.

As in the first set, participants were asked to estimate the probability that the team would win its next game, providing their response as a probability value between 0 and 100 to the nearest integer. The wording of the scenarios remained standardized, with only the outcomes, streak lengths, and margin trends varying.

An example scenario from the second set was:

> ”A basketball team is currently on a four-game winning streak. Normally, this team wins 50% of its games. Over the past 8 games, the team has the following results:
>
> W (+8), W (+6), L (-5), L (-4), W (+15), W (+8), W (+10), W (+2) (W = win, L = loss; the numbers in parentheses represent the margin of victory or loss).
>
> Based on this information, what is the probability that the team will win its next game?”

Probabilistic Models for First Set of Scenarios

For the first set of 10 scenarios, I utilized three distinct probabilistic approaches. These models spanned varying levels of complexity, serving as both baseline comparisons and more nuanced methods to capture potential sequential dependencies in the data.

1. Windowed Average Model. The windowed average model provided a simplistic baseline. This model took two inputs:

- (1) A list of binary game outcomes representing wins and losses (x_i, where $x_i = 1$ for a win and $x_i = 0$ for a loss), and
- (2) A window size, denoted as w, specifying the number of recent events to include in the calculation.

The predicted probability of a win, P_{win}, was computed as the average of the outcomes within the window:

$$P_{\text{win}} = \frac{\sum_{i=1}^{w} x_i}{w}.$$

This model assigned equal weight to all events within the window, meaning that older and more recent events contributed equally to the final probability prediction.

2. Exponentially Weighted Average Model. The exponentially weighted average model introduced a more sophisticated method by weighting recent events more heavily than older ones. This model took three inputs:

- (1) A list of binary game outcomes (x_i),
- (2) A window size, denoted as w, specifying the number of recent events to consider, and
- (3) An exponential decay factor, α, ranging between 0 and 1, which determined the rate at which the weights decreased exponentially with the age of the event.

The predicted probability of a win, P_{win}, was computed as:

$$P_{\text{win}} = \frac{\sum_{i=1}^{w} \alpha^{i-1} x_i}{\sum_{i=1}^{w} \alpha^{i-1}} = \sum_{i=1}^{w} \alpha^{i-1} x_i,$$

where x_i represents the outcome of the i-th game (1 for a win, 0 for a loss), and α^{i-1} is the weight assigned to the i-th game, decreasing exponentially with increasing i. The weights were normalized to a scale of 0 to 1, meaning that the sum of the weights in the denominator of the original formula was equal to 1. As a result, the formula simplified to the numerator.

3. Bayesian Inference Model. The Bayesian inference model was implemented using the PyMC library to estimate the probability of a win. This model treated the binary game outcomes as observed Bernoulli trials and used a Beta distribution as a prior for the win probability. The model consisted of three steps:

1. **Prior Distribution:** The win probability, P_{win}, was modeled as a random variable drawn from a Beta distribution:

$$P_{\text{win}} \sim \text{Beta}(\alpha, \beta),$$

where the parameters α and β represented prior beliefs about the proportion of wins to losses. In this study, α and β were both set to 3 to center the prior distribution around 0.5, reflecting the assumption that the team's base win probability was 50%. This choice corresponded to what participants were told about the team's neutral performance baseline.

2. **Likelihood Function:** The observed game outcomes (x_i, where $x_i = 1$ for a win and $x_i = 0$ for a loss) were modeled as samples from a Bernoulli distribution:

$$x_i \sim \text{Bernoulli}(P_{\text{win}}).$$

This likelihood function quantified how well the observed outcomes aligned with the win probability. The observed outcomes that were used were a sequence of the team's 8 most recent game outcome, with specific constraints imposed by the winning streak condition described in the scenario.

3. **Posterior Inference:** Using PyMC, posterior sampling was performed via No U-Turn Sampler (NUTS), a Hamiltonian Monte Carlo Method, to approximate the posterior distribution of P_{win}:

$$P(P_{\text{win}} \mid \text{data}) \propto P(\text{data} \mid P_{\text{win}}) \cdot P(P_{\text{win}}).$$

The posterior combined the prior beliefs with the observed outcomes to refine the estimate of P_{win}. The posterior mean was used as the predicted probability for the next game.

These three models provided a diverse framework for analyzing sequential dependencies in game outcomes. The windowed average model served as a straightforward baseline, while the exponentially weighted average model and Bayesian inference model captured more complex temporal and probabilistic dynamics.

Probabilistic Models for Second Set of Scenarios

For the second set of 10 scenarios, which included information about the margin of victory or loss for each game, I adapted the three probabilistic models to incorporate this additional variable. The updated models were designed to capture both binary win/loss outcomes and the influence of performance margins in predicting the probability of a win.

1. Windowed Average Model with Margin Weighting. The windowed average model was updated to account for the margin of victory or loss by using the absolute values of the margins as weights for the binary outcomes. This ensured that games with larger margins contributed more to the final probability estimate. The predicted probability of a win, P_{win}, was computed as follows:

$$P_{\text{win}} = \frac{\sum_{i=1}^{w} |m_i| \cdot x_i}{\sum_{i=1}^{w} |m_i|} = \sum_{i=1}^{w} |m_i| \cdot x_i,$$

where:

- x_i represents the binary outcome of the i-th game (1 = win, 0 = loss),
- m_i is the margin of victory or loss for the i-th game, and
- w is the window size, specifying the number of recent games to include.

The absolute margins ($|m_i|$) were normalized to a scale of 0 to 1, ensuring that the sum of the weights in the denominator of the original formulation, $\sum_{i=1}^{w} |m_i|$, was equal to 1. As a result, the fraction in the formulation simplified to just the numerator, with the predicted probability determined directly by the weighted sum of the binary outcomes.

2. Exponentially Weighted Average Model with Sigmoid-Transformed Margins. The exponentially weighted average model was updated to combine the effects of recency and game performance margins. In this version, the binary outcomes were weighted by both the exponential decay factor (α) and a sigmoid-transformed margin of victory or loss. The sigmoid function mapped each margin to a weight between 0 and 1:

$$w_{\Delta p_i} = \frac{1}{1 + e^{-\Delta p_i}}$$

The combined weight for each game was then given by:

$$w_i = \alpha^{i-1} \cdot w_{\Delta p_i},$$

where α is the exponential decay factor and Δp_i is the margin of victory or loss for the i-th game. The predicted probability of a win, P_{win}, was computed as a weighted average:

$$P_{\text{win}} = \frac{\sum_{i=1}^{w} w_i \cdot x_i}{\sum_{i=1}^{w} w_i} = \sum_{i=1}^{w} w_i \cdot x_i$$

This formulation allowed recent games with larger victory margins to have a greater influence on the final prediction, while still accounting for recency through the exponential decay factor. The weights, $w_i = \alpha^{i-1} \cdot |m_i|$, were normalized to a scale of 0 to 1, ensuring that the sum of the weights in the denominator, $\sum_{i=1}^{w} w_i$, was equal to 1. This normalization simplified the formulation, as the denominator effectively became redundant, leaving the final prediction determined solely by the weighted sum of the outcomes.

The third approach was a Bayesian inference model, implemented using the PyMC library, to model the distribution of observed game margins and their influence on win probabilities. The model structure and parameter settings were as follows:

1. **Prior Distributions:** The mean and standard deviation of the margins (m_i) were modeled using the following priors:

$$\mu_m \sim \text{Normal}(0, 10), \quad \sigma_m \sim \text{HalfNormal}(10),$$

   where μ_m represents the mean margin and σ_m represents the standard deviation of the margins. These priors reflect an initial belief that the margins are symmetrically distributed around 0, with a wide spread to capture potential variability. The choice of $\sigma = 10$ ensures coverage of a broad range of margin values, consistent with the variability observed in real-world sports outcomes.

2. **Likelihood Function:** The observed margins, m_i, were modeled as samples from a Gaussian distribution:

$$m_i \sim \text{Normal}(\mu_m, \sigma_m),$$

   where μ_m and σ_m are latent variables inferred from the data. This formulation allowed the model to capture the natural variability in margins across the team's games, as presented to the participants.

3. **Posterior Sampling:** Posterior distributions for μ_m and σ_m were estimated using No-U-Turn Sampling (NUTS), a variant of Hamiltonian Monte Carlo (HMC). NUTS was selected for its efficiency in sampling from complex posterior distributions. These settings balanced computational efficiency with the need for robust posterior estimates.

4. **Win Probability Estimation:** To estimate the probability that the team would win its next game, we summed the posterior probability mass corresponding to positive values of μ_m. This aligned with the experimental design, as the margin of victory was hypothesized to influence participants' perceptions of momentum and, consequently, the likelihood of winning the next game.

This Bayesian framework leveraged principled probabilistic inference to integrate prior beliefs, observed data, and model assumptions. The parameter settings, including the priors and sampling method, were chosen to align with the study's neutral assumptions and experimental goals while capturing the variability inherent in sports outcomes.

Results

Participant Responses Across Scenarios

Participants' estimates for the probability of the team winning its next game varied systematically across scenarios. For the first set of scenarios (without margin information), predictions showed a general increase with streak length, consistent with the belief that longer winning streaks improve a team's chances. For example:

- For a 2-game winning streak, the mean estimated probability was 60% (SD = 6%).
- For a 4-game winning streak, the mean estimated probability increased to 68% (SD = 6%).
- For an 8-game winning streak, the mean estimated probability reached 82% (SD = 8%).

In the second set of scenarios, which included margin information, participants demonstrated greater sensitivity to performance trends:

- For a 5-game winning streak with increasing margins, the mean estimated probability was 85% (SD = 7%).
- For a 5-game winning streak with decreasing margins, the mean estimated probability was 72% (SD = 6%).

These results suggest that participants interpreted margins of victory as indicative of a team's performance trajectory, assigning higher probabilities of winning when recent victories were by increasingly larger margins.

Model Outputs Across Scenarios

The probabilistic models produced systematic predictions that reflected the constraints of their respective formulations. Key trends in the model outputs are summarized below.

Scenarios Without Margin Information For the first set of scenarios, the models generally aligned with the trend observed in participant responses, showing increased win probabilities with longer streaks. The windowed average model and Bayesian model exhibited gradual increases, while the exponentially weighted average models with higher α values (e.g., $\alpha = 0.9$) were more sensitive to recent outcomes, producing sharper changes in win probabilities. For instance:

- For a 2-game winning streak, the Bayesian model predicted a win probability of 57.1% (SD = 6.0%).
- For an 8-game winning streak, the Bayesian model predicted a win probability of 78.9% (SD = 0.0%).

The windowed average model showed steadier increases, while the exponentially weighted models emphasized the most recent outcomes, resulting in higher variability for shorter streaks. Full results are presented in Table 1.

Scenarios With Margin Information In the second set of scenarios, the models leveraged margin data to make more nuanced predictions. For example:

- For a 5-game winning streak with increasing margins, the Bayesian model predicted a win probability of 78.7% (SD = 0.0%).
- For a 5-game winning streak with decreasing margins, the Bayesian model predicted a win probability of 70.2% (SD = 0.0%).

These results demonstrate the models' sensitivity to performance trends, aligning with the participant responses in the second set. Full results are presented in Table 2.

Win Streak	Mean	SD
1	0.2500	0.1768
2	0.5000	0.1021
3	0.4583	0.0589
4	0.5833	0.1179
5	0.5833	0.0589
6	0.6875	0.0625
7	0.7500	0.0
8	0.8750	0.0

Table 1: Model 1 performance across win streak lengths without margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.7256	0.0265
2	0.9261	0.0102
3	0.9772	0.0026
4	0.9930	0.0010
5	0.9979	0.0002
6	0.9994	0.0001
7	0.9998	0.0
8	1.0	0.0

Table 2: Model 2 ($\alpha = 0.3$) performance across win streak lengths without margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.5115	0.0389
2	0.7684	0.0528
3	0.8490	0.0233
4	0.9161	0.0223
5	0.9481	0.0078
6	0.9753	0.0057
7	0.9886	0.0
8	1.0	0.0

Table 3: Model 3 ($\alpha = 0.6$) performance across win streak lengths without margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.3934	0.1266
2	0.6447	0.0921
3	0.6428	0.0522
4	0.7566	0.0907
5	0.7781	0.0420
6	0.8647	0.0420
7	0.9160	0.0
8	1.0	0.0

Table 4: Model 4 ($\alpha = 0.9$) performance across win streak lengths without margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.4288	0.1006
2	0.5710	0.0599
3	0.5503	0.0350
4	0.6206	0.0682
5	0.6188	0.0334
6	0.6789	0.0356
7	0.7154	0.0
8	0.7891	0.0

Table 5: Model 5 performance across win streak lengths without margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.7769	0.0169
2	0.8909	0.0195
3	0.7953	0.0208
4	0.9343	0.0264
5	0.8875	0.0445
6	0.9542	0.0163
7	0.9438	0.0158
8	0.8695	0.0752

Table 6: Model 1 performance across win streak lengths with margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.9895	0.0046
2	0.9975	0.0016
3	0.9987	0.0008
4	0.9997	0.0003
5	0.9999	0.0001
6	0.99998	0.00003
7	0.999997	0.0000
8	1.0	0.0

Table 7: Model 2 ($\alpha = 0.3$) performance across win streak lengths with margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.9666	0.0116
2	0.9863	0.0086
3	0.9920	0.0040
4	0.9974	0.0029
5	0.9968	0.0020
6	0.9991	0.0010
7	0.9998	0.0002
8	0.9887	0.0001

Table 8: Model 3 ($\alpha = 0.6$) performance across win streak lengths with margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.9671	0.0112
2	0.9738	0.0161
3	0.9782	0.0076
4	0.9940	0.0067
5	0.9818	0.0125
6	0.9951	0.0051
7	0.9983	0.0019
8	0.9163	0.0008

Table 9: Model 4 ($\alpha = 0.9$) performance across win streak lengths with margin information (mean and standard deviation).

Win Streak	Mean	SD
1	0.7865	0.0
2	0.7865	0.0
3	0.7865	0.0
4	0.7865	0.0
5	0.7865	0.0
6	0.7865	0.0
7	0.7865	0.0
8	0.7865	0.0

Table 10: Model 5 performance across win streak lengths with margin information (mean and standard deviation).

Discussion

The findings highlight the different ways participants and models integrate sequential information and scoring data when predicting game outcomes. Participants' probability estimates in the first set of scenarios, which lacked margin information, demonstrated a trend: as the length of the winning streak increased, so did their predicted probability of a win. This pattern aligns with heuristic-based reasoning, resembling simpler models such as the windowed average, where outcomes are averaged without incorporating additional contextual data. However, the relatively smaller increases in probability estimates across streak lengths suggest that participants may not fully internalize the implications of longer streaks, as probabilistic models generally predicted sharper increases in win probabilities for the same scenarios.

In the second set of scenarios, the inclusion of margin information provided a clearer picture of how participants integrated sequential and contextual data. Participants displayed a greater sensitivity to trends in the margins of victory, assigning higher probabilities of winning to streaks with increasing margins compared to those with decreasing or inconsistent margins. This nuanced response suggests that participants may apply a mental model akin to logistic regression, using performance trends to adjust their predictions.

The probabilistic models provided a useful benchmark for understanding human predictions. Simpler models, such as the windowed average, captured basic trends but lacked the sophistication needed to incorporate margin-based adjustments. In contrast, the exponentially weighted average and Bayesian models offered more refined predictions, with the latter successfully integrating prior beliefs and observed margins. The divergence between participant predictions and model outputs in certain scenarios underscores the limitations of human reasoning, particularly in situations where statistical priors or recency effects play a significant role.

The observed alignment between participants' estimates and model predictions in the second set of scenarios points to the potential for training or interventions to improve human probabilistic reasoning. For example, exposing individuals to probabilistic models that incorporate contextual factors, such as margins of victory, could help bridge the gap between intuitive reasoning and more systematic, data-driven approaches. This alignment also suggests that human predictions are not entirely heuristic-driven but can be adaptive and context-sensitive when provided with richer information.

The results of this study provide a foundation for exploring additional aspects of sequential dependencies in sports contests. The distinct ways in which participants modeled winning streaks, particularly in the presence of margin information, raise important questions about how similar factors might influence perceptions of losing streaks or sequences involving ties. By expanding the scope of research to include these scenarios, future studies can further illuminate the psychological mechanisms underlying momentum perception and sequential reasoning.

Future Directions

Having investigated how individuals model winning streaks and compare their predictions with probabilistic models, an important next step is to explore how people perceive and model losing streaks in sports contests. While much of the focus in momentum research has been on the belief in "hot hands" and winning streaks, losing streaks—where a team is believed to have a diminished chance of winning—are just as pervasive in sports culture. Understanding how individuals adjust their predictions in the face of consecutive losses could shed light on the psychological aspects of perceived momentum and how it affects decision-making in betting, fan predictions, and even player performance. By applying similar probabilistic models to those used for winning streaks, we could test whether people perceive the likelihood of a team's next loss as increasing during a losing streak, or if they see losses as independent events with no impact on future outcomes. Comparing human models with probabilistic models, like those based on a uniform or Beta distribution, would allow for a deeper understanding of how sequential dependencies are conceptualized when outcomes are negative.

Another area for future exploration is the impact of ties, which are common in sports such as soccer, on how people perceive sequential dependencies. In sports where ties can occur, such as soccer, a game outcome is not strictly a win or a loss but a neutral outcome. This could introduce a new dynamic in how people model and perceive sequences of events, as ties may disrupt the perception of momentum. For instance, a team's performance following a tie might be viewed differently compared to a win or loss, and this perception could influence predictions for future games. Investigating how ties affect the probability estimates people assign to future outcomes would provide valuable insights into how individuals perceive and process incomplete information, especially in sports with non-binary outcomes. Additionally, comparing the role of ties in human predictions to the performance of probabilistic models could help refine models that account for more complex, real-world event sequences, where outcomes are not always strictly binary.

By exploring these areas, we can better understand the broader cognitive and psychological factors at play in how individuals assess and predict sports outcomes, furthering our knowledge of sequential dependencies in sports and decision-making under uncertainty.

Author Contributions

I was solely responsible for the conception, design, analysis, and writing of this paper. All aspects of the project, including data collection, model development, analysis, and manuscript preparation, were carried out independently.

References

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. In *Proceedings of the seventeenth annual conference of the cognitive psychology society* (pp. 295–314). San Diego, CA: Elsevier.

Paul, R., Weinbach, A., & Humphreys, B. (2011, January). The belief in the "hot hand" in the nfl: Evidence from betting volume data. *Journal of Sports Economics*.

Vergin, R. C. (2000, June). Winning streaks in sports and the misperception of momentum. *Journal of Sport Behavior*, *23*(2), 181. Retrieved from https://link.gale.com/apps/doc/A62990408/AONE?u=mlin$_o websid = googleScholarxid = 835f7723$ (Accessed 15 Dec. 2024)