UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

BAYESIAN ANALYSIS OF STRIKE LIKELIHOOD IN BASEBALL: EVALUATING THE IMPACT OF PITCH DYNAMICS AND BATTER CHARACTERISTICS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirement for the

Degree of

MASTER OF SCIENCE

By

EBENEZER OLUBAYODE

Norman, Oklahoma

2025

BAYESIAN ANALYSIS OF STRIKE LIKELIHOOD IN BASEBALL: EVALUATING THE IMPACT OF PITCH DYNAMICS AND BATTER CHARACTERISTICS

A THESIS APPROVED FOR THE

DEPARTMENT OF HEALTH AND EXERCISE SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Daniel Larson, Chair

Dr. Christopher Black

Dr. Tyler Ransom

© Copyright by EBENEZER OLUBAYODE 2025

All Rights Reserved

Acknowledgments

I would like to express my deepest gratitude to my primary supervisor, Dr. Daniel Larson, whose support, patience, and guidance made this research possible. His timely feedback and encouragement at every stage of this journey not only eased challenges but also shaped the direction of this project.

I am also very grateful to my committee members, Dr. Christopher Black and Dr. Tyler Ransom, for their invaluable expertise in baseball analytics and their thoughtful contributions to this process. And a special thanks to Dr. Ransom, it was during his inspiring class that the initial idea for this project came from.

To my wife, friends and family, thank you for your constant motivation, understanding, and encouragement throughout this research. Your support kept me focused and driven, even during the most demanding phases.

Table of Contents

Abstract

In baseball, a pitch resulting in strikes is very important for pitchers as it shapes count leverage, controls games, and also dictates overall command. To maximize these advantages, pitchers must understand how pitch characteristics and context influence the likelihood of a strike outcome. This study examines those factors by focusing on pitch selection, movement, and matchup dynamics. Using a Bayesian logistic regression model, a stratified sample of 4,593 pitches from over 1.4 million MLB records was analyzed to determine how pitch characteristics and batter-pitcher handedness influence the probability of a pitch resulting in a strike. The findings show that vertical location (plate_z) is the most important factor with pitches higher in the zone, more likely to result in strikes. Sliders and Sweepers were particularly effective at generating strikes, especially in full counts or two-strike situations, and their effectiveness varied with batter side-of-the-plate matchups. While batter side of the plate does influence strike likelihood, its impact depends heavily on pitch type, count pressure, and other contextual variables.

On the other hand, pitches at the bottom edge of the zone (sz_bot) had a lower chance of resulting in strikes. By using Bayesian posterior distributions and contrast analysis, this study goes beyond simple correlations by providing a probabilistic, interpretable framework to understand pitch outcomes. These insights provide real-world applications for pitch design, scouting, and in-game strategy. Unlike previous research, this work incorporates strike zone variables, pitch movement, and handedness matchups through a causal lens in providing a richer and more actionable understanding of what really drives pitches to result in strikes.

Keywords: ‘Strike Likelihood’, ‘Bayesian Logistic Regression,’ ‘Pitch Dynamics’, ‘Causal Inference’, ‘Umpire Decision-Making’, ‘Pitch Location’, ‘Batter-Pitcher Matchups’, and ‘Baseball Analytics’.

Introduction **Chapter 1**

Background and Importance of Baseball Analytics

The baseball game is one of the most popular sports in North America and Japan. The sport is far more than a game of bats and balls; pitching is as much about strategy as it is about mechanics. Every throw is shaped by quick judgment and intent and for pitchers, landing a strike isn't just a routine goal, it sets the tone for the count, keeps pitch counts low, and helps maintain control of the game. Choosing how to place a pitch, whether by adjusting its location, type, speed, or movement is less about repetition and more about reading the moment and able to respond with precision. Looking at strike probability from the pitcher's side helps connect pitch features to what actually happens. It becomes a way to think through how pitches are planned, how they're used, and how decisions shift as the game unfolds. Using tools like Bayesian models gives pitchers and coaches a way to make decisions grounded in data and tailored to performance (Vock & Boehm Vock, 2018)

Advancements in Baseball Analytics

Baseball is a sport that is often described as a game of strategy, precision, and milliseconds, which makes it well suited for data-driven analysis. While most research has focused on batting results, recent work in baseball analytics has started to explore what shapes a pitch that result in a strike. Pitchers make quick decisions about what to throw and where to place it, trying to get a swing or earn a strike. Umpires, meanwhile, judge each pitch as it comes, relying on both where the ball lands and their own interpretation of the zone, which can vary. These interactions provides opportunity for statistical modeling, particularly through Bayesian methods that can quantify the likelihood of a strike based on pitch characteristics and contextual variables (Brantley & Körding, 2022). For instance, Bayesian integration models can be used to synthesize prior information with observational data to estimate pitch trajectories and understand how pitch type and location affect the likelihood of a strike under uncertainty (Brantley & Körding, 2022; Gray, 2002).

Research has shown that umpire strike zone judgments in Major League Baseball are not entirely consistent with rulebook definitions and are influenced by contextual and perceptual factors. For instance, umpires tend to contract the strike zone in two-strike counts and expand it in three-ball counts, leading to shifts in strike probabilities by as much as 19 percentage points (Green & Daniels, 2014). Additionally, the shape and size of the strike zone can vary with batter handedness and individual umpire tendencies (Huang & Hsu, 2020). Umpires have also been found to lower the top of the strike zone by an average of 2.64 inches, deviating from official specifications (Rainey & Larsen, 1988). While improvements in strike zone standardization have occurred, particularly between 2008 and 2016 (Zimmerman et al., 2019), these findings explain the complex, subjective nature of strike zone enforcement and highlight the importance of empirical modeling of strike likelihood across different pitch types and locations.

The Challenge of Correlation vs. Causation

The incorporation of data analytics in baseball has greatly improved how sports scientists model baseball performance, yet to distinguish correlation from causation in research remains a big challenge, as traditional statistical methods often blur the line between the two, leading to misinterpretations. Shrier (2007) shows that traditional statistical methods, such as regression and stratification, often fail to provide a solid foundation for causal inference and can lead to misleading conclusions. While machine learning and artificial intelligence show promise in predicting player performance and injury prevention (Mizels et al., 2022), relying solely on correlation can result in errors, particularly in complex, nonlinear dynamics (Cao & Qu, 2023)

As statistical methods and data visualization tools become more sophisticated, it's becoming harder to distinguish between causality and correlation in many studies Dong & Heineke (2016). Shrier (2007) highlights that recent advances of causal inference in epidemiology for sports injury prevention have shown traditional approaches, like regression, may not accurately adjust for confounding, explaining the need for improved methods to identify true cause-and-effect relationships. This problem shows the need for better methods to clearly identify cause-and-effect relationships in observational data. For instance, Rohrer (2018) study emphasized the significance of causal inference techniques like Directed Acyclic Graphs (DAGs) in mapping and understanding causal pathways. Observational studies, like those in physical activity research, often face similar difficulties in establishing causal links due to confounding and measurement errors. Lynch et al. (2020) highlight contemporary methods, such as directed acyclic graphs (DAGs), propensity scores, and causal mediation analysis, that can help improve causal inference in such contexts. Similarly, Pearl (2009) advocated for Bayesian causal frameworks to address these challenges, especially in domains like baseball analytics, where controlled experiments are often infeasible.

Observational Data and the Need for Causal Analysis

Controlled experiments are difficult to implement in baseball analytics due to the dynamic and uncontrolled nature of real-world games, where numerous variables, such as weather, player fatigue, and strategy, naturally interact. For example, it is impractical to replicate identical scenarios, such as repeatedly placing the same batter against the same pitcher under identical conditions, to isolate specific outcomes like the effect of a particular pitch type. It's also not fair or ethical to ask players to intentionally perform badly or stick to one specific strategy just for research. Similarly, in physical activity research, rare or long-term outcomes make randomized trials infeasible. Lynch et al. (2020) suggest that advanced causal inference methods, like g-methods and Mendelian randomization, can help tackle these challenges in observational studies.

The work of Myers et al. (2023) also highlights the value of quasi-experiments and causal modelling tools, including directed acyclic graphs (DAGs), to address these limitations and derive meaningful insights from observational data. Additionally, baseball is a complex game, things like team dynamics and player performance are connected in many ways, making it difficult to control or isolate a single variable.

Since controlled experiments are often not feasible, researchers depend on observational data drawn from real game situations. Distinguishing cause from correlation in this kind of data remains a major challenge. To address this, causal tools are needed. Kalkhoven et al. (2024) point out that using tools like causal diagrams, including directed acyclic graphs (DAGs) can help clarify causal assumptions, guide study design, and reduce issues like data dredging or p-hacking, leading to more meaningful and reliable insights.

Advanced techniques like Directed Acyclic Graphs (DAGs) and Bayesian causal inference as noted by Rohrer (2018), help address these challenges by modeling and inferring cause-and-effect relationships in complex datasets. Myers et al. (2023) emphasize the importance of understanding causal effects through advanced techniques like Bayesian inference and Directed Acyclic Graphs (DAGs), which can model cause-and-effect relationships from observational data.

These complexities underscore the need for strong causal inference in baseball analytics. Without robust modeling, teams may rely on surface-level patterns and miss key insights about how pitch characteristics and context influence umpire decisions. Tools such as Bayesian models and Directed Acyclic Graphs (DAGs) offer a structured way to identify confounders and draw more reliable conclusions about strike zone dynamics.

Recent Advances in Bayesian Methods for Baseball

Recent research has extended these methodologies to baseball. For example, Yee and Deshpande (2023) employed Bayesian Additive Regression Trees (BART) to evaluate plate discipline, integrating contextual variables like game state, player characteristics, and umpire tendencies. Their work gives a perfect example of how advanced Bayesian techniques can quantify uncertainties and optimize decision-making in baseball scenarios.

Objectives and Contributions of this Study

This study builds on existing work by applying Bayesian econometric models to analyze how pitch characteristics such as pitch type, release speed, and contextual features like batter's side of the plate and pitch location (plate_x and plate_z) affects the chance that a pitch is called a strike. Instead of modeling the exact shape of the strike zone, this analysis looks at how umpires respond to pitches in different parts of or near the zone. The goal is to move beyond surface-level associations and explore possible causal relationship in the pitch level observational data, and the results are presented from the pitcher's point of view, with the aim of informing pitch choices and location strategies to improve strike outcomes against different profile of batters.

Purpose Of Study

The purpose of this study is to analyze the factors that influence whether a pitch results in a strike. Specifically, it examines the effects of pitch dynamics (type, speed), batter's side of the plate and pitch location within the strike zone using Bayesian econometric modeling. This research seeks to provide pitchers and coaches with strategic insights to optimize pitch placement and selection based on strike likelihood.

Specific Objectives

Use Bayesian statistical methods to:

1. Determine the isolated, direct effects of key variables (batter handedness, pitch location, strike zone dimensions, pitch type, and speed) on strike Likelihood.
2. Analyze interaction effects by investigating whether the relationship between two variables changes depending on the presence or level of a third variable, highlighting complex dependencies e.g., Does the effect of pitch type on strike likelihood varies by batter's side of the plate.

Variables of Interest

Dependent Variables:

- Strike Likelihood (Y): The probability that a pitch result in a strike. This binary outcome (1 = strike, 0 = ball) is influenced by pitch type, location, and other factors.

Independent Variables:

- Stand (Batter Handedness): A categorical variable which indicates whether the batter is on the left-hand or right-hand batting i.e. batter's side of the plate.
- Release Speed: The speed that the ball travels starting from the point of the released from the hand of the pitcher which can affect the batter's timing and decision.
- Pitch location (plate_x and plate_z):

  plate_x: The horizontal position of the pitch as it crosses the plate, indicating whether it is inside or outside from the batter's perspective.

  plate_z: The vertical position of the pitch as it crosses the plate, indicating whether it is high, low, or in the strike zone

- Pitch Movement: This is the **horizontal (pfx_x)** and **vertical (pfx_z)** deviation of the pitch as it approaches the plate.
- Strike Zone (sz_top and sz_bot)

  sz_top: This refers to the top boundary of the strike zone for a specific batter, measured in feet. It varies from batter to batter, depending on their stance and height.

  sz_bot: This is the bottom boundary of the strike zone for a specific batter, and it is also measured in feet. Like sz_top, it adjusts to the batter's stance and height.

- Pitch Type: The different classes of the pitch (e.g., fastball, curveball, slider).

Research Problem and Research Questions

Main Research Question:

- The primary aim of this research is to analyze what factors determine the likelihood of a pitch result in a strike. This study addresses the following research questions:

Specific Research Questions:

- How do different pitch types influence strike likelihood?
- How does pitch location (horizontal and vertical) affect strike likelihood?
- How does batter handedness influence how pitch type affects the likelihood of a pitch resulting in a strike?

Delimitations

This research is focused on specific boundaries to maintain a clear and manageable scope. The study is restricted to modeling the likelihood of a pitch resulting in a strike, based on pitch type, pitch location (plate_x, plate_z), pitch movement, and contextual features such as batter handedness and pitch dynamics which keep the scope of the research precise and aligned with its objectives. Additionally, the data is drawn from two specific seasons, providing a consistent timeframe for the study.

The dataset is delimited to a stratified random sample of 4,593 pitches, drawn from a larger dataset of over 1.4 million records, using a uniform sampling fraction of approximately 0.354% across pitch type and batter combinations. Pitch types with extremely low representation (e.g., SC, PO, KN, CS, EP, FA) were excluded due to insufficient frequency during stratified sampling.

Therefore, this analysis focuses on the most frequently used pitch types in Major League Baseball, providing a detailed foundation for identifying consistent patterns in pitches that result in strikes.

Limitations

This research has several limitations that may impact its scope, despite the efforts to ensure how robust it could be. Mizels et al. (2022) research works explain that the complexity of confounding factors, such as ballpark effects and individual player tendencies, poses significant challenges to causal inference in baseball analytics. While advanced methods are used to mitigate the effects of confounding, it is impossible to completely eliminate all confounding variables, which may influence the conclusions. Moreover, inaccuracies in measuring pitch dynamics are still a concern since motion capture systems have limitations in reliably recording certain metrics Mizels et al., (2022)

To ensure fair representation, the study used stratified sampling based on batter and pitch type combinations. While this helped balance the sample, it also meant that rare pitch types were excluded from the final dataset. As a result, the model doesn't generalize as well to pitch types that don't show up often in competitive games.

Also, the way a batter positions themselves, crouching lower, or subtly shifting their position in the box, might shape how an umpire sees the strike zone because it introduces further variability that's difficult to quantify but could still shape strike outcomes. Finally, even though pitch-tracking technologies provide detailed data, they're not perfect. Small inaccuracies in capturing pitch location, especially for borderline calls can affect precision.

Assumptions

The study is based on several assumptions to facilitate analysis. This study assumes that the effects of the predictors on the outcome are modeled using weakly informative normal priors: Normal (0, 10). A mean of 0 indicates no strong prior assumption about the effect's size or direction, while the standard deviation of 10 provides enough flexibility for the data to guide the posterior estimates. These priors also provide regularization to prevent extreme or unrealistic parameter values in the absence of strong evidence.

The use of weakly informative priors is supported by their growing adoption in Bayesian analysis as a way to balance prior information with data-driven inference. For instance, Student-t priors (centered at 0 with a scale of 2.5) are often recommended as defaults for modeling logistic regression by Gelman et al. (2008). This is because they perform better than Gaussian or Laplace priors in cross-validation. Rover et al. (2020) highlighted the benefits of using weakly informative priors for heterogeneity parameters in hierarchical models, particularly when data is limited. These priors are particularly useful in causal analysis, as they help avoid overfitting and ensure that variation is maintained in a reasonable way. It is a great approach that demonstrates a careful balance between flexibility and robustness in the methods used for this study. Therefore, this analysis assumes that the prior distributions in the Bayesian models are chosen to appropriately reflect the underlying data.

The model controls for batter and pitcher specific characteristics using fixed effects, which account for player-specific traits such as tendencies, mechanics, or general decision styles, that are assumed to remain stable over the dataset. While real-world performance can shift due to factors like fatigue or in-game strategy, these unobserved traits are absorbed through the fixed-effects structure. In addition, the model incorporates observable variables such as batter handedness (stand), pitch type, pitch location (plate_x, plate_z), and personalized strike zone dimensions (sz_top, sz_bot) to show meaningful differences between matchups. The model are able to capture the variability in how different batters and pitchers influence strike through these features.

The relationship between pitch characteristics and strike likelihood is modeled with the understanding that it may vary across different batters and situations. Including fixed effects, random effects, and interaction terms between batter and pitch type allows the model to account for how specific pitch–batter combinations relate to strike outcomes. This structure captures how pitch location, type, and player characteristics interact in determining whether a pitch results in a strike.

Finally, while the model captures a wide range of variability through fixed effects, random effects, and observable predictors but some factors remain unmeasured. These include catcher framing, weather, and changes in batter posture. Such factors may affect borderline strike calls but are assumed not to introduce systematic bias. Future research could build on this analysis by incorporating such contextual features, offering a more precise understanding of how certain pitch characteristics consistently generate strike outcomes across different game conditions.

Operational Definitions

- Strike Likelihood (Y): is the probability or chance that a pitch is classified as a strike, based on the pitch description and the batter's current strike count. It is calculated as a binary outcome (Strike = 1 or Not Strike = 0) where a pitch is considered a strike (Strike = 1) if its description is 'called_strike', 'swinging_strike', 'foul_tip', or 'foul_bunt'. A foul pitch is counted as a strike only if the batter has two or more strikes; otherwise, it is not counted as a strike. Pitches with descriptions such as 'ball', 'hit_into_play', 'hit_by_pitch', 'blocked_ball', or 'pitchout' are considered not strikes (Not Strike = 0). Strike likelihood is defined as the chance that a pitch is either in the strike zone or is missed when the batter swings. It is based on the batter's current strike count and the type of pitch. This approach shows a pitcher-centric perspective, where any strike, whether it is called by the umpire or generated through batter action, is considered a favorable outcome. From the pitcher's perspective, the goal is to get ahead in the count or generate outs. Whether the strike is called or swung at, it benefits the pitcher.

- Bayesian Model: A Bayesian model is a statistical framework that incorporates prior knowledge or beliefs (expressed as prior distributions) and updates these beliefs using observed data to estimate the posterior distributions of parameters i.e. it estimates the probability of different outcomes. Bayesian models were used in this study to understand the causal impact of variables of interests on strike likelihood by combining prior assumptions with evidence from pitch-level data.

- **Observational Research:** Observational research involves studying and analyzing data without manipulating the environment or variables. In this context, observational research examines the relationships between variables like pitch dynamics, batter handedness, and strike likelihood as they naturally occur in baseball games.

- **Causal Directed Acyclic Graph (DAG):** A causal DAG is a graphical representation of variables and their causal relationships, where nodes represent variables (e.g., pitch type, release speed) and directed edges (arrows) indicate causal influence. The "acyclic" nature ensures no feedback loops, meaning causality flows in one direction. DAGs help identify confounders, mediators, and causal pathways in this study.

- **Confounding:** Confounding occurs when an external variable influences both an independent variable (e.g., pitch type) and a dependent variable (e.g., strike likelihood), creating a spurious association. For example, pitch type affects both release speed and strike likelihood, making it a confounder that must be controlled for accurate causal inference.

- **Confounder:** A confounder is a variable that affect both the independent and outcome variables. In this study, pitch type can be identified as confounders that influence both the predictors and strike likelihood.

- **Direct Effect:** The direct effect represents the causal impact of one variable on another without involving mediators. For example, the direct effect of batter handedness on strike likelihood bypasses intermediate variables like pitch dynamics.

- **Indirect Effect:** The indirect effect refers to the influence of one variable on another through mediators. In this study, the effect of pitch type on strike likelihood through release speed and movement dynamics (pfx_x, pfx_z) is an indirect effect.

Literature Review **Chapter 2**

This chapter aims to comprehensively review existing literature pertinent to understanding the Bayesian analysis of pitch dynamics and batter characteristics in baseball. Building on the foundational concepts introduced in Chapter 1, this section reviews empirical work that uses Bayesian and causal models to study how pitch characteristics and game context relate to strike outcomes. It will explore the nuanced relationships between the variables under study specifically, how pitch dynamics (such as velocity, movement, and location) interact with batter handedness and pitch count to influence umpire decision-making. This review also highlights the role of confounding variables such as pitch type, and game context that may bias or obscure these relationships.

In delineating the scope of this literature review, emphasis will be placed on dissecting the layers of research that contribute to our understanding of these phenomena. The chapter will explore the theoretical underpinnings Bayesian analysis and causal inference in baseball analytics, with a focus to investigate the effects of pitch and batter characteristics on strike likelihood. Specifically, show how modern analytics can move beyond correlation and offer actionable insights for pitchers and strategists.

To enrich our exploration and ensure a thorough examination of the topic, a search was conducted using various keywords such as "batter's handedness", "pitch characteristics", "umpire bias", "Strikes Outcomes", "pitch type and speed", "batter's strike zone", "pitcher and batter dynamics", "Bayesian analytics in baseball", "Causal and correlation effects in baseball", and "baseball analytics" within sport literature articles across different databases. Articles containing these keywords and clear connections to the topic were reviewed. While none directly address the specific gap this study investigates, they offer usefu insights and understanding into this topic. Furthermore, this review will identify gaps within the current body of research, paving the way for the exploration of uncharted areas in the subsequent chapters of this study, particularly the lack of integrated Bayesian econometric approaches that model strike likelihood in relation to pitch characteristics and batter abilities. Pointing out these gaps helps show where more research is needed and it also opens up ways to think more clearly about how pitcher and umpire interactions affect what happens during a game.

2.1 Pitcher-Umpire Interactions in Baseball

Understanding how umpires judge pitches isn’t just about tracking the strike zone but it’s also about unpacking the hidden influences that turn a borderline pitch into a strike or a ball. While much of the research on umpire decision-making focuses on batter behavior, this section examines how pitchers, through the pitches they throw and the situations they create, shape the calls umpires make. By using insights from sports analytics and cognitive psychology, it becomes easier to start to see how the strike zone variability arises and where opportunities for modelling exist. Recent studies further highlight how pitcher characteristics beyond count and context shape umpire decisions. For instance, pitch sequencing (e.g., velocity/movement correlations) significantly predicts strikeout rates (Healey & Zhao, 2017). These findings suggest that pitcher-specific traits, combined with cognitive biases, create a complex landscape for modeling strike likelihood.

Umpire Biases: The Human Element Behind the Call

Umpires might strive for objectivity, but their decisions are undeniably human. One of the clearest patterns is how the strike zone bends depending on the count. When a pitcher gets to a 3-ball count, the zone suddenly seems to expand, umpires become more likely to call a strike, as if reluctant to hand out a walk (Green & Daniels, 2015; Huang & Hsu, 2020). On the other hand, with two strikes, that same zone tightens up, almost as if umpires are avoiding the finality of strikeouts (Green & Daniels, 2015). This reflects a form of impact aversion, where umpires tend to hold back on calls that carry more weight. And then there’s the streakiness of it all. Just like a gambler’s fallacy (negatively autocorrelated calls after streaks), umpires tend to avoid calling consecutive strikes, especially when they’re less experienced (Chen et al., 2016; Green & Daniels, 2014). For pitchers, these biases aren't just quirks, they can be used as part of a strategy. Knowing how umpires respond in certain counts can shape pitch sequencing and sometimes mean the difference between a strikeout and a walk. These biases are compounded by broader contextual factors. Home batters, for example, receive more favorable calls (Hsu, 2023), and umpires may call more strikes to hasten game pace (MacMahon & Starkes, 2008). Such tendencies show the need for econometric models that control for non-pitch variables (e.g., home/away status) to isolate the true impact of pitcher-umpire dynamics.

What Pitches Get the Call? The Unanswered Questions

Most of the research on umpire decisions emphasizes counts and game situations, but what about the pitches themselves? We know that expert umpires are better at tracking ball flight and pitcher kinematics (Chen & Huang, 2024), but does that mean a 98 mph fastball gets more borderline calls than a looping curveball? And why do umpires struggle so much with left-handed pitchers? (Chen & Huang, 2024). Maybe it’s the with their release points or ball movement, However, it is a gap this research aims to address. There’s surprisingly little work on whether certain pitch types like sliders, sinkers, etc. confuse umpires or make them more likely to call a pitch wrong, or if high-spin fastballs (movement profiles) sneak strikes that slower pitches don’t. Those are gap worth exploring, especially for pitchers looking for every possible edge.

Experience Matters But Not How You’d Think

Experience umpires have a reputation for being more stubborn with their strike zones, and the research supports this. They tend to call more strikes, maybe as a way to keep control of the game (MacMahon & Starkes, 2008; Chen & Huang, 2024). Over time, they also get better at sticking to the rulebook definition of the zone (Zimmerman et al., 2019), though every umpire still has their own individual variability on the strike zone. But here’s the interesting part: even people with no baseball experience fall into the same biases as professional umpires when judging pitches (MacMahon & Starkes, 2008). That suggests some of these tendencies, like the count effects are rooted in how humans process uncertainty, not just something learned through years behind the plate. For pitchers, that raises a bigger question about whether pitchers can exploit umpire-specific tendencies (e.g., targeting high-break pitches to certain umpires)? All of this points to a bigger idea that the strike likelihood isn’t just about where the pitch crosses the plate. It’s a product of the pitch’s type, movement, the count, the umpire’s experience, and even subconscious biases that nobody likes to admit. That’s where probabilistic models like Bayesian approaches and causal inference become especially useful because they help show how different variables combine to influence strike outcomes. For pitchers, this goes beyond just aiming for the zone. It’s about knowing when a pitch is more likely to result in a strike, and adjusting based on what’s happening in the moment.

2.2 Bayesian and Econometric Approaches in Baseball Analytics

Bayesian models are useful for handling uncertainty, which makes them a good fit for studying strike likelihood in baseball. For example, Bayesian models have been applied to evaluate fielding performance using hierarchical structures (Jensen et al., 2008) and to predict hitting outcomes by incorporating covariates such as player age and position (Jensen et al., 2009). Similarly, Brantley and Körding (2022) demonstrated how Bayesian integration allows batters to synthesize sensory input with prior information, improving their ability to estimate pitch trajectories under uncertain conditions.

Econometric approaches complement Bayesian methods by providing tools to estimate the probability that a pitch result in a strike, accounting for both observable pitch characteristics (e.g., pitch type, velocity, location) and unobserved player-specific effects through hierarchical structures. For example, Bendtsen (2017) used gated Bayesian networks to track changes in player performance over a career, linking major events to shifts in productivity. This kind of approach moves beyond basic descriptions and helps explore cause-and-effect patterns that matter for real decisions.

Causal Inference in Observational Data

It's often difficult in sports analytics to tell if one thing actually causes another or if they just happen to be related, especially when using observational data. Traditional statistical methods can identify patterns, but this method doesn't always provide the evidence that is needed to show clear cause-and-effect relationships, which are key to making informed decisions. Traditional statistical methods often reveal associations but fail to establish causal relationships that are essential for actionable insights. Directed Acyclic Graphs (DAGs) have emerged as a useful tool in addressing this issue as it enables researchers to visualize causal pathways, identify confounders such as pitcher handedness or count and guide model specification to improve causal interpretation (Rohrer, 2018; Krieger & Davey Smith, 2016). For example, Park (2019) in their research work used DAGs to derive new metrics for player evaluation by emphasizing their utility in baseball analytics.

Bayesian causal frameworks expand on these techniques by combining probabilistic reasoning with causal diagrams, offering a systematic approach to modeling uncertainties and dependencies. These frameworks are very useful in handling challenges like unobserved confounding, selection bias, and heterogeneous populations (Hunermund & Bareinboim, 2019). Researchers have integrated econometric methods with Bayesian frameworks to create models that help measure both direct and indirect effects, that provides a clearer connection between descriptive insights and causal analysis making them well suited for analyzing Strikes behavior. While traditional models often capture only associations, Bayesian econometrics enables better interpretations such as understanding how pitch type might affect strike probability differently depending on batter handedness or pitch location. These capabilities make Bayesian approaches a strong methodological choice for modeling strike zone dynamics from the pitcher's perspectives.

2.3 Applications for Game Strategy and Pitching Dynamics

Bayesian methods provide deeper insights into how players perform in different situations having been used to analyze decision making, including pitcher performance and strategic adjustments over time. Brill et al. (2022) used Bayesian analysis to study the "Time Through the Order Penalty" (TTOP), showing how pitcher performance can change over the course of a game.

A similar pattern may apply to umpires, whose decisions might shift later in the game due to fatigue, repetition, or situational pressure.

Similarly, Yang and Swartz (2021) developed a two-stage Bayesian model to predict game outcomes by incorporating player ability and starting pitchers to predict game outcomes. Though focused on performance forecasting, these models show the flexibility of Bayesian methods to handle layered, context-driven decision variables, which is a quality this study uses to model how pitch type, location, and player traits influence the likelihood of a strike

Adoption and Impact of Bayesian Methods

The adoption of Bayesian methods in sports analytics has grown significantly in recent years. Between 2013 and 2017, the number of studies leveraging Bayesian frameworks rivaled the output of the previous three decades (Santos-Fernandez et al., 2019). This rapid growth shows the increasing ability of these methods to handle complex datasets, quantify uncertainty, and integrate prior knowledge with real-time data, critical when trying to understand umpire judgments that are shaped by both rules and human variability. Machine learning techniques, such as regularization and ensemble learning, have further enhanced Bayesian models by enabling researchers to tackle high-dimensional datasets without sacrificing interpretability (Silahtaroğlu, 2024).

Big data and machine learning provides powerful tools for analysis, but they fall short when it comes to identifying causal relationships between variables. As Todri & Stern (2015) note, without structured causal frameworks, prediction models can miss the *why* behind the *what*. This is where Bayesian causal models excel in a way it uncovers direct and mediated effects, such as whether a slider low and inside is systematically more or less likely to result in a strike, depending on the pitcher's handedness or the count. The integration of econometrics, causal inference, and

Bayesian thinking supports a more structured and interpretable way to understand how pitch characteristics, contextual features, and batter-pitcher interactions influence strike probability. Rather than simply modeling umpire tendencies, it helps pitchers and teams assess the conditions under which a pitch is most likely to result in a strike. This intersection of statistical modeling and game strategy enables more informed pitch sequencing, matchup planning, and player development. As Hunermund & Bareinboim (2019) suggest, integrating causal reasoning with Bayesian frameworks deepens prediction accuracy, and also the interpretability needed for actionable insights in high-stakes decision environments like baseball

This synergy between econometrics, machine learning, and Bayesian methods enables researchers to enhance model complexity, predictive accuracy, and causal inference in advancing the understanding of complex phenomena and supporting better decision-making (Zhao et al., 2024; Silahtaroğlu, 2024).

2.4 Bridging Econometrics and Bayesian Causal Inference

Econometric models grounded in Bayesian causal frameworks help untangle complex relationships in baseball data. In this study, they are used to examine how pitch type, location, and pitcher handedness affect the chance that a pitch results in a strike. The study by Hunermund and Bareinboim (2019) introduced advanced methods to address selection bias and unobserved confounding, which are particularly important when analyzing observational data where experimental control is not possible.

Through this approach, researchers can go beyond identifying statistical associations and instead estimate the *true effect* of specific pitch features on the likelihood of a strike. Rather than treating strikes as purely objective outcomes, this framework acknowledges that strike likelihood is shaped by layered factors which include the visual profile of a pitch and its contextual placement within the count or sequence.

In doing so, this research supports a broader shift in sports analytics toward models that not only predict outcomes but also explain the mechanisms that are behind it. In this way, the model helps identify how context matters, measure pitch-level effects, and better understand how pitch design and game situations influence strike outcomes from the pitcher's point of view.

2.5 Batter Handedness and Pitcher Handedness

Batter handedness plays a key role in strike variability. Prior research shows that left-handed batters tend to face more right-handed pitchers which create a unique visual angles and pitch trajectories that may influence how pitches are perceived by both batters and umpires (Chance & Maymin, 2023). These differences can alter framing opportunities and strike zone presentation, making handedness a meaningful factor to include in models of pitch-level strike likelihood. It also considers the side of the plate a batter stands on (left or right) and how that may interact with pitch type, location, and pitcher handedness to influence Strike Likelihood. By controlling for batter handedness (stand), the model accounts for this structural variability in batter-pitcher matchups which can help improve the precision of estimated strike probabilities.

Research further suggests that handedness modifies pitch movement perception. For instance, left-handed sliders tend to break away from right-handed batters and move in toward lefties, creating asymmetric pitch shapes that may influence umpire calls differently based on batter-pitcher orientation (Adler, 2019). Since left- and right-handed pitchers also differ in arm angles and release points, the same pitch type may appear to move differently depending on who throws it. These differences add to the visual complexity for umpires and show the importance of accounting for pitcher handedness and pitch movement profiles when modeling strike likelihood.

Pitcher handedness influences not just batter performance, but also how pitches are perceived by umpires. Research shows that left-handed pitchers are particularly effective due to their relative rarity, which introduces unfamiliar visual angles and pitch trajectories (Chance & Maymin, 2023). This unfamiliarity may also affect umpire judgment, as supported by studies showing decreased accuracy in calls involving left-handed pitchers (Chen & Huang, 2024).

From a modeling perspective, handedness creates structural variability in batter-pitcher matchups and pitch presentation. A slider thrown by a left-handed pitcher, for instance, might breaks away from a right-handed batter, which may be visually and spatially different than the same pitch from a right-hander. These interactions explains the importance of including pitcher handedness in models of strike likelihood, especially when accounting for perceptual and positional factors influencing umpire decisions.

2.6 Pitch Type as a Confounding Variable in Baseball Analysis.

The influence of confounding variables, such as pitch type, pitcher's throwing hand, has been extensively studied to understand their impact. For example, the type of pitch thrown (e.g., fastball, curveball, etc.) is often dictated by the count and game state, with fastballs being more prevalent in early innings and breaking balls increasing in frequency later in games (Ganeshapillai & Guttag, 2012; Whiteside et al., 2016).

Pitch type matters when it comes to strike likelihood, as it shapes how the pitch looks in terms of speed, movement, and trajectory. Fastballs, which tend to follow a straighter path, are often result in strikes more frequently, especially near the edges of the zone compared to breaking or off-speed pitches, which introduce greater lateral or vertical deviation. These variations can influence how a pitch is perceived relative to the strike zone, particularly in contexts where pitch movement interacts with batter or pitcher handedness. Modeling strike likelihood must therefore account for pitch type as a confounding variable that shapes both trajectory and contextual interpretation of zone boundaries.

Studies indicate that pitch type impacts batter timing errors, with prior knowledge of an upcoming pitch reducing mistakes (Kidokoro et al., 2020). Over the course of a game, pitchers tend to shift their strategy by throwing fewer hard pitches and relying more on breaking balls and off-speed pitches, leading to decreased velocity and increased vertical movement as early as the fifth inning (Whiteside et al., 2016). This shows that pitch type contributes to the flow of the game by affecting both how batters react and how pitchers make adjustments over the course of a game.

Researchers have used Bayesian methods for pitch classification, reinforcing the importance of pitch type in statistical modeling. Umemura et al. (2020) used Variational Bayesian

Gaussian Mixture Models (VBGMM) to classify pitch types with improved accuracy when factoring in fastball characteristics. Similarly, Gomaz et al. (2023) applied machine learning techniques to classify pitch types based on pelvis and trunk kinematics, achieving 71% accuracy in fastball identification. These studies show how statistical models rely on pitch type classification in improving predictive accuracy, reinforcing its role as a confounding factor in baseball analytics.

Furthermore, pitch quality measured as the expected number of bases conceded is influenced by pitch count, location, type, and speed, making pitch type an essential factor in modeling batter behavior and strike likelihood (Swartz et al., 2017). Pitch type influences factors like movement, speed, release mechanics, in strike likelihood. Because of this, pitch type should be included in causal models to avoid distorted results. Recognizing it as a confounder helps clarify its influence on other variables and supports more accurate analysis of performance.

2.7 The Role of Pitch Location and Strike Zone Boundaries on Strike Likelihood

Umpires evaluate pitch location relative to the strike zone boundary, and prior research shows that both pitch height (plate_z) and horizontal location (plate_x) strongly influence strike. However, strike zone enforcement is not always consistent, particularly near the edges, where cognitive biases, prior pitch sequencing, and visual limitations can affect decision accuracy.

Pitches above sz_top or below sz_bot are expected to fall outside the rulebook strike zone, yet umpires occasionally call strikes at these locations, depending on contextual cues like pitch type or count. Similarly, horizontal boundaries (plate_x near the edges) may lead to more variability in judgment. Zimmerman et al. (2019) observed that pitches near the zone edge are more likely to receive inconsistent calls, especially when combined with vertical extremity.

While much of the existing work focuses on descriptive trends, this study uses causal inference to identify the direct effects of pitch height and width on strike calls, while adjusting for confounders like pitch type and batter handedness. This approach helps clarify how pitch location, zone boundaries, and matchup context affect strike outcomes, insights that matter especially from the pitcher's point of view.

2.8 The Role of Spin, Speed, and Ball Movement in Strike Likelihood

The lateral and vertical movement of pitches which is also shaped by spin rate and velocity, plays an important role whether a pitch is likely to result in a strike. For example, two-seam fastballs show more horizontal movement than four-seamers, which may cause borderline pitches to appear off the plate (Keeley et al., 2010). Spin-induced lift or drop, explained by the Magnus effect (Nagami et al., 2013), affects vertical movement, potentially pushing a pitch above or below the expected strike zone. Elite pitchers use high spin rates to manipulate trajectory, and such unexpected movement can increase the likelihood of a strike, especially on pitches near the edges of the zone. When a pitch breaks sharply at the last moment, it can deceive both batter and umpire, enhancing the pitcher's chances of earning a strike due to its unexpected path.

Visual perception studies suggest that umpires, just the same way batters, rely on early pitch cues and anticipate trajectories. When pitches deviate from standard movement patterns due to spin, arm angle, or velocity can increase or decrease strike likelihood depending on pitch type and context (Higuchi et al., 2013). This justifies the inclusion of spin-related variables like pitch movement (pfx_x, pfx_z) in strike likelihood models and supports examining whether certain pitch types are more prone to result in strike.

2.9 Research Gaps in Batter-Pitcher Dynamics and Performance Analytics

Despite the growing use of analytics to understand strike zone behavior, but important research gaps remain, especially in relation to modeling the probability that a pitch result in a strike based on its physical traits and contextual variables.

Limited Integration of Causal Inference

While Bayesian methods and advanced predictive models have gained traction in baseball analytics, most studies focus on correlation-based analyses rather than causation. The use of causal inference techniques, such as Directed Acyclic Graphs (DAGs) and Bayesian causal frameworks, remains sparse in batter-pitcher research. As Rohrer (2018) and Pearl (2009) have emphasized, distinguishing correlation from causation is essential for generating actionable insights in baseball decision-making.

Traditional models, which include machine learning techniques, fall short in identifying the causal effects. This study contributes by leveraging Bayesian causal inference to build interpretable models that improve both prediction and understanding of strike likelihood behavior, especially from the pitcher's perspective.

Incomplete Integration of Batter and Pitcher Handedness

While batter and pitcher handedness are commonly included as control variables, their interactive effect on pitch presentation and strike likelihood remains largely unexplored. For instance, a left-handed pitcher throwing to a right-handed batter creates a different visual profile and pitch trajectory than a same-handed matchup but few studies examine how these combinations may influence strikes outcomes across pitch types or locations. This study closes that gap by directly modeling handedness dynamics within strike likelihood.

Lack of Joint Modeling of Strike Zone Boundaries and Pitch Characteristics

Although many studies show that pitches on the edge of the strike zone are called inconsistently, few go further to investigate how those edge effects vary based on pitch speed, spin, or movement. This gap limits the understanding of how pitch-level features can amplify or mitigate strike zone uncertainty. A fastball just above the top of the zone may be judged differently than a slow curveball in the same spot, but this information is often missed. This research addresses that gap by evaluating the interaction between pitch location and physical pitch traits in predicting strike outcomes.

Methods **Chapter 3**

Research Designs and Data Collection Method

This study adopts an econometric modeling framework using Bayesian logistic regression to evaluate the likelihood of a pitch being a strike outcome. The analysis focuses on how pitch traits, zone location, and player matchups affect the chances of a strike. It also offers a way to think about how pitchers can adjust their pitch selection and placement to improve outcomes. Specifically, the model assesses how factors such as pitch type, release speed, pitch movement (horizontal and vertical), pitch location at the plate, batter handedness (stand), and pitcher handedness (p_throws) contribute to the probability of a pitch resulting in strike.

To ensure manageable computation and focused analysis, a stratified random sample was drawn from the original dataset containing over 1.4 million MLB pitch records from the MLB Statcast data provided by the Miami Marlins (data strategy team), a professional baseball team. Sampling was performed at the pitch type and batter group level using a uniform sampling fraction of approximately 0.354%, resulting in a final dataset of 4,593 pitches. This stratified approach preserved the underlying batter-pitch type distribution while reducing computational complexity for Bayesian estimation. The resulting dataset includes 598 unique batters, ensuring sufficient representation across different player profiles and minimizing the risk of individual-level skew and 832 pitchers.

Due to the low representation of some pitch types (e.g., SC, PO, KN, CS, EP, FA), these were excluded during the sampling process as they did not meet the minimum threshold for inclusion. These rare pitch types accounted for a negligible fraction of the full dataset and are uncommon in competitive gameplay. As a result, the final analysis focuses on the most frequently observed pitch types, which includes FF, SL, SI, CH, CU, FC, KC, ST, and FS and they form the foundation for modeling consistent patterns in strike decisions.

Statistical Analysis

Before implementing the Bayesian model, preliminary exploratory analysis was conducted to assess potential relationships between release speed and Strike outcomes and a panel logistic regression to examine associations between pitch characteristics, batter/pitcher traits, and strike outcomes across multiple batters. These initial analyses provided the initial insights into key associations and guided variable selection for the final Bayesian specification.

The primary modeling approach employs Bayesian logistic regression, with the binary outcome indicating whether a pitch results in a strike (Strike = 1) or not (Strike = 0). The Bayesian logistic regression model includes several predictors:

- Pitch dynamics: pitch_type, release_speed, pfx_x, pfx_z
- Strike zone variables: plate_x, plate_z, sz_top, sz_bot
- Batter and pitcher characteristics: stand (batter handedness), p_throws (pitcher handedness)

These selected variables are based on prior literature and variables like pitcher handedness are treated as potential confounders, as they can alter how pitch location is perceived and interpreted by umpires. For instance, the same pitch on the inside edge of the plate may be judged differently depending on whether it comes from a left-handed or right-handed pitcher.

The pitch movement components (pfx_x, pfx_z) may also function as mediators of pitch type effects, capturing perceptual changes that influence umpire judgment. Additionally, batter handedness (stand) is modeled to assess interactions with pitch location and type, given documented evidence that strike zones shift with batter orientation. Continuous variables release_speed were standardized, (plate_x, plate_z, pfx_x, pfx_z) were standardized and squared to capture non-linear effects while sz_top and sz_bot were standardized only, as reference strike zone thresholds. This standardization was done using z-scores transformation to ensure there is consistent scale across predictors.

To capture the causal and interaction structure among variables, Directed Acyclic Graphs (DAGs) were developed to identify adjustment sets and guide model specification. Markov Chain Monte Carlo (MCMC) methods were done for estimation, particularly the No-U-Turn Sampler (NUTS), which efficiently explores the posterior distribution of model parameters. Credible intervals were derived from High-Density Intervals (HDI), and contrast means were computed to assess the strength and direction of key effects. The analysis was carried out using Python, with Bayesian modeling which was implemented via the PyMC4 library for probabilistic inference and posterior estimation.

Data Sampling Justification

The sample size of 4,419 observations, covering 590 unique batters, was chosen to balance computational efficiency with the precision required for Bayesian estimation. Prior research in Bayesian inference supports the adequacy of sample sizes in the range of 2,000–5,000 observations for reliable posterior estimation when using efficient samplers like NUTS. This subset supports deeper analysis while preserving the variation in batter-pitcher matchups across pitch types and handedness. For future predictive modeling, the full dataset can be used to improve generalizability and model performance.

Research Method: Bayesian Inference Workflow for Strike Likelihood

Step I: Defining the Generative Model of the Sample

The analysis began by building a generative model which represents the factors thought to influence strike probability. This model was used to describe how the distribution of outcomes relates to observed variables, forming the basis for applying Bayesian inference.

In this study, we model the probability of a strike decision using logistic regression, which is well-suited for binary classification problems where the outcome is either a strike (1) or no strike (0). The key variables considered in this study to influence strike likelihood include:

- X_1: Pitch speed
- X_2: Pitch type (numerically encoded)
- X_3: Horizontal movement of the pitch (pfx_x)
- X_4: Vertical movement of the pitch (pfx_z)
- X_5: Pitcher handedness (p_throws)
- X_6: Pitch location (plate_z)
- X7: Pitch location (plate_x)
- X8: Strike Zone (sz_top)
- X9: Pitch location (sz_bot)

This formulation follows the foundational approach of generative modeling in probabilistic inference, where the distribution of the response variable (strike likelihood) is expressed as a function of predictor variables.

Generative models are essential for capturing how different variables interact, especially when dealing with uncertainty and causal analysis (Wan & Wei, 2019). They help in mapping out probability distributions in a structured way, giving us a clearer picture of the factors influencing an outcome. Despite the facts generative models offer a great foundation for understanding data patterns, they aren't enough on their own to establish causal relationships or inference. Additional steps, such as making informed assumptions about the data-generating process, are necessary to draw meaningful causal conclusions (Megiddo, 2023).

A key advantage of using a generative approach is its ability to handle variability in pitch characteristics while preserving the relationships among variables. Compared to discriminative models, which can directly estimate the decision boundary between strike and no-strike, generative models offer more flexibility in incorporating prior knowledge and handling missing data through Bayesian updates (Salojärvi et al., 2005). Additionally, they allow us to integrate semi-supervised learning techniques, that leverage information from labeled and unlabeled data sources to refine posterior distributions (Argouarc'h et al., 2024).

Despite their strengths, generative models come with challenges. One such limitation is their reliance on correctly specifying the probability distribution of input variables. If the model assumptions are incorrect or very simple, estimates can become biased, particularly in cases of misclassified outcomes (Gerlach & Stamey, 2007). To mitigate this, we ensure the logistic regression framework used is detailed by evaluating model fit and exploring alternative specifications where necessary.

Moreover, graphical models can be incorporated to explicitly represent sampling assumptions and assess whether consistent estimation of outcomes is feasible under different observational settings (Didelez et al., 2010). This is particularly relevant when dealing with pitch types that appear less frequently in the dataset, where important sampling techniques may improve model accuracy (Wan & Wei, 2019). While our approach focuses on Bayesian inference with a generative logistic regression model to analyze strike likelihood, other methods have been explored in research to tackle challenges like unobserved confounders. For example, some studies have used variational autoencoders and adversarial training to refine estimates and improve causal effect analysis (Zhu et al., 2018). These techniques weren't part of our analysis, but they show how the field is evolving. In future work, approaches like these could be useful for further improving how we model strike likelihood, especially when dealing with hidden variables that might influence a batter's choices.

A generative framework for modeling strike likelihood provides the basis for Bayesian inference which allows posterior estimates to be updated as new data are introduced. This is crucial for understanding how contextual and pitch-level features shape the likelihood of a pitch resulting in a strike, especially under conditions of uncertainty. Capturing these patterns will allow teams to adjust pitch framing, anticipate changes in zone judgment, and adapt pitcher strategy to improve strike outcomes.

Generative Model for Strike Likelihood

The generative model for the strike likelihood Y can be defined using a logistic regression framework: $P(Y = 1 \mid X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\beta_4 X_4+\beta_5 X_5)}}$

Strike Probability**:** P (Y=1 | X) This represents the probability that a pitch results in a strike (Y = 1), given a set of observed predictor variables.

- Y is the dependent variable where Y = 1 if the pitch is a strike and Y = 0 if the pitch is not a strike.
- X is the set of independent variables influencing the strike likelihood, which include X_1, X_2, X_3, X_4, X_5, X_6………. Xn.

Logistic Function (Probability Mapping)

The right-hand side of the equation is the logistic function (also called the sigmoid function):

$$\frac{1}{1 + e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\beta_4 X_4+\beta_5 X_5)}}$$

This function converts any real number into a value between 0 and 1, which makes it well-suited for binary classification problems such as predicting whether a pitch results in a strike. The natural base logarithm is e, which shapes the curve of the function, β_0 is the intercept term, and β_1, β_2, β_3, β_4, β_5, β_6…. βn are coefficients that determine how strongly each predictor variable influences the probability of a strike.

How to Interpret the Model

The expression $\beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$……. n is called the linear predictor. It combines all the weighted influences of the predictor variables into a single value. The logistic function transforms this linear predictor into probability. If the linear predictor is very large and positive, the probability approaches 1 (high chance of a strike), similarly if the linear predictor is very large and negative, the probability approaches 0 (low chance of a strike) but if the linear predictor is close to zero, the probability is around 0.5. making the model to be uncertain about the decision.

The probability that a pitch result in a strike was modeled using logistic regression, drawing on pitch movement, location, and the batter-pitcher pairing. Rather than predicting individual outcomes, the approach helps make sense of how certain pitch traits are linked to strike outcomes in different games.

Step II: Define Estimands

Understanding Estimands and Their Relevance to Sports Analytics

Estimands are precise target quantities that guide the focus of a statistical analysis, making sure the study's results truly reflect its theoretical goals (Lundberg et al., 2020, 2021).

Originally formalized in clinical research under the ICH E9 R1 guideline, estimands address issues of ambiguity in treatment effect estimation by accounting for intercurrent events such as patient dropouts or treatment changes that can impact how results are interpreted (Gogtay et al., 2021).

Beyond clinical trials, estimands serve a broader role in quantitative research by connecting statistical evidence to theoretical concepts, improving clarity in study design, and guiding appropriate analytical methods (Kahan et al., 2023; Lundberg et al., 2021). Their adoption in fields such as economics, epidemiology, and social sciences has helped refine how researchers define and interpret causal relationships (Lundberg et al., 2020). While estimands have been widely used in medical and regulatory research, their relevance extends to sports analytics, where precise definitions of player performance metrics are essential.

Different multiple interacting factors, such as pitch type, speed, and movement affect the likelihood that a pitch resulting in a strike outcome. The use of estimands in this research ensures that the analysis remains consistent, interpretable, and theoretically grounded, such that it provides a structured way to measure how different pitch characteristics influence strike probability. These methods connect both statistical modelling and everyday sports strategy, making it easier for coaches and pitchers with actionable insights for game planning. To implement estimands in this study, we follow a three-step framework adapted from Lundberg et al. (2021):

a. Setting the theoretical estimand: Defining what we aim to measure (the impact of batter characteristics and pitch characteristics on strike likelihood).

b. Linking it to an empirical estimand: Quantifying this relationship through statistical modeling (Bayesian logistic regression).

c. Learning from data: Using posterior distributions to estimate and interpret strike probabilities under different pitch conditions and interactions.

This process provides a clear, reproducible, and probabilistic interpretation of how pitch characteristics influence strike likelihood, helping coaches and teams make more informed decisions about pitch selection and in-game strategy

Key Estimands in This Study

In this study, the two main estimands are Posterior Distributions of Model Coefficients (β_i) and Probability of a Strike Likelihood ($P(Y=1|X)$). These estimands are very useful in understanding how different pitch characteristics influence the likelihood of a strike outcome and provide a probabilistic framework to quantify these relationships.

Posterior Distributions of Model Coefficients (β_i)

The posterior distributions of the model coefficients provide insights into how each predictor variable (X_1, X_2, X_3, X_4, X_5, X_6…. n) affects strike probability. Unlike point estimates obtained through traditional maximum likelihood estimation, Bayesian logistic regression incorporates uncertainty into the estimation process, yielding full probability distributions for each coefficient rather than single fixed values (Erp & Gelder, 2013). The estimated coefficients β_0, β_1, β_2, β_3, β_4, β_5…… n describe the strength and direction of each predictor variable's impact on strike likelihood. A wider posterior distribution suggests greater uncertainty about a variable's influence, whereas a narrower distribution indicates stronger confidence in its effect. This method allows for a probabilistic interpretation of effect sizes, helping us not only determine whether a factor matters but also assess the certainty of its impact (Pham & Pham, 2021).

Bayesian estimation methods, such as Markov Chain Monte Carlo (MCMC) sampling, are widely used to approximate posterior distributions in logistic regression models (Emenyonu & Adam, 2019). Compared to classical methods, Bayesian inference provides greater flexibility, particularly in scenarios with small sample sizes or cases involving separation in logistic regression (Gordóvil-Merino et al., 2012). The posterior means of coefficients depend on both the probability function and the choice of prior distribution, with certain priors leading to more stable estimates (Pham & Pham, 2021).

Because Bayesian logistic regression directly models the probability of strike likeloihood, it provides a way to assess how different pitch characteristics influence decision-making while also capturing the uncertainty inherent in the data (Lukman et al., 2021). Through the posterior distributions, we get a better insight of how factors like batter handedness, pitch dynamics and its interaction with strike zone can impact the probability of a pitch resulting in a strike.

Posterior Distribution Formula

Combining these components using Bayes' theorem gives us the posterior distribution:

$$P(\beta \mid Y, X) \propto \left[\prod_{i=1}^{N}\left(\frac{1}{1+\exp(-X_i\beta)}\right)^{Y_i}\left(1-\frac{1}{1+\exp(-X_i\beta)}\right)^{1-Y_i}\right]P(\beta)$$

Mathematical Representation

The mathematical function for the posterior distribution for normal can be written as:

$$P(\beta \mid Y, X) \propto \left[\prod_{i=1}^{N}\left(\frac{1}{1+\exp(-X_i\beta)}\right)^{Y_i}\left(1-\frac{1}{1+\exp(-X_i\beta)}\right)^{1-Y_i}\right]\cdot\prod_{j=0}^{p}\mathrm{Normal}(\beta_j \mid 0,10)$$

where:

- β is the vector of coefficients (including the intercept).
- Xi is the vector of predictor variables for observation iii.
- Yi is the observed outcome for observation iii.
- N is the total number of observations.
- P is the number of predictor variables.

Where the product of the individual likelihoods for each observation $P(Yi \mid Xi, \beta)$ represents the combined likelihood of observing the entire dataset given the parameters. Prior is the product of the normal priors $P(\beta j)$, for each coefficient represents the prior beliefs about the parameters and the posterior distribution $P(\beta \mid Y, X)$ combines the prior beliefs and the likelihood of the observed data to update our beliefs about the parameters after observing the data.

S*trike Probability (P(Y=1|X))*

The second estimand, $P(Y=1|X)$, gives the probability that a pitch result in a strike under given conditions. It is computed using the logistic function, which transforms a sum of inputs into a value between 0 and 1. The probability is based on the posterior distributions from the model. These show how the inputs are linked to strike calls (Lukman et al., 2021). Strike probability estimations allow for a probabilistic rather than deterministic interpretation of umpire decision-making, acknowledging the inherent uncertainty in how different pitch features are judged. The use of Bayesian posterior distributions allows us to see the full range of possible strike probabilities instead of just a single predicted value. This gives us a clearer picture of the different factors that impact strike probability (Shin, 2015). This method is especially useful for making strategic adjustments in baseball analytics, helping coaches and analysts estimate how likely a pitch is result in a strike outcome in different situations, like facing specific pitch types or speeds.

Bayesian logistic regression is very useful and recognizable for its ability to incorporate uncertainty into probability estimations, making it particularly useful for modeling decision-making under variable conditions (Zhan et al., 2023). This is a key advantage over frequentist approaches, which provide point estimates without accounting for uncertainty in model parameters (Gray & Ferson, 2021). Estimands are used here to link the statistical analysis to the main questions about strike likelihood. The use of posterior distributions of model coefficients and strike probability estimates, creates a probabilistic framework to analyze umpire decision-making. This approach makes the findings to be precise and applicable to real-world game strategy, scouting, and pitcher development.

For logistic regression, the likelihood for a single observation is given by:

$$P(Y_i = 1 \mid X_i, \beta) = \frac{1}{1 + \exp(-X_i\beta)}$$

$$P(Y\_i = 0 \mid X\_i, \beta) = 1 - P(Y\_i = 1 \mid X\_i, \beta$$

For the entire dataset: $P(Y \mid \beta, X) = \prod_{i=1}^{N} \left(\frac{1}{1+\exp(-X_i\beta)}\right)^{Y_i} \left(1 - \frac{1}{1+\exp(-X_i\beta)}\right)^{1-Y_i}$

Step III: Designing the Estimator

The process of using the Bayesian logistic regression to estimate model coefficients, with priors for the coefficients βi is known as designing the estimator. The prior for the coefficients βi is set as Normal (0,10) based on weakly informative where (i=0,1,2,3,4,5,6,7,8....... n). In this study, the goal of designing an estimator in this study is to accurately capture how different pitch features and context affect whether a pitch results in a strike. An estimator is used to learn about a value we can't observe directly, based on the data we have. It helps us make educated guesses about relationships within our data, providing estimates that can be used for prediction and analysis.

An estimator works by applying a specific mathematical formula or model to the data, which transforms raw observations into meaningful statistical measures. In this case, we employ Bayesian logistic regression, a method well-suited for binary outcomes, such as whether a pitch is a strikes outcome (1) or not (0), and known for its flexibility in incorporating uncertainty through probability distributions.

The Bayesian logistic regression allows us to incorporate both prior knowledge and observed data to improve model accuracy (Casella & Berger, 2005). In baseball, decision-making is inherently probabilistic. Batters often rely on both prior experiences and real-time pitch information to make split-second decisions. Interestingly, this mirrors Bayesian thinking such that prior beliefs are updated as new data (pitch characteristics) become available (Brantley & Kording, 2024).

The Bayesian Framework: Priors, Likelihood, and Posterior

Bayesian logistic regression revolves around a simple, yet powerful idea called Bayes' theorem: we start with what we know (or assume) about the data, update this understanding based on new observations, and end up with a refined picture of what's likely true. The mathematical formulation includes three key components:

Prior Distribution ($P(\beta)$), Likelihood Function ($P(Y|\beta, X)$), and Posterior Distribution ($P(\beta|Y, X)$)

Why Normal (0,10) for prior distribution?

- Centering Around Zero (Mean = 0): This reflects a neutral assumption, which shows that no strong initial belief that any predictor has a positive or negative effect. It allows the data to speak for itself without biasing the results.
- Large Spread (Standard Deviation = 10): A standard deviation of 10 provides flexibility that covers a wide range of plausible values that prevents extreme estimates unless strongly supported by the data, offering numerical stability without overly constraining the model.
- Expert Knowledge: The Normal (0,10) prior we use is based on expert insights suggesting that, without strong prior evidence about the effect sizes, it's reasonable to assume that the coefficients are centered around zero. This reflects a neutral stance in a way we don't initially favor either a positive or negative effect. At the same time, the relatively large standard deviation of 10 provides the necessary flexibility by allowing a wide range of plausible values. In other words, this choice is informed by expert recommendations in the literature by Gelman et al., 2008; van Zwet, 2018, which suggests that a Normal (0,10)

prior is a robust, weakly informative option. It helps prevent the model from being overly constrained while still providing some regularization to stabilize the estimates, particularly when working with complex or noisy data such as that in baseball strike likelihood (Chen et al., 1999; Kuhnert et al., 2010).

Priors Relevance in Baseball

Applying Bayesian logistic regression in baseball is not just a theoretical exercise. Brantley and Kording (2024) demonstrated that professional batters integrate prior beliefs with real-time pitch information, which is essentially performing Bayesian updates subconsciously. Our estimator mirrors this cognitive process statistically, helping coaches and analysts understand the probability of strike likelihood under varying pitch dynamics conditions.

Moreover, Bayesian estimators are particularly useful in sports analytics because they handle small sample sizes and noisy data effectively. Gordòvil-Merino et al. (2012) showed that weakly informative priors stabilize estimates in small datasets, which is often the case when analyzing specific batter-pitcher matchups.

The estimator was defined using Bayesian logistic regression. This allowed prior information to be combined with observed data. Uncertainty in the model's parameters was accounted for. The setup reflects how decisions are made during a game. Priors were chosen and Bayesian methods applied to examine how certain variables relate to strike calls, rather than only predicting outcomes.

Step IV: Testing the Estimator

After the designation of the estimator, the next thing is to evaluate its performance to ensure it captures the relationship between the batter handedness and its interaction with the pitch characteristics and strike likelihood accurately. This phase involves generating synthetic data, fitting the Bayesian logistic regression model, and assessing whether the model can recover the true parameter values used in the simulation.

Why Test the Estimator?

Testing the estimator helps verify its robustness before applying it to real-world baseball data. It's not just about confirming the model works in theory, it's about seeing if it can deliver reliable results when put to the test. This step matters a lot in Bayesian analysis because the way priors and data interact can really shape the final outcomes (Gordóvil-Merino et al., 2012). The process includes:

a. Generating Synthetic Data: Synthetic data is created using the generative model established in earlier steps and this phase involves:
b. Assigning Realistic Coefficients (β_i): We assign values to the model coefficients that reflect plausible effects of pitch characteristics on strike likelihood.
c. Simulating Predictor Variables (X_1, X_2, X_3, X_4, X_5, X_6): These variables are generated based on expected ranges or distributions commonly observed in baseball data (e.g., pitch speed, type, movement, etc.).
d. Generating Binary Strike Likelihood (Y): Probabilities of strike outcomes are calculated using the logistic function and the strike outcomes are then drawn from these probabilities using a Bernoulli distribution. This approach allows us to create a dataset with known underlying relationships, serving as a benchmark for model evaluation.

e. Fitting the Bayesian Logistic Regression Model: With the synthetic data prepared, we fit the Bayesian logistic regression mode and applied it to the synthetic dataset and also set the priors to weakly informative. This step reflects what happens in real data analysis, making sure the estimator performs in simulations the same way it would in real-world situations.

f. Evaluating Estimator Performance: After fitting the model, we assess its ability to recover the true values of the coefficients by comparing the estimated posterior distributions of β_i with the true values used when generating the data. To assess accuracy and precision, we check if the credible intervals capture the true values and see how closely the posterior means match them.

Past simulation work has found that Bayesian methods tend to give more stable estimates when sample sizes are small (Gordóvil-Merino et al., 2012; Merino et al., 2012). These studies also point to some limits, including how results can change with skewed data or depending on the prior. Grounding the model in both theory and earlier work helps support its use in this setting.

Step V: Analyze Sample and Summarize

After validating the estimator, the next important step in the methodology is to apply the model to the actual dataset and interpret the findings. This is done through the use of Markov Chain Monte Carlo (MCMC) to sample from the posterior distributions of the coefficients of the model, allowing for detailed estimation and insights underlying how batter handedness and its interaction with different pitch characteristics influence strike likelihood. Below are the processes involved:

a. Fitting the Bayesian Logistic Regression Model: The Bayesian logistic regression model is fitted to the real dataset, employing MCMC techniques (e.g., using PyMC3) samples were taken from the posterior distributions of the parameter of the model called coefficients (β_i). MCMC is particularly effective in Bayesian inference as it generates samples that represent the posterior distribution, offering a detailed view of parameter uncertainty (Chen, 2024).
b. Sampling from Posterior Distributions: The use of the MCMC algorithms in sampling from posterior distributions provides:

- Posterior Means: These represent the average effect of each predictor on strike likelihood.
- 95% Credible Intervals: This gives an interval that shows where the parameter values tend to land, based on the data and the model, providing a more intuitive measure of uncertainty compared to traditional confidence intervals (Hespanhol et al., 2019).
- Summary Statistics: Other metrics like the mean contrasts are also calculated. The contrast means helps in assessing both the strength and direction of each predictor's impact on strike likelihood.

c. Evaluating Model Performance: For the Bayesian estimates in this analysis to be reliable, we need to check for the following:

- Convergence Diagnostics: In Bayesian analysis, getting the model to converge is more important than just having a large sample size. We use tools like trace plots, Gelman-Rubin statistics (R-hat), and Effective Sample Size (ESS) to check if the MCMC chains have stabilized (Bates & Campbell, 2001). ESS helps gauge model reliability by showing how many independent samples were effectively used after accounting for autocorrelation. A high ESS (over 1000) means the estimates are solid and have good convergence, while a low ESS (under 200) could signal issues. Ideally, a well-converged model hits around 2,000–5,000 ESS, which boosts the precision and stability of parameter estimates (Zitzmann & Hecht, 2019).
- Sensitivity Analysis: This step tests how robust the model is by tweaking prior settings or adjusting data variations. It helps ensure the results aren't overly sensitive to specific assumptions (Zuur et al., 2002).
- Comparison to True Values: When possible, we compare model outputs to known benchmarks or true values to see how well the model performs in real scenarios.

d. Contrast Analysis and Effect Estimation

Beyond standard regression results, contrast analysis helps test specific hypotheses, like whether pitch type or batter handedness affects strike probability. Bayesian methods shine here because they provide credible intervals instead of p-values, giving a clearer picture of effect sizes and uncertainties (Wiens & Nilsson, 2016).

When dealing with complex interactions, like pitch location combined with batter shandedness, we use bias-corrected bootstrap methods. These methods help assess indirect effects more accurately, reducing bias and increasing the power to detect real relationships (Williams & Mackinnon, 2008).

Contrast Mean and Interpretation: The contrast means show the average difference between groups, such as comparing strike probabilities for Left handedness vs. Right handedness. A positive contrast means the first group has a higher effect, while a negative contrast suggests the opposite. If the contrast means close to zero, it means there's not much difference between the groups (Wiens & Nilsson, 2016).

Dealing with Values Close to 0, 1, or -1:

- Near 0: Suggests a minimal effect.
- Near 1 (or -1): Indicates a strong positive (or negative) effect, especially in models predicting probabilities.
- Exactly 0 or 1: Rare in Bayesian models; if these pop up, it could point to model issues or overly strict constraints (Magee et al., 2021).

e. High-Density Intervals (HDI) and Credible Intervals.

High-Density Intervals (HDIs) are key in Bayesian analysis. They show the range where the most credible values for a parameter lie. Unlike traditional confidence intervals, HDIs directly tell you the probability that the true value falls within that range (Hespanhol et al., 2019). The HDI Calculation is performed using Monte Carlo simulations, HDIs help measure uncertainty around effect sizes (Le et al., 2018). In Interpreting HDIs, a 95% HDI means there's a 95% chance the true value falls within that interval. If the HDI excludes zero, it suggests a credible effect. A narrow HDI indicates high precision, while a wide HDI points to more uncertainty (Hespanhol et al., 2019). In sports analytics, HDIs help make sense of player performance and strategic decisions by offering clear, data-backed insights.

Other Key Metrics for Bayesian Interpretation

Beta Coefficients (β): These show the effect size and direction for each predictor. A positive β means as the predictor increases, so does the strike probability. A negative β shows the opposite. If β is close to 0, it implies little to no effect, while larger values (positive or negative) indicate stronger effects (Hespanhol et al., 2019).

By applying Bayesian logistic regression to real data and using MCMC techniques, we achieve a robust parameter estimation by capturing both the average effects and uncertainties of predictors and make clear interpretation using credible intervals and HDIs to present findings in an easy way to understand. It also helps in identifying key factors that influence strike likelihood, thereby helping in coaching strategies and player development. This approach makes the analysis not only to be statistically analyzed but also meaningful in a way it turns complex data into practical insights for baseball analytics.

Results **Chapter 4**

This section reports the results by examining the factors linked to whether a pitch is called a strike. The analysis starts with descriptive statistics, then moves to correlation checks to look at basic relationships. From there, panel regression and Bayesian comparisons between selected pitch types are used to explore differences in more detail.

These methods are used to examine how pitch traits and location relate to strike calls. The analysis moves from basic associations toward identifying possible causal effects on pitch outcomes. This step-by-step progression is used to support claims about causality and to show how pitch features and location relate to pitch resulting in strikes.

Descriptive Statistics

The dataset comprises 4,593 observations detailing pitch characteristics, batter-pitcher interactions, and the outcome of each pitch. The Strike variable records how many strikes were on the batter at the time of the pitch, with 0, 1, or 2 strikes observed. The distribution of strike counts is relatively balanced, with 39.97% of pitches occurring when there were 0 strikes, 29.65% at 1 strike, and 30.37% at 2 strikes. For the Strike Outcome variable, among the 4,593 observed pitches, 65.86% did not result in strikes, while 34.14% did *(see Table 1)*. The distribution is reasonably balanced for a binary model, as both have a sufficient proportion of strike and non-strike observations to support model inference. This framing captures all pitcher-favorable strike possible events which align with the study's pitcher-centered focus on overall strike effectiveness.

Regarding pitch types, the dataset includes a diverse distribution of pitch categories (see Table 1). The Four-Seam Fastball was the most frequent pitch (n = 1,731), followed by Slider (n = 870), Sinker (n = 758), Changeup (n = 546), Curveball (n = 345), Cutter (n = 316), and Sweeper (n = 27). Strike rates varied by pitch type: Four-Seam Fastballs resulted in strikes on 35.6% of occasions, while Sliders and Sinkers produced strikes at slightly lower rates of 36% and 33.1%, respectively. Less commonly thrown pitches such as the Sweeper still maintained reasonable representation in the dataset of 40.7%. This distribution across pitch types offers critical variation for analyzing how pitch selection influences strike outcomes

Several continuous variables describe pitch dynamics with the average pitch speed at release is 89.47 mph (SD = 5.88), ranging from 65.1 mph to 102.2 mph, which shows a wide variation across pitch types and pitchers. Horizontal pitch movement (pfx_x) averages to 0.14 inches, while vertical movement (pfx_z) also averages to 0.71 inches, which indicates slight lateral and upward deviations from a gravity-only trajectory.

The horizontal pitch location across the plate (plate_x) has a mean of 0.03 feet (SD = 0.85), suggesting that pitches are generally centered but with broad lateral variability (–3.28 to 3.04 ft). Vertical location (plate_z) averages 2.31 feet above the ground and ranging from –1.37 to 5.74 feet. Also, the vertical strike zone dimensions vary by batter, with sz_top (top of strike zone) averaging 3.39 ft and sz_bot (bottom) with a mean of 1.59 ft.

All these variables offer key insight into the spatial and dynamic characteristics of pitches that may influence strike outcomes. The variation in speed, location, and movement provides a good basis for modeling the probability that a given pitch will result in a strike.

Table 1 *Descriptive Statistics - Continuous*

Variables	*Observations*	*Mean*	*SD*	*Minimum*	*Maximum*
Speed (*mph*)	4593	89.455	5.880	65.100	102.200
pfx_x (*in*)	4593	-0.139	0.861	-2.0400	1.920
pfx_z (*in*)	4593	0.711	0.717	-1.890	2.260
Plate_x (*ft*)	4593	0.033	0.847	-3.280	3.040
Plate_z (*ft*)	4593	2.312	0.980	-1.370	5.740
sz_top (*ft*)	4593	3.387	0.178	2.640	4.090
sz_bot (*ft*)	4593	1.586	0.103	0.970	2.010

Categorical; Frequency of Strike Count

Variables		*Observations*	*Percentage (%)*
Strikes Count	0	1836	39.97
	1	1362	29.65
	2	1395	30.37

Categorical: Frequency of Strike Outcome

Variables		*Observations*	*Percentage (%)*
Strikes Outcome	0	3025	65.86
	1	1568	34.14

Strike = 1 | No Strike = 0

Pitch Type Frequency

Pitch Types	*Observation*	*Strike*	*No strike*
Four-Seam Fastball	1731	617	1114
Slider	870	313	557
Sinker	758	251	507
Changeup	546	162	384
Curveball	345	113	232
Cutter	316	101	215
Sweeper	27	11	16

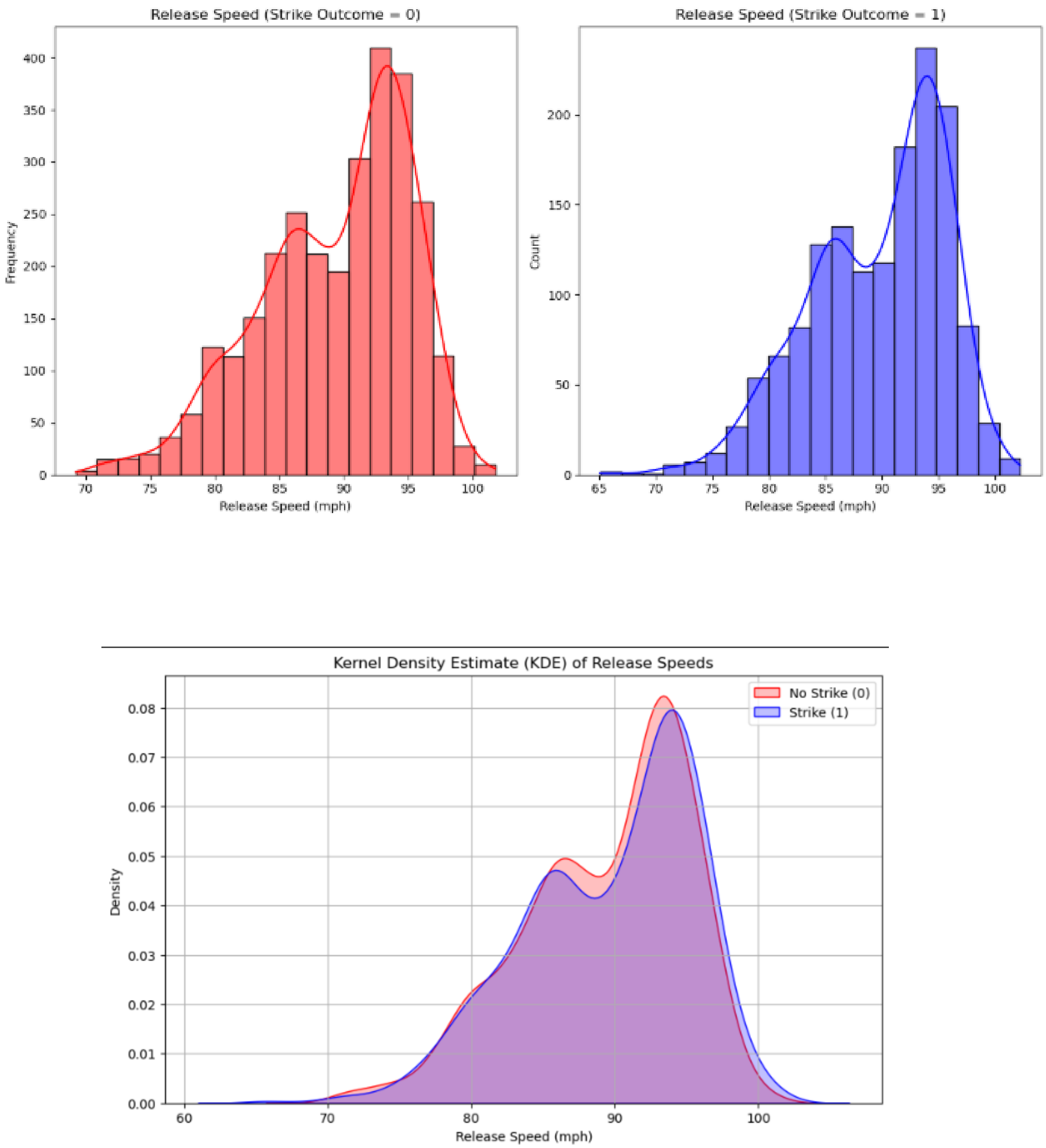
Exploring Release Speed and Strike Likelihood

The descriptive analysis shows that not all pitches are thrown at the same speed and this indicates the presence of different pitch types and throwing styles in the data. Release speed is important for pitch identification, for example it separates fastballs from breaking balls and off-speed pitches. More importantly, it plays a central role in shaping how pitches are perceived by both batters and umpires. Because release speed helps signal the intent and deception behind a pitch, understanding its relationship with strike outcomes provides pitchers with insights into how their velocity affects game control.

In *Figure 1*, the graph on the right represents pitches that resulted in strikes showing a higher concentration around the 92–95 mph range. By contrast, non-strike pitches (left side) are more dispersed and skewed toward slower velocities and this suggests that pitches with higher release speed may have a slight edge in producing strike outcomes. The kernel density estimates (KDEs) then support this with the distribution of strike outcomes peaks more sharply near **93 mph**. The overlap between the KDEs for strikes and non-strike pitches explains that release speed alone doesn't fully determine the outcome. This shows an association between the two-variable playing a moderating role when combined with other pitch features such as vertical drop, horizontal break, or placement near the edges of the strike zone, but not the sole determinant of strike likelihood.

From a strategic standpoint, pitchers who throw harder with command get better results, not just from speed, but by disrupting timing. Fast pitches cut reaction time, and when paired with off-speed, they squeeze more strikes, even on close calls.

Figure 1 *Release Speed and Strike Likelihood*



Panel Logistic Regression Results

To further examine the relationship between pitch characteristics and strike likelihood, a panel regression model was conducted using standardization and squared terms for plate_x and plate_z to capture non-linearities, standardized continuous predictors, and categorical pitch type dummies. This method accounts for the batter-level differences by including fixed effects, which allows the model to isolate the effects of pitch dynamics from individual batting tendencies.

The regression results (*Table 2*) show that horizontal pitch location is part of the determinant of strike outcomes. Specifically, the squared horizontal coordinate ($plate\_x^2$) has a significant negative coefficient (–0.1214, $p < 0.000$), which indicate that pitches placed farther from the center of the plate are less likely to result in strikes, whether swinging or taken. This supports the idea that pitching toward the edges increases risk, particularly if the pitch misses the optimal zone.

The standardized coefficient ($\beta = 0.0180$, $p = 0.1892$) for speed was positive but not significant, showing that while faster pitches may slightly increase the odds of a strike, speed alone doesn't offer a consistent advantage. Vertical location and strike zone boundaries were not statistically significant, suggesting pitchers may face less predictable strike outcomes when targeting height or low-zone edges. When evaluating pitch types, Sliders (SL) significantly increase the likelihood of achieving a strike ($\beta = 0.0947$, $p < 0.002$), while Curveballs (CU) approach significance ($\beta = 0.0611$, $p = 0.0807$). Other pitch types did not show statistically significant effects. Although the model's within R^2 is modest (0.070), the results provide great insights for pitchers, placement and pitch type selection matter, particularly when aiming to generate strike outcomes.

Table 2 *Panel Logistic Regression Results (standardizied)*

Variables	*Coefficients*	*Std Error*	*T-Statistics*	*P- Value*	*[0.03*	*0.97]*
const	0.3217	0.2350	1.3690	0.1711	-0.1390	0.7825
release_speed	0.0180	0.0137	1.3132	0.1892	-0.0089	0.0448
plate_z ²	-0.0026	0.0017	-1.5273	0.1268	-0.0059	0.0007
plate_x²	-0.1214	0.0072	-16.815	0.0000	-0.1356	-0.1072
sz_top	0.1355	0.0883	1.5343	0.1250	-0.0377	0.3087
sz_bot	-0.2371	0.1344	-1.7641	0.0778	-0.5007	0.0264
pitch_type_CU	0.0611	0.0350	1.7472	0.0807	-0.0075	0.1297
pitch_type_FC	0.0261	0.0343	0.7595	0.4476	-0.0412	0.0934
pitch_type_FF	0.0323	0.0320	1.0076	03136	-0.0305	0.0951
pitch_type_SI	0.0174	0.0325	0.5375	0.5310	-0.0462	0.0811
pitch_type_SL	0.0947	0.0255	3.7135	0.0002	0.0447	0.1447
pitch_type_ST	0.1342	0.0930	1.4421	0.1494	-0.0482	0.3166

F-Statistics=27.43, p <0.000|P- value<0.005| Predictors except pitch type were standardized

pitch_type_CU: **Curveball**, pitch_type_FC: **Cutter**, pitch_type_FF: **Four-Seam Fastball**, pitch_type_SI : **Sinker**, pitch_type_SL: **Slider**, pitch_type_ST: **Sweeper**

Bayesian Analysis Results

Building on the relationships and insights from the panel logistic regression, a Bayesian analysis was used to understand more about the potential causal effects of pitch characteristics on strike outcomes. Through this method, it is easier to understand how pitch dynamics truly impact the likelihood of a pitch resulting in a strike, helping to move closer to identifying more reliable causal relationships. The DAG was drawn to show the causal relationship and pathways as seen in *Figure 2*.

Posterior Distributions of Coefficients

The Bayesian model yielded posterior distributions for the fixed effects coefficients associated with pitch dynamics and batter-pitcher interactions. These coefficients represent the expected effect of each predictor variable on the likelihood of a pitch resulting in a strike, such that it accounts for the uncertainty in the estimates.

Plate Location Effects on Strike Outcome

From *Figure 3*, the coefficient for plate_z (vertical pitch location) is 0.33 (95% of Highest Density Intervals: 0.075 to 0.58), indicating that higher pitches are significantly more likely to be strikes outcome. This shows the strong influence of vertical location on strike outcomes.

In contrast, the coefficient for plate_x (horizontal location) is -0.20 (95% Highest Density Intervals (HDI): -0.49 to 0.092), suggesting that pitches on the outer edges, particularly the outside corner are slightly less likely to be strikes outcomes. However, the HDI includes zero, indicating some uncertainty in this effect. Also, the top of the strike zone (sz_top) has a positive coefficient (0.58, 95% HDI: 0.13 to 1.0), indicating that pitches near the upper boundary are more likely to result in strikes. Meanwhile, the bottom of the zone (sz_bot) shows a weak negative effect (-0.25, 95% HDI: -0.66 to 0.15), implying that lower pitches may be slightly disfavored, though this effect is less pronounced than for plate_z.

Pitch Speed

The coefficient for release_speed is 0.29 (95% HDI: -0.16 to 0.75; *Figure 3*), suggesting faster pitches may be slightly more likely to result in strikes. However, this effect is both minimal compared to location variables and statistically uncertain (as evidenced by the HDI including zero). The results demonstrate that while pitch location strongly affects strike outcomes, pitch speed shows only marginal influence.

Contrast Between Higher and Lower plate_z on Strike Likelihood

While both pitch speed and location show meaningful relationships with strike outcomes, the most pronounced effect emerges when examining vertical location (*plate_z*) contrasts. The Bayesian analysis reveals a striking difference: the mean contrast of 0.64 (95% HDI: 0.125 to 1.166, *excluding zero*; *Figure 4*) indicates not only that higher pitches are substantially more likely to result in strikes and this effect is *statistically significant*. With the HDI entirely above zero, we can be highly confident in this directional advantage for high-zone pitches. This robust effect quantifies what the coefficient analysis suggested about vertical location dominating strike outcomes, with high pitches having a systematic advantage.

Figure 2 *Directed Acyclic Graph*

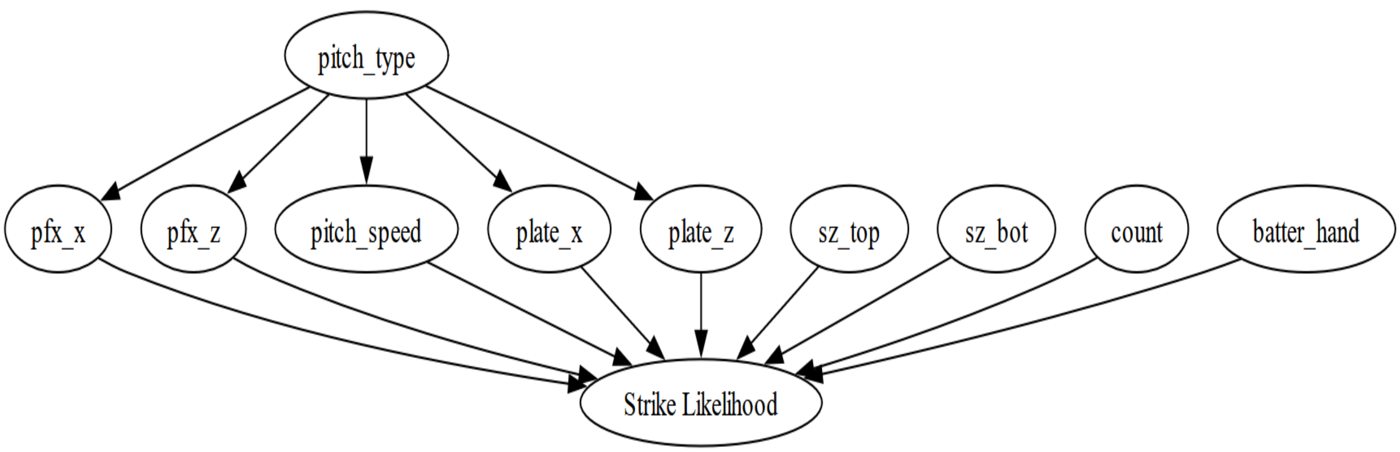


Figure 3 *Posterior Distributions of Coefficients of Pitch characteristics on Strike Likelihood*

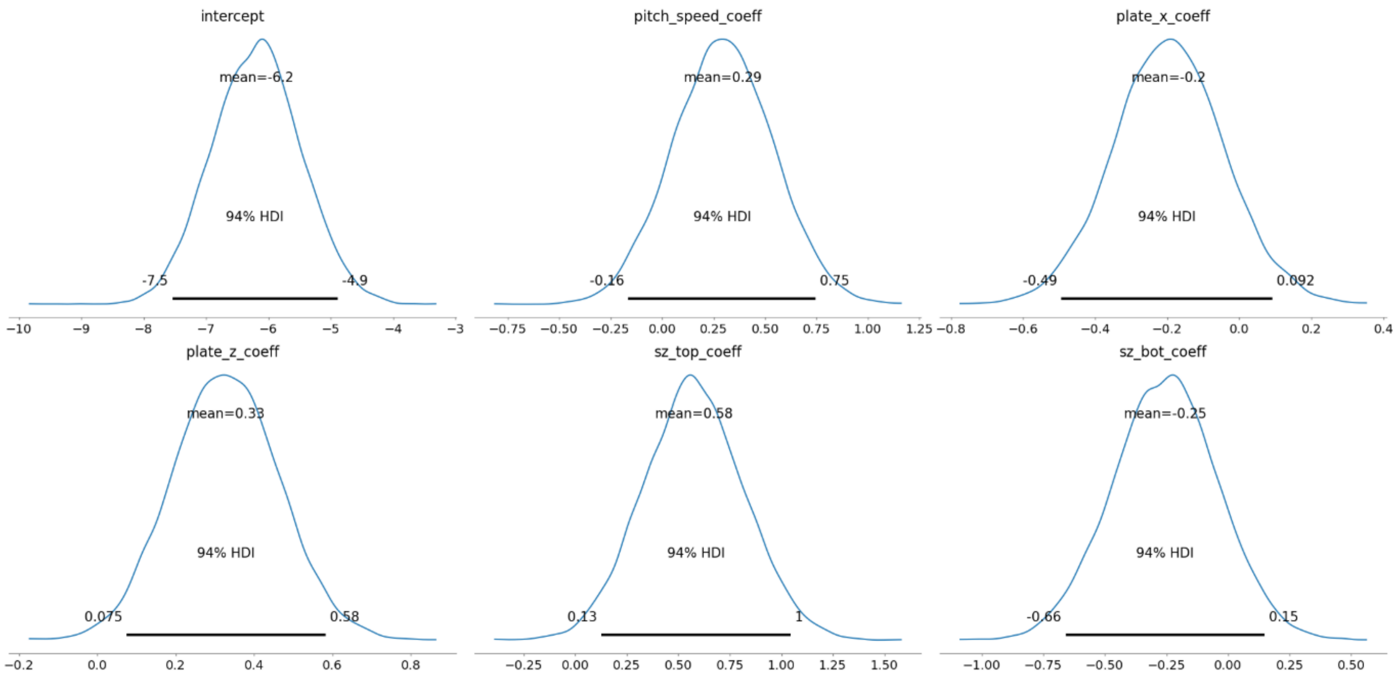
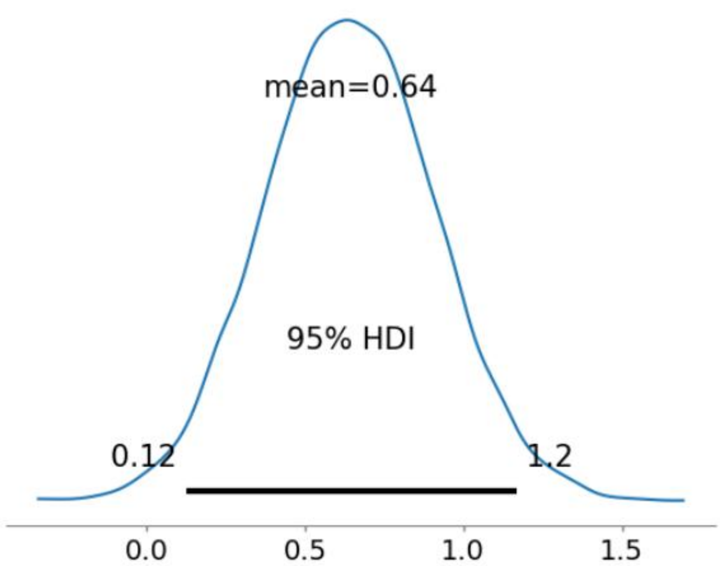


Figure 4 *Posterior Distribution of Contrast Between Higher and Lower Plate_Z on Strike Likelihood*



Pitch Type Impact on Strike Probability

Figure 6 shows the posterior distributions of pitch type coefficients, which shows how each pitch influences strike likelihood. While mean coefficients indicate directionality, the HDIs reflect uncertainty (not significant) in these estimates but still insightful.

Changeups (-2.2 mean, HDI [-9.1,4.3]) and Sinkers (-1.31, [-8.2,5.3]) typically reduce strike outcomes due to their deceptive movement, though their wide HDIs indicate inconsistent results. Forkballs (-2.3, [-9,4.6]) show the strongest negative effects, as their unpredictable breaks often miss the zone.

Fastballs (-0.51, [-7.4,6]) and Cutters (-0.65, [-7,7.6]) demonstrate neutral effects, serving as reliable baseline pitches. Curveballs (0.21, [-6.4,6.9]) show slight positive potential, while Sliders (2.5, [-5.1,11]) show that a pitch is likely to increase the likelihood of a strike, this may be due to the fact that a well-placed Slider often breaks into the strike zone, making it more likely to be a strike outcome through sharp lateral movement.

The Sweeper (2.8, [–4.9, 11.0]) had the highest estimated strike rate among the pitches analyzed. Its movement tends to stay near the zone late in the pitch. These results point to a tradeoff between pitches that are more likely to be called strikes (like Sliders and Sweepers) and those designed to draw swings outside the zone (like Forkballs).

The analysis suggests that the Slider (SL) and Sweeper (ST) are most likely to increase strike likelihood, while the Changeup (CH), Sinker (SI), and Forkball (FS), are more likely to decrease it, with the Fastball (FF) and Cutter (FC) showing neutral effects.

Contrast Analysis Between Slider (SL) and Sweeper (ST)

After analyzing the individual effects of Slider (SL) and Sweeper (ST) on strike likelihood, both pitch types have shown a positive relationship with strike outcome, but it is useful to compare how these two pitch types differ in terms of their likelihood to results in a strike. The contrast analysis yields a mean difference of -0.004 (HDI: [-1.085, 1.149]) as illustrated in *Table 3* and *Figure 5*, indicating no meaningful advantage for either pitch type in generating strikes.

The moderate standard deviation (0.579) and wide HDI reflect substantial uncertainty in their relative performance. This suggests contextual factors like pitcher skill or batter tendencies may influence outcomes more than inherent pitch characteristics. The analysis demonstrates strong computational reliability, with minimal Monte Carlo error (MCSE = 0.006) and high Effective Sample Sizes (bulk ESS = 8495; tail ESS = 7282). While both pitches increase strike probability similarly, their equivalent performance leaves room for situational selection based on other strategic considerations.

Figure 5 *Overall Average Contrast: Slider vs Sweeper*

.

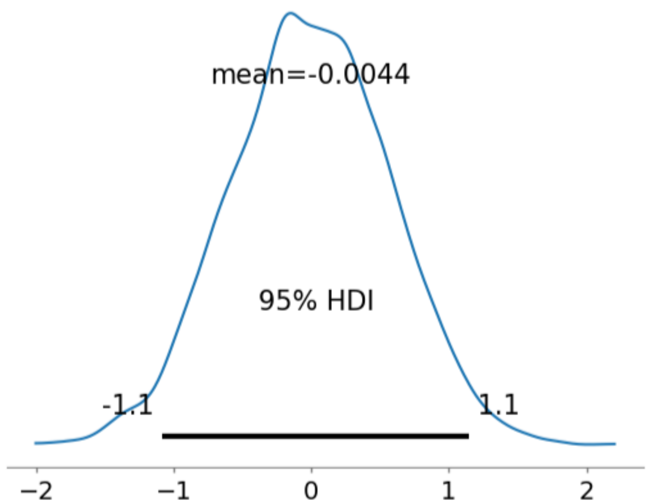
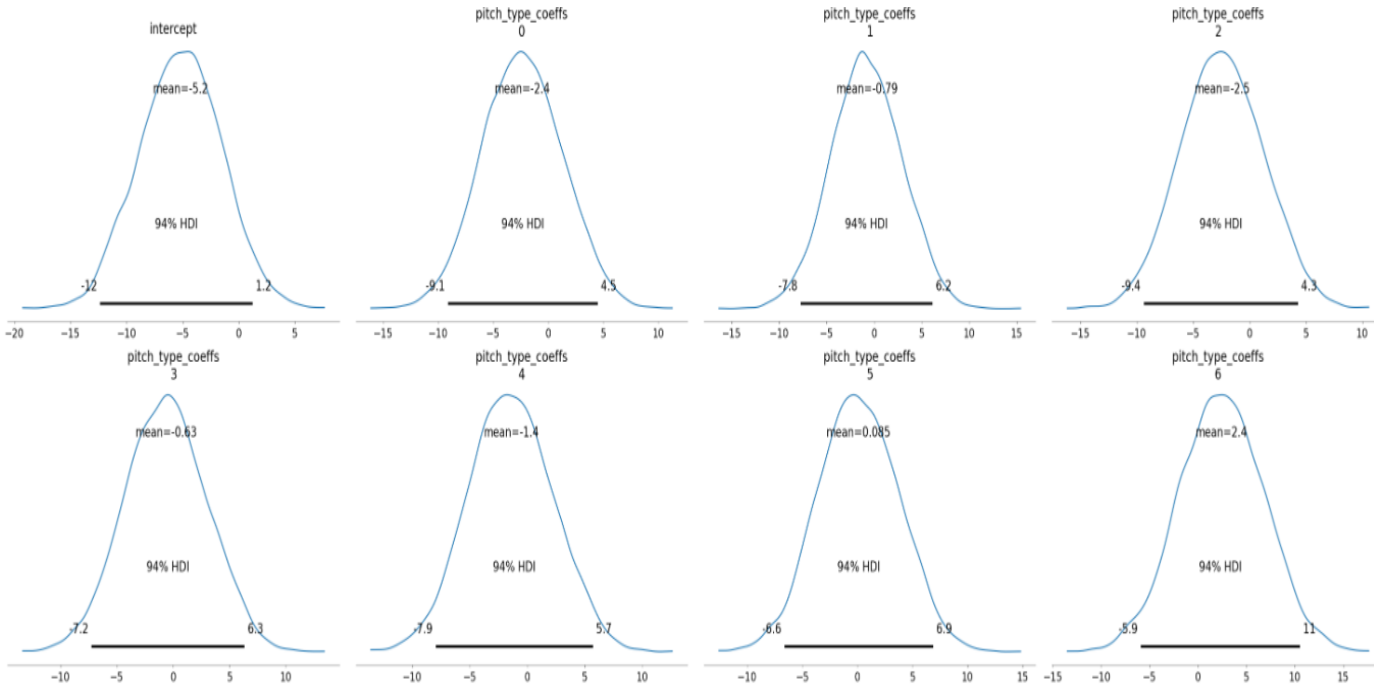


Table 3 *Overall Average Contrast Metrics for Slider vs Sweeper*

	Mean	*SD*	*HDI [0.03]*	*HDI [0.97]*	*Mcse mean*	*Mcse sd*	*Ess bulk*	*Ess tail*
Contrast_SL_vs_ST_avg	0.566	0.808	-0.920	2.241	0.007	0.007	15206	7660

Figure 6 *Posterior Distributions of Coefficients of Pitch Types on Strike Likelihood*



Pitch Type Encoding Mapping

0: CH (Changeup), 1: CU (Curveball), 2: FC (Four-Seam Fastball), 3: FF (Fastball), 4: FS (Fastball-Sinker Hybrid), 5: KC (Knuckle-Curve), 6: SI (Sinker), 7: SL (Slider), 8: ST (Sweeper)

Effect of Batter Handedness on Strike Likelihood

The analysis of batter handedness (denoted as ‘stand’) reveals an interesting relationship between a batter's side of the plate and strike likelihood. The coefficient for the batter's side of the plate has a positive mean of 0.866 as shown in the results in *Table 4* and *Figure 7*, which means pitches against batters batting right-handed are more likely to result in strikes on average compared to hitters batting left-handed, therefore pitchers have a slight advantage when facing right-handed batters. However, this finding comes with important caveats.

The sample size shows some class imbalance with the dataset having more right-handed batters (n = 2,763) than left-handed batters (n = 1,830) which might be driving the observed effect. The moderate standard deviation (0.32) further confirms that while there's a general trend, individual batter skills and specific pitch locations probably play bigger roles than the batter’s side of the plate alone.

These findings are reliable because the model shows excellent convergence (R-hat = 1.0) and stability (MCSE ≤ 0.005). With substantial effective sample sizes (ESS > 4000), the analysis has enough data to reliably detect these patterns, as detailed in *Table 4*.

In practical terms, while umpires might call strikes slightly more often on right-handed batters overall, coaches shouldn't overemphasize this factor. The class imbalance between the them suggests platoon advantages may depend more on pitcher-batter matchups and pitch selection than simple batter's side of the plate.

Contrast Analysis: Slider vs Sweeper by Batter Handedness (Batter Side of the Plate)

Given the earlier finding that both Sliders (SL) and Sweeper (ST) increase strike likelihood, we now find this effect varies significantly by batter's side of the plate showing how batter batter's side of the plate changes the Slider/Sweeper advantage as illustrated in *Figure 8* and *Table 5*.

Against hitters batting left-handed, the Slider shows a stronger advantage (mean +0.54, HDI: -0.99 to 2.10) likely due to its sweeping movement away from their vision creating a larger perceived strike zone, though it has a substantial variability (SD=0.80) which indicates this isn't guaranteed as some hitters batting left-handed still punish poorly located Sliders. For batting hitters right-handed, the Slider maintains a more modest edge (mean +0.24, HDI: -0.98 to 1.57) with slightly more consistent results (SD=0.65), though the HDI crossing zero means pitchers can’t fully rely on this effect. These results confirm two critical insights: first, the traditional platoon advantage holds as Sliders gain an extra +0.30 mean effect against hitters batting left-handed, and second, the wide HDIs for both groups prove pitch quality (e.g. location and break) ultimately outweighs batter batting side alone in determining strike outcomes, this means other variables and their interaction with batter side of the plate matter more than plate side alone, showing that effect varies substantially by pitch and context.

This explains why elite pitchers dominate both sides of the plate, their premium command overrides these baseline tendencies, while average pitchers must carefully consider other variables interacting with these batter’s side of the plate effects when selecting pitches and not just the batter's side of the plate alone.

Figure 7 *Effect of Batter Handedness (stand)*

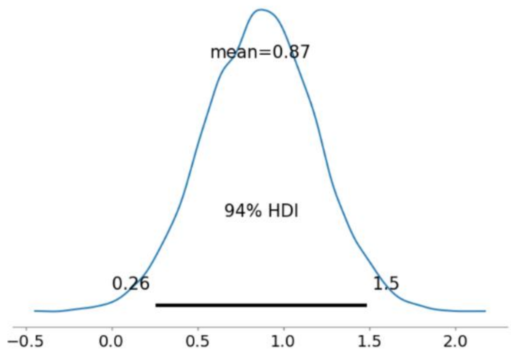


Table 4 *Coefficient of Batter Handedness (Denoted by stand)*

Mean	*SD*	*HDI [0.03]*	*HDI [0.97]*	*Mcse mean*	*Mcse sd*	*Ess bulk*	*Ess tail*	*r_hat*
0.866	0.328	0.232	1.507	0.005	0.003	4618.0	5740.0	1.0

Figure 8 *Contrast: Slider vs Sweeper by Batter Handedness*

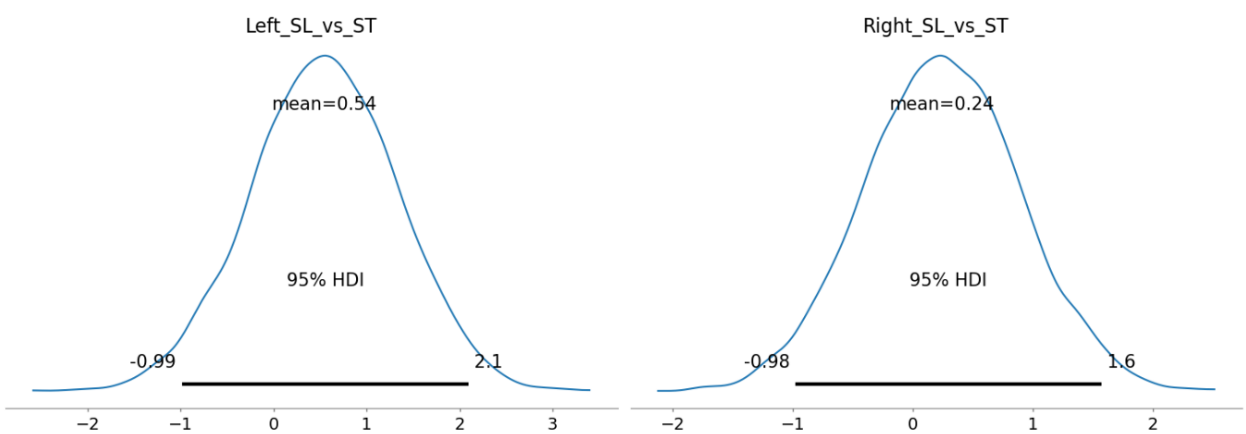


Table 5 *Contrast Metrics for Slider vs Sweeper by Batter Handedness*

	Mean	*SD*	*HDI [0.03]*	*HDI [0.97]*	*Mcse mean*	*Mcse sd*	*Ess bulk*	*Ess tail*
Left_SL_vs_ST	0.541	0.791	-0.988	2.095	0.006	0.007	15172	6871
Right_SL_vs_ST	0.238	0.648	-0.978	1.569	0.006	0.006	13279	7866

Analysis of Slider vs. Sweeper in 2-Strike Counts

Since the batter's position interacts to affects strike outcome, it's important to see if these patterns hold in high-count situations, especially with two strikes. In these moments, batters often shift to a more defensive approach, and pitchers may change their pitch selection to go for strikeouts. Two-strike counts change how pitch choices result to strike outcome. This makes it useful to compare the Slider and Sweeper in these situations.

From *Figure 9* and *Table 6*, against hitters batting left-handed, the slider maintains a clear +0.53 run advantage on average, though with significant variability (SD =0.867; HDI: -1.15 to 2.27). This wide HDI shows what is regularly observed in real game, a perfectly executed backdoor slider paints the black for a strike outcome or generates a weak swing, while one that catches too much plate often gets driven into the gap.

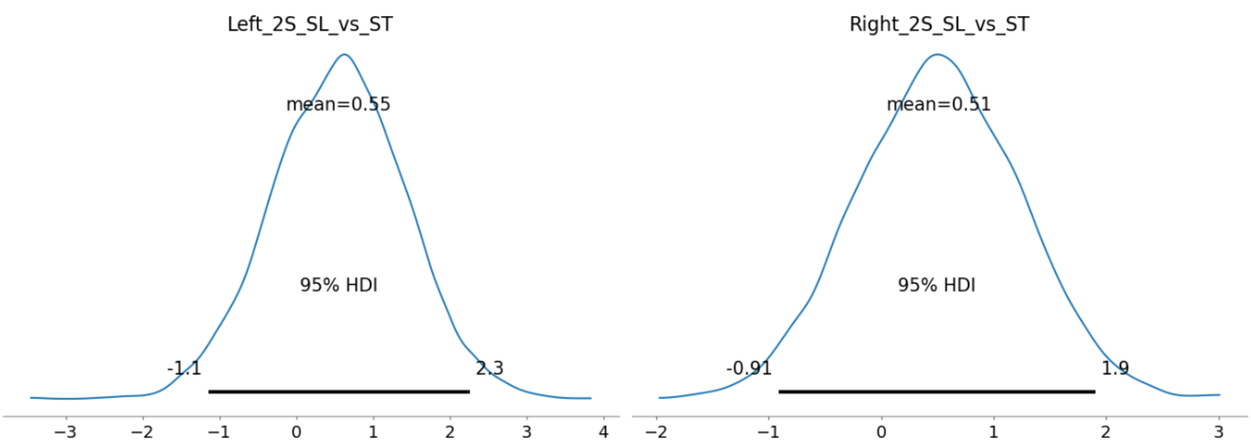
The advantage becomes similar against hitters batting right-handed, where sliders show a stronger +0.51 mean effect with more consistent results (SD = 0.727; HDI: -0.91 to 1.91). This enhanced effectiveness comes from the pitch's natural movement profile, as the slider breaks sharply away from a hitter batting right-handed swing path, it creates particularly uncomfortable at-bats when the hitter must protect with two strikes. These findings show that the Slider remains the premium decisive strike pitch regardless of batter handedness in two-strike count, the margin for success is slim, which makes execution quality, particularly pitch location and movement sharpness, a key to successful strike outcomes. Though small platoon tendencies still exist but command remains the key differentiator between elite and average pitchers.

The data ultimately validates the timeless pitching adage: with two strikes, it's not just about what you throw, but where you throw it.

Table 6 *Contrast Metrics for Slider vs Sweeper in 2-Strike Counts by Batter Handedness*

	Mean	*SD*	*HDI [0.03]*	*HDI [0.97]*	*Mcse mean*	*Mcse sd*	*Ess bulk*	*Ess tail*
Left_2S_SL_vs_ST	0.553	0.867	-1.146	2.265	0.007	0.007	13782	7261
Right_2S_SL_vs_ST	0.514	0.727	-0.912	1.905	0.006	0.006	12880	8073

Figure 9 *Slider vs Sweeper in 2-Strike Counts by Batter Handedness*



Pitch Performance Analysis: Full Count (3-2) Situation

Following the examination of the effects of the Slider compared to Sweeper when batters face 2 strikes, it is critical to explore how the Slider behaves in high-stakes pitch-count scenarios, such as full counts (3-2). This situation stands out because it includes a higher likelihood of batter action (swinging), and it also provides a clear context for understanding how pitch type and batter handedness influence strike outcomes when the pressure is higher.

Primary Findings

The contrast analysis in *Table 7* reveals significant differences in Slider effectiveness based on batter's side of the plate and other pitch types:

Against hitters batting left-handed (L_SL_vs_CH), the Slider shows essentially no systematic advantage over the Changeup, with a mean contrast of **+0.043** and a wide HDI (–2.092 to +2.142) that fully crosses zero. This shows that for left-handed batters in full counts, the choice between a Slider and a Changeup does not consistently favor one pitch type over the other and may depend heavily on location, movement, and batter tendencies.

Against hitters batting right-handed (R_SL_vs_CH), the Slider maintains a large and statistically significant advantage, with a mean contrast of **+2.591** (HDI: **0.763 to 4.376**). The HDI fully excludes zero, which indicates strong confidence that Sliders are significantly more likely than Changeups to result in strikes when facing right-handed batters in full counts.

Secondary Pitch Comparisons

The Slider maintains moderate advantages compared to other pitch types (Curveball, Forkball, Sinker) when a pitcher faces right-handed hitters in full counts, although the results are less definitive against left-handed hitters. Against lefties, the strike outcomes appear to be more dependent on pitch quality rather than just pitch type alone, with no single pitch clearly outperforming the others.

These results show that the Slider serves as a particularly effective option in full-count situations, especially against batter batting right-handed where its advantage is both large and statistically significant. Also, against left-handed batters, pitch selection may need to be more carefully designed to individual batter weaknesses rather than assuming the Slider is the superior option. This analysis explains the timeless pitching adage that in full counts, it's not just about what you throw, but where and how you throw it.

Table 7 *Contrast Metrics for Slider vs Other Pitch Types by Batter Handedness (Full Count:3-2)*

	Mean	*SD*	*HDI [0.03]*	*HDI [0.97]*	*Mcse mean*	*Mcse sd*	*Ess bulk*	*Ess tail*
L_SL_vs_CH	0.043	1.087	-2.092	2.142	0.010	0.010	10785	7756
R_SL_vs_CH	2.591	0.928	0.763	4.376	0.009	0.007	10755	8446
L_SL_vs_CU	-1.234	1.167	-3.457	1.113	0.011	0.009	11461	8005
R_SL_vs_CU	1.246	0.950	-0.597	3.103	0.009	0.007	11045	8444
L_SL_vs_FC	0.241	1.180	-2.100	2.534	0.011	0.010	1619	8392
R_SL_vs_FC	1.257	0.979	-0.597	3.201	0.009	0.007	10848	7531
L_SL_vs_FF	-1.355	1.010	-3.270	0.472	0.010	0.007	9626	7552
R_SL_vs_FF	0.374	0.773	-1.174	1.811	0.008	0.006	7994	8006
L_SL_vs_SI	1.230	1.129	-0.959	3.325	0.011	0.008	10351	8343
R_SL_vs_SI	0.731	0.859	-0.892	2.438	0.010	0.008	8050	6833
L_SL_vs_ST	-0.150	1.373	-2.918	2.431	0.012	0.014	13262	7447
R_SL_vs_ST	0.997	1.131	-1.278	3.139	0.010	0.009	13459	7676

Discussion of Results

This aim of this study is to examine the likelihood of a pitch resulting in a strike using a Bayesian econometric model by focusing on the interactions that exist between pitch characteristics, pitch location, and batter/pitcher handedness. The findings contribute new empirical insights that help bridge the gap between descriptive observations and causal understanding of pitch location with strike zone enforcement. This discussion integrates the results with the prior literature and aligns some of the important findings with practical applications and explains the implications for strategic baseball decisions from a pitcher's perspective.

Strike Zone Location

The results show that vertical pitch location (plate_z) is one of the strongest and consistent predictors of strikes outcomes, and this finding is in line with Zimmerman et al. (2019), who emphasized the variability of umpire calls near the strike zone edges, particularly at vertical extremes. However, unlike previous studies that mostly described this behavior, the use of Bayesian modeling helps by quantifying this effect probabilistically and establishes that higher pitches (including those near sz_top) are significantly more likely to result in strikes.

This insight supports what many in the baseball world have seen play out visually, especially with high-zone fastballs and breaking balls seen to be more likely to draw a favorable call. A significant difference was found between high and low pitches. This supports earlier results and gives a basis for how pitch height might be used in two-strike situations.

Horizontal Location:

Horizontal pitch location (plate_x) demonstrated weaker and more uncertain effects on strike probability, with the 95% HDI including zero. This information confirms the findings from Gray and Cãnal-Bruland (2018), who observed that visual tracking challenges may impact perception of outside pitches but unlike their findings which suggested that higher swing rates on high-and-outside pitches, the results in this study focus on strike likelihood and indicate umpires are more hesitant to reward these outer-edge locations with strike calls, especially when compared to vertical positioning. This insight shows the risk pitchers take when targeting the outer edge (horizontal dimension) for strike efficiency. This reinforces the idea that command in the vertical dimension is more actionable for pitchers than attempting to live on the corners, especially given the inconsistency of calls across horizontal boundaries.

Pitch Type: Strategic Insights for Pitchers

While prior work like that of Chen and Huang (2024) focused on umpire difficulties with certain pitch types (e.g., left-handed sliders or high-spin fastballs), this study directly evaluates the relationship between certain frequent pitch types and strike likelihood. The Bayesian analysis shows that sliders (SL) and sweeper (ST) are consistent in demonstrate higher strike probabilities than other pitch types such as cutters (FC), and forkballs (FS) etc.

The results suggest that in high-pressure moments like a full count or two-strike scenarios, pitchers might benefit from relying more on sliders, because sliders provide a probabilistic advantage in generating strikes outcomes, particularly against hitters batting right-handed. This result is statistically significant and also in line with real world preferences seen in pitch calling and scouting reports.

Batter Handedness and Contextual Factors

The relationship between batter handedness and strike outcomes shows some insights and patterns that are context-dependent across models. When analyzing batter side of the plate directly, pitchers appear to have a slight but reliable advantage when facing right-handed batters, with pitches more likely to result in strikes overall. This baseline builds on Chance and Maymin (2023), who noted the visual and perceptual challenges created by handedness matchups.

However, when isolating specific pitch types like the Slider and Sweeper, batter side effects become more subtle. In general count situations, Sliders show a modestly stronger advantage against left-handed batters, consistent with traditional platoon expectations, but these effects exhibit wide uncertainty, which emphasizes that pitch execution matters more than batter side alone.

In two-strike counts, the influence of batter side appears to weaken. Sliders show similar results against both left- and right-handed hitters. This may point to pitch location and movement becoming more important than batter side of the plate in these moments. In full-count situations (3-2), however, the platoon advantage re-emerges clearly with sliders having a large and statistically significant advantage against right-handed batters compared to Changeups, but no consistent advantage is observed against left-handed batters. This pattern shows that batter handedness can reassert importance under certain pitch types, count scenarios and other game variables.

Overall, these findings explain that while batter side of the plate does play a role in strike likelihood, its influence is very dependent on pitch type, count pressure and other variables.

Coaches and pitchers should consider batter handedness strategically, but not rigidly as execution quality remains the primary determinant of success in high-leverage situations.

Contrasting Basic Logistic Regression and Bayesian Insights

While the panel logistic regression model provided important baseline insights into factors influencing strike likelihood, there was a notable significant negative effect of horizontal plate location (plate_x²) and the positive effect of using Sliders. The Bayesian analysis revealed deeper and more detailed patterns not immediately apparent in the basic logit approach. Specifically, the panel model suggested that vertical location and strike zone boundaries (sz_top, sz_bot) were not significant predictors. However, the Bayesian analysis, by modeling full uncertainty distributions rather than relying solely on p-values, showed that higher vertical pitch locations (plate_z) and pitches near the upper boundary of the strike zone (sz_top) substantially increase the probability of a strike, with HDIs excluding zero and providing strong directional confidence.

Furthermore, while the basic model detected the significance of pitch type (e.g., Sliders), the Bayesian contrasts provided more information about pitch type effectiveness, such as quantifying the comparative advantage of Sliders versus Changeups in full counts considering different game variables as mediators and confounders, therefore highlighting large uncertainty around some pitch types (e.g., Forkballs) that a simple logit analysis could not capture. All in all, the Bayesian framework allowed for the discovery of credible but subtle effects, accounting for how variables interact under specific situations insights that the frequentist model, which focuses narrowly on point estimates and fixed hypothesis testing, might have overlooked. These results add support to the findings and help explain how pitch strategy relates to strike outcomes.

Strategic Relevance and Application

These findings can be applied in areas like pitch selection, player development, and game planning. Todri and Stern (2015) noted that many machine learning models are difficult to interpret in causal terms. In contrast, the use of Bayesian methods here is aimed at making the results easier to apply in real decisions. For example, pitchers can confidently use Sliders in full counts against batters batting right-handed, not because they always succeed, but because they are statistically more likely to draw get a strike when well-executed.

Similarly, since plate_z and sz_top consistently perform well which is a clear advantage of targeting high in the zone and coaches should emphasize high-zone command, particularly for pitchers with average velocity but sharp location up in the zone. This proves that when stats account for real-world cause and effect, they only don't just describe the game but also change how game is played.

Broader Applications of the Approach

Beyond the specific questions explored in this study, the modeling framework developed here by combining panel logistic regression and Bayesian analysis provides a flexible way that can be adapted to address a variety of other questions in baseball research and decision-making. One possible use of these methods is to examine how the order of pitches affects strike outcomes depending on the count or the type of batter. This kind of analysis can support the coach and pitcher decisions about pitch selection throughout a game. Likewise, this approach could be extended to analyze how individual players behave against opponent players, also it can be used to analyze situational context such as base runners, inning pressure, or pitcher's fatigue level, how they interacts with pitch type and location to affect strike outcomes. Through the ability to model uncertainty along with contextual interactions, this framework used in this work provides a foundation for answering deeper tactical questions in baseball without being restricted to the narrow assumptions of traditional models.

Limitations and Future Work

This study has some key limitations. First, the study didn't include rare pitch types like knuckleballs or screwball pitches, because they appeared so infrequently in the data. While leaving them out helped keep the analysis statistically great and made the models easier to train, it also means the findings here might not fully apply to the wide variety of pitches that are thrown in professional baseball. Some of these rare pitches could have certain unusual movement patterns that can affect a pitch resulting in a strike. Future studies could address this to better understand their impact on strike outcomes.

Another limitation is that the model didn't account for individual umpire differences. Adding umpire-specific effects could have helped control for personal biases in improving the model's accuracy and providing better insights into how different umpires react to pitches that can result in strike outcomes. From a strategic perspective, future research could examine game theory to analyze pitch selection. Instead of just looking at pitch characteristics, this approach would treat pitchers and umpires as players with objectives, such that the pitchers try to exploit the strike zone while umpires adjust (consciously or not) to pitch patterns, game context, and visual cues. A game-theoretic model could be used to explore which pitch sequences work best in different counts or against certain batters. When combined with probabilistic models, this may help explain how pitch selection works not just to avoid hits, but also to influence how umpires call the zone. Knowing this, pitchers may change their approach based on the count, the umpire's tendencies, or a batter's known patterns. In that case, strike probability depends not only on the pitch but also on when and why it's used. Looking at this, through game theory could help explain how pitchers and umpires respond to each other during the game.

Conclusion **Chapter 5**

Previous work on strike outcome has limited depth and practical application but this study aimed to fill those gaps and make the findings more useful in real games. Existing literature has heavily relied on descriptive or correlational approaches, with minimal application of causal inference approaches. Also, batter and pitcher handedness have been included in past models, their interactive influence on pitch presentation and strike likelihood has remained something that is less explored.

Through posterior distributions and contrast estimates from Bayesian logistic regression framework, these gaps were addressed using a carefully stratified MLB pitch-level dataset that can provide estimation of direct and interaction effects. The analysis showed that vertical pitch location (plate_z) with higher strike zone (sz_top) has the most consistent and significant effect on strike likelihood. Pitch type, particularly Sliders and Sweepers, also plays a meaningful role, with their effects varying by batter handedness and count context (e.g., full counts or two-strike scenarios). These findings confirm that strike likelihood is shaped by a combination of location precision, pitch type, and pitcher-batter matchups, therefore it provides strategic implications for pitch selection and game planning.

In doing so, this study directly answers the central research questions: (1) How do different pitch types influence strike likelihood? (2) How does pitch location both horizontally and vertically affect the probability of a pitch resulting in a strike? (3) How does batter handedness (batter side of plate) influence how pitch type affects strike outcomes? By answering these questions, this analysis helps show how pitchers might adjust pitch choice and location based on pitch movement, batter side, and game context. It also looks at how these factors relate to whether a pitch will result in a strike. Moving beyond simple correlations, this model provides a cause-and-effect understanding of how pitch dynamics and context interact to influence strike outcomes. Bayesian methods were used to build on earlier descriptive results and make the findings easier to interpret and apply. This research supports future directions in pitch design, strategic matchup planning, and the development of pitcher-focused tools that improve strike efficiency and game impact.

References

Adler, D. (2019, June 1). Identifying pitch types: *A fan's guide. MLB.com.* https://www.mlb.com/news/identifying-pitch-types-a-fan-s-guide

Bates, B., & Campbell, E. (2001). A Markov Chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Water Resources Research, 37(4), 937–947*. https://doi.org/10.1029/2000WR900363

Brantley, J. A., & Körding, K. P. (2024). Bayesball: Bayesian integration in professional baseball batters. *bioRxiv.* https://doi.org/10.1101/2022.10.12.511934

Briggs, L. J. (1959). Effect of spin and speed on the lateral deflection (curve) of a baseball; and the Magnus effect for smooth spheres. *American Journal of Physics, 27(7), 589–596*. https://doi.org/10.1119/1.1934921

Cao, Z., & Qu, H. (2023). Review on causality detection based on empirical dynamic modeling. *arXiv*. https://doi.org/10.48550/arXiv.2312.15919

Casella, G., & Berger, R. (2005). Estimation with selected binomial information or do you really believe that Dave Winfield is batting .471 *Anthology of Statistics in Sports.* https://doi.org/10.1137/1.9780898718386.ch13

Chance, D. M., & Maymin, P. Z. (2023). A new look at the left-handed advantage in baseball. *International Journal of Performance Analysis in Sport*. https://doi.org/10.1080/24748668.2023.2255806

Chen, M.-H., & Shao, Q. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics, 8(1), 69–92.* https://doi.org/10.1080/10618600.1999.10474802

Chen, M. H., Ibrahim, J. G., & Yiannoutsos, C. (1999). Prior elicitation, variable selection, and Bayesian computation for logistic regression models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61(1), 223–242*. https://doi.org/10.1111/1467-9868.00173

Chen, H. T., Tsai, W. J., & Lee, S. Y. (2009). Stance-based strike zone shaping and visualization in broadcast baseball video: Providing reference for pitch location positioning. *IEEE International Conference on Multimedia and Expo.* https://doi.org/10.1109/ICME.2009.5202495

Chen, Y.-H., & Huang, S.-K. (2024). The influence of pitcher handedness on pitch-calling behavior: Insights from fMRI study on baseball umpires. *Psychophysiology*. https://doi.org/10.1111/psyp.14501

Dong, X., & Heineke, J. (2016). Correlation or causation?: The sorry state of inference in empirical modeling. *ERN: Other Econometrics: Econometric & Statistical Methods*. https://doi.org/10.2139/ssrn.2933053

Emenyonu, S. C., & Adam, M. (2019). Bayesian analysis via Markov Chain Monte Carlo algorithm on logistic regression model. *International Journal of Statistics and Applications, 9(1), 1–8*. https://doi.org/10.22610/JSDS.V4I4.751

Erp, H. V., & Gelder, P. V. (2013). Bayesian logistic regression analysis. *Journal of Physics: Conference Series, 1725(1), 012010*. https://doi.org/10.1063/1.4819994

Ganeshapillai, G., & Guttag, J. V. (2012). Predicting the next pitch (*Sloan Analytics Conference*)

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2(4), 1360–1383*. https://doi.org/10.1214/08-AOAS191

Gogtay, N., Ranganathan, P., & Aggarwal, R. (2021). Understanding estimands. *Perspectives in Clinical Research*. https://doi.org/10.4103/picr.picr_384_20

Gomaz, L., Bouwmeester, C., van der Graaff, E., van Trigt, B., & Veeger, D. (2023). Machine learning approach for pitch type classification based on pelvis and trunk kinematics captured with wearable sensors. *Italian National Conference on Sensors.* https://doi.org/10.3390/s23239373

Gordóvil-Merino, A., Guàrdia-Olmos, J., & Peró-Cebollero, M. (2012). Estimation of logistic regression models in small samples: A simulation study using a weakly informative default prior distribution. Quality & Quantity, 46(4), 1131–1141.

Gray, N., & Ferson, S. (2021). Logistic regression through the veil of imprecise data. International *Journal of Approximate Reasoning, 135, 1–15.* https://doi.org/10.48550/arXiv.2106.00492

Green, E. A., & Daniels, D. P. (2014). What does it take to call a strike? Three biases in umpire decision making. *SSRN*. https://doi.org/10.2139/ssrn.2391558

Healey, G. (2015). Modeling the probability of a strikeout for a batter/pitcher matchup. *IEEE Transactions on Knowledge and Data Engineering, 27(9), 2415–2423.* https://doi.org/10.1109/TKDE.2015.2416735

Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy, 23(4), 290–300*. https://doi.org/10.1016/j.bjpt.2018.12.006

Higuchi, T., Morohoshi, J., Nagami, T., Nakata, H., & Kanosue, K. (2013). The effect of fastball backspin rate on baseball hitting accuracy. *Journal of Applied Biomechanics, 29*(3), 279-285. https://doi.org/10.1123/JAB.29.3.279

Hsu, M. (2023). Umpire home bias in Major League Baseball. *Journal of Sports Economics.* https://doi.org/10.1177/15270025231222631

Huang, J.-H., & Hsu, H.-J. (2020). Approximating strike zone size and shape for baseball umpires under different conditions. *Journal of Sports Analytics*, *6*(1), 15–27. https://doi.org/10.1080/24748668.2020.1726156

Hunermund, P., & Bareinboim, E. (2019). Causal inference and data fusion in econometrics. *Econometrics Journal*. https://doi.org/10.1093/ectj/utad008

Ito, H., & Gotoh, Y. (2018). Technical support system for baseball beginners by analyzing batting stance with sensors. *Advances in Mobile Multimedia.* https://doi.org/10.1145/3282353.3282364

Jensen, S. T., McShane, B. B., & Wyner, A. J. (2009). Hierarchical Bayesian modeling of hitting performance in baseball. *Bayesian Analysis, 4(4), 631–652*. https://doi.org/10.1214/09-BA424

Kahan, B., Cro, S., Li, F., & Harhay, M. (2023). Eliminating ambiguous treatment effects using estimands. *American Journal of Epidemiology.* https://doi.org/10.1093/aje/kwad036

Kalkhoven, J. T., et al. (2024). Athletic injury research: Frameworks, models, and the need for causal knowledge. *Sports Medicine, 54(5), 1121–1137*. https://doi.org/10.1007/s40279-024-02008-1

Keeley, D. W., Wicke, J., Alford, K., & Oliver, G. D. (2010). Biomechanical analysis of forearm pronation and its relationship to ball movement for the two-seam and four-seam fastball pitches. *Journal of Strength and Conditioning Research, 24(9), 2366–2371.* https://doi.org/10.1519/JSC.0b013e3181b22aea

Kidokoro, S., Matsuzaki, Y., & Akagi, R. (2020). Does the combination of different pitches and the absence of pitch type information influence timing control during batting in baseball? *PLOS ONE, 15(3), e0230385*. https://doi.org/10.1371/journal.pone.0230385

Lukman, P., Abdullah, S., & Rachman, A. (2021). Bayesian logistic regression and its application for hypothyroid prediction in post-radiation nasopharyngeal cancer patients. *Journal of Physics: Conference Series, 1725(1), 012010.* https://doi.org/10.1088/1742-6596/1725/1/012010

Lundberg, I., Johnson, R. A., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review.* https://doi.org/10.1177/00031224211004187

Lynch, B. M., Dixon-Suen, S. C., Ramirez Varela, A., Yang, Y., English, D. R., Ding, D., Gardiner, P. A., & Boyle, T. (2020). Approaches to improve causal inference in physical activity epidemiology. *Journal of Physical Activity and Health, 17(1), 80–84.* https://doi.org/10.1123/jpah.2019-0515

MacMahon, C., & Starkes, J. (2008). Contextual influences on baseball ball-strike decisions in umpires, players, and controls. *Journal of Sports Sciences*, *26*(7), 751–760. https://doi.org/10.1080/02640410701813050

Magee, A. F., Karcher, M. D., Matsen, F. A., & Minin, V. M. (2021). How trustworthy is your tree? Bayesian phylogenetic effective sample size through the lens of Monte Carlo error. *Bayesian Analysis, 16(2), 503–529*. https://doi.org/10.1214/22-ba1339

Manzi, J., Dowling, B., Krichevsky, S., Roberts, N., Sudah, S. Y., Moran, J., Chen, F. R., Quan, T., Morse, K. W., & Dines, J. (2023). Pitch-classifier model for professional pitchers utilizing 3D motion capture and machine learning algorithms. *Journal of Orthopaedics*. https://doi.org/10.1016/j.jor.2023.12.007

Mizels, J., Erickson, B., & Chalmers, P. N. (2022). Current state of data and analytics research in baseball. *Current Reviews in Musculoskeletal Medicine, 15(2), 120–130.* https://doi.org/10.1007/s12178-022-09763-6

Myers, T., Gallagher, I. J., & Reading, J. (2023). Understanding quantitative analysis in sport and exercise science: Concepts, tools, and approaches. *Sport and Exercise Science Review.* https://doi.org/10.51224/SRXIV.418

Nakata, H., Miura, A., Yoshie, M., Kanosue, K., & Kudo, K. (2013). Electromyographic analysis of lower limbs during baseball batting. *Journal of Strength and Conditioning Research, 27(5), 1179–1187*. https://doi.org/10.1519/JSC.0b013e3182653ca9

Pham, H. T. T., & Pham, H. (2021). On the existence of posterior mean for Bayesian logistic regression. Monte Carlo Methods and Applications, 27(1), 1–12. https://doi.org/10.1515/mcma-2021-2089

Rainey, D., & Larsen, J. (1988). Balls, strikes, and norms: Rule violations and normative rules among baseball umpires. *Journal of Sport and Exercise Psychology*, *10*(1), 75–80. https://doi.org/10.1123/JSEP.10.1.75

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science, 1(1), 27–42*. https://doi.org/10.1177/2515245917745629

Röver, C., Bender, R., Dias, S., Schmid, C. H., Sturtz, S., Trikalinos, T. A., & Friede, T. (2020). On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods, 11(2), 246–262.* https://doi.org/10.1002/jrsm.1475

Santos-Fernandez, E., Wu, P. P., & Mengersen, K. L. (2019). Bayesian statistics meets sports: A comprehensive review. *Journal of Quantitative Analysis in Sports, 15(4), 289–312.* https://doi.org/10.1515/JQAS-2018-0106

Shrier, I. (2007). Understanding causal inference: The future direction in sports injury prevention. *Clinical Journal of Sport Medicine, 17(3), 220–224.* https://doi.org/10.1097/JSM.0b013e3180385a8c

Swartz, P. M., Grosskopf, M., Bingham, D., & Swartz, T. (2017). The quality of pitches in Major League Baseball. *The American Statistician, 71(1), 40–47.* https://doi.org/10.1080/00031305.2016.1264313

Todri, V., & Stern, L. N. (2015). Big data: From correlation to causation. Economics, Computer Science.

Umemura, K., Yanai, T., & Nagata, Y. (2020). Application of VBGMM for pitch type classification: Analysis of TrackMan's pitch tracking data. *Japanese Journal of Statistics and Data Science, 3(2), 201–215*. https://doi.org/10.1007/s42081-020-00079-8

Vock, D. M., & Boehm Vock, L. F. (2018). Estimating the effect of plate discipline using a causal inference framework: An application of the G-computation algorithm. *Journal of Quantitative Analysis in Sports, 14(2), 37–56*. https://doi.org/10.1515/jqas-2016-0029

Whiteside, D., Martini, D. N., Zernicke, R. F., & Goulet, G. C. (2016). Changes in a starting pitcher's performance characteristics across the duration of a Major League Baseball game. *International Journal of Sports Physiology and Performance, 11(2), 247–254.* https://doi.org/10.1123/ijspp.2015-0121

Wiens, S., & Nilsson, M. E. (2016). Performing contrast analysis in factorial designs: From NHST to confidence intervals and beyond. *Educational and Psychological Measurement, 77(4), 612–635.* https://doi.org/10.1177/0013164416668950

Williams, J., & MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling, 15*(1), 23–51. https://doi.org/10.1080/10705510701758166

Yee, R., & Deshpande, S. K. (2023). Evaluating plate discipline in Major League Baseball with Bayesian additive regression trees. *Journal of Quantitative Analysis in Sports, 20(1), 5–20.* https://doi.org/10.1515/jqas-2023-0048

Zhan, W., Baise, L., & Moaveni, B. (2023). An uncertainty quantification framework for logistic regression-based geospatial natural hazard modeling. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4392324

Zhao, Y., Liu, W., Yan, Y., & Li, F. (2024). Navigating the confluence of econometrics and data science: Implications for economic analysis and policy. *Theoretical and Natural Science.* https://doi.org/10.54254/2753-8818/38/20240551

Zimmerman, D., Tang, J., & Huang, R. (2019). Outline analyses of the called strike zone in Major League Baseball. *Annals of Applied Statistics*, *13*(4), 2223–2244. https://doi.org/10.1214/19-aoas1285

Zitzmann, S., & Hecht, M. (2019). Going beyond convergence in Bayesian estimation: Why precision matters too and how to assess it. Structural Equation Modeling: *A*

Multidisciplinary Journal, 26(4), 646–661.
https://doi.org/10.1080/10705511.2018.1545232

Zuur, A. F., Ieno, E. N., & Smith, G. M. (2008). Analysing ecological data. https://doi.org/10.1111/j.1541-0420.2008.00962_2.x

Appendix

The full code, and additional materials on how to run the code related to this work are available at the following GitHub repository:

https://github.com/Olubayode/Bayesian-Analysis-Baseball/tree/main