

Discrete Margin Modeling and Beyond: A Toolkit for Football Betting Analysis

Richard Oldham

September 25, 2025

Abstract

We extend our earlier development of a discrete-margin, key-number probability model for football games into a comprehensive toolkit for betting analysis. This extended framework covers alternative wagering strategies inspired by Wong teasers, machine-learning pipelines for probability modeling in NFL and NCAA football, and a Bayesian hierarchical modeling layer for dynamic team strength estimation. We provide concrete modeling recipes and highlight scenarios where each approach outperforms traditional methods. The unified methodology enables accurate pricing of exotic bets (teasers, alt-spread parlays, pleasers), robust probability estimates for in-game and same-game combinations, and principled uncertainty quantification for bankroll management.

1 Introduction

Point-spread betting in American football often revolves around a few *key numbers* – notably 3 and 7 – which correspond to the most common victory margins due to the scoring system (field goals and touchdowns). In the NFL, for example, roughly 15% of games end with a 3-point margin and about 10% with a 7-point margin, making these outcomes far more frequent than neighboring values [7]. These discrete probability spikes mean that moving a betting line across 3 or 7 (for instance from +2.5 to +3.5) has outsized value. Betting strategies like the well-known *Wong teaser* were designed to exploit this phenomenon by teasing point spreads through 3 and 7 to capture those key margins [8]. In earlier work, we constructed a discrete-margin probability model that explicitly quantifies the probability mass at each possible scoring margin. That model provides “fair” win probabilities for alternate spreads and teaser legs by accounting for the empirical frequency of key margins.

Armed with this discrete-margin engine, we now expand our analysis in several directions. First, we examine alternative wagering strategies that leverage the same price-for-probability edge as classic Wong teasers, but in different wrappers (e.g. parlays of alternate spreads, correlated same-game bets, or even reverse teasers). These strategies are motivated by recent market trends: sportsbooks have adjusted payout pricing on standard NFL teasers, shrinking the margins that bettors enjoyed a decade ago [1]. We describe how a bettor can adapt by shopping for outlier lines and constructing custom combinations that replicate a teaser’s value

at better odds. Second, we outline machine-learning (ML) pipelines to generate and validate the underlying probability distributions for game margins. These include dynamic rating systems (to produce a baseline expected score difference and total points) combined with calibration mechanisms to distribute probability mass across discrete scores. We emphasize handling NFL and NCAA games, noting that college football’s higher variance and scoring distribution call for different treatment of key numbers. Third, we incorporate a Bayesian hierarchical modeling layer to improve estimates in data-sparse situations and to quantify uncertainty. A hierarchical Bayesian approach, with partial pooling of team strengths and a dynamic time component, offers a principled way to update predictions as new information arrives (e.g. team performance shifts or injuries) while avoiding overfitting to noisy outcomes.

The remainder of this paper is organized as follows. In Section 2, we detail several alternative strategies akin to Wong teasers and explain how to evaluate their expected value using the discrete-margin model. In Section 3, we present ML-driven workflows for simulating game outcomes and pricing various bets, from straightforward spread bets to complex same-game parlays. Section 4 introduces a Bayesian hierarchical model for football scores which provides a flexible, high-end layer for our predictions. Section 5 then synthesizes these elements into an end-to-end betting evaluation workflow and offers practical guidelines on when to deploy each strategy. We also outline some prototype tools and “recipes” that practitioners can derive from our framework (such as half-point value tables and teaser pricing calculators). Finally, Section 6 concludes and the Appendix includes full references to key literature and data sources.

2 Alternative Strategies Similar to Wong Teasers

Wong teasers (also called basic strategy NFL teasers) refer to two-team 6-point teasers that “cross” the key numbers 3 and 7 on both legs [8]. For instance, teasing an underdog of +2.5 up to +8.5, or a favorite of −7.5 down to −1.5, converts a bet that would have lost by a small margin into one that wins in the frequent cases where games land on 3 or 7. Historically, these specific teaser legs won around 76% of the time, easily clearing the break-even threshold when two such legs are parlayed [1]. Sportsbooks eventually caught on and adjusted payouts (often charging -120 or worse instead of the old -110 for a two-team teaser) and even dynamically vary teaser odds by team and point spread [1]. In today’s market, bettors seeking an edge must be more creative in how they exploit discrete scoring margins. We outline several strategies that leverage the same core math in different forms:

1. **Price-based custom teasers across books.** Modern books vary teaser pricing by number of legs, points teased, and sometimes by the specific game or team. The classic two-leg, 6-point teaser at -110 is no longer universal; you might see -120 or -130, or require three legs for a decent payout. Using our model, we can calculate the *fair win probability* for any candidate teaser leg (by summing the probability the final margin falls within the teased range). Comparing the product of leg win probabilities to the offered payout reveals if the teaser is +EV (positive expected value). Often the best opportunities come from *shopping* multiple books: some sportsbooks might misprice certain legs or still offer -110 on two-leg teasers while others are -130. The bettor’s

advantage lies in locating these outliers and even mix-and-matching legs from different books or days. For example, if Book A offers a generous price on a 6.5-point teaser and Book B is slow to adjust odds on a particular underdog, using one leg from each could yield a higher combined edge than either book's teaser alone.

2. **Parlays of alternate spreads instead of fixed teasers.** A standard teaser essentially parlays two (or more) alternate lines with a fixed 6-point shift and standardized price. However, sometimes you can do better by directly parlaying two alternate spreads that achieve a similar net move. For instance, suppose a favorite is -8.5; a 6-point teaser would take them to -2.5. If the sportsbook's alternate line market lets you take -2.5 at better than the implied -300 moneyline (which is roughly the fair price if the true win chance is around 75%), then parlaying two such alt-spread bets might pay, say, -125 instead of the -140 you might get in a worse teaser payout table. Our discrete-margin model is used to price each alternate spread: e.g. if team X -2.5 has a 74% win probability (covering -2.5), the fair single-leg moneyline is about -285. By examining the alt-spread ladders, we often find cases where sportsbooks "shade" the teaser lines (knowing many bettors blindly play teasers) but leave the alt spreads priced by a simpler formula. We then parlay two alt spreads that correspond to the same moves a teaser would give. This approach, essentially a do-it-yourself teaser, can outperform a traditional teaser in payout. The methodology mirrors the logic presented by Unabated's teaser tools [1], which allow bettors to input a line and see the derived price for alternate spreads.
3. **Totals-aware teaser leg selection.** Classic Wong teaser logic performs best when games have low totals (fewer expected points imply tighter possible margins and relatively higher probability mass on 3 and 7). If a game's over/under is, say, 38, a 6-point swing covers a larger fraction of likely outcomes than in a game with total 55. Our strategy is to rank potential teaser legs by the game's total in ascending order (low-total games first), and then by how many key numbers (3 and 7 primarily) the teaser would cross. Using our half-point value table (see Section 6) we can quantify the value of each 0.5 point move for a given spread/total combination. Empirically, NFL games with totals below about 42 have significantly higher density on 3 and 7 [7]. Thus an underdog +2.5 in a 38-total game is a prime teaser candidate (it will cover +8.5 a very high percentage of the time), whereas a favorite -7.5 in a 55-total game is less attractive (the game's higher scoring volatility reduces the relative frequency of winning by 1-7 points). In short: focus teasers on low-total games, and use your model's fine-grained probability outputs to cherry-pick only those legs that gain the most from the additional points.
4. **"Reverse" teasers (pleasers).** A *pleaser* is the opposite of a teaser: you lay points (move the spread against you) in exchange for a bigger payout. For example, a 6-point pleaser with two legs might turn an underdog +2.5 into -3.5 on each, paying out at perhaps +600 if both cover the new line. Pleasers are generally sucker bets because you are giving away the key numbers instead of capturing them. However, there are niche cases where they make sense. If our margin model indicates that a certain underdog has very little chance to lose by a small margin (say this team either wins outright

or gets blown out badly), then taking a pleaser that moves them from +2.5 to -3.5 could be +EV. Essentially, if the distribution of outcomes is bimodal (a lot of weight on either team X winning big or losing big, with little in between), then selling points around the middle can be advantageous. These situations are rare, but they can occur with teams that are highly volatile or when a favorite's true strength is in question (e.g. a high-variance offense). Because sportsbooks pay out pleasers at long odds and many bettors avoid them, the pricing models may not be as sharp for tails of the distribution. Our engine can calculate the probability a team wins by more than the pleaser line; if the implied odds are better than the book's payout, we have an edge.

5. Derivative markets that monetize the same distribution shape. Beyond spreads, there are other bets where key-number logic applies:

- *Half-quarter spreads*: Betting on 1st half, 2nd half, or even quarter lines that hover around 3 or 7. For example, a first-half line of -3 is analogous to a full-game -6 or -7 in some respects. If a team is -6 for the game, the first half might be -3 or -3.5. A model that properly accounts for the distribution of half-time scores (which also show spikes at 3 and 7, though typically lower frequencies) can uncover value in teasing or buying points in half-lines.
- *In-game (live) middles*: A middle is when you hold two opposite positions and can win both if the final margin falls in between. Live betting allows dynamic middling opportunities. For instance, if a favorite was -7 pre-game and falls behind early, the live line might swing to the opponent -3. Now there's a chance to have one bet on the original favorite -7 and another on the opponent -3. If the favorite wins by 4, 5, or 6, you win both bets. The probability of hitting such a middle is small, but if it involves key numbers (in this example, a middle covering 3,4,5,6,7), a smart model can estimate it and identify when the combined odds justify a middle attempt. Particularly, if a sudden change (like a lowered total due to injuries or weather mid-game) increases the likelihood of a tight finish, the distribution "clusters" more and a middle around those keys becomes more likely.
- *Correlated same-game parlays (SGPs)*: Sportsbooks often allow SGPs where you can parlay a side with a total or props from the same game. If done naively, books use approximate correlation adjustments or none at all, leading to potential pricing errors. A classic correlation play is parlaying an underdog with the under (if the game is low-scoring, points are at a premium which helps the underdog cover or win outright). Using a joint distribution of scores, you can compute the probability that both legs hit (e.g. Team Y +7 *and* Under 41). This essentially captures a "teaser-like" effect: low score magnifies the value of +7. If the book's parlay price undervalues this correlation, the bet is +EV. Our approach is to simulate or integrate over the bivariate distribution of (points for, points against) for the game, using the discrete margin model as a marginal on the difference and some model for the total, to get these joint probabilities. We can then compare to the same-game parlay payout. In essence, same-game parlays can recreate teaser moves with extra leverage from correlation, and a careful pricing (not relying on the sportsbook's black-box) will tell us when it's advantageous.

6. **College football (NCAA) adjustments.** Thus far, most examples focus on the NFL where scoring is more constrained. In college football (CFB), key numbers are less “key.” The distribution of winning margins is more diffuse due to higher scoring games, more variance, and uneven matchups (a 30-point favorite is common in CFB, unheard of in NFL). While 3 and 7 are still the most frequent margins in college, they occur at much lower rates (e.g. only about 9% of FBS games end with a 3-point margin, vs 14–15% in the NFL; and combined 16% for 3 or 7 in college vs 24% in NFL) [5]. This means basic Wong teasers (6 points) are far less effective in college; you simply don’t gain as much probability by teasing through 3 and 7. Instead, successful CFB betting strategies might focus on alternative key numbers (like 10, 14, 17 which are somewhat common in college blowouts) or use larger teasers (e.g. 10-point three-team teasers) selectively. Our model can be trained separately on college data to produce the college-specific margin frequency curve. We then identify if there are any “Wong-like” opportunities (for instance, teasing an underdog of +4.5 to +10.5 might capture 7 and 10 in one move, which could be valuable if those frequencies are meaningful). Additionally, college games might yield better pleaser opportunities because underdogs can either lose close or lose by 30+. If a book naively offers standard teaser/pleaser pricing for college, a tailored model will highlight that those prices are off given the distribution differences. In summary, treat college differently: use a distribution calibrated to college scores, and expect fewer edges on 6-point teasers but potentially some on bigger moves or different alt lines.

Key Points from Market and Literature:

- **Wong teaser definition & profitability:** A “Wong teaser” traditionally means a two-team, 6-point NFL teaser that moves favorites of -7.5 to -8.5 down below -3, or underdogs of +1.5 to +2.5 up above +7 [8]. These capture the key numbers 3 and 7 and historically were very profitable when priced at -110 for two legs. Many bettors hit upwards of 75% on such teasers, well above the 72.3% break-even rate for -110 parlays [1]. Wong (2001) noted that including other points (e.g. teasing -10 to -4) did not perform as well, hence the strict definition.
- **Key number frequencies in NFL:** About 30% of NFL games end with a margin of exactly 3 or 7 points (with 3 being the most common margin by far) [7]. The next most common margins are 6, 4, 10, etc., each much lower in frequency (on the order of 5–8%). This distribution underpins point-buying and teaser value: the reason getting +3.5 instead of +3 is so valuable is that 9–10% of games land on exactly 3, turning pushes into wins. Similarly, teasing a +2 up to +8. (crossing 3 and 7) targets those frequent margins. In college, as noted, the top two margins (3 and 7) sum to only 16% of games [5], so six-point moves are relatively less potent.
- **Sportsbook adjustments:** Sportsbooks have adapted to these realities. Many now price NFL teasers dynamically: the payout for a two-team teaser might be -120 or worse, and they may disallow certain “basic strategy” legs or require extra juice to include them [1]. Some books (like those using the Kambi platform) effectively treat teasers as parlays of alternate spreads, removing the fixed pricing advantage [1]. There

are even cases of books changing teaser rules (e.g. making a push count as a loss) to deter advantage play. The bettor’s counter is to either find the rare books that still offer value pricing or to replicate the bets via alt spreads as discussed. Tools such as the Unabated Teaser Calculator are invaluable for this, as they let you input your own win probability for each leg and see if the teaser’s implied odds are favorable.

In summary, while the NFL betting market has grown more efficient around key numbers, opportunities remain for those who use robust probability models and are willing to be flexible in their strategy. Whether through custom parlays, selective teasers, or correlated plays, the core idea is the same: exploit the “discrete margin” effect where the chance of a bet winning jumps disproportionately at certain score differentials. Our discrete-margin model provides the foundation for quantifying those jumps and identifying +EV situations in the modern landscape.

3 Machine-Learning Pipelines to Simulate NFL and NCAA Outcomes

Having a solid probability model for game outcomes is essential for pricing exotic bets and evaluating strategies. In this section, we outline two complementary modeling approaches (and a hybrid third) that leverage machine learning and simulation to produce calibrated win probability distributions. The aim is to generate a full probability mass function (pmf) over possible score differentials for each game, which our betting strategies can then plug into. We specifically address differences between NFL and NCAA (college) contexts, as the data distributions and optimal modeling choices can differ.

3.1 Ratings Backbone with a Discrete-Margin Front End

One effective approach is a two-stage model. In Stage A, we predict the expected outcome of the game in terms of point spread and total (the mean and variance of the score difference, implicitly). In Stage B, we translate those predictions into a full discrete distribution of the point margin, with special adjustments for key numbers.

Team Rating Models (Stage A). We start by constructing power ratings for each team as a baseline. Classic methods include Elo ratings or Bradley-Terry models that update team strengths based on wins/losses and margins, accounting for opponent strength. A more advanced approach uses a Bayesian dynamic model where each team i has attack and defense parameters (α_i, δ_i) that evolve throughout the season (e.g. via a random walk week to week) [3]. We can also incorporate known covariates: for example, a separate home-field advantage term h (possibly varying by team), or adjustments for short rest, travel distance, etc. The output of this ratings backbone is an estimate of the expected point spread μ for a matchup (team A vs team B on a neutral field might have $\mu = \text{rating}_A - \text{rating}_B + h \cdot (\text{home advantage})$). Similarly, an expected total points λ can be predicted using team offensive and defensive strengths (e.g. sum of expected points by each offense against the opposing defense). We can train these ratings using historical data: minimize prediction error of game results, or

use Bayesian inference to update distributions of team strengths each week [3]. Techniques like gradient boosted trees or neural networks can also be applied here using features such as team stats, injuries, weather, etc., to predict μ and λ . For instance, a gradient boosted model might take as input the teams’ recent efficiencies (points per drive, yards per play), starting quarterback status, and other situational variables, and output a predicted spread. Crucially, we also estimate the uncertainty or dispersion σ around that expected spread for Stage B.

From Continuous Outcomes to a Discrete PMF (Stage B). Given an estimated mean point spread μ and some measure of variance σ^2 for the matchup, we need to allocate probability mass to each possible score margin (including negative margins for team B wins). A simple approach might assume the point difference is approximately normally distributed around μ [6]. However, a Normal distribution will vastly underpredict the probability of exact outcomes like exactly 3 or 7. Instead, we use a **hybrid empirical approach**: Start with a continuous distribution (say Normal or a scaled t distribution) for baseline probabilities, then *re-weight* specific margins according to historical frequencies. For example, if our continuous model says the probability of a 3-point margin is 5%, but we know empirically it should be closer to 9% in an average NFL game, we can boost that probability and adjust neighboring probabilities downward to compensate. We do this for all key margins (3, 7, and to a lesser extent 6, 4, 10, 14, etc. for NFL; and similar but smaller adjustments for NCAA). Essentially, Stage B is a calibration layer. One implementation is to use a flexible discrete distribution family: for instance, a Skellam distribution (difference of two Poissons) or a Conway-Maxwell Poisson, fitted to the predicted mean and variance [2]. These distributions allow a parameter for dispersion that can be tuned. We then multiply certain probabilities by weights: w_3 for 3-point games, w_7 for 7-point games, etc., which are learned by fitting to historical data. The weights could depend on the total λ ; e.g. when λ is low, w_3 is high (field goals decide more games), and when λ is high (college shootouts), w_3 is lower. We ensure the pmf sums to 1 after re-weighting. This yields a categorical distribution $\{P(\text{margin} = m)\}$ for $m = -\infty \dots +\infty$ (practically from -60 to +60 or so). We verify calibration by checking that, for instance, in games where our model gave a 75% chance to a -2.5 favorite after teasing, about 75

Calibration and Validation. We use reliability diagrams and Brier score or log-loss to evaluate the model’s probabilistic predictions. For example, take all games where our model says the favorite has a 60% chance to cover -6. Did about 60% of them cover? Any systemic bias can be corrected via isotonic regression on the predicted probabilities. For validation, we back-test on past seasons: simulate betting strategies using these probabilities and see if the edges materialize after accounting for bookmaker vigorish. We also compare our distributional forecasts to simpler baselines: e.g. Stern’s model of a normal with fixed variance [6] or to an exact score frequency model like Baker and McHale’s NFL score predictor [2]. A good model should have sharper predictions (lower entropy) while maintaining calibration.

Once this pipeline is in place, we can price nearly anything:

- The probability a team covers a spread x is just $\sum_{m \geq -x+0.5} P(m)$ (if x is positive in favor of team A, and we consider half-point increments to avoid pushes).

- Teaser leg win probability is similar: if you tease a team from -8 to -2, that leg wins unless the team wins by 1 or loses the game. So $P(\text{leg wins}) = P(\text{margin} \geq -1)$ for that favorite.
- Parlay win probability for independent games is the product of win probabilities; for same-game combos, we use joint distributions or copulas to estimate the correlation.
- Expected value of a bet = win probability \times payout odds – loss probability. Using our model’s win probability and the book’s offered odds, we can flag bets with positive EV.

Many public tools (e.g. from Unabated, Action Network, etc.) essentially automate these calculations [1]. Our contribution is the underlying probability engine that feeds such tools with more accurate numbers, especially in situations the market may misestimate (like low-total games or unusual spreads).

3.2 Drive-Based Monte Carlo Simulation

An alternative (and complementary) modeling approach is to simulate the actual gameplay of a football contest. Instead of predicting final score directly, we simulate drives and scoring sequences. The idea is to naturally incorporate game mechanics such as possessions, field position, time, and decision-making (punts, field goal attempts, etc.), which in turn produce a distribution of final scores. For example, consider a simple Monte Carlo model:

- Represent each team’s offense and defense strengths, perhaps derived from our ratings (Stage A) or historical averages.
- Simulate quarter by quarter or drive by drive: starting from kickoff, sample the starting field position (touchback at 25, etc.), then simulate a drive outcome (touchdown, field goal, turnover, punt) using probabilities that depend on the offensive vs defensive strength and field position. Each simulated score adds to the game total, and we track time remaining.
- Include special teams and other quirks: higher chance of two-point conversions when down late, field goal range depending on kicker, etc. These details ensure that, for instance, a team down by 14 in Q4 is more likely to attempt risky plays (increasing chances of either scoring quickly or turning over, which could widen the final margin).
- Stop when the game time is exhausted. We then have a simulated final score.

By running, say, 50,000 Monte Carlo trials, we get a distribution of outcomes. This method is computationally heavier but has a key advantage: it can capture the tail behavior and correlation structure of scores better than a static distribution. It tends to naturally produce spikes at common margins without explicitly forcing them. For instance, a 3-point margin often arises in simulation because a team in field goal range on the final drive will tie or win by 3 if successful, etc. If our simulation properly reflects typical coaching decisions (e.g. when teams settle for FG vs go for TD), the resulting histogram of margins should show peaks at 3 and 7 similar to real data. We then calibrate the simulation to match observed

frequencies (adjust, for example, the likelihood of exactly 7 which might be a touchdown + extra point in simulation vs sometimes missed extra points in reality).

This drive-based model is especially useful for **live betting and derivative markets**. We can restart the simulation from any game state (score and time remaining) to get updated odds. It also can output joint distributions for things like (Team wins and total < 40) which helps price SGPs. For NCAA games, one might tweak the simulation with different pace (more plays, faster scoring) and wider variance in outcomes (a higher chance a drive results in a long TD or a turnover due to the skill disparity).

3.3 Same-Game Parlay Correlation Engine

While the two approaches above handle either independent game outcomes or the full distribution of one game, we sometimes need the joint probability of multiple bets within the same game. The aforementioned same-game parlay of side and total is a common example: the probability that the underdog covers +7 *and* the game stays Under 41 points. A straightforward method is to simulate games as in the drive model and record whether each leg of the parlay wins. By averaging over thousands of simulations, we directly get the joint probability (and hence the correlation between the events). Alternatively, one can use a bivariate probability model: model the game’s total points and margin jointly, perhaps with a correlation parameter. For instance, a high total might slightly reduce the chance of a very low margin (because more scoring tends to create separation, unless teams match each other). However, capturing these nuances analytically is difficult, so simulation is often preferred.

In practice, our ML pipeline can incorporate a correlation engine by extending the Stage B distribution. Instead of a univariate pmf for margin, we produce a two-dimensional pmf for (margin, total). This could be done by first modeling the total points distribution (often roughly normal or Poisson mixture) and then conditioning the margin on the total. We might assume that given the total $T = t$, the margin follows a certain discrete distribution (with perhaps smaller variance for very low t , since if few points are scored, margin is likely small too). By fitting such a model to historical play-by-play data, we can simulate correlated outcomes quickly. For example, the probability of (covering +7 & Under 41) can be computed by summing $P(\text{margin} \geq -7, \text{total} \leq 41)$ over the joint pmf.

The benefit of explicitly modeling correlation is that it exposes when sportsbooks misprice parlays. Books often use heuristic correlation adjustments or even cap payouts for popular combos. Our approach quantifies the true correlation. If the book overestimates correlation (thus lowering payout too much), there may be an edge in playing the SGP because the true joint probability is higher than assumed. Conversely, if they underestimate it, some SGPs might be very lucrative (though this is rarer since books fear correlated parlays).

In summary, our ML toolkit provides multiple avenues to evaluate bets:

- A fast analytical model (ratings + calibrated distribution) for quick pricing of standard bets.
- A detailed Monte Carlo simulation for scenario analysis and complex/time-dependent bets.

- A joint outcome simulator for correlated multi-leg bets in one game.

Each of these is validated against historical outcomes to ensure realistic probabilities. In the next section, we will add a Bayesian hierarchical layer that strengthens these models, especially in terms of stability and coherence over time.

4 Bayesian Hierarchical Modeling for Dynamic Team Strength

So far, we have discussed predictive models largely in a frequentist or machine-learning context. We now turn to a Bayesian hierarchical approach, which offers several advantages for sports betting models: (1) it allows partial pooling of information across teams and seasons (improving estimates for teams with sparse data, such as early in the season or smaller college programs), (2) it naturally incorporates parameter uncertainty into predictions (yielding probability distributions for our model outputs themselves), and (3) it provides a formal framework to update predictions as new data comes in (via Bayesian updating).

4.1 Hierarchical Model Structure

At the core of our Bayesian model, we assume that each team i in each week t has latent offensive and defensive strength parameters, $\alpha_{i,t}$ and $\delta_{i,t}$, which evolve over time. A simple structure could be:

$$\alpha_{i,t} = \alpha_{i,t-1} + \epsilon_{i,t}^{(O)}, \quad \delta_{i,t} = \delta_{i,t-1} + \epsilon_{i,t}^{(D)},$$

where $\epsilon_{i,t}^{(O)}$, $\epsilon_{i,t}^{(D)}$ are zero-mean random disturbances (reflecting changes in team performance due to roster changes, injuries, random variance, etc.). This is a state-space or random walk model for team strengths [9]. One can also center these around a global mean or a conference mean for college, to allow mean reversion. Home field advantage can be another parameter h_t that might itself evolve slowly over years (some have observed that NFL home field advantage has been declining, for example).

Now, given these latent strengths, we model game outcomes. A common approach is a *bivariate Poisson* or similar count model for (points scored by Team A, points scored by Team B). However, an equivalent formulation is directly modeling the point difference. We can write:

$$Y_{i,j,t} = (\text{score of team } i - \text{score of team } j) \text{ in game at time } t.$$

We model $Y_{i,j,t}$ as coming from a discrete distribution whose mean is $(\alpha_{i,t} - \delta_{j,t}) - (\alpha_{j,t} - \delta_{i,t}) + H \cdot I(\text{home} = i)$, where H is a home field parameter (so effectively $\alpha_i - \delta_j$ is the expected points i scores on j , and subtract the expected points j scores on i). This simplifies to $\mu_{i,j,t} = (\alpha_{i,t} + \delta_{i,t}^{(opp)}) - (\alpha_{j,t} + \delta_{j,t}^{(opp)})$ where $\delta^{(opp)}$ might be interpreted as defensive adjustments. It's easier conceptually to think in terms of offense vs defense: expected points for team i = exp(some base + $\alpha_{i,t} - \beta_{j,t}$) if using a Poisson model, where $\beta_{j,t}$ is opponent j 's defensive strength. Many authors (e.g., Baio and Blangiardo 2010 [11]) have used such hierarchical models for soccer, and similar ideas apply to football.

For the distribution of $Y_{i,j,t}$, one choice is a Skellam distribution (difference of two Poissons) [10], which inherently models the score difference. Another is to directly use a custom likelihood: e.g.,

$$P(Y = m) = \frac{\exp(\eta_m)}{\sum_k \exp(\eta_k)},$$

where η_m is a parameter (or linear predictor) for each possible margin m . This is like treating the outcome as a categorical with logits η_m that we want to learn. We could set $\eta_m = \theta \cdot f(m; \mu, \text{total})$ where $f(m; \mu, \text{total})$ is some basis function that captures that outcomes near μ are most likely, but also has extra peaks at key m . However, modeling each margin explicitly can be infeasible if we allow a wide range of m .

Instead, we typically assume points are scored via Poisson processes. For example, in the Bayesian framework, we might say:

$$\begin{aligned} X_{i,j,t} &\sim \text{Poisson}(\Lambda_{i,j,t}), \\ X_{j,i,t} &\sim \text{Poisson}(\Lambda_{j,i,t}), \end{aligned}$$

for points by i and j respectively, with

$$\log \Lambda_{i,j,t} = \text{attack}_{i,t} + \text{defense}_{j,t} + \log(\text{baseline pace}),$$

and similarly for $\Lambda_{j,i,t}$ [3]. The baseline pace might be something like a function of the total points market or league average. The sum of two independent Poissons is Poisson (with mean = sum), and the difference $X_i - X_j$ follows a Skellam distribution [10]. This approach inherently accounts for variance: if the teams are mismatched, the Λ 's will be far apart, yielding a skewed distribution of Y ; if they're equal, Y has a lot of mass near 0 (possibly a lot of ties in soccer, but in NFL ties at 0 margin are rare due to overtime; however, overtime rules mean a margin of 3 or 6 are quite likely in tied games).

We integrate this model in a Bayesian inference framework (using MCMC via Stan or PyMC). Priors are given for all parameters:

- Team strengths α, β (offense, defense) at season start could have broad priors around 0. We might set an NFL prior that most teams are within, say, ± 7 points of average.
- The random walk evolution variance: we treat it as a hyperparameter to be estimated. If the data suggests team strengths fluctuate a lot week to week, the posterior for this variance will be higher; if not, it will shrink toward zero changes.
- Home field advantage H gets a prior around 3 points (with uncertainty).
- For college, we may introduce a hierarchical prior by conference: each conference has a mean offense and defense level relative to an overall mean. A Sun Belt team and an SEC team might both be average within their conference, but the SEC conference mean might be higher than Sun Belt's, reflecting non-conference results.
- If using market spreads/totals as inputs, we can even treat the closing spread as an additional noisy observation of $\mu_{i,j,t}$ to "anchor" our predictions. Some recent work has explored dynamic models that incorporate betting market data as prior information [3].

The model is fit to historical data (e.g., several seasons of NFL or college results). The output is a posterior distribution for all unknowns: team strengths each week, home field, and crucially the predictive distribution of future games. From the posterior samples, we can directly compute probabilities: e.g., probability team i beats team j by at most 7 points (cover +7) is just $P(Y_{i,j} \leq 7)$ under the model, which we can approximate by simulation from the posterior predictive.

4.2 Benefits and Use-Cases

Why go through this complexity? Bayesian hierarchical models shine in situations with limited data or structural uncertainty:

- **Early season and small sample:** In Week 1–3 of NFL, or for a small-school college team, we don’t have much direct data. A hierarchical model leverages information from similar teams or league averages. It will shrink extreme predictions towards the mean until enough evidence accumulates. This prevents us from overreacting to one blowout in Week 1 (the model’s credible interval for that team’s strength will be wide, reflecting uncertainty).
- **Dynamic updating:** As the season progresses, the model automatically adjusts team ratings. If a star quarterback is injured and the offense struggles, the posterior for that team’s offensive strength will drop. Moreover, the Bayesian approach gives us a distribution over that strength each week, not just a point estimate, which we can propagate into our betting decisions (we might reduce bet sizes when there’s high uncertainty).
- **Joint outcome distributions:** The Bayesian model can naturally produce joint distributions for multiple games or multiple legs because it captures correlations through common parameters. For example, if we want the probability that *all* 6 teaser legs in a weekend win, we can simulate from the posterior predictive across all games and count how often all six cover. While one could also do this with a frequentist model, the Bayesian framework makes it straightforward to include parameter uncertainty in that simulation (so we’re not conditioning on one fixed estimate of team strength, but averaging over possible true strengths).
- **Bankroll and EV distribution:** Instead of a single EV for a bet, we get a distribution of EV given the uncertainty in probabilities. For instance, our model might say a teaser has an expected value of +2% of stake, but there is a 20% chance our probabilities are off and it’s actually negative EV (perhaps because one team is much better than we think). The Bayesian output would allow us to calculate that distribution, which could be used in a Kelly betting formula that accounts for parameter risk.

In practice, implementing this hierarchical model might involve some compromises for computational efficiency. We might run a full Bayesian update weekly (which can be slow for hundreds of teams in college). One approach is to use variational inference or assumed density filtering to approximate the posterior updates faster [4]. Another is to pre-compute a large set of prior distributions for key numbers frequencies (e.g., incorporate a prior that

around 15% of NFL games end in 3, with some uncertainty, and let the data slightly adjust it if needed for a given season’s conditions like new scoring rules).

Several studies in sports analytics have demonstrated improved prediction accuracy with such hierarchical models [3, 9]. We take inspiration from those but tailor the model to our betting use-cases: in particular, we ensure that the output is a full discrete distribution over margins (not just win probabilities or expected points). We also explicitly model the tail outcomes (blowouts vs close games) by allowing heavier-tailed distributions (e.g. a Student- t component for scoring, or using the Conway-Maxwell Poisson which can accommodate overdispersion [10]).

The result of our Bayesian modeling stage is a richer understanding of the game outcome distribution. It augments the ML pipeline by providing well-calibrated probabilities and by tying together games in a coherent way (so our probabilities for all games in a week are based on a common set of team ratings, which lends internal consistency; e.g., there won’t be two different estimates of Team A’s strength in two different games because one model is used globally).

Illustrative Example: Suppose after Week 4 our Bayesian model has learned that Team X has an elite defense (say $\delta_X = +5$ points above average) and mediocre offense ($\alpha_X = 0$), and Team Y has a good offense ($\alpha_Y = +3$) but poor defense ($\delta_Y = -2$). If these teams meet at Team X’s home, our model might predict something like a 21-17 win for Team X on average. But it will also yield a distribution: perhaps a 10% chance of a low-scoring game where Team X’s defense completely stifles Y (X wins by ≥ 10), a 15% chance Y’s offense shines and Y wins outright by a small margin, etc. The discrete margin probabilities might show 17% on X winning by 3, 12% on X by 7, 5% on Y by 3, etc. We would use these for any bets involving that game. Now imagine Team X’s quarterback gets injured in practice before Week 5; we can adjust the prior or input for α_X downward and run an update with one week of data or even a subjective input, yielding an updated distribution in time to evaluate bets for the weekend. This is the power of a Bayesian approach: it is flexible enough to incorporate new information at any point (even if not purely data-driven, one could input an informed prior for the backup QB’s effect).

5 Integrated Workflow and Strategy Deployment

We have described various components: alternative betting strategies, a discrete probability engine, ML prediction pipelines, and a Bayesian model. In this section, we outline how these pieces come together in an end-to-end workflow for a practitioner. This serves as a blueprint for implementing the system and using it to find +EV bets.

5.1 End-to-End Modeling and Betting Workflow

1. **Data Collection and Preparation:** Gather historical data for model training:

- Game results (scores, margins) for NFL and college, ideally with context (home/away, closing spreads, totals).

- Play-by-play or drive data if using the simulation approach, including time stamps and play outcomes.
- Team statistics, player injuries, depth charts, weather conditions for each game.
- Sportsbook lines (spreads, moneylines, totals, teaser payouts, alt-line prices) across multiple books, both current and historical (for analyzing how often books diverge and by how much).

Preprocess this data to create features for the ML models (e.g., rolling average yards/play, efficiency metrics) and to compute empirical frequency tables of score margins (for different total ranges, league, etc.).

2. **Predictive Model Training (Stage A and B):** Train the rating model and calibrate the distribution:

- Use data up to last season to fit the team rating model (this could be done in the offseason). For example, use 2018–2022 data to train, then test on 2023.
- If using Bayesian, either sample from the posterior for past seasons to ensure it fits well, or use MAP estimates for speed.
- Derive the continuous baseline distribution of margins (e.g., determine that historically the standard deviation of NFL score differences is about 13.5 points [6], or fit a small model that predicts variance as a function of total).
- Compute the adjustment weights for key numbers by comparing the baseline distribution to actual outcomes. This can be a regression: predict $\mathbf{1}_{\{\text{margin}=m\}}$ from features like “is $m = 3$ ” etc., to find the average probability uplift for each key number.
- Validate that the combined Stage A + Stage B accurately reproduces margin distributions on held-out seasons.

3. **Bet Pricing Module:** With the predictive model in place, build functions to price any bet:

- Given a game and a line (spread or total), return win probability (for a spread, probability favorite covers; for total, probability over hits, etc.).
- Given multiple games and lines (for parlays/teasers), return joint win probability (assuming independence, multiply, or if same-game or correlated, use appropriate joint model).
- Given a teaser specification (e.g., 2-team, 6-point), compute break-even probability per leg (as $\text{BE}^{1/n} = (1/\text{payout})^{1/n}$ in decimal terms) [1]. Then flag which legs have model win probability above that threshold.
- Incorporate sportsbook-specific payout schedules. Some books might pay +100 for a 3-team teaser, others +180, etc. The module should ingest those to compute EV.

4. **Line Shopping and Identification:** In real time (say each Tuesday when opening lines come, and then continuously):
 - Pull current odds from various sportsbooks APIs or websites for spreads, totals, alt lines, teaser payouts, SGP offers.
 - For each available bet, use the pricing module to compute the model's win probability and hence EV.
 - Create a ranked list of potential +EV bets. For example, you might find that teasing Team A +2.5 (with total 38) and Team B +2 (total 41) at Book X's -120 teaser is +5% EV according to your model. Or that parlaying Team C -2.5 alt and Team D +8.5 alt yields +3% EV over the offered odds.
 - Also consider derivative strategies: Are there good middling opportunities? The program can check if any pair of lines across books suggests a middle (e.g., one book has Team E -7.5 -110, another has Team F +10.5 -110 for the same game, giving a 3-point middle; our model might estimate maybe a 10% chance of the game landing 8,9,10 for a double win, which could be profitable if combined cost is low).
 - Output these opportunities with details.
5. **Bet Selection and Sizing:** For each positive EV opportunity, decide whether to bet and how much:
 - We might apply a Kelly criterion based on the model's EV and variance. Since we have Bayesian outputs, we could compute not just an EV point estimate but a distribution. A conservative approach might bet a fraction of Kelly to account for estimation error.
 - Also consider portfolio correlation: if many of our edges rely on similar outcomes (e.g., a lot of bets on the same underdog in different forms), we might limit exposure.
 - In practice, we might limit bet size to, say, 1–2% of bankroll on any single teaser or parlay due to their higher variance.
6. **Monitoring and Update:** As games are played:
 - Update the team ratings (Stage A). If using Bayesian, run the posterior update after each week (or even use particle filters to update live). If using an ML model, retrain or at least incorporate the new data point.
 - Track the performance of bets: are the edges realized? If not, investigate if the model is miscalibrated or if variance is just playing out. For example, if we consistently miss on games with high totals, maybe our adjustment for totals is off.
 - Adjust the model if needed (recalibrate, retrain certain components with more data).

- Keep historical logs to further refine. For instance, maybe certain books react faster, so by the time we try to bet, the line has moved. Our system could learn to predict line moves (i.e., if our model is way different from current line, often the line will move toward our number).

This workflow essentially turns the theory into practice. Each step can be automated to a large degree, creating a sort of “betting bot” guided by the discrete-margin model and its enhancements. Human judgment remains important, especially in interpreting when to trust the model vs when to override (e.g., major injury news might not be fully captured quantitatively).

5.2 When to Use Each Strategy

Not all the strategies from Section 2 are optimal at all times. Here we provide a quick guide on which approach tends to shine in which scenario:

- **Classic Wong teaser (2-team, 6-pt):** Best in NFL when totals are low (say ≤ 45) and you can cross both 3 and 7. Only play if you can get a favorable price (around -120 or better for 2-teamers, or in a 3+ team teaser that lowers the per-leg break-even). Our model will highlight these by giving high cover probabilities on those teased lines. If a book still offers -110, these are usually auto-bets [8].
- **Alt-spread parlays:** Great when standard teasers are overpriced. For example, if books universally moved to -140 for a 2-team teaser, it’s likely more EV to identify two alt spreads that replicate a teaser and parlay them, often yielding -130 or so. This also shines when one leg you want to tease isn’t allowed or is off the board, but an alt can be found. Our tools check both options and we take whichever has higher EV or lower variance.
- **College football betting:** We almost never do 6-pt two-team teasers in college; instead, we might consider 13-pt super teasers (3 teams) if anything, but generally the variance is too high. Strategies focusing on correlation (like pairing an underdog with under total) tend to pay off more in college, because sometimes books price those correlations naively. Also, live betting or quarter betting edges can be bigger in college due to slower bookmaker adjustments on Saturday when dozens of games are live.
- **Live middling:** Particularly useful on NFL primetime games or playoffs where liquidity is high. If the game starts low-scoring (total plummets) but one team still has, say, a 10-point lead, the chance of a middle around 3 or 7 increases. We use our simulator to quantify this quickly. This is an opportunistic strategy; you can’t plan pre-game, but our model is running during the game to alert if, say, “Take Team A +4.5 live now while holding Team B +7 from pre-game” has a high overlay.
- **Pleasers:** Only in rare cases, maybe Week 17 NFL where a huge favorite might either cover easily or lose outright if resting players. Or if you strongly handicap that a particular spread is very mispriced and the underdog is much better than thought, a pleaser could maximize the payoff of that opinion. We only fire on pleasers if our model

probability for the moved line is significantly above what the book implies (which is uncommon, but not impossible).

- **Same-Game Parlays:** Use when you identify a strong game script angle. For instance, you think a certain underdog will either keep it very close in a defensive battle or get blown out in a shootout. Instead of a straight bet, you might do a small parlay on (Underdog +7 and Under 42) as well as (Favorite -14 and Over 52) to capitalize on the bimodal belief. Books often can't perfectly price such combinations, especially involving props (like a RB over yards with his team + points, etc.). Our correlation engine guides these picks by quantifying reasonable ranges. SGPs are a high-variance, high-reward play and should be sized smaller, but can be +EV if done selectively.

Naturally, the best approach is to compare expected values. Our system does that, so many of these guidelines are built into the recommendations it spits out. But the above highlights the intuition of where the “easy money” used to be (Wong teasers), and where it may still be hiding (creative parlays and in-game bets).

5.3 Practical Tools and Recipes

Finally, we outline a few concrete tools that can be derived from our modeling framework. These are useful both for understanding the model's implications and for day-to-day betting decisions:

- **Half-Point Value Table:** Using the discrete distribution, we calculate the value of buying or selling a half-point at each integer spread. For example, what is the change in win probability when moving from -3 to -2.5 (buying the hook off 3)? In NFL, that might be about +4-5% win probability (because pushes on 3 become wins) [7]. We produce tables for each key number and for different total ranges. This can be used to evaluate if buying a half-point at the book's price is worth it, or to rank teaser legs. For instance, if going from +2.5 to +8.5 yields +20% win probability and the next best leg is +1.5 to +7.5 with +18%, you'd prioritize the former.
- **Fair Teaser Pricing Calculator:** A small tool where you input two (or more) legs and it returns the fair odds for the teaser. For each leg i , compute $p_i = P(\text{leg } i \text{ covers teased spread})$ from the model. Then the fair parlay probability for all legs is $\prod_i p_i$. Convert that to fair odds: $\text{odds}_{\text{fair}} = \frac{1}{\prod_i p_i} - 1$ (in decimal, subtract 1 then convert to American if needed). For example, if $p_1 = 0.75$, $p_2 = 0.74$, $\prod p_i = 0.555$, fair odds = $1/0.555 - 1 = 0.80$ in decimal, which is -125 in American odds. If a book offers -120, we have an edge. This calculator essentially replicates what some betting tools do [1], but using our custom probabilities.
- **Alt-Line Parlay Comparator:** An extension of the above, this tool takes a target teaser scenario and finds equivalent alt spreads. For example, you input: “I want to tease Team X -8.5 to -2.5 and Team Y +2 to +8” and the tool finds the closest alternate lines: maybe Team X -2.5 and Team Y +8.5, and then computes the parlay odds and compares to the teaser odds. It will show which is better and by how much. It can also

suggest slightly different alt lines if those are mispriced; e.g. maybe Team Y +7.5 is actually a bigger jump in probability than the payout drop vs +8. This helps exploit any inefficiencies in the alt market relative to the fixed teaser pricing.

- **Bayesian Update Script (Weekly):** A reproducible routine (possibly in Stan or PyMC3) that takes the latest week’s games, updates the posterior on team strengths, and outputs the new distribution of margins for upcoming games. It could run overnight after Sunday games and produce Tuesday morning odds that incorporate Monday Night Football as well. In practice, this could augment the ML predictions: for example, use the Bayesian model’s median predictions as inputs to the Stage B distribution instead of pure ML. The script would also output interesting quantities like the probability each team makes playoffs, etc., but for our purposes, its main output is improved game win probabilities. We might integrate this with market odds: e.g., use the closing Vegas spreads of week t as an input to the Bayesian update for week $t + 1$ (informative prior), essentially blending wisdom of crowds with our model.

Each of these recipes is intended to be almost “plug-and-play.” A trader or analyst could use the half-point table and fair teaser calculator on the fly when considering a bet, without running a full simulation each time. Meanwhile, the heavy lifting (generating the table or calibrating the calculator) is done by our comprehensive model in the background.

6 Conclusion

We have presented an integrated framework for analyzing point-spread related bets in American football, built around a discrete margin probability model and enhanced by both machine learning methods and Bayesian statistics. This toolkit addresses the challenges of the modern betting market – where simple strategies like Wong teasers have become less profitable due to bookmaker adjustments – by offering more nuanced and flexible approaches. By modeling the game outcome distribution at a granular level (every possible score margin), we unlock the ability to price a wide array of bets consistently. The inclusion of alternative strategies (like alt-line parlays and correlated plays) ensures that wherever the value shifts due to market changes, our framework can identify it. The machine learning pipelines provide the scalability and adaptability to ingest a multitude of factors (from team stats to weather to live game states), while the Bayesian layer adds stability and rigorous uncertainty quantification.

Moving forward, one could extend this work to other sports or bet types. The general principle of exploiting discrete outcome distributions applies to any sport with key score differentials (e.g., basketball’s 5-point plays or soccer’s 1-goal games in totals). Furthermore, the synergy between data-driven modeling and betting strategy suggests that as more data becomes available (player tracking, play-level success rates), models will improve, but so will the market efficiency. Our approach is positioned to continually recalibrate and find edges in the ever-evolving landscape of sports betting. In summary, success in this domain comes from marrying deep domain knowledge (like knowing why 3 and 7 matter) with sophisticated modeling and relentless shopping for value. This paper provides a roadmap and tools for doing exactly that.

References

- [1] Andrews, Jack. (2024). *How To Build Profitable NFL Teaser Bets*. Unabated.com, Aug 21, 2024.
- [2] Baker, Ryan D., and Ian G. McHale. (2013). Forecasting exact scores in National Football League games. *International Journal of Forecasting*, **29**(1), 122–130.
- [3] Egidi, Leonardo, Francesco Pauli, and Nicola Torelli. (2018). Combining historical data and bookmakers’ odds in modelling football scores. *Statistical Modelling*, **18**(5-6), 436–459.
- [4] Macrì-Demartino, Roberto, Leonardo Egidi, and Nicola Torelli. (2025). Bayesian weighted discrete-time dynamic models for association football prediction. *arXiv preprint arXiv:2508.05891*.
- [5] Moxley, Chris. (2023). *College Football Betting 101 – Understanding The Fundamentals*. Campus2Canton.com, May 3, 2023.
- [6] Stern, Hal S. (1991). On the probability of winning a football game. *The American Statistician*, **45**(3), 179–183.
- [7] Walsh, P.J. (2024). *Key Numbers in NFL Betting, Explained*. Action Network, Aug 09, 2024.
- [8] Wong, Stanford. (2001). *Sharp Sports Betting*. Pi Yee Press.
- [9] Owen, Alun. (2011). Dynamic Bayesian forecasting models of football match outcomes with estimation of the evolution variance parameter. *IMA Journal of Management Mathematics*, **22**(2), 99–113.
- [10] Karlis, Dimitris, and Ioannis Ntzoufras. (2009). Bayesian modelling of football outcomes: using the Skellam’s distribution for the goal difference. *IMA Journal of Management Mathematics*, **20**(2), 133–145.
- [11] Baio, Gianluca, and Marta Blangiardo. (2010). Bayesian hierarchical model for the prediction of football results. *Journal of Applied Statistics*, **37**(2), 253–264.