Summary of Considerations, Hypotheses, Methods, and Principles Used in the Model

1. Core Framework

Monte Carlo Simulation: Simulates a large number of race outcomes (100,000 iterations) to estimate empirical probabilities, incorporating randomness and dynamic race conditions.

Bayesian Updating: Uses a Beta distribution as a prior belief about each horse's win probability, continuously updated after each simulated race to generate posterior probabilities.

2. Probability Distributions

Levy Distribution: Models random variation in horse performance, capturing heavy tails and occasional extreme outcomes (suitable for unpredictable athletic performance).

Poisson Distribution: Models recent form and performance trends, where a horse's past win rate influences expected future performance via Poisson adjustments.

Beta Distribution: Models win probabilities for Bayesian updating. The distribution is updated dynamically based on simulated outcomes, reflecting evolving knowledge.

3. Dynamic Race Conditions

Weather and Track Bias: Randomly applies a performance adjustment based on weather conditions (e.g., dry or wet). Each horse has an individualized track bias parameter.

Temperature Variability: Incorporates the effect of environmental conditions through a normal distribution (mean 20°C, standard deviation 5°C).

Fatigue Model: Introduces a penalty proportional to the horse's fatigue rate and the number of previous races.

4. Performance Modifiers

Jockey Skill: Adjusts horse performance by a jockey skill coefficient, affecting overall speed.

Pace Factor: Allows dynamic variation in the horse's speed during the race, reflecting tactical differences in pacing (higher factors speed up performance).

5. Betting Structures

Fixed-Odds Market (Bookmaker): Computes expected value (EV) and optimal bet sizing using the Kelly Criterion, adjusting for dynamic bookmaker odds.

Pari-Mutuel (Tote) System: Simulates pari-mutuel payouts where total winnings depend on the betting pool, with a 15% house takeout.

6. Bayesian Probabilities and Model Refinements

Prior Belief: Derived from past performance (win count and races run) using a Beta distribution.

Posterior Update: Adjusts belief after each simulated race. Winners increase their success count, and others increase their failure count.

Implied Odds Calculation: Converts Bayesian posterior probabilities into fair decimal odds.

7. Optimization Techniques

Kelly Criterion: Maximizes long-term bankroll growth by calculating the optimal bet fraction based on the discrepancy between true probability and bookmaker odds.

Expected Value Calculation: Evaluates the profitability of each bet by comparing true probability against market odds.

8. Statistical and Mathematical Considerations

Heavy-Tailed Modeling: Uses Levy and Cauchy distributions to account for extreme outcomes and unexpected results.

Conditional Probabilities: Each horse's win likelihood is adjusted based on dynamic and race-specific conditions, improving prediction accuracy.

This model integrates advanced probability theory (Bayesian inference), dynamic performance adjustments, and multiple betting frameworks, providing a comprehensive simulation and strategic optimization tool for horse racing analysis

---

To incorporate Bayesian and conditional probabilities into the model, we will:

Update prior beliefs based on past performance and external conditions.

Compute likelihoods using race-specific factors (fatigue, track bias, pace).

Apply Bayes' Theorem to adjust posterior probabilities dynamically

Pari-Mutuel Betting (Bettor vs. Bettor)

How it works: All bets of a particular type (e.g., win, place, show) are pooled together. The house (track operator) takes a percentage (called the "takeout"), and the remaining pool is divided among the winning bettors.

Bettors compete against each other, not the house or bookmaker.

Odds fluctuate based on how much money is wagered on each horse.

Common in: Most legal horse racing tracks worldwide (e.g., USA, France, Japan).

Fixed-Odds Betting (Bettor vs. Bookmaker)

How it works: You lock in a fixed price (odds) when you place your bet. The bookmaker sets the odds and takes on the risk.

Bettors compete against the bookmaker.

Odds are fixed when you place the bet, so your payout won't change.

Common in: UK and other regions with private bookmakers or online sportsbooks.

I've added track bias and jockey skill adjustments to the simulation. These modifiers directly affect each horse's performance:

Track Bias: Positively or negatively influences race time based on the horse's affinity for specific track conditions.

Jockey Skill: Adjusts for rider proficiency, where higher values reduce time (better performance).

I have implemented a Poisson adjustment where each horse's performance is modified based on their historical win rate. This adjustment reflects the impact of past wins on their race time.

I expanded the simulation to include:

Poisson Adjustment: Performance is now influenced by past wins, simulating the impact of race experience.

Expected Value Calculation: Compares simulated probabilities against market odds to assess profitable bets.

Next Steps:

Model Construction: Incorporate logistic regression and Poisson models to predict win probabilities from historical data.

Further Analysis: Extend the model to account for track conditions, jockey performance, and hedging strategies.

Monte Carlo Simulation for Race Scenarios

Monte Carlo simulations estimate probabilities by running thousands of simulated races.

Logistic Regression (Horse Ranking)

Models the relationship between independent variables (e.g., speed, weight) and the binary outcome (win/lose)

Poisson Distribution (Race Outcomes)

Models the probability of rare events (e.g., a horse winning a race).

Useful when estimating the likelihood of outcomes from past performance.

Statistical Modeling for Odds Calculation

A robust betting model evaluates factors affecting performance. Common techniques:

Poisson Distribution (Race Outcomes)

Models the probability of rare events (e.g., a horse winning a race).

Useful when estimating the likelihood of outcomes from past performance.

If a horse has won 3 out of 10 races, the expected wins (λ) = 0.3.

Probability the horse wins next race:

$\mathrm{P}(\mathrm{X} = 1) = \lambda\, e^{-\lambda}$= 0.22 or 22%

Monte Carlo Simulation for Race Scenarios

Monte Carlo simulations estimate probabilities by running thousands of simulated races.

Example: Simulate a 3-Horse Race

Assign performance parameters (speed, variance).

Simulate the finish order across 10,000 trials.

Estimate winning probabilities and compare them to market odds.

Advanced Betting Considerations

Follow Professional Money: Identify horses with late odds compression.

Diversify Bets: Spread wagers across races using fractional Kelly.

Market Inefficiencies: Exploit mispriced horses by combining historical and real-time data.

Timing Strategy: Place bets as late as possible to incorporate sharp bettor insights.

By combining these methodologies and understanding the dynamics of different bettor categories, a mathematically sound and strategically advanced betting plan emerges for maximizing long-term profitability.

In reality, initial odds in pari-mutuel horse racing are determined by the track's oddsmaker through the 'morning line,' which is an estimate of how the public is expected to bet. These odds are not binding but serve as a guide for early bettors. At the true starting point—before any bets are placed—each horse has equal implied probabilities, meaning if there are 10 horses, each would theoretically start at 9/1 (implying a 10% chance of winning, ignoring the house's takeout). As wagers accumulate, the odds dynamically adjust purely based on the distribution of total money wagered on each horse. This bettor-against-bettor structure means the odds reflect collective bettor sentiment rather than a house-fixed edge. Bookmakers do not take positions against the bettors but instead collect a fixed percentage of the betting pool (takeout). Thus, the odds continuously update to reflect the current weight of money in the pool. Early odds are often inefficient and reflect casual sentiment, while sharp bettors exert influence closer to race time, compressing odds on horses with perceived value.

In a bettor-against-bettor structure, such as pari-mutuel betting in horse racing, odds are determined by the proportion of total wagers placed on each horse. As more money flows to a particular horse, its odds shorten (decreasing the potential payoff), while less-favored horses see lengthened odds. This structure reflects collective market sentiment rather than the bookmaker's opinion. The track earns revenue via a takeout percentage, meaning there is no risk exposure for the house. Late professional bets significantly influence odds compression, while early odds often reflect casual bettor sentiment and are more volatile.

In contrast, in a structure where bookmakers take positions against bettors, such as fixed-odds betting, the bookmaker sets initial odds based on their proprietary models and market data. These odds are adjusted to manage the bookmaker's risk and maximize profitability. If too much money is wagered on a specific outcome, the bookmaker adjusts the odds to reduce liability. In this model, bookmakers face potential losses if their odds are inaccurate. Odds movements in this system are driven by both public sentiment and the bookmaker's risk management strategies. Professional bettors exploiting mispriced odds can create asymmetric risks for bookmakers, leading to sharper adjustments in real-time.

Dynamic Odds Adjustment by Bookmakers

Bookmakers actively manage their risk exposure by adjusting the odds based on the flow of wagers and external information. This process, known as odds shading or line movement,

ensures that the bookmaker maintains a profit margin (the "overround" or "vig") while balancing their liabilities.

Key factors driving dynamic odds adjustments:

Betting Volume: If a significant amount of money is wagered on one horse, its odds will shorten (decreasing payoffs) while the odds of other horses lengthen (increasing payoffs). This adjustment reduces the bookmaker's liability on popular outcomes.

Market Sentiment: Odds can shift if influential bettors (like sharps or syndicates) place large, informed bets, signaling new information.

External Information: Changes in track conditions, jockey substitutions, or unexpected news about a horse's health can cause odds to fluctuate.

Competitive Response: Bookmakers monitor other markets and adjust their odds to remain competitive and attract balanced betting.

How Odds Are Adjusted in Practice

Opening Line: Based on proprietary models, past performance, and market conditions.

Live Adjustments: As bets come in, odds change to balance exposure and reflect real-time information.

Closing Line: Final odds before betting closes—typically the most accurate reflection of true probabilities due to late sharp money.

Factors Influencing Odds Changes

In the days leading up to the race, odds fluctuate based on betting activity. In pari-mutuel betting systems, common in horse racing, odds are determined by the proportion of total bets placed on each horse. As more money is wagered on a particular horse, its odds shorten, indicating a higher probability of winning. Conversely, less popular horses see their odds lengthen.

Betting Timeline and Cut-off:

Bettors can place wagers up until a few minutes before the race begins. Typically, betting closes approximately 2-3 minutes prior to the start to allow for final calculations and odds adjustments. This window provides an opportunity to observe late betting trends, which can be indicative of insider information or professional betting activity.

Optimal Betting Strategies:

Drawing inspiration from renowned mathematicians Claude Shannon and Edward Thorp, the Kelly Criterion offers a systematic approach to bankroll management. The formula calculates the optimal bet size to maximize logarithmic growth of wealth.

Advanced Hedging Techniques

In-Race Hedging (Live Betting)

Hedge positions by wagering on different outcomes during a race as odds fluctuate.

Useful when your selected horse takes an early lead and the market shortens its odds.

Advanced Betting Strategies in Horse Racing

Introduction

Horse racing, particularly in prestigious events like the Kentucky Derby or other Grand Prix-style tournaments, provides a fascinating landscape for the mathematically inclined bettor. This case study examines the evolution of betting odds from five days before the race to the final 30 seconds, integrates professional betting strategies, and develops an optimal allocation of a fixed betting bankroll across multiple races using advanced probabilistic methodologies.

Structure of the Horse Race

Key Parameters:

Number of Horses: 10 horses in the race.

Betting Duration: Bets open until approximately 30 seconds before the race begins.

Odds Calculation: Parimutuel betting system (bettors wager against each other, not against the house).

Grand Prix Format: Typically, 12 to 14 races occur in a single race day.

How Odds Are Formed and Updated

Parimutuel System Mechanics:

Odds reflect the proportion of total money wagered on each horse.

Time Window for Odds Calculation: Odds are continuously updated based on incoming wagers. X = 7 to 14 days represents the typical window where public money shapes early odds. Professional money heavily influences the final 15 to 60 minutes before betting closes.

Statistical Factors Influencing Results

Horse-Specific Factors: Breed genetics, past performance on similar track conditions, injury history.

Jockey Influence: Jockey win rate, prior experience with the horse, and performance under high-stakes conditions.

Track & Weather Conditions: Wet vs. dry track, temperature, wind speed.

Betting Market Movement: Last-minute professional wagers that drive significant odds shifts.

Insights from Professional Handicappers

Andy Beyer: Developed the Beyer Speed Figures, a quantitative measure of horse performance.

Steven Crist: Advocated for vertical and horizontal wagers using exotic bets.

Edward Thorp: Applied the Kelly Criterion to parimutuel markets.

Andy Beyer

Andy Beyer developed a quantitative system known as the Beyer Speed Figures, which measures a horse's performance in past races. This system adjusts for the race's inherent difficulty by accounting for the track conditions and other environmental factors. Beyer's

methodology begins by calculating the horse's raw running time and then adjusting it to reflect the track's speed on the day of the race. This adjustment is crucial because identical times on different days may not reflect equal performances if the track conditions vary.

He introduced the concept that not all victories or losses should be valued equally. Instead, a horse that runs well on a slow track may deserve more credit than one that performs similarly on a fast track. By standardizing the performance across different conditions, Beyer Speed Figures provide a clearer comparison of horse abilities. This insight allows professional bettors to identify underappreciated horses whose past performances may have been underestimated due to external variables. Beyer's approach emphasizes that analyzing performance requires adjusting for underlying conditions, leading to more accurate predictions of future races.

Steven Crist: Advocated for vertical and horizontal wagers using exotic bets.

Edward Thorp: Applied the Kelly Criterion to parimutuel markets.

Insights from Professional Handicappers

Andy Beyer: Andy Beyer developed a quantitative system known as the Beyer Speed Figures, which measures a horse's performance in past races. This system adjusts for the race's inherent difficulty by accounting for the track conditions and other environmental factors. Beyer's methodology begins by calculating the horse's raw running time and then adjusting it to reflect the track's speed on the day of the race. This adjustment is crucial because identical times on different days may not reflect equal performances if the track conditions vary.

He introduced the concept that not all victories or losses should be valued equally. Instead, a horse that runs well on a slow track may deserve more credit than one that performs similarly on a fast track. By standardizing the performance across different conditions, Beyer Speed Figures provide a clearer comparison of horse abilities. This insight allows professional bettors to identify underappreciated horses whose past performances may have been underestimated due to external variables. Beyer's approach emphasizes that analyzing performance requires adjusting for underlying conditions, leading to more accurate predictions of future races.

Steven Crist: Steven Crist was a leading advocate for using complex betting structures known as vertical and horizontal wagers. Vertical wagers involve predicting the exact order of finish within a single race. Examples include exactas, where the bettor must select the first two horses in the correct order, and trifectas, which require predicting the first three. Horizontal wagers span multiple races, requiring bettors to correctly choose the winner across a sequence of races, such as in the pick-three or pick-six format.

Crist emphasized these exotic bets because they offer significantly higher payouts compared to simple win, place, or show bets. He recommended a methodical approach where bettors identify undervalued horses to include in their combinations. By focusing on races where the public underestimates specific outcomes, a bettor can create multiple combinations to increase the probability of hitting a lucrative payout. This approach involves balancing potential rewards with the risk of spreading the bankroll across many combinations, while prioritizing situations where the public's misjudgment skews the odds in favor of the informed bettor.

Edward Thorp: Applied the Kelly Criterion to parimutuel markets.

Code Language

The horse racing simulation is written in Python.

Libraries and Packages Used

numpy

Numerical operations, random sampling, and simulations.

pandas

Data manipulation and storage (for potential extensions or output handling).

matplotlib.pyplot

Data visualization (for plotting Bayesian win probabilities).

scipy.stats

Advanced statistical distributions (e.g., Poisson, Levy, Cauchy, Beta, Zipf).

Key Functions and Methods Used (Function/Method; Source; Purpose)

np.random.seed()

numpy

Ensures reproducibility by fixing the random seed.

np.random.choice()

numpy

Randomly selects values with given probabilities (e.g., weather conditions).

np.random.normal()

numpy

Samples from a normal (Gaussian) distribution (e.g., temperature modeling).

np.random.poisson()

numpy

Samples from a Poisson distribution (recent form influence).

levy.rvs()

scipy.stats

Samples from a Levy distribution (models heavy-tailed randomness).

beta.mean()

scipy.stats

Computes the mean of the Beta distribution (posterior probabilities).

plt.bar() and plt.show()

matplotlib.pyplot

Visualizes Bayesian win probabilities.

Here are additional libraries and methods you might consider incorporating for further refinement and extension of the model:

NumPy (numpy):

Core library for numerical computing and random sampling.

Functions used:

np.random.seed() – Ensures reproducibility.

np.random.choice() – Random sampling with specified probabilities (e.g., weather).

np.random.normal() – Samples from a normal distribution (e.g., temperature variation).

np.random.poisson() – Samples from a Poisson distribution (e.g., recent form adjustments).

np.random.uniform() – Samples from a uniform distribution (e.g., pace adjustments).

SciPy (scipy.stats):

Provides access to advanced statistical distributions and methods.

Distributions used:

levy.rvs() – Generates random variables from the Levy distribution (models heavy-tailed performance).

cauchy.rvs() – Samples from a Cauchy distribution (another heavy-tailed alternative).

zipf.rvs() – Samples from the Zipf-Mandelbrot distribution (for rank-based modeling).

beta.mean() – Computes the mean of the Beta distribution (for Bayesian updating).

Pandas (pandas):

Flexible data manipulation and analysis library.

Used here to manage and store horse data efficiently.

Matplotlib (matplotlib.pyplot):

Visualization library for plotting results and distributions.

Functions used:

plt.bar() – Plots Bayesian win probabilities as a bar chart.

plt.show() – Displays the generated plot.

Additional Libraries for Further Extensions:

sympy: For symbolic math and closed-form solution checking.

scikit-learn: For machine-learning-based predictive modeling if you want to add ML predictions.

tensorflow or pytorch: For neural networks to model complex horse-racing patterns.

HORSE RACING SIMULATION

https://chatgpt.com/canvas/shared/67b73529bb5c8191ab39eebd5380dfdf