The Math Behind Bookmaking

Bookmaking is the process of turning a game of probabilities into a profitable outcome for businesses. In this handout, we will explore why betting is a game you're designed to lose in and explore the mathematics behind this process.

1 Terminology

First, we introduce common terms used in betting.

- **Straight bet**: a wager on a single outcome of an event.
- **Parlay**: a combination of straight bets that increases payouts if most or all bets are correct.
- **Legs**: the individual events that make up a parlay (ex. a six-leg is a parlay consisting of 6 straight bets).
- **Favorite**: the more probable outcome.
- **Underdog**: the less probable outcome.
- **Lines**: what the bookmaker sets a projection of a player at. For example, Player A to score 25.5 points.
- **Odds**: the probability of an event happening. There are three main types of odds: fractional, decimal, and moneyline (aka American) odds.
    1. *Fractional odds*: listed as a fraction a/b (a to b). a is the amount of money you win by wagering b dollars. If the people who came up with the term fractional odds were mathematicians, they might have used the term *expected value* instead. However, this is the term people in the gambling community came up with.
    2. *Decimal odds*: the factor by which you multiply your wager to get your total payout, including both your original and winning amount.
    3. *American odds*: favorites are indicated by a negative sign (-) and underdogs are indicated by a positive sign (+). For favorites, the amount that follows the negative sign is the amount you need to place to profit \$100. For underdogs, the amount that follows the positive sign is the amount you profit by betting \$100.

Problem 1. Suppose we have the following fractional odds for what the weather will be tomorrow:

- sunny: $5/4$
- cloudy: $2/1$
- rainy: $7/2$

You submit an entry for \$10 on cloudy, and it turns out that the weather is cloudy tomorrow. How much is your return, which is the amount you win plus the amount you place?

2 Expected Value Revisited

Recall that we previously learned about the concept of *expected value*. Define a **contest** to be a competition with a set prize pool such that the events are disjoint and cover the sample size. In other words, a contest satisfies the following conditions:

- S is a sample space consisting of a finite number of disjoint events A, B, C, ... such that the union of these events is S
- The player is only allowed to choose one event

Problem 2. Suppose there are a finite number of events in our contest. How can we compensate for those who take a chance on the less likely events?

Problem 3. Suppose that the fractional odds for event A is x/y. If we assume that our bookmakers are fair (which we call running a *fair book*, defined by setting odds so that the probabilities of every possible independent event adds up to 100%). Based on your answer to Problem 2, write a formula in terms of x, y, and $P(A)$ to ensure that on average we should be getting back the same amount that we invest.

Problem 4. The value of $P(A)$ from Problem 3 is known as the **implied probability**. How do we convert from fractional odds to implied probability?

Problem 5. Using the odds from Problem 1, prove or disprove that the odds form a contest as defined above.

Problem 6. How do we convert from fractional odds to decimal odds? Convert the following and determine its implied probability:

- 7/2
- 1/4
- 8/3

Problem 7. Convert the following from American odds to decimal odds. *Hint: Use fractional odds. When necessary, round to two decimal places.*

- +1000
- +120
- -115

Problem 8. Building a parlay involves combining independent events. Continue to assume that the book is fair. Which form of odds is the easiest to use to determine the total payout of a parlay, and how should we do so?

Problem 9. Using the idea you developed in Problem 8, determine the American odds for the following parlays:

- +300, -120
- +120, +200, +1000

3 (Un)fair Books

As mentioned in the previous problem, assuming that the events in a given book are independent is important. In the real world, however, some events are hard to separate from each other. For example, if we think that a certain basketball player will score 30+ points, it is unlikely that his teammate will do the same, and hence pairing the 'over' on one player and the 'under' on another player is a good strategy (here, *over* and *under* refer to the player scoring more than or less than their projection, respectively). To increase users' capacity and flexibility, books have developed the ability for users to combine picks from the same event. Unfortunately, the primary goal of bookmakers isn't to make everything convenient for you - but to earn money.

Problem 10. A person has to spend time and money to create a book. If a person or business is only seeking to maximize the amount of money they make with their time, will they run a fair book? What do you think this person or business might want to do if they want to offer a book and maximize the amount of money they make?

Let's explore an example illustrating Problem 10. Suppose we have a soccer match that has a $\frac{1}{2}$ chance of team A winning, $\frac{1}{6}$ chance of team B winning, and $\frac{1}{3}$ chance of the game ending in a draw. As a bookmaker, suppose we wanted to maximize our profit. We decide to increase the implied probabilities of team A winning to $\frac{3}{5}$. (First, understand why increasing the implied probabilities helps achieve this goal.)

Problem 11. What do the implied probabilities of team B winning and the match ending in a tie have to be and what is the sum of all of the implied probabilities?

We call the value by which the implied probability exceeds 100% a **bookmaker margin** or **overround**. In an ideal situation, the bookmaker expects the amount of money wagered by users to match the implied probability ratios.

Problem 12. If the total amount placed on this match is \$120, how much do the bookmakers profit?

4 You Are The Underdog

In this section, we will examine how a typical fixed book functions. A **fixed book** can be described as follows. To help streamline the process, companies have chosen to have fixed payout models and try to maintain projections for events to be at 50-50 odds. We will explore how much these changes affect the users. In our example, assume that the bookmaker has a large amount of 50-50 events and the number of possible events on a parlay ranges from 2 to 4.

Standard payout multipliers for these parlays are 2-legs at 3x, 3-legs at 5x, and 4-legs at 10x (note: total payout includes both entry fee and winning amount and these multipliers act on your entry amount).

For the following questions, suppose that the industry standard for straight bet "even" odds is -115. Answer Problems 13 through 15 for 2-legs, 3-legs, and 4-legs of fixed books.

Problem 13. Calculate the overround on the different types of parlays using -115 odds.

Problem 14. Compare your expected winnings when playing the 3 types of parlays on a fixed book versus placing them using standard straight bet odds. Which is more profitable and how much are you expected to lose?

Problem 15. Suppose that each event in the fair book is independent. At what percent do users need to pick the correct outcomes (also known as the user's **hit rate**) for all 3 types of parlays to ensure a long-term profit?

5 March Madness

During every March, college basketball teams participate in an annual single-elimination tournament beginning with 64 teams, for a total of 6 rounds to determine the winner. There are many upsets in this tournament, which are typically determined by the **seedings** or placements of these teams. There are 4 separate conferences (for the sake of simplicity, we will call them north, south, east, and west), and 4 total sets of 1-seeds to 16-seeds. In the first round, the 1-seed plays the 16-seed, the 2-seed plays the 15-seed, and so on.

We will define an **upset** as a game played between a single-digit seed and a double-digit seed.

Problem 16. How many potential upsets are there per tournament?

Problem 17. In terms of odds, we define an upset as the favorite having at least -150 odds and an underdog having at most +150 odds. Suppose that every non-upset game has -100 odds for either team and all potential upsets have a -150/+150 split. What is the probability of having no upsets in the first round[1] ?

Problem 18. Calculate the probability of having less than 3 upsets in the first round.[2]

[1] Recently, the NCAA has also added a few games that precede the first round to determine the 10-seed and the 16-seed. How does this affect our answer?

[2] This problem shows why following the 'People's Bracket' is generally a bad idea if you want to gain an edge over your group.

Problem 19. Suppose we wanted to predict the entire first round of March Madness. On average, there are 5 upsets in the first round. Using the odds from Problem 17, find an upper and lower bound of the American odds to choose all 32 winners correctly.

6 Hedging

A strategy that involves placing bets involving the opposite choice you have already selected is known as **hedging**.

Problem 20. Suppose you have a fair book 3-leg parlay that already has two correct events, while the third event hasn't started yet. Explain why hedging the last event as a straight bet can be a safe move.

Problem 21. Under what conditions should a user *not* hedge their bets?

This problem illustrates the reasoning behind why the fixed books do not allow players to submit straight bet entries.

7 Hard Combinatorics Problems

Problem 22. Let S be the set of sequences of length 2018 whose terms are in the set $\{1, 2, 3, 4, 5, 6, 10\}$ and sum to 3860. Prove that the cardinality of S is at most

$$2^{3860} \cdot (\frac{2018}{2048})^{2018}$$

Problem 23. Recall that a regular icosahedron is a convex polyhedron having 12 vertices and 20 faces; the faces are congruent equilateral triangles. On each face of a regular icosahedron is written a nonnegative integer such that the sum of all 20 integers is 39. Show that two faces share a vertex and have the same integer written on them.

Problem 24. A round-robin tournament among $2n$ teams lasted for $2n - 1$ days, as follows. On each day, every team played one game against another team, with one team winning and one team losing in each of the n games. Throughout the tournament, each team played every other team exactly once. Can one necessarily choose one winning team from each day without choosing any team more than once?