Optimal Kelly Wagering for Multiple Mutually Exclusive Outcomes

Christopher D. Long
Headlamp Software

April 17, 2025

Abstract

The Kelly criterion provides a framework for optimal capital allocation in scenarios with favorable odds, aiming to maximize the long-term logarithmic growth rate of wealth. While the single-outcome case is well-known, the extension to multiple, mutually exclusive outcomes (such as horse races or predicting winners in a tournament) presents greater complexity. This paper provides a clear and rigorous derivation of the optimal Kelly wagering strategy for $n \geq 2$ outcomes. We derive explicit formulas for the optimal wager fractions and the resulting expected log-growth, interpreting the latter in terms of Kullback-Leibler divergence. We analyze the behavior of optimal wagers as the total invested fraction varies, leading to efficient algorithms for determining the optimal betting portfolio both for the unconstrained (full Kelly) case and for scenarios with a fixed total wager fraction (including fractional Kelly betting). Our exposition aims to be accessible. We highlight properties such as the conditions under which outcomes are included in the optimal portfolio, including the possibility of wagering on outcomes with individually negative expectation if they serve as conditional hedges.

1 Introduction

In his seminal 1956 paper [1], J. L. Kelly Jr. proposed a strategy derived from information theory principles that maximizes the expected logarithm of wealth over repeated plays. This strategy, now known as the Kelly criterion, advises wagering a fraction of one's capital equal to the expected net winnings divided by the net odds if successful. Its key property is asymptotic optimality: under certain conditions, it achieves a higher long-term growth rate of capital than any other essentially different strategy.

The original formulation often focuses on binary outcomes (win/loss) or scenarios where only one favorable bet is available. However, many real-world situations involve multiple, mutually exclusive potential outcomes, such as betting on different horses in a race, different teams in a match, or different candidates in an election. Each outcome i has associated odds d_i offered by a bookmaker (implying a probability $\hat{p}_i = 1/d_i$) and a subjective probability p_i estimated by the bettor. The bettor seeks to allocate fractions f_i of their bankroll across these outcomes to maximize expected logarithmic growth.

This multi-outcome problem has been addressed previously, notably by Smoczynski and Tomkins [2]. They presented an explicit solution for the optimal allocation. However, this exposition can be challenging for readers not familiar with the mathematical language and underlying optimization techniques. This paper revisits the n-outcome Kelly problem with the goal of providing a more accessible, step-by-step derivation based on standard constrained optimization theory (specifically, Karush-Kuhn-Tucker conditions).

We will:

- Formally define the problem and notation.
- Derive the optimal betting fractions f_i for the set of "active" outcomes (those receiving positive wagers).
- Express the optimal total fraction wagered, $F = \sum f_i$.
- Calculate the maximum expected log-growth and relate it to the Kullback-Leibler divergence between the bettor's probability distribution and the market-implied distribution.

- Analyze how the set of active bets and their corresponding fractions change as the total wagered fraction F is constrained.
- Characterize the conditions under which an outcome is included in the optimal betting set, including the concept of "conditional positive expectation".
- Present computationally efficient algorithms for finding the optimal betting portfolio for both the unconstrained Kelly optimum and for a fixed total wager fraction F.
- Briefly discuss fractional Kelly betting within this framework.

Our approach emphasizes clarity and builds the results logically, aiming to make Kelly optimization for multiple outcomes understandable and applicable.

2 Notation and Problem Setup

We consider a betting scenario with $n \geq 2$ mutually exclusive outcomes, denoted by the set $O = \{1, 2, \ldots, n\}$. Exactly one of these outcomes will occur.

Definition 2.1 (Notation). *We define the following quantities:*

- p_i: *The bettor's subjective probability estimate for outcome i occurring. We assume $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$.*
- d_i: *The decimal odds offered for outcome i. If the bettor wagers an amount x on outcome i and wins, they receive back xd_i (for a net profit of $x(d_i - 1)$). We assume $d_i \geq 1$.*
- $\hat{p}_i$: *The probability implied by the odds, defined as $\hat{p}_i = 1/d_i$. Note that typically $\sum_{i=1}^{n} \hat{p}_i > 1$ due to the bookmaker's margin (vigorish).*
- f_i: *The fraction of the bettor's bankroll wagered on outcome i. We require $f_i \geq 0$.*
- *Bankroll: Assumed to be normalized to 1 without loss of generality.*
- A: *The set of* active *outcomes, $A = \{i \in O \mid f_i > 0\}$.*
- F: *The total fraction of the bankroll wagered, $F = \sum_{i=1}^{n} f_i = \sum_{i \in A} f_i$. We must have $F \leq 1$.*

For any subset of outcomes $S \subseteq O$:

- $P_S = \sum_{i \in S} p_i$ *(the bettor's total probability for outcomes in S).*
- $Q_S = 1 - P_S = \sum_{i \notin S} p_i$ *(the bettor's total probability for outcomes not in S).*
- $\hat{P}_S = \sum_{i \in S} \hat{p}_i$ *(the sum of implied probabilities for outcomes in S).*
- $\hat{Q}_S = 1 - \hat{P}_S$ *(Note: This is based on a total implied probability of 1, which might not hold if $\sum \hat{p}_i \neq 1$. A more careful definition might be needed if we strictly interpret $\hat{P}_S + \hat{Q}_S = 1$. However, the term $1 - \hat{P}_S$ appears naturally in the derivations, so we retain this notation.) We specifically use $Q_A = 1 - P_A$ and $\hat{Q}_A = 1 - \hat{P}_A$ frequently, relating to the active set A.*

Definition 2.2 (Objective Function). *If outcome i occurs (with probability p_i), the bettor's bankroll changes from 1 to W_i. The amount not wagered is $1 - F$. The wager f_i on the winning outcome returns $f_i d_i$. All other wagers f_j ($j \neq i$) are lost. Thus, the final wealth if outcome i occurs is:*

$$W_i = (1 - F) + f_i d_i = 1 - \sum_{j=1}^{n} f_j + f_i d_i$$

The Kelly criterion seeks to maximize the expected logarithm of the final wealth:

$$G(f_1, \ldots, f_n) = E[\log(W)] = \sum_{i=1}^{n} p_i \log(W_i) = \sum_{i=1}^{n} p_i \log\left(1 - \sum_{j=1}^{n} f_j + f_i d_i\right)$$

The optimization problem is:

$$\max_{f_1,\dots,f_n} G(f_1,\dots,f_n)$$

subject to the constraints:

$$f_i \geq 0 \quad \forall i \in O$$
$$\sum_{i=1}^{n} f_i \leq 1$$

3 Main Results

We present the main results concerning the optimal wager fractions f_i, the total fraction F, the resulting expected growth, and the structure of the active set A. Proofs will follow in the subsequent section.

Let A denote the set of active outcomes at the global optimum.

Theorem 3.1 (Optimal Wager Fractions). *At the global optimum, the fraction wagered on an active outcome $i \in A$ is given by:*

$$f_i = p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A} \tag{1}$$

where $Q_A = 1 - P_A = 1 - \sum_{j\in A} p_j$ and $\hat{Q}_A = 1 - \hat{P}_A = 1 - \sum_{j\in A} \hat{p}_j$. For outcomes $k \notin A$, $f_k = 0$.

Proposition 3.2 (Optimal Total Wager Fraction). *At the global optimum, the total fraction wagered $F = \sum_{i\in A} f_i$ is:*

$$F = P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A} \tag{2}$$

Equivalently, using $P_A + Q_A = 1$ and $\hat{P}_A + \hat{Q}_A = 1$:

$$F = 1 - \frac{Q_A}{\hat{Q}_A} \tag{3}$$

Proposition 3.3 (Optimal Expected Log-Growth). *At the global optimum, the maximum expected log-growth G^* is:*

$$G^* = \sum_{i\in A} p_i \log\left(\frac{p_i}{\hat{p}_i}\right) + Q_A \log\left(\frac{Q_A}{\hat{Q}_A}\right) \tag{4}$$

Remark 3.4 (KL Divergence Interpretation). *The first term in G^*, $\sum_{i\in A} p_i \log(p_i/\hat{p}_i)$, is the Kullback-Leibler (KL) divergence from the implied probability distribution $\hat{p}$ to the bettor's distribution p, restricted to the active set A. It measures the "information gain" or "edge" the bettor perceives over the market odds for the outcomes they choose to bet on. The second term, $Q_A \log(Q_A/\hat{Q}_A)$, reflects the relative probability mass assigned to the non-active outcomes by the bettor versus the market.*

Theorem 3.5 (Characterization of the Active Set A). *The set of active outcomes A at the global optimum consists precisely of those outcomes $i \in O$ satisfying:*

$$p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A} > 0 \tag{5}$$

Equivalently, $i \in A$ if and only if $p_i/\hat{p}_i > Q_A/\hat{Q}_A$. This implies:

(a) If no outcome i has positive expected value (i.e., $p_i d_i \leq 1$ or $p_i \leq \hat{p}_i$ for all i), then the optimal strategy is to wager nothing ($F = 0$, $A = \emptyset$). (Claim 7)

(b) Any outcome i with strictly positive expected value ($p_i d_i > 1$) must belong to the active set A at the global optimum. (Claim 6)

(c) Outcomes k with negative expected value ($p_k d_k < 1$) may still be included in the active set A if their relative probability $p_k/\hat{p}_k$ is sufficiently high compared to the ratio $Q_A/\hat{Q}_A$. Such bets can act as conditional hedges. (Claim 9)

Theorem 3.6 (Behavior under Fixed Total Wager F)**.** *Consider optimizing the wager fractions f_i subject to $\sum f_i = F$ for a fixed $F \in [0,1]$, and $f_i \geq 0$. Let the outcomes be sorted such that $p_1 d_1 \geq p_2 d_2 \geq \cdots \geq p_n d_n$.*

(a) If $F = 1$, all outcomes i with $p_i > 0$ are active ($f_i = p_i$). (Claim 8)

(b) As F decreases from 1 to 0, outcomes become inactive (transition from $f_i > 0$ to $f_i = 0$). The order in which outcomes become inactive is precisely the order of increasing $p_i d_i$ (i.e., outcome n becomes inactive first, then $n-1$, etc.). (Claim 4)

(c) Within any interval of F values where the active set A remains constant (an "interior region"), the optimal fractions f_i for $i \in A$ scale linearly with F. Specifically:

$$f_i(F) = \frac{p_i F}{P_A} + (1-F)\left(\frac{p_i \hat{P}_A}{P_A} - \hat{p}_i\right) \quad \text{for } i \in A \tag{6}$$

(Claims 5, 13, 14)

(d) Let $F_{crit,m}$ be the value of F at which outcome m transitions from active to inactive (assuming outcomes with higher $p_k d_k$ are still active). This critical value is given by:

$$F_{crit,m} = \frac{P_A - p_m d_m \hat{P}_A}{P_A + p_m d_m \hat{Q}_A} \tag{7}$$

where A is the active set just before m becomes inactive. (Claim 11)

(e) At such a critical point $F = F_{crit,m}$, the fractions for the remaining active outcomes $i \in A \setminus \{m\}$ are given by:

$$f_i = (1 - F_{crit,m})\left[\frac{p_i}{p_m d_m} - \hat{p}_i\right] \tag{8}$$

(Claim 12)

4 Proofs

We use the Karush-Kuhn-Tucker (KKT) conditions for constrained optimization to find the optimal f_i. The objective is $G = \sum_{k=1}^n p_k \log(W_k)$ where $W_k = 1 - F + f_k d_k$. The constraints are $f_i \geq 0$ and $F = \sum f_i \leq 1$.

We form the Lagrangian:

$$\mathcal{L}(f_1, \ldots, f_n, \mu_1, \ldots, \mu_n, \nu) = G(f) - \sum_{i=1}^n \mu_i(-f_i) - \nu\left(\sum_{i=1}^n f_i - 1\right)$$

where $\mu_i \geq 0$ are the multipliers for $f_i \geq 0$, and $\nu \geq 0$ is the multiplier for $F \leq 1$.

The KKT conditions include:

1. Stationarity: $\frac{\partial \mathcal{L}}{\partial f_i} = 0$ for all $i = 1, \ldots, n$.
2. Primal Feasibility: $f_i \geq 0$, $\sum f_i \leq 1$.
3. Dual Feasibility: $\mu_i \geq 0$, $\nu \geq 0$.
4. Complementary Slackness: $\mu_i f_i = 0$, $\nu(\sum f_i - 1) = 0$.

First, let's calculate the partial derivative of G:

$$\frac{\partial G}{\partial f_i} = \sum_{k=1}^n p_k \frac{1}{W_k} \frac{\partial W_k}{\partial f_i} = \sum_{k=1}^n p_k \frac{1}{W_k}(-1 + \delta_{ik} d_i)$$

$$\frac{\partial G}{\partial f_i} = \frac{p_i d_i}{W_i} - \sum_{k=1}^n \frac{p_k}{W_k}$$

where δ_{ik} is the Kronecker delta. Let $C = \sum_{k=1}^n \frac{p_k}{W_k} = E[1/W]$. The stationarity condition becomes:

$$\frac{p_i d_i}{W_i} - C + \mu_i - \nu = 0$$

4.1 Proof of Theorem 3.1 (Optimal f_i)

Assume an optimal solution exists where the total wager $F < 1$. By complementary slackness, $\nu = 0$. The stationarity condition simplifies to $\frac{p_i d_i}{W_i} - C + \mu_i = 0$.

Consider an outcome $i \in A$. By definition, $f_i > 0$. By complementary slackness, $\mu_i = 0$. Thus, for $i \in A$:

$$\frac{p_i d_i}{W_i} - C = 0 \implies W_i = \frac{p_i d_i}{C}$$

Recall $W_i = 1 - F + f_i d_i$. So, $1 - F + f_i d_i = \frac{p_i d_i}{C}$. Solving for f_i:

$$f_i d_i = \frac{p_i d_i}{C} - (1 - F) \implies f_i = \frac{p_i}{C} - \frac{1 - F}{d_i} = \frac{p_i}{C} - \hat{p}_i(1 - F)$$

Now consider an outcome $k \notin A$. By definition, $f_k = 0$. Thus $W_k = 1 - F$. The stationarity condition is $\frac{p_k d_k}{1-F} - C + \mu_k = 0$. Since $\mu_k \geq 0$, this implies:

$$\frac{p_k d_k}{1 - F} \leq C \quad \text{for } k \notin A$$

Let's determine the constant C.

$$C = \sum_{j=1}^{n} \frac{p_j}{W_j} = \sum_{j \in A} \frac{p_j}{W_j} + \sum_{j \notin A} \frac{p_j}{W_j}$$

Substitute W_j for $j \in A$ and $j \notin A$:

$$C = \sum_{j \in A} \frac{p_j}{p_j d_j / C} + \sum_{j \notin A} \frac{p_j}{1 - F} = C \sum_{j \in A} \frac{1}{d_j} + \frac{1}{1 - F} \sum_{j \notin A} p_j$$

$$C = C\hat{P}_A + \frac{Q_A}{1 - F}$$

$$C(1 - \hat{P}_A) = \frac{Q_A}{1 - F} \implies C\hat{Q}_A = \frac{Q_A}{1 - F}$$

If $Q_A = 0$ (i.e., $P_A = 1$), then either $C = 0$ or $\hat{Q}_A = 0$. If $P_A = 1$, G involves $\log(f_i d_i)$ which requires $f_i > 0$. If $Q_A > 0$ and $\hat{Q}_A > 0$, we have:

$$1 - F = \frac{Q_A}{C\hat{Q}_A}$$

Now substitute this back into the expression for f_i ($i \in A$):

$$f_i = \frac{p_i}{C} - \hat{p}_i(1 - F) = \frac{p_i}{C} - \hat{p}_i \frac{Q_A}{C\hat{Q}_A} = \frac{1}{C}\left(p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A}\right)$$

Sum these fractions over $i \in A$:

$$F = \sum_{i \in A} f_i = \frac{1}{C} \sum_{i \in A} \left(p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A}\right) = \frac{1}{C}\left(P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A}\right)$$

We have two expressions relating F and C: 1. $F = 1 - \frac{Q_A}{C\hat{Q}_A}$ 2. $F = \frac{1}{C}\left(P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A}\right)$

Equating them:

$$1 - \frac{Q_A}{C\hat{Q}_A} = \frac{1}{C}\left(P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A}\right)$$

Multiply by C:

$$C - \frac{Q_A}{\hat{Q}_A} = P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A}$$

$$C = P_A + \frac{Q_A}{\hat{Q}_A}(1 - \hat{P}_A) = P_A + \frac{Q_A}{\hat{Q}_A}\hat{Q}_A = P_A + Q_A = 1$$

So, at the optimum (assuming $F < 1$), we must have $C = 1$.

Substituting $C = 1$ back into the expressions: For $i \in A$: $f_i = \frac{1}{1}\left(p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A}\right) = p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A}$. This proves (1). For $k \notin A$: $\frac{p_k d_k}{1-F} \leq C = 1$. From $1 - F = Q_A/(C\hat{Q}_A)$, we get $1 - F = Q_A/\hat{Q}_A$.

What if the optimum occurs at $F = 1$? Then $\nu \geq 0$. The stationarity condition is $\frac{p_i d_i}{W_i} - C + \mu_i = \nu$. If $F = 1$, then $W_i = f_i d_i$. If $f_i > 0$, $\mu_i = 0$. $\frac{p_i d_i}{f_i d_i} - C = \nu \implies \frac{p_i}{f_i} = C + \nu$. So $f_i = p_i/(C+\nu)$. Summing over all i (assuming $f_i > 0$ for all i with $p_i > 0$): $F = \sum f_i = 1 = \sum p_i/(C+\nu) = P_{all}/(C+\nu) = 1/(C+\nu)$. This implies $C + \nu = 1$. So $f_i = p_i/1 = p_i$. In this case ($F = 1$), $A = \{i \mid p_i > 0\}$, $P_A = 1$, $Q_A = 0$. The formula (1) gives $f_i = p_i - \hat{p}_i(0/\hat{Q}_A) = p_i$, provided $\hat{Q}_A \neq 0$. If $\hat{Q}_A = 0$ (i.e. $\hat{P}_A = 1$), the market odds sum to 1 on the active set. The formula needs careful interpretation, but the result $f_i = p_i$ holds when $F = 1$. Thus, Theorem 3.1 holds generally.

4.2 Proof of Proposition 3.2 (Optimal F)

Summing the optimal f_i from (1) over the active set A:

$$F = \sum_{i \in A} f_i = \sum_{i \in A} \left(p_i - \hat{p}_i \frac{Q_A}{\hat{Q}_A}\right) = \left(\sum_{i \in A} p_i\right) - \left(\sum_{i \in A} \hat{p}_i\right) \frac{Q_A}{\hat{Q}_A}$$

$$F = P_A - \hat{P}_A \frac{Q_A}{\hat{Q}_A}$$

This proves (2). Using $P_A = 1 - Q_A$ and $\hat{P}_A = 1 - \hat{Q}_A$:

$$F = (1 - Q_A) - (1 - \hat{Q}_A)\frac{Q_A}{\hat{Q}_A} = 1 - Q_A - \frac{Q_A}{\hat{Q}_A} + \hat{Q}_A \frac{Q_A}{\hat{Q}_A} = 1 - Q_A - \frac{Q_A}{\hat{Q}_A} + Q_A = 1 - \frac{Q_A}{\hat{Q}_A}$$

This proves (3).

4.3 Proof of Proposition 3.3 (Optimal Growth)

The expected log-growth is $G = \sum_{i=1}^{n} p_i \log(W_i) = \sum_{i \in A} p_i \log(W_i) + \sum_{i \notin A} p_i \log(W_i)$. At the optimum, we found $C = 1$. For $i \in A$, $W_i = p_i d_i / C = p_i d_i = p_i/\hat{p}_i$. For $k \notin A$, $f_k = 0$, so $W_k = 1 - F$. We also found $1 - F = Q_A/(C\hat{Q}_A) = Q_A/\hat{Q}_A$. Substituting these into the expression for G:

$$G^* = \sum_{i \in A} p_i \log\left(\frac{p_i}{\hat{p}_i}\right) + \sum_{k \notin A} p_k \log\left(\frac{Q_A}{\hat{Q}_A}\right)$$

$$G^* = \sum_{i \in A} p_i \log\left(\frac{p_i}{\hat{p}_i}\right) + \left(\sum_{k \notin A} p_k\right) \log\left(\frac{Q_A}{\hat{Q}_A}\right)$$

$$G^* = \sum_{i \in A} p_i \log\left(\frac{p_i}{\hat{p}_i}\right) + Q_A \log\left(\frac{Q_A}{\hat{Q}_A}\right)$$

This proves (4).

4.4 Proof of Theorem 3.5 (Active Set)

From the KKT analysis, we established: 1. For $i \in A$, $f_i = p_i - \hat{p}_i Q_A/\hat{Q}_A$. Since $f_i > 0$, this requires $p_i - \hat{p}_i Q_A/\hat{Q}_A > 0$. 2. For $k \notin A$, we required $\frac{p_k d_k}{1-F} \leq C = 1$. Substituting $1 - F = Q_A/\hat{Q}_A$, this becomes $\frac{p_k d_k}{Q_A/\hat{Q}_A} \leq 1$, or $p_k d_k \hat{Q}_A \leq Q_A$. Dividing by $d_k \hat{Q}_A$ (assuming $d_k \geq 1, \hat{Q}_A > 0$), we get $p_k \leq (1/d_k) Q_A/\hat{Q}_A = \hat{p}_k Q_A/\hat{Q}_A$. This is equivalent to $p_k - \hat{p}_k Q_A/\hat{Q}_A \leq 0$.

Combining these, an outcome i is in the active set A if and only if $p_i - \hat{p}_i Q_A/\hat{Q}_A > 0$, proving (5). The equivalent form $p_i/\hat{p}_i > Q_A/\hat{Q}_A$ follows directly.

Now let's prove the specific claims: (a) If $p_i d_i \leq 1$ (i.e., $p_i \leq \hat{p}_i$) for all i. Suppose A is non-empty. Then $P_A > 0$ and $\hat{P}_A > 0$. If $p_i/\hat{p}_i > Q_A/\hat{Q}_A$, note that since $p_i/\hat{p}_i \leq 1$, this would require $Q_A/\hat{Q}_A < 1$, which means $Q_A < \hat{Q}_A \implies 1 - P_A < 1 - \hat{P}_A \implies P_A > \hat{P}_A$. However, if $p_j \leq \hat{p}_j$ for all j, then $P_A = \sum_{j \in A} p_j \leq \sum_{j \in A} \hat{p}_j = \hat{P}_A$. The only way to have $P_A > \hat{P}_A$ is if this inequality doesn't hold, which contradicts the premise. Therefore, the condition $p_i/\hat{p}_i > Q_A/\hat{Q}_A$ cannot be met if $p_j \leq \hat{p}_j$ for all $j \in A$. The only possibility is $A = \emptyset$. If $A = \emptyset$, then $P_A = 0, Q_A = 1, \hat{P}_A = 0, \hat{Q}_A = 1$. The condition becomes $p_i - \hat{p}_i(1/1) > 0 \implies p_i > \hat{p}_i$, which contradicts $p_i \leq \hat{p}_i$. So no outcome can enter A, A remains empty, and $F = 0$.

(b) If $p_i d_i > 1$ (i.e., $p_i > \hat{p}_i$) for some i. Let A be the optimal active set. Assume $i \notin A$. Then $p_i - \hat{p}_i Q_A/\hat{Q}_A \leq 0$. This means $p_i/\hat{p}_i \leq Q_A/\hat{Q}_A$. Since $p_i/\hat{p}_i > 1$, this requires $Q_A/\hat{Q}_A > 1 \implies Q_A > \hat{Q}_A \implies 1 - P_A > 1 - \hat{P}_A \implies P_A < \hat{P}_A$. This is possible. However, consider the algorithm for constructing A (see Algorithm 1 below). If we start with $A = \emptyset$, then $Q_A = 1, \hat{Q}_A = 1$. The condition for adding i is $p_i - \hat{p}_i(1/1) > 0$, which is true since $p_i > \hat{p}_i$. So i must be included in A at some point. Can it be removed later? An outcome k is removed if $p_k - \hat{p}_k Q_A/\hat{Q}_A$ becomes non-positive. If this happen to our i where $p_i/\hat{p}_i > 1$, it requires $Q_A/\hat{Q}_A \geq p_i/\hat{p}_i > 1$, which implies $P_A < \hat{P}_A$. The set A generated by the algorithm (which we will show finds the optimum) includes all i satisfying the condition relative to the final A. If $p_i d_i > 1$, its f_i value might be small if $Q_A/\hat{Q}_A$ is large, but it must be positive. So i must be in A.

(c) If $p_k d_k < 1$ (i.e., $p_k < \hat{p}_k$). Outcome k is active if $p_k - \hat{p}_k Q_A/\hat{Q}_A > 0$, which means $p_k/\hat{p}_k > Q_A/\hat{Q}_A$. Since $p_k/\hat{p}_k < 1$, this requires $Q_A/\hat{Q}_A < 1 \implies P_A > \hat{P}_A$. This can happen if other outcomes $j \in A$ have $p_j d_j \gg 1$. Thus, negative expectation outcomes can be active.

4.5 Proof of Theorem 3.6 (Fixed F)

We now optimize G subject to $\sum f_i = F$ (equality) and $f_i \geq 0$. The Lagrangian is $\mathcal{L} = G(f) + \lambda(F - \sum f_i) - \sum \mu_i(-f_i)$. KKT stationarity: $\frac{\partial \mathcal{L}}{\partial f_i} = \frac{p_i d_i}{W_i} - C - \lambda + \mu_i = 0$, where $W_i = 1 - F + f_i d_i$ and $C = \sum p_k/W_k$. Note C is not necessarily 1 here. For $i \in A$, $f_i > 0 \implies \mu_i = 0$. So $\frac{p_i d_i}{W_i} = C + \lambda$. Let $\Lambda = C + \lambda$. $W_i = p_i d_i/\Lambda$. $1 - F + f_i d_i = p_i d_i/\Lambda$. $f_i = \frac{p_i}{\Lambda} - \frac{1-F}{d_i} = \frac{p_i}{\Lambda} - \hat{p}_i(1-F)$. Sum over $i \in A$: $F = \sum_{i \in A} f_i = \frac{P_A}{\Lambda} - \hat{P}_A(1-F)$. Solve for Λ: $\Lambda = \frac{P_A}{F + \hat{P}_A(1-F)}$. Substitute Λ back into f_i:

$$f_i(F) = p_i \frac{F + \hat{P}_A(1-F)}{P_A} - \hat{p}_i(1-F)$$

$$f_i(F) = \frac{p_i F}{P_A} + \frac{p_i \hat{P}_A(1-F)}{P_A} - \frac{\hat{p}_i P_A(1-F)}{P_A}$$

$$f_i(F) = \frac{p_i F}{P_A} + (1-F)\left(\frac{p_i \hat{P}_A - \hat{p}_i P_A}{P_A}\right)$$

This proves the linear dependence on F for a fixed active set A, proving (6) and thus Claims 5 and 13.

(a) If $F = 1$, $W_i = f_i d_i$. $G = \sum p_i \log(f_i d_i)$. Maximize subject to $\sum f_i = 1, f_i \geq 0$. Lagrangian $\mathcal{L} = \sum p_i \log(f_i d_i) + \lambda(1 - \sum f_i) - \sum \mu_i(-f_i)$. $\partial\mathcal{L}/\partial f_i = p_i/f_i - \lambda + \mu_i = 0$. If $f_i > 0$, $\mu_i = 0$, so $p_i/f_i = \lambda \implies f_i = p_i/\lambda$. Summing: $\sum f_i = 1 = \sum p_i/\lambda = (\sum p_i)/\lambda = 1/\lambda$. So $\lambda = 1$. Thus $f_i = p_i$. This holds for all i with $p_i > 0$. So if $F = 1$, $A = \{i \mid p_i > 0\}$ and $f_i = p_i$. This proves Claim 8.

(d, e) Critical points: An outcome m becomes inactive when $f_m(F) = 0$. This defines $F_{crit,m}$. Set (6) to 0 for $i = m$:

$$\frac{p_m F_{crit,m}}{P_A} + (1 - F_{crit,m})\left(\frac{p_m \hat{P}_A - \hat{p}_m P_A}{P_A}\right) = 0$$

$$p_m F_{crit,m} + (1 - F_{crit,m})(p_m \hat{P}_A - \hat{p}_m P_A) = 0$$

$$F_{crit,m}(p_m - p_m \hat{P}_A + \hat{p}_m P_A) = \hat{p}_m P_A - p_m \hat{P}_A$$

$$F_{crit,m}(p_m(1 - \hat{P}_A) + \hat{p}_m P_A) = \hat{p}_m P_A - p_m \hat{P}_A$$

$$F_{crit,m}(p_m \hat{Q}_A + \hat{p}_m P_A) = \hat{p}_m P_A - p_m \hat{P}_A$$

$$F_{crit,m} = \frac{\hat{p}_m P_A - p_m \hat{P}_A}{p_m \hat{Q}_A + \hat{p}_m P_A} = \frac{P_A/\hat{p}_m^{-1} - p_m \hat{P}_A}{p_m \hat{Q}_A + P_A/\hat{p}_m^{-1}}$$

Multiply numerator and denominator by $d_m = 1/\hat{p}_m$:

$$F_{crit,m} = \frac{d_m P_A - p_m d_m \hat{P}_A}{p_m d_m \hat{Q}_A + d_m P_A \hat{p}_m} = \frac{P_A - p_m d_m \hat{P}_A}{p_m d_m \hat{Q}_A + P_A}$$

This proves (7) and Claim 11. To prove (8), we use the alternative form $f_i = \frac{p_i}{\Lambda} - \hat{p}_i(1-F)$. At $F = F_{crit,m}$, we have $f_m = 0$, so $0 = \frac{p_m}{\Lambda_{crit}} - \hat{p}_m(1 - F_{crit,m})$. This implies $\Lambda_{crit} = \frac{p_m}{\hat{p}_m(1-F_{crit,m})} = \frac{p_m d_m}{1-F_{crit,m}}$. Substitute this Λ_{crit} into the expression for f_i ($i \in A \setminus \{m\}$):

$$f_i = \frac{p_i}{\Lambda_{crit}} - \hat{p}_i(1 - F_{crit,m}) = p_i \frac{1 - F_{crit,m}}{p_m d_m} - \hat{p}_i(1 - F_{crit,m})$$

$$f_i = (1 - F_{crit,m}) \left[\frac{p_i}{p_m d_m} - \hat{p}_i \right]$$

This proves (8) and Claim 12.

(b) Ordering: We need to show that $F_{crit,m}$ is highest for the outcome m with the lowest $p_m d_m$. Let $x = p_m d_m$.

$$F_{crit}(x) = \frac{P_A - x\hat{P}_A}{P_A + x\hat{Q}_A}$$

We compute the derivative with respect to x (treating $P_A, \hat{P}_A, \hat{Q}_A$ as constants, which holds for a given set A):

$$\frac{dF_{crit}}{dx} = \frac{-\hat{P}_A(P_A + x\hat{Q}_A) - (P_A - x\hat{P}_A)\hat{Q}_A}{(P_A + x\hat{Q}_A)^2}$$

$$\text{Numerator} = -\hat{P}_A P_A - x\hat{P}_A\hat{Q}_A - P_A\hat{Q}_A + x\hat{P}_A\hat{Q}_A = -P_A(\hat{P}_A + \hat{Q}_A) = -P_A$$

Assuming $P_A > 0$, the derivative $dF_{crit}/dx = -P_A/(P_A + x\hat{Q}_A)^2 < 0$. Therefore, $F_{crit,m}$ decreases as $p_m d_m$ increases. This means the outcome with the smallest $p_m d_m$ has the largest $F_{crit,m}$ and becomes inactive first as F decreases from 1. This proves Claim 4.

Claim 14 follows from the linearity established in (6) and the boundary conditions derived.

5 Algorithms

The characterization of the active set A leads directly to algorithms for finding the optimal wagers.

5.1 Algorithm for Global Optimum (Full Kelly)

Theorem 3.5 states $i \in A \iff p_i - \hat{p}_i Q_A/\hat{Q}_A > 0$. This condition depends on the set A itself. This suggests an iterative approach. Sorting outcomes by some measure of attractiveness seems beneficial. A natural candidate is $p_i d_i = p_i/\hat{p}_i$.

Algorithm 1 Find Optimal Active Set A and Full Kelly Fractions f_i

Sort outcomes such that $p_1 d_1 \geq p_2 d_2 \geq \cdots \geq p_n d_n$.
Initialize $A = \emptyset$, $P_A = 0$, $\hat{P}_A = 0$, $Q_A = 1$, $\hat{Q}_A = 1$.
Set 'converged = false'.
while not 'converged' **do**
    Set 'changed = false'.
    Let $A_{prev} = A$.
    Define potential active set $A_{cand} = \{i \in O \mid p_i\hat{Q}_A > \hat{p}_i Q_A\}$.
    if $A_{cand} == A_{prev}$ **then**
        Set 'converged = true'.
    else
        Set $A = A_{cand}$.
        Recalculate $P_A = \sum_{i \in A} p_i$, $\hat{P}_A = \sum_{i \in A} \hat{p}_i$.
        Recalculate $Q_A = 1 - P_A$, $\hat{Q}_A = 1 - \hat{P}_A$. ▷ Handle $\hat{Q}_A = 0$ case if necessary
        Set 'changed = true'.
▷ The optimal active set A is found.
for all $i \in O$ **do**
    if $i \in A$ **then**
        Calculate $f_i = p_i - \hat{p}_i Q_A/\hat{Q}_A$. ▷ Ensure $\hat{Q}_A \neq 0$. If $\hat{Q}_A = 0$, $F = 1, f_i = p_i$.
    else
        Set $f_i = 0$.
return $f_1, \ldots, f_n$

Remark 5.1. *The iterative process in Algorithm 1 converges because adding outcomes to A tends to decrease Q_A and increase $\hat{P}_A$ (decreasing $\hat{Q}_A$), making the ratio $Q_A/\hat{Q}_A$ potentially smaller, which could allow more outcomes to satisfy the condition $p_i/\hat{p}_i > Q_A/\hat{Q}_A$. Conversely, removing outcomes increases $Q_A/\hat{Q}_A$, potentially excluding more. The process stabilises at the unique optimal set A. The initial sorting is not strictly necessary for convergence but aligns with intuition. A greedy approach (Claim 10) - adding outcomes one by one in order of $p_i d_i$ as long as the condition holds - also works and finds the same optimal set A.*

5.2 Algorithm for Fixed Total Wager F

Theorem 3.6 provides the condition $f_i > 0$ for a fixed F and active set A, based on equation (6).

$$f_i(F) = \frac{p_i F}{P_A} + (1 - F)\left(\frac{p_i \hat{P}_A - \hat{p}_i P_A}{P_A}\right) > 0$$

Multiplying by P_A (assume $P_A > 0$):

$$p_i F + (1 - F)(p_i \hat{P}_A - \hat{p}_i P_A) > 0$$

This condition again depends on A. We use an iterative procedure similar to Algorithm 1, but using this F-dependent condition.

Algorithm 2 Find Optimal Fractions f_i for Fixed Total Wager F

Require: Total wager fraction $F \in [0, 1]$.
Sort outcomes such that $p_1 d_1 \geq p_2 d_2 \geq \cdots \geq p_n d_n$.
Initialize $A = \emptyset$, $P_A = 0$, $\hat{P}_A = 0$.
Set 'converged = false'.
while not 'converged' **do**
    Set 'changed = false'.
    Let $A_{prev} = A$.
    Define potential active set $A_{cand} = \{i \in O \mid p_i F + (1 - F)(p_i \hat{P}_A - \hat{p}_i P_A) > 0\}$. ▷ Handle $P_A = 0$ case
    if $P_A == 0$ **then** ▷ Check condition for $i = 1$
        if $p_1 d_1 > 1$ **then** $A_{cand} = \{1\}$
        else$A_{cand} = \emptyset$
    if $A_{cand} == A_{prev}$ **then**
        Set 'converged = true'.
    else
        Set $A = A_{cand}$.
        Recalculate $P_A = \sum_{i \in A} p_i$, $\hat{P}_A = \sum_{i \in A} \hat{p}_i$.
        Set 'changed = true'.
▷ The optimal active set A for the given F is found.
▷ Handle case $P_A = 0$ (results in $f_i = 0$ for all i)
if $P_A > 0$ **then**
    for all $i \in O$ **do**
        if $i \in A$ **then**
            Calculate $f_i = \frac{p_i F}{P_A} + (1 - F)\left(\frac{p_i \hat{P}_A - \hat{p}_i P_A}{P_A}\right)$.
        else
            Set $f_i = 0$.
else
    Set $f_i = 0$ for all i.
return $f_1, \ldots, f_n$

Remark 5.2. *Note this (Claim 15) suggests a greedy approach, adding outcomes one by one based on the condition. This is equivalent to the iterative refinement shown here and also converges to the correct active set for the given F. The condition used in Claim 15, $F[p_i \hat{Q}_A / P_A] + [p_i \hat{P}_A / P_A - \hat{p_i}] > 0$, is algebraically equivalent to the one used in Algorithm 2. Once A is determined, the fractions f_i are computed using (6).*

6 Fractional Kelly Betting

Often, bettors choose to wager only a fraction, say $k \in (0, 1)$, of the optimal Kelly bet sizes. This is known as fractional Kelly betting. There are two common interpretations:

1. Scale the individual optimal fractions: $f_{i,frac} = k \cdot f_{i,opt}$, where $f_{i,opt}$ is calculated using Algorithm 1. The total wagered amount becomes $F_{frac} = kF_{opt}$.

2. Set the total wagered amount to a fraction of the optimal total: $F_{target} = kF_{opt}$, and then find the optimal allocation for this fixed total F_{target} using Algorithm 2.

Interpretation 2 is generally preferred as it re-optimizes the allocation for the reduced risk level (lower total F). Simply scaling the fractions (Interpretation 1) might lead to suboptimal growth for the target risk level F_{frac}.

Algorithm 2 directly handles Interpretation 2. We first compute the optimal full Kelly fractions using Algorithm 1 to find $F_{opt} = \sum f_{i,opt}$. Then we set $F_{target} = kF_{opt}$ and run Algorithm 2 with this F_{target}.

The note in Claim 16 suggests an alternative way to find the active set for F_{target}: start with the full Kelly active set A_{opt} and remove outcomes m if $f_m(F_{target}) < 0$, typically starting with the one with the lowest $p_m d_m$ among the candidates for removal. This works because, as shown in Theorem 3.6(b), outcomes leave the active set in order of increasing $p_i d_i$ as F decreases. So, starting from A_{opt} and removing outcomes in reverse order of $p_i d_i$ (lowest first) if their calculated $f_i(F_{target})$ using (6) with the *current set A becomes negative is another valid way to find the correct active set for F_{target}. This avoids starting from an empty set if k is close to 1.*

7 Conclusion

We have presented a clear and rigorous derivation of the optimal Kelly wagering strategy for $n \geq 2$ mutually exclusive outcomes. By applying KKT conditions, we derived explicit formulas for the optimal wager fractions (1), the total fraction wagered (2), and the expected logarithmic growth (4), connecting the latter to KL divergence.

We characterized the active set A using the condition $p_i/\hat{p}_i > Q_A/\hat{Q}_A$, explaining why positive expectation bets are always included, zero expectation bets are never included, and negative expectation bets may be included as conditional hedges.

Furthermore, we analyzed the behavior of the optimal strategy when the total wager F is fixed, showing the linear scaling of fractions within regions of constant A (6) and determining the critical F values (7) where outcomes become inactive, ordered by $p_i d_i$.

These results lead to efficient algorithms (Algorithms 1 and 2) for computing optimal full Kelly and fixed-F (including fractional Kelly) portfolios. This work provides a comprehensive and accessible framework for applying the Kelly criterion to betting scenarios involving multiple outcomes.

References

[1] Kelly, J. L., Jr. (1956). A New Interpretation of Information Rate. Bell System Technical Journal, 35(4), 917–926.

[2] Smoczynski, Peter; Tomkins, Dave (2010). An explicit solution to the problem of optimizing the allocations of a bettor's wealth when wagering on horse races. Mathematical Scientist, 35 (1): 10–17.