

The Kelly Criterion in the Modern Day

Nang Dang, Timothy Woods

April 2026

Abstract

This paper explores the mathematical foundations and applications of the Kelly Criterion, a strategy used to maximize long-term compounded growth under uncertainty. We first develop the prerequisite concepts of probability needed to understand the Kelly Criterion, before deriving the Kelly Criterion and examining its applications in fields such as gambling, sports betting, and the stock market. Furthermore, we will explain its importance in modern optimization mathematics, something that has not been heavily researched yet. This paper is meant to provide a foundational knowledge base of the Kelly Criterion to permit a reader to apply the Kelly Criterion to real-world scenarios.

Contents

1	Introduction	3
2	Basics of Probability	3
2.1	Definitions and Notation	3
2.2	Axioms	4
2.3	Combinatorics	4
2.3.1	Fundamental Principle of counting	4
2.3.2	Permutations	5
2.3.3	Combinations	5
3	Prerequisites	5
3.1	Random Variables	5
3.1.1	Discrete and Continuous Random Variables	6
3.1.2	Function of a Random Variable	6
3.1.3	Bernoulli Variable and Binomial Distribution	6
3.1.4	Geometric Variable	7
3.2	Expected Value	7
3.2.1	Linearity of Expectation	8
3.3	Variance	8
4	The Kelly Criterion	9
4.1	Introducing the Kelly Criterion	9
4.2	Deriving the Kelly Criterion	9
4.3	Fractional Kellyes	10
4.4	Kelly Criterion for multiple outcomes	11
5	Applications of the Criterion	12
5.1	Simple game	12
5.2	Sports Betting	14
5.3	Stock Market	15
5.4	Quant and algorithmic applications	15
6	Avenues for further research	16
7	Conclusion	16

1 Introduction

The Kelly Criterion is a foundational element of modern financial operations and has numerous potent applications, some of which include gambling, algorithmic investment, and deep-learning investment. It was created in 1956 by John Larry Kelly Jr, working along with Claude Shannon. Touching on the topic of applications, we decided to write this paper in order to bridge the gap between old and behind-the-times literature on the Kelly Criterion and its new useful applications in the modern world. Hence, a majority of this paper will be spent examining the various applications of the Kelly Criterion in modern day real world scenarios.

2 Basics of Probability

This section covers the basics of probability and combinatorics, so that the rest of the prerequisites section is comprehensible to readers. The most basic definition of probability is the measure of how likely an event is to occur.

In mathematics, probability is formally defined as

$$P = \frac{\# \text{ of favorable outcomes}}{\# \text{ of total outcomes}}$$

It is common to use the notation $P(A)$ instead of just P , where A can be any term representing a random occurrence. This will be elaborated upon shortly.

2.1 Definitions and Notation

To further build upon this basic definition of probability, we need to discuss some more common definitions and notations used.

Definition 1 *The sample space is the set of all possible outcomes. For example, the sample space of rolling a six sided die is $S = \{1, 2, 3, 4, 5, 6\}$.*

Definition 2 *An event is a set consisting of possible outcomes in an experiment. It is a subset of the sample space. It is usually denoted with a capital letter like A . For example, if E is the event that rolling a dice lands on an even number, $E = \{2, 4, 6\}$ and if O is the event that rolling a dice lands on an odd number, $O = \{1, 3, 5\}$.*

Definition 3 *$P(A)$ is read “probability of event A ”. In other words $P(A)$ represents the probability that event A occurs.*

Definition 4 *$E \cup O$ is called the union of events E and O . The union of two events consists of all outcomes that are either in E or in O or in both E and O . Revisiting our dice example where E is the event that the dice lands on an even number and O is the event that the dice lands on an odd number, then $E \cup O = \{1, 2, 3, 4, 5, 6\}$ which is the whole sample space S .*

Definition 5 $E \cap O$ is called the intersection of events E and O . The intersection of two events is the outcomes which occur in both events. Again, going back to our dice example, $E \cap O = \{\}$. Note that this is an empty set because there is no outcomes which occur in both events.

Definition 6 A *mutually exclusive event* is an event that does not share any common outcomes with another event, which we just saw with E and O .

Definition 7 N factorial is denoted as $n! = n(n-1)(n-2)\dots(1)$.

2.2 Axioms

Building upon the basic definitions and the notations we have just introduced, we will now present the three axioms of probability. Note that an axiom is a fundamental principle that is accepted as true without proof, serving as the starting point for further reasoning. These axioms are quite intuitive.

Axiom 1

$$0 \leq P(E) \leq 1$$

Axiom 2

$$P(S) = 1$$

Axiom 3

For any sequence of mutually exclusive events E_1, E_2, \dots

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

2.3 Combinatorics

Combinatorics is the field of mathematics that is used for counting. Its most prominent use is in probability. It is essential for probability because it is used for counting the number of outcomes in an experiment, whether those are the favorable outcomes or total outcomes. To start, we will present a foundational question in combinatorics:

Suppose we want to select k objects from a set of n objects. How many different selections are possible?

2.3.1 Fundamental Principle of counting

To tackle this question, we need to first understand the fundamental principle of counting, which states

If one event can occur in m ways, and another event can occur in n ways, then the total outcomes of both events occurring is $m \times n$.

2.3.2 Permutations

Now, suppose that the order in which we select the k objects is important. This is known as a **permutation**.

When we select the first object, there are n choices. When we select the second object there are $n - 1$ choices. This continues on until we select the k th object, where the amount of choices will be $n - k + 1$. Hence, from the fundamental principle of counting we get,

$$n(n - 1)(n - 2)(n - 3)\dots(n - k + 1)$$

Rewriting using factorials we get

$$\frac{n!}{(n-k)!}, \text{ which is known as } {}_n P_k.$$

This counts the ordered selection of k objects from n total objects.

2.3.3 Combinations

Now suppose in our question we don't care about the order in which the k objects are selected. For any k objects selected, there are $k!$ ways to order them because for the first object there are k options, and for the second we must take into account that one object has already been selected, thus there are $k - 1$ options, et cetera.

Therefore, each combination is counted $k!$ times in the ordered count. Hence, we have to divide by $k!$ to correct the overcounting:

$$\frac{\frac{n!}{(n-k)!}}{k!} = \frac{n!}{(n-k)!(k!)}$$

This is commonly denoted as $\binom{n}{k}$ or ${}_n C_k$, which means the amount of ways in which we can select k objects from a total of n objects.

3 Prerequisites

3.1 Random Variables

Random Variables are functions whose output depends on a random process. They are very useful as a way to define situations with a random outcome because they give mathematical values to each outcome and allow us to perform operations and extract information. The outputs of a random variable can be any of a set of events, each with a requisite probability. These probabilities may be the same or differ. To formally define a random variable X ,

$$X : \Omega \rightarrow \mathbb{R}$$

where Ω represents all possible results in the sample space, and \mathbb{R} is a set of real numbers which are the quantified possible outcomes of the random process.

To use the previous example of dice, a random variable equivalent to a fair dice roll is defined with six outputs $\{1, 2, \dots, 6\}$ each with a probability of $\frac{1}{6}$, which would be Ω ; one could consider it a blank cube. \mathbb{R} is the number depicted on each face of the cube, making it a die.

3.1.1 Discrete and Continuous Random Variables

Ω can be broken down further. If it consists of a discrete set, with countable values, then $X : \Omega \rightarrow R$ is a discrete random variable, and if it is part of a continuous set, with uncountable values, then it is part of a continuous random variable. They have many differences in their characteristics.

A function describing the likelihood of a particular output for a discrete random variable is called a probability mass function, while for a continuous random variable it is called a probability density function. Both mass and density functions define the shape of the distribution. Since the density function is only a model, each individual outcome has an infinitesimal probability. Thus, they must be used to find the probability of an interval of outcomes.

The cumulative distribution can be used to find the probability of any value less than a :

$$\int_{-\infty}^a f(x)dx$$

If $f(X)$ is a probability density function, then the probability of values between a and b is

$$\int_a^b f(x)dx$$

3.1.2 Function of a Random Variable

A function (eg $f(x) = x^2, g(a) = 2x$) applied to a random variable becomes another random variable because it meets the requirement of mapping a set of outcomes to a set of probabilities.

3.1.3 Bernoulli Variable and Binomial Distribution

The Bernoulli Variable is a discrete random variable with only two possible values, 0 or 1, representing success and failure.

The Binomial distribution is the sum of multiple Bernoulli trials. It has two parameters, the chance of success p and the number of trials to run, n . Its probability mass function is

$$P(X = K) = \binom{n}{k} p^k (1 - p)^{n-k}$$

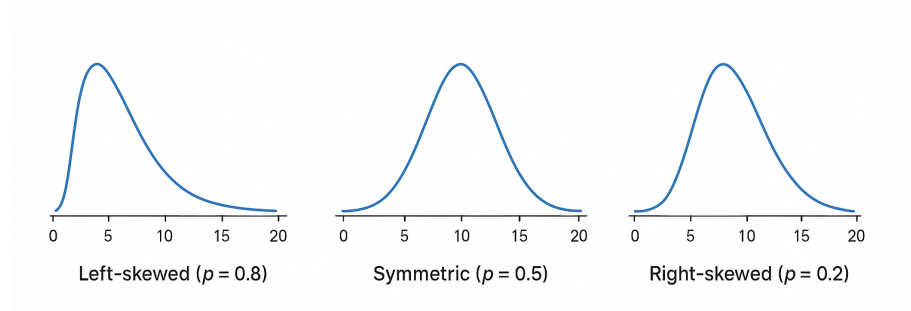
where k is the number of successes.

The distribution of a binomial variable can be described as

$$P(X \leq i) = \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}$$

where $\lfloor i \rfloor$ is the greatest possible value up to i , which is the upper bound of the amount of successes we measure.

Shape of Binomial Distribution



Source: hodausama.github.io

3.1.4 Geometric Variable

The geometric random variable is a descendent of the Bernoulli variable. It describes the number of Bernoulli trials needed to achieve the first success. The probability that a random variable X is equal to n trials is

$$(1-p)^{n-1}p$$

3.2 Expected Value

One of the most essential topics we need to discuss before diving into the Kelly Criterion is the expected value of a random variable. We will use our previous definition of what a random variable is, and the probability mass function to introduce the formula. For a discrete random variable we have the formula,

$$E[X] = \sum_i^n x_i p(x_i)$$

Where x_i is the random variable, and $p(x_i)$ is the probability mass function. You will see the use of expected value come into play later in the derivation section of the apper where we need to take the expected value of the logarithm of wealth.

The expected value for a continuous variable differs. The formula goes

$$\int_{-\infty}^{\infty} xf(x)dx$$

where x is the value of the random variable and $p(x)$ is the probability density function.

3.2.1 Linearity of Expectation

Expected value has properties of linearity, meaning

$$E[X_1 + X_2 + X_3 \dots + X_{n-1} + X_n] = E[X_1] + E[X_2] + E[X_3] + \dots E[X_{n-1}] + E[X_n]$$

and also

$$E[cX] = cE[X]$$

where c is a constant.

3.3 Variance

Variance is the measure of how often and how much the observed value differs from the expectation. It can be calculated in different ways depending on the variable and distribution, and is symbolized as σ^2 . However, it is always expressed in square units.

The Standard Deviation is more convenient to use because it is expressed in the same units as the data, and is symbolized as σ . It can be used to compute how many observed values fall within a range of the expectation. As such, standard deviation is very commonly used as a measure of reliability, for example in the sense of measuring what proportion of products are within an acceptable quality tolerance of the mean.

However, the variance can always be generalized as $\text{Var}(X) = E[(X - E(X))^2]$. Therefore:

1. $\text{Var}(X) \geq 0$
2. $\text{Var}(a) = 0$ where a is a constant
3. $\text{Var}(X + a) = \text{Var}(X)$
4. $\text{Var}(aX) = a^2\text{Var}(X)$
5. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ when X and Y are independent

The **variance** of a continuous random variable is still $E[X^2] - (E[X])^2$, and can be computed as follows, where $f(x)$ is the variable's probability density function:

$$\int_{-\infty}^{\infty} x^2 f(x)dx - E(X)^2$$

Since in future sections we will be talking about the stock market and volatility, it is important to note that $\text{Volatility} = \sqrt{\sigma^2} = \text{Variance}$.

4 The Kelly Criterion

4.1 Introducing the Kelly Criterion

The Kelly Criterion, discovered in 1956 by John L. Kelly, [2] is a formula for the optimal fraction of one's current wealth to apply to some game, which will yield the greatest expected value of logarithmic growth of wealth. It must be applied to a situation wherein $P(\lim_{n \rightarrow \infty} X_n = +\infty) > 0$, where X_n is the wealth possessed by the player after n trials. To put it simply, this means that the situation does not have a case in which the player always fails to gain money. [1]

$$f^* = \frac{p}{l} - \frac{q}{g}$$

where f^* is the fraction of current wealth to bet, p is the probability of gaining value, l is the fraction of the bet lost when losing money, q is the probability of losing value, and g is the fraction of the bet gained when p is true. [2]

Particularly, the Kelly optimizes for geometric growth rate. This is because optimizing for the expected value of wealth is unrealistic in that it calculates that one should always invest all of one's money, which while in theory can produce the greatest amount of money, in practice is almost guaranteed to bring complete loss eventually.

To put it simply, geometric growth rate is the growth rate of a variable with a value which changes via repeated multiplication, just as how with games one's bet can be multiplied by a certain amount for winning or taken away or otherwise reduced for losing. Another term for this is the logarithm of wealth. [2]

4.2 Deriving the Kelly Criterion

The Criterion can be derived into various forms for various applications such as gambling with different games, and for investing inside the stock market. However, the most basic form described above can be found as follows: Given that the objective is to optimize the expected value of the growth of our wealth, we use f to represent the fraction of our wealth that we bet, starting with 1 unit of wealth. Then, as standard characteristics of a game, we have probability of favorable outcome p , unfavorable outcome q , ratio of loss to profit a and ratio of profit to loss b . If we win, we now have $1 + fb$ units of wealth, and if we lose, $1 - fa$. This indicates that our geometric growth rate, the percentage change per period, r , can be found as:

$$r = (1 + fb)^p \cdot (1 - fa)^q$$

This formula modifies the winnings by the chance of winning and the losses by the chance of loss, then takes the product of the two to describe the direction and rate of growth. In order to make it logarithmic, we take the logarithm of both sides:

$$\log(r) = p \log(1 + fb) + q \log(1 - fa)$$

We do this in order to derive more conveniently, and we can note $\log(r)$ as E . Now in order to learn the optimal f which earns the greatest $\log(r)$, we can take the derivative of E with respect to f and set the derivative equal to zero in order to get

$$\left. \frac{dE}{df} \right|_{f=f^*} = \frac{pb}{1+f^*b} + \frac{-qa}{1-f^*a} = 0$$

Finally, we can rearrange this to solve for f^* , the value of f that gives the greatest geometric growth rate.

4.3 Fractional Kellies

Although the Kelly Criterion makes money in the long run, there is still volatility, risk of ruin, and the model does not take into account risk aversion. Hence, many investors feel safer, and better emotionally betting a fraction of what the Kelly Criterion suggests. For example, a half Kelly means you bet half the amount the Kelly Criterion tells you to bet. Hence, we develop the equation

$$f = cf^*$$

Where c is a constant between 0 and 1 and f^* is the original Kelly we discussed before. The fractional Kelly is very significant because great investors such as Warren Buffet, and Bill Gross take into account fractional Kellies when they are trying to manage their risk while investing or gambling. [2] Now in order to look at a situation where fractional Kellies are useful, we must talk about two important figures, that being Edward O. Thorp and Harry Markowitz.

Edward O. Thorp, a pioneer of the Kelly Criterion, and a name you will read numerous times throughout this paper, goes into great depth as to why fractional Kellies should be utilized. Harry Markowitz on the other hand, is a famous American economist who Thorp mentions in order to make an argument as to why fractional Kellies have significance. He first compares the Kelly Criterion to Harry Markowitz's mean-variance portfolio theory. This theory focuses on one-period returns, and its goal is to identify portfolios on the efficient frontier. A portfolio in this state means that no other portfolio yields both higher returns and lower variance for the given investment. This is applicable to fields such as investing and gambling. Note that the Kelly Criterion contrasts this theory because it is used for multi-period returns since it aims to optimize compounded growth. Thorp introduces Markowitz's theory in order to propose the question [2]

Is there an efficient frontier for long-term compounded growth, similar to Markowitz's frontier for one-period returns?

That is, is there a strategy such that no other strategies offer both higher expected returns and lower variance for multi-period compounded growth?

Thorp concluded after many experiments that portfolios for long-term compounding growth lying on the efficient frontier consists of **fractional kelly strategies**.

Why is this true?

1. Fractional Kelly allows an individual to better manage risk.

For $0 \leq c \leq 1$

As c increases expected value increases, but so does variance. Hence, one could see why changing the constant depending on the security you are investing in is beneficial. For example, if you are investing into a small biotech stock with great upside potential but very high volatility, you should probably use a smaller c value as opposed to investing into an established company like Costco.

2. Constant fractions outperform changing bet sizes.

If an investor changes bet sizes over time but keeps the same average exposure, (e.g. betting 10%, 10%, 10% is the same average exposure as betting 20%, 10%, 0%) using a constant fraction yields higher expected returns. This is due to the concavity in the logarithmic function used in the analysis of the Kelly Criterion, so uneven betting sizes will yield lower expected returns.

3. Overbetting is inefficient.

For $c \geq 1$

Thorp proved that expected value decreases, while variance increases. A primary reason of this is due to **Black Swans**, which are high-impact unpredictable events that can lead to major losses. When one of these events occur and you are betting higher than what the Kelly Criterion suggests, losses are devastating and take a long time to compound back again, in which case another Black Swan event may occur. [2]

4.4 Kelly Criterion for multiple outcomes

The basic Criterion only accounts for betting on one item, and does not account for how much should be bet when there are multiple favorable bets and outcomes, and there is no single optimal choice.

To begin with, we can establish a set of outcomes:

n number of outcomes p_1, p_2, \dots, p_n probabilities, or p_i , for each outcome b_i net odds (the multiplier of profit excluding the wager) for each outcome b_1, b_2, \dots, b_n

A wager fraction of x , meaning that we bet xM units of currency

Based on the expected value definition

$$E(X) = \sum_i^n p_i x_i$$

we can take the expected value of the logarithm of our wealth after making any bet: $\sum p_i \log((1-x)M + xM + b_i xM)$ This is because we multiply each probability of an outcome p_i by the logarithm of our total amount of money; $(1-x)M$ is the money we did not wager, xM is our wager, and $b_i xM$ is our profit. This can be simplified down to $\sum p_i \log(1 + b_i x) + \log(M)$ because by the product property of logarithms we can extract M from every term within the log and then add like terms. In order to maximize the expected value of growth of our wealth, we must maximize $\sum p_i \log(1 + b_i x)$, which is the expected value of gross growth on our bet.

5 Applications of the Criterion

This section will talk about the appliance of the Kelly Criterion to real world scenarios.

5.1 Simple game

Lets familiarize ourselves with the original Kelly Criterion formula.

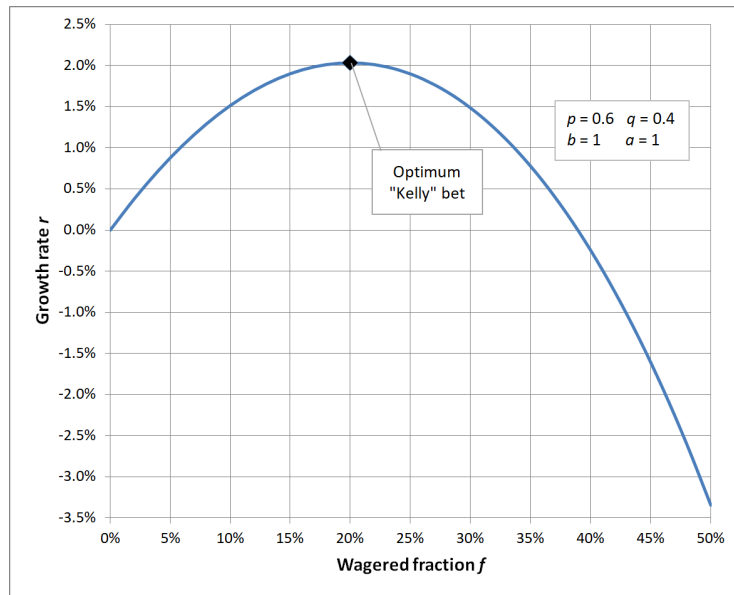
$$f^* = \frac{p}{l} - \frac{q}{g}$$

Suppose that there is a betting game with these specific circumstances

- You have a 60% chance of winning $\rightarrow p = 60\%$
- You have a 40% chance of losing $\rightarrow q = 40\%$
- You gain 100% of your bet when you win $\rightarrow g = 1.0$
- You lose 100% of your bet when you lose $\rightarrow l = 1.0$

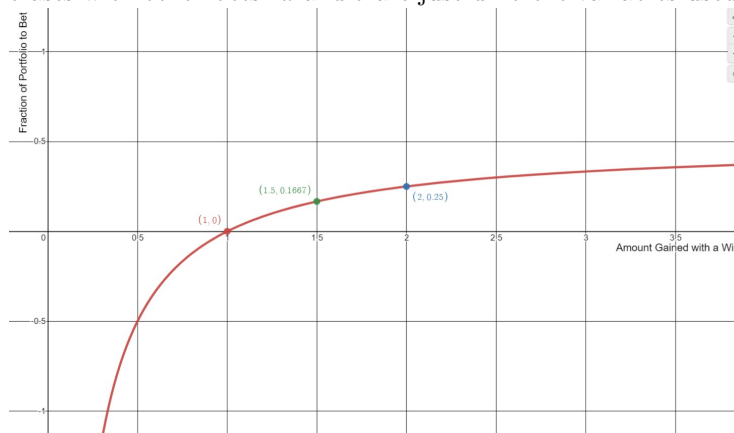
Plugging in the values to our formula we get $f^* = .2$

Which means that with these circumstances the Kelly Criterion tells us to bet 20% of our wealth with each turn.



Source: Wikipedia

Note this graph from our experiment that shows how the Optimal Kelly bet contrasts with other bets. a and b are just different variables used for g and l .



Source: Wikipedia

This graph assumes that $p = q = .5$ and the amount gained and loss is 100% of the bet. Notice how the graph trails off at the end, showing that even when you are expected to earn a ridiculous amount of money, it is not wise to bet a large fraction of your wealth due to the risk of ruin.

5.2 Sports Betting

One prominent example of the Kelly Criterion being super successful in the real world is when Edward O. Thorp applied it to sports betting. In 1994, Thorp used a computer-based model with the Kelly Criterion to size wagers. He started with 50,000\$ and in 101 days, he turned it into 173,000\$. [3] Though many might attribute this to luck, this can not be the entire truth because the total profit graph matched very closely with the expected profit graph, never tailoring a few thousand off of it.

Another thing to note is that in sports betting, in contrast to many other wagers, bettors are often placing multiple bets simultaneously. This makes the math much more interesting. Let us see how that works within a scenario where we aim to optimize two bets simultaneously. Note that this example can of course be extended to n amount of simultaneous bets.

First, let us define a few variables.

- f_1, f_2 = fractions of bankroll placed on bet 1 and 2
- R_1, R_2 = returns of each bet (random variables)
- W_o = initial bankroll \$
- W = final bankroll \$
- $g(f_1, f_2)$ = maximizing the growth of two bets

Using similar steps to derive the original Kelly formula we have

$$W = W_0(1 + f_1R_1)(1 + f_2R_2)$$

Each bet independently scales wealth and they compound multiplicatively. Then, we take the natural log, just as before when we derived the original formula.

$$g(f_1, f_2) = E[\ln W]$$

Doing substitution and some algebra we arrive at

$$g(f_1, f_2) = E[\ln(1 + f_1R_1)(1 + f_2R_2)]$$

and finally

$$g(f_1, f_2) = E[\ln(1 + f_1R_1) + \ln(1 + f_2R_2)]$$

There are two important cases we must consider when working with simultaneous bets.

Case 1: Independent bets: This is the simpler case. When you place independent bets the problem splits and you can just apply Kelly to each bet. By split, that means $\ln a * b = \ln a + \ln b$, which is much easier to optimize than our next case.

Case 2: Dependent bets: Now what if the bets correlate with each other? For example, the bet that LeBron scores 50 points is not independent of the bet that the Lakers win a game because if that were to occur the probability that the Lakers win the game is drastically increased. In this scenario, Thorp states that

$$E[\ln((1 + f_1 R_1)(1 + f_2 R_2))] \neq E[\ln(1 + f_1 R_1) + \ln(1 + f_2 R_2)]$$

That is the bets are not separable. [3]

A quick understanding as to why dependent bets are not able to be optimized individually is because their outcomes will affect each other, namely that is R_1 is not independent of R_2 . Thus, correlated bets will spike variance greatly.

From here, you would need to optimize a joint probability distribution in order to find f_1 and f_2 . Optimizing joint probability distributions require an extensive amount of calculus, so therefore it will not be covered in this paper. Generally though, after extensive optimization, you will find that the Kelly fraction must be reduced in order to account for the huge spike in variance.

5.3 Stock Market

One of the most prominent uses of the Kelly Criterion is in the stock market. In mathematical finance, when a security maximizes the logarithm of wealth, then a portfolio is growth optimal, which we know that the Kelly Criterion aims to do. Through manipulation of the original Kelly Formula, the Kelly Criterion shows that for a given security with volatility, the growth optimal portfolio is achieved when [1]

$$f^* = \frac{\mu - r}{\sigma^2}$$

Where f^* is the fraction of available capital invested in order to maximize the growth rate of a security, μ is the expected growth rate coefficient (e.g. it represents the estimated growth in a stock's price per period, usually per year), σ^2 is the variance of the growth rate coefficient, and r is the risk-free rate of return (e.g. the U.S 1 year Treasury Bill which sits at around 3.71%).

This formula essentially tells you

- high expected returns \rightarrow invest more
- high variance \rightarrow invest less

Note that when using this version of the Kelly Criterion for the stock market one has to make an educated guess about what the variance, expected growth rate coefficient, etc, are. This means that these computations can suffer garbage in, garbage out problems as Thorp states [1]. Since one has to make an educated guess about what the inputted variables are, estimates that are off slightly can lead to huge changes in f^* and thus tremendous losses. Hence, that is why many investors when using the Kelly Criterion for investing use fractional Kellys in order to manage risk. [1].

5.4 Quant and algorithmic applications

The Kelly Criterion's derivatives are used heavily in a number of financial fields, the most promising of which being algorithmic investment. One example of this

application is what is described in Jincheng Zhang's paper *A Dynamic Robust Kelly Optimization Framework via Dual Semidefinite Programming* which describes that the Kelly Criterion lacks a way to account for the influence of certain investments on others as well as for uncertain constraints of the investor and investment. Zhang uses a Kelly model in the form of a complex optimization method involving advanced matrix mathematics called Dual Semidefinite Programming in order to account for these issues. [4]

It is for this reason, generally speaking, that the Kelly Criterion is so useful. It is a foundational tool to optimize gains and though it may be altered to fit its applications, the logic behind it remains.

6 Avenues for further research

There is presently much new research in the use of the Kelly Criterion in the field of investment, but there are still many unanswered questions. The most basic among them are how to compensate for the Criterion's weaknesses; it requires quality input to produce quality output. Machine Learning is already being put to use to enhance our inputs, but what scope is most effective? Can data from the economy as a whole, or perhaps the news, be incorporated to make useful predictions ?

Another weakness is that, although generally applicable in many cases, it is still a simple principle and may not be the most efficient option. Can we use machine learning to modify the Criterion and tailor it to individual situations, so that it is more efficient? Is it possible to discover a better pattern using the predictions of neural networks?

7 Conclusion

The Kelly Criterion is extremely useful as a foundational tool and regarded as almost universal within the financial field. However, it is not without issues; among them, that although it maximizes growth in the long term, it is volatile and risky to use in the short term. [5]. Another criticism of the method is that, according to Paul Samuelson, when one determines one's wealth with any one of $H(w) = \frac{-1}{w}$, $K(W) = \log(W)$, $T(W) = W^{1/2}$ and uses another of the three to determine f^* with Kelly, one will lose money. This criticism is true because, according to Thorp and Whitley, comparing through Kelly two functions defined and differentiable over the interval $(0, \infty)$ and having a positive and decreasing derivative, if they always have the same f^* , they are equivalent. Thus, under Samuelson's conditions, they will decrease over time.

References

- [1] Louis M. Rotando and Edward O. Thorp. The kelly criterion and the stock market. *The American Mathematical Monthly*, 99(10):922-931, 1992.

- [2] Edward O. Thorp. *Understanding the Kelly Criterion*, chapter 36, pages 509–523.
- [3] Edward O. Thorp. Chapter 9 - the kelly criterion in blackjack sports betting, and the stock market*. In S.A. Zenios and W.T. Ziemba, editors, *Handbook of Asset and Liability Management*, pages 385–428. North-Holland, San Diego, 2008.
- [4] Jincheng Zhang. A dynamic robust kelly optimization framework via dual semidefinite programming.
- [5] William T. Ziemba and Leonard C. MacLean. *Using the Kelly Criterion for Investing*, pages 3–20. Springer New York, New York, NY, 2011.