

Optimal f and the Kelly Criterion

by Ralph Vince

Abstract

Keywords: Geometric growth optimization, Kelly criterion, risk, gambling, markets.

Widely-accepted in the gambling and trading community, the Kelly Criterion, named after John L. Kelly, whose 1956 Bell Labs Technical Journal paper presented the criterion resulting in wagering a constant, optimal fraction of the gambler's stake which results in maximizing the growth of the gambler's stake in the case of the gambler possessing inside information, is compared to Optimal f .ⁱ

Repeatedly in literature and commentary the notions of the Kelly Criterion and Optimal f are mistakenly conflated. The two are different and the former should not be used in assessing trading quantity except under certain circumstances. Optimal f yields the correct optimal fraction of an account to wager in all cases. This paper will attempt to distinguish the two, as well as provide means for translating between them.

The Optimal f calculation provides a bounded context for studying the nature of the curve whose optimal point, i.e. peak, represents the correct fraction of a stake to risk to result in the greatest geometric growth asymptotically. It is the nature of the curve whose bounding allows us to study the different phenomena of the curve, as well as provide a context from which we pursue criteria other than mere growth maximization.

Since Optimal f affords us this context, this paper seeks to examine another phenomenon inherent in Optimal f , which becomes evident when contrasted to the Kelly Criterion.

Introduction

The Kelly Criterion does not yield the Optimal Fraction to Risk in Trading Except in a Special Case

The Kelly Criterion does **not** solve for the optimal "fraction" to allocate to a trading situation except in a special case, whereas Optimal f does solve for the optimal fraction to risk in all cases.

Optimal f does **not** satisfy the Kelly Criterion (except in the special case).

The two notions are similar (their mathematical relation forthcoming) but different. To conflate the two is a mistake, and doing so in trading applications often leads to the unintended (and often dangerous) miscalculation of the quantities which one should assume so as to maximize asymptotic geometric growth.

Kelly discusses discerning fractions of a gambler's stake to risk in maximizing what is a gambling outcome, i.e. a binomial outcome (which implicitly may be extended to more than two outcomes) and he presents the required mathematics (a). In his conclusion he asserts that geometric growth is maximized by the gambler betting a **fraction** such that, 'At every bet he maximizes **the expected value of the logarithm** of his capital.'ⁱⁱ

Therein is the Kelly Criterion. The **fraction** of one's stake to bet, in order to maximize the long-run growth of one's capital, is that fraction which maximizes the expected value of the logarithm of his capital (or sum of the logs of the returns when the probability associated with each data point is the same). In other words, if we look at a stream of n returns on our capital, $A_1...A_n$, where each return is weighted by a variable, f , with a probability associated with each of the n returns, $P_1 .. P_n$, the expected value of the logarithm of our capital, the Objective Function in (1), is:

(1)

$$ObjectiveFunction = \sum_{i=1}^n \ln(1 + A_i * f) * P_i$$

According to Kelly, the value for f that maximizes the objective function is the fraction that results in the greatest long-run growth of capital to a gambler. Thus, the value for f that maximizes (1) is that value which is said to satisfy the Kelly Criterion (b).

Rather than taking the sum of the logs of the returns, we can take the product of those returns. Thus, the value for f that maximizes (1) will also maximize:

(1a)

$$ObjectiveFunction = \prod_{i=1}^n (1 + A_i * f)^{P_i}$$

Contrast this, the formula that satisfies the Kelly Criterion, with the formula for Optimal f , where the objective function, G , is the Geometric Mean Holding Period multiple:

(2)

$$G = \prod_{i=1}^n \left(1 + \frac{X_i}{\left(\frac{-W}{f} \right)} \right)^{P_i}$$

Whereas the Kelly Criterion solution uses returns, A_i , the Optimal f solution uses actual outcomes, X_i , based on a user-defined, consistent quantity (e.g. 100 share lot) and W is the largest losing

data point of the $X_1 \dots X_n$ data points.
 $W = \min\{X_1 \dots X_n\}$.

Clearly, the Kelly Criterion when restated in terms of products (1a) so that it is compared formulaically on an apples to apples basis with Optimal f (2), rather than sums of logarithms (1), is not the same. They do not yield the same answers for the values that maximize them except in the special case.

The value for f which maximizes (1,1a,1b|r=0) is the same as the f which maximizes (2) only in what is referred to herein as the “special case” in trading defined as:

1. $-W$ = the price of the underlying instrument when purchased, **and**
2. The position to be assumed is a long position only.

When one or both of these conditions are **not** met, the Kelly Criterion (1,1a,1b|r=0) not only results in a different value (for the optimal fraction to bet) than does the Optimal f solution (2), but can often result in a number that is greater than unity. This is because, as explained later, the Kelly Criterion doesn’t produce an “optimal fraction to bet,” but rather a leveraging factor. These numbers are identical only in the “special case.”

In the more common cases, the value that solves for the Kelly Criterion is not the optimal “fraction” of a trading account to risk. In all cases, the Optimal f solution will yield the correct growth-optimal fraction to wager. Thus, the Optimal f solution is a more generalized solution of which the Kelly Criterion is a subset, applicable in trading only when both conditions of the “special case” are satisfied. When these conditions are not both met (as is typically the case in trading) one must rely on the more generalized Optimal f solution (2) to yield the optimal fraction to risk.

Both conditions of the special case are met in a gambling situation. In such situations, the value for f which maximizes (1,1a,1b|r=0) is the same as the f which maximizes (2), and thus the Kelly Criterion yields the same value as the answer provided by the Optimal f solution.

Let us consider the ubiquitous case of a fair coin which when tossed will pay \$2 on heads and -\$1 on tails. This

situation meets both criteria required for the value for f which maximizes (1,1a,1b|r=0) being the same as the f which maximizes (2).

We find the objective function in (1) maximized where $f = .25$ wherein we have:

$$\begin{aligned} &= \ln(1+2*.25) * .5 + \ln(1+-1*.25) * .5 \\ &= \ln(1.5) * .5 + \ln(.75) * .5 \\ &= .405465 * .5 + -.28768 * .5 \\ &= .202733 - .14384 \\ &= .058892 \end{aligned}$$

The expected value of the logs of returns in this case is .058892 and maximized at $f = .25$.

(Substituting (1a) for (1), we find the objective function still maximized at a value where $f=.25$.)

Similarly, solving for f to maximize (2) again yields an f value of .25:

$$\begin{aligned} &= (1 + 2 / (-1 / .25))^5 * (1 + -1 / (-1 / .25))^5 \\ &= (1 + 2 / (1 / .25))^5 * (1 + -1 / (1 / .25))^5 \\ &= (1 + 2 / 4)^5 * (1 + -1 / 4)^5 \\ &= (1 + .5)^5 * (1 + -.25)^5 \\ &= 1.5^5 * .75^5 \\ &= 1.224745 * .866025 \\ &= 1.06066 \end{aligned}$$

The result of the objective function for (2) is the geometric average return per play as a multiple. That is, it represents the multiple made on our stake, on average, each play (or compounding period) when we reinvest profits and losses.

An analog situation in trading (of the “special case” i.e., the gambling case) is that where:

1. $-W$ = the price of the underlying instrument when purchased, **and**
2. The position to be assumed is a long position only.

We find the Kelly Criterion and Optimal f yield the same optimal fraction of our stake to risk(c).

This two to one coin toss gambling game is analogous to a trading situation where the price of the stock is \$1 per share and the worst-case loss is \$1 per share. The distribution of outcomes of what might happen to this trade is entirely described by the two simple scenarios. Either we exit the trade at \$2 per share or we lose the entire investment.

Now, let us consider the case where the price of the stock is \$1 per share but the most we can lose (the “worst-case outcome”) is -.8 rather than -1.0. We are now faced with two possible scenarios: exit at .2 or 2.0. This is equivalent to a coin toss scenario where we either win two or lose .8.

The Kelly Criterion in this case would have us wager .375 of our stake to optimize growth in such a situation (whether using (1, 1a, 1b|r=0) as all give the same value for f as that which optimizes each objective function).

Optimal f , on the other hand, has us wager .3 of our stake to maximize growth (d).

Now let us examine what happens as the size of the loss continues to shrink, from minus one, which qualifies as a “special case” where the optimal fraction determined by both methods is the same to -.1. See Table 1.

Table 1

		Optimal fractions given by:	
Heads p (.5)	Tails p (.5)	(1) Kelly Criterion	(2) Optimal f
2	-1	0.25	0.25
2	-0.8	0.375	0.3
2	-0.5	0.75	0.375
2	-0.25	1.75	0.4375
2	-0.1	4.75	0.475

Notice how in all but the special case the growth optimal fractions returned by the objective functions for the Kelly Criterion and Optimal *f* are **not** the same and the values that optimize the objective functions differ. To be a “fraction” implies a number bounded at zero and one inclusively. We see here that when we deviate from the special case, the objective function of the Kelly Criterion is maximized by a value greater than one (on the last two rows) and, in all but the special case, the Kelly Criterion not only fails to yield the optimal fraction (to be demonstrated later) but doesn’t even yield a fraction.

Let us assume now a three-scenario trading situation (where, for the sake of simplicity, a stock is priced at \$100 per share). Since the Kelly Criterion, (1,1a,1b|r=0), requires percentage returns as input and the more general Optimal *f* solution, (2), requires raw data points, we then have outcomes of ten, one, and minus five with corresponding probabilities of occurrence of .1, .6, and .3 respectively. We will designate these three outcomes as A, B and C, See Table 2.

Again, the values for *f* which maximize the objective functions given by the Kelly Criterion, (1,1a,1b|r=0) versus that given by the Optimal *f* solution, (2), are disparate indeed. Note the “fraction” of one’s stake to bet that maximizes the expected value of the logs of the returns, the Kelly Criterion, (1,1a,1b|r=0), is not a fraction as the loss diminishes.

The reconciliation of the two notions, in trading, can be found by determining the relative quantities one should assume.

The Optimal *f* solution is converted into a number of “units” to trade in by dividing the largest losing outcome, W,

by the optimal fraction (*f*) returned in (2), and taking this resulting quotient (herein as *f*\$) as the divisor of the total equity. The individual data points used in the Optimal *f* calculation, since it is based on the raw data points as opposed to returns (as in the Kelly Criterion solution) are based on the notion of a single, user-determined, consistently-sized “unit,” as is the largest losing data point, W.

For example, in the second row, the

(3)

$$f\$ = -W / f$$

row where the outcome of C is minus three (corresponding to the largest losing outcome, W) with a probability of .3, we find the optimal “fraction” as determined by Optimal *f*, (2), to be 0.213519068. From this, we can solve for (3):

$$f\$ = -W / f$$

$$f\$ = -3 / 0.2135191$$

$$f\$ = 3 / 0.2135191$$

$$f\$ = 14.0502653$$

Therefore, we should capitalize

each “unit” (be it one share or 100 shares or any other arbitrary but consistent, user-defined amount) by (3) in order to be at a “fraction” of our stake consistent with the *f* value used to calculate (3). In other words, when the worst-case loss manifests (outcome C in this example), where we have one unit (which experiences an outcome of minus three in this example) for every

14.0502653 in account equity, our loss will be that optimal fraction of our account, or 0.2135191:

$$3 / 14.0502653 = 0.2135191$$

The manifestation of the worst-case

outcome is equivalent to losing a fraction, *f* of our stake in the Optimal *f* calculation, (2). Thus, Optimal *f* provides us with the fraction of our stake at risk (provided we have adequately determined the worst-case scenario) and the corresponding quantity to put on to be consistent with that fraction at risk.

*Note: It is specifically because the Optimal *f* calculation incorporates worst-case outcomes that it is bounded between zero and one inclusively.*

The Kelly Criterion solution is clearly unbounded “to the right”. The disparate results given by the Kelly Criterion and Optimal *f* are reconciled through (3). If we take the price of the stock (S), or the wager (always unity, in gambling), and divide it by the quotient given in (3), we obtain the result given by the Kelly Criterion (1, 1a, 1b|r=0):

Formula (4) represents not an optimal

(4)

$$\text{Kelly Criterion Solution} = S / f\$$$

fraction to “bet” in trading, but rather a “leverage factor” to apply in trading. In other words, what we are referring to herein as the Kelly Criterion Solution is that value for *f* which maximizes (1,1a,1b|r=0). So for the three scenario example used, and for the case where outcome C = -3, we found our *f*\$, (3), to be 14.0502674. Therefore, for a stock priced at 100 (S):

This corresponds to the value that

$$\text{Kelly Criterion Solution} = S / f\$$$

$$\text{Kelly Criterion Solution} = 100 / 14.05026529$$

$$\text{Kelly Criterion Solution} = 7.1173$$

maximizes the Kelly Criterion for this row, the fraction that maximizes the

Table 2

			Optimal fractions given by:	
A p(.1)	B p(.6)	C p(.3)	(1) Kelly Criterion	(2) Optimal <i>f</i>
10	1	-5	0.5623922	0.0281196
10	1	-3	7.1173022	0.2135191
10	1	-1	48.053266	0.4805327
10	1	-0.1	674.28384	0.6742838
10	1	-0.01	6973.8987	0.6973899

expected value of the logs of the returns. Thus, the Kelly Criterion, except in the special case, does **not** yield an optimal fraction. It is shown to be mathematically related to the optimal fraction, the fraction at risk (by (3) and (4), converting Optimal f to the value returned by the Kelly Criterion), but it is **neither the optimal fraction nor** even a “fraction,” by definition.

Rather, the Kelly Criterion Solution, equivalent to the value for f which maximizes (1, 1a, 1b[r=0]), tells us how many shares to have on by virtue of the fact that it is a “leverage factor” (a.k.a. the misnomer “fraction” which satisfies the Kelly Criterion).

Kelly’s Oversight: *Arguably, even in the gambling situation (where W equals minus unity), the solution that satisfies the Kelly Criterion is **not** a fraction, appearances to the contrary, but is in fact a leverage factor and this becomes evident when we begin to move W (or, essentially in trading, $-S$) away from minus unity.*

Simply for any number, f to be zero $\leq f \leq$ one in certain instances does not make it a fraction when it is shown that number can at times exceed one. In all such cases, f is a leverage factor, including the case where zero $\leq f \leq$ one. The answer that satisfies the Kelly Criterion is not evidently what Kelly and others thought it to be, a fraction, but instead it is a leverage factor (e). *It is only in the special case that the leverage factor [as determined by the Kelly Criterion (1, 1a, 1b[r=0])] is the same value as the optimal fraction [as determined by the Optimal f calculation (2)].*

Returning to our three-scenario example, to assume a long position at \$100 per share, the Kelly Criterion Solution calls to (growth) optimally lever at 7.1173022 to one. At such a factor of leverage, when the largest losing scenario manifests (minus three per unit) the resultant loss will be 7.1173022 * -3 = 21.35191. Dividing this outcome by the \$100 per share gives us the resultant Optimal f value (2) to maximize this scenario set.

*If one wants to consider the value that satisfies the Kelly Criterion in terms of the optimal fraction to bet or risk in trading (i.e. converting the number that maximizes the expected value of the logs of the returns to a tradable quantity to assume), one is **de facto incorporating the largest losing outcome (f)** (and consequently, when one utilizes the Kelly Criterion in trading, the calculation becomes contingent on the underlying price).*

The incorporation (and necessity) of the biggest loss, W , (a data point with the worst outcome of all data points being employed) is not as problematic as the reader may be inclined to regard it.

Returning to the two to one coin toss example, an instance of the special case, the optimal fraction to risk regardless of calculation method is .25.

Since the largest loss is minus one, we have an $f\$$ given by (3) of \$4 (-1/.25 = 4), to make one bet for every \$4 in our stake. Now, if we arbitrarily say that our W parameter is -\$2 (leaving both scenarios the same, a loss of \$1 and a gain of \$2, but using a new W parameter of \$2 in (2)) we find that our optimal f value is now .5. Then we subsequently divide the absolute value of our largest loss by the optimal f value, and obtain an $f\$$ of $-2/.5 = 4$. Again, we trade one unit; make one bet, for every \$4 in our stake. The following table (table 3) demonstrates this for varying values of our biggest loss, W , wherein the optimal f for each row is determined using that row’s W in (2) in determining the optimal f at that row. See Table 3.

Notice that a different largest loss, though it unbounds the solution, does not result in a different optimal quantity

to assume ($f\$$). The incorporation of largest loss into the objective function for Optimal f , (2), serves solely to bound the solution for f between zero and one inclusively.

It would seem then that the Kelly Criterion and Optimal f can be used interchangeably, and, in theory, given the translations for both, they could be. Optimal f is easier to employ particularly when one considers quantities in short positions and pre-leveraged positions such as futures. Further, but most importantly, a bounded solution, such as what Optimal f provides directly, since zero \leq Optimal $f \leq$ one (as opposed to zero $\leq f$ value satisfying the Kelly Criterion $< \infty$), opens up a broad spectrum of possibilities.

Only in a gambling situation is the optimal fraction to wager equal to the leverage factor which satisfies the Kelly Criterion. In a trading situation, one must translate this back into the fraction dictated by Optimal f (unless it meets both criteria of the special case).

Most importantly, Optimal f is germane to the trading situation because it is bound between zero and one inclusively. Bounding permits us to:

- 1 Examine a bevy of geometrical relationships in context (g) and consider various points along the curve, giving these points context and meaning that an unbounded solution would not have (e.g. inflection points, f values as minimum expected drawdown, points x percent to the left and the right of the peak having the same return but different drawdowns, etc.). These points open up a legitimate study of the nature of the curve, the tenets of money management and position sizing.

Table 3

W	f	f\$	(2)
-0.6	0.15	4	1.125
-1	0.25	4	1.125
-2	0.5	4	1.125
-5	1.25	4	1.125
-29	7.25	4	1.125

2 Combine assets into a portfolio on an apples-to-apples basis, allowing such models as the Leverage Space Portfolio Model^{iii,iv} to permit us to:

3 Satisfy criteria other than mere geometric growth maximization via “Migration Paths” through this uniformly-bounded-for-all-components leverage space.

Relationship to Technical Analysis

Let us further consider point 1, specified earlier. There is a perceived point to the right of the peak where $G(f) < 1$. In our two to one coin toss example, the point where $G(f) < 1$ occurs at $f = .5$. This can be seen in Figure 1.

Here we see at $f = .5$ that point where $G(f)$, the average factor of growth per play on our stake, drops below 1.0. In other words, at each play, we expect to make $G(f) \cdot$ our current stake. If $G(f)$ therefore is less than one, we expect at such levels of quantity to be multiplying our stake by a value less than one. In such cases, we expect our stake to diminish with each play, and approach zero. We go broke at such levels.

Employing (3), we find that at $f = .5$:

$$f\$ = -.5/1.5 = 2$$

Thus, $f = .5$ corresponds to making a \$1 wager for every \$2 in our stake. We are not borrowing to assume these wagers, we have ample funds in our

stake to cover the wager (i.e. this is not a margin account, or leveraged in any manner). Note that even with an edge wildly in our favour as in this two to one coin toss, we can unwittingly bet in a manner aggressive enough to insure our demise as we continue to trade without being so aggressive that we must borrow.

Market analysis is a discipline that seeks to find the edge. Through the study of price, volume, and other data, we seek those circumstances that provide us an edge.

However, whenever we assume a position, whenever we take on a trade, we are ineluctably at some level for f , and are somewhere on the function $G(f)$ at a coordinate between $f =$ zero and one inclusively. We can therefore find advantageous trading situations via technical analysis but sabotage our efforts by misappropriating quantity whether we acknowledge it or not.

It is precisely these kinds of unforeseen pitfalls that make the study of market analysis - timing and selection - subordinate to this material(h).

Singularities and Discontinuities in Geometric Growth

As a discipline in its own right, the study of this material necessitates its known precepts be catalogued.

Alluding again to point 1 above, the bounded solution, (2), permits us to *examine a bevy of geometrical*

relationships in context and consider various points along the curve.

Here we will add to this sub-discipline with yet another phenomenon that comports with the differences between the Kelly Criterion and Optimal f .

A negative expectation set of data points has no optimal fraction to bet. If the expected value of the data points is negative, we assume $f =$ zero (i.e. do not wager anything so as to “maximize” growth).

Similarly, if all data points are positive (i.e. no losing data points) we have no possibility of loss at any play, and thus, in order to maximize growth, we wager 100% of our stake on each play ($f = 1.0$).

But a peculiar thing happens. We would expect that when we further diminish the loss in our two to one coin toss game, our value for Optimal f approaches 1.0. But this does **not** occur, as shown in Table 4.

Notice that instead of approaching 1.0 for the optimal fraction to wager, we approach .5

Let us look at the three-scenario situation mentioned earlier, in Table 5, wherein we will further diminish loss.

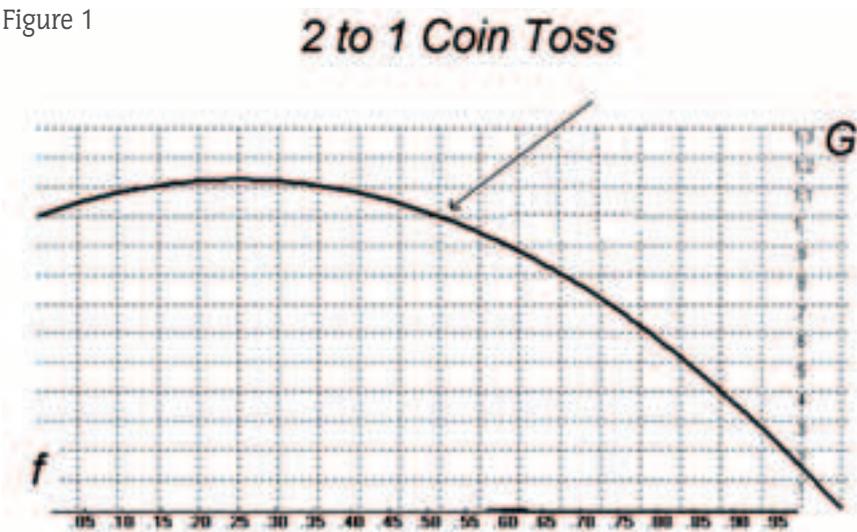
Yet again, we approach a singularity for the value for Optimal f , rather than approach 1.0.

Unequivocally, however, when there are no losses, growth is maximized by risking 100% of our stake ($f = 1.0$). Yet we find that as loss diminishes and approaches zero, the value for f approaches a **singularity**, and this singularity is less than 1.0. We see the value for f emerge again at 1.0 when all losses disappear, resulting in a **discontinuity**. Therefore, as loss approaches zero, the **optimal fraction** to wager approaches a singularity(i).

This seemingly unusual phenomenon is explained when we consider that Optimal f is bounded. If we convert to its unbounded analog, the Kelly Criterion solution (the “leverage factor” given by $(1, 1a, 1b[r=0])$ as f therein, to maximize $(1, 1a, 1b[r=0])$), it is clarified.

Equation (5) allows us to convert from the answer for the leverage factor given by the Kelly Criterion solution $(1, 1a, 1b[r=0])$, to the optimal fraction as determined by the Optimal f means, (2) as:

Figure 1



(5)

$$\text{Optimal } f = (\text{Kelly Criterion Solution} * -W) / S$$

Table 4

Heads p(.5)	Tails p(.5)	Optimal <i>f</i>
2	-1	0.25
2	-0.8	0.3
2	-0.5	0.375
2	-0.25	0.4375
2	-0.1	0.475
2	-0.001	0.49974999844375100000
2	-0.0001	0.49997499938974100000
2	-0.00001	0.49999749899514000000
2	-0.000001	0.49999974853450900000
2	-0.0000001	0.49999974853450900000
2	-0.00000001	0.49999974853450900000 <singularity>
<discontinuity>		
2	0	1.0

Table 5

A p(.1)	B p(.6)	C p(.3)	Optimal <i>f</i>
10	1	-5	0.0281196
10	1	-3	0.2135191
10	1	-1	0.4805327
10	1	-0.1	0.6742838
10	1	-0.01	0.6973899
10	1	-0.001	0.69973849290211600000
10	1	-0.0001	0.69997373914116300000
10	1	-0.00001	0.69999726369202300000
10	1	-0.000001	0.69999961737115700000
10	1	-0.0000001	0.69999985261339300000
10	1	-0.00000001	0.69999985261339300000
10	1	-0.000000001	0.69999985261339300000 <singularity>
<discontinuity>			
10	1	0	1.0

Because *f* is bounded to the left, at zero, by either the Kelly Criterion calculations or the Optimal *f* method, we find there is no singularity left of the peak, but only to the right, where the unbounding occurs.

The Kelly Criterion solution approaches infinity at a rate where *W* diminishes and *S* remains constant in (5), providing the Optimal *f* solution to approach a singular value.

The singularity makes sense when, for example, we consider the case in our two to one coin toss of 2, -0.00000001. At such small loss, our answer for (3) would be so high ($f = -0.00000001 / .49999974853450900000$ or make one bet for every .0000002000001005862 in our stake!) as to result in a percentage loss to our stake equal to the singularity itself (j). In other words, it is the Optimal *f*, as given by (2), that truly is the percentage, the fraction, of our stake at risk (i.e. betting one unit for every .0000002000001005862 in our stake results in a percentage loss of the singularity as a percent, or .49999974853450900000 of the stake, when the loss of , -0.00000001 manifests).

As it happens, the singularity in near-lossless Optimal *f* scenario sets occurs at $f = 1 -$ the probability of the losing scenario.

Conclusions

The above findings have important implications for a trader wishing to implement Optimal *f* in his future trading. One of the major impediments to implementing the usage of Optimal *f* for geometric growth in trading is the lack of knowledge as to where the optimal point will be in the future.

Since the Optimal *f* case will necessarily bound the future optimal point between zero and *p* (the sum of the probabilities of the winning scenarios), the trader need only perceive what *p* will be in the future. From there, trading a value for *f* of *p*/2 will minimize the cost of missing the peak of the Optimal *f* curve in the future.

This occurs because each point along the Optimal *f* curve varies with the increase in the number of plays (time), *T*, as G^T , where *G* is the geometric mean holding period multiple as given in Equation (2). Thus, at *T*=2, the price

paid for being at any future f value other than the optimal value is squared, at $T=3$, the penalty is cubed. Just as with the measure of statistical variance, outliers cost proportionally more. Although the trader cannot judge what will be the future value for Optimal f , by using the value of $p/2$ as the future estimate of the Optimal f , the trader minimizes this cost and is able to make a “best guess” estimate of what the future value for Optimal f will be.

Note that the trader uses a predicted value for p in determining his future “best guess” for f . The *greatest* amount the trader *might* miss actually is the optimal point in the future and is the greater of $p/2$ or what we call p' , which is what p actually comes in as in the future window, $p' - p/2$. These extreme cases manifest when the trader opts for $f = p/2$ and the future Optimal $f=0$, or, the trader opts for $f = p/2$ and the future Optimal $f=p'$. Thus, the greatest outlier, when the trader is opting to use a “best guess” for his future Optimal $f = p / 2$ is minimized as the greater of $p/2$ and $p'-p/2$.

Because the Kelly Criterion Solution is unbounded to the right, we are not afforded this outcome unless, we convert it to its Optimal f analog.

At no losses, the Kelly Criterion solution is infinitely high, and only by convention can we conclude that the corresponding Optimal f is 1.0. The point of singularity we witness in Optimal f is mathematical, the discontinuity, by convention. **IFTA**

Notes

- (a) Whether known by Kelly or not, the notion of a variable as the regulator which will maximize geometric growth was first introduced by Daniel Bernoulli in 1738^v. It is also likely that Bernoulli was not the originator of the idea, either. Bernoulli's 1738 paper was translated into English in 1954, two years before Kelly's paper. In fairness to Kelly, his paper was presented as a solution to a technological problem that did not exist in Daniel Bernoulli's day. Further in fairness to Kelly, he never presented his criterion as being optimal in a trading context. This fallacy has been perpetuated by others. Kelly discusses the gambling context, and hence the largest loss is always -1 , and hence the optimal value, f , is always a “fraction,” $0 \leq f \leq 1$. The differences, however subtle, between gambling and trading render the Kelly Criterion inapplicable in determining growth optimal quantities to risk in trading except in the special case.
- (b) Vince^{vi} and independently Thorp^{vii} provide a solution that satisfies the Kelly Criterion for the continuous finance case, often quoted in the financial community to the effect that “ f should equal the expected excess return of the strategy divided by the expected variance of the excess return.”

(1b)

$$f = (m-r) / s^2$$

where m =return (an expected value of return), r = the so-called risk-free rate, and s =the standard deviation in the expected excess returns comprising $(m-r)$. It should be noted that when $r=0$, all three forms for satisfying the Kelly Criterion, (1,1 a,1 b|r=0), will yield the same value for f .

- (c) Regardless of the means used to determine the optimal fraction, whether by the Kelly Criterion in the special case, or the Optimal f means in all cases, the optimal fraction returned is never really optimal as noted by Samuelson in 1971^{viii}. Rather, it is optimal in the long run sense, i.e. as the number of plays approach infinity; the optimal fraction approaches what we deem as this optimal fraction. For a single play, the expected growth is optimized for a positive expectancy game by betting 100% of the stake (optimal fraction =1.0). As the number of plays increase, the optimal fraction approaches that amount deemed the optimal fraction asymptotically, never really reaching the optimal fraction and thus the optimal fraction is actually always sub-optimal; the real optimal fraction will always be a more aggressive risk posture than that deemed as the optimal fraction.

- (d) Mathematical proof of Optimal f providing for geometric growth optimality.^{ix, x}
- (e) To see this, consider Table 1, row 3, where the player wins two or loses -.5 with probability .5 each. The optimal fraction to wager is .375, whereas the Kelly Criterion solution is .75. If the player uses .75 as a leverage factor, he will be growth optimal. However, if he uses .75 as the fraction of his stake to risk, he will be far too aggressive – well beyond that which is growth optimal, and will go broke with certainty as he continues to play.
- (f) Which is why the Kelly Criterion calculation of maximizing the expected values of the logs of the returns, (1, 1a,1 b|r=0) is applicable only when considering long positions. The largest loss is assumed to be the value of the position itself. To apply it equally to short positions, assumes that the worst-case outcome is a doubling of price. Thus in our three scenario example the Kelly Criterion makes the assumption that the worst that can happen is that the stock goes to 200 per share on our short position.
- (g) The same mathematical relations hold in an “unbounded to the right” situation such as that provided by the Kelly Criterion Solution, but context becomes ambiguous if not lost altogether, akin to a map without a distance scale. Each separate set of data points providing a curve between zero and some ambiguous point to the right. When we get into $N+1$ dimensional space, where N is the number of components considered in a portfolio, each component, thus each axis, has a different scale. Opting for a messy, nearly-untenable solution such as this wherein we opt for the Kelly Criterion as opposed to Optimal f gains us nothing; the Kelly Criterion, in real-world applicability to trading still utilizes the largest losing data point *de facto*. Nothing is gained by opting for the messier solution it entails over the Optimal f solution.
- (h) Particularly when the inputs to this discipline of position sizing and money management are exactly the very inputs used by the analyst; the data points used as inputs to (2), the “scenarios,” are essentially the distribution of price transformed by the analyst's trading rules.
- (i) This is a serendipitous phenomenon for the investor. Typically, one pays a steep price when one attempts to be at the growth optimal point in the future, and finds oneself having missed it as a result of market characteristics having changed when applying the optimal

allocation versus market characteristics from which the optimal allocation was derived. There is a small range of possible values for the future optimal point. Rather than being bound zero \leq Optimal $f \leq$ 1.0 it is rather bound between zero \leq Optimal $f \leq$ a singularity and that singularity $<$ 1.0.

- (j) The objective function solution to the Optimal f calculation provides not only the geometric growth multiple per play, but the value for f itself dictates the percentage loss on the stake when W manifests.

References

- i J L Kelly Jr, 'A new interpretation of information rate', *Bell System Technical Journal*, vol.35, 1956, pp.917-926.
- ii *ibid.*
- iii R Vince, *The New Money Management*, John Wiley & Sons, New York, 1995.
- iv R Vince, *The Leverage Space Trading Model*, John Wiley & Sons, New York, 2009.
- v D Bernoulli, 'Specimen Theoriae Novae de Mensura Sortis' (Exposition of a New Theory on the Measurement of Risk), *Commentarii academiae scientiarum imperialis Petropolitanae*, vol.5, 1738, pp.175-192, trans. L. Sommer, 1954.
- vi R Vince, *The Mathematics of Money Management*, John Wiley & Sons, New York 1992, pp.289.
- vii E O Thorp, *The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market*, presentation at the 10th International Conference on Gambling and Risk Taking, Montreal, June 1997.
- viii P A Samuelson, 'The "Fallacy" of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling', *Proceedings of the National Academy of Sciences of the United States of America*, vol.68, 1971, pp.2493-2496.
- ix R Vince, *The New Money Management*, John Wiley & Sons, New York, 1995.
- x Vince, 2009, loc.cit.

SAVE THE DATE



24th Annual IFTA Conference October 2011 ■ Sarajevo, Bosnia and Herzegovina

—Hosted by the Society for Market Studies <http://trzisnestudije.org>

 SOCIETY FOR MARKET STUDIES

Check the web site
for announcements.
www.ifta.org


International Federation of Technical Analysts