# Optimal $f$:Calculating the expected growth-optimal fraction for discretely-distributed outcomes

Ralph Vince*

March 12, 2018

First version: January, 2012

*The president of Vince Strategies, LLC, The Chrysler Building 26th Fl., 405 Lexington Ave., New York, NY 10174, *Email: ralph@ralphvince.com*

# Optimal $f$: Calculating the expected growth-optimal fraction for discretely-distributed outcomes

**Abstract**

Presented is the formulation for determining the exact, expected growth-optimal fraction of equity to risk for all conditions, rather than merely the asymptotic growth-optimal fraction. The formulation presented represents the surface in the leverage space manifold, wherein the loci at the peak of the surface are those fractions for maximizing expected growth. Other criteria can be solved for upon the surface in the leverage space manifold utilizing the equation specified here.

**Keywords.** Growth-optimal portfolio, risk management, Kelly Criterion, finite investment horizon, drawdown.

**AMS Classifications.** 91G10, 91G60, 91G70, 62C, 60E.

## 1 Discussion

Repeatedly in both the gambling and trading communities the expected growth optimal fraction is employed, often referred to as the Kelly Criterion, and referring to the 1956 paper of John L. Kelly[1] [2]. Though applicable in many gaming situations (though not all, blackjack being a prime example [12]) where the most that can be lost is the amount wagered, the Kelly Criterion solution is not directly applicable to trading situations which are more complicated.

Take for example a short sale. Clearly what can be lost is not the same as the price of the stock. Forex transactions incur a different analysis as one is long a currency specified in another currency (and thus vice-versa in effect, short that other currency against the long currency). Commodity futures represent a different problem than simply being long an equity in terms of a floor price (what, specifically, is the lowest, say, wheat can go? Clearly it has value and thus a price of zero may not be realistic, the risk on a long wheat futures transaction therefore actually less than the immediately priced contract amount), and they too are represented in a base currency. CD swaps, complicated strategies involving options, warrants, leaps, trades in volatility where the limiting function of zero as a price and maximum loss is distorted, interest rate products and derivatives all create a situation where the Kelly Criterion solution does not result in the expected growth-optimal fraction to risk, and, in fact, to do so is often to risk more than the actual expected growth-optimal fraction would call for[10]. Capital market situations. given their inherent complexity, are not the same as gambling situations and the Kelly Criterion cannot be applied directly in determining proper expected growth-optimal fractions.

Further, as pointed out by Samuelson[4] (though he does not provide a computational solution), the Kelly Criterion solution (in those cases where it is applicable, i.e. wagering situations where what can be lost is equivalent to what is put up) seeks the *asymptotic* expected growth-optimal fraction as the number of trials approaches infinity. For example, in a single proposition where the probability-weighted expected outcome is greater than 0, the expected growth for a participant whose horizon is one trial is $f = 1.0$, or to risk his entire equity. If his horizon were longer (but *necessarily finite*) the expected growth optimal fraction would be greater than the Kelly Criterion solution but less than one. The handicapper who goes to the track for a ten-race card day, seeking to maximize his return for the day employs a horizon, $Q$, of ten. The portfolio manager, depending on his criteria, has a similarly finite number of periods. The Kelly Criterion solution (when applicable) is

[1] The Kelly Criterion was merely ***posutlated*** by Kelly *et al* in [2] (so as to determine the *fraction* which *asymptotically* maximizes expected growth), but not solved for in [2] (and thus only postulated). The solution presented therein to satisfying this criterion posited in [2] is not a fraction but rather a stake *multiplier* (a number which *can* be greater than 1). This multiplier *equals* the expected growth-optimal fraction when losses equal the amount which is put up by the entity assuming the risk. This does not make it a fraction, only equal to the expected growth-optimal fraction in this *special case*, demonstrated in [9] and [10]. The first solution to determining the expected growth-optimal fraction (for binomially distributed data, asypmtotically) appears in [5], and for a more general solution, asymptotically, in [6].

only an asymptote; it is never the expected growth-optimal fraction, but rather approached asymptotically as the number of trials approaches infinity. The Kelly Criterion solution, as well as closed-form formulations that seek to solve it, are asymptotic and always less than what is the *actual*, expected-growth optimal fraction.

As Hirashita [1] points out (referring to the asymptotic case), unless the payoff occurs immediately, the cost of assuming the risk is germane to the calculation of the expected growth-optimal fraction. Since we are discussing the finite case here, with a specific horizon, we must include the cost of the wager to the specific horizon as is expressed in our equation. Finally, the actual expected growth-optimal fraction formula should incorporate multiple, simultaneous propositions. The Kelly Criterion solution is therefore merely a subset of this more generalized formula, and represents only the asymptote to the special case where it is applicable.

Presented here is the formulation for $N$ multiple, simultaneous propositions at a horizon of $Q$ trials, for all potential propositions. The solution is presented in the context of the $N + 1$ dimensional leverage space manifold where, for each of the $N$ components, a surface bound between $f\{0, 1\}$ along each axis represents the fraction allocated to that individual proposition.

$$G(f_1, ...f_N)^Q =$$

$$1 + (\sum_{x=0}^{n^Q-1} (((\prod_{q=1}^{Q} (1 + (\sum_{i=1}^{N} f_i(-(a_{i,k} - C_i)/w_i)))) - 1) \prod_{q=1}^{Q} p_k)) \tag{1.1}$$

where:
$Q$ = the horizon.
$N$ = the number of components in the portfolio.
$n$ = the number of possible outcomes in the outcomes copula.
$k = int(x/(n^{(Q-q)}))\%n$.
$a_{N,n}$ = the $N$ $x$ $n$ matrix of outcomes in the outcomes copula.
$p_n$ = the $n$-lengthed vector of probabilities associated with each $n$ in the outcomes copula.
$w_N$ = the worst case outcome of the all discrete outcomes associated with each of the $N$ components.
$C_i$ = The one period opportunity-cost of the risk in component $i$, per [10]. This is explained further, below.

Note $k$ is a function of the iterators $x$ and $q$, returning a specific zero'th-based row in the outcomes copula.

The drawdown constraint or other risk constraint as proposed in [7, 8] for a given $Q$ can be employed upon the surface mapped by this equation in the leverage space manifold.

To compute the amount to allocate to replicate the percentage to allocation to a given individual proposition, $i$ for a given $f_i$:

$$f\$_i = w_i / f_i \tag{1.2}$$

Then for a given *equity* amount, the number of units to assume of component $i$ to represent wagering $f_i$ percent:

$$NumberOfUnits_i = equity / f\$_i \tag{1.3}$$

The one period opportunity-cost of the risk in component $i$, used in (1.1) as $C_i$, is given in [10]:

$$C_i = exp(rt)S_i - S_i - d_i \tag{1.4}$$

where:
$r$ = the current risk-free rate.
$t$ = the percentage of a year for one period to transpire .
$d_i$ = dividends, disbursements or costs (negative) assosiated over one period with one unit of the $i$th component .
$S_i$ = The maximum of $\{f\$_i$, regulatory (i.e. *margin*) requirement of the $i$th component$\}$ where:
$f\$_i$ = the amount to allocate to replicate the percentage to allocation to a given individual proposition, $i$ for a given $f_i$ as given in (1.2)

We examine how to collect the data for determining the surface in leverage space at a given $Q$. Assume we have two separate propositions we wish to engage in simultaneously. For the sake of simplicity, suppose one is a coin toss paying 2:1, where we win with a probability of .5, and the other, a flawed coin paying 1:1 where we win with a probability of .6.

We create the copula to use as input from this data:

| Coin 1 ($a_1$) | Coin 2 ($a_2$) | Probability (p) |
|---|---|---|
| 2 | 1 | .3 |
| 2 | -1 | .2 |
| -1 | 1 | .3 |
| -1 | -1 | .2 |

We note here that the number of components, $N$, is two since we have two separate, simultaneous propositions. The number of rows, $n$ representing the space of what can happen for each discrete interval, is 4. $w$, the worst-case outcome for each $N$, each column, is $-1$ for both Coin 1 and Coin 2.

To see the equation graphically, consider that for the first interval, the possible outcomes are represented by the table. For the second interval, each node from the first interval now further branches by the number of rows in the table such that at a given period, $q$ toward a horizon, $Q$, we have $n^q$ nodes. The formula can thus account for dependency by permitting the probabilities for the various rows in the copula to change at each subsequent interval based on the previous outcome(s) along a given branch being traversed.

The equation represents the surface in the leverage space manifold for given $Q, f_1, ...f_N$. This surface represents what one would expect in terms of return, as a multiple on equity, after $Q$ trials (hence, $G(f_1, ...f_N)$ represents what one would expect in terms of a multiple on equity, on average, per trial). The maximum of this surface (*i.e.* the greatest $G(f_1, ...f_N)^Q$ or $G(f_1, ...f_N)$), provides the expected growth-optimal fractions. The equation represents the *actual* (*i.e.* non-asymptotic) fractions to risk of $N$ components ($N >= 1$) for all possible propositions; thus, all other expected growth-optimal solutions tend to be subsets of the asymptote (*i.e.* $Q \to \infty$) of this more generalized equation, which is expressed here as

$$\lim_{Q\to+\infty} G(f_1, ...f_N)^Q = (\prod_{j=1}^{n}(1 + (\sum_{i=1}^{N} f_i(-(a_{i,j} - C_i)/w_i)))^{p_j})^Q \tag{1.5}$$

Equation (1.5) represents the asymptotic manifestation of equation (1.1). The proof of this is found in [11] and [3]. Since (1.1) yields the surface in the leverge space manifold after $Q$ trials, (1.5) represents what this surface tends to asymptotically as $Q \to +\infty$. Since (1.5) is far less computationally expensive than (1.1) we can use (1.5) as a reasonable proxy of (1.1) after even a relatively small $Q$, thus for many calculations, including the expected risk-adjusted maximizing loci on this surface, $\nu$ and $\zeta$ as proposed in [3, 11, 12] as well as $\psi$ proposed in [3] can be reasonably determined from (1.5).

Acceptance of the underlying proposition in [10], that what the sole player *expects* to make over $Q$ discrete trials or periods is *not* the classical expectation, defined as the probability-weighted mean, but rather the median of the ranked cumulative outcomes made along each branch at $Q$ discrete trials or periods, we see a different approach must be undertaken to represent the expectation of the individual over $Q$ discrete trials or periods. Rather what we have thus far considered with Equation (1.1) represents such a campaign performed many times over by the same individual, or, by many individuals over $Q$ discrete trials or periods each. It does *not* represent what the individual, a sole individual, would *expect* to occur, for his solitary campaign across $Q$ discrete trials or periods. *Ergodicity diminishes as the number of discrete trials or periods approach 1 from the right.* I take this to be innately evident, the mathematical proof left for those who are more mathematically adroit than I am.

To reject the underlying proposition in [10], is to accept the following. Consider a game with a million-to-one chance of losing one million units, and wins one unit the other 999,999 times. Such a game has a negative expectation in the classic sense, but the individual looking to make one play then leave, would erroneously accept such as his expectation and decline to participate in such a proposition.[2] Equation (1.1) in its raw form suggests an $f = 0$, that is, the player should wager nothing in the classical sense. Clearly though, the median of the sorted outcomes suggest that the individual expects to make one unit profit from such an attempt, hence, his expected growth-optimal wager is to risk 100%, *i.e.* $f = 1$.

Similarly, rejection of this proposition makes the assumption that a game which pays $1 : -1,000,000$ with corresponding probabilities of $.999999 : .000001$ is the equivalent of a proposition which pays $1 : -1$ with corresponding probabilities of $0.4999995 : .0.5000005$ as both have classical (asymptotic) probability-weighted means (mathematical expectations) of -0.000001, yet they are clearly not the same proposition to someone with a finite horizon of participation. This difference becomes evermore apparent the shorter the horizon.

Here we see that Equation (1.1) does, in fact, yield the correct altitude in the leverage space manifold for given $f$ coordinates at $Q$ trials, but what must be amended is the copula used as input to achieve these corrected values. This will be demonstrated by way of a step-by-step example. Consider a proposition as depicted in Figure 1 of three possible outcomes and their associated probabilities. Here, we have a classical coin toss example, wherein the player wins 2 units on heads, loses 1 unit on tails, and in the unlikely (1 in 1,000 chance) of the coin stopping on its side, costs the player 1,000,000 units.

The classical, probability-weighted mean outcome calls for a loss of 99.50005 units per play, and hence the individual assessing this proposition should reject participating.

However, in this example, we are considering participating only for 2 plays, $Q = 2$, and we arrive at a dramatically different conclusion.

In Figure 2 we see the possible outcomes for two independent trials of this proposition, and their cumulative outcomes.

Figure 3 is the same as Figure 2 but with the associated probabilities of the outcomes in Figure 2. Here, the *probability product* is the product of all probabilities along a given branch.

In Figure 4 we see the summary of the cumulative outcomes and corresponding probability product at the terminus of each branch.

In Figure 5 we sort the values of the outcomes and probabilities at the terminus ($Q = 2$) of each branch. From here, we can determine the median outcome, representing the

[2]Similarly, this criterion, that of using the classical mathematical expectation for acceptance or rejection of a given proposition, would have the individual accept the reverse of the given proposition. That is, the individual acting on this criterion would accept a one-in-a-million chance of winning one million with a corresponding 999,999 chances in one million of losing one unit. The reader may argue this is precisely what individuals do when they purchase a lottery ticket, but that is an altogether different reasoning process. There, one knows the classical expectation is negative when they purchase a lotter ticket. In this sense, we are speaking of a positive expectation lottery purchase, where one continues buying tickets in an attempt to make the positive expectation ultimately manifest. The argument here, is that such manifestation occurs at an unrealistically far horion for the player.

| | Copula: |
|---|---|
| Outcome | Probability |
| 2 | 0.4999500 |
| -1 | 0.4999500 |
| -1000000 | 0.0001000 |
| | |
| Expectation= | -99.5000500 |

Figure 1: 2:1 coin toss proposition with the possibility of landing on the coins side.

| Q = | 1 | 2 | | Cumulative Outcome |
|---|---|---|---|---|
| | | | | |
| | | 2 | | 4 |
| | 2 | -1 | | 1 |
| | | -1000000 | | -999998 |
| | | 2 | | 1 |
| | -1 | -1 | | -2 |
| | | -1000000 | | -1000001 |
| | | 2 | | -999998 |
| | -1000000 | -1 | | -1000001 |
| | | -1000000 | | -2000000 |

Figure 2: Possible outcomes for two independent trials and their cumulative outcomes

| Q = | 1 | 2 | | Probability Product |
|---|---|---|---|---|
| | | 0.49995 | | 0.2499500025 |
| | 0.49995 | 0.49995 | | 0.2499500025 |
| | | 0.0001 | | 0.0000499950 |
| | | 0.49995 | | 0.2499500025 |
| | 0.49995 | 0.49995 | | 0.2499500025 |
| | | 0.0001 | | 0.0000499950 |
| | | 0.49995 | | 0.0000499950 |
| | 0.0001 | 0.49995 | | 0.0000499950 |
| | | 0.0001 | | 0.0000000100 |

Figure 3: Associated probabilities of the outcomes in Figure 2. The probability product is the product of all probabilities along a given branch.

| Cumulative Outcome | Probability Product |
|---|---|
| 4 | 0.2499500025 |
| 1 | 0.2499500025 |
| -999998 | 0.0000499950 |
| 1 | 0.2499500025 |
| -2 | 0.2499500025 |
| -1000001 | 0.0000499950 |
| -999998 | 0.0000499950 |
| -1000001 | 0.0000499950 |
| -2000000 | 0.0000000100 |

Figure 4: Cumulative outcomes and probabilities at the end of each branch

| Cumulative Outcome | Probability Product | Probability Sum | |
|---|---|---|---|
| 4 | 0.2499500025 | 1.0000000000 | |
| 1 | 0.2499500025 | 0.7500499975 | |
| 1 | 0.2499500025 | 0.5000999950 | <---Approximate Median {-1,2} |
| -2 | 0.2499500025 | 0.2501499925 | |
| -999998 | 0.0000499950 | 0.0001999900 | |
| -999998 | 0.0000499950 | 0.0001499950 | |
| -1000001 | 0.0000499950 | 0.0001000000 | |
| -1000001 | 0.0000499950 | 0.0000500050 | |
| -2000000 | 0.0000000100 | 0.0000000100 | |

Figure 5: Outcomes sorted with their probabilities and median (the individual's expectation) determined.

expectation the solitary individual *expects* as outcome at the end of $Q$ trials. We find this value to be a cumulative outcome of 1, characterized by the branch $2, -1($ or $-1, 2)$. We select these two rows from the initial copula, and we amend the probabilities of each row to be $1/Q = 1/2 = .5$ :

| Outcome | Probability |
|---|---|
| 2 | .5 |
| -1 | .5 |

and use this modified copula to determine our altitude in the leverage space manifold (and hence the expected growth-optimal peak) as specified by Equation (1.1).[3]

Astonishingly, we find that when we consider the effects of diminishing ergodicity with respect to ever-shorter horizon, for the individual player and his solitary path of outcomes, the shape of leverage space is altered profoundly. Whereas in this example, not only does the classical mathematical expectation indicate a negative proposition that we should therefore not engage in, the Kelly Criterion buttresses this conclusion[4] by its satisfaction

[3] When multiple, simultaneous propositions are involved, the process for finding the altitude in leverage space is the same. We amend the input copula by eliminating those rows whose set of outcomes do not lie on the median-sorted-outcome path, and we assign a probability of $1/Q$ to each of these remaining rows, and performing Equation (1.1) as specified.

[4] In fact, the Kelly Criterion, when presented with a one-in-ten chance of winning ten units, and a nine-in-ten chance of losing one unit, would specify that .05 as the expected growth optimal wager, *regardless* of horizon. If one were to play this game, say, one time, the correct expected growth optimal fraction would be to wager nothing as the correct expected growth-optimal wager is a function of horizon always. The Kelly Criterion bettor, to a horizon of one play, should expect to lose .05 of his stake. The characteristic becomes more obvious as the distribution becomes more pronounced. Consider again our one-in-a- million chance of winning one million units, and all other plays losing one unit. The Kelly Criterion solution would

with an $f = 0$, *i.e.* to risk nothing on this proposition. Yet, we see that for the solitary player, participating to a horizon of $Q = 2$. he not only expects to profit, but he maximizes his expected profit in this particular proposition (via Equation 1.1, with inputs of 2, -1 and corresponding probabilities of $1/Q, 1/Q$) at $f = .5$.

Since all campaigns are ultimately finite, this is the correct calculation for expected optimal growth as well as, more generally, the defined shape of leverage space and any and all other points germane to the individual (*e.g.* $\nu, \zeta, \psi$) within that manifold.

This means of assessing the expectation for the individual, experiencing a necessarily finite horizon, as pointed out in [10], is how human beings innately assess a proposition, often assuming what seem to be risky ventures aligned with short-horizon bounty. It is why a commuter will drive to his work as well as why early hominids descended to the savanna despite the risks. It is this innate mechanism of risk-assessment, contrary to our classical understanding of it, that belies our flourishing as a species, absent which, stunted, we would be cowering agoraphobically in the shadows of a primeval world.

## References


[1] Y. Hirashita. Game pricing and double sequence of random variables. *Journal of Modern Mathematics Frontier*, 2:33–55, 2013.

[2] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926, 1956.

[3] Marcos Lopez de Prado, Ralph Vince, and Qiji Jim Zhu. Optimal risk budgeting under a finite investment horizon. *SSRN*, 2364092, 2013.

[4] Paul 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*, 68:2493–2496, 1971.

[5] Edward O. Thorp and Sheen T. Kassouf. *Beat the Market*. Random House, New York, 1967.

[6] R. Vince. *Portfolio Management Formulas*. John Wiley and Sons, New York, 1990.

[7] R. Vince. *The Handbook of Portfolio Mathematicsl*. John Wiley and Sons, Hoboken, NJ, 2007.

[8] R. Vince. *The Leverage Space Trading Model*. John Wiley and Sons, Hoboken, NJ, 2009.


require the player to risk one two-milliointh of his stake per play, regardless of horizon. However, for a finite horizon, to a horizon whose length is less than the critical length given in [10], where the positive expectation can be considered to manifest, the expected growth optimal amount to wager is zero, not the positive amount as specified by the asymptotically-dependant, *ergodicity-dependant* Kelly Criterion.

[9] R. Vince. Optimal f and the kelly criterion. *Journal of the International Federation of Technical Analysts*, 11:21–28, 2011.

[10] R. Vince. *Risk-Opportunity Analysis*. Createspace Division of Amazon, 2012.

[11] R. Vince and Q. J. Zhu. Inflection point significance for the investment size. *SSRN*, 2230874, 2013.

[12] R. Vince and Q. J. Zhu. Optimal betting sizes for the game of blackjack. *SSRN*, 2324852, 2013.