generalization of this formula to the situation where a unit wager wins b with probability $p > 0$ and loses a with probability q. Then if the expectation $m \equiv bp - aq > 0$, $f^* > 0$ and $f^* = m/ab$. The generalization does stand up to the objection. One can buy on credit in the financial markets and lose much more than the amount bet. Consider buying commodity futures or selling short a security (where the loss is potentially unlimited). See, e.g., Thorp and Kassouf (1967) for an account of the E.L. Bruce short squeeze.

For purists who insist that these payoffs are not binary, consider selling short a binary digital option. These options are described in Browne (1996).

A criticism sometimes applied to the Kelly strategy is that capital is not, in fact, infinitely divisible. In the real world, bets are multiples of a minimum unit, such as \$1 or \$.01 (penny "slots"). In the securities markets, with computerized records, the minimum unit can be as small as desired. With a minimum allowed bet, "ruin" in the standard sense is always possible. It is not difficult to show, however, (see Thorp and Walden, 1966b) that if the minimum bet allowed is small relative to the gambler's initial capital, then the probability of ruin in the standard sense is "negligible" and also that the theory herein described is a useful approximation. This section follows Rotando and Thorp (1992).

3 Optimal growth: Kelly criterion formulas for practitioners

Since the Kelly criterion asymptotically maximizes the expected growth rate of wealth, it is often called the optimal growth strategy. It is interesting to compare it with the other fixed fraction strategies. I will present some results that I have found useful in practice. My object is to do so in a way that is simple and easily understood. These results have come mostly from sitting and thinking about "interesting questions". I have not made a thorough literature search but I know that some of these results have been previously published and in greater mathematical generality. See e.g. Browne (1995, 1996) and the references therein.

(a) The probability of reaching a fixed goal on or before n trials. We first assume coin tossing. We begin by noting a related result for standard Brownian motion. Howard Tucker showed me this in 1974 and it is probably the most useful single fact I know for dealing with diverse problems in gambling and in the theory of financial derivatives.

For standard Brownian motion $X(t)$, we have

$$P(\sup[X(t)-(at+b)] \geq 0,\ 0 \leq t \leq T) = N(-\alpha-\beta)+e^{-2ab}N(\alpha-\beta) \tag{3.1}$$

where $\alpha = a\sqrt{T}$ and $\beta = b/\sqrt{T}$. See Figure 2. See Appendix 2 for Tucker's derivation of (3.1).

In our application $a < 0$, $b > 0$ so we expect $\lim_{T\to\infty} P(X(t) \geq at+b,\ 0 \leq t \leq T) = 1$.

Let f be the fraction bet. Assume independent identically distributed (i.d.d.) trials Y_i, $i = 1, ..., n$, with $P(Y_i = 1) = p > 1/2$, $P(Y_i = -1) = q < 1/2$; also assume $p < 1$ to avoid the trivial case $p = 1$.

Bet a fixed fraction f, $0 < f < 1$, at each trial. Let V_k be the value of the gambler or investor's bankroll after k trials, with initial value V_0. Choose initial stake $V_0 = 1$ (without loss of generality); number of trials n; goal $C > 1$.

What is the probability that $V_k \geq C$ for some k, $1 \leq k \leq n$? This is the same as the probability that $\log V_k \geq \log C$ for some k, $1 \leq k \leq n$. Letting $ln = log_e$ we have:

$$\begin{aligned}
V_k &= \prod_{i-1}^{k}(1+Y_i f) \quad \text{and} \\
\ln V_k &= \sum_{i=1}^{k}\ln(1+Y_i f) \\
E\ln V_k &= \sum_{i=1}^{k}E\ln(1+Y_i f) \\
\mathrm{Var}(\ln V_k) &= \sum_{i=1}^{k}\mathrm{Var}(\ln(1+Y_i f)) \\
E\ln(1+Y_i f) &= p\ln(1+f)+q\ln(1-f) \equiv m \equiv g(f) \\
\mathrm{Var}[\ln(1+Y_i f)] &= p[\ln(1+f)]^2+q[\ln(1-f)]^2-m^2 \\
&= (p-p^2)[\ln(1+f)]^2+(q-q^2)[\ln(1-f]^2-2pq\ln(1+f)\ln(1-f) \\
&= pq\{[\ln(1+f)]^2-2\ln(1+f)\ln(1-f)+[\ln(1-f)]^2\} \\
&= pq\{\ln[(1+f)/(1-f)]\}^2 \equiv s^2
\end{aligned}$$

Drift in n trials: mn

Variance in n trials: $s^2 n$

$\ln V_k \geq \ln C,\ 1 \leq k \leq n$ iff

$\sum_{i=1}^{k} \ln(1 + Y_i f) \geq \ln C,\ 1 \leq k \leq n$ iff

$S_k \equiv \sum_{i=1}^{k} [\ln(1 + Y_i f) - m] \geq \ln C - mk,\ 1 \leq k \leq n$

$E(S_k) = 0 \quad \text{Var}(S_k) = s^2 k$

We want $\text{Prob}(S_k \geq \ln C - mk,\ 1 \leq k \leq n)$.

Now we use our Brownian motion formula to approximate S_n by $\text{Prob}(X(t) \geq \ln C - mt/s^2,\ 1 \leq t \leq s^2 n)$ where each term of S_n is approximated by an $X(t)$, drift 0 and variance s^2 $(0 \leq t \leq s^2, s^2 \leq t \leq 2s^2, \ldots, (n-1)s^2 \leq t \leq ns^2)$. Note: the approximation is only "good" for "large" n.

Then in the original formula (3.1):

$$
\begin{aligned}
T &= s^2 n \\
b &= \ln C \\
a &= -m/s^2 \\
\alpha &= a\sqrt{T} = -m\sqrt{n}/s \\
\beta &= b/\sqrt{T} = \ln C / s\sqrt{n}
\end{aligned}
$$

Example 3.1

$$
\begin{aligned}
C &= 2 \\
n &= 10^4 \\
p &= .51 \\
q &= .49 \\
f &= .0117 \\
m &= .000165561 \\
s^2 &= .000136848 \\
\text{Then} \quad P(\cdot) &= .9142
\end{aligned}
$$

Example 3.2

Repeat with

$$
\begin{aligned}
f &= .02 \\
\text{then} \quad m &= .000200013 \\
s^2 &= .000399947 \\
\text{and} \quad P(\cdot) &= .9214
\end{aligned}
$$

(b) **The probability of ever being reduced to a fraction x of this initial bankroll.**

This is a question that is of great concern to gamblers and investors. It is readily answered, approximately, by our previous methods.

Using the notation of the previous section, we want $P(V_k \leq x$ for some k, $1 \leq k \leq \infty)$. Similar methods yield the (much simpler) continuous approximation formula:

$\text{Prob}(\cdot) = e^{2ab}$ where $a = -m/s^2$ and $b = -\ln x$ which can be rewritten as

(3.2) $\text{Prob}(\cdot) = x\wedge(2m/s^2)$ where $\wedge$ means exponentiation.

Example 3.3.

$p = .51 \quad f = f^* = .02$

$2m/s^2 = 1.0002$

$\text{Prob}(\cdot) \doteq x$

We will see in section 7 that for the limiting continuous approximation and the Kelly optimal fraction f^*, $P(V_k(f^*) \leq x$ for some $k \geq 1) = x$.

My experience has been that most cautious gamblers or investors who use Kelly find the frequency of substantial bankroll reduction to be uncomfortably large. We can see why now. To reduce this, they tend to prefer somewhat less than the full betting fraction f^*. This also offers a margin of safety in case the betting situations are less favorable than believed. The penalty in reduced growth rate is not severe for moderate underbetting. We discuss this further in section 7.

(c) **The probability of being at or above a specified value at the end of a specified number of trials.**

Dr. Richard Hecht (1995) suggested setting this probability as the goal and used a computerized search method to determine optimal (by this criterion) fixed fractions for $p - q = .02$ and various c, n and specified success probabilities.

This is a much easier problem than the similar sounding (a). We have for the probability that $X(T)$ at the end exceeds the goal:

$$\begin{aligned} P(X(T) \geq aT + b) &= \frac{1}{\sqrt{2\pi T}} \int_{aT+b}^{\infty} exp\{-x^2/2T\}dx \\ &= \frac{1}{\sqrt{2\pi T}} \int_{aT^{1/2}+bT^{-1/2}}^{\infty} exp\{-u^2/2\}du \end{aligned}$$

where $u = x/\sqrt{T}$ so $x = aT + b$ gives $u\sqrt{T} = aT + b$ and $U = aT^{1/2} + bT^{-1/2}$. The integral equals

$$(3.3) \qquad \begin{aligned} 1 - N(aT^{1/2} + bT^{-1/2}) &= N(-(aT^{1/2} + bT^{-1/2})) \\ &= 1 - N(\propto + \beta) = N(-\propto - \beta). \end{aligned}$$

For example (3.1) $f = .0117$ and $P = .7947$. For example (3.2) $P = .7433$. Example (3.1) is for the Hecht optimal fraction and example (3.2) is for the Kelly optimal fraction. Note the difference in P values.

Our numerical results are consistent with Hecht's simulations in the instances we have checked.

Browne (1996) has given an elegant continuous approximation solution to the problem: What is the strategy which maximizes the probability of reaching a fixed goal C on or before a specified time n and what is the corresponding probability of success? Note that the optimal strategy will in general involve betting varying fractions, depending on the time remaining and the distance to the goal.

As an extreme example, just to make the point, suppose $n = 1$ and $C = 2$. If $X_0 < 1$ then no strategy works and the probability of success is 0. But if $1 \le X_0 < 2$ one should bet at least $2 - X_0$, thus any fraction $f \ge (2 - X_0)/X_0$, for a success probability of p. Another extreme example: $n = 10$, $C = 2^{10} = 1024$, $X_0 = 1$. Then the only strategy which can succeed is to bet $f = 1$ on every trial. The probability of success is p^{10} for this strategy and 0 for all others (if $p < 1$), including Kelly.

(d) **Continuous approximation of expected time to reach a goal.**

According to Theorem 1(v), the optimal growth strategy asymptotically minimizes the expected time to reach a goal. Here is what this means. Suppose for goal C that $m(C)$ is the greatest lower bound over all strategies for the expected time to reach C. Suppose $t^*(C)$ is the expected time using the Kelly strategy. Then $\lim_{C \to \infty} (t^*(c)/m(c)) = 1$.

The continuous approximation to the expected number of trials to reach the goal $C > X_0 = 1$ is

$$(3.4) \qquad n(C, f) = (\ln C)/g(f)$$

where f is any fixed fraction strategy. Appendix III has the derivation. Now $g(f)$ has a unique maximum at $g(f^*)$ so $n(C, f)$ has a unique minimum at $f = f^*$. Moreover, we can see how much longer it takes, on average, to reach C if one deviates from f^*.

(e) **Comparing fixed fraction strategies: the probability that one strategy leads another after n trials.**

Theorem 1(iv) says that wealth using the Kelly strategy will tend, in the long run, to an infinitely large multiple of wealth using any "essentially different" strategy. It can be shown that any fixed $f \neq f^*$ is an "essentially different" strategy. This leads to the question of how fast the Kelly strategy gets ahead of another fixed fraction strategy, and more generally, how fast one fixed fraction strategy gets ahead of (or behind) another.

If W_n is the number of wins in n trials and $n - W_n$ is the number of losses,

$$G(f) = (W_n/n)\ln(l + f) + (1 - W_n/n)\ln(1 - f)$$

is the actual (random variable) growth coefficient.

As we saw, its expectation is

$$g(f) = E(G(f)) = p\log(1 + f) + q\log(1 - f) \tag{3.5}$$

and the variance of $G(f)$ is

$$\mathrm{Var}G(f) = ((pq)/n)\{\ln((1 + f)/(1 - f))\}^2 \tag{3.6}$$

and it follows that $G(f)$, which has the form $G(f) = a(\sum T_k)/n + b$, is approximately normally distributed with mean $g(f)$ and variance $\mathrm{Var}G(f)$. This enables us to give the distribution of X_n and once again answer the question in 3(c). We illustrate this with an example.

Example 3.3 $\quad p = .51 \quad q = .49 \quad f^* = .02 \quad N = 10,000$ and

s = standard deviation of $G(f)$

g/s	f	g	s	$Pr(G(f) \leq 0)$
1.5	.01	.000150004	.0001	.067
1.0	.02	.000200013	.0002	.159
0.5	.03	.000149977	.0003	.309

Continuing, we find the distribution of $G(f_2) - G(f_1)$. We consider two cases.

Case 1. The <u>same</u> game.

Here we assume both players are betting on the same trials, e.g. betting on the same coin tosses, or on the same series of hands at blackjack, or on the same games with the same odds at the same sports book. In the stock market, both players could invest in the same "security" at the same time, e.g. a no–load S&P 500 index mutual fund.

We find

$$E(G(f_2) - G(f_1)) = p\log((1+f_2)/(1+f_1)) + q\log((1-f_2)/(1-f_1))$$

and $\mathrm{Var}(G(f_2) - G(f_1)) =$

$$(pq/n)\left\{\log\left[\left(\frac{1+f_2}{1-f_2}\right)\left(\frac{1-f_1}{1+f_1}\right)\right]\right\}^2$$

where $G(f_2) - G(f_1)$ is approximately normally distributed with this mean and variance.

Case 2. Identically distributed independent games.

This corresponds to betting on two different series of tosses with the same coin. $E(G(f_2) - G(f_1))$ is as before. But now $\mathrm{Var}(G(f_2) - G(f_1)) = \mathrm{Var}(G(f_2)) + \mathrm{Var}(G(f_1))$ because $G(f_2)$ and $G(f_1)$ are now independent. Thus $\mathrm{Var}(G(f_2) - G(f_1)) =$

$$(pq/n)\left\{\left[\log\left(\frac{1+f_2}{1-f_2}\right)\right]^2 + \left[\log\left(\frac{1+f_1}{1-f_1}\right)^2\right]\right\}.$$

Let

$$a = \log\left(\frac{1+f_1}{1-f_1}\right) \quad b = \log\left(\frac{1+f_2}{1-f_2}\right).$$

Then in case 1, $V_1 = (pq/n)(a-b)^2$ and in case 2, $V_2 = (pq/n)(a^2+b^2)$ and since $a, b > 0$, $V_1 < V_2$ as expected. We can now compare the Kelly strategy with other fixed fractions to determine the probability that Kelly leads after n trials. Note that this probability is always greater than 1/2 (within the accuracy limits of the continuous approximation, which is the approximation of the binomial distribution by the normal, with its well known and thoroughly studied properties) because $g(f^*) - g(f) > 0$ where $f^* = p - q$ and $f \neq f^*$ is some alternative. This can fail to be true for small n, where the approximation is poor. As an extreme example to make the point, if $n = 1$, any

$f > f^*$ beats Kelly with probability $p > 1/2$. If instead $n = 2$, $f > f^*$ wins with probability p^2 and $p^2 > 1/2$ if $p > 1/\sqrt{2} \doteq .7071$. Also, $f < f^*$ wins with probability $1 - p^2$ and $1 - p^2 > 1/2$ if $p^2 < 1/2$, i.e. $p < 1/\sqrt{2} = .7071$. So when $n = 2$, Kelly always loses more than half the time to some other f unless $p = 1/\sqrt{2}$.

We now have the formulas we need to explore many practical applications of the Kelly criterion.

4 The Long Run: When Will The Kelly Strategy "Dominate"?

The late John Leib wrote several articles for Blackjack Forum which were critical of the Kelly criterion. He was much bemused by "the long run". What is it and when, if ever, does it happen?

We begin with an example.

Example 4.1 $p = .51, \quad n = 10,000$

V_i and s_i, $i = 1, 2$ are the variance and standard deviation, respectively, for 3(e) Cases 1 and 2, and $R = V_2/V_1 = (a^2 + b^2)/(a - b)^2$ so $s_2 = s_1\sqrt{R}$. Table 4.1 summarizes some results. We can also approximate $\sqrt{R}$ with a power series estimate using only the first term of a and of b: $a \doteq 2f_1, b \doteq 2f_2$ so $\sqrt{R} \doteq \sqrt{f_1^2 + f_2^2}/ \mid f_1 - f_2 \mid$. The approximate results, which agree extremely well, are 2.236, 3.606 and 1.581, respectively.

TABLE 4.1 Comparing strategies

f_1	f_2	$g_2 - g_1$	s_1	$(g_2 - g_1)/s_1$	$\sqrt{R}$
.01	.02	.00005001	.00010000	.50	2.236
.03	.02	.00005004	.00010004	.50	3.604
.03	.01	.00000003	.00020005	.00013346	1.581

The first two rows show how nearly symmetric the behavior is on each side of the optimal $f^* = .02$. The column $(g_2 - g_1)/s_1$ shows us that $f^* = .02$ only has a .5 standard deviation advantage over its neighbors $f = .01$ and $f = .03$ after $n = 10,000$ trials. Since this advantage is proportional to $\sqrt{n}$, the column $(g_2 - g_1)/s_1$ from Table 4.1 gives the results of Table 4.2: