Optimal Nonmyopic Gambling Strategy for the Generalized Kelly Criterion

S. Cetinkaya,[1] M. Parlar[2]

[1] *Department of Management Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

[2] *DeGroote School of Business, McMaster University, Hamilton, Ontario L8S 4M4, Canada*

Received May 1996; revised 6 February 1997; accepted 4 April 1997

Abstract: We consider the optimal wagers to be made by a gambler who starts with a given initial wealth. The gambler faces a sequence of two-outcome games, i.e., ''win'' vs. ''lose,'' and wishes to maximize the expected value of his terminal utility. It has been shown by Kelly, Bellman, and others that if the terminal utility is of the form $\log x$, where x is the terminal wealth, then the optimal policy is myopic, i.e., the optimal wager is always to bet a *constant* fraction of the wealth provided that the probability of winning exceeds the probability of losing. In this paper we provide a critique of the simple logarithmic assumption for the utility of terminal wealth and solve the problem with a more general utility function. We show that in the general case, the optimal policy is *not* myopic, and we provide analytic expressions for optimal wager decisions in terms of the problem parameters. We also provide conditions under which the optimal policy reduces to the simple myopic case. © 1997 John Wiley & Sons, Inc. Naval Research Logistics **44:** 639–654, 1997

Keywords: stochastic dynamic programming; gambling; nonmyopic optimal policy

1. INTRODUCTION AND REVIEW OF PREVIOUS RESEARCH

In his seminal paper entitled ''A New Interpretation of Information Rate,'' Kelly [11] considered an unscrupulous gambler who was receiving prior information concerning the outcomes of sporting events over a noisy communication channel (e.g., a phone line) before the results became common knowledge. After receiving the words ''win'' or ''lose''—which he could hear incorrectly due to communication difficulties—the gambler would place his bet (a nonnegative amount up to his present fortune) on the original odds. With p as the probability of correct transmission and $q = 1 - p$ the probability of an incorrect transmission, Kelly showed using calculus techniques that if $p > q$ (i.e., the superfair case),

Correspondence to: M. Parlar
Contract grant sponsor: Natural Sciences and Engineering Research Council of Canada; contract grant number: A5872

 CCC 0894-069X/97/070639-16

the gambler's optimal wager α^* (i.e., the fraction of his capital bet each time) was simply $\alpha^* = p - q$. If $p \le q$ (i.e., the subfair case), then $\alpha^* = 0$. He also showed that the maximum growth of capital occurred at a rate equal to the capacity of the channel given as $G_{\max} = \log_2 2 + p \log_2 p + q \log_2 q$. (Interestingly, this was the same result obtained by Shannon [16] from considerations of coding of information.)

Kelly's work attracted the attention of a large number of researchers including the economists (Arrow [1]), the psychologists (Edwards [8]) and the applied mathematicians (Bellman and Kalaba [3] and [4], Bellman [2]) who extended and reinterpreted Kelly's results. In particular, Bellman and Kalaba re-solved the same problem after formulating it as a dynamic program (DP) with a logarithmic utility function for the terminal wealth of the gambler. They assumed that the gambler is allowed n plays (bets) and at each play of the gamble, he could bet any nonnegative amount up to his present fortune. Defining $V_n(x)$ as the maximal expected return if the gambler has a present fortune of x and is allowed n more gambles, the optimality equation

$$V_n(x) = \max_{0 \le \alpha \le 1} \{pV_{n-1}(x + \alpha x) + qV_{n-1}(x - \alpha x)\}$$

was formed with the boundary condition $V_0(x) = \log x$. It was then shown (as in Kelly [11]) that for this problem with $p > q$ the optimal positive wager fraction is $\alpha^* = p - q$ and the maximum expected return is $V_n(x) = nC + \log x$, where $C = \log 2 + p \log p + q \log q$. This formulation and solution also appears in Ross [14, Section 1.2]. (Bellman, Kalaba, and Ross use the natural logarithm without loss of generality. We will follow this assumption in the subsequent development.)

One of the most interesting features of this problem is the nature of its solution: The optimal strategy is myopic (invariant) in the sense that regardless of the number of bets left to place (n) and the current wealth (x), the optimal (nonnegative) fraction to bet in each period is the same, i.e., $\alpha^* = p - q$ in the superfair case, and $\alpha^* = 0$ in the subfair case. Optimal myopic policies of this type are usually difficult to come by, but they exist for some models. For example, in some periodic-review inventory problems with random demand it has been shown that provided that some terminal cost is appended to the objective function, the optimal order quantity is the same for all periods, i.e., the policy is myopic (Veinott [18], Heyman and Sobel [10, Chapter 3]). In some dynamic portfolio problems (Mossin [13], Bertsekas [5, Section 2.3]) myopic policy is also optimal provided that the amount invested is not constrained and the investor's utility function satisfies certain conditions. For the utility function

$$V_0(x) = \begin{cases} 1, & \text{if } x \ge 1, \\ 0, & \text{if } 0 \le x < 1, \end{cases}$$

studied by Dubins and Savage [7], it has been shown that myopic play is optimal in the subfair case (i.e., if $p \le \frac{1}{2}$ or equivalently $p \le q$), but it is not optimal in the superfair case (i.e., if $p > \frac{1}{2}$ or equivalently $p > q$). However, for this latter case, the optimal strategy has been found by Kulldorf [12].

For the sequential decision problems described above, (e.g., the gambler's problem, the inventory problem, etc.) myopic policies—when optimal—substantially reduce the computational burden since they only need the solution of a single-period optimization problem. It must be noted, however, that such convenience is obtained at the expense of major simplifications in the model. In the inventory problem one must append an unnatural terminal cost term to the objective function. In the gambling and portfolio problems, the decision maker's terminal utility function must be of the form $V_0(x) = \log x$, etc.

In this paper, we argue that the assumption of the simple logarithmic function in the Kelly–Bellman–Kalaba model is somewhat unrealistic and the problem should ideally be solved with a more general logarithmic utility function. We thus assume that, in general, the gambler's terminal utility is given as $V_0(x) = \log(b + x)$, where b is a positive constant. As might be expected, this generalization results in the disappearance of the myopic nature of the solution and the optimal strategy assumes a form that *depends* on the stage (n) and the state (x) of the gambling process. We then find analytic expressions for the nonmyopic optimal strategy.

Although the simple logarithmic terminal utility $V_0(x) = \log x$ results in a convenient myopic solution, if the gambler loses all his bets resulting in a terminal wealth of $x = 0$, then his utility would be $V_0(0) = \log 0 = -\infty$. For Kelly's risk-averse gambler, losing all his wealth is tantamount to financial ''death''; thus his model with $b = 0$ prescribes the proper course of action for such individuals. However, bankruptcy may not be the ''end of the world'' for many people, and the optimal strategy for the $b > 0$ case needs to be obtained for these decision makers.

Arrow [1] has stated that the relative risk aversion index $\rho = -xV_0''(x)/V_0'(x)$ should ''hover around 1, being, if anything, somewhat less for low wealths and somewhat higher for high wealths.'' For the case of general logarithmic utility $V_0(x) = \log(b + x)$, we have $\rho = x/(b + x) < 1$ for finite x; thus generalizing the definition of the gambler's terminal utility still satisfies Arrow's requirement for the relative risk aversion index. For additional discussions on the general logarithmic utility $\log(b + x)$, see Freimer and Gordon [9] and Samson [15].

Kelly's results have been extended in a different direction by Breiman [6] who considers the case of multiple gambling/investment opportunities existing at each stage of the process. Breiman has shown that for this problem the optimal strategy is also myopic when the gambler's terminal utility is again given by $\log x$. See Thorp [17] for a good discussion of Breiman's results.

The rest of this paper is concerned with the solution of the problem for the generalized Kelly criterion. The optimality equation resulting from the DP formulation is

$$V_n(x) = \max_{0 \le \alpha \le 1} \{pV_{n-1}(x + \alpha x) + qV_{n-1}(x - \alpha x)\}, \qquad n = 1, 2, \ldots, \tag{1}$$

with the boundary condition given as the general logarithmic terminal utility

$$V_0(x) = \log(b + x), \qquad b > 0, \tag{2}$$

and

$$p + q = 1. \tag{3}$$

We show that this more general formulation leads to a nonmyopic strategy which we obtain analytically as a function of the stage n, the state x, and the win probability p.

2. SINGLE-STAGE AND TWO-STAGE PROBLEMS

In this section we will analyze some of the mathematical properties of the dynamic program that we introduced in the previous section. First, in Section 2.1 we will discuss the solution of the single-stage problem (i.e., the single-play game), and we will show that the optimal decision depends on gambler's fortune x at the beginning of the game *and* the win probability p. Next, we will present the solution for the 2-stage problem in Section 2.2. We will then show that the optimal strategy is not necessarily myopic for the 2-stage problem, i.e., if the win probability p is greater than $\frac{1}{2}$ (the superfair case), then the single-stage strategy will no longer be optimal.

We will first prove a lemma in which we discuss the properties of the following functions:

$$u_{k+1}(\alpha, x) = p \log(b + 2^k x + 2^k \alpha x) + q \log(b + 2^k x - 2^k \alpha x), \qquad (4)$$
$$k = 0, 1, \ldots, \quad 0 \le \alpha \le 1, \quad 0 \le x < \infty.$$

LEMMA 1: Functions $u_{k+1}(\alpha, x)$, $k = 0, 1, \cdots$ are concave over α. For any given x, the maximizer of $u_{k+1}(\alpha, x)$, $k = 0, 1, \cdots$ over α is denoted by α_{k+1}, $k = 0, 1, \ldots$, where

$$\alpha_{k+1} = (p - q)\left(1 + \frac{b}{2^k x}\right). \qquad (5)$$

Furthermore,

$$\max_{0 \le \alpha \le 1} u_{k+1}(\alpha, x) = \begin{cases} u_{k+1}(0, x), & \text{if } p \le \dfrac{1}{2}, \\[2ex] u_{k+1}(\alpha_{k+1}, x), & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2^{k+1}x}{2(b + 2^k x)}, \\[2ex] u_{k+1}(1, x), & \text{if } p > \dfrac{b + 2^{k+1}x}{2(b + 2^k x)}, \end{cases} \qquad (6)$$

where

$$u_{k+1}(0, x) = \log(b + 2^k x), \qquad (7)$$

$$u_{k+1}(\alpha_{k+1}, x) = \log(b + 2^k x) + C, \qquad (8)$$

$$u_{k+1}(1, x) = p \log(b + 2^{k+1}x) + D, \tag{9}$$

and

$$C \equiv p \log(1 + p - q) + q \log(1 - p + q), \tag{10}$$

$$D \equiv q \log b. \tag{11}$$

PROOF: The first two derivatives of $u_{k+1}(\alpha, x)$, $k = 0, 1, \cdots$ with respect to α are given by

$$\frac{\partial u_{k+1}(\alpha, x)}{\partial \alpha} = \frac{2^k px}{b + 2^k x + 2^k \alpha x} - \frac{2^k qx}{b + 2^k x - 2^k \alpha x}, \tag{12}$$

and

$$\frac{\partial^2 u_{k+1}(\alpha, x)}{\partial \alpha^2} = -\frac{2^{2k} px^2}{(b + 2^k x + 2^k \alpha x)^2} - \frac{2^{2k} qx^2}{(b + 2^k x - 2^k \alpha x)^2}. \tag{13}$$

The right-hand side of (13) implies that $u_{k+1}(\alpha, x)$ is a concave function of α. It then follows that α_{k+1} that solves $\partial u_{k+1}(\alpha, x)/\partial \alpha = 0$ maximizes $u_{k+1}(\alpha, x)$. Thus, using (12) it can easily be shown that

$$\alpha_{k+1} = (p - q)\left(1 + \frac{b}{2^k x}\right). \tag{14}$$

Inserting $q = 1 - p$ in the above expression, we can also write

$$\alpha_{k+1} = (2p - 1)\left(1 + \frac{b}{2^k x}\right). \tag{15}$$

Here we note that if

$$\frac{1}{2} \le p \le \frac{b + 2^{k+1}x}{2(b + 2^k x)},$$

then (15) implies that $0 \le \alpha_{k+1} \le 1$. Thus, the result in (6) follows since $u_{k+1}(\alpha, x)$ is concave over α.

Now, inserting $\alpha = 0$ in (4), and using (3), we have

$$u_{k+1}(0, x) = \log(b + 2^k x).$$

Also inserting (5) for α in (4), and using (3) and (10), we obtain

$$u_{k+1}(\alpha_{k+1}, x) = \log(b + 2^k x) + C.$$

Finally, inserting $\alpha = 1$ in (4) and using (3) and (11) gives

$$u_{k+1}(1, x) = p \log(b + 2^{k+1} x) + D.$$

This completes the proof. ■

2.1. Single-Stage Problem

In this section we present a theorem that gives the optimal solution for the single-stage problem where the gambler has only one chance to play.

THEOREM 1: If the gambler has a single play of the game, then the optimal percentage to bet denoted by α_1^* and the maximal expected return $V_1(x)$ are specified as follows depending on the initial fortune x, and the probability of winning p:

p:	$p \leq \frac{1}{2}$	$\frac{1}{2} < p \leq \frac{b + 2x}{2(b + x)}$	$p > \frac{b + 2x}{2(b + x)}$
$\alpha_1^* =$	0	$\alpha_1 = (p - q)\left(1 + \frac{b}{x}\right)$	1
$V_1(x) =$	$\log(b + x)$	$\log(b + x) + C$	$p \log(b + 2x) + D$

PROOF: Suppose the gambler is allowed a single bet. The optimality equation (1) leads to $V_1(x) = \max_{0 \leq \alpha \leq 1} \{pV_0(x + \alpha x) + qV_0(x - \alpha x)\}$. Utilizing (2) in this equation, we have

$$V_1(x) = \max_{0 \leq \alpha \leq 1} \{p \log(b + x + \alpha x) + q \log(b + x - \alpha x)\}. \tag{16}$$

If we now use (4) in (16), we can write

$$V_1(x) = \max_{0 \leq \alpha \leq 1} u_1(\alpha, x). \tag{17}$$

Let α_1^* denote the *optimal* percentage to bet if the gambler has only one play. Lemma 1 states that $u_1(\alpha, x)$ is a concave function of α. Since the maximizer of $u_1(\alpha, x)$ over α is denoted by α_1, its expression can be obtained by substituting $k = 0$ in (5). We can also obtain an alternative expression for α_1 by substituting $k = 0$ in (15). Furthermore, if

$$\frac{1}{2} \le p \le \frac{b + 2x}{2(b + x)},$$

then (15) implies that $0 \le \alpha_1 \le 1$. Thus, results of Lemma 1 and (17) imply that

$$\alpha_1^* = \begin{cases} 0, & \text{if } p \le \dfrac{1}{2}, \\ \alpha_1, & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2x}{2(b + x)}, \\ 1, & \text{if } p > \dfrac{b + 2x}{2(b + x)}, \end{cases} \tag{18}$$

where α_1 is obtained by substituting $k = 0$ in (5) or (15).

Observe that if $p \le \frac{1}{2}$, then the optimal decision is not to bet at all, i.e., $\alpha_1^* = 0$. Otherwise, depending on gambler's fortune x at the beginning of the game and the probability of winning p, the optimal decision may be either to bet some fraction of the current fortune denoted by α_1, or the entire fortune, i.e., $\alpha_1^* = 1$.

Using (18), we now express $V_1(x)$ explicitly. It follows from (17) and the concavity of $u_1(\alpha, x)$ that $V_1(x) = u_1(\alpha_1^*, x)$. Thus,

$$V_1(x) = \begin{cases} u_1(0, x), & \text{if } p \le \dfrac{1}{2}, \\ u_1(\alpha_1, x), & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2x}{2(b + x)}, \\ u_1(1, x), & \text{if } p > \dfrac{b + 2x}{2(b + x)}. \end{cases} \tag{19}$$

Substituting $k = 0$ in (7), (8), and (9), and utilizing these in (19), we get

$$V_1(x) = \begin{cases} \log(b + x), & \text{if } p \le \dfrac{1}{2}, \\ \log(b + x) + C, & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2x}{2(b + x)}, \\ p \log(b + 2x) + D, & \text{if } p > \dfrac{b + 2x}{2(b + x)}. \end{cases} \tag{20}$$

This completes the proof. ■

So far we computed the optimal strategy for the single-stage problem. Using the optimal percentage to bet, denoted by α_1^* and given by (18), we obtained an explicit expression for the maximal expected return $V_1(x)$ given by (20).

Note that for either $b = 0$ (i.e., the Kelly criterion) or *very large* initial wealth ($x \to \infty$), the solution we have found reduces to the one found by Kelly, i.e., $\alpha_1^* \to p - q$ for $p > q$, and $\alpha_1^* = 0$ for $p \le q$.

Next, we utilize $V_1(x)$ given by (20) in order to compute the optimal strategy for the 2-stage problem.

2.2. Two-Stage Problem and the Nonmyopic Optimal Policy

Using (1), we can write $V_2(x) = \max_{0 \le \alpha \le 1} \{pV_1(x + \alpha x) + qV_1(x - \alpha x)\}$. Inserting $V_1(\cdot)$ of (20) in this expression and using (3), we have

$$V_2(x) = \max_{0 \le \alpha \le 1} \begin{cases} p \log(b + x + \alpha x) \\ \quad + q \log(b + x - \alpha x), & \text{if } p \le \dfrac{1}{2}, \\ p \log(b + x + \alpha x) \\ \quad + q \log(b + x - \alpha x) + C, & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2x}{2(b + x)}, \\ p^2 \log(b + 2x + 2\alpha x) \\ + qp \log(b + 2x - 2\alpha x) + D, & \text{if } p > \dfrac{b + 2x}{2(b + x)}. \end{cases} \tag{21}$$

If we use the functions defined by (4), then (21) can be expressed as

$$V_2(x) = \max_{0 \le \alpha \le 1} \begin{cases} u_1(\alpha, x), & \text{if } p \le \dfrac{1}{2}, \\ u_1(\alpha, x) + C, & \text{if } \dfrac{1}{2} < p \le \dfrac{b + 2x}{2(b + x)}, \\ pu_2(\alpha, x) + D, & \text{if } p > \dfrac{b + 2x}{2(b + x)}. \end{cases} \tag{22}$$

Let α_2^* denote the optimal percentage to bet at the beginning of the game if the gambler has two more plays, i.e., if the gambler has two more stages to go. Let us recall Lemma 1 which states that $u_1(\alpha, x)$ and $u_2(\alpha, x)$ are concave functions of α. Note that the maximizers of $u_1(\alpha, x)$ and $u_2(\alpha, x)$ are denoted by α_1 and α_2 and the expressions for α_1 and α_2 are provided by substituting $k = 0$ and $k = 1$ in (5) [or in (15)], respectively. As we noted in the previous section, for $k = 0$, (15) implies that if

$$\frac{1}{2} \le p \le \frac{b + 2x}{2(b + x)},$$

then $0 \le \alpha_1 \le 1$. Similarly, for $k = 1$, it follows from (15) that if

$$\frac{1}{2} \le p \le \frac{b + 4x}{2(b + 2x)},$$

then $0 \le \alpha_2 \le 1$. Thus, using (22), we have

$$\alpha_2^* = \begin{cases} 0, & \text{if } p \le \dfrac{1}{2}, \\[2ex] \alpha_1, & \text{if } \dfrac{1}{2} < p \le \dfrac{b+2x}{2(b+x)}, \\[2ex] \alpha_2, & \text{if } \dfrac{b+2x}{2(b+x)} < p \le \dfrac{b+4x}{2(b+2x)}, \\[2ex] 1, & \text{if } p > \dfrac{b+4x}{2(b+2x)}. \end{cases} \tag{23}$$

Furthermore, (22) and (23) lead to the following expression for $V_2(x)$:

$$V_2(x) = \begin{cases} u_1(0, x), & \text{if } p \le \dfrac{1}{2}, \\[2ex] u_1(\alpha_1, x) + C, & \text{if } \dfrac{1}{2} < p \le \dfrac{b+2x}{2(b+x)}, \\[2ex] pu_2(\alpha_2, x) + D, & \text{if } \dfrac{b+2x}{2(b+x)} < p \le \dfrac{b+4x}{2(b+2x)}, \\[2ex] pu_2(1, x) + D, & \text{if } p > \dfrac{b+4x}{2(b+2x)}. \end{cases} \tag{24}$$

Now, using (7), (8), and (9) in (24), we obtain

$$V_2(x) = \begin{cases} \log(b+x), & \text{if } p \le \dfrac{1}{2}, \\[2ex] \log(b+x) + 2C, & \text{if } \dfrac{1}{2} < p \le \dfrac{b+2x}{2(b+x)}, \\[2ex] p\log(b+2x) + pC + D, & \text{if } \dfrac{b+2x}{2(b+x)} < p \le \dfrac{b+4x}{2(b+2x)}, \\[2ex] p^2\log(b+4x) + pD + D, & \text{if } p > \dfrac{b+4x}{2(b+2x)}. \end{cases} \tag{25}$$

Our results indicate that optimality of a myopic play holds for the subfair gamble with $p \le \frac{1}{2}$. As Kelly–Bellman–Kalaba (see [11], [3], and [4]), we observe that it is optimal to bet nothing in this case. However, the optimal strategy for the 2-stage game is not myopic in general since the single-stage solution is not optimal for the 2-stage problem in the superfair case: If $p > \frac{1}{2}$, the optimal strategy depends on the gambler's wealth x at the beginning of the game, the probability of winning p, and the number of plays left n $(=2)$.

It is worth noting that the present nonmyopic solution converges to the myopic one (given by Kelly–Bellman–Kalaba) for the special case of very large initial wealth. That is, as $x \to \infty$ the optimal wager $\alpha_2^* \to p - q$ for $p > q$, and $\alpha_2^* = 0$ for $p \leq q$. It follows that if the starting wealth is sufficiently large before a gamble, α_n^* $(n = 1, 2)$ is approximated by $p - q$ for the superfair case.

For the single-stage problem there are three possible values for the optimal percentage to bet α_1^*. However, for the 2-stage problem we find four possible values for the optimal percentage to bet α_2^*. In general, for the n-stage gamble there are $n + 2$ possible values for the optimal percentage to bet, denoted by α_n^*, depending on gambler's initial fortune x, and the probability of winning p. In the following section we provide analytical expressions for the optimal strategy α_n^* and the maximal expected return $V_n(x)$ for the n-stage game.

3. NONMYOPIC OPTIMAL POLICY FOR THE *n*-STAGE PROBLEM

Let α_n^* denote the optimal percentage to bet if there are n stages left in the process, i.e., if the gambler has n more plays of the game. The theorem we will present below summarizes our results for the optimal solution α_n^* and the maximal expected return $V_n(x)$ for the n-stage problem where $n \geq 2$.

THEOREM 2: If the gambler has $n \geq 2$ plays of the game left, then depending on the initial fortune x, the probability of winning p and the number of plays left n; the optimal percentage to bet α_n^* and the maximal expected return $V_n(x)$ are given as

p:	$p \leq \frac{1}{2}$	$\frac{1}{2} < p \leq \frac{b + 2x}{2(b + x)}$	$\frac{b + 2^k x}{2(b + 2^{k-1}x)} < p \leq \frac{b + 2^{k+1}x}{2(b + 2^k x)}$ for $k = 1, 2, \ldots, n - 1$	$p > \frac{b + 2^n x}{2(b + 2^{n-1}x)}$
$\alpha_n^* =$	0	α_1	α_{k+1}, $k = 1, 2, \ldots, n - 1$	1
$V_n(x) =$	$\log(b + x)$	$\log(b + x) + nC$	$p^k \log(b + 2^k x) + (n - k)p^k C + \sum_{j=0}^{k-1} p^j D$	$p^n \log(b + 2^n x) + \sum_{j=0}^{n-1} p^j D$

where

$$\alpha_j = (p - q)\left(1 + \frac{b}{2^{j-1}x}\right), \qquad j = 1, 2, \ldots, n. \tag{26}$$

(See Figure 1 for a graphical description of the policy regions in terms of the wealth x and probability of winning p.)

PROOF: The proof is by induction on $n = 2, 3, \ldots$. We analyze the following four cases:

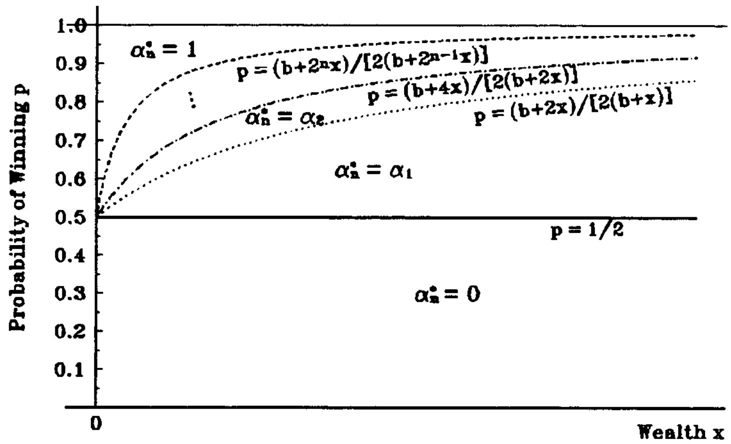

Figure 1. Optimal Policy Regions.

Case 1 $[p \leq \frac{1}{2}]$

First we show that if $p \leq \frac{1}{2}$, then

$$\alpha_n^* = 0 \qquad \text{and} \qquad V_n(x) = \log(b + x).$$

Recall that according to our results in Section 2.2, if $p \leq \frac{1}{2}$, then $\alpha_2^* = 0$ and $V_2(x) = \log(b + x)$. Thus, if $p \leq \frac{1}{2}$, then the theorem is already proved for $n = 2$.

Now suppose that for $p \leq \frac{1}{2}$, the maximal expected return for the $n - 1$ period problem is $V_{n-1}(x) = \log(b + x)$. It follows from (1) that

$$V_n(x) = \max_{0 \leq \alpha \leq 1} \{p \log(b + x + \alpha x) + q \log(b + x - \alpha x)\}.$$

Using functions (4), the above equation can be rewritten as

$$V_n(x) = \max_{0 \leq \alpha \leq 1} u_1(\alpha, x).$$

Thus, using Lemma 1, we find that if $p \leq \frac{1}{2}$, then

$$V_n(x) = \max_{0 \leq \alpha \leq 1} u_1(\alpha, x) = u_1(0, x).$$

That is, $\alpha_n^* = 0$ for this case. Substituting $k = 0$ in (7), we obtain $V_n(x) = u_1(0, x) = \log(b + x)$.

Case 2 [$\frac{1}{2} < p \le (b + 2x)/2(b + x)$]

Now we show that if $\frac{1}{2} < p \le (b + 2x)/2(b + x)$, then

$$\alpha_n^* = \alpha_1 \qquad \text{and} \qquad V_n(x) = \log(b + x) + nC.$$

where

$$\alpha_1 = (p - q)\left(1 + \frac{b}{x}\right).$$

Based on our analysis in Section 2.2, $\alpha_2^* = \alpha_1$ for this case, where α_1 is obtained by substituting $k = 0$ in (5). Also recall (25) which states that for this case $V_2(x) = \log(b + x) + 2C$. Thus, our results for Case 2 are proved if $n = 2$.

Suppose now that $V_{n-1}(x) = \log(b + x) + (n - 1)C$. Using (3) and (1), we can write

$$V_n(x) = \max_{0\le\alpha\le1} \{p \log(b + x + \alpha x) + q \log(b + x - \alpha x) + (n - 1)C\}.$$

If we use (4), the above equation can be rewritten as

$$V_n(x) = \max_{0\le\alpha\le1} \{u_1(\alpha, x) + (n - 1)C\}.$$

It follows from Lemma 1 that

$$\max_{0\le\alpha\le1} u_1(\alpha, x) = u_1(\alpha_1, x)$$

for Case 2. Thus, $\alpha_n^* = \alpha_1$, and α_1 is given by (26). Using (8), we see that

$$V_n(x) = u_1(\alpha_1, x) + (n - 1)C = \log(b + x) + nC.$$

Case 3 [$(b + 2^k x)/2(b + 2^{k-1}x) < p \le (b + 2^{k+1}x)/2(b + 2^k x)$ *for Some* k *Such That* $k \in \{1, 2, \dots, n - 1\}$]

Now, we show that if

$$\frac{b + 2^k x}{2(b + 2^{k-1}x)} < p \le \frac{b + 2^{k+1}x}{2(b + 2^k x)} \tag{27}$$

for some value of $k \in \{1, 2, \dots, n - 1\}$, then

$$\alpha_n^* = \alpha_{k+1} \qquad \text{and} \qquad V_n(x) = p^k \log(b + 2^k x) + (n - k)p^k C + \sum_{j=0}^{k-1} p^j D,$$

where α_{k+1} is given by (26).

Consider $n = 2$. Observe that the only possible value for k is 1 for $n = 2$. If we insert $k = 1$ in (27), then we have

$$\frac{b + 2x}{2(b + x)} < p \leq \frac{b + 4x}{2(b + 2x)}.$$

We need to find α_2^* if p lies on the above interval. According to (23), we have $\alpha_2^* = \alpha_2$, where α_2 is obtained by substituting $k = 1$ in (5). Also, Eq. (25) states that $V_2(x) = p \log(b + 2x) + pC + D$ for this case. Thus, our results for Case 3 are proved if $n = 2$.

Suppose that if (27) holds for some value of $k \in \{1, 2, \ldots, n - 2\}$, then

$$V_{n-1}(x) = p^k \log(b + 2^k x) + (n - 1 - k)p^k C + \sum_{j=0}^{k-1} p^j D.$$

We now fix k, and complete the induction on n. Using (1) and (3), we can write

$$V_n(x) = \max_{0\leq\alpha\leq 1} \Bigg\{ p^{k+1} \log(b + 2^k x + 2^k \alpha x) + qp^k \log(b + 2^k x - 2^k \alpha x) + (n - 1 - k)p^k C + \sum_{j=0}^{k-1} p^j D \Bigg\}. \tag{28}$$

If we utilize (4) in (28), we get

$$V_n(x) = \max_{0\leq\alpha\leq 1} \Bigg\{ p^k u_{k+1}(\alpha, x) + (n - 1 - k)p^k C + \sum_{j=0}^{k-1} p^j D \Bigg\}. \tag{29}$$

By induction, assumption (27) holds, and it follows from Lemma 1 that $\alpha_n^* = \alpha_{k+1}$. Furthermore, using this result in (29) leads to

$$V_n(x) = p^k u_{k+1}(\alpha_{k+1}, x) + (n - 1 - k)p^k C + \sum_{j=0}^{k-1} p^j D.$$

If we use (8) in the above expression, we have

$$V_n(x) = p^k \log(b + 2^k x) + (n - k)p^k C + \sum_{j=0}^{k-1} p^j D.$$

In order to complete the proof for Case 3, now we need to show that if p lies in the range defined by (27) where $k = n - 1$, then

$$\alpha_n^* = \alpha_n \qquad \text{and} \qquad V_n(x) = p^{n-1} \log(b + 2^{n-1} x) + p^{n-1} C + \sum_{j=0}^{n-2} p^j D, \tag{30}$$

where α_n is given by (26). Suppose that (27) holds for $k = n - 1$. Then our induction assumption is

$$V_{n-1}(x) = p^{n-1} \log(b + 2^{n-1}x) + \sum_{j=0}^{n-2} p^j D.$$

It follows from (1) and (3) that

$$V_n(x) = \max_{0\le\alpha\le 1} \Bigg\{ p^n \log(b + 2^{n-1}x + 2^{n-1}\alpha x) + qp^{n-1} \log(b + 2^k x - 2^k \alpha x) + \sum_{j=0}^{n-2} p^j D \Bigg\}. \tag{31}$$

If we use (4), then (31) can be expressed as

$$V_n(x) = \max_{0\le\alpha\le 1} \Bigg\{ p^{n-1} u_n(\alpha, x) + \sum_{j=0}^{n-2} p^j D \Bigg\}. \tag{32}$$

Based on Lemma 1 it is easy to see $\alpha_n^* = \alpha_n$ where α_n is given by (26) for $j = n$. Then (32) leads to

$$V_n(x) = p^{n-1} u_n(\alpha_n, x) + \sum_{j=0}^{n-2} p^j D. \tag{33}$$

If we insert $k = n - 1$ in (8), we obtain the expression of $u_n(\alpha_n, x)$. Thus, using this expression in (33), we have $V_n(x) = p^{n-1} \log(b + 2^{n-1}x) + p^{n-1}C + \sum_{j=0}^{n-2} p^j D$.

Case 4 $[p > (b + 2^n x)/2(b + 2^{n-1}x)]$

Finally, we show that if $p > (b + 2^n x)/2(b + 2^{n-1}x)$, then

$$\alpha_n^* = 1 \qquad \text{and} \qquad V_n(x) = p^n \log(b + 2^n x) + \sum_{j=0}^{n-1} p^j D.$$

Based on our analysis in Section 2.2, we know that $\alpha_2^* = 1$ and $V_2(x) = p^2 \log(b + 4x) + pD + D$. Thus, the proposition is proved for $n = 2$.

Suppose that

$$V_{n-1}(x) = p^{n-1} \log(b + 2^{n-1}x) + \sum_{j=0}^{n-2} p^j D.$$

Thus, using (1) and (3), we can write

$$V_n(x) = \max_{0 \le \alpha \le 1} \left\{ p^n \log(b + 2^{n-1}x + 2^{n-1}\alpha x) + qp^{n-1} \log(b + 2^{n-1}x - 2^{n-1}\alpha x) + \sum_{j=0}^{n-2} p^j D \right\}.$$

If we utilize (4) in the above equation, we have

$$V_n(x) = \max_{0 \le \alpha \le 1} \left\{ p^{n-1} u_n(\alpha, x) + \sum_{j=0}^{n-2} p^j D \right\}.$$

Lemma 1 then implies that $\alpha_n^* = 1$; thus

$$V_n(x) = p^{n-1} u_n(1, x) + \sum_{j=0}^{n-2} p^j D.$$

Substituting $k = n - 1$ in (9), we obtain the expression of $u_n(1, x)$. If we insert this in the above expression, we obtain

$$V_n(x) = p^n \log(b + 2^n x) + \sum_{j=0}^{n-1} p^j D.$$

This completes the proof. ■

As in the two-stage problem discussed in Section 2.2, optimality of myopic play holds for the subfair case; i.e., if $p \le \frac{1}{2}$, then $\alpha_1^* = \alpha_2^* = \cdots = \alpha_n^* = 0$. For the superfair case, however, the optimal strategy is nonmyopic, and depends on the initial wealth (x) and the stage (n). We observe that if the gambler starts with a very large initial wealth $(x \to \infty)$, then his optimal decision becomes independent of the stage n and the state x, i.e., $\alpha_n^* \to p - q$ for $p > q$, and $\alpha_n^* = 0$ for $p \le q$; thus, Kelly–Bellman–Kalaba's myopic play is optimal. Provided that the player starts gambling with a very large initial wealth and continues to win his bets, this myopic policy remains optimal. On the other hand, once a bet is lost, the myopic play is not necessarily optimal for the entire game.

4. CONCLUDING COMMENTS

In this study we computed the optimal strategy for a gambler who had a general logarithmic utility function for his terminal wealth. The gambler faced a sequential game with n plays. We restricted our attention to a simple sequential decision making problem such that a single investment opportunity with two outcomes arises at each stage. We showed that maximizing the utility of final wealth (fortune) leads to a nonmyopic policy for the superfair case if the utility function is given by $\log(b + x)$, where $b > 0$. (The case of $b = 0$ as discussed by Kelly [11] and others leads to a myopic policy for both subfair and superfair cases.) We also discussed that if the game is subfair, then it is always optimal to bet nothing, i.e., optimality of a myopic play holds much more generally in this case.

Since the optimal strategy is nonmyopic for the two-outcome case when $b > 0$, we expect that the optimal strategy would also be nonmyopic for the more general cases of this gambling situation. For example, Breiman's [6] results for the multiple gambling/investment opportunity would not hold. Therefore, restricting attention to the class of myopic strategies may be suboptimal, especially if the main interest is the dynamic portfolio choice for an investor with a general logarithmic utility function. Computing the optimal nonmyopic policies for these more general problems appears to be fruitful research endeavor.

ACKNOWLEDGMENT

Research supported by the Natural Sciences and Engineering Research Council of Canada under Grant No. A5872 and conducted while the first author was a doctoral student in Management Science/Systems at McMaster University. The authors thank an anonymous referee for useful comments and suggestions on an earlier draft of this paper.

REFERENCES

[1] Arrow, K.J., *Essays in the Theory of Risk-Bearing,* Markham, Chicago, 1971.
[2] Bellman, R., *Adaptive Control Processes: A Guided Tour,* Princeton University Press, Princeton, NJ, 1961.
[3] Bellman, R., and Kalaba, R.E., ''Dynamic Programming and Statistical Communication Theory,'' *Proceedings of the National Academy of Sciences, USA,* **43,** 749–751 (1957).
[4] Bellman, R., and Kalaba, R.E., ''On the Role of Dynamic Programming in Statistical Communication Theory,'' *IRE Transactions on Information Theory,* **IT-3,** 197–203 (1957).
[5] Bertsekas, D., *Dynamic Programming: Deterministic and Stochastic Models,* Prentice-Hall, Englewood Cliffs, NJ, 1987.
[6] Breiman, L., ''Investment Policies for Expanding Businesses Optimal in the Long Run Sense,'' *Naval Research Logistics Quarterly,* **4,** 647–651 (1960).
[7] Dubins, L.E., and Savage, L.J., *Inequalities for Stochastic Processes (How to Gamble If You Must),* Dover, New York, 1976.
[8] Edwards, W., ''Dynamic Decision Theory and Probabilistic Information Processing,'' *Human Factors,* **4,** 59–73 (1962).
[9] Freimer, M., and Gordon, M.J., ''Investment Behaviour with Utility a Concave Function of Wealth,'' in K. Borch and J. Mossin (Eds.), *Risk and Uncertainty,* Macmillan, New York, 1968, pp. 94–119.
[10] Heyman, D.P., and Sobel, M.J., *Stochastic Models in Operations Research, Volume II: Stochastic Optimization,* McGraw-Hill, New York, 1984.
[11] Kelly, J.L., ''A New Interpretation of Information Rate,'' *Bell System Technical Journal,* **35,** 917–926 (1956).
[12] Kulldorf, M., ''Optimal Control of Favorable Games with a Time Limit,'' *SIAM Journal on Control and Optimization,* **31,** 52–69 (1993).
[13] Mossin, J., ''Optimal Multi-Period Portfolio Policies,'' *Journal of Business,* **41,** 215–229 (1968).
[14] Ross, S.M., *Introduction to Stochastic Dynamic Programming,* Academic, New York, 1983.
[15] Samson, D., *Managerial Decision Analysis,* Irwin, Homewood, 1988.
[16] Shannon, C., ''A Mathematical Theory of Communication,'' *Bell System Technical Journal,* **27,** 379–423 (1948).
[17] Thorp, E.O., ''Portfolio Choice and the Kelly Criterion,'' in W.T. Ziemba and R.G. Vickson (Eds.), *Stochastic Optimization Models in Finance,* Academic, New York, 1975.
[18] Veinott, A.F., ''Optimal Policy for a Multi-Product, Dynamic Nonstationary Inventory Problem,'' *Management Science,* **12,** 206–222 (1965).