Portfolio Choice and the Bayesian Kelly Criterion

Sid Browne [1]
Columbia University

Ward Whitt [2]
AT&T Bell Laboratories

Original: March 4, 1994
Final Version: August 3, 1995

Appeared in *Advances in Applied Probability*, **28**, 4: 1145-1176, December 1996

[1]Postal address: 402 Uris Hall, Graduate School of Business, Columbia University, New York, NY 10027
[2]Postal address: AT&T Bell Laboratories, Room 2C-178, Murray Hill NJ 07974-0636

Abstract

We derive optimal gambling and investment policies for cases in which the underlying stochastic process has parameter values that are unobserved random variables. For the objective of maximizing logarithmic utility when the underlying stochastic process is a simple random walk in a random environment, we show that a state-dependent control is optimal, which is a generalization of the celebrated Kelly strategy: The optimal strategy is to bet a fraction of current wealth equal to a linear function of the posterior mean increment. To approximate more general stochastic processes, we consider a continuous-time analog involving Brownian motion. To analyze the continuous-time problem, we study the diffusion limit of random walks in a random environment. We prove that they converge weakly to a Kiefer process, or tied-down Brownian sheet. We then find conditions under which the discrete-time process converges to a diffusion, and analyze the resulting process. We analyze in detail the case of the natural conjugate prior, where the success probability has a beta distribution, and show that the resulting limiting diffusion can be viewed as a rescaled Brownian motion. These results allow explicit computation of the optimal control policies for the continuous-time gambling and investment problems without resorting to continuous-time stochastic-control procedures. Moreover they also allow an explicit quantitative evaluation of the financial value of randomness, the financial gain of perfect information and the financial cost of learning in the Bayesian problem.

Key words: betting systems; proportional gambling; Kelly criterion; portfolio theory; logarithmic utility; random walks in a random environment; Kiefer process; time-changed Brownian motion; conjugate priors; Bayesian control.

AMS 1991 Subject Classification: Primary: 90A09. Secondary: 60G40, 62C10, 60J60.

1 Introduction

Suppose you are faced with a sequence of *favorable games*, and you decide to bet repeatedly on these games using *proportional betting*. If you bet x on the n^{th} game, then your return is xZ_n, where $\{Z_n : n \geq 1\}$ is a sequence of i.i.d. random variables with $EZ_n > 0$. Let V_n be your fortune after n bets, and let f_n denote the proportion of your wealth that you wager on the nth bet. Your fortune then evolves as

$$V_n = V_{n-1} + (f_n V_{n-1})Z_n = V_{n-1}(1 + f_n Z_n) = V_0 \prod_{i=1}^{n}(1 + f_i Z_i) \ , \ n \geq 1 \ . \tag{1}$$

Of considerable interest is the special case where you always bet the same *constant proportion*, $f_n = f$, for all $n \geq 1$. For such a policy, let the *growth rate* of your strategy be defined by

$$G_n(f) \equiv \frac{1}{n} \ln(V_n/V_0) \equiv \frac{1}{n} \sum_{i=1}^{n} \ln(1 + fZ_i) \ .$$

By the Law of Large Numbers,

$$G_n(f) \to G(f) \equiv E(\ln(1 + fZ_1)) \text{ w.p.1 as } n \to \infty \ .$$

You can optimize your long-run growth rate by choosing f to maximize $G(f)$; this is commonly referred to as the *Kelly criterion*, and the resulting constant proportional strategy is commonly referred to as the Kelly gambling scheme, in honor of the seminal paper by Kelly (1956). Proportional betting or Kelly gambling has since been quite extensively studied; see Breiman (1961), Thorp (1969), Bell and Cover (1980), Finkelstein and Whitely (1981), Ethier and Tavare (1983), Ethier (1988), Algoet and Cover (1988) and Cover and Thomas (1991).

Kelly (1956) treated the special case in which Z_n is the increment in a *simple random walk*, i.e.,

$$P(Z_n = 1) = \theta = 1 - P(Z_n = -1) \, .$$

For this case,

$$G(f) = \theta \ln(1 + f) + (1 - \theta) \ln(1 - f) \ ,$$

from which we see that the optimal fixed fraction is just the mean increment in the random walk

$$f^* = 2\theta - 1 = EZ_n \tag{2}$$

and the optimal win rate is

$$G(f^*) = \Lambda(\theta) := \ln 2 + \theta \ln \theta + (1 - \theta) \ln(1 - \theta) \ , \tag{3}$$

which is $\ln 2 - H$, where H is the *entropy* of the distribution of Z_1.

For more general random walks, where Z_n has an arbitrary distribution, it is clear that the optimal policy is still $f^* = \arg\sup_f E(\ln(1 + fZ_1))$, although computation of the optimal policy is

in general much more difficult. Some approximations to the optimal policy for discrete-time models are developed in Ethier and Tavare (1983), Ethier (1988) and Thorp (1971).

It is now known that proportional betting has many good properties, besides maximizing the growth rate. For example, Breiman (1961) proved that f^* also *asymptotically* minimizes the expected time to reach a fixed fortune, and asymptotically dominates any other strategy. There are optimality properties associated with this strategy for *finite horizons* as well. For example, Bell and Cover (1980) proved that this strategy is also optimal in a game theoretic sense for finite horizons. Bellman and Kalaba (1957) also considered the problem for a simple random walk over a finite horizon and proved that this policy is optimal for the equivalent problem of maximizing the utility of terminal wealth at any fixed terminal time, when the utility function is logarithmic. Proportional gambling and the Kelly problem have also been considered directly in a continuous-time setting in Pestien and Sudderth (1985), Gottlieb (1985) and Heath, Orey, Pestien and Sudderth (1987). In the continuous-time model, the underlying random walk is replaced by a Brownian motion (with given positive drift). Many of the discrete-time optimality properties of proportional gambling also carry over into the continuous-time model as well.

The problem of optimal gambling in repeated favorable games is intimately related to the problem of optimal multi-period *investment* in financial economics. The only essential difference in fact is the option of investing in a risk free security that pays a nonstochastic interest rate $r > 0$. We consider here the simple case in which there is only one risky stock available for investment. In this case Z_n denotes the return of the risky stock on day n, and if the investor decides to invest a fraction f_n of his wealth in the risky stock on day n, with the remainder of his wealth wealth invested in the riskless security, then his fortune evolves as

$$V_n = V_{n-1}\left[1 + r(1 - f_n) + f_n Z_n\right] = V_0 \prod_{i=1}^{n} \left[1 + r(1 - f_i) + f_i Z_i\right] .$$

Thus, all our results, while stated mostly in the picturesque language of gambling, are in fact equally applicable to investment problems. The Kelly criterion in this context (usually referred to as the *optimal-growth* criterion) was studied in discrete-time in Latane (1959) and Hakansson (1970), and in continuous-time in Merton (1990) and Karatzas (1989). An adaptive portfolio strategy that performs asymptotically as well as the best constant proportion strategy, for an arbitrary sequence of gambles, was introduced in Cover (1991) (see also Cover and Gluss (1986)) in discrete-time and was extended to continuous-time in Jamishidian (1992).

In this paper, we consider the *Bayesian* version of both the discrete and continuous-time gambling and investment problems; where certain parameters of the distribution of the increment of the underlying stochastic process are unobserved random variables. As the underlying stochastic process evolves, the investor observes the outcomes and thus obtains information as to the true value of the stochastic parameters. The approach taken in this paper is to first solve the discrete-time problem, and then use those results to solve the continuous-time problem by treating it as the (diffusion) limit of the discrete-time problem.

We study first the discrete-time problem. What we would really like to do is treat the case where Z_n has a general distribution, as well as allow some weak dependence in the sequence $\{Z_n\}$. However, while the optimal policy for the general case turns out to be relatively easy to characterize (see (25) below), explicit computation of the policy is even more difficult than for the non-Bayesian case. There is one case where the policy can be explicitly calculated, and that is the simple random walk with a random success probability, which is the generalization of the original Kelly (1956) problem. For this case, the optimal (Bayesian) policy turns out to be the *certainty equivalent* of its deterministic counterpart for the ordinary simple random walk, whereby the expected value of the increment is replaced by the conditional expected value. Since the explicit computation of the optimal policy for more general random walks appears to be intractable, it is of interest therefore to develop approximations to the optimal policy. One way to approximate the optimal policy is to approximate the random walk by a continuous process that follows a stochastic differential equation. That is the approach taken here, where we propose approximating the general discrete-time Bayesian problem by a continuous-time Bayesian problem involving Brownian motion. However, it is desirable for the approximation to be based on a precise limit theorem. We provide a such limit theorem (Theorem 2 below) for the simple random walk with random success probability, also known as a random walk in a random environment (RWIRE). This diffusion approximation, besides allowing explicit computation of the optimal policy, also allows us to evaluate and compare the performance of the optimal Bayesian policy with the non-Bayesian case. Specifically, it allows us to give an *explicit quantitative evaluation of the financial value of randomness, the financial gain of perfect information and the financial cost of learning needed in the Bayesian problem.*

In the non-Bayesian case, the continuous-time problem was treated directly in the papers cited above, independently of the discrete-time problem. By the central limit theorem, the diffusion limit of the random walk with arbitrary distributions is Brownian motion, so that the continuous-time model implicitly generates approximations for random walks whose increments have a general distribution. Indeed we contend that the continuous-time results should properly be viewed as corollaries of the earlier discrete-time results. We in fact show below (in Section 2) how to obtain the continuous-time results from the corresponding discrete-time results, although we are not yet able to give a complete proof of optimality via this route. The general idea is to apply arguments such as in Chapter 10 of Kushner and Dupuis (1992), but this step remains open.

However, our primary concern here is the continuous-time Bayesian problem. The first issue is to properly formulate an appropriate continuous-time Bayesian problem. We propose *Brownian motion with known diffusion coefficient and unknown random drift coefficient.* As we show below, this model is in fact the diffusion limit of the RWIRE. In the context of the general sequence of gambles $\{Z_n\}$ of interest in applications we thus assume that Z_n has known variance σ^2, but *unknown random mean.* (In general though, it is much more complicated to precisely represent our uncertainty when the random walk is not simple.) This is consistent with current continuous-time models in financial economics, which assumes that stock prices evolve as a stochastic integral

involving Brownian motion. The quadratic variation (and hence the diffusion coefficient) of such a process can be estimated precisely from the sample path, but not the drift. Thus it seems reasonable to assume in applications that the diffusion coefficient is known, but not the drift. Our analysis of the continuous-time problem supports using the optimal policy for simple random walks with random success probabilities as an approximation for more general sequences of gambles. The reason for this is that it turns out that the optimality of the certainty equivalent (i.e., using the structure of the deterministic policy with the unknown parameter replaced by the posterior expected value of the parameter) carries forth to the continuous-time case as well.

A major thrust of this paper is showing how the discrete case goes to the continuous case in the Bayesian problem. While it turns out that the continuous-time Bayesian problem has a relatively simple direct solution by a martingale argument, as we show in Section 5, we are primarily interested in approaching the continuous-time Bayesian problem as a limit of discrete-time Bayesian problems. The Bayesian setting is substantially harder than the non-Bayesian setting, so that it should come as no surprise that our results are incomplete. Nevertheless, we do establish limit theorems that provide additional support for considering the particular continuous-time Bayesian problem we do and for using the natural extension of the Bayesian policy for simple random walks. To do this, we prove that random walks in a random environment (RWIRE's) converge weakly to a Keifer process or tied-down Brownian sheet, with two-dimensional argument. Moreover, under a proper normalization, we prove that the RWIREs converge weakly to a Brownian motion (BM) with a random drift, which is still a diffusion process. Our most explicit results are for the natural conjugate case, where the success probability in the random walk has a beta distribution. In this case, the random drift in the resulting diffusion limit has a normal distribution, which is the natural conjugate prior for the Brownian motion. We also prove that such a diffusion has the interesting property of being distributionally equivalent to a rescaled Brownian motion under a *deterministic* time change. This limit theorem is of independent interest since the beta mixed random walk is used quite often in modeling various physical and economic phenomena. These limit theorems allow us to determine the appropriate continuous-time control problem and its optimal policy.

The remainder of the paper is organized as follows: In section 2, we review the theory for non-Bayesian proportional gambling, relating discrete-time proportional betting to a continuous-time control problem via classical weak convergence results for simple random walks. We then quantify the notion of the financial value of randomness. For completeness, we also review connections to both discrete-time and continuous-time financial portfolio theory. In Section 3 we establish the discrete-time Bayesian proportional gambling results. We establish limits to continuous-time models in Section 4. In Section 5 we determine the optimal control for the continuous-time problem and compute the financial cost of learning. In Section 6 we examine the case of *power utilities*, where the objective is to maximize a fractional power of terminal wealth. Bellman and Kalaba (1957) proved that this is the most general form of a utility function that admits an optimal betting strategy that is a fixed proportion. For the discrete-time version of this problem, we show

that the optimal strategy in the Bayesian case is *not the certainty equivalent* of the corresponding result for the case of a fixed probability of success. We have yet to determine a continuous-time limit.

2 The Non-Bayesian Case

In this section, we give background on the non-Bayesian case. We first review the simple arguments yielding the results of Bellman and Kalaba (1957) for maximizing the expected logarithm of terminal wealth in discrete-time, since we later extend this to the case of random parameters. We also relate their result via a weak-convergence argument to the continuous-time result which was obtained independently via classical Hamilton-Jacobi-Bellman (HJB) methods.

2.1 Maximizing the Logarithm of Terminal Wealth

Let N denote a fixed terminal time, and suppose you wish to gamble on the outcomes of the increments of a random walk in such a manner as to maximize $E(\ln V_N)$, where V_j is your fortune at time j, and satisfies (1). While Bellman and Kalaba (1957) used a dynamic programming argument to solve this problem, a simpler argument (see Breiman (1961) or Algoet and Cover (1988)) is presented here.

If we let $F_N(x)$ denote the maximal expected value of $\ln V_N$, with $V_0 = x$, then by linearity of expectations it follows that we may write

$$F_N(x) = \max_{f_1,\ldots,f_N} E \ln \left(x \prod_{i=1}^{N} (1 + f_i Z_i) \right) = \ln(x) + \sum_{i=1}^{N} \max_{f_i} E \left(\ln(1 + f_i Z_i) \right) ,$$

from which it is clear by inspection that a myopic strategy is optimal. Furthermore, since the increments are iid, it follows that $f_i^* = f^*$ for $i = 1, \ldots, N$, where

$$f^* = \arg \max_f E \ln(1 + f Z_1)$$

and that

$$F_N(x) = \ln(x) + N E \ln(1 + f^* Z_1) . \tag{4}$$

Here we consider only the case where the random walk is *simple*. Thinking about the rescaling needed to approach continuous-time gambling, we will let the step size of the random walk be $\pm\Delta$ instead of ± 1, thus, $P(Z_i = \Delta) = \theta = 1 - P(Z_i = -\Delta)$. For this case we have

$$E \ln(1 + f Z_1) = \theta \ln(1 + f\Delta) + (1 - \theta) \ln(1 - f\Delta)$$

from which a simple computation shows that

$$f^* = \frac{2\theta - 1}{\Delta} , \tag{5}$$

and then $E \ln(1 + f^* Z_1) = \Lambda(\theta)$ where $\Lambda(\theta)$ is defined earlier in (3). Placing this into (4) shows that for this case we have

$$F_N(x) = \ln x + N\Lambda(\theta) . \tag{6}$$

2.2 The Continuous-Time Analog

In the continuous-time analog of the Kelly problem (see Pestien and Sudderth (1985), Heath et al. (1987), Ethier (1988)), your fortune evolves as the controlled stochastic differential equation

$$dV_t = f_t V_t \left(\mu\, dt + \sigma\, dW_t\right) , \tag{7}$$

where f_t is an admissible control, μ and σ are given, positive constants, and W_t is an independent standard Brownian motion. The objective is to maximize $E(\ln V_T)$, for a fixed deadline T. This control problem can be solved directly, independently of the discrete-time results, by using the HJB equations of stochastic control, which in this case reduces to solving a second order nonlinear partial differential equation. However, as a prelude to our Bayesian analysis, we use the discrete-time results to find the solution to the continuous-time problem. To do this, first realize that the diffusion governed by the stochastic differential equation, $dX_t = \mu\, dt + \sigma dW_t$, arises as the diffusion limit of the simple random walk with constant probability θ, when we rescale time and space appropriately, and send θ to $1/2$ in the appropriate way.

Specifically, for each $n \geq 1$, let $\{\xi_i^n : i \geq 1\}$ denote a sequence of i.i.d. random variables with $P(\xi_i^n = \delta_n) = \theta_n = 1 - P(\xi_i^n = -\delta_n)$, where the step size, and success probability are, respectively,

$$\delta_n := \frac{\sigma}{\sqrt{n}}, \quad \text{and} \qquad \theta_n := \frac{1}{2} + \frac{\mu}{2\sigma\sqrt{n}}.$$

Let $X_m^n = \sum_{i=1}^m \xi_i^n$ denote the random walk associated with the nth sequence. It is well known that the sequence of random walks converge weakly to a (μ, σ^2)-Brownian motion, i.e.,

$$X_{[nt]}^n \Rightarrow \mu t + \sigma W_t \quad \text{as } n \to \infty,$$

where W_t is a standard Brownian motion, and $\Rightarrow$ denotes (throughout) weak convergence of processes, as described in Billingsley (1968). Here we are only considering the case where the increments have positive expectation ($\theta_n > 1/2$), so we have $\mu > 0$. To connect the diffusion control problem with the discrete-time result described previously, note that for the nth random walk, the optimal Kelly fraction, by (5), is simply

$$f_n^* = \frac{2\theta_n - 1}{\delta_n} \equiv \frac{\mu}{\sigma^2} \quad \text{for all } n\ . \tag{8}$$

The invariance principle suggests that f_n^* should be the optimal control for the diffusion control problem that occurs in the limit as $n \to \infty$, and after doing the calculations from the HJB equations, we do find that $f_t^* = \frac{\mu}{\sigma^2}$ for all t. The optimal value function in this case also follows directly from the corresponding limiting result for the random walk. By (6), the optimal value of the objective function at a terminal time $[nT]$ for the nth random walk is

$$F_{[nT]}^n(x) = \ln x + [nT]\frac{1}{2}\left[\left(1 + \frac{\mu}{\sigma\sqrt{n}}\right)\ln\left(1 + \frac{\mu}{\sigma\sqrt{n}}\right) + \left(1 - \frac{\mu}{\sigma\sqrt{n}}\right)\ln\left(1 - \frac{\mu}{\sigma\sqrt{n}}\right)\right]. \tag{9}$$

Then, using the expansion (valid for all $0 < z < 2$) $\ln z = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}(z-1)^i}{i}$, we find that

$$(1+y)\ln(1+y) + (1-y)\ln(1-y) = y^2 + o(y^2),$$

which with (9) yields

$$\lim_{n\to\infty} F^n_{[nT]}(x) = F_T(x) = \ln x + \frac{T}{2}\frac{\mu^2}{\sigma^2}, \tag{10}$$

corresponding with the result obtained from the HJB equations.

Pestien and Sudderth (1985) proved that this policy for maximizing the expected log of terminal wealth in fact also minimizes the expected time until a given level of wealth is reached, thus extending the discrete-time asymptotic results of Breiman (1961), (see also Heath et al.(1987)).

We also note that with constant proportional gambling the stochastic difference equation (1) (with $f_j = f$) converges to the stochastic differential equation (7) with $f_t = f$, whose solution is a geometric Brownian motion with drift coefficient $f\mu x$ and variance coefficient $f^2\sigma^2x^2$, i.e.,

$$V_t = V_0 \exp\left\{\left(f\mu - \frac{f^2\sigma^2}{2}\right)t + f\sigma W_t\right\},$$

see Example 3.6 of Kurtz and Protter (1991).

Since

$$E\ln V_t = E\ln V_0 + \left(f\mu - \frac{f^2\sigma^2}{2}\right)t + f\sigma E W_t\,, \tag{11}$$

we see that it suffices to maximize $\left(f\mu - \frac{f^2\sigma^2}{2}\right)$, which is done by $f^* \equiv \mu/\sigma^2$. Since the optimality of f^* in (11) holds for all V_0 and all t, it is intuitively clear that f^* should be optimal among all admissible continuous-time controls. The optimal value function (10) can then be obtained from (11).

In the sequel, we will consider cases in which the underlying stochastic process has random parameters which the gambler obtains information about as the gambling proceeds. In discrete-time, we will randomize the success probability of the simple random walk, while in continuous-time we will randomize over the drift coefficient. One of our goals is to quantify how much this learning costs. Therefore, before we proceed, we consider the question of whether any randomness in the underlying model is beneficial or not.

2.3 The Financial Value of Randomness

Suppose now that the gambler is offered the chance to choose between the following two scenarios: gambling on the respective stochastic process with given *constant* parameters, or allowing the values of the parameters to be randomized from an arbitrary distribution with the mean given by the appropriate constant, and then be told the *actual value of the random variable*. We will refer to the first instance as *constant* gambling, and to the second as *perfect information*. To be completely general, we allow the gambler to borrow an unlimited amount of money and to bet

against the games as well. This allows us to consider any $0 < \theta < 1$ in discrete-time as well as any $-\infty < \mu < \infty$ in continuous-time, in that we then allow $-\infty < f^* < \infty$ in both cases.

In both discrete and continuous-time, *the gambler should always choose to randomize.* In discrete-time this follows from (6), i.e., suppose that Θ is a random variable with support (0,1), with $E(\Theta) = \theta$, then since $\Lambda(\theta)$ is a *convex* function of θ (see (3)), Jensen's inequality gives $E(\Lambda(\Theta)) \geq \Lambda(\theta)$, which implies that the value function under the randomization with perfect information is at least that of the value function with constant parameters. The expected financial gain from the randomization is clearly $N\left[E(\Lambda(\Theta)) - \Lambda(\theta)\right]$, which for general distributions is quite difficult to compute.

The situation is simpler in continuous-time. Here the gambler is offered the chance to randomizing his drift coefficient (μ) and then playing with the new (random) value obtained – which will be told to him immediately after the "random draw". In this case, the gambler (who is trying to maximize the expected logarithm of his terminal wealth) should choose to randomize the drift. To quantify how much the gambler could gain from this randomization, suppose Z is a random variable with mean μ and variance c: Under perfect information, the gambler's optimal control, conditional upon the information $Z = z$, is $f^* = z/\sigma^2$ by (8), and therefore by (10) the conditional optimal fortune is $\ln x + Tz^2/(2\sigma^2)$. Clearly then, if we let $F_T^P(x)$ denote the (apriori) optimal value function under the randomization, we have

$$F_T^P(x) = \ln x + \frac{T}{2\sigma^2} E(Z^2) = \ln x + \frac{T}{2\sigma^2}(\mu^2 + c) \tag{12}$$

and therefore the *expected gain from the randomization* is clearly $F_T^P(x) - F_T(x) = cT/(2\sigma^2)$.

Since it is the continuous-time case that allows an explicit evaluation of the financial gain, we will wait until we consider the Bayesian version of the continuous-time problem in Section 5 to compare these results with the cost of learning.

2.4 Connections with Portfolio Theory

Suppose now that the investor has a choice at each gamble of splitting his wager between the risky bet described above (with step size Δ) and a *sure* bet (e.g., a bond), which has a fixed return, say r, per unit time. This is essentially the discrete-time portfolio problem considered by Hakansson (1970) and many others.

It is straightforward to see that in this case the fortune evolves as

$$V_j = V_{j-1}\left(1 + r(1 - f_j) + f_j Z_j\right) ,$$

where f_j is the proportion of the fortune invested in the risky stock on the j-th bet.

An argument similar to that used above then shows that a constant proportional strategy is again optimal since

$$F_N(x) \equiv \max_{f_1,\ldots,f_N} E \ln V_N = \cdots = \ln(x) + N \max_f E \ln(1 + r(1 - f) + f Z_1) .$$

Since

$$E\ln(1+r(1-f)+fZ_1) = \theta\ln\left(1+r(1-f)+f\Delta\right)+(1-\theta)\ln\left(1+r(1-f)-f\Delta\right)\,,$$

we obtain the optimizer

$$f^* = \frac{(1+r)\left[\Delta(2\theta-1)-r\right]}{\Delta^2-r^2} \tag{13}$$

and optimal value $F_N(x) = \ln x + NK$ with

$$\begin{aligned} K &= \theta\ln\left(\frac{(1+r)2\Delta\theta}{\Delta+r}\right)+(1-\theta)\ln\left(\frac{(1+r)2\Delta(1-\theta)}{\Delta-r}\right) \\ &\equiv \Lambda(\theta)+\ln(1+r)-\left(\theta\ln\left(1+r/\Delta\right)+(1-\theta)\ln\left(1-r/\Delta\right)\right)\,, \end{aligned} \tag{14}$$

for $\Lambda(\theta)$ in (3). Note that if $r=0$, then this reduces to the previous result.

The classical continuous-time portfolio problem, first introduced and studied directly in Merton (1971), can be obtained as the limit of discrete-time problems. To study the limiting case, consider the optimal portfolio strategy for maximizing the expected logarithm of terminal fortune in the nth random walk, i.e., substitute θ_n and δ_n for θ and Δ appropriately in (13), and rearrange, to obtain the optimal fraction to be invested in the risky stock

$$f^{*(n)} = \frac{(1+r)\left[\delta_n(2\theta_n-1)-r\right]}{\delta_n^2-r^2} = \frac{(1+r)(\mu-rn)}{\sigma^2-r^2n}\,. \tag{15}$$

To complete the limit, recall that in the nth random walk, n steps are being taken every unit time, hence we must replace the interest rate, r, by the rate per step time, i.e., say $r\equiv\frac{\gamma}{n}$. When this is substituted into (15) we get

$$f^{*(n)} \to f^* = \frac{\mu-\gamma}{\sigma^2} \quad \text{as } n\to\infty. \tag{16}$$

The limit of the optimal value function is then $\ln x + \lim_{\tau\to\infty} NK$ for K in (14), and $N=[nT]$. To obtain this limit, first recall that $N\Lambda \to \frac{T}{2}\frac{\mu^2}{\sigma^2}$, and note that since $\ln(1+r)=\ln(1+\gamma/n) = \gamma/n+o(n^{-2})$, clearly we have $\lim_{n\to\infty} N\ln(1+r)=\gamma T$. By (14), it remains to determine the limit of the term

$$N\left(\theta_n\ln\left(1+r/\delta_n\right)+(1-\theta_n)\ln\left(1-r/\delta_n\right)\right).$$

This is obtained by using the expansion of the logarithm given earlier, i.e.,

$$\theta_n\ln\left(1+r/\delta_n\right)+(1-\theta_n)\ln\left(1-r/\delta_n\right) = \frac{2\mu\gamma-\gamma^2}{2\sigma^2 n}+o(n^{-3/2})\,,$$

so that

$$F_{[nT]}(x) \equiv \ln x + [nT]K \to \ln x + T\left(\frac{(\mu-\gamma)^2}{2\sigma^2}+\gamma\right) \quad \text{as } n\to\infty. \tag{17}$$

The corresponding diffusion control problem which Merton (1971) first considered directly using techniques of continuous-time stochastic control, independently of the earlier discrete-time results, was that of an investor whose fortune evolves according to

$$dV_t = V_t\left(\left[(1-f_t)\gamma+f_t\mu\right]dt+f_t\sigma dW_t\right)\,, \tag{18}$$

where W_t is an ordinary Brownian motion, and f_t a suitable admissible control. In Merton's continuous-time model, the price of the risky stock, say S_t, and the price of the riskless bond, say B_t, were assumed to evolve respectively according to

$$dS_t = \mu S_t dt + \sigma S_t dW_t \ , \quad \text{and} \quad dB_t = \gamma B_t dt \, ,$$

with $\mu > \gamma$. If f_t is the fraction of the investor's wealth invested in the risky stock, the investor's wealth, V_t, then evolves as

$$dV_t = f_t V_t \frac{dS_t}{S_t} + (1 - f_t) V_t \frac{dB_t}{B_t}$$

which, upon substitution, is equivalent to (18). Among many other results, Merton (1971) showed that the control in (16) is the optimum for the problem of maximizing $E \ln V_T$ (see also Karatzas (1989)). Eq.(17) is a special case of (9.15) of Karatzas (1989), who studies the control of more general diffusions. The derivation of (17) from the limiting form of the terminal wealth in the discrete-time case is apparently new.

This completes our review of the known results for Kelly gambling and investing. In the next section, we consider the Bayesian control problem in discrete time.

3 Bayesian Gambling in Discrete time

We begin by discussing the random walk in a random environment. Then we focus on the special case of a beta prior. Then we consider the Bayesian gambling and investment problem in discrete-time.

3.1 A Random Walk in a Random Environment

By a random walk in a random environment (RWIRE), we mean a simple random walk where the success probability θ, is a random variable with a given density $f_\theta(\cdot)$, on (0,1). Let $S_0 = 0$ and $S_n = \sum_{i=1}^n Z_i$, where $P(Z_i = 1|\theta) = \theta = 1 - P(Z_i = -1|\theta)$, and $Y_n = (S_n + n)/2 \equiv \sum_{i=1}^n W_i$, where $W_i = (Z_i + 1)/2$. The posterior distribution of θ, conditioned by observing $(W_1, \ldots, W_n)$ depends only on the sufficient statistic $Y_n \equiv \sum_{i=1}^n W_i$, as can be seen from a direct application from Bayes' formula. Specifically, the posterior distribution is

$$dP(\theta \le u | Y_n = y) = \frac{\binom{n}{y} u^y (1-u)^{n-y} f_\theta(u) \, du}{\int_0^1 \binom{n}{y} u^y (1-u)^{n-y} f_\theta(u) \, du} \, , \tag{19}$$

with posterior moments

$$E(\theta^j | Y_n) = \frac{\int_0^1 u^{y+j} (1-u)^{n-y} f_\theta(u) \, du}{\int_0^1 u^y (1-u)^{n-y} f_\theta(u) \, du} \, .$$

De Finetti's theorem states that this is the only possible (joint) distribution for a set of exchangeable random variables that take on only the values 0 or 1 (e.g., p.229 of Feller (1971)). The random

walk case therefore gives the posterior distribution

$$dP(\theta \le u|S_n) = \frac{u^{(S_n+n)/2}(1-u)^{(n-S_n)/2} f_\theta(u)du}{\int_0^1 u^{(S_n+n)/2}(1-u)^{(n-S_n)/2} f_\theta(u)du},$$

so that clearly $\{S_n\}$ is a Markov process with transition probability $P(S_{n+1} = S_n+1|S_n) = E(\theta|S_n)$.

An interesting case of this RWIRE arises when the prior density, $f_\theta(u)$, is a *beta* density, because it is the natural conjugate prior, which means that the posterior distribution is once again beta.

3.2 Natural Conjugate: Beta Prior

It is easiest to first describe the beta/binomial process. Let $P(W_i = 1|\theta) = \theta = 1 - P(W_i = 0|\theta), i \ge 1$, where θ is a random variable with a beta distribution, i.e., $\theta \sim \mathsf{Be}(\alpha, \beta)$, by which we mean

$$f_\theta(u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1}(1-u)^{\beta-1} \quad \text{for } u \in (0,1), \tag{20}$$

see DeGroot (1970). The prior mean and variance of the success probability θ are

$$E(\theta) = \frac{\alpha}{\alpha+\beta}, \quad V(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)},$$

and the posterior distribution of θ is then once again a beta. By (19) and (20), $\theta|Y_n \sim \mathsf{Be}(Y_n + \alpha, n - Y_n + \beta)$, i.e.,

$$dP(\theta \le u|Y_n = y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} u^{y+\alpha-1}(1-u)^{n-y+\beta-1}\, du,$$

with posterior mean and variance

$$E(\theta|Y_n) = \frac{Y_n+\alpha}{n+\alpha+\beta}, \quad V(\theta|Y_n) = \frac{(Y_n+\alpha)(n-Y_n+\beta)}{(n+\alpha+\beta)^2(n+\alpha+\beta+1)}.$$

We can rewrite the posterior mean as the following convex combination of the prior and observed means:

$$E(\theta|Y_n) = \frac{\alpha+\beta}{n+\alpha+\beta}E(\theta) + \frac{n}{n+\alpha+\beta}\left(\frac{Y_n}{n}\right).$$

By the law of large numbers, $Y_n/n \to \theta$ w.p.1 as $n \to \infty$, so that $E(\theta|Y_n) \to \theta$ w.p.1 as $n \to \infty$.

The random walk process is now obtained simply by letting $Z_i = 2W_i - 1$. Then $S_n = \sum_{i=1}^n Z_i \equiv 2Y_n - n$ is a Markov process with one-step transition probability

$$P(S_{n+1} = S_n + 1|S_n) = E(\theta|S_n) = \frac{S_n + n + 2\alpha}{2(n+\alpha+\beta)}. \tag{21}$$

Note that the (*state and time-dependent*) mean drift of this random walk is

$$E(S_{n+1} - S_n|S_n) = E(Z_{n+1}|S_n) = 2E(\theta|S_n) - 1 \equiv \frac{S_n + \alpha - \beta}{n+\alpha+\beta}. \tag{22}$$

The conditional variance of the increment is

$$V(S_{n+1} - S_n|S_n) = V(Z_{n+1}|S_n) = \frac{(S_n+n+2\alpha)(n-S_n+2\beta)}{4(n+\alpha+\beta)^2(n+\alpha+\beta+1)}.$$

The m-step transition probability can be shown to be

$$P(S_{n+m} = z|S_n = y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma\left(\frac{y+n+2\alpha}{2}\right)\Gamma\left(\frac{n-y+2\beta}{2}\right)}\binom{m}{\left(\frac{z+y+m}{2}\right)}\frac{\Gamma\left(\frac{z+m+n+2\alpha}{2}\right)\Gamma\left(\frac{m+n-z+2\beta}{2}\right)}{\Gamma(m+n+\alpha+\beta)}.$$

Note that the unconditional mean and variance of the random walk satisfy

$$\begin{aligned} E(S_n) &\equiv E\left(E(S_n|\theta)\right) = n\left(2E(\theta)-1\right) = n\frac{\alpha-\beta}{\alpha+\beta}, \qquad (23) \\ V(S_n) &\equiv E\left(V(S_n|\theta)\right) + V\left(E(S_n|\theta)\right) \\ &= E\left(n4\theta(1-\theta)\right) + V\left(n(2\theta-1)\right) = 4nE\left(\theta(1-\theta)\right) + 4n^2V(\theta) \\ &= \frac{4n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}. \qquad (24) \end{aligned}$$

The optimal gambling policy for such a random walk is now derived by a simple extension of the previous argument.

3.3 Gambling

Our objective is to maximize $E(\ln V_N)$, over all nonanticipating strategies, where $\{Z_n\}$ is a RWIRE and our fortune evolves as $V_j = V_{j-1} + f_j V_{j-1} Z_j$. This means that if f_n is the proportion bet on day n, then f_n is adapted to $\mathcal{F}_{n-1}$, where $\mathcal{F}_j = \sigma\{Z_i : i = 1, \ldots, j\}$, i.e., f_n can depend at most on the current information and hence the (observed) values of $(Z_1, \ldots, Z_{n-1})$. At first we consider a general prior density. Since we obtain more information about the true value of the unknown random θ as play continues, we should not expect an optimal policy to be a *constant* proportion for this case. While we could incorporate this learning into a dynamic programming argument that extends that of Bellman and Kalaba (1957), it is simpler to use results of Algoet and Cover (1988) (see also Section 15.5 in Cover and Thomas (1991)) who show that in general

$$\max_{f_1,\ldots,f_N} E\ln V_N = E\sum_{i=1}^{N}\max_{f_i} E\left(\ln(1+f_iZ_i)|\mathcal{F}_{i-1}\right). \qquad (25)$$

For our Markovian model, this reduces to

$$\max_{f_1,\ldots,f_N} E\ln V_N = E\sum_{i=1}^{N}\max_{f_i} E\left(\ln(1+f_iZ_i)|S_{i-1}\right),$$

where S_j is the value of the random walk after the jth step. With a view towards the rescaling that will be needed to go to continuous-time, we will consider again the case here where the RWIRE takes steps of size $\pm\Delta$ instead of ± 1. Furthermore, since the transition probability for this random walk is both time and space dependent (see (21)), for the sequel we will denote it by $P(S_{k+1} = S_k + \Delta|S_k) = E(\theta|S_k, k) = 1 - P(S_{k+1} = S_k - \Delta|S_k)$.

To proceed, suppose the random walk has been observed already for k steps, and let

$$F_m(x, S_k, k) = \max_{f_{k+1},\ldots,f_{k+m}} E\left(\ln V_{m+k}|V_k = x, S_k, k\right).$$

We can use the cited results above to write this as

$$F_m(x, S_k, k) = \ln(x) + E \sum_{i=k}^{m+k-1} \max_{f_i} E\left(\ln(1 + f_{i+1}Z_{i+1})|S_i, i\right) .$$

Since for any i we have

$$E\left(\ln(1 + f_{i+1}Z_{i+1})|S_i, i\right) = E(\theta|S_i,i)\ln(1 + f_{i+1}\Delta) + (1-E(\theta|S_i,))\ln(1 - f_{i+1}\Delta)\,, \qquad (26)$$

it is clear that this is optimized by taking $f^*_{i+1}(S_i, i) = (2E(\theta|S_i, i) - 1)/\Delta$. This shows that for any prior, the optimal strategy is to bet the *certainty equivalent* of the deterministic counterpart (compare with f^* of (14)). (Note that we allow negative bets here, i.e., if $f^*_i < 0$, the gambler is betting against the next outcome.)

For the special case in which the prior is a beta distribution, (22) shows that this becomes

$$f^*_{i+1}(S_i, i) = \frac{1}{\Delta}\left[\frac{S_i + \alpha - \beta}{i + \alpha + \beta}\right] .$$

In general, when we put the optimal control f^*_{i+1} back into (26), we obtain

$$E\left(\ln(1 + f^*_{i+1}Z_{i+1})|S_i, i\right) \equiv \Lambda(S_i, i) = \ln 2 + E(\theta|S_i,i)\ln E(\theta|S_i,i) + (1-E(\theta|S_i,i))\ln(1-E(\theta|S_i,i))\,, \qquad (27)$$

which in the beta case reduces to

$$\Lambda(S_i, i) = \frac{S_i + i + 2\alpha}{2(i + \alpha + \beta)} \ln\left(\frac{S_i + i + 2\alpha}{i + \alpha + \beta}\right) + \frac{i - S_i + 2\beta}{2(i + \alpha + \beta)} \ln\left(\frac{i - S_i + 2\beta}{i + \alpha + \beta}\right) .$$

Noting that

$$E\left[\Lambda(S_{i+1}, i + 1)|S_i, i\right] = E(\theta|S_i, i)\Lambda(S_i + 1, i + 1) + (1 - E(\theta|S_i, i))\Lambda(S_i - 1, i + 1)\,,$$

we summarize this as follows.

Theorem 1 *Suppose the current fortune is x, and the RWIRE has been observed for k-steps, with a current value S_k and there remains another m steps remain to play. Then at each step, the optimal policy is the certainty equivalent of the deterministic counterpart (with the current posterior expected value of the random probability replacing the probability in the latter), i.e., for $j = 0, \ldots, m - 1$, the optimal policy bets $\left(2E(\theta|S_{k+j}, k + j) - 1\right)/\Delta$ percent of the fortune on the $k + j + 1$-st step.*

Furthermore, under this policy, the expected logarithm of the terminal fortune is

$$\begin{aligned} F_m(x, S_k, k) &\equiv \ln x + E\left[\sum_{i=k}^{m+k-1} \Lambda(S_i, i)|S_k, k\right] \\ &= \ln x + C_m(S_k, k)\,, \end{aligned} \qquad (28)$$

where the function $C_m(i, l)$ is the solution to the difference equation

$$C_j(i, l) = \Lambda(i, l) + E(\theta|i, l)C_{j-1}(i + 1, l + 1) + (1 - E(\theta|i, l))C_{j-1}(i - 1, l + 1)\,,\ \ i \le l, \qquad (29)$$

for $\Lambda(i, l)$ in (27) and $C_0 = 0$.

For the beta case, (29) simplifies to

$$\begin{aligned} 2(l+\alpha+\beta)C_j(i,l)+\ln(l+\alpha+\beta) &= (i+l+2\alpha)\left[C_{j-1}(i+1,l+1)+\ln(i+l+2\alpha)\right] \\ &\quad +(l-i+2\beta)\left[C_{j-1}(i-1,l+1)+\ln(l-i+2\beta)\right]. \end{aligned}$$

It is important to note that the policy obtained above also asymptotically *maximizes the asymptotic growth rate* when used over an infinite-horizon. This follows from Algoet and Cover's (1988) infinite-horizon result, for very general discrete processes, that maximizing the conditional expected log return given all the currently available information at each stage is asymptotically optimal for maximizing the asymptotic growth rate. In our Markovian model, this reduces to precisely the policy obtained. However, unlike Algoet and Cover (1988), our main focus is on characterizing and computing optimal policies *explicitly* for finite-horizons, which is simplified in the Markovian setting.

3.4 Adaptive Portfolios

The portfolio problem in discrete time for the RWIRE is also easily solved from the adaptive gambling results just obtained. Once again the investor has a choice at each step in the RWIRE of splitting his wager between the next step in the RWIRE and the sure bet (or bond) which has a fixed return of r per unit time. Analogously to the pure gambling problem, if we allow the random walk to have step size Δ, in this case we have

$$\begin{aligned} F_m(x,S_k,k) &= \max_{f_{k+1},\ldots,f_{k+m}} E\left(\ln V_{m+k}|V_k=x,S_k,k\right) \\ &= \cdots = \ln(x)+E\sum_{i=k}^{m+k-1}\max_{f_i} E\left(\ln(1+r(1-f_{i+1})+f_{i+1}Z_{i+1})|S_i,i\right). \end{aligned}$$

As before, it is fairly easy to show that the optimal control for the portfolio problem after the RWIRE has taken i steps is simply to invest the proportion $f^*(S_i,i)$ in the random walk, where (see (13))

$$f^*(S_i,i)=\frac{(1+r)\left(\Delta(2E(\theta|S_i,i)-1)-r\right)}{\Delta^2-r^2}. \tag{30}$$

Having established that the optimal control for the discrete-time case is simply the certainty equivalent of the case with a fixed constant probability, we would now like to use this to treat the continuous-time problem with random parameter values. For such processes, the techniques of continuous-time stochastic control with *partial information* become quite complicated (see e.g. Section VI.10 of Fleming and Rishel (1975)), so rather than studying the continuous-time dynamic programming optimality equations directly, we would like to employ once again a limiting argument to the discrete-time process just studied.

To employ this approach, we need a limit theorem that describes precisely how such a random walk approaches a diffusion. This is the content of the next section.

4 Continuous-Time Diffusion Limits

In this section we first establish limits for RWIREs and then we show that the limit, which is a Brownian motion with a normally distributed random mean, is equivalent to a deterministically scaled Brownian motion.

4.1 Limits for RWIREs

Recall now, that to send a simple random walk with θ being constant to a simple Brownian motion with drift, we had to rescale time and space appropriately, and send θ to $1/2$ in the appropriate way. However, if the success probability is a *random* variable, it is no longer clear how to standardize, and how to take the limit, and what the resulting limiting process should be. For example, the central limit theorem for exchangeable random variables shows that for a general density $f_\theta(\cdot)$, the normalized random walk converges to a *mixed* normal distribution, i.e.,

$$\lim_{n\to\infty} P\left(\frac{S_n - E(S_n)}{\sqrt{n}} \le z\right) = \int_0^1 \Phi\left(\frac{z}{2\sqrt{u(1-u)}}\right) f_\theta(u)\, du\,,$$

where Φ denotes the standard normal cdf. No particular simplification occurs here for the case in which f_θ is a beta density. This suggests that to get a reasonable diffusion limit, we need to in fact take double limits, by allowing the distribution of θ to vary in the appropriate manner with n as well. We make this precise for the beta case in the following theorem, which we will prove in this section and make ample use of later.

Theorem 2 *Let $\{Z_i^n\}$ denote a sequence of i.i.d. random variables, with*

$$P(Z_i^n = 1|\theta_n) = \theta_n = 1 - P(Z_i^n = -1|\theta_n)\,,$$

where $\theta_n \sim \mathsf{Be}(\alpha_n, \beta_n)$ and the parameters α_n, β_n are given by

$$\alpha_n := \frac{\sigma^2 n - (\mu^2 + c)}{2c\sqrt{n}}\left(\sqrt{n} + \frac{\mu}{\sigma}\right), \tag{31}$$

$$\beta_n := \frac{\sigma^2 n - (\mu^2 + c)}{2c\sqrt{n}}\left(\sqrt{n} - \frac{\mu}{\sigma}\right). \tag{32}$$

Let $S_m^n = \sum_{i=1}^m Z_i^n$. Then

$$\frac{\sigma S_{[nt]}^n}{\sqrt{n}} \Rightarrow X_t := (\sigma^2 + ct) W_{\frac{t}{\sigma^2+ct}} + \mu t\,, \tag{33}$$

where W_s is a standard Brownian motion.

Note that under the parameterization above, the prior mean satisfies $E(\theta_n) \equiv \frac{\alpha_n}{\alpha_n+\beta_n} = 1/2 + \mu/2\sigma\sqrt{n} \to 1/2$, while the prior variance satisfies $V(\theta_n) \equiv \frac{\alpha_n\beta_n}{(\alpha_n+\beta_n)^2(\alpha_n+\beta_n+1)} = c/4\sigma^2 n \to 0$. Furthermore, the unconditional mean and variance of an increment in the n-th random walk satisfies $E(Z_i^n) = \mu/\sigma\sqrt{n}$, and $V(Z_i^n) = 1 - \frac{\mu^2}{4\sigma^2 n}$, which are the same as the mean and variance for the n-th random walk with a *constant* probability, (see (23) and (24)).

At first glance, the limit obtained in (33) does not appear intuitive at all, since we expected a Brownian motion with a possibly random drift and diffusion coefficient, whereas we ended up with a rescaled time-changed Brownian motion without any randomization. However, we will show that, in fact, the limiting process in this case, X_t is equivalent to a Brownian Motion with a *normally* distributed drift term, i.e.,

$$X_t = Zt + \sigma W_t \,,$$

where $Z \sim N(\mu, c)$, and W_t is an independent Brownian motion.

First, however, we describe the limiting behavior of an arbitrary RWIRE (with arbitrary mixing distribution). To get started, we need some preliminary results and definitions:

Definition 1. A "tied down Brownian Sheet", or a *Kiefer Process*, is a continuous two-parameter process, $\{B(t,x) : 0 \le t < \infty\,, 0 \le x \le 1\}$ defined by (see Csörgö and Révész (1981), page 80)

$$B(t,x) = W(t,x) - xW(t,1)\,, \tag{34}$$

where $\{W(t,x) : t \ge 0, x \ge 0\}$ is the *Brownian Sheet* (see Révész (1990), Chapter 12, for a complete description of the Brownian Sheet).

The property of the Kiefer process that we need is the following:

For any fixed $0 < x_0 < 1$, the process $W_t^1 := \frac{B(t,x_0)}{\sqrt{x_0(1-x_0)}}\,, t \ge 0$, is a standard Brownian Motion.

The main utility of the Kiefer process for our problem comes from the following well known result.

Theorem 3 (Kiefer) *Suppose that $\{U_n : n \ge 1\}$ is a sequence of i.i.d. uniform $(0,1)$ random variables. Define the process that counts the number of uniforms below x in the first nt observations by*

$$F_n(t,x) := \sum_{i=1}^{[nt]} \mathbf{1}_{\{U_i \le x\}}\,.$$

Then

$$\frac{F_n(t,x) - ntx}{\sqrt{n}} \Rightarrow B(t,x) \quad \text{as } n \to \infty,$$

where $B(t,x)$ is the tied down Brownian sheet (or Kiefer Process) in (34).

As noted above, for a fixed x, $B(t,x)$ is distributionally equivalent to a Brownian motion with variance parameter $x(1-x)$, which implies that for a fixed x, the process $\{B(t,x) : t \ge 0\}$ is statistically identical to a rescaled simple Brownian motion, $\sqrt{x(1-x)}W_t$, i.e.,

$$\{B(t,x) \,: t \ge 0\} \stackrel{\mathrm{d}}{=} \left\{\sqrt{x(1-x)}\, W_t \,: t \ge 0\right\} \qquad \text{for each } x\,. \tag{35}$$

We now establish one more preliminary result.

Theorem 4 *Suppose now that* $\{X_n : n \geq 1\}$ *is a sequence of random variables taking on values in* $[0,1]$*, with*

$$\sqrt{n}(X_n - x) \Rightarrow L\,. \tag{36}$$

If the sequence $\{X_n\}$ *is independent of the sequence of uniforms,* $\{U_n\}$*, then*

$$\frac{F_n(t, X_n) - ntx}{\sqrt{n}} \Rightarrow B(t,x) + tL\,.$$

Proof: Since the sequences are independent, Theorems 3.2 and 4.4 of Billingsley (1968) imply the joint convergence

$$\left(\frac{F_n(t,x) - ntx}{\sqrt{n}},\, X_n,\, \sqrt{n}(X_n - x)\right) \Rightarrow (B(t,x),\, x,\, L)\ .$$

The continuous mapping theorem (Theorem 5.1 of Billingsley (1968)) applied to the mapping

$$g\left(f(t,x), y, z\right) = f(t,y) + tz$$

gives therefore

$$\frac{F_n(t,X_n) - ntx}{\sqrt{n}} = \frac{F_n(t,X_n) - ntX_n}{\sqrt{n}} + t\sqrt{n}(X_n - x) \Rightarrow B(t,x) + tL\,. \qquad \blacksquare$$

To see the connection to the gambling problem, note that for a fixed x, the random variable $\mathbf{1}_{\{U_i \leq x\}}$, is distributionally equivalent to a Bernoulli random variable with parameter x. Therefore, $F_n(t,x) - (nt - F_n(t,x)) \equiv 2F_n(t,x) - nt$, is distributionally equivalent to the position of a simple random walk with success probability x, i.e.,

$$2F_n(t,x) - nt \stackrel{\mathrm{d}}{=} \sum_{i=1}^{[nt]} \xi_i$$

where $\{\xi_i : i \geq 1\}$ is a sequence of i.i.d. random variables, with $P(\xi_i = 1) = x = 1 - P(\xi_i = -1)$. Therefore, if for each n, $\{\xi_i^n\}$ is an independent sequence where ξ_i^n take on the values $+1, -1$, with respective *random* probability $X_n, 1 - X_n$, it is clear that the position of this random walk after $[nt]$ steps satisfies

$$\sum_{i=1}^{[nt]} \xi_i^n \equiv S_{[nt]}^n \stackrel{\mathrm{d}}{=} 2F_n(t,X_n) - nt\,. \tag{37}$$

This distributional equality is all we need to obtain the limiting behavior of a RWIRE under the convergence condition (36).

Theorem 5 *Suppose* $\{\xi_i^n : i \geq 1, n \geq 1\}$ *are a sequence of independent random variables with* $P(\xi_i^n = 1) = X_n = 1 - P(\xi_i^n = -1)$ *for each* n*, with* $\sqrt{n}(X_n - x) \Rightarrow L$*. Then, if we let* $S_m^n = \sum_{i=1}^m \xi_i^n$*, we have*

$$\frac{S_{[nt]}^n - nt(2x-1)}{\sqrt{n}} \Rightarrow 2B(t,x) + 2tL \quad \text{as } n \to \infty\,.$$

Proof: Subtracting $nt(2x-1)$ from both sides of (37), and dividing by $\sqrt{n}$, gives

$$\frac{S^n_{[nt]} - nt(2x-1)}{\sqrt{n}} \stackrel{\mathrm{d}}{=} \frac{2F_n(t, X_n) - 2ntx}{\sqrt{n}} \Rightarrow 2\left[B(t,x) + tL\right] ,$$

by the previous Theorem 4. ∎

We now relate the limit in Theorem 5 to standard Brownian motion, W_t, under the condition that the limiting mean of the random probabilities is 1/2.

Corollary 1 *If $x = \frac{1}{2}$ in the setting of Theorem 5, then*

$$\frac{S^n_{[nt]}}{\sqrt{n}} \Rightarrow W_t + 2tL\,,$$

where W_t denotes a standard Brownian motion.

Proof: Apply the previous Theorem 5, with $x = 1/2$, to see that in this case we have

$$\frac{S^n_{[nt]}}{\sqrt{n}} \Rightarrow 2B\left(t, \frac{1}{2}\right) + 2tL\,.$$

But since $B(t,x) \stackrel{\mathrm{d}}{=} \sqrt{x(1-x)}\, W_t$, for each x, $0 \le x \le 1$, it is clear that $2B\left(t, \frac{1}{2}\right) \stackrel{\mathrm{d}}{=} W_t$. ∎

The case in which the success probability has a beta prior, as in Theorem 2, now follows directly.

Corollary 2 *In the setting of Theorem 5, if $X_n \sim \mathsf{Be}(\alpha_n, \beta_n)$, where α_n, β_n are given in (31) and (32), then (36) holds, and thus Corollary 1 holds, with L having a* normal *distribution, specifically, $L \sim N\left(\frac{\mu}{2\sigma}, \frac{c}{4\sigma^2}\right)$.*

Proof: First note that under (31) and (32), the mean and variance of X_n satisfy

$$\begin{aligned} E(X_n) \qquad &= \tfrac{\alpha_n}{\alpha_n+\beta_n} \qquad &&\equiv \frac{1}{2} + \frac{\mu}{2\sigma\sqrt{n}}\,, \\ V(X_n) \quad &= \tfrac{\alpha_n\beta_n}{(\alpha_n+\beta_n)^2(\alpha_n+\beta_n+1)} \quad &&\equiv \frac{c}{4\sigma^2 n}\,, \end{aligned}$$

so that

$$E\left(\sqrt{n}\left(X_n - \frac{1}{2}\right)\right) = \frac{\mu}{2\sigma}\,, \;\; V\left(\sqrt{n}\left(X_n - \frac{1}{2}\right)\right) = \frac{c}{4\sigma^2}.$$

Since the beta distribution converges to the normal when $\alpha_n = k_1 + an + o(n), \beta_n = k_2 + bn + o(n)$, we have $\sqrt{n}\left(X_n - \frac{1}{2}\right) \Rightarrow L$, where $L \sim N\left(\frac{\mu}{2\sigma}, \frac{c}{4\sigma^2}\right)$. ∎

Since in this case, $2L \sim N\left(\frac{\mu}{\sigma}, \frac{c}{\sigma^2}\right)$, it now follows directly, that for the random walk displayed earlier in Theorem 2, with $X_n \equiv \theta_n$ for all $n \ge 1$, we have

$$\frac{\sigma S^n_{[nt]}}{\sqrt{n}} \Rightarrow \sigma W_t + tZ\,, \qquad \text{where } \; Z \sim N(\mu, c).$$

So the diffusion limit of the mixed, or weighted, random walk described earlier is in fact a Brownian motion with a random mean. The fact that the distribution of the random drift term, Z, is normally

distributed, is a consequence of the parameterization and the convergence of the beta distribution to the normal. However, it is important to note that the beta distribution is the natural conjugate prior for the success probability of the Bernoulli random variable, while the natural conjugate prior for the mean of a normal random variable is once again a normal. Thus under the parameterization given above, the limit shows that we go from one natural conjugate pair (beta-binomial) to another (normal-normal).

To complete the proof of Theorem 2, it now only remains to prove that

$$\sigma W_t + tZ \stackrel{\mathrm{d}}{=} (\sigma^2 + ct) W_{\frac{t}{\sigma^2+ct}} + \mu t \,. \tag{38}$$

There are a few ways to show this, for example, one way would be to recognize that both sides of (38) are Gaussian processes, and hence we would then only need to evaluate the means and covariances for each side and show that they are equal. Here, we prefer to take a constructive approach: we will first analyze the process $\sigma W_t + tZ$, and then derive the righthand side of (38) from first principles.

4.2 Brownian Motion with a Random Mean

Suppose we have a diffusion, X_t that follows the stochastic differential equation

$$dX_t = Z dt + \sigma \, dW_t \tag{39}$$

where W_t is a standard Brownian motion, and Z is an independent random variable with c.d.f. $G(z)$. Let $\mathcal{F}_t^X = \sigma\{X_s : 0 \le s \le t\}$ denote the filtration generated by the diffusion in eq.(39). The *filtering theorem* of Fujisaki, Kallianpur and Kunita (see e.g. Kallianpur (1980)), states that X_t has a representation as the diffusion

$$dX_t = E(Z|\mathcal{F}_t^X)\, dt + \sigma \, d\hat{W}_t \,, \tag{40}$$

where $\hat{W}_t$ is another, independent standard Brownian motion (the *innovations process*). Since $X_t|Z \sim N(Zt, \sigma^2 t)$, Bayes' formula shows that the posterior distribution for Z, given the filtration $\mathcal{F}_t^X$, can be written as

$$dP(Z \le z|\mathcal{F}_t^X) = \frac{\exp\left(-\frac{1}{2\sigma^2 t}(X_t - zt)^2\right) \, dG(z)}{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2 t}(X_t - zt)^2\right) \, dG(z)} \,,$$

and the posterior mean is

$$E(Z|\mathcal{F}_t^X) = \frac{\int_{-\infty}^{\infty} z \exp\left(-\frac{1}{2\sigma^2 t}(X_t - zt)^2\right) \, dG(z)}{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2 t}(X_t - zt)^2\right) \, dG(z)} \,. \tag{41}$$

If $Z \sim N(\mu, c)$ (which is the natural conjugate prior), then we can show that

$$Z|\mathcal{F}_t^X \sim N\left(\frac{cX_t + \sigma^2\mu}{\sigma^2 + ct}, \frac{\sigma^2 c}{\sigma^2 + ct}\right) \,. \tag{42}$$

Therefore, in this case, the filtering theorem implies that X_t is equivalent to the diffusion determined by the stochastic differential equation

$$dX_t = \frac{cX_t + \sigma^2\mu}{\sigma^2 + ct}dt + \sigma\, d\hat{W}_t\,, \tag{43}$$

where $\hat{W}_t$ is a standard Brownian motion. This is a linear stochastic differential equation, which has the strong solution

$$X_t \equiv \mu t + (\sigma^2 + ct)\left[X_0 + \int_0^t \frac{\sigma}{\sigma^2 + cs} d\hat{W}_s\right]\,, \tag{44}$$

(see Karatzas and Shreve, (1988) section 5.6). The mean and covariance functions for this process are

$$\begin{aligned} m(t) &:= E(X_t) = \mu t + (\sigma^2 + ct)E(X_0) \\ \rho(s,t) &:= E(X_t - m(t))(X_s - m(s)) = (\sigma^2 + cs)(\sigma^2 + ct)\left[\rho(0,0) + \frac{s \wedge t}{\sigma^2 + c(s \wedge t)}\right]. \end{aligned}$$

We of course have $\rho(t,t) := V(t) = (\sigma^2 + ct)[t + V(0)(\sigma^2 + ct)]$.

If X_0 is constant, and $c = 0$, then it is clear that X_t is an ordinary Brownian motion with drift μ and diffusion parameter σ. If not, then if either X_0 is constant, or normally distributed, X_t is a Gaussian process. *For notational ease we will, without any loss of generality, always take* $X_0 = 0$ *for the remainder of the paper.*

We now show that the randomized diffusion displayed above has a representation as a simple *time changed* Brownian motion, without any randomization. This will complete the proof of Theorem 2.

Proposition 1 *The diffusion process* X_t *in (44) has the representation*

$$X_t \stackrel{\mathrm{d}}{=} (\sigma^2 + ct)\tilde{W}_{\frac{t}{\sigma^2+ct}} + \mu t$$

where $\{\tilde{W}_s\}$ *denotes an independent Brownian motion.*

Proof: Suppose we have a diffusion Y_t obtained from a standard Brownian motion by the following transformation:

$$Y_t = a(t)W_{g(t)} + f(t)$$

where $a(t), g(t), f(t)$, are all continuous differentiable functions, with $g(t) \geq 0$, $b(0) = f(0) = 0$. Ito's formula applied to the above shows that

$$dY_t = [a'(t)W_{g(t)} + f'(t)]dt + a(t)dW_{g(t)}.$$

Since $W_{g(t)} = \frac{Y_t - f(t)}{a(t)}$, and since $dW_{g(t)} = \sqrt{g'(t)}\, dU_t$, where U_t is another independent standard Brownian motion, we can rewrite this last equation as

$$dY_t = \frac{a'(t)Y_t + a(t)f'(t) - a'(t)f(t)}{a(t)}\, dt + a(t)\sqrt{g'(t)}\, dU_t\,. \tag{45}$$

If we equate the drifts and diffusion parameters of the two stochastic differential equations (45) and (43), we find

$$\frac{a'(t)Y_t + a(t)f'(t) - a'(t)f(t)}{a(t)} = \frac{cX_t + \sigma^2\mu}{\sigma^2 + ct}\,, \tag{46}$$

$$a(t)\sqrt{g'(t)} = \sigma\,. \tag{47}$$

If we now set $Y_t = X_t$, and then equate the coefficients in (46), we get

$$a(t) = \sigma^2 + ct \quad , a'(t) = c\,, \tag{48}$$

as well as

$$a(t)f'(t) - a'(t)f(t) = \sigma^2\mu\,. \tag{49}$$

When we put (48) into (49), we get the linear ordinary differential equation

$$f'(t) = \frac{c}{\sigma^2 + ct}f(t) + \frac{\sigma^2\mu}{\sigma^2 + ct} \;\; ; \; f(0) = 0\,.$$

The solution to this is simply $f(t) = \mu t$. Eq.(47) now shows that

$$g'(t) = \sigma^2\frac{1}{a(t)^2} \equiv \sigma^2\left(\frac{1}{\sigma^2 + ct}\right)^2\,,$$

which is easily solved to yield $g(t) = \frac{t}{\sigma^2 + ct}$, and so we've established the proposition. ∎

Now, having proved Theorem 2, we can relate the discrete-time Bayesian control problem to the continuous-time Bayesian control problem.

5 Continuous-Time Bayesian Gambling

We now apply the discrete-time results in section 3 to treat the continuous-time Bayesian gambling problem, much as we did for the non-Bayesian case in section 2. We first consider gambling and the financial value of information, then the portfolio problem, and finally we present a final martingale argument.

5.1 Gambling

Now consider the continuous-time optimal control problem in which your fortune evolves according to

$$dV_t = f_t V_t\, dX_t\,, \tag{50}$$

where, as before, f_t is an admissible control, but now X_t is the Brownian motion with a random drift term, described above. Thus V_t is not a Markov process, although the two dimensional process (V_t, X_t) is. Here we will consider the case in which the prior distribution on the drift is a normal with mean μ, and variance c, which was analyzed previously. What is the optimal policy to maximize $E\ln V_T$? The answer does not seem evident from the HJB approach, but it seems intuitively clear that the optimal control should be the *certainty equivalent* of the deterministic drift case. This result can be obtained from the discrete-time results in section 3.

Theorem 6 *The optimal control to maximize the (conditional) expected log of terminal fortune, for any terminal time is to invest at time t the fraction*

$$f_t^* = \frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right) \tag{51}$$

which is simply the posterior mean drift of the underlying diffusion, X_t, divided by its diffusion coefficient, i.e., the certainty equivalent of (8).

Proof: By (42) and (43), we see that (51) is the certainty equivalent of (8). The optimality of (51) follows from the same type of limiting argument as we used for the simple random walk with constant success probability. We will first show that f_t^* is the appropriate limit of the controls of the discrete-time optimal control. We will then verify that f_t^* is in fact the optimal control by a martingale argument. To proceed, the development in section 3 shows that if we consider a sequence of RWIREs, the n-th of which has success probability θ_n, and step size $\sigma/\sqrt{n}$, then the optimal policy is to bet

$$f^*(S_k^n, k) = \frac{2E(\theta_n|S_k^n) - 1}{\sigma/\sqrt{n}}$$

at the k-th step. To show that the limit of the controls for the random walks converge to the optimal control for the corresponding diffusion, i.e., that $f^*(S_{[nt]}^n, [nt]) \Rightarrow f_t^*$, where f_t^* is given by (51), we need the following corollary to Theorem 2, which gives the weak convergence result for the drift of the sequence of RWIREs.

Corollary 3 *Under the conditions of Theorem 2,*

$$\frac{2E(\theta_n|S_{[nt]}^n) - 1}{\sigma/\sqrt{n}} \Rightarrow \frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right). \tag{52}$$

Proof: By (22),

$$2E(\theta_n|S_k^n) - 1 = \frac{S_k^n + \alpha_n - \beta_n}{k + \alpha_n + \beta_n}. \tag{53}$$

By (53), (31), (32) and Theorem 2

$$\frac{2E(\theta_n|S_{[nt]}^n) - 1}{\sigma/\sqrt{n}} = \frac{\frac{S_{[nt]}^n}{\sqrt{n}} + \frac{\mu\sigma}{c} - \frac{\mu}{c\sigma}\left(\frac{\mu^2+c}{n}\right)}{\sigma\left(t + \frac{\sigma^2}{c} - \frac{(\mu^2+c)}{cn}\right)} \Rightarrow \frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right). \quad \blacksquare$$

Of course, we have not yet proved that f_t^* given by (51) is in fact the optimal policy for the continuous time problem, but only that it is the weak limit of the discrete-time optimal controls. We will in fact verify this (in greater generality) after we discuss the portfolio problem. We note now however, that under the policy given above, the optimal fortune evolves as

$$dV_t^* = V_t^*\left[\frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right)^2 dt + \frac{1}{\sigma}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right) d\hat{W}_t\right], \quad V_0^* = x, \tag{54}$$

which can be verified by placing (51) and (43) into (50). This allows us to compare the Bayesian gambler with the *constant* gambler and the gambler with *perfect information* as in Section 2.3. Barron and Cover (1988) have developed general bounds on the growth rate of wealth for various stages of information in terms of conditional entropy measures. Due to the specific parametric form of the model considered here, (54) will enable us to characterize these quantities to a very explicit degree.

5.2 The Financial Value of Randomness and Information

Under the optimal Bayesian policy, the gambler's fortune evolves as (54), and therefore by Ito's formula we find

$$\ln V_T^* = \ln x + \frac{1}{2\sigma^2}\int_0^T \left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right)^2 dt + \frac{1}{\sigma}\int_0^T \left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right) d\hat{W}_t\,, \tag{55}$$

which is the continuous-time limit of (28). If we let $F_T^B(x)$ denote the optimal value function for the Bayesian gambler, then taking expectations on (55) using (44) shows

$$F_T^B(x) := E\ln V_T^* = \ln x + \frac{\mu^2}{2\sigma^2}T + \frac{c}{2\sigma^2}T + \frac{1}{2}\ln\left(\frac{\sigma^2}{\sigma^2 + cT}\right)\,. \tag{56}$$

Comparing this with (10) and (12) shows that for any $T > 0$, the optimal value in the Bayesian case is always sandwiched between the optimal value of the *constant* gambler and the gambler who randomizes with *perfect information*, i.e.,

$$F_T(x) \le F_T^B(x) \le F_T^P(x).$$

Thus while the randomness in the model is helping the Bayesian gambler do better than the constant gambler, he of course cannot do better than the gambler with perfect information. However, it appears that in fact, he does substantially better than the constant gambler since

$$F_T^B(x) - F_T(x) = \frac{c}{2\sigma^2}T + \frac{1}{2}\ln\left(\frac{\sigma^2}{\sigma^2 + cT}\right)\,, \tag{57}$$

which is convex increasing in the horizon T, while only slightly less than the gambler with perfect information. In particular, we see from (12) and (56) that the financial cost of learning is logarithmic since

$$F_T^P(x) - F_T^B(x) = \frac{1}{2}\ln\left(1 + \frac{c}{\sigma^2}T\right)\,. \tag{58}$$

This should be considered the *financial value of perfect information*, since it is how much the Bayesian gambler should be willing to pay to learn the true value of Z.

We note that $F_T(x)$ given by (10) is also the value under the Bayesian model for a gambler who always invests according to the *prior* expected value, i.e., a gambler who always invests the constant fraction μ/σ^2, and whose fortune therefore evolves as $dV_t = (\mu/\sigma^2)V_t\,dX_t$, for dX_t in (43).

Therefore (57) can be considered the *financial gain of using the posterior over the prior*. From (57) we find that the expected *marginal benefit* of increasing the horizon T to the gambler using the posterior over the prior is

$$\frac{\partial}{\partial T}\left(F_T^B(x) - F_T(x)\right) = \frac{c^2}{2\sigma^2}\left(\frac{T}{\sigma^2 + cT}\right) \to \frac{c}{2\sigma^2}\,, \qquad \text{as } T \to \infty$$

while from (58) we find that the expected marginal benefit of increasing the horizon to the gambler with perfect information over the Bayesian gambler is

$$\frac{\partial}{\partial T}\left(F_T^P(x) - F_T^B(x)\right) = \frac{c}{2\sigma^2}\left(\frac{1}{\sigma^2 + cT}\right) \to 0\,, \qquad \text{as } T \to \infty\,.$$

5.3 Adaptive Portfolios

Similarly, if we consider the portfolio problem associated with a sequence of RWIREs, where the nth random success probability is θ_n, and the step size is δ_n, it follows that the optimal control is to invest, at time $[nt]$ in the nth random walk, the fraction (see (30))

$$f^{*(n)}(S^n_{[nt]}, [nt]) = \frac{(1+r)\left(\delta_n(2E(\theta_n|S^n_{[nt]}) - 1) - r\right)}{\delta_n^2 - r^2}. \tag{59}$$

The optimal control for the resulting diffusion control problem, i.e., for the portfolio problem where the drift of the Brownian motion has a normal distribution, can now be obtained with the aid of the previous results.

Suppose an investor is faced with the diffusion control problem of maximizing $E(\ln V_T)$, where the *return process* of the risky stock, X_t, is a Brownian motion with a random, normally distributed drift. I.e., in the context of Merton's (1971) model, we will assume that the price of the risky stock, S_t, and the price of the riskless bond, B_t, evolve respectively as

$$dS_t = ZS_t dt + \sigma S_t dW_t \quad \text{and} \quad dB_t = \gamma B_t dt$$

where W_t is a standard Brownian motion, and Z is an independent *unobserved* random variable with a normal distribution, as before. If we define the return process, X_t, associated with the stock price process S_t by

$$X_t := \int_0^t \frac{1}{S_u} dS_u \tag{60}$$

then it follows from our previous results that in this case, the extension of Merton's control problem is equivalent to solving for the optimal admissible control process $\{f_t : 0 \le t \le T\}$, to maximize $E \ln V_T$, where now

$$dV_t = V_t(1 - f_t)\gamma\, dt + V_t f_t\, dX_t \tag{61}$$

$$dX_t = \frac{cX_t + \sigma^2\mu}{\sigma^2 + ct} dt + \sigma\, d\hat{W}_t. \tag{62}$$

(Note that a simple application of Ito's rule shows that $X_t = \ln S_t + t\sigma^2/2$, where $S_0 = 1$, since we've assumed that $X_0 = 0$.)

In this case, the problem is clearly more complicated than in the nonrandom case. However, an appeal to our previous results will again give us the optimal control without resorting to the HJB techniques. The optimal control is the certainty equivalent of the optimal control (16) first obtained by Merton (1971) for the nonrandom case.

Theorem 7 *In the context of (61) and (62) the optimal control is*

$$f_t^* = \frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct} - \gamma\right). \tag{63}$$

Proof: We will show that $f^{*(n)}(S^n_{[nt]}, [nt]) \Rightarrow f_t^*$. To see that this is true, substitute for the step size, $\delta_n = \sigma/\sqrt{n}$, as well as for the interest rate, in terms of the interest rate per step, i.e., $r = \gamma/n$, in (59), to rewrite it as

$$\begin{aligned} f^{*(n)}(S^n_{[nt]}, [nt]) &= \left(1 + \frac{\gamma}{n}\right)\left(\frac{\frac{\sigma}{n}\left(2E(\theta_n|S^n_{[nt]}) - 1\right) - \frac{\gamma}{n}}{\frac{\sigma^2}{n} - \frac{\gamma^2}{n^2}}\right) \\ &= \frac{(1 + \frac{\gamma}{n})}{\left(1 - \frac{\gamma^2}{\sigma^2 n}\right)}\left(\frac{2E(\theta_n|S^n_{[nt]}) - 1}{\sigma/\sqrt{n}} - \frac{\gamma}{\sigma^2}\right) \\ &\Rightarrow \frac{1}{\sigma^2}\left(\frac{cX_t + \mu\sigma^2}{\sigma^2 + ct}\right) - \frac{\gamma}{\sigma^2} \end{aligned}$$

by (52).

5.4 The Final Martingale Argument

It remains now to verify that the limiting controls obtained above are in fact the optimal ones for the continuous-time problems. In fact, more is true. We have concentrated on the natural conjugate case (where Z was assumed to have a normal distribution) only for the sake of analytical tractability. However, as we now show, the optimality of the certainty equivalent holds for arbitrary prior distributions, i.e., *regardless of the parametric form of the distribution of the unobserved random variable, Z, the optimal control is to invest*

$$f_t^* = \frac{1}{\sigma^2}\left[E(Z|\mathcal{F}_t^X) - \gamma\right] \tag{64}$$

in the risky stock, where $\{\mathcal{F}_t^X : t \geq 0\}$ denotes the filtration of the return process X defined in (60). Note that (64) is the general certainty equivalent of (16), and that $E(Z|\mathcal{F}_t^X)$ can be computed for arbitrary prior distributions from (41).

To verify the optimality of the policy f_t^* of (64), apply Ito's rule to the semi-martingale $\ln V_t$, where V_t is given by (61) but where dX_t is given by the more general (40) to get

$$d\ln V_t = \frac{1}{V_t}dV_t - \frac{1}{2V_t^2}d < V >_t$$

where $< V >_t$ is the quadratic variation of the semi-martingale V_t, and note that $< V >_t = f_t^2 V_t^2 \sigma^2\, t$. Therefore, we have

$$\begin{aligned} d \ln V_t &= (1 - f_t)\gamma\, dt + f_t dX_t - \frac{1}{2} f_t^2 \sigma^2\, dt \\ &= \left[(1 - f_t)\gamma + f_t\, E(Z|\mathcal{F}_t^X) - \frac{1}{2} f_t^2 \sigma^2 \right] dt + f_t \sigma\, d\hat{W}_t\,, \end{aligned}$$

where the last equality follows from (40). Integrating from 0 to T, and then taking conditional expectations – recognizing that the posterior mean of Z, $E(Z|\mathcal{F}_t^X)$, is a function of just X_t alone (see (41)) – shows that

$$E\left(\ln V_T \,|\, \mathcal{F}_{T-}^X\right) = \int_0^T \left[(1 - f_t)\gamma + f_t\, E(Z|\mathcal{F}_t^X) - \frac{1}{2} f_t^2 \sigma^2 \right] dt\,. \tag{65}$$

Since the right-hand-side of (65) does *not* contain the process V_t, it is clear that it suffices to maximize the integrand *pointwise* with respect to f_t, which will result in the optimal control f_t^* given above in (64). For the special case where the prior is normal, this of course reduces to the policy of (63). Thus we may conclude that the policies given above are in fact optimal for the continuous-time problems. ∎

6 Other Utility Functions

For the objective of maximizing log utility, we have shown that the optimal policy for the RWIRE as well as for the Brownian motion with a random mean, is the certainty equivalent of respectively the ordinary random walk and the ordinary Brownian motion, i.e., the optimal control for the case of a random parameter is obtained by just substituting the current conditional expected value of the random parameter (μ) in place of the known parameter in the control policy for the case of the known parameter.

This raises the question of whether a similar result holds for other objective functions as well. While we do not have the full answer to this question, we do observe that this is *not* true for the case of a power utility function. In fact, Bellman and Kalaba (1957) proved that a purely proportional betting scheme is optimal only for the case where the gambler is interested in maximizing $E(V_T^\eta/\eta + \kappa)$ for $0 \le \eta \le 1$, and for some constant κ. (The logarithmic case occurs in the limit as $\eta \to 0$ when $\kappa = -1/\eta$.) A somewhat more general result for continuous-time problems was found by Merton (1971), for what he referred to as the HARA (hyperbolic-absolute-risk-aversion) class of utility functions.

Now, since for $\eta > 0$ the strategy that maximizes $E(V_T^\eta/\eta + \kappa)$ is the same as the strategy that maximizes $E(V_T^\eta)$, without any loss in generality, we will suppose that the investor's objective is to maximize $E(V_T^\eta)$, for $0 < \eta < 1$. For notational ease, we will consider only the case of $r = \gamma = 0$. We can no longer use the simple approach of Sections 2 and 3 to determine the optimal utility as well as the optimal policy. For this case, we must use the technique of dynamic programming. If

we let $F_m(x)$ denote the optimal value function with m bets to go and with a current fortune of x, then when all parameters are known constants, the optimality equation for this problem is

$$F_m(x) = \max_f \{\theta F_{m-1}(x + fx\Delta) + (1-\theta)F_{m-1}(x - fx\Delta)\} \ ,$$

with the boundary condition $F_0(x) = x^\eta$. For $m = 1$, we have

$$\begin{aligned} F_1(x) &= \max_f \{\theta(x + fx\Delta)^\eta + (1-\theta)(x - fx\Delta)^\eta\} \\ &= x^\eta \max_f \{\theta(1 + f\Delta)^\eta + (1-\theta)(1 - f\Delta)^\eta\} \end{aligned}$$

from which a simple computation yields the optimizer

$$f^* = \frac{1}{\Delta}\left[\frac{\theta^{\frac{1}{1-\eta}} - (1-\theta)^{\frac{1}{1-\eta}}}{\theta^{\frac{1}{1-\eta}} + (1-\theta)^{\frac{1}{1-\eta}}}\right] , \tag{66}$$

and associated optimal value

$$F_1(x) = x^\eta \left[C(\theta, \eta)\right]$$

for

$$C(\theta, \eta) = 2^\eta \left(\theta^{\frac{1}{1-\eta}} + (1-\theta)^{\frac{1}{1-\eta}}\right)^{1-\eta} . \tag{67}$$

Induction then shows that for a horizon of length N, the optimal policy is in fact to invest the same fixed fraction f^* of (66) at each stage, with a final terminal value

$$F_N(x) = x^\eta \left[C(\theta, \eta)\right]^N \ .$$

The optimal policy for the continuous time case can be obtained from this by letting f_n^* denote the optimal control for the nth random walk, with $\theta_n = \frac{1}{2}\left(1 + \frac{\mu}{\sigma\sqrt{n}}\right)$, and $\Delta = \delta_n = \sigma/\sqrt{n}$ in (66).

Letting $\epsilon = \frac{1}{1-\eta}$, and then expanding around 0, we see that

$$\begin{aligned} f_n^* &= \frac{1}{\delta_n}\left(\frac{\theta_n^\epsilon - (1-\theta_n)^\epsilon}{\theta_n^\epsilon + (1-\theta_n)^\epsilon}\right) \\ &= \frac{\sqrt{n}}{\sigma}\left(\frac{2\epsilon(\frac{\mu}{\sigma\sqrt{n}}) + o(n^{-2})}{2 + o(n^{-2})}\right) \\ &\to \frac{\epsilon\mu}{\sigma^2} \equiv \frac{\mu}{\sigma^2(1-\eta)} , \end{aligned}$$

which is in fact the optimal control for the continuous-time problem (see Merton (1971) or Karatzas (1989)).

Since the optimal control for the case of a power utility function is again a fixed fraction, it is natural to expect that the Bayesian controls might then be the certainty equivalent of these constants. However, the Bayesian case breaks down already for a two-step horizon. In the Bayesian discrete-time problem, the dynamic programming equation is given by

$$\begin{aligned} &F_m(x, S_k, k) = \\ &\quad \max_f \{E(\theta|S_k,k)F_{m-1}(x+fx, S_k+1, k+1) + (1-E(\theta|S_k,k))F_{m-1}(x-fx, S_k-1, k+1)\} , \end{aligned}$$

with the terminal boundary condition $F_0(x, S_k, k) = x^\eta$. Thus for a one-step horizon, the Bellman equation is

$$F_1(x, S_k, k) = \max_f \{E(\theta|S_k, k)\,(x + fx\Delta)^\eta + (1 - E(\theta|S_k, k))\,(x - fx\Delta)^\eta\} \tag{68}$$

and the optimizer is

$$f^*(S_k, k) = \frac{1}{\Delta}\left[\frac{[E(\theta|S_k, k)]^{\frac{1}{1-\eta}} - [1 - E(\theta|S_k, k)]^{\frac{1}{1-\eta}}}{[E(\theta|S_k, k)]^{\frac{1}{1-\eta}} + [1 - E(\theta|S_k, k)]^{\frac{1}{1-\eta}}}\right], \tag{69}$$

which is the certainty equivalent of the nonrandom case. When the optimal control is placed back into the Bellman equation (68), we get the optimal value function

$$F_1(x, S_k, k) = x^\eta \cdot C\,(\,E(\theta|S_k, k), \eta)$$

where the function C is given by (67). Since $F_1(x, S_k, k) = F_0(x, S_k, k)C$, it is clear that for a two-step horizon the optimality equation can be written as

$$\begin{aligned} F_2(x, S_k, k) \;\; = \;\; & x^\eta \max_f \{E(\theta|S_k, k) \cdot (1 + f)^\eta C\,(E(\theta|S_k + 1, k + 1), \eta) \\ & + (1 - E(\theta|S_k, k)) \cdot (1 - f)^\eta C\,(E(\theta|S_k - 1, k + 1), \eta)\}\,, \end{aligned}$$

resulting in the optimizer

$$f_2^*(x, S_k, k) =$$

$$\left(\frac{1}{\Delta}\right)\frac{\left[\hat\theta_k(S_k)\hat\theta_{k+1}(S_k{+}1)\right]^\epsilon + \left[\hat\theta_k(S_k)\overline{\hat\theta_{k+1}}(S_k{+}1)\right]^\epsilon - \left[\hat\theta_{k+1}(S_k{-}1)\overline{\hat\theta_k}(S_k)\right]^\epsilon - \left[\overline{\hat\theta_{k+1}}(S_k{-}1)\overline{\hat\theta_k}(S_k)\right]^\epsilon}{\left[\hat\theta_k(S_k)\hat\theta_{k+1}(S_k{+}1)\right]^\epsilon + \left[\hat\theta_k(S_k)\overline{\hat\theta_{k+1}}(S_k{+}1)\right]^\epsilon + \left[\hat\theta_{k+1}(S_k{-}1)\overline{\hat\theta_k}(S_k)\right]^\epsilon + \left[\overline{\hat\theta_{k+1}}(S_k{-}1)\overline{\hat\theta_k}(S_k)\right]^\epsilon}$$

where $\epsilon := 1/(1 - \eta)$, $\hat\theta_j(y) := E(\theta|S_j = y, j)$, and $\overline{z} := 1 - z$.

Clearly, the optimal control in this case is *not* the certainty equivalent, and thus it appears that the continuous-time optimal control will also not be the certainty equivalent.

We may in fact continue the iteration to show that the Bayesian optimal control, after observing the random walk for k steps, with j steps left to go, is to invest the fraction

$$f_j^*(x, S_k, k) = \left(\frac{1}{\Delta}\right)\frac{\left[\hat\theta_k(S_k)g_{j-1}(k + 1, S_k + 1)\right]^\epsilon - \left[\overline{\hat\theta_k}(S_k)g_{j-1}(k + 1, S_k - 1)\right]^\epsilon}{\left[\hat\theta_k(S_k)g_{j-1}(k + 1, S_k + 1)\right]^\epsilon + \left[\overline{\hat\theta_k}(S_k)g_{j-1}(k + 1, S_k - 1)\right]^\epsilon} \tag{70}$$

where the functions $g_l(r, v)$ are defined by the nonlinear difference equations

$$g_l(r, v) = \left(\left[\hat\theta_r(v)g_{l-1}(r + 1, v + 1)\right]^\epsilon + \left[\overline{\hat\theta_r}(v)g_{l-1}(r + 1, v - 1)\right]^\epsilon\right)^{1/\epsilon} \tag{71}$$

with $g_0 = 1$.

A similar analysis can be given for other utility function such as the exponential, where the investor is interested in maximizing $E\,(a - b\exp\{-\lambda V_T\})$. In this case too the discrete-time optimal

control in the Bayesian case is also *not* the certainty equivalent. However, since an exponential utility function does not lead to a proportional strategy for the case in which the parameters are known constants, we will not pursue this here.

It would appear from the development in Section 3 that a *sufficient* condition for the Bayesian control to be the certainty equivalent of the deterministic control is that the optimal value function be completely separable, i.e., that

$$F_j(x, S_k, k) = F_{j-1}(x, S_k, k) + \psi(S_k, k)$$

where $\psi(\cdot)$ is independent of the wealth x. This is the case for the logarithmic utility, but not for power utilities nor exponential.

The question of what the optimal controls of (70) converge to remains to be determined.

7 Concluding Remarks

Besides the problems just outlined for utility functions other than the logarithmic, four important issues still remain open: First, we need to prove for both the Bayesian and non-Bayesian problems that the discrete-time results directly imply optimality for the limiting controls. Second, we need to establish limits in the Bayesian setting for more general gambles than simple random walks. Third, we need to determine useful conditions on general priors (and even priors for simple random walks (beyond (36)) for the posterior means to be well behaved, i.e., to be in the domain of attraction of a normally distributed random drift of a Brownian motion. Finally, it would be nice to consider more general models in which the success probability of the simple random walk is itself an unobserved stochastic process instead of just a fixed random variable.

References

[1] Algoet, P.H. and Cover, T.M. (1988) Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment. *Ann. of Probab.*, **16**, 876-898.

[2] Barron, A. and Cover, T.M. (1988) A Bound on the Financial Value of Information. *IEEE Trans. Info. Theory*, **34**, 1097-1100.

[3] Bell, R.M., and Cover, T.M. (1980) Competitive Optimality of Logarithmic Investment. *Math. Oper. Res.*, **5**, 161-166.

[4] Bellman, R. and Kalaba, R. (1957) On the Role of Dynamic Programming in Statistical Communication Theory. *IRE (now IEEE) Trans. Info. Theory*, Vol. IT, **3**, 197-203.

[5] Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley.

[6] Breiman, L. (1961) Optimal Gambling Systems for Favorable Games. *Fourth Berkeley Symp. Math. Stat. and Prob.*, **1**, 65 - 78.

[7] Cover, T.M. (1991). Universal Portfolios, *Mathematical Finance*, **1**, 1-29.

[8] Cover, T.M. and Gluss, D.H. (1986). Empirical Bayes Stock Market Portfolios, *Adv. Appl. Math.*, **7**, 170-181.

[9] Cover, T.M. and Thomas, J. (1991) *Elements of Information Theory*, Wiley.

[10] Csörgö, M. and Révész, P. (1981) *Strong Approximations in Probability and Statistics*, Academic Press.

[11] DeGroot, M.H. (1970) *Optimal Statistical Decisions*, McGraw-Hill.

[12] Ethier, S.N. (1988). The proportional bettor's fortune. *Proc. 5th International Conference on Gambling and Risk Taking*, **4**, 375-383.

[13] Ethier, S.N. and Tavare, S. (1983) The proportional bettor's return on investment. *J. Appl. Prob.*, **20**, 563 - 573.

[14] Feller, W. (1971) *An Introduction to Probability Theory and its Applications, Vol.2*, 2nd ed., Wiley.

[15] Finkelstein, M. and Whitely, R. (1981) Optimal Strategies for repeated games. *Adv. Appl. Prob.*, **13**, 415 - 428.

[16] Gottlieb, G. (1985) An optimal betting strategy for repeated games. *J. Appl. Prob.*, **22**, 787 - 795.

[17] Hakansson, N. (1970) Optimal investment and consumption strategies under risk for a class of utility functions. *Econometrica*, **38**, 587 - 607.

[18] Heath, D., Orey,S., Pestien, V. and Sudderth, W. (1987) Minimizing or Maximizing the expected time to reach zero. *SIAM J. Cont. and Opt.*, **25**, 195-205.

[19] Jamishidian, F. (1992). Asymptotically Optimal Portfolios, *Mathematical Finance*, **2**, 131-150.

[20] Kallianpur, G. (1980) *Stochastic Filtering Theory*, Springer-Verlag.

[21] Karatzas, I. (1989) Optimization Problems in the Theory of Continuous Trading. *SIAM J. Cont. and Opt.*, **7**, 1221-1259.

[22] Karatzas, I. and Shreve, S. (1988) *Brownian Motion and Stochastic Calculus*, Springer-Verlag.

[23] Kelly, J. (1956) A New Interpretation of Information Rate. *Bell Sys. Tech. J.*, **35**, 917 - 926.

[24] Kurtz, T.G. and Protter, P. (1991) Weak Limit Theorems for Stochastic Integrals and Stochastic Differential Equations. *Ann. Probab.*, **19**, 1035-1070.

[25] Kushner, H.J. and Dupuis, P.G. (1992) *Numerical Methods for Stochastic Control Problems in Continuous-Time*, Springer-Verlag.

[26] Latane, H. (1959) Criteria for choice among risky assets. *J. Polit. Econ.*, **35**, 144 - 155.

[27] Merton, R. (1971) Optimum Consumption and Portfolio Rules in a Continuous Time Model. *J. Econ. Theory*, **3**, 373-413.

[28] Merton, R. (1990) *Continuous Time Finance*, Blackwell.

[29] Pestien, V.C., and Sudderth, W.D. (1985) Continuous-Time Red and Black: How to control a diffusion to a goal. *Math. Oper. Res.*, **10**, 599-611.

[30] Révész, P. (1990) *Random Walk in Random and Non-Random Environments*, World Scientific.

[31] Thorp, E.O. (1969) Optimal gambling systems for favorable games. *Rev. Int. Stat. Inst.*, **37**, 273 - 293.

[32] Thorp, E.O. (1971). Portfolio Choice and the Kelly Criterion, *Bus. and Econ. Stat. Sec., Proc. Amer. Stat. Assoc.*, 215-224.