Quantitative Investment

S. G. Kou
Questrom School of Business
Boston University

This version Spring 2020

Chapter 7

Dynamic Money Management (II): Kelly Criterion

7.1 The Problem Setting

Consider a sequence of returns R_i associated with a risky asset. More precisely, R_i is the return of the investment at time i. For simplicity, we shall assume that R_i, $i \geq 1$, are i.i.d. random variables. The central question here is to study what fractional of money π_i for an investor to put into the investment at time i. Note that the wealth of the investor X_n at time n is given by

$$X_n = X_{n-1} + (\pi_{n-1} X_{n-1}) R_n + (1 - \pi_{n-1}) X_{n-1} r,$$

where r is the risk-free rate, because $(1 - \pi_n) X_{n-1}$ is the amount of money invested in the money market account. The above recursion can be easily solved to get

$$X_n = X_0 \prod_{i=1}^{n} (1 + r + \pi_{i-1} (R_i - r)).$$

The investor aims at maximizing the utility function

$$\max_{\pi} E[U(X_n)]$$

For simplicity we shall first discuss the constant proportion investment strategy, i.e.

$$\pi_i = \pi,$$

and attempt to choose π optimally. Note the constant proportion strategy is a contrarian strategy, i.e. it will buy more shares of the asset when the asset price goes down, and sell shares of the asset when the asset price goes up, to maintain the constant proportion. In addition, it was shown in N. H. Hakansson (1970, Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions, Econometrica, Vol., 38, pp. 585-607) that the constant proportion strategy is optimal if and only if the utility function is either $U(x) = x^{1-\gamma}/(1-\gamma)$, $\gamma > 0$, or $U(x) = \ln(x)$.

7.2 An Example

Suppose $r = 0$ and we have investment such that

$$R = \left\{ \begin{array}{ll} \text{100\%} & \text{with prob. } 70\% \\ \text{-100\%} & \text{with prob. } 30\% \end{array} \right\}.$$

Should you invest 100% of your money in this to play it for n times, where n is not a small number? Clearly the answer is no, especially for large n, because if one invests 100%, then losing one time means one loses everything. In particular, if one invest 100% of money, then the probability of still no bankruptcy after the nth investment is 0.7^n, which may be quite small (e.g. when $n = 10$, $0.7^{10} = 2.8\%$). However, things becomes less clear with the following example.

Example 1. Suppose we have an investment such that

$$R = \left\{ \begin{array}{ll} \text{50\%} & \text{with prob. } 55\% \\ \text{-50\%} & \text{with prob. } 45\% \end{array} \right\}.$$

Note that the expected return for each time is

$$0.55 * 0.5 - 0.45 * 0.5 = 5\%.$$

Should you invest 100% of your money to play it for 100 times?

In this example, one cannot lose everything, if one just loses once. Thus, the above argument fails. To answer the question in this case, we have to study median and various objective functions

7.3 Mean and Median of the Investment

Suppose we have investment such that

$$R = \left\{ \begin{array}{ll} u & \text{with prob. } p \\ d & \text{with prob. } q = 1-p \end{array} \right\},$$

where $u > 0 > d$, and one invests π fractional wealth to play it for n times. Assume $r = 0$. Then the wealth after the n plays is

$$X_n = X_0 \prod_{i=1}^{n} (1+\pi R_i) = X_0 (1+\pi u)^{B_n} (1+\pi d)^{n-B_n} = X_0 \left(\frac{1+\pi u}{1+\pi d} \right)^{B_n} (1+\pi d)^n,$$

where B_n is a binomial random variable with parameters n and p. We are interested in compute the mean and the median of X_n.

Clearly, the mean of X_n is given by

$$\begin{aligned} E[X_n] &= X_0 \prod_{i=1}^{n} E(1+\pi R_i) = X_0 \left\{(1+\pi u)p + (1+\pi d)q\right\}^n \\ &= X_0 \left\{1 + \pi\left(pu + qd\right)\right\}^n. \end{aligned}$$

For Example 1, with $\pi = 100\%$, $n = 100$, and $X_0 = 10{,}000$, this yields

$$10000 * \left\{1 + (0.55 * 0.5 - 0.45 * 0.5)\right\}^{100} = 1{,}315{,}000.$$

Thus, with 10,000 after 100 times it becomes over 1.3 million!

To compute the median of X_n, i.e. $M(X_n)$, note that for our purpose, we define median to be the smallest number such that there are at most 50% of the number less than or equal to such number. This will avoid the standard problem of taking the average if there are two median numbers. For example, if we have 1, 3, 4, 5, then the median will be 3 by our definition, rather than (3+4)/2.

Also note that

(i) The ln of median is the median of the ln, due to the monotonicity of the ln function

(ii) $M(aX+b) = aM(X) + b$ for constants a and b.

Thus, we have

$$\begin{aligned} \ln(M(X_n)) &= M(\ln(X_n)) = M\left(\ln X_0 + B_n \ln\left(\frac{1+\pi u}{1+\pi d}\right) + n\ln(1+\pi d)\right) \\ &= M\left(B_n\right) \cdot \ln\left(\frac{1+\pi u}{1+\pi d}\right) + \ln X_0 + n\ln(1+\pi d). \end{aligned}$$

To compute the median of the binomial distribution, note the following result from K. Hamza (1995, The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions, Statistics and Probability Letters, Vol 23, pp. 21-25): with $\gamma_n = M(B_n) - np$ we have

$$|\gamma_n| < \ln 2 = 0.693...$$

For our Example 1, with $\pi = 100\%$, $n = 100$, and $X_0 = 10{,}000$, we have

$$|M(B_n) - 100 * 0.55| < 0.693...,$$

yielding $|M(B_n) - 55| < 0.693...$, which further implies that

$$M(B_n) = 55,$$

because in our definition $M(B_n)$ is a unique number for the binomial distribution and must be an integer. Thus

$$\ln(M(X_n)) = 55 * \ln\left(\frac{1+0.5}{1-0.5}\right) + \ln 10000 + 100 * \ln(1-0.5).$$

Take exponential functions on both sides we have

$$\begin{aligned} M(X_n) &= \left(\frac{1+0.5}{1-0.5}\right)^{55} * 10000 * (1-0.5)^{100} \\ &= (1+0.5)^{55} * 10000 * (1-0.5)^{45} \\ &= 1.37\,62. \end{aligned}$$

In summary, with full investment, i.e. $\pi = 100\%$ and the initial investment of \$10,000, we have the mean in Example 1 is about 1.3 million dollars, but the median is only about \$1!

The truth is that the distribution is highly skewed to the right, and while most of people lose, the few people make a lot of money, making the mean very high. This can be seen in two ways.

(1) Suppose one win 100 times. Of course, this happens with a tiny but non-zero probability. If this happens, the final payoff with the initial investment of \$10,000 is

$$10000 * 1.5^{100} = 4.065\,6 \times 10^{21}.$$

Remember that the money for a billionaire is only about 10^9. Now the money in the whole world is perhaps only a tiny fraction of 10^{21}! The mean is quite large compare to median, partly because there are few cases, one makes really a lot of money.

(2) Suppose one starts with \$1, and win one lose one. Then the money becomes

$$1 * 1.5 * 0.5 = 0.75.$$

Thus, due to the multiplication (not the addition) nature of the money, the median tends to be smaller!

This example shows the importance of choosing an optimal π.

7.4 Brownian Approximation

Instead of computing things exactly in the discrete setting, it is easier to obtain analytical solution in the continuous setting. For example, we can use the geometric Brownian motion to approximate the original wealth process. In fact, in the continuous setting the wealth process X_t^π using the strategy π becomes

$$dX_t^\pi = (r + \pi(\mu - r))\, X_t^\pi dt + \sigma\pi X_t^\pi dW_t.$$

The Ito formula yields

$$\ln(X_t^\pi / X_0) = \left(r + \pi\,(\mu - r) - \frac{\pi^2\sigma^2}{2}\right) t + \pi\sigma W_t. \tag{7.1}$$

Thus, the mean and median are given by

$$E[X_t^\pi] = X_0 \exp\{(r + \pi\,(\mu - r))\, t\}$$

$$M[X_t^\pi] = X_0 \exp\left\{\left(r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}\right)t\right\}.$$

For example, in Example 1, we have an approximation by the geometric Brownian motion with

$$\mu = E[R] = 0.05,\ \sigma^2 = Var[R] = 0.5^2 - (0.05)^2 = 0.2475.$$

In summary, we can use an approximation with the geometric Brownian motion with

$$\mu = 0.05, \quad \sigma = \sqrt{0.2475} = 0.497\,49. \tag{7.2}$$

Thus, if we take $\pi = 100\%$ in Example 1, we have

$$E[X_t^\pi] = X_0 \exp\{0.05 * t\}$$

$$M[X_t^\pi] = X_0 \exp\left\{\left(0.05 - \frac{0.2475}{2}\right)t\right\} = X_0 \exp\{-0.07375 * t\}.$$

In particular, after 100 plays, we have

$$E[X_{100}^\pi] = 10000 * \exp\{0.05 * 100\} = 1.\,484\,1 \times 10^6$$

$$M[X_t^\pi] = 10000 * \exp\{-0.07375 * 100\} = 6.2673.$$

These approximations appear to be good compared to the true values computed in the previous section.

Furthermore,

$$\begin{aligned} P(X_t^\pi &\leq X_0) = P\left(\exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\right\} \leq 1\right) \\ &= P\left((\mu - \frac{1}{2}\sigma^2)t + \sigma W_t \leq 0\right) \\ &= P\left(W_t \leq -\frac{(\mu - \frac{1}{2}\sigma^2)t}{\sigma}\right). \end{aligned}$$

With $T = 100$, we have

$$\begin{aligned} P(X_T^\pi &\leq X_0) = \Phi\left(-\frac{\left(0.05 - \frac{1}{2}0.2475\right) * 100}{\sqrt{0.2475 * 100}}\right) \\ &= \Phi(1.4824) \\ &= 93.09\%; \end{aligned}$$

i.e. about 93% of people lose money, if they invest 100% of their money.

7.5 Kelly Criterion

7.5.1 Random Walk Model

Under the discrete random walk mode, the goal is to maximize the expected long run growth rate (or using the logarithm utility)

$$\max_{\pi} E[\ln(X_n)].$$

For example, if

$$R = \left\{ \begin{array}{ll} u & \text{with prob. } p \\ d & \text{with prob. } q = 1-p \end{array} \right\},$$

then

$$E[\ln X_n] = \ln X_0 + nE\ln(1+\pi R_i) = \ln X_0 + n\left\{p\ln(1+\pi u) + (1-p)\ln(1+\pi d)\right\},$$

from which we can get the optimal π^*.

Indeed we have the first order condition

$$0 = \frac{\partial}{\partial \pi} E[\ln X_n] = \frac{un}{1+\pi u} p + \frac{dn}{1+\pi d}(1-p).$$

Therefore, the optimal π^* satisfies

$$up(1+\pi^* d) = -d\,(1+\pi^* u)\,(1-p),$$

or

$$\pi^* = -\frac{up + d(1-p)}{dup + du(1-p)} = -\frac{up + d(1-p)}{du}.$$

In addition,

$$\frac{\partial^2}{\partial \pi^2} E[\ln X_n] = -\left(\frac{1}{1+\pi u}\right)^2 u^2 np - \left(\frac{1}{1+\pi d}\right)^2 d^2 n(1-p) < 0,$$

which implies that the objective function is concave and achieves the unique maximum point at π^*.

In Example 1, we have

$$R = \left\{ \begin{array}{ll} \text{50\%} & \text{with prob. 55\%} \\ \text{-50\%} & \text{with prob. 45\%} \end{array} \right\}.$$

Thus, with $u = 0.5$, $d = -0.5$, $p = 0.55$, we have

$$\pi^* = -\frac{0.5 * 0.55 - 0.5 * 0.45}{0.5 * (-0.5)} = 0.20.$$

7.5.2 Geometric Brownian motion model

Note that under the geometric Brownian motion model

$$E[\ln(X_t^\pi/X_0)] = \left(r + \pi\,(\mu - r) - \frac{\pi^2\sigma^2}{2}\right)t.$$

Maximizing the above expression yields the optimal π for the logarithm utility, i.e., under the Kelly Criterion

$$\pi^* = \frac{\mu - r}{\sigma^2}.$$

If we use the approximation parameter (7.2), we have

$$\pi^* = \frac{0.05 - 0}{0.2475} = 20.2\%,$$

which is similar to the case that we study before.

7.6 Fractional Kelly Criterion

By using the HJB equation, later we shall show that if we have a power type utility function $U(x) = x^{1-\gamma}/(1-\gamma)$, $\gamma > 0$, then the solution to the portfolio optimization problem

$$V(t,x) = \sup_{(\pi,c)\in\mathcal{A}} \mathsf{E}\left[\int_t^T e^{-\rho(s-t)}U(c(s))ds | X(t) = x\right]$$

is given by using the fractional wealth

$$\pi^* = \frac{1}{\gamma}\frac{\mu - r}{\sigma^2}$$

to invest in the risky asset. Thus, the solution to the power utility maximization problem is also called the fractional Kelly criterion.

Note that the logarithm utility used in Kelly criterion is a special case of power utility of the form

$$u(x) = \frac{x^{1-\gamma} - 1}{1-\gamma}, \quad \gamma > 0, \quad \gamma \neq 1.$$

Indeed, when $\gamma \to 1$, we have the logarithm utility

$$u(x) = \frac{x^{1-\gamma} - 1}{1-\gamma} \to \lim_{\gamma\to 1}\frac{x^{1-\gamma}(-\ln(x))}{-1} = \ln(x),$$

by L'Hôpital's rule.

As we mention before a typical value of γ is about 1.5 to 3.5, thus

$$\frac{1}{\gamma} \in [\frac{1}{3.5}, \frac{1}{1.5}] = [0.28571, 0.66667].$$

Just for illustration, for Example 1, Kelly criterion would give about 20%. With the power utility, this leads to a range of π^*

$$[5.714\%, \quad 13.333\%].$$

7.7 Problems with Kelly criterion

Below we will discuss problems related to Kelly criterion. See two papers:

Browne (2000), Can you do better than Kelly in the short run?

Browne (1999), The risk and rewards of minimizing shortfall probability.

The two papers are almost identical. So we focus on the first one, Browne (2000). Note that some numerical calculations in Browne (2000) are wrong, and were corrected in Browne (1999).

7.7.1 Background on the Laplace Transforms of First Passage Times

Recall that there is a one-to-one correspondence between a nonnegative random variable X and its Laplace transform $\mathsf{E}[e^{-sX}]$, $s > 0$. Indeed, if we can find the Laplace transform, then the distribution of X can be obtained either by the analytical inversion or numerical inversion (typically by using Fourier inversion). Very often it is much easier to compute the Laplace transform. For example, for the distribution of the sum of independent random variables is a convoluation of the each individual distributions, leading to high dimensional integrals. However, the Laplace transform of the sum is simply the products of each individual Laplace transforms.

Instead of using the probability transform, another way to compute the first passage times is to compute Laplace transforms via martingale methods. Without loss of generality, we can assume that $b > 0$. Let

$$s = \theta\mu + \frac{1}{2}\theta^2\sigma^2, \tag{7.3}$$

for any $s > 0$. Then it is easy to see that

$$\exp\{\theta W_{\mu,\sigma}(t) - st\}$$

is a martingale. Here $W_{\mu,\sigma}(t)$ denotes a Brownian motion with drift μ and volatility σ, i.e.

$$W_{\mu,\sigma}(t) = \mu t + \sigma W(t),$$

where $W(t)$ is a standard Brownian motion.

Heuristically we may appeal to the optional sampling theorem to find $\mathsf{E}[\exp\{-s\tau_b\}]$ as

$$1 = \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(\tau_b) - s\tau_b\}] = \mathsf{E}[\exp\{\theta b - s\tau_b\}],$$

yielding

$$\mathsf{E}[\exp\{-s\tau_b\}] = e^{-\theta b}.$$

However, there are some problems with the heuristic argument; for example, why the optional sampling theorem holds, and which θ to use, as there are two roots to the quadratic equation (7.3).

The above equation can be justified rigorously and θ can be chosen properly as fellows: By the optional sampling theorem, for any $t > 0$,

$$1 = \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(t \wedge \tau_b) - s(\tau_b \wedge t)\}] \tag{7.4}$$

With $s > 0$ and

$$\theta > 0,$$

the term inside the expectation is dominated by

$$\begin{aligned} 0 &\leq \exp\{\theta W_{\mu,\sigma}(t \wedge \tau_b) - s(\tau_b \wedge t)\} \\ &\leq \exp\{\theta b - s(\tau_b \wedge t)\} \leq \exp\{\theta b\}, \end{aligned}$$

independent of t. Therefore, by the dominated convergence theorem, we can let $t \to \infty$ in (7.4) to get

$$\begin{aligned} 1 &= \lim_{t\to\infty} \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(t \wedge \tau_b) - s(t \wedge \tau_b)\} \cdot I\{\tau_b = \infty\}] \\ &\quad + \lim_{t\to\infty} \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(t \wedge \tau_b) - s(t \wedge \tau_b)\} \cdot I\{\tau_b < \infty\}] \\ &= \lim_{t\to\infty} \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(t) - s(t)\} \cdot I\{\tau_b = \infty\}] \\ &\quad + \lim_{t\to\infty} \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(t \wedge \tau_b) - s(t \wedge \tau_b)\} \cdot I\{\tau_b < \infty\}] \\ &= \mathsf{E}\left[\lim_{t\to\infty} (\exp\{\theta W_{\mu,\sigma}(t) - st\} I\{\tau_b = \infty\})\right] \\ &\quad + \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(\tau_b) - s(\tau_b)\} \cdot I\{\tau_b < \infty\}] \\ &= \mathsf{E}[\exp\{\theta W_{\mu,\sigma}(\tau_b) - s\tau_b\} \cdot I\{\tau_b < \infty\}]; \end{aligned}$$

note in the second to the last equation the first term is zero because $\text{limsup} W(t) \leq b$ on the set $\{\tau_b = \infty\}$ and, again, $\theta > 0$. Since $W_{\mu,\sigma}(\tau_b) = b$ on the set $\tau_b < \infty$, we further have

$$\mathsf{E}[\exp\{\theta b - s\tau_b\}] = \mathsf{E}[\exp\{\theta b - s\tau_b\} \cdot I\{\tau_b < \infty\}] = 1.$$

In summary, we have

$$\mathsf{E}[e^{-s\tau_b}] = e^{-\theta b}.$$

However, since θ is a positive solution of a quadratic equation we have

$$\theta = \frac{1}{\sigma^2}\left(\sqrt{\mu^2 + 2\sigma^2 s} - \mu\right) > 0.$$

Thus,

$$\mathsf{E}[e^{-s\tau_b}] = \exp\{-\frac{b}{\sigma^2}\left(\sqrt{\mu^2+2\sigma^2 s}-\mu\right)\}. \tag{7.5}$$

Note that

$$\lim_{s\to 0}\mathsf{E}[e^{-s\tau_b}] = \lim_{s\to 0}\mathsf{E}[e^{-s\tau_b}\cdot I\{\tau_b<\infty\}] = \mathsf{E}[\lim_{s\to 0} e^{-s\tau_b}\cdot I\{\tau_b<\infty\}] = \mathsf{P}\left(\tau_b<\infty\right),$$

where we have used the dominated convergence theorem to interchange the limit and the expectation. As a by-product, we also have

$$\mathsf{P}\left(\tau_b<\infty\right) = \lim_{s\to 0}\mathsf{E}[e^{-s\tau_b}] = \exp\{-b\left(\sqrt{\mu^2}-\mu\right)/\sigma^2\}.$$

Since $\sqrt{\mu^2}=|\mu|$, we get it again that

$$\mathsf{P}\left(\tau_b<\infty\right) = \begin{cases} \exp\left(2\mu b/\sigma^2\right) < 1, & \text{if } \mu b<0; \\ 1, & \text{otherwise.} \end{cases}$$

One can also get the moments from the Laplace transform such as

$$\mathsf{E}[\tau_b] = -\lim_{s\to 0}\frac{d}{ds}\mathsf{E}[e^{-s\tau_b}].$$

Generally speaking, one can only numerically invert the Laplace transform to compute the density function. However, in this special case of the first passage time, the Laplace transform can be inverted analytically.

Exercise 1. Consider a geometric Brownian motion

$$Z_t = \exp\{\gamma t+\beta W_t\}, \quad \gamma>0.$$

Define

$$\tau = \inf\{t\geq 0 : Z_t = 1+\lambda\}, \quad \lambda>0.$$

Show that

$$E[\tau] = \frac{1}{\gamma}\ln\left(1+\lambda\right).$$

Note that we need the assumption $\gamma>0$ in the above statement; otherwise, if $\gamma\leq 0$, then $E[\tau]=\infty$.

The Laplace inversion table (available in many handbooks of mathematical tables) gives the Laplace inversion of the function $\exp\{-a\sqrt{s}\}$, $a>0$, as

$$\frac{a}{2\sqrt{\pi}}\frac{1}{x^{3/2}}\exp\{-\frac{a^2}{4x}\}.$$

Use this result to invert the Laplace transform in (7.5) one can show that the density of the τ_b is given by

$$f_{\tau_b}(x) = \frac{b}{\sigma\sqrt{2\pi x^3}}\exp\{-\frac{(b-x\mu)^2}{2x\sigma^2}\}.$$

The derivation of the Laplace transform only relies on finding a martingale, applying the optional sampling theorem, and the fact $W_{\mu,\sigma}(\tau_b) = b$. This makes it more suitable to solve more complied first passage time problems. In particular, we shall see later that we can derive Laplace transforms to first passage times of some jump diffusion processes.

7.7.2 Expected Time to Reach a Goal

Now we ask the following question: If we use three strategies, all cash ($\pi = 0$), all stock ($\pi = 1$), and the optimal π^*, how long does it take for us to double our investment.

Note that from (7.1) we have

$$X_t^\pi / X_0 = \exp\left\{\left(r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}\right)t + \pi\sigma W_t\right\}.$$

Thus, if the goal is to double the original investor mean, we have $\lambda = 1$, and

$$\gamma = r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}$$

in Exercise 1.

Use

$$\mu = 0.15,\ \sigma = 0.30,\ r = 0.07,$$

we have

$$\pi^* = \frac{0.15 - 0.07}{0.3^3} = 0.89.$$

We have the expected time to double, $E[\tau]$, according to Exercise 1 as follows.

Strategy	Expected Time to Double
Cash ($\pi = 0$)	9.90 years
All Stock ($\pi = 1$)	6.60 years
Kelly Optimal ($\pi = \pi^* = 0.89$)	6.57 years

Next we ask the second question, how long does it take for the optimal strategy to beat the all cash or all stock strategy by 10%. To do this, we need to simplify $X_t^{\pi^*}/X_t^\pi$.

Note that

$$
\begin{aligned}
& X_t^{\pi^*}/X_t^{\pi} \\
= \; & \exp\left\{\left(r+\pi^*(\mu-r)-\frac{(\pi^*)^2\sigma^2}{2}\right)t+\pi^*\sigma W_t-\left(r+\pi(\mu-r)-\frac{\pi^2\sigma^2}{2}\right)t-\pi\sigma W_t\right\} \\
= \; & \exp\left\{\left((\pi^*-\pi)(\mu-r)-\frac{\left\{(\pi^*)^2-\pi^2\right\}\sigma^2}{2}\right)t+(\pi^*-\pi)\sigma W_t\right\} \\
= \; & \exp\left\{(\pi^*-\pi)\left((\mu-r)-\frac{\{\pi^*+\pi\}\sigma^2}{2}\right)t+(\pi^*-\pi)\sigma W_t\right\} \\
= \; & \exp\left\{(\pi^*-\pi)\left(\sigma^2\pi^*-\frac{\{\pi^*+\pi\}\sigma^2}{2}\right)t+(\pi^*-\pi)\sigma W_t\right\} \\
= \; & \exp\left\{(\pi^*-\pi)\left(\frac{\{\pi^*-\pi\}\sigma^2}{2}\right)t+(\pi^*-\pi)\sigma W_t\right\} \\
= \; & \exp\left\{\frac{1}{2}(\pi^*-\pi)^2\sigma^2 t+(\pi^*-\pi)\sigma W_t\right\}.
\end{aligned}
$$

Thus,

$$
X_t^{\pi^*}/X_t^{\pi}=\exp\left\{\frac{1}{2}(\pi^*-\pi)^2\sigma^2 t+(\pi^*-\pi)\sigma W_t\right\}. \tag{7.6}
$$

Therefore, to answer this question we can use Exercise 1 with

$$
\gamma=\frac{1}{2}(\pi^*-\pi)^2\sigma^2.
$$

More precisely, the expected time for the Kelly strategy to beat the strategy π by ε is given by

$$
\frac{2}{(\pi^*-\pi)^2\sigma^2}\ln(1+\varepsilon).
$$

Using $\varepsilon = 10\%$, we have the following table.

Strategy	Expected time to beat by 10%
All Cash ($\pi = 0$)	2.7 years
All Stock ($\pi = 1$)	172.0 years

7.7.3 High Probability to Reach a Goal

The above calculation only cares about the expected time to beat a strategy. However, if we ask a different question about how long does it take for us to make sure with 95% chance to beat a strategy by 10%, we will get a totally different answer.

By (7.6)

$$\begin{aligned}
& P(X_t^{\pi^*}/X_t^{\pi} \geq 1+\varepsilon) \\
= & P\left(\exp\left\{\frac{1}{2}(\pi^*-\pi)^2\sigma^2 t + (\pi^*-\pi)\sigma W_t\right\} \geq 1+\varepsilon\right) \\
= & P\left((\pi^*-\pi)\sigma W_t \geq -\frac{1}{2}(\pi^*-\pi)^2\sigma^2 t + \ln(1+\varepsilon)\right) \\
= & \Phi\left(\frac{\frac{1}{2}(\pi^*-\pi)^2\sigma^2 t - \ln(1+\varepsilon)}{\sqrt{(\pi^*-\pi)^2\sigma^2 t}}\right) \\
= & \Phi\left(\frac{1}{2}M - \frac{\ln(1+\varepsilon)}{M}\right),
\end{aligned}$$

where $M = \sqrt{(\pi^*-\pi)^2\sigma^2 t}$.

To ensure the above probability to be $1-\alpha$, we can set

$$\frac{1}{2}M - \frac{\ln(1+\varepsilon)}{M} = z_\alpha,$$

where $\Phi(z_\alpha) = 1-\alpha$. Solving this quadratic equation and finding the positive root we have

$$M = z_\alpha + \sqrt{z_\alpha^2 + 2\ln(1+\varepsilon)}.$$

This implies that

$$t = \left(\frac{z_\alpha + \sqrt{z_\alpha^2 + 2\ln(1+\varepsilon)}}{(\pi^*-\pi)\sigma}\right)^2.$$

The next table reports t, the time need for the optimal Kelly strategy to beat another strategy π by 10% with high probability.

Probability		Time (in years) need to beat π by 10%	
$1-\alpha$	z_α	All cash ($\pi = 0$)	All stock ($\pi = 1$)
0.90	1.28	98	6,250
0.95	1.645	158	10,080
0.99	2.33	310	19,824

7.8 Solving a Goal Problem

Consider a problem of starting from \$$x$ and try to beat the goal of \$$b$ at time T. More precisely, we have the following stochastic control problem

$$V(t,x;b) = \max_{\pi} P\left(X_T^{\pi} \geq b | X_t = x\right).$$

To make the problem nontrivial, we must have

$$xe^{rT} \leq b;$$

otherwise, if the inequality fails, one can simply put all the money in cash, and have a payoff xe^{rT} at time T.

Solving this control problem yields the optimal proportional amount of money invested in the stock is

$$\pi_t^* = \frac{be^{-r(T-t)}}{x}\phi\left(\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right)\right)\frac{1}{\sigma\sqrt{T-t}}, \tag{7.7}$$

which leads to the optimal value function

$$V(t,x) = \Phi\left(\Phi^{-1}\left(\frac{x}{b}e^{r(T-t)}\right) + \frac{\mu - r}{\sigma}\sqrt{T-t}\right).$$

This result can be proved in several different ways, e.g. solving the HJB nonlinear PDE (Browne 1999), using the Neyman-Pearson lemma (see Kulldorff 1993 and Heath 1993), and the duality results (Spivak and Cvitanic 1999).

By using (7.6) we can solve the following more general result

$$V(t,w;\varepsilon,\pi) = \max_f P\left(X_T^f \geq (1+\varepsilon)X_T^\pi | \left(X_t^f / X_t^\pi\right) = w\right),$$

for any constant proportion stratergy π. Indeed the solution has the optimal proportion

$$f^* = \pi + \left(\frac{1+\varepsilon}{w}\right)\phi\left(\Phi^{-1}\left(\frac{w}{1+\varepsilon}\right)\right)\frac{1}{\sigma\sqrt{(T-t)}},$$

and the optimal value function

$$V(t,w;\varepsilon,\pi) = \Phi\left(\Phi^{-1}\left(\frac{w}{1+\varepsilon}\right) + \sigma\sqrt{(\pi^* - \pi)^2(T-t)}\right). \tag{7.8}$$

For a given strategy π and a given excess level ε, we set the optimal value to be $1-\alpha$ and get the implied T to achieve the desired probability α. Indeed from (7.8) we get

$$\Phi\left(\Phi^{-1}\left(\frac{w}{1+\varepsilon}\right) + \sigma\sqrt{(\pi^* - \pi)^2(T-t)}\right) = 1-\alpha,$$

i.e.

$$\Phi^{-1}\left(\frac{w}{1+\varepsilon}\right) + \sigma\sqrt{(\pi^* - \pi)^2(T-t)} = \Phi^{-1}(1-\alpha),$$

$$T - t = \left(\frac{\Phi^{-1}(1-\alpha) - \Phi^{-1}\left(\frac{w}{1+\varepsilon}\right)}{\sigma(\pi^* - \pi)}\right)^2.$$

Letting $t = 0$ and $w = 1$, we have

$$T = \left(\frac{\Phi^{-1}(1-\alpha) - \Phi^{-1}\left(\frac{1}{1+\varepsilon}\right)}{\sigma(\pi^* - \pi)}\right)^2.$$

The next table reports T, the time need for the optimal goal beating strategy to beat another strategy π by 10% with high probability.

Probability	Time (in years) need to beat π by 10%	
$1-\alpha$	All cash ($\pi = 0$)	All stock ($\pi = 1$)
0.90	0.04 (15 days)	2.6
0.95	1.35	86
0.99	14.0	884

This table is a significant improvement over the previous table.

7.9 Connection with Digital Options

There is an interesting connection between the above goal problem and a digital option, as pointed out in Browne (2000) and also Spivak and Cvitanic (1999), mainly due to the duality between utility maximization and option pricing via Legendre-Fenchel transform, which we will study in a later chapter. For now, we simply briefly point out such a connection.

Recall that the goal problem is

$$V(t, x; b) = \max_{\pi} P\left(X_T^{\pi} \geq b | X_t = x\right).$$

Now consider a digital option that pay \$$b$ at time T if $X_T \geq K$, where K is the strike price. Note that b is given but K needs to be determined.

The price of the digital option with strike price K is given by

$$C(t, X_t) = be^{-r(T-t)}\Phi\left(\frac{\ln(X_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right),$$

and the hedging strategy is to hold

$$\begin{aligned} & \frac{\partial}{\partial X_t} C(t, X_t) \\ = \quad & be^{-r(T-t)}\phi\left(\frac{\ln(X_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)\frac{1}{X_t\sigma\sqrt{T-t}} \end{aligned}$$

shares of the stock at time t, where ϕ is the density function of $N(0,1)$.

Choose K^* such that the price of this is exactly equal to the initial wealth x, we have

$$x = be^{-r(T-t)}\Phi\left(\frac{\ln(X_t/K^*) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right),$$

i.e.

$$\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right) = \frac{\ln(X_t/K^*) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$K^* = X_t \exp\left\{(r - \frac{1}{2}\sigma^2)(T-t) - \sigma\sqrt{T-t}\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right)\right\}.$$

Note that

$$\ln(X_t/K^*) = -(r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right). \tag{7.9}$$

Then the hedging strategy using K^* is to hold

$$be^{-r(T-t)}\phi\left(\frac{\ln(X_t/K^*) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)\frac{1}{X_t\sigma\sqrt{T-t}}$$

shares of the stock at time t, i.e. to hold

$$\begin{aligned} & be^{-r(T-t)}\phi\left(\frac{-(r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}\right)\frac{1}{X_t\sigma\sqrt{T-t}} \\ = \; & be^{-r(T-t)}\phi\left(\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right)\right)\frac{1}{X_t\sigma\sqrt{T-t}} \end{aligned}$$

shares of the stock at time t, via (7.9). Thus, the total amount of money invest in the stock is

$$be^{-r(T-t)}\phi\left(\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right)\right)\frac{1}{\sigma\sqrt{T-t}}.$$

Since the hedging strategy is self financing, the total wealth is exactly x. Thus, the proportional amount of money invested in the stock is

$$\pi_t^* = \frac{be^{-r(T-t)}}{x}\phi\left(\Phi^{-1}\left(\frac{xe^{r(T-t)}}{b}\right)\right)\frac{1}{\sigma\sqrt{T-t}},$$

which is exactly (7.7).

7.10 Borrowing Region

A main drawback of the goal beating strategy is that it may lead to high leverage, as will be shown below.

Introduce z as the percentage of the goal reached by time t,

$$z = \frac{x}{be^{-r(T-t)}},$$

and

$$\tau = \sigma^2\,(T-t)\,, \quad v = \Phi^{-1}(z).$$

where τ is called risk-adjusted remaining time and $z = \Phi(v)$. Then we have from (7.7)

$$\pi_t^* = \frac{1}{z}\phi\left(\Phi^{-1}\left(z\right)\right)\frac{1}{\sqrt{\tau}} = \frac{1}{\sqrt{\tau}}\frac{1}{z}\phi(v) = \frac{1}{\sqrt{\tau}}\frac{\phi(v)}{\Phi(v)}.$$

Since $\phi(v)/\Phi(v)$ is a decreasing function in v, we have

$$\pi_t^* > 1 \Longleftrightarrow z < z^*,$$

i.e. the trading strategy is to borrow money ($\pi_t^* > 1$) if and only if

$$z < z^*,$$

where z^* solves the equation

$$1 = \frac{1}{\sigma\sqrt{T-t}} \left(\frac{\phi(\Phi^{-1}(z))}{\Phi(\Phi^{-1}(z))} \right).$$

If we define v^* as the root (which is unique due to the monotonicity of $\phi(v)/\Phi(v)$) of

$$\frac{\phi(v)}{\Phi(v)} = \sigma\sqrt{T-t},$$

then

$$z^* = \Phi(v^*).$$

Below is a table, taken from Browne (1999), for the numerical values of z^* versus $\tau = \sigma\sqrt{T-t}$, which shows the borrowing region.

τ	z^*	τ	z^*	τ	z^*
0.001	0.99	0.50	0.56	1	0.38
0.05	0.88	0.55	0.54	1.5	0.27
0.10	0.82	0.60	0.51	2.00	0.19
0.15	0.77	0.65	0.49	2.50	0.14
0.20	0.73	0.70	0.48	3.00	0.10
0.25	0.70	0.75	0.46	3.50	0.08
0.30	0.67	0.80	0.44	4.00	0.06
0.35	0.64	0.85	0.43	4.50	0.04
0.40	0.61	0.90	0.41	5.00	0.03
0.45	0.58	0.95	0.39		

As shown above, the borrowing region can be very significant when the time horizon is short.

In summary, with different objective functions, the optimal trading strategies are different. The Kelly criterion may look very good in terms of maximizing the log utility or the median of the investment outcome, but may look bad from other objectives, e.g. maximizing the chance of the beating a benchmark goal. The strategy of doing the latter one may involve heavy leverage, which may make the strategy more suitable for investors who can access the power of leverage easier, e.g. institutional investors.