Contributions to the Theory of Kelly Betting with Applications to Stock Trading: A Control-Theoretic Approach

by

Chung-Han Hsieh

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Electrical and Computer Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2019

Date of final oral examination: 12/17/2019

The dissertation is approved by the following membmers of the Final Oral Committee:

B. Ross Barmish, Profeesor, Electrical and Computer Engineering

John A. Gubner, Professor, Electrical and Computer Engineering

James A. Primbs, Associate Professor, Finance, California State University, Fullerton

David P. Brown, Professor, School of Business

Andreas Seeger, Professor, Mathematics.

To Shun-Ying, Ya-Nuo and Ya-Jie

ACKNOWLEDGMENTS

In my later undergraduate years in Taiwan, my dream was to become a control theorist. The path to the realization of this dream was indeed full of challenges, but I was lucky enough to have great help from many wonderful people along the way.

I would like to first express my sincere appreciation to my dissertation advisor, Professor B. Ross Barmish, for his advice, guidance, and unending patience during the years we worked together. I vividly remember uncountably many discussions, debates, and brainstorming sessions on many exciting research problems. From him, I am still learning how to think critically and communicate effectively. His wisdom, integrity and high standards make him my role model as a researcher.

I am grateful to Professor John Gubner for his considerable help during my graduate studies. It was my good fortune to collaborate with him on several research projects during which I learned a lot from his thoughtful feedback and the discussions we had. I especially wish to thank him for his willingness to read many of my research drafts and provide extensive comments. His kindness, patience, and knowledge will always be fondly remembered.

I am also grateful to Professors David Brown, James Primbs and Andreas Seeger for their years of continuous service on my dissertation committee. Each of them also played an important role in my academic life at UW-Madison. Professor Brown taught the first course which I took on Investment Theory. As a student in engineering, many concepts were entirely new to me at the time and it was his teaching which made it possible for me to learn the material. His wisdom on the subject and thoughtful feedback will be remembered. I also wish to thank Professor James Primbs for his valuable feedback and support. Some of his papers with Professor Barmish using control-theoretic tools in financial engineering problems inspired me in the early stages of my research. Finally, I wish to thank Professor

Andreas Seeger who taught me Measure Theory and Real Analysis. These courses were challenging and motivated me to pursue an additional M.A. degree in mathematics along the way to my doctoral dissertation.

I am also grateful to many others who supported me in the course of my graduate work. The Departments of Mathematics and Electrical and Computer Engineering provided financial support through teaching assistantships and instructorships. Specifically, I would like to thank Professors Bing Wang, Alexandra Kjuchukova, Saverio Spagnolie, Yu Hen Hu, Barry Van Veen, Varun Jog, Kangwook Lee, and Matthew Malloy, who gave me the opportunity to serve as their Teaching Assistant with special gratitude to Professor Barry Van Veen for his mentorship during my university teaching career. He showed me the power of the flipped classroom, but more importantly, he showed me the importance of caring about students' development and future.

Throughout the entire course of my studies, my beloved family members in Taiwan, my parents, my aunt and uncles, sisters and brothers were highly supportive. My deepest and greatest appreciation goes to my wife, Shun-Ying, for her unconditional love and support throughout these years; she gave me the courage to persevere through tough times. Without her encouragement, tolerance and support, my graduate work would not have been possible. This dissertation is dedicated to her, and our son Ya-Nuo and our newest family member Ya-Jie.

Finally, I would like to thank God for leading me through this long journey and making the seemingly impossible become possible.

Chung-Han Hsieh
Madison, Wisconsin

TABLE OF CONTENTS

Abstract

Kelly Betting is a prescription for optimal resource allocation among a sequence of gambles which are typically repeated in an independent and identically distributed manner. Within this setting, the theory is aimed at maximizing the expected value of the logarithmic growth of wealth. Many papers in the existing literature indicate that such a maximization leads to a number of desirable properties. These include *superior long-term growth of wealth*, *competitive optimality* and a certain *myopic property*. This betting scheme has also been criticized as being too aggressive with respect to various risk metrics. To address this, many papers suggest ad-hoc ways for scaling down the bet size. In our first collection of results, we provide a new perspective on this aggressiveness issue. That is, we show that in some cases, the Kelly optimum may actually lead to bets which are too conservative rather than too aggressive. To make this more precise, we provide a result which we call the *Restricted Betting Theorem*. Subsequently, we point out some additional negatives of the Kelly-based theory by quantifying what difficulties are encountered with various approximations which are used in some of the literature. Throughout this dissertation, we emphasize the feedback control system point of view and the ramification of our results in the context of stock trading.

Following the initial results above, we report on our research aimed at improving the existing Kelly-based theory. Our second collection of results, which we call *Drawdown-Modulated Betting*, is focused on mitigating the potentially large drawdown for a rather general class of betting schemes including the classical Kelly Betting scheme as special case. Motivated by the fact that this issue is of paramount concern from a risk management perspective, we prove a result, called the *Drawdown Modulation Lemma*, which characterizes investment strategies guaranteeing that the percentage drawdown is no greater than a prespecified level for all sequences of admissible returns. With the aid of this lemma, we show that investment functions can be expressed as a linear time-varying feedback control parameterized by a feedback gain and leading to satisfaction of the drawdown specification. Subsequently, a generalization of the lemma to the portfolio setting is also provided. In addition, with the risk-reward pair being drawdown and expected return, we prove that

the drawdown-modulated feedback strategy "dominates" the classical linear time-invariant (LTI) feedback strategy. In the parlance of finance, the LTI strategy is said to be *inefficient*.

The third collection of results in this dissertation, called *Frequency-Based Betting*, is focused on investigating how optimization and expected logarithmic growth performance vary with respect to betting frequency and on how our formulation and results apply to the stock market. Going beyond existing literature, in this part of the work, the *frequency*, or equivalently the number of stages n between trades, is included as an additional optimization parameter in our analysis. For a single stock, in the absence of transaction costs, we show that high-frequency trading is *unbeatable* in the sense of expected logarithmic growth. Moreover, we prove that if a stock satisfies a certain "sufficient attractiveness" condition, then the buy-and-hold strategy with $n > 1$ can match the performance of the high-frequency strategy with $n = 1$. Subsequently, when we generalize the notion of sufficient attractiveness from the single-stock case to a portfolio with multiple risky assets, a similar result is obtained. One highlight in this part of the dissertation involves the notion of a "dominant asset" which we define. When such an asset is present in the portfolio, we prove that the optimal performance requires putting "all eggs in one basket." As a consequence, we see that the performance of the high-frequency trader is matched by that of the buy and holder.

The final collection of results in this dissertation is motivated by the fact that a trader's interactions with the market are not instantaneous. This leads us to extend our frequency-based framework to include *delay* in trade execution. For the case when a single unit of delay is present, in contrast to existing literature on Kelly Betting, it turns out that bankruptcy is a distinct possibility. This leads to a problem formulation in which the no-bankruptcy issue is cast as a *state positivity* problem. Subsequently, we prove two theorems. The first theorem gives sufficient conditions for avoidance of bankruptcy and the second gives necessary conditions. Some other technical results regarding state positivity are given as enrichments to the theory; e.g., we provide an example which suggests that when delay is present, the buy-and-hold strategy can achieve strictly higher performance than high-frequency trading.

Chapter 1

Overview of Dissertation Research

In the mid-1950s, John Kelly at Bell Labs published a seminal paper [1] which uses the Expected Logarithmic Growth (ELG) criterion to study gambling systems.[1] His work provides a prescription for optimal betting when faced with a sequence of gambles which are modeled in terms of independent and identically distributed (i.i.d.) random variables; see [1–5]. With the ELG criterion serving as the takeoff point, this chapter provides an overview of the dissertation research, which generalizes various aspects of the existing theory of Kelly Betting. Many of the formulations, methods and results is developed from a control-theoretic point of view. In Section 1.1, we provide a brief literature survey and an introduction to classical scalar Kelly Betting and its motivation. In Sections 1.2 and 1.3, a general Kelly Betting framework and its extension to a more general stock-trading portfolio case is discussed, with an emphasis on a control system point of view. Subsequently, in Section 1.4, an overview of the contributions of the dissertation is provided. Finally, some concluding remarks are provided.

1.1 Introduction to Kelly Betting

Kelly Betting is a prescription for optimal resource allocation among a sequence of gambles which are typically repeated in an i.i.d. manner. The theory involves maximization of the expected value of the logarithmic growth of a bettor's account value. To provide one of the simplest possible illustrations of the key ideas, we consider a game consisting of N repeated

[1] John Kelly was a colleague and collaborator of Claude Shannon, the founder of information theory. The history to this work and wonderful anecdotes behind the development of the theory can be found in [2].

independent coin flips. On the k-th flip, the random variable $X(k)$ corresponds to an even-money payoff; i.e., we have $X(k) \in \{-1, 1\}$ for $k = 0, 1, \ldots, N-1$, where $X(k) = 1$ stands for heads and $X(k) = -1$ stands for tails. In addition, it is assumed that the coin is biased with probability of heads being $p > 1/2$ which is known. Hence, the gambler has an edge and specifies in advance the fraction $K \in [0, 1]$ of the account value wagered at each stage. Letting $V(k)$ denote this account value after k flips, the next wager is $KV(k)$ and the update from k to $k+1$ is given by the recursion

$$V(k+1) = (1 + KX(k))V(k).$$

Now, for this simple special case of the Kelly-based theory, a basic question drives the analysis to follow: *What is the optimal fraction $K \in [0, 1]$ maximizing the expected logarithmic growth*

$$g(K) \doteq \frac{1}{N}\mathbb{E}\left[\log \frac{V(N)}{V(0)}\right]?$$

This simple special case of Kelly Betting is easy to solve. The optimal betting fraction K, call it K^*, is given by

$$K^* = 2p - 1.$$

For example, if the probability of heads is $p = 0.6$, then $K^* = 0.2$ implies that 20% of the account value should be bet on each flip.

1.1.1 Remarks: It is also important to observe that the expected logarithmic growth $g(K)$ is a concave function of K and and the same holds true when the formulation above is generalized to multiple fractions K_i and rather arbitrary probability distributions. The maximization of $g(K)$ is also called the *Kelly Criterion.* Whereas concavity is arguably unimportant when K is scalar, it becomes very important when the more general multi-coin gambling or a portfolio optimization in the stock market is considered in this dissertation. That is, when K is a vector, concavity facilitates computation of the optimal components K_i.

1.1.2 Big Picture in the Literature: Following the pioneering work in [1], many applications were developed, and a number of properties of Kelly Betting schemes were studied over the subsequent decades. For example, it is well-known that it has "superior" long-term growth properties for $V(k)$; e.g., see [3, 6–13]. In addition, finding the Kelly optimum is accomplished in a so-called *myopic manner*, which aids the maximization process; e.g., see [14] and [15]. Roughly speaking, this means that at each stage k, one can carry out a single-stage logarithmic growth optimization with respect to K as if it is the only bet under consideration. It is also known that the expected logarithmic growth criterion produces a strategy which leads to satisfaction of a so-called *competitive optimality* condition; i.e., under mild hypotheses, a bettor maximizing the expected logarithmic growth will outperform a competitor using an "essentially different" strategy with probability tending to one as the number of stages N goes to infinity; see [6–9, 16–18].

The expected logarithmic growth criterion is not only fundamental to gambling theory, but also a starting point for a line of research on portfolio optimization in the stock market; e.g., see [2, 4, 9–13, 18–23]. The problem of sequential investment in the stock market was first studied in [24] and a survey of Kelly-based strategies which also includes blackjack, sports betting and as well as the stock market can be found in [25]. The Kelly criterion in a continuous-time framework using Ornstein-Uhlenbeck process is studied in [26] and continuous-time long-run logarithmic growth issues with transaction costs consideration are considered in [27] and [28]. The fact that the logarithmic growth criterion can be used for pricing portfolio is discussed in [29] and [30] where the risky assets with prices governed by geometric Brownian motion. Finally, a sampling of more recent papers on the topic of Kelly Betting includes [31–40].

1.1.3 Further Motivation: The study of expected logarithmic growth is further motivated via a comparison with results obtained via the classical optimization of the expected value of the terminal account value $\mathbb{E}[V(N)]$. Indeed, for the coin-flipping gamble discussed above, with $p > 1/2$, instead of using logarithmic growth, suppose we seek to find an optimal fraction K which maximizes $\mathbb{E}[V(N)]$. Then, it is straightforward to show that the optimum irrespective of the value of p, is $K = 1$. Such a strategy is arguably far too aggressive to use for a game which is being played over and over again. With N large, it is almost certain that gambler's ruin will occur. In contrast to maximizing $\mathbb{E}[V(N)]$, the Kelly criterion, maximizing $\mathbb{E}[\log V(N)]/N$, "automatically" factors in some degree of risk. For example, for the case of the coin above with $p = 1/2 + \varepsilon$ and small $\varepsilon > 0$, the optimal betting fraction is $K^* = 2\varepsilon$, thereby making it much less aggressive and more in line with common sense.

1.1.4 Kelly Criterion Versus Mean-Variance Criterion: In existing literature, the analysis of mean and variance is typically used to compare various bets; e.g., as illustrated by [41–43], in modern portfolio theory, these metrics are widely used. The following simple example illustrates why some bettors prefer to use the Kelly criterion rather than the mean-variance criterion.

Consider two different gambles defined as follows: For the first gamble, we let random variable $X_1 \in \{-1,\ 3/2\}$ denote the *return* with probability

$$P(X_1 = -1) = 1/5 \text{ and } P(X_1 = 3/2) = 4/5.$$

Now, suppose a gambler decides to enter this first gamble with initial account value $V(0)$. Then the associated account value for next stage becomes $V(1) = V(0) + 3V(0)/2$ if gambler wins; otherwise, $V(1) = 0$. On the other hand, for the second gamble, we let random variable $X_2 \in \{-2/3, 8/5\}$ denote the return with probability

$$P(X_2 = -2/3) = 9/34 \text{ and } P(X_2 = 8/5) = 25/34.$$

Now, with initial account value $V(0) > 0$, assume that we are only allowed to bet a fraction $K \in [0, 1]$ for each of these gambles and assume that $N = 1$; i.e., only a single bet is to be made. Our goal is to decide which one is preferable to the gambler. We begin by noting that a straightforward calculation leads to

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] = 1$$

and

$$\mathrm{var}[X_1] = \mathrm{var}[X_2] = 1.$$

This shows that these two bets are indistinguishable from a mean-variance point of view. However, if we use the Kelly criterion, we claim that the first gamble with return X_1 is preferable to the second gamble with return X_2. To establish this, for X_1, the associated expected logarithmic growth, call it $g_1(K)$, is calculated as

$$\begin{aligned} g_1(K) &= \mathbb{E}\left[\log \frac{V(1)}{V(0)}\right] \\ &= \mathbb{E}\left[\log\left(1 + KX_1\right)\right] \\ &= \frac{1}{5}\log\left(1 - K\right) + \frac{4}{5}\log\left(1 + \frac{3K}{2}\right). \end{aligned}$$

Then, using the fact that $g(K)$ is concave in K, a straightforward calculation leads to a unique global maximum, call it $K = K_1^*$, given by

$$K_1^* = \frac{2}{3}.$$

Finally, to complete our analysis, we substitute K_1^* into $g_1(K)$. It is readily verified that the optimal value, call it g_1^*, is given by $g_1^* \approx 0.3347$. Similarly, for X_2, the associated expected logarithmic growth, call it $g_2(K)$, is calculated as

$$\begin{aligned} g_2(K) &= \mathbb{E}\left[\log\left(1 + KX_2\right)\right] \\ &= \frac{9}{34}\log\left(1 - \frac{2K}{3}\right) + \frac{25}{34}\log\left(1 + \frac{8K}{5}\right). \end{aligned}$$

The optimal fraction, call it K_2^* in $[0,1]$, is given by

$$K_2^* = \frac{15}{16},$$

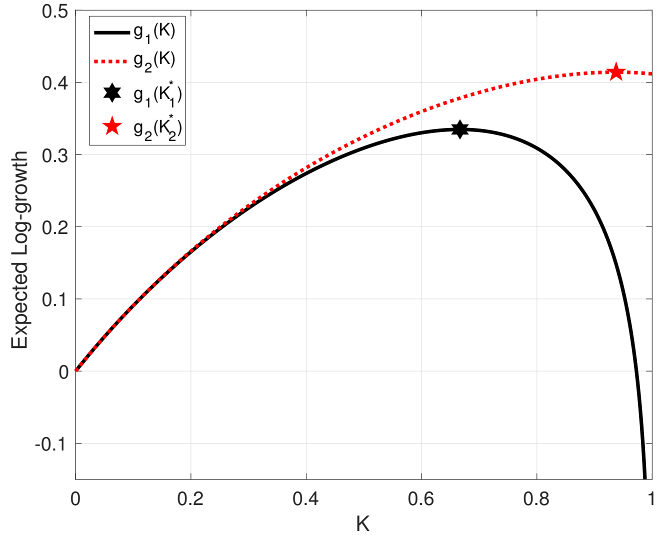


Figure 1.1: Expected Logarithmic Growth for the Two Bets

and the corresponding optimal value, call it g_2^*, is given by $g_2^* \approx 0.4141$. Noting that $g_1^* < g_2^*$, in the Kelly Betting framework, the second bet is deemed preferable to the first. Figure 1.1 shows the plots for $g_1(K)$ and $g_2(K)$ with the optima indicated.

While both K_1^* and K_2^* are within the feasible set $[0, 1]$, this is not always the case for zero-derivative points in the more general Kelly optimization problem. As mentioned in the preceding sections, results from the theory of convex programming are helpful to facilitate computation of the optimum in more general settings.

1.1.5 Some Well-Known Negatives of Kelly Betting: The maximization of expected logarithmic growth often leads to betting which is "too aggressive" with respect to various risk metrics; e.g., see [12, 18, 44]. To remedy this issue, some authors resort to a so-called *fractional* betting scheme which involves scaling down the betting fraction K, which in turn reduces the size of the bet; e.g., see [11, 12, 25, 45]. Other authors resort to incorporation of constraints into the logarithmic growth formulation aimed at reducing the *downside risk* effect; e.g., see [46] where the Value-at-Risk (VaR) or Conditional VaR is incorporated as a constraint into the optimization problem. One of the main negatives associated with such constrained optimization approaches is that the problem to be solved may no longer be concave. Thus, for a strategy with many parameters to be optimized, computational tractability might become a significant issue.

1.2 General Kelly Betting Problem Formulation

Our goal in this section is to describe a generalization of the simple coin-flipping scenario appearing in many of the previously cited papers. The formulation for both betting games and stock trading are provided. Indeed, for $k = 0, 1, 2, \ldots, N-1$, we take the returns

$$X(k) = \begin{bmatrix} X_1(k) & X_2(k) & \cdots & X_m(k) \end{bmatrix}^T,$$

specified by the "house", to be i.i.d. random vectors in $\mathbb{R}^m$ having probability density function, denoted by f_X, and let $\mathcal{X} \subseteq \mathbb{R}^m$ to denote the common support for the $X(k)$. To make the exposition simple, in the sequel, whenever convenient, we use X as a shorthand instead of $X(k)$.

Now, with initial account value $V(0) > 0$ and letting $V(k)$ being the account value at stage k, the bettor's wager for the i-th gamble is $K_i V(k)$ with K_i being the constant gain. Now, we take

$$K \doteq \begin{bmatrix} K_1 & K_2 & \cdots & K_m \end{bmatrix}^T$$

and assume the vector K satisfying the constraint

$$K \in \mathcal{K} \doteq \Big\{ K \in \mathbb{R}^m : K \text{ is such that } V(k) \geq 0 \text{ for all } k \Big\}.$$

The constraint set $\mathcal{K}$ involves the notion of *survival*; i.e., bankruptcy avoidance. That is, any K-value that can potentially lead to $V(k) < 0$ is disallowed. In addition, consistent with existing literature; e.g., see [4, 25, 47], in many cases, as explained later in the dissertation, we often allow the possibility that $K_i < 0$ for $i \in \{1, \ldots, m\}$ so that the theory is flexible enough for the bettor to take either side of the bet or, in the case of the stock market, sell short. Then the dynamics of the account value is described by the recursion

$$\begin{aligned} V(k+1) &= V(k) + K^T X(k) V(k) \\ &= (1 + K^T X(k)) V(k). \end{aligned}$$

1.2.1 Modifications for Stock Market: The formulation above for betting games can be readily adapted to the stock trading. Indeed, we first consider a single stock case. Let $S(k) > 0$ be the underlying stock price at stage k. Then the corresponding return is given by

$$X(k) = \frac{S(k+1) - S(k)}{S(k)}.$$

We assume that the returns are i.i.d. with known bounds satisfying $X_{\min} \leq X(k) \leq X_{\max}$, and $X_{\min}$ and $X_{\max}$ being support points satisfying $-1 < X_{\min} < 0 < X_{\max}$. We further assume that stock trading occurs within an "idealized market." That is, we assume zero transaction costs, zero interest rates and perfect liquidity conditions. These assumptions arise in the finance literature in the context of "frictionless" markets; see [48–50].

Now, a *long* trade is initiated when the trader purchases shares from the broker in the hope of making a profit from a subsequent rise in the price of the stock. To model this mathematically, we begin at some initial account value $V(0) > 0$ and let $V(k)$ be the account value at stage k. Then, for each stage k, the trader selects constant gain $K \geq 0$ and purchases the number of shares $N(k)$ given by

$$N(k) \doteq \frac{KV(k)}{S(k)}.$$

Then, the associated account value dynamics is described by the recursion

$$\begin{aligned} V(k+1) &= V(k) + N(k)(S(k+1) - S(k)) \\ &= V(k) + KV(k)X(k) \\ &= (1 + KX(k))V(k). \end{aligned}$$

Alternatively, a short sale is initiated when the trader borrows shares from a broker in the hope of making a profit from a subsequent fall in the price of the stock. After borrowing the shares, the trader sells them at the market price immediately. Later, the resulting short position is "covered" when the trader repurchases the same number of shares in the market and delivers them back to the broker. If the price declines before covering, the short trader will profit. Conversely, the short position will result in a loss if the stock price rises prior to repurchase. While short selling is considered quite risky by many investors due to the unlimited potential for loss, for example see [30] and [51], it has the potential to improve a portfolio's risk-return trade-off; e.g., see [52–54]. Similar to the long-only case, we again consider single stock with prices $S(k)$ and returns $X(k)$ as described above. In contrast to the long trader with $K > 0$, the short trader selects constant gain $K < 0$ and shorting the number of shares $N(k)$ given by

$$N(k) \doteq \frac{|K|V(k)}{S(k)}.$$

Then, with $K < 0$, the associated account value dynamics updated for $V(k+1)$ is readily shown to be the same as the one derived for the case above.

1.2.2 More General Multi-Stock Case: For a portfolio which involves m stocks with prices $S_i(k) > 0$ for $i = 1, 2, \dots, m$, similar to the above, the associated return is given by

$$X_i(k) = \frac{S_i(k+1) - S_i(k)}{S_i(k)}.$$

Let $V(k)$ be the trader's account value at stage k. Then, for each i-th stock, the trader place an order with the shares

$$N_i(k) \doteq \frac{|K_i|V(k)}{S_i(k)}$$

where $K_i > 0$ means the trader going long and $K_i < 0$ means going short. Now assuming that $K \in \mathcal{K}$, the corresponding account value dynamics recursion is given by

$$\begin{aligned} V(k+1) &= V(k) + \sum_{i=1}^{m} \operatorname{sgn}(K_i) N_i(k)(S_i(k+1) - S_i(k)) \\ &= V(k) + \sum_{i=1}^{m} K_i V(k) X_i(k) \\ &= (1 + K^T X(k)) V(k) \end{aligned}$$

which is the same formula for the account value as seen earlier in gambling situation. As seen in the later sections, whenever convenient, we work with investment of the feedback form $I_i(k) = K_i V(k)$ rather than the number of shares invested $N_i(k)$.

1.2.3 General Kelly Betting Problem: Our objective is to select $K \in \mathcal{K}$ maximizing the expected logarithmic growth

$$g(K) \doteq \frac{1}{N}\mathbb{E}\left[\log\left(\frac{V(N)}{V(0)}\right)\right].$$

We denote the associated optimal value[2]

$$g^* \doteq \sup_{K \in \mathcal{K}} g(K),$$

and $K^* \in \mathcal{K}$ satisfying $g(K^*) = g^*$ is called an *optimal element*. Using the recursion for $V(k)$ and the fact that the $X(k)$ are i.i.d, via a straightforward calculation, as seen in many papers in the literature, the expected logarithmic growth function reduces to

$$\begin{aligned} g(K) &= \frac{1}{N}\mathbb{E}\left[\log\left(\prod_{k=0}^{N-1}\left(1 + K^T X\left(k\right)\right)\right)\right] \\ &= \frac{1}{N}\sum_{k=0}^{N-1}\mathbb{E}[\log\left(1 + K^T X\left(k\right)\right)] \\ &= \mathbb{E}[\ \log(1 + K^T X(0))\ ]. \end{aligned}$$

[2]The supremum can be replaced by maximum operator when the constraint set $\mathcal{K}$ is closed and bounded. This is the typical setting we considered in the later chapters.

Recalling our earlier remarks that the Kelly-based optimization problem is a concave program and the use of some readily available codes such as CVX are available to solve the problem efficiently; e.g., see [55] and [56]. The reader is referred to [9, 34, 46, 57–59] where concave programming is used in the multi-stock portfolio case.

1.2.4 Remarks on Additional Constraint on K: In practice, other than the survival requirement $K \in \mathcal{K}$ as introduced earlier, some additional constraints on K which occur in practice may be imposed; for example, in stock trading context, one may require that $K_i \geq 0$, which corresponds to a so-called *long-only* requirement; e.g., see [9, 30, 60–62]. As a second example, one may also impose the so-called *cash-financing* constraint

$$\sum_{k=1}^{m} |K_i| \leq 1.$$

The condition above assures that the trader's overall investment level can not exceed its own account value. That is,

$$\left| \sum_{k=1}^{m} K_i V(k) \right| \leq \sum_{k=1}^{m} |K_i V(k)| \leq V(k).$$

The use of such constraint can be found in papers such as [60, 63–66]. In the sequel, as seen later in the subsequent chapters, the cash-financing constraint is often imposed when the analysis to follow involved use of historical data and back-testing.

1.3 Feedback Control System Point of View

In this dissertation, the approach which we take to many of the problems under consideration involves a control-theoretic point of view. In this regard, the language we use is consistent with a growing body of the literature addressing finance problems from the control community's point of view; e.g., see [65–81]. Although most of our analysis to follow can be carried out without reference to control theory, the use of feedback systems not only enables us to make the work accessible to engineering researchers with little background in finance, but also allows us to bring ideas from feedback systems into play; good references for control systems can be found in [82] and [83].

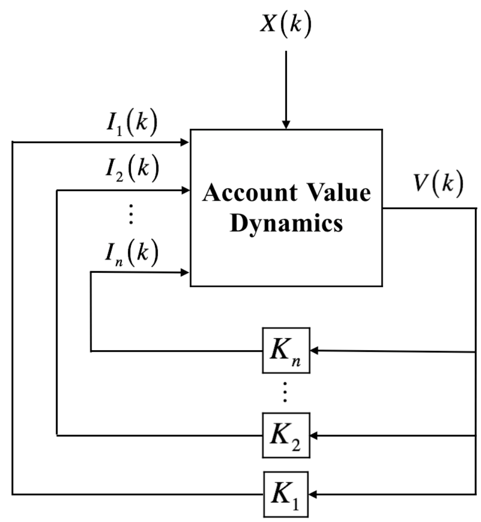


Figure 1.2: Feedback Control Equivalent for Kelly Betting

Specifically, the setup in Section 1.2 can be readily interpreted in terms of a classical feedback control loop. That is, we view $V(k)$ as the *state* of a system with linear time-invariant feedback control. To make this more precise, consistent with the formulation in Section 1.1, for $i = 1, 2, \ldots, m$, the i-th *investment level* is given by

$$I_i(k) = K_i V(k).$$

That is, the i-th component of the control vector $I(k)$ with *feedback gain* K_i. Subsequently, the update of the account value for this stochastic system is again recursive equation $V(k+1) = (1 + K^T X(k))V(k)$ with associated feedback-control configuration depicted in Figure 1.2; see [48, 49, 68, 81, 84–87] where this control-theoretic paradigm is pursued in detail.

1.4 Overview of Dissertation Results

We now provide a brief overview of our results with their details relegated to the chapters to follow.

1.4.1 Conservatism of Kelly Betting: In Chapter 2, we provide a result that Kelly Betting may lead to investment levels which are "too conservative." This is in contrast to the previously mentioned existing literature which only concentrates on situations when this betting scheme is too "aggressive." Our formal result demonstrating this is called the *Restricted Betting Theorem*. The takeoff point for this theorem is motivated by a simple special case: Consider i.i.d. scalar random *returns* $X(k)$ satisfying $X_{\min} \leq X(k) \leq X_{\max}$ with $X_{\min} < 0$ and $X_{\max} > 0$ being points in the support set $\mathcal{X}$. Then, since negative account value cannot be optimal and typically disallowed, the optimal feedback gain $K = K^*$ must satisfy the condition

$$-\frac{1}{X_{\max}} \leq K^* \leq -\frac{1}{X_{\min}}.$$

For the extreme case when the support of the distribution is unbounded both from above and below, we obtain a result indicating that the optimum is not to bet at all. More generally, the Restricted Betting Theorem is apropos in many other scenarios when the Kelly bettor uses "pure theory" in lieu of empirical data. As seen in Chapter 2, the bettor who uses a theoretical model may make wagers which contradict common sense in real-world considerations.

To be more specific, when a formula for the probability density function f_X for the i.i.d. returns $X(k)$ is not available, it is typical to obtain data samples $x_1, x_2, \ldots, x_M$ from f_X and work with the *empirical probability mass function* given by the sum of Dirac Delta functions

$$\widehat{f}_X(x) \doteq \frac{1}{M} \sum_{i=1}^{M} \delta(x - x_i).$$

The reader is referred to [88, 89] as early papers providing an empirical study of stock returns. When dealing with an empirically derived probability mass function, one might ask the following question: If we derive an optimal betting gain based on $\widehat{f}_X$ rather than f_X, how will this optimum, call it $\widehat{K}^*$, compare with a true optimum K^*? Also, how will the optimal performance using $\widehat{K}^*$ compare with the optimum using K^*? In Chapter 2, we show that the optimum K^* using the true theoretical distribution f_X may deviate considerably from that using the empirical distribution $\widehat{f}_X$. An interesting example provided in Section 2.1 shows that if a bet based on empirical data which appears to be a "golden" opportunity is with $\widehat{K}^* \approx 1$, it may be the case that this bet would be rejected when the underlying theoretical model is used; i.e., one obtains $K^* = 0$.

1.4.2 Research on Account Drawdown: In Chapter 3, an important downside risk metric called *drawdown*, which measures the drops in account value $V(k)$ over time from peaks to subsequent lows, is studied. Consistent with the body of existing literature on drawdown, for example, see [43, 90–92], the definition which we use is as follows: For $k = 0, 1, 2, \ldots, N$, we let $V(k)$ be the corresponding account value. Then, as k evolves, the *percentage drawdown to date* is defined as

$$d(k) \doteq \frac{V_{\max}(k) - V(k)}{V_{\max}(k)}$$

where

$$V_{\max}(k) \doteq \max_{0 \le i \le k} V(i).$$

This leads to the overall *maximum percentage drawdown*

$$d^* \doteq \max_{0 \le k \le N} d(k).$$

Note that $0 \le d^* \le 1$. To emphasize the dependence on the constant gain K, below, we may write $d^* = d^*_K$. Although not considered here, there is another well-known drawdown-based measure, called the *maximum absolute drawdown* which is explained in the example to follow. The reader is referred to [74,93,94] for work on this topic. The literature also includes various methodologies to address different types of drawdown-based metrics; see [95–97].

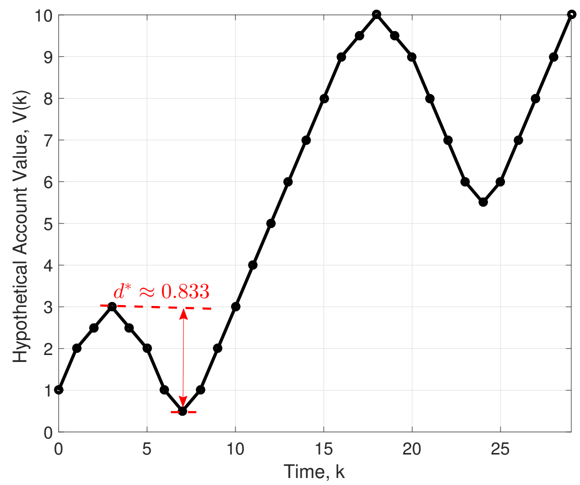


Figure 1.3: Maximum Percentage Drawdown

To illustrate the ideas above, we consider an example with a hypothetical account value $V(k)$ shown in Figure 1.3. From the plot, the overall maximum percentage drawdown

$$d^* = \frac{3 - 0.5}{3} \approx 0.833$$

occurs at stage $k = 7$. Note that this maximum percentage drawdown is not necessarily equal to the maximum absolute drawdown which has value 4.5 and occurs at stage $k = 24$. Percentage drawdown, is generally preferred to absolute drawdown because the scale of betting is taken into account.

To get a rough idea on how bad the drawdown can be, we consider betting on N times on flips of a coin for which $X(k) = 1$ with probability p and $X(k) = -1$ with probability $1 - p$ for $k = 0, 1, \ldots, N - 1$. Assuming that constant gain $K \in [0, 1]$, it is easy to see that the probability of the maximum percentage drawdown being greater than or equal to K is given by

$$P(d_K^* \geq K) = 1 - p^N.$$

To see why this is true, we note that the only way to avoid maximum percentage drawdown K or more is if a gambler wins all N coin flips. So if the gambler loses even once, the drawdown is immediately K. Now, with the aid of the formula above, we can make the following observation: If we take N large, say $N = 252$, and take $p = 0.99$, using the optimal Kelly's fraction $K^* = 2p - 1 = 0.98$, it follows that

$$P(d_K^* \geq 0.98) = 1 - (0.98)^{252} = 0.92.$$

That is, there is at least a 92% chance that the maximum drawdown exceeds 98%. That is, even for a bet which is indisputably attractive, a large drawdown occurs with high probability.

1.4.3 Notion of Complementary Drawdown and Convexity: In Chapter 3, we also provide some preliminary results aimed at mitigating the drawdown problems. Specifically, we work with the *complementary drawdown*

$$\begin{aligned} \overline{D}_K &\doteq 1 - d_K^* \\ &= \min_{0 \leq l \leq k \leq N} \frac{V(k)}{V(l)} \end{aligned}$$

where the subscript K on $\overline{D}$ is used to emphasize the dependence on the feedback gain K. With the aid of the complementary drawdown above, for any $0 < \varepsilon < 1$, we prove that the *surrogate expected maximum drawdown constraint set*, call it $\mathcal{D}_K$, defined by

$$\mathcal{D}_K \doteq \left\{ \ K \in \mathcal{K} : \mathbb{E}[\log \overline{D}_K] \geq \log(1 - \varepsilon) \ \right\}$$

is convex. Therefore, it enables us to formulate a drawdown-constrained Kelly-based optimization problem as concave program which can be solved in an efficient way.

1.4.4 Negative Associated with Approximation: In addition to the negatives associated with drawdown, in Chapter 3, we also fill a void in existing literature by showing that the approximate Kelly optimum obtained via the Taylor-based approximation, for example, see [25, 31], may lead to infeasibility or poor performance compared to the true optimum obtained by concave programming. To provide a simple example, we consider returns $X(k)$ that are scalar i.i.d. random variables. Then recalling that $g(K) = \mathbb{E}[\log(1 + KX(0))]$, according to [25] and [31], one approximation involves treating $X(k)$ as a geometric Brownian motion with drift $\mu = \mathbb{E}[X]$ and variance $\sigma^2 = \text{var}[X]$. Subsequently, using Taylor approximation leads to the associated "optimal" constant gain K, call it $\mathcal{K}_{\text{GBM}}$, is given by

$$\mathcal{K}_{\text{GBM}} = \frac{\mathbb{E}[X(0)]}{\text{var}[X(0)]}.$$

While this approximation technique leads to a closed-form solution, as seen in Chapter 3, it may not be feasible; i.e., the optimum $\mathcal{K}_{\text{GBM}}$ may fail to be in $\mathcal{K}$. One standard way to "fix" this infeasibility problem is to apply a saturation operation to the solution above. For example with $\mathcal{K} = [0, 1]$, instead of using $\mathcal{K}_{\text{GBM}}$ above, one may use

$$K_{\text{GBM}} = SAT\left[\frac{\mathbb{E}[X(0)]}{\text{var}[X(0)]}\right].$$

where $SAT[x]$ is given by

$$SAT\,[x] = \begin{cases} 0 & x < 0; \\ x & 0 \leq x \leq 1; \\ 1 & x > 1. \end{cases}$$

However, suppose we take $X(0) = 0.15$ with probability 0.95 and $X(0) = -0.95$ with probability 0.05. Then it is straightforward to see that $K_{\text{GBM}} = 1$ which leads to a negative expected logarithmic growth $g(K_{\text{GBM}}) = -0.017 < 0$. Ironically, the approximation-based results yields the "minimum" growth of $g(K)$ rather than the desired maximum. Suffice it to say, the combination of approximation and saturation due to constraint violation can lead to very poor performance. The detailed discussion on this topic is given in Chapter 3.

1.4.5 Drawdown-Modulation Control Systems: As mentioned in Section 1.4.2, the control of drawdown is of great concern from a risk management perspective. Suffice it to say, this issue has received considerable attention in the finance literature; e.g., see [93,94,98–101]. Of these papers, references [98–100] are most relevant to our new drawdown-modulation control systems to follow. Although the problem setups and assumptions in these papers differ from ours, they include one basic consideration which is central to our new "drawdown modulation" control scheme described in Chapter 4. That is, the investment level is dynamically controlled as a function of "drawdown to date" $d(k)$ which is defined in Section 1.4.2.

The scheme which we study in Chapter 4 involves controlling the drawdown in the almost-sure sense. Given the returns $X(k)$ satisfying some mild technical conditions, a *maximum acceptable drawdown level* $d_{\max} \in (0,1)$ is specified and we focus on conditions on the investment level $I(k)$ under which satisfaction of the constraint $d(k) \leq d_{\max}$ is assured for all k and all sample paths. In particular, our result, which we call the Drawdown Modulation Lemma, given in Section 4.2, provides a necessary and sufficient condition for this requirement to be satisfied.

With the aid of the Drawdown Modulation Lemma, we construct a linear time-varying feedback control parameterized by a gain γ and leading to satisfaction of the drawdown specification. Specifically, for each stage $k = 0, 1, 2, \ldots, N-1$, we define

$$M(k) \doteq \frac{d_{\max} - d(k)}{1 - d(k)}$$

which we call the *modulator*. Now, using $M(k)$, we express $I(k)$ in the feedback form

$$I(k) \doteq \gamma M(k) V(k)$$

and, as detailed in Section 4.3, γ can be selected without regard for the modulator $M(k)$. This idea has a similar flavor to that of the celebrated Separation Theorem in linear system theory; e.g., see [102]. We refer to the investment $I(k)$ above as defining a *drawdown-modulated feedback controller*; see Figure 1.4.

1.4.6 Efficiency Property Related to Drawdown-Modulated Feedback: In Chapter 5, we continue the study of *drawdown*. Motivated by the fact that drawdown is of paramount concern to conservative investors, we dispense with the classical variance as the risk metric and work with drawdown and expected return as the risk-reward pair. Our

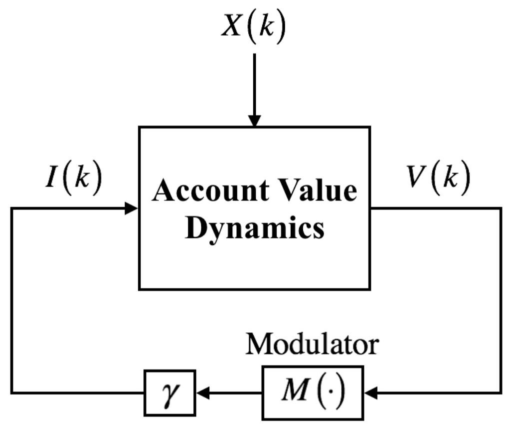


Figure 1.4: Drawdown-Modulated Feedback Controller

analysis begins with a well-known principles which is widely used in finance. An investment opportunity is said to be *efficient* if the following condition holds: There is no other opportunity available with lower risk and at least as high return; e.g., see [30, 103–105]. In this framework, with risk-reward pair being expected drawdown and expected return, we establish the following: Given any linear time-invariant (LTI) feedback $I(k) = KV(k)$ for the investment, there exists a modulated feedback controller

$$I_{\mathcal{M}}(k) \doteq K(k)V(k)$$

with its time-varying gain $K(k) \doteq \gamma M(k)$ leading to a return which is at least as high as that of its LTI competitor and no greater risk. That is, from an efficiency point of view, in a design scenario, there is no loss of generality neglecting the LTI controller in favor of the drawdown-modulated controller. As a bonus, it is also seen that the modulator assures a worst-case level of drawdown protection for all sequences of admissible returns.

1.4.7 Frequency-Based Kelly Betting: In Chapter 6, our analysis of Kelly Betting is extended to include consideration of the *frequency* with which wagers are being made. Suffice it to say, this topic has not been heavily considered in the literature and some initial results along these lines, given in [38] and [106], are either deficient or in need of improvement.

Notably, in [106], we explain what we view to be a serious weakness: The constant gain K is selected without regard for the frequency with which the portfolio rebalancing is done. Subsequently, when this same constant gain K is used to find an optimal rebalancing period, the resulting levels of logarithmic growth are suboptimal. Said another way, in the standard Kelly framework, the frequency at which one updates bets should determine the constant feedback gain K and not vice versa.

In contrast to [38] and [106], we consider the entire range of frequencies from low to high and analyze, in discrete time, the more general case when both the probability distribution of the returns and the time interval between updates are arbitrary. We deal with what we view to be a more appropriate frequency-based Kelly Betting formulation and seek to find an optimum which is frequency dependent. In our new *frequency-based betting* context, we investigate how optimal performance varies with respect to frequency. Specifically, letting $V(k)$ denote the bettor's account value at stage k, our formulation begins with the gambler initially declaring which side of the bet is being taken. The gambler then wagers $KV(0)$ with $K \geq 0$ at stage $k = 0$ and waits $n \geq 1$ steps before updating the bet size.

After each bet, the "house" takes its share and the balance of the money is left to "ride" with resulting profits or losses viewed as "unrealized" until stage n is reached. When n is small, this is viewed as the high-frequency case, and when n is large, one might use the term "bet and hold" in anticipation of the stock market analysis to follow. In this regard, the initial question we address is as follows: Does the high-frequency strategy always lead to the best performance?

When studying in the context of frequency-based framework, as seen in Chapter 6, the compound returns called $\mathcal{X}_n$, between controller updates play an important role. Specifically, for each integer $n \geq 1$ denoting the number of waiting periods between bets, under appropriate boundedness assumptions given in the chapter, we work with the *compound return* given by

$$\mathcal{X}_n \doteq \prod_{k=0}^{n-1}(1 + X(k)) - 1.$$

In our new frequency-based formulation, fixed waiting period $n \geq 1$ and initial account value $V(0) > 0$, the corresponding account value at stage n, given by $V(n) = (1+K\mathcal{X}_n)V(0)$ is considered subject to $K \in \mathcal{K}_n$ where $\mathcal{K}_n$, described in Sections 6.2 and 6.7, is imposed to guarantee survival; i.e., $V(k) < 0$ for all k is disallowed. Now, our objective is to select $K \in \mathcal{K}_n$ maximizing the expected logarithmic growth

$$g_n(K) \doteq \frac{1}{n}\mathbb{E}\left[\log\left(\frac{V_n(n)}{V(0)}\right)\right],$$

and obtain the associated optimal value

$$g_n^* \doteq \max_{K \in \mathcal{K}_n} g_n(K)$$

and $K_n^* \in \mathcal{K}$ satisfying $g_n(K_n^*) = g^*$ is called an *optimal Kelly feedback gain.*

1.4.8 Preview of Contributions in Frequency-Based Betting: As a first step, in Chapter 6, we provide rather complete analysis for the important special case when $X(k)$ is a Bernoulli random variable corresponding to an even-money bet with probability p of winning. This leads us to the analysis of more general probability distributions for $X(k)$. Subsequently, we prove a result, which we call *High-Frequency Maximality Theorem*, which tells us that, with no transaction cost, high-frequency betting is unbeatable in the sense of expected logarithmic growth. This result raises the following question: Under what condition that $g_n^* = g_1^*$ for all $n \geq 1$? To address this, we introduce the technical condition which we call the *sufficient attractiveness inequality* given by

$$\mathbb{E}\left[\frac{1}{1+X(0)}\right] \leq 1$$

and prove that satisfaction of this inequality is sufficient to guarantee that the low-frequency bettor using any $n > 1$ can match the performance of the high-frequency bettor using $n = 1$.

Next, later in Chapter 6, we shift our focus from betting to stock trading and aim at extending our frequency-based theory to portfolio case. Along this direction, we generalize the frequency-based framework to a portfolio case, which involve multi-stock for long-only case. A notion of *relative attractiveness* and *dominant asset*, generalizations of the notion of sufficient attractiveness, are introduced. Then, we prove a result, which we call *Dominant Asset Theorem*, generalizing the Sufficient Attractiveness Theorem from single risky asset scenario to a multi-asset portfolio.

1.4.9 Extended Formulation with Execution Delay: In Chapter 7, motivated by the fact that a trader's interactions with the market are not instantaneous, we extend the frequency-based framework to include delay in execution. In this section, our aim is to preview results on this topic. Indeed, given investment $K \geq 0$, aiming to have stock valued at $KV(0)$, the trader places an order for

$$N(0) \doteq \frac{KV(0)}{S(0)}$$

shares. Due to the one-step execution delay, in our model, these $N(0)$ shares are purchased at stage $k = 1$ at price $S(1)$. Hence the cost of the shares is

$$N(0)S(1) = KV(0)(1 + X(0)).$$

The factor $1 + X(0)$, which involves the random return $X(0)$ implies that the cost of the executed the trade, and therefore the dollar amount invested, is uncertain. It is noted that this investment amount is distinctly different from the previous analysis where no delay was present. In the chapter, for the high-frequency trader with investment at every stage, cash-financing dictates that for $k \geq 1$, with $V_1(0) = V_1(1) = V(0) > 0$, the corresponding investment

$$I_1(k) \doteq N_1(k-1)S(k)$$

is seen to satisfy $0 \leq I_1(k) \leq V_1(k)$. On the other hand, for the buy and holder, without the ability to update until $k = n$, for $k \geq 1$, the initial investment is

$$I_n(1) \doteq N_n(0)S(1)$$

where $N_n(0) \doteq KV_n(0)/S(0)$ satisfying $0 \leq I_n(1) \leq V_n(1)$.

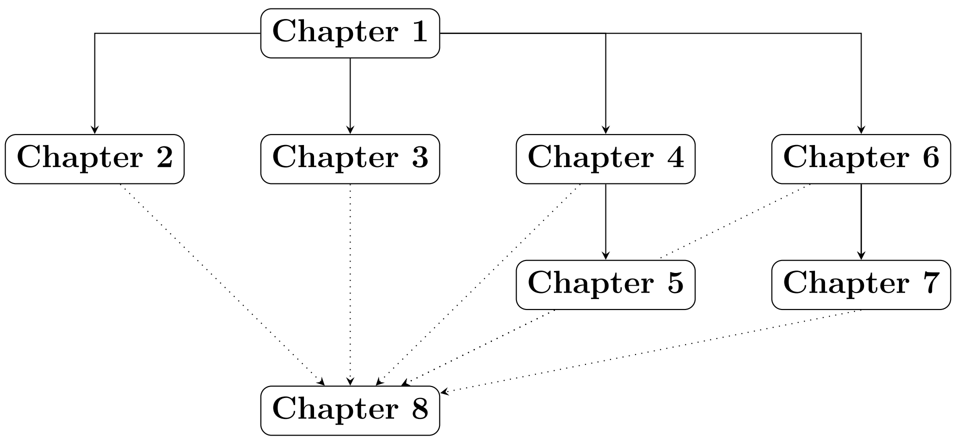
1.4.10 Overview of Results Obtained: Beginning with the formulation above, in Section 7.1, we first provide a result, which we call the Cash-Financing Theorem, which gives a necessary and sufficient condition for a trade to stay long-only and cash-financed in terms of feedback gain K when execution delay is present. Subsequently, similar to the case without delay, for expected logarithmic growth purposes, we use the notation $g_1(K)$ and $g_n(K)$ to denote the performance, as a function of K, achieved by high-frequency trading and buy and hold, respectively. In addition, we denote optima by K_1^* and K_n^* and the associated optimal values by g_1^* and g_n^*. Then, in this framework which includes execution delay, we see that the low-frequency bettor may *strictly* outperform the high-frequency bettor. That is, we obtain $g_n^* > g_1^*$. This result and its technical ramifications are discussed in Sections 7.1–7.4.

As explained in the latter sections of the chapter, the survival issue is the main theme. In contrast to the case without delay, obtaining simple conditions under which the account value remains positive appears to be highly non-trivial. Accordingly, we formulate the so-called *all-time state positivity* problem, which is closely related to the existing positive system theory; e.g., see [107–112]. Then, we develop conditions on the feedback gain $K \geq 0$ such that the condition $V(k) > 0$ is guaranteed for all $k \geq 0$ in an almost sure sense.

To establish this, we introduce two thresholds, K_- and K_+, depending on these bounds, and prove that for $K < K_-$, all-time state positivity is guaranteed. On the other hand, for $K > K_+$, we prove that it is not; i.e., we construct a sequence of asset returns for which the state fails to be positive for all time. Said another way, along this sequence, bankruptcy is certain and the solution of the state equation ceases to be meaningful after some finite time. Finally, the chapter includes a conjecture, with support, which says that for the "gap" interval $K_- \leq K \leq K_+$, all-time state positivity is also guaranteed.

1.5 Concluding Remarks

In summary, the primary objective of this dissertation research is generalizing and extending the theory of Kelly Betting with emphases on its applications to the stock market. The figure below shows the inter-dependence of the technical chapters of this dissertation with solid lines and related future work with dotted lines.



Chapter 1: Overview of Dissertation Research

Chapter 2: Conservatism of Kelly Betting

Chapter 3: Some Limitations of Kelly Betting

Chapter 4: Drawdown-Modulated Control Systems

Chapter 5: An Efficiency Result for Drawdown-Modulated Feedback

Chapter 6: On Frequency-Based Kelly Betting and Stock Trading

Chapter 7: Frequency-Based Framework Involving Delay

Chapter 8: Conclusion and Future Work

Figure 1.5: The Dependence of the Chapters

Chapter 2

Conservatism of Kelly Betting

The starting point in this chapter is the fact that in the existing literature, Kelly Betting often leads to bet sizes which are arguably too aggressive. In contrast, we now describe scenarios when the Kelly-based theory may actually lead to bets which are too conservative rather than too aggressive.[1] Below, we begin our analysis with toy examples which illustrate how overly conservative betting arises. This leads to our "Restricted Betting Theorem," and its corollaries and generalization.

2.1 How Overly Conservative Bets Arise

Our objective in this section is to describe scenarios where the Kelly bettor who uses "pure theory" in lieu of empirical data may reach a conclusion about the optimal bet size which entirely contradicts common sense real-world considerations. That is, we first give a toy example which demonstrates how formal application of the Kelly theory can lead to a bet size which is far smaller than that merited by analysis of risk versus return. Then we provide a more realistic numerical example showing that this pathology which we describe is realizable using real data.

2.1.1 Pathology Explained for a Toy Example: We consider one of the simplest possible Kelly Betting problems described by a Bernoulli random variable X whose probability mass function (PMF) is described as follows: $P(X = 1) = 1 - \varepsilon$ and $P(X = -x_0) = \varepsilon$

[1]The work reported in this chapter has been published in [35].

where $x_0 \gg 1$ and

$$0 < \varepsilon < \frac{1}{1+x_0}.$$

This simple probability mass function is depicted in Figure 2.1.

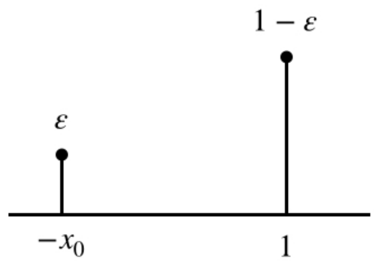


Figure 2.1: Example of Bernoulli Random Variable

For this scenario, the Kelly Betting problem is easily solved via existing literature. That is, we calculate the expected logarithmic growth

$$\begin{aligned} g(K) &= \mathbb{E}\left[\log(1+KX)\right] \\ &= (1-\varepsilon)\log(1+K) + \varepsilon \log(1-Kx_0) \end{aligned}$$

and maximize it and note that this function is readily maximized with respect to K using ordinary calculus. Via a lengthy but straightforward calculation, we obtain the optimal Kelly fraction $K = K^*$ given by

$$K^* = \frac{1-\varepsilon(1+x_0)}{x_0}.$$

It is readily verified to satisfy

$$0 < K^* < \frac{1}{x_0}.$$

This is consistent with the observation that $K \geq 1/x_0$ leads to $\log(1-Kx_0) = -\infty$ irrespective of the size of ε. The key point to note is the following: *No matter how small ε is, the size of K^* is bounded from above by $1/x_0$.* In other words, even when the risk ε of losing becomes negligible, for the Kelly bettor using this theoretical model, the size of the bet will be inappropriately small. For example, with $x_0 = 100$, no matter how small ε is, the betting fraction K can never be more than 1% of the account value. In summary, situations can arise with common sense dictating that one should wager almost all of one's account. However, the Kelly theory may force the betting fraction to be far too small; i.e., overly conservative.

To complete the arguments related to this toy example, we now imagine this same scenario except for the fact that the underlying probability mass function (PMF) above is unknown and has to be estimated by the practitioner. Will an overly conservative bet, based on the estimated PMF, still be overly conservative? The answer depends on the number of data points n which are available. When ε is extremely small, unless n being unacceptably large, it is virtually certain that the bettor will see $x_i = 1$ for $i = 1, 2, \ldots, n$. Hence, the empirically derived PMF for the estimated random variable, call it $\widehat{X}$, is trivially described. Namely, $\widehat{X} = 1$ with probability one, and the resulting expected logarithmic growth maximizer, namely $\widehat{K}^* = 1$ is more consistent with the common-sense maxim: "When conditions are right, bet the farm."

On the other hand, in practice, there is a limitation on n, say $n \leq M$, which can arise for various reasons. For example, if $X(k)$ represents daily returns on a stock, then it would typically be the case that n is strongly limited because the underlying assumption of independent and identically distributed returns becomes questionable when M is too large. For example, if $X(k)$ represents daily returns, many traders use $M = 50$ in the belief that larger M-values requires processing of "old data" which may not reflect current market conditions; i.e., nonstationarity considerations, not included in the model, are in play. For the case of the random variable X under consideration, we ask: What is the probability, call it p_{bad}, that the practitioner will see a "bad" sample; i.e., $x_i = -x_0$ for some $i \leq M$. For this simple problem, we obtain

$$p_{bad} = 1 - (1 - \varepsilon)^M.$$

For example, if $\varepsilon = 0.001$ and $M = 50$, then we obtain $p_{bad} \approx 0.05$ and if $\varepsilon = 0.0001$, we obtain $p_{bad} \approx 0.005$. Note that if such a bad sample is "seen," the behavior of the practitioner becomes similar to that of the theoretical Kelly bettor. That is, once again, an overly-conservative bet results.

2.1.2 A More Realistic Example: To study the issue of conservatism using more realistic data, we consider a family of random variables each of which is governed by the normal distribution with fixed standard deviation $\sigma = 1$. However, we consider their means $0 \leq \mu \leq 4$ as a parameter. Our goal here is to see that the empirical Kelly bettor who relies on this "more realistic data" obtains an optimum $\widehat{K}^*$, which is consistent with common sense; i.e., when μ increases, the associated $\widehat{K}^*$ increases too. Alternatively, as seen in the analysis in next section, a purely theoretical analysis using the normal distribution with its unbounded support leads to $K^* = 0$, which is too conservative. Said another way, a theoretical analysis using a normal distribution $\mathcal{N}(\mu, \sigma)$ leads to no betting regardless of the relative size of the mean μ and the standard deviation σ.

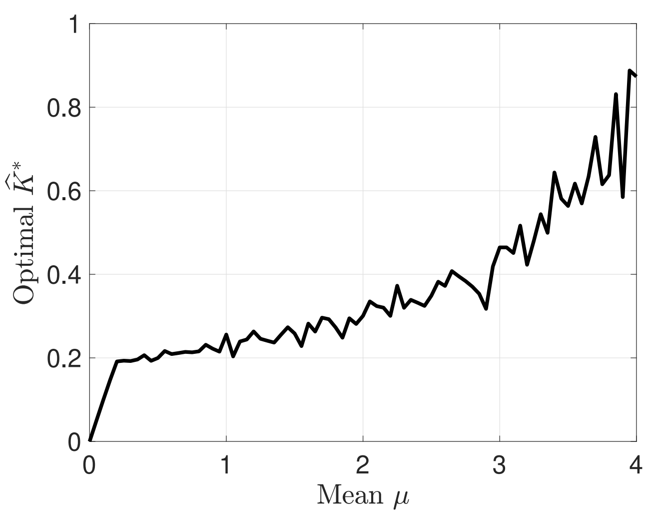


Figure 2.2: Optimal Kelly Fraction $\widehat{K}^*$ Versus μ

Specifically, for each value of μ in this range, we let X_μ denote the random variable of interest and construct an empirically-derived probability mass function by drawing $m = 1,000,000$ sample points. Next, for each μ, we find the optimal Kelly fraction, call it $\widehat{K}^* = \widehat{K}^*(\mu)$;

see Figure 2.2 where this function is plotted. Looking at the plot, we now argue that this result is consistent with common sense considerations. Indeed, when μ is at the low end of the range, it is no surprise to see that $\widehat{K}^*(\mu)$ is small because the probability of $X_\mu < 0$ is significant. For example, when $\mu = 1$, the optimum is to wager about 20% of one's wealth on each bet. Similarly, when μ is at the high end of the range, we see that $\widehat{K}^*(\mu)$ is large because the probability of $X_\mu < 0$ becomes small. For example, when $\mu = 4$, the optimum is to wager about 90% of one's wealth on each bet because the chance of losing is vanishingly small. In the next section, we see that this analysis using real data is entirely at odds with the use of a purely theoretical model.

2.2 Restricted-Betting: The Scalar Case

In this section, we provide an analysis which is apropos to the motivating examples in the preceding section. In rough terms, for a scalar random variable X, we see that the minimum and maximum values of points x in the support lead to fundamental restrictions on the size of the bet allowed by the Kelly theory — the larger these values, the smaller the Kelly fraction is forced to be. Moreover, this restriction holds true whether the probability of these maximal deviations is significant or not.

Since the key ideas driving the analysis to follow, similar to what was discussed in the toy example above, are most simply understood when X is a scalar random variable, we first consider this case. To begin, suppose $x_0 < 0$ is a point in the support set of X. Then, to avoid $g(K) = -\infty$, Kelly theory forces the betting fraction to satisfy $K \leq -1/x_0$. This holds true even when the probability that X gets close to x_0 is vanishingly small. Similarly, for a point $x_0 > 0$ in the support, similar reasoning forces $K \geq -1/x_0$. As a result of this aspect of the theory, as in the toy example, many bets which are "excellent" from a common sense point of view lead to unduly small bets. To summarize, in the Kelly theory, large values of X, whether rare or not, lead to dramatic restrictions in the bet size.

In the lemma below, we formalize the ideas above. An extreme case of the result occurs when the support of X is the entire real line; e.g., X is normally distributed. For such cases, as seen below, $K = 0$ is forced. That is, no betting is allowed. This result holds true regardless of the relative sizes of the mean μ and standard deviation σ. We note that this outcome of Kelly theory is clearly at odds with practical considerations. Even when the ratio μ/σ is very large, synonymous with an excellent bet, the theory nevertheless forces $K = 0$. The lemma below is a special case of the Restricted Betting Theorem given in the next section. Accordingly, its proof is deferred until then.

Lemma 2.2.1 (Scalar Restricted Betting): *Let X be a random variable with $\mathbb{E}[|X|] < \infty$, probability density function $f_X(x)$ and support $\mathcal{X}$ with extremes*

$$X_{\min} \doteq \inf\{x : x \in \mathcal{X}\} \quad \textit{and} \quad X_{\max} \doteq \sup\{x : x \in \mathcal{X}\}$$

satisfying $X_{\min} < 0$ and $X_{\max} > 0$. Then any optimizing Kelly fraction K maximizing $g(K)$ satisfies the condition

$$K \in \mathcal{K} \doteq \left[-\frac{1}{X_{\max}}, -\frac{1}{X_{\min}}\right].$$

2.2.2 Remarks: For the extreme case when the support of the distribution is unbounded both from above and below, the lemma indicates that the optimal fraction is $K^* = 0$. To see this, we observe that $K^* \leq 0$ when $X_{\min} = -\infty$ and $K^* \geq 0$ when $X_{\max} = +\infty$. It follows that $K = 0$ is forced. In other words, the best bet is no bet at all. Surprisingly, as indicated via the counterexample below, the satisfaction of the condition on K above is not sufficient for finiteness of $g(K)$.

2.2.3 Counterexample Demonstrating Lack of Sufficiency: We provide an example of a random variable X and a constant $K > 0$ satisfying the interval confinement condition above but having the property that $g(K) = -\infty$. Indeed, let $0 < K < 1$ be arbitrary and held fixed in the calculations to follow. We now consider a random variable X constructed as follows. Let

$$\theta \doteq \frac{1}{2} + \sum_{k=1}^{\infty} \frac{1}{k^2}$$
$$= \frac{1}{2} + \frac{\pi^2}{6},$$

take $X = x_0 = 1$ with probability $p_0 = 1/(2\theta)$ and for $k \geq 1$, take

$$X = x_k \doteq \frac{1}{K}(e^{-k} - 1)$$

with probability $p_k \doteq 1/(k^2\theta)$. Note that the definition of θ above assures that the p_k define a probability mass function; i.e., $p_k \geq 0$ and $\sum_{k=0}^{\infty} p_k = 1$. Now, for this random variable, we have $X_{\min} = -1/K$, and $X_{\max} = 1$. Furthermore, since $0 < K < 1$, the interval confinement condition above is satisfied. To complete the analysis, it remains to show that $g(K) = -\infty$. Indeed, we calculate

$$\begin{aligned}
g\left(K\right) &= \mathbb{E}[\log\left(1 + KX\right)] \\
&= \sum_{k=0}^{\infty} \log\left(1 + Kx_k\right) p_k \\
&= \log\left(1 + Kx_0\right) p_0 + \sum_{k=1}^{\infty} \log\left(1 + Kx_k\right) p_k \\
&= \frac{1}{2\theta} \log\left(1 + K\right) + \frac{1}{\theta} \sum_{k=1}^{\infty} \frac{1}{k^2} \log\left(1 + Kx_k\right) \\
&= \frac{1}{2\theta} \log\left(1 + K\right) + \frac{1}{\theta} \sum_{k=1}^{\infty} \frac{1}{k^2} \log\left(e^{-k}\right) \\
&= \frac{1}{2\theta} \log\left(1 + K\right) - \frac{1}{\theta} \sum_{k=1}^{\infty} \frac{1}{k} \\
&= -\infty.
\end{aligned}$$

2.3 The Restricted-Betting Theorem

This section provides a generalization of the scalar random variable results above to the case of an m-dimensional random vector X whose support set $\mathcal{X}$ can be rather arbitrary. To obtain the theorem below, we make use of the classical support function which is heavily used in convex analysis; e.g., see [113]. That is, given a set $\mathcal{X} \subseteq \mathbb{R}^m$, the *support function* on $\mathcal{X}$ is the mapping $h : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ defined as follows: For $y \in \mathbb{R}^m$,

$$h_{\mathcal{X}}(y) \doteq \sup_{x \in \mathcal{X}} y^T x.$$

After establishing the theorem below, we consider a number of special cases to show that there are large classes of Kelly Betting problems for which checking for satisfaction of the requirements of the theorem is highly tractable.

Theorem 2.3.1 (Restricted Betting): *Given an m-dimensional random vector X with probability density function f_X, support $\mathcal{X}$, and $\mathbb{E}[\|X\|] < \infty$, any optimizing Kelly fraction vector K satisfies the condition*

$$h_{\mathcal{X}}(-K) \leq 1.$$

Furthermore, whether the support $\mathcal{X}$ is convex or not, the set

$$\mathcal{K} \doteq \{K \in \mathbb{R}^m : h_{\mathcal{X}}(-K) \leq 1\}$$

is nonempty, convex and closed.

Proof. In the arguments to follow, we work with the extended logarithmic function which takes value $\log x = -\infty$ for $x \leq 0$. Proceeding by contradiction, suppose K is optimal but fails to satisfy the support function condition above. Then

$$\sup_{x \in \mathcal{X}} [-K]^T x > 1.$$

Equivalently, there exists some $x^K \in \mathcal{X}$ such that $-K^T x^K > 1$. Hence

$$1 + K^T x^K < 0.$$

Now noting that $1 + K^T x$ is continuous in x and that x^K is in the support, there exists a suitably small neighborhood of x^K, call it $\mathcal{N}(x^K)$, such that

$$1 + K^T x < 0$$

for $x \in \mathcal{N}(x^K)$ and

$$P(X \in \mathcal{N}(x^K)) > 0.$$

We now claim that the existence of such a neighborhood implies that $g(K) = -\infty$. To prove this, we first observe that

$$\begin{aligned} g(K) &= \mathbb{E}[\log(1 + K^T X)] \\ &= \int \log(1 + K^T x) f_X(x) dx \\ &= \int_{1+K^T x \leq 0} \log(1 + K^T x) f_X(x)\, dx + \int_{1+K^T x > 0} \log(1 + K^T x) f_X(x)\, dx. \end{aligned}$$

Using the property of logarithmic function that

$$\log(1 + K^T x) \leq |K^T x|$$

for all x satisfying $1 + K^T x > 0$, we obtain an upper bound for $g(K)$. That is,

$$\begin{aligned} g(K) &\leq \int_{1+K^T x \leq 0} \log(1 + K^T x) f_X(x)\, dx + \int_{1+K^T x > 0} \left|K^T x\right| f_X(x)\, dx \\ &\leq \int_{1+K^T x \leq 0} \log(1 + K^T x) f_X(x)\, dx + \int \|K\|\, \|x\| f_X(x)\, dx \\ &\leq \int_{1+K^T x \leq 0} \log(1 + K^T x) f_X(x)\, dx + \|K\|\, \mathbb{E}[\|X\|]. \end{aligned}$$

Since $\mathbb{E}[\|X\|] < \infty$, it suffices to show that the last integral above has value $-\infty$. Beginning with the fact that

$$P(X \in \mathcal{N}(x^K)) > 0$$

and noting that $\mathcal{N}(x^K) \subseteq \{x : 1 + K^T x \leq 0\}$, the density function f_X must assign positive probability to the set $\{x : 1 + K^T x \leq 0\}$. Furthermore, since $\log(1 + K^T x) = -\infty$ for x

satisfying $1 + K^T x \leq 0$, it follows that

$$\int_{1+K^T x \leq 0} \log(1 + K^T x) f_X(x)\, dx = -\infty$$

and we conclude that $g(K) = -\infty$ as required.

To complete the proof, we first note that $0 \in \mathcal{K}$, which shows that $\mathcal{K}$ is nonempty. To establish closedness and convexity of $\mathcal{K}$, we use a rather standard convex analysis argument: For each fixed $x \in \mathcal{X}$, we define the linear function

$$L_x(K) \doteq -K^T x$$

and associated set

$$\mathcal{K}_x \doteq \{K \in \mathbb{R}^m : L_x(K) \leq 1\}.$$

Note, that $\mathcal{K}_x$, being a halfspace, is a closed convex set. Now, using the definition of the support function, it follows that

$$\mathcal{K} = \bigcap_{x \in \mathcal{X}} \mathcal{K}_x.$$

Hence, since $\mathcal{K}$ is the intersection of an indexed collection of closed convex sets, it is also closed and convex. $\square$

2.3.2 Existence of Solution: Note that the theorem above does not guarantee the existence of an optimal element K^*; i.e., it simply provides a necessary condition that any optimum must satisfy. Since $g(K)$ is concave and the set $\mathcal{K}$ is non-empty convex and closed, it follows that the set of all maximizers, call it $\mathcal{K}^*$, is convex; see [114]. In addition, when $\mathcal{K}$ is bounded, as considered in the remaining chapters, then using the fact that a continuous function on a non-empty compact set attains its supremum; for example, see [115], it follows that the $\mathcal{K}^*$ is nonempty.

2.3.3 Scalar Result as a Special Case: To see that the Scalar Restricted Betting Lemma 2.2.1 in Section 2.2 is a special case of the Restricted Betting Theorem 2.3.1 above, we assume $X_{\min} < 0$ and $X_{\max} > 0$ as in the lemma. Now, for $K > 0$, the support function in the theorem becomes $h_{\mathcal{X}}(-K) = -KX_{\min}$ and for $K < 0$, it becomes $h_{\mathcal{X}}(-K) = -KX_{\max}$. Hence the requirement of the theorem $h_{\mathcal{X}}(-K) \leq 1$ leads to the interval condition of Lemma 2.2.1.

2.3.4 Example (Hypercube Support Set): One m-dimensional generalization of the scalar situation above is obtained when the convex hull of the support of X, $\text{conv}\mathcal{X}$, is a hypercube. Suppose this hypercube has center x^0 and components x_i satisfying $|x_i - x_i^0| \leq \delta_i$ where $\delta_i > 0$ for $i = 1, 2, \ldots, m$. Then using a basic fact about support functions, for example, see [116, p. 269], that

$$h_{\mathcal{X}}(y) = h_{\text{conv}\mathcal{X}}(y)$$

for all $y \in \mathbb{R}^n$, a straightforward calculation leads to

$$h_{\mathcal{X}}(-K) = \sum_{i=1}^{m} |K_i|\delta_i - \sum_{i=1}^{m} K_i x_i^0.$$

Hence, application of the Restricted Betting Theorem 2.3.1 leads to the requirement that any optimizing Kelly fraction vector K satisfy the condition

$$\sum_{i=1}^{m} |K_i|\delta_i - \sum_{i=1}^{m} K_i x_i^0 \leq 1.$$

2.3.5 Example (Hypersphere Support Set): As a final example, suppose the convex hull of the support set $\mathcal{X}$ is a hypersphere in $\mathbb{R}^m$ with description $\|x - x^0\| \leq r$ with Euclidean norm used above, center x^0, and radius $r > 0$. Then using an argument which is similar to that used for the hypercube example above, it is easily shown that any optimizer K must satisfy

$$r\|K\| - K^T x^0 \leq 1.$$

We note that the constraint sets

$$\mathcal{K}_r \doteq \{K \in \mathbb{R}^m : r\|K\| - K^T x^0 \leq 1\}$$

are nested. That is, if radii $r_1 \leq r_2$, then the set $\mathcal{K}_{r_2} \subseteq \mathcal{K}_{r_1}$. In Figure 2.3, these sets are depicted for $x^0 = (1/2, 1/2)$ and various radii $r_1 = 1$, $r_2 = 1.25$, $r_3 = 2$, $r_4 = 3$ and $r_5 = 5$.

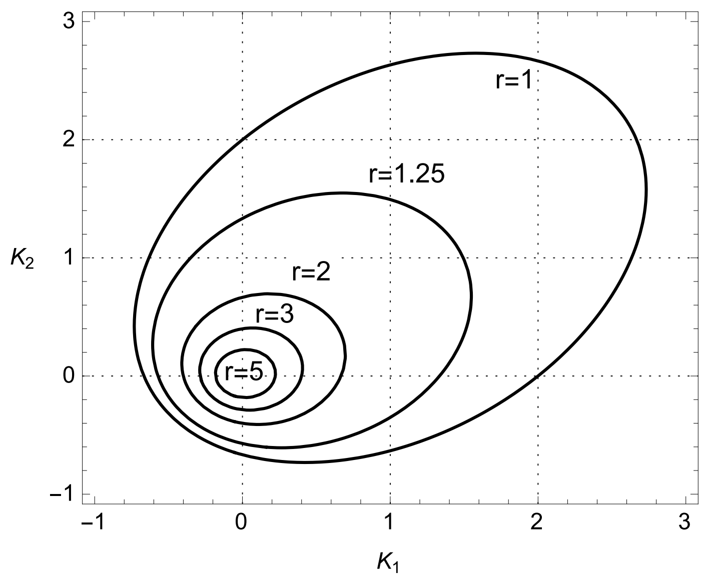


Figure 2.3: Constraint Sets $\mathcal{K}_r$ for Optimal Fraction K

2.4 Concluding Remarks and Future Work

In this chapter, we began by comparing the size of Kelly bets which are derived using a purely theoretical probability distribution versus those which are obtained from its empirically-obtained counterpart. In this regard, the Restricted Betting Theorem 2.3.1 tells us that when the expected logarithmic growth function $g(K)$ is maximized, an "excessively" conservative K-value may result.

The results in this chapter open the door to a new line of research which might be appropriately called "data-driven Kelly Betting." In such an empirical framework, new problems involving the sample size m will be of fundamental importance. Given that many betting processes involve non-stationary stochastic processes, there is typically a bound $m \leq M$ which must be respected when deriving the empirical distribution. That is, when the analysis involves sequential betting based on i.i.d. random variables, the use of "untrustworthy old data" from far in the past may be inappropriate to use; see also Section 6.14 for an example related to this direction.

In next chapter, we fill a void in existing literature by pointing out further limitations associated with application to some of the Kelly Betting results in the literature. We first show that the Taylor-based approximation technique used in some papers may lead to a so-called *inefficient* solution which is suboptimal for the original problem. Subsequently, we discuss a negative issue associated with the vast preponderance of existing papers on Kelly Betting: namely, the drawdown. Aimed at finding a remedy for this issue, we introduce a "drawdown surrogate" which enables us to formulate a drawdown-constrained Kelly-based optimization problem as a concave program, which can be solved in an efficient way.

Chapter 3

Some Limitations of Kelly Betting

The primary goal of this chapter is to fill some voids in the existing literature involving limitations associated with Kelly Betting theory and the way it has been applied to date.[1] To this end, we first analyze Taylor-based approximation methods used in some papers to optimize the expected logarithmic growth. As shown in next section, when such approximations are used, the associated "optimal solution" may be either infeasible or lead to performance which is significantly lower than that attained by the true optimum. In addition, we show that approximate solutions may have a certain "inefficiency" property which is also undesirable.

Following the analysis above, we address the important issue of wealth drawdown and provide some specific examples for which drawdown can be computed. After providing some preliminary results aimed at mitigating the drawdown problem, we state two related conjectures. If true, they facilitate the study of drawdown-constrained Kelly-based optimization problems because convex programming tools become possible to use. More specifically, these two conjectures tell us that two drawdown constraint sets of practical interest are convex polytopes. This may prove to be useful when incorporated into some asset allocation optimization problems in finance. Given the possibility that the two conjectures may ultimately prove to be false, in Section 3.4, we introduce a surrogate drawdown measure which enables us to formulate an "approximate" drawdown-constrained Kelly-based optimization problem which also lends itself to solution via convex programming. However, it is important to note that the surrogate is only valuable when the quality of the approximation associated with its use is high. This is seen to be the case for the example which we consider.

---

[1] The work reported in this chapter has been published in [34].

3.1 Negatives Associated with Approximation

To obtain the optimal logarithmic growth rate g^*, one approach in the literature involves approximation. That is, either a multivariate Taylor expansion to the logarithmic growth function is used or $X(k)$ is treated as an approximation of a Geometric Brownian Motion and low order expansion terms are used; e.g., see [23, 25, 31–33]. The main objective in this section is to point out some pitfalls associated with the use of such approximations. While it is arguable that approximation-based solutions provide a degree of insight into the risk-return tradeoffs, it is demonstrated via numerical examples that approximation methods may lead to poor results. That is, when the range of variation of $X(k)$ can be large, the true optimum $K = K^*$ and associated expected logarithmic growth $g^* = g(K^*)$ can differ considerably from their corresponding quantities obtained by approximation.

3.1.1 Example Involving Approximation: We consider a somewhat attractive coin-flipping style gamble as follows: For $k = 0, 1, ..., N-1$, we assume i.i.d. returns $X(k) = 0.15$ with probability $p = 0.95$ and $X(k) = -0.95$ with probability $p = 0.05$. It is noted that this bet is "attractive" in an asymptotic sense; i.e., since $\mathbb{E}[X(0)] = 0.095 > 0$, repeated i.i.d. trials, will almost certainly lead to success if N is large. Now, recalling that the expected logarithmic growth, defined in Section 1.2, is given by

$$g(K) = \mathbb{E}[\log(1 + KX(0))],$$

for $K \in [0, 1]$, according to [32] and [33], using the Taylor approximation

$$\mathbb{E}[\log(1 + KX(0))] \approx K\mathbb{E}[X(0)] - \frac{1}{2}K^2\mathbb{E}[X^2(0)],$$

it is straightforward to see that the associated optimal investment fraction K, call it $\mathcal{K}_{\text{Taylor}}$, is given by

$$\mathcal{K}_{\text{Taylor}} = \frac{\mathbb{E}[X(0)]}{\mathbb{E}[X^2(0)]} \approx 1.4286.$$

Note that this solution is infeasible because $K \in [0, 1]$ is required. Hence, to guarantee feasibility, one possibility, pursued in [32], is to introduce a saturation function applied to approximate solution above. Specifically, the optimal approximate solution with saturation and the associated expected logarithmic growth are given by

$$SAT\left[\mathcal{K}_{\text{Taylor}}\right] = 1;$$

$$g(SAT\left[\mathcal{K}_{\text{Taylor}}\right]) \approx -0.017$$

where $SAT[x]$ is given by

$$SAT[x] = \begin{cases} 0 & x < 0; \\ x & 0 \leq x \leq 1; \\ 1 & x > 1. \end{cases}$$

Since $g(K_{\text{sat}}) < 0$, a loss is to be expected. However, one can always use $K = 0$ to achieve better performance; i.e., $g(0) = 0$.

An alternative approach, for example, see [25] and [31], involves treating $X(k)$ as approximate Geometric Brownian Motion with drift $\mu = \mathbb{E}[X(0)]$ and variance $\sigma^2 = \text{var}(X(0))$. Then, for some *market* with typical returns which are small, some authors; e.g., see [31], use the approximation $\text{var}[X(0)] \approx \mathbb{E}[X^2(0)]$, to obtain solution, call it $\mathcal{K}_{\text{GBM}}$, given by

$$\mathcal{K}_{\text{GBM}} = \frac{\mathbb{E}[X(0)]}{\text{var}[X(0)]} = \frac{\mu}{\sigma^2} \approx 1.6529.$$

Note that the same solution above is also obtained in [25]. Similarly, since this solution is infeasible, one possible way to resolve this issue, discussed in [25], is to apply saturation to $\mathcal{K}_{\text{GBM}}$. This approach leads to

$$K_{\text{GBM}} \doteq SAT\left[\mathcal{K}_{\text{GBM}}\right] = 1;$$

$$g(K_{\text{GBM}}) \approx -0.017.$$

As in our Taylor series approximation analysis, the negative expected logarithmic growth again implies an expected loss, and $K = 0$ is a better choice.

3.1.2 Remarks: In contrast to the two approximate solutions above, the true optimum, obtained by maximizing the expected logarithmic growth

$$g(K) = 0.95 \log(1 + 0.15K) + 0.05 \log(1 - 0.95K),$$

is found by differentiation. This leads to an optimal feasible solution in $[0, 1]$ and associated expected logarithmic growth given by $K^* \approx 0.6667$ and $g(K^*) \approx 0.0404$, respectively.

A summary of all three solutions and the $g(K)$ plot are shown in Figure 3.1. Ironically, the approximation-based results minimize expected logarithmic growth rather than achieving the desired maximization. Suffice it to say, the combination of approximation and saturation due to constraint violation can lead to significant error.

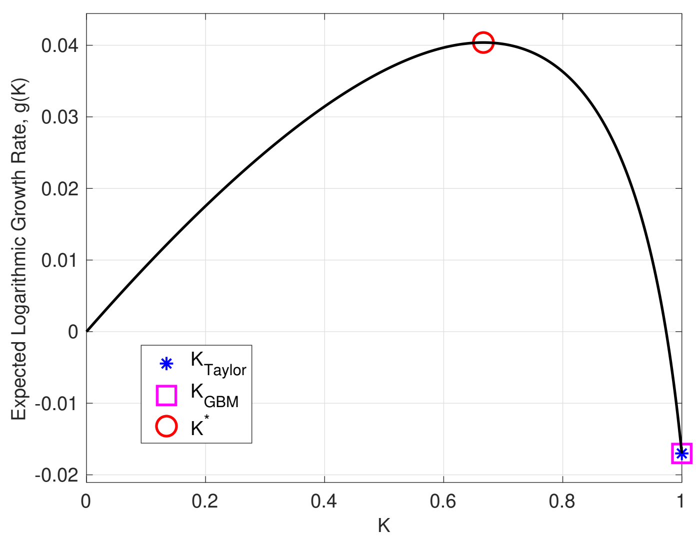


Figure 3.1: Expected Logarithmic Growth Rate

3.1.3 More Realistic Example with Real Stock Data: In this example, we further consider the problems associated with approximation. Here, however, instead of using a theoretical model, we work with real data for two stocks: Tesla Motors and IBM. This data covers the ninety-day period January 2, 2013 until May 13, 2013; see Figure 3.2 for daily closing prices.

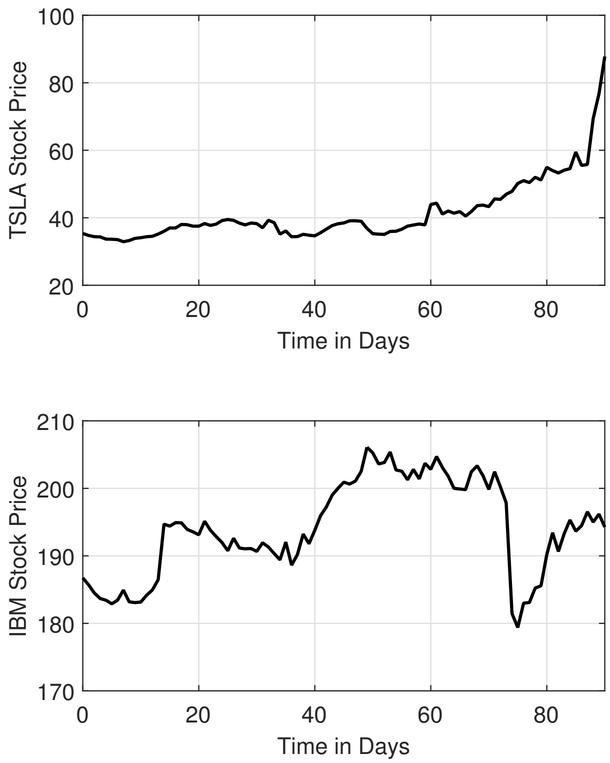


Figure 3.2: Two Stock Prices: TSLA and IBM

Using $s_1(k)$ and $s_2(k)$ for the k-th daily *realized prices* of Tesla and IBM respectively, for $k = 0, 1, \ldots, 89$, we first calculate the *realized return* vector

$$x(k) \doteq [x_1(k)\ x_2(k)]$$

where

$$x_i(k) \doteq \frac{s_i(k+1) - s_i(k)}{s_i(k)}$$

for $i = 1, 2$. Subsequently, we obtain an estimate of the joint probability mass function (PMF) as the sum of Dirac Delta functions

$$\widehat{f}_X = \frac{1}{90}\sum_{j=0}^{89}\delta\left(x - x(j)\right)$$

and use it as input to the optimization to be carried out. Next, we solve an in-sample maximization of $g(K)$ subject to the constraints $K_1 \geq 0, K_2 \geq 0$ and $K_1 + K_2 \leq 1$. This leads to solution $K_1^* = 1$ and $K_2^* = 0$. That is, the optimum in-sample log-growth solution requires all funds invested in Tesla and no funds in IBM.

Now, suppose instead that one computes the Taylor-based solutions for the random vector, call it X, with joint probability mass function $\widehat{f}_X$ as described above. Then, according to [31–33], a straightforward calculation leads to

$$\mathcal{K}_{\text{Taylor}} = \Sigma^{-1}(X)\mathbb{E}[X] \approx [5.321 \;\; 2.725]^T,$$

or for suitably small return, one obtains

$$\mathcal{K}_{\text{GBM}} = \overline{\Sigma}^{-1}(X)\mathbb{E}[X] \approx [5.599 \;\; 2.681]^T$$

where

$$\mathbb{E}[X] \doteq \frac{1}{90}\left[\begin{array}{cc} \sum\limits_{j=0}^{89} x_1(j) & \sum\limits_{j=0}^{89} x_2(j) \end{array}\right]^T \approx \left[\begin{array}{cc} 0.0111 & 0.0005 \end{array}\right]^T$$

and

$$\Sigma(X) \doteq \frac{1}{90}\left[\begin{array}{cc} \sum\limits_{j=0}^{89} x_1^2(j) & \sum\limits_{j=0}^{89} x_1(j)\,x_2(j) \\ \sum\limits_{j=0}^{89} x_2(j)\,x_1(j) & \sum\limits_{j=0}^{89} x_2^2(j) \end{array}\right] \approx \left[\begin{array}{cc} 0.002075 & 0.000011 \\ 0.000011 & 0.000172 \end{array}\right]$$

is the second moment matrix for X, and

$$\overline{\Sigma}(X) \doteq \left[\begin{array}{cc} \operatorname{var}(x_1) & \operatorname{cov}(x_1, x_2) \\ \operatorname{cov}(x_2, x_1) & \operatorname{var}(x_2) \end{array}\right] \approx \left[\begin{array}{cc} 0.001974 & 0.000005 \\ 0.000005 & 0.000174 \end{array}\right]$$

is the covariance matrix for X with component

$$\overline{\Sigma}_{ij} \doteq \frac{1}{89}\sum_{k=0}^{89}(x_i(k) - m_i)(x_j(k) - m_j)$$

and

$$m_i \doteq \frac{1}{90}\sum_{k=0}^{89} x_i(k)$$

for $i,j \in \{1,2\}$. Note that the approximate solutions $\mathcal{K}_{\text{Taylor}}$ and $\mathcal{K}_{\text{GBM}}$ are infeasible since the constraints are violated.

Given the violation of the constraint $K_1 + K_2 \leq 1$ for both the Taylor and GBM approximate solutions, one standard remedy, a generalization of the saturation concept, involves projecting these solutions onto the constraint satisfaction set. That is, we take

$$Proj(\mathcal{K}_{\text{Taylor}}) \approx [0.661 \;\; 0.339]^T;$$

$$Proj(\mathcal{K}_{\text{GBM}}) \approx [0.661 \;\; 0.339]^T$$

where the projection function $Proj(\cdot)$ above is given by

$$Proj(K_1, K_2) \doteq \left[\frac{K_1}{K_1+K_2} \quad \frac{K_2}{K_1+K_2}\right]^T$$

and is defined for all nonnegative $(K_1, K_2) \neq (0,0)$. Although the projection procedure leads to a feasible solution, it is important to note that these solutions differ considerably from the previously found true optimum $K^* = (K_1^*, K_2^*) = (1,0)$. Moreover, the associated optimal expected logarithmic growth is given by $g^* = g(K^*) \approx 0.0101$ but the expected logarithmic growth obtained by using $Proj(\mathcal{K}_{\text{Taylor}})$ and $Proj(\mathcal{K}_{\text{GBM}})$ are

$$g\left(Proj(\mathcal{K}_{\text{Taylor}})\right) \approx g(Proj(\mathcal{K}_{\text{GBM}})) \approx 0.0071$$

which is strictly less than the optimal value g^*. The true optimum K^* along with approximate solutions $\mathcal{K}_{\text{Taylor}}$ and $\mathcal{K}_{\text{GBM}}$, and their projections are seen in Figure 3.3. The associated expected logarithmic growths are shown in Figure 3.4.

3.1.4 Inefficiency of Approximate Solution: We point out another danger associated with the use of approximate solutions which may arise. This is based on the following principle which is widely used in finance: If two investments have the same *risk*, but one

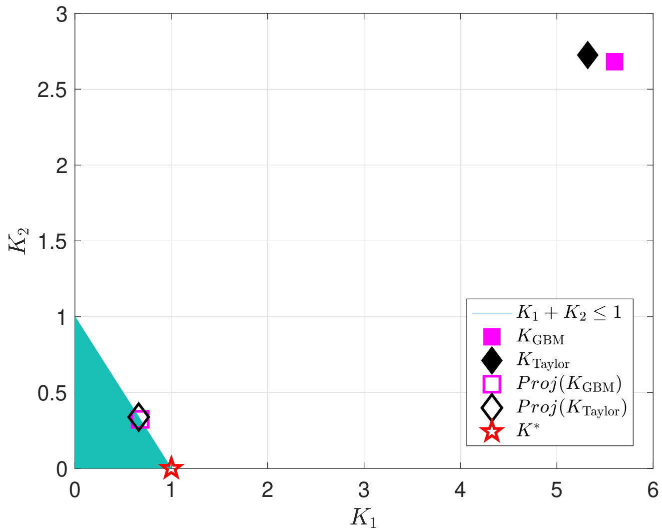


Figure 3.3: Approximate Solutions and its Projections

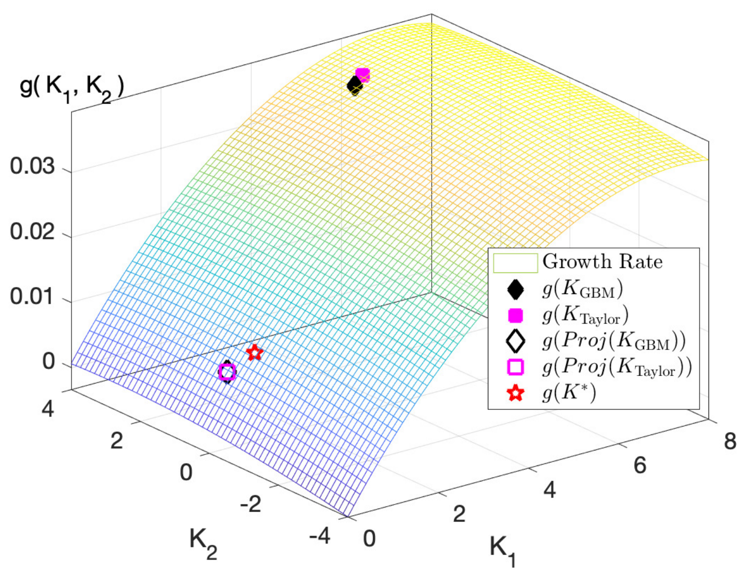


Figure 3.4: Constraint Violation Example for Two-Stock Case

has smaller return, it will be discarded and deemed to be "inefficient." We now argue that the approximate solutions $\mathcal{K}_{\text{Taylor}}$ or $\mathcal{K}_{\text{GBM}}$, even when feasible, might be inefficient. To this end, we now provide an example illustrating inefficiency using $\mathcal{K}_{\text{Taylor}}$ and note that the same example can be used for $\mathcal{K}_{\text{GBM}}$. Indeed, we consider a random variable X described as follows: Given $\gamma > 0$ as the return, we take $X = \gamma$ with probability $p > 0$ and $X = -1$ with probability $1 - p$. Using the Taylor approximation, as a function of reward level γ, we obtain fraction $\mathcal{K}_{\text{Taylor}} = \mathcal{K}_{\text{Taylor}}(\gamma)$ given by

$$\begin{aligned}\mathcal{K}_{\text{Taylor}}(\gamma) &= \frac{\mathbb{E}[X]}{\mathbb{E}[X^2]} \\ &= \frac{p\gamma + p - 1}{p\gamma^2 - p + 1}.\end{aligned}$$

In order for an investment based on $\mathcal{K}_{\text{Taylor}}(\gamma)$ to be efficient from an economic risk-taking point of view, it should have the following property: When $\gamma_2 \geq \gamma_1 \geq 0$, we require

$$K(\gamma_2) \geq K(\gamma_1).$$

That is, if the bet associated with γ_2 offers more return with the same probabilities of success and failure as those for γ_1, a rational gambler should invest at least as much in the γ_2 bet as in the γ_1 bet. We claim that the K-value, call it $K(\gamma)$, obtained from the Taylor-based approximation scheme, may be feasible and fail to satisfy this condition. To establish this claim, it suffices to show that for some $\gamma > 0$, $d\mathcal{K}_{\text{Taylor}}/d\gamma$ can be negative. Furthermore, with $\mathcal{K}_{\text{Taylor}}(\gamma)$ in $(0, 1)$, we calculate

$$\frac{d\mathcal{K}_{\text{Taylor}}}{d\gamma} = -p\,\frac{p\gamma^2 + 2(p-1)\gamma + p - 1}{(p\,\gamma^2 - p + 1)^2}$$

and note that the denominator cannot vanish. Hence, we see $d\mathcal{K}_{\text{Taylor}}/d\gamma < 0$ for

$$\gamma > \gamma^*(p) \doteq \frac{1 - p + \sqrt{1-p}}{p}$$

which corresponds to the zero-crossing of the numerator. In Figure 3.5, the plot of $\mathcal{K}_{\text{Taylor}}(\gamma)$ is given for $p = 0.8$. It is readily apparent that the claimed inefficiency occurs for the parameter range $\gamma > \gamma^*(0.8) \approx 0.809$.

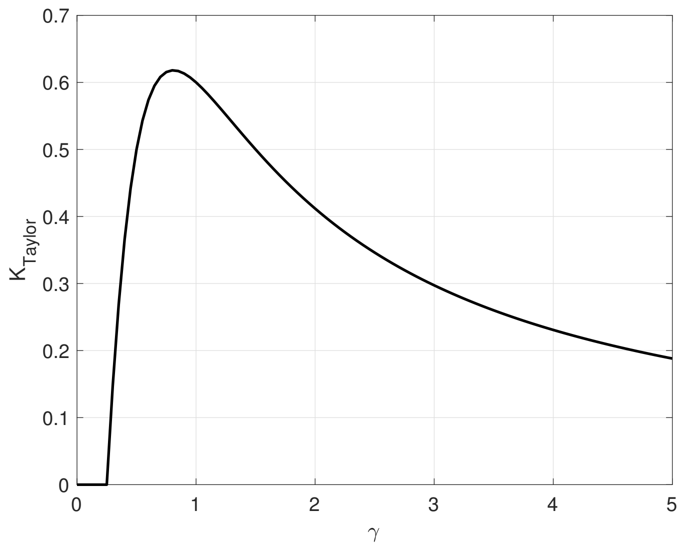


Figure 3.5: $\kappa_{\text{Taylor}}(\gamma)$ Plot for $p = 0.8$

3.2 Negatives Associated with Drawdown

The issue of drawdown, that is, drops in wealth from peaks to subsequent lows, is of great concern from a risk management perspective. In a stock trading context, large drawdown is unsatisfactory to many traders and fund managers; it can be psychologically painful and may lead to a strategy being abandoned. Suffice it to say, the issue of drawdown control has received considerable attention in the finance literature. In this regard, in trading a single stock, one standard approach is to include drawdown limits as a constraint in an optimization problem which involves maximization of some performance index such as the return; e.g., see [98–100, 117, 118] which focus on a single-stock scenario. There are also some papers dealing with modifications and extensions of these results for the single-stock case to address a multi-asset portfolio under various stochastic modeling assumptions; e.g., see [119–122]. To provide further context for this chapter, we mention a sampling of some other papers in the existing literature using *risk* measures other than drawdown. Examples of such measures include Value at Risk (VaR), Conditional Value at Risk, Expected Shortfall and the celebrated mean-variance criterion; e.g., see [30, 41, 104, 123].

In this sections to follow, we demonstrate how Kelly Betting often results in excessively large drawdown. This includes a discussion of some of our preliminary work to date aimed at mitigating the drawdown issue within the Kelly Betting framework. To this end, for $k = 0, 1, \ldots, N$, with $V(k)$ being the account value at stage k, we take being the *overall percentage drawdown*

$$d^* \doteq \max_{0 \le k \le N} d(k)$$

where

$$V_{\max}(k) \doteq \max_{0 \le i \le k} V(i)$$

and

$$d(k) \doteq \frac{V_{\max}(k) - V(k)}{V_{\max}(k)}$$

is the *percentage drawdown to date.* Note that maximum percentage drawdown d^* defined above is equivalent to

$$d^* = \max_{0 \le l \le k \le N} \frac{V(l) - V(k)}{V(l)}.$$

In the sequel, for simplicity, we often drop the word "percentage" in reference to this quantity and write d_K^* instead of d^* to emphasize that the drawdown is dependent on the fraction K. In next subsection, we address the following question: In what sense can we control drawdown within the Kelly Betting framework? Noting that one can trivially reduce the drawdown to any desired level by making K suitably small. The key issue is the extent to which expected logarithmic growth can be guaranteed with considerations of d^* in play.

3.2.1 Control of Drawdown: To properly control drawdown, one possibility is to add a probabilistic constraint to the Kelly-based optimization problem. To be more specific, given $0 < \varepsilon < 1$ and $0 < \delta < 1$, consider the so-called *chance constraint*

$$P(d_K^* \le \varepsilon) \ge 1 - \delta.$$

Thus, the corresponding constrained Kelly Betting problem is described as follows: We seek to find an admissible K maximizing the expected logarithmic growth subject to this probabilistic constraint. This chance constraint has a similar flavor to some of the well-known constraints involving risk measures in finance such as the Value at Risk (VaR) and Conditional Value at Risk; e.g., see [30, 104, 123, 124]. This type of optimization problem with a chance constraint, while very appealing, is often difficult to solve in practice due to high computational complexity; see [125] and [126] for a detailed discussion. To this end, instead of using the probabilistic constraint above, we consider a Kelly Betting problem subject to a constraint involving expected maximum drawdown. To be more specific, given target drawdown limit $0 < \varepsilon < 1$, we work with constraint

$$\mathbb{E}[d_K^*] \leq \varepsilon.$$

The reader is referred to [74] and [96] where the expected drawdown in a stock trading context is discussed.

3.2.2 A Simple Drawdown Control Example via Fractional Kelly: The literature on drawdown also includes a well-known method called the *Fractional Kelly Strategy.* This involves scaling down the size of investment for the purpose of mitigating this risk; e.g., see [12] and [25]. To demonstrate how a fractional Kelly strategy works, we now revisit the example used in Section 3.1.1. That is, we take random variable $X = 0.15$ with probability $p = 0.95$ and $X = -0.95$ with probability $p = 0.05$. Using $N = 252$, we now plot the expected maximum percentage drawdown $\mathbb{E}[d_K^*]$ versus the Kelly fraction K in Figure 3.6.

In this example, using the optimum fraction $K^* \approx 0.667$ already found, we see from Figure 3.6 that the corresponding expected maximum drawdown is

$$\mathbb{E}[d_K^*]\big|_{K=K^*} \approx 0.903.$$

In other words, the Kelly optimum leads to about 90.3% expected maximum drawdown, which is significantly large. We now imagine a gambler who is enamored with Kelly theory and decides to reduce K from its optimum $K^* = 0.667$ by included the constraint, say $\mathbb{E}[d_K^*] \leq 0.2$, as part of the optimization with respect to K. Using the fact that $g(K)$

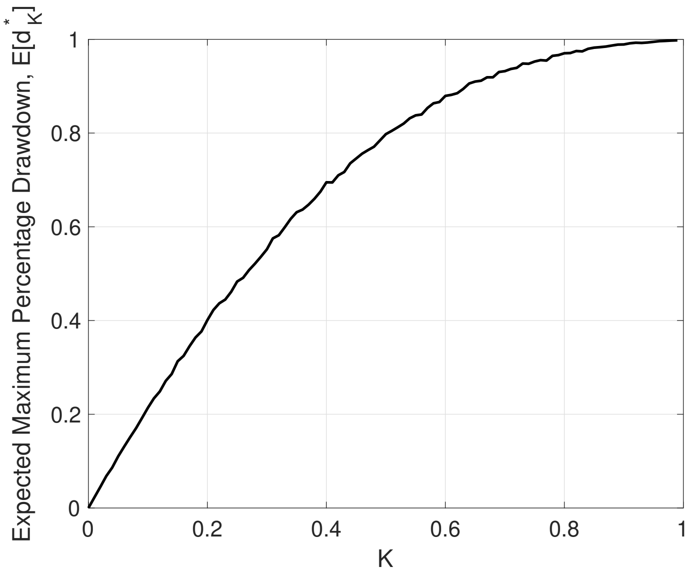


Figure 3.6: Expected Maximum Percentage Drawdown Versus K

is concave for all K and is increasing for $K < K^*$, with the aid on Figure 3.6 again, it is clear to see that the optimal fraction K reduces from 0.667 to $K = K^* \approx 0.1$. Similarly, if one use $\mathbb{E}[d_K^*] \leq 0.4$ as the drawdown constraint, then the associated optimal investment fraction K reduces from $K = 0.667$ to $K = K^* \approx 0.2$. To summarize, the use of the expected drawdown constraint is seen to systematically reduce K making its use less risky versus using the true optimum $K^* = 0.667$.

3.3 Two Conjectures Involving Drawdown

When the vector K is multi-dimensional, it is desirable to have a convex drawdown constraint so that the logarithmic growth optimization problem can be treated as convex program and thereby solved in an efficient way. To this end, in this section, we provide two conjectures involving convex dependence of the drawdown on K. For motivation, we first consider the case when K is a scalar. Indeed, it is clear in Figure 3.6 that expected drawdown is an increasing function of K. Thus, for $0 < \varepsilon < 1$, the constraint $\mathbb{E}[d_K^*] \leq \varepsilon$ leads to an interval restriction on K which is trivially convex.

To analyze a less trivial situation, we now consider two independent and identical coin-flipping type gambles for which $X_i(k) = 1$ with probability $p = 0.9$ and $X_i(k) = -1$ with probability $1 - p$ for $i = 1, 2$ and take $N = 252$. Now, to estimate the drawdown constraining set in the plane, we perform a Monte-Carlo simulation as follows: Given $0 < \varepsilon < 1$, for each pair $K = (K_1, K_2)$, we generate $1,000,000$ sample paths for $X(k)$, and then calculate the associated expected drawdown $\mathbb{E}[d_K^*]$ and estimate the associated constraint set $\{K = (K_1, K_2) : \mathbb{E}[d_K^*] \leq \varepsilon\}$. The result indicates that the constraint set for expected maximum drawdown appears not only to be convex but also polytopic; see Figure 3.7.

A similar finding is obtained for some multiple-coin scenarios which we considered. To be more specific, we take the return for the i-th coin be $X_i(k) \in \{-1/i, 1/i\}$ for $i = 1, 2, 3$, with probability $P(X_i(k) = 1/i) =\ 0.9$ and $N = 10$. Then, again, given $0 < \varepsilon < 1$, for each triple $K = (K_1, K_2, K_3)$, we generate $100,000$ sample paths for $X(k)$, and then we estimate the set $\{K = (K_1, K_2, K_3) : \mathbb{E}[d_K^*] \leq \varepsilon\}$, which is again a convex polytope. These findings lead to the following conjecture, which we believe holds not only for $\mathbb{R}^2$, but also for $\mathbb{R}^m$.

Conjecture 3.3.1 (Polytopic Expected Drawdown Constraint): *Given $0 < \varepsilon < 1$, the expected maximum drawdown constraint set given by*

$$\mathcal{K}_E \doteq \{K \in \mathcal{K} \subseteq \mathbb{R}^m : \mathbb{E}[d_K^*] \leq \varepsilon\}$$

is a convex polytope.

3.3.2 Probabilistic Drawdown Constraint: As mentioned in Section 3.2.1, we can also consider a probabilistic constraint requiring the maximum drawdown to stay below some prescribed level $\varepsilon > 0$ with probability $p = 1 - \delta$ or more. For the same two-coin example above, with $\varepsilon = \delta = 0.2$, a Monte-Carlo simulation again indicates that the probabilistic

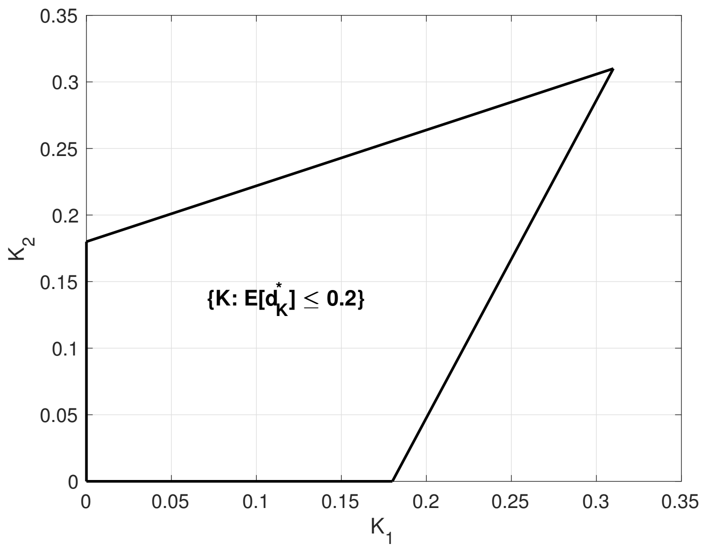


Figure 3.7: Example of Expected Maximum Drawdown Constraint

drawdown constraint set appears to be a convex polytope; see Figure 3.8. A similar finding was obtained for the three-coin scenario. This leads to the following conjecture.

Conjecture 3.3.3 (Polytopic Probabilistic Drawdown Constraint): *Given* $0 < \varepsilon < 1$, *and* $0 < \delta < 1$, *the probabilistic drawdown constraint set given by*

$$\mathcal{K}_P \doteq \{K \in \mathcal{K} \subseteq \mathbb{R}^m : P(d_K^* \leq \varepsilon) \geq 1 - \delta\}$$

is a convex polytope.

3.3.4 Remarks: The computational complexity required to generate both constraint sets is fairly high when Monte-Carlo simulation is used. Thus, obtaining a "clever" algorithm which estimates these sets would be of interest to pursue in future research. Although we have carried out several additional Monte-Carlo simulations supporting these conjectures, our lack of a proof does not rule out the possibility that the two conjectures above are false. In view of this possibility, in the next section, we introduce a "surrogate" for drawdown constraint which is an approximation. This leads to a new problem formulation which is a concave program. However, it is important to note that the surrogate is valuable only when it tightly bounds the ordinary drawdown.

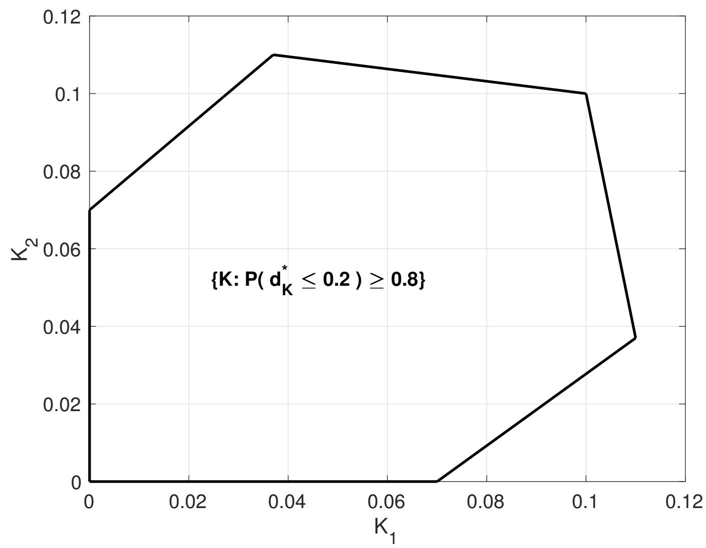


Figure 3.8: Example of Probability of Maximum Drawdown Constraint

3.4 Surrogate for Expected Maximum Drawdown

Given that the conjectures above may not be true, we now describe an alternative approach for handling expected drawdown constraints. We begin by noting that satisfaction of the desired stochastic inequality $d_K^* \leq \varepsilon$ is equivalent to $1 - d_K^* \geq 1 - \varepsilon$. Taking logarithms on both sides and using the monotonicity of logarithmic function, the desired inequality becomes $\log(1 - d_K^*) \geq \log(1 - \varepsilon)$. Now recalling Section 1.4.3, where the *complementary drawdown*

$$\begin{aligned}\overline{D}_K &\doteq 1 - d_K^* \\ &= \min_{0 \leq l \leq k \leq N} \frac{V(k)}{V(l)}\end{aligned}$$

is defined, the desired inequality becomes

$$\log \overline{D}_K \geq \log(1-\varepsilon).$$

This condition is central to the following lemma.

Lemma 3.4.1 (Convexity of Surrogate Drawdown): *Given* $0 < \varepsilon < 1$*, the surrogate expected maximum drawdown constraint set*

$$\mathcal{D}_K \doteq \left\{ \ K \in \mathcal{K} \subseteq \mathbb{R}^n : \mathbb{E}[\log \overline{D}_K] \geq \log(1-\varepsilon) \ \right\}$$

is convex.

Proof. Given $0 < \varepsilon < 1$, we have

$$\begin{aligned}
\mathbb{E}[\log \overline{D}_K] &= \mathbb{E}\left[\log\left(\min_{0\leq l\leq k\leq N} \frac{V(k)}{V(l)}\right)\right] \\
&= \mathbb{E}\left[\min_{0\leq l\leq k\leq N} \log \frac{V(k)}{V(l)}\right] \\
&= \mathbb{E}\left[\min_{0\leq l\leq k\leq N} \sum_{i=l}^{k-1} \log\left(1 + K^T X(i)\right)\right]
\end{aligned}$$

Observe that the function

$$\sum_{i=l}^{k-1} \log\left(1 + K^T x\right)$$

is concave in K. Using the fact that the minimum over an indexed collection of concave functions is concave, and the fact that the expectation preserves the concavity, it follows that $\mathbb{E}[\log \overline{D}_K]$ is a concave function. Hence, the surrogate expected maximum drawdown constraint set

$$\mathcal{D}_K = \{K \in \mathcal{K} : \mathbb{E}[\log \overline{D}_K] \geq \log(1-\varepsilon)\}$$

is convex. $\square$

3.4.2 Remarks: As mentioned in the introduction in this chapter, the surrogate is only valuable when the quality of this approximation is high. This is seen to be the case for the example which we consider below. To demonstrate this, We begin by noting that concavity of the log function enables us to use Jensen's inequality to obtain the bound

$$\mathbb{E}[\log \overline{D}_K] \leq \log \mathbb{E}[\overline{D}_K].$$

Now exponentiating on both sides, we obtain

$$\mathbb{E}[\overline{D}_K] \geq \exp(\mathbb{E}[\log \overline{D}_K]),$$

and, to consider the tightness of this bound, we revisit the single coin-flipping gamble again with probability $p = 0.9$ and $N = 252$. Figure 3.9, obtained via Monte-Carlo simulation, provides a comparison between $\mathbb{E}[\overline{D}_K]$ and $\exp(\mathbb{E}[\log \overline{D}_K])$. For this simple case, it is clear that $\mathbb{E}[\overline{D}_K]$ is very close to $\exp(\mathbb{E}[\log \overline{D}_K])$. In other words, this example suggests that the surrogate for the maximum expected drawdown may be useful in practice.

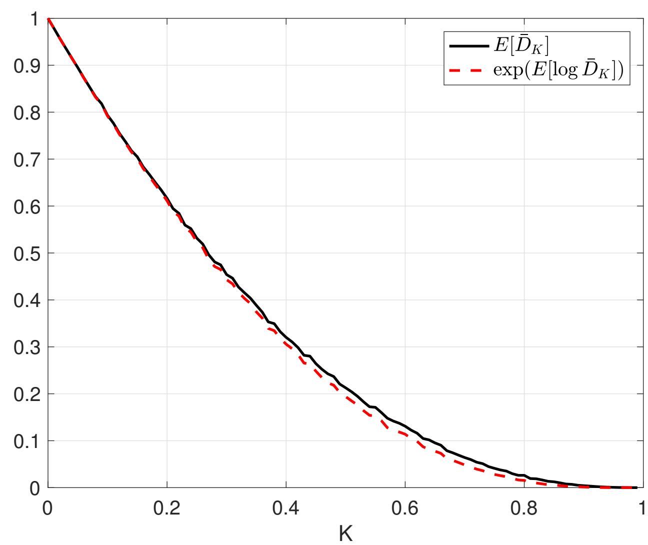


Figure 3.9: Expected Complementary Drawdown and its Surrogate

3.5 Concluding Remarks and Future Work

In this chapter, we pointed out some of the limitations in the existing literature associated with Kelly's theory. We first showed that a Taylor-based approximation technique, used in some papers, may lead to an approximate solution which may not be feasible. In addition, we also demonstrated that such an approximation may lead to an *inefficient strategy*. Subsequently, we considered the fact that using Kelly-based theory may incur potentially large drawdown. To study the drawdown issue, we first provided two conjectures with limited supporting evidence. That is, for both the expected drawdown constraint and the probabilistic drawdown constraint, we conjectured that the resulting constraint sets are convex polytopes. A proof for these two conjectures are relegated to future research. Finally, we introduced a surrogate drawdown measure which enables us to formulate an approximate drawdown-constrained Kelly-based optimization problem as a concave program and thereby solved efficiently.

Our study of drawdown issues, conjectures and Lemma 3.4.1 suggest a direction for future research which involves approximation of the expected logarithmic growth optimization problem to bring convex programming into play. In the next chapter, we further pursue the issue of drawdown. To this end, we describe a new control-theoretic framework which deals with the drawdown issue in an almost sure sense.

Chapter 4

Drawdown-Modulated Control Systems

As discussed in previous chapters, control of drawdown, that is, control of the drops in wealth over time from peaks to subsequent lows, is of great concern from a risk management perspective; e.g., see the discussion of references [30, 41, 74, 93–100, 104, 119–123] in Sections 1.4.2 and 3.2–3.4. In this chapter, we provide some new results on drawdown which are aimed at "probability one" guarantees that drawdown will not exceed a prescribed maximum level $d_{\max}$.[1] Specifically, beginning with a control-theoretic formulation in a stock trading context, we first provide a result, which we call the *Drawdown Modulation Lemma*, serving as a stepping stone to the remainder of the chapter. This lemma characterizes investments which guarantee that the *percentage drawdown* is no greater than a prespecified level $0 < d_{\max} < 1$ for all sequences of returns. As an immediate consequence, an investment satisfying the requirement of the lemma guarantees survival. That is, the no-bankruptcy condition $V(k) > 0$ for all k is assured.

Subsequently, based on this lemma, a new trading scheme which we call *drawdown-modulated feedback control* is introduced in Section 4.3. Then, using this control scheme, in Section 4.4, we consider a drawdown-constrained problem of maximizing the expected logarithmic growth. Then, in Section 4.5, a numerical example using historical data is provided, and in Section 4.6, a further generalization of the Drawdown Modulation Lemma for the multi-asset portfolio case is provided. In the sequel, we use the definitions of percentage drawdown to date $d(k)$, and overall percentage drawdown d^* as defined in Section 3.2. For the sake of completeness, we repeat the definition below again. That is, for $k = 0, 1, 2, \ldots, N$,

[1] Part of the work reported in this chapter has been published in [90].

with $V(k)$ being the account value at stage k, we take

$$d(k) \doteq \frac{V_{\max}(k) - V(k)}{V_{\max}(k)}$$

where

$$V_{\max}(k) \doteq \max_{0 \leq i \leq k} V(i)$$

and overall percentage drawdown

$$d^*_{\max} \doteq \max_{0 \leq k \leq N} d(k).$$

Note that the percentage drawdown satisfies $0 \leq d(k) \leq 1$.

4.1 Stock-Trading Formulation

Consistent with the conventions in the preceding chapters, for $k = 0, 1, \ldots, N$, we let $S(k) > 0$ denote the stock price and take $X(k)$ to be the associated returns with the standing assumption that $X_{\min} \leq X(k) \leq X_{\max}$ with $X_{\min}$ and $X_{\max}$ being points in the support, denoted by $\mathcal{X}$, and satisfying $-1 < X_{\min} < 0 < X_{\max} < \infty$. As in the preceding chapters, we also assume that stock trading occurs within an "idealized market" with zero transaction costs, zero interest rates and perfect liquidity. As previously mentioned, these assumptions arise in the finance literature in the context of "frictionless" markets; see [48] and [50].

In addition, with $I(k)$ being the investment at stage k, and with the convention that $I(k) < 0$ corresponds to short selling, beginning at initial account value $V(0) > 0$, we recall that the account value update is given by

$$V(k+1) = V(k) + I(k)X(k).$$

Given a *maximum acceptable drawdown level* $d_{\max}$ satisfying $0 < d_{\max} < 1$, we focus on conditions on $I(k)$ under which satisfaction of the constraint $d(k) \leq d_{\max}$ is assured along all sequences of returns and all k.

4.2 The Drawdown Modulation Lemma

In this section, the *Drawdown Modulation Lemma*, discussed in the previous section, is given. This lemma provides a necessary and sufficient condition on the investment $I(k)$ which guarantees that the percentage drawdown is no greater than a given level $d_{\max}$ for all sequences of returns with $X(k)$ satisfying $X_{\min} \le X(k) \le X_{\max}$.

Lemma 4.2.1 (Scalar Drawdown Modulation): *An investment function $I(k)$ guarantees a maximum acceptable drawdown of $d_{\max}$ or less for all sequences of returns if and only if for all k, the condition*

$$-\frac{d_{\max} - d(k)}{(1 - d(k))\, X_{\max}} V(k) \le I(k) \le \frac{d_{\max} - d(k)}{(1 - d(k))\, |X_{\min}|} V(k)$$

is satisfied for all sequences of returns.

Proof. To prove necessity, assuming that $d(k) \le d_{\max}$ for all k and all sequences of returns, we must show the required condition on $I(k)$ holds along all sequences of returns. Letting k be given, since both $d(k) \le d_{\max}$ and $d(k+1) \le d_{\max}$ for all sequences of returns, we claim this forces the required inequalities on $I(k)$. Without loss of generality, we provide a proof of the right-hand inequality for the case $I(k) \ge 0$ and note that a nearly identical proof is used for $I(k) < 0$. To establish the condition on $I(k)$ for all sequences of returns, it suffices to consider the path with worst loss $|X_{\min}|I(k)$. In this case, we have $V_{\max}(k+1) = V_{\max}(k)$. Hence,

$$d(k+1) = d(k) + \frac{|X_{\min}|I(k)}{V_{\max}(k)} \le d_{\max}.$$

We now substitute

$$V_{\max}(k) = \frac{V(k)}{1 - d(k)} > 0$$

into the inequality above and obtain

$$|X_{\min}|\, I(k) \le \frac{d_{\max} - d(k)}{1 - d(k)} V(k),$$

which implies that

$$I(k) \leq \frac{d_{\max} - d(k)}{(1 - d(k))\,|X_{\min}|} V(k).$$

To prove sufficiency, assuming that the stated bounds on $I(k)$ hold along all sequences of returns, we must show $d(k) \leq d_{\max}$ for all k and all sequences of returns. Proceeding by induction, for $k = 0$, we trivially have $d(0) = 0 \leq d_{\max}$. To complete the inductive argument, we assume that $d(k) \leq d_{\max}$ for all sequences of returns, and must show $d(k+1) \leq d_{\max}$ for all sequences of returns. Without loss of generality, we again provide a proof for the case $I(k) \geq 0$ and note that a nearly identical proof is used for $I(k) < 0$. Indeed, by noting that

$$d(k+1) = 1 - \frac{V(k+1)}{V_{\max}(k+1)},$$

and $V_{\max}(k) \leq V_{\max}(k+1)$ for all sequences of returns, we split the argument into two cases: If $V_{\max}(k) < V_{\max}(k+1)$, then $V_{\max}(k+1) = V(k+1)$ and we have $d(k+1) = 0 \leq d_{\max}$. On the other hand, if $V_{\max}(k) = V_{\max}(k+1)$, with the aid of the dynamics of account value, we have

$$\begin{aligned} d(k+1) &= 1 - \frac{V(k) + I(k)X(k)}{V_{\max}(k)} \\ &\leq 1 - \frac{V(k) - I(k)\,|X_{\min}|}{V_{\max}(k)}. \end{aligned}$$

Using the upper bound on $I(k)$, we obtain $d(k+1) \leq d_{\max}$ which completes the proof. □

4.2.2 Remarks: We note that if one has a stochastic model; i.e., $X(k)$ being random processes satisfying the known bounds on $X_{\min}$ and $X_{\max}$, then the drawdown guarantee $d(k) \leq d_{\max}$ holds with probability one. However, the lemma above tells us that even if no stochastic model is available, we still have the guarantee. Any investment $I(k)$ satisfying the two inequalities in the statement of lemma assures *survival*; i.e., no bankruptcy. That is, along all sample paths, the account value $V(k) > 0$ with probability one for $k = 0, 1, 2, \ldots, N$.

This is an easy consequence of the fact that at each stage k, drawdown of 100% from some previous maximum for $V(k)$ never occurs. Second, the lemma above opens a door to solution of new drawdown-constrained optimization problems involving parameters entering into an investment $I(k)$; see Section 4.4.

4.3 Drawdown-Modulated Feedback Control Realization

With the aid of the Drawdown Modulation Lemma, we now describe a class of investment functions $I(k)$ expressed as a linear feedback control parameterized by a gain γ and leading to satisfaction of the drawdown specification. To be more specific, for $k = 0, 1, \ldots, N-1$, we define

$$M\left(k\right) \doteq \frac{d_{\max} - d\left(k\right)}{1 - d\left(k\right)}$$

which we call *modulator*. Now, using this function, we work with investment

$$I(k) \doteq \gamma M(k) V(k)$$

where

$$\gamma \in \Gamma \doteq \left[-\frac{1}{X_{\max}}, \frac{1}{|X_{\min}|} \right]$$

which guarantees survival. Note that the feedback gain γ can be selected without regard for the modulator $M(k)$. This has a similar flavor to that of the celebrated Separation Theorem in linear control theory; e.g., see [102]. Henceforth, we call the investment $I(k)$ described above a *drawdown-modulated feedback controller*. The closed-loop configuration which results is depicted in Figure 4.1.

4.3.1 Remarks: The limiting case obtained by letting $d_{\max} \to 1$ leads to $M(k) \to 1$ which brings us back to the classical Kelly optimization considered in previous chapters. Recalling the discussion in Section 1.1.5, since the Kelly solution is often deemed to be

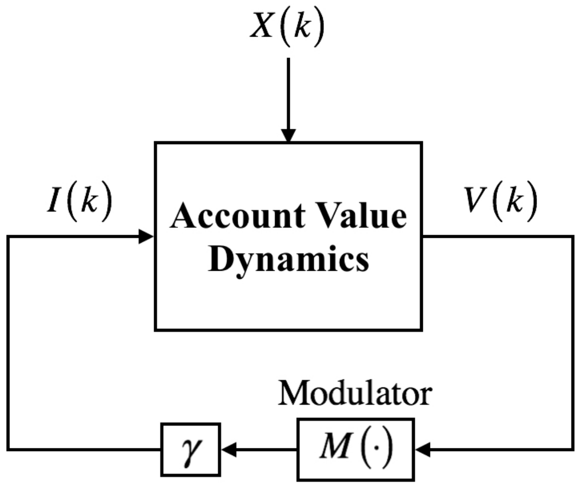


Figure 4.1: Drawdown-Modulated Feedback Configuration

too risky, the often-used Fractional Kelly strategy is arguably designed in an ad-hoc manner. In contrast, the drawdown-modulated feedback control provides a new and systematic way to address drawdown with the added advantage that is time-varying and adapts the investment level based on the sample path being encountered. Closer inspection of the controller above indicates that the adjustment of the investment level based on drawdown to date can be viewed as a form of *trend following*; e.g., see [127–132]. It is readily verified that $0 \leq M(k) \leq d_{\max}$ guarantees

$$0 \leq \mathbb{E}[M(k)] \leq d_{\max}.$$

We also note that the feedback-control realization above is not the only one leading to satisfaction of the drawdown requirement $d(k) \leq d_{\max}$ with probability one. A more general class of feedback controllers can be formulated with a time-varying gain $\gamma(k)$. That is, one can equally well take

$$I(k) = \gamma(k)M(k)V(k)$$

with $\gamma(k) \in \Gamma$ and still satisfy the drawdown requirement; see the conclusion of this chapter for further discussion of promising research along these lines.

4.4 Drawdown-Modulated Kelly Optimization

In this section, we consider one of the possible ways to bring the Drawdown Modulation Lemma into the Kelly optimization framework. To this end, we now assume that the returns $X(k)$ are i.i.d. random variables with probability density function f_X, and we work with the drawdown-modulated feedback controller

$$I(k) = \gamma M(k)V(k)$$

with $\gamma \in \Gamma$ treated as an optimization variable. Now letting $V_\gamma(k)$ be the account value at stage k induced by the feedback gain γ, the associated recursion is described by

$$V_\gamma(k+1) = (1 + \gamma M(k)X(k))V_\gamma(k).$$

We seek an optimizer $\gamma = \gamma^*$ which attains

$$g^* \doteq \max_{\gamma \in \Gamma} \frac{1}{N} \mathbb{E}\left[\log \frac{V_\gamma(N)}{V(0)}\right].$$

In the next section, we consider a numerical example, based on historical data, to compare the trading performance obtained via drawdown-modulated feedback control with that obtained via a classical Kelly solution.

4.5 Numerical Example

In this section, we provide a numerical example involving real stock data to illustrate how drawdown modulation might be used in practice. We consider the Kelly Optimization Problem described in the last section and use historical price data for Tesla Motors (ticker: TSLA) covering the period December 31, 2013 to March 28, 2014; see Figure 4.2 where these underlying stock prices are plotted. The figure also shows an additional sixty-one days within the period from March 28, 2014 to June 24, 2014 of stock prices which will be used in the out-of-sample test described below.

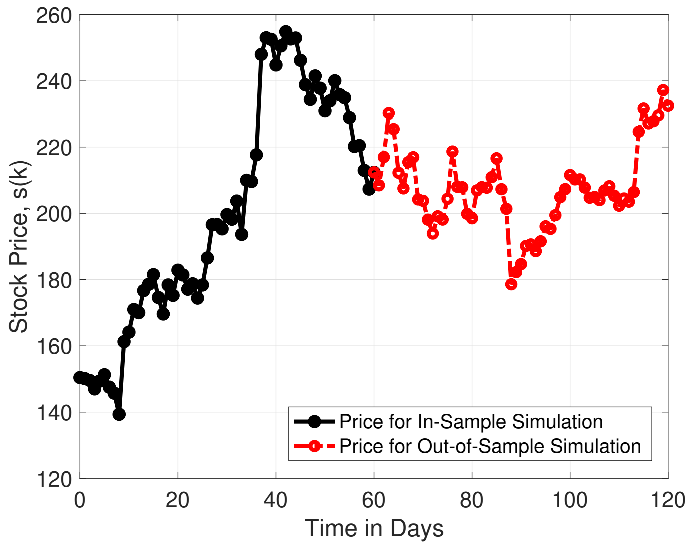


Figure 4.2: TSLA Stock Price, $s(k)$

As a first step, we use the price data $s(k)$ to estimate a probability mass function. That is, we first calculate the corresponding *realized returns*, called $x(k)$, where

$$x(k) \doteq \frac{s(k+1) - s(k)}{s(k)}.$$

Now letting $x_i = x(i-1)$ denote the i-th calculated return for $i = 1, 2, \ldots, 60$, we obtain the estimated PMF of the returns as the sum of Dirac Delta functions

$$\widehat{f}_X(x) = \frac{1}{60} \sum_{i=1}^{60} \delta(x - x_i)$$

which is used as input to the optimization to be carried out. This PMF, plotted in Figure 4.3, has $X_{\min} \approx -0.049$ and $X_{\max} \approx 0.157$. Hence, the constraint set on γ is given by

$$\Gamma = [-6.354,\ 20.248]$$

for the Kelly optimization problem to be solved.

Now using this probability mass function, we seek to find an optimal feedback gain $\gamma^* \in \Gamma$ which maximizes the objective function

$$J(\gamma) \doteq \frac{1}{N}\mathbb{E}\left[\log \frac{V_\gamma(N)}{V(0)}\right]$$

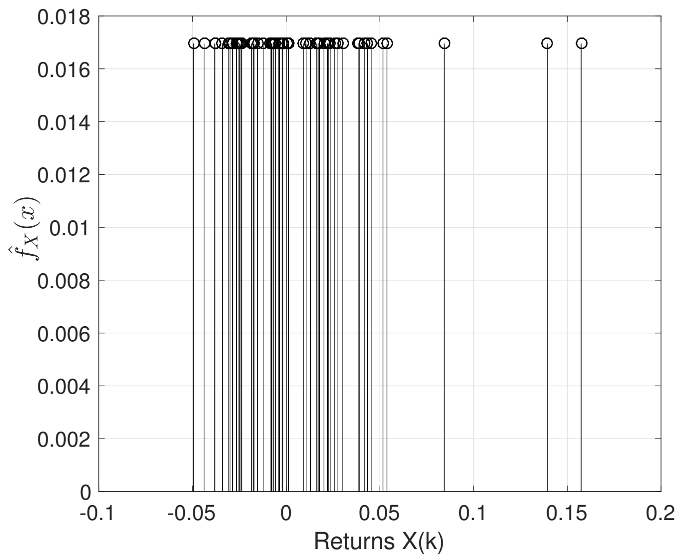


Figure 4.3: Estimated PMF $\widehat{f}_X(x)$ of Returns

with $V_\gamma(N)$ being the account value resulting from feedback gain γ. In this example we use $d_{\max} \doteq 0.05$ as the maximum acceptable drawdown level. Next, to evaluate $J(\gamma)$ for each fixed γ, we perform Monte-Carlo simulation to generate $100,000$ sample paths for $S(k)$ from the PMF. Then we use these to estimate $\mathbb{E}[\log V_\gamma(N)]$ which is needed for the J function evaluation. We note that use of the modulator automatically assures that the drawdown requirement is satisfied. Using the plot of the expected logarithmic growth in Figure 4.4, we obtain a maximizer for $J(\gamma)$ given by $\gamma^* \approx 11.15$.

4.5.1 Cash-Financing Considerations and Trading Performance: Once γ^* is determined, we compare the trading performance obtained via a drawdown-modulated feedback control with that obtained via the classical Kelly solution. To make the comparison fair, in the simulation below, we impose a cash-financing constraint on both strategies; i.e., we saturate each of the two strategies so that $0 \leq I(k) \leq V(k)$ is satisfied. That is, for drawdown-modulated feedback controller, along sample paths, instead of using

$$I^*(k) \doteq \gamma^* M(k) V(k)$$

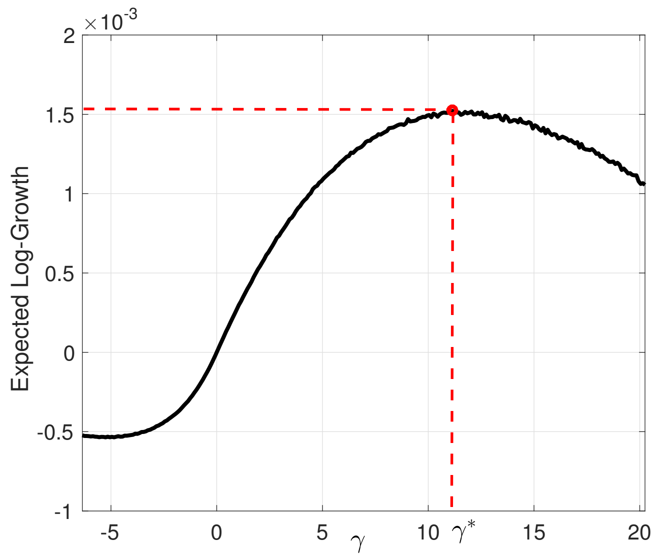


Figure 4.4: Expected Growth Versus γ via Drawdown Modulation

with $\gamma^* \approx 11.15$, we work with

$$I_{sat,mod}(k) \doteq \min\{I^*(k), V(k)\}.$$

Similarly, for the classical Kelly strategy, instead of using $I(k) = K^*V(k)$ with optimum $K^* \approx 5.318$, we use $I_{sat,Kelly}(k) \doteq 1 \cdot V(k)$.

Now, to examine the trading performance, we take the initial account value $V(0) = 10,000$ and in Figure 4.5, we depict the sixty-one days in-sample simulation result for the account value $V(k)$. In particular, we see that the drawdown-modulated feedback controller $I_{sat,mod}(k)$ leads to account value given by $V(60) = 1.146 \times 10^4$ with the overall percentage drawdown $d^*_{\max} \approx 0.047$ which is within the allowed upper limit of 5%. As mentioned earlier, for comparison purposes, we also computed the classical Kelly solution which is obtained as a special case of our formulation as $d_{\max} \to 1$. In this case, the Kelly strategy $I_{sat,Kelly}(k)$ leads to the account value $V(60) = 1.412 \times 10^4$ and the overall percentage drawdown $d^*_{\max} \approx 0.186$ which is much larger than the allowed 5% specification for $d_{\max}$.

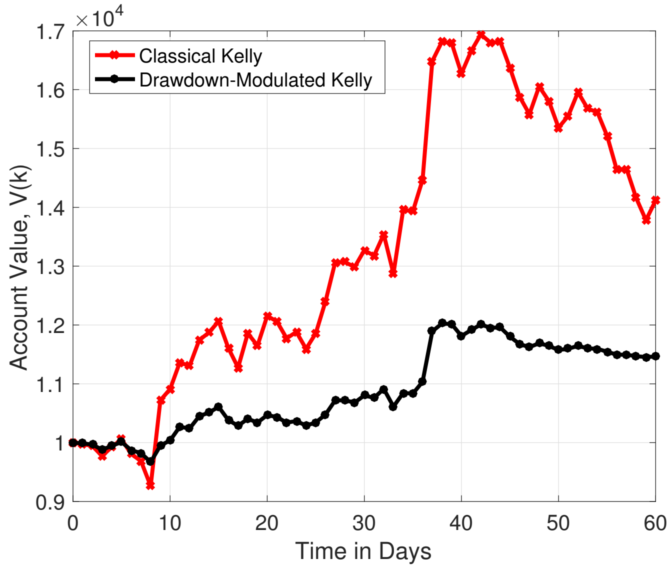


Figure 4.5: Account Value Under Two Strategies (In-Sample)

Next, we evaluated the out-of-sample performance for the second segment of stock prices given in Figure 4.2. Again, beginning with $V(60) = 10,000$, the resulting account value $V(k)$ for $k \geq 60$ is depicted in Figure 4.6. We see that the drawdown-modulated control $I_{sat,mod}(k)$ leads to the terminal account value $V(120) \approx 1.005 \times 10^4$ with the overall percentage drawdown $d^*_{\max} \approx 0.05$ as required. In contrast, the classical Kelly strategy $I_{sat,Kelly}(k)$ leads to the terminal account value $V(120) \approx 1.136 \times 10^4$ and the overall percentage drawdown $d^*_{\max} \approx 0.225$ which corresponds to 22.5%.

While this level is much higher than the specification $d_{\max} = 0.05$, the terminal account value is considerably higher. This higher return is to be expected since the drawdown risk was neglected. There is one interesting observation that can be made from Figure 4.6. That is, when the maximum drawdown acceptable level $d_{\max}$ is hit, then the investment $I_{sat,mod}(k)$ becomes zero. Thus, we see that the drawdown-modulated feedback control yields a "stop-loss" type of behavior.

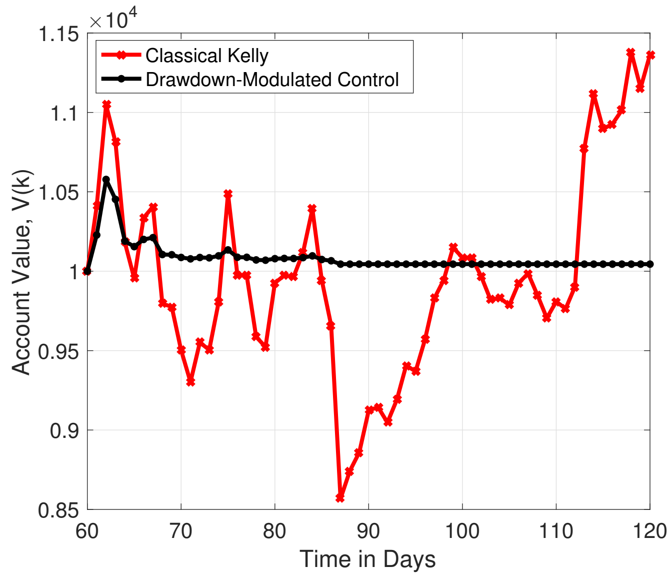


Figure 4.6: Account Value Under Two Strategies (Out-of-Sample)

4.6 The Generalized Drawdown Modulation Theorem

In this section, the Drawdown Modulation Lemma of Section 4.2 is generalized to address a portfolio of m assets. For the sake of simplicity, in the sequel, we consider m stocks. However, we note that if one or more of the assets is riskless, such as a bond, the analysis to follow is trivially modified. Now, with $S_i(k)$ being the i-th stock price, we take the returns to be

$$X_i(k) \doteq \frac{S_i(k+1) - S_i(k)}{S_i(k)}$$

for $i = 1, 2, \ldots, m$ and vectorize it as

$$X(k) \doteq [X_1(k)\ X_2(k)\ \cdots X_m(k)]^T.$$

As in the scalar case, we assume known bounds $X_{\min,i} \leq X_i(k) \leq X_{\max,i}$ with

$$-1 < X_{\min,i} < 0 < X_{\max,i}$$

for $i = 1, 2, \ldots, m$ and $k = 0, 1, 2, \ldots, N-1$. We let $\mathcal{V}$ denote the 2^m vertices of the hypercube defined by $X_{\min,i}$ and $X_{\max,i}$ above. Henceforth, for simplicity of notation, we take $X_{\min}$ and $X_{\max}$ to be vectors with i-th components $X_{\min,i}$ and $X_{\max,i}$, respectively and we let $|X_{\min}|$ denote the vector with i-th component $|X_{\min,i}|$. Now letting $I_i(k)$ be the i-th component of the investment vector $I(k)$, the associated account value is updated as

$$V(k+1) = V(k) + I^T(k)X(k).$$

We also denote the positive and negative parts of $I(k)$ component-wise by

$$I_i^+(k) \doteq \max\{I_i(k), 0\}$$

and

$$I_i^-(k) \doteq \min\{I_i(k), 0\}$$

respectively, we are now prepared to this m-stock case. The proof, although more general, proceeds along similar lines to that of the Lemma 4.2.1.

Theorem 4.6.1 (Generalized Drawdown Modulation): *An investment function $I(\cdot)$ guarantees a maximum acceptable drawdown of $d_{\max}$ or less for all sequences of returns if and only if for all k, the condition*

$$\left|X_{\min}^T\right| I^+(k) - X_{\max}^T I^-(k) \leq M(k)V(k)$$

is satisfied for all sequences of returns.

Proof. To prove necessity, assuming that $d(k) \leq d_{\max}$ for all k and all sequences of returns, we must show the required condition on $I(k)$ holds along all sequences of returns. Letting k be given, since both $d(k) \leq d_{\max}$ and $d(k+1) \leq d_{\max}$ for all sequences of return, we claim this forces the required inequalities on $I(k)$. With the i-th component $I_i(k)$ being either $I_i^+(k)$ or $I_i^-(k)$, there exists a vertex $v \in \mathcal{V}$ such that

$$v^T I(k) = X_{\max}^T I^-(k) + X_{\min}^T I^+(k) \leq 0.$$

Noting that in this case with loss $|v^T I(k)|$, we also have $V_{\max}(k+1) = V_{\max}(k)$. It follows that

$$d(k+1) = d(k) - \frac{v^T I(k)}{V_{\max}(k)} \leq d_{\max}.$$

Now, substituting

$$V_{\max}(k) = \frac{V(k)}{1 - d(k)} > 0$$

into the inequality above and obtain

$$-v^T I(k) \le \frac{d_{\max} - d(k)}{1 - d(k)} V(k)$$

for any sequences of returns. This implies that

$$|X_{\min}^T| I^+(k) - X_{\max}^T I^-(k) \le M(k) V(k).$$

To prove sufficiency, we assume the condition on $I(k)$ holds for all sequences of returns. We must show $d(k) \le d_{\max}$ for all k and all sequences of return. Proceeding by induction, for $k = 0$, we trivially have $d(0) = 0 \le d_{\max}$ for all sequences of returns. To complete the inductive argument, we assume that $d(k) \le d_{\max}$ for all sequences of return, and must show $d(k+1) \le d_{\max}$ for all sequences of return. Now, by noting that

$$d(k+1) = 1 - \frac{V(k+1)}{V_{\max}(k+1)}$$

and $V_{\max}(k) \le V_{\max}(k+1)$, we split the argument into two cases: If $V_{\max}(k) < V_{\max}(k+1)$, then $V_{\max}(k+1) = V(k+1)$. Thus, we have $d(k+1) = 0 \le d_{\max}$. On the other hand, if $V_{\max}(k) = V_{\max}(k+1)$, with the aid of the dynamics of account value, we have

$$\begin{aligned} d(k+1) &= 1 - \frac{V(k) + I^T(k) X(k)}{V_{\max}(k)} \\ &\le 1 - \frac{V(k) - \left(\left|X_{\min}^T\right| I^+(k) - X_{\max}^T I^-(k)\right)}{V_{\max}(k)}. \end{aligned}$$

Using the given inequality condition on $I(k)$, we obtain $d(k+1) \le d_{\max}$ which completes the proof. □

4.6.2 Remarks: The drawdown-modulated feedback realization described for the single-stock case is generalized to this multi-stock case as follows: For the i-th stock, we use feedback gain γ_i and take the investment to be

$$I_i(k) \doteq \gamma_i M(k) V(k).$$

Then, the condition in the Generalized Drawdown Modulation Theorem 4.6.1 above leads to the constraint on γ as follows:

$$|X_{\min}^T|\gamma^+ - X_{\max}^T\gamma^- \leq 1$$

where γ^- and γ^+ have i-th component

$$\gamma_i^- \doteq \min\{\gamma_i, 0\}; \quad \gamma_i^+ \doteq \max\{\gamma_i, 0\}.$$

4.7 Concluding Remarks and Future Work

In this chapter we introduced a new trading scheme, which we called drawdown-modulated feedback control. It enables us to express the investment $I(k)$ as a linear time-varying feedback realization which leads to satisfaction of a given percentage drawdown specification for all sequences of returns. We also provided an illustration of our theory in the context of classical Kelly Betting. The resulting drawdown-modulated feedback control which we obtain can be viewed as a time-varying fractional Kelly strategy which includes a guarantee that the drawdown requirements will be satisfied.

Further to the optimization issue considered in this chapter, we now draw attention to some possible future research directions here. Indeed, as seen in this chapter, the optimal feedback γ^* we obtained is simply a constant gain. Thus, one immediate direction would be to allow a time-varying feedback gain $\gamma(k)$ and compare the performance with the case of constant γ. That is, this time-varying feedback gain must satisfy

$$\gamma(k) \in \Gamma_{tv} \doteq \{\gamma(\cdot) : \gamma(k) \in \Gamma \text{ for all } k \geq 0\}.$$

Then the problem is formulated by finding a function $\gamma \in \Gamma_{tv}$ achieving

$$V_{tv}^* \doteq \max_{\gamma \in \Gamma_{tv}} \frac{1}{N}\mathbb{E}\left[\log \frac{V_\gamma(N)}{V(0)}\right].$$

Since a constant γ defines a subset of Γ_{tv} above, we envision that superior performance can be achieved in this time-varying setting.

4.7.1 New Drawdown-Based Portfolio Design Problems: Another interesting problem for future research involves drawdown-based portfolio optimization. That is, we consider a risk-return pair obtained using *expected* maximum drawdown in combination with the *expected* return. Then, with constraint $\gamma \in \Gamma$ as in Section 4.6, we consider an optimization problem which involves selection of

$$\mathcal{M} \doteq (\gamma, d_{\max}) = (\gamma_1, \dots, \gamma_m, d_{\max}) \in \Gamma \times (0, 1)$$

and note that the $d_{\max}$ above is now taken to be a design parameter. Now, we express the expected return as

$$\overline{R}_{\mathcal{M}} \doteq \mathbb{E}\left[\frac{V_{\mathcal{M}}(N) - V(0)}{V(0)}\right]$$

where the notation $V_{\mathcal{M}}(N)$ is used to emphasize the dependence of the final account value on drawdown-modulated feedback $\mathcal{M}$. Similarly, the expected maximum percentage drawdown is written as $\overline{d}_{\mathcal{M}} \doteq \mathbb{E}[d^*_{\mathcal{M}}]$. Then, for any given target level of expected return, call it $\widehat{R}$, we seek an admissible drawdown-modulated feedback $\mathcal{M} \in \Gamma \times (0, 1)$ minimizing $\overline{d}_{\mathcal{M}}$ subject to the constraint $\overline{R}_{\mathcal{M}} = \widehat{R}$. That is, we seek

$$\begin{aligned} &\inf_{\mathcal{M} \in \Gamma \times (0,1)} \overline{d}_{\mathcal{M}} \\ &\text{subject to } \overline{R}_{\mathcal{M}} = \widehat{R}. \end{aligned}$$

There is also a dual problem formulation described as follows: Given any target level of expected drawdown, call it $\widehat{d} \in (0, 1)$, we seek an admissible $\mathcal{M} \in \Gamma \times (0, 1)$ maximizing $\overline{R}_{\mathcal{M}}$ subject to the constraint $\overline{d}_{\mathcal{M}} = \widehat{d}$. It is also important to note that while our new problem is focused on controlling expected drawdown, as we saw in Theorem 4.6.1, once $d_{\max}$ is determined, the associated drawdown protection for all sequences of returns is also assured.

In the next chapter, we further pursue the drawdown-modulated scheme when dealing with a single risky asset. With risk-reward pair being expected drawdown and expected return, we show that the drawdown-modulated feedback strategy has a certain efficiency property when in competition with the classical linear time-invariant (LTI) controller. To be more specific, we show that given a target risk level, the drawdown-modulated feedback strategy can achieve a higher reward level than *any* LTI feedback strategy. This includes the classical Kelly strategy as a special case.

Chapter 5

An Efficiency Result for Drawdown-Modulated Feedback

In this chapter, we continue the study of *drawdown* in recursive betting scenarios and related applications in the stock market. As in Chapter 4, we dispense with the classical variance and the expected return as the risk-reward pair. Here we work with expected drawdown and mean return instead; e.g., see [41, 104] in the finance literature for the study using classical variance and mean return as the risk-reward pair. While mean-variance is quite useful, it relies on the assumption that the returns are normally distributed. Thus, if the distribution of returns is skewed, then the use of such a risk-return metric may be misleading; e.g., see [30, 74, 85] for more detailed discussion. Some studies account for higher order moments can be found in [133] and [134] and, as mentioned in Chapter 4, the issue of drawdown has received a considerable attention in the finance literature; e.g., see [94, 96, 98–100, 122].

Our analysis begins with a well-known principle which is widely used in finance. An investment opportunity is said to be *efficient* if the following condition holds: There is no other opportunity available with lower risk and at least as high return. In this framework, with risk-reward pair being expected drawdown and expected return, we establish the following: Given any linear time-invariant (LTI) feedback $I(k) = KV(k)$ for the investment, there exists a modulated feedback controller $I_{\mathcal{M}}(k) = K(k)V(k)$ with its time-varying gain $K(k) \doteq \gamma M(k)$ leading to a return which is at least as high as that of its LTI competitor and with no greater risk.[1] That is, from an efficiency point of view, in a design scenario, there is no loss of generality neglecting the LTI controller in favor of the drawdown-modulated controller. As a

[1] The work reported in this chapter has been published in [91].

bonus, it is also seen that the modulator assures a worst-case level of drawdown protection for all sequences of admissible returns.

After we dispense with preliminaries in Sections 5.1–5.2, the driver for the remainder of this chapter is a simple result, which we call the *Efficiency Lemma*, which is made possible by the fact that the modulator $M(k)$ includes memory of account values $V(0), V(1), \ldots, V(k-1)$ whereas the constant feedback investment strategy is memoryless. Subsequently, some examples, both theoretical and numerical, are provided. In the examples we actually see that the LTI feedback controller is *dominated* by the modulator. By this we mean the following: Given an LTI controller, there exists a modulator which leads to the same level of risk with strictly greater return.

5.1 Introduction

As previously stated, in this chapter, our risk-reward pair is expected maximum percentage drawdown and the expected return, respectively. In the sequel, our understanding is that $I(k)$ denotes either an "investment" or a "bet." As in previous chapters, we take $V(k)$ to be the account value of an investor or bettor at stage k. For the case of the drawdown modulator, we work with investments of the form

$$I(k) = \gamma M(k)V(k),$$

in competition with the modulator, a linear time-invariant feedback has

$$I(k) = KV(k)$$

with the constant K which represents proportion of the account invested. As indicated at the beginning of this chapter, given two investment opportunities, if one of them has larger risk and possibly lower return, it will be deemed to be *inefficient* and generally rejected in the marketplace.

5.2 Preliminaries for Efficiency Analysis

As in previous chapters, we take the returns $X(k)$ to be identically distributed (i.i.d.) random variables with bounds $-1 < X_{\min} \leq X(k) \leq X_{\max}$ with $X_{\min}$ and $X_{\max}$ being points in the support, denoted by $\mathcal{X}$, and satisfying $X_{\min} < 0 < X_{\max} < \infty$. Now, with linear feedback investment at stage k is given by $I(k) = KV(k)$ and, as discussed in Chapter 2, the constraint

$$K \in \mathcal{K} \doteq \left[\frac{-1}{X_{\max}}, \frac{1}{|X_{\min}|}\right]$$

is imposed due to bankruptcy considerations. In the sequel, we also allow short selling with $I(k) < 0$. Thus, at stage k, $I(k) > 0$ leads to a profit when $X(k) > 0$ and $I(k) < 0$ leads to a profit when $X(k) < 0$. Now beginning at initial account value $V(0) > 0$, as in preceding chapters, for $k = 1, 2, \ldots, N$, we work with the account value

$$V(k) = \prod_{i=0}^{k-1} (1 + KX(i))\, V(0).$$

Whenever convenient, we write $V_K(k)$ instead of $V(k)$ to emphasize the dependence on the feedback gain K.

5.2.1 Attainable Set for the Linear Time-Invariant Feedback: Given the LTI feedback investment $I(k) = KV(k)$ with $K \in \mathcal{K}$, we let R_K denote the *overall return* by

$$R_K \doteq \frac{V_K(N) - V(0)}{V(0)},$$

which we write more explicitly in terms of the sample path $X = (X(0), \ldots, X(N-1))$; i.e.,

$$R_K = \prod_{k=0}^{N-1} (1 + KX(k)) - 1$$

and we denote the corresponding maximum percentage drawdown by

$$d_K^* \doteq \max_{0 \leq k \leq N} d_K(k),$$

where $d_K(k)$ is defined in Section 3.2. The subscript K in R_K and d_K^* above is used to emphasize the dependence on the feedback gain K. In the sequel, we use the expected values of R_K and d_K^* as the reward-risk pair, and write

$$\overline{R}_K \doteq \mathbb{E}[R_K]$$

and

$$\overline{d}_K^* \doteq \mathbb{E}[d_K^*].$$

With the setup above, the *attainable risk-return performance curve* in the plane is

$$\mathscr{R}_K \doteq \{(\overline{R}_K, \overline{d}_K^*) : K \in \mathcal{K}\}$$

where

$$\overline{R}_K = \mathbb{E}\left[\prod_{k=0}^{N-1} (1 + KX(k))\right] - 1$$

and, using the i.i.d. assumption on the $X(k)$, we arrive at

$$\begin{aligned}\overline{R}_K &= (1 + K\mathbb{E}\left[X(k)\right])^N - 1\\ &= (1 + K\mu)^N - 1\end{aligned}$$

where $\mu \doteq \mathbb{E}[X(k)]$. As far as calculation of $\overline{d}_K^*$ is concerned, except for small values of N, Monte-Carlo simulations will be used to calculate this quantity.

5.2.2 Remarks: Further to $\overline{d}_K^*$, it is straightforward to see that the worst-case percentage drawdown can get as bad as

$$d_{K,worst} \doteq 1 - (1 - |K| \max\{|X_{\min}|, X_{\max}\})^N$$

which is somewhat less useful than the expected percentage drawdown $\overline{d}_K^*$ since it corresponds to losing every bet and typically occurs with very low probability. For example, for a simple even-money payoff coin-flipping gamble with $N = 10$ and probability of heads

$$P(X(k) = 1) = p = 0.6,$$

the celebrated Kelly optimum $K = 2p-1 = 0.2$ obtained in papers such as [1] and [25], leads to $d_{K,worst} = 1-(1-0.2)^{10} \approx 0.89$ which corresponds to 89% drawdown as the worst-case.

5.2.3 Attainable Set for Drawdown-Modulated Feedback: Recalling the drawdown-modulated feedback $I(k) = \gamma M(k)V(k)$ in Section 4.3 with $\gamma \in \Gamma$ and $d_{\max} \in [0,1]$, in contrast to the theory in Chapter 4, we treat $d_{\max}$ as a design parameter. Thus, an admissible drawdown modulation controller is defined by a pair

$$\mathcal{M} \doteq (\gamma, d_{\max}) \in \Gamma \times [0,1].$$

Similar to the calculation for the linear feedback scheme, we now use with modulated feedback controller $I(k) = \gamma M(k)V(k)$ and obtain the recursion for the resulting account value

$$\begin{aligned} V(k+1) &= V(k) + I(k)X(k) \\ &= (1+\gamma M(k)X(k))V(k) \end{aligned}$$

which leads to terminal account value $V_{\mathcal{M}}(N)$, where the subscript $\mathcal{M}$ is used to emphasize the dependence on the design $\mathcal{M}$. Then, along any sample path $X = (X(0), \ldots, X(N-1))$, we arrive at the associated overall return

$$\begin{aligned} R_{\mathcal{M}} &\doteq \frac{V_{\mathcal{M}}(N) - V(0)}{V(0)} \\ &= \prod_{k=0}^{N-1} (1+\gamma M(k)X\,(k)) - 1 \end{aligned}$$

and the corresponding maximum percentage drawdown, call it $d_{\mathcal{M}}^*$. Now, proceeding exactly as in the linear feedback case, using the expected return

$$\overline{R}_{\mathcal{M}} \doteq \mathbb{E}[R_{\mathcal{M}}]$$

and expected maximum drawdown

$$\overline{d}_{\mathcal{M}}^* \doteq \mathbb{E}[d_{\mathcal{M}}^*],$$

we obtain the *attainable risk-return performance set* in the plane described by

$$\mathscr{R}_{\mathcal{M}} \doteq \left\{ \left(\overline{R}_{\mathcal{M}}, \overline{d}^*_{\mathcal{M}} \right) : \mathcal{M} \in \Gamma \times [0,1] \right\}.$$

5.3 Drawdown-Modulated Feedback Efficiency and Optimization

As previously stated at the beginning of the chapter, the simple lemma below drives the examples to follow. At the heart of its straightforward proof is the fact that the class of admissible LTI feedback controllers can be viewed as a subclass of the set of admissible drawdown-modulated feedbacks.

Lemma 5.3.1 (Efficiency): *For any admissible $K \in \mathcal{K}$, there exists a drawdown modulator pair $\mathcal{M} = (\gamma, d_{\max}) \in \Gamma \times [0,1]$ such that $\overline{R}_{\mathcal{M}} \geq \overline{R}_K$ and $\overline{d}^*_{\mathcal{M}} = \overline{d}^*_K$.*

Proof. Let $K \in \mathcal{K}$ be arbitrary. Now, taking the target level of drawdown $\widehat{d} \doteq \overline{d}^*_K$, we must show that there is an admissible pair $\mathcal{M} = (\gamma, d_{\max})$ which leads to $\overline{d}^*_{\mathcal{M}} = \widehat{d}$ and $\overline{R}_{\mathcal{M}} \geq \overline{R}_K$. Indeed, taking $\gamma = K$ and letting $d_{\max} = 1$, we first replicate the performance of linear feedback investment scheme; i.e., we obtain $\overline{d}^*_{\mathcal{M}} = \overline{d}^*_K$ and $\overline{R}_{\mathcal{M}} = \overline{R}_K$. Hence the maximization of $\overline{R}_{\mathcal{M}}$ over $\mathcal{M} \in \Gamma \times [0,1]$ with constraint $\overline{d}^*_{\mathcal{M}} = \overline{d}^*_K$ must be at least as large as $\overline{R}_K$. $\square$

5.3.2 Modulator Optimization: Motivated by the lemma above, the modulator can be optimized using the same idea as found in the celebrated Markowitz risk-return theory in finance; e.g., see [41,43,135,136]. That is, as described in Section 4.7.1, given any target level of expected drawdown, call it $\widehat{d} \in (0,1]$, we seek an admissible pair $\mathcal{M} = (\gamma, d_{\max}) \in \Gamma \times [0,1]$ maximizing $\overline{R}_{\mathcal{M}}$ subject to the constraint $\overline{d}^*_{\mathcal{M}} \leq \widehat{d}$. That is,

$$\sup_{\mathcal{M} \in \Gamma \times [0,1]} \overline{R}_{\mathcal{M}}$$

$$\text{subject to } \overline{d}_{\mathcal{M}} \leq \widehat{d}.$$

In our case, this is found by solving a two-dimensional optimization over the rectangle constraining γ and $d_{\max}$ above. The Efficiency Lemma 5.3.1 above suggests that the drawdown-modulated feedback leads to a "better" solution than that the one for linear feedback. As a bonus, as previously stated, we also see that the modulator assures a pre-specified worst-case level of drawdown protection for all sequences of returns.

5.4 Illustrative Examples

In the sequel, we begin our efficiency analysis with the simple case when $N = 2$ where calculations can be carried out in closed-form. Then we study the more general scenario with $N > 2$ where Monte-Carlo simulation is used. To make the comparison fair, in the examples to follow, we impose long-only and cash-financing constraint for both drawdown modulated feedback and LTI feedback. That is, for drawdown modulated feedback, to guarantee this condition is satisfied, the constraint on γ described in Section 4.3 is augmented to include $0 \leq \gamma M(k) \leq 1$. Similarly, for a LTI feedback, to guarantee the cash financing condition, we augment the constraint on K to include $0 \leq K \leq 1$. In the examples below, we actually show that the LTI feedback controller is actually *dominated* by the modulator; i.e., there exists a modulator which leads to the same level of risk with strictly higher return.

5.4.1 Demonstration of Domination for $N = 2$: We establish efficiency using a drawdown-modulated feedback strategy. For $N = 2$, we consider a single coin-flipping gamble having even-money payoff described by independent and identically distributed random variables $X(k) \in \{-1, 1\}$ and $P(X(k) = 1) \doteq p > 1/2$. Beginning with $\mu = \mathbb{E}[X(k)] = 2p - 1$, for the linear feedback betting strategy with fraction $0 \leq K \leq 1$, the associated expected return is readily calculated to be

$$\overline{R}_K = (1 + K\,(2p - 1))^2 - 1.$$

and the expected maximum percentage drawdown, found by a straightforward calculation is given by

$$\overline{d}_K^* = K(1-p)\,(2-K+Kp)\,.$$

For drawdown modulator pair $\mathcal{M} = (\gamma, d_{\max})$, a lengthy but straightforward computation leads to expected return and expected maximum percentage drawdown given by

$$\overline{R}_{\mathcal{M}} = \gamma d_{\max}(2p-1)(\gamma d_{\max} p + \gamma p - \gamma + 2)$$

and

$$\overline{d}_{\mathcal{M}}^* = \gamma d_{\max}\,(1-p)\,(2-\gamma+\gamma p)\,.$$

Lemma 5.4.2 (Domination for $N = 2$): *Consider the single coin-flipping gamble having even-money payoff with $N = 2$ as described above. Then for $0 < K < 1$, there exists a modulator $\mathcal{M} = (\gamma, d_{\max})$ such that $\overline{R}_K < \overline{R}_{\mathcal{M}}$ and $\overline{d}_{\mathcal{M}}^* = \overline{d}_K^*$.*

Proof. To prove this, it is sufficient to take $\gamma \doteq 1$ and

$$d_{\max} \doteq \frac{K\,(2-K+Kp)}{1+p}$$

which is obtained by setting $\overline{d}_{\mathcal{M}}^* = \overline{d}_K^*$ above. It is readily verified that $0 < d_{\max} < 1$ and by substitution of $d_{\max}$ and γ into $\overline{R}_{\mathcal{M}}$, after a lengthy but straightforward calculation, we obtain

$$\overline{R}_{\mathcal{M}} = \frac{K(2p-1)(2-K+Kp)f\,(K,p)}{(1+p)^2}$$

where

$$f(K,p) \doteq 2Kp - K^2p + K^2p^2 + p^2 + 2p + 1.$$

Now, to establish the desired domination, we now claim that $\overline{R}_{\mathcal{M}} > \overline{R}_K$. To prove this, we show that the difference between left and right hand sides above is positive. Indeed, via a lengthy but straightforward calculation, we obtain

$$\overline{R}_{\mathcal{M}} - \overline{R}_K = \frac{K^2(1-K)(1-p)p(2p-1)(3+p+Kp-K)}{(1+p)^2}.$$

Noting that $0 < K < 1$ and $p > 1/2$ above, it is immediate that both numerator and denominator for the difference described above are strictly positive. Thus, $\overline{R}_{\mathcal{M}} > \overline{R}_K$. $\square$

5.4.3 Remark: In Figure 5.1 we provide a plot which shows the degree of the strict domination in the difference based on our calculation of $\overline{R}_{\mathcal{M}} - \overline{R}_K$ above.

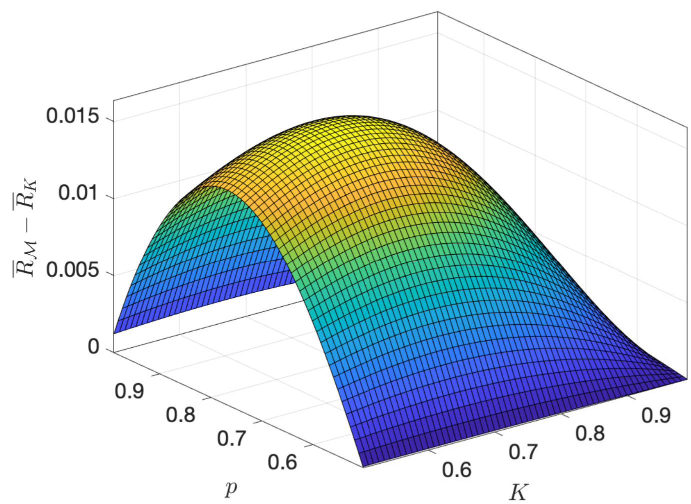


Figure 5.1: Degree of Strict Domination in Expected Return

5.4.4 Example of Domination with Larger N: Again, we consider a single coin-flipping scenario with even-money payoff described by independent and identically distributed random variables $X(k) \in \{-1/30, 1/30\}$ and $P(X(k) = 1/30) = 0.6$ with corresponding mean $\mu = \mathbb{E}[X(k)] = 1/150$. We choose $N = 252$ and view this as a trading problem for a binomial stock-price model over one year with daily returns varies around $\pm 3.3\%$ corresponding to

$X(k) = 1/30$ above. Note that this scenario is more biased on $X(k) = 1/30$. Hence, we study domination for the case when $0 \leq K \leq 1$.

To establish efficiency, for the linear feedback, we first obtain the expected return

$$\overline{R}_K = \left(1 + \frac{K}{150}\right)^{252} - 1.$$

As far as the expected maximum percentage drawdown $\overline{d}_K^*$ is concerned, this quantity is computed via performing a large number of Monte-Carlo simulations. Our finding is that for $0 \leq K \leq 1$, we have $\overline{d}_K^* \approx 0.25K$. For the drawdown-modulated feedback with the cash-financing condition imposed, we proceed as follows: As previously discussed earlier in this chapter, for a given target level of drawdown $\widehat{d} \in (0, 1]$, we seek to find a pair $\mathcal{M} = (\gamma^*, d_{\max})$ maximizing $\overline{R}_{\mathcal{M}}$ subject to $\overline{d}_{\mathcal{M}}^* = \widehat{d}$. This two-parameter optimization is solved with a large Monte-Carlo simulation. Then, letting $\overline{R}_{\mathcal{M}}^*(\widehat{d})$ denote the approximate optimal value obtained, we generate the dotted line in the Figure 5.2. We see that for any given risk level, the drawdown-modulated feedback leads to a certifiably higher expected return than the linear feedback investment scheme. In other words, the linear feedback investment scheme is inefficient as seen in the figure.

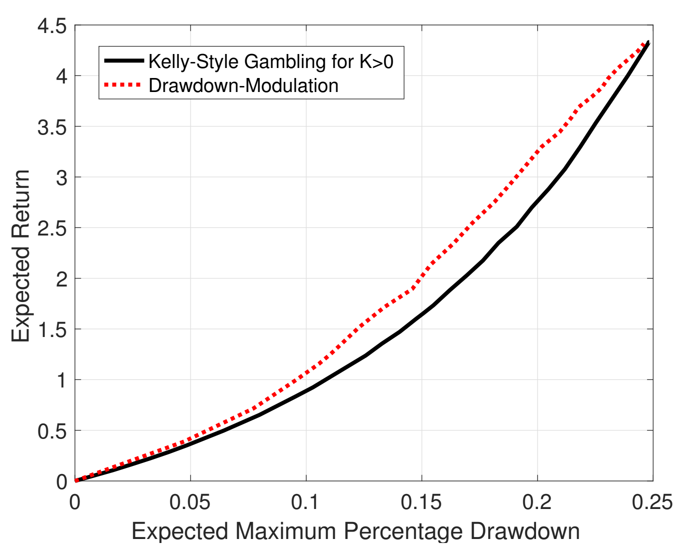


Figure 5.2: Risk-Reward Efficiency Plot for $N = 252$

5.5 Concluding Remarks and Future Work

In this chapter, using expected maximum percentage drawdown and mean return as the risk-reward pair, we demonstrated inefficiency of linear time-invariant feedback investment schemes. This was accomplished using our drawdown-modulated feedback controller, which is not only used to establish inefficiency but also provides a prescribed level of drawdown protection for all sequences of admissible returns. By way of extending the results given in this chapter, it is interesting to note that a drawdown-modulated controller can be used to obtain very similar efficiency results for other return metrics as well. For example, if $\overline{R}_K$ is replaced by the expected logarithmic growth $\mathbb{E}[\log(V(N)/V(0))]$, performance comparisons obtained are very similar to that given in Figure 5.2 result. When the returns $X(k)$ has dimension m which is large, finding the set of efficient points in the attainable set, namely, the so-called the *efficient frontier*, may typically involves high computational complexity. Another interesting extension of this work involves revisiting the Efficiency Lemma 5.3.1 in order to provide conditions under which strict inequality can be guaranteed. In this regard, the examples following the lemma suggest that this direction of research may be fruitful.

In next chapter, we consider a new problem formulation which extends the existing Kelly framework to include consideration of "betting frequency." Within this context, we focus on addressing the following questions: How does the optimal performance, expected logarithmic growth, depend on the betting frequency? Does the high-frequency case always lead to the best performance? Is there a condition under which the performance of the low frequency bettor can match that of high-frequency bettor?

Chapter 6

On Frequency-Based Kelly Betting and Stock Trading

In this chapter, using the expected logarithmic growth (ELG) criterion as the performance metric, we first formulate the problem of optimization with respect to the *betting frequency* in a dynamic game setting.[1] Then, in later sections, we extend our results so as to be relevant to the stock market. Along the lines of the preceding chapters, we begin with the analysis of a coin-flipping game defined by a sequence of bets with independent and identically distributed (i.i.d.) returns $X(k)$. Letting $V(k)$ denote the account value at stage k, our formulation begins at $k = 0$ with a wager $KV(0)$ with $K \geq 0$. Then the bettor waits $n \geq 1$ steps before updating the bet size. After each play, the "house" takes its share and the balance of the money is left to "ride" with resulting profits or losses viewed as "unrealized" until stage n is reached. When n is small, this is viewed as the high-frequency case, and when n is large, one might use the term "bet and hold" in anticipation of the stock market analysis to follow. As seen later in this chapter, when suitably large transaction costs are imposed, consistent with common sense, a low-frequency bettor may perform better than a high-frequency bettor.

After the motivating coin-flipping example, in Section 6.2, the frequency-based formulation is given for a general class of probability distributions for $X(k)$. Our goal is to study the optimal performance as a function of frequency; equivalently as a function of the waiting period n. Within this context, we answer the following questions: How does the optimal performance, we call it g_n^*, change with waiting period n? Does the high-frequency case, $n = 1$, always lead to the best performance? Is there a condition under which the performance obtained with

[1] Most of the research reported in this chapter has been published in [63] and [64].

low-frequency betting matches that of high-frequency betting? To address these questions, in Section 6.4, we provide a result, which we call the *High-Frequency Maximality Theorem*. This theorem tells us that the high-frequency bettor is *unbeatable*; i.e., $g_1^* \geq g_n^*$ for $n \geq 1$ and leads to the following interesting question: Under what conditions does it follow that $g_n^* = g_1^*$? Roughly speaking, if such a result is true, it suggests that high-frequency betting is a "waste of time." Further to this, in Sections 6.5 and 6.6, we work with the inequality

$$\mathbb{E}\left[\frac{1}{1+X(0)}\right] \leq 1,$$

which is called the *sufficient attractiveness* condition. As it name suggests, it is interpreted to be an indicator of the "attractiveness" of the bet. We subsequently prove that satisfaction of this condition guarantees $g_n^* = g_1^*$ for all $n \geq 1$.

6.0.1 From Betting Games to Stock Trading: Beginning in Section 6.7, when transitioning from betting games to stock trading, we make the following terminology changes: The *bettor* is replaced by a *trader*, the *betting frequency* is replaced by the *rebalancing frequency* and the bettor's *wager* $KV(k)$ becomes trader's *investment* $I(k)$. Another change is that we allow the feedback gain K to take either sign, where $K > 0$ means the trader is going long and $K < 0$ means the trader is going short. When no update on the trade occurs until stage $k = n$, we use the term "buy and hold." When $K < 0$ is used and no update occurs until $k = n$, loosely speaking, this can be called "short and hold." For $K < 0$, as explained in Section 6.7, the possibility of bankruptcy arises.

When dealing with stocks, similar to the betting scenario, the returns $X(k)$ are assumed to be i.i.d. with rather arbitrary distribution, and we study trading performance g_n^* as a function of frequency; equivalently as a function of the waiting period n. First, some preliminary results regarding optimality and bankruptcy issues are provided and then, our results, the High-Frequency Maximality Theorem and the Sufficient Attractiveness Theorem obtained for betting, are modified so as to apply to the stock market. In the final sections of the chapter, working in a control-theoretic framework, we extend the frequency-based framework to the case of a multi-asset portfolio with the restriction that trades are long-only. In this context, we generalize the Sufficient Attractiveness Theorem and show that if there is an asset satisfying a certain *dominance* condition, then an optimal portfolio consists of this asset alone; i.e., the trader puts "all eggs in one basket," and it is seen that performance becomes a constant function of rebalancing frequency.

6.1 Frequency-Based Formulation for Coin-Flipping Games

In this section, we use a simple even-money payoff coin-flipping game as an entry point for the more general analysis to follow. Indeed, beginning with $k = 0$, we let $X(k) \in \{-1, 1\}$ be a sequence of i.i.d. Bernoulli random variables with probability

$$p \doteq P(X(k) = 1)$$

where $X(k) = 1$ represents the return on a "win" and $X(k) = -1$ represents the return on a "loss." We assume $p \geq 1/2$ and, without loss of generality, this is the probability of heads. Therefore, in the analysis below, the constant $0 \leq K \leq 1$ represents the "betting fraction" on heads and we set up a frequency-dependent optimization in this variable. Now, for each integer $n \geq 1$ denoting the number of coin flips between bets, it is convenient to work with the random variable

$$\mathcal{X}_n \doteq \prod_{k=0}^{n-1}(1 + X(k)) - 1,$$

which corresponds to the *compound return*. For this simple scenario it is immediate to see that $\mathcal{X}_n = 2^n - 1$ with probability p^n and $\mathcal{X}_n = -1$ with probability $1 - p^n$.

6.1.1 Frequency-Based Kelly Betting: Now, for $k = 0, 1, \ldots, n-1, n$, letting $V(k)$ be the bettor's account value at stage k with initial account value $V(0) > 0$, we assume that each wager can be no more than the account value $V(k)$; i.e., this is the cash-financed situation which requires $K \in \mathcal{K} \doteq [0, 1]$ in the optimization to follow. With the initial wager $KV(0)$, the terminal account value at stage n is given by

$$V(n) = (1 + K\mathcal{X}_n)V(0)$$

and $K \in \mathcal{K}$ assures $V(k) \geq 0$ for all $k \geq 0$ along every sample path of returns

$$X \doteq (X(0), X(1), \ldots, X(n-1)).$$

To emphasize the dependence of $V(k)$ on the feedback gain K, we sometimes write $V(k, K)$ instead. Now, we seek $K \in \mathcal{K}$ maximizing

$$\begin{aligned} g_n(K) &\doteq \frac{1}{n}\mathbb{E}\left[\log \frac{V(n,K)}{V(0)}\right] \\ &= \frac{1}{n}\mathbb{E}[\log(1 + K\mathcal{X}_n)] \\ &= \frac{1}{n}\left[p^n \log(1 + (2^n - 1)K) + (1 - p^n)\log(1 - K)\right]. \end{aligned}$$

We note that the factor $1/n$ is used so that the expected logarithmic growth is on a per-period basis. The associated optimal expected logarithmic growth is now obtained as

$$g_n^* \doteq \max_{K \in \mathcal{K}} g_n(K)$$

and any $K_n^* \in \mathcal{K}$ satisfying $g_n(K_n^*) = g_n^*$ is called an *optimal Kelly fraction*.

In the theorem below, we note that the special case $n = 1$ corresponds to the classical Kelly solution, which is given by $K_1^* = 2p - 1$ with associated optimal expected logarithmic growth

$$\begin{aligned} g_1^* &= p\log(1 + K_1^*) + (1 - p)\log(1 - K_1^*) \\ &= p\log(2p) + (1 - p)\log(2 - 2p). \end{aligned}$$

Theorem 6.1.2 (Kelly Optimum for Frequency-Based Betting): *For $n \geq 1$, the coin-flipping game with $p \geq 1/2$ and even-money payoff described above has optimal Kelly fraction*

$$K_n^* = \frac{2^n p^n - 1}{2^n - 1}$$

and associated optimal expected logarithmic growth is given by

$$g_n^* = p^n \log p + \left(\frac{1 - p^n}{n}\right)\log\left[\frac{1 - p^n}{2^n - 1}\right] + \log 2.$$

Proof. We first observe that $\mathcal{X}_n = 2^n - 1$ with probability p^n and $\mathcal{X}_n = -1$ with probability $1 - p^n$. Hence, the expected logarithmic growth is

$$\begin{aligned} g_n(K) &= \frac{1}{n}\mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right] \\ &= \frac{1}{n}\left[p^n \log(1 + (2^n - 1)K) + (1 - p^n)\log(1 - K)\right] \end{aligned}$$

and setting its derivative

$$\frac{\partial g_n}{\partial K} = \frac{1}{n}\left[\frac{p^n(2^n - 1)}{1 + (2^n - 1)K} - \frac{1 - p^n}{1 - K}\right]$$

to zero, we obtain a unique candidate for the maximizer given by

$$K_n^* \doteq \frac{2^n p^n - 1}{2^n - 1}.$$

We first observe that $p \geq 1/2$ implies that K_n^* is feasible; i.e., $K_n^* \in \mathcal{K} = [0, 1]$ for all $n \geq 1$. Now using the fact that $g_n(K)$ is concave, in combination with uniqueness of the zero-derivative point above, it follows that K_n^* is *the* global maximizer; e.g., see [137]. Finally, to complete our analysis, we substitute K_n^* into $g_n(K)$. A lengthy but straightforward calculation leads to the formula given for g_n^*. $\square$

6.1.3 Illustrative Example: To demonstrate the frequency dependence of the performance, in Figure 6.1, for various p values, we plot g_n^* as a function of n. Since g_n^* is strictly decreasing in n, it suggests that betting "faster is better."

6.1.4 Transaction Cost Considerations: We now extend the problem formulation to include consideration of transaction costs. Indeed, we consider the case when a percentage charge $0 < \varepsilon < 1$ is assessed by the "house" each time the bet is updated. If this occurs at stage k, the charge is $\varepsilon KV(k)$.[2] Common sense tells us that as ε increases, we should

[2] There are many other models for transaction costs. One possibility is to simply consider a fixed charge at each step. Another is at stage k, to assess a charge only on *new money* being brought into play. In the case of stocks, at stage k, this corresponds to commission only on newly acquired or shares being sold.

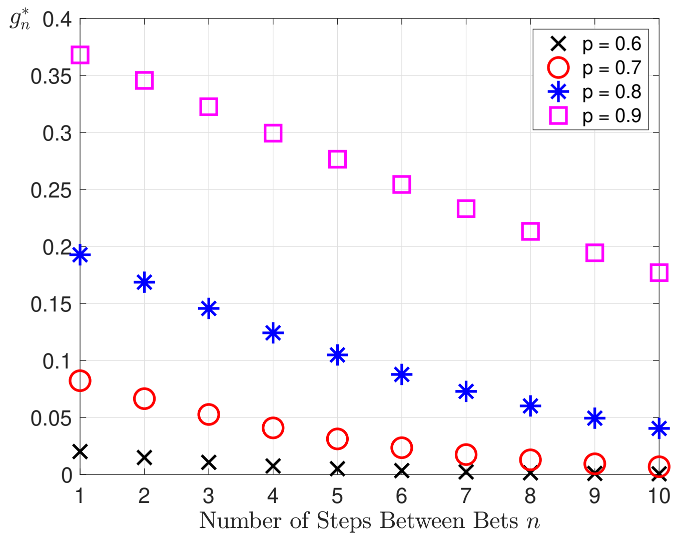


Figure 6.1: Optimal Expected Logarithmic Growth Versus n for $p > 1/2$

see that the benefits of high-frequency betting are reduced. Now, with $V(0) > 0$, via a straightforward calculation, we first obtain the account value at stage n given by

$$V(n, K) = (1 + K\mathcal{X}_n - K\varepsilon)V(0).$$

Note that for $K > 1/(1+\varepsilon)$, the situation $V(k) < 0$ can arise, which is taken to mean infinitely negative expected logarithmic growth. Hence, without loss of generality, in the analysis to follow, we work with the subset of $\mathcal{K}$ where the survival issue does not arise; i.e.,

$$K \in \mathcal{K}_\varepsilon \doteq \left[0, \frac{1}{1+\varepsilon}\right]$$

and we note that any K value violating the constraint above cannot possibly be optimal. Thus, our goal is to find $K_n^* \in \mathcal{K}_\varepsilon$ maximizing the modified expected logarithmic growth given by

$$g_{n,\varepsilon}(K) = \frac{1}{n}\mathbb{E}\left[\log\left(1 + K\mathcal{X}_n - K\varepsilon\right)\right].$$

To illustrate how transaction costs reduce the benefits of high-frequency betting, we revisit the even-money coin-flipping game considered in the previous subsection. We now generalize Theorem 6.1.2 to include these costs.

Theorem 6.1.5 (Kelly Optimum with Transaction Costs): *Given $n \geq 1$, consider the coin-flipping game described above with $p > 1/2$, even-money payoff and transaction cost $0 < \varepsilon < 1$. Then, if*

$$p \geq p_{n,\varepsilon} \doteq \left(\frac{1+\varepsilon}{2^n}\right)^{1/n},$$

the optimal Kelly fraction is given by

$$K^*_{n,\varepsilon} = \frac{2^n p^n - \varepsilon - 1}{(1+\varepsilon)(2^n - \varepsilon - 1)}$$

with the associated expected logarithmic growth given by

$$g^*_{n,\varepsilon} = \frac{p^n}{n} \log\left(\frac{2^n p^n}{1+\varepsilon}\right) + \left(\frac{1-p^n}{n}\right) \log\left(\frac{2^n(1-p^n)}{2^n - \varepsilon - 1}\right).$$

Alternatively, if $p < p_{n,\varepsilon}$, then $K^*_{n,\epsilon} = 0$ and $g^*_{n,\varepsilon} = 0$.

Proof. We first note that the expected logarithmic growth is given by

$$\begin{aligned} g_{n,\varepsilon}(K) &= \frac{1}{n}\mathbb{E}\left[\,\log\left(1 + K\mathcal{X}_n - K\varepsilon\,\right)\right] \\ &= \frac{1}{n}\left[p^n \log\left(1 + K\left(2^n - 1\right) - K\varepsilon\right) + (1-p^n)\log\left(1 - K - K\varepsilon\right)\right] \end{aligned}$$

and that its first derivative is calculated to be

$$\frac{\partial g_{n,\varepsilon}}{\partial K} = \frac{1}{n}\left[p^n \frac{2^n - 1 - \varepsilon}{1 + K(2^n - 1) - K\varepsilon} + (1-p^n)\frac{-1-\varepsilon}{1 - K - K\varepsilon}\right].$$

Noting that this derivative vanishes if and only if

$$K = K^*_{n,\varepsilon} \doteq \frac{2^n p^n - \varepsilon - 1}{(1+\varepsilon)(2^n - \varepsilon - 1)}$$

and observing that $2^n - \varepsilon - 1 > 0$ and $p > 1/2$, it is straightforward to see that $K^*_{n,\varepsilon} \in \mathcal{K}_\varepsilon$. Now, in combination with the fact that $g_{n,\varepsilon}(K)$ is concave, arguing as in the proof of Theorem 6.1.2, it follows that $K^*_{n,\varepsilon}$ is the global maximizer for $p \geq p_{n,\varepsilon}$. Therefore, substituting $K^*_{n,\varepsilon}$ into $g_{n,\varepsilon}(K)$, a straightforward calculation leads to $g^*_{n,\varepsilon}$ as required.

To complete the proof, it remains to show that $p \leq p_{n,\varepsilon}$ implies $K^*_{n,\epsilon} = 0$. To see this, it suffices to show that for any $K > 0$, $g_{n,\varepsilon}(K) \leq g_{n,\varepsilon}(0)$. Indeed, for $K > 0$, by Jensen's inequality and the fact that $\mathbb{E}[\mathcal{X}_n] = (2p)^n - 1$, we obtain upper bound

$$\begin{aligned} g_{n,\varepsilon}(K) &\leq \frac{1}{n}\log(1 + K\mathbb{E}[\mathcal{X}_n] - K\varepsilon]) \\ &= \frac{1}{n}\log(1 + K[(2p)^n - 1 - \varepsilon]) \end{aligned}$$

Now observe that if $p < p_{n,\varepsilon}$, it implies that the quantity $(2p)^n - 1 - \varepsilon < 0$. In combination with the assumption that $K > 0$, it follows that the $g_{n,\varepsilon}(K) < 0 = g_{n,\varepsilon}(0)$. $\square$

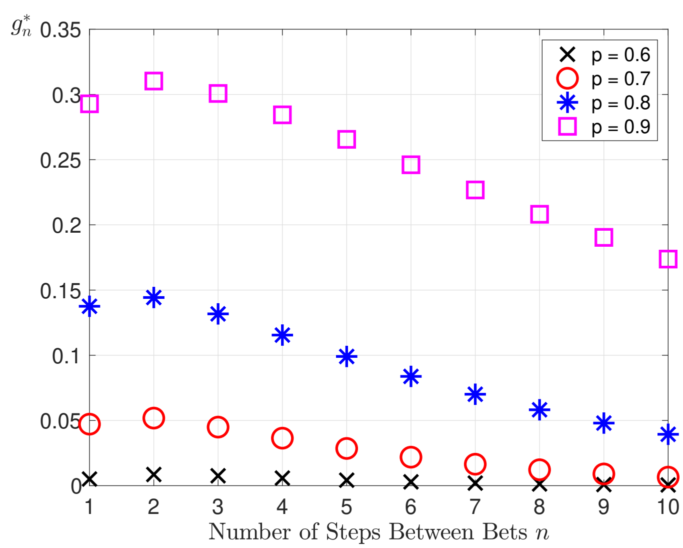


Figure 6.2: Transaction Cost Effects on g_n^* for $p > 1/2$

6.1.6 Remarks: Since $\varepsilon \to 0$ implies $p_{n,\varepsilon}$ converges to $1/2$, it is straightforward to see that the zero-transaction cost result in Theorem 6.1.2 is a special case; i.e.,

$$\lim_{\varepsilon \to 0} K^*_{n,\varepsilon} = \frac{2^n p^n - 1}{2^n - 1} = K^*_n.$$

To illustrate use of the results above, for $p = 0.6, 0.7, 0.8$ and 0.9 and $\varepsilon = 0.1$, we plot the optimal expected logarithmic growth g_n^* versus n in Figure 6.2. From the figure, we see that the optimal waiting period is $n^* = 2$. In other words, the optimum is achieved by waiting two time units rather than betting at each step. Said another way, in the presence of transaction costs, betting at the highest possible frequency, namely using $n = 1$, is not necessarily optimal; i.e., in some cases we obtain $g_2^* > g_1^*$. In the absence of transaction costs, in the sequel, we typically see $g_1^* > g_n^*$ and in some cases, as described in the first section of this chapter, when the sufficient attractiveness condition is satisfied, we obtain $g_1^* = g_n^*$.

6.2 More General Betting Games

In this section, we extend our frequency-based formulation to allow for rather general distributions for sequence of i.i.d. random variables $X(k)$ representing the *returns* specified by the "house" and assume, as in previous chapters, that $X_{\min} \leq X(k) \leq X_{\max}$ with $X_{\min}$ and $X_{\max}$ satisfying $-1 < X_{\min} < 0 < X_{\max} < \infty$. If the bettor likes these returns offered by house and decides to bet, then we interpret $X(k) > 0$ to be the return on a "win" and $X(k) < 0$ for a "loss." In the sequel, we say that this bettor *takes the* $X > 0$ *side*. We also analyze the case when the house allows the bettor to take the opposite side of the bet. In this case, we say that the bettor *takes* $X < 0$ *side*. By this we mean $-X(k)$ represents the return on a "win" and $X(k)$ represents the return on a "loss." For this case, it is easy to show that by replacing $\widetilde{X}(k) \doteq -X(k)$, we obtain an equivalent game with the bettor on the $\widetilde{X} > 0$ side and new bounds obtained from original bounds as

$$\widetilde{X}_{\min} \doteq -X_{\max};$$
$$\widetilde{X}_{\max} \doteq -X_{\min}.$$

Considering the above, without loss of generality, we henceforth assume the bettor takes the $X > 0$ side with the following caveat: Whereas, the original problem has lower bounds $X_{\min} > -1$ corresponding to a 100% loss, the new problem might have $\widetilde{X}_{\min} < -1$. In other words, a loss of more than 100% is possible. This introduces technical complications if one tries to extend the analysis earlier in the chapter which relies on the compound returns $\mathcal{X}_n$.

This is because the possibility of "sign switching" of the compound returns confounds the mathematical analysis which was used for betting on the $X > 0$ side. In view of these considerations, we henceforth assume $X_{\max} < 1$ for the case of betting on the $X < 0$ side; see suggestions for future research described at the end of this chapter.

6.2.1 Frequency-Based Kelly Betting: Following the progression in Section 6.1.1, we now apply the frequency-based formulation to a betting game with returns having general distributions. For each integer $n \geq 1$, we again use the compound returns $\mathcal{X}_n$ as defined in Section 6.1 whose bounds are readily seen to be

$$\mathcal{X}_n \leq (1 + X_{\max})^n - 1 \doteq \mathcal{X}_{n,\max};$$
$$\mathcal{X}_n \geq (1 + X_{\min})^n - 1 \doteq \mathcal{X}_{n,\min}$$

with $-1 < \mathcal{X}_{n,\min} < 0 < \mathcal{X}_{n,\max} < \infty$. Now, the bettor who takes the $X > 0$ side and wagers $KV(0)$ with $K \in \mathcal{K} = [0, 1]$ at stage $k = 0$ waits n steps before updating the bet size and arrives at account value $V(n, K) = (1 + K\mathcal{X}_n)V(0)$. Then we seek $K \in \mathcal{K}$ maximizing the expected logarithmic growth

$$\begin{aligned} g_n(K) &\doteq \frac{1}{n}\mathbb{E}\left[\log \frac{V(n,K)}{V(0)}\right] \\ &= \frac{1}{n}\mathbb{E}[\log(1 + K\mathcal{X}_n)]. \end{aligned}$$

6.3 On Which Side Should the Bettor Bet?

This far, we have not addressed the following question: Should the bettor take the $X > 0$ side or the $X < 0$ side? For the special case of even-money coin-flipping games with a biased coin, suppose a bet on heads offered by the house corresponds to the $X > 0$ side. Then common sense dictates that the bettor should take this side if $p > 1/2$. Otherwise, the $X < 0$ side corresponding to tails is better. More generally, as seen in the lemma below, betting on which side is determined by the sign of $\mathbb{E}[X(0)]$.

Lemma 6.3.1 (Preliminaries): *For the frequency-based betting scenario described in Section 6.2, if $\mathbb{E}[X(0)] > 0$, it is optimal to bet on the $X > 0$ side with $K_n^* > 0$. If $\mathbb{E}[X(0)] < 0$, it is optimal to bet on the $X < 0$ side with $K_n^* > 0$. If $\mathbb{E}[X(0)] = 0$, the optimal Kelly fraction $K_n^* = 0$.*

Proof. We only give a proof for the two cases covered by $\mathbb{E}[X(0)] \geq 0$ since a nearly identical proof works for $\mathbb{E}[X(0)] < 0$. Indeed, to address the case $\mathbb{E}[X(0)] > 0$, we proceed by contradiction. Suppose that it is optimal to bet on the $X < 0$ side with $K_n^* > 0$. Note that in this case, $X_{\max} < 1$. Applying Jensen's inequality, for $n \geq 1$, the expected logarithmic growth has upper bound

$$g_n(K_n^*) = \frac{1}{n}\mathbb{E}\left[\log\left(1 + K_n^*\widetilde{\mathcal{X}}_n\right)\right] \leq \frac{1}{n}\log(1 + K_n^*\mathbb{E}[\widetilde{\mathcal{X}}_n])$$

where

$$\widetilde{\mathcal{X}}_n = \prod_{k=0}^{n-1}(1 - X(k)) - 1.$$

Since returns $X(k)$ are i.i.d. with $X_{\max} < 1$ and $\mathbb{E}[X(0)] > 0$, we have $0 < \mathbb{E}[X(0)] < 1$. Thus, it follows that

$$\mathbb{E}[\widetilde{\mathcal{X}}_n] = (1 - \mathbb{E}[X(0)])^n - 1 < 0.$$

In combination with the fact that $K_n^* > 0$, this implies that $g_n(K_n^*) < 0 = g_n(0)$, which contradicts the optimality of K_n^*. Hence, the bettor must bet on the $X > 0$ side to be optimal.

We now show that $K_n^* > 0$. To see this, it suffices to show that $g_n(K) > 0$ for some $K > 0$ in some neighborhood of zero. Taking the derivative

$$\frac{d}{dK}g_n(K) = \frac{1}{n}\frac{d}{dK}\mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right]$$

and noting that $K \in [0,1]$, $\mathcal{X}_n$ is bounded and $1 + K\mathcal{X}_n > 0$, results in probability theory, for example, see [138], allow us to commute the differentiation and expectation operators above. Thus, if we evaluate the derivative at $K = 0$, it follows that

$$\left.\frac{d}{dK}g_n(K)\right|_{K=0} = \frac{1}{n}\mathbb{E}\left[\frac{\mathcal{X}_n}{1+K\mathcal{X}_n}\right]\Bigg|_{K=0} = \frac{1}{n}\mathbb{E}[\mathcal{X}_n] > 0$$

where the last inequality holds since $\mathbb{E}[X(0)] > 0$. Therefore, we obtain

$$\left.\frac{d}{dK}g_n(K)\right|_{K=0} > 0.$$

Since $g_n(0) = 0$, it follows that there exists some $K > 0$ suitably small so that $g_n(K) > 0$, which implies that there is an optimum $K_n^* > 0$.

To address the case $\mathbb{E}[X(0)] = 0$, we first note that this implies $\mathbb{E}[\mathcal{X}_n] = 0$. Now, using Jensen's inequality, for all admissible K, we obtain,

$$g_n(K) \leq \frac{1}{n}[\log(1 + K\mathbb{E}[\mathcal{X}_n])] \leq 0.$$

Now note that if $K = 0$, the upper bound is attained since $g_n(0) = 0$. Hence, $K_n^* = 0$ and the proof is complete. $\square$

6.4 The High-Frequency Maximality Theorem

In this section we address the following question: Does the high-frequency case, $n = 1$, always lead to the best expected logarithmic growth? To this end, we prove a result called the *High-Frequency Maximality Theorem*, which tells us that $g_n^* \leq g_1^*$ for all $n \geq 1$. Said another way, high-frequency betting is *unbeatable*. As mentioned in Section 6.2, while the theorem below is proven by assuming that the bettor takes the $X > 0$ side, the same result holds true for the $X < 0$ side provided $X_{\max} < 1$.

Theorem 6.4.1 (High-Frequency Maximality): *For the frequency-based betting scenario described in Section 6.2, for all $n \geq 1$, it follows that $g_n^* \leq g_1^*$.*

Proof. We begin by noting that when $n = 1$, the statement holds trivially. Hence, in the sequel, we assume $n > 1$. Using the shorthand X_k for $X(k)$, for $K \in [0, 1]$, we recall from Section 6.2 that the account value of the bettor is given by $V(n, K) = (1 + K\mathcal{X}_n)\, V(0)$. To show $g_n^* \leq g_1^*$, we use the smoothing property of conditional expectation to write the expected logarithmic growth as

$$\begin{aligned}
g_n(K) &= \frac{1}{n}\mathbb{E}\left[\log \frac{V(n,K)}{V(0)}\right] \\
&= \frac{1}{n}\mathbb{E}[\log(1 + K\mathcal{X}_n)] \\
&= \frac{1}{n}\mathbb{E}\left[\,\mathbb{E}[\log(1 + K\mathcal{X}_n)|X_{n-1}]\,\right] \\
&= \frac{1}{n}\mathbb{E}\left[\,\mathbb{E}\left[\log \frac{1 + K\mathcal{X}_n}{1 + KX_{n-1}}\middle| X_{n-1}\right]\right] + \frac{1}{n}\mathbb{E}\left[\,\mathbb{E}\left[\log(1 + KX_{n-1})|X_{n-1}\right]\right] \\
&= \frac{1}{n}\mathbb{E}\left[\,\mathbb{E}\left[\log \frac{1 + K\mathcal{X}_n}{1 + KX_{n-1}}\middle| X_{n-1}\right]\right] + \frac{1}{n}\mathbb{E}\left[\log(1 + KX_{n-1})\right] \\
&= \frac{1}{n}\mathbb{E}\left[\,\mathbb{E}\left[\log \frac{1 + K\mathcal{X}_n}{1 + KX_{n-1}}\middle| X_{n-1}\right]\right] + \frac{1}{n}g_1(K)
\end{aligned}$$

where the last step follows from the fact that the X_k are i.i.d. To simplify the inner conditional expectation above, we use the independence of $\mathcal{X}_{n-1}$ and X_{n-1} and the fact that

$$\mathcal{X}_n - X_{n-1} = (1 + X_{n-1})\mathcal{X}_{n-1}.$$

Conditioning on $X_{n-1} = x$ for $x > -1$, we write

$$\begin{aligned}
\mathbb{E}\left[\log \frac{1 + K\mathcal{X}_n}{1 + KX_{n-1}}\middle| X_{n-1} = x\right] &= \mathbb{E}\left[\log\left(1 + \frac{K(\mathcal{X}_n - X_{n-1})}{1 + KX_{n-1}}\right)\middle| X_{n-1} = x\right] \\
&= \mathbb{E}\left[\log\left(1 + \frac{K(1 + X_{n-1})}{1 + KX_{n-1}}\mathcal{X}_{n-1}\right)\middle| X_{n-1} = x\right] \\
&= \mathbb{E}\left[\log\left(1 + \frac{K(1 + x)}{1 + Kx}\mathcal{X}_{n-1}\right)\middle| X_{n-1} = x\right] \\
&= \mathbb{E}\left[\log\left(1 + \frac{K(1 + x)}{1 + Kx}\mathcal{X}_{n-1}\right)\right] \\
&= (n-1)g_{n-1}\left(K_x\right)
\end{aligned}$$

where

$$K_x \doteq \frac{K(1+x)}{1+Kx}.$$

Since $K \in [0,1]$ and $x > -1$, it follows that $K_x \in [0,1]$. Hence,

$$g_{n-1}(K_x) \leq \max_{K \in [0,1]} g_{n-1}(K).$$

Since the right-hand side is equal to g_{n-1}^*, it follows that

$$\mathbb{E}\left[\log \frac{1+K\mathcal{X}_n}{1+KX_{n-1}} \middle| X_{n-1} = x\right] \leq (n-1)g_{n-1}^*.$$

We now have

$$g_n(K) \leq \frac{n-1}{n} g_{n-1}^* + \frac{1}{n} g_1(K).$$

Taking the supremum over $K \in [0,1]$ leads to

$$g_n^* \leq \frac{n-1}{n} g_{n-1}^* + \frac{1}{n} g_1^*,$$

and, to complete the proof, it is noted that the foregoing argument for g_n^* also applies to any g_m^* for $m > 1$. Hence,

$$g_m^* \leq \frac{m-1}{m} g_{m-1}^* + \frac{1}{m} g_1^*.$$

Now with $m = 2$, we have $g_2^* \leq g_1^*$. Similarly, for $m = 3$, it follows that

$$g_3^* \leq \frac{2}{3} g_2^* + \frac{1}{3} g_1^* \leq g_1^*,$$

Continuing in this way we arrive at $g_n^* \leq g_1^*$. $\square$

6.5 Notion of Sufficient Attractiveness

The High-Frequency Maximality Theorem above leads to the following question: Under what conditions might low-frequency betting achieve the same performance as high-frequency betting? That is, we already know that $g_n^* \leq g_1^*$ for $n \geq 1$ but might there be important cases when equality occurs? Loosely speaking, under what conditions is high-frequency betting a waste of time? To address this, below, we introduce a notion which we call the *sufficient attractiveness* condition. When this condition holds, we prove that $g_n^* = g_1^*$ for all $n \geq 1$.

6.5.1 Sufficient Attractiveness: With returns $X(k)$ being i.i.d. random variables satisfying the bounds given in Section 6.2, we say that the bet on the $X > 0$ side is *sufficiently attractive* if

$$\mathbb{E}\left[\frac{1}{1+X(0)}\right] \leq 1.$$

For the case when the bettor takes the $X < 0$ side and $X_{\max} < 1$, the corresponding sufficient attractiveness inequality is

$$\mathbb{E}\left[\frac{1}{1-X(0)}\right] \leq 1.$$

In the sequel, unless otherwise stated, the standing assumption is that we are referring to the $X > 0$ side of the bet and the first inequality is relevant.

6.5.2 Implications of the Sufficient Attractiveness Inequality: To provide some insight as to the meaning of the sufficient attractiveness inequality above, we first consider the extreme case when $X(0) > 0$ with probability one. That is, an arbitrage is available, and the sufficient attractiveness inequality is trivially satisfied. More generally, when $X(k)$ are i.i.d. returns and the sufficient attractiveness inequality holds, we claim that $\mathbb{E}[X(0)] \geq 0$. To establish this, we begin by noting that the function $f(z) \doteq 1/(1+z)$ is convex for $z > -1$. Applying Jensen's inequality leads to

$$\frac{1}{1+\mathbb{E}[X(0)]} \leq \mathbb{E}\left[\frac{1}{1+X(0)}\right],$$

which implies that if $\mathbb{E}[1/(1+X(0))] \leq 1$, we have $\mathbb{E}[X(0)] \geq 0$. More generally, the quantity

$$\theta \doteq \mathbb{E}\left[\frac{1}{1+X(0)}\right]$$

can be viewed as the "degree of favorability" of the game in the following sense: For $0 < \theta \leq 1$, Jensen's inequality leads to

$$\mathbb{E}[X(0)] \geq \frac{1}{\theta} - 1,$$

Thus, smallness of $\mathbb{E}[1/(1+X(0))]$ is desirable. Below we provide two examples to further illustrate the notion of sufficient attractiveness.

6.5.3 Example (Sufficient Attractiveness for a Bernoulli Distribution): For i.i.d. returns $X(k) = \gamma > 0$ with probability p and $X(k) = -\gamma$ with probability $1-p$, it is easy to verify that the sufficient attractiveness inequality is satisfied if and only if

$$p \geq \frac{1+\gamma}{2}.$$

For example, for $\gamma = 1/2$, we require $p \geq 0.75$. Note that the extreme case $\gamma = 1$ corresponds to $p = 1$ and it follows that all g_n^* are equal. More generally, for binary lattice problems with $X(k) \in \{X_{\min}, X_{\max}\}$ and $P(X(k) = X_{\max}) = p$, the sufficient attractiveness inequality reduces to

$$p \geq \frac{|X_{\min}|\,(1 + X_{\max})}{X_{\max} - X_{\min}}.$$

This tells us that the smallest value of the probability p should be higher than the threshold above so that sufficient attractiveness is assured.

6.5.4 Example (Sufficient Attractiveness for the Uniform Distribution): As a second example, we consider the case of a uniform distribution. Suppose the i.i.d. returns $X(k)$ are governed by the uniform distribution on $[a, b]$ with $-1 < a < 0$ and $b > 0$. To study the sufficient attractiveness inequality, we calculate

$$\mathbb{E}\left[\frac{1}{1+X(0)}\right] = \frac{1}{b-a}\log\left(\frac{1+b}{1+a}\right).$$

Now, for a in its allowed range, we define

$$b_{\min}(a) \doteq \min\left\{b > 0 : \mathbb{E}\left[\frac{1}{1+X\,(0)}\right] \leq 1\right\}.$$

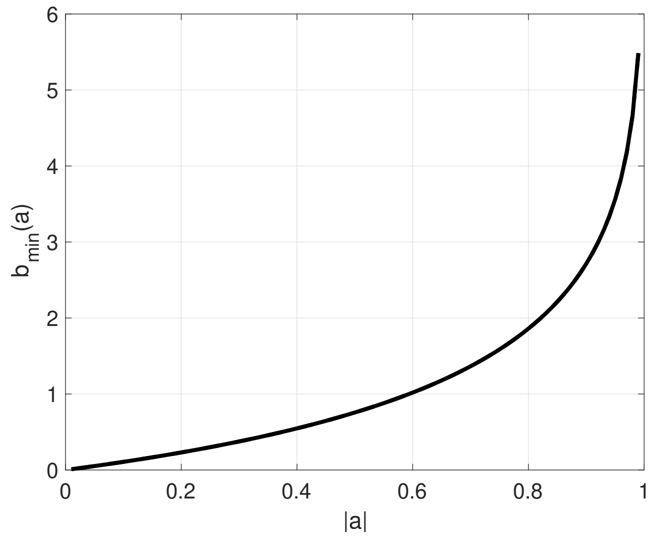


Figure 6.3: Sufficient Attractiveness: $b_{\min}(a)$ Versus $|a|$

Note that the sufficient attractiveness inequality for the uniform distribution case is easily seen to be equivalent to

$$\frac{e^a}{1+a} \leq \frac{e^b}{1+b}.$$

Recalling that for $-1 < a < 0 < b$, conditions on the pair (a, b) under which this inequality holds, it is straightforward to verify that the desired sufficiently attractive inequality holds if and only if

$$b \geq b_{\min}(a),$$

where $b_{\min}(a) > 0$ is given in terms of the Lambert function; see [139]. Namely,

$$b_{\min}(a) = -1 - W_{_{-1}}\left(-\left(1+a\right)e^{-(1+a)}\right)$$

where $W_{_{-1}}(\cdot)$ is the Lambert function with lower branch satisfying $W_{_{-1}}(\cdot) \leq -1$. Figure 6.3 showing $b_{\min}(a)$ versus $|a|$ for $a \in (-1, 0)$ indicates that as a decreases, the b needed to sustain sufficient attractiveness increases monotonically.

6.6 The Sufficient Attractiveness Theorem

The theorem below tells us that when a bet is sufficiently attractive, the "bet-and-hold" strategy achieves the *same* performance as the high-frequency strategy. That is, $g_n^* = g_1^*$ for all $n \geq 1$, and, as previously stated, high-frequency betting can be viewed as a waste of time.

Theorem 6.6.1 (Sufficient Attractiveness): *Satisfaction of the sufficient attractiveness inequality guarantees that for all $n \geq 1$, the condition*

$$g_1^* = g_n^*$$

holds with corresponding optimum $K_n^* = 1$.

Proof. Assuming that the sufficient attractiveness inequality holds, we first claim that $g_n(K)$ is nondecreasing. Beginning with

$$\frac{d}{dK} g_n(K) = \frac{1}{n} \frac{d}{dK} \mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right]$$

and noting that $\mathcal{X}_n$ is bounded, results in probability theory, for example, see [140], allow us to commute the differentiation and expectation operators above. Hence,

$$\begin{aligned}\frac{d}{dK} \mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right] &= \mathbb{E}\left[\frac{d}{dK} \log(1 + K\mathcal{X}_n)\right] \\ &= \mathbb{E}\left[\frac{\mathcal{X}_n}{1 + K\mathcal{X}_n}\right].\end{aligned}$$

Now noting that the inequality

$$\frac{z}{1 + Kz} \geq 1 - \frac{1}{1 + z}$$

holds for all $K \in [0, 1]$ and all $z > -1$, we obtain

$$\frac{d}{dK} \mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right] \geq 1 - \mathbb{E}\left[\frac{1}{1 + \mathcal{X}_n}\right].$$

Using the fact that the $X(k)$ are i.i.d., we observe that

$$\mathbb{E}\left[\frac{1}{1+\mathcal{X}_n}\right] = \left(\mathbb{E}\left[\frac{1}{1+X\left(0\right)}\right]\right)^n$$
$$\leq 1.$$

Combining this with the previous inequality, we have

$$\frac{d}{dK}g_n(K) \geq 0$$

which shows that $g_n(K)$ is non-decreasing in K. Hence $g_n(K)$ is maximized at $K = 1$. Therefore, for all n, we have $g_n^* = g_n\left(1\right)$. It only remains to observe that

$$\begin{aligned} g_n\left(1\right) &= \frac{1}{n}\mathbb{E}[\log\left(1+\mathcal{X}_n\right)] \\ &= \frac{1}{n}\mathbb{E}\left[\log\left(\prod_{k=0}^{n-1}\left(1+X\left(k\right)\right)\right)\right] \\ &= \frac{1}{n}\mathbb{E}\left[\sum_{k=0}^{n-1}\log\left(1+X\left(k\right)\right)\right] \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}[\log\left(1+X\left(k\right)\right)] \\ &= \mathbb{E}[\log\left(1+X\left(0\right)\right)] \\ &= g_1^*. \quad \square \end{aligned}$$

6.7 Frequency-Based Formulation for Stock Trading

In the earlier sections of this chapter, our frequency-based theory is formulated in a way that a bettor first declares on which side to bet, then wagers $KV(0)$ with $K \geq 0$. In this section, we now shift our attention from betting games to the stock market and allow K to be negative to accommodate "short selling" which, from a mathematical point of view, is fundamentally different[3] from betting on the $X < 0$ side in the game setting; see Section 1.2.

Consistent with the notation in the preceding chapters, $S(k) > 0$ represents the price for an underlying stock at stage k and then the corresponding returns $X(k)$ are assumed to be i.i.d. random variables with known bounds satisfying $X_{\min} \leq X(k) \leq X_{\max}$ with $-1 < X_{\min} < 0 < X_{\max} < \infty$. Again, we take n to be the number of steps between updates and generally assumed to be arbitrary but fixed. Similar to the formulation for betting games in Section 6.2, given initial account value $V(0) > 0$, the trader places an initial investment $I(0) \doteq KV(0)$ which represents

$$\begin{aligned} N(0) &\doteq \frac{I(0)}{S(0)} \\ &= \frac{|K|V(0)}{S(0)} \end{aligned}$$

shares where $K > 0$ corresponds to going long and $K < 0$ corresponds going short. Then, the trader waits n steps before updating the investment with the account value at stage n is given by

$$\begin{aligned} V(n, K) &= V(0) + \mathrm{sgn}(K)N(0)(S(n) - S(0)) \\ &= V(0) + KV(0)\left(\frac{S(n)}{S(0)} - 1\right) \\ &= V(0) + KV(0)\left[\prod_{k=0}^{n-1}(1 + X(k)) - 1\right] \\ &= (1 + K\mathcal{X}_n)V(0) \end{aligned}$$

[3]As mentioned in Section 1.2, short trader with $K \in [-1, 0]$ borrows shares of a certain underlying stock from a broker then sell it in the hope of making a profit from a subsequent fall in the price of the stock. In our frequency-based framework for stocks, when $n > 1$, the trader who has gone short using $K < 0$ faces a potential bankruptcy issue. On the other hand, in the betting games scenario, as seen in previous sections, when betting on the $X < 0$ side, bankruptcy is ruled out by the assumption $X_{\max} < 1$.

where, we recall $\mathcal{X}_n$ to be the compound return, defined in Section 6.1, with known bounds

$$-1 < \mathcal{X}_{n,\min} \leq \mathcal{X}_n \leq \mathcal{X}_{n,\max}$$

where

$$\mathcal{X}_{n,\max} = (1 + X_{\max})^n - 1;$$

$$\mathcal{X}_{n,\min} = (1 + X_{\min})^n - 1.$$

It is also readily verified that $V(k, K) \geq 0$ for every sample path of returns X and all $k \geq 0$ if and only if

$$K \in \mathcal{K}_n \doteq \left[\frac{-1}{\mathcal{X}_{n,\max}}, \frac{1}{|\mathcal{X}_{n,\min}|} \right].$$

Similar to the arguments for betting games in Section 6.2, this constraint guarantees that the bankruptcy issues and the infinitely negative expected logarithmic growth associated with $V(k, K) < 0$ is avoided. Note that the constraint sets $\mathcal{K}_n$ is nonincreasing; i.e., the nesting condition $\mathcal{K}_{n+1} \subseteq \mathcal{K}_n$ is satisfied for $n \geq 1$. As $n \to \infty$, the condition $K \in \mathcal{K}_n$ reduces to $0 \leq K \leq 1$. Hence, $[0, 1] \subset \mathcal{K}_n$ for all $n \geq 1$. This implies that $K \in [0, 1]$ guarantees no-bankruptcy for all $k \geq 0$. Now, we proceed as we already have many times in this chapter and its predecessors. That is, we seek an optimal feedback gain K_n^* maximizing the concave function

$$\begin{aligned} g_n(K) &\doteq \frac{1}{n}\mathbb{E}\left[\log \frac{V(n, K)}{V(0)} \right] \\ &= \frac{1}{n}\mathbb{E}[\log(1 + K\mathcal{X}_n)] \end{aligned}$$

subject to $K \in \mathcal{K}_n$ and again denote the optimal expected logarithmic growth by g_n^*. We should also note that the problem formulation above can be modified to accommodate additional constraints which occur in practice. For example, for a *cash-financed* trader who is long only, we argument the constraint $\mathcal{K}_n$ with $K \in [0, 1]$. For the cash-financed trader who is either long or short, then we require $K \in \mathcal{K}_n \cap [-1, 1]$.

6.8 On Which Side Should Trader Trade?

The following preliminary lemma can be viewed as the stock-trading version of the Lemma 6.3.1 which arose in the context of a betting game.

Lemma 6.8.1 (Preliminaries): *For the frequency-based stock-trading scenario described above with $K \in \mathcal{K}_n$, if $\mathbb{E}[X(0)] > 0$, then the optimal Kelly feedback gain $K_n^* > 0$. If $\mathbb{E}[X(0)] < 0$, then $K_n^* < 0$. If $\mathbb{E}[X(0)] = 0$, then $K_n^* = 0$.*

Proof. The same idea used to prove Lemma 6.3.1 is easily adapted to the situation at hand. Hence, for the sake of brevity, we only prove that $\mathbb{E}[X(0)] > 0$ implies $K_n^* > 0$. Proceeding by contradiction, we assume the optimum $K_n^* < 0$. With the aid of Jensen's inequality, we have

$$g_n(K_n^*) = \frac{1}{n}\mathbb{E}[\log(1 + K_n^*\mathcal{X}_n)] \leq \frac{1}{n}\log(1 + K_n^*\mathbb{E}[\mathcal{X}_n]).$$

Using the fact that returns $X(k)$ are i.i.d. and $\mathbb{E}[X(0)] > 0$, we observe that

$$\begin{aligned}\mathbb{E}[\mathcal{X}_n] &= \mathbb{E}\left[\prod_{k=0}^{n-1}(1 + X(k))\right] - 1 \\ &= (1 + \mathbb{E}[X(0)])^n - 1 > 0.\end{aligned}$$

However, since $K_n^* < 0$, and $\mathbb{E}[\mathcal{X}_n] > 0$, it follows that $g_n(K_n^*) < 0 = g_n(0)$, which contradicts the optimality of K_n^*. Hence, $K_n^* \geq 0$.

To complete the proof, it remains to show that $K_n^* > 0$. To see this, it suffices to show that $g_n(K) > 0$ for some $K > 0$ in some neighborhood of zero. Consider a neighborhood of zero by restricting $K \in \mathcal{K}_n \cap [-a, a]$ for some sufficiently small $0 < a < 1$ so that $1 + K\mathcal{X}_n > 0$. Taking the derivative

$$\frac{d}{dK}g_n(K) = \frac{1}{n}\frac{d}{dK}\mathbb{E}\left[\log(1 + K\mathcal{X}_n)\right]$$

and noting that since $\mathcal{X}_n$ is bounded and $1 + K\mathcal{X}_n > 0$, results in probability theory, for example, see [138], allow us to commute the differentiation and expectation operators above.

Thus, if we evaluate the derivative at $K = 0$, it follows that

$$\frac{d}{dK}g_n(K)\bigg|_{K=0} = \frac{1}{n}\mathbb{E}\left[\frac{\mathcal{X}_n}{1+K\mathcal{X}_n}\right]\bigg|_{K=0} = \frac{1}{n}\mathbb{E}[\mathcal{X}_n] > 0$$

where the last inequality holds since $\mathbb{E}[X(0)] > 0$. Therefore, we obtain

$$\frac{d}{dK}g_n(K)\bigg|_{K=0} > 0.$$

Since $g_n(0) = 0$, it follows that there exists some $K > 0$ suitably small so that $g_n(K) > 0$, which implies that there is an optimum $K_n^* > 0$. $\square$

6.9 The High-Frequency Maximality Theorem for Stock Trading

The theorem, given in Section 6.4 for betting games, can be readily modified to address the stock trading scenario under consideration. In the sequel, we assume that $K \in \mathcal{K}_n \cap [-1, 1]$; i.e., survival and cash-financing are required. Then, the following theorem tells us that when the i.i.d. returns of the underlying stock satisfy $\mathbb{E}[X(0)] > 0$, we have $g_n^* \leq g_1^*$; i.e., high-frequency trading is *unbeatable*. On the other hand, for the case where $\mathbb{E}[X(0)] < 0$, after the theorem, we conjecture that the same conclusion holds and provide a supporting example.

Theorem 6.9.1 (High-Frequency Maximality for Stock Trading): *For the frequency-based stock-trading scenario, if* $\mathbb{E}[X(0)] > 0$*, it follows that* $g_n^* \leq g_1^*$ *for all* $n \geq 1$*.*

Proof. When $\mathbb{E}[X(0)] > 0$, Lemma 6.8.1 tells us that the optimum $K_n^* > 0$. Thus, in combination with the cash-financing requirement, it suffices to consider the expected logarithmic growth optimization problem for $K \in [0, 1]$. Then, the same proof used in High-Frequency Maximality Theorem 6.4.1 leads to the desired result. $\square$

6.9.2 A Conjecture on High-Frequency Maximality for Short Selling: Note that the Theorem 6.9.1 does not address the case $\mathbb{E}[X(0)] \leq 0$; i.e., we have no proof available. Nevertheless, we conjecture that $g_n^* \leq g_1^*$ for $n \geq 1$ still holds. We first note that for the trivial case if $\mathbb{E}[X(0)] = 0$, by Lemma 6.8.1, we obtain $K_n^* = 0$, which implies that $g_n^* = g_1^* = 0$. Support for this conjecture is given in the example below.

6.9.3 Example (Binary Lattice Model): We consider a binary lattice model with returns $X_{\max} = -X_{\min} = \gamma$ with $0 < \gamma < 1$ with probability

$$p \doteq P(X(k) = \gamma) < 1/2.$$

Hence, it is readily verified that $\mathbb{E}[X(k)] < 0$. To support the conjecture above, our goal is to show that $g_1^* > g_n^*$ for $n \geq 1$. Indeed, noting that with $0 < \gamma < 1$, the random variable $\mathcal{X}_n$ has point masses located at

$$x_i \doteq (1+\gamma)^i (1-\gamma)^{n-i} - 1$$

for $i = 0, 1, 2, \dots, n$, with associated probabilities

$$p_i \doteq P\left(\mathcal{X}_n = x_i\right) = \binom{n}{i} p^i (1-p)^{n-i},$$

we obtain

$$g_n\left(K\right) = \frac{1}{n} \sum_{i=0}^{n} p_i \log\left(1 + K x_i\right).$$

We maximized the function above with respect to $K \in [-1, 1]$ with $\gamma \doteq 1/10$ and $p \doteq j/10$ for $j = 1, \dots, 4$. A lengthy but straightforward calculation leads to

$$K_n^* = \max\left\{ -1, \frac{-1}{(1+\gamma)^n - 1} \right\}.$$

A plot of the expected logarithmic growth $g_n(K)$ versus K for $p = 1/10$ and $n = 1, 2, 5, 10$ is shown in Figure 6.4. Consistent with the conjecture, the figure indicates that

$$g_1^* = g_1(-1) > g_n^*$$

for various choices of $n = 2, 5, 10$. Although not shown in the figure, similar results were obtained for probability $p = 2/10$, $3/10$, $4/10$.

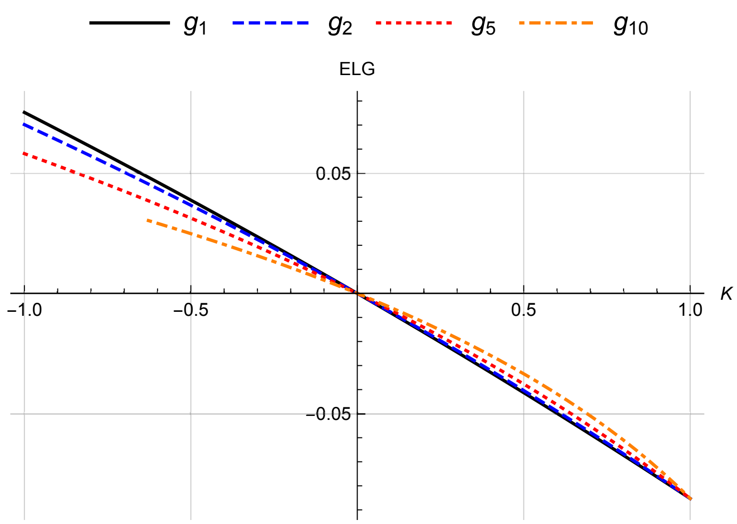


Figure 6.4: Illustrative Example with $\mathbb{E}[X(0)] < 0$ and $p = 1/10$

6.9.4 More Practical Example Related to Implementation: We now consider an example aimed at demonstrating some of the issues which arise in practice when one seeks to use the linear feedback controllers above and study high-frequency maximality. In particular, we use the high-frequency historical intra-day tick data for Apple (ticker: AAPL); see [141] for the period 9:30:00 AM to 2:13:47 PM on December 2, 2015; see Figure 6.5. During this period, we have $N = 110,000$ ticks. Each "tick" corresponds to a transaction which takes the realized stock price from $s(k)$ to $s(k + 1)$. The average time between arrivals of ticks is about one tenth of a second.

The first step in our analysis is to use the time series of prices $s(k)$ to calculate the corresponding realized returns

$$x(k) \doteq \frac{s(k+1) - s(k)}{s(k)}$$

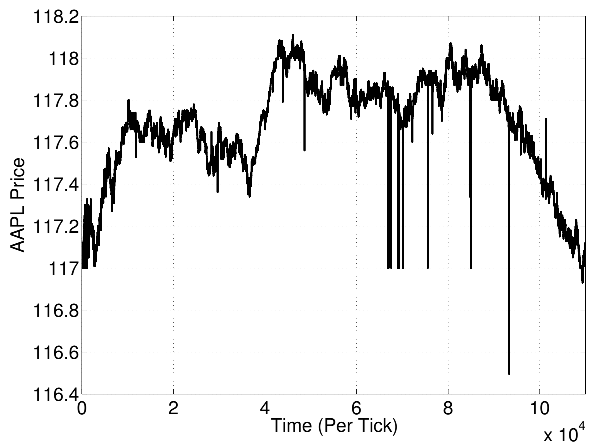


Figure 6.5: AAPL Tick-by-Tick Price of Trade

for $k = 0, 1, \ldots, N - 1$. Note that the smallness of the inter-tick price changes and returns leads to the probability masses largely concentrated between $x = -0.0002$ and $x = 0.0002$. Using this data, we now create an empirical probability mass function (PMF), call it $\widehat{f}_X(x)$, for the random variables $X(k)$, presumably i.i.d., entering into the theory. Namely,

$$\widehat{f}_X(x) = \frac{1}{N} \sum_{j=0}^{N-1} \delta(x - x(j))$$

with $\delta(x - x(j))$ being the Dirac Delta function at sample point $x(j)$. This serves as input to the maximization problem for $g(K)$.

To carry out this maximization, we generated a plot of $g(K)$, which is shown in Figure 6.6. The solid line in the figure represents the performance for the high-frequency case in the sense that the waiting period $n\Delta t$ is obtained with $n = 1$ where Δt is the average time between arrivals of ticks. From the figure, the corresponding Kelly optimum and expected logarithmic growth are $K_1^* \approx 0.824$ and $g_1^* = 6.71 \times 10^{-9}$, respectively. Interestingly, in this example, the price has no obvious "bullish" pattern but yet we see that the theory leads to a rather aggressive bet size which is more than 80% of the account value.

Now, to study g_n^* for $n > 1$, per Section 6.7, we work with the empirical estimates of the compound returns. For example, for $n = 10$, we use the *realized compound returns*

$$\chi_{10}(m) \doteq \prod_{k=0}^{9}(1 + x(k+m)) - 1$$

for $m = 0, 10, 20, \ldots, N-10$ to estimate an empirical PMF. Based on the data, the Kelly optimum and the associated expected logarithmic growth are estimated as $K_{10}^* \approx 1$ and $g_{10}^* \approx 6.2 \times 10^{-9}$. In Figure 6.6, we see that $g_1^* > g_{10}^*$, which is consistent with our High-Frequency Maximality Theorem.

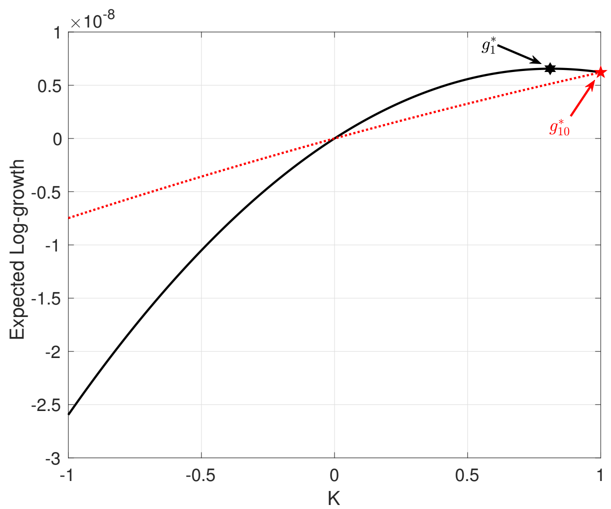


Figure 6.6: Expected Logarithmic Growth Versus K with $n = 1$ and $n = 10$

6.10 Sufficient Attractiveness Theorem in Stock Trading Context

Theorem 6.6.1 and its proof, developed for betting games in Section 6.2, are readily modified to the context of stock trading for a long-only trader. To avoid redundancy, we simply state the result without proof below.

Theorem 6.10.1 (Sufficient Attractiveness in Stock Trading): *For the frequency-based stock-trading scenario, if the sufficient attractiveness inequality*

$$\mathbb{E}\left[\frac{1}{1+X(0)}\right] \leq 1,$$

holds, then for all $n \geq 1$, $K_n^* = 1$ *and the condition* $g_1^* = g_n^*$ *holds.*

6.10.2 Remarks: The theorem above tells us that when a trade is sufficiently attractive, the buy-and-hold strategy achieves the *same* performance as the high-frequency strategy. Loosely speaking, high-frequency trading is a waste of time. We also remind the reader about the condition

$$\mathbb{E}\left[\frac{1}{1-X(0)}\right] \leq 1$$

which is explained in Sections 6.5.1 in the context of betting games on the $X < 0$ side. For the case of stock trading, if $X_{\max} < 1$ and the inequality above is satisfied, we were only able to obtain a much weaker result: For the case $n = 1$ and $K \in [-1, 0]$, it can be shown that $K_1^* = -1$ is the optimum. Said another way, if the inequality above holds, the underlying stock *is* sufficiently attractive to a high-frequency short trader and the optimum is, at each k, to short $V(k)$ worth of stock.

6.11 From a Single Stock to a Multi-Stock Portfolio

We begin with a trader who is forming a portfolio consisting of $m \geq 2$ assets and assume that at least one of them is riskless with nonnegative rate of return $r \geq 0$. That is, if an asset is riskless, its return is deterministic and is treated as a degenerate random variable with value r for all k with probability one. It is noted that this formulation also allows the trader to maintain cash in the account; i.e., this corresponds to the case $r = 0$. Alternatively, if Asset i is a stock whose price at time k is $S_i(k)$, then its return is

$$X_i(k) = \frac{S_i(k+1) - S_i(k)}{S_i(k)}.$$

In the sequel, for stocks, we assume that the return vectors $X(k)$ have a known distribution and have components $X_i(k)$ which can be arbitrarily correlated. It is also assumed that these vectors are i.i.d. with components satisfying $X_{\min,i} \leq X_i(k) \leq X_{\max,i}$ with known bounds above and with $X_{\max,i}$ being finite and $X_{\min,i} > -1$.

Our objective now is to extend the frequency-based formulation from trading a single stock to trading a multi-stock portfolio. We provide some initial results along these lines aimed at optimal selection of portfolio weights. To preview the main result to follow, we consider the case when the portfolio is comprised of two or more potentially investable assets with each having i.i.d. returns $X_i(k)$ and possibly correlated, Asset j is said to be *dominant* if

$$\mathbb{E}\left[\frac{1+X_i(0)}{1+X_j(0)}\right] \leq 1$$

holds for all $i \neq j$. In this case, our main result, which we call the *Dominant Asset Theorem*, tells us that when this condition is satisfied, an optimal strategy is obtained by investing all of the trader's funds in Asset j. Figuratively speaking, this result says that an optimal portfolio is obtained by putting all eggs in one basket. Of equal importance, as a consequence of the theorem, it is seen that the performance of the high-frequency trader and the buy and holder are identical. That is, $g_n^* = g_1^*$ for all $n \geq 1$. Thus, there is no benefit associated with trading often; it suffices to buy and hold.

6.11.1 Feedback Control Formulation: As discussed in the previous chapters, the finance problems considered in this dissertation can be reformulated in terms of feedback control systems. In this case, the system output at stage k is taken to be the trader's time-varying account value $V(k)$ and the i-th feedback gain $0 \leq K_i \leq 1$ represents the fraction of the account allocated to the i-th asset. Said another way, the i-th controller is a linear feedback of the form $I_i(k) = K_iV(k)$. Since $K_i \geq 0$, the trader is going *long*. In view of the above and recalling that there is at least one riskless asset available, without loss of generality, we consider the unit simplex constraint

$$K \in \mathcal{K} \doteq \left\{ K \in \mathbb{R}^m : K_i \geq 0 \text{ for all } i, \ \sum_{i=1}^{m} K_i = 1 \right\}$$

which is classical in finance. That is, with $K \in \mathcal{K}$, we have a guarantee that 100% of the account is invested.

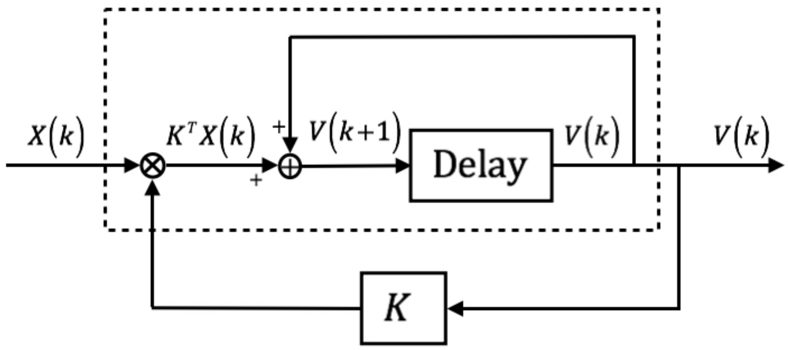


Figure 6.7: Feedback Configuration for Trading

With the setup above, the update in account value from stage k to $k+1$ for the resulting closed-loop system, depicted in Figure 6.7, is $V(k+1) = (1 + K^T X(k))V(k)$. Letting n be the number of steps between rebalancings, at time k, preceding as in the single-stock case, the trader begins with initial investment

$$I(0) = \sum_{i=1}^{m} K_i V(0)$$

and waits n steps in the spirit of "buy and hold." Then, when $k = n$, the investment is updated to be

$$I(n) = \sum_{i=1}^{m} K_i V(n).$$

Now, to study performance as a function of frequency, for $i = 1, 2, \ldots, m$, we use the compound returns

$$\mathcal{X}_{n,i} \doteq \prod_{k=0}^{n-1} (1 + X_i(k)) - 1$$

which are readily seen to satisfy $\mathcal{X}_{n,i} > -1$ for all n and we work with the random vector $\mathcal{X}_n$ having i-th component $\mathcal{X}_{n,i}$. Then, for any initial account value $V(0) > 0$ and $n \geq 1$, the corresponding account value at stage n is given by $V(n) \doteq (1 + K^T\mathcal{X}_n)V(0)$.

6.11.2 Frequency-Dependent Optimization Problem in the Portfolio Case: Generalizing on the analysis in Section 6.2, for arbitrary $n \geq 1$, we study the problem of maximizing the expected logarithmic growth

$$g_n(K) = \frac{1}{n}\mathbb{E}\left[\log(1 + K^T\mathcal{X}_n)\right],$$

which is concave in K with the associated optimal expected logarithmic growth g_n^*. Furthermore, any $K_n^* \in \mathcal{K}$ satisfying

$$g_n(K_n^*) = g_n^*$$

is called an *optimal Kelly feedback gain vector* for the rebalancing period of length n.

6.12 Relative Attractiveness and Dominant Asset

In this section, we generalize the definition of sufficient attractiveness, given in Section 6.5 for a single asset, to the multi-asset case. Subsequently, we explore the use of the new definition.

6.12.1 Definition (Relative Attractiveness and Dominant Asset): Given the collection of $m \geq 2$ assets, we say that Asset j is *relatively more attractive* than Asset i if

$$\mathbb{E}\left[\frac{1 + X_i(0)}{1 + X_j(0)}\right] \leq 1.$$

Equivalently, Asset j is relatively more attractive than Asset i if the correlation between $[1 + X_j(0)]^{-1}$ and $1 + X_i(0)$ is at most one. Asset j is said to be *dominant*[4] if it is relatively more attractive than every other asset $i \neq j$.

[4]There is a difference between the definition of dominant asset and the definition of stochastic dominance. Note that stochastic dominance involves only the marginal distributions of two random variables, while the

6.12.2 Remarks: The definition above works for the single asset case by viewing cash with zero return as a second asset. When $m = 2$ and the other asset is riskless with zero return, then the dominance is equivalent to the notion of sufficient attractiveness as defined in Section 6.5 for the single-stock case. Also, when $m = 2$, we note that Asset j is dominant if and only if it is relatively more attractive than the other asset. If $m \geq 2$, the definition above tells us that a riskless Asset j with return r is easily seen to be relatively more attractive than risky Asset i if and only if

$$\mathbb{E}[X_i(0)] \leq r.$$

For a risky Asset j to be relatively more attractive than the riskless Asset i, we require more than just $\mathbb{E}[X_j(0)] > r$. For example, suppose the returns $X_j(k) \in \{-1/2, 1/2\}$ with

$$P(X_j(k) = 1/2) = 0.6.$$

Then with $X_i(k) = r = 0.05$, a straightforward calculation leads to $\mathbb{E}[X_j(0)] > r$, but

$$\mathbb{E}\left[\frac{1 + X_i(0)}{1 + X_j(0)}\right] = 1.26$$

which violates the relative attractiveness inequality.

Although the condition $\mathbb{E}[X_j(0)] \geq r$ is not sufficient for a risky Asset j to be relatively more attractive than a riskless Asset i, the condition is necessary. This can be seen by applying Jensen's inequality to obtain

$$\frac{1 + r}{\mathbb{E}[1 + X_j(0)]} \leq \mathbb{E}\left[\frac{1 + r}{1 + X_j(0)}\right].$$

If Asset j is relatively more attractive than riskless Asset i, then the right hand side above is one at most, and we obtain

$$\frac{1 + r}{\mathbb{E}[1 + X_j(0)]} \leq 1$$

dominant asset definition involves the correlation between $[1 + X_j(0)]^{-1}$ and $1 + X_i(0)$, which depends on the joint distribution of $X_j(0)$ and $X_i(0)$. For the background theory related to stochastic dominance, the reader is referred to [125].

from which it follows that $\mathbb{E}[X_j(0)] \geq r$. Thus, $\mathbb{E}[X_j(0)] \geq r$ is necessary, but not sufficient for risky Asset j to be relatively more attractive than riskless Asset i.

6.13 The Dominant Asset Theorem

The theorem below tells us that the satisfaction of the dominant asset inequality leads to an optimal portfolio which involves investing 100% of available funds in a single asset. Figuratively speaking, this result says that an optimal portfolio is obtained by putting all eggs in one basket.

Theorem 6.13.1 (Dominant Asset): *Given a collection of m assets, if Asset j is dominant, then, for all $n \geq 1$, $g_n(K)$ is maximized by $K_n^* = e_j$ where e_j is the unit vector in the j-th coordinate direction. Furthermore, the resulting optimal expected logarithmic growth rate is given by*

$$g_n^* = g_1^* = \mathbb{E}\left[\log(1 + X_j(0))\right].$$

Proof. In order to prove $K_n^* = e_j$, it suffices to show that $g_n(K) \leq g_n(e_j)$ for $K \in \mathcal{K}$. Letting

$$\mathbf{1} \doteq [1 \;\; 1 \;\; \cdots \;\; 1]^T \in \mathbb{R}^m,$$

for notational convenience, we work with the random vector

$$\mathcal{R}_n \doteq \mathcal{X}_n + \mathbf{1}$$

representing the total return with i-th component $\mathcal{R}_{n,i}$. Since $K^T\mathbf{1} = 1$ for $K \in \mathcal{K}$, it follows that

$$\begin{aligned} g_n(K) &= \frac{1}{n}\mathbb{E}[\log(1 + K^T\mathcal{X}_n)] \\ &= \frac{1}{n}\mathbb{E}[\log(K^T\mathcal{R}_n)]. \end{aligned}$$

Hence, by applying Jensen's inequality to the concave logarithm function above, we obtain the chain of inequalities

$$\begin{aligned}
g_n(K) - g_n(e_j) &= \frac{1}{n}\mathbb{E}\left[\log \frac{K^T\mathcal{R}_n}{\mathcal{R}_{n,j}}\right] \\
&\leq \frac{1}{n}\log \mathbb{E}\left[\frac{K^T\mathcal{R}_n}{\mathcal{R}_{n,j}}\right] \\
&= \frac{1}{n}\log\left(\sum_{i=1}^{m} K_i \mathbb{E}\left[\frac{\mathcal{R}_{n,i}}{\mathcal{R}_{n,j}}\right]\right) \\
&= \frac{1}{n}\log\left(\sum_{i=1}^{m} K_i \mathbb{E}\left[\prod_{k=0}^{n-1}\frac{1+X_i\left(k\right)}{1+X_j\left(k\right)}\right]\right) \\
&= \frac{1}{n}\log\left(\sum_{i=1}^{m} K_i \left(\mathbb{E}\left[\frac{1+X_i\left(0\right)}{1+X_j\left(0\right)}\right]\right)^n\right) \\
&\leq \frac{1}{n}\log 1 \\
&= 0
\end{aligned}$$

where the last inequality follows from the dominance of Asset j and the fact that $K \in \mathcal{K}$. Now, it follows that $g_n(K) \leq g_n(e_j)$ and $g_n^* = g_n(e_j)$. To complete the proof, it remains to show that $g_n^* = g_1^*$. This is readily obtained by recalling that the $X_i(k)$ are i.i.d. and observing that

$$\begin{aligned}
g_n^* &= g_n(e_j) \\
&= \frac{1}{n}\mathbb{E}[\log \mathcal{R}_{n,j}] \\
&= \frac{1}{n}\mathbb{E}\left[\log \prod_{k=0}^{n-1}(1+X_j(k))\right] \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}\left[\,\log\left(1+X_j(k)\right)\,\right] \\
&= \mathbb{E}\big[\log\big(1+X_j(0)\big)\big] \\
&= g_1(e_j) \\
&= g_1^*. \quad \square
\end{aligned}$$

6.14 Application to Stock-Market Data

In this section, our objective is to illustrate some of the practical issues which arise when studying asset dominance using historical data. To this end, we consider three assets as portfolio candidates. Asset 1 is Netflix (ticker: NFLX), Asset 2 is Facebook (ticker: FB) and Asset 3 is a riskless asset with daily interest rate $r \geq 0$. We consider the problem of rebalancing our positions in these assets over the four-year period beginning on January 24, 2013 and work with the adjusted daily closing prices. In this setting, we demonstrate how the Dominant Asset Theorem might apply in practice. The price plot for the two stocks in Figure 6.8 begins with the 126-day period prior to the start of trading. This data was used as a training set to initialize the analysis to follow. In the sequel, using $s_1(k)$ and $s_2(k)$ for the kth daily *realized prices* of Netflix and Facebook respectively, we first calculate the associated *realized return*, call it $x_i(k)$, where

$$x_i(k) \doteq \frac{s_i(k+1) - s_i(k)}{s_i(k)}$$

for $i = 1, 2$, and, as mentioned above, $x_3(k) = r$.

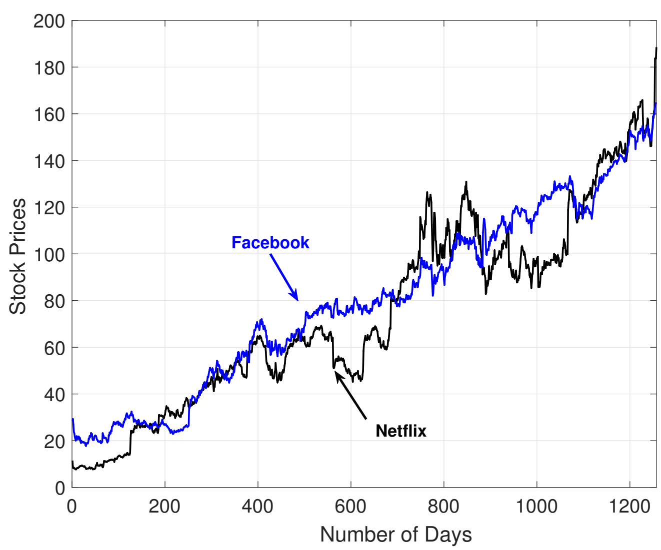


Figure 6.8: Stock Prices of Facebook and Netflix

When testing for satisfaction of the relative attractiveness inequalities, we work with a sliding window consisting of the most recent N trading days. Hence, at day k, we use an empirical estimate of the expected value for the attractiveness ratio involving the i-th and j-th assets which is given by

$$R_{ij}(k) \doteq \frac{1}{N}\sum_{\ell=0}^{N-1}\frac{1+x_i(k-\ell)}{1+x_j(k-\ell)}.$$

The simulations to follow use a window size of $N = 126$, which corresponds to about six months. Beginning with the initial condition $R_{ij}(0)$ established using the training set, we generate the $R_{ij}(k)$ over the period of interest. In view of the nonstationarity of the returns, as seen in the simulation to follow, the $R_{ij}(k)$ are time-varying. Hence, an asset which is dominant at one point in time may no longer be dominant at a later time.

With the considerations above, we begin our analysis with $r = 0$ and consider the following two questions. *Question 1*: In a zero interest rate environment, over what time periods is Netflix the dominant asset? During such periods, in accordance with the theorem, the trader would be non-diversified with the entirety of $V(k)$ in this stock. *Question 2*: At stage k, how large must the interest rate r be so that the riskless asset is dominant? That is, when the interest rate is suitably high, our theory dictates that the trader has a portfolio which is 100% in fixed income with no positions in Netflix and Facebook.

To answer the first question, we define the relative attractiveness estimator for Netflix

$$R_1(k) \doteq \max\{R_{21}(k), R_{31}(k)\}.$$

A sample path of $R_1(k)$ versus k is shown in Figure 6.9, and, consistent with the theorem, we deem Netflix to be *dominant* over the subset of time periods for which $R_1(k) \leq 1$. Over the periods when $R_1(k) > 1$, there are various additional scenarios which can be studied with the given data. For example, sometimes there is no dominant asset and at other times either Facebook or the riskless asset is dominant.

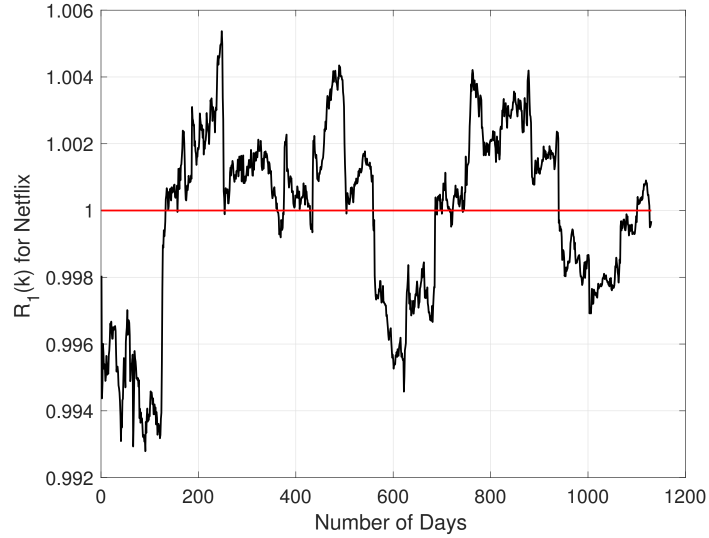


Figure 6.9: Relative Attractiveness Estimator for Netflix

To answer the second question, we observe that the theorem tells us that Asset 3, the riskless asset, is dominant if and only if

$$\max\{\mathbb{E}[X_1(0)], \mathbb{E}[X_2(0)]\} \leq r.$$

Noting that the nonstationarity of the $X(k)$ process is implicit when working with our sliding window, we let $r^*(k)$ denote the estimated value of the left hand size above based on the most recent N-day window. That is, we define the interest rate estimator for dominance of the riskless asset by

$$r^*(k) \doteq \frac{1}{N} \max\left\{\sum_{\ell=0}^{N-1} x_1(k-\ell), \sum_{\ell=0}^{N-1} x_2(k-\ell)\right\}.$$

In Figure 6.10, we see that there are time periods when the market is performing quite well and it takes a remarkably high interest rate in order to forego investing in the stocks.

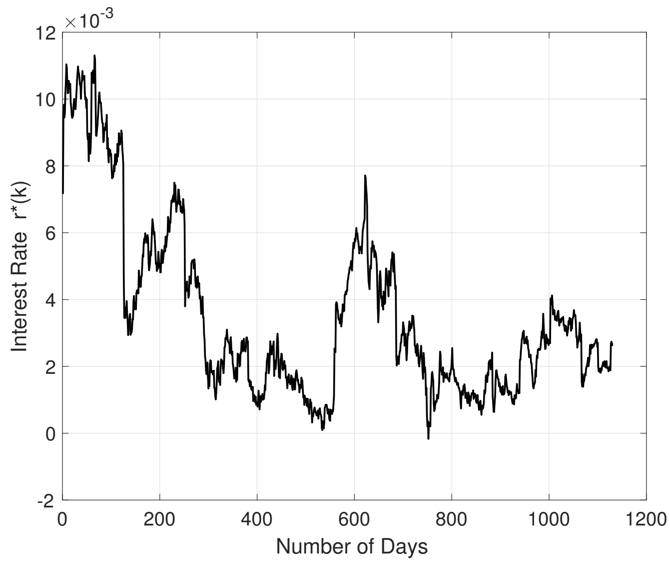


Figure 6.10: Interest Rate Estimator for Dominance of Riskless Asset

6.15 Concluding Remarks and Future Work

In this chapter, we studied the problem of optimizing the betting frequency in a dynamic setting using Kelly's expected logarithmic growth criterion as the performance criterion. We proved that the performance for the high-frequency case, $n = 1$, absent transaction costs, always leads to the best performance. Then, we investigated conditions under which betting with arbitrarily low frequency can still achieve the same performance as betting with very high frequency. To this end, we showed that satisfaction of the sufficient attractiveness inequality assures that $g_n^* = g_1^*$ for all $n \geq 1$. Later in this chapter, we shifted our focus from betting games to a stock-trading scenario. To this end, results obtained for betting games were modified, whenever appropriate, to be apropos to the stock market. Finally, we extended our formulation from a single stock to a multi-stock portfolio case, and we showed that if an asset is dominant, an optimal trading strategy is to invest all available funds in it. For such cases, rebalancing becomes moot and the performance, g_n^*, is constant.

6.15.1 Further Research for $X_{\max} > 1$: As discussed in Section 6.2, when betting on the $X < 0$ side, by switching the signs of the $X(k)$, the theory for the $X > 0$ side can be used provided that $X_{\max} < 1$. When this inequality is violated, working with compound returns $\mathcal{X}_n$ becomes problematic because the possibility of "sign switching" of the compound returns confounds the mathematical analysis, which was used for betting on the $X > 0$ side.

To motivate future research, we now analyze a simple example involving betting on $X < 0$ with $X_{\max} > 1$. Given that the results given in this chapter cannot be used to address this example, we now carry out a "proper" analysis. We consider a coin-flipping game with returns $X(k) \in \{-1/2, 2\}$ with $P(X(k) = 2) = 1/10$ being the probability of seeing heads and

$$p \doteq P(X(k) = -1/2) = 9/10$$

being probability of seeing tails.

For the setup above, it is readily seen that betting on tails is more attractive than on heads. Thus, beginning with $V(0) > 0$, a bettor wagers $KV(0)$ on tails for some $K \geq 0$. We now consider $n = 2$ and proceed to find the optimal Kelly feedback gain K_2^* and the associated expected logarithmic growth g_2^*. The account value at stage $n = 2$ is analyzed using the following cases: For the sample path *(Tails, Tails)*, the bettor wins both the first and second flips with probability p^2, and the account value is $V(2) = (1 + 5K/4)V(0)$, and for the sample path *(Tails, Heads)*, the bettor wins at the first flip, but loses the second flip with probability $p(1-p)$, and the associated account value is $V(2) = (1-5K/2)V(0)$. Finally, for either of the sample paths *(Heads, Tails)* or *(Heads, Heads)*, the bettor loses at the first flip with probability $1 - p$. Given the initial loss is $2KV(0)$, the amount $KV(0)$ must be paid to the house to settle the debt. Subsequently, the second flip has no effect on the account value because no money is left on the table. Hence, it follows that $V(2) = V(1) = (1 - 2K)V(0)$.

Taking all three cases above into account, a straightforward calculation leads to the expected logarithmic growth

$$g_2(K) = \frac{1}{2}\left[p^2 \log\left(1 + \frac{5K}{4}\right) + p(1-p)\log\left(1 - \frac{5K}{2}\right) + (1-p)\log\left(1 - 2K\right)\right].$$

Maximizing this concave function by setting the derivative to zero, via a lengthy but straightforward calculation, we obtain a maximizer $K_2^* \approx 0.2063$ and associated optimal expected logarithmic growth $g_2^* \approx 0.0336$. We should also note that in this example, number of bet which can be made depends on the sample path. As seen above, if the bettor losses the first flip, then there is no possibility to continue.

In the next chapter, the motivation for the analysis is the fact that a trader's interactions with the market are not instantaneous. That is, our analysis to date does not account for delay effects, often referred to as *latency*, which is attributable to factors such as order transmission and trade execution. With this consideration in mind, we extend the frequency-based theory in this chapter to include delay in execution.

Chapter 7

Frequency-Based Framework Involving Delay

Motivated by the fact that a trader's interactions with the stock market are not instantaneous, in this chapter, our goal is to generalize the frequency-based formulation to include delay in trade execution.[1] To our knowledge, no papers in the literature address this issue. In next section, we describe a trading system with unit delay and the associated Kelly optimization problem for the case of single asset case. As in Chapter 6, we distinguish between the high-frequency trader and the buy and holder. In contrast to the no-delay case, the trader who uses long-only feedback gain $0 \leq K \leq 1$ cannot guarantee that a cash-financing requirement can be satisfied when delay is present. In this regard, in Section 7.3.1, we first prove a result which we call the *Cash-Financing Theorem.* This theorem tells us that the trader who is long-only is cash-financed if and only if

$$0 \leq K \leq \frac{1}{1+X_{\max}}$$

with $X_{\max} > 0$ being the known upper bound for the returns $X(k)$.

In Section 7.4, in contrast to the results in Chapter 6, we demonstrate, using a binary lattice model, that the buy-and-hold strategy can actually outperform high-frequency trading when execution delay is in play. Subsequently, we provide a detailed analysis on the issue of *survival*, that is, avoidance of bankruptcy, for the trading system involving delay. We formulate the survival issue as a *all-time state positivity* problem and then we provide results, for both buy and holder and high-frequency trader, which indicate conditions under which $V(k) > 0$ for all k can be assured. Finally, some concluding remarks are provided.

[1]Part of the work reported in this chapter has been published to the IEEE Conference on Decision and Control, and to the IEEE Transaction on Automatic Control; see [60] and [61], respectively.

7.1 Delay Considerations

Motivated by the fact that a trader's interactions with the market are not instantaneous, in this section, our first goal is to extend the formulation in Chapter 6 to incorporate a one-step delay in trade execution. With $S(k) > 0$ being the price for a single stock for $k = 0, 1, \ldots, n$, as in the preceding chapters, we work with the returns $X(k)$ for $k = 0, 1, \ldots, n-1$ which are assumed to be independent and identically distributed random variables satisfying

$$X_{\min} \leq X(k) \leq X_{\max}$$

with $-1 < X_{\min} < 0 < X_{\max} < \infty$ being known bounds. The assumption $X_{\min} > -1$ excludes the case that the underlying company's stock have zero price.

Additionally, consistent with preceding chapters, we assume that stock-trading occurs within an idealized market. That is, we assume zero transaction costs, zero interest rates and perfect liquidity conditions. There is no gap between the bid and ask prices, and the trader can buy or sell any number of shares including fractions at the traded price $S(k)$.

As previously mentioned, we primarily focus on the performance of two traders. The first is a high-frequency trader who submits an order at each stage, and the other is a buy and holder who only submits one order at $k = 0$ and waits n steps before updating. In the sequel, we require all traders to be long-only with investment $I(k) \geq 0$, and cash-financed; i.e., $I(k) \leq V(k)$. In the sequel, the trader's investment level is described by a linear time-invariant feedback

$$I(k) = KV(k)$$

with $K \geq 0$ and the cash-financing requirement will lead to a restriction on the range of K.

7.2 Formulation with Unit Execution Delay

We now extend the frequency-based formulation of Chapter 6 to include a unit execution delay.[2] One complication is that while in the delay-free case, an order specified in dollars is equivalent to an order specified in shares, when orders are delayed this is no longer true. To illustrate the difference between orders in dollars versus shares, we first suppose that a broker were willing to accept an order in dollars. Then an order made at stage $k = 0$ to buy shares worth $KV(0)$ dollars is executed at stage $k = 1$, at price $S(1)$. This results in a stock holding at $k = 1$, whose value is exactly $KV(0)$. In contrast, consider an order made at stage $k = 0$ for

$$N(0) \doteq \frac{KV(0)}{S(0)}$$

shares, which is executed at stage $k = 1$ at price $S(1)$. The cost of these shares is

$$N(0)S(1) = KV(0)(1 + X(0)).$$

The factor $1+X(0)$, which involves the random return $X(0)$, implies that the cost of executed the trade, and therefore the dollar amount invested, is uncertain. It is noted that this investment amount is distinctly different from the previous analysis where no delay was present. Also, it should be emphasized here that the analysis above at $k = 0$ holds for both the high-frequency trader and for the buy and holder.

For the case of the high-frequency trader, similarly, our convention is that at stage k, the trader places an order for

$$N_1(k) \doteq \frac{KV_1(k)}{S(k)}$$

shares which are purchased at stage $k + 1$ at price $S(k + 1)$. Consistent with Chapter 6, the subscript 1 is used to indicate the trade is made by the high-frequency trader.

[2] Although not considered here, it is also possible to model and analyze the effect of delays in various other parts of the trading systems. For example, a delay might be present in transmission of information from the exchange to the trader. Further extensions of the theory to follow are relegated to future research.

7.2.1 Account Value Dynamics: For the high-frequency trader, we require that the corresponding investment executed at stage k to be

$$I_1(k) \doteq N_1(k-1)S(k)$$

and must satisfy the cash-financing requirement $0 \leq I_1(k) \leq V_1(k)$ for all $k \geq 1$. Then the evolution of the account value is described by

$$V_1(k+1) = V_1(k) + N_1(k-1)(S(k+1) - S(k))$$

for $k \geq 1$ with $V_1(0) = V_1(1) = V(0) > 0$.

On the other hand, for the buy and holder, since only one order is executed at stage $k = 1$, the long-only and cash-financing conditions force the corresponding investment to be

$$I_n(1) \doteq N_n(0)S(1)$$

where

$$N_n(0) \doteq \frac{KV_n(0)}{S(0)}$$

and must satisfy $0 \leq I_n(1) \leq V_n(1)$. Then the corresponding account value is readily shown to satisfy the recursion

$$V_n(k+1) = V_n(k) + N_n(0)(S(k+1) - S(k))$$

for $k \geq 1$ with $V_n(0) = V_n(1) = V(0) > 0$. Given the fact that the number of shares never changes, it is straightforward to obtain the closed-form

$$V_n(n) = \left(1 + K(1 + X(0))\left(\prod_{k=1}^{n-1}(1 + X(k)) - 1\right)\right)V_n(0).$$

This formula is useful when we study the survival issue in Section 7.5.

7.2.2 Expected Logarithmic Growths for Two Traders: Similar to the case without delay described in Chapter 6, for expected logarithmic growth (ELG) purposes, we use the notation $g_1(K)$ and $g_n(K)$ to denote the performance, as a function of K, achieved by high-frequency trading and buy and hold, respectively. That is, we consider

$$g_1(K) = \frac{1}{n}\mathbb{E}\left[\log\frac{V_1(n)}{V(0)}\right];$$

$$g_n(K) = \frac{1}{n}\mathbb{E}\left[\log\frac{V_n(n)}{V(0)}\right].$$

In addition, we denote maximizers by K_1^* and K_n^* and the associated maximal values by g_1^* and g_n^*, respectively.

7.3 On Cash-Financing with Delay

In contrast to the no-delay case, $K \leq 1$ does not guarantee cash-financing. For the buy and holder at stage $k = 1$, cash-financing requires $I_n(1) \leq V_n(1)$, which is equivalent to

$$\frac{KV_n(0)}{S(0)}S(1) \leq V_n(1).$$

Now using the fact that $S(1)/S(0) = 1 + X(0)$ and $V_n(0) = V_n(1)$, the inequality above holds for all possible values of $X(0)$ if and only if

$$K \leq \frac{1}{1 + X_{\max}}.$$

Combining this with the long-only constraint that $I_n(1) \geq 0$, we have

$$0 \leq K \leq \frac{1}{1 + X_{\max}}.$$

For the case of the high-frequency trader, as seen in the theorem below, once again, the same restriction on K results, but a lengthier argument is required.

Theorem 7.3.1 (Cash-Financing): *For the case of one-step delay in execution, the high-frequency trader is long-only and cash-financed if and only if*

$$0 \leq K \leq \frac{1}{1 + X_{\max}}.$$

Furthermore, when K satisfies the inequality above, the trader's account value satisfies $V_1(k) \geq 0$ for all $k \geq 0$.

Proof. The necessity of the conditions on K are established via the same argument used for the buy and holder given preceding the statement of the lemma. To prove sufficiency, we first establish the preliminary result that $V_1(k) \geq 0$ for all k. To see this, we note that $V_1(0) = V_1(1) > 0$. Using the assumed inequality on K, we have that

$$\begin{aligned} V_1(2) &= V_1(1) + N_1(0)(S(2) - S(1)) \\ &= (1 + K(1 + X(0))X(1))V_1(0) \\ &\geq \left(1 + \frac{1}{1 + X_{\max}}(1 + X_{\max})X_{\min}\right)V_1(0) \\ &= (1 + X_{\min})V_1(0) \geq 0. \end{aligned}$$

Continuing by induction, it follows that $V_1(k) \geq 0$ for all k. To complete the proof of sufficiency, we first use the nonnegativity of $V_1(k)$ to see that

$$\begin{aligned} I_1(k) &= N_1(k-1)S(k) \\ &= (1 + X(k-1))KV_1(k-1) \geq 0. \end{aligned}$$

To show that $I_1(k) \leq V_1(k)$ for all $k \geq 1$, we proceed by induction by noting that for $k = 1$,

$$\begin{aligned} I_1(1) &= (1 + X(0))KV_1(0) \\ &\leq (1 + X_{\max})KV_1(0) \leq V_1(0) = V_1(1). \end{aligned}$$

We next fix any $k \geq 1$, and suppose that for all paths $(X(0), X(1), \ldots, X(k-1))$, we have $I_1(i) \leq V_1(i)$ for $i \leq k$. We now show $I_1(k+1) \leq V_1(k+1)$ using two cases:

Case 1: If $X(k) \geq 0$, then using the assumed bound on K and the fact that $I_1(k) \geq 0$, we obtain

$$\begin{aligned} I_1(k+1) &\leq (1+X_{\max})KV_1(k) \\ &\leq V_1(k) \\ &\leq V_1(k) + I_1(k)X(k) = V_1(k+1). \end{aligned}$$

Case 2: If $X(k) < 0$, then, with the aid of the standing inductive hypothesis $I_1(k) \leq V_1(k)$, we have

$$I_1(k)X(k) \geq V_1(k)X(k).$$

Now, using the facts that $0 \leq K \leq 1/(1+X_{\max}) < 1$ and $X(k) > -1$, we observe that

$$\begin{aligned} V_1(k+1) &= V_1(k) + I_1(k)X(k) \\ &\geq V_1(k) + V_1(k)X(k) \\ &= 1 \cdot (1+X(k))V_1(k) \\ &> K(1+X(k))V_1(k) \\ &= I_1(k+1). \end{aligned}$$

This completes the proof of the theorem. $\square$

7.4 Performance: Buy and Hold Versus High-Frequency

In this section, in contrast to the results in Chapter 6, we show that trade execution delay can lead to better performance for a buy and holder versus that of the high-frequency trader. To this end, we provide examples involving the simple case of a binary lattice model for the stock returns. The rationale for this simplification in lieu of a general model is that the computations are easy to follow and are straightforward. In addition this model also has the property that as the time Δt between stages becomes small, one obtains an approximation of classical geometric Brownian motion; e.g., see [30, 103, 142, 143].

Before we provide our main example with $n = 100$ steps and with returns that are a somewhat reasonable facsimile of real-world trading, we first analyze a toy example with only three trades and unrealistic returns. For this simple case, it is easy to show mathematically, rather than by simulation, that execution delay in combination with the cash-financing requirement leads to $g_n^* > g_1^*$.

7.4.1 Toy Binary Lattice Example: We take $n = 3$ and use returns $X_{\max} = 0.8$ and $X_{\min} = -0.2$ with equal probability. Since n is small, a straightforward calculation allows one to obtain both $g_1(K)$ and $g_n(K)$ in closed-form. First restricting K to guarantee cash-financing, we find that

$$K_1^* = K_n^* = \frac{1}{1 + X_{\max}} \approx 0.556$$

with associated optimal expected logarithmic growths given by $g_1^* \approx 0.1009$ and $g_n^* \approx 0.1104$. Hence, the buy and holder outperforms the high-frequency trader by about 9.44%. In fact, even if this cash-financing constraint is removed and leverage is allowed, say by allowing $K \in [0, 1]$, then a straightforward calculation leads to optimal expected logarithmic growths $g_1^* \approx 0.1237$ and $g_n^* \approx 0.1262$. Hence, in this case, the buy and holder outperforms the high-frequency trader by about 2.11%. This shows that delay alone, rather than in combination with the cash-financing constraint, can lead to the buy and holder outperforming the high-frequency trader.

7.4.2 Example (More Realistic Binary Lattice Model): Our goal in this example is to argue that when execution delay is present, in real markets, it might also be the case that the buy and holder sometimes outperforms the high-frequency trader. Below, we now work with smaller returns, $n = 100$ and we consider a binary lattice model with returns $X_{\max} = 0.02$ with probability $p = 0.6$ and $X_{\min} = -0.01$ with probability $1 - p = 0.4$. When there is no delay, we recall the sufficient attractiveness inequality from Section 6.5 and note that for the more general binary lattice model parameterized in $X_{\min}$, $X_{\max}$ and p, the inequality reduces to

$$p \geq \frac{X_{\min}(1+X_{\max})}{X_{\min}-X_{\max}}.$$

For the lattice under consideration, the sufficient attractiveness condition in Section 6.5 reduces to the requirement that $p \geq 0.34$. Since the assumed value is $p = 0.6$, the requirement is therefore satisfied. Hence, the Sufficient Attractiveness Theorem 6.10.1 tells us that without delay, the optimal fractions are $K_1^* = K_n^* = 1$ and the optimal expected logarithmic growths satisfy $g_1^* = g_n^*$.

We now consider the effect of a unit delay in execution. According to the Cash-Financing Theorem in the previous section, we require

$$0 \leq K \leq \frac{1}{1+X_{\max}} \approx 0.9804.$$

Now, for the buy and holder, we use the closed-form solution in Section 7.2 to compute $g_n(K)$. Indeed, a lengthy but straightforward calculation leads to

$$g_n(K) = \frac{1}{n} p \sum_{i=0}^{n-1} p_i \log\left(1 + K(1+X_{\max})z_i\right) + \frac{1}{n}(1-p) \sum_{i=0}^{n-1} p_i \log\left(1 + K(1+X_{\min})z_i\right)$$

where

$$z_i \doteq (1+X_{\max})^i (1+X_{\min})^{n-1-i} - 1$$

and

$$p_i \doteq \binom{n-1}{i} p^i (1-p)^{n-1-i}.$$

Then, by plotting $g_n(K)$ versus K, we see in Figure 7.1 that the optimal fraction $K_n^* \approx 0.9804$ corresponds to the limit imposed by cash-financing. We also obtain the associated optimal expected logarithmic growth $g_n^* \approx 0.007719$; see the dash-dotted line in Figure 7.1.

On the other hand, for the high-frequency trader, since a closed-form for $g_1(K)$ is unavailable, we perform a Monte-Carlo simulation using 500,000 sample paths. In Figure 7.1, from the

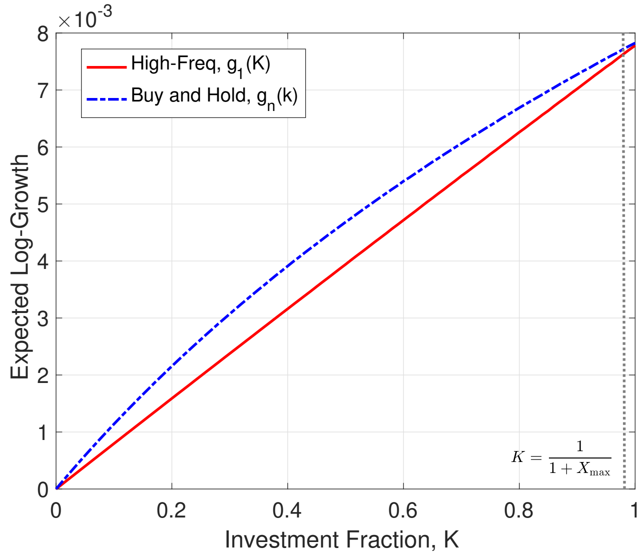


Figure 7.1: Expected Log-Growth with $n = 100$

plots of $g_1(K)$ and $g_n(K)$ versus K, we obtain $K_1^* = K_n^* \approx 0.9804$, which leads to the optimal expected logarithmic growth $g_1^* \approx 0.0076$. Recalling that $g_n^* \approx 0.007719$, the optimal expected logarithmic growth for the buy-and-hold strategy exceeds that of the high-frequency trading strategy by about 1.1%. The difference $g_n^* - g_1^* > 0$ is consistently observed when one carries out many repetitions of the simulation. It is also noted that, if one drops the cash-financing constraint, then the optimal fractions become $K_1^* = K_n^* = 1$, which corresponds to allowing leverage as we saw in the $n = 3$ case. That is, the optimal investments satisfy

$$
\begin{aligned}
I_i^*(k) &= K_i^*(1 + X(k))V_i(k) \\
&\leq (1 + X_{\max})V_i(k) = 1.02V_i(k)
\end{aligned}
$$

for $i \in \{1, n\}$. In this case, we obtain $g_1^* \approx 0.0077$ and $g_n^* \approx 0.007826$, which shows that the buy-and-hold strategy outperforms the high-frequency strategy by about 0.56%. This shows again that delay alone, rather than in combination with the cash-financing constraint can lead to $g_n^* > g_1^*$.

7.4.3 Binary Lattice with Variable Probability: We consider the binary lattice example above with the same parameters $X_{\max} = 0.02$, $X_{\min} = -0.01$ and $n = 100$, but now, we study performance parameterized in the probability p. Recalling the analysis in the previous subsection, when there is no delay, the trade is sufficiently attractive if $p > 0.34$. Thus, within this range of p, for the no-delay case, the buy and holder matches the performance of the high-frequency trader. However, when a unit execution delay is in play, via a lengthy computation, it is readily verified that the percentage difference $(g_n^* - g_1^*)/g_1^* \times 100\%$ versus p shows an increasing "margin of victory" for the buy and holder as p varies over its range. At $p = 0.35$, this difference is about 0.1% and as $p \to 1$, it reaches about 1.9%.

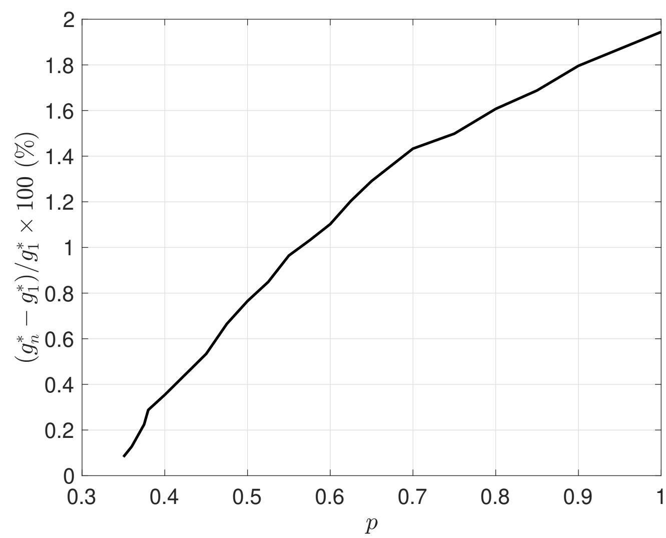


Figure 7.2: $(g_n^* - g_1^*)/g_1^* \times 100\%$ Versus p

7.4.4 Remark on Fractional Kelly Strategies: The optimum above for $p = 0.6$ in Section 7.4.2 requires that almost all funds be invested in the underlying stock. Since this might be viewed as far too aggressive for many traders, as previously discussed in Section 3.2 and references [12, 25, 45, 144], many authors suggest using a so-called *fractional* Kelly strategy. This is obtained by scaling down the fraction K so that the investment level is lower. We note that if one uses a fractional Kelly strategy for the binary lattice model above, as seen in Figure 7.1, the "margin of victory" for the buy and holder can be larger. In fact, $g_n(K) > g_1(K)$ for the entire open interval $0 < K < 1$.

7.5 Delay: Survival Analysis and All-Time Positivity

As discussed in the introduction to this chapter, the issue of *survival*, that is bankruptcy avoidance, arises when the delay is included in our model. Unlike the analysis in Chapter 6, this can occur even when cash-financing is required. With this consideration as the starting point, we now pursue this issue for the both high-frequency trader and the buy and holder. When $V(k) > 0$ for all k and all possible ample paths for the returns, we refer to this as *all-time positivity.* Related to our results to follow on all-time positivity are papers in the mathematics literature which deal with difference equations with multiple delays and provide conditions under which solutions are either eventually positive or eventually negative. That is, the state $V(k)$ has one sign for k suitably large; e.g., see [111] and [112] and their bibliographies. As noted in Section 7.9, conditions in the aforementioned literature under which eventual positivity and negativity fail can be viewed as a special case of Theorem 7.7.3 to follow which provides a necessary condition for all-time positivity.

For the buy and holder with $K \geq 0$, a necessary and sufficient survival condition is easy to establish. That is, the fact that $X_{\min} \leq X(k) \leq X_{\max}$ with $-1 < X_{\min} < 0 < X_{\max}$, noting that the account value dynamics of buy and holder satisfies

$$\begin{aligned} V_n(n) &= \left(1 + K(1 + X(0))\left(\prod_{k=1}^{n-1}(1 + X(k)) - 1\right)\right)V_n(0) \\ &\geq \left(1 + K(1 + X_{\max})\left(\prod_{k=1}^{n-1}(1 + X_{\min}) - 1\right)\right)V_n(0), \end{aligned}$$

it is readily verified that the account value $V_n(n) > 0$ for all n if and only if

$$K < \frac{-1}{(1 + X_{\max})\mathscr{R}_{n,\min}}$$

where

$$\mathscr{R}_{n,\min} \doteq \prod_{k=1}^{n-1}(1 + X_{\min}) - 1.$$

7.5.1 Survival Condition for High-Frequency Trader: For the high-frequency trader updating the investment at every step using $n = 1$, the analysis of survival is more complicated than that of the buy and holder using $n > 1$. Henceforth, for the sake of notational simplicity, throughout this section, we use $V(k)$ instead of $V_1(k)$ to denote the account value of high-frequency trader.

In the analysis to follow, we introduce two thresholds, K_- and K_+, depending on $X_{\min}$ and $X_{\max}$, and prove that for $K < K_-$, $V(k) > 0$ is guaranteed for all time and all asset-return sequences; i.e., bankruptcy is ruled out and positive solutions of the account-value recursion for $V(k)$ are continuable indefinitely. On the other hand, for $K > K_+$, we show that there is always a sequence of asset returns $X(k)$ for which $V(k)$ fails to be positive for all time; i.e., along this sequence, bankruptcy is certain and the recursion for $V(k)$ ceases to be meaningful after some finite number of steps. Finally, we provide a conjecture which says that for the "gap" interval

$$K_- \leq K \leq K_+,$$

positivity of $V(k)$ is also guaranteed for all time. Support for the conjecture, both theoretical and computational, is provided.

7.6 Survival as a State-Space Positivity Problem

We begin by taking $K \geq 0$ representing the specified percentage of a trader's account $V(k)$ to be invested in the shares of risky asset. Then, at stage k, in the presence of one unit of delay, instead of seeing $KV(k)$ as the value invested, it will be $K(1+X(k-1))V(k-1)$. To arrive at a state-space formulation, we view $V(k)$ as a state, K as a feedback gain and $u(k)$

as defining a feedback control which begins with $u(0) \doteq 0$, and for $k > 0$,

$$u(k) \doteq K(1 + X(k-1))V(k-1)$$

to account for the delay. Now the account value dynamics, viewed as a closed-loop state equation, is described by

$$\begin{aligned} V(k+1) &= V(k) + u(k)X(k) \\ &= V(k) + K(1 + X(k-1))V(k-1)\, X(k) \end{aligned}$$

with positive initial conditions $V(0) = V(1) = V_0 > 0$. In the sequel, a time-varying sequence of risky asset returns

$$X \doteq \{X(k)\}_{k=0}^{\infty}$$

is called a *path*, and is said to be *admissible* if, for all k, it satisfies the conditions stated many times in preceding chapters; i.e., $X_{\min} \leq X(k) \leq X_{\max}$ where $-1 < X_{\min} < 0 < X_{\max} < \infty$. We take $\mathscr{X}$ to be the set of all admissible paths and often emphasize the state dependence on $X \in \mathscr{X}$ by writing $V(X, k)$ instead of $V(k)$. Additionally, we take $\mathscr{X}^N$ to be the set of all $X = (X(0), X(1), \ldots, X(N-1))$ such that for $k = 0, 1, \ldots, N-1$, $X(k)$ stays within the known bounds above. Elements of $\mathscr{X}^N$ are called *admissible partial paths*, or simply admissible paths when there is no confusion.

As is standard in control theory, we now eliminate the delay term in the state equation above and work with a two-state system. That is, defining the state vector

$$x(k) \doteq [V(k) \;\; V(k-1)]^T,$$

we obtain the linear time-varying state-space system

$$x(k+1) = A(X, k)\, x(k)$$

where

$$A(X, k) \doteq \begin{bmatrix} 1 & K(1 + X(k-1))X(k) \\ 1 & 0 \end{bmatrix}.$$

As mentioned previously, we work with the specific initial conditions $V(0) = V(1) = V_0 > 0$. Although our goal, state positivity, is the same as in existing positive system theory, for example, see [107–110, 145, 146], this body of work is not in play because the matrix $A(X, k)$ can have a negative entry.

7.7 All-Time Positivity Results

Although the solution to the state equation above exists for all k, since bankruptcy precludes future trading, the analysis ceases to be meaningful if $V(X, k) \leq 0$. With this as motivation, the focal point in this section is the issue of all-time positivity. As mentioned earlier in the chapter, this is the issue of existence and continuability of positive solutions for infinitely many steps. Indeed, for a given feedback gain $K \geq 0$, we say that the *all-time positivity* condition holds if $V(X, k) > 0$ for all $X \in \mathscr{X}$ and all $k \geq 0$.

The results to follow involve two critical thresholds, K_- and K_+. The first of these is motivated by considering $k = 2$ and noting that

$$\begin{aligned} V(2) &= V(1) + K(1 + X(0))X(1)\,V(0) \\ &\geq [1 + K(1 + X(0))X_{\min}]V_0 \\ &\geq [1 + K(1 + X_{\max})X_{\min}]V_0. \end{aligned}$$

This lower bound is positive if and only if

$$K < \frac{1}{|X_{\min}|(1 + X_{\max})}.$$

To show $V(k) > 0$ for all k rather than just $k = 2$, it requires the stronger assumption that $K < K_-$, where

$$K_- \doteq \frac{1}{1 + X_{\max}}.$$

Since the proof of the theorem below requires lengthy arguments, it is relegated to Section 7.9.

Theorem 7.7.1 (Sufficient Condition for All-Time Positivity): *The condition*

$$0 \leq K < K_-$$

is sufficient for all-time positivity. That is, if $K < K_-$, given any admissible path $X \in \mathscr{X}$, it follows that $V(X,k) > 0$ for all k.

7.7.2 Necessary Condition for All-Time Positivity: Our necessary condition for all-time positivity is motivated by studying the state equation in response to a *distinguished path* of returns X^*. This path is defined by $X^*(0) = X_{\max}$ and $X^*(k) = X_{\min}$ for $k \geq 1$. Along this path, since the first trade is executed at stage $k = 1$, the return $X(0) = X_{\max}$ can be viewed as "baiting" the trader with a large positive return and then, a worst-case scenario of sorts occurs because the account loses value on every subsequent trade. To motivate the definition of the threshold K_+ entering into our analysis of necessity, we take $X = X^*$. Let

$$K_s \doteq \frac{1}{4|X_{\min}|(1+X_{\min})}.$$

Then with $K > K_s$, consistent with the fact that the matrix

$$A(X^*,k) = \begin{bmatrix} 1 & K(1+X_{\min})X_{\min} \\ 1 & 0 \end{bmatrix}$$

has a pair of complex eigenvalues, the state $V(X^*,k)$ is oscillatory about zero. Hence, the value of state can be negative. This becomes a special case of the theorem to follow. For $K \leq K_s$, the solution is nonoscillatory, and our analysis shows that the state is negative for large k when $X_{\max} > 1 + 2X_{\min}$ and $K > K^\star$, where

$$K^\star \doteq \frac{X_{\max} - X_{\min}}{|X_{\min}|(1+X_{\max})^2}.$$

The theorem below, whose lengthy proof is also relegated to Section 7.9, brings these ideas to fruition. The threshold K_+ defined below is readily verified to exceed K_-.

Theorem 7.7.3 (Necessary Condition for All-Time Positivity): *With*

$$K_+ \doteq \begin{cases} K^\star & \text{if } X_{\max} > 1 + 2X_{\min}; \\ K_s & \text{if } X_{\max} \leq 1 + 2X_{\min}, \end{cases}$$

the condition $K \leq K_+$ is necessary for all-time positivity. Equivalently, if $K > K_+$, then there exists an admissible path $X \in \mathscr{X}$ such that $V(X, k) \leq 0$ for some k.

7.7.4 Graphical Depiction of Bounds: In Figure 7.3, the dependencies of K_- and K_+ on $X_{\min}$ and $X_{\max}$ are displayed over the range $-1 < X_{\min} < 0 < X_{\max} = 2$. The black lower surface is obtained by using the formula for K_-, the red part of the upper surface is obtained by using the formula for $K_+ = K_s$, and the larger green part of the upper surface is obtained using the formula for $K_+ = K^\star$.

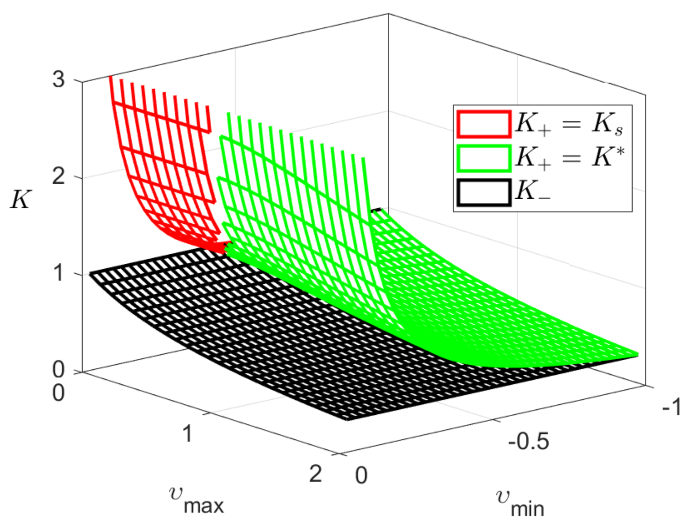


Figure 7.3: Two Critical Thresholds K_- and K_+

Figure 7.3 can be used to better understand which triples

$$(K, X_{\min}, X_{\max}) \in [0, \infty) \times (-1, 0) \times (0, \infty)$$

lead to all-time positivity. If the triple falls below the surface given by K_-, then, according to the Sufficiency Theorem, all-time positivity holds. Alternatively, if the triple lies above the surface given by K_+, then, according to the Necessity Theorem, all-time positivity fails. Finally, when the triple lies between the two surfaces, we conjecture later in the chapter that all-time positivity still holds.

7.8 Preliminary Technical Results

This section provides the technical lemmas underlying the main results, Theorems 7.7.1 and 7.7.3 above. The proofs for these lemmas are also relegated to Section 7.9. In the first two lemmas below, we focus on the behavior of the state $V(X^*, k)$ along the path X^*.

Lemma 7.8.1 (Closed-Form for $V(X^*, k)$): *If $K \neq K_s$, for $k \geq 2$, the state along the distinguished path X^* is given by*

$$V(X^*, k) = \frac{V_0}{2\sqrt{\theta}}\left(\lambda_+^{k-1} g_+ + \lambda_-^{k-1} g_-\right)$$

where

$$\theta \doteq 4KX_{\min}(1 + v_{\min}) + 1,$$

$$g_\pm \doteq \sqrt{\theta} \pm (2K(v_{\max} + 1)v_{\min} + 1)$$

and

$$\lambda_\pm \doteq \frac{1}{2}\left(1 \pm \sqrt{\theta}\right)$$

are the eigenvalues of $A(X^*, k)$. For the singular case, $K = K_s$, the state solution is given by

$$V(X^*, k) = \frac{2^{-k} V_0((1 - X_{\max} + 2X_{\min})k + 1 + X_{\max})}{1 + X_{\min}}.$$

Lemma 7.8.2 (Distinguished Path Properties):

(a) If $K > K_s$, then $V(X^*, k)$ is oscillatory about zero and is therefore negative for some k.

(b) If $K^\star < K \leq K_s$ and $X_{\max} > 1 + 2X_{\min}$, then $V(X^*, k)$ is negative for sufficiently large k.

(c) If $K^\star < K \leq K_s$ and $X_{\max} < 1 + 2X_{\min}$, then $V(X^*, k)$ is positive for all $k \geq 0$.

(d) If $0 \leq K \leq K^\star$, then $V(X^*, k)$ is positive for all $k \geq 0$.

7.8.3 Remarks on State Along Distinguished Path X^*: In its own right, it is interesting to study the asymptotic behavior of $V(X^*, k)$, since its tending to zero signifies "practical bankruptcy," even for cases when all-time positivity is assured. For the case $0 < K < K_s$ in Lemma 7.8.1, it is clear that $0 < \theta < 1$, which implies $|\lambda_\pm| < 1$. Thus, by the well-known unit-circle stability criterion, for example see [147], $V(X^*, k) \to 0$ as $k \to \infty$. The closed-form solution for the singular case $K = K_s$ also tends to zero, and it is also readily verified, using l'Hôpital's rule, that the closed-form expression of the state solution $V(X^*, k)$ is continuous at $K = K_s$.

7.9 Proofs of Preliminary Results and Two Theorems

This section may be skipped by readers who are not interested in the technical details of the proofs of the two theorems and preliminary lemmas in Sections 7.7 and 7.8.

Proof of Sufficiency Theorem 7.7.1. Since the case $K = 0$ is trivial, we assume $0 < K < K_-$ and note that it suffices to prove all-time positivity with returns $X(k)$ allowed to range over the larger interval $-1 \leq X(k) \leq X_{\max}$. We proceed by induction on k. First recall that $K < K_-$ was shown to guarantee positivity of $V(2)$ in Section 7.7. Next, for $k \geq 2$, we assume $V(i) > 0$ for $i = 0, 1, \ldots, k$ and all $X(0), \ldots, X(k-1)$. Then for arbitrary $X(0), X(1), \ldots, X(k)$, we must show $V(k+1) > 0$. Indeed, noting $1 + X(k-1) \geq 0$ and

$V(k-1) > 0$ by the induction hypothesis, we obtain lower bound

$$V(k+1) = V(k) + K(1 + X(k-1))X(k)V(k-1)$$
$$\geq V(k) - K(1 + X(k-1))V(k-1).$$

To further lower bound the right hand side above, for $-1 \leq w \leq X_{\max}$, let $V_w(k)$ be the value of $V(k)$ with $X(k-1)$ replaced by w. With this notation, we can write

$$V(k+1) \geq \min_{w} \left\{ V_w(k) - K(1+w)V(k-1) \right\}.$$

Since the function to be minimized on the right-hand side above is affine linear in w, its minimum value is achieved by $w = -1$ or $w = X_{\max}$. We now analyze what happens to the minimum in each of case.

For $w = -1$, the preceding lower bound of $V(k+1)$ leads to

$$V(k+1) \geq V_{-1}(k) > 0$$

by the induction hypothesis. For $w = X_{\max}$, we obtain

$$V(k+1) \geq V_{X_{\max}}(k) - K(1 + X_{\max})V(k-1).$$

Since $V_{X_{\max}}(k) = V(k-1) + K(1+X(k-2))X_{\max}V(k-2)$, using the facts that $1+X(k-2) \geq 0$, $K > 0$, and $V(k-2)$ is positive by the induction hypothesis, it follows that

$$V_{X_{\max}}(k) \geq V(k-1) > 0$$

where last inequality holds by induction hypothesis again. Hence, the $V(k+1)$ is further lower bounded as

$$V(k+1) \geq [1 - K(1 + X_{\max})]V_{X_{\max}}(k).$$

Now, applying the assumed inequality $0 < K < K_{-}$ and the fact that $V_{X_{\max}}(k) > 0$ by the induction hypothesis, we obtain $V(k+1) > 0$. $\square$

Proof of Lemma 7.8.1. Recall the state space representation introduced in Section 7.1. Using the standard state augmentation

$$x(k) \doteq [V(k) \;\; V(k-1)]^T,$$

we obtain the linear time-varying system

$$x(k+1) = A(X,k)\, x(k)$$

where $A(X,k)$ is the 2×2 matrix defined in Section 7.1. Starting from initial conditions $V(X^*,0) = V(X^*,1) = V_0$ and $V(X^*,2) = (1 + K(1+X_{\max})X_{\min})V_0$, in state-space form, we have for $k \geq 2$,

$$\begin{bmatrix} V(X^*,k+1) \\ V(X^*,k) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & K(1+X_{\min})X_{\min} \\ 1 & 0 \end{bmatrix}}_{=A(X^*,k)} \begin{bmatrix} V(X^*,k) \\ V(X^*,k-1) \end{bmatrix},$$

and we obtain

$$V(X^*,k) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & K(1+X_{\min})X_{\min} \\ 1 & 0 \end{bmatrix}^{k-1} \begin{bmatrix} V(X^*,2) \\ V(X^*,1) \end{bmatrix}.$$

We consider two cases: For the generic case, $K \neq K_s$, a lengthy but straightforward computation leads to

$$\begin{aligned} V(X^*,k) &= \frac{2^{-k} V_0 \left(\left(1+\sqrt{\theta}\right)^{k-1} g_+ + \left(1-\sqrt{\theta}\right)^{k-1} g_- \right)}{\sqrt{\theta}} \\ &= \frac{V_0}{2\sqrt{\theta}} \left(\lambda_+^{k-1} g_+ + \lambda_-^{k-1} g_- \right). \end{aligned}$$

Another lengthy but straightforward computation shows that $\lambda_\pm$ are the eigenvalues of $A(X^*,k)$. For the singular case, $K = K_s$, we find that

$$V(X^*,k) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/4 \\ 1 & 0 \end{bmatrix}^{k-1} \begin{bmatrix} V(X^*,2) \\ V(X^*,1) \end{bmatrix}$$

which, again, following a third lengthy but straightforward computation, results in

$$V(X^*,k) = \frac{2^{-k}V_0(k(1-X_{\max}+2X_{\min})+1+X_{\max})}{1+X_{\min}}. \quad \square$$

Proof of Lemma 7.8.2. A proof of part (a) that does not use the closed-form of $V(X^*,k)$ can be given immediately by applying Theorem 2.2 in [111]. However, for the sake of self-containment, we provide a first-principles proof here. Assuming that $K > K_s$, we must show the state $V(X^*,k)$ is oscillatory about zero and is negative for some values of k. By Lemma 7.8.1, we have

$$V(X^*,k) = \frac{V_0}{2\sqrt{\theta}}\left(\lambda_+^{k-1}g_+ + \lambda_-^{k-1}g_-\right)$$

for $k \geq 2$. With $K > K_s$, it is readily shown that $\theta < 0$, which implies that the two eigenvalues $\lambda_\pm$ are complex conjugates. It follows that these eigenvalues can be written in polar form as $\lambda_+ = re^{j\omega}$ and $\lambda_- = re^{-j\omega}$, where $r = |\lambda_\pm| > 0$ and

$$\omega = \tan^{-1}(\sqrt{|\theta|}) \in \left(0, \frac{\pi}{2}\right).$$

Next, substituting the polar form of $\lambda_\pm$ into $V(X^*,k)$ above, a lengthy but straightforward calculation shows that

$$V(X^*,k) = Br^{k-1}\cos((k-1)\omega+\varphi)$$

where B and φ are constants, with $B > 0$. Since $\omega \in (0,\pi/2)$, it is straightforward to find a value of k such that the argument of the cosine lies in $(\pi/2, 3\pi/2)$, thus making the cosine negative. This completes the proof of part (a).

To prove part (b), we first consider the case $K = K_s$. Then using the formula

$$V(X^*,k) = \frac{2^{-k}V_0(k(1-X_{\max}+2X_{\min})+1+X_{\max})}{1+X_{\min}}$$

for the singular case in Lemma 7.8.1, for $X_{\max} > 1+2X_{\min}$ and k sufficiently large,

$$V(X^*,k) < 0.$$

Next, for the case $K^\star < K < K_s$, we assume again $X_{\max} > 1 + 2X_{\min}$. Since $\lambda_+ > \lambda_-$, the state $V(X^*, k)$ will be negative for sufficiently large k if we can show that $g_+ = \sqrt{\theta} + q < 0$ where

$$q \doteq 2K(X_{\max} + 1)X_{\min} + 1.$$

To establish this, since $K \in (K^\star, K_s)$, we have

$$0 < \theta < 4K^\star X_{\min}(1 + X_{\min}) + 1 = \frac{(X_{\max} - 2X_{\min} - 1)^2}{(1 + X_{\max})^2}.$$

Since the square root is an increasing function, the inequality on θ above implies that

$$\sqrt{\theta} < \frac{X_{\max} - 2X_{\min} - 1}{1 + X_{\max}}.$$

In addition, we also have

$$q < 2K^\star(1 + X_{\max})X_{\min} + 1 = \frac{1 - X_{\max} + 2X_{\min}}{1 + X_{\max}}.$$

Thus, it follows that

$$\begin{aligned} g_+ &= \sqrt{\theta} + q \\ &< \frac{X_{\max} - 2X_{\min} - 1}{1 + X_{\max}} + \frac{1 - X_{\max} + 2X_{\min}}{1 + X_{\max}} = 0. \end{aligned}$$

Hence, the proof of part (b) is complete.

To prove part (c), we first note that the desired positivity holds trivially for $k = 0, 1$. For $k \geq 2$, assuming that $K = K_s$ and $X_{\max} < 1 + 2X_{\min}$, the singular case formula given in Lemma 7.8.1 leads that

$$V(X^*, k) > \frac{2^{-k}V_0(1 + X_{\max})}{1 + X_{\min}}$$

which is positive for all $k \geq 2$ because $X_{\min} > -1$, $V_0 > 0$ and $X_{\max} > 0$. It remains to treat the case $K^\star < K < K_s$ and $X_{\max} < 1 + 2X_{\min}$. To show $V(X^*, k) > 0$ for all $k \geq 2$, substitute $g_\pm = \sqrt{\theta} \pm q$ and $\lambda_\pm = (1 \pm \sqrt{\theta})/2$ into $V(X^*, k)$ and note that $\theta \in (0, 1)$. Then the formula for $X(X^*, k)$ reduces to

$$V(X^*, k) = \frac{V_0}{2^k\sqrt{\theta}}\left[\sqrt{\theta}\left(\left(1 + \sqrt{\theta}\right)^{k-1} + \left(1 - \sqrt{\theta}\right)^{k-1}\right) + q\left(\left(1 + \sqrt{\theta}\right)^{k-1} - \left(1 - \sqrt{\theta}\right)^{k-1}\right)\right].$$

Since $X_{\max} < 1 + 2X_{\min}$ and $X_{\min} > -1$, we obtain

$$q \geq 2K_s(v_{\max} + 1)v_{\min} + 1 = \frac{1 - X_{\max} + 2X_{\min}}{2(1 + X_{\min})} > 0.$$

Since $\sqrt{\theta} > 0$, $q > 0$ and

$$(1 + \sqrt{\theta})^{k-1} > (1 - \sqrt{\theta})^{k-1}$$

for all $k \geq 2$, it follows that $V(X^*, k) > 0$. This completes the proof of part (c).

Finally, to prove part (d), since the result trivially follows for $K = 0$, we assume $K > 0$. Note that the inequality $K_s \geq K^\star$ is readily shown to be equivalent to

$$\big((1 + X_{\max}) - 2(1 + X_{\min})\big)^2 \geq 0.$$

Furthermore the above inequalities are both strict if and only if $X_{\max} \neq 1 + 2X_{\min}$. Suppose $X_{\max} \neq 1 + 2X_{\min}$. Then $0 < K \leq K^\star$ implies $K < K_s$, and so in Lemma 7.8.1, we have $0 < \theta < 1$ and $\lambda_\pm > 0$. It suffices to prove that $g_\pm \geq 0$ and that one of g_+ or g_- is strictly positive. In the formula for $g_\pm = \sqrt{\theta} \pm q$, the quantity $q = 1 + 2K(1 + X_{\max})X_{\min}$ is either negative or nonnegative. If it is nonnegative, then $g_+ > 0$, and $g_- \geq 0$ on account of the fact that $K \leq K^\star$ is equivalent to

$$\theta \geq [1 + 2K(1 + X_{\max})X_{\min}]^2.$$

Similarly, if the quantity q above is negative, then $g_- > 0$, while $g_+ \geq 0$ on account of the fact that $K \leq K^\star$ again.

Suppose $X_{\max} = 1 + 2X_{\min}$. Then for the case $K = K^\star = K_s$, the state $V(X^*, k)$ for this singular case given in Lemma 7.8.1 applies and is clearly positive for all k. Alternatively, for the case $0 < K < K^\star = K_s$, we argue as in the preceding paragraph and obtain $0 < \theta < 1$, $\lambda_\pm > 0$. Moreover, since $X_{\max} = 1 + 2X_{\min}$, we have $q = \theta$, which leads to

$$g_\pm = \sqrt{\theta} \pm \theta > 0.$$

This completes the proof of part (d). $\square$

Proof of Necessity Theorem 7.7.3. Given $K > K_+$, it suffices to exhibit a path v for which the state $V(X,k)$ is not positive for some k. We claim that the distinguished path X^* is such a path. To establish this, we split our analysis into two cases:

Case 1: For $X_{\max} \leq 1 + 2X_{\min}$, we have $K_+ = K_s$. Thus, it suffices to prove $V(X^*,k) < 0$ for some k when $K > K_s$. Using part (a) of Lemma 7.8.2, we obtain that the state $V(X^*,k)$ oscillates and takes negative values for some k.

Case 2: For $X_{\max} > 1 + 2X_{\min}$, we have $K_+ = K^\star$. Note that if $K > K_s$, the negativity of $V(X^*,k)$ is again established by part (a) of Lemma 7.8.2. Thus, it suffices to prove $V(X^*,k) < 0$ for some k when $K^\star < K \leq K_s$. Since $X_{\max} > 1 + 2X_{\min}$, using part (b) of Lemma 7.8.2, we obtain that the state $V(X^*,k)$ is negative for all sufficiently large k. Hence, the proof is complete. $\square$

7.10 All-Time Positivity Conjecture and Support

The conjecture to follow addresses the "gap" between the lower and upper bounds, K_- and K_+, for all-time positivity provided by the theorems in Section 7.7. Subsequently, we support the conjecture with analysis and simulations for various cases involving a finite time horizon. As seen below, the notion of "extreme paths" plays an important role.

Conjecture 7.10.1 (All-Time Positivity): *The all-time positivity condition, $V(k) > 0$ for all k, holds for the gap interval $K_- \leq K \leq K_+$.*

7.10.2 Extreme Paths: To study the conjecture, for given $N \geq 0$, we consider the 2^N *extreme paths* $X^i \in \mathscr{X}^N$, defined by $X^i(k)$ being either $X_{\min}$ or $X_{\max}$ for $k = 0, 1, \ldots, N-1$. For example, $(X_{\min}, X_{\max}, X_{\min})$ is an extreme path in $\mathscr{X}^3$. First noting that the positivity condition $V(X,k) > 0$ for all $k \le N$ and all $X \in \mathscr{X}^N$ is equivalent to

$$\min_{v \in \mathscr{X}^N} V(X,k) > 0$$

for $k \le N$, we make use of the fact that $V(X,k)$ is *multilinear* in v; i.e., affine linear in each component $X(k)$. For example,

$$V(v,3) = [1 + X(2) + X(1)X(2) + K\,(X(1) + X(0)X(1))]\,V_0$$

is multilinear in $X(0), X(1)$ and $X(2)$. We now use the well-known fact that the minimum of a multilinear function over a hypercube is attained at one of the vertices; e.g., see [148]. This implies that $V(X,k)$ is minimized by one of the extreme paths X^i. Hence, $V(X,k)$ is positive for all $v \in \mathscr{X}^N$ and all $k \le N$ if and only if

$$\min_{i \in \{1,2,\ldots,2^N\}} V(X^i,k) > 0$$

for $k = 0,1,2,\ldots,N$. For small N, checking this condition is feasible, but for large N, the number of "checks," namely 2^N, becomes too large. For example, in the stock market, we can easily have $N = 100$, but it is computationally prohibitive to check 2^{100} extreme paths.

7.10.3 Examples for Various N: Taking $V_0 = 1$, $X_{\max} = 0.9$, and $X_{\min} = -0.8$, we then have $X_{\max} > 1 + 2X_{\min}$, and the gap interval is computed to be

$$[K_-,\, K_+] \approx [0.5263,\, 0.5888].$$

To support the conjecture, we took $N = 10$ and chose $n = 100$ equally-spaced values of K from the gap interval and, for each K, we used Matlab to check state positivity of each of the $2^N = 1024$ extreme paths. We found that state positivity held for all of them. In Figure 7.4, the $V(X^i,k)$ are shown for $K = 0.54$, which lies within the gap interval above. We also ran many other simulations for various choices of $X_{\min}$, $X_{\max}$ and $N \le 15$, and consistently observed that state positivity held in the corresponding gap interval.

Given the motivation for the distinguished path X^* in terms of a "worst-case" trading scenario in Section 7.7, it is natural to ask if $V(X^*, k)$ might be the minimum value of $V(X, k)$ for all $k \leq N$. However, as seen in Figure 7.4, this proves not to be the case for $7 \leq k \leq 10$.

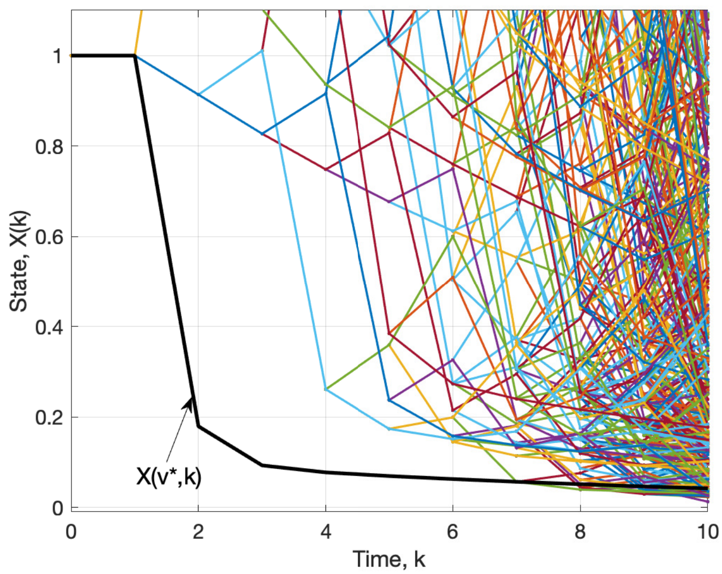


Figure 7.4: Simulation Supporting the Conjecture for $K = 0.54$

To provide further support for the conjecture, we also studied $N = 100$, $V_0 = 1$, $X_{\max} = 0.2$ and $X_{\min} = -0.3$. For $n = 100$ equally spaced values of K in the gap interval

$$[K_-, K_+] \approx [0.8333, 1.1905],$$

we generated $200,000$ of the 2^{100} extreme paths for each K. The positivity condition was seen to be satisfied in all cases. Finally, in support of the conjecture, we also ran other simulations for various choices of $X_{\min}$ and $X_{\max}$, including smaller values of these bounds to more closely model values found in stock trading, and consistently observed that the desired state positivity held within the gap interval.

7.10.4 Theoretical Result for $N \leq 3$: In this subsection and the lemma below, we prove that if $K \leq K_+$, then state positivity holds for all partial paths of length $N \leq 3$. We begin by noting that the cases $N = 0$ and $N = 1$ are immediate since $V(0) = V(1) = V_0 > 0$ are the initial conditions. Next, for $N = 2$, as shown in the beginning of Section 7.7,

$$V(X,2) \geq [1 + K(1 + X_{\max})X_{\min}]V_0.$$

Thus, $V(v,2) > 0$ if and only if

$$K < K_{\max}(2) \doteq \frac{1}{|X_{\min}|(1 + X_{\max})}.$$

Since it is also easily verified that $K_+ < K_{\max}(2)$, it follows that $V(X,k) > 0$ for $X \in \mathscr{X}^2$ and $k \leq 2$ when $K \leq K_+$. The case $N = 3$, per lemma below, requires a lengthier derivation to show that $V(v,3) > 0$ if and only if $K < K_{\max}(3)$ where

$$K_{\max}(3) \doteq \frac{1}{|X_{\min}|(2 + X_{\max} + X_{\min})}.$$

Then a straightforward calculation shows that $K_+ < K_{\max}(3)$.

Lemma 7.10.5 (Finite-Time Positivity for $N = 3$): *If $K < K_{\max}(3)$, then $V(X,k) > 0$ for all $X \in \mathscr{X}^3$ and all $k \leq 3$.*

Proof. For any $(X(0), X(1))$, we observe that

$$V(X,3) = V(X,2) + K(1 + X(1))X(2)V_0$$

is minimized with $X(2) = X_{\min}$. It follows that

$$\begin{aligned} V(X,3) &\geq V(X,2) + K(1 + X(1))X_{\min}V_0 \\ &= [1 + K\left((1 + X_{\min} + X(0))X(1) + X_{\min}\right)]\, V_0. \end{aligned}$$

Since the right-hand side is multilinear in $X(0)$ and $X(1)$, the minimum must occur when they take the values $X_{\min}$ or $X_{\max}$. If $X(1) = X_{\max}$, then to minimize the right-hand side, $X(0)$ must be $X_{\min}$. In this case, using a lower bound for the right-hand side above, we obtain

$$V(X,3) \geq [1 + K\,(\,(1 + X_{\min} + X_{\min})X_{\max} + X_{\min})]\,V_0.$$

Similarly, if $X(1) = X_{\min}$, then $X(0)$ must be $X_{\max}$, which leads to lower bounds

$$V(X,3) \geq [1 + K\,(\,(1 + X_{\min} + X_{\max})X_{\min} + X_{\min})]\,V_0.$$

It is easy to check that this second bound is strictly smaller than the first. Furthermore, if $K < K_{\max}(3)$, this second lower bound is positive which implies $V(X,3) > 0$ as required. $\square$

7.10.6 Finite-Time Positivity Set: We let $\mathscr{K}(N)$ denote the set of all feedback gains K assuring state positivity up to stage N and define

$$K_{\max}(N) \doteq \sup\{K \geq 0 : [0, K) \subseteq \mathscr{K}(N)\}.$$

Then we have already seen above that $\mathscr{K}(2) = [0,\, K_{\max}(2))$ and $\mathscr{K}(3) = [0,\, K_{\max}(3))$ with $K_{\max}(3) < K_{\max}(2)$ readily verified. Beyond these two simple cases, one can in principle determine whether or not a given feedback parameter K belongs to $\mathscr{K}(N)$ by checking all extreme paths. We also know, by the Sufficiency Theorem 7.7.1, that

$$[0,\, K_{-}) \subseteq \mathscr{K}(N).$$

If the All-Time Positivity Conjecture is true, we must have $[0,\, K_{+}] \subseteq \mathscr{K}(N)$ as well. Moreover, since $\mathscr{K}(N+1) \subseteq \mathscr{K}(N)$ for all N, the $K_{\max}(N)$ are nonincreasing, and since they are bounded below by K_{-}, they converge to a limit

$$K_{\infty} \doteq \lim_{N \to \infty} K_{\max}(N).$$

It is also readily verified that $K_{\infty} \leq K_{+}$; otherwise, there would exist an $K \in (K_{+}, K_{\infty})$ assuring all-time positivity, which contradicts the Necessity Theorem 7.7.3. Finally, if the All-Time Positivity Conjecture is true, then $K_{\infty} \geq K_{+}$, in which case it would follow that $K_{\infty} = K_{+}$.

7.11 Concluding Remarks and Future Work

In this chapter, we extended our frequency-based formulation to involve delay in execution. We then provided conditions under which cash-financing is assured and showed that when delay is present, the buy and holder may outperform high-frequency trader in terms of expected logarithmic growth. However, the result obtained is based on the binary lattice model with returns $X(k)$ which are somewhat unrealistic in that they are larger than those typically seen seen with real "high-frequency" trading data. Thus, it is important to conduct experiments and see if the buy and holder can outperform the high-frequency trader using real-world historical data. This issue is relegated to future research.

Subsequently, we considered a state positivity problem aimed at obtaining conditions under which survival is guaranteed; i.e., bankruptcy is avoided. To this end, we characterized survival in terms of two critical thresholds, K_- and K_+ with $K_- < K_+$. Finally, we conjectured that state positivity is guaranteed for the "gap" interval

$$K_- \leq K \leq K_+.$$

Support for this conjecture, both theoretical and computational, was also provided.

Regarding further research, we mention another two attractive directions: The first is obviously to pursue a proof of the conjecture in Section 7.10. The second direction, motivated by the multi-asset portfolio case, would be to study the state positivity problem when $X(k)$ is vector-valued rather than a scalar. That is, if $X(k) \in \mathbb{R}^m$ with $X_i(k)$ being the i-th component satisfying

$$X_{\min,i} \leq X_i(k) \leq X_{\max,i}$$

with $-1 < X_{\min,i} < 0 < X_{\max,i}$ for $i = 1, 2, \ldots, m$, then, when an execution delay is present, the more general state equation

$$V(k+1) = V(k) + \sum_{i=1}^{m} K_i(1 + X_i(k-1))X_i(k)V(k-1)$$

arises where the $K_i \geq 0$ are scalar constant feedback gains. In this case, generalization of the theory in this chapter would be of interest. To this end, one result along these lines is that the condition

$$4\sum_{i=1}^{m} K_i(1 + X_{\min,i})|X_{\min,i}| > 1,$$

leads to oscillation and failure of all-time positivity. This can be established using arguments similar to those given in the proof of Lemma 7.8.2 and the related literature.

Chapter 8

Conclusion and Future Research

The main objective in this dissertation has been to extend and generalize various aspects of the theory of Kelly Betting with an emphasis on applications to the stock market. In many places in the dissertation, we adopted a control-theoretic point of view. The topics which we covered include issues in classical gambling theory, risk management, performance comparisons between high and low-frequency trading and the effects of delay in trade execution. At the end of each chapter, some directions for future research were described in detail. We now describe some of the end-of-chapter highlights.

8.0.1 Data-Driven Kelly Betting: In classical Kelly Betting theory, the optimum is selected to maximize the expected logarithmic growth with an assumption that the probability distributions for the returns are well known. In practice, it is typically the case that this distribution is empirical, based on historical data and perhaps highly unreliable for prediction purposes. With this consideration in mind, in Chapter 2, we briefly described a new research direction called data-driven Kelly Betting. In this setting, many research issues arise. For example, when deriving empirical distributions, non-stationarity of the underlying stock-price process motivates the study of sample-size issues. Simply put, data used for empirical distribution which goes too far back in time may not be relevant going forward.

8.0.2 Two Convexity Conjectures and Its Ramifications: In Chapter 3, we provided two conjectures involving convexity of sets associated with drawdown. A modification of the expected logarithmic growth optimization problem to include convex constraints limiting K is another attractive direction for future research. In addition, we introduced a surrogate drawdown measure which enables us to formulate an approximate drawdown-constrained

Kelly-based optimization problem as a concave program. A closer look at the usefulness of the surrogate is another interesting direction to pursue.

8.0.3 Drawdown With Time-Varying Feedback: In Chapter 4, we mentioned the possibility of extending our theory of drawdown-modulated control to the case of time-varying feedback $\gamma(k)$ rather than constant feedback. That is, one can take

$$I(k) = \gamma(k)M(k)V(k)$$

with $\gamma(k) \in \Gamma$ and still satisfy the drawdown requirement. Given that controllers with constant feedback gain defines a subset of the class of modulators above, we envision superior performance being achieved being possible to achieve.

8.0.4 Drawdown-Based Portfolio Design: Using expected return and drawdown as return-risk pair, a new drawdown-based portfolio optimization problem was formulated in Chapter 4 as being attractive to pursue. Recalling the formulation, given any target level of expected return, call it $\widehat{R}$, an admissible drawdown-modulated feedback $\mathcal{M} \in \Gamma \times [0,1]$ is sought to minimize $\overline{d}_{\mathcal{M}}$ subject to the constraint $\overline{R}_{\mathcal{M}} = \widehat{R}$. That is, we seek

$$\inf_{\mathcal{M} \in \Gamma \times [0,1]} \overline{d}_{\mathcal{M}}$$
$$\text{subject to } \overline{R}_{\mathcal{M}} = \widehat{R}.$$

8.0.5 Efficiency Issues: With the expected return and drawdown as the return-risk pair, the Efficiency Lemma 5.3.1 in Chapter 5 tells us that for a given linear time-invariant feedback controller, we can find drawdown-modulated feedback strategy leading to the same expected drawdown and possibly higher expected return. However, the lemma does not guarantee domination. Hence, an interesting extension of this work would be to provide conditions under which returns for the modulator can be guaranteed to be strictly higher.

8.0.6 Frequency and Delay Considerations: In Chapter 6 , we shifted our focus to a new class of Kelly Betting and stock trading problems involving frequency and the delay issues. For the frequency-based stock-trading scenario, if $\mathbb{E}[X(0)] \leq 0$, the arguments used for High-Frequency Maximality Theorem 6.9.1 fail to generalize to the case of short selling. To this end, we provided a conjecture for this case which would be interesting to pursue. Additionally, as mentioned in the conclusion, when betting on the $X < 0$ side with $X_{\max} > 1$, working with the compound returns $\mathcal{X}_n$ becomes problematic. Hence, another direction to pursue is to extend our frequency-based formulation for betting games which allow $X_{\max} > 1$. Finally, in Chapter 7, motivated by the fact that a trader's interactions with the market are not instantaneous, extension of our single time delay theory to *multiple* time delays would be an important topic to study.

8.0.7 Survival Issues: In Chapter 7, we studied the issue of survival when delay in trade execution is included in our frequency-based framework; i.e., bankruptcy is a possibility. Based on this, we formulated a state positivity problem. Our theory for positivity of the state, namely the account value in our case, involved two critical thresholds, K_- and K_+ with $K_- < K_+$ with a "gap" in between. We proved that $K < K_-$ is sufficient for all-time positivity and $K > K_+$ is necessary. For the gap interval $K_- \leq K \leq K_+$, we conjectured that state positivity is still guaranteed. Thus, a proof of the conjecture would be an obvious research direction to pursue.

References

[1] J. L. Kelly, “A New Interpretation of Information Rate,” *Bell System Technical Journal*, pp. 917–926, 1956.

[2] W. Poundstone, *Fortune’s Formula: The Untold Story of the Scientific Betting System that Beat the Casinos and Wall Street*. Hill and Wang, 2010.

[3] T. M. Cover, “An Algorithm for Maximizing Expected Log Investment Return,” *IEEE Transactions on Information Theory,* IT-30, pp. 369–373, 1984.

[4] E. O. Thorp, “Optimal Gambling Systems for Favorable Games,” *Review of the International Statistical Institute*, vol. 37, pp. 273–293, 1969.

[5] E. O. Thorp, *The Mathematics of Gambling*. Gambling Times, 1984.

[6] R. M. Bell and T. M. Cover, “Competitive Optimality of Logarithmic Investment,” *Mathematics of Operations Research*, vol. 5, pp. 161–166, 1980.

[7] P. H. Algoet and T. M. Cover, “Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment,” *The Annals of Probability*, vol. 16, pp. 876–898, 1988.

[8] T. M. Cover and E. Ordentlich, “Universal Portfolios with Side Information,” *IEEE Transactions on Information Theory,* IT-42, pp. 348–363, 1996.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2012.

[10] L. C. Maclean, W. T. Ziemba, and G. Blazenko, “Growth Versus Security in Dynamic Investment Analysis,” *Management Science*, vol. 38, pp. 1562–1585, 1992.

[11] L. C. Maclean and W. T. Ziemba, “Growth Versus Security Tradeoffs in Dynamic Investment Analysis,” *Annals of Operations Research*, vol. 85, pp. 193–227, 1999.

[12] L. C. Maclean, E. O. Thorp, and W. T. Ziemba, "Long-Term Capital Growth: The Good and Bad Properties of The Kelly and Fractional Kelly Capital Growth Criteria," *Quantitative Finance*, vol. 10, pp. 681–687, 2010.

[13] L. C. MacLean, E. O. Thorp, and W. T. Ziemba, *The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific, 2011.

[14] N. H. Hakansson, "On Optimal Myopic Portfolio Policies With and Without Serial Correlation of Yields," *Journal of Business*, vol. 44, pp. 324–334, 1971.

[15] J. Mossin, "Optimal Multiperiod Portfolio Policies," *Journal of Business*, vol. 41, pp. 215–229, 1968.

[16] H. A. Latane, "Criteria for Choice Among Risky Ventures," *Journal of Political Economy*, vol. 67, pp. 144–155, 1959.

[17] L. Breiman, "Optimal Gambling Systems for Favorable Games," *Fourth Berkeley Symposium*, vol. 1, pp. 65–78, 1961.

[18] E. O. Thorp, "Understanding the Kelly criterion," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*, pp. 509–523, World Scientific, 2011.

[19] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[20] M. Finkelstein and R. Whitley, "Optimal Strategies for Repeated Games," *Advanced Applied Probability*, vol. 13, pp. 415–428, 1981.

[21] E. O. Thorp, "Portfolio Choice and the Kelly Criterion," *Proceedings of the Business and Economics Section of the American Statistical Association*, pp. 215–224, 1971.

[22] L. M. Rotando and E. O. Thorp, "The Kelly Criterion and the Stock Market," *The American Mathematical Monthly*, vol. 99, pp. 922–931, 1992.

[23] N. H. Hakansson, "Capital Growth and the Mean-Variance Approach to Portfolio Selection," *Journal. of Financial and Quantitative Analysis*, vol. 6, pp. 517–557, 1971.

[24] T. M. Cover and D. H. Gluss, "Empirical Bayes Stock Market Portfolios," *Advances in Applied Mathematics*, vol. 7, no. 2, pp. 170–181, 1986.

[25] E. O. Thorp, "The Kelly Criterion in Blackjack Sports Betting, and The Stock Market," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*, pp. 789–832, World Scientific, 2011.

[26] Y. Lv and B. K. Meister, "Application of the Kelly Criterion to Ornstein-Uhlenbeck Processes," in *International Conference on Complex Sciences*, pp. 1051–1062, 2009.

[27] T. Duncan, B. P. Duncan, and L. Stettner, "Growth Optimal Portfolio Selection Under Proportional Transaction Costs with Obligatory Diversification," *Applied Mathematics and Optimization*, vol. 63, no. 1, pp. 107–132, 2011.

[28] T. E. Duncan, B. P. Duncan, and L. Stettner, "Parameter Continuity of the Ergodic Cost for a Growth Optimal Portfolio with Proportional Transaction Costs," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 4275–4279, 2008.

[29] J. B. Long Jr, "The Numeraire Portfolio," *Journal of Financial economics*, vol. 26, no. 1, pp. 29–69, 1990.

[30] D. G. Luenberger, *Investment Science*. Oxford University Press, 2013.

[31] J. K. Rising and A. J. Wyner, "Partial Kelly Portfolios and Shrinkage Estimators," in *IEEE Intl. Symposium on Information Theory Proceedings*, pp. 1618–1622, 2012.

[32] V. Nekrasov, "Kelly Criterion for Multivariate Portfolios: A Model-Free Approach," *Available at SSRN 2259133*, 2014.

[33] V. Nekrasov, *Knowledge Rather Than Hope: a Book for Retail Investors and Mathematical Finance Students*. Vasily Nekrasov, 2014.

[34] C. H. Hsieh and B. R. Barmish, "On Kelly Betting: Some Limitations," in *Proceedings of Allerton Conf. on Communication, Control, and Computing*, pp. 165–172, 2015.

[35] C. H. Hsieh, B. R. Barmish, and J. A. Gubner, "Kelly Betting Can be Too Conservative," in *Proceedings of IEEE Conf. on Decision and Control*, pp. 3695–3701, 2016.

[36] M.-E. Wu and W.-H. Chung, "A Novel Approach of Option Portfolio Construction Using the Kelly Criterion," *IEEE Access*, vol. 6, pp. 53044–53052, 2018.

[37] M.-E. Wu, W.-H. Chung, and C.-J. Lee, "On the Analysis of Kelly Criterion and Its Application," in *Asian Conference on Intelligent Information and Database Systems*, pp. 165–172, Springer, 2019.

[38] D. Kuhn and D. G. Luenberger, "Analysis of the Rebalancing Frequency in Log-Optimal Portfolio Selection," *Quantitative Finance*, vol. 10, no. 2, pp. 221–234, 2010.

[39] M.-E. Wu and P.-J. Hung, "A Framework of Option Buy-Side Strategy with Simple Index Futures Trading Based on Kelly Criterion," in *2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing*, pp. 210–212, IEEE, 2018.

[40] A. W. Lo, H. A. Orr, and R. Zhang, "The Growth of Relative Wealth and the Kelly Criterion," *Journal of Bioeconomics*, vol. 20, no. 1, pp. 49–67, 2018.

[41] H. Markowitz, "Portfolio Selection," *The Journal of Finance*, vol. 7, pp. 77–91, 1952.

[42] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, 1959.

[43] H. Markowitz and K. Blay, *Risk-Return Analysis: The Theory and Practice of Rational Investing*. McGraw-Hill, 2013.

[44] P. Samuelson, "The "Fallacy" of Maximizing The Geometric Mean in Long Sequences of Investing or Gambling," in *Proceedings of the National Academy of Sciences*, vol. 68, pp. 2493–2496, 1971.

[45] M. Davis and S. Lleo, "Fractional Kelly Strategies for Benchmarked Asset Management," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*, pp. 385–407, World Scientific, 2011.

[46] L. C. Maclean, R. Sanegre, Y. Zhao, and W. T. Ziemba, "Capital Growth with Security," *Journal of Economic Dynamics and Control*, vol. 28, pp. 937–954, 2004.

[47] T. M. Cover and E. Ordentlich, "Universal Portfolios with Short Sales and Margin," in *Proceedings. IEEE International Symposium on Information Theory*, p. 174, 1998.

[48] B. R. Barmish and J. A. Primbs, "On a New Paradigm for Stock Trading Via a Model-Free Feedback Controller," *IEEE Transactions on Automatic Control,* AC-61, pp. 662–676, 2016.

[49] S. Malekpour, J. A. Primbs, and B. R. Barmish, "On Stock Trading Using a PI Controller in an Idealized Market: The Robust Positive Expectation Property," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 1210–1216, 2013.

[50] R. C. Merton, *Continuous-Time Finance*. Blackwell Cambridge, 1990.

[51] J. White, "More Institutional Investors Selling Short: But Tactic Is Part of Wider Strategy," *Wall Street Journal*, vol. 7, pp. 493–501, 1990.

[52] C. C. Kwan, "Portfolio Selection Under Institutional Procedures for Short Selling: Normative and Market-Equilibrium Considerations," *Journal of Banking and Finance*, vol. 21, no. 3, pp. 369–391, 1997.

[53] G. J. Alexander, "Short Selling and Efficient Sets," *The Journal of Finance*, vol. 48, no. 4, pp. 1497–1506, 1993.

[54] G. J. Alexander, "Efficient Sets, Short-selling, and Estimation Risk," *Journal of Portfolio Management*, vol. 21, no. 2, p. 64, 1995.

[55] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming."

[56] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley and Sons, 1997.

[57] S. Boyd, E. Busseti, S. Diamond, R. N. Kahn, K. Koh, P. Nystrup, and J. Speth, "Multi-Period Trading via Convex Optimization," *Foundations and Trends in Optimization*, vol. 3, no. 1, pp. 1–76, 2017.

[58] L. E. Ghaoui, M. Oks, and F. Oustry, "Worst-Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach," *Operations research*, vol. 51, no. 4, pp. 543–556, 2003.

[59] D. P. Brown, "Multiperiod Financial Planning," *Management Science*, vol. 33, no. 7, pp. 848–875, 1987.

[60] C. H. Hsieh, B. R. Barmish, and J. A. Gubner, "The Impact of Execution Delay on Kelly-Based Stock Trading: High-Frequency Versus Buy and Hold," *Proceedings of the IEEE Conference on Decision and Control*, pp. 2580–2585, 2019.

[61] C. H. Hsieh, B. R. Barmish, and J. A. Gubner, "On Positive Solutions of a Delay Equation Arising When Trading in Financial Markets," *IEEE Transactions on Automatic Control,* in press, 2019.

[62] T. M. Cover, "Universal Portfolios," *Mathematical Finance*, vol. 1, pp. 1–29, 1991.

[63] C. H. Hsieh, B. R. Barmish, and J. A. Gubner, "At What Frequency Should Kelly Bettor Bet?," in *Proceedings of the American Control Conference*, pp. 5485–5490, 2018.

[64] C. H. Hsieh, J. A. Gubner, and B. R. Barmish, "Rebalancing Frequency Considerations for Kelly-Optimal Stock Portfolios in a Control-Theoretic Framework," in *Proceedings of IEEE Conference on Decision and Control*, pp. 5820–5825, 2018.

[65] A. Deshpande and B. R. Barmish, "A General Framework for Pairs Trading with a Control-Theoretic Point of View," in *Proceedings of the IEEE Conference on Control Applications*, pp. 761–766, 2016.

[66] J. A. Primbs and B. R. Barmish, “On Robustness of Simultaneous Long-Short Stock Trading Control with Time-Varying Price Dynamics,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 12267–12272, 2017.

[67] J. A. Primbs, “Portfolio Optimization Applications of Stochastic Receding Horizon Control,” in *Proceedings of the American Control Conference*, pp. 1811–1816, 2007.

[68] B. R. Barmish, “On Trading of Equities: A Robust Control Paradigm,” *IFAC Proceedings Volumes*, vol. 41, pp. 1621–1626, 2008.

[69] G. C. Calafiore, “Multi-period Portfolio Optimization with Linear Control Policies,” *Automatica*, vol. 44, no. 10, pp. 2463–2473, 2008.

[70] G. C. Calafiore, “An Affine Control Method for Optimal Dynamic Asset Allocation with Transaction Costs,” *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2254–2274, 2009.

[71] G. C. Calafiore and B. Monastero, “Triggering Long-Short Trades on Indexes,” *International Journal of Trade, Economics and Finance*, vol. 1, pp. 289–296, 2010.

[72] B. R. Barmish and J. A. Primbs, “On Arbitrage Possibilities Via Linear Feedback in an Idealized Brownian Motion Stock Market,” in *Proceedings of the IEEE Conference on Decision and Control and European Control Conference*, pp. 2889–2894, 2011.

[73] B. R. Barmish and J. A. Primbs, “On Market-Neutral Stock Trading Arbitrage Via Linear Feedback,” in *Proceedings of the American Control Conf.*, pp. 3693–3698, 2012.

[74] S. Malekpour and B. R. Barmish, “Drawdown Formula for Stock Trading Via Linear Feedback in a Market Governed by Brownian Motion,” in *Proceedings of the European Control Conference*, pp. 87–92, 2013.

[75] M. H. Baumann, “On Stock Trading Via Feedback Control When Underlying Stock Returns Are Discontinuous,” *IEEE Transactions on Automatic Control,* AC-62, no. 6, pp. 2987–2992, 2016.

[76] A. Deshpande and B. R. Barmish, "A Generalization of the Robust Positive Expectation Theorem for Stock Trading via Feedback Control," in *Proceedings of the European Control Conference*, pp. 514–520, 2018.

[77] D. P. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, 2004.

[78] G. Maroni, S. Formentin, and F. Previdi, "A Robust Design Strategy for Stock Trading via Feedback Control," in *Proceedings of the European Control Conference*, pp. 447–452, 2019.

[79] I. Khademi and K. Zhou, "How to Invest in Uncertain Markets?," in *Proceedings of the Chinese Control Conference*, pp. 8617–8622, 2015.

[80] S. Iwarere and B. R. Barmish, "On Stock Trading Over a Lattice Via Linear Feedback," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7799–7804, 2014.

[81] S. Iwarere and B. R. Barmish, "A Confidence Interval Triggering Method for Stock Trading via Feedback Control," in *Proceedings of the American Control Conference*, pp. 6910–6916, 2010.

[82] R. C. Dorf and R. H. Bishop, *Modern Control Systems*. Pearson, 2011.

[83] K. J. Aström and R. M. Murray, *Feedback Systems: an Introduction for Scientists and Engineers*. Princeton University Press, 2010.

[84] B. R. Barmish, "On Performance Limits of Feedback Control-Based Stock Trading Strategies," in *Proceedings of the American Control Conference*, pp. 3874–3879, 2011.

[85] S. Malekpour and B. R. Barmish, "How Useful are Mean-Variance Considerations in Stock Trading via Feedback Control," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 2110–2115, 2012.

[86] S. Malekpour, *Towards Demystification of Model-Free Technical Analysis: The Power of Feedback Control*. PhD thesis, University of Wisconsin-Madison, 2016.

[87] S. Malekpour, J. A. Primbs, and B. R. Barmish, "A Generalization of Simultaneous Long-Short Stock Trading to PI Controllers," *IEEE Transactions on Automatic Control,* AC-63, pp. 3531–3536, 2018.

[88] E. F. Fama and K. R. French, "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, vol. 47, no. 2, pp. 427–465, 1992.

[89] E. F. Fama, "The Behavior of Stock-Market Prices," *The journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.

[90] C. H. Hsieh and B. R. Barmish, "On Drawdown-Modulated Feedback Control in Stock Trading," *IFAC-PapersOnline*, vol. 50, pp. 952–958, 2017.

[91] C. H. Hsieh and B. R. Barmish, "On Inefficiency of Markowitz-Style Investment Strategies When Drawdown is Important," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 3075–3080, 2017.

[92] P. M. Möller, "Drawdown Measures and Return Moments," *International Journal of Theoretical and Applied Finance*, vol. 21, no. 7, p. 1850042, 2018.

[93] B. T. Hayes, "Maximum Drawdowns of Hedge Funds with Serial Correlation," *The Journal of Alternative Investments*, vol. 8, no. 4, p. 26, 2006.

[94] M. Ismail, A. Atiya, A. Pratap, and Y. Abu-Mostafa, "On the Maximum Drawdown of a Brownian Motion," *Journal of Applied Probability*, pp. 147–161, 2004.

[95] R. T. Rockafellar, S. Uryasev, and M. Zabarankin, "Generalized Deviations in Risk Analysis," *Finance and Stochastics*, vol. 10, no. 1, pp. 51–74, 2006.

[96] L. R. Goldberg and O. Mahmoud, "Drawdown: From Practice to Theory and Back Again," *Mathematics and Financial Economics*, vol. 74, pp. 1–23, 2016.

[97] R. T. Rockafellar, S. P. Uryasev, and M. Zabarankin, "Deviation Measures in Risk Analysis and Optimization," *University of Florida, Department of Industrial and Systems Engineering Working Paper*, no. 2002-7, 2002.

[98] S. J. Grossman and Z. Zhou, "Optimal Investment Strategies for Controlling Drawdowns," *Mathematical Finance*, vol. 3, no. 3, pp. 241–276, 1993.

[99] D. P. Foster and S. Hart, "An Operational Measure of Riskiness," *Journal of Political Economy,*, vol. 117, pp. 785–814, 2009.

[100] J. Cvitanic and I. Karatzas, "On Portfolio Optimization Under Drawdown Constraints," *Institute for Mathematics and its Applications (IMA) Lecture Notes in Mathematics and Applications*, vol. 65, pp. 77–88, 1994.

[101] G. C. Calafiore and B. Monastero, "Data-Driven Asset Allocation with Guaranteed Short-Fall Probability," in *Proceedings of the American Control Conference*, pp. 3687–3692, 2012.

[102] C. T. Chen, *Linear System Theory and Design*. Oxford University Press, 2009.

[103] G. Campolieti and R. N. Makarov, *Financial Mathematics: A Comprehensive Treatment*. CRC Press, 2014.

[104] Z. Bodie, A. Kane, and A. J. Marcus, *Investments*. McGraw-Hill, 2010.

[105] Z. Bodie, A. Kane, and A. J. Marcus, *Investments and Portfolio Management*. McGraw-Hill, 2013.

[106] S. R. Das, D. Kaznachey, and M. Goyal, "Computing Optimal Rebalance Frequency for Log-Optimal Portfolios," *Quantitative Finance*, vol. 14, no. 8, pp. 1489–1502, 2014.

[107] L. Farina, "Positive Systems in the State Space Approach: Main Issues and Recent Results," in *Proceedings of International Symposium on Mathematical Theory of Networks and Systems*, 2002.

[108] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, vol. 50. John Wiley and Sons, 2011.

[109] T. Kaczorek, "Positivity and Stability of Time-Varying Discrete-Time Linear Systems," in *Proceedings of Asian Conference on Intelligent Information and Database Systems*, pp. 295–303, 2015.

[110] D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications*. John Wiley and Sons, 1979.

[111] L. H. Erbe and B.-G. Zhang, "Oscillation of Discrete Analogues of Delay Equations," *Differential Integral Equations*, vol. 2, no. 3, pp. 300–309, 1989.

[112] L. Berezansky and E. Braverman, "On Existence of Positive Solutions for Linear Difference Equations with Several Delays," *Advances in Dynamical Systems and Applications*, vol. 1, no. 1, pp. 29–47, 2006.

[113] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[114] A. Cambini and L. Martein, *Generalized convexity and optimization: Theory and applications*, vol. 616. Springer Science and Business Media, 2008.

[115] W. Rudin, *Principles of Mathematical Analysis*. McGraw-hill, 1976.

[116] H. Witsenhausen, "Some Aspects of Convexity Useful in Information Theory," *IEEE Transactions on Information Theory,* IT-26, no. 3, pp. 265–271, 1980.

[117] M. J. Klass and K. Nowicki, "The Grossman and Zhou Investment Strategy is Not Always Optimal," *Statistics and Probability Letters*, vol. 74, pp. 245–252, 2005.

[118] J. Estrada, "Geometric Mean Maximization: An Overlooked Portfolio Approach?," *Journal of Investing*, pp. 134–147, 2010.

[119] J. Zhou and J. Shang, "Drawdown Control With a Risk Floor," *The Journal of Alternative Investments*, vol. 17, pp. 79–94, 2015.

[120] V. Cherny and J. Obłój, "Portfolio Optimisation under Non-linear Drawdown Constraints in a Semimartingale Financial Model," *Finance and Stochastics*, vol. 17, no. 4, pp. 771–800, 2013.

[121] A. Chekhlov, S. Uryasev, and M. Zabarankin, "Portfolio Optimization with Drawdown Constraints," in *Supply Chain and Finance*, pp. 209–228, World Scientific, 2004.

[122] A. Chekhlov, S. Uryasev, and M. Zabarankin, "Drawdown Measure in Portfolio Optimization," *International Journal of Theoretical and Applied Finance*, vol. 8, no. 1, pp. 13–58, 2005.

[123] J. Philippe, "Value at Risk: the New Benchmark for Managing Financial Risk," *NY: McGraw-Hill Professional*, 2001.

[124] N. Wesselhöfft and W. K. Härdle, "Risk-Constrained Kelly Portfolios Under Alpha-Stable Laws," *Computational Economics*, pp. 1–26, 2019.

[125] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, 2009.

[126] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009.

[127] M. Dai, Q. Zhang, and Q. J. Zhu, "Trend Following Trading under a Regime Switching Model," *SIAM Journal on Financial Mathematics*, vol. 1, no. 1, pp. 780–810, 2010.

[128] M. Covel, *Trend Following: Learn to Make Millions in Up or Down Markets*. Financial Times Press, 2009.

[129] D. Darius, A. Ilhan, J. Mulvey, K. D. Simsek, and R. Sircar, "Trend-Following Hedge Funds and Multi-Period Asset Allocation," *Journal of Quantitative Finance*, vol. 2, no. 5, pp. 354–361, 2002.

[130] M. Davidsson, "Expected Return and Portfolio Rebalancing," *International Journal of Economics and Finance*, vol. 3, no. 3, pp. 15–26, 2011.

[131] M. Qi and Y. Wu, "Technical Trading-rule Profitability, Data Snooping, and Reality Check: Evidence from the Foreign Exchange Market," *Journal of Money, Credit and Banking*, pp. 2135–2158, 2006.

[132] D. P. Brown and R. H. Jennings, "On Technical Analysis," *The Review of Financial Studies*, vol. 2, no. 4, pp. 527–551, 1989.

[133] P. A. Samuelson, "The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments," *The Review of Economic Studies*, vol. 37, no. 4, pp. 537–542, 1970.

[134] C. R. Harvey, J. C. Liechty, M. W. Liechty, and P. Müller, "Portfolio Selection with Higher Moments," *Quantitative Finance*, vol. 10, no. 5, pp. 469–485, 2010.

[135] B. I. Jacobs, K. N. Levy, and H. M. Markowitz, "Portfolio Optimization with Factors, Scenarios, and Realistic Short Positions," *Jour. of Operations Research*, vol. 53, pp. 586–599, 2005.

[136] M. C. Steinbach, "Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis," *SIAM Review*, vol. 43, no. 1, pp. 31–85, 2001.

[137] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.

[138] J. S. Rosenthal, *A First Look at Rigorous Probability Theory*. World Scientific Publishing Company, 2006.

[139] E. W. Weisstein, "Lambert W-function," *MathWorld, A Wolfram Web Resource*, 2002.

[140] P. Billingsley, *Probability and Measure*. John Wiley and Sons, 2008.

[141] Wharton School, "Wharton Research Data Services." avaiable at `https://wrds-web.wharton.upenn.edu/wrds/`.

[142] S. E. Shreve, *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Science and Business Media, 2004.

[143] I. Karatzas and S. E. Shreve, "Brownian Motion," in *Brownian Motion and Stochastic Calculus*, pp. 47–127, Springer, 1998.

[144] N. Rujeerapaiboon, B. R. Barmish, and D. Kuhn, "On Risk Reduction in Kelly Betting Using the Conservative Expected Value," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 5801–5806, 2018.

[145] X. Liu, L. Wang, W. Yu, and S. Zhong, "Constrained Control of Positive Discrete-Time Systems with Delays," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 2, pp. 193–197, 2008.

[146] T. Kaczorek, *Positive 1D and 2D Systems*. Springer Science and Business Media, 2012.

[147] E. I. Jury, *Theory and Application of the z-Transform Method*. Wiley New York, 1964.

[148] B. R. Barmish, *New Tools for Robustness of Linear Systems*. Macmillan Publishing Company, 1994.