KELLY CRITERION: FROM A SIMPLE RANDOM WALK TO LÉVY PROCESSES

SERGEY LOTOTSKY* AND AUSTIN POLLOK†

Abstract. The original Kelly criterion provides a strategy to maximize the long-term growth of winnings in a sequence of simple Bernoulli bets with an edge, that is, when the expected return on each bet is positive. The objective of this work is to consider more general models of returns and the continuous time, or high frequency, limits of those models. The results include an explicit expression for the optimal strategy in several models with continuous time compounding.

Key words. Logarithmic Utility, Processes with Independent Increments, Weak Convergence

AMS subject classifications. 60G50, 60F05, 60F17, 60G51, 91A60

1. Introduction. Consider repeatedly engaging in a game of chance where one side has an edge and seeks to optimize the betting in a way that ensures maximal long-term growth rate of the overall wealth. This problem was originally posed and analyzed for some special cases by John Kelly [9] at Bell Labs in 1956; the solution was implemented and tested in a variety of setting by a successful mathematician, gambler, and hedge fund manager Ed Thorp over the period from the 60's to the early 00's [20, 21, 23, etc.] Reference [14] provides a comprehensive survey of the Kelly criterion and its applications.

As a motivating example, going back to Kelly [9], consider betting on a biased coin toss where the return r is a random variable with distribution

$$\mathbb{P}(r = 1) = p, \quad \mathbb{P}(r = -1) = 1 - p; \tag{1.1}$$

in what follows, we refer to this as the *simple Bernoulli model*. The condition to have an edge in this setting becomes $1/2 < p \leq 1$ or, equivalently,

$$\mathbb{E}[r] = 2p - 1 > 0. \tag{1.2}$$

We plan on being able to make a large sequence of bets on this biased coin, resulting in an iid sequence of returns $\{r_k\}_{k\geq 1}$ with the same distribution as r, and ask how much we should bet so as to maximize long term wealth, given that we are compounding our returns. Assume we are betting with a fixed exposure f, that is, each bet involves a fixed fraction f of the overall wealth, and $f \in [0, 1]$. Practically, $f \geq 0$ means **no shorting** and $f \leq 1$ means **no leverage**, which we refer to as the **NS-NL** condition. Then, starting with the initial amount W_0, the total wealth at time $n = 1, 2, 3, \ldots$ is the following function of f:

$$W_n^f = W_0 \prod_{k=1}^{n} \big(1 + f r_k\big).$$

For the long-term compounder wishing to maximize their long term wealth, a natural and equivalent goal would be to find the strategy $f = f^*$ maximizing the long-term

*Department of Mathematics, University of Southern California, Los Angeles, CA 90089 (lototsky@usc.edu, https://dornsife.usc.edu/sergey-lototsky/).

†Department of Mathematics, University of Southern California, Los Angeles, CA 90089 (pollok@usc.edu).

growth rate

$$g_r(f) := \lim_{n\to\infty} \frac{1}{n} \ln \frac{W_n^f}{W_0}. \tag{1.3}$$

By direct computation,

$$g_r(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \ln(1 + fr_k) = \mathbb{E} \ln(1 + fr) = p \ln(1 + f) + (1 - p) \ln(1 - f),$$

where the second equality follows by the law of large numbers, and therefore, after solving $g_r'(f^*) = 0$,

$$f^* = 2p - 1, \quad \max_{f\in[0,1]} g_r(f) = g_r(f^*) = p \ln \frac{p}{1 - p} + (2 - p) \ln(2 - 2p); \tag{1.4}$$

note that the edge condition (1.2) ensures that f^* is an admissible strategy and $g_r(f^*) > 0$. For more discussions of this result see [21].

Our objective in this paper is to derive analogues of (1.4) in the following situations:

1. the distribution of returns is a more general random variable (Section 2);
2. the compounding is continuous in time (Section 3);
3. the compounding is high frequency, leading to a continuous-time limit (Sections 4 and 5).

In particular,

1. We show that, for a large class of returns, the optimal strategy in the high-frequency limit only depends on the mean μ and the variance σ^2 of the return:

   $$f^* = \frac{\mu}{\sigma^2}. \tag{1.5}$$

   In fact, as we show in Section 5, relation (1.5) holds even in some models with random time horizon. These results contribute to the exiting literature on the subject, such as [2] and [21].
2. We show in Section 3 that, for a *non-random* long-term growth rate to exist in continuous time, the corresponding return process, up to a small correction, must be a semi-martingale with independent increments; the result also leads to a non-trivial generalization of (1.5).
3. We derive in Section 4 a completely new version of (1.5) when the returns are described by infinitely divisible distributions other than normal, and thus address some of Thorp's questions regarding fat-tailed distributions in finance [23].

To illustrate (1.5), consider a classical problem in portfolio theory asking for an optimal mix of equities and risk-free assets. For simplicity, assume that the portfolio has only two components, a risky asset and cash. If the annual rate of return on the risky asset has mean 30% and the standard deviation 65%, and if the investor is frequently re-balancing the portfolio with the goal of maximizing the long-term wealth, then, according to (1.5), the optimal portfolio should be

$$f^* = \frac{0.3}{(0.65)^2} \approx 0.7 \text{ or } 70\%$$

the risky asset and $100 - 70 = 30\%$ cash. Note that the Markowitz portfolio theory can lead to the same result, but using a different argument; we compare and contrast the two approaches in Section 6. This classical problem has been re-branded by the asset management industry as *goal-based investing* which seeks to build multi-asset, or model, portfolios with a mix of assets to achieve a particular goal, such as maximizing income or minimizing shortfall risk of reaching a particular target. Another possible goal is a growth-optimal portfolio [13, Ch. 15], which is the subject of our investigation. More precisely, the goal is to maximize long term growth (1.3), and then, according to (1.5), the optimal portfolio should use a mix prescribed by $(1-f^*, f^*)$ in cash and equities, respectively.

It has been well documented [3, 4, 18, etc.] that many financial time series, especially at a higher frequency, exhibit heavy tails and skewness, both at the statistical and risk-neutral levels. Formula (1.5) does not account for risky assets with heavy-tailed or skewed distributions, and therefore any portfolio based on (1.5) is likely to be sub-optimal in many situations. In the current work, we take a more realistic view of financial markets by allowing non-normally distributed returns. This approach leads us to several new formulas expressing the long-term growth rate and the corresponding optimal allocation (see, for example, (3.7) and (3.28) below). Not surprisingly, these new formulas contain known results, such as (1.5), as a particular case.

The Kelly criterion, such as (1.4) or (1.5), is also related to expected utility theory when the utility function is logarithmic, even though this relation is not an easy one. While the idea of logarithmic utility goes back to Daniel Bernoulli's resolution of the St. Petersburg paradox, there are many other choices of the utility function in modern economic and financial literature. Moreover, Kelly's use of the logarithm function should be interpreted as a *prescriptive* goal-based utility function coming from the need to maximize long term growth rate of an investment, as opposed to a (more traditional) *descriptive* utility function characterizing preferences. For more details on this part of the story, see [22].

In what follows, we write $\xi \overset{d}{=} \eta$ to indicate equality in distribution for two random variables, and $X \overset{\mathcal{L}}{=} Y$ to indicate equality in law (as function-valued random elements) for two random processes. For $x > 0$, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. To simplify the notations, we always assume that $W_0 = 1$.

2. Discrete Compounding: General Distribution of Returns. Assume that the returns on each bet are independent random variables $r_k,\ k \geq 1$, with the same distribution as a given random variable r, and let

$$W_n^f = \prod_{k=1}^{n} \big(1 + f r_k\big),\ n = 1, 2, \dots, \tag{2.1}$$

denote the corresponding wealth process. We also keep the NS-NL condition on admissible strategies: $f \in [0, 1]$.

For the wealth process W^f to be a non-trivial object worthy of investigation, we need the random variable r to have the following properties:

$$\mathbb{P}(r \geq -1) = 1; \tag{2.2}$$

$$\mathbb{P}(r > 0) > 0,\ \mathbb{P}(r < 0) > 0; \tag{2.3}$$

$$\mathbb{E}|\ln(1 + r)| < \infty. \tag{2.4}$$

Condition (2.2) quantifies the idea that a loss in a bet should not be more than 100%. Condition (2.3) is basic non-degeneracy: both gains and losses are possible. Condition (2.4) is a minimal requirement to define the long-term growth rate of the wealth process.

The key object in this section will be the function

$$g_r(f) = \mathbb{E}\ln(1+fr). \tag{2.5}$$

In particular, the following result shows that $g_r(f)$ is the long term growth rate of the wealth process W^f.

Proposition 2.1. *If (2.2) and (2.4) hold and $g_r(f) \neq 0$, then, for every $f \in [0,1]$, the wealth process W^f has an asymptotic representation*

$$W_n^f = \exp\Big(ng_r(f)\big(1+\varepsilon_n\big)\Big), \tag{2.6}$$

where

$$\lim_{n\to\infty} \varepsilon_n = 0 \tag{2.7}$$

with probability one.

Proof. By (2.1), we have (2.6) with

$$\varepsilon_n = \frac{1}{ng_r(f)} \sum_{k=1}^{n} \Big(\ln(1+fr_k) - g_r(f)\Big), \tag{2.8}$$

and then (2.7) follows by (2.4) and the strong law of large numbers. □

A stronger version of (2.4) leads to a more detailed asymptotic of W_n^f.

Theorem 2.2. *Assume that (2.2) holds and*

$$\mathbb{E}|\ln(1+r)|^2 < \infty. \tag{2.9}$$

Then then, for every $f \in [0,1]$, the wealth process W^f has an asymptotic representation

$$W_n^f = \exp\Big(ng_r(f) + \sqrt{n}\big(\sigma_r(f)\zeta_n + \epsilon_n\big)\Big), \tag{2.10}$$

where $\zeta_n,\ n \geq 1$, are standard Gaussian random variables,

$$\sigma_r(f) = \Big(\mathbb{E}\big[\ln^2(1+fr)\big] - g_r^2(f)\Big)^{1/2},$$

and

$$\lim_{n\to\infty} \epsilon_n = 0$$

in probability.

Proof. With ε_n from (2.8), the result follows by the Central Limit Theorem:

$$ng_r(f)\,\varepsilon_n = \sqrt{n}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\Big(\ln(1+fr_k) - g_r(f)\Big)\right) = \sqrt{n}\big(\sigma_r(f)\zeta_n + \epsilon_n\big).$$

Because the Central Limit Theorem gives convergence in distribution, the random variables ζ_n in (2.10) can indeed depend on n. Additional assumptions about the distribution of r, such as existence of higher-order moments, lead to higher-order asymptotic expansions and a possibility to have $\lim_{n\to\infty} \epsilon_n = 0$ with probability one [24, Theorem 1]. We do not pursue this direction: our goal is to keep the conditions on the distribution of r as general as possible.

The following properties of the function g_r are immediate consequences of the definition and the assumptions (2.2)–(2.4):

PROPOSITION 2.3. *The function $f \mapsto g_r(f)$ is continuous on the closed interval $[0, 1]$ and infinitely differentiable in $(0, 1)$. In particular,*

$$\frac{dg_r}{df}(f) = \mathbb{E}\left[\frac{r}{1+fr}\right], \quad \frac{d^2 g_r}{df^2}(f) = -\mathbb{E}\left[\frac{r^2}{(1+fr)^2}\right] < 0. \tag{2.11}$$

COROLLARY 2.4. *The function g_r achieves its maximal value on $[0, 1]$ at a unique point $f^* \in [0, 1]$ and $g_r(f^*) \geq 0$.*

Proof. Note that $g_r(0) = 0$ and, by (2.11), the function g_r is strictly concave on $[0, 1]$. □

While concavity of g_r implies that g_r achieves a unique global maximal value at a point f^{**}, it is possible that the domain of the function g_r is bigger than the interval $[0, 1]$ and $f^{**} \notin [0, 1]$. A simple way to exclude the possibility $f^{**} < 0$ is to consider returns r that are not bounded from above: $\mathbb{P}(r > c) > 0$ for all $c > 0$: in this case, the function $g_r(f) = \mathbb{E} \ln(1 + fr)$ is not defined for $f < 0$. Similarly, if $\mathbb{P}(r < -1 + \delta) > 0$ for all $\delta > 0$, then the function g_r is not defined for $f > 1$, excluding the possibility $f^{**} > 1$.

Below are more general sufficient conditions to ensure that the point $f^* \in [0, 1]$ from Corollary 2.4 is the point of global maximum of g_r: $f^* = f^{**}$.

PROPOSITION 2.5. *If*

$$\lim_{f\to 0+} \mathbb{E}\left[\frac{r}{1+fr}\right] > 0 \qquad \text{and} \tag{2.12}$$

$$\lim_{f\to 1-} \mathbb{E}\left[\frac{r}{1+fr}\right] < 0, \tag{2.13}$$

then there is a unique $f^* \in (0, 1)$ such that

$$g_r(f) < g_r(f^*)$$

for all f in the domain of g_r.

Proof. Together with the intermediate value theorem, conditions (2.12) and (2.13) imply that there is a unique $f^* \in (0, 1)$ such that

$$\frac{dg_r}{df}(f^*) = 0.$$

It remains to use strong concavity of g_r. □

Because $r \geq -1$, the expected value $\mathbb{E}[r]$ is always defined, although $\mathbb{E}[r] = +\infty$ is a possibility. Thus, by (2.11), condition (2.12) is equivalent to the intuitive idea of an edge:

$$\mathbb{E}[r] > 0,$$

which, similar to (1.2), guarantees that $g_r(f) > 0$ for some $f \in (0, 1)$. Condition (2.13) can be written as

$$\mathbb{E}\left[\frac{r}{1+r}\right] < 0,$$

with the convention that the left-hand side can be $-\infty$. This condition does not appear in the simple Bernoulli model, but is necessary in general, to ensure that the edge is not too big and leveraged gambling ($f^* > 1$) does not lead to an optimal strategy.

As an example, consider the *general Bernoulli model* with

$$\mathbb{P}(r=-a) = 1-p,\ \ \mathbb{P}(r=b) = p,\ \ 0 < a \le 1,\ b > 0,\ 0 < p < 1. \tag{2.14}$$

The function

$$g_r(f) = p\ln(1+fb) + (1-p)\ln(1-fa)$$

is defined on $(-1/b, 1/a)$, achieves the global maximum at

$$f^* = \frac{p}{a} - \frac{1-p}{b},$$

and

$$g_r(f^*) = p\ln\frac{p(a+b)}{a} + (1-p)\ln\frac{(1-p)(a+b)}{b}.$$

Note that $g_r(f^*)$ is equal to the *Kullback-Leibler Divergence* $D_{KL}(P \parallel Q)$, where P is the Bernoulli distribution with probability of success p and Q is the Bernoulli distribution with probability of success $a/(a+b)$. This observation confirms that $g_r(f^*) \ge 0$ and provides another example of an information-theoretic interpretation of the optimal rate of return; cf. [9].

The NS-NL condition $f^* \in [0,1]$ becomes

$$\frac{a}{a+b} \le p \le \min\left(\frac{ab}{a+b}\left(1+\frac{1}{b}\right), 1\right),$$

and it is now easy to come up with a model in which $f^* > 1$: for example, take

$$a = 0.1,\ b = 0.5,\ p = 0.5$$

so that $f^* = 4$. Given that a gain and a loss in each bet are equally likely, but the amount of a gain is five times as much as that of a loss, a large value of f^* is not surprising, although economical and financial implications of this type of leveraged betting are potentially very interesting and should be a subject of a separate investigation.

Because of the logarithmic function in the definition of g_r, the distribution of r can have a rather heavy right tail and still satisfy (2.4). For example, consider

$$r = \eta^2 - 1, \tag{2.15}$$

where η has standard Cauchy distribution with probability density function

$$h_\eta(x) = \frac{1}{\pi(1+x^2)},\quad -\infty < x < +\infty.$$

Then

$$g_r(f) = \frac{2}{\pi}\int_0^{+\infty} \frac{\ln\big((1-f)+fx^2\big)}{1+x^2}\,dx = 2\ln\big(\sqrt{f}+\sqrt{1-f}\big),$$

where the second equality follows from [7, Formula (4.295.7)]. As a result, we get a closed-form answer

$$f^* = \frac{1}{2},\ g_r(f^*) = \ln 2.$$

A general way to ensure (2.2)–(2.4) is to consider

$$r = e^{\xi} - 1 \tag{2.16}$$

for some random variable ξ such that $\mathbb{P}(\xi > 0) > 0$, $\mathbb{P}(\xi < 0) > 0$, and $\mathbb{E}|\xi| < \infty$; note that (2.15) is a particular case, with $\xi = \ln \eta^2$. Then (2.12) and (2.13) become, respectively,

$$\mathbb{E}e^{\xi} > 1 \quad \text{and} \tag{2.17}$$

$$\mathbb{E}e^{-\xi} > 1. \tag{2.18}$$

For example, if ξ is normal with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, then

$$\mathbb{E}e^{\xi} = e^{\mu + (\sigma^2/2)}, \quad \mathbb{E}e^{-\xi} = e^{-\mu + (\sigma^2/2)},$$

and (2.17), (2.18) are equivalent to

$$-\frac{\sigma^2}{2} < \mu < \frac{\sigma^2}{2}, \tag{2.19}$$

which, when interpreted in terms of returns, can indeed be considered as a "reasonable" edge condition: large values of $|\mu|$ do create a bias in one direction. For this and many other models of the form (2.16), the corresponding f^* is not available in closed form but can be evaluated numerically by approximating the function $g_r(f)$ for several values of f using (2.5) and Monte Carlo simulations.

Except for Theorem 2.2, all other results in this section continue to hold when the sequence $\{r_k,\ k \geq 1\}$ of returns is strictly stationary and ergodic [10, Section 20.2].

3. Continuous Compounding and a Case for Lévy Processes. Continuous time compounding includes discrete compounding as a particular case and makes it possible to consider more general types of return processes. The objective of this section is to show that continuous time compounding that leads to a non-trivial and non-random long-term growth rate of the resulting wealth process effectively forces the return process to have independent increments. The two main examples of such process are sums of iid random variables from the previous section and the Lévy processes.

Writing (2.1) as

$$W^f_{n+1} - W^f_n = \left(fW^f_n\right) r_{n+1}, \tag{3.1}$$

we see that a natural continuous time version of (3.1) is

$$dW^f_t = fW^f_t\, dR_t \tag{3.2}$$

for a suitable process $R = R_t,\ t \geq 0$ on a stochastic basis

$$\mathbb{F} = \Big(\Omega, \mathcal{F},\ \{\mathcal{F}_t\}_{t\geq 1}, \mathbb{P}\Big)$$

satisfying the usual conditions [17, Section I.1]. We interpret (3.2) as an integral equation

$$W^f_t = 1 + f\int_0^t W^f_s\, dR_s; \tag{3.3}$$

recall that $W_0^f=1$ is the standing assumption. Then the Bichteler-Dellacherie theorem [17, Theorem III.47] implies that the process R must be a semi-martingale (a sum of a martingale and a process of bounded variation) with trajectories that, at every point, are continuous from the right and have limits from the left. Furthermore, if we allow the process R to have discontinuities, then, by [17, Theorem II.36], we need to modify (3.3) further:

$$W_t^f = 1 + f\int_0^t W_{s-}^f\, dR_s,$$

where

$$W_{s-}=\lim_{\varepsilon\to 0,\varepsilon>0} W_{s-\varepsilon},$$

and, assuming $R_0=0$, the process W^f becomes the Doléans-Dade exponential

$$W_t^f = \exp\left(fR_t - \frac{f^2\langle R^c\rangle_t}{2}\right)\prod_{0<s\leq t}(1+f\triangle R_s)\,e^{-f\triangle R_s}; \tag{3.4}$$

cf. [12, Theorem 2.4.1]. In (3.4), $\langle R^c\rangle$ is the quadratic variation process of the continuous martingale component of R and $\triangle R_s = R_s - R_{s-}$.

A natural analog of (2.2) is

$$\triangle R_s \geq -1, \tag{3.5}$$

and then (3.4) becomes

$$W_t^f = \exp\left(fR_t - \frac{f^2\langle R^c\rangle_t}{2} + \sum_{0<s\leq t}\Big(\ln(1+f\triangle R_s) - f\triangle R_s\Big)\right). \tag{3.6}$$

To proceed, let us assume that the trajectories of R are continuous: $\triangle R_s=0$ for all s so that

$$W_t^f = \exp\left(fR_t - f^2\langle R^c\rangle_t\right).$$

If, similar to (1.3), we define the long-term growth rate $g_R(f)$ by

$$g_R(f)=\lim_{t\to\infty}\frac{\ln W_t^f}{t}, \tag{3.7}$$

then we need the limits

$$\mu := \lim_{t\to\infty}\frac{R_t}{t},\quad \sigma^2 := \lim_{t\to\infty}\frac{\langle R^c\rangle_t}{t} \tag{3.8}$$

to exist with probability one and with non-random numbers μ,σ^2. Being a semi-martingale without jumps, the process R has a representation

$$R_t = A_t + R_t^c, \tag{3.9}$$

where A is process of bounded variation; cf. [8, Theorem II.2.34]. Then (3.8) imply that, as $t\to+\infty$,

$$A_t=\mu t\big(1+o(1)\big),\quad \langle R^c\rangle_t = \sigma^2 t\big(1+o(1)\big), \tag{3.10}$$

A natural way to achieve (3.8) is to ignore the $o(1)$ terms in (3.10), and then, by the Lévy characterization of the Brownian motion, the process R becomes

$$R_t = \mu t + \sigma B_t,$$

where $\sigma > 0$ and $B = B_t$ is a standard Brownian motion. Then

$$W_t^f = \exp\left(f\mu t + f\sigma B_t - \frac{f^2\sigma^2 t}{2}\right) \tag{3.11}$$

is a geometric Brownian motion.

The long-term growth rate (3.7) becomes

$$g_R(f) = f\mu - \frac{f^2\sigma^2}{2}, \tag{3.12}$$

so that

$$f^* = \frac{\mu}{\sigma^2}, \quad g_R(f^*) = \frac{\mu^2}{2\sigma^2},$$

and the NS-NL condition is

$$0 < \mu < \sigma^2.$$

Even though these results are not especially sophisticated, we will see in the next section (Theorem 4.1) that the process (3.11) naturally appears as the continuous-time, or high frequency, limit of discrete-time compounding for a large class of returns.

On the other hand, if we assume that the process R is purely discontinuous, with jumps $\triangle R_k = r_k$ at times $s = k \in \{1, 2, 3, \ldots\}$, then

$$R_t = 0,\ t \in (0,1), \quad R_t = \sum_{k=1}^{\lfloor t \rfloor} r_k = \sum_{0<s\leq t} \triangle R_s,\ t \geq 1,$$

and (3.4) becomes (2.1). Accordingly, we will now investigate the general case (3.4) when the process R has both a continuous component and jumps. To this end, we use [8, Proposition II.1.16] and introduce the jump measure $\mu^R = \mu^R(dx, ds)$ of the process R by putting a point mass at every point in space-time where the process R has a jump:

$$\mu^R(dx, ds) = \sum_{s>0} \delta_{(\triangle R_s, s)}(dx, ds); \tag{3.13}$$

note that both the time s and size $\triangle R_s$ of the jump can be random. In particular, with (3.5) in mind,

$$\sum_{0<s\leq t} \Big(\ln(1 + f\triangle R_s) - f\triangle R_s\Big) = \int_0^t \fint_{-1}^{+\infty} \big(\ln(1 + fx) - fx\big)\mu^R(dx, ds); \tag{3.14}$$

here and below,

$$\fint_a^b, \quad a < 0 < b, \tag{3.15}$$

stands for

$$\int\limits_{(a,0)\bigcup(0,b)}.$$

By [8, Proposition II.2.9 and Theorem II.2.34], and keeping in mind (3.5), we get the following generalization of (3.9):

$$\begin{aligned} R_t = A_t + R_t^c + \int_0^t \int_1^{+\infty} x\mu^R(dx, ds) \\ + \int_0^t \int_{-1}^1 x\big(\mu^R(dx, ds) - \nu(dx, s)da_s\big), \end{aligned} \tag{3.16}$$

where $a = a_t$ is a predictable non-decreasing process and $\nu = \nu(dx, t)$ is the non-negative random time-dependent measure on $(-1, 0)\bigcup(0, +\infty)$ with the property

$$\int_{-1}^{+\infty} \min(1, x^2)\nu(dx, t) \leq 1$$

for all $t \geq 0$ and $\omega \in \Omega$. Moreover

$$A_t = \int_0^t \mu_s\, da_s \ \text{ for some predictable process } \mu = \mu_t, \tag{3.17}$$

$$\langle R^c\rangle_t = \int_0^t \sigma_s^2\, da_s \ \text{ for some predictable process } \sigma = \sigma_t, \tag{3.18}$$

and the process

$$t \mapsto \int_0^t \int_{-1}^{+\infty} h(x)\big(\mu^R(dx, ds) - \nu(dx, s)da_s\big)$$

is a martingale for every bounded measurable function h such that $\limsup\limits_{x\to 0} \frac{|h(x)|}{|x|} < \infty$.

To proceed, we assume that

$$\mathbb{E}\int_0^t \int_{-1}^{+\infty} |\ln(1 + x)|\, \nu(dx, s)\, da_s < \infty,\ t > 0,$$

which is a generalization of condition (2.4). Then, by [8, Theorem II.1.8], the process

$$t \mapsto \int_0^t \int_{-1}^{+\infty} \ln(1 + x)\big(\mu^R(dx, ds) - \nu(dx, s)da_s\big)$$

is a martingale.

Next, we combine (3.6), (3.14), and (3.16), and re-arrange the terms so that the logarithm of the wealth process becomes

$$\begin{aligned} \ln W_t^f = fA_t + fR_t^c - \frac{f^2}{2}\langle R_t^c\rangle - f\int_0^t \int_{-1}^1 x\nu(dx, s)da_s \\ + \int_0^t \int_{-1}^{+\infty} \ln(1 + fx)\nu(dx, s)da_s + M_t^f, \end{aligned} \tag{3.19}$$

where

$$M_t^f = \int_0^t \int_{-1}^{+\infty} \ln(1 + fx)\big(\mu^R(dx, ds) - \nu(dx, s)da_s\big).$$

In general, for equality (3.19) to hold, we need to make an additional assumption

$$\int_{-1}^{1} x\nu(dx,t) < \infty \tag{3.20}$$

for all $t \geq 0$ and $\omega \in \Omega$.

In the particular case (2.1),

- $a_s = \lfloor s \rfloor$ is the step function, with unit jumps at positive integers, so that da_s is the collection of point masses at positive integers;
- $\nu(dx,s) = F^R(dx)$, where F^R is the cumulative distribution function of the random variable r, so that (3.20) holds automatically;
- $\mu_t = g_f(r) + \int_{-1}^{1} x\, F^R(dx)$, $R^c_t = 0$, $\sigma_t = 0$;
- $M^f_t = \sum_{0<k\leq t} \big(\ln(1+fr_k) - g_r(f) \big)$;
- condition (3.25) is (2.9).

A natural way to reconcile (3.10) with (3.17), (3.18) is to take $\mu_t = \mu$, $\sigma_t = \sigma$ for some non-random numbers $\mu \in \mathbb{R}$, $\sigma \geq 0$, and a non-random non-decreasing function $a = a_t$ with the property

$$\lim_{t\to+\infty} \frac{a_t}{t} = 1. \tag{3.21}$$

Then, to have a non-random almost-sure limit

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \int_{-1}^{+\infty} \varphi(x)\nu(dx,s)da_s$$

for a sufficiently rich class of non-random test functions φ, we have to assume that there exists a non-random non-negative measure $F^R = F^R(dx)$ on $(-1,0)\bigcup(0,+\infty)$ such that

$$\int_{-1}^{+\infty} \min(|x|,1)\, F^R(dx) < \infty \tag{3.22}$$

and, for large s,

$$\nu(dx,s) \approx F^R(dx).$$

As a result, if

$$\nu(dx,s) = F^R(dx) \tag{3.23}$$

for all s, then

$$A_t = \mu a_t,\ \langle R^c \rangle_t = \sigma^2 a_t,\ \nu(dx,t) = F^R(dx)a_t \tag{3.24}$$

are all non-random, and [8, Theorem II.4.15] implies that R is a *process with independent increments*. Furthermore, (3.24) and the strong law of large numbers for martingales imply

$$\mathbb{P}\left(\lim_{t\to\infty} \frac{R^c_t}{t} = 0 \right) = 1;$$

cf. [12, Corollary 1 to Theorem II.6.10]. Similarly, if

$$\int_{-1}^{+\infty} \ln^2(1+x)\, F^R(dx) < \infty, \tag{3.25}$$

then M^f is a square-integrable martingale and

$$\mathbb{P}\left(\lim_{t\to\infty}\frac{M_t^f}{t}=0\right)=1.$$

Writing

$$\bar{\mu}=\mu-\int_{-1}^{1}xF^R(dx)$$

the long-term growth rate (3.7) becomes

$$g_R(f)=f\bar{\mu}-\frac{f^2\sigma^2}{2}+\int_{-1}^{\infty}\ln(1+fx)F^R(dx), \tag{3.26}$$

which does include both (2.5) and (3.12) as particular cases. By direct computation, the function $f\mapsto g_R(f)$ is concave and the domain of the function contains $[0,1]$.

Similar to Proposition 2.5, we have the following result.

THEOREM 3.1. *Consider continuous-time compounding with return process*

$$\begin{aligned} R_t=A_t+R_t^c&+\int_0^t\int_1^{+\infty}x\mu^R(dx,ds)\\ &+\int_0^t\int_{-1}^{1}x\big(\mu^R(dx,ds)-\nu(dx,s)da_s\big), \end{aligned} \tag{3.27}$$

where the random measure μ^R is from (3.13), and assume that equalities (3.21) and (3.24) hold. If F^R satisfies (3.22), (3.25), and

$$\lim_{f\to0+}\int_{-1}^{\infty}\frac{x}{1+fx}F^R(dx)>-\bar{\mu},$$
$$\lim_{f\to1-}\int_{-1}^{\infty}\frac{x}{1+fx}F^R(dx)<\sigma^2-\bar{\mu},$$

then the long-term growth rate is given by (3.26), and there exists a unique $f^*\in(0,1)$ such that

$$g_R(f)<g_R(f^*)$$

for all f in the domain of g_R. The number f^* is the unique solution of the equation $g_R'(f)=0$, that is,

$$\bar{\mu}-f^*\sigma^2+\int_{-1}^{\infty}\frac{x}{1+xf^*}F^R(dx)=0. \tag{3.28}$$

By the Lebesgue decomposition theorem, the measure corresponding to the function $a=a_t$ has a discrete, absolutely continuous, and singular components. With (3.21) in mind, a natural choice of the discrete component is $a_t=\lfloor t\rfloor$, which, as we saw, corresponds to discrete compounding discussed in the previous section. A natural choice of the absolutely continuous component is $a_t=t$. Then $A_t=\mu t$, $R_t^c=\sigma B_t$, $\nu(dx,t)da_t=F^R(dx)\,dt$, where B is a standard Brownian motion. By [8, Corollary II.4.19], we conclude that the process R has independent and stationary increments, that is, *R is a Lévy process.* In this case, equality (3.27) is known as the Lévy-Itô decomposition of the process R; cf. [19, Theorem 19.2].

We do not consider the singular case in this paper and leave it for future investigation.

4. Continuous Limit of Discrete Compounding.

4.1. A (Simple) Random Walk Model. Following the methodology in [2] and [21, Section 7.1], we assume compounding a sufficiently large number n of bets in a time period $[0, T]$. The returns $r_{n,1}, r_{n,2}, \ldots$ of the bets are

$$r_{n,k} = \frac{\mu}{n} + \frac{\sigma}{\sqrt{n}}\,\xi_{n,k} \tag{4.1}$$

for some $\mu > 0$, $\sigma > 0$ and independent identically distributed random variables $\xi_{n,k},\ k = 1, 2, \ldots,$ with mean 0 and variance 1. The classical simple random walk corresponds to $\mathbb{P}(\xi_{n,k} = \pm 1) = 1/2$ and can be considered a *high frequency* version of (1.1). Similar to (2.2), we need $r_{n,k} \geq -1$, which, in general, can only be achieved with *uniform boundedness* of $\xi_{n,k}$:

$$|\xi_{n,k}| \leq C_0, \tag{4.2}$$

and then, with no loss of generality, we assume that n is large enough so that

$$|r_{n,k}| \leq \frac{1}{2}. \tag{4.3}$$

Similar to (1.2), a condition to have an edge is

$$\mathbb{E}[r_{n,k}] = \frac{\mu}{n} > 0,$$

and, similar to (2.1), given n bets per unit time period, with exposure $f \in [0, 1]$ in each bet, we get the following formula for the total wealth $W_t^{n,f}$ at time $t \in (0, T]$ assuming $W_0 = 1$:

$$W_t^{n,f} = \prod_{k=1}^{\lfloor nt \rfloor} \big(1 + f r_{n,k}\big); \tag{4.4}$$

$\lfloor nt \rfloor$ denotes the largest integer less than nt. Denote by $g^n(f)$ the corresponding long-term growth rate:

$$g^n(f) = \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{\lfloor nt \rfloor} \ln\big(1 + f r_{n,k}\big).$$

Using the law of large numbers, and with notation (2.5) in mind,

$$g^n(f) = n\mathbb{E}\ln(1 + f r_{n,1}) = n g_{r_n,1}(f). \tag{4.5}$$

Let

$$f_n^* = \arg\max_{f\in[0,1]} g^n(f) \tag{4.6}$$

be the value of f maximizing g^n. What can we say about f_n^* as $n \to \infty$?

As a motivation, consider the high-frequency version of the simple Bernoulli model (1.1):

$$\mathbb{P}\left(r_{n,k} = \frac{\mu}{n} \pm \frac{\sigma}{\sqrt{n}}\right) = \frac{1}{2}, \tag{4.7}$$

which, for fixed n, is is a particular case of the general Bernoulli model (2.14) with $p=q=1/2$,

$$a=\frac{\sigma}{\sqrt{n}}-\frac{\mu}{n},\ b=\frac{\sigma}{\sqrt{n}}+\frac{\mu}{n}.$$

Then, by direct computation,

$$f_n^*=\frac{\mu}{\sigma^2-(\mu^2/n)}\to\frac{\mu}{\sigma^2},\ n\to\infty,$$

and

$$\lim_{n\to\infty} g^n(f_n^*)=\frac{\mu^2}{2\sigma^2}.$$

We will now show that all models of the form (4.1) behave this way in the high frequency limit.

Let $B=B_t,\ t\geq 0$, be a standard Brownian motion on a stochastic basis $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P})$ satisfying the usual conditions, and define the process

$$W_t^f=\exp\left(\left(f\mu-\frac{f^2\sigma^2}{2}\right)t+f\sigma B_t\right). \tag{4.8}$$

Note that (4.8) is the same as (3.11).

Similar to (1.3), we define the long-term continuous time growth rate

$$g(f)=\lim_{t\to\infty}\frac{1}{t}\ln W_t^f.$$

Then a simple computation show that

$$g(f)=f\mu-\frac{f^2\sigma^2}{2},$$

and so

$$f^*=\frac{\mu}{\sigma^2} \tag{4.9}$$

achieves the maximal long-term continuous time growth rate

$$g(f^*)=\frac{\mu^2}{2\sigma^2}. \tag{4.10}$$

The NS-NL condition $f^*\in(0,1)$ holds if $0<\mu<\sigma^2$, which, to the order $1/n$, is consistent with (2.12) and (2.13), when applied to (4.1):

$$\mathbb{E}[r_{n,k}]=\frac{\mu}{n},\ \ \mathbb{E}\left[\frac{r_{n,k}}{1+r_{n,k}}\right]=\frac{\mu-\sigma^2}{n}+o(n^{-1}).$$

The wealth process (4.8) is that of someone who is "continuously" placing bets, that is, adjusts the positions instantaneously, and, for large n, turns out to be a good approximation of high frequency betting (4.4).

THEOREM 4.1. *For every $T>0$ and every $f\in[0,1]$, the sequence of processes $\big(W_t^{n,f},\ n\geq 1,\ t\in[0,T]\big)$ converges in law to the process $W^f=W_t^f,\ t\in[0,T]$, and the convergence is uniform in f on compact subsets of $(0,1)$. In particular, with f_n^* defined in (4.6), if $0<\mu<\sigma^2$, then*

$$\lim_{n\to\infty} f_n^*=\frac{\mu}{\sigma^2},\qquad \lim_{n\to\infty} g^n(f_n^*)=\frac{\mu^2}{2\sigma^2}. \tag{4.11}$$

Proof. Writing

$$Y_t^{n,f} = \ln W_t^{n,f},$$

the objective is to show weak convergence, as $n \to \infty$, of $Y^{n,f}$ to the process

$$Y_t^f = \left(f\mu - \frac{f^2\sigma^2}{2} \right) t + f\sigma B_t, \quad t \in [0,T],$$

and that the convergence is uniform in f over compact subsets of $(0,1)$. The proof relies on the method of predictable characteristics for semimartingales from [8]. More specifically, we make suitable changes in the proof of Corollary VII.3.11.

By (4.8)

$$Y_t^{n,f} = \sum_{k=1}^{\lfloor nt \rfloor} \ln(1 + f r_{n,k}).$$

Then (4.1) and (4.2) imply

$$\mathbb{E}\left(Y_t^{n,f} - \mathbb{E}Y_t^{n,f} \right)^4 \le \frac{C_0^4\sigma^4}{n^2}\big(nT + 3nT(nT-1)\big) \le 3C_0^4\sigma^4T^2,$$

from which uniform integrability of the family $\{Y_t^{n,f},\ n \ge 1,\ t \in [0,T]\}$ follows.

Then, by [8, Theorem VII.3.7], it suffices to confirm that the following results hold uniformly in f over compact subsets of $(0,1)$:

$$\lim_{n\to\infty} \sup_{t \le T} \left| \lfloor nt \rfloor \mathbb{E}\big[\ln(1+fr_{n,1})\big] - \left(f\mu - \frac{f^2\sigma^2}{2} \right) t \right| = 0, \tag{4.12}$$

$$\lim_{n\to\infty} \lfloor nt \rfloor \left(\mathbb{E}\big(\ln(1+fr_{n,1})\big)^2 - \Big(\mathbb{E}\big[\ln(1+fr_{n,1})\big] \Big)^2 \right) = f^2\sigma^2 t,\ t \in [0,T], \tag{4.13}$$

$$\lim_{n\to\infty} \lfloor nt \rfloor \mathbb{E}\big[\phi\big(\ln(1+fr_{n,1})\big)\big] = 0,\ t \in [0,T]. \tag{4.14}$$

Equality (4.14) must hold for all functions $\phi = \phi(x),\ x \in \mathbb{R}$, that are continuous and bounded on $\mathbb{R}$ and satisfy $\phi(x) = o(x^2),\ x \to 0$, that is,

$$\lim_{x\to 0} \frac{\phi(x)}{x^2} = 0. \tag{4.15}$$

Equalities (4.12) and (4.13) follow from

$$r_{n,1}^2 = \frac{\sigma^2}{n}\,\xi_{n,1}^2 + \frac{2\mu\sigma\xi_{n,1}}{n^{3/2}} + \frac{\mu^2}{n^2},$$

together with (4.3) and an elementary inequality

$$\left| \ln(1+x) - x - \frac{x^2}{2} \right| \le |x|^3,\ |x| \le \frac{1}{2}.$$

In particular,

$$\mathbb{E}\big[\big(\ln(1+fr_{n,1})\big)^2\big] = \frac{f^2\sigma^2}{n} + o(1/n),\ n \to +\infty.$$

To establish (4.14), note that (4.15) and (4.1) imply

$$\phi\big(\ln(1+fr_{n,1})\big) = o(1/n),\ n \to +\infty.$$

Now both equalities in (4.11) follow because $g^n(f) = \mathbb{E}Y_1^{n,f}$ by (4.5), whereas $g(f) = \mathbb{E}Y_1^f$ and so $\lim_{n\to\infty} g^n(f) = g(f)$ uniformly in f on compact subsets of $(0,1)$. □

With natural modifications, Theorem 4.1 extends to the setting (2.16).

THEOREM 4.2. *Assume that*

$$r_{n,k} + 1 = \exp\left(\frac{b}{n} + \frac{\sigma}{\sqrt{n}}\,\xi_{n,k}\right), \tag{4.16}$$

where $b \in \mathbb{R}$, $\sigma > 0$, and, for each $n \geq 1$, $k \leq n$, the random variables $\xi_{n,k}$ are independent and identically distributed, with zero mean, unit variance, and, for every $a > 0$,

$$\lim_{n\to\infty} n\,\mathbb{E}\big[|\xi_{n,1}|^2 I(|\xi_{n,1}| > a\sqrt{n})\big] = 0. \tag{4.17}$$

Define

$$\mu = b + \frac{\sigma^2}{2}.$$

Then the conclusions of Theorem 4.1 hold.

Proof. Even though a formal Taylor expansion suggests

$$r_{n,k} = \frac{\mu}{n} + \frac{\sigma}{\sqrt{n}}\,\xi_{n,k} + o(1/n),$$

we cannot apply Theorem 4.1 directly because the random variables $\xi_{n,k}$ are not necessarily uniformly bounded. Still, condition (4.17) makes it possible to verify conditions (4.12)–(4.14). □

Comparing Theorems 4.1 and 4.2, we see at least two reasons why Theorem 4.2 could be more useful in the study of high-frequency compounding: (a) an analytic expression for $g^n(f)$ using (4.5) can often exist for $r_{n,k}$ from (4.1) but not for $r_{n,k}$ from (4.16); (b) The class of admissible random variables $\xi_{n,k}$ is bigger in the case (4.16).

Condition (4.17) is clearly satisfied when $\xi_{n,k}$, $k = 1, 2, \ldots$, are iid standard normal, which corresponds to

$$r_{n,k} = \frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}} \tag{4.18}$$

and

$$P_t = e^{bt+\sigma B_t}. \tag{4.19}$$

Thus, while the exponential model (2.16) with log-normal returns is not solvable in closed form, the high-frequency version leads to the (approximately) optimal strategy

$$f^* = \frac{b}{\sigma^2} + \frac{1}{2}, \tag{4.20}$$

and, under (2.19), the NS-NL condition holds: $f^* \in (0,1)$. Monte Carlo simulations with $\sigma = 1$ and $n = 10$ show that the values of the corresponding optimal f_{10}^* are very close to those given by (4.20) for all $b \in (-1/2, 1/2)$.

Informally, both Theorems 4.1 and 4.2 can be considered as particular cases of the delta method for the Donsker theorem with drift: if the sequence of processes

$$t \mapsto \sum_{k=1}^{\lfloor nt \rfloor} \xi_{n,k}$$

converges, as $n \to \infty$, to the processes $t \mapsto bt + \sigma B_t$ and $\varphi = \varphi(x)$ is a suitable function with $\varphi(0) = 0$, then one would expect the sequence of processes

$$t \mapsto \sum_{k=1}^{\lfloor nt \rfloor} \varphi\big(\xi_{n,k}\big)$$

to converge to the process

$$t \mapsto \left(\varphi'(0)b + \frac{\varphi''(0)\sigma^2}{2}\right) t + |\varphi'(0)|\sigma B_t.$$

4.2. Beyond the Log-Normal Limit. With the results of Section 3 in mind, we consider the following generalization of (4.18): (4.19):

$$r_{n,k} = \frac{P_{k/n} - P_{(k-1)/n}}{P_{(k-1)/n}},\ k = 1, 2, \dots,$$

where the process $P = P_t,\ t \geq 0$, has the form $P_t = e^{R_t}$, and $R = R_t$ is a Lévy process. In other words,

$$r_{n,k} = e^{R_{k/n} - R_{(k-1)/n}} - 1. \tag{4.21}$$

As in (3.27), the process $R = R_t$ can be decomposed into a drift, diffusion/small jump, and large jump components according to the Lévy-Itô decomposition [19, Theorem 19.2]:

$$R_t = \mu t + \sigma\, B_t + \int_0^t \fint_{-1}^{1} x\big(\mu^R(dx, ds) - F^R(dx)ds\big) + \int_0^t \int_{|x|>1} x\mu^R(dx, ds); \tag{4.22}$$

we continue to use the notation $\fint$ first introduced in (3.15).

The function F^R in (4.22) is a non-random non-negative measure on $(-\infty, 0) \cup (0, +\infty)$ such that

$$\fint_{-\infty}^{+\infty} \min(x^2, 1) F^R(dx) < \infty.$$

Now that the process R_t is exponentiated,

- there is no need to assume that $\triangle R_t \geq -1$;
- the analog of (3.25) becomes $\mathbb{E}|R_1| < \infty$, that is

$$\int_{|x|>1} |x| F^R(dx) < \infty.$$

Equality (4.22) has a natural interpretation in terms of financial risks [20]: the drift represents the edge ("guaranteed" return), diffusion and small jumps represent small fluctuations of returns, and the large jump component represents (sudden) large changes in returns. Similar to (4.4), the corresponding wealth process is

$$W_t^{n,f} = \prod_{k=1}^{\lfloor nt \rfloor} \big(1 + f r_{n,k}\big). \tag{4.23}$$

Denote by $\mathbb{D}((0,T))$ the Skorohod space on $(0,T)$ [12, Section 6.1]. We have the following generalization of Theorem 4.2.

THEOREM 4.3. *Consider the family of processes $W^{n,f} = W_t^{n,f}$, $t \in [0,T]$, $n \ge 1$, $f \in [0,1]$, defined by (4.23). If $r_{n,k}$ is given by (4.21), with $P_t = e^{R_t}$, and $R = R_t$ is a Lévy process with representation (4.22) and $\mathbb{E}|R_1| < \infty$, then, for every $f \in [0,1]$ and $T > 0$,*

$$\lim_{n\to\infty} W^{n,f} \overset{\mathcal{L}}{=} W^f$$

in $\mathbb{D}((0,T))$, where

$$\begin{aligned} W_t^f = \exp\bigg(& f R_t + \frac{f(1-f)\sigma^2}{2}\, t \\ & + \int_0^t \int_{-\infty}^{+\infty} \Big[\ln\big(1 + f(e^x - 1)\big) - f x\Big] \mu^R(dx, ds)\bigg). \end{aligned} \tag{4.24}$$

The convergence is uniform in f on compact subsets of $(0,1)$.

Proof. By (4.21) and (4.23),

$$\ln W_t^{n,f} = \sum_{k=1}^{\lfloor nt \rfloor} \ln\Big(1 + f\big(e^{R_{k/n} - R_{(k-1)/n}} - 1\big)\Big).$$

Step 1: For $s \in \big(\frac{k-1}{n}, \frac{k}{n}\big]$, let

$$r_s^{n,k} = e^{R_s - R_{(k-1)/n}} - 1, \tag{4.25}$$

and apply the Itô's formula [17, Theorem II.32] to the process

$$s \mapsto \ln\big(1 + f r_s^{n,k}\big), \quad s \in \bigg(\frac{k-1}{n}, \frac{k}{n}\bigg].$$

The result is

$$\begin{aligned} \ln\big(1 + f r_s^{n,k}\big) = & \int_{\frac{k-1}{n}}^{s} \frac{f(1 + r_{u-}^{n,k})}{1 + f r_{u-}^{n,k}}\, dR_u + \frac{\sigma^2}{2} \int_{\frac{k-1}{n}}^{s} \frac{f(1-f)(1 + r_{u-}^{n,k})}{\big(1 + f r_{u-}^{n,k}\big)^2}\, du \\ & + \int_{\frac{k-1}{n}}^{s} \int_{-\infty}^{+\infty} \Big[\ln\big(1 - f + f e^x (r_{u-}^{n,k} + 1)\big) \\ & - \ln(1 + f r_{u-}^{n,k}) - x \frac{f(1 + r_{u-}^{n,k})}{1 + f r_{u-}^{n,k}}\Big] \mu^R(dx, du). \end{aligned}$$

Step 2: Putting $s = \frac{k}{n}$ in the above equality and summing over k, we derive the following expression for $\ln W_t^{n,f}$:

$$\begin{aligned}\ln W_t^{n,f} &= \sum_{k=1}^{\lfloor nt\rfloor}\Big(\int_{\frac{k-1}{n}}^{\frac{k}{n}} h_{n,k}^{(1)}(s)\,dR_s + \int_{\frac{k-1}{n}}^{\frac{k}{n}} h_{n,k}^{(2)}(s)\,ds \\ &\quad + \int_{\frac{k-1}{n}}^{\frac{k}{n}} \fint_{-\infty}^{+\infty} h_{n,k}^{(3)}(s,x)\mu^R(dx,du)\Big) \\ &= \int_0^t H_{n,t}^{(1)}(s)\,dR_s + \int_0^t H_{n,t}^{(2)}(s)\,ds + \int_0^t \fint_{-\infty}^{+\infty} H_{n,t}^{(3)}(s,x)\mu^R(dx,ds)\,, \qquad (4.26)\end{aligned}$$

where

$$h_{n,k}^{(1)}(s) = \frac{f(1+r_{s-}^{n,k})}{1+fr_{s-}^{n,k}}\,, \quad h_{n,k}^{(2)}(s) = \frac{\sigma f(1-f)}{2}\frac{1+r_{s-}^{n,k}}{(1+fr_{s-}^{n,k})^2}\,,$$

$$h_{n,k}^{(3)}(s,x) = \ln\big(1-f+fe^x(r_{s-}^{n,k}+1)\big) - \ln(1+fr_{s-}^{n,k}) - fx\frac{1+r_{s-}^{n,k}}{1+fr_{s-}^{n,k}}\,;$$

$$H_{n,t}^{(i)}(s) = \sum_{k=1}^{\lfloor nt\rfloor} h_{n,k}^{(i)}(s)\mathbf{1}_{(\frac{k-1}{n},\frac{k}{n}]}(s),\ i=1,2;\quad H_{n,t}^{(3)}(s,x) = \sum_{k=1}^{\lfloor nt\rfloor} h_{n,k}^{(3)}(s,x)\mathbf{1}_{(\frac{k-1}{n},\frac{k}{n}]}(s).$$

<u>Step 3:</u> Because

$$\lim_{n\to\infty,\,k/n\to s} R_{(k-1)/n} = R_{s-},$$

equality (4.25) implies

$$\lim_{n\to+\infty,\,k/n\to s} r_{s-}^{n,k} = 0$$

for all s. Consequently, we have the following convergence in probability:

$$\lim_{n\to+\infty} H_{n,t}^{(1)}(s) = f,\ \lim_{n\to+\infty} H_{n,t}^{(2)}(s) = \frac{\sigma^2 f(1-f)}{2},$$
$$\lim_{n\to+\infty} H_{n,t}^{(2)}(s,x) = \ln\big(1+f(e^x-1)\big) - fx.$$

To pass to the corresponding limits in (4.26), we need suitable bounds on the functions $H^{(i)}$, $i=1,2,3$.

Using the inequalities

$$0 < \frac{1+y}{1+ay} \le \frac{1}{a},\quad 0 < \frac{1+y}{(1+ay)^2} \le \frac{1}{4a(1-a)},\qquad y>-1,\ a\in(0,1),$$

we conclude that

$$0 < h_{n,k}^{(1)}(s) \le 1,\ 0 < h_{n,k}^{(2)}(s) \le \sigma^2,$$

and therefore

$$0 < H_{n,t}^{(1)}(s) \le 1,\ 0 < H_{n,t}^{(2)}(s) \le \sigma^2. \qquad (4.27)$$

Similarly, for $f\in(0,1)$ and $y>-1$,

$$\left|\ln\frac{1-f+fe^x(y+1)}{1+fy} - fx\frac{1+y}{1+fy}\right| \le 2\big(|x|\wedge|x|^2\big), \qquad (4.28)$$

so that

$$|h^{(3)}_{n,k}(s,x)| \le 2\big(|x| \wedge |x|^2\big)$$

and

$$\big|H^{(3)}_{n,t}(s)\big| \le 2\big(|x| \wedge |x|^2\big). \tag{4.29}$$

To verify (4.28), fix $f \in (0,1)$ and $y > -1$, and define the function

$$z(x) = \ln \frac{1 - f + fe^x(y+1)}{1+fy},\ x \in \mathbb{R}.$$

By direct computation,

$$\begin{aligned} z(0) &= 0,\\ z'(x) &= \frac{fe^x(y+1)}{1-f+fe^x(y+1)} = 1 - \frac{1-f}{1-f+fe^x(y+1)},\\ z'(0) &= \frac{f(y+1)}{1+fy}, \end{aligned}$$

so that, using the Taylor formula,

$$\ln \frac{1-f+fe^x(y+1)}{1+fy} - fx\frac{1+y}{1+fy} = z(x) - z(0) - xz'(0) = \int_0^x (x-u)z''(u)du. \tag{4.30}$$

It remains to notice that

$$0 \le z'(x) \le 1,\quad 0 \le z''(x) \le 1,$$

and then (4.28) follows from (4.30).

With (4.27) and (4.29) in mind, the dominated convergence theorem [17, Theorem IV.32] makes it possible to pass to the limit in probability in (4.26); the convergence in the space $\mathbb{D}$ then follows from the general results of [8, Section IX.5.12]. ◻

The following is a representation of the long-term growth rate of the limiting wealth process W^f.

THEOREM 4.4. *Let $R = R_t$ be a Lévy process with representation (4.22). If $\mathbb{E}|R_1| < \infty$, then the process $W^f = W^f_t$ defined in (4.24) satisfies*

$$\begin{aligned} \lim_{t\to+\infty} \frac{\ln W^f_t}{t} &= f\Big(\mu + \int_{|x|>1} xF^R(dx)\Big) + \frac{f(1-f)\sigma^2}{2}\\ &\quad + \int_{-\infty}^{+\infty} \big[\ln\big(1 + f(e^x-1)\big) - fx\big]F^R(dx). \end{aligned} \tag{4.31}$$

Proof. By (4.24),

$$\frac{\ln W^f_t}{t} = f\frac{R_t}{t} + \frac{f(1-f)\sigma^2}{2} + \frac{1}{t}\int_0^t \int_{-\infty}^{+\infty} \Big[\ln\big(1+f(e^x-1)\big) - fx\Big]\mu^R(dx,ds).$$

It remains to apply the law of large numbers for Lévy processes [19, Theorem 36.5].◻

Similar to Proposition 2.5, we also have the following result.

THEOREM 4.5. *In the setting of Theorem 4.4, denote the right-hand side of (4.31) by $g_R(f)$ and assume that*

$$\lim_{f\to 0+} \int_{-\infty}^{+\infty} \left(\frac{e^x-1}{1+f(e^x-1)} - x\right) F^R(dx) > -\left(\mu + \frac{\sigma^2}{2} + \int_{|x|>1} xF^R(dx)\right),$$
$$\lim_{f\to 1-} \int_{-\infty}^{+\infty} \left(\frac{e^x-1}{1+f(e^x-1)} - x\right) F^R(dx) < -\left(\mu + \int_{|x|>1} xF^R(dx)\right),$$

Then there exists a unique $f^* \in (0,1)$ such that

$$g_R(f) < g_R(f^*)$$

for all f in the domain of g_R. The number f^* is the unique solution of the equation

$$\mu + \int_{|x|>1} x dF^R(x) + \frac{\sigma^2}{2} - \sigma^2 f^* + \int_{-\infty}^{+\infty} \left(\frac{e^x-1}{1+f^*(e^x-1)} - x\right) dF^R(x) = 0. \tag{4.32}$$

In particular, if $\sigma > 0$ and

$$f_0^* = \frac{\mu}{\sigma^2} + \frac{1}{2},$$

then (4.32) becomes a generalization of (4.20):

$$f^* = f_0^* + \frac{1}{\sigma^2}\int_{|x|>1} x dF^R(x) + \frac{1}{\sigma^2}\int_{-\infty}^{+\infty} \left(\frac{e^x-1}{1+f^*(e^x-1)} - x\right) dF^R(x). \tag{4.33}$$

To conclude the section, let us compare and contrast Theorems 3.1 and 4.4. In Theorem 3.1, the process R is a semimartingale with independent but not necessarily stationary increments, that is, more general than a Lévy process. The key technical assumption (3.22) means that the jump component of R has bounded variation, something that is not required in Theorem 4.4. If, in the setting of Theorem 4.4, we additionally assume that the (small-jump) component of R has bounded variation:

$$\int_{-1}^{1} |x| F^R(dx) < \infty,$$

then, after a logarithmic substitution and re-arrangement of terms, (4.31) becomes (3.26); in particular, the finite moment condition in Theorem 4.4 becomes equivalent to (3.25). On the other hand, equality (3.26) is derived for a wider class of return processes that includes Lévy processes as a particular case. To summarize, Theorem 4.4 provides a stronger result in the case of Lévy processes, whereas Theorems 3.1 covers a bigger class of processes under an additional integrabililty condition.

5. Continuous Limit of Random Discrete Compounding. The objective of this section is to analyze high frequency limits for betting *in business time.* In other words, the number of bets is not known a priori, so that a natural model of the corresponding wealth process is

$$W_t^{n,f} = \prod_{k=1}^{\lfloor \Lambda_{n,t} \rfloor} (1 + f r_{n,k}) \tag{5.1}$$

where, for each n, the process $t \mapsto \Lambda_{n,t}$ is a subordinator, that is, a non-decreasing Lévy process, independent of all $r_{n,k}$. In a typical application, for example, investment for retirement, $\Lambda_{n,t}$ represents an uncertain time horizon.

To study (5.1), we will follow the methodology in [11], where convergence of processes is derived after *assuming* a suitable convergence of the random variables. The main result in this connection is as follows.

THEOREM 5.1. *Consider the following objects:*

- *random variables* $X_{n,k},\ n,k \geq 1$ *such that* $\{X_{n,k},\ k \geq 1\}$ *are iid for each* n, *with mean zero and, for some* $\beta \in [0,1]$, $m_n := \Big(\mathbb{E}|X_{n,1}|^{\beta}\Big)^{1/\beta} < \infty$;
- *random processes* $\Lambda_n = \Lambda_{n,t},\ n \geq 1,\ t \geq 0$, *such that, for each* n, Λ_n *is a subordinator independent of* $\{X_{n,k},\ k \geq 1\}$ *with the properties* $\Lambda_{n,0} = 0$, *and for some numbers* $0 < \delta, \delta_1 \leq 1$ *and* $C_n > 0$, $\Big(\mathbb{E}\Lambda_{n,t}^{\delta}\Big)^{1/\delta} \leq C_n t^{\delta_1/\delta}$.

Assume that there exist infinitely divisible random variables Y *and* U *such that*

$$\lim_{n\to\infty} \sum_{k=1}^{n} X_{n,k} \overset{d}{=} \bar{Y},\ \lim_{n\to\infty} \frac{\Lambda_{n,1}}{n} \overset{d}{=} \bar{U}.$$

If

$$\sup_n \Big(C_n m_n^{\beta}\Big) < \infty, \tag{5.2}$$

then, as $n \to \infty$, *the sequence of processes*

$$t \mapsto \sum_{k=1}^{\lfloor \Lambda_{n,t} \rfloor} X_{n,k},\ t \in [0,T],$$

converges, in the Skorokhod topology, to the process $Z = Z_t$ *such that* $Z_t = Y_{U_t}$, *where* Y *and* U *are independent Lévy processes satisfying* $Y_1 \overset{d}{=} \bar{Y}$ *and* $U_1 \overset{d}{=} \bar{U}$.

The proof is a word-for-word repetition of the arguments leading to [11, Theorem 1]: the result of [6], together with the assumptions of the theorem, implies

$$\lim_{n\to\infty} \sum_{k=1}^{\lfloor \Lambda_{n,1} \rfloor} X_{n,k} \overset{d}{=} Z_1,$$

and therefore the convergence of finite-dimensional distributions for the corresponding processes; together with condition (5.2), this implies the convergence in the Skorokhod space. Because we deal exclusively with Lévy processes, it is possible to avoid the heavy machinery from [8].

We now consider the wealth process (5.1) and apply Theorem 5.1 with

$$X_{n,k} = \ln(1 + f r_{n,k}) - \mathbb{E}\ln(1 + f r_{n,k}).$$

On the one hand, convergence to infinitely divisible distributions other than normal is a very diverse area, with a variety of conditions and conclusions; cf. [5, Chapter XVII, Section 5] or a summary in [10, Section 16.2]. On the other hand, optimal strategy (4.9) seems to persist.

For example, assume that the returns $r_{n,k}$ are as in (4.1), and let $\Lambda_{n,t} = S_{n^{\alpha}t}$, where $\alpha \in (0,1]$ and $S = S_t$ is the Lévy process such that S_1 has the α-stable distribution with both scale and skewness parameters equal to 1. Recall that an

α-stable Lévy process $L^\alpha = L^\alpha_t$ satisfies the following equality in distribution (as processes):

$$L^\alpha_{\gamma t} \overset{\mathcal{L}}{=} \gamma^{1/\alpha} L^\alpha_t,\ \gamma > 0. \tag{5.3}$$

Then

$$\Lambda_{n,t} \overset{\mathcal{L}}{=} n S_t$$

and, in the notations of Theorem 5.1, $\bar{Y}$ is normal with mean zero and variance σ^2. Keeping in mind that

$$\mathbb{E} \ln(1 + f r_{n,k}) = \mathbb{E} \ln(1 + f r_{n,1}) = \Big(f\mu - \frac{f^2\sigma^2}{2}\Big) n^{-1} + o(n^{-1}),$$

we repeat the arguments from [11, Example 1] to conclude that

$$\lim_{n\to\infty} \ln W^{n,f}_t \overset{\mathcal{L}}{=} \Big(f\mu - \frac{f^2\sigma^2}{2}\Big) S_t + Z_t,$$

where Z_1 has symmetric 2α-stable distribution. By (5.3),

$$S_t \overset{d}{=} t^{1/\alpha} S_1,\ \lim_{t\to+\infty} t^{-1/\alpha} Z_t \overset{d}{=} \lim_{t\to+\infty} t^{-1/(2\alpha)} Z_1 \overset{d}{=} 0,$$

and the "natural" long term growth rate becomes

$$\lim_{t\to\infty} t^{-1/\alpha} \Big(\lim_{n\to\infty} \ln W^{n,f}_t\Big) \overset{d}{=} \Big(f\mu - \frac{f^2\sigma^2}{2}\Big) S_1,$$

which is random, but, for each realization of S, is still maximized by f^* from (4.9). Therefore, if the time interval over which we compound our wealth is random, then the growth rate is also random as we do not know when compounding stops, yet, in the high frequency limit, this rate is still maximized by a deterministic fraction. Note that, for the purpose of this computation, the (stochastic) dependence between the processes S and Z is not important.

6. Conclusions And Further Directions. In the classical Markowitz portfolio theory [15], risk is formalized as the standard deviation σ of portfolio's return and reward is formalized as the mean of a portfolio's return μ. Through the construction of the Markowitz set of efficient portfolios, one can see a trade-off between risk and reward, namely, more risk implies more reward. A refinement of this idea is provided by the utility score

$$U = \mu - \frac{A}{2}\sigma^2,$$

with parameter $A \geq 0$ representing the degree of risk-aversion. Direct computations then show that, for a portfolio with fraction f invested in the risky asset, the corresponding utility score is maximized by

$$f^* = \frac{\mu}{A\sigma^2},$$

which, for $A = 1$ [a moderately risk-averse investor] coincides with (1.5). The same utility score, when combined with the original Kelly criterion, leads to optimization of the function

$$g_A(f) = \lim_{t\to+\infty} \left(\frac{1}{t}\left(\ln W^{n,f}_t - \frac{A}{2}\mathrm{Var}\Big(\ln W^{n,f}_t\Big)\right)\right)$$

and a formalization of the *fractional* Kelly criterion [21, Section 7.3]. In particular, when the wealth process is geometric Brownian motion (4.8), the maximum of the corresponding function g_A is achieved at the point

$$f_A^* = \frac{\mu}{(1+A)\sigma^2}. \tag{6.1}$$

Still, the traditional Markowitz formalization of risk versus reward is a static theory, as it only applies to a single bet with no consideration of compounding. The Kelly criterion is an example of a dynamic theory that takes into account compounding of returns and losses. Compared to straightforward optimization of the long-term growth [16], the parameter f in the Kelly criterion and the NS-NL condition $f^* \in [0,1]$ ensure that the corresponding strategy avoids ruin and the total wealth never reaches zero. Reference [1, Ch 5] provides a detailed, if somewhat informal, comparison between Markotitz and Kelly-type approaches to investment.

The NS-NL condition can fail in many situations. Even in the simple Bernoulli model, if $p < 1/2$, then the short position $f^* = 2p-1$ achieves positive long-time wealth growth:

$$g_r(f^*) = p\ln\frac{p}{1-p} + (2-p)\ln(2-2p) = \ln 2 + p\ln p + (1-p)\ln(1-p) > 0.$$

Note that $-p\ln p-(1-p)\ln(1-p)$ is the Shannon entropy of the Bernoulli distribution, and the largest value of the entropy is $\ln 2$, corresponding to $p = 1/2$. When the edge is too big (cf. (2.14)), then $f^* > 1$, that is, leveraged gambling leads to bigger long-time wealth growth than any NS-NL strategy. The economical and financial implications of $f^* \notin [0,1]$ are beyond the scope of our investigation and must be studied in a broader context of risk tolerance: even when $f^* \in (0,1)$, some investors might prefer the strategy based on a certain fraction of f^*, such as (6.1); cf. [21, Section 7.3].

A related observation, to be further studied in the future, is that high-frequency betting can lead to a more aggressive strategy than the "low frequency" counterpart. For example, comparing (1.1) and (4.7), we see that $\mu = 2p-1$ and $\sigma^2 = 4p(1-p) < 1$ when $p \neq 1/2$. As a result, by (4.9), the optimal strategy for (4.7) with large n is $f^* \approx (2p-1)/(4p(1-p)) > 2p-1$; recall that $f^* = 2p-1$ is the optimal strategy for the simple Bernoulli model (1.1). On the other hand, numerical simulations suggest that, in the log-normal model (4.18), (4.19), high-frequency compounding does not always lead to larger f^*.

Other problems warranting further investigation to build a theory which is also a practical guide to investment include

1. A dynamic strategy $f = f(t)$ with a predictable process f;
2. A portfolio of bets, with a vector of strategies $\mathbf{f} = (f_1, \ldots, f_N)$;
3. Convergence rates, both in the approximation as $n \to \infty$, and in the long run as $t \to \infty$.

7. Acknowledgements. The authors are grateful to the referees for numerous helpful and insightful comments that lead to a major improvement of the manuscript.

REFERENCES

[1] A. Brown, *Red-Blooded Risk: The Secret History of Wall Street*, Wiley, 2011.
[2] S. Browne and W. Whitt, *Portfolio choice and the Bayesian Kelly criterion*, Adv. in Appl. Probab., 28 (1996), pp. 1145–1176.

[3] P. Carr, H. Geman, D. B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business, 75 (2002), pp. 305–332.

[4] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman and Hall, 2003.

[5] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, second ed., 1971.

[6] B. V. Gnedenko and G. Fahim, *A certain transfer theorem*, Dokl. Akad. Nauk SSSR, 187 (1969), pp. 15–17.

[7] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Elsevier/Academic Press, eighth ed., 2015.

[8] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, second ed., 2003.

[9] J. L. Kelly, *A new interpretation of information rate*, Bell System Technical Journal, (1956), pp. 917–926.

[10] A. Klenke, *Probability theory*, Springer, second ed., 2014.

[11] V. Y. Korolev, L. M. Zaks, and A. I. Zeifman, *On convergence of random walks generated by compound Cox processes to Lévy processes*, Statist. Probab. Lett., 83 (2013), pp. 2432–2438.

[12] R. S. Liptser and A. N. Shiryayev, *Theory of martingales*, vol. 49 of Mathematics and its Applications (Soviet Series), Kluwer, 1989.

[13] D. Luenberger, *Investment Science*, Oxford University Press, 1998.

[14] L. C. MacLean, W. T. Ziemba, and E. O. Thorp, eds., *The Kelly Capital Growth Investment Criterion: Theory and practice*, vol. 3 of World Scientific Handbook in Financial Economics Series, World Scientific, 2011.

[15] H. M. Markowitz, *Portfolio selection: Efficient diversification of investments*, Cowles Foundation for Research in Economics at Yale University, Monograph 16, Wiley-Chapman & Hall, 1959.

[16] H. M. Markowitz, *Investment for the long run: New evidence for an old rule*, Journal of Finance, 31 (1976), pp. 1273–1286.

[17] P. E. Protter, *Stochastic integration and differential equations*, vol. 21 of Stochastic Modelling and Applied Probability, Springer, second ed., 2005.

[18] S. Rachev, ed., *Handbook of Heavy Tailed Distributions in Finance*, vol. 1 of Handbooks in Finance, Elsevier, 2003.

[19] K. Sato, *Lévy processes and infinitely divisible distributions*, vol. 68 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2013.

[20] E. O. Thorp, *A perspective on quantitative finance: models for beating the market*, in The Best of Wilmott 1: Incorporating the Quantitative Finance Review, P. Wilmott, ed., Willey, 2005, pp. 33–38.

[21] E. O. Thorp, *The Kelly criterion in blackjack, sports betting, and the stock market*, in The Kelly Capital Growth Investment Criterion: Theory and practice, L. C. MacLean, W. T. Ziemba, and E. O. Thorp, eds., vol. 3 of World Scientific Handbook in Financial Economics Series, World Scientific, 2011, pp. 789–832.

[22] E. O. Thorp, *Portfolio choice and the Kelly criterion*, in The Kelly Capital Growth Investment Criterion: Theory and practice, L. C. MacLean, W. T. Ziemba, and E. O. Thorp, eds., vol. 3 of World Scientific Handbook in Financial Economics Series, World Scientific, 2011, pp. 81–90.

[23] E. O. Thorp, *Understanding the Kelly criterion*, in The Kelly Capital Growth Investment Criterion: Theory and practice, L. C. MacLean, W. T. Ziemba, and E. O. Thorp, eds., vol. 3 of World Scientific Handbook in Financial Economics Series, World Scientific, 2011, pp. 509–524.

[24] V. M. Zolotarev, *Asymptotic expansions with probability* 1 *of sums of independent random variables*, Teor. Veroyatnost. i Primenen., 36 (1991), pp. 770–772.