

Article

Expectation and Optimal Allocations in Existential Contests of Finite, Heavy-Tail-Distributed Outcomes

Ralph Vince

Exsuperatus, Sarasota, FL 34231, USA; rmvince@exsuperatus.com

Abstract: Financial time series and other human-driven, non-natural processes are known to exhibit fat-tailed outcome distributions. That is, such processes demonstrate a greater tendency for extreme outcomes than the normal distribution or other natural distributional processes would predict. We examine the mathematical expectation, or simply “expectation”, traditionally the probability-weighted outcome, regarded since the seventeenth century as the mathematical definition of “expectation”. However, when considering the “expectation” of an individual confronted with a finite sequence of outcomes, particularly existential outcomes (e.g., a trader with a limited time to perform or lose his position in a trading operation), we find this individual “expects” the median terminal outcome over those finite trials, with the classical seventeenth-century definition being the asymptotic limit as trials increase. Since such finite-sequence “expectations” often differ in values from the classic one, so do the optimal allocations (e.g., growth-optimal). We examine these for fat-tailed distributions. The focus is on *implementation*, and the techniques described can be applied to all distributional forms. We make no assertion that the empirical data for any financial time series comports to the generalized hyperbolic distribution (GHD), which we will use as a proxy of any heavy-tailed distribution herein. Rather, we have selected the GHD to highlight the process for determining expectation and other important time-dependent metrics in existential contests, using the GHD as a generic proxy for the specific distributional form an implementor of the presented technique might ascribe to the empirical data.



Citation: Vince, R. Expectation and Optimal Allocations in Existential Contests of Finite, Heavy-Tail-Distributed Outcomes. *Mathematics* **2024**, *12*, 11. <https://doi.org/10.3390/math12010011>

Academic Editors: Maria C. Mariani, Javier Perote, Andrés Mora-Valencia and Trino-Manuel Níguez

Received: 7 November 2023

Revised: 13 December 2023

Accepted: 17 December 2023

Published: 20 December 2023

Keywords: stopping times; expectation; finite sequences; non-asymptotic; optimal allocation; Kelly criterion; stable pareto; fat tails; generalized hyperbolic distribution

MSC: 62E17; 62G32; 62L15; 62P05; 91A35; 91B28; 91B32

1. Introduction and Review of the Literature

In order to adequately apply methods of non-asymptotic expectation and allocation to fat-tailed-distributed outcomes, we must examine the history of all three of these facets. First, we will review the history of fat-tailed distributions in economic time series, leading to the proposal of the generalized hyperbolic distribution (GHD). Next, we will explore the history of mathematical expectation as well as optimal allocation methods, which we will apply to stable fat-tailed-distributed outcomes.

1.1. Fat-Tailed Distributions in Economic Time Series—History

While explicit mentions of fat-tailed distributions in the financial price context may have been limited prior to the cited references, observations and discussions regarding market fluctuation irregularities and extreme event occurrences existed in financial history.

Historical examples, such as the financial crises and market panics occurring throughout history, provide insights into understanding market dynamics and extreme event presence. Documented observations of events, such as the 17th century Tulip Mania and 18th century South Sea Bubble, offer glimpses into the irregular, non-normal behaviors exhibited in financial prices.



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

The field of speculative trading and risk management, with roots in ancient civilizations, involved assessing trading risks and managing uncertainties in commodity prices and exchange rates. While formal conceptualization of fat-tailed distributions may not have been explicitly articulated during these early periods, experiences and observations related to market irregularities and extreme price movements contributed to gradually understanding the complexities and non-standard behaviors present in financial markets and asset prices.

Exploring historical accounts of financial crises and the evolution of trading practices in ancient and medieval civilizations provides valuable insights into the early recognition of irregular market behaviors and consideration of extreme events in financial transactions and risk management.

One of the early explicit mentions of fat-tailed distributions in the financial context is found in the work of economist Alfred Marshall. His contributions to supply–demand theory and goods pricing included discussions on market price variability and economic outlier occurrences. While Marshall’s work focused on microeconomic principles, his insights into market fluctuation nature provided a preliminary understanding of heavy-tailed distribution’s potential presence in economic phenomena.

The field of actuarial science and insurance, predating modern finance theory, involved risk analysis and extreme event occurrence, laying foundations for understanding tail risk and accounting for rare but significant events in risk management.

While these earlier references did not explicitly focus on fat-tailed distributions in finance, they contributed to understanding variability and risk in economic/financial systems and provided insights leading to modern heavy-tailed distribution theories and finance applications.

Marshall’s 1890 “Principles of Economics” [1] laid foundations for modern microeconomic theory and analysis. While focused on general economic principles, it included discussions on market dynamics, supply–demand, and price fluctuation factors.

Later, in his 1919 work “Industry and Trade” [2], Marshall examined industrial organization dynamics and market functioning, providing insights into price variability and factors influencing economic fluctuations.

Marshall’s works provided insights into economic principles and market dynamics, laying groundwork for modern economic theory development. While not explicitly discussing finance-related fat-tailed distributions, they contributed to understanding market variability and economic analysis, providing context on risk and uncertainty ideas in economic systems.

Louis Bachelier’s 1900 thesis “Théorie de la spéculation” [3] delved into mathematical stock price modeling and introduced random walks, playing a significant role in efficient market hypothesis development and understanding financial price movements. While not focused on fat-tailed distributions, Bachelier’s work is considered an early reference laying groundwork for understanding financial market stochastic processes.

The work of French mathematician Paul Lévy [4,5] made significant contributions to probability theory and stochastic processes, including those related to finance. Specifically, Lévy’s research on random variable addition and general stochastic processes greatly influenced modern probability theory development and understanding stochastic dynamics across fields. Lévy’s contributions to Brownian motion theory and introduction of Lévy processes provided groundwork for understanding asset price dynamics and modeling market fluctuations. Through foundational work on topics such as Brownian motion, random processes, and stable distributions, Lévy helped establish mathematical tools and probability frameworks later applied to quantitative finance and modeling.

While explicit mentions of finance-related fat-tailed distributions may have been limited up to this time, Bachelier and Lévy’s contributions provided important insights into the probabilistic nature of finance and asset price movement governing processes.

In Feller’s work [6], we begin to find discussions on heavy-tailed distributions and their applications across fields. Feller’s work contributed to understanding heavy-tailed distribution.

Additionally, the notion of fat-tailed distributions in economics/finance was discussed by prominent statisticians and economists. Notably, Vilfredo Pareto [7] introduced the Pareto distribution in the early 20th century, closely related to the stable Pareto distribution, though it may not have been directly applied to price data then.

The concept of stable Pareto distributions in the finance context was notably introduced by mathematician Benoit Mandelbrot in his influential fractal and finance modeling work. Mandelbrot extensively discussed heavy-tailed distribution's presence and implications for finance modeling.

In particular, Mandelbrot [8] discussed stable Pareto distributions as a potential model for certain financial data. This seminal paper laid foundations for his research on the financial market's fractal nature and non-Gaussian model applications describing market dynamics.

Additionally, Mandelbrot [9] expanded on these ideas, delving into stable distribution concepts and their significance in understanding market fluctuations and risk dynamics.

Differences between the Pareto distribution, originally described by Vilfredo Pareto, and the stable Pareto distribution, referred to by Benoit Mandelbrot in 1963, lie in their specific characteristics and properties.

The Pareto distribution, introduced by Vilfredo Pareto in the early 20th century, is a power-law probability distribution describing wealth and income distribution in society. It exhibits a heavy tail, indicating a small number of instances with significantly higher values than the majority. It follows a specific mathematical form, reflecting the "Pareto principle" or the "80-20 rule", suggesting that roughly 80% of effects come from 20% of causes.

The stable Pareto distribution, referred to by Benoit Mandelbrot, represents an extension of the traditional Pareto distribution. Mandelbrot's 1963 work marks a pivotal point in applying stable distributions, including stable Pareto, to financial data and market behavior modeling. It incorporates additional parameters, allowing greater flexibility in modeling extreme events and heavy-tailed phenomena across fields such as finance and economics. The stable Pareto distribution accounts for distribution shape and scale stability across scales, making it suitable for analyzing complex systems with long-term dependencies and extreme fluctuations.

While both distributions share heavy-tailed characteristics and implications for analyzing extremes and rarities, the stable Pareto distribution, as conceptualized by Mandelbrot, provided a more nuanced and adaptable framework for modeling complex phenomena, especially in finance and other dynamic systems where traditional distributions typically fell short in capturing real-world data intricacies.

1.1.1. Subsequent Historical Milestones in Heavy-Tailed Distributions Applied to Financial Time Series

With time, many other fat-tailed modeling techniques found their way into quantitative departments. We examine several historically influential contributions within the realm of heavy-tailed distributions in finance.

Reference [10] made the area of non-Gaussian stable processes accessible to researchers, graduate students, and practitioners. This book presented the similarity between Gaussian and non-Gaussian stable multivariate distributions and introduced one-dimensional stable random variables. It also discussed the most basic sample path properties of stable processes, including sample boundedness and continuity.

The problem of parametrically specifying distributions suitable for asset-return models is reconsidered in [11]. The authors described alternative distributions, showing how they could be estimated and applied to stock-index and exchange-rate data. The book also investigates the implications for options pricing, making it a valuable resource for understanding the application of stable Paretian models in finance.

Further along the line of essential study in heavy-tailed distributions in finance, Reference [12] was a comprehensive resource that presented current research focusing on

heavy-tailed distributions with specific applications in financial time series. The book covered methodological issues, including probabilistic, statistical, and econometric modeling under non-Gaussian assumptions. It also explored the applications of stable and other non-Gaussian models in finance and risk management, making it a valuable resource for finance and economics professors, researchers, and graduate students.

Finally, Reference [13] provided an accessible introduction to stable distributions for financial modeling. This article highlighted the need for better models for financial returns, as the normal (or Gaussian) model did not capture the large fluctuations seen in real assets. The author introduced stable laws, a class of heavy-tailed probability distributions that could model large fluctuations and allow more general dependence structures, offering a more accurate model for financial returns, and even focused on pitfalls in parameter estimation of such forms.

1.1.2. The Generalized Hyperbolic Distribution (GHD)

The generalized hyperbolic distribution (GHD) was introduced in [14], examining aeolian processes.

The GHD constitutes a flexible continuous probability law defined as the normal variance-mean mixture distribution, where the mixing distribution assumes the generalized inverse Gaussian form. Its probability density function can be expressed in terms of modified Bessel functions of the second kind, commonly denoted as BesselK functions in the literature. The specific functional form of the GHD density involves these BesselK functions along with model parameters governing aspects such as tail behavior, skewness, and kurtosis. While not possessing a simple closed-form analytical formula, the density function can be reliably evaluated through direct numerical calculation of the BesselK components for given parameter values. The mathematical tractability afforded by the ability to compute the GHD density and distribution function underpins its widespread use in applications spanning economics, finance, and the natural sciences.

Salient features of the GHD include closure under affine transformations and infinite divisibility. The latter property follows from its constructability as a normal variance-mean mixture based on the generalized inverse Gaussian law. As elucidated in [15], infinite divisibility bears important connections to Lévy processes, whereby the distribution of a Lévy process at any temporal point manifests infinite divisibility. While many canonical infinitely divisible families (e.g., Poisson and Brownian motion) exhibit convolution-closure, whereby the process distribution remains within the same family at all times, Podgórski et al. showed the GHD does not universally possess this convolution-closure attribute.

Owing to its considerable generality, the GHD represents the overarching class for various pivotal distributions, including the Student's t , Laplace, hyperbolic, normal-inverse Gaussian, and variance-gamma. Its semi-heavy tail properties, unlike the lighter tails of the normal distribution, enable modeling of far-field behaviors. Consequently, the GHD finds ubiquitous applications in economics, particularly in financial market modeling and risk management contexts, where its tail properties suit the modeling of asset returns and risk metrics, and by 1995, the GHD appeared in financial markets application in [16]. They applied the hyperbolic subclass of the GHD to fit German financial data.

This work was later extended by Prause in 1999 [17], who applied GHDs to model financial data on German stocks and American indices. Since then, the GHD has been widely used in finance and risk management to model a wide range of financial and economic data, including stock prices, exchange rates, and commodity prices. See [18–25] for applications of the GHD to economic and share price data.

The GHD has been shown to provide a more realistic description of asset returns than the other classical distributional models and has been used to estimate the risk of financial assets and to construct efficient portfolios in energy and stock markets.

The generalized hyperbolic distribution has emerged as an indispensable modeling tool in modern econometrics and mathematical finance due to its advantageous mathemati-

cal properties and empirical performance. Salient features underscoring its suitability for analyzing economic and asset price data include:

- Substantial flexibility afforded by multiple shape parameters, enabling the GHD to accurately fit myriad empirically observed, non-normal behaviors in financial and economic data sets. Both leptokurtic and platykurtic distributions can be readily captured.
- Mathematical tractability, with the probability density function, cumulative distribution function, and characteristic function expressible in closed analytical form. This facilitates rigorous mathematical analysis and inference.
- Theoretical connections to fundamental economic concepts, such as utility maximization. Various special cases also share close relationships with other pivotal distributions, such as the variance-gamma distribution.
- Empirical studies across disparate samples and time horizons consistently demonstrate a superior goodness-of-fit compared to normal and stable models when applied to asset returns, market indices, and other economic variables.
- Ability to more accurately model tail risks and extreme events compared to normal models, enabling robust quantification of value-at-risk, expected shortfall, and other vital risk metrics.
- Despite the lack of a simple analytical formula, the distribution function can be reliably evaluated through straightforward numerical methods, aiding practical implementation.

In summary, the mathematical tractability, theoretical foundations, empirical performance, and remarkable flexibility of the GHD render it effective for modeling the non-normal behaviors ubiquitous in economic and financial data sets.

Although we will be using the symmetric GHD for ease-of-use in demonstrating the techniques we will illuminate in the discussion herein (the techniques being applicable to any distributional form), and despite the GDH possessing numerous favorable characteristics, its flexibility is limited in comparison to other extensions of the normal distribution that provide improved adaptability.

1.1.3. Extensions of the Normal Distribution for Greater Flexibility (GHD)

The skew two-piece, skew-normal distribution is a generalization of the normal distribution with an additional shape parameter. This distribution can model both left-skewed (negatively skewed) and right-skewed (positively skewed) data, depending on the value of the skewness parameter [26–32]. It becomes a normal distribution when the skewness is equal to one.

This is particularly useful in economic time series, where data often exhibit asymmetry and non-normality [33]. The skew two-piece, skew-normal distribution becomes a normal distribution when the skewness is equal to one, making it a more general model that can capture a wider range of data patterns [34]. Additionally, using a distribution that directly handles skewness can perform better than data transformation, which may have limitations, such as reduced information, difficulty in interpreting transformed variables, and no guarantee of joint normality [33].

The skew two-piece, skew-normal distribution and the skew normal distribution are not the same, but they are related. Both distributions are generalizations of the normal distribution that can accommodate skewed data, but they do so in different ways.

The skew normal distribution is a type of distribution that introduces skewness into a normal distribution. It is characterized by an asymmetry in the probability distribution of a real-valued random variable about its mean [35]. The skewness value can be positive, zero, negative, or undefined. For a unimodal distribution, negative skew commonly indicates that the tail is on the left side of the distribution, and positive skew indicates that the tail is on the right.

On the other hand, the skew two-piece, skew-normal distribution, also known as the Fernández-Steel Skew Normal distribution, is a more general model that introduces an additional shape parameter to the normal distribution. This distribution can model both left-skewed (negatively skewed) and right-skewed (positively skewed) data, depending on

the value of the skewness parameter. It becomes a normal distribution when the skewness is equal to one. This distribution is particularly useful for modeling data that are concentrated in two directions in roughly equal proportions [34].

In summary, while both distributions can model skewed data, the skew two-piece, skew-normal distribution provides a more flexible model that can capture a wider range of data patterns.

Overall, the skew two-piece, skew-normal distribution provides a more flexible and accurate model for capturing the specific characteristics of economic time series data [36].

We now turn our attention to the odd log-logistic-type distributions.

The odd log-logistic normal distribution, the odd log-logistic geometric normal distribution, and the odd log-logistic skew normal distribution are all different, albeit related probability distributions.

The odd log-logistic normal distribution is a three-parameter model that is symmetric, platykurtic, leptokurtic, and may be unimodal or bimodal. It is a new distribution that has been studied for its various structural properties, including explicit expressions for the ordinary and incomplete moments, generating function, and mean deviations. It has been shown to provide more realistic fits than other special regression models in certain cases [37].

The odd log-logistic geometric normal distribution is a new regression model based on the odd log-logistic geometric normal distribution for modeling asymmetric or bimodal data with support in \mathbb{R} . This distribution is designed to model data that are not symmetric and may exhibit bimodal characteristics [38].

The odd log-logistic skew normal distribution, on the other hand, is a different distribution that combines the odd log-logistic and skew normal distributions. It is designed to model data that are not symmetric and may exhibit skewness characteristics. This distribution is particularly useful for modeling data that are concentrated in two directions in roughly equal proportions [37].

In summary, the odd log-logistic normal distribution, the odd log-logistic geometric normal distribution, and the odd log-logistic skew normal distribution are all different distributions that have been developed to model specific characteristics of data, including symmetry, kurtosis, bimodality, and skewness. Each distribution is designed for different purposes and may be more or less suitable depending on the specific characteristics of the data being modeled.

1.2. Mathematical Expectation—History

The genesis of mathematical expectation lies in the mid-17th century study of the “problem of points”, which concerns the equitable division of stakes between players forced to conclude a game prematurely. This longstanding conundrum was posed to eminent French mathematician Blaise Pascal in 1654 by compatriot Chevalier de Méré, an amateur mathematician. Galvanized by the challenge, Pascal commenced correspondence with eminent jurist Pierre de Fermat, and the two deduced identical solutions rooted in the fundamental tenet that a future gain’s value must be proportional to its probability.

Pascal’s 1654 work [39], compiled amidst his collaboration with Fermat, explores sundry mathematical concepts, including foundational probability theory and expectation calculation for games of chance. Through rigorous analysis of gambling scenarios, Pascal delved into mathematical quantification of expected values, establishing core principles that seeded probability theory’s growth and diverse applications. The “Pascal–Fermat correspondence” offers insights into the germination of mathematical expectation and its integral role in comprehending the outcomes of stochastic phenomena.

Contemporaneously in the mid-17th century, the Dutch mathematician, astronomer, and physicist Christiaan Huygens significantly advanced the field of probability and study of mathematical expectation. Building upon the contributions of Pascal and Fermat, Huygens’ 1657 work [40] incisively examined numerous facets of probability theory, with particular emphasis on games of chance and the calculation of fair expectations in contexts

of gambling. He conceived an innovative methodology for determining expected value, underscoring the importance of elucidating average outcomes and long-term behaviors of random processes.

Huygens' mathematical expectation framework facilitated subsequent scholars in expanding on his ideas to construct a comprehensive system for disentangling stochastic phenomena, quantifying uncertainty, and illuminating the nature of randomness across diverse scientific realms.

Essentially, the calculation for mathematical expectation, or "expectation", remained unchanged for centuries, until the work in [41]. It was Vince who decided to look at a different means of calculating it to reflect an individual or group of individuals who are in an "existential" contest of finite length. Consider a young trader placed on an institution's trading desk and given a relatively short time span to "prove himself" or be let go, or the resources used by a sports team in an "elimination-style" playoff contest, or such resources of a nation state involved in an existential war.

Such endeavors require a redefinition of the "expectation"—what the party "expects" to happen—and can readily be defined as that outcome, where half the outcomes are better or the same and half are worse or the same at the end of the specified period or number of trials.

Thus, the answer to this definition of "expectation" is no longer the classic, centuries-old definition of the probability weighted mean outcome, but instead is the mean sorted cumulative outcome over the specified number of trials.

For example, consider the prospect of a game where one might win 1 unit with a probability of 0.9, and lose -10 units with a probability of 0.1. The classical expectation is to lose -0.1 per trial.

Contrast this to the existential contest, which is only 1 trial long. Here, what one "expects" to happen is to win 1 unit. In fact, it can be shown that if one were to play and quit this game after 6 trials, one would "expect" to make 6 units (that is, half the outcomes of data distributed as such would show better or the same results, and half would show worse or the same results). After 7 trials, with these given parameters, however, the situation turns against the bettor.

Importantly, as demonstrated in [41], this non-asymptotic expectation converges to the asymptotic, centuries-old "classic" expectation in the limit as the number of trials approaches infinity. Thus, the "classic" expectation can be viewed as the asymptotic manifestation of Vince's more general non-asymptotic expression.

This redefined "expectation" then links to optimal resource allocations in existential contests. The non-asymptotic expectation provides a modern conceptualization of "expectation" tailored to modeling human behavior in high-stakes, finite-time scenarios and, in the limit as the number of trials becomes ever-large, converges on the classical definition of mathematical expectation.

1.3. Optimal Allocations—History

The concept of geometric mean maximization originates with Daniel Bernoulli, who made the first known reference to it in 1738. Prior to that time, there is no known mention in any language of even generalized optimal reinvestment strategies by merchants, traders, or in any of the developing parts of the earth. Evidently, no one formally codified the concept. If anyone contemplated it, they did not record it.

Bernoulli's 1738 paper [42] was originally published in Latin. A German translation appeared in 1896, and it was referenced in John Maynard Keynes' work [43]. In 1936, John Burr Williams' paper [44], pertaining to trading in cotton, posited that one should bet on a representative price. If profits and losses are reinvested, the method of calculating this price is to select the geometric mean of all possible prices.

Bernoulli's 1738 paper was finally translated into English in *Econometrica* in 1954. When game theory emerged in the 1950s, concepts were being widely examined by numerous economists, mathematicians, and academics. Against this fertile backdrop, John L. Kelly

Jr. [45] demonstrated that to achieve maximum wealth, a gambler should maximize the expected value of the logarithm of his capital. This is optimal because the logarithm is additive in repeated bets and satisfies the law of large numbers. In his paper, Kelly showed how Claude Shannon's information theory [46] could determine growth-optimal bet sizing for an informed gambler.

Maximizing the expected end wealth is known as the Kelly criterion. Whether Kelly knew it or not, the cognates to his paper are from Daniel Bernoulli. In all fairness, Bernoulli was likely not the originator either. Kelly's paper presented this as a solution to a technological problem absent in Bernoulli's day.

Kelly's paper makes no reference to applying the methods to financial markets. The gambling community embraced the concepts but applying them to various capital market applications necessitated formulaic alterations, specifically scaling for the absolute value of the worst-case potential outcome, which become particularly important considering fat-tailed probabilities. In fairness, neither Kelly nor Shannon were professional market traders, and the work presented did not claim applicability to finance. However, in actual practice, scaling to worst-case outcomes is anything but trivial. It is this author's contention that Kelly and Shannon (who signed off on it [28]) would have caught this oversight (they would not have intentionally opted to make the application scope of this work knowingly less general) had the absolute value of worst-case not simply been unity, thus cloaking it. The necessary scaling was provided in [47].

In subsequent decades after Reference [45], many papers by numerous researchers expanded on geometric growth optimization strategies in capital market contexts, notably Bellman and Kalaba [48], Breiman [49], Latane [50,51], Tuttle [51], Thorp [52,53], and others. Thorp, a colleague of Claude Shannon, developed a winning strategy for Blackjack using the Kelly criterion and presented closed-form formulas to determine the Kelly fraction [52].

The idea of geometric mean maximization was also well-critiqued. Samuelson [54,55], Goldman [56], Merton [57], and others argued against universally accepting it as the investor criterion. Samuelson [54] highlighted that the Kelly fraction is only optimal asymptotically, as the number of trials approaches infinity, not for finite trials, whereas, in fact, it would always be sub-optimal.

The formulation that yields the growth-optimal allocation to each component in a portfolio of multiple components for a finite number of trials is provided in [41]. This approach incorporates the net present value of amounts wagered and any cash flows over the finite timespan, acknowledging that real-world outcomes typically do not manifest instantly.

Growth-optimality, however, is not always the desired criterion. The formulaic framework for determining it can be used to discern other "optimal" allocations. In Vince and Zhu's work [58], we find two fractional allocations less than the growth-optimal allocation for maximizing the various catalogued return-to-risk ratios. This was further expanded upon by de Prado, Vince, and Zhu [59].

Finally, Reference [60] provides the case of using the formula for geometric growth maximization as a "devil's advocate" in contexts where the outcome from one period is a function of available resources from previous periods, and one wishes to diminish geometric growth. Examples include certain biological/epidemiological applications and "runaway" cost functions, such as national debts.

2. Characteristics of the Generalized Hyperbolic Distribution (GHD)—Distributional Form and Corresponding Parameters

As discussed, although the two-piece, skew-normal distributions and the odd log-logistic-type distributions offer greater flexibility than the GHD and, therefore, are generally considered to better represent financial time series data, we will be illustrating our methods on the symmetrical GHD for simplicity.

These methods can be applied to any distributional form. The focus herein is not on the specific distributional form employed to model financial time series, but on the techniques involved in discerning a sample set of probable outcomes based on the particular

distribution of an existential contest of finite length, from which numerous time-dependent metrics can be determined.

In addition to having the characteristics of being closed under affine transformations and possessing an infinitely divisible property, which enables connections to Lévy processes, the GHD provides some practical implantation advantages to the study of economic data series. Among these, the GHD possesses:

- A more general distribution class that includes Student’s t, Laplace, hyperbolic, normal-inverse Gaussian, and variance-gamma as special cases. Its mathematical tractability provides the ability to derive other distribution properties.
- Semi-heavy tail behavior, which allows modeling data with extreme events and fat-tailed probabilities.
- A density function, involving modified Bessel functions of the second kind (BesselK functions). Thus, although there is a lack of a closed-form density function, it can still be numerically evaluated in a straightforward manner.

We examine the derivation of the probability density function (PDF) of the symmetrical GHD, and then derive the first integral of the PDF with respect to the random variable to obtain the cumulative density function (CDF) of the symmetrical GHD.

A characteristic function is a mathematical tool that provides a way to completely describe the probability distribution of a random variable. It is defined as the Fourier transform of the PDF of the random variable and is particularly useful for dealing with linear functions of independent random variables. We, therefore, derive the PDF of the symmetrical GHD:

1. Starting with the characteristic function of the symmetrical GHD, as derived from [61] as:

$$\varphi(t) = \exp \left\{ i\mu t + \frac{\alpha}{\delta} \left[\left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 - 2i\beta t - t^2} \right) - \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \right) \right] \right\} \tag{1}$$

where:

α = concentration

δ = scale

μ = location

λ = shape

γ = skewness

$\beta = \sqrt{\gamma^2 - \delta^2} / \delta$

2. Using the inversion formula, the PDF of the GHD can be obtained from the characteristic function:

n.b. In the characteristic function provided in the paper by Klebanov and Rachev (1996) [15], t is a variable used to represent the argument of the characteristic function. It is not the same as x , which is a variable used to represent the value of the random variable being modeled by the GHD.

The characteristic function of a distribution is a complex-valued function that is defined for all values of its argument, t . It is the Fourier transform of the probability density function (PDF) of the distribution, and it provides a way to calculate the moments of the distribution and other properties, such as its mean, variance, and skewness.

In the case of the GHD, the characteristic function is given by (1).

The argument t of the characteristic function is a complex number that can take on any value. It is not the same as the value of the random variable being modeled by the GHD, which is represented by the variable x . The relationship between the characteristic function and the PDF of the GHD is given by the Fourier inversion formula:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

where $f(x)$ is the PDF of the GHD and x is the value of the random variable being modeled.

3. Substituting the expression for the characteristic function into the Fourier inversion formula:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \exp \left\{ i\mu t + \frac{\alpha}{\delta} \left[\left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 - 2i\beta t - t^2} \right) - \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \right) \right] \right\} dt$$

4. Simplifying the expression inside the exponential:

$$i\mu t + \frac{\alpha}{\delta} \left[\left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 - 2i\beta t - t^2} \right) - \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \right) \right]$$

$$= i\mu t + \frac{\alpha}{\delta} \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \right) \left[K_\lambda \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right) - K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 - 2i\beta t - t^2} \right) \right]$$

5. Using the identity: $K_\lambda(z) = \frac{1}{2} \left(\frac{z}{2} \right)^\lambda \int_0^\infty e^{-z \cosh(u)} \cos h(\lambda u) du$, to simplify the expression:

$$f(x) = \frac{\alpha}{2\pi\delta} \left(\frac{\gamma}{\delta} \right)^\lambda \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^{\lambda-1} \int_0^\infty e^{-\alpha \sqrt{\delta^2 + (x - \mu)^2} \cosh(u)} \cos h(\lambda u) du$$

6. Substituting: $v = \alpha \sqrt{\delta^2 + (x - \mu)^2} \sinh(u)$, to simplify the integral:

$$f(x) = \frac{\alpha}{\pi\delta} \left(\frac{\alpha}{2} \right)^{\lambda/2} \frac{1}{\sqrt{\delta^2 + (x - \mu)^2}} K_{\lambda-1/2} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right)$$

7. Finally, setting the skewness parameter γ to zero to obtain the simplified PDF of the symmetric GHD:

$$f(x) = \frac{\alpha}{\delta} \left(\delta^2 + (x - \mu)^2 \right)^{-\lambda/2} K_\lambda \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right)$$

The derivation of the CDF of the symmetrical GHD is as follows:

1. Integrating the PDF from negative infinity to x to obtain the CDF:

$$F(x) = \int_{-\infty}^x f(t) dt$$

2. Substituting the PDF into the integral and simplifying:

$$F(x) = \int_{-\infty}^x \frac{\alpha}{\delta} \left(\delta^2 + (t - \mu)^2 \right)^{-\lambda/2} K_\lambda \left(\alpha \sqrt{\delta^2 + (t - \mu)^2} \right) dt$$

3. Substituting $u = \sqrt{\delta^2 + (t - \mu)^2}$ and $du = \frac{t - \mu}{\sqrt{\delta^2 + (t - \mu)^2}} dt$ to obtain:

$$F(x) = \int_{-\infty}^x \frac{\alpha}{\delta} u^{-\lambda} K_\lambda(\alpha, u) \frac{du}{t - \mu}$$

4. With the identity: $K_\lambda(x) = \frac{1}{2} \left(\frac{x}{2} \right)^\lambda \int_0^\infty e^{-\frac{x}{2} \left(t + \frac{1}{t} \right)} t^{\lambda-1} dt$, to obtain:

$$F(x) = \frac{\alpha}{2\delta} \left(\frac{\alpha}{2} \right)^\lambda \int_{-\infty}^x \int_0^\infty u^{-\lambda} e^{-\frac{\alpha}{2} (u^2 + v^2 - 2uv \cos(\theta))} v^{\lambda-1} dv d\theta \frac{du}{t - \mu}$$

5. Substituting $v = u \sin(\theta)$ and $dv = u \cos(\theta) d\theta$ to obtain:

$$F(x) = \frac{\alpha}{2\delta} \left(\frac{\alpha}{2} \right)^\lambda \int_{-\infty}^x \int_0^{\pi/2} e^{-\frac{\alpha}{2} u^2 (1 + \sin^2(\theta) - 2 \sin(\theta) \cos(\theta))} u^{2\lambda-2} \sin^{2\lambda-2}(\theta) d\theta du \frac{1}{t - \mu}$$

6. Finally, simplifying the expression by noting that $\sin^2(\theta) - 2\sin(\theta)\cos(\theta) + 1 = (\sin(\theta) - \cos(\theta))^2$ and taking the limit as $x \rightarrow \infty$ to obtain:

$$\text{CDF}(x) = F(x) = \frac{1}{2} \left[1 + \text{sgn}(x - \mu) \left\{ 1 - \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^\alpha \right\}^{\frac{\lambda-1}{\alpha}} \right]$$

where:

α = concentration

δ = scale

μ = location

λ = shape

$\text{sgn}(x - \mu)$ = sign function, the sign of $(x - \mu)$, or simply $\text{sgn}(x - \mu) = \frac{x - \mu}{|x - \mu|}$.

Thus far, we have discussed the symmetrical GHD, where we are not accountable for skew (γ).

Implementing skew can become an unwieldy process in the GHD. Our focus is on implementation in this text, and the reader is reminded that although we present our work herein on the symmetrical GHD, it is applicable to all distributional forms, with the symmetrical GHD chosen so as to be current and relatively clear in terms of implementation of the techniques. It is the techniques that are important here—not so much the distributional form selected to demonstrate them. As such, the balance of the text will be using the symmetrical form of the GHD.

Parameter Description

The symmetric generalized hyperbolic distribution (GHD) is parameterized by four variables governing salient distributional properties:

1. The concentration parameter (α) modulates tail weight, with larger values engendering heavier tails and hence elevating the probability mass attributable to extreme deviations. α typically assumes values on the order of 0.1 to 10 for modeling economic phenomena.
2. The scale parameter (δ) controls the spread about the central tendency, with larger values yielding expanded variance and wider dispersion. Applied economic analysis often utilizes δ ranging from 0.1 to 10.
3. The location parameter (μ) shifts the distribution along the abscissa, with positive values effecting rightward translation. In economic contexts, μ commonly falls between -10 and 10 .
4. The shape parameter (λ) influences peakedness and flatness, with larger values precipitating more acute modes and sharper central tendencies. λ on the order of 0.1 to 10 frequently appears in economic applications.

While these indicated ranges provide first approximations, the precise parametric tuning warrants meticulous examination of the data properties and phenomenon under investigation. Moreover, complex interdependencies between the four variables potentially exist, necessitating holistic assessment when calibrating the symmetric GHD for applied modeling tasks.

3. Parameter Estimation of the Symmetrical GHD

The task is now specified as follows. We have a set of empirical observations that represent the returns or percentage changes in an economic time series from one period to the next. This can be stock prices, economic data, the output of a trading technique (These are often expressed as $d_0/d_1 - 1$, where d_0 represents the most recent data and d_1 the data from the previous period. In the case of outputs of a trading approach, the divisor is problematic—the results must be expressed in log terms, in terms of a percentage change. Such results must be expressed, therefore, in terms of an amount put at risk to assume such results, which are generalized to the largest potential loss. The notion of “potential loss” is

another Pandora's box worthy of a separate discussion), or any other time series data that comport to an assumed probability density function whose parameters we wish to fit to this sample data.

In the context of finding the concentration, scale, location, and shape parameters that best fit the PDF(x) or the symmetrical GHD, we can use maximum likelihood estimation (MLE) to estimate these parameters.

Maximum likelihood estimation (MLE) stands as a widely used statistical technique for estimating probability distribution parameters based on observed data. This method involves determining parameter values that optimize the likelihood of observing the data within a given statistical model assumption. MLE can be viewed as a special instance of maximum a posteriori estimation (MAP) under the assumption of a uniform parameter prior distribution. Alternatively, it can be considered a variant of MAP that disregards the prior, resulting in an unregularized approach. The MLE estimator selects the parameter value that maximizes the probability (or probability density, in continuous cases) of observing the given data. When the parameter comprises multiple components, individual maximum likelihood estimators are defined for each component within the complete parameter's MLE. MLE is recognized as a consistent estimator, meaning that under certain conditions and as the sample size increases, the estimates derived from MLE tend toward the true parameter values. Put differently, with larger sample sizes, the likelihood of obtaining accurate parameter estimates increases.

Mathematically, we can express MLE as follows:

1. Beginning with the likelihood function for the symmetrical GHD, which is the product of the PDF evaluated at each data point:

$$L(\theta|x) = \prod_{i=1}^n f(x_i|\theta)$$

2. Taking the natural logarithm of the likelihood function to obtain the log-likelihood function:

$$\ln L(\theta|x) = \sum_{i=1}^n \ln f(x_i|\theta)$$

3. Taking the partial derivative of the log-likelihood function with respect to each parameter and setting the result equal to zero:

$$\frac{\partial \ln L(\theta|x)}{\partial \theta} = 0$$

4. Solving the resulting system of equations to obtain the maximum likelihood estimates of the parameters.

5. Checking the second-order conditions to ensure that the estimates are indeed maximum likelihood estimates.

6. Using the estimated parameters to obtain the best-fit symmetrical GHD.

In step 1, θ represents the vector of parameters to be estimated, which includes the concentration, scale, location, and shape parameters. In step 2, we take the natural logarithm of the likelihood function to simplify the calculations and avoid numerical issues. In step 3, we take the partial derivative of the log-likelihood function with respect to each parameter and set the result equal to zero to find the values of the parameters that maximize the likelihood function. In step 4, we solve the resulting system of equations to obtain the maximum likelihood estimates of the parameters. In step 5, we check the second-order conditions to ensure that the estimates are indeed maximum likelihood estimates. Finally, in step 6, we use the estimated parameters to obtain the best-fit symmetrical GHD.

We shall now apply the discussed method of fitting the symmetrical GHD parameters to a real-world, financial time series. We implemented the parameter fitting techniques to the determination of the values for the real-world data set of S&P500 Index percent daily price changes for the first six months of 2023.

By way of this example, we obtained a time-ordered vector of daily percentage changes in the S&P500 for the first six months of 2023. We proffer these results as follows in Table 1, which we will use to fit our symmetrical GHD parameters to.

Table 1. Daily percentage changes in the S&P500 for the first six months of 2023.

Date	S&P500	% Change	Date	S&P500	% Change
20221230	3839.5		20230403	4124.51	0.003699
20230103	3824.14	−0.004	20230404	4100.6	−0.0058
20230104	3852.97	0.007539	20230405	4090.38	−0.00249
20230105	3808.1	−0.01165	20230406	4105.02	0.003579
20230106	3895.08	0.022841	20230410	4109.11	0.000996
20230109	3892.09	−0.00077	20230411	4108.94	-4.1×10^{-5}
20230110	3919.25	0.006978	20230412	4091.95	−0.00413
20230111	3969.61	0.012849	20230413	4146.22	0.013263
20230112	3983.17	0.003416	20230414	4137.64	−0.00207
20230113	3999.09	0.003997	20230417	4151.32	0.003306
20230117	3990.97	−0.00203	20230418	4154.87	0.000855
20230118	3928.86	−0.01556	20230419	4154.52	-8.4×10^{-5}
20230119	3898.85	−0.00764	20230420	4129.79	−0.00595
20230120	3972.61	0.018918	20230421	4133.52	0.000903
20230123	4019.81	0.011881	20230424	4137.04	0.000852
20230124	4016.95	−0.00071	20230425	4071.63	−0.01581
20230125	4016.22	−0.00018	20230426	4055.99	−0.00384
20230126	4060.43	0.011008	20230427	4135.35	0.019566
20230127	4070.56	0.002495	20230428	4169.48	0.008253
20230130	4017.77	−0.01297	20230501	4167.87	−0.00039
20230131	4076.6	0.014642	20230502	4119.58	−0.01159
20230201	4119.21	0.010452	20230503	4090.75	−0.007
20230202	4179.76	0.014699	20230504	4061.22	−0.00722
20230203	4136.48	−0.01035	20230505	4136.25	0.018475
20230206	4111.08	−0.00614	20230508	4138.12	0.000452
20230207	4164	0.012873	20230509	4119.17	−0.00458
20230208	4117.86	−0.01108	20230510	4137.64	0.004484
20230209	4081.5	−0.00883	20230511	4130.62	−0.0017
20230210	4090.46	0.002195	20230512	4124.08	−0.00158
20230213	4137.29	0.011449	20230515	4136.28	0.002958
20230214	4136.13	−0.00028	20230516	4109.9	−0.00638
20230215	4147.6	0.002773	20230517	4158.77	0.011891
20230216	4090.41	−0.01379	20230518	4198.05	0.009445
20230217	4079.09	−0.00277	20230519	4191.98	−0.00145
20230221	3997.34	−0.02004	20230522	4192.63	0.000155
20230222	3991.05	−0.00157	20230523	4145.58	−0.01122

Table 1. *Cont.*

Date	S&P500	% Change	Date	S&P500	% Change
20230223	4012.32	0.005329	20230524	4115.24	−0.00732
20230224	3970.04	−0.01054	20230525	4151.28	0.008758
20230227	3982.24	0.003073	20230526	4205.45	0.013049
20230228	3970.15	−0.00304	20230530	4205.52	1.66×10^{-5}
20230301	3951.39	−0.00473	20230531	4179.83	−0.00611
20230302	3981.35	0.007582	20230601	4221.02	0.009854
20230303	4045.64	0.016148	20230602	4282.37	0.014534
20230306	4048.42	0.000687	20230605	4273.79	−0.002
20230307	3986.37	−0.01533	20230606	4283.85	0.002354
20230308	3992.01	0.001415	20230607	4267.52	−0.00381
20230309	3918.32	−0.01846	20230608	4293.93	0.006189
20230310	3861.59	−0.01448	20230609	4298.86	0.001148
20230313	3855.76	−0.00151	20230612	4338.93	0.009321
20230314	3919.29	0.016477	20230613	4369.01	0.006932
20230315	3891.93	−0.00698	20230614	4372.59	0.000819
20230316	3960.28	0.017562	20230615	4425.84	0.012178
20230317	3916.64	−0.01102	20230616	4409.59	−0.00367
20230320	3951.57	0.008918	20230620	4388.71	−0.00474
20230321	4002.87	0.012982	20230621	4365.69	−0.00525
20230322	3936.97	−0.01646	20230622	4381.89	0.003711
20230323	3948.72	0.002985	20230623	4348.33	−0.00766
20230324	3970.99	0.00564	20230626	4328.82	−0.00449
20230327	3977.53	0.001647	20230627	4378.41	0.011456
20230328	3971.27	−0.00157	20230628	4376.86	−0.00035
20230329	4027.81	0.014237	20230629	4396.44	0.004474
20230330	4050.83	0.005715	20230630	4450.38	0.012269
20230331	4109.31	0.014437			

We performed the MLE best-fit analysis on our returns sample data. One of the benefits of the GHD, as mentioned earlier, is the flexibility afforded by multiple shape parameters. We do not necessarily have to fit per MLE or any other “best-fit” algorithm—we could fit by eye if we so desired, the GHD is that extensible. In our immediate case, we decided to fit but set the location parameter equal to the mean of our sample returns (0.0029048), and fit the other three parameters, arriving at a best-fit set of parameters in Table 2.

Table 2. Best-fit parameter set to Table 1, only optimizing three parameters to the first six months in 2023 of daily S&P500 closing price changes. Location was set to the mean (and median, being symmetrical) of the sample data of 0.0029048 and the scale was kept at the sample standard deviation of 0.0105381.

	Name	Values
α	Concentration	2.1
δ	Scale	0.01053813
μ	Location	0.0029048
λ	Shape	1.64940218

This resulted in the following PDF and CDF, in Figure 1a,b, respectively, of our sample data set.

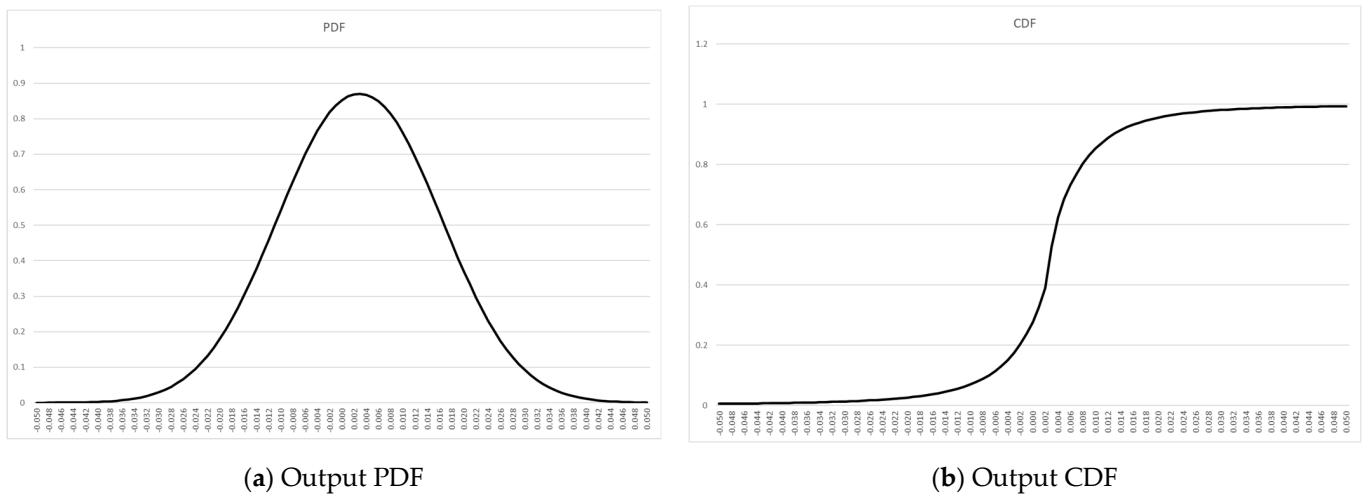


Figure 1. PDF and CDF to fitted sample returns data from the first half of 2023 for the S&P500.

For the shape of the PDF in the first 6 months of the 2023 daily price changes in the S&P500 index, the distribution is symmetrical solely because we used the symmetrical incarnation of the GHD. What has been fit here is the scale parameter and the shape parameter, reflected in Figure 1 for both the PDF and CDF of our sample data set. In working with the GHD, it is often necessary to use the flexibility of the parameters, in addition to a best-fit algorithm such as MLE, to obtain a satisfactory fit, not only to the sample data, but also to what the implementor may be anticipating in the future. Best-fit routines such as MLE are, in actual practice, a starting point.

4. Discussion

This far, we have laid the structural foundation to how we will use a given probability distribution and its parameters, which are fitted to a sample set of data. The presented technique allows any distributional shape to be applied. We chose the symmetrical GHD for its ease-of-use. In actual practice, it will be necessary to account for skewness in the distribution of the sampled data (this has been intentionally avoided in this paper to maintain simplicity).

In actual implementation, we would use the full, non-symmetrical GHD. Deriving the PDF from the characteristic function of the GHD:

Starting with the characteristic function of the GHD:

$$\varphi(t) = \exp \left\{ i\mu t + \frac{\alpha}{\delta} \left[\left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 - 2i\beta t - t^2} \right) - \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \right) \right] \right\}$$

where μ is the location parameter, α is the concentration parameter, δ is the scale parameter, λ is the shape parameter, γ is the skewness parameter, and $\beta = \sqrt{\gamma^2 - \delta^2} / \delta$.

Using the inversion formula to obtain the PDF of the GHD:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

Substituting the characteristic function into the inversion formula and simplifying:

$$f(x) = \frac{\alpha}{\pi\delta} \int_0^{\infty} \left(\frac{\gamma}{\delta} \right)^\lambda K_\lambda \left(\alpha \sqrt{\gamma^2 - \delta^2} \sqrt{1 + t^2} \right) \cos(\mu t + \beta\delta t^2 - xt) dt$$

1. Changing the variable $u = \alpha\sqrt{\gamma^2 - \delta^2}t$:

$$f(x) = \frac{\alpha}{\pi\delta} \left(\frac{\gamma}{\delta}\right)^\lambda \int_0^\infty K_\lambda \left(u \sqrt{1 + \frac{\delta^2}{\alpha^2(\gamma^2 - \delta^2)}(x - \mu)^2} \right) \cos \left(\frac{\beta\delta}{\alpha^2(\gamma^2 - \delta^2)}(x - \mu)u^2 - ux \right) du$$

2. Using the identity: $K_\lambda(z) = \frac{1}{2} \left(\frac{z}{2}\right)^\lambda \int_0^\infty e^{-z\cosh(u)} \cos h(\lambda u) du$, to simplify the integral:

$$PDF(x) = f(x) = \frac{\alpha}{\pi\delta} \left(\frac{\gamma}{\delta}\right)^\lambda \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^{\lambda-1} K_{\lambda-1} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right)$$

This is the full PDF with the skewness parameter (γ) for the GHD.

The integral required to compute the CDF(x) of the GHD with skewness (γ) does not have a closed-form solution, so numerical methods must be used to compute the CDF(x). However, there are software packages, such as R 4.2.0, MATLAB R2021a 9.10, or SAS 9.4M8, that have built-in functions for computing the CDF(x) of the GHD. Once the CDF(x) is computed, the inverse CDF(x) can be obtained by solving the equation: $F(x) = p$ for x , where p is a probability value between 0 and 1. This equation can be solved numerically using methods such as bisection, Newton–Raphson, or the secant method. Alternatively, software packages, such as R, MATLAB, or SAS, have built-in functions for computing the inverse CDF of the GHD.

Our focus is on finite sequences of outcomes from a stochastic process. Although in our discussion, we are making the assumption that these outcomes are independent and identically distributed (IID), we assume this for simplicity of argument only; absent this IID assumption, the probabilities need only be amended at each subsequent trial in the finite, N-length sequence of outcomes based on previous outcomes. We refer to these finite sequences as being “existential contests”, realizing they exist in the milieu of life itself, as part of a longer sequence of a “lifetime-long” sequence of stochastic outcomes, or in the milieu of many players confronted with a finite sequence of stochastic outcomes.

The gambler wandering through the casino does not see this sequence stop and start as it continues when he continues gambling.

These are “existential contests” because we are walling-off these principles of ultimate ergodicity. We are examining what the single, individual (in the context of one man, or one team, one nation state—all one “individual”) player whose existence hangs on the outcome of the N-length, finite sequence of outcomes—the milieu it resides in be damned. Examples of existential contests might include a young trader placed on an institution’s trading desk and given a relatively short time span to “prove himself” or be let go, or the resources used by a sports team in an “elimination-style” playoff contest, or such resources of a nation state involved in an existential war. Even our gambler wandering through a Las Vegas casino, placing the minimum sized bets, who, suddenly, decides to wager his life savings on numbers 1–24 at the roulette wheel before heading for the airport, is now confronted with an N = 1 existential contest.

The notion of the existential contest does not fly in the face of The Optional Stopping Theorem, which still applies in the limit as N gets ever-larger. This theorem proffers that, under certain conditions, the expected value of a martingale at a stopping time is equal to its initial value. A martingale is a stochastic process that satisfies a certain property related to conditional expectations. The Optional Stopping Theorem is an important result in probability theory and has many applications in finance, economics, and other fields. The theorem is based on the idea that, if a gambler plays a fair game and can quit whenever they like, then their fortune over time is a martingale, and the time at which they decide to quit (or go broke and are forced to quit) is a stopping time. The theorem says that the expected value of the gambler’s fortune at the stopping time is equal to their initial fortune. The Optional Stopping Theorem has many generalizations and extensions, and it is an active area of research in probability theory. The theorem becomes less applicable as N approaches 0 from the right.

Not all human endeavors are regarded with the same, isotropic importance by the individual as traditional analysis assumes—some are more important than others, and some are absolutely critical to the individual, deserving of analysis in their own, critical enclosure.

4.1. *The Goal of Median-Sorted Outcome after N Trials—The Process of Determining the Representative “Expected” Set of Outcomes*

Let us return to our gambler in Las Vegas, ready to head for the airport, who decides to wager his life savings—half of it on the roulette ball falling on the numbers 1–12, “First Twelve” on the board, and the other half on 13–24, “Second Twelve” on the board. The payoff is 2:1. If the ball falls in any slot numbered 1–24, he will be paid 2:1 on one half of his life savings, keeping the half of his life savings wagered, and losing the other half. In effect, he has risked his entire life savings to increase it by half should the ball fall 1–24.

Here, 24 of the 38 slots will produce this result for him, so there is $24/38$, or approximately 63.2% chance, he will see a 50% increase in his life savings. On one, single play, what one expects is the median outcome, which in this case, given the 0.632 probability of winning, he “expects” to win (in a European casino, he would have a better probability, as there is only one green zero, and his probabilities, therefore, increase to $24/37$, or approximately 64.9%).

This is a critical notion here—what one “expects” to occur is that outcome, over N trials, where half of the outcomes after N trials are better, and half are worse. This is the definition of “expectation” in an existential contest.

However, it is always the expectation, as demonstrated in [41], that, as N increases, this definition of expectation converges on the classical one accepted since the seventeenth century, as the sum of the probability-weighted outcomes. This “classical” definition of expectation is the expectation in the limit as the number of trials increases. Over a finite sequence of trials, the expectation is the median-sorted outcome, which converges to the classic definition as the number of trials, N , becomes ever greater. Thus, we can state, in all instances, the finite-sequence expectation is the actual expectation.

Examining this further, we find the divergence between the finite-sequence definition and the classical one is a function of skewness in the distribution of outcomes, whereas the finite-sequence definition (expectation as the median-sorted outcome) is always the correct “expectation” amount, regardless of N . The divergence between this result and the classical definition (the sum of the probability-weighted outcomes) is a function of skewness in the distribution of outcomes.

Most games of chance offer the player a probability < 0.5 of winning, with various payoffs depending on the player’s chances, even if the payoff is even money. However, in any game where the probability of profitability on a given play is < 0.5 , and the classical expectation is negative, so too will the finite-sequence expectation be negative. Convergence is rapidly attained and never in the player’s favor for any period of plays.

Earlier, we referred to a hypothetical proposition where one might win 1 unit with a probability of 0.9 and lose -10 units with a probability of 0.1. The classical expectation is to lose -0.1 per trial.

This is in contrast to the existential contest, which is only 1 trial long. Here, what one “expects” to happen is to win 1 unit. In fact, it can be shown that if one were to play and quit this game after 6 trials, one would “expect” to make 6 units (that is, half the outcomes of data distributed as such would show better or the same results, and half would show worse or the same results). After 7 trials, with these given parameters, however, the situation turns against the bettor.

We find that similar situations arise in capital markets. It is not at all uncommon for commodity traders to employ trend-following models. Such models suffer during periods of price congestion and generally profit handsomely when prices break out and enter a trend. It is not at all uncommon for such approaches to see only 15% to 30% winning trades and yet be profitable in the “long run.” Thus, an investor in such programs would be wise to have as large an “ N ” as possible. The early phase of being involved in such an endeavor

is likely to be a losing one, yet something that would payoff well in the long run—as the finite period becomes ever longer, the finite expectation converges on the classical one.

As we are discussing finite-length outcomes—“existential contests”, we wish to:

1. Determine what these finite-length expectations are, given a probability distribution we believe the outcomes comport to, so as to base on our sample data and the fitted parameters of that distribution.
2. Determine a likely N-length sequence of outcomes for what we “expect”; that is, a likely set of N outcomes representative of the outcome stream that would see half of the outcome streams of N length be better than, and half less than, such a stream of expected outcomes. It is this “expected, representative stream” that we will use as input to other functions, such as determining growth-optimal allocations.

Regarding point #2, returning to our hypothetical game of winning one unit with $p = 0.9$, and losing 10 units otherwise, here, the classical expectation is to lose 0.1 per play, with the finite expectation a function of N, and positive for the first 6 plays (assuming IID).

If one were to apply, say, the Kelly criterion, which requires a positive classical expectation, one would wager nothing. However, this asymptotic assumption under the finite-period expectation is found to be precisely the opposite, which would have the player start out risking a “fraction” equal to 1.0.

4.2. Determining Median-Sorted Outcome after N Trials

4.2.1. Calculating the Inverse CDF

To determine the median-sorted outcome, we need to first construct the inverse CDF function. An inverse cumulative distribution function (inverse CDF) is a mathematical function that provides the value of the random variable for which a given probability has been reached. It is also referred to as the quantile function or the percent-point function. In simpler terms, the inverse CDF essentially allows you to determine the value at which a specified probability is attained in a given distribution.

To derive the inverse CDF of the symmetrical GHD, we first start with the CDF of the symmetrical GHD, which is given by:

$$CDF(x) = F(x) = \frac{1}{2} \left[1 + \operatorname{sgn}(x - \mu) \left\{ 1 - \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^\alpha \right\}^{\frac{\lambda-1}{\alpha}} \right]$$

where:

α = concentration

δ = scale

μ = location

λ = shape

$\operatorname{sgn}(x - \mu)$ = sign function, the sign of $(x - \mu)$, or simply $\operatorname{sgn}(x - \mu) = \frac{x - \mu}{|x - \mu|}$.

Rearranging the equation to isolate the term inside the curly braces:

$$\left\{ 1 - \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^\alpha \right\}^{\frac{\lambda-1}{\alpha}} = \frac{2(F(x) - 1/2)}{\operatorname{sgn}(x - \mu)}$$

We raise both sides to the power of $\alpha/(\lambda - 1)$:

$$1 - \left(\frac{\delta}{\sqrt{\delta^2 + (x - \mu)^2}} \right)^\alpha = \left[\frac{2(F(x) - 1/2)}{\operatorname{sgn}(x - \mu)} \right]^{\frac{\alpha}{\lambda-1}}$$

Solving for $\sqrt{\delta^2 + (x - \mu)^2}$:

$$\sqrt{\delta^2 + (x - \mu)^2} = \delta \left[1 - \left\{ \left[\frac{2(F(x) - 1/2)}{\text{sgn}(x - \mu)} \right]^{\frac{\alpha}{\lambda-1}} \right\}^{1/\alpha} \right]$$

Solving for x :

$$x = \begin{cases} \mu - \delta \left[1 - \left\{ \left[\frac{2(F(x) - 1/2)}{-1} \right]^{\frac{\alpha}{\lambda-1}} \right\}^{1/\alpha} \right] & \text{if } F(x) < \frac{1}{2} \\ \mu + \delta \left[1 - \left\{ \left[\frac{2(F(x) - 1/2)}{1} \right]^{\frac{\alpha}{\lambda-1}} \right\}^{1/\alpha} \right] & \text{if } F(x) \geq \frac{1}{2} \end{cases}$$

The inverse CDF function takes a probability value, p , as input and returns the corresponding value of x , such that:

$$F^{-1}(p) = \begin{cases} \mu - \delta \left(1 - \left(\frac{2(1-p)}{-1} \right)^{1/\alpha} \right)^{1/\lambda} & \text{if } p < \frac{1}{2} \\ \mu + \delta \left(1 - \left(\frac{2(p-1/2)}{1} \right)^{1/\alpha} \right)^{1/\lambda} & \text{if } p \geq \frac{1}{2} \end{cases}$$

Simplifying the expressions inside the parentheses:

$$F^{-1}(p) = \begin{cases} \mu - \delta \left(\frac{2p-1}{2p-1} \right)^{1/\alpha\lambda} & \text{if } p < \frac{1}{2} \\ \mu + \delta \left(\frac{2p-1}{2-2p} \right)^{1/\alpha\lambda} & \text{if } p \geq \frac{1}{2} \end{cases}$$

Simplifying the expressions using the fact that $\text{SIGN}(p - 0.5) = \pm 1$:

$$x = F^{-1}(p) = \mu + \text{SIGN}(p - 0.5) \times \delta \times \left(\frac{2|p - 0.5|}{1 - 2|p - 0.5|} \right)^{1/\alpha\lambda}$$

This is our inverse CDF for the symmetrical GHD.

4.2.2. Worst-Case Expected Outcome

Since we sought to construct a “representative” stream of outcomes, a stream of outcomes we would “expect” to be typical (half better, half worse) of what we would expect to happen over N trials, and given the critical nature of the “worst-case outcome” in [41,47,58–60,62] for which we wish to apply this expected stream of outcomes to, we thus wished to find that point on the distribution that we would expect be seen, typically, over N trials.

Using the inverse CDF, we can find this point, this “worst-case expected outcome,” as the value of the inverse CDF for the input:

$$\text{worst-case expected outcome} = (1/N)^N$$

We can use this to examine against the results of our representative stream of outcomes at the median-sorted outcome.

4.2.3. Calculating the Representative Stream of Outcomes from the Median-Sorted Outcome

We proceeded to discern the median-sorted outcome. We generated a large sequence of random variables (0..1) to process via our inverse CDF to yield the random outcomes from our distribution, which we took N at a time:

- a. For each of N outcomes, we determined the corresponding probability by taking the value returned by the inverse CDF as input to the CDF. When using the inverse CDF function, the random number generated is not the probability of the value returned by the inverse CDF. Instead, the random number generated is used as the input to the inverse CDF function to obtain a value from the distribution. This value is a random variable that follows the distribution specified by the CDF. To determine the probability of this value, one would need to use the CDF function and evaluate it at the value returned by the inverse CDF function. The CDF function gives the probability that a random variable is less than or equal to a specific value. Therefore, if the CDF function is evaluated at the value returned by the inverse CDF function, the probability of that value will be obtained.
- b. For our purposes, however, we took any of the probabilities that were >0.5 and took the absolute value of their difference to 0.5. We called this p' :
 - i. $p' = p$ if $p \leq 0.5$
 - ii. $p' = |1 - p|$ if $p > 0.5$
- c. We took M sets of N outputs, corresponding to outputs drawn from our fitted distribution as well as the corresponding p' probabilities.
- d. For each of these M sets of N outcome— p' combinations, we summed the outcomes, and took the product of all of the p' values.
- e. We sorted these M sets based on their summed outcomes.
- f. We then took the sum of these M product of p' values. We called this the “SumOfProbabilities”.
- g. For each M set, we proceeded sequentially through them, where each M has a cumulative probability, specified as the cumulative probability of the previous M in the (outcome-sorted) sequence, plus the probability of that M divided by the SumOfProbabilities.
- h. It is the final calculation, we sought that M whose final calculation was closest to 0.5. This is our median-sorted outcome, and the values for N that comprise it are our typical, median-sorted outcome, our expected, finite-sequence set of outcomes.

4.2.4. Example—Calculating the Representative Stream of Outcomes from the Median-Sorted Outcome

Here, we demonstrate the steps from Section 4.2.3. We assumed a finite stream of 7 outcomes in length (since we used a distribution of daily S&P500 returns for the first half of 2023, we assume we are modeling an existential contest of 7 upcoming days in the S&P500). Thus, N = 7.

We calculated 1000 such sets of these 7 outcomes and, therefore, M = 1000, and we proceeded to create, for each M, 7 random values (0..1), representing the inputs to the inverse CDF, and the output for that N, then deriving the probability by taking the output for that N as the input to the CDF. This final probability number was processed into a p' value.

For each set of 7 probability–outcome pairs, we took the product of the p' values and the sum of the outcomes (Table 3).

Table 3. Sorted outcomes. Outcomes are labeled as columns o1..o7, and each row is one of 1000 of the M sets of N = 7 outcomes.

o1	o2	o3	o4	o5	o6	o7	Σo	$\prod p$	Cum Sum Probs	
0.01898	-0.01143	-0.02245	-0.00848	-0.00245	0.00988	-0.04847	-0.064417766	0.000000000791	0.000367729532	
-0.01148	-0.01351	-0.01114	-0.00849	-0.00428	-0.0135	-0.00353	-0.065923437	0.000000015425	0.000361021142	
-0.023	-0.00209	-0.00094	-0.01424	-0.0132	-0.01449	-0.00144	-0.069414377	0.00000020100	0.000230227464	
-0.00535	-0.01323	-0.00791	-0.01091	0.00768	-0.02873	-0.01846	-0.076917579	0.000000002948	0.000059794713	
-0.00817	-0.0109	-0.00531	-0.01709	-0.01115	-0.0179	-0.00953	-0.080049157	0.000000003151	0.000034795240	
0.01658	-0.0653	-0.00424	-0.01429	-0.01636	-0.01301	0.01241	-0.084203152	0.000000000228	0.000008072898	
-0.00544	-0.11487	0.01205	-0.01222	-0.00734	0.01483	-0.00471	-0.117685371	0.000000000724	0.000006137397	<-- Previous Cum Sum Probs + current $\prod p$ /SumOfProbs
									0	
									0.000117933	<-- Sum of $\prod p$'s (SumOfProbs)

We next calculated the cumulative sum of the probabilities for each M. This is simply the previous M in the sorted sequence, plus the p' product of the N p' values in the outcome— p' pairs, divided by the sum of such p' values over all M (“SumOfProbs”).

Lastly, as shown in Table 4, we found the M in the sorted set of M sets that corresponds to a cumulative sum of probabilities value closest to 0.5. This is the median-sorted outcome. Its output values (o1...o7) represent the “typical” expected outcomes over N trials (Table 4).

Table 4. Sorted outcomes with median selected.

o1	o2	o3	o4	o5	o6	o7	$\sum o$	$\prod p$	Cum Sum Probs
−0.01076	−0.0096	0.01428	0.03664	−0.00401	−0.01816	0.00873	0.017127894	0.000000003902	0.503768567507
−0.00529	−0.00806	0.01028	0.01246	−0.00586	−0.00624	0.0196	0.016892532	0.000000096482	0.503735481158
−0.01023	−0.00529	−0.00461	0.01046	0.00852	0.01987	−0.00196	0.01676511	0.000000268421	0.502917368488
−0.00667	−0.00638	−0.00662	−0.00515	0.01318	0.01693	0.01147	0.016764245	0.000000106581	0.500641320047
0.00863	0.02348	−0.00931	−0.00549	0.01595	−0.00801	−0.00849	0.01676081	0.000000026072	0.499737576467
−0.00721	−0.01002	0.02994	−0.01062	−0.00915	−0.01151	0.03524	0.016677462	0.000000000506	0.499516499541
−0.00352	0.01047	0.02715	−0.00517	−0.01393	−0.01018	0.01182	0.016631721	0.000000023517	0.499512210603
−0.0099	0.01316	−0.00314	0.0184	−0.00324	0.00952	−0.00819	0.01661394	0.000000135827	0.499312798

In this instance of 1000 samples of sequences of N outcomes per the parameters for our symmetrical GHD representation of S&P500 closing percentage changes, we found the sample set, the “typical” set of seven closing prices—what we would expect, where half the time we would see a better sequence, and half the time a worse one, as:

−0.006670152, −0.00638097, −0.006615831, −0.005151115, 0.013180558, 0.01693477, 0.011466985

Of note, our expected worst-case outcome was calculated as −0.00642. As we obtained an outcome that is at least as negative as this, we are not deluding ourselves on what might be the worst-case outcome for the “typical”, i.e., the median outcome stream of 7 outputs.

This set of outcomes represents the expected—the typical set of outcomes from data distributed in the first half of 2023 for daily percentage changes in the S&P500.

Further, for 7 trials in this distribution of outcomes, we can expect to see the $\sum o$ outcome of 0.016764 as our non-asymptotic expectation. Of note, we used a symmetrical distribution in this paper and example, and as mentioned earlier, the divergence between the actual, finite expectation and the classical, asymptotic expectation is a function of the skewness of the distribution, which was 0 in our example. Nevertheless, the concept remains, and the procedure illuminated this for more perverse and skewed distributions.

4.3. Optimal Allocations

4.3.1. Growth-Optimal Fraction

Earlier, mention was made of determining the growth-optimal fraction from a sample data set, and how such results might diverge from its non-asymptotic determination.

Individuals and individual entities enmeshed in existential conflicts may very well be interested in finite-length growth optimization.

Consider, for example, a proposition with p (0.5) of winning 2 units versus losing 1 unit. The traditional method (even by way of the Kelly criterion, as the absolute value of the largest loss here is 1 unit) would determine asymptotic growth maximized by risking 0.25 of the stake on each play. However, when examined under the lens of finite-length existential contests, terminal growth is maximized if the length of the contest is 1 play at a fraction of 1.0, and if $N = 2$ plays, at a fraction of 0.5 per play, ad infinitum, to the asymptotic 0.25 risk of stake as N grows ever larger.

In determining optimal fractions of finite-length existential contests, the input data (the values of −0.006670152, −0.00638097, −0.006615831, −0.005151115, 0.013180558, 0.01693477, and 0.011466985 in our example, above) are what ought to be used for determining the optimal fraction. The data determine the expectation in such finite-length trials; that is, the data are germane to also determining these other facets—the data one “expects” over those N trials—and seeking to maximize them.

4.3.2. Return/Risk Optimal Points

The amount to risk that is growth-optimal is return/risk-optimal only asymptotically. For finite-length streams—existential contests—the authors of [58] demonstrated the two points on the spectrum of percentage risk that maximize the ratio of return to risk both as a tangent to the function on this spectrum as well as the inflection point “left” of the peak of the function.

Later, in [59], this was extended to the determination of multiple, simultaneous contests (portfolios).

Particularly in determining these critical points of return/risk maximization, which are length-dependent, utilizing the input data that would be “expected” over such length of trials, as demonstrated herein, is more appropriate to the individual engaged in an existential contest.

4.3.3. Risk of Ruin/Drawdown Calculations

Similarly, the determination of the risk of ruin, defined as drawdown on cumulative equity of a given percent of capital from an initial starting point, as well as the risk of drawdown, defined as drawdown on cumulative equity of a given percent of capital from a high watermark point, are more accurately calculated for the course of existential contests using a sample of outcomes determined by this method, as opposed to the universe of potential outcomes, as dictated by the full distribution of outcomes. In [63], it is demonstrated that these calculations depend on a specific-length horizon (and, in fact, demonstrate that risk of drawdown = 1.0, certainty, for any defined percentage for drawdown as the length of time approaches infinity). As these calculations are also time-specific, the accurate calculation requires the use of a realistic time-specific sample set, as provided herein.

4.3.4. Growth Diminishment in Existential Contests

In [60], a technique was presented, whereby examining an isomorphic mapping of the fraction of risk to the cosine of the standard deviation in percentage inputs to their arithmetic average allows one to examine growth functions sans a “risk fraction”, and instead in terms of variance in outcomes. The presented technique examined geometric growth functions, where growth was a negative property. Such functions might be the cumulative debt of a nation, the growth of a pathogen in a population or within an individual organism, etc.

On the function plotted in terms of the fraction of a stake to risk versus growth, as there is a growth-optimal point, there is necessarily a point where the function begins decreasing and where growth becomes negative by risking too much (for example, repeatedly risking beyond 0.5 in the 2:−1 payoff game mentioned earlier with p (0.9), one is certain to go broke as they continue to risk 0.5 of their stake or more, even in this very favorable game).

The isomorphic mapping provided in this paper allows one to examine such traits of the growth function without regard to the notion of a percentage of stake to risk, but rather of tweaking the variance.

As most endeavors to break malevolent growth functions are inherently time-sensitive (i.e., of finite N —true “existential contests”), the notion of using the median-expected outcome stream over N trials is very germane to this exercise, as well.

5. Further Research

We have demonstrated how to fit an extensible, albeit symmetrical distribution to economic data series, and done so with techniques applicable to other distributional forms. As discussed, although the two-piece, skew-normal distributions and the odd log-logistic-type distributions offer greater flexibility than the GHD and, therefore, are generally considered to better represent financial time series data, we have illustrated our methods on the symmetrical GHD for simplicity, but these methods can be applied to any distributional form.

However, it is the very parameter of skewness—the differences between mean and median—that accounts for the difference between conclusions drawn for existential contests and the classical, asymptotic assessments. Generally speaking, the greater the skewness, the more likely these differences will be highlighted, and the more likely to see cases where, for example, the expectation may be positive for the existential contest, but negative asymptotically, and vice versa. Thus, this paper leaves wide open the ripened area of applying the techniques to data modeled per the two-piece, skew-normal distributions and the odd log-logistic-type distributions.

6. Conclusions

We were concerned with “individuals” (as one or a collection of individuals with a common goal) involved in finite-length sequences of stochastic outcomes—existential contests—where the distribution of such outcomes is characterized by fat tails.

Once fitted, we demonstrated the process to determine the “expected” terminal outcome over the finite-length sequence, and the “expected” or “typical” outcomes that characterize such. We have demonstrated how to obtain the expected set of discrete outcomes for the given form for a given length of an existential contest.

This is important in analyzing the expectation of finite sequences, which differ from the classical expectation of the sum of the probabilities times their outcomes, with the former being what is germane to the “individual” in an existential contest, whereby the former is always the correct expectation, one that approaches the classical one asymptotically.

Such “expectation,” what the “individual” can “expect”, as that which has half the outcomes better half worse, also yields a typical stream of outputs achieving this expected outcome.

It is this stream of outputs which, to the individual in the existential contest, are the important inputs, considering that they are what the individual “expects” for such finite-length calculations as growth-optional fractions, maximizing return to risk, calculating ruin or drawdown in the existential contest, or even for those whose charge is growth diminishment with a finite deadline, given that life itself, despite the ever-present asymptotes, is rife with existential contests.

Funding: This research received no external funding.

Data Availability Statement: Data are also available upon request from the author.

Conflicts of Interest: The author was employed by Exsuperatus. The author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Exsuperatus had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Marshall, A. *Principles of Economics*; Macmillan and Co.: New York, NY, USA, 1890.
2. Marshall, A. *Industry and Trade: A Study of Industrial Technique and Business Organization*; Macmillan and Co.: New York, NY, USA, 1919.
3. Bachelier, L. Théorie de la spéculation. *Ann. Sci. de l'École Norm. Supérieure* **1900**, *17*, 21–86. [[CrossRef](#)]
4. Lévy, P. *Théorie de L'addition des Variables Aléatoires*; Gauthier-Villars: Paris, France, 1937.
5. Lévy, P. *Processus Stochastiques et Mouvement Brownien*; Gauthier-Villars: Paris, France, 1948.
6. Feller, W. *Theory of Probability*; John Wiley & Sons: New York, NY, USA, 1950.
7. Pareto, V. Pareto Distribution and its Applications in Economics. *J. Econ. Stud.* **1910**, *5*, 128–145.
8. Mandelbrot, B. The Variation of Certain Speculative Prices. *J. Bus.* **1963**, *36*, 394. [[CrossRef](#)]
9. Mandelbrot, B.; Hudson, R.L. *The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward*; Basic Books: New York, NY, USA, 2004.
10. Samorodnitsky, G.; Taqqu, M.S. *Stable Non-Gaussian Random Processes*; Chapman and Hall: Boca Raton, FL, USA, 1994.
11. Rachev, S.; Mittnik, S. *Stable Paretian Models in Finance*; Wiley: New York, NY, USA, 2000.
12. Rachev, S.T. *Handbook of Heavy Tailed Distributions in Finance*; Elsevier Science B.V.: Amsterdam, The Netherlands, 2003.
13. Nolan, J.P. Financial Modeling with Heavy-Tailed Stable Distributions. *WIREs Comput. Stat.* **2014**, *6*, 45–55. [[CrossRef](#)]

14. Barndorff-Nielsen, O.E. Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **1977**, *353*, 401–409. [[CrossRef](#)]
15. Podgórski, K.; Wallin, J. Convolution-Invariant Subclasses of Generalized Hyperbolic Distributions. *Commun. Stat. Theory Methods* **2015**, *45*, 98–103. [[CrossRef](#)]
16. Eberlein, E.; Keller, U. Hyperbolic distributions in finance. *Bernoulli* **1995**, *1*, 281–299. [[CrossRef](#)]
17. Prause, K. The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures. Ph.D. Thesis, University of Freiburg, Freiburg im Breisgau, Germany, 1999.
18. Küchler, U.; Neumann, K. Stock Returns and Hyperbolic Distributions. *ASTIN Bull. J. Int. Actuar. Assoc.* **1999**, *29*, 3–16. [[CrossRef](#)]
19. Eberlein, E.; Keller, U.; Prause, K. New Insights into Smile, Mispricing, and Value at Risk: The Hyperbolic Model. *J. Bus.* **1998**, *71*, 371–405. [[CrossRef](#)]
20. Núñez-Mora, J.A.; Sánchez-Ruenes, E. Generalized Hyperbolic Distribution and Portfolio Efficiency in Energy and Stock Markets of BRIC Countries. *Int. Bus. Econ. Res. J.* **2014**, *13*, 299–310. [[CrossRef](#)]
21. Huang, C.K.; Chinhamu, K.; Huang, C.S.; Hammujuddy, J. Generalized Hyperbolic Distributions and Value-at-Risk Estimation for the South African Mining Index. *Int. Bus. Econ. Res. J.* **2014**, *13*, 265–274. [[CrossRef](#)]
22. Wang, C.; Liu, K.; Li, B.; Tan, K. Portfolio optimization under multivariate affine generalized hyperbolic distributions. *Int. Rev. Econ. Financ.* **2022**, *80*, 49–66. [[CrossRef](#)]
23. Kostovetsky, L.; Kostovetska, B. Distribution Analysis of S&P 500 Financial Turbulence. *J. Math. Financ.* **2023**, *13*, 67–88.
24. Hu, W.; Kercheval, A. Risk Management with Generalized Hyperbolic Distributions. Available online: <https://www.math.fsu.edu/e-prints/archive/paper321.pdf> (accessed on 1 November 2023).
25. Klebanov, L.; Rachev, S.T. ν -Generalized Hyperbolic Distributions. *J. Risk Financ. Manag.* **2023**, *16*, 251. [[CrossRef](#)]
26. Jamalizadeh, A.; Arabpour, A.R.; Balakrishnan, N. A Generalized Skew Two-Piece Skew-Normal Distribution. *Stat. Papers* **2011**, *52*, 431–446. [[CrossRef](#)]
27. Shafiei, S.; Doostparast, M.; Jamalizadeh, A. The Alpha-Beta Skew Normal Distribution: Properties and Applications. *Statistics* **2016**, *50*, 338–349. [[CrossRef](#)]
28. Rasekhi, M.; Chinipardaz, R.; Alavi, S.M.R. A Flexible Generalization of the Skew Normal Distribution Based on a Weighted Normal Distribution. *Stat. Methods Appl.* **2016**, *25*, 375–394. [[CrossRef](#)]
29. Rasekhi, M.; Hamedani, G.G.; Chinipardaz, R. A Flexible Extension of Skew Generalized Normal Distribution. *Metron* **2017**, *75*, 87–107. [[CrossRef](#)]
30. Fernández, C.; Steel, M.F.J. On Bayesian Modeling of Fat Tails and Skewness. *J. Am. Stat. Assoc.* **1998**, *93*, 359–371. [[CrossRef](#)]
31. Castillo, N.O.; Gómez, H.W.; Leiva, V.; Sanhueza, A. On the Fernández-Steel Distribution: Inference and Application. *Comput. Stat. Data Anal.* **2011**, *55*, 2951–2961. [[CrossRef](#)]
32. Rigby, R.A.; Stasinopoulos, M.D.; Heller, G.Z.; De Bastiani, F. *Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2019.
33. Farzammehr, M.A.; Zadkarami, M.R.; McLachlan, G.J.; Lee, S.X. Skew-Normal Bayesian Spatial Heterogeneity Panel Data Models. *J. Appl. Stat.* **2019**, *47*, 804–826. [[CrossRef](#)] [[PubMed](#)]
34. Kim, H.-J. On a Class of Two-Piece Skew-Normal Distributions. *Statistics* **2005**, *39*, 537–553. [[CrossRef](#)]
35. Charenza, W.; Díaz, C.; Makarova, S. *Choosing the Right Skew Normal Distribution: The Macroeconomist's Dilemma*; Working Paper No. 15/08; University of Leicester: Leicester, UK, May 2015.
36. Neethling, A.; Ferreira, J.; Bekker, A.; Naderi, M. Skew Generalized Normal Innovations for the AR(p) Process Endorsing Asymmetry. *Symmetry* **2020**, *12*, 1253. [[CrossRef](#)]
37. da Braga, A.S.; Cordeiro, G.M.; Ortega, E.M.; da Cruz, J.N. The Odd Log-Logistic Normal Distribution: Theory and Applications in Analysis of Experiments. *J. Stat. Theory Pract.* **2016**, *10*, 311–335. [[CrossRef](#)]
38. Prataviera, F.; Cordeiro, G.M.; Ortega, E.M.M.; Suzuki, A.K. The Odd Log-Logistic Geometric Normal Regression Model with Applications. *Adv. Data Sci. Adapt. Anal.* **2019**, *11*, 1950003. [[CrossRef](#)]
39. Pascal, B.; Fermat, P. Traité du triangle arithmétique. *Corresp. Blaise Pascal Pierre Fermat* **1654**, *2*, 67–82.
40. Huygens, G. *Libellus de Ratiociniis in ludo aleae (A Book on the Principles of Gambling)*; Original Latin transcript translated and published in English; S. KEIMER for T. WOODWARD: London, UK, 1714.
41. Vince, R. Expectation and Optimal f: Expected Growth with and without Reinvestment for Discretely-Distributed Outcomes of Finite Length as a Basis in Evolutionary Decision-Making. *Far East J. Theor. Stat.* **2019**, *56*, 69–91. [[CrossRef](#)]
42. Bernoulli, D. Specimen theoriae novae de mensura sortis (Exposition of a new theory on the measurement of risk). *Comment. Acad. Sci. Imp. Petropolitanae* **1738**, *5*, 175–192, Translated into English: *Sommer, L. Econometrica* **1954**, *22*, 23–36.
43. Keynes, J.M. *A Treatise on Probability*; Macmillan: London, UK, 1921.
44. Williams, J.B. Speculation and the carryover. *Q. J. Econ.* **1936**, *50*, 436–455. [[CrossRef](#)]
45. Kelly, J.L., Jr. A new interpretation of information rate. *Bell Syst. Tech. J.* **1956**, *35*, 917–926. [[CrossRef](#)]
46. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656. [[CrossRef](#)]
47. Vince, R. *Portfolio Management Formulas*; John Wiley & Sons, Inc.: New York, NY, USA, 1990.
48. Bellman, R.; Kalaba, R. On the role of dynamic programming in statistical communication theory. *IEEE Trans. Inf. Theory* **1957**, *3*, 197–203. [[CrossRef](#)]

49. Breiman, L. Optimal gambling systems for favorable games. In *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability*; Neyman, J., Ed.; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 65–78.
50. Latane, H.A. Criteria for choice among risky ventures. *J. Political Econ.* **1959**, *67*, 144–155. [[CrossRef](#)]
51. Latane, H.; Tuttle, D. Criteria for portfolio building. *J. Financ.* **1967**, *22*, 362–363. [[CrossRef](#)]
52. Thorp, E.O. *Beat the Dealer: A Winning Strategy for the Game of Twenty-One*; Vintage: New York, NY, USA, 1962.
53. Thorp, E.O. The Kelly Criterion in blackjack, sports betting, and the stock market. In *Proceedings of the 10th International Conference on Gambling and Risk Taking*, Montreal, QC, Canada, 31 May–4 June 1997.
54. Samuelson, P.A. The “fallacy” of maximizing the geometric mean in long sequences of investing or gambling. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 2493–2496. [[CrossRef](#)]
55. Samuelson, P.A. Why we should not make mean log of wealth big though years to act are long. *J. Bank. Financ.* **1979**, *3*, 305–307. [[CrossRef](#)]
56. Goldman, M.B. A negative report on the “near optimality” of the max-expected-log policy as applied to bounded utilities for long-lived programs. *J. Financ. Econ.* **1974**, *1*, 97–103. [[CrossRef](#)]
57. Merton, R.C.; Samuelson, P.A. Fallacy of the lognormal approximation to optimal portfolio decision-making over many periods. *J. Financ. Econ.* **1974**, *1*, 67–94. [[CrossRef](#)]
58. Vince, R.; Zhu, Q. Optimal betting sizes for the game of blackjack. *J. Invest. Strateg.* **2015**, *4*, 53–75. [[CrossRef](#)]
59. Lopez de Prado, M.; Vince, R.; Zhu, Q.J. Optimal Risk Budgeting Under a Finite Investment Horizon. *Risks* **2019**, *7*, 86. [[CrossRef](#)]
60. Vince, R. Diminution of Malevolent Geometric Growth Through Increased Variance. *J. Econ. Bus. Mark. Res.* **2020**, *1*, 35–53.
61. Klebanov, L.; Rachev, S.T. *The Global Distributions of Income and Wealth*; Springer: Berlin/Heidelberg, Germany, 1996.
62. Vince, R. *The Mathematics of Money Management*; John Wiley & Sons, Inc.: New York, NY, USA, 1992.
63. Vince, R. *The Handbook of Portfolio Mathematics: Formulas for Optimal Allocation & Leverage*; John Wiley & Sons, Inc.: New York, NY, USA, 2007.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.