A NOVEL EXTENSION OF ENTROPY MEASURES IN QUANTIFYING FINANCIAL MARKETS UNCERTAINTY: THEORY AND APPLICATIONS.

Williema Nangolo[1*], Rodrique Gnitchogna[2]

[1]Department of Computing, Mathematical & Statistical Science, University of Namibia, Private Bag 13301 Windhoek, Namibia, e-mail: wnangolo@unam.na. Telephone: +264-61-2063949, Fax: +264612063791*
[2]Department of Computing, Mathematical & Statistical Science, University of Namibia, Private Bag 13301 Windhoek, Namibia, e-mail: rgnitchoga@unam.na. Telephone: +264-61-2063962, Fax: +264612063791

Corresponding Author:
**E mail: wnangolo@unam.na*

Abstract
The global financial market is characterized by inherent and evolving uncertainty. Measuring this uncertainty plays a crucial role in managing risk associated with financial derivatives. Various mathematical models, including robust risk measures, model risk measures, and locally risk-minimizing strategies, have been employed to quantify this uncertainty. This paper contributes to this ongoing research by proposing novel approaches to quantify uncertainty in financial derivatives, specifically by leveraging entropy measures with stochastic probability density functions. Traditionally, entropy models have relied on Gaussian probability density functions. This paper proposes an alternative approach using stochastic probability density functions, to capture the inherent randomness of uncertainty in financial markets. Furthermore, the use of this developed stochastic density function will achieve linear and sub-linear scaling without relying on the sparsity of the density matrix nor on the design of the subsystem interaction in embedding schemes. We demonstrate that this approach adheres to key entropy properties and can be extended to various entropy families. Empirical results show that the proposed model using stochastic probabilities outperforms models using normal probabilities, potentially representing a significant advancement in quantifying uncertainty with entropy measures.

KEY WORDS: *Quantifying Uncertainty, Entropy Measure, Normal Density Function, Stochastic Density Function. Financial derivatives.*

A.M.S. SUBJECT CLASSIFICATION: 28D20, 81S07, 91G45, 05Cxx, 60Exx

The authors declare that there is no conflict of interest regarding the publication of this paper.

Introduction

In the ever-changing world of finance, where fortunes and losses are made under the shadow of uncertainty, entropy emerges as a powerful tool for measuring the unknowable. In the context of financial derivatives, these complex instruments designed to manage risk, ironically, carry their own inherent risk stemming from the unpredictable nature of the underlying assets [1], [2]. Accurately gauging and accounting for this uncertainty is crucial for navigating the financial tightrope with confidence, and entropy steps in as a reliable guide [3].

Why Measure Uncertainty in Derivatives?

Financial derivatives like options, futures, swaps, and forwards derive their value from the price of the underlying asset. Acknowledging that predicting the future with perfect accuracy seems to be impossible, because a multitude of factors, from geopolitical tensions to economic data releases, can send markets into unpredictable gyrations. This inherent volatility translates into uncertainty about the ultimate payoff of a derivative contract [4], [5].

Measuring uncertainty allows market participants to:

Quantify risk: By understanding the potential range of outcomes associated with a derivative position, investors can make informed decisions about hedging strategies and risk mitigation.

Price derivatives accurately: Accurately reflecting uncertainty in pricing models leads to fairer and more efficient markets.

Improve portfolio management: Uncertainty measures can be used to optimize portfolio allocation and diversification, ensuring a balance between risk and return.

Entropy to the Rescue: Measuring the uncertainty

So, how exactly does entropy help us grapple with the uncertainty in derivatives? Imagine entropy as a gauge on your car's dashboard. A low entropy reading, with the needle barely budging, indicates a smooth, predictable highway ahead. Conversely, a high entropy reading, with the needle bouncing wildly, warns of a treacherous, obstacle-ridden road. In the financial world, entropy acts as that needle, quantifying the" bumpiness" or unpredictability of the market landscape.

There are several methods that exist for measuring uncertainty in derivatives using entropy [6],[7], each with own strengths and limitations, this study acknowledges a few below:

1. **Volatility:** This widely used metric measures the degree of price fluctuations in the underlying asset. Higher volatility implies greater uncertainty about future prices.
2. **Value at Risk (VaR):** VaR estimates the maximum potential loss of a portfolio within a given confidence level over a specific time horizon. It provides a quantifiable snapshot of the portfolio's exposure to uncertainty.
3. **Scenario analysis:** This qualitative approach involves constructing various hypothetical scenarios representing potential market outcomes and assessing their impact on derivative positions.
4. **Monte Carlo simulations:** This stochastic method simulates thousands of possible price paths for the underlying asset, generating a probability distribution of potential payoffs for the derivative.

Uncertainty is an inescapable reality in the world of finance, and nowhere is its presence more keenly felt than in the realm of derivatives. By employing effective methods for measuring and managing uncertainty using entropy, market participants can navigate the ever-shifting landscape of financial markets with greater confidence and make informed decisions that protect their capital and maximize their returns. Remember, while uncertainty may cast a long shadow, it is by embracing its presence and developing strategies to mitigate its impact that we can truly unlock the potential of financial derivatives as powerful tools for managing risk and achieving financial goals.

Concept of entropy model

Consider a stochastic process defined by a collection of random variables indexed by time. In a discrete time, a stochastic process $X = \{x_n, n = 1,2,3, \dots\}$ can be used to define the information entropy with the probability mass function $f(\cdot)$ given by

$$H(X) = \sum_{x} f(x) \log f(x), \tag{1}$$

and in a continuous time, stochastic process $X = \{x_t, 0 \leq t \leq \infty\}$ extend the information entropy to

$$H(X) = \int_R f(x) \log f(x) dx, \tag{2}$$

where $H(X)$ is the information entropy, $f(x)$ is the probability distribution and $\log$ is a natural logarithm or a logarithm to base 2 [1]. The information entropy what is commonly referred to as Shannon entropy by Schwill and Shannon (1948) [8], whom stated "Entropy is the measure of uncertainty in random variables".

Many extensions of Shannon entropy have been introduced. Related to Shannon entropy is Relative entropy or Kullback-Leibler divergence,

$$H(X||Y) = \int_R g(x) \log \frac{g(x)}{f(x)} dx \quad (3)$$

where $f(x)$ and $g(x)$ are probability density functions of probability measures X and Y respectively [9]. Other notable entropy is Renyi and Tsallis entropy which is specified as

$$H_q(X) = \frac{1}{1-\gamma} \log \int_R f(x)^\gamma dx \quad (4)$$

where γ is the order of the entropy [10]. Many other entropy measures such as Sample, Mutual information, and transfer, also exist.

Common properties of all these entropies are:

- The Applicability in both discrete and continuous cases.
- The range is from 0 to 1 unity. The measures are normalized to 0 in the case of independency, and the modulus of the measure is 1 in the case of measurable exact relationship between the random variables.
- In the case of a bivariate normal distribution, the measure of dependence has a simple relationship with the correlation coefficient.
- It measures not only the distance but also the divergence.

Literature Review

Entropy, a fundamental concept in information theory. It attempts to quantifies the inherent uncertainty associated with a probability distribution. It does this by capturing both the randomness within the distribution and the information content embedded in its higher-order moments [11],[1],[3]. While entropy has been used for uncertainty quantification, existing literature often relies solely on the normal probability density function (NPDF) as a penalty term [12]. This approach, while established, has limitations in capturing complex real-world uncertainties. Notably, it treats weights as discrete normal probabilities and employs entropy as a penalty to push them towards an equally weighted distribution [13], potentially overlooking valuable information embedded in other probability density functions [14]. Instead of the traditional penalty term approach using the normal distribution, there is a leverage of entropy directly as a quantifier of uncertainty in financial derivative markets. This aligns with information theory principles [15], capturing richer information content beyond the limitations of the normal distribution. Given its capacity to quantify inherent randomness and information content, Shannon entropy has garnered widespread recognition as a valuable measure in financial derivative management [16], uncertainty quantification [11], and utility theory [17].

Shannon Entropy: A Pillar of Uncertainty Quantification

Shannon entropy, introduced by Claude Shannon in 1948 [8], has become a cornerstone of information theory and a widely used tool in diverse academic disciplines. It quantifies the uncertainty associated with a random variable, essentially measuring the average information needed to predict its outcome [18],[3]. Defined differently for continuous and discrete cases [19], Shannon entropy captures the inherent randomness of systems ranging from thermodynamics to financial markets [18]. However, it's crucial to acknowledge Shannon entropy's limitations. Its dependence on chosen parameters and lack of an invariant measure can lead to inconsistencies [20]. Notably, Kullback-Leibler divergence [21] highlights this shortcoming, demonstrating that Shannon entropy is not inherently symmetric in comparing two probability distributions [9]. Despite these limitations, Shannon entropy remains a valuable tool due to its simplicity and interpretability. Furthermore, several alternative measures address its shortcomings. Renyi entropy [9], a generalization with parameter 'γ', offers additional flexibility. When 'γ' approaches 1, it converges to Shannon entropy. Tsallis entropy [10], on the other hand, yields power-law distributions, making it suitable for situations where traditional exponential distributions fall short. Moreover, the Maasoumi-Racine [22] measure presents a valuable alternative, particularly for time series analysis [23]. Applicable to both discrete and continuous data, it ranges from 0 to 1 and excels at capturing non-linear dependencies between random variables [24],[25], something Shannon entropy cannot directly address. Nevertheless, while Shannon entropy has limitations, its simplicity and interpretability make it a foundational tool in diverse academic fields [26]. The development of alternative measures like Renyi [9], Tsallis [10], and Maasoumi-Racine [22], enriches the toolbox for quantifying uncertainty, allowing researchers to choose the most appropriate measure for their specific needs and data characteristics [27]. These measures continue to push the boundaries of uncertainty quantification, leading to deeper insights across various academic disciplines.

Kullback-Leibler Divergence: Unveiling Uncertainty through Informational Disparity

While not explicitly measuring uncertainty, Kullback-Leibler [28] (KL) divergence offers a critical lens into quantifying this elusive concept within diverse academic disciplines. Its ability to assess the informational disparity between a presumed distribution (X) and the true underlying one (Y) unlocks valuable insights into uncertainty [9]. KL divergence as an Informational Gap [29]: at its core, KL divergence, denoted by KL $(X||Y)$, quantifies the additional information (in bits) required to encode samples from the true distribution X using the code designed for our assumed distribution Y. This" extra information" acts as a proxy for the degree of uncertainty associated with relying on Y to represent X [30].

Unveiling Uncertainty through Model Evaluation:

Machine Learning: In model training, minimizing KL divergence between predicted and true data distributions becomes a key objective. This ensures the model captures the underlying data accurately, effectively reducing prediction uncertainty [31].

Hypothesis Testing: Comparing statistical models involves minimizing KL divergence between their predicted and observed distributions. This identifies the model that best represents the observed data, aiding in uncertainty reduction within the chosen model framework [32].
Beyond Traditional Metrics:
Unlike typical error metrics focused on point estimates, KL divergence tackles the entire distribution [33]. It captures not only the average prediction error but also the spread and shape of the error distribution, providing a richer understanding of uncertainty [34],[31].

Academic Nuances and Considerations:
Non-negativity: KL divergence is always non-negative, indicating how much" worse" Y is than X, not how" good" X is on its own. This necessitates complementary metrics for absolute uncertainty assessment [35],[36].
Computational Cost: Calculating KL divergence can be computationally expensive for complex distributions, demanding trade-offs between accuracy and efficiency [37].
Interpretability: While conceptually powerful, interpreting KL divergence in specific contexts requires domain knowledge and additional analysis to extract meaningful insights [38].
In conclusion, KL divergence, despite not directly measuring uncertainty, offers a unique perspective by quantifying the informational disparity induced by our assumption. By minimizing this disparity, we effectively reduce uncertainty, making KL divergence a valuable tool for model evaluation, hypothesis testing, and various uncertainty informed decision-making processes across diverse academic domains [39],[40].

Renyi and Tsallis Entropy: Bridging the Gaps in Uncertainty Quantification
While Shannon entropy serves as a foundational tool for quantifying uncertainty, its limitations [20], particularly its dependence on chosen parameters and lack of an invariant measure, necessitate exploring alternative measures. Renyi and Tsallis entropy [41],[9],[10] provide valuable solutions, offering greater flexibility and adaptability in diverse academic disciplines.
Renyi Entropy: Unveiling Hidden Aspects of Uncertainty:
Proposed by Alfred Renyi [42], Renyi entropy ($H_q(X)$) presents a generalization of Shannon entropy by introducing the parameter 'q'. This parameter grants flexibility, enabling the capture of different facets of uncertainty depending on its value [1]. As 'q' approaches 1, Renyi entropy converges to Shannon entropy, maintaining compatibility with existing applications [43].

Beyond Traditional Applications:
Renyi entropy transcends Shannon's limitations, finding applications in diverse fields like information theory, statistical mechanics, and complex systems analysis [44]. Its strength lies in capturing power-law distributions and heavy-tailed behaviour, effectively handling non-uniform information content where Shannon entropy struggles [45]. The tenable parameter 'q' empowers researchers to focus on specific aspects of uncertainty, tailoring the measure to their data and research questions.

Tsallis Entropy: Embracing Correlations and Complexity:
Introduced by Constantino Tsallis in 1988 [46], Tsallis entropy ($H_q(X)$) incorporates an additional parameter 'q' to account for potential correlations between system elements. This makes it particularly suitable for systems with long-range correlations or non-extensive interactions, where conventional entropy measures fall short [47]. Tsallis entropy generates power-law distributions, offering valuable insights into complex systems exhibiting self-similar or fractal behaviour [48].

Academic Nuances and Rigor:
Both Renyi and Tsallis entropy necessitate careful parameter selection. Inappropriate choices can lead to misinterpretations and inconsistencies in uncertainty quantification [49]. A thorough understanding of their theoretical underpinnings and limitations is crucial before application in specific academic contexts. Comparative analysis with Shannon entropy is essential, along with clear justifications for chosen parameters, to enrich research methodologies and deepen the understanding of uncertainty within specific disciplines [50].
Renyi and Tsallis entropy represent crucial advancements in quantifying uncertainty, offering researchers wider applicability and adaptability beyond the conventional limitations of Shannon entropy [51]. By delving into their theoretical foundations, understanding their applications, and acknowledging their limitations, researchers can leverage these powerful tools to gain deeper insights into complex systems and phenomena characterized by non-uniform information content, intricate correlations, and multifaceted interactions [48],[50].

Maasoumi-Racine Entropy: A Rigorous Exploration Beyond Shannon Entropy
Within the realm of information theory, Shannon entropy reigns supreme as the foundational measure of uncertainty associated with random variables [52]. However, its scope is limited to individual variables, neglecting the crucial realm of dependence between them [20]. To address this shortcoming, Maasoumi-Racine (MR) entropy [25] emerges as a more versatile and academically rigorous extension, offering a deeper understanding of information content in complex systems.

Fundamental Principles:
Joint Uncertainty: Unlike Shannon's focus on individual probabilities, MR entropy explicitly incorporates pairwise and higher-order dependencies, enabling the analysis of intricate interactions within datasets. This makes it immensely valuable for tackling problems where variables exhibit non-trivial relationships
[53].
Generalized Dependence Structures: MR entropy transcends limitations of traditional measures by accommodating a wider spectrum of dependence structures [54]. It seamlessly handles linear, nonlinear, and even nonparametric dependencies [55], providing a more general framework for diverse data and applications.

Enhanced Information Capture: By delving into the realm of dependence, MR entropy offers richer information compared to Shannon entropy [22]. It unveils crucial insights into how variables intertwine, empowering researchers with a deeper understanding of system dynamics and facilitating improved modeling and prediction capabilities [56].

Academic Insights and Rigor:
Renowned scholars commend MR entropy for its flexibility [52], generality, and information richness, highlighting its advantages in various academic disciplines [56]. Theoretical, MR entropy is firmly grounded in information theory principles, drawing upon well-established axioms and mathematical frameworks [22]. This ensures theoretical soundness and rigor. Extensive research showcases the practical utility of MR entropy in diverse academic fields, including finance, econometrics, machine learning, signal processing, image analysis, bioinformatics, and social network analysis [56],[22]. While offering enhanced information, MR entropy can involve higher computational costs compared to Shannon entropy. However, ongoing research explores efficient algorithms and approximations to mitigate this challenge [57]. Mathematically sound, MR entropy measures can sometimes lack intuitive interpretations [22]. Recent efforts focus on developing more interpretable formulations that bridge the gap between theory and practical application.
Moreover, MR entropy stands as a powerful extension to Shannon entropy, offering a rigorous and flexible framework for information analysis in complex systems [52]. Its ability to capture dependence structures and provide richer information content makes it a valuable tool for academics across various disciplines [52]. By addressing computational and interpretability considerations, MR entropy is poised to play an increasingly critical role in advancing our understanding of information and dependence in diverse academic pursuits [22].

Methodology

Stochastic Probability Density Function Theory

Introduction
Stochastic (Non-deterministic) probability density function theory (SPDF) encompasses a rich and diverse set of methods used to describe and analyse systems governed by random processes. Understanding these systems often requires going beyond traditional deterministic approaches and embracing the inherent randomness present. SPDF theory provides a powerful framework for doing just that.

- *Stochastic Processes:* Systems often evolve in a random way over time. SPDF theory models such systems using stochastic processes, which describe the evolution of the system's probability distribution over time. These processes can be discrete (e.g., coin flips) or continuous (e.g., Brownian motion).

- *Probability Densities:* SPDF relies heavily on the concept of probability density functions (PDFs). A PDF assigns probabilities to different possible values of a random variable, allowing us to quantify the likelihood of each outcome.
- *Equation-Based Analysis:* Instead of deterministic equations, SPDF theory utilizes equations that involve probabilities and random fluctuations. These equations, such as the Fokker-Planck equation, describe how the probability distribution of the system evolves over time.

Stochastic Process
Stochastic processes describe dynamical systems whose time-evolution is of probabilistic nature.

Definition 1. *Let* T *be an ordered set,* $(\Omega, \mathcal{F}, \mathrm{P})$ *a probability space and* $(\mathcal{E}, \mathcal{G})$ *a measurable space. A stochastic process is a collection of random variables* $\mathrm{X} = \{\mathrm{X}_{\mathrm{t}}; \mathrm{t} \in \mathrm{T}\}$ *where, for each fixed* $\mathrm{t} \in \mathrm{T}$ *,* X_{t} *is a random variable from* $(\Omega, \mathcal{F}, P)$ *to* $((\mathcal{E}, \mathcal{G})$. Ω *is known as the sample space, where* $\mathcal{E}$ *is the state space of the stochastic process* X_{t}

[58].

The set T can be either discrete, for example the set of positive integers Z^+, or continuous $T = R^+$. The state space $\mathcal{E}$ will usually be R^d equipped with the σ-algebra of Borel sets. A stochastic process X may be viewed as a function of both $t \in T$ and $\omega \in \Omega$. The notations are sometimes used, $X(t), X(t,\omega)$ or $X_t(\omega)$ instead of X_t. For a fixed sample point $\omega \in \Omega$, the function $X_t(\omega): T \longmapsto \mathcal{E}$ is called the path of the process X.
One of the most important continuous-time stochastic process is **Brownian motion**. Brownian motion is a process with almost surely continuous paths and independent Gaussian increments. A process X_t has independent increments if for every sequence $t_0 < t_1 < \cdots < t_n$ the random variables

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

are independent. If, furthermore, for $t_1, t_2, s \in T$ and Borel set $\mathcal{B} \subset R$.

$$P\left(X_{t_2+s} - X_{t_1+s} \in \mathcal{B}\right) = P\left(X_{t_2} - X_{t_1} \in \mathcal{B}\right)$$

[59], then the process X_t has stationary independent increments.

Definition 2. *A one-dimensional standard Brownian motion* $W(t): R^+ \longmapsto R$ *is a real valued stochastic process with almost surely continuous paths such that* $W(0) = 0$*, it has independent increments and for every* $t > s \geq 0$*, the increment* $W(t) - W(s)$ *has a Gaussian distribution with mean 0 and variance* $t - s$*, i.e. the density of the random variable* $W(t) - W(s)$ *is*

$$g(x; t, s) = \left(\pi(t-s)\right)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2(t-s)}\right) \quad [53]$$

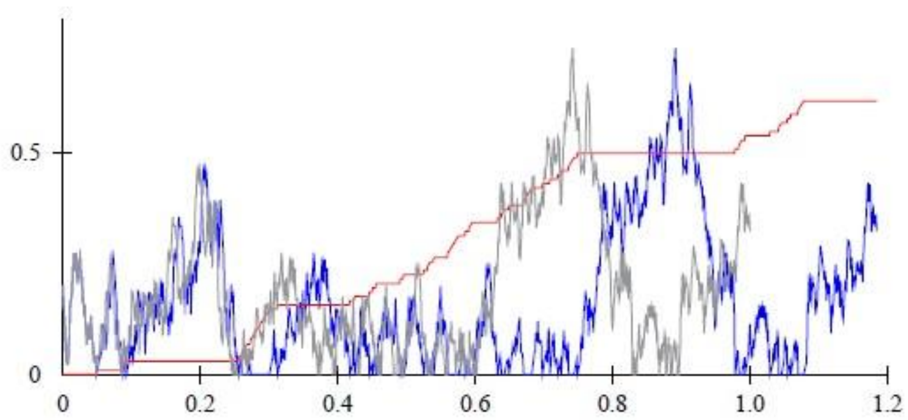


Figure 1: A path of a "sticky" Brownian motion (blue) constructed from the path of a reflecting Brownian motion (grey). The local time at 0 of the "sticky" paths is in red [60]

.

A standard d-dimensional standard Brownian motion $W(t): R^+ \longmapsto R^d$ is a vector of d independent one-dimensional Brownian motions:

$$W(t) = \left(W_1(t), \dots, W_d(t)\right),$$

where $W_i(t), i = 1, \dots, d$ are independent one-dimensional Brownian motions. The density of the Gaussian random vector $W(t) - W(s)$ is thus

$$g(x; t, s) = \left(\pi(t-s)\right)^{-\frac{d}{2}} \exp\left(-\frac{\left||x|\right|^2}{2(t-s)}\right)$$

Acknowledging that Brownian motion with almost surely continuous paths, also has a continuous modification. Consider two stochastic processes X_t, and $Y_t, t \in T$, that are defined on the same probability space $(\Omega, \mathcal{F}, P)$. The process Y_t, is said to be a modification of X_t, if $P(X_t = Y_t) = 1$ for all $t \in T$. The fact that there is a continuous modification of Brownian motion follows from the following result known as Kolmogorov theorem, see Figure 2.

Theorem 1. *Let* X_t *,* $t \in [0, \infty)$ *be a stochastic process on a probability space* $(\Omega, \mathcal{F}, P)$. *Suppose that there are positive constants* α *and* β, *and for each* $T \geq 0$ *there is a constant* $C(T)$ *such that*

$$E|X_t - Xs|^{\alpha} \leq C(T)|t-s|^{1+\beta}, 0 \leq s, t \leq T. \quad (5)$$

Then there exists a continuous modification and Y_t *of the process* of X_t,

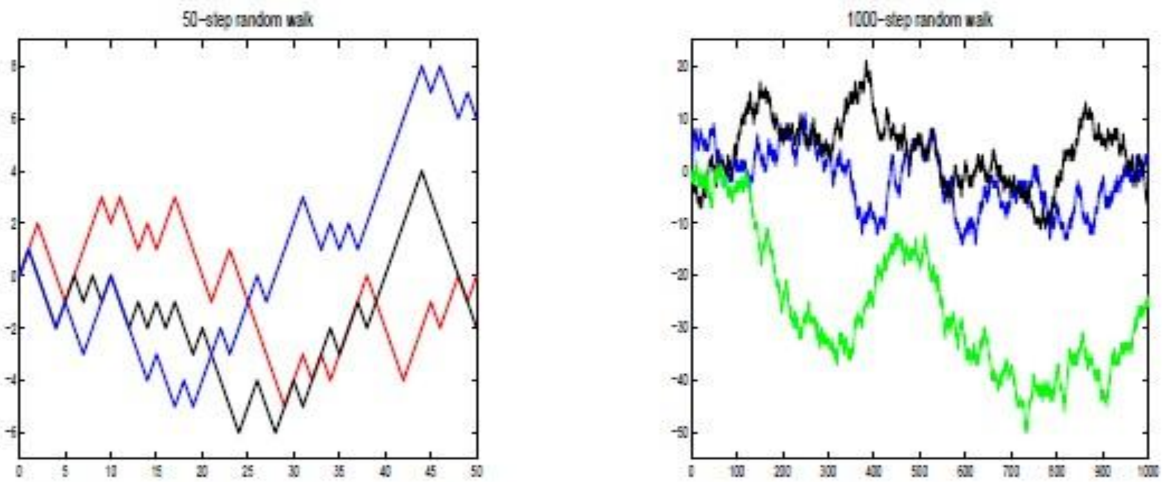
Brownian motion is also referred to as the *Wiener process*. It is possible to prove the existence of the Wiener process (Brownian motion) as shown in the theorem below:

Theorem 2. *There exists an almost surely continuous process* W_t *with independent increments such and* $W_t = 0$, *such that for each* $t \geq 0$ *the random variable* W_t *is* $\mathcal{N}(0, t)$. *Furthermore,* W_t *is almost surely locally Hölder continuous with exponent* α *for any* $\alpha \in \left(0, \frac{1}{2}\right)$.

Proof: let X_1, X_2 , ... be iid random variables on a probability space $(\Omega, \mathcal{F}, P)$. with mean 0 and variance 1. Define the discrete time stochastic process S_n with $S_0 = 0$, $S_n = \sum_{j=1} X_j$, $n \geq 1$. Define now a continuous time stochastic process with continuous paths as the linearly interpolated, appropriately rescaled random walk:

$$W_t^n = \frac{1}{\sqrt{n}} S_{[nt]} + (nt + [nt]) \frac{1}{\sqrt{n}} X_{[nt]+1} \quad (6)$$

where $[\cdot]$ denotes the integer part of a number. Then W_t^n converges weakly, as $n \longmapsto +\infty$ to a one-dimensional standard Brownian motion, see Figure 2.



a. $n = 50$ b. $n = 1000$

Figure 2: Sample paths of the random walk of length *n* = 50 and *n* = 1000.

Furthermore, the definition of the one-dimensional standard Brownian motion is that of a Gaussian stochastic process on a probability space $(\Omega, \mathcal{F}, P)$ with continuous paths for almost all $\omega \in \Omega$, and finite dimensional distributions with zero mean and covariance $\mathrm{E}\left(W_{t_i}, W_{t_j}\right) = \min(t_i, t_j)$.
For the d-dimensional Brownian motion we have [60],

$$\mathrm{E}(W_t) = 0, \forall t \geq 0 \qquad (7)$$

and

$$\mathrm{E}\big((W_t - W_s) \otimes (W_t - W_s)\big) = (t - s)I, \qquad (8)$$

Where I represent the identity matrix. And hence

$$\mathrm{E}(W_t \otimes W_s) = \min(t, s)I \qquad (9)$$

The probability density of the one-dimensional Brownian motion is

$$f(x,t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \qquad (10)$$

We can easily calculate all moments:

$$\mathrm{E}(W_t^n) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} e^{-x^2/2t}\, dt \qquad (11)$$

$$= \begin{cases} 1.3 \ldots (n-1)t\text{^\{n/2\}},\ n\ \text{"\{even\}} \\ 0,\ n\ \text{"\{odd\}} \end{cases}, \ [59]$$

One can see that the mean square displacement of Brownian motion grows linearly in time and Brownian motion is invariant under various transformations in time.

Proposition 1. *Let* W_t *denote a standard Brownian motion in* $\mathbb{R}$. *Then,* W_t *has the following properties:*

(1) *(Rescaling). For each* $k > 0$ *define* $X_t = \frac{1}{\sqrt{k}} W_{ct}$. *Then* $(X\_t\$, \$t \geq 0) = (W\_t\$, \$t \geq 0)$ *in law.*
(2) *(Shifting). For each* $c > 0, W_{c+t} - W_c, t \geq 0$ *is a Brownian motion which is independent of* $W_a, a \in [0, a]$.
(3) *(Time reversal). Define* $X_t = W_{1-t} - W_1, t \in [0,1]$. *Then* $(X_t, t \in [0,1]) = (W_t, t \in [0,1])$ *in law.*
(4) *(Inversion). Let* $X_t, t \geq 0$ *defined by* $X_0 = 0, X_t = tW_{(1/t)}$. *Then* $(X_t, t \geq 0) = (W_t, t \geq 0)$ *in law.*

Proof

(1) Consider t be a specific time point, where it can be rewrite as $t = cs$ for some $s \geq 0$. Using change of variable formula, we have

$$X_t = \frac{1}{\sqrt{k}} W_{ct} = \frac{1}{\sqrt{k}} W_{cs(1)} = \frac{1}{\sqrt{k}} \sqrt{c} W_s.$$

√_

Since W_s is normal distributed with mean 0 and variance s, then $W_s/\sqrt{c}$ is also normal distributed with mean 0 and variance s/c. If we multiply through with $\frac{1}{\sqrt{K}}$, we get $s/ks = t/k$. So $X_t = \frac{1}{\sqrt{k}} W_{ct}$ is normal distributed with mean 0 and variance t/k. Since X_t matches the distribution of W_t for any $t \geq 0$, then we conclude that $(X_t = W_t$. for any $t \geq 0)$ in law.

(2) Let $S = \inf\{t \geq 0 :\ W_t = c\}$. S is stopping time, and $W_c = W_S$. Therefore, the sigma-algebra generator by $\{W_a : 0 \leq a \leq S\} = \{W_a : 0 \leq a \leq S\}$ is independent of the sigma-algebra generated by $\{W_{c+t} : t \geq 0\} = \{\{W_s : S \leq s \leq S + t\}$. Since $W_{c+t} - W_c = (W_{c+t} - W_S) + (W_S - W_c)$, we can see that:
- $W_{c+t} - W_S$ is independent of $\{W_a : 0 \leq a \leq S\}$ due to the strong Markov Property.
- $W_S - W_S$ is independent of $\{W_a : 0 \leq a \leq S\}$ because W_c is a function of $\{W_a : 0 \leq a \leq S\}$.

•

Therefore $W_{c+t} - W_S$ is a sum of independent random variables and hence independent of $\{W_a: 0 \leq a \leq S\}$. By stationary of increments, it is also independent of any W_a for $a \in [0, c]$. Since $W_{c+t} - W_c$ has same distribution as $W_t - W_0$, for any $c \geq 0$ and, also independent of any W_a for $a \in [0, c]$, then $W_{c+t} - W_c$ is a Brownian motion independent of W_a for $a \in [0, c]$.

(3) (Time reversal). Having $X_t = W_{1-t} - W_1$, apply the reflection principal with $a = 1$ and $t = 1 - t$ to get:

$$X_t = W_{(1-t)-1} - W_1 = W_{t/2} - W_1$$

This means X_t is the reflection of a Brownian motion at 0 across the line $t = \frac{1}{2}$. To show that the distribution of X_t matches the distribution of W_t for $t \in [0,1]$ we take note of the following:

- Both W_t and the reflection of a Brownian motion start at 0 (since $W_0 = 0$).
- Both have same variance for any given *t*.
- The reflection simply changes the direction of the movement but preserves the magnitude.

Since the distribution X_t matches the distribution of W_t for any $t \in [0,1]$, we conclude that $X_t = W_t$ for any $t \in [0,1]$ in law.

(4) Define a transformation $s = \frac{1}{t}$. Then $t = \frac{1}{s}$ and $dX_t = \frac{ds}{s^2}$. Using transformation in the Itô formula on X_t,

$$dX_t = t dW_{1/t} + \frac{1}{2} t^2 d\left(W_{1/t}\right)_t$$

substituting ds and simplifying

$$dX_t = -dW_s + \frac{1}{2s} t^2 ds$$

The standard Brownian motion W_t satisfies the stochastic differential equation (SDE)

$$dW_t = dB_t$$

where B_t is a standard Brownian motion with variance *t*. Comparing the two (SDE):

- both have the same drift term (0).
- the diffusion term for X_t volves $\frac{1}{s}$ instead of s compared to W_t .

Although the diffusion terms differ, they are related by a simple transformation:

- for $t > 1, \frac{1}{s} < s$
- for $0 < t < 1, \frac{1}{s} < s$

Therefore, the diffusion term of X_t scales the diffusion term of W_t in a predictable way depending on the time range. This scaling does not affect the distribution of the process if the scaling factor is non-random and deterministic. Since both X_t and W_t satisfy SDEs with the same drift term and equivalent diffusion terms they have the same distribution. Therefore, ($X_t = W_t$ for any $t \geq 0$) in law.

One can also add a drift and change the diffusion coefficient of the Brownian motion: Let's define a Brownian motion with drift μ and variance σ^2 as the process

$$X_t = \mu t + \sigma W_t. \tag{12}$$

The mean and variance of X_t are

$$E(X_t) = \mu t, E(X_t) - E(X_t)^2 = \sigma^2 t. \tag{13}$$

Equation (12) above satisfies the equation

$$dX_t = \mu dt + \sigma dW_t. \qquad (14)$$

Which is a stochastic differential equation. Now in the probability density equation (10), we insert the drift μ and the variance σ^2 as controller of a stochastic behaviours. Hence the stochastic probability density function or a Gaussian Random Field distribution can be express as in [61]

$$f(x,t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{[\log x_t - \mu]^2}{2\sigma^2}\right). \qquad (15)$$

Quantifying uncertainty

Here we lay the groundwork for a robust discussion on quantifying uncertainty within diverse mathematical frameworks. Let's delve deeper into the mathematical aspects:

- Let U denote the set of all uncertainty measures offered by various theories (e.g., probability theory, possibility theory, Dempster-Shafer theory [62]).
- Each theory operates on a space of evidence representations, denoted by C. This space typically comprises subsets of the universal set of possibilities X.
- The crucial aspect is to assign a function u as a representation in the theory μ, defined as follows:

$$u: U(\mu) \longmapsto R^+, \qquad (16)$$

where

$$\mu: C \longmapsto R \qquad (17)$$

associating a non-negative real value (uncertainty magnitude) with every uncertainty measure $\mu \in U$ [63]. This function translates theoretical uncertainty representations into quantifiable values.

The mapping function u must adhere to specific axioms ensuring consistency and meaningful interpretation of assigned uncertainty values. Some essential axioms could include:

1. $u(\mu) \geq 0$ for all $\mu \in U$.
2. $u(\mu_0) = 0$ for a specific reference measure μ_0 representing perfect certainty (e.g., measure concentrated on a single outcome).
3. Higher uncertainty (less evidence) should have a higher numerical value: $\mu_1 \leq \mu_2 \rightarrow u(\mu_1) \geq u(\mu_2)$
4. Compatibility with fundamental operations on measures (e.g., union, intersection) might be desirable depending on the application.

The fundamental measure of uncertainty based on Hartley's measure quantifies [64] $H(\mu) = -\log_2(|A|)$, where A represents the set of possibilities consistent with the evidence μ. The higher cardinality $(|A|)$ implies more ambiguity and results in a higher uncertainty value. This measure solely considers the number of possibilities, ignoring potential information within μ about their relative likelihoods. In the context of possibility theory, the finite set X of conceived alternatives includes only one alternative in each situation that is true. This uncertainty measure can be interpreted as a measure of diagnostic uncertainty. The level of uncertainty increases as the number of alternatives increases [63]. We now investigate the Shannon entropy.

The Shannon Entropy with stochastic probability

The standard Shannon entropy $H(x)$ defined in equation (2) uses a uniform probability distribution, assuming all possible outcomes are equally likely. This works well for simple scenarios with no inherent bias, but it might not reflect the true uncertainty in many real-world situations. This study considers that stochastic or non-deterministic probability distributions, can incorporate specific information about the variability and skewness present in the data.

Let $M(x)$ be the newly modified Shannon entropy measure, with all the attributes of uncertainty quantifier defined by equation (16), so

$$M(X) = -\int_R f(x,t)\log_2 f(x,t)dx \tag{18}$$

in case where X is a continuous random variable, or

$$M(X) = -\sum_{i=1}^{n} f(x,t)\log_2 f(x,t) \tag{19}$$

when X is a discrete random variable and $f(x,t)$ is the stochastic probability density function given in equation (15).

$M(x)$ leads to:

- *More accurate entropy values*: Reflecting the actual uncertainty characteristics of your system.
- *Deeper insights*: Revealing how uncertainty changes over time, depends on external factors, or exhibits specific patterns.
- *Tailored analysis*: Matching the distribution to your specific problem domain and research questions.

Theorem 3. *For any valid PDF* $f(x,t) \geq 0,\ \ H(f) = \int_R f(x,t)\log_2 f(x,t)dx \geq 0$

Proof
By *PDF* definition, $f(x,t) > 0$ for all $x \geq 0$. Which results to $f(x,t)\cdot[-\log \_2(f(x,t)] > 0$. Subsequently, it follows that $H(f) > 0$. The product $f(x,t)\log_2 f(x,t) = 0$ if and only if $f(x,t) = 1$. Therefore $H(f) = \int_R f(x,t)\log_2 f(x,t)dx = 0$ if $f(x,t) = 1$ everywhere. Putting all together, we have $H(f) \geq 0$.

Theorem 4. *(Sub-additivity Property)*
Let $f(x,t)$ *be a proper PDF i.e.,* $\int_R f(x,t)dt = 1$ *suppose* $H(f) \leq 0$ *is finite and* $f(x,t) + g(x,t) = 1$ *where f and g are proper PDF. Then* $H(f+g) \leq H(f) + H(g)$.

Proof
We can utilize Jensen's inequality, which states that, for a convex function $u(x)$ with the probability distribution $p(x,t)$, we have $\int_R u(E[x])p(x,t)dx \leq E[u(x)]$. Consider the convex function $u(x) = x\log_2 x$ for all $x \geq 0$ and $p(x,t) = f(x,t) + g(x,t)$. Then,

$$\int_R \big(f(x,t) + g(x,t)\log_2\big(f(x,t) + g(x,t)\big)\big)dx \leq \big(f(x,t) + g(x,t)\big) + g(x,t)E(1)dx$$

note that, $E[1] = \int_R f(x,t) + g(x,t)dx = 1$ due to the given condition. Rearranging and using the definition $H(f)$ and $H(g)$, then $H(f+g) \leq f(x,t) + g(x,t) = H(f) + H(g)$. Therefore, $H(f+g) \leq H(f) + H(g)$ holds.

Theorem 5. *(Maximum Entropy Principle)*
Let $E[h(X)]$ *be the expected value of a function* $h(X)$ *with respect to the random variable* X *and let* C *be a set of constraints on these expected values. Then, the PDF* $f(x,t)$ *that maximizes* $H(f)$ *subject to constraints in C is the solution to:*

$$ArgMax_f\ \{\ H(f)\ \mid E[h(X)]\ \in C\} \tag{20}$$

Theorem 6. (*Data Processing Inequality)*
Let $Y = g(X)$ *be a deterministic function of* X *with PDF* $f(x,t)$. *Then:*

$$H(Y) \leq H(X) \tag{21}$$

Theorem 7. (*Chain Rule)*
If X *and* Y *are independent random variables with PDFs* $f(x,t)$ *and* $g(y,t)$*:*

$$H(X,Y) = H(X) + H(Y) \tag{22}$$

Proof
Let X and Y be two random variables with joint PDF $f(x,t)$ defined as

$$\mathrm{H(X,Y)} = -\int_{\mathrm{R}}\int_{\mathrm{R}} \mathrm{f(x,y)} \log_2 f(x,y) d\,\mathrm{xdy}.$$

Since X and Y are independent, their PDF can be express as $\mathrm{f(x,y)} = \mathrm{f(x,y)} \cdot \mathrm{g(y,t)}$ where $f(x,t)$ and $g(y,t)$are marginal PDF of X and Y respectively. Using product rule into joint entropy, we have

$$\mathrm{H(X,Y)} = -\int_{\mathrm{R}}\int_{\mathrm{R}} \big(\mathrm{f(x,t)} \cdot \mathrm{g(y,t)}\big) \log_2 \big(\mathrm{f(x,t)} \cdot \mathrm{g(y,t)}\big)\, \mathrm{dxdy} \tag{23}$$

Separating equation (23) using the logarithmic properties, we get

$$\mathrm{H(X,Y)} = -\int_{\mathrm{R}}\int_{\mathrm{R}} [\mathrm{f(x,t)} \cdot \log_2 \mathrm{f\,(x,t)}] + \mathrm{f(x,t)} \cdot \log_2[\mathrm{g(y,t)} + \mathrm{g(y,t)} \cdot \log_2 \mathrm{f\,(x,t)}]\, \mathrm{dxdy}$$

Therefore, grouping the terms based on the functions $f(x,t)$ and $g(y,t)$

$$\mathrm{H(X,Y)} = -\left(\left(\int_{\mathrm{R}} \mathrm{f(x,t)} \log_2 f(x,t)\, \mathrm{dx}\right) + \left(-\int_{\mathrm{R}} \mathrm{g(y,t)} \log_2 g(y,t)\, \mathrm{dy}\right) = \mathrm{H(X)} + \mathrm{H(Y)}\right.$$

Hence proven.

Results and Discussion

Introduction
This section details the application of the proposed entropy measure to analyse data from various financial instruments, including stocks, futures contracts, and exchange rates. Data was retrieved from a publicly available financial database [65] and subsequently analysed using Python as the primary programming language. Specific details regarding the chosen instruments are provided in Table 1. The data encompasses a nine-year timeframe, spanning from December 2014 to December 2023. Adjusted closing prices were utilized, with an average of up to 2265 observations per instrument. This selection was based on data availability within the chosen database. To ensure consistency in the analysis, all instruments were examined over the identical time-period, specifically, the first month, first year, third year, fifth year, seventh year, and ninth year. Notably, the proposed quantifier is applicable to both discrete and continuous data series. Furthermore, the study compared the traditional Shannon entropy, which utilizes a normal probability distribution, with a modified version employing a stochastic probability distribution.

Periods		One Month	One Year	Three Years	Five years	Seven Years	Nine Years
Counts		21	252	756	1258	1763	2265
Mean (μ)	Bitcoins	251.8562	271.9979	1606.9022	3951.6407	11189.1130	15041.2365
	Facebook	76.7020	88.5732	120.5631	142.8207	181.3183	190.0512
	SP500	2029.6414	2061.1271	2201.0824	2451.7115	2821.1732	3125.0128
	Rusell2000	1181.6657	1205.8931	1266.6228	1386.9670	1528.6443	1602.5650
	Tesla	13.6112	15.3316	16.7503	17.9175	63.6369	102.8474
	Amazon	15.1555	23.8344	35.7203	55.6913	82.8309	91.9168
Standard Deviation (σ^2)	Bitcoins	38.4433	59.0922	2875.2343	3970.3189	15898.3456	16257.2142
	Facebook	1.1113	10.1734	30.5622	38.0501	75.1616	76.3271
	SP500	22.4612	54.8753	197.9467	356.9490	732.7045	871.4692
	Rusell2000	13.5689	47.9205	133.3196	187.2190	345.9905	339.7825
	Tesla	0.5592	1.5845	3.6168	3.7331	89.3987	110.9273
	Amazon	0.7128	5.5269	11.2922	26.6005	50.3666	48.6703

Table 1: Financial derivatives and their descriptions per period

Testing for novel Shannon entropy

Prior to analysis, all data pertaining to the financial instruments outlined in Table 1 underwent a rigorous cleaning process. This process ensured the removal of any inconsistencies or errors within the data. Subsequently, histograms and normal probability density functions (PDFs) were generated to visualize the distribution of the cleaned data. These visualizations are presented in Figures 5 and 6, respectively. Figure 6 depicts the histograms of the organized data. By visually inspecting these histograms, a tendency towards normality can be observed in the data distribution. Conversely, Figure 6 employs kernel density estimation (KDE) to capture the inherent randomness within the data. This technique provides a smoother representation of the underlying distribution compared to histograms.

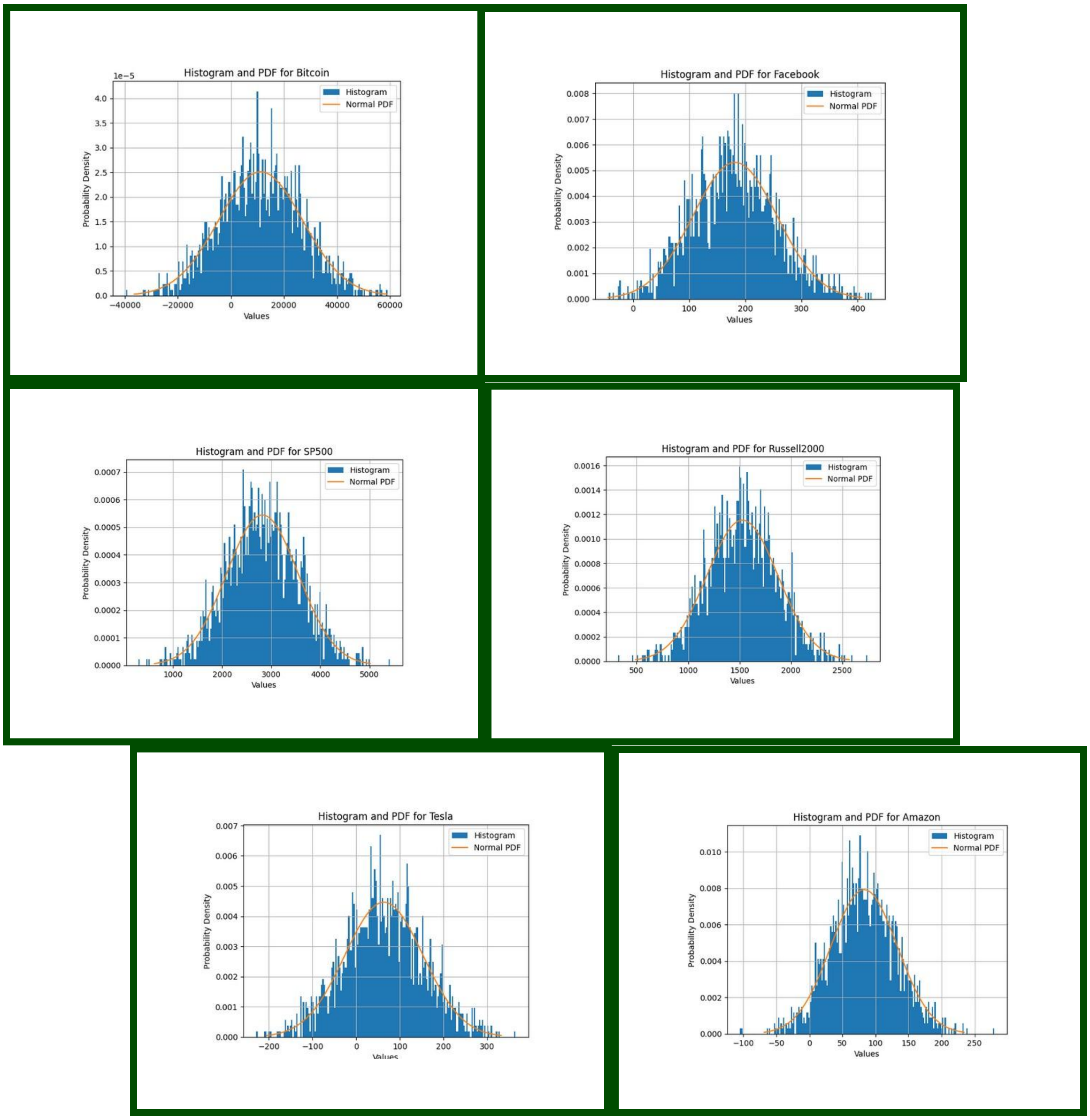


Figure 5: Histograms and normal PFD of financial derivatives

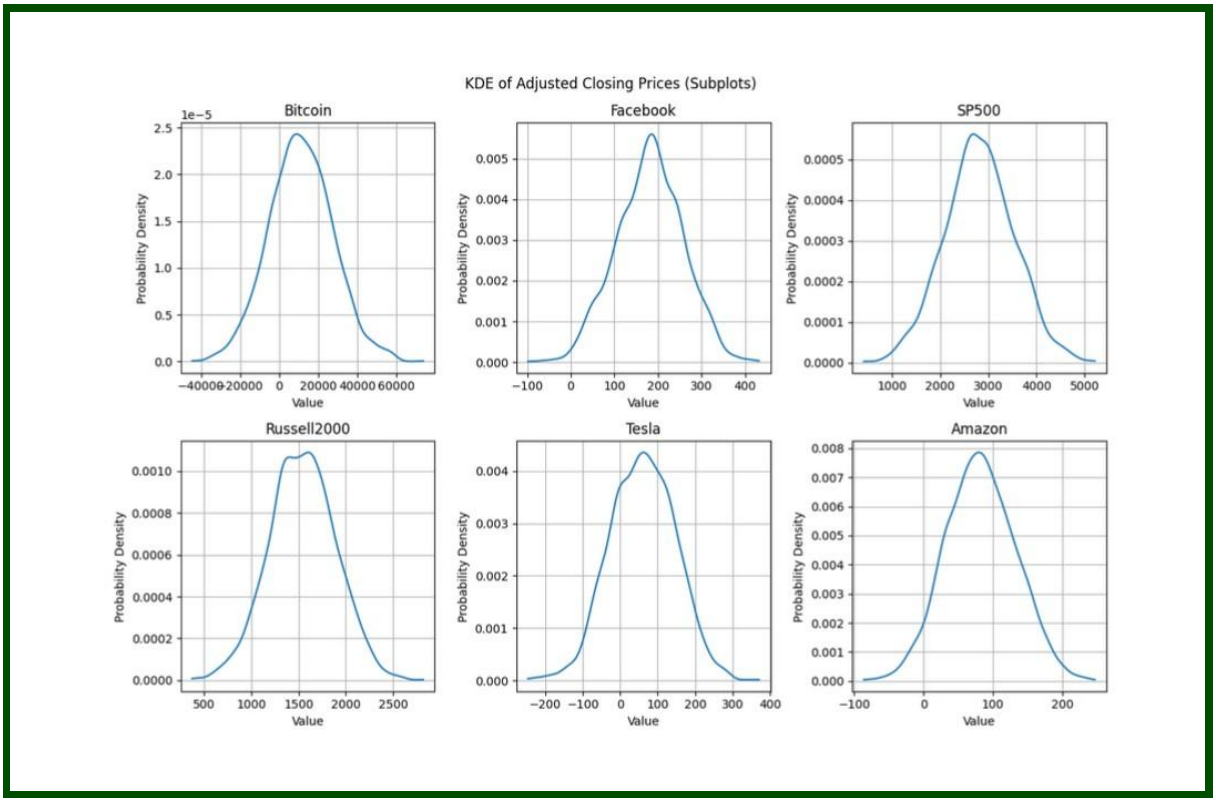


Figure 6: PDF of the Random Distribution using KDE function

Results

This section details the evaluation of information entropy for the adjusted closing prices of the aforementioned financial instruments. The proposed entropy measure, as presented in Equation ([19]), was employed for this analysis. Additionally, traditional Shannon entropy, based on Equation ([1]), was calculated for comparative purposes. The results of this evaluation are summarized in Table 2.

Periods		One Month	One Year	Three Years	Five Years	Seven Years	Nine Years
Shannon Entropy	Bitcoins	4.3923	7.9773	9.5596	10.2937	10.7815	11.1435
	Facebook	4.2018	7.8820	9.4723	10.1838	10.6940	11.0459
	SP500	4.3923	7.9535	9.5411	10.2794	10.7713	11.1347
	Russell200	4.3923	7.9455	9.5384	10.2747	10.7634	11.1250
	Tesla	4.3923	7.9187	9.5030	10.2343	10.7380	11.1052
	Amazon	4.3923	7.9614	9.5490	10.2762	10.7657	11.1114
Proposed Entropy	Bitcoins	4.16E-07	0.005698	1.549536	0.858404	0.306457	0.254947
	Facebook	0	1.9E-12	0.16549	0.23663	3.918133	3.284718
	SP500	0	0	0	2.46E-09	0.007357	0.015511
	Russell200	0	0	0	1.14E-10	0.001731	0.000492
	Tesla	0	1.7E-10	0.916559	0.479705	28.94218	21.00299
	Amazon	0	0.91687	6.492424	19.81093	20.46706	15.04915

Table 2: Entropies outcome results

Discussions

The entropy risk measure results presented in Table 2 suggest a potential advantage over traditional Shannon entropy for certain assets. While traditional entropy might indicate relatively high risk across all considered instruments (Equation [1]), the proposed entropy measure offers a potentially different perspective (Equation [19]).

Table 3 provides further insights. It suggests that all assets might offer a relatively safe environment for short-term investments, with Bitcoin appearing slightly riskier than the remaining derivatives. Conversely, SP500 and Russell2000 emerge as the most favourable options within this timeframe. For long-term investments, however, the study indicates that Tesla, Amazon, and Facebook might be less suitable based on the proposed entropy measure

Periods	One Month	One Year	Three Years	Five Years	Seven Years	Nine Years
Shannon Entropy	Bitcoins	Bitcoins	Bitcoins	Bitcoins	Bitcoins	Bitcoins
	Amazon	Amazon	Amazon	Amazon	SP500	SP500
	SP500	SP500	SP500	SP500	Amazon	Amazon
	Russell2000	Russell200	Russell200	Russell200	Russell200	Russell200
	Tesla	Tesla	Tesla	Tesla	Tesla	Tesla
	Facebook	Facebook	Facebook	Facebook	Facebook	Facebook
Proposed Entropy	Bitcoins	Amazon	Bitcoins	Amazon	Tesla	Tesla
	Amazon	Bitcoins	Amazon	Bitcoins	Amazon	Amazon
	Tesla	Tesla	Facebook	Tesla	Facebook	Facebook
	Facebook	Facebook	Tesla	Facebook	Bitcoins	Bitcoins
	Russell200	Russell200	Russell200	SP500	SP500	SP500
	SP500	**SP500**	**SP500**	**Russell200**	**Russell200**	**Russell200**

Table 3: Ranking according to entropy information

Conclusion

This article proposed a novel modified Shannon entropy that incorporates the stochastic behaviour of market derivatives to quantify their associated uncertainty and risk. It demonstrates that traditional Shannon entropy theory can be enhanced by considering the inherent randomness of data, particularly for highly volatile assets, as evidenced by the results. This work contributes to the field of risk analysis by employing Shannon entropy to quantify the disparity arising from stochasticity within the distribution.

The article also identified several key areas for future research. Further investigation is needed to address potential inconsistencies identified in the formulation. Despite these limitations, the results are promising for various applications. Future work should explore the applicability of incorporating the stochastic nature of data into other entropy families, such as Kullback-Liebler divergence, mutual information, R´enyi entropy, Tsallis entropy, and others. This novel approach has the potential to significantly benefit future research by highlighting the importance of accounting for randomness in data. It paves the way for addressing the fundamental question of" what is the optimal distribution of data?" Ultimately, the proposed novel Shannon entropy perspective offers an alternative way to quantify uncertainty in financial derivatives.

References

[1] A. Namdari, Z. Li, A review of entropy measures for uncertainty quantification of stochastic processes, Advances in Mechanical Engineering 11 (6) (2019) 1687814019857350.
[2] A. Namdari, Z. S. Li, Integrating fundamental and technical analysis of stock market through multi-layer perceptron, in: 2018 IEEE technology and engineering management conference (TEMSCON), IEEE, 2018, pp. 1–6.
[3] A. Barchielli, M. Gregoratti, A. Toigo, Measurement uncertainty relations for position and momentum: relative entropy formulation, Entropy 19 (7) (2017) 301.
[4] P. Glasserman, X. Xu, Robust risk measurement and model risk, Quantitative Finance 14 (1) (2014) 29–58.
[5] C. Neri, L. Schneider, Maximum entropy distributions inferred from option portfolios on an asset, Finance and Stochastics 16 (2012) 293–318.
[6] G. N. Saridis, Entropy formulation of optimal and adaptive control, IEEE Transactions on Automatic Control 33 (8) (1988) 713–721.
[7] V. BRATIAN, Evaluation of options using the monte carlo method and the entropy of information, Expert Journal of Economics 6 (2) (2018).
[8] C. E. Shannon, A mathematical theory of communication, The Bell system technical journal 27 (3) (1948) 379–423.
[9] J. Lin, Divergence measures based on the shannon entropy, IEEE Transactions on Information theory 37 (1) (1991) 145–151.
[10] J. D. Ramshaw, H-theorems for the tsallis and renyi entropies, Physics Letters A 175 (3-4) (1993) 169–170.
[11] S. Mishra, B. M. Ayyub, Shannon entropy for quantifying uncertainty and risk in economic disparity, Risk Analysis 39 (10) (2019) 2160–2181.
[12] N. Misra, H. Singh, E. Demchuk, Estimation of the entropy of a multivariate normal distribution, Journal of multivariate analysis 92 (2) (2005) 324–342.
[13] K. Conrad, Probability distributions and maximum entropy, Entropy 6 (452) (2004) 10.
[14] H. Joe, Estimation of entropy and other functionals of a multivariate density, Annals of the Institute of Statistical Mathematics 41 (1989) 683–697.
[15] G. Gour, M. Tomamichel, Entropy and relative entropy from information theoretic principles, IEEE Transactions on Information Theory 67 (10) (2021) 6313–6327.
[16] M. Sheraz, S. Dedu, V. Preda, Entropy measures for assessing volatile markets, Procedia Economics and Finance 22 (2015) 655–662.

[17] A. E. Abbas, Entropy methods for adaptive utility elicitation, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 34 (2) (2004) 169–178.
[18] T. Seidenfeld, Entropy and uncertainty, Philosophy of Science 53 (4) (1986) 467– 491.
[19] C. Marsh, Introduction to continuous entropy, Department of Computer Science, Princeton University 1034 (2013).
[20] R. Rajaram, B. Castellani, A. Wilson, et al., Advancing shannon entropy for measuring diversity in systems, Complexity 2017 (2017).
[21] A. Ullah, Uses of entropy and divergence measures for evaluating econometric approximations and inference, Journal of Econometrics 107 (1-2) (2002) 313–326.
[22] E. Properzi, Genome characterization through a mathematical model of the genetic code: an analysis of the whole chromosome 1 of a. thaliana (2013).
[23] P. K. Narayan, S. Popp, Investigating business cycle asymmetry for the g7 countries: Evidence from over a century of data, International Review of Economics & Finance 18 (4) (2009) 583–591.
[24] W. Orzeszko, Measuring nonlinear serial dependencies using the mutual information coefficient (2010).
[25] A. C. Annegues, E. A. d. Figueiredo, W. P. S. d. F. Souza, Determinants of unfair inequality in brazil, 1995 and 2009, CEPAL Review (2015).
[26] A. Lesne, Shannon entropy: a rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics, Mathematical Structures in Computer Science 24 (3) (2014) e240311.
[27] J. Liang, X. Zhao, D. Li, F. Cao, C. Dang, Determining the number of clusters using information entropy for mixed data, Pattern Recognition 45 (6) (2012) 2251–2265.
[28] S. Kullback, Kullback-leibler divergence (1951).
[29] S.-H. Lin, Y.-M. Yeh, B. Chen, Leveraging kullback–leibler divergence measures and information-rich cues for speech summarization, IEEE transactions on audio, speech, and language processing 19 (4) (2010) 871–882.
[30] E. H. Aoki, A. Bagchi, P. Mandal, Y. Boers, A theoretical look at information driven sensor management criteria, in: 14th International Conference on Information Fusion, IEEE, 2011, pp. 1–8.
[31] A. Clim, R. D. Zota, G. TinicA, The kullback-leibler divergence used in machine˘ learning algorithms for health care applications and hypertension prediction: a literature review, Procedia Computer Science 141 (2018) 448–453.
[32] P. Yang, B. Chen, Robust kullback-leibler divergence and universal hypothesis testing for continuous distributions, IEEE Transactions on Information Theory 65 (4) (2018) 2360–2373.
[33] Z. Yang, H. Zhang, Z. Yuan, E. Oja, Kullback-leibler divergence for nonnegative matrix factorization, in: International Conference on Artificial Neural Networks, Springer, 2011, pp. 250–257.
[34] S. Ji, Z. Zhang, S. Ying, L. Wang, X. Zhao, Y. Gao, Kullback–leibler divergence metric learning, IEEE transactions on cybernetics 52 (4) (2020) 2047–2058.
[35] M. Ka´rny`, J. Andry`sek, Use of kullback–leibler divergence for forgetting, International Journal of Adaptive Control and Signal Processing 23 (10) (2009) 961–975.
[36] Y. Xiang, L. Shi, J. L. Højvang, M. H. Rasmussen, M. G. Christensen, A speech enhancement algorithm based on a non-negative hidden markov model and kullback-leibler divergence, EURASIP Journal on Audio, Speech, and Music Processing 2022 (1) (2022) 22.
[37] S. Claici, M. Yurochkin, S. Ghosh, J. Solomon, Model fusion with kullback-leibler divergence, in: International conference on machine learning, PMLR, 2020, pp. 2038–2047.
[38] Z.-Y. Ran, B.-G. Hu, Determining parameter identifiability from the optimization theory framework: A kullback–leibler divergence approach, Neurocomputing 142 (2014) 307–317.
[39] M. Ponti, J. Kittler, M. Riva, T. de Campos, C. Zor, A decision cognizant kullback–leibler divergence, Pattern Recognition 61 (2017) 470–478.
[40] S. Filippi, O. Capp´e, A. Garivier, Optimism in reinforcement learning and kullback-leibler divergence, in: 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2010, pp. 115–122.
[41] S.-B. LIU, C.-Y. RONG, Z.-M. WU, T. LU, R´enyi entropy, tsallis entropy and onicescu information energy in density functional reactivity theory, Acta PhysicoChimica Sinica 31 (11) (2015) 2057–2063.
[42] A. R´enyi, On measures of entropy and information, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, Vol. 4, University of California Press, 1961, pp. 547–562.
[43] A. G. Bashkirov, Renyi entropy as a statistical entropy for complex systems, Theoretical and Mathematical Physics 149 (2) (2006) 1559–1573.
[44] K. E. Hild II, D. Erdogmus, J. C. Principe, An analysis of entropy estimators for blind source separation, Signal processing 86 (1) (2006) 182–194.
[45] H. Pham, C. Shahabi, Y. Liu, Ebm: an entropy-based model to infer social strength from spatiotemporal data, in: Proceedings of the 2013 ACM SIGMOD international conference on management of data, 2013, pp. 265–276.
[46] C. Tsallis, The nonadditive entropy sq and its applications in physics and elsewhere: Some remarks, Entropy 13 (10) (2011) 1765–1804.
[47] C. Tsallis, Approach of complexity in nature: Entropic non-uniqueness, Axioms 5 (3) (2016) 20.
[48] A. Iliopoulos, Complex systems: Phenomenology, modelling, analysis, Int. J. Appl. Exp. Math 1 (2016) 105.
[49] K. P. Nelson, Reduced perplexity: A simplified perspective on assessing probabilistic forecasts, arXiv preprint arXiv:1603.08830 (2016).

[50] Karmeshu, Entropy measures, maximum entropy principle and emerging applications (2003).

[51] G. Balasis, R. V. Donner, S. M. Potirakis, J. Runge, C. Papadimitriou, I. A. Daglis, K. Eftaxias, J. Kurths, Statistical mechanics and information-theoretic perspectives on complexity in the earth system, Entropy 15 (11) (2013) 4844–4888.

[52] C. Adami, Introduction to artificial life, Springer Science & Business Media, 1998.

[53] T. F. Varley, Uncovering higher-order structures in complex systems with multivariate information theory, Ph.D. thesis, Indiana University (2023).

[54] G. H. Johnstone, M. U. Gonza´lez-Rivas, K. M. Taddei, R. Sutarto, G. A. Sawatzky, R. J. Green, M. Oudah, A. M. Hallas, Entropy engineering and tunable magnetic order in the spinel high-entropy oxide, Journal of the American Chemical Society 144 (45) (2022) 20590–20600.

[55] E. Avcı, Artifical neural network models in finance: Application of the artificial decision-making turkey, Ph.D. thesis, Marmara Universitesi (Turkey) (2006).

[56] M. G. Zoia, P. Biffi, F. Nicolussi, Value at risk and expected shortfall based on gram-charlier-like expansions, Journal of Banking & Finance 93 (2018) 92–104.

[57] S. Anyfantaki, E. Maasoumi, J. Ren, N. Topaloglou, Evidence of uniform inefficiency in market portfolios based on dominance tests, Journal of Business & Economic Statistics 40 (3) (2022) 937–949.

[58] G. Zitkovi´c, Introduction to stochastic processes-lecture notes, Department of˘ Mathematics, The University of Texas at Austin (2010).

[59] M. Annunziato, A. Borzi, Optimal control of probability density functions of stochastic processes, Mathematical Modelling and Analysis 15 (4) (2010) 393– 407.

[60] V. Kostrykin, J. Potthoff, R. Schrader, Brownian motions on metric graphs: Feller Brownian motions on intervals revisited, arXiv preprint arXiv:1008.3761 (2010).

[61] G. A. Pavliotis, Stochastic processes and applications, Springer, 2016.

[62] B. M. Ayyub, G. J. Klir, Uncertainty modeling and analysis in engineering and the sciences, CRC Press, 2006.

[63] A. Feutrill, M. Roughan, A review of shannon and differential entropy rate estimation, Entropy 23 (8) (2021) 1046.

[64] R. V. Hartley, Transmission of information 1, Bell System technical journal 7 (3) (1928) 535–563.

[65] Y. Finance, Yahoo finance, Retrieved from finance. yahoo. com: https://finance.yahoo. com/recent quotes (2020).