**Stochastic Models in Probability and Statistics**
Vol. 2, No. 1, 2025, pp. 101–117. Research Article

# Varextropy measure with application

**Santosh Kumar Chaudhary***

*Department of Statistics, Central University of Jharkhand, Cheri-Manatu, Ranchi, Jharkhand, 835222, India*

*Email(s): skchaudhary1994@gmail.com; santosh.chaudhary@cuj.ac.in*

**Abstract.** In statistical analysis, understanding and quantifying uncertainty is fundamental. Measures such as entropy, extropy, varentropy, and varextropy provide valuable insights into the characteristics of probability distributions. This paper focuses on the concept of varextropy and presents a novel characterization of the uniform distribution, showing that the varextropy of a random variable is zero if and only if the variable is uniformly distributed on the unit interval. Building on this property, we propose a new goodness-of-fit test for uniformity based on a nonparametric estimator of varextropy, denoted by $\widehat{\Delta}$, as introduced by Noughabi and Noughabi (2024). The test statistic is shown to be consistent, and its distribution under the null hypothesis is explored via Monte Carlo simulations. Critical values are tabulated for various sample sizes and tuning parameters, and the test's power is empirically evaluated against alternatives such as the Beta(1,2) distribution, demonstrating superior performance in detecting departures from uniformity. The proposed method is further applied to a real-world environmental dataset of vinyl chloride concentrations, where the transformed data, via the probability integral transform, are shown to conform to a uniform distribution. Overall, this study not only extends the theoretical understanding of varextropy but also introduces a practical and effective tool for uniformity testing in both simulated and real data contexts.

*Keywords*: Entropy, Extropy, Goodness-of-fit, Monte Carlo simulation, Nonparametric estimator, Order statistics, Uniformity test, Varextropy.

## 1 Introduction

Quantifying uncertainty in random variables is a central theme in probability theory, information theory, and statistical inference. Various measures have been developed to capture different

*Corresponding author

Received: 11 March 2025 / Accepted: 07 June 2025
DOI: 10.22067/smps.2025.92576.1044

 https://smps.um.ac.ir

aspects of uncertainty, variability, and information content associated with probability distributions. Among the most widely known and utilized is Shannon's entropy, introduced by Shannon (1948), which provides a foundational measure of the average uncertainty or information content in a random variable. For an absolutely continuous random variable, entropy reflects the expected value of the negative logarithm of its probability density function (pdf) and plays a critical role in areas such as data compression, communication theory, and statistical modeling. Complementing entropy is the concept of *extropy*, proposed by Lad et al. (2015) as a dual measure of uncertainty. While entropy captures average surprise or unpredictability, extropy is designed to assess the regularity and concentration in the distribution of a continuous random variable. Defined in terms of the squared density function, extropy offers a different perspective on information, with applications in decision theory and statistical diagnostics. To understand not just the mean behavior but also the variability of information content, the notion of *varentropy* is employed. Varentropy, introduced by Arıkan (2016), is the variance of the information content (i.e., the log-density). It quantifies the dispersion around the average uncertainty and is particularly useful in finite blocklength information theory, where variability in data coding and transmission must be accounted for. Varentropy has also gained attention in statistical contexts as a more sensitive alternative to classical measures such as kurtosis, especially when analyzing continuous distributions. Building on these concepts, the measure of *varextropy* has recently been introduced to extend the idea of extropy by incorporating variability. Analogous to varentropy, varextropy captures the variance of the density function itself, providing insights into the fluctuation of distribution concentration. This new measure broadens the information-theoretic toolkit for studying distributional properties and can offer useful characterizations of specific distributions, such as the uniform distribution. The interplay between these four measures—entropy, extropy, varentropy, and varextropy—opens up new avenues for theoretical exploration and practical applications, particularly in statistical testing, distribution characterization, and information processing. This paper focuses on the properties and applications of varextropy, particularly in the context of testing for uniformity.

Testing for uniformity is a fundamental problem in statistical analysis with wide-ranging applications across various fields, including quality control, cryptography, simulation, and goodness-of-fit testing. The uniform distribution often serves as a benchmark or null model in many statistical procedures. For example, in simulation studies, ensuring that random number generators produce values that are uniformly distributed is essential for the validity of results. Similarly, in goodness-of-fit testing, the uniform distribution is commonly used to assess whether observed data deviate significantly from a theoretical model. Moreover, many statistical transformations and procedures assume an underlying uniformity, especially in the context of probability integral transforms. As such, reliable tests for uniformity are crucial for validating assumptions, detecting structure in data, and supporting the development of robust statistical methodologies. This motivates the exploration of new approaches, such as those based on information-theoretic measures like varextropy, to enhance the sensitivity and applicability of uniformity tests.

Although varextropy is a relatively recent addition to the family of information-theoretic measures, it has begun to draw interest for its potential applications in characterizing probability distributions. Previous studies have explored the mathematical properties of varextropy and demonstrated its sensitivity to distributional shape and concentration. However, its use in formal hypothesis testing, particularly for assessing uniformity, remains limited in the literature.

Existing uniformity tests are primarily based on classical approaches such as the Kolmogorov–Smirnov test, Cramér-von Mises criterion, and entropy-based methods. In contrast, this work introduces a novel test procedure that leverages the variance of the squared density, *varextropy*, as a means to detect deviations from the uniform distribution. By establishing a new characterization of the uniform distribution through varextropy, we extend its utility beyond descriptive analysis and into inferential statistics. Our method differs from earlier work in that it provides a nonparametric, information-theoretic framework for uniformity testing, offering a potentially more sensitive alternative to traditional approaches. Furthermore, we evaluate the effectiveness of the proposed test through both theoretical derivations and empirical analyses using real-world data, thereby demonstrating its practical relevance. The entropy of a discrete probability distribution $P = \{p_1, \ldots, p_n\}$ is defined as (Shannon, 1948) $H(P) = -\sum_{i=1}^{n} p_i \ln p_i$. The varentropy of a discrete probability distribution $P = \{p_1, \ldots, p_n\}$ is defined as (see Arıkan (2016); De Crescenzo et al. (2025); Maadani et al. (2022))

$$VH(P) = \sum_{i=1}^{n} p_i (\ln p_i)^2 - \left( \sum_{i=1}^{n} p_i \ln p_i \right)^2.$$

Varentropy serves as a measure of the variability in the information content.
Lad et al. (2015) introduced the concept of *extropy*, which is the complement of Shannon entropy. The extropy of a discrete probability distribution $P = \{p_1, \ldots, p_n\}$ is defined as

$$J(P) = -\sum_{i=1}^{n} (1 - p_i) \ln(1 - p_i).$$

Let $X$ be an absolutely continuous random variable with common cumulative distribution function (cdf) $F_X$ and probability density function (pdf) $f_X$. Let $l_X = \inf\{x \in \mathbb{R} : F_X(x) > 0\}, u_X = \sup\{x \in \mathbb{R} : F_X(x) < 1\}$ and $S_X = (l_X, u_X)$. Then, Shannon (1948) defined differential entropy as a measure of uncertainty

$$H(X) = -\int_{S_X} f_X(x) \log f_X(x)\, dx.$$

Varentropy of $X$ is defined as (Arıkan, 2016; Maadani et al., 2022)

$$\begin{aligned} VH(X) &= \mathrm{Var}[-\log f_X(X)] \\ &= \int_{S_X} f_X(x) (\log f_X(x))^2\, dx - \left( \int_{S_X} f_X(x) \log f_X(x)\, dx \right)^2. \end{aligned}$$

This varentropy measure is widely used in data compression, finite blocklength information theory, and statistics, as it aids in determining ideal code lengths, source dispersion, and other relevant quantities. In statistics, it has proven to be a superior alternative to the kurtosis measure for continuous density functions; see (Arıkan, 2016; Dudewicz and van der Meulen, 1981; Hazeb et al., 2021; Maadani et al., 2022) studied entropy- and extropy-based goodness-of-fit tests for uniformity. An alternative measure of uncertainty, *extropy*, for a nonnegative absolutely continuous random variable $X$, defined by Lad et al. (2015), is given by

$$J(X) = \mathbb{E}\left( -\frac{1}{2} f_X(X) \right) = -\frac{1}{2} \int_{S_X} f_X^2(x)\, dx.$$

The primary objective of this study is to investigate the properties of varextropy and demonstrate its potential in testing for the uniformity of continuous probability distributions. Specifically, we aim to derive and explore the theoretical properties of varextropy, provide a characterization of the uniform distribution based on this measure, and develop a nonparametric estimator for varextropy from observed data. Additionally, we propose a novel test for uniformity that leverages varextropy, extending its applicability beyond descriptive analysis into inferential statistics. Finally, we evaluate the performance of the proposed test through both theoretical analysis and empirical validation using real-world data to assess its effectiveness in detecting deviations from the uniform distribution.

The main purpose of this paper is to obtain a test of uniformity using the derived characterization of the uniform distribution based on the varextropy of a continuous random variable. This paper is organized as follows. Section 2 contains some properties of varextropy. A characterization of the uniform distribution using varextropy is given in Section 3. A nonparametric estimator is given in Section 4. A test of uniformity is presented in Section 5, and Section 6 contains an application to real data.

## 2 Varextropy

The varextropy of a discrete probability distribution $P = \{p_1, \ldots, p_n\}$ is defined as (see, Goodarzi (2024); Vaselabadi et al. (2021))

$$VJ(P) = \sum_{i=1}^{n} (1-p_i)\left(\ln((1-p_i))\right)^2 - \left(\sum_{i=1}^{n} (1-p_i)\ln((1-p_i))\right)^2.$$

Varextropy also serves as a measure of the variability in the information content. Varextropy of absolutely continuous random variables $X$ is defined as

$$\begin{aligned} VJ(X) = \mathrm{Var}\left(-\frac{1}{2}f_X(X)\right) &= E\left(-\frac{1}{2}f_X(X) - J(X)\right)^2 \\ &= \frac{1}{4}E(f_X^2(X)) - \frac{1}{4}[E(f_X(X))]^2 \\ &= \frac{1}{4}\int_{S_X} f_X^3(x)\,dx - \frac{1}{4}\left(\int_{S_X} f_X^2(x)\,dx\right)^2. \end{aligned}$$

Note that $VJ(X) \geq 0$, for any random variable $X$. Vaselabadi et al. (2021) obtained several varextropy properties, as well as conditional varextropy properties based on order statistics, record values, and proportional hazard rate models. The article contains some comparative results regarding varextropy and varentropy. Goodarzi (2024) provided lower bounds for varextropy, obtained the varextropy of a parallel system, and used the varextropy of order statistics to construct a symmetry test. Zaid et al. (2022) computed the entropy, varentropy, and varextropy measures in closed form for generalized and $q$-generalized extreme value distributions. Varentropy is sometimes independent of the model parameters, whereas the varextropy measure is more adaptable, for example, when $X$ has a normal distribution with mean $\mu$ and variance $\sigma^2$ (see Vaselabadi et al. (2021)).

Chacko and Grace (2024) investigated the varextropy measure for the $n$th upper and lower $k$-record values, deriving expressions for both the measure and its residual and past forms. They applied this to estimate the varextropy of a two-parameter Weibull distribution using maximum likelihood estimations (MLEs) and Bayes estimates based on upper $k$-record values, with MCMC used for the Bayes estimates. Their simulation results showed that mean squared errors (MSEs) decreased as $n$ increased, and Bayes estimates outperformed MLEs. Among the Bayes estimators, those using the SEL function performed better, and the lowest MSE was achieved using Prior 1. Goodarzi (2022) derived the conditional covariance and variance for a parallel system with $n$ identical, independent components, assuming all components are still functioning at time $x$. A lower bound for the conditional variance was also provided. Additionally, lower bounds for varextropy were established, and the varextropy of a parallel system was calculated. The results were applied to create a symmetry test, with a real dataset used to illustrate the test statistics. Vaselabadi et al. (2021) explored several properties of the varextropy measure $VJ$, highlighting its use in quantifying information volatility in residual and past lifetimes. They examined its behavior in relation to order statistics, record values, and proportional hazard rate models. An approximate expression for $VJ(X)$ was also derived using a Taylor series expansion. Additionally, they introduced the concept of conditional varextropy and proposed a new stochastic order called varextropy ordering. Noughabi and Noughabi (2024) investigated the varextropy of a random variable and introduced consistent estimators for it, highlighting their location-invariant variance and mean squared error. Through Monte Carlo simulations, they evaluated the estimators' bias and RMSE under different distributions, showing that the proposed methods performed reliably across various scenarios.

In some situations, two random variables can have the same extropy, which prompts the age-old question, "Which of the extropies is a more appropriate criterion for measuring the uncertainty?" For example, consider random variables $U$ and $V$ (see, Balakrishnan et al. (2020)) with pdfs

$$f_U(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise}, \end{cases} \qquad \text{and} \qquad f_V(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{otherwise}. \end{cases}$$

We get $J(U) = J(V) = -1/2$, $VJ(U) = 0$, and $VJ(V) = 1/12$. This is the motivation behind considering the variance of $-\frac{1}{2}f(x)$, which is known as the varextropy of a random variable $X$. So, varextropy can also play a role in measuring uncertainty. The varextropy for some standard distributions are given in Table 1; for more example, see Vaselabadi et al. (2021).

Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed observations. An observation $X_j$ will be called an upper record value if its value exceeds that of all previous observations. Thus, $X_j$ is an upper record if $X_j > X_i$ for every $j > i$. See, Arnold et al. (1998) for more details about record values. A random variable $X$ is said to be smaller than $Y$ in the dispersive ordering ($X \leq_{disp} Y$) if $F_Y^{-1}(F_X(x)) - x$ is increasing in $x \geq 0$. Belzunce et al. (2001) showed that if $X \leq_{disp} Y$, then $U_n^X \leq_{disp} U_n^Y$, where $U_n^X$ and $U_n^Y$ are the $n$th upper records of $X$ and $Y$, respectively. Qiu (2017) showed that if $X \leq_{disp} Y$, then $J(X) \leq J(Y)$ and $J(U_n^X) \leq J(U_n^Y)$. Vaselabadi et al. (2021) showed that if $X \leq_{disp} Y$, then $VJ(X) \geq VJ(Y)$. In view of these results, it is conclude that $X \leq_{disp} Y$, then $VJ(U_n^X) \geq VJ(U_n^Y)$, for $n \geq 1$. It is obvious that if $X$ and $Y$ are identically distributed, that is, $X \stackrel{d}{=} Y$, then $VJ(X) = VJ(X)$, $VJ(X_{i:n}) = VJ(Y_{i:n})$ and $VJ(U_n^X) =$

Table 1: Expression for $VJ(X)$.

| Distribution | pdf | $VJ(X)$ |
|---|---|---|
| Uniform | $\frac{1}{b-a}, \quad a<x<b$ | 0 |
| Exponential | $\lambda e^{-\lambda x}, \quad x \geq 0, \lambda > 0$ | $\lambda^2/48$ |
| Weibull distribution | $2xe^{-x^2}, \quad x>0$ | $\frac{\sqrt{\pi}}{2^{3/2}} - \frac{\pi}{8}$ |
| Normal | $\frac{1}{\sqrt{2\pi}\sigma} e^{(-\frac{(x-\mu)^2}{\sigma^2})}, \quad -\infty<x<\infty$ | $\frac{2-\sqrt{3}}{16\pi\sigma^2\sqrt{3}}$ |
| Laplace distribution | $\frac{1}{2}e^{-\lvert x\rvert} \quad -\infty<x<\infty$ | $\frac{1}{24}$ |
| Logistic distribution | $\frac{e^{-x}}{(1+e^{-x})^2}, \quad -\infty<x<\infty$ | $\frac{1}{8}$ |
| Cauchy distribution | $\frac{1}{\pi(1+x^2)}, \quad -\infty<x<\infty$ | $\frac{1}{8\pi} - \frac{1}{16\pi^2}$ |

$VJ(U_n^Y)$, where $X_{i:n}$ is the $i$th order statistic in a random sample of size $n$. Vaselabadi et al. (2021) showed that varextropy is location-invariant but not scale-invariant, that is, if $Y = aX + b$, where $a > 0$ and $-\infty < b < \infty$, then $VJ(Y) = \frac{1}{a^2} VJ(X)$.

We have the following result for varextropy of order statistics of symmetric distribution.

**Lemma 1.** *Let $X_1, X_2, \ldots, X_n$ be random sample from continuous distribution with symmetric around a finite $\mu$ with sample size $n$. Then*

$$VJ(X_{i:n}) = VJ(X_{n-i+1:n}), \quad 1 \leq i \leq n.$$

*Proof.* The result follows by location-invariant property of varextropy. □

## 3 Weighted varextropy

Applications of weighted distributions include distribution theory, dependability, probability, ecology, biostatistics, and applied statistics. Two random variables can have the same extropy as well as the same varextropy in some situations. For example, consider random variables $X$ and $Y$ with pdfs, respectively:

$$f_X(x) = \begin{cases} 2x, & 0<x<1, \\ 0, & \text{otherwise,} \end{cases} \qquad f_Y(x) = \begin{cases} 2(1-x), & 0<x<1, \\ 0, & \text{otherwise.} \end{cases}$$

We get $J(X) = J(Y) = -2/3$, $VJ(X) = VJ(Y) = 1/18$, but $VJ^w(X) = 1/12$ and $VJ^w(Y) = 1/180$. So here, weighted varextropy can also play a role as a measure of uncertainty. Gupta and Chaudhary (2023) defined general weighted extropy with nonnegative weight $w(x)$ as

$$J^w(X) = -\frac{1}{2}\int_0^\infty w(x) f_X^2(x)\, dx.$$

Table 2: Expression for $VJ^x(X)$.

| Distribution | pdf | $VJ^x(X)$ |
|---|---|---|
| Uniform | $\frac{1}{b-a},\quad a<x<b$ | $\frac{1}{4}\left[\frac{b^3-a^3}{3(b-a)^3}-\frac{(b^2-a^2)^2}{4(b-a)^4}\right]$ |
| Exponential | $\lambda e^{-\lambda x},\quad x\geq 0,\lambda>0$ | $\frac{5}{1728}$ |
| Weibull distribution | $2xe^{-x^2},\quad x>0$ | $\frac{1}{4}\left(\frac{1}{3^{3/2}}-1\right)$ |
| Normal | $\frac{1}{\sqrt{2\pi}\sigma}e^{\left(-\frac{(x-\mu)^2}{\sigma^2}\right)},\quad -\infty<x<\infty$ | $\frac{1}{4}\left(\frac{1}{(2\pi\sigma^2)^{3/2}}\left(\mu^2+\frac{1}{6\sigma^2}\right)-\mu^2\right)^2$ |
| Laplace distribution | $\frac{1}{2}e^{-\lvert x\rvert}\quad -\infty<x<\infty$ | $\frac{1}{216}$ |
| Logistic distribution | $\frac{e^{-x}}{(1+e^{-x})^2},\quad -\infty<x<\infty$ | $\frac{1}{8}$ |
| Cauchy distribution | $\frac{1}{\pi(1+x^2)},\quad -\infty<x<\infty$ | $\frac{1}{16\pi^2}$ |

We define the general weighted varextropy of a discrete probability distribution $P=\{p_1,\ldots,p_n\}$ with $X=\{x_1,x_2,\ldots,x_n\}$ and weights $w=\{w_1,w_2,\ldots,w_n\}$ as

$$VJ^w(P)=\sum_{i=1}^{n}w_i^2(1-p_i)\left(\ln((1-p_i))\right)^2-\left(\sum_{i=1}^{n}w_i(1-p_i)\ln((1-p_i))\right)^2.$$

When $w_i=x_i,\ \forall\ i=1,2,\ldots,n$, then the weighted varextropy is given as

$$VJ^x(P)=\sum_{i=1}^{n}x_i^2(1-p_i)\left(\ln((1-p_i))\right)^2-\left(\sum_{i=1}^{n}x_i(1-p_i)\ln((1-p_i))\right)^2.$$

We define general weighted varextropy for an absolutely continuous random variable as

$$\begin{aligned}VJ^w(X)&=\operatorname{Var}\left(-\frac{1}{2}w(X)f_X(X)\right)\\&=\frac{1}{4}\left[E(w^2(X)f^2(X))-(E(w(X)f_X(X)))^2\right]\\&=\frac{1}{4}\left[\int_{S_X}w^2(x)f^3(x)\,dx-\left(\int_{S_X}w(x)f^2(x)\,dx\right)^2\right].\end{aligned}$$

When $w(x)=x$, then weighted varextropy is given as

$$VJ^x(X)=\frac{1}{4}\left[\int_{S_X}x^2f^3(x)\,dx-\left(\int_{S_X}xf^2(x)\,dx\right)^2\right].$$

The weighted varextropy $VJ^x(X)$ for some standard distributions are given in Table 2.

## 4 A characterization of uniform distribution

In many practical problems, the goodness-of-fit test may be reduced to the problem of testing uniformity. Since the varextropy of $X$ is the variance of $-\frac{1}{2}f_X(x)$, the varextropy is nonnegative for any random variable $X$. Among all distributions with support on $[0,1]$, the uniform distribution has the maximum extropy. An important property of the uniform distribution is that it obtains the minimum varextropy among all distributions having support on $[0,1]$ (see, Qiu and Jia (2018)).

The characterization provided in Theorem 1 is significant because it establishes a clear and definitive criterion for identifying a uniform distribution based on varextropy. By showing that a random variable $X$ has zero varextropy if and only if it is uniformly distributed over $[0,1]$, it offers a direct and precise method for testing uniformity without needing complex parametric assumptions. This improves upon existing characterizations by linking uniformity to an easily computable quantity, varextropy, which is grounded in variance, making it more practical for statistical analysis. Previous methods might have relied on more complex or indirect approaches, but the varextropy-based test is simple, theoretically sound, and offers a direct comparison for uniformity. This approach fills a gap by providing a nonparametric and computationally feasible solution to uniformity testing, making it a more accessible tool in both theoretical and applied statistics.

The characterization in Theorem 1 makes a few key assumptions. First, it assumes that the random variable $X$ is continuous and has support on the interval $[0,1]$. This is crucial because the result specifically applies to distributions confined to this interval, such as the uniform distribution. Second, the characterization assumes that the pdf $f_X(x)$ is well-defined and continuous over this support. This ensures that the varextropy formula, which relies on the second moment of the pdf, can be computed without encountering issues related to discontinuities or undefined behavior. Additionally, the proof assumes that $f_X(x)$ integrates to 1 over $[0,1]$, which is a fundamental property of any valid probability density function. These assumptions are necessary to guarantee the correctness and applicability of the characterization, ensuring it is valid for continuous distributions on the unit interval and can be used as a reliable test for uniformity. Noughabi and Noughabi (2023) applied varentropy to test for uniformity. They showed that the varentropy of $X$ is zero if and only if $X$ follows the standard uniform distribution, and they used their proposed varentropy estimators as test statistics for conducting goodness-of-fit tests for uniformity. Following result is a characterization of the uniform distribution using varextropy (see (Chaudhary and Gupta, 2024, Theorem 11)).

**Theorem 1.** *Let $X$ be a continuous random variable with support on $[0,1]$. Then $VJ(X) = 0$ if and only if $X$ has a uniform distribution on the interval $[0,1]$.*

*Proof.* Let random variable $X$ have a uniform distribution on the interval $[0,1]$; then $f_X(x) = 1$ for $0 \leq x \leq 1$, and

$$VJ(X) = \frac{1}{4}\int_0^1 f^3(x)\,dx - \frac{1}{4}\left[\int_0^1 f^2(x)\,dx\right]^2 = 0.$$

Conversely, $VJ(X) = 0$ implies $\operatorname{Var}(f_X(X)) = 0$, that is, $f_X(x) = c$. Since

$$\int_0^1 f_X(x)\,dx = 1, \quad \text{therefore} \quad f_X(x) = 1, \quad 0 \le x \le 1.$$

Hence, the proof is complete. □

## 5 Nonparametric estimators

Suppose that $X_{1:n}, X_{2:n}, X_{3:n}, \ldots, X_{n:n}$ are order statistics of random sample $X_1, X_2, \ldots, X_n$ from cdf $F$. Then, then empirical distribution function of cdf $F$ is given by

$$F_n(x) = \begin{cases} 0, & x < X_{1:n} \\ \frac{i}{n}, & X_{i:n} \le x < X_{i+1:n}, \quad i = 1, 2, \ldots, n-1, \\ 1, & x \ge X_{n:n}. \end{cases}$$

Noughabi and Noughabi (2024) provided various estimators of $VJ(X)$. $VJ(X)$ can be expressed as

$$VJ(X) = \frac{1}{4}\left[\int_0^1 \left(\frac{d}{dp}(F^{-1}(p))\right)^{-2} dp - \left(\int_0^1 \left(\frac{d}{dp}(F^{-1}(p))\right)^{-1} dp\right)^2\right].$$

Following the idea of Vasicek (1976), Noughabi and Noughabi (2024) proposed the estimator $\Delta$ for $VJ(X)$ as

$$\Delta = \frac{1}{4n}\sum_{i=1}^{n}\left(\frac{2m/n}{X_{i+m:n} - X_{i-m:n}}\right)^2 - \frac{1}{4}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{2m/n}{X_{i+m:n} - X_{i-m:n}}\right)\right)^2.$$

Here, the window size $m$ is a positive integer less than or equal to $\frac{n}{2}$. If $i + m > n$, then we consider $X_{i+m:n} = X_{n:n}$, and if $i - m < 1$, then we consider $X_{i-m:n} = X_{1:n}$. The proposed estimator for varextropy, $\Delta$, calculates the weighted variance of order statistics based on sample data, using a window size parameter $m$. It is defined as

$$\Delta = \frac{1}{4n}\sum_{i=1}^{n}\left(\frac{2m/n}{X_{i+m:n} - X_{i-m:n}}\right)^2 - \frac{1}{4}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{2m/n}{X_{i+m:n} - X_{i-m:n}}\right)\right)^2.$$

This estimator is consistent, meaning it converges to the true value of varextropy as the sample size increases and is flexible for a range of distributions. Its primary advantages include its practical applicability for goodness-of-fit tests, such as testing uniformity, and its ability to offer consistent results for large datasets. However, it is sensitive to the choice of the window size $m$, and for large samples, it can become computationally intensive. Additionally, for small sample sizes, the estimator may not be highly accurate, and its distribution under the null hypothesis requires Monte Carlo simulations to determine critical values. Despite these limitations, the estimator improves upon existing methods by directly utilizing order statistics and providing a reliable approach to test uniformity.

The proposed estimator, $\Delta$, for varextropy offers several advantages when compared to other estimators used in statistical tests for uniformity. One key feature is its use of order statistics, which captures more nuanced information about the distribution of data, particularly in non-parametric contexts. Compared to traditional estimators like the sample variance or methods based on moments, $\Delta$ incorporates weighted variations in the sample, making it more sensitive to the underlying distribution, especially for detecting deviations from uniformity.

When compared to earlier estimators like the ones proposed by Noughabi and Noughabi (2023), which are based on varentropy for uniformity testing, the proposed estimator has a distinct advantage in terms of flexibility and consistency. While their estimator is also consistent, it is based on a more complex approach, potentially requiring additional assumptions about the distribution shape. The $\Delta$ estimator, on the other hand, relies on empirical distributions and requires fewer assumptions about the underlying data, making it more adaptable.

However, one limitation of $\Delta$ is its reliance on the window size parameter $m$, which requires tuning and may impact its performance in smaller datasets. Other methods, like those using bootstrap resampling techniques, can offer an alternative, providing robust estimates without relying on window size. Overall, the proposed estimator provides a more robust and flexible approach than many existing alternatives, particularly when testing for uniformity in real-world data.

## 6 A test of uniformity

In this section, we introduce a statistical test for uniformity based on the concept of varextropy, specifically using the estimator $\Delta$ proposed by Noughabi and Noughabi (2024). It has been established that the varextropy of a random variable $X$ is zero if and only if $X$ follows a standard uniform distribution. Using this property, we can utilize the proposed varextropy estimators as test statistics for conducting goodness-of-fit tests to determine whether a given sample follows a uniform distribution. The hypothesis of interest is framed as follows:

- Null hypothesis ($H_0$): The random variable $X$ is uniformly distributed.
- Alternative hypothesis ($H_1$): The random variable $X$ is not uniformly distributed.

We propose using $\Delta$, an estimator of the varextropy $VJ(X)$, as the test statistic. The estimator $\Delta$ is consistent, meaning that as the sample size $n$ increases, $\Delta$ converges in probability to the true value of $VJ(X)$. Under the null hypothesis $H_0$, if $X$ follows a uniform distribution, $\Delta$ converges in probability to zero. On the other hand, if $X$ is not uniformly distributed (under $H_1$), $\Delta$ converges to a nonzero value. This distinction allows us to use large values of $\Delta$ as evidence of nonuniformity. Therefore, we reject the null hypothesis when $\Delta$ exceeds a certain threshold.

Since the distribution of the test statistic $\Delta$ under the null hypothesis is too complex to derive analytically, we use Monte Carlo simulation to empirically determine the critical values and power of the test. The critical region for the test is defined as $\Delta \geq C_{1-\alpha}$, where $C_{1-\alpha}$ is the critical value corresponding to the significance level $\alpha$. For a given sample size $n$ and significance level $\alpha$, we compute $C_{1-\alpha}$ using a Monte Carlo simulation. This approach allows us to determine the appropriate threshold for rejecting the null hypothesis based on simulated data from the uniform distribution.

Table 3: Critical values at significance level $\alpha = 0.05$.

| $m \backslash n$ | 10 | 20 | 30 | 40 | 50 | 80 | 100 |
|---|---|---|---|---|---|---|---|
| 2 | 4.7570 | 3.1388 | 2.2838 | 1.9478 | 1.6402 | 1.1355 | 1.0451 |
| 3 | 1.4909 | 1.2126 | 0.7925 | 0.6502 | 0.5729 | 0.4235 | 0.3559 |
| 4 | 0.7064 | 0.6089 | 0.4724 | 0.3841 | 0.3434 | 0.2541 | 0.2121 |
| 5 | 0.4074 | 0.4252 | 0.3525 | 0.2881 | 0.2396 | 0.1800 | 0.1541 |
| 9 | | 0.1551 | 0.1703 | 0.1542 | 0.1399 | 0.1039 | 0.0947 |
| 14 | | | 0.0869 | 0.0973 | 0.1006 | 0.0829 | 0.0731 |
| 19 | | | | 0.0637 | 0.0722 | 0.0722 | 0.0665 |
| 24 | | | | | 0.0528 | 0.0639 | 0.0619 |
| 30 | | | | | | 0.0514 | 0.0546 |
| 39 | | | | | | 0.0380 | 0.0436 |
| 49 | | | | | | | 0.0336 |

## 6.1 Critical points

We define a function to calculate the value of $\Delta$. A sample of size $n$ is generated from the $U(0,1)$ distribution, and the test statistic is computed for the sample data. After 10,000 replications, the $(1-\alpha)^{\text{th}}$ quantile of the test statistics is determined as the critical value at significance level $\alpha$. Critical values for $\alpha = 0.05$ are given in Table 3 for different values of $m$ and $n$.

To derive the critical values of the proposed test statistic $\widehat{\Delta}$, we employ a Monte Carlo simulation approach due to the analytical intractability of its sampling distribution under the null hypothesis. Specifically, we generate 10,000 independent random samples of size $n$ from the standard uniform distribution $U(0,1)$, which represents the null hypothesis $H_0$. For each simulated sample, we compute the value of the test statistic $\widehat{\Delta}$ using the nonparametric estimator that involves a fixed window size $m$. After obtaining 10,000 such values of $\widehat{\Delta}$, we determine the empirical $(1-\alpha)$-th quantile to serve as the critical value $C_{1-\alpha}$ at a given significance level $\alpha$. These critical values are summarized in Table 3 for various combinations of $n$ and $m$, thus providing practical benchmarks for implementation.

## 6.2 Power of test

We used the following procedure to estimate the power of the test. For each sample size $n$, we generate 10,000 random samples of size $n$ from the alternative distribution. The test statistic is then computed for each sample. The power of the test at a significance level $\alpha$ is estimated as the proportion of these 10,000 samples that fall within the corresponding critical region.

The estimated power of the test is obtained as the proportion of samples for which the test statistic exceeds the critical value, leading to the rejection of $H_0$. This empirical procedure provides a consistent and practical method for evaluating the effectiveness of the test. The results, presented in Table 4, demonstrate that the proposed test performs well in detecting deviations from uniformity and exhibits higher power compared to existing tests for standard alternatives like the Beta(1,2) distribution. The pdf of the beta distribution with parameters $a$

Table 4: Power at significance level $\alpha = 0.05$.

| $m \backslash n$ | 10 | 20 | 30 | 40 | 50 | 80 | 100 |
|---|---|---|---|---|---|---|---|
| 2 | 0.0817 | 0.0966 | 0.0960 | 0.1038 | 0.1137 | 0.1492 | 0.1746 |
| 3 | 0.1197 | 0.1343 | 0.1558 | 0.1689 | 0.1836 | 0.2522 | 0.2924 |
| 4 | 0.1546 | 0.1882 | 0.2082 | 0.2302 | 0.2570 | 0.3536 | 0.4467 |
| 5 | 0.1786 | 0.2370 | 0.2621 | 0.2874 | 0.3359 | 0.4451 | 0.5483 |
| 9 | | 0.3962 | 0.4360 | 0.4654 | 0.5072 | 0.6540 | 0.7211 |
| 14 | | | 0.5991 | 0.6221 | 0.6566 | 0.7380 | 0.8187 |
| 19 | | | | 0.7752 | 0.7734 | 0.8311 | 0.8700 |
| 24 | | | | | 0.8875 | 0.8829 | 0.9128 |
| 30 | | | | | | 0.9397 | 0.9459 |
| 39 | | | | | | 0.9883 | 0.9850 |
| 49 | | | | | | | 0.9983 |

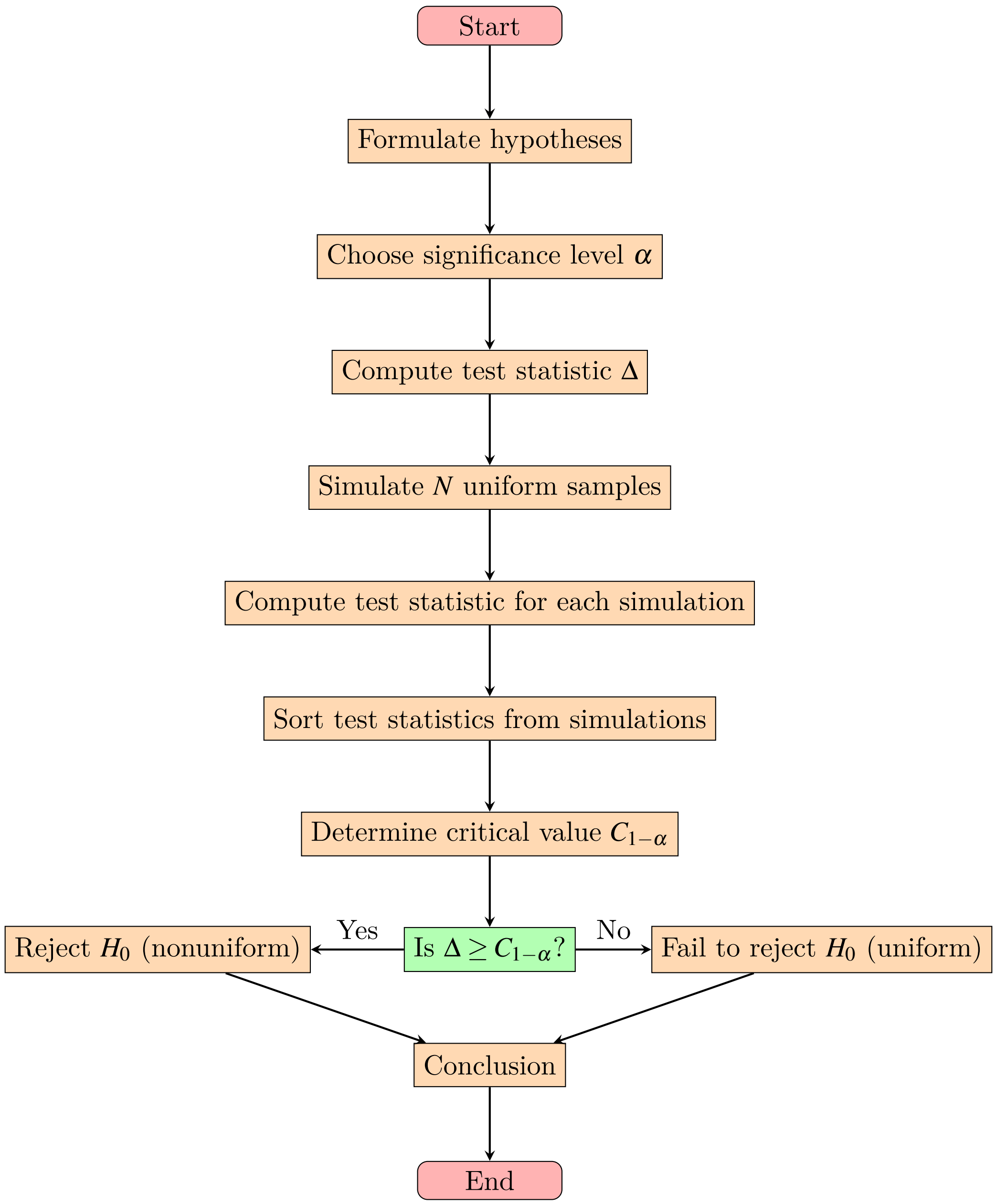
and $b$ is given by

$$f_x(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}, \quad 0 < x < 1,$$

where $B(a,b)$ is the complete beta function.

To estimate the power of the test, we again utilize a Monte Carlo simulation, this time under the alternative hypothesis $H_1$. We generate 10,000 random samples of size $n$ from a nonuniform distribution, such as the Beta(1,2) distribution, which is a common alternative to $U(0,1)$. For each sample, the value of $\widehat{\Delta}$ is calculated and compared to the corresponding critical value derived under $H_0$.

The estimated power against the alternative Beta(1,2) distribution is given in Table 4 at the significance level $\alpha = 0.05$. Our test performs well in detecting nonuniform data. Note that Beta(1,1) is identically distributed with $U(0,1)$. The power of this test against the alternative Beta(1,1) is approximately $\alpha$, so the test achieves its level of significance. The power of our test is higher than the power of the test proposed by Noughabi and Noughabi (2023) for the common alternative Beta(1,2).




## 6.3 Application to real data

The dataset used in this application comes from a real-world environmental study involving vinyl chloride concentrations. Vinyl chloride is a toxic substance, and understanding its distribution is critical for assessing environmental risks and regulatory compliance. In particular, the dataset represents a sample of vinyl chloride measurements that have been transformed to fit a uniform distribution using the probability integral transformation, as proposed by Xiong et al. (2022). This transformation is commonly used to standardize nonuniform data so that it can be tested for uniformity.

The choice of this dataset is directly linked to our theoretical findings in the previous sections, particularly the application of the test of uniformity based on the varextropy estimator $\widehat{\Delta}$. As we shown earlier, the test statistic $\widehat{\Delta}$ can effectively detect whether a dataset conforms to a uniform distribution, which is important for validating the uniformity of the transformed data. Given that uniformity is a key assumption in many statistical procedures, it is essential to verify whether the transformation of the vinyl chloride concentrations actually results in data that adheres to the uniform distribution.

For this dataset, we computed the value of the proposed test statistic $\widehat{\Delta}$ using a window size $m = 16$ and sample size $n = 34$. The computed statistic was found to be $\widehat{\Delta} = 0.0329$. The critical value for the test at a significance level of $\alpha = 0.05$ was obtained through Monte Carlo simulation, resulting in a critical value of 0.0733 for $m = 16$ and $n = 34$. Since the test statistic $\widehat{\Delta} = 0.0329$ is less than the critical value of 0.0733, the observed statistic lies within the acceptance region.

This outcome suggests that the transformed data conforms to the uniform distribution, and therefore, we fail to reject the null hypothesis of uniformity. In other words, our proposed test successfully verifies that the transformation applied to the vinyl chloride data indeed resulted in a uniform distribution, aligning with the expectations of the transformation method.

**Dataset 1:** 0.0518, 0.0518, 0.1009, 0.1009, 0.1917, 0.1917, 0.1917, 0.2336, 0.2336, 0.2336, 0.2733, 0.2733, 0.3467, 0.3805, 0.3805, 0.4126, 0.4431, 0.4719, 0.4719, 0.4993, 0.6162, 0.6550, 0.6550, 0.7059, 0.7211, 0.7356, 0.7623, 0.7863, 0.8178, 0.8810, 0.9337, 0.9404, 0.9732, 0.9858.

The dataset represents vinyl chloride concentrations transformed into a uniform distribution using the probability integral transformation (Xiong et al., 2022). The value of the test statistic $\widehat{\Delta}$ is 0.0329 when the window size $m = 16$ and the sample size $n = 34$. The critical point is 0.0733 at the 5% level of significance, based on Monte Carlo simulations for $m = 16$ and $n = 34$. Since the estimated value of the test statistic lies in the acceptance region, our test based on $\widehat{\Delta}$ fails to reject the null hypothesis. Therefore, the test verifies that the data is fitted to a uniform distribution.

The results of our uniformity test based on the varextropy estimator $\widehat{\Delta}$ indicate that the transformed vinyl chloride data fits well with a uniform distribution. The calculated test statistic ($\widehat{\Delta} = 0.0329$) was smaller than the critical value (0.0733) at the 5% significance level, suggesting that we failed to reject the null hypothesis of uniformity.

However, there are several limitations and potential biases in our analysis. First, the critical values were derived using Monte Carlo simulations, which, while accurate, are approximations and depend on the number of replications used (10,000 in this case). Furthermore, the performance of the test is influenced by the sample size, and our results may not generalize well to smaller or larger samples. Another limitation is the assumption that the data under the null hypothesis is perfectly uniform, which may not always hold in practice, especially with real-world data where small deviations from uniformity can occur. Additionally, the window size used in the calculation of $\widehat{\Delta}$ could impact the test's power and its sensitivity to nonuniformity. While the test performed well in detecting significant departures from uniformity in this case, its ability to detect subtle differences might be limited. Moreover, the Monte Carlo method, while effective, can introduce bias if the number of replications is not large enough or if the underlying assumptions about the test statistic are inaccurate.

Lastly, the choice of dataset, in this case, vinyl chloride concentrations, may not be representative of other datasets, and the conclusions drawn here might not be directly applicable to different contexts. Despite these limitations, the test demonstrates its utility in assessing uniformity, and future research can refine its performance and extend its application to other distributions.

## 7 Conclusion

The results of our study demonstrate that the proposed test based on the varextropy estimator $\widehat{\Delta}$ is an effective tool for assessing the uniformity of datasets. We applied the test to real-world vinyl chloride concentration data that had been transformed to fit a uniform distribution using the probability integral transformation. The test statistic $\widehat{\Delta} = 0.0329$ was found to be smaller than the critical value of 0.0733 at a 5% significance level, leading to the conclusion that the transformed data conforms to a uniform distribution. This outcome is consistent with our theoretical expectation that $\widehat{\Delta}$ should be small when the data follows a uniform distribution.

Moreover, the test showed strong performance in detecting deviations from uniformity when applied to simulated data from a Beta(1,2) distribution, a common alternative to the uniform distribution. As expected, the test's power increased with sample size, and the critical values, derived through Monte Carlo simulations, provided a reliable framework for determining decision thresholds for uniformity testing at various levels of significance. These results demonstrate the robustness and practicality of the proposed test, especially in the context of assessing uniformity in real-world datasets.

However, the study also highlights certain limitations. The Monte Carlo simulation approach, while effective, relies on approximations that depend on the number of replications used, and the performance of the test can be influenced by the sample size and the choice of window size. Additionally, the test assumes that the null hypothesis represents perfectly uniform data, which may not always hold in practice, particularly when small deviations from uniformity exist in real-world data. These factors suggest that while the test performs well under the given conditions, its generalizability to other datasets and scenarios may require further investigation.

Future research could address several areas for improvement. First, exploring the test's performance with smaller sample sizes and more diverse datasets would provide insights into its robustness and lead to better calibration of critical values. Investigating the test's sensitivity to a wider range of nonuniform distributions, beyond Beta(1,2), could help evaluate its applicability to different data patterns. Enhancing the Monte Carlo simulation process through more efficient sampling techniques or parallel processing could improve both computational speed and scalability. Additionally, examining the test's robustness to different distributional assumptions, such as normal or skewed distributions, would further validate its flexibility.

Comparing the varextropy-based test with other established goodness-of-fit tests, such as the Kolmogorov–Smirnov or Anderson-Darling tests, could provide valuable insights into its relative strengths and weaknesses. Expanding the test's application to multivariate or time-series data could broaden its utility in domains such as finance, ecology, and other fields that deal with complex data structures. Furthermore, implementing dynamic window size selection methods might enhance the accuracy and adaptability of the test across various datasets.

Lastly, applying the test to additional real-world datasets from diverse fields would help assess its practical applicability and identify domain-specific challenges or opportunities for refinement. Addressing these areas in future research could improve the accuracy, efficiency, and broad applicability of the proposed uniformity test, making it a valuable tool for detecting uniformity across a variety of statistical and applied contexts

## Acknowledgement

The authors are grateful to the editors and anonymous referees for their insightful comments and valuable suggestions, which have significantly improved the quality of this paper.

## References

Arıkan, E. (2016). Varentropy decreases under the polar transform. *IEEE Transactions on Information Theory*, **62**, 3390–3400.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998). *Records*. John Wiley & Sons, New York.

Balakrishnan, N., Buono, F., and Longobardi, M. (2020). On weighted extropies. *Communications in Statistics - Theory and Methods*, **51**, 6250–6267.

Belzunce, F., Guillamón, A., Navarro, J., and Ruiz, J. M. (2001). Kernel estimation of residual entropy. *Communications in Statistics - Theory and Methods*, **30**, 1243–1255.

Chacko, M., and Grace, A. (2024). Varextropy of $k$-record values and its applications. *Communications in Statistics - Theory and Methods*, 1–24.

Chaudhary, S. K., and Gupta, N. (2024). Extropy and varextropy estimators with applications. *arXiv preprint* arXiv:2401.13065.

De Crescenzo, A., Paolillo, L., and Suárez-Llorens, A. (2025). Stochastic comparisons, differential entropy and varentropy for distributions induced by probability density functions. *Metrika*, **88**, 43–59.

Dudewicz, E. J., and van der Meulen, E. C. (1981). Entropy-based tests of uniformity. *Journal of the American Statistical Association*, **76**, 967–974.

Goodarzi, F. (2022). Results on conditional variance in parallel system and lower bounds for varextropy.*Communications in Statistics - Theory and Methods*, **53**, 1590–1610.

Goodarzi, F. (2024). Results on conditional variance in parallel system and lower bounds for varextropy. *Communications in Statistics - Theory and Methods*, **53**, 1590–1610.

Gupta, N., and Chaudhary, S. K. (2023). Some characterizations of continuous symmetric distributions based on the extropy of record values. *Statistical Papers*, 1–18.

Hazeb, R., Bayoud, H., and Raqab, M. (2021). Kernel and CDF-based estimation of extropy and entropy from progressively type-II censoring with application for goodness of fit problems. *Stochastics and Quality Control*, **36**(1), 73–83.

Lad, F., Sanfilippo, G., and Agro, G. (2015). Extropy: Complementary dual of entropy. *Statistical Science*, **30**, 40–58.

Maadani, S., Mohtashami Borzadaran, G. R., and Rezaei Roknabadi, A. H. (2022). Varentropy of order statistics and some stochastic comparisons. *Communications in Statistics - Theory and Methods*, **51**, 6447–6460.

Noughabi, H. A., and Noughabi, M. S. (2023). Varentropy estimators with applications in testing uniformity. *Journal of Statistical Computation and Simulation*, **93**, 2582–2599.

Noughabi, H. A., and Noughabi, M. S. (2024). On the estimation of varextropy. *Statistics*, **58**, 827–841.

Qiu, G. (2017). The extropy of order statistics and record values. *Statistics and Probability Letters*, **120**, 52–60.

Qiu, G., and Jia, K. (2018). Extropy estimators with applications in testing uniformity. *Journal of Nonparametric Statistics*, **30**, 182–196.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.

Vaselabadi, N. M., Tahmasebi, S., Kazemi, M. R., and Buono, F. (2021). Results on varextropy measure of random variables. *Entropy*, **23**, 356.

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **38**, 54–59.

Xiong, P., Zhuang, W., and Qiu, G. (2022). Testing exponentiality based on the extropy of record values. *Journal of Applied Statistics*, **49**, 782–802.

Zaid, E. O. A., Attwa, R. A. E. W., and Radwan, T. (2022). Some measures information for generalized and q-generalized extreme values and its properties. *Fractals*, **30**, 2240246.