frontiers | Frontiers in Applied Mathematics and Statistics

TYPE Original Research
PUBLISHED 02 February 2026
DOI 10.3389/fams.2026.1733600

Further aspects of information-generating function of order statistics with health application in symmetry of chronic disease management

Mohamed Said Mohamed[1], Manal Al-Labadi[2], Eman Almuhur[3] and Hanan H. Sakr[4]*

[1]Department of Mathematics, College of Science and Humanities, Prince Sattam bin Abdulaziz, University, Hawtat Bani Tamim, Saudi Arabia, [2]Department of Mathematics, Faculty of Arts and Sciences, University of Petra, Amman, Jordan, [3]Department of Mathematics, Faculty of Science, Applied Science Private University, Amman, Jordan, [4]Department of Management Information Systems, College of Business Administration in Hawtat Bani Tamim, Prince Sattam Bin Abdulaziz University, Hawtat Bani Tamim, Saudi Arabia

Check for updates

OPEN ACCESS

EDITED BY
Han-Ying Liang,
Tongji University, China

REVIEWED BY
Zakariya Yahya Algamal,
University of Mosul, Iraq
Shuji Ando,
Tokyo University of Science, Japan

*CORRESPONDENCE
Hanan H. Sakr
h.sakr@psau.edu.sa

RECEIVED 27 October 2025
REVISED 24 December 2025
ACCEPTED 05 January 2026
PUBLISHED 02 February 2026

CITATION
Mohamed MS, Al-Labadi M, Almuhur E and Sakr HH (2026) Further aspects of information-generating function of order statistics with health application in symmetry of chronic disease management. *Front. Appl. Math. Stat.* 12:1733600. doi: 10.3389/fams.2026.1733600

This investigation aimed to explore novel theoretical aspects and applications of the information-generating function measure for order statistics. We developed fundamental properties and established stochastic ordering relationships based on this information-theoretic measure. Our analysis demonstrated that when two order statistics share identical information-generating measures, their underlying parent distributions can be uniquely identified. We implemented our proposed measure to characterize the exponential distribution. Moreover, we derived bounds and investigated monotonicity properties for these functional measures. The study further examined how information-generating functions characterize distributional symmetry, with particular applications to uniform and normal distributions for identifying symmetry points of order statistics. Building on these theoretical foundations, we proposed a new symmetry test statistic derived from the information-generating properties of the order statistics. Using comprehensive Monte Carlo simulations, we evaluated the test's statistical power against existing alternatives. The present results demonstrated superior performance across various asymmetric distributional alternatives. The practical utility of our methodology is illustrated through an empirical analysis of chronic disease prevalence data.

KEYWORDS

information-generating function, non-parametric estimation, order statistics, stochastic order comparison, symmetry testing

1 Introduction and background

Several criteria have been proposed in information theory to gauge a probabilistic model's degree of uncertainty. The most significant information measurement that has been applied in several scientific and technical fields is the Shannon entropy. It started with Shannon's groundbreaking research [1], which examined how systems behaved when characterized by probability density or mass functions (pdf or pmf). Assuming that the variable X^* has a pdf $h(x)$ in the continuous case, the differential entropy, often known as the Shannon entropy, is analogously provided by

$$En(X^*) = -\int_{-\infty}^{\infty} h(x) \ln h(x)\, dx. \tag{1}$$

One practical technique for assessing the variance, mean, and other moments of a probability distribution is its moment-generating function. If there are successive moments in the probability distribution, they may be found by taking the sequential derivatives of the moment-generating function at zero. To calculate information quantities like extropy, Kullback-Leibler divergence, and Shannon information, generating functions for PDFs have been defined in information theory. As long as the integral remains in existence, the information-generating function of a random variable X^* was suggested by Golomb [2], who was inspired by the ideas of moments and probabilities of generating functions. It is defined as

$$GEn_\delta(X^*) = E\left(e^{(\delta-1)\ln h(x)}\right) = \int_{-\infty}^{\infty} h^\delta(x)\, dx, \tag{2}$$

for any $\delta > 0$. Golomb [2] then demonstrated the following features of the information-generating function as

1. $GEn_1(X^*) = 1$
2. $\frac{\partial}{\partial\delta} GEn_\delta(X^*)|_{\delta=1} = -En(X^*)$ (the negative of Shannon's entropy in Equation 1).

Because information-generating functions are important in information theory, several authors have recently investigated them. For a list of information-generating functions and their many features and uses, see Kharazmi and Balakrishnan [3–7], Zamani et al. [8], Kharazmi et al. [9], and Kayal and Balakrishnan [10].

Specifically, the information-generating function measure is simplified to $GEn_2(X^*)$, sometimes referred to as the informative-energy function, when $\delta = 2$. Using the example of kinetic energy in mechanics, Onicescu [11] introduced a discrete version of the informative-energy measurement into information theory. Bhatia [12] provides further information.

In many statistical methodologies, it is commonly assumed that the distribution of the population under study is symmetric. For example, the validity of regression models often hinges on the assumption that the residuals exhibit symmetry. This makes it critical to rigorously assess whether the symmetry assumption holds in practice. Consider that the support of the cumulative distribution function (cdf) H is denoted by S_X^*. Assume further that there exists a constant μ^* such that for all $x \in S_X^*$, the equation $H(\mu^* - x) + F(\mu^* + x) = 1$ is satisfied. When this condition is met, the distribution of X^* is considered symmetric about the point μ^*.

Symmetry is a concept of substantial theoretical and practical importance in both probability and statistics. It underpins many models and inferential procedures and has been explored extensively across various contexts. Researchers have introduced a range of characterizations for symmetric distributions, often using ordered samples such as order statistics, record values, and sequential statistics. For instance, Balakrishnan and Selvitella [13] showed that, for a sample of size m, the distributional identity $X_{i,m} \overset{DI}{=} X_{m-i+1,m}$ holds for a fixed $i = 1, \ldots, m$ if and only if the underlying distribution H is symmetric about zero. In this notation, $\overset{DI}{=}$ signifies that the two random variables have identical distributions.

Furthermore, Ahmadi [14] introduced innovative formulations of symmetry for continuous distributions by leveraging the properties of k-record values. Building on this foundation, Mahdizadeh and Zamanzade employed ranked set sampling techniques to construct nonparametric estimators of symmetric distribution functions [15]. Broadly speaking, assessing symmetry often involves developing criteria tailored to its specific structural features. This task is frequently carried out using goodness-of-fit tests, as demonstrated by Dai et al. [16] and Bozin et al. [17].

In this study, we explore several stochastic orderings that are useful for comparing random variables in a meaningful way. Suppose X_1^* and X_2^* are two continuous random variables with pdfs h_1 and h_2, and corresponding cdfs H_1 and H_2. Their generalized inverses (also known as left-continuous quantile functions) are defined as $H_1^{-1}(x) = \inf\{v : H_1(v) \geq x\}$ and $H_2^{-1}(x) = \inf\{v : H_2(v) \geq x\}$ for $0 < x < 1$.

Based on these definitions, we say that X_1^* is smaller than X_2^* in various stochastic orders if the following conditions hold for all $x \geq 0$:

(1) Likelihood Ratio Order ($X_1^* \leq^{lr} X_2^*$): This ordering holds if the ratio $\frac{h_1(x)}{h_2(x)}$ is a decreasing function of x.
(2) Hazard Rate Order ($X_1^* \leq^{hr} X_2^*$): This comparison holds if the hazard rate function of X_1^* is greater than or equal to that of X_2^* for all x. That is, $\Lambda_{X_1^*}(x) \geq \Lambda_{X_2^*}(x)$.
(3) Usual Stochastic Order ($X_1^* \leq^{st} X_2^*$): This relation holds when the survival function of X_1^* is less than or equal to that of X_2^*, i.e., $\overline{H}_1(x) \leq \overline{H}_2(x)$.
(4) Super-Additive Order ($X_1^* \leq^{su} X_2^*$): This order applies if the composition $H_2^{-1}(H_1(x))$ defines a super-additive function.
(5) Dispersive Order ($X_1^* \leq^{disp} X_2^*$): This ordering is satisfied if the difference $H_2^{-1}(H_1(x)) - x$ increases with x.

Notably, the hazard rate function for a random variable X_i^* is given by $\Lambda_{X_i^*}(v) = \frac{h_i(v)}{1-H_i(v)}$ for $v \geq 0$, where the survival function is denoted by $\overline{H}_i(v) = 1 - H_i(v)$ for $i = 1, 2$. For a comprehensive treatment of these stochastic orders and their properties, readers are encouraged to consult Shaked and Shanthikumar [18].

Kharazmi and Balakrishnan [6] explored the information-generating function for ordered random variables, specifically order statistics. In their study, they derived several properties of mixed systems built from independent and identically distributed components. Building on this foundation, we present a comparative analysis of mixed systems using these information metrics.

In a separate study on record values, Zamani et al. [8] investigated comparative outcomes linked to the information-generating (IG) measure. A key finding was that if two upper record value sequences share an identical IG function, the underlying distributions from which they originate must be the same. Their research also offers a rigorous characterization of the exponential distribution, demonstrating that its IG function for record values is either maximized or minimized under specific constraints.

This study aims to further explore the properties of the information-generating function for order statistics and to demonstrate its application in testing for symmetry. The remainder of the paper is structured as follows: Section 2 develops characterizations and examines monotonicity properties using ordered variables. Section 3 investigates stochastic ordering results based on the information-generating function of order statistics and establishes bounds for this measure. Section 4 analyzes the symmetric properties of the information-generating function model for order statistics, proposes a nonparametric test for symmetry, and illustrates the methodology using chronic disease management data.

2 Properties of information-generating function

In the following scenario, we will discuss some stochastic arrangements of the information-generating functional model for the entropy measure. Shaked and Shanthikumar's Theorem 4.B.2 [18] enables us to examine the following findings:

1. If $X_1^* \leq^{lr} X_2^*$, then $X_1^* \leq^{hr} X_2^*$ implies $X_1^* \leq^{st} X_2^*$.
2. If $X_1^* \leq^{st} X_2^*$, then $X_1^* \leq^{su} X_2^*$ implies $X_1^* \leq^{disp} X_2^*$.

Lemma 2.1. Assume that $X_1^* \leq^{disp} X_2^*$. Then the following inequality holds: $GEn_\delta(X_1^*) \geq (\leq)\, GEn_\delta(X_2^*)$ for $\delta \geq 1$ (respectively, $0 < \delta \leq 1$).

Proof. Starting from Equation 2, we express the information-generating functional entropy as:

$$GEn_\delta(X^*) = \int_{-\infty}^{\infty} [h(x)]^\delta\, dx = \int_0^1 \left[h\left(H^{-1}(v)\right)\right]^{\delta-1} dv.$$

Given that $X_1^* \leq^{disp} X_2^*$, it follows that $h_1(H_1^{-1}(v)) \geq h_2(H_2^{-1}(v))$ holds for every v in the interval $(0, 1)$. Consequently, we derive:

$$GEn_\delta(X_1^*) = \int_0^1 \left[h(H_1^{-1}(v))\right]^{\delta-1} dv \geq (\leq) \int_0^1 \left[h(H_2^{-1}(v))\right]^{\delta-1} dv = GEn_\delta(X_2^*),$$

which confirms the result for $\delta \geq 1$ (respectively, $0 < \delta \leq 1$).

2.1 Employing ordered variables, characterizations redesigned

With cdf H and pdf h, presume that the m occurrences $X_1^*, ..., X_m^*$ are independent and have the same distributions. Therefore, $X_{1,m}^* \leq X_{2,m}^* \leq\leq X_{m,m}^*$ are the order of statistics of the sample. The pdf of a random sample of size m, drawn from a distribution denoted by X^*, which includes the ith order statistic $X_{i,m}^*$ for $1 \leq i \leq m$, is expressed as:

$$h_{i,m}(x) = \frac{1}{\Delta_h(i, m-i+1)} H^{i-1}(x)\overline{H}^{m-i}(x)h(x), \tag{3}$$

where the normalizing constant is given by $\Delta_h(i, m-i+1) = \frac{\Gamma(i)\Gamma(m-i+1)}{\Gamma(m+1)}$. Therefore, from Equation 2, we can define the information-generating function measure for the ith order statistic $X_{i,m}^*$ as:

$$\begin{aligned} GEn_\delta(X_{i,m}^*) &= \int_{-\infty}^{\infty} h_{i,m}^\delta(x)\, dx \\ &= \left(\frac{1}{\Delta_h(i, m-i+1)}\right)^\delta \int_{-\infty}^{\infty} H^{\delta i-\delta}(x)\, \overline{H}^{\delta m-\delta i} (x)\, h^\delta(x)\, dx, \end{aligned} \tag{4}$$

for any $\delta > 0$, $1 \leq i \leq m$.

To support the main conclusions of this section, we refer to a corollary derived from the Stone–Weierstrass Theorem, as presented by Aliprantis and Burkinshaw [19]. This yields the following lemma:

Lemma 2.2. Let ζ^* be a continuous function on the interval $[0, 1]$. If it satisfies the integral condition $\int_0^1 z^m \zeta^*(z)\, dz = 0$ for all integers $m \geq 0$, then it follow that $\zeta^*(z) = 0$ for every $z \in [0, 1]$.

The next theorem shows that the characteristics of the information-generating function associated with the order statistic $X_{i,m}^*$ uniquely identify the distribution of the parent.

Theorem 2.1. Assume that h_1 and h_2 are two pdfs, with corresponding cdfs H_1 and H_2, for the random variables X_1^* and X_2^*, respectively. Fix a value of i, with $1 \leq i \leq m$, and let $\delta > 0$. Then the following equivalence holds:

$$X_1^* \overset{\mathrm{DI}}{=} X_2^* \iff GEn_\delta(X_{1;i,m}^*) = GEn_\delta(X_{2;i,m}^*), \forall m \geq i.$$

Proof. We only need to establish sufficiency, since necessity is immediate. Assume that

$$GEn_\delta(X_{1;i,m}^*) = GEn_\delta(X_{2;i,m}^*), \qquad \forall\, m \geq i.$$

Using Equations 2, 3, 4, this is equivalent to

$$\begin{aligned} &\int_{-\infty}^{\infty} H_1^{\delta i-\delta}(x)\, \overline{H}_1^{\delta m-\delta i}(x)\, h_1^\delta(x)\, dx \\ &= \int_{-\infty}^{\infty} H_2^{\delta i-\delta}(x)\, \overline{H}_2^{\delta m-\delta i}(x)\, h_2^\delta(x)\, dx. \end{aligned} \tag{5}$$

Step 1: Change of variables. Note that $d\overline{H}_k^\delta(x) = -\delta\, \overline{H}_k^{\delta-1}(x)h_k(x)\, dx$. Rewriting Equation 5 yields

$$\begin{aligned} &\int_{-\infty}^{\infty} H_1^{\delta i-\delta}(x)\, \overline{H}_1^{\delta m-\delta i}(x)\, \Lambda_{X_1^*}^{\delta-1}(x)\, d\overline{H}_1^\delta(x) \\ &= \int_{-\infty}^{\infty} H_2^{\delta i-\delta}(x)\, \overline{H}_2^{\delta m-\delta i}(x)\, \Lambda_{X_2^*}^{\delta-1}(x)\, d\overline{H}_2^\delta(x), \end{aligned}$$

where $\Lambda_{X_k^*}^{\delta-1}(x) = h_k^{\delta-1}(x)/\overline{H}_k^{\delta-1}(x)$.

Let

$$v = \overline{H}_k^\delta(x), \qquad k = 1, 2.$$

Since $\overline{H}_k$ is continuous and strictly decreasing, the mapping is bijective and sends $x \in (-\infty, \infty)$ to $v \in [0, 1]$. The identity becomes

$$\begin{aligned} &\int_0^1 (1 - v^{1/\delta})^{\delta i-\delta} v^{m-i} \Lambda_{X_1^*}^{\delta-1}\left(H_1^{-1}(1 - v^{1/\delta})\right) dv \\ &= \int_0^1 (1 - v^{1/\delta})^{\delta i-\delta} v^{m-i} \Lambda_{X_2^*}^{\delta-1}\left(H_2^{-1}(1 - v^{1/\delta})\right) dv. \end{aligned} \tag{6}$$

Step 2: Application of Lemma 2.2. Let

$$\zeta^*(v) = \Lambda_{X_1^*}^{\delta-1}\left(H_1^{-1}(1 - v^{1/\delta})\right) - \Lambda_{X_2^*}^{\delta-1}\left(H_2^{-1}(1 - v^{1/\delta})\right).$$

Equation 6 implies

$$\int_0^1 (1 - v^{1/\delta})^{\delta i-\delta} \zeta^*(v)\, v^l\, dv = 0, \qquad \forall\, l = m - i \ge 0.$$

The prefactor $(1 - v^{1/\delta})^{\delta i-\delta}$ is continuous and strictly positive for $v \in (0, 1)$; hence the above is equivalent to

$$\int_0^1 v^l\, \zeta^*(v)\, dv = 0, \qquad \forall\, l \ge 0.$$

Since ζ^* is continuous, Lemma 2.2 implies

$$\zeta^*(v) = 0, \qquad \forall\, v \in [0, 1].$$

Therefore,

$$\Lambda_{X_1^*}^{\delta-1}\left(H_1^{-1}(p)\right) = \Lambda_{X_2^*}^{\delta-1}\left(H_2^{-1}(p)\right), \qquad \forall\, p \in [0, 1]. \tag{7}$$

Step 3: Deduction of equality of the densities at corresponding quantiles. Recall that

$$\Lambda_{X_k^*}^{\delta-1}(x) = \frac{h_k^{\delta-1}(x)}{\overline{H}_k^{\delta-1}(x)}.$$

Since for the argument $x = H_k^{-1}(p)$ we have $\overline{H}_k(x) = 1 - p$, Equation 7 gives

$$h_1(H_1^{-1}(p)) = h_2(H_2^{-1}(p)), \qquad \forall\, p \in [0, 1].$$

Step 4: Equality of derivatives of inverse cdfs. Using the identity

$$h_k(H_k^{-1}(p)) = \frac{1}{(H_k^{-1})'(p)},$$

we obtain

$$(H_1^{-1})'(p) = (H_2^{-1})'(p), \qquad \forall\, p \in (0, 1).$$

Integrating over $[0, p]$ yields

$$H_1^{-1}(p) = H_2^{-1}(p) + C,$$

for some constant C.

Step 5: Determination of the constant. Both inverse cdfs satisfy

$$\lim_{p\to 0} H_k^{-1}(p) = \inf\{x : H_k(x) > 0\},$$

which is finite and equal for the two distributions, because equality of information-generating functions implies identical lower-support endpoints. Hence, the limit of the difference is zero, implying $C = 0$. Thus,

$$H_1^{-1}(p) = H_2^{-1}(p), \qquad \forall\, p \in [0, 1].$$

Therefore, $H_1 = H_2$, which completes the proof.

Remark 2.1. By taking $i = 1$ in Theorem 2.1, we have

$$X_1^* \stackrel{\text{DI}}{=} X_2^* \iff GEn_\delta(X_{1;1,m}^*) = GEn_\delta(X_{2;1,m}^*),\ \forall m \ge 1.$$

It is well established that the exponential distribution plays a significant role in reliability theory. In what follows, we present a novel characterization of this distribution.

Theorem 2.2. Let the exponential distribution be defined by $\bar{H}(x) = e^{-\theta x}$, where $\theta > 0$ and $x > 0$. This distribution is uniquely identified by the condition

$$GEn_\delta(X_{1,m}^*) = m^{\delta-1} GEn_\delta(X^*), \quad \forall m \ge 1.$$

With noting that $\delta > 0$.

Proof. We first verify the forward implication, then prove the converse.

(i) If X^* is exponential, then the IGF identity holds. If $\bar{H}(x) = e^{-\theta x}$ $(\theta > 0)$, a direct computation using Equations 2, 3 (the expression for GEn_δ of an order statistic and the definition of Λ_{X^*}) yields

$$GEn_\delta(X_{1,m}^*) = \frac{\theta^{\delta-1} m^{\delta-1}}{\delta} = m^{\delta-1}\left(\frac{\theta^{\delta-1}}{\delta}\right) = m^{\delta-1} GEn_\delta(X^*),$$

for every integer $m \ge 1$. Thus, the displayed identity holds for the exponential distribution.

(ii) Converse: the IGF identity implies an exponential parent.

Assume

$$GEn_\delta(X_{1,m}^*) = m^{\delta-1} GEn_\delta(X^*), \qquad \forall m \ge 1.$$

Using the integral representations in Equations 2, 3, this equality can be written as

$$\int_{-\infty}^{\infty} m^\delta\, \overline{H}^{\delta m-\delta}(x)\, h^\delta(x)\, dx = m^{\delta-1} \int_{-\infty}^{\infty} h^\delta(x)\, dx, \qquad \forall m \ge 1.$$

Bring all terms to one side and perform the change of variable

$$v = \overline{H}^\delta(x), \qquad v \in [0, 1].$$

As in the proof of Theorem 2.1, this substitution is admissible because $\overline{H}$ is continuous and monotone on the support, and it yields, for every integer $m \ge 1$,

$$\int_0^1 \left[\frac{1}{\delta}\, \Lambda_{X^*}^{\delta-1}\left(H^{-1}(1 - v^{1/\delta})\right) - GEn_\delta(X^*)\right] v^{m-1}\, dv = 0.$$

Define the continuous function on $[0, 1]$

$$\zeta(v) := \frac{1}{\delta}\, \Lambda_{X^*}^{\delta-1}\left(H^{-1}(1 - v^{1/\delta})\right) - GEn_\delta(X^*).$$

The previous displayed family of equalities says that $\int_0^1 \zeta(v)\, v^{m-1}\, dv = 0$ for every integer $m \ge 1$. Reindex by letting $l = m - 1$ (so $l \ge 0$) and apply Lemma 2.2; we conclude $\zeta(v) \equiv 0$ on $[0, 1]$. Hence

$$\Lambda_{X^*}^{\delta-1}\left(H^{-1}(1 - v^{1/\delta})\right) = \delta\, GEn_\delta(X^*) \qquad \text{for all } v \in [0, 1].$$

Equivalently, with $p := 1 - v^{1/\delta} \in [0, 1]$,

$$\Lambda_{X^*}\big(H^{-1}(p)\big) =: C \quad \text{for all } p \in [0, 1],$$

where $C := \big(\delta\, GEn_\delta(X^*)\big)^{1/(\delta-1)}$ is a positive constant. Thus the composed function $\Lambda_{X^*} \circ H^{-1}$ is constant on $[0, 1]$, and therefore

$$\Lambda_{X^*}(x) = C \qquad \text{for all } x \text{ in the (interior of the) support.}$$

(iii) From constant Λ_{X^*} to constant hazard (and hence exponential).

We now use the explicit relation between Λ_{X^*} and the parent density/hazard given in Equation 3 of the manuscript. *(Insert here the explicit formula for $\Lambda_{X^*}(x)$ from Equation 3.)* In the form needed below that formula expresses $\Lambda_{X^*}(x)$ as a continuously differentiable function of the hazard rate

$$\lambda(x) := \frac{h(x)}{\overline{H}(x)}.$$

Write this relation as

$$\Lambda_{X^*}(x) = \Phi\big(\lambda(x)\big),$$

where Φ is an explicit, continuously differentiable function (determined by Equation 3). The explicit algebra in the manuscript shows that Φ is one-to-one on $(0, \infty)$; hence, $\Lambda_{X^*}(x) = C$ for all x implies $\lambda(x) = \Phi^{-1}(C)$ for all x. Denote $\theta := \Phi^{-1}(C) > 0$. Therefore, the hazard is constant:

$$\lambda(x) = \theta, \qquad x \text{ in the support.}$$

A distribution with constant hazard $\lambda(x) \equiv \theta$ has survival function

$$\overline{H}(x) = \exp\Big(-\int_0^x \lambda(t)\, dt\Big) = \exp(-\theta x),$$

Thus, H is the exponential distribution with rate θ. Substituting $C = \big(\delta\, GEn_\delta(X^*)\big)^{1/(\delta-1)}$ and tracing back $\theta = \Phi^{-1}(C)$ yields the explicit relation between θ and $GEn_\delta(X^*)$ stated in the theorem. This completes the proof.

2.2 Monotonous characteristics

Ebrahimi et al. [20], Zamani et al. [8], and other related studies have reviewed the monotonic behavior of information measures for ordered variables. This section covers the monotonic characteristics of the information-generating function of ordered statistics of order δ.

Lemma 2.3. (Adapted from Shaked and Shanthikumar [18]) Let X^*_{i,m_1} and X^*_{j,m_2} be the ith and jth order statistics drawn from independent samples of sizes m_1 and m_2, respectively, drawn from a distribution H with a monotone non-increasing failure rate. Then,

$$X^*_{i,m_1} \leq^{disp} X^*_{j,m_2} \quad \text{whenever} \quad i \leq j \quad \text{and} \quad m_1 - i \geq m_2 - j.$$

An immediate consequence of Lemma 2.3 is that if $X^*_1, X^*_2, \ldots, X^*_m$ are independent and equally distributed observations from a monotone non-increasing failure rate distribution, then for any $i = 1, \ldots, m$, it holds that

$$X^*_{i,m+1} \leq^{\text{disp}} X^*_{i,m} \leq^{\text{disp}} X^*_{i+1,m+1}. \tag{8}$$

Utilizing this result alongside Lemma 2.3, we can now establish the following theorem.

Theorem 2.3. (1) Suppose X^* follows a distribution with a monotone non-increasing failure rate. Then for a fixed index i satisfying $1 \leq i \leq m$, the generalized entropy $GEn_\delta(X^*_{i,m})$ increases with m.
(2) Under the same distributional assumption, for a fixed sample size m with $m \geq i \geq 1$, $GEn_\delta(X^*_{i,m})$ decreases as i increases.
With noting that $\delta \in \mathbb{N}^+$.

Proof. The proof follows from Lemma 2.3 and the Equation 8.

Let us recall that a random variable X^* is said to have an increasing reversed hazard rate if the function $\tilde{\Lambda}_{X^*}(x) = h(x)/H(x)$ is non-decreasing in x. Under this alternative assumption, we now present the reversed implications of Theorem 2.3.

Theorem 2.4. (1) If X^* has an increasing reversed hazard rate, then for a fixed i within $1 \leq i \leq m$, the quantity $GEn_\delta(X^*_{i,m})$ decreases with increasing m.
(2) Under the same condition, if m is fixed and $m \geq i \geq 1$, then $GEn_\delta(X^*_{i,m})$ increases as i becomes larger.
With noting that $\delta \in \mathbb{N}^+$.

Proof. According to Equations 2, 3, we have

$$\begin{aligned}\frac{GEn_\delta(X^*_{i,m})}{GEn_\delta(X^*_{i,m+1})} &= \left(\frac{m-i+1}{m+1}\right)^\delta \frac{\int_{-\infty}^{\infty} H^{\delta i-\delta}(x)\overline{H}^{\delta m-\delta i}(x)h^\delta(y)\, dx}{\int_{-\infty}^{\infty} H^{\delta i-\delta}(x)\overline{H}^{\delta m+\delta-\delta i}(x)h^\delta(y)\, dx} \\ &= J(m; i; \delta)\frac{\int_0^1 \frac{1}{\Delta_h(\delta i, \delta m-\delta i+1)} v^{\delta i-1}(1-v)^{\delta m-\delta i}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\, dv}{\int_0^1 \frac{1}{\Delta_h(\delta i, \delta m-\delta i+\delta+1)} v^{\delta i-1}(1-v)^{\delta m-\delta i+\delta}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\, dv},\end{aligned} \tag{9}$$

where

$$J(m; i; \delta) = \left(\frac{m-i+1}{m+1}\right)^\delta \cdot \frac{\Gamma(\delta(m-i)+1)\Gamma(\delta(m+1)+1)}{\Gamma(\delta m+1)\Gamma(\delta(m-i+1)+1)}.$$

We introduce $t = m - i$ to simplify the notation $J(m; i; \delta)$. The gamma function terms can be rewritten using the property of the gamma function for shifted arguments:

$$\frac{\Gamma(\delta m+\delta+1)}{\Gamma(\delta m+1)} \cdot \frac{\Gamma(\delta t+1)}{\Gamma(\delta t+\delta+1)}$$

This can be expressed as a product of terms:

$$(\delta m+1)(\delta m+2)\cdots(\delta m+\delta) \cdot \frac{1}{(\delta t+1)(\delta t+2)\cdots(\delta t+\delta)}$$

Combining these products with the initial term $\left(\frac{m-i+1}{m+1}\right)^{\delta}$, we get a product over k from 1 to δ:

$$J(m;\, i;\, \delta) = \prod_{k=1}^{\delta}\left(\frac{(m-i+1)(\delta m+k)}{(m+1)(\delta(m-i)+k)}\right), \tag{10}$$

where $\delta \in \mathbb{N}^{+}$, $1 \leq i \leq m$. Substituting from Equation 10 in Equation 9, we obtain

$$\begin{aligned}\frac{GEn_{\delta}(X^{*}_{i,m})}{GEn_{\delta}(X^{*}_{i,m+1})} &= J(m;\, i;\, \delta)\frac{\int_0^1 \frac{1}{\Delta_h(\delta i,\delta m-\delta i+1)} v^{\delta i-1}(1-v)^{\delta m-\delta i}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\,dv}{\int_0^1 \frac{1}{\Delta_h(\delta i,\delta m-\delta i+\delta+1)} v^{\delta i-1}(1-v)^{\delta m-\delta i+\delta}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\,dv}\\ &\geq \frac{\mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i,\delta m}))\right]}{\mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i,\delta m+\delta}))\right]},\end{aligned} \tag{11}$$

where $W^{*}_{i,m}$ represent the ith order statistic derived from a sample of size m drawn from a uniform distribution. The corresponding pdf is given by $h_{i,m}(w) = \frac{1}{\Delta_h(i,m-i+1)} w^{i-1}(1-w)^{m-i}$ for $w \in [0,1]$, and $i = 1, 2, \ldots, m$. Shaked and Shanthikumar [18] state that Theorem 1.B.28 states that $W^{*}_{\delta i,\delta m} \geq^{hr} W^{*}_{\delta i,\delta m+\delta}$. This means that $W^{*}_{\delta i,\delta m} \geq^{st} W^{*}_{\delta i,\delta m+\delta}$ is also implied. Given that $\delta \in \mathbb{N}^{+}$, the assumption leads to the inequality:

$$\mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i,\delta m}))\right] \geq \mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i,\delta m+\delta}))\right],$$

which in turn implies that $\frac{GEn_{\delta}(X^{*}_{i,m})}{GEn_{\delta}(X^{*}_{i,m+1})} \geq 1$. Similarly, for Part (2), we have

$$\begin{aligned}\frac{GEn_{\delta}(X^{*}_{i,m})}{GEn_{\delta}(X^{*}_{i+1,m})} &= J^{*}(m;\, i;\, \delta)\frac{\int_0^1 \frac{1}{\Delta_h(\delta i,\delta m-\delta i+1)} v^{\delta i-1}(1-v)^{\delta m-\delta i}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\,dv}{\int_0^1 \frac{1}{\Delta_h(\delta i+\delta,\delta m-\delta i-\delta+1)} v^{\delta i+\delta-1}(1-v)^{\delta m-\delta i-\delta}\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(v))\,dv}\\ &\leq \frac{\mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i,\delta m}))\right]}{\mathbb{E}\left[\tilde{\Lambda}^{\delta-1}_{X^*}(H^{-1}(W^{*}_{\delta i+\delta,\delta m}))\right]},\end{aligned} \tag{12}$$

where

$$J^{*}(m;\, i;\, \delta) = \prod_{k=1}^{\delta}\left(\frac{i(\delta(m-i)+k)}{(m+i)(\delta i+k)}\right),$$

where $\delta \in \mathbb{N}^{+}$, $m \geq i \geq 1$. Thus, $\frac{GEn_{\delta}(X^{*}_{i,m})}{GEn_{\delta}(X^{*}_{i+1,m})} \leq 1$, with noting that $W^{*}_{\delta i,\delta m} \leq^{st} W^{*}_{\delta i+\delta,\delta m}$.

Theorem 2.5. (1) If X^{*} has a decreasing reversed hazard rate, then for a fixed i within $1 \leq i \leq m$, the quantity $GEn_{\delta}(X^{*}_{i,m})$ increases with increasing m.

(2) Under the same condition, if m is fixed and $m \geq i \geq 1$, then $GEn_{\delta}(X^{*}_{i,m})$ decreases as i becomes larger.

With noting that $\delta \in \mathbb{N}^{+}$.

Proof. The steps are similar to those in the proof of the previous theorem.

Recalling the Pareto distribution's diminishing reversed hazard rate, represented by the CDF $1 - x^{-\alpha}$, $x \geq 1$, and $\alpha > 0$. With rising m and increasing i, respectively, for the Pareto distribution and $\delta = 2, 3$, Figures 1, 2 illustrate the information-generating function model of $X^{*}_{i,m}$, which guarantees the monotonous qualities of Theorem 2.5 when $\delta \in \mathbb{N}^{+}$.

3 Ordering outcomes based on the information-generating function of order statistics

In this section, we present some stochastic comparison results for the information-generating function measure of order statistics. The information-generating function of order statistics can be rewritten as follows lemma.

Lemma 3.1. The information-generating function measure of the ith order statistics, $X^{*}_{i,m}$, can be written as

$$GEn_{\delta}(X^{*}_{i,m}) = \frac{\Delta_h(\delta i-\delta+1,\delta m-\delta+1)}{(\Delta_h(i,m-i+1))^{\delta}}\mathbb{E}\left[h^{\delta-1}\left(H^{-1}(V^{*})\right)\right], \tag{13}$$

where the random variable V^{*} has the pdf

$$h_{V^*}(v) = \frac{1}{\Delta_h(\delta i-\delta+1,\delta m-\delta+1)} v^{\delta i-\delta}(1-v)^{\delta m-\delta i}, \tag{14}$$

$v \in [0, 1]$.

Proof. From Equations 2, 3, and making use of the transformation $v = H(x)$, we can express the information-generating function measure of the ith order statistics, $X^{*}_{i,m}$, as

$$GEn_{\delta}(X^{*}_{i,m}) = \frac{\Delta_h(\delta i-\delta+1,\delta m-\delta+1)}{\Delta_h(\delta i-\delta+1,\delta m-\delta+1)(\Delta_h(i,m-i+1))^{\delta}} \int_0^1 v^{\delta i-\delta}(1-v)^{\delta m-\delta i}h^{\delta-1}(H^{-1}(V^{*}))\,dv,$$

and the result follows.

The impact of monotonic transformations on the information-generating function measure of order statistics is examined in the following theorem.

Theorem 3.1. Assume that φ is a strictly increasing function satisfying $\varphi(-\infty) = 0$ and $\varphi(\infty) = \infty$. Then, for the ith order statistic of the transformed random variable $Y^{*} = \varphi(X^{*})$, the information-generating function measure is expressed as

$$GEn_{\delta}(Y^{*}_{i,m}) = \frac{\Delta_h(\delta i-\delta+1,\delta m-\delta+1)}{(\Delta_h(i,m-i+1))^{\delta}}\mathbb{E}\left[\frac{h(H^{-1}(V^{*}))}{\varphi'(H^{-1}(V^{*}))}\right]^{\delta-1},$$

where V^{*} denotes a random variable whose pdf is defined in Equation 14.

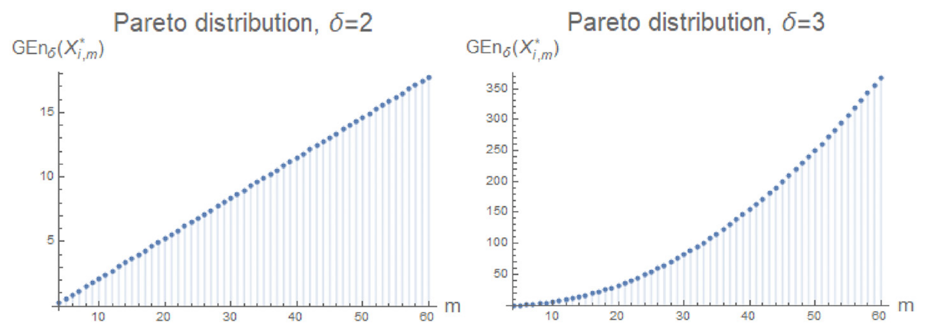

FIGURE 1
Information-generating function of $X^*_{4,m}$ for the Pareto distribution (with parameter $\alpha = 2$), with increasing m and $\delta = 2, 3$.

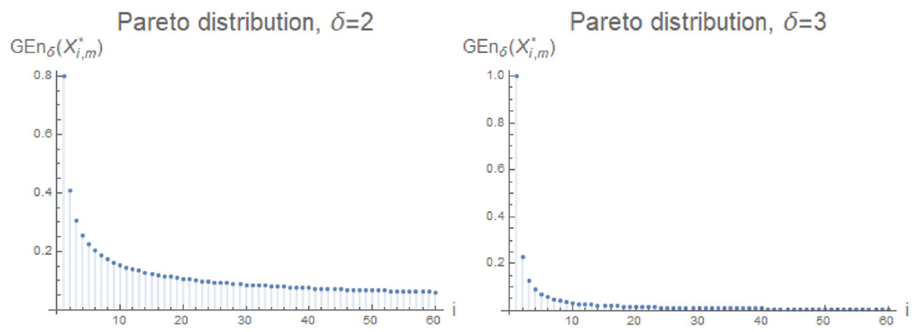

FIGURE 2
Information-generating function of $X^*_{i,60}$ for a Pareto distribution (with parameter $\alpha = 2$), with increasing i and $\delta = 2, 3$.

Proof. Given the transformation $Y^* = \varphi(X^*)$, the cdf and pdf of Y^* become $F(y) = H(\varphi^{-1}(y))$ and $f(y) = \frac{h(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))}$, respectively. Using the definition of the information-generating function for the ith order statistic, along with the substitutions $x = \varphi^{-1}(y)$ and $v = H(x)$, we derive

$$GEn_\delta(Y^*_{i,m}) = \frac{1}{(\Delta_h(i, m-i+1))^\delta} \int_{-\infty}^{\infty} H^{\delta i-\delta}(\varphi^{-1}(y)) \bar{H}^{\delta m-\delta i}(\varphi^{-1}(y)) \left[\frac{h(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))}\right]^\delta dy.$$

Next, using the change of variables $x = \varphi^{-1}(y)$, we obtain

$$GEn_\delta(Y^*_{i,m}) = \frac{1}{(\Delta_h(i, m-i+1))^\delta} \int_{-\infty}^{\infty} H^{\delta i-\delta}(x) \bar{H}^{\delta m-\delta i}(x) \left[\frac{h^\delta(x)}{(\varphi'(x))^{\delta-1}}\right] dx$$
$$= \frac{\Delta_h(\delta i-\delta+1, \delta m-\delta+1)}{(\Delta_h(i, m-i+1))^\delta} \mathbb{E}\left[\frac{h(H^{-1}(V^*))}{\varphi'(H^{-1}(V^*))}\right]^{\delta-1}.$$

Theorem 3.2. Let X^* be a random variable with pdf h, and let φ be a strictly increasing and convex function satisfying $\varphi(0) = 0$ and $\varphi(x) \to \infty$ as $x \to \infty$. Assume further that $\varphi'(x)$ exists, is non-decreasing, and fulfills the condition $\varphi'(0) \geq 1$. Then:

(1) If $\delta \geq 1$, then

$$GEn_\delta(\varphi(X^*_{i,m})) \leq GEn_\delta(X^*_{i,m}).$$

(2) If $0 < \delta \leq 1$, then

$$GEn_\delta(\varphi(X^*_{i,m})) \geq GEn_\delta(X^*_{i,m}).$$

Proof. Since φ is convex and strictly increasing, its derivative $\varphi'(x)$ is non-decreasing and satisfies

$$\varphi'(x) \geq \varphi'(0) \geq 1, \qquad \forall x \geq 0.$$

Let $Y = \varphi(X^*)$. By a standard change-of-variable argument, the pdf of Y is given by

$$h_Y(\varphi(x)) = \frac{h(x)}{\varphi'(x)} \leq h(x),$$

because $\varphi'(x) \geq 1$.

Equation 2.3 gives the IGF representation

$$GEn_\delta(X^*) = \mathbb{E}\Big[\Lambda_{X^*}^{\delta-1}(H^{-1}(V^*))\Big],$$

and therefore

$$\Lambda_Y(\varphi(x)) = \frac{h(x)}{\varphi'(x)} \leq h(x) = \Lambda_{X^*}(x).$$

When $\delta \geq 1$, the function $u \mapsto u^{\delta-1}$ is increasing, which implies

$$\Lambda_Y^{\delta-1}(\varphi(x)) \leq \Lambda_{X^*}^{\delta-1}(x).$$

For $0 < \delta \leq 1$, the same function is decreasing, hence

$$\Lambda_Y^{\delta-1}(\varphi(x)) \geq \Lambda_{X^*}^{\delta-1}(x).$$

Lemma 3.1 together with Theorem 3.1 ensures that these inequalities carry over to the IGF evaluated at the order statistic $X^*_{i,m}$. Consequently:
- If $\delta \geq 1$, then

$$GEn_\delta(\varphi(X^*_{i,m})) \leq GEn_\delta(X^*_{i,m}).$$

- If $0 < \delta \leq 1$, then

$$GEn_\delta(\varphi(X^*_{i,m})) \geq GEn_\delta(X^*_{i,m}).$$

This completes the proof.

Remark 3.1. The additional requirement $\varphi'(0) \geq 1$ is not intended to restrict the class of admissible convex transformations. Its role is to ensure that the map φ does not locally contract the distribution near the origin. Since the IGF involves powers of the hazard function; such a contraction would reverse the direction of the inequalities in Theorem 3.2. The condition $\varphi'(0) \geq 1$ is therefore a convenient and sufficient way to guarantee that

$$\Lambda_{\varphi(X^*)}(\varphi(x)) = \frac{h(x)}{\varphi'(x)} \leq h(x) = \Lambda_{X^*}(x),$$

which is the key step in applying Lemma 3.1 and Theorem 3.1. We note that this assumption may be relaxed to $\varphi'(x) \geq 1$ on a neighborhood of the origin, without altering the main results. In this sense, the condition is mild and does not significantly reduce the applicability of the theorem.

The information-generating function measurements associated with the ith order statistics of two continuously generated random variables are compared as follows. Theorem 3.B.26 by Shaked and Shanthikumar [18] states that $X^*_{1;i,m} \leq^{disp} X^*_{2;i,m}$, if $X^*_1 \leq^{disp} X^*_2$, where $i = 1, 2, ..., m$. Therefore, using Lemma 2.1, we can easily get the following conclusion.

Proposition 3.1. Assume that $X^*_1 \leq^{disp} X^*_2$. Then, it holds that $GEn_\delta(X^*_{1;i,m}) \geq (\leq) GEn_\delta(X^*_{2;i,m})$ for $\delta \geq 1$ (respectively, $0 < \delta \leq 1$).

Proof. From Lemma 2.1 and Equation 13, let $X^*_1 \leq^{disp} X^*_2$. Then, for any $\delta \geq 1$ $(0 < \delta \leq 1)$, we have

$$\mathbb{E}\left[h_1^{\delta-1}\left(H_1^{-1}(V^*)\right)\right] \geq (\leq) \mathbb{E}\left[h_2^{\delta-1}\left(H_2^{-1}(V^*)\right)\right],$$

and the result follows.

The following theorem compares the information-generating functions of related ith order statistics by measuring the information-generating functions of two variables.

Theorem 3.3. Consider two continuous random variables, X^*_1 and X^*_2, associated with cdfs H_1 and H_2, and corresponding pdfs h_1 and h_2. Suppose that the condition $\inf \Psi^*_1 \geq \sup \Psi^*_2$ holds, where

$$\Psi^*_1 = \left\{ v^* > 0 \;\middle|\; \frac{h_2(H_2^{-1}(v^*))}{h_1(H_1^{-1}(v^*))} \leq 1 \right\},$$
$$\Psi^*_2 = \left\{ v^* > 0 \;\middle|\; \frac{h_2(H_2^{-1}(v^*))}{h_1(H_1^{-1}(v^*))} > 1 \right\}.$$

Then, the following statements are true:

(1) If $0 < \delta \leq 1$ and $GEn_\delta(X^*_1) \leq GEn_\delta(X^*_2)$, then it follows that $GEn_\delta(X^*_{1,m}) \leq GEn_\delta(X^*_{2,m})$.
(2) If $\delta \geq 1$ and $GEn_\delta(X^*_1) \geq GEn_\delta(X^*_2)$, then it follows that $GEn_\delta(X^*_{1,m}) \geq GEn_\delta(X^*_{2,m})$.

Proof. When either of the sets Ψ^*_1 or Ψ^*_2 is empty, the conclusion holds trivially. Therefore, we assume both sets are non-empty. Given the assumption that $GEn_\delta(X^*_1) \leq GEn_\delta(X^*_2)$, we can write

$$\int_{-\infty}^{\infty} h_1^\delta(x)\,dx - \int_{-\infty}^{\infty} h_2^\delta(x)\,dx = \int_0^1 \left[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\right] dv \leq 0.$$

Since $0 \leq (1-v) \leq 1$ for $v \in [0,1]$, it follows that

$$\int_0^1 (1-v)^{\delta(m-i)} \left[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\right] dv \leq 0, \quad (15)$$

where $m - i \geq 0$ for $i = 1, 2, \ldots, m$.

Now, applying the definition of the information-generating function of the i-th order statistic, we obtain

$$GEn_\delta(X^*_{1;i,m}) - GEn_\delta(X^*_{2;i,m}) = \int_{-\infty}^{\infty} h^\delta_{1;i,m}(x)\,dx - \int_{-\infty}^{\infty} h^\delta_{2;i,m}(x)\,dx = \Omega(x),$$

where we define

$$\Omega(x) = \int_{-\infty}^{\infty} h^\delta_{1;i,m}(x)\,dx - \int_{-\infty}^{\infty} h^\delta_{2;i,m}(x)\,dx.$$

To verify the first part of the theorem, it suffices to show that $\Omega(x) \leq 0$. Using the substitution $v = H_i(x)$ for $i = 1, 2$, we rewrite $\Omega(x)$ as follows:

$$\begin{aligned}(\Delta_h(i, m-i+1))^\delta \Omega(x) &= \int_0^1 v^{\delta i-\delta}(1-v)^{\delta m-\delta i}\Big[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\Big]\,dv \\ &= \int_{\Psi^*_1} v^{\delta i-\delta}(1-v)^{\delta m-\delta i}\Big[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\Big]\,dv \\ &\quad + \int_{\Psi^*_2} v^{\delta i-\delta}(1-v)^{\delta m-\delta i}\Big[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\Big]\,dv.\end{aligned}$$

From the given condition $\inf \Psi_1^* \geq \sup \Psi_2^*$ and the boundedness of $v^{\delta(i-1)}$ on $[0,1]$, we obtain:

$$(\Delta_h(i,m-i+1))^\delta \Omega(x) \leq (\inf \Psi_1^*)^{\delta(i-1)} \int_{\Psi_1^*} (1-v)^{\delta(m-i)} \left[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\right] dv$$
$$+ (\sup \Psi_2^*)^{\delta(i-1)} \int_{\Psi_2^*} (1-v)^{\delta(m-i)} \left[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\right] dv$$
$$\leq (\inf \Psi_1^*)^{\delta(i-1)} \int_0^1 (1-v)^{\delta(m-i)} \left[h_1^{\delta-1}(H_1^{-1}(v)) - h_2^{\delta-1}(H_2^{-1}(v))\right] dv \leq 0.$$

The last inequality follows directly from Equation 15 and the assumption that $\inf \Psi_1^* \geq \sup \Psi_2^*$. A similar argument can be applied to prove the second part.

3.1 Bounds for information-generating function measure of order statistics

Theorem 3.4. Let X^* be a random variable with cdf H and pdf h. If $M_d^* = f(m) < \infty$, where $m_d^* = \sup\{x : h(x) \leq M_d^*\}$ is the mode of X^*, then

$$GEn_\delta(X_{i,m}^*) \leq \max\left\{ \frac{(M_d^*)^{\delta-1}}{(\Delta_h(i,m-i+1))^\delta} D^*(\delta i;\, \delta m), \frac{(i-1)^{\delta i-\delta}(m-i)^{\delta m-\delta i}}{(m-1)^{\delta m-\delta}(\Delta_h(i,m-i+1))^\delta} GEn_\delta(X^*) \right\} \tag{16}$$

where $D^*(\delta i;\, \delta m) = \int_0^1 v^{\delta i-\delta}(1-v)^{\delta m-\delta i}\, dv$, and under the condition $\delta \geq 1$.

Proof. From Equations 2, 3, and under the condition $\delta \geq 1$, we can use the transformation $v = H(x)$ to express the information-generating function measure for the ith order statistics, $X_{i,m}^*$, as

$$GEn_\delta(X_{i,m}^*) = \frac{1}{(\Delta_h(i,m-i+1))^\delta} \int_0^1 v^{\delta i-\delta}(1-v)^{\delta m-\delta i} h^{\delta-1} (H^{-1}(v))\, dv.$$

Given $h(x) \leq M_d^*$, it follows that

$$GEn_\delta(X_{i,m}^*) \leq \frac{(M_d^*)^{\delta-1}}{(\Delta_h(i,m-i+1))^\delta} D^*(\delta i;\, \delta m).$$

Conversely, since the beta distribution with pdf $\frac{1}{\Delta_h(i,m-i+1)} \int_0^1 v^{i-1}(1-v)^{m-i}\, dv$ has the mode $\frac{i-1}{m-1}$, we can say that

$$GEn_\delta(X_{i,m}^*) \leq \frac{1}{(\Delta_h(i,m-i+1))^\delta} \left(\frac{i-1}{m-1}\right)^{\delta i-\delta} \left(1-\frac{i-1}{m-1}\right)^{\delta m-\delta i} \int_0^1 h^{\delta-1}(H^{-1}(v))\, dv$$
$$= \frac{(i-1)^{\delta i-\delta}(m-i)^{\delta m-\delta i}}{(m-1)^{\delta m-\delta}(\Delta_h(i,m-i+1))^\delta} GEn_\delta(X^*).$$

Example 3.1. Suppose X^* follows a Pareto distribution with pdf given by $h(x) = \frac{\alpha s^\alpha}{x^{\alpha+1}}$, $x \geq s > 0$, $\alpha > 0$. It can be shown that the transformed density becomes $h(H^{-1}(v)) = \frac{\alpha}{s}(1-v)^{\alpha+\frac{1}{\alpha}}$, $0 < v < 1$. Taking $\alpha = 1$ and $s = 1$, we find $M_d^* = 1$, and hence, $h(H^{-1}(v)) = (1-v)^2$. Furthermore, we compute

$$GEn_\delta(X^*) = \int_1^\infty x^{-2\delta}\, dx = \frac{1}{2\delta-1}, \quad \delta > \frac{1}{2}.$$

According to Theorem 3.4, we deduce that

$$GEn_\delta(X_{i,m}^*) \leq \max\left\{ \frac{D^*(\delta i;\, \delta m)}{(\Delta_h(i,m-i+1))^\delta}, \frac{(i-1)^{\delta i-\delta}(m-i)^{\delta m-\delta i}}{(2\delta-1)(m-1)^{\delta m-\delta}(\Delta_h(i,m-i+1))^\delta} \right\}.$$

Letting $m = 20$ and $i = 15$, for $\delta = 3$, we evaluate

$$GEn_3(X_{15,20}^*) \leq \max\{9.83131, 13.6021\} = 13.6021.$$

4 Information-generating function model symmetric features of the order statistics

A number of interesting features of the information-generating function of order statistics appear when the pdf of the underlying system, aside from the independent distributed random variables, is symmetric. We begin with two lemmas, the proof of which follows immediately from the symmetry assumption and the definition of $h_{i,m}$ in Equation 3.

Lemma 4.1. (Fashandi and Ahmadi [21]) Let X^* be a continual random variable defined over the support $S_{X^*}^*$, with pdf h and cdf H. If the following condition holds:

$$h\left(H^{-1}(v)\right) = h\left(H^{-1}(1-v)\right), \quad \text{for all } v \in (0,1),$$

then the cdf $H(x)$ is symmetric with respect to some point $c^* \in S_{X^*}^*$.

Lemma 4.2. (Balakrishnan and Selvitella [13]) Suppose the order statistic $X_{i,m}^*$, for $i = 1,\ldots,m$, arises from a distribution whose pdf h satisfies the symmetry condition $h(\mu^*+x) = h(\mu^*-x)$ for $x \geq 0$, where μ^* denotes the mean of X^*. Under this assumption, the following identities are satisfied:

$$H(\mu^*+x) = \overline{H}(\mu^*-x), \quad h_{i,m}(\mu^*+x) = h_{m-i+1,m}(\mu^*-x).$$

Theorem 4.1. Assume that $X_1^*,\ldots,X_m^*$ are iid samples drawn from a distribution with pdf h that is symmetric about its mean μ^*. Then, the following properties hold:

1. If the sample size m is odd, then for every $i = 1, \ldots, m$,

$$GEn_\delta(X^*_{i,m}) = GEn_\delta(X^*_{m-i+1,m}).$$

2. The pdf h is symmetric (about some point) if and only if

$$GEn_\delta(X^*_{1,m}) = GEn_\delta(X^*_{m,m}) \quad \text{for all integers } m \geq 1.$$

Moreover, if the first moment exists, the center of symmetry equals the mean μ^*.

Proof. (1) *(Symmetry implies equality of GEn for reflected order statistics).* By Lemma 4.2, we have the pointwise identity

$$h_{i,m}(\mu^* + x) = h_{m-i+1,m}(\mu^* - x), \qquad x \in \mathbb{R}.$$

Using this identity and the substitution $y = \mu^* + x$ (whose Jacobian is $dy = dx$), we obtain

$$\begin{aligned} GEn_\delta(X^*_{i,m}) &= \int_{-\infty}^{\infty} h^\delta_{i,m}(y)\, dy = \int_{-\infty}^{\infty} h^\delta_{i,m}(\mu^* + x)\, dx \\ &= \int_{-\infty}^{\infty} \left(h_{m-i+1,m}(\mu^* - x)\right)^\delta dx \\ &= \int_{-\infty}^{\infty} h^\delta_{m-i+1,m}(t)\, dt = GEn_\delta(X^*_{m-i+1,m}), \end{aligned}$$

where in the penultimate equality we used the change of variable $t = \mu^* - x$. This proves (1).

(2) *(Necessity).* Part (1) with $i = 1$ gives immediately $GEn_\delta(X^*_{1,m}) = GEn_\delta(X^*_{m,m})$ for all odd m. Because the identity for all $m \geq 1$ is stronger, necessity is immediate.

(Sufficiency). Assume that

$$GEn_\delta(X^*_{1,m}) = GEn_\delta(X^*_{m,m}) \qquad \text{for every } m \geq 1. \tag{17}$$

Using the representations of GEn_δ and proceeding exactly as in the proof of Theorem 2.1, the Equation 17 yields, after the standard change of variable $v = \overline{H}^\delta(x)$ and grouping factors, an identity of the form

$$\int_0^1 w(v)\left[h(H^{-1}(v)) - h(H^{-1}(1-v))\right] v^\ell\, dv = 0 \qquad \text{for all } \ell \geq 0,$$

where $w(v) = (1 - v^{1/\delta})^{\delta i - \delta}$ is continuous and strictly positive on $(0, 1)$. Dividing by $w(v)$ and using the continuity of the integrand, we obtain

$$\int_0^1 v^\ell \left[h(H^{-1}(v)) - h(H^{-1}(1-v))\right] dv = 0 \qquad \text{for all } \ell \geq 0.$$

By Lemma 2.2 (Stone–Weierstrass corollary), the continuous function

$$\zeta(v) := h(H^{-1}(v)) - h(H^{-1}(1-v))$$

must vanish identically on $[0, 1]$; hence

$$h(H^{-1}(v)) = h(H^{-1}(1-v)), \qquad \forall v \in (0, 1). \tag{18}$$

Now Lemma 4.1 (Fashandi and Ahmadi) implies that the cdf H is symmetric about some point $c^* \in \mathbb{R}$ (that is, $H(c^* + x) = 1 - H(c^* - x)$ for all x). Consequently h is symmetric about c^*.

To identify the center c^* with the mean μ^*, note that for any distribution symmetric about c^* with a finite first moment, we necessarily have

$$\mathbb{E}[X] = c^*.$$

Therefore, when the first moment exists, the center of symmetry equals the mean, and the pdf is symmetric about μ^*. This completes the proof of sufficiency and hence of the theorem.

Corollary 4.1. As a direct consequence of Theorem 4.1, let the forward difference operator with respect to i be defined as $\Xi GEn_\delta(X^*_{i,m}) = GEn_\delta(X^*_{i+1,m}) - GEn_\delta(X^*_{i,m})$ for $1 \leq i \leq m-1$. Then, it follows that $\Xi GEn_\delta(X^*_{i,m}) = -\Xi GEn_\delta(X^*_{m-i,m})$ for $i = 1, \ldots, m-1$.

Remark 4.1. Define $\Theta_m = GEn_\delta(X^*_{1,m}) - GEn_\delta(X^*_{m,m})$. Then, $\Theta_m = 0$ if and only if X^* is symmetric. Hence, Θ_m serves as a potential measure of symmetry and can be used as a test statistic for assessing symmetry.

Based on the conditions outlined in Corollary 4.1, the information-generating function $GEn_\delta(X^*_{i,m})$ attains either a local maximum or a minimum at the median position. This behavior can be illustrated using the uniform $U(-1, 1)$ and standard normal $N(0, 1)$ distributions. Specifically, for the median case ($i = 4$) when the sample size is $m = 7$, we can observe (refer to Figures 3, 4):

(1) Under the $U(-1, 1)$ distribution, the function reaches a minimum value of 3.263403 for $\delta = 2$, 5.940808 for $\delta = 3$, and 11.36502 for $\delta = 4$.
(2) Under the $N(0, 1)$ distribution the function reaches a maximum value of 0.6147224 for $\delta = 2$, 0.43858655 for $\delta = 3$, and 0.33141763 for $\delta = 4$.

4.1 Symmetry test using nonparametric estimation

Nonparametric approaches to testing symmetry have been extensively explored in the literature; notable contributions include those by Xiong et al. [22], Noughabi and Jarrahiferiz [23], and Mohamed and Almuqrin [24]. In this section, we focus on a nonparametric estimation framework for the information-generating function inspired by the methodology proposed by Vasicek [25]. This formulation is then employed to assess symmetry in a distribution. Consider a random sample $X^*_1, \ldots, X^*_m$ drawn from a continuous distribution $H(x)$ with associated density function $h(x)$. The hypothesis under investigation is:

$$Hy_0 : H(\mu^* - x) = 1 - H(\mu^* + x), \quad \text{for all } x.$$

where the parameter μ^* is unspecified. The alternative hypothesis is expressed as:

$$Hy_1 : H(\mu^* - x) \neq 1 - H(\mu^* + x).$$

When the underlying random variables are equally distributed and independent and have a symmetric

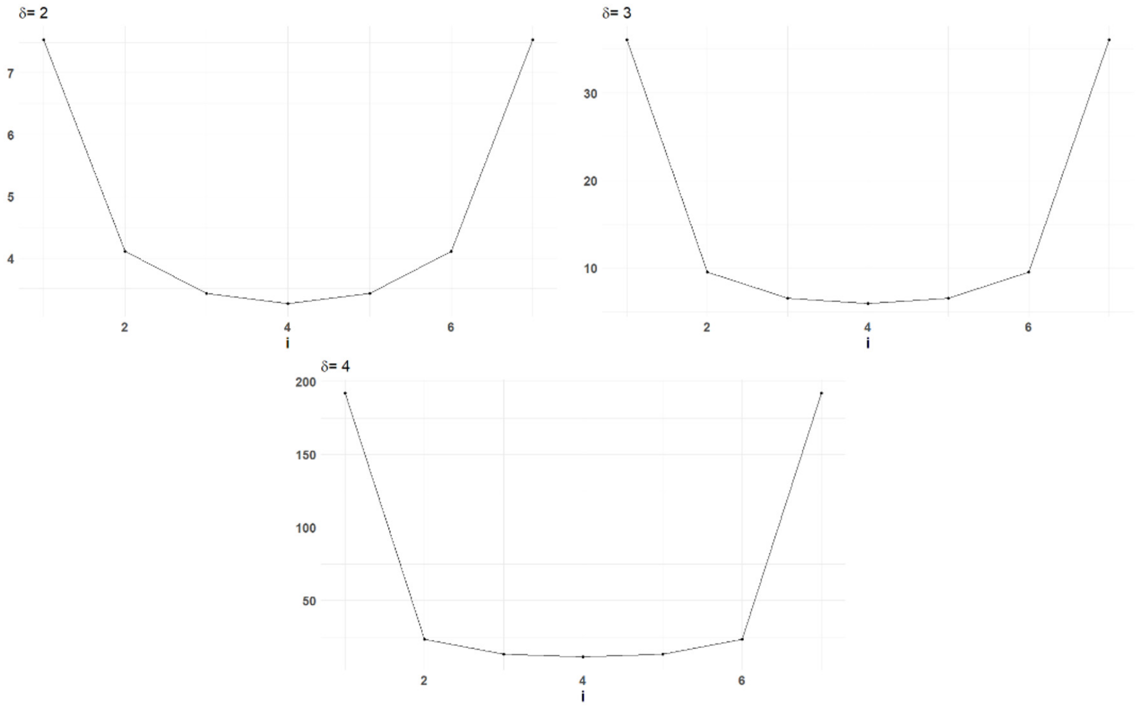

FIGURE 3
Information-generating function of the ith order statistics of $U(-1, 1)$ distribution.

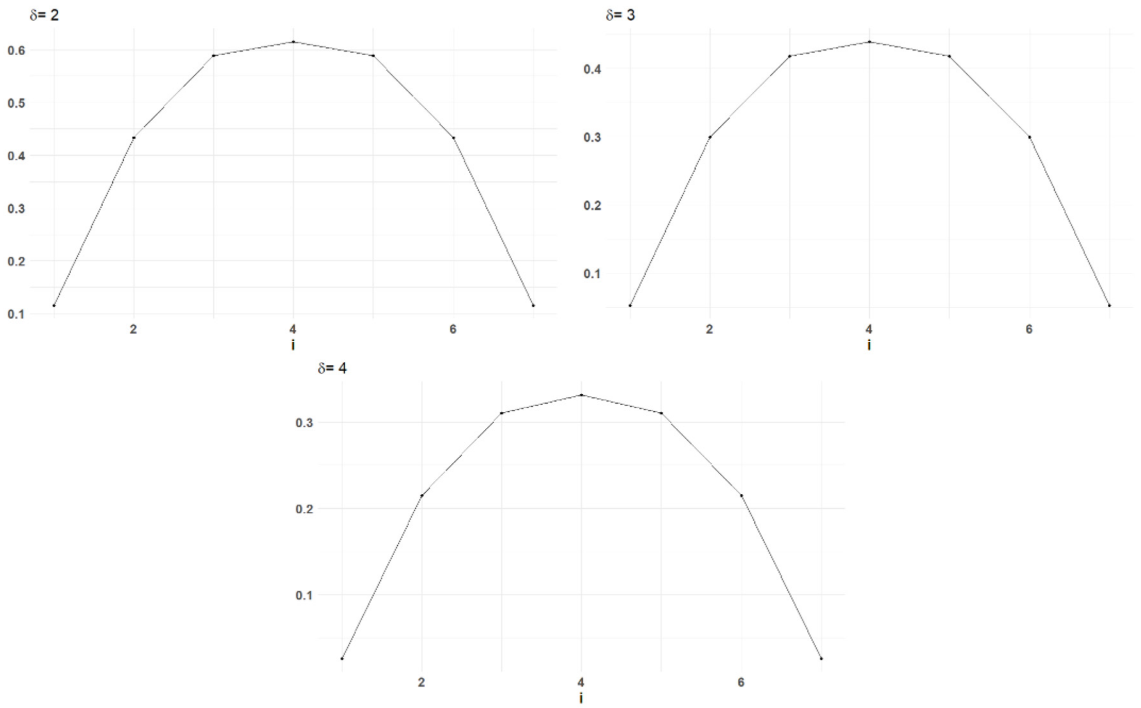

FIGURE 4
Information-generating function of the ith order statistics of $N(0, 1)$ distribution.

pdf, the information-generating function derived from their order statistics exhibits several notable properties. The Vasicek entropy estimator, originally introduced in Equation 1, has been instrumental in the progression of statistical analysis techniques. Its formulation is given by:

$$En(h_m) = -\int_{-\infty}^{\infty} h(x)\ln h(x)\,dx = -\int_0^1 \ln\left[\left(\frac{d}{d\chi}H^{-1}(\chi)\right)^{-1}\right] d\chi$$
$$= \frac{1}{m}\sum_{i=1}^{m}\ln\left[\frac{m}{2u^*}\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)\right], \tag{19}$$

Here, u^* is a positive integer satisfying $u^* < \frac{m}{2}$. For boundary handling, the values are extended such that $X_i = X_1$ when $i < 1$, and $X_i = X_m$ when $i > m$. The generalized entropy expressions for the smallest and largest ordered statistics can be reformulated as follows:

$$GEn_\delta(X^*_{1,m}) = \int_0^1 \left[m(1-u)^{m-1}\right]^\delta f^{\delta-1}(F^{-1}(y))\,du,$$

$$GEn_\delta(X^*_{m,m}) = \int_0^1 \left[mu^{m-1}\right]^\delta f^{\delta-1}(F^{-1}(y))\,du.$$

Park [26], expanding on the foundation laid by Vasicek [25], proposed a test for symmetry based on entropy derived from order statistics. Following this approach, sample-based estimators of $GEn_\delta(X^*_{1,m})$ and $GEn_\delta(X^*_{m,m})$ for a sample size m and $k = 1, 2, \ldots$, can be expressed as:

$$\widehat{GEn_\delta(X^*_{1,k})} = \frac{k^\delta}{m}\sum_{i=1}^{m}\left(1-\frac{i}{m+1}\right)^{k\delta-\delta}\left(\frac{2u^*}{m\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)}\right)^{\delta-1},$$

$$\widehat{GEn_\delta(X^*_{k,k})} = \frac{k^\delta}{m}\sum_{i=1}^{m}\left(\frac{i}{m+1}\right)^{k\delta-\delta}\left(\frac{2u^*}{m\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)}\right)^{\delta-1}.$$

Accordingly, the expression $\widehat{\Theta_k} = \widehat{GEn_\delta(X^*_{1,k})} - \widehat{GEn_\delta(X^*_{k,k})}$, defined for $k = 1, 2, \ldots$, can be approximated through the following empirical estimator:

$$\widehat{\Theta_k} = \frac{k^\delta}{m}\sum_{i=1}^{m}\left(\frac{2u^*}{m\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)}\right)^{\delta-1}\left[\left(1-\frac{i}{m+1}\right)^{k\delta-\delta} - \left(\frac{i}{m+1}\right)^{k\delta-\delta}\right].$$

To simplify the analysis, we fix $k = 2$ in what follows and suggest employing the estimator:

$$\widehat{\Theta_2} = \frac{2^\delta}{m}\sum_{i=1}^{m}\left(\frac{2u^*}{m\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)}\right)^{\delta-1}\left[\left(1-\frac{i}{m+1}\right)^{\delta} - \left(\frac{i}{m+1}\right)^{\delta}\right]$$
$$= \frac{2^\delta}{m}\sum_{i=1}^{m}\left(\frac{2u^*}{m\left(X^*_{i+u^*,m} - X^*_{i-u^*,m}\right)}\right)^{\delta-1}\Phi\left(\frac{i}{m+1}\right),$$

in which $\Phi(v) = -\Phi(1-v)$, and $\Phi(v)$ is both continuous and limited. This estimator corresponds to $\Theta_2 = GEn_\delta(X^*_{1,2}) - GEn_\delta(X^*_{2,2})$ and is utilized to evaluate whether the distribution of the random variable X^* is symmetric. Substantial deviations of Θ_2, whether in a positive or negative direction, can be interpreted as evidence of asymmetry in the underlying distribution.

Theorem 4.2. Let $X^*_1, \ldots, X^*_m$ be an equally distributed and independent random variables, and define $Y^*_i = aX^*_i + b^*$ for constants $a > 0$ and $b^* \in \mathbb{R}$, for each $i = 1, \ldots, m$. Denote the estimators of Θ_2 based on the sequences $\{X^*_i\}$ and $\{Y^*_i\}$ as $\widehat{\Theta_2}^{X^*}$ and $\widehat{\Theta_2}^{Y^*}$, respectively. Then, the following relationships (expectation, variance, and mean square error, respectively) hold:

(1) $\mathbb{E}[\widehat{\Theta_2}^{Y^*}] = \frac{\mathbb{E}[\widehat{\Theta_2}^{X^*}]}{a^{\delta-1}}$,
(2) $\mathrm{Var}[\widehat{\Theta_2}^{Y^*}] = \frac{\mathrm{Var}[\widehat{\Theta_2}^{X^*}]}{a^{2\delta-2}}$,
(3) $\mathrm{MSE}[\widehat{\Theta_2}^{Y^*}] = \frac{\mathrm{MSE}[\widehat{\Theta_2}^{X^*}]}{a^{2\delta-2}}$.

Proof. We begin by expressing the estimator for Θ_2 based on the transformed variables:

$$\widehat{\Theta_2}^{Y} = \frac{2^\delta}{m}\sum_{i=1}^{m}\left(\frac{2u^*}{m\left(Y^*_{i+u^*,m} - Y^*_{i-u^*,m}\right)}\right)^{\delta-1}\left[\left(1-\frac{i}{m+1}\right)^{\delta} - \left(\frac{i}{m+1}\right)^{\delta}\right]$$
$$= \frac{2^\delta}{m}\sum_{i=1}^{m}\left(\frac{2u^*}{m\left(aX^*_{i+u^*,m} - aX^*_{i-u^*,m}\right)}\right)^{\delta-1}\left[\left(1-\frac{i}{m+1}\right)^{\delta} - \left(\frac{i}{m+1}\right)^{\delta}\right].$$

This transformation directly leads to the stated scaling properties, completing the proof.

However, the estimator $\widehat{\Theta_2}$ depends not only on the observed sample, but also varies with the chosen window size u^*. Determining its exact distribution under the null hypothesis presents significant analytical challenges. Consequently, Monte Carlo simulation is used to estimate the critical values. Following prior studies (e.g., McWilliams [27] and Corzo and Babativa [28]), the generalized lambda distribution is selected as an alternative model. From this distribution, samples of sizes $m = 20, 30, 50,$ and 100 are generated across nine different parameter settings. The simulated data are defined as

$$x_i = \eta_1 + \frac{v_i^{\eta_3} - (1-v_i)^{\eta_4}}{\eta_2}, \quad 0 \le v_i \le 1, \quad i = 1, 2, \ldots, m.$$

Table 1 presents the parameter values $\eta_1, \eta_2, \eta_3,$ and η_4, originally chosen by McWilliams [27]. For each parameter combination, 1,000 samples are produced for each sample size. To determine the optimal u^*, we utilize a heuristic formula suggested by Crzcgorzewski and Wirczorkowski [29] for entropy estimation, given by

$$u^* = [\sqrt{m} + 0.5], \tag{20}$$

where $[\cdot]$ denotes the floor function. Figure 5 illustrates the empirical distributions of the test statistic $\widehat{\Theta_2}$, based on 10,000 replications from the standard normal distribution. These distributions are shown for sample sizes $m = 25, 40, 50, 70,$ and 100, with u^* selected via Equation 20. Sample generation and computation of the test statistic were performed using Wolfram Mathematica (version 13), chosen for its efficient random number generation and symbolic computation features. Further statistical analysis and visualization were conducted in R, leveraging its

TABLE 1 Parameter configurations of the generalized lambda distribution used in the Monte Carlo simulations, categorized into nine distinct cases.

Case of study	η_1	η_2	η_3	η_4	Value of skewness	Value of kurtosis
1	0.0000	0.1975	0.1349	0.1349	0.0000	3.0000
2	−0.1167	−0.3517	−0.1300	−0.1600	0.8000	11.4000
3	0.0000	−1.0000	−0.1000	−0.1800	2.0000	21.2000
4	3.5865	0.0431	0.0252	0.0940	0.9000	4.2000
5	0.0000	−1.0000	−0.0075	−0.0300	1.5000	7.5000
6	0.0000	1.0000	1.4000	0.2500	0.5000	2.2000
7	0.0000	1.0000	0.0001	0.1000	1.5000	5.8000
8	0.0000	−1.0000	−0.0010	−0.1300	3.1600	23.8000
9	0.0000	−1.0000	−0.0001	−0.1700	3.8800	40.7000

Each case specifies values for skewness and kurtosis.

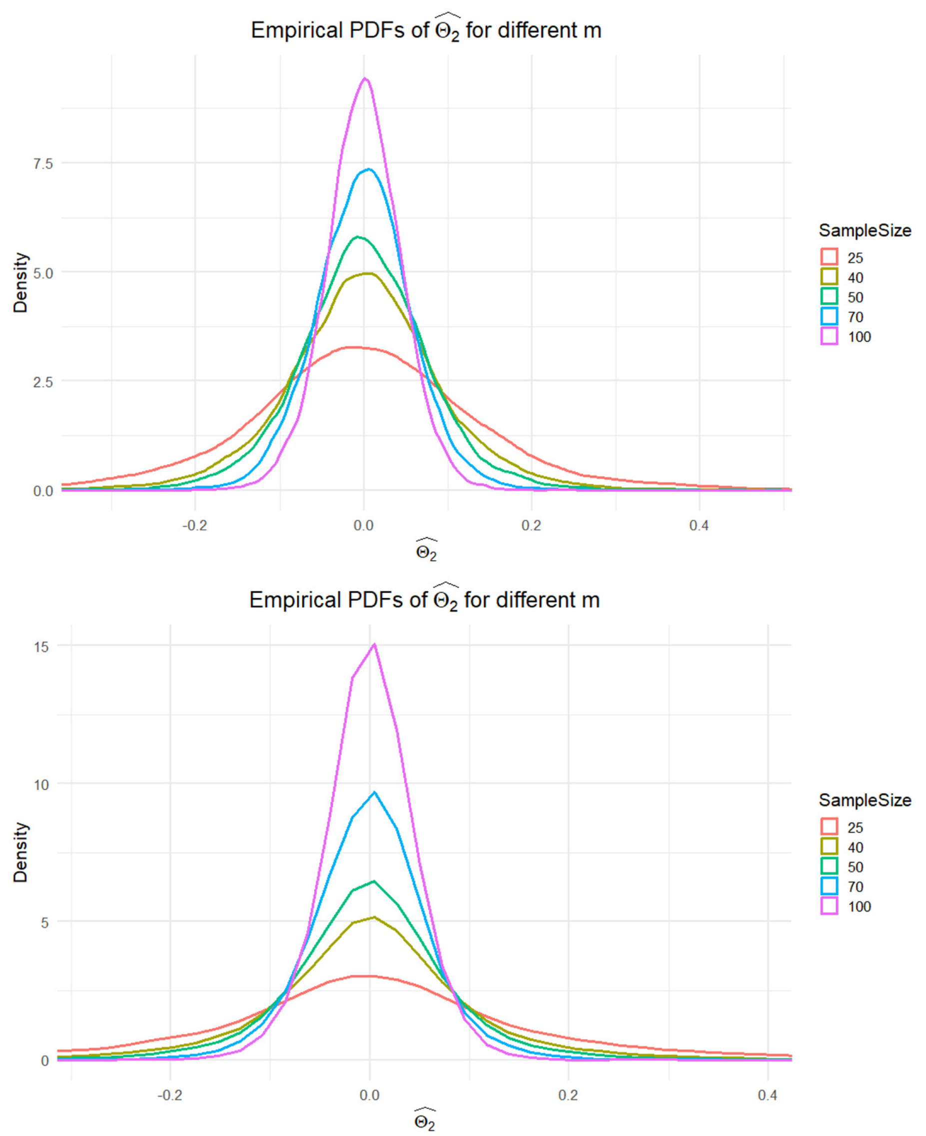

FIGURE 5
Empirical density plots of the test statistic based on 50,000 samples generated under the null distribution for sample sizes $m = 25, 40, 50, 70$, and 100, with $\delta = 2$ (top panel) and $\delta = 2$ (bottom panel).

advanced capabilities in statistical computing and graphical presentation. In Figure 5, as the sample size m increases, the empirical pdf of the statistic $\widehat{\Theta_2}$ becomes increasingly concentrated around its central value. Specifically, larger sample sizes yield steeper and more sharply peaked curves, reflecting a reduction in variability due to the greater amount of information contained in the sample. Conversely, smaller sample sizes produce flatter, more dispersed distributions, indicating greater variability. This behavior is consistent with the general principles of asymptotic theory, where statistics based on larger samples tend to exhibit reduced variance and greater stability.

Using a 1,000-reiteration Monte Carlo simulation, Table 2 presents the exact critical quantities of the examined statistic $\widehat{\Theta_2}$ for varying sample sizes, which correspond to the statistically significant level $\alpha^* = 0.05$. According to Table 2, we observe that the value of zero lies within the critical intervals as both m and δ increase. Furthermore, the length of these intervals decreases significantly, converging closely around zero.

Furthermore, the power of the test is calculated as the percentage of the 1,000 samples in the important range that reject the symmetrical null assumption at the level of significance $\alpha^* = 0.05$. The expected power levels for the proposed test are shown in Table 3.

The determination of the critical values and the power for our proposed symmetry test at a significance level of $\alpha^* = 0.05$ was carried out as follows:

(1) Generate a random sample of size m from the standard normal distribution, and then calculate the corresponding test statistic for the sample;
(2) Repeat Step 1 a total of 1,000 times and define the critical values based on the 25th and 975th percentiles of the obtained test statistics (that is, the 25th and 975th ordered statistics, $\widehat{\Theta_2^{(25)}}$ and $\widehat{\Theta_2^{(975)}}$, are used to set the thresholds. Specifically, the critical values are given by $\widehat{\Theta_2^{\alpha^*=0.05}} = \widehat{\Theta_2^{(975)}}$ and $\widehat{\Theta_2^{\alpha^*=0.05}} = \widehat{\Theta_2^{(975)}}$, considering that for $\alpha^* = 0.05$, we have $\frac{\alpha^*}{2} = 0.025 = \frac{25}{1{,}000}$ and $1 - \frac{\alpha^*}{2} = 0.975 = \frac{975}{1{,}000}$: Hence, the null hypothesis is rejected if $\widehat{\Theta_2}$ falls below $\widehat{\Theta_2^{(25)}}$ or exceeds $\widehat{\Theta_2^{(975)}}$, and accepted otherwise when $\widehat{\Theta_2^{(25)}} < \widehat{\Theta_2} < \widehat{\Theta_2^{(975)}}$);
(3) Draw another sample of size m under the null distribution, then verify whether the absolute value of the test statistic crosses the critical thresholds;
(4) Estimate the test's power as the proportion of rejections over 1,000 repetitions of Step 3.

4.1.1 Performance assessment using Monte Carlo methods

To rigorously evaluate the proposed testing methodology, we implement Monte Carlo simulation techniques. The comparative analysis examines statistical power across multiple competing tests, with detailed results presented in Tables 3, 4.

4.1.1.1 Comparative test procedures

The study incorporates the following established testing approaches for benchmarking purposes:

1. McWilliams' runs-based examination [27] utilizes the counting measure $At^{(1)}$ as its fundamental test statistic, quantifying total sequence runs.
2. Baklizi's modified runs analysis [30] introduces an adjusted formulation of the runs test, operationalized through statistic $At^{(2)}$.
3. Signed-Rank Wilcoxon procedure [31], developed by Gibbons and Chakraborti, employs the test measure $At^{(3)}$ for distribution-free inference.
4. Tajuddin's rank-sum approach [32] adapts the Wilcoxon two-sample framework using test statistic $At^{(4)}$.
5. Cheng-Balakrishnan rank methodology [33] implements the testing criterion $At^{(5)}$ for nonparametric analysis.

TABLE 2 The test statistic's critical intervals $\widehat{\Theta_2}$ at the significance level of 0.05.

$m\backslash\delta$	2	3	4	5
20	(−0.629547, −0.117521)	(−0.76081, 0.0950801)	(−1.2442, 0.124146)	(−1.97969, 0.196312)
30	(−0.492156, −0.107647)	(−0.526785, 0.0502732)	(−0.738742, 0.0748229)	(−0.947342, 0.110406)
50	(−0.363772, −0.0906927)	(−0.337903, 0.0470503)	(−0.370311, 0.0707895)	(−0.389405, 0.0949751)
75	(−0.296201, −0.083219)	(−0.258534, 0.0181001)	(−0.265887, 0.0351457)	(−0.25885, 0.0469076)
100	(−0.241448, −0.0573265)	(−0.211483, 0.0241135)	(−0.187036, 0.0341597)	(−0.166362, 0.0496393)
150	(−0.192845, −0.0382762)	(−0.165892, 0.0213619)	(−0.144363, 0.0306424)	(−0.120414, 0.0311807)
$m\backslash\delta$	6	7	8	9
20	(−3.15239, 0.288594)	(−5.67084, 0.449217)	(−9.96943, 0.673464)	(−17.3659, 0.995891)
30	(−1.27852, 0.138963)	(−1.73192, 0.156962)	(−2.54621, 0.188728)	(−3.6984, 0.221285)
50	(−0.427708, 0.0944352)	(−0.478008, 0.0888199)	(−0.550025, 0.0913595)	(−0.637507, 0.0878994)
75	(−0.254051, 0.0466358)	(−0.239656, 0.0411638)	(−0.252259, 0.0392213)	(−0.252353, 0.0359931)
100	(−0.149328, 0.0472344)	(−0.135345, 0.0430554)	(−0.136509, 0.037766)	(−0.133282, 0.0346512)
150	(−0.10202, 0.0269356)	(−0.090705, 0.0243265)	(−0.0824416, 0.0210497)	(−0.0745419, 0.0187016)

TABLE 3 Comparative analysis of the power examination for the test at the 0.05 significance threshold.

Alternatives	m	$\widehat{\Theta_2}$							
		$\delta = 2$	3	4	5	6	7	8	9
Case-1(Hy0)	20	0.053	0.05	0.052	0.053	0.053	0.05	0.049	0.029
	30	0.056	0.055	0.053	0.055	0.059	0.057	0.054	0.041
	50	0.054	0.051	0.057	0.054	0.06	0.053	0.05	0.029
	100	0.057	0.05	0.058	0.049	0.047	0.049	0.047	0.016
Case-2	20	0.124	0.133	0.153	0.153	0.152	0.148	0.147	0.106
	30	0.124	0.144	0.162	0.18	0.187	0.194	0.188	0.135
	50	0.123	0.162	0.189	0.209	0.228	0.244	0.249	0.168
	100	0.137	0.172	0.244	0.277	0.314	0.346	0.348	0.205
Case-3	20	0.731	0.796	0.886	0.936	0.953	0.967	0.975	0.965
	30	0.726	0.834	0.918	0.961	0.98	0.991	0.993	0.993
	50	0.736	0.857	0.96	0.981	0.991	0.999	0.999	1
	100	0.613	0.872	0.964	0.986	0.996	0.999	1.	1.
Case-4	20	0.07	0.013	0.006	0.002	0.002	0.001	0.0	0.0
	30	0.06	0.039	0.001	0.003	0.001	0.0	0.0	0.0
	50	0.04	0.032	0.001	0.0	0.0	0.0	0.0	0.0
	100	0.04	0.028	0.003	0.0	0.0	0.0	0.0	0.0
Case-5	20	0.975	0.997	1.	1.	1.	1.	1.	1.
	30	0.974	0.999	1.	1.	1.	1.	1.	1.
	50	0.961	0.999	0.999	1.	1.	1.	1.	1.
	100	0.954	1.	1.	1.	1.	1.	1.	1.
Case-6	20	0.43	0.734	0.8	0.858	0.88	0.895	0.903	0.873
	30	0.531	0.819	0.872	0.923	0.942	0.967	0.972	0.959
	50	0.678	0.893	0.948	0.972	0.985	0.994	0.999	0.996
	100	0.933	0.987	0.997	0.999	1.	1.	0.999	0.999
Case-7	20	0.887	0.998	0.999	1.	1.	1.	1.	1.
	30	0.916	1.	1.	1.	1.	1.	1.	1.
	50	0.96	1.	1.	1.	1.	1.	1.	1.
	100	0.999	1.	1.	1.	1.	1.	1.	1.
Case-8	20	0.877	0.993	0.999	1.	1.	1.	1.	1.
	30	0.949	0.999	1.	1.	1.	1.	1.	1.
	50	0.987	1.	1.	1.	1.	1.	1.	1.
	100	1.	1.	1.	1.	1.	1.	1.	1.
Case-9	20	0.885	0.991	0.998	0.999	1.	1.	1.	1.
	30	0.955	0.998	1.	1.	1.	1.	1.	1.
	50	0.996	0.999	1.	1.	1.	1.	1.	1.
	100	1.	1.	1.	1.	1.	1.	1.	1.

6. Modarres' trimmed statistical measure [34] incorporates a proportional trimming factor q within its test statistic $At_q^{(6)}$.
7. Baklizi's size-adaptive test [35] accounts for both sample dimensionality m and trimming proportion q through statistic $At_{m;q}^{(7)}$.
8. Baklizi's secondary testing framework [35] presents an alternative formulation based on $At^{(8)}$.
9. Baklizi's extended testing protocol [36] features an enhanced version using evaluation metric $At^{(9)}$.
10. Corzo-Babativa nonparametric technique [28] establishes its testing procedure on the foundation of $At^{(10)}$.

TABLE 4 Comparison of the test's power analysis at the significance level 0.05.

Alternatives	m	$At^{(1)}$	$At^{(2)}$	$At^{(3)}$	$At^{(4)}$	$At^{(5)}$	$At^{(6)}_{25}$	$At^{(6)}_{60}$	$At^{(7)}_{m;0}$	$At^{(7)}_{m;0.8}$	$At^{(8)}$	$At^{(9)}$	$At^{(10)}$	$At^{(11)}$
Case-1(Hy0)	20	0.044	0.052	0.043	0.045	0.048	0.047	0.046	0.051	0.055	0.055	0.046	0.056	0.049
	30	0.049	0.053	0.052	0.051	0.052	0.051	0.054	0.051	0.054	0.048	0.046	0.047	0.048
	50	0.051	0.054	0.050	0.052	0.051	0.049	0.049	0.049	0.051	0.058	0.051	0.046	0.049
	100	0.051	0.047	0.052	0.048	0.052	0.053	0.051	0.055	0.054	0.049	0.051	0.048	0.049
Case-2	20	0.052	0.057	0.051	0.051	0.054	0.054	0.053	0.057	0.062	0.046	0.058	0.070	0.097
	30	0.052	0.051	0.055	0.056	0.061	0.053	0.055	0.051	0.063	0.058	0.062	0.061	0.130
	50	0.055	0.056	0.052	0.060	0.070	0.062	0.066	0.058	0.062	0.053	0.075	0.068	0.201
	100	0.054	0.051	0.055	0.071	0.091	0.057	0.062	0.053	0.066	0.065	0.106	0.084	0.324
Case-3	20	0.067	0.075	0.055	0.079	0.080	0.079	0.087	0.057	0.088	0.070	0.114	0.112	0.667
	30	0.074	0.075	0.062	0.097	0.119	0.094	0.109	0.069	0.128	0.088	0.156	0.125	0.809
	50	0.089	0.094	0.064	0.131	0.204	0.120	0.153	0.075	0.145	0.141	0.253	0.206	0.920
	100	0.113	0.109	0.088	0.224	0.366	0.169	0.217	0.122	0.228	0.233	0.486	0.356	0.988
Case-4	20	0.090	0.103	0.061	0.106	0.118	0.122	0.142	0.072	0.138	0.087	0.187	0.177	0.038
	30	0.114	0.122	0.070	0.149	0.219	0.166	0.199	0.100	0.229	0.142	0.287	0.243	0.071
	50	0.143	0.154	0.085	0.209	0.428	0.234	0.301	0.144	0.303	0.314	0.499	0.443	0.160
	100	0.216	0.209	0.127	0.385	0.757	0.406	0.522	0.333	0.572	0.595	0.818	0.750	0.567
Case-5	20	0.115	0.131	0.067	0.133	0.155	0.162	0.190	0.095	0.165	0.120	0.254	0.235	0.992
	30	0.151	0.160	0.080	0.194	0.309	0.232	0.287	0.131	0.333	0.219	0.404	0.343	0.998
	50	0.197	0.213	0.103	0.287	0.587	0.342	0.437	0.230	0.457	0.455	0.668	0.602	1.000
	100	0.321	0.316	0.166	0.522	0.890	0.566	0.696	0.556	0.769	0.784	0.939	0.885	1.000
Case-6	20	0.200	0.234	0.072	0.160	0.256	0.346	0.396	0.136	0.267	0.191	0.420	0.468	0.454
	30	0.303	0.330	0.095	0.231	0.606	0.558	0.671	0.256	0.649	0.469	0.653	0.715	0.610
	50	0.497	0.524	0.122	0.364	0.950	0.825	0.920	0.642	0.908	0.914	0.894	0.972	0.742
	100	0.782	0.782	0.198	0.633	1.000	0.989	0.998	0.995	1.000	1.000	0.994	1.000	1.000
Case-7	20	0.311	0.358	0.096	0.281	0.421	0.511	0.578	0.226	0.330	0.314	0.593	0.644	0.997
	30	0.457	0.490	0.123	0.393	0.797	0.750	0.828	0.444	0.823	0.689	0.854	0.868	0.999
	50	0.683	0.707	0.185	0.600	0.991	0.941	0.977	0.860	0.978	0.980	0.980	0.994	1.000
	100	0.928	0.927	0.358	0.883	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Case-8	20	0.373	0.426	0.105	0.330	0.494	0.594	0.656	0.295	0.366	0.389	0.666	0.715	0.999
	30	0.539	0.570	0.150	0.484	0.861	0.819	0.878	0.555	0.876	0.790	0.913	0.915	1.000
	50	0.761	0.782	0.233	0.697	0.996	0.970	0.989	0.930	0.991	0.991	0.993	0.998	1.000
	100	0.966	0.965	0.420	0.947	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Case-9	20	0.399	0.452	0.112	0.351	0.530	0.631	0.696	0.322	0.359	0.428	0.692	0.752	0.998
	30	0.580	0.614	0.152	0.498	0.877	0.848	0.900	0.608	0.898	0.821	0.924	0.929	1.000
	50	0.802	0.821	0.241	0.725	0.997	0.979	0.992	0.953	0.993	0.995	0.995	0.999	1.000
	100	0.980	0.980	0.441	0.956	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

11. Noughabi-Jarrahiferiz extropy-based method [23] develops a novel approach using order statistic extropy measure, formalized as $At^{(11)}$.

A symmetric distribution is shown by Case-1 in Table 3, where we can observe that all of the values of δ have powers of the testing statistic $\widehat{\Theta_2}$ that are near 0.05 as expected. The corresponding distribution is asymmetric in the next 8 examples (situations 2 and 3 are almost symmetrical). Test statistics with varying δ values, particularly as they grow, exhibit comparable powers in cases 5, 7, 8, and 9. The weakened power values in instance 4 may be explained by the fact that η_1 is much larger than 0, whereas it is nearly 0 in the other examples. We may conclude that our suggested test, based on the information-generating function of order statistics,

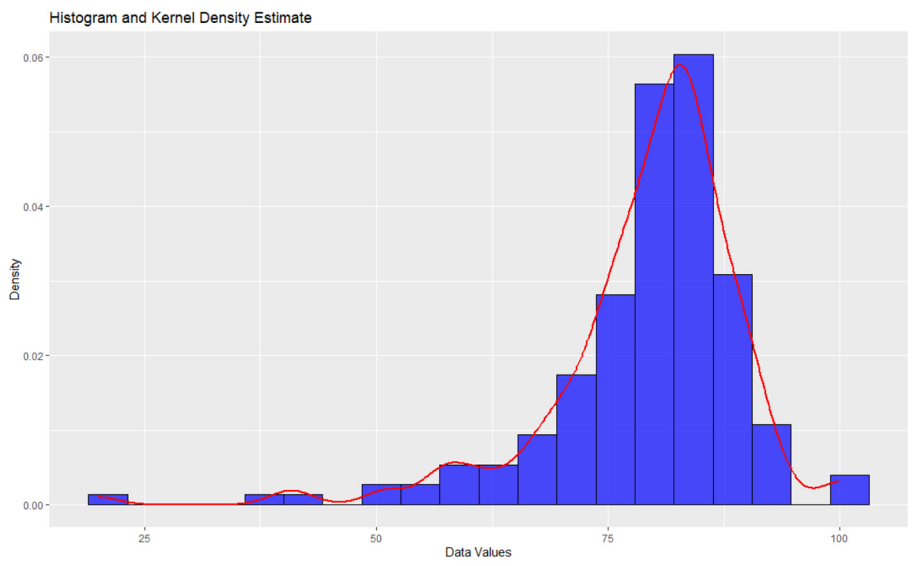

FIGURE 6
The kernel density estimate and histogram of the data set.

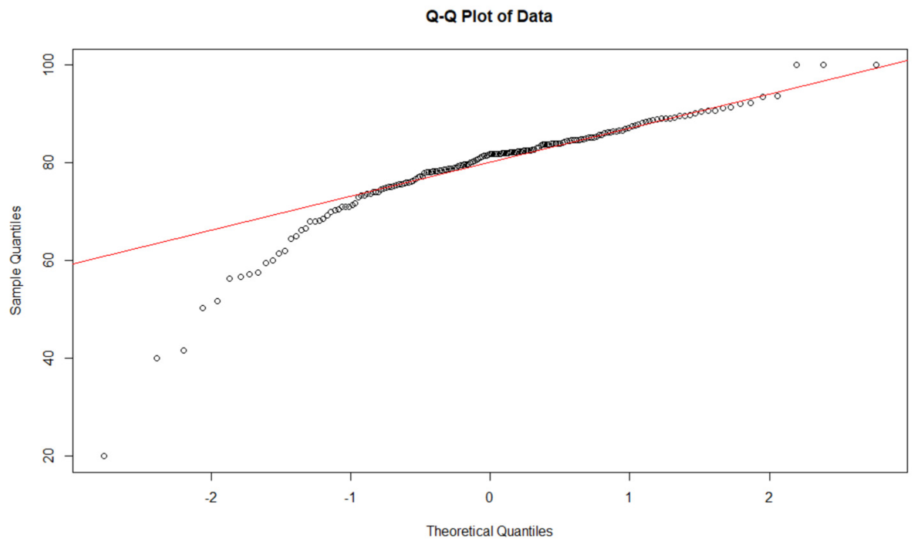

FIGURE 7
Q–Q diagram of all the data.

performs well in the simulation study as the values of δ increase, compared with the other tests in Table 4. Therefore, we anticipate that the suggested test will outperform the competing tests across a wide range of real-world applications.

4.2 Real data set

To demonstrate the applicability of our methodology, we used data from the health statistics bulletin published by the General Authority for Statistics in the Kingdom of Saudi Arabia. This comprehensive dataset captures key health indicators, including:

1. Prevalence of chronic diseases,
2. Mental health status.

The statistical population encompasses all households—both Saudi and non-Saudi—permanently residing in the Kingdom of Saudi Arabia. The survey covers 13 administrative regions and 151 governorates, with 2023 as the base year for calculating indicators. Health status among adults (aged 15 years and above) is assessed using the Visual Analog Scale (VAS), with scores ranging from 0 (worst possible health) to 100 (excellent health). The data is stratified by administrative region and age group, enabling detailed demographic and geographic analysis. The complete dataset is publicly available through the General Authority

TABLE 5 The symmetry test results for different sensitivity parameters δ.

δ	$\lvert\widehat{\Theta}_2\rvert$	p-value
2	0.0189	< 0.001
3	0.00180	0.012
4	0.000173	0.051
5	0.0000161	0.103

for Statistics portal at: https://www.stats.gov.sa/statistics-tabs?tab=436312&category=417594. Figure 6 shows visualizations of the data sets histogram and the kernel density estimates, while Figure 7 shows the Q–Q diagram.

4.2.1 Bootstrap procedure

Since the null distribution of $\widehat{\Theta}_2$ is non-pivotal, we employ a reflection bootstrap:

1. Symmetrize the data by generating $\mathcal{X}_{\text{sym}} = \{X_1^*, \dots, X_m^*\} \cup \{2\tilde{X} - X_1^*, \dots, 2\tilde{X} - X_m^*\}$, where $\tilde{X}$ is the sample median.
2. For each $b = 1, \dots, B$: (i) Resample $\mathcal{X}_b^*$ uniformly from $\mathcal{X}_{\text{sym}}$. (ii) Compute $\widehat{\Theta}_{2,b}^*$.
3. The p-value is $\frac{1}{B}\sum_{b=1}^{B} \mathbb{I}(\widehat{\Theta}_{2,b}^* \geq \widehat{\Theta}_2^{\text{obs}})$.

Results. The sample exhibits negative skewness (-1.89) and high kurtosis (9.55), indicating:

- A left-skewed distribution with a longer left tail
- Heavy tails and peakedness relative to a normal distribution

The symmetry test results for different sensitivity parameters δ are given in Table 5.
Key findings:

- Strong evidence against symmetry for $\delta = 2$ and 3 ($p < 0.05$).
- Marginal evidence at $\delta = 4$ ($p = 0.051$).
- Insufficient evidence to reject symmetry at $\delta = 5$ ($p = 0.103$).

Interpretation:

- The negative skewness suggests potential outliers in the left tail.
- The decreasing p-values with higher δ indicate the test's reduced sensitivity to asymmetry.

5 Conclusion

This research advances the theoretical understanding of information-generating functions for order statistics through several key contributions. We have systematically investigated monotonicity properties and derived bounds for the proposed measure. The study establishes important stochastic ordering results based on this information-theoretic framework, demonstrating that equality of information-generating function measures for order statistics uniquely determines their parent distributions. Furthermore, we have developed novel characterization theorems for the exponential distribution using this approach. For symmetric distributions, our analysis shows that the information-generating function $GEn_\delta(X_{i,m}^*)$ exhibits extremal behavior (either a local maximum or a minimum) at the median position. This theoretical finding is substantiated through explicit computations for both uniform and standard normal distributions. Building on these theoretical insights, we have formulated a nonparametric symmetry test based on the proposed measure, whose effectiveness increases with δ. The practical utility of our methodology is validated through comprehensive simulation studies and an application to chronic disease management data. Both theoretical and empirical results consistently show that higher values of δ significantly improve the test's performance, confirming the robustness of our approach.

Future studies will include a comprehensive performance comparison with a broader set of established symmetry tests, such as the Baringhaus–Henze, Ahmad–Li, and Bonett–Seier tests, to further situate our method within the broader literature. While this study provides initial validation of the test's power, a more comprehensive investigation against a wider array of alternatives, including heavy-tailed and bounded-support distributions, is a priority for future research.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

MM: Writing – original draft, Investigation, Software, Formal analysis, Funding acquisition, Visualization, Resources, Supervision, Validation, Project administration, Conceptualization, Writing – review & editing, Data curation, Methodology. MA-L: Writing – review & editing, Methodology, Formal analysis. EA: Writing – review & editing, Methodology, Formal analysis. HS: Funding acquisition, Data curation, Visualization, Resources, Conceptualization, Formal analysis, Validation, Project administration, Methodology, Writing – review & editing, Software, Investigation, Writing – original draft, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was funded by Prince Sattam bin Abdulaziz University (PSAU/2025/02/35141).

Acknowledgments

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/02/35141).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. Shannon C. A mathematical theory of communication. *Bell Syst Tech J*. (1948) 27:379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

2. Golomb S. The information generating function of a probability distribution (corresp.). *IEEE Trans Inform Theory*. (1966) 12:75–7. doi: 10.1109/TIT.1966.1053843

3. Kharazmi O, Balakrishnan N. Cumulative and relative cumulative residual information generating measures and associated properties. *Commun Stat Theory Methods*. (2021) 52:5260–73. doi: 10.1080/03610926.2021.2005100

4. Kharazmi O, Balakrishnan N. Cumulative residual and relative cumulative residual Fisher information and their properties. *IEEE Trans Inf Theory*. (2021) 67:6306–12. doi: 10.1109/TIT.2021.3073789

5. Kharazmi O, Balakrishnan N. Jensen-information generating function and its connections to some well-known information measures. *Stat Prob Lett*. (2021) 170:108995. doi: 10.1016/j.spl.2020.108995

6. Kharazmi O, Balakrishnan N. Information generating function for order statistics and mixed reliability systems. *Commu Stat Theory Methods*. (2021) 51:7846–55. doi: 10.1080/03610926.2021.1881123

7. Kharazmi O, Balakrishnan N. Generating function for generalized Fisher information measure and its application to finite mixture models. *Hacet J Math Stati*. (2022) 51:1472–83. doi: 10.15672/hujms.1094273

8. Zamani Z, Kharazmi O, Balakrishnan N. Information generating function of record values. *Math Methods Stat*. (2022) 31:120–33. doi: 10.3103/S1066530722030036

9. Kharazmi O, Balakrishnan N, Ozonur D. Jensen-discrete information generating function with an application to image processing. *Soft Comput*. (2023) 27:4543–52. doi: 10.1007/s00500-023-07863-0

10. Kayal S, Balakrishnan N. Quantile-based information generating functions and their properties and uses. *Prob Eng Inf Sci*. (2024) 38:1–19. doi: 10.1017/S0269964824000068

11. Onicescu O. *The Informational Energy, Component of the Statistical Barometer Concerning the Systems*. Bucharest: Technical Publishing House (1966).

12. Bhatia PK. On measures of information energy. *Inf Sci*. (1997) 97:233–40. doi: 10.1016/0020-0255(94)00071-9

13. Balakrishnan N, Selvitella A. Symmetry of a distribution via symmetry of order statistics. *Stat Probabil Lett*. (2017) 129:367–72. doi: 10.1016/j.spl.2017.06.023

14. Ahmadi J. Characterization results for symmetric continuous distributions based on the properties of k-records and spacings. *Stat Probabil Lett*. (2020) 162:108764. doi: 10.1016/j.spl.2020.108764

15. Mahdizadeh M, Zamanzade E. Estimation of a symmetric distribution function in multistage ranked set sampling. *Stat Papers*. (2020) 61:851–67. doi: 10.1007/s00362-017-0965-x

16. Dai XJ, Niu CZ, Guo X. Testing for central symmetry and inference of the unknown center. *Comput Stat Data An*. (2018) 127:15–31. doi: 10.1016/j.csda.2018.05.007

17. Bozin V, Milosevic B, Nikitin YY, Obradovic M. New characterization-based symmetry tests. *Bull Malays Math Sci Soc*. (2020) 43:297–320. doi: 10.1007/s40840-018-0680-3

18. Shaked M, Shanthikumar JG. *Stochastic Orders and Their Applications*. San Diego, CA: Academic Press (1994).

19. Aliprantis CD, Burkinshaw O. *Principles of Real Analysis*. London: Edward Arnold (1981).

20. Ebrahimi N, Soofi ES, Zahedi H. Information properties of order statistics and spacings. *IEEE Trans Inform Theory*. (2004) 50:177–83. doi: 10.1109/TIT.2003.821973

21. Fashandi M, Ahmadi J. Characterizations of symmetric distributions based on Renyi entropy. *Stat Probabil Lett*. (2012) 82:798–804. doi: 10.1016/j.spl.2012.01.004

22. Xiong PH, Zhuang WW, Qiu GX. Testing symmetry based on the extropy of record values. *J Nonparametr Stat*. (2021) 33:134–55. doi: 10.1080/10485252.2021.1914338

23. Noughabi HA, Jarrahiferiz J. Extropy of order statistics applied to testing symmetry. *Commun Stat-Simul C*. (2022) 51:3389–99. doi: 10.1080/03610918.2020.1714660

24. Mohamed MS, Almuqrin MA. Properties of fractional generalized entropy in ordered variables and symmetry testing. *AIMS Math*. (2025) 10:1116–41. doi: 10.3934/math.2025053

25. Vasicek O. A test for normality based on sample entropy. *J R Stat Soc B*. (1976) 38:54–9. doi: 10.1111/j.2517-6161.1976.tb01566.x

26. Park S. A goodness-of-fit test for normality based on the sample entropy of order statistics. *Stat Probabil Lett*. (1999) 44:359–63. doi: 10.1016/S0167-7152(99)00027-9

27. McWilliams TP. A distribution-free test for symmetry based on a runs statistic. *J Am Stat Assoc*. (1990) 85:1130–3. doi: 10.1080/01621459.1990.10474985

28. Corzo J, Babativa G. A modified runs test for symmetry. *J Stat Comput Sim*. (2013) 83:984–91. doi: 10.1080/00949655.2011.647026

29. Crzcgorzewski P, Wirczorkowski R. Entropy-based goodness-of-fit test for exponentiality. *Commun Stat-Theor M*. (1999) 28:1183–202. doi: 10.1080/03610929908832351

30. Baklizi A. A conditional distribution runs test for symmetry. *J Nonparametr Stat*. (2003) 15:713–8. doi: 10.1080/10485250310001634737

31. Gibbons JD, Chakraborti SM. *Non-Parametric Statistical Inference*. New York, NY: Dekker (1992).

32. Tajuddin IH. Distribution-free test for symmetry based on the Wilcoxon two-sample test. *J Appl Stat*. (1994) 21:409–15. doi: 10.1080/757584017

33. Cheng WH, Balakrishnan N. A modified sign test for symmetry. *Commun Stat-Simul C*. (2004) 33:703–9. doi: 10.1081/SAC-200033302

34. Modarres R, Gastwirth JL. A modified runs test for symmetry. *Stat Probabil Lett*. (1996) 31:107–12. doi: 10.1016/S0167-7152(96)00020-X

35. Baklizi A. Testing symmetry using a trimmed longest run statistic. *Aust N Z J Stat*. (2007) 49:339–47. doi: 10.1111/j.1467-842X.2007.00485.x

36. Baklizi A. Improving the power of the hybrid test. *Int J Contemp Math Sciences*. (2008) 3:497–9.