REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000–000
https://doi.org/00.00000/revstat.v00i0.000

Varentropy Properties of Record Values

Authors: Manoj Chacko ✉
– Department of Statistics, University of Kerala,
Trivandrum, Kerala, India.
manojchacko02@gmail.com

Annie Grace
– Department of Statistics, University of Kerala,
Trivandrum, Kerala, India.
anniegracees@gmail.com

Received: March 2024 Revised: August 2024 Accepted: September 2024

Abstract:

- In this paper, we study the varentropy measure for record values and examine some of its properties. We establish some bounds for the varentropy measure of record values. In addition, the varentropy measure of residual and past lifetimes based on record values are also considered. As an application of varentropy measure, we consider the problem of estimation of the varentropy for the Pareto model based on the record values. Maximum likelihood estimation and Bayes estimation are considered. A simulation study is performed to determine the performance of the estimators developed in this paper.

Keywords:

- *Record values, Entropy, Varentropy, Pareto distribution.*

AMS Subject Classification:

- 54C70, 62F15, 62B10, 94A17.

✉ Corresponding author

1. Introduction

Information theory, a pivotal discipline in applied mathematics and computer science, delves into data quantification, transmission, and comprehension. It seeks to unveil underlying patterns and structures within the system. The differential entropy introduced by Shannon (1948) is a well-known information measure that represents the expectation of the information content of an absolutely continuous random variable. Then, Shannon entropy of a continuous random variable X, with probability density function(pdf) $f(x)$, denoted by $H(X)$, is defined by

$$H(X) = E(-log\ f(X)) \tag{1.1}$$

$$= -\int_{\mathcal{X}} f(x) log\ f(x) dx. \tag{1.2}$$

The variance of the entropy, known as varentropy, studied by Fradelizi et al. (2016), represents the dispersion of information content around the entropy of a random variable X. This measure serves as a pivotal measurement in understanding how information is distributed within the probability distribution of X, depicting the degree of deviation or scattering of information from its expected value captured by entropy. Varentropy's significance lies in its ability to quantify the variability and uncertainty inherent in the information carried by X, holding vital implications across information theory, computer sciences, and statistics. Its applications encompass aiding in the assessment of communication system uncertainties in information theory, facilitating pattern recognition in computer sciences, and contributing to a deeper understanding of variability within probability distributions in statistical analyses. Nowadays, it is pointed out that this information measure has wide applications in reliability and survival analysis also. The varentropy measure, denoted by $\mathrm{VH}(X)$, is defined by

$$\mathrm{VH}(X) = Var(-log\ f(X)). \tag{1.3}$$

Some contributions on the varentropy can be found in various papers by Bobkov and Madiman (2011), Arıkan (2016), Kontoyiannis and Verdú (2013) and Di Crescenzo et al. (2024). Bobkov and Madiman (2011) demonstrated a concentration property of the information content when an n-dimensional random vector has a log-concave density function. Arıkan (2016) established that the sum of the varentropies at the output of the polar transform is less than or equal to the sum of the varentropies at the input, with equality if and only if at least one of the inputs has zero varentropy. Fradelizi et al. (2016) provided a bound for the varentropy of random vectors with log-concave densities, which is sharper than that proposed by Bobkov and Madiman (2011).

Recently, Maadani et al. (2020) introduced a new generalized varentropy based on the Tsallis entropy and obtained some bounds for it. Maadani et al. (2022) have introduced a method for calculating varentropy measure for the ith order statistic and a new stochastic order based on the varentropy and its relationships with the other stochastic orders. Raqab et al. (2022) studied the varentropy of the inactivity time of a random variable and its applications. Di Crescenzo and Paolillo (2021) introduced residual varentropy and studied its various properties with some applications related to the proportional hazards model and the first-passage times of an Ornstein-Uhlenbeck jump-diffusion process. Sharma and Kundu

(2023) introduced the concept of varentropy in a doubly truncated random variable and examined several theoretical properties. Alizadeh and Shafaei (2023) introduced some non-parametric estimates of the varentropy with some theoretical properties. Alizadeh and Shafaei (2024) constructed some tests for normality based on varentropy estimators. Saha and Kayal (2024) introduced two uncertainty measures, weighted past varentropy and weighted paired dynamic varentropy, and studied their properties.

If X represents the lifetime of a system, then the residual lifetime of the system when it is still operating at time t is $X_t = [X - t|X > t]$, which has the pdf given by

$$f_t(x) = \frac{f(x)}{\overline{F}(t)},\ x \geq t > 0,$$

where $\overline{F}(t) = 1 - F(t) > 0$ is the survival function of X . Muliere et al. (1993) and Ebrahimi (1996) defined the entropy of the random variable X_t as the residual entropy at time t and is given by

$$H(X_t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} log\, \frac{f(x)}{\overline{F}(t)} dx,\ t \geq 0. \tag{1.4}$$

Di Crescenzo and Paolillo (2021) explored the concept of residual varentropy, which represents the varentropy related to the residual lifetime distribution and suggest that when combined with residual entropy, this measure allows to analyze the dynamical information content of time-varying systems conditional on being active at current time. The residual varentropy of X at time t, denoted by $\text{RVH}(X_t)$, and is defined as

$$\text{RVH}(X_t) = \int_t^\infty \frac{f(x)}{\overline{F}(t)} \left[log\, \frac{f(x)}{\overline{F}(t)}\right]^2 dx - H^2(X_t). \tag{1.5}$$

In reliability theory, the past lifetime is studied extensively, which represents a variable dependent on failure occurring before a specified inspection time t. It holds importance in exploring the uncertainty surrounding the duration $(0, t)$ when observing systems that have already experienced failure by time t. If X denotes the lifetime of a system, then the past lifetime of a system at time t is $X_{[t]} = [X|X \leq t]$, which has the pdf given by

$$f_{[t]}(x) = \frac{f(x)}{F(t)},\ 0 \leq x < t.$$

Then the past entropy is defined as (see, Di Crescenzo and Longobardi (2002))

$$H(X_{[t]}) = -\int_0^t \frac{f(x)}{F(t)} log\, \frac{f(x)}{F(t)} dx. \tag{1.6}$$

Buono et al. (2022) introduced the concept of past varentropy, defined as the dynamic measure of variability of information for past lifetimes. The past varentropy of X at time t, denoted by $\text{PVH}(X_{[t]})$ is given by (see, Buono et al. (2022))

$$\text{PVH}(X_{[t]}) = \int_0^t \frac{f(x)}{F(t)} \left[log\, \frac{f(x)}{F(t)}\right]^2 dx - H^2(X_{[t]}). \tag{1.7}$$

Extensive investigation over the last two decades has been dedicated to understand the role of information measures concerning residual and past lifetimes in reliability modeling, with the initial groundwork laid out by Muliere et al. (1993) and Di Crescenzo and Longobardi (2002).

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables having a common cumulative distribution function (cdf) $F(x)$, which is absolutely continuous. An observation X_j is called an upper record if its value exceeds that of all preceding observations. Thus, X_j is an upper record if $X_j > X_i$ for every $i < j$. In an analogous way, one can also define lower record values.

The times at which upper record values appear are given by the random variables T_j which are called record times and are defined by $T_1 = 1$ with probability 1 and, for $j \geq 2$, $T_j = Min\{i : X_i > X_{T_{j-1}}\}$. The waiting time between the ith upper record value and the $(i+1)$th upper record value is called the inter-record time, and is denoted by $\Delta_i = T_{i+1} - T_i$, $i = 1, 2, \ldots$. In an analogous way, one can also define record times and inter-record times for lower record values.

Let $U_1, U_2, \ldots$ be a sequence of upper record values from a distribution with the cdf $F(x)$ and the pdf $f(x)$. Then the pdf of U_n (the nth upper record value, $n \geq 1$) is given by,

$$f_{U_n}(x) = \frac{1}{\Gamma(n)}[-log\ (1 - F_{U_n}(x))]^{n-1} f_{U_n}(x), -\infty < x < \infty. \tag{1.8}$$

Also, the corresponding survival function of U_n is given by

$$\overline{F}_{U_n}(x) = \frac{1}{\Gamma(n)} \Gamma_{-log\ \overline{F}(x)}(n). \tag{1.9}$$

where $\Gamma_x(a)$ is the incomplete gamma function given by

$$\Gamma_x(a) = \int_x^\infty u^{a-1} e^{-u} du,\ a, x > 0. \tag{1.10}$$

Let $L_1, L_2, \ldots$ be a sequence of lower record values from a distribution with the cdf $F(x)$ and the pdf $f(x)$. Then the pdf of L_n (the nth lower record value, $n \geq 1$) is given by,

$$f_{L_n}(x) = \frac{1}{\Gamma(n)}[-log\ (f_{L_n}(x))]^{n-1} f_{L_n}(x), -\infty < x < \infty. \tag{1.11}$$

Also, the corresponding survival function of L_n is given by

$$\overline{F}_{L_n}(x) = \frac{1}{\Gamma(n)} \overline{\Gamma}_{-log\ F(x)}(n), \tag{1.12}$$

where $\overline{\Gamma}_x(a)$ is given by

$$\overline{\Gamma}_x(a) = \int_0^x u^{a-1} e^{-u} du,\ a, x > 0. \tag{1.13}$$

For more details on record values one can refer Arnold et al. (1992). Let $R_i, i = 1, 2, \ldots, n$ be the first n upper record values arising from a distribution with cdf $F(x)$ and pdf $f(x)$. Let $D_n = (R_1, R_2, \ldots, R_n)$. Then the joint pdf of D_n is given by

$$f(r_1, r_2, \ldots, r_n) = f(r_n) \prod_{i=1}^{n-1} \frac{f(r_i)}{1 - F(r_i)}, -\infty < r_1 < r_2 < \ldots < r_n < \infty. \tag{1.14}$$

Record values have been studied extensively in the literature. Record values arise naturally in problems such as industrial stress testing, meteorological analysis, hydrology, sporting, stock markets, athletic events and seismology. For some recent applications of record values, see Baratpour et al. (2007), Abbasnejad and Arghami (2011), Chacko and Asha (2021), Muraleedharan and Chacko (2022a), Muraleedharan and Chacko (2022b) and so on.

In this paper, our central aim is to delve into the concept of varentropy concerning record values and establish its mathematical expression for some distributions. Our main goal is to provide a foundational comprehension of varentropy, emphasizing its possible uses and importance within the domain of record values. Varentropy holds importance across finance, healthcare, environmental science, and manufacturing, where it aids in assessing volatility in portfolios, treatment variability, climate fluctuations, and quality control processes. Despite its importance, the existing literature still needs to thoroughly explore the estimation of varentropy using record values, inspiring us to develop a novel estimation approach. This gap highlights the need for new methodologies to understand better and utilize varentropy in practical applications. Furthermore, our study explores practical uses by investigating how maximum likelihood estimation and Bayesian methods can be applied to estimate varentropy, particularly highlighting its significance within the framework of the Pareto distribution. Additionally, real-life data is used for illustration, demonstrating the applicability and relevance of our proposed methods.

The rest of this paper is orgnized as follows: Section 2 examines varentropy for both the nth upper and lower record values and considers an upper bound for the varentropy measure. Section 3 focuses on residual varentropy for the nth upper and lower record values. Moving to Section 4, it explores varentropy of past lifetimes for the nth upper and lower record values. Section 5 covers the estimation of varentropy for the Pareto distribution based on upper record values, presenting the maximum likelihood estimators (MLEs) and Bayes estimators. Finally, in section 6, a real data set is used to illustrate the application of the varentropy measure based on record values. Some concluding remarks are obtained in Section 7.

2. Varentropy of Record Values

In this section, we first examine the varentropy measure for upper and lower record values and then establish some results, including an upper bound for the varentropy measure for these record values.

Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Then analogous to (1.3), we define the varentropy measure of nth upper record value, denoted by $\mathrm{VH}(U_n)$, and is given by

$$\mathrm{VH}(U_n) = Var(-log\ f_{U_n}(U_n)),$$

where $f_{U_n}(x)$ is the pdf of nth upper record given in (1.8).

Similarly, we can express the varentropy measure of nth lower record value, denoted by

VH(L_n), as

$$\text{VH}(L_n) = Var(-log\ f_{L_n}(L_n)), \tag{2.1}$$

where $f_{L_n}(x)$ is the pdf of nth lower record given in (1.11).

In the following theorem, we formulate the varentropy measure of nth upper record.

Theorem 2.1. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let U_n denote the nth upper record. Then the varentropy measure of U_n is given by,

$$\text{VH}(U_n) = (n-1)^2\psi'(n) + C''_n(0) - \left[C'_n(0)\right]^2, \tag{2.2}$$

where $C_n(z) = E[f^z(F^{-1}(1-e^{-U_{n(z+1)+1,1}}))]$, $V_{\alpha,1} \sim G(\alpha, 1)$, ψ' is the trigamma function, $C'_n(0)$ and $C''_n(0)$ are the first and second derivative of $C_n(z)$ with respect to z at $z = 0$. Here $G(a, b)$ denotes a gamma distribution with pdf given by

$$g(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx};\ x > 0, a, b > 0. \tag{2.3}$$

Proof. The moment generating function (mgf) of the $log\ f_{U_n}(U_n)$ is given by,

$$\begin{aligned} M_X(z) &= E(e^{z\ log f_{U_n}(U_n)}) \\ &= E([f_{U_n}(U_n)]^z) \\ &= \int_z [f_{U_n}(x)]^{z+1} dx \\ &= \frac{1}{(\Gamma(n))^{z+1}} \int [-log\ (1-F(x))]^{z(n-1)+n-1} f^{z+1}(x) dx. \end{aligned}$$

On putting $u = -log\ (1-F(x))$, we get

$$\begin{aligned} M_X(z) &= \frac{1}{(\Gamma(n))^{z+1}} \int u^{z(n-1)+n-1} e^{-u} f^z(F^{-1}(1-e^{-u})) du \\ &= \frac{\Gamma(z(n-1)+n)}{(\Gamma(n))^{z+1}} C_n(z), \end{aligned}$$

where $C_n(z) = E[f^z(F^{-1}(1-e^{-U_{n(z+1)+1,1}}))]$. Taking logarithm, we get the cumulant generating function given by

$$K_X(z) = log\ M_X(z) = log\ \Gamma(z(n-1)+n) - (z+1)\ log\ \Gamma(n) + log\ C_n(z).$$

Now, the second derivative of $K_X(z)$ with respect to z at $z = 0$ is given by,

$$\frac{d^2K_X(z)}{dz^2}|_{z=0} = (n-1)^2\psi'(n) + C''_n(0) - \left[C'_n(0)\right]^2.$$

Since $\text{VH}(U_n) = \frac{d^2K_X(z)}{dz^2}|_{z=0}$, we get the result.

Example 2.1. Let $\{X_i, i \geq 1\}$ be a sequence of iid random variables having a uniform distribution with pdf given by

$$f(x) = \frac{1}{b-a},\ a < x < b. \tag{2.4}$$

Here

$$F^{-1}(x) = x(b-a) + a.$$

Therefore

$$f^z(F^{-1}(1-e^{-u})) = \frac{1}{(b-a)^z}$$

Hence

$$C_n'(0) = log\ \frac{1}{b-a} \text{ and } C_n''(0) = \left[log\ \frac{1}{b-a}\right]^2$$

Thus we have,

$$\text{VH}(U_n) = (n-1)^2\psi'(n). \tag{2.5}$$

Therefore, the varentropy of nth upper record for uniform distribution is independent of the distribution parameters.

Example 2.2. Let $\{X_i, i \geq 1\}$ be a sequence of iid random variables having an exponential distribution with pdf given by

$$f(x) = \lambda e^{-\lambda x}; x > 0; \lambda > 0.$$

Here

$$F^{-1}(x) = -\frac{1}{\lambda}log(1-x).$$

Therefore

$$f^z(F^{-1}(1-e^{-u})) = \lambda^z e^{-uz}.$$

Hence

$$C_n'(0) = log\ \lambda - n,$$

and

$$C_n''(0) = [log\ \lambda]^2 - 2n\ log\ \lambda + n^2 - n + 2.$$

Thus we have,

$$\text{VH}(U_n) = (n-1)^2\psi'(n) - n + 2.$$

Therefore, the varentropy of nth upper record for exponential distribution is independent of the scale parameter.

We have drawn the graph of varentropy measure of nth upper record for exponential distribution for different values of n and is given in Figure 1.

Example 2.3. Let $\{X_i, i \geq 1\}$ be a sequence of iid random variables having a common Pareto distribution with pdf given by

$$f(x) = \frac{\lambda}{\sigma}\left(\frac{x}{\sigma}\right)^{-\lambda-1}, x \geq \sigma > 0, \lambda > 0.$$

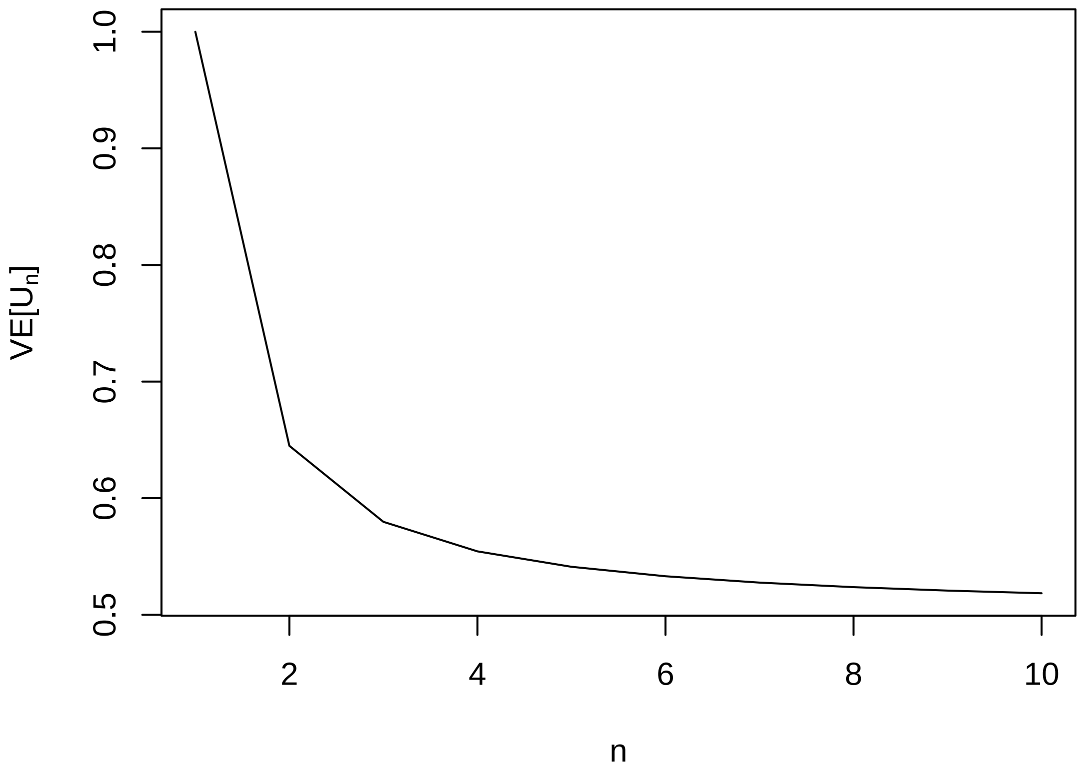

Figure 1: Graph of the varentropy measure of nth upper record for exponential distribution.

Here

$$F^{-1}(x) = \sigma\,(1-x)^{-\frac{1}{\lambda}}.$$

Therefore

$$f^{z-1}(F^{-1}(1-e^{-u})) = \left(\frac{\lambda}{\sigma}\right)^{z} e^{-uz(1+\frac{1}{\lambda})},$$

and hence

$$C_n'(0) = log\,\frac{\lambda}{\sigma} - n\left(1+\frac{1}{\lambda}\right),$$

and

$$C_n''(0) = \left(log\,\frac{\lambda}{\sigma}\right)^2 + n(n+1)\left(1+\frac{1}{\lambda}\right)^2 - 2(n-1)\left(1+\frac{1}{\lambda}\right) - 2n\left(log\,\frac{\lambda}{\sigma}\right)\left(1+\frac{1}{\lambda}\right).$$

Thus we have,

$$\text{VH}(U_n) = (n-1)^2\psi'(n) + n\left(1+\frac{1}{\lambda}\right)^2 - 2(n-1)\left(1+\frac{1}{\lambda}\right).$$

We have drawn the graph of varentropy measure of nth upper record for Pareto distribution for different values of λ and are given in Figure 2.

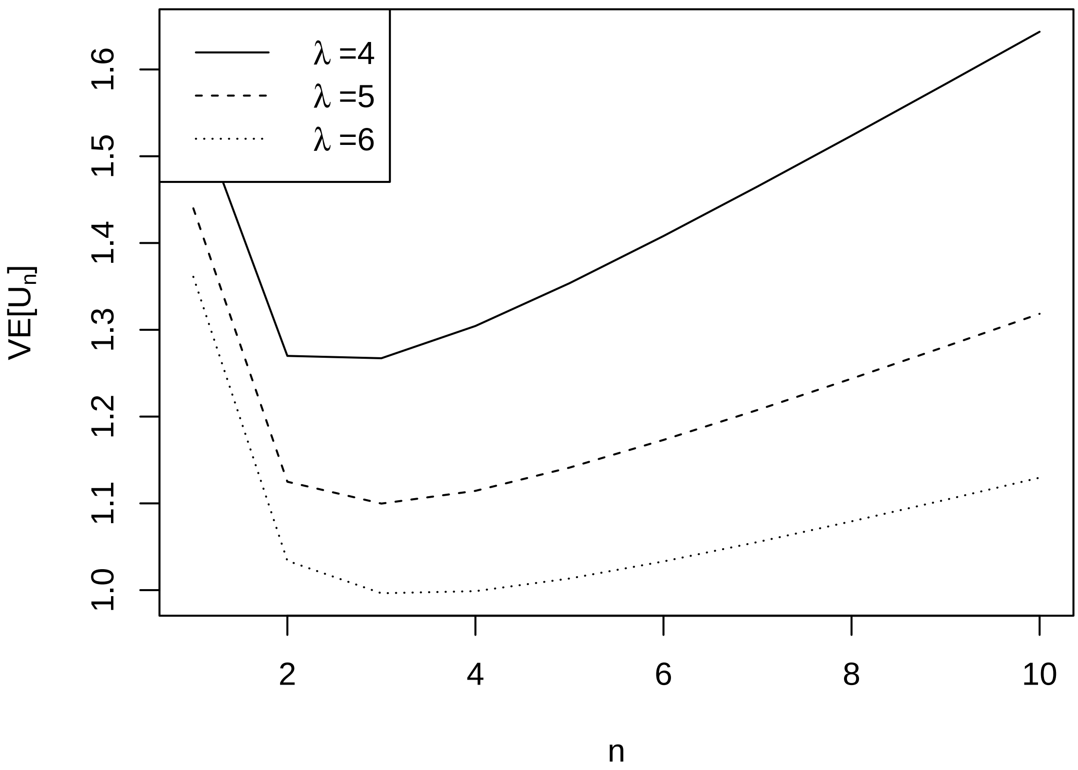

Figure 2: Graph of the varentropy measure of nth upper record for Pareto distribution for different values of λ.

In the following theorem, we obtain an upper bound of the varentropy measure nth upper record.

Theorem 2.2. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common distribution function $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let U_n denote the nth upper record. Then an upper bound for the varentropy measure of U_n is given by,

$$\text{VH}(U_n) \leq \left(\int (log\ f_{U_n}(U_n) + H(U_n))^4 dx \right)^{\frac{1}{2}} (E(f_{U_n}(U_n)))^{\frac{1}{2}}, \tag{2.6}$$

where $H(U_n)$ is the entropy of U_n.

Proof. We have

$$\begin{aligned} \text{VH}(X) &= E[-log\ f_{U_n}(U_n) - H(U_n)]^2 \\ &= \int (log\ f_{U_n}(x) + H(U_n))^2 f_{U_n}(x) dx. \end{aligned}$$

By applying Cauchy-Schwartz inequality, we have

$$\begin{aligned}\mathrm{VH}(U_n) &\leq \left(\int (log\ f_{U_n}(x) + H(U_n))^4 dx\right)^{\frac{1}{2}} \left(\int f_{U_n}^2(x)dx\right)^{\frac{1}{2}} \\ &= \left(\int (log\ f_{U_n}(x) + H(U_n))^4 dx\right)^{\frac{1}{2}} \left(E(f_{U_n}(U_n))\right)^{\frac{1}{2}}.\end{aligned}$$

Hence the theorem.

In the following theorem, we obtain the varentropy measure of nth lower record.

Theorem 2.3. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let L_n denote the nth lower record. Then the varentropy measure of L_n is given by,

$$\mathrm{VH}(L_n) = (n-1)^2\psi'(n) + D_n^{''}(0) - \left[D_n'(0)\right]^2, \tag{2.7}$$

where $D_n(z) = E[f^z(F^{-1}(e^{-V_{n(z+1)+1,1}}))]$, $V_{\alpha,1} \sim G(\alpha, 1)$, ψ' is the trigamma function, $D_n^{'}(0)$ and $D_n^{''}(0)$ are the first and second derivative of $D_n(z)$ with respect to z at $z = 0$. Here $G(a, b)$ denotes a gamma distribution with pdf given in (2.3).

Proof. The mgf of the $log\ f_{L_n}(L_n)$ is given by,

$$\begin{aligned}M_X(z) &= E(e^{z\ log f_{L_n}(L_n)}) \\ &= E([f_{L_n}(L_n)]^{z+1}) \\ &= \int_z [f_{L_n}(x)]^{z+1} dx \\ &= \frac{1}{(\Gamma(n))^{z+1}} \int [-log\ (1-F(x))]^{z(n-1)+n-1} f^z(x)dx.\end{aligned}$$

On putting $v = -log\ (F(x))$, we get

$$\begin{aligned}M_X(z) &= \frac{1}{(\Gamma(n))^{z+1}} \int v^{z(n-1)+n-1} e^{-v} f^z(F^{-1}(e^{-v}))dv \\ &= \frac{\Gamma(z(n-1)+n)}{(\Gamma(n))^{z+1}} D_n(z),\end{aligned}$$

where $D_n(z) = E[f^z(F^{-1}(e^{-V_{n(z+1)+1,1}}))]$. Taking logarithm, we get the cumulant generating function as

$$K_X(z) = log\ M_X(z) = log\ \Gamma(z(n-1)+n) - (z+1)\ log\ \Gamma(n) + log\ D_n(z).$$

Now, the second derivative of $K_X(z)$ with respect to z at $z = 0$ is given by,

$$\frac{d^2K_X(z)}{dz^2}|_{z=0} = (n-1)^2\psi'(n) + D_n''(0) - \left[D_n'(0)\right]^2.$$

Since $\mathrm{VH}(L_n) = \dfrac{d^2K_X(z)}{dz^2}|_{z=0}$, we get the result.

Example 2.4. Let $\{X_i, i \geq 1\}$ be a sequence of iid random variables having a standard power distribution with pdf given by

$$f(x) = \beta x^{\beta-1}, 0 < x < 1.$$

Here

$$F^{-1}(x) = x^{\frac{1}{\beta}}.$$

Therefore

$$f^z(F^{-1}(e^{-u})) = \beta^z e^{-uz(1-\frac{1}{\beta})}.$$

Hence

$$D_n'(0) = log\ \beta - n\left(1 - \frac{1}{\beta}\right),$$

and

$$D_n''(0) = (log\ \beta)^2 + n(n+1)\left(1 - \frac{1}{\beta}\right)^2 - 2(n-1)\left(1 - \frac{1}{\beta}\right) - 2n\,(log\ \beta)\left(1 - \frac{1}{\beta}\right).$$

Thus we have,

$$\mathrm{VH}(L_n) = (n-1)^2\psi'(n) + n\left(1 - \frac{1}{\beta}\right)^2 - 2(n-1)\left(1 - \frac{1}{\beta}\right).$$

We have drawn the graph of varentropy measure of nth lower record for standard power distribution for different values of β and are given in Figure 3.

Example 2.5. Let $\{X_i, i \geq 1\}$ be a sequence of iid random variables having a uniform distribution with pdf given in (2.4). By solving, we have,

$$\mathrm{VH}(L_n) = (n-1)^2\psi'(n). \tag{2.8}$$

Remark 1. From (2.5) and (2.8), for uniform distribution, the varentropy of nth upper record is same as the varentropy of nth lower record and is independent of the distribution parameters.

In the following theorem, we obtain an upper bound of the varentropy measure of nth lower record.

Theorem 2.4. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let L_n denote the nth lower record. Then an upper bound for the varentropy measure of L_n is given by,

$$\mathrm{VH}(L_n) \leq \left(\int (log\ f_{L_n}(x) + H(L_n))^4 dx\right)^{\frac{1}{2}} (E(f_{L_n}(L_n)))^{\frac{1}{2}}, \tag{2.9}$$

where $H(L_n)$ is the entropy of L_n.

Proof. The proof is omitted since it is similar to that of Theorem 2.2.

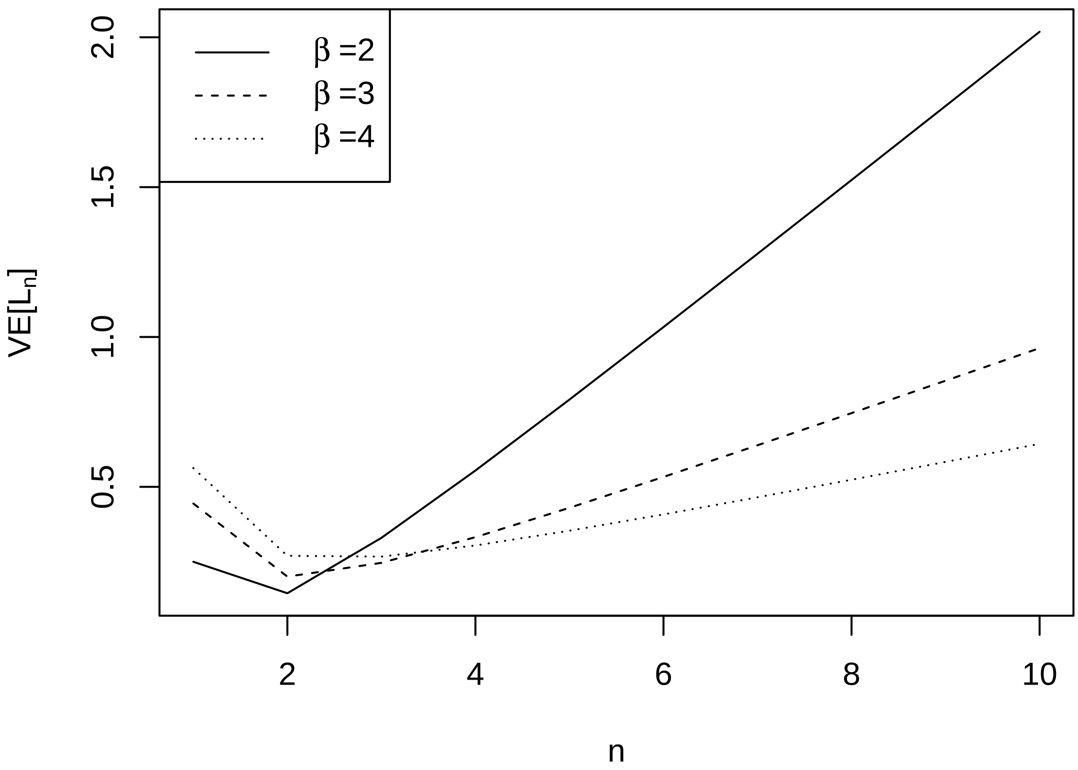

Figure 3: Graph of the varentropy measure of nth lower record for standard power distribution for different values of β.

3. Residual Varentropy of Record Values

In this section, we examine the residual varentropy measure of lower and upper record values and then establish some results for this measure.

Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Then analogous to (1.5), we define the residual varentropy measure of nth upper record value, denoted by RVH(U_n), and is given by

$$\mathrm{RVH}(U_n) = Var\left(-log\,\frac{f_{U_n}(U_n^{(t)})}{\overline{F}_{U_n}(t)}\right),$$

where $U_n^{(t)}$ is the residual lifetime of the random variable U_n at time t, with density function given by

$$f_{U_n^{(t)}}(x) = \frac{f_{U_n}(x)}{\overline{F}_{U_n}(t)},\ x \geq t > 0,$$

and $\overline{F}_{U_n}(t)$ is the corresponding survival function defined in (1.9).

Similarly, we can express the residual varentropy measure of nth lower record value, denoted by RVH(L_n), as

$$\mathrm{RVH}(L_n) = Var\left(-log\,\frac{f_{L_n}(L_n^{(t)})}{\overline{F}_{L_n}(t)}\right),$$

where $L_n^{(t)}$ is the residual lifetime of the random variable L_n at time t, with density function given by

$$f_{L_n^{(t)}}(x) = \frac{f_{L_n}(x)}{\overline{F}_{L_n}(t)},\ x \geq t > 0,$$

and $\overline{F}_{L_n}(t)$ is the corresponding survival function defined in (1.12).

In the following theorem, we obtain the residual varentropy measure of nth upper record.

Theorem 3.1. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let U_n denote the nth upper record. Then the residual varentropy measure of U_n is given by,

$$\mathrm{RVH}(U_n) = \eta'' + C_n''(0) - \left[C_n'(0)\right]^2, \tag{3.1}$$

where η'' is the second derivative of $log\,\Gamma_{-log\,\overline{F}(t)}(z(n-1)+n, 1)$, $C_n(z) = E[f^{z-1}(F^{-1}(1 - e^{-V^{(t)}_{z(n-1)+n,1}}))]$, $V^{(t)}_{\alpha,1} \sim G_t(\alpha, 1)$, $C_n'(0)$ and $C_n''(0)$ are the first and second derivative of $C_n(z)$ with respect to z at $z = 0$. Here $G_t(a, 1)$ denotes a truncated gamma distribution with pdf given by

$$h_1(x) = \frac{1}{\Gamma_t(a,1)} x^{a-1} e^{-x};\ 0 < t < x, a > 0, \tag{3.2}$$

and $\Gamma_t(a, 1)$ is given by

$$\Gamma_t(a,1) = \int_t^\infty x^{a-1} e^{-x} dx. \tag{3.3}$$

Proof. The mgf of $log\,\frac{f_{U_n}(U_n^{(t)})}{\overline{F}_{U_n}(t)}$ is given by

$$\begin{aligned}
M_{X_t}(z) &= E\left[e^{z\,log\frac{f_{U_n}(U_n^{(t)})}{\overline{F}_{U_n}(t)}}\right] \\
&= E\left[\left(\frac{f_{U_n}(U_n^{(t)})}{\overline{F}_{U_n}(t)}\right)^z\right] \\
&= \int_t^\infty \left(\frac{f_{U_n}(x)}{\overline{F}_{U_n}(t)}\right)^{z+1} dx \\
&= \frac{1}{[\Gamma_{-log\,\overline{F}(t)}(n,1)]^{z+1}} \int_t^\infty [-log\,(1-F(x))]^{z(n-1)+n-1} f^{z+1}(x) dx.
\end{aligned}$$

On putting $u = -log\,(1 - F(x))$, we get

$$\begin{aligned}
M_{X_t}(z) &= \frac{1}{[\Gamma_{-log\,\overline{F}(t)}(n,1)]^{z+1}} \int_{-log\,\overline{F}(t)}^\infty u^{z(n-1)+n-1} e^{-u} f^z(F^{-1}(1-e^{-u})) du \\
&= \frac{\Gamma_{-log\,\overline{F}(t)}(z(n-1)+n, 1)}{[\Gamma_{-log\,\overline{F}(t)}(n,1)]^{z+1}} C_n(z).
\end{aligned}$$

where $C_n(z) = E[f^z(F^{-1}(1-e^{-V^{(t)}_{z(n-1)+n,1}}))]$. Taking logarithm, we get the cumulant generating function as

$$K_{X_t}(z) = log\ M_{X_t}(z) = log\ \Gamma_{-log\ \overline{F}(t)}(z(n-1)+n,1) \\ -(z+1)\ log\ (\Gamma_{-log\ \overline{F}(t)}(n,1)) + log\ C_n(z).$$

Now, the second derivative of $K_{X_t}(z)$ with respect to z at $z=0$ is given by,

$$\frac{d^2 M_{X_t}(z)}{dz^2}|_{z=0} = \eta'' + C''_n(0) - \left[C'_n(0)\right]^2,$$

where η'' is the second derivative of $log\ \Gamma_{-log\ \overline{F}(t)}(z(n-1)+n,1)$.
Since RVH$(U_n) = \dfrac{d^2 K_{X_t}(z)}{dz^2}|_{z=0}$, we get the result.

In the following theorem, we derive the expression for residual varentropy measure of nth lower record.

Theorem 3.2. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let L_n denote the nth lower record. Then the residual varentropy measure of L_n is given by,

$$\text{RVH}(L_n) = \xi'' + D''_n(0) - \left[D'_n(0)\right]^2, \tag{3.4}$$

where ξ'' is the second derivative of $log\ \overline{\Gamma}_{-log\ F(t)}(z(n-1)+n,1)$,
$D_n(z) = E[f^{z-1}(F^{-1}(e^{-V^{(t)}_{z(n-1)+n,1}}))]$, $V^{(t)}_{\alpha,1} \sim \overline{G}_t(\alpha,1)$, $D'_n(0)$ and $D''_n(0)$ are the first and second derivative of $D_n(z)$ with respect to z at $z=0$. Here $\overline{G}_t(a,1)$ denotes a truncated gamma distribution with pdf given by

$$h_2(x) = \frac{1}{\overline{\Gamma}_t(a,1)} x^{a-1} e^{-x};\ 0 < x < t, a > 0, \tag{3.5}$$

and $\overline{\Gamma}_t(a,1)$ is given by

$$\overline{\Gamma}_t(a,1) = \int_0^t x^{a-1} e^{-x} dx. \tag{3.6}$$

Proof. The mgf of $log\ \dfrac{f_{L_n}(L_n^{(t)})}{\overline{F}_{L_n}(t)}$ is given by

$$\begin{aligned} M_{X_t}(z) &= E\left[e^{z\ log \frac{f_{L_n}(L_n^{(t)})}{\overline{F}_{L_n}(t)}}\right] \\ &= E\left[\left(\frac{f_{L_n}(L_n^{(t)})}{\overline{F}_{L_n}(t)}\right)^z\right] \\ &= \int_t^\infty \left(\frac{f_{L_n}(x)}{\overline{F}_{L_n}(t)}\right)^{z+1} dx \\ &= \frac{1}{(\overline{\Gamma}_{-log\ F(t)}(n,1))^{z+1}} \int_t^\infty [-log\ F(x)]^{z(n-1)+n-1} f^{z+1}(x) dx. \end{aligned}$$

On putting $v = -log\ F(x)$, we get

$$M_{X_t}(z) = \frac{1}{(\overline{\Gamma}_{-log\ F(t)}(n,1))^{z+1}} \int_0^{-log\ F(t)} v^{z(n-1)+n-1} e^{-v} f^z(F^{-1}(e^{-v}))dv$$
$$= \frac{\overline{\Gamma}_{-log\ F(t)}(z(n-1)+n,1)}{(\overline{\Gamma}_{-log\ F(t)}(n,1))^{z+1}} D_n(z).$$

where $D_n(z) = E[f^z(F^{-1}(e^{-V^{(t)}_{z(n-1)+n,1}}))]$. Taking logarithm, we get the cumulant generating function as

$$K_{X_t}(z) = log\ M_{X_t}(z) = log\ \overline{\Gamma}_{-log\ F(t)}(z(n-1)+n,1) - (z+1)\ log\ (\overline{\Gamma}_{-log\ F(t)}(n,1)) + log\ D_n(z).$$

Now, the second derivative of $K_{X_t}(z)$ with respect to z at $z=0$ is given by,

$$\frac{d^2K_{X_t}(z)}{dz^2}|_{z=0} = \xi'' + D''_n(0) - \left[D'_n(0)\right]^2,$$

where ξ'' is the second derivative of $log\ \overline{\Gamma}_{-log\ F(t)}(z(n-1)+n,1)$.

Since $\text{RVH}(L_n) = \frac{d^2K_{X_t}(z)}{dz^2}|_{z=0}$, we get the result.

4. Past Varentropy of Record Values

In this section, we examine the past varentropy measure of lower and upper record values and then establish some results for this measure.

Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common distribution function $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Then analogous to (1.7), we define the past varentropy measure of nth upper record value, denoted by $\text{PVH}(U_n)$, and is given by

$$\text{PVH}(U_n) = Var\left(-log\ \frac{f_{U_n}(U_n^{[t]})}{F_{U_n}(t)}\right),$$

where $U_n^{[t]}$ is the past lifetime of the random variable U_n at time t, with density function given by

$$f_{U_n^{[t]}}(x) = \frac{f_{U_n}(x)}{F_{U_n}(t)},\ 0 \leq x < t,$$

and $F_{U_n}(t)$ is the corresponding cdf.

Similarly, we can express the past varentropy measure of nth lower record value, denoted by $\text{PVH}(L_n)$, as

$$\text{PVH}(L_n) = Var\left(-log\ \frac{f_{L_n}(L_n^{(t)})}{F_{L_n}(t)}\right),$$

where $L_n^{[t]}$ is the past lifetime of the random variable L_n at time t, with density function given by

$$f_{L_n^{[t]}}(x) = \frac{f_{L_n}(x)}{F_{L_n}[t]},\ 0 \leq x < t,$$

and $F_{L_n}(t)$ is the corresponding cdf.

In the following theorem, we obtain the past varentropy measure of nth upper record.

Theorem 4.1. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let U_n denote the nth upper record. Then the past varentropy measure of U_n is given by,

$$\text{PVH}(U_n) = \eta'' + C_n^{''}(0) - \left[C_n^{'}(0)\right]^2, \tag{4.1}$$

where η'' is the second derivative of $log\ \overline{\Gamma}_{-log\ \overline{F}(t)}(z(n-1)+n, 1)$,
$C_n(z) = E[f^{z-1}(F^{-1}(1 - e^{-V^{(t)}_{z(n-1)+n,1}}))]$, $V^{(t)}_{\alpha,1} \sim \overline{G}_t(\alpha, 1)$, $C_n^{'}(0)$ and $C_n^{''}(0)$ are the first and second derivative of $C_n(z)$ with respect to z at $z = 0$. Here $\overline{G}_t(a, 1)$ denotes a truncated gamma distribution with pdf given in (3.5).

Proof. The mgf of $log\ \dfrac{f_{U_n}(U_n^{[t]})}{F_{U_n}(t)}$ is given by

$$\begin{aligned}
M_{X_{[t]}}(z) &= E\left[e^{z\ log \frac{f_{U_n}(U_n^{[t]})}{F_{U_n}(t)}}\right] \\
&= E\left[\left(\frac{f_{U_n}(U_n^{[t]})}{F_{U_n}(t)}\right)^z\right] \\
&= \int_0^t \left(\frac{f_{U_n}(x)}{F_{U_n}(t)}\right)^{z+1} dx \\
&= \frac{1}{[\overline{\Gamma}_{-log\ \overline{F}(t)}(n,1)]^{z+1}} \int_0^t [-log\ (1-F(x))]^{z(n-1)+n-1} f^{z+1}(x) dx.
\end{aligned}$$

On putting $u = -log\ (1 - F(x))$, we get

$$\begin{aligned}
M_{X_{[t]}}(z) &= \frac{1}{[\overline{\Gamma}_{-log\ \overline{F}(t)}(n,1)]^{z+1}} \int_0^{-log\ \overline{F}(t)} u^{z(n-1)+n-1} e^{-u} f^z(F^{-1}(1-e^{-u})) du \\
&= \frac{\overline{\Gamma}_{-log\ \overline{F}(t)}(z(n-1)+n,1)}{[\overline{\Gamma}_{-log\ \overline{F}(t)}(n,1)]^{z+1}} C_n(z).
\end{aligned}$$

where $C_n(z) = E[f^z(F^{-1}(1 - e^{-V^{(t)}_{z(n-1)+n,1}}))]$. Taking logarithm, we get the cumulant generating function as

$$\begin{aligned}
K_{X_{[t]}}(z) = log\ M_{X_t}(z) &= log\ \overline{\Gamma}_{-log\ \overline{F}(t)}(z(n-1)+n, 1) \\
&\quad - (z+1)\ log\ (\overline{\Gamma}_{-log\ \overline{F}(t)}(n,1)) + log\ C_n(z).
\end{aligned}$$

Now, the second derivative of $K_{X_{[t]}}(z)$ with respect to z at $z = 0$ is given by,

$$\frac{d^2 M_{X_{[t]}}(z)}{dz^2}|_{z=0} = \eta'' + C_n''(0) - \left[C_n'(0)\right]^2,$$

where η'' is the second derivative of $log\ \overline{\Gamma}_{-log\ \overline{F}(t)}(z(n-1)+n, 1)$.

Since $\text{PVH}(U_n) = \dfrac{d^2 K_{X_{[t]}}(z)}{dz^2}|_{z=0}$, we get the result.

In the following theorem, we derive the expression for past varentropy measure of nth lower record.

Theorem 4.2. Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables from a distribution with common cdf $F(x)$, pdf $f(x)$ and quantile function $F^{-1}(\cdot)$. Let L_n denote the nth lower record. Then the past varentropy measure of L_n is given by,

$$\text{PVH}(L_n) = \xi'' + D_n^{''}(0) - \left[D_n^{'}(0)\right]^2, \tag{4.2}$$

where ξ'' is the second derivative of $log\, \Gamma_{-log\, F(t)}(z(n-1)+n, 1)$,
$D_n(z) = E[f^{z-1}(F^{-1}(e^{-V^{(t)}_{z(n-1)+n,1}}))]$, $V^{(t)}_{\alpha,1} \sim G_t(\alpha, 1)$ and $D_n^{'}(0)$ and $D_n^{''}(0)$ are the first and second derivative of $D_n(z)$ with respect to z at $z = 0$. Here $G_t(a, 1)$ denotes a truncated gamma distribution with pdf given in (3.2).

Proof. The mgf of $log\, \dfrac{f_{L_n}(L_n^{[t]})}{F_{L_n}(t)}$ is given by

$$\begin{aligned}
M_{X_{[t]}}(z) &= E\left[e^{z\, log \frac{f_{L_n}(L_n^{[t]})}{F_{L_n}(t)}}\right] \\
&= E\left[\left(\frac{f_{L_n}(L_n^{[t]})}{F_{L_n}(t)}\right)^z\right] \\
&= \int_0^t \left(\frac{f_{L_n}(x)}{F_{L_n}(t)}\right)^{z+1} dx \\
&= \frac{1}{(\Gamma_{-log\, F(t)}(n,1))^{z+1}} \int_0^t [-log\, F(x)]^{z(n-1)+n-1} f^{z+1}(x) dx.
\end{aligned}$$

On putting $v = -log\, F(x)$, we get

$$\begin{aligned}
M_{X_{[t]}}(z) &= \frac{1}{(\Gamma_{-log\, F(t)}(n,1))^{z+1}} \int_{-log\, F(t)}^{\infty} v^{z(n-1)+n-1} e^{-v} f^z(F^{-1}(e^{-v})) dv \\
&= \frac{\Gamma_{-log\, F(t)}(z(n-1)+n, 1)}{(\Gamma_{-log\, F(t)}(n,1))^{z+1}} D_n(z).
\end{aligned}$$

where $D_n(z) = E[f^z(F^{-1}(e^{-V^{(t)}_{z(n-1)+n,1}}))]$. Taking logarithm, we get the cumulant generating function as

$$\begin{aligned}
K_{X_{[t]}}(z) = log\, M_{X_{[t]}}(z) = log\, \Gamma_{-log\, F(t)}(z(n-1)+n, 1)& \\
- (z+1)\, log\, (\Gamma_{-log\, F(t)}(n,1)) + log\, D_n(z)&.
\end{aligned}$$

Now, the second derivative of $K_{X_{[t]}}(z)$ with respect to z at $z = 0$ is given by,

$$\frac{d^2 K_{X_{[t]}}(z)}{dz^2}|_{z=0} = \xi'' + D_n''(0) - \left[D_n'(0)\right]^2,$$

where ξ'' is the second derivative of $log\, \Gamma_{-log\, F(t)}(z(n-1)+n, 1)$.
Since $\text{PVH}(L_n) = \dfrac{d^2 K_{X_{[t]}}(z)}{dz^2}|_{z=0}$, we get the result.

5. Estimation of varentropy measure for Pareto Distribution based on record values

In this section, we consider the estimation of varentropy measure for Pareto distribution based on upper record values. We obtain the maximum likelihood estimators(MLEs) and Bayes estimators of varentropy measure using upper record values. The Pareto distribution has cdf given by

$$F(x|\lambda,\sigma) = 1 - \left(\frac{x}{\sigma}\right)^{-\lambda}, \lambda > 0, x > \sigma. \tag{5.1}$$

The pdf corresponding to the above cdf is given by

$$f(x|\lambda,\sigma) = \frac{\lambda}{\sigma}\left(\frac{x}{\sigma}\right)^{-(\lambda+1)}. \tag{5.2}$$

The Pareto distribution is widely used to model situations where a small number of extreme events or observations significantly influence the overall distribution, prevalent in various fields due to its ability to capture heavy-tailed distributions. Also, it aids in understanding the variability and uncertainty inherent in systems where extreme events play a crucial role.The varentropy measure of the nth upper record value for the Pareto distribution with cdf given in (5.1) is given by

$$\text{VH}(\lambda,\sigma) = (n-1)^2\psi'(n) + n\left(1+\frac{1}{\lambda}\right)^2 - 2(n-1)\left(1+\frac{1}{\lambda}\right). \tag{5.3}$$

Recently, Chacko and Grace (2023) obtained estimators for the information generating function of a Weibull distribution based on upper record values. Chacko and Asha (2021) obtained estimators for the entropy functions of a Weibull distribution based on upper record values and Chacko and Asha (2018) obtained estimators for the entropy functions of a generalized exponential distribution based on lower record values.

5.1. Maximum Likelihood Estimation

In this subsection, we obtain the MLEs of varentropy measure for the Pareto distribution using upper record values. Let $R_i, i = 1, 2, \ldots, n$ be the first n upper record values arising from Pareto distribution with cdf given in (5.1). Let $D_n = (R_1, R_2, \ldots, R_n)$. Then from (1.14) the likelihood function is given by

$$L(\lambda,\sigma|d_n) = \lambda^n\left(\frac{\sigma}{r_n}\right)^{\lambda}\prod_{i=1}^{n}\frac{1}{r_i},\ r_1 < r_2 < \ldots, r_n,$$

where $d_n = (r_1, r_2, \ldots, r_n)$. The logarithm of the likelihood function is given by

$$log\ L(\lambda,\sigma|d_n) = n\ log\ \lambda + \lambda\ log\ \sigma - \lambda\ log\ r_n - \sum_{i=1}^{n} log\ r_i.$$

Since $log\ \sigma$ is monotonically increasing, we can maximize the likelihood by setting $\hat{\sigma}$ as high as possible. Since $\sigma \leq r_1 < r_2 < \ldots < r_n$, we have $\hat{\sigma} = r_1$, the first record value.

For λ, we differentiate $log\ L(\lambda, \sigma|d_n)$ with respect to λ and equates to zero, we get,

$$\frac{\partial logL}{\partial \lambda} = \frac{n}{\lambda} + log\ \hat{\sigma} - log\ r_n = 0. \tag{5.4}$$

Therefore the MLE of λ is given by

$$\hat{\lambda} = \frac{n}{log\ \frac{r_n}{\hat{\sigma}}}.$$

Then by the invariant property of MLE, the MLE of varentropy measure for Pareto distribution based on nth upper record values is given by

$$\widehat{\mathrm{VH}}_{MLE} = (n-1)^2\psi'(n) + n\left(1 + \frac{1}{\hat{\lambda}}\right)^2 - 2(n-1)\left(1 + \frac{1}{\hat{\lambda}}\right). \tag{5.5}$$

5.2. Bayesian Estimation

In this subsection, we consider the Bayesian estimation of the varentropy measure for the Pareto distribution based on upper record values. Here, we consider Bayesian estimation of varentropy measure under symmetric as well as asymmetric loss functions. For a symmetric loss function we consider the squared error loss (SEL) function and for assymetric loss functions we consider LINEX loss function. The Bayes estimate of any parameter μ under SEL is the posterior mean of μ. The Bayes estimate of μ under LINEX loss function can be obtained as

$$\hat{\mu}_{LB} = -\frac{1}{h} log\{E_\mu(e^{-h\mu}|\underline{x})\}, h \neq 0,$$

provided $E_\mu(.)$ exists. Let $R_i, i = 1, 2, \ldots, n$ be the first n upper record values arising from Pareto distribution with pdf given in (5.2). Then from (1.14) the likelihood function is given by

$$L(\lambda, \sigma|d_n) = \lambda^n \left(\frac{\sigma}{r_n}\right)^\lambda \prod_{i=1}^{n} \frac{1}{r_i},$$

where $d_n = (r_1, r_2, \ldots, r_n)$. Assume that the prior distributions of λ and σ follow independent gamma distributions with density functions respectively given by

$$\pi_1(\lambda|a, b) = \frac{b^a}{\Gamma a}\lambda^{a-1}e^{-b\lambda}; a > 0, b > 0,$$

and

$$\pi_2(\sigma|c, d) = \frac{d^c}{\Gamma c}\sigma^{c-1}e^{-d\sigma}; c > 0, d > 0.$$

Thus, the joint prior distribution of λ and σ is given by

$$\pi(\lambda, \sigma) = \frac{b^a d^c}{\Gamma a \Gamma c}\lambda^{a-1}\sigma^{c-1}e^{-b\lambda}e^{-d\sigma}.$$

Then, the joint posterior density of λ and σ given $D_n = d_n$ can be written as

$$\pi^*(\lambda, \sigma|d_n) = \frac{L(\lambda, \sigma|d_n)\pi(\lambda, \sigma)}{\int\int L(\lambda, \sigma|d_n)\pi(\lambda, \sigma)d\lambda d\sigma}. \tag{5.6}$$

Therefore, the Bayes estimate of VH(λ, σ) of λ and σ under SEL and LL are respectively given by

$$\widehat{\text{VH}}_S = \frac{\int\int \text{VH}(\lambda) L(\lambda, \sigma|d_n)\pi(\lambda, \sigma) d\lambda d\sigma}{\int\int L(\lambda, \sigma|d_n)\pi(\lambda, \sigma) d\lambda d\sigma}, \tag{5.7}$$

and

$$\widehat{\text{VH}}_L = -\frac{1}{h} log \left[\frac{\int\int e^{-h\text{VH}(\lambda)} L(\lambda, \sigma|d_n)\pi(\lambda, \sigma) d\lambda d\sigma}{\int\int L(\lambda, \sigma|d_n)\pi(\lambda, \sigma) d\lambda d\sigma} \right]. \tag{5.8}$$

It is not possible to compute (5.7)-(5.8) explicitly. Thus we propose MCMC method to find the Bayes estimates for the varentropy measure given in (5.3).

5.3. MCMC Method

In this subsection, we consider the MCMC method to generate samples from the posterior distributions and then find the Bayes estimates for varentropy measure. The joint posterior distribution given in (5.6) can be written as

$$\pi^*(\lambda, \sigma|d_n) \propto \lambda^{n+a-1} \sigma^{\lambda+c-1} e^{-(\lambda b + d\sigma + \lambda \, log \, r_n + \sum_{i=1}^{n} log \, r_i)}. \tag{5.9}$$

From (5.9) the conditional posterior distribution of λ given σ and d_n is given by

$$\pi_1^*(\lambda|\sigma, d_n) \propto \lambda^{n+a-1} e^{-\lambda(b + log \, r_n - log \, \sigma)}. \tag{5.10}$$

Again from (5.9), the conditional posterior distribution of σ given λ and d_n is given by

$$\pi_2^*(\sigma|\lambda, d_n) \propto \sigma^{\lambda+c-1} e^{-d\sigma}. \tag{5.11}$$

Thus from (5.10) we can see that, the conditional posterior distribution of λ follows a Gamma distribution with parameters $(n+a)$ and $(b+log\ r_n - log\ \sigma)$. That is, $\lambda \sim Gamma(n+a, b+ log\ r_n - log\ \sigma)$. Also from (5.11) we can see that, the conditional posterior distribution of σ follows a Gamma distribution with parameters $(\lambda+c)$ and (d). That is, $\sigma \sim Gamma(\lambda+c, d)$. Therefore one can easily generate sample from the conditional posterior distributions of λ and σ.

By setting initial values $\sigma^{(0)}$ and $\lambda^{(0)}$, let $\sigma^{(t)}$ and $\lambda^{(t)}$, $t = 1, 2, \ldots, N$ be the observations generated from (5.11) and (5.10) respectively. Then the Bayes estimator of varentropy measure given in (5.3) under SEL and LL, by taking first m iterations as burn-in period, are respectively given by

$$\widehat{\text{VH}}_{SEL} = \frac{1}{N-m} \sum_{t=m+1}^{N} \text{VH}(\lambda^{(t)}), \tag{5.12}$$

and

$$\widehat{\text{VH}}_{LL} = -\frac{1}{h} log \left[\frac{1}{N-m} \sum_{t=m+1}^{N} e^{-h\text{VH}(\lambda^{(t)})} \right], \tag{5.13}$$

where VH$(\lambda^{(t)})$ is given in (5.3).

5.4. Simulation Study

In this subsection, we carry out a simulation study for finding the efficiency of different estimates developed in subsection (5.3). First we obtain the MLEs for varentropy measure using (5.5). We have obtained the MLEs and the corresponding MSE of MLEs for different values of n using 1000 simulated samples for different combinations of λ and fixed $\sigma = 1$ and are given in Table 1. For the simulation studies for Bayes estimators we take the hyper parameters for the prior distributions of λ and σ as $a = 1, b = 0.5, c = 1$ and $d = 0.5$. We have obtained the Bayes estimators for varentropy measure of Pareto distribution using upper record values under SEL and LL functions using MCMC method. For that we use the following algorithm.

1. Generate upper record values from Pareto distribution with parameters λ and σ.
2. Calculate estimators of varentropy measure using the generated upper record values using MCMC method as describe below.
   (a) Start with initial values $\sigma^{(0)}$ and $\lambda^{(0)}$.
   (b) Set $t = 1$.
   (c) Generate $\lambda^{(t)}$ from $Gamma(n + a, b + log\, r_n - log\, \sigma^{(t-1)})$.
   (d) Generate $\sigma^{(t)}$ from $Gamma(\lambda^{(t)} + c, d)$.
   (e) Calculate $\widehat{\text{VH}}(\lambda^{(t)})$ using (5.3).
   (f) Set $t = t + 1$.
   (g) Repeat steps (c) to (f) for $N = 50,000$ times.
   (h) Calculate the Bayes estimators for the varentropy measure $\text{VH}(\lambda)$ using (5.12) to (5.13) by taking burn-in-period $m = 5000$.
3. Repeat the steps 1 and 2 for 500 times.
4. Calculate the Bayes estimates and the corresponding MSEs of the estimators.

Repeat the simulation study for $n = 10, 15, 20$ and for different values of λ and σ. The ML estimates, the Bayes estimators and the corresponding MSE for varentropy measure under SEL and LL functions for classical records are given in Tables 1, 2 and 3 . From the tables we have the following inference:

1. The MSEs of all estimators decrease when n increases.
2. The MSEs corresponding to the MLEs are smaller than that of Bayes estimates.
3. Among the Bayes estimators, estimators under LL function have the least MSE.

Table 1: The estimate and the corresponding MSE for maximum likelihood estimator and Bayes estimator for varentropy measure of Pareto distribution when $\sigma = 1$

n	λ	VH	MLE		SEL		LL	
			$\hat{VH}_{MLE}$	MSE	$\hat{VH}_{SEL}$	MSE	$\hat{VH}_{LL}$	MSE
10	4	1.70413	2.21894	0.44516	2.65510	1.45847	1.45605	0.71655
10	4.5	1.50429	1.91995	0.30703	2.54445	1.36692	1.26845	0.70024
10	5	1.35663	1.68450	0.21193	2.52839	1.80706	1.06833	0.62262
10	5.5	1.24391	1.50942	0.14833	2.26984	1.13474	1.03711	0.55441
10	6	1.15552	1.37523	0.10709	1.92525	0.61113	1.01083	0.49893
15	4	2.01110	1.91747	0.41386	2.46065	1.03190	1.74845	0.65502
15	4.5	1.74567	1.66541	0.26617	2.09618	0.78338	1.35057	0.61040
15	5	1.55110	1.47885	0.18570	1.80896	0.53952	1.19306	0.47415
15	5.5	1.40366	1.34657	0.13277	1.81051	0.38764	1.03158	0.50684
15	6	1.28888	1.23806	0.09740	1.59559	0.23990	0.98130	0.40290
20	4	2.32083	1.57816	0.32663	2.37018	0.81684	1.98445	0.59461
20	4.5	1.98981	1.40687	0.22877	1.96733	0.70116	1.62088	0.58914
20	5	1.74833	1.27501	0.15922	1.80017	0.49777	1.44937	0.34961
20	5.5	1.56618	1.16892	0.11358	1.58437	0.19391	1.16005	0.31807
20	6	1.42500	1.09720	0.08816	1.34147	0.11997	1.03021	0.25591

Table 2: The estimate and the corresponding MSE for maximum likelihood estimator and Bayes estimator for varentropy measure of Pareto distribution when $\sigma = 1.5$

n	λ	VH	MLE		SEL		LL	
			$\hat{VH}_{MLE}$	MSE	$\hat{VH}_{SEL}$	MSE	$\hat{VH}_{LL}$	MSE
10	4	1.70413	2.21989	0.46697	2.66324	1.54455	1.73646	0.50602
10	4.5	1.50429	1.91009	0.30800	2.60121	1.63102	1.51832	0.45477
10	5	1.35663	1.68491	0.21082	2.38211	1.40016	1.39104	0.31763
10	5.5	1.24391	1.51528	0.14851	2.34710	1.46904	1.42697	0.20658
10	6	1.15552	1.37785	0.10797	2.18581	1.16687	1.14634	0.16418
15	4	2.01110	1.91043	0.41243	2.54690	1.06764	1.97112	0.48333
15	4.5	1.74567	1.66354	0.26574	2.35480	1.13801	1.28683	0.32683
15	5	1.55110	1.47888	0.18372	1.91580	0.37910	1.28683	0.26827
15	5.5	1.40366	1.34342	0.13200	1.54734	0.30583	1.19281	0.18493
15	6	1.28888	1.23393	0.09679	1.79671	0.35970	1.23470	0.15229
20	4	2.32083	1.58400	0.33754	2.41008	0.98868	2.27783	0.43932
20	4.5	1.98981	1.40160	0.22300	2.23748	0.88120	1.76027	0.32421
20	5	1.74833	1.27191	0.15826	1.79843	0.27150	1.61082	0.20016
20	5.5	1.56618	1.17073	0.11425	1.72403	0.18625	1.15691	0.15428
20	6	1.42500	1.09136	0.08608	1.49738	0.15636	1.38267	0.14782

6. Illustration Using Real Data

In this section, we consider the real data set given in Volovskiy and Kamps (2020) for the maximum product of spacings predictions of the next record for the Pareto distribution. The data set contains hourly measurements (in cm) of water level for the period from January

Table 3: The estimate and the corresponding MSE for maximum likelihood estimator and Bayes estimator for varentropy measure of Pareto distribution when $\sigma = 2$

n	λ	VH	MLE		SEL		LL	
			$\hat{VH}_{MLE}$	MSE	$\hat{VH}_{SEL}$	MSE	$\hat{VH}_{LL}$	MSE
10	4	1.70413	2.22723	0.46664	2.97497	2.17147	2.00548	0.67238
10	4.5	1.50429	1.90900	0.30861	2.69485	1.90355	1.64239	0.49655
10	5	1.35663	1.68706	0.21082	2.28303	1.08104	1.39104	0.31763
10	5.5	1.24391	1.51442	0.14902	2.34904	1.52808	1.43228	0.23237
10	6	1.15552	1.38051	0.10805	2.07611	1.08567	1.18425	0.19581
15	4	2.01110	1.91200	0.40607	2.76514	1.23153	2.29336	0.62905
15	4.5	1.74567	1.65782	0.26030	2.28238	1.00877	1.83485	0.38615
15	5	1.55110	1.48156	0.18476	1.79359	0.34462	1.56350	0.27825
15	5.5	1.40366	1.34398	0.13058	1.93490	0.80539	1.50808	0.15050
15	6	1.28888	1.23621	0.09586	1.65005	0.30465	1.40755	0.18986
20	4	2.32083	1.58370	0.34194	2.47386	0.77371	2.77457	0.59009
20	4.5	1.98981	1.41232	0.22756	2.22127	0.52853	2.05994	0.32281
20	5	1.74833	1.27348	0.15839	1.93213	0.43787	1.75522	0.25379
20	5.5	1.56618	1.16955	0.11338	1.68607	0.19747	1.17456	0.14990
20	6	1.42500	1.09115	0.08651	1.62995	0.27030	1.21483	0.15758

1918 to February 2019 collected at the measurement site Cuxhaven-Steubenhoft located at the river Elbe. The sequence of upper record values extracted from the dataset of weekly maximum water levels exceeding are 713, 781, 880, 885, 901, 914, 915, 993 and 1010.
We have obtained the estimators for varentropy measure using different methods and are given below.

MLE		0.6108267
Bayes estimator	SEL	0.6761488
	LL	0.6757579

7. Conclusion

In this paper, we considered the varentropy measure for the nth upper and lower record values. The expressions for the varentropy measure for the nth upper and lower record values were obtained. We also derived the expressions for residual varentropy measure and past varentropy measure for the nth upper and lower record values. Then, as an application of varentropy measure in estimation, we obtained the MLEs and the Bayes estimates for the varentropy measure for Pareto model based on classical record values. MCMC method was applied to obtain the Bayes estimates. Based on the simulation study, we found that the MSEs of all estimators decrease when n increases. Also, the MLEs for the varentropy measure performed better than Bayes estimates with respect to MSEs. Among the Bayes estimators, estimators under LL function performed better that of SEL function.

ACKNOWLEDGMENTS

The authors would like to thank the editor and the reviewers for their constructive comments, which helped to improve the quality of the paper.

REFERENCES

Abbasnejad, M. and Arghami, N. R. (2011). Rényi entropy properties of records. *Journal of Statistical Planning and Inference*, 141(7):2312–2320.

Alizadeh, N. H. and Shafaei, M. N. (2023). Varentropy estimators with applications in testing uniformity. *Journal of Statistical Computation and Simulation*, 93(15):2582–2599.

Alizadeh, N. H. and Shafaei, M. N. (2024). Monte carlo comparison of normality tests based on varentropy estimators. *Journal of Statistical Computation and Simulation*, pages 1–17.

Arıkan, E. (2016). Varentropy decreases under the polar transform. *IEEE Transactions on Information Theory*, 62(6):3390–3400.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1992). A first course in order statistics, vol. 54. *SIAM, Philadelphia*.

Baratpour, S., Ahmadi, J., and Arghami, N. R. (2007). Entropy properties of record statistics. *Statistical Papers*, 48:197–213.

Bobkov, S. and Madiman, M. (2011). Concentration of the information in data with log-concave distributions. *The Annals of Probability*, 39:1528–1543.

Buono, F., Longobardi, M., and Pellerey, F. (2022). Varentropy of past lifetimes. *Mathematical Methods of Statistics*, 31(2):57–73.

Chacko, M. and Asha, P. (2018). Estimation of entropy for generalized exponential distribution based on record values. *Journal of the Indian Society for Probability and Statistics*, 19:79–96.

Chacko, M. and Asha, P. (2021). Estimation of entropy for Weibull distribution based on record values. *Journal of Statistical Theory and Applications*, 20(2):279–288.

Chacko, M. and Grace, A. (2023). Information generating function of k-record values and its applications. *Mathematical Methods of Statistics*, 32(3):176–196.

Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied probability*, 39(2):434–440.

Di Crescenzo, A. and Paolillo, L. (2021). Analysis and applications of the residual varentropy of random lifetimes. *Probability in the Engineering and Informational Sciences*, 35(3):680–698.

Di Crescenzo, A., Paolillo, L., and Suárez-Llorens, A. (2024). Stochastic comparisons, differential entropy and varentropy for distributions induced by probability density functions. *Metrika*, pages 1–17.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 48–56.

Fradelizi, M., Madiman, M., and Wang, L. (2016). Optimal concentration of information content for log-concave densities. In *High Dimensional Probability VII: The Cargèse Volume*, pages 45–60. Springer.

Kontoyiannis, I. and Verdú, S. (2013). Optimal lossless compression: Source varentropy and dispersion. In *2013 IEEE International Symposium on Information Theory*, pages 1739–1743. IEEE.

Maadani, S., Mohtashami Borzadaran, G. R., and Rezaei Roknabadi, A. (2020). A new generalized varentropy and its properties. *Ural Mathematical Journal*, 6(1 (10)):114–129.

Maadani, S., Mohtashami Borzadaran, G. R., and Rezaei Roknabadi, A. (2022). Varentropy of order statistics and some stochastic comparisons. *Communications in Statistics-Theory and Methods*, 51(18):6447–6460.

Muliere, P., Parmigiani, G., and Polson, N. G. (1993). A note on the residual entropy function. *Probability in the Engineering and Informational Sciences*, 7(3):413–420.

Muraleedharan, L. and Chacko, M. (2022a). Inference and prediction of order statistics based on k-record values from a weibull distribution. *IAPQR Transactions*, 47(1-2):86.

Muraleedharan, L. and Chacko, M. (2022b). Interval prediction of order statistics and record values using concomitants of order statistics and record values for morgenstern family of distributions. *Journal of Statistical Research*, 56(1):55–73.

Raqab, M. Z., Bayoud, H. A., and Qiu, G. (2022). Varentropy of inactivity time of a random variable and its related applications. *IMA Journal of Mathematical Control and Information*, 39(1):132–154.

Saha, S. and Kayal, S. (2024). Weighted past and paired dynamic varentropy measures, their properties and usefulness. *arXiv preprint arXiv:2405.06428*.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.

Sharma, A. and Kundu, C. (2023). Varentropy of doubly truncated random variable. *Probability in the Engineering and Informational Sciences*, 37(3):852–871.

Volovskiy, G. and Kamps, U. (2020). Maximum product of spacings prediction of future record values. *Metrika*, 83(7):853–868.