Akash Sharma; Chanchal Kundu

On stochastic properties of past varentropy with applications

Applications of Mathematics, Vol. 69 (2024), No. 3, 373–394

Persistent URL: `http://dml.cz/dmlcz/152355`

69 (2024) APPLICATIONS OF MATHEMATICS No. 3, 373–394

ON STOCHASTIC PROPERTIES OF PAST VARENTROPY WITH APPLICATIONS

Akash Sharma, Chanchal Kundu, Jais

Received August 10, 2023. Published online March 11, 2024.

Abstract. To have accuracy in the extracted information is the goal of the reliability theory investigation. In information theory, varentropy has recently been proposed to describe and measure the degree of information dispersion around entropy. Theoretical investigation on varentropy of past life has been initiated, however details on its stochastic properties are yet to be discovered. In this paper, we propose a novel stochastic order and introduce new classes of life distributions based on past varentropy. Further, we illustrate some of its applications in reliability modeling and in the diversity measure of Boltzmann distribution.

Keywords: ageing classes; past varentropy order; stochastic orders; varentropy

MSC 2020: 94A17, 62B10, 60E15

1. Introduction

In order to encode the notion of uncertainty linked with a random variable, the idea of entropy provides a mathematical way. It serves as a fundamental measure in information theory and statistical mechanics. Since uncertainty is involved in many reasonable situations such as lifetime data analysis and econometrics, studying the predictability of occurrences is important. To provide a quantitative measure to uncertainty for time-to-event random variables, the standard measure of uncertainty, i.e., differential entropy, is extensively used.

Let the lifetime of a system be denoted by a non-negative random variable X with an absolutely continuous distribution function (df) F and probability density function (pdf) f. Then

$$H_s(X) = E[-\log f(X)] = -\int_0^\infty f(x)\log f(x)\,\mathrm{d}x, \tag{1.1}$$

DOI: 10.21136/AM.2024.0163-23

is the defined Shannon [21], [22] entropy, where log symbolizes base 'e' logarithm. It quantifies the expected randomness in X concerning the likelihood of an outcome and gauges the information associated with probability distribution. The higher value of entropy is symbolic of greater uncertainty or higher volume of information content. To simplify, the outcomes induce significant information upon happenings. Moreover, it picturizes the distribution of information density, on an average in the overall range of the variable (cf. Cover and Thomas [6]).

Shannon's measure is found to be functional in characterizing the quality of informational sources, communication channels and data transmission. It is fruitful in ascertaining the channel capacity for a reliable data transmission in varied pragmatic method of information coding techniques such as Huffman coding. In fact, such techniques leverage the probabilistic complexion of data achieving efficient compression without any information loss. In such cases, average length of the encoded message is minimized using random arrangements of symbols following the specific technique. Further, it also intensifies the cryptographic system generating safe keys since high-entropy keys are less predictable and thus, much secure. These cryptographic keys are randomly arranged codes providing foremost unpredictability challenging adversaries to guess the key. Moreover, increased precision in speech recognition and natural language processing model are developed using the random arrangements based on entropy, assessing the uncertainty of words or sequence of words in language modeling. Therefore, the measure is crucial for comprehending, analyzing, and communicating the information pertaining to several aspects including pattern recognition, language processing and data analysis paving the path to make more informed decisions.

The random variable constrained on the fact that the failure happens before a given inspection time t is the subject of many studies in reliability theory. This is known as past lifetime. In many rational contexts such as in forensic sciences, uncertainty can have reference to the past where right truncated lifetime distributions are the focus of study. Suppose that a system that is only monitored during specific predetermined inspection intervals is discovered to be down at time t, then the uncertainty of system's life depends on the past, namely, which instant in time $(0,t)$ it failed. As it is inappropriate to use (1.1) when the system is ascertained to be failing at time t, the interpretation of past entropy was put forward by Di Crescenzo and Longobardi [7] as

$$H_s(X_{(t)}) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} \,\mathrm{d}x, \tag{1.2}$$

which quantifies the uncertainty in the inactivity time $X_{(t)} = (t - X \mid X \leqslant t)$. In the case of a system that is determined to fail at time $t > 0$, past entropy signifies the uncertainty concerning the system's failure instant in $(0,t)$.

One may find voluminous literature on the information measures related to its dynamic measures and their generalizations, yet the analysis about the variation of the information measures is limited. Studying variance of information measures, called varentropy, outlines the expected volatility in the information contained in the random variable X. It pictures the dispersion of information content around the entropy and indicates the amount of variation to expect in the entropy, which is useful in extracting significant information from it. The average information content of the random variable is represented by the differential entropy, whereas varentropy indicates the deviation we could anticipate in the information content from entropy. Mathematically, it is given by (cf. Song [24])

$$\begin{aligned}(1.3)\qquad V_s(X) &= \operatorname{Var}(-\log f(X)) = E[\log^2 f(X)] - [H_s(X)]^2 \\ &= \int_0^\infty f(x)\{\log f(x)\}^2\,\mathrm{d}x - \left[\int_0^\infty f(x)\log f(x)\,\mathrm{d}x\right]^2.\end{aligned}$$

It measures the extent of scatterness we may expect in the information content of X. Some of the applications of this measure can be found in calculating the dispersion of sources, in data compression, in estimating the functioning of schemes of optimal coding and in computing channel capacity. For instance, see Liu [13], Bobkov and Madiman [3] and Arıkan [1]. Some recent study on the variance of information measure and its materiality can be seen in Fradelizi et al. [11] and Di Crescenzo et al. [9].

Because of the deficient investigation on varentropy, Di Crescenzo and Paolillo [8] came up with the study on varentropy of residual lifetimes to put forward some of its properties. Recently, Buono et al. [4] defined the past varentropy for a system observed to have some inactivity period at time 't' as

$$\begin{aligned}V_s(X_{(t)}) &= \operatorname{Var}(-\log f_{X_{(t)}}(X_{(t)})) = E[\log^2 f_{X_{(t)}}(X_{(t)})] - [H_s(X_{(t)})]^2 \\ &= \int_0^t \frac{f(x)}{F(t)}\Big[\log\frac{f(x)}{F(t)}\Big]^2\,\mathrm{d}x - \left[\int_0^t \frac{f(x)}{F(t)}\log\frac{f(x)}{F(t)}\,\mathrm{d}x\right]^2,\end{aligned}$$

with $f_{X_{(t)}}$ as the pdf of $X_{(t)}$. Further results on varentropy of past lifetimes can be found in Raqab et al. [18]. Sharma and Kundu [23] analyzed the varentropy for a random variable truncated on both ends and found its effectiveness in selecting acceptable system.

Uncertainty measures are fruitful in predicting the failure instant of a system. It is fair to consider a system with less uncertainty to be more reliable than a system with larger uncertainty. The order reliant on the measure of uncertainty was first initiated in Ebrahimi and Kirmani [10], where they stated X to be less uncertain than Y if $E[-\log f(X)] \leqslant E[-\log g(Y)]$. In many life testing investigations, often we are provided with the time a system is found to be failed. The measure given by (1.2) is functional in identifying systems with lesser inactivity time. Certain systems were distinguished by Di Crescenzo and Longobardi [7] as having equal entropies but different past entropies. Intuitively, it is practical to consider a system with less uncertainty in $(0, t)$ to have lesser chance of early failure than that of a system with more uncertainty in that interval. Motivated from these aspects, Nanda and Paul [15] advanced with an order using past entropy. They considered the comparison of two systems having lifetimes in $(0, t)$, where t is the observed time of failure, with regards to past entropy. Moreover, they defined a novel category of lifetime distribution characterised by increasing uncertainty in past life (IUPL) in order to study its relation with the existing ageing classes. The usefulness of past entropy order is notable in choosing a better system.

The efficacy of the past entropy order cannot be ignored but at times it may fail or may not be of worth. Thus, on comparing lifetime distributions on account of uncertainty measures, it is judicious to think "if not entropy then why not varentropy?" Varentropy is instrumental in analyzing the spread of information making it a constructive tool to compare life distributions where the dispersion measure is of gravity. Moreover, for the life distributions where entropy order is insufficient, comparison in terms of varentropy may be influential. The varentropy order is found to be effective compared to the kurtosis measure traditionally used in contrasting distribution weights in the tail region. Recently, Maadani et al. [14] came up with varentropy order and explored how the existing stochastic orders are related to it. But the stochastic orders on dynamic measures of varentropy and its related ageing classes are yet to be examined. Motivated from the above and considering the importance of past lifetime, the present paper aims to propose past varentropy order and its corresponding ageing classes.

The structure for the remaining portion is as follows: In the first part of Section 2, we put forth a novel stochastic order constructed utilising past varentropy and compare it to the pre-existing orders. Some results using the defined order are produced. The non- parametric classes based on past varentropy are introduced and their relation with the IUPL class is investigated in the latter part of Section 2. In Section 3, we examine the applications of the PVE order in reliability modeling to identify the best fitted model for a given data set and in comparing the variability in the diversity measure of Boltzmann distribution. Finally, a brief summary of the present analysis has been added in Section 4.

The notion of "increasing" and "decreasing" in this article are not used in strict sense.

2. Stochastic properties of varentropy for past lifetimes

Nanda and Paul [15], [16] examined stochastic orders and an additional class of life distribution with regard to past entropy. Drawing inspiration from the work of Di Crescenzo and Longobardi [7], they defined the following order, which assisted in the construction of beneficial results.

Definition 2.1. Suppose the random variables X and Y denote the survival times of two systems. Then X is considered greater than Y in past entropy order (denoted by $X \geqslant_{\mathrm{PE}} Y$) if

$$H_s(X_{(t)}) \leqslant H_s(Y_{(t)}) \quad \forall\, t \geqslant 0. \tag{2.1}$$

The above definition signifies that the predicted uncertainty concerning the indication of the failure instant in the distribution of X, provided $X \leqslant t$ is less in comparison to that of Y when $Y \leqslant t$. It is found to be useful in choosing the most acceptable system.

2.1. Past varentropy order.

We know that the varentropy is much useful under certain cases examining the dispersion of information in risk theory. This strikes the motivation to define the order on varentropy in past life.

Definition 2.2. With regard to Definition 2.1, the random variable X is considered greater than random variable Y in past varentropy order (denoted by $X \geqslant_{\mathrm{PVE}} Y$) if

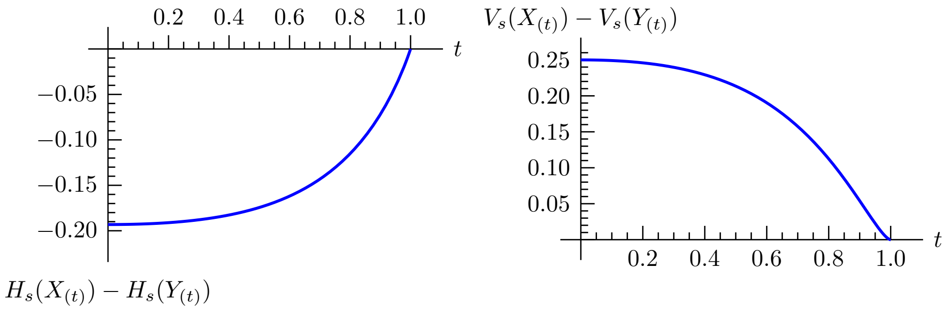
$$V_s(X_{(t)}) \leqslant V_s(Y_{(t)}) \quad \forall\, t \geqslant 0. \tag{2.2}$$

The physical understanding of $X \geqslant_{\mathrm{PVE}} Y$ extends more of the definition. That is, if $V_s(X_{(t)})$ and $V_s(Y_{(t)})$ are the varentropy measures of two systems X and Y in terms of inactivity life and $X \geqslant_{\mathrm{PVE}} Y$ holds, then after both components are observed out of order on time 't', the scatterness of information about the predicted uncertainty concerning the indication of the failure time carried by the distribution function of X, provided $X \leqslant t$ is less in comparison to scatterness of information around the predicted uncertainty content in the distribution function of Y when $Y \leqslant t$. Thus, X is more precise in calculating uncertainty for inactivity time than Y.

Remark 2.1. The reflexive, antisymmetric and transitive nature of the ordering described above may be proven. As a result, it is a partial order.

It is to be noted that $\geqslant_{\mathrm{PVE}}$ is based on information deviation in inactivity time whereas $\geqslant_{\mathrm{PE}}$ order on information expectation. A natural thought is to look for their interrelation. We present two examples to support the distinctiveness of the $\geqslant_{\mathrm{PVE}}$ order. The example below depicts that $X \geqslant_{\mathrm{PE}} Y \nRightarrow X \geqslant_{\mathrm{PVE}} Y$.

E x a m p l e 2.1. Consider the pdfs $f(x) = 2x$ and $g(x) = 2(1-x)$, $0 < x < 1$ of the random variables X and Y, respectively. From Figure 1, it is clear that $X \geqslant_{\mathrm{PE}} Y$ but $X \ngeqslant_{\mathrm{PVE}} Y$.

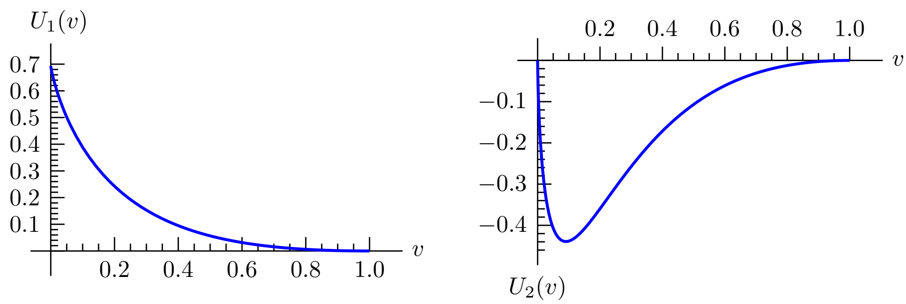


(a) Plot of $H_s(X_{(t)}) - H_s(Y_{(t)})$. (b) Plot of $V_s(X_{(t)}) - V_s(Y_{(t)})$.

Figure 1. Difference of past entropy and varentropy (Example 2.1).

The case below allows us to conclude $X \geqslant_{\mathrm{PVE}} Y \nRightarrow X \geqslant_{\mathrm{PE}} Y$.

E x a m p l e 2.2. Suppose the random variables X and Y have exponential distribution with means 1 and $\frac{1}{2}$, respectively. Let $U_1(t) = H_s(X_{(t)}) - H_s(Y_{(t)})$ and $U_2(t) = V_s(X_{(t)}) - V_s(Y_{(t)})$. Then $U_1(t) \geqslant 0$ but $U_2(t) \leqslant 0$ as can be seen from Figure 2. Consequently, $X \geqslant_{\mathrm{PVE}} Y$ holds but $X \ngeqslant_{\mathrm{PE}} Y$. Take a note of the transformation $t = -\log v$, where $v \in (0,1)$ is used while plotting the curves which results in $U_1(t) = U_1(v)$ and $U_2(t) = U_2(v)$, say.



(a) Plot of $U_1(v)$. (b) Plot of $U_2(v)$.

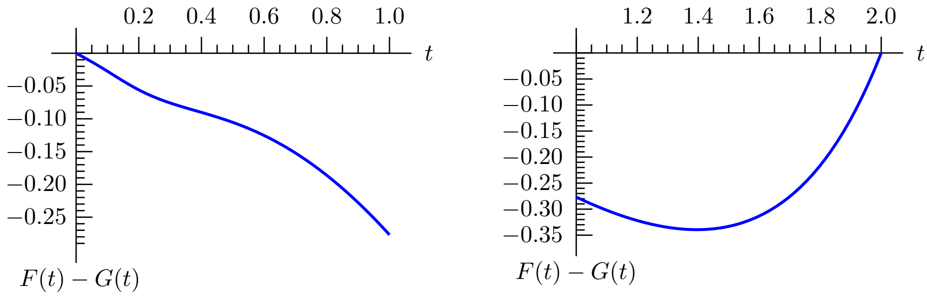
Figure 2. The past entropy and varentropy difference (Example 2.2).

Nanda and Paul [16] investigated that no relation exists between PE order and usual stochastic order. We substantiate this claim for past varentropy order. The following counterexample demonstrates that $X \geqslant_{ST} Y$ but $X \ngeqslant_{\mathrm{PVE}} Y$. In-depth study of usual dispersion (DISP), failure rate (FR), likelihood ratio (LR) and stochastic (ST) orders can be found in Shaked and Shanthikumar [19].

C o u n t e r e x a m p l e 2.1. Consider the distribution function of two systems having lifetimes given by X and Y as

$$F(t) = \begin{cases} \mathrm{e}^{-1/2-1/t}, & 0 \leqslant t \leqslant 1, \\ \mathrm{e}^{-2+t^2/2}, & 1 \leqslant t \leqslant 2, \end{cases} \quad \text{and} \quad G(t) = \begin{cases} \dfrac{t^2+t}{4}, & 0 \leqslant t \leqslant 1, \\ \dfrac{t}{2}, & 1 \leqslant t \leqslant 2, \end{cases}$$

respectively. Figure 3 shows that $X \geqslant_{ST} Y$. Let $U_3(t) = V_s(X_{(t)}) - V_s(Y_{(t)})$, then $U_3(1.6) = -0.035 < 0$ and $U_3(1.8) = 0.091 > 0$. Consequently, $U_3(t)$ passes the x-axis for $t \in (1,2)$. Therefore, X and Y fails to be in PVE order.



(a) $F(t) - G(t)$ for $t \in (0,1)$.

(b) $F(t) - G(t)$ for $t \in (1,2)$.

Figure 3. Graph of $F(t) - G(t)$, where $t \in (0,2)$ (Counterexample 2.1).

An example which indicates that $X \geqslant_{\mathrm{PVE}} Y$ but $X \ngeqslant_{ST} Y$ can be seen below.

C o u n t e r e x a m p l e 2.2. Let the random variables X and Y be given by dfs

$$F(t) = 1 - \mathrm{e}^{-t} \quad \text{and} \quad G(t) = 1 - \frac{1}{(1+t)^2} \quad \text{for } t > 0 \text{ correspondingly.}$$

If $U_4(t) = F(t) - G(t)$, then $U_4(2) = -0.0242 < 0$ and $U_4(3) = 0.0127 > 0$. The horizontal axis is thus clearly crossed by $U_4(t)$ at some $t \in (2,3)$. Thus, X and Y fails to be in ST order. Let $U_5(t) = V_s(Y_{(t)}) - V_s(X_{(t)})$. It is clear from Figure 4 that $U_5(v) \geqslant 0$, where $U_5(t) = U_5(v)$, say when $t = -\log v$ is substituted. Thus, $X \geqslant_{\mathrm{PVE}} Y$.

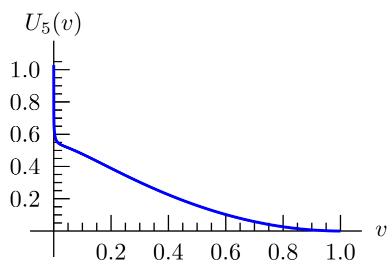


Figure 4. Plot of $U_5(v)$ (Counterexample 2.2).

It is noteworthy to highlight that varentropy expression in compact form is either difficult to achieve or demands a great deal of effort. Nevertheless, one can assess with ease the required varentropy in respect of known varentropy of initial distribution utilizing specific transformations. In fact, examining varentropy underneath transformations proves to be imperative in investigating characteristics that are inherited. Amongst the inherited characteristics, one is varentropy order. Put simply, if one has details regarding pre-existing past varentropy order and specific transformations, the goal of comparing volatility in past lifetime is achieved saving humongous computational work. We recollect that for

$$Z = \alpha_1 X + \beta_1, \quad \alpha_1 > 0,\ \beta_1 \geqslant 0;$$

the past varentropy of X and Z are coupled by the relation (cf. Buono et al. [4])

$$V_s(Z_{(t)}) = V_s(X_{((t-\beta_1)/\alpha_1)}) \quad \forall\, t \geqslant \beta_1. \tag{2.3}$$

The PVE order is closed under certain conditions of increasing affine transformation, as shown by the following theorem.

Theorem 2.1. *Let $X_1 = aX + b$ and $X_2 = cY + d, a, c > 0$ and $b, d \geqslant 0$. Further:*

(i) *Suppose* (a) $X \geqslant_{\rm PVE} Y$, (b) $a = c$ *and* $b = d$. *Then* $X_1 \geqslant_{\rm PVE} X_2$.

(ii) *Let* (a) $X \geqslant_{\rm PVE} Y$, (b) $a \geqslant c$ *and* $b \geqslant d$, (c) $V(X_{(t)})$ *or* $V(Y_{(t)})$ *be increasing in* $t > b$. *Then* $X_1 \geqslant_{\rm PVE} X_2$.

Proof. We prove (ii) as (i) is straightforward. Assume $V(X_{(t)})$ to increase with t. Since $(t-b)/a \leqslant (t-d)/c$, we have

$$V_s(X_{((t-b)/a)}) \leqslant V_s(X_{((t-d)/c)}) \leqslant V_s(Y_{((t-d)/c)}),$$

where the last inequality comes from $X \geqslant_{\rm PVE} Y$. The proof when $V(Y_{(t)})$ is increasing is similarly obtained. The result therefore follows. □

The following counterexample demonstrates that the monotonic condition of the past varentropy cannot be dropped in the preceding theorem.

Counterexample 2.3. Let X and Y have dfs

$$F(t) = \frac{3t^2 - t^3}{4} \quad \text{for } 0 \leqslant t \leqslant 2 \quad \text{and} \quad G(t) = \begin{cases} \dfrac{t^2}{2} & \text{if } 0 \leqslant t \leqslant 1, \\ \dfrac{t^2+2}{6} & \text{if } 1 \leqslant t \leqslant 2. \end{cases}$$

respectively. It can be verified from Figure 5 and Figure 6 that $V_s(X_{(t)})$ and $V_s(Y_{(t)})$ are non-increasing, correspondingly. Further, $X \geqslant_{\rm PVE} Y$ is inferred from Figure 7. Now, if $X_1 = X + 0.01, X_2 = Y$ and $U_6(t) = V_s(X_{2(t)}) - V_s(X_{1(t)})$, then $U_6(1.99) = -0.0019 < 0$. Thus, $X_1 \ngeqslant_{\rm PVE} X_2$.

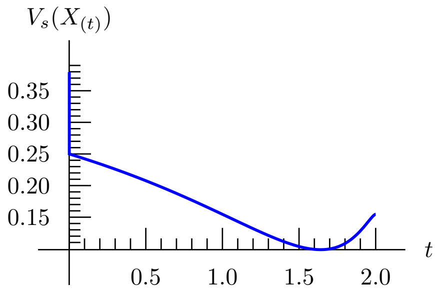
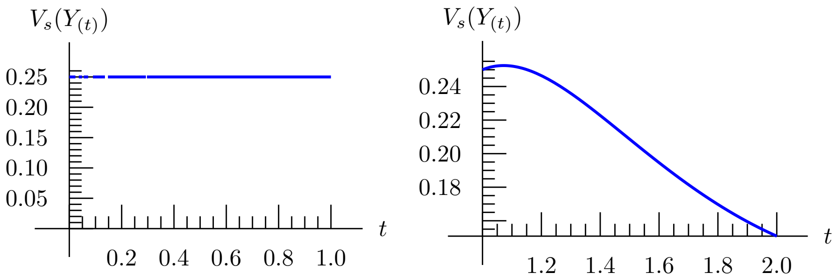
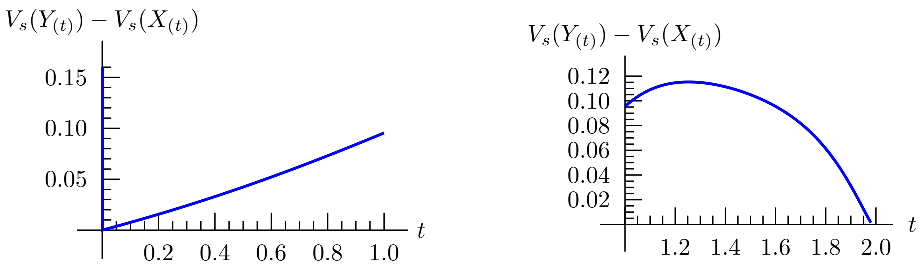


Figure 5. Plot of $V_s(X_{(t)})$ (Counterexample 2.3).



(a) $V_s(Y_{(t)})$ for $t \in (0,1)$. (b) $V_s(Y_{(t)})$ for $t \in (1,2)$.

Figure 6. Plot of $V_s(Y_{(t)})$ (Counterexample 2.3).



(a) Plot of $V(Y_{(t)}) - V(X_{(t)})$ for $t \in (0,1)$. (b) Plot of $V(Y_{(t)}) - V(X_{(t)})$ for $t \in (1,2)$.

Figure 7. Plot of varentropy difference (Counterexample 2.3).

In the next theorem we show the closure of PVE order under increasing convex transformation. Before moving forward, we note from Proposition 2.3 of Sharma and Kundu [23] that

(2.4)
$$\begin{aligned}V_s(Y_{(t)}) = V_s(X_{(\varphi^{-1}(t))}) - 2E\Big[\log \frac{f_X(X)}{F_X(\varphi^{-1}(t))} \log \varphi'(X) \mid X < \varphi^{-1}(t)\Big] \\ + \mathrm{Var}[\log \varphi'(X) \mid X < \varphi^{-1}(t)] - 2H_s(X_{(\varphi^{-1}(t))}) E[\log \varphi'(X) \mid X < \varphi^{-1}(t)],\end{aligned}$$

when $Y = \varphi(X)$, for a strictly increasing function φ.

Theorem 2.2. *Suppose that $X \geqslant_{\rm PE} Y$ and $X \leqslant_{\rm PVE} Y$ hold for the non-negative random variables X and Y. Assume that a convex function φ is strictly increasing with $\varphi(0)=0$. Then $\varphi(X) \leqslant_{\rm PVE} \varphi(Y)$, provided the derivative of φ is continuous with $0<\varphi'(x)\leqslant 1$.*

Proof. Let φ' represent the derivative of φ. From (2.4) we have

(2.5)

$$\begin{aligned} &V_s(\varphi(X)_{(t)}) - V_s(\varphi(Y)_{(t)}) \\ &= V_s(X_{(\varphi^{-1}(t))}) - V_s(Y_{(\varphi^{-1}(t))}) - 2E\Big[\log \frac{f(X)}{F(\varphi^{-1}(t))} \log \varphi'(X) \mid X<\varphi^{-1}(t)\Big] \\ &\quad + 2E\Big[\log \frac{g(Y)}{G(\varphi^{-1}(t))} \log \varphi'(Y) \mid Y<\varphi^{-1}(t)\Big] + \mathrm{Var}[\log\varphi'(X) \mid X<\varphi^{-1}(t)] \\ &\quad - \mathrm{Var}[\log\varphi'(Y) \mid Y<\varphi^{-1}(t)] - 2H_s(X_{(\varphi^{-1}(t))})E[\log\varphi'(X) \mid X<\varphi^{-1}(t)] \\ &\quad + 2H_s(Y_{(\varphi^{-1}(t))})E[\log\varphi'(Y) \mid Y<\varphi^{-1}(t)]. \end{aligned}$$

Proceeding with term by term comparison, observe that the first term comparison follows directly from the assumption $V_s(X_{(t)})-V_s(Y_{(t)}) \geqslant 0$. Next, we take account of

$$\begin{aligned} &2E\Big[\log \frac{f(X)}{F(\varphi^{-1}(t))} \log \varphi'(X) \mid X<\varphi^{-1}(t)\Big] \\ &\qquad\qquad - 2E\Big[\log \frac{g(Y)}{G(\varphi^{-1}(t))} \log \varphi'(Y) \mid Y<\varphi^{-1}(t)\Big] \\ &\qquad = 2\int_0^{\varphi^{-1}(t)} (-\log\varphi'(x))\Big\{-\frac{f(x)}{F(\varphi^{-1}(t))}\log\frac{f(x)}{F(\varphi^{-1}(t))} \\ &\qquad\qquad + \frac{g(x)}{G(\varphi^{-1}(t))}\log\frac{g(x)}{G(\varphi^{-1}(t))}\Big\}\,\mathrm{d}x \\ &\qquad \leqslant 2(-\log\varphi'(0))\{H_s(X_{(\varphi^{-1}(t))}) - H_s(Y_{(\varphi^{-1}(t))})\} \leqslant 0. \end{aligned}$$

The assumption used to get the last inequality is $X \geqslant_{\rm PE} Y$. Further, observe

$$\begin{aligned} &E[(\log\varphi'(Y))^2 \mid Y<\varphi^{-1}(t)] - E[(\log\varphi'(X))^2 \mid X<\varphi^{-1}(t)] \\ &\quad = \int_0^{\varphi^{-1}(t)} (\log\varphi'(x))^2 \Big\{\frac{g(x)}{G(\varphi^{-1}(t))} - \frac{f(x)}{F(\varphi^{-1}(t))}\Big\}\,\mathrm{d}x \\ &\quad \leqslant (\log\varphi'(0))^2 \int_0^{\varphi^{-1}(t)} \Big\{\frac{g(x)}{G(\varphi^{-1}(t))} - \frac{f(x)}{F(\varphi^{-1}(t))}\Big\}\,\mathrm{d}x = 0. \end{aligned}$$

The fact that $\phi'(x)$ increases in x leads to the inequality mentioned above. On similar course of action, we obtain

$$-(E[\log\varphi'(X) \mid X<\varphi^{-1}(t)])^2 + (E[\log\varphi'(Y) \mid Y<\varphi^{-1}(t)])^2 \geqslant 0.$$

Bringing together, we have

$$\mathrm{Var}[\log \varphi'(X) \mid X < \varphi^{-1}(t)] - \mathrm{Var}[\log \varphi'(Y) \mid Y < \varphi^{-1}(t)] \geqslant 0.$$

Moreover, it is easy to establish

$$\begin{aligned}2H_s(Y_{(\varphi^{-1}(t))})E[\log \varphi'(Y) \mid Y < \varphi^{-1}(t)]\\ - 2H_s(X_{(\varphi^{-1}(t))})E[\log \varphi'(X) \mid X < \varphi^{-1}(t)] \geqslant 0.\end{aligned}$$

Thus, from (2.5), $V_s(\varphi(X)_{(t)}) \geqslant V_s(\varphi(Y)_{(t)})$ for all $t \geqslant 0$, i.e., $\varphi(X) \leqslant_{\mathrm{PVE}} \varphi(Y)$. $\square$

We provide an example to verify the above theorem.

Example 2.3. Suppose that X and Y are the random variables having dfs

$$F(t) = 1 - \mathrm{e}^{-2t} \quad \text{and} \quad G(t) = 1 - \mathrm{e}^{-t} \quad \text{for } t \geqslant 0, \text{ respectively.}$$

From Figure 2, it is evident that $X \geqslant_{\mathrm{PE}} Y$ and $X \leqslant_{\mathrm{PVE}} Y$ hold. Choose an increasing convex function $\varphi(x) = 0.5x$. Clearly, φ satisfies the condition stated in Theorem 2.2. Denote $U_6(t) = V_s(\varphi(X)_{(t)}) - V_s(\varphi(Y)_{(t)})$. It can be verified from Figure 8 that $\varphi(X) \leqslant_{\mathrm{PVE}} \varphi(Y)$. Note that $U_6(t) = U_6(v)$, say for the substitution $t = -\log v$.

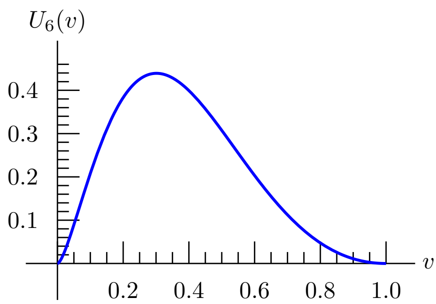


Figure 8. Plot of $U_6(v)$ for $v \in (0,1)$ (Example 2.3).

Here, we provide two counterexamples to stress that the convexity condition and the derivative condition in Theorem 2.2 are merely sufficient requirements.

Counterexample 2.4. For the random variables X and Y in Example 2.3, if we choose the increasing function $\varphi(x) = 1 - \mathrm{e}^{-x}$, then it satisfies all the conditions on φ function stated in Theorem 2.2 except the convexity condition. Denote $U_7(t) = V_s(\varphi(X)_{(t)}) - V_s(\varphi(Y)_{(t)})$. After substituting $t = -\log v$, where $v \in (0,1)$, write $U_7(t) = U_7(v)$. It is seen from Figure 9 that $U_7(v) \geqslant 0$. Hence, $\phi(X) \leqslant_{\mathrm{PVE}} \phi(Y)$.

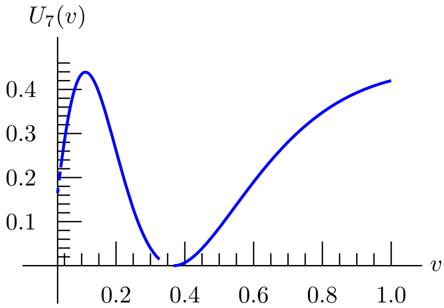


Figure 9. Plot of $U_7(v)$ for $v \in (0,1)$ (Counterexample 2.4).

Counterexample 2.5. For the random variables X and Y in Example 2.3, take an increasing convex function $\varphi(x) = \mathrm{e}^x - 1$. Then it is evident that $\varphi(0) = 0$ but $\varphi'(x) > 1$, for $x > 0$. Now, if $U_8(t) = V_s(\varphi(X)_{(t)}) - V_s(\varphi(Y)_{(t)})$, then it is clear from Figure 10 that $U_8(v) \geqslant 0$. Note that $U_8(t) = U_8(v)$, say for the substitution $t = -\log v$, $v \in (0,1)$. Thus, $\varphi(X) \leqslant_{\mathrm{PVE}} \varphi(Y)$.

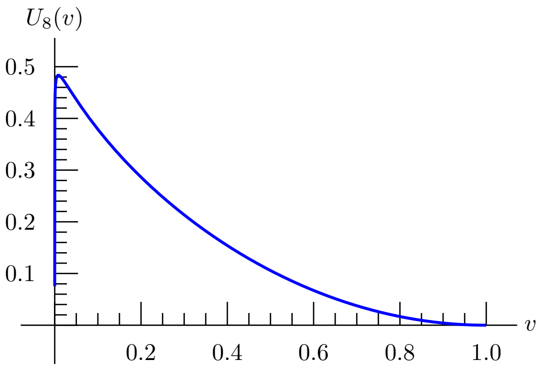


Figure 10. Plot of $U_8(v)$ $v \in (0,1)$ (Counterexample 2.5).

2.2. New classes of lifetime distributions. Barlow and Proschan [2] and Zacks [25] both provided several classifications and applications of life distributions of systems. The concept of ageing serves the foundation for nearly all the existing classes of life distributions. Such classifications play a pivotal role in survival data modeling, broadly reliability engineering. As the notion of ageing is inappropriate for classifying life distributions on inactivity time, Nanda and Paul [15] proposed an additional class of life distribution conditional on past entropy. It is desired for a system found to be broken-down upon observing to have lesser inactivity period. Thus, the class of increasing uncertainty in past lifetime is significant to group distributions. Intuitively, it is recommended to analyze the varentropy of systems having increasing uncertainty in past lifetime and classify them over and above. Thus, to further group distributions, new classes of life distributions ought to be initiated, since the past entropy class is inadequate to extract information on dispersion. Besides, it is equally important to analyze the variability aspect of the existing classes. This might be productive to establish relation among the pre-existing classes functional to further refine life distributions. We expand the study by presenting novel classes of life distributions constructed using past varentropy.

Definition 2.3. A non-negative, absolutely continuous random variable X is said to have increasing (decreasing) past varentropy, written as IPVE (DPVE), if $V_s(X_{(t)})$ increases (decreases) in t.

The following example exhibits that the above defined classes are not void.

Example 2.4. (i) Suppose that X follows standard exponential distribution. Then the past varentropy of X increases with t as asserted in Figure 11. Thus, X belongs to the IPVE class. While plotting the curve, the substitution $t = -\log v$ is used.

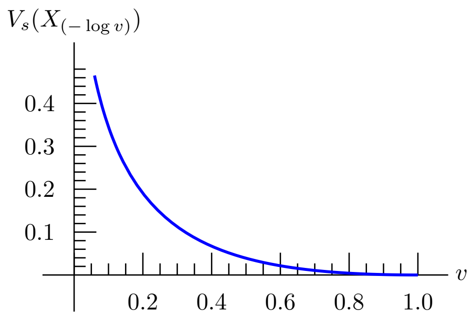


Figure 11. Plot of past varentropy (Example 2.4 (i)).

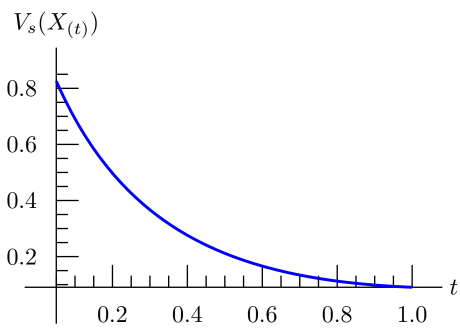


Figure 12. Plot of past varentropy (Example 2.4 (ii)).

(ii) Let X be a random variable with df $F(t) = \mathrm{e}^{(1-1/t)}\mathbf{I}_{(0,1)}(t)$. Then X belongs to DPVE class as shown in Figure 12.

Definition 2.4. A random variable X is said to have constant past varentropy, written as CPVE, if $V_s(X_{(t)})$ is constant.

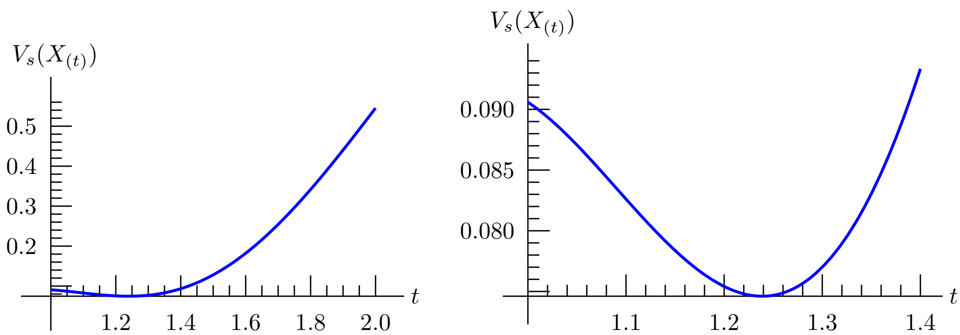
One may find several distributions that belong to the CPVE class. For instance, the uniform distribution, the standard power distribution with the shape parameter $\beta = 2$ belong to the CPVE class.

Looking at the example below, we acknowledge that perhaps not every distribution has a monotone $V_s(X_{(t)})$.

Example 2.5. Let X be the random variable having df given by

$$F(t) = \begin{cases} \exp\left(-\dfrac{1}{2} - \dfrac{1}{t}\right) & \text{for } 0 \leqslant t \leqslant 1, \\ \exp\left(-2 + \dfrac{t^2}{2}\right) & \text{for } 1 \leqslant t \leqslant 2. \end{cases}$$

It can be clearly distinguished from Figure 13 that $V_s(X_{(t)})$ is not monotone with respect to t.



(a) $V_s(X_{(t)})$ for $t \in (1,2)$.

(b) Magnified view of $V_s(X_{(t)})$ about $t = 1.25$.

Figure 13. Graph of $V_s(X_{(t)})$ (Example 2.5).

As mentioned earlier, varentropy evaluation is not only cumbersome at times, but also exhausting. For the distributions of interest that are transformation of one with known class of its volatility and the concern is restricted to having knowledge on its classification of volatility, the following theorem applies to show that the proposed classes are inherited under affine transformations. The proof is dropped being straightforward.

Theorem 2.3. *Suppose $Z = \alpha X + \beta$, where $\alpha > 0$, $\beta \geqslant 0$ and X is an absolutely continuous non-negative random variable. Then*

(i) *X belonging to IPVE class $\Longrightarrow$ Z belongs to IPVE class;*

(ii) *X belonging to DPVE class $\Longrightarrow$ Z belongs to DPVE class.*

Nanda and Paul [15] defined the increasing past entropy (IUPL) class of life distribution. It is of the purpose to investigate the connection of the IUPL class with the DPVE and IPVE classes. The following two counterexamples show that IPVE and DPVE classes are different from the IUPL class.

Counterexample 2.6. Suppose X is defined as in Example 2.4 (ii). Then it can be seen from Figure 14 and Figure 12, respectively, that $X \in$ IUPL but $X \notin$ IPVE.

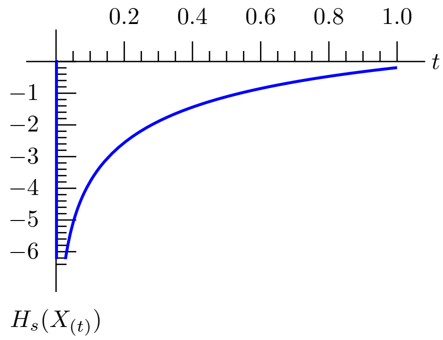


Figure 14. Plot of past entropy (Counterexample 2.6).

Counterexample 2.7. Suppose X is defined as in Example 2.4 (i). Then $X \in$ IUPL but $X \notin$ DPVE as seen from Figure 15 and Figure 11, respectively. The substitution $t = -\log v$, $v \in (0, 1)$ has been applied to plot the past entropy.

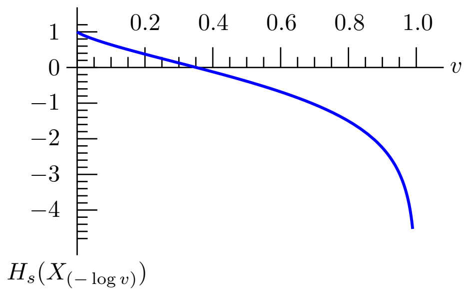


Figure 15. Plot of past entropy (Counterexample 2.7).

3. Applications

In this section, we first provide an illustration of how the PVE order can be used in reliability modeling to determine the distribution that is best fitted. Next, we consider PVE order of the particles in the case of Boltzmann distribution at different temperature levels to study the diversity of distributions.

3.1. PVE order in reliability modeling. In an effort to assess the efficacy of past varentropy order in reliability modeling, we take into consideration two real-life data sets. It is to be noted that our method is parametric involving certain families of distributions. The strength of 1.5 cm glass fibres was measured in the first data set, while the remission times (in months) of a random sample of 128 bladder cancer patients are shown in the second data set (cf. Shanker et al. [20]). The data sets are reproduced below.

Data set 1. 0.55, 0.74, 0.77, 0.81, 1.24, 0.93, 1.04, 1.11, 1.13, 1.30, 1.25, 1.27, 1.28, 1.29, 1.48, 1.36, 1.39, 1.42, 1.48, 1.51, 1.49, 1.49, 1.50, 1.50, 1.55, 1.52, 1.53, 1.54, 1.55, 1.61, 1.58, 1.59, 1.60, 1.61, 1.63, 1.61, 1.61, 1.62, 1.62, 1.67, 1.64, 1.66, 1.66, 1.66, 1.70, 1.68, 1.68, 1.69, 1.70, 1.78, 1.73, 1.76, 1.76, 1.77, 1.89, 1.81, 1.82, 1.84, 1.84, 2.00, 2.01, 2.24 and 0.84.

Data set 2. 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 6.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63 and 22.69.

The following two distributions were fitted by Shanker et al. [20] to analyse the data sets.

(i) The Lindley distribution

$$f(x) = \frac{\mu^2}{\mu+1}(1+x)\mathrm{e}^{-\mu x}, \quad \mu, x > 0$$

with parameter values 0.996116 and 0.196045 to the data sets 1 and 2, respectively.

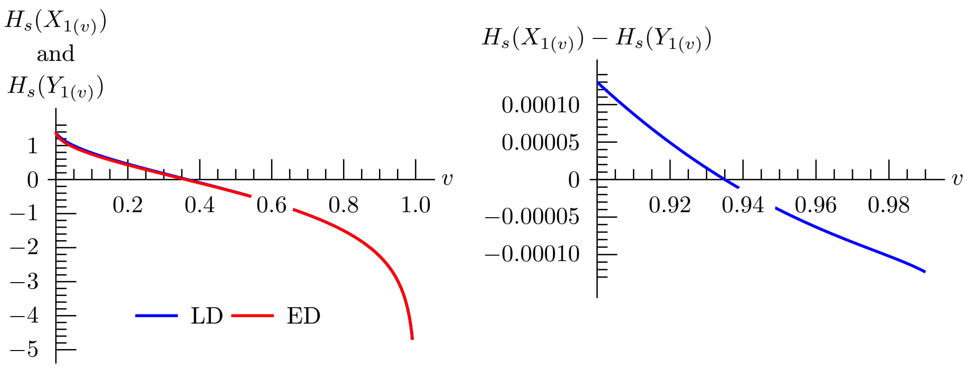
(ii) The Exponential distribution

$$f(x) = \theta \mathrm{e}^{-\theta x}, \quad \theta, x > 0$$

with parameter values 0.663647 and 0.106773 to the data sets 1 and 2, respectively.

In data set 1, they asserted that the Lindley distribution fit the data better than the exponential distribution, whereas in data set 2, they asserted that the exponential distribution fit the data better than the Lindley distribution. We aim to investigate the above fitting of statistical models to the given data sets in terms of past varentropy.

It is well known that the entropy measures the uncertainty or randomness and the probability distribution with maximum entropy is the one that most accurately reflects the state of current knowledge. Traditionally, one must first compute negative log-likelihood criterion, Alkaikes-information criterion (AIC), Bayesian information criterion (BIC), Cramer-von Mises (C^*) measure and Anderson-Darling (A^*) measure along with p-values in order to apply goodness of fit test. The distribution for which the values of C^* and A^* are smaller and p-values greater, fits the data set better than the others. In an effort to avoid such a drawn-out process, recently Kundu and Singh [12] applied the novel idea of using entropy to select the better fitted distribution following Burnham and Anderson [5]. According to the information-theoretic method for choosing a model due to Burnham and Anderson [5], the model with the maximum entropy corresponding to it is the one that best fits a given data set. Thus, if we assume to have more than one distribution that fits to a data set, then the best fitted distribution is the one with maximum entropy.

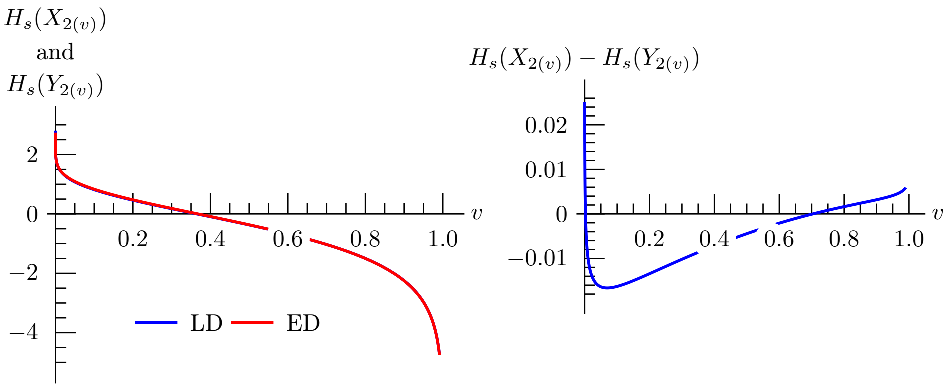


(a) Plot of $H_s(X_{1(v)})$ and $H_s(Y_{1(v)})$.

(b) Plot of $H_s(X_{1(v)}) - H_s(Y_{1(v)})$.

Figure 16. Plot of past entropies and their differences for data set 1.

Let X_1, Y_1 be the non-negative random variables that follow Lindley distribution (LD) and Exponential distribution (ED) with the specified parameter values for data set 1, respectively. Then Figure 16 (a) shows the plots of corresponding past entropies. It is clear from Figure 16 (b) that neither the past entropy of X_1 (LD) nor the past entropy of Y_1 (ED) dominates the other, i.e., the difference $H_s(X_{1(t)}) - H_s(Y_{1(t)})$ passes through the x-axis. In other words, X_1 and Y_1 fails to be in PE order. Similarly, let X_2, Y_2 be the random variables following LD and ED with specified parameters for the data set 2. Then the plots of corresponding past entropy are given in Figure 17 (a). Again, the past entropy of X_2 is neither less than nor greater than that of Y_2 for the data set 2 as can be seen from Figure 17 (b). Thus, X_2 and Y_2 fails to be in PE order. As a consequence, the information-theoretic method for choosing the best fitted model fails. However, in order to determine the best fitted model, it is encouraged to bring past varentropy into consideration. Intuitively speaking, it is desirable for the best fitted model to have maximum entropy but least varentropy as well. Figure 18 and Figure 19 show the plots of past varentropies for the distributions specified for data sets 1 and 2, respectively. Clearly, the past varentropy of LD is less when compared to the past varentropy of ED for the data set 1 (i.e., $X_1 \geqslant_{\mathrm{PVE}} Y_1$), whereas it is greater than that of ED for the data set 2 (i.e., $Y_2 \geqslant_{\mathrm{PVE}} X_2$) that enables us to choose the best fitted model as X_1 and Y_2, respectively. Take a note on the transformation $t = -\log v$, where $v \in (0, 1)$ is used while plotting the curves. As a result, even if the past entropy fails to choose the best fitted model, the past varentropy serves the purpose. The above observation validates the choice of the best fitted model as identified by Shanker et al. [20]. Hence, the PVE order is found to have conclusive applications in reliability modeling.



(a) Plot of $H_s(X_{2(v)})$ and $H_s(Y_{2(v)})$. (b) Plot of $H_s(X_{2(v)}) - H_s(Y_{2(v)})$.

Figure 17. Plot of past entropies and their differences for data set 2.

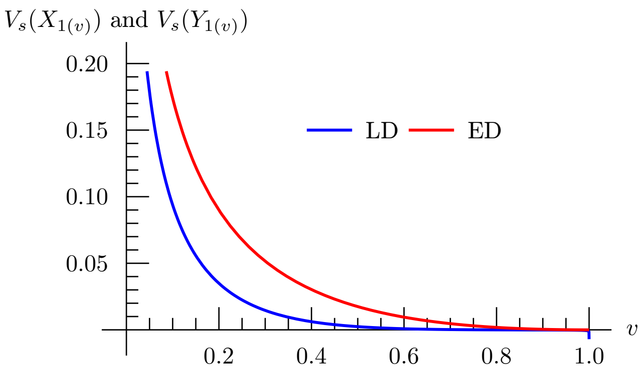


Figure 18. Plot of $V_s(X_{1(v)})$ and $V_s(Y_{1(v)})$.

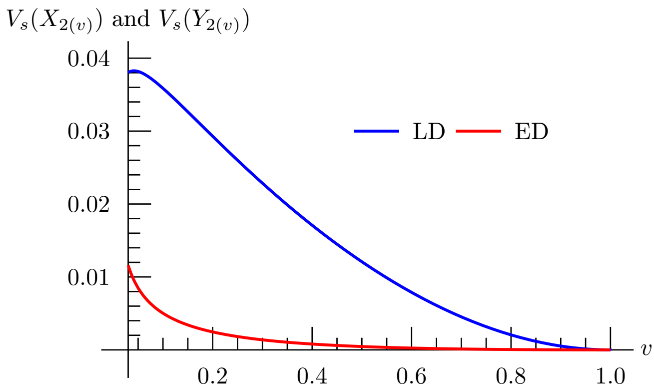


Figure 19. Plot of $V_s(X_{2(v)})$ and $V_s(Y_{2(v)})$.

3.2. PVE order of Boltzmann distribution in one dimension. The number of equivalent equiprobable types necessary to maintain a given level of entropy, or the quantity of Shannon-equivalent equiprobable forms, is referred to as a probability distribution's diversity. The true diversity measure D is related to entropy as $H = \log D$. It determines the range of equiprobable values of x giving the same amount of entropy. Rajaram et al. [17] measured the distribution of diversity by calculating diversity fractional contribution D_c upto cumulative probability 'c' by replacing entropy with the conditional entropy, i.e., past entropy. It indicates the energy states that are needed to explain information up to 'c' within the distribution. The Shannon past entropy has been useful as it simplifies the diversity index into a single value. The comparison of concentration of information around the entropy may find its usefulness to analyze the scatterness in the fractional diversity contribution C_c (or case-based entropy) and advance the study to better analyze evenness and richness. We compare the Boltzmann distribution's varentropy of past life in one dimension.

The Boltzmann distribution law explains how the states of a system are populated based on the energy of the particles which is given by

$$p_B(E) = \Big(\frac{1}{k_B T}\Big)\mathrm{e}^{-E/k_B T} = \frac{\beta}{\mathrm{e}^{\beta E}}, \tag{3.1}$$

where p_B, k_B, T and E are the probability density of particle's kinetic energy, Boltzmann constant, temperature and energy, respectively. It is worth noting that $\beta = 1/k_B T$, which expresses the response of entropy to an increase in energy. When a small amount of energy is applied to the system, β defines how much the system will randomize with regard to temperature. As the temperature rises, the most probable kinetic energy increases, shifting the entire distribution towards higher kinetic energies. Furthermore, the fraction of gas molecules with kinetic energies greater than a certain threshold also rises significantly as the temperature rises. The generalized Boltzmann distribution is often derived from the concept of maximum entropy. As β indicates the amount of randomness to be expected in the system depending on temperature, it results to variation in the fractional diversity measure. In other words, since the diversity measure is dependent on the average uncertainty of a section of the distribution not beyond c, which refers to past life, it is natural to expect dissimilarity in it. Moreover, small changes in the entropy values may result in large variations in the diversity measure because of it being an exponential function of past entropy. Thus, the study of varentropy is useful to picture out the concentration region of the diversity measure. Alternatively, putting the comparison of diversity measure of two distributions would not be suitable at time points, where they have nearly equal entropy values or when comparing this measure for a distribution at sufficiently close time points t_1 and t_2 such that $t_1 < t_2$, since small difference in entropy values would produce large differences in the diversity measure. In order to estimate the variability in the measure and the case based entropy, it is advantageous to study varentropy. Moreover, as the rise in temperature results in a shift of the complete distribution towards higher kinetic energy, we anticipate less variance in the diversity contribution factor D_c which further results to less variation in the fractional diversity contribution C_c. It can be seen from Figure 20 that the past varentropy decreases on increasing temperature. That is, $V_s(X_{(t)})$ (at T_2) $\leqslant V_s(X_{(t)})$ (at T_1) if the temperature $T_1 < T_2$. Thus, the dispersion in the diversity contribution factor D_c reduces at higher temperatures. Hence, it can be concluded that the variation in the case-based entropy C_c is more at lower temperature as compared to the variation in C_c at higher temperature. The transformation $E = -\log v$, $v \in (0, 1)$ has been used to plot Figure 20.

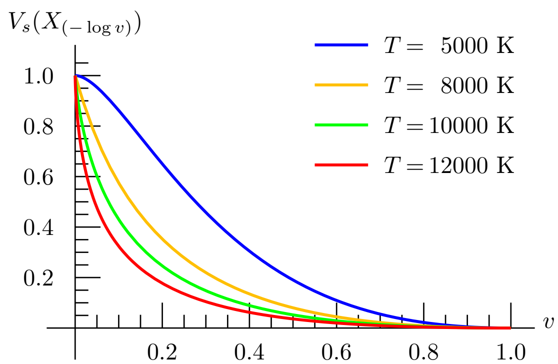


Figure 20. Past varentropy of particles energy at different temperatures.

4. Conclusions

In this study, we have initiated a novel stochastic order based on varentropy in the past life called PVE order. The distinctiveness of the order has been elucidated and its interrelation with the past entropy order and the classical stochastic orders has been researched. Further, we determined the conditions leading to the inheritance of the PVE order in the case of increasing affine transformations. The PVE ordering is also found to be preserved under increasing convex transformation subject to some conditions. Moreover, the life distributions have been classified based on the monotonicity of past varentropy resulting in new classes of life distributions. These classes have been examined under affine transformations and further it has been established that these classes differ from increasing past entropy class. Finally, the efficacy of the PVE order has been reflected in choosing the best fitted distribution among two or more distributions when the entropy method for model selection as discussed in Kundu and Singh [12] fails. Furthermore, it has also been used in analyzing the dispersion of case-based entropy for Boltzmann distribution.

References

[1] *E. Arıkan*: Varentropy decreases under polar transform. IEEE Trans. Inf. Theory *62* (2016), 3390–3400. zbl MR doi

[2] *R. E. Barlow, F. Proschan*: Statistical Theory of Reliability and Life Testing. Holt, Rinehart and Winston, New York, 1975. zbl MR

[3] *S. Bobkov, M. Madiman*: Concentration of the information in data with log-concave distributions. Ann. Probab. *39* (2011), 1528–1543. zbl MR doi

[4] *F. Buono, M. Longobardi, F. Pellerey*: Varentropy of past lifetimes. Math. Methods Stat. *31* (2022), 57–73. zbl MR doi

[5] *K. P. Burnham, D. R. Anderson*: Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York, 2002. zbl MR doi

[6] *T. M. Cover, J. A. Thomas*: Elements of Information Theory. Wiley Series in Telecommunications. John Wiley & Sons, New York, 1991. zbl MR doi

[7] *A. Di Crescenzo, M. Longobardi*: Entropy-based measure of uncertainty in past lifetime distributions. J. Appl. Probab. *39* (2002), 434–440. zbl MR doi

[8] *A. Di Crescenzo, L. Paolillo*: Analysis and applications of the residual varentropy of random lifetimes. Probab. Eng. Inf. Sci. *35* (2021), 680–698. zbl MR doi

[9] *A. Di Crescenzo, L. Paolillo, A. Suárez-Llorens*: Stochastic comparisons, differential entropy and varentropy for distributions induced by probability density functions. Available at `https://arxiv.org/abs/2103.11038` (2021), 21 pages. doi

[10] *N. Ebrahimi, S. N. U. A. Kirmani*: Some results on ordering of survival functions through uncertainty. Stat. Probab. Lett. *29* (1996), 167–176. zbl MR doi

[11] *M. Fradelizi, M. Madiman, L. Wang*: Optimal concentration of information content for log-concave densities. High Dimensional Probability VII. Progress in Probability 71. Birkhäuser, Basel, 2016, pp. 45–60. zbl MR doi

[12] *C. Kundu, S. Singh*: On generalized interval entropy. Commun. Stat., Theory Methods *49* (2020), 1989–2007. zbl MR doi

[13] *J. Liu*: Information Theoretic Content and Probability: Ph.D. Thesis. University of Florida, Gainesville, 2007. MR
[14] *S. Maadani, G. R. Mohtashami Borzadaran, A. H. Rezaei Roknabadi*: Varentropy of order statistics and some stochastic comparisons. Commun. Stat., Theory Methods *51* (2022), 6447–6460. zbl MR doi
[15] *A. K. Nanda, P. Paul*: Some properties of past entropy and their applications. Metrika *64* (2006), 47–61. zbl MR doi
[16] *A. K. Nanda, P. Paul*: Some results on generalized past entropy. J. Stat. Plann. Inference *136* (2006), 3659–3674. zbl MR doi
[17] *R. Rajaram, B. Castellani, A. N. Wilson*: Advancing Shannon entropy for measuring diversity in systems. Complexity *2017* (2017), Article ID 8715605, 10 pages. zbl doi
[18] *M. Z. Raqab, H. A. Bayoud, G. Qiu*: Varentropy of inactivity time of a random variable and its related applications. IMA J. Math. Control Inf. *39* (2022), 132–154. zbl MR doi
[19] *M. Shaked, J. G. Shanthikumar*: Stochastic Orders. Springer Series in Statistics. Springer, New York, 2007. zbl MR doi
[20] *R. Shanker, H. Fesshaye, S. Selvaraj*: On modeling of lifetimes data using exponential and lindley distributions. Biometrics Biostat. Int. J. *2* (2015), 140–147. doi
[21] *C. E. Shannon*: A mathematical theory of communication. Bell Syst. Tech. J. *27* (1948), 379–423. zbl MR doi
[22] *C. E. Shannon*: A mathematical theory of communication. Bell Syst. Tech. J. *27* (1948), 623–656. zbl MR doi
[23] *A. Sharma, C. Kundu*: Varentropy of doubly truncated random variable. Probab. Eng. Inf. Sci. *37* (2023), 852–871. MR doi
[24] *K.-S. Song*: Rényi information, loglikelihood and an intrinsic distribution measure. J. Stat. Plann. Inference *93* (2001), 51–69. zbl MR doi
[25] *S. Zacks*: Introduction to Reliability Analysis: Probability Models and Statistical Methods. Springer Texts in Statistics. Springer, New York, 1992. zbl MR doi

Authors' address: *Akash Sharma*, *Chanchal Kundu* (corresponding author), Department of Mathematical Sciences, Rajiv Gandhi Institute of Petroleum Technology, Jais 229 304, U.P., India, e-mail: `20bs0002@rgipt.ac.in`, `chanchal_kundu@yahoo.com`, `ckundu@rgipt.ac.in`.