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

Some Information Properties of Order Statistics of Skew-normal Distribution

Authors: Parisa Hasanalipour
– Department of Statistics, Ferdowsi University of Mashhad, Iran
p.hasanalipour116919@gmail.com

Mostafa Razmkhah ✉
– Department of Statistics, Ferdowsi University of Mashhad, Iran
razmkhah_m@um.ac.ir

Gholamreza Mohtashami Borzadaran
– Department of Statistics, Ferdowsi University of Mashhad, Iran
grmohtashami@um.ac.ir

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- The skew-normal distribution and some of its extensions have been considered in the last two decades in view of distribution theory and the associated properties. However, less attention has been paid to other aspects of this family of distributions. In this paper, we focus on the information properties of this distribution and the distributions of order statistics of a simple random sample from the skew-normal distribution. The Shannon's entropy as well as Kullback-Leibler divergence between the order statistics of two independent skew-normal distributions are studied. Some interesting properties of the information measures of different order statistics are presented.

Keywords:

- *Kullback-Leibler information; Order statistic; Shannon's entropy.*

AMS Subject Classification:

- 62B10, 62F10, 62G30.

✉ Corresponding author

1. Introduction and Preliminaries

The celebrated normal distribution has been known in all fields of data analysis for centuries. Its popularity has been derived from its analytical simplicity and the associated Central Limit Theorem. There are numerous situations in which the assumption of normality is not validated by the data. So, some families of near-normal distributions have played a crucial role in data analysis. Azzalini [7] introduced the skew-normal (SN) distribution and studied some of its properties. This class of distributions includes the normal distribution and possesses several properties which coincide or are close to the properties of the normal family. The random variable X is said to have skew-normal distribution, denoted by $X \sim SN(\lambda)$, if it has the following probability density function (pdf)

$$\phi(x;\lambda) = 2\phi(x)\Phi(\lambda x), \qquad x \in (-\infty, \infty), \tag{1.1}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cumulative distribution function (cdf) of standard normal distribution, respectively. The skewness parameter λ varies on the real line and controls the skewness of the distribution. From (1.1), the cdf of the $SN(\lambda)$ distribution can be expressed as

$$\Phi(x;\lambda) = \int_{-\infty}^{x} 2\phi(t)\Phi(\lambda t)dt = \Phi(x) - 2\int_{x}^{\infty}\int_{0}^{\lambda t}\phi(u)\phi(t)dudt. \tag{1.2}$$

The standard normal distribution is a special case of the SN distribution, such that $SN(0)$ coincide with the normal distribution. Moreover, as λ tends to infinity, $\phi(x;\lambda)$ tends to the half-normal density. Also, the pdf of SN distribution is a log-concave function. Many extensions of skew-normal distribution have been proposed by different authors, and some inferential aspects of them have been investigated; see, for example, Arellano-Valle et al. [3], Gómez et al. [13], Hasanalipour and Sharafi [16], Azzalini and Capitanio [8], Hasanalipour et al. [17], Hasanalipour and Razmkhah [14, 15] and Arnold et al. [6].

Let $X_1, \ldots, X_n$ be a random sample of $SN(\lambda)$ distribution; moreover, let $X_{1:n} \le \cdots \le X_{n:n}$ denote the corresponding order statistics. Then, for $1 < i < n$, the pdf of $X_{i:n}$ is given by,

$$\phi_{i:n}(x;\lambda) = c_{i,n}\phi(x;\lambda)\Phi^{i-1}(x;\lambda)[1-\Phi(x;\lambda)]^{n-i}, \tag{1.3}$$

where $c_{i,n} = i\binom{n}{i}$, also, $\phi(x;\lambda)$ and $\Phi(x;\lambda)$ are as defined in (1.1) and (1.2), respectively. For more details about order statistics and their applications, one may refer to David and Nagaraja [9] and Arnold et al. [5].
The mathematical theory of communication introduced by Shannon [20] describes logarithmic measures of information and has stimulated a tremendous amount of study in engineering fields. It is a branch of applied probability and statistics relevant to statistical inference and therefore, it should be of essential interest to statisticians. The Shannon's Entropy (SE) of a random variable X with pdf $f(\cdot)$ is given by:

$$H(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx. \tag{1.4}$$

Entropy is a measure of average uncertainty in a random variable, and also it is considered as a measure of the randomness of a probabilistic system.

The KL divergence measuring the degree of divergence between two probability distributions, is another information index considered in this paper. By assuming X and Y have pdfs $f(\cdot)$ and $g(\cdot)$, respectively, the KL divergence of $f(\cdot)$ with respect to $g(\cdot)$ is defined as

$$K(X \mid Y) = \int_{-\infty}^{\infty} f(x) \log\left(\frac{f(x)}{g(x)}\right) dx. \tag{1.5}$$

Note that $K(X \mid Y)$ becomes zero when $f(x) = g(x)$, almost everywhere. Several authors have studied the properties of the information measures of ordered data in the fields of estimation, reliability analysis, quality control, goodness of fit tests, characterization of probability distributions, and many other problems. See, for example, Ebrahimi et al. [10, 11, 12], Zarezadeh and Asadi [21], Arellano-Valle et al. [2, 4], Kayal and Kumar [19], Ardakani et al. [1], and Jose and Abdul Sathar [18].
In this paper, we study some information properties of the order statistics of simple random samples from the SN distribution. First, Shannon's entropy of the SN distribution is studied. It is proved that this measure is symmetric to the skewness parameter, such that the maximum entropy occurs when the skewness parameter is zero or equivalently when the distribution is normal. The results are extended to the distributions of order statistics of a simple random sample from the SN distribution. The relation between the entropies of lower and upper order statistics from different SN distributions with opposite signs skewness parameters is stated. The average entropy of distributions of order statistics and data distribution is also compared. Then, the Kullback-Leibler (KL) divergence between the distribution of different order statistics is investigated, and some interesting results are obtained, theoretically or numerically.

The rest of this paper is organized as follows. The Shannon's entropy of SN distribution and the distribution of order statistics are investigated in Section 2. The KL divergence and some results are studied in Section 3. Eventually, some conclusions are stated in Section 4.

2. The entropy of the skew-normal distribution

Let X have $SN(\lambda)$ distribution. Then, using (1.4), the Shannon's entropy of X is given by

$$H(X;\lambda) = -\int_{-\infty}^{\infty} \phi(x;\lambda) \log(\phi(x;\lambda)) dx. \tag{2.1}$$

Theorem 2.1. *The maximum entropy model in the skewed normal family is the normal distribution.*

Proof: Assuming $X \sim SN(\lambda)$, the proof includes three parts:
(i) $H(X;\lambda) = H(X;-\lambda)$;
(ii) $\lim_{\lambda\to\pm\infty} H(X;\lambda) = \frac{1}{2} + \log\sqrt{\frac{\pi}{2}}$;
(iii) $H(X;\lambda)$ is increasing for $\lambda < 0$ and it is decreasing for $\lambda > 0$.

A well-known property of SN distribution is that if $X \sim \phi(x;\lambda)$, then $-X = Y \sim \phi(y;-\lambda)$, hence, $H(X;\lambda) = H(X;-\lambda)$ and the proof of part (i) is complete. That is, the entropy of SN distribution does not depend on the sign of the skewness parameter.

To prove part (ii), note that equation (6) of Arellano-Valle et al. [4] with $\Phi(\lambda x)$ gives the following relationship between the entropies of $\phi(x;\lambda)$ and $\phi(x)$:

$$H(X;\lambda) - H(X;\lambda = 0) = -\log 2 - E[\log\Phi(\lambda X)],$$

where $H(X;\lambda = 0) = \frac{1}{2}\log(2\pi e)$ is the Shannon entropy of standard normal distribution. Hence,

$$H(X;\lambda) = \frac{1}{2} + \log\sqrt{\frac{\pi}{2}} - E[\log\Phi(\lambda X)]. \tag{2.2}$$

On the other hand,

$$E[\log\Phi(\lambda X)] = \int_{-\infty}^{0}\phi(x;\lambda)\log\Phi(\lambda x)dx + \int_{0}^{\infty}\phi(x;\lambda)\log\Phi(\lambda x)dx. \tag{2.3}$$

Note that for $x > 0$ as $\lambda \to \infty$, we get $\Phi(\lambda x) \to 1$, hence, the second term in (2.3) tends to zero when $\lambda \to \infty$. Moreover, for $x < 0$ as $\lambda \to \infty$, we have $\Phi(\lambda x) \to 0$ and by using L'Hopital's rule $\Phi(\lambda x)\log\Phi(\lambda x) \to 0$, hence, the first term in (2.3) also tends to zero. Therefore, $E[\log\Phi(\lambda X)] \to 0$ as $\lambda \to \infty$. Further, according to part (i), it is deduced that $\lim_{\lambda\to-\infty} H(X;\lambda) = \lim_{\lambda\to\infty} H(X;\lambda)$; hence, part (ii) is also proved.

Finally, to prove part (iii), using (2.2), we get

$$\begin{aligned}\frac{\partial}{\partial\lambda}H(X;\lambda) &= -\int_{-\infty}^{\infty} 2x\phi(x)\phi(\lambda x)\log\Phi(\lambda x)dx - \int_{-\infty}^{\infty} 2x\phi(x)\phi(\lambda x)dx \\ &= -\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{1+\lambda^2}}E\big(Y\log\Phi(\lambda Y)\big),\end{aligned}$$

where $Y \sim N(0, \frac{1}{1+\lambda^2})$. By Stein's lemma, if $Y \sim N(0,\sigma^2)$, then $E[Yg(Y)] = \sigma^2 E[g'(Y)]$, provided that $g(\cdot)$ is a function for which both expectations $E[Yg(Y)]$ and $E[g'(Y)]$ exist. Therefore,

$$\frac{\partial}{\partial\lambda}H(X;\lambda) = -\sqrt{\frac{2}{\pi}}\frac{\lambda}{(1+\lambda^2)^{\frac{3}{2}}}E[\frac{\phi(\lambda Y)}{\Phi(\lambda Y)}].$$

On the other hand, $E[\frac{\phi(\lambda Y)}{\Phi(\lambda Y)}] > 0$ for all λ. Hence, $\frac{\partial}{\partial\lambda}H(X;\lambda)$ is negative (positive) for $\lambda > (<)0$. Therefore, the entropy of $SN(\lambda)$ distribution is a unimodal symmetric function of λ that maximizes at $\lambda = 0$. This completes the proof. $\square$

Corollary 2.1. The skewness parameter orders the entropy and variance in the skewed normal family similarly, in that they both increase for $\lambda < 0$ and decrease for $\lambda > 0$. Such ordering behavior which holds for parts of supports of parameter was studied by Ebrahimi et al. [11] for the beta family.

Remark 2.1. It is clear that when $\lambda \to \infty$ (or $\lambda \to -\infty$), the $\phi(x;\lambda)$ tends to the positive (or negative) half-normal distribution with pdf $2\phi(x)$, for $x > 0$ (or $x < 0$) (Azzalini [7]). It is not difficult to show that $H(X^+) = H(X^-) = \frac{1}{2} + \log\sqrt{\frac{\pi}{2}}$, where X^+ and X^- stand for positive and negative half-normal distributions, respectively. Comparing to part (ii) of the proof of Theorem 2.1, it can be concluded that limit of entropy is equal to the entropy of a limiting distribution; precisely, $\lim_{\lambda\to\infty} H(X;\lambda) = H(X^+)$ and $\lim_{\lambda\to-\infty} H(X;\lambda) = H(X^-)$. On the other hand, from part (iii) of the proof of Theorem 2.1, it is concluded that the entropy of $SN(\lambda)$ distribution decreases to entropy of the limiting positive (or negative) half-normal case for $\lambda > 0$ (or $\lambda < 0$). Such information properties were investigated in details by Ardakani et al. [1] for symmetric families that include the normal distribution as special cases.

Using (2.1) and employing the numerical computations, the behavior of $H(X;\lambda)$ with respect to λ is shown in Figure 1. This figure confirms that the maximum entropy occurs for the case of $\lambda = 0$, which coincides with the case of standard normal distribution.

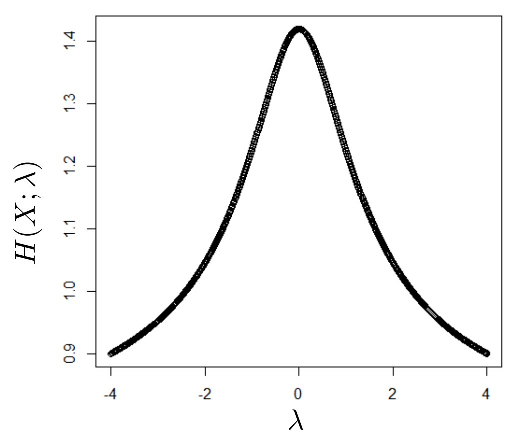

Figure 1: Plot of $H(X;\lambda)$ with respect to λ.

Now, we focus on entropy of order statistics of SN distribution. Using (1.1) and (1.3) and doing some algebraic calculations, one can show that the entropy of the ith order statistic of the SN distribution is

$$
\begin{aligned}
H(X_{i:n};\lambda) \;=\; & -\log c_{i,n} - \log\frac{2}{\sqrt{2\pi}} \\
& - E\big(\log\Phi\big(\lambda\Phi^{-1}(W;\lambda)\big)\big) + \frac{1}{2}E\big((\Phi^{-1}(W;\lambda))^2\big) \\
& - i(i-1)\Big(\psi(i) - \psi(n+1)\Big) - i(n-i)\Big(\psi(n-i+1) - \psi(n+1)\Big), \qquad (2.4)
\end{aligned}
$$

where the random variable W has the beta distribution with parameters i and $(n-i+1)$, denoted by $W \sim Beta(i, n-i+1)$.

Remark 2.2. Similar to part (i) of the proof of Theorem 2.1, it is easy to deduce that entropy of the ith order statistic from a SN distribution equals the entropy of the $(n-i+1)$th order statistic from a different SN distribution with opposite sign skewness parameter. That is, for given i and λ, we have

$$H(X_{i:n}; \lambda) = H(X_{n-i+1:n}; -\lambda). \tag{2.5}$$

From (2.4) and by using numerical computations, the behavior of $H(X_{i:n}; \lambda)$ with respect to λ is shown in Figure 2 for $n = 5$ and $i = 1, \ldots, 5$; for other values of n and i, similar figures are obtained which are omitted due to similarity. From this figure and similar ones, the following results are deduced:

- For given n and a fixed i, the entropy $H(X_{i:n}; \lambda)$ is a symmetric increasing-decreasing function of λ, such that the maximizer decreases when i goes from 1 up to n. For example, the order statistics $X_{1:5}, X_{2:5}, X_{3:5}, X_{4:5}, X_{5:5}$ get their maximum entropy at $\lambda = 1.4, 0.6, 0, -0.6, -1.4$, respectively.
- For given n, the entropies $H(X_{i:n}; \lambda)$ and $H(X_{n-i+1:n}; \lambda)$ have the same maximum entropy that confirms the relation (2.5).
- For given n, the maximum entropy of $X_{i:n}$ (or equivalently $X_{n-i+1:n}$) is decreasing in i for $i \leq \frac{n}{2}$.

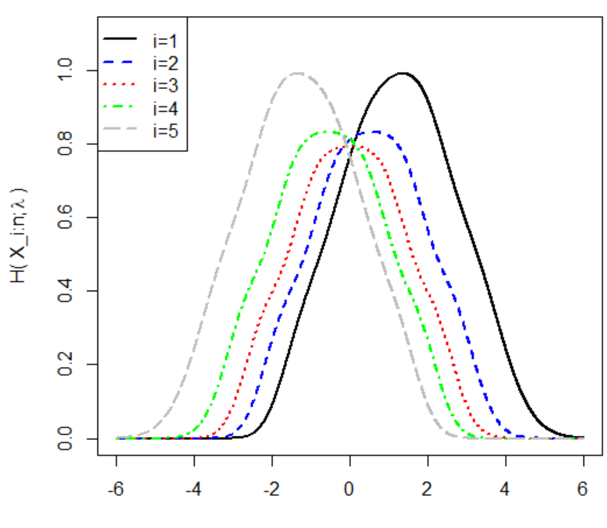

Figure 2: Plot of $H(X_{i:5}; \lambda)$ with respect to λ for $i = 1, \ldots, 5$.

For more investigation about the entropy of order statistics of a simple random sample of size n from the $SN(\lambda)$ distribution, let us define the average uncertainty of these statistics as follows:

$$\bar{H}_n(X;\lambda) = \frac{1}{n}\sum_{i=1}^{n} H(X_{i:n};\lambda). \tag{2.6}$$

This measure can be used to compare the average entropy of distributions of order statistics of a simple random sample of size n with the entropy of a single observation or data distribution.

Using (2.6), the values of $\bar{H}_n(X;\lambda)$ are calculated for some positive values of λ and some choices of n. To compare them with entropy of the SN distribution, numerical values of $H(X;\lambda)$ are also computed. The results are presented in Table 1. From this table, it is deduced that,

- Since, according to Remark 2.2, we get $\bar{H}_n(X;\lambda) = \bar{H}_n(X;-\lambda)$, it is concluded that both $\bar{H}_n(X;\lambda)$ and $H(X;\lambda)$ are symmetric increasing-decreasing functions in λ for $-\infty < \lambda < \infty$, such that $\max_\lambda \bar{H}_n(X;\lambda) = \bar{H}_n(X;0)$ and $\max_\lambda H(X;\lambda) = H(X;0)$.
- $H(X;\lambda) > \bar{H}_n(X;\lambda)$, for all values of n and λ.

Table 1. Values of $\bar{H}_n(X;\lambda)$ for some choices of λ and n.

	λ								
n	0.5	1	2	3	4	5	6	7	7.5
2	1.1576	1.0326	0.9391	0.8462	0.7570	0.6128	0.2274	0.0718	0.0265
3	1.0211	0.9961	0.8036	0.7208	0.6705	0.5446	0.2189	0.0636	0.0241
5	0.8384	0.8034	0.7123	0.6395	0.5774	0.4483	0.1941	0.0537	0.0211
10	0.5486	0.5136	0.4243	0.4114	0.4087	0.3033	0.1397	0.0411	0.0171
$H(X;\lambda)$	1.3507	1.2257	1.0456	0.9528	0.9001	0.7094	0.3461	0.0862	0.0307

3. Kullback–Leibler divergence

This section discusses some information between distributions of the ith and the jth order statistics from SN distributions with different skewness parameters. Ebrahimi et al. [10] showed that the discrimination information between a given order statistic and data distribution of the same population is distribution free. They also proved that the discrimination information among different order statistics of the same distribution is distribution free. The question arises here is that what relationship exists between the order statistics of two different distributions. To study this important subject, suppose that $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_m$ are simple random samples from different SN distributions with parameters λ_1 and λ_2, respectively. Also, let $X_{i:n}$ be the ith order statistic from X, and $Y_{j:m}$

be the jth order statistic from Y. Using (1.5), the discrimination information between the distribution of $X_{i:n}$ and $Y_{j:m}$ is obtained as

$$\begin{aligned} K_{\lambda_1,\lambda_2}(X_{i:n} \mid Y_{j:m}) &= \int_{-\infty}^{\infty} \phi_{i:n}(x;\lambda_1) \log \frac{\phi_{i:n}(x;\lambda_1)}{\phi_{j:m}(x;\lambda_2)} dx \\ &= -H(X_{i:n};\lambda_1) - \log c_{j,m} - E\Big(\log \phi\big(\Phi^{-1}(W,\lambda_1);\lambda_2\big)\Big) \\ &\quad -(j-1)E\Big(\log \Phi\big(\Phi^{-1}(W,\lambda_1);\lambda_2\big)\Big) \\ &\quad -(m-j)E\Big(\log\big(1-\Phi\big(\Phi^{-1}(W,\lambda_1);\lambda_2\big)\big)\Big), \end{aligned} \tag{3.1}$$

where $W \sim Beta(i, n-i+1)$ and $\Phi^{-1}(\cdot,\lambda)$ stands for the inverse function of $\Phi(\cdot;\lambda)$.

Remark 3.1. When $\lambda_1 = \lambda_2$, that is, in the situation in which both samples come from the same distribution, it is trivial that in the spacial case of $n = m$, the discrimination information between the distribution of $X_{i:n}$ and $Y_{i:n}$ is zero, i.e., $K_{\lambda_1,\lambda_1}(X_{i:n}, Y_{i:n}) = 0$, for $1 \leq i \leq n$. But, when $\lambda_1 \neq \lambda_2$, the KL information between the distribution of the order statistics of different distributions is positive.

Using (3.1), the KL information between the distribution of sample maxima may be obtained as

$$\begin{aligned} K_{\lambda_1,\lambda_2}(X_{n:n} \mid Y_{n:n}) &= (n-1)E(\log W) + E\Big(\log \phi(\Phi^{-1}(W,\lambda_1);\lambda_1)\Big) \\ &-E\Big(\log \phi(\Phi^{-1}(W,\lambda_1);\lambda_2)\Big) - (n-1)E\Big(\log \Phi(\Phi^{-1}(W,\lambda_1);\lambda_2)\Big). \end{aligned}$$

Similarly, the KL information between the distribution of sample minima is given by

$$\begin{aligned} K_{\lambda_1,\lambda_2}(X_{1:n} \mid Y_{1:n}) &= (n-1)E(\log(1-W)) + E\Big(\log \phi(\Phi^{-1}(W,\lambda_1);\lambda_1)\Big) \\ &-E\Big(\log \phi(\Phi^{-1}(W,\lambda_1);\lambda_2)\Big) - (n-1)E\Big(\log\Big(1-\Phi(\Phi^{-1}(W,\lambda_1);\lambda_2)\Big)\Big). \end{aligned}$$

Remark 3.2. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be two independent random samples from $SN(\lambda_1)$ and $SN(\lambda_2)$ distributions, respectively. Then, using the fact that if $X \sim \phi(x;\lambda)$, then $-X = Y \sim \phi(y;-\lambda)$, it can be simply shown that the KL divergence between two lower sample quantiles from the SN distribution with given skewness parameter equals that of upper sample quantiles from the SN distribution with opposite sign skewness parameter. More precisely, for each $1 \leq i \leq n$ and $1 \leq j \leq m$, we get

$$\begin{aligned} K_{\lambda_1,\lambda_2}(X_{i:n} \mid Y_{j:m}) &= K_{-\lambda_1,-\lambda_2}(X_{n-i+1:n} \mid Y_{m-j+1:m}), \\ K_{\lambda_1,-\lambda_2}(X_{i:n} \mid Y_{j:m}) &= K_{-\lambda_1,\lambda_2}(X_{n-i+1:n} \mid Y_{m-j+1:m}). \end{aligned}$$

The values of $K_{\lambda_1,\lambda_2}(X_{i:n} \mid Y_{j:m})$ may be numerically obtained using (3.1). Figure 3 shows the behavior of $K_{\lambda_1,\lambda_2}(X_{i:n} \mid Y_{j:m})$ to λ_2 when $\lambda_1 = 1$, $n = 6$ and $m = 3$. In fact, the KL between the pdfs of $X_{1:6}$ (or $X_{6:6}$) and $Y_{j:3}$ for $j = 1, 2, 3$ are plotted in the left (right) hand side of this figure. Analogous results may be obtained for other values of n and m.

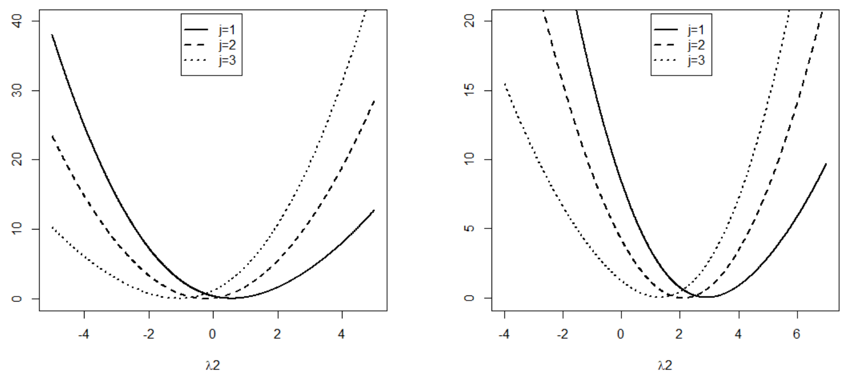

Figure 3: Plots of $K_{1,\lambda_2}(X_{i:6} \mid Y_{j:3})$ with respect to λ_2 for $i = 1$ (the left plot) and $i = 6$ (the right plot).

From Figure 3, the following results are deduced:

- For given λ_1, the KL divergence is a decreasing-increasing function of λ_2.
- From the left plot, it is observed that for given λ_1, the KL divergence between the minimum of X sample and the minimum of Y sample tends to zero when λ_2 tends to λ_1. Though, the KL of $X_{1:n}$ and $Y_{j:m}$, for $j > 1$, becomes zero for smaller λ_2.
- From the right plot, it is observed that for given λ_1, the KL between $X_{n:n}$ and $Y_{m:m}$ becomes zero when λ_2 tends to λ_1, however, the KL of $X_{n:n}$ and $Y_{j:m}$, for $j < m$, is zero for larger λ_2.
- The above results mean that for given λ_1, there exists a value such λ_2^j that the pdf of a sample quantile $X_{i:n}$ of $SN(\lambda_1)$ distribution closes to the pdf of the sample quantile $Y_{j:m}$ of $SN(\lambda_2^j)$ distribution, such that λ_2^j decreases with respect to j.

4. Conclusions

In this paper, some information properties of SN distribution and its order statistics were studied. Shannon entropy and KL criteria were investigated, and some theoretical and numerical results were obtained. The behavior of entropy of the $SN(\lambda)$ distribution to λ was studied, and it was shown that the maximum entropy model in the skewed normal family is the normal distribution. Moreover, it was deduced that limit entropy equals the entropy of the limiting distribution when λ tends to infinity. It was also shown that for fixed sample size, the entropy of a given order statistic is symmetric an increasing-decreasing function of λ in which the maximizer of entropy of $X_{i:n}$ decreases when i goes from 1 up to n; further, the maximum entropy of sample quantiles are decreasing to the maximum entropy of sample median. The maximum entropy plays an important role in choosing the best order statistics in specifying the outliers or determining the control limits in statistical quality control. Also, it is possible to compare the uncertainty of the distribution of k-out-of-n systems for different values of k or n. Some relations were also obtained for the KL divergence between distributions of the order statistics of two independent $SN(\lambda_1)$ and $SN(\lambda_2)$ distributions with respect to the variations of skewness parameters and the ranks of order statistics for given sample sizes. It was shown that for given λ_1, the KL divergence is a decreasing-increasing function of λ_2; moreover, for any given λ_1, there exists a value λ_2^j such that the pdf of the sample quantile $X_{i:n}$ of $SN(\lambda_1)$ distribution closes to the pdf of the sample quantile $Y_{j:m}$ of $SN(\lambda_2^j)$ distribution, such that λ_2^j decreases for j.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees and the associate editor for their useful comments and constructive criticisms on the original version of this manuscript which led to this considerably improved version. The research of the corresponding author is supported by a grant from Ferdowsi University of Mashhad (N. 2/55138).

REFERENCES

[1] Ardakani, O. M.; Ebrahimi, N. and Soofi, E. S. (2018). Ranking forecasts by stochastic error distance, information, and reliability measures, *International Statistical Review*, **83**, 442–468.

[2] Arellano-Valle, R. B.; Contreras-Reyes, J. E. and Genton, M. G. (2013). Shannon entropy and mutual information for multivariate skew-elliptical distributions, *Scandinavian Journal of Statistics*, **40**, 42–62.

[3] Arellano-Valle, R. B.; Cortes, M. A. and Gomez, H. W. (2010). An extension of the epsilon-skew-normal distribution, *Communications in Statistics-Theory and Methods*, **39**, 912–922.

[4] Arellano-Valle, R. B.; Contreras-Reyes, J. E. and Stehlk, M. (2017).

Generalized skew-normal negentropy and its application to fish condition factor time series, *Entropy*, **19**, 528–546.

[5] Arnold, B. C.; Balakrishnan, N. and Nagaraja, H. N. (2008). A First Course in Order Statistics, *SIAM, Philadelphia*.

[6] Arnold, B. C.; Gallardo, D. I. and Gómez, H. W. (2021). A new bivariate Birnbaum-Saunders type distribution based on the skew generalized normal model, *REVSTAT-Statistical Journal*, **21**, 1–20.

[7] Azzalini, A. (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

[8] Azzalini, A. and Capitanio, A. (2014). *The skew-normal and related families*, Cambridge University Press.

[9] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, John Wiley, New York.

[10] Ebrahimi, N.; Soofi, E. S. and Zahedi, H. (2004). Information properties of order statistics and spacing, *IEEE Transactions on Information Theory*, **50**, 177–183.

[11] Ebrahimi, N.; Maasoumi, E. and Soofi, E.S. (1999). Ordering univariate distributions by entropy and variance, *Journal of Econometrics*, **90**, 317–336.

[12] Ebrahimi, N.; Soofi, E.S. and Soyer, R. (2010). Information measures in perspective, *International Statistical Review*, **78**, 383–412.

[13] Gómez, H. W. and Salinas, H. S. (2010). Information matrix for generalized skew-normal distributions, *Proyecciones (Antofagasta)*, **29**, 83–92.

[14] Hasanalipour, P. and Razmkhah, M. (2020). Testing skew-symmetry based on extreme ranked set sampling, *Statistical Papers*, **62**, 2311–2332.

[15] Hasanalipour, P. and Razmkhah, M. (2022). Inference on skew-normal distribution based on Fisher information in order statistics, *Communications in Statistics - Simulation and Computation*, **51**, 1525–1541.

[16] Hasanalipour, P. and Sharafi, M. (2012). A new generalized Balakrishnan skew-normal distribution, *Statistical Papers*, **53**, 219–228.

[17] Hasanalipour, P.; Hassani Oskouei, L.; Jabbari Khamnei, H. and Khorshidian, K. (2017). The Beta skew-generalized normal distribution, *Communications in Statistics-Theory and Methods*, **46**, 2270–2276.

[18] Jose, J. and Abdul Sathar, E. I. (2021). Renyi Entropy of k-Records: Properties and Applications. *REVSTAT-Statistical Journal*, **20**, 481–500.

[19] Kayal, S. and Kumar, S. (2017). Estimating Renyi entropy of several exponential distributions under an asymmetric loss function, *REVSTAT-Statistical Journal*, **15**, 501-522.

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

[21] Zarezadeh, S. and Asadi, M. (2010). Results on residual Renyi entropy of order statistics and record values, *Information Sciences*, **180**, 4195–4206.