ENTROPY FOR PARETO-TYPES AND ITS ORDER STATISTICS DISTRIBUTIONS

GHOLAM-HOSSEIN YARI* AND G. R. MOHTASHAMI BORZADARAN†

Abstract. The aim of this paper is deriving the exact analytical expressions of entropy for the Pareto-types and related distributions. Entropy for i^{th} order statistics of these distributions corresponding to the random sample size n is introduced. We have shown that all the expressions related to finding corresponding entropy for these families and their order statistics distributions are obtained through some particular techniques via integration. Indeed, these techniques for that improper integrals has its own importance also.

AMS Classification: Primary 62B10; Secondary 94A17

Keywords: Entropy, Pareto distributions, Gamma functions, Beta functions, Polygamma functions, Order Statistics, Burr distributions.

1. Introduction. The idea of the entropy of random variables is developed by Claude Shannon (1948), for the first time in information theory. The applications of entropy originated in the nineteenth century in the field of Statistical Mechanics and Thermodynamics. During the last sixty years or so, a number of research papers and monographs discussing and extending Shannon's original work have appeared. In this paper, the exact form of the entropy for the Pareto-types and related distributions is determined. Entropy for the ith order statistics corresponding to the independent random variable with sample size n for these distributions is given also. These distributions have important role as parametric models in reliability, actuarial science, economics, finance and telecommunications. The Shannon entropy of a random variable X taking its values in R with probability density function $f_X(x)$, is defined by

$$H_X = -\int_R f(x)\ln f(x)dx,$$

where existence of the integral is one of our main conditions. Analytical expression for the entropy of univariate distributions are discussed in references such as Cover and Thomas (1991), Verdugo Lazo and Rathie (1978) and so on. Recently, Bercher and Vignat (2008) derived generalized Pareto distribution as the Renyi Tsallis maximum entropy under some constraints. This distribution is widely used in practice for modeling the tails of the distributions.

*School of Mathematics, Iran University of Science and Technology, Tehran-IRAN. E-mail: Yari@iust.ac.ir

†The author is member of the "Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad". Department of Statistics, Ferdowsi University of Mashhad, Mashhad-IRAN. E-mail : grmohtashami@um.ac.ir, gmb1334@yahoo.com

In this paper, the entropy expression for Pareto (IV) distribution is given in section 2, entropy for the ith order statistics of this distribution is given by section 3, and entropy expressions for Pareto-type (III), (II), (I) and i^{th} order statistics of these distributions corresponding to the random samples are discussed in section 4. The entropy expression of Burr (XII) distribution and i^{th} order statistics of it, is given in section 5.

   2. Pareto (IV) distribution. A hierarchy Pareto distribution is established by starting with the classical Pareto (I) distribution. Extended version of it, is introducing additional parameters related to location, scale, shape and inequality (Gini index) via the Pareto (IV) distribution, as discussed in chapter 3 of Arnold (1983). Such an approach leads to a general family of distributions, called the Pareto (IV) family, with the cumulative distribution function,

$$(1) \qquad F_X(x) = 1 - \left(1 + \left(\frac{x - \mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{-\alpha}, \quad x > \mu,$$

where $-\infty < \mu < +\infty$, $\theta > 0$, $\gamma > 0$ and $\alpha > 0$ are respectively, location parameter, scale parameter, inequality parameter and the shape parameter which characterizes the tail of the distribution. We denote this distribution as Pareto (IV) $(\mu, \theta, \gamma, \alpha)$. γ is called the inequality parameter because of its interpretation in the economics context. If we choose $\alpha = 1$ and $\mu = 0$ in expression (1), the parameter $(\gamma \leq 1)$ is precisely the Gini index of the inequality. For the Pareto (IV) $(\mu, \theta, \gamma, \alpha)$ distribution, we have the density function as,

$$(2) \qquad f_X(x) = \frac{\alpha \left(\frac{x-\mu}{\theta} \right)^{\frac{1}{\gamma}-1}}{\theta \gamma \left(1 + \left(\frac{x-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{\alpha+1}}, \quad x > \mu,$$

such that the entropy of this density can be found via :

$$(3) \quad H_X = -\int_x f(x) \ln f(x) dx = -\ln \left(\frac{\alpha}{\theta \gamma} \right)$$
$$+ \left(1 - \frac{1}{\gamma} \right) E \left[\ln \left(\frac{X - \mu}{\theta} \right) \right] + (\alpha + 1) E \left[\ln \left(1 + \left(\frac{X - \mu}{\theta} \right)^{\frac{1}{\gamma}} \right) \right].$$

We need to find the expressions $E \left[\ln \left(\frac{X-\mu}{\theta} \right) \right]$ and $E \left[\ln \left(1 + \left(\frac{X-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right) \right]$. First we calculate the expectation of

$$E \left[\left(\frac{X - \mu}{\theta} \right)^r \right] = \int_x \left(\frac{x - \mu}{\theta} \right)^r f(x) dx,$$

by using the change of variable,

$$1 + \left(\frac{X - \mu}{\theta} \right)^{\frac{1}{\gamma}} = \frac{1}{1 - t} \quad 0 < t < 1,$$

we obtain,

$$(4) \qquad E\left[\left(\frac{X-\mu}{\theta}\right)^r\right] = \alpha\frac{\Gamma(r\gamma+1)\Gamma(\alpha-r\gamma)}{\Gamma(\alpha+1)} \quad,$$
$$\alpha - r\gamma \neq 0, -1, -2, \cdots.$$

Differentiating both sides of (4) with respect to r leads to,

$$(5) \quad \frac{d}{dr}E\left[\left(\frac{X-\mu}{\theta}\right)^r\right] = E\left[\left(\frac{X-\mu}{\theta}\right)^r \ln\left(\frac{X-\mu}{\theta}\right)\right]$$
$$= \frac{1}{\Gamma(\alpha)}\left[\gamma\Gamma'(r\gamma+1)\Gamma(\alpha-r\gamma) - \gamma\Gamma'(\alpha-r\gamma)\Gamma(r\gamma+1)\right].$$

Using relation (5), and $r = 0$, we obtain,

$$(6) \qquad E\left[\ln\left(\frac{X-\mu}{\theta}\right)\right] = \gamma\left[\Psi(1) - \Psi(\alpha)\right],$$

where Ψ is the digamma function that is defined by $\Psi(z) = \frac{d}{dz}\ln\Gamma(z)$ such that Γ is the gamma function.

Differentiating both sides of the following relation with respect to α,

$$(7) \qquad 1 = \int_\mu^{+\infty} f(x)dx,$$

leads to

$$(8) \qquad E\left[\ln\left(1 + \left(\frac{X-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)\right] = \frac{1}{\alpha}.$$

Substitute (6) and (8) in relation (3) implies that :

$$(9) \qquad H_X = \ln\left(\frac{\gamma\theta}{\alpha}\right) + (\gamma-1)\left[\Psi(1) - \Psi(\alpha)\right] + \frac{\alpha+1}{\alpha}.$$

3. Entropy for Order Statistics. Let $X_1, X_2, \cdots, X_n$ be a random sample of the probability density function (2) and let $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n-1:n} \leq Y_{n:n}$ denotes the corresponding order statistics; then

$$g_{i:n}(y) = n\binom{n-1}{i-1}[F_{i:n}(y)]^{i-1}[1 - F_{i:n}(Y)]^{n-i}f_X(y)$$
$$= \frac{n\alpha}{\theta\gamma}\binom{n-1}{i-1}\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}\left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha(n-i+1)-1}$$
$$(10) \qquad \left[1 - \left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right]^{i-1} \quad, y > \mu.$$

The entropy expression of the $g_{i:n}(y)$ is :

$$H_{i:n}(Y) = -\ln\left[\frac{n\alpha}{\theta\gamma}\binom{n-1}{i-1}\right] + \left(1 - \frac{1}{\gamma}\right)E\left[\ln\left(\frac{Y-\mu}{\theta}\right)\right]$$

$$+ \left[\alpha(n-i+1)+1\right]E\left[\ln\left(1+\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right] +$$

(11)
$$(1-i)E\left[\ln\left(1-\left(1+\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right].$$

We need to calculate the expressions of :

$$E\left[\ln\left(\frac{Y-\mu}{\theta}\right)\right], \quad E\left[\ln\left(1+\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]$$

and

$$E\left[\ln\left(1-\left(1+\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right],$$

by deriving based on the following statements. So, firstly,

$$E\left[\left(\frac{Y-\mu}{\theta}\right)^r\right] = \int_\mu^\infty \left(\frac{y-\mu}{\theta}\right)^r g_{i:n}(y)dy =$$

$$\frac{n\alpha}{\theta\gamma}\binom{n-1}{i-1}\int_\mu^\infty\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}+r-1}\left(1+\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha(n-i+1)-1}$$

(12)
$$\left[1-\left(1+\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right]^{i-1} dy,$$

and on choosing $t = 1+\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}$, $\quad 0 < t < 1$, we have :

$$E\left[\left(\frac{Y-\mu}{\theta}\right)^r\right] = n\alpha\binom{n-1}{i-1}\sum_{l=0}^{i-1}(-1)^l\binom{i-l}{l}$$

(13)
$$\frac{\Gamma(r\gamma+1)\Gamma(-r\gamma+\alpha(n-i+1+l))}{\Gamma(1+\alpha(n-i+1+l))}.$$

Differentiating both sides of (13) with respect to r, and putting $r = 0$, we obtain :

$$E\left[\ln\left(\frac{Y-\mu}{\theta}\right)\right] = n\alpha\gamma\binom{n-1}{i-1}$$

(14)
$$\sum_{l=0}^{i-1}(-1)^l\binom{i-1}{l}\left[\frac{\Psi(1)-\Psi(\alpha(n-i+1+l))}{\alpha(n-i+1+l)}\right].$$

Secondly,

$$\phi(r) = E\left[\left(1 + \left(\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^r\right] = \frac{n\alpha}{\theta\gamma}\binom{n-1}{i-1}\int_\mu^\infty \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}$$

$$(15) \qquad \left[\left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{r-\alpha(n-i+1)-1}\right]\left[1 - \left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right]^{i-1} dy,$$

using the change of variable, $t = \left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}$, $\quad 0 < t < 1$, we have,

$$\phi(r) = n\binom{n-1}{i-1}\int_0^1 t^{n-i-\frac{r}{\gamma}}(1-t)^{i-1}dt$$

$$(16) \qquad = \frac{\Gamma(n+1)\Gamma\left(n-i-\frac{r}{\alpha}+1\right)}{\Gamma(n-i+1)\Gamma\left(n-\frac{r}{\alpha}+1\right)},$$

and

$$\frac{d\phi(r)}{dr}\|_{r=0} = E\left[\ln\left(1 + \left(\frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)\right]$$

$$(17) \qquad = \frac{\Psi(n+1) - \Psi(n-i+1)}{\alpha}.$$

Thirdly, it can be shown that :

$$k(r) = E\left[\ln\left(1 - \left(1 + \frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right]^r$$

$$(18) \qquad = \frac{\Gamma(n+1)\Gamma(r+i)}{\Gamma(i)\Gamma(r+n+1)},$$

and

$$(19) \qquad \frac{dk(r)}{dr}\|_{r=0} = E\left[\ln\left(1 - \left(1 + \frac{Y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}\right] = \Psi(i) - \Psi(n+1).$$

Putting from (14), (17) and (19) in relation (11) implies that,

$$H_{i:n}(Y) = -\ln\left[\frac{n\alpha}{\theta\gamma}\binom{n-1}{i-1}\right] + (\gamma-1)n\binom{n-1}{i-1}$$

$$\sum_{l=0}^{i-1}(-1)^l\binom{i-1}{l}\left[\frac{\Psi(1) - \Psi(\alpha(n-i+1+l))}{n-i+1+l} + [\alpha(n-i+1)+1]\right]$$

$$(20) \qquad \left[\frac{\Psi(n+1) - \Psi(n-i+1)}{\alpha}\right] + (i-1)\left[\Psi(n+1) - \Psi(i)\right].$$

In particular cases,

$$H_{n:n}(Y) = -\ln\left(\frac{n\alpha}{\theta\gamma}\right) + n(\gamma-1)\sum_{l=0}^{n-1}(-1)^l\binom{n-1}{l}$$

$$(21) \qquad \left[\frac{\Psi(1) - \Psi(\alpha(1+l))}{1+l}\right] + \left(\frac{n-1}{n}\right) + \left(\frac{\alpha+1}{\alpha}\right)\left[\Psi(n+1) - \Psi(1)\right],$$

and

$$(22) \qquad H_{1:n}(Y) = -\ln\left[\frac{n\alpha}{\theta\gamma}\right] + (\gamma - 1)\left[\Psi(1) - \Psi(n\alpha)\right] = H_X(n\alpha).$$

Also, the entropy for median where $n = 2m + 1, i = m + 1$ is the following expression:

$$H_{m+1:n}(Y) = -\ln\left[\frac{n\alpha}{\theta\gamma}\binom{n-1}{m}\right] + (\gamma - 1)n\binom{n-1}{m}$$

$$\sum_{l=0}^{m}(-1)^l\binom{m}{l}\left[\frac{\Psi(1) - \Psi(\alpha(n-m+l))}{n-m+l}\right] +$$

$$\left(\frac{\alpha(n-m)+1}{\alpha}\right)\left[\Psi(n+1) - \Psi(n-m)\right] +$$

$$(23) \qquad m\left[\Psi(n+1) - \Psi(m+1)\right].$$

4. Pareto (III), (II) and (I) Distributions. By setting $\alpha = 1$ in relations (1) and (2), we obtain,

$$(24) \qquad F_X(x) = 1 - \left(1 + \left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-1}, \quad x > \mu,$$

and

$$(25) \qquad f_X(x) = \frac{1}{\gamma\theta}\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}\left(1 - \left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-2}, \quad x > \mu,$$

as the distribution function and the probability density function of the Pareto type (III) respectively. The probability density function of the ith order statistics of the independent random variable with sample size n of the Pareto type (III) distribution is :

$$g_{i:n}(y) = \frac{n}{\gamma\theta}\binom{n-1}{i-1}\left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}\left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-(n-i+2)}$$

$$(26) \qquad \left[1 - \left(1 + \left(\frac{y-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right)^{-1}\right]^{i-1}, \quad y > \mu,$$

where its entropy expression is equal to :

$$H_{i:n}(Y) = -\ln\left[\frac{n}{\gamma\theta}\binom{n-1}{i-1}\right] + (\gamma - 1)n\binom{n-1}{i-1}$$

$$\sum_{l=0}^{i-1}(-1)^l\binom{i-1}{l}\left[\frac{\Psi(1) - \Psi(n-i+l+1)}{n-i+l+1}\right] +$$

$$(n-i+2)\left[\Psi(n+1) - \Psi(n-i+1)\right] + (i-1)\left[\Psi(n+1) - \Psi(i)\right].$$

In particular cases :

- If $i = 1$ and $i = n$, then we have,

$$(27) \quad H_{1:1}(Y) = -\ln\left(\frac{n}{\gamma\theta}\right) + (\gamma - 1)\left[\Psi(1) - \Psi(n)\right] = H_X(\alpha = n),$$

and

$$H_{n:n}(Y) = -\ln\left(\frac{n}{\gamma\theta}\right) + n(\gamma - 1)\sum_{l=0}^{n-1}(-1)^l\binom{n-1}{l}$$
$$\left[\frac{\Psi(1) - \Psi(l+1)}{l+1}\right] + 2\left[\Psi(n+1) - \Psi(1)\right] + \left(\frac{n-1}{n}\right),$$

respectively.

The best known Pareto distributions are types (II) and (I). A Pareto distribution of type (II) is obtained by setting $\gamma = 1$ in relation (1) with the entropy $H_X = \ln(\frac{\theta}{\alpha}) + \frac{\alpha+1}{\alpha}$. The corresponding entropy expression for the ith order statistics of this distribution is obtained by setting $\gamma = 1$ in relation (20) as :

$$H_{i:n}(Y) = -\ln\left[\frac{n\alpha}{\theta}\right] + \left[\alpha(n-i+1) + 1\right]$$
$$\left[\frac{\Psi(n+1) - \Psi(n-i+1)}{\alpha}\right] + (i-1)\left[\Psi(n+1) - \Psi(i)\right].$$

In particular cases :

- The entropy of order statistics for $i = n$ and $i = 1$ are as following :

$$H_{n:n} = -\ln\left(\frac{n\alpha}{\theta}\right) + \left[\frac{\Psi(1) - \Psi(\alpha(1+l))}{1+l}\right] +$$
$$(28) \qquad \left(\frac{n-1}{n}\right) + \left(\frac{\alpha+1}{\alpha}\right)\left[\Psi(n+1) - \Psi(1)\right],$$

$H_{1:n}(Y) = \ln\left(\frac{\theta}{n\alpha}\right) + \frac{n\alpha+1}{n\alpha} = H_X(n\alpha)$, respectively.

If one sets $\gamma = 1$ and $\mu = \theta$ in relation (1), then one gets a Pareto distribution of type (I). Since the entropy expression of the Pareto family does not depend on location parameter μ, one finds that the entropy of Pareto(II) and Pareto(I) and the entropy of their i^{th} order statistics are equal.

5. Burr (XII) Distribution. This distribution is a special case of Pareto (IV) with $\mu = 0$, and replacing γ by $\frac{1}{\gamma}$ in relations (1) and (2), where

$$(29) \qquad F_X(x) = 1 - \left(1 + \left(\frac{x}{\theta}\right)^\gamma\right)^{-\alpha}, \quad x > 0, \quad \alpha, \gamma > 0,$$

$$(30) \qquad f_X(x) = \left(\frac{\alpha\gamma}{\theta}\right)\left(\frac{x}{\theta}\right)^{\gamma-1}\left(1 + \left(\frac{x}{\theta}\right)^\gamma\right)^{-(\alpha+1)}, \quad x > 0, \quad \alpha, \gamma > 0,$$

are the cumulative distribution and the probability density function of Burr type (XII) distribution. By replacing γ by $\frac{1}{\gamma}$ in relation (9) the entropy expression of Burr type (XII) distribution is

$$(31) \qquad H_X = \ln\left(\frac{\theta}{\alpha\gamma}\right) + (\gamma - 1)\left[\frac{\Psi(\alpha) - \Psi(1)}{\gamma}\right] + \left(\frac{\alpha + 1}{\alpha}\right),$$

and

$$H_{i:n}(Y) = -\ln\left[\frac{\alpha\gamma}{\theta}n\binom{n-1}{i-1}\right] + (\gamma - 1)n\binom{n-1}{i-1}\sum_{l=0}^{i-1}(-1)^l\binom{i-1}{l}$$

$$\left[\frac{\Psi(1) - \Psi(\alpha(n-i+1+l))}{\gamma(n-i+1+l)}\right] + [\alpha(n-i+1) + 1]$$

$$(32) \qquad \left[\frac{\Psi(n+1) - \Psi(n-i+1)}{\alpha}\right] + (i-1)\left[\Psi(n+1) - \Psi(i)\right],$$

is the entropy expression of the ith order statistics corresponding to the sample size n of this distribution.

In particular case,

$$H_{n:n}(Y) = -\ln\left(\frac{n\alpha\gamma}{\theta}\right) + (\gamma - 1)n\sum_{l=0}^{n-1}(-1)^l\binom{n-1}{l}$$

$$\left[\frac{\Psi(1) - \Psi(\alpha(1+l))}{\gamma(1+l)}\right] + \left(\frac{\alpha + 1}{\alpha}\right)\left[\Psi(n+1) - \Psi(1)\right] + \left(\frac{n-1}{n}\right),$$

and $H_{1:n}(Y) = H_X(n\alpha)$.

REMARK 5.1. When $Y = w(X)$ is an invertible function of X, then $H_Y = H_X$ in the discrete case, and under some mild conditions, $H_Y = H_X + E\left(\ln|w'(X)|\right)$, in the continuous case. It is obvious that $w(x) = ax + b$ simplifies it to $H_Y = H_X + \ln|a|$. Using this idea may help in finding the entropy of the transformation of the random variables via the formula that is mentioned. In some few cases may find the results easier.

6. Conclusions. In this paper, we obtained the exact form of the entropy expression for Pareto (IV) and related distributions and the entropy for i^{th} order statistics of these distributions corresponding to the independent random samples size n. Obtaining entropy expression for members of the Burr family in addition to Burr (XII) and i^{th} order statistics of them is the future of this work.

REFERENCES

[1] B. C. ARNOLD, *Pareto distributions, International Cooperative Publishing House.* Fairland, Maryland, 1983.

[2] M. ASADI, N. EBRAHIMI, G. G. HAMEDANI, AND E. S. SOOFI, *Information measures for Pareto distributions and order statistics.* Statistics for Industry and Technology, part II, page 207-223, 2006.

[3] J. F. BERCHER AND C. VIGNAT, *An entropic view of Pickands theorem.* In: International Symposium on Information Theory ISIT2008, Toronto, 2008.

[4] T. COVER AND J. THOMAS, *Elements of information Theory.* Wiley, New York, 1991.

[5] C. SHANNON, *A Mathematical theory of communication.* Bell system technical journal, 27(1948), pp. 379-423.

[6] A. VERDUGO LAZO AND P. RATHIE, *On the Entropy of continuous probability distributions*, IEEE Trans. Inform Theory, 24(1978), pp. 120-122.