Statistics and Probability Letters 80 (2010) 1512–1519

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

β-entropy for Pareto-type distributions and related weighted distributions

Mehdi Yaghoobi Avval Riabi [a], G.R. Mohtashami Borzadaran [b,*,1], G.H. Yari [c]

[a] Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran

[b] Department of Statistics, Ferdowsi University of Mashhad, Iran

[c] Iran University of Science and Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 12 September 2009
Received in revised form 2 June 2010
Accepted 3 June 2010
Available online 16 June 2010

Keywords:
Renyi entropy
Pareto distributions
Weighted distributions
Burr distributions
Order statistics

ABSTRACT

In this paper, we derive the β-entropy for Pareto-type and related distributions. Further, the β-entropy for some weighted versions of these distributions, such as order statistics, proportional hazards, proportional reversed hazards, probability weighted moments, upper record and lower record, is obtained.

1. Introduction

The origin of the term entropy goes back to the works of Clausius (1864) and Boltzmann (1872) in Thermodynamics. The idea of information-theoretic entropy was first introduced by Shannon (1948) and later by Weiner (1949) in Cybernetics. Over the past 60 years, after Shannon (1948) introduced his measure of entropy, a huge number of papers, books and monographs have been published on its extensions and applications, among which are Renyi (1961), Havrda and Charvat (1967), Tsallis (1988), Kapur (1989), Ullah (1996), Dragomir (2003), Cover and Thomas (2006), Asadi et al. (2006) and Harremoes (2006). A well-known parametric extension of the Shannon entropy is β-entropy, which was defined by Havrda and Charvat (1967) and later studied in more detail by Tsallis (1988). Although β-entropy was first introduced by Havrda and Charvat in the context of cybernetics theory, it was Tsallis who exploited its non-extensive features and placed it in a physical setting. Hence β-entropy is also known as Tsallis entropy. In recent years, authors have shown more interest in studying the properties and applications of Tsallis entropy. For more details, one can see Plastino and Plastino (1999), Tsallis (2002), Tsallis and Brigatti (2004), Jizba and Arimitsu (2004), Suyari (2004), Bercher (2008), Singh Pharwaha and Singh (2009) and Herrmann (2009).

The concept of a weighted distribution, which was introduced by Rao (1965), has many applications in different areas of statistics such as survey sampling, reliability, biostatistics, etc. Many well-known probability models, such as probability density functions of order statistics, record values, proportional hazards and proportional reversed hazards, can be considered as weighted distributions. Motivated by this, we focus on the relation between the β-entropy of the parent

* Corresponding author at: Department of Statistics, Ferdowsi University of Mashhad, Iran.
E-mail addresses: m_yaghoobiawal@yahoo.com (M.Y.A. Riabi), gmb1334@yahoo.com (G.R. Mohtashami Borzadaran), Yari@iust.ac.ir (G.H. Yari).

[1] Member of Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad, Mashhad, Iran.

0167-7152/$ – see front matter
doi:10.1016/j.spl.2010.06.003

distribution and the β-entropy of corresponding weighted distributions, especially the distribution of order statistics, record values, proportional hazards and proportional reversed hazards. Recently, Park (1995), Ebrahimi et al. (2004) and Oluyede (2006) obtained various results on the information properties of order statistics. Yari and Mohtashami Borzadaran (in press) obtained the Shannon entropy for Pareto-type distributions and their order statistics. Baratpour et al. (2007) obtained some results related to the Shannon entropy and Renyi entropy for record values. Hegde et al. (2005) found an order statistics based estimator for the Renyi entropy. Belzunce et al. (2004) and Di Crescenzo and Longobardi (2006) investigated the residual and past entropy for weighted distributions.

Pareto-type distributions are flexible parametric models with applications in many branches of science such as reliability, actuarial science, economics, finance and telecommunications.

In this paper, we investigate the β-entropy of Pareto-type distributions and different associated weighted distributions such as those for order statistics, record values, proportional hazards, proportional reversed hazards and probability weighted moments. This paper is organized as follows. In Section 2, we present some preliminary results which will be used in subsequent sections. Section 3 contains the main results of the paper. In this section, we obtain the β-entropy for a general Pareto-type distribution, which has four parameters, and then we tabulate values for some specialized versions. Further, we derive the β-entropy for the order statistics, record values, proportional hazards and proportional reversed hazards related to Pareto-type distributions. Finally, Section 4 is devoted to the conclusions.

2. Preliminaries

Pareto distributions provide models for many applications in social, natural and physical sciences and are related to many other families of distributions.

A hierarchy of the Pareto distributions has been established starting from the classical Pareto (I) distribution, and subsequently additional parameters related to location, scale, shape and inequality are introduced. A general version of this family of distributions is called the Pareto (IV) distribution, which is discussed in Arnold (1983). The cumulative distribution function of the Pareto (IV) distribution is

$$F(x) = 1 - \left[1 + \left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{-\alpha}, \quad x > \mu, \tag{1}$$

where $-\infty < \mu < \infty, \theta > 0, \gamma > 0$ and $\alpha > 0$ are location, scale, inequality and shape parameters, respectively. This distribution is denoted by Pareto (IV) $(\mu, \theta, \gamma, \alpha)$, and its density function is as follows:

$$f(x) = \frac{\alpha\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}}{\theta\gamma\left[1 + \left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{\alpha+1}}, \quad x > \mu. \tag{2}$$

- Setting $(\alpha = 1)$, $(\gamma = 1)$ and $(\gamma = 1, \mu = \theta)$ in relations (1) and (2), one at a time, leads to the cumulative distribution function and probability density function of the Pareto (III), Pareto (II) and Pareto (I) distributions, respectively.
- A special case of the Pareto (IV) distribution in which $\mu = 0$, $\gamma \to \frac{1}{\gamma}$ implies the Burr (XII) distribution with the following cumulative distribution and probability density functions:

$$F(x) = 1 - \left[1 + \left(\frac{x}{\theta}\right)^{\gamma}\right]^{-\alpha}, \quad x > 0, \ \alpha, \gamma > 0, \tag{3}$$

$$f(x) = \left(\frac{\alpha\gamma}{\theta}\right)\left(\frac{x}{\theta}\right)^{\gamma-1}\left[1 + \left(\frac{x}{\theta}\right)^{\gamma}\right]^{-(\alpha+1)}, \quad x > 0, \ \alpha, \gamma > 0. \tag{4}$$

2.1. Weighted distributions

Let X be a non-negative random variable with a probability density function $f(x; \theta)$, where the natural parameter is $\theta \in \Omega$ (Ω is the parameter space). The weight function $w(x, \beta)$ is a non-negative function with the parameter β representing the recording (sighting) mechanism. Corresponding to $w(x, \beta)$, we have a probability density function

$$f^w(x; \theta, \beta) = \frac{w(x, \beta) f(x; \theta)}{E[w(X, \beta)]},$$

where $E[w(X, \beta)]$ is the normalizing factor. The random variable X^w is called the weighted version of the random variable X and its distribution is called the weighted distribution with weight function w, as mentioned in Patil (2002). Some special weight functions are as follows.

- $w(x) = x^k e^{lx} F^i(x) \bar{F}^j(x)$. Setting $(l = 0)$, $(k = j = i = 0)$, $(l = i = j = 0)$, $(k = l = 0, i \to i - 1, j = n - i)$, $(k = l = i = 0)$ and $(k = l = j = 0)$ in this weight function, one at a time, implies probability weighted moments, moment-generating functions, moments, order statistics, proportional hazards and proportional reversed hazards, respectively, where $F(x) = P(X \leq x)$ and $\bar{F}(x) = 1 - F(x)$.

- $w(x) = (-\ln \bar{F}(x))^n(-\ln F(x))^m$. Setting $(n = 0)$ and $(m = 0)$, one at a time, implies lower record and upper record, respectively.

2.2. β-entropy

Let X be a continuous random variable with a probability density function $f(x)$. The Shannon entropy is defined as

$$H(f) = -\int_{\chi} f(x)\log f(x)\mathrm{d}x, \tag{5}$$

where χ is the support of the random variable.

The Shannon entropy is the expected value of the function $g(f) = -\log f$, in which $g(1) = 0$ and $g(0) = \infty$. In general, we can choose any convex function $g(f)$ as a measure of information content, provided that $g(1) = 0$ (Khinchin, 1957). The expected information content is then given by

$$H_g(f) = E[g(f)] = \int_{\chi} g(f)f(x)\mathrm{d}x,$$

and we refer to this as a class of g-entropies (Ullah, 1996). A class of smooth functions that can be presented is as follows:

$$g(f) = \begin{cases} \dfrac{1}{\beta-1}(1-f^{\beta-1}), & \beta \neq 1,\ \beta > 0, \\ -\log(f), & \beta = 1, \end{cases} \tag{6}$$

where β is a non-stochastic constant. The entropy measure of this class is as follows:

$$H_\beta(f) = \begin{cases} \dfrac{1}{\beta-1}\left[1-\displaystyle\int_{\chi} f^\beta \mathrm{d}x\right], & \beta \neq 1,\ \beta > 0, \\ -E[\log(f)] = H(f), & \beta = 1, \end{cases} \tag{7}$$

which is called β-class entropy. β-entropy was originally introduced by Havrda and Charvat (1967) and later applied to physical problems by Tsallis (1988). Tsallis exploited its non-extensive features and placed it in a physical setting (hence it is also known as Tsallis entropy). It is currently fruitfully used in many statistical systems: three-dimensional fully developed hydrodynamic turbulence, two-dimensional turbulence in pure electron plasma, Hamiltonian systems with long-range interactions, granular systems, systems with strange non-chaotic attractors, peculiar velocities in galactic clusters, etc. (for more details, see Jizba and Arimitsu, 2004). Moreover, β-entropy is a one-parameter generalization of the Shannon entropy which can lead to models or statistical results that are different from those obtained by using the Shannon entropy. Bear in mind that the β-entropy is a monotonic function of the Renyi entropy (Ullah, 1996). On the other hand, Tsallis distributions (those derived from the maximization of Tsallis entropy) are of great interest in many physical systems because they can exhibit heavy tails and model power-law phenomena. Indeed, power laws are especially interesting since they appear widely in physics, biology, economics and many other fields. In addition, Tsallis distributions are similar to generalized Pareto distributions and appear as the limit distribution of excesses over a threshold (see Bercher, 2008). Also, note that β-entropy is non-extensive and β can be seen as measuring the degree of non-extensivity. Recently, many papers have been published about Tsallis entropy and its applications, such as Plastino and Plastino (1999), Ullah (1996), Tsallis and Brigatti (2004), Suyari (2004), Singh Pharwaha and Singh (2009) and Herrmann (2009). A review of its successful applications was presented in Tsallis (2002).

In the next section, we concentrate on the β-entropy for different versions of the Pareto-type distributions and its properties.

3. Main results

In this section, we derive the β-entropy for Pareto-type distributions and for some of their weighted versions such as probability weighted moments, order statistics, record values, proportional hazards and proportional reversed hazards.

3.1. β-entropy for Pareto-type distributions

Let X be a random variable with the probability density function (2); then we have

$$H_\beta(f) = \frac{1}{\beta-1}\left\{1-\frac{\alpha^\beta \Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta(\alpha+\gamma)-\gamma]}{(\gamma\theta)^{\beta-1}\Gamma[\beta(\alpha+1)]}\right\}, \tag{8}$$

where $H_\beta(f)$ is β-entropy obtained via the type-2 beta integral. Further,

$$\lim_{\beta\to 1} H_\beta(f) = \ln\left(\frac{\gamma\theta}{\alpha}\right) + (\gamma-1)\left[\psi(1)-\psi(\alpha)\right] + \frac{\alpha+1}{\alpha} = H(f), \tag{9}$$

Table 1
β-entropy for some particular values of the parameters for Pareto distributions.

Distribution	Density function	H_β
$Pa\left(\mu,\alpha,\gamma,\frac{1}{\gamma}\right)$	$\frac{\alpha[\gamma(x-\mu)]^{\frac{1}{\gamma}-1}}{\{1+[\gamma(x-\mu)]^{\frac{1}{\gamma}}\}^{\alpha+1}}$	$\frac{1}{\beta-1}\left\{1-\alpha^\beta\frac{\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta\alpha+\beta\gamma-\gamma]}{\Gamma[\beta(\alpha+1)]}\right\}$
$Pa(0,\alpha,1,\lambda^\alpha;\alpha\to\infty)$ (Weibull)	$\frac{1}{\gamma}x^{\frac{1}{\gamma}-1}e^{-x^{\frac{1}{\gamma}}}$	$\frac{1}{\beta-1}\left\{1-\frac{\lambda^{1-\beta}[n\binom{n-1}{i-1}]^\beta\Gamma[\beta(n-i+1)]\Gamma[\beta(i-1)+1]}{\Gamma[\beta n+1]}\right\}$
$Pa(0,\alpha,1,\theta)$ (Lomax)	$\frac{\alpha\theta^\alpha}{(x+\theta)^{(\alpha+1)}}$	$\frac{1}{\beta-1}\left(1-\frac{\alpha^\beta\theta^{1-\beta}}{\beta\alpha+\beta-1}\right)$
$Pa(0,1,1,\theta)$ (Log-logistic)	$\frac{\theta}{(x+\theta)^2}$	$\frac{1}{\beta-1}\left(1-\frac{\theta^{1-\beta}}{2\beta-1}\right)$
$P(\mu,1,1,1)$	$\frac{1}{(1+x-\mu)^2}, x>\mu$	$\frac{2}{2\beta-1}$
$Pa(0,\alpha,1,\lambda\alpha;\alpha\to\infty)$ (Exponential)	$\frac{1}{\lambda}e^{\frac{-x}{\lambda}}$	$\frac{\beta-\lambda^{1-\beta}}{\beta(\beta-1)}$

where ψ is the digamma function. This result is the same as that achieved by Yari and Mohtashami Borzadaran (in press) from the viewpoint of the Shannon entropy.

- For Pareto (III) and Pareto (II) distributions, we have

$$H_\beta(f)=\frac{1}{\beta-1}\left\{1-\frac{\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta(1+\gamma)-\gamma]}{(\gamma\theta)^{\beta-1}\Gamma[2\beta]}\right\}, \tag{10}$$

$$H_\beta(f)=\frac{1}{\beta-1}\left\{1-\frac{\alpha^\beta\Gamma[\beta(\alpha+1)-1]}{\theta^{\beta-1}\Gamma[\beta(\alpha+1)]}\right\}, \tag{11}$$

as their β-entropy, respectively.

Since β-entropy expressions of the Pareto family are not dependent on μ, we have

$$H_\beta(Pareto(I))=H_\beta(Pareto(II)). \tag{12}$$

- Note that $\lim_{\beta\to1}H_\beta(f)$ for Pareto (III), Pareto (II) and Pareto (I) distributions is the same as achieved by Yari and Mohtashami Borzadaran (in press) for the Shannon entropy.
- By replacing γ with $\frac{1}{\gamma}$ in (8), we have the β-entropy of the Burr (XII) distribution as follows:

$$H_\beta(f)=\frac{1}{\beta-1}\left\{1-\frac{\theta}{\gamma}\left(\frac{\alpha\gamma}{\theta}\right)^\beta\frac{\Gamma\left[\frac{\beta(\gamma-1)+1}{\gamma}\right]\Gamma\left[\frac{\beta(\alpha\gamma+1)-1}{\gamma}\right]}{\Gamma[\beta(\alpha+1)]}\right\}, \tag{13}$$

for which the Shannon entropy is obtained as

$$\lim_{\beta\to1}H_\beta(f)=\ln\left(\frac{\theta}{\alpha\gamma}\right)+(\gamma-1)\left[\frac{\psi(\alpha)-\psi(1)}{\gamma}\right]+\left(\frac{\alpha+1}{\alpha}\right)=H(f). \tag{14}$$

The β-entropy values for particular values of the parameters for some versions of the Pareto-type distributions have been derived and are summarized in Table 1.

3.2. β-entropy for a weighted version of the Pareto-type distributions

Let X be a random variable with the probability density function (2) and

$$w(x)=x^kF^i(x)\bar{F}^j(x), \tag{15}$$

as a weight function; then we have

$$f^w(x)=\frac{x^kF^i(x)\bar{F}^j(x)f(x)}{E[X^kF^i(X)\bar{F}^j(X)]}. \tag{16}$$

By taking $1+(\frac{x-\mu}{\theta})^{\frac{1}{\gamma}}=t$ in (2), we have

$$E[X^kF^i(X)\bar{F}^j(X)]=\alpha\mu^k\int_1^\infty\left[1+\frac{\theta}{\mu}(t-1)^\gamma\right]^k[1-t^{-\alpha}]^it^{-\alpha j-\alpha-1}dt. \tag{17}$$

As mentioned in Nadarajah and Kotz (2007), we know that

$$(1-x)^p=\sum_{m=0}^{\infty}(-1)^mA(p,m)x^m \tag{18}$$

and

$$(1+x)^p = \sum_{m=0}^{\infty} A(p,m)x^m, \tag{19}$$

where

$$A(p,m) = \frac{\Gamma(p+1)}{\Gamma(p+1-m)\Gamma(m+1)}. \tag{20}$$

Hence,

$$(1-t^{-\alpha})^i = \sum_{q=0}^{\infty}(-1)^q A(i,q)t^{-\alpha q}. \tag{21}$$

By substituting (21) in (17), we obtain

$$E[X^k F^i(X)\bar{F}^j(X)] = \alpha\mu^k \sum_{m=0}^{k}\sum_{q=0}^{\infty}(-1)^q\left(\frac{\theta}{\mu}\right)^m A(k,m)A(i,q)\frac{\Gamma(\gamma m+1)\Gamma(\alpha j+\alpha+\alpha q-\gamma m)}{\Gamma(\alpha j+\alpha q+\alpha+1)}. \tag{22}$$

Via a similar process, when $k\beta$ is integer we derive

$$\int_{\mu}^{\infty} x^{k\beta}[F(x)]^{i\beta}[\bar{F}(x)]^{j\beta}[f(x)]^{\beta}(x)\mathrm{d}x = \gamma\theta\left[\frac{\alpha\mu^k}{\gamma\theta}\right]^{\beta}\sum_{m=0}^{k\beta}\sum_{q=0}^{\infty}(-1)^q\left(\frac{\theta}{\mu}\right)^m A(k\beta,m)A(i\beta,q) \times \frac{\Gamma[\beta(1-\gamma)+\gamma m+\gamma]\Gamma[\beta(\gamma+\gamma j+\alpha)-\gamma-\gamma m+\alpha q]}{\Gamma[\beta+\beta\alpha j+\beta\alpha+\alpha q]}. \tag{23}$$

Hence,

$$H_\beta(f^w) = \frac{1}{\beta-1} \times \left\{1 - \frac{\sum_{m=0}^{k\beta}\sum_{q=0}^{\infty}\left(\frac{\theta}{\mu}\right)^m(-1)^q\frac{A(k\beta,m)A(i\beta,q)\Gamma[\beta(1-\gamma)+\gamma m+\gamma]\Gamma[\beta(\gamma+\alpha j+\alpha)-\gamma-\gamma m+\alpha q]}{\Gamma[\beta(\alpha j+\alpha+1)+\alpha q]}}{(\gamma\theta)^{\beta-1}\left[\sum_{m=0}^{k}\sum_{q=0}^{\infty}\left(\frac{\theta}{\mu}\right)^m(-1)^q\frac{A(k,m)A(i,q)\Gamma(\gamma m+1)\Gamma(\alpha j+\alpha+\alpha q-\gamma m)}{\Gamma(\alpha j+\alpha+\alpha q+1)}\right]^{\beta}}\right\}. \tag{24}$$

- For Pareto (III) and Pareto (II) distribution, $H_\beta(f^w)$ is obtained via (24) by setting $\alpha = 1$ and $\gamma = 1$, respectively.
- Let X be a random variable with the probability density function (4); we have

$$E[X^k F^i(X)\bar{F}^j(X)] = \alpha\theta^k\sum_{m=0}^{\infty}(-1)^m A(i,m)\int_1^{\infty}(t-1)^{\frac{k}{\gamma}}t^{-\alpha j-\alpha-1-\alpha m}\mathrm{d}t$$
$$= \alpha\theta^k\sum_{m=0}^{\infty}(-1)^m A(i,m)\frac{\Gamma\left(\frac{k}{\gamma}+1\right)\Gamma\left(\alpha j+\alpha+\alpha m-\frac{k}{\gamma}\right)}{\Gamma(\alpha j+\alpha+\alpha m+1)}. \tag{25}$$

Hence,

$$H_\beta(f^w) = \frac{1}{\beta-1} \times \left\{1-\left(\frac{\gamma}{\theta}\right)^{\beta-1}\frac{\sum_{m=0}^{\infty}(-1)^m\frac{A(i\beta,m)\Gamma\left(\frac{k\beta}{\gamma}+\beta-\frac{\beta}{\gamma}+\frac{1}{\gamma}\right)\Gamma\left(\alpha\beta j+\alpha\beta+\alpha m-\frac{k\beta}{\gamma}+\frac{\beta}{\gamma}-\frac{1}{\gamma}\right)}{\Gamma(\alpha\beta j+\alpha\beta+\alpha m+\beta)}}{\left[\sum_{m=0}^{\infty}(-1)^m\frac{A(i,m)\Gamma\left(\frac{k}{\gamma}+1\right)\Gamma\left(\alpha j+\alpha+\alpha m-\frac{k}{\gamma}\right)}{\Gamma(\alpha j+\alpha+\alpha m+1)}\right]^{\beta}}\right\}. \tag{26}$$

- Let X be a random variable with the probability density function (2) and

$$w(x) = [-\ln F(x)]^m[-\ln\bar{F}(x)]^n; \tag{27}$$

then we have

$$H_\beta(f^w) = \frac{1}{\beta-1}\left\{1-\int_{\mu}^{\infty}\frac{[-\ln F(x)]^{m\beta}[-\ln\bar{F}(x)]^{n\beta}[f^{\beta}(x)]}{E^{\beta}[(-\ln F(X))^m(-\ln\bar{F}(X))^n]}\mathrm{d}x\right\}. \tag{28}$$

The explicit form of (28) cannot be obtained, but it simplifies to the β-entropy for upper record and lower record of the Pareto-type distributions when $m=0$ and $n=0$, respectively. These are obtained in Section 3.4.

3.3. β-entropy for the order statistics of the Pareto-type distributions

Let $X_1, X_2, \ldots, X_n$ be a random sample of probability density function (2) and $X_{(1:n)} \le X_{(2:n)} \le \cdots \le X_{(n:n)}$ denote the corresponding order statistics; then

$$g_{i:n}(x) = \frac{n\alpha}{\gamma\theta}\binom{n-1}{i-1}\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}-1}\left[1+\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{-\alpha(n-i+1)-1}\left[1-\left[1+\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{-\alpha}\right]^{i-1}, \quad x>\mu, \tag{29}$$

where $g_{i:n}(x)$ is the probability density function of $X_{(i:n)}$. The β-entropy for $g_{i:n}(x)$ is obtained via (24) by replacing i, j and k with $i-1$, 0 and $n-i$, respectively, and is of the form:

$$\begin{aligned} H_\beta(g_{i:n}) = \frac{1}{\beta-1}\Bigg[1-\gamma\theta\left[\frac{n\alpha}{\gamma\theta}\binom{n-1}{i-1}\right]^\beta \\ \times \sum_{q=0}^{\infty}(-1)^q A[\beta(i-1),q]\frac{\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\alpha\beta(n-i+1)+\beta\gamma+\alpha q-\gamma]}{\Gamma[\alpha\beta(n-i+1)+\alpha q+\beta]}\Bigg]. \end{aligned} \tag{30}$$

Also,

$$\begin{aligned} \lim_{\beta\to 1} H_\beta(g_{i:n}) &= -\ln\left[\frac{n\alpha}{\gamma\theta}\binom{n-1}{i-1}\right]+[\alpha(n-i+1)+1]\left[\frac{\psi(n+1)-\psi(n-i+1)}{\alpha}\right] \\ &\quad + (i-1)[\psi(n-1)-\psi(i)]+(\gamma-1)n\binom{n-1}{i-1} \\ &\quad \times \sum_{q=0}^{i-1}(-1)^q\binom{i-1}{q}\left[\frac{\psi(1)-\psi[\alpha(n-i+q+1)]}{n-i+1+q}\right] = H(g_{i:n}). \end{aligned} \tag{31}$$

This result is the same as that achieved by Yari and Mohtashami Borzadaran (in press).

Note that, in the process of finding (31), we have used the following relations:

$$\sum_{m=0}^{\infty}(-1)^m A(p,m)\frac{1}{m+n-i+1} = \frac{\Gamma(p+1)\Gamma(n-i+1)}{\Gamma(n-i+p+2)} \tag{32}$$

and

$$\sum_{m=0}^{\infty}(-1)^m A(p,m)\frac{1}{(m+n-i+1)^2} = \frac{\Gamma(p+1)\Gamma(n-i+1)}{\Gamma(n-i+p+2)}[\psi(n-i+p+2)-\psi(n-i+1)]. \tag{33}$$

- For the Pareto (III) and Pareto (II) distributions, $H_\beta(g_{i:n})$ is obtained by setting ($\alpha=1$) and ($\gamma=1$) in (30), respectively. The β-entropy for the Pareto (I) and Pareto (II) distributions is the same. In addition, for the Burr (XII) distribution one can find the β-entropy via a similar process.

3.4. β-entropy for record values of the Pareto-type distributions

Suppose $X_1, X_2, \ldots, X_n$ to be a random sample of probability density function (2); then probability density function of the upper record is

$$u(x) = \frac{\left[-\ln\left[1+\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{-\alpha}\right]^{n-1}}{\Gamma(n)}f(x), \tag{34}$$

and the probability density function of the lower record is

$$l(x) = \frac{\left[-\ln\left(1-\left[1+\left(\frac{x-\mu}{\theta}\right)^{\frac{1}{\gamma}}\right]^{-\alpha}\right)\right]^{n-1}}{\Gamma(n)}f(x). \tag{35}$$

Table 2
β-entropy for the order statistics and proportional hazards for some particular values of the parameters for Pareto distributions.

Distribution	H_β for Order statistics	H_β for proportional hazards
$Pa(\mu, 1, \gamma, \theta)$ (Pareto (III))	$\frac{1}{\beta-1}\left\{1-\frac{\gamma\theta[\frac{1}{i\gamma\theta}\binom{n}{i}]^\beta\Gamma[\beta(n-i+1+\gamma)-\gamma]\Gamma[\beta(i-\gamma)+\gamma]}{\Gamma[\beta(n+1)]}\right\}$	$\frac{1}{\beta-1}\left\{1-\frac{(j+1)^\beta\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta(j+1+\gamma)-\gamma]}{(\gamma\theta)^{\beta-1}\Gamma[\beta(j+2)]}\right\}$
$Pa(0, \alpha, 1, \theta)$ (Lomax)	$\frac{1}{\beta-1}\left\{1-\theta[\frac{\alpha}{i\theta}\binom{n}{i}]^\beta\frac{\Gamma[\beta(n+1-i)+\frac{\beta-1}{\alpha}]\Gamma[\beta(i-1)+1]}{\alpha\Gamma[\beta n+1+\frac{\beta-1}{\alpha}]}\right\}$	$\frac{1}{\beta-1}\left\{1-\frac{[\alpha(j+1)]^\beta}{\theta^{\beta-1}[\beta(\alpha j+\alpha+1)-1]}\right\}$
$Pa(0, \alpha, 1, \lambda\alpha; \alpha\to\infty)$ (Exponential)	$\frac{1}{\beta-1}\left\{1-\frac{[n\binom{n-1}{i-1}]^\beta\Gamma[\beta(n+1-i)]\Gamma[\beta(i-1)+1]}{\lambda^{\beta-1}\Gamma[\beta n+1]}\right\}$	$\frac{1}{\beta-1}\left\{1-\frac{(j+1)^{\beta-1}}{\beta\lambda^{\beta-1}}\right\}$
$Pa(0, 1, 1, \theta)$ (Log-logistic)	$\frac{1}{\beta-1}\left\{1-\frac{\theta[\frac{1}{i\theta}\binom{n}{i}]^\beta\Gamma[\beta(n-i+2)-1]\Gamma[\beta(i-1)+1]}{\Gamma[\beta(n+1)]}\right\}$	$\frac{1}{\beta-1}\left\{1-\frac{(j+1)^\beta}{\theta^{\beta-1}[\beta(j+2)-1]}\right\}$
$Pa(\mu, 1, 1, 1)$	$\frac{1}{\beta-1}\left\{1-\frac{[\frac{1}{i}\binom{n}{i}]^\beta\Gamma[\beta(n-i+2)-1]\Gamma[\beta(i-1)+1]}{\Gamma[\beta(n+1)]}\right\}$	$\frac{1}{\beta-1}\left\{1-\frac{(j+1)^\beta}{\beta(j+2)-1}\right\}$

Thus by setting ($m = 0$ and replacing n with $n-1$) and ($n = 0$ and replacing m with $m-1$) in (28), one at a time, we derive

$$H_\beta(u) = \frac{1}{\beta-1}\left\{1-\gamma\theta\left[\frac{\alpha^n}{\gamma\theta\Gamma(n)}\right]^\beta\sum_{m=0}^{\infty}(-1)^m\frac{\Gamma(\beta-\beta\gamma+\gamma)}{\Gamma(\beta-\beta\gamma+\gamma-m)\Gamma(m+1)}\frac{\Gamma(\beta(n-1)+1)}{(\beta\alpha+\beta\gamma-\gamma)^{\beta(n-1)+1}}\right\} \tag{36}$$

and

$$H_\beta(l) = \frac{1}{\beta-1}\left\{\frac{\gamma\theta}{\alpha}\left[\frac{\alpha}{\gamma\theta\Gamma(n)}\right]^\beta\sum_{m=0}^{\infty}\sum_{q=0}^{\infty}(-1)^{m+q}\frac{\Gamma(\beta-\beta\gamma+\gamma)}{\Gamma(\beta-\beta\gamma+\gamma-m)\Gamma(m+1)}\right.$$
$$\left.\times\frac{\Gamma\left(\frac{\beta\gamma}{\alpha}-\frac{\gamma}{\alpha}+\frac{m}{\alpha}+\beta\right)\Gamma(\beta(n-1)+1)}{\Gamma\left(\frac{\beta\gamma}{\alpha}-\frac{\gamma}{\alpha}+\frac{m}{\alpha}+\beta-q\right)\Gamma(q+1)(1+q)^{\beta(n-1)+1}}\right\}, \tag{37}$$

as the β-entropy for the upper and lower records of the Pareto-type distributions, respectively.

- For the Pareto (III), Pareto (II), Pareto (I) and Burr (XII) distributions, the β-entropy can easily be found for upper and lower records.

3.5. β-entropy for proportional hazards and proportional reversed hazards of Pareto-type distributions

Consider the Pareto (IV) distribution with cumulative distribution function (1); then, by setting ($i = 0, k = 0$) and ($j = 0, k = 0$), one at a time, in (26), we obtain

$$H_\beta\left[\frac{\bar{F}^j(x)f(x)}{E(\bar{F}^j(X))}\right] = \frac{1}{\beta-1}\left\{1-\gamma\theta\left[\frac{\alpha(j+1)}{\gamma\theta}\right]^\beta\frac{\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta(\alpha j+\alpha+\gamma)-\gamma]}{\Gamma[\beta(\alpha j+\alpha+1)]}\right\} \tag{38}$$

and

$$H_\beta\left[\frac{F^i(x)f(x)}{E(F^i(X))}\right]$$
$$= \frac{1}{\beta-1}\left\{1-\gamma\theta\left[\frac{\alpha(i+1)}{\gamma\theta}\right]^\beta\sum_{m=0}^{\infty}(-1)^m\frac{\Gamma(\beta i+1)\Gamma[\beta(1-\gamma)+\gamma]\Gamma[\beta(\alpha+\gamma)+\alpha m-\gamma]}{\Gamma(\beta i+1-m)\Gamma(m+1)\Gamma[\beta(\alpha+1)+\alpha m]}\right\}, \tag{39}$$

which give the β-entropy for the proportional hazards and proportional reversed hazards for the Pareto (IV) distribution, respectively.

- For the Pareto (III), Pareto (II), Pareto (I) and Burr (XII) distributions, $H_\beta[\frac{\bar{F}^j(x)f(x)}{E(\bar{F}^j(X))}]$ and $H_\beta[\frac{F^i(x)f(x)}{E(F^i(X))}]$ can be derived by setting ($\alpha = 1$), ($\gamma = 1$), ($\gamma = 1, \mu = \theta$) and ($\mu = 0, \gamma \to \frac{1}{\gamma}$) in (38) and (39) as the β-entropy for proportional hazards and proportional reversed hazards, respectively.

The β-entropy values for the order statistics and proportional hazards for some particular values of the parameters for Pareto distributions have been derived and are summarized in Table 2.

Remark 1. Based on H_β for the order statistics, presented in Table 2, the β-entropy for the first-order statistics from a sample of $Pa(\mu, \alpha, 1, \theta)$, $Pa(\mu, 1, 1, \theta)$, $Pa(\mu, 1, 1, 1)$, Lomax, exponential and log-logistic distributions has a simple form. Also, the β-entropy for the upper record for the above distributions can be shown to have a simple form.

4. Conclusions

In this paper, the β-entropy for Pareto-type distributions and their related distributions has been obtained. Also, the β-entropy for the weighted versions of these distributions and their special cases such as order statistics, proportional hazards, proportional reversed hazards, probability weighted moments, upper and lower records has been obtained. Furthermore, for some particular values of parameters, the β-entropy values for Pareto-type distributions, their order statistics and proportional hazards have been derived and summarized in Tables 1 and 2. We have shown that our results reduce to the Shannon entropy as β tends to one. Some of these Shannon entropy results were derived by Yari and Mohtashami Borzadaran (in press).

Acknowledgements

The authors are grateful to the editor and the referees for their helpful comments, which have greatly improved the paper.

References

Arnold, B.C., 1983. Pareto Distributions. International Cooperative Publishing House, Fairland, Maryland.
Asadi, M., Ebrahimi, N., Hamedani, G.G., Soofi, E.S., 2006. Information measures for Pareto distributions and order statistics. Statistics for Industry and Technology, Part II, 207–223.
Baratpour, S., Ahmadi, J., Arghami, N.R., 2007. Some characterizations based on entropy of order statistics and record values. Communications in Statistics — Theory and Methods 36 (1), 47–57.
Belzunce, F., Navarro, J., Ruiz, J.M., Aguila, Y., 2004. Some results on residual entropy function. Metrika 59, 147–161.
Bercher, J.F., 2008. Possible rationales for Renyi–Tsallis entropy maximization. In: International Workshop on Applied Probabilities, IWAP'08, electronic version.
Boltzmann, L., 1872. Neitere Studien uber das Warmegleichgewicht unter Gasmolekulen. K. Akad. Wiss. (Wein) Sitzb. 66, 275.
Clausius, R., 1864. Abhaudlungen uber die mechanische Warmetheorie Friedrich. Vieweg, Braunschweig.
Cover, T., Thomas, J., 2006. Elements of Information Theory. Wiley, New York.
Di Crescenzo, A., Longobardi, M., 2006. On weighted residual and past entropies. Scientiae Math. Japan 64, 255–266.
Dragomir, S.S., 2003. On the p-logarithmic and α-power divergence measures in information theory. Panamerican Mathematics Journal 13 (3), 1–10.
Ebrahimi, N., Soofi, E.S., Zahedi, H., 2004. Information properties of order statistics and spacing. IEEE Transactions on Information Theory 50, 177–183.
Harremoes, P., 2006. Interpretations of Renyi entropies divergences. Physica A 365, 57–62.
Havrda, J., Charvat, F., 1967. Quantification method in classification processes: concept of structural α-entropy. Kybernetika 3, 30–35.
Hegde, A., Lan, T., Erdogmus, D., 2005. Order statistics based estimator for Renyi entropy. In: Machine Learning for Signal Processing, IEEE Workshop on Machine Learning for Signal Processing, pp. 335–339.
Herrmann, K., 2009. Non-extensivity vs. informative moments for financial models — a unifying framework and empirical results. A Letters Journal Exploring the Frontiers of Physics 88 (3), 30007 (5pp.).
Jizba, P., Arimitsu, T., 2004. The world according to Renyi: thermodynamics of multifractal systems. Annals of Physics 312, 17–59.
Kapur, J.N., 1989. Maximum Entropy Models in Science and Engineering. Wiley Eastern, New Delhi.
Khinchin, A.I., 1957. Mathematical Foundation of Information Theory. Dover, New York.
Nadarajah, S., Kotz, S., 2007. Two generalized beta distributions. Applied Economics 39, 1743–1751.
Oluyede, B.O., 2006. Some general notions of stochastic orderings for weighted reliability and uncertainty measures with applications. International Mathematical Forum 1 (2), 51–60.
Park, S., 2003–2007. The entropy of consecutive order statistics. IEEE Transactions on Information Theory 41.
Patil, G.P., 2002. Weighted distribution. Encyclopedia of Environ. metrics 4, 2369–2377.
Plastino, A., Plastino, A.R., 1999. Tsallis entropy and Jaynes' information theory formalism. Brazilian Journal of Physics 29 (1), 50–60.
Rao, C.R., 1965. On discrete distributions arising out of methods of ascertainment. Sankhya: The Indian Journal of Statistics, Series A 27 (2–4), 311–324.
Renyi, A., 1961. On Measures of Entropy and Information. Univ. California Press, Berkeley, Calif, pp. 547–561.
Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423.
Singh Pharwaha, A.P., Singh, B., 2009. Shannon and non-Shannon measures of entropy for statistical texture feature extraction in digitized Mammograms. In: Proceedings of the World Congress on Engineering and Computer Science, vol. II.
Suyari, H., 2004. Generalization of Shannon–Khinchin axioms to non-extensive systems and the uniqueness theorem for the non-extensive entropy. IEEE Transactions on Information Theory 50, 1783–1787.
Tsallis, C., 1988. Possible generalization of Boltzmann–Gibbs statistics. Journal of Statistical Physics 52, 479–487.
Tsallis, C., 2002. Entropic nonextensivity: a possible measure of complexity. Chaos, Fractals and Solitons 13, 371–391.
Tsallis, C., Brigatti, E., 2004. Non-extensive statistical mechanics: a brief introduction. Continuum Mechanics and Thermodynamics 16, 223–235.
Ullah, A., 1996. Entropy, divergence and distance measures with econometric applications. Journal of Statistical Planning and Inference 49, 137–162.
Weiner, N., 1949. Cybernetics. Wiley, New York.
Yari, G.H., Mohtashami Borzadaran, G.R., 2010. Entropy for Pareto-types and its order statistics distributions. Communications in Information and Systems (in press).