Some Properties of Entropy for the Exponentiated Pareto Distribution (EPD) Based on Order Statistics

J. Behboodian
Islamic Azad University-Shiraz Branch

S. Tahmasebi
Shiraz University

Abstract. In this paper, we derived the exact form of the entropy for Exponentiated Pareto Distribution (EPD). Some properties of the entropy and mutual information are presented for order statistics of EPD. Also, the bounds are computed for the entropies of the sample minimum and maximum for EPD.

AMS Subject Classification: 94A17; 60E05
Keywords and Phrases: Differential entropy, entropy bounds, exponentiated Pareto Distribution, order Statistics, mutual information.

1. Introduction

We will first introduce the concept of differential entropy which is the entropy of a continuous random variable. Let $X$ be a random variable with cumulative distribution function $F_X(x) = P(X \leq x)$ and density
The differential entropy $H(X)$ of a continuous random variable $X$ with a density $f_X(x)$ is defined as

$$H(X) = -\int_{-\infty}^{\infty} f_X(x) \log f_X(x) \, dx = -\int_{0}^{1} \log f_X(F_X^{-1}(u)) \, du,$$

where $u = F_X(x)$. Now, let us consider the exponentiated Pareto distribution (EPD) with probability density function (pdf)

$$f_X(x) = \theta \lambda \left[ 1 - (x + 1)^{-\lambda} \right]^{\theta-1} (x + 1)^{-(\lambda+1)}, \quad x > 0, \quad \lambda > 0, \quad \theta > 0,$$

and cumulative distribution function (cdf)

$$F_X(x) = \left[ 1 - (x + 1)^{-\lambda} \right]^\theta, \quad x > 0, \quad \lambda > 0, \quad \theta > 0,$$

where $\theta$ and $\lambda$ are two shape parameters. When $\theta = 1$, the above distribution corresponds to the standard Pareto distribution of the second kind.

Analytical expression for the entropy of univariate continuous distributions are discussed by Cover and Thomas [3], Lazo and Rathie [8], Nadarajah and Zagrafos [9]. Also, the information properties of order statistics have been studied by a few authors. Among them Wong and Chen [15], Park [10], Ebrahimi et al. [5] provided several results and some characterizations of shannon entropy for order statistics.

The rest of this paper is organized as follows. In Section 2, we derived the exact form of the entropy for exponentiated pareto distribution.
SOME PROPERTIES OF ENTROPY FOR THE EXPONENTIATED...

(EPD). In Section 3, we present shannon entropy for jth order statistic of EPD, some properties of the entropy, and mutual information for order statistics of EPD.

2. Entropy for EPD

Suppose $X$ is a random variable with $EP(\theta, \lambda)$ and density function (2). Now using (1), the log-density of (2) is

$$\log f_X(x) = \log(\theta \lambda) + (\theta - 1) \log \left(1 - (x + 1)^{-\lambda}\right) - (\lambda + 1) \log(x + 1), \quad (4)$$

and the entropy is

$$H(X) = E \left( -\log f_X(x) \right)$$

$$= -\int_0^\infty f_X(x) \log f_X(x) dx$$

$$= -\log(\theta \lambda) + (1 - \theta) E \left[ \log \left(1 - (x + 1)^{-\lambda}\right) \right]$$

$$+ (\lambda + 1) E \left[ \log(x + 1) \right]. \quad (5)$$

So, we need to find $E \left[ \log \left(1 - (x + 1)^{-\lambda}\right) \right]$ and $E \left[ \log(x + 1) \right]$ to obtain the Shannon entropy.

Derivation of these two expectations are based on the following strategy:

$$k(r) = E \left[ (x + 1)^r \right] = \int_0^\infty \theta \lambda (x + 1)^{r-(\lambda + 1)} \left[1 - (x + 1)^{-\lambda}\right]^{\theta - 1} dx. \quad (6)$$
By the change of variable \([1 - (x + 1)^{-\lambda}] = t, 0 < t < 1\), we obtain:

\[
k(r) = E[(x + 1)^r] = \int_0^1 \theta(t)^{\theta-1} (1 - t)^{-\frac{r}{\lambda}} \, dx = \frac{\Gamma(\theta) \Gamma(1 - \frac{r}{\lambda})}{\Gamma(\theta + 1 - \frac{r}{\lambda})}, \quad 1 - \frac{r}{\lambda} \neq 0, -1, -2, \ldots.
\]

Differentiating both sides of \((7)\) with respect to \(r\) we obtain:

\[
k'(r) = E[(x + 1)^r \log(x + 1)] = \frac{\theta \Gamma(\theta) \left[ -\frac{1}{\lambda} \Gamma'(1 - \frac{r}{\lambda}) \Gamma(\theta + 1 - \frac{r}{\lambda}) + \frac{1}{\lambda} \Gamma(1 - \frac{r}{\lambda}) \Gamma'(\theta + 1 - \frac{r}{\lambda}) \right]}{\left( \Gamma(\theta + 1 - \frac{r}{\lambda}) \right)^2}.
\]

From relation \((8)\), at \(r = 0\) we obtain

\[
E[\log (x + 1)] = \frac{1}{\lambda} \left[ \psi(\theta + 1) - \psi(1) \right],
\]

where \(\psi\) is the digamma function defined by \(\psi(\theta) = \frac{d}{d\theta} \ln \Gamma(\theta)\).

Now we calculate

\[
t(r) = E\left[ \left(1 - (x + 1)^{-\lambda}\right)^r \right] = \int_0^\infty \theta \lambda \left(1 - (x + 1)^{-\lambda}\right)^{r+\theta-1} (x + 1)^{-(\lambda+1)} \, dx = \frac{\theta}{\theta + r}
\]

\[
\left. \frac{dt(r)}{dr} \right|_{r=0} = E\left[ \log \left(1 - (x + 1)^{-\lambda}\right) \right] = -\frac{1}{\theta}.
\]

Putting \((9)\) and \((11)\) in relation \((5)\) we have:

\[
H(X) = -\log (\lambda \theta) + \frac{\lambda + 1}{\lambda} \left[ \psi(\theta + 1) - \psi(1) \right] + \frac{1 + \theta}{\theta},
\]

where \(-\psi(1) = 0.5772\ldots\) is the Euler constant.
3. Some Properties of Entropy Based on Order Statistics EPD

Let $X_1, \ldots, X_n$ be a random sample from a distribution $F_X(x)$ with density $f_X(x) > 0$. The order statistics of this sample is defined by the arrangement of $X_1, \ldots, X_n$ from the smallest to the largest, by $Y_1 < Y_2 < \ldots < Y_n$. The density of $Y_j$, $j = 1, \ldots, n$, is

$$f_{Y_j}(y) = \frac{n!}{(j-1)!(n-j)!} f_X(y) [F_X(y)]^{j-1} [1 - F_X(y)]^{n-j}. \quad (13)$$

Now, let $U_1, U_2, \ldots, U_n$ be a random sample from $U(0,1)$ with the order statistics $W_1 < W_2 < \ldots < W_n$. The density of $W_j$, $j = 1, \ldots, n$, is

$$f_{W_j}(w) = \frac{1}{B(j, n-j+1)} w^{j-1} [1 - w]^{n-j}, \quad 0 < w < 1, \quad (14)$$

where $B(j, n-j+1) = \frac{\Gamma(j)\Gamma(n-j+1)}{\Gamma(n+1)} = \frac{(j-1)!(n-j)!}{n!}$.

The entropy of the beta distribution is

$$H_n(W_j) = -(j-1)[\psi(j) - \psi(n+1)] - (n-j)$$

$$\quad [\psi(n+1-j) - \psi(n+1)] + \log B(j, n-j+1),$$

where $\psi(t) = \frac{d \log \Gamma(t)}{dt}$, $\psi(n+1) = \psi(n) + \frac{1}{n}$.

Using the fact that $W_j = F_X(Y_j)$ and $Y_j = F_X^{-1}(W_j)$, $j = 1, 2, \ldots, n$, are one to one transformations, the entropies of order statistics can be computed by
\[ H(Y_j) = H_n(W_j) - E_{g_j}\left[\log f_X(F_X^{-1}(W_j))\right] \]  
\[ = H_n(W_j) - \int f_j(y) \log f_X(y) dy, \]  
\[ (15) \]

Now, we can have an application of (16) for the EPD. Let \( X \) be a random variable having the EPD(\( \theta, \lambda \)). For computing \( H(Y_j) \), we have

\[ F_X^{-1}(W_j) = \left[ 1 - (W_j)^{\frac{1}{\theta}} \right] \frac{1}{\lambda} - 1, \]

and the expectation term in (15) is obtained as follows:

\[ E_{g_j}\left[\log f_X(F_X^{-1}(W_j))\right] = E_{g_j}\left[\log(\theta \lambda) + \frac{\lambda + 1}{\lambda} \log(1 - W_j^{\frac{1}{\theta}}) \right] + (\theta - 1) \log((W_j)^{\frac{1}{\theta}})] \]

\[ = \log(\theta \lambda) + \frac{\lambda + 1}{\lambda} E_{g_j}\left[\log(1 - W_j^{\frac{1}{\theta}})\right] + \theta E_{g_j}\left[\log(W_j)\right] \]

\[ = \log(\theta \lambda) + \frac{\lambda + 1}{\lambda} \left[ \frac{n!}{(j - 1)!} \sum_{k=0}^{n-j} \frac{(-1)^k(\psi(1) - \psi(\theta k + j \theta + 1))}{k!(n - j - k)!(k + j)} \right] 
+ \frac{\theta - 1}{\theta} (\psi(j) - \psi(n + 1)). \]  
\[ (17) \]

Therefor, by (15) and (17) the entropy of \( j \) th order statistic is

\[ H(Y_j) = H_n(W_j) - \log(\theta \lambda) + \frac{\lambda + 1}{\lambda} \]

\[ \left[ \frac{n!}{(j - 1)!} \sum_{k=0}^{n-j} \frac{(-1)^k(-\psi(1) + \psi(\theta k + j \theta + 1))}{k!(n - j - k)!(k + j)} \right] 
+ \frac{1 - \theta}{\theta} (\psi(j) - \psi(n + 1)). \]  
\[ (18) \]
For the sample minimum \( j = 1 \), \( H_n(W_1) = 1 - \log n - \frac{1}{n} \) and

\[
H(Y_1) = 1 - \log n - \frac{1}{n} - \log(\theta \lambda) + \\
\frac{\lambda + 1}{\lambda} \times \left[ \sum_{u=1}^{n} (-1)^{u-1} \binom{n}{u} \left( \psi(u \theta + 1) + \gamma \right) \right] + \frac{\theta - 1}{\theta} \left( \psi(n) + \frac{1}{n} + \gamma \right),
\]

where \( \gamma = -\psi(1) = 0.5772... \) is the Euler constant.

The distribution function of \( Y_n \) is \( F_n(y) = \left[ 1 - (y + 1)^{-\lambda} \right]^{n\theta} I(0, \infty)(y) \) and the density is \( f_n(y) = n \theta \lambda \left[ 1 - (y + 1)^{-\lambda} \right]^{n\theta-1} (y+1)^{-(\lambda+1)} I(0, \infty)(y) \).

Noting that \( H_n(W_n) = 1 - \log n - \frac{1}{n} \), the formula (18) gives

\[
H(Y_n) = 1 - \log n - \frac{1}{n} - \log(\theta \lambda) + \\
\frac{\lambda + 1}{\lambda} \left[ \psi(n \theta + 1) + \gamma \right] + \frac{\theta - 1}{\theta} \left( \frac{1}{n} \right).
\]

For any random variable \( X \) with \( H(X) < \infty \), Ebrahimi et al.\cite{10} showed that the entropy of order statistics \( Y_j \), \( j = 1, 2, ..., n \), is bounded as follows:

\[
H(Y_j) \geq H_n(W_j) - \log M,
\]

and

\[
H(Y_j) \leq H_n(W_j) - \log M + n B_j(H(X) + \log M),
\]

where \( M \) is the mode of the distribution and \( B_j \) denotes the \( j \) th term of binomial probability \( Bin(n - 1, \frac{j - 1}{n - 1}) \). Therefore, we can compute
the bounds for the entropies of the sample minimum and maximum for EPD with parameters $\lambda$ and $\theta$. We have $M = \left( \frac{\lambda \theta + 1}{\lambda + 1} \right)^{\frac{1}{\lambda}} - 1$. So,

$$1 - \log n - \frac{1}{n} - \log \left( \left( \frac{\lambda \theta + 1}{\lambda + 1} \right)^{\frac{1}{\lambda}} - 1 \right) \leq H(Y_1),$$

(23)

and

$$H(Y_n) \leq 1 - \log n - \frac{1}{n} + (n - 1) \log \left( \left( \frac{\lambda \theta + 1}{\lambda + 1} \right)^{\frac{1}{\lambda}} - 1 \right)$$

$$+ n \left( -\log (\lambda \theta) + \frac{\lambda + 1}{\lambda} [\psi(\theta + 1) - \psi(1)] + \frac{1 + \theta}{\theta} \right).$$

(24)

The lower bound and upper bound for the entropies of the sample minimum and maximum for EPD are useful when $n$ is small.

Information theory provides some concepts of extensive use in statistics, one of which is mutual information of two random variables. It is a generalization of the coefficient of determination, for a bivariate random vector $(X, Y)$ with joint density function $f(x, y)$ and marginal density functions, $f_Y(y)$ and $f_X(x)$. The mutual information is defined as

$$I(X, Y) = \int_S f(x, y) \log \frac{f(x, y)}{f_Y(y)f_X(x)} \, dx \, dy$$

$$= H(X) + H(Y) - H(X, Y),$$

(25)

where $S$ is the region $f(x, y) \geq 0$ and $H(X, Y)$ is the entropy of $(X, Y)$. Mutual information for order statistics have an important role in statis-
SOME PROPERTIES OF ENTROPY FOR THE EXPONENTIATED...

tical sciences. In view of Ebrahimi et al. [10], the degree of dependency among \(Y_1, ..., Y_n\) is measured by the mutual information between consecutive order statistics, defined by

\[
I_n(Y_j, Y_{j+1}) = -\log \left( \binom{n}{j} \right) + n\psi(n) - j\psi(j) - (n - j)\psi(n - j) - 1. \tag{26}
\]

For given \(n\), \(I_n(Y_j, Y_{j+1})\) is symmetric in \(j\) and \(n - j\); increases in \(j\) for \(j < \frac{n}{2}\), and decreases for \(j > \frac{n}{2}\). \(I_n(Y_j, Y_{j+1})\) is increasing in \(n\). Thus, \(I_n(Y_j, Y_{j+1})\) is maximum at the median and is symmetric about the median. Now, suppose \(Y_1, ..., Y_n\) denote the order statistics of a random sample \(X_1, ..., X_n\) from EPD, then we can calculate the mutual information between \(Y_1\) and \(Y_n\). Thus, we have

\[
I(Y_1, Y_n) = H(Y_1) + H(Y_n) - H(Y_1, Y_n) \\
= -\log n(n - 1) - \frac{(n - 2)}{n(n - 1)^2} + 4\left(1 - \frac{1}{n}\right) - 2 - 2\log(\theta\lambda) \\
+ \frac{\lambda + 1}{\lambda} \left[ \sum_{u=1}^{n} (-1)^{u-1} \binom{n}{u} (\psi(\theta u + 1) + \gamma) + \psi(n\theta + 1) + \gamma \right] \\
+ \frac{\theta - 1}{\theta} \left( \psi(n) + \frac{2}{n} + \gamma \right). \tag{27}
\]

Noting that \(H(Y_1, Y_n)\) can be computed by

\[
\int_{0}^{\infty} \int_{0}^{z} -f_{Y_1, Y_n}(y, z) \log f_{Y_1, Y_n}(y, z) \, dy \, dz.
\]

Conclusion

We have derived the exact form of shannon entropy for the Exponentiated Pareto Distribution (EPD) and its order statistics. This distribu-
tion is applied in reliability, actuarial sciences, economics, and telecommunications. We have also presented some properties of the entropy and mutual information for order statistics of EPD.

Acknowledgements

We would like to thank Executive Editor and two referees for their careful reading of our manuscript and helpful suggestions which improved the paper.

References


**Javad Behboodian**  
Department of Mathematics  
Islamic Azad University - Shiraz Branch  
Shiraz, Iran.  
E-mail: behboodian@susc.ac.ir

**Saeid Tahmasebi**  
Department of Mathematics and Statistics  
Shiraz University  
Shiraz 71454, Iran  
E-mail: Stahmasby@yahoo.com