Applied Mathematical Sciences, Vol. 7, 2013, no. 65, 3247 - 3253
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2013.3296

Shannon Entropy for the Generalized Feller-Pareto (GFP) Family and Order Statistics of GFP Subfamilies

Mahmoud Riad Mahmoud [1] and Amani S. Abd El Ghafour [2]

[1] Department of mathematical statistics, I.S.S.R., Cairo University
5 Dr.Ahmed Zweil St. Orman, Giza, Egypt, Zip code 12613

Abstract

In this paper, we derive the exact analytical expressions of entropy for the Generalized Feller-Pareto (GFP) and order statistics of GFP subfamilies. The GFP family traces its roots back to Zandonatii 2001 as mentioned in Keleiber and Kotz [2]. But it is defined and investigated by Paranaíba et al. [4] under different name which is called Beta Burr XII (BBXII) distribution. The GFP family is a very general distribution which includes some known distributions such as FP and beta Weibull distributions. Also, The GFP family includes some new distributions such as beta-Lomax and generalized Burr XII distributions. These distributions are applied in reliability, medicine, actuarial science, and economics.

Keywords: entropy, generalized Feller-Pareto, beta Burr XII, generalized Burr XII, beta Lomax, order statistics.

1. Introduction

The concept of entropy originated in the nineteenth century by Shannon [1]. The Shannon entropy for a continuous random variable X with probability density function $f(x)$ is defined as

$$H(x) = E\left(-lnf(x)\right) = -\int_{-\infty}^{\infty} f(x)\, lnf(x)\, dx \quad (1)$$

[2]e-mail: inamaamani@yahoo.com

This is a mathematical measure of information which measures the average reduction of uncertainty of X. The aim of this paper is to determine the exact form of the Shannon information for generalized Feller-Pareto family and order statistics of GFP subfamilies. The GFP family traces its roots back to Zandonatii 2001 as mentioned in Keleiber and Kotz [2].

Let a random variable V distributed $Beta(q,p)$. For some $a,\theta,b > 0$, use $x = b(v^{-\frac{1}{\theta}} - 1)^{1/a}$ when make transformation to obtain

$$f(x) = \frac{a\theta x^{a-1}}{b^a B(p,q)}\left(1+\left(\frac{x}{b}\right)^a\right)^{-\theta q-1}\left(1-\left(1+\left(\frac{x}{b}\right)^a\right)^{-\theta}\right)^{p-1} \quad ,x>0 \qquad (2)$$

Distribution (2) was introduced as defined in Keleiber and Kotz [2] and X is said to have a generalized Feller-Pareto distribution (GFP) and can be denoted by $GFP(a,\theta,b,p,q)$. The four parameters a,θ,p,q are shape parameters and b is a scale parameter. Some known distributions can be special case from GFP distribution (2). For example, Feller-Pareto distribution is obtained for $\theta = 1$ and a Singh–Maddala (or Burr) distribution, which is obtained for $p = 1$.

A new distribution includes in GFP family is the generalized Burr XII distribution, which is obtained as

$$f(x) = \frac{a\theta p x^{a-1}}{b^a}\left(1+\left(\frac{x}{b}\right)^a\right)^{-\theta-1}\left(1-\left(1+\left(\frac{x}{b}\right)^a\right)^{-\theta}\right)^{p-1} \quad ,x>0 \qquad (3)$$

Distribution (3) was defined by Zandonatti 2001 as mentioned in Keleiber and Kotz [2] and is obtained by putting $q = 1$ in (2). Another new distribution can be obtained from GFP family is the beta Lomax distribution, which has density function defined by

$$f(x) = \frac{\theta}{bB(p,q)}\left(1+\left(\frac{x}{b}\right)\right)^{-\theta q-1}\left(1-\left(1+\left(\frac{x}{b}\right)\right)^{-\theta}\right)^{p-1} \quad ,x>0 \qquad (4)$$

Distribution (4) is obtained by putting $a = 1$ in (2) and can be defined by using a generalized class of distributions defined by Eugene et al. [3].

This paper is organized as follows. In section 2, the exact form of the Shannon information for generalized Feller-Pareto family and related distributions are obtained. In section 3, the Shannon information for order statistics of GFP subfamilies are also obtained.

2. Shannon Entropy for the Generalized Feller-Pareto Distribution

Now, suppose X is a random variable with $GFP(a,\theta,b,p,q)$ distribution with density function of GFP given in (2). The Shannon entropy (1) for a continuous random variable X with probability density function (2) will be

$$H(x) = -\ln(a\theta) + \ln\left(bB(p,q)\right) - (a-1)E\left(ln\left(\frac{x}{b}\right)\right) + (\theta q+1)E\left(ln\left(1+\left(\frac{x}{b}\right)^{a}\right)\right) - (p-1)E\left(ln\left(1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}\right)\right) \quad (5)$$

To determine an expression (5) for $H(x)$, we need to find $E\left(ln\left(\frac{x}{b}\right)\right)$, $E\left(ln\left(1+\left(\frac{x}{b}\right)^{a}\right)\right)$ and $E\left(ln\left(1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}\right)\right)$. Derivations of these expressions are based on three steps. To obtain $E\left(ln\left(\frac{x}{b}\right)\right)$, First computing

$$E\left(\frac{x}{b}\right)^{k} = \int_{0}^{\infty} \frac{a\theta(\frac{x}{b})^{a+k-1}}{bB(p,q)}\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta q-1}\left(1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}\right)^{p-1} dx$$

$$= \frac{\theta}{B(p,q)}\int_{0}^{1} y^{q\theta-\frac{k}{a}-1}\left(1-y^{\theta}\right)^{p-1}(1-y)^{\frac{k}{a}}\, dy \quad ; y = \left(1+\left(\frac{x}{b}\right)^{a}\right)^{-1}$$

$$= \frac{\theta\, \Gamma(p)}{B(p,q)}\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}}{\alpha!\ \Gamma(p-\alpha)}B(\theta q+\theta\alpha-\frac{k}{a}, 1+\frac{k}{a}) \quad ; \theta q+\theta\alpha-\frac{k}{a} > 0$$

Then, differentiating both sides with respect to k

$$\frac{d}{dk}E\left(\frac{x}{b}\right)^{k} = E\left(\left(\frac{x}{b}\right)^{k}\ln\left(\frac{x}{b}\right)\right)$$

$$= \frac{\theta\, \Gamma(p)}{aB(p,q)}\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\, \Gamma\left(1+\frac{k}{a}\right)\Gamma\left(\theta q+\theta\alpha-\frac{k}{a}\right)}{\alpha!\ \Gamma(p-\alpha)\Gamma(\theta q+\theta\alpha+1)}[\Psi\left(1+\frac{k}{a}\right) - \Psi\left(\theta q+\theta\alpha-\frac{k}{a}\right)]$$

Finally, put $k = 0$ in both sides

$$E\left(\ln\left(\frac{x}{b}\right)\right) = \frac{\Gamma(p)}{aB(p,q)}\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}(\Psi(1)-\Psi(\theta q+\theta\alpha))}{\alpha!\ \Gamma(p-\alpha)(q+\alpha)} \quad (6)$$

where $\Psi(\,)$ is the digamma function defined by $\Psi(\lambda) = \frac{d}{d\lambda}ln\Gamma(\lambda)$.

Now, by the same way we can compute

$$E\left(1+\left(\frac{x}{b}\right)^{a}\right)^{k} = \int_{0}^{\infty}\frac{a\theta x^{a-1}}{b^{a}B(p,q)}\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta q+k-1}\left(1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}\right)^{p-1} dx$$

$$= \frac{1}{B(p,q)}\int_{0}^{1} v^{q-\frac{k}{a}-1}\,(1-v)^{p-1}\, dv \quad ; v = \left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}$$

$$= \frac{1}{B(p,q)}B\left(p, q-\frac{k}{\theta}\right)$$

Then, differentiate both sides with respect to k

$$\frac{d}{dk}E\left(\left(1+\left(\frac{x}{b}\right)^{a}\right)^{k}\right) = \frac{1}{\theta B(p,q)}B\left(p,q-\frac{k}{\theta}\right)\left(\Psi\left(p+q-\frac{k}{\theta}\right)-\Psi\left(q-\frac{k}{\theta}\right)\right)$$

and put $k = 0$ in both sides to obtain

$$E\left(ln\left(1+\left(\frac{x}{b}\right)^{a}\right)\right)=\frac{1}{\theta}\left(\Psi(p+q)-\Psi(q)\right) \quad (7)$$

Finally, using the same strategy enables us to obtain

$$E\left(ln\left(1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}\right)\right)=\Psi(p)-\Psi(q+p) \quad (8)$$

and substituting (6), (7) and (8) into (5) to get the exact analytic expression for the Shannon entropy for GFP family

$$H(x)=\ln\left(\frac{bB(p,q)}{a\theta}\right)-(a-1)\frac{\Gamma(p)}{aB(p,q)}\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\left(\Psi(1)-\Psi(\theta q+\theta\alpha)\right)}{\alpha!\ (q+\alpha)\Gamma(p-\alpha)}$$
$$+\left(\frac{1}{\theta}+(p+q-1)\right)\Psi(p+q)-(p-1)\Psi(p)-\left(\frac{1}{\theta}+q\right)\Psi(q) \quad (9)$$

The Shannon entropy (9) associated with $\theta = 1$ will be the Shannon entropy for Feller-Pareto distribution as in Tahmasebi and Behboodian [5]. From (9) the Shannon entropy for a Singh–Maddala (or Burr) distribution, which is obtained for $p = 1$ in (9), will be

$$H(x)=\ln\left(\frac{b}{a\theta q}\right)-\frac{(a-1)}{a}\left(\Psi(1)-\Psi(\theta q)\right)+\frac{(1+\theta q)}{\theta q}$$

Moreover, the Shannon entropy for the generalized Burr XII distribution (3) is given by putting $q = 1$ in (9) as

$$H(x)=\ln\left(\frac{b}{a\theta p}\right)-\frac{(a-1)}{a}\Gamma(p+1)\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\left(\Psi(1)-\Psi(\theta(\alpha+1)\right)}{(\alpha+1)!\,\Gamma(p-\alpha)}$$
$$+\frac{\theta p+1}{\theta p}+\left(\frac{1}{\theta}+1\right)\left(\Psi(p)-\Psi(1)\right) \quad (10)$$

Furthermore, the beta Lomax distribution (4) has the following Shannon entropy, choose $a = 1$ in (9),

$$H(x)=\ln\left(\frac{bB(p,q)}{\theta}\right)+\left(\frac{1}{\theta}+(p+q-1)\right)\Psi(p+q)-(p-1)\Psi(p)$$
$$-\left(\frac{1}{\theta}+q\right)\Psi(q)$$

Hence the Shannon entropy for Transformed beta, beta of second kind, Pareto IV, inverse burr, Fisk, Lomax and more other distributions can be obtained from (9) or (10) as special cases.

3. Shannon Entropy for Order Statistics of GFP Subfamilies

Starting with random variables $X_1, X_2, \ldots, X_m$ which are usually assumed to be independent and identically distributed with their continuous density function $f(x)$. Let $X_{(1)}$ denotes the smallest of the set $X_1, X_2, \ldots, X_n$, $X_{(2)}$denotes the second smallest, etc., and $X_{(m)}$ denotes the largest. Thus $X_{(1)}, X_{(2)}, \dots, X_{(m)}$ are called the order statistics of the sample and $X_{(i)}$, for $1 \le i \le m$, is called the i^{th} order statistics. Consider a random sample $x_1, x_2, \dots, x_n$ with probability density function $f(x)$ and cumulative distribution function $F(x)$. Then the probability density function of the i^{th} order statistic $X_{(i)}$ is given as

$$f_m\left(x_{(i)}\right) = \frac{m!}{(i-1)!(m-i)!}\left[F\left(x_{(i)}\right)\right]^{i-1} f\left(x_{(i)}\right)\left[1-F\left(x_{(i)}\right)\right]^{m-i}, -\infty < x_{(i)} < \infty \quad (11)$$

The cumulative distribution function of GFP was given by Paranaíba et al. [4] as

$$F(x) = \int_0^{1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}} \frac{1}{B(p,q)} w^{q-1}(1-w)^{p-1}\, dw = I_{1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}}[p,q] \quad (12)$$

It is based on incomplete beta function so it is complicated to compute the Shannon entropy of (11) for (2). But for special case $q = 1$ in (12), the Shannon entropy (1) for generalized Burr XII distribution (3) is easier to obtain as the following

$$H_m\left(x_{(i)}\right) = ln\left(\frac{(i-1)!\,(m-i)!}{m!}\right) - (i-1)E\left(ln\left(F\left(x_{(i)}\right)\right) + H\left(x_{(i)}\right) - (m-i)E\left(ln\left(1-F\left(x_{(i)}\right)\right)\right) \quad (13)$$

where $H\left(x_{(i)}\right)$ is the Shannon entropy for generalized Burr XII given in (10) and the cumulative distribution function (12) will be reduced to be

$$F(x) = p\int_0^{1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta}} (1-w)^{p-1}\, dw = 1-\left(1+\left(\frac{x}{b}\right)^{a}\right)^{-\theta p}$$

Thus, we can compute

$$E\left(ln\left(F\left(x_{(i)}\right)\right) = \Gamma(p+1)\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}}{(\alpha+1)!\,\Gamma(p-\alpha)}\left(\Psi(1) - \Psi\left(1+\frac{\alpha+1}{p}\right)\right)$$

and

$$E\left(ln\left(1-F\left(x_{(i)}\right)\right) = p\left(\Psi(1) - \Psi(1+p)\right)$$

Finally, the Shannon entropy (13) for generalized Burr XII distribution (3) will be

$$
\begin{aligned}
H_m(x_{(i)}) \\
&= l\,n\left(\frac{(i-1)!\,(m-i)!\,b}{m!\,a\theta p}\right) + \frac{(1+\theta p)}{\theta p} - \left(1+\frac{1}{\theta}+(m-i)p\right)\Psi(1) \\
&- (i-1)\sum_{\alpha=0}^{\infty}\frac{\Gamma(p+1)(-1)^{\alpha}}{(\alpha+1)!\,\Gamma(p-\alpha)}\left(\Psi(1)-\Psi\left(1+\frac{\alpha+1}{p}\right)\right) \\
&- \frac{(a-1)}{a}\Gamma(p+1)\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\left(\Psi(1)-\Psi(\theta(1+\alpha)\right)}{(\alpha+1)!\,\Gamma(p-\alpha)} \\
&+ \left(1+\frac{1}{\theta}+(m-i)p\right)\Psi(p) + (m-i)
\end{aligned} \tag{14}
$$

In particular cases, the Shannon entropy for the sample minimum and maximum are

$$
\begin{aligned}
H_m(x_{(1)}) \\
&= l\,n\left(\frac{b}{ma\theta p}\right) + \frac{(1+\theta p)}{\theta p} - \left(1+\frac{1}{\theta}+(m-1)p\right)\Psi(1) \\
&- \frac{(a-1)}{a}\Gamma(p+1)\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\left(\Psi(1)-\Psi(\theta(1+\alpha)\right)}{(\alpha+1)!\,\Gamma(p-\alpha)} \\
&+ \left(1+\frac{1}{\theta}+(m-1)p\right)\Psi(p) + (m-1)
\end{aligned}
$$

and

$$
\begin{aligned}
H_m(x_{(i)}) \\
&= l\,n\left(\frac{b}{ma\theta p}\right) + \frac{(1+\theta p)}{\theta p} + \left(1+\frac{1}{\theta}\right)(\Psi(p)-\Psi(1)) \\
&- (m-1)\sum_{\alpha=0}^{\infty}\frac{\Gamma(p+1)(-1)^{\alpha}}{(\alpha+1)!\,\Gamma(p-\alpha)}\left(\Psi(1)-\Psi\left(1+\frac{\alpha+1}{p}\right)\right) \\
&- \frac{(a-1)}{a}\Gamma(p+1)\sum_{\alpha=0}^{\infty}\frac{(-1)^{\alpha}\left(\Psi(1)-\Psi(\theta(1+\alpha)\right)}{(\alpha+1)!\,\Gamma(p-\alpha)}
\end{aligned}
$$

Furthermore, the Shannon entropy of order statistics subfamilies in Tahmasebi and Behboodian [5] is obtained, (we omit the additional location parameters), by putting $p = 1$ in (14).

Conclusion

We have derived the exact form of Shannon entropy for the generalized Feller-Pareto (GFP) family and order statistics of GFP subfamily which is generalized Burr XII distribution. These families cover a wide spectrum of areas such as reliability, medicine, actuarial science, and economics. So our results in this paper will be important as a reference for scientists and engineers from many areas.

References

[1] C.E. Shannon, A Mathematical Theory of Communication, Bell System Technical Journal, 27 (1948), 379-432.

[2] C. Kleiber and S. Kotz, Statistical Size Distributions in Economics and Actuarial Sciences, John Wiley, Hoboken, NJ, 2003.

[3] N. Eugene, C. Lee and F. Famoye, Beta-Normal Distribution and Its Applications. Communication in Statistics-Theory and Methods, 31 (2002), 497-512.

[4] P.F. Paranaiba, E.M.M. Ortega, G.M. Cordeiro and R.R. Pescim, The Beta Burr XII Distribution with Application to Lifetime Data, Computational Statistics and Data Analysis,55 (2011), 1118-1136.

[5] S. Tahmasebi and J. Behboodian, Shannon Entropy for the Feller-Pareto (FP) Family and Order Statistics of FP Subfamilies. Applied Mathematical Sciences, 4(2010), 10, 495-504.

Received: February 19. 2013