ITM Web of Conferences **59**, 02009 (2024) https://doi.org/10.1051/itmconf/20245902009
HMMOCS-II 2023

Information entropy of the generalized beta distribution

Vitaly G Polosin*

Medical Cybernetics and Computer Science Department, Medical Institute of Penza State University, 40 Krasnaya Street, Penza, 440026, Russian Federation

Abstract. The output distribution of statistical variables characterizes the level of uncertainty of a complex system, the state of which is given by probabilistic changes in internal relationships. Information entropy is an effective tool for analysing the distribution of data in complex systems. Subfamilies of generalized beta distributions of the first and second types includes many simple distributions as special cases. These subfamilies are often used to model the output state of the system by distributions of random variables. There are conditions in the paper, that are imposed on the parameters of the generalized beta distribution to obtain relationships with simpler types. The paper contains an expression for calculating the information entropy, the generalized beta distribution, which includes distributions of the first and second kind as special cases. Expressions of information entropies for the most famous subfamilies of generalized beta distribution also are given.

1 Introduction

Probability distributions serve as the basis for the analysis of experimental data and arrays of values at the output of the investigated and controlled objects. The experiment results depend on the degree to which the assumed distribution model matches the probabilistic data set. In this case, both the distribution model and data sets should have similar probabilistic characteristics, such as mean, variance, skewness, and kurtosis of data. When modelling probable processes, beta distributions of the first and second types are often used. These are very flexible distributions with many independent shapes that are characterized by the following features. The beta distribution of the first type is specified for the X random variable in the interval from 0 to 1, while the beta distribution of the second type is defined on the entire positive axis of values for a random variable from 0 to ∞.

Often generalized distributions are used to modelling, that are include the beta distributions of the first and second types as special cases. Since generalized distributions of the first and second types do not include each other as a special case, when modelling these distributions are considered as separate independent families of distributions. Despite the fact that the distributions are built on the basis of the beta function, any comparison of these families does not take place, since they include various subfamilies of distributions. McDonald J.B. was summarized in paper [1] the details of many of the relationships between

* Corresponding author: polosin-vitalij@yandex.ru

different distribution subfamilies. In a later work [2] given a generalized unbiased beta distribution that contains each of the beta distributions of the first and second types as special cases.

This is to combine independent distributions the c parameter was used to select distribution types. A simplified formula of combining standardized beta distributions of the first and second types is

$$f_B(x;c,u,v) = \frac{x^{u-1}}{B(u,v)}(1-(1-c)x)^{v-1}(1+cx)^{-(u+v)}. \tag{1}$$

Where u and v are shape parameters of the beta function.

If the c parameter is equal to 0, then the random variable X is set in the interval from 0 to 1. If the c parameter is equal to 1, then the random variable X is set in the interval from 0 to ∞. The distribution density (1) includes beta distributions of the first and second types with the values of the c parameter equal to 0 or 1, respectively. The generalized beta distribution can be obtained by assuming that the random variables X and Y are related by the relation that it is given a

$$x = \left(\frac{y-\zeta}{\vartheta}\right)^a. \tag{2}$$

Where ζ is the displacement parameter, ϑ is the scale parameter, a is the power parameter.

Then the density function for the generalized beta distribution was obtained after substituting relation (2) into expression (1). This is given as

$$f_{GB}(y;\zeta,\vartheta,a,c,u,v) = \frac{1}{B(u,v)}\frac{a}{\vartheta}\left(\frac{y-\zeta}{\vartheta}\right)^{au-1}\cdot \left(1-(1-c)\left(\frac{y-\zeta}{\vartheta}\right)^a\right)^{v-1}\left(1+c\left(\frac{y-\zeta}{\vartheta}\right)^a\right)^{-(u+v)}. \tag{3}$$

Where a, c, u and v are the parameters of the distribution shape. A y random variable is specified on the interval

$$\zeta < y < \frac{\zeta+\vartheta}{1-c}. \tag{4}$$

It follows from Ex. (4) if the value of the c parameter is equal to zero that the random variable Y is specified in the interval from ζ to $\zeta + \vartheta$, and if the value of the c parameter is equal to one that the random variable Y is specified in the interval from ζ to infinity. If the values of the c parameter are different from 0 and 1, then the density function of (3) is a multiplicative mixture from generalized beta equation partitions of the first and second types.

The density function of (3) is the continuous probability density of generalized beta distribution of the Y random variable, that it is including more than thirty known distributions as limiting or special cases. In addition, the exponential generalized beta distribution follows directly from generalized beta. It includes other common distributions.

2 Subfamilies of generalized beta distribution

The shape variety of the generalized beta distribution of (3) is determined by the presence of six parameters, of which two parameters are for the bias and the scale of the distribution, other four parameters determine the distribution shape. The density function for the generalized beta distribution of (3) gives a description for the most complete beta-distributions family. This includes both non-shifted and shifted distributions. If the c parameter is equal null that expression of (3) defines a subfamily of the generalized beta distribution of the first type, that includes such distributions as beta of the first kind, standard beta, arcsine, Kumaraswamy, inverse Beta, Pearson XII and etc.

If the c parameter is one that expression (3) description a subfamily of generalized beta prime distributions, that known as the Feller-Pareto family also. This family is defined by five parameters. The Feller-Pareto family includes distribution with four parametric such as a Pareto IV distributions and the transformed beta family [3]. The subfamily of the Pareto IV includes the two-parameter distribution Pareto I and both three parametric distributions Pareto II and Pareto III. The paper [4] contains the hierarchy of the Pareto distribution. The classification of the distributions is given in the form of Table 1 for the family of generalized beta distribution. There subfamilies with four or more parameters are shown in bold.

Since two-parameter distributions is included in different subfamilies of three-parameter distributions, these distributions are given separately with indicate of the most common subfamily. There are also given special cases that can be used to calculate two-parameter distributions from subfamilies of three-parameter distributions. There are one-parameter distributions with a known shape, that is implemented on the basis of a generalized beta distribution.

Table 1. Generalized beta distribution.

Subfamily name	**Conditions imposed on the parameter**
Generalized beta of the first kind	$c = 0$
Beta of the first kind	$c = 0, a = 1$
Standard beta	$c = 0, a = 1, \vartheta = 1, \zeta = 0$
Kumaraswamy	$c = 0, \zeta = 0, \vartheta = 1, u = 1$
Sine wave distribution for $\zeta<y<(\zeta+\vartheta)$	$c = 0, a = 1, u = 0.5, v = 0.5$
Sine wave distribution for $(-\zeta<y<\zeta)$	$c = 0, a = 1, u = 0.5, v = 0.5, \vartheta = 2\zeta$
Inverse Beta of the first kind	$c = 0, a = -1$
Generalized beta of the second kind	$c = 1$
Feller-Pareto family	$c = 1, a = \gamma^{-1}$
Pareto IV	$c = 1, u = 1, a = \gamma^{-1}$
Pareto I	$c = 1, u = 1, a = 1, \zeta = \vartheta$
Pareto II	$c = 1, u = 1, a = 1,$
Pareto III	$c = 1, u = 1, v = 1, a = \gamma^{-1}$
Transformed beta family (TB)	$c = 1, \zeta = 0$
Generralised Pareto	$c = 1, \zeta = 0, a = 1$
Beta of the second kind	$c = 1, \zeta = 0, \vartheta = 1, a = 1$
Loglogistic	$c = 1, \zeta = 0, u = 1, v = 1$
Paralogistic	$c = 1, \zeta = 0, v = a, u = 1, \vartheta = 1$
F-istributions	$c = 1, \zeta = 0, a = 1, \vartheta=\omega v^{-1}$, u=0.5ω, v=0.5$v$
Burr type XII	$c = 1, \zeta = 0, u = 1, \vartheta = 1$
Burr type III	$c = 1, \zeta = 0, v = 1, \vartheta = 1$
Lomax	$c = 1, \zeta = 0, a = 1, u = 1$
Inverse Lomax	$c = 1, \zeta = 0, a = 1, v = 1$
Half Cauchy	$c = 1, a = 2, u = v = 0.5$
Inverse Pareto II	$c = 1, \zeta = 0, a = 1, u = 1$
Half Pearsons VII for $\zeta<y$□□	$c = 1, a = 2, u = 0.5$, v = m−0.5

As specific cases, the generalized beta distribution of the second type includes the *F* distribution, half-Pearsons VII and half-Cauchy. For these distributions, Table 1 also gives the conditions for setting the parameters of the generalized beta distribution and possible options for communication with other subfamilies.

3 Information measures

More information about the shape of the distribution can be obtained from the analysis of the beta distribution information measure. Shannon's entropy for a continuous random variable X with a probability density function f(x) is defined as

$$H(Y) = -\int_{-\infty}^{\infty} f(x)\ln f(x)dx. \tag{5}$$

Shannon's entropy is a mathematical measure that estimates the reduction in the uncertainty of the resulting random variable. In modern literature, Shannon's entropy is used as an independent property of data distribution that is characterizing the information content of the obtained sample values.

Despite the huge number of publications on the study of Shannon's entropy of various distributions, the number of distributions for which analytical formulas for calculating the entropy are obtained are very limited. Formulas are the most informative if it is given. For a family of distributions. On the basis of expression (3), an analytical expression for the information entropy always the family of generalized beta distribution was obtained. The formula for calculating the entropy is given as

$$\begin{aligned} H_{GB}(\vartheta,a,c,u,v) &= \ln B(u,v) + \ln(\vartheta a^{-1}) + \\ &+ (a^{-1}-u)[\psi(u) - \psi(v+u-cu)] + \\ &+ (v+c(u+1)-1)[\psi(u+v)-\psi(v)]. \end{aligned} \tag{6}$$

It is possible obtained from expression (6) special cases of information entropies for subfamilies of distributions that it's included in the family of generalized beta distribution. The entropies of the most famous subfamilies are given in Table 2.

Since many information entropy formulas for different subfamilies are often used in modern literature, then it is many formulas for the distributions entropy that you can be found in independent sources. The famous papers [5, 6] contain expressions for the entropies a lot of subfamilies of the generalized beta distribution. There are formulas for calculating the entropies of such well-known distribution subfamilies as Kumaraswamy, Sine wave distribution for ($-\zeta<y<\zeta$), Beta of the first kind, F-distribution, Cauchy and other. For the Pareto family of distributions, entropy formulas are given in the paper [3, 6]. In papers [7, 8, 9] you can find entropy formulas for such subfamilies as Burr type XII, loglogistic and Lomax. Burr's type XII distribution is widely used in technical and physical systems in the analysis of survival as a more flexible alternative than the Weibull distribution [7]. The relationship of Burr distributions and various other distributions such as Lomax, Weibull family, lognormal family, normal distribution, logistic distribution, bell and J-shaped beta distributions were illustrated by P. R. Tadikamalla in [10]. There, in particular, it is noted that the Burr distributions of type III and type XII can be used to fit almost any unimodal data and are comparable to the systems of Pearson and Johnson distributions.

The universality of expression (6), it illustrated the possibility of obtaining formulas for a number of subfamilies of distributions that is used in the known studies of various authors. Particular cases of formulas for determining the entropy of subfamilies of the generalized beta distribution in Table 2 are given.

The formulas were obtained by substitution on the expression (6) the conditions from Table 1, that were used to distinguish subfamilies from the generalized beta distribution of function (3).

If a distribution contains one or two shape parameters, then such distributions can be obtained by simplifying various subfamilies of distributions that contain three or four parameters. For such subfamilies, the distributions of the differential entropy equation are given in Table 3.

Table 2. Information entropies of subfamilies generalized beta distribution.

Distributions	**Differential entropy equation**
Subfamily for generalized beta distribution of the first type (c=0)	
Generalized beta of the first kind	$H_{GB1}(\vartheta,a,u,v)=\ln(\vartheta a^{-1}B(u,v))+(a^{-1}-u)\cdot$ $\cdot[\psi(u)-\psi(v+u)]+(v-1)[\psi(u+v)-\psi(v)]$
Beta of the first kind	$H_{BI}(u,v)=\ln B(u,v)-(u-1)\psi(u)-$ $-(v-1)\psi(v)+(u+v-2)\psi(u+v)$
Kymarawamy	$H_{Kw}(a,v)=-\ln(av)+\left(1-v^{-1}\right)+\left(1-a^{-1}\right)+[v^{-1}+\psi(v)-\psi(1)]$
Sine wave distribution	$H_{\sin}(\vartheta)=\ln(0.25\pi\vartheta)$
Subfamily for generalized beta distribution of the second type (c=1)	
Generalized beta of the second kind	$H_{GBII}(\vartheta,a,u,v)=\ln(\vartheta a^{-1}B(u,v))+(a^{-1}-u)\cdot$ $\cdot[\psi(u)-\psi(v)]+(v+u)[\psi(u+v)-\psi(v)]$
Beta of the second kind	$H_{BII}(u,v)=\ln B(u,v)+(1-u)\psi(u)-(1+v)\psi(v)+(v+u)\psi(u+v)$
Feller-Pareto family	$H_{FP}(\vartheta,\gamma,u,v)=\ln(\vartheta\gamma B(u,v))+(\gamma-u)[\psi(u)-\psi(v)]+$ $+(v+u)[\psi(u+v)-\psi(v)]$
Pareto IV	$H_{Pareto\mathrm{IV}}(\vartheta,\gamma,v)=\ln(\vartheta\gamma v^{-1})+(\gamma-1)[\psi(1)-\psi(v)]+1+v^{-1}$
Transformed beta	$H_{TB}(\vartheta,a,u,v)=\ln(\vartheta a^{-1}B(u,v))+(a^{-1}-u)\cdot$ $\cdot[\psi(u)-\psi(v)]+(v+u)[\psi(u+v)-\psi(v)]$
Generralised Pareto	$H_{GenP}(\vartheta,u,v)=\ln(\vartheta B(u,v))+(1-u)\psi(u)-$ $-(1+v)\psi(v)+(v+u)\psi(u+v)$

Currently, the concept of "entropy" is widely used in various fields of science. Information entropy is used in various scientific disciplines as a measure of disorganization of the any complexity system. The entropy estimate is considered as a measure of the uncertainty of any experiment or test, for which different outcomes are possible. Entropy is a measure of the set of those states of the system about being in which the system must forget for its development, at the same time useful information must be stored and allocated for the development of the system [11].

Entropy estimates are used to control technological processes, to analyse data in information systems, to protect information in communication systems, to solve queuing problems and in other areas. It should also be noted that the concept of combining information and probabilistic methods of data analysis is promising. This combination allows you to create effective criteria for the study of radon sources [12], for assessing the state of the cardiovascular system [13], for the study of statistical patterns [14].

In complex systems of biology, medicine, cybernetics, economics, etc., entropy is used to independently assess the uncertainty of the state of the system. Since the uncertainty of the system is associated with the distribution law of the data that is observed in the system, then for the theoretical analysis of such systems it will be useful differential entropy equations of both generalized beta distribution and subfamilies that contain only one or two shape parameters, which are contained in this paper.

Table 3. Information entropies of subfamilies that contain only one or two shape parameters.

Distributions	Differential entropy equation
Burr type XII	$H_{Burr\text{XII}}(a,v) = -\ln(va) + (a^{-1} - u)[\psi(1) - \psi(v)] + 1 + v^{-1}$
Burr type III	$H_{Bur\text{III}}(a,u) = -\ln(ua) + (a^{-1} - u)[\psi(u) - \psi(1)] + 1 + u^{-1}$
Pareto I & Pareto II	$H_{Pareto\text{I}}(\vartheta,v) = H_{Pareto\text{II}}(\vartheta,v) = \ln\vartheta - \ln v + 1 + v^{-1}$
Pareto III	$H_{Pareto\text{III}}(\vartheta,\gamma) = \ln(\vartheta\gamma) + 2$
Loglogistic	$H_{Loglog}(\vartheta,a) = \ln(\vartheta) - \ln(a) + 2$
Paralogistic	$H_{Paralog}(a) = -2\ln(a) + (a^{-1} - 1)[\psi(1) - \psi(a)] + 1 + a^{-1}$
F-distribution	$H_F(\omega,\nu) = \ln\left(\frac{\omega}{\nu}B\left(\frac{\omega}{2},\frac{\nu}{2}\right)\right) + \left(1 - \frac{\omega}{2}\right)\psi\left(\frac{\omega}{2}\right) + + \frac{\omega+\nu}{2}\psi\left(\frac{\omega+\nu}{2}\right) - \left(1 - \frac{\nu}{2}\right)\psi\left(\frac{\nu}{2}\right)$
Lomax	$H_{Lomax}(\vartheta,v) = \ln(\vartheta) - \ln(v) + 1 + v^{-1}$
Half Cauchy	$H_{halfCauchy}(\vartheta) = \ln(2\pi\vartheta)$
Cauchy	$H_{Cauchy}(\vartheta) = \ln(4\pi\vartheta)$
Inverse Lomax	$H_{InvLomax}(\vartheta,u) = \ln\vartheta - \ln u + (1-u)\psi(u) + (1+u)\psi(u+1) - 2\psi(1)$
Inverse Pareto II	$H_{InvPareto\text{II}}(\vartheta,v) = \ln\vartheta - \ln v + (v+1)[\psi(1+v) - \psi(v)]$
Half Pearsons VII for $y = [\zeta,\ \infty]$	$H_{GB\text{II}}(\vartheta,m) = \ln(0.5\vartheta B(0.5, m-0.5)) + m[\psi(m) - \psi(m-0.5)]$

4 Conclusions

Thus, the differential entropy of the generalized beta distribution is proportional to the logarithm of the distribution scale parameter. It is possible to use it to analyse the shape of data distributions of complex systems, you should use data transformation according to expression (2), that allows you to normalize the distribution sample with respect to the scale parameter. The application of the data normalization procedure allows the use of standardized expressions for the informational entropy of beta distribution when assessing the uncertainty of the state of complex systems.

References

1. J.B. McDonald, Econometric **52(3)**, 647-63 (1984)
2. J.B. McDonald, Y.J. Xu, Journal of Econometrics **66(1-2)**, 133-135 (1995)
3. S. Tahmasebi1, J. Applied Mathematical Sciences **4(10)**, 495-504 **(**2010)
4. B.C. Arnold, *Pareto Distributions* (International Cooperative Publishing House, Silver Spring, Maryland, 1983)
5. J.V. Michalowicz, J.M. Nichols, F. Bucholtz, *Handbook of differential entropy, CRC Press* (Taylor & Francis Group, Boca Raton, 2013)

6. T.M. Cover, J.A. Thomas, *Elements of information theory* (Published by John Wiley & Sons, New Jersey, 1991)
7. X. Wang, W. Gui, Mathematics **9**, 313 (2021)
8. F. Brouers, Open Journal of Statistics **5**, 730-741 (2015)
9. B. Devi, P. Kumar, K. Kour, International Journal of Statistics and Systems **12(2)**, 175-181 **(**2017)
10. P.R. Tadikamalla, Int. Stat. Rev. **48(3)**, 337-344 **(**1980)
11. A.A. Borodin, Vestnik RosNOU, Complex systems: models, analysis, management **3**, 33-36 (2010)
12. V.G. Polosin, S.V. Tertychnaya, University proceedings. Volga region. Physics and mathematics **4**, 71-78 (2008)
13. V.G. Polosin, University proceedings. Volga region. Technical science **51(3)**, 31-44 (2019)
14. V.G. Polosin, J. Phys.: Conf. Ser. **1515**, 032044 (2020)