Shape measures for the generalized beta exponential distribution

V G Polosin
Medical Cybernetics and Computer Science Department, Penza State University, 40, Krasnay Str., Penza, 440026, Russia
E-mail: polosin-vitalij@yandex.ru

Abstract. This paper contains parametric and informational shape measures for various families of the generalized beta exponential distribution since it is important to determine the distribution shape for analysing an experimental data set. A logistic parameter is used to select independent types of beta exponential distributions, that it allows to combine the distributions of different subfamilies. In this paper the use of parametric shape measures to pre-define distribution shape is discusses. In particular, the initial and standard central moments for the main types of generalized beta exponential distribution are given. In the paper it is proposes to use the entropy coefficient of unshifted distribution as an independent information measure of the shape of unshifted generalized beta exponential distributions. In order to increase the reliability of the preliminary determination of the shape of the model, expressions for the entropy coefficient of shifted families both the generalized beta exponential distributions of the first and second types, and the generalized gamma exponential distribution were obtained. For practical applied the entropy coefficients of unshifted distributions for various subfamilies of generalized beta exponential distributions can be useful.

1. Introduction
Exponential models are extensively applied in experimental research in physics, economics, biology and other areas. Different families of distributions exponentials are commonly used for data analysis when observing complex systems. Probabilistic modeling plays a critical role in modern image processing problems [1]. The task of obtaining the result of a measurement experiment of real estimates of physical quantities is also often based on the choice of the distribution density [2]. Exponential distribution is also used in medicine for modeling response probability and dosage in quantum reaction bioanalysis [3]. Many distributions have been considered as descriptive models for the distributions of income [4]. Certain cumulative exponential distribution functions were obtained in the nineteenth century by B. Gompertz [5] and P. F. Verhulst [6] when solving the equations of population growth and used to compare the known tables of changes in mortality. These models remain relevant to modern research.

Various distribution models provide the possibility of constructing models with sequential and parallel inclusion of elements. Generalized exponential distributions include such families as gamma, Weibula, lognormal distribution [7]. The versatile class of exponential families of distributions includes many well-known distributions such as Gaussian, Poisson, Dirichlet, polynomial, Gamma, Beta and many others distributions. The most general is the family of generalized beta exponential distributions, since it includes many known models as special or limiting cases.
With a wide variety of families of distributions, it becomes relevant to choose the shape of the model for the array of experimental data. In modern studies, the preliminary choice of the form of an alternative model is based on the use of probabilistic measures obtained on the basis of statistical moments of distributions. It is of interest to use information measures to increase the recognizability of models.

2. Density function of generalized beta exponential distribution
The generalized beta exponential distribution is pioneered by J. B. McDonald [8], where the $c$ logistic parameter was used to combine and to select the independent distributions types. The generalized beta exponential distribution is built on the basis of the transformation of beta distributions of the first and second types. The essence of the transformation can be formulated as follows. A random variable $Z$ corresponds to a generalized beta exponential distribution if a random variable $X$ corresponds to a beta distribution of the first or second type and if the random variables $X$ and $Z$ are related by the relation that is given by

$$Z = -\sigma \ln(X) - \delta.$$  

(1)

Where the letters of $\delta$ and $\sigma$ are location and scale parameters.

For the random variable $Z$ that is defined by relation (1), the density function of the generalized beta exponential distribution is given as

$$f_{\text{GBE}}(z; \delta, \sigma, c, u, v) = \frac{1}{B(u, v)} \frac{1}{\sigma} e^{-\frac{z-\delta}{\sigma}} \left(1 - (1-c) e^{-\frac{z-\delta}{\sigma}} \right)^{v-1} \left(1 + c e^{-\frac{z-\delta}{\sigma}} \right)^{-u-v}.$$  

(2)

Where $B(u, v)$ is the beta function, the letters of $u$ and $v$ are first and second shape parameters, respectively. This is under condition (1), a random variable $Z$ given on the interval

$$\ln(1-c)^{-1} - \sigma - \delta < z < \infty.$$  

(3)

The random variable $Z$ is set in the interval $[(-\sigma-\delta), +\infty]$, if the value of the parameter $c$ is equal to zero, and in the interval $[-\infty, +\infty]$ if the value of the parameter $c$ is equal to one. The logistic function of $(\ln(1-c)-1)$ allowed us to combine the two types of beta distribution. This makes also possible to include other types of distributions in the family of the generalized beta exponential distribution (2).

3. Parametric shape measures for the generalized beta exponential distribution
In the modern literature, probabilistic models are used to construct experimental regressions. The result of modeling the experimental data strongly depends on the flexibility of alternative models and the possibility of pre-selecting its shape. For this reason, the generalized beta exponential distribution is of interest, as this allows us to consider the continuous range of values of possible shape of the including exponential families of distributions as special cases. Usually, for the preliminary selection of the probabilistic model of the subfamily, various methods of moments are used, which are well were considered in the works [9, 10, 11, 12].

The theoretical moments of the probability distribution provide useful information about the flexibility of alternative models. Due to the fact that formula (2) unites families using a logical function, this contains properties of the generalized beta exponential for individual most complete subfamilies of distributions that we used often to analyze. For the subfamily of generalized beta exponential distribution, the expressions of the initial and standardized moments in [13, 14] is given. Formulas of the initial moments of the most significant subfamilies of distributions in table 1 are given. For these random variables, interval is given by inequality (4). In table 1, the initial and standardized moments are defined through polygammas functions of first, second, third and four order, that are denotes as $\psi(\xi)$, $\psi'(\xi)$, $\psi''(\xi)$ and $\psi'''(\xi)$, respectively.
As you can see at table 1, the standardized moments only depend on the first and second shape parameters. This allows us to use these moments as for a preliminary estimate of the shape of the generalized beta exponential distribution, and as well as for analyze the proximity of the distributions of different subfamilies.

An effective tool for analyzing distributions is the use of moment various diagrams, that were first proposed in the work [15]. Typically, moment ratio diagrams are used to preselect the distribution type. [15, 16, 17]. Standard moment diagrams allow you to investigate the closeness of shapes between different univariate distributions and create a short list of potential probabilistic models from a dataset [18]. The standardized moment of degree k is the ratio of the k-th moment about the mean to the k-th power of the standard deviation. The third and fourth standardized moment are measures of skewness and kurtosis. For this reason, the coefficients of skewness and kurtosis for main subfamilies of beta distributions in table 1 is given.

As shown in [13], in the space of standard moments, there is an overlap of the distributions included in the generalized beta exponential distribution. For this reason, these curves are only used to define the behavior of the models. Moreover, the location of the sampled moments in the moment ratio diagram can be used to pre-select alternative models.

4. Entropy coefficients for families of generalized beta exponential distribution
It was shown in [19, 20], the ratio of information and probability intervals allows one to obtain the entropy coefficients for shifted and unshifted distributions, that are depend only on the shape parameters. New shape measures can be used for preliminary analysis of distributions shape together with known probabilistic shape measures such as asymmetry and kurtosis.

It was proposed in [21] to use the ratio of the Δdf(Y) entropy uncertainty interval to the square root of the $m_2(Y)$ second initial moment of the same random value as an informational characteristic of the shape.
for unshifted distributions. For the entropy coefficients of the unshifted generalized beta exponential distributions, the system of expressions are given as

\[
\begin{align*}
K_{\text{HoEGB}}(u,v) &= \frac{B(u,v) \cdot \exp\left((u+v-1)\psi(v+u)-(v-1)\psi(v)-uv\psi(u)\right)}{\sqrt{\psi'(u)-\psi'(u+v)+(\psi(u))^2+(\psi(u+v))^2-2\psi(u)\psi(u+v)}}; \\
K_{\text{HoEGB}}(u,v) &= \frac{B(u,v) \exp\left((u+v)\psi(v+u)-v\psi(v)-u\psi(u)\right)}{\sqrt{\psi'(v)+\psi'(u)+(\psi(v))^2-2\psi(u)\psi(v)+(\psi(u))^2}}; \\
K_{\text{HoGE}}(u) &= \frac{\Gamma(u) \cdot \exp\left(-uv(u)+u\right)}{\sqrt{\psi'(u)+(\psi(u))^2}}.
\end{align*}
\]  

(4)

At system of (4) the first and second expressions correspond to the entropy coefficients of unshifted distributions of generalized beta exponential distributions of the first and second types, respectively, the third expression of system is the entropy coefficients of unshifted distributions of the generalized exponential gamma distribution.

Table 2 contains expressions for the entropy coefficients of unshifted distributions, that were obtained by reducing the expressions of system (4) for the subfamilies of the generalized beta exponential distribution. It is convenient to use simplified expressions for the entropy coefficients when you carrying out research work. As you can see, a number of special cases do not contain beta functions.

**Table 2.** Entropy coefficients of unshifted distributions for the generalized beta exponential distribution.

| Distributions                      | Entropy coefficients of unshifted distributions |
|-----------------------------------|-----------------------------------------------|
| Generalized beta exponential      |                                               |
| of the first kind                 |                                               |
| Generalized Gompertz-Verhulst     |                                               |
| Generalized exponential           |                                               |
| распределение                     |                                               |
| Exponential                       |                                               |
| Standard Beta exponential of the  |                                               |
| first kind                        |                                               |
| Nadarajah-Kotz distributions      |                                               |

\[K_{\text{HoEGB}}(u,v) = \frac{B(u,v) \cdot \exp\left((u+v-1)\psi(v+u)-(v-1)\psi(v)-uv\psi(u)\right)}{\sqrt{\psi'(u)-\psi'(u+v)+(\psi(u))^2+(\psi(u+v))^2-2\psi(u)\psi(u+v)}}; \]

\[K_{\text{HoEGB}}(u,v) = \frac{B(u,v) \exp\left((u+v)\psi(v+u)-v\psi(v)-u\psi(u)\right)}{\sqrt{\psi'(v)+\psi'(u)+(\psi(v))^2-2\psi(u)\psi(v)+(\psi(u))^2}}; \]

\[K_{\text{HoGE}}(u) = \frac{\Gamma(u) \cdot \exp\left(-uv(u)+u\right)}{\sqrt{\psi'(u)+(\psi(u))^2}}.
\]
Generalized beta exponential distribution of the second type ($c=1$)

$$K_{HnEGB}(u,v) = \frac{B(u,v)\exp\left((u+v)\psi(v+u) - v\psi(v) - u\psi(u)\right)}{\sqrt{\psi'(v) + \psi'(u) + (\psi(v) - \psi(u))^2}}$$

Generalized logistic distribution type IV

$$K_{HnGLogistIV}(u,v) = \frac{B(u,v)\exp\left((u+v)\psi(v+u) - v\psi(v) - u\psi(u)\right)}{\sqrt{\psi'(v) + \psi'(u) + (\psi(v) - \psi(u))^2}}$$

Generalized logistic distribution type III

$$K_{HnGLogistIII}(u) = \frac{\exp(2u(\psi(2u) - \psi(u)))}{\sqrt{2\psi'(u)}}$$

Generalized logistic distribution type II

$$K_{HnGLogistII}(u) = \frac{1}{u} \frac{\exp(\psi(1+u) - \psi(1)+1)}{\psi'(1) + \psi'(u) + (\psi(1) - \psi(u))^2}$$

Generalized logistic distribution type I

$$K_{HnGLogistI}(v) = \frac{1}{v} \frac{\exp(\psi(v+1) - \psi(1)+1)}{\psi'(v) + \psi'(1) + (\psi(v) - \psi(1))^2}$$

Logistic distributions

Hyperbolic tangents

$$K_{HnLogist} = e^2 \frac{\sqrt{3}}{\pi} = 4.074$$

Hyperbolic secant

$$K_{HnSech} = \pi \frac{\exp(\psi(1) - \psi(0.5))}{\sqrt{2\psi'(0.5)}} = 4$$

Generalized exponential gamma distribution ($v \to \infty$)

Gamma exponential

$$K_{HnGE}(u) = \frac{\Gamma(u) \cdot \exp(-u\psi(u) + u)}{\sqrt{\psi'(u) + (\psi(u))^2}}$$

Standard Gamma exponential

$$K_{HnSGE}(u) = \frac{\Gamma(u) \cdot \exp(-u\psi(u) + u)}{\sqrt{\psi'(u) + (\psi(u))^2}}$$

Chi square exponential

$$K_{HnChSqE} = \frac{\Gamma(0.5) \cdot \exp(0.5(1+\gamma) + \ln 2)}{\sqrt{0.5 \cdot \pi^2 + (\ln 4 + \gamma)^2}} = 2.631$$

Generalized Gumbel

$$K_{HnGG}(n) = \frac{\Gamma(n) \cdot \exp(n(1 - \psi(n)))}{\sqrt{\psi'(n) + (\psi(n))^2}}$$

Minimum and maximum extreme value distributions

$$K_{HnGE} = \frac{\exp(1+\gamma)}{\sqrt{6^{-1} \pi^2 + \gamma^2}} = 3.442$$

5. Conclusions

It follows from table 2 that the entropy coefficient of unbiased distributions for generalized beta exponential distributions is completely specified by two shape parameters. Wherein for most of the subfamilies, the entropy coefficients are determined by only one of the shape parameters or by a specific numerical value. Therefore, it is possible to use the entropy coefficients of unbiased distributions as an information measure of the forms for unbiased families of a generalized beta exponential distribution. It should be noted that the entropy coefficients of equations system of (4) is subject to the condition which the position parameter of unshifted distributions is zero.
Reference

[1] Nielsen F and Nock R 2010 Entropies and cross-entropies of exponential families Proceedings of the International Conference on Image Processing ICIP 3621-2

[2] Svistunov B L, Zabkov A F and Gusyninavalue Yu S 2019 Method of selecting the density of distribution of a random value of evaluating the value of the measurement result on the basis of the principle of the maximum Measurement Monitoring Control 1(27) 95-102

[3] Prentice R L 1976 A generalization of entropy and logit methods for dose response curves Biometrics 32(4) 761-76

[4] McDonald J B 1984 Some generalized functions for the size distributions of income Econometrica 52(3) 647-63

[5] Gompertz B 1832 On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies Philosophical Transactions of the Royal Society of London 123 513-85

[6] Verhulst P F 1845 Mathematical research on the law of population growth New Memoirs of the Académie Royale des Sciences et Belles-Lettres de Bruxelles 18 14-54

[7] Gupta R D and Kundu D 2007 Generalized exponential distribution: Existing results and some recent developments Journal of Statistical Planning and Inference 137(11) 1-24

[8] McDonald J B & Xu Y J 1995 A generalization of the beta distribution with applications Journal of Econometrics 66(1-2) 133-15

[9] Prokhorov S A 2002 Modelling and analysis of random processes Laboratory practice (Samara: Samara Scientific Centre of the Russian Academy of Sciences)

[10] Vadzinsky R N 2001 Handbook of probability distributions (St. Petersburg: Nauka Publishing House of the Russian Academy of Sciences)

[11] Hall A R 2005 Generalized method of moments Advanced texts in econometrics (Oxford: University Press)

[12] Kobzar A I 2006 Applied mathematical statistics For engineers and scientists (Moskow: Publishing House of Physical, mathematical and technical literature)

[13] Kerman S C and McDonald J B 2015 Skewness-Kurtosis Bounds for the EGB1 EGB2 and Special cases Communication in Statistics-Theory and Methods 44(18) 3857-64

[14] Nadarajaha S and Kotz S (2006) The beta exponential distribution Reliability Engineering and System Safety 91 689-69

[15] Craig C C 1936 A new exposition and chart for the Pearson system of frequency curves Annals of mathematical statistics 7 16-28

[16] Johnson N L, Kotz S and Balakrishnan N 1994 Continuous Univariate Distributions (New York, NY: John Wiley & Sons) 1

[17] Meeker W Q and Escobar L A 1998 Statistical Methods for Reliability Data (New York: Wiley-Interscience)

[18] Vargo E, Pasupathy R and Leemis L 2010 Moment-Ratio Diagrams for Univariate Distributions Journal of Quality Technology 42(3) 1-10

[19] Polosin V G 2020 Topographic space for distributions of the Amoroso family J. Phys.: Conf. Ser. 1679 022010

[20] Polosin V G 2021 Informational measures of the shape of the Amoroso distributions for the research of dispersed materials Material science forum 1031 58-66

[21] Polosin V G 2020 Space of entropy-parametric characteristics for control of the non-symmetric distributions form University proceedings Volga region 3(55) 44-57