A New Class of Gamma Distribution

ABSTRACT. This paper presents a new class of probability distributions generated from the gamma distribution. For the new class proposed, we present several statistical properties, such as the risk function, the density expansions, moment generating function, characteristic function, the moments of order m, central moments of order m, the log likelihood and its partial derivatives and also entropy, kurtosis, skewness and variance. These same properties are indicated for a particular distribution within this new class that is used to illustrate the capability of the proposed new class through an application to a real data set. The data set presented in Choulakian and Stephens (2001) was used. Six models are compared and for the selection of these models were used the Akaike Information Criterion (AIC), the Akaike Information Criterion corrected (AICc), Bayesian Information Criterion (BIC), Hannan Quinn Information Criterion (HQIC) and tests of Cramer-Von Mises and Anderson-Darling to assess the models fit. Finally, we present the conclusions from the analysis and comparison of the results obtained and the directions for future work.

Keywords: gamma distribution, probability distributions class, model fit.

Uma Nova Classe de Distribuições Gama

RESUMO. Este artigo apresenta uma nova classe de distribuição de probabilidades gerada a partir da distribuição gama. Para a classe proposta apresentamos algumas propriedades estatísticas tais como função de risco, expansões para densidade e acumulada, função geratriz de momentos, função característica, momentos ordinários, momentos centrais, medidas de curtose e assimetria, entropia de Rényi, função de log-verossimilhança e suas respectivas derivadas parciais. Algumas dessas propriedades são determinadas para uma distribuição base particular dentro desta nova classe para ilustrar a potencialidade da classe proposta através de uma aplicação a um conjunto de dados reais. O conjunto de dados apresentado em Choulakian and Stephens (2001) foi usado. Seis modelos são comparados e para a seleção destes foram usados os critérios de informação de Akaike, Akaike corrigido, Baysiano e Hannan Quinn. Os testes de Cramer-Von Mises e Anderson-Darling foram usados para avaliar o ajuste aos modelos. Finalmente, apresentamos as conclusões a partir da análise e comparação dos resultados obtidos e sugerimos trabalhos futuros.

Palavras-chave: distribuição gama, classe de distribuição de probabilidade, ajuste de modelos.

INTRODUCTION

The gamma distribution is used in a variety of applications including queue, financial and weather models. It can naturally be considered as the distribution of the waiting time between events distributed according to a Poisson process. It is a biparamétrica distribution, whose density is given by:

$$f(t) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}, t > 0,$$

where $\alpha > 0$ is a shape parameter and $\beta > 0$ is the reciprocal of a scale parameter.

Due to the importance of this distribution recently some new distributions as well as families of probability distributions based on generalizations of the gamma distribution have been proposed. The first is based on the family of exponentiated distribution defined by Mudholkar, Srivastava and Freimer (1995). Given a distribution with continuous distribution function Fd its *generalization* or *exponentiated form* $G(x)$ is obtained by $F(x) = G^{a}(x)$, with $a > 0$ (power parameter). Gupta, Gupta and Gupta (1998) proposed and studied some properties exponentiated gamma distribution.

Cordeiro, Ortega and Silva (2011) extended the exponentiated gamma distribution defining a new distribution called Exponentiated Generalized gamma Distribution with four parameters, which is capable of modeling bathtub shaped failure rate phenomena.

Zografos and Balakrishnan (2009) defined a family of probability distributions based on the integration of a gamma distribution as follows:

$$F(x) = \frac{1}{\Gamma(\alpha)} \int_0^{-ln\left(1-G(x)\right)} t^{\alpha-1} e^{-t} dt,$$

where $G(x)$ is an arbitrary distribution function. When $\alpha = n + 1$ this distribution coincides with the distribution of the *n*th highest value record (Alzaatreh, Famoye & Lee, 2014).

Alternatively, Ristic and Balakrishnan (2012) have proposed a new family of probability distributions, which is also based on the integration of the gamma distribution. They defined this new family as follows:

$$F(x) = 1 - \frac{1}{\Gamma(\alpha)} \int_0^{-ln\left(G(x)\right)} t^{\alpha-1} e^{-t} dt,$$

where $G(x)$ is an arbitrary distribution function. Similarly, when $\alpha = n + 1$ this distribution coincides with the distribution of the *n*th smallest value record (Alzaatreh et al. 2014).

Following the line of work of Zografos and Balakrishnan (2009) and Ristic and Balakrishnan (2012), our goal in this work is to propose a new family of distributions based on gamma distribution. The family of distributions proposed here is the following:

$$H_G(x) = \int_{\frac{1-G(x)}{G(x)}}^{+\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt,$$

where $G(x)$ is an arbitrary distribution function and $H_G(x)$ and has the same support as the distribution $G(x)$. We shall call this new family *Gamma*-$[(1-G)/G]$ *class*. The statistical properties of this new class, such as mean, variance, standard deviation, mean deviation, kurtosis, skewness, moment generating function, characteristic function and graphical analysis, are derived.

Then, to illustrate the applicability of the proposed new family, we consider the particular case of the distribution obtained when considering that $G(x)$ is the distribution function of an exponential random variable. By presenting mathematical structures for Gamma-$[(1-G)/G]$ class we also derived statistical properties of this new distribution and to illustrate its potentiality, an application to a set of real data is performed. For this, we used the data set presented in the work Choulakian and Stephens (2001) to see if the models are well adjusted to this data. As comparative criteria of fitness of the models, it was considered: the Akaike (AIC) (Akaike, 1972), the Akaike Fixed (AIC) (Burnham & Anderson, 2002), the Bayesian information criterion (BIC) (Schwartz, 1978), the Hannan-Quinn information criterion (HQIC) (Hannan & Quinn, 1979), and the Cramer-von Mises (Darling, 1957) and Anderson-Darling (Anderson & Darling, 1952) tests. Both hypothesis tests, Anderson-Darling and Cramér-von Mises, are discussed in detail by Chen and Balakrishnan (1995) and belong to the class of quadratic statistics based on the empirical distribution function, because they work with the squared differences between the empirical distribution and the hypothetical.

MATERIAL AND METHODS

Obtaining a class of probability distributions

The Gamma-$[(1-G)/G]$ class is defined by the cumulative distribution function (cdf) (for $x > 0$)**:**

$$H_G(x) = \int_{\frac{1-G(x)}{G(x)}}^{+\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt,$$

which is equivalent to

$$H_G(x) = 1 - \int_0^{\frac{1-G(x)}{G(x)}} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt. \quad (1)$$

If the distribution $G(x)$ has density $g(x)$ the class will have a probability density function (pdf) given by

$$h_G(x) = \frac{g(x)}{G^2(x)} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1-G(x)}{G(x)}\right)^{\alpha-1} exp\left(-\beta\left(\frac{1-G(x)}{G(x)}\right)\right). \quad (2)$$

The Equations (1) and (2) can be rewritten as a sum of *exponentiated* distributions. These distributions have been studied by some authors in recent years, see for example, Mudholkar and Srivastava (1993) for exponentiated Weibull, Gupta and Kundu (1999) for exponentiated exponential, among others.

Using the power series exponential, we rewrite (2) as

$$h_G(x) = g(x)\frac{\beta^\alpha}{\Gamma(\alpha)}\sum_{k=0}^{\infty}\frac{(-1)^k\beta^k}{k!}G^{-\alpha-k-1}(x)\big(1-G(x)\big)^{\alpha+k-1}.$$

Furthermore, as

$$\big(1-G(x)\big)^{k+\alpha-1} = \sum_{j=0}^{\infty}\binom{k+\alpha-1}{j}(-1)^j\,G^j,$$

it follows that

$$h_G(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}g(x)G^{j-\alpha-k-1}(x). \quad (3)$$

Since $H_G(x) = \int_{-\infty}^{x} h_G(t)dt$ we can rewrite the distribution function as

$$H_G(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}\int_{-\infty}^{x}g(x)G^{j-\alpha-k-1}(x)dt.$$

Therefore,

$$H_G(x) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,(j-\alpha-k)\Gamma(\alpha)}\binom{k+\alpha-1}{j}G^{j-\alpha-k}(x). \quad (4)$$

The following we presented an expansion to Gamma-$[(1-G)/G]$ class when G is discrete. If the distribution $G(x)$ is discrete, $H_G(x)$ is also discrete and we have that $P(X = x_l) = F(x_l) - F(x_{l-1})$. Therefore,

$$P(X = x_l) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,(j-\alpha-k)\Gamma(\alpha)}\binom{k+\alpha-1}{j}\left(G^{j-\alpha-k}(x_l) - G^{j-\alpha-k}(x_{l-1})\right).$$

■

In addition, we can obtain the risk function of the new Gamma-$[(1-G)/G]$ class as follows:

$$R_G(x) = \frac{\frac{g(x)}{G^2(x)}\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1-G(x)}{G(x)}\right)^{\alpha-1} exp\left(-\beta\left(\frac{1-G(x)}{G(x)}\right)\right)}{\int_0^{\frac{1-G(x)}{G(x)}}\frac{\beta^\alpha}{\Gamma(\alpha)}t^{\alpha-1}e^{-\beta t}dt}. \quad ■$$

Using the density and distribution function expansions, we can get the statistical properties of the new class, as discussed below. Equations (3) and (4) are the main results of this subsection.

Moments and moment generating function

Several of the interesting characteristics and features of a probability model can be obtained using moments such as tendency, dispersion, skewness and kurtosis. The following is the development of the expansion calculations for the moments of order m for the Gamma-$[(1-G)/G]$ class. The nth moment of a random variable having cdf (1) can be easily obtained from Equation (3). Hence, we have

$$\mu_m = \int_{-\infty}^{+\infty} x^m \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)} \binom{k+\alpha-1}{j} g(x) G^{j-\alpha-k-1}(x)dx.$$

Therefore,

$$\mu_m = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)} \binom{k+\alpha-1}{j} \tau_{m,0,j-\alpha-k-1},$$

where

$$\tau_{m,\eta,r} = E(X^m g^\eta(X) G^r(X)) = \int_{-\infty}^{+\infty} x^m g^\eta(x) G^r(x) dG(x). \qquad (5)$$

The expression (5) is important because it generalize the well-established probability weighted moments. ■

In particular, we have the following expansion of the mean for the Gamma-$[(1-G)/G]$ class

$$\mu = \mu_1 = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)} \binom{k+\alpha-1}{j} \tau_{1,0,j-\alpha-k-1}.$$

■

The following is the development of the expansion calculations for the moment generating function for the Gamma-$[(1-G)/G]$ class. We have from Equation (3),

$$M_X(t) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)} \binom{k+\alpha-1}{j} \int_{-\infty}^{+\infty} e^{tx} g(x)\, G^{j-\alpha-k-1}(x)dx.$$

Using the fact that

$$e^{tx} = \sum_{m=0}^{\infty} \frac{t^m x^m}{m!},$$

we can rewrite

$$M_X(t)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}t^m}{k!\,m!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}\int_{-\infty}^{+\infty}x^m g(x)G^{j-\alpha-k-1}(x)dx.$$

Therefore, using (5) we can express the last equation as

$$M_X(t)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}t^m}{k!\,m!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}\tau_{m,0,j-\alpha-k-1}.$$

■

Similarly, one can establish the following expansion for the characteristic function for the Gamma-$[(1-G)/G]$ class.

$$\varphi_X(t)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}i^m t^m}{k!\,m!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}\tau_{m,0,j-\alpha-k-1}.$$

Central moments and general coefficient

We'll look at the development of the expansion calculations for central moments of order m to the Gamma-$[(1-G)/G]$ class. This measure can be calculated as

$$\mu'_m=E[(X-\mu)^m]=\int_{-\infty}^{+\infty}(x-\mu)^m dF(x),$$

or equivalently

$$\mu'_m=\sum_{r=0}^{m}\binom{m}{r}(-1)^r\mu^r\mu_{m-r}.$$

Since

$$\mu_{m-r}=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\beta^{\alpha+k}}{k!\,\Gamma(\alpha)}\binom{k+\alpha-1}{j}\tau_{m-r,0,j-\alpha-k-1},$$

it follows that

$$\mu'_m=\sum_{r=0}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j+r}\beta^{\alpha+k}\mu^r}{k!\,\Gamma(\alpha)}\binom{m}{r}\binom{k+\alpha-1}{j}\tau_{m-r,0,j-\alpha-k-1}. \quad (6)$$

■

In particular, by expanding the range of variance for the Gamma-$[(1-G)/G]$ class we have:

$$\sigma^2=\mu'_2=\sum_{r=0}^{2}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j+r}\beta^{\alpha+k}\mu^r}{k!\,\Gamma(\alpha)}\binom{2}{r}\binom{k+\alpha-1}{j}\tau_{2-r,0,j-\alpha-k-1} \quad (7)$$

■

A new generalization called *general coefficient*, which extends the skewness and kurtosis, is given by

$$C_g(m) = \frac{E[(X-\mu)^m]}{\sqrt{\{E[(X-\mu)^2]\}^m}} = \frac{E[(X-\mu)^m]}{\sigma^m} = \frac{\mu'_m}{\sigma^m}. \tag{8}$$

Substituting (6) and (7) in Equation (8), we obtain

$$C_g(m) = \frac{\sum_{r=0}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j+r}\beta^{\alpha+k}\mu^r}{k!\,\Gamma(\alpha)}\binom{m}{r}\binom{k+\alpha-1}{j}\tau_{m-r,0,j-\alpha-k-1}}{\left(\sum_{r=0}^{2}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j+r}\beta^{\alpha+k}\mu^r}{k!\,\Gamma(\alpha)}\binom{2}{r}\binom{k+\alpha-1}{j}\tau_{2-r,0,j-\alpha-k-1}\right)^{\frac{m}{2}}}.$$

Note that, in particular, as $m = 3$ and $m = 4$ in $C_g(m)$ we obtain expansions to skewness and kurtosis measures, respectively.

Maximum likelihood estimation and Rényi entropy

Once met some regularity conditions, the maximum likelihood estimates (MLEs) can be obtained by equating the derivative of the log-likelihood function with respect to each parameter to zero. We determine the MLEs of the parameters of the Gamma-$[(1-G)/G]$ class from complete samples only. Let $x_1, \dots, x_n$ be a random sample of size n from the new class, where $\underline{\theta}$ is a vector of unknown parameters in the parent distribution $G(x;\underline{\theta})$. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\alpha, \beta, \underline{\theta}^{T})^{T}$ can be obtained as

$$l(\boldsymbol{\theta}) = n\,log\left(\frac{\lambda\beta^{\alpha}}{\Gamma(\alpha)}\right) + \sum_{i=1}^{n} log\left(\frac{g(x_i;\underline{\theta})}{G^2(x_i;\underline{\theta})}\right) +$$
$$+ (\alpha-1)\sum_{i=1}^{n} log\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right) - \beta\sum_{i=1}^{n}\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right). \tag{9}$$

The log-likelihood can be maximized, for example, either directly by using the SAS (ProcNLMixed) or by using the nonlinear likelihood expressions obtained by differentiating (9). The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$U_{\alpha}(\boldsymbol{\theta}) = n\,log\beta - n\,\psi(\alpha) + \sum_{i=1}^{n} log\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right),$$

$$U_{\beta}(\boldsymbol{\theta}) = \frac{n\alpha}{\beta} - \sum_{i=1}^{n}\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right),$$

and

$$U_{\underline{\theta}_j}(\boldsymbol{\theta}) = \sum_{i=1}^{n}\frac{\partial log}{\partial\theta_j}\left(\frac{g(x_i;\underline{\theta})}{G^2(x_i;\underline{\theta})}\right) + (\alpha-1)\sum_{i=1}^{n}\frac{\partial log}{\partial\theta_j}\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right) - \beta\sum_{i=1}^{n}\frac{\partial}{\partial\theta_j}\left(\frac{1-G(x_i;\underline{\theta})}{G(x_i;\underline{\theta})}\right),$$

where $\psi(\alpha) = \frac{d\Gamma(\alpha)}{d\alpha}$ is the digamma function. ■

Entropy is a measure of uncertainty in the sense that the higher the entropy value the lowest the information and the greater the uncertainty, or the greater the randomness or disorder. The following is the expansion entropy calculations for the Gamma-$[(1-G)/G]$ class, using the *Rényi entropy*, which is given by

$$L_R(\eta)=\frac{1}{1-\eta}log\left(\int_{-\infty}^{+\infty} f^{\eta}\,(x)dF(x)\right).$$

Substituting the expressions of density and cumulative distribution function given by Equations (2) and (1), respectively, we have

$$L_R(\eta)=\frac{1}{1-\eta}log\left(\int_{-\infty}^{+\infty}\left(\frac{g(x)}{G^2(x)}\frac{\beta^{\alpha}}{\Gamma(\alpha)}\left(\frac{1-G(x)}{G(x)}\right)^{\alpha-1}exp\left(-\beta\left(\frac{1-G(x)}{G(x)}\right)\right)\right)^{\eta}dx\right).$$

By expanding the exponential function in Taylor series as

$$exp\left(-\eta\,\beta\,\frac{1-G(x)}{G(x)}\right)=\sum_{k=0}^{\infty}\frac{(-1)^k\eta^k\beta^k}{k!}G^{-k}\big(1-G(x)\big)^k,$$

we have

$$h_G^{\eta}(x)=g^{\eta}(x)\frac{\beta^{\eta\alpha}}{\Gamma^{\eta}(\alpha)}\sum_{k=0}^{\infty}\frac{(-1)^k\eta^k\beta^k}{k!}G^{-\eta(\alpha+1)-k}(x)\big(1-G(x)\big)^{\eta(\alpha-1)+k}$$

Now, using the following binomial expansion

$$\big(1-G(x)\big)^{\eta(\alpha-1)+k}=\sum_{j=0}^{\infty}\binom{\eta(\alpha-1)+k}{j}(-1)^j\,G^j(x),$$

it follows that

$$h_G^{\eta}(x)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\eta^k\beta^{\eta\alpha+k}}{k!\,\Gamma^{\eta}(\alpha)}\binom{\eta(\alpha-1)+k}{j}g^{\eta}(x)\,G^{-\eta(\alpha+1)-k+j}(x).$$

Thus, an explicit expression for Rényi entropy can be write as

$$L_R(\eta)=\frac{1}{1-\eta}log\left(\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\eta^k\beta^{\eta\alpha+k}}{k!\,\Gamma^{\eta}(\alpha)}\binom{\eta(\alpha-1)+k}{j}\int_{-\infty}^{+\infty}g^{\eta}(x)\,G^{-\eta(\alpha+1)-k+j}(x)\,dx\right),$$

which, in turn, implies that (using Equation (5))

$$L_R(\eta)=\frac{1}{1-\eta}log\left(\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{k+j}\eta^k\beta^{\eta\alpha+k}}{k!\,\Gamma^{\eta}(\alpha)}\binom{\eta(\alpha-1)+k}{j}\tau_{0,\eta-1,-\eta(\alpha+1)-k+j}\right).$$

RESULTS AND DISCUSSION

Special model

In this section, we will examine a particular distribution of the Gamma-$[(1-G)/G]$ class proposed here. We will consider the particular case in which $G(x)=1-e^{-\lambda x}$, $x>0$, that is called the Gamma-$[(1-Exp)/Exp]$ distribution.

Gamma-$[(\mathbf{1}-\mathbf{Exp})/\mathbf{Exp}]$ distribution

Considering $G(x)$ the cdf of the exponential distribution with parameter λ in Equation (1), we have the Gamma-$[(1-Exp)/Exp]$ distribution:

$$H(x) = 1 - \int_0^{\frac{e^{-\lambda x}}{1-e^{-\lambda x}}} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt, x > 0.$$

Differentiating $H(x)$, we get the density function of the Gamma-$[(1-Exp)/Exp]$ distribution:

$$h(x) = \frac{\lambda\beta^\alpha}{\Gamma(\alpha)} \frac{e^{-\lambda x}}{(1-e^{-\lambda x})^2} \left(\frac{e^{-\lambda x}}{1-e^{-\lambda x}}\right)^{\alpha-1} e^{-\beta\left(\frac{e^{-\lambda x}}{1-e^{-\lambda x}}\right)}. \quad \blacksquare$$

Figure 1 show the graph of the Gamma-$[(1-Exp)/Exp]$ distribution probability density functions and cumulative distribution, for some values of the parameters.

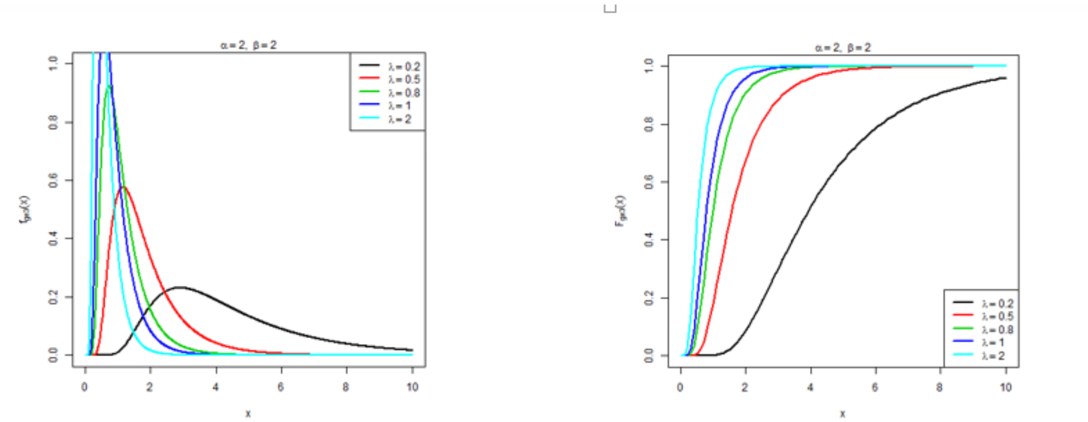

Figure 1. In right pdf and left cdf of the Gamma-$[(1-Exp)/Exp]$ distribution for some values of λ.

We can also obtain the risk function using the Gamma-$[(1-Exp)/Exp]$ distribution as follows:

$$R(x) = \frac{\frac{\lambda\beta^\alpha}{\Gamma(\alpha)} \frac{e^{-\lambda x}}{(1-e^{-\lambda x})^2} \left(\frac{e^{-\lambda x}}{1-e^{-\lambda x}}\right)^{\alpha-1} e^{-\beta\left(\frac{e^{-\lambda x}}{1-e^{-\lambda x}}\right)}}{\int_0^{\frac{e^{-\lambda x}}{1-e^{-\lambda x}}} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt}.$$

Figure 2 show the graph of the risk function using the Gamma-$[(1-Exp)/Exp]$ distribution generated from some values assigned to parameters.

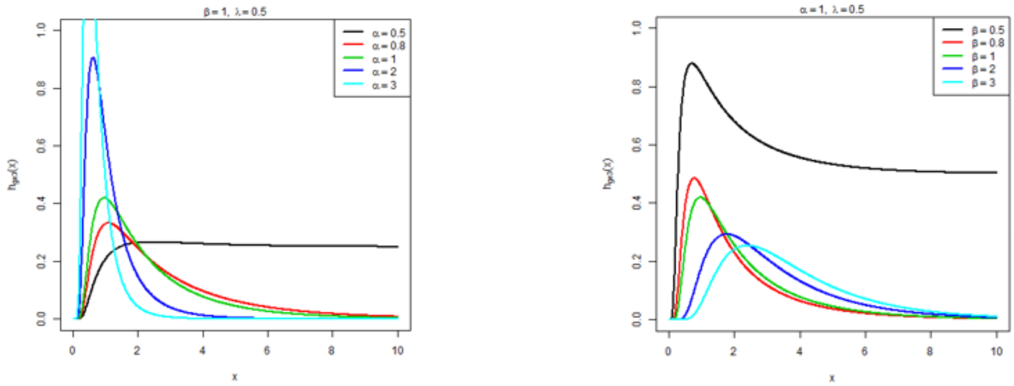

Figure 2. Plots of the risk function for some parameter values.

Using procedure similar to what was done in pdf and cdf expantions, we can rewrite the pdf and cdf of the Gamma- $[(1-Exp)/Exp]$ distribution as a sum of exponentiated exponentials, as follows:

$$h(x)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\binom{-k-\alpha-1}{j}\frac{(-1)^{k+j}\lambda\beta^{k+\alpha}}{k!\,\Gamma(\alpha)}e^{-\lambda(k+\alpha+j)x} \tag{10}$$

and

$$H(x)=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\binom{-k-\alpha-1}{j}\frac{(-1)^{k+j+1}\beta^{k+\alpha}}{k!\,(k+\alpha+j)\Gamma(\alpha)}\left(e^{-\lambda(k+\alpha+j)x}-1\right). \tag{11}$$

Various properties of the exponentiated exponential can be obtained from Gupta and Kundu (1999). Using expansions (10) and (11), we can obtain mathematical quantities of the special model such as ordinary and central moments, moment generating and characteristic functions, general coefficient, Rényi entropy and some others from quantities exponentiated exponential distribution. For example, we consider only moments for reasons of space. The *m*th ordinary moment of the special model can be expressed as

$$\mu_m=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\binom{-k-\alpha-1}{j}\frac{(-1)^{k+j}\lambda\beta^{k+\alpha}}{k!\,\Gamma(\alpha)}\frac{\Gamma(m+1)}{\left(\lambda(k+\alpha+j)\right)^{m+1}}.$$

In particular, we have that the mean of the Gamma-$[(1-Exp)/Exp]$ distribution is given by

$$\mu_1=\mu=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\binom{-k-\alpha-1}{j}\frac{(-1)^{k+j}\lambda\beta^{k+\alpha}}{k!\,\Gamma(\alpha)}\frac{1}{\left(\lambda(k+\alpha+j)\right)^{2}}.$$

Let $x_1,\dots,x_n$ be a sample of the size n from $X\sim$ Gamma-$[(1-Exp)/Exp]$ (α,β,λ). The log-likelihood function for the vector of parameters $\boldsymbol{\theta}=(\alpha,\beta,\lambda^T)^T$ can be obtained as

$$l(\boldsymbol{\theta})=n\,log\left(\frac{\lambda\beta^{\alpha}}{\Gamma(\alpha)}\right)-n\alpha\lambda-(\alpha+1)\sum_{i=1}^{n}log\left(1-e^{-\lambda x_i}\right)-\beta\sum_{i=1}^{n}\left(\frac{e^{-\lambda x_i}}{1-e^{-\lambda x_i}}\right).$$

The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$U_\alpha(\boldsymbol{\theta}) = nlog\beta - n\psi(\alpha) - n\lambda - \sum_{i=1}^{n} log\left(1 - e^{-\lambda x_i}\right),$$

$$U_\beta(\boldsymbol{\theta}) = \frac{n\alpha}{\beta} - \sum_{i=1}^{n}\left(\frac{e^{-\lambda x_i}}{1 - e^{-\lambda x_i}}\right)$$

and

$$U_\lambda(\boldsymbol{\theta}) = \frac{n}{\lambda} - n\alpha - (\alpha + 1)\sum_{i=1}^{n}\left(\frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}}\right) + \beta\sum_{i=1}^{n}\left(\frac{x_i e^{-\lambda x_i}}{(1 - e^{-\lambda x_i})^2}\right),$$

where $\psi(\alpha) = \frac{d\Gamma(\alpha)}{d\alpha}$.

Application

In this section, we will show an application to real data for the proposed gamma distribution. The data used in this research are from the excesses of flood peaks (in m ^{3} / s) Wheaton River near Carcross in the Yukon Territory, Canada. 72 exceedances of the years 1958 to 1984 were recorded, rounded to one decimal place. These data were analyzed by Choulakian and Stephens (2001), and are presented in Table 1.

Table 1. Full excess peaks in m ^{3} / s Rio Wheaton.

Excess flood peaks of Rio Wheaton (m ^{3} / s)											
1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7	1.9	13.0	12.0	9.3
1.4	18.7	8.5	25.5	11.6	14.1	22.1	1.1	2.5	14.4	1.7	37.6
0.6	2.2	39.0	0.3	15.0	11.0	7.3	22.9	1.7	0.1	1.1	0.6
9.0	1.7	7.0	20.1	0.4	2.8	14.1	9.9	10.4	10.7	30.0	3.6
5.6	30.8	13.3	4.2	25.5	3.4	11.9	21.5	27.6	36.4	2.7	64.0
1.5	2.5	27.4	1.0	27.1	20.2	16.8	5.3	9.7	27.5	2.5	27.0

It is worth mentioning that this data set has also been analyzed by means of the distributions of Pareto, Weibull three parameters, the generalized Pareto and beta - Pareto (Akinsete, Famoye & Lee, 2008).

In Table 2, we can see the maximum likelihood estimates obtained by the Newton-Raphson implemented in SAS 9.1 statistical software, parameters, standard errors, Akaike information criterion, corrected Akaike, Bayesian, Hannan-Quinn and Anderson-Darling statistics (A*) and Cramér von Mises (W*) to the Gamma-[$-ln(1-Exp)$] distributions (M1), Gamma-[$(1-Exp)/Exp$] distribution (*proposed model*, M2), exponentiated Weibull (M3), modified Weibull (M4), beta Pareto (M5) and Weibull (M6).

Table 2. Estimated maximum likelihood parameter, errors (standard errors in parentheses) and calculations of AIC statistics, AIC, BIC, HQIC, tests A* and W* for the M1 to M6 distributions.

Models	Estimates				Statistics					
					AIC	AIC	BIC	HQIC	A*	W*
M1	0.838 (0.121)	0.035 (0.007)	1.96 (<E-3)	---- ----	508.689	509.042	515.519	511.408	0.7519	0.1306
M2	0.131 (0.053)	0.179 (0.07)	0.539 (0.251)	---- ----	505.030	505.383	511.860	507.749	0.4516	0.0757
M3	1,387 (0.59)	0.519 (0.312)	0.05 (0.021)	---- ----	508.050	508.403	514.880	510.769	1.4137	0.2534
M4	0.776 (0.124)	0.124 (0.035)	0.01 (0.008)	---- ----	507.343	507.696	514.173	510.062	0.5938	0.0978
M5	84.682 (<E-3)	65.574 (<E-3)	0.063 (0.005)	0.01 (<E-3)	524.398	524.995	533.504	528.023	2.0412	0.3516
M6	0.901 (0.086)	0.086 (0.012)	---- ----	---- ----	506.997	507.171	511.551	508.810	0.7855	0.1380

For the six distributions shown in Table 2, the data applied to Wheaton River flooding, it was observed that beta-Pareto model (M5), which was described by Akinsete, Famoye and Lee (2008) as the best fitted model, in our studies had a lower performance with AIC = 524.398, AICc = 524.995, BIC=533.504, HQIC = 528.023, A = 2.0412 and W = 0.3516, when compared to the proposed Gamma-[$(1-Exp)/Exp$] model (M2) that obtained AIC = 505.030, AICc = 505.383, BIC=511.860, HQIC = 507.749, A = 0.4516 and W = 0.0757. Also according to Table 2, the proposed distribution model M2 is the best tested once the lowest values of AIC, AICc, BIC HQIC, A* and W* are from such distribution, and only according to the BIC criterion was exceeded solely by the model M6.

In the Figures 3 and 4 below, there are the graphs of density functions and distributions of M1 to M6 models fitted to the data and their corresponding histograms. The graph shows that the Gamma-[$(1-Exp)/Exp$] model has similar behavior to that of other distributions, except that of the beta Pareto which distances itself from the others.

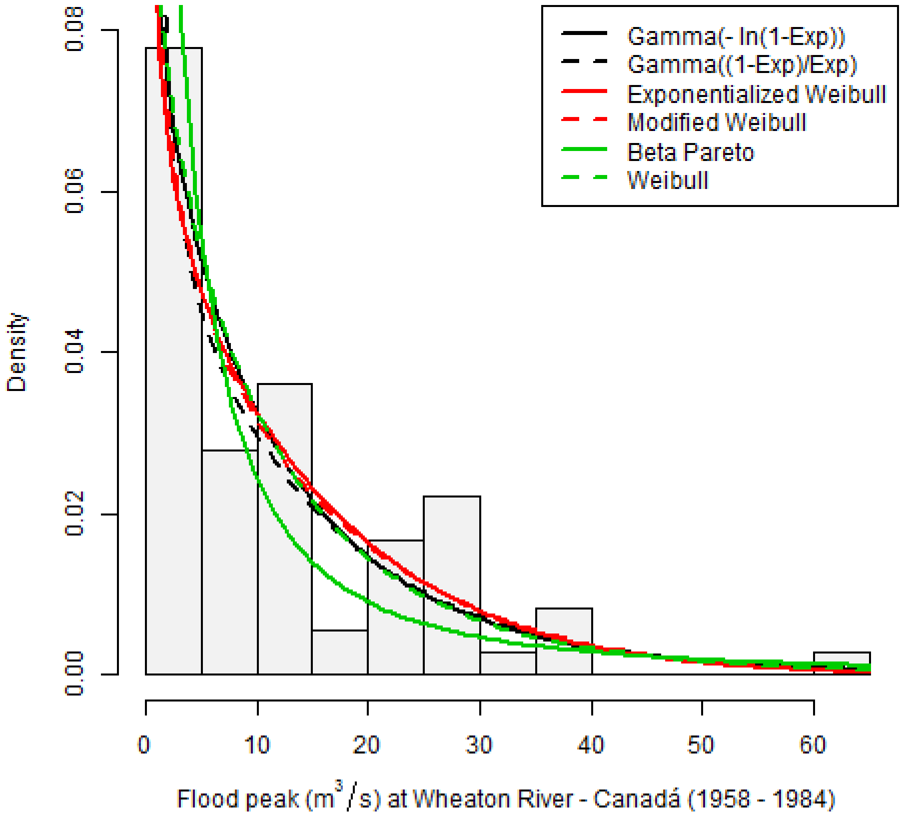

Figure 3. Fitted distributions to the mass data of flood peaks in river Wheat

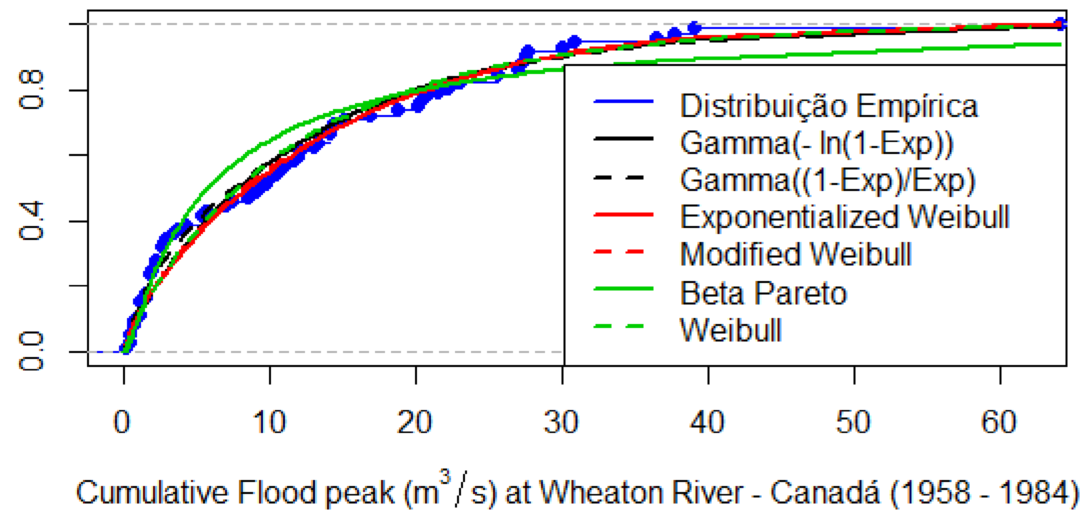

Figure 4. Fitted cdf´s to the mass data of flood peaks in river Wheaton

CONCLUSION

As concluding remarks, we note that the class of Gamma-$[(1-G)/G]$ probability distributions developed in this work is a novel way of generalizing the gamma distribution and can be applied in different areas depending on the choice of the distribution G. As future work, we intend to carry out more detailed comparisons between the novel distribution family proposed in this paper and the family of distributions investigated in Zografos and Balakrishnan (2009), which is also based on the integration of the gamma distribution.

In this work, we study in detail only a distribution of the Gamma-$[(1-G)/G]$ class, namely the Gamma-$[(1-Exp)/Exp]$ distribution. We derive some properties of this distribution and applied to a set of real data obtaining better fit than that obtained in a previous study by Akinsete *et al.* (2008). We intend to conduct the study of new distributions within this class as future work.

We note that after adding several parameters to a model it can better be adjusted to a particular phenomenon due to its greater flexibility. On the other hand, one should not forget that there may be a problem for the estimation of the parameters since it can occur both computational and identifiability problems in parameter estimation. Thus, the ideal is to choose a model that reflects well the phenomenon / experiment with the minimum number of parameters. In the case of the proposed class, only two additional parameters are added to the set of parameters of the G distribution.

REFERENCES

Akaike, H. (1972). Use of an information theoretic quantity for statistical model identification. Proc. 5th Hawaii International Conference on System Sciences. *Hawaii, Western Periodicals Co.*, 249-250.

Akinsete, A., Famoye, F., & Lee, C. (2008). The beta-Pareto distribution. *Statistics*, 42 (6), 547-563.

Alzaatreh, A., Famoye, F., & Lee, C. (2014). The Gamma-Normal distribution: Properties and Applications. *Computational Statistics and Data Analysis*, 69, 67-80.

Anderson, T. W., & Darling, D. A. (1952). Certain asymptotic theory of goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23, 193-212.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: the information-theoretic approach practical.* [Sl]: Springer.

Chen, G., & Balakrishnan, N. (1995). The general purpose approximate goodness-of-fit test. *Journal of Quality Technology,* 27, 154-161.

Choulakian, V., & Stephens, M. A. (2001). Goodness-of-fit for the generalized Pareto distribution. *Technometrics,* 43(4), 478-484.

Cordeiro, G. M., Ortega, E. M. M., & Silva, G. O. (2011). The exponentiated generalized gamma distribution with application to lifetime date. *Journal of Statistical Computation and Simulation*, 81(7), 827-842.

Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramer-von Mises Tests. *Annals of Mathematical Statistics*, 28(4), 823-838.

Gupta R. C., Gupta, P. L., & Gupta, R. D. (1998). Modeling failure time data by Lehman Alternative. *Communication in Statistics - Theory and Methods*, 27(4), 877-904.

Gupta, R. D., & Kundu, D. (1998). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, 41, 173-188.

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an Autoregression. *Journal of the Royal Statistical Society. Series B (Methodological),* JSTOR, 190-195.

Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated weibull family. *Technometrics,* 37, 436-445.

Ristic, M. M., & Balakrishnan, N. (2012). The gamma exponentiated exponential distribution, *Journal of Statistical Computation and Simulation,* 82, 1191-1206.

Zografos, K., & Balakrishnan, N. (2009). On the families of beta-and gamma-generated generalized distribution and associated inference, *Statistical Methodological,* 6, 344-362.