J. of Ramanujan Society of Math. and Math. Sc. **ISSN : 2319-1023**
Vol.6, No.1 (2017), pp. 07-24

ON THE GENERALIZED GAMMA-GENERATED DISTRIBUTIONS AND APPLICATIONS

Pushpa N. Rathie and Paulo H. D. Silva*

Department of Statistics and Applied Mathematics
Federal University of Ceara
Fortaleza, CE, Brazil.
E-mail: pushpanrathie@yahoo.com

*Department of Statistics
University of Brasilia
Brasilia, Brazil.
E-mail: paulodourado.unb@gmail.com

Dedicated to Prof. K. Srinivasa Rao on his 75^{th} ***Birth Anniversary***

Abstract: In this paper, we propose two new generalized gamma-generated distributions with any base distribution. In particular, we obtain the generalized gamma-generated exponential Weibull distribution (GG-EW) and generalized gamma-generated Dagum distribution (GG-Dagum). We study some mathematical properties of the new distributions, including explicit formulas for the probability density function, cumulative distribution function, by using Gauss' hypergeometric and Meijer G functions. It is shown, in general, that the generalized gamma-generated distributions are infinite linear combinations of the powers of the base distribution. Incorrect results given earlier by other authors are pointed out. We applied the distributions in the following data sets: (a) spending on public education in various countries in 2003 and (b) total expenditure on health in 2009 in various countries. The results show that the distributions fit well the data sets. The general R codes for fitting the distributions introduced in this paper are given in Appendix.

Keywords and Phrases: Exponentiated Weibull distribution, Dagum distribution, gamma-generated distribution, special functions.

2010 Mathematics Subject Classification: 60E05, 62B15, 33C60, 60E10.

1. Introduction

In this section, we give two known distributions which will be used later on. The exponentiated Weibull (EW) distribution and density functions for a random variable $Y > 0$ are defined by [4]

$$F_{01}(y) = \left[1 - e^{-uy^v}\right]^\alpha, \quad y, u, v, \alpha > 0, \tag{1}$$

and

$$f_{01}(y) = uv\alpha y^{v-1} e^{-uy^v} \left[1 - e^{-uy^v}\right]^{\alpha-1}, \quad y, u, v, \alpha > 0. \tag{2}$$

The Dagum distribution and density functions [2] for a random variable $Y > 0$ are given respectively by

$$F_{02}(y) = \left[1 + \lambda y^{-\alpha}\right]^{-\beta}, \quad y, \alpha, \beta, \lambda > 0, \tag{3}$$

and

$$f_{02}(y) = \alpha\beta\lambda y^{-\alpha-1} \left[1 + \lambda y^{-\alpha}\right]^{-\beta-1}, \quad y, \alpha, \beta, \lambda > 0. \tag{4}$$

The paper is divided as follows: Section 2 deals with the two definitions of the generalized gamma-generated distributions. In Theorems 2.1 and 2.4, the two distributions are expressed in terms of confluent hypergeometric function ${}_1F_1$. Theorem 2.2 gives a series expansion to be used in Theorems 2.3 and 2.5, where the generalized distributions are expressed as an infinite linear combinations of powers of base distributions. Incorrect results given by other authors are pointed out. In Section 3, GG-EW and GG-Dagum distributions are introduced which were used in Section 4 to analyse two data sets involving expenditures on education and health.

2. Generalized Gamma-generated distributions

Let a random variable X has base distribution and density functions $F(x)$ and $f(x)$ respectively. Then the generalized gamma-generated distribution and density functions for a random variable X are defined respectively by

$$H_1(x) = \frac{cb^{a/c}}{\Gamma(a/c)} \int_0^{-\ln(1-F(x))} w^{a-1} e^{-bw^c} dw, \quad a, b, c > 0, \tag{5}$$

and

$$h_1(x) = \frac{cb^{a/c}}{\Gamma(a/c)} \frac{(-\ln(1-F(x)))^{a-1} e^{-b(-\ln(1-F(x)))^c} f(x)}{1-F(x)}. \tag{6}$$

Another definition produces the following generalized gamma-generated distribution and density functions

$$H_2(x) = 1 - \frac{cb^{a/c}}{\Gamma(a/c)} \int_0^{-\ln(F(x))} w^{a-1} e^{-bw^c} dw, \quad a, b, c > 0, \tag{7}$$

and

$$h_2(x) = \frac{cb^{a/c}}{\Gamma(a/c)} \frac{(-\ln(F(x)))^{a-1} e^{-b(-\ln(F(x)))^c} f(x)}{F(x)}. \tag{8}$$

Theorem 2.1. *The first generalized gamma-generated class of distribution corresponding to base distribution $F_0(x)$ are given by*

$$H_1(x) = \frac{b^{a/c}}{\Gamma(1+a/c)} \left(-ln(1-F_0(x))\right)^a {}_1F_1\left(\frac{a}{c}; 1+\frac{a}{c}; -b\left(-ln(1-F_0(x))\right)^c\right). \tag{9}$$

The corresponding density function is

$$h_1(x) = \frac{cb^{a/c}}{\Gamma(a/c)} (-\ln(1-F_0(x)))^{a-1} e^{-b(-\ln(1-F_0(x)))^c} \frac{f_0(x)}{1-F_0(x)}. \tag{10}$$

Proof. The generalized-generated class of distributions corresponding to base distribution $F_0(x)$ are given by

$$\begin{aligned}
H_1(x) &= \frac{cb^{a/c}}{\Gamma(a/c)} \int_0^{-\ln(1-F_0(x))} u^{a-1} e^{-bu^c} du \\
&= \frac{cb^{a/c}}{\Gamma(a/c)} \int_0^{-\ln(1-F_0(x))} G_{0,1}^{1,0}\left[bu^c \middle| \begin{matrix} \\ (a-1)/c \end{matrix}\right] \\
&= \frac{cb^{a/c}}{\Gamma(a/c)} \frac{1}{2\pi i} \int_L \Gamma\left(\frac{a-1}{c} - s\right) b^s \int_0^{-\ln(1-F_0(x))} u^{cs} du ds \\
&= \frac{cb^{a/c}}{\Gamma(a/c)} \frac{1}{2\pi i} \int_L \Gamma\left(\frac{a-1}{c} - s\right) b^s \frac{[-\ln(1-F_0(x))]^{cs+1}}{c(s+1/c)} ds, \quad Re(cs+1) > 0 \\
&= \frac{b^{a/c}}{\Gamma(a/c)} \frac{[-\ln(1-F_0(x))]}{2\pi i} \int_L \Gamma\left(\frac{a-1}{c} - s\right) \frac{\Gamma(s+1/c)}{\Gamma(s+1+1/c)} (\beta[-\ln(1-F_0(x))]^c)^s ds \\
&= \frac{b^{a/c}}{\Gamma(a/c)} [-\ln(1-F_0(x))]^a G_{1,2}^{1,1}\left[\beta[-\ln(1-F_0(x))]^c \middle| \begin{matrix} 1-a/c \\ 0, -a/c \end{matrix}\right],
\end{aligned} \tag{11}$$

using equation (4) p.150 of [3], and

$$= \frac{b^{a/c}}{\Gamma(1+a/c)} [-\ln(1-F_0(x))]^a {}_1F_1\left(a/c; 1+a/c; \beta[-\ln(1-F_0(x))]^c\right), \tag{12}$$

using equation (1) p. 230 of [3].

The corresponding density function is given by

$$h_1(x) = \frac{cb^{a/c}}{\Gamma(a/c)} [-\ln(1-F_0(x))]^{a-1} e^{-b[-\ln(1-F_0(x))]^c} \frac{f_0(x)}{1-F_0(x)}. \tag{13}$$

Alternatively

$$h_1(x) = \frac{dF(x)}{dx}.$$

Theorem 2.2. *For* $0 < x < 1$

$$\left[\frac{-ln(1-x)}{x}\right]^{\delta} = \sum_{r=0}^{\infty}\left[\sum_{k=0}^{r}\frac{(-1)^{r-k}(-\delta)_{r-k}}{(r-k)!}C_{(k)(r-k)}\right]x^r,$$

where

$$C_{(0)(t)} = \left(\frac{1}{2}\right)^t,$$

and

$$C_{(m)(t)} = \frac{2}{m}\sum_{l=1}^{m}\frac{(l(t+1)-m)}{l+2}C_{(m-l)(t)}, \quad m \geq 1.$$

Proof.

Using Taylor expansion of function $\frac{-\ln(1-x)}{x}$ we have

$$S = \left[\frac{-\ln(1-x)}{x}\right]^{\delta} = \left(1 + \frac{x}{2} + \frac{x^2}{3} + \dots\right)^{\delta}$$

$$= \left(1 + x\sum_{l=0}^{\infty} b_l x^l\right)^{\delta} = (1 + xS_1)^{\delta}, \tag{14}$$

where $b_l = \frac{1}{l+2}$ and $S_1 = \sum_{l=0}^{\infty} b_l x^l$. Then

$$S = \sum_{r=0}^{\infty}\frac{(-1)^r(-\delta)_r}{r!}x^r S_1^r, \tag{15}$$

But

$$S_1^r = \sum_{k=0}^{\infty} C_{(k)(r)}x^k,$$

with

$$C_{(0)(r)} = b_0^r = \left(\frac{1}{2}\right)^r,$$

and

$$C_{(m)(r)} = \frac{2}{m}\sum_{k=1}^{m}\frac{(k(r+1)-m)}{k+2}C_{(m-k)(r)}, \quad m \geq 1,$$

Using

$$\sum_{r=0}^{\infty}\sum_{k=0}^{\infty} A(r,k)x^{r+k} = \sum_{r=0}^{\infty}\left(\sum_{k=0}^{r} A(r-k,k)\right)x^r, \tag{16}$$

we finally get

$$S = \sum_{r=0}^{\infty}\left(\sum_{k=0}^{r} \frac{(-1)^{r-k}(-\delta)_{r-k}}{(r-k)!} C_{(k)(r-k)}\right)x^r, \quad 0 < x < 1. \tag{17}$$

It may be observed that the corresponding expression used by Pinho et al. [6] is incorrect.

Theorem 2.3 *$H_1(x)$ is an infinite linear combination of powers of the base distribution $F_0(x)$.*

Proof. We can write $H_1(x)$ (9) in the following form

$$\begin{aligned}
H_1(x) &= \frac{b^{a/c}}{\Gamma(1+a/c)}\sum_{r=0}^{\infty}\frac{(a/c)_r(-b)^r}{r!(1+a/c)_r}[-\ln(1-F_0(x))]^{a+rc} \\
&= \frac{b^{a/c}}{\Gamma(1+a/c)}\sum_{r=0}^{\infty}\frac{(a/c)_r(-b)^r}{r!(1+a/c)_r}(F_0(x))^{a+cr}[\frac{-\ln(1-F_0(x))}{F_0(x)}]^{a+rc} \\
&= \frac{b^{a/c}}{\Gamma(1+a/c)}\sum_{r=0}^{\infty}\sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{(a/c)_r(-b)^r}{r!(1+a/c)_r}\frac{(-1)^{m-k}(-a-cr)_{m-k}}{(m-k)!}C_{(k)(m-k)}\times \\
&\times (F_0(x))^{a+cr+m},
\end{aligned} \tag{18}$$

where

$$C_{(0)(m)} = \left(\frac{1}{2}\right)^m,$$

and

$$C_{(n)(m)} = \frac{2}{n}\sum_{l=1}^{n}\frac{(l(m+1)-n)}{l+2}C_{(n-l)(m)}, \quad n \geq 1.$$

A special case of Theorem 2.3 was incorrectly proved by Pinho et al. [6] where they used a wrong series expansion. This was also pointed out by Castellares and Lemonte [1] who gave another series expansion different from the result given in Theorem 2.2.

Hence, the generalized gamma-generated distribution function is always an infinite linear combination of the "new" distribution function

$$G_0(x) = [F_0(x)]^{a+cr+m}.$$

Theorem 2.4. *The second generalized gamma-generated class of distribution function corresponding to base distribution $F_0(x)$ are given by*

$$H_2(x) = 1 - \frac{b^{a/c}}{\Gamma(1+a/c)} \left(-ln(F_0(x))\right)^a {}_1F_1\left(\frac{a}{c}; 1+\frac{a}{c}; -b\left(-ln(F_0(x))\right)^c\right) \tag{19}$$

The corresponding density function is

$$h_2(x) = \frac{cb^{a/c}}{\Gamma(a/c)}(-ln(F_0(x)))^{a-1}e^{-b(-ln(F_0(x)))^c}\frac{f_0(x)}{F_0(x)}. \tag{20}$$

Proof. We have

$$\begin{aligned} H_2(x) &= 1 - \frac{cb^{a/c}}{\Gamma(a/c)}\int_0^{-\ln(F_0(x))} w^{a-1}e^{-bw^c}dw \\ &= 1 - \frac{b^{a/c}}{\Gamma(1+a/c)} \left(-\ln(F_0(x))\right)^a {}_1F_1\left(\frac{a}{c}; 1+\frac{a}{c}; -b\left(-\ln(F_0(x))\right)^c\right), \end{aligned} \tag{21}$$

as before.

Theorem 2.5. *$H_2(x)$ is an infinite linear combination of powers of the base distribution $F_0(x)$.*

Proof. Using expansion of (19), we arrive at

$$H_2(x) = 1 - \frac{b^{a/c}}{\Gamma(1+a/c)}\sum_{r=0}^{\infty}\frac{(a/c)_r(-b)^r}{r!(1+a/c)_r}\left(-\ln(F_0(x))\right)^{a+cr}.$$

Now,

$$\begin{aligned} (-\ln(F_0(x)))^{a+cr} &= (1-F_0(x))^{a+cr}\left[-\frac{\ln(1-(1-F_0(x)))}{1-F_0(x)}\right]^{a+rc} \\ &= \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{(-1)^{m-k}(-a-cr)_{m-k}}{(m-k)!}C_{(k)(m-k)}(1-F_0(x))^{a+cr+m} \\ &= \sum_{m=0}^{\infty}\sum_{k=0}^{m}\sum_{s=0}^{\infty}\frac{(-1)^{m-k}(-a-cr)_{m-k}}{(m-k)!}C_{(k)(m-k)}\frac{(-a-cr-m)_s}{s!}\times \\ &\times F_0(x)^s, \end{aligned} \tag{22}$$

by using the expansion $(1+x)^\alpha = \sum_{s=0}^{\infty} \frac{(-x)^s(-\alpha)_s}{s!}$.

Oluyede et al. [5] proved incorrectly a particular case of Theorem 2.5 by using a wrong series expansion employed by Pinho et al. [6].

3. GG-EW and GG-D distributions

Using $F_{01}(x)$ of (1) and $f_{01}(x)$ of (2) in (10) for $F_0(x)$ and $f_0(x)$ respectively we define the Gamma-generated exponential Weibull density function (GG-EW) as

$$\begin{aligned} h_2(x) &= \frac{cb^{a/c}(-\ln F_{01}(x))^{a-1}e^{-b(-\ln F_{01}(x))^c}f_{01}(x)}{\Gamma(a/c)F_{01}(x)} \\ &= \frac{cb^{a/c}uv\alpha^a x^{v-1}e^{-ux^v}\left[-\ln\left(1-e^{-ux^v}\right)\right]^{a-1}e^{-b\left[-\alpha\ln\left(1-e^{-ux^v}\right)\right]^c}}{\Gamma(a/c)\left[1-e^{-ux^v}\right]}. \end{aligned} \tag{23}$$

Substituting $B = b\alpha^c$, which removes the non-identifiability of (23), $h_2(x)$ is rewritten as

$$h_2(x) = \frac{cB^{a/c}uvx^{v-1}e^{-ux^v}\left[-\ln\left(1-e^{-ux^v}\right)\right]^{a-1}e^{-B\left[-\ln\left(1-e^{-ux^v}\right)\right]^c}}{\Gamma(a/c)\left[1-e^{-ux^v}\right]}. \tag{24}$$

The model (24) is modified by replacing x by $\frac{x-\mu}{\sigma}$ where μ is a location parameter and σ a scale parameter. This modified model h_{2m} is used in the next section to analyse two data sets:

$$h_{2m}(x) = \frac{h_2\left(\frac{x-\mu}{\sigma}\right)}{\sigma}. \tag{25}$$

Using $F_{02}(x)$ of (3) and $f_{02}(x)$ of (4) in (10) for $F_0(x)$ and $f_0(x)$ respectively, the GG-Dagum density function is defined by

$$\begin{aligned} h_1(x) &= \frac{cb^{a/c}\left(-\ln(1-F_{02}(x))\right)^{a-1}e^{-b\left(-\ln(1-F_{02}(x))\right)^c}f_{02}(x)}{\Gamma(a/c)\left[1-F_{02}(x)\right]} \\ &= \frac{cb^{a/c}\left(-\ln(1-(1+\lambda x^{-\alpha})^{-\beta})\right)^{a-1}e^{-b\left(-\ln(1+\lambda x^{-\alpha})^{-\beta}\right)^c}\alpha\beta\lambda x^{-\alpha-1}(1+\lambda x^{-\alpha})^{-\beta-1}}{\Gamma(a/c)\left[1-(1+\lambda x^{-\alpha})^{-\beta}\right]} \\ &= \frac{cb^{a/c}\alpha\beta\lambda x^{-\alpha-1}(1+\lambda x^{-\alpha})^{-\beta}\left(-\ln(1-(1+\lambda x^{-\alpha})^{-\beta})\right)^{a-1}e^{-b\left(-\ln(1+\lambda x^{-\alpha})^{-\beta}\right)^c}}{\Gamma(a/c)\left[1-(1+\lambda x^{-\alpha})^{-\beta}\right]}. \end{aligned} \tag{26}$$

By replacing x by $\frac{x-\mu}{\sigma}$ in (26) where μ and σ are location and scale parameters respectively, the modified model h_{1m} is given by

$$h_{1m}(x) = \frac{h_1\left(\frac{x-\mu}{\sigma}\right)}{\sigma}. \tag{27}$$

The density functions in (25) and (27) are used in analysing two data sets in the next section.

4. Applications

In this section, we apply the two distributions introduced in this paper to two real data sets. The informations criterion AIC, BIC and AICC are given by

$$\begin{aligned} AIC &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + 2p; \\ BIC &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + p\log(n); \\ AICC &= -2\log(f(\mathbf{x}|\boldsymbol{\theta})) + 2\frac{p(p+1)}{n-p-1}, \end{aligned} \tag{28}$$

where $\log(f(\mathbf{x}|\boldsymbol{\theta}))$ is the log-likelihood function, p the number of parameters of models and n the sample size. The models that have lower AIC, BIC and AICC values are the best.

The MSE, MDA and MaxD are given by

$$\begin{aligned} MSE &= \frac{\sum_{i=1}^{n}(F_e(x_i) - \hat{F}(x_i))^2}{n} \\ MAD &= \frac{\sum_{i=1}^{n}|F_e(x_i) - \hat{F}(x_i)|}{n} \\ MaxD &= \max(|F_e(x_i) - \hat{F}(x_i)|), \ \ i = 1, \ldots, n, \end{aligned} \tag{29}$$

where $F_e(x_i)$ is the empirical cumulative distribution and $\hat{F}(x_i)$ the fitted cumulative distribution of the data. The models that have minimum values of MSE, MAD and MaxD (close to zero) are the best. A general R code for fitting the distributions introduced in this paper is given in Appendix.

4.1 Application 1: Expenditure on Education

The first data set giving the total spending on public education (% of GDP - Gross Domestic Product) in various countries in 2003, has unimodal and asymmetrical behavior. These data are given in the site www.worldbank.org [8]. The expenditure on public education includes the current and capital spending by private and government agencies, educational institutions both public and private, as well as educational administration and subsidies to private (student / family) entities.

The software R is used in (27) and (25) respectively to calculate the estimates of the parameters through maximum likelihood method and the R function *constrOptim* [7] is used to maximize the log-likelihood function. The maximum likelihood estimates for the parameters of the models are given by:

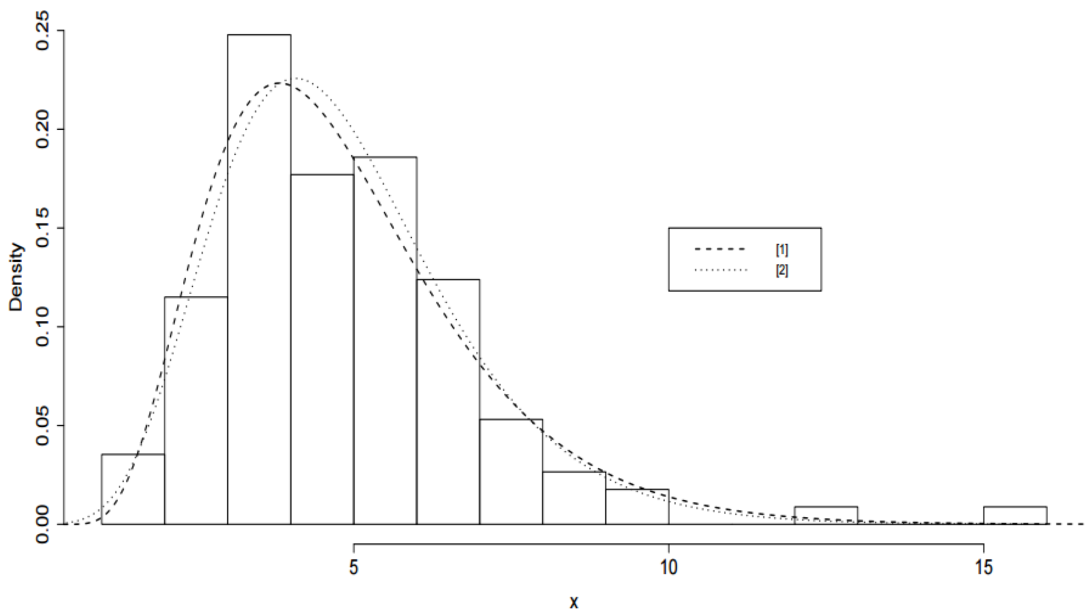
- GG-EW: $\hat{u} = 1.64 \times 10^{-11}, \hat{v} = 4.22, \hat{\alpha} = 2.16, \hat{a} = 19.49, \hat{B} = 0.3233, \hat{c} = 1.55, \hat{\mu} = 0.21$ and $\hat{\sigma} = 0.13$;

- GG-Dagum: $\hat{\beta} = 0.0087, \hat{\lambda} = 0.0018, \hat{\alpha} = 14.42, \hat{a} = 13.16, \hat{b} = 0.0004, \hat{c} = 2.92, \hat{\mu} = -2.19$ and $\hat{\sigma} = 2.41$.

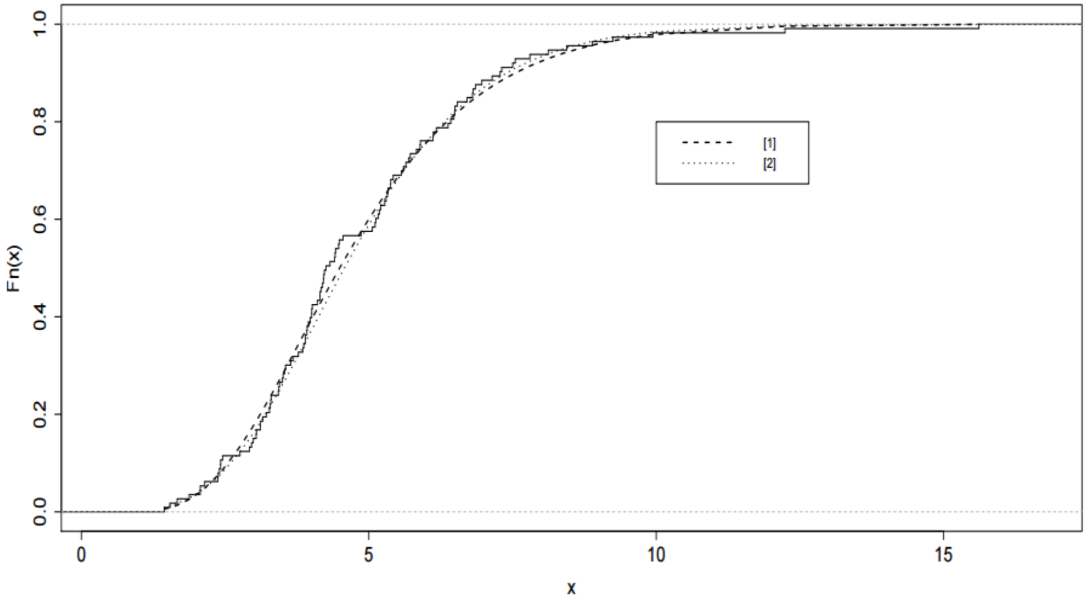
The Figure 1 illustrates the fit of the distributions. The Figure 2 illustrates the *pp plot* of the two distributions. The performance of the two fitted distributions is given in Table 1. The p-value of the KS test tells us that both the distributions can be used to model the data. Both the distributions presented same values of AIC, BIC and AICC. The GG-EW distribution indicated better results as compared to the GG-Dagum distribution (smaller values of MSE, MAD and MaxD).

Table 1: Performance and accuracy of the distributions.

Model	**AIC**	**BIC**	**AICC**	**KS-Test (p-value)**	**MSE**	**MAD**	**MaxD**
GG-EW	448.233	426.413	462.848	0.9818	0.0004	0.0171	0.0591
GG-Dagum	448.233	426.413	462.848	0.866	0.0007	0.0188	0.0773

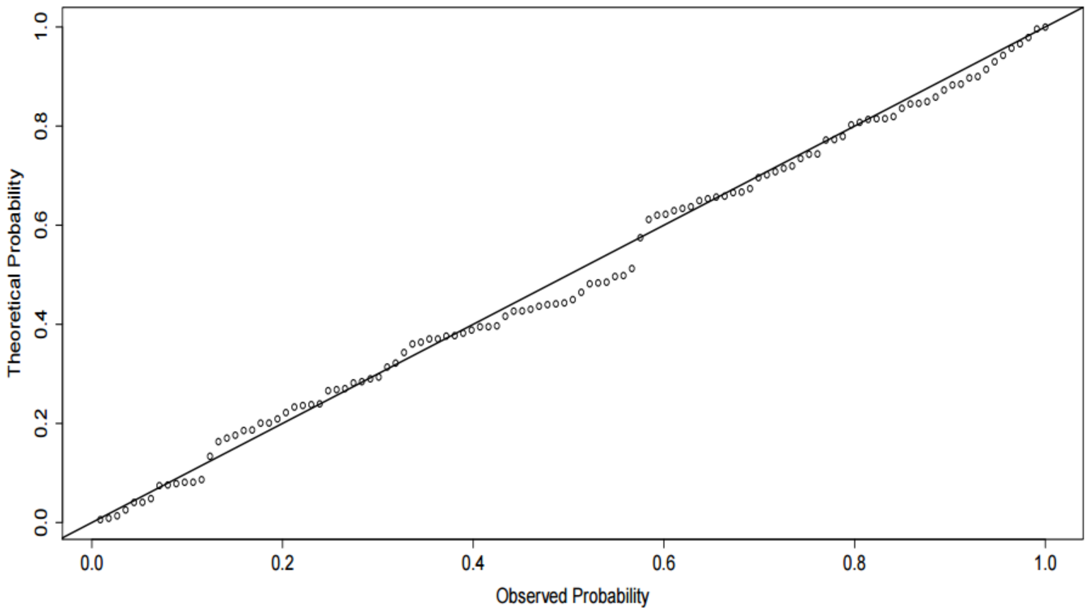


(a) Probability density function.

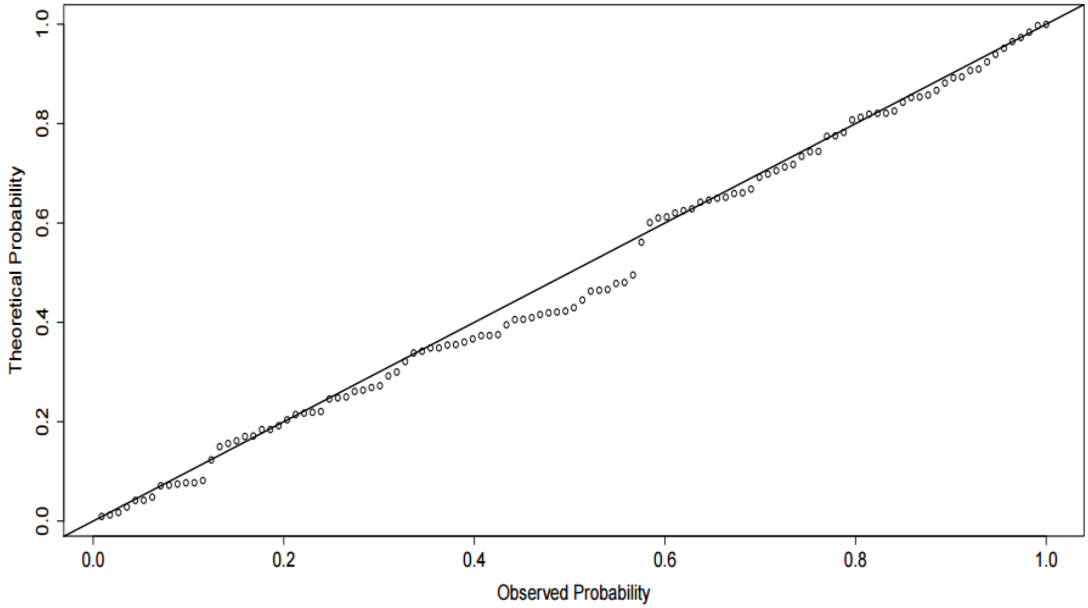


(a) Cumulative distribution.

Figure 1: Education data - Fitted distributions. [1] Generalized Gamma-generated Exponential Weibull distribution. [2] Generalized Gamma-generated Dagum distribution.



(a) Skew Normal-GL.



(a) Skew GL-Normal.

Figure 2: Education data - PP-Plot.

4.2 Application 2: Expenditure on Health

The second data set has assymetric behavior, and relates to the total expenditure, in 2009, on health (% of GDP - Gross Domestic Product) in various countries. These data are obtained from the site data.worldbank.org [9]. Total health expenditure is the sum of expenses with public and private health and also covers the provision of health services (preventive and curative), family planning activities, nutrition activities and emergency aid designated for health but does not include water supply and sanitation.

The maximum likelihood method is used on (27) and (25) respectively to estimate the model parameters. The software R is used to calculate estimates of the parameters by using the R function *constrOptim* [7] to maximize the log-likelihood function. The maximum likelihood estimates for the parameters of the models are given by:

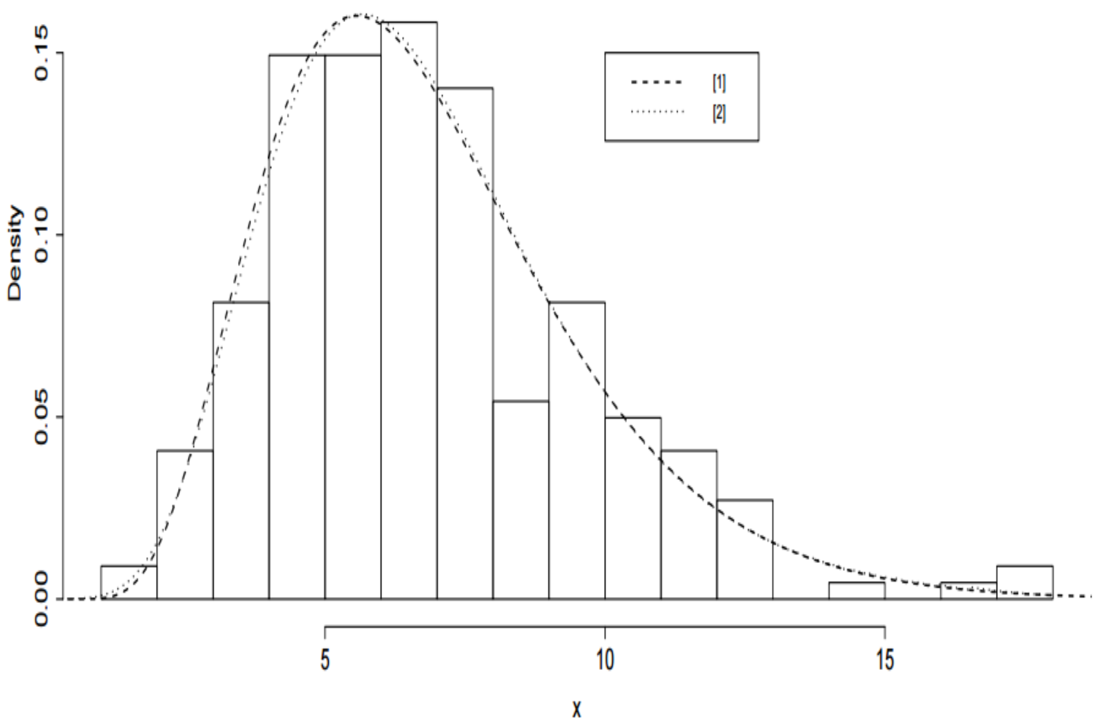
- GG-EW: $\hat{u} = 3.83 \times 10^{-11}, \hat{v} = 3.97, \hat{\alpha} = 2.16, \hat{a} = 15.85, \hat{B} = 0.3662, \hat{c} = 1.55, \hat{\mu} = 0.208$ and $\hat{\sigma} = 0.118$;

- GG-Dagum: $\hat{\beta} = 0.10, \hat{\lambda} = 0.004, \hat{\alpha} = 13.81, \hat{a} = 13.26, \hat{b} = 0.0003, \hat{c} = 2.90, \hat{\mu} = -2.05$ and $\hat{\sigma} = 2.32$.

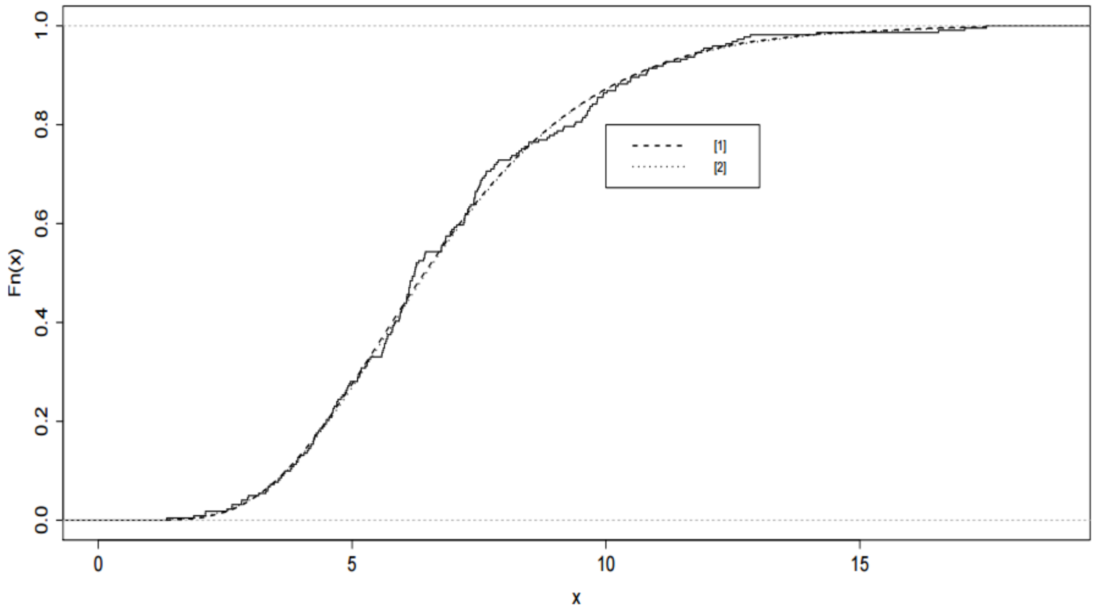
The Figure 3 illustrates the fit of the distributions. The Figure 4 illustrates the *pp plot* of both the distributions. The performance of both the fitted distributions is included in Table 2. Looking the p-value of the KS test, both the distributions can be used to model the data. The GG-EW distribution indicated better results as compared to the GG-Dagum distribution (smaller values of MSE, MAD and MaxD). Both the distributions presented same values of AIC, BIC and AICC.

Table 2: Performance and accuracy of the distributions.

Model	**AIC**	**BIC**	**AICC**	**KS-Test (p-value)**	**MSE**	**MAD**	**MaxD**
GG-EW	448.233	421.047	463.553	0.9774	0.0002	0.0116	0.0451
GG-Dagum	448.233	421.047	463.553	0.9004	0.0003	0.0127	0.0513

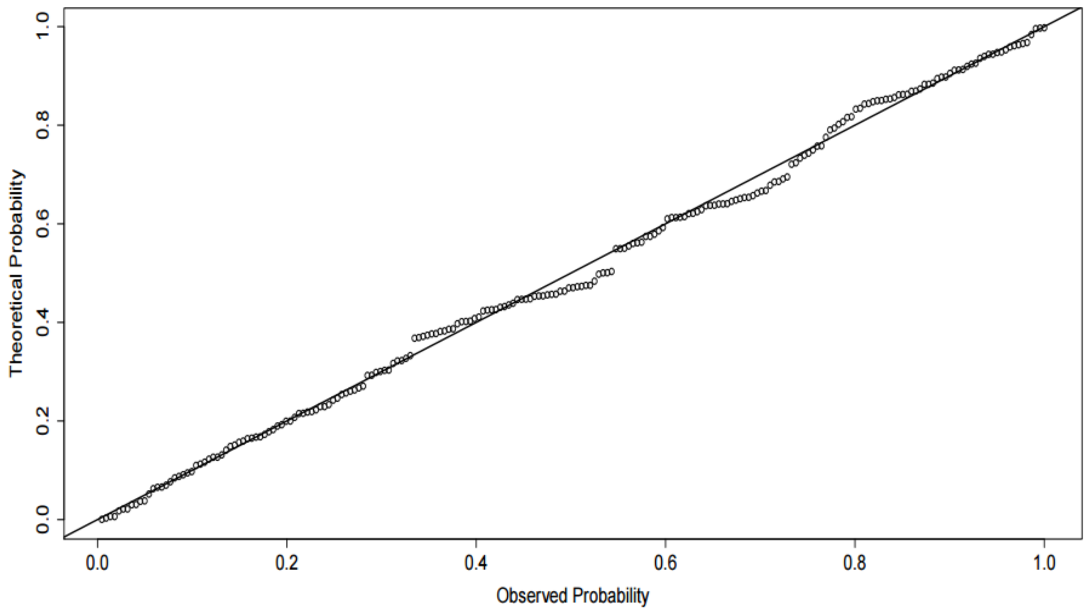


(a) Probability density function.

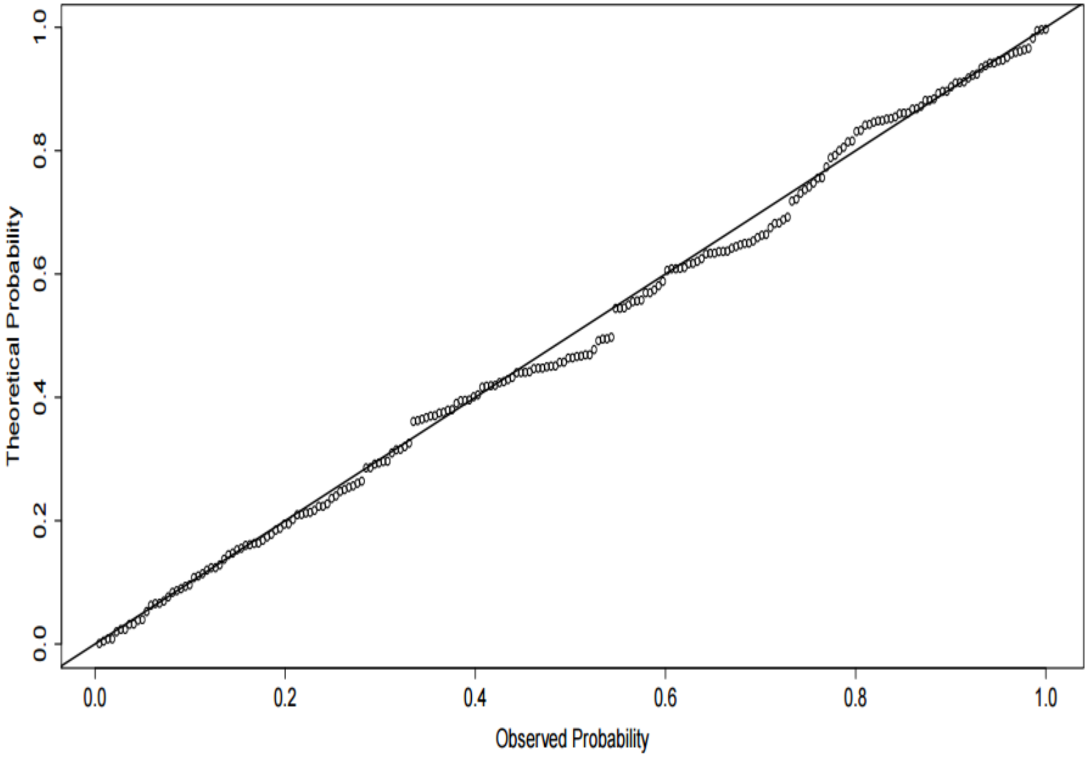


(b) Cumulative distribution.

Figure 3: Health data - Fitted distributions. [1] Generalized Gamma-generated Exponential Weibull distribution. [2] Generalized Gamma-generated Dagum distribution.



(a) Skew Normal-GL.



(b) Skew GL-Normal.

Figure 4: Health data - PP-Plot.

5. Conclusions

Two forms of generalized gamma-generated distributions with any base distribution are given along with two of its particular cases (GG-EW and GG-Dagum) involving exponentiated Weibull and Dagum distributions as base distributions. It is shown, in general, that the generalized gamma-generated distributions are infinite linear combinations of distributions which are powers of the base distributions. As applications, the GG-EW and GG-Dagum distributions are used to analyse the two real data sets involving expenditures on education and health, showing that the distributions are flexible to adjust asymmetric data with unimodal behaviour.

Acknowledgement

P. N. Rathie thanks CAPES, Brazil, for supporting his Senior National Visiting Professorship.

References

[1] F. Castellares, A.J. Lemonte, *A new generalized Weibull distribution generated by gamma random variables*, Journal of Egyptian Mathematical Society, 23(2) (2015), 382-390.

[2] C. Dagum, *A new model of personal income distribution: specification and estimation*, Economie Applique'e, 30 (1977), 413-437.

[3] Y.L. Luke, *The Special Functions and Their Approximations*, Academic Press, New York, 1969.

[4] G.S. Mudholkar, D.K. Srivastava, *Exponentiated Weibull family for analysing bathtub failure-rate*, IEEE Trans. Reliab., 42 (1993), 299-302.

[5] B.O. Oluyede, S. Huang, M. Pararai, *A new class of generalized Dagum distribution with applications to income and lifetime data*, Journal of Statistical and Econometric Methods, 3(2) (2014), 125-151.

[6] L.G.B. Pinho, G.M. Cordeiro, J.S. Nobre, *The gamma-exponentiated Weibull distribution*, Journal of Statistical Theory and Application, 11(4) (2012), 379-395.

[7] R Core Team, R, *A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015. *URL: http:// www.R-project.org/. constrOptim URL: https:// stat.ethz.ch/ R-manual/ R-devel/ library/ stats/ html/ constrOptim.html.*

[8] World Bank web site, *URL: http://data.worldbank.org/ indicator/ SE. XPD. TOTL. GD. ZS*, accessed on 15/01/2017.

[9] World Bank web site, *URL: http://data.worldbank.org/ indicator/ SH. XPD. TOTL. ZS*, accessed on 15/01/2017.

Appendix: General R codes

The general R codes for fitting the two distributions introduced in this paper are given below.

GG-EW distribution

```
#READ THE DATA
data <- read.csv(file.choose(), header=T, stringsAsFactor=F, sep=';')
################GG-WE - DENSITY################
dwe <- function(y,u,v,alpha,mu,sigma){
z <- (y-mu)/sigma
dens=((u*v*alpha)/sigma)*(z**(v-1))*exp(-u*(z**v))*
((1-exp(-u*(z**v)))**(alpha-1))
return(dens)
}
pwe <- function(y,u,v,alpha,mu,sigma){
z <- (y-mu)/sigma
pdens=((1-exp(-u*(z**v)))**alpha)
return(pdens)
}
dggwe <- function(y,u,v,alpha,a,b,c,mu,sigma){
z <- (y-mu)/sigma
const=(c*(b**(a/c)))/gamma(a/c)
r=dwe(y=y, mu=mu, sigma=sigma, u=u, v=v, alpha=alpha)/
pwe(y=y, mu=mu, sigma=sigma,
u=u, v=v, alpha=alpha)
dens=(const)*((-log(pwe(y=y, mu=mu, sigma=sigma, u=u,
 v=v, alpha=alpha)))**(a-1))*
exp(-b*((-log(pwe(y=y, mu=mu, sigma=sigma, u=u, v=v,
alpha=alpha)))**c))*r
return(dens)
}
#GIVE THE INITIAL PARAMETERS HERE
theta <- theta0
#LOG-LIKELIHOOD FUNCTION
loglik <- function(pars){
u <- pars[1]
v <- pars[2]
alpha <- pars[3]
a <- pars[4]
b <- pars[5]
c <- pars[6]
```

```
mu <- pars[7]
sigma <- pars[8]
logl <-sum(log(dggwe(x,u=u,v=v,alpha=alpha,a=a,b=b,c=c,
mu=mu,sigma=sigma)))
return(-logl)
}
#FIT
fit=constrOptim(theta=theta, f=vero,
ui=rbind(c(1, 0, 0, 0, 0, 0, 0, 0),
c(0, 1, 0, 0, 0, 0 ,0 ,0),
c(0, 0, 1, 0, 0, 0, 0, 0),
c(0, 0, 0, 1, 0, 0, 0, 0),
c(0, 0, 0, 0, 1, 0, 0, 0),
c(0, 0, 0, 0, 0, 1, 0, 0),
c(0, 0, 0, 0, 0, 0, 0, 1)), ci=c(0, 0, 0, 0, 0, 0, 0)
, method="Nelder-Mead", outer.iterations=300)
```

GG-Dagum distribution

```
#READ THE DATA
data <- read.csv(file.choose(), header=T, stringsAsFactor=F, sep=';')
################GG-Dagum - DENSITY################
ddagum <- function(y,beta,lambda,alpha,mu,sigma){
z <- (y-mu)/sigma
dens=((beta*lambda*alpha)/sigma)*(z**(-alpha-1))*((1+lambda*
(z**(-alpha)))**(-beta-1))
return(dens)
}
sdagum <- function(y,beta,lambda,alpha,mu,sigma){
z <- (y-mu)/sigma
pdens=1-(1+lambda*(z**(-alpha)))**(-beta)
return(pdens)
}
dggdagum <- function(y,beta,lambda,alpha,a,b,c,mu,sigma){
z <- (y-mu)/sigma
const=(c*(b**(a/c)))/gamma(a/c)
r=ddagum(y=y, beta=beta, lambda=lambda, alpha=alpha,mu=mu,
sigma=sigma)/ sdagum(y=y, beta=beta, lambda=lambda,
```

```
alpha=alpha,mu=mu,sigma=sigma)
dens=(const)*((-log(sdagum(y=y, beta=beta, lambda=lambda,
alpha=alpha,mu=mu,sigma=sigma)))**(a-1))*exp(-b*((-log(
sdagum(y=y, beta=beta, lambda=lambda, alpha=alpha,
mu=mu,sigma=sigma)))**c))*r
return(dens)
}
#GIVE THE INITIAL PARAMETERS HERE
theta <- theta0
#LOG-LIKELIHOOD FUNCTION
loglik <- function(pars){
beta <- pars[1]
lambda <- pars[2]
alpha <- pars[3]
a <- pars[4]
b <- pars[5]
c <- pars[6]
mu <- pars[7]
sigma <- pars[8]
logl <-sum(log(dggdagum(x, beta=beta, lambda=lambda, alpha =alpha,
 a=a, b=b, c=c, mu=mu, sigma=sigma)))
return(-logl)
}
#FIT
fit=constrOptim(theta=theta, f=vero,
ui=rbind(c(1, 0, 0, 0, 0, 0, 0, 0),
c(0, 1, 0, 0, 0, 0 ,0 ,0),
c(0, 0, 1, 0, 0, 0, 0, 0),
c(0, 0, 0, 1, 0, 0, 0, 0),
c(0, 0, 0, 0, 1, 0, 0, 0),
c(0, 0, 0, 0, 0, 1, 0, 0),
c(0, 0, 0, 0, 0, 0, 0, 1)), ci=c(0, 0, 0, 0, 0, 0, 0)
, method="Nelder-Mead", outer.iterations=300)
```