Indian Institute Of Management Kozhikode Working Paper

COMPOUNDED GENERALIZED WEIBULL DISTRIBUTIONS - A UNIFIED APPROACH

Shovan Chowdhury

IIMK/WPS/148/QM&OM /2014/06

March 2014

IIMK/WPS/148/QM&OM /2014/06

COMPOUNDED GENERALIZED WEIBULL DISTRIBUTIONS - A UNIFIED APPROACH

Shovan Chowdhury[1]

[1]Indian Institute of Management Kozhikode, IIMK Campus PO, Kozhikode- 673570,
email: shovanc@iimk.ac.in

Compounded Generalized Weibull Distributions - A Unified Approach

Shovan Chowdhury

Department of Quantitative Methods and Operations Management,

Indian Institute of Management; Kozhikode; India

E-mail: shovanc@iimk.ac.in

Abstract: A unified approach is proposed in this paper to study a family of lifetime distributions of a system consisting of random number of components in series and in parallel. While the lifetimes of the components are assumed to follow generalized (exponentiated) Weibull distribution, a zero-truncated Poisson is assigned to model the random number of components in the system. The resulting family of compounded distributions describes several well-known distributions as well as some new models with some of their statistical and reliability properties. Various ageing classes of life distributions including increasing, decreasing, bath-tub, upside-down-bathtub and roller coaster shaped failure rates are covered by the family of compounded distributions. The simplest algorithm for maximum likelihood method of estimation of the model parameters is discussed. Some numerical results are obtained via Monte-Carlo Simulation. The asymptotic variance-covariance matrices of the estimators are also obtained. Five different real data sets are used to validate the distributions and the results demonstrate that the family of distributions can be considered as a suitable model under several real situations.

Keywords: Unified approach, Compounding, Generalized Weibull Distribution; Hazard Function; ML Estimation; Zero-Truncated Poisson Distribution.

AMS 2010 subject classification: Primary 62E99; Secondary 62N05, 62F10, 62-07

1 Introduction

Lifetime data modeling in the literature of reliability analysis is studied extensively by several researchers with the help of different lifetime distributions which are mainly based on some modifications and generalizations of exponential or Weibull distributions. While this modification is carried out in some of the life distributions through exponentiation viz. Exponentiated or Generalized Exponential (GE) distribution (e.g. Gupta and Kundu, 1999), Exponentiated Weibull (EW) distribution (e.g. Mudholkar and Srivastava, 1993; Mudholkar et al., 1995; Mudholkar and Hutson, 1996; Nassar and Eissa, 2003; Nadarajah and Kotz, 2006), there are others where lifetime distributions are compounded with distribution of unknown number of components yielding a new class of life distributions viz. Exponential-Geometric (EG) distribution (Adamidis and Loukas, 1998), Extended Exponential-Geometric (EEG) distribution (Adamidis et al., 2005), Exponential-Poisson (EP) distribution (Kus, 2007), Exponential-Logarithmic (EL) distribution (Tahmasbi and Rezaei, 2008), Exponentiated Exponential-Poisson (EEP) distribution (Barreto-Souza and Cribari-Neto, 2009), Generalized Exponential-Geometric (GEG) distribution (Rodrigo et al., 2010), Weibull-Geometric (WG) distribution (Barreto-Souza et al., 2011), Weibull-Poisson (WP) distribution (Hemmati et al., 2011), Exponentiated Weibull-Poisson (EWP) Distribution (Mahmoudi and Sepahdar, 2013).

Consider a system consists of K components in series such that the system fails if at least one of the K units fails. Hence, the failure time distribution of the system is the distribution of the failure time of the first (the one with minimum lifetime) out of the K components. In contrast, if the components are in parallel, the failure time distribution of the system is the failure time distribution of the last (the one with maximum lifetime) of the K components. In this context, it is interesting to note that most of the works on compounded distributions associated with exponential or Weibull models are based on the assumption that the components are in series. The compounded exponential or Weibull type distributions with components in parallel are largely overlooked. Unlike the previous works, in this paper, we consider both systems with components in series and parallel to study a family of compounded distribution. As in Mahmoudi and Sepahdar (2013), we consider that the lifetime of each of the system components is independently and identically distributed as exponentiated Weibull and the number of components in the system follows a zero truncated Poisson model. It is shown that irrespective of the system alignments, namely, series or parallel, the compounded distribution can be represented in a unified manner with the help of an additional parameter.

Moreover, several well-known distributions can be obtained as special cases of the proposed family of distributions viz. EP (Kus, 2007), EEP(Barreto-Souza and Cribari-Neto,

2009) and WP (Hemmati et al., 2011) in the series system and several other new distributions in parallel system, which, to the best of our knowledge, are not studied in the literature before. This family of distributions has increasing, decreasing, bath-tub, up-side-down bathtub and roller coaster hazard functions. The proposed family of distributions encompasses better fits to many real data sets including one used by Hemmati et al., 2011.

The rest of the paper is organized as follows. In Section 2, compounded generalized Weibull Poisson type distributions are obtained for series and parallel systems and are presented in a unified manner. In Section 3, various properties of these distributions are discussed. Parameters of the distributions are estimated in Section 4 by the maximum likelihood method through a simulation study, and asymptotic variances and covariances of maximum likelihood estimators (MLEs) are obtained. Five real life examples are provided in Section 5 with a detailed compar- ison with the other competitive distributions. Finally, Section 6 concludes the manuscript.

2 The Generalized Weibull Poisson Distribution

The Exponentiated Weibull distribution with scale parameters $\lambda > 0$ and shape parameters $\gamma > 0, \alpha > 0$, denoted as $EW(\lambda, \gamma, \alpha)$, has the following probability density function (PDF):

$$g_w(w; \lambda, \gamma, \alpha) = \alpha\gamma\lambda^\gamma w^{\gamma-1} e^{-(\lambda w)^\gamma} \left[1 - e^{-(\lambda w)^\gamma}\right]^{\alpha-1}; w, \lambda, \gamma, \alpha > 0 \tag{2.1}$$

with cumulative distribution function (CDF)

$$G_W(w \mid \lambda, \gamma, \alpha) = \left[1 - e^{-(\lambda w)^\gamma}\right]^{\alpha}; w, \lambda, \gamma, \alpha > 0 \tag{2.2}$$

Let $W_1, W_2, ..., W_K$ be independently and identically distributed (iid) failure times of K components in series following $EW(\lambda, \gamma, \alpha)$ and K be a zero-truncated Poisson variable with pmf given by

$$P(k; \mu) = \frac{e^{-\mu}\mu^k}{k!(1 - e^{-\mu})}; k \in \mathcal{N}, \mu > 0 \tag{2.3}$$

where $\mathcal{N}$ is the set of natural numbers. The motivation behind the use of the zero-truncated Poisson distribution (to model counts) is that it has been used in many popular applications, viz., to model (i) number of illegal immigrants in four large cities in the Netherlands (Heijden et al., 2003), (ii) mental health services data (Elhai et al., 2008), (iii) word or species frequency count data (Ginebra and Puig, 2010), and (iv) fertility trait phenotypes (Xu and Hu, 2011). Assuming that random variables W and K are independent and $\boldsymbol{\theta} = (\mu, \lambda, \gamma, \alpha)$, we define

$$X = \min_{1 \le i \le K} W_i$$

to model failure time distribution of a system of components in series.
Then the conditional distribution of $(X \mid K = k)$ is given by

$$f_X(x \mid k) = z\alpha\gamma\lambda^\gamma x^{\gamma-1} e^{-(\lambda x)^\gamma} \left(1 - e^{-(\lambda x)^\gamma}\right)^{\alpha-1} \left[1 - (1 - e^{-(\lambda x)^\gamma})^\alpha\right]^{k-1}; x, \lambda, \gamma, \alpha > 0 \quad (2.4)$$

with the unconditional probability density function (pdf) of X as

$$f(x;\theta) = \mu\alpha\gamma\lambda^\gamma (1 - e^{-\mu})^{-1} x^{\gamma-1} e^{-(\lambda x)^\gamma} \left[1 - e^{-(\lambda x)^\gamma}\right]^{\alpha-1} e^{-\mu\left[1-e^{-(\lambda x)^\gamma}\right]^\alpha}; x, \lambda, \gamma, \alpha > 0 \quad (2.5)$$

Hereafter, the distribution of X is referred to as the Generalized Weibull Poisson distribution of Type-I and is denoted by GWP(I).
In this context, it may be interesting to note that, for a system of parallel components, writing

$$Y = \max_{1 \le i \le K} W_i,$$

we get another type of GWP distribution, as obtained by Mahmoudi and Sepahdar, 2013, say GWP distribution of Type-II and may be denoted by GWP(II). Using Equations 2.1-2.3, we get conditional pdf of GWP(II) as

$$g_Y(x \mid k) = z\alpha\gamma\lambda^\gamma x^{\gamma-1} e^{-(\lambda x)^\gamma} \left[1 - e^{-(\lambda x)^\gamma}\right]^{\alpha k-1}; x, \lambda, \gamma, \alpha > 0 \quad (2.6)$$

One can easily check that the unconditional pdf of GWP(II) is given by

$$g(x;\theta) = \mu\alpha\gamma\lambda^\gamma e^{-\mu} (1 - e^{-\mu})^{-1} x^{\gamma-1} e^{-(\lambda x)^\gamma} \left[1 - e^{-(\lambda x)^\gamma}\right]^{\alpha-1} e^{\mu\left[1-e^{-(\lambda x)^\gamma}\right]^\alpha}; x, \lambda, \gamma, \alpha > 0 \quad (2.7)$$

At this junction, we have unified GWP(I) and GWP(II) as derived in (2.5) and (2.7), through a parameter c to yield the following distribution and will be named GWP hereafter.

$$f^*(x;\theta) = \mu\alpha\gamma\lambda^\gamma e^{-c\mu} (1 - e^{-\mu})^{-1} x^{\gamma-1} e^{-(\lambda x)^\gamma} \left[1 - e^{-(\lambda x)^\gamma}\right]^{\alpha-1} e^{(-1)^{c+1}\mu\left[1-e^{-(\lambda x)^\gamma}\right]^\alpha};$$
$$x, \lambda, \gamma, \alpha > 0;\ c = 0, 1 \quad (2.8)$$

We get GWP(I) and GWP(II) distributions for $c = 0$ and $c = 1$ respectively. Now onwards, distribution(I) and distribution(II) would refer to the distribution of failure time of a system consisting of components in series and parallel respectively. Moreover, GWP(I) and GWP(II) together will be called GWP distribution as mentioned earlier. All the results derived in the next sections are based on GWP and hence we get properties of GWP(I) and GWP(II) on substituting c by 0 and 1 respectively. It is to be noted that all the results obtained at $c = 1$ can be found in Mahmoudi and Sepahdar, 2013.

The following values of the parameters are of particular interest as they produce many existing and new lifetime distributions: (i) $\gamma = 1$, GWP reduces to EEP distribution, hence GWP(I) reduces to EEP(I) and GWP(II) reduces to EEP(II); (ii) $\gamma = 2$, the GWP(I) and GWP(II) reduce to generalized Rayleigh-Poisson GRP(I) and GRP(II) respectively, (iii) $\alpha = 1$, GWP(I) corresponds to WP(I) (Hemmati et al., 2011) and GWP(II) reduces to WP(II), (iv) $\alpha = 1$ and $\gamma = 1$, GWP(I) and GWP(II) reduce to EP(I) and EP(II) respectively, (v) $\alpha = 1$ and $\gamma = 2$, GWP(I) and GWP(II) reduce to Rayleigh-Poisson RP(I) and RP(II) respectively and (vi) as μ approaches zero, both of GWP(I) and GWP(II) reduce to EW; Moreover, results in (i) and (vi) together yield exponentiated exponential distribution and (ii) and (vi) yield generalized Rayleigh distribution. It is now evident that the present paper not only proposes GWP distribution, but introduces several others including EEP(II), GRP(I), GRP(II), WP(II), EP(II) and RP(II).

It can be easily shown that the GWP distribution can be expressed as an infinite mixture of exponentiated Weibull distribution with the same scale parameter λ and shape parameters γ and $\alpha(j + 1)$, j=0,1,2,...

It is evident that both the pdf's tend to zero as $x \to \infty$. Further, it is clear from Figure 2. that both GWP(I) and GWP(II) are either decreasing or unimodal. The shape of the pdfs of GWP(I) and GWP(II) are also illustrated in this figure.

<Figure 2. HERE.>

3 Properties of the distribution

3.1 Distribution function and moments

The CDF of GWP is derived from (2.8) and is given by

$$F(x;\theta) = \frac{(-1)^{c+1}e^{-c\mu}\left[e^{(-1)^{c+1}\mu\left(1-e^{-(\lambda x)^{\gamma}}\right)^{\alpha}}-1\right]}{1-e^{-\mu}};\ x>0,\ c=0,1 \tag{3.1}$$

As an immediate consequence, quantile function (ξ_p), of the GWP can be derived as

$$\xi_p = \frac{-1}{\lambda}\left[\ln\left\{1-\left(\frac{(-1)^{c+1}}{\mu}\ln\left(1+(-1)^{c+1}e^{c\mu}(1-e^{-\mu})p\right)\right)^{1/\alpha}\right\}\right]^{1/\gamma},\ c=0,1 \tag{3.2}$$

For $r \in \mathcal{N}$, the rth order raw moments of the GWP can be obtained from (2.8) as follows:

$$\mu_r^{'} = \frac{\kappa\Gamma(r/\gamma+1)}{\lambda^r}\sum_{j=0}^{\infty}\left[\frac{(-1)^{(c+1)j}\mu^j}{j!}\left\{1+\sum_{i=1}^{\infty}\left(\eta_i(i+1)^{-(r/\gamma+1)}\right)\right\}\right],\ c=0,1 \tag{3.3}$$

where $\kappa = \frac{\mu\alpha e^{-c\mu}}{(1-e^{-\mu})}$, $\eta_i = \prod_{l=0}^{\infty}\frac{(-1)^i(\upsilon-l)}{i!}$; $i=1,2,...$ and $\upsilon=\alpha(1+j)-1$. $\sum_{i=1}^{\infty}\left(\eta_i(i+1)^{-(r/\gamma+1)}\right)$ is convergent for $\alpha(1+j)>0$. Expression in (3.3) allows us to derive coefficient of variation (CV), measure of skewness (γ_1) and measure of kurtosis (γ_2) of the GWP distribution as follows:

$$\begin{aligned}
CV &= \kappa^{-1/2}\left[\frac{(\Gamma_2S_2(\mu,\gamma)-\kappa(\Gamma_1)^2(S_1(\mu,\gamma))^2)^{1/2}}{(\Gamma_1S_1(\mu,\gamma))}\right]\\
\gamma_1 &= \kappa^{-1/2}\left[\frac{\Gamma_3S_3(\mu,\gamma)-3\kappa\Gamma_2\Gamma_1S_1(\mu,\gamma)S_2(\mu,\gamma)+2\kappa^2\left(\Gamma_1\right)^3\left(S_1(\mu,\gamma)\right)^3}{\left(\Gamma_2S_2(\mu,\gamma)-\kappa\left(\Gamma_1\right)^2\left(S_1(\mu,\gamma)\right)^2\right)^{3/2}}\right]\\
\gamma_2 &= \kappa^{-1}\left[\frac{\Gamma_4J_4(\mu,\gamma)-4\kappa\Gamma_3\Gamma_1S_3(\mu,\gamma)+6\kappa^2\Gamma_2\left(\Gamma_1\right)^2S_2(\mu,\gamma)\left(S_1(\mu,\gamma)\right)^2-3\kappa^3\left(\Gamma_1\right)^4}{\left(\Gamma_2S_2(\mu,\gamma)-\kappa\left(\Gamma_1\right)^2\left(S_1(\mu,\gamma)\right)^2\right)^2}\right]-3
\end{aligned} \tag{3.4}$$

where $\Gamma_r=\Gamma(r/\gamma+1)$ and $S_r(\mu,\gamma)=\sum_{j=0}^{\infty}\left[\frac{(-1)^{(c+1)j}\mu^j}{j!}\left\{1+\sum_{i=1}^{\infty}\left(\eta_i(i+1)^{-(r/\gamma+1)}\right)\right\}\right]$, $r=1,2,3,4$. It is to be noted that properties of GWP(I) and GWP(II) are obtained by taking $c=0$ and 1 respectively.

3.2 Survival and Hazard functions

Survival functions of the GWP is derived as follows.

$$S(x;\theta) = \frac{1-e^{-\mu}-e^{-c\mu}(-1)^{c+1}\left\{e^{(-1)^{c+1}\mu(1-e^{-(\lambda x)^{\gamma}})^{\alpha}}-1\right\}}{1-e^{-\mu}}; x>0,\ c=0,1. \tag{3.5}$$

Consequently, hazard function of the GWP is obtained as

$$h(x;\theta) = \frac{\mu\alpha\gamma\lambda^{\gamma}x^{\gamma-1}e^{-\{(\lambda x)^{\gamma}+c\mu\}}\left[1-e^{-(\lambda x)^{\gamma}}\right]^{\alpha-1}e^{(-1)^{c+1}\mu\left[1-e^{-(\lambda x)^{\gamma}}\right]^{\alpha}}}{1-e^{-\mu}-e^{-c\mu}(-1)^{c+1}\left\{e^{(-1)^{c+1}\mu(1-e^{-(\lambda x)^{\gamma}})^{\alpha}}-1\right\}};\ x>0,\ c=0,1. \tag{3.6}$$

Mean Residual life (MRL) of the GWP distribution is given by

$$m(x_0;\theta) = E_{\theta}(X-x_0 \mid X \geq x_0) = \frac{1}{S(t)}\int_{x_0}^{\infty}(x-x_0)f(x)dx \tag{3.7}$$

which yields

$$\begin{aligned} m(x_0;\theta) = \frac{\kappa}{\lambda}\sum_{j=0}^{\infty}\Bigg[\frac{(-1)^{(c+1)j}\mu^{j}}{j!}&\left\{\Gamma_{(\lambda x_0)^{\gamma}}(1/\gamma+1)\right.\\ +\sum_{i=1}^{\infty}\left(a_i(i+1)^{-(1/\gamma)}\Gamma_{(i+1)(\lambda x_0)^{\gamma}}(1/\gamma+1)\right)&;\ c=0,1, \end{aligned} \tag{3.8}$$

where $\Gamma_x(n)$ represents the incomplete gamma integral given by

$$\Gamma_x(n) = \int_0^x e^{-v}v^{n-1}dv \tag{3.9}$$

The shape of the hazard functions of GWP(I) and GWP(II) are illustrated in Figure 3. It is observed that the failure rate of GWP(I) can take all shapes viz. increasing, decreasing, bath-tub, upside-down bath-tub and roller coaster depending on the values of the parameters as opposed to the mostly used lifetime distributions available in the literature. On the other hand, GWP(II) can exhibit increasing, decreasing and upside-down bath-tub shaped failure rates.

<Figure 3. HERE.>

4 Estimation of the parameters

Here, we consider estimation of the unknown parameters of the GWP distribution by the method of maximum likelihood. Let $x_1, x_2, ..., x_n$be a random sample of size n drawn from (2.8) with parameters $\boldsymbol{\theta} = (\mu, \lambda, \gamma, \alpha)$. Then the log-likelihood functions, $L(\theta)$ for GWP can be written as

$$L(\theta) = n\log\mu + n\log\alpha + n\log\gamma + n\gamma\log\lambda - nc\mu + \sum_{i=1}^{n}(\gamma-1)\log x_i - \sum_{i=1}^{n}(\lambda x_i)^{\gamma} + \sum_{i=1}^{n}(\alpha-1)\log(1-e^{-(\lambda x_i)^{\gamma}}) + (-1)^{c+1}\mu\sum_{i=1}^{n}\left(1-e^{-(\lambda x_i)^{\gamma}}\right)^{\alpha} - n\log(1-e^{-\mu});\ \ c=0,1 \tag{4.1}$$

The likelihood equations are

$$\begin{aligned}
\frac{\delta L}{\delta\mu} &= n/\mu - n\left(c + \frac{1}{e^{\mu}-1}\right) + (-1)^{c+1}\sum_{i=1}^{n} C_i \\
\frac{\delta L}{\delta\lambda} &= n\gamma/\lambda + \gamma\sum_{i=1}^{n}\left[(\lambda x_i)^{\gamma-1}x_i\left(-1 + \frac{(\alpha-1)B_i}{1-B_i} + (-1)^{c+1}\mu\alpha\frac{B_iC_i}{1-B_i}\right)\right] \\
\frac{\delta L}{\delta\gamma} &= n/\gamma + n\log\lambda + \sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\left[A_iD_i\left(-1 + \frac{(\alpha-1)B_i}{1-B_i} + (-1)^{c+1}\mu\alpha B_i(1-B_i)^{\alpha-1}\right)\right] \\
\frac{\delta L}{\delta\alpha} &= n/\alpha + \sum_{i=1}^{n}\left[(1 + (-1)^{c+1}\mu C_i)\log(1-B_i)\right]
\end{aligned} \tag{4.2}$$

where $c = 0,1$; $A_i = (\lambda x_i)^{\gamma}$; $B_i = e^{-A_i}$; $C_i = (1-B_i)^{\alpha}$; $D_i = \log(\lambda x_i)$
The maximum likelihood estimators (MLE) of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, are the simultaneous solutions of the likelihood equations (4.2). Since no closed form expressions for the mle's of $\boldsymbol{\theta}$ are available from the expressions (4.2), we go for simulation study, which is detailed in subsection 4.2.

4.1 Asymptotic variance-covariance matrix of the MLE's

The MLE's of $\boldsymbol{\theta} = (\mu, \lambda, \gamma, \alpha) = (\theta_1, \theta_2, \theta_3, \theta_4)$, say $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\lambda}, \hat{\gamma}, \hat{\mu})$ can be considered to follow approximately multivariate normal with mean $\boldsymbol{\theta}$ and a variance-covariance matrix $\boldsymbol{I}^{-1}$ where

$\boldsymbol{I} = \boldsymbol{I(\theta, x_{obs})}$ is the Fisher information matrix for the estimators of $(\mu, \lambda, \gamma, \alpha)$ with elements $I_{i,j} = E\left(M_{i,j}\right)$ with $i, j = 1, 2, 3, 4$ where $M_{i,j} = -\frac{\delta^2 L}{\delta\theta_i\theta_j}$. By differentiating expressions in (4.2) we get $M_{i,j}$'s as

$$
\begin{aligned}
M_{\mu\mu} &= \frac{n}{\mu^2} - \frac{ne^{\mu}}{(e^{\mu}-1)^2} \\
M_{\lambda\lambda} &= \frac{n\gamma}{\lambda^2} + \frac{\gamma(\gamma-1)}{\lambda^2}J_{200(000)0} + \frac{\gamma(\alpha-1)}{\lambda^2}J_{210(101)0} \\
&\quad - \frac{\mu\alpha\gamma(2\gamma-1)}{\lambda^2}J_{210(111)0} - \frac{\gamma^2(\alpha-1)(1-\mu\alpha)}{\lambda^2}J_{220(102)0} \\
M_{\gamma\gamma} &= \frac{n}{\gamma^2} + J_{102(000)0} - (\alpha-1)(J_{112(101)0} - J_{212(101)0} + J_{222(111)0}) - \mu\alpha J_{112(111)0} \\
M_{\alpha\alpha} &= \frac{n}{\alpha^2} - \mu J_{000(110)2} \\
M_{\mu\lambda} &= \frac{-\alpha\gamma}{\lambda}J_{110(111)0} \\
M_{\mu\gamma} &= -\alpha J_{111(111)0} \\
M_{\mu\alpha} &= -J_{000(110)1} \\
M_{\lambda\gamma} &= \frac{-n}{\lambda} + \frac{1}{\lambda}J_{100(000)0} - \frac{\alpha-1}{\lambda}J_{110(101)0} - \frac{\mu\alpha}{\lambda}J_{110(111)0} + \frac{\gamma}{\lambda}J_{101(000)0} - \frac{(\alpha-1)\gamma}{\lambda}J_{111(101)0} \\
&\quad - \frac{\mu\alpha\gamma}{\lambda}J_{111(111)0} + \frac{(\alpha-1)\gamma}{\lambda}J_{211(101)0} - \frac{(\alpha-1)\gamma}{\lambda}J_{221(102)0} - \frac{\mu\alpha\gamma}{\lambda}(J_{211(111)0} + J_{110(111)1}) \\
M_{\lambda\alpha} &= \frac{-\gamma}{\lambda}J_{110(101)0} - \frac{\gamma\mu\alpha}{\lambda}J_{110(111)1} - \frac{\mu\gamma}{\lambda}J_{110(111)0} \\
M_{\gamma\alpha} &= J_{111(101)0} - \mu J_{111(111)0} - \alpha\mu J_{111(111)1}
\end{aligned}
\tag{4.3}
$$

where $J_{jkl(mpq)r} = (-1)^{c+1}\sum_{i=1}^{n}(\lambda x_i)^{j\gamma}e^{-(\lambda x_i)^{k\gamma}}(\log(\lambda x_i))^l(1-e^{-(\lambda x)^{\gamma}})^{m(p\alpha-r)}(\log(1-e^{-(\lambda x)^{\gamma}}))^r$, $c = 0, 1; j, k, l, q, r \in \{0, 1, 2\}$ and $m, p \in \{0, 1\}$. Expectation is taken with respect to the distribution of GWP as in (2.8) and can be easily computed by R software. Then the inverse of $\boldsymbol{I(\theta)}$ evaluated at $\boldsymbol{\theta}$ provides the asymptotic variance-covariance matrix of the MLE's.

4.2 Simulation study

As mentioned earlier, it is not feasible to solve the equations in (4.2) explicitly in order to get MLE's for the family of GWP distributions. However, one can easily find the numerical solution applying some suitable optimization techniques. In the present context, we use in-built *spg* function in R.3.0 software for numerical minimization of negative of log-likelihood function. We carry out detailed simulation studies to capture the means and the standard deviations (SD) of the MLE's of $\boldsymbol{\theta} = (\mu, \lambda, \gamma, \alpha)$. Five thousand replicates of Monte-Carlo experiments of size 50, 100 and 500 are considered in the present investigation for each of the six sets of $\boldsymbol{\theta}$. The results from simulated data sets are reported in Tables 4.1. and 4.2. The results show that the estimates are quite stable around the assumed values of $\boldsymbol{\theta}$ and moreover, standard errors of the MLEs decrease when sample size increases.

<Tables 4.1. and 4.2. HERE.>

5 Applications

This paper is greatly motivated by some empirical findings based on real data sets through which our proposed models outperform many other models present in the literature. For comparison purposes, we consider some well known real data sets from different literature on lifetime distributions viz. 100 uncensored observations on breaking stress of carbon fibers(in Gba)(Nichols and Padgett (2006)), 63 records on strengths of 1.5 cm glass fibres measured at the National Physical Laboratory, England (Smith and Naylor), 101 observations of the fatigue life of 6061-T6 aluminium coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second (Birnbaum and Saunders (1969)) and data on the number of million revolutions before failure for each of the 23 ball bearings (Lawless; 1986, page 228). Moreover, we consider remission times (in months) of a random sample of 128 bladder cancer patients (Lee and Wang (2003)), probably not used in the context of lifetime distribution before. To carry out comparison of the performance of our proposed models, we have considered some alternative models viz. WP ($\boldsymbol{\theta} = (\mu, \lambda, \gamma)$), GEP ($\boldsymbol{\theta} = (\mu, \lambda, \alpha)$) and EW ($\boldsymbol{\theta} = (\lambda, \gamma, \alpha)$). For each data set, we derive the maximum likelihood estimates, Akaike information criterion (AIC), Bayesian information criterion (BIC), Kolmogorov-Smirnov statistic and the corresponding p-value for each of the distributions. Next, we show the results in detail for each of the data sets individually.

First Application:

Here, we consider a real data set from Nichols and Padgett (2006) consisting of 100 uncensored observations on breaking stress of carbon fibers(in Gba). The obtained results are presented in the following table, where the values in the body of the table represents the parameter estimates and related statistics obtained from the fit of each of the four distributions for the breaking stress data.

Distribution	Estimates	Log-likelihood	AIC	BIC	K-S	p-value
GWP(I)	(4.629, 0.479, 6.066, 0.202)	-141.142	288.3	296.1	0.0609	0.8524
GEP(I)	(1.186,5.734,0.893)	-147.432	300.9	308.7	0.1054	0.2162
WP(I)	(10.887, 2.987,0.157)	-141.281	288.6	296.4	0.0634	0.8159
EW	(9.183,0.995,1.087)	-146.550	299.1	306.9	0.1108	0.1715
GWP(II)	(15.808, 0.407, 2.189, 0.110)	-141.317	288.6	296.4	0.0647	0.7969
GEP(II)	(3.473,4.599,1.238)	-142.837	291.7	299.4	0.0872	0.4353
WP(II)	(3.373,1.810,0.509)	-141.544	289.1	296.9	0.0615	0.8442

Second Application:

Another real data set is considered here, which represents remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang (2003). Bladder cancer is a disease in which abnormal cells multiply without control in the bladder. The most common type of bladder cancer recapitulates the normal histology of the urothelium and is known as transitional cell carcinoma. The obtained results are presented in the following table, where the values in the body of the table represents the parameter estimates and related statistics obtained from the fit of each of the four distributions for the remission time data.

Distribution	Estimates	Log-likelihood	AIC	BIC	K-S	p-value
GWP(I)	(3.569, 1.672, 0.888, 0.065)	-409.745	825.5	834.0	0.0373	0.9942
GEP(I)	(1.559,1.458,0.101)	-411.006	828.0	836.6	0.0474	0.9361
WP(I)	(6.614,1.236,0.026)	-410.668	827.3	835.9	0.0471	0.9387
EW	(6.346,0.236,4.841)	-451.085	908.2	916.7	0.1985	0.0084
GWP(II)	(5.413,2.989, 0.407, 2.613)	-409.627	825.2	833.8	0.0293	0.9992
GEP(II)	(6.237,1.218, 0.121)	-413.078	832.2	840.7	0.0725	0.5121
WP(II)	(7.222,0.506,0.852)	-409.785	825.6	834.1	0.0342	0.9983

Other Applications:

For the next three data sets, distributions with higher p-values are considered for brevity. First block of the following table gives results from strength data. Results on the basis of data from Lawless are shown in the second block of the same table whereas the last block demonstrates results from over dispersed data on the fatigue life. Details of these data sets are already discussed earlier.

	Distribution	Estimates	Log-likelihood	AIC	BIC	K-S	p-value
Block1	GWP(I)	(26.815,0.664,8.781)	-15.187	36.4	42.8	0.1508	0.1052
	GWP(II)	(2.784,0.577,5.503)	-12.981	31.9	38.4	0.1145	0.3538
	WP(II)	(2.386,4.484,0.694)	-13.474	32.9	39.4	0.1216	0.2851
	EW	(0.676,7.255,0.582)	-14.676	35.3	41.8	0.1445	0.1302
Block2	GWP(II)	(11.026,0.831,0.0524)	-113.035	232.1	235.5	0.1052	0.9380
	WP(II)	(2.582,6.692,0.772,0.070)	-113.087	232.2	235.6	0.1069	0.9301
Block3	GWP(II)	(1.417,6.596,2.663,0.012)	-455.8907	915.8	924.6	0.0702	0.7017
	WP(II)	(9.516,2.733,0.011)	-455.7802	917.6	925.4	0.0839	0.4751

As we can see, the smallest values of the AIC, BIC and Kolmogorov-Smirnov statistic, and the largest value of p are obtained only for the GWP distribution. It is also observed in the case of data set from Nichols and Padgett (2006) that GWP(I) outperforms all other distributions including WP(I) which was shown to give best fit to the same data as reported earlier by Hemmati et. al. (2011). Even, WP(II), being another new distribution as proposed in this paper performs more or less as competitive as GWP for all the data sets in comparison with the other distributions. While for the case of remission data, GWP(I), GWP(II) and WP(II) perform nearly similarly, it is GWP(II) which performs better for the rest of the data sets. So, it may be easily concluded that this family of unified compounded distributions provide the best fit among the other considered distributions.

7 Conclusion

Acknowledgements:

References

1. Adamidis, K. and Loukas, S. (1998). A life time distribution with decreasing failure rate, *Statistics and Probability Letters*, **39**, pp. 35-42

2. Adamidis, K., Dimitrakopoulou, T. and Loukas, S. (2005). On a generalization of the exponential-geometric distribution, *Statistics and Probability Letters*, **73**, pp. 259-269.

3. Barreto-Souza W. and Cribari-Neto F. (2009). A generalization of the exponential-Poisson distribution, *Statistics and Probability Letters*, **79**, pp. 2493-2500

4. Barreto-Souza W., A.L. de Morais and G.M. Cordeiro (2011). The Weibull-geometric distribution, *Journal of Statistical Computation and Simulation*, **81(5)**, pp. 645-657.

5. Birnbaum, Z.W. and Saunders, S.C. (1969). Estimation for a family of life distributions with applications to fatigue, *Journal of Applied Probability*, **6**, pp. 328-347.

6. Elhai, J.D., Calhoun, P.S. and Ford, J.D. (2008). Statistical procedures for analyzing mental health services data, *Psychiatry Research*, **160**, pp. 129-136

7. Ginebra, J. and Puig, X. (2010). On the measure and the estimation of evenness and diversity, *Computational Statistics and Data Analysis*, **54**, pp. 2187-2201

8. Gupta, R.D., Kundu, D. (1999). Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, **41 (2)**, pp. 173-188.

9. Hemmati, H., Khorram, E. and Rezakhah, S. (2011): A new three-parameter ageing distribution, *Journal of Statistical Planning and Inference*, **141**, pp. 2266-2275.

10. Kus, C. (2007). A new life time distribution, *Computational Statistics and Data Analysis*, **51**, pp. 4497-4509.

11. Lawless, J.F. (1982). Statistical Models and Methods for Lifetime Data, John Wiley and Sons, New York.

12. Lee, E.T. and Wang, J. (2003). Statistical Methods for survival data analysis, John Wiley and Sons, New York.

13. Mahmoudi, E. and Sepahdar, A. (2013). Exponentiated Weibull-Poisson distribution: Model, properties and applicatios, *Journal of Mathematics and Computer in Simulation*, **92**, pp. 76-97

14. Mudholkar, G.S. and Hutson, A.D. (1996). The Exponentiated Weibull Family: Some Properties and a Flood Data Application, *Communications in Statistics-Theory and Methods*, **25(12)**, pp. 3059-3083

15. Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, **42**, pp. 299 - 302.

16. Mudholkar, G.S., Srivastava, D.K. and Friemer, M. (1995). The exponential Weibull family: a re analysis of the bus-motor-failure data, *Technometrics*, **37**, pp. 436-445.

17. Nadarajah, S. and Kotz, S (2006). The beta exponential distribution, *Reliability Engineering and System Safety*, **91**, pp. 689-697

18. Nassar, M.M. and Eissa, F.H. (2003). On the exponentiated Weibull distribution, *Communications in Statistics-Theory and Methods*, **32**, pp. 1317-1333.

19. Nichols, M.D. and Padgett, W.J. (2006). A bootstrap control chart for Weibull percentiles, *Quality and Reliability Engineering International*, **22**, pp. 141-151.

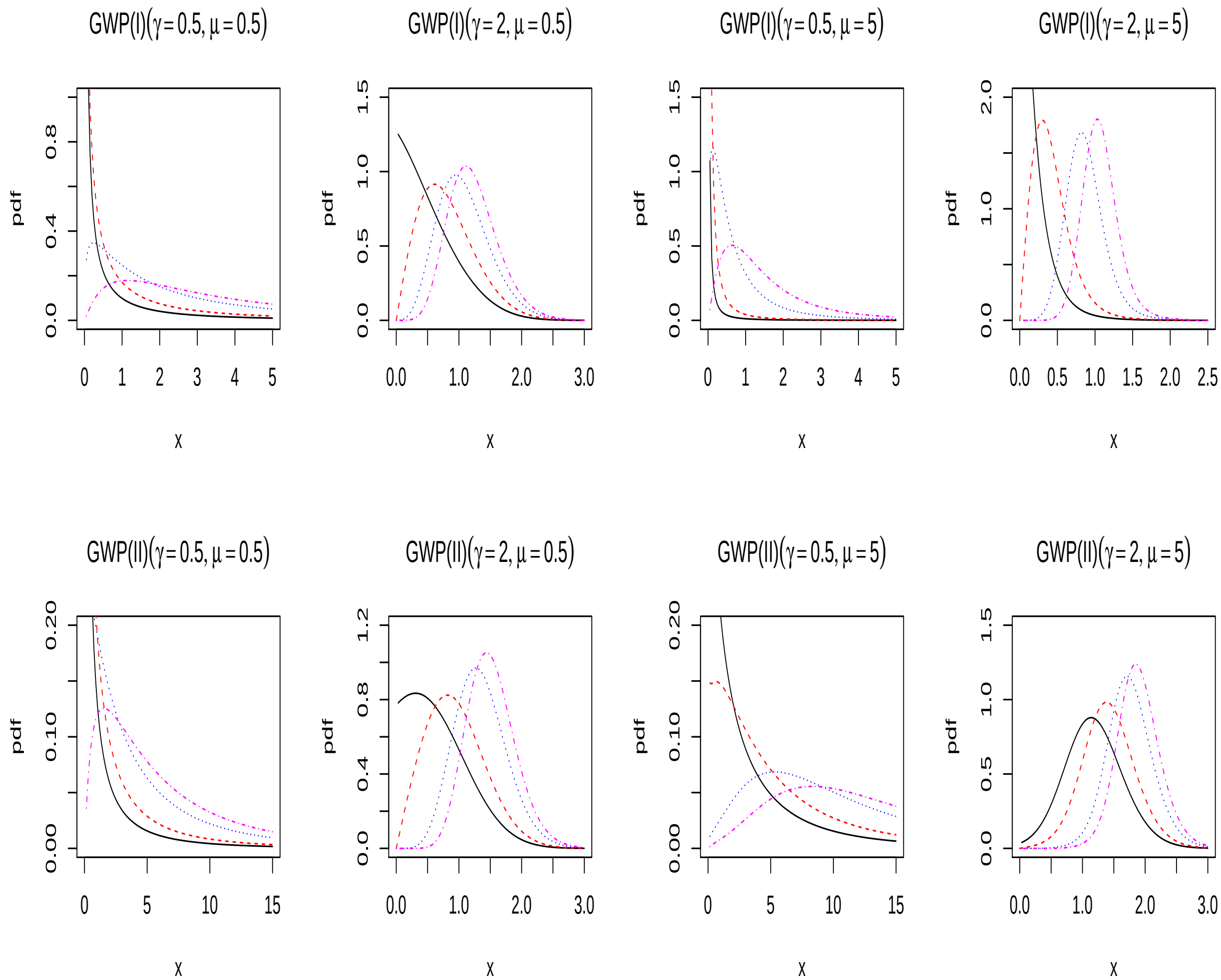
20. R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

21. Rodrigo, B. Silva., Wagner, Barreto-Souza. and Gauss, M. Cordeiro. (2010). A new distribution with decreasing, increasing and upside-down bathtub failure rate, *Computational Statistics and Data Analysis*, **54**, pp. 935-944.

22. Smith, R.L. and Naylor, J.C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution, *Applied Statistics*, **36**, pp. 358-1987.

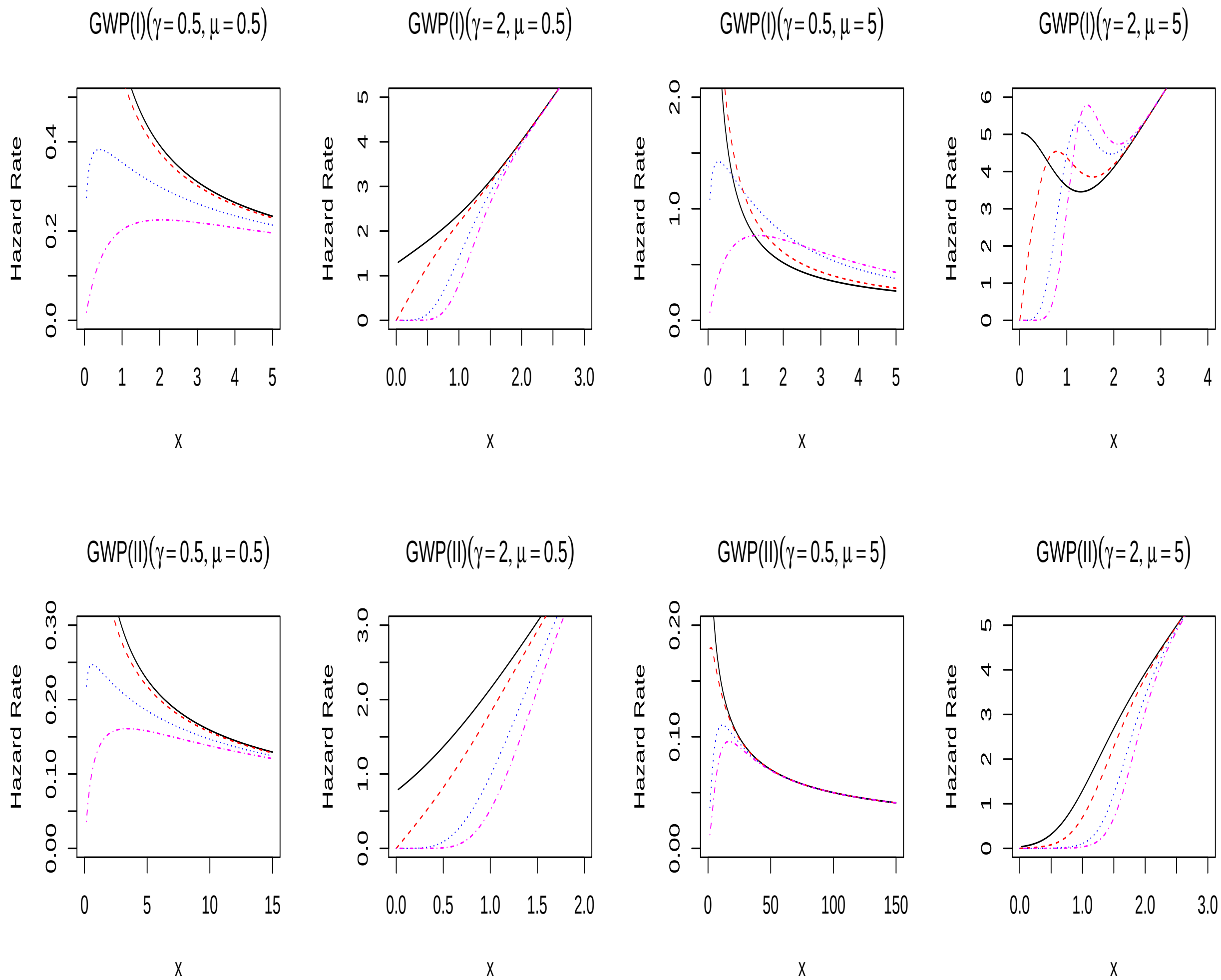
23. Tahmasbi, R. and Rezaei, S. (2008). A two-parameter life time distribution with decreasing failure rate, *Computational Statistics and Data Analysis*, **52**, pp. 3889-3901.

24. Xu, S. and Hu, Z. (2011). Mapping quantitative trait loci using the MCMC procedure in SAS, *Heredity*, **106**, pp. 357-369



$\alpha = 0.5$; $\alpha = 1$; $\alpha = 3$; $\alpha = 5$

Figure 2. : Probability density functions of the GWP(I) and GWP(II) distributions



—— $\alpha = 0.5$; ----- $\alpha = 1$; $\cdots\cdots$ $\alpha = 3$; $\cdot - \cdot -$ $\alpha = 5$

Figure 3. : Hazard functions of the GWP(I) and GWP(II) distributions

TABLE 4.1. Comparison of MLEs of parameters of GWP(I) Distribution Based on Monte Carlo Simulation

$\mu, \alpha, \gamma, \lambda$	n	MLEs	SD of MLEs
5, 0.5, 0.5, 0.5	50	4.9304; 0.4869; 0.4884; 0.4869	0.001492475; 0.000279887; 0.00039050479; 0.0002803135
	100	4.9042; 0.4901; 0.4902; 0.4903	0.000633463; 0.000207306; 0.00020730691; 0.0002073631
	500	4.9559; 0.4955; 0.4955; 0.4955	0.0001249365; 9.466046e-05; 9.466145e-05; 9.468005e-05
5, 2, 0.5, 0.5	50	4.9766; 1.9582; 0.4903; 0.5295	0.0006979332; 0.0008281978; 0.0002095827; 0.000653262
	100	4.9772; 1.9766; 0.4940; 0.5300	0.0004911315; 0.0004693609; 0.0001221766; 0.000639997
	500	4.9915; 1.9951; 0.4979; 0.5195	0.0001790427; 0.0001014522; 6.853689e-05; 0.000420816
5, 2, 2, 0.5	50	4.9785; 1.9703; 2.0118; 0.5112	0.0012358932; 0.001557318; 0.0009062982; 0.0003839686
	100	4.9936; 1.9929; 2.0042; 0.5107	0.0006781852; 0.0006859042; 0.0004574815; 0.000259358
	500	4.9978; 1.9989; 2.0042; 0.5052	4.661443e-05; 2.213895e-05; 8.957011e-05; 0.0001106684
5, 2, 2, 2	50	4.9764; 1.9622; 1.9612; 2.0310	0.0006020727; 0.0007454828; 0.0007543795; 0.000748331
	100	4.9768; 1.9794; 1.9777; 2.0305	0.0004708203; 0.0004017499; 0.0004341177; 0.000669263
	500	4.9907; 1.9946; 1.9929; 2.0180	0.0001871316; 0.0001097178; 0.0001794691; 0.000370626
5, 0.5, 2, 2	50	4.9299; 0.4867; 1.9391; 1.9396	0.0014381562; 0.0002617809; 0.0012350544; 0.001236148
	100	4.9038; 0.4896; 1.9522; 1.9534	0.000402392; 0.0002070967; 0.0009686258; 0.0009651375
	500	4.9557; 0.4951; 1.9778; 1.9785	0.000134965; 9.531116e-05; 0.0004455848; 0.0004437698
5, 0.5, 0.5, 2	50	4.9326; 0.4874; 0.4874; 1.9464	0.003378481; 0.0002589772; 0.0002578994; 0.001108993
	100	4.9027; 0.4899; 0.4899; 1.9535	0.002122958; 0.0002133881; 0.0002133881; 0.001015551
	500	4.9537; 0.4952; 0.4952; 1.9779	0.000987379; 9.857832e-05; 9.857825e-05; 0.0004689608

TABLE 4.2. Comparison of MLEs of parameters of GWP(II) Distribution Based on Monte Carlo Simulation

$\mu, \alpha, \gamma, \lambda$	n	MLEs	SD of MLEs
5, 0.5, 0.5, 0.5	50	5.0419; 0.4829; 0.4830; 0.4829	0.000941907; 0.000377071; 0.000376683; 0.000376864
	100	5.0320; 0.4869; 0.4870; 0.4870	0.000688391; 0.000275845; 0.000275379; 0.000243517
	500	5.0618; 0.4939; 0.4939; 0.4939	0.000133042; 0.000127179; 0.000127168; 0.000127178
5, 2, 0.5, 0.5	50	5.0323; 1.9810; 0.4928; 0.4869	0.000828137; 0.000543329; 0.000368399; 0.000340591
	100	5.0197; 1.9886; 0.4928; 0.4921	0.0004175504; 0.000290250; 0.00018884; 0.000191740
	500	5.0365; 1.9837; 0.4966; 0.4966	0.000186694; 0.00014888; 7.339843e-05; 7.338202e-05
5, 2, 2, 0.5	50	5.0283; 1.9869; 2.0037; 0.4886	0.000567146; 0.000261966; 7.347964e-05; 0.00026858
	100	5.0091; 1.9944; 2.0007; 0.4964	0.000188039; 0.000114927; 2.44261e-05; 7.52159e-05
	500	5.0489; 1.9783; 2.0092; 0.4954	0.000053631; 0.000013631; 1.7326e-05; 5.928426e-05
5, 2, 2, 2	50	5.0131; 1.9934; 1.9766; 1.9750	0.00032259; 0.000162217; 0.000595815; 0.000612326
	100	5.0127; 1.9926; 1.9791; 1.9757	0.000190676; 0.000102971; 0.000316516; 0.000361372
	500	5.0166; 1.9926; 1.9926; 1.9926	0.00005544; 0.000087631; 0.000157595; 0.000157532
5, 0.5, 2, 2	50	5.0224; 0.4904; 1.9633; 1.9570	0.00048816; 0.000210494; 0.000837212; 0.000931920
	100	5.0169; 0.4928; 1.9731; 1.9676	0.000357676; 0.000148830; 0.000584700; 0.00068110
	500	5.0354; 0.4963; 1.9838; 1.9834	0.00011589; 6.974112e-05; 0.000317487; 0.000318759
5, 0.5, 0.5, 2	50	5.0474; 0.4809; 0.4814; 1.9709	0.001041872; 0.000416765; 0.00035215; 0.000654382
	100	5.0367; 0.4852; 0.4852; 1.9729	0.000778875; 0.000311278; 0.000311278; 0.00047476
	500	5.0580; 0.4937; 0.4937; 1.9739	0.000126836; 0.000129405; 0.000129405; 0.000261608

Indian Institute of Management Kozhikode

Type of Document: (**Working Paper/Case/ Teaching Note, etc.)** **Working Paper**	*Ref. No.: (to be filled by RCP office)* ***IIMK/WPS/148/QM&OM /2014/06***
Title: **COMPOUNDED GENERALIZED WEIBULL DISTRIBUTIONS - A UNIFIED APPROACH**	
Author(s):	*Institution(s)*
Shovan Chowdhury	Indian Institute of Management Kozhikode, IIMK Campus PO, Kozhikode- 673570, Kerala email: shovanc@iimk.ac.in
Subject Areas : ***Statistics***	*Subject Classification Codes, if any:*
Supporting Agencies, if any:	*Research Grant/Project No.(s):* *SGRP/2014/72*
Supplementary Information, if any:	*Date of Issue: (to be filled by RCP office)* **March 2014**
Full text or only abstract to be uploaded on website: (please choose one)	*Number of Pages:* ***18***

Abstract:

A unified approach is proposed in this paper to study a family of lifetime distributions of a system consisting of random number of components in series and in parallel. While the lifetimes of the components are assumed to follow generalized (exponentiated) Weibull distribution, a zero-truncated Poisson is assigned to model the random number of components in the system. The resulting family of compounded distributions describes several well-known distributions as well as some new models with some of their statistical and reliability properties. Various ageing classes of life distributions including increasing, decreasing, bath-tub, upside-down-bathtub and roller coaster shaped failure rates are covered by the family of compounded distributions. The simplest algorithm for maximum likelihood method of estimation of the model parameters is discussed. Some numerical results are obtained via Monte-Carlo Simulation. The asymptotic variance-covariance matrices of the estimators are also obtained. Five different real data sets are used to validate the distributions and the results demonstrate that the family of distributions can be considered as a suitable model under several real situations.

Key Words/Phrases: Unified approach, Compounding, Generalized Weibull Distribution; Hazard Function; ML Estimation; Zero-Truncated Poisson Distribution

Referencing Style Followed:

Research, Conference And Publication Office
Indian Institute Of Management Kozhikode
IIMK Campus P.O., Kozhikode 673 570
Kerala, India
Telephone +91 495 2809 238
E-mail rcp@iimk.ac.in
website www.iimk.ac.in