REVSTAT – Statistical Journal
Volume 14, Number 1, February 2016, 29–48

GOODNESS OF FIT TESTS AND POWER COMPARISONS FOR WEIGHTED GAMMA DISTRIBUTION

Authors: Neetu Singla
– Department of Statistics, Panjab University, Chandigarh-160014, India
neetu.singla12@gmail.com

Kanchan Jain
– Department of Statistics, Panjab University, Chandigarh-160014, India
jaink14@gmail.com

Suresh Kumar Sharma
– Department of Statistics, Panjab University, Chandigarh-160014, India
ssharma643@yahoo.co.in

Received: December 2013 Revised: July 2014 Accepted: September 2014

Abstract:

- In this paper, a weighted version of Gamma distribution known as Weighted Gamma (WG) distribution has been considered. Various tests of goodness of fit viz Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling have been applied to this family. Monte Carlo simulations have been carried out for power calculations. The powers of these tests have been compared which helps in ranking of these goodness of fit tests.

Key-Words:

- *Weighted Gamma; Kolmogorov–Smirnov; Cramér–von Mises and Anderson–Darling; power.*

AMS Subject Classification:

- 62E15, 62F03.

1. INTRODUCTION

Goodness of fit tests (GOFTs) validate the closeness of the theoretical distribution function to the empirical distribution function. They are also known as empirical distribution function tests. These tests determine how well the distribution under study fits to a data set. They are used to test simple hypothesis which completely specifies the model and also the composite hypotheses where only the name of the model/distribution is stated but not its parameters. In the latter case, the parameters are estimated from the data. The common GOFTs are Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling.

In literature, many authors have studied the goodness of fit tests. Nikulin [21,22] studied Chi-squared test for continuous distributions. Rao and Robson [25] studied Chi-squared statistic for exponential family. Power of a series of goodness of fit tests for simple and complex hypotheses have been analyzed by Lemeshko *et al.* [14,15]. Lemeshko *et al.* [16] analyzed the goodness of fit test for Inverse Gaussian family. Goodness of fit tests for testing composite hypotheses, using maximum likelihood estimators (MLEs) of double exponential distribution, have been given in Lemeshko and Lemeshko [17].

The idea of weighted distributions was conceptualized by Fisher [6] and studied by Rao [24] in a unified manner who pointed out that in many situations, the recorded observations cannot be considered as a random sample from the original distribution. This can be due to one or the other reason viz non-observability of some events, damage caused to original observations and adoption of unequal probability sampling. In observational studies for human, wild-life, insect, plant or fish population, it is not possible to select sampling units with equal probabilities. In such cases, there are no well-defined sampling frames and recorded observations are biased. These observations do not follow the original distribution and hence their modelling uses the theory of weighted distributions. It is, therefore, important to study the stochastic orderings and ageing properties of the weighted random variables with respect to the original random variables.

For a non-negative random variable X with pdf $f(x)$, the weighted random variable X^w has the pdf given by

$$f^w(x) = \frac{w(x)\, f(x)}{E\big[w(X)\big]}\ , \tag{1.1}$$

where $w(x)$ is a non-negative weight function such that $E[w(X)]$ is non-zero and finite. The distribution of X^w is called the weighted distribution corresponding to X.

The weighted distribution with $w(x) = x$ is called the length-biased (size-biased) distribution which finds various applications in biomedical areas such as early detection of a disease. Rao [24] used this distribution in the study of human families and wild-life populations. Various other important weighted distributions and their properties have been discussed by Mahfoud and Patil [19], Jain *et al.* [12], Gupta and Kirmani [10], Nanda and Jain [20], Patil [23] and Gupta and Kundu [11].

A brief discussion of weighted version of Gamma distribution labelled as Weighted Gamma (WG) distribution is provided in Section 2. This distribution has been introduced by Jain *et al.* [13]. The Weighted Gamma (WG) distribution has Weighted exponential, Gamma and Exponential distributions as its submodels. This distribution can also be interpreted as a hidden upper truncation model as in case of skew-normal distribution (Arnold and Beaver [2]). The pdf of WG distribution is also expressible as a linear combination of two Gamma pdfs. This distribution accommodates increasing and upside-down bathtub shaped failure rate function and hence has wider applicability in reliability and survival analysis.

The motive of this study is to carry out goodness of fit tests viz Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling and to compare their powers for Weighted Gamma and some competing distributions namely Weighted Weibull, Weighted Exponential and Gamma distributions. Using the calculated powers of these goodness of fit tests, we can determine the sample size at which these various closely related distributions can be distinguished from each other.

The paper is organized as follows. In Section 2, we provide a brief description of Weighted Gamma (WG) distribution. Various goodness of fit tests have been described in Section 3. Testing of simple and composite hypotheses for WG versus Weighted Weibull (WW), Weighted Exponential (WE) and Gamma is presented in Section 4. This section also consists of results and power studies based on simulations and real data set analysis. Section 5 includes the concluding remarks.

2. WEIGHTED GAMMA DISTRIBUTION

The random variable X is said to follow Weighted Gamma distribution with scale parameter λ and shape parameters α and β if the probability density function (pdf) of X is given by

$$f_X(x;\alpha,\beta,\lambda) = k\,\frac{(1-\mathrm{e}^{-\alpha\lambda x})\,\lambda^\beta\,x^{\beta-1}\,\mathrm{e}^{-\lambda x}}{\Gamma(\beta)}\,, \qquad x>0\,,\quad \alpha,\beta,\lambda>0\,, \tag{2.1}$$

where $k^{-1} = 1-\left(\frac{1}{1+\alpha}\right)^\beta$.

If X is a random variable with pdf given in (2.1), we use the notation $X \sim WG(\alpha,\beta,\lambda)$.

The distribution function of X can be written as

$$F(x) = \left[\frac{(1+\alpha)^\beta}{(1+\alpha)^\beta - 1}\right]\left[G(x;\beta,\lambda) - \frac{1}{(1+\alpha)^\beta}\,G(x;\beta,\lambda(1+\alpha))\right], \tag{2.2}$$

where $G(x;a,b) = \dfrac{b^a \int_0^x \mathrm{e}^{-bt} t^{a-1} dt}{\Gamma(a)}$ is the cumulative distribution function of Gamma distribution with shape parameter a and scale parameter b.

Remark 2.1. (2.1) is the weighted version of the Gamma pdf with weight function

$$w(x) = 1 - \mathrm{e}^{-\alpha\lambda x}, \qquad \alpha, \lambda > 0 .$$

The choice of the weight function has been made so that Weighted Exponential (Gupta and Kundu [11]), Gamma and Exponential distributions can be obtained as special cases of WG distribution for particular values of parameters. The special cases are:

- Weighted Exponential (WE) distribution obtained by putting $\beta = 1$,
- the Gamma distribution when $\alpha \to \infty$.
- For $\alpha \to \infty$ and $\beta = 1$, Exponential distribution can be obtained.

Suppose X follows WG distribution and let $\boldsymbol{\theta} = (\alpha, \beta, \lambda)^T$ be the parameter vector. The log likelihood based on the observed sample $(x_1, x_2, ..., x_n)$ is

$$\begin{aligned} l &= l(\alpha,\beta,\lambda) \\ &= n\Big\{\log(1+\alpha)^\beta - \log\big\{(1+\alpha)^\beta - 1\big\}\Big\} + \sum_{i=1}^n \log\big(1 - \mathrm{e}^{-\alpha\lambda x_i}\big) + n\,\beta\log\lambda \\ &\quad + (\beta-1)\sum_{i=1}^n \log x_i - \lambda\sum_{i=1}^n x_i - n\log\big\{\Gamma(\beta)\big\} . \end{aligned} \tag{2.3}$$

The first derivative of the log likelihood function is called Fisher's score function and is written as

$$\boldsymbol{u}(\boldsymbol{\theta}) = \frac{\partial l}{\partial \boldsymbol{\theta}} .$$

Score is a vector of first partial derivatives, one for each element of $\boldsymbol{\theta}$. If the log likelihood is concave, then MLEs can be obtained by solving the system of equations

$$\boldsymbol{u}(\boldsymbol{\theta}) = \mathbf{0} ,$$

where elements of $\boldsymbol{u}(\boldsymbol{\theta})$ are given by

$$\frac{\partial l}{\partial \alpha} = -n\,\beta(1+\alpha)^{\beta-1}\left\{\frac{1}{\big\{(1+\alpha)^\beta - 1\big\}\,(1+\alpha)^\beta}\right\} + \lambda\sum_{i=1}^n \frac{x_i\,\mathrm{e}^{-\alpha\lambda x_i}}{1 - \mathrm{e}^{-\alpha\lambda x_i}} , \tag{2.4}$$

$$\frac{\partial l}{\partial \beta} = n\log(1+\alpha) - \frac{n(1+\alpha)^\beta \log(1+\alpha)}{\{(1+\alpha)^\beta - 1\}} + n\log\lambda + \sum_{i=1}^{n} \log x_i - n\,\psi(\beta)\,, \tag{2.5}$$

$$\frac{\partial l}{\partial \lambda} = \alpha \sum_{i=1}^{n} \frac{x_i\, \mathrm{e}^{-\alpha\lambda x_i}}{1-\mathrm{e}^{-\alpha\lambda x_i}} + \frac{n\,\beta}{\lambda} - \sum_{i=1}^{n} x_i\,, \tag{2.6}$$

where $\psi(\cdot)$ denotes the digamma function, the logarithmic derivative of the gamma function.

As these equations are difficult to be solved, Newton–Raphson method can be used for finding ML estimates. Using this method, the score function is evaluated at the MLE $\hat{\boldsymbol{\theta}}$ around an initial value $\boldsymbol{\theta}_0$, using a first order Taylor series which gives

$$\boldsymbol{u}(\hat{\boldsymbol{\theta}}) \approx \boldsymbol{u}(\boldsymbol{\theta}_0) + \frac{\partial \boldsymbol{u}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\,. \tag{2.7}$$

Equating (2.7) to zero and solving for $\hat{\boldsymbol{\theta}}$ leads to first approximation:

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 - \boldsymbol{H}^{-1}(\boldsymbol{\theta}_0)\, \boldsymbol{u}(\boldsymbol{\theta}_0)\,, \tag{2.8}$$

where

$$\boldsymbol{H}(\boldsymbol{\theta}) = \frac{\partial^2 l}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}'} = \frac{\partial \boldsymbol{u}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

denotes the Hessian matrix.

Given a trial value, (2.8) is employed for obtaining an improved estimate and the process is repeated until the differences between successive estimates are sufficiently close to zero. The estimates obtained are considered as maxima if the Hessian matrix is negative definite, that is, all its eigenvalues are negative.

As sometimes, it is computationally difficult to invert the Hessian matrix, hence we use the quasi Newton method in R for finding the ML estimates as this method usually generates an estimate of H^{-1} directly. The results have been included in Table 2 of Section 4.

3. GOODNESS OF FIT TESTS

For a random sample of size n, let $x_{(1)}, ..., x_{(n)}$ be ordered observations. The empirical distribution function (edf) $F_n(x)$ is a step function with a step of height $\frac{1}{n}$ at each ordered sample observation. Empirical Distribution Function (EDF) tests measuring the distance between the edf and theoretical cdf are described by Dufour *et al.* [6]. Arshad *et al.* [3] and Seier [26] claimed that the widely used EDF tests are Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling tests.

For a random variable X, we let $F(x)$ to be the theoretical cumulative distribution function (cdf). $F(x,\theta)$ denotes the cdf for a particular distribution with parameter θ. The focus shall be on testing the following types of null hypotheses:

- **Simple null hypothesis:**

  H_0: $F(x) = F(x,\theta)$,

  where the form of $F(x,\theta)$ is completely specified;

- **Composite null hypothesis:**

  H_0: $F(x) \in \{F(x,\theta),\ \theta \in \Theta\}$,

  where Θ is the domain of unknown parameter θ which is replaced by its estimator.

We will use the tests explained in the subsequent discussion.

Kolmogorov–Smirnov Test:

This test is based upon the largest vertical distance between empirical distribution function $F_n(x)$ and theoretical distribution function $F(x,\theta)$. The statistic is

$$D_n = \sup_{|n|<\infty} \big|F_n(x) - F(x,\theta)\big| , \qquad \theta \in \Theta . \tag{3.1}$$

If the value of KS statistic is greater than critical point, we reject the null hypothesis (Gibbons and Chakraborti [9]).

Cramér–von Mises and Anderson–Darling statistics belong to the class of quadratic EDF statistics (Stephens [28]) defined as

$$n \int_{-\infty}^{\infty} \big(F_n(x) - F(x)\big)^2\, w(x)\, dF(x) , \tag{3.2}$$

where $w(x)$ is a weighting function.

Cramér–von Mises Test:

For $w(x) = 1$, (3.2) gives n times the Cramér–von Mises (CVM) statistic. This statistic can be computed using the sum of squared differences between the empirical distribution function (EDF) and theoretical CDF (Anderson and Darling [1]) and is defined as

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(x_i,\theta) - \frac{2i-1}{2n}\right)^2 . \tag{3.3}$$

If the value of CVM test statistic is greater than the critical point, we reject the null hypothesis. According to Conover [5], CVM is more powerful than KS test because it uses more sample data.

Anderson–Darling test:

It is a modification of the CVM Test. It gives more weightage to the tails of the distribution (Farrel and Stewart [7]).

By taking $w(x) = [F(x)\,(1-F(x))]^{-1}$ in (3.2), Anderson–Darling (AD) test statistic (Anderson and Darling [1]) is obtained as

$$n \int_{-\infty}^{\infty} \frac{\big(F_n(x) - F(x)\big)^2}{\big[F(x)\,\big(1-F(x)\big)\big]}\, dF(x)\ .$$

It can also be written as

$$\text{(3.4)} \qquad AD \,=\, -n - 2\sum_{i=1}^{n} \left\{ \frac{2i-1}{2n} \ln F(x_i,\theta) + \left(1 - \frac{2i-1}{2n}\right) \ln\!\big(1 - F(x_i,\theta)\big) \right\}$$

(Lewis [18]).

If the value of AD test statistic is greater than critical point, we reject the null hypothesis.

The critical points (C.P.) of these tests have been calculated by generating random samples from the distribution under null hypothesis, calculating value of test statistics and arranging values of test statistic in increasing order. $(1-\alpha)^{\text{th}}$ largest order test statistic gives the critical point corresponding to α level of significance. These values have been calculated for sample sizes $n = 50, 100, 200, 500, 1000$ and 2000 at $\alpha = .20, .15, .10$ and $.05$ and are shown in Table 1.

Table 1: Critical points for Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling tests.

n	**Kolmogorov–Smirnov**				**Cramér–von Mises**				**Anderson–Darling**			
	Level of significance				Level of significance				Level of significance			
	.20	.15	.10	.05	.20	.15	.10	.05	.20	.15	.10	.05
50	.151	.161	.172	.192	.241	.281	.344	.455	1.427	1.619	1.900	2.422
100	.107	.114	.122	.136	.244	.286	.361	.475	1.388	1.603	1.909	2.412
200	.076	.081	.086	.096	.237	.282	.339	.453	1.39	1.59	1.92	2.49
500	.048	.051	.055	.061	.232	.275	.334	.444	1.405	1.609	1.932	2.500
800	.038	.040	.043	.048	.241	.286	.347	.469	1.410	1.617	1.904	2.399
1000	.034	.036	.039	.043	.245	.286	.347	.449	1.423	1.638	1.945	2.514
2000	.024	.026	.027	.030	.241	.287	.349	.476	1.395	1.588	1.914	2.438

4. APPLICATION

4.1. Simulations for estimation and applying GOFTs

Weighted Exponential and Gamma distributions are considered as competing distributions for WG. The Weighted Weibull (WW) distribution with three parameters α, β and λ (Shahbaz *et al.* [27]) has also been considered as one of the competing distributions for WG. The cdf and pdf of WW are

$$F(x;\alpha,\beta,\lambda) = \frac{(1+\alpha)\left[1-\mathrm{e}^{-\lambda x^{\beta}} - \frac{\left(1-\mathrm{e}^{-(1+\alpha)\lambda x^{\beta}}\right)}{1+\alpha}\right]}{\alpha}$$

and

$$f(x;\alpha,\beta,\lambda) = \frac{(1+\alpha)\,\lambda\,\beta\,x^{\beta-1}\,\mathrm{e}^{-\lambda x^{\beta}}\left(1-\mathrm{e}^{-\alpha\lambda x^{\beta}}\right)}{\alpha}\ .$$

A random sample of size 200 from Weighted Gamma (WG) distribution with parameters $\alpha=5$, $\beta=2.5$ and $\lambda=2$ is generated. The empirical cumulative distribution function (ecdf) based on the data and the theoretical cdf of WG distribution are plotted in Figure 1. This figure depicts that ecdf and exact cdf of WG distribution are quite close to each other.

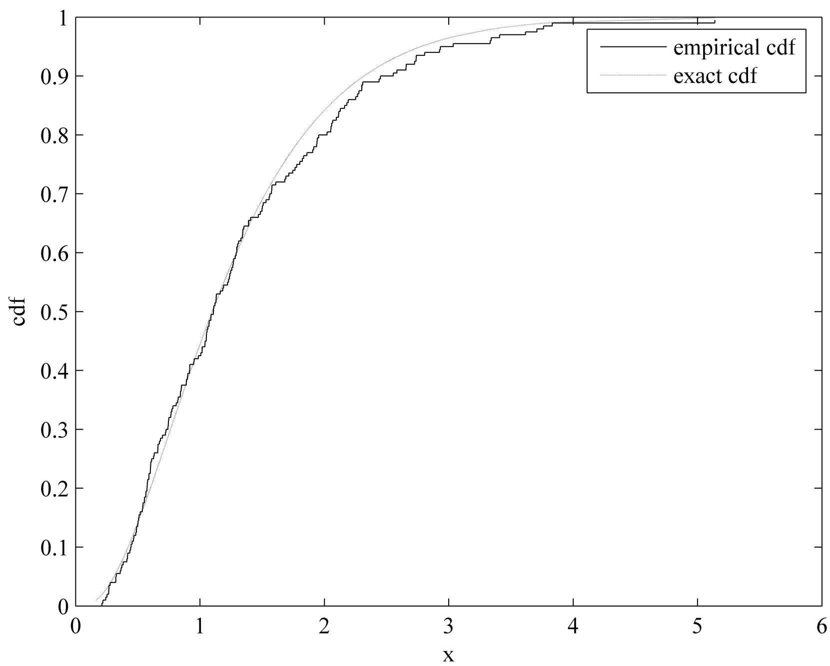

Figure 1: Plots for ecdf and exact cdf of WG distribution.

For the generated data set, maximum likelihood estimates (MLEs) of parameters of WG, WW, WE and Gamma distributions and the corresponding AIC and AICc values are given in Table 2. In quasi Newton algorithm in R, the Broyden–Fletcher–Goldfarb Shanno (BFGS) method has been used by applying optim routine. Hessian matrices have been checked for all the distributions and found to be negative definite as all the eigenvalues of each Hessian matrix come out to be negative. This implies that the estimates obtained are maximum likelihood estimates.

Table 2: Estimates of the parameters and AIC and AICc values for different distributions.

Distribution	MLE			AIC	AICc
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\lambda}$		
WG	3.0230	2.1688	1.7399	432.9696	433.092
WW	53.0360	1.6025	0.5560	437.9989	438.121
WE	0.0006	1	1.5444	438.4732	438.534
Gamma	—	2.4871	1.9212	433.4271	433.488

From the above table, we can conclude that:

a) Since AIC and AICc values are the lowest for WG distribution, it can be considered to be the best fit.

b) Since AIC values of WG and Gamma distributions are close, hence a large sample size shall be required to distinguish between WG and Gamma distributions.

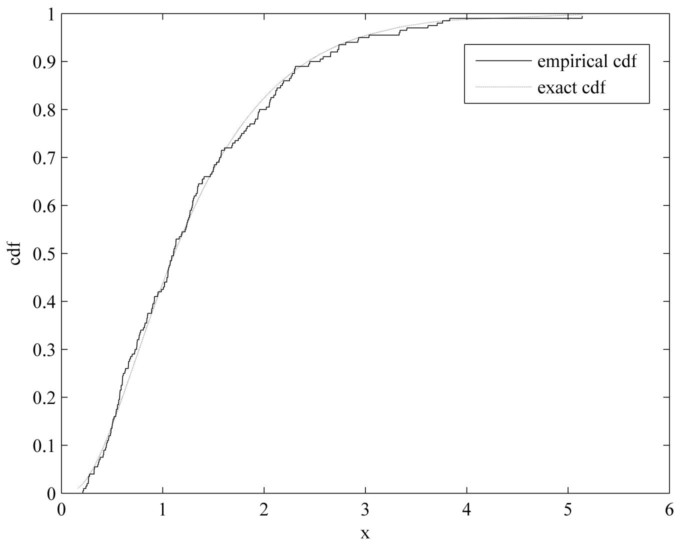

Figure 2: Plots of ecdf and estimated cdf of WG distribution.

Figure 2 displays the plots of empirical cdf and theoretical cdf using estimates of parameters of WG distribution for generated data set.

Weighted Weibull (WW), Weighted Exponential (WE) and Gamma distributions are taken up as the competing distributions for WG distribution. The estimates of parameters for all these distributions are found for generated data set.

To check whether the generated data set fits well to WG distribution (with assumed and estimated parameters), WW, WE and Gamma distributions, the simple and composite hypotheses have been tested in the sequel.

Testing of simple hypothesis:

The aim is to test the simple hypothesis

H_{01}: WG (α=5, β=2.5, λ=2) distribution fits well to the generated data set versus

H_{11}: It does not fit well.

The values of KS, CVM and AD test statistics and critical points (extracted from Table 1) are given in Table 3.

Table 3: Values of test statistic and critical points for testing H_{01} versus H_{11}.

Test	Statistic values	C.P. at 0.05 level of significance
Kolmogorov–Smirnov	0.0447	.096
Cramér–von Mises	0.0779	.453
Anderson–Darling	0.5654	2.49

It is observed that for all the tests, the null hypothesis is not rejected at 0.05 level of significance implying that WG distribution fits well to the generated data set under all testing procedures.

Testing of composite hypotheses:

We consider testing of composite hypotheses

H_{02}: WG ($\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\lambda}$) distribution fits the generated data well versus

H_{12}: It does not fit the data well,

where $\widehat{\alpha} = 3.0230$, $\widehat{\beta} = 2.1688$ and $\widehat{\lambda} = 1.7399$.

The following table gives calculated values of test statistics.

Table 4: Values of test statistic and critical points for testing H_{02} versus H_{12}.

Test	Statistic values	C.P. at 0.05 level of significance
Kolmogorov–Smirnov	0.0422	.096
Cramér–von Mises	0.0546	.453
Anderson–Darling	0.3480	2.49

It is observed that values of test statistics for KS, CVM and AD are less than C.Ps. at 0.05 level of significance for $n = 200$. This means that WG distribution with estimated parameters fits the data well.

Next, we consider testing of composite hypotheses:

i) H_{03}: Generated data set is fitted well by WW ($\widehat{\alpha} = 53.0360$, $\widehat{\beta} = 1.6025$, $\widehat{\lambda} = .5566$)
versus
H_{13}: compliment to H_{03}, that is, data is not fitted well.

ii) H_{04}: WE ($\widehat{\alpha} = .0006$, $\widehat{\lambda} = 1.5444$) distribution fits the generated data well
versus
H_{14}: WE ($\widehat{\alpha} = .0006$, $\widehat{\lambda} = 1.5444$) distribution does not fit the data well.

iii) H_{05}: Gamma ($\widehat{\beta}$=2.4871, $\widehat{\lambda}$=1.9212) distribution fits the generated data well
versus
H_{15}: Gamma distribution with estimated parameters does not fit the data well.

Tables 5–7 display the values of test statistics for KS, CVM and AD tests and corresponding critical points at 0.05 level of significance for testing the composite hypotheses.

Table 5: Values of test statistic and critical points for testing H_{03} versus H_{13}.

Test	Statistic values	C.P. at 0.05 level of significance
Kolmogorov–Smirnov	0.0576	.096
Cramér–von Mises	0.1489	.453
Anderson–Darling	0.8541	2.49

Table 6: Values of test statistic and critical points for testing H_{04} versus H_{14}.

Test	Statistic values	C.P. at 0.05 level of significance
Kolmogorov–Smirnov	0.0540	.096
Cramér–von Mises	0.0786	.453
Anderson–Darling	0.8263	2.49

Table 7: Values of test statistic and critical points for testing H_{05} versus H_{15}.

Test	Statistic values	C.P. at 0.05 level of significance
Kolmogorov–Smirnov	0.0464	.096
Cramér–von Mises	0.0822	.453
Anderson–Darling	0.5203	2.49

The values in Tables 3–7 help us to conclude that:

a) All the distributions fit well to the given data set at 0.05 level of significance because the values of test statistics are less than critical points.

b) WG distribution fits best to the data set because the values of test statistics are lowest in case of WG distribution.

In the next subsection, we find the powers of goodness of fit tests viz KS, CVM and AD for comparing WG distribution with WW, WE and Gamma distributions. The values of power for GOFTs help us in differentiating among the distributions under consideration and also in determining the optimal sample size for differentiation.

4.1.1. Powers of goodness of fit tests for WG

To differentiate among different distributions, we carry out the power study for testing of hypotheses about belonging of the sample to WG distribution, considering WW, WE and Gamma distributions as competing distributions.

For power analysis, we use the technique of Bootstrapping to generate the samples. We generate 10,000 copies of random sample under alternative hypotheses. The values of the test statistics have been calculated using estimates of parameters for different distributions. The power analysis has been carried out for sample sizes $n = 50, 100, 200, 500, 1000, 2000$ at .20, .15, .10, .05 levels of significance.

Using estimated parameters, Tables 8–10 give the power of KS, CVM and AD tests for testing about belonging of the samples to WG distribution against that sample is from WW, WE and Gamma distributions.

Power of Anderson–Darling test is more than those of Cramér–von Mises and Kolmogorov–Smirnov tests in all cases. Hence, AD is the most powerful and KS is the least powerful test.

From Tables 8–10, it is observed that at 0.10 level of significance to obtain low probability of type II error (less than or equal to 0.1):

a) A sample size greater than or equal to 2000 is required to differentiate WG distribution from WW distribution, since the power of AD test is .9630 implying that probability of type II error is .0370;

b) A sample of at least 2000 observations is required to distinguish WG distribution from WE and Gamma distributions.

Table 8: Power of tests for testing goodness of fit of WG versus WW with estimated parameters.

Level of significance	Sample size					
	$n=50$	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling						
.20	.4993	.5834	.6253	.7351	.9005	.9899
.15	.4321	.4995	.5535	.7032	.8622	.9869
.10	**.2557**	**.3072**	**.4993**	**.5938**	**.8123**	**.9630**
.05	.1540	.2505	.3857	.4887	.7945	.8756
Power of Cramér–von Mises						
.20	.4286	.4993	.5547	.5790	.8750	.9666
.15	.4274	.4740	.5038	.5732	.8443	.9311
.10	.1571	.2946	.4586	.5606	.7801	.8959
.05	.1243	.2815	.2783	.4043	.7278	.8322
Power of Kolmogorov–Smirnov						
.20	.4078	.4551	.5013	.5485	.8539	.9521
.15	.3451	.4738	.5008	.5308	.8123	.9222
.10	.1526	.2574	.4165	.5243	.6959	.7898
.05	.1182	.2299	.2439	.2858	.5557	.6345

Table 9: Power of tests for testing goodness of fit of WG versus WE with estimated parameters.

Level of significance	Sample size					
	$n=50$	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling						
.20	.5947	.6543	.7686	.8504	.9404	.9969
.15	.4928	.5689	.7038	.8153	.8935	.9851
.10	**.3853**	**.4537**	**.6583**	**.7328**	**.8589**	**.9708**
.05	.1549	.3839	.5841	.6685	.8040	.9146
Power of Cramér–von Mises						
.20	.4899	.5251	.6493	.8039	.8991	.9784
.15	.4518	.5103	.5552	.6751	.8599	.9485
.10	.2538	.3840	.4993	.5998	.8328	.9113
.05	.1959	.3014	.3547	.4853	.7943	.8993
Power of Kolmogorov–Smirnov						
.20	.4286	.4865	.5878	.6438	.8689	.9663
.15	.3945	.4793	.5584	.5991	.8402	.9365
.10	.2090	.2940	.4738	.5344	.7556	.8734
.05	.1547	.2591	.2973	.3905	.6938	.7488

Table 10: Power of tests for testing goodness of fit of WG versus Gamma with estimated parameters.

Level of significance	Sample size					
	$n=50$	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling						
.20	.4037	.4270	.4408	.6728	.8875	.8993
.15	.3018	.3363	.3401	.5556	.8543	.8775
.10	**.2134**	**.2627**	**.2800**	**.3959**	**.8024**	**.8543**
.05	.1172	.1498	.2268	.3463	.7738	.8345
Power of Cramér–von Mises						
.20	.3535	.3889	.4229	.5389	.8198	.8856
.15	.3008	.3232	.4113	.4458	.7993	.8691
.10	.1418	.2138	.2542	.3304	.7583	.8434
.05	.1004	.1184	.2034	.3183	.7234	.8138
Power of Kolmogorov–Smirnov						
.20	.3038	.3359	.3947	.5126	.8057	.8535
.15	.2857	.3015	.3998	.4032	.7328	.8119
.10	.1218	.2028	.2238	.3123	.6888	.7735
.05	.0926	.1039	.1727	.2485	.5311	.6188

Further, it can also be concluded on the basis of Tables 8–10 that:

a) Power in case of testing goodness of fit of WG versus WE distribution is more than in other cases. Hence, the tests are detecting the gap between WG and WE distributions with high power and hence a small sample is sufficient to differentiate WG from WE.

b) The power of all GOFTs for all sample sizes and levels of significance is least when comparing WG and Gamma distributions. This means that the GOFTs are not detecting the difference between these two distributions as efficiently as in other cases. It implies that these distributions are quite close to each other. So, large sample sizes are required to differentiate these distributions.

4.2. Real data set illustration

We consider a data set consisting of survival times of guinea pigs injected with different amount of tubercle bacilli and studied by Bjerkedal [4]. The observations in the data set are: 12 15 22 24 24 32 32 33 34 38 38 43 44 48 52 53 54 54 55 56 57 58 58 59 60 60 60 60 61 62 63 65 65 67 68 70 70 72 73 75 76 76 81 83 84 85 87 91 95 96 98 99 109 110 121 127 129 131 143 146 146 175 175 211 233 258 258 263 297 341 341 376.

This data set was also considered by Gupta and Kundu [9] for fitting of Weighted Exponential (WE) distribution. The estimates of parameters, AIC and AICc values for above considered data set are reported in Table 11.

Table 11: Estimates of the parameters, AIC and AICc values for different distributions.

Distribution	MLE			AIC	AICc
	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\lambda}$		
WG	2.274	1.513	.0172	791.438	791.784
WW	139.4	1.39	.0014	799.271	799.624
WE	1.624	1	.0138	791.138	791.312
Gamma	—	2.081	.0208	792.495	792.669

From Table 11, it is seen that there is not a significant difference in AIC and AICc values for WG and WE models, hence both the models can be considered for fitting to this real data set. As WG provides generalization to many existing distributions viz WE, Gamma and Exponential distributions, hence it can be considered as a better choice for this data set.

4.2.1. Powers of goodness of fit tests for real data set

For power calculation, we generate random samples of sizes 100, 200, 500, 1000 and 2000 under alternative hypothesis. Test statistics are calculated using the estimates of parameters. By comparing these values with critical points, we either reject or do not reject the null hypothesis. Repeating this process 10,000 times and dividing the total number of rejections by 10,000, gives power.

Powers for goodness of fit tests for the following hypotheses have been reported in Tables 12, 13 and 14 respectively:

H_{06}: WG ($\widehat{\alpha}$=2.274, $\widehat{\beta}$=1.513, $\widehat{\lambda}$=.0172) fits the data set well

versus

i) H'_{16}: WW ($\widehat{\alpha}$=139.4, $\widehat{\beta}$=1.39, $\widehat{\lambda}$=.0014) fits the data set well,

ii) H''_{16}: WE ($\widehat{\alpha}$=1.624, $\widehat{\lambda}$=.0138) distribution fits the data well,

iii) H'''_{16}: Gamma ($\widehat{\beta}$=2.081, $\widehat{\lambda}$=.0208) distribution fits the data well.

From the Tables 12–14, it can be concluded that:

a) Anderson–Darling (AD) is the most powerful and Kolmogorov–Smirnov (KS) is the least powerful test.

b) Power for testing GOF of WG versus WW is more than for testing in other cases.

c) Power is least when comparing WG distribution versus WE distribution. This means that for the considered data set, the GOFT's are not detecting the difference between these two models.

Table 12: Power of tests for testing goodness of fit of WG versus WW with estimated parameters.

Level of significance	Sample size				
	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling					
.20	.4829	.6238	.7812	.8524	.9423
.15	.4458	.5744	.7123	.8047	.8850
.10	.3943	.5209	.6838	.7773	.8595
.05	.3451	.4753	.6552	.7239	.8391
Power of Cramér–von Mises					
.20	.4467	.5924	.6874	.7955	.8620
.15	.4139	.5251	.6193	.7338	.8354
.10	.3533	.4435	.5366	.6940	.7889
.05	.2669	.3981	.4921	.6569	.7495
Power of Kolmogorov–Smirnov					
.20	.3999	.5099	.6434	.7809	.8345
.15	.3458	.4875	.5701	.7051	.8003
.10	.3049	.4223	.4959	.6532	.7448
.05	.2225	.3801	.4153	.6034	.7115

Table 13: Power of tests for testing goodness of fit of WG versus WE with estimated parameters.

Level of significance	Sample size				
	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling					
.20	.4145	.5125	.6720	.7518	.8498
.15	.3509	.4809	.6548	.7285	.8156
.10	.2877	.4053	.5740	.6893	.7632
.05	.2329	.3069	.4169	.6673	.7253
Power of Cramér–von Mises					
.20	.3595	.4430	.5407	.6863	.8002
.15	.3250	.4018	.4933	.6545	.7803
.10	.2589	.2944	.4356	.6187	.7234
.05	.2055	.2882	.3204	.5522	.6868
Power of Kolmogorov–Smirnov					
.20	.3486	.3997	.4407	.6562	.7259
.15	.3058	.3449	.4113	.5328	.6885
.10	.2137	.2507	.3876	.4935	.6138
.05	.1851	.2187	.3092	.4580	.5609

Table 14: Power of tests for testing goodness of fit of WG against Gamma with estimated parameters.

Level of significance	Sample size				
	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$
Power of Anderson–Darling					
.20	.4277	.5459	.7032	.8089	.8927
.15	.3639	.4994	.6633	.7746	.8558
.10	.3073	.4227	.6118	.7268	.7982
.05	.2857	.3998	.5844	.6934	.7639
Power of Cramér–von Mises					
.20	.4008	.4935	.6632	.7604	.8239
.15	.3401	.4349	.6110	.7093	.7994
.10	.2831	.3970	.5256	.6859	.7530
.05	.2217	.3239	.5012	.6221	.7126
Power of Kolmogorov–Smirnov					
.20	.3603	.4158	.5728	.6953	.7649
.15	.3041	.3945	.4592	.5889	.7325
.10	.2859	.3567	.4182	.5234	.7049
.05	.2130	.3018	.3993	.5008	.6532

ACKNOWLEDGMENTS

The authors are grateful to the referees for their suggestions that helped in improving the paper. The first author is also thankful to University Grants Commission, Government of India, for providing financial support for this work.

REFERENCES

[1] Anderson, T.W. and Darling, D.A. (1954). A Test of Goodness of Fit, *Journal of American Statistical Association*, **29**, 765–769.

[2] Arnold, B.C. and Beaver, R.J. (2000). The Skew Cauchy distribution, *Statistics and Probability Letters*, **49**, 285–290.

[3] Arshad, M.; Rasool, M.T. and Ahmad, M.I. (2003). Anderson–Darling and Modified Anderson–Darling Tests for Generalized Pareto Distribution, *Pakistan Journal of Applied Sciences*, **3**(2), 85–88.

[4] Bjerkedal, T. (2003). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli, *Amer. J. Hyg.*, **72**, 130–148.

[5] Conover, W.J. (1999). *Practical Nonparametric Statistics*, John Wiley & Sons, New York.

[6] Dufour, J.M.; Farhat, A.; Gardiol, L. and Khalaf, L. (1998). Simulation-based Finite Sample Normality Tests in Linear Regressions, *Econometrics Journal*, **1**, 154–173.

[7] Farrel, P.J. and Stewart, K.R. (2006). Comprehensive Study Of Tests For Normality And Symmetry: Extending The Spiegelhalter Test, *Journal of Statistical Computation and Simulation*, **76**(9), 803–816.

[8] Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, **6**, 13–25.

[9] Gibbons J.D. and Chakraborti S. (2011). *Nonparametric Statistical Inference*, Chapman and Hall / CRC, Florida.

[10] Gupta, R.C. and Kirmani, S.N.U.A. (1990). The role of weighted distributions in stochastic modelling, *Communications in Statistics – Theory and Methods*, **19**(9), 3147–3162.

[11] Gupta, R.D. and Kundu, D. (2009). A new class of weighted exponential distributions, *Statistics*, **43**(6), 621–634.

[12] Jain, K.; Singh, H. and Bagai, I. (1989). Relations for Reliability measures of weighted distribution, *Communications in Statistics – Theory and Methods*, **18**(12), 4393–4412.

[13] Jain, K.; Singla, N. and Gupta, R.D. (2014). A weighted version of Gamma distribution, *Discussiones Mathematicae Probability and Statistics*, **34**, 89–111.

[14] Lemeshko, B.Yu.; Lemeshko, S.B. and Postovalov, S.N. (2009). Comparative Analysis of the Power of Goodness-of-Fit Tests for Near Competing Hypotheses. I. The Verification of Simple Hypotheses, *Journal of Applied and Industrial Mathematics*, **3**(4), 462–475.

[15] Lemeshko, B.Yu.; Lemeshko, S.B. and Postovalov, S.N. (2010a). Comparative Analysis of the Power of Goodness-of-Fit Tests for Near Competing Hypotheses. II. Verification of Complex Hypotheses, *Journal of Applied and Industrial Mathematics*, **4**(1), 79–93.

[16] Lemeshko, B.Yu.; Lemeshko, S.B.; Akushkina, K.A.; Nikulin, M.S. and Saaidia, N. (2010b). *Inverse Gaussian Model and Its Applications in Reliability and Survival Analysis.* In "Mathematical and Statistical Models and Methods in Reliability: Applications to Medicine, Finance and Quality Control" (V. Rykov, N. Balakrishanan and M. Nikulin, Eds.), Birkhäuser, Boston, 433–453.

[17] Lemeshko, B.Yu. and Lemeshko, S.B. (2011). Models of Statistic Distributions of Nonparametric Goodness-of-Fit Tests in Composite Hypotheses Testing for Double Exponential Law Cases, *Communications in Statistics – Theory and Methods*, **40**, 2879–2892.

[18] Lewis, P.A.W. (1961). Distribution of the Anderson–Darling Statistic, *The Annals of Mathematical Statistics*, **32**(4), 1118–1124.

[19] Mahfoud, M. and Patil, G.P. (1982). *On weighted distributions.* In "Statistics and Probability: Essays in Honor of C.R. RAO" (G. Kallianpur, P.R. Krishnaiah and J.K. Ghosh, Eds.), North-Holland, Amsterdam, 479–492.

[20] Nanda, A.K. and K. Jain (1999). Some weighted distribution results on univariate and bivariate cases, *Journal of Statistical Planning and Inference*, **77**(2), 169–180.

[21] Nikulin, M.S. (1973). On a chi-squared test for continuous distributions, *Theory of Probability and Its Applications*, **18**(3), 638–639.

[22] Nikulin, M.S. (1973). Chi-squared test for continuous distributions with shift and scale parameters, *Theory of Probability and Its Applications*, **18**(3), 559–568.

[23] Patil, G.P. (2002). *Weighted Distributions*. In "Encyclopaedia of Environmetrics" (Abdel H. El-Shaarawi and Walter W. Piegorsch, Eds.), John Wiley & Sons, Ltd, Chichester, 2369–2377.

[24] Rao, C.R. (1965). *On discrete distributions arising out of methods of ascertainment*. In "Classical and Contagious Discrete Distributions" (G.P. Patil, Eds.), Pergamon Press, Oxford and Statistical Publishing Society, Calcutta, 320–332.

[25] Rao, K.C. and Robson, D.S. (1974). A Chi-squared statistic for goodness-of-fit tests within the exponential family, *Communication in Statistics*, **3**, 1139–1153.

[26] Seier, E. (2002). Comparison of Tests for Univariate Normality, *Interstat Statistical Journal*, **1**, 1–17.

[27] Shahbaz, S.; Shahbaz, M.Q. and Butt, N.S. (2010). A class of Weighted Weibull Distributions, *Pakistan Journal of Statistical Operation Research*, **6**(1), 53–59.

[28] Stephens, M.A. (1986). *Tests based on EDF statistics*. In "Goodness-of-Fit Techniques" (R.B. D'Agostino and M.A. Stephens, Eds.), Marcel Dekker, New York, 97–193.