Far East Journal of Mathematical Sciences (FJMS)

Published Online: June 2015
http://dx.doi.org/10.17654/FJMSJul2015_655_665
Volume 97, Number 5, 2015, Pages 655-665 **ISSN: 0972-0871**

METHOD OF MOMENTS PARAMETER ESTIMATION FOR BETA INVERSE WEIBULL DISTRIBUTION

Mezbahur Rahman, Dayeong An and Mohammad Shaha Alam Patwary

Minnesota State University
Mankato, MN 56001
U. S. A.
e-mail: mezbahur.rahman@mnsu.edu
dayeong.an@mnsu.edu
mohammad-shaha.patwary@mnsu.edu

Abstract

The inverse Weibull distribution is one of the widely applied distributions. In this paper, we are going to extend the study of the beta inverse Weibull distribution (BIW). The method of moment estimation (MME) procedure is implemented in the BIW distribution with three parameters. A comparison is made with the maximum likelihood estimation (MLE) procedure using simulation. In computations, for both MME and MLE, the Newton-Raphson method is applied and the simulation results are given. Finally, practical use of the model is demonstrated by dint of an application to real data. For the real data computation, using the Fisher information matrix, the variance-covariance matrix is estimated for MLE. Also, bootstrap standard deviations are computed for both MME and MLE for the application data.

Received: January 22, 2015; Revised: April 9, 2015; Accepted: April 20, 2015

2010 Mathematics Subject Classification: 62F10.

Keywords and phrases: beta function, di-gamma function, gamma function, tri-gamma function.

Communicated by K. K. Azad

1. Introduction

The beta inverse Weibull (BIW) distribution is derived from a generalized class of probability distributions discussed by Eugene et al. [3]. Let $F(x)$ be the cumulative distribution function (cdf) of a random variable X. The cdf for a generalized class of distributions for the random variable X, defined by Eugene et al. [3] as the logit of beta random variable, is given as:

$$G(x) = I_{F(x)}(a, b), \quad a > 0 \text{ and } b > 0, \tag{1}$$

where $I_{F(x)}(a, b) = \dfrac{B_{F(x)}(a, b)}{B(a, b)}$, $B_{F(x)}(a, b) = \int_0^{F(x)} t^{a-1}(1-t)^{b-1}dt$, and $B(a, b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

The BIW distribution is introduced by taking the $F(x)$ in (1) to be cdf of inverse Weibull distribution which is given as:

$$F(x) = e^{-x^{-\beta}}. \tag{2}$$

The probability density function (PDF) of the BIW distribution is obtained by taking derivative of (1) and using (2):

$$g(x) = \frac{1}{B(a, b)}[e^{-x^{-\beta}}]^a[1 - e^{-x^{-\beta}}]^{b-1}\beta x^{-(\beta+1)}, \quad x > 0. \tag{3}$$

The random variables $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are defined as an ordered random sample from the BIW distribution (3) which has been used in computational purposes.

In this paper, the method of the maximum likelihood estimation (MLE) is implemented in estimating parameters in a three parameter BIW distribution using the Newton-Raphson method of solving a system of equations. The method of moments estimate is compared with the method of maximum likelihood estimate using simulation.

Beta inverse Weibull distribution was also introduced and its important properties were summarized by Khan [8] and Hanook et al. [6].

The moments of beta inverse distribution can be derived from (3). The rth moment of BIW is given as:

$$E(x^r) = \frac{1}{B(a, b)} \int_0^\infty x^r [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} dx. \tag{4}$$

2. Parameter Estimation

In this section, we employ two widely used methods for parameter estimation. The outlines of those two methods MME and MLE are introduced. Simulation results will be given in the section followed by their descriptions.

2.1. Method of moments estimate

The method of moments to find the estimate of the parameters from nonlinear system of equations is new discovery in modern statistical analysis. Sen and Singer [10, p. 219] have discussed the detail outline of the method. Readers are also referring Wooldridge [11] for more detail understanding of the method. Based on the moment of BIW given by (4), we can set up a system of nonlinear equations as:

$$\frac{1}{n} \sum_{i=1}^n X_i^j = \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} dx;$$

$$j = 1, 2, 3.$$

The left sides of the nonlinear system are coming from sampling, while the right sides are coming from direct calculation of the integrals. By solving the nonlinear system above, we can get an estimation of a, b and β. From a system of nonlinear equations, let f be:

$$f_j(x) = \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} dx - \frac{1}{n} \sum_{i=1}^n X_i^j. \tag{5}$$

For solving the system of nonlinear equations, the Newton-Raphson method can be applied and it requires a matrix which is known as the Jacobian of the system. The Jacobian is defined as:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial a} & \frac{\partial f_1}{\partial b} & \frac{\partial f_1}{\partial \beta} \\ \frac{\partial f_2}{\partial a} & \frac{\partial f_2}{\partial b} & \frac{\partial f_2}{\partial \beta} \\ \frac{\partial f_3}{\partial a} & \frac{\partial f_3}{\partial b} & \frac{\partial f_3}{\partial \beta} \end{bmatrix}. \tag{6}$$

The elements of the Jacobian matrix or Hessian matrix J (Gradshteyn and Ryzhik [5, p. 1069]) are commonly known as second order Fréchet derivatives (Fréchet [4]) and are defined as:

$$\frac{\partial f_j}{\partial a} = \frac{\Psi(a+b) - \Psi(a)}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} dx$$

$$+ \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} \log e^{-x^{-\beta}} dx,$$

$$\frac{\partial f_j}{\partial b} = \frac{\Psi(a+b) - \Psi(b)}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} dx$$

$$+ \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(\beta+1)} \log(1 - e^{-x^{-\beta}}) dx$$

and

$$\frac{\partial f_j}{\partial \beta} = \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} x^{-(\beta+1)} dx$$

$$+ \frac{1}{B(a, b)} \int_0^\infty x^j [e^{-x^{-\beta}}]^a [1 - e^{-x^{-\beta}}]^{b-1} \beta x^{-(2\beta+1)}$$

$$\cdot \log x \left(a - x^\beta - \frac{(b-1) e^{-x^{-\beta}}}{(1 - e^{-x^{-\beta}})} \right) dx,$$

where $\Psi(a) = \frac{\Psi'(a)}{\Psi(a)}$ and $\Psi'(a)$ is the derivative of $\Psi(a)$ with respect to *a*. $\Psi(a)$ is known as the di-gamma function and $\Psi'(a)$ is known as the tri-gamma function.

Three parameters can be easily estimated by the Newton-Raphson method. The multivariate Newton-Raphson iteration is performed as

$$\begin{bmatrix} \hat{a}_M^{(l+1)} \\ \hat{b}_M^{(l+1)} \\ \hat{\beta}_M^{(l+1)} \end{bmatrix} = \begin{bmatrix} \hat{a}_M^{(l)} \\ \hat{b}_M^{(l)} \\ \hat{\beta}_M^{(l)} \end{bmatrix} - \begin{bmatrix} \frac{\partial f_1^{(l)}}{\partial a} & \frac{\partial f_1^{(l)}}{\partial b} & \frac{\partial f_1^{(l)}}{\partial \beta} \\ \frac{\partial f_2^{(l)}}{\partial a} & \frac{\partial f_2^{(l)}}{\partial b} & \frac{\partial f_2^{(l)}}{\partial \beta} \\ \frac{\partial f_3^{(l)}}{\partial a} & \frac{\partial f_3^{(l)}}{\partial b} & \frac{\partial f_3^{(l)}}{\partial \beta} \end{bmatrix}^{-1} \begin{bmatrix} f_1^{(l)}(x) \\ f_2^{(l)}(x) \\ f_3^{(l)}(x) \end{bmatrix}, \tag{7}$$

where *l* is the index for the iterations.

Apart from univariate case, Newton-Raphson method allows dealing with multivariate cases too. Though this method is not new in advance mathematical and statistical researches, the multivariate Newton-Raphson method (Andrews et al. [2], Harville [7, p. 309] and Michel [9]) is outlined below in brief:

For the multivariate case where $f : \mathbb{R}^n \to \mathbb{R}^n$.

The estimate of the parameters can be obtained by solving the following equation:

$$\hat{\Lambda}^{(t+1)} = \hat{\Lambda}^{(t)} - (H(\hat{\Lambda}^{(t)}))^{-1} D(\hat{\Lambda}^{(t)}),$$

where the column vector $D^{(t)}$ is computed using first order Fréchet derivatives, the Hessian matrix $H^{(t)}$ computed using second order Fréchet derivatives. For the initial value of $\hat{\Lambda}^{(t)}$, $\hat{\Lambda}^{(t)} = \mathbf{1}$, a unit column vector is an obvious choice.

3. Maximum Likelihood Estimates (MLE)

The maximum likelihood estimation is a method of estimating the parameters for the model. To use this method, one first specifies the joint density function for all observations and this function is called the *likelihood*:

$$g(x_1, x_2, \ldots, x_n \mid a, b, \beta) = \prod_{i=1}^{n} g(x_i \mid a, b, \beta).$$

Let its natural logarithm be

$$L = \log(g(x_1, x_2, \ldots, x_n \mid a, b, \beta)).$$

The log-likelihood function is:

$$L = -n \log(B(a, b)) + n \log(\beta) + \sum_{i=1}^{n} \log(x_i^{-(\beta+1)}) - a \sum_{i=1}^{n} x_i^{-\beta}$$

$$+ (b-1) \sum_{i=1}^{n} \log(1 - e^{-x_i^{-\beta}}) + \sum_{i=1}^{n} \log(x_i^{-(\beta+1)}). \tag{8}$$

The first derivatives of the log-likelihood function (8) with respect to a, b, and β are, respectively,

$$L'_a = n\Psi(a+b) - \sum_{i=1}^{n} x_i^{-\beta} - n\Psi(a),$$

$$L'_b = n\Psi(a+b) + \sum_{i=1}^{n} \log(1 - e^{-x_i^{-\beta}}) - n\Psi(b)$$

and

$$L'_\beta = \frac{n}{\beta} - \sum_{i=1}^{n} \log x_i + a \sum_{i=1}^{n} (x_i^{-\beta} \log x_i) - (b-1) \sum_{i=1}^{n} \frac{x_i^{-\beta} e^{-x_i^{-\beta}} \log x_i}{1 - e^{-x_i^{-\beta}}},$$

where $\Psi(a) = di\text{-}gamma(a) = \frac{\partial(\log(\Gamma(a)))}{\partial a}$. The second derivatives of the log-likelihood function (8) with respect to a, b and β are, respectively,

$$L''_{aa} = n\Psi_1(a+b) - n\Psi_1(a),$$

$$L''_{ab} = n\Psi_1(a+b),$$

$$L''_{a\beta} = \sum_{i=1}^{n} (x_i^{-\beta} \log x_i),$$

$$L''_{bb} = n\Psi_1(a+b) - n\Psi_1(b),$$

$$L''_{b\beta} = \sum_{i=1}^{n} \frac{e^{-x_i^{-\beta}} x_i^{-\beta} \log x_i}{e^{-x_i^{-\beta}} - 1}$$

and

$$L''_{\beta\beta} = -\frac{n}{\beta^2} - a\sum_{i=1}^{n} (x_i^{-\beta}(\log x_i)^2) + (b-1)\sum_{i=1}^{n} (\log x_i)^2 x_i^{-\beta} e^{-x_i^{-\beta}}$$

$$\cdot \left[\frac{1}{1-e^{-x_i^{-\beta}}} - \frac{x_i^{-\beta}}{(1-e^{-x_i^{-\beta}})^2}\right],$$

where $\Psi_1(a) = \frac{\Gamma(a)\Gamma''(a) - (\Gamma'(a))^2}{(\Gamma(a))^2}$ and $\Gamma''(a)$ is the derivative of $\Gamma'(a)$ with respect to a. $\Psi_1(a)$ is also known as the tri-gamma function.

Three parameters of BIW distribution can then be easily estimated by the Newton-Raphson method (An and Rahman [1]). The multivariate Newton-Raphson iteration is performed as:

$$\begin{bmatrix} \hat{a}_L^{(l+1)} \\ \hat{b}_L^{(l+1)} \\ \hat{\beta}_L^{(l+1)} \end{bmatrix} = \begin{bmatrix} \hat{a}_L^{(l)} \\ \hat{b}_L^{(l)} \\ \hat{\beta}_L^{(l)} \end{bmatrix} - \begin{bmatrix} L''^{(l)}_{aa} & L''^{(l)}_{ab} & L''^{(l)}_{a\beta} \\ L''^{(l)}_{ab} & L''^{(l)}_{bb} & L''^{(l)}_{b\beta} \\ L''^{(l)}_{a\beta} & L''^{(l)}_{b\beta} & L''^{(l)}_{\beta\beta} \end{bmatrix}^{-1} \begin{bmatrix} L'^{(l)}_{a} \\ L'^{(l)}_{b} \\ L'^{(l)}_{\beta} \end{bmatrix}, \tag{9}$$

where l is the index for the iterations.

Table 1. Simulation results

MEAN	SD	BIAS	MAD	MSE	MEAN	SD	BIAS	MAD	MSE
MLE	$n = 20$	$a = 2$	$b = 2$	$\beta = 2$	*MLE*	$n = 20$	$a = 3$	$b = 2$	$\beta = 1$
1.6354	0.9208	–0.3646	0.8438	0.9809	2.1164	0.9320	–0.8836	1.1008	1.6494
1.5508	0.9981	–0.4492	0.9482	1.1981	1.1286	0.6321	–0.8714	0.9670	1.1588
3.0303	1.3027	1.0303	1.1780	2.7585	1.7904	0.6970	0.7904	0.8135	1.1105
MLE	$n = 50$	$a = 2$	$b = 2$	$\beta = 2$	*MLE*	$n = 50$	$a = 3$	$b = 2$	$\beta = 1$
1.7220	0.7327	–0.2780	0.6686	0.6142	2.2071	0.6477	–0.7929	0.8735	1.0482
1.6757	0.8279	–0.3243	0.7685	0.7905	1.2524	0.5288	–0.7476	0.8063	0.8385
2.5885	0.8890	0.5885	0.7527	1.1367	1.5091	0.4537	0.5091	0.5210	0.4650
MLE	$n = 100$	$a = 2$	$b = 2$	$\beta = 2$	*MLE*	$n = 100$	$a = 3$	$b = 2$	$\beta = 1$
1.8409	0.6461	–0.1591	0.5624	0.4428	2.3782	0.5328	–0.6218	0.6868	0.6705
1.8233	0.7518	–0.1767	0.6572	0.5965	1.4111	0.4628	–0.5889	0.6425	0.5609
2.3311	0.6276	0.3311	0.5179	0.5035	1.3267	0.3086	0.3267	0.3381	0.2019
MEAN	SD	BIAS	MAD	MSE	MEAN	SD	BIAS	MAD	MSE
MME	$n = 20$	$a = 2$	$b = 2$	$\beta = 2$	*MME*	$n = 20$	$a = 3$	$b = 2$	$\beta = 1$
2.0876	0.7108	0.0876	0.4463	0.5130	3.5355	1.4545	0.5355	1.0610	2.4024
2.2687	0.8603	0.2687	0.5625	0.8124	2.5718	1.5259	0.5718	0.9786	2.6552
2.1487	0.2638	0.1487	0.1789	0.0917	1.0025	0.3101	0.0025	0.2360	0.0961
MME	$n = 50$	$a = 2$	$b = 2$	$\beta = 2$	*MME*	$n = 50$	$a = 3$	$b = 2$	$\beta = 1$
1.6886	10.7760	–0.3114	0.5438	116.2185	3.5128	1.1774	0.5128	0.9065	1.6493
1.9048	0.4918	–0.0952	0.3553	0.2509	2.4849	1.3994	0.4849	0.9272	2.1935
2.0690	0.1707	0.0690	0.1020	0.0339	0.9812	0.2944	–0.0188	0.2303	0.0870
MME	$n = 100$	$a = 2$	$b = 2$	$\beta = 2$	*MME*	$n = 100$	$a = 3$	$b = 2$	$\beta = 1$
1.7775	0.4342	–0.2225	0.3776	0.2381	3.3274	1.0998	0.3274	0.7859	1.3168
1.8129	0.3543	–0.1871	0.3160	0.1606	2.2963	1.2040	0.2963	0.8229	1.5375
2.0382	0.1340	0.0382	0.0787	0.0194	0.9988	0.3066	–0.0012	0.2316	0.0940

4. Simulation

Ten thousand samples are generated for two different parameter settings $(a = 2, b = 2, \beta = 2)$ and $(a = 3, b = 2, \beta = 1)$ and for three different sample sizes $n = 20,$ $n = 50$ and $n = 100$ for the both MLE and MME methods. Means (MEAN), standard deviations (SD), biases (BIAS), mean of the absolute deviations from the mean (MAD) and mean squared errors (MSE) are computed and displayed in Table 1. MATLAB software is used in all computations and the codes are readily available.

Table 1 results show that the MLEs are consistent and are asymptotically unbiased without any exception. MMEs are also consistent and asymptotically unbiased. There is an exception in MME for $a = 2$ and $n = 50,$ the standard deviation became exceptionally high which caused the MSE also high. Other than this exception, standard errors are usually smaller but biases are slightly higher in some instances for MME. It is to be noted that due to the computational complexities, the computations are done independently which can explain some discrepancies. In computations for both the methods, built-in MATLAB codes for the di-gamma and the psi-gamma functions are used. Some samples were discarded as they were ill-conditioned for our computations. But in practice, by changing the starting values of the parameters and/or adjusting the tolerance limits of the built-in functions, a successful computation can be performed.

5. Application

The following data in Table 2 represents failure times of machine parts from a manufacturer and are taken from

http://v8doc.sas.com/sashtml/stat/chap29/sect44.htm.

Table 2. Failure times

620	470	260	89	388	242	103	100	39	460	284	1285	218	393	106
158	152	477	403	103	69	158	818	947	399	1274	32	12	134	660
548	381	203	871	193	531	317	85	1410	250	41	1101	32	421	32
343	376	1512	1792	47	95	76	515	72	1585	253	6	860	89	1055
537	101	385	176	11	565	164	16	1267	352	160	195	1279	356	751
500	803	560	151	24	689	1119	1733	2194	763	555	14	45	776	1

The MLEs of the parameters are: $\hat{a} = 1.3838$, $\hat{b} = 7.9252$, $\hat{\beta} = 0.3334$ and their respective standard error estimates are computed using the information matrix as 2.0811, 6.8664 and 0.1872. The MMEs of the parameters are: $\hat{a} = 0.1401$, $\hat{b} = 5.5426$, $\hat{\beta} = 0.9248$. For the MMEs, the standard errors can be estimated using the bootstrap sampling. Note that the data were scaled by dividing the maximum value to avoid numerical computational complexities.

References

[1] D. An and M. Rahman, Maximum likelihood parameter estimation for beta inverse Weibull distribution, Far East J. Math. Sci. (FJMS) 97(2) (2015), 131-137.

[2] D. F. Andrews, R. Gnanadesikan and J. L. Warner, Transformations of multivariate data, Biometrics 27 (1971), 825-840.

[3] N. Eugene, C. Lee and F. Famoye, Beta-normal distribution and its applications, Comm. Statist. Theory Methods 31(4) (2002), 497-512.

[4] M. R. Fréchet, Sur la loi de répartition de certaines grandeurs géographiques, Journal de la Société de Statistiques de Paris (in French) 82 (1941), 114-122.

[5] I. S. Gradshteyn and I. M. Ryzhik, Hessian Determinants, § 14.314 in Tables of Integrals, Series, and Products, 6th ed., Academic Press, San Diego, CA, 2000.

[6] S. Hanook, M. Q. Shahbaz, M. Mohsin and B. M. G. Kibria, A note on beta inverse-Weibull distribution, Comm. Statist. Theory Methods 42(2) (2013), 320-335.

[7] D. A. Harville, Matrix Algebra from a Statistician's Perspective, Springer, New York, 1999.

[8] M. S. Khan, The beta inverse Weibull distribution, International Transactions in Mathematical Sciences and Computer 3(1) (2010), 113-119.

[9] R. Michel, A multidimensional Newton-Raphson method and its applications to the existence of asymptotic F_n-estimators and their stochastic expansions, J. Multivariate Anal. 7 (1977), 235-248.

[10] P. K. Sen and J. M. Singer, Large Sample Methods in Statistics: An Introduction with Applications, Chapman and Hall, New York, London, 1993.

[11] J. M. Wooldridge, Applications of generalized method of moments estimation, Journal of Economic Perspectives 15(4) (2001), 87-100.