$$G(x) = \frac{1}{B(a,b)} \int_0^{F(x)} t^{a-1}(1-t)^{b-1}dt, \quad a > 0, \quad b > 0$$

Investigation of Beta Inverse Weibull Distribution

Md Nur Islam

Department of Math & Statistics

MOTIVATION

- **An and Rahman (2015) explored the parameters estimation of the BIW distribution using Newton-Raphson method of system of likelihood equations.**

- **Newton-Raphson method does not guarantee the global maximum and its sensitive on the starting values.**

MOTIVATION

- So in these cases, we tried to investigate BIW distribution to get better estimation.

- We applied Newton-Raphson method and Steepest Descent method with single and multiple initial guess.

INTRODUCTON

- **Maximum Likelihood Estimation (MLE) method is used to estimate our targeted three parameters.**

- **In computation The Newton-Raphson method and Steepest Descent method are used with single and multiple initial guess.**

- **Finally we investigated the application of *BIW* distribution with a real data set.**

- Let $F(x)$ be the cdf of a random variable X then the cdf for general class of distribution for the random variable X, as the $logit$ of beta random variables is as,

$$G(x) = I_{F(x)}(a,b), \qquad a > 0, b > 0$$

where $I_{F(x)}(a,b) = \frac{B_{F(x)}(a,b)}{B(a,b)}$

and $B_{F(x)}(a,b) = \int_0^{F(x)} t^{a-1}(1-t)^{b-1}dt, \; B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

- Thus the cdf of *BIW* distribution is as,

$$G(x) = \frac{1}{B(a,b)} \int_0^{e^{-x^{-\beta}}} t^{a-1}(1-t)^{b-1}dt, \qquad a > 0, b > 0$$

- The pdf of *BIW* distribution is obtained by taking derivative of $cdf(G(x))$ of *BIW* distribution as,

$$g(x) = \frac{1}{B(a,b)} \left[e^{-x^{-\beta}}\right]^a \left[1 - e^{-x^{-\beta}}\right]^{b-1} \beta x^{-(\beta+1)}$$

- **MLE is a method of estimating the parameters for the model, to use this method need to specify the Joint pdf for all observations called likelihood function,**

$$\boldsymbol{g(x_1, x_2, \dots, x_n|a, b, \beta) = \sum_{i=1}^{n} g(x_i|a, b, \beta)}$$

- **The log-likelihood function is,**

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

$$+(b-1)\sum_{i=1}^{n} \log\left(1 - e^{-x_i^{-\beta}}\right) + \sum_{i=1}^{n} \log\left(x_i^{-(\beta+1)}\right)$$

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

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

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

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

❑ The second derivatives of the log-likelihood function,

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

$$L''_{a\beta} = \sum_{i=1}^{n}\left(x_i^{-\beta}\log x_i\right), \qquad L''_{bb} = n\Psi_1(a,b) - n\Psi_1(b),$$

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

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

$\Psi(a) = \frac{\partial\left(\log(\Gamma(a))\right)}{\partial a}$ **and** $\Psi_1(a) = \frac{\Gamma(a)\Gamma''(a)-\left(\Gamma'(a)\right)^2}{\left(\Gamma(a)\right)^2}$ **are *digamma* and *tri-gamma* functions respectively**

- **Newton’s method is widely used to find successively better approximation to the roots of a real-valued function.**

- **In the one-dimensional problem, Newton's method attempts to construct a sequence x_n from an initial guess x_0 that converges towards some value x_k satisfying $f'(x_k) = 0$.**

- **Newton-Raphson method is mainly generated from the expansion of Taylor series.**

- **Three parameters $a, b,$ and β of *Beta Inverse Weibull* Distribution can then be easily estimated by the Newton-Raphson method. For our objective function the multivariate Newton-Raphson iteration is performed as,**

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

STEEPEST DESCENT METHOD

- A weakness of Newton’s and Quasi-Newton’s methods is that an accurate initial approximation to the solution is needed to ensure convergence. The Steepest Descent method implemented in this research usually converge even with poor initial approximations.

- For the quantities of the estimation of the parameters, such as for the estimation of the standard deviation, the Steepest Descent method provides the most accurate and stable estimates.

- The method of Steepest Descent for finding a local minimum for an arbitrary function f from $\mathbb{R}^n \to \mathbb{R}$ can be intuitively described as follows,

1. Evaluate f at an initial approximation $X^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})'$.

2. Determine a direction from $X^{(0)}$ that results in a decrease in the value of f.

3. Move an appropriate amount in this direction and call the new vector $X^{(1)}$.

4. Repeat steps 1 through 3 with $X^{(0)}$ replaced by $X^{(1)}$.

SIMULATION

- In simulation we used MATLAB and generated ten thousand samples for two different parameter settings $(a = 2, b = 2, \beta = 2)$ and $(a = 3, b = 2, \beta = 1)$ and considered three different sample sizes $n = 20, n = 50,$ and $n = 100.$

- Means $(MEAN)$, standard deviations (SD), biases $(BIAS)$, mean of the absolute deviations from the mean (MAD) and mean squared errors (MSE) are computed.

- For the multiple starts, five different arbitrary starting values are considered and then the solution among the five leads to the maximum likelihood value is accepted as a solution.

SIMULATION

- **The estimated values for $a = 2, b = 2,$ and $\beta = 2$ by using Newton-Raphson method with one initial guess,**

		$a = 2$	$b = 2$	$\beta = 2$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		2.7031	2.3175	0.7031	1.5376	5.8652
		2.9343	3.1030	0.9343	1.9643	10.5015
		2.5788	1.4670	0.5788	1.1057	2.4871
50		2.6849	1.9961	0.6849	1.3314	4.4537
		2.9297	2.6855	0.9297	1.6955	8.0761
		2.1796	0.9048	0.1796	0.7186	0.8510
100		2.5010	1.6167	0.5010	1.0516	2.8648
		2.6867	2.1483	0.6867	1.3283	5.0866
		2.0801	0.6820	0.0801	0.5480	0.4715

SIMULATION

- For the parameters $a = 3, b = 2$, and $\beta = 1$,

		$a = 3$	$b = 2$	$\beta = 1$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		3.0173	1.9647	0.0173	1.4476	3.8603
		2.0895	2.0421	0.0895	1.4106	4.1781
		1.5085	0.7746	0.5085	0.6601	0.8586
50		3.0386	1.4131	0.0386	1.0879	1.9983
		2.0962	1.4793	0.0962	1.0795	2.1975
		1.2346	0.4790	0.2346	0.3885	0.2844
100		3.1643	1.2451	0.1643	0.9671	1.5773
		2.2091	1.2794	0.2091	0.9622	1.6806
		1.1165	0.3688	0.1165	0.2921	0.1496

SIMULATION

- **we have seen that for the estimated values of Newton's method the values of SD, BIAS, MAD, and MSE are decreased as of sample size increased.**
- **That means we have gotten better estimation for higher sample size by using Newton-Raphson method with a single initial guess.**
- **So the best estimation for the parameters** $a = 2, b = 2,$ **and** $\beta = 2$ **are** $a = 2.5010, b = 2.6867,$ **and** $\beta = 2.0801$. **Similarly for the parameters** $a = 3, b = 2,$ **and** $\beta = 1$ **the better estimation are** $a = 3.1643, b = 2.2091,$ **and** $\beta = 1.1165$.

SIMUATION

- **The estimated values for $a = 2, b = 2,$ and $\beta = 2$ by using Newton-Raphson method with multiple start,**

		$a = 2$	$b = 2$	$\beta = 2$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		2.6123	1.6617	0.6123	1.2201	3.1361
		2.7290	2.1272	0.7290	1.4751	5.0565
		2.2219	0.8911	0.2219	0.7139	0.8433
50		2.4684	1.3756	0.4684	1.0176	2.1116
		2.6172	1.7730	0.6172	1.2761	3.5243
		2.0925	0.6977	0.0925	0.5635	0.4953
100		2.4684	1.3756	0.4684	1.0176	2.1116
		2.6172	0.7730	0.6172	1.2761	3.5243
		2.0925	0.6977	0.0925	0.5635	0.4953

SIMULATION

❑ For the parameters $a = 3, b = 2$, and $\beta = 1$,

		$a = 3$	$b = 2$	$\beta = 1$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		2.8182	1.3934	-0.1818	1.1406	1.9746
		1.8776	1.3144	-0.1224	1.0650	1.7427
		1.3793	0.6162	0.3793	0.5179	0.5235
50		2.9298	1.1960	-0.0702	0.9688	1.4354
		1.9436	1.1401	-0.0565	0.9254	1.3030
		1.2273	0.4658	0.2273	0.3695	0.2687
100		3.0253	1.0558	0.0253	0.8417	1.1154
		2.0346	1.0382	0.0346	0.8190	1.0790
		1.1268	0.3375	0.1268	0.2717	0.1300

SIMULATION

- **For the parameter $a = 2, b = 2,$ and $\beta = 2$ the values of SD, BIAS, MAD, and MSE are same for the sample size 50 and 100. That means in this case multiple start of Newton-Raphson gives better estimation when sample size is 50.**

- **On the other hand, for the parameters $a = 3, b = 2,$ and $\beta = 1$ accuracy is increased due to the sample size increased.**

- **So the best estimation for the parameters $a = 2, b = 2,$ and $\beta = 2$ are $a = 2.4684, b = 2.6172,$ and $\beta = 2.0925$. Similarly for the parameters $a = 3, b = 2,$ and $\beta = 1$ the better estimation are $a = 3.0253, b = 2.0346,$ and $\beta = 1.1268$.**

SIMULATION

- The estimated values for $a = 2, b = 2$, and $\beta = 2$ by using Steepest Descent method with single start,

		$a = 2$	$b = 2$	$\beta = 2$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		1.7317	0.4995	-0.2683	0.4577	0.3215
		1.6065	0.4359	-0.3935	0.4814	0.3448
		2.5338	0.5881	0.5338	0.5863	0.6309
50		1.6775	0.2908	-0.3225	0.3679	0.1886
		1.5928	0.2812	-0.4072	0.4320	0.2449
		2.3585	0.3039	0.3538	0.3775	0.2209
100		1.6603	0.2068	-0.3397	0.3509	0.1582
		1.5956	0.2055	-0.4044	0.4093	0.2058
		2.3245	0.2160	0.3245	0.3319	0.1520

❑ For the parameters $a = 3, b = 2$, and $\beta = 1$,

		$a = 3$	$b = 2$	$\beta = 1$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		2.2382	1.4005	-0.7618	1.0406	2.5416
		1.1356	0.9483	-0.8644	0.9694	1.6464
		1.6181	1.8426	0.6181	0.7539	3.7772
50		2.1085	0.5052	-0.8915	0.9366	1.0501
		1.1392	0.3777	-0.8608	0.8809	0.8836
		1.4884	0.3448	0.4884	0.4916	0.3574
100		2.0823	0.3769	-0.9177	0.9379	0.9842
		1.1385	0.3073	-0.8615	0.8726	0.8366
		1.4336	0.2283	0.4336	0.4352	0.2401

SIMULATION

- **We have gotten better estimation for higher sample size by using Steepest-descent method with a single initial guess. So the best estimation for the parameters** $a = 2, b = 2,$ **and** $\beta = 2$ **are** $a = 1.6603, b = 1.5956,$ **and** $\beta = 2.3245$**.**

Similarly for the parameters $a = 3, b = 2,$ **and** $\beta = 1$ **the better estimation are** $a = 2.0823, b = 1.1385,$ **and** $\beta = 1.4336.$

SIMULATION

- The estimated values for $a = 2, b = 2,$ and $\beta = 2$ by using Steepest Descent method with multiple start,

		$a = 2$	$b = 2$	$\beta = 2$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		1.8067	0.4499	-0.1933	0.3922	0.2397
		1.6884	0.4100	-0.3116	0.4160	0.2652
		2.4358	0.4750	0.4358	0.5052	0.4156
50		1.7374	0.3051	-0.2626	0.3229	0.1621
		1.6725	0.3082	-0.3275	0.3630	0.2023
		2.3211	0.3100	0.3211	0.3555	0.1992
100		1.7303	0.2363	-0.2697	0.2898	0.1286
		1.6718	0.2661	-0.3282	0.3426	0.1785
		2.2746	0.2530	0.2746	0.2958	0.1394

SIMULATION

❑ For the parameters $a = 3, b = 2$, and $\beta = 1$,

		$a = 3$	$b = 2$	$\beta = 1$		
n	θ	$MEAN$	SD	$BIAS$	MAD	MSE
20		2.2042	0.7112	-0.7958	0.9449	1.1392
		1.1122	0.3865	-0.8878	0.8938	0.9377
		1.6139	0.5173	0.6139	0.6203	0.6445
50		2.1280	0.4018	-0.8720	0.8913	0.9218
		1.1414	0.2663	-0.8586	0.8608	0.8081
		1.4540	0.3050	0.4540	0.4548	0.2991
100		2.1117	0.2498	-0.8883	0.8892	0.8515
		1.1673	0.1896	-0.8327	0.8327	0.7293
		1.3790	0.2008	0.3790	0.3792	0.1840

SIMULATION

- **we have gotten better estimation for higher sample size by using Steepest-descent method with multiple initial guess. So the best estimation for the parameters $a = 2, b = 2,$ and $\beta = 2$ are $a = 1.7303, b = 1.6718,$ and $\beta = 2.2746$.**

Similarly for the parameters $a = 3, b = 2,$ and $\beta = 1$ the better estimation are $a = 2.1117, b = 1.1673,$ and $\beta = 1.3790$.

❑ Observation for Simulations:

- Throughout the table 1 - table 4, we noticed that the SD and the MSE decreased as sample size increased, so we can say better estimation has come from higher sample size.

- In table 1 - table 2, we notice that the SD and the MSE are lower for multiple starts. Similar can be seen in table 3 - table 4, that means multiple start for both methods provide better estimation.

- **By comparing table 1 – table 2 with table 3 - table 4, we notice that the SD and the MSE are lower for the Steepest Descent method compared to the Newton-Raphson method. Thus, Steepest Descent method provides best estimation.**

- **Exception:**

For instance, for the parameters $a = 3, b = 2,$ and $\beta = 1$ single start of Newton-Raphson provides better estimation compare to the Steepest Descent method with multiple start for the parameter β when sample size is 50, and 100 based on MAD, and MSE.

APPLICATION

- **The following data represents failure times of machine parts from manufacturer A and these are taken from http://v8doc.sas.com/sashtml/stat/chap29/sect44.htm:**

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

- **Estimated parameters for the failure time data set by using Newton-Raphson and Steepest Descent method with single and multiple start,**

θ	Newton-Raphson Single Start	Newton-Raphson Multiple Start	Steepest Descent Single Start	Steepest Descent Multiple Start
a	1.3838	1.3838	0.1479	0.1625
b	7.9252	7.9252	1.4137	1.3870
β	0.3334	0.3334	0.7202	0.6988

CONCLUDING REMARKS

- **The Steepest Descent method provides lower standard deviations and mean square errors as compared to the Newton-Raphson method.**

- **Multiple start for both the methods provides better estimation based on standard deviation and mean square errors compare to the single initial guess.**

- **So we can conclude that the Steepest Descent Method with multiple start is one of the best method to estimate the parameters for *Beta Inverse Weibull* Distribution.**

REFERENCES

[1] An, Dayeong and Rahman, Mezbahur (2015). Maximum Likelihood Parameter Estimation for Beta Inverse Weibull Distribution. *Far East Journal of Mathematical Sciences (FJMS),* **97**(2), 131-137.

[2] Burden, R. L. and Faires, J. D. (2011). *Numerical Analysis.* Boston, MA: BROOKS/COLE, CENGAGE Learning

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

[4] Hanook, S., Shahbaz, M. Q., Mohsin, M., and Kibria, B. M. G. (2013). A Note On Beta Inverse-Weibull Distribution. *Communications in Statistics - Theory and Methods,* **42**(2), 320-335.

[5] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing. Second Edition.* Cambridge University Press.

[6] Khan, M. S. (2010). The Beta Inverse Weibull Distribution. *International T ransactions in Mathematical Sciences and Computer,* **3**(1), 113-119.