Prediction under modified Weibull distribution with applications

Saieed F. Ateya

Mathematics & Statistics Department
Faculty of Science, Taif University, Taif, Saudi Arabia
Permanent Address: Mathematics Department,
Faculty of Science, Assiut University, Assiut , Egypt
Email: 4270176 @gmail.com

Mohamed M. Rizk

Mathematics & Statistics Department, Faculty of Science
Taif University, Taif, Saudi Arabia
Permanent Address: Mathematics Department, Faculty of Science,
Menoufia University, Shebin El-Kom, Egypt
Email: mhm96@yahoo.com

Abstract

In this paper, the Bayesian prediction intervals ($BPI's$) of future observations are obtained under Modified Weibull Distribution (MWD) in case of one and two-sample prediction schemes. Based on a type-II censored sample from a real data set, the $BPI's$ of the remaining observations are obtained.

Key words and phrases: Generalized exponential distribution; Bayesian prediction intervals; one and two-sample prediction schemes; modified Weibull distribution.
Mathematics Subject Classification (2000) 62F10; 62F15; 62N01; 62N02.

1 Introduction

The Weibull distribution is one of the most popular and widely used models of failure time in life testing and reliability theory. The Weibull distribution has been shown to be useful for modeling and analysis of life time data in medical, biological and engineering sciences. Applications of the Weibull distribution in various fields are given in Zaharim et al. [21], Gotoh et al. [7], Shamilov et al. [16], Vicen-Bueno et al. [20], Niola et al. [15] and Green et al. [8]. A great deal of research has been done on estimating the parameters of the Weibull distribution using both classical and Bayesian techniques, and a very good summary of this work can be found in Johnson et al. [11]. Hossain and Zimmer [9] have discussed some comparisons of estimation methods for Weibull parameters using complete and censored samples. Jaheen and Harbi[10] studied the Bayesian estimation of the exponentiated Weibull distribution using Markov chain Monte Carlo simulation. The modified Weibull distribution was proposed by Lai et al. [12] as a new lifetime distribution. They have shown the capability of the model for modeling a bathtub-shaped hazard-rate function. In addition, they characterized the model through the Weibull plot paper. Further, they have shown that the modified Weibull model compares well with other competing models to fit data that exhibit a bathtub-shaped hazard-rate function. Sultan [17] studied the record values from the modified Weibull distribution and studied its applications. Vasile et al. [19] used the Bayes method to estimate the parameters of the modified Weibull distribution and Upadhyaya and Gupta [18] studied the Bayes analysis of the modified Weibull distribution using Markov chain Monte Carlo simulation. Ateya[4] study the estimation problem under a censored sample of generalized order statistics from MWD.

Bayesian prediction of future observations play a fundamental role in disciplines such as medicine, economics, engineering, education, metrology and politics, among others. Moreover, Geisser [6] mentioned that the problem of prediction can be solved fully within the Bayesian framework. For more details, see Ateya[1-3], and Ateya and Rizk[5].

A random variable X is said to have a MWD with vector of parameters $\boldsymbol{\theta} = (\beta, \tau, \lambda)$ if its probability density function (pdf) is given by

$$f(x;\boldsymbol{\theta}) = \tau\,(\beta+\lambda\,x)\,x^{\beta-1}\,\exp(\lambda\,x)\,\exp(-\tau\,x^{\beta}\,e^{\lambda\,x}),\; x \geq 0, (\tau > 0, \beta \geq 0, \lambda \geq 0). \tag{1.1}$$

The reliability function (rf) and the hazard rate function (hrf) of this distribution can be written, respectively as

$$R(x) = \exp(-\tau\,x^{\beta}\,e^{\lambda\,x}), \tag{1.2}$$

$$h(x) = \tau\,(\beta+\lambda\,x)\,x^{\beta-1}\,\exp(\lambda\,x), \tag{1.3}$$

2 Bayesian Prediction Intervals of Future Observations in Case of One-sample Scheme

Suppose that $X_1 < X_2 < ... < X_r$ is the informative sample, representing the first r ordered lifetimes of a random sample of size n drawn from a population with pdf $f_X(x)$, cumulative distribution function (cdf) $F_X(x)$ and rf $R(x)$. In one-sample scheme, the $BPI's$ for the remaining unobserved future $(n-r)$ lifetimes are sought based on the first r observed ordered lifetimes. For the remaining $(n-r)$ components, let $Y_s = X_{r+s}$ denote the future lifetime of the s^{th} component to fail, $1 \leq s \leq (n-r)$. The conditional density function of Y_s given that the r components had already failed is

$$g_1(y_s|\boldsymbol{\theta}) \propto [R(x_r) - R(y_s)]^{(s-1)}[R(y_s)]^{n-r-s}[R(x_r)]^{-(n-r)}f_X(y_s|\boldsymbol{\theta}), y_s > x_r, \tag{2.1}$$

where $\boldsymbol{\theta}$ is the vector of parameters. The predictive density function is given by

$$g_1^*(y_s|x) = \int_{\boldsymbol{\theta}} g_1(y_s|\boldsymbol{\theta})\pi^*(\boldsymbol{\theta}|x)d\boldsymbol{\theta}, y_s > x_r, \tag{2.2}$$

$\pi^*(\boldsymbol{\theta}|x)$ is the posterior density function of $\boldsymbol{\theta}$ given $\boldsymbol{x}$ and $\boldsymbol{x} = (x_1, ..., x_r)$.
To obtain $(1-\tau)100\%$ BPI for x_a^*, say (L, U), we solve the following two nonlinear equations, numerically,

$$P(X_a^* > L|\ \boldsymbol{x}) = \int_L^\infty f^*(x_a^*|\ \boldsymbol{x})dx_a^* = 1 - \frac{\tau}{2}, \tag{2.3}$$

$$P(X_a^* > U|\ \boldsymbol{x}) = \int_U^\infty f^*(x_a^*|\ \boldsymbol{x})dx_a^* = \frac{\tau}{2}. \tag{2.4}$$

Eqs. (2.3) and (2.4) can be solved by using Newton- Raphson iteration form as follows

$$L_{j+1} = L_j - \frac{\int_{L_j}^\infty f^*(x_a^*|\ \boldsymbol{x})dx_a^* - (1 - \frac{\tau}{2})}{-f^*(L_j|\ \boldsymbol{x})}, \tag{2.5}$$

$$U_{j+1} = U_j - \frac{\int_{U_j}^\infty f^*(x_a^*|\ \boldsymbol{x})dx_a^* - \frac{\tau}{2}}{-f^*(U_j|\ \boldsymbol{x})}, \tag{2.6}$$

where the initial values L_0, U_0 can be taken equal to x_r.
The integrals in (2.5) and (2.6) can be obtained using $Mathematica8$.

3 Bayesian Prediction Intervals of Future Observations in Case of Two-sample Scheme

Let $X_1 < X_2 < ... < X_r$ and $Z_1 < Z_2 < ... < Z_m$ represent informative (type II censored) sample from a random sample of size n and a future ordered sample of size m, respectively. It is assumed that the two samples are independent and drawn from a population with $pdf\, f_X(x)$, $cdf\, F_X(x)$ and $rf\, R(x)$. Our aim is to obtain the $BPI's$ for $Z_s, s = 1, 2, ..., m$. The conditional density function of Z_s, given the vector of parameters $\boldsymbol{\theta}$, is

$$g_2(z_s|\boldsymbol{\theta}) \propto [1 - R(z_s)]^{(s-1)}[R(z_s)]^{m-s} f_X(z_s|\boldsymbol{\theta}), z_k > 0, \tag{3.1}$$

where $\boldsymbol{\theta}$ is the vector of parameters.
The predictive density function is given by

$$g_2^*(z_s|x) = \int_{\boldsymbol{\theta}} g_2(z_s|\boldsymbol{\theta})\pi^*(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta}, z_s > 0, \tag{3.2}$$

where $\pi^*(\boldsymbol{\theta}|\boldsymbol{x})$ is the posterior density function of
$\boldsymbol{\theta}$ given $\boldsymbol{x}$ and $\boldsymbol{x} = (x_1, ..., x_r)$.
A $(1-\tau)100\%$ BPI for z_s is an interval (L, U) such that

$$P(Z_s > L|\boldsymbol{x}) = \int_L^\infty g_2^*(z_s|\boldsymbol{x})dz_s = 1 - \frac{\tau}{2}, \tag{3.3}$$

$$P(Z_s > U|\boldsymbol{x}) = \int_U^\infty g_2^*(z_s|\boldsymbol{x})dz_s = \frac{\tau}{2}. \tag{3.4}$$

Eqs. (3.3) and (3.4) can be solved by using Newton- Raphson iteration form as follows

$$L_{j+1} = L_j - \frac{\int_{L_j}^\infty g_2^*(z_s|\ \boldsymbol{x})dz_s - (1 - \frac{\tau}{2})}{-g^*(L_j|\ \boldsymbol{x})}, \tag{3.5}$$

$$U_{j+1} = U_j - \frac{\int_{U_j}^{\infty} g_2^*(z_s|\, \boldsymbol{x}) dz_s - (1 - \frac{\tau}{2})}{-g^*(U_j|\, \boldsymbol{x})}, \tag{3.6}$$

where the initial values L_0, U_0 can be taken equal to z_{s-1}.
The integrals in (3.5) and (3.6) can be obtained using $Mathematica 8$.

In the previous two sections, we will use the likelihood function of the form

$$L(\beta, \tau, \lambda|\boldsymbol{x}) \propto \tau^r \exp[-(n - r + 1)\, \tau\, x_r^{\beta}\, e^{\lambda\, x_r}] \left(\prod_{i=1}^{r} (\beta + \lambda\, x_i)\, x_i^{\beta-1} \exp(\lambda\, x_i)\right) \times \left(\prod_{i=1}^{r-1} \exp[-\tau\, x_i^{\beta}\, e^{\lambda\, x_i}]\right), \tag{3.7}$$

and the trivariate prior *pdf* suggested by Ateya[3] which of the form

$$\pi(\beta, \tau, \lambda) \propto \frac{1}{\Gamma(\beta)} \beta^{c_1+c_3-1}\, \tau^{\beta+c_3-1}\, \lambda^{\beta-1} \exp[-\beta\, (\tau + c_2) - \tau\, \lambda], \quad \beta > 0,\ \tau > 0,\ \lambda > 0,\ (c_1 > 0,\ c_2 > 0,\ c_3 > 0), \tag{3.8}$$

where c_1, c_2 and c_3 are the prior parameters (also known as hyperparameters).
Then, the posterior density function can be written in the form

$$\pi^*(\beta, \tau, \lambda|\, \boldsymbol{x}) = \frac{A}{\Gamma(\beta)} \beta^{c_1+c_3-1}\, \tau^{\beta+r+c_3-1}\, \lambda^{\beta-1} \exp[-\beta\, (\tau + c_2) - \tau\, (\lambda + (n - r + 1)\, x_r^{\beta}\, e^{\lambda\, x_r})] \times \left(\prod_{i=1}^{r} (\beta + \lambda\, x_i)\, x_i^{\beta-1} \exp(\lambda\, x_i)\right) \left(\prod_{i=1}^{r-1} \exp[-\tau\, x_i^{\beta}\, e^{\lambda\, x_i}]\right), \quad \beta > 0,\ \tau > 0,\ \lambda > 0,\ (c_1 > 0,\ c_2 > 0,\ c_3 > 0), \tag{3.9}$$

where A is a normalizing constant.

4 Results

4.1 Simulated Results

A 95% one-sample BPI of the future observations, $x_a^*, a = 1, 2, 3$, are obtained. Also, a 95% two-sample BPI for the future observations, $z_s, s = 1, 2, 3$, are obtained as follows:

1. For given set of prior parameters ($c_1 = 1.5, c_2 = 0.5, c_3 = 2.5$), the generated population parameters are $\beta = 2.1, \tau = 1.77$ and $\lambda = 3.4$,

2. Using the generated parameters, we generate a sample of size 40 of upper ordered observations and we will take the censoring values $r = 10, 15, 20$,

3. A 95% one-sample BPI for the future ordered observation $x_a^*, a = 1, 2, 3$, are obtained,

4. Using the generated ordered sample in 2, a 95% two-sample BPI for the future ordered observation $z_s, s = 1, 2, 3$ are obtained,

In our study,
Table(1) displays the one-sample BPI and the coverage percentage of the a^{th} future ordered observation ,
Table(2) displays the two-sample BPI and the coverage percentage of the s^{th} future ordered observation,
Table(1): 95 % one-sample BPI for future upper ordered observation $x_a^*, a = 1, 2, 3$.
1-Percentage coverage of the BPI.
2-BPI for x_a^*.
3- Length of the BPI.

(n, r)		x_1^*	x_2^*	x_3^*
	1	96.22%	97.19%	98.23%
(40, 10)	2	(0.1432,0.5421)	(0.4658,1.0104)	(0.9446,1.8142)
	3	0.3989	0.5446	0.8696
	1	95.87%	96.38%	97.47%
(40, 15)	2	(0.8007,1.1133)	(1.1047,1.5511)	(1.3421,2.1210)
	3	0.3126	0.4464	0.7789
	1	95.24%	95.85%	96.54%
(40, 20)	2	(1.3309,1.6118)	(1.5400,1.9164)	(1.8342,2.5340)
	3	0.2809	0.3764	0.6998

Table(2): 95% two-sample BPI for future upper ordered observation $z_s, s = 1, 2, 3$.
1-Percentage coverage of the BPI.
2-BPI for z_s.
3- Length of the BPI.

(n, r, N)		z_1	z_2	z_3
	1	97.21%	98.64%	98.98%
(40, 10, 15)	2	(0.0918,0.3105)	(0.2917,0.6034)	(0.5906,1.2220)
	3	0.2187	0.3117	0.6314
	1	96.77%	97.22%	98.15%
(40, 15, 15)	2	(0.1905,0.3732)	(0.3203,0.5551)	(0.4809,0.8590)
	3	0.1827	0.2351	0.3781
	1	95.94%	96.00%	97.25%
(40, 20, 15)	2	(0.2053,0.3144)	(0.3011,0.4820)	(0.4614,0.6915)
	3	0.1091	0.1809	0.2301

In one and two-sample schemes, observe the following:

1. For fixed sample size n and a certain x_a^*, the length of the BPI and its Percentage coverage decrease, by increasing r,

2. For fixed sample size n and r, and for a certain x_a^*, the length of the BPI and its Percentage coverage increase, by increasing a,

3. For fixed sample size n and a certain z_s, the length of the BPI and its Percentage coverage decrease, by increasing r,

4. For fixed sample size n and r, and for a certain z_s, the length of the BPI and its Percentage coverage increase, by increasing s.

4.2 Data Analysis And Applications

In this subsection, two real data sets are introduced as illustrative applications. These real data sets are analyzed using Weibull and modified Weibull distributions. A comparison study is carried out between the two models based on $K-S$ test statistic, AIC and BIC. 95% One and two-sample BPI of the first three unknown observations are obtained.

4.2.1 Application 1

The first real data set is from Lawless [13] and represents the breakdown time of an insulating fluid between electrodes at a voltage of 34 kv (minutes). The 19 times to breakdown are: 0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 3.16, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71 and 72.89
A model suggested by engineering considerations is that, for a fixed voltage level, time to breakdown has a Weibull distribution. These real data are analyzed using $Weibull(\alpha, \beta)$ distribution and using $MWD(\beta, \tau, \lambda)$. The $K-S$, AIC and BIC are computed for each model in Table 3. Observe that the $MWD(\beta, \tau, \lambda)$ model has smaller $K-S$, AIC and BIC values which means that the $MWD(\beta, \tau, \lambda)$ model fits the real data set better than $Weibull(\alpha, \beta)$ model. Based on type-II censored sample with censoring value $r = 12$, the 95% one-sample BPI for the unknown observations x_{13}, x_{14} and x_{15} are $(7.9165, 8.7414), (11.1083, 13.6201)$ and $(30.4132, 37.7612)$, respectively. Also, the 95% two-sample BPI for the unknown observations z_1, z_2 and z_3 are $(0.2314, 0.4102), (0.4415, 0.7712)$ and $(0.79109, 1.5015)$.

Table 3:-$MLE's$ of the parameters, the associated $K-S$, AIC and BIC values.

Model	$MLE's$	$K-S$	AIC	BIC
$Weibull(\alpha, \beta)$	$\hat{\alpha} = 0.7708, \hat{\beta} = 0.1452$	0.1981	78.991	81.824
$MWD(\beta, \tau, \lambda)$	$\hat{\beta} = 0.7793, \hat{\tau} = 0.1416, \hat{\lambda} = 0.0015$	0.1319	76.111	78.944

4.2.2 Application 2

The second real data set from Nicholas and Padgett[14]. The data concerning tensile strength of 100 observations of carbon fibers, they are: 3.7, 3.11, 4.42, 3.28, 3.75, 2.96, 3.39, 3.31, 3.15, 2.81, 1.41, 2.76, 3.19, 1.59, 2.17, 3.51, 1.84, 1.61, 1.57, 1.89, 2.74, 3.27, 2.41, 3.09, 2.43, 2.53, 2.81, 3.31, 2.35, 2.77, 2.68, 4.91, 1.57, 2.00, 1.17, 2.17, 0.39, 2.79, 1.08, 2.88, 2.73, 2.87, 3.19, 1.87, 2.95, 2.67, 4.20, 2.85, 2.55, 2.17, 2.97, 3.68, 0.81, 1.22, 5.08, 1.69, 3.68, 4.70, 2.03, 2.82, 2.50, 1.47, 3.22, 3.15, 2.97, 2.93, 3.33, 2.56, 2.59, 2.83, 1.36, 1.84, 5.56, 1.12, 2.48, 1.25, 2.48, 2.03, 1.61, 2.05, 3.60, 3.11, 1.69, 4.90, 3.39, 3.22, 2.55, 3.56, 2.38, 1.92, 0.98, 1.59, 1.73, 1.71, 1.18, 4.38, 0.85, 1.80, 2.12, 3.65.
These real data are analyzed using $Weibull(\alpha, \beta)$ distribution and using $MWD(\beta, \tau, \lambda)$. The $K-S$, AIC and BIC are computed for each model in Table 4. Observe that the $MWD(\beta, \tau, \lambda)$ model has smaller $K-S$, AIC and BIC values which means that the $MWD(\beta, \tau, \lambda)$ model fits the real data set better than $Weibull(\alpha, \beta)$ model.

Based on type-II censored sample with censoring value $r = 10$, the 95% one-sample BPI for the unknown observations x_{11}, x_{12} and x_{13} are $(1.2172, 1.3662), (1.2322, 1.4301)$ and $(1.2032, 1.4812)$, respectively. Also, the 95% two-sample BPI for the unknown observations z_1, z_2 and z_3 are $(0.1022, 0.1902), (0.1702, 0.29155)$ and $(0.25654, 0.5408)$.

Table 4:-$MLE's$ of the parameters, the associated $K-S$, AIC and BIC values.

Model	$MLE's$	$K-S$	AIC	BIC
$Weibull(\alpha,\beta)$	$\hat{\alpha}=0.0391, \hat{\beta}=3.0629$	0.2103	258.659	266.475
$MWD(\beta,\tau,\lambda)$	$\hat{\beta}=2.2800, \hat{\tau}=0.0323, \hat{\lambda}=0.3477$	0.1211	255.036	262.852

5 Conclusions

In this paper, the Bayesian prediction problem of future observations are studied under MWD. A simulation study is carried out to study the behavior of the length and the percentage coverage of the $BPI's$ under various n and r.

Two real data sets from Weibull(α, β) distribution are introduced and analyzed using MWD. A comparison is carried out between the mentioned distributions based on the corresponding Kolmogorov-Smirnov $(K-S)$ test statistic to emphasize that the MWD model fits the data better than the other model. Based on a type-II censored sample from the real data, the $BPI's$ of the remaining observations are obtained in case of the one and two- sample schemes.

References

[1] **S.F. Ateya**, Bayesian prediction intervals under bivariate truncated generalized Cauchy distribution, *J. Statist. Res. Iran 7(2010), 133-153.*

[2] **S.F. Ateya**, Prediction under generalized exponential distribution using MCMC algorithm, *International Mathematical Forum 6(2012), 3111-3119.*

[3] **S.F. Ateya**, Bayesian Prediction Intervals of Future Nonadjacent Generalized Order Statistics from Generalized Exponential Distribution Using Markov Chain Monte Carlo Method, *Applied Mathematical Sciences, 6(2012), 1335 - 1345.*

[4] **S.F. Ateya**, Estimation Under Modified Weibull Distribution Based on Right Censored Generalized Order Statistics, *J. Applied Statistics, 40(2013), 2720-2734.*

[5] **S.F. Ateya and M.M. Rizk**, Bayesian Prediction Intervals of Future Generalized Order Statistics Under a Finite Mixture of Generalized Exponential Distributions, *Applied Mathematical Sciences, 7(2013), 1575 - 1592*

[6] **S. Geisser**, *Predictive Inference: An Introduction*, Chapman and Hall, London, 1993.

[7] **T. Gotoh, M. Fukuhara, and K. I. Kikuchi**, Mathematical model for change in size distribution of baculovirus-infected Sf-9 insect cells, The 3rd WSEAS International Conference on Cellular and Moleculaz Biology, Biophysics and Bioengineering, 2007, 25-29.

[8] **E.J. Green, F.A. Jr. Roesh, A.F.M. Smith and W. E. Strawderman**, Bayes estimation for the three parameter Weibull distribution with tree diameters data, *Biometrics, 50(1994), 254-269.*

[9] **A.M. Hossain and W.J. Zimmer**, Comparison of estimation methods for Weibull parameters: complete and censored samples, *J. Statist. Comput. Simulation, 73 (2003), 145-153.*

[10] **Z.F. Jaheen and M.M. Al Harbi**, Bayesian Estimation for the Exponentiated Weibull Model via Markov Chain Monte Carlo Simulation, *Commun. Statist.-Simul. Comput., 40(2011), 532-543.*

[11] **N.L. Johnson, S. Kotz and N. Balakrishnan**, *Continuous Univariate Distributions*,Wiley, NewYork,(1994).

[12] **C.D. Lai, M. Xie and D.N. Murthy**, A modified Weibull distribution, *IEEE Trans. Reliab., 52(2003), 33-37.*

[13] **J. F. Lawless**, *Statistical Model and Methods for Lifetime Data*, Wiley, New York (1982).

[14] **M.D. Nicholas and W. J. Padgett**, A bootstrap control chart for Weibull percentiles, *Quality and Reliability Engineering International, 22(2006), 141-151.*

[15] **V. Niola, R. Oliviero and G. Quaremba**, The application of wavelet transform for estimating the shape parameter of a Weibull pdf, Proceedings 5th Wseas International Conference on Signal Processing, (2005), 126-130.

[16] **A. Shamilov, I. Usta and Y.M. Kantar**, The Distribution of Minimizing Maximum Entropy: Alternative to Weibull distribution for Wind Speed,WSEAS Transactions on Mathematics, 5(2006), 695-700.

[17] **K.S. Sultan**, Record values from the modified Weibull distribution and applications, *International Mathematical Forum, 2(2007), 2045-2054.*

[18] **S.K. Upadhyaya, and A. Gupta**, A Bayes analysis of modified Weibull distribution via Markov chain Monte Carlo simulation, *J. Statist. Comput. Simul. 80 (2010), 241-254.*

[19] **P. Vasile, P. Eugenia and C. Alina**, Bayes estimators of modified Weibull distribution parameters using Lindley's approximation, WSEAS Transactions on Mathematics, 9(2010), 539-549.

[20] **R. Vicen-Bueno, M. Rosa-Zurera, L. Cuadra- Rodriguez and D. De La Mata-Moya**, Models of radar clutter from the Weibull distribution, Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation, (2006), 376-380.

[21] **A. Zaharim, S.K. Najid, A.M. Razali and K. Sopian**, The Suitability of Statistical Distribution in Fitting Wind Speed Data, WSEAS Transactions on Mathematics, 7(2008), 718-727.