

The Conditional Estimation for related Weibull parameters Under Type-II Censoring

J. kazempoor¹, A. Habibirad^{1*}

1. Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

In this paper, the conditional estimation of the Weibull and its related parameters are introduced. Some interesting properties of this estimator in contrast with the well-known maximum likelihood estimators have been investigated. This task is done under the famous sampling plan type-ii censoring scheme. Because of the complex behavior in the calculation of the likelihood function of the presented scheme in this situation without loss of generality, this problem fixed with the Gumbel (log-Weibull) model. The one to one transformation between these models and satisfying in their parameters enabling us for utilizing this alternative model. Finally, the comparison of this method and maximum likelihood estimation are provided through some numerical results.

Keywords: conditional estimation, Gumbel distribution, Weibull mode.

Introduction

Weibull distribution is commonly used in reliability studies. It is used for modeling the observed failures of many different types of components and phenomena. In some situations, statistical inference about its parameters in quality control and some engineering problems is of interest. Depending on a kind of sampling plan, and based on the maximum likelihood estimation, the performance of estimating in this model can differ. On one hand, the observation number in the engineering problems is almost few. On the other hand, it is clear that the maximum likelihood estimators have not a good performance in such situations. The behavior of Bayesian estimators is often like the maximum likelihood estimators. One of the most applicable sampling strategies in engineering problems is the type-II censored order statistics scheme. For a comprehensive, perfect, and complete discussion on censored sampling scheme, the reader can refer to Balakrishnan and Cramer [2]. The main problem of estimating parameters under this scheme is a major fault occurrence, especially for large scale parameters in confrontation with Weibull models. A routine solution of this end is utilizing Gumbel or extreme value distribution, which have the capacity of one to one transformation between corresponding random variables and related parameters. Yet, the problem of bad performance of maximum likelihood estimators (MLEs) is still holding on but a little better than previous estimating.

However, in the situations of having a location-scale family of distributions for sampling, some authors proposed a conditional method of estimating. This method was introduced by Lawless [10] and was applied in some inferences problem by Lawless [11] and Lawless [9]. A generalize conditional strategy of estimating in the location-scale family of distribution done by Viveros and Balakrishnan [17]. In continuing, the conditional Bayes estimators have been discussed in some quality control disciplines by Haghghi [5]. In addition, in this concept the conditional strategy has been used alone in Haghghi et.al. [7] and Haghghi and Costagliola [6]. The aforementioned paper uses the conditional method for the estimation of the Weibull parameters and introduces a new control chart for the Weibull quantile function. Utilizing conditional estimators arising from the point that if $U(\cdot)$ and $T(\cdot)$ be two statistics proposed for estimating the parameter, Thus $E[TU]$ is a new unbiased statistics that have a small variance than $T(\cdot)$. It can be shown that the variance can be small and smaller until the $U(\cdot)$, be sufficient statistics, but in the case of non-existence sufficient statistics like the Weibull model, it is straightforward that conditioning on an ancillary variable can improve our estimations.

The combined Bayes and conditional method have effectively used in Haghghi [5] for quality control disciplines focusing on constructing the new control charts for the Weibull shape and quantile parameter. In the Bayes method, choosing a good prior distribution is so important but the presented method has focused on improper Jeffreys prior to or proportions this prior. The extension of this prior has been considered in this study and the best of these priors has been identified. In

* Corresponding Author Email: ahabibi@um.ac.ir

continue, the efficient performances of the presented strategy are established by some comparison results through mean squared errors (MSE) and mean absolute errors (MAE). In addition, this method is applicable for estimating any related parameters for the Weibull distribution and has not an exception for estimating particular parameters.

For text shortening and page boundaries, only formulas and relationships are presented in the Gumbel distribution. The derivation of the results for the Weibull distribution is accomplished with a simple transformation of linear, and conversions are kept one by one in the parameters (Haghighi [5]). It is also worth mentioning that if one can find a strategy with the better performances in the Gumbel model, it is also held on for the Weibull distribution. Moreover, the Gumbel model belongs to the location-scale family of distributions which help us in constructing the ancillary statistics. Therefore, without diminishing loss of generality, in section 2, the required relationships in the Gumbel distribution are shown, and in Section 3, the numerical comparison of the two methods presented in deducing the parameters of the location and the scale for $u = 5, 0, 5, 10$ and $\sigma = 1, 5, 10$

Notations and Models

Suppose that Y_1, Y_2, \dots, Y_n are some random variables following to the Weibull distribution with shape parameter α and scale parameter β . The probability density function of this distribution is

$$f_{Y_i(Y)} = \frac{\alpha y^{\alpha-1}}{\beta^\alpha} e^{-\left(\frac{y}{\beta}\right)^\alpha}$$

$$i = 1, 2, \dots, n, (x, \alpha, \beta) \in R^+$$

For simplicity the notation of these representation is considered as $Y \sim W(\alpha, \beta)$.

Taking into account the one-to-one transformation $X_i = \log(Y_i)$. Here, suppose that X_1, X_2, \dots, X_n are some random variables following to the Gumbel distribution with parameters μ and σ . The probability density function of this distribution is

$$f_{X_i=\frac{1}{\sigma}} = \frac{x-\mu}{\sigma} e^{-\frac{x-\mu}{\sigma}}$$

$$i = 1, 2, \dots, n, (x, \mu) \in R, \sigma \in R^+$$

Where $\mu = \log(\alpha)$, and $\sigma = \frac{1}{\beta}$

For simplicity the notation of these representation is considered as $X \sim G(\mu, \sigma)$.

The joint probability density function of first r ordinary order statistics $X_{1:n}, X_{2:n}, \dots, X_{r:n}$ arising from independent and identical random variables X_1, X_2, \dots, X_n , is (Arnold et.al. [1])

$$f_{X_{1:n}, X_{2:n}, \dots, X_{r:n}}(x_1, x_2, \dots, x_r) = \frac{1}{\sigma^r} \exp\left(\sum_{j=1}^r \frac{x_j - \mu}{\sigma}\right) \times \exp\left(-\left((n-r)e^{\frac{x_r - \mu}{\sigma}} + \sum_{j=1}^r e^{\frac{x_j - \mu}{\sigma}}\right)\right) \quad (1)$$

It is noteworthy that this sampling method can be considered as a special case of progressively type-II censored order statistics which named as type-II censored order statistics sampling scheme. However, the maximum likelihood estimators (MLEs) based on maximizing likelihood function 1, called by μ and $\hat{\sigma}$ respectively referred as MLE of μ and σ . In continuing another method of estimating these parameters according to the Bayesian conditional strategy are introduced. The performance of this manner of estimating is constructed as follows.

Let $\pi(\mu, \sigma) = \frac{1}{\sigma^m}$, $m \in R^+$ denote the improper prior distribution for two location and scale parameters and immediately from 1, the posterior distribution $\pi(\mu, \sigma | x)$ satisfies.

$$\pi(\mu, \sigma | x) = \frac{1}{\sigma^{r+m}} \exp\left(\sum_{j=1}^r \frac{x_j - \mu}{\sigma}\right) \times \exp\left(-\left((n-r)e^{\frac{x_r - \mu}{\sigma}} + \sum_{j=1}^r e^{\frac{x_j - \mu}{\sigma}}\right)\right)$$

Now, consider two ancillary statistics $Z_1 = \frac{\hat{\mu} - \mu}{\hat{\sigma}}$ and $Z_2 = \frac{\hat{\sigma}}{\sigma}$ and moreover pivotal quantity $a_i = \frac{x_i - \hat{\mu}}{\hat{\sigma}}$, $i = 1, 2, \dots, r$ According to these notations

$$f(Z_1, Z_2) | (a_1, a_2, \dots, a_r)(Z_1, Z_2) \propto Z_2^{r+m-2} \exp\left(\sum_{j=1}^r (Z_1 + a_j) Z_2\right) \times \exp\left(-\left((n-r)e^{(Z_1+a_r)Z_2} + \sum_{j=1}^r e^{(Z_1+a_j)Z_2}\right)\right),$$

or equivalently

$$f(Z_1, Z_2) | (a_1, a_2, \dots, a_r)(Z_1, Z_2) = CZ_2^{r+m-2} \exp\left(\sum_{j=1}^r (Z_1 + a_j) Z_2\right) \times \exp\left(-\left((n-r)e^{(Z_1+a_r)Z_2} + \sum_{j=1}^r e^{(Z_1+a_j)Z_2}\right)\right)$$

Where C is normalizing constant such that

$$\frac{1}{C} = \int_R \int_{R^+} Z_2^{r+m-2} \exp\left(\sum_{j=1}^r (Z_1 + a_j) Z_2\right) \times$$

$$\exp\left(-\left((n-r)e^{(Z_1+a_r)Z_2} + \sum_{j=1}^r e^{(Z_1+a_j)Z_2} dz_2 dz_1\right)\right).$$

In order to calculate the Bayesian estimator of Z_1 and Z_2 , the full conditional probability density function of these parameters are required which can be derived in the following

$$f_{Z_1}(Z_1) \propto e^{rZ_1Z_2} - e^{Z_1Z_2} [(n-1)e^{a_1Z_2} + \sum_{j=1}^r e^{a_jZ_2}], Z_1 \in R$$

And

$$f_{Z_2}(Z_2) \propto Z_2^{r+m-2} \times e^{Z_2(\sum_{j=1}^r (Z_1+a_j))} - [(n-r)e^{Z_2(Z_1+a_r)} - \sum_{j=1}^r e^{Z_2(Z_1+a_j)}], Z_2 \in R^+$$

It is obvious that finding a closed normalizing constant for these distribution cannot be manipulating essay and we should apply the Gibbs sampling method for generating random variables from these two probability density functions. Moreover, because Z_1 and Z_2 are the linear combination of parameters μ and σ , the Bayes estimator of these new parameters can keep their Bayesian features under the first and the last loss functions. However, under the second loss function,

some challenges have appeared. It is perfectly understood that with some slight mathematical calculations the Bayes estimator of a linear combination $\delta_{2B}^*(a\theta + b)$ under the LinEx loss function is $\delta_{2B}^*(a\theta + b)$ (Lehman and Casella [12]).

Finally, it is worth mentioning that the similar distribution of Z_1 is

$$u(x) = \frac{\beta\lambda^\beta}{\Gamma(\frac{\alpha}{\beta})} e^{ax} - \lambda e^{Bx}, \quad x \in R, (\alpha, \beta, \lambda) \in R^+,$$

where it's cumulative distribution function is

$$U(x) = \int_0^{e^{Bx}} \frac{\beta\lambda^\beta}{\Gamma(\frac{\alpha}{\beta})} t^{\frac{\alpha}{\beta}-1} e^{-\lambda t} dt \tag{2}$$

Hence, the generating random samples from f_{z_1} , (z_1) can be easily done but for f_{z_2} , (z_2) some complicated problems has been appeared. For calculation conditional Bayes estimator of Z_1 and Z_2 , three loss functions have considered and these estimators have compared with their MLE's.

- Absolute Error Loss Functions (AEL)
 $L(\delta, \theta) = |\delta - \theta|$
 $\delta_{1B}^*(\theta) =$ median of θ in posterior density function
- Linear Exponential Loss Functions (LinEx)
 $L(\delta, \theta) = e^{c(\delta-\theta)} - c(\delta - \theta) - 1, c \in R^+$
 $\delta_{2B}^*(\theta) = \frac{-\log(E[e^{-c\theta}])}{c}$

In present study $c = 1$ has been considered.

- Squared Error Loss Functions (SE)
 $L(\delta, \theta) = (\delta - \theta)^2$
 $\delta_{3B}^*(\theta) =$ mean of posterior density function

In each case, $\delta_B^*(\theta)$, represent the corresponding Bayesian estimator to each of loss functions (Lehman and Casella [12]).

Numerical comparison of two presented methods

In this section, we provide some comparison results for the performance of two presented methods of estimating parameters in Gumbel distribution. The location and scale parameters of the Gumbel model or equivalently the shape and scale parameters of the corresponding Weibull distribution is our target parameters for estimating. Obviously the same task is also applied for estimating another important parameter like as reliability, density, survival, and hazard rate functions. The MLE estimators of these parameters can be easily derived by maximizing the likelihood function 1, but the calculation of Bayes estimators have some conflicts.

In continue, the famous algorithm is described and the method of derivation conditional estimation is explained through statistical software R, Core Team, [16].

Metropolis et al. [13], introduced the Metropolis-Hastings (M-H) algorithmic rule in the program as a general Mont Carlo Markov Chain (MCMC) technique

and afterward Hastings (Hastings [8]), expanded the M-H algorithm. One can apply the M-H algorithm to get a random sample from any subjectively complicated target distribution of any dimension that is known up to a normalizing constant. Gibbs sampling method is a particular instance of the MCMC method. It can be utilized to generate a sample from the full conditional probability distributions of two or more random variables. Gibbs sampling requires decomposing the joint posterior distribution into full conditional distributions for each parameter and then sampling from them. We propose using the Gibbs sampling plan to generate a sample from the posterior density functions f_{z_1} (z_1) and f_{z_2} (z_2) in turn compute the Bayesian estimates under given loss functions (Soliman, et.al. [14,15]).

Generating random samples such that distributed as similar as Z_1 is expressed previously. For such a same task in Z_2 we propose the following steps:

- I. fix values $n = 15, r = 8, R_1 = 0, R_1 = R_2 = \dots = R_7 = 0$, and $R_8 = 7$.
- II. Utilizing the given algorithm in Balakrishnan and Sandhu [3], to generating type-II censored order statistics arising from independent and identical Gumbel random variables under the scheme which introduced in previous step.
- III. Calculate MLEs of parameters μ and σ based on samples that generated in previous step, say $\hat{\mu}$ and $\hat{\sigma}$ respectively.
- IV. Construct $a_i = \frac{x_i - \hat{\mu}}{\hat{\sigma}}, i = 1, 2, \dots, 8$
- V. Assume that a fix value of z_2 say $z_2 = 1$.
- VI. Generate a random sample z_1 based on a cdf 2, with parameters $\alpha = 8Z_2, \beta = Z_2$ and $\lambda = 7e^{\alpha_8 - z_2} + \sum_{j=1}^8 e^{\alpha_j z_2}$
- VII. Generate a new random sample t based on a cdf 2, with parameters $\alpha = \sum_{j=1}^8 (z_1 + a_j), \beta = z_1 + a_8$ and $\lambda = 7$.
- VIII. For $M = 20000$ and $N = 100000$ calculate CC

$$\frac{\sum_{j=20001}^{100000} t^{m+6} e^{-\sum_{j=1}^8 e^{t(z_1+a_j)}}}{\sum_{j=20001}^{N-M} z_2^{r+m-2}}. \text{Hence,}$$

$$f_{z_2}(z_2) = \frac{1}{CC} z_2^{r+m-2} \times e^{z_2(\sum_{j=1}^r (z_1+a_j)) - [(n-r)e^{z_2(z_1+a_r)} - \sum_{j=1}^r e^{z_2(z_1+a_j)}]}, z_2 \in R^+$$

IX. It is clear that z_2 belongs to the log concave family of distribution. Therefore, using the given algorithm of this family of distributions in Devroye [4], and generate one random variable from this pdf.

X. Call the generated random variable Z_{1s} and Z_{2s} . Moreover repeat this step for $S = 1, 2, \dots, 100000$ and afterward delete the first 20000th and construct new vectors Z_{1s} and Z_{2s} with length 80000. It is straightforward to see that

$$\delta_{1B}^*(Z_i) = \frac{Z_{(i(60000))} + Z_{(i(60001))}}{2}, \delta_{2B}^*(Z_i) = -\log \left[\sum_{s=20001}^{100000} \frac{e^{-Z_{1s}}}{80000} \right], \text{ and } \delta_{3B}^*(Z_i) = \sum_{s=20001}^{100000} \frac{Z_{1s}}{80000}, i = 1, 2$$

Now, based on the above- mentioned examples, the Bayes estimators of Z_1 and Z_2 , can be calculated and subsequently the Bayes estimators of μ and σ .will be calculated .The number of repeated in each tables is 1000000.

Table 1. Estimation criterion 1

True values of parameters	Bias	MSE	MAE
$\mu = -5$ $\sigma = 1$	0.68680050 -0.04166361	0.63839000 0.21870780	0.70501050 0.38150950
$\mu = -5$ $\sigma = 5$	-0.93424835 0.05166332	4.32319800 4.49697100	1.62188400 1.62527700
$\mu = -5$ $\sigma = 10$	-2.88157680 0.10755720	22.1155700 18.0756800	3.71012200 3.25669600
$\mu = 0$ $\sigma = 1$	0.68716310 -0.04246840	0.63873880 0.21780420	0.70513560 0.38138020
$\mu = 0$ $\sigma = 5$	-0.93734297 0.05174331	4.32370200 4.50535600	1.62178800 1.62830300
$\mu = 0$ $\sigma = 10$	-2.87358300 0.10011000	22.0749800 17.9499900	3.70312600 3.24967700
$\mu = 5$ $\sigma = 1$	0.68562740 -0.04257320	0.63654430 0.21866210	0.70387780 0.38165130
$\mu = 5$ $\sigma = 5$	-0.93823280 0.05137040	4.33715900 4.52099100	1.62496700 1.62886800
$\mu = 5$ $\sigma = 10$	-2.87603700 0.10698600	22.0890500 18.0351600	3.70735800 3.25654700
$\mu = 10$ $\sigma = 1$	0.65110020 1.13773740	0.56575350 1.44024000	0.67037940 1.13773740
$\mu = 10$ $\sigma = 5$	-0.93614990 1.05268570	4.32951100 5.60223000	1.62322800 1.71455800
$\mu = 10$ $\sigma = 10$	-2.87460300 1.11072600	22.0702600 19.2965000	3.70402900 3.24181500

Table 2. Estimation criterion 2

True values of parameters	Bias	MSE	MAE
$\mu = -5$ $\sigma = 1$	0.8722585 -0.0479138	0.791762 0.4196371	0.8722585 0.1637934
$\mu = -5$ $\sigma = 5$	-0.89537 -0.04946586	0.8310668 0.04220974	0.89537 0.1640008
$\mu = -5$ $\sigma = 10$	-1.176859 0.09995236	1.385527 0.04999095	1.176859 0.1791974
$\mu = 0$ $\sigma = 1$	0.6835907 -0.04048524	0.5058638 0.4148003	0.6836182 0.1622532
$\mu = 0$ $\sigma = 5$	-0.9004636 0.04918261	0.839695 0.04233654	0.9004636 0.1642744
$\mu = 0$ $\sigma = 10$	-1.17571 0.09990332	1.382866 0.04973746	1.17571 0.1786718
$\mu = 5$ $\sigma = 1$	0.6806113 -0.04001441	0.5016124 0.04170052	0.6806389 0.163145
$\mu = 5$ $\sigma = 5$	-0.900369 0.05015226	0.8393726 0.04250714	0.900369 0.1644665
$\mu = 5$ $\sigma = 10$	-1.176583 0.01089497	1.384879 0.4598741	1.176583 0.5843977
$\mu = 10$ $\sigma = 1$	0.0729195 0.125118	0.4612334 0.4633722	0.5853499 0.5870599
$\mu = 10$ $\sigma = 5$	-0.1040864 0.1218691	0.4626862 0.4643177	0.586616 0.5874308
$\mu = 10$ $\sigma = 10$	-0.3145018 0.1265605	0.4995298 0.4652878	0.6156897 0.5883557

The last table denotes the behavior of Conditional Bayes estimators of the corresponding parameters under LinEx loss function. The similar tables for another loss functions are omitted. Finally, as you can see the new strategy of estimating has so good performance in comparison with MLES.

References

- [1] M. Modarres, M. Kaminskiz, and V. Krivstov, Reliability Engineering and Risk Analysis: A Practical Guide. Vol. 55: CRC press., 2009.
- [2] W. R. Blischke, M. D. N. Prabhakar, Case Studies in Reliability and Maintenance, Wiley Series in Probability and Statistics, 2003.
- [3] P.Fahlstrom, T. Gleason, Introduction to UAV Systems, Wiley press, 2012.
- [4] G.W. Hanks, G.T. Katt, R.H. Edwards, R.D. Shannon, 747 primary flight control system reliability and Maintenance, NASA, 1980.
- [5] Leigh P. Ackart, An evaluation of markov chain modeling for fa-18 aircraft readiness, M. Sc. Thesis, 1998.
- [6] M. Nanda and S. Rao, A Formal Method Approach to Analyze the Design of Aircraft Flight Control Systems, 3rd Annual IEEE International Systems Conference, Vancouver, Canada, 2009.
- [7] M. Pourgl-Mohammad, K. Sepanloo, K. Karimi, Hybrid Fault Tree Markov Chain (Hft-Mc) Probabilistic Risk Assessment Methodology with Application, ANS PSA 2011 International Topical Meeting on Probabilistic Safety Assessment and Analysis, 2011.
- [8] L. Xing and G. Levitin, BDD-based Reliability Evaluation of Phased-Mission Systems with Internal-External Common-Cause Failures, Reliab. Eng. Syst. Saf., vol. 112, pp. 145–153, 2013.
- [9] J. Mi, Y. Li, H. Huang, Y. Lio, X. Zhang, Reliability Analysis of Multi-State System with Common-Cause Based on Bayesian Networks, Maintenance and Reliability, vol. 15, no. 2, pp. 169–175, 2013.
- [10] J. Asghari, M. Pourgol-Mohammad, F. Salehpour, Improving Dynamic Fault Tree Method for Complex System Reliability Analysis: Case Study of a Wind Turbine, ASME 2015 International Mechanical Engineering Congress and Exposition, vol. 14, Houston, Texas, USA, 2015.
- [11] C. Kanga, L. Zhonga, Z. Haijun, Research on probabilistic safety analysis approach of flight control system based on Bayesian network, Procedia Engineering vol. 99, pp. 180-184, 2015.
- [12] R. Duan and J. Fan, Reliability evaluation of data communication system based on DFT under epistemic uncertainty, Hindawi Publishing Corporation, 2014.
- [13] H. R. Garoarsdottir, Reliability Analysis of the RB-211 Jet Engines Operated by Icelandair, M. Sc. Thesis, 2014.
- [14] E. G. Okafor & I. H. Eze, Failure analysis of a UAV flight control system using markov analysis, Nigerian Journal of Technology (NIJOTECH), vol. 35, no. 1, pp. 167-173, 2016.

- [15] H. Boudali, J.B. Dugan, a discrete-time Bayesian network reliability modeling and analysis framework, *Reliab. Eng. Syst. Saf.*, vol. 87, pp. 337-349, 2005.
- [16] K. Verma, A. Srividya, D. R. Karanki, *Reliability and Safety engineering*, Springer, 2010.
- [17] R. Duan, H. Zhou, J. Fan, Diagnosis strategy for complex systems based on reliability analysis and MADM under epistemic uncertainty, *Eksploatacja i Niezawodność – Maintenance and Reliability*, vol. 17, no. 3, pp. 345-354, 2015.
- [18] H. Boudali, J.B. Dugan, A New Bayesian Network Approach to Solve DFT, *Reliability and Maintainability Symposium (RAMS)*, IEEE, 2005.
- [19] K. D. Rao, V. Gopika, V. V. S. S. Rao, H. S. Kushwaha, A. K. Verma, and A. Srividya, Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. *Reliab. Eng. Syst. Saf.*, vol. 94, no. 4, pp. 872-883, 2009.