Some Statistical inferences on the records Weibull distribution using Shannon Entropy and Renyi Entropy

Gholamhossein Yari* and Rezvan Rezaei
Iran University of Science and Technology
e-mail: Yari@iust.ac.ir
e-mail: Rezvan_Rezaei@mathdep.iust.ac.ir

Abstract

In this paper, we discuss different estimators of the records Weibull distribution parameters and also we apply the Kullback-Leibler divergence of survival function method to estimate record Weibull parameters. Finally, these methods have been compared using Monte Carlo simulation and suggested good estimators.

1 Introduction

In the common practice, the communication theory is necessary in order to quantify and fully study the information mathematically. Information theory is a branch of applied mathematics, electrical engineering and computer science involving the quantification of information. limited information-theoretic ideas had been developed by Boltzmann (1896), Nyquist (1924), Hartley (1928, 1930) and etc. What is known today as the information theory has been founded by Shannon in 1948 through introducing $H(X)$ as the entropy of random variable X. He for the first time introduced the qualitative and quantitative model of communication as a statistical process underlying information theory. Later on, other measures were introduced such as Renyi entropy, relative entropy and Mutual information. Relative entropy (also Kullback-Leibler divergence) was first defined by Kullback and Leibler in 1951 as the directed divergence between two distributions.

Also, record values and associated statistics have widely been used in many real life applications, for example in sports, weather, business, etc. Record values are very important in case when observations are difficult to obtain or when they are being destroyed when observations are subjected to an experimental test. Observations are obtained by observing successive maximum (minimum) values. The term record value was first introduced by Chandler (1952).

In statistics, the Weibull distribution is one of the most important continuous probability distributions. It was, first introduced by Weibull in 1939 when he was studying the issue of structural strength. Teymori et al. (2012) introduce a point estimator for shape parameter of upper record Weibull distribution and use maximum-likelihood estimation (MLE) of scale parameter in their method. Also some inferential statistics based on Weibull distribution have been discussed by Yari et al. (2013).

In this paper we use Kullbak-Leibler method for estimation parameters of upper record Weibull and we discuss on point estimator method. The results show moment method estimation (MME) or Kullback-Leibler divergence of survival function estimation (DLS) of β are better than maximum-likelihood estimation (MLE) of β, for estimation α.

Now we introduce some basic definitions that play a central role the present paper, using the notation of Arnold et al. (1998) and Yari et al. (2013).

Definition 1.1 *Let* $Y_n = max\{X_1, X_2, ..., X_n\}$, $n \geq 1$. X_i *is an upper record of* $X_1, X_2, ..., X_n$ *if* $Y_i > Y_{i-1}$, $i \geqslant 1$. *By this definition* X_1 *is an upper record values.* we use the notation $X_{U(i)}$ for the i^{th} upper record statistic. The pdf of the $X_{U(i)}$ is obtained by

$$f_{X_{U(i)}}(x) = \frac{(-log[1 - F_X(x)])^{i-1} f_X(x)}{\Gamma(i)}, \quad x \in R, \quad i = 1, 2, ..., n, \tag{1}$$

where Γ is the gamma function.

Definition 1.2 *Suppose that* X *is a random variable with a pdf* $f(x)$ *and support* S_x *,Shannon entropy of* X *is defined*

$$h(X) = -\int_S f(x) log f(x) dx. \tag{2}$$

Definition 1.3 *Suppose that* X *is a random variable with a pdf* $f(x)$ *and support* S_x *,Renyi entropy of* X *is defined*

$$h(X, \gamma) = \frac{1}{1-\gamma} log \int_S f^{\gamma}(x) dx, \quad 0 < \gamma \neq 1. \tag{3}$$

Definition 1.4 *Let $X_1, X_2, ...$ be a sequence of positive, independent and identically distributed (iid) random variable with a non-increasing survival function $\overline{F}(x,\Theta) = P_\Theta(X > x)$ with support S_x and vector of parameters Θ. Define the* ***empirical survival function*** *of a random sample of size n by*

$$G_n(x) = \sum_{j=0}^{n-1}(1 - \frac{j}{n})I_{[X_{(j)},X_{(j+1)})}, \quad (4)$$

where I is the indicator function and $(0 = X_{(0)} \leqslant)X_{(1)} \leqslant X_{(2)} \leqslant ... \leqslant X_{(n)}$ is the ordered sample.[1]

Definition 1.5 *Let $\overline{F}(x,\Theta)$ be the true survival function with unknown parameters Θ and $G_n(x)$ be the empirical survival function of a random sample of size n from $F(x,\Theta)$. Define the* ***Kullback-Leibler divergence of Survival function*** *G_n and $\overline{F}$ by [1]*

$$DLS(G_n(x)\|\overline{F}(x)) = \int_0^\infty G_n(x)ln\frac{G_n(x)}{\overline{F}(x)} - [G_n(x) - \overline{F}(x)]dx. \quad (5)$$

2 Some properties of record statistics

In current section we examine some properties of upper record statistics. See Arnold et al. (1992) and Teimouri et al. (2012).

a) Using (1), the chi-square distribution with 2i degree of freedom and conversion of $U = -2log(1 - F(x))$, we have

$$F_X^{-1}(1 - exp(-\frac{\chi^2_{(2i)}}{2})) \stackrel{d}{=} X_{U(i)}.$$

b) The $100(1 - \varepsilon)\%$ confidence interval for $X_{U(i)}$ is

$$(F_X^{-1}(1 - exp(-\frac{\chi^2_{(2i,\frac{\alpha}{2})}}{2})), F_X^{-1}(1 - exp(-\frac{\chi^2_{(2i,1-\frac{\alpha}{2})}}{2}))).$$

c) The p^{th} quintile for $0 < p < 1$ of $X_{U(i)}$ is

$$F^{-1}_{X_{U(i)}}(p) = F_X^{-1}(1 - exp(-\frac{F^{-1}_{\chi^2_{(2i)}}(p)}{2})).$$

In the remainder of this specification, we will state a lemma and two theorems which has been presented in the Baratpour et al. (2007).

Lemma 2.1 *(Ahmadi(2000))For* $n \geqslant 1$, *we have*

$$\psi(n) = \int x^n e^{-x} logx dx = n![\sum_{i=1}^{n} \frac{1}{i} - \gamma], \tag{6}$$

where γ *is the Euler's constant.*

Theorem 2.2 *Let* $X_1, X_2, \ldots$ *be a sequence of iid random variables from cdf* $F(x)$ *with pdf* $f(x)$ *and entropy* $H(X) < \infty$. *We have*

$$H(X_{U(i)}) = \sum_{j=1}^{i-1} (logj - \frac{i-1}{j}) + (i-1)\gamma - \phi_f(i-1), \quad i \geqslant 1 \tag{7}$$

where $\phi_f(i) = \int_0^\infty \frac{z^i}{i!} e^{-z} logf(F^{-1}(1-e^{-z}))dz$.

Theorem 2.3 *Let* Y *is a random variable from cdf* $F(x)$ *with pdf* $f(x)$. *Under the assumption of theorem 1, we have:*

i)$H(X_{U(i)}) \leqslant H(X^*_{U(i)}) - i - B_i E[\frac{log f_Y(y)}{1 - F_Y(y)}] - logM$ *and*
ii)$H(X_{U(i)}) \geqslant H(X^*_{U(i)}) - i - logM$,
where $M = f(m) < \infty$, m *is the mode of the distribution, and* $Y = MX$,
$X^*_{U(i)} \sim \Gamma(i, 1)$ *and* $B_i = \frac{(i-1)^{(i-1)}}{\Gamma(i)} e^{-(i-1)}$.

3 The upper record value of Weibull distribution and its Entropy

Now we examine the Shannon's Entropy on the upper record of the Weibull distribution. A continuous random variable X is said to be a two-parameter Weibull distribution with shape parameter α and scale parameter β, denoted by $X \sim W(\alpha, \beta)$, if its cdf is

$$F(x) = 1 - exp(-(\frac{x}{\beta})^{\alpha}), \quad x > 0.$$

Now according to section 2 and $F^{-1}(y) = \beta[-log(1 - F_X(x))]^{\frac{1}{\alpha}}$ we have:

a) $$\beta(\frac{\chi^2_{2i}}{2})^{\frac{1}{\alpha}} \stackrel{d}{=} X_{U(i)},$$

$$f_{X_{U(i)}}(x) = \frac{1}{\Gamma(i)}(\frac{\alpha}{\beta})(\frac{x}{\beta})^{\alpha i-1} e^{-(\frac{x}{\beta})^{\alpha}} \tag{8}$$

and

$$\bar{F}_{X_{U(i)}}(x) = \sum_{j=0}^{i-1} \frac{e^{-(\frac{x}{\beta})}(\frac{x}{\beta})^j}{j!}. \tag{9}$$

b) The $100(1-\varepsilon)\%$ confidence interval for $X_{U(i)}$ is

$$(\beta(\frac{\chi^2_{(2i,\frac{\varepsilon}{2})}}{2})^{\frac{1}{\alpha}},\ \beta(\frac{\chi^2_{(2i,1-\frac{\varepsilon}{2})}}{2})^{\frac{1}{\alpha}}).$$

c) The p^{th} quintile for $0<p<1$ of $X_{U(i)}$ is $q_{U(i)}(i)=\beta$.

d) $H(X_{U(i)}) = ln(\frac{\alpha}{\beta}) - i + \sum_{j=1}^{i-1}(logj - \frac{i-1}{j}) + (i-1)\gamma + (1-\frac{1}{\alpha})(\sum_{j=1}^{i-1}\frac{1}{j} - \gamma)$,

$$H(X_{U(i)},\gamma) = \frac{1}{1-\gamma}[-(\Gamma(i)\beta(\gamma+1)) - \frac{\alpha i\gamma+1}{\alpha}log\gamma - log\alpha - log\Gamma(\frac{\alpha i\gamma+1}{\alpha})]$$

and

$$\begin{aligned} DLS(\overline{G}_n(x)\|\overline{F}_n() &= \sum_{j=1}^{n-1}(1-\frac{i}{n})ln(1-\frac{i}{n})\Delta x_{i+1} + \frac{1}{2n\beta}\sum_{i=1}^{n} x^2_{(i)} \\ &- \sum_{i=1}^{n-1}(1-\frac{i}{n})\int_{x_{(i)}}^{x_{(i+1)}} ln(\sum_{i=1}^{n-1}\frac{x^i}{\beta^i i!})dx \\ &- (\overline{x} - \frac{\beta\Gamma(n+\frac{1}{\alpha})}{\Gamma(n)}), \end{aligned} \tag{10}$$

where $\Delta x_i = x_{i+1} - x_i, x_0 = 0$.

4 Estimation of parameters of the upper records Weibull distribution

In this paper, we use the notation $i^{th} - UW$ for the i^{th} upper record Weibull distribution and in current section we estimate shape parameter α and scale parameter β by five methods. Suppose that $X_1, X_2, ..., X_n$ is a random sample from (7) with sample size n.

a) Moment method (MME): Here we provide the MME method of the parameters of a $i^{th} - UW$ distribution. To this purpose, one can show that its mean and variance of the sample are respectively:

$$(a) \quad \overline{x} = \frac{\beta\Gamma(i+\frac{1}{\alpha})}{\Gamma(i)}, \quad (b) \quad s^2 = \frac{\beta^2\Gamma(i+\frac{2}{\alpha})}{\Gamma(i)} - \frac{\beta^2\Gamma^2(i+\frac{1}{\alpha})}{\Gamma^2(i)}.$$

We consider three cases:

- When α is known, then from (a) we have an estimator for β, denoted by $\tilde{\beta}_{\alpha known}$:

$$\tilde{\beta}_{\alpha known} = \frac{\overline{x}\Gamma(i)}{\Gamma(i+\frac{1}{\alpha})}. \tag{11}$$

- When β is known, in this case we need to solve (a) with respect to α, denoted by $\tilde{\alpha}_{\beta known}$:

$$\Gamma(i + \frac{1}{\tilde{\alpha}_{\beta known}}) = \frac{\overline{x}\Gamma(i)}{\beta}. \tag{12}$$

- When both α and β are unknown, first we obtain the population coefficient of variation (CV) from (a) and (b)

$$\frac{\sqrt{Var(X)}}{E(X)} = \frac{[\Gamma(i)\Gamma(i+\frac{2}{\alpha}) - \Gamma^2(i+\frac{1}{\alpha})]^{\frac{1}{2}}}{\Gamma(i+\frac{1}{\alpha})}.$$

Then, equating the sample CV with the population CV(We observe that the population CV is independent of β) we have

$$\frac{s}{\overline{x}} = \frac{[\Gamma(i)\Gamma(i+\frac{2}{\alpha}) - \Gamma^2(i+\frac{1}{\alpha})]^{\frac{1}{2}}}{\Gamma(i+\frac{1}{\alpha})}, \tag{13}$$

where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ and $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. We need to solve (12) to obtain MME of α, denoted by $\tilde{\alpha}$. Then substituting α in (10) we have an estimator for β, denoted by $\tilde{\beta}$.

b) Maximum-likelihood method (MLE):
Here the maximum likelihood estimators of $i^{th} - UW$ are considered. The log-likelihood function is given by

$$lnL(\alpha,\beta) \propto nln(\frac{\alpha}{\beta}) + (\alpha i - 1)[ln(\prod_{i=1}^{n} x_i) - ln\beta] - \sum_{i=1}^{n}(\frac{x_i}{\beta})^{\alpha}. \tag{14}$$

We consider three different cases:

• α is known: differentiate (13) with respect to β then equating to zero and solving with respect to β. Then the MLE of β, denoted by $\widehat{\beta}_{MLE}$, will be

$$\widehat{\beta}_{\text{MLE}} = (\frac{\sum_{j=1}^{n} \alpha x_j^{\alpha}}{1+n-\alpha i})^{\frac{1}{\alpha}}. \tag{15}$$

• β is known: again differentiate (13) with respect to α and then equating by zero and solving with respect to α. Then the MLE of α, denoted by $\widehat{\alpha}_{MLE}$, will be

$$\frac{\partial lnL}{\partial \alpha} = \frac{n}{\widehat{\alpha}_{MLE}} + iln(\prod_{i=1}^{n} x_j) - iln\beta - \sum_{i=1}^{n}(\frac{x_j}{\beta})^{\widehat{\alpha}_{MLE}} ln(\frac{x_j}{\beta}) = 0. \tag{16}$$

• Both α and β are unknown: in this case, replacing (14) in (13) and then differentiate on it with respect to α then equating to zero and solving with respect to α. Here the MLE of α, denoted by $\widehat{\alpha}$, will be

$$\frac{n}{\alpha} - n\frac{\sum_{j=1}^{n} x_j^{\alpha}\ln x_j}{\sum_{j=1}^{n} x_j^{\alpha}} + [\frac{ni}{\alpha(\alpha i - 1 + n)} - \frac{n}{\alpha^2}\ln(\alpha i - 1 + n)] + \sum_{j=1}^{n} i\ln x_j$$

$$- \frac{1}{\alpha^2}\ln(\alpha\sum_{j=1}^{n} x_j^{\alpha}) - (i - \frac{1}{\alpha})(\frac{1}{\alpha} + \frac{\sum_{j=1}^{n} x_j^{\alpha}\ln x_j}{\sum_{j=1}^{n} x_j^{\alpha}}) + \frac{1}{\alpha^2}\ln(\alpha i - 1 + n)$$

$$+(i- \frac{1}{\alpha})\frac{i}{\alpha i - 1 + n} - \frac{n-1}{\alpha^2} = 0. \qquad (17)$$

c) Bayesian method:

Here, we assume the previous information of α and β are independent of each other, so $\pi(\alpha, \beta) = \pi(\alpha)\pi(\beta)$. We represent in this approch for normal, uniform and triangular prior distributions by Monte Carlo simulation.

d) Point estimator for the shape parameter (PE):

Theorem 4.1 *(Teimouri et al. (2012))Suppose that a sequence of i^{th} upper record from Weibull family are observed. A simple estimator of shape parameter is given by*

$$\widehat{\alpha}_{PE} = \frac{log(0.000002i^2 + 0.9998i - 0.3271)}{log(\widehat{m}_{U(i)}) - log(\beta)}, \qquad (18)$$

where $\widehat{m}_{U(i)}$ is the sample median of i^{th} upper record values.[5]

e) Kullback-Leibler divergence of survival function method (DLS): To estimate the parameters by this method, we set (8) in (4). Then we have

$$\mathrm{DLS}(\mathrm{G}_n(x)\|\overline{F}_n(x)) = \sum_{i=1}^{n-1}(1-\frac{i}{n})ln(1-\frac{i}{n})\Delta x_{i+1} + \frac{1}{2n\beta}\sum_{i=1}^{n} x_{(i)}^2$$

$$-\sum_{i=1}^{n-1}(1-\frac{i}{n})\int_{x_{(i)}}^{x_{(i+1)}} ln(\sum_{i=1}^{n}\frac{x^i}{\beta^i i!})dx-(\overline{x}_n-\frac{\beta\Gamma(n+\frac{1}{\alpha})}{\Gamma(n)}), \qquad (19)$$

where $\Delta x_i = x_{i+1} - x_i, x_0 = 0.$

Now differentiate (18) with respect to α then equating to zero and solving with respect to α. Then, $\frac{\partial}{\partial\alpha}\Gamma(i+\frac{1}{\alpha})=0$. But we don't have a good estimator for α by this equation since it does not depend on the sample values.

Now we suppose that α is known. Now equation (18) should be minimized respect to β. For this propose, we have

$$\frac{\Gamma(i+\frac{1}{\alpha})}{\Gamma(i)}-\frac{\sum_{i=1}^{n}x_{(j)}^{2}}{2n\beta^{2}}-\sum_{l=1}^{n}(1-\frac{l-1}{n})\int_{x_{(l-1)}}^{x_{(l)}}\frac{\sum_{k=0}^{n-1}\frac{(\frac{x}{\beta})^{k}}{(k-1)!}}{\sum_{k=0}^{n-1}\frac{(\frac{x}{\beta})^{k}}{k!}}dx=0.(20)$$

We observe that a close form solution of (19) for β is not possible.

5 Simulation study

Since the MME, MLE, DLS and Bayesian estimation of the parameters have not closed form so checking the performance of them, theoretically is a difficult task. Therefore, they must be solved numerically. We have done this work by Matlab software. First, we have generated $x_1, x_2, ..., x_n$ from a $i^{th} = 4$ upper record Weibull distribution (for more details, see Teymori and Gupta). In this study we assumed $\alpha_{true} = 2$, $\beta_{true} = 3$ and both α and β are unknown. This sample simulated was used to estimate the $4^{th} - UW$ using the MME, MLE, DLS, Baysian and point methods. The above process was repeated 10,000 times.

Consequently, we have a set of $M = 10000$ $4^{th} - UW$ parameter estimations using each method. Then the mean values $\bar{\alpha}$ and $\bar{\beta}$, and sample variances S_{α}^2 and S_{β}^2 were computed using:

$$\overline{\alpha} = \frac{1}{M}\sum_{k=1}^{M}\alpha_k, \quad \overline{\beta} = \frac{1}{M}\sum_{k=1}^{M}\beta_k,$$

$$S_{\alpha}^2 = \frac{1}{M-1}\sum_{k=1}^{M}(\alpha_k - \overline{\alpha}), \quad S_{\beta}^2 = \frac{1}{M-1}\sum_{k=1}^{M}(\beta_k - \overline{\beta}),$$

where α_k, β_k are the estimated $4^{th} - UW$ shape and scale parameters from k^{th} sample. To illustrate the effect of sample size, we carried out a simulation study for some levels of $n = 5, 6, \cdots, 30$. First, we study the behavior of β given in (10), (14), (19) and Bayesian method with four prior distribution. Figure 1 shows the simulation results of β.

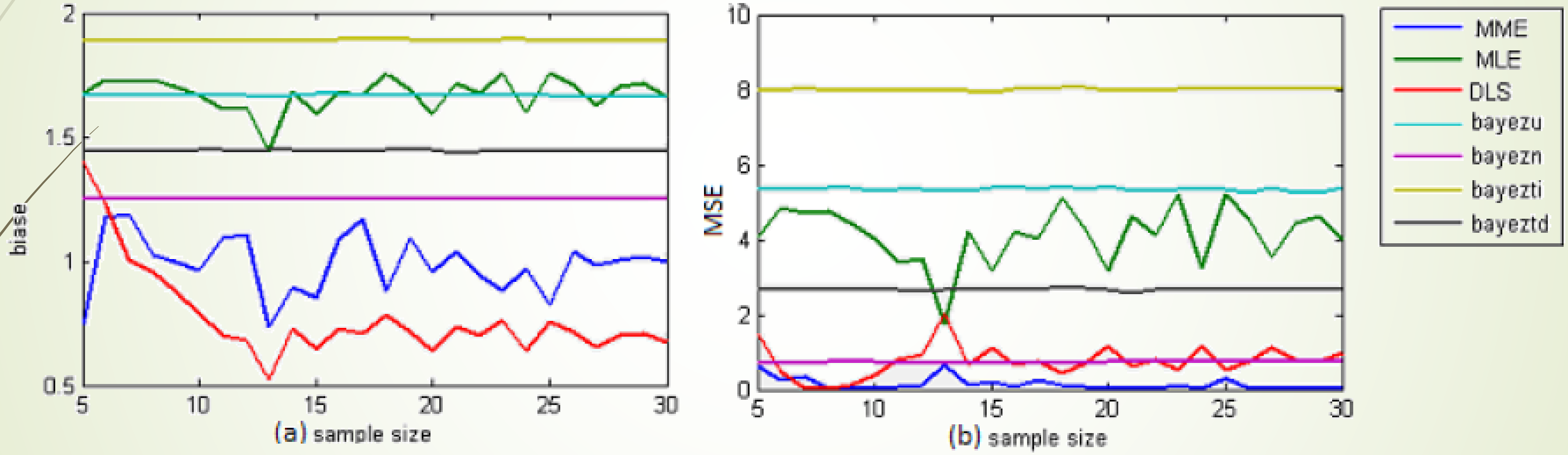

Figure 1: (a) $\frac{\overline{\beta}}{\beta}$. (b) $\widehat{MSE}_\beta$ as functions of sample size

The following results can be concluded from this figure:

1) According to figure 1.a, the MME, DLS and Bayesian with prior normal methods act better than the rest in terms of bias. MME and Bayesian estimation for β are overestimated whereas DLS is underestimated.

2) According to figure 1.b, mean squared error (MSE) of MME method is best for every n.

Now we study the behavior of α given in (12), (16), (17) and Bayesian method with four prior distribution. Figure 2 shows the simulation results for α. The following results can be concluded from this figure:

1) According to figure 2.a, the MME and point estimator methods act better than the rest and are so similar in terms of bias. Also MLE values for α are closed to α_{true}, then this method can be a good estimator for α.

2) figure 2.b shows MSE of MME, MLE and point estimator methods are best for every n and very closed to zero.

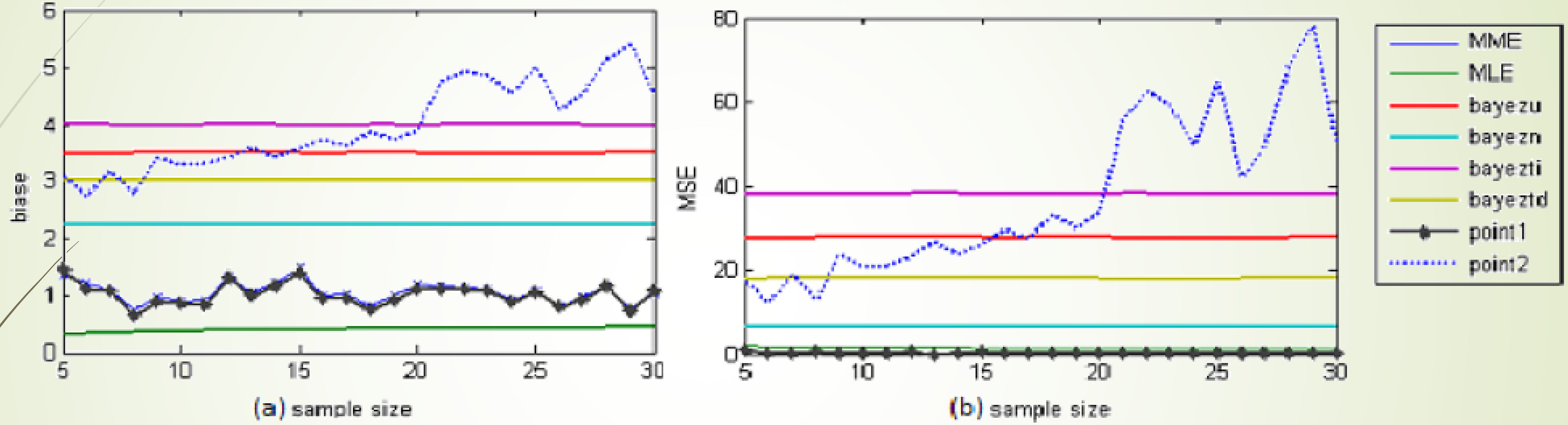

Figure 2: (a) $\frac{\overline{\alpha}}{\alpha}$. (b) $\widehat{MSE}_{\alpha}$as functions of sample size

6 Conclusion

The current paper concerned with different estimators of the records Weibull distribution parameters and discussed how appropriate and inappropriate these estimators are. The simulation process, suggests MME and DLS methods for estimating the parameter β. MME and proposed point estimators based on estimated β by MME indicated to be appropriate estimators for α.

References

Ahmadi, J. (2009). Entropy properties of Certain Record Statistics and Some Characterization results, JIRSS Vol.7, Nos. 1-2, pp 1-13.

Ahmadi, J. and Doostparast, M. (2008). Satistical inference based on K-records, Mashhad R.J. Math.Sci., vol.1, pp 67-82.

Ahmadi, J. and Balakrishnan, N. (2004). Confidence intervals for quantiles in terms of record range. Statistics probability Letters 68, 395-405.

Balakrishnan, N. and Chan, P.s. (1993). Record values from Rayleigh and Weibull distribution and associated inference.proceedings of the confidence on Extreme Value Theory and applications, volume 3 Gaithersbarg, Maryland.

Baratpour, J., Ahmaidi, J. and Arghami, R. (2007). Entropy properties of record statistics, statistical papers 48.197-213.

Cover, T.M. and Thomas, J.A. (2006). Elements of information theory, 2nd ed., Wiley, New york.

Casella, G. and Berger, R.L. (2007). Statistical Information, 2nd ed., Duxbury Press.

Kubong Mbah, A. (2007). On the theory of records and applications , USF Granduate School, Florida.

Lio, J. (2001). Information theoretic content and probability, Ph.D. thesis, University of Florida, USA.

Perez-Cruz, F. (2008). Kullbak-Leibler divergence estimation of continuous distributions, IEEE International symposium on Information Theory, Toronto, Canada.

Rinne, H. (2009) .The Weibull distribution: handbook, CRC Press. pp. 275.

Scholz, F. (2008). Inference for the Weibull distribution stat 498B Industrial statistics.

Teimouri, M. and Gupta, A.K. (2012). On the Weibull record statistics and associated inferences, statistica, anno LXXII, n.2.

Wuu Wu, J. and Chiao, T. (2006). Satistical inference about the shape parameter of the Weibull distribution by upper record values, statistical papers 48,95-129.

Yari, G.H. and Saghafi, A. (2012). Unbiased Weibull modulus estimation using Differential Cumulative Entropy, communication in statistics simulation and computation , 41:1-7.

Yari, G.H. and Saghafi, A. (2009). Past Renyi entropy of relaibility distributions, Proceeding of the 7 the Seminar on Probability Stochastics Processes, Isfahan, August 13-14, p.240-247.

Yari, G.H. and Saghafi, A. (2010). Past Renyi entropy of Beta distributions, Proceeding of the 1 th International Workshop on Stochastic Processes, Tehran, Decebmer 1-3.

Yari, G.H., Mirhabibi, A. and Saghafi, A.(2013). Estimation of the Weibull parameters by Kullback-Leibler divergence of Survival functions, Appl. Math. Inf. Sci. 7, No. 1, 187-192.

Yuan, Ch,J. and Druzdzel, M. (2007). Theoretical analysis and practical insights on importance sampling in Bayesian networks. International Journal of Approximate Reasoning 46(2007)320-333.

END

Thank you for your attention