REVSTAT – Statistical Journal
Volume 19, Number 1, January 2021, 99–110

MINIMUM AREA CONFIDENCE REGION FOR WEIBULL DISTRIBUTION BASED ON RECORDS

Authors: Junmei Zhou
– School of Mathematics and Statistics, Hainan Normal University, Haikou, China

Fen Jiang
– School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China

Jin Zhang
– School of Mathematics and Statistics, Yunnan University, Kunming, China
jinzhang@ynu.edu.cn

Received: September 2017 Revised: August 2018 Accepted: October 2018

Abstract:

- Record values are commonly seen in real life applications, and many important studies on record values relate to Weibull distributions. Based on record values, we establish the minimum area confidence region for the two-parameter Weibull distribution, which is shown to be superior to the classical confidence regions for having smaller expected area.

Keywords:

- *order statistic; parameter; record value; simulation; sufficient statistic.*

AMS Subject Classification:

- 62F25.

1. INTRODUCTION

Weibull distribution has wide applications in survival analysis, reliability engineering, weather forecasting, hydrology, meteorology and insurance (e.g, Murthy *et al.* [16], Ye *et al.* [26]). The cumulative distribution function (cdf) of the two-parameter Weibull distribution, denoted by Weibull(β, η), is

$$F(x; \beta, \eta) = 1 - e^{-(x/\eta)^{\beta}}, \; x > 0,$$

where $\beta > 0$ is the shape parameter and $\eta > 0$ is the scale parameter. In particular, if $\beta = 1$, then the Weibull distribution simplifies as the exponential distribution Exp(η) with mean η, and it becomes the Rayleigh distribution when $\beta = 2$. In the case of $\beta \geq 10$, the shape of Weibull distribution is close to that of the smallest extreme value distribution (e.g, Nelson [17]).

Record values were first introduced by Chandler [11] as special order statistics from random samples, which can be simply described as follows. (For more description, refer to Ahsanullah [1] and Arnold *et al.* [2].) Let $\{X_n, \; n = 1, 2, ...\}$ be an iid (independent and identically distributed) sequence of continuous random samples. Observation X_j is called an upper record if $X_j > X_i$ for each $i < j$. In addition, the record times sequence $\{U_n, \; n \geq 1\}$ is defined by $U_1 = 1$ with probability 1 and $U_n = \min\{j : j > U_{n-1}, \; X_j > X_{U_{n-1}}\}$ for $n \geq 2$. Thus, the sequence $\{X_{U_n}, \; n \geq 1\}$ is called a sequence of upper record statistics. Lower record statistics can be defined analogously.

Record values are commonly seen in real life applications, such as those in meteorology, sports, economics and life tests (e.g., Ahsanullah [1] and Arnold *et al.* [2]), where joint confidence region for unknown parameters is of great practical significance. In the recent years, joint confidence regions based on records were investigated by many authors, and most of their studies on record values are related to Weibull distributions. For references, see, for example, Chan [10], Chen [12], Murthy *et al.* [16], Soliman *et al.* [21], Wu and Tseng [25], Soliman and Al-Aboud [20], Asgharzadeh *et al.* [7], Asgharzadeh and Abdi [3, 4, 5], Teimouri and Nadarajah [22], Wang and Shi [23], Jafari and Zakerzadeh [13], Wang and Ye [24], Zakerzadeh and Jafari [27], and Zhao *et al.* [30].

In the next section, we discuss the classical methods to build joint confidence regions for parameters of Weibull(β, η) distribution, based on (upper) record values. Then the minimum area confidence region (MACR) for (β, η) based on records is established in Section 3 and Section 4. Comparison of these confidence regions is given in Section 5, showing that the proposed MACR is superior to the classical confidence regions for having smaller expected area.

2. CLASSICAL CONFIDENCE REGIONS BASED ON RECORDS

Let $X_{U_1} < X_{U_2} < \cdots < X_{U_n}$ be the upper record values coming from Weibull(β, η). For simplicity, we write X_{U_i} as R_i and let $Y_i = (R_i/\eta)^{\beta}$ $(i = 1, 2, ..., n)$. Then $Y_1 < Y_2 < \cdots < Y_n$ are the first n upper record values from the standard exponential distribution. Arnold *et al.* [2]

showed that $Z_1, ..., Z_n$ are iid from Exp(1), that is, $Z_1, ..., Z_n \overset{\text{iid}}{\sim} \text{Exp}(1)$, where $Z_i = Y_i - Y_{i-1}$ $(i = 1, 2, ..., n;\ Y_0 \equiv 0)$. It follows that for $j = 1, 2, ..., n-1$,

- **(i)** $U_j = 2\sum_{i=1}^{j} Z_i = 2(\frac{R_j}{\eta})^\beta \sim \chi^2_{2j}$, $V_j = 2\sum_{i=j+1}^{n} Z_i = 2[(\frac{R_n}{\eta})^\beta - (\frac{R_j}{\eta})^\beta] \sim \chi^2_{2(n-j)}$ and the two pivotal quantities are independent, where χ^2_m denotes the chi-square distribution with m degrees of freedom;
- **(ii)** $U_j + V_j = 2(\frac{R_n}{\eta})^\beta \sim \chi^2_{2n}$, $\frac{V_j/2(n-j)}{U_j/2j} = \frac{j}{n-j}[(\frac{R_n}{R_j})^\beta - 1] \sim F_{2(n-j),2j}$, and the two pivotal quantities are independent (see Asgharzadeh and Abdi [4], Johnson *et al.* [15], p. 350) where F_{n_1,n_2} stands for the F-distribution with n_1 and n_2 degrees of freedom.

To build a joint confidence region for β and η, we have from (ii) that

$$P\left[F_{2(n-j),2j}(\alpha_1) \le \frac{j}{n-j}[(\frac{R_n}{R_j})^\beta - 1] \le F_{2(n-j),2j}(\alpha_2)\right] = \sqrt{1-\alpha}$$

for $j = 1, 2, ..., n-1$, and

$$P\left[\chi^2_{2n}(\alpha_1) \le 2(\frac{R_n}{\eta})^\beta \le \chi^2_{2n}(\alpha_2)\right] = \sqrt{1-\alpha},$$

where $\alpha_1 = \frac{1-\sqrt{1-\alpha}}{2}$, $\alpha_2 = \frac{1+\sqrt{1-\alpha}}{2}$, $F_{n_1,n_2}(p)$ is the p quantile of F_{n_1,n_2} and $\chi^2_m(p)$ is the p quantile of χ^2_m. Then one type of the classical level $1-\alpha$ confidence region for (β, η) is given by (Asgharzadeh and Abdi [4])

$$A_j : \begin{cases} \dfrac{\log[1 + \frac{n-j}{j}F_{2(n-j),2j}(\alpha_1)]}{\log(R_n/R_j)} \le \beta \le \dfrac{\log[1 + \frac{n-j}{j}F_{2(n-j),2j}(\alpha_2)]}{\log(R_n/R_j)}, \\ R_n[\dfrac{2}{\chi^2_{2n}(\alpha_2)}]^{\frac{1}{\beta}} \le \eta \le R_n[\dfrac{2}{\chi^2_{2n}(\alpha_1)}]^{\frac{1}{\beta}}, \end{cases} \tag{2.1}$$

where $j = 1, 2, ..., n-1$, and each A_j produces a level $1-\alpha$ confidence region for (β, η). Based on Monte Carlo simulation, Asgharzadeh and Abdi [4] observed that $A_{\lfloor \frac{n}{5} \rfloor}$ and $A_{\lfloor \frac{n}{5} \rfloor + 1}$ provide the smallest confidence areas in most cases, where $\lfloor x \rfloor$ denotes the largest integer value smaller than x.

Noticing that $U = 2\beta\sum_{i=1}^{n}\log(R_n/R_i) \sim \chi^2_{2n-2}$ and $V = 2(R_n/\eta)^\beta \sim \chi^2_{2n}$, which are independent, Jafari and Zakerzadeh [13] derived another type of the classical level $1-\alpha$ confidence region for (β, η):

$$B : \begin{cases} \dfrac{\chi^2_{2n-2}(\alpha_1)}{2\sum_{i=1}^{n}\log(R_n/R_i)} \le \beta \le \dfrac{\chi^2_{2n-2}(\alpha_2)}{2\sum_{i=1}^{n}\log(R_n/R_i)}, \\ R_n[\dfrac{2}{\chi^2_{2n}(\alpha_2)}]^{\frac{1}{\beta}} \le \eta \le R_n[\dfrac{2}{\chi^2_{2n}(\alpha_1)}]^{\frac{1}{\beta}}, \end{cases} \tag{2.2}$$

where $\alpha_1 = \frac{1-\sqrt{1-\alpha}}{2}$ and $\alpha_2 = \frac{1+\sqrt{1-\alpha}}{2}$. By simulation study, Jafari and Zakerzadeh [13] concluded that the expected area of the confidence region in (2.2) is smaller than that in (2.1) proposed by Asgharzadeh and Abdi [4].

3. A BASIC THEOREM ON THE MACR

Let $T = T(X)$ be a sufficient statistic of parameter $\theta = (\beta, \eta)$ with pdf (probability density function) $f(t; \theta)$, where $t \in T(\mathcal{X})$, $\theta \in \Theta$. Here, X denotes the random sample with sample space $\mathcal{X}$, and Θ is the parameter space.

According to the Sufficiency Principle in mathematical statistics (e.g., Bickel and Doksum [8], Casella and Berger [9]), we only need to consider the confidence region $C(T)$ based on sufficient statistic $T = T(X)$, without loss of information from the sample X. The purpose of using the sufficient statistic to simplify or reduce the sample X to $T = T(X)$, so that we have $T(\mathcal{X}) = \Theta$ to be used in the following theorem. This theorem creates the MACR for θ under some restriction, where $|C|$ denotes the area of any confidence region C.

Theorem 3.1. *Suppose that for any $\theta \in \Theta$,*

1. *$T = T(X)$ is a sufficient statistic of θ with pdf $f(t; \theta)$, such that $T(\mathcal{X}) = \Theta$;*
2. *There exists some $p(\theta) > 0$, such that $\tilde{f}(T; \theta) = f(T; \theta)/p(\theta)$ is a pivotal quantity;*
3. *The confidence region $C_k(T)$ is defined by*

$$C_k(T) = \{\theta :\ \tilde{f}(T; \theta) \geq k,\ \theta \in \Theta\},$$

where $k > 0$ is the critical value determined by $P[\theta \in C_k(T)] = 1 - \alpha$ for any $\alpha \in (0, 1)$.

Then $C_k(T)$ is the level $1 - \alpha$ MACR of θ, under restriction

$$\int_{\theta \in C(t)} dt \leq r_k(\theta)|C(\theta)| \tag{3.1}$$

for any $C(T)$ and $\theta \in \Theta$, where $r_k(\theta) = \int_{\theta \in C_k(t)} dt/|C_k(\theta)|$.

Proof: Let $C(T)$ be any level $1-\alpha$ confidence region of θ, satisfying $\int_{\theta \in C(t)} dt \leq r_k(\theta)|C(\theta)|$. Then for any $\theta \in \Theta$,

$$1 - \alpha \ \leq\ P[\theta \in C(T)] = \int_{\theta \in C(t)} [f(t; \theta) - kp(\theta)]dt + kp(\theta) \int_{\theta \in C(t)} dt.$$

It follows from $P[\theta \in C_k(T)] = 1 - \alpha$ that

$$\begin{aligned} 0 \ &\leq\ P[\theta \in C(T)] - P[\theta \in C_k(T)] \\ &=\ d_k(\theta) + kp(\theta)\Big[\int_{\theta \in C(t)} dt - \int_{\theta \in C_k(t)} dt\Big] \\ &\leq\ kp(\theta)r_k(\theta)\big[|C(\theta)| - |C_k(\theta)|\big], \end{aligned}$$

where

$$d_k(\theta) = \Big(\int_{\theta\in C(t)} dt - \int_{\theta\in C_k(t)} dt\Big)[f(t;\theta) - kp(\theta)]dt$$
$$= \Big(\int_{\theta\in C(t)\cap\overline{C_k(t)}} - \int_{\theta\in C_k(t)\cap\overline{C(t)}}\Big)[f(t;\theta) - kp(\theta)]dt \le 0$$

where $\overline{C}$ denotes the complementary set of C, and $f(t;\theta) - kp(\theta) \ge$ or ≤ 0 if $\theta \in C_k(t)$ or $\theta \in \overline{C_k(t)}$. It follows that $|C(\theta)| - |C_k(\theta)| \ge 0$ or $|C_k(\theta)| \le |C(\theta)|$ for any $\theta \in \Theta$, which, together with $T(\mathcal{X}) = \Theta$, implies that $|C_k(T)| \le |C(T)|$ for any T. The proof is complete. $\square$

This theorem extends the basic theorems in Zhang [28, 29], which are valid for building the MACRs of parameters for normal and exponential distributions, but are not for the Weibull(β, η) distribution. By Theorem 3.1, $C_k(T)$ is the level $1-\alpha$ optimal confidence region of θ, minimizing the area of any level $1-\alpha$ confidence region $C(T)$ under the restriction in (3.1). This restricted condition may look complicated, but the MACR $C_k(T)$ does satisfy this condition, due to

$$\int_{\theta\in C_k(t)} dt = r_k(\theta)|C_k(\theta)|.$$

Moreover, there is no need to check which $C(T)$ is under the restriction. The situation is like that of using Lehmann-Scheffé theorem to build the UMVUE (uniformly minimum variance unbiased estimator), without need to check which estimator is unbiased (e.g., Bickel and Doksum [8], Casella and Berger [9]).

A similar theorem was established in Jeyaratnam [14]. The minimum volume confidence region built by Jeyaratnam is based on a pivotal quantity $T(X, \theta)$ such that for each x, $T(x, \theta)$ is a one-to-one map on Θ whose Jacobian J does not depend on θ, and it is optimal for any level $1-\alpha$ confidence region based on the special pivotal quantity.

4. FORMULATION OF THE MACR BASED ON RECORDS

Based on n record values $R_1 < R_2 < \cdots < R_n$ from Weibull(β, η), we now apply Theorem 3.1 to derive the MACR for parameter $\theta = (\beta, \eta)$. Let

$$Z = \sum_{i=1}^{n-1} \log(R_n/R_i).$$

Then (Z, R_n) is a sufficient statistic for (β, η), according to Wang and Ye [24]. Being its equivalent statistic, $T = (Z, \log R_n/Z)$ is also sufficient for (β, η). By Section 2, $U = 2\beta Z \sim \chi^2_{2n-2}$ and $V = 2(\frac{R_n}{\eta})^\beta \sim \chi^2_{2n}$, which are independent. Thus, (U, V) has pdf $f_{\chi^2_{2n-2}}(u) f_{\chi^2_{2n}}(v)$, $u, v > 0$, and the pdf of $T = (T_1, T_2)$ is

$$f(t_1, t_2; \beta, \eta) = f_{\chi^2_{2n-2}}(2\beta t_1) f_{\chi^2_{2n}}[2(\frac{e^{t_1 t_2}}{\eta})^\beta]\Big|\frac{\partial(u,v)}{\partial(t_1,t_2)}\Big|, \ t_1 > 0,$$

where $f_{\chi^2_m}(x)$ $(F_{\chi^2_m}(x))$ denotes the pdf (cdf) of χ^2_m, $U = 2\beta T_1$, $V = 2(\frac{e^{T_1 T_2}}{\eta})^\beta$ and Jacobian $|\frac{\partial(u,v)}{\partial(t_1,t_2)}| = 4\beta^2 t_1 (\frac{e^{t_1 t_2}}{\eta})^\beta$.

Treating $R = (R_1, R_2, ..., R_n)$ as the random sample X in Section 3, we can see that $T = (Z, \log R_n/Z)$ is a sufficient statistic for $\theta = (\beta, \eta)$, satisfying the conditions 1 and 2 in Theorem 3.1, where the pivotal quantity is

$$\tilde{f}(t_1, t_2; \beta, \eta) = f_{\chi^2_{2n-2}}(2\beta t_1) f_{\chi^2_{2n}}[2(\frac{e^{t_1 t_2}}{\eta})^\beta] \cdot (2\beta t_1) \cdot 2(\frac{e^{t_1 t_2}}{\eta})^\beta, \; t_1 > 0.$$

Hence, the level $1-\alpha$ MACR for (β, η) is $C_k(T) = \{(\beta, \eta) : \; \tilde{f}(T_1, T_2; \beta, \eta) \geq k\}$ or

$$C_k(T) = \{(\beta, \eta) : \; g(\beta Z) + h((R_n/\eta)^\beta) \leq k_\alpha\}, \tag{4.1}$$

where $g(x) = x - (n-1)\log x$ and $h(y) = y - n \log y$ are both convex functions, and k_α is a critical value to be determined.

Let $k(x) \equiv k_\alpha - g(x)$, $\tilde{k} \equiv k_\alpha - h_{\min}$ and $h_{\min} = h(n)$. Then the confidence region in (4.1) can be equivalently expressed as

$$C_k(T) = \begin{cases} g(\beta Z) \leq \tilde{k}, \\ h((R_n/\eta)^\beta) \leq k(\beta Z), \end{cases}$$

for computational purpose. From the property of convex function, $g(\beta Z) \leq \tilde{k}$ is equivalent to $k_1 \leq \beta Z \leq k_2$ with $g(k_1) = g(k_2) = \tilde{k}$, and $h((R_n/\eta)^\beta) \leq k(\beta Z)$ means $k_{11}(\beta Z) \leq (R_n/\eta)^\beta \leq k_{12}(\beta Z)$ with $h(k_{11}(\beta Z)) = h(k_{12}(\beta Z)) = k(\beta Z)$. Finally, the level $1-\alpha$ MACR for (β, η) in (4.1) can be written as

$$C_k(T) = \begin{cases} k_1/Z \leq \beta \leq k_2/Z, \\ R_n/[k_{12}(\beta Z)]^{\frac{1}{\beta}} \leq \eta \leq R_n/[k_{11}(\beta Z)]^{\frac{1}{\beta}}, \end{cases} \tag{4.2}$$

where $g(x) = x - (n-1)\log x$ with $g(k_1) = g(k_2) = \tilde{k}$, $h(y) = y - n \log y$ with $h(k_{11}(\beta Z)) = h(k_{12}(\beta Z)) = k(\beta Z)$, and the critical value k_α is determined by

$$\begin{aligned} 1 - \alpha &= P[(\beta, \eta) \in C_k(T)] \\ &= P[g(\beta Z) + h((R_n/\eta)^\beta) \leq k_\alpha] \\ &= \underset{g(x)+h(y) \leq k_\alpha}{\int_0^\infty \int_0^\infty} 4 f_{\chi^2_{2n-2}}(2x) f_{\chi^2_{2n}}(2y) dx dy \\ &= \int_{k_1}^{k_2} \int_{k_{11}(x)}^{k_{12}(x)} 4 f_{\chi^2_{2n-2}}(2x) f_{\chi^2_{2n}}(2y) dx dy \\ &= \int_{k_1}^{k_2} 2 f_{\chi^2_{2n-2}}(2x) [F_{\chi^2_{2n}}(2k_{12}(x)) - F_{\chi^2_{2n}}(2k_{11}(x))] dx, \end{aligned}$$

where $k_\alpha > g_{\min} + h_{\min}$ and $h(k_{11}(x)) = h(k_{12}(x)) = k_\alpha - g(x)$. A short R code (R Core Team [18]) for computing k_α, k_1, k_2, $k_{11}(x)$, $k_{12}(x)$ in (4.2) is given in Appendix A, where the last integral in the above equation is computed by using Simpson's rule for numerical integration (the interval $[k_1, k_2]$ is split up into 1000 subintervals).

5. COMPARISON OF CONFIDENCE REGIONS

In the statistical literature, the commonly used measure of accuracy for a confidence region is its volume (area). Clearly, the smaller the volume (area), the more accurate the confidence region. To compare the MACR in (4.1) or (4.2) with the classical confidence regions in (2.1) and (2.2), we now discuss their areas as follows.

Given the sample data of upper record values: $R = (R_1, R_2, ..., R_n)$, the area of the classical confidence region in (2.1) is

$$|A_j| = \int_{\frac{\log[1+\frac{n-j}{j}F_{2(n-j),2j}(\alpha_1)]}{\log(R_n/R_j)}}^{\frac{\log[1+\frac{n-j}{j}F_{2(n-j),2j}(\alpha_2)]}{\log(R_n/R_j)}} R_n[(\frac{2}{\chi^2_{2n}(\alpha_1)})^{\frac{1}{\beta}} - (\frac{2}{\chi^2_{2n}(\alpha_2)})^{\frac{1}{\beta}}]d\beta,$$

where the integral can be computed by using Simpson's rule for numerical integration.

Similarly, the area of the classical confidence region in (2.2) is

$$|B| = \int_{\frac{\chi^2_{2n-2}(\alpha_1)}{2\sum_{i=1}^{n}\log(R_n/R_i)}}^{\frac{\chi^2_{2n-2}(\alpha_2)}{2\sum_{i=1}^{n}\log(R_n/R_i)}} R_n[(\frac{2}{\chi^2_{2n}(\alpha_1)})^{\frac{1}{\beta}} - (\frac{2}{\chi^2_{2n}(\alpha_2)})^{\frac{1}{\beta}}]d\beta,$$

and the area of the MACR in (4.1) or (4.2) is

$$\begin{aligned}|C_k(T)| &= \int_{k_1/Z}^{k_2/Z} R_n[(\frac{1}{k_{11}(\beta Z)})^{\frac{1}{\beta}} - (\frac{1}{k_{12}(\beta Z)})^{\frac{1}{\beta}}]d\beta \\ &= \frac{R_n}{Z}\int_{k_1}^{k_2}[(\frac{1}{k_{11}(x)})^{\frac{Z}{x}} - (\frac{1}{k_{12}(x)})^{\frac{Z}{x}}]dx.\end{aligned}$$

Monte Carlo simulation is conducted to compute the expected areas of confidence regions in (2.1), (2.2) and (4.1). Since η is the scale parameter of Weibull(β, η), we can set $\eta = 1$ without loss of generality. We generate $N = 1000$ independent upper record values $R^{(i)} = (R_1^{(i)}, R_2^{(i)}, ..., R_n^{(i)})$ from Weibull$(\beta, 1)$, where $i = 1, 2, ..., N$. Then $\sum_{i=1}^{N}|C(R^{(i)})|/N$ is used to simulate $E|C(R)|$, the expected area of $C(R)$.

Table 1 lists the expected areas of confidence regions in (2.1), (2.2) and (4.1), where A_* stands for the smallest-area confidence region in (2.1), B represents the confidence region in (2.2), and $C_k(T)$ is the MACR in (4.1). We see from Table 1 that the MACR is always the best for having the smallest expected area.

Example 5.1. Roberts [19] gave the monthly maximal of one-hour average concentration of sulfur dioxide in pphm (parts per hundred million) from Long Beach, California, for the years 1956 to 1974. The related upper record values for the month of October is 26, 27, 40 and 41, where $n = 4$ and $R_4 = 41$.

Table 1: Expected areas of confidence regions for (β, η) with $\eta = 1$.

$1-\alpha$	n	Region	β							
			0.25	0.5	1.0	1.2	1.5	2.0	3.0	5.0
0.90	5	A_*	403.8	17.92	6.110	5.250	4.788	4.275	4.054	3.911
		B	336.7	14.93	5.347	4.660	4.127	3.876	3.620	3.565
		$C_k(T)$	315.1	14.01	5.113	4.403	3.890	3.569	3.415	3.364
	10	A_*	214.6	10.17	2.982	2.525	2.253	2.064	1.926	1.893
		B	140.0	6.644	2.384	2.123	1.881	1.756	1.648	1.618
		$C_k(T)$	108.7	6.392	2.232	1.889	1.793	1.651	1.549	1.546
	15	A_*	152.4	6.399	1.970	1.717	1.503	1.364	1.283	1.264
		B	73.04	4.232	1.491	1.291	1.174	1.122	1.068	1.069
		$C_k(T)$	67.13	4.129	1.464	1.267	1.128	1.075	1.031	1.003
	20	A_*	108.1	4.788	1.434	1.233	1.110	1.028	0.973	0.952
		B	55.59	2.741	1.067	0.962	0.900	0.854	0.788	0.780
		$C_k(T)$	45.46	2.605	1.035	0.908	0.876	0.789	0.767	0.744
	30	A_*	76.49	2.806	0.906	0.812	0.741	0.666	0.635	0.629
		B	37.69	1.712	0.684	0.617	0.571	0.539	0.521	0.518
		$C_k(T)$	21.18	1.505	0.659	0.601	0.550	0.517	0.491	0.494
0.95	5	A_*	697.9	27.58	8.752	7.528	6.373	5.800	5.417	5.228
		B	578.3	22.50	7.254	6.413	5.559	5.167	4.787	4.743
		$C_k(T)$	509.4	21.82	6.919	6.050	5.258	4.756	4.471	4.408
	10	A_*	420.0	13.81	4.003	3.490	3.108	2.673	2.538	2.520
		B	228.6	8.852	3.135	2.772	2.490	2.289	2.184	2.142
		$C_k(T)$	198.0	8.648	2.931	2.668	2.342	2.129	2.048	2.004
	15	A_*	336.2	9.370	2.714	2.326	2.006	1.817	1.682	1.670
		B	138.0	6.083	2.057	1.756	1.615	1.494	1.412	1.393
		$C_k(T)$	111.1	5.490	1.911	1.685	1.530	1.420	1.311	1.299
	20	A_*	179.8	7.024	1.998	1.694	1.457	1.356	1.271	1.251
		B	78.08	4.305	1.429	1.275	1.195	1.097	1.038	1.022
		$C_k(T)$	70.62	3.659	1.328	1.218	1.112	1.034	0.988	0.965
	30	A_*	112.8	3.929	1.278	1.093	0.987	0.890	0.845	0.830
		B	50.17	2.508	0.925	0.825	0.766	0.714	0.693	0.675
		$C_k(T)$	39.10	2.265	0.911	0.780	0.722	0.676	0.646	0.643
0.99	5	A_*	3731	71.42	16.96	13.87	11.68	9.785	8.808	8.413
		B	2566	56.98	14.70	12.25	10.13	8.881	8.004	7.626
		$C_k(T)$	1541	45.17	11.68	10.10	8.757	7.533	6.943	6.690
	10	A_*	1022	29.63	7.443	5.990	5.039	4.394	4.002	3.943
		B	568.6	19.05	5.510	4.588	4.046	3.675	3.424	3.349
		$C_k(T)$	434.5	16.75	5.099	4.293	3.778	3.368	3.147	3.054
	15	A_*	750.5	19.30	4.685	3.835	3.263	2.866	2.687	2.600
		B	367.2	11.27	3.356	2.927	2.561	2.321	2.225	2.178
		$C_k(T)$	309.7	10.25	3.104	2.684	2.396	2.160	2.013	2.002
	20	A_*	477.4	13.45	3.479	2.863	2.507	2.159	1.993	1.946
		B	212.1	7.883	2.496	2.174	1.889	1.741	1.634	1.600
		$C_k(T)$	179.1	7.291	2.292	1.968	1.772	1.624	1.533	1.482
	30	A_*	306.3	8.630	2.204	1.892	1.597	1.460	1.330	1.300
		B	115.9	4.887	1.497	1.370	1.239	1.141	1.070	1.068
		$C_k(T)$	79.47	4.170	1.471	1.290	1.125	1.044	1.002	0.981

Chan [10] showed that Weibull(β, η) is a reasonable model for the data set. Then the level 95% MACR for (β, η) in (4) is given by

$$C_k(T) = \{(\beta, \eta) : 0.8979\beta - 3\log(0.8979\beta) + (41/\eta)^{\beta} - 4\beta\log(41/\eta) \leq k_\alpha\}$$

with area 154.908, where $k_\alpha = 1.297$, $k_1 = 0.451$ and $k_2 = 9.640$ are obtained by using the R code in Appendix A.

The level 95% confidence regions for (β, η) in (2.1) are

$$A_1 = \{(\beta, \eta) : 0.5826 \le \beta \le 11.9955,\ 41(0.1029)^{\frac{1}{\beta}} \le \eta \le 41(1.1318)^{\frac{1}{\beta}}\},$$

$$A_2 = \{(\beta, \eta) : 0.1646 \le \beta \le 6.4905,\ 41(0.1029)^{\frac{1}{\beta}} \le \eta \le 41(1.1318)^{\frac{1}{\beta}}\},$$

$$A_3 = \{(\beta, \eta) : 0.1720 \le \beta \le 58.9824,\ 41(0.1029)^{\frac{1}{\beta}} \le \eta \le 41(1.1318)^{\frac{1}{\beta}}\},$$

with areas 195.118, 166.671 and 369.396 respectively.

The level 95% confidence region for (β, η) in (2.2) is

$$B = \{(\beta, \eta) : 0.5305 \le \beta \le 9.0277,\ 41(0.1029)^{\frac{1}{\beta}} \le \eta \le 41(1.1318)^{\frac{1}{\beta}}\}$$

with area 172.502. The plots of the confidence regions for MACR, A_2 and B are displayed in Figure 1, where the MACR has the smallest area and better shape.

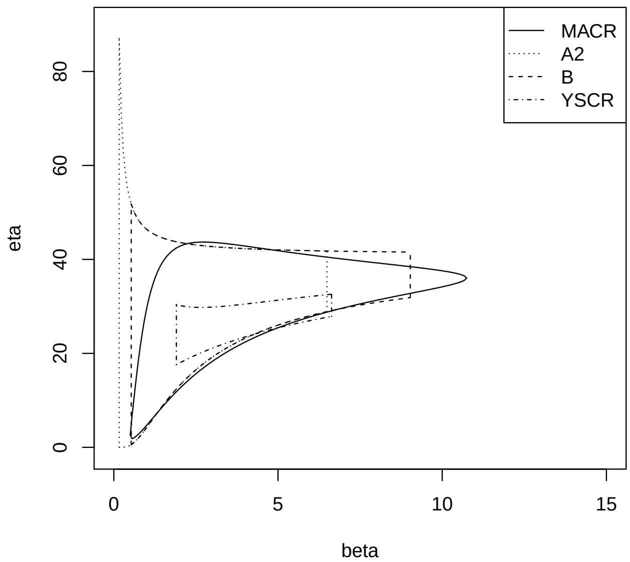

Figure 1: 95% confidence regions MACR, A_2, B and YSCR for (β, η).

For comparison, consider the confidence region (YSCR) of (β, η) in Chen [12] for a complete sample $X = (X_1, X_2, ..., X_n)$. Here the original data set of X is $(n = 19)$

$$26, 14, 27, 15, 16, 16, 11, 10, 14, 12, 15, 40, 29, 13, 20, 41, 31, 28, 11.$$

Then Chen's level 95% confidence region (YSCR) for (β, η) is

$$\begin{cases} 1.9056 \le \beta \le 6.6327, \\ \left(\dfrac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{60.0972}\right)^{\frac{1}{\beta}} \le \eta \le \left(\dfrac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{21.2138}\right)^{\frac{1}{\beta}}, \end{cases}$$

which has area 34.2436 and is also plotted in Figure 1. Clearly, the YSCR is much more accurate, but it is based on a complete sample with $n = 19$.

A. APPENDIX: R code for computing k_α, k_1, k_2, $k_{11}(x)$, $k_{12}(x)$ in (4.2)

```
# Compute the critical value k, k1, k2, k11(x), k12(x) at level Pk=1-c.
# n= sample size
f <- function(n,c) {a <- (n-1)*(1-log(n-1))+n*(1-log(n)); b <- 50
g <- function(x) x-(n-1)*log(x)
h <- function(y) y-n*log(y)
# Step 1: find k1 < k2 so that g(k1)=g(k2)=k-h(n).
k1k2 <- function(n,k) {a <- 0; b <- n-1
 kk<- k-(n-n*log(n))
 for (i in 1:50) if (g((a+b)/2)<kk) b <- (a+b)/2 else a <- (a+b)/2
 k1 <- (a+b)/2; a <- n-1; b <- n+100
 for (i in 1:50) if (g((a+b)/2)<kk) a <- (a+b)/2 else b <- (a+b)/2
 k2 <- (a+b)/2; c(k1,k2)}
# Step 2: find k11(x) < k12(x) so that h(k11(x))=h(k12(x))=k-g(x).
k11k12 <- function(n,k,x) {a <- 0; b <- n
 kk<- k-g(x)
 for (i in 1:50) if (h((a+b)/2)<kk) b <- (a+b)/2 else a <- (a+b)/2
 k11x <- (a+b)/2; a <- n; b <- n+100
 for (i in 1:50) if (h((a+b)/2)<kk) a <- (a+b)/2 else b <- (a+b)/2
 k12x <- (a+b)/2; c(k11x,k12x)}
# Step 3: find k so that Pk=1-c.
Int<- function(n,k) {N <- 1000
 K <- k1k2(n,k)
 H <- (K[2]-K[1])/N; df <- 2*(n-1)
 P <- function(x) {
 KK <-k11k12(n,k,x)
 2*dchisq(2*x,df)*(pchisq(2*KK[2],2*n)-pchisq(2*KK[1],2*n))}
 x1 <- K[1]+((1:N)-0.5)*H ; x2 <- K[1]+(1:(N-1))*H
 s1<-0; s2<-0
 for (j in 1:N) s1<- s1+P(x1[j])
 for (j in 1:(N-1)) s2<- s2+P(x2[j])
 Pk <- H/6*(P(K[1])+P(K[2])+4*s1+2*s2); c(Pk,K) }
 for (i in 1:100) {
 R <- Int(n,(a+b)/2)
 if (R[1]<1-c) a <-(a+b)/2 else b<-(a+b)/2}
 k <- (a+b)/2; list(k=k,k1=R[2],k2=R[3])}
```

ACKNOWLEDGMENTS

This research is supported by the National Natural Science Foundation of China (NSFC, Grant No. 11561073, 11461079), and by the Young Teachers Development Program at School of Mathematics and Statistics, Hainan Normal University, China (hnkfj201804).

REFERENCES

[1] Ahsanullah, M. (1995). *Introduction to Record Statistics*, NOVA Science Publishers Inc., Huntington, New York.

[2] Arnold, B.C.; Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*, John Wiley & Sons, New York.

[3] Asgharzadeh, A. and Abdi, M. (2011a). Exact confidence intervals and joint confidence regions for the parameters of the Gompertz distribution based on records, *Pakistan Journal of Statistics*, **27**(1), 55–64.

[4] Asgharzadeh, A. and Abdi, M. (2011b). Joint confidence regions for the parameters of the Weibull distribution based on records, *ProbStat Forum*, **4**, 12–24.

[5] Asgharzadeh, A. and Abdi, M. (2011c). Confidence intervals and joint confidence regions for the Two-Parameter Exponential distribution based on records, *Communications of the Korean Statistical Society*, **18**(1), 103–110.

[6] Asgharzadeh, A. and Abdi, M. (2012). Confidence intervals for the parameters of the Burr Type XII distribution based on records, *International Journal of Statistics and Economics*, **8**, 96–104.

[7] Asgharzadeh, A.; Abdi, M. and Kuş, C. (2011). Interval estimation for the Two-Parameter Pareto distribution based on record values, *Selçuk Journal of Applied Mathematics*, special issue, 149–161.

[8] Bickel, P.J. and Doksum, K.A. (2001). *Mathematical Statistics: Basic Ideas and Selected Topics*, Vol. I, 2nd ed., New Jersey, Prentice Hall.

[9] Casella, G. and Berger, R.L. (2002). *Statistical Inference*, 2nd ed., Pacific Grove, CA, Duxbury Press.

[10] Chan, P.S. (1998). Interval estimation of location and scale parameters based on record values, *Statistics and Probability Letters*, **37**, 49–58.

[11] Chandler, K.N. (1952). The distribution and frequency of record values, *Journal of the Royal Statistical Society Series B*, **14**, 220–228.

[12] Chen, Z.M. (1998). Joint estimation for the parameters of Weibull distribution, *Journal of Statistical Planning and Inference*, **66**, 113–120.

[13] Jafari, A.A. and Zakerzadeh, H. (2015). Inference on the parameters of the Weibull distribution using records, *SORT*, **39**(1), 3–18.

[14] Jeyaratnam, S. (1985). Minimum volume confidence regions, *Statistics & Probability Letters*, **3**, 307–308.

[15] Johnson, N.L.; Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., Wiley & Sons, New York.

[16] Murthy, D.N.P.; Xie, M. and Jiang, R. (2004). *Weibull Models*, Wiley, Hoboken.

[17] Nelson, W. (1982). *Applied Life Data Analysis*, John Wiley & Sons, INC., New York.

[18] R Core Team (2014). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

[19] Roberts, E. (1979). Review of statistics of extreme values with applications to air quality data: Part II. Applications, *Journal of the Air Pollution Control Association*, **29**, 733–740.

[20] Soliman, A.A. and Al-Aboud, F.M. (2008). Bayesian inference using record values from Rayleigh model with application, *European Journal of Operational Research*, **185**, 659–672.

[21] Soliman, A.A.; Abd Ellah, A.H. and Sultan, K.S. (2006). Comparison of estimates using record statistics from Weibull model: bayesian and non-bayesian approaches, *Computational Statistics & Data Analysis*, **51**, 2065–2077.

[22] Teimouri, M. and Nadarajah, S. (2013). Bias corrected MLEs for the Weibull distribution based on records, *Statistical Methodology*, **13**, 12–24.

[23] Wang, L. and Shi, Y.M. (2013). Reliability analysis of a class of exponential distribution under record values, *Journal of Computational and Applied Mathematics*, **239**, 367–379.

[24] Wang, B.X. and Ye, Z.S. (2015). Inference on the Weibull distribution based on record values, *Computational Statistics* & *Data Analysis*, **83**, 26–36.

[25] Wu, J.W. and Tseng, H.C. (2006). Statistical inference about the shape parameter of the Weibull distribution by upper record values, *Statistical Papers*, **48**, 95–129.

[26] Ye, Z.S.; Hong, Y. and Xie, Y. (2013). How do heterogeneities in operating environments affect field failure predictions and test planning, *The Annals of Applied Statistics*, **7**(4), 2249–2271.

[27] Zakerzadeh, H. and Jafari, A.A. (2015). Inference on the parameters of two Weibull distributions based on record values, *Statistical Methods and Applications*, **24**, 25–40.

[28] Zhang, J. (2017). Minimum-volume confidence sets for parameters of normal distributions, *AStA-Advances in Statistical Analysis*, **101**, 309–320.

[29] Zhang, J. (2018). Minimum volume confidence sets for two-parameter exponential distributions, *The American Statistician*, **72**(3), 213–218.

[30] Zhao, X.; Cheng, W.B; Zhang, Y.; Zhang, Q.N. and Yang, Z.H. (2015). New statistical inference for the Weibull distribution, *The Quantitative Methods for Psychology*, **11**(3), 139–147.