The Sampling Distribution of the Maximum Likelihood Estimators for the Parameters of Weibull Distribution Based on Upper Record Values

Ashour, S. K.[*] Amin, E. A.[*]

Abstract

Record values and associated statistics are of great important in several real-live problems involving weather, economic, and sport data. Based on upper record values and for the two parameters Weibull distribution, Ashour and Amin (2005) obtained the maximum likelihood estimators for the unknown parameters. In this paper the sampling distributions and their properties for these estimators will be investigated numerically.

Keywords and phrase: Weibull distribution; Upper record values; Maximum likelihood estimation; Sampling distribution;

1-Introduction

A random variable is said to have the Weibull(θ, β) distribution if its probability density function (pdf) is given by

$$f(x;\theta,\beta) = \theta\beta x^{\beta-1} e^{-\theta x^{\beta}} \qquad x > 0, \theta > 0, \beta > 0 \tag{1.1}$$

and its cumulative distribution function will be

$$F(x;\theta,\beta) = 1 - e^{-\theta x^{\beta}} \qquad x > 0,\ \theta > 0,\ \beta > 0\,. \tag{1.2}$$

Since the hazard function of distribution (1.1) is a decreasing function when the shape parameter β is less than 1, a constant when β equal 1, and an increasing function when β greater than 1, the distribution becomes suitable for the area of reliability, life testing and quality control (see Johnson et al (1995)).

[*] Department of Mathematical Statistics, I.S.S.R Cairo University

Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed (i.i.d) random variables with cumulative function $F(x)$ and corresponding pdf $f(x)$. Set $Y_n = \max\{X_1, X_2, \cdots, X_n\}$ for $n \geq 1$, we say X_j is an upper record value of $\{X_n\}$, if $Y_{j+1} > Y_j$, by definition, X_1 is an upper record values. The indices at which the upper record values occur are called upper record times $\{U(m), m \geq 0\}$, where $U(0) = 1$ and $U(m) = \min\{j : j > U(m-1), X_j > X_{U(m-1)}\}$. Then $R_m = X_{U(m)}$, $m \geq 0$ are called the upper record values.

The statistical study of record values started with Chandler (1952) and has now spread in different directions. Interested readers may refer to Foster and Stuart (1954),Galambos (1978), Dunsmore (1983), Resnick (1987), Nagaraja (1988), Ahsanullah (1994) and Arnold et al. (1992, 1998) for a review of developments in this area of research.. Ahsanullah (1994) studied the record values based on the Weibull distribution and obtained the best linear unbiased estimators for its parameters. El-Qasem (1996) used the upper record values to obtain the maximum likelihood estimator for the uniform, the exponential and the Pareto distribution with one parameter. Ashour and Amin (2005) obtained the maximum likelihood estimators for the parameters of Weibull distribution based on upper record values.

Present work is devoted to obtain, numerically, the sampling distribution of the maximum likelihood estimators for the unknown parameters of Weibull distribution (1.1) based on the upper record values. Mathcad 2001 package will be used to obtain such distribution.

2- Sampling Distribution of the Maximum Likelihood Estimators.

Let $X_1, X_2, \ldots$ be an infinite sequence of independent and identically distributed random variables having the Weibull (θ, β) distribution (1.1). Consider $R_0, R_1, R_2, \ldots, R_m$ represent the first (m+1) upper records from Weibull (θ, β), the likelihood function based on $(m+1)$ upper record values $R_0, R_1, \ldots, R_m$ is given by

$$L(\theta,\beta)=\frac{\theta^{m+1}\,\beta^{m+1}}{m!}\left(\prod_{i=0}^{m} r_i\right)^{\beta-1} e^{-\theta\, r_m^{\beta}} \qquad \theta,\beta>0\,. \tag{2.1}$$

Ashour and Amin (2005) obtained the maximum likelihood estimators for the parameters of Weibull distribution (1.1) as follows:

Taking the logarithm of the likelihood function (2.1) then we have

$$\ln L \propto (m+1)\ln\theta+(m+1)\ln\beta+(\beta-1)\sum_{i=0}^{m}\ln r_i-\theta\, r_m^{\beta}\,. \tag{2.2}$$

Differentiate (2.2) with respect to θ and β respectively we get

$$\frac{\partial \ln L}{\partial\theta}=\frac{(m+1)}{\theta}-r_m^{\beta} \tag{2.3}$$

and

$$\frac{\partial \ln L}{\partial\beta}=\frac{(m+1)}{\beta}+\sum_{i=0}^{m}\ln r_i-\theta\, r_m^{\beta}\ln r_m \tag{2.4}$$

Equating (2.3) and (2.4) by zero and solving with respect to $\hat{\theta}$ and $\hat{\beta}$ then the maximum likelihood estimators $\hat{\theta}$ and $\hat{\beta}$ for θ and β will be

$$\hat{\beta}=\frac{(m+1)}{m\ln r_m-\sum_{i=0}^{m-1}\ln r_i} \tag{2.5}$$

and

$$\hat{\theta}=\frac{(m+1)}{r_m^{\hat{\beta}}} \tag{2.6}$$

Although estimators (2.5) and (2.6) are in exact form but there are some difficulties to obtain the exact distributions for these estimators, so we will try to derive such distributions numerically. STATGRAPHICS package will be used to obtain the best fitted distributions that is, we obtained the best fit for each case and tested the goodness of fit using the Chi-square test and the Kolmogrov-Smirnove test.

The following steps will be used to obtain the sampling distribution for the maximum likelihood estimators $\hat{\beta}$ and $\hat{\theta}$.

(1) Generate $n = 100$ independent and identically distributed random variates $X_1, X_2, \dots X_n$ from Weibull (1.1) with θ=2,3,5,7,9 and β=0.5,1,(1.5),2,(2.5),3,4,5.

(2) Do step (1) 1000 times, therefore we have 1000 vectors of (i.i.d) random variates.

(3) Choose from each vector the first $m = 3$ upper record values, then we have k vectors containing four upper record values where k<1000.

(4) Compute the values of $\hat{\beta}$ and $\hat{\theta}$ in each vector, then we have two vectors of length k containing the random variables from the distribution of $\hat{\beta}$ and $\hat{\theta}$.

(5) Do steps 1 to 5, for $m = 5$ and $m = 7$.

(6) The distribution of each vector was fitted using STATGRAPHICS package and the results recorded in Tables 1 to 4 with significance level $\alpha = 0.01$.

Tables (1 to 4) contain the results which show

(i) the values of the parameters θ and β

(ii) the sampling distribution of the random variables $\hat{\beta}$ and $\hat{\theta}$.

(iii) the mean and standard deviation for each distribution.

(iv) The fitted distribution for the estimators $\hat{\beta}$ and $\hat{\theta}$.

(v) The P-value for the Chi-square test and the Kolmogorov-Smirnov test.

From these tables we conclude the following

(1) In all considered values of the parameters and for all numbers of record values the fitted distribution of the maximum likelihood estimator for the shape parameter $\hat{\beta}$ is lognormal distribution.

(2) The fitted distribution for the maximum likelihood estimator $\hat{\theta}$ is always lognormal distribution except some cases when $m = 3$, but when m is large then the fitted distribution is lognormal.

Table (1)

Mean, standard deviation, P-value and sampling distribution of the MLE's $\hat{\theta}$ when m=3 for Different Values of β

Parameters		Sampling distribution of $\hat{\theta}$				
	β	Fitted	Mean	St.dev	P-Value	
					Chi	Kolm
$\theta=2$	0.5	Gam	1.629*	0.642+	0.023	0.21
	1	Gam	1.756*	0.629+	0.056	0.29
	1.5	Gam	1.881*	0.723+	0.077	0.19
	2	Chi-Square with 3.239 df			0.0174	0.001
	2.5	Chi-Square with 2.702 df			0.041	0.063
	3	Gam	1.584*	0.646+	0.0163	0.074
	4	Gam	1.480*	0.545+	0.037	0.032
	5	Chi-Square with 3.031 df			0.012	0.008
$\theta=3$	0.5	Lognormal	4.613	3.791	0.012	0.074
	1		5.065	4.348	0.289	0.348
	1.5		4.12	3.16	0.04	0.11
	2		4.667	3.906	0.04	0.15
	2.5		4.251	3.25	0.014	0.37
	3		4.36	3.23	0.011	0.067
	4		4.97	4.27	0.011	0.253
	5		5.026	4.153	0.025	0.168
$\theta=5$	0.5	Lognormal	6.63	3.16	0.047	0.387
	1		7.866	4.633	0.013	0.11
	1.5		7.57	4.424	0.104	0.077
	2		7.56	4.47	0.004	0.122
	2.5		8.06	4.936	0.004	0.099
	3		7.982	5.041	0.0334	0.0355
	4		7.78	4.27	0.001	0.17
	5		8.64	5.78	0.096	0.099
$\theta=7$	0.5	Lognormal	13.04	9.2686	0.0012	0.289
	1		13.257	9.856	0.0013	0.144
	1.5		14.746	11.922	0.001	0.02
	2		13.07	9.655	0.003	0.057
	2.5		12.712	8.336	0.017	0.264
	3		16.455	14.609	0.004	0.37
	4		12.56	8.367	0.083	0.213
	5		11.313	7.2236	0.004	0.054
$\theta=9$	0.5	Lognormal	15.188	8.827	0.004	0.272
	1		14.805	9.389	0.008	0.387
	1.5		14.72	9.235	0.001	0.264
	2		13.65	8.0754	0.002	0.175
	2.5		17.593	12.926	0.003	0.134
	3		15.141	9.034	0.026	0.497
	4		13.719	8.443	0.004	0.03
	5		15.497	10.161	0.01	0.139

* Shape parameter

+ Scale parameter

Table (2)

Mean, standard deviation, P-value and sampling distribution of the MLE's $\hat{\theta}$ when m=5 for Different Values of β

Parameters		Sampling distribution of $\hat{\theta}$				
	β	Fitted	Mean	St.dev	P-Value	
					Chi	Kolm
$\theta = 2$	0.5	Erlang	4*	1.300+	0.027	0.179
	1	Erlang	4*	1.404+	0.129	0.115
	1.5	Weib	2.311*	3.448+	0.349	0.417
	2	Gam	4.962*	1.687+	0.252	0.208
	2.5	Weib	2.352*	3.291+	0.376	0.605
	3	Gam	3.209*	1.071+	0.011	0.093
	4	Logn	3.1329	1.771	0.166	0.38
	5		3.5158	3.155	0.166	0.657
$\theta = 3$	0.5	Gamm	4.958*	3.280+	0.006	0.014
	1	Lognormal	4.6688	2.028	0.048	0.59
	1.5		4.9218	2.368	0.01	0.045
	2		4.9563	2.387	0.042	0.235
	2.5		5. 08	2.51	0.27	0.073
	3		5.046	2.406	0.071	0.031
	4		4.48	1.86	0.037	0.073
	5		5.21	2.665	0.128	0.213
$\theta = 5$	0.5	Lognormal	8.062	2.99	0.012	0.022
	1		8.226	3.463	0.004	0.024
	1.5		9. 041	4.42	0.046	0.204
	2		8.582	3.306	0.018	0.312
	2.5		9.3042	3.663	0.02	0.11
	3		9.107	4.071	0.003	0.056
	4		9.1418	3.970	0.231	0.788
	5		8.256	3. 699	0.003	0.068
$\theta = 7$	0.5	Lognormal	12.474	5.975	0.002	0.147
	1		12.912	6.434	0.164	0.282
	1.5		11.64	4.793	0.012	0.146
	2		12.08	5.306	0.056	0.075
	2.5		12.599	6.552	0.013	0.144
	3		13.726	7.382	0.11	0.105
	4		13.208	7.259	0.005	0.115
	5		12.722	6.346	0.004	0.227
$\theta = 9$	0.5	Lognormal	13.415	5.951	0.043	0.29
	1		13.478	4.804	0.138	0.721
	1.5		13.380	4.968	0.176	0.547
	2		13.749	5.066	0.702	0.904
	2.5		15.056	6.063	0.013	0.427
	3		14.377	6.435	0.29	0.61
	4		13.524	5.099	0.086	0.473
	5		14.855	7.274	0.43	0.866

* Shape parameter

\+ Scale parameter

Table (3)

Mean, standard deviation, P-value and sampling distribution of the MLE's $\hat{\theta}$ when m=7 for Different Values of β

Parameters		Sampling distribution of $\hat{\theta}$				
	β	Fitted	Mean	St.dev	P-Value	
					Chi	Kolm
$\theta = 2$	0.5	Lognormal	3.9161	1.4415	0.055	0.55
	1		3.8528	1.4424	0.22	0.65
	1.5		3.9958	1.4729	0.29	0.46
	2		3.8798	1.4088	0.688	0.651
	2.5		3.8525	1.3672	0.242	0.474
	3		3.8873	1.5914	0.18	0.129
	4		3.9768	1.6428	0.053	0.56
	5		4.0041	1.3389	0.43	0.56
$\theta = 3$	0.5	Lognormal	5.8872	1.6343	0.041	0.44
	1		5.8158	1.6255	0.23	0.576
	1.5		5.6889	1.6129	0.604	0.774
	2		5.6603	1.8313	0.007	0.342
	2.5		5.7421	1.6628	0.116	0.289
	3		6.0767	1.9657	0.098	0.083
	4		5.9840	1.8694	0.211	0.741
	5		5.6471	1.7654	0.142	0.851
$\theta = 5$	0.5	Lognormal	10.153	3.7632	0.097	0.296
	1		10.590	3.8985	0.044	0.195
	1.5		10.255	3.8334	0.023	0.547
	2		10.412	3.9257	0.278	0.675
	2.5		10.072	3.2848	0..016	0.463
	3		10.153	3.7632	0.097	0.296
	4		10.918	4.3540	0.018	0.241
	5		10.229	3.7090	0.643	0.261
$\theta = 7$	0.5	Lognormal	15.090	6.9889	0.003	0.026
	1		14.631	5.0489	0.004	0.038
	1.5		13.068	4.5442	0.031	0.094
	2		13.259	4.4540	0.021	0.147
	2.5		13.819	5.0238	0.167	0.558
	3		13.351	4.2381	0.365	0.758
	4		14.311	5.5735	0.014	0.091
	5		13.481	4.7478	0.144	0.661
$\theta = 9$	0.5	Lognormal	17.291	5.4885	0.01	0.142
	1		16.363	5.0423	0.036	0.402
	1.5		16.337	5.3426	0.017	0.313
	2		15.993	5.2219	0.004	0.326
	2.5		16.587	5.492	0.007	0.327
	3		16.657	4.8891	0.263	0.873
	4		16.223	5.3261	0.007	0.194
	5		16.389	5.2471	0.447	0.837

Table (4)

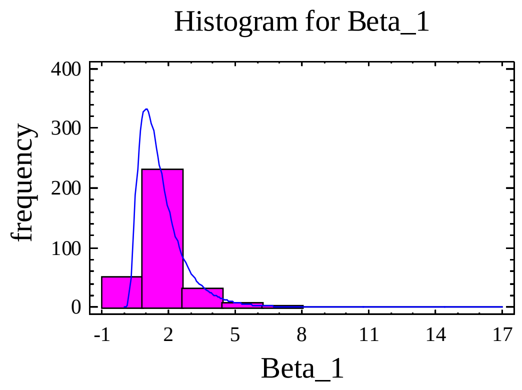
Mean, standard deviation, P-value and sampling distribution of the MLE's $\hat{\beta}$

Parameters		Sampling distribution of $\hat{\beta}$				
	β	Fitted	Mean	St.dev	P-Value	
					Chi	Kolm
$m = 3$	0.5	Lognormal	0.9059	0.5721	0.048	0.35
	1		1.6904	1.1157	0.35	0.12
	1.5		2.5237	1.6180	0.17	0.44
	2		3.5737	2.3628	0.29	0.43
	2.5		4.7255	3.2357	0.034	0.06
	3		5.5139	4.9432	0.016	0.015
	4		7.331	5.0201	0.016	0.61
	5		8.5234	5.4829	0.019	0.21
$m = 5$	0.5	Lognormal	0.6381	0.2898	0.183	0.431
	1		1.3827	0.7164	0.422	0.767
	1.5		1.8719	0.9651	0.803	0.682
	2		2.4629	1.0652	0.055	0.533
	2.5		3.2644	1.4972	0.028	0.183
	3		4.0561	2.035	0.44	0.32
	4		5.0724	2.7672	0.853	0.566
	5		6.3973	2.7672	0.101	0.404
$m = 7$	0.5	Lognormal	0.556	0.222	0.258	0.305
	1		1.062	0.366	0.223	0.262
	1.5		1.584	0.53	0.422	0.185
	2		2.153	0.789	0.735	0.997
	2.5		2.647	0.928	0.339	0.612
	3		3.325	1.225	0.124	0.487
	4		4.366	1.567	0.397	0.425
	5		5.418	1.936	0.705	0.505

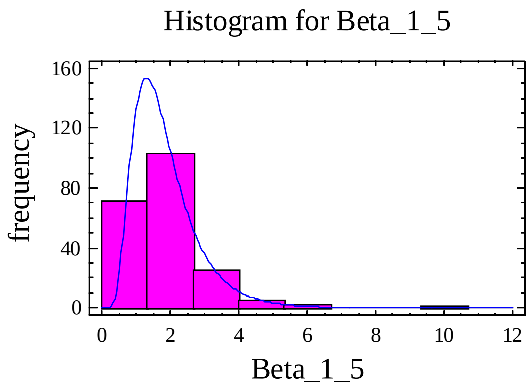
* Shape parameter

\+ Scale parameter

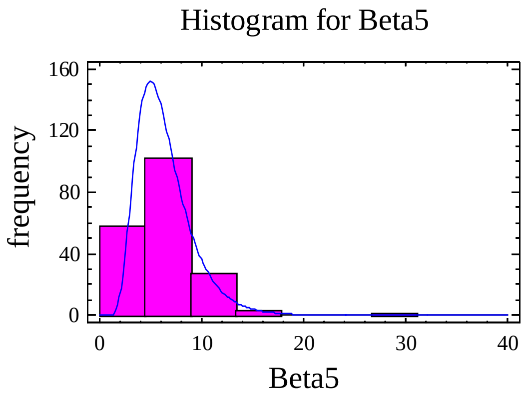
Graph (1)
Histogram for Distribution of $\beta = 1$ when $m = 3$



Graph (2)
Histogram for Distribution of $\beta = 1.5$ when $m = 5$

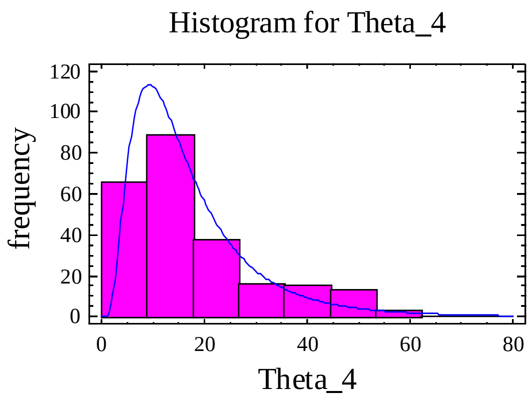


Graph (3)
Histogram for Distribution of $\beta = 5$ when $m = 7$



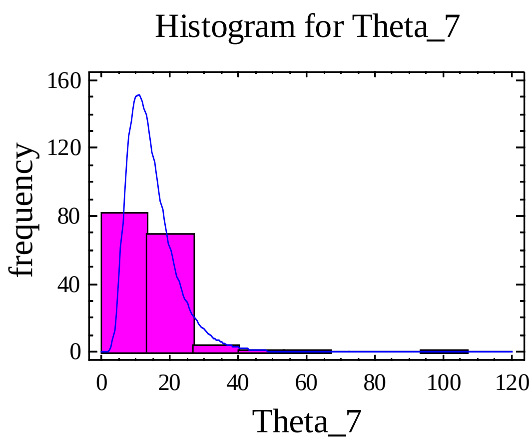
Graph (4)

Histogram for Distribution of $\theta = 7$ **when** $\beta = 2.5$ **and** $m = 3$



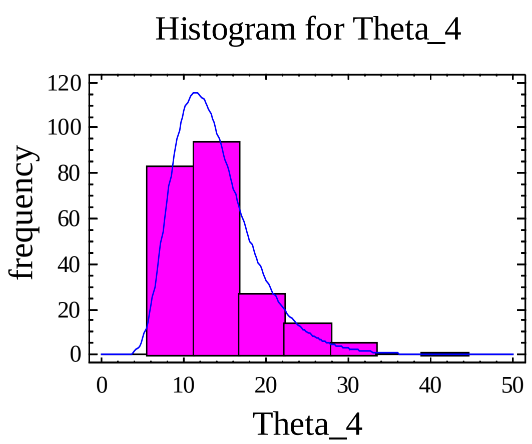
Graph (5)

Histogram for Distribution of $\theta = 7$ **when** $\beta = 5$ **and** $m = 5$



Graph (6)

Histogram for Distribution of $\theta = 7$ **when** $\beta = 2.5$ **and** $m = 7$



References

Ahsanullah, M. (1994). Record statistic. Nova Science Publishers, Inc.

Arnold, C. B., Balakrishnan, N. & Nagaraja, H. N. (1992). A first course in order statistics. John Wiley & Sons, Inc.

Arnold, C. B., Balakrishnan, N. & Nagaraja, H. N. (1998). Records. John Wiley & Sons, Inc.

Ashour, S. K. and Amin, E. A. (2005). Bayesian and Non-Bayesian Estimation About Weibull Distribution Based on Upper Record Values. Under Publication.

Chandler, K. N. (1952). The distribution and frequency of record values, Journal of royal statistics society, series B, 14, 220-228.

Dunsmore, I. R. (1983). The future occurrence of records, Annals of the Institute Statistical Mathematics, 51, 267-277.

El-Qasem, A. A. (1996). Estimation via record values. Journal of Information & Optimization Sciences. 17, 541-548.

Foster, F. G. and Stuart, A. (1954). Distribution free tests in time series based on the braking of records, Journal of royal statistics society, series B, 16, 1-22.

Galambos, J. (1978). The asymptotic theory of extreme of extreme order statistics, New York: John Wiley & Sons.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). Continuous univariate distribution, Volume 2, Second edition, John Wiley & Sons, New York.

Nagaraja, H. N. (1988). Record values and related statistics. A review. Communication in Statistics- Theory and Methods,7, 7, 2223-2238.

Resnick, S. I. (1973). Record values and maxima. Annals of Probability, 1, 650-662