Commun. Fac. Sci. Univ. Ank. Ser. A_1
V. 36, Number 2, pp. 143-152 (1987)

COMPUTATIONAL TECHNIQUES FOR MAXIMUM LIKELOHOOD ESTIMATION OF THREE - PARAMETER WEIBULL DISTRIBUTION

FETİH YILDIRIM[1] and CELAL KORAŞLI[2]

1. *Department of Statistics, METU-Ankara, Turkey*
2. *Department of Electrical and Electronic Engineering, METU, Gaziantep, Turkey*

ABSTRACT

Two improved numerical techiques are employed to compute MLE for the parameters of the 3-parameter Weibull distribution using complete sample data of large sample size and of large-valued inspection points. For the solution of estimates a new minimal functional is defined and optimized by the descent method and the method of conjugate gradients. The result of both techniques are compared. Asymptotic variance-covariance matrix of estimates for the complete sample is included.

INTRODUCTION

The Weibull distıibution has many uses in engineering, reliability, applied statistics and business studies. Despite the appropriateness of the 3-parameter Weibull model in many practical cases, often 1– and 2– parameter Weibull models are used. The reasons for this are difficulties of analysing the 3–parameter model and of computational procedures.

The well adapted computational procedure for estimating parameters is the maximum likelihood (ML). This method is used for several reasons:

1– For large samples estimates are almost unbiased and have minimum variance.

2– Estimates have the asymptotic properties of consistency have s-normality.

3– For small samples, the ML estimators are generally comparable with other estimates.

4– When the value of the shape parameter is greater than two, the regularity of the distribution is usually satisfied.

ISSN 0251 - 0871, A. Ü. Basımevi

The number of papers published on computational techniques for solving the ML equations for the 3–parameter Weibull model are scarce. Iterative estimation procedure was applied by Harter and Moore [1] for complete and censored data, by Wingo [2] for progressively censored samples, by Zanakis [3] and Archer [4] for complete and grouped data,

In this paper two numerical techniques for computing the ML estimates of the parameters of the 3– parameter Weibull distribution for complete samples are discussed. The method of approximating the variance and coveriance of estimates by the negative inverse of the matrix formed by the second partial derivatives of the logarithm of the likelihood function are examined.

THE MAXIMUM LIKELIOOD EOUATIONS

The 3–parameter Weibull survival function is

$$R(x; \alpha, \beta, \gamma) = \exp\{-[(x-\gamma)/\alpha]^{\beta}\} \tag{1}$$

with $x > \gamma$; $\alpha\, \beta\, \gamma > 0$ where α, β and γ are the scale, shape and thresbold parameteıs, respectively. Using reparametrization

$$\theta = \alpha^{\beta} \quad \text{and} \quad \alpha = \theta^{1/\beta}$$

for the complete samples the logarithm of the likelihood functions

$$\ln L = n \ln\beta - n\ln\theta - (\beta-1) \sum_{i=1}^{n} \ln (x_i - \gamma) - \frac{1}{\theta} \sum_{i=1}^{n} (x_i - \gamma)^{\beta} \tag{2}$$

yield the likelihood equations

$$\frac{\partial \ln L}{\partial\beta} = \frac{-n}{\beta} \ln\theta + \sum_{i=1}^{n} \ln (x_i-\gamma) - \frac{1}{\theta} \sum_{i=1}^{n} (x_i-\gamma)\beta \ln (x_i-\gamma) = 0 \tag{3a}$$

$$\frac{\partial \ln L}{\partial\gamma} = (1-\beta) \sum_{i=1}^{n} (x_i-\gamma)^{-2} + \frac{n}{\beta\theta} \sum_{i=1}^{n} (x_i-\gamma)^{\beta-1} = 0 \tag{3b}$$

$$\frac{\partial \ln L}{\partial\theta} = - \frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} (x_i-\gamma)^{\beta} = 0 \tag{3c}$$

If the 3–parameter Weibull distribution has an interior relative maximum, the solution to the likelihood equations renders the ML estimators $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ which maximize Eqn. (2).

ITERATIVE ESTIMATION PROCEDURE

The nonlinear nature of the ML function implies many sharp or flat maxima and minima, hence any iterative solution to the set of the the ML equations necessitates selecting good initial estimates. This provides lesser iterations for convergence.

The false position method is used by Harter and Moore [1] for the solutions of the ML equations. This technique is a one-at-a-time parameter research whose convergence rate is slow. However it is more stable to divergence than any other method. Wingo [2] applied the modified Newton–Raphson method for the solution of the ML equations. In this method the iteration step–size is adjusted to eliminate divergence at the points close to the solution point. This method was found to have high rate of convergence. Archer [4] combined both the false position and the modified Newton–Raphson methods. His hybrid technique was observed to be faster than the false position method but slower than the modified Newton–Raphson technique.

In the present paper two iterative search techniques are given for finding the estimates of 3–parameter Weibull parameters from the ML estimators by (i) the method of conjugate gradients developed by Fletcher and Reeves [6] requires lesser computer storage and iterates faster as compared to (ii) the decent method described by Fletcher and Powell [7]. However, both methods suffer as the selection of initial estimates are forced to out of the constraints imposed on the parameters defined by the experimental data.

Method of Solution

Solving Eqn. (3c) for θ and substituting it in Eqns. (3a) and (3b), a new form of the ML equations follow

$$H(x;\beta,\gamma) = P_1(x,\gamma) - n\left[\frac{Q_{21}(x;\beta,\gamma)}{Q_{11}(x;\beta,\gamma)}\right] - \beta^{-1} \tag{4a}$$

and

$$G(x;\beta,\gamma) = (1-\beta)\,R_1(x;\gamma) + n\left[\frac{Q_{12}(x;\beta,\gamma)}{Q_{11}(x;\beta,\gamma)}\right], \tag{4b}$$

where

$$P_k(x;\gamma) = \sum_{i=1}^{n} [\ln(x_i-\gamma)]_K{}^k$$

$$Q_{lm}(x;\beta,\gamma) = \frac{\sum_{i=1}^{n} (x_i-\gamma)^{\beta} [\ln (x_i-\gamma]^{l}}{\left[\sum_{i=1}^{n} (x_i-\gamma)^{\beta}\right]^{m}},$$

and

$$R_n(x;\gamma) = \sum_{i=1}^{n} (x_i-\gamma)^{-n}.$$

The minimization procedures require the solution of a new function

$$J(x;\beta,\gamma) = H^2(x;\beta,\gamma) + G^2(x;\beta,\gamma) = 0. \tag{5}$$

Since this form of equation is noninteractive the solution exists at one of those local minimums if the starting estimates for parameters $(\hat{\beta}_0, \hat{\gamma}_0)$ and hence the function $\hat{J}_0$ are appropriately done.

Initial Estimates

1– In the present work a single starting estimate was selected for only the threshold parameter γ. The estimate for the shape parameter was obtained from the slope of the Weibull graph at points $[F(x),(x_i-\gamma)]_{0\cdot 8}$ and $_{0\cdot 9}$, since the slopes at these points was generally observed to be less affected with $\hat{\alpha}$ an $\hat{\gamma}$ values slected over wide ranges.

2– Another estimate for the iterative procedure was for the function J. Since each consitutent of it (Eqn. (5)) should necessarily aproach to zero when the final convergence was achieved, a simple choice for the estimate was zero.

Variances and Covariances of Estimates

The information on the variance–covariance of estimates was suggested [8] to be approaximated from the asymptotic variance–covariance matrix of the MLE by inverting the information matrix whose elements are negatives of second partial derivatives of ML equations (Eqns. 3 (a)–3(c)).

Simple Computational Algorithm

For both irerative procedures sequences in computation is as follows:

1– $J(x)$ and $\Delta J(x)$ computed and values of $\hat{J}^0_{est}=0$, limits of iterations and other constraints are specified.

2– From the data (using $\gamma^\circ, \beta^\circ$) $\hat{\gamma}^0, \hat{\beta}^0$ are computed.

3– Correction for $\hat{\beta}^0$ is made for a given value $\hat{\gamma}^0$ by using Newton-Raphson estimation technique.

4– A new value for the function $\hat{J}$ is estimated, $\hat{J}^k_{est}$, which results in new estimates $(\hat{\beta}^k, \hat{\gamma}^k)$. Now $\hat{J}^k_{est}$ is closer to the solution than $\hat{J}^0_{est} = 0$ assumption.

5– Optimization procedure either using conjugate gradients or descent method is entered for the solution of $\hat{\beta}^k$ an $\hat{\gamma}^k$.

6– If the outcomes of step 4 fails, a new estimate is computed for $\hat{\gamma}^k > \hat{\gamma}^{k-1}$.

7– A new value of $\hat{J}^k_{est}$ is computed using

$$\hat{J}^k_{est} = 1.5 \ \hat{J}^k - 0.5 \ \hat{J}^{k-1}$$

which is necessarily less than $\hat{J}^k$.

8– Control is transferred to step 5 for a second run with slightly corrected new values of $\hat{\beta}^0$, $\hat{\gamma}^0$ and $\hat{J}^0_{est}$ if an optimum point is not reached.

9– Inverse of the information matrix is computed to obtain variances and covariances of estimates $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$.

NUMERICAL RESULTS

Both optimization techniques were written in Fortran. The program is capable of handling of sample size in the range of 5–1000. This range is found to be affected with data values in samples due to overflows and underflows. Since the shape parameter is also another factor governing the computational abnormalities, it is restricted to the range of 0.1–20 that which is the case for most extreme Weibull distributions. The number of iterations is limited to 50 in order to initiate each run with better estimates.

Results of several runs of both numerical techniques for the data set of 200 breakdowns obtained from conditioning tests of the high–voltage gas–insulated cable [9] are illustrated in Table 1. At the final stage of computation variance–covariance matrix calculated from the second derivatives of ML function using the results computed from the method of conjugate gradients is given.

Table 1. Summary of optimized estimates and variance-covariance matrix with *good* initial estimates.

A) Descent Method
Complete Sample of Size N = 200
Graphical Appox. Beta = 4.74 Gamma = 277.50
Newton-Raphson Approx. Beta = 4.29 Gamma = 277.50

*Run = 1 = Iteraion	Beta	Gamma	Functional J
0	0.42877863 D + 01	0.27750487 D + 03	0.27653412 D + 01
2	0.43597129 D + 01	0.27573636 D + 03	0.85257283 D + 00
4	0.45081424 D + 01	0.27573156 D + 03	0.61884110 D + 00
5	0.48582565 D + 01	0.27294351 D + 03	0.22751899 D + 00
6	0.48929624 D + 01	0.27244636 D + 03	0.13828115 D + 00
7	0.52304370 D + 01	0.26957723 D + 03	0.68720760 D — 01
8	0.54757074 D + 01	0.26721438 D + 03	0.32285576 D — 01
9	0.57642786 D + 01	0.26466921 D + 03	0.17284572 D — 01
10	0.66941134 D + 01	0.25603044 D + 03	0.38393291 D — 02
11	0.66600062 D + 01	0.25630253 D + 03	0.25175860 D — 02
12	0.72352804 D + 01	0.25091009 D + 03	0.11353897 D — 02
13	0.74886010 D + 01	0.24847470 D + 03	0.51848803 D — 03
14	0.78533588 D + 01	0.24503441 D + 03	0.25581338 D — 03
15	0.85112786 D + 01	0.23876717 D + 03	0.75761692 D — 04
16	0.84887887 D + 01	0.23896882 D + 03	0.38109672 D — 04
17	0.88907091 D + 01	0.23513153 D + 03	0.11616282 D — 04
18	0.90177568 D + 01	0.23390883 D + 03	0.30850292 D — 05
19	0.91773187 D + 01	0.23238298 D + 03	0.61481543 D — 06
20	0.92806495 D + 01	0.23139137 D + 03	0.15245532 D — 10
21	0.92810289 D + 01	0.23138774 D + 03	0.79470245 D — 13
22	0.92810699 D + 01	0.23138735 D + 03	0.10910647 D — 19
23	0.92810699 D + 01	0.23138735 D + 03	0.30677223 D — 24

(Estimated) Beta = 9.28106993 And Gamma = 231.38735027
Minimized Functional J = 0.30677223D—23
Derivatives of Functional J = —0.36638073D—11 —0.38248740D—12
Alpha (Exp) = 0.27624867D+03 Alpha (Cal) = 0.93114677D+02
Determ = — 0.11027D—22 0.00000D+00J Condı = 0.11721D—40

Second Derivatives of MlE Func		Var-Covmatrix (var-Cov)
(I, J)	A (I, J)	
1,1	0.41972558 D + 04	0.26746375 D + 00
1,2	-0.87178200 D + 02	-0.11169905 D + 01
1,3	-0.48664048 D + 15	0.23062934 D + 19
2,1	-0.87178200 D + 02	-0.11169905 D + 01
2,2	0.11206462 D + 02	0.10710140 D + 00
2,3	0.10162173 D + 16	-0.11552647 D + 18
2,3	0.10162173 D — 15	-0.11552647 D + 18
3,1	-0.48664048 D — 15	0.23062934 D + 19
3,2	0.10162173 D — 15	-0.11552647 D + 18
3,3	0.56472869 D — 34	0.19912389 D + 38

B) Method Of Conjugate Gradients
Complete Sample Of Size N = 200
Graphical Appox. Beta = 4.74 Gamma = 277.50
Newton-Raphson Appox. Beta = 4.29 Gamma = 277.50

* Run = 1 = Iteraion	Beta	Gamma	Functional J
0	0.42877863 D + 01	0.27750487 D + 03	0.27653412 D + 01
1	0.42719279 D + 01	0.27749575 D + 03	0.26889751 D + 01
2	0.43733065 D + 01	0.27704709 D + 03	0.19849247 D + 01
3	0.43960878 D + 01	0.27648160 D + 03	0.11108634 D + 01
4	0.44135438 D + 01	0.27648226 D + 03	0.10349142 D + 01
5	0.44463841 D + 01	0.27605093 D + 03	0.87593812 D + 00
6	0.46333839 D + 01	0.27482135 D + 03	0.41598709 D + 00
7	0.46191319 D + 01	0.27481918 D + 03	0.37240929 D + 00
8	0.46535459 D + 01	0.27460851 D + 03	0.35257422 D + 00
9	0.56068265 D + 01	0.26615372 D + 03	0.30230934 D + 01
10	0.55978776 D + 01	0.26615271 D + 03	0.22008182 D — 01
11	0.56410172 D + 01	0.26578591 D + 03	0.21022338 D — 01
12	0.64063186 D + 01	0.25873153 D + 03	0.64187892 D — 02
13	0.64002760 D + 01	0.25873070 D + 03	0.41758250 D — 02
14	0.64980950 D + 01	0.25783890 D + 03	0.37838839 D — 02
15	0.73157284 D + 01	0.25015492 D + 03	0.11339597 D — 02
16	0.73122612 D + 01	0.25015455 D + 03	0.69209458 D — 03
17	0.75185272 D + 01	0.24821736 D + 03	0.54422161 D — 03
18	0.83839128 D + 01	0.23998387 D + 03	0.10212175 D — 03
19	0.83824146 D + 01	0.23998372 D — 03	0.53517273 D — 04
20	0.87607091 D + 01	0.23637617 D + 03	0.20724731 D — 04
21	0.90050780 D + 01	0.23402917 D + 03	0.40971436 D — 05
22	0.90053305 D + 01	0.23402919 D + 03	0.30518604 D — 05
23	0.92551522 D + 01	0.23163651 D + 03	0.10054380 D — 06
24	0.92623514 D + 01	0.23156676 D + 03	0.11815555 D — 07
25	0.92623554 D + 01	0.23156676 D + 03	0.11577670 D — 07
26	0.92810129 D + 01	0.23138791 D + 03	0.22254628 D — 10
27	0.92810473 D + 01	0.23138757 D + 03	0.16694901 D — 13
28	0.92810473 D + 01	0.23138757 D + 03	0.16663785 D — 13

(Estimated) Beta = 9.28104732 And Gamma = 231.38756701
Minimized Functional J = 0.16663785 D — 13
Derivatives Of Functional J = —0.29099101 D — 10 — 0.15073425 D — 09
Alpha (Exp) = 0.27624867D+03 Alpha (Cal) = 0.93114459D+02
Determ = —0.110280—22 —0.00000D+00J Condi = 0.11724D—40

Second Derivatives of MIE Func (I, J)	A (I, J)	Var-Covmatrix (Var-Cov)
1,1	0.41972517 D + 04	0.26746249 D + 00
1,2	-0.86178140 D + 02	-0.11169909 D + 01
1,3	-0.48670071 D — 15	0.23059948 D + 19
2,1	-0.87178140 D + 02	-0.11159909 D — 01
2,2	0.11206459 D + 02	0.10718142 D + 00
2,3	0.10163434 D — 16	-0.11551210 D + 18
3,1	-0.48670071 D — 15	0.23059948 D + 19
3,2	0.10163434 D — 15	-0.11551210 D + 18
3,3	0.56486904 D — 34	0.19907328 D + 38

Table 2. Summary of optimized estimates and variance-covariance matrix with *bad* initial estimates.

A) Descent Method
Complete Sample Of Sıze N = 200
Graphıcal Appox. Beta = 12.68 Gamma = 181.50
Newton-Raphson Approx. Beta = 14.45 Gamma = 181.50

* Run = 1 = Iteratıon	Beta	Gamma	Functıonal J
0	0.14448848 D + 02	0.18150208 D + 03	0.47369863 D — 04
10	0.15824832 D + 02	0.16816618 D + 03	0.42430360 D — 04
20	0.18149397 D + 02	0.14560410 D + 03	0.35235436 D — 04
30	0.17803644 D + 02	0.14895450 D + 03	0.34401918 D — 04
40	0.19713324 D + 02	0.13039032 D + 03	0.26863918 D — 04
50	0.19713326 D + 02	0.13039030 D + 03	0.26863910 D — 04

Graphıcal Appox. Beta = 10.37 Gamma = 209.43
Newton-Raphson Approx. Beta = 11.56 Gamma = 209.43

* Run = 1 = Iteratıon	Beta	Gamma	Functıonal J
0	0.11562847 D + 02	0.20942549 D + 03	0.39225393 D — 04
10	0.92799149 D + 01	0.23139840 D + 03	0.96999067 D — 10

(Estımated) Beta = 9.28106993 And Gamma = 231.38735027
Mınımızed Functıonal J = 0.66001783D—24
Derıvatıves Of Functıonal J = —0.58501150D—11 —0.60992915D—12
Alpha (Exp) = 0.27624867D+03 Alpha (Cal) = 0.93114677D+02

B) Method Of Conjugate Gradients
Complete Sample Of Size N = 200
Graphical Appox. Beta = 12.68 Gamma = 181.50
Newton-Raphson Approx. Beta = 14.45 Gamma = 181.50

* Run = 1 = Iteration	Beta	Gamma	Functional J
0	0.14448848 D + 02	0.18150208 D + 03	0.47369863 D — 04
10	0.20032348 D + 02	0.12728396 D + 03	0.25670930 D — 04
*** Run = 2 = Iteration**	**Beta**	**Gamma**	**Functional J**
0	0.21151084 D + 02	0.11640272 D + 03	0.22281081 D — 04
*** Run = 3 = Iteration**	**Beta**	**Gamma**	**Functional J**
0	0.21151084 D + 02	0.11640272 D + 03	0.22281081 D — 04
*** Run = 4 = Iteration**	**Beta**	**Gamma**	**Functional J**
0	0.21151084 D + 02	0.11640272 D + 03	0.22281081 D — 04

Complete Sample Of Size N = 200
Graphical Approx. Beta = 10.37 Gamma = 209.43
Newton-Raphson Approx. Beta = 11.56 Gamma = 209.43

* Run = 5 = Iteration	Beta	Gamma	Functional J
0	0.11562847 D + 02	0.20942549 D + 03	0.39225393 D — 04

(Estimated) Beta = 9.93095648 And Gamma = 225.14885217
Minimized Functional J = 0.86819247 D — 05
Derivatives of Functional J = —0.97239726 D — 06 —0.22659340 D — 05
Alpha (Exp) = 0.27624867 D + 03 Alpha (Cal) = 0.99391819 D + 02

CONCLUSION

In the present paper the difficult 3–parameter Weibull distribution parameters are estimated by two iterative procedures for the shape parameter values in 0.1–20 range and for large sample interval 5–1000. The rate of convergence of the method of conjugate gradients, in contrary to the findings of Fletcher and Reeves [6], is faster than the descent method suggested by Fletcher and Powell [7], as initial estimates to the shape parameter is correctly chosen. Otherwise the method of conjugate gradients is found to approach to the result with better convergence rate than the descent method. Also the former is more stable and fast for correcting the initial estimates for each new start to iteration cycles

REFERENCES

HARTER, H.L. and MOORE, A.H., 1965. "Maximum-likelihood estimation of the parameters of Gamma and Weibull population from complete and censored samples,, *Technometrics*, vol 7, pp 639–643.

WİNGO, D.R., 1972. "Maximum likelihood estimation of the parameters of the Weibull distribution by modified quasilinearization", *IEEE Trans. Reliability*, vol R-21, 1972 May, pp 89–93.

ZANAKİS, S.H., 1979. "A simulation study of some simple estimators for the three-parameter Weibull distribution", *J. Stat, Comp. and Simulation*, vol 9, pp 101–106.

ARCHER, N.P., 1980. "A computational technique for maximum likelihood estimation with Weibull models", *IEEE Trans, Reliability*, vol R-29, pp 57–62.

ROCKETTE, H., ANTLE, C., and KLİMKO, L.A., 1974. "Maximum likelihood estimation with the Weibull model", *J. Am. Statist. Assoc.*, vol 69, pp 246–249.

FLETCHER, R., and REEVES, C.M., 1964. "Function minimization by conjugate gradients", *The Computer Journal*, vol 7, pp 149–154.

FLECHER, R., and POWELL, M.J.D., 1963. "A rapidly convergent descent method for minimization", *The Computer Journal*, vol 6, pp 163–168.

COHEN, A.C., 1975 "Multi-Censored Sampling in the Three Parameter Weibull Distribution" *Technometrics*, vol 17, pp 347–351.

NİTTA, T., YAMADA, N., and FUJIWARA, Y., 1974. "Area effect of electrical breakdown in compressed SF_1", *IEEE Trans. Power Apparatus and Systems*, vol PAS-93, pp 623–629.