Regular Paper

Parameter Estimation in the Extreme-Value Distributions Using the Continuation Method

HIDEO HIROSE†

An efficient and stable maximum likelihood parameter estimation scheme is introduced for the three kinds of extreme-value distribution (Weibull, Gumbel, and Fréchet) using the generalized extreme-value distribution and the continuation method. As the proposed algorithm can almost always obtain the existing local maximum likelihood estimates automatically, it is of considerable practical value. This paper focuses on the Weibull distribution parameter estimation and shows that it is better to use the generalized extreme-value distribution than the Weibull distribution itself, and that the continuation method is more efficient than the grid search method in searching for parameters globally. The paper also shows that when there are no finite local maximum likelihood estimates in the Weibull distribution, it is probable that there are finite local maximum likelihood estimates in the Fréchet distribution, and vice versa. Only complete data sets are considered in this paper, but the algorithm can easily be applied to censored data.

1. Introduction

It is well-known that maximum likelihood parameter estimations in the three-parameter Weibull distribution (W3P) are extremely tedious, and many researchers have long been pursuing algorithms that find the estimates (see Panchang and Gupta[15] and Hirose[9]). The probability distribution function of the Weibull model is expressed by

$$F^1(x;\eta,\beta,\gamma)=1-\exp\left\{-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right\},$$
$$x\geq\gamma,\ \eta>0,\ \beta>0, \qquad (1)$$

where η, β, and γ are scale, shape, and location parameters, respectively. There are mainly two difficulties in the Weibull parameter estimation; one is the non-regularity and other is the parameter diverging problem.

In the non-regular condition, which is treated anomalously in the distribution, there are no maximum likelihood estimates (MLEs) when $0<\beta<1$; MLEs do exist, but are not asymptotically normal, when $1<\beta\leq 2$ (see Duby[5] or Smith[19]). In order to avoid these difficulties, especially for non-regular cases, a variety of methods have been proposed. Among them are modified moment estimation, proposed by Cohen and Whitten[3] and Cohen, Whitten, and Ding[4]; modified maximum likelihood estimation, proposed by Cohen and Whitten[3]; and closed form methods, proposed by Kappenmann[11]; Newby[14]; Wingo[20]; Wyckoff, Bain, and Engelhardt[21]; Zanakis[22]; and Zanakis and Mann[23]. However, the estimates obtained by these methods have some biases, as is shown by the results of the Monte Carlo simulation (Hirose[9]), when the shape parameter is large, and such cases are not rare. Therefore, obtaining the MLEs is much better than using other temporary expedient estimates.

In some cases the Weibull parameters diverge in the Newton-Raphson process, and the three diverging parameters in the Weibull distribution correspond to the parameters in the two-parameter Gumbel distribution (G2P). This is the second problem. In such a case, the three-parameter Fréchet distribution (F3P) is much more appropriate than the other two distributions, in a likelihood sense. This is also true for the Fréchet parameter diverging pattern. Thus, the GEV, which gives us a better understanding of the features of the sampled data, is recommended.

Here, a stable and efficient maximum likelihood estimation method in the W3P is proposed, using the generalized extreme-value distribution (GEV) and the continuation method (CM). This approach is quite new. The algorithm turns out to be simple, but the results are aston-

† Info. Math. Res. Lab., Takaoka Electric; Fac. Info. Sci., Hiroshima City University.

Table 1 Dielectric breakdown voltage data of epoxy resin.

Data Items	Case 1	Case 2	Case 3	Case 4	Case 5
1	24.54	27.15	27.66	27.98	28.04
2	28.00	29.13	26.54	27.49	28.57
3	25.69	28.28	26.96	27.85	26.33
4	27.72	27.74	26.15	27.93	29.61
5	28.05	28.87	25.26	24.19	28.17
6	27.53	26.42	29.44	25.01	27.14
7	27.34	24.46	28.32	27.06	29.17
8	26.80	30.88	27.66	27.62	25.44
9	26.51	29.11	28.21	28.94	28.49
10	27.28	27.31	27.80	29.09	27.46
11	28.16	27.54	27.59	27.63	27.31
12	28.86	27.98	26.63	28.28	26.95
13	26.67	28.49	28.08	27.63	27.88
14	28.37	26.25	28.83	28.20	27.30
15	28.37	28.50	27.96	27.95	29.02
16	28.44	25.61	28.13	28.33	29.52
17	28.05	29.50	29.06	27.11	26.89
18	24.61	28.04	26.78	26.47	27.89
19	27.54	27.94	28.00	28.17	28.08
20	26.85	26.66	26.28	27.35	27.75

ishing. For cases in the Monte Carlo simulation where the random data are drawn in the way shown by Hirose,[9] this algorithm obtained all the existing local MLEs. Another aspect of this method is that we can automatically see the properties of the estimates in all three of the extreme-value type distributions.

To obtain the MLEs, grid search methods have been used in searching for initial points of the Newton-Raphson iteration (see Panchang and Gupta[15]), but they are time-consuming and tedious. As an improvement on this method, it is shown in Section 3 that adopting the GEV is effective, but still imperfect. Combining the GEV with the CM turns out to be a perfect method for finding the existing local MLEs. The algorithm is given in Section 4.

Although the method focuses on the Weibull parameter estimation, the Fréchet and Gumbel parameters are obtained simultaneously. Only complete data cases are cosidered in this paper, but the proposed algorithm can be applied to censored data.

2 Another Issue in the Three-Parameter Weibull Distribution

Table 1 shows the dielectric breakdown voltages of one hundred epoxy resin test pieces that were tested in our experimental laboratory (Hirose[9]); it should be noted that some experiments (such as that for a generator) are very expensive, and therefore such a large number of test pieces is unusual. Twenty test pieces were tested by applying an ascending voltage to each in turn; this was repeated five times. Thus, we have five sets of data for the same underlying distribution. Since failure data usually fit the Weibull distribution, the MLEs of the Weibull parameters were obtained for the five data sets. The computing results obtained by the proposed method are shown in **Table 2**. We see that the first and fourth data sets have infinite estimates in the table, and this phenomenon becomes much clearer if we plot the likelihood function for case 4 (**Fig. 1**). This diverging property is discussed in the next section.

Table 2 Estimates of the parameters for various distributions.

Distribution	Estimates	Case 1	Case 2	Case 3	Case 4	Case 5
W3P	$\hat{\eta}$	∞	6.239	5.051	∞	4.725
	$\hat{\beta}$	∞	4.529	5.267	∞	4.811
	$\hat{\gamma}$	$-\infty$	22.092	22.921	$-\infty$	23.523
	$\log L$	−28.979	−35.375	−28.652	−27.613	−28.824
G2P	$\hat{\sigma}$	0.8473	1.3643	0.9183	0.7994	0.9391
	$\hat{\mu}$	27.790	28.491	28.066	27.997	28.356
	$\log L$	−28.979	−36.446	−29.165	−27.613	−29.503
F3P	$\hat{\eta}$	5.181	∞	∞	9.596	∞
	$\hat{\beta}$	6.553	∞	∞	12.329	∞
	$\hat{\gamma}$	33.037	$-\infty$	$-\infty$	37.628	$-\infty$
	$\log L$	−28.672	−36.446	−29.165	−27.444	−29.503
GEV	$\hat{\sigma}$	0.7906	1.3775	0.9589	0.7783	0.9821
	$\hat{\mu}$	27.567	28.331	27.972	28.032	28.248
	$\hat{k}$	−0.1526	0.2208	0.1898	−0.0811	0.2079
	$\log L$	−28.672	−35.375	−28.652	−27.444	−28.824

3. Reparameterization to the GEV Distribution

In order to figure out the diverging property, the Newton-Raphson iteration process was traced by using carefully selected initial values; the result is shown in **Table 3** for the data from case 4. In the table, we can see that both η/β and $\eta+\gamma$ converge, to 0.7994 and 27.997, respectively, and that the log-likelihood function converges to -27.613. This reminds us of the generalized extreme-value distribution (GEV) derived by von Mises[13]. Using the reparameterization.

$$\sigma=\eta/\beta,\quad \mu=\eta+\gamma,\quad k=1/\beta, \qquad (2)$$

the probability distribution function of the GEV

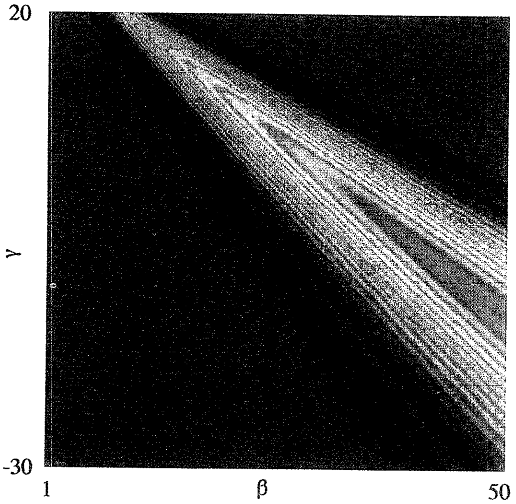


Fig. 1 Contour plot of the likelihood function for the Weibull distribution based on data case 4. The scale parameter is optimized.

Table 3 Newton-Raphson iterations based on the data from case 4.

Iteration Number	Log-Likelihood	$\hat{\eta}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\eta}/\hat{\beta}$	$\hat{\eta}+\hat{\gamma}$
0	−28.642	8.0	10.0	20.0	0.8000	28.0000
1	−28.259	10.1	12.4	17.9	0.8139	27.9687
2	−28.032	13.3	16.4	14.7	0.8133	27.9758
3	−27.886	18.1	22.3	9.9	0.8104	27.9817
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
18	−27.613	5844.4	7310.8	−5816.4	0.7994	27.9972
19	−27.613	8762.7	10961.5	−8734.7	0.7994	27.9972

is expressed by

$$F(x;\sigma,\mu,k)=1-\exp\left[-\left\{1+k\left(\frac{x-\mu}{\sigma}\right)\right\}^{1/k}\right],$$
$$\sigma>0,\quad 1+k\left(\frac{x-\mu}{\sigma}\right)>0, \qquad (3)$$

when $k\neq 0$. The Weibull distribution (1) corresponds to the case $k>0$. The cases $k\to 0$ and $k<0$ correspond to Gumbel (4) and Fréchet (5) distributions, respectively.

$$F^2(x;\sigma,\mu)=1-\exp\left\{-\exp\left(\frac{x-\mu}{\sigma}\right)\right\},$$
$$\sigma>0. \qquad (4)$$
$$F^3(x;\eta,\beta,\gamma)=1-\exp\left\{-\left(\frac{\gamma-x}{\eta}\right)^{-\beta}\right\},$$
$$x\leq\gamma,\quad \eta>0,\quad \beta>0. \qquad (5)$$

Distributions (1), (4), and (5), which are called extreme-value distributions for the minimum value, were first discovered by Fisher and Tippett[6], and the extreme types theorem was later completely proved by Gnedenko[8] (see Castillo[2] and Galambos[7]). Historically, the GEV has been mainly treated as a distribution representing the extreme values for maxima (see Jenkinson[10]) ; here, however, the extreme-value distributions for minima are discussed.

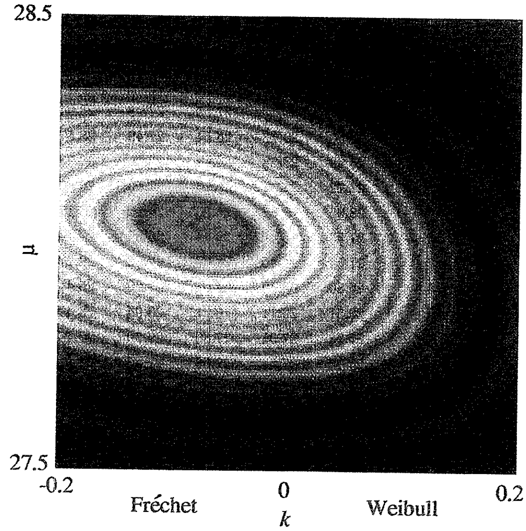


Fig. 2 Contour plot of the likelihood function for the GEV distribution based on data case 4. The scale parameter is optimized.

It is interesting to imagine what will happen if we fit the F3P (or GEV) to data case 4. Surprisingly, we can obtain the finite MLEs in the F3P and the corresponding value of the log-likelihood ($=-27.444$) is greater than that ($=-27.613$) obtained by the limiting form of the Weibull or Gumbel distribution. The MLEs of the Fréchet parameters are $\hat{\eta}=9.596$, $\hat{\beta}=12.329$, and $\hat{\gamma}=37.628$ (for the GEV, $\hat{\sigma}=0.7783$, $\hat{\mu}=28.032$, and $\hat{k}=-0.0811$). This suggests that it is better to adopt the F3P as the failure probability function than the W3P in a pure likelihood sense in such cases, and that it will be more advantageous to use the extreme-value distributions together. This means that we should use the GEV as a representative distribution for all three distributions.

It is intriguing that in **Fig. 2**, a contour plot similar to that in Fig. 1, showing the likelihood function of the GEV, seems to be smooth around the local maximum point and that the value of the location parameter μ does not change much even if the shape parameter k varies greatly on a equi-contour line. The latter property is very different from the Weibull parameter pattern (see Fig. 1). The location parameter γ varies a lot as the value of the scale parameter β changes. This suggests that by adopting the GEV a search for the local maximum of the likelihood function in the Weibull or Fréchet distribution would be easier, because of the more stable values.

4. Parameter Estimation in the GEV Distribution Using the CM

4.1 Using the G2P's MLEs

As mentioned above, the location parameter μ does not change the value of the log-likelihood function in the vicinity of the local maximum point in the GEV as much as the shape parameter k does. This also holds for the scale parameter σ. Thus, these two parameters might be useful for initial guesses in the iterative process of searching for the MLEs of the parameters in the GEV, provided we know the value of k. For example, the following moment estimators might be used:

$$\sigma = k\cdot s\{\Gamma(1+2k)-\Gamma^2(1+k)\}^{-1/2},$$
$$\mu = m+\frac{\sigma}{k}\{1-\Gamma(1+k)\}, \qquad (6)$$

where m and s denote the mean and standard deviation of the samples, but unfortunately we do not know k. Only in the case where $k\to 0$ can we obtain the MLEs of the parameters in the G2P, in which case equations (6) become

$$\sigma = \frac{\sqrt{6}}{\pi}s = 0.779697s,$$
$$\mu = m+0.577216\sigma, \qquad (7)$$

Using these initial values in the Newton-Raphson process, we can obtain the MLEs, $\hat{\sigma}_{G2P}$ and $\hat{\mu}_{G2P}$, easily.

The MLEs of the parameters in the G2P may be useful as initial guesses for estimating the MLEs of the three parameters in the GEV, but this is not a very powerful technique. An example will be given in Section 6. Rather, these estimates will be useful as a starting point for the continuation method.

4.2 Continuation Method

Continuation methods are known to be useful tools in solving a system of m nonlinear equations $f(\theta)=0$ (see Kincade and Cheney[12] and Allgower and Georg[1]). The Newton-Raphson algorithm defined by an iteration formula such as

$$\theta^{i+1} = \theta^i - (J^i)^{-1} f(\theta^i), \quad i=0, 1, \cdots \qquad (8)$$

will often fail because poor starting values are likely to be chosen, where J^i denotes a Jacobian of $f(\theta^i)$. Let us define a homotopy deformation $h: [0, 1]\times\Re^m\to\Re^m$ such that

$$h(0, \theta)=g(\theta), \quad h(1, \theta)=f(\theta), \qquad (9)$$

where $g: \Re^m\to\Re^m$ is a trivial smooth map having known zero points, and h is also smooth. For example, h is defined as follows:

$$h(t, \theta) = tf(\theta)+(1-t)\{f(\theta)-f(\theta^{(0)})\}, \qquad (10)$$

where $\theta^{(0)}$ is a solution when $t=0$; thus we can expect a zero point at $t=1$. We pursue a smooth curve

$$c(s)=(t(s), \theta(s))\in h^{-1}(0), \qquad (11)$$

with starting point $c(0)=(0, \theta(0))$ for a given critical point $\theta(0)$ of g, and starting tangent $\dot{c}(0)=(\dot{t}(0), \dot{\theta}(0))$; s corresponds to the arc-length on the curve c, and differentiation is done with respect to s. By differentiating $h=0$ with respect to s, we obtain a system of equations

$$h'(c(s))\,\dot{c}(s)=0, \qquad (12)$$

which is expressed more concretely as:

$$\frac{\partial f_1}{\partial t}\frac{dt}{ds}+\frac{\partial f_1}{\partial\theta_1}\frac{d\theta_1}{ds}+\frac{\partial f_1}{\partial\theta_2}\frac{d\theta_2}{ds}+\cdots+\frac{\partial f_1}{\partial\theta_m}\frac{d\theta_m}{ds}=0,$$
$$\frac{\partial f_2}{\partial t}\frac{dt}{ds}+\frac{\partial f_2}{\partial\theta_1}\frac{d\theta_1}{ds}+\frac{\partial f_2}{\partial\theta_2}\frac{d\theta_2}{ds}+\cdots+\frac{\partial f_2}{\partial\theta_m}\frac{d\theta_m}{ds}=0,$$
$$\vdots$$
$$\frac{\partial f_m}{\partial t}\frac{dt}{ds}+\frac{\partial f_m}{\partial\theta_1}\frac{d\theta_1}{ds}+\frac{\partial f_m}{\partial\theta_2}\frac{d\theta_2}{ds}+\cdots+\frac{\partial f_m}{\partial\theta_m}\frac{d\theta_m}{ds}=0. \qquad (13)$$

In (13), $d\theta_i/ds$ can be obtained by setting $\|\dot{c}\|=1$ and $\det\begin{pmatrix}h'(c(s))\\ \dot{c}(s)^T\end{pmatrix}>0$ (or <0), but it can also be obtained by setting $dt/ds=\delta$ (where δ is a small number); δ plays an important role, as will be shown later. Then, we continue to trace the curve $c(s)$ until $t(s)\geq 1$.

Obtaining a series of $\theta^{(j)}$ by marching, using differential equations (13), is very similar to obtaining a series of θ^i by using the Newton-Raphson algorithm, but equation (8) is totally different from equation (14) of the continuation algorithm in which a wide convergence region can be expected (see Section 5).

$$\theta^{(j+1)}=\theta^{(j)}-\delta\cdot(J^{(j)})^{-1}f(\theta^{(0)}),\quad j=0,1,\cdots, \tag{14}$$

Thus, adopting the continuation method described in this paper is quite advantageous, because minor changes to the computer codes of the Newton-Raphson algorithm yield major improvements in convergence.

4.3 Solving the Likelihood Equations

In order to solve a system of likelihood equations for the GEV, a slight modification of the above method is applied. The form of equation (10) is

$$(h(t,\sigma,\mu,k))=\begin{pmatrix}\partial\log L/\partial\sigma\\ \partial\log L/\partial\mu\\ \partial\log L/\partial k\end{pmatrix}+(t-1)\begin{pmatrix}\partial\log L/\partial\sigma\\ \partial\log L/\partial\mu\\ \partial\log L/\partial k\end{pmatrix}_{(\sigma^{(0)},\mu^{(0)},k^{(0)})}, \tag{15}$$

and equation (13) at step j becomes

$$\begin{pmatrix}\frac{\partial^2\log L}{\partial\sigma^2} & \frac{\partial^2\log L}{\partial\mu\partial\sigma} & \frac{\partial^2\log L}{\partial k\partial\sigma}\\ \frac{\partial^2\log L}{\partial\sigma\partial\mu} & \frac{\partial^2\log L}{\partial\mu^2} & \frac{\partial^2\log L}{\partial k\partial\mu}\\ \frac{\partial^2\log L}{\partial\sigma\partial k} & \frac{\partial^2\log L}{\partial\mu\partial k} & \frac{\partial^2\log L}{\partial k^2}\end{pmatrix}_{(\sigma^{(j)},\mu^{(j)},k^{(j)})}\cdot\begin{pmatrix}d\sigma(s)/ds\\ d\mu(s)/ds\\ dk(s)/ds\end{pmatrix}=-\delta\begin{pmatrix}\partial\log L/\partial\sigma\\ \partial\log L/\partial\mu\\ \partial\log L/\partial k\end{pmatrix}_{(\sigma^{(0)},\mu^{(0)},k^{(0)})}, \tag{16}$$

where the likelihood function is $L=\prod_{i=1}^{n}\frac{d}{dx_i}F(x_i)$ (n: sample size) and the likelihood equations are

$$\begin{aligned}&\partial\log L/\partial\sigma=0,\\ &\partial\log L/\partial\mu=0,\\ &\partial\log L/\partial k=0.\end{aligned} \tag{17}$$

Thus, $c(s)$ can be traced by solving (16) and updating the parameters ($\theta^{(j+1)}=\theta^{(j)}+\text{sign}(\delta)\cdot(d\theta/ds)^{(j)}$) successively until $t(s)\geq 1$. The sign of δ at step j should be chosen so that $\log L^{(j-1)}\leq\log L^{(j)}$; this is a new treatment for solving likelihood equations, and is not seen in the ordinary continuation method.

If the sign of the Hessian, $D=\det(\partial^2\log L/\partial\theta_a\partial\theta_b)$, induced from the log-likelihood function does not change from the starting point $c(0)$ until the final point $c(s_{\text{final}})$, h is called *regular*. For example, h is regular and the sign of D is always positive when we consider the two-parameter Weibull distribution. However, h is not always regular in the case of the GEV, and the determinant sometimes changes its sign. Thus, certain treatments are necessary under such conditions. When $D>0$, $c(s)$ is likely to climb down the likelihood function; so we are forced to change the direction of $c(s)$ by setting the sign of δ to be negative until D becomes negative. When $D^{(l)}>0$ and $D^{(l+1)}<0$, the sign of δ should be switched to the reverse side, and the starting point $\theta^{(0)}$ should be replaced with the point $\theta^{(l+1)}$ as a new starting point. Then, $c(s)$ can continue to climb up the likelihood function. This is the point of the algorithm.

4.3 Selecting the Starting Point

Where the starting point is selected in the continuation method is important. The use of the MLEs in the G2P as the starting point seems to be natural, since the G2P is the embedded model of the GEV, and the estimation of the MLEs in the G2P is rather easy. However, they cannot be used directly as a starting point for ($\sigma^{(0)}$, $\mu^{(0)}$, $k^{(0)}$), because the Hessian is not defined at $k=0$. Instead, we find the MLEs of the GEV in the very vicinity of $k=0$, say $k=k^{(0)}=0.01$. The initial value of the GEV for $k^{(0)}$ in the Newton-Raphson iteration are $\hat{\sigma}_{\text{G2P}}$ and $\hat{\mu}_{\text{G2P}}$.

Therefore, the algorithm is as follows:

4.4 Algorithm

Step 1; Obtain the MLEs, $\hat{\sigma}_{\text{G2P}}$ and $\hat{\mu}_{\text{G2P}}$, in the G2P, using the initial guesses σ^0_{G2P} and μ^0_{G2P} given by (7). We alter the notation from $\hat{\sigma}_{\text{G2P}}$ and $\hat{\mu}_{\text{G2P}}$ to σ_0 and μ_0, respectively, for simplicity.

Step 2; Use σ_0 and μ_0 as initial values of the starting point, and alter the notation (σ_0, μ_0, ε) to ($\sigma^{(0)}$, $\mu^{(0)}$, $k^{(0)}$)$=\theta^{(0)}$, where ε is a small number, say 0.01.

Step 3; Solve equation (16). If $D^{(j)}<0$, then set δ to be positive; otherwise, set δ to be negative. Obtain the next point by calculating

$$\theta^{(j+1)}=\theta^{(j)}+\text{sign}(\delta)\cdot(d\theta/ds)^{(j)},$$
$$t^{(j+1)}=t^{(j)}+|\delta|.$$

If $D^{(j)}<0$, $D^{(j+1)}<0$ and $t^{(j+1)}<1$, then repeat this step. If $D^{(j)}<0$, $D^{(j+1)}<0$ and $t^{(j+1)}\geq 1$, then search for the MLEs of $\hat{\theta}$ in the GEV model by the three-parameter Newton-Raphson method, using the initial point $\theta^{(j+1)}$. If the parameters converge, then cease; other-

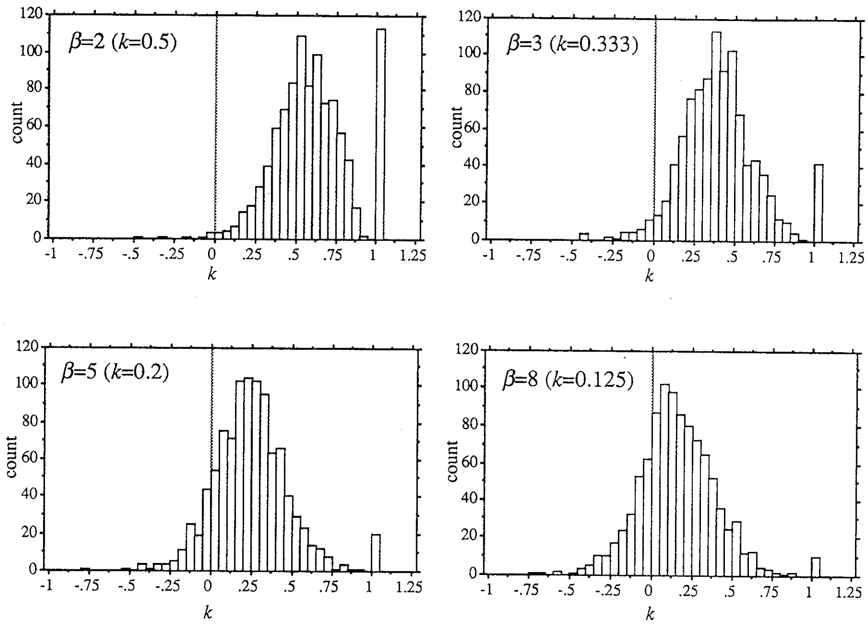


Fig. 3 Frequency of estimate k obtained by Monte Carlo simulation.
Number of simulation trials $=1000$, $n=20$

wise go to Step 4. If $D^{(j)}>0$ and $D^{(j+1)}>0$, then repeat this step. If $D^{(j)}>0$ and $D^{(j+1)}<0$, then go to Step 4.

Step 4 ; Reset $t^{(j+1)}\equiv 0$ and go to Step 3.

In these steps, the first and second derivatives of $\log L$ and Fisher's information matrix can be obtained as in the method described by Prescott and Walden[16),17)].

5. Computational Results

It was of no interest to recompute the MLEs of data sets from the literature, because we already knew the results. Instead, a Monte Carlo simulation was used, in which random data are drawn exactly as in Hirose[9)], where $n=10, 20, 50, 100$, $\beta=2, 3, 5, 8$ (Weibull), and the number of trials for each case is 1000. The MLEs could not be obtained for some data samples from the literature, even with better initial values. Failures in finding the MLEs seem to be caused by either non-regular conditions or diverging conditions. However, the computing results obtained by the proposed method are astonishing. For all the data sets tried, the local MLEs are always obtained if they exist, on condition that $k<1$. See **Fig. 3** for the frequency distributions of the estimates of the parameter k when $n=20$. This is confirmed by comparing the results obtained by the continuation method with those obtained by the grid search method. When $n=20$, for example, and $\beta=8$ ($k=0.125$), 22.1% of the estimates are regarded as being from the Fréchet distribution, even if the original random variables are taken from the Weibull distribution; when $\beta=2$ ($k=0.5$), 11.3% of the estimates are at the corner, $\hat{\beta}=1$.

6. Discussion

The proposed method is of great value from a practical point of view, because it allows the local MLEs to be obtained automatically from any data. Other optimization schemes have sometimes failed to consistently obtain the local MLEs, even if the computing time is faster. The other algorithms tried are as follows:

(1) Direct use of the three-parameter Newton-Raphson method from the initial point $(\sigma_1, \mu_1, k_1)_{k_1=0.01}$. When $|k|\leq 0.5$ and $n\geq 20$ the optimal value may be obtained, but the method often fails. When $|k|>0.5$ it usually fails to obtain the MLEs. For instance, when $n=20$ and

Table 4 Success rate of convergency using the direct method.

n	$\beta=2$	$\beta=3$	$\beta=5$	$\beta=8$
10	0.158	0.327	0.454	0.518
20	0.064	0.238	0.516	0.650
50	0.006	0.104	0.504	0.742
100	0.0	0.039	0.432	0.819

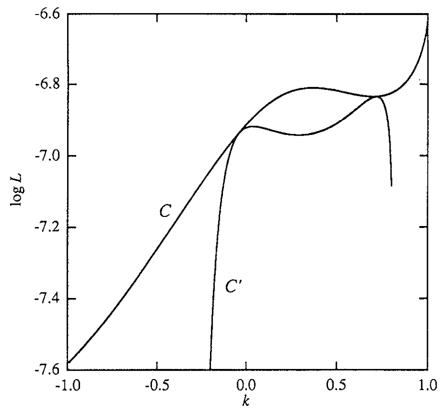


Fig. 4 Log-likelihood functions from curves C and C'.
C : $\log L|_{\sigma,\,\mu:\text{optimized}}$
C' : $\log L|_{\sigma,\,\mu:\text{from equations(0)}}$

$\beta=2$ ($k=0.5$), the success rate was only 6.4% (see **Table 4**).

(2) Use of a curve C' obtained by using equations (6) for a given k. C' consists of a set $\{(\sigma, \mu, k)\in\Re^3, \sigma>0, 1+k\{(x-\mu)/\sigma)\}>0 \mid$ equations (6) hold$\}$. However, it is not a good approximation of the curve C, where C consists of a set $\{(\sigma, \mu, k)\in\Re^3, \sigma>0, 1+k\{(x-\mu)/\sigma)\}>0 \mid \partial \log L/\partial\sigma=0, \partial\log L/\partial\mu=0\}$. For instance, **Fig. 4** shows the difference in the log-likelihood functions for the curve C and the curve C', where samples are taken from Rockette et al.[18]

(3) Use of a grid search method parameterized by k. First, we construct a curve C using the curve C' as initial points for the two-parameter Newton-Raphson method. After finding the grid point on the curve C at which the log-likelihood is maximum, we search for the optimal point by using the three-parameter Newton-Raphson method. However, this is time-consuming and tedious. For example, when $n=100$ and $\beta=3$, the computing time required by the grid search method is more than 2.5 times longer than that required by the continuation method, where the number of grids is 201 in $-1\leq k\leq 1$.

7. Concluding Remarks

A stable and efficient parameter estimation method using the generalized extreme-value distribution and the continuation method has been explored. From the initial starting point given by Gumbel's maximum likelihood estimates, an approximated optimal point in the generalized extreme-value distribution is obtained by the continuation method, and far more accurate local maximum likelihood estimates are obtained by the three-parameter Newton-Raphson method. Using this newly developed scheme, we can automatically obtain the existing local maximum likelihood estimates in three kinds of extreme-value distribution without numerical failure. In addition, the most appropriate distribution in the sense of likelihood is determined automatically. Although the method explained is applied only to complete data, this algorithm can easily be extended to censored data.

References

1) Allgower, E. L. and Georg, K.: *Numerical Continuation Methods,* Springer-Verlag (1990).
2) Castillo, E.: *Extreme Value Theory in Engineering,* Academic Press, New York (1988).
3) Cohen, A. C. and Whitten, B. J.: Modified Maximum Likelihood and Modified Moment Estimators for the Three-Parameter Weibull Distribution, *Communications in Statistics-Theory and Methods,* Vol. 11, pp. 2631-2656 (1982).
4) Cohen A. C., Whitten, B. J. and Ding, Y.: Modified Moment Estimation for the Three-Parameter Weibull Distribution, *Journal of Quality Technology,* Vol. 16, pp. 159-167 (1984).
5) Dubey, S. D.: Hyper-efficient Estimator of the Location Parameter of the Weibull Laws, *Naval Research Logistics Quarterly,* Vol. 13, pp. 253-264 (1966).
6) Fisher, R. A. and Tippett, L. H. C.: Limiting Forms of the Frequency Distributions of the Largest or Smallest Member of a Sample, *Proceedings of the Cambridge Philosophical Soci-*

ety, Vol. 24, pp. 180-190 (1928).

7) Galambos, J.: *The Asymptotic Theory of Extreme Order Statistics,* Robert Krieger Publishing Company (1987).

8) Gnedenko, P. B.: Sur la Distribution Limite du Terme Maximum d'une Sĕrie Alĕatore, *Annals of Mathematics,* Vol. 44, pp. 423-453 (1943).

9) Hirose, H.: Percentile Point Estimation in the Three-Parameter Weibull Distribution by the Extended Maximum Likelihood Estimates, *Computational Statistics and Data Analysis,* Vol. 11, pp. 309-331 (1991).

10) Jenkinson, A. F.: The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements, *Quarterly Journal of the Royal Meteorological Society,* Vol. 81, pp. 158-171 (1955).

11) Kappenmann, R. F.: Estimation for the Three-Parameter Weibull, Lognormal and Gamma Distributions, *Computational Statistics and Data Analysis,* Vol. 3, pp. 11-23 (1985).

12) Kincade, D. and Cheney, W.: *Numerical Analysis,* Brooks/Cole Publishing Company (1991).

13) Mises, R. von: La Distribution de la Plus Grande de *n* Valuers, *Rev. Math. Union Interbalcanique,* Vol. 1, pp. 141-160 (1936).

14) Newby, M.: Properties of Moment Estimators for the 3-Parameter Weibull Distribution, *IEEE Transactions on Reliability,* Vol. 33, pp. 192-195 (1984).

15) Panchang, V. G. and Gupta, R. C.: On the Determination of the Three-Parameter Weibull MLE's, *Communications in Statistics—Simulation and Computation,* Vol. 18, pp. 1037-1057 (1989).

16) Prescott, P. and Walden, A. T.: Maximum Likelihood Estimation of the Parameters of the Generalized Extreme-Value Distribution, *Biometrika,* Vol. 67, pp. 723-724 (1980).

17) Prescott, P. and Walden, A. T.: Maximum Likelihood Estimation of the Parameters of the Three-Parameter Generalized Extreme-Value Distribution from Censored Samples, *Journal of Statistical Computation and Simulation,* Vol. 16, pp. 241-250 (1983).

18) Rockette, A., Antle, C. E. and Klimko, L. A.: Maximum Likelihood Estimation with the Weibull Model, *Journal of the American Statistical Association,* Vol. 69, pp. 246-249 (1974).

19) Smith, R. L.: Maximum Likelihood Estimation in a Class of Nonregular Cases, *Biometrika,* Vol. 72, pp. 67-90 (1985).

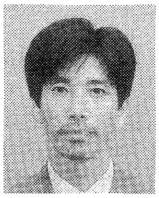
20) Wingo, D. R.: Solutions of the Three-Parameter Weibull Equations by Constrained Modified Quasilinearization, *IEEE Transactions on Reliability,* Vol. 22, pp. 96-102 (1973).

21) Wyckoff, J., Bain, L. J. and Engelhardt, M.: Some Complete and Censored Sampling Results for the Three-Parameter Weibull Distribution, *Journal of Statistical Computation and Simulation,* Vol. 11, pp. 139-151 (1980).

22) Zanakis, S. H.: A Simulation Study of Some Simple Estimators for the Three-Parameter Weibull Distribution, *Journal of Statistical Computation and Simulation,* Vol. 9, pp. 101-116 (1979).

23) Zanakis, S. H. and Mann, N. R.: A Good Simple Percentile Estimator of the Weibull Shape Parameter for Use When All Three Parameters Are Unknown, *Naval Research Logistics Quarterly,* Vol. 29, pp. 419-428 (1982).

(Received June 14, 1993)
(Accepted May 12, 1994)



Dr. Hideo Hirose Information Mathematics Research Lab, Takaoka Electric Mfg. Co., Ltd., Nishibiwajima, Aichi 452; Faculty of Information Sciences, Hiroshima City University, Hiroshima 731-31, JAPAN.

Hideo Hirose is Vice Director of Research at Takaoka Electric Mfg. Co., Ltd. He was born 1951 December 2. He studied electrical engineering at Miyakonojo Technical College from 1967 to 1972 and received a BS in Mathematics from Kyushu University (1977). Since then he has been working for Takaoka Electric Mfg. Co., Ltd. He received his Ph. D. from Nagoya University (1988) and received a prize from the Institute of Electrical Engineers of Japan for having written an excellent paper in 1987. He was a visiting scholar in statistics at Stanford University from 1989 to 1990. He has been a lecturer at Mie University since 1992, and at Hiroshima City University since 1994. He is interested in a variety of numerical computations such as finite element analysis, computational fluid dynamics, transient analysis of electrical networks, and electrical reliability. He is a member of IPS of Japan, IEEJ, IEEE, ASA, IMS, SIAM, AMS, and New York Academy of Sciences.