1. Introduction

In two-parameter reliability models such as the Weibull, Gumbel (extreme-value), log-normal and gamma distributions, the maximum likelihood parameter estimation is not a difficult task. Even one parameter addition to these models, however, makes the estimation difficult; thus, numerous researchers have studied appropriate estimation methods corresponding to each probability distribution model. The difficulties in three-parameter models such as the Weibull, log-normal and gamma distributions have gradually been overcome (Johnson, Kotz and Balakrishnan (1994)). However, four-parameter models such as the generalized gamma distribution still have difficulties in parameter estimation. This paper proposes a novel technique to solve the likelihood equations for such a distribution. The idea is simple: enlarge the parameter space from a simpler model to a more complex model in parameter estimation, and connect both models continuously in enlarged parameter space. This is called the model augmentation here. For example, consider the case of estimating the maximum likelihood estimates (MLE) of parameters in the two-parameter Weibull distribution. The one-parameter exponential distribution is a special case in the Weibull distribution (shape parameter is 1). The MLE of the exponential distribution is trivial. If the MLE of the exponential distribution and that of the Weibull distribution can be connected continuously in the Weibull parameter space, the MLE of the Weibull distribution can be obtained. This is the idea of the model augmentation. This model augmentation can be done by using the continuation method. Thus, the continuation method is first explained in this paper, and the model augmentation specific to the generalized gamma distribution is later described.

Although this paper applies the augmentation method to parameter estimation in the four-parameter generalized gamma distribution as a typical model augmentation method, this new procedure will be widely applicable because of its generality. For example, an appropriate finite mixture distribution model might be obtained by augmenting a simpler model to a more complex model successively.

This paper consists of eight sections including this introduction section. First, Section 2 introduces the general idea of the continuation method because this is central for the model augmentation. Section 3 proposes an extended model for the four-parameter generalized gamma distribution; the extension is intended to make the numerical estimation procedure stable. Then, Section 4 shows cases of model augmentation between a simpler model and a more complex model by using the continuation method; this paper specifically demonstrates the case for the four-parameter generalized gamma distribution. Section 5 describes boundary and corner solutions in the extended four-parameter generalized gamma distribution. Section 6 illustrates some typical examples in the distribution, and Section 7 is devoted to discussions.

2. Continuation method

Suppose that the parameter vector to be estimated is $\theta = (\theta_1, \theta_2, \cdots, \theta_m)$, where m is the number of unknown parameters. The Newton-Raphson method in solving the log-likelihood equations, $\nabla \log L(\theta) = 0$, often fails to find a solution, unless an appropriate starting point is carefully selected. It would be beneficial if some additional tools were available to find such a starting point for the Newton-Raphson method. The continuation method (Allgower and Georg (1990)) is one such tool for this purpose. Since applications of the continuation methods to statistical problems are not well known, a brief explanation seems useful.

2.1 Naive continuation method

Suppose that the functions $\nabla \log L(\theta)$ and $g(\theta)$ are smooth in $\mathbf{R}^m$, and g has a trivial solution such that $g(\theta_0) = 0$ for some θ_0. The principle idea of the continuation method is to first enlarge the parameter space from $\{\theta |\ \theta \in \mathbf{R}^m\}$ to $\{(t, \theta)|\ t \in \mathbf{R},\ \theta \in \mathbf{R}^m\}$ and make a smooth function $h(t, \theta(t))$ in this enlarged parameter space $\mathbf{R}^{m+1}$. Let's define $h(0, \theta(0)) = g(\theta(0)) \equiv g(\theta_0) = 0$ and $h(1, \theta(1)) = \nabla \log L(\theta(1)) = 0$. The set $C = \{(t, \theta(t))|\ h(t, \theta(t)) = 0\}$ parameterized by t becomes a smooth curve (path) in $\mathbf{R}^{m+1}$. Then, the maximum likelihood estimates can be obtained by pursuing the path from the trivial starting point $(0, \theta(0)) \in \mathbf{R}^{m+1}$ to a target solution point $(1, \theta(1)) \in \mathbf{R}^{m+1}$ such that $h(t, \theta(t)) = 0$ continuously.

There are many connection methods from the trivial starting function to the solution function, but the linear connection between the two functions is one of the most popular method. If $g(\theta)$ is defined by

$$g(\theta) = \nabla \log L(\theta) - \nabla \log L(\theta_0), \tag{1}$$

then $g(\theta)$ has a trivial solution in a sense that θ_0 can be selected arbitrarily as long as θ_0 is defined in the parameter space and is a regular point. If a smooth function $h(t, \theta)$ is defined by connecting the two functions, $g(\theta)$ and $\nabla \log L(\theta)$, linearly, then $h(t, \theta)$ becomes

$$\begin{aligned} h(t, \theta) =& t \cdot \nabla \log L(\theta) + (1 - t) \cdot g(\theta) \\ =& \nabla \log L(\theta) + (t - 1) \nabla \log L(\theta_0), \end{aligned} \tag{2}$$

where,

$$\begin{aligned} h(0, \theta(0)) =& \nabla \log L(\theta(0)) - \nabla \log L(\theta_0) = 0, \\ h(1, \theta(1)) =& \nabla \log L(\theta(1)). \end{aligned} \tag{3}$$

Thus, the target solution for $\nabla \log L(\theta) = 0$ can be obtained by tracing the points of $h^{-1}(0)$ from a starting point $(0, \theta(0))$ to a final point $(1, \theta(1))$; $\theta(1)$ can be the maximum likelihood estimates $\hat{\theta}$.

By differentiating $h = 0$ with respect to t, a differential equation,

$$\frac{d}{dt}\theta(t) = -[h_\theta(t, \theta(t))]^{-1} h_t(t, \theta(t)), \tag{4}$$

is obtained. By applying an Eulerian method to (4), a successive scheme

$$\theta^{(j+1)} = \theta^{(j)} - \delta(J^{(j)})^{-1} \nabla \log L(\theta_0), \quad j = 0, 1, ..., \tag{5}$$

will find a solution, where J denotes a Jacobian for $\nabla \log L(\theta)$, and δ is a small number. Since this iterative method is very similar to the usual Newton-Raphson iterative scheme,

$$\theta^{(i+1)} = \theta^{(i)} - (J^{(i)})^{-1} \nabla \log L(\theta^{(i)}), \quad i = 0, 1, ..., \tag{6}$$

we no longer have to develop a particular code if the Newton-Raphson scheme is already available (Hirose (1994a)). This is called the *naive continuation* method. Note that t should increase monotonically.

2.2 Introducing the arclength as a monotone increasing function

The naive continuation method will fail in solution finding at possible turning points of t; the turning point means a point where t cannot increase (this will be illustrated later, e.g., by Fig.2). To circumvent this inconvenience, arclength s on the curve C which consists of points of $h^{-1}(0)$ is introduced; s is monotone increasing. Parameter t and the curve C are parameterized by s, and (2) can be denoted as $h(C(s))$.

By differentiating $h = 0$ with respect to s,

$$h'(C(s)) \cdot \dot{C}(s) = 0 \tag{7}$$

is obtained, where $\dot{C}(s) = dC/ds$. To reduce one free parameter a constraint

$$\|\dot{C}(s)\| = 1, \tag{8}$$

should be imposed, where $\|\cdot\|$ denotes a Euclidian norm (l^2 norm). With an assumption that $\mathrm{rank}(h'(C(s))) = m$, an augmented Jacobian matrix,

$$A(s) = \begin{pmatrix} h'(C(s)) \\ \dot{C}(s)^T \end{pmatrix}, \tag{9}$$

becomes nonsingular, because $h'(C(s))$ is orthogonal to $\dot{C}(s)$ (see (7)). Thus, the direction of traversing the curve $C(s)$ at each iteration should be determined by a constraint,

$$\det(A(0)) \cdot \det(A(s)) > 0. \tag{10}$$

The starting direction of the curve $C(s)$ is defined such that $t(s) > 0$ in general cases, but it is convenient to determine the starting direction in such a manner that $\log L(\theta^{(1)})$ is greater than $\log L(\theta^{(0)})$ in the maximum likelihood parameter estimation procedure by experience.

2.3 Predictor-corrector continuation method

Using (8)-(10), increments $(dC(s))^{(j)}$ are obtained by solving a system of linear equations (7). Then, a new point $(\check{C}(s))^{(j)}$ is obtained by $((C(s))^{(j)} + (dC(s))^{(j)})$. However, this point is not necessarily on the curve $C(s)$. A correction process is needed for finding a point $(C(\tilde{s}^{(j)}))$ such that it is on the curve $C(s)$. For the correction to make the vector $((C(\tilde{s}))^{(j)} - (\check{C}(s))^{(j)})$ perpendicular to the vector $(dC(s))^{(j)}$ is used. The correction point $(C(\tilde{s}))^{(j)}$ is obtained by solving $h((C(\tilde{s}))^{(j)}) = 0$. This procedure which consists of these two steps is called the *predictor-corrector continuation* method. In contrast with the need of more than 100 steps in the naive continuation method, the predictor-corrector continuation method requires 1/10 times as many steps of the naive continuation for usual cases in reliability distributions .

3. Generalized gamma distribution

3.1 Historical background

The generalized gamma distribution, proposed by Stacy (1962),

$$\frac{p}{a}\frac{1}{\Gamma(b)}\left(\frac{x}{a}\right)^{bp-1}\exp\left\{-\left(\frac{x}{a}\right)^{p}\right\}, \quad (x \geq 0;\ a,\ b,\ p > 0), \tag{11}$$

is attractive because of its properties; it includes the exponential, Weibull, gamma, half-normal (Stacy and Mihram (1965)). Rayleigh and χ^2 distributions are also included because they are in a sub-family of the gamma distribution. Moreover, by using the logarithmic transformation to (11), the limiting form of the transformed distribution as $b \to \infty$

becomes a normal distribution (Bartlett and Kendall (1946)), and consequently (11) becomes a log-normal distribution as $b \to \infty$ (Prentice (1974) and Lawless (1980)). Lienhard and Meyer (1967) show a physical basis for the model, and Mees and Gerald (1984) apply the model to study the distributional shape of seed germination curves.

Maximum likelihood estimation for (11) is described in Parr and Webster (1965). Harger and Bain (1970) show that the three simultaneous likelihood equations can be reduced to a single non-linear equation in a single unknown p, but they mention that the Newton-Raphson method does not work well. Lawless (1980) also reports the difficulty in iterative numerical computation. Stacy (1973) reports the existence of multiple solutions for the non-linear equation, and Wingo (1987) consents to his claim by using the root isolation method of Jones, Waller and Feldman (1978). Cohen and Whitten (1988) use a trial-and-error procedure.

Harter (1967) proposes a four-parameter generalized gamma distribution which includes a location parameter c (this is also called a threshold parameter),

$$\frac{p}{a}\frac{1}{\Gamma(b)}\Big(\frac{x-c}{a}\Big)^{bp-1}\exp\Big\{-\Big(\frac{x-c}{a}\Big)^{p}\Big\}, \quad (x \geq c;\ a,\ b,\ p > 0), \tag{12}$$

and show an iterative procedure which consists of the rule of false position and the Newton-Raphson method for solving the four simultaneous likelihood equations. Along with the numerical difficulties in the three-parameter case of (11), inclusion of the location parameter will cause further difficulties in maximum likelihood parameter estimation. This can easily be understood by a similar interpretation to Smith (1985); the difficulties are due to non-regular problem. Johnson, Kotz and Balakrishnan (1994) suggest that this model is not recommended for analysis of sample data, unless the sample size is large enough to group the data in a frequency table. Because of these the literature on the four-parameter generalized gamma model seems to be limited. However, parameter estimation by using the continuation technique has been proved to be successful and it has brought us new features of this distribution as will be shown later; it would be useful and helpful if the estimation procedure by using the continuation technique becomes much easier than that by using a conventional method such as the Newton-Raphson method, when we refer to this four-parameter model as one of the distributions for reliability analysis.

3.2 Extended model

In some situation, enlarging the parameter space makes the numerical estimation more stable. For instance, estimation in the generalized extreme-value (GEV) distribution parameter space is more stable than that in the three-parameter Weibull (W3) space (Hirose (1994a)). This is because the GEV includes the W3; the embedding problem (Cheng and Iles (1990)) which occurs in the W3 vanishes in the GEV. This paper intends to make the parameter estimation in larger parameter space to the four-parameter generalized gamma (GGM4) distribution for stable computation.

Stacy and Mihram's model (1965) which allows us a negative value of p is one of the convenient extensions. A transformation of $b = 1/\lambda^2$ makes the treatment of the limiting form at $b \to \infty$ easy (Prentice (1974)). Besides, reparameterization of $p = 1/k$ seems convenient in making the estimation stable; this is similar to the GEV reparameterization of the three-parameter Weibull distribution. Therefore, an extended (and reparameterized) model of the GGM4 with the density function,

$$f(x; \lambda, k, \sigma, \mu) = \frac{|\lambda|}{\sigma} \frac{1}{\Gamma(1/\lambda^2)} \Big\{1 + \lambda k \Big(\frac{x-\mu}{\sigma}\Big)\Big\}^{1/(\lambda^2 k)-1} \exp\Big[-\Big\{1 + \lambda k \Big(\frac{x-\mu}{\sigma}\Big)\Big\}^{1/k}\Big],$$
$$(1 + \lambda k \Big(\frac{x-\mu}{\sigma}\Big) \geq 0;\ \lambda \neq 0,\ k \neq 0,\ \sigma > 0), \tag{13}$$

is proposed by the reparameterization,

$$a = \frac{\sigma}{|\lambda k|}, \quad c = \mu - \frac{\sigma}{|\lambda k|}, \quad p = \frac{1}{|k|}, \quad b = \frac{1}{\lambda^2}. \tag{14}$$

The model (13) is called the extended four-parameter generalized gamma distribution (EGGM4) here.

The function (13) becomes a GEV density function when $\lambda = 1$, and this restricted density becomes a Gumbel (GB) density function as $k \to 0$. This extended model (13) includes various probability distribution models as the model (12) includes a variety of types of distribution. However, the relationships among the three distributions, the extended four-parameter generalized gamma (EGGM4), the GEV, and the GB distributions, are enough to construct the model connection which will be used in Section 4.

The log-likelihood function for (13) is

$$\begin{aligned}\log L = {} & n\{\log|\lambda| - \log\sigma - \log\Gamma(1/\lambda^2)\} \\ & + \sum_{i=1}^{n} \Big[\Big(\frac{1}{\lambda^2 k} - 1\Big) \log\Big\{1 + \lambda k\Big(\frac{x_i - \mu}{\sigma}\Big)\Big\} - \Big\{1 + \lambda k\Big(\frac{x_i-\mu}{\sigma}\Big)\Big\}^{1/k}\Big],\end{aligned} \tag{15}$$

where n denotes the number of samples. When $\lambda^2 k > 1$, function (15) becomes infinity as $x \to \mu - \sigma/(\lambda k)$, thus a boundary of the maximum likelihood parameter space is $\lambda^2 k = 1$.

Another formulation for the EGGM4 is briefly discussed in Section 6.

4. Model augmentation

The GEV is a special case in the EGGM4 ($\lambda = 1$ in (13)). Thus, the solution of the likelihood equations in the EGGM4 can be traced from that in the GEV continuously by using the continuation method when the solution in the GEV has already been obtained in the interior of the EGGM4 parameter space. That is, the solutions in the GEV and EGGM4 models correspond to the trivial solution $\theta(0)$ and the target solution $\theta(1)$ respectively in Section 2. This method, on the other hand, can be interpreted as the model connection from the simpler (easier) model to the more complex (more difficult) model. In other words, the GEV enlarges the probability distribution model from three-dimensional space to four-dimensional space, and the solution of the likelihood equations in the EGGM4 can be obtained by enlarging the GEV model. If the solutions in the GEV and EGGM4 are in the interior of the parameter space and these solution points are regular, it is possible to find the solution in the EGGM4 by tracing the curve $C(s)$ defined by the continuation method from the GEV solution point to the EGGM4 solution point continuously as long as the curve $C(s)$ does not have bifurcation points.

The GB is not included in the GEV, but it is a limiting distribution of the GEV ($\lambda = 1,\ k \to 0$ in (13)). However, a small perturbation of k from 0 leaving other two parameters, σ and μ, fixed to the GB's maximum likelihood estimates, allows an *approximate* GB model exist in the GEV space. Then, the solution in the interior of the parameter space in the GEV can be traced from the solution in the approximate GB continuously when the solution in the GB has already been obtained.

As a result, the solution in the EGGM4 can be obtained by tracing the two curves from a point which is an approximate solution of the GB whenever the solution in the GB is obtained. The solution of the likelihood equations in the GB always exists and is unique; Pike (1966) has shown this for the two-parameter Weibull case, and by a logarithmic transformation the Weibull variates become the Gumbel variates. Thus we can always find a starting point in the EGGM4 parameter space as long as the random variable X is positive.

However, the GEV does not always have a solution in the interior of the parameter space; the maximum likelihood estimates may be located on the boundary of the parameter space, i.e., the location parameter c in (12) approaches the minimum order statistics in the sample and (12) becomes an exponential density function when $b = 1$. Similarly, the solution in the EGGM4 is not always located in the interior of the parameter space, it may be located on the boundary. The continuation method seems incompetent in such cases. However, it leads us to the boundary even in such situations. Therefore, the continuation method leads us to the interior, the boundary or the corner of the parameter space in the EGGM4 corresponding to the sampled data. This will be shown in Section 5.

There are many routes to find a solution in the EGGM4 from a trivial starting point to the final point. Fig.1 shows some typical routes. This paper selects a route from the GB to the EGGM4 via the GEV, because it seems the most stable way from a computational view-point by experience.

(INSERT FIG.1 ABOUT HERE.)

5. Boundary and corner solutions

5.1 Boundary solution

A locally maximum point of (12) is defined when $bp \geq 1$, and the boundary of the parameter space is $bp = 1$. At the boundary, the maximum likelihood of c tends to the smallest sample of the data, $x_{\min}$, because $\exp\{-(x - x_{\min})/a\} > \exp\{-(x - c)/a\}$ for $x \geq c$. Consequently, $\hat{\mu} \to x_{\min} + \hat{\lambda}\hat{\sigma}$ when k is eliminated in the reparameterized model. Then (13) becomes at the boundary

$$f(x; \lambda, \sigma) = \frac{|\lambda|}{\sigma} \frac{1}{\Gamma(1/\lambda^2)} \exp\Big\{-\Big(\frac{x - x_{\min}}{\lambda\sigma}\Big)^{\lambda^2}\Big\}, \quad (\lambda \neq 0,\ \sigma > 0). \tag{16}$$

5.2 Corner solution

If $\lambda \to \infty$ ($\sigma \to 0$) with $\lambda\sigma = \tau$ (τ:const.) in (16), (16) becomes a uniform distribution

$$f(x; \tau) = \frac{1}{\tau}, \quad (x_{\min} \leq x < x_{\max};\ \tau > 0), \tag{17}$$

where $x_{\max}$ is the largest sample of the data. This is easily seen by the fact that

$$\begin{aligned} \lambda^2/\Gamma(1/\lambda^2) &\to 1, \\ \exp\Big\{-\Big(\frac{x - x_{\min}}{\tau}\Big)^{\lambda^2}\Big\} &\to 1 \quad (0 \leq x - x_{\min} < \tau), \\ &\to 1/e \quad (x - x_{\min} = \tau), \\ &\to 0 \quad (\text{otherwise}), \end{aligned}$$

when $\lambda \to \infty$.

6. Examples

The generalized gamma distribution includes many types of the distribution according to the values of the parameters as mentioned in Stacy and Mihram (1965). In this section, however, only the typical degenerated cases which have not been introduced so far are mainly dealt with. That is, boundary solutions and a corner solution relevant to such degenerated distributions are treated.

6.1 Interior solution: example 1

Before degenerated cases are dealt with, a useful four-parameter case which has a solution in the interior of the parameter space is introduced first. Hirose and Lai (1997), treat a difficult Weibull analysis due to shape parameter divergence. That is, the log-likelihood function is maximized as the shape parameter tends to infinity. In the literature of Hirose and Lai (1997) the data are dealt with as grouped and are considered in the enlarged distribution model, the GEV, because of circumvention of non-regularity and the divergent problem. However, the problem for the positive endpoint still remains in the Weibull or GEV distributions unless some other techniques such as the Bayesian method (Smith and Naylor (1987)) is introduced. Such an annoying problem vanishes when the EGGM4 is applied to the same data (data #2) as in Kako (1986).

The parameter estimation method in the extreme-value distributions by using the maximum likelihood estimates in the GB as a starting point for the continuation method has already been illustrated by Hirose (1994a), thus how successfully the continuation method can search the solution of the log-likelihood equations in the EGGM4 is introduced here. The maximum likelihood estimates of the parameters in the GEV,

$$\begin{aligned} f(x;\sigma,\mu,k) =& \frac{1}{\sigma}\Big\{1+k\big(\frac{x-\mu}{\sigma}\big)\Big\}^{1/k-1}\exp\Big[-\Big\{1+k\big(\frac{x-\mu}{\sigma}\big)\Big\}^{1/k}\Big], \\ & (1+k\big(\frac{x-\mu}{\sigma}\big)\geq 0;\ k\neq 0,\ \sigma>0), \end{aligned} \tag{18}$$

are $\hat{\sigma} = 0.2730$, $\hat{\mu} = 3.364$, $\hat{k} = -0.1593$, and $\log L_{\max} = -7.4936$ when the sampled data are treated as continuous data. By using these values in addition to $\lambda = 1$ as a starting point in the EGGM4 parameter space, the predictor-corrector continuation method successfully traces the curve $C(s)$ to each optimum parameter as shown in Fig.2. In the figure, parameter s is implicitly hidden; one turning point is seen around $t = -4$, and this would not be traced by the naive continuation method. The maximum likelihood estimates are $\hat{\lambda} = 2.751$, $\hat{k} = 0.03747$, $\hat{\sigma} = 0.2246$, $\hat{\mu} = 3.661$; $\log L_{\max} = -6.8060$ which is slightly larger than that of the GEV. The Newton-Raphson method by using the same starting point as in the continuation method fails to search this local maximum point; instead, it finds a stationary but not a local maximum point, $\lambda = 0.8941$, $k = -0.2291$, $\sigma = 0.2940$, $\mu = 3.282$, where $\log L = -7.5003$.

The origin (endpoint) of the distribution is obtained by using the parameter transformation (14) as $\hat{c} = 1.482$, and it is intriguing that this value is strictly positive. The problem of negative endpoint vanishes accidentally here, although its confidence interval includes 0 (the standard error computed by using the delta method is 1.266). The shapes of the density functions of the EGGM4 and the GEV are shown in Fig.3.

(INSERT FIGS.2 AND 3 ABOUT HERE.)

6.2 Boundary solutions

The EGGM4 may not have a solution in the interior of the parameter space, but we can give a boundary optimum solution. Such a case in the W3 is shown in Rockette, Antle, and Klimko (1974).

(a) Semi half-normal: example 2

Using the data case 1 in Table 1 which are taken from the data case 4 in Kako (1986), the continuation method provides a curve approaching the boundary in the EGGM4 parameter space; this is shown in Fig.4 in which $\lambda^2 k \to 1$ and $c \to x_{\min}$ can be seen. By using a point, $\lambda = 1.5, \sigma = 0.65$, near the boundary, as a new starting point for the degenerated density (16), the continuation method can find a local maximum point on the boundary as shown in Fig.5. They are $\hat{\lambda} = 2.000, \hat{\sigma} = 0.6544$ (consequently, $\hat{k} = 0.2501$ and $\hat{\mu} = 3.909$), and $\log L_{\max} = -8.4147$. Since the shape of this density function is similar to the half-normal distribution of $(\lambda, k) = (\sqrt{2}, 1/2)$, the distribution of this example can be denoted as *semi half-normal*.

(INSERT TABLE 1 ABOUT HERE.)

(INSERT FIGS.4 AND 5 ABOUT HERE.)

(b) Reverse J-shaped: example 3

When the maximum likelihood parameter estimation procedure in the GEV locates an optimum point at the corner $k = 1$, the search in the EGGM4 should be begun at the boundary, $\lambda = k = 1$, $\mu = x_{\min} + \sigma$. The data case shown in Engelhardt and Bain (1979) corresponds to such a case. Along the boundary, the continuation method searches a solution successfully as $\hat{\lambda} = 0.8842$, $\hat{\sigma} = 49.94$ (consequently, $\hat{k} = 1.279$ and $\hat{\mu} = 196.9$), and $\log L_{\max} = -52.088$. Fig.6 shows a difference between the corner solution in the GEV which is a two-parameter exponential distribution and the boundary solution in the EGGM4; the solution in the EGGM4 can express a steeper tangent at the endpoint, which cannot be expressed in the GEV distribution.

(INSERT FIG.6 ABOUT HERE.)

6.3 Corner solution: example 4

There are some cases in which local maxima exist neither for the density (13) nor for the density (16). For the data case 2 in Table 1 which are taken from the data case 1 in Kako (1986), the continuation method in the EGGM4 parameter space gives an approximate boundary maximum point, $\lambda = 2.09, \sigma = 0.475$. Using this point as a starting point for the degenerated model (16), the continuation procedure traces a curve which approaches the boundary $t = 1$ but will never cross the boundary as shown in Fig.7 on the left; parameter λ seems to diverge. By using a reparameterization $\tau = \lambda\sigma$ and k (because $k \to 0$ as $\lambda \to \infty$), the continuation method reaches an optimum point, $\hat{\tau} = 1.2$, and the maximum log-likelihood value is -3.6464 (Fig.7 on the right). This degenerated model is a uniform distribution in (17) which has two endpoints, $x_{\min} = 2.6$ and $x_{\max} = 3.8$.

(INSERT FIG.7 ABOUT HERE.)

7. Discussions

7.1 Alternative formulation for EGGM4

In Section 3, the existence of an alternative extension for the EGGM4 model is mentioned. This formulation is expressed as:

$$f(x;\lambda,k,\sigma,\mu) = \frac{1}{\sigma|\lambda|}\frac{1}{\Gamma(1/\lambda^2)}\Big\{\frac{1}{\lambda^2}+\frac{k}{\lambda}\big(\frac{x-\mu}{\sigma}\big)\Big\}^{1/(\lambda^2 k)-1}\exp\Big[-\Big\{\frac{1}{\lambda^2}+\frac{k}{\lambda}\big(\frac{x-\mu}{\sigma}\big)\Big\}^{1/k}\Big],$$
$$(1+\lambda k\big(\frac{x-\mu}{\sigma}\big)\geq 0;\ \lambda\neq 0,\ k\neq 0,\ \sigma>0), \tag{19}$$

by the reparameterization,

$$a = \frac{|\lambda|\sigma}{|k|}, \quad c = \mu - \frac{\sigma}{|\lambda k|}, \quad p = \frac{1}{|k|}, \quad b = \frac{1}{\lambda^2}. \tag{20}$$

When $k = 1$, (19) becomes the extended three-parameter gamma distribution (EGM3) which is introduced by Hirose (1994b and 1998). Thus, an alternative route from the normal distribution to the EGGM4 via the EGM3 is possible (see Fig.1). This model also includes the GEV model ($\lambda = 1$). However, (19) is not recommended as an extended model for the GGM4 from a numerical stability viewpoint.

Table 2 shows some computational results obtained by the EGGM4 models using (13) and (19); the table also includes the results from the GEV and EGM3 for comparison. All the estimates are located in the interior of the parameter space. An indication that the locations of the estimates of $\hat{\sigma}$ and $\hat{\mu}$ in the EGGM4 using (19) are far from those obtained by the GEV and EGM3 models can be seen in some cases. On the other hand, these estimates in the EGGM4 using (13) do not seem to show this tendency very much. This suggests that parameter estimation using the model (13) is more stable than that using (19) as long as the starting point is set to the optimum values of the GEV or the EGM3. Practical experience supports this. Therefore, the model (13) is recommended for the extension of the GGM4.

(INSERT TABLE 2 ABOUT HERE.)

7.2 Multiple solutions

Stacy (1973) finds multiple solutions for the log-likelihood equations in the GGM3 to the data in Menon (1963). These solutions are, θ_1=(0.029, 0.001, −5.300), θ_2= (3.986, 0.004, 0.247), and θ_3=(0.029, 61.542, 7.604); $\theta=(a, b, p)$. Wingo (1987) gives another solution, θ_4=(3.7717, 0.0059, 0.2531), to the same data. The continuation method can lead us to the local maximum point $\hat{\theta}_5$=(3.9611, 0.0043021, 0.24734) using the optimum point (a, p)=(1.8435, 0.50573) in the W2 with $b = 1$. However, it does not direct us to other local maximum points. I have a strong belief that θ_2 and θ_4 should be identical to θ_5. The Newton-Raphson method shows that the point θ_6=(0.029436, 0.00098393, −5.3008) is a stationary point, but it is not a local maximum point. This indicates the difficulty of the numerical computation in parameter estimation even in the GGM3, as well as the existence of the multiple solutions for the likelihood equations in the EGGM4.

8. Concluding remarks

The numerical difficulties in maximum likelihood parameter estimation in the four-parameter generalized gamma distribution have hindered us from obtaining the estimates. Due to the nonlinearity of the likelihood equations we ought to use some iterative methods such as the Newton-Raphson method for finding a solution. In maximum likelihood estimation in reliability engineering, the exponential family, especially extreme-value types of the distribution, has been used, and it is known that the initial value setting is subtle. The continuation method has made us free from the initial value search as shown in this paper, typically in the four-parameter generalized gamma distribution.

The starting point for the continuation method can arbitrarily be selected if the point is a regular point in the interior of the parameter space. However, computational experience still suggests that we use an appropriately selected point because of stability in computing. The idea of the model augmentation from a simpler model to a more complex model seems to help this, and practically the model augmentation works well. This estimation method is so general, and it can be applied to many cases even if they are more complex.

Aimed at stable computation, the four-parameter generalized gamma distribution is reparameterized and extended as the Weibull model is extended to the generalized extreme-value distribution in this paper. This extended model includes the generalized extreme-value distribution which includes the Gumbel distribution as the limiting case. These relationships make us use the model connection from the Gumbel distribution to the four-parameter generalized gamma distribution.

In obtaining the maximum likelihood estimates of the parameters, the continuation method shows paths to optimum solutions. In a certain case, the path reaches the boundary of the parameter space, and the continuation method finds the optimum point along the boundary. It can find even the degenerated distribution functions including the uniform distribution. This characteristics cannot be realized by the Newton-Raphson method and related root finding methods.

References

Allgower, E.L. and Georg, K. (1990), *Numerical continuation methods*, Berlin: Springer-Verlag.

Bartlett, M.S. and Kendall, D.G. (1946), "The statistical analysis of variance-heterogeneity and the logarithmic transformation", *Journal of the Royal Statistical Society, Supplement*, **8**, 128-138.

Cheng, R.C.H. and Iles, T.C. (1990), "Embedded models in three-parameter distributions and their estimation", *Journal of the Royal Statistical Society, Series B*, **52**, 135-149.

Cohen, A. C. and Whitten, B. J. (1988), *Parameter Estimation in Reliability and Life Span Models*, New York: Marcel Dekker.

Engelhardt, M. and Bain, L.J. (1979), "Prediction limits and two sample problems with complete or censored Weibull data", *Technometrics*, **21**, 233-237.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distributions, vol.1*, New York: Wiley.

Jones, B., Waller, W.G and Feldman, A. (1978), "Root isolation using function values", *BIT*, **18**, 311-319.

Hager, H.W and Bain, L.J. (1970), "Inferential procedures for the generalized gamma distribution", *Journal of the American Statistical Society*, **65**, 1601-1609.

Harter, H.L. (1967), "Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population from complete and censored samples", *Technometrics*, **9**, 159-165.

Harter, H.L. and Moore, A.L. (1966), "Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples", *Journal of the American Statistical Association*, **61** 842-851.

Hirose, H. (1994a), "Parameter estimation in the extreme-value distributions using the continuation method", *Transactions of Information Processing Society of Japan*, **35**, 1674-1681.

Hirose, H. (1994b), "Maximum likelihood parameter estimation in the three-parameter gamma distribution", *Computational Statistics and Data Analysis*, **20**, 343-354.

Hirose, H. (1998), "Erratum to Maximum likelihood parameter estimation in the three-parameter gamma distribution", *Computational Statistics and Data Analysis*, **27**, 113.

Hirose, H. and Lai, T.L., (1997), "Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments", *Technometrics*, **39**, 199-210.

Kako, Y. (1986), "Estimation of Weibull parameters for BDV data of transformer oil by most likelihood methods", *Conference Internationale des Grands Reseaux Electriques*, Working Group **15-01**, Task Force 02.

Lambert, J.A. (1964), "Estimation of parameters in the three-parameter lognormal distribution", *Australian Journal of Statistics*, **6** 29-32.

Lawless, J.F. (1980), "Inference in the generalized gamma and log gamma distributions", *Technometrics*, **22**, 409-419.

Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: Wiley.

Lieblein, J., Zelen, M. (1956), "Statistical investigation of the fatigue lift of deep-groove ball bearings", *Journal of Research of the National Bureau of Standards*, **47** 273-316.

Lienhard, J.H. and Meyer, P.L. (1967), "A physical basis for the generalized gamma distribution", *Quarterly of Mathematics*, **25**, 330-334.

Mees, E. and Gerald, G. (1984), "La famille gamma generalisee: un modele de courbes de germination", *Biometrie-Praximetrie*, **24**, 101-115.

Menon, M.V. (1963), "Estimation of shape and scale parameters of the Weibull distribution", *Technometrics*, **5**, 175-182.

Parr, v.B., and Webster, J.T. (1965), "A method for discriminating between failure density functions used in reliability predictions", *Technometrics*, **7**, 1-10.

Pike, M. (1966), "A suggested method of analysis of a certain class of experiments in carcinogenesis", *Biometrics*, **22**, 142-161.

Prentice, R.L. (1974), "A log gamma model and its maximum likelihood estimation", *Biometrika*, **61**, 539-544.

Rockette, H., Antle, C. E., and Klimko, L. A. (1974), "Maximum likelihood estimation with the Weibull model", *Journal of the American Statistical Society*, **69**, 246-249.

Smith, R. L. (1985), "Maximum likelihood estimation in a class of nonregular cases", *Biometrika*, **72**, 67-90.

Smith, R. L. and Naylor, J. C. (1987), "A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution", *Biometrika*, **73**, 67-90.

Stacy, E.W. (1962), "A generalization of the gamma distribution", *Annals of Mathematical Statistics*, **33**, 1187-1192.

Stacy, E.W. and Mihram, G.A. (1965), "Parameter estimation for a generalized gamma distribution", *Technometrics*, **7**, 349-358.

Stacy, E.W. (1973), "Quasimaximum likelihood estimators for two-parameter gamma distributions", *IBM Journal Research & Development*, **17**, 115-124.

Wingo, D.R. (1987), "Computing maximum-likelihood parameter estimates of the generalized gamma distribution by numerical root isolation", *IEEE Transactions on Reliability*, **R-36**, 586-590.

Table 1. Data *1

case #	data					
1	3.0	4.2	3.8	3.8	3.3	2.7
	2.7	3.3	3.2	3.5	3.5	2.6
	3.1	3.4	2.8	2.9	3.1	3.1
	3.6	3.4				
2	3.4	3.8	3.1	2.6	2.9	3.0
	3.2	3.5	3.5	3.7	3.1	3.6
	3.5	2.8	2.9	3.1	2.9	2.9
	2.7	3.6				

*1: Kako (1986)

Table 2. Computational Results by the EGGM4, GEV, and EGM3

data source	model	estimates				
		$\hat{\lambda}$	$\hat{k}$	$\hat{\sigma}$	$\hat{\mu}$	$\log L$
Kako (1986)*2	EGG4$^{(3)}$	2.751	0.03747	0.2246	3.661	−6.80599
	EGG4$^{(19)}$	2.751	0.03747	0.02968	1.770	−6.80599
	GEV	1	−0.1593	0.2730	3.364	−7.49358
	EGM3	−0.7177	1	0.4172	3.160	−7.00027
Smith et. al. (1987)*3	EGG4$^{(3)}$	0.7457	0.01849	0.2755	1.399	−14.2394
	EGG4$^{(19)}$	0.7457	0.01849	0.4955	17.36	−14.2394
	GEV	1	0.08435	0.2729	1.642	−14.2853
	EGM3	−0.2781	1	0.3145	1.507	−14.8607
Smith et. al. (1987)*4	EGG4$^{(3)}$	1.432	0.08907	0.2255	1.394	−2.00359
	EGG4$^{(19)}$	1.432	0.08907	0.1100	0.4882	−2.00359
	GEV	1	0.04692	0.2215	1.249	−2.08236
	EGM3	−0.4158	1	0.2687	1.130	−2.04942
Harter et. al. (1966)*5	EGG4$^{(3)}$	1.666	0.2758	63.31	151.0	−205.833
	EGG4$^{(19)}$	1.666	0.2758	22.82	62.87	−205.833
	GEV	1	0.4547	46.28	104.3	−206.258
	EGM3	0.2961	1	43.81	92.68	−206.759
Harter et. al. (1966)*6	EGG4$^{(3)}$	1.359	0.3235	44.43	138.0	−196.390
	EGG4$^{(19)}$	1.359	0.3235	24.05	91.68	−196.390
	GEV	1	0.4292	35.84	114.3	−196.518
	EGM3	0.2650	1	34.03	104.8	−196.889
Lambert (1964)	EGG4$^{(3)}$	0.8781	0.6360	157.4	398.4	−564.072
	EGG4$^{(19)}$	0.8781	0.6360	204.1	476.0	−564.072
	GEV	1	0.6088	192.3	459.2	−564.101
	EGM3	0.5609	1	177.5	426.2	−564.167
Lieblein et.al. (1956)	EGG4$^{(3)}$	1.191	0.5300	49.89	95.38	−112.841
	EGG4$^{(19)}$	1.191	0.5300	35.18	72.07	−112.841
	GEV	1	0.6272	40.07	78.76	−112.850
	EGM3	0.5913	1	37.24	72.22	−112.914

*2: Data Case #2 in Kako (1986)
*3: Data Case #1 in Smith and Naylor (1987)
*4: Data Case #2 in Smith and Naylor (1987)
*5: Data Case #1 in Harter and Moore (1966)
*6: Data Case #2 in Harter and Moore (1966)

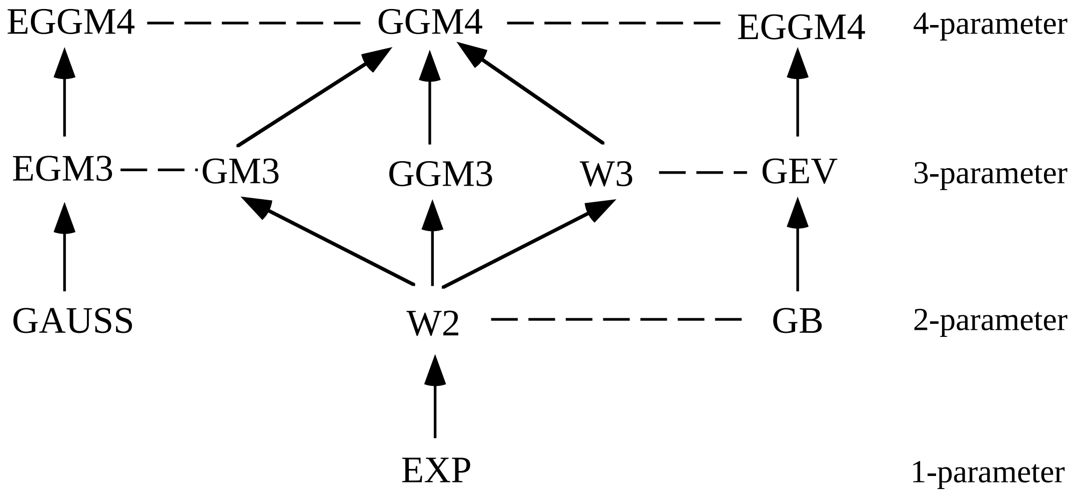

Fig.1 Various Routes to the Generalized Gamma Distribution

EGGM4: 4-parameter extended generalized gamma
GGM4: 4-parameter generalized gamma
GEV: generalized extreme-value
EGM3: 3-parameter extended gamma
GGM3: 3-parameter generalized gamma
W3: 3-parameter Weibull
GM3: 3-parameter gamma
W2: 2-parameter Weibull
GAUSS: normal
GB: 2-parameter Gumbel
EXP: 1-parameter exponential

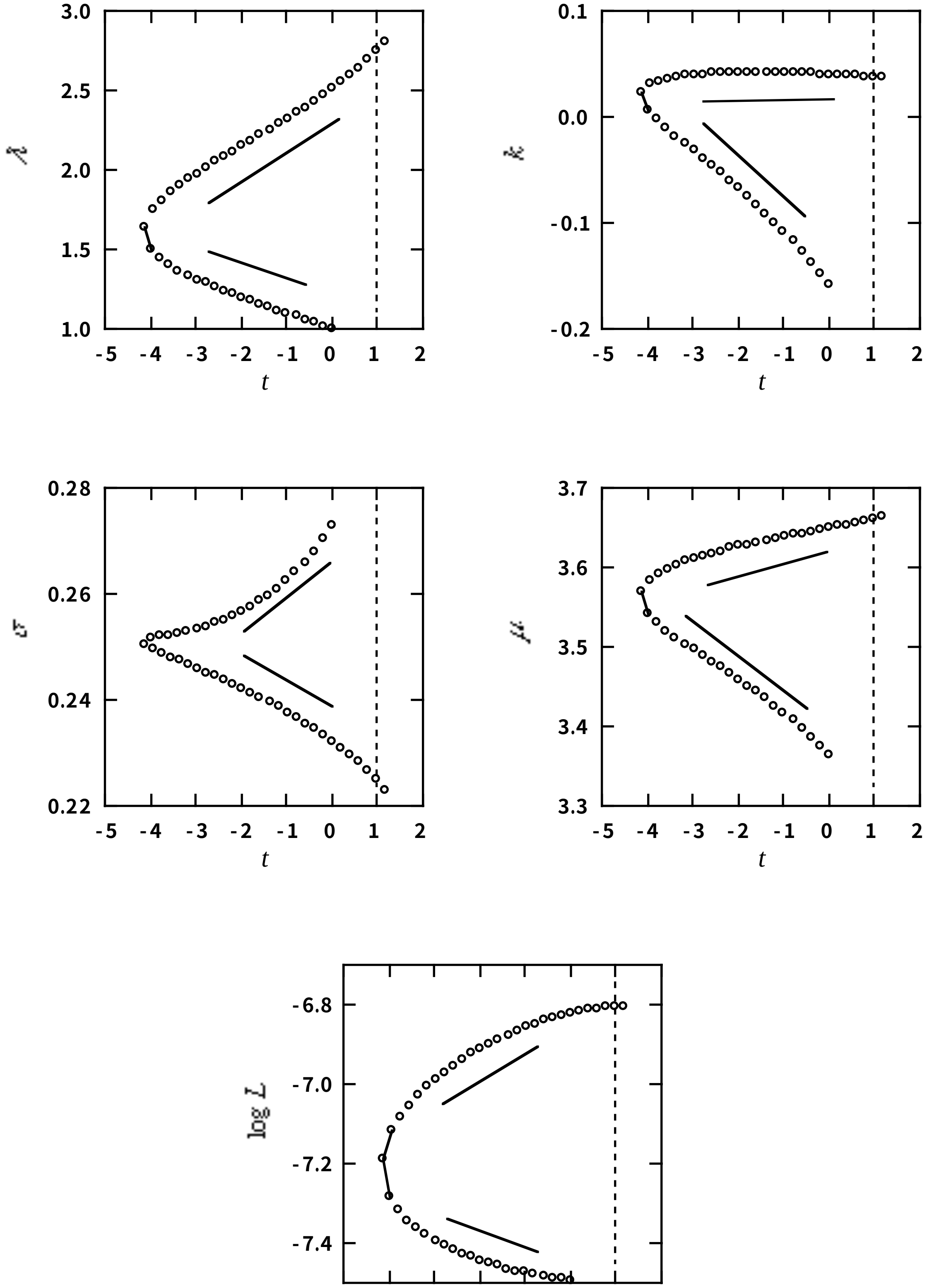

Fig. 2 The Continuation Traces for Example 1

Data Case #2 in Kako (1986)

A turning point is seen around t=-4

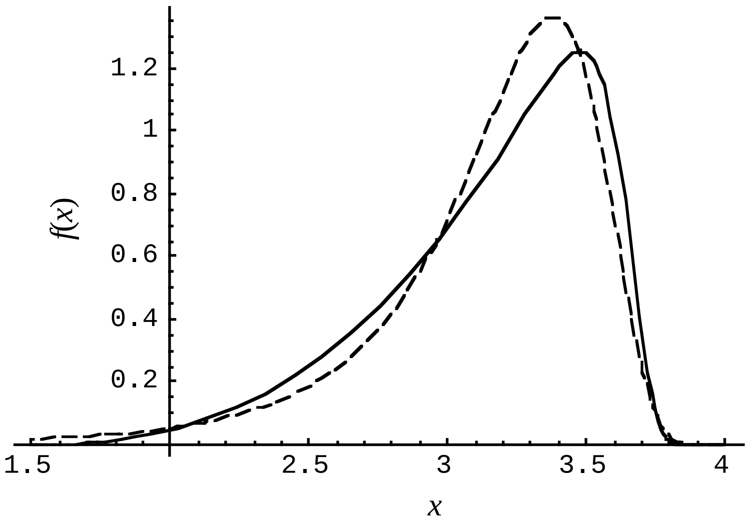

Fig. 3 The Density Functions for Example 1

Solid line: The EGGD4
Dashed line: The GEV
Data case #2 in Kako (1986)

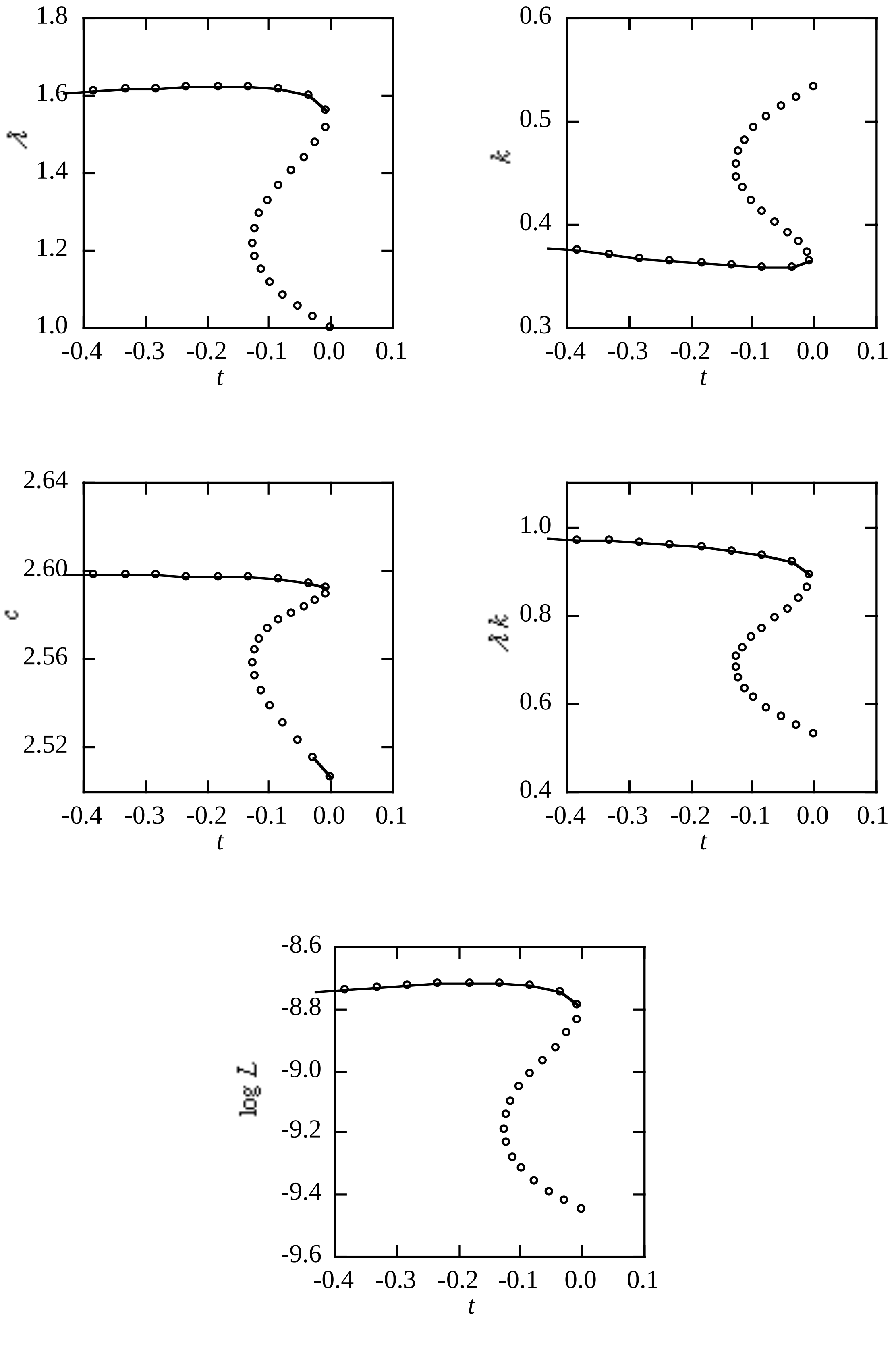

Fig. 4 The Continuation Traces for Example 2

Data Case #4 in Kako (1986)

Starting at the GEV Optimum Point

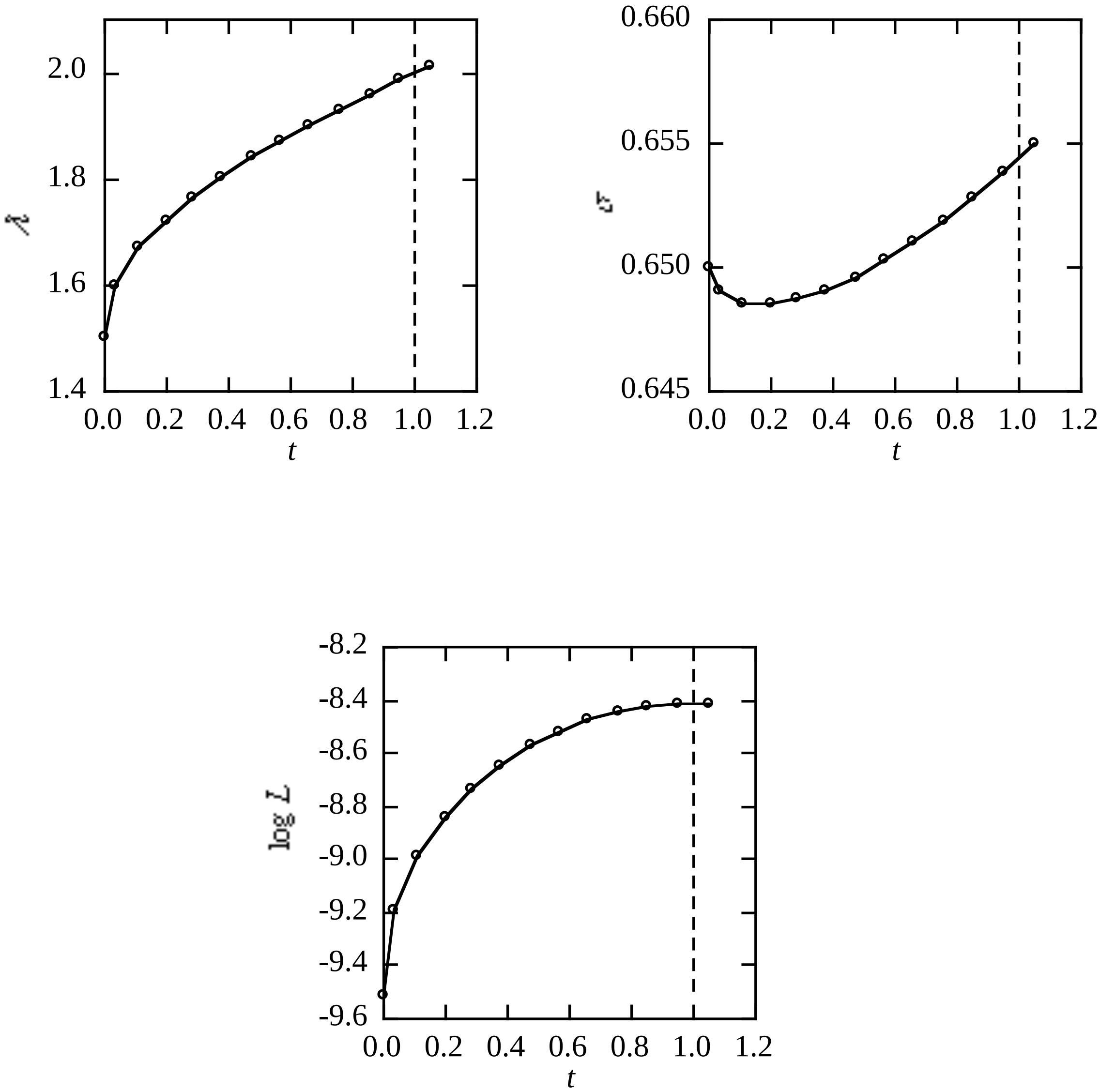

Fig. 5 The Continuation Traces for Example 2

Data Case #4 in Kako (1986)

Tracing Along the Boudadry l^2k=1

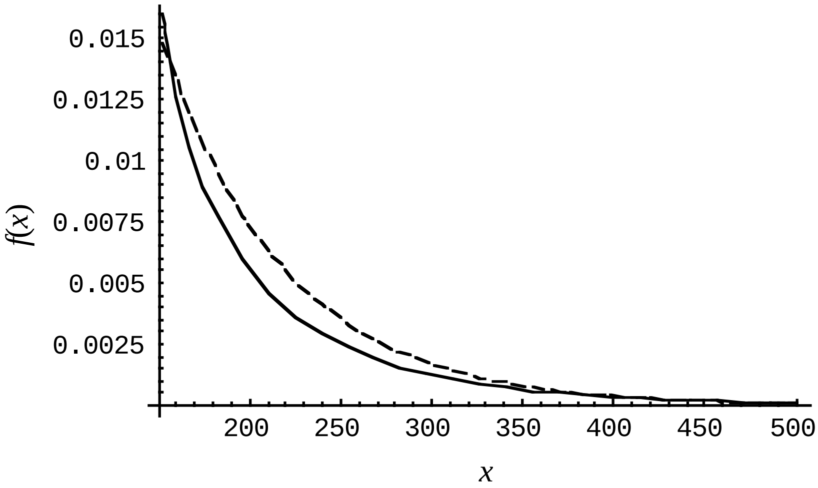

Fig. 6 The Density Functions for Example 3

Solid line: The EGGD4
Dashed line: The GEV Corner Solution
Data Case in Engelhardt and Bain (1979)

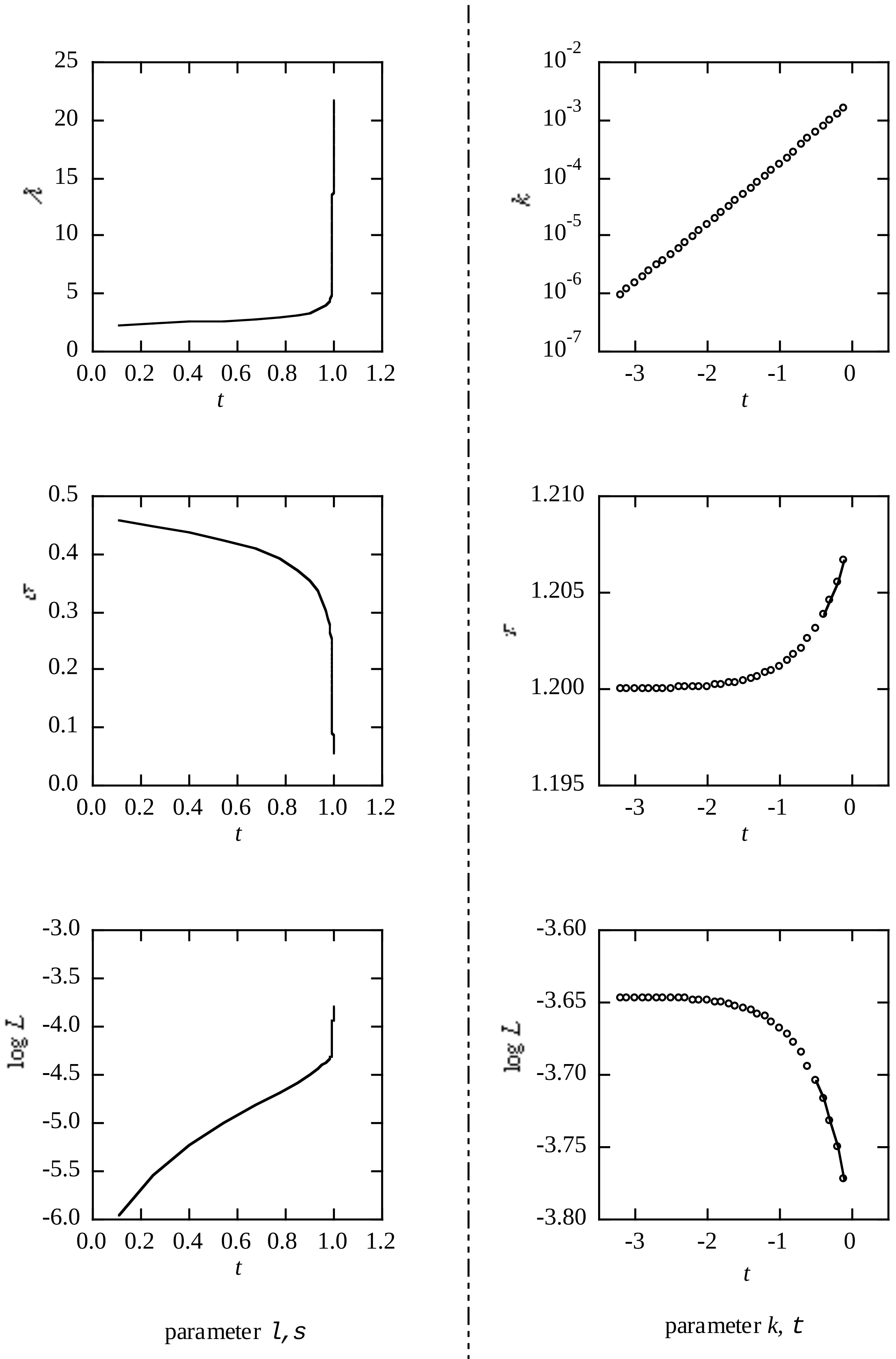

Fig. 7 The Continuation Traces for Example 4

Data Case 1 in Kako (1986)

Tracing Along the Boundary ¿ k=1