Math. Commun., Vol. **15**, No. 2, pp. 539-553 (2010)

Least squares fitting the three-parameter inverse Weibull density*

Miljenko Marušić[1,†], Darija Marković[2] and Dragan Jukić[2]

[1] *Department of Mathematics, University of Zagreb, Bijenička 30, HR-10 000 Zagreb, Croatia*
[2] *Department of Mathematics, University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia*

Received September 12, 2009; accepted May 20, 2010

Abstract. The inverse Weibull model was developed by Erto [10]. In practice, the unknown parameters of the appropriate inverse Weibull density are not known and must be estimated from a random sample. Estimation of its parameters has been approached in the literature by various techniques, because a standard maximum likelihood estimate does not exist.
To estimate the unknown parameters of the three-parameter inverse Weibull density we will use a combination of nonparametric and parametric methods. The idea consists of using two steps: in the first step we calculate an initial nonparametric density estimate which needs to be as good as possible, and in the second step we apply the nonlinear least squares method to estimate the unknown parameters. As a main result, a theorem on the existence of the least squares estimate is obtained, as well as its generalization in the l_p norm ($1 \leq p < \infty$). Some simulations are given to show that our approach is satisfactory if the initial density is of good enough quality.
AMS subject classifications: 65D10, 62J02, 62G07, 62N05
Key words: three-parameter inverse Weibull density, least squares, least squares estimate, existence problem, data fitting

1. Introduction

The probability density function of the random variable T having a three-parameter inverse Weibull distribution (IWD) with location parameter $\alpha \geq 0$, scale parameter $\eta > 0$ and shape parameter $\beta > 0$ is given by

$$f(t;\alpha,\beta,\eta) = \begin{cases} \frac{\beta}{\eta}\left(\frac{\eta}{t-\alpha}\right)^{\beta+1} \mathrm{e}^{-\left(\frac{\eta}{t-\alpha}\right)^{\beta}} & t > \alpha \\ 0, & t \leq \alpha. \end{cases} \tag{1}$$

If $\alpha = 0$, the resulting distribution is called the two-parameter inverse Weibull distribution. The inverse Weibull model was developed by Erto [10].

*Presented at the Sixth Conference on Applied Mathematics and Scientific Computing, Zadar, Croatia, September 2009. This work is supported by the Ministry of Science, Education and Sports, Republic of Croatia, through research grants 037-0982913-2762 and 235-2352818-1034.
†Corresponding author. *Email addresses:* `miljenko.marusic@math.hr` (M. Marušić), `darija@mathos.hr` (D. Marković), `jukicd@mathos.hr` (D. Jukić)

The IWD is very flexible and by an appropriate choice of the shape parameter β the density curve can assume a wide variety of shapes (see Figure 1). The density function is strictly increasing on $(\alpha, t_m]$ and strictly decreasing on $[t_m, \infty)$, where $t_m = \alpha + \eta(1 + 1/\beta)^{-1/\beta}$. This implies that the density function is unimodal with the maximum value at t_m. This is in contrast to the standard Weibull model where the shape is either decreasing (for $\beta \leq 1$) or unimodal (for $\beta > 1$). When $\beta = 1$, the IWD becomes an inverse exponential distribution; when $\beta = 2$, it is identical to the inverse Rayleigh distribution; when $\beta = 0.5$, it approximates the inverse Gamma distribution. That is the reason why the IWD is very often used as a model in reliability and lifetime studies (see e.g. Cohen and Whitten [6], Lawles [18], Murthy et al. [21], Nelson [22]).

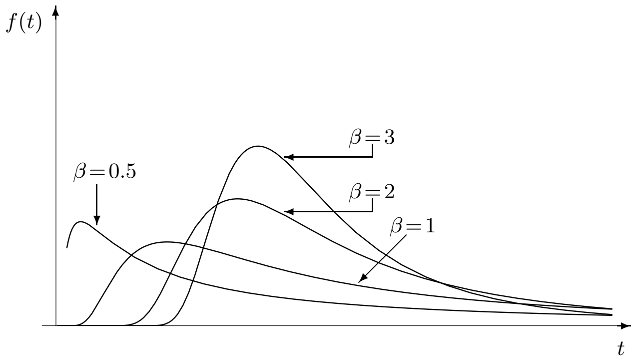

Figure 1. *Plots of the inverse Weibull density for some values of β and by assuming $\alpha = 0$ and $\eta = 1.2$*

In practice, unknown parameters of the appropriate density are not known and must be estimated from a random sample $t_1, \dots, t_n$ consisting of n observations of the random variable T. There is no unique way to perform density reconstruction and many different methods have been proposed in literature. Density estimation methods can be categorized into parametric and nonparametric approaches. Parametric density estimators make assumptions about the functional form of the empirical density, and the estimate is constructed by finding the best parameters for the density function, given the data. In contrast, nonparametric estimators use the characteristics of the data to arrive at an optimal shape without making assumptions about a particular functional form. There are many ways to estimate the density function nonparametrically. These include histograms, kernel estimates, nearest neighbor estimates, and orthogonal series estimates, among others (see e.g. Silverman [25], Tapia and Thompson [28]).

Maximum likelihood (ML) estimation is a traditional parametric method for parameter estimation since it possesses beneficial properties such as asymptotic normality and consistency. For the two-parameter inverse Weibull distribution a standard ML estimate exists and it is unique (see e.g. Calabria and Pulcini [4]). But when the parameter α is added to the inverse Weibull distribution, there is no standard ML estimate over the parameter space $\{(\alpha, \beta, \eta) \in \mathbb{R}^3 : \alpha \geq 0; \beta, \eta > 0\}$ for any observation from the density function (see Jukić and Marković [13]). In the literature, considerable effort has been devoted to such difficulties with the maximum likelihood approach (see e.g. Cheng and Iles [5], Smith and Naylor [27]). There are several other statistical methods for estimating density parameters such as the method of moments, the method of percentiles and the Bayesian method. Unfortunately, none of these methods (excluding the Bayesian one) is appropriate for small data sets (see e.g. Lawless [18], Murthy et al. [21], Nelson [22]). Because of that, some authors suggest to use other methods to determine the unknown parameters (see e.g. Abbasi et al. [1], Jukić et al. [14], Murthy et al. [21], Smith and Naylor [27, 26]).

A very popular method for parameter estimation is the least squares (LS) method. The nonlinear weighted LS fitting problem for the three-parameter Weibull density is considered by Marković et al. [19].

In this paper we consider the parameter estimation problem for the three-parameter inverse Weibull density function. Our approach to density estimation is a combination of nonparametric and parametric methods. The basic idea is to start with the initial nonparametric density estimate $\hat{f}$ which needs to be as good as possible, and then apply a nonlinear LS fit procedure to estimate the unknown parameters α, β and η. The structure of the paper is as follows. In Section 2 we briefly describe the LS method and present our main result (Theorem 1) which guarantees the existence of the LS estimate for the three-parametric inverse Weibull density. Its proof is given in Section 3. In Section 4, some simulation results are given. They show that our approach can give a good density estimate if the initial density is of good enough quality.

2. LS existence theorem for the three-parameter inverse Weibull density

The least squares method required the initial nonparametric density estimates $\hat{f}$ which should be as good as possible. Suppose we are given the points (t_i, y_i), $i = 1, \ldots, n$, $n > 3$, where

$$0 < t_1 < t_2 < \ldots < t_n$$

are observations of the nonnegative inverse Weibull random variable T and $y_i := \hat{f}(t_i)$ are the respective density estimates.

The goal of the LS method is to choose the unknown parameters of density function (1) such that the weighted sum of squared distances between the model and the data is as small as possible. To be more precise, let $w_i > 0$, $i = 1, \ldots, n$, be the data weights which describe the assumed relative accuracy of the data. The unknown parameters α, β and η have to be estimated by minimizing the functional

$$S(\alpha, \beta, \eta) = \sum_{i=1}^{n} w_i [f(t_i; \alpha, \beta, \eta) - y_i]^2 \tag{2}$$

on the set

$$\mathcal{P} := \{(\alpha, \beta, \eta) \in \mathbb{R}^3 : \alpha \geq 0, \beta > 0, \eta > 0\}.$$

A point $(\alpha^\star, \beta^\star, \eta^\star) \in \mathcal{P}$ such that

$$S(\alpha^\star, \beta^\star, \eta^\star) = \inf_{(\alpha,\beta,\eta)\in\mathcal{P}} S(\alpha, \beta, \eta)$$

is called the least squares estimate (LSE), if it exists (see e.g. Bates and Watts [2], Björck [3], Gill et al. [11], Ross [23], Seber and Wild [24]).

Numerical methods for solving the nonlinear LS problem are described in Dennis and Schnabel [9] and Gill et al. [11]. Before starting an iterative procedure one should ask whether a LSE exists. For nonlinear LS problems this question is difficult to answer. As we have already mentioned, the nonlinear weighted LS fitting problem for the three-parameter Weibull density is considered by Marković et al. [19]. Results on the existence of the LSE for some other special classes of functions can be found in [2, 3, 7, 8, 12, 14, 15, 17].

Now we state our main result (Theorem 1) which guarantees the existence of the LSE for the three-parameter inverse Weibull density function. This theorem is also applicable in a classical nonlinear regression problem with the model function of the form (1). Its corresponding generalization in the l_p norm ($p \geq 1$) is given in Remark 1.

Theorem 1. *If the data* (w_i, t_i, y_i), $i = 1, \ldots, n$, $n > 3$, *are such that* $0 < t_1 < t_2 < \ldots < t_n$ *and* $y_i > 0$, $i = 1, \ldots, n$, *then the LSE for the three-parametric inverse Weibull density exists.*

3. Proof of Theorem 1

Before starting the proof of Theorem 1, we need some preliminary results.

Lemma 1. *Suppose we are given data* (w_i, t_i, y_i), $n > 3$, *such that* $0 < t_1 < t_2 < \ldots < t_n$ *and* $w_i, y_i > 0$, $i = 1, \ldots, n$. *Let* (w_r, t_r, y_r) *be a datum such that* $w_r y_r^2$ *is the greatest, i.e.* $w_r y_r^2 = \max_{i=1,\ldots,n} w_i y_i^2$. *Then there exists a point in* $\mathcal{P}$ *at which functional* S *defined by (2) attains a value smaller than*

$$S_r := \sum_{\substack{i=1 \\ i \neq r}}^{n} w_i y_i^2.$$

Proof. Let us first associate with each real $b \in (0, 1)$ a three-parametric inverse Weibull density function

$$f(t; 0, \beta(b), \eta(b)) = \begin{cases} \frac{\beta(b)}{t} \left(\frac{\eta(b)}{t}\right)^{\beta(b)} \mathrm{e}^{-\left(\frac{\eta(b)}{t}\right)^{\beta(b)}}, & t > 0 \\ 0, & t \leq 0 \end{cases} \tag{3}$$

where

$$\beta(b) := t_r y_r \frac{\mathrm{e}^b}{b}, \quad \eta(b) := t_r b^{1/\beta(b)}, \qquad b > 0.$$

This function has maximum at the point $\eta(b)(1 + 1/\beta(b))^{-1/\beta(b)} = t_r$ and it is strictly increasing on $(0, t_r]$ and strictly decreasing on $[t_r, \infty)$. By a straightforward calculation, it can be verified that

$$f(t_r; 0, \beta(b), \eta(b)) = y_r, \tag{4}$$

$$\lim_{b \to 0} \beta(b) = \infty, \tag{5}$$

$$\lim_{b \to 0} \eta(b) = t_r. \tag{6}$$

Now we are going to show that

$$\lim_{b\to 0} f(t;0,\beta(b),\eta(b)) = 0, \qquad t \neq t_r. \tag{7}$$

In view of (6), we obtain

$$\lim_{b\to 0}\Big(\frac{\eta(b)}{t}\Big) = \frac{t_r}{t}.$$

If $t > t_r$, then from (5) it follows readily that

$$\lim_{b\to 0} \mathrm{e}^{-(\frac{\eta(b)}{t})^{\beta(b)}} = 1 \quad \text{and} \quad \lim_{b\to 0} \beta(b)\Big(\frac{\eta(b)}{t}\Big)^{\beta(b)} = 0,$$

and therefore

$$\lim_{b\to 0} f(t;0,\beta(b),\eta(b)) = \lim_{b\to 0}\Big[\frac{\beta(b)}{t}\Big(\frac{\eta(b)}{t}\Big)^{\beta(b)} \mathrm{e}^{-(\frac{\eta(b)}{t})^{\beta(b)}}\Big] = 0.$$

If $t < t_r$, then there exists a sufficiently great $k_0 \in \mathbb{N}$ such that

$$\mathrm{e} < \Big(\frac{\eta(b)}{t}\Big)^{k_0}$$

for every sufficiently small $b > 0$. Now, by using the inequality $x < \mathrm{e}^x$ $(x \geq 0)$ we obtain

$$\beta(b) < \mathrm{e}^{\beta(b)} < \Big(\frac{\eta(b)}{t}\Big)^{k_0\beta(b)}, \qquad b \approx 0,$$

and therefore, for any $b \approx 0$ we have

$$0 < f(t;0,\beta(b),\eta(b)) = \frac{\beta(b)}{t}\Big(\frac{\eta(b)}{t}\Big)^{\beta(b)} \mathrm{e}^{-(\frac{\eta(b)}{t})^{\beta(b)}} < \frac{1}{t}\Big(\frac{\eta(b)}{t}\Big)^{(k_0+1)\beta(b)} \mathrm{e}^{-(\frac{\eta(b)}{t})^{\beta(b)}}.$$

Since

$$\lim_{b\to 0}\Big(\frac{\eta(b)}{t}\Big)^{(k_0+1)\beta(b)} \mathrm{e}^{-(\frac{\eta(b)}{t})^{\beta(b)}} = 0,$$

then from the above-mentioned inequality it follows that

$$\lim_{b\to 0} f(t;0,\beta(b),\eta(b)) = 0, \qquad t < t_r.$$

Thus, we proved the desired limits (7).

Let b_0 be sufficiently small, so that

$$0 < f(t_i;0,\beta(b_0),\eta(b_0)) \leq y_i,$$

whereby the equality holds only if $t_i = t_r$. Due to (4) and (7), such b_0 exists. Then

$$S(0,\beta(b_0),\eta(b_0)) = \sum_{i=1}^{n} w_i\big[f(t_i;0,\beta(b_0),\eta(b_0)) - y_i\big]^2 < \sum_{\substack{i=1\\ i\neq r}}^{n} w_i y_i^2 = S_r.$$

This completes the proof of the lemma. □

Proof of Theorem 1. Since functional S is nonnegative, there exists $S^\star := \inf_{(\alpha,\beta,\eta)\in\mathcal{P}} S(\alpha,\beta,\eta)$. It should be shown that there exists a point $(\alpha^\star,\beta^\star,\eta^\star)\in\mathcal{P}$, such that $S(\alpha^\star,\beta^\star,\eta^\star)=S^\star$.

Let $(\alpha_k,\beta_k,\eta_k)$ be a sequence in $\mathcal{P}$, such that

$$\begin{aligned} S^\star &= \lim_{k\to\infty} S(\alpha_k,\beta_k,\eta_k) = \lim_{k\to\infty}\sum_{i=1}^n w_i[f(t_i;\alpha_k,\beta_k,\eta_k)-y_i]^2 \\ &= \lim_{k\to\infty}\Big\{\sum_{t_i\le\alpha_k} w_i y_i^2 + \sum_{t_i>\alpha_k} w_i\Big[\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i-\alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} - y_i\Big]^2\Big\}. \end{aligned} \tag{8}$$

The summation $\sum\limits_{t_i>\alpha_k}$ (or $\sum\limits_{t_i<\alpha_k}$) is to be understood as follows: The sum over those indices $i\le n$ for which $t_i>\alpha_k$ (or $t_i<\alpha_k$). If there are no such points t_i, the sum is empty; following the usual convention, we define it to be zero.

Without loss of generality, in further consideration we may assume that sequences (α_k), (β_k) and (η_k) are monotone. This is possible because the sequence $(\alpha_k,\beta_k,\eta_k)$ has a subsequence $(\alpha_{l_k},\beta_{l_k},\eta_{l_k})$, such that all its component sequences (α_{l_k}), (β_{l_k}) and (η_{l_k}) are monotone; and since $\lim_{k\to\infty} S(\alpha_{l_k},\beta_{l_k},\eta_{l_k}) = \lim_{k\to\infty} S(\alpha_k,\beta_k,\eta_k) = S^\star$.

Since each monotone sequence of real numbers converges in the extended real number system $\bar{\mathbb{R}}$, define

$$\alpha^\star := \lim_{k\to\infty}\alpha_k, \quad \beta^\star := \lim_{k\to\infty}\beta_k, \quad \eta^\star := \lim_{k\to\infty}\eta_k.$$

Note that $0\le\alpha^\star,\beta^\star,\eta^\star\le\infty$, because $(\alpha_k,\beta_k,\eta_k)\in\mathcal{P}$.

To complete the proof it is enough to show that $(\alpha^\star,\beta^\star,\eta^\star)\in\mathcal{P}$, i.e. that $0\le\alpha^\star<\infty$ and $\beta^\star,\eta^\star\in(0,\infty)$. The continuity of functional S will then imply that $S^\star=\lim_{k\to\infty} S(\alpha_k,\beta_k,\eta_k) = S(\alpha^\star,\beta^\star,\eta^\star)$.

It remains to show that $(\alpha^\star,\beta^\star,\eta^\star)\in\mathcal{P}$. The proof will be done in five steps. In step 1 we will show that $\alpha^\star<t_n$. In step 2 we will show that $\beta^\star\ne0$. The proof that $\eta^\star\ne\infty$ will be done in step 3. In step 4 we prove that $\eta^\star\ne0$. Finally, in step 5 we show that $\beta^\star\ne\infty$. Before continuing with the proof, let us note that Lemma 1 implies that

$$S^\star < \sum_{\substack{i=1\\ i\ne r}}^n w_i y_i^2 =: S_r \tag{9}$$

where the index r is such that $w_r y_r^2 = \max\limits_{i=1,\ldots,n} w_i y_i^2$.

Step 1. If $\alpha^\star\ge t_n$, from (8) it follows that $S^\star=\sum_{i=1}^n w_i y_i^2 > S_r$, which contradicts (9). Thus, we have proved that $\alpha^\star<t_n$.

By taking an appropriate subsequence of $(\alpha_k,\beta_k,\eta_k)$, if necessary, we may assume that if $t_i<\alpha^\star$, then $t_i<\alpha_k$ for every $k\in\mathbb{N}$. Similarly, if $t_i>\alpha^\star$, we may assume that $t_i>\alpha_k$ for every $k\in\mathbb{N}$. Due to this, now it is easy to show that from (8) it follows that

$$S^\star \ge \sum_{t_i<\alpha^\star} w_i y_i^2 + \lim_{k\to\infty}\Big\{\sum_{t_i>\alpha^\star} w_i\Big[\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i-\alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} - y_i\Big]^2\Big\}. \tag{10}$$

Step 2. If $\beta^\star = 0$, then by using the inequality $x < \mathrm{e}^x$ $(x \geq 0)$ we obtain

$$0 < \frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} < \frac{\beta_k}{t_i - \alpha_k}, \qquad t_i > \alpha^\star,$$

wherefrom it follows readily that

$$\lim_{k\to\infty}\Big[\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}\Big] = 0, \qquad t_i > \alpha^\star.$$

Now, from (10), it follows that $S^\star \geq \sum_{t_i \neq \alpha^\star} w_i y_i^2 \geq S_r$. This contradicts (9). Thus, we have proved that $\beta^\star \neq 0$.

Step 3. Let us show that $\eta^\star \neq \infty$. We prove this by contradiction. Suppose on the contrary that $\eta^\star = \infty$. Without loss of generality, we may then assume that if $t_i > \alpha^\star$, then $\mathrm{e} < \frac{\eta_k}{t_i-\alpha_k}$ for all $k \in \mathbb{N}$. Then from the inequality $x < \mathrm{e}^x$ $(x \geq 0)$ it follows that if $t_i > \alpha^\star$, then

$$\beta_k < \mathrm{e}^{\beta_k} < \Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k}, \qquad k \in \mathbb{N}.$$

Thus, if $t_i > \alpha^\star$, then

$$\begin{aligned} 0 &< \frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} = \frac{\beta_k}{t_i - \alpha_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} \\ &< \frac{1}{t_i - \alpha_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{2\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}. \end{aligned} \tag{11}$$

Furthermore, since $\lim_{k\to\infty}\big(\frac{\eta_k}{t_i-\alpha_k}\big) = \infty$ and $\beta^\star \neq 0$, we have $\lim_{k\to\infty}\big(\frac{\eta_k}{t_i-\alpha_k}\big)^{\beta_k} = \infty$ and therefore $\lim_{k\to\infty}\big(\frac{\eta_k}{t_i-\alpha_k}\big)^{2\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} = 0$, so that from (11) it follows that

$$\lim_{k\to\infty}\Big[\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}\Big] = 0, \qquad t_i > \alpha^\star.$$

Putting the above limits into (10), we immediately obtain $S^\star \geq \sum_{t_i \neq \alpha^\star} w_i y_i^2 \geq S_r$, which contradicts (9). Hence we proved that $\eta^\star \neq \infty$.

So far we have shown that $\alpha^\star < t_n$, $\beta^\star \neq 0$ and $\eta^\star \neq \infty$. By using this, in the next step we will show that $\eta^\star \neq 0$.

Step 4. Let us show that $\eta^\star \neq 0$. To see this, suppose on the contrary that $\eta^\star = 0$. Then only one of the following two cases can occur: (i) $\eta^\star = 0$ and $\beta^\star \in (0, \infty)$, or (ii) $\eta^\star = 0$ and $\beta^\star = \infty$. Now, we are going to show that functional S cannot attain its infimum in either of these two cases, which will prove that $\eta^\star \neq 0$.

Case (i): $\eta^\star = 0$ and $\beta^\star \in (0, \infty)$. In this case we would have

$$\begin{aligned} \lim_{k\to\infty}\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} &= \lim_{k\to\infty}\frac{\beta_k}{t_i - \alpha_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} \\ &= 0, \qquad t_i > \alpha^\star \end{aligned}$$

and hence from (10) it would follow that

$$S^\star \geq \sum_{t_i \neq \alpha^\star} w_i y_i^2 \geq S_r$$

which contradicts assumption (9).

Case (ii): $\eta^\star = 0$ and $\beta^\star = \infty$. Since $\eta_k \to 0$, there exists a real number $q > 1$ and sufficiently great $k_0 \in \mathbb{N}$ such that if $t_i > \alpha^\star$ and $k > k_0$, then $\eta_k/(t_i - \alpha_k) < 1/q$. Without loss of generality, we may assume that $k_0 = 1$. Thus, if $t_i > \alpha^\star$, then

$$\begin{aligned} 0 &< \frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} = \frac{\beta_k}{t_i - \alpha_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} \\ &< \frac{1}{t_i - \alpha_k}\Big(\frac{\beta_k}{q^{\beta_k}}\Big) \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}. \end{aligned} \tag{12}$$

Furthermore, since

$$\lim_{k\to\infty}\Big(\frac{\beta_k}{q^{\beta_k}}\Big) = 0 \qquad \text{and} \qquad \lim_{k\to\infty} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} = 1,$$

from (12) it follows that

$$\lim_{k\to\infty}\Big[\frac{\beta_k}{\eta_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k+1} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}\Big] = 0, \qquad t_i > \alpha^\star.$$

Finally, from (10) we obtain $S^\star \ge \sum_{t_i \ne \alpha^\star} w_i y_i^2 \ge S_r$, which contradicts assumption (9). This means that in this case functional S cannot attain its infimum.

Thus, we have proved that $\eta^\star \ne 0$.

Step 5. It remains to show that $\beta^\star \ne \infty$. We prove this by contradiction. Suppose that $\beta^\star = \infty$.

Arguing as in case (ii) from step 4, it can be shown that

$$\lim_{k\to\infty}\Big[\frac{\beta_k}{t_i - \alpha_k}\big(\frac{\eta_k}{t_i - \alpha_k}\big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}\Big] = 0, \qquad \text{if } 0 < \frac{\eta^\star}{t_i - \alpha^\star} < 1. \tag{13}$$

If $\frac{\eta^\star}{t_i - \alpha^\star} > 1$, then there exists a sufficiently great $k_0 \in \mathbb{N}$ such that $\mathrm{e} < \big(\frac{\eta_k}{t_i - \alpha_k}\big)^{k_0}$. Now, by using the inequality $x < \mathrm{e}^x$ $(x \ge 0)$ we obtain

$$\beta_k < \mathrm{e}^{\beta_k} < \Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{k_0\beta_k}, \qquad k \in \mathbb{N},$$

and therefore

$$0 < \frac{\beta_k}{t_i - \alpha_k}\Big(\frac{\eta_k}{t_i - \alpha_k}\Big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} < \frac{1}{t_i - \alpha_k}\big(\frac{\eta_k}{t_i - \alpha_k}\big)^{(k_0+1)\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}. \tag{14}$$

Since $\lim_{k\to\infty}\big(\frac{\eta_k}{t_i-\alpha_k}\big)^{\beta_k} = \infty$, we have that $\lim_{k\to\infty}\big(\frac{\eta_k}{t_i-\alpha_k}\big)^{(k_0+1)\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}} = 0$ and therefore from (14) it follows that

$$\lim_{k\to\infty}\Big[\frac{\beta_k}{t_i - \alpha_k}\big(\frac{\eta_k}{t_i - \alpha_k}\big)^{\beta_k} \mathrm{e}^{-(\frac{\eta_k}{t_i-\alpha_k})^{\beta_k}}\Big] = 0, \qquad \text{if } \frac{\eta^\star}{t_i - \alpha^\star} > 1. \tag{15}$$

From (10), (13) and (15) we would obtain $S^\star \ge \sum_{t_i \ne \alpha^\star} w_i y_i^2 \ge S_r$, which contradicts (9). Thus, we have proved that $\beta^\star \ne \infty$ and completed the proof. □

Remark 1. *Given* $1 \leq p < \infty$*, let*

$$S_p(\alpha, \beta, \eta) = \sum_{i=1}^{n} w_i |f(t_i; \alpha, \beta, \eta) - y_i|^p.$$

Arguing in a similar way as in proofs of Lemma 1 and Theorem 1, it can be easily shown that there exists a point $(\alpha_p^\star, \beta_p^\star, \eta_p^\star) \in \mathcal{P}$ *such that* $S_p(\alpha_p^\star, \beta_p^\star, \eta_p^\star) = \inf_{(\alpha,\beta,\eta)\in\mathcal{P}} S_p(\alpha, \beta, \eta)$*. It suffices to replace the* l_2 *norm with the* l_p *norm. Thereby all parts of the proof remain the same.*

4. Simulation study

In this section we exemplify our method on generated data. For given parameters we generate random data with the three-parameter inverse Weibull distribution. To obtain density data we use symmetric and adaptive kernel estimates, common nonparametric density estimates. After that, we estimate model parameters by the nonlinear least squares method.

4.1. Nonparametric estimates for initial density

Symmetric kernel estimator. Firstly, we consider the symmetric kernel estimates for a density function, i.e. a function of the form

$$\hat{f}_{sk}(t) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right),$$

where K is a kernel and h is the window width (also called the smoothing parameter or bandwidth [25]). For a kernel K we choose normal density

$$K(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}}.$$

The bandwidth h is set to the ideal value for the inverse Weibull distribution and the chosen kernel K in the sense of minimizing the approximate mean integrated square error (see Silverman [25], page 40):

$$\begin{aligned} h_{opt} &= n^{-1/5} \left(\int t^2 K(t)\, dt\right)^{-2/5} \left(\int K^2(t)\, dt\right)^{1/5} \left(\int [f''(t)]^2\, dt\right)^{-1/5} \\ &= \sigma^{-1} \left(2n\sqrt{\pi} \int [f''(t)]^2\, dt\right)^{-1/5}. \end{aligned}$$

Here $f(t) = f(t; \alpha, \beta, \eta)$ is the inverse Weibull density given by (1). Since we choose normal density for kernel:

$$K\left(\frac{t - t_i}{h}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(\frac{t-t_i}{h}\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t - t_i)^2}{2(h\sigma)^2}\right),$$

the symmetric kernel estimates reads

$$\hat{f}_{sk}(t) = \frac{1}{nh\sigma\sqrt{2\pi}} \sum_{i=1}^{n} \exp\left(-\frac{(t-t_i)^2}{2(h\sigma)^2}\right).$$

To determine symmetric kernel estimates we have to define the product $h_{opt}\sigma$ instead of h_{opt} and σ separately. We still have to estimate a term $\int [f''(t)]^2\,dt$. For the simulation purposes we calculate this integral for the original density function.

Adaptive kernel estimator. The adaptive kernel approach (see e.g. Silverman [25]) is a two-stage procedure. Instead of using constant bandwidth h, for each point t_i a different bandwidth is used.

In the first stage, a pilot estimate $\tilde{f}$ is used to get a rough idea of the density and to yield a pattern of bandwidths corresponding to various observations. After that, local bandwidth factors λ_i are defined by

$$\lambda_i = \left(\frac{1}{g}\tilde{f}(t_i)\right)^{-\gamma},$$

where $g = \left(\prod_{i=1}^{n} \tilde{f}(t_i)\right)^{1/n}$ is the geometric mean of numbers $\tilde{f}(t_1), \ldots, \tilde{f}(t_n)$ and $\gamma \in [0,1]$ is the sensitivity parameter. As a pilot estimate we use a symmetric kernel estimate $\hat{f}_{sk}$.

The adaptive kernel estimate $\hat{f}_{ak}$ is given by

$$\hat{f}_{ak}(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h\lambda_i} K\left(\frac{t-t_i}{h\lambda_i}\right),$$

where K is the normal density with variance σ^2, as in the symmetric kernel approach.

The product of smoothing parameters h and σ is kept the same as in the symmetric kernel approach and the value for γ was obtained subjectively. (The value $\gamma = 0.2$ was applied throughout all simulations.)

Once the nonparametric density estimate $\hat{f}_{sk}$ or $\hat{f}_{ak}$ was obtained, the parameters α, β and η were estimated by solving the following LS problem:

$$\min_{(\alpha,\beta,\eta)\in\mathcal{P}} \sum_{i=1}^{n} [f(t_i;\alpha,\beta,\eta) - y_i]^2,$$

where

$$y_i = \hat{f}_{sk}(t_i)$$

in the symmetric kernel approach or

$$y_i = \hat{f}_{ak}(t_i)$$

in the adaptive kernel approach.

Since the kernel K is strictly positive, Theorem 1 guarantees the existence of the LSE $(\hat{\alpha}_{sk}, \hat{\beta}_{sk}, \hat{\eta}_{sk})$ (i.e. $(\hat{\alpha}_{ak}, \hat{\beta}_{ak}, \hat{\eta}_{ak})$). For the nonlinear minimization we use a package for nonlinear regression used in [20].

4.2. Results

In this simulation study we present results based on 1000 simulations. In each simulation the values of parameters α, β and η were fixed at $\alpha = 10$, $\beta = 2$ and $\eta = 10$. Using a random number generator we generated a large number of data points distributed according to the inverse Weibull distribution with prescribed parameters (we compared results obtained for 500, 1000 and 2000 data points). After that, we calculated density data using a nonparametric estimate for initial density. We applied both symmetric and adaptive kernel estimators. Approximate density values were calculated at each data point. Typical density data are shown in Figure 2.

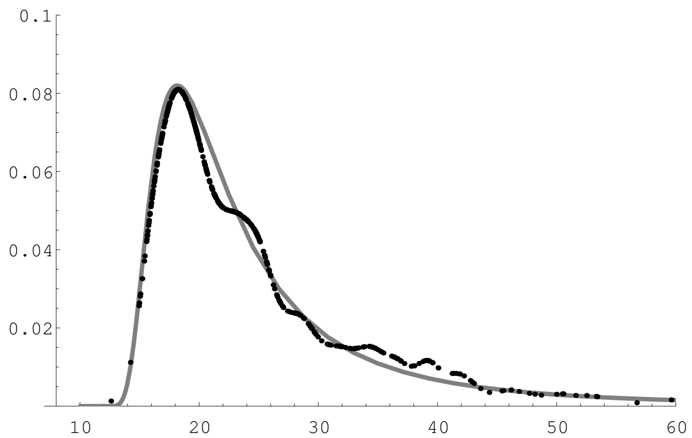

Figure 2. *Typical random density data (black dots) and theoretical IW density curve (gray line)*

To such data we fitted an IW model. The fitted curve is shown in Figure 3.

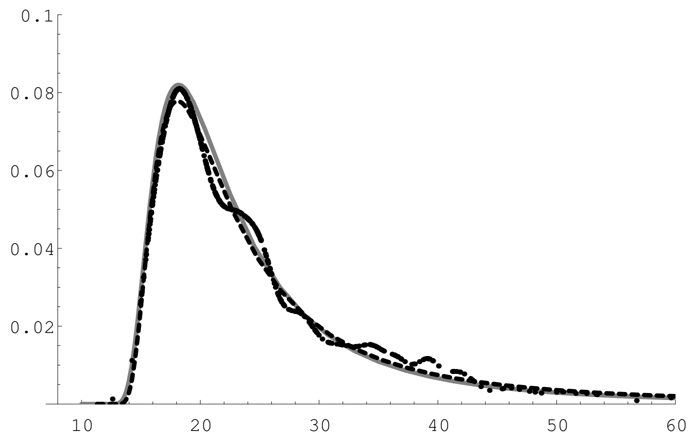

Figure 3. *Typical fitted IW curve (dashed line) against random density data (black dots) and theoretical IW density curve (gray line)*

The average of estimated parameters $\hat{\alpha}, \hat{\beta}, \hat{\eta}$ and $\sqrt{S/n}$ for 1000 simulations are given in Table 1. Here S stands for the sum of squared errors:

$$S = S(\hat{\alpha}, \hat{\beta}, \hat{\eta}) = \sum_{i=1}^{n} [f(t_i; \hat{\alpha}, \hat{\beta}, \hat{\eta}) - y_i]^2,$$

where n is the number of data points.

	Number of data points					
	$n = 500$		$n = 1000$		$n = 2000$	
	mean	standard deviation	mean	standard deviation	mean	standard deviation
$\hat{\alpha}$	7.7	7.2	8.6	1.9	8.7	1.2
$\hat{\beta}$	2.5	1.5	2.3	0.4	2.2	0.2
$\hat{\eta}$	12.4	7.1	11.6	1.9	11.4	1.2
$\sqrt{\frac{1}{n}S}$	$3.2 \cdot 10^{-3}$		$2.3 \cdot 10^{-3}$		$1.7 \cdot 10^{-3}$	

Table 1. *Results of the simulation study ($\alpha = 10$, $\beta = 2$, $\eta = 10$, $N = 1000$ simulations) for symmetric kernel initial density*

We can see from Table 1 that there is a bias in parameter estimation. Bias is decreasing with an increase in the number of data points. Still, there is a relatively large standard deviation in parameters estimation. It is also decreasing with an increase in the number of data points. The use of adaptive kernel initial density somehow improves result, but not substantially (Table 2).

	Number of data points					
	$n = 500$		$n = 1000$		$n = 2000$	
	mean	standard deviation	mean	standard deviation	mean	standard deviation
$\hat{\alpha}$	8.4	8.5	9.4	1.8	9.5	1.1
$\hat{\beta}$	2.3	1.8	2.1	0.4	2.1	0.2
$\hat{\eta}$	11.6	8.5	10.7	1.8	10.6	1.1
$\sqrt{\frac{1}{n}S}$	$3.8 \cdot 10^{-3}$		$2.8 \cdot 10^{-3}$		$2.0 \cdot 10^{-3}$	

Table 2. *Results of the simulation study ($\alpha = 10$, $\beta = 2$, $\eta = 10$, $N = 1000$ simulations) for adaptive kernel initial density*

Checking dependency between estimated parameters, we observed a high correlation; Pearson correlation coefficients were $r_{\alpha\beta} = -0.978$, $r_{\alpha\eta} = -0.995$ and $r_{\beta\eta} = 0.970$ (for 1000 simulations and the use of symmetric kernel initial density). Dispersion diagrams for estimated parameters are presented on Figures 4–6. This correlation indicates that a large standard deviation in parameters estimation is not due the inadequacy of the method but rather due to overparametrization. To check this hypothesis we performed simulations with one parameter fixed to the original value and other two parameters determined by minimization. The result is shown in Figures 7 and 8. Resulting estimations are almost unbiased with a much smaller standard deviation.

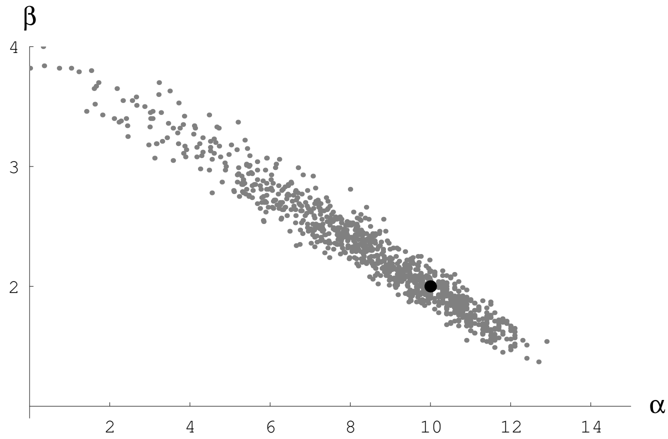

Figure 4. *Dispersion diagram for parameters α and β (500 simulations)*

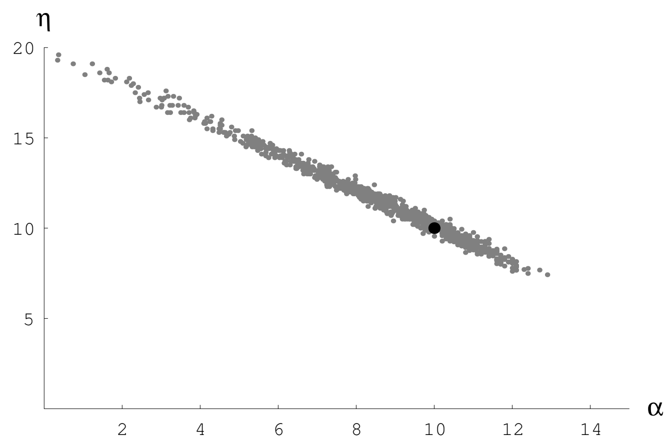

Figure 5. *Dispersion diagram for parameters α and η (500 simulations)*

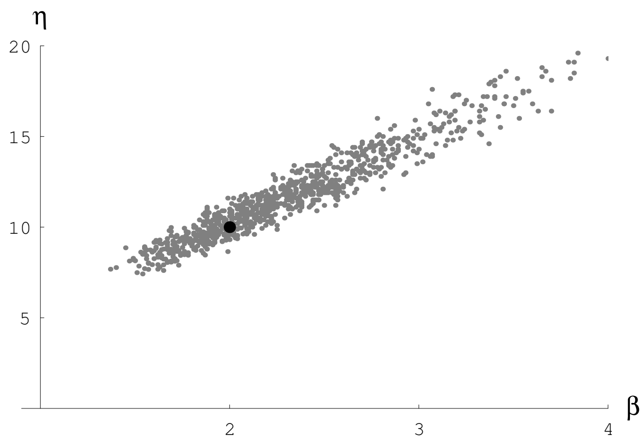

Figure 6. *Dispersion diagram for parameters β and η (500 simulations)*

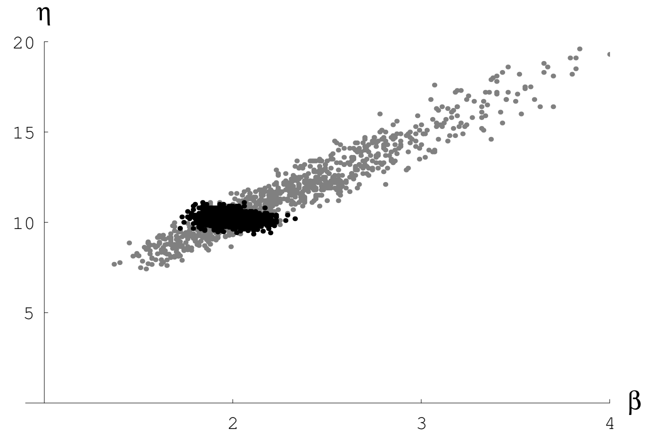

Figure 7. *Dispersion diagram for parameters β and η (500 simulations). Comparison of IW model (gray) and IW model with fixed parameter α (black)*

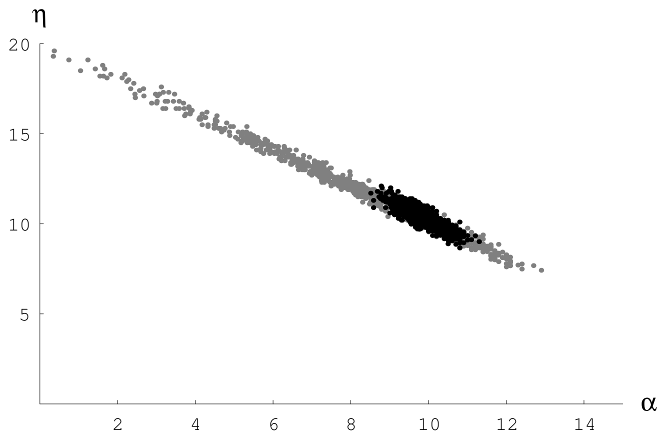

Figure 8. *Dispersion diagram for parameters α and η (500 simulations). Comparison of IW model (gray) and IW model with fixed parameter β (black)*

References

[1] B. Abbasi, A. H. E. Jahromi, J. Arkat, M. Hosseinkouchack, *Estimating the parameters of Weibull distribution using simulated annealing algorithm*, Appl. Math. Comput. **183**(2006), 85–93.

[2] D. M. Bates, D. G. Watts, *Nonlinear regression analysis and its applications*, Wiley, New York, 1988.

[3] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[4] R. Calabria, G. Pulcini, *On the maximum likelihood and least-squares estimation in the inverse Weibull distribution*, Statist. Appl. **2**(1990), 53–63.

[5] R. C. H. Cheng, T. C. Iles, *Embedded models in three-parameter distributions and their estimation* J. R. Stat. Soc. Ser. B **52**(1990), 135–149.

[6] A. C. Cohen, B. J. Whitten, *Parameter Estimation in Reliability and Life Span Models*, Marcel Dekker Inc., New York and Basel, 1988.

[7] E. Z. Demidenko, *Criteria for global minimum of sum of squares in nonlinear regression*, Comput. Statis. Data Anal. **51**(2006), 1739–1753.

[8] E. Z. Demidenko, *On the existence of the least squares estimate in nonlinear growth curve models of exponential type*, Commun. Statist.-Theory Meth. **25**(1996), 159–182.

[9] J. E. Dennis, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.

[10] P. Erto, *New Practical Bayes estimators for the 2-Parameter Weibull distribution*, IEEE Transactions on Reliability **R-31**(1982), 194–197.

[11] P. E. Gill, W. Murray, M. H. Wright, *Practical Optimization*, Academic Press, London, 1981.

[12] K. P. Hadeler, D. Jukić, K. Sabo, *Least squares problems for Michaelis Menten kinetics*, Math. Meth. Appl. Sci. **30**(2007), 1231–1241.

[13] D. Jukić, D. Marković, *On nonlinear weighted least squares fitting of the three-parameter inverse Weibull distribution*, Math. Commun. **15** (2010), 13–24.

[14] D. Jukić, M. Benšić, R. Scitovski, *On the existence of the nonlinear weighted least squares estimate for a three-parameter Weibull distribution*, Comput. Statis. Data Anal. **52**(2008), 4502–4511.

[15] D. Jukić, G. Kralik, R. Scitovski, *Least squares fitting Gompertz curve*, J. Comput. Appl. Math. **169**(2004), 359–375.

[16] D. Jukić, R. Scitovski, H. Späth, *Partial linearization of one class of the nonlinear total least squares problem by using the inverse model function*, Computing **62**(1999), 163–178.

[17] D. Jukić, R. Scitovski, *Existence results for special nonlinear total least squares problem*, J. Math. Anal. Appl. **226**(1998), 348–363.

[18] J. F. Lawless, *Statistical models and methods for lifetime data*, Wiley, New York, 1982.

[19] D. Marković, D. Jukić, M. Benšić, *Nonlinear weighted least squares estimation of a three-parameter Weibull density with a nonparametric start*, J. Comput. Appl. Math. **228**(2009), 304–312.

[20] M. Marušić, Ž. Bajzer, J.P. Freyer, S. Vuk-Pavlović, *Analysis of Growth of Multicellular Tumor Spheroids by Mathematical Models*, Cell Proliferations **27**(1994), 73–94.

[21] D. N. P. Murthy, M. Xie, R. Jiang, *Weibull Models*, Wiley, New York, 2004.

[22] W. Nelson, *Applied Life Data Analysis*, Wiley, New York, 1982.

[23] G. J. S. Ross, *Nonlinear Estimation*, Springer, New York, 1990.

[24] G. A. F. Seber, C. J. Wild, *Nonlinear Regression*, Wiley, New York, 1989.

[25] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall/CRC, Boca Raton, 2000.

[26] R. L. Smith, J. C. Naylor, *Statistics of the three-parameter Weibull distribution*, Annals of Operations Research **9**(1987), 577–587.

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

[28] R. A. Tapia, J. R. Thompson, *Nonparametric Probability Density Estimation*, Johns Hopkins University Press, Baltimore, 1978.