MATHEMATICAL COMMUNICATIONS

Math. Commun., Vol. **15**, No. 1, pp. 13-24 (2010)

On nonlinear weighted least squares fitting of the three-parameter inverse Weibull distribution*

Dragan Jukić[1] and Darija Marković[1,†]

[1] *Department of Mathematics, University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia*

Received January 5, 2009; accepted April 14, 2009

Abstract. In this paper we consider nonlinear least squares fitting of the three-parameter inverse Weibull distribution to the given data (w_i, t_i, y_i), $i = 1, \dots, n$, $n \geq 3$. As the main result, we show that the least squares estimate exists provided that the data satisfy just the following two natural conditions: (i) $0 < t_1 < t_2 < \dots < t_n$ and (ii) $0 < y_1 < y_2 < \dots < y_n < 1$. To this end, an illustrative numerical example is given.

AMS subject classifications: 65D10, 62N05, 62J02, 62G07

Key words: three-parameter inverse Weibull distribution, data fitting, nonlinear least squares, least squares estimate, existence problem

1. The three-parameter inverse (or reverse) Weibull distribution

Let X denote the three-parameter Weibull model with distribution function

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

where $\alpha \geq 0$ is a location parameter, $\gamma > 0$ the scale, and $\beta > 0$ the shape parameter (see e.g. Murthy et al. [17]). Then a random variable T defined by

$$T = \alpha + \frac{\eta^2}{X - \alpha}$$

has the distribution function

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

*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 grant 235-2352818-1034.
†Corresponding author. *Email addresses:* `jukicd@mathos.hr` (D. Jukić), `darija@mathos.hr` (D. Marković)

and the density function

$$f(t;\alpha,\beta,\eta) = \begin{cases} \frac{\beta}{\eta}\left(\frac{\eta}{t-\alpha}\right)^{\beta+1} \mathrm{e}^{-(\frac{\eta}{t-\alpha})^{\beta}}, & t > \alpha \\ 0, & t \le \alpha. \end{cases} \tag{2}$$

The distribution of T is known as the three-parameter inverse Weibull distribution (IWD). If $\alpha = 0$, the resulting distribution is called a two-parameter IWD. Drapella [8] calls the IWD the complementary Weibull distribution, while Mudholkar and Kollia [16] call it the reciprocal Weibull distribution. The IWD is also known as a type 2 extreme value or the Fréchet distribution (Johnson et al. [10]). For more details on the IWD, see e.g. Johnson et al. [10] and Murthy et al. [17].

The IWD plays an important role in reliability and lifetime studies (see [15, 17, 18]). The reliability function $R(t)$ and the hazard (failure rate) function $h(t)$ for the IWD are given by

$$R(t;\alpha,\beta,\eta) = 1 - F(t;\alpha,\beta,\eta) = 1 - \mathrm{e}^{-(\frac{\eta}{t-\alpha})^{\beta}}, \quad t > \alpha$$

and

$$h(t;\alpha,\beta,\eta) = \frac{f(t;\alpha,\beta,\eta)}{1 - F(t;\alpha,\beta,\eta)} = \frac{\beta\eta^{\beta}(t-\alpha)^{-\beta-1}\,\mathrm{e}^{-(\frac{\eta}{t-\alpha})^{\beta}}}{1 - \mathrm{e}^{-(\frac{\eta}{t-\alpha})^{\beta}}}, \quad t > \alpha.$$

It can be shown that the hazard function is similar to that of the log-normal and inverse Gaussian distributions.

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 $\beta \le 1$) 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 one of the most widely used models in reliability and lifetime studies.

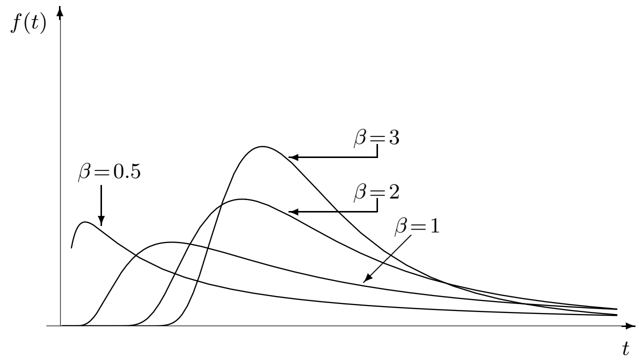

Figure 1. *Plots of the inverse Weibull density for some values of β and by assuming $\alpha = 0$ and $\eta = 1.2$*

There is no unique way to perform density reconstruction from the observed data and many different methods have been proposed in the literature. The maximum likelihood (ML) method is a traditional method since it possesses beneficial properties such as asymptotic normality and consistency. Assuming n independent observations $t_1, \ldots, t_n$ from the density $p(t;\theta)$, the likelihood function of this sample is given by $L(\theta) = \prod_{i=1}^{n} p(t_i;\theta)$. The maximum likelihood estimate (MLE) of parameters θ is the value $\hat{\theta}$ that maximizes the likelihood function. But the MLE does not necessarily exist, and it is not necessarily unique. For the two-parameter IWB a standard MLE exists and it is unique (see e.g. Calabria and Pulcini [3]). Now we are going to show that for the three-parameter IWB the likelihood function

$$L(\alpha,\beta,\eta) = \prod_{i=1}^{n} f(t_i;\alpha,\beta,\eta) = \prod_{i=1}^{n} \frac{\beta}{\eta}\Big(\frac{\eta}{t_i-\alpha}\Big)^{\beta+1} \mathrm{e}^{-(\frac{\eta}{t_i-\alpha})^{\beta}}$$

is unbounded from above so that a standard MLE does not exist. In order to verify this, we may assume, without loss of generality, that $0 < t_1 < t_2 < \ldots < t_n$. Fix $\eta \in (0,\infty)$. Then

$$\begin{aligned} L(t_1-\beta^{n+1},\beta,\eta) &= \prod_{i=1}^{n} \frac{\beta}{\eta}\Big(\frac{\eta}{t_i-t_1+\beta^{n+1}}\Big)^{\beta+1} \mathrm{e}^{-(\frac{\eta}{t_i-t_1+\beta^{n+1}})^{\beta}} \\ &= \frac{\eta^{1-n}}{\beta}\Big(\frac{\eta}{\beta^{n+1}}\Big)^{\beta} \mathrm{e}^{-(\frac{\eta}{\beta^{n+1}})^{\beta}} \prod_{i=2}^{n}\Big(\frac{\eta}{t_i-t_1+\beta^{n+1}}\Big)^{\beta+1} \mathrm{e}^{-(\frac{\eta}{t_i-t_1+\beta^{n+1}})^{\beta}} \quad (3) \end{aligned}$$

Since

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

and

$$\lim_{\beta\to 0}\prod_{i=2}^{n}\Big(\frac{\eta}{t_i-t_1+\beta^{n+1}}\Big)^{\beta+1} \mathrm{e}^{-(\frac{\eta}{t_i-t_1+\beta^{n+1}})^{\beta}} = \Big(\frac{\eta}{\mathrm{e}}\Big)^{n-1}\prod_{i=2}^{n}\frac{1}{t_i-t_1},$$

from (3) it follows that $\lim_{\beta\to 0} L(t_1-\beta^{n+1},\beta,\eta) = \infty$, and hence the MLE does not exist. In the literature, considerable effort has been devoted to such difficulties with the maximum likelihood approach (see e.g. Cheng and Iles [4], Smith and Naylor [21]). There are several other statistical methods for estimating model parameters such as the method of moments, the method of percentile and the Bayesian method. Unfortunately, none of these methods (excluding Bayesian) is appropriate for small data sets (see e.g. Lawless [15], Murthy et al. [17], Nelson [18]).

A very popular method for parameter estimation is the least squares (LS) method. The method can be described as follows: Suppose we are given the data (w_i, t_i, y_i), $i = 1,\ldots,n$, $n > 3$, where t_i denotes the values of the independent variable, y_i are the respective measured function values and $w_i > 0$ are the data weights which describe the assumed relative accuracy of the data. The unknown parameters α, β and η of function (1) have to be estimated by minimizing the functional

$$S(\alpha,\beta,\eta) = \sum_{i=1}^{n} w_i\big[F(t_i;\alpha,\beta,\eta)-y_i\big]^2 = \sum_{\substack{i=1\\ t_i\le\alpha}}^{n} w_i y_i^2 + \sum_{\substack{i=1\\ t_i>\alpha}}^{n} w_i\Big[\mathrm{e}^{-(\frac{\eta}{t_i-\alpha})^{\beta}} - y_i\Big]^2$$

on the set $\mathcal{P} := \{(\alpha, \beta, \eta) \in \mathbb{R}^3 : \alpha \geq 0; \beta, \eta > 0\}$. 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 Björck [2], Gill et al. [9], Ross [19], Seber and Wild [20]).

Numerical methods for solving the nonlinear LS problem are described in Dennis and Schnabel [7] and Gill et al. [9]. Before the iterative minimization of the sum of squares it is still necessary to ask whether the least squares estimate (LSE) exists. In the case of nonlinear LS problems it is still extremely difficult to answer this question (see Bates and Watts [1], Björck [2], Demidenko [5, 6], Hadeler et al [12], Jukić et al. [11, 13, 14]).

The problem of nonlinear weighted LS fitting of the three-parameter Weibull distribution to the given data (w_i, t_i, y_i), $i = 1, \ldots, n$, is considererd by Jukić et al. [11]. They showed that the LSE exists provided that the data satisfy just the following two conditions: (i) $0 < t_1 < t_2 < \ldots < t_n$ and (ii) $0 < y_1 < y_2 < \ldots < y_n < 1$. Since the Weibull random variable T is nonegative and since numbers y_i usually denote empirical CDF values, these two conditions are natural. Surprisingly, in spite of the many papers on Weibull models, we have not managed to find any paper dealing with the existence problem of a solution to a nonlinear LS problem for the three-parametric inverse Weibull distribution. The structure of the paper is as follows. In Section 2 we present our main result (Theorem 1) which guarantees the existence of the LSE for the three-parametric inverse Weibull distribution, provided the data satisfy conditions (i) and (ii). An illustrative numerical example is given in Section 3.

2. The existence theorem

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

Lemma 1. *Suppose we are given the data (w_i, t_i, y_i), $i = 1, \ldots, n$, $n > 3$, such that*

(i) $0 < t_1 < t_2 < \ldots < t_n$,

(ii) $0 < y_1 < y_2 < \ldots < y_n < 1$

and $w_i > 0$, $i = 1, \ldots, n$. Given any two real numbers $\theta \geq 0$ and $A \geq 0$, let

$$\Sigma_{\theta,A} := \sum_{t_i<\theta} w_i y_i^2 + \sum_{t_i>\theta} w_i (y_i - A)^2.$$

Then there exists a point in $\mathcal{P}$ at which functional S attains a value less than $\Sigma_{\theta,A}$.

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

Proof. It is easy to verify by definition of $\Sigma_{\theta,A}$ that $\Sigma_{0,A} \geq \Sigma_{t_1,A}$ for any $A \geq 0$ and that $\Sigma_{\theta,A} \geq \Sigma_{\theta,y_n} > \Sigma_{t_{n-1},y_n}$ for all $\theta > t_{n-1}$, $A \geq 0$. Thus, it will suffice to consider the case $\theta \in [t_1, t_{n-1}]$.

Assume that $\theta \in [t_1, t_{n-1}]$. Let $k \in \{1, \ldots, n-1\}$ such that

$$\theta \in (t_{k-1}, t_k],$$

where $t_0 = 0$ by definition. Now, define

$$\xi_0 := \begin{cases} y_k, & \text{if } \theta = t_k \\ \frac{1}{2}(y_{k-1} + y_k), & \text{if } \theta \neq t_k. \end{cases}$$

Furthermore, let a point (τ_1, ξ_1) be defined in the following way:

$$\xi_1 := \frac{\sum\limits_{t_i > \theta} w_i y_i}{\sum\limits_{t_i > \theta} w_i} \qquad \text{and} \qquad \tau_1 := \begin{cases} t_l, & \text{if } y_l = \xi_1 \\ \frac{t_{l-1}+t_l}{2}, & \text{if } y_l > \xi_1, \end{cases}$$

where

$$l := \min\{i : y_i \geq \xi_1\}.$$

Since the sequences $\{t_i\}_{i=1}^n$ and $\{y_i\}_{i=1}^n$ are strictly increasing, it is easy to show that $\theta < \tau_1$ and $\xi_0 < \xi_1$.

Now, we are going to construct a class of inverse Weibull distributions whose graph will contain points (θ, ξ_0) and (τ_1, ξ_1). For this purpose, we first define a function $K : (0, \infty) \to (0, \infty)$ by formula

$$K(\beta) := \frac{\left(\frac{\ln \xi_1}{\ln \xi_0}\right)^{1/\beta}}{1 - \left(\frac{\ln \xi_1}{\ln \xi_0}\right)^{1/\beta}}.$$

Function K is strictly increasing on $\mathbb{R}$, $\lim_{\beta \to 0} K(\beta) = 0$ and $\lim_{\beta \to \infty} K(\beta) = \infty$. Since $\tau_1 > \theta > 0$, there exists $B > 0$ such that $\theta - K(B)(\tau_1 - \theta) = 0$, so that for all $\beta \in (0, B)$, $\theta - K(\beta)(\tau_1 - \theta) > 0$. Note that

$$(\alpha(\beta), \beta, \eta(\beta)) := \left(\theta - K(\beta)(\tau_1 - \theta), \beta, \left(\ln\left(\frac{1}{\xi_1}\right)\right)^{1/\beta}\left(\tau_1 - \theta + K(\beta)(\tau_1 - \theta)\right)\right) \in \mathcal{P}$$

for all $\beta \in (0, B)$. Let us now associate with each real $\beta \in (0, B)$ an inverse Weibull distribution

$$F(t; \alpha(\beta), \beta, \eta(\beta)) = \begin{cases} \mathrm{e}^{-\ln\left(\frac{1}{\xi_1}\right)\left(\frac{\tau_1 - \theta + K(\beta)(\tau_1 - \theta)}{t - \theta + K(\beta)(\tau_1 - \theta)}\right)^{\beta}}, & t > \theta - K(\beta)(\tau_1 - \theta) \\ 0, & t \leq \theta - K(\beta)(\tau_1 - \theta). \end{cases}$$

It is not difficult to show that

$$F(\theta; \alpha(\beta), \beta, \eta(\beta)) = \xi_0, \qquad F(\tau_1; \alpha(\beta), \beta, \eta(\beta)) = \xi_1 \tag{4}$$

and

$$\lim_{\beta\to 0} F(t;\alpha(\beta),\beta,\eta(\beta)) = \xi_1, \quad t > \theta. \tag{5}$$

Note that only one of the following two cases can occur: (i) $\theta < t_{n-1}$, or (ii) $\theta = t_{n-1}$.

Case (i): $\theta < t_{n-1}$. Let β_0 be an arbitrary but fixed point of $(0, B)$ such that

$$\theta - K(\beta)(\tau_1 - \theta) > t_{k-1} \geq 0.$$

Since $\theta > t_{k-1} \geq 0$ and $\lim_{\beta\to 0} K(\beta) = 0$, such β_0 exists. Then for all $\beta \in (0, \beta_0)$ we have

$$F(t_i;\alpha(\beta),\beta,\eta(\beta)) = 0, \quad t_i < \theta. \tag{6}$$

The function $t \mapsto F(t;\alpha(\beta),\beta,\eta(\beta))$, $\beta \in (0,\beta_0)$, is strictly increasing on the interval $(\theta - K(\beta)(\tau_1 - \theta), \infty)$. Due to this fact and (5), we may assume that $\beta_0 > 0$ is sufficiently small, so that for all $\beta \in (0, \beta_0)$,

$$\begin{aligned} &y_i < F(t_i;\alpha(\beta),\beta,\eta(\beta)) < \xi_1, && \text{if } \theta < t_i < \tau_1 \\ &y_i = F(t_i;\alpha(\beta),\beta,\eta(\beta)) = \xi_1, && \text{if } t_i = \tau_1 \\ &\xi_1 < F(t_i;\alpha(\beta),\beta,\eta(\beta)) < y_i, && \text{if } t_i > \tau_1. \end{aligned} \tag{7}$$

Since $\theta < t_{n-1}$, there are at least two indices i for which $t_i > \theta$. Furthermore, since the sequence $\{y_i\}_{i=1}^n$ is strictly increasing, the equality $y_i = \xi_1$ can hold for at most one index i. Hence, for every $\beta \in (0, \beta_0)$, it follows from (6) and (7) that for every $\beta \in (0, \beta_0)$,

$$\begin{aligned} S(\alpha(\beta),\beta,\eta(\beta)) &= \sum_{i=1}^n w_i \big[F(t_i;\alpha(\beta),\beta,\eta(\beta)) - y_i\big]^2 \\ &\leq \sum_{t_i<\theta} w_i \big[F(t_i;\alpha(\beta),\beta,\eta(\beta)) - y_i\big]^2 + \sum_{t_i>\theta} w_i \big[F(t_i;\alpha(\beta),\beta,\eta(\beta)) - y_i\big]^2 \\ &< \sum_{t_i<\theta} w_i y_i^2 + \sum_{t_i>\theta} w_i (y_i - \xi_1)^2 \\ &\leq \sum_{t_i<\theta} w_i y_i^2 + \sum_{t_i>\theta} w_i (y_i - A)^2 = \Sigma_{\theta,A}. \end{aligned}$$

The last inequality follows directly from a well-known fact that the quadratic function $x \mapsto \sum_{t_i>\theta} w_i(y_i - x)^2$ attains its minimum $\sum_{t_i>\theta} w_i(y_i - \xi_1)^2$ at point $\xi_1 = \sum_{t_i>\theta} w_i y_i / \sum_{t_i>\theta} w_i$.

Case (ii): $\theta = t_{n-1}$. First note that in this case it must be $\tau_1 = t_n$ and $\xi_1 = y_n$. Hence $\Sigma_{t_{n-1},A} = \sum_{i=1}^{n-2} w_i y_i^2$.

Let β_1 be a point of $(0, B)$ such that $t_{n-1} - K(\beta_1)(t_n - t_{n-1}) = t_{n-2}$. Then, since K is a strictly increasing function, for every $\beta \in (\beta_1, B)$ we have that

$$t_{n-1} - K(\beta)(t_n - t_{n-1}) < t_{n-2}.$$

Thus

$$F(t_{n-2};\alpha(\beta),\beta,\eta(\beta)) = \mathrm{e}^{-\ln\left(\frac{1}{y_n}\right)\left(\frac{t_n - t_{n-1} + K(\beta)(t_n - t_{n-1})}{t_{n-2} - t_{n-1} + K(\beta)(t_n - t_{n-1})}\right)^{\beta}}, \quad \beta \in (\beta_1, B),$$

from where it follows that $F(t_{n-2}; \alpha(\beta), \beta, \eta(\beta)) \to 0$ as $\beta \to \beta_1$ from the right. Then by definition of the limit there exists a point $\hat{\beta} \in (\beta_1, B)$ such that

$$0 < F(t_{n-2}; \alpha(\hat{\beta}), \beta, \eta(\hat{\beta})) < y_{n-2}. \tag{8}$$

Without loss of generality, we may suppose that $\hat{\beta}$ is sufficiently close to β_1, so that $0 \leq t_{n-3} < t_{n-1} - K(\hat{\beta})(t_n - t_{n-1}) < t_{n-2}$. Then

$$F\big(t_i; \alpha(\hat{\beta}), \beta, \eta(\hat{\beta})\big) = 0, \quad i = 1, \ldots, n-3. \tag{9}$$

As shown earlier (see (4)),

$$F\big(t_{n-1}; \alpha(\hat{\beta}), \hat{\beta}, \eta(\hat{\beta})\big) = y_{n-1} \quad \text{and} \quad F\big(t_n; \alpha(\hat{\beta}), \hat{\beta}, \eta(\hat{\beta})\big) = y_n. \tag{10}$$

From (8), (9) and (10) it follows that $S(\alpha(\hat{\beta}), \hat{\beta}, \eta(\hat{\beta})) < \sum_{i=1}^{n-1} w_i y_i^2$. This completes the proof of the lemma. $\square$

Now we state our main result (Theorem 1) which guarantees the existence of the LSE for the three-parameter IWD. This theorem is also applicable in a classical nonlinear regression problem with the model function of the form (1).

Theorem 1. *Let the data* (w_i, t_i, y_i), $i = 1, \ldots, n$, $n > 3$, *be such that*

(i) $0 < t_1 < t_2 < \ldots < t_n$,

(ii) $0 < y_1 < y_2 < \ldots < y_n < 1$

and $w_i > 0$, $i = 1, \ldots, n$. *Then, there exists the LSE for the three-parametric inverse Weibull distribution.*

Proof. 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

$$S^\star = \lim_{k\to\infty} S(\alpha_k, \beta_k, \eta_k) = \lim_{k\to\infty} \sum_{i=1}^{n} w_i \big[F(t_i; \alpha_k, \beta_k, \eta_k) - y_i\big]^2 \tag{11}$$

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 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 $\bar{\mathbb{R}}$, denote

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

Note that $0 \leq \alpha^\star, \beta^\star, \eta^\star \leq \infty$, 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 the 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 $\alpha^\star \neq \infty$. In step 2 we will show that $\eta^\star \neq 0$. The proof that $\eta^\star \neq \infty$ will be done in step 3. In step 4 we prove that $\beta^\star \neq \infty$. Finally, in step 5 we show that $\beta^\star \neq 0$.

Step 1. If $\alpha^\star = \infty$, then $F(t_i; \alpha_k, \beta_k, \eta_k) = 0$ for every sufficiently great $k \in \mathbb{N}$, and therefore from (11) it follows that $S^\star = \sum_{i=1}^n w_i y_i^2$. Since $\sum_{i=1}^n w_i y_i^2 > \Sigma_{t_{n-1}, y_n}$ and since according to Lemma 1 there exists a point in $\mathcal{P}$ at which functional S attains a value smaller than Σ_{t_{n-1}, y_n}, this means that in this way $(\alpha^\star = \infty)$ functional S cannot attain its infimum. Thus, we have proved that $\alpha^\star \neq \infty$.

Step 2. Let us show that $\eta^\star \neq 0$. We prove this by contradiction. Suppose on the contrary that $\eta^\star = 0$. Then only one of the following two cases can occur: (i) $\eta^\star = 0$ and $\beta^\star = 0$ or (ii) $\eta^\star = 0$ and $\beta^\star > 0$. 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 $\beta^\star = 0$. Since $\eta_k \to 0$, for every sufficiently great $k \in \mathbb{N}$, $0 < \eta_k^{\beta_k} < 1$. This means that sequence $(\eta_k^{\beta_k})$ is bounded. We may assume that it is convergent. Let $\eta_k^{\beta_k} \to L \in [0, 1]$. Since

$$\lim_{k\to\infty} \Big(\frac{\eta_k}{t - \alpha_k}\Big)^{\beta_k} = \lim_{k\to\infty} \eta_k^{\beta_k} (t - \alpha_k)^{\beta_k} = L, \quad t > \alpha^\star,$$

in this case we would have

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = \begin{cases} 0, & \text{if } t < \alpha^\star \\ \mathrm{e}^{-L}, & \text{if } t > \alpha^\star \end{cases} \tag{12}$$

and hence from (11) it would follow that

$$S^\star \ge \sum_{t_i < \alpha^\star} w_i y_i^2 + \sum_{t_i > \alpha} w_i (\mathrm{e}^{-L} - y_i)^2 = \Sigma_{\alpha^\star, \mathrm{e}^{-L}}.$$

According to Lemma 1, there exists a point in $\mathcal{P}$ at which functional S attains a value smaller than $\Sigma_{\alpha^\star, \mathrm{e}^{-L}}$. This means that in this case functional S cannot attain its infimum.

Case (ii): $\eta^\star = 0$ and $\beta^\star > 0$. In this case we would have

$$\lim_{k\to\infty} \Big(\frac{\eta_k}{t - \alpha_k}\Big)^{\beta_k} = 0, \quad t \ge \alpha^\star$$

and therefore

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = 1, \quad t \ge \alpha^\star.$$

Arguing now as in case (i), it can be shown that $S^\star \ge \Sigma_{\alpha^\star, 1}$. Again, according to Lemma 1, we conclude that in this case functional S cannot attain its infimum.

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

Step 3. Now, we are going to show that $\eta^\star \neq \infty$. We prove this by contradiction. Suppose $\eta^\star = \infty$. Then, without loss of generality, we may assume that $\eta_k > 1$ for all $k \in \mathbb{N}$. Since in that case $\eta_k^{\beta_k} \geq 1$, only one of the following can occur: (i) $\eta^\star = \infty$ and $\eta_k^{\beta_k} \to \infty$ or (ii) $\eta^\star = \infty$ and $\eta_k^{\beta_k} \to L \in [1, \infty)$. Let us show that functional S cannot attain its infimum in either of these two cases, which will prove that $\eta^\star \neq 0$.

Case (i): $\eta^\star = \infty$ and $\eta_k^{\beta_k} \to \infty$. In this case we have

$$\lim_{k\to\infty} \big(\frac{\eta_k}{t-\alpha_k}\big)^{\beta_k} = \infty, \quad t > \alpha^\star$$

and therefore

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = \lim_{k\to\infty} \mathrm{e}^{-(\frac{\eta_k}{t-\alpha_k})^{\beta_k}} = 0, \quad t > \alpha^\star.$$

Indeed, if $t > \alpha^\star$ and $\beta^\star = 0$, then

$$\lim_{k\to\infty} \big(\frac{\eta_k}{t-\alpha_k}\big)^{\beta_k} = \lim_{k\to\infty} \eta_k^{\beta_k}(t-\alpha_k)^{\beta_k} = \infty \cdot 1 = \infty.$$

If $t > \alpha^\star$ and $\beta^\star > 0$, then obviously $\lim_{k\to\infty} \big(\frac{\eta_k}{t-\alpha_k}\big)^{\beta_k} = \infty$.

Case (ii): $\eta^\star = \infty$ and $\eta_k^{\beta_k} \to L \in [1, \infty)$. In this case, sequence (β_k) must converge to 0, because by assumption $\eta_k \to \infty$. Therefore,

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = \begin{cases} 0, & \text{if } t < \alpha^\star \\ \mathrm{e}^{-L}, & \text{if } t > \alpha^\star. \end{cases}$$

Arguing now similarly to case (i) from step 2, in both cases it can be shown that in this way ($\eta^\star = 0$) functional T cannot attain its infimum.

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

Step 4. Let us show that $\beta^\star \neq \infty$. To see this, suppose on the contrary that $\beta^\star = \infty$. Then

$$\lim_{k\to\infty} \big(\frac{\eta_k}{t-\alpha_k}\big)^{\beta_k} = \begin{cases} \infty, \text{ if } \alpha^\star < t < \alpha^\star + \eta^\star \\ 0, \text{ if } t > \alpha^\star + \eta^\star \end{cases}$$

and therefore

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = \begin{cases} 0, \text{ if } \alpha^\star < t < \alpha^\star + \eta^\star \\ 1, \text{ if } t > \alpha^\star + \eta^\star. \end{cases}$$

Arguing now similarly to case (i) from step 2, it can be shown that $S^\star \geq \Sigma_{\alpha^\star+\eta^\star,1}$. By Lemma 1, there exists a point in $\mathcal{P}$ at which functional S attains a value smaller than $\Sigma_{\alpha^\star+\eta^\star,1}$. Therefore, in this way ($\beta^\star = \infty$) functional S cannot attain its infimum. Thus, we proved that $\beta^\star < \infty$.

Step 5. It remains to be shown that $\beta^\star \neq 0$. If $\beta_k \to 0$, then

$$\lim_{k\to\infty} \big(\frac{\eta_k}{t-\alpha_k}\big)^{\beta_k} = 1, \quad t > \alpha^\star$$

and therefore

$$\lim_{k\to\infty} F(t; \alpha_k, \beta_k, \eta_k) = \begin{cases} 0, & \text{if } t < \alpha^\star \\ \mathrm{e}^{-1}, & \text{if } t > \alpha^\star. \end{cases}$$

Arguing similarly to case (i) from step 2, it can be shown that $S^\star \geq \Sigma_{\alpha^\star, 1-\mathrm{e}^{-1}}$. According to Lemma 1, in this way functional S cannot attain its infimum. Thus, we proved that $\beta^\star > 0$ and herewith we completed the proof. $\square$

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

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

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 $S_p(\alpha_p^\star,\beta_p^\star,\eta_p^\star) = \inf_{(\alpha,\beta,\eta)\in\mathcal{P}} S_p(\alpha,\beta,\eta)$.

3. Numerical illustration

The three-parameter IWB has been extensively used in modelling failure times. For example, consider a real data set from Murthy et al. [17, p. 291] concerning failure of the photocopier cleaning web. The observed 14 failure times (in days) are displayed in Table 1, where t_i denotes the i^{th} failure time. Suppose that the failure time T is a random variable following an inverse Weibull distribution (1).

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
t_i	99	269	166	159	194	100	95	245	56	36	66	69	26	31
$t_{(i)}$	26	31	36	56	66	69	95	99	100	159	166	194	245	269

Table 1. *Failure times t_i and the order statistics $t_{(i)}$*

There are many different ways of computing the empirical cumulative distribution function $\hat{F}$ corresponding to the sample data $t_1,\ldots,t_n$. They all involve arranging the data in an ascending order so that $t_{(1)} < t_{(2)} < \ldots < t_{(n)}$, which is also shown in Table 1. Most commonly used estimators can be expressed in the following form (see Lawless [15] and Nelson [18]):

$$\hat{F}(t_{(i)}) = \frac{i-c}{n+1-2c} =: y_i, \qquad 0 \le c < 1. \tag{13}$$

Some alternatives are as follows: $y_i = \frac{i}{n+1}$ (*mean rank* estimator, $c=0$), $y_i = \frac{i-0.5}{n}$ (*median rank* estimator, $c=0.5$), $y_i = \frac{i-0.3}{n+0.4}$ (*Benard's median rank* estimator, $c=0.3$).

Our data for least squares estimation are $\big(w_i, t_{(i)}, y_i\big)$, $i = 1,\ldots,14$, where $w_i > 0$ are data weights, $t_{(i)}$ are the values from Table 1 and y_i are calculated by using (13). It is easy to verify that these data satisfy the conditions of Theorem 1, therefore an LSE exists.

Numerical methods for minimizing the sum of squares require an initial approximation $(\alpha_0,\beta_0,\eta_0) \in \mathcal{P}$, which needs to be as good as possible. We suggest to do this in the following way: For α_0 take $t_{(1)}/2$, and then calculate β_0 and η_0 by transforming Weibull distribution $F(t;\alpha_0,\beta,\eta)$ to the form

$$y = -\ln\big[-\ln F(t;\alpha_0,\beta,\eta)\big] = \beta\ln(t-\alpha_0) - \beta\ln\eta$$

and then fit a straight line y on $x = \ln(t-\alpha_0)$ using least squares. The above transformation was first proposed by Drapella [8]. A plot of y versus x is called the inverse Weibull probability paper (IWPP) plot.

In this numerical illustration we are going to use mean rank, median rank and Benard's approach. For all weights w_i we took 1. Our results obtained by using the Levenberg-Marquart method (see e.g. Dennis and Schnabel [7]) are given in Table 2.

Approach	Parameter estimates	Sum of squares (SS)
mean rank	$\alpha^\star = 0$, $\beta^\star = 1.11430$, $\eta^\star = 68.7775$	$SS = 0.0223877$
median rank	$\alpha^\star = 0$, $\beta^\star = 1.22757$, $\eta^\star = 71.0390$	$SS = 0.0249334$
Benard	$\alpha^\star = 0$, $\beta^\star = 1.17909$, $\eta^\star = 70.1115$	$SS = 0.0238659$

Table 2. *Least squares parameter estimates based on failure times*

Murthy et al. [17] used a mean rank estimator and fitted a two-parameter IWD to the data and obtained the following estimate of unknown parameters: $\hat{\beta} = 1.23$, $\hat{\eta} = 61.8$. The corresponding sum of squares $SS = 0.0528198$ is greater than the one we obtained. The reason for this is that they used the so-called Weibull probability paper (WPP) plot.

In Figure 2 we show the data $\big(t_{(i)}, \hat{F}(t_{(i)})\big)$, $i = 1, \ldots, 14$, calculated by using a mean rank estimator, the graph of function $F(t; \alpha^\star, \beta^\star, \eta^\star)$ and the graph of function $F(t; 0, \hat{\beta}, \hat{\eta})$.

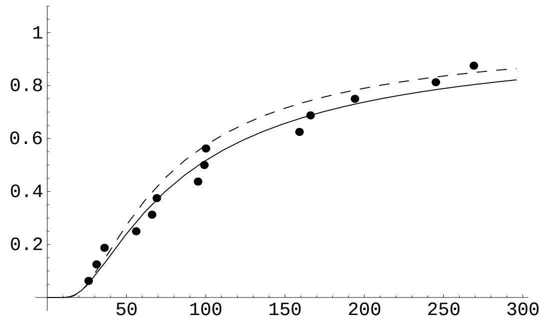

Figure 2. *The data,* ___ $F(t; \alpha^\star, \beta^\star, \eta^\star)$, - - - $F(t; 0, \hat{\beta}, \hat{\eta})$

References

[1] D. M. Bates, D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley, New York, 1988.

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

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

[4] 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.

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

[6] E. Z. Demidenko, *Is this the least squares estimate?* Biometrika **87**(2000), 437-452.

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

[8] A. Drapella, *Complementary Weibull distribution: Unknown or just forgotten*, Quality and Reliability Engineering International **9**(1993), 383-385.

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

[10] N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, Vol.2, second edition, John Wiley & Sons, New York, 1995.

[11] D. Jukić, R. Scitovski, M. Benšić, *On the existence of the nonlinear weighted least squares estimate for a three-parameter Weibull distribution*, Comput. Statis. Data Anal. **52**(2008), 4502-4511. doi:10.1016/j.csda.2008.03.001

[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, G. Kralik, R. Scitovski, *Least squares fitting Gompertz curve*, J. Comput. Appl. Math. **169**(2004), 359-375.

[14] 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.

[15] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 1982.

[16] G. S. Mudholkar, G. D. Kollia, *Generalized Weibull family: A structural analysis*, Communications in Statistics: Theory and Methods **23**(1994), 1149-1171.

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

[18] W. Nelson, *Applied life data analysis*, Wiley, New York, 1982.

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

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

[21] 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.