Journal of Computational and Applied Mathematics 228 (2009) 304–312

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Nonlinear weighted least squares estimation of a three-parameter Weibull density with a nonparametric start[☆]

Darija Marković, Dragan Jukić[*], Mirta Benšić

Department of Mathematics, University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia

ARTICLE INFO

Article history:
Received 6 June 2008

MSC:
65D10
62J02
62G07
62N05

Keywords:
Three-parameter Weibull density
Data fitting
Nonlinear least squares
Least squares estimate
Existence problem

ABSTRACT

This paper is concerned with the parameter estimation problem for the three-parameter Weibull density which is widely employed as a model in reliability and lifetime studies. Our approach is a combination of nonparametric and parametric methods. The basic idea is to start with an initial nonparametric density estimate which needs to be as good as possible, and then 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. Some simulations are given to show that our approach is satisfactory if the initial density is of good enough quality.

1. Introduction

A commonly used model in reliability and lifetime studies (see [6,15–17]) is the three-parameter Weibull distribution with a probability density function (PDF) given by

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

Here $\alpha \geq 0$, $\beta > 0$ and $\eta > 0$ are a location, a shape, and a scale parameter, respectively [24,25]. The corresponding cumulative distribution function (CDF) reads

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

This distribution was introduced by the Swedish statistician Waloddi Weibull who used it for the first time in 1939 in connection with his studies on the strength of materials (see [24]). If $\alpha = 0$, the resulting distribution is called the two-parameter Weibull distribution. A nonnegative random variable T is said to follow the three-parameter Weibull distribution if its PDF and CDF are given by (1) and (2), respectively. This correspondence is written as $T \sim W(\alpha, \beta, \eta)$.

The Weibull three-parameter distribution is very flexible and by an appropriate choice of the shape parameter β the density curve can assume a wide variety of shapes (see Fig. 1). If $\beta \in (0, 1]$, the density function is decreasing on (α, ∞). In

☆ This work is supported by the Ministry of Science, Education and Sports, Republic of Croatia, through research grants 235-2352818-1034 and 235-2352818-1039.

* Corresponding author. Tel.: +385 31 224 800; fax: +385 31 224 801.
E-mail addresses: darija@mathos.hr (D. Marković), jukicd@mathos.hr (D. Jukić), mirta@mathos.hr (M. Benšić).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.09.025

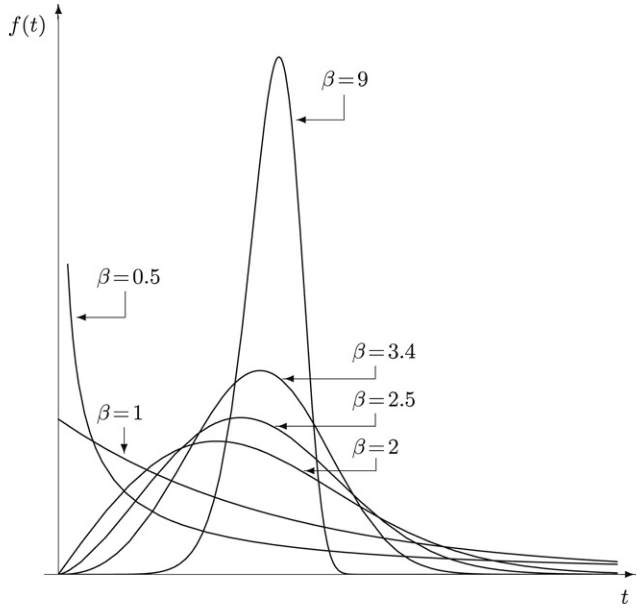

Fig. 1. Plots of the Weibull PDF for some values of β and by assuming $\alpha = 0$ and $\eta = 1.2$.

the case when $\beta > 1$, the density curve is bell-shaped with the maximum value at $\alpha + \eta(1 - 1/\beta)^{1/\beta}$. It is often proposed that the Weibull distribution with a value of β lying between 3 and 4 gives a fair approximation to the normal distribution; the most common value of β used to approximate the normal distribution is $\beta = 3.4$. When $\beta = 1$, the Weibull distribution becomes the two-parameter exponential distribution; when $\beta = 2$, it is identical to the Rayleigh distribution; when $\beta = 2.5$, it approximates the lognormal distribution. That is the reason why the Weibull distribution is one of the most widely used models in reliability and lifetime studies (see [6,15–17]). There are many applications of the Weibull distribution in medicine, biology, agriculture, forestry and engineering sciences. In addition, the Weibull distribution has found applications as a trend curve in many areas of applied research (see [2,19,20]). We refer the reader to the book in [16] for a review of different Weibull models and their applications in reliability.

In practice, the unknown parameters α, β and η of the appropriate Weibull density (distribution) are not known and must be estimated from a random sample $t_1, \ldots, t_n$ consisting of n observations of the random variable $T \sim W(\alpha, \beta, \eta)$. 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. In parametric density estimation, the parameters are estimated by maximum likelihood (ML), least squares (LS), or other methods. ML is a traditional method since it possesses beneficial properties such as asymptotic normality and consistency. For the two-parameter Weibull distribution, the results of Pratt [18] and Burridge [5] guarantee the existence of a unique ML estimate. However, for the three–parameter Weibull distribution the likelihood function is unbounded from above so that a standard ML estimate does not exist (see e.g. [11,15]). Some of the existing results regarding parameter estimation of the three-parameter Weibull distribution are based on finding a local maximum of the likelihood function, if it exists. However, it is shown that the asymptotic behavior of this estimator is completely different for $\beta > 2$ and $\beta \leq 2$ (see e.g. [23]) and that even a local extreme does not always exist. Such difficulties with the ML approach contributed to the usage of other methods to determine the unknown parameters in the three-parameter Weibull model (see e.g. [1,11,16,22,23]).

Different nonparametric density estimation approaches are available, including histograms, kernel estimates, nearest neighbor estimates, and orthogonal series estimates, among others (see e.g. [21]).

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 least squares (LS) fit procedure to estimate the unknown parameters α, β and η. To be more precise, suppose we are given a random sample

$$0 < t_1 < t_2 < \cdots < t_n$$

consisting of n observations of the random variable $T \sim W(\alpha, \beta, \eta)$. Let $\hat{f}$ be a nonparametric density estimate constructed by using this sample. Our data for LS estimation are (w_i, t_i, y_i), $i = 1, \ldots, n$, where $y_i = \hat{f}(t_i)$ 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 density function (1) have to be estimated by minimizing the functional

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