Applied Mathematics and Computation 247 (2014) 803–814

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Kernel density estimation of three-parameter Weibull distribution with neural network and genetic algorithm

Fan Yang *, Zhufeng Yue

Department of Engineering Mechanics, Northwestern Polytechnical University, Xi'an, PR China

ARTICLE INFO

Keywords:
Neural network model
Genetic algorithm
Weibull distribution
Optimization algorithm
Maximum likelihood method
Grey model

ABSTRACT

Three-parameter Weibull distribution is widely employed as a model in reliability and lifetime studies due to its good fit to data. It is important to estimate the unknown parameters exactly for modeling. There are many methods to estimate the parameters of three-parameter Weibull distribution and the kernel density estimation method is one of them. The smoothing parameter has a significant influence on the estimation accuracy. In this paper, the neural network and genetic algorithm were used to get the best smoothing parameter and the result was compared with other methods. The Monte Carlo simulations were carried out to show the feasibility of our approach for estimation of three-parameter Weibull distribution.

1. Introduction

Weibull distribution was first introduced by the Swedish statistician Waloddi Weibull [1] who used it in his studies on the strength of materials. Because it is extremely flexible in fitting random data and adaptable in different shape distribution, a wide range of applications can be described as Weibull model, such as data processing in biology [2], medicine [3], weather [4], economics [5] even the large-scale analysis [6–10] if possible. In reliability analysis, Weibull distribution is one of the most statistical distributions. The problem of reliability is widespread in aircraft reliability engineering. Single crystal superalloy cooled blade is one of the most critical components in advanced aero-engine. It operates at high temperature, high pressure and high rotation conditions. The strength, life and reliability of cooled blade determine the life and reliability of aero-engine directly, and even whole aircraft. According to engineers' experiences and fatigue life data, it indicates that the fatigue life data obeys Weibull distribution. Once the certain known continues distribution function of life data, the statistical methods can be taken to estimate the fatigue life under conditions of confidence and reliability.

The three-parameter Weibull distribution [11] shows better fitting precision than two-parameter Weibull distribution. But to estimate three-parameter Weibull density is a more complicated and difficult task. When its parameters take different values, three-parameter Weibull distribution can be equal or approximate to some common distributions such as exponential distribution, Raleigh distribution and normal distribution. The parameter estimation means to obtain the location, shape and scale parameter, so the statistical properties of experimental data carried can be better illustrated by the three parameter determined distribution.

The result of parameter estimation directly affects the accuracy of the reliability analysis. And the selection of parameter estimation methods directly affects the results of the estimation. Therefore, it is important to take exact parameter estimation

* Corresponding author.
E-mail address: yangfan3124@gmail.com (F. Yang).

http://dx.doi.org/10.1016/j.amc.2014.09.065
0096-3003/

of the probability density function of three-parameter Weibull distribution. But it is complicated to do it, and many researches have been done [12].

So far, there are two main methods for parameter estimation [13–15], namely, graphic methods [16–18] and statistic methods [19–25]. The graphic parameter estimation method is easy to understand and use. But this kind of method relies on observation of data curves. The drawing needs strict quality requirements and the estimation results have large deviation compared with statistic methods.

Statistic methods include histograms [19] moment estimation [20], maximum likelihood estimation (ML) method [21–23], kernel method [24,25] etc. Moment estimation needs to calculate the sample standard deviation, mean value, coefficient of skewness. Then, it needs to use the available table to take the point estimates and use given shape parameter to reduce order of likelihood equations and obtain the three-parameter estimation by method of bisectors. The precision of this method is not good because the explicit expression is difficult to obtain for the existence Gamma function in expectation and variance.

Regression analysis methods [26–30] can handle linear problems. It is adopted in two-parameter estimation of Weibull distribution. Maximum likelihood (ML) method is of high precision, but it is complicated to solve three transcendental equations simultaneity and requires suitable initial values. It brings multiple solutions or zero sometimes. Meanwhile, the simulated algorithm (SA) was adopted to get the three parameters of Weibull distribution [31]. It is a method to search the minimum value of likelihood function using SA. Some scholars use bilinear regression estimation and maximum likelihood estimation method to solve three-parameter of Weibull distribution. Probability-weighted moment method (PWM) [12] is suitable for the probabilistic model which inverse of distribution function is explicit formulation. But this method is not suitable for small sample data and the calculation of correlation coefficient is painstaking.

Kernel density estimation is nonparametric estimation method which has been widely used in parameter estimation. In this paper, we proposed a new method based on the kernel density estimation using artificial neural network (ANN) and genetic algorithm (GA). Meanwhile, the result of the new method is better than other methods (grey model, Rank regression method (RRM), ML, kernel method).

2. Theory and calculation

The three-parameter Weibull distribution with a probability density function (PDF) is given by

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

where $\alpha \geqslant 0,\ \beta > 0$ and $n > 0$ are a location, a shape and a scale parameter respectively. The corresponding cumulative distribution function (CDF) is

$$F(t,\alpha,\beta,\eta)=\begin{cases}1-e^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}}, & t>\alpha\\ 0, & t\leqslant\alpha\end{cases} \tag{2}$$

First, the Rank regression method (RRM) estimation is adopted to get the initial value of location, scale and shape parameter. Then, kernel estimation method is adopted for further research. Meanwhile grey model method for location parameter estimation is taken to compare with the other methods.

2.1. Rank regression method (RRM)

Let $t_1, t_2, \ldots, t_n$ is a group of random sequence satisfies a three-parameter Weibull distribution, they are sorted in ascending order, namely, $0 < t_1 < t_2 < \cdots < t_n$. The cumulative distribution function of Weibull,

$$F(t,\alpha,\beta,\eta)=1-e^{-\left(\frac{t-\alpha}{\eta}\right)^{\beta}} \tag{3}$$

Transformed,

$$y=\ln\left\{\ln\left[\frac{1}{1-F(t,\alpha,\beta,\eta)}\right]\right\}=\beta\ln(t-\alpha)-\beta\ln\eta, \tag{4}$$

The location parameter can be obtained by the following expression [32].

$$\alpha_0=t_1-1/n. \tag{5}$$

The procedures to estimate the scale and shape parameters are as follows,

Step 1: For every sample, $t_1, \cdots, t_n$, calculate the cumulative frequency, $\widehat{F}$

$$\widehat{F}(t_i)=\frac{i}{n+1},\quad i=1,\ldots,n. \tag{6}$$