Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 48-3 (2015) 219–224

Bootstrap and Maximum Entropy Based Small-sample Product Lifetime Probability Distribution

Hailong Suo,, Jianmin Gao, Zhiyong Gao,Hongqvan Jiang,Rongxi Wang
Xi' an Jiaotong University, Xi' an ,Shaanxi,China
State Key Laboratory for Manufacturing System Engineering
(email:hl_suo@163.com)

ABSTRACT: How to build the small-sample data distribution is a major issue to be addressed for analyzing large power equipment lifetime. To find a better product lifetime probability distribution function, a new method based on bootstrap and maximum entropy is proposed to study on the problem of small-sample product life probability distribution. The result shows that this new method is more accurate and feasible compared with bootstrap and maximum entropy respectively. Finally, an actual example of small-sample product lifetime distribution shows that this new approach has better guiding significance for practice.

Keywords: bootstrap; maximum entropy; nonlinear programming; small sample; probability density function

1. INTRODUCTION

Spare parts inventory management plays an important supporting role in guaranteeing stable operation of equipment. In the process of building a spare parts inventory model, different empirical distribution functions are chosen to analyze and predict depending on the product life data. However, failure lifetime data shows small-sample characteristics, and several distribution hypotheses may be validated by the method of Kolmogolov-Smirnov, therefore, this traditional approach is invalid and not feasible. In this case, to study on the personalized and reasonable product lifetime probability density function is necessary according to the actual lifetime data. The purpose is to provide scientific and objective guidance or advices to inventory management.

Srinivasan, V.S.(Srinivasan, V.S.2004)describes a theoretical method of estimating the most general form of probability density functions of random variables, and the estimated probability density function depends on the mean value of exponential distribution of the initial random variable. Lv Zhen(Lv Zhen et al. 2008) presents a method for determining fatigue life distribution under the assumption that fatigue life obeys three-parameter Weibull distribution, and an experimental data with 200 specimens is employed to demonstrate the good agreement with the large samples. Radu Florescu(Radu Florescu et al. 2006)exposes an original technique to determine the empirical probability density function and the empirical cumulative distribution function, and to estimate moments of any order for a small size random sample (m=3 - 10) of a continuous random variable.

In summary, there are two main methods for the present study of the small-sample probability distribution problem：
1) Suppose the probability distribution and verify; 2) Build a probability distribution function by oneself. For the first method, the error is often large for the small-sample problem of product lifetime probability distribution; for the second method, practical application is more difficult because it is based on the particular sample data. In this paper, a new method based on bootstrap and maximum entropy is proposed to study on the problem of small-sample product life probability distribution. It aims to make an objective and reliable analysis of unknown small-sample product life probability distribution for large power equipment.

10.1016/j.ifacol.2015.06.084

2. BOOTSTRAP METHOD

2.1 Methods and Principles

Bootstrap method is proposed as a new statistical analysis by professor B.Efron in 1979, and in the next years it is being developed. Gray bootstrap method (GBM)(Wang Yanqing et al. 2014) is proposed to solve the problem of estimating frequency-varying RVS with small samples for environment testing. Ouysse, Rachida(Ouysse, Rachida. 2013)describes a fast approximation for the small sample bias correction of the iterated bootstrap. Its approximation adapts existing fast approximation techniques of the bootstrap p-value and quantile functions to the problem of estimating the bias function. The characteristics of the bootstrap is: there is no need to make any assumptions to overall distribution, and it can be inferred with the sample data by the computer technology.

The main idea of bootstrap is:

1) $X_i\,(i=1,2,3,\cdots,n)$ are independent identically distributed samples, $X_i \sim F(X)$,sample distribution function F_n is built by X_i ;

2) Random sampling with replacement from F_n ,generating N bootstrap samples, $X^*_{(1)}, X^*_{(2)}, \cdots, X^*_{(N)}$

$$X^*_{(i)} = (x^*_{i1}, x^*_{i2}, \cdots, x^*_{im})\ (i=1,2,3,\cdots,N), \tag{1}$$

m is the bootstrap sample size;

3) For every $X^*_{(i)}$, estimate value of Statistics $\hat{\theta}$ should be calculated

$$\hat{\theta}^*(i) = S(X^{*i}),\ i=1,2,3,\cdots,N \tag{2}$$

4) The estimated standard error $\hat{se}_N$ from bootstrap is the standard deviation of $\hat{\theta}^*(i)$

$$\hat{se}_N = \left(\sum_{i=1}^{N} \frac{[\hat{\theta}^*(i) - \hat{\theta}^*(\bullet)]}{N-1}\right)^{\frac{1}{2}} \text{ Here, } \hat{\theta}^*(\bullet) = \sum_{i=1}^{N} \frac{\hat{\theta}^*(i)}{N} \tag{3}$$

2.2 Analysis of Examples

To verify the accuracy of the probability density function by the bootstrap method, three subsamples(sample size:15、30、60) are in turn selected from the overall sample of standard normal distribution, and results of 1000 re-sampling process by the bootstrap method are analyzed and compared with standard normal probability density function. The results are as follows:

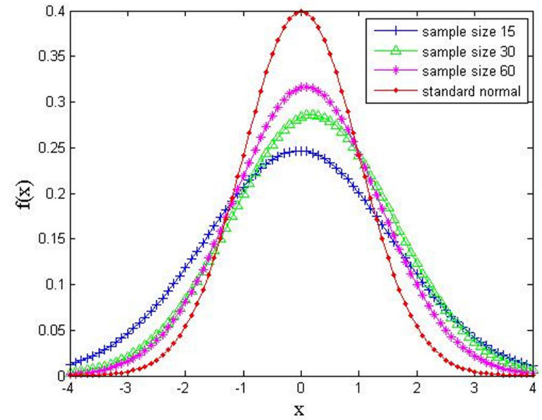

Figure 1 The probability density functions of different sample sizes based on bootstrap

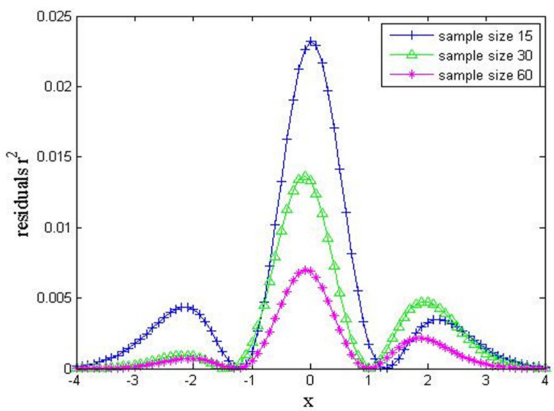

Figure 2 The probability density residuals of different sample sizes based on bootstrap

Figure 1 and figure 2 show: for the construction of standard normal probability density function with bootstrap method, the accuracy of the constructed sample density function has a positive correlation with sample size. However, the disadvantage of this approach is that whether the statistics is choiced in an appropriate way will affect the analysis for overall distribution of the samples.

3. MAXIMUM ENTROPY

3.1 Methods and Principles

The maximum entropy method is a reasoning view proposed by E.T.Jaynes. To infer the system state in the case of only grasping partial information, the legitimate state should be the state where entropy is the largest under the constraints, and it is the only impartial choice. In the case of only