Journal of Polish CIMAC
Gdansk University of Technology
The Faculty of Ocean Engineering and Ship Technology

STATISTICAL ANALYSIS OF FAILURES

Leszek Knopik

University of Technology and Life Science in Bydgoszcz
Faculty of Management Science
Fordonska 424 st, 85-789 Bydgoszcz, Poland
e-mail: knopikl@utp.edu.pl

Abstract

The lifetime distribution is important in reliability studies. There are many situations in lifetime testing, where an item (technical object) fails instantaneously and hence the observed lifetime is reported as a small real positive number. The investigation in this article was motivated by an extended and generalized Weibull distribution. We suggest a mixture of a singular distribution and Weibull distribution. We apply the maximum likelihood method to estimate the parameters of the mixture. The method is illustrated by a numerical example for the time between the failures of bus engines.

Keywords: *Weibull distribution, mixture of distributions, maximum likelihood method, confidence interval, early failures*

1. Introduction

In reliability and quality control, the probability density and failure rate function provides the valuable information about the distribution of the failure times. The important thing in the field of the lifetime analysis is the most appropriate lifetime distribution that describes the time to a failure of a component, assembly or system.
Occurrence of instantaneous or early failures in lifetime testing is observed in sets of failures of machines. These occurrences may be due to faulty constructions or inferior quality. Some failures result from natural damages of the machine while the other failures may be caused by inefficient repairs of previous failures resulting from incorrect organization of the repairs. These situations can be modeled by modifying commonly used parametric models such as exponential, gamma and Weibull distributions.
In the papers [9], [10] the set of failures of a machine is divided into two subsets, namely into the set of primary failures and the set of secondary failures. This division suggests that the population of lifetime is heterogeneous. The population of time before failures can be described by using the statistical concept of mixture. This mixture, in a particular case, has the unimodal failure rate function. In the paper [9], special attention has been paid to determination of shape of the failure rate function from mixture of the exponential distribution and distribution with linear increasing failure rate function. It is clear that instantaneous failures can be primary failures or secondary failures. In this paper, the mixture of a singular and Weibull distributions is considered.
In this paper, we show that the mixture of a singular and Weibull distributions is useful to describe the lifetime of the machines. The numerical example is also provided to illustrate the practical

impact of this approach. In this example n = 193 failures of the bus engine is studied. This example shows that in this case a mixture of the singular distribution and the exponential distribution is sufficient.

2. The model of lifetime distribution

We consider the continuous time failure cumulative distribution function F(t, a, b) with F(0, a, b) = 0. In this paper we assume that F is two parameter Weibull distribution with parameters (a, b). The distribution function is the following:

$$F(t, a, b) = 1 - \exp(- x^{b} / a) \qquad \text{for } x \geq 0 \tag{1}$$

and the density function

$$f(t, a, b) = (b/a)\, t^{b-1} \exp(-t^{b})$$

We assume that the early failures are recorded as a class with notational failure time c so that the modified distribution has the density function given by

$$g(t, p, a, b) = \begin{cases} 0 & \text{for} \quad x < c \\ 1 - p + pF(c, a, b) & \text{for} \quad x = c \\ pf(t, a, b) & \text{for} \quad x > c \end{cases} \tag{2}$$

The purpose of this paper is to consider the probability density function given by (2), when F is two parameters Weibull distribution. The problem of statistical inference about (Θ, p) has received considerable attention particularly when T is exponential. Some of the early references are: Aitchison [1], Kleyle and Dahiya [4], Jayade and Parasad [2], Muralidharan [5, 6], Kale and Muralidharan [3] and the references contained therein. Muralidharan and Kale [7] considered the case where F is a two parameters gamma distribution with shape parameter β and scale parameter α, and they obtained confidence interval for δ = pαβ assuming α as being known and unknown respectively.

3. The maximum likelihood estimation

In the paper [8] these distributions are considered and the maximum likelihood estimation of parameters (p, a, b) is obtained. Here, we give the equations for the maximum likelihood estimate for (p, a, b).

$$\widehat{p} = \frac{n - n_0}{n} \exp(c^{\hat{b}} / \widehat{a}), \tag{3}$$

where n_0 is the number of observations reordered as c.

$$\widehat{a} = \frac{1}{n - n_0} \sum_{x_i > 0} x_i^{\hat{b}} - c^{\hat{b}} \tag{4}$$

and $\widehat{b}$ is the solution of the equation

$$\frac{n_1 c^b \ln c - s_3}{s_1 - n_1 c^b} + \frac{1}{b} + \frac{s_2}{n_1} = 0 \qquad (5)$$

where

$$n_1 = n - n_0$$

$$s_1 = \sum_{x_i > 0} x_i^b$$

$$s_2 = \sum_{x_i > 0} \ln x_i$$

$$s_3 = \sum_{x_i > 0} x_i^b \ln x_i$$

The Fisher information matrix is given by

$$I(p, a, b) = \begin{bmatrix} I_{pp} & I_{pa} & I_{pb} \\ I_{ap} & I_{aa} & I_{ab} \\ I_{bp} & I_{ba} & I_{bb} \end{bmatrix} \qquad (6)$$

where

$$I_{pp} = A/(p\ B)$$

$$I_{pa} = I_{ap} = (c^b A)/(a^2 B)$$

$$I_{pb} = I_{bp} = -\ c^b A/(aB^2)$$

$$I_{aa} = (p\ A/a^2)\ (c^{2b}/(a^2 B)+1)$$

$$I_{ab} = I_{ba} = (-\ p/a)(c^{2b} \ln(c)\ A/(a^2(\ 1- B) + A(\ln(c) + 1/b - I_1/b))$$

$$I_{bb} = p\ A\ ((c^b \ln(c))^2 / (a^2 B) + 1/b^2 + \ln(c)/b + (\ln(c))^2) + (p/b^2)\ (I_1 + I_2),$$

where

$$A = \exp(-\ c^b / a)$$

$$B = 1 - \exp(-c^b / a)$$

$$I_1 = \int_{c^b}^{\infty} y^{-1} \exp(-y/a)dy = - \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} (\frac{c^b}{a})^j$$

$$I_2 = \int_{c^b}^{\infty} y^{-1} \ln y \exp(-y/a)dy = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} (\frac{c^b}{a})^j \left[\frac{1}{j} - b \ln c\right]$$

By $I^{-1}(p, a, b)$ we denote the inverse of the matrix $I(p, a, b)$.
Let

$$K(p, a, b) = \frac{1}{n} I^{-1}(p, a, b)$$

Using the matrix K(p, a, b) we estimate the confidence interval for (p, a, b). The approximate $1 - \alpha$ confidence interval for p is the following:

$$(\widehat{p} - u_\alpha \sqrt{K_{pp}}, \widehat{p} + u_\alpha \sqrt{K_{pp}})$$

for a

$$(\widehat{a} - u_\alpha \sqrt{K_{aa}}, \widehat{a} + u_\alpha \sqrt{K_{aa}})$$

for b

$$(\widehat{b} - u_\alpha \sqrt{K_{bb}}, \widehat{b} + u_\alpha \sqrt{K_{bb}})$$

where K_{pp}, K_{aa}, K_{bb} are the diagonal elements of the invert matrix $K(p, a, b)$, and u_α is the value of standard Gauss random variable U such that

$$P\{|U| < u_\alpha\} = 1 - \alpha .$$

4. Numerical example

The object of the investigation is a real municipal bus transport system within large agglomeration. The analyzed system operates and maintains 190 municipal buses of various types and makes. In this section, we consider a real time data on failure of bus. The data set contains n = 193 times between failures of a bus. This set contains $n_0 = 50$ times equal to one and we assume that c=1.

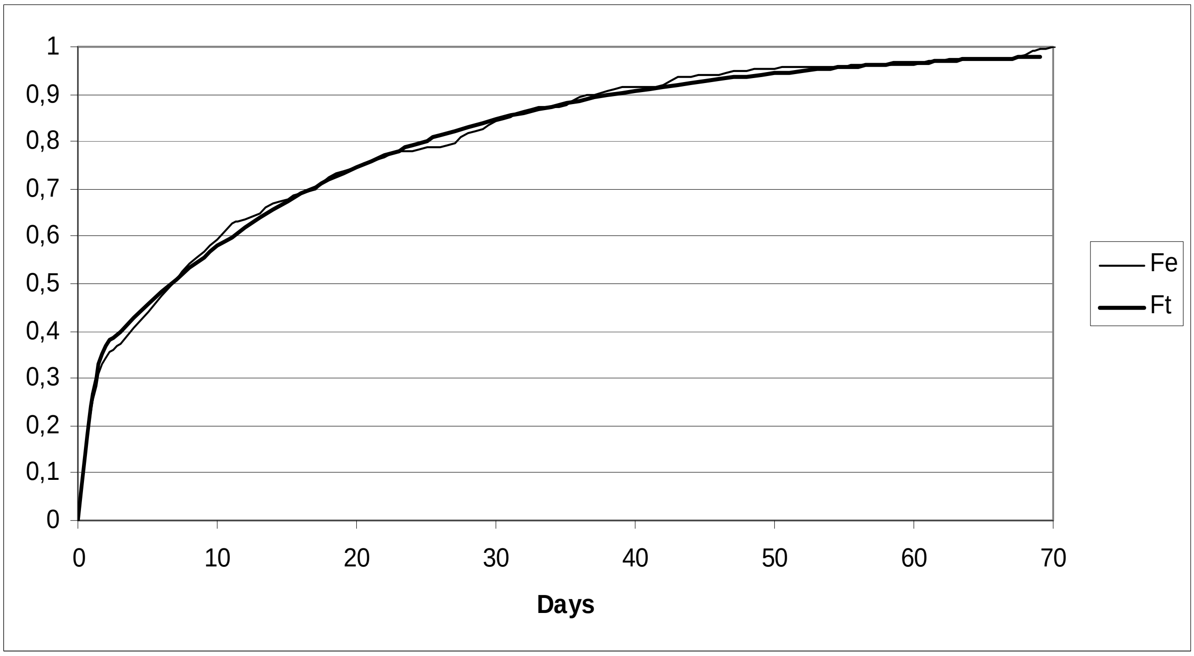

Fig. 1. Plots of empirical and mixture distribution functions

We apply the maximum likelihood estimates of the parameters p, a and b. As the initial solution of the equation (5), we give b = 1.We calculate the values of the parameters p = 0.26, b = 1.02 and a = 17.4, and the corresponding confidence interval, for p we obtain (0.23, 0.29), for b (0.94, 1.2) and a (10.2, 11.2). For these values of parameters, we prove the Pearson's test of fit and compute the associated p-value = 0.54. It shows a good conformity of the empirical data with the mixture distributions. Fig. 1 shows the empirical distribution function (Fe) and mixture distribution function (Ft) for this example.

5. Conclusions

In this paper, we study the lifetime model for instantaneous and early failures of bus engines. The

estimation of parameters is approached by the method of maximum likelihood and the expected information matrix is derived. The estimate for b can be obtained as the solution of equation (5). The confidence intervals for p, a and b depend on the information matrix
I(p, a, b). Furthermore, these confidence intervals can be used for large sample. When the parameters are estimated, it is possible to apply for further calculations, such as MTTF (Mean Time to Failure), burn-in time, the failure rate function and the replacement time. The development of efficient parameters estimation methods for this mixture distribution and their application for times to failure modeling are topics for further study. An application to a real data set shows that this model may be applicable in practice.

References

1. Aitchison I.: (1955) On the distribution of a positive random variable having a discrete probability mass at origin, Journal of the American Statistical Associations, vol. 50, pp. 901-908.
2. Jayade V. P. and Parasad M. S.: (1996) Estimations of parameters of mixed failure time distribution. Communications Statistics, Theory and Method, vol19, pp. 4667-4677.
3. Kale B. K. and Muralidharan K.: (2000) Optimal estimating equations in mixture distributions accommodating instantaneous or early failures, Journal Indian Statistical Associations, vol. 38, pp.317-329.
4. Kleyle R.M. and Dahiya R.L.: (1975) Estimation of parameters of mixed failure time distribution from censored data, Communications Statistics, Theory and Method, vol.4, pp. 873-882.
5. Muralidharan K.: (1999) Test for mixing proportion in mixture of a degenerate and exponential distributions, Journal Indian Statistical Associations, vol. 37, pp. 105-119.
6. Muralidharan K.: (2000) The UMVUE and Bayes estimate of reliability of mixed failure time distribution, Communications Statistics, Simulation Computer, vol. 29, No 2, pp. 603-158.
7. Muralidharan K. and Kale B.K.: (2002) Modified gamma distribution with singularity at zero, Communications Statistics, Simulation Computer, vol. 31, No1, pp.143-158.
8. Muralidharan K. and Lathika P.: (2006) Analysis of instantaneous and early failures in Wiebull distribution, Metrika vol. 64, pp.305-316.
9. Knopik L.: (2010) Mixture of distributions as a lifetime distribution of a technical object, Scientific Problems of Machines Operation and Maintenance, 2010, vol.45, 2(165), pp. 53-60.
10. Wdzięczny A., Woropay M. and Muślewski Ł.: Division and investigation of damages to technical object and their influence on the reliability of operation and maintenance systems Scientific Problems of Machines Operation and Maintenance, 2008, vol.44, 2(154), 2008, pp. 31-43.