First Seminar on
Reliability Theory and its Applications
27-28, May 2015

Nonparametric and Parametric Estimation of Survival Function

Mireh, S. [1] and Khodadadi, A. [2]

Department of Statistics, University of Shahid Beheshti

Abstract

This paper considers a general degradation path model and failure time data with traumatic failure mode. It provides a review of the nonparametric estimator of survival function, studied by Bagdonavicius, and considers the parametric estimation of survival function of failure times with a hazard rate in the degradation space. In addition, we discuss the comparison of both parametric and nonparametric methods according to simulated and real data.

Keywords: Degradation models, Failure times, Hazard rate, Nonparametric and parametric estimation, Survival function.

1 Introduction

Analyzing survival data is historically based on $(T_1, \ldots, T_n)$ each measuring an individual time to event. It is difficult to assess reliability with traditional life tests that record only time to the failure. In some cases, degradation is measured directly by passage of time. Thus, it is necessary to define a level of degradation at which a failure is said to have occurred. We define soft and hard failures in terms of a specified level of degradation and traumatic failures.

Usually, one attempts to conditionally define the hazard rate such as Bagdonavicus[3] that define $\lambda(t|A) = \lambda_0(t) \times \lambda(g(t, A))$ where g is a given non-decreasing function.

Statistical analysis of linear degradation and multiple failure modes using nonparametric method are discussed by Bagdonavicuset al.[1]. They have presented reliability characteristics using a semiparametric method[2]. In this work, we estimate the survival

[1] s_mireh@sbu.ac.ir
[2] khodadadi@sbu.ac.ir

function from the degradation and failure time data using parametric and nonparametric methods.

This paper is organized as follows. Section 2 defines the joint models for degradation and failure time. Section 3 is devoted to a review on the estimation of survival function. Section 4 deals with the performance of the two methods through a set of real data. In section 5, a simulation study is performed for the comparison of both methods usingin various sample sizes.

2 Joint models for degradation and failure time

Assume that the degradation of an item is given by stochastic process $Z(t)$. We denote the true degradation path of particular unit by $g(t)$, but the observed degradation processes is a degradation path plus error: $Z(t) = g(t, A) + e$, where A is a vector of unknown parameters. Bagdonavicus[1] has studied a linear degradation model with multiple failure modes.

Suppose the life time T^0 is the first time of crossing a ultimate threshold z_0 for $Z(t)$. If we denote h for the inverse function of g and h' for its partial derivative then: $T^0 = h(z_0, A)$. Let T^1 be traumatic failure time. Thus the moment of the observed failure is: $T = \min(T^0, T^1)$.

Suppose the random variable T^1 has the intensity $\lambda^{(1)}(z)$ and the cumulative intensity $\Lambda^{(1)}(z)$, depending on the degradation level. The conditional survival function of T^1 given A is:

$$S^{(1)}(t|A) = \exp\Big\{ -\int_0^t \lambda^{(1)}(g(y,a))dy \Big\} = \exp\Big\{ -\int_0^{g(t,a)} h'(z,a)d\Lambda^{(1)}(z) \Big\}.$$

We can obtain the survival function of the random variable T:

$$S(t) = \int_{g(t,a)<z_0} \exp\Big\{ -\int_0^{g(t,a)} h'(z,a)d\Lambda^{(1)}(z) \Big\} d\pi(a) \tag{1}$$

where π is the distribution function of A.

3 Estimation of the reliability functions

Suppose the data are collected from n unit: $(T_1, Z_1, \delta_1), \ldots, (T_n, Z_n, \delta_n)$ where T_i is the failure time, Z_i is the degradation level and δ_i is the indicator of the failure modes. In the parametric method, we set a distribution on A and the parameters are estimated using MLE. However in the nonparametric method, the estimators are given by the following: The estimation of the distribution function and the cumulative hazard function:

$$\hat{\pi}(a) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{A_i \le a\}} \quad , \quad \hat{\Lambda}(z) = \sum_{Z_i \le z, \delta_i = 1} \frac{1}{\sum_{j, Z_j \le Z_i} h'(Z_j, A_i)}$$

4 Estimation by real data

The real data are the wear and failure time data of 79 bus tires. The critical tire wear value is $z_0 = 15mm$. Set $g(T, A) = T/A$. We have used the Exponentiated Weibull family

Table 1: Maximum Likelihood Estimators: Exponentiated Weibull distribution and intensity functions

Parameter	Estimation
β	1.4836
σ	1.4772
θ	130.439
α_1	0.0425
ν_1	6.4332

as the parametric family of π and let $\lambda^{(1)}(z) = (\alpha_1 z)^{\nu_1}$. The MLE of the parameters are summarized in Table 1.

We obtain a parametric estimation of survival function by substituting parameter estimates in (1). In addition, we calculate the nonparametric estimation. Figure 1 gives graphs of empirical cdf of A and the estimators of $S(t)$.

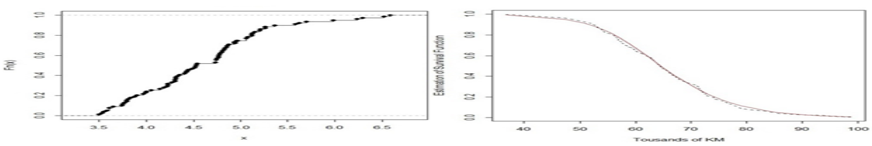

Figure 1: (Left) Empirical cdf of A; (Right) parametric(solid line) and nonparametric(dotted line) estimators of $S(t)$

5 Simulation study

Example 1. *In this example, we compare the parametric and nonparametric estimations by using small, moderate, and large sample sizes. We have generated vector A from the Weibull distribution with parameters (5, 2) and set* $z_0 = 10$.

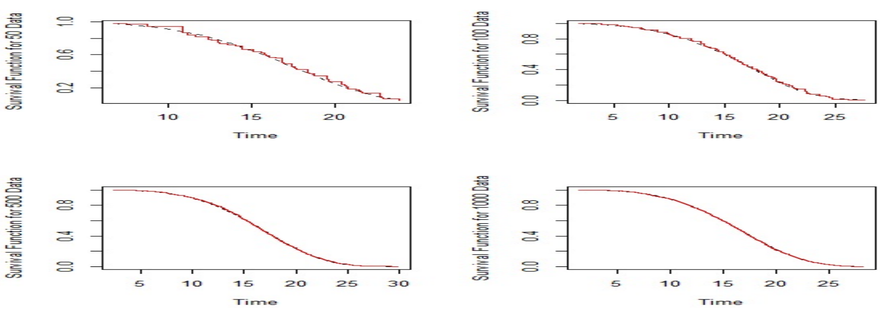

Figure 2: Parametric(dotted line) and nonparametric(solid line) estimators of $S(t)$ in different sample sizes

Example 2. *We consider simulations of n=100 degradation curves* $Z(t, \theta_1, \theta_2) = e^{\theta_1}(1 + t)^{\theta_2}, t \in [0, 12]$ *with a hazard rate in the degradation space of Weibull-type(*$\alpha = 5, \beta = 2.5$*) and* $A = (\theta_1, \theta_2)$ *is a Gaussian vector with mean (-2,2) and* $Var\theta_1 = Var\theta_2 = 0.1^2$.

Figure 3 shows the distribution function of θ_2 *and the nonparametric estimation of the cumulative hazard rate.*

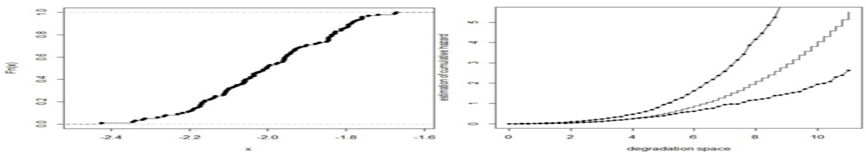

Figure 3: (Left) Empirical cdf of θ_1; (Right) Nonparametric hazard rate and 95% confidence band

References

[1] Bagdonavicus, V., Bikelis, A. and Kasakevicius, V. (2004), Statistical Analysis of Linear Degradation and Failure Time Data with Multiple Failure Modes. *Lifetime Data Analysis*, **10**, 65-81.

[2] Bagdonavicus, V., Haghighi, F. and Nikulin, M. (2005), Statistical Analysis of Degradation and Failure Time Data using the General Path Model. *Communication in Statistics-Theory and Methods*, **34**, 1771-1793.

[3] Bagdonavicius, V. and Nikulin, M. (2004), *Semiparametric Analysis of Degradation and Failure Times Data with Covariates*, Birkauser.