Journal of Multivariate Analysis 101 (2010) 340–351

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Wiener processes with random effects for degradation data

Xiao Wang

Department of Mathematics and Statistics, University of Maryland, 1000 Hilltop Circle, Baltimore, MD 21250, USA

ARTICLE INFO

Article history:
Available online 24 December 2008

AMS subject classifications:
62G20
62M05
62N05

Keywords:
Bootstrap
Degradation
EM algorithm
Empirical processes
Random effects
Reliability
Wiener process

ABSTRACT

This article studies the maximum likelihood inference on a class of Wiener processes with random effects for degradation data. Degradation data are special case of functional data with monotone trend. The setting for degradation data is one on which n independent subjects, each with a Wiener process with random drift and diffusion parameters, are observed at possible different times. Unit-to-unit variability is incorporated into the model by these random effects. EM algorithm is used to obtain the maximum likelihood estimators of the unknown parameters. Asymptotic properties such as consistency and convergence rate are established. Bootstrap method is used for assessing the uncertainties of the estimators. Simulations are used to validate the method. The model is fitted to bridge beam data and corresponding goodness-of-fit tests are carried out. Failure time distributions in terms of degradation level passages are calculated and illustrated.

1. Introduction

In some studies where the subjects are put on test at time zero, these subjects degrade over time. Usually continuous observation of degradation for these subjects is not possible. The degradation of each subject may be observed at each of several times, where the number of observation times and observation times themselves are allowed to vary across the subjects. Data of this type are called *degradation data*. Degradation data are special case of functional data with monotone trend. Degradation data have applications in many fields, such as HIV study and industrial reliability. For example, it is suggested that the immune system of a person infected with the HIV virus degrades over time [1]. CD4 counts constitute a critical assessment of the status of the immune system and CD4 counts are commonly used markers for the health status of HIV infected persons. In reliability area, the data from fatigue crack growth subject to loading cycles [2,3] and bridge beams subject to erosion of chloride ion ingression [4] are typical degradation data. Degradation data are a very rich source of survival information. In many tests, the failure time data are supplemented by degradation data and degradation data offer many advantages over failure time data. For general discussion of degradation models, see [5–7].

Wiener processes and extensions to them have been used as models for degradation data (e.g. [8–11]). Let $\Lambda(t)$ be a nondecreasing function. The nonhomogeneous Wiener process $Y(t)$ has independent increments $\Delta Y(t) = Y(t+\Delta t) - Y(t)$, where $\Delta Y(t)$ has a normal distribution with mean $\Delta\Lambda(t) = \Lambda(t+\Delta t) - \Lambda(t)$ and variance $\sigma^2\Delta\Lambda(t)$. If letting $U(t) = t + \sigma W(t)$ be the Wiener process with drift t and diffusion σ, then

$$Y(t) = U(\Lambda(t)) = \Lambda(t) + \sigma W(\Lambda(t)) \tag{1}$$

is just a time-transformed Wiener process. When the amount of degradation reaches a pre-specified critical level D, failure occurs. Let T denote the failure time, then $T = \inf\{t : Y(t) \geq D\}$. The level-crossing of the cumulative degradation threshold D by a nonhomogeneous Wiener process $Y(t)$ can be obtained in terms of the inverse Gaussian (IG) distribution [12,13,31] as

E-mail address: wangxiao@umbc.edu.

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2008.12.007

$T_\Lambda = \Lambda^{-1}(T_{IG})$, where T_{IG} is inverse Gaussian (IG) random variable $IG(D, D^2/\sigma^2)$ with mean D and variance σ^2/D. Note the similarities between nonhomogeneous Wiener process model and hazard rate model. The failure time in hazard rate model can be obtained by $T_\Lambda = \Lambda^{-1}(T_{\exp})$ where Λ is the integrated hazard rate and $T_{\exp}$ is the standard exponential distribution. Detailed nonparametric inference of this model is discussed by [14].

In most degradation applications, there is substantial subject-to-subject variability among the degradation processes of different individuals. The unit-specific random effects can be incorporated into the process to represent such heterogeneity in the degradation paths. One such model can be specified by allowing both drift and diffusion of the above Wiener process model to be random. Mathematically tractable distribution results if we adopt a model as follows:

$$\text{Given } (\nu, \sigma), \quad Y(t) = \nu\Lambda(t) + \sigma W(\Lambda(t)), \tag{2}$$

$$w = \sigma^{-2} \sim \mathrm{Gam}(r^{-1}, \delta), \qquad \nu | w \sim N(1, \theta/w). \tag{3}$$

Here, w has mean δ/r and variance δ/r^2 and thus σ^2 has finite mean $r/(\delta-1)$ for $\delta > 1$ and finite variance $r^2/((\delta-1)^2(\delta-2))$ for $\delta > 2$. Letting the conditional mean of ν be 1 is for model identification since any constant can be incorporated into function $\Lambda(t)$. An alternative parameterization that is often used in such situations is to let $r = \delta$ in the distribution of w. There is no physical meaning for choosing the random effect distributions in (3). The idea of making these choices of the distributions (3) for random effects is borrowed from Bayesian linear regression [15] for computational convenience. Semiparametric maximum likelihood method is developed to estimate the unknown parameter $(\Lambda(t), \delta, r, \theta)$ of the underlying process. Here, $\Lambda(t)$ is estimated nonparametrically, which leads naturally to an infinite dimensional statistical problem. We show that the maximum likelihood estimator (MLE) is consistent and we also derive the convergence rate of the MLE. Bootstrap method is used for assessing the uncertainty of the MLE. Simulation results suggest that this method works well and we apply our method to the degradation data of a civil engineering structure to estimate its reliability.

The remainder of the paper is as follows. Section 2 introduces the Wiener process model with random effects. Section 3 establishes the consistency and convergence rate of the maximum likelihood estimator and uses bootstrap method to assess the uncertainties of the estimators. Section 4 uses EM algorithm to compute the MLE. Section 5 presents Monte Carlo studies to validate the methods and we also fit the Wiener process model to bridge beam data in Section 6. Section 7 makes some concluding remarks.

2. Wiener process with random effects: *t* process

Model (2) shows that the conditional distribution of $Y(t)$ given w and ν is normal, and the marginal density of $Y(t)$ follows as

$$\begin{aligned} f(y) &= \int_{-\infty}^{\infty}\int_0^{\infty} f(y; \nu, w) g_1(\nu; \theta, w) g_2(w; r, \delta)\,\mathrm{d}w\,\mathrm{d}\nu \\ &= \frac{\Gamma(\delta + \frac{1}{2})}{\sqrt{2\pi r}\,\Gamma(\delta)[\Lambda^2\theta + \Lambda]^{1/2}} \left[1 + \frac{(y-\Lambda)^2}{2r(\Lambda^2\theta + \Lambda)}\right]^{-\delta-\frac{1}{2}}. \end{aligned} \tag{4}$$

Note that $\sqrt{\frac{\delta}{r(\Lambda^2\theta+\Lambda)}}(Y(t) - \Lambda(t))$ has a t distribution with degrees of freedom 2δ. Thus, $Y(t)$ has finite mean $\Lambda(t)$ and finite variance

$$\mathrm{Var}(Y(t)) = [\Lambda(t)^2\theta + \Lambda(t)]\frac{r}{\delta - 1}, \quad \text{for } \delta > 1.$$

An extreme situation of the above model is when the variances of the random effects are zero and it becomes the general Wiener process model (1). This situation can be realized by letting $\theta \to 0$ and $r \to \infty$ with $\delta/r = c$ fixed. Conditionally on their common random effects, ν and w, the level-crossing of the cumulative degradation threshold D by the process $Y(t)$ in (2) follows IG distribution $IG(D/\nu, D^2 w)$. Hence, $P(T \le t) = E_{\nu,w} G(\Lambda(t); \nu, w)$, where $G(t, \nu, w)$ is the inverse Gaussian distribution function with parameters $(D/\nu, D^2 w)$. If the degradation paths are monotonic, the failure time distribution has explicit form and is given by

$$P(T \le t) = P(Y(t) > d) = F_{2\delta}\left[\sqrt{\frac{\delta}{r}}\frac{\Lambda(t) - d}{\sqrt{\theta\Lambda(t)^2 + \Lambda(t)}}\right], \tag{5}$$

where $F_{2\delta}$ is the t distribution function with degrees of freedom 2δ.

Suppose that we observe $Y(t)$ for a subject at times $t_1, \ldots, t_m$, yielding observations $Y_1, \ldots, Y_m$. Let $y_j = Y_j - Y_{j-1}$ and $\lambda_j = \Lambda(t_j) - \Lambda(t_{j-1})$ with $Y_0 = 0$ and $\Lambda(t_0) = 0$. Given the common random effects ν and w, the increments y_j are independently normally distributed. Thus the joint density of the y_j can be obtained as

$$\begin{aligned} f(y_1, \ldots, y_m) &= \int_{-\infty}^{\infty}\int_0^{\infty} \phi(y_1, \ldots, y_m; \nu, w) g_1(\nu; \theta, w) g_2(w; r, \delta)\,\mathrm{d}w\,\mathrm{d}\nu \\ &= \frac{\Gamma(\delta + \frac{m}{2})}{\Gamma(\delta)(\sqrt{2\pi r})^{m/2}|A|^{1/2}}\left[1 + \frac{1}{2r}(y - \lambda)' A^{-1}(y - \lambda)\right]^{-\delta-\frac{m}{2}}, \end{aligned} \tag{6}$$