Analyzing Accelerated Degradation Data By Nonparametric Regression

Jyh-Jen Horng Shiau
National Chiao Tung University, Hsinchu

Hsin-Hua Lin
National Chiao Tung University, Hsinchu

Key Words — Accelerated degradation test, Acceleration factor, Accelerated life-stress degradation model, Local linear regression smoother, Nonparametric regression, Stochastic process.

Summary & Conclusions — This paper presents a nonparametric regression accelerated life-stress (NPRALS) model for accelerated degradation data wherein the data consist of groups of degrading curve data. In contrast to the usual parametric modeling, a nonparametric regression model relaxes assumptions on the form of the regression functions and lets data speak for themselves in searching for a suitable model for data. NPRALS assumes that various stress levels affect only the degradation rate, but not the shape of the degradation curve. An algorithm is presented for estimating the components of NPRALS. By investigating the relationship between the acceleration factors and the stress levels, the mean time to failure estimate of the product under the usual use condition is obtained. The procedure is applied to a set of data obtained from an accelerated degradation test for a light emitting diode product. The results look very promising. The performance of NPRALS is further checked by a simulated example and found satisfactory. We anticipate that NPRALS can be applied to other applications as well.

1. INTRODUCTION

Acronyms[1]

LED	light emitting diode
ADT	accelerated degradation test
LLR	local linear regression
PC	principal component
PCA	PC analysis
PCBF	PC basis function
MTTF	mean time to failure
SAFT	scale accelerated failure time
SL	stress level
SLI	standardized light intensity
NPRALS	the model in this paper

In many experiments in life sciences and engineering, the collected data are samples of response curves. Growth data of children is an example, and degradation data of a product-performance measure is another. How to analyze such data has become important in recent statistical research.

Due to rapid advances of technology and to quality improvement efforts, many products are so reliable that traditional life tests are not feasible in estimating the lifetimes of these products. In such circumstances, accelerated tests are widely used to shorten the life of products or hasten the degradation of their performance. The aim of such tests is to collect data quickly so that desirable information on product life or on performance under usual use can be obtained in a reasonable time by appropriate modeling and analysis.

This study focuses on accelerated degradation tests (ADT) data analysis. Most of the ADT analyses use parametric regression models to estimate the lifetime of the product under usual use. To relax the assumptions on the form of regression functions and let data speak for themselves in searching for a suitable model for data, we propose a nonparametric regression model to analyze ADT data in this paper.

Nonparametric regression techniques are useful in obtaining a smooth fit to noisy data, to describe the relationship between response variables and *s*-independent variables. These smoothing techniques are powerful tools in statistical data-analysis because of the,

- model flexibility,
- appealing look of the fitted curves (or surfaces).

Several books were published in recent years on smoothing techniques, *eg*, [2 – 6, 13].

1.1 Relevant Works

Recently, many researchers have paid much attention to modeling & analyzing degradation data. Nelson [11] provided a fairly thorough survey on ADT, which included areas of applications, statistical models, and data analyses. Meeker & Escobar [8] reviewed recent research in accelerated testing. Lu & Meeker [7] used a non-linear mixed-effects model with degradation data to estimate the life distribution. Meeker, Escobar, Lu [10] presented methods for analysis of accelerated degradation data. They used approximate maximum likelihood estimation to estimate model parameters from the assumed mixed-effects

[1] The singular & plural of an acronym are always spelled the same.

nonlinear regression model; and they suggested methods for estimating lifetime distributions for various situations. Meeker & Escobar [9] presented many up-to-date statistical methods for analyzing reliability data.

In life sciences, Capra & Müller [1] proposed an accelerated time model for cohorts of female Mediterranean fruit flies that may age faster or slower, depending on inherent genetic dispositions or external factors like varying temperature and humidity. A nonparametric regression method was used to estimate the mean mortality function under usual conditions.

For more research on this topic, see [1, 7 – 11].

1.2 The Motivated Application

Our research was motivated by a set of ADT data on LED. LED have been widely used in many fields, ranging from consumer electronics to optical fiber transmission systems. LED has many nice features, such as less power consumption, small volume, good visual effect, and long lifetime. Applications include electronic boards on highways and on streets, smoke sensors on ceilings, night lights, and traffic lights. LED products are gradually replacing traditional light bulbs in many places. Because of the very-high-reliability of LED products, it is difficult to obtain information on product life under usual use in a short time.

Yu & Tseng [14] proposed an on-line procedure for terminating an ADT experiment. They applied the procedure on a set of ADT data of an LED product with electric current as the accelerating variable. This study analyzes another set of ADT data of the same LED product with temperature as the accelerating variable. Our goal is to estimate the MTTF of the product under usual use. Section 4 describes this LED data set in more detail.

1.3 Overview

Section 2 introduces the NPRALS model. Section 3 proposes an algorithm for estimating the components of this model. Section 4 performs a data analysis on a set of ADT data of an LED product based on NPRALS. Section 5 has a simulation study to explore the effectiveness of NPRALS and reports the results. Section 6 discusses the resulting situation.

Notation

m	number of SL
i	index for SL, $i = 1, \ldots, m$
n_i	number of test items for SL i
j	index for test items, $j = 1, \ldots, n_i$
p	number of measurements for each test item
k	index for the measurements, $k = 1, \ldots, p$
t_k	measurement time for measurement k of each test item
$X_{i,j,k}$	measurement for item j under SL i at t_k
$t_{i,j,k}$	measurement time for $X_{i,j,k}$; $t_{i,j,k} = t_k$ for each i, j
T	a closed interval of time
$\hat{}$	implies an estimate
a_i	acceleration factor (or relative acceleration factor after scaling) for SL i
$t'_{i,j,k}$	rescaled time for $t_{i,j,k}$ according to $\hat{a}_i$
$X'_{i,j,k}$	measurement corresponding to $t'_{i,j,k}$; $X'_{i,j,k} = X_{i,j,k}$
$X(\cdot)$	stochastic process of quality characteristic, $\{X(t), t \in T\}$
$\mu(\cdot)$	mean function of $X(\cdot)$
$W(\cdot)$	stochastic process with mean zero, and $r(\cdot,\cdot)$
$r(\cdot,\cdot)$	covariance of $W(\cdot)$
$\epsilon(\cdot)$	measurement/experimental error
q	index for PCBF; $q = 1, \ldots, L$
L	number of PCBF
$\rho_{i,q}$	PCBF q of $X(\cdot)$ for SL i
$\epsilon_{i,j,q}$	random coefficient of $\rho_{i,q}$ for curve j
V_i	sample covariance matrix for SL i
M_i	MTTF under SL i
T_i	absolute temperature for SL i
(α_0, α_1)	unknown parameters of the Arrhenius relationship between M_i and T_i
(β_0, β_1)	unknown parameters of the Arrhenius relationship between a_i and T_i
$\tilde{X}_{i,j,k}$	simulated data of $X_{i,j,k}$
$K(\cdot)$	kernel function in the LLR model
h	bandwidth
$K_h(\cdot)$	$\frac{1}{h} \cdot K\left(\frac{\cdot}{h}\right)$
(b_0, b_1)	coefficients of the LLR model
w_i	weight of observation i in LLR smoothing
$S_{n,l}$	$\sum_{i=1}^{n} K_h(t_i - t) \cdot (t_i - t)^l, \quad l = 1, 2$
$X_{i,k}$	sum of n_i measurements at t_k at SL i
$\bar{X}_{i,k}$	average of n_i measurements at t_k at SL i

2. AN ACCELERATED LIFE-STRESS DEGRADATION MODEL

NPRALS has 2 important aspects:

- degradation path of the product characteristic (modeled as a stochastic-process).[2]
- relationship between the rate of acceleration and the SL of the product characteristic.

2.1 Stochastic-Process Model for Curve Data

Assumptions

1. The sample degradation path of each experimental subject (or test item) is a realization of an underlying $\{X(t), t \in T\}$, $T > 0$.
2. The model for $X(t)$ is:

$$X(t) = \mu(t) + W(t) + \epsilon(t), \tag{1}$$

$\mu(t) \equiv \mathrm{E}[X(t)]$;[3]

[2] In recent years, this approach has been used to model curve data in other applications. For more information, see [12] and references therein.

[3] Mean degradation curve of the product.

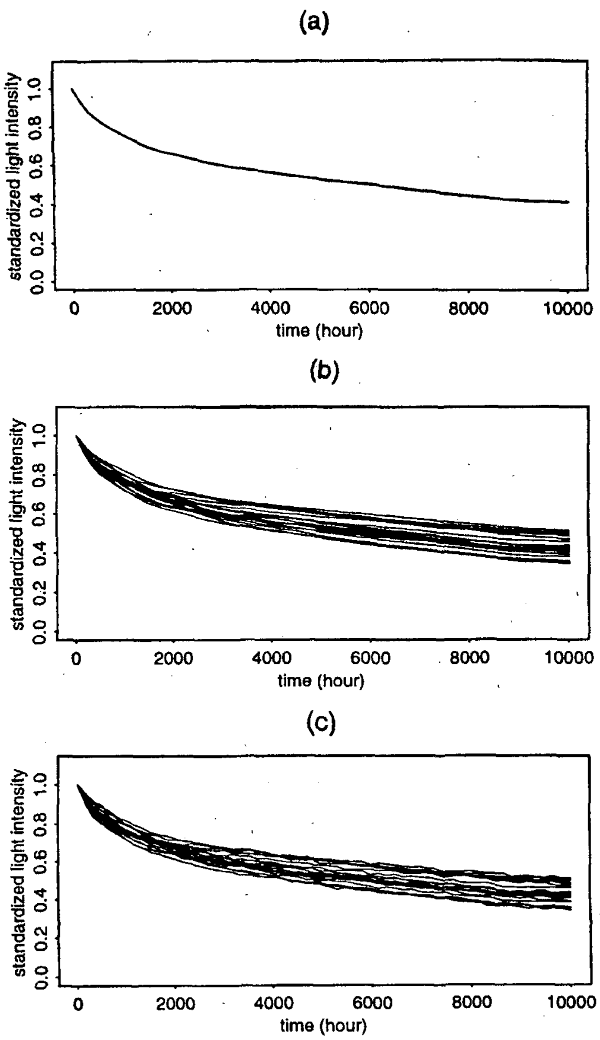

Figure 1: a. $\mu(t)$, b. $\mu(t) + W(t)$, c. $\mu(t) + W(t) + \epsilon(t)$

$W(t) \equiv$ a stochastic process with mean 0;[4]

$r(s,t) \equiv \text{Cov}[W(s), W(t)]$;

$\epsilon(t) \equiv$ uncorrelated error terms with $\text{E}[\epsilon(t)] = 0$ and $\text{Var}[\epsilon(t)] = \sigma^2$.[5] ◀

NPRALS is illustrated by a simulated example in figure 1. Figure 1a shows $\mu(t)$; figure 1b shows a group of sample paths of $\mu(t) + W(t)$; figure 1c shows the same group of sample paths as in figure 1b, but with experimental/measurement errors added.

2.2 An Acceleration Model and Acceleration Factor

With time & cost limitations, failures of highly reliable products are not always observable. Hence, ADT is a useful technique to speed up the degradation process.

[4]The subject effect.

[5]Experimental and/or measurement error at time t.

Assumptions

3. The acceleration stress affects only the degradation rate, not the shape of the degradation curve. ◀

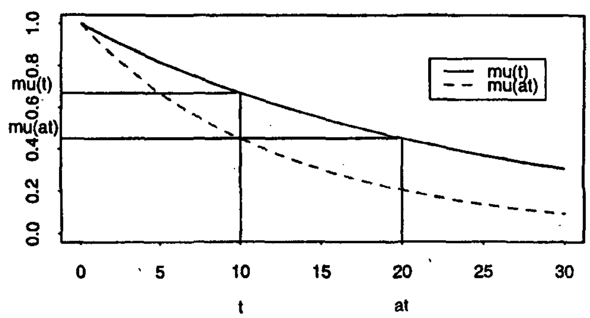

Figure 2: $\mu(t)$ (solid) and $\mu(a \cdot t)$ (dashed), for some $a > 1$

Figure 2 illustrates assumption #3. The solid curve is $\mu(t)$. The dashed curve is $\mu(a \cdot t)$, where a is the acceleration factor. The usual baseline (unaccelerated) process for defining the acceleration factor is the usual-use process; thus usually $a > 1$ for an accelerated process. The same value of the quality characteristic is observed at,

- t for the accelerated curve,
- $t' = a \cdot t$ for the unaccelerated curve.

Thus the lifetime — defined as the time the degradation path crosses a specified degradation level — of the accelerated-test item is the lifetime of the product under usual use divided by a. Consequently, investigation of the relationship between the SL and the acceleration factors, results in the mean lifetime of highly reliable products under usual use. This model is called the scale accelerated failure time (SAFT) model in [10].

The form of the acceleration time in assumption #3 is a special kind of acceleration model that might not be adequate to describe the effect of the accelerating variable on time for some problems. See, for example, [9: chapter 18] for other models.

2.3 An Accelerated Life-Stress Degradation Model

There are m SL in the experiments. For SL i, n_i experimental units are used. The product characteristic is measured at $t_1, \ldots, t_p$ for each experimental unit. Using the functional PC technique, consider the accelerated life-stress degradation model in (2) to describe data:

$$X_{i,j,k} = \mu(a_i \cdot t_k) + \sum_{q=1}^{L} \epsilon_{i,j,q} \cdot \rho_{i,q}(t_k), \tag{2}$$

$$1 \le i \le m, \ 1 \le j \le n_i, \ 1 \le k \le p, \ 1 \le q \le L.$$

Thus the random component of (1), $W(t)+\epsilon(t)$, is modeled by a random combination of functional PC, $\{\rho_{i,q}(\cdot), \ q = 1, \ldots, L\}$.

3. ESTIMATION OF THE MODEL COMPONENTS

This section presents an algorithm for estimating $\mu(\cdot)$ and $a_1, \ldots, a_m$. These a_i are not estimable since we do not have data from the unaccelerated process (usual use). We choose the group with the largest acceleration factor as the baseline process. The reason for this choice is that we then can use all the data points in the analysis in the pooling step of the algorithm. Then without loss of generality, let $0 \le a_1 \le \ldots \le a_m = 1$. After scaling, the a_i are no longer acceleration factors. In this section, they are called 'relative acceleration factors'.

Estimate a_i, for each $i = 1, \ldots, m$, by solving the minimization problem:

$$\min_{a_i} \left[\sum_{j=1}^{n_i} \sum_{k=1}^{p} [X_{i,j,k} - \hat{\mu}_m(a_i \cdot t_k)]^2 \right]. \tag{3}$$

$\hat{\mu}_m(a_i \cdot t)$ is an estimate of $\mu_i(t)$, for $i = 1, \ldots, m$. Estimates of $\mu_m(\cdot)$ and $\{a_i;\ i = 1, \ldots, m\}$ are obtained as follows.

1. Smooth
$\{X_{m,j,k},\ j = 1, \ldots, n_m,\ k = 1, \ldots, p\}$ on
$\{t_{m,j,k}, j = 1, \ldots, n_m,\ k = 1, \ldots, p\}$ to obtain an initial estimate of $\mu_m(\cdot)$. This step can be done more efficiently by smoothing
$\{\bar{X}_{m,k},\ k = 1, \ldots, p\}$ on
$\{t_{m,j,k},\ k = 1, \ldots, p\}$,
$$\bar{X}_{m,k} \equiv \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} X_{m,j,k}.$$
The derivation is given in the appendix.

2. To estimate a_i, search over (0,1) to find the minimizer of (3); this estimate is $\hat{a}_i$.

3. Since all $\mu_i(\cdot)$ are just time-scaled functions of $\mu_m(\cdot)$, pool all the data points in m groups together to estimate $\mu_m(\cdot)$. The data are compiled by mapping each data point
$(t_{i,j,k},\ X_{i,j,k})$ to $(t'_{i,j,k},\ X'_{i,j,k})$,
$$t'_{i,j,k} \equiv t_k \cdot (\frac{\hat{a}_i}{\hat{a}_m}),$$
$$X'_{i,j,k} = X_{i,j,k}.$$

4. Then obtain $\hat{\mu}_m(\cdot)$ by smoothing
$\{X'_{i,j,k},\ i = 1, \ldots, m,\ j = 1, \ldots, n_i,\ k = 1, \ldots, p\}$ on
$\{t'_{i,j,k},\ i = 1, \ldots, m,\ j = 1, \ldots, n_i,\ k = 1, \ldots, p\}$.
Figure 3 shows the result of the compilation of all data points for an example.

5. Solve (3) to obtain a new set of $\hat{a}'_i$. Repeat steps 2 – 4 until values of $\hat{a}_i$ converge. ◄

Next, investigate the covariance structure of $X(\cdot)$. Compute:

$$(V_i)_{s,r} = \frac{1}{n_i - 1} \cdot \sum_{j=1}^{n_i} [X_{i,j,s} - \hat{\mu}_i(t_s)] \cdot [X_{i,j,r} - \hat{\mu}_i(t_r)],$$

for $1 \le k, s, r \le p$.

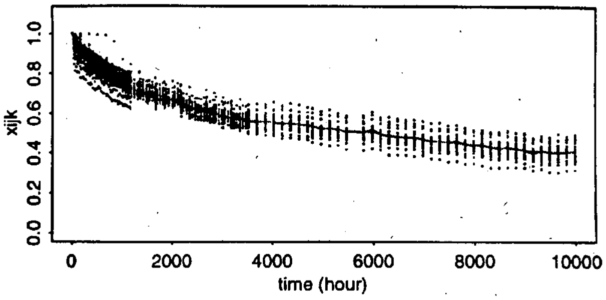

Figure 3: Result of Compiling All Data into Group 3

Let

$$\hat{\rho}_{i,q} \equiv (\hat{\rho}_{i,q}(t_1), \ldots, \hat{\rho}_{i,q}(t_p))^T,$$
for $i = 1, \ldots, m$

be the eigenvector corresponding to the q^{th} largest eigenvalue $\lambda_{i,q}$ of V_i, for $q = 1, \ldots, p$.

Estimate:

- $\sigma^2_{i,q}$ by $\lambda_{i,q}$,
- $\rho_{i,q}(\cdot)$ by smoothing $\{\hat{\rho}_{i,q}(t_k),\ k = 1, \ldots, p\}$ on $\{t_k, k = 1, \ldots, p\}$.

Remarks

1. In practice there is no need to estimate a_m since a_m is known to be 1. The $\hat{a}_m$ is computed simply to check whether the proposed estimating procedure works. An $\hat{a}_m$ too far off from the true value 1 indicates that the estimates obtained should not be trusted.

2. The smoothing technique in this paper is the LLR smoothing for ease of implementation; it is briefly described in the appendix. Other techniques such as smoothing splines can be used as well. The bandwidth selection in smoothing is important in practice. We presume that any reasonable method applies here.

3. Smoothing $\{X'_{i,j,k}\}$ on $\{t'_{i,j,k}\}$ directly, might be slow because there are many data points in the compiled data. The appendix derives an efficient way to reduce the computing time by a large factor. ◄

This estimation procedure is summarized in the Algorithm.

Algorithm.

1. Obtain an initial estimate $\hat{\mu}_m$ of μ_m by smoothing $\{\bar{X}_{m,k}, k = 1, \ldots, p\}$ on $\{t_k, k = 1, \ldots, p\}$.
2. For each $i = 1, \ldots, m$, obtain initial estimates $\hat{a}_i$ by minimizing (3).
3. Compile m groups of data by time-scaling according to $\hat{a}_i$:
$$X'_{i,j,k} = X_{i,j,k} \quad \text{and} \quad t'_{i,j,k} = \frac{\hat{a}_i}{\hat{a}_m} \cdot t_{i,j,k}.$$
4. Obtain $\hat{\mu}_m$ by smoothing $\{X'_{i,j,k}\}$ on $\{t'_{i,j,k}\}$.
5. Repeat steps 2 – 4 until all $\hat{a}_i$ converge.

End_Algorithm

4. REAL-DATA ANALYSIS

The Algorithm is demonstrated by analyzing a set of real ADT data of an LED product. Items of an LED product were randomly selected for this ADT experiment. The experiment was conducted up to 9998 hours for each SL. The key quality characteristic is the SLI which degrades over time. This quality characteristic was measured at 59 time points for each test item. Thus a response curve was collected for each test item.

The accelerating variable is temperature. Three levels of temperature, 25°C, 65°C, 105°C, were chosen by engineers. The usual use condition is 20°C.

The goal is to estimate the MTTF of the product at the usual use condition. The 16, 18, 19 response curves were collected for the accelerated conditions, 25°C, 65°C, 105°C, respectively. Figure 4 shows the data. The step-by-step data analysis is shown.

1. The fastest degrading group is group #3 in figure 4c. Let $a_3 = 1$. Apply the Algorithm:
$\hat{a}_1 = 0.116726$, $\hat{a}_2 = 0.353575$, $\hat{a}_3 = 0.999995$.
$\hat{\mu}_3(t)$.

2. Estimate the mean curves of group #1 by

$$\hat{\mu}_1(t) = \hat{\mu}_3 \cdot \left(\frac{\hat{a}_1}{\hat{a}_3} \cdot t\right)$$

and that of group #2 by

$$\hat{\mu}_2(t) = \hat{\mu}_3 \cdot \left(\frac{\hat{a}_2}{\hat{a}_3} \cdot t\right).$$

To see how well the estimation is, compare $\hat{\mu}_i$, $i = 1, 2, 3$, with the corresponding averaged curve obtained by averaging n_i curve data at each of the 59 time points for each group. The results are in figure 5.

Figure 5 shows that: $\hat{\mu}_i$ fits the averaged-curve i quite well.

3. Let a_0 be the relative acceleration factor under the usual use condition (20°C). To estimate a_0, consider finding a regression relationship between $\{a_i\}$ and the corresponding SL $\{T_i\}$. According to [11], consider the following Arrhenius rate relationship:

$$\log(M_i) = \alpha_0 + \alpha_1 \cdot \frac{1}{T_i}. \tag{4}$$

The Arrhenius relationship [9, 11], obtained through empirical observation[6], is widely used to describe the effect that temperature has on acceleration. Of course, the constant activation-energy does not apply to all temperature-acceleration problems; it is adequate over only a limited temperature range depending on the application. Under the SAFT model, $M_3/M_i = a_i$. Then by (4), it is reasonable that the relative acceleration factor and the SL have the linear relationship:

$$\log(a_i) = \beta_0 + \beta_1 \cdot \frac{1}{T_i}, \tag{5}$$

for some real numbers β_0 and β_1. Regressing $\{\hat{a}_i\}$ on $\{1/T_i\}$ using model (5), gives $\hat{\beta}_0 = 7.9438$ and $\hat{\beta}_1 = -3017.0264$. P-values for $\hat{\beta}_0$ and $\hat{\beta}_1$ are less than 0.05, and R^2=0.9975.

[6] In principle, it merely defines the activation energy.

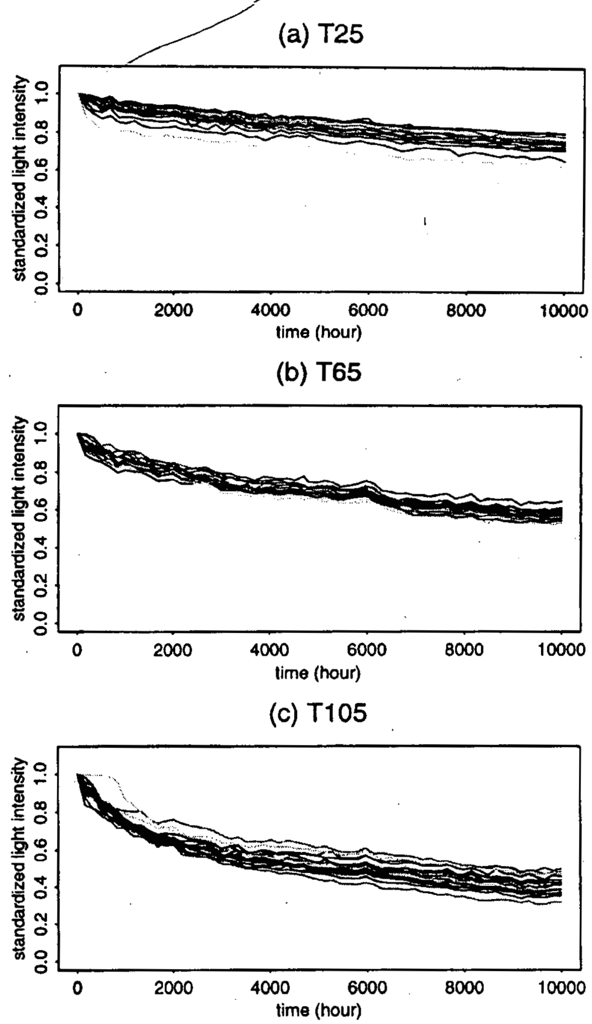

Figure 4: ADT Data of an LED Product

a. 16 SLI curves for 25°C,
b. 18 SLI curves for 65°C,
c. 19 SLI curves for 105°C

These evidences support the potential adequacy of (5). Further, since the usual use condition (20°C) is not too far away from the temperature range tested, we then can estimate the relative acceleration factor at the usual use condition by

$$\hat{a}_0 = \exp\left[\hat{\beta}_0 + \frac{\hat{\beta}_1}{273.16 + 20}\right] = 0.09559846.$$

4. By an industrial standard, an LED product is declared failed when its SLI degrades to 1/2. Consequently, we can estimate the MTTF of the LED product in group 3 by finding t_3^* when $\hat{\mu}_3(t_3^*) = 1/2$. Then the MTTF under

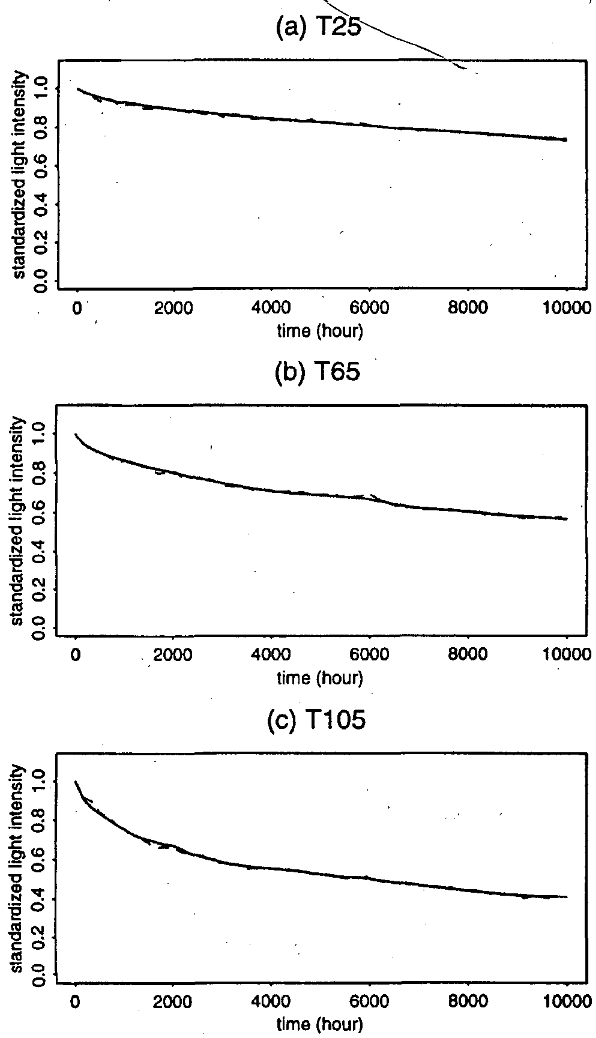

Figure 5: Solid Curves Are Estimated Mean-Degradation Curves; Dashed Curves are Averaged-Data Curves for:
a. 25°C, b. 65°C, c. 105°C

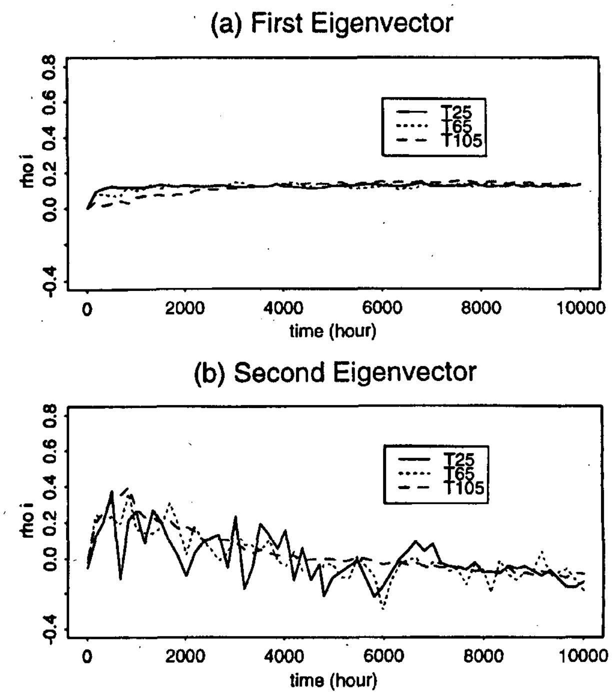

Figure 6: The Eigenvectors of the Sample Covariance Matrix Associated with: a. the largest, and b. the second largest, eigenvalue for each group.
Solid curves are 25°C,
Dotted curves are 65°C,
Dashed curves are 105°C

the usual use condition can be estimated by:

$$t_0^* = \frac{\hat{a}_3}{\hat{a}_0} \cdot t_3^* = 52211.94 \text{ hours } = 5.96 \text{ years}.$$

5. The stochastic component is estimated for each group. We show only the first 2 eigenvectors in figure 6; under different SL they have some features in common.

5. SIMULATION RESULTS

A data analysis on real data cannot provide enough information on how well the procedure does in recovering the truth. To investigate the performance of NPRALS, we conduct a simulation study using the estimated parameters, from the LED data analysis in section 4, to generate the (hypothetical) truth. This truth is generated as follows.

1. Let,

$$a_i = \exp\left[\hat{\beta}_0 + \hat{\beta}_1 \cdot \frac{1}{T_i}\right], \text{ for } i = 0, 1, 2, 3.$$

The $\hat{\beta}_0$ and $\hat{\beta}_1$ are from section 4. Rescale $\{a_i\}$ such that $a_3 = 1$ (divide each a_i by a_3). Then these rescaled a_i are regarded as the true relative acceleration factors.

2. Let $\mu_i(t) = \hat{\mu}_3 \cdot (a_i \cdot t)$, $i = 0, 1, 2, 3$. The $\hat{\mu}_3(\cdot)$ is from section 4, and the a_i are from step 1. These $\mu_i(t)$ are regarded as the true mean function.

3. Compute V_i, for $i = 1, 2, 3$:

$$(V_i)_{s,r} = \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} [X_{i,j,s} - \mu_i(t_s)] \cdot [X_{i,j,r} - \mu_i(t_r)],$$

for $1 \le i \le 3, \quad 1 \le j \le n_i, \quad 1 \le s, r, k \le 59.$

Notation

$\sigma_{i,q}^2$ q^{th} largest eigenvalue of V_i

$\rho_{i,q}$ eigenvector associated with $\sigma_{i,q}^2$

These 2 parameters are regarded as the true parameters.

The simulated data are constructed from (2) as follows. Let,

$n_1 = 16$, $n_2 = 18$, $n_3 = 19$. Then for $i = 1, 2, 3$:

1. generate $\epsilon_{i,j,q}$ from $N(0, \sigma_{i,q}^2)$, $1 \le j \le n_i$, $1 \le q \le 59$;

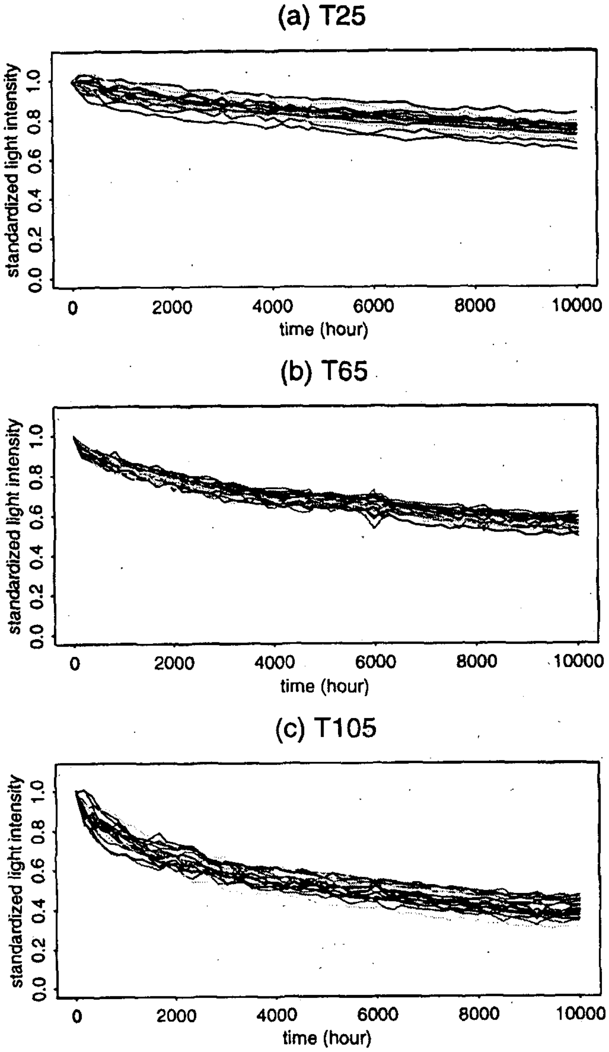

Figure 7: Simulated Data Generated from the Hypothetical True Model

a. 16 SLI curves for 25°C,
b. 18 SLI curves for 65°C,
c. 19 SLI curves for 105°C

2. generate $\{\tilde{X}_{i,j,k}\}$:

$$\tilde{X}_{i,j,k} = \mu_i(t_k) + \sum_{q=1}^{59} \epsilon_{i,j,q} \cdot \rho_{i,q}(t_k),$$

for $1 \le j \le n_i$, $1 \le k \le 59$.

Figure 7 shows the generated data; these simulated data have many features of the real data.

Perform the procedure, described in section 3, on the simulated data $\{\tilde{X}_{i,j,k}\}$. The estimates of a_i are summarized in table 1.

· Figure 8 shows $\hat{\mu}_i(\cdot)$ and the hypothetical true mean curves.

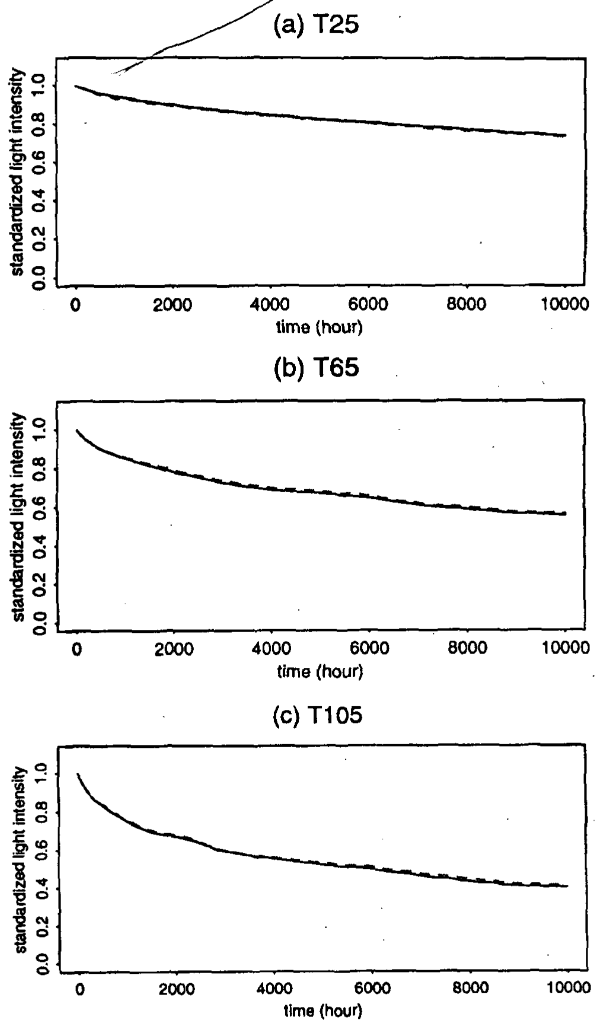

Figure 8: Mean Degradation Curves (MDC) for Simulated Data

Dashed curves: estimated MDC
Solid curves: (hypothetical) true MDC
a. 25°C, b. 65°C, c. 105°C

· Figure 9 shows $\sigma^2_{i,q}$, the hypothetical true eigenvalues, and $\hat{\sigma}^2_{i,q}$, the estimated eigenvalues.

· Figures 10, 11 show the hypothetical true and the estimated eigenvectors associated with the largest and the second-largest eigenvalues, respectively.

Table 1: True & Estimated Values of the a_i

	a_0	a_1	a_2	a_3
True	0.09559894	0.1175793	0.3891789	1
Est.	0.09052029	0.1063360	0.3987680	0.999993

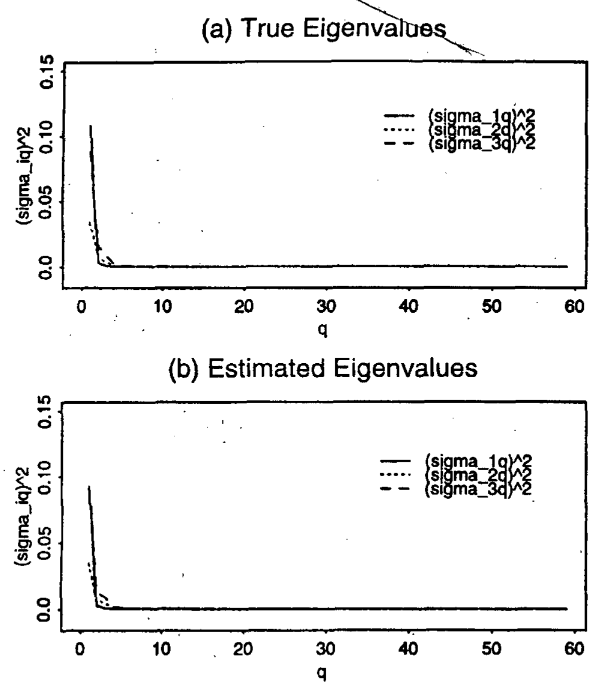

Figure 9: Eigenvalues

a. Hypothetical true eigenvalues
b. Estimated eigenvalues in the simulation.
Solid curves: $\sigma^2_{1,q}$,
Dotted curves: $\sigma^2_{2,q}$,
Dashed curves: $\sigma^2_{3,q}$

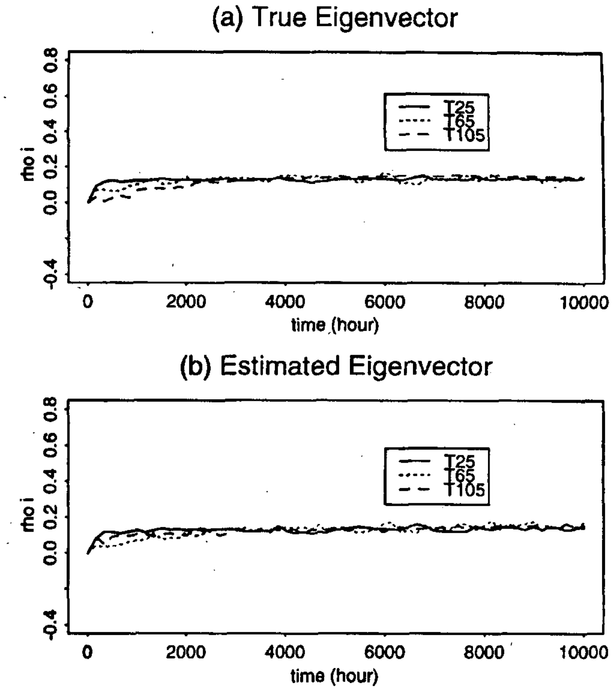

Figure 10: Eigenvector Associated with Largest Eigenvalue of:

a. Hypothetical true V_i
b. Estimated V_i.
Solid curves: 25°C,
Dotted curves: 65°C,
Dashed curves: 105°C

The results show that most of the estimated values are fairly close to the true values. NPRALS seems to perform quite well for this simulated example. This also indicates that NPRALS describes the LED data fairly well.

6. DISCUSSION

Meeker, Escobar, Lu [10] presented some good methods, based on parametric models, for analyzing accelerated degradation data similar to the data analyzed in this paper. A nonparametric regression alternative, such as NPRALS can free analysts from the burden of specifying models in the usual parametric modeling. The tradeoff is the slight inefficiency, which means one probably needs more data to get the same accuracy of the estimates using a parametric model (assuming that the parametric model is correct). On the other hand, a nonparametric regression method performs much better than a wrongly specified parametric model. A popular data-analysis strategy is to explore the data *via* nonparametric regression techniques first; if the results suggest a suitable parametric model, then perform the usual parametric methods for:

- better efficiency,
- possibly easier interpretation.

Nonparametric regression is a fast growing research area in statistics. The data-analyses developed from this research are getting more mature and advanced. The potential applications of these techniques are in any area that needs regression techniques, such as the application in this paper.

In this paper, the goal of estimating MTTF for the LED product under usual use is achieved by finding the relationship between the relative acceleration factors and the SL. The important issues of obtaining an interval estimate of MTTF and estimating the lifetime distribution are topics for future research.

ACKNOWLEDGMENT

We are pleased to thank the editor, an associate editor, two anonymous referees, and Prof. Sheng-Tsaing Tseng of National Tsing Hua University, whose helpful suggestions led to a great improvement on the quality of the paper. We are very grateful to Prof. Tseng for providing the LED data to us. Jyh-Jen Shiau is grateful to the Department of Statistics at Stanford University for their hospitality; this paper was revised during her visit there. This research was partially supported by the National Science Council of the Republic of China, Grant No. NSC88-2118-M-009-009.

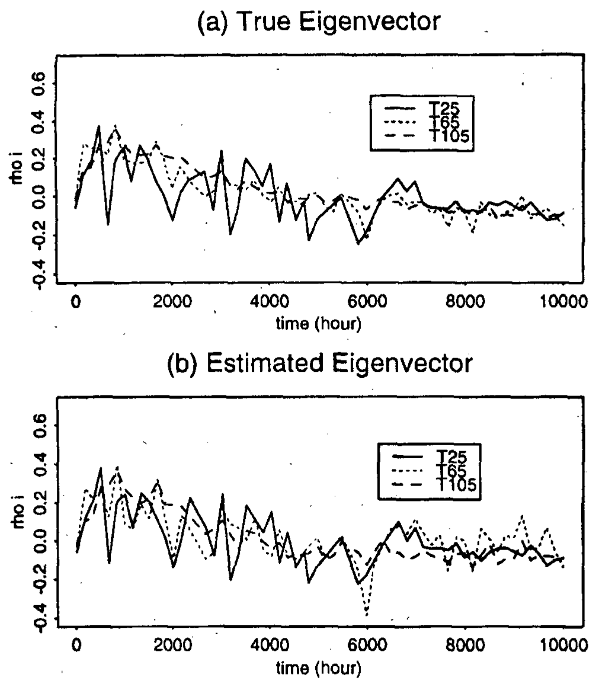

Figure 11: Eigenvector Associated with Second-Largest Eigenvalue of:

a. Hypothetical true V_i
b. Estimated V_i.
Solid curves: 25°C,
Dotted curves: 65°C,
Dashed curves: 105°C

APPENDIX

Appendix A.1 briefly describes the LLR smoother method to smoothing one curve data. Appendix A.2 extends the method to estimating efficiently the mean curve of a group of curve data. Appendix A.3 further extends the method to estimating efficiently the mean curve from the complied data of several groups of curve data.

A.1 Local Linear Regression Smoother

This is a locally weighted least squares estimation. Given n observations $\{(t_k, y_k)\}_{k=1}^{n}$, consider the model:

$$y_k = \mu(t_k) + \epsilon_k, \quad k = 1, \ldots, n,$$

$\mu(\cdot)$ is a smooth function

$\{\epsilon_k\}$ are uncorrelated errors with zero mean and standard deviation σ.

We obtain a $\hat{\mu}(t)$ by the minimizer $\hat{b}_0$ of:

$$\min_{b_0, b_1} \left[\sum_{k=1}^{n} (y_k - b_0 - b_1(t_k - t))^2 \cdot K_h(t_k - t) \right],$$

$$K_h(\cdot) \equiv \frac{1}{h} \cdot K\left(\frac{\cdot}{h}\right).$$

$K(\cdot)$ = a kernel function,

h = the bandwidth for the local smoothing.

$\hat{\mu}(t)$ is explicitly expressed by,

$$\hat{\mu}(t) = \frac{\sum_{k=1}^{n} w_k \cdot y_k}{\sum_{k=1}^{n} w_k},$$

$$w_k \equiv K_h(t_k - t)\left[S_{n,2} - (t_k - t) \cdot S_{n,1}\right],$$

$$S_{n,l} \equiv \sum_{k=1}^{n} K_h(t_k - t) \cdot (t_k - t)^l, \quad l = 1, 2.$$

For this study, we chose Epanechnikov kernel for $K(\cdot)$:

$$K(t) = \tfrac{3}{4} \cdot (1 - t^2) \cdot I(|t| \le 1).$$

Epanechnikov kernel is one of the most popular kernels in nonparametric regression; it is the optimal kernel in the sense that it minimizes asymptotic mean squared error of the function estimator over a group of kernels [2: theorem 3.4]. For more information on LLR smoothers, see, *eg*, [2, 4, 5].

A.2 Method Extension

For group m, the LLR estimator of $\mu_m(t)$ is:

$$\hat{\mu}_m(t) = \frac{\sum_{j=1}^{n_m} \sum_{k=1}^{p} w_{m,j,k} \cdot X_{m,j,k}}{\sum_{j=1}^{n_m} \sum_{k=1}^{p} w_{m,j,k}};$$

$$w_{m,j,k} \equiv K_h(t_{m,j,k} - t) \cdot \left[S_{n,2} - (t_{m,j,k} - t) \cdot S_{n,1}\right].$$

$$S_{n,l} \equiv \sum_{j=1}^{n_m} \sum_{k=1}^{p} K_h(t_{m,j,k} - t) \cdot (t_{m,j,k} - t)^l,$$

$$\text{for } l = 1, 2, \quad n = n_m \cdot p.$$

$$X_{m,k} \equiv \sum_{j=1}^{n_m} X_{m,j,k},$$

$$\bar{X}_{m,k} \equiv \frac{X_{m,k}}{n_m}.$$

Since $w_{m,j,k} = w_{m,j',k}$, for all $j, j' = 1, \ldots, n_m$, then by setting $w_{m,k} = w_{m,j,k}$,

$$\hat{\mu}_m(t) = \frac{\sum_{k=1}^{p} w_{m,k} \cdot X_{m,k}}{n_m \cdot \sum_{k=1}^{p} w_{m,k}} = \frac{\sum_{k=1}^{p} w_{m,k} \cdot \bar{X}_{m,k}}{\sum_{k=1}^{p} w_{m,k}}.$$

Thus, smoothing all the data

$\{X_{m,j,k},\ j = 1, \ldots, n_m,\ k = 1, \ldots, p\}$

can be simplified to smoothing

$\{\bar{X}_{m,k},\ k = 1, \ldots, p\}$.

A.3 Further Method Extension

For the compiled data $\{(t'_{i,j,k},\ X'_{i,j,k})\}$, the LLR estimator of $\mu_m(t)$ is:

$$\hat{\mu}_m(t) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{p} w_{i,j,k}\cdot X'_{i,j,k}}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{p} w_{i,j,k}};$$

$$w_{i,j,k} = K_h(t'_{i,j,k}-t)\cdot\left[S_{n,2}-(t'_{i,j,k}-t)\cdot S_{n,1}\right],$$

$$t'_{i,j,k} = \frac{\hat{a}_i}{\hat{a}_m}\cdot t_k,$$

$$S_{n,l} = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{p} K_h(t'_{i,j,k}-t)\cdot(t'_{i,j,k}-t)^l,$$

$$l = 1, 2,\quad n = (n_1+\ldots+n_m)\cdot p.$$

Notation

$X'_{i,k}$ $\quad \sum_{j=1}^{n_i} X'_{i,j,k}$

$\bar{X}'_{i,k}$ $\quad \frac{X'_{i,k}}{n_i}$

$w_{i,k}$ $\quad w_{i,j,k}$

Since $w_{i,j,k} = w_{i,j',k}$, for all $j, j' = 1,\ldots,n_i$,

$$\hat{\mu}_m(t) = \frac{\sum_{i=1}^{m}\sum_{k=1}^{p} w_{i,k}\cdot X_{i,k}}{\sum_{i=1}^{m}\sum_{k=1}^{p} w_{i,k}\cdot n_i} = \frac{\sum_{i=1}^{m}\sum_{k=1}^{p} w_{i,k}\cdot n_i\cdot \bar{X}'_{i,k}}{\sum_{i=1}^{m}\sum_{k=1}^{p} w_{i,k}\cdot n_i}.$$

Thus, smoothing all the data:

$\{X'_{i,j,k},\ \ i=1,\ldots,m,\ \ j=1,\ldots,n_i,\ \ k=1,\ldots,p\}$ can be simplified to smoothing:

$\{\bar{X}'_{i,k},\ \ i=1,\ldots,m,\ \ k=1,\ldots,p\}$ with adjusted weights on $\{\bar{X}'_{i,k}\}$. ◀

REFERENCES

[1] W.B. Capra, H.-G. Müller, "An accelerated-time model for response curves", *J. American Statistical Assoc*, vol 92, 1997 Mar, pp 72 – 83.

[2] J. Fan, I. Gijbels, *Local Polynomial Modeling and Its Applications*, 1996; Chapman & Hall.

[3] P.J. Green, B.W. Silverman, *Nonparametric Regression and Generalized Linear Model: A Roughness Penalty Approach*, 1994; Chapman & Hall.

[4] W. Härdle, *Applied Nonparametric Regression*, 1990; Cambridge Univ. Press.

[5] W. Härdle, *Smoothing Techniques: With Implementation in S*, 1991; Springer-Verlag.

[6] T. Hastie, R. Tibshirani, *Generalized Additive Models*, 1990; Chapman & Hall.

[7] C.J. Lu, W.Q. Meeker, "Using degradation measures to estimate a time-to-failure distribution", *Technometrics*, vol 35, 1993 May, pp 161 – 174.

[8] W.Q. Meeker, L.A. Escobar, "A review of recent research and current issues in accelerated testing", *Int'l Statistical Review*, vol 61, 1993 Apr, pp 147 – 168.

[9] W.Q. Meeker, L.A. Escobar, *Reliability Data Analysis*, 1998; John Wiley & Sons.

[10] W.Q. Meeker, L.A. Escobar, C.J. Lu, "Accelerated degradation tests: Modeling and analysis", *Technometrics*, vol 40, 1998 May, pp 89 – 99.

[11] W. Nelson, *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*, 1990; John Wiley & Sons.

[12] J.O. Ramsey, B.W. Silverman, *Functional Data Analysis*, 1997; Springer.

[13] G. Wahba, *Spline Models for Observational Data*, 1990; Society for Industrial and Applied Mathematics.

[14] H.F. Yu, S.T. Tseng, "On-line procedure for terminating an accelerated degradation test", *Statistica Sinica*, vol 8, 1998 Apr, pp 207 – 220.

AUTHORS

Dr. Jyh-Jen Horng Shiau; Institute of Statistics; National Chiao Tung Univ.; Hsinchu 300 TAIWAN – R.O.C.
Internet (e-mail): jyhjen@stat.nctu.edu.tw

Jyh-Jen Horng Shiau received her BS in Mathematics from the National Taiwan University, MS in Applied Mathematics from the University of Maryland Baltimore County, MS in Computer Science and PhD in Statistics from the University of Wisconsin – Madison. She taught in Southern Methodist University, University of Missouri at Columbia, and National Tsing Hua University, and worked for Engineering Research Center of AT&T Bell Labs before she moved to Hsinchu, Taiwan, where she is an Associate Professor in the Institute of Statistics at the National Chiao Tung University. Her interests include industrial statistics and nonparametric function estimation. She is a member of Institute of Mathematical Statistics.

Hsin-Hua Lin; Institute of Statistics; National Chiao Tung Univ.; Hsinchu 300 TAIWAN – R.O.C.

Hsin-Hua Lin is a graduate student in the Institute of Statistics at the National Chiao Tung University. She received her BS in Mathematics from Chung Yuan Christian University. She is interested in industrial statistics.

Manuscript TR1998-136 received: 1998 September 8;
revised: 1998 October 7, 1999 February 23
Responsible editor: H.-K. Hsieh
Publisher Item Identifier S 0018-9529(99)06711-1