Nonparametric Estimation of a Lifetime Distribution When Censoring Times Are Missing

X. Joan Hu
Department of Biostatistics
Harvard School of Public Health
Boston, MA 02115

Jerald F. Lawless
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
(jlawless@setosa.uwaterloo.ca)

Kazuyuki Suzuki
Department of Communications and Systems Engineering
University of Electro-Communication
Chofu-City, Tokyo 182
Japan

We consider datasets for which lifetimes associated with the units in a population are observed if they occur within certain time intervals but for which lengths of the time intervals, or censoring times of unfailed units, are missing. We consider nonparametric estimation of the lifetime distribution for the population from such data; a maximum likelihood estimator and a simple moment estimator are obtained. An example involving automobile warranty data is discussed at some length.

KEY WORDS: Incomplete data; Moment estimates; Warranty reports.

There are several contexts in the analysis of failure time or lifetime data in which censoring times for unfailed units are missing. The area that motivated the current research concerns the estimation of failure-time distributions or rates from product warranty data. If under warranty a product may experience a certain type of event, or "failure," then we can estimate the distribution of time to failure (or the intensity function for recurrent events) over the warranty period from warranty reports. We typically have to deal with missing censoring times, however, as we now discuss.

Suppose that T_i is the time to failure for product unit i in a population of M manufactured units. In some applications T_i is measured in calendar time since the sale time of the unit. For many types of products the manufacturers do not know the date of sale for most units, and therefore the censoring time (i.e., the elapsed time between the sale of the item and when the data are assembled) for most unfailed items is unknown. For units that fail under warranty, the failure time and the potential censoring time are known because the date of sale is verified as part of the warranty claims process.

Similar problems arise when T_i is some type of usage, or operational time. A familiar example is in connection with automobiles, where T_i represents the mileage at failure. If failure data are collected up to some current date, the censoring time is the minimum of the vehicle's current mileage and the mileage at which it passes out of the warranty plan. For example, for a two-year/24,000-mile warranty, this latter mileage is the lesser of the vehicle's mileage at two years and 24,000 miles. Because exact mileage accumulation data are not available for most cars, the exact censoring times are in general unknown. By making the crude but generally satisfactory assumption that mileage accumulation is linear over time, however, they may be estimated for cars experiencing a failure because the date of sale, date of failure, and mileage at failure are all observed. An application involving automobile warranty data is discussed at some length in Section 5.

Suppose that the lifetime variable T has distribution function $F(t) = \Pr(T \leq t)$ and that the population of M units has independent lifetimes $t_1, \ldots, t_M$ generated from that distribution. There are also censoring times $\tau_1, \ldots, \tau_M$ associated with the units, and we assume that the τ_i's are independent of each other, with common distribution function $G(\tau) = \Pr(T_i \leq \tau)$. The distribution $G(\tau)$ is determined by the random process by which units are sold and by the random process by which units accumulate usage over calendar time, when T is a usage time. A reviewer asked whether finite population assumptions might be used instead. This is feasible but more difficult to implement and, given the large sample sizes typical of this area, unlikely to give results that differ much from ours. The observed data are as follows: If $t_i \leq \tau_i$, we observe t_i (and possibly also τ_i), but if $t_i > \tau_i$, we know only that fact and not the value of τ_i or t_i. Our objective is to estimate the distribution $F(t)$ from such data, avoiding any parametric assumptions.

Suzuki (1985), Kalbfleisch and Lawless (1988), and Hu and Lawless (1996a) discussed the use of supplementary follow-up samples of unfailed units as a way to compensate for the missing censoring times. In many circumstances it is possible to estimate the censoring-time distribution, however, and this provides another approach. We present in this article two nonparametric estimation methods for the case in which the censoring-time distribution $G(\tau)$ is known, or at least estimated from other sources. The main assumptions that we make initially are (a) the number of product units M in service is known, (b) all failures are reported under the data-collecting scheme, (c) the censoring-time distribution is known, and (d) censoring times are statistically independent of failure times. The assumptions are discussed further in the article, and ways to handle departures from them are presented.

Section 1 presents nonparametric maximum likelihood and moment estimators for $F(t)$, assuming that $G(\tau)$ is known. Section 2 reports on a small simulation study comparing the two estimators. Section 3 considers cases in which $G(\tau)$ is estimated, and Section 4 examines the assumption of independence of failure and censoring times. Section 5 presents a detailed example involving automobiles. Section 6 outlines extension of the methodology to deal with multiple failure modes and recurrent events, and Section 7 presents some concluding remarks.

1. MAXIMUM LIKELIHOOD AND SIMPLE MOMENT ESTIMATORS

A nonparametric method of lifetime distribution estimation was previously given by Suzuki (1988). Suzuki and Kasashima (1993), however, showed that method to be inferior to maximum likelihood, so we will not discuss it here.

To develop nonparametric estimators, it is convenient and customary to work with discrete distributions; finite-sample estimates of continuous cumulative distribution functions $F(t)$ are discrete anyway and may be obtained from the discrete-time framework. Thus, we assume that lifetime T and censoring time $\mathcal{T}$ may each take on values $1, 2, \ldots,$ and $f(t) = \Pr(T_i = t), g(\tau) = \Pr(\mathcal{T}_i = \tau)$. The corresponding cumulative distribution functions are $F(t) = f(1) + \cdots + f(t)$ and $G(\tau) = g(1) + \cdots + g(\tau)$. In this section, we assume that T_i and $\mathcal{T}_i$ are independent and that $G(\tau)$ is known.

1.1 Maximum Likelihood Estimation

With known population size M and censoring-time distribution $G(\tau)$, the likelihood function based on the probability of the observed data for the population is of the familiar censored-data form (Lawless 1982, chap. 1),

$$\prod_{t_i \le \tau_i} f(t_i) \prod_{t_i > \tau_i} \Pr(T_i > \tau_i). \tag{1}$$

The difference with the usual situation is that τ_i is not observed for the unfailed units. Thus, treating it as a random variable with distribution $G(\cdot)$,

$$\begin{aligned}\Pr(T_i > \tau_i) &= \sum_{\tau=1}^{\tau_{\max}} \left[1 - \sum_{t \le \tau} f(t)\right] g(\tau) \\ &= 1 - \sum_{t=1}^{\tau_{\max}} f(t)\bar{G}(t),\end{aligned}$$

where $\bar{G}(\tau) = \Pr(\mathcal{T}_i \ge \tau)$ and $\tau_{\max} = \sup\{\tau : \bar{G}(\tau) > 0\}$. We assume that $\tau_{\max} < \infty$, which is unrestrictive in the present context.

The likelihood function (1) may be written as

$$\prod_{t=1}^{\tau_{\max}} f(t)^{n_t} \left[1 - \sum_{t=1}^{\tau_{\max}} f(t)\bar{G}(t)\right]^{M-m}, \tag{2}$$

where n_t is the number of failures at time t that are observed and m is the total number of failures observed. Notice that $\sum_{t=1}^{\tau_{\max}} n_t = m$. Using $\#A$ to represent the number of elements in set A, we have $n_t = \#\{t_i : t_i \le \tau_i, t_i = t\}$ and $m = \#\{i : t_i \le \tau_i\}$. Estimates of $f(t), t = 1, \ldots, \tau_{\max}$, can be obtained by maximizing (2) with respect to $f(1), \ldots, f(\tau_{\max})$ under the constraints $f(t) \ge 0$ and $f(1) + \cdots + f(\tau_{\max}) = F(\tau_{\max}) \le 1$. Denote the log-likelihood function by

$$l(\underline{f}) = \sum_{t=1}^{\tau_{\max}} n_t \log f(t) + (M - m) \log\left[1 - \sum_{t=1}^{\tau_{\max}} f(t)\bar{G}(t)\right].$$

The solution of the equations

$$\frac{\partial l}{\partial f(t)} = \frac{n_t}{f(t)} - (M - m)\frac{\bar{G}(t)}{1 - \sum_{s=1}^{\tau_{\max}} f(s)\bar{G}(s)} = 0, \qquad t = 1, \ldots, \tau_{\max}, \tag{3}$$

maximizes $l(\underline{f})$ with no constraints on the $f(t)$'s. By rewriting (3) as

$$(M - m)\bar{G}(t)f(t) = n_t\left\{1 - \sum_{s=1}^{\tau_{\max}} f(s)\bar{G}(s)\right\}$$

and summing both sides over $t = 1, \ldots, \tau_{\max}$, we see that $\sum_{s=1}^{\tau_{\max}} f(s)\bar{G}(s)$ equals m/M. Inserting this into (3), we then find that (3) is solved uniquely by

$$\hat{f}_{\mathrm{ML}}(t) = \frac{n_t}{M\bar{G}(t)}, \qquad t = 1, \ldots, \tau_{\max}. \tag{4}$$

This then gives the maximum likelihood estimates, provided that $\hat{f}_{\mathrm{ML}}(1) + \cdots + \hat{f}_{\mathrm{ML}}(\tau_{\max}) \le 1$. This is virtually always the case when $F(\tau_{\max})$ is not too close to 1, which is satisfied in most applications. In the warranty-reports context, for example, $\tau_{\max}$ is the maximum failure time observable and not larger than the warranty time limit, and the probability that a unit fails while under warranty is considerably less than 1. If the constraint is not met, then doubt may be cast on the validity of the assumed function $G(\tau)$. When the estimates (4) sum to slightly over 1, a reasonable approach is to simply rescale them so that they sum to 1, but this may differ slightly from the maximum likelihood estimator (MLE). To get the MLE, we need to maximize (2) under constraints $f(t) \ge 0$ and $f(1) + \cdots + f(\tau_{\max}) = 1$ in that situation.

The nonparametric maximum likelihood estimate of $F(t), t = 1, \ldots, \tau_{\max}$, is $\hat{F}_{\mathrm{ML}}(t) = \hat{f}_{\mathrm{ML}}(1) + \cdots + \hat{f}_{\mathrm{ML}}(t)$. Similar to the discussion for the nonparametric MLE of Hu and Lawless (1996a), for example, arguments can be given to establish the consistency and asymptotic normality of this estimator. This also follows from the next section.

1.2 A Simple Moment Estimator

We see that the number of observed units with failures at t in (2) is $n_t = \sum_{i=1}^{M} \mathrm{I}\ (t_i = t, \tau_i \geq t)$, where $\mathrm{I}(A)$ is the indicator of event A (i.e., it equals 1 if A is true and 0 if not). Let $u_i(t) = \mathrm{I}\ (t_i = t, \tau_i \geq t)/\bar{G}(t), t = 1, 2, \ldots$. A simple moment estimator of $f(t)$,

$$\hat{f}_{\mathrm{SM}}(t) = \frac{1}{M}\sum_{i=1}^{M} u_i(t) = \frac{n_t}{M\bar{G}(t)}, \qquad t = 1, \ldots, \tau_{\max}, \quad (5)$$

is obtained by noting that $\mathrm{E}\{n_t\} = M\bar{G}(t)f(t), t = 1, \ldots, \tau_{\max}$. Notice that $\hat{f}_{\mathrm{SM}}(t)$ is the same as the nonparametric MLE $\hat{f}_{\mathrm{ML}}(t)$ in (4) in the current situation. Under our assumptions, n_t is binomial$(M, f(t)\bar{G}(t))$ and it is easy to see that $\hat{f}_{\mathrm{SM}}(t)$ is unbiased with variance

$$\mathrm{var}\{\hat{f}_{\mathrm{SM}}(t)\} = \frac{f(t)}{M\bar{G}(t)}\,[1 - f(t)\bar{G}(t)], \qquad t = 1, \ldots, \tau_{\max}. \quad (6)$$

The sample variance estimator

$$\begin{aligned} \widehat{\mathrm{var}}\{\hat{f}_{\mathrm{SM}}(t)\} &= \frac{1}{M^2}\sum_{i=1}^{M}\{u_i(t) - \bar{u}(t)\}^2 \\ &= \frac{n_t(M - n_t)}{M^3\bar{G}(t)^2}, \end{aligned}$$

with $\bar{u}(t) = \sum_{i=1}^{M} u_i(t)/M = \hat{f}_{\mathrm{SM}}(t)$, is the same as the consistent estimate for the variance (6) achieved by replacing $f(t)$ with its estimate $\hat{f}_{\mathrm{SM}}(t)$.

Then the estimator for $F(t), t = 1, \ldots, \tau_{\max}$, based on $\hat{f}_{\mathrm{SM}}(\cdot)$, is

$$\hat{F}_{\mathrm{SM}}(t) = \sum_{s=1}^{t}\hat{f}_{\mathrm{SM}}(s), \qquad t = 1, \ldots, \tau_{\max}. \quad (7)$$

By noting that the n_t's are multinomial ($t = 1, \ldots, \tau_{\max}$), we have

$$\mathrm{cov}\{\hat{f}_{\mathrm{SM}}(s_1), \hat{f}_{\mathrm{SM}}(s_2)\} = \frac{f(s_1)}{M\bar{G}(s_1)}[\mathrm{I}(s_1 = s_2) - f(s_2)\bar{G}(s_1)]$$

and a consistent estimate for the variance of $\hat{F}_{\mathrm{SM}}(t)$ as

$$\begin{aligned} \widehat{\mathrm{var}}\{\hat{F}_{\mathrm{SM}}(t)\} &= \frac{1}{M^2}\sum_{i=1}^{M}\left\{\sum_{s=1}^{t}[u_i(s) - \bar{u}(s)]\right\}^2 \\ &= \sum_{s=1}^{t}\frac{n_s(M - n_s)}{M^3\bar{G}(s)^2} - \sum_{s_1 \neq s_2}\frac{n_{s_1}n_{s_2}}{M^3\bar{G}(s_1)\bar{G}(s_2)}. \end{aligned} \quad (8)$$

1.3 Extension

We generalize the preceding situation slightly to allow the distribution of $\mathcal{T}_i$ to depend on a discrete covariate or group indicator x_i. Suppose that x_i takes on values $x_1^0, \ldots, x_K^0$ and is observable. This is useful because with automobiles, for example, the censoring time for a car may depend on what time it entered service and that time is usually provided by the dealer. In such cases it is not possible to estimate $G(\tau)$ without considering the sales pattern.

We denote $g_k(\tau) = \Pr(\mathcal{T}_i = \tau | x_i = x_k^0)$ and $G_k(\tau) = \Pr(\mathcal{T}_i \leq \tau | x_i = x_k^0)$. Let $M_k = \#\mathcal{P}_k$, with $\mathcal{P}_k = \{i : x_i = x_k^0, i = 1, \ldots, M\}, k = 1, \ldots, K$. We assume that the distribution of T_i does not depend on x_i now. With known subpopulation sizes M_k and censoring-time distributions $G_k(\tau), k = 1, \ldots, K$, the likelihood function of the failure-time distribution based on the data available is

$$\begin{aligned} &\prod_{t_i \leq \tau_i} f(t_i) \prod_{t_i > \tau_i} \Pr(T > \mathcal{T}_i | x_i) \\ &= \prod_{t=1}^{\tau_{\max}} f(t)^{n_t} \prod_{k=1}^{K}\left[1 - \sum_{s=1}^{\tau_{\max}} f(s)\bar{G}_k(s)\right]^{M_k - m_k}, \end{aligned} \quad (9)$$

where $\bar{G}_k(t) = \Pr(\mathcal{T}_i \geq \tau | x_i = x_k^0), \tau_{\max} = \max_k \sup\{\tau : \bar{G}_k(\tau) > 0\}$, and $m_k = \#\{x_i : x_i = x_k^0, t_i \leq \tau_i\}$. A nonparametric MLE of $f(t)$ can be obtained similarly to the procedure in Section 1.1. In this case the maximum likelihood equations reduce to

$$\hat{f}_{\mathrm{ML}}(t) = n_t\left\{\sum_{k=1}^{K}\frac{(M_k - m_k)\bar{G}_k(t)}{1 - \sum_{s=1}^{\tau_{\max}}\hat{f}_{\mathrm{ML}}(s)\bar{G}_k(s)}\right\}^{-1}, \qquad t = 1, \ldots, \tau_{\max}. \quad (10)$$

There is no closed form for $\hat{f}_{\mathrm{ML}}(t)$ when $K \geq 2$. Equation (10) can be used to provide an iteration scheme for obtaining the $\hat{f}_{\mathrm{ML}}(t)$'s.

We can, however, obtain a moment estimator. The fact that

$$\mathrm{E}\{n_t | x_i, i = 1, \ldots, M\} = \sum_{k=1}^{K} M_k\bar{G}_k(t)f(t), \qquad t = 1, \ldots, \tau_{\max},$$

gives an unbiased estimator of the $f(t)$:

$$\hat{f}_{\mathrm{ESM}}(t) = \frac{n_t}{\sum_{k=1}^{K} M_k\bar{G}_k(t)}, \qquad t = 1, \ldots, \tau_{\max}. \quad (11)$$

This also arises if we approximate $1 - \sum_{s=1}^{\tau_{\max}}\hat{f}_{\mathrm{ML}}(s)\bar{G}_k(s)$ in (10) with $(M_k - m_k)/M_k, k = 1, \ldots, K$. Notice that the three estimators $\hat{f}_{\mathrm{SM}}(\cdot)$ in (5), $\hat{f}_{\mathrm{ML}}(\cdot)$ from (10), and $\hat{f}_{\mathrm{ESM}}(\cdot)$ in (11) have the same numerator n_t but different denominators: $M\bar{G}(t), \sum_{k=1}^{K}(M_k - m_k)\bar{G}_k(t)/(1 - \sum_{s=1}^{\tau_{\max}}\hat{f}_{\mathrm{ML}}(s)\bar{G}_k(s))$, and $\sum_{k=1}^{K} M_k\bar{G}_k(t)$. Because

$$\begin{aligned} \bar{G}(t) &= \sum_{k=1}^{K}\bar{G}_k(t)\Pr(X = x_k^0) \\ &= \sum_{k=1}^{K}\bar{G}_k(t)\frac{\Pr(T > \mathcal{T}, X = x_k^0)}{\Pr(T > \mathcal{T} | X = x_k^0)}, \end{aligned}$$

the difference between $\hat{f}_{\mathrm{SM}}$ and $\hat{f}_{\mathrm{ESM}}$ is due to the difference between $\Pr(X = x_k^0)$ and its moment estimator

$M_k/M, k = 1, \ldots, K$, which is small when M is large enough; the difference between $\hat{f}_{\mathrm{ML}}$ and $\hat{f}_{\mathrm{ESM}}$ is due to the difference of the MLE $1 - \sum_{s=1}^{\tau_{\max}} \hat{f}_{\mathrm{ML}}(s)\bar{G}_k(s)$ and the moment estimator $(M_k - m_k)/M_k$ for $\Pr(T > \mathcal{T} | X = x_k^0)$, which is also small when M_k is large, $k = 1, \ldots, K$. Thus, the three estimators are virtually the same for cases with large M and M_k's.

We have

$$\operatorname{cov}\{\hat{f}_{\mathrm{ESM}}(s_1), \hat{f}_{\mathrm{ESM}}(s_2)\} = \frac{f(s_1)}{\sum_{k=1}^{K} M_k \bar{G}_k(s_1)} \times \left[\mathrm{I}(s_1 = s_2) - f(s_2) \frac{\sum_{k=1}^{K} M_k \bar{G}_k(s_1) \bar{G}_k(s_2)}{\sum_{k=1}^{K} M_k \bar{G}_k(s_2)} \right]. \quad (12)$$

The variance of $\hat{F}_{\mathrm{ESM}}(t) = \sum_{s=1}^{t} \hat{f}_{\mathrm{ESM}}(s)$ is

$$\operatorname{var}\{\hat{F}_{\mathrm{ESM}}(t)\} = \sum_{s_1=1}^{t} \sum_{s_2=1}^{t} \operatorname{cov}\{\hat{f}_{\mathrm{ESM}}(s_1), \hat{f}_{\mathrm{ESM}}(s_2)\}$$

and can be consistently estimated by

$$\widehat{\operatorname{var}}\{\hat{F}_{\mathrm{ESM}}(t)\} = \sum_{s_1=1}^{t} \sum_{s_2=1}^{t} \frac{n_{s_1}}{\{\sum_{k=1}^{K} M_k \bar{G}_k(s_1)\}^2} \times \left[\mathrm{I}(s_1 = s_2) - n_{s_2} \frac{\sum_{k=1}^{K} M_k \bar{G}_k(s_1) \bar{G}_k(s_2)}{\{\sum_{k=1}^{K} M_k \bar{G}_k(s_2)\}^2} \right], \quad (13)$$

obtained by replacing the $f(\cdot)$ in (12) with the estimate $\hat{f}_{\mathrm{ESM}}(\cdot)$. It is easily seen that $\sqrt{M}\{\hat{F}_{\mathrm{ESM}}(t) - F(t)\}$ has a limiting normal distribution as $M \to \infty$ and that to construct tests or confidence intervals it can be treated as normal with mean 0 and variance estimated by (13).

Estimation of the asymptotic variance of $\hat{f}_{\mathrm{ML}}(t)$ may be obtained through the standard procedure for an MLE. The information matrix is now

$$\mathrm{INFO}(\underline{f}) = -\frac{1}{M} \left(\frac{\partial^2 l}{\partial f(s_1) \partial f(s_2)} \right)_{\tau_{\max} \times \tau_{\max}}, \quad (14)$$

where the elements are

$$\frac{\partial^2 l}{\partial f(s_1) \partial f(s_2)} = -\mathrm{I}(s_1 = s_2) \frac{n_{s_1}}{f(s_1)^2} - \sum_{k=1}^{K} (M_k - m_k) \frac{\bar{G}_k(s_1) \bar{G}_k(s_2)}{\{1 - \sum_{s=1}^{\tau_{\max}} f(s) \bar{G}_k(s)\}^2},$$

$s_1, s_2 = 1, \ldots, \tau_{\max}$. Thus, an estimate of the asymptotic variance of $\hat{F}_{\mathrm{ML}}(t) = \sum_{s=1}^{t} \hat{f}_{\mathrm{ML}}(s)$ may be obtained as well. To implement the procedure is complex when $\tau_{\max}$ is large, however, because inversion of large matrices is involved. In addition, when censoring is heavy, $\mathrm{INFO}(\underline{f})$ could be singular and the method cannot be applied. The discussion by Hu and Lawless (1996a) for estimation of asymptotic variances of nonparametric MLE's applies here.

2. SIMULATION

Comparison of the nonparametric maximum likelihood and moment estimators presented in Section 1 was investigated through a simulation study. Being motivated by warranty data, we chose the following simulation setup. Consider a product with a one-year warranty; suppose that there are $M = 4{,}000$ units sold within a year and the warranty data have been collected over one and a half years since the first unit was sold (we take its sale time as 0). Time is calendar time, or age of a unit, measured from the date of sale. Suppose further that the times of these units to their first failures are independent from each other, identically Weibull distributed, and independent of their sale times and the sale times are uniformly distributed over the one-year period. The censoring time associated with unit i is now $\tau_i = \min(1, 1.5 - x_i)$ year, where x_i is its sale time and $\tau_{\max} = 1$.

We generated sale times x_i and failure times $t_i, i = 1, \ldots, M$, respectively, from discrete approximations to the uniform distribution over (0, 1] and the Weibull distribution

$$f(t) = \frac{\delta}{\alpha} \left(\frac{t}{\alpha} \right)^{\delta - 1} \exp\left\{ - \left(\frac{t}{\alpha} \right)^{\delta} \right\}$$

with $\delta = 2.0$ and $\alpha = 3.95, 1.85$. Discrete random variables were obtained through discretizing time as follows. We divided the time period into intervals, $((k - 1)/120, k/120], k = 1, 2, \ldots$, and assigned variables having values in the kth interval to the value $k/120$. The values $t = 1, 2, \ldots$ then correspond to $k/120, k = 1, 2, \ldots$. The values of the parameters were chosen to make the simulation realistic. The corresponding values of m/M for cases of $\alpha = 3.95, 1.85$ are about .05, .20, respectively, which allows us to study situations with heavy and moderately heavy censoring. From the simulated data, we evaluated the three nonparametric estimates of $F(t) = 1 - \exp\{-(t/\alpha)^{\delta}\}$, $t \in (0, 1]$: (1) $\hat{F}_{\mathrm{SM}}(t)$, based on (5); (2) $\hat{F}_{\mathrm{ESM}}(t)$, based on (11) and assuming that the numbers of units sold in each quarter of the year (M_1, M_2, M_3, M_4) are known; (3) $\hat{F}_{\mathrm{ML}}$, the MLE based on (10) for the same stratified-sample situation as (2).

We used S-Plus (©1988 MathSoft, Inc.) for generating the random variables needed and all the computing. The maximum likelihood estimate was evaluated by using the iteration procedure based on (10); we took $\tilde{f}^{(0)}(k/120) = f(k/120)$ and terminated the iterations when $\sum_{k=1}^{120} |\tilde{f}^{(j+1)}(k/120) - \tilde{f}^{(j)}(k/120)| < 10^{-4}$, where $\tilde{f}^{(j)}(t)$ is the jth iterate toward $\hat{f}_{\mathrm{ML}}(t)$.

Figure 1, (a)–(d), shows estimates for $F(.8)$ in the cases $\alpha = 3.95$ and $\alpha = 1.85$ based on the three estimators from 1,000 simulation repetitions. The corresponding sample means and standard errors are essentially the same. The sample means are very close to the true values of $F(.8)$ in the two cases, respectively. This indicates that there is almost no difference in the three estimators, and they all estimate $F(t)$ well in the situations we consider.

We repeated the preceding simulation with $M = 400$ instead, to study situations with sample size fairly small. The corresponding sample means and standard errors of the estimates for $F(.8)$ are once again essentially the same. We present the estimates for case $\alpha = 1.85$ in Figure 1, (e)–(f).

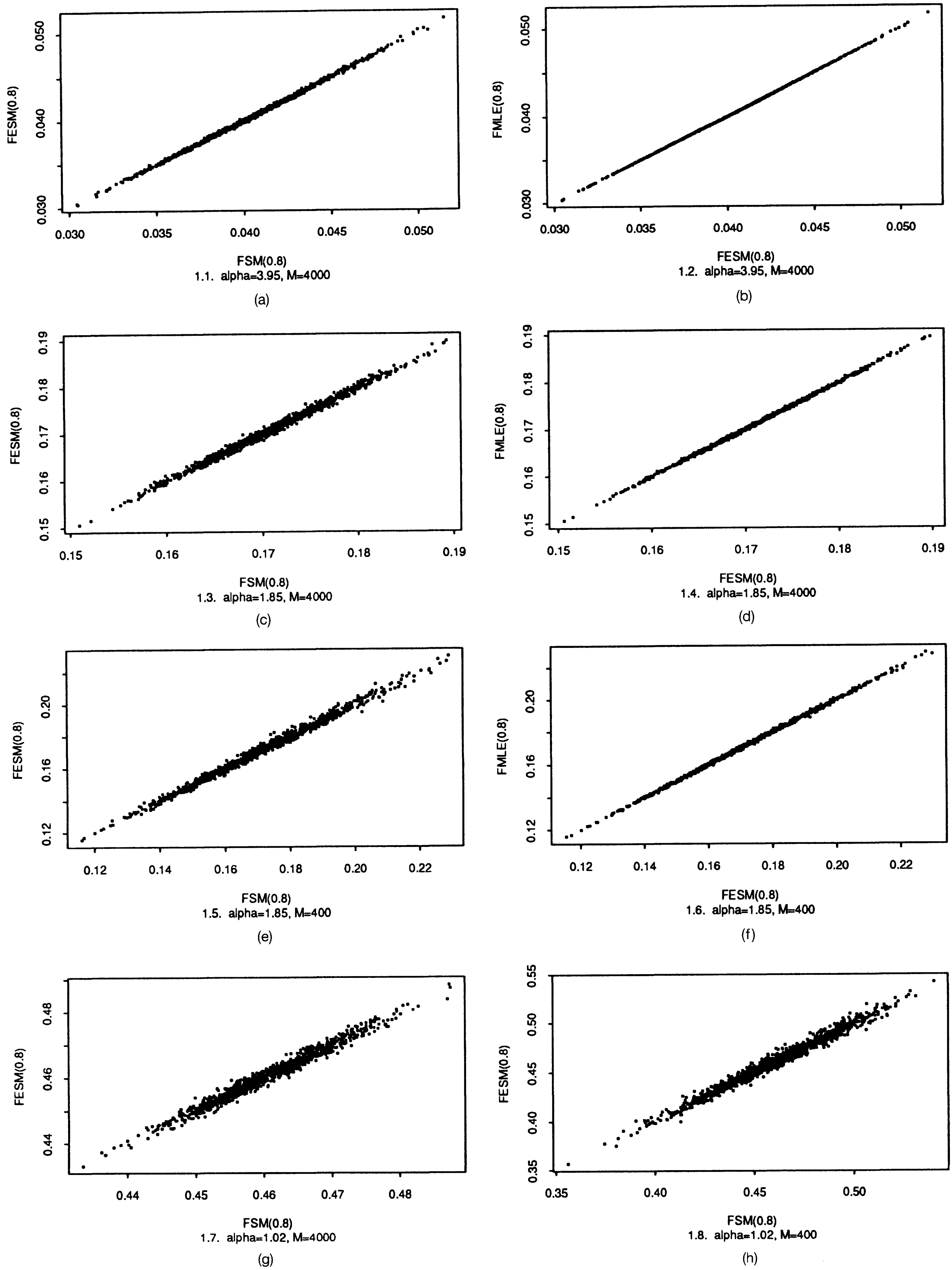

Figure 1. Estimates of 80% Quantiles of $F(t)$: FSM = $\hat{F}_{SM}(\cdot)$, FESM = $\hat{F}_{ESM}(\cdot)$, FMLE = $\hat{F}_{ML}(\cdot)$.

Table 1. Sample Means and Standard Errors (in brackets) of the Estimates for F(.5) and F(.8) From 1,000 Simulation Repetitions: α = 1.02.

	F(.5) = .2136		F(.8) = .4594	
Cases	M = 4,000	M = 400	M = 4,000	M = 400
$\hat{F}_{SM}$	.2139 (.0066)	.2127 (.0204)	.4609 (.0082)	.4595 (.0278)
$\hat{F}_{ESM}$	.2139 (.0066)	.2127 (.0204)	.4609 (.0081)	.4596 (.0276)
$\hat{F}_{ML}$	.2135 (.0062)	.2106 (.0207)	.4603 (.0076)	.4564 (.0275)

To investigate the estimators further, we conducted the two preceding simulations for the case $\alpha = 1.02$ (m/M in this case is about .5). The sample means and standard errors of the estimates are now slightly different and are presented in Table 1. Figure 1, (g)–(h), shows the estimates for $F(.8)$ based on $\hat{F}_{\mathrm{SM}}(t)$ and $\hat{F}_{\mathrm{ESM}}(t)$ in this case with $M = 4{,}000$ and $M = 400$, respectively.

Table 2 shows standard errors of estimates $\hat{F}_{\mathrm{SM}}(.8)$ and $\hat{F}_{\mathrm{ESM}}(.8)$ in different cases and the sample means of estimates for their standard deviations based on (8) and (13), respectively. We can see that (8) and (13) estimate well the variances of $\hat{F}_{\mathrm{SM}}(t)$ and $\hat{F}_{\mathrm{ESM}}(t)$, respectively.

As discussed in Section 1.3 and shown by the simulations, the three estimators agree closely, especially in the situations with heavy censoring. The practical consequence of this is that it is satisfactory to use the easily computed moment estimates.

3. EFFECT OF USING AN ESTIMATE OF THE CENSORING–TIME DISTRIBUTION

The estimation procedures in Section 1 assume that the censoring-time distribution $G(\tau)$ is known. In most practical situations, however, $G(\tau)$ is estimated, or only roughly known. Hu and Lawless (1996b) investigated likelihood-based parametric estimation for this situation; their approach can be extended to a nonparametric setting. We focus here on the extension of the simple moment estimator in Section 1.2; the estimator in Section 1.3 can be extended similarly. An example is presented in Section 5.

Suppose that $\tilde{G}(\tau)$ is a consistent estimate of $G(\tau)$, and for convenience let $\tilde{G}_b(t) = 1 - \tilde{G}(t)$ denote the estimate of $\bar{G}(t)$. In that case, the estimates analogous to (5),

$$\tilde{f}_{\mathrm{SM}}(t) = \frac{n_t}{\tilde{G}_b(t)}, \qquad t = 1, \ldots, \tau_{\max}, \tag{15}$$

and to $\hat{F}_{\mathrm{SM}}(t)$ in (7),

$$\tilde{F}_{\mathrm{SM}}(t) = \sum_{s=1}^{t} \tilde{f}_{\mathrm{SM}}(s), \qquad t = 1, \ldots, \tau_{\max}, \tag{16}$$

are both consistent.

Behavior of the estimator $\tilde{f}_{\mathrm{SM}}(t)$ depends on how well $\tilde{G}(\tau)$ estimates $G(\tau)$ and whether $\tilde{G}(\tau)$ is related to the primary data—that is, the n_t's. In this article, we assume that $\tilde{G}(\tau)$ is independent of the primary data. The covariance of the $\tilde{f}_{\mathrm{SM}}(t)$'s is then

$$\begin{aligned}
&\operatorname{cov}\{\tilde{f}_{\mathrm{SM}}(s_1), \tilde{f}_{\mathrm{SM}}(s_2)\} \\
&\quad = \mathrm{E}\{\operatorname{cov}[\tilde{f}_{\mathrm{SM}}(s_1), \tilde{f}_{\mathrm{SM}}(s_2) | \tilde{G}(\tau)]\} \\
&\qquad + \operatorname{cov}\{\mathrm{E}[\tilde{f}_{\mathrm{SM}}(s_1)|\tilde{G}(\tau)],\ \mathrm{E}[\tilde{f}_{\mathrm{SM}}(s_2)|\tilde{G}(\tau)]\} \\
&\quad = \mathrm{E}\left\{ \frac{f(s_1)\bar{G}(s_1)[\mathrm{I}(s_1 = s_2) - f(s_2)\bar{G}(s_2)]}{\tilde{G}_b(s_1)\tilde{G}_b(s_2)} \right\} \\
&\qquad + \operatorname{cov}\left\{ \frac{\bar{G}(s_1)f(s_1)}{\tilde{G}_b(s_1)}, \frac{\bar{G}(s_2)f(s_2)}{\tilde{G}_b(s_2)} \right\},
\end{aligned} \tag{17}$$

which can be estimated by

$$\begin{aligned}
&\frac{\tilde{f}_{\mathrm{SM}}(s_1)}{\tilde{G}_b(s_2)}\, [\mathrm{I}(s_1 = s_2) - \tilde{f}_{\mathrm{SM}}(s_2)\tilde{G}_b(s_2)] \\
&\qquad + \frac{\tilde{f}_{\mathrm{SM}}(s_1)\tilde{f}_{\mathrm{SM}}(s_2)}{\tilde{G}_b(s_1)\tilde{G}_b(s_2)}\, \widehat{\operatorname{cov}}\{\tilde{G}_b(s_1), \tilde{G}_b(s_2)\},
\end{aligned} \tag{18}$$

assuming that an estimate $\widehat{\operatorname{cov}}\{\tilde{G}_b(s_1), \tilde{G}_b(s_2)\}$ is available. The second term in (17) accounts for variation due to $G(\tau)$ having been estimated. An estimate for the variance of $\tilde{F}_{\mathrm{SM}}(t)$ can be obtained from

$$\operatorname{var}\{\tilde{F}_{\mathrm{SM}}(t)\} = \sum_{s_1=1}^{t} \sum_{s_2=1}^{t} \operatorname{cov}\{\tilde{f}_{\mathrm{SM}}(s_1), \tilde{f}_{\mathrm{SM}}(s_2)\}.$$

In Section 5, we will discuss this further based on the example there.

4. NONINDEPENDENT CENSORING

Section 1 assumes that censoring times $\mathcal{T}_1, \ldots, \mathcal{T}_M$ are independent of lifetimes $T_1, \ldots, T_M$. This assumption may sometimes be questionable: For example, if the lifetime of an automobile component depends on both the age of the car and the number of miles it is driven, then the fact that warranty plans have age and mileage limitations (e.g., two years and 24,000 miles) implies a dependence between T_i

Table 2. Comparisons of Standard Errors ($\widetilde{sd}$) and Average Estimated Standard Deviations ($\widehat{sd}$) of $\hat{F}_{SM}(.8)$ and $\hat{F}_{ESM}(.8)$.

		$\hat{F}_{SM}(.8)$			$\hat{F}_{ESM}(.8)$		
Cases		α = 3.95	α = 1.85	α = 1.02	α = 3.95	α = 1.85	α = 1.02
M = 4,000	$\widetilde{sd}$	.00340	.00623	.00817	.00339	.00622	.00812
	$\widehat{sd}$	.00333	.00639	.00862	.00332	.00636	.00852
M = 400	$\widetilde{sd}$	.01098	.01952	.02777	.01097	.01944	.02756
	$\widehat{sd}$	.01039	.02008	.02731	.01040	.02005	.02705

and $\mathcal{T}_i$. Our objective here is to briefly consider the effect of nonindependent censoring on the estimator of Section 1.2. We also present a version of the simple moment estimator for a special case in this situation.

4.1 Effect on $\hat{f}_{SM}(\cdot)$

The estimator $\hat{f}_{\mathrm{SM}}(\cdot)$ of (5) can be written in the form

$$\hat{f}_{\mathrm{SM}}(t) = \frac{\sum_{i=1}^{M} \mathrm{I}(t_i = t, \tau_i \geq t)}{\sum_{i=1}^{M} \Pr(\mathcal{T}_i \geq t)}.$$

Now $\mathrm{E}\{\mathrm{I}(t_i = t, \tau_i \geq t)\} = f(t)\Pr(\mathcal{T}_i \geq t | T_i = t)$, so if $\Pr(\mathcal{T}_i \geq t|T_i = t) \neq \Pr(\mathcal{T}_i \geq t)$, then $\hat{f}_{\mathrm{SM}}(t)$ is biased, with

$$\mathrm{E}\{\hat{f}_{\mathrm{SM}}(t)\} = f(t)\left\{\frac{\sum_{i=1}^{M} \Pr(\mathcal{T}_i \geq t|T_i = t)}{\sum_{i=1}^{M} \Pr(\mathcal{T}_i \geq t)}\right\}. \quad (19)$$

The extent of the bias may be assessed by hypothesizing models for the dependence of T_i and $\mathcal{T}_i$, and in many cases we may find that the bracketed term in (19) is close to 1. If it is not, there may be little motivation to estimate the marginal distribution $f(t)$; what is needed instead is a model that accounts for the dependence of T and $\mathcal{T}$. With automobiles, this usually means that a lifetime model that incorporates both age and mileage is needed. Lawless, Hu, and Cao (1995) discussed such models and indicated how to test independence of lifetime and censoring time from automobile warranty data. If there is a serious concern in a practical situation about dependence, then such methods should be employed.

4.2 A Special Case

In some situations T_i and $\mathcal{T}_i$ are related only through a covariate (or covariates), say x_i, such that T_i and $\mathcal{T}_i$ are independent given $x_i, i = 1, \ldots, M$. This was considered in different contexts, for example, by Kalbfleisch and Lawless (1991) and Hu and Lawless (1996b). We extend the model of Section 1.3 slightly to deal with this.

As in Section 1.3, we suppose that x_i takes on values $x_k^0, k = 1, \ldots, K$, and is observed for every unit i. Then

$$\hat{f}_{\mathrm{SM}}(t|x_k^0) = \frac{n_{t,k}}{M_k \tilde{G}_k(t)}, \qquad t = 1, \ldots, \tau_{\max}, \quad (20)$$

is an unbiased estimator of $f(t|x_k^0), k = 1, \ldots, K$, where $n_{t,k} = \#\{t_i : t_i \leq \tau_i, t_i = t, x_i = x_k^0\}$. Noting that $f(t) = \sum_{k=1}^{K} f(t|x_k^0)\Pr(X = x_k^0)$, we have an estimator for $f(t)$ and also for $F(t)$, provided that $\Pr(X = x_k^0)$ is known or estimated, $k = 1, \ldots, K$. The changing pattern of $\hat{F}_{\mathrm{SM}}(t|x_k^0) = \sum_{s=1}^{t} \hat{f}_{\mathrm{SM}}(s|x_k^0)$ when the value of x_k^0 varies may help us see how lifetime is related to censoring time. If the dependences between T_i and x_i and $\mathcal{T}_i$ and x_i can be specified parametrically, we can see how the dependence of the failure time and the censoring time affects the simple moment estimator from (19). Parametric models also allow us to handle continuous covariates. Hu and Lawless (1996b) considered this approach.

For a slightly different situation in which only the x_i's associated with units having observed failures can be observed, we may consider the estimator for $f(t|x_k^0)$,

$$\tilde{f}(t|x_k^0) = \frac{n_{t,k}}{\tilde{M}_k \tilde{G}_k(t)}, \qquad t = 1, \ldots, \tau_{\max}, \quad (21)$$

if an estimate $\tilde{M}_k$ is available. We address this in Section 5 through the example. This idea may be applied to situations in which the number of product units in service M is unknown but there is an estimate for it. That is, for example, we consider $\tilde{f}_{\mathrm{SM}}(t) = n_t/\tilde{M}\tilde{G}(t)$ instead of (5).

Similarly, as in Sections 1 and 2, we can also consider variance estimation of $\hat{F}(t|x_k^0), t = 1, \ldots, \tau_{\max}$ and $k = 1, \ldots, K$.

5. AN EXAMPLE

Some real warranty data for a specific system on a particular car model are considered for illustration. The data include warranty claims from 823 cars among 8,394 cars produced during a two-month period. The warranty plan in question was for one year or 12,000 miles; the data here cover up to 18 months after the first car was sold. The numbers of cars sold within the first, second, ..., and the sixth three-month periods were 5,699, 1,593, 725, 227, 129, and 21, respectively, among which 479, 232, 83, 18, 10, and 1 cars had warranty claims recorded. We examine the distribution of the time to the first failure (claim) of the cars. For illustration we consider "time" both as mileage in miles and as age in years (i.e., calendar time), although for engineering purposes mileage is more relevant.

Let t_i and $\tau_i, i = 1, \ldots, M = 8{,}394$, be the first failure times and censoring times, respectively, and let s_i denote the time of sale for car i, where the first car sold has a sale time of 0. Calendar time will be expressed in years, and the warranty data therefore will include follow-up of cars up to time 1.5 years. The censoring time τ_i for car i may be described as follows. Let the age of the car (i.e., time since the car was sold) when it reaches 12,000 miles be a_i years, and define $u_i = 12{,}000/a_i$ as the average mileage accumulation rate over the age interval $(0, a_i]$, in miles per year. Then for the case in which t_i and τ_i represent age (i.e., years since sale), $\tau_i = \min(1.5 - s_i, 1, a_i)$. In the case in which t_i and τ_i represent mileage, $\tau_i = \min(\min[1.5 - s_i, 1]u_i, 12{,}000)$.

The values of t_i's are observed only for those cars with $t_i \leq \tau_i$—that is, for the $m = 823$ cars with their warranty claims recorded. The numbers of first failures experienced in the cars' first, second, ..., twelfth month were 71, 81, 77, 92, 84, 75, 98, 69, 42, 57, 37, and 40, and within their first, second, ..., twelfth 1,000 miles were 80, 87, 70, 80, 65, 77, 59, 65, 45, 66, 63, and 66, respectively. The following investigation analyzes the data more thoroughly.

Although the sale dates s_i's are known for all 8,394 cars, the values of τ_i's are not. If we are willing to make the simplifying assumption that mileage accumulation is linear over $(0, a_i]$, then u_i may be evaluated for cars that fail because the mileage as well as the age at failure is recorded. In this case we would thus have the τ_i's for the cars that fail but not for those that do not. The simple estimators used here do not require any censoring times, but they do require an estimate of their distribution, which we now discuss.

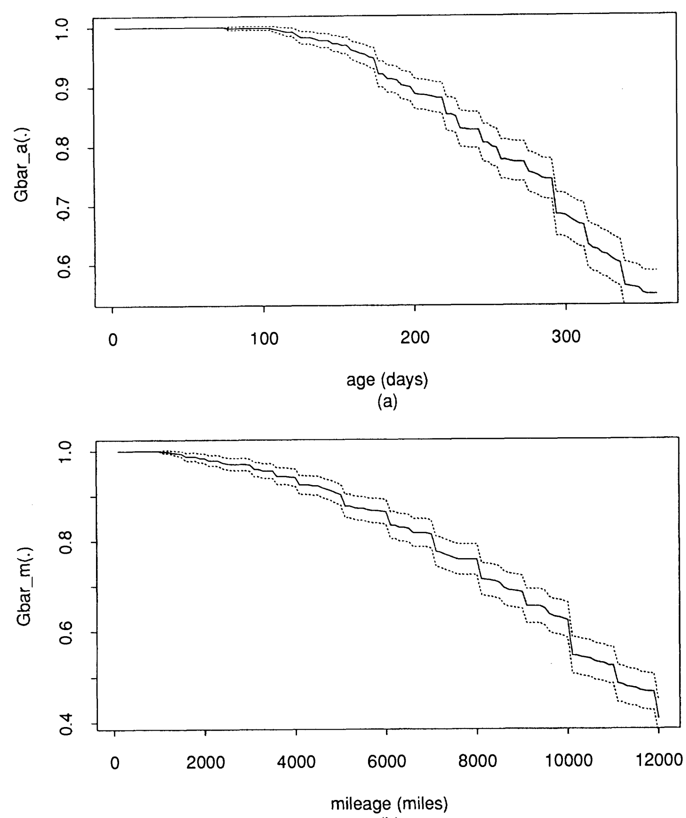

Figure 2. Estimates and Approximate 95% Confidence Intervals for Survival Functions of Censoring Time: (a) Time = Age, (b) Time = Mileage.

A customer survey of 607 cars of the same type and approximate geographic location as those in the warranty database was taken, wherein the approximate mileages at one year after sale were obtained for each car. We assume that mileage accumulation occurs at a constant rate u_i for car i over the first year after sale; this is obviously an oversimplification but is satisfactory for practical purposes in this case. Sale date and mileage-accumulation rate can reasonably be assumed independent, and we know that the survival functions of the censoring time are

$$\bar{G}_a(\tau) = \mathrm{I}(\tau \leq 1)\Pr\left(1.5 - s_i \geq \tau, U_i \leq \frac{12{,}000}{\tau}\right)$$

when T is age at failure and

$$\bar{G}_m(\tau) = \mathrm{I}(\tau \leq 12{,}000)\Pr(U_i \min[1.5 - s_i, 1] \geq \tau)$$

when T is mileage at failure. Then we can estimate the survival distribution of censoring time $\bar{G}(\tau)$ in the warranty database population by using the empirical distribution of sale dates $s_i (i = 1, \ldots, 8{,}394)$ along with the empirical distribution of U_i based on the customer survey. Figure 2 shows estimates and approximate pointwise 95% confidence intervals for survival functions of the censoring time in both time-scale cases.

The moment estimate (7) may thus be computed and is shown in Figure 3, for the case in which failure "times" are measured in miles. Figure 3 also shows approximate pointwise 95% confidence intervals for the failure-time distribution function $F(t)$, obtained as $\hat{F}_{\mathrm{SM}}(t) \pm 1.96\hat{V}(t)^{1/2}$, where $\hat{V}(t)$ is the estimated variance of $\hat{F}_{\mathrm{SM}}(t)$. These intervals are based on the fact that, as M increases, the distribution of $[\hat{F}_{\mathrm{SM}}(t) - F(t)]\hat{V}(t)^{-1/2}$ approaches a standard normal distribution for a given t. Two sets of confidence limits are shown: Intervals I use the variance estimate (8), which assumes that $G(\tau)$ is known; intervals II are based on (18) and account for the fact that $G(\tau)$ has been estimated by using the car survey. The second set of intervals is considerably wider and provides a more valid assessment of uncertainty. We could similarly produce estimates of the failure-time distribution in terms of car age.

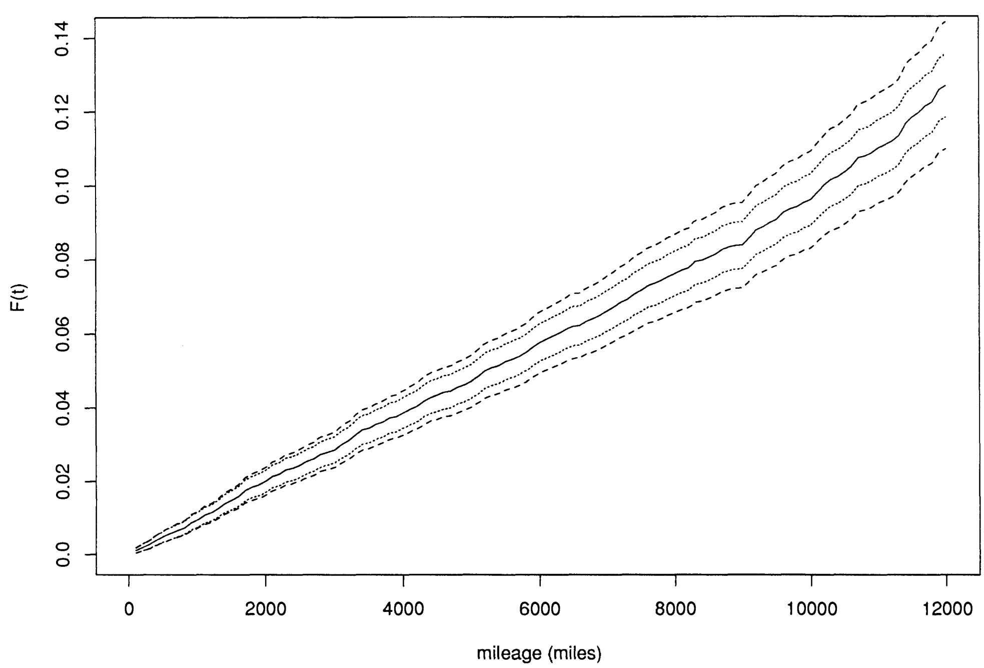

Figure 3. Approximate 95% Confidence Intervals of Failure-Time Distribution (time = mileage): ———, SM Estimate; - - - -, Confidence Interval I; – – –, Confidence Interval II.

We remark that an alternative approach is to stratify cars according to their time of sale and then to use the approach in Section 1.3. This produces an estimate of $F(t)$ that is indistinguishable from that in Figure 3 in this case, which was based on the unstratified data.

It is possible here that failure may be related to both age (time since sale) and mileage. To investigate this we formed a covariate x based on mileage-accumulation rates, as follows. We divided mileage rates into five classes—(0, 6,000], (6,000, 12,000], (12,000, 18,000], (18,000, 24,000], and (24,000, ∞) miles per year—and let $x = k$ denote the kth class ($k = 1, \ldots, 5$). The numbers of failures for cars in the five classes are 92, 266, 245, 109, and 111, respectively. From the customer survey of 607 cars and the car sales data, we estimated the censoring-time distributions $G_k(\tau) = \Pr(T_i \le \tau | x_i = k), k = 1, \ldots, 5$, through

$$\bar{G}_{a,k}(\tau) = \Pr(\min[1.5 - s_i, 1] \ge \tau) \Pr\left(U_i \le \frac{12{,}000}{\tau} \middle| x_i = k\right)$$

for the age case and

$$\bar{G}_{m,k}(\tau) = \mathrm{I}(\tau \le 12{,}000) \times \int \Pr\left(\min[1.5 - s_i, 1] \ge \frac{\tau}{u}\right) d\Pr(U_i \le u | x_i = k)$$

for the mileage case; from the survey data alone, we estimated M_k by $M \Pr(X_i = k)$. The numbers of cars from the survey sample falling into the five groups are 96, 271, 148, 53, and 39, respectively. Finally, we imputed a value of u_i, and thus x_i, for each car that experienced a failure under warranty by dividing the mileage at failure by the age at failure. We then estimate $F_k(t) = \Pr(T_i \le t | x_i = k)$ as

$$\tilde{F}_k(t) = \sum_{s=1}^{t} \frac{n_{s,k}}{\tilde{M}_k \tilde{G}_{b,k}(s)}, \qquad t = 1, \ldots, \tau_{\max,k}, \tag{22}$$

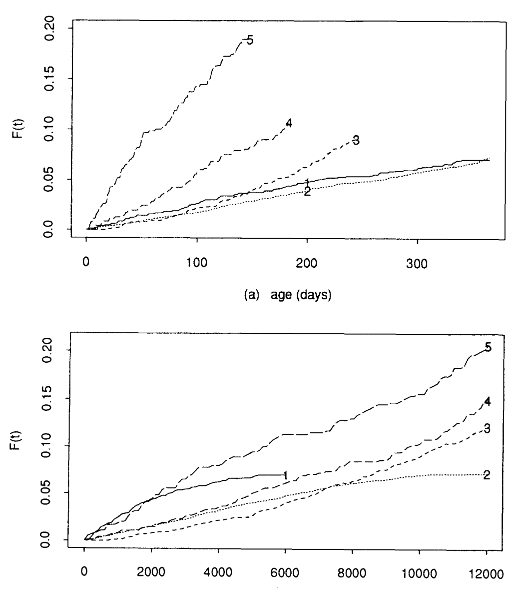

Figure 4. Estimates of Failure-Time Distributions With Different Usage Rates: (a) Time = Age, (b) Time = Mileage.

where $n_{s,k} = \#\{i : t_i = s, x_i = k, \tau_i \ge s\}, \tilde{M}_k = M \Pr(X_i = k)$ with $M = 8{,}394$, and $\tau_{\max,k} = \sup\{\tau : \tilde{G}_{b,k}(\tau) > 0\}, k = 1, \ldots, 5$. Estimates of $\tilde{F}_k(t), k = 1, \ldots, 5$, are presented in Figure 4, (a) and (b), for the cases in which failure time is measured as car age and car mileage, respectively. Bearing in mind that the estimates are not very precise, in part because the estimates of $\tilde{G}_k(\tau)$ and $\Pr(X_i = k)$ are based on rather small samples, Figure 4 suggests that failure times measured in miles do not depend much on the mileage-accumulation rate but that failure times measured as car age do. This suggests that mileage is the more relevant time scale for this type of failure. Lawless et al. (1995) reached a similar conclusion by using parametric models for failure that incorporate both age and mileage as factors.

6. RECURRENT EVENTS AND MULTIPLE FAILURE MODES

Products under warranty are usually repairable systems in which there are multiple types of failure that may occur more than once. The problem discussed in this article can be studied in this broader context, and methods based on maximum likelihood and on moment estimation may be developed. We will merely mention the main ideas, which are discussed elsewhere.

The modeling of recurrent events often uses Poisson or renewal processes (Ascher and Feingold 1984; Lawless 1995). More generally, the mean and rate functions for the recurrent events or failures are of interest. They are defined as follows: Let $N_i(t)$ denote the number of events occurring on unit i over the time interval $(0, t]$. Then $\Lambda(t) = \mathrm{E}\{N_i(t)\}$ is called the mean function and $\lambda(t) = d\Lambda(t)/dt$ is called the rate (or rate of occurrence) function. If the recurrent events follow a Poisson process, then $\lambda(t)$ is also the intensity function.

In the case of recurrent events the "censoring" time τ_i refers to the time period $(0, \tau_i]$ over which unit i is observed. Hu and Lawless (1996a) discussed maximum likelihood estimation under a Poisson model when censoring times are missing for units not experiencing any failures. They also presented a moment estimator for $\lambda(t)$ that is analogous to the ones given for failure-time distributions in Section 1.2 and 1.3 and is of exactly the same form as (5),

$$\hat{\lambda}_{\mathrm{SM}}(t) = \frac{n_t}{M\bar{G}(t)}, \qquad t = 1, \ldots, \tau_{\max},$$

where now, however, n_t is the total number of recurrent events observed at time t across all product units. Hu and Lawless (1996a) gave variance estimates for $\hat{\lambda}_{\mathrm{SM}}(t)$ and $\hat{\Lambda}_{\mathrm{SM}}(t) = \hat{\lambda}_{\mathrm{SM}}(1) + \cdots + \hat{\lambda}_{\mathrm{SM}}(t)$ and discussed their properties.

Multiple failure modes may also be dealt with. For simplicity we consider two modes, A and B, and the case of failure times; recurrent events can also be considered. Let T_i^A and T_i^B represent the times to failure of modes A and B, respectively, let $f_A(t) = \Pr(T_i^A = t)$ and $f_B(t) = \Pr(T_i^B = t)$ denote the marginal probability functions, and let $f_{AB}(s, t) = \Pr(T_i^A = s, T_i^B = t)$ denote the joint probability function of T_i^A and T_i^B. Under the assumption of independent censoring times τ_i, the following are unbiased estimates of $f_A(t)$ and $f_B(t)$:

$$\hat{f}_A(t) = \frac{n^A(t)}{M\bar{G}(t)}, \quad \hat{f}_B(t) = \frac{n^B(t)}{M\bar{G}(t)}, \tag{23}$$

where $n^A(t) = \sum_{i=1}^M \mathrm{I}(T_i^A = t, \tau_i \geq t)$ and $n^B(t) = \sum_{i=1}^M \mathrm{I}(T_i^B = t, \tau_i \geq t)$, and once again we assume that $\bar{G}(\tau) = \Pr(\mathcal{T}_i \geq \tau)$ is known. It is also possible to give a simple moment estimator of $f_{AB}(s,t)$:

$$\hat{f}_{AB}(s,t) = \frac{n^{AB}(s,t)}{M\bar{G}(s \vee t)}, \tag{24}$$

where $n^{AB}(s,t) = \sum_{i=1}^M \mathrm{I}(T_i^A = s, T_i^B = t, \tau_i \geq s \vee t)$ and $s \vee t$ denotes the maximum of s and t. In applications in which the probability of a failure of any given mode is fairly small over the observation period, however, the probability of getting failures on two or more modes is usually very small, so (24) may not be very precise. In many situations it may be adequate simply to consider the different failure modes separately, in which case the estimates (23) are all that are needed. Variance estimates are then given by the expressions for $\hat{f}_{\mathrm{SM}}(t)$ in Section 3.2. If, however, we wish to gain insight into how failure times for different modes are related, (24) can be used. If this is too imprecise to be useful, then one can adopt a parametric model to get more precise (but model-dependent) estimates.

The preceding discussion of multiple failure modes assumes that, when a failure of one type occurs, it does not preclude failures of other types. In some situations the failure modes may be competing so that this does happen.

7. COMMENTS AND RECOMMENDATIONS

When censoring times are missing, standard methods of estimating lifetime distributions are not available. If the censoring-time distribution $G(\tau)$ is known or estimated from additional data, however, then either maximum likelihood or moment estimation may be used to obtain nonparametric estimates. The methods in this article depend on the validity of the assumed $G(\tau)$, and it is important in practice to be confident that $G(\tau)$ is suitable. The use of the easily computed simple estimators $\hat{F}_{\mathrm{SM}}$ and $\hat{F}_{\mathrm{ESM}}$ is entirely satisfactory in practice, as demonstrated by our simulation results. We also recommend the use of confidence limits for the lifetime distribution that account for uncertainty in $G(\tau)$. When $G(\tau)$ is estimated, the method of Section 3 can be employed. If standard errors for the estimate of $G(\tau)$ are not available, we recommend varying $G(\tau)$ in a sensible way around the estimate and examining the range of confidence limits obtained.

The estimates in Section 3 also require that the censoring times be independent of lifetimes. This can be a problem for some applications. As shown in Sections 4 and 5, we can often handle dependent censoring by using a covariate x such that lifetimes and censoring times are roughly independent conditional on x. In the case of automobile warranty data, the mileage-accumulation (or usage) rate fulfills this function.

If censoring times are available for units that fail, then inferences about the lifetime distribution may be obtained by considering the distribution of t_i given that $t_i \leq \tau_i$ for failed units. This gives the truncated data-likelihood function

$$L_T = \prod_{t_i \leq \tau_i} \frac{f(t_i)}{F(\tau_i)} \tag{25}$$

instead of (1). It is well known that for parametric models $f(t;\theta)$ the likelihood (25) gives much less precise estimates of θ than do methods that use information about the censoring times for unfailed units (Kalbfleisch and Lawless 1988; Hu and Lawless 1996b). Thus, the use of (1) to estimate θ with a parametric model would be much preferred to (25). The same holds true for nonparametric estimation of $f(t)$; Kalbfleisch and Lawless (1991) discussed nonparametric estimation based on (25), but the methods of this article are to be preferred. The price, of course, is that a good estimate of the censoring time distribution $G(\tau)$ must be obtained.

Finally, in contexts such as manufacturing, one may sometimes wish to make estimates for a finite population of units. The estimates of $f(t)$ given here can be used to do this. In principle, finite-population corrections to variance estimates can be made, but given the large size of the typical populations, it makes no practical difference if these are ignored.

ACKNOWLEDGMENTS

This research was supported in part by grants to the second author from General Motors Canada, the Manufacturing Research Corporation of Ontario, and the Natural Sciences and Engineering Research Council of Canada (NSERC), and by a fellowship to the first author from NSERC. We thank K. Adachi, T. Kasashima, J. MacKay, two referees, and an associate editor for helpful comments.

[Received June 1996. Revised March 1997.]

REFERENCES

Ascher, H., and Feingold, H. (1984), *Repairable System Reliability*, New York: Marcel Dekker.

Hu, X. J., and Lawless, J. F. (1996a), "Estimation of Rate and Mean Functions From Truncated Recurrent Event Data," *Journal of the American Statistical Association*, 91, 300–310.

——— (1996b), "Estimation From Truncated Lifetime Data With Supplementary Information on Covariates and Censoring Times," *Biometrika*, 83, 747–761.

Kalbfleisch, J. D., and Lawless, J. F. (1988), "Estimation of Reliability From Field Performance Studies" (with discussion), *Technometrics*, 30, 365–388.

——— (1991), "Regression Models for Right Truncated Data With Applications to AIDs Incubation Times and Reporting Lags," *Statistica Sinica*, 1, 19–32.

Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: Wiley.

——— (1995), "The Analysis of Recurrent Events for Multiple Subjects," *Applied Statistics*, 44, 487–498.

Lawless, J. F., Hu, X. J., and Cao, J. (1995), "Methods for the Estimation of Failure Distributions and Rates From Automobile Warranty Data," *Lifetime Data Analysis*, 1, 227–240.

Suzuki, K. (1985), "Estimation of Lifetime Parameters From Incomplete Field Data," *Technometrics*, 27, 263–271.

——— (1988), "Analysis of Field Lifetime Data Drawn From Different Observation Periods," *Hinshitsu (Quality)*, 18, 4–14 (in Japanese).

Suzuki, K., and Kasashima, T. (1993), "Estimation of Lifetime Distribution From Incomplete Field Data With Different Observational Periods," Technical Report UEC-CAS-93-02, University of Electro-Communications, Japan, Dept. of Communications and Systems Engineering.