ESTIMATES AND CONFIDENCE INTERVALS FOR MEDIAN AND MEAN LIFE IN THE PROPORTIONAL HAZARD MODEL

by

Dorota M. Dabrowska

Carnegie-Mellon University and University of California, Berkeley

and

Kjell A. Doksum

University of California, Berkeley

Technical Report No. 67
July (1986)
(revised November 1986)

Research Partially Supported by
National Science Foundation Grant DMS-85-41787 and
National Institute of General Medical Sciences
Grant SSS-Y1R01GM35416-01

Department of Statistics
University of California
Berkeley, California

Estimates and confidence intervals for median and mean life in the proportional hazard model

By Dorota M. Dabrowska

Department of Statistics

Carnegie-Mellon University

Pittsburgh, Pennsylvania 15213 USA

and

Kjell A. Doksum

Department of Statistics

University of California

Berkeley, California 94720

SUMMARY

Estimation procedures and confidence intervals are given for the median and mean survival time in the proportional hazard regression model. For median survival, the methods apply to censored data. The procedures are based on Cox's partial likelihood estimates of the linear model parameters in the loglinear proportional hazard model and on Breslow's estimate of the baseline hazard function. The asymptotic properties of these semiparametric estimates are developed and they are compared with the optimal parametric estimates for the Weibull regression model. For the parameter values considered, the more generally valid semiparametric estimate of mean survival loses little efficiency relative to the optimal parametric estimates in this model unless the Weibull shape parameter is close to zero. The efficiency loss for the semiparametric estimate of median survival is greater but not severe unless censoring is heavy. We also compare the optimal parametric estimates of mean and median survival with the Weibull shape parameter known and unknown and find that the efficiency loss for the unknown shape estimate is small.

Key words: median and mean regression, partial likelihood, proportional hazard model, Weibull model.

1. INTRODUCTION

Cox's (1972,1975) proportional hazard model with loglinear hazard specifies the failure rate for the survival time Y of an individual with covariate vector $x = (x_1, \cdots, x_r)^T$ as

$$\lambda(t\ ; x) = \lambda_0(t)\exp(\beta^T x). \tag{1}$$

Here $\beta = (\beta_1, \cdots, \beta_r)^T$ is a vector of unknown regression coefficients and $\lambda_0(t)$ is the unknown baseline hazard function.

The influence of the covariates on survival is measured by β since β_i represents the increase in log hazard as x_i is increased one unit. In some applications, it is also useful to consider how the median or mean survival time is affected by the covariates and in this case we need methods for estimating or predicting conditional median or mean survival time given a value of the covariate vector. For instance, a smoker may be interested in by how much stopping smoking will increase his or her mean survival time and researchers may want to know by how much a certain medication increases the median life of patients.

Miller & Halpern (1982) considered the estimate of the conditional median survival time obtained by solving $\hat{F}(t \mid x) = 1/2$ for t, where

$$\hat{F}(t \mid x) = \exp\{-e^{\hat{\beta}^T x} \int_0^t \hat{\lambda}_0(s)ds\}$$

and $\hat{\beta}$ and $\hat{\lambda}_0$ are Cox's (1972,1975) and Breslow's (1972,1974) estimates of β and λ_0, respectively. They computed this estimate for the Stanford heart transplant data. In the case of random right censoring, we derive the asymptotic distribution for this estimate as well as for estimates of the p-th quantile and, for uncensored data, for the mean of the conditional survival time. These results are used to construct approximate confidence intervals for the pth quantile and the mean of the survival distribution.

The semiparametric regression methods developed by Cox (1972, 1975), Breslow (1972, 1974), and others are now frequently used to analyse survival data. However, many studies still use parametric methods, in particular methods based on the Weibull model. Thus it is of interest to compare the more generally valid semiparametric estimates with the optimal parametric estimates for the Weibull model, both when the Weibull shape parameter α is known and when it is unknown.

The expressions for the efficiencies that provide these comparisons are too cumbersome to yield much insight except when there is only one covariate ($p = 1$) and when the regression parameter β is near zero. In this case we find that the more generally valid semiparametric estimate of the mean of the conditional survival time loses very little efficiency in fact, it is fully efficient in the exponential case ($\alpha = 1$), and the efficiency is not less than 92% for $\alpha \geq 0.5$, and not less than 99% for $\alpha \geq 1$ when compared with the α unknown Weibull model estimate.

In the case of the semiparametric estimate of median survival based on censored data, we are able to evaluate the efficiency relative to the α known Weibull model estimate and we find that the efficiency is never below 48% when censoring does not exceed 67%. Let $\mu = E(X)$, $\sigma^2 = var(X)$ and $z = (x - \mu)\sigma^{-1}$. Then as $|z| \to \infty$, the efficiency tends to one. In the uncensored case, the efficiency relative to the α unknown Weibull estimate is at least 77% for all $z \geq 1$, and at least 88% for all $z \geq 2$.

In the one sample case, confidence intervals for the median survival time based on right censored observations have been considered by Brookmeyer & Crowley (1982), Slud, Byar & Green (1984), and Jennison & Turnbull (1985), among others. The asymptotic theory of the one-sample estimate of the mean based on the Kaplan-Meier (1958) estimate of the survival function has been developed by Susarla & van Ryzin (1980) and Gill (1983).

In the proportional hazard model with censored data and a parametric baseline hazard function, Borgan (1984) has developed the asymptotic estimation theory. See Borgan for further references to the parametric case.

2. MEDIAN AND QUANTILE SURVIVAL TIME

2.1. *Estimates and confidence intervals*

Let $(Y_1, X_1^T)^T, \cdots, (Y_n, X_n^T)^T$ be a sampie of independent vectors and suppose that given $X_i = x_i$, $i = 1, \cdots, n$, $Y_1, \cdots, Y_n$ follow model (1). We do not observe the Y_i's, but $T_i = \min\{Y_i, C_i\}$ and $\delta_i = I\{T_i = Y_i\}$. Define $N_i(t) = I\{T_i \le t, \delta_i = 1\}$ and

$J_i(t) = I\{T_i \ge t\}$. Conditionally on $X_i = x_i$, $i = 1, \cdots, n$, Cox's (1972, 1975) partial likelihood can be written

$$L(\beta) = \prod_{i=1}^{n} \prod_{s \ge 0} \left\{ \frac{J_i(s)\exp(\beta^T x_i)}{\sum_{j=1}^{n} J_j(s)\exp(\beta^T x_j)} \right\}^{dN_i(s)}$$

Cox's partial likelihood estimator $\hat{\beta}$ is the value of β which maximizes $L(\beta)$.

Let $y_{(1)} < \cdots < y_{(k)}$ denote the observed ordered distinct survival times; let $y_{(0)} = 0$, $y_{(k+1)} = \infty$ and set $N(s) = \sum_{j=1}^{n} N_j(s)$. For $y_{(i-1)} < t \le y_{(i)}$, Breslow's (1972, 1974) estimate of the baseline hazard rate $\lambda_0(t)$ is

$$\hat{\lambda}_0(t) = \frac{N(y_{(i)}) - N(y_{(i-1)})}{\{y_{(i)} - y_{(i-1)}\} \sum_{j=1}^{n} J_j(t)\exp(\hat{\beta}^T x_j)}, \quad i = 1, \cdots, k+1.$$

To estimate the integrated hazard

$$\Lambda_0(t) = \int_0^t \lambda_0(s)\,ds,$$

we use

$$\hat{\Lambda}_0(t) = \int_0^t \hat{\lambda}_0(s)\,ds.$$

Following Andersen & Gill (1982), set

$$S^{(0)}(\beta,t) = \frac{1}{n}\sum_{i=1}^{n} J_i(t)\exp(\beta^T X_i)$$
$$S^{(1)}(\beta,t) = \frac{1}{n}\sum_{i=1}^{n} X_i J_i(t)\exp(\beta^T X_i)$$
$$S^{(2)}(\beta,t) = \frac{1}{n}\sum_{i=1}^{n} X_i^{\otimes 2} J_i(t)\exp(\beta^T X_i)$$
$$E(\beta,t) = \{S^{(0)}(\beta,t)\}^{-1} S^{(1)}(\beta,t)$$
$$V(\beta,t) = \{S^{0}(\beta,t)\}^{-1} S^{(2)}(\beta,t) - E(\beta,t)^{\otimes 2}$$

where for any r-vector $a = (a_1, \cdots, a_r)^T$, $a^{\otimes 2}$ is an $r \times r$ matrix with (i,j) entry equal to $a_i a_j$. Assume conditions A-D of Andersen & Gill (1982), in particular that $S^{(k)}(\beta,t)$ converges in probability to $s^{(k)}(\beta,t)$, $0 \le t \le T$, $k = 0,1,2$, uniformly in t and in a neighbourhood of β_0, where β_0 denotes the true value of β. Here $[0, T]$ is the time interval over which the individuals are observed. Set

$$e(\beta,t) = \{s^{(0)}(\beta,t)\}^{-1} s^{(1)}(\beta,t)$$
$$v(\beta,t) = \{s^{(0)}(\beta,t)\}^{-1} s^{(2)}(\beta,t) - e(\beta,t)^{\otimes 2}.$$

For any covariate value x_0 and $p \in (0,1)$, let $Q(p \mid x_0) = \Lambda_0^{-1}[\{-\log(1-p)\}\exp(-\beta^T x_0)]$ be the p-th quantile of the conditional distribution of Y given x_0, where $\Lambda_0^{-1}(u) = inf\{t : \Lambda_0(t) \ge v\}$. A natural predictor of the p-th quantile is

$$\hat{Q}(p \mid x_0) = \hat{\Lambda}_0^{-1}[\{-\log(1-p)\}\exp(-\hat{\beta}^T x_0)].$$

Using the results of Begun, Hall, Huang & Wellner (1983), it follows that this estimate is asymptotically optimal in the sense of having the smallest possible asymptotic variance in a given class of regular estimates. We next give this asymptotic variance.

For $0 < p_1 \le p_2 < 1$, let $q_i = Q(p_i \mid x_0)$, $r_i = \Lambda_0(q_i)$,

$$\xi_i = \int_0^{q_i} e(\beta_0,u)\, d\Lambda_0(u),$$

and

$$\Sigma = \int_0^T v(\beta_0, v) s^{(0)}(\beta_0, u) d\Lambda_0(u).$$

PROPOSITION 2.1. *Assume that* $0 < a < b < 1$ *are points such that* $\lambda_0(t)$ *is continuous and bounded away from zero for* $t \in [Q(a \mid x_0),\ Q(b \mid x_0)]$. *Then* $\hat{U}(p \mid x_0) = n^{1/2}\{\hat{Q}(p \mid x_0) - Q(p \mid x_0)\}$ *converges weakly in* $D[a, b]$ *to a mean zero Gaussian process* $U(p \mid x_0)$ *with covariance*

$$\tau^2(p_1, p_2 \mid x_0) = [\lambda_0(q_1)\lambda_0(q_2)]^{-1}\gamma(q_1, q_2 \mid x_0)$$

where

$$\gamma(q_1, q_2 \mid x_0) = \{\int_0^{q_1} [s^0(\beta_0, u)]^{-1} d\Lambda_0(u) \tag{2}$$

$$+ \xi_1^T \Sigma^{-1} \xi_2 - r_2 x_0^T \Sigma^{-1} \xi_1 - r_1 x_0^T \Sigma^{-1} \xi_2 + r_1 r_2 x_0^T \Sigma^{-1} x_0\}$$

The proof is deferred to the appendix.

Note that, using Tsiatis (1981) and Andersen & Gill (1982), a consistent estimate $\hat{\tau}^2(p_1, p_2 \mid x_0)$ of $\tau^2(p_1, p_2 \mid x_0)$ results if in (2) the unknowns λ_0, Λ_0, $s^{(0)}$, e, and v are replaced by their estimates $\hat{\lambda}_0$, $\hat{\Lambda}_0$, $S^{(0)}$, E and V as defined earlier. Thus, we can conclude

COROLLARY 2.1. *Under the conditions of Proposition 2.1, an approximate level* $(1 - \alpha)$ *confidence interval for the conditional quantile survival time* $Q(p \mid x_0)$ *is* $\hat{Q}(p \mid x_0) \pm z_{\alpha/2}\hat{\tau}(p, p \mid x_0)$, *where* $z_{\alpha/2}$ *is the* $(1 - \alpha/2)$*th quantile of the standard normal distribution.*

We let $\nu(x_0)$ denote the median $Q\left(1/2 \mid x_0\right)$ of Y given x_0, and we let $\tau^2(\hat{\nu} \mid x_0)$ denote the asymptotic variance of the estimate $\hat{\nu}(x_0) = \hat{Q}\left(1/2 \mid x_0\right)$. Thus $\tau^2(\hat{\nu} \mid x_0) = \tau^2\left(1/2, 1/2 \mid x_0\right)$ can be obtained from (2) by substituting $p_1 = p_2 = 1/2$, and an approximate level $(1-\alpha)$ confidence interval for $\nu(x_0)$ is $\hat{\nu}(x_0) \pm z_{\alpha/2}\hat{\tau}(\hat{\nu} \mid x_0)$, where $\hat{\tau}(\hat{\nu} \mid x_0) = \hat{\tau}^2\left(1/2, 1/2 \mid x_0\right)$.

Proposition 2.1 implies that the efficiency of the semiparametric estimate with respect to parametric estimates is positive. We now look closer at such efficiencies.

2.2. *Comparisons with parametric methods in the Weibull model with known shape parameter.*

We consider experiments with one covariate $(r = 1)$ which we denote X. The performance of the semiparametric median life estimate $\hat{\nu}(x_0)$ will be compared with parametric estimate appropriate for the Weibull model with baseline hazard function $\lambda_0(t) = \alpha\lambda^\alpha t^{\alpha-1}$, $t > 0$, $\alpha > 0$, $\lambda > 0$. Although our results in Section 2.1 are valid for arbitrary β_0, we find that the expression for the approximate variance of $\hat{\nu}(x_0)$ is too cumbersome to yield any useful insights except in the case $\beta_0 = 0$. Thus, as is the case with many investigations of properties of partial likelihood methods, we restrict attention to the case $\beta_0 = 0$. The results for this case can be expected to be approximately correct for β_0 in a neighbourhood of zero.

Furthermore, we suppose that the conditional survival function $G(t \mid x_0)$ of the censoring variable C given x_0 has form $G(t \mid x_0) = \exp\{-\rho^\alpha t^\alpha\}$, where $\rho > 0$. Let

μ and σ^2 denote the mean and variance of X, and set $z = |x_0 - \mu|\,\sigma^{-1}$. Then the asymptotic variance of $\hat{\nu}(x_0)$ is

$$\tau^2(\hat{\nu} \mid x_0) = \alpha^{-2}\lambda^{-2}(\log 2)^{2/\alpha}\ \{\theta^{-1}(2^{\theta} - 1)(log\ 2)^{-2} + \theta z^2\},$$

where $\theta = 1 + (\rho/\lambda)^{\alpha}$.

In the case when it is known that $\lambda_0(t) = \alpha\lambda^{\alpha}t^{\alpha-1}$, where α is known and λ in unknown, an asymptotically optimal predictor of the median of Y given x_0 is $\hat{\nu}_{\alpha}(x_0) = \hat{\lambda}^{-1}\ exp\,(-\hat{\beta}x_0/\alpha)\log 2)^{1/\alpha}$. Here $\hat{\beta}$ and $\hat{\lambda}$ are the maximum likelihood estimates of β and λ. The asymptotic variance of $\hat{\nu}_{\alpha}(x_0)$ is given by

$$\tau^2(\hat{\nu}_{\alpha} \mid x_0) = \alpha^{-2}\lambda^{-2}(\log 2)^{2/\alpha}\ \theta(1 + z^2).$$

Comparing the variance of the semiparametric estimate $\hat{\nu}(x_0)$ to the variance of the Weibull model estimate $\hat{\nu}_{\alpha}(x_0)$ in the Weibull model, we find that the asymptotic efficiency $e\,(\hat{\nu}, \hat{\nu}_{\alpha}) = \tau^2(\hat{\nu}_{\alpha} \mid x_0)/\tau^2(\hat{\nu} \mid x_0)$ depends on λ, ρ and α through θ only. This parameter measures heaviness of censoring in the sense that the probability of an individual being censored is $p = 1 - \theta^{-1}$. Furthermore, the asymptotic efficiency depends on the underlying distribution of the covariate X and the value x_0 through $z = |x_0 - \mu|\sigma^{-1}$ only.

As $z \rightarrow \infty$, $e\,(\hat{\nu}, \hat{\nu}_{\alpha})$ tends to 1 so that no efficiency is lost by using the more generally valid estimate $\hat{\nu}$. On the other hand, as $z \rightarrow 0$, $e\,(\hat{\nu}, \hat{\nu}_{\alpha})$ tends to $\theta^2(\log 2)^2\ (2^{\theta} - 1)^{-1}$. This limit is equal to 0.4805 when there is no censoring ($\theta = 1$) and tends to 0 as $\theta \rightarrow \infty$.

Next we consider a table of the asymptotic efficiencies.

Table 1 about here

Table 1 shows that the loss in efficiency in using the more generally valid semiparametric estimate is at most 52% when the probability that an individual is censored is at most 67%. Only when the censoring is extremely heavy does the parametric procedure perform considerably better than the semiparametric procedure. For large values of z, the efficiency of the semiparametric procedure is quite high. It does not fall below 82% for $z \geq 2$ and $p \leq 2/3$.

2.3 Comparison with parametric methods with unknown shape parameter.

In this section we compare the semiparametric estimate with the maximum likelihood estimate in the case where both the shape parameter α and the scale parameter λ, as well as the regression parameter β, are unknown. For simplicity, we assume that the data are uncensored.

We compare the semiparametric estimate of median regression with the parametric estimate

$$\hat{\nu}_W(x_0) = \hat{\lambda}^{-1} exp\,(-\hat{\beta} x_0/\hat{\alpha})(\log 2)^{1/\hat{\alpha}},$$

where $\hat{\beta}$, $\hat{\lambda}$, and $\hat{\alpha}$ are the maximum likelihood estimates of β, λ and α.

The asymptotic variance of this estimate is given by

$$\tau^2(\hat{\nu}_W \mid x_0) = \alpha^{-2}\, \lambda^{-2}(\log 2)^{2/\alpha}\, (dc^{-1} + 1 + z^2)$$

where $z = |x_0 - \bar{\mu}|\, \sigma^{-1}$, $c = 1 + \Gamma''(2) - \{\Gamma'(2)\}^2$ and $d = \{\log(\log 2) - \Gamma'(2)\}^2$.

We find that the asymptotic efficiency $e(\hat{\nu}, \hat{\nu}_W)$ evaluated at $\beta = 0$, does not depend on λ and α.

As $z \to \infty$, $e(\hat{\nu}, \hat{\nu}_W)$ tends to one so that no efficiency is lost by using the more generally valid semiparametric estimate. As $z \to 0$, $e(\hat{\nu}, \hat{\nu}_W)$ tends to $(1 + dc^{-1})(\log\ 2)^2 = 0.6624$, which is much higher than the value 0.4805 in the case of $e(\hat{\nu}, \hat{\nu}_\alpha)$. For $z = 0, 0.5, 1.0, 2.0, 4.0$, and 10.0, the values of the efficiency $e(\hat{\nu}, \hat{\nu}_W)$ are 0.6624, 0.6986, 0.7720, 0.8845, 0.9611 and 0.9931, respectively. Thus for $z \geq 1$, the loss in efficiency using the more generally valid semiparametric estimate is at most 23%.

Finally, note that the Weibull model can be thought of as a power transformation model where $h(Y) = Y^\alpha$ follow an exponential model with survival function $F(t \mid x_0) = exp(-t\, \lambda^\alpha e^{\beta x_0})$. From the earlier results of this section, we find that for $z = 0, 0.5, 1.0, 2.0, 4.0, 10.00$, the asymptotic efficiency $e(\hat{\nu}_W, \hat{\nu}_\alpha)$ is 0.7253, 0.7675, 0.8407, 0.9296, 0.9782, and 0.9963, respectively. Thus for $z \geq 1$, the loss in efficiency when estimating α is at most 7%.

3. MEAN SURVIVAL TIME

3.1 *Estimates and confidence intervals*

In this section we assume that the data are uncensored. For any covariate value x_0

let $\mu(x_0)$ denote the mean of Y given x_0, i.e.

$$\mu(x_0) = \int_0^\infty F(s \mid x_0) ds$$

where $F(s \mid x_0)$ is the conditional survival function of Y given $X = x_0$. A natural estimator of $\mu(x_0)$ is

$$\hat{\mu}(x_0) = \int_0^\infty \hat{F}(s \mid x_0) ds$$

where $\hat{F}(s \mid x_0) = \exp[-\exp\{\hat{\beta}^T x_0\} \hat{\Lambda}_0(s)]$. From Tsiatis (1981) and Andersen & Gill (1982), it follows that $n^{1/2}\{\hat{\mu}(x_0) - \mu(x_0)\}$ is asymptotically mean zero normal with variance

$$\tau^2(\hat{\mu} \mid x_0) = \int_0^\infty \int_0^\infty F(q_1 \mid x_0) F(q_2 \mid x_0) e^{2\beta^T x_0} \gamma(q_1, q_2 \mid x_0) dq_1 dq_2$$

where $\gamma(q_1, q_2 \mid x_0)$ is defined in equation (2).

It follows from this that a level $(1-\alpha)$ confidence interval for $\mu(x_0)$ is $\hat{\mu}(x_0) \pm z_{\alpha/2}$ $\hat{\tau}^2(\hat{\mu} \mid x_0)$, where $\hat{\tau}^2(\hat{\mu} \mid x_0)$ is obtained from $\tau^2(\hat{\mu} \mid x_0)$ by replacing $s^{(0)}$, $s^{(1)}$, $s^{(2)}$, e v, $F(s \mid x_0)$ and β_0 by their sample counterparts defined earlier.

3.2 *Comparison with parametric methods in the Weibull model with known shape parameter.*

We compare the performance of $\hat{\mu}$ with that of the parametric estimate appropriate for the Weibull model with α known. As in Section 2.2, consider the case of one covariate and $\beta_0 = 0$. The asymptotic variance of $\hat{\mu}(x_0)$ is then given by

$$\tau^2(\hat{\mu} \mid x_0) = \lambda^{-2}\, \alpha^{-2}\{2\alpha\Gamma(2/\alpha) - \Gamma(1/\alpha)^2(1 - \alpha^{-2}z^2)\},$$

where $z = |x_0 - \mu|\, \sigma^{-1}$. In the case when it is known that $\lambda_0(t) = \alpha\lambda^\alpha t^{\alpha-1}$, an asymptotically optimal predictor of the mean of Y given x_0 is $\hat{\mu}_\alpha = \Gamma(1/\alpha)\exp(-\hat{\beta}x_0/\alpha)/\hat{\lambda}\alpha$ where $\hat{\beta}$ is the maximum likelihood estimate of β. Its asymptotic variance is given by

$$\tau^2(\hat{\mu}_\alpha \mid x_0) = \lambda^{-2}\, \alpha^{-4}\Gamma(1/\alpha)^2(1 + z^2)$$

Comparing the variance of the semiparametric estimate $\hat{\mu}(x_0)$ to the variance of the Weibull model estimate $\hat{\mu}_\alpha(x_0)$ in the Weibull model, we find that the asymptotic efficiency $e(\hat{\mu}, \hat{\mu}_\alpha) = \tau^2(\hat{\mu}_\alpha \mid x_0)/\tau^2(\hat{\mu} \mid x_0)$ does not depend on λ. Moreover, the asymptotic efficiency depends on the underlying distribution of the covariate and the value x_0 through $z = |x_0 - \mu|\sigma^{-1}$ only.

We note that in the exponential model $(\alpha = 1)$, the more generally valid semiparametric estimate $\hat{\mu}(x_0)$ of the mean regression is fully efficient for any value of z.

As $z \to \infty$, $e(\hat{\mu}, \hat{\mu}_\alpha)$ tends to 1 so that again no efficiency is lost by using the more generally valid estimate $\hat{\mu}$. On the other hand, as $z \to 0$, $e(\hat{\mu}, \hat{\mu}_\alpha)$ tends to $\alpha^{-2}\{2\alpha\Gamma(2/\alpha)\Gamma(1/\alpha)^{-2} - 1\}^{-1}$. This limit is equal to 1 for $\alpha = 1$, tends to 0 as $\alpha \to 0$ or $\alpha \to \infty$. However, this convergence is very slow and for α in the range $0.5 \leq \alpha \leq 100$, the efficiency is in the interval from 0.6166 to 1.

Next we give a table showing asymptotic efficiencies of the semiparametric and parametric estimates of conditional mean survival

Table 2 about here

Table 2 shows that the efficiency of the semiparametric estimate of the mean regression is quite high and much higher than the semiparametric estimate of the median regression. In fact for $z \geq 1$ and $0.5 \leq \alpha \leq 1.5$, the efficiency is 0.89 or higher.

3.3 *Comparison with the parametric methods in the Weibull model with unknown shape parameter.*

We compare the semiparametric estimate of the mean regression with the parametric estimate $\hat{\mu}_W(x_0) = \Gamma(1/\hat{\alpha}) exp\,(-\hat{\beta} x_0/\hat{\alpha})/\hat{\lambda}\hat{\alpha}$, where $\hat{\beta}$, $\hat{\lambda}$ and $\hat{\alpha}$ are the maximum likelihood estimates of β, λ and α. The asymptotic variance of this estimate is given by

$$\tau^2(\hat{\mu}_W \mid x_0) = \lambda^{-2}\alpha^{-4}\Gamma^2(1/2)[1 + z^2 + \{\Gamma'(2) - \psi(1 + 1/\alpha)\}^2 c^{-1}]$$

where $\psi(x)$ is the digamma function $\psi(x) = \Gamma'(x)/\Gamma(x)$ and c is defined as in Section 2.3.

The asymptotic efficiency $e\,(\hat{\mu}, \hat{\mu}_W)$ does not depend on λ. For varying α and z, the entries of Table 3 show that the more generally valid semiparametric estimate is very efficient for α values in the range from 0.5 to 10. In fact, when $\alpha = 1$ no efficiency is lost, and for α between 1 and 10 the efficiency loss is less than 1% for all values of z.

As $z \rightarrow \infty$, the asymptotic efficiency $e\,(\hat{\mu}, \hat{\mu}_W)$ tends to 1 for all values of α. As $z \rightarrow 0$, the efficiency tends to

$\Gamma^2(1/\alpha)[1 + \{\Gamma'(2) - \psi(1 + 1/\alpha)\}^2/c\,]\ \alpha^{-2}\{2\alpha\Gamma(2/\alpha) - \Gamma(1/\alpha)^2\}^{-1}$, values of which are given in Table 2.

Finally, the efficiency tends to 0 as $\alpha \rightarrow 0$ or $\alpha \rightarrow \infty$. However, for $\alpha \rightarrow \infty$, this convergence is very slow

3.4. *Comparison between the parametric methods in the Weibull model.*

Finally, we consider the asymptotic efficiency $e(\hat{\mu}_W, \hat{\mu}_\alpha)$ of the parametric estimate with the power parameter α unknown with the parametric estimate with α known. As $z \to \infty$, this efficiency tends to 1. Further, for $\alpha = 1$ it is exactly equal to 1 for all z. For $z \to 0$, the asymptotic efficiency tends to $[1 + \{\Gamma'(2) - \psi(1 + 1/\alpha)\}^2/c]^{-1}$. For $\alpha \to 0$, the asymptotic efficiency tends to 0. For $\alpha \to \infty$, the asymptotic efficiency tends to $(1 + z^2)\{1 + z^2 + c^{-1}\}^{-1}$. For $z = 0.0, 0.5, 1.0, 2.0, 4.0$ and 10, this limit is equal to 0.6219, 0.6728, 0.7669, 0.8916, 0.9655, 0.9940, respectively.

Thus the loss in efficiency due to estimating the power parameter is small to moderate unless α is small, in which case the loss is severe. Similar findings have been obtained by Carrol & Ruppert (1981) and Taylor (1986) in the case of power transformations to a linear regression model.

APPENDIX

Proof of Proposition 2.1

From Tsiatis (1981) and Andersen & Gill (1982), $\hat{L}(t) = n^{1/2}\{\hat{\Lambda}_0(t) - \Lambda_0(t)\}$ converges weakly in $D[0,T]$ to a mean zero Gaussian process $L(t)$ with covariance

$$cov\{L(s), L(t)\} = \int_0^s \frac{\lambda_0(u)}{s^{(0)}(\beta_0, u)} du + \Big[\int_0^s e(\beta_0, u)\lambda_0(u) du\Big]^T \Sigma^{-1}\Big[\int_0^t e(\beta_0, u)\lambda_0(u) du\Big]$$

for $s \le t$. Without loss of generality let us assume that the sample paths of $L(t)$ are continuous and $\hat{L}(t)$ converges to $L(t)$ almost surely in the supremum norm on $[0,T]$.

Set $a' = \Lambda_0(Q(a \mid x_0))$ and $b' = \Lambda_0(Q(b \mid x_0))$. We shall show that $n^{1/2}(\hat{\Lambda}_0^{-1} - \Lambda_0^{-1}) = \hat{U}$ converges weakly to $-\left[\lambda_0\{\Lambda_0^{-1}(u)\}\right]^{-1} L\{\Lambda_0^{-1}(u)\}$ in $D[a',b']$. The Proposition follows then by Taylor expansion and Theorems 3.2 and 3.4 in Andersen & Gill (1982). We can write

$$\hat{U}(u) = -\frac{\hat{V}(u)}{\lambda_0\{\Lambda_0^{-1}(u)\}} + \hat{V}(u)\left[\frac{\hat{U}(u)}{\hat{V}(u)} + \frac{1}{\lambda_0\{\Lambda_0^{-1}(u)\}}\right] \tag{3}$$

where $\hat{V}(u) = n^{1/2}[\Lambda_0\{\hat{\Lambda}_0^{-1}(u)\} - u]$. By the uniform continuity of $\lambda_0\{\Lambda_0^{-1}(u)\}^{-1}$, it is enough to show that $\hat{V}(u)$ converges in probability to $V(u) = L\{\Lambda_0^{-1}(u)\}$ and the second term in (3) is asymptotically negligible.

Adding and subtracting terms

$$\sup|\hat{V}(u) - V(u)| \le \sup|\hat{L}\{\hat{\Lambda}_0^{-1}(u)\} - L\{\hat{\Lambda}_0^{-1}(u)\}| \tag{4}$$

$$+ \sup n^{1/2}\,|\hat{\Lambda}_0\{\hat{\Lambda}_0^{-1}(u)\} - u| + \sup|L\{\hat{\Lambda}_0^{-1}(u)\} - L\{\Lambda_0^{-1}(u)\}|.$$

The first term is bounded above by $\sup|\hat{L}(t) - L(t)| \to 0$ almost surely, with the sup taken over $[0,T]$. Further

$$\sup|\Lambda_0\{\hat{\Lambda}_0^{-1}(u)\} - u| \le \sup|\hat{\Lambda}_0\{\hat{\Lambda}_0^{-1}(u)\} - \Lambda_0\{\hat{\Lambda}_0^{-1}(u)\}| \tag{5}$$

$$+ \sup|\hat{\Lambda}_0\{\hat{\Lambda}_0^{-1}(u)\}u| \le \sup|\hat{\Lambda}_0(t) - \Lambda_0(t)| + R_n$$

where $R_n = \inf S^{(0)}(\hat{\beta},t)^{-1}\sup(\Delta N/n)$, which converges in probability to zero. The consistency of $\hat{\Lambda}_0$ entails that (5) converges in probability to 0. Furthermore

$\sup n^{1/2} |\hat{\Lambda}_0\{\hat{\Lambda}_0^{-1}(u)\} - u| \leq \sup \ n^{1/2} |\hat{\Lambda}_0(t) - \hat{\Lambda}_0(t-)| \leq n^{1/2} R_n$ which converges in probability to zero. Further, for $u \in [a', b']$, we have $L\{\hat{\Lambda}_0^{-1}(u)\} = L\{\Lambda_0^{-1}(u')\}$ where $u' = \Lambda_0\{\hat{\Lambda}_0^{-1}(u)\}$. The uniform continuity of the sample paths of $L\{\Lambda_0^{-1}(u)\}$ and consistency of $\Lambda_0\{\hat{\Lambda}_0^{-1}(u)\}$ implies therefore that the third term in (4) converges in probability to 0.

Finally the second term in (3) converges in probability to 0 by (4) and boundedness of the sample paths of $\hat{V}(u)$.

ACKNOWLEDGEMENT

This paper was prepared with the partial support of the National Science Foundation, the National Institute of General Medical Sciences and the University of California Presidential Fellowship.

REFERENCES

ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10,** 1100-1120.

BEGUN, J. M., HALL, W. J., HUANG, W. M. & WELLNER, J. A. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **11,** 432-52.

BORGAN , O. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data. *Scand. J. Statist.* **11,** 1-16.

BRESLOW, N. E. (1972). Contribution to discussion of paper by D.R. Cox. *J. R. Statist. Soc., B* **34,** 216-217.

BRESLOW, N. E. (1974). Covariance analysis of censored survival data. *Biometrics* **30,** 89-100.

BROOKMEYER, R. & CROWLEY, J. J. (1982). A confidence interval for the median survival time. *Biometrics* **38,** 29-41.

CARROLL, R. J. & RUPPERT, D. (1981). On prediction and the power transformation family. *Biometrika* **68,** 609-615.

COX, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B* **34,** 187-202.

COX, D. R. (1975). Partial likelihood. *Biometrika* **62,** 269-276.

GILL, R. (1983). Large sample behaviour of the product limit estimator on the whole line. *Ann. Statist.* **11,** 49-58.

JENNISON, C. & TURNBULL, B. W. (1985). Repeated confidence intervals for median. *Biometrika* **72,** 619-625.

KAPLAN, E. L. & MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* **53,** 457-481.

MILLER, R. G. & HALPERN, J. (1982). Regression with censored data. *Biometrika* **69,** 521-532.

SLUD, E. V., BYAR, D. P. & GREEN, S. B. (1984). A comparison of reflected versus test-based confidence intervals for the median survival time based on censored data. *Biometrics* **40,** 587-600.

SUSARLA, V. & VAN RYZIN, J. (1980). Large sample theory for an estimator of the mean survival time from censored samples. *Ann. Statist.* **8,** 1002-1016.

TAYLOR, J. M. G. (1986). The retransformed mean after a fitted power transformation. *J. Amer. Statist. Assoc.,* **81,** 114-8.

TSIATIS, A. A. (1981). A large sample study of Cox's regression model. *Ann. Statist.* **9,** 93-108.

Table 1. Asymptotic efficiency $e(\hat{\nu}, \hat{\nu}_\alpha)$ of the semiparametric estimate of the conditional median survival time in the Weibull model with known shape and unknown scale parameter. The probability of an individual being censored is $p = 1 - \theta^{-1}$.

p \ z	0.0	0.5	1.0	2.0	4.0	10.0
0.0	0.481	0.536	0.649	0.822	0.940	0.989
1/3	0.591	0.644	0.743	0.879	0.961	0.993
2/3	0.618	0.669	0.764	0.890	0.965	0.994
0.9	0.047	0.058	0.090	0.198	0.456	0.833

Table 2. Asymptotic efficiency $e(\hat{\mu}, \hat{\mu}_\alpha)$ of the semiparametric estimate of the conditional mean survival time in the Weibull model for uncensored data.

z \ α	Known shape parameter					Unknown shape parameter				
	0.1	0.5	1.0	1.5	10	0.1	0.5	1.0	1.5	10
0.0	0.001	0.800	1.0	0.964	0.691	0.002	0.922	1.0	0.997	0.992
0.5	0.001	0.833	1.0	0.971	0.736	0.002	0.935	1.0	0.998	0.993
1.0	0.001	0.889	1.0	0.982	0.817	0.002	0.957	1.0	0.999	0.995
2.0	0.003	0.952	1.0	0.993	0.918	0.004	0.981	1.0	0.999	0.998
4.0	0.009	0.986	1.0	0.998	0.974	0.010	0.994	1.0	0.999	0.999
10.0	0.052	0.998	1.0	0.999	0.996	0.053	0.999	1.0	0.999	0.999