The Annals of Statistics
1981, Vol. 9, No. 6, 1276–1288

REGRESSION ANALYSIS WITH RANDOMLY RIGHT-CENSORED DATA

BY H. KOUL[1], V. SUSARLA,[2,3] AND J. VAN RYZIN[2]

Michigan State University; Michigan State University; and Columbia University

This paper proposes a new estimator of the parameter vector in a linear regression model when the observations are randomly censored on the right and when the error distribution is unknown. This estimator is explicitly defined and easily computable. The paper contains sufficient conditions under which this estimator is mean square consistent and asymptotically normal. A numerical example is given.

1. Introduction and summary. This paper is concerned with the estimation of the regression parameters in a linear model when the data is randomly right censored. Often in medical studies when patients are entering a study randomly for a fixed time period, the observation on the survival time of a patient is incomplete in the sense that it is right censored. This censoring can be due to a number of causes: the patient was alive at the termination of the study, the patient withdrew alive during the study, or the patient died of causes other than those under study.

Formally the above type of situation can be described by the following random censorship linear model. Let $\{T_i, i = 1, \cdots, n\}$ be n independent random variables (rv's) satisfying

(1.1) $$T_i = \alpha + \beta x_i + \varepsilon_i, \qquad 1 \le i \le n,$$

where $x_1, \cdots, x_n$ are known input variables and

(1.2) $\varepsilon_1, \cdots, \varepsilon_n$ are independent and identically distributed (i.i.d.) rv's with zero mean.

The parameters of interest are α and β. One observes not $\{T_i\}$ but

(1.3) $$\delta_i = [T_i < Y_i] \quad \text{and} \quad Z_i = \min(T_i, Y_i), \qquad 1 \le i \le n,$$

where $[A]$ denotes the indicator of the set A and where

(1.4) $Y_1, \cdots, Y_n$ are i.i.d. rv's which are independent of $\varepsilon_1, \cdots, \varepsilon_n$.

The rv's $Y_1, \cdots, Y_n$ are called the censoring variables. When dealing with survival time data, one can take T_i to be $\log_{10}$ or ln of the survival time. The problem considered here is that of the estimation of (α, β) based on $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$.

Miller (1976) introduced an estimator of (α, β), called the Kaplan-Meier Least Squares (KMLS) estimator, which is obtained by minimizing the weighted sum of squares of the residuals with the weights determined by the Kaplan-Meier (1958) estimator of the error distribution based on the residuals. More recently Buckley and James (1979) suggested another estimator of (α, β), herein called the BJ estimator, based on an expectation identity. Both of these estimators are computed using iteration methods. In both cases, as these authors point out, the iterations may eventually settle down to oscillation between

Received October 1979; revised March 1981.
[1] Research supported by NIH research grant No. 1 R01 GM25449.
[2] Research supported in part by NIH research grant No. 1 R01 GM28405.
[3] Research supported in part by NSF research grant No. MCS77-26809.
AMS 1970 *subject classification.* Primary 62G05, 62J05; secondary 62P10, 62N05.
Key words and phrases. Least squares, Kaplan-Meier, consistent, asymptotically normal.

two values. According to Buckley and James, these values are closer to each other for the BJ estimator than for the KMLS estimator. Miller gives a heuristic argument for the asymptotic normality of the KMLS estimator under assumptions including that $Y_i = \alpha + \beta x_i + \varepsilon_i'$ for i.i.d. $\{\varepsilon_i'\}$, independent of $\{\varepsilon_i\}$. Buckley and James provide some idea of the large sample behavior of their estimator but do not provide any mathematical justification.

In this paper we introduce a new estimator $(\hat{\alpha}, \hat{\beta})$ of (α, β) and study its asymptotic distribution theory. Section 2 contains the definition of $(\hat{\alpha}, \hat{\beta})$. In Section 3 we provide sufficient conditions for the mean square consistency and the asymptotic normality of $(\hat{\alpha}, \hat{\beta})$ with the proofs deferred to Section 7. Section 4 contains a detailed discussion about the sufficient conditions of Section 3, implications of our results, and an estimator of the asymptotic variance of $\hat{\beta}$. Section 5 provides details concerning extension of $(\hat{\alpha}, \hat{\beta})$ to the multiple linear regression model while Section 6 illustrates the estimator using the heart transplant data of Miller.

Some of the important features of $(\hat{\alpha}, \hat{\beta})$ are that they are explicitly defined, easy to compute and do not require any iteration scheme. Another important feature of this estimator is that its computational and theoretical aspects are easily extendable to the multiple linear regression model as is done in Section 5 whereas the same can not be said for the KMLS and the BJ estimators.

NOTATION. Throughout for any real number t, $G(t) = P(Y > t)$, $F_i(t) = P(X_i > t) = F(t - \alpha - \beta x_i)$, $1 \le i \le n$, where $F(t) = P(\varepsilon_1 > t)$; the index i in the summation and maximization runs from 1 through n; $\bar{F} = n^{-1} \sum F_i$, $H_j = F_j G$, $1 \le j \le n$, $\bar{H} = \bar{F}G$ and $\hat{H}_n(\cdot) = n^{-1} \sum [Z_i > \cdot]$; H^{-r} will stand for $(1/H)^r$ for any distribution or survival function H and for $r > 0$; all limits are taken as $n \to \infty$, $\to_d$ means 'convergence in distribution,' $o(1)$ means a term which converges to zero and $o_p(1)$ means a term which converges to zero in probability; $N_k(\boldsymbol{\mu}, \Sigma)$ will stand for the k-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ, and consistent will stand for "mean square consistent." Moreover, in the sequel, all distributional calculations are carried out under the assumptions (1.1), (1.2) and (1.4), besides any other assumptions that will be mentioned as we proceed. Finally, $\bar{x} = n^{-1} \sum x_i$ and $\tau_x^2 = \sum (x_i - \bar{x})^2$.

2. Definition of estimators. Observe that under the assumptions (1.1), (1.2), (1.4) and the assumption that $G(t) > 0$ for all $-\infty < t < +\infty$,

$$E(\delta_i Z_i) = -\int tG(t)\, dF(t - \alpha - \beta x_i), \qquad 1 \le i \le n, \tag{2.1}$$

which in turn suggests the relation

$$E\delta_i Z_i \{G(Z_i)\}^{-1} = -\int t\, dF(t - \alpha - \beta x_i) = \alpha + \beta x_i, \qquad 1 \le i \le n. \tag{2.2}$$

Hence the variables $\{\delta_i Z_i \{G(Z_i)\}^{-1}, 1 \le i \le n\}$ obey a linear regression model with the same parameters as in (1.1) except that the errors may not be identically distributed. Thus, if G were known, $\tau_x^{-2} \sum (x_i - \bar{x})\, \delta_i Z_i \{G(Z_i)\}^{-1}$ would be the least squares estimator of β. But usually G is unknown and a natural thing to do is to replace G by an estimator in this quantity. From Susarla and Van Ryzin (1980), an estimator of $G^{-1}(t)$ is given by the second factor of (1.10) there, which upon inversion yields an estimator of $G(t)$ given by

$$\hat{G}(t) = \prod_{j=1}^{n} \{(1 + N^+(Z_j))/(2 + N^+(Z_j))\}^{[\delta_j = 0,\, Z_j \le t]}, \qquad -\infty < t < +\infty \tag{2.3}$$

where $N^+(z)$ is the number of Z_i exceeding z, i.e.

$$N^+(z) = \sum [Z_i > z]. \tag{2.4}$$

Asymptotically, $\hat{G}$ behaves like the product-limit estimator of G. We have chosen $\hat{G}$ over the product-limit estimator because we need to take the logarithm of the product in $\hat{G}$ in our proof. Details in Section 7 show that $\hat{G}(t)$ is a consistent estimator of $G(t)$ under the conditions of this paper.

Because of the explosive behavior of the asymptotic variance of $\hat{G}(t)$ for large t, we are motivated to define an estimator of β to be

$$\hat{\beta} = \tau_x^{-2} \sum (x_i - \bar{x})\delta_i Z_i \{\hat{G}(Z_i)\}^{-1}[Z_i \le M_n] \tag{2.5}$$

where M_n is a sequence of real numbers tending to ∞ at a rate to be specified later on. Similarly, an estimator of α is given by

$$\hat{\alpha} = n^{-1} \sum \delta_i Z_i \{\hat{G}(Z_i)\}^{-1}[Z_i \le M_n] - \hat{\beta}\bar{x}. \tag{2.6}$$

3. Statement of theorems. This section contains the formal statements of the consistency and the asymptotic normality results for $(\hat{\alpha}, \hat{\beta})$. From (2.5) and (2.6), one can write

$$\begin{aligned} \hat{\alpha} &= \sum \{n^{-1} - \bar{x}\, d_i \tau_x^{-2}\} W_i \{\hat{G}(Z_i)\}^{-1} = \sum a_{ni} \hat{W}_i \\ \hat{\beta} &= \tau_x^{-2} \sum d_i W_i \{\hat{G}(Z_i)\}^{-1} = \sum b_{ni} \hat{W}_i \end{aligned} \tag{3.1}$$

where

$$W_i = \delta_i Z_i [Z_i \le M_n], \quad d_i = x_i - \bar{x} \quad \text{and} \quad \hat{W}_i = W_i \{\hat{G}(Z_i)\}^{-1}, \quad 1 \le i \le n, \tag{3.2}$$

and where

$$a_{ni} = n^{-1} - \bar{x}\, d_i \tau_x^{-2} \quad \text{and} \quad b_{ni} = \tau_x^{-2} d_i, \quad 1 \le i \le n. \tag{3.3}$$

Since both $\hat{\alpha}$ and $\hat{\beta}$ are linear combinations of $\hat{W}_i$ we shall prove consistency for such a linear combination, and then prove the joint asymptotic normality for the two linear combinations.

In order to state our theorems we introduce the following notation and assumptions. Let

$$\begin{aligned} L_{n1}(t) &= \sum a_{ni} F_i(t), & L_{n2}(t) &= \sum b_{ni} F_i(t), \\ K_{n1}(t) &= \sum |a_{ni}| F_i(t), & K_{n2}(t) &= \sum |b_{ni}| F_i(t), \quad -\infty < t < +\infty \end{aligned} \tag{3.4}$$

and

$$\bar{L}_{nj} = n^{-1} L_{nj}, \quad j = 1, 2. \tag{3.5}$$

Various subsets of the following assumptions will be needed for the consistency and asymptotic normality of $(\hat{\alpha}, \hat{\beta})$, with $e_{ni} = a_{ni}$ and b_{ni} as appropriate.

A1. $$G(t) > 0, \quad -\infty < t < +\infty.$$

A2. $$\sum e_{ni} \int_{M_n}^{\infty} t \, dF_i(t) \to 0.$$

A3. $$\sum e_{ni}^2 \left[-\int_{-\infty}^{M_n} t^2 G^{-1}(t)\, dF_i(t) - \left\{ -\int_{-\infty}^{M_n} t\, dF_i(t) \right\}^2 \right] \to 0.$$

A4. $$\sum e_{ni}^2 \int_{-\infty}^{M_n} t^2 \{\bar{H}(t)\}^{-3} \left\{ -\int_{-\infty}^{t} \bar{F}(s) \{\bar{H}(s)\}^{-8}\, dG(s) \right\}^{1/2} G(t)\, dF_i(t) \to 0.$$

A5. $$\liminf n\, \bar{H}(M_n) \ge b \quad \text{for some} \quad b > 2.$$

A6. $$n^{-1/2} \int_{-\infty}^{M_n} |t| G(t) \{\bar{H}(t)\}^{-2} \left[-\int_{-\infty}^{t} \bar{F}(s) \{\bar{H}(s)\}^{-6}\, dG(s) \right]^{1/2} dK_{nj}(t) \to 0, \quad j = 1, 2.$$

A7. $\sum \sigma_{ni}^2 \le c$ for all $n \ge 1$ and $\max \sigma_{ni}^2 \to 0$,

where $\sigma_{ni}^2 = \mathrm{Var}(A_{ni})$, A_{ni} as in (7.27), $1 \le i \le n$ and c is a finite constant. We now state

THEOREM 3.1. *Under assumptions* A1–A5

$$E\{\sum e_{ni}\hat{W}_i - \sum e_{ni}(\alpha + \beta x_i)\}^2 = o(1).$$

COROLLARY 3.1. *Suppose* A1–A5 *hold, with both* $e_{ni} = b_{ni}$ *and* a_{ni} *then*

$$E(\hat{\alpha} - \alpha)^2 + E(\hat{\beta} - \beta)^2 = o(1).$$

The joint asymptotic normality is given by the following

THEOREM 3.2. *Assume that* A1, A5, A6, A7 *hold and that* A4 *holds with both* $e_{ni} = a_{ni}$ *and* b_{ni}, $1 \le i \le n$. *Then*

$$n^{1/2}\left(\hat{\alpha} + \int_{-\infty}^{M_n} t\, dL_{n1}(t), \hat{\beta} + \int_{-\infty}^{M_n} t\, dL_{n2}(t)\right) \to_d N_2(\mathbf{0}, \Sigma)$$

where $\Sigma = ((\sigma_{jk}))$, $j, k = 1, 2$ *with*

$$\sigma_{11} = \lim n\left[\sum a_{ni}^2 \mathrm{Var}\{W_i G^{-1}(Z_i)\} + \int_{-\infty}^{M_n} \bar{F}(t)\{\bar{H}(t)\}^{-2}\left\{\int_t^{M_n} s\, d\bar{L}_{n1}(s)\right\}^2 dG(t)\right] \tag{3.6}$$

(3.7) σ_{22} *is the same as* σ_{11} *with* $\{b_{ni}\}$ *in place of* $\{a_{ni}\}$

and

$$\begin{aligned}\sigma_{12} = \lim n\Big[&\sum a_{ni}b_{ni} \mathrm{Var}\{W_i G^{-1}(Z_i)\} \\ &+ \int_{-\infty}^{M_n} \bar{F}(t)\{\bar{H}(t)\}^{-2}\left\{\int_t^{M_n} s\, dL_{n1}(s)\right\}\left\{\int_t^{M_n} s\, dL_{n2}(s)\right\} dG(t)\Big].\end{aligned} \tag{3.8}$$

4. Discussion of asymptotic results. In the previous section, we have given the general asymptotic results for the estimators $(\hat{\alpha}, \hat{\beta})$ introduced in Section 2. The purpose here is to see how the results specialize and what the conditions of the theorems mean from the viewpoint of applications.

REMARK 4.1. If $G(x) \equiv 1$, and $M_n = \infty$, then A2–A4 are trivially satisfied provided $\tau_x^2 \to \infty$. Also, $\hat{\alpha}$ and $\hat{\beta}$ reduce to the usual least squares estimators for the uncensored data situation.

REMARK 4.2. Assumption A2 can be interpreted as saying that the truncated weighted average of $\{ET_i\}$, when weighted by $\{e_{ni}\}$, goes to zero. A sufficient condition for this to happen, when $e_{ni} = a_{ni}$ or b_{ni}, $1 \le i \le n$, is that $\tau_x^2 \to \infty$ and $\{E\,|\,T_i\,|\}$ be bounded. If $\{x_i\}$ are bounded, and if ε_i are i.i.d. $N_1(0, \gamma^2)$ for some $\gamma^2 < \infty$, then $\{E\,|\,T_i\,|\}$ are bounded. In much the same way, A3 can be interpreted as a condition on the variances of $\{W_i\{G(Z_i)\}^{-1}, 1 \le i \le n\}$. If these variances are bounded, then A3 would be satisfied with $e_{ni} = a_{ni}$ or b_{ni} provided $\tau_x^2 \to \infty$. An example where these variances are bounded is when $\{x_i\}$ are bounded, F is $N_1(0, \gamma^2)$ for some $\gamma^2 < \infty$ and G is double exponential with scale paramter $\theta > 0$. (Note that in proving the asymptotic normality of the least squares estimators based on uncensored data, one generally assumes that $\tau_x^2 \to \infty$.)

It is clear from the assumptions of Theorems 3.1 and 3.2 that the choice of M_n (preferably depending on the data) satisfying these assumptions is an important and interesting question by itself. Just to get an idea concerning the magnitude of $\{M_n\}$, we point out here that if $M_n = c(\ln n)^r$, with $0 < 2r < 1$ and $c > 0$ known constants, then conditions A2–A7 hold provided that $\{x_i\}$ is bounded with $\tau_x^2 \to \infty$, and that for an $\eta > 0$,

$\{F(x) + G(x)\}\exp(\eta x^2/2)$ is bounded as $x \to \infty$. In particular, the conclusions of Theorem 3.1 and 3.2 hold whenever F and G are normal or have right tails lighter than a normal distribution. If one were to choose an M_n depending on data, we conjecture that $M_n = c(\ln \sum |Z_i|)^r$, with $0 < 2r < 1$ and $c > 0$, would include several interesting possible choices for F and G under which Theorems 3.1 and 3.2 would hold.

REMARK 4.3. The results of Corollary 3.1 and Theorem 3.1 extend to the situation where the ε_i are not necessarily indentically distributed; they need only be independent with zero means. From the application point of view this means that these results are applicable to survival data analysis with covariates even when error distributions may be different from different patients—a phenomenon which can easily arise in practice. A theoretical example would be when ε_i are $N(0, \gamma_i^2)$, $\{\gamma_i\}$ bounded, and G is an exponential distribution.

REMARK 4.4. As in Susarla and Van Ryzin (1980), it can be shown that it is not possible to replace the centering constants in Theorem 3.2 by α and β.

REMARK 4.5. *On the asymptotic variance of $\hat{\beta}$ and its estimator.* Note that if there is no censoring then the second term on the right hand side of (3.7) is zero and σ_{22} reduces to the variance of the least squares estimator of β. In general this second term is negative thereby implying that the asymptotic variance of the standardized $\hat{\beta}$ is smaller than that of $\hat{\beta}_0 = \tau_x^{-2} \sum d_i W_i G^{-1}(Z_i)$, an estimator of β when G is known. One reason for this phenomenon is that $\hat{\beta}_0$ uses only uncensored observations whereas $\hat{\beta}$ uses all the observations including the censored ones via $\hat{G}$.

In some interesting situations the second term in (3.7) tends to zero. To see this, recall that $b_{ni} = d_i\tau_x^{-2}$, $1 \le i \le n$ in (3.7). The Cauchy-Schwarz inequality and Fubini's theorem yield

$$n \int_{-\infty}^{M_n} \bar{F}(t)\{\bar{H}(t)\}^{-2}\left\{-\int_t^{M_n} s\, d\bar{L}_{n2}(s)\right\}^2 d(-G(t)) \le n(\max d_i^2)\tau_x^{-2} \cdot \tau_x^{-2} \int_{-\infty}^{M_n} t^2\{1 - G(t)\}G^{-1}(t)\, d(-\bar{F}(t)). \tag{4.1}$$

Thus if $\{x_i\}$, G and F satisfy

$$\limsup n(\max d_i^2)\tau_x^{-2} < \infty \tag{4.2}$$

and

$$\limsup \tau_x^{-2} \int_{-\infty}^{M_n} t^2\{1 - G(t)\}G^{-1}(t)\, d(-\bar{F}(t)) = 0, \tag{4.3}$$

then the asymptotic variance of $n^{1/2}\hat{\beta}$ (see (3.7)) becomes $\lim \sigma_{n22}$ where

$$\sigma_{n22} = n\tau_x^{-4} \sum d_i^2 \operatorname{Var}(\delta_i Z_i \{G(Z_i)\}^{-1}[Z_i \le M_n]). \tag{4.4}$$

Conditions (4.2) and (4.3) are satisfied, for example, when

$$\{x_i\} \text{ are bounded}, \qquad n^{-1}\tau_x^2 \to a^2, \qquad 0 < a^2 < \infty \tag{4.5}$$

and

$$\int_{-\infty}^{\infty} t^2 G^{-1}(t)\, d(-\bar{F}(t)) < \infty \qquad \text{for all} \quad n \ge 1. \tag{4.6}$$

Under (4.5) and (4.6) one actually has

$$\sigma_{n22} = n\tau_x^{-4} \sum d_i^2 \operatorname{Var}(\delta_i Z_i \{G(Z_i)\}^{-1}) + o(1). \tag{4.7}$$

Note that (4.6) is essential for the integrals in σ_{n22} to exist.

Now (4.7) suggests an estimator of σ_{n22} to be

$$\hat{\sigma}_{n22} = n\tau_x^{-4} \sum d_i^2(\delta_i Z_i\{\hat{G}(Z_i)\}^{-1} - \hat{\alpha} - \hat{\beta}x_i)^2[Z_i \le M_n]. \tag{4.8}$$

Using Corollary 3.1, one can show, under (4.5), (4.6) and the conditions of Corollary 3.1, that $|\hat{\sigma}_{n22} - \sigma_{n22}| = o_p(1)$.

In general, one may use $\hat{\gamma}_{n22}$ as an estimator of σ_{22}, where

$$\hat{\gamma}_{n22} = \hat{\sigma}_{n22} - n\hat{\Delta}_{n22},$$

$$\hat{\Delta}_{n22} = \int_{-\infty}^{M_n} \{\hat{H}_n(t)\}^{-2}\left\{\int_t^{M_n} s\hat{G}^{-1}(s)\, d\hat{S}_{n2}(s)\right\}^2 d\hat{S}_{n3}(t),$$

with

$$\hat{S}_{n2}(t) = \tau_x^{-2}n^{-1}\sum d_j\delta_j[Z_j \le t], \qquad \hat{S}_{n3}(t) = n^{-1}\sum(1-\delta_j)[Z_j \le t], \qquad -\infty < t < +\infty.$$

Note that $\hat{\Delta}_{n22} \ge 0$ and hence $\hat{\gamma}_{n22} \le \hat{\sigma}_{n22}$.

Similarly one can construct the estimators of the asymptotic variance σ_{11} and covariance σ_{12}.

One may use $\hat{\beta}$ to test $H_0 : \beta = 0$. The test would reject H_0 at asymptotic level $2t$ if $|n^{1/2}\hat{\beta}| > (\hat{\sigma}_{n22})^{1/2}z_{1-t}$ where z_t is the tth percentile of $N_1(0, 1)$.

5. Extension to multiple regression.

Here we briefly discuss the extension of $(\hat{\alpha}, \hat{\beta})$ to the multiple linear regression model. Suppose

$$T_i = \mathbf{c}_i\boldsymbol{\beta} + \varepsilon_i, \qquad 1 \le i \le n \tag{5.1}$$

where $\mathbf{c}_i = (1, x_{i1}, \cdots, x_{ik})$ is the ith row of the design matrix C_n, $\boldsymbol{\beta}' = (\beta_0, \beta_1, \cdots, \beta_k)$ is the parameter vector and $\{\varepsilon_i\}$ and $\{Y_i\}$ are as in (1.2), (1.4) and A1. We assume that $(C_n'C_n)^{-1}$ exists for large n.

A generalization of $(\hat{\alpha}, \hat{\beta})$ of Section 3 is given by

$$\hat{\boldsymbol{\beta}} = (C_n'C_n)^{-1}C_n'\hat{\mathbf{W}} \tag{5.2}$$

with $\hat{\mathbf{W}}' = (\hat{W}_1, \cdots, \hat{W}_n)$, where $\{\hat{W}_i, 1 \le i \le n\}$ are defined in (3.2).

For the sake of brevity, we state an extension of Theorem 3.2 only. For any matrix A, let $A_{i,j}$ denote its (i, j)th element. Let

$$\mathbf{F}(t) = (F_1(t), \cdots, F_n(t))', \qquad B_n = (C_n'C_n)^{-1}C_n'$$

$$\mathbf{v}_n'(t) = \mathbf{F}'(t)B_n'(t) = (v_{n0}(t), \cdots, v_{nk}(t))$$

$$\mu_{nj} = \int_{-\infty}^{M_n} t\, dv_{nj}(t), \qquad j = 0, 1, \cdots, k,$$

$$\boldsymbol{\mu}_n' = (\mu_{n0}, \cdots, \mu_{nk}) \quad \text{and} \quad \bar{v}_{nj} = n^{-1}v_{nj}, \qquad 0 \le j \le k.$$

Then

THEOREM 5.1. *Under* (5.1), (1.2), (1.4), A1 *and* A4 *through* A7 *with* $\{(a_{ni}, b_{ni})\}$ *there replaced by the* $(k+1)$ *row elements of* B_n, *one has*

$$n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\mu}_n) \to_d N_{k+1}(\mathbf{0}, \Sigma)$$

where for $0 \le i, j \le k$ *the* (i, j)th *element of* Σ *is given by*

$$\Sigma_{i,j} = \lim\left\{(B_nD_nB_n')_{i,j} + \int_{-\infty}^{M_n} \bar{F}(t)\{\bar{H}(t)\}^{-2}\left(\int_t^{M_n} s\, d\bar{v}_{ni}(s)\right)\left(\int_t^{M_n} s\, d\bar{v}_{nj}(s)\right) dG(t)\right\}$$

with $D_n = \operatorname{diag}(\operatorname{Var}\{W_r G^{-1}(Z_r)\})\ n \times n$. □

6. Numerical example. This section illustrates the computation of $(\hat{\alpha}, \hat{\beta})$ as defined by (2.5) and (2.6) and that of its extension (5.2) using the heart transplant data of Miller (1976). In the following, M_n is chosen to be $c(\ln n)^r$ with $c = 3$ and $r = 0.4$ (see Remark 4.2 above). The dependent variable is $T = \log_{10}$ (survival time).

When regressing T on the mismatch score ($T5$), we changed the observation number 60 from 65 to 50 as in Miller (1976) and obtained

$$\hat{\alpha} = 1.31102, \quad \hat{\beta} = 0.25801, \quad \text{s.d.}(\hat{\beta}) = 0.34673, \quad \hat{\sigma}^2 = .61217. \tag{6.1}$$

Here non-rejection death is treated as censoring.

When regressing T on the age (A), we left the 60th observation unchanged as in Miller (1976) and obtained

$$\hat{\alpha} = -0.83484, \quad \hat{\beta} = 0.05355, \quad \text{s.d.}(\hat{\beta}) = 0.02039, \quad \hat{\sigma}^2 = 0.64379. \tag{6.2}$$

Here being alive by the termination of the study is treated as being censored.

In (6.1) and (6.2), $\text{s.d.}(\hat{\beta}) = (\hat{\sigma}_{n22}/n)^{1/2}$ where $\hat{\sigma}_{n22}$ is defined by (4.8), and $\hat{\sigma}^2$ is an estimator of the error variance $\sigma^2 = \text{Var}(\varepsilon_i)$ given by the equation

$$\hat{\sigma}^2 = n^{-1} \sum (Z_i - \hat{\alpha} - \hat{\beta} x_i)^2 \delta_i \{\hat{G}(Z_i)\}^{-1} [Z_i \le M_n]. \tag{6.3}$$

Using the details of Section 7 one can show that $\hat{\sigma}^2$ is a consistent estimator of σ^2 under conditions similar to those in Theorem 3.1.

When regressing T jointly on A and $T5$, where non-rejection related death is being treated as censoring, we changed the 60th observation from 65 to 50 and obtained

$$\hat{\alpha} = -1.98386, \quad \hat{\beta}(A) = 0.07663, \quad \text{and} \quad \hat{\beta}(T5) = 0.05154. \tag{6.4}$$

When censoring is being alive at the end of the study, we again changed 65 to 50 and obtained

$$\hat{\alpha} = -0.86862, \quad \hat{\beta}(A) = 0.05567, \quad \hat{\beta}(T5) = -0.10211.$$

The reader should be cautioned that in order to use the results of this paper to draw any further conclusions about the heart transplant data from (6.1), (6.2), (6.4) and (6.5) one should first ensure the validity of the assumptions (1.4) and A1 through A7. The example given here is merely for illustrative purposes.

7. Proofs of Theorems 3.1 and 3.2. The proofs here follow the pattern of those of Sections 3 and 4 of Susarla and Van Ryzin (1980). However the proofs are somewhat different because here $\{T_i, 1 \le i \le n\}$ are not identically distributed.

In the sequel C will denote a generic constant which will be different in different contexts, but never will depend on n or any distributions. E_i and $E_{i,j}$ will denote the conditional expectation, given (δ_i, Z_i) and $\{(\delta_i, Z_i), (\delta_j, Z_j)\}$, respectively. M will stand for M_n.

PROOF OF THEOREM 3.1. Let $S_n = \sum e_{ni} \hat{W}_{ni}$. Then

$$\begin{aligned} S_n - \sum e_{ni}(\alpha + \beta x_i) &= \sum e_{ni} W_i \{\hat{G}^{-1}(Z_i) - G^{-1}(Z_i)\} \\ &\quad + \sum e_{ni} \{W_i G^{-1}(Z_i) - (\alpha + \beta x_i)\} \\ &= I + II, \qquad \text{say}. \end{aligned} \tag{7.1}$$

Using (2.2) and (3.2), one observes that

$$E(II) = -\sum e_{ni} \int_M^\infty t \, dF_i(t)$$

and

$$\mathrm{Var}(II) = \sum e_{ni}^2 \left\{-\int_{-\infty}^{M} t^2 G^{-1}(t)\, dF_i(t) - \left(-\int_{-\infty}^{M} t\, dF_i(t)\right)^2\right\}.$$

Hence by A2 and A3

$$E(II^2) \to 0. \tag{7.2}$$

Now we deal with the more difficult term I. The Cauchy-Schwarz inequality applied to the sum and expectation gives

$$E(I^2) \le n \sum e_{ni}^2 E[W_i^2 E_i\{\hat{G}^{-1}(Z_i) - G^{-1}(Z_i)\}^2]. \tag{7.3}$$

Write $\hat{G}^{-1}$ and G^{-1} as $e^{-\ln \hat{G}}$ and $e^{-\ln G}$, use the Mean Value Theorem, argue as in Lemma 3.1 of Susarla and Van Ryzin (1980) and apply the Cauchy-Schwarz inequality to conclude

$$\begin{aligned} E_i\{\hat{G}^{-1}(Z_i) - G^{-1}(Z_i)\}^2 &\le CE_i(\{\ln \hat{G}(Z_i) - \ln G(Z_i)\}^2 \\ &\quad \cdot [G^{-2}(Z_i) + \{(N^+(Z_i)+1)/(n+1)\}^{-2}]) \\ &\le C[G^{-2}(Z_i) + E_i^{1/2}\{(N^+(Z_i)+1)/(n+1)\}^{-4}] \\ &\quad \cdot E_i^{1/2}\{\ln \hat{G}(Z_i) - \ln G(Z_i)\}^4. \end{aligned} \tag{7.4}$$

The following lemmas are found useful in bounding the upper bounds of (7.4) and in the proof of Theorem 3.2.

MOMENT LEMMA. *Let V be a Binomial (k, p) rv. Then for $r \ge 1$ an integer,*

$$E(1+V)^r \le r!(kp)^{-r}.$$

PROOF. For any $v \ge 0$ and integer $\ell \ge 1$, $\ell + v \le \ell(1+v)$. Hence

$$\begin{aligned} E(1+V)^{-r} &\le r! E\{(1+V)(2+V)\cdots(r+V)\}^{-1} \\ &\le r!\,(kp)^{-r} \sum_{\ell=r}^{k+r} \binom{k+r}{\ell} p^{\ell}(1-p)^{k+r-\ell} \le r!\,(kp)^r. \quad \square \end{aligned}$$

LEMMA 7.1. *Let B_i, $i = 1, \cdots, k$ be independent Bernoulli* rv.'s *with $P[B_i = 1] = p_i$, $i = 1, \cdots, k$ and let $S = B_1 + \cdots + B_k$. Then for any $r \ge 1$,*

$$E(1+S)^{-r} \le C(k\bar{p})^{-r}, \qquad \bar{p} = k^{-1} \sum_{k=1}^{k} p_i$$

where C is a constant depending only on r.

PROOF. Fix an $r \ge 1$. Since $f(x) = (1+x)^{-r}$ is a convex function on $[0, \infty)$, Theorem 3 of Hoeffding (1956) shows that $E(1+S)^{-r} \le E(1+S^*)^{-r}$ where S^* is a Binomial $(k, \bar{p})$ rv. Application of the above Moment Lemma yields the inequality. $\square$

COROLLARY 7.1.

$$E_i\{1 + N^+(Z_i)\}^{-r} \le Cn^{-r}\{\bar{H}(Z_i) - n^{-1}\}^{-r}, \tag{7.5}$$

$$E_{i,j}\{1 + N^+(Z_j)\}^{-r} \le Cn^{-r}\{\bar{H}(Z_j) - 2n^{-1}\}^{-r}, \qquad 1 \le i, j \le n, \tag{7.6}$$

where C depends only on r.

PROOF. Given (δ_i, Z_i), $N^+(Z_i)$ is the sum of $n-1$ Bernoulli independent rv's with probability of success of ℓth rv equal to $H_\ell(Z_i)$. Applying Lemma 7.1 with $k = n-1$ and

$p_\ell = H_\ell(Z_i)$, one gets

$$E_i\{1 + N^+(Z_i)\}^{-r} \le C\,\{\textstyle\sum_{\ell\ne i} H_\ell(Z_i)\}^{-r} \le Cn^{-r}\{\bar{H}(Z_i) - n^{-1}\}^{-r}.$$

For (7.6), observe that $N^+(Z_i) \ge N_i^+(Z_j) = \sum_{k\ne i,j}[Z_k > Z_j]$ and that $N_i^+(Z_j)$, given $\{(\delta_i, Z_i), (\delta_j, Z_j)\}$, is the sum of $n - 2$ independent Bernoulli rv's. Now argue as for (7.5). □

LEMMA 7.2.

$$E_i\{\ln\hat{G}(Z_i) - \ln G(Z_i)\}^4 \le C\Big[-n^{-2}\int_{-\infty}^{Z_i} \bar{F}(t)\{\bar{H}(t) - 2n^{-1}\}^{-8}\,dG(t) \tag{7.7}$$
$$+ n^{-4}\bar{H}^{-4}(Z_i)\{[\delta_i = 0] + (1 - G(Z_i))\}^4\Big].$$

PROOF. Write

$$-\ln\hat{G}(Z_i) = -\textstyle\sum_j[\delta_j = 0, Z_j \le Z_i]\ln(1 - \{2 + N^+(Z_j)\}^{-1}). \tag{7.8}$$

Expansion of $-\ln(1 - x)$ for $x < 1$ and the c_4-inequality (see Loève, 1963, page 155) yield

$$\{-\ln\hat{G}(Z_i) + \ln G(Z_i)\}^4 \le 8(n^{-1}\textstyle\sum_j\{[\delta_j = 0, Z_j \le Z_i]n(2 + N^+(Z_j))^{-1} + \ln G(Z_i)\})^4$$
$$+ 8(\textstyle\sum[\delta_j = 0, Z_j \le Z_i]\sum_{k=2}^{\infty}\{2 + N^+(z_j)\}^{-k}k^{-1})^4.$$

Bound the infinite series in the second term by the geometric series with common ratio $\{2 + N^+(Z_j)\}^{-1}$ to get

$$E_i\{-\ln\hat{G}(Z_i) + \ln G(Z_i)\}^4 \le 8E_i(I_1 + I_2), \tag{7.9}$$

where

$$I_1 = (n^{-1}\textstyle\sum_j\{[\delta_j = 0, Z_j \le Z_i]n\{2 + N^+(Z_j)\}^{-1} + \ln G(Z_i)\})^4$$

and

$$I_2 = (n^{-1}\textstyle\sum_j[\delta_j = 0, Z_j \le Z_i]n\{1 + N^+(Z_j)\}^{-2})^4.$$

Applying the moment inequality to the average and taking expectation yields

$$E_i(I_2) \le n^{-1}\textstyle\sum_j E_i([\delta_j = 0, Z_j \le Z_i]E_{i,j}\{n^4(1 + N^+(Z_j))^{-8}\}).$$

An application of (7.6) yields

$$E_i(I_2) \le -C\,n^{-4}\int_{-\infty}^{Z_i}\bar{F}(t)\{\bar{H}(t) - 2n^{-1}\}^{-8}\,dG(t). \tag{7.10}$$

Next, adding and subtracting $\bar{H}^{-1}(Z_j)$, which is the approximate centering for $n\{2 + N^+(Z_j)\}^{-1}$ given $\{(\delta_i, Z_i), (\delta_j, Z_j)\}$, in the summands of I_1 and using the c_4-inequality one gets

$$I_1 \le 8(I_{11} + I_{12}), \tag{7.11}$$

where

$$I_{11} = \{n^{-1}\textstyle\sum_j[\delta_j = 0, Z_j \le Z_i]\bar{H}^{-1}(Z_j) + \ln G(Z_i)\}^4$$

and

$$I_{12} = [n^{-1}\textstyle\sum_j[\delta_j = 0, Z_j \le Z_i]\{n(2 + N^+(Z_j))^{-1} - \bar{H}^{-1}(Z_j)\}]^4.$$

Applying the moment inequality to the average and taking expectations gives

$$E_i(I_{12}) \le n^{-1}\textstyle\sum_j E_i[\delta_j = 0, Z_j \le Z_i]E_{i,j}\{n(2 + N^+(Z_j))^{-1} - \bar{H}^{-1}(Z_j)\}^4.$$

Now

$$E_{i,j}[n\{2 + N^+(Z_j)\}^{-1} - \bar{H}^{-1}(Z_j)]^4$$
$$= \bar{H}^{-4}(Z_j)E_{i,j}[\{2 + N^+(Z_j)\}^{-1}\{(2 + N^+(Z_j)) - n\bar{H}(Z_j)\}]^4.$$

Applying the Cauchy-Schwarz inequality under $E_{i,j}$, a bound of order n^4 to the 8th moment of the sum of centered Bernoulli independent rv's and (7.6), one gets

$$E_i(I_{12}) \le -n^{-2}\int_{-\infty}^{Z_i} \bar{F}(t)\{\bar{H}(t) - 2n^{-1}\}^{-8}\, dG(t). \tag{7.12}$$

To deal with I_{11}, notice that

$$E_i\{n^{-1}\textstyle\sum_j [\delta_j = 0, Z_j \le Z_i]\bar{H}^{-1}(Z_j)\}$$
$$= -\ln G(Z_i) + n^{-1}\{\bar{H}^{-1}(Z_i)[\delta_i = 0] + \int_{-\infty}^{Z_i} F_i\bar{H}^{-1}\, dG\}.$$

This relation together with the c_4-inequality yields

$$E_i(I_{11}) \le 8[E_i(n^{-1}\textstyle\sum_{j\ne i}\alpha_j)^4 + n^{-4}\{\bar{H}^{-1}(Z_i)[\delta_i = 0] + \displaystyle\int_{-\infty}^{Z_i} F_i\bar{H}^{-1}\, dG\}^4] \tag{7.13}$$

where α_j is the conditionally centered, given (δ_i, Z_i), rv $[\delta_j = 0, Z_j \le Z_i]\bar{H}^{-1}(Z_j)$. Since $\{\alpha_j, j \ne i\}$ are conditionally independent and centered, given (δ_i, Z_i), (7.13) yields

$$E_i(I_{11}) \le C\left[-n^{-2}\int_{-\infty}^{Z_i} \bar{F}(\bar{H})^{-4}\, dG + n^{-4}\{\bar{H}^{-1}(Z_i)[\delta_i = 0] + \int_{-\infty}^{Z_i} F_i\bar{H}^{-1}\, dG\}^4\right]. \tag{7.14}$$

The proof of (7.7) is completed upon observing that

$$-\int_{-\infty}^{Z_i} F_i\bar{H}^{-1}\, dG \le \bar{H}^{-1}(Z_i)\{1 - G(Z_i)\} \qquad \text{and} \qquad \bar{H}^{-4} \le \bar{H}^{-8} \le (\bar{H} - 2n^{-1})^{-8}$$

for sufficiently large n and upon combining these observations with (7.8) through (7.14). □

We now return to the term I in (7.1). From (7.3), (7.4), (7.5) and (7.7) and the facts that $W_i^2[\delta_i = 0] = 0$ and $(a + b)^{1/2} \le a^{1/2} + b^{1/2}$ for $a > 0$, $b > 0$, it follows that

$$\begin{aligned}
\limsup EI^2 &\le C \limsup \{\textstyle\sum e_{ni}^2 \displaystyle\int_{-\infty}^{M} t^2[G^{-2}(t) + (\bar{H}(t) - n^{-1})^{-2}] \\
&\quad\cdot\left[\left(-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s) - 2n^{-1}\}^{-8}\, dG(s)\right)^{1/2}\right. \\
&\quad\left. + n^{-1}\{\bar{H}(t)\}^{-2}(1 - G(t))^2\right] d\tilde{H}_i(t)\} \\
&\le Cb^* \limsup\left\{\textstyle\sum e_{ni}^2 \displaystyle\int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-2}\right. \\
&\quad\cdot\left[\left(-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s)\}^{-8}\, dG(s)\right)^{1/2}\right. \\
&\quad\left.\left. + n^{-1}\{\bar{H}(t)\}^{-2}(1 - G(t))^2\right] d\tilde{H}_i(t)\right\}
\end{aligned} \tag{7.15}$$

$$\leq Cb^*(1+b^{-1})\limsup\left\{\sum e_{ni}^2 \int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-3} \cdot\left(-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s)\}^{-8}\,dG(s)\right)^{1/2} d\tilde{H}_i(t)\right\}$$
$$= 0,$$

by A4. Here $d\tilde{H}_i = Gd\,(-F_i)$. In the second inequality $b^* = \{b(b-2)^{-1}\}^{-4}$. The second inequality follows because A5 and the monotonicity of $\bar{H}$ imply that for sufficiently large n, $\bar{H}(s) - 2n^{-1} > 0$, and that

$$\bar{H}(s)\{\bar{H}(s) - 2n^{-1}\}^{-1} \leq 1 + 2\{n\bar{H}(M) - 2\}^{-1} \leq b(b-2)^{-1}$$

for all $s \leq M$ and because $G^{-2} \leq (\bar{H})^{-2}$.

The last inequality in (7.15) follows because

$$b^{-1}\limsup \sum e_{ni}^2 \int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-3}\left\{-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s)\}^{-8}\,dG(s)\right\}^{1/2} d\tilde{H}_i(t)$$

(7.16)
$$\begin{aligned}
&\geq \limsup b^{-1}\sum e_{ni}^2 \int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-3}\left(7^{-1}\int_{-\infty}^{t} dG^{-7}(s)\right)^{1/2} d\tilde{H}_i(t),\\
&\geq 7^{1/2}\limsup\left\{n\bar{H}(M)\right\}^{-1}\sum e_{ni}^2 \int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-3}(1-G(t))^2\, d\tilde{H}_i(t),\\
&\geq C\limsup n^{-1}\sum e_{ni}^2 \int_{-\infty}^{M} t^2\{\bar{H}(t)\}^{-4}(1-G(t))^2\, d\tilde{H}_i(t),
\end{aligned}$$

where we have used $(\bar{F})^{-7} \geq 1$, $G^{-7/2} \geq 1$, $(1-G^7)^{1/2} \geq 1 - G$, assumption A5 and the fact that $(\bar{H})^{-1}$ is increasing. The proof of Theorem 3.1 is now complete in view of (7.15), (7.2) and (7.1). □

PROOF OF THEOREM 3.2. We need to show that $n^{1/2}(\lambda_1\hat{\alpha} + \lambda_2\hat{\beta})$ is asymptotically normally distributed with suitable parameters for every pair (λ_1, λ_2) of real numbers. Let

(7.17)
$$\mathscr{V}_n = n^{1/2}\left\{\lambda_1\left(\hat{\alpha} + \int_{-\infty}^{M} t\,dL_{n1}(t)\right) + \lambda_2\left(\hat{\beta} + \int_{-\infty}^{M} t\,dL_{n2}(t)\right)\right\}.$$

The proof consists of approximating $\mathscr{V}_n$ by a U-statistic $\mathscr{U}_n$ and then applying Hoeffding (1948) to $\mathscr{U}_n$. Write

(7.18)
$$\begin{aligned}
\mathscr{V}_n &= n^{1/2}\left[\sum c_{ni}W_i\{\hat{G}^{-1}(Z_i) - G^{-1}(Z_i)\} + \sum c_{ni}\left\{W_iG^{-1}(Z_i) + \int_{-\infty}^{M} t\,dF_i(t)\right\}\right]\\
&= n^{1/2}(B_{n1} + B_{n2}), \qquad \text{say,}
\end{aligned}$$

where $c_{ni} = \lambda_1 a_{ni} + \lambda_2 b_{ni}$, $1 \leq i \leq n$. Note that B_{n2} is already a sum of centered independent rv's. We sketch a proof of approximating B_{n1} by a U-statistic.

LEMMA 7.3. *Under* A5 *and* A6

(7.19)
$$n^{1/2}\left|B_{n1} - \sum c_{ni}W_iG^{-1}(Z_i)\{\ln G(Z_i) - \ln \hat{G}(Z_i)\}\right| = o_p(1).$$

PROOF. Once again write $\hat{G}^{-1}$ and G^{-1} as $\exp(-\ln \hat{G})$ and $\exp(-\ln G)$, use Taylor's expansion up to the second term of e^x function and argue as in Susarla and Van Ryzin (1980) to get

$E\{\text{LHS}(7.19)\}$

$$\le n^{1/2}\sum |c_{ni}| E|W_i| E_i(\{\ln \hat{G}(Z_i) - \ln G(Z_i)\}^2[G^{-1}(Z_i) + \{1 + N^+(Z_i)\}^{-1}(n+1)]).$$

Now use the Cauchy-Schwarz inequality on E_i and (7.7) to conclude that

$$\begin{aligned}\limsup E\{\text{LHS}(7.19)\} &\le C \limsup \Bigg\{ n^{1/2}\sum |c_{ni}| \int_{-\infty}^{M} |t| \\ &\quad \cdot \{\bar{H}(t)\}^{-1}\Bigg[n^{-1}\bigg(-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s)\}^{-8}\, dG(s)\bigg)^{1/2} b^* \\ &\quad + n^{-2}\{\bar{H}(t)\}^{-3}(1 - G(t))^2\Bigg] G(t)\, d(-F_i(t)\Bigg\} \\ &\le Cb^{**} \limsup n^{-1/2}\int_{-\infty}^{M} |t|\{\bar{H}(t)\}^{-2} \\ &\quad \cdot \bigg(-\int_{-\infty}^{t} \bar{F}(s)\{\bar{H}(s)\}^{-6}\, dG(s)\bigg)^{1/2} G(t)\, d(-K_n(t)) \\ &= 0\end{aligned}$$

by A6. Here $b^{**} = b^* + 1$, and $K_n = |\lambda_1| K_{n1} + |\lambda_2| K_{n2}$. The second inequality follows because $(\bar{H})^{-1}$ is increasing and because of an inequality similar to (7.16) holds here also. □

LEMMA 7.4. *Under* A5 *and* A6

$$n^{1/2}|\sum c_{ni} W_i G^{-1}(Z_i)[\ln \hat{G}(Z_i) + \sum_{j=1}^{n} [\delta_j = 0, Z_j \le Z_i]\{2 + N^+(Z_j)\}^{-1}]| = o_p(1). \tag{7.20}$$

PROOF. Use (7.8), argue as just before (7.9) and use $W_i[\delta_i = 0] = 0$ to get

$$\begin{aligned}&E\{\text{LHS}(7.20)\}\\ &\quad \le Cn^{1/2}\sum |c_{ni}| E[|W_i| G^{-1}(Z_i) E_i\{\textstyle\sum_{j\ne i} [\delta_j = 0, Z_j \le Z_i] E_{i,j}(1 + N^+(Z_j))^{-2}\}].\end{aligned} \tag{7.21}$$

Now use A5, (7.6) and an equality like (7.16) together with A6 to conclude the lemma. □

LEMMA 7.5. *If* A6 *holds then*

$$\begin{aligned}&n^{1/2}\textstyle\sum_i c_{ni} W_i G^{-1}(Z_i)\sum_j[\delta_j = 0, Z_j \le Z_i]\{2 + N^+(Z_j)\}^{-1}\\ &\quad = n^{1/2}\textstyle\sum_i c_{ni} W_i G^{-1}(Z_i)\sum_j[\delta_j = 0, Z_j \le Z_i]\{2(\bar{H}(Z_j))^{-1} - \hat{H}_n(Z_j)(\bar{H}(Z_j))^{-2}\} + o_p(1).\end{aligned}$$

PROOF. The proof follows from the Corollary 7.1, the Cauchy-Schwarz inequality and

$$\begin{aligned}&n(2 + N^+(Z_j))^{-1} - \{2(\bar{H}(Z_j))^{-1} - \hat{H}_n(Z_j)(\bar{H}(Z_j))^{-2}\}\\ &\qquad = n^{-1}\{\bar{H}(Z_j)\}^{-2}(2 + N^+(Z_j))^{-1}(n\bar{H}(Z_j) - N^+(Z_j))^2\\ &\qquad\quad - 2n^{-2}\{\bar{H}(Z_j)\}^{-2}(2 + N^+(Z_j))^{-1} - 4\{\bar{H}(Z_j)\}^{-2}(2 + N(Z_j))^{-1} \quad □\end{aligned}$$

Therefore, Lemmas 7.3, 7.4 and 7.5 show that the asymptotic distribution of $\mathscr{V}_n$ is the same as that of

$$\begin{aligned}\mathscr{U}_n = n^{1/2}(B_{n2} &+ \textstyle\sum_i c_{ni} W_i G^{-1}(Z_i)\ln G(Z_i)\\ &+ n^{-1}\textstyle\sum_i c_{ni} W_i G^{-1}(Z_i)\sum_j[\delta_j = 0, Z_j \le Z_i](2\{\bar{H}(Z_j)\}^{-1}\\ &- \hat{H}_n(Z_j)\{\bar{H}(Z_j)\}^{-2}).\end{aligned} \tag{7.22}$$

To obtain a U-statistic approximation, write

$$\mathscr{U}_n = n^{1/2}\{B_{n2} + \sum c_{ni} W_i G^{-1}(Z_i) \ln G(Z_i) - \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \Phi_{n1}((\delta_i, Z_i), (\delta_j, Z_j)) - \binom{n}{3}^{-1} \sum_{1\le i<j<k\le n} \Phi_{n2}((\delta_i, Z_i), (\delta_j, Z_j), (\delta_k, Z_k))\}, \tag{7.23}$$

where

$$\Phi_{n1}((\delta_i, Z_i), (\delta_j, Z_j)) = (n-1) c_{ni} W_i (1-\delta_j)[Z_j \le Z_i \le M]\{G(Z_i)\bar{H}(Z_j)\}^{-1} + c_{nj} W_j (1-\delta_i)[Z_i \le Z_j \le M]\{G(Z_j)\bar{H}(Z_i)\}^{-1} \tag{7.24}$$

and

$$\Phi_{n2}((\delta_i, Z_i), (\delta_j, Z_j), (\delta_k, Z_k)) = (n-1)(n-2)(6n)^{-1}\Sigma^* c_{ni} W_i (1-\delta_j)[Z_j \le Z_i \le M][Z_k > Z_j]\{G(Z_i)\bar{H}^2(Z_j)\}^{-1} \tag{7.25}$$

where the Σ^* is the summation over all 6 permutations of i, j and k.

Since $\mathscr{U}_n$ is a U-statistic, we can now use the techniques of Hoeffding (1948) to show that the distribution of $\mathscr{U}_n$ is the same as that of

$$\hat{\mathscr{U}}_n = n^{1/2} \sum_i A_{ni} \tag{7.26}$$

where

$$A_{ni} := c_{ni}\Big(W_i G^{-1}(Z_i) + \int_{-\infty}^{M} t\, dF_i(t)\Big) - (1-\delta_i)\bar{H}^{-1}(Z_i) \int_{Z_i}^{M} t\, d\bar{L}_n(t) - \int_{-\infty}^{Z_i \wedge M} \bar{F}(t)\bar{H}^{-2}(t)\left\{\int_t^M s\, d\bar{L}_n(s)\right\} dG(t), \qquad 1 \le i \le n \tag{7.27}$$

and where $\bar{L}_n = \lambda_1 \bar{L}_{n1} + \lambda_2 \bar{L}_{n2}$ and $x \wedge y = \min(x, y)$.

The conditions needed for approximating $\mathscr{U}_n$ by $\hat{\mathscr{U}}_n$ are implied by A4, A5, and A6. Assumption A7 in turn enables one to apply the Lindeberg-Feller CLT to $\hat{\mathscr{U}}_n$. Finally, tedious but straightforward calculations obtain the expressions (3.6)–(3.8) for the asymptotic variances. □

REFERENCES

BUCKLEY, J. and JAMES, I. (1979). Linear regression with censored data. *Biometrika* **66** 429–436.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **16** 293–325.

HOEFFDING, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **27** 713–721.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Stat. Assoc.* **53** 457–481.

LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, New York.

MILLER, R. G. (1976). Least squares regression with censored data. *Biometrika* **63** 449–464.

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

H. KOUL
DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

V. SUSARLA
DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

J. VAN RYZIN
DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10025