Biometrika (1990), 77, 4, *pp.* 845–51
Printed in Great Britain

Linear regression analysis of censored survival data based on rank tests

BY L. J. WEI
Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706, U.S.A.
Z. YING
Department of Statistics, University of Illinois, Champaign, Illinois 61820, U.S.A.
AND D. Y. LIN
Department of Biostatistics, University of Washington, Seattle, Washington 98195, U.S.A.

SUMMARY

Recently linear rank statistics with censored data have been used as the estimating functions for the regression parameters in the linear model with an unspecified error distribution. The resulting rank estimators are consistent and asymptotically normal. However, the asymptotic variances of these estimators are complicated and are difficult to estimate well with censored data. In this paper, we propose some simple methods for making inference about a subset of the regression coefficients while regarding others as nuisance parameters. A lack-of-fit test for the linear model is also presented. The proposed procedures are illustrated with an example.

Some key words: Accelerated failure time model; Estimating function; Generalized Wilcoxon statistic; Goodness-of-fit; Log rank statistic; Partial likelihood.

1. INTRODUCTION

The Cox proportional hazards model (Cox, 1972) has been the most successful regression model for analyzing censored failure time data. Simple and efficient inference procedures for this semiparametric model can be obtained through the partial likelihood function (Cox, 1975). Their asymptotic properties have been justified by elegant martingale theory (Andersen & Gill, 1982).

An alternative to the Cox regression model is the accelerated failure time model, which relates covariates linearly to the logarithm of the survival time; compare Kalbfleisch & Prentice (1980, pp. 32–4, 143–62). This approach is quite appealing to medical investigators due to its ease of interpretation. However, when the distribution of the error term in this linear model is unspecified, the development of statistical methods which have practical usage and theoretical justification has been relatively slow. Among the important contributors in this area are Miller (1976), Prentice (1978), Buckley & James (1979) and Koul, Susarla & Van Ryzin (1981). In particular, Prentice (1978) proposed a class of linear rank statistics with censored data for testing the entire set of the regression coefficients based on the marginal probability of a generalized rank vector with various specified distributions of the error term in the accelerated failure time model. Again, martingale theory can be used to show that these statistics are asymptotically normal regardless of the actual sampling distribution for the error term (Cuzick, 1985).

Recently, Lai & Ying (1990) and Tsiatis (1990) have used the linear rank statistics with censored data as the estimation functions for the regression coefficients in the linear model with an unspecified error distribution. The resulting rank estimators are consistent and asymptotically normal. The asymptotic variances of these estimators depend on the density function and its derivative of the error term. The variance estimators suggested by Tsiatis (1990) involve nonparametric kernel estimation of the underlying density function. These estimators are rather complicated and unstable with censored data. In addition, their consistency has not been established. Therefore, Wald statistics based on the rank estimators and their estimated variances may be inappropriate for practical use.

In this paper, we propose some simple methods with theoretical justification for inference about a subset of the regression coefficients in the linear model while regarding others as nuisance parameters without using complicated and unreliable variance estimators. Our approach is similar to an informal proposal of Prentice (1978, p. 176). In addition, we present a simple lack-of-fit test for the linear model. The proposed methods are illustrated with a real-life example.

2. Inference procedures based on linear rank statistics

Let random variables T and C be the failure and censoring times of which the minimum X is observed. The failure indicator Δ takes the value of 1 if $X=T$ and 0 otherwise. Furthermore, let Z be a $p\times 1$ vector of covariates. Conditional on Z, random variables T and C are assumed to be independent. The data consist of n triplets (X_i, Δ_i, Z_i) $(i=1,\ldots,n)$ which are generated from n independent replicates of (T, C, Z).

We assume that the logarithm of the failure time T is linearly related to Z; that is

$$\log T_i = \beta' Z_i + \varepsilon_i \quad (i=1,\ldots,n), \tag{2·1}$$

where β is a $p\times 1$ vector of regression coefficients, and the ε_i's are independent and identically distributed with an unspecified distribution function F. The logarithm may be replaced by other strictly increasing functions.

If the true value of β is 0, then the linear rank statistic $n^{\frac{1}{2}}S_n$ is asymptotically normal with mean 0, where

$$S_n = n^{-1}\sum_{i=1}^{n} \Delta_i \phi\{\hat{F}(\log X_i)\}(Z_i - \bar{Z}_i), \tag{2·2}$$

ϕ is a twice continuously differentiable function on $[0, 1]$, $\hat{F}(.)$ is the left-continuous version of the Kaplan–Meier estimator for $F(.)$,

$$\bar{Z}_i = \sum_{j=1}^{n} Z_j I(X_j \geq X_i) \Big/ \sum_{j=1}^{n} I(X_j \geq X_i), \tag{2·3}$$

and $I(.)$ is the indicator function. Note that if $\phi(.)=1$, (2·2) is the Cox partial likelihood score statistic. If $\phi(u)=1-u$, S_n corresponds to the Peto–Prentice generalization of the Wilcoxon statistic (Peto & Peto, 1972; Prentice, 1978).

To avoid some technical difficulties in the derivation of the asymptotic theory for our inference procedures based on (2·2), the observations in the extreme upper tail of the data should be excluded from the analysis. This truncation can be achieved by including in the summation over i for (2·2) only those i's for which X_i satisfies

$$\frac{\log n}{n}\sum_{j=1}^{n} I(X_j \geq X_i) > \eta, \tag{2·4}$$

where η is some positive constant. In applications, however, the sample size n is finite, and with an appropriate choice of η all data points could be included in the analysis.

Now, for a given β, let $e_i(\beta) = \log X_i - \beta' Z_i$ $(i = 1, \ldots, n)$. Also, let $\hat{F}_\beta(.)$ be the left-continuous version of the Kaplan–Meier estimator based on the $e_i(\beta)$'s. Furthermore, let $S_n(\beta)$ be the resulting S_n by replacing $\log X_i$ and $\hat{F}(.)$ on the right-hand sides of (2·2), (2·3) and (2·4) with $e_i(\beta)$ and $\hat{F}_\beta(.)$, respectively. If the true value of β is β_0, then $n^{\frac{1}{2}} S_n(\beta_0)$ is also asymptotically normal with mean 0. The corresponding covariance matrix is asymptotically equivalent to $\Lambda(\beta_0)$, where

$$\Lambda(\beta) = n^{-1} \sum_{i=1}^{n} \phi^2[\hat{F}_\beta\{e_i(\beta)\}]\Delta_i \left[\frac{\Sigma_j I\{e_j(\beta) \geq e_i(\beta)\} Z_j^{\otimes 2}}{\Sigma_j I\{e_j(\beta) \geq e_i(\beta)\}} - \left\{ \frac{\Sigma_j I\{e_j(\beta) \geq e_i(\beta)\} Z_j}{\Sigma_j I\{e_j(\beta) \geq e_i(\beta)\}} \right\}^{\otimes 2} \right], \quad (2{\cdot}5)$$

where the sums Σ_j are over the range $j = 1, \ldots, n$, and $a^{\otimes 2}$ denotes aa' for a column vector a (Andersen & Gill, 1982; Andersen, Borgan et al., 1982).

The above results suggest that a reasonable estimator for β_0 is a value of β, denoted by $\hat{\beta}$, which solves the system of equations $\{S_n(\beta) = 0\}$. In general, no exact solutions can be obtained because $S_n(\beta)$ is a p-dimensional step function of β. We choose $\hat{\beta}$ which minimizes $\|S_n(\beta)\|$. Such estimators have been studied by Louis (1981) and Wei & Gail (1983) in the two-sample setting.

Suppose that we are interested in $\beta^{(1)}$, the first q $(1 \leq q \leq p)$ components of $\beta = [\beta^{(1)\prime}, \beta^{(2)\prime}]'$. The vectors $\hat{\beta}$ and β_0 are similarly partitioned. We show in Appendix 1 that, under some mild conditions, $n^{\frac{1}{2}}(\hat{\beta}^{(1)} - \beta_0^{(1)})$ is asymptotically normal with mean 0. The corresponding covariance matrix involves the density function f and its derivative f' of F. With censored data, it is rather difficult to obtain a useful estimator for this covariance matrix. Here, we propose some simple methods for making inference about $\beta^{(1)}$ without estimating the covariance matrix of $\hat{\beta}^{(1)}$.

Consider testing the hypothesis H_0: $\beta^{(1)} = \beta_0^{(1)}$. Define

$$Q(\beta^{(2)}; \beta_0^{(1)}) = nS_n'(\tilde{\beta})\Lambda^{-1}(\hat{\beta})S_n(\tilde{\beta}),$$

where $\tilde{\beta} = [\beta_0^{(1)\prime}, \beta^{(2)\prime}]'$. In Appendix 2, we show that the statistic $G_n(\beta_0^{(1)}) = \min Q(\beta^{(2)}; \beta_0^{(1)})$ is asymptotically distributed as χ_q^2 under H_0, where the minimization is taken with respect to $\beta^{(2)}$ around a neighbourhood of $\hat{\beta}^{(2)}$.

Confidence regions for $\beta^{(1)}$ with level $(1-\alpha)$ can also be constructed based on $G_n(\beta^{(1)})$. One such region is $\{\beta^{(1)}: G_n(\beta^{(1)}) \leq \chi_q^2(\alpha)\}$, where $\chi_q^2(\alpha)$ is the upper α point of the χ_q^2 distribution.

3. Lack-of-fit test for the linear model

In this section, we use an interesting idea of Gill & Schumacher (1987) for testing the proportional hazards assumption in the two-sample case to develop a simple test for the adequacy of the linear model. If the assumed model (2·1) is valid, two rank estimators of the parameter vector, say $\hat{\beta}_1$ and $\hat{\beta}_2$ with different weight functions ϕ_1 and ϕ_2, respectively, should be close to each other. Thus, a significant difference between $\hat{\beta}_1$ and $\hat{\beta}_2$ suggests model misspecification. However, the covariance matrix of $(\hat{\beta}_1 - \hat{\beta}_2)$ depends on the density function f and its derivative f'. Therefore, a lack-of-fit test for model (2·1) based on $(\hat{\beta}_1 - \hat{\beta}_2)$ would be unreliable in practice especially with censored observations.

Now, let $S_{nk}(\beta)$ be the statistic $S_n(\beta)$ defined in § 2 with weight function ϕ_k $(k=1,2)$. If model (2·1) is correct, then there exists a β such that both $S_{n1}(\beta)$ and $S_{n2}(\beta)$ converge in probability to zero as $n\rightarrow\infty$. Thus, if there does not exist a common value of β such that both $S_{n1}(\beta)$ and $S_{n2}(\beta)$ are close to 0, then the assumed model may be inappropriate. This motivates the following lack-of-fit test based on $S_{nk}(\beta)$ $(k=1,2)$.

Under model (2·1), there exists a β such that the statistic $n^{\frac{1}{2}}[S'_{n1}(\beta), S'_{n2}(\beta)]'$ is asymptotically normal with mean 0 and with a covariance matrix that is asymptotically equivalent to

$$\begin{bmatrix}\Lambda_{11}(\beta) & \Lambda_{12}(\beta)\\ \Lambda_{21}(\beta) & \Lambda_{22}(\beta)\end{bmatrix},$$

where $\Lambda_{kl}(\beta)$ $(k,l=1,2)$ are obtained from $\Lambda(\beta)$ defined in (2·5) with $\phi^2(.)$ being replaced by $\phi_k(.)\phi_l(.)$. By the arguments presented in Appendices 1 and 2, under model (2·1), which implies that the limit of $(\hat\beta_1-\hat\beta_2)$, say δ, is 0, the statistic

$$H(\phi_1,\phi_2)=\min\left\{n\begin{bmatrix}S_{n1}(\beta)\\ S_{n2}(\beta+\delta)\end{bmatrix}'\begin{bmatrix}\Lambda_{11}(\hat\beta_1) & \Lambda_{12}(\hat\beta_1)\\ \Lambda_{21}(\hat\beta_1) & \Lambda_{22}(\hat\beta_1)\end{bmatrix}^{-1}\begin{bmatrix}S_{n1}(\beta)\\ S_{n2}(\beta+\delta)\end{bmatrix}\right\} \qquad (3\cdot1)$$

is asymptotically distributed as χ^2_p, where the minimization is taken with respect to β in a neighbourhood around $\hat\beta_1$. A large value of $H(\phi_1,\phi_2)$ suggests model misspecification.

It is easy to see from the arguments in Appendix 2 that the test statistic $H(\phi_1,\phi_2)$ is asymptotically equivalent to the Wald statistic based on $(\hat\beta_1-\hat\beta_2)$ under model (2·1). Thus, the lack-of-fit test based on $H(\phi_1,\phi_2)$ can be interpreted in a very natural way. In addition, it involves little extra work to perform such a simple test for the fit of the assumed linear model after the inference about the regression coefficients has been done with various ϕ functions. Naturally, the power of the lack-of-fit test against a given alternative depends on the choice of ϕ_1 and ϕ_2. A formal investigation into the power properties of the test statistic $H(\phi_1,\phi_2)$ needs to be undertaken along the same lines as Gill & Schumacher (1987).

4. An example

As an illustration, we applied the proposed methods to the well-known Stanford heart transplant data as of February 1980 reported by Miller & Halpern (1982). These data contain the survival times of 184 heart-transplanted patients along with their ages at the time of the first transplant and T5 mismatch scores. The 27 patients who did not have T5 mismatch scores were excluded from our analysis. Out of the remaining 157 patients, 55 were censored as of February 1980. These data were previously analyzed by Miller & Halpern (1982) using the Cox and Buckley–James regression methods.

Since $n^{-1}\log n\simeq 0\cdot03$ in this example, we first chose η in condition (2·4) to be 0·03 for our analysis. For this particular η, all observations were included in the calculation of $S_n(\beta)$ and $\Lambda(\beta)$. Several different values of η were also tried. The results of analysis remained virtually unchanged for $\eta\leqslant 0\cdot3$ because most of the large survival times in this data set were censored. For simplicity, only the results for $\eta=0\cdot03$ are reported here.

Following Miller & Halpern (1982), we first regressed the base 10 logarithm of the survival time on age and T5 mismatch score. The estimation results are presented in Table 1. The Buckley–James estimates given by Miller & Halpern (1982) are also included in the table. The confidence intervals based on the new methods indicate that the covariate age is significant at the 5% level whereas T5 mismatch score is not. The point estimates

Table 1. *Point estimates and* 95% *confidence intervals for the linear regression of* $\log_{10}$ *of survival time on age and* T5 *mismatch score with* $n = 157$ *Stanford heart transplant patients*

Methods	Age Estimate	Age Confidence interval	T5 Estimate	T5 Confidence interval
Proposed, $\phi(u) = 1$	−0·025	(−0·047, −0·007)	−0·124	(−0·395, 0·197)
Proposed, $\phi(u) = 1 - u$	−0·021	(−0·041, −0·004)	−0·062	(−0·331, 0·248)
Buckley-James	−0·015	(−0·031, 0·001)	−0·003	(−0·266, 0·260)

Table 2. *Point estimates and* 95% *confidence intervals for the linear regression of* $\log_{10}$ *of survival time on age and age*^{2} *with* $n = 152$ *Stanford heart transplant patients who survived at least* 10 *days*

Methods	Age Estimate	Age Confidence interval	Age^2 Estimate	Age^2 Confidence interval
Proposed, $\phi(u) = 1$	0·099	(−0·007, 0·189)	−0·0016	(−0·0028, −0·0003)
Proposed, $\phi(u) = 1 - u$	0·102	(0·003, 0·207)	−0·0016	(−0·0030, −0·0004)
Buckley-James	0·107	(0·034, 0·180)	−0·0017	(−0·0027, −0·0007)

using the log rank weight function are considerably smaller than those of the Peto-Prentice weight function, which suggests that this additive model may be inadequate. Indeed, the lack-of-fit test $H(\phi_1, \phi_2)$ in (3·1) with these two weight functions yields a value of 5·03, the corresponding p-value being 0·08. The residual plots of Miller & Halpern (1982) also discredited this model.

In an attempt to achieve a better fit, a quadratic age model without T5 mismatch score was tried by Miller & Halpern (1982). For their analysis, Miller & Halpern (1982) deleted the 5 patients with survival times less than 10 days to symmetrize the data. For comparison, we also excluded these 5 observations in our quadratic age model although the new procedures are less sensitive to outliers. The results for this model are displayed in Table 2. The observed value of $H(\phi_1, \phi_2)$ based on the log rank versus Peto-Prentice weight functions is 0·81 with 2 degrees of freedom, yielding a p-value of 0·67. This reflects the closeness of the parameter estimates with these two weight functions and confirms the graphical findings of Miller & Halpern (1982). The estimates based on the new approach are comparable to the Buckley-James estimates except that the latter approach gives a considerably larger lower bound of the confidence interval for the linear age effect.

One should be cautious in interpreting the estimates of the linear age effect in Table 2 since age and age^2 are highly correlated. To avoid this multicollinearity, we also fitted the quadratic age model with age being centred around 42, the approximate sample mean of the patients' ages. The results for the covariate $(\text{age} - 42)^2$ with this reparameterization are identical to those of age^2 in Table 2. However, the point estimates for the covariate (age − 42) are −0·038 and −0·036 based on the log rank and Peto-Prentice weight functions, respectively. The corresponding 95% confidence intervals are (−0·061, −0·020) and (−0·058, −0·018), which clearly indicate a negative linear age effect.

Acknowledgements

This research was supported by the National Institutes of Health and the National Science Foundation. The authors are grateful to a referee for helpful comments and to Dr A. A. Tsiatis for providing them with his paper before its publication.

APPENDIX 1

Asymptotic properties of $S_n(\beta)$ and $\hat{\beta}$

The asymptotic properties of the rank estimator $\hat{\beta}$ have been carefully studied by Lai & Ying (1990) and Tsiatis (1990) under somewhat different setups. However, Lai & Ying (1990) only dealt with the problem of a single regression coefficient, and Tsiatis (1990) did not verify the consistency of $\Lambda(\hat{\beta})$. Here we generalize the results of Lai & Ying (1990) to the multidimensional case.

Using results of Andersen & Gill (1982) and Andersen, Borgan et al. (1982), one can show that, even without condition (2·4), $n^{\frac{1}{2}}S_n(\beta_0)$ is asymptotically normal with mean 0. In addition, the corresponding limiting covariance matrix is asymptotically equivalent to $\Lambda(\beta_0)$.

To study the consistency of $\hat{\beta}$ and the asymptotic linearity of $S_n(\beta)$ around β_0, we require the following assumptions. The covariate vector Z has bounded support. The parameter space $\mathscr{B}$ of β is compact. The density function f of the error term ε is bounded. The random variables log C and ε have finite second moments. The density functions of $(\log C-\beta'Z)$ are bounded uniformly in β. Furthermore, suppose that, for some constant $\xi>0$,

$$\int_{-\infty}^{\infty}\sup_{|t|\leqslant\xi}\left\{\frac{f'(s+t)}{f(t)}\right\}^2 f(s)ds<\infty, \tag{A·1}$$

$$\sup_{0\leqslant u\leqslant 1}\left|\frac{d^2\phi(u)}{du^2}\right|<\infty. \tag{A·2}$$

Then, as $n\to\infty$, almost surely,

$$\sup_{\beta\in\mathscr{B}}\|S_n(\beta)-S(\beta)\|\to 0, \tag{A·3}$$

$$S_n(\beta)-S_n(\beta_0)=\Gamma(\beta-\beta_0)+o_p\{\max(n^{-\frac{1}{2}},\|\beta-\beta_0\|)\}, \tag{A·4}$$

where

$$S(\beta)=\int_{-\infty}^{T^*(\beta)}\phi\{F_\beta(t)\}\left\{d\tau_1(t,\beta)-\frac{\mu_1(t,\beta)}{\mu_0(t,\beta)}d\tau_0(t,\beta)\right\},$$

$$\Gamma=\int_{-\infty}^{T^*(\beta_0)}\phi\{F(t)\}\left\{\frac{f'(t)}{f(t)}+\frac{f(t)}{1-F(t)}\right\}\left\{\mu_2(t,\beta_0)-\frac{\mu_1(t,\beta_0)^{\otimes 2}}{\mu_0(t,\beta_0)}\right\}\frac{dF(t)}{1-F(t)},$$

$$\mu_r(t,\beta)=E[Z^{\otimes r}I\{e(\beta)\geqslant t\}]\quad(r=0,1,2),$$

$$\tau_r(t,\beta)=E[Z^{\otimes r}I\{e(\beta)\leqslant t,\Delta=1\}]\quad(r=0,1),$$

$$T^*(\beta)=\sup\{t:\mu_0(t,\beta)>0\},$$

and $F_\beta(t)$ is the limit of $\hat{F}_\beta(t)$. Results (A·3) and (A·4) are straightforward multivariate generalizations of Theorem 2 of Lai & Ying (1990).

Now, if $\{S(\beta)=0\}$ has a unique root at β_0, then it follows from (A·3) and the continuity of $S(\beta)$ in β that $\hat{\beta}\to\beta_0$ almost surely as $n\to\infty$. Furthermore, the arguments that led to (4·7), (4·8) and (4·37) of Lai & Ying (1990) can be used to show that $\Lambda(\hat{\beta})-\Lambda(\beta_0)\to 0$ in probability as $n\to\infty$.

Remark 1. Condition (A·1) is rather mild and is satisfied by most familiar distributions.

Remark 2. Condition (A·2) is satisfied by at least two weight functions, $\phi(u)=1$ and $\phi(u)=1-u$, which correspond to the most commonly used log rank test and Peto-Prentice-Wilcoxon test, respectively.

Remark 3. Condition (2·4) of § 2 is essential to establish (A·4).

Remark 4. If Γ is nonsingular, then $n^{\frac{1}{2}}(\hat{\beta}-\beta_0)$ is asymptotically normal with mean 0 and with a covariance matrix that is asymptotically equivalent to $\Gamma^{-1}\Lambda(\beta_0)\Gamma^{-1}$.

Appendix 2

Asymptotic distribution of $G_n(\beta_0^{(1)})$

Let $\mathcal{N}(\hat\beta^{(2)})$ be a small neighbourhood of $\hat\beta^{(2)}$ in the sense that

$$\mathcal{N}(\hat\beta^{(2)}) = (\beta^{(2)}: \|\beta^{(2)} - \hat\beta^{(2)}\| < n^{-1/3}d)$$

for some given $d > 0$. It follows from (A·4) that

$$n^{\frac{1}{2}}S_n(\hat\beta) - n^{\frac{1}{2}}S_n(\tilde\beta) = n^{\frac{1}{2}}\Gamma(\hat\beta - \tilde\beta) + o_p(n^{\frac{1}{2}}\|\hat\beta - \beta_0\|) + o_p(n^{\frac{1}{2}}\|\tilde\beta - \beta_0\|),$$

where $\tilde\beta = [\beta_0^{(1)\prime}, \beta^{(2)\prime}]'$. By the definition of $\hat\beta$, $n^{\frac{1}{2}}S_n(\hat\beta) = o_p(1)$. This implies that, for $\beta^{(2)} \in \mathcal{N}(\hat\beta^{(2)})$,

$$n^{\frac{1}{2}}S_n(\tilde\beta) = -n^{\frac{1}{2}}\Gamma(\hat\beta - \tilde\beta) + o_p(n^{\frac{1}{2}}\|\hat\beta^{(2)} - \beta^{(2)}\|) + o_p(1),$$

where both $o_p(.)$ terms do not depend on $\beta^{(2)}$. Hence, for $\beta^{(2)} \in \mathcal{N}(\hat\beta^{(2)})$,

$$Q(\beta^{(2)}; \beta_0^{(1)}) = n\begin{bmatrix}\hat\beta^{(1)} - \beta_0^{(1)}\\ \hat\beta^{(2)} - \beta^{(2)}\end{bmatrix}'\Gamma\Lambda^{-1}(\beta_0)\Gamma\begin{bmatrix}\hat\beta^{(1)} - \beta_0^{(1)}\\ \hat\beta^{(2)} - \beta^{(2)}\end{bmatrix} + o_p(n\|\hat\beta^{(2)} - \beta^{(2)}\|^2) + o_p(1).$$

This implies that the statistic $G_n(\beta_0^{(1)}) = \min Q(\beta^{(2)}; \beta_0^{(1)})$, where the minimization is taken with respect to $\beta^{(2)}$ in $\mathcal{N}(\hat\beta^{(2)})$, is asymptotically equivalent to

$$\min\left\{n\begin{bmatrix}\hat\beta^{(1)} - \beta_0^{(1)}\\ \hat\beta^{(2)} - \beta^{(2)}\end{bmatrix}'\Gamma\Lambda^{-1}(\beta_0)\Gamma\begin{bmatrix}\hat\beta^{(1)} - \beta_0^{(1)}\\ \hat\beta^{(2)} - \beta^{(2)}\end{bmatrix}\right\}, \tag{A·5}$$

where the minimization is taken with respect to $\beta^{(2)}$. It can be shown through some matrix algebra that (A·5) equals

$$n(\hat\beta^{(1)} - \beta_0^{(1)})'([I_q, 0]\{\Gamma\Lambda^{-1}(\beta_0)\Gamma\}^{-1}[I_q, 0]')^{-1}(\hat\beta^{(1)} - \beta_0^{(1)}),$$

where I_q is the $q \times q$ identity matrix. Note that the term $[I_q, 0]\{\Gamma\Lambda^{-1}(\beta_0)\Gamma\}^{-1}[I_q, 0]'$ of this quadratic form is asymptotically equivalent to the limiting covariance matrix of $n^{\frac{1}{2}}(\hat\beta^{(1)} - \beta_0^{(1)})$. Therefore, $G_n(\beta_0^{(1)})$ is asymptotically distributed as χ^2 with q degrees of freedom.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. & Keiding, N. (1982). Linear nonparametric tests for comparison of counting processes, with applications to censored data. *Int. Statist. Rev.* **50**, 219-58.

Andersen, P. K. & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10**, 1100-20.

Buckley, J. & James, I. (1979). Linear regression with censored data. *Biometrika* **66**, 429-36.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *J. R. Statist. Soc.* B **34**, 187-220.

Cox, D. R. (1975). Partial likelihood. *Biometrika* **62**, 269-76.

Cuzick, J. (1985). Asymptotic properties of censored linear rank tests. *Ann. Statist.* **13**, 133-41.

Gill, R. D. & Schumacher, M. (1987). A simple test of the proportional hazards assumption. *Biometrika* **74**, 289-300.

Kalbfleisch, J. D. & Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Koul, H., Susarla, V. & Van Ryzin, J. (1981). Regression analysis with randomly right censored data. *Ann. Statist.* **9**, 1276-88.

Lai, T. L. & Ying, Z. (1990). Linear rank statistics in regression analysis with censored or truncated data. *J. Mult. Anal.* To appear.

Louis, T. A. (1981). Nonparametric analysis of an accelerated failure time model. *Biometrika* **68**, 381-90.

Miller, R. G. (1976). Least squares regression with censored data. *Biometrika* **63**, 449-64.

Miller, R. & Halpern, J. (1982). Regression with censored data. *Biometrika* **69**, 521-31.

Peto, R. & Peto, J. (1972). Asymptotically efficient rank invariance test procedures (with discussion). *J. R. Statist. Soc.* A **135**, 185-206.

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65**, 167-79.

Tsiatis, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18**, 354-72.

Wei, L. J. & Gail, M. H. (1983). Nonparametric estimation for a scale-change with censored observations. *J. Am. Statist. Assoc.* **78**, 382-8.

[*Received November* 1989. *Revised May* 1990]