The Annals of Statistics
1981, Vol. 9, No. 1, 109–121

THE COX REGRESSION MODEL, INVARIANCE PRINCIPLES FOR SOME INDUCED QUANTILE PROCESSES AND SOME REPEATED SIGNIFICANCE TESTS[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

For the Cox regression model, the partial likelihood functions involve linear combinations of induced order statistics. Some invariance principles pertaining to such linear combinations of induced order statistics are studied and the theory is incorporated in the formulation of some repeated significance tests (for the hypothesis of no regression) based on these partial likelihoods.

1. Introduction. In the Cox (1972) regression model for survival data, it is assumed that the ith subject (having *survival time* Y_i and a set of *covariates* $\mathbf{Z}_i = (Z_{i1}, \cdots, Z_{ip})'$ for some $p \geq 1$) has the *hazard rate* (given $\mathbf{Z}_i = \mathbf{z}_i$)

$$h_i(t) = h_0(t)\exp(\boldsymbol{\beta}'\mathbf{z}_i), \qquad i = 1, \cdots, n, t \geq 0, \tag{1.1}$$

where $h_0(t)$, the hazard rate for $\mathbf{z}_i = \mathbf{0}$, is an unknown, arbitrary nonnegative function (for which $\int_0^\infty h_0(t)\,dt = \infty$) and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)'$ parameterizes the regression of survival time on the covariates. We assume that $h_0(t)$ is continuous in t almost everywhere (a.e.), so that ties among the Y_i may be neglected, with probability 1. Let $\mathbf{Q}_n = (Q_1, \cdots, Q_n)'$, the vector of antiranks, be defined by

$$Y_{Q_i} = Y_{ni} \quad \text{for} \quad i = 1, \cdots, n, \tag{1.2}$$

where $Y_{n1} < \cdots < Y_{nn}$ are the order statistics corresponding to $Y_1, \cdots, Y_n$. Then, following Bhattacharya (1974), $\mathbf{Z}_{Q_1}, \cdots, \mathbf{Z}_{Q_n}$ are termed the *induced order statistics.* In the event of no loss in the follow-up, the *partial* (log−) *likelihood function* when all the failures have been observed (cf. Cox (1972, 1975)) is given by

$$\log L_n = \sum_{i=1}^n \{\boldsymbol{\beta}'\mathbf{Z}_{Q_i} - \log(\sum_{j=i}^n \exp\{\boldsymbol{\beta}'\mathbf{Z}_{Q_j}\})\}. \tag{1.3}$$

We consider here the following scheme where all the n subjects enter into the study at a common point of time, so that the failures are observed in order. However, to incorporate possible withdrawals (drop-outs) of subjects from the scheme, we conceive of a set of withdrawal (censoring) times $W_1, \cdots, W_n$ where the W_i are independent and identically distributed random variables (i.i.d.rv's) with a distribution function (df) $G(t)$, $t \geq 0$. Then, the observable rv's are $(Y_i^0, \delta_i, \mathbf{Z}_i)$ where $Y_i^0 = Y_i \wedge W_i = \min(Y_i, W_i)$ and $\delta_i = 1$ or 0 according as Y_i^0 is $= Y_i$ or not, for $i = 1, \cdots, n$. Note that by assumption the W_i are independent of the Y_i and $\mathbf{Z}_i$. Also, for the Y_i^0, the hazard rates are given by

$$g_0(t) + h_i(t), i = 1, \cdots, n; \quad g_0(t) = -(\partial/\partial t)\log[1 - G(t)], \qquad t \geq 0, \tag{1.4}$$

where the $h_i(t)$ are defined by (1.1). Thus, if $Y_{n1}^0 < \cdots < Y_{nn}^0$ be the order statistics

Received January 1979; revised September 1979.

[1] Work supported by the National Heart, Lung and Blood Institute, Contract No. NIH-HNLBI-71-2243 from the National Institutes of Health.

AMS 1970 *subject classification.* 60B10, 60F05, 62G99.

Key words and phrases. Analysis of covariance, contiguity of probability measures, Cox regression model, hazard rate, induced order statistics, invariance principles, quantile processes, repeated significance tests, survival functions, Wiener processes, withdrawals.

corresponding to $Y_1^0, \cdots, Y_n^0$ and if $Y_{Q_i^0}^0 = Y_{ni}^0$, $i = 1, \cdots, n$, then the partial log-likelihood function when all the Y_i^0 have been observed is given by

$$\log L_n^0 = \sum_{i=1}^n \{\log[g_0(Y_{ni}^0) + h_0(Y_{ni}^0)\exp\{\boldsymbol{\beta}'\mathbf{Z}_{Q_i^0}\}] \tag{1.5}$$
$$- \log(\sum_{j=i}^n [g_0(Y_{ni}^0) + h_0(Y_{ni}^0)\exp\{\boldsymbol{\beta}'\mathbf{Z}_{Q_j^0}\}])\}.$$

As (1.4) invalidates the proportionality of the hazard function, (1.5) depends on the unknown g_0, h_0 (as well as the Y_i^0) and is of not much use. However, if $T = \{t_1 < \cdots < t_m\} = \{Y_i^0: \delta_i = 1, i = 1, \cdots, n\}$ be the set of *failure points* (for which W_i exceeds Y_i), then a second partial likelihood function may be defined as follows. At time $t_j - 0$, there is a risk set $\mathscr{R}_j$ or r_j individuals which have neither failed nor dropped out by that time, for $j = 1, \cdots, m$, so that $\mathscr{R}_m \subset \cdots \subset \mathscr{R}_1$. Considering the risk set $\mathscr{R}_j$ and the conditional probability of a failure at time t_j, for $j = 1, \cdots, m$, we obtain on using (1.4) the partial log-likelihood function

$$\log L_m^* = \sum_{j=1}^m \{\boldsymbol{\beta}'\mathbf{Z}_{Q_j^*} - \log(\sum_{i\in\mathscr{R}_j} \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\})\}. \tag{1.6}$$

where $\mathbf{Q}^* = (Q_1^*, \cdots, Q_m^*)'$ is a (random) subvector of $\mathbf{Q}$. This corresponds to the model of Cox (1972), though Cox has primarily in mind the case of a staggered entry and a fixed point of termination, leading to possibly different censoring times for the n subjects. In the sequel we shall refer to the Cox model in the set up of this nonstaggering entry and random withdrawal model. A discrete version of (1.6) has also been considered by Cox (1972) and we shall refer to that in Section 3.

For testing the hypothesis of no regresson viz.,

$$H_0: \boldsymbol{\beta} = \mathbf{0} \quad \text{vs.} \quad H_1: \boldsymbol{\beta} \neq \mathbf{0}, \tag{1.7}$$

Cox(1972) considered the test statistic

$$\mathscr{L}_{nm}^* = \mathbf{U}_{nm}^{*\prime}\mathbf{J}_{nm}^{*-}\mathbf{U}_{nm}^*, \tag{1.8}$$

where

$$\mathbf{U}_{nm}^* = (\partial/\partial\boldsymbol{\beta})\log L_{nm}^*|_{\boldsymbol{\beta}=\mathbf{0}}, \qquad \mathbf{J}_{nm}^* = -(\partial^2/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}')\log L_{nm}^*|_{\boldsymbol{\beta}=\mathbf{0}}, \tag{1.9}$$

and $\mathbf{A}^-$ stands for the generalized inverse of $\mathbf{A}$. Cox (1972, 1975) argued heuristically that under H_0, $\mathscr{L}_{nm}^*$ has asymptotically chi-square distribution with p degrees of freedom (DF). In a variety of situations, relating to clinical trials and life-testing experimentations, one may be interested in monitoring the study from the very beginning with the objective of an early termination if H_0 in (1.6) is not tenable. Such a plan is known as a *progressively censored scheme* (PCS) (vis., Chatterjee and Sen (1973) and Sen (1976, 1979). Thus, in a PCS, instead of making a *terminal test* at the mth failure t_m, one may like to review the process at each failure t_j, $j > 1$ and stop experimentation as soon as $\mathscr{L}_{nj}^*$ (defined as in (1.8)–(1.9), but, based on $\mathbf{U}_{nj}^*$ and $\mathbf{J}_{nj}^*$), leads to the rejection of H_0, for some $j \leq m$; if $\mathscr{L}_{n1}^*, \cdots, \mathscr{L}_{nm}^*$ are all insignificant, then H_0 is accepted along with the termination of the study at the preplanned time. Hence, a *repeated significance testing* (RST) procedure is involved in a PCS. We may note that by (1.6) and (1.9),

$$\mathbf{U}_{nk}^* = \sum_{j=1}^k \{\mathbf{Z}_{Q_j^*} - r_j^{-1}\sum_{i\in\mathscr{R}_j} \mathbf{Z}_i\} \qquad \text{for} \quad k = 1, \cdots, m, \tag{1.10}$$

and hence, these are all linear combinations of induced order statistics. We first study (in Sections 2, 3 and 4) some invariance principles relating to these induced order statistics (under the null as well as local alternative hypotheses) and in the concluding section, we incorporate these invariance principles for the study of the (asymptotic) properties of some RST procedures.

2. Weak convergence of some induced quantile processes. For convenience of presentation, we first consider the uncensored case, and by analogy to (1.3), (1.6) and (1.9), we let for every $k: 1 \leq k \leq n$,

$$\log L_{nk} = \sum_{j=1}^k \{\boldsymbol{\beta}'\mathbf{Z}_{Q_j} - \log\{\sum_{i=j}^n \exp\{\boldsymbol{\beta}'\mathbf{Z}_{Q_i}\}\}\}, \tag{2.1}$$

$$\mathbf{U}_{nk} = (\partial/\partial\boldsymbol{\beta})\log L_{nk}|_{\boldsymbol{\beta}=\mathbf{0}} \tag{2.2}$$
$$= \sum_{j=1}^{k} \{\mathbf{Z}_{Q_j} - (n-j+1)^{-1} \sum_{i=j}^{n} \mathbf{Z}_{Q_i}\},$$

$$\mathbf{J}_{nk} = -(\partial^2/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}')\log L_{nk}|_{\boldsymbol{\beta}=\mathbf{0}} \tag{2.3}$$
$$= \sum_{j=1}^{k} (n-j+1)^{-1}(n-j)\mathbf{S}_{nj}, \quad \text{say,}$$

where $S_{nn} = \mathbf{0}$ and

$$\mathbf{S}_{nj} = (n-j)^{-1} \sum_{i=j}^{n} (\mathbf{Z}_{Q_i} - \bar{\mathbf{Z}}_j^*)(\mathbf{Z}_{Q_i} - \bar{\mathbf{Z}}_j^*)' \quad \text{for } j = 1, \cdots, n-1, \tag{2.4}$$

$$\bar{\mathbf{Z}}_j^* = (n-j+1)^{-1} \sum_{i=j}^{n} \mathbf{Z}_{Q_i}, \qquad j = 1, \cdots, n. \tag{2.5}$$

Conventionally, we let $\mathbf{U}_{n0} = \mathbf{0}$ and $\mathbf{J}_{n0} = \mathbf{0}$, for every $n > 1$. Also, throughout this paper, the covariates $\mathbf{Z}_1, \cdots, \mathbf{Z}_n$ are assumed to be stochastic vectors; there are some simplifications when these are nonstochastic and these will be briefly considered later on. In the usual custom of an analysis of covariance model, we assume that $\mathbf{Z}_1, \cdots, \mathbf{Z}_n$ are i.i.d. rv's with $\boldsymbol{\mu} = E\mathbf{Z}_i$ and $\boldsymbol{\Gamma} = E(\mathbf{Z}_i - \boldsymbol{\mu})(\mathbf{Z}_i - \boldsymbol{\mu})'$, and assume that

$$\boldsymbol{\Gamma} \text{ is positive definite (p.d.) with } \det \boldsymbol{\Gamma} < \infty. \tag{2.6}$$

Note that the $\mathbf{Z}_i$ are all observable at the beginning of the experimentation, $Q_1, \cdots, Q_k$ are observable at the kth failure with $\sum_{i=k}^{n} \mathbf{Z}_{Q_i} = \sum_{i=1}^{n} \mathbf{Z}_i - \sum_{j=1}^{k-1} \mathbf{Z}_{Q_j}$ is also observable at the kth failure, and hence, there is no difficulty in computing the $\mathbf{U}_{nk}$ and $\mathbf{J}_{nk}$ at the successive failures. We are primarily interested in the asymptotic behavior of the partial sequences $\{\mathbf{U}_{nk}: 0 \le k \le n\}$ and $\{\mathbf{J}_{nk}: 0 \le k \le n\}$.

For every $n(\ge 1)$, we consider a stochastic process $\boldsymbol{\xi}_n = \{\boldsymbol{\xi}_n(t), t \in E = [0, 1]\}$ by letting

$$\boldsymbol{\xi}_n(t) = \mathbf{J}_{nn}^{-1/2}\mathbf{U}_{n[nt]}, \qquad 0 \le t \le 1, \tag{2.7}$$

where $[s]$ denotes the largest integer contained in s and $\mathbf{J}_n^{-1/2} = \mathbf{B}_n$ is defined by $\mathbf{B}_n\mathbf{J}_{nn}\mathbf{B}_n' = \mathbf{I}_p$. Then, $\boldsymbol{\xi}_n$ belongs to the space $D^p[0, 1]$, endowed with the Skorokhod J_1-topology. Also, let $\boldsymbol{\xi}(t) = (\xi_{(1)}(t), \cdots, \xi_{(p)}(t))'$, $t \in E$ and $\xi_{(j)} = \{\xi_{(j)}(t), t \in E\}$, $j = 1, \cdots, p$ be independent copies of a standard Wiener process on E. Then the main theorem of this section is the following

THEOREM 2.1. *Under H_0: $\boldsymbol{\beta} = \mathbf{0}$ and* (2.6), *$\boldsymbol{\xi}_n$ weakly converges to $\boldsymbol{\xi} = \{\boldsymbol{\xi}(t), t \in E\}$.*

The proof of Theorem 2.1 rests on a martingale characterization of $\{\mathbf{U}_{nk}\}$ (and $\{\mathbf{S}_{nk}\}$) and some invariance principles for such martingales studied by Scott (1973) and McLeish (1974), among others. For this reason, first, we consider the following. Let $\mathscr{B}_{nk} = \mathscr{B}(\mathbf{Z}_1, \cdots, \mathbf{Z}_n; Q_1, \cdots, Q_k)$ be the sigma-field generated by $\mathbf{Z}_1, \cdots, \mathbf{Z}_n$ and $Q_1, \cdots, Q_k$ for $k = 1, \cdots, n$ and let $\mathscr{B}_{n0} = \mathscr{B}(\mathbf{Z}_1, \cdots, \mathbf{Z}_n)$. Then, for every $n(\ge 1)$, $\mathscr{B}_{nk}$ is nondecreasing in $k(\le n)$.

LEMMA 2.2. *Under H_0: $\boldsymbol{\beta} = \mathbf{0}$, for every $n(>1)$, $\{\mathbf{U}_{nk}, \mathscr{B}_{nk}; 0 \le k \le n\}$ is a martingale.*

PROOF. Note that by (2.2), for every k: $1 \le k \le n$,

$$\mathbf{U}_{nk} - \mathbf{U}_{nk-1} = \mathbf{Z}_{Q_k} - (n-k+1)^{-1} \sum_{i=k}^{n} \mathbf{Z}_{Q_i}. \tag{2.8}$$

Now, given $\mathscr{B}_{nk-1}$, under H_0, Q_k can take on any one value in the set $(1, \cdots, n\backslash(Q_1, \cdots, Q_{k-1})$ with the equal conditional probability $(n-k+1)^{-1}$, so that $E(\mathbf{Z}_{Q_k} | \mathscr{B}_{nk-1}) = (n-k+1)^{-1} [\sum_{i=1}^{n} \mathbf{Z}_i - \sum_{i=1}^{k-1} \mathbf{Z}_{Q_i}] = (n-k+1)^{-1} \sum_{i=k}^{n} \mathbf{Z}_{Q_i}$. Thus, $E(\mathbf{U}_{nk} - \mathbf{U}_{nk-1} | \mathscr{B}_{nk-1}, H_0) = \mathbf{0}$ a.e., for $k = 1, \cdots, n$. □

LEMMA 2.3. *Under H_0: $\boldsymbol{\beta} = \mathbf{0}$, for every $n(\ge 2)$, $\{\mathbf{S}_{nk} - \boldsymbol{\Gamma}, \mathscr{B}_{nk}; 0 \le k \le n\}$ is a martingale.*

PROOF. By the same arguments as in the proof of Lemma 2.2, under H_0,

$$E\{(\mathbf{U}_{nk} - \mathbf{U}_{nk-1})(\mathbf{U}_{nk} - \mathbf{U}_{nk-1})' | \mathscr{B}_{nk-1}\} = (n-k)(n-k+1)^{-1}\mathbf{S}_{nk} \ \forall\ k = 1, \cdots, n, \tag{2.9}$$

where the $\mathbf{S}_{nk}$ are defined by (2.4). Note that if we let

$$\phi(\mathbf{a}, \mathbf{b}) = \tfrac{1}{2}(\mathbf{a} - \mathbf{b})(\mathbf{a} - \mathbf{b})', \qquad \mathbf{a}, \mathbf{b} \in R^p, \tag{2.10}$$

then, we have by (2.4),

$$\mathbf{S}_{nk} = \binom{n-k+1}{2}^{-1} \textstyle\sum_{k \le i < j \le n} \phi(\mathbf{Z}_{Q_i}, \mathbf{Z}_{Q_j}) \qquad \text{for} \quad k = 1, \cdots, n-1. \tag{2.11}$$

As such, by the same technique as in the proof of Lemma 2.2, we have under H_0,

$$E(\mathbf{S}_{nk} \mid \mathscr{B}_{nk-1}) = \mathbf{S}_{nk-1} \quad \text{a.e.,} \qquad \text{for every} \quad k = 1, \cdots, n. \tag{2.12}$$

Also, note that

$$\mathbf{S}_{n1} = \binom{n}{2}^{-1} \textstyle\sum_{1 \le i < j \le n} \phi(\mathbf{Z}_i, \mathbf{Z}_j)$$

is a U-statistic of degree 2 (based on the i.i.d. rv's $\mathbf{Z}_1, \cdots, \mathbf{Z}_n$), and hence, $E\mathbf{S}_{n1} = E\phi(\mathbf{Z}_1, \mathbf{Z}_2) = \mathbf{\Gamma}$. Thus, by (2.12) and the above, we have $E(\mathbf{S}_{nk} - \mathbf{\Gamma} \mid \mathscr{B}_{nk-1}, H_0) = \mathbf{S}_{nk-1} - \mathbf{\Gamma}$ a.e., $\forall\, k = 1, \cdots, n$. □

For a $p \times p$ matrix $\mathbf{A} = ((a_{ij}))$, we let $\|\mathbf{A}\| = \max\{|a_{ij}|: 1 \le i, j \le p\}$. Then, by Lemma 2.3, we have for every $n(\ge 2)$, under H_0: $\boldsymbol{\beta} = \mathbf{0}$,

$$\{\|\mathbf{S}_{nk} - \mathbf{\Gamma}\|, \mathscr{B}_{nk};\ 1 \le k \le n-1\} \quad \text{is a nonnegative submartingale.} \tag{2.13}$$

LEMMA 2.4. *Under* (2.6) *and* H_0: $\boldsymbol{\beta} = \mathbf{0}$,

$$\max_{1 \le k \le n} \|n^{-1}(\mathbf{J}_{nk} - k\Gamma)\| \to_p 0, \text{ as } n \to \infty. \tag{2.14}$$

PROOF. Let $\{k_n\}$ be any sequence of positive integers ($k_n \le n$), such that

$$k_n \to \infty \qquad \text{but} \quad n^{-1}k_n^2 \to 0 \qquad \text{as} \quad n \to \infty. \tag{2.15}$$

Then, by (2.13), (2.15) and the generalized Kolmogorov-inequality for submartingales, under H_0: $\beta = \mathbf{0}$ and (2.6), for every $t > 0$,

$$p\{\max_{1 \le k \le n-k_n} \|\mathbf{S}_{nk} - \mathbf{\Gamma}\| > t\} \le t^{-1}E\|\mathbf{S}_{nn-k_n} - \mathbf{\Gamma}\|, \tag{2.16}$$

$$P\{\max_{n-k_n+1 \le k \le n-1} \|\mathbf{S}_{nk} - \mathbf{\Gamma}\| > t\} \le t^{-1}E\|\mathbf{S}_{nn-1} - \mathbf{\Gamma}\|. \tag{2.17}$$

Note that by Lemma 1 of Bhattacharya (1974), under H_0: $\boldsymbol{\beta} = \mathbf{0}$, $\mathbf{Z}_{Q_1} = \mathbf{Z}_1^*, \cdots, \mathbf{Z}_{Q_n} = \mathbf{Z}_n^*$ are i.i.d. rv's and $\mathbf{Z}_i^*$ and $\mathbf{Z}_i$ both have the same df. Hence, by (2.11), we have

$$\begin{aligned} E\|\mathbf{S}_{nn-k} - \mathbf{\Gamma}\| &= E\left\| \binom{k+1}{2}^{-1} \textstyle\sum_{1 \le i < j \le k+1} \{\phi(\mathbf{Z}_i^*, \mathbf{Z}_j^*) - \mathbf{\Gamma}\} \right\| \\ &= E\|\mathbf{U}(\mathbf{Z}_1^*, \cdots, \mathbf{Z}_{k+1}^*) - \mathbf{\Gamma}\| \qquad \text{for every} \quad k = 1, \cdots, n-1, \end{aligned} \tag{2.18}$$

where $\mathbf{U}(\mathbf{Z}_1^*, \cdots, \mathbf{Z}_m^*)$ is a matrix of Hoeffding's (1948) U-statistics, for every $m \ge 2$. Let $\mathscr{C}_m$ be the sigma-field generated by the unordered collection $\{\mathbf{Z}_1^*, \cdots, \mathbf{Z}_m^*\}$ and by $\mathbf{Z}_{m+1}^*, \mathbf{Z}_{m+2}^*, \cdots$, so that $\mathscr{C}_m$ is nonincreasing in $m(\ge 1)$. Then (cf. Berk (1966)). $\{\mathbf{U}(\mathbf{Z}_1^*, \cdots, Z_m^*), \mathscr{C}_m;\ m \ge 2\}$ forms a reverse martingale sequence, so that $\{\|\mathbf{U}(\mathbf{Z}_1^*, \cdots, \mathbf{Z}_m^*) - \mathbf{\Gamma}\|, \mathscr{C}_m;\ m \ge 2\}$ forms a reverse submartingale sequence, and hence, by the reverse submartingale convergence theorem,

$$E\|\mathbf{U}(\mathbf{Z}_1^*, \cdots, \mathbf{Z}_m^*) - \mathbf{\Gamma}\| \qquad \text{converges to} \quad 0 \qquad \text{as} \quad m \to \infty. \tag{2.19}$$

Further,

$$\max_{1 \le k \le n} \|n^{-1}\mathbf{J}_{nk} - n^{-1}k\mathbf{\Gamma}\|$$

$$
\begin{aligned}
&= \max_{1\le k\le n} \| n^{-1} \textstyle\sum_{i=1}^{k} \{(n-i)(n-i+1)^{-1}(\mathbf{S}_{ni} - \mathbf{\Gamma}) + (n-i+1)^{-1}\mathbf{\Gamma}\} \| \\
(2.20) \quad &\le \max_{1\le k\le n-k_n} \| \mathbf{S}_{nk} - \mathbf{\Gamma} \| + n^{-1}k_n\{\max_{n-k_n+1\le k\le n-1} \| \mathbf{S}_{nk} - \mathbf{\Gamma} \|\} \\
&\quad + \| \mathbf{\Gamma} \| \{\textstyle\sum_{j=1}^{n} j^{-1}\}/n.
\end{aligned}
$$

Hence, using (2.16) (for $t = \epsilon/3$), (2.17) (for $t = k_n$), (2.15), (2.19) and (2.20), we conclude that for every $\epsilon > 0$, $P\{\max_{1\le k\le n} \| n^{-1}\mathbf{J}_{nk} - n^{-1}k\mathbf{\Gamma} \| > \epsilon\} \to \mathbf{0}$ as $n \to \infty$. □

$$(2.21) \qquad n^{-1}\mathbf{J}_{nn} \to_p \mathbf{\Gamma} \quad \text{and is p.d., in probability.}$$

Let $I(A)$ stand for the indicator function of the set A and let $\tilde{Z}_n = \max_{1\le k\le n} | \mathbf{Z}_k | = \max_{1\le k\le n} (\mathbf{Z}_k'\mathbf{Z}_k)^{1/2}$. Then, we have the following

LEMMA 2.5. *Under H_0: $\boldsymbol{\beta} = \mathbf{0}$, for every $\epsilon > 0$, as $n \to \infty$,*

$$(2.22) \quad \| n^{-1} \textstyle\sum_{i=1}^{n} \{(\mathbf{U}_{ni} - \mathbf{U}_{ni-1})(\mathbf{U}_{ni} - \mathbf{U}_{ni-1})' I(| \mathbf{U}_{ni} - \mathbf{U}_{ni-1} | > \epsilon \sqrt{n}) | \mathscr{B}_{ni-1}\}) \| \to_p 0.$$

PROOF. Note that by (2.8),

$$(2.23) \qquad \max_{1\le k\le n} | \mathbf{U}_{nk} - \mathbf{U}_{nk-1} | \le 2\{\max_{1\le k\le n} | \mathbf{Z}_k |\} = 2\tilde{Z}_n.$$

Also, by (2.6), $| \mathbf{Z}_i |$, $i = 1, \cdots, n$ are i.i.d. rv's with a finite second moment, so that

$$(2.24) \qquad P\{\max_{1\le k\le n} | \mathbf{Z}_k | > \epsilon n^{1/2}\} \to 0 \quad \text{as} \quad n \to \infty \qquad \forall \epsilon > 0.$$

Finally, as in the proof of Lemma 2.2, the left-hand side of (2.22) is given by

$$
\begin{aligned}
&\| n^{-1} \textstyle\sum_{i=1}^{n} (n-i+1)^{-1} \sum_{j=i}^{n} (\mathbf{Z}_{Q_j} - \bar{\mathbf{Z}}_i^*)(\mathbf{Z}_{Q_j} - \bar{\mathbf{Z}}_i^*)' I(| \mathbf{Z}_{Q_j} - \bar{\mathbf{Z}}_i^* | > \epsilon \sqrt{n}) \| \\
(2.25) \quad &\le \{n^{-1} \textstyle\sum_{i=1}^{n} \mathrm{Trace}[(n-i)(n-i+1)^{-1}\mathbf{S}_{ni}]\} I(\tilde{Z}_n > \tfrac{1}{2} \epsilon \sqrt{n}) \\
&= [\mathrm{Trace}(n^{-1}\mathbf{J}_{nn})] I(\tilde{Z}_n > \tfrac{1}{2} \epsilon \sqrt{n}).
\end{aligned}
$$

Hence, (2.22) follows from (2.23), (2.24), (2.25) and (2.21). □

We are now in a position to prove Theorem 2.1. By virtue of Lemma 2.2 and (2.25), we are tempted to use invariance principles for martingales (viz., Scott (1973) and McLeish (1974) and, for this purpose, we consider the following lemma which is a direct multivariate generalization of Theorem (3.8) of McLeish (1974) (and hence, the proof is omitted). Let $\{\mathbf{X}_{ni}, i = 1, \cdots, m_n; n \ge 1\}$ be a triangular array of random vectors, $\{k_n(t)\}$ be a sequence of integer valued, nonnegative, right-continuous functions on $E = [0, 1]$ ($k_n(0) = 0$) and let $\mathbf{W}_n = \{\mathbf{W}_n(t), t \in E\}$ be defined by

$$(2.26) \qquad \mathbf{W}_n(t) = \textstyle\sum_{i\le k_n(t)} \mathbf{X}_{ni}, \qquad t \in E.$$

Let $\mathbf{E}_{ni}$ denote the conditional expectation given $\mathbf{X}_{nk}$, $k \le i$, for $i > 0$ and, finally, let ξ be defined as in Theorem 2.1.

LEMMA 2.6. *Suppose that for every $t \in E$ and $\epsilon > 0$,*

$$(2.27) \qquad \textstyle\sum_{i\le k_n(t)} E_{ni-1}\{\mathbf{X}_{ni}\mathbf{X}_{ni}' I(| \mathbf{X}_{ni} | > \epsilon)\} \to_p \mathbf{0},$$

$$(2.28) \qquad \textstyle\sum_{i\le k_n(t)} | E_{ni-1}\mathbf{X}_{ni} | \to_p \mathbf{0} \text{ and } \sum_{i\le k_n(t)} E_{ni-1}\mathbf{X}_{ni}\mathbf{X}_{ni}' \to_p t\mathbf{I}_p.$$

Then, $\mathbf{W}_n$ converges in distribution to ξ on $D^p[0, 1]$.

If we now let $\mathbf{X}_{ni} = n^{-1/2}(\mathbf{U}_{ni} - \mathbf{U}_{ni-1})$, $i = 1, \cdots, n$ and $k_n(t) = [nt]$, for $t \in E$, then, we have $\sum_{i\le k_n(t)} E_{ni-1}\mathbf{X}_{ni}\mathbf{X}_{ni}' = n^{-1}\mathbf{J}_{n[nt]}$, $\forall\, t \in E$. As such, (2.28) follows from Lemmas 2.2 and 2.4, while (2.27) follows from (2.22). Consequently, Theorem 2.1 follows from Lemmas 2.2, 2.4, 2.5 and 2.6. □

Note that by virtue of the Courant theorem (on the ratio of two quadratic forms) and by (2.14),

$$\max_{1\le k\le n} |\mathbf{U}'_{nk}\mathbf{J}_{nn}^{-1}\mathbf{U}_{nk}/\mathbf{U}'_{nk}(n\boldsymbol{\Gamma})^{-1}\mathbf{U}_{nk} - 1| \le \max\{|\mathrm{ch}_1(n\boldsymbol{\Gamma}\mathbf{J}_{nn}^{-1}) - 1|, |\mathrm{ch}_p(n\boldsymbol{\Gamma}\mathbf{J}_{nn}^{-1}) - 1|\} \to_p 0 \quad \text{as} \quad n\to\infty, \tag{2.29}$$

where ch_1 and ch_p stand for the largest and smallest characteristic roots. On the other hand, by Lemma 2.2, under (2.6) and H_0: $\boldsymbol{\beta} = \mathbf{0}$,

$$\{n^{-1}\mathbf{U}'_{nk}\boldsymbol{\Gamma}^{-1}\mathbf{U}_{nk}, \mathscr{B}_{nk}; 0 \le k \le n\} \quad \text{is a nonnegative submartingale.} \tag{2.30}$$

Also, note that for every $k: 1 \le k \le n$,

$$\begin{aligned} E\{n^{-1}\mathbf{U}'_{nk}\boldsymbol{\Gamma}^{-1}\mathbf{U}_{nk}\} &= E\{n^{-1}\mathrm{Trace}[\boldsymbol{\Gamma}^{-1}\mathbf{U}_{nk}\mathbf{U}'_{nk}]\} \\ &= E\{n^{-1}\mathrm{Trace}[\boldsymbol{\Gamma}^{-1}\mathbf{J}_{nk}]\} \\ &= n^{-1}\textstyle\sum_{i=1}^{k}(n-i)(n-i+1)^{-1}E(\mathrm{Trace}[\boldsymbol{\Gamma}^{-}s\mathbf{S}_{ni}]) \\ &= pn^{-1}\textstyle\sum_{i=1}^{k}(n-i)(n-i+1)^{-1} \le pk/n, \quad \text{as} \quad E\mathbf{S}_{ni} = \boldsymbol{\Gamma} \quad \text{for} \quad i = 1, \cdots, n. \end{aligned} \tag{2.31}$$

Therefore, by (2.30), (2.31) and Theorem 2.1 of Birnbaum and Marshall (1961), for any $\{a_{n1} \ge a_{n2} \ge \cdots \ge a_{nn} > 0\}$, $0 < \delta < 1$, $\epsilon > 0$,

$$P\{\max_{1\le k\le[n\delta]} a_{nk}|n^{-1}\mathbf{U}'_{nk}\boldsymbol{\Gamma}^{-1}\mathbf{U}_{nk}| > \epsilon\} \le \textstyle\sum_{k=1}^{[n\delta]}(a_{nk} - z_{nk+1})pk/n\epsilon \le pn^{-1}(a_{n1} + \cdots a_{n[n\delta]})/\epsilon. \tag{2.32}$$

Thus, if $q = \{q(t), t \in E\}$ be a nonnegative, nondecreasing and continuous function of t such that

$$\int_0^1 [q(t)]^{-2}\,dt < \infty, \tag{2.33}$$

then $n^{-1}(q^{-2}(1/n) + \cdots + q^{-2}([n\delta]/n)) \le \int_0^\delta [q(t)]^{-2}\,dt$, $\forall\ 0 < \delta < 1$. By choosing $\delta(>0)$ adequately small and using (2.33), it follows that the right-hand side of (2.32) can be made small when $a_{ni} = q^{-2}(i/n)$, $i = 1, \cdots, n$. On the other hand, for $t > \delta(>0)$, $q^{-2}(t) \le q^{-2}(\delta) < \infty$, so that for $\delta < t \le 1$, the weak convergence of $\boldsymbol{\xi}_n$ in the Skorokhod metric insures the same under the *sup-norm metric*

$$\rho_q(x, y) = \sup\{|x(t) - y(t)|/q(t), t \in E\}. \tag{2.34}$$

This leads us to the following

THEOREM 2.7. *Under* (2.6), (2.33) *and* $H_0: \boldsymbol{\beta} = \mathbf{0}$, *the weak convergence in Theorem* 2.1 *holds in the sup-norm metric* ρ_q.

Let us now consider the censored case. Let $h_0(t)$ be defined as in (1.1) and let $F_0(x) = 1 - \exp\{-\int_0^x h_0(t)\,dt\}$ be the df of Y_i under $H_0: \boldsymbol{\beta} = \mathbf{0}$. Then, defining T and its cardinality m as in before (1.6), we have

$$m/n \to \Pi = \int_0^\infty F_0(x)\,dG(x) > 0, \text{ a.s.}, \tag{2.35}$$

whenever the supports of the df F_0 and G are overlapping, as will be assumed in the sequel. Thus, m a.s. goes to ∞ as $n \to \infty$. We define $\mathscr{R}_k$, r_k, L^*_{nk} and $\mathbf{U}^*_{nk}$ as in Section 1 (for $k = 1, \cdots, m$), and let $\mathscr{B}^*_{nk}$ be the sigma-field generated by the risk set $\mathscr{Q}_k$, $k = 1, \cdots, m$. Since under H_0, Y_i, W_i and $\mathbf{Z}_i$ are mutually independent and $(Y_i, W_i, \mathbf{Z}_i)$ are i.i.d. rv's, defining the

Q_j^* as in (1.6), we claim that under H_0, Q_j^* has a uniform conditional distribution (given $\mathscr{R}_j$) over the set of r_j realizations in $\mathscr{R}_j$, so that

$$E\{\mathbf{Z}_{Q_j^*} - r_j^{-1} \sum_{i\in\mathscr{R}_j} \mathbf{Z}_i \mid \mathscr{B}_{nj}^*\} = \mathbf{0} \quad \text{for every} \quad j = 1, \cdots, m. \tag{2.36}$$

Similarly,

$$\begin{aligned} &E\{[\mathbf{Z}_{Q_j^*} - r_j^{-1} \textstyle\sum_{i\in\mathscr{R}_j} \mathbf{Z}_i][\mathbf{Z}_{Q_j^*} - r_j^{-1} \textstyle\sum_{i\in\mathscr{R}_j} \mathbf{Z}_i]' \mid \mathscr{B}_{nj}^*\} \\ &\quad = r_j^{-1} \textstyle\sum_{i\in\mathscr{R}_j} [\mathbf{Z}_i - r_j^{-1} \textstyle\sum_{k\in\mathscr{R}_j} \mathbf{Z}_k][\mathbf{Z}_i - r_j^{-1} \textstyle\sum_{k\in\mathscr{R}_j} \mathbf{Z}_k]' \\ &\quad = r_j^{-1}(r_j - 1)\, \mathbf{S}_{nj}^*, \quad \text{say, for} \quad j = 1, \cdots, m. \end{aligned} \tag{2.37}$$

As such, Lemmas 2.2, 2.3, 2.4 and 2.5 all hold for the censored case provided we replace the $\mathbf{U}_{nk}$, $\mathbf{S}_{nk}$, $\mathbf{J}_{nk}$ and $\mathscr{B}_{nk}$ by $\mathbf{U}_{nk}^*$, $\mathbf{S}_{nk}^*$, $\mathbf{J}_{nk}^*$ and $\mathscr{B}_{nk}^*$, respectively. For intended brevity and similarity of the techniques, the proofs are omitted. Thus, if we define $\xi_n^* = \{\xi_n^*(t),\, t \in E\}$ by letting $\xi_n^*(t) = \mathbf{J}_{nm}^{*-1/2}\mathbf{U}_{n[mt]}^*$, $t \in E$, then, parallel to Theorems 2.1 and 2.7, we have the following.

THEOREM 2.8. *The weak convergence of ξ_n^* to ξ holds under $H_0: \boldsymbol{\beta} = \mathbf{0}$ (both in the Skorokhod metric and the sup-norm metric).*

3. Weak convergence in the discrete time model. As in Section 1, we conceive of m risk sets $\mathscr{R}_1 \supset \cdots \supset \mathscr{R}_m$ where $\mathscr{R}_k$ has r_k subjects whose survival times are $\geq t_k$, $k \geq 1$, $0 = t_0 < t_1 < \cdots < t_m < t_{m+1} = \infty$ and we denote by

$$D_j = [t_j, t_{j+1}), \qquad j = 0, \cdots, m. \tag{3.1}$$

Let there be s_j failures in the interval D_j, $j \geq 0$ and let

$$\Omega_{ij} = I(Y_i \in D_j),\ \Omega_{ij}^* = \textstyle\sum_{k\geq j} \Omega_{ik} \quad \text{for} \quad j = 0, \cdots, m, \quad i = 1, \cdots, n. \tag{3.2}$$

Then, proceeding as in Section 6 of Cox (1972), we obtain the derivatives of the partial log-likelihood functions (evaluated at $\boldsymbol{\beta} = \mathbf{0}$) as

$$\mathbf{U}_{nk}^{**} = \textstyle\sum_{j=1}^k \{\sum_{i=1}^n (\boldsymbol{\Omega}_{ij} - \boldsymbol{\Omega}_{ij}^* s_j/r_j)\mathbf{Z}_i\}, \qquad k = 1, \cdots, m, \tag{3.3}$$

$$\begin{aligned} \mathbf{J}_{nk}^{**} &= \textstyle\sum_{j=1}^k [r_j(r_j - 1)]^{-1} s_j(r_j - s_j) \sum_{i=1}^n (\mathbf{Z}_i\Omega_{ij}^* - \bar{\mathbf{Z}}_j^*)(\mathbf{Z}_i\Omega_{ij}^* - \bar{\mathbf{Z}}_j^*)' \\ &= \textstyle\sum_{j=1}^k [s_j(r_j - s_j)/r_j]\mathbf{S}_{nj}^{**} \quad \text{say,} \quad k = 1, \cdots, m, \end{aligned} \tag{3.4}$$

where

$$\bar{\mathbf{Z}}_j^* = (\textstyle\sum_{i=1}^n \mathbf{Z}_i\Omega_{ij}^*)/r_j \quad \text{for} \quad j = 1, \cdots, m. \tag{3.5}$$

(The close relationship between (2.4)–(2.5) and (3.4)–(3.5) need not be overemphasized.) In this case, we let $\mathscr{B}_{nk}^{**} = \mathscr{B}(\mathbf{Z}_1, \cdots, \mathbf{Z}_n, s_1, \cdots, s_m, r_1, \cdots, r_m, \Omega_{ij}, j \leq k, i = 1, \cdots, n)$ for $k = 1, \cdots, m$, while $\mathscr{B}_{n0}^{**} = \mathscr{B}(\mathbf{Z}_1, \cdots, \mathbf{Z}_n, s_1, \cdots, s_m, r_1, \cdots, r_m)$. Then, by arguments very similar to those in the proof of Lemma 2.2, we arrive at the following.

LEMMA 3.1. *Under* (2.6) *and $H_0: \boldsymbol{\beta} = \mathbf{0}$, for every $n(\geq 1)$, $\{\mathbf{U}_{nk}^{**}, \mathscr{B}_{nk}^{**};\ 1 \leq k \leq m\}$ is a martingale.*

Also, parallel to Lemma 2.4, we have under (2.6) and $H_0: \boldsymbol{\beta} = \mathbf{0}$,

$$\max_{1\leq k\leq m} \| n^{-1}(\mathbf{J}_{nk}^{**} - \textstyle\sum_{j=1}^k s_j(r_j - s_j)r_j^{-1}\boldsymbol{\Gamma})\| \to_p 0 \quad \text{as} \quad n \to \infty. \tag{3.6}$$

Now, to study the desired weak convergence results, we consider two different cases:

(I). m is fixed, so that as $n \to \infty$, s_j and r_j both increase for every $j (= 1, \cdots, m)$. This situation arises when we have a given number of ordered categories, while n is large.

(II). $m = m(n) \to \infty$ but $\max_{1 \le j \le m(n)} n^{-1} s_j \to 0$ as $n \to \infty$. This situation arises when the width of D_j ($j \ge 1$) is small, so that there are a large number of cells with small probabilities, but the possibility of ties is no longer negligible.

In case (I), by virtue of Lemma 3.1, $\mathbf{U}_{nk}^{**} - \mathbf{U}_{nk-1}^{**}$, $k \ge 1$ are all uncorrelated. Moreover, given $\mathscr{B}_{nk-1}^{**}$, the conditional distribution of

$$n^{-1/2}(\mathbf{U}_{nk}^{**} - \mathbf{U}_{nk-1}^{**}) = n^{-1/2}\{\textstyle\sum_{i=1}^{n} \mathbf{Z}_i(\Omega_{ik} - s_k r_k^{-1} \Omega_{ik}^{*})\} \tag{3.7}$$

is generated by the $r_k!$ equally likely realizations of the Ω_{ik} (over the set $\mathscr{R}_k = \{i: \Omega_{ik}^{*} = 1\}$) and by an appeal to the classical permutational central limit theorem (Hájek (1961)), we conclude that this conditional distribution is asymptotically (in probability) multinormal with null mean vector and dispersion matrix

$$s_k(r_k - s_k) r_k^{-1} \mathbf{S}_{nk}^{**} = \mathbf{\Sigma}_{(k)} \quad \text{say}, \quad (k = 1, \cdots, m). \tag{3.8}$$

Thus, using a chain of conditioning (for $k = m - 1, \cdots, 1$), it follows by some routine steps that given $\mathscr{B}_{n0}^{**}$, the joint conditional distribution of $\{n^{-1/2}(\mathbf{U}_{nk}^{**} - \mathbf{U}_{nk-1}^{**}),\ 1 \le k \le m\}$ is asymptotically (in probability) multinormal with null mean vector and a dispersion matrix which is block-diagonal with the matrices $\mathbf{\Sigma}_{(k)}$, $k = 1, \cdots, m$. This, in turn, insures the asymptotic multinormality of $\{n^{-1/2}\, \mathbf{U}_{nk}^{**},\ k = 1, \cdots, m\}$ when (2.6) and H_0: $\boldsymbol{\beta} = \mathbf{0}$ hold.

In case (II), $m = m(n) \to \infty$ as $n \to \infty$. Hence, using Lemma 3.1, (3.6) and proceeding on the same line as in the proof of Theorem 2.1, it follows that a similar invariance principle holds when we replace $\mathbf{U}_{nk}$, $\mathbf{S}_{nk}$ and $\mathbf{J}_{nk}$ by $\mathbf{U}_{nk}^{**}$, $\mathbf{S}_{nk}^{**}$ and $\mathbf{J}_{nk}^{**}$, respectively.

4. Invariance principles under local alternatives. We like to extend the invariance principles studied in earlier sections when H_0: $\boldsymbol{\beta} = \mathbf{0}$ may not hold. We conceive of a sequence $\{K_n\}$ of local (Pitman-type) alternative hypotheses, where

$$K_n\colon \boldsymbol{\beta} = \boldsymbol{\beta}_{(n)} = n^{-1/2}\boldsymbol{\lambda} \quad \text{for some} \quad \boldsymbol{\lambda} \in R^P, \tag{4.1}$$

and desire to study the weak convergence results under $\{K_n\}$. Let us define the rv Y_i^0 as in Section 1 and let $\Psi(y)$ be the df of Y_i^0 under H_0: $\boldsymbol{\beta} = \mathbf{0}$. Then $[1 - \Psi(y)] = [1 - G(y)][1 - F_0(y)]$, $\forall\, y \ge 0$, where G and F_0 are defined in Sections 1 and 2. Also, let $h_0(t)$ and $g_0(t)$ be defined as in Section 1 and let $\pi(t) = h_0(t)/[h_0(t) + g_0(t)]$ for $t \ge 0$. Further let

$$\Pi(\alpha) = \int_0^{\Psi^{-1}(\alpha)} \pi(z)\, d\Psi(z), \qquad \forall\, \alpha \in [0, 1], \tag{4.2}$$

so that $\Pi(1) = \Pi$ is defined by (2.35). Finally, let $t^* = \inf\{u: \Pi(u)/\Pi(1) \ge t\}$, $t \in E$,

$$\zeta^* = \{\zeta^*(t) = \Pi^{-1/2}\Pi(t^*)\mathbf{\Gamma}^{1/2}\boldsymbol{\lambda} = \Pi^{1/2} t \mathbf{\Gamma}^{1/2}\boldsymbol{\lambda},\ t \in E\} \tag{4.3}$$

and we define the processes ξ_n and ξ as in Theorem 2.8. Then, we have the following.

THEOREM 4.1. *Under* (1.1), (2.6) *and* $\{K_n\}$ *in* (4.1), ξ_n^* *converges weakly to* $\xi + \zeta^*$.

Note that if the W_i are all equal to $+\infty$, with probability 1 (i.e., there is no withdrawal from the scheme), then $\pi(t) = 1$, $\forall\, t \ge 0$, so that ζ^* reduces to $\zeta = \{\zeta(t) = t\mathbf{\Gamma}^{1/2}\boldsymbol{\lambda},\ t \in E\}$. Thus, if we define ξ_n and ξ as in Theorem 2.1, from Theorem 4.1 we arrive at the following.

COROLLARY 4.1. *Under* (1.1), (2.6) *and* $\{K_n\}$ *in* (4.1), ξ_n *converges weakly to* $\xi + \zeta$.

For the proofs of these results, we employ the concept of contiguity of probability measures (as in Chapter VI of Hájek and Šidák (1967)). Let P_n and P_n^* be respectively the joint distributions of $(Y_i^0, \delta_i, \mathbf{Z}_i)$, $i = 1, \cdots, n$ under H_0: $\boldsymbol{\beta} = \mathbf{0}$ and K_n. Then, as a first step for the proof of Theorem 4.1, we consider the following.

THEOREM 4.2. *Under* (1.1), (2.6) *and* (4.1), $\{P_n^*\}$ *is contiguous to* $\{P_n\}$.

PROOF. Let $H_0(x)$ be a nondecreasing and nonnegative function defined by $(d/dx)H_0(x) = h_0(x)$. Then, the conditional density of (Y_i^0, δ_i), given $\mathbf{Z}_i$ is

$$[1 - G(Y_i^0)][\exp\{-H_0(Y_i^0)e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}]\{g_0(Y_i^0)\}^{1-\delta_i}\{h_0(Y_i^0)e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}^{\delta_i} \qquad i = 1, \cdots, n. \tag{4.4}$$

Since the marginal distribution of the $\mathbf{Z}_i$ does not depend on $\boldsymbol{\beta}$, we obtain from (4.4) that for testing H_0: $\boldsymbol{\beta} = \mathbf{0}$ vs. K_n: $\boldsymbol{\beta} = n^{-1/2}\boldsymbol{\lambda}$, the log-likelihood ratio statistic is

$$\log \mathscr{L}_n^0 = n^{-1/2}\textstyle\sum_{i=1}^n \{\delta_i \boldsymbol{\lambda}'\mathbf{Z}_i + H_0(Y_i^0)n^{1/2}(1 - e^{n^{-1/2}\boldsymbol{\lambda}'\mathbf{Z}_i})\}. \tag{4.5}$$

Note that by definition, $1 - \Psi(y) = (1 - G(y))\exp\{-H_0(y)\}$, so that $d\Psi(y) = [\exp\{-H_0(y)\}] \cdot [(1 - G(y))\, dF_0(y) + dG(y)]$. Hence, for every $r > 0$, under H_0: $\boldsymbol{\beta} = \mathbf{0}$, $E[H_0(Y_i^0)]^r < \infty$. Therefore, by the Khintchine law of large numbers, under H_0: $\boldsymbol{\beta} = \mathbf{0}$,

$$n^{-1} \textstyle\sum_{i=1}^n [H_0(Y_i)]^r \to E[H_0(Y_1^0)]^r, \text{ a.s.} \quad \text{as} \quad n \to \infty. \tag{4.6}$$

Further, by (2.24), uniformly in $i (1 \le i \le n)$,

$$|n^{1/2}(1 - e^{n^{-1/2}\boldsymbol{\lambda}'\mathbf{Z}_i}) + \boldsymbol{\lambda}'\mathbf{Z}_i + n^{-1/2}(\boldsymbol{\lambda}'\mathbf{Z}_i)^2| = o_p(n^{-1/2})(\boldsymbol{\lambda}'\mathbf{Z}_i)^2 \quad \text{as} \quad n \to \infty. \tag{4.7}$$

Also, using the independence of Y_i^0 and $\mathbf{Z}_i$ [under H_0: $\boldsymbol{\beta} = \mathbf{0}$] and the Khintchine law of large numbers, we have under H_0,

$$\begin{aligned} n^{-1} \textstyle\sum_{i=1}^n H_0(Y_i^0)(\boldsymbol{\lambda}'\mathbf{Z}_i)^2 &\to EH_0(Y_1^0)E(\boldsymbol{\lambda}'\mathbf{Z}_i\mathbf{Z}_i'\boldsymbol{\lambda}) = EH_0(Y_1^0)[\boldsymbol{\lambda}'(\boldsymbol{\Gamma} + \boldsymbol{\mu}\boldsymbol{\mu}')\boldsymbol{\lambda}] \text{ a.s.} \\ &= \nu^2, \text{ say.} \end{aligned} \tag{4.8}$$

Finally, note that by (4.5) through (4.8), as $n \to \infty$,

$$\log \mathscr{L}_n^0 = n^{-1/2} \textstyle\sum_{i=1}^n (\boldsymbol{\lambda}'\mathbf{Z}_i)(\delta_i - H_0(Y_i^0)) - \tfrac{1}{2}\nu^2 + o_p(1), \tag{4.9}$$

where under H_0: $\boldsymbol{\beta} = \mathbf{0}$, for every $i (= 1, \cdots, n)$,

$$\begin{aligned} E[H_0(Y_i^0)] &= -\int_0^\infty H_0(y)\, d[1 - \Psi(y))] = \int_0^\infty [1 - \Psi(y)]\, dH_0(y) \\ &= \int_0^\infty (1 - G(y))\, dF(y) \\ &= E[\delta_i] = \Pi, \text{ defined by (2.35),} \end{aligned} \tag{4.10}$$

$$\begin{aligned} E[\delta_i H_0(Y_i^0)] &= \int_0^\infty H_0(y)(1 - G(y))\, dF(y) = \int_0^\infty H_0(y)[1 - \Psi(y)]\, dH_0(y) \\ &= \frac{1}{2}\int_0^\infty [1 - \Psi(y)]\, dH_0^2(y) = \frac{1}{2}\int_0^\infty H_0^2(y)\, d\Psi(y) = \frac{1}{2}E[H_0^2(y)]. \end{aligned} \tag{4.11}$$

Thus, under H_0: $\boldsymbol{\beta} = \mathbf{0}$, $E(\delta_i - H_0(Y_i^0)) = 0$ and $E(\delta_i - H_0(Y_i^0))^2 = E\delta_i = \Pi$. Also, by the classical central limit theorem, under H_0, $n^{-1/2} \sum_{i=1}^n (\boldsymbol{\lambda}'\mathbf{Z}_i)(\delta_i - H_0(Y_i^0))$ is asymptotically normal with mean 0 and variance $\Pi E(\boldsymbol{\lambda}'\mathbf{Z}_1\mathbf{Z}_1'\boldsymbol{\lambda})$. Since by (4.8), $\nu^2 = \Pi E(\boldsymbol{\lambda}'\mathbf{Z}_1\mathbf{Z}_1'\boldsymbol{\lambda})$, from (4.9) and the above, we conclude that under H_0,

$$\log \mathscr{L}_n^0 \to_{\mathscr{D}} \mathscr{N}(-\tfrac{1}{2}\nu^2, \nu^2) \quad \text{as} \quad n \to \infty. \tag{4.12}$$

By (4.12) and the corollary to Le Cam's first lemma (cf Hájek and Šidák (1967, page 204)), we conclude that $\{P_n^*\}$ is contiguous to $\{P_n\}$. □

Let us now return to the proof of Theorem 4.1. By the same arguments as in the proof of Theorem 2 of Sen (1976), we conclude that the tightness of $\{\xi_n^*\}$, under H_0, insured by Theorem 2.8, and the contiguity of $\{P_n^*\}$ with respect to $\{P_n\}$, insured by Theorem 4.2, imply that $\{\xi_n^*\}$ remains tight under $\{K_n\}$ as well. Hence, to prove Theorem 4.1, it suffices to show that under $\{K_n\}$, the finite dimensional distributions (f.d.d.) of $\{\xi_n^*\}$ converge to those of ξ

$+ \zeta^*$. By using (4.4), we obtain by some standard steps that the conditional density of $\mathbf{Z}_i$ given $Y_i^0 = y$ and $\delta_i = \delta$, when K_n holds, is given by

$$l(\mathbf{z})[1 + n^{-1/2}(\delta - H_0(y))(\mathbf{z} - \boldsymbol{\mu})'\boldsymbol{\lambda} + o(n^{-1/2})], \tag{4.13}$$

where $l(\mathbf{z})$ is the marginal density of $\mathbf{Z}_i$. Thus, under K_n,

$$E(\mathbf{Z}_i \mid Y_i^0 = y, \delta_i = \delta) = \boldsymbol{\mu} + n^{-1/2}(\delta - H_0(y))\boldsymbol{\Gamma\lambda} + o(n^{-1/2}). \tag{4.14}$$

Now, for every $\alpha \in [0, 1]$, let us define

$$\mathbf{V}_n(\alpha) = n^{-1/2} \textstyle\sum_{i\le[n\alpha]} I(\delta_{Q_i^0} = 1)(n - i + 1)^{-1}(n - i)[\mathbf{Z}_{Q_i^0} - (n - i)^{-1}\sum_{j>i} \mathbf{Z}_{Q_j^0}], \tag{4.15}$$

where the Q_j^0 are defined after (1.4). Then, by (4.14) and (4.15), we have

$$\begin{aligned} E(\mathbf{V}_n(\alpha) \mid K_n) &= n^{-1}\textstyle\sum_{i\le[n\alpha]}(n - i + 1)^{-1}(n - i)\boldsymbol{\Gamma\lambda}[E\{I(\delta_{Q_i^0} = 1)(\delta_{Q_i^0} - H_0(Y_{ni}^0))\} \\ &\quad - (n - i)^{-1}\textstyle\sum_{j>i}E\{I(\delta_{Q_i^0} = 1)(\delta_{Q_j^0} - H_0(Y_{nj}^0))\}], \end{aligned} \tag{4.16}$$

where the Y_{ni}^0 are defined after (1.4). Note that for each $i(= 1, \cdots, n)$,

$$\begin{aligned} &E\{I(\delta_{Q_i^0} = 1)(\delta_{Q_i^0} - H_0(Y_{ni}^0))\} \\ &\quad = n\tbinom{n-1}{i-1} \int_0^\infty [1 - H_0(y)]\pi(y)[\Psi(y)]^{i-1}[1 - \Psi(y)]^{n-i}\, d\Psi(y), \end{aligned} \tag{4.17}$$

and hence, using the moment-convergence (of continuous functions) of sample quantiles (cf Sen (1959)), we claim that (4.17) is convergent equivalent to

$$\pi(\Psi^{-1}(i/(n + 1)))[1 - H_0(\Psi^{-1}(i/(n + 1)))]. \tag{4.18}$$

In a similar manner, it can be shown that $(n - i)^{-1} \sum_{j>i} E\{1(\delta_{Q_i^0} = 1)(\delta_{Q_j^0} - H_0(Y_{nj}^0))\}$ is convergent equivalent to

$$\begin{aligned} &\pi\left(\Psi^{-1}\left(\frac{i}{n+1}\right)\right)\left(1 - \frac{i}{n+1}\right)^{-1} \int_{\Psi^{-1}(i/(n+1))}^\infty \\ &\qquad \{[1 - H_0(y)]f_0(y)(1 - G(y)) - H_0(y)(1 - F_0(y))g(y)\}\, dy \\ &= \pi\left(\Psi^{-1}\left(\frac{i}{n+1}\right)\right)\left(1 - \frac{i}{n+1}\right)^{-1} \int_{\Psi^{-1}(i/(n+1))}^\infty \\ &\qquad \{(1 - G(y))\, dF_0(y) - H_0(y)\, d\Psi(y)\} \\ &= \pi(\Psi^{-1}\left(\frac{i}{n+1}\right)\Big)\left(1 - \frac{i}{n+1}\right)^{-1}\left\{\int_{\Psi^{-1}(i/n+1)}^\infty\right. \\ &\qquad \left.[1 - \Psi(y)]\, dH_0(y) - \int_{\Psi^{-1}(i/n+1)}^\infty H_0(y)\, d\Psi(y)\right\} \\ &= -\pi\left(\Psi^{-1}\left(\frac{i}{n+1}\right)\right)H_0(\Psi^{-1}\left(\frac{i}{n+1}\right)\Big). \end{aligned} \tag{4.19}$$

Hence, (4.16) is convergent equivalent to

$$\boldsymbol{\Gamma\lambda} \int_0^{\Psi^{-1}(\alpha)} \pi(z)\, d\Psi(z) = \Pi(\alpha)\boldsymbol{\Gamma\lambda}. \tag{4.20}$$

Now, by (1.10) and (4.15), $n^{-1/2}\mathbf{U}_{nm}^* = \mathbf{V}_n(1)$. Let $\omega_k = \inf\{r: \sum_{i=1}^r I(\delta_{Q_i^0} = 1) = k\}$, for $k = 1, \cdots, m$. Then, $n^{-1/2}\mathbf{U}_{nk}^* = \mathbf{V}_n(n^{-1}\omega_k)$, $k = 1, \cdots, m$. Also, adapting the proof of the Glivenko-Cantelli lemma and using the fact that $\Pi(z)$ is strictly monotonic in $z(\ge 0)$, it follows that $\max_{1\le k\le m} | n^{-1}\omega_k - \Pi^{-1}(k/(n + 1))| \to 0$, in probability, as $n \to \infty$. Similarly, proceeding as in the proof of Lemma 2.4, we have $n^{-1}\mathbf{J}^*_{nm} \to_p \Pi(1)\mathbf{\Gamma}$. Hence, using the tightness property of ξ^*_n [under H_0], insured by Theorem 2.8, we obtain that for any (fixed) $q(\geq 1)$ and $0 \leq t_1 < \cdots < t_q \leq 1$, under H_0, $\max_{1\leq j\leq q} | \xi^*_n(t_j) - \Pi^{-1/2}\mathbf{\Gamma}^{-1/2}\mathbf{V}_n(t^*_j) | \to_p 0$, where $\Pi(t^*_j) = t_j\Pi(1)$, for $j = 1, \cdots, q$. Because of the contiguity of $\{P^*_n\}$ with respect to $\{P_n\}$, established in Theorem 4.2, from the above, we obtain that

$$\max_{1\leq j\leq q} | \xi^*_n(t_j) - \Pi^{-1/2}\mathbf{\Gamma}^{-1/2}\mathbf{V}_n(t^*_j) | \to_p 0 \qquad \text{under } \{K_n\} \text{ as well.} \tag{4.21}$$

Thus, it suffices to show that under $\{K_n\}$, $\{\Pi^{-1/2}\mathbf{\Gamma}^{-1/2}\mathbf{V}_n(t^*_j), j = 1, \cdots, q\}$ has asymptotically a multinormal df with mean vector $\{\zeta^*(t_1), \cdots, \zeta^*(t_q)\}$ and dispersion matrix $((t_j \wedge t_{j'})) \otimes \mathbf{I}_p$, where $\otimes$ stands for the Kronecker product of the matrices. Now, (4.16) and (4.20) insure that under $\{K_n\}$, the asymptotic mean of $\Pi^{-1/2}\mathbf{\Gamma}^{-1/2}\mathbf{V}_n(t^*_j)$ is $\zeta^*(t_j)$, for $j = 1, \cdots, q$. Also, under H_0 and (1.1), $\mathbf{Z}_i$ is independent of (Y^0_i, δ_i), so that the conditional distribution of $\mathbf{Z}_i$ given (Y^0_i, δ_i), does not depend on Y^0_i and δ_i. Thus, using Lemma 1 of Bhattacharya (1974), we claim that under H_0 and given $\delta_1, \cdots, \delta_n$, $\mathbf{V}_n(t^*_j)$ involve a linear combination of i.i.d. rv's and hence a version of a theorem of Behnen and Neuhaus (1975) and our Theorem 4.2 insure that under $\{K_n\}$, the asymptotic multinormality of $\{\Pi^{-1/2}\mathbf{\Gamma}^{-1/2}\mathbf{V}_n(t^*_j), j = 1, \cdots, q\}$ holds. This concludes the proof of Theorem 4.1. The Corollary 4.1 follows directly by letting $\pi(z) = 1$ for all $z \in [0, \infty)$.

We conclude this section with the remark that similar weak convergence results hold for the discrete time case treated in Section 3.

5. Asymptotic relative efficiency results and some RST procedures.

Note that by virtue of (1.8), Theorem 2.8 and Theorem 4.1, $\mathscr{L}^*_{nm}$ in (1.8) has asymptotically (under H_0: $\boldsymbol{\beta} = \mathbf{0}$) chi-square distribution with p DF, and under $\{K_n\}$ in (4.1), it has asymptotically a noncentral chi-square distribution with p DF and noncentrality parameter

$$\Delta^* = \Pi(\boldsymbol{\lambda}'\mathbf{\Gamma}\boldsymbol{\lambda}), \ \Pi = \Pi(1) \text{ being defined by (2.35).} \tag{5.1}$$

In the uncensored case where the probability of withdrawal is 0 (i.e., the W_i are equal to $+\infty$ with probability 1), the parallel statistic is $\mathscr{L}_{nn} = \mathbf{U}'_{nn}\mathbf{J}^{-}_{nn}\mathbf{U}_{nn}$ and under $\{K_n\}$, it has asymptotically noncentral chi-square distribution with p DF and noncentrality parameter $\Delta = (\boldsymbol{\lambda}'\mathbf{\Gamma}\boldsymbol{\lambda})$; naturally, under H_0, the asymptotic df is central chi-square with p DF. Thus, the asymptotic relative efficiency (A.R.E.) of $\mathscr{L}^*_{nm}$ test with respect to the test based on $\mathscr{L}_{nn}$ is

$$e^* = e(\mathscr{L}^*, \mathscr{L}) = \Delta^*/\Delta = \Pi = \int_0^\infty [1 - G(x)]\, dF_0(x). \tag{5.2}$$

This indicates that whenever Π is close to 1, the A.R.E. of the censored case is also so, that is, there is not much loss in information due to censoring—this result has been obtained for the discrete time model by Efron(1977) from a somewhat different consideration.

Suppose now that we may not want to continue experimentation until all the failures have occurred, but desire to make a terminal test based on the partial set $(Y^0_{ni}, \delta_{Q^0_i}, \mathbf{Z}_{Q^0_i}$, $i = 1, \cdots, r$ where $r = [n\alpha]$ and α is some (fixed) number $(0 < \alpha < 1)$. In such a case, defining $m_\alpha = \sum_{j=1}^{[n\alpha]} I(\delta_{Q^0_j} = 1)$, we may construct a test statistic $\mathscr{L}^*_{nm_\alpha} = \mathbf{U}^{*\prime}_{nm_\alpha}\mathbf{J}^{*-1}_{nm_\alpha}\mathbf{U}^*_{nm_\alpha}$. Under H_0: $\boldsymbol{\beta} = \mathbf{0}$, $\mathscr{L}^*_{nm_\alpha}$ has asymptotically chi-square distribution with p DF, while under $\{K_n\}$ in (4.1), it has asymptotically noncentral chi-square distribution with p DF and noncentrality parameter $\Pi(\alpha)(\boldsymbol{\lambda}'\mathbf{\Gamma}\boldsymbol{\lambda})$. Thus, the A.R.E. of this censored test with respect to the ideal test based on $\mathscr{L}_{nn}$ (if the Y_i were observable) is given by

$$e^*_\alpha = e(\mathscr{L}^*_{nm_\alpha}, \mathscr{L}_{nn}) = \Pi(\alpha) = [\Pi(\alpha)/\alpha]\alpha, \tag{5.3}$$

where the first factor is bounded above by 1 and the second factor accounts for the intrinsic A.R.E. for censoring the experiment at the $[n\alpha]$th order statistic. $\Pi(\alpha)/\alpha$ represents the A.R.E. of $\mathscr{L}^*_{nm_\alpha}$ with respect to $\mathscr{L}_{n[n\alpha]}$ and reflects the loss due the incorporation of withdrawals in the scheme, when experimentation is curtailed at the $[n\alpha]$th order statistic Y^0_{ni}.

A common feature of these tests is that they demand the experimentation be continued until all the m (or m_α) failures occur. In the RST procedures, we like to update the picture as successive failures occur and thereby consider some *time-sequential procedures*. These procedures are based on the partial sequence $\{\mathscr{L}^*_{nk},\ k \le m\}$ of partial likelihood ratio statistics. Among other possibilities, we consider the following test statistic:

$$T^*_n = \max\{\mathscr{L}^*_{nk}: [n\epsilon] \le k \le m\} \tag{5.4}$$

where m is equal to $\sum_{j=1}^n I(\delta_j = 1)$ and ϵ $(0 < \epsilon < 1)$ is some prefixed (small) positive number. If we denote by τ^*_α, the upper $100\alpha\%$ point of the null (under $\boldsymbol{\beta} = \mathbf{0}$) distribution of T^*_n, then operationally the RST procedure consists in computing $\mathscr{L}^*_{nk}$ at each failure and curtailing experimentation along with the rejection of the null hypothesis as soon as for some k:$[n\epsilon] \le k \le m$, $\mathscr{L}^*_{nk}$ exceeds τ^*_α; if no such k exists, then experimentation is continued till the end and H_0 is accepted. If we define the vector valued Wiener process ξ as in Theorem 2.1 and let

$$\xi^*(\rho) = \sup\{t^{-1}[\xi(\mathbf{t})]'[\xi(\mathbf{t})]: \rho \le t \le 1] \qquad \forall\ \rho \in (0, 1), \tag{5.5}$$

then, by virtue of (2.35) and Theorem 2.8, we have under H_0: $\boldsymbol{\beta} = \mathbf{0}$,

$$T^*_n \to_{\mathscr{D}} \xi^*(\Pi^{-1}(\epsilon)) \qquad \text{where} \quad \Pi^{-1}(\epsilon) \ge \epsilon \qquad \forall\ \epsilon > 0. \tag{5.6}$$

Though the analytical form of the df of $\xi^*(\epsilon)$ is quite complicated, for various typical values of ϵ, the critical values of $\xi^*(\epsilon)$ have been studied by Majumdar (1978). Thus, if $\xi^*_\alpha(\epsilon)$ be the upper $100\alpha\%$ point of the distribution of $\xi^*(\epsilon)$, then,

$$\lim_{n\to\infty}\tau^*_{\alpha n} = \tau^*_\alpha \le \xi^*_\alpha(\epsilon) \qquad \text{for every} \quad 0 < \alpha < 1 \quad \text{and} \quad \epsilon > 0. \tag{5.7}$$

Hence, a (somewhat conservative) large sample RST can be constructed using $\xi^*_\alpha(\epsilon)$ as its critical value. Theorem 4.1 can be used to provide an expression for the asymptotic power function of this RST procedure when $\{K_n\}$ in (4.1) holds. The usual definition of the Pitman-A.R.E. is not applicable to compare the RST with the earlier ones and extensive simulation studies are needed to achieve this goal. These will be considered in a subsequent issue. We conclude this section with the remark that for the RST procedure, one need not wait (for accepting the null hypothesis) until the end of the experiment. We may prefix a positive number $\gamma(0 < \gamma < 1)$ and in (5.4), limit the range of k to $[n\epsilon] \le k \le m_\gamma$, where m_γ is defined as in before (5.3). In such a case, in (5.5), we need to limit the range of t to $(\rho < t < \gamma)$. In (5.7), we need to take $\gamma\xi^*_\alpha(\epsilon)$ for the limiting critical point and the rest of the procedure remains the same. Further, in (5.4), we have chosen $\epsilon > 0$. For t close to 0, $t^{-1}[\xi(t)]'[\xi(t)]$ does not behave very regularly (in fact, it blows up, with probability 1) and also the weak convergence of $t^{-1}[\xi^*_n(t)]'[\xi^*_n(t)]$ for t near 0 may not hold, as for $q(t) = t$, (2.33) does not hold. However, exclusion of a small neighbourhood at the origin eliminates this problem and enables us to use the invariance principles studied earlier for approximating the critical values by those of the process derived from ξ. Finally, we have considered the case where the $\mathbf{Z}_i$ are stochastic vectors. For nonstochastic $\mathbf{Z}_i$, a somewhat different approach formulated in Sen (1979, 1981) works out well.

6. Acknowledgments. The author is grateful to the referees for their critical readings and most helpful comments on the manuscript.

REFERENCES

BEHNEN, K. and NEUHAUS, G. (1975). A central limit theorem under contiguous alternatives. *Ann. Statist.* **3** 1349–1353.

BERK, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

BHATTACHARYA, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. *Ann. Statist.* **2** 1034–1039.

BIRNBAUM, Z. W. and MARSHALL, A. W. (1961). Some multivariate Chebyshev inequalities with extensions to continuous parameter processes. *Ann. Math. Statist.* **32** 687–703.

CHATTERJEE, S. K. and SEN, P. K. (1973). Nonparametric testing under progressive censoring. *Calcutta Statist. Asso. Bull.* **22** 13–50.

COX, D. R. (1972). Regression models and life tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

COX, D. R. (1975). Partial likelihoods. *Biometrika* **62** 269–276.

EFRON, B. (1977). The efficiency of Cox's likelihood function for censored data. *J. Amer. Statist. Asso.* **72** 557–565.

HÁJEK, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann. Math. Statist.* **32** 506–523.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic, New York.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

MAJUMDAR, H. (1978). Rank order tests for multiple regression for grouped data under progressive censoring. *Calcutta Statist. Assoc. Bull.* **26** 1–16.

MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probability* **2** 620–628.

SCOTT, D. J. (1973). Central limit theorems for martingales and for processes with stationary increments using a skorokhod representation approach. *Advances Appl. Probability* **5** 119–137.

SEN, P. K. (1959). On the moments of sample quantiles. *Calcutta Statist. Assoc. Bull.* **9** 1–19.

SEN, P. K. (1976). Weak convergence of progressively censored likelihood ratio statistics and its role in asymptotic theory of life testing. *Ann. Statist.* **4** 1247–1257.

SEN, P. K. (1979). Weak convergence of some quantile processes arising in progressively censored tests. *Ann. Statist.* **7** 414–431.

SEN, P. K. (1981). Asymptotic theory of some time-sequential tests based on progressively censored quantile processes. In *Statistics and Probability: Essays in Honor of C. R. Rao.* (G. Kallianpur et al. ed.) North Holland Pub., Amsterdam (in press).

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF NORTH CAROLINA
ROSENAN HALL 201 H
CHAPEL HILL, NORTH CAROLINA 27514