THE COX REGRESSION MODEL, RANDOM CENSORING AND LOCALLY OPTIMAL RANK TESTS

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1376

February 1982

THE COX REGRESSION MODEL, RANDOM CENSORING AND LOCALLY OPTIMAL RANK TESTS

by

Pranab Kumar Sen

University of North Carolina, Chapel Hill

Key Words and Phrases: Asymptotic efficiency; covariates; hazard function; LMPR test; local optimality; log-rank test; survival function; withdrawal.

ABSTRACT

Conditions on the hazard functions under which the usual log-rank test remains locally optimal for the Cox regression model under random censoring (withdrawal) are examined. In the light of these, the asymptotic efficiency results pertaining to the Cox partial likelihood statistic and the log-rank statistic are studied.

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* $\underset{\sim}{Z}_i = (Z_{i1},\ldots,Z_{ip})'$, for some $p \geq 1$) has the *hazard rate* (given $\underset{\sim}{Z}_i = \underset{\sim}{z}_i$)

$$h_i(t) = h_0(t) \cdot \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i), \quad i=1,\ldots,n, \qquad (1.1)$$

where $h_0(t)$, the hazrad rate for $\underset{\sim}{z}_i = 0$, is an unknown, arbitrary, nonnegative function (for which $\int_0^\infty h_0(t)dt = +\infty$) and $\underset{\sim}{\beta} = (\beta_1,\ldots,\beta_p)'$ parameterizes the regression of survival times on the covariates. If the conditional *distribution function* (d.f.) and, its complement, the *survival function* (s.f.) of Y_i, given $\underset{\sim}{Z}_i = \underset{\sim}{z}_i$, are denoted by $F_i(y|\underset{\sim}{z}_i)$

and $\bar{F}_i(y|\underset{\sim}{z}_i)$, respectively, then by (1.1), we have

$$\bar{F}_i(y|\underset{\sim}{z}_i) = [\bar{F}_0(y)]^{\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)}, \quad i=1,\ldots,n , \qquad (1.2)$$

where

$$h_0(y) = -(d/dy)\log \bar{F}_0(y) . \qquad (1.3)$$

In the particular case of $p = 1$ and binary z_i (assuming the values 0 or 1), (1.2) reduces to the Lehmann (1953) model, so that the Cox model includes the Lehmann model as a special case. Also, for scalar z_i and β, it follows from a general theorem in Hájek and Sidák (1967, pp. 70-72) that for testing H_0: $\beta = 0$ against non-null β close to 0, a *locally most powerful rank* (LMPR) test is based on the statistic

$$T_n^0 = \Sigma_{i=1}^n (Z_i - \bar{Z}_n) a_n^0(R_{ni}), \quad \bar{Z}_n = n^{-1}\Sigma_{i=1}^n Z_i , \qquad (1.4)$$

where R_{ni} is the *rank* of Y_i among $Y_1,\ldots,Y_n$, for $i=1,\ldots,n$, and the *scores* $a_n^0(1),\ldots,a_n^0(n)$ are defined by

$$a_n^0(k) = -1 - E\{\log(1-U_{n,k})\}$$
$$(= -1 + \Sigma_{j=1}^k (n-j+1)^{-1}) , \quad \text{for } k=1,\ldots,n , \qquad (1.5)$$

where $U_{n,1} <\ldots< U_{n,n}$ are the random variables of a sample of size n from the uniform (0,1) distribution. The scores in (1.5) are known as the *log-rank* scores and T_n^0 as the *log-rank statistic* or the (generalized) *Savage statistic*. When $p \geq 1$, we may define

$$\underset{\sim}{T}_n^0 = \Sigma_{i=1}^n (\underset{\sim}{Z}_i - \underset{\sim}{\bar{Z}}_n) a_n^0(R_{ni}), \quad \underset{\sim}{\bar{Z}}_n = n^{-1}\Sigma_{i=1}^n \underset{\sim}{Z}_i , \qquad (1.6)$$

$$\underset{\sim}{V}_n = \Sigma_{i=1}^n (\underset{\sim}{Z}_i - \underset{\sim}{\bar{Z}}_n)(\underset{\sim}{Z}_i - \underset{\sim}{\bar{Z}}_n)' \qquad (1.7)$$

and

$$\mathcal{L}_n^0 = (\underset{\sim}{T}_n^0)'(\underset{\sim}{V}_n)^-(\underset{\sim}{T}_n^0) . \qquad (1.8)$$

Then L_n^0 provides a *locally maximin rank test* for H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$ against non-null $\underset{\sim}{\beta}$ close to $\underset{\sim}{0}$.

In survival analysis, random censoring due to withdrawals is not uncommon. Here, we conceive of a set $W_1,\ldots,W_n$ of independent random variables (censoring times) and assume that the W_i and Y_i are independent. The ovservable r.v.'s are then

$$X_i = \min(W_i,Y_i) \text{ and } \delta_i = \begin{cases} 1, & X_i = Y_i , \\ 0, & X_i \neq Y_i ; \quad i \geq 1 . \end{cases} \quad (1.9)$$

If G and $\bar{G}$ be respectively the d.f. and s.f. for the W_i (assumed to be identically distributed) and if $S_i(x|\underset{\sim}{z}_i)$ and $\bar{S}_i(x|\underset{\sim}{z}_i)$ be respectively the d.f. and s.f. of X_i, given $\underset{\sim}{Z}_i = \underset{\sim}{z}_i$ (ignoring δ_i), then by (1.2) and (1.9),

$$\begin{aligned} \bar{S}_i(x|\underset{\sim}{z}_i) &= \bar{G}(x)\bar{F}_i(x|\underset{\sim}{z}_i) \\ &= \bar{G}(x)[\bar{F}_0(x)]^{\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)}, \; i=1,\ldots,n. \quad (1.10) \end{aligned}$$

Note that under H_0: $\underset{\sim}{\beta} = 0$, $\bar{S}_1,\ldots,\bar{S}_n$ are all the same (i.e., equal to $\bar{G}\,\bar{F}_0$), so that if one ignores the information contained in $\underset{\sim}{\delta}_n = (\delta_1,\ldots,\delta_n)'$, rank tests based on $X_1,\ldots,X_n$ are genuinely distribution-free (under H_0). The test based on L_n^0 or $\underset{\sim}{T}_n^0$ (replacing the Y_i by X_i, $1 \leq i \leq n$) is therefore a valid test for H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$ under random censoring, though it may not be locally optimal.

If g is the density function corresponding to the d.f. G, then the hazard rate for the W_i is $h_G(t) = g(t)/\bar{G}(t)$, so that the hazard rate for the X_i (conditional on $\underset{\sim}{Z}_i = \underset{\sim}{z}_i$) are

$$\begin{aligned} h_i^*(t) &= -(d/dt)\log \bar{S}_i(t|\underset{\sim}{z}_i) = h_G(t) + h_i(t) \\ &= h_G(t) + h_0(t)\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i), \quad 1 \leq i \leq n . \quad (1.11) \end{aligned}$$

This, in general, vitiates the proportionality assumption in (1.1), and hence, the log-rank test may not be locally optimal.

Note that the likelihood function (of the X_i and δ_i, given $\underset{\sim}{Z}_i$) is given by

$$\prod_{i=1}^{n}\{[\overline{F}_0(x_i)]^{\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)}\overline{G}(x_i)([h_0(x_i)]\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i))^{\delta_i}(h_G(x_i))^{1-\delta_i}\}\ , \tag{1.12}$$

so that the joint distribution of the ranks and $\underset{\sim}{\delta}_n$ depends on the unknown $\overline{F}_0$ and $\overline{G}$, even under H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$. To eliminate this problem, Cox (1972, 1975) considered a partial likelihood function (which takes into account the indicator variables $\delta_1,\ldots,\delta_n$) and obtained some (non-rank) test statistics which depend on the covariates $\underset{\sim}{Z}_i$, $i=1,\ldots,n$ in a more involved way. Peto and Peto (1972) considered the two-sample problem [as a special case of (1.10) with binary z_i] and, under a somewhat different setup, concluded that the log-rank statistic is LMPR even under random censoring. As we shall see in Section 2 that under the model (1.10)-(1.11), the log-rank statistic is not generally LMPR (or maximin), even for the special case of the two-sample problem. For this reasom, we will investigate the conditions (on the hazard rates $h_0(t)$ and $h_G(t)$ or equivalently on $\overline{F}_0$ and $\overline{G}$) under which the log-rank statistic is locally optimal among the rank based tests on $X_1,\ldots,X_n$ (gnoring $\underset{\sim}{\delta}_n$). The question that naturally arises about the gain in efficiency of the Cox procedure (due to incorporation of $\underset{\sim}{\delta}_n$) over the log-rank procedure (ignoring $\underset{\sim}{\delta}_n$), and this will be addressed here too .

Section 2 is devoted to the study of locally optimal rank tests (for testing H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$), under the model (1.10). Section 3 deals with the local optimality of the log-rank test. Asymptotic efficiency results on the Cox procedure are presented in Section 4, and the concluding section is devoted to some general remarks.

2. LOCALLY OPTIMAL RANK PROCEDURES UNDER RANDOM CENSORING

Ignoring $\delta_1,\ldots,\delta_n$ and based on the ranks of $X_1,\ldots,X_n$, we like to construct suitable tests for H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$, having some local optimality properties.

Let $s_i(x|\underset{\sim}{z}_i)$ be the probability density function (p.d.f.) corresponding to the d.f. $S_i(x|\underset{\sim}{z}_i)$ in (1.10), for i=1,...,n. Then, by (1.10) and (1.11), we have

$$\begin{aligned}
\log S_i(x|\underset{\sim}{z}_i) &= \log \bar{S}_i(x|\underset{\sim}{z}_i) + \log h_i^*(x) \\
&= \log \bar{G}(x) + \log \bar{F}_i(x|\underset{\sim}{z}_i) + \log h_i^*(x) \\
&= \log \bar{G}(x) + \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)\log \bar{F}_0(x) + \\
&\qquad \log[h_G(x) + \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)h_0(x)],\ 1 \le i \le n \qquad (2.1)
\end{aligned}$$

which leads to the log-likelihood function (of the X_i given $\underset{\sim}{Z}_i = \underset{\sim}{z}_i$, i=1,...,n), ignoring $\underset{\sim}{\delta}_n$, as

$$\begin{aligned}
\log L_n^*(\underset{\sim}{\beta}) &= \Sigma_{i=1}^n \log S_i(X_i|\underset{\sim}{z}_i) \\
&= \Sigma_{i=1}^n \log \bar{G}(X_i) + \Sigma_{i=1}^n \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)\log \bar{F}_0(X_i) + \\
&\qquad \Sigma_{i=1}^n \log[h_G(X_i) + \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)h_0(X_i)] \ . \qquad (2.2)
\end{aligned}$$

Note that by (2.2),

$$(\partial/\partial\underset{\sim}{\beta})\log L_n^*(\underset{\sim}{\beta})\Big|_{\underset{\sim}{\beta}=0} = \sum_{i=1}^n \underset{\sim}{Z}_i\{\log \bar{F}_0(X_i) + \pi(X_i)\} \qquad (2.3)$$

where

$$\pi(x) = h_0(x)/\{h_G(x) + h_0(x)\},\ x \in E \ . \qquad (2.4)$$

Further note that under H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$, $\bar{S}_1(x|\underset{\sim}{z}_1) = \ldots = \bar{S}_n(x|\underset{\sim}{z}_n) = \bar{S}_0(x) = \bar{G}(x)\bar{F}_0(x)$ and this does not depend on $\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_n$. Thus, under H_0, $X_1,\ldots,X_n$ are independent and identically distributed random variables (i.i.d.r.v.), so that $\underset{\sim}{R}_n = (R_{n1},\ldots,R_{nn})$, the vector of ranks of $X_1,\ldots,X_n$ among themselves, has the (discrete) uniform distribution over the set of n! possible permutations of $(1,\ldots,n)$. Let then $\psi^* = \psi_1 + \psi_2$, where

$$\psi_1(u) = \log \bar{F}_0(x)\Big|_{\bar{S}_0(x)=1-u}, \quad 0 < u < 1, \tag{2.5}$$

$$\psi_2(u) = \pi(x)\Big|_{\bar{S}_0(x)=1-u}, \quad 0 < u < 1, \tag{2.6}$$

define the ordered r.v.'s $U_{n,1},\ldots,U_{n,n}$ as in after (1.5), and let

$$a_{n,j}(k) = E\psi_j(U_{n,k}), \text{ for } 1 \leq k \leq n \text{ and } j=1,2 . \tag{2.7}$$

Also, let $a_n^*(k) = a_{n,1}(k) + a_{n,2}(k)$, for $k=1,\ldots,n$. Then, by (2.3), (2.4), (2.5), (2.6) and (2.7), we obtain that

$$E_0\{(\partial/\partial\underset{\sim}{\beta})\log L_n^*(\beta)\Big|_{\underset{\sim}{\beta}=\underset{\sim}{0}}|R_n\} = \sum_{i=1}^{n} \underset{\sim}{Z}_i a_n^*(R_{ni}) = \underset{\sim}{T}_n^*, \text{ say,} \tag{2.8}$$

where E_0 denotes the expectation under H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$.

Note that by (2.4) and (2.6),

$$\begin{aligned}
\bar{\psi}_2 &= \int_0^1 \psi_2(u)du = \int\pi(x)dS_0(x) = -\int \pi(x)d\bar{S}_0(x) \\
&= \int \bar{F}_0(x)\bar{G}(x)h_0(x)dx \\
&= -\int \bar{F}_0(x)\bar{G}(x)d \log \bar{F}_0(x) \qquad \text{[by (1.3)]} \\
&= \int\log \bar{F}_0(x)d\cdot\bar{S}_0(x) = -\int \log \bar{F}_0(x)dS_0(x) \\
&= -\int_0^1 \psi_1(u)du = -\bar{\psi}_1 .
\end{aligned} \tag{2.9}$$

Further, by (2.7)

$$\bar{a}_{n,j} = \frac{1}{n} \Sigma_{i=1}^n a_{n,j}(i) = \int_0^1 \psi_j(u)du = \bar{\psi}_j, \text{ for } j=1,2. \tag{2.10}$$

Hence, from (2.9) and (2.10), we have

$$\bar{a}_n^* = n^{-1} \Sigma_{i=1}^n a_n^*(i) = \bar{a}_{n,1} + \bar{a}_{n,2} = \bar{\psi}_1 + \bar{\psi}_2 = 0 \ . \quad (2.11)$$

Thus, by (2.8) and (2.11), we may rewrite $\underset{\sim}{T}_n^*$ as

$$\underset{\sim}{T}_n^* = \Sigma_{i=1}^n (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n) a_n^*(i) \ . \quad (2.12)$$

At this stage, we may note that (i) $h_0(x)\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i)/\{h_0(x)\exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i) + h_G(x)\}$ is a bounded and continuous function of $\underset{\sim}{\beta}$, and (ii) for every x, $0 \le -\log \bar{F}_0(x) = -\log \bar{S}_0(x) + \log \bar{G}(x) \le -\log \bar{S}_0(x)$, where the right hand side is square integrable (with respect to $S_0(x)$). Hence, for p = 1, we may appeal directly to a theorem in Hájek and Šidák (1967, p.71), verify their basic conditions (1) and (2) (on page 70) and conclude that T_n^* is a LMPR test statistic for testing H_0: $\beta = 0$ (based on the X_i and ignoring the δ_i). For $p \ge 1$, we let

$$L_n^* = (\underset{\sim}{T}_n^*)' \underset{\sim}{V}_n^- (\underset{\sim}{T}_n^*) \ , \quad (2.13)$$

where $\underset{\sim}{V}_n$ is defined by (1.7). Then, by an appeal to the Union-Intersection principle, as in Sen (1982), or the maximin theory as in Hájek and Šidák (1967), we claim that L_n^* is a locally maximin rank test for H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$, when $\underset{\sim}{\delta}_n$ is ignored. Let us define

$$A^{*2} = \int_0^1 \{\psi_1(u) + \psi_2(u)\}^2 du = \int_0^1 \psi^{*2}(u) du \ , \quad (2.14)$$

$$A_n^{*2} = (n-1)^{-1} \Sigma_{i=1}^n [a_n^*(i)]^2 \ . \quad (2.15)$$

Note that $A^* < \infty$ and $A_n^* \to A^*$ as $n \to \infty$. Thus, when the $\underset{\sim}{Z}_i$ satisfy the (generalized) Noether condition:

$$\max_{1 \le i \le n} (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n)' \underset{\sim}{V}_n^- (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n) \to 0 \text{ (a.s.), as } n \to \infty, \quad (2.16)$$

then, by an appeal to the permutational central limit theorem and the Cochran theorem, we conclude that under H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$,

$$L_n^*/A_n^{*2} \underset{\mathcal{D}}{\to} \chi_p^2 , \quad \text{as} \quad n \to \infty , \tag{2.17}$$

where χ_p^2 has the (central) chi square distribution with p degrees of freedom. In particualr, for p = 1, U_n is a scalar (nonnegative) quantity, and under H_0,

$$A_n^{*-1} V_n^{-\frac{1}{2}} T_n^* \sim N(0,1) . \tag{2.18}$$

We may note that by (2.5) and (2.6),

$$\begin{aligned} \int_0^1 \psi_1(u)\psi_2(u)du &= -\int \{\log \overline{F}_0(x)\}\pi(x)d\overline{S}_0(x) \\ &= -\int \{\log \overline{F}_0(x)\}h_0(x)\overline{F}_0(x)\overline{G}(x)dx \\ &= \tfrac{1}{2}\int \overline{F}_0(x)\overline{G}(x)d(-\log \overline{F}_0(x))^2 \\ &= -\tfrac{1}{2} \int(-\log \overline{F}_0(x))^2 dS_0(x) = -\tfrac{1}{2} \int_0^1 \psi_1^2(u)du . \end{aligned} \tag{2.19}$$

Hence, by (2.14) and (2.19)

$$\begin{aligned} A^{*2} &= \int_0^1 [\psi_1^2(u) + \psi_2^2(u) + 2\psi_1(u)\psi_2(u)]du \\ &= \int_0^1 \psi_2^2(u)du = \int \pi^2(x)dS_0(x) . \end{aligned} \tag{2.20}$$

Consider now a sequence $\{K_n\}$ of alternative hypotheses, where for each n,

$$K_n\text{: (1.10) holds with } \underset{\sim}{\beta} = n^{-\frac{1}{2}}\underset{\sim}{\lambda} , \tag{2.21}$$

for some (fixed) $\underset{\sim}{\lambda} \in E^b$, and assume that

$$n^{-1}\underset{\sim}{V}_n \to \underset{\sim}{\nu} \text{ (a.s.), as } n \to \infty , \tag{2.22}$$

where $\underset{\sim}{\nu}$ is posotive definite (p.d.). Then, in (2.2), writing

$\exp(\beta' z_i) = 1 + n^{-\frac{1}{2}}(\lambda' z_i) + \frac{1}{2}n^{-1}(\lambda' z_i)^2 + O(n^{-3/2})$, we obtain by some routine steps that

$$\log \{L_n^*(n^{-\frac{1}{2}}\lambda)/L_n^*(0)\} = n^{-\frac{1}{2}} \Sigma_{i=1}^n (\lambda' z_i)[\log \bar{F}_0(X_i) + \pi(X_i)]$$

$$+ \frac{1}{2n} \Sigma_{i=1}^n (\lambda' z_i)^2 [\log \bar{F}_0(X_i) + \pi(X_i)] -$$

$$\frac{1}{2n} \Sigma_{i=1}^n (\lambda' z_i)^2 \pi^2(X_i) + O_p(n^{-\frac{1}{2}}) , \tag{2.23}$$

where, under H_0: $\beta = 0$, the first term on the right hand side of (2.23) is asymptotically normal with 0 mean and variance A^{*2} $(\lambda'\nu\lambda)$ [$= \sigma^2$ (say)], the second term converges in probability to 0, while the third term to $\frac{1}{2}(\lambda'\nu\lambda)$ A^{*2} $(=\frac{1}{2}\sigma^2)$. Thus, the left hand side is asymptotically normal with mean $-\frac{1}{2}\sigma^2$ and variance σ^2. This, according to LeCam's first Lemma [viz., Hájek and Šidák (1967, p. 204)], establishes the contiguity of the sequence of probability measures under $\{K_n\}$ to that under H_0. As such, using LeCam's third lemma along with the usual projection of T_n^*, it follows that under $\{K_n\}$,

$$n^{-\frac{1}{2}} T_n^* \sim N(\nu\lambda A^{*2}, A^{*2}\nu) \quad . \tag{2.24}$$

Thus, under $\{K_n\}$, L_n^*/A_n^{*2} has asymptotically a noncentral chi-square distribution with p degrees of freedom (D.F.) and noncentrality parameter

$$\Delta^* = A^{*2}(\lambda'\nu\lambda) = (\int_0^1 \psi_2^2(u)du)(\lambda'\nu\lambda) \quad . \tag{2.25}$$

Note that, in practice, to use the statistic L_n^*, one needs to know the score function ψ^*, which [by(2.5)-(2.6)] depends on the unknown F_0 and G. Thus, in general, L_n^* is not an adaptable test statistic. Nevertheless, the above result provides a convenient means for studying the asymptotic efficiency of other tests, and this will be taken up in the next section.

3. LOCAL OPTIMALITY OF THE LOG-RANK TEST

In this section, we shall study the asymptotic efficiency and optimality properties of the log-rank test. By virtue of the contiguity results of Section 2 and the usual projection results on $\underset{\sim}{T}_n^0$, parallel to (2.24), we obtain that under $\{K_n\}$ and the regularity conditions of Section 2,

$$n^{-\frac{1}{2}}\underset{\sim}{T}_n^0 \sim N(\gamma^2 \underset{\sim}{\nu}\underset{\sim}{\lambda},\ A_\phi^2 \underset{\sim}{\nu}) \ , \tag{3.1}$$

where

$$A_\phi^2 = \int_0^1 \{-1 - \log(1-u)\}^2 du = 1 \ , \tag{3.2}$$

$$\begin{aligned} \gamma &= \int_0^1 \{-1 - \log(1-u)\}\{-\psi_1(u) - \psi_2(u)\}du \\ &= \int_0^1 \psi^*(u)\log(1-u)du \ . \end{aligned} \tag{3.3}$$

Thus, under $\{K_n\}$, L_n^0 has asymptotically a noncentral chi-square distribution with p D.F. and noncentrality parameter

$$\Delta^0 = \gamma^2(\underset{\sim}{\lambda}'\underset{\sim}{\nu}\underset{\sim}{\lambda}) \ . \tag{3.4}$$

By (2.25) and (3.4), the Pitman-efficiency of the log-rank test with respect to the locally optimal one is

$$e(L^0, L^*) = \Delta^0/\Delta^* = \gamma^2/A^{*2} \ . \tag{3.5}$$

If we write $\phi^*(u) = \log(1-u)$, $0 < u < 1$, then by (2.14), (2.20), (3.2) and (3.3), we have

$$\begin{aligned} \gamma^2/A^{*2} &= (\int_0^1 \phi^*(u)\psi^*(u)du)^2/\{(\int_0^1 \psi^{*2}(u)du)A_\phi^2\} \\ &= \frac{(\int_0^1 (\phi^*(u) - \bar{\phi}^*)\psi^*(u)du)^2}{(\int_0^1 \psi^{*2}(u)du)(\int_0^1 [\phi^*(u) - \bar{\phi}^*]^2 du)} \\ &= \rho^2(\phi^*, \psi^*) \ , \end{aligned} \tag{3.6}$$

where

$$\rho^2(\phi^*,\psi^*) \leq 1, \text{ with the strict equality sign holding only for } \phi^*(u) = a\psi^*(u) + b,\ 0 < u < 1, \text{ for some real } a\ (\neq 0) \text{ and } b. \tag{3.7}$$

By (2.5), (2.6) and (3.7), we conclude that $\rho^2(\phi^*,\psi^*) = 1$ only when, for some real k_1, k_2 ($\neq 0$),

$$\log \overline{S}_0(x) = k_1 + k_2 \log \overline{F}_0(x) + k_2 h_0(x)/\{h_0(x) + h_G(x)\}\ , \tag{3.8}$$

for almost all x, and since, $\log \overline{S}_0 = \log \overline{F}_0 + \log \overline{G}$, (3.8) may be written equivalently as

$$\log \overline{G}(x) = k_1 + (k_2-1) \log \overline{F}_0(x) + k_2 h_0(x)/\{h_0(x) + h_G(x)\}\ , \tag{3.9}$$

for almost all x. Since $h_0/\{h_0+h_G\}$ is nonnegative and bounded between 0 and 1, and $\log \overline{G} \to -\infty$ as $x \to +\infty$, in (3.9), k_2 has to be different from 1. (3.9) specifies the interrelationship of F_0 and G for which the log-rank test is a locally optimal rank test under the model (1.10).

An important class of distributions for which (3.9) holds may be characterized by the two hazard functions $h_0(x)$ and $h_G(x)$ as follows. Suppose that

$$h_G(x) \equiv c\, h_0(x), \text{ for some } c \neq 0\ . \tag{3.10}$$

Then, by integration on both sides, we have

$$\log \overline{G}(x) \equiv c \log \overline{F}_0(x) + c',\ c' \text{ real}\ , \tag{3.11}$$

while by (3.10), $h_0/\{h_0+h_G\} = (1+c)^{-1}$. Hence, it is easy to verify that (3.9) holds with $k_2 = c+1$. Thus, *if the hazard rates for the d.f.*

F_0 *and* G *are proportional to each other, then the log-rank test is a locally optimal rank test for the Cox model under random censoring.*

A second situation where (3.9) holds is the degenerate case where $\overline{G}(x) = 1, \forall x < \infty$, so that $\overline{S}_0(x) = \overline{F}_0(x), \forall x < \infty$, $h_G(x) = 0, \forall x < \infty$, and hence, (3.8) holds with $k_2=1$. Thus, if the withdrawal distribution lies completely to the right of the d.f. F_0, then the log-rank test is locally optimal -- as it is in the case where there is no (random) censoring.

In passing, we may remark that Peto and Peto (1972) considered the two-sample problem, where for some n_1 $(= N-n_2)$, $\overline{S}_1 = \ldots = \overline{S}_{n_1} = \overline{S}_0$ and $\overline{S}_{n_1+1} = \ldots = \overline{S}_N = [\overline{S}_0]^{1+\lambda}$ and showed that the locally most powerful rank test (for $\lambda = 0$, vs. $\lambda \neq 0$) is the log-rank test. Their model differs from (1.10) [in the sense that $\overline{G}$ does not remain the same under alternatives], and hance, their conclusions may not hold for the model (1.10) even for the special case of the two-sample problem.

So far, we have considered the local optimality and efficiency of the log-rank test relative to L_n^*, where the information contained in $\underset{\sim}{\delta}_n$ has not been incorporated in the testing procedures. We like to study the loss of efficiency due to this. For this, we may note that the joint density of the (X_i, δ_i) in (1.9) is given by

$$L_n(\underset{\sim}{\beta}) = \prod_{i=1}^{n} \{[f_i(X_i|\underset{\sim}{z}_i)\overline{G}(X_i)]^{\delta_i}[g(X_i)\overline{F}_i(X_i|\underset{\sim}{z}_i)]^{1-\delta_i}\}$$

$$= \prod_{i=1}^{n} \{\overline{F}_i(X_i|\underset{\sim}{z}_i)\overline{G}(x_i)[h_i(X_i)]^{\delta_i}[h_G(X_i)]^{1-\delta_i}\}, \quad (3.12)$$

so that by (1.1), (1.2) and (3.12), we have

$$\log L_n(\underset{\sim}{\beta}) = \Sigma_{i=1}^{n}\{\log \overline{G}(X_i) + \exp(\underset{\sim}{\beta}'\underset{\sim}{z}_i) \log \overline{F}_0(X_i) + \delta_i\underset{\sim}{\beta}'\underset{\sim}{z}_i + \delta_i \log h_0(X_i) + (1-\delta_i) \log h_G(X_i)\}. \quad (3.13)$$

Thus, under $\{K_n\}$ in (2.21),

$$\log\{L_n(n^{-\frac{1}{2}}\lambda)/L_n(0)\} = n^{-\frac{1}{2}} \Sigma_{i=1}^{n}(\lambda' z_i)\{\log \bar{F}_0(X_i) + \delta_i\}$$

$$+ \frac{1}{2n} \Sigma_{i=1}^{n}(\lambda' z_i)^2\{\log \bar{F}_0(X_i)\} + O_p(n^{-\frac{1}{2}}) . \qquad (3.14)$$

Thus, if $n^{-1} \Sigma_{i=1}^{n} Z_i Z_i' \to \nu^*$ (in probability, if the Z_i are stochastic), then by (3.14), under H_0: $\beta = 0$, as $n \to \infty$,

$$\log\{L_n(n^{-\frac{1}{2}}\lambda)/L_n(0)\} \sim N(-\tfrac{\pi}{2}\lambda'\nu^*\lambda,\ \pi\lambda'\nu^*\lambda) \qquad (3.15)$$

where

$$\pi = \int \bar{G}(x)dF_0(x) = \int \bar{G}(x)\bar{F}_0(x)h_0(x)dx$$

$$= \int (h_0(x)/\{h_0(x) + h_G(x)\})dS_0(x)$$

$$= \int \pi(x)dS_0(x) , \qquad (3.16)$$

and $\pi(x)$ is defined by (2.4). Now (3.15) establishes the contiguity of the probability measure under $\{K_n\}$ with respect to that under H_0. Further, by (3.13),

$$U_n = (\partial/\partial\beta)\log L_n(\beta)\Big|_{\beta=0} = \Sigma_{i=1}^{n} Z_i\{\log \bar{F}_0(X_i) + \delta_i\} , \qquad (3.17)$$

and hence, it follows by some standard steps that if $\bar{L}_n$ stands for the likelihood ratio test statistic [for testing H_0: $\beta = 0$ vs H: $\beta \neq 0$ on the model (3.12)], then under H_0,

$$-2 \log \bar{L}_n \sim \chi_p^2 , \qquad (3.18)$$

while under $\{K_n\}$, $-2 \log \bar{L}_n$ has asymptotically a non-central chi square distribution with p D.F. and noncentrality parameter

$$\bar{\Delta} = \pi(\lambda'\nu^*\lambda) . \qquad (3.19)$$

Note that by (1.7), (2.21) and the definition of ν^*, $\nu^* - \nu$ is positive semi-definite (of rank 1 at most), and hence

$$\underset{\sim}{\lambda}'\underset{\sim}{\nu}^*\underset{\sim}{\lambda} - \underset{\sim}{\lambda}'\underset{\sim}{\nu}\underset{\sim}{\lambda} \geq 0, \ \forall \ \underset{\sim}{\lambda} . \tag{3.20}$$

[Actually, if the $\underset{\sim}{Z}_i$ are non-stochastic, the model in (1.1) may be so chosen that $\overline{\underset{\sim}{Z}}_n = \underset{\sim}{0}$ and this will lead to $\underset{\sim}{\nu} = \underset{\sim}{\nu}^*$. Even, otherwise, in (1.1), $\underset{\sim}{Z}_i$ may be replaced by $(\underset{\sim}{Z}_i - \overline{\underset{\sim}{Z}}_n)$, $1 \leq i \leq n$, and this will lead to $\underset{\sim}{\nu}^* = \underset{\sim}{\nu}$. Thus, we may assume without any essential loss of generality that $\underset{\sim}{\nu} = \underset{\sim}{\nu}^*$, so that the equality sign in (3.20) holds.]

By (2.25), (3.5) and (3.19)-(3.20), we obtain that the asymptotic relative efficiency of the log-rank test relative to the likelihood ratio test [for the model (3.12)] is

$$\begin{aligned} e(L^0,\overline{L}) &= \Delta^0/\overline{\Delta} = (\Delta^0/\Delta^*)(\Delta^*/\overline{\Delta}) \\ &\leq [\rho^2(\phi^*,\psi^*) \int_0^1 \psi_2^2(u)du]/\pi \\ &= [\rho^2(\phi^*,\psi^*)]\{(\int\pi^2(x)dS_0(x))/\int\pi(x)dS_0(x)\} \\ &= \rho^2(\phi^*,\psi^*)\}\{\rho_2\} \end{aligned} \tag{3.21}$$

where $\rho^2(\phi^*,\psi^*) \leq 1$, with the equality sign holding under (3.7)-(3.9) and

$$0 < \rho_2 = \int\pi^2(x)dS_0(x)/\int\pi(x)dS_0(x) \leq 1 , \tag{3.22}$$

with the equality sign on the right hand side holding only when $\pi(x) = 1$ almost everywhere (S_0), i.e., $h_G(x) = 0$ almost everywhere. Note that even if (3.9) holds [viz., (3.10)], (3.21) will be generally less than 1, due to ρ_2, so that there is always some inherent loss or efficiency due to ignoring $\underset{\sim}{\delta}_n$ and using a rank test based on the X_i alone.

4. ASYMPTOTIC EFFICIENCY OF THE COX PROCEDURE

If $T = \{t_1<\ldots<t_m\} = \{X_i \text{ (ordered)}: \delta_i=1, i=1,\ldots,n\}$ be the set of *failure points* (for which W_i exceeds Y_i), then a partial likelihood

function may be defined as in Cox (1972, 1975) as follows. At time $t_j - 0$, there is a risk set R_j of r_j individuals which have neither failed nor dropped out by that time, for $j=1,\dots,m$, note that $R_m \subset \dots \subset R_1$. Considering the risk set R_j and the conditional probability of a failure at time t_j $(1 \le j \le m)$, we obtain the partial log-likelihood function (from (1.11))

$$\log L_m^{**} = \Sigma_{j=1}^m \{\underset{\sim}{\beta}' \underset{\sim}{Z}_{Q_j^*} - \log(\Sigma_{i \in R_j} \exp\{\underset{\sim}{\beta}' \underset{\sim}{Z}_i\})\} \tag{4.1}$$

where $\underset{\sim}{Q}^* = (Q_1^*, \dots, Q_m^*)$ is a sub-vector of the anti-ranks, relating to the indices of the observations corresponding to the ordered failures (preceding withdrawals). For testing H_0: $\underset{\sim}{\beta} = \underset{\sim}{0}$ against $\underset{\sim}{\beta} \neq \underset{\sim}{0}$, Cox (1972) considered the test statistic

$$L_{nm} = \underset{\sim}{U}'_{nm} \underset{\sim}{J}^{-}_{nm} \underset{\sim}{U}_{nm} \tag{4.2}$$

where

$$\underset{\sim}{U}_{nm} = \Sigma_{j=1}^m \{\underset{\sim}{Z}_{Q_j^*} - r_j^{-1} \Sigma_{k \in R_j} \underset{\sim}{Z}_k\} , \tag{4.3}$$

$$\underset{\sim}{J}_{nm} = \Sigma_{j=1}^m r_j^{-1} \Sigma_{i \in R_j} (\underset{\sim}{Z}_i - r_j^{-1} \Sigma_{k \in R_j} \underset{\sim}{Z}_k)(\underset{\sim}{Z}_i - r_j^{-1} \Sigma_{k \in R_j} \underset{\sim}{Z}_k)' . \tag{4.4}$$

Whenever, π, defined by (3.16) is > 0, under H_0, L_{mn} has asymptotically chi-square distribution with p D.F. Also, it follows from Sen (1981, Sec. 4) that under $\{K_n\}$ in (2.21), L_{nm} has asymptotically a noncentral chi-square distribution with p D.F. and noncentrality parameter

$$\Delta = \pi(\underset{\sim}{\lambda}' \underset{\sim}{\nu} \underset{\sim}{\lambda}) . \tag{4.5}$$

By virtue of the remarks made after (3.20), (4.4) is quite comparable to (3.19), and this reveals the asymptotic optimality of L_{nm}, under (2.21). We may note that the information on $\underset{\sim}{\delta}_n$ is incorporated in the Cox procedure through the construction of the risk sets R_j, $1 \le j \le m$, and this explains the better efficiency when the model in (1.1) holds. Finally, unlike the log-rank test, L_{nm} is not a rank statistic.

5. SOME GENERAL REMARKS

The results in Section 3 reveal the loss in efficiency of the log-rank test (or L_n^*) [due to ignoring the information in δ_n and restricting to the ranks of the X_i] when the Cox model in (1.1) holds; the Cox procedure remains asymptotically locally optimal for the same model and it incorporates the information in δ_n. However, it may be remarked that whenever under the null hypothesis, $X_1,\ldots,X_n$ are i.i.d.r.v., the log-rank test is a genuinely distribution-free test. This is particularly true when instead of the Cox model in (1.1), the $F_i(x|z_i)$ are given by the conventional regression model $F(x-\beta' z_i)$, so that $\bar{S}_i(x) = \bar{G}(x)\bar{F}(x-\beta' z_i)$, $1 \leq i \leq n$, where under H_0: $\beta = 0$, $\bar{S}_1,\ldots,\bar{S}_n$ are the same. In such a case, the log-rank test remains valid for a general class of $\bar{F}$, $\bar{G}$. On the other hand the rationality and/or optimality of L_{nm} for this conventional regression model may be open to questions. Thus, if we have random censoring, it may be a basic issue whether to stick to the Cox model and adapt the locally optimal Cox procedure (which may not be robust for departures from the Cox model) or to use the log-rank test which remains valid (under more general setups) and reasonably efficient for a broad class of models. In any case, if π, defined by (3.16) is small, the (random) censoring results in substantial loss of efficiency, and, by (3.22), a greater loss is incurred for the log-rank procedure. Hence, when π is close to 0, rank procedures may not be recommended.

ACKNOWLEDGEMENT

This work was partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

BIBLIOGRAPHY

Cox, D.R. (1972). Regression models and life tables. *Jour. Roy. Statist. Soc. Ser. B. 34,* 187-220.

Cox, D.R. (1975). Partial likelihoods. *Biometrika 62,* 269-276.

Hájek, J. & Šidák, Z. (1967). *Theory of Rank Tests.* New York: Academic Press.

Lehmann, E.L. (1953). The power of rank tests. *Ann. Math. Statist. 24,* 23-43.

Peto, R. (1972). Rank tests for maximal power against Lehmann-type alternatives. *Biometrika 59,* 472-474.

Peto, R. & Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Jour. Roy. Statist. Soc. Ser. A. 34,* 185-207.

Sen. P.K. (1981). The Cox regression model, invariance principles for some induced quantile processes and some repeated significance tests. *Ann. Statist. 9,* 109-121.

Sen, P.K. (1982). The UI-principle and LMP rank tests. In *Colloquium on Nonparametric Statistical Inference* (ed: B.V. Gnedenko et.al.). Amsterdam: North Holland (in press).