

SUBHYPOTHESES TESTING AGAINST RESTRICTED ALTERNATIVES
FOR THE COX REGRESSION MODEL

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1425

January 1983

Subhypotheses Testing Against Restricted Alternatives for the Cox Regression Model

BY PRANAB KUMAR SEN

*Department of Biostatistics, University of North Carolina, Chapel Hill,
North Carolina, U.S.A .*

SUMMARY

For the Cox regression model involving both design and concomitant variates, testing of subhypotheses against restricted alternatives based on the partial likelihood scores is considered. Roy's union-intersection principle plays a vital role in this context. Properties of the proposed tests are also studied.

Some key words : *chi-bar distribution; Kuhn-Tucker-Lagrange formula; partial likelihood; several sample problem; survival function; union-intersection principle.*

1. INTRODUCTION

We consider here the general regression model, due to Cox(1972), involving both design variables (controlled and nonstochastic) and concomitant variables (stochastic). It is assumed that the i th subject (having survival time Y_i) has the hazard rate (given the design variate c_i and the covariate z_i)

$$h_i(t; c_i, z_i) = h_0(t) \cdot \exp(\beta^T c_i + \gamma^T z_i), \quad i=1, \dots, N, \quad (1.1)$$

where $h_0(t)$, the hazard rate for $c_i = 0$, $z_i = 0$, is an unknown, arbitrary non-negative function (for which $\int_0^\infty h_0(t) dt = +\infty$), β is the vector of unknown parameters relating to the design effects and γ is the unknown parameter vector relating to the effects of the concomitant variates. Generally, we are interested in the null hypothesis : $H_0: \beta = 0$ against $\beta \neq 0$, where we treat γ as a nuisance parameter. In many problems of practical interest, though the null hypothesis remains the same, the alternative hypotheses may

be restricted in nature. For example, if we consider the $k (=r+1)$ sample problem involving a control and $r (\geq 1)$ treatment groups, then the c_i^T may be either of the vectors $(0,0,\dots,0)$, $(1,0,\dots,0)$, $(0,1,\dots,0), \dots, (0,\dots,0,1)$, so that if we write $\beta^T = (\beta_1, \dots, \beta_r)$, then β_j stands for the j th treatment effect (over the control), for $j=1, \dots, r$. In such a case, against the null hypothesis of no treatment effect (i.e., $\beta = 0$), we may be interested in the alternatives that none of the treatment is inferior to the control, i.e.,

$$H^> : \beta_j \geq 0, j=1, \dots, r, \text{ with at least one strict inequality. } (1.2)$$

This may be termed an orthant alternative, as in (1.2), β belongs to the positive orthant. A second hypothesis of interest may be the ordering of the treatment effects when H_0 may not hold, i.e.,

$$H^* : \beta_1 \leq \dots \leq \beta_r, \text{ with at least one strict inequality. } (1.3)$$

This may be termed an ordered alternative. We are basically interested in testing for $H_0: \beta = 0$ against an orthant or ordered alternative. A third case of interest is the following :

$$H^{*>} : 0 \leq \beta_1 \leq \dots \leq \beta_r, \text{ with at least one strict inequality, } (1.4)$$

and this may be termed an ordered orthant alternative.

As is usual the case, the survival times Y_i may not be observable; the observable random variables are the $X_i = \min(Y_i, T_i) = Y_i \wedge T_i$ and $\delta_i = I(Y_i \leq T_i)$, where the censoring variables T_i are either fixed or independent of Y_i , given c_i, z_i . Further, the covariates z_i are assumed to be given at the start, so that c_i, z_i are observable. Therefore, based on the set $(X_i, \delta_i, c_i, z_i)$, $i=1, \dots, N$, our problem is to test for $H_0: \beta = 0$ against $H^>$, H^* or $H^{*>}$. For this testing problem, the Cox (1972) omnibus test based on the partial likelihood is a valid one, but is generally not very efficient. The situation is very much comparable with the classical parametric case, where the global likelihood ratio test may not be very efficient for restricted

alternatives , for which there are other versions having better power properties [see, for example, Barlow et. al. (1972, ch.3 & 4)]. In the multivariate non-parametric case, Chinchilli and Sen (1981 a,b) have used the Roy (1953) union-intersection (UI-) principle for generating rank order tests for restricted alternatives. In the current problem, as in Cox (1972), we shall make use of the partial likelihood function and incorporate the UI-principle to generate alternative tests for the restricted alternative problems in (1.2), (1.3) or (1.4). Along with the preliminary notions, the partial likelihood, UI-principle and the proposed test procedures are considered in Section 2. Section 3 is devoted to the study of the distribution theory of the tests under the null hypothesis. General comments on the non-null case are made in the concluding section.

2. THE PROPOSED TESTS

Note that the partial likelihood function is given by

$$L_N^P(\beta, \gamma) = \prod_{i=1}^N \left\{ \frac{\exp(\beta^T c_i + \gamma^T z_i)}{\sum_{j=1}^N I(X_j > X_i) \exp(\beta^T c_j + \gamma^T z_j)} \right\}^{\delta_i} \quad (2.1)$$

Thus, under $H_0: \beta = 0$, $\hat{\gamma}_N^o$, the maximum partial likelihood estimator (MPLE) of γ is obtained as the solution of the system of equations:

$$\sum_{i=1}^N \delta_i \left\{ z_i - \left(\frac{\sum_{j=1}^N I(X_j > X_i) z_j e^{\gamma^T z_j}}{\sum_{j=1}^N I(X_j > X_i) e^{\gamma^T z_j}} \right) \right\} = 0 \quad (2.2)$$

Consider then the partial likelihood scores :

$$\begin{aligned} \hat{U}_N &= N^{-1/2} (\partial/\partial\beta) \log L_N^P(\beta, \gamma) \Big|_{\beta=0, \gamma=\hat{\gamma}_N^o} \\ &= N^{-1/2} \sum_{i=1}^N \delta_i \left\{ c_i - \left(\frac{\sum_{j=1}^N I(X_j > X_i) c_j e^{z_j^T \hat{\gamma}_N^o}}{\sum_{j=1}^N I(X_j > X_i) e^{z_j^T \hat{\gamma}_N^o}} \right) \right\} \end{aligned} \quad (2.3)$$

Also, let

$$\bar{V}_{11} = -N^{-1} (\partial^2/\partial\beta\partial\beta^T) \log L_N^P(\beta, \gamma) \Big|_{\beta=0, \gamma=\hat{\gamma}_N^o} \quad (2.4)$$

$$\bar{V}_{12} = \bar{V}_{21}^T = -N^{-1} (\partial^2/\partial\beta\partial\gamma^T) \log L_N^P(\beta, \gamma) \Big|_{\beta=0, \gamma=\hat{\gamma}_N^o} \quad (2.5)$$

$$\bar{V}_{22} = -N^{-1} (\partial^2/\partial\gamma\partial\gamma^T) \log L_N^P(\beta, \gamma) \Big|_{\beta=0, \gamma=\hat{\gamma}_N^o} ; \quad (2.6)$$

$$\bar{V}_{11.2} = \bar{V}_{11} - \bar{V}_{12} \bar{V}_{22}^{-1} \bar{V}_{21} \quad (2.7)$$

Note that if we let $w_{ij} = I(X_j > X_i) \exp(z_j^T \hat{\gamma}_N^0)$, for $i, j=1, \dots, N$, then $\bar{V}_{11} = N^{-1} \sum_{i=1}^N \delta_i \{ \sum_{j=1}^N w_{ij} c_j c_j^T / \sum_{j=1}^N w_{ij} - (\sum_{j=1}^N w_{ij} c_j) (\sum_{j=1}^N w_{ij} c_j^T) / (\sum_{j=1}^N w_{ij})^2 \}$, and similar expressions hold for \bar{V}_{12} (with the c_j^T being replaced by z_j^T) and \bar{V}_{22} (with the c_j and c_j^T being replaced by z_j and z_j^T , respectively).

The omnibus partial likelihood ratio test statistic (for testing $H_0: \beta = 0$ against $\beta \neq 0$), due to Cox(1972), is given by

$$\mathcal{L}_N^0 = \hat{U}_N^T (\bar{V}_{11.2})^- \hat{U}_N, \quad (2.8)$$

where A^- stands for a generalized inverse of A . To motivate the proposed tests,

we may note that by definition $\mathcal{L}_N^0 = \sup \{ (a^T \hat{U}_N) / (a^T \bar{V}_{11.2} a)^{1/2} : a \neq 0 \}$. Or, in other words, it is the global maximum of the normalized linear combinations

of the elements of \hat{U}_N . For testing H_0 vs $H^>$ in (1.2), we make use of the union-intersection principle. Note that $\beta > 0 \iff a^T \beta > 0$, for every $a > 0$.

An asymptotically efficient test for H_0 against the specific alternative $H_a^+ : a^T \beta > 0$ is based on the test statistic $a^T \hat{U}_N / (a^T \bar{V}_{11.2} a)^{1/2}$, rejecting the null hypothesis for large positive values of the statistic. Also, for each a , it

follows from the results of Cox(1972), Tsiatis(1981), Sen(1981) and Anderson

and Gill (1982), among others, that under H_0 , $a^T \hat{U}_N / (a^T \bar{V}_{11.2} a)^{1/2}$ has closely

the normal distribution with 0 mean and unit variance, so that the normal

percentile point may be used to demarcate the critical region. Hence, according

to the Roy (1953) UI-principle, it seems appropriate to choose the test

statistic for testing H_0 vs $H^>$ in (2.1) as

$$\mathcal{L}_N^{(1)} = \sup \{ a^T \hat{U}_N / (a^T \bar{V}_{11.2} a)^{1/2} : a > 0 \}. \quad (2.9)$$

Thus, we need to maximise $a^T \hat{U}_N$ subject to $a > 0$ and $a^T \bar{V}_{11.2} a = \text{constant}$. If

we let $h(a) = -a^T \hat{U}_N$, $h_1(a) = -a$ and $h_2(a) = (a^T \bar{V}_{11.2} a - 1)$, then the problem reduces to minimize $h(a)$ subject to the constraints: $h_1(a) \leq 0$ and $h_2(a) = 0$.

For this non-linear programming problem, the Kuhn-Tucker-Lagrange (KTL-) point formula theorem applies, and, we arrive at the following solution.

Note that \hat{U}_N has $r(\geq 1)$ components. Let a be any subset of $P = \{1, \dots, r\}$ and \bar{a} be the complementary subset ($\emptyset \subseteq a \subseteq P$). For each a , we partition (following possible rearrangements) \hat{U}_N^T and $\bar{V}_{11.2}$ as

$$\left(\hat{U}_{N(a)}^T, \hat{U}_{N(\bar{a})}^T \right) \text{ and } \begin{pmatrix} \bar{V}_{11.2(aa)} & \bar{V}_{11.2(a\bar{a})} \\ \bar{V}_{11.2(\bar{a}a)} & \bar{V}_{11.2(\bar{a}\bar{a})} \end{pmatrix}. \quad (2.10)$$

Also, for each a , let

$$\hat{U}_{N(a)}^* = \hat{U}_{N(a)} - \bar{V}_{11.2(a\bar{a})} \bar{V}_{11.2(\bar{a}\bar{a})}^{-1} \hat{U}_{N(\bar{a})}^*, \quad (2.11)$$

$$\bar{V}_{11.2(a)}^* = \bar{V}_{11.2(aa)} - \bar{V}_{11.2(a\bar{a})} \bar{V}_{11.2(\bar{a}\bar{a})}^{-1} \bar{V}_{11.2(\bar{a}a)}. \quad (2.12)$$

Then, the UI-test statistic in (2.9) is given by

$$\sum_{\emptyset \subseteq a \subseteq P} \{ (\hat{U}_{N(a)}^* \bar{V}_{11.2(a)}^*{}^{-1} \hat{U}_{N(a)}^*)^{\frac{1}{2}} I(\hat{U}_{N(a)}^* \geq 0) I(\bar{V}_{11.2(\bar{a}\bar{a})}^{-1} \hat{U}_{N(\bar{a})}^* \leq 0) \}. \quad (2.13)$$

Consider next the test for $H_0: \beta = 0$ vs. $H^>$ in (1.4). Let us write $\beta^T = (\beta_1, \dots, \beta_r)$, $\lambda^T = (\lambda_1, \dots, \lambda_r)$ and $\lambda_j = \beta_j - \beta_{j-1}$, $j=1, \dots, r$, where $\beta_0 = 0$. Also, we write $c_i^T = (c_{i1}, \dots, c_{ir})$, $i=1, \dots, N$ and $d_i^T = (d_{i1}, \dots, d_{ir})$, $i=1, \dots, N$, where, we take $d_{ij} = \sum_{s=j}^r c_{is}$, for $j=1, \dots, r$ and $i=1, \dots, N$. With these changes in notations, we may rewrite (1.1) as

$$h_0(t) \cdot \exp(\lambda^T d_i + \gamma^T z_i), \quad i=1, \dots, N. \quad (2.14)$$

In terms of this reparameterized model in (2.14), the null hypothesis $H_0: \beta = 0$ reduces to that of $H_0: \lambda = 0$, while the ordered orthant alternative in (1.4) reduces to the ordinary orthant alternative $H^>: \lambda > 0$. As such, if in (2.1) through (2.7) as well as in (2.10) through (2.12), we replace the c_i by the d_i , then the resulting statistic in (2.13) would be the test statistic for testing H_0 against the ordered orthant alternative.

Finally, we consider the case of the ordinary ordered alternative hypothesis in (1.3), where, under the alternative, β_1 need not be non-negative. With the reparameterization in (2.14), our null hypothesis is $H_0: \lambda = 0$, while, the alternative is $\lambda_j \geq 0$, for $j=2, \dots, r$, while λ_1 is unrestricted. As such, in this case, in (2.9), we allow for $a^T = (a_1, \dots, a_r)$, a_1 to be

unrestrained, while the $a_j, j=2, \dots, r$ are all non-negative. For the single element set $a = 1$, $\bar{a} = \{2, \dots, r\} = P^0$, say, we define $\hat{U}_{N(1)}^*$ and $\bar{V}_{11.2(1)}^*$ (scaler) as in (2.11) and (2.12), and let $Q_{N(1)}^2 = (\hat{U}_{N(1)}^*)^2 / \bar{V}_{11.2(1)}^*$. Further, we let $\hat{U}_N^T = (\hat{U}_{N1}, \hat{U}_N^{oT})$, $\hat{U}_N^{oT} = (\hat{U}_{N2}, \dots, \hat{U}_{Nr})$ and the minor of $\bar{V}_{11.2}$ comprising the last $r-1$ rows and columns is denoted by $\bar{V}_{11.2}^0$. Then, for every $a \in P^0$, we define $\hat{U}_{N(a)}^{o*}$, $\hat{U}_{N(\bar{a})}^{o*}$, $\bar{V}_{11.2(a)}^{o*}$ etc., as in (2.11)-(2.12) (but for the $r-1$ dimensional case). Note that we are to work here with the reparameterized model in (2.14), so that for all these statistics, we use the vectors d_i instead of c_i . Then, $\chi_N^{(2)}$, the proposed test statistic, for testing H_0 against H^* , is given by

$$\sum_{\beta \in \bar{a} \in P^0} \{ (Q_{N(1)}^2 + \hat{U}_{N(a)}^{o*T} \bar{V}_{11.2(a)}^{o*-1} \hat{U}_{N(a)}^{o*})^{1/2} I(\hat{U}_{N(a)}^{o*} \geq 0) I(\bar{V}_{11.2(\bar{a})}^{o*-1} \hat{U}_{N(\bar{a})}^{o*} \leq 0) \} \quad (2.15)$$

In passing, we may remark that the UI-test statistic in (2.13) or (2.15) may be readily extended to test for $H_0: \beta = 0$ against $H^{**}: \{\beta^T = (\beta_{(1)}^T, \beta_{(2)}^T): \beta_{(2)} \geq 0, \beta_{(1)} \neq 0\}$ or having an ordered relationship among the elements of $\beta_{(2)}$, leaving $\beta_{(1)}$ unrestrained. The only difference would be to replace in (2.15) $Q_{N(1)}^2$ by a quadratic form in $\hat{U}_{N(\ell)}^*$ with discriminant $\bar{V}_{11.2(\ell)}^{*-1}$ (where ℓ is the set of indices of $\beta_{(1)}$) and, restricted to the subset of indices of $\beta_{(2)}$, the other statistics are to be computed as in (2.11), (2.12) and (2.15). Further, if instead of testing $H_0: \beta = 0$, we like to test for $H_0: \beta_{(1)} = 0$ against an orthant or ordered alternative relating to $\beta_{(1)}$ alone, then the testing problem may be handled similarly by the following reparameterization. Write $c_i^T = (c_{i(1)}^T, c_{i(2)}^T)$, $i=1, \dots, N$, $\xi^T = (\beta_{(2)}^T, \gamma^T)$ and $z_i^{*T} = (c_{i(2)}^T, z_i^T)$, for $i=1, \dots, N$. Then, the model in (1.1) may be written as

$$h_0(t) \exp(\beta_{(1)}^T c_{i(1)} + \xi^T z_i^*) \quad , \quad i=1, \dots, N \quad , \quad (2.16)$$

where the z_i^* involve partly nonstochastic and partly conditionally fixed stochastic elements. Treating ξ as a nuisance parameter vector, we may then work with the statistic in (2.13) or (2.15) with obvious modifications on the sets P, a and the allied component vectors and matrices, as defined in (2.11)-(2.12).

3. DISTRIBUTION THEORY UNDER THE NULL HYPOTHESIS

Note that if $W^T = (W_{(1)}^T, W_{(2)}^T)$ has a multinormal distribution with a dispersion matrix $\Sigma = ((\Sigma_{ij}))_{i,j=1,2}$ and if, as in (2.11)-(2.12), we define $W_{(1)}^* = W_{(1)} - \Sigma_{12}\Sigma_{22}^{-1}W_{(2)}$, $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, then the pair of events $I(W_{(1)}^* \geq 0)$ and $I(\Sigma_{22}^{-1}W_{(2)} \leq 0)$ are independent, irrespective of whether the mean vector of W is null or not. Also, if we let $Q_{(1)}^2 = W_{(1)}^{*T}\Sigma_{11.2}^{-1}W_{(1)}^*$, then, it follows from Lemma 3.2 of Kudo (1963) that when $EW = 0$, the three events $I(Q_{(1)} \leq x)$, $I(W_{(1)}^* \geq 0)$ and $I(\Sigma_{22}^{-1}W_{(2)} \leq 0)$ are mutually independent, for every $x \geq 0$. Thus, when $EW = 0$, for every $x \geq 0$,

$$\begin{aligned} & P\{ Q_{(1)} I(W_{(1)}^* \geq 0) I(\Sigma_{22}^{-1}W_{(2)} \leq 0) \leq x \} \\ &= P\{ Q_{(1)} \leq x \} \cdot P\{ W_{(1)}^* \geq 0 \} \cdot P\{ \Sigma_{22}^{-1}W_{(2)} \leq 0 \} \\ &= P\{ \chi_k \leq x \} \cdot P\{ W_{(1)}^* \geq 0 \} \cdot P\{ \Sigma_{22}^{-1}W_{(2)} \leq 0 \} , \end{aligned} \quad (3.1)$$

where k is the dimension of the vector $W_{(1)}$ and χ_k has the central chi distribution with k degrees of freedom (DF). For $k = 0$, χ_0 is equal to 0 with probability 1. Further, both $P\{ W_{(1)}^* \geq 0 \}$ and $P\{ \Sigma_{22}^{-1}W_{(2)} \leq 0 \}$ are normal orthant probabilities, for two independent normal distributions with null mean vectors and dispersion matrices $\Sigma_{11.2}$ and Σ_{22}^{-1} , respectively. Now, as in (2.10)-(2.12), we define the sets a in P and the corresponding statistics $W_{(a)}^*$ and Σ_{aa}^* , for every a in P . Also, we define a statistic as in (2.13) but involving these $W_{(a)}^*$ and $W_{(\bar{a})}$. Then, we may note that of the 2^P possible terms in (2.13) only one term will have both the indicator functions different from 0. As such, using (3.1) and arguing as in Kudo (1963), we may conclude that based on the vector W and the corresponding Σ , the UI-test statistic in (2.13) will have the distribution (when $EW = 0$):

$$\sum_{\emptyset \subseteq a \subseteq P} P\{ \chi_{k_a} \leq x \} \cdot P\{ W_{(a)}^* \geq 0 \} \cdot P\{ \Sigma_{aa}^{-1} W_{(\bar{a})} \leq 0 \}, \quad (3.2)$$

where k_a is the cardinality of the set a , $a \in P$. Regrouping the 2^P terms in (3.2) in terms of the cardinality, we obtain that (7.2) is equivalently

$$\sum_{k=0}^r \omega_k P\{ \chi_k \leq x \} , x \geq 0, \quad (3.3)$$

where the ω_k are nonnegative quantities with $\sum_{k=0}^r \omega_k = 1$. In the statistical literature, (3.3) is referred to as the *chi-bar distribution*. The weights ω_k can be computed from the normal orthant probabilities; some tables for the same are given in Gupta(1963), Barlow et. al (1972), and other places.

We shall make use of (3.1)-(3.3) in our study of the large sample distribution theory of the UI-test statistics under the null hypothesis. It follows from the general theory of partial likelihood functions relating to the Cox (1972) model [see for example, Tsiatis(1981), Sen(1981), Slud(1982) and Anderson and Gill(1982), among others] that under $H_0: \beta = 0$ and some general regularity conditions (on the c_i and z_i), as stated in these papers,

$$\hat{U}_N \text{ is asymptotically normal with null mean vector and dispersion matrix } v_{11.2} , \quad (3.4)$$

and

$$\bar{V}_{11.2} \text{ converges in probability to } v_{11.2} , \quad (3.5)$$

where $v_{11.2}$ is a positive definite and finite matrix of rank r . Given these basic convergence results, we may justify the steps in (3.1) through (3.3).

For each $\alpha \in P$, we may use (2.10)-(2.12) along with (3.1), (3.4) and (3.5),

and claim that under H_0 , for every $x \geq 0$, $P\{[(\hat{U}_{N(\alpha)}^* T \bar{V}_{11.2(\alpha)}^* - 1) \hat{U}_{N(\alpha)}^*]^{\frac{1}{2}} I(\hat{U}_{N(\alpha)}^* \geq 0) \cdot I(\bar{V}_{11.2(\alpha\alpha)}^{-1} U_{N(\alpha)} \leq 0) \leq x\}$ converges to $P\{\chi_{k_\alpha} \leq x\} \cdot P(\alpha)$, where k_α is the

cardinality of the set α and $P(\alpha)$ is the product of two normal orthant probabilities with respective dispersion matrices $v_{11.2(\alpha)}^*$ and $v_{11.2(\alpha\alpha)}^{-1}$. Also,

of the 2^r terms in (2.13), only for one term the two indicator functions are simultaneously different from 0. Hence, we conclude that under the null

hypothesis $H_0: \beta = 0$, the UI-test statistic $\mathcal{L}_N^{(1)}$ in (2.13) has asymptotically the chi-bar distribution in (3.3), where the weights ω_k are to be computed

from the multinormal distribution with dispersion matrix $v_{11.2}$. In practice,

because of (3.5), $\bar{V}_{11.2}$ may be used instead of $v_{11.2}$ to determine these weights.

Let us next consider the null distribution of $\hat{\mathcal{L}}_N^{(2)}$ in (2.15). Note that by virtue of (3.4), $\hat{U}_{N(1)}^*$ and \hat{U}_N^0 , defined before (2.15) are uncorrelated and asymptotically independent. Further, $Q_{N(1)}^2$ has asymptotically (under H_0) the chi square distribution with 1 DF, and in (2.15), the rest of the arguments all depend on \hat{U}_N^0 and the corresponding $\bar{V}_{11.2}^0$. Thus, in this case, we need to partition W into three components i.e., $W^T = (W_1, W_{(1)}^T, W_{(2)}^T)$, where W_1 is independent of $W_{(1)}$ or $W_{(2)}$, and then, (3.1) holds even when we replace the $Q_{(1)}$ by $bW_1^2 + Q_{(1)}$, $b \geq 0$. Thus, we may claim that under H_0 , for every $\alpha \in P^0$, the three events $I(Q_{N(1)}^2 + \hat{U}_{N(\alpha)}^{0*T} \bar{V}_{11.2(\alpha)}^{0* -1} \hat{U}_{N(\alpha)}^{0*} \leq x^2)$, $I(\hat{U}_{N(\alpha)}^{0*} \geq 0)$ and $I(\bar{V}_{11.2(\alpha\alpha)}^{0* -1} \hat{U}_{N(\alpha)}^0 \leq 0)$ are asymptotically independent, where, under H_c , $Q_{N(1)}^2 + \hat{U}_{N(\alpha)}^{0*T} \bar{V}_{11.2(\alpha)}^{0* -1} \hat{U}_{N(\alpha)}^{0*}$ has asymptotically chi square distribution with $k_\alpha + 1$ DF. Hence, proceeding on parallel lines, we conclude that under $H_0: \beta = 0$, the large sample distribution of $\hat{\mathcal{L}}_N^{(2)}$ is given by

$$\sum_{k=1}^r \omega_k^* P\{ \chi_k \leq x \}, \quad x \geq 0, \quad (3.6)$$

where the weights ω_k^* are nonnegative and $\sum_{k=1}^r \omega_k^* = 1$; these may be computed by using the matrix $\bar{V}_{11.2}^0$ [of order $(r-1) \times (r-1)$] instead of the unknown $v_{11.2}^0$.

We conclude this section with the following two remarks. First, both the distributions in (3.3) and (3.6) are dominated (in the upper tail) by the chi distribution with r DF. Hence, the percentile points of (3.3) or (3.6) are dominated by those of the chi distribution with r DF. This is not surprising as the restricted maximum is always smaller than or equal to the unrestricted one, so the test statistics $\hat{\mathcal{L}}_N^{(1)}$ and $\hat{\mathcal{L}}_N^{(2)}$ are always smaller than or equal to $(\hat{\mathcal{L}}_N^0)^{1/2}$. On the other hand, under the specific alternative considered, $\hat{\mathcal{L}}_N^{(1)}$ (or $\hat{\mathcal{L}}_N^{(2)}$) will be close to $(\hat{\mathcal{L}}_N^0)^{1/2}$, so that with smaller critical values, they would have better power properties. Secondly, based on some simulation studies on normally distributed vectors, it has been observed that the chi-bar approximation in (3.3) or (3.6) is satisfactory only for moderately large or large sample sizes; the situation is somewhat better with the chi square approximation

for the null distribution of the unrestricted test statistic \mathcal{L}_N^0 .

4. SOME GENERAL REMARKS

Under local alternatives, \mathcal{L}_N^0 in (2.8) has asymptotically non-central chi square distribution with r DF and some appropriate noncentrality parameter. The situation is more complicated with the restricted alternative tests. The contiguity of probability measures under such local alternatives to those under the null hypothesis may be established as in Sen(1981, Sec.4). However, even for such contiguous alternatives, the asymptotic distribution of \hat{U}_N , though normal, would have non-null mean vectors. For $EW \neq 0$, the independence in (3.1) does not hold, and, as a result, we may not have a noncentral chi-bar distribution for either of the statistics in (2.13) or (2.15), that is, we may not have an average of several noncentral chi distributions for the asymptotic distribution of $\mathcal{L}_N^{(1)}$ or $\mathcal{L}_N^{(2)}$. This situation is quite comparable to the parametric case, where the chi-bar approximation works out only for the null hypothesis case. In the parametric case, the power superiority of the restricted alternative tests to that of the global test has mostly been proved in various specific cases by particular constructions or by numerical studies. Chinchilli and Sen(1981b) have done some similar studies for the multivariate nonparametric case in some specific simpler models. The picture is expected to be the same in this case of partial likelihood statistics. However, this requires an extensive amount of simulation work, and is intended to be covered in a follow up study. For non-local alternatives, $N^{-\frac{1}{2}} \hat{U}_N$ converges to some non-null μ , so that under the alternative hypothesis of the restrictive type, we would expect a much better performance of the restricted tests. More numerical studies are needed to cover the case of small or moderate sample sizes and non-local alternatives.

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

R E F E R E N C E S

- ANDERSON, P.K. & GILL, R.D. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Statist.* 10, 1100-1120.
- BARLOW, R.E., BARTHOLOMEW, D.J., BREMNER, J.M. & BRUNK, H.D. (1972). *Statistical Inference Under Order Restrictions*. New York : Wiley.
- CHINCHILLI, V.M. & SEN, P.K. (1981a). Multivariate linear rank statistics and the union-intersection principle for hypothesis testing under restricted alternatives. *Sankhya, Ser. B* 42, 135-151.
- CHINCHILLI, V.M. & SEN, P.K. (1981b). Multivariate linear rank statistics and the union-intersection principle for the orthant restriction problem. *Sankhya, Ser. B* 42, 152-171.
- COX, D.R. (1972). Regression models and life-tables (with discussion). *J. R. Statist. Soc. Ser. B* 34, 187-202
- GUPTA, S.S. (1963). Bibliography on the multivariate normal integrals and related topics. *Ann. Math. Statist.* 34, 829-838.
- KUDO, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* 50, 403-418.
- ROY, S.N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Statist.* 24, 220-238.
- 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.
- SLUD, E.V. (1982). Consistency and efficiency of inferences with the partial likelihood. *Biometrika* 69, 547-552
- TSIATIS, A. (1981). A large sample study of Cox's regression model. *Ann. Statist.* 9, 109-121.