PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 13, Fasc. 1 (1992), pp. 127–137

TESTS OF FIT FOR COX'S REGRESSION MODEL

BY

M. D. BURKE* (CALGARY) AND E. GOMBAY** (EDMONTON, ALBERTA)

Abstract. An omnibus test of fit for Cox's proportional hazards regression model is proposed for continuous data. The procedure is extended to a random censorship model. Density estimation methods are used.

1. Introduction. One of the principal models of failure time data analysis is the proportional hazards model of Cox [6], [7]. This semiparametric model has been assumed as the underlying structure in numerous instances, and so it is important to have a test which can be used to determine whether it is appropriate in a given situation or not. Bednarski [3] shows how the Cox estimator can misbehave if the model is not correct. We propose an omnibus test procedure in which the test statistic is asymptotically normal under the null hypothesis that Cox's model is true. The results are generalized to the random censorship case in Section 4 (cf. Theorem 4.1).

Our approach is based on density function estimates. Although their convergence rate is slower than that of the sample distribution function, they enjoy the desirable property that their limiting distribution does not depend on the fact that parameters of the model must be estimated. For an approach based on the sample distribution function, the results of Durbin [8] and Burke et al. [4] indicate that the limiting behavior would depend on the parametric family of distribution functions underlying the model and possibly on the values of the unknown parameters.

Previous approaches are mostly based on data analytic techniques (e.g., Kay [12], Andersen [1] and Schoenfeld [15]). Schoenfeld [14] proposed a class of chi-squared tests where p-dimensional Euclidean space is partitioned into a finite number of classes. His approach, thus, discretizes the data and, by choosing different partitions, one arrives at different tests in the continuous case. While there are many ingredients in the density approach which can be varied (kernel function, bandwidth), this approach seems more natural in view of the model's definition in the continuous case. The monograph by Prakasa Rao [13] gives a good survey of density estimation results. Horváth [10] obtained asymptotic normality for L_p-norms of multivariate densities.

* Research partially supported by a Canadian NSERC grant.

** Research partially supported by the NSERC grant of M. D. Burke.

We define the *hazard rate function* of a random variable T given Z as

(1.1) $$\lambda(t, z) = \lim_{\Delta t \to 0} (\Delta t)^{-1} P\{t \leqslant T < t+\Delta t \mid T \geqslant t, Z = z\},$$

where T denotes the failure time and Z is the $(p-1)$-dimensional covariate or regressor variable. Our null hypothesis is that Cox's model is true:

(1.2) $$\mathrm{H}_0\colon \lambda(t, z) = \lambda_0(t)e^{z\beta},$$

where β is an unknown $(p-1)$-vector of regression parameters and $\lambda_0(t)$ an unknown base-line hazard function. Our results will also be true if we replace $e^{z\beta}$ by a known function $\eta(z, \beta)$ for which $\eta(0, \beta) = 1$.

Our test procedure will be based on the fact that H_0 is equivalent to

(1.3) $$\lambda(t, z)e^{-\beta z} = \lambda_0(t),$$

being a function of t only.

Let $F(t, z)$ denote the *joint survival function* of (T, Z), that is,

(1.4) $$F(t, z) = P\{T \geqslant t, Z \leqslant z\}.$$

We assume that the corresponding density $f(t, z)$ exists. Hence $\lambda(t, z)$ of (1.1) can be written as

(1.5) $$\lambda(t, z) = f(t, z)[g(t, z)]^{-1},$$

where

(1.6) $$g(t, z) = (\partial^{p-1}/\partial z_1, \ldots, \partial z_{p-1})F(t, z).$$

Statement (1.5) is well defined if the denominator is not zero. Our approach is to estimate β by $\hat{\beta}_n$, Cox's [7] partial likelihood estimator, the density f by f_n, a p-variate kernel estimate, and the derivative g by the estimator g_n of (1.10) below. We then arrive at the process

(1.7) $$X_n(t, z, w) = f_n(t, z)[g_n(t, z)]^{-1}\exp\{-\hat{\beta}_n z\} - f_n(t, w)[g_n(t, w)]^{-1}\exp\{-\hat{\beta}_n w\}.$$

Under H_0 and in view of (1.3), each term in the difference (1.7) is an estimate of the base-line hazard rate $\lambda_0(t)$. We will establish the asymptotic normality of

(1.8) $$W_n^2 = \iiint_D X_n^2(t, z, w)\,dt\,dz\,dw,$$

where $D = (0, Q) \times M^2$ (cf. Condition 2.1 (a)).

Let $(T_1, Z_1), (T_2, Z_2), \ldots, (T_n, Z_n)$ be independent random vectors with survival function (1.4) and let

$$F_n(u, v) = n^{-1} \sum_{i=1}^{n} I\{T_i > u, Z_i \leqslant v\}$$

denote the empirical survival function, $u \in R$, $v \in R^{p-1}$. For the kernel function $K(u, v)$ satisfying Condition 2.1 (c), we define

$$(1.9)\qquad f_n(t, z) = -\int b^{-p} K\big[b^{-1}[(t, z)-(u, v)]\big]\, dF_n(u, v)$$
$$= -(nb^p)^{-1} \sum_{i=1}^{n} K[b^{-1}(t-T_i, z-Z_i)],$$

where the "bandwidth" sequence of constants $\{b = b_n\}$ satisfies Condition 2.1 (e). Next, with the kernel function K_2 satisfying Condition 2.1 (d), we define

$$(1.10)\qquad g_n(t, z) = \int b^{-(p-1)} K_2[b^{-1}(z, v)]\, d_v F_n(t, v)$$
$$= (nb)^{-(p-1)} \sum_{i=1}^{n} K_2[b^{-1}(z-Z_i)] I\{T_i > t\}.$$

Lastly, let $\hat{\beta}_n$ be the sequence of estimators obtained by maximizing the partial likelihood (Cox [6], [7]):

$$L(\beta) = \prod_{i \in S} \{\exp\{\beta Z_i\}\big(\sum_{j \in R(t_i)} \exp\{\beta Z_j\}\big)^{-1}\},$$

where S is the set of indices $1, 2, \ldots, n$ corresponding to individuals who died (failed), t_i is the failure time of the i-th individual, and $R(t_i)$ is the set of indices corresponding to individuals who survived until time t_i.

In Section 2 we give the main results for the uncensored case. The proofs are indicated in Section 3. Although these results may be considered as preliminary to the results on randomly censored data (Section 4), they are of interest in their own right. (The behavior of Cox's partial likelihood estimator under a sequence of local alternatives is treated in Burke and Gombay [5].) We follow the approach of Hall [9] in our handling of density-type estimators.

2. The uncensored case. We will assume the following conditions:

CONDITION 2.1. (a) Let $\mathscr{Y} = (0, Q) \times M$ be the support of (T, Z), where M is a bounded subset of R^{p-1} having (finite) Lebesgue measure λ_M.

(b) Let f be the joint density of (T, Z). Assume that all partial derivatives of order 2 are bounded and uniformly continuous on R^p.

(c) Let K be a p-variate density function satisfying

$$\int u_i K(u)\,du = 0, \qquad \int u_i u_j K(u)\,du = C\delta_{ij} < \infty$$

for each $i, j = 1, 2, \ldots, p$, where the constant C does not depend on i and $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ otherwise.

(d) Let K_2 be a $(p-1)$-variate density function satisfying

$$\int u_i K_2(u)\,du = 0, \qquad \int u_i u_j K_2(u)\,du = C\delta_{ij} < \infty$$

for each $i, j = 1, 2, \ldots, p-1$, where C is independent of i.

(e) $b = b_n$ is a nonincreasing sequence of positive numbers such that

$$nb^p \to \infty \quad \text{and} \quad nb^{p+4} \to 0 \quad \text{as } n \to \infty.$$

The main result of this paper is

THEOREM 2.2. *Under Conditions* 2.1 *and the null hypothesis* H_0 *defined by* (1.2), *we have*

$$nb^{p/2}\sigma^{-1}(W_n^2 - \mu) \underset{D}{\longrightarrow} N(0, 1), \tag{2.1}$$

where W_n^2 *is defined by* (1.8), $\mu = \mu_{1n} + \mu_{2n} + \mu_{3n}$, *with*

$$\mu_{1n} = (nb^p)^{-1} \cdot 2\lambda_M \int_{\mathscr{Y}} \int_{R^p} K^2(v)a^2(x)f(x - bv)\,dv\,dx,$$

$$\mu_{2n} = (nb^{p-1})^{-1} \cdot 2\lambda_M \iint_{\mathscr{Y}} \int_{R^{p-1}} K_2^2(v_2)g(t, z - bv_2)\,dv_2\,dz\,dt, \tag{2.2}$$

$$\mu_{3n} = -4(nb^{p-1})^{-1} \int_{\mathscr{Y}} a(x)r(x) \int_{-\infty}^{0} \int_{R^{p-1}} K(v)K_2(v_2)f(x - bv)\,dv_2\,dv_1\,dx,$$

and

$$\sigma^2 = 8\lambda_M \iint_{\mathscr{Y}} \lambda_0^2(t)a^2(t, z) \cdot \int \left[\int K(u)K(u - v)\,dv\right]^2 du, \tag{2.3}$$

with

$$a(t, z) = [g(t, z)e^{z\beta}]^{-1}. \tag{2.4}$$

Remark 2.3. To use the result of Theorem 2.2 as a test of the null hypothesis H_0 of (1.2), one can estimate μ, σ of (2.2), (2.3) by $\hat{\mu}$, $\hat{\sigma}$, where $\hat{\mu}$ and $\hat{\sigma}$ are defined like μ and σ but with f, g and β replaced by f_n, g_n and $\hat{\beta}$, respectively. It is easy to show that

$$nb^{p/2}\hat{\sigma}^{-1}(W_n^2 - \hat{\mu}) \underset{D}{\longrightarrow} N(0, 1)$$

under H_0. Hence H_0 would be rejected if

$$nb^{p/2}\hat{\sigma}^{-1}(W_n^2 - \hat{\mu}) \geqslant z_{1-\alpha},$$

where $z_{1-\alpha}$ is the $(1-\alpha)100$ percentile of the standard normal distribution.

As an alternative to a test based on W_n^2, one can also consider the vector

$$\xi_n = [X_n(t_1, z_1, w_1), \ldots, X_n(t_k, z_k, w_k)]$$

and establish

THEOREM 2.4. *Assume that Conditions* 2.1 *hold and that the support of* K *is finite. Then, as* $n \to \infty$,

$$(nb^p)\xi_n \underset{D}{\longrightarrow} N,$$

where N *is a* k*-variate normal distribution with zero mean and covariance matrix*

Σ *having entries*

$$\sigma_{ij} = [a(t_i, z_i)a(t_j, z_j)f(t_i, z_i) - a(t_i, z_i)a(t_j, w_j)f(t_i, z_i)$$
$$- a(t_i, w_i)a(t_j, z_j)f(t_i, w_i)$$
$$+ a(t_i, w_i)a(t_j, w_j)f(t_i, w_i)] \int K^2(v)dv.$$

As a consequence of Theorem 2.4, by replacing Σ by its estimator $\hat{\Sigma}$ as in Remark 2.3, we have

$$(nb^p)\xi_n \hat{\Sigma}^{-1} \xi_n' \xrightarrow[D]{} \chi^2(k),$$

where $\chi^2(k)$ is a chi-square distribution with k degrees of freedom. The test: reject H_0 of (1.2) if

$$(nb^p)\xi_n \hat{\Sigma}^{-1} \xi_n' \geq \chi^2_{1-\alpha,k},$$

where

$$P\{\chi^2(k) \leq \chi^2_{1-\alpha,k}\} = 1-\alpha$$

is an asymptotically α-level test which would detect departures from H_0 at a finite number of points.

3. Proof of the uncensored results. We herewith sketch the proofs of the results. Details of the proofs can be found in the technical report of Burke and Gombay [5].

We will consider a closely related statistic to that of W_n^2, namely

$$[W_n^{(1)}] = \iiint_D [X_n^{(1)}(t, z, w)]^2 dtdzdw, \tag{3.1}$$

where

$$X_n^{(1)}(t, z, w) = a(t, z)[f_n(t, z) - f(t, z)] - a(t, w)[f_n(t, w) - f(t, w)] \tag{3.2}$$
$$- r(t, z)[g_n(t, z) - g(t, z)] + r(t, w)[g_n(t, w) - g(t, w)],$$

$a(t, z)$ is defined by (2.4), and $r(t, z) = f(t, z)a(t, z)g(t, z)^{-1}$. We will prove the following

THEOREM 3.1. *Under the conditions of Theorem* 2.2,

$$nb^{p/2}\sigma^{-1}([W_n^{(1)}]^2 - \mu) \xrightarrow[D]{} N(0, 1) \quad \text{as } n \to \infty,$$

where μ *and* σ *are defined by* (2.2) *and* (2.3), *respectively, and* $W_n^{(1)}$ *is defined by* (3.1).

We have the expansion

$$[W_n^{(1)}]^2 = \sum_{i=1}^{4} T_i,$$

where

$$T_1 = \iiint_D \{a(t, z)[f_n(t, z)-f(t, z)]-a(t, w)[f_n(t, w)-f(t, w)]\}^2 dtdzdw,$$

$$T_2 = \iiint_D \{r(t, z)[g_n(t, z)-g(t, z)]-r(t, w)[g_n(t, w)-g(t, w)]\}^2 dtdzdw,$$

(3.3)

$$T_3 = -4\lambda_M \iint_{\mathscr{Y}} a(t, z)r(t, z)[f_n(t, z)-f(t, z)][g_n(t, z)-g(t, z)]dtdz,$$

$$T_4 = 4\iiint_D a(t, z)[f_n(t, z)-f(t, z)]r(t, w)[g_n(t, w)-g(t, w)]dtdzdw.$$

We will first consider T_1 and write

$$T_1 = \sum_{i=1}^{4} T_{1i},$$

where

$$T_{11} = \iiint_D \{a(t, z)[f_n(t, z)-\mathrm{E}f_n(t, z)]-a(t, w)[f_n(t, w)-\mathrm{E}f_n(t, w)]\}^2 dtdzdw,$$

$$T_{12} = \iiint_D \{a(t, z)[\mathrm{E}f_n(t, z)-f(t, z)]-a(t, w)[\mathrm{E}f_n(t, w)-f(t, w)]\}^2 dtdzdw,$$

(3.4)

$$T_{13} = 4\lambda_M \iint_{\mathscr{Y}} a^2(t, z)[f_n(t, z)-\mathrm{E}f_n(t, z)][\mathrm{E}f_n(t, z)-f(t, z)]dtdz,$$

$$T_{14} = 4\iiint_D a(t, z)a(t, w)[f_n(t, z)-\mathrm{E}f_n(t, z)][\mathrm{E}f_n(t, w)-f(t, w)]dtdzdw.$$

Under Conditions 2.1,

(3.5) $$\sup_{\mathscr{Y}} |\mathrm{E}f_n(t, z)-f(t, z)-b^2 C\nabla^2 f(t, z)| \to 0,$$

where $\nabla^2 f$ is the Laplacian and C is a constant. Hence, as $n \to \infty$, we obtain

(3.6) $$T_{12} = O(b^4).$$

LEMMA 3.2. *Under Conditions* 2.1,

$$T_{13} \xrightarrow[D]{} (4\lambda_M b^2 n^{-1/2} C\sigma_{13})\cdot Z,$$

where Z is a standard normal (0, 1) *random variable, C is a constant, and*

$$\sigma_{13}^2 = \int_{\mathscr{Y}} a^4[\nabla^2 f]^2 f - [\int_{\mathscr{Y}} a^2(\nabla^2 f) f]^2.$$

LEMMA 3.3. *Under Conditions* 2.1,

$$T_{14} \xrightarrow[D]{} (4b^2 n^{-1/2} C\sigma_{14})\cdot Z,$$

where Z is a standard normal random variable, C is a constant, and

$$\sigma_{14}^2 = \iint_{\mathscr{Y}} a^2(t, z) m^2(t) f(t, z) dtdz - \left[\iint_{\mathscr{Y}} a(t, z) m(t) f(t, z) dtdz\right]^2$$

with

$$m(t) = \int_{\mathscr{Y}} a(t, w) \nabla^2 f(t, w) dw. \tag{3.7}$$

As a consequence of (3.6) and Lemmas 3.2 and 3.3 we have

$$nb^{p/2}(T_{12}+T_{13}+T_{14}) \xrightarrow[P]{} 0 \quad \text{as } n \to \infty. \tag{3.8}$$

Hence the term T_{11} determines the asymptotic distribution of T_1 which is described by

LEMMA 3.4. *Under Conditions 2.1, $nb^{p/2}\sigma^{-1}(T_1-\mu_{1n}) \xrightarrow[D]{} Z$, where Z is a standard normal random variable, μ_{1n} and σ are defined by* (2.2) *and* (2.3), *respectively.*

We can treat the term T_2 of (3.3) in a similar manner to that of T_1 and write

$$T_2 = \sum_{i=1}^{4} T_{2i},$$

where T_{2i} is defined like T_{1i} in (3.4) but with g and r replacing f and a, respectively. We then obtain

$$nb^{p/2}(T_{22}+T_{23}+T_{24}) \xrightarrow[P]{} 0 \quad \text{as } n \to \infty \tag{3.9}$$

in a similar way to (3.8).

LEMMA 3.5. *Under Conditions* 2.1,

$$nb^{p/2}(T_2-\mu_{2n}) \xrightarrow[P]{} 0 \quad \text{as } n \to \infty,$$

where μ_{2n} is defined by (2.2).

Since g_n is a $(p-1)$-dimensional kernel estimator, the deviation of T_2 from its mean is asymptotically negligible as compared to T_1. Similarly, we have

LEMMA 3.6. *Under Conditions* 2.1,

$$nb^{p/2}(T_3-\mu_{3n}) \xrightarrow[P]{} 0, \quad nb^{p/2}T_4 \xrightarrow[P]{} 0 \quad \text{as } n \to \infty,$$

where T_3 and T_4 are defined by (3.3) *and μ_{3n} is defined by* (2.2).

Proof of Theorem 3.1. The theorem follows directly from (3.8), (3.9) and Lemmas 3.4, 3.5 and 3.6.

Proof of Theorem 2.2. We have

$$W_n^2 = \iiint_D [X_n^{(1)}(t, z, w)+R_n(t, z, w)]^2 dtdzdw$$
$$= [W_n^{(1)}]^2 + 2\iiint_D X_n^{(1)} R_n + \iiint_D R_n^2,$$

where

$$R_n(t, z, w) = \sum_{i=1}^{4} [R_{in}(t, z) - R_{in}(t, w)]$$

and

$$R_{1n}(t, z) = f_n(t, z) g(t, z)^{-1} [\exp\{-\hat{\beta} z\} - \exp\{-\beta z\}],$$

$$R_{2n}(t, z) = [g_n(t, z) g(t, z) \exp\{\hat{\beta} z\}]^{-1} [f_n(t, z) - f(t, z)][g(t, z) - g_n(t, z)],$$

(3.10) $$R_{3n}(t, z) = f(t, z)[g_n(t, z) g(t, z)]^{-1} [\exp\{-\hat{\beta} z\} - \exp\{-\beta z\}][g(t, z) - g_n(t, z)],$$

$$R_{4n}(t, z) = f(t, z)[g_n(t, z) g(t, z) \exp\{-\hat{\beta} z\}]^{-1} [g(t, z) - g_n(t, z)]^2.$$

Under conditions weaker than ours, Tsiatis [16] has shown that $n^{1/2}(\hat{\beta} - \beta)$ is asymptotically normal with zero mean and finite variance. For another approach, see Andersen and Borgan [2]. Hence, by the mean value theorem,

$$\exp\{\hat{\beta} z\} - \exp\{\beta z\} = O_p(n^{-1/2}),$$

uniformly in $z \in M$. Since f_n is a uniformly consistent estimator of f, we have

$$\sup_{\mathscr{Y}} |R_{1n}(t, z)| = O_p(n^{-1/2}).$$

Consequently,

$$\iiint_D a(t, z)[f_n(t, z) - f(t, z)] R_{1n}(t, z) dt dz dw = o_p(nb^{p/2}).$$

Using similar calculations to those above, we obtain

(3.11) $$\iiint_D X_n^{(1)} R_n = o_p(nb^{p/2}), \quad \iiint_D R_n^2 = o_p(nb^{p/2}).$$

Hence Theorem 2.2 follows from (3.11) and Theorem 3.1. ∎

Remark 3.7. We have assumed throughout that $nb^{p+4} \to 0$. The cases $nb^{p+4} \to c$ and $nb^{p+4} \to \infty$ can also be treated with an asymptotic normal result. However, in these cases the terms T_{i2} and T_{i3} ($i = 1, 2, 3, 4$) are the ones determining the asymptotic behavior of $[W_n^{(1)}]^2$ (cf., e.g., Lemmas 3.2 and 3.3). The resulting asymptotic variance would be too complicated for this approach to be practical.

Proof of Theorem 2.4. The proof follows as in the proof of Theorem 2.2 above. We can replace ξ_n by

$$\xi_n^{(1)} = [X_n^{(1)}(t_1, z_1, w_1), \ldots, X_n^{(1)}(t_k, z_k, w_k)],$$

that is,

$$(nb^p)^{1/2} \|\xi_n - \xi_n^{(1)}\| \xrightarrow[P]{} 0.$$

The vector $\xi_n^{(1)}$ is a sum of independent random vectors with zero mean and covariance matrix $(nb^p)^{-1}\Sigma+o((nb^p)^{-1})$. Note that

$$(nb^p)^{1/2}r(t, z)[g_n(t, z)-g(t, z)] \xrightarrow[P]{} 0.$$

On applying a central limit theorem the theorem is proved. ∎

4. The censored case. Suppose that the survival times $T_1, T_2, \ldots, T_n$ of n individuals are subject to random censoring by the random variables $C_1, C_2, \ldots, C_n$, which are assumed to be independent. Moreover, T_i and C_i are assumed to be conditionally independent given the covariate vector Z_i (cf. Tsiatis [16]). The observable time until death will be denoted by $Y_i = \min\{T_i, C_i\}$ and let $\delta_i = I\{Y_i = T_i\}$, $i = 1, 2, \ldots, n$.

Let F^* denote a joint "survival" function of Y and Z,

$$F^*(t, z) = P\{Y \geqslant t, Z \leqslant z\},$$

where $0 \leqslant t \leqslant Q$ and $z \in M \subset R^{p-1}$ (cf. Condition 2.1). Let

$$g^*(t, z) = (\partial^{p-1}/\partial z_1, \ldots, \partial z_{p-1}]F^*(t, z). \tag{4.1}$$

Then, if f_Z is the marginal density of Z and if

$$F(t\,|\,z) = P\{T_i > t \mid Z_i = z\}, \quad G(t\,|\,z) = P\{C_i > t \mid Z_i = z\},$$

we have

$$g^*(t, z) = f_Z(z)F(t\,|\,z)G(t\,|\,z)$$

by the conditional independence of T_i and C_i, given Z_i. Also,

$$\lambda(t, z) = \tilde{f}(t, z)[g^*(t, z)]^{-1}, \tag{4.2}$$

where

$$\tilde{f}(t, z) = -(\partial^p/\partial t, \partial z_1, \ldots, \partial z_{p-1})P\{Y_i > t, \delta_i = 1, Z_i \leqslant z\}$$
$$= f(t, z)G(t\,|\,z) \tag{4.3}$$

is the joint subdensity of Y_i and Z_i with $Y_i = T_i$ (uncensored), and f is the joint density of (T_i, Z_i).

To proceed with our test of H_0 of (1.2) in this random censorship case, we estimate λ by

$$\lambda_n^*(t, z) = \tilde{f}_n(t, z)[g_n^*(t, z)]^{-1},$$

where

$$\tilde{f}_n(t, z) = -b^{-p}\int K(b^{-1}[(t, z)-(u, v)])d\tilde{F}_n(u, v),$$
$$g_n^*(t, z) = b^{-(p-1)}\int K_2(b^{-1}[z-v])d_v F_n^*(t, v) \tag{4.4}$$

and

$$\tilde{F}_n(u, v) = n^{-1} \sum_{i=1}^{n} I\{Y_i > u, Z_i \leqslant v, \delta_i = 1\},$$

$$F_n^*(u, v) = n^{-1} \sum_{i=1}^{n} I\{Y_i > u, Z_i \leqslant v\}.$$

Note that both $\tilde{F}_n$ and F_n^* are based on the observed data (Y_i, Z_i, δ_i), $i = 1, 2, \ldots, n$.

Let $\hat{\beta}_n$ denote the Cox estimator for the $(p-1)$-vector β (cf. Tsiatis [16]). We arrive at the process corresponding to (1.7):

$$X_n^*(t, z, w) = \tilde{f}_n(t, z)[g_n^*(t, z)]^{-1} \exp\{-\hat{\beta}z\} - \tilde{f}_n(t, w)[g_n^*(t, w)]^{-1} \exp\{-\hat{\beta}w\} \tag{4.5}$$

and to the statistic corresponding to (1.8):

$$(W_n^*)^2 = \iiint_D [X_n^*(t, z, w)]^2 \, dt dz dw, \tag{4.6}$$

where $\tilde{f}_n$ and g_n^* are defined by (4.4).

We have

THEOREM 4.1. *Assume Conditions 2.1 hold with f and g replaced by $\tilde{f}$ and g^*, respectively. Then the conclusions of Theorems 2.2 and 2.4 and Remark 2.3 hold for $(W_n^*)^2$ and X_n^* with f and g replaced by $\tilde{f}$ and g^*, respectively.*

The proof of Theorem 4.1 follows from the arguments in Section 3 and on noting (4.2).

REFERENCES

[1] P. K. Andersen, *Testing goodness of fit of Cox's regression and life model*, Biometrics 38 (1982), pp. 67–77.

[2] — and O. Borgan, *Counting process model for life history data: a review*, Scand. J. Statist. 12 (1985), pp. 97–118.

[3] T. Bednarski, *On sensitivity of Cox's estimator*, Statist. Decisions 7 (1989), pp. 215–228.

[4] M. D. Burke, M. Csörgö, S. Csörgö and P. Révész, *Approximations of the empirical process when parameters are estimated*, Ann. Probab. 7 (1979), pp. 790–810.

[5] M. D. Burke and E. Gombay, *A test of fit for Cox's regression model*, Res. Paper 678 (1989), Math. Statist. Dept., University of Calgary.

[6] D. R. Cox, *Regression models and life tables*, J. Roy. Statist. Soc. Ser. B 34 (1972), pp. 187–226.

[7] — *Partial likelihood*, Biometrika 62 (1975), pp. 269–276.

[8] J. Durbin, *Weak convergence of the sample df when parameters are estimated*, Ann. Statist. 1 (1973), pp. 279–290.

[9] P. Hall, *Central limit theorem for integrated square error of multivariate nonparametric density estimators*, J. Multivariate Anal. 14 (1984), pp. 1–16.

[10] L. Horváth, *On L_p-norms of multivariate density estimators*, Ann. Statist. 19 (1991), pp. 1933–1949.

[11] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, Wiley, New York 1980.

[12] R. Kay, *Proportional hazards regression models and the analysis of censored survival data*, Appl. Statist. 26 (1977), pp. 227–237.

[13] B. L. S. Prakasa Rao, *Nonparametric Functional Estimation*, Academic Press, New York 1983.

[14] D. Schoenfeld, *Chi-squared goodness-of-fit tests for the proportional hazards regression model*, Biometrika 67 (1980), pp. 145–153.

[15] — *Partial residuals for the proportional hazards regression model*, ibidem 69 (1982), pp. 239–241.

[16] A. A. Tsiatis, *A large sample study of Cox's regression model*, Ann. Statist. 9 (1981), pp. 93–108.

Department of Mathematics
and Statistics
University of Calgary
Calgary, Alberta, Canada
T2N 1N4

Department of Statistics
and Applied Probability
University of Alberta
Edmonton, Alberta, Canada
T6G 2G1

Received on 15.6.1991;
revised version on 8.1.1991