AD-A202 440

DTIC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 23699.21-MA	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Goodness-of-Fit Tests For Additive Hazards And Proportional Hazards Models		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ian W. McKeague and Klaus J. Utikal		8. CONTRACT OR GRANT NUMBER(s) DAAL03-86-K-0094
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University, Department of Statistics, Tallahassee, FL 32306-3033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office, Post Office Box 12211, Research Triangle Park, NC 27709		12. REPORT DATE October, 1988
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

N/A

DTIC SELECTE DEC 0 6 1988 S D E

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Censored survival data, counting processes, martingale methods.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Goodness-of-fit tests for Cox's proportional hazards model and Aalen's additive risk model, in which each model is compared on an equal footing with the best fitting fully nonparametric model, are developed. The goodness-of-fit statistics are based on differences between estimates of the doubly cumulative hazard function $A(t,z) = \int_0^t \int_0^z \lambda(s,x)\,ds\,dx$, under each model, with a fully nonparametric estimator of A recently introduced by the authors. Here $\lambda(\cdot,z)$ denotes the conditional hazard function of the survival time of an individual with covariate vector z. Comparison of the results of the tests makes it possible to decide whether Cox's proportional hazards or Aalen's additive risk model gives a better fit to the data. In addition, a goodness-of-fit test for Cox's model within the family of all proportional hazards models $\lambda(t,z) = \lambda_0(t)r(z)$, where λ_0 is a baseline hazard function and r is a general relative risk function, is developed.

88 12

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARO 23699.21-MA

GOODNESS-OF-FIT TESTS FOR ADDITIVE HAZARDS AND PROPORTIONAL HAZARDS MODELS

by

Ian W. McKeague[1] and Klaus J. Utikal

FSU Technical Report No. M-793
USARO Technical Report No. D-105

October, 1988

Department of Statistics
The Florida State University
Tallahassee, FL 32306-3033

Department of Statistics
University of Kentucky
Lexington, KY 40506

[1] Research supported by the Army Research Office under Grant DAAL03-86-K-0094.

AMS 1980 subject classifications. 62G10, 62M99, 62P10.

Key words and phrases. Censored survival data, counting processes, martingale methods.

GOODNESS-OF-FIT TESTS FOR ADDITIVE HAZARDS AND PROPORTIONAL HAZARDS MODELS

Ian W. McKeague
The Florida State University

and

Klaus J. Utikal
University of Kentucky

Abstract

Goodness-of-fit tests for Cox's proportional hazards model and Aalen's additive risk model, in which each model is compared on an equal footing with the best fitting fully nonparametric model, are developed. The goodness-of-fit statistics are based on differences between estimates of the doubly cumulative hazard function $\mathcal{A}(t,z) = \int_0^z \int_0^t \lambda(s,x)\,ds\,dx$, under each model, with a fully nonparametric estimator of $\mathcal{A}$ recently introduced by the authors. Here $\lambda(\cdot,z)$ denotes the conditional hazard function of the survival time of an individual with covariate vector z. Comparison of the results of the tests makes it possible to decide whether Cox's proportional hazards or Aalen's additive risk model gives a better fit to the data. In addition, a goodness-of-fit test for Cox's model within the family of all proportional hazards models $\lambda(t,z) = \lambda_0(t)\,r(z)$, where λ_0 is a baseline hazard function and r is a general relative risk function, is developed.

1. Introduction

Additive hazards and proportional hazards regression models used in the analysis of censored survival data can give substantially different results. For instance, in connection with a study of cancer mortality among Japanese atomic bomb survivors, Muirhead and Darby (1987) have noted that the two models give substantially different estimates of the age-specific probability that an individual will develop radiation induced cancer. Muirhead and Darby ~~(see also Aranda-Ordaz, 1983)~~ introduced a generalized parametric model which contains parametric additive hazards and proportional hazards models as special cases. The goodness-of-fit of each model is then obtained by comparing with the best fitting model within the generalized family, allowing the two special models to be treated on an equal footing.

Beyond the parametric setting, much effort has been devoted to the development of goodness-of-fit tests for Cox's (1972) proportional hazards model

$$\lambda(t,z) = \lambda_0(t)\, e^{\beta_0' z}, \tag{1.1}$$

where $\lambda(t,z) = \lambda(t|z)$ is the conditional hazard function of the survival time of an individual who has a covariate vector $z = (z_1, \ldots, z_p)'$, say, at time t, β_0 is a vector of p parameters and λ_0 is an arbitrary baseline hazard function. Many papers deal with graphical methods and their interpretation—see Arjas (1988) for a recent contribution. These methods are useful for detecting and diagnosing possible departures from the model, but their interpretation in the absence of formal significance tests is largely subjective. A formal test for Cox's model was introduced by Schoenfeld (1980) who developed a test of Cox's model against the fully nonparametric alternative that $\lambda(t,z)$ is arbitrary by using a chi-squared test based on the observed and expected frequencies that data points fall into cells that partition the product of the time and covariate state spaces. Using a similar approach, Andersen (1982) introduced a test of whether the inclusion of a new covariate z_{p+1} gives rise to a Cox model when $(z_1, \ldots, z_p)'$ already does. Moreau, O'Quigley and Mesbah (1985) considered testing whether β_0 varies with time.

Recently, Aalen (1988) discussed some graphical methods for examining the goodness-of-fit of the additive risk model (Aalen, 1980)

$$\lambda(t,z) = \sum_{j=1}^{p} \alpha_j(t)\, z_j, \tag{1.2}$$

where $\alpha_1, \ldots, \alpha_p$ are arbitrary functions of time.

The purpose of ~~the present~~ paper is to develop formal goodness-of-fit tests for the models of Aalen and Cox in which each model is compared on an equal footing with the best fitting fully nonparametric model. Our test statistics are based on differences between estimates of the doubly cumulative hazard function $\mathcal{A}(t,z) = \int_0^z \int_0^t \lambda(s,x)\, ds\, dx$, under each model, with a nonparametric estimator $\tilde{\mathcal{A}}$ of $\mathcal{A}$ introduced by McKeague and Utikal (1988). Comparison of the results of our tests makes it possible to decide whether Cox's proportional hazards or Aalen's additive risk model gives a better fit to the data.

Goodness-of-fit statistics based on a comparison of estimates of cumulative hazard functions, allowing the application of powerful counting process and martingale techniques, have been previously studied by Hjort (1984). He constructed tests for the hypothesis that the baseline hazard function in Cox's model follows a given parametric form, where the relative risk function $r(z) = e^{\beta_0' z}$ is assumed to be correctly specified. Hjort obtained a weak convergence result for the difference between nonparametric and parametric estimates of the cumulative baseline hazard function, and used his result to construct a chi-squared statistic based on a division of the time domain into cells. We have extended this approach to the full Cox model.

The general model (in which $\lambda(t,z)$ is fully nonparametric) is described in Section 2. Our goodness-of-fit tests for the models of Cox and Aalen are presented in Sections 3 and 5 respectively. In the case of Cox's model, with $p = 1$, we compare $\tilde{\mathcal{A}}$ with the semiparametric estimator $\hat{\mathcal{A}}$ of $\mathcal{A}$ given by

$$\hat{\mathcal{A}}(t,z) = \hat{\Lambda}(t)\int_0^z e^{\hat{\beta}x}dx, \quad (t,z) \in [0,1]^2 \tag{1.3}$$

where $\hat{\Lambda}$ is an estimator of the baseline hazard function and $\hat{\beta}$ is Cox's maximum partial likelihood estimator. Under Cox's model the two estimators should be close to one another. We show that $\sqrt{n}(\tilde{\mathcal{A}} - \hat{\mathcal{A}})$ converges weakly to a certain Gaussian random field which is represented as a sum of stochastic integrals with respect to a Brownian sheet process. This result leads to the construction of our goodness-of-fit test for Cox's model against the general alternative. In Section 4 we develop a test for Cox's model against the more restrictive alternative of general proportional hazards: $\lambda(t,z) = \lambda_0(t)\,r(z)$, where r is an unspecified relative risk function.

Proofs of all our results are collected in Section 6.

2. The general model

Let $N(t) = (N_1(t),\ldots,N_n(t))'$, $t \in [0,1]$, be a multivariate counting process with respect to a right-continuous filtration $(\mathcal{F}_t)$, i.e. N is adapted to the filtration and has components N_i which are right-continuous step functions, zero at time zero, with jumps of size $+1$ such that no two components jump simultaneously. Here $N_i(t)$ records the number of observed failures (0 or 1) in $[0,1]$ for the ith individual. Suppose that N_i has intensity of the general form

$$\lambda_i(t) = Y_i(t)\,\lambda(t, Z_i(t)), \quad i = 1,\ldots,n$$

where $Y_i(t)$ is a predictable $\{0,1\}$-valued process, indicating that the ith individual is at risk when $Y_i(t) = 1$, and $Z_i(t)$ is a predictable covariate process. The function $\lambda(t,z)$ represents the failure rate for an individual at risk at time t with covariate $Z_i(t) = z$. We assume throughout that (N_i, Y_i, Z_i), $i = 1,\ldots,n$ are i.i.d. replicates of (N,Y,Z) and Z is scalar valued. Note that (see Andersen and Borgan, 1985) the processes

$$M_i(t) = N_i(t) - \int_0^t Y_i(s)\,\lambda(s, Z_i(s))\,ds, \quad i = 1,\ldots,n \tag{2.1}$$

are orthogonal local square integrable martingales with predictable variation process

$$\langle M_i, M_i\rangle_t = \int_0^t Y_i(s)\,\lambda(s, Z_i(s))\,ds, \quad i = 1,\ldots,n.$$

The goodness-of-fit tests introduced in this paper involve a certain estimator for $\mathcal{A}(t,z) = \int_0^z \int_0^t \lambda(s,x)\,ds\,dx$ over the unit square $[0,1]^2$. Let $x_r = r/d_n$ and $\mathcal{I}_r = [x_{r-1}, x_r)$ for $r = 1, \ldots, d_n$, where d_n is an increasing sequence of positive integers. Let $N_{ir}(t)$ be the counting process which registers the jumps of $N_i(t)$ when $Z_i(t) \in \mathcal{I}_r$, so that

$$N_{ir}(t) = \int_0^t I(Z_i(s) \in \mathcal{I}_r)\,dN_i(s). \tag{2.2}$$

Beran (1981) suggested that the cumulative conditional hazard function $\Lambda(t,z) = \int_0^t \lambda(s,z)\,ds$ could be estimated by the Nelson-Aalen type estimator

$$\tilde{\Lambda}(t,z) = \int_0^t \frac{1}{Y_r^{(n)}(s)}\,dN_r^{(n)}(s), \quad \text{for } z \in \mathcal{I}_r$$

where

$$Y_r^{(n)}(s) = \sum_{i=1}^n I(Z_i(s) \in \mathcal{I}_r)\,Y_i(s)$$

and $N_r^{(n)} = \sum_{i=1}^n N_{ir}$. McKeague and Utikal (1988), subsequently referred to as MU, proposed the following estimator for $\mathcal{A}$

$$\tilde{\mathcal{A}}(t,z) = \int_0^z \tilde{\Lambda}(t,x)\,dx,$$

and they obtained a weak convergence result for $\tilde{\mathcal{A}}$.

Before stating that result we need to introduce some more notation and some conditions on Y and Z. Let $\int_0^t \int_0^z \phi(s,x)\,dW(s,x)$ denote a continuous version of the Wiener integral of a function $\phi \in L^2([0,1]^2, ds\,dx)$ with respect to a Brownian sheet W, see Wong and Zakai (1974). Suppose that for each $t \in [0,1]$, the random vector (Z_t, Y_t) is absolutely continuous with respect to the product of Lebesgue measure on $[0,1]$ and counting measure, and denote the corresponding density by $f_{Z(t)\,Y(t)}(z,y)$. Also, assume that $f_{Z(t)\,Y(t)}(z,1)$ is a positive, continuous function of $(t,z) \in [0,1]^2$. Let D_2 denote the extension of Skorohod space $D[0,1]$ to functions on $[0,1]^2$, as defined in Neuhaus (1971), and let $D[0,1]^p$ denote the product of p copies of $D[0,1]$.

In the present setting we may state Theorem 3.1 of MU as follows.

Proposition 3.1. *Suppose that λ is Lipschitz, $d_n^2/n \to \infty$ and $d_n = o(n^\delta)$ for some $\delta \in (\frac{1}{2}, 1)$. Then $\sqrt{n}(\tilde{\mathcal{A}} - \mathcal{A}) \xrightarrow{\mathcal{D}} m$ in D_2 as $n \to \infty$, where $m = (m(t,z),\ (t,z) \in [0,1]^2)$ is given by*

$$m(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\,dW(s,x), \tag{2.3}$$

$$h(s,x) = \frac{\lambda(s,x)}{f_{Z(s)\,Y(s)}(x,1)}.$$

In the sequel we shall denote $H(t,z) = \int_0^t \int_0^z h(u,x)\,dx\,du$ and denote the corresponding measure on $[0,1]^2$ by H as well.

3. Cox's proportional hazards model vs. the general model

Inference for β_0 in (1.1) can be based on the partial likelihood function

$$L(\beta) = \prod_{i=1}^{n} \left\{ \frac{e^{\beta Z_i(T_i)}}{\sum_{j \in \mathcal{R}_i} e^{\beta Z_j(T_i)}} \right\}^{\delta_i}, \tag{3.1}$$

where δ_i and T_i are the indicator of noncensorship and the survival time for the ith individual respectively, and $\mathcal{R}_i$ is the risk set consisting of all individuals who are observed to be at risk at time T_i. This approach was proposed by Cox (1972, 1975). Let $\hat{\beta}$ be the value that maximizes $L(\beta)$ and estimate the cumulative baseline hazard $\Lambda_0(t) = \int_0^t \lambda_0(s)\, ds$ by the Breslow (1972, 1974) type estimator

$$\hat{\Lambda}(t) = \sum_{T_i \leq t} \frac{\delta_i}{\sum_{j \in \mathcal{R}_i} e^{\hat{\beta} Z_j(T_i)}}. \tag{3.2}$$

We are interested in testing the null hypothesis H_0: Cox's proportional hazards model (1.1) holds over the region $(t, z) \in [0,1]^2$. The natural estimator of $\mathcal{A}$ under H_0 is

$$\hat{\mathcal{A}}(t,z) = \hat{\Lambda}(t) \int_0^z e^{\hat{\beta} x} dx, \qquad (t,z) \in [0,1]^2$$

where, if $(T_i, Z_i(T_i))$ falls outside $[0,1]^2$, the survival time T_i is regarded as being censored (i.e. δ_i is set to 0). Introduce some notation (cf. Andersen and Gill, 1982):

$$S^{(j)}(\beta,t) = \frac{1}{n} \sum_{i=1}^{n} Z_i(t)^j Y_i(t)\, I(0 \leq Z_i(t) \leq 1)\, e^{\beta Z_i(t)},$$

$$s^{(j)}(\beta,t) = E\, S^{(j)}(\beta,t),$$

for $j = 0,1,2$, where $0^0 = 1$, and

$$e = s^{(1)}/s^{(0)}, \qquad v = s^{(2)}/s^{(0)} - e^2,$$

$$\Sigma = \int_0^t v(\beta,t)\, s^{(0)}(\beta_0,t)\, \lambda_0(t)\, dt.$$

Formulated in terms of the counting processes, the estimate $\hat{\beta}$ is the unique solution to $\frac{\partial}{\partial \beta} \log L(\beta) = U(\beta,1) = 0$, where

$$U(\beta,t) = \sum_{i=1}^{n} \int_0^t \left\{ Z_i(u) - \frac{S^{(1)}(\beta,u)}{S^{(0)}(\beta,u)} \right\} I(0 \leq Z_i(u) \leq 1)\, dN_i(u), \tag{3.3}$$

and the estimator (3.2) is given by

$$\hat{\Lambda}(t) = \int_0^t \frac{d\bar{N}(u)}{nS^{(0)}(\hat{\beta},u)},$$

where $\bar{N} = \sum_{r=1}^{d_n} N_r^{(n)}$, and $N_r^{(n)}$ is defined in Section 2.

Theorem 3.1. *Suppose that Y and Z are left-continuous with right hand limits, Σ is positive, λ_0 is Lipschitz, $d_n^2/n \to \infty$ and $d_n = o(n^\delta)$ for some $\delta \in (\frac{1}{2}, 1)$. Then, under Cox's proportional hazards model (1.1), $\sqrt{n}(\tilde{A} - \hat{A}) \xrightarrow{\mathcal{D}} m'$ in D_2 as $n \to \infty$, where*

$$\begin{aligned} m'(t,z) &= \int_0^t \int_0^z \sqrt{h(u,x)}\, dW(u,x) - b(z) \int_0^t \int_0^1 \frac{\sqrt{g(u,x)}}{s^{(0)}(\beta_0,u)}\, dW(u,x) \\ &\quad - c(t,z) \int_0^1 \int_0^1 \left\{ x - \frac{s^{(1)}(\beta_0,u)}{s^{(0)}(\beta_0,u)} \right\} \sqrt{g(u,x)}\, dW(u,x), \\ h(u,x) &= \frac{\lambda_0(u)\, e^{\beta_0 x}}{f_{Z(u)\,Y(u)}(x,1)}, \\ g(u,x) &= \lambda_0(u)\, e^{\beta_0 x} f_{Z(u)\,Y(u)}(x,1), \\ b(z) &= \int_0^z e^{\beta_0 x} dx, \\ c(t,z) &= \Sigma^{-1}\Big(\Lambda_0(t) \int_0^z x e^{\beta_0 x} dx - b(z) \int_0^t e(\beta_0,u)\, \lambda_0(u)\, du \Big). \end{aligned}$$

In order to test H_0 against the alternative that λ has the general form of Section 2 we might consider using statistics of Kolmogorov–Smirnov type or Cramér–von Mises type:

$$\sqrt{n} \sup_{(t,z)\in[0,1]^2} |\tilde{A}(t,z) - \hat{A}(t,z)| \qquad \text{or} \qquad \sqrt{n} \int_0^1 \int_0^1 (\tilde{A}(t,z) - \hat{A}(t,z))^2\, dt\, dz$$

which have asymptotic distributions $\sup_{(t,z)\in[0,1]^2} |m'(t,z)|$ and $\int_0^1 \int_0^1 (m'(t,z))^2\, dt\, dz$ respectively. However, general tables for these distributions are not available. MU suggested that critical values for such distributions be obtained by simulation of the process m'. A more feasible approach might be to bootstrap the estimators $\tilde{A}$ and $\hat{A}$ in some way [cf. the papers of Akritas (1986), Horváth and Yandell (1987) and Lo and Singh (1986) on the bootstrapped Kaplan–Meier estimator], but we shall not pursue that possibility here. Rather, our present approach [following Schoenfeld (1980)] is to derive a chi-squared test based on a partition of the product of the time and covariate state spaces into cells.

Let $0 = t_0 < \cdots < t_R = 1$ and $0 = z_0 < \cdots < z_L = 1$ and denote $\mathcal{T}_r = (t_{r-1}, t_r]$ and $\mathcal{Z}_l = (z_{l-1}, z_l]$ so that the cells $\mathcal{J}_{rl} = \mathcal{T}_r \times \mathcal{Z}_l$ partition $[0,1]^2$. The increment of $X \equiv \sqrt{n}(\tilde{A} - \hat{A})$ over $\mathcal{J}_{rl}$ is given by $Q_{rl}^{(n)} = X(\mathcal{J}_{rl}) = X(t_r, z_l) - X(t_r, z_{l-1}) - X(t_{r-1}, z_l) + X(t_{r-1}, z_{l-1})$. Under H_0 and the conditions of Theorem 3.1 we have that $Q^{(n)} = (Q_{rl}^{(n)},\ r = 1, \ldots, R;\ l = 1, \ldots, L)$ converges in distribution to the Gaussian random array $Q = (Q_{rl},\ r = 1, \ldots, R;\ \ l = 1, \ldots, L)$ with mean zero and covariance

$$\mathrm{Cov}(Q_{rl}, Q_{r'l'}) = H(\mathcal{J}_{rl} \cap \mathcal{J}_{r'l'}) - b(\mathcal{Z}_l)\, b(\mathcal{Z}_{l'}) \int_{\mathcal{T}_r \cap \mathcal{T}_{r'}} \frac{d\Lambda_0(u)}{s^{(0)}(\beta_0,u)} - c(\mathcal{J}_{rl})\, c(\mathcal{J}_{r'l'}) \Sigma,$$

where $b(\mathcal{Z}_l)$ and $c(\mathcal{J}_{rl})$ denote increments of b and c. A consistent estimator for this covariance can be obtained by inserting the usual estimates of β_0, Λ_0, $s^{(0)}$, Σ and $e(\beta_0, \cdot)$ in the last two terms above and estimating the first term by $\hat{H}(\mathcal{J}_{rl})$, where

$$\hat{H}(t,z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} e^{\hat{\beta} x_r} \int_0^t \frac{d\hat{\Lambda}_0(s)}{Y_r^{(n)}(s)}. \tag{3.4}$$

The estimator $\hat{H}$ is similar to the general estimator of H employed in MU (Section 3). Routine modifications to the proof of Lemma 9 of MU show that $\hat{H}$ is consistent.

If we write $Q^{(n)}$ and Q in the form of column vectors $U^{(n)}$ and U, respectively, (by stacking columns one on top of each other, say) and let $\hat{C}^{(n)}$ denote the corresponding estimate of the covariance matrix C of U, then our test statistic is given by

$$\hat{\Gamma}^{(n)} = U^{(n)\prime} \hat{C}^{(n)-1} U^{(n)}.$$

Under H_0 and the conditions of Theorem 3.1 we obtain that $\hat{\Gamma}^{(n)}$ has a limiting χ_q^2 distribution, where $q = \mathrm{rank}(C)$. Usually we would expect that C is of full rank, in which case $q = RL$.

4. Cox's model vs. general proportional hazards

The general proportional hazards model $\lambda(t,z) = \lambda_0(t)\, r(z)$, where r is an unknown relative risk function, was proposed by Thomas (1983). This model admits nonlinear covariate effects while preserving the proportional hazards form. Tibshirani (1984), Hastie and Tibshirani (1986) and O'Sullivan (1986a, 1986b) have studied various estimators for the log relative risk function $\log r(z)$. MU studied an estimator of the cumulative relative risk function $\int_0^{\cdot} r(x)\,dx$ and developed a goodness-of-fit test for the general proportional hazards model.

Now consider testing the Cox model null hypothesis H_0 of Section 3 vs. the alternative that the general proportional hazards model holds. Our test statistic compares two estimators of the normalized cumulative relative risk function

$$R(z) = \frac{\int_0^z r(x)\,dx}{\int_0^1 r(x)\,dx}.$$

Since $r(z) = e^{\beta_0 z}$ under Cox's model, the natural estimator of R under H_0 is

$$\hat{R}(z) = \frac{\int_0^z e^{\hat{\beta} x}\,dx}{\int_0^1 e^{\hat{\beta} x}\,dx}.$$

An estimator for R under the general proportional hazards model is

$$\tilde{R}(z) = \frac{\tilde{A}(1,z)}{\tilde{A}(1,1)}.$$

The following result gives the asymptotic distribution of $\sqrt{n}(\tilde{R} - \hat{R})$ under H_0.

Theorem 4.1. *Suppose that the conditions of Theorem 3.1 hold. Then, under Cox's proportional hazards model (3.1), $\sqrt{n}(\tilde{R} - \hat{R}) \xrightarrow{\mathcal{D}} \rho m''$ in $D[0,1]$ as $n \to \infty$, where $\rho = \left(\Lambda_0(1)\, b^2(1)\right)^{-1}$,*

$$m''(z) = b(1) \int_0^1 \int_0^z \sqrt{h(u,x)}\, dW(u,x) - b(z) \int_0^1 \int_0^1 \sqrt{h(u,x)}\, dW(u,x)$$
$$- \varphi(z) \int_0^1 \int_0^1 \left\{ x - \frac{s^{(1)}(\beta_0, u)}{s^{(0)}(\beta_0, u)} \right\} \sqrt{g(u,x)}\, dW(u,x),$$
$$\varphi(z) = \Sigma^{-1} \Lambda_0(1) \Big(b(1) \int_0^z x e^{\beta_0 x} dx - b(z) \int_0^1 x e^{\beta_0 x}\, dx \Big).$$

We shall omit the proof of this result since it is very similar to the proof of Theorem 3.1. A chi-squared goodness-of-fit test can be derived using Theorem 4.1, much as it was done using Theorem 3.1.

Let $0 = z_0 < \cdots < z_L = 1$ so that the intervals $\mathcal{Z}_l = (z_{l-1}, z_l]$ partition $[0,1]$. Let $X = \sqrt{n}(\tilde{R} - \hat{R})$ and let $Q_l^{(n)}$ denote the increment of X over $\mathcal{Z}_l$. Under H_0 and the conditions of Theorem 4.1 we have that $Q^{(n)} = (Q_l^{(n)},\ l = 1, \ldots, L)$ converges in distribution to the Gaussian random vector $Q = (Q_l,\ l = 1, \ldots, L)$ with mean zero and covariance

$$\mathrm{Cov}(Q_l, Q_{l'}) = \rho^2 \,[b(1)^2 H_1(\mathcal{Z}_l \cap \mathcal{Z}_{l'}) - b(1)\, b(\mathcal{Z}_l)\, H_1(\mathcal{Z}_{l'}) - b(1)\, b(\mathcal{Z}_{l'})\, H_1(\mathcal{Z}_l)$$
$$+ b(\mathcal{Z}_l)\, b(\mathcal{Z}_{l'})\, H(1,1) + 3\, \varphi(\mathcal{Z}_l)\, \varphi(\mathcal{Z}_{l'})\, \Sigma],$$

where $H_1(z) = H(1,z)$. This covariance can be estimated consistently by inserting the usual estimates of β_0, Λ_0 and Σ; and estimating H_1 by $\hat{H}_1(\cdot) = \hat{H}(1,\cdot)$, where $\hat{H}$ is given by (3.4).

5. Aalen's additive risk model vs. the general model

For simplicity, we shall only consider the following special case of Aalen's model (1.2):

$$\lambda(t,z) = \alpha_1(t) + \alpha_2(t)\, z, \tag{5.1}$$

where z is scalar valued, but our approach could easily be extended to the full Aalen model (1.2). Weighted least squares estimators $\tilde{A}_j$ for the functions $A_j(t) = \int_{t_0}^t \alpha_j(s)\, ds$, $t \in [t_0, 1]$, $j = 1,2$ were introduced by Huffer and McKeague (1987) and McKeague (1988a). Here t_0 is fixed, $0 < t_0 < 1$. The reason for restricting estimation to the time interval $[t_0, 1]$ is that it is not possible to estimate the correct weights uniformly over the whole of $[0,1]$, as required for the asymptotic theory developed in McKeague (1988a). In this section D_2 is restricted to functions on $[t_0, 1] \times [0,1]$.

The null hypothesis that we intend to test is H_0: Aalen's additive risk model (5.1) holds over the region $[t_0, 1] \times [0, 1]$. Define $\mathcal{A}$ and $\bar{\mathcal{A}}$ as in Section 2, except with the range of integration for the time variable going from t_0 to t. Since $\mathcal{A}(t,z) = z\,A_1(t) + \frac{1}{2}z^2 A_2(t)$ under H_0, a reasonable estimator for $\mathcal{A}$ under H_0 is

$$\tilde{\mathcal{A}}(t,z) = z\,\bar{A}_1(t) + \tfrac{1}{2}z^2\bar{A}_2(t).$$

Let $Y_{i1}(t) = Y_i(t)\,I(0 \le Z_i(t) \le 1)$ and $Y_{i2}(t) = Y_i(t)\,Z_i(t)\,I(0 \le Z_i(t) \le 1)$. Then the intensity of the counting process N_i is given by

$$\lambda_i(t) = \alpha_1(t)\,Y_{i1}(t) + \alpha_2(t)\,Y_{i2}(t)$$

under the additive risk model. Using similar notation to McKeague (1988a), the weighted least squares estimator $\bar{A}$ is defined by

$$\bar{A}(t) = \int_{t_0}^{t} Y^{-}(s)\,dN(s),$$

where $Y^{-}(s) = (Y'(s)\hat{W}(s)Y(s))^{-1}Y'(s)\hat{W}(s)$ and $Y(s) = (Y_{ij}(s))$ is the $n \times 2$ matrix of covariate processes, $\hat{W}(t)$ is the $n \times n$ diagonal matrix with ith diagonal entry $(\hat{\lambda}_i(t))^{-1}$,

$$\hat{\lambda}_i(t) = \hat{\alpha}_1(t)\,Y_{i1}(t) + \hat{\alpha}_2(t)\,Y_{i2}(t)$$

is an estimate of the intensity $\lambda_i(t)$, and $\hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2)'$ is the smoothed least squares estimator

$$\hat{\alpha}(t) = \frac{1}{b_n}\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,d\hat{A}(s).$$

Here $\hat{A} = (\hat{A}_1, \hat{A}_2)'$ is Aalen's least squares estimator

$$\hat{A}(t) = \int_0^t (Y'(s)Y(s))^{-1}Y(s)\,dN(s),$$

K is a left-continuous kernel function of bounded variation having integral 1, support $[0,1]$ and $b_n > 0$ is a bandwidth parameter. Let $L(t)$ and $V(t)$ denote the 2×2 matrices with entries $L_{jk}(t) = EY_{1j}(t)Y_{1k}(t)$, $V_{jk}(t) = EY_{1j}(t)Y_{1k}(t)\,\lambda_1^{-1}(t)$ respectively. Also, for any square matrix D, let D^{-1} denote the inverse of D if D is invertible, the zero matrix otherwise.

Theorem 5.1. *Suppose that the processes Y_t and Z_t are left-continuous with right hand limits, α_1 and α_2 are Lipschitz, the matrix functions $L(\cdot)$ and $V(\cdot)$ are continuous, $L(t)$ and $V(t)$ are nonsingular for all $t \in [0,1]$, $\inf_{(t,z)\in[0,1]^2}\lambda(t,z) > 0$, $b_n \to 0$, $nb_n^2 \to \infty$, $d_n^2/n \to \infty$ and $d_n = o(n^{\delta})$ for some $\delta \in (\frac{1}{2}, 1)$. Then, under Aalen's additive risk model (5.1), $\sqrt{n}(\hat{\mathcal{A}} - \tilde{\mathcal{A}}) \xrightarrow{\mathcal{D}} m'$ in D_2 as $n \to \infty$, where*

$$m'(t,z) = \int_0^t\int_0^z \sqrt{h(s,x)}\,dW(s,x) - z\int_0^t\int_0^1 \frac{[(V^{-1}(s))_{11} + x\,(V^{-1}(s))_{12}]}{\sqrt{h(s,x)}}\,dW(s,x)$$
$$- \tfrac{1}{2}z^2\int_0^t\int_0^1 \frac{[(V^{-1}(s))_{21} + x\,(V^{-1}(s))_{22}]}{\sqrt{h(s,x)}}\,dW(s,x),$$

$$h(s,x) = \frac{\alpha_1(s) + \alpha_2(s)\,x}{f_{Z(s)\,Y(s)}(x,1)}.$$

Define a chi-squared statistic $\hat{\Gamma}^{(n)}$ for testing H_0 in the same way that $\hat{\Gamma}^{(n)}$ was defined in Section 2, but with $X = \sqrt{n}(\tilde{\mathcal{A}} - \mathcal{A})$. Under H_0 and the conditions of Theorem 5.1 we have that $Q^{(n)} = (Q_{rl}^{(n)},\ r = 1,\ldots,R;\ l = 1,\ldots,L)$ converges in distribution to the Gaussian random array $Q = (Q_{rl},\ r = 1,\ldots,R;\ \ l = 1,\ldots,L)$ with mean zero and covariance

$$\begin{aligned}\mathrm{Cov}(Q_{rl}, Q_{r'l'}) &= H(\mathcal{J}_{rl} \cap \mathcal{J}_{r'l'}) \\ &- \Delta_l \Delta_{l'} \int_{\mathcal{I}_r \cap \mathcal{I}_{r'}} [(V^{-1}(s))_{11} + \tfrac{1}{2}(\Delta_l + \Delta_{l'})(V^{-1}(s))_{12} + \tfrac{1}{4}\Delta_l \Delta_{l'}(V^{-1}(s))_{22}]\, ds,\end{aligned}$$

where $\Delta_l = z_l - z_{l-1}$. A consistent estimator of this covariance can be obtained by estimating the first term by $\hat{H}(\mathcal{J}_{rl})$, where

$$\hat{H}(t,z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \int_0^t \frac{d\hat{A}_1(s) + x_r\, d\hat{A}_2(s)}{Y_r^{(n)}(s)},$$

and estimating the remaining terms using the following estimator (see McKeague, 1988a) of $\int_{t_0}^t (V^{-1}(s))_{jk}\, ds$:

$$n \sum_{i=1}^{n} \int_{t_0}^{t} (Y^{-}(s))_{ji} (Y^{-}(s))_{ki}\, dN_i(s).$$

A chi-squared statistic $\hat{\Gamma}^{(n)}$ for testing H_0 can then be developed as in Section 3.

6. Proofs

We shall make repeated use of the notation $(R_n,\ n \geq 1)$ for a generic sequence of processes which converge uniformly in probabilty to zero as $n \to \infty$. Also, the processes

$$\begin{aligned}M_r^{(n)}(t) &= \sum_{i=1}^{n} \int_0^t I\{Z_i(s) \in \mathcal{I}_r\}\, dM_i(s), \\ \tilde{M}^{(n)}(t,z) &= \sqrt{n} \sum_{r=1}^{d_n} \int_0^z \int_0^t \frac{1}{Y_r^{(n)}(s)}\, dM_r^{(n)}(s)\, I(x \in \mathcal{I}_r)\, dx, \\ \widetilde{M}^{(n)}(t,z) &= \frac{\sqrt{n}}{d_n} \sum_{r=1}^{[zd_n]} \int_0^t \frac{1}{Y_r^{(n)}(s)} dM_r^{(n)}(s),\end{aligned}$$

play an important role in the proofs.

Proof of Theorem 3.1. By the proof of Theorem 3.1 of MU we have (in our notation)

$$\sqrt{n}(\tilde{\mathcal{A}} - \mathcal{A}) = \tilde{M} + R_n. \tag{6.1}$$

The Lipschitz condition on λ required for that theorem to be applicable is satisfied for Cox's proportional hazards model (3.1), since λ_0 is assumed to be Lipschitz.

The next step in the proof is to decompose $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A})$ using the results of Andersen and Gill (1982), subsequently referred to as AG. Conditions A to D of AG can be checked under conditions of the present theorem (cf. the proof of Theorem 4.1 of AG). Then write

$$\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A})(t, z) = \sqrt{n}(\hat{\Lambda}(t) - \Lambda_0(t))\, b(z) + \sqrt{n} \int_0^z (e^{\hat{\beta}x} - e^{\beta_0 x})\, dx\, \hat{\Lambda}(t). \tag{6.2}$$

By AG (p.1104)

$$\sqrt{n}(\hat{\Lambda}(t) - \Lambda_0(t)) = \widehat{M}_0(t) - \sqrt{n}(\hat{\beta} - \beta_0) \int_0^t e(\beta_0, u)\, \lambda_0(u)\, du + R_n(t) \tag{6.3}$$

where

$$\widehat{M}_0(t) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^t \frac{1}{S^{(0)}(\beta_0, u)} I(0 \le Z_i(u) \le 1)\, dM_i(u).$$

Also, since $\hat{\beta}$ is a consistent estimator of β_0 and $\hat{\Lambda}$ is uniformly consistent estimator of Λ_0, by a Taylor series expansion the second term in (6.2) can be written as

$$\sqrt{n}(\hat{\beta} - \beta_0)\, \Lambda_0(t) \int_0^z x e^{\beta_0 x} dx + R_n(t, z). \tag{6.4}$$

By AG (proof of Theorem 3.2)

$$\sqrt{n}(\hat{\beta} - \beta_0) = n^{-\frac{1}{2}} U(\beta_0, 1)\, \Sigma^{-1} + o_P(1). \tag{6.5}$$

Note that by (3.3) and (2.1)

$$U(\beta_0, t) = \sum_{i=1}^{n} \int_0^t \Big\{ Z_i(u) - \frac{S^{(1)}(\beta_0, u)}{S^{(0)}(\beta_0, u)} \Big\} I(0 \le Z_i(u) \le 1)\, dM_i(u). \tag{6.6}$$

Let $x_r = r/d_n$, and introduce the martingale

$$\begin{aligned}
\widehat{M}_1(t) &\equiv n^{-\frac{1}{2}} \sum_{r=1}^{d_n} \int_0^t \Big\{ x_r - \frac{S^{(1)}(\beta_0, u)}{S^{(0)}(\beta_0, u)} \Big\}\, dM_r^{(n)}(u) \\
&= n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^t \Big\{ \sum_{r=1}^{d_n} x_r I(Z_i(u) \in \mathcal{I}_r) - \frac{S^{(1)}(\beta_0, u)}{S^{(0)}(\beta_0, u)} \Big\} I(1 \le Z_i(u) \le 1)\, dM_i(u).
\end{aligned}$$

By Doob's inequality and (6.6)

$$\begin{aligned}
E \sup_t |n^{-\frac{1}{2}} U(\beta_0, t) - \widehat{M}_1(t)|^2 &\le 4\, E \int_0^1 \Big\{ \sum_{r=1}^{d_n} (x_r - Z_1(u))\, I(Z_1(u) \in \mathcal{I}_r) \Big\}^2 d\langle M_1 \rangle_u \\
&= O\Big(\frac{1}{d_n^2}\Big) \to 0.
\end{aligned} \tag{6.7}$$

Thus, combining (6.1)-(6.5), we obtain the decomposition

$$\sqrt{n}(\tilde{A} - \hat{A})(t,z) = \tilde{M}(t,z) - b(z)\widehat{M}_0(t) - c(t,z)\widehat{M}_1(1) + R_n(t,z). \tag{6.8}$$

Set

$$\begin{aligned} m_0(t) &= \int_0^t \int_0^1 \frac{\sqrt{g(u,x)}}{s^{(0)}(\beta_0,u)}\, dW(u,x), \\ m_1(t) &= \int_0^t \int_0^1 \left\{ x - \frac{s^{(1)}(\beta_0,u)}{s^{(0)}(\beta_0,u)} \right\} \sqrt{g(u,x)}\, dW(u,x). \end{aligned} \tag{6.9}$$

Then m_0 and m_1 are independent zero mean Gaussian martingales with predictable variation processes

$$\langle m_0 \rangle_t = \int_0^t \frac{\lambda_0(u)}{s^{(0)}(\beta_0,u)}\, du,$$

$$\langle m_1 \rangle_t = \int_0^t v(\beta_0,u)\, s^{(0)}(\beta_0,u)\, \lambda_0(u)\, du.$$

Suppose that $(\widetilde{M}, \widehat{M}_0, \widehat{M}_1) \xrightarrow{\mathcal{D}} (m, m_0, m_1)$ jointly in $D' \equiv D_2 \times D[0,1]^2$. Define a map $\pi_n' : D' \to D'$ by $\pi_n'(f_1, f_2, f_3) = (\pi_n(f_1), f_2, f_3)$, where π_n is defined by $\pi_n(f)(t,x) = f(t, x_{r-1}) + d_n(x - x_{r-1}) f(t, x_r)$ for $x \in \mathcal{I}_r$, where $x_r = r/d_n$. Here $\pi_n(f)(t,\cdot)$ is a piecewise linear approximation to $f(t,\cdot)$ based on the points x_r, $r = 1, \ldots, d_n$, for each t. Note that $\tilde{M}^{(n)} = \pi_n(\widetilde{M}^{(n)})$. Also, appealing to a D' version of Lemma 4.1 of McKeague (1988b) we have $\pi_n'(\widetilde{M}, \widehat{M}_0, \widehat{M}_1) \xrightarrow{\mathcal{D}} (m, m_0, m_1)$ in D', where we have used the fact that m (defined by (2.3)), m_0 and m_1 have continuous sample paths. Thus $(\tilde{M}, \widehat{M}_0, \widehat{M}_1) \xrightarrow{\mathcal{D}} (m, m_0, m_1)$ jointly in D' and by the continuous mapping theorem and (6.8) we may conclude that

$$\sqrt{n}(\tilde{A} - \hat{A}) \xrightarrow{\mathcal{D}} m - b\, m_0 - c\, m_1(1) = m'.$$

It remains to show that $(\widetilde{M}, \widehat{M}_0, \widehat{M}_1) \xrightarrow{\mathcal{D}} (m, m_0, m_1)$ jointly in D'. By (6.7) and the proof of Theorem 3.4 of AG we have $(\widehat{M}_0, \widehat{M}_1) \xrightarrow{\mathcal{D}} (m_0, m_1)$ in $D[0,1]^2$. Also, by the proof of Theorem 3.1 of MU we have $\widetilde{M} \xrightarrow{\mathcal{D}} m$ in D_2. If we can show that the finite dimensional distributions of $(\widetilde{M}, \widehat{M}_0, \widehat{M}_1)$ converge to those of (m, m_0, m_1) then we are finished. It suffices to show that for any $0 \le z_0 < z_1 < \cdots < z_q \le 1$, $q \ge 1$,

$$\left((\widetilde{M}(\cdot, z_j) - \widetilde{M}(\cdot, z_{j-1}))_{j=1}^q,\ \widehat{M}_0(\cdot),\ \widehat{M}_1(\cdot) \right) \xrightarrow{\mathcal{D}} \left((m(\cdot, z_j) - m(\cdot, z_{j-1}))_{j=1}^q,\ m_0(\cdot),\ m_1(\cdot) \right)$$

in $D[0,1]^{q+2}$. This is done using Rebolledo's (1980) martingale central limit theorem. Since $M_r^{(n)}$, $r = 1, \ldots, d_n$ are orthogonal martingales and $\widehat{M}_0$ can be written in the form

$$\widehat{M}_0(t) = n^{-\frac{1}{2}} \sum_{r=1}^{d_n} \int_0^t \frac{1}{S^{(0)}(\beta_0,u)}\, dM_r^{(n)}(u),$$

we have (cf. Lemma 9 of MU)

$$\langle \widetilde{M}(\cdot,z),\, \widehat{M}_0(\cdot)\rangle_t = \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\int_0^t \frac{1}{Y_r^{(n)}(u)}\frac{1}{S^{(0)}(\beta_0,u)}d\langle M_r^{(n)}\rangle_u$$
$$\xrightarrow{P} \int_0^t\int_0^z \frac{e^{\beta_0 x}\lambda_0(u)}{s^{(0)}(\beta_0,u)}dx\,du = \langle m(\cdot,z),\, m_0(\cdot)\rangle_t.$$

Also, directly from the definitions of $\widetilde{M}$ and $\widehat{M}_1$

$$\langle \widetilde{M}(\cdot,z),\, \widehat{M}_1(\cdot)\rangle_t = \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\int_0^t \frac{1}{Y_r^{(n)}(u)}\left\{x_r - \frac{S^{(1)}(\beta_0,u)}{S^{(0)}(\beta_0,u)}\right\} d\langle M_r^{(n)}\rangle_u$$
$$\xrightarrow{P} \int_0^t\int_0^z \left\{x - \frac{s^{(1)}(\beta_0,u)}{s^{(0)}(\beta_0,u)}\right\} e^{\beta_0 x}\lambda_0(u)\,dx\,du = \langle m(\cdot,z),\, m_1(\cdot)\rangle_t.$$

Now apply the version of Rebolledo's central limit theorem given by AG (Theorem I.2) with $p = q+2$ and d_n playing the role of n. There are $q+2$ Lindeberg conditions to check. For the q components involving $\widetilde{M}(\cdot,z)$, these conditions follow from Lemma 6 of MU. The same approach works for the $\widehat{M}_1$ component, and the $\widehat{M}_0$ component is treated in AG (proof of Theorem 3.4). □

Proof of Theorem 5.1. First note that by the proof of Theorem 3.2 of McKeague (1988a) we can decompose $\sqrt{n}(\bar{\mathcal{A}} - \mathcal{A})$ as

$$\begin{aligned}\sqrt{n}(\bar{\mathcal{A}} - \mathcal{A})(t,z) &= z\sqrt{n}(\bar{A}_1 - A_1)(t) + \frac{1}{2}z^2\sqrt{n}(\bar{A}_2 - A_2)(t)\\ &= z\,\bar{M}_1(t) + \tfrac{1}{2}z^2\bar{M}_2(t) + R_n(t,z),\end{aligned} \tag{6.10}$$

where

$$\bar{M}_j(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \int_{t_0}^t \bar{G}_{ij}^{(n)}(s)\,dM_i(s),$$
$$\bar{G}_{ij}^{(n)}(s) = \begin{cases} G_{ij}^{(n)}(s) & \text{if } |G_{ij}^{(n)}(s)| \le C \\ C & \text{otherwise,}\end{cases}$$
$$G_{ij}^{(n)}(s) = \sum_{k=1}^2 (\hat{V}^{-1}(s))_{jk} Y_{ik}(s)\,\hat{\lambda}_i^{-1}(s),$$
$$\hat{V}(s) = (\hat{V}_{jk}(s)), \quad (\text{a } 2\times 2 \text{ matrix})$$
$$\hat{V}_{jk}(s) = \frac{1}{n}\sum_{i=1}^n Y_{ij}(s)\,Y_{ik}(s)\,\hat{\lambda}_i^{-1}(s)$$

and C is a positive constant such that

$$P\left(G_{ij}^{(n)}(s) = \bar{G}_{ij}^{(n)}(s) \text{ for all } i = 1,\ldots,n,\ s \in [t_0,1]\right) \to 1. \tag{6.11}$$

By (3.1) and (6.10) we obtain the following decomposition

$$\sqrt{n}(\hat{A} - \bar{A})(t,z) = \tilde{M}(t,z) - z\,\bar{M}_1(t) - \tfrac{1}{2}z^2\bar{M}_2(t) + R_n(t,z). \tag{6.12}$$

Set

$$m_j(t) = \int_{t_0}^{t}\int_0^1 \frac{[(V^{-1}(s))_{j1} + x\,(V^{-1}(s))_{j2}]}{\sqrt{h(s,x)}}dW(s,x), \qquad t_0 \le t \le 1$$

for $j = 1,2$. Then (m_1, m_2) is a bivariate Gaussian martingale with zero mean and, as routine calculations show, predictable covariation processes

$$\langle m_j, m_k\rangle_t = \int_{t_0}^{t}(V^{-1}(s))_{jk}ds, \qquad j,k = 1,2.$$

By the proof of Theorem 3.2 of McKeague (1988a) it follows that $(\bar{M}_1, \bar{M}_2)\xrightarrow{\mathcal{D}}(m_1, m_2)$ in $D[t_0,1]^2$. Also, by the proof of Theorem 3.1 of MU, we have $\widetilde{M}\xrightarrow{\mathcal{D}}m$ in D_2, where m is defined in the statement of Proposition 2.1. Thus, from the representation (6.12), to complete the proof (cf. the proof of Theorem 3.1) it suffices to show that for any $0 \le z_0 < z_1 < \cdots < z_q \le 1$, $q \ge 1$,

$$\left((\widetilde{M}(\cdot,z_j) - \widetilde{M}(\cdot,z_{j-1}))_{j=1}^q, \bar{M}_1(\cdot), \bar{M}_2(\cdot)\right)\xrightarrow{\mathcal{D}}\left((m(\cdot,z_j) - m(\cdot,z_{j-1}))_{j=1}^q, m_1(\cdot), m_2(\cdot)\right)$$

in $D[t_0,1]^{q+2}$. As in the proof of Theorem 3.1, we apply the version of Rebolledo's martingale central limit theorem given in AG (Theorem I.2) to do this. Note that $\bar{M}_j$, $j = 1,2$ are square integrable martingales and

$$\begin{aligned}
\langle \widetilde{M}(\cdot,z), \bar{M}_j(\cdot)\rangle_t &= \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\sum_{i=1}^{n}\int_{t_0}^{t}\frac{1}{Y_r^{(n)}(s)}\bar{G}_{ij}^{(n)}(s)\,d\langle M_i, M_r^{(n)}\rangle_s \\
&= \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\int_{t_0}^{t}\frac{1}{Y_r^{(n)}(s)}\Big[\sum_{i=1}^{n}G_{ij}^{(n)}(s)\,I(Z_i(s)\in\mathcal{I}_r)\lambda_i(s)\Big]\,ds + o_P(1) \\
&= \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\int_{t_0}^{t}[(\hat{V}^{-1}(s))_{j1} + x_r(\hat{V}^{-1}(s))_{j2}]\,ds + O_P(d_n^{-1}) + o_P(1) \\
&\xrightarrow{P}\int_0^z\int_{t_0}^{t}[(\hat{V}^{-1}(s))_{j1} + x\,(\hat{V}^{-1}(s))_{j2}]\,ds\,dx = \langle m(\cdot,z), m_j(\cdot)\rangle_t,
\end{aligned}$$

where we have used (6.11), Lemma 4.3 of McKeague (1988a) and the fact that $|z - x_r| < d_n^{-1}$, when $z \in \mathcal{I}_r$. Once again there are $q+2$ Lindeberg conditions to check. They have been checked for $\bar{M}_j$, $j = 1,2$ in the proof of Theorem 3.2 of McKeague (1988a), and for the q components involving $\widetilde{M}(\cdot,z)$ in MU (Lemma 6). $\square$

References

Aalen, O. O. (1980). A model for nonparametric regression analysis of counting processes. *Lecture Notes in Statistics* **2**. Springer-Verlag, New York.

Aalen O. O. (1988). A linear regression model for the analysis of life times. Tech. Report, Section of Medical Statistics, University of Oslo.

Akritas, M. G. (1986). Bootstrapping the Kaplan-Meier estimator. *J. Amer. Stat. Assoc.* **81** 1032–1038.

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

Andersen, P. K., Borgan, Ø. (1985). Counting process models for life history data: a review. *Scand. J. Statist.* **12** 97–158.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **10** 1100–1120.

Aranda-Ordaz, F. J. (1983). An extension of the proportional hazards model for grouped data. *Biometrics* **39** 109–117.

Arjas, E. (1988). A graphical method for assessing goodness-of-fit in Cox's proportional hazards model. *J. Amer. Statist. Assoc.* **83** 204–212.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Tech. Report, Dept. of Statistics, University of California, Berkeley.

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

Cox, D. R. (1975). Partial likelihood. *Biometrika* **62** 269–276.

Hastie, T. J. and Tibshirani, R. J. (1986). Generalized additive models (with discussion). *Stat. Sci.* **1** 297–310.

Hjort, N. L. (1984). Weak convergence of cumulative intensity process when parameters are estimated, with applications to goodness-of-fit tests in models with censoring. Tech. Report, Norwegian Computing Center.

Horváth, L. and Yandell, B. S. (1987). Convergence rates for the bootstrapped product-limit process. *Ann. Statist.* **15** 1155–1173.

Huffer, F. W. and McKeague, I. W. (1988). Survival analysis using additive risk models. Tech. Report, Dept. of Statistics, Florida State University, Tallahassee.

Lo, S.-H. and Singh, K. (1986). The product-limit estimator and the bootstrap: Some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.

McKeague, I. W. (1988a). Asymptotic theory for weighted least squares estimators in Aalen's additive risk model. To appear in the Proceedings of the Conference on Inference from Stochastic Processes, Cornell University, August, 1987 (ed., N. U. Prabhu), *Contemporary Mathematics.*

McKeague, I. W. (1988b). A counting process approach to the regression analysis of grouped survival data. *Stoch. Process. Appl.* **28** 221–239.

McKeague, I. W. and Utikal, K. J. (1988). Identifying nonlinear covariate effects in semimartingale regression models. Tech. Report, Dept. of Statistics, Florida State University, Tallahassee.

Moreau, T., O'Quigley, J. and Mesbah, M. (1985). A goodness-of-fit statistic for the proportional hazards model. *Appl. Statist.* **34** 212–218.

Muirhead, C. R. and Darby S. C. (1987). Modelling the relative and absolute risks of radiation-induced cancers. *J. R. Statist. Soc. A* **150** 83–118.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

O'Sullivan, F. (1986a). Nonparametric estimation in the Cox proportional hazards model. Tech. Report, Dept. of Statistics, University of California, Berkeley.

O'Sullivan, F. (1986b). Relative risk estimation. Tech. Report, Dept. of Statistics, University of California, Berkeley.

Rebolledo, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. verw. Gebiete* **51** 269–286.

Schoenfeld, D. (1980). Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika* **67** 145–53.

Thomas, D. C. (1983). Non-parametric estimation and tests of fit for dose response relations. *Biometrics* **39** 263–268.

Tibshirani, R. J. (1984). Local likelihood estimation. Tech. Report and unpublished Ph.D. dissertation, Dept. of Statistics, Stanford University.

Wong, E. and Zakai, M. (1974). Martingales and stochastic integrals for processes with a multidimensional parameter. *Z. Wahrsch. verw. Gebiete* **51** 109–122.