Kybernetika

Petr Volf

A large sample study of nonparametric proportional hazard regression model

Kybernetika, Vol. 26 (1990), No. 5, 404--414

Persistent URL: http://dml.cz/dmlcz/124596

A LARGE SAMPLE STUDY OF NONPARAMETRIC PROPORTIONAL HAZARD REGRESSION MODEL

PETR VOLF

The proportional hazard model is considered for the regression of survival time on a covariate x. The cumulative hazard function generally acquires the form $L(t, x) = A(t) \,.\, B(x)$. We suggest a method for estimating the functions A, B, provided that the function A is normalized by the condition $\int_0^{\mathcal{T}} A(t) \, d\mu(t) = 1$. Then $B(x) = \int_0^{\mathcal{T}} L(t, x) \, d\mu(t)$. The cumulative hazards are estimated by a simple kernel procedure, meanwhile the estimator for the function $A(t)$ is obtained by the ML method. The resulting estimates are strongly consistent and most of them asymptotically normal.

1. INTRODUCTION AND NOTATION

During the past few years the methods for survival analysis with censored data have been studied intensively. In order to model the regression of such data, the proportional hazard model has been developed. Its parametrized form has been suggested by D. R. Cox and examined (under various conditions) by several authors, among them by Tsiatis [5]. In our paper, where the general nonparametric form of the model is considered, the cumulative hazard function of survival time is supposed to depend on a covariate x in the following factorisable way:

$$L(t, x) = A(t) \, B(x) \,. \tag{1}$$

In order to avoid ambiguity, one of the functions A, B has to be normalized by some condition. An attempt to identify such a model was made in [6], where the values of covariate were divided into several levels and the function A characterized an average cumulative hazard. In the present contribution, the funtion A will be estimated by the method developed by Tsiatis [5]. The result coincides with the nonparametric maximum-likelihood estimate. The function B may be regarded as a function describing the cumulative influence of given x through the domain of time, namely

$$B(x) = \int_0^{\mathcal{T}} L(t, x) \, d\mu(t), \tag{2}$$

with a finite measure μ chosen on $(\mathbb{R}_1, \mathscr{B}_1)$ and some upper bound for time, $\mathscr{T}$. Under proportional hazard model this relation corresponds to the normalizing condition $\int_0^{\mathscr{T}} A(t)\, \mathrm{d}\mu(t) = 1$. Our estimate of $A(t)$ will fulfil this condition asymptotically only.

Denote by the positive, independent random variables $Y(x_i)$ the survival times of N cases. They are censored from the right side by means of random variables $V(x_i)$, which are independent mutually as well as of $Y(x_i)$'s. Therefore we observe the realization of $T(x_i) = \min(Y(x_i), V(x_i))$ and $\delta(x_i) = I[Y(x_i) \leqq V(x_i)]$, $i = 1, 2, \ldots, N$. Denote the distribution functions of $Y(x)$ and $V(x)$ by $F(t, x)$, $G(t, x)$, their survival functions by $P(t, x) = 1 - F(t, x) = \exp\{-L(t, x)\}$, $Q(t, x) = 1 - G(t, x)$.

Let us imagine that the values of the regressor are realizations of a random variable X distributed with a density $h(x)$ on an interval $\mathscr{X} \subseteq \mathbb{R}_1$. The extension of the results to a vector valued set of covariates is possible and straightforward.

We shall examine the behaviour of the cumulative hazards up to some finite time $\mathscr{T}$ such that $P(\mathscr{T}, x)\, Q(\mathscr{T}, x) > 0$ for every $x \in \mathscr{X}$ and, naturally, $P(\mathscr{T}, x) < 1$.

Let us now establish other notation useful throughout the paper:

$$H(t, x) = \mathsf{P}\{T(x) \geqq t\} = P(t, x)\, Q(t, x)$$
$$R(t, x) = \mathsf{P}\{T(x) \geqq t\,, \quad \delta(x) = 1\}$$
$$R(t) \quad = \int_{\mathscr{X}} R(t, x)\, h(x)\, \mathrm{d}x$$

and in accordance with [5], for a function $g(x)$ continuous on $\mathscr{X}$

$$\mathsf{E}(g(x), t) = \int_{\mathscr{X}} g(x)\, H(t, x)\, h(x)\, \mathrm{d}x\,.$$

Whichever is the procedure for identification of the functions A, B, quality of results depends on the properties of the cumulative hazard functions $L(t, x)$ estimates. A method of estimation of $L(t, x)$ has been developed in [6], let us now remind the main results.

For a given point $x \in \mathscr{X}$ define its neighbourhood $O_{d_N}(x) = \{z \in \mathscr{X}: |z - x| \leqq d_N\}$. Denote by $M_N(x)$ the number of points x_i in $O_{d_N}(x)$. With increasing extent of sample N, the sequence d_N is chosen such that $\mathrm{d}_N \to 0$ and $Nd_N \to \infty$. Then $M_N(x)/(2Nd_N)$ tends to $h(x)$, almost surely. It follows from the theory of nonparametric estimation of density. The estimator of the cumulative hazard function at x is constructed standardly by the nonparametric ML method – cf. [2], but only from the realizations at the points $x_i \in O_{d_N}(x)$:

$$L_N(t, x) = 0 \quad \text{for} \quad t < \min\{T(x_i): x_i \in O_{d_N}(x)\}\,, \tag{3}$$
$$= \sum \frac{\delta(x_i)\, I[T(x_i) \leqq t]}{M_N(T(x_i), x)}\, I[x_i \in O_{d_N}(x)] \quad \text{otherwise}\,.$$

There $M_N(s, x) = \sum I[T(x_j) \geqq s]\,.\,I[x_j \in O_{d_N}(x)]$ denotes the number of observations in $O_{d_N}(x)$ with the results no less than s.

2. ESTIMATION OF PROPORTION OF HAZARDS

Let us now formulate several assumptions:

A1. Function $B(x)$ is positive, bounded and continuous on $\mathscr{X}$, function $A(t)$ is nonegative, nondecreasing and continuous on $[0, \mathscr{T}]$.

A2. Distribution function $G(t, x)$ is continuous in both arguments.

A3. The distribution of the covariate random variable X possesses a continuous, positive and bounded density $h(x)$ on $\mathscr{X}$.

A4. Functions $B(x)$ and $G(t, x)$ are Lipschitz-continuous with regard to variable x on $\mathscr{X}$.

A5. Sequence d_N is such that $\lim_{N \to \infty} (N d_N^3) = 0$.

The third assumption states that $M_N(x)$ is proportional to $N d_N$, the last assumption means that $\sqrt{M_N(x)}\, d_N \to 0$, a.s.. When d_N is chosen proportionally to $N^{-\alpha}$, by A5 we demand $\alpha > \frac{1}{3}$.

In [6] for every $x \in \mathscr{X}$, strong consistency of $L_N(t, x)$ is proved under A1 – A3, uniformly in $t \in [0, \mathscr{T}]$. It means that in the sup norm

$$\sup_{t \in [0, \mathscr{T}]} |L_N(t, x) - L(t, x)| \to 0 \quad \text{a.s.}$$

Asymptotic normality is there proved, too, this time under A1 – A5: Random functions $Z_N(t, x) = \sqrt{[M_N(x)]}\,(L_N(t, x) - L(t, x))$ converge weakly to a Gaussian random function $Z(t, x)$, which has zero mean and the covariance function given for $0 \leqq s \leqq t \leqq \mathscr{T}$ by

$$\operatorname{cov}(Z(t, x), Z(s, x)) = C(s, x) = \int_0^s H^{-1} P^{-1}\, \mathrm{d}F .$$

These results together with relation (2) suggest immediately the estimate for function $B(x)$, namely

$$B_N(x) = \int_0^{\mathscr{T}} L_N(t, x)\, \mathrm{d}\mu(t) . \tag{4}$$

$L_N(t, x)$ is a nondecreasing, finite stepwise function of t, therefore (4) is well defined. Let $T_{k1} \leqq T_{k2} \leqq \ldots \leqq T_{kM}$, with $M = M_N(x)$, be the ordered realizations observed at points $x_i \in O_{d_N}(x)$, let T_{kl} be the greatest from them, less than $\mathscr{T}$. Then

$$B_N(x) = \sum_{i=1}^{l-1} L_N(T_{ki}, x) \int_{T_{ki}}^{T_{k(i+1)}} \mathrm{d}\mu(t) + L_N(T_{kl}, x) \int_{T_{kl}}^{\mathscr{T}} \mathrm{d}\mu(t) .$$

Theorem 1. Let assumptions A1 – A3 hold. Then $B_N(x)$ is a strongly consistent estimate of function $B(x)$ defined by (2).

Proof. According to our definition connected with (2), μ is a finite measure. The estimator $L_N(t, x)$ of $L(t, x)$ is a strongly consistent one, uniformly in $t \in [0, \mathscr{T}]$, as it has been demonstrated in [6]. Therefore

$$|B_N(x) - B(x)| \leqq \int_0^{\mathcal{T}} |L_N(t, x) - L(t, x)| \, \mathrm{d}\mu(t) \leqq$$
$$\leqq \sup_{t \leqq \mathcal{T}} |L_N(t, x) - L(t, x)| \, \mu[0, \mathcal{T}] ,$$

which tends to zero almost surely. □

Remark. Let us remind that $\sup_{t \leqq \mathcal{T}} |L_N(t, x) - L(t, x)| = \mathcal{O}\{(\ln M/M)^{1/2}\}$, confer for example the work of L. Rejtö [4].

Corollary 1. Let assumptions A1 – A3 hold and function $B(x)$ have a finite variation on $\mathcal{X}$. Then a strongly, uniformly in $x \in \mathcal{X}$, consistent variant of the estimate for $B(x)$ can be constructed.

Proof. Choose an arbitrary sequence of positive numbers $\varepsilon_1 > \varepsilon_2 > \ldots > \varepsilon_n \to 0$. For every ε_n interval $\mathcal{X}$ can be divided into a finite number (K_n) of disjoint intervals I_{nj}, $j = 1, \ldots, K_n$, such that the variation of $B(x)$ is less than $\varepsilon_n/2$ throughout I_{nj}. Choose one point z_{nj} in every interval I_{nj}. In accordance with Theorem 1 there exists *almost surely* N_n such that $N \geqq N_n$ implies

$$\sup \{|B_N(z_{nj}) - B(z_{nj})|, j = 1, \ldots, K_n\} \leqq \varepsilon_n/2 .$$

Evidently, for $m > n$ the points z can be chosen in such a way that $\{z_{nj}\} \subseteqq \{z_{mj}\}$.

For given N, n define now our desired estimate by the following relation:

$$B_{N,n}(x) = B_N(z_{nj}) \quad \text{for} \quad x \in I_{nj} . \tag{5}$$

Let us imagine that both N and n tend to infinity. Then

$$\sup_{x \in \mathcal{X}} |B_{N,n}(x) - B(x)| \leqq$$
$$\leqq \sup_j \{|B_N(z_{nj}) - B(z_{nj})| + \sup_{x \in I_{nj}} |B(z_{nj}) - B(x)|\} \leqq (\varepsilon_n/2) + (\varepsilon_n/2) = \varepsilon_n$$

for $N \geqq N_n$.

4. ASYMPTOTIC NORMALITY

Throughout this part our interest is concentrated on examination of asymptotic distribution of random variable $W_N(x) = \sqrt{(M_N(x))}\,(B_N(x) - B(x))$, for a fixed $x \in \mathcal{X}$. $W_N(x)$ is obtained by integration, $W_N(x) = \int_0^{\mathcal{T}} Z_N(t, x) \, \mathrm{d}\mu(t)$. Almost all trajectories of $Z_N(t, x)$ are bounded and integrable, therefore the integral may be understood as a result of integration of the trajectories. On the other hand, the limiting process $Z(t, x)$ is Gaussian, with zero mean and with finite and continuous covariance function, therefore $W(x) = \int_0^{\mathcal{T}} Z(t, x) \, \mathrm{d}\mu(t)$ is Gaussian again, $\mathsf{E}W(x) = 0$, var $W(x) = D(x)$, where

$$D(x) = \int_0^{\mathcal{T}} \int_0^{\mathcal{T}} \operatorname{cov}(Z(t, x), Z(s, x)) \, \mathrm{d}\mu(t) \, \mathrm{d}\mu(s) =$$
$$= \int_0^{\mathcal{T}} \int_0^{\mathcal{T}} C(\min(t, s), x) \, \mathrm{d}\mu(t) \, \mathrm{d}\mu(s) .$$

Theorem 2. Let assumptions A1 – A5 hold. Then random variables $W_N(x)$ converge weakly to the Gaussian random variable $W(x)$.

Proof. The definition and the distribution of $W(x)$ has been presented above. In [6] the weak convergence of the processes $Z_N(t, x)$ to $Z(t, x)$ has been proved provided that A1 – A5 hold. The step from it to the weak convergence of $W_N(x)$ to $W(x)$ can be accomplished, using the propositions from [3], Chap. IX. 1. □

Remark. Let x, y be two different points from $\mathscr{X}$. If N is sufficiently large, the neighbourhoods of x and of y are disjoint, $W_N(x)$ is then independent of $W_N(y)$. Therefore $W(x)$ and $W(y)$ are independent random variables, too.

Up to now, two asymptotic covariance functions have been derived, $C(t, x)$ and $D(x)$. The method of their estimation can be developed easily, their consistency is the subject of our following examination. Denote $H_N(t, x) = M_N(t, x)/M_N(x)$ the empirical estimate of $H(t, x)$.

$$P_N(t, x) = \prod_i \left\{\frac{M_N(T(x_i), x) - 1}{M_N(T(x_i), x)}\right\}^{\delta(x_i)\, I[T(x_i) \leqq t]\, I[x_i \in O_{d_N}(x)]}$$

$$\text{for} \quad t \geqq \min \{T(x_i)\colon x_i \in O_{d_N}(x)\}\,,$$

$$P_N(t, x) = 1 \quad \text{otherwise}\,,$$

is the product-limit estimate of $P(t, x)$. Both H_N and P_N are strongly consistent, uniformly in $t \in [0, \mathscr{T}]$ – cf. also [6]. Let us define

$$C_N(t, x) = \int_0^t \frac{-\mathrm{d}P_N}{P_N H_N} = \sum \frac{\delta(x_i)\, M_N(x)\, I[T(x_i) \leqq t]}{[M_N(T(x_i), x) - 1]^2}\, I[x_i \in O_{d_N}(x)]\,.$$

Sometimes the values in the denominator are replaced by $M_N(T(x_i), x)^2$, because they are always positive, the asymptotic result remains unchanged. The estimator of $D(x)$ can be expressed as

$$D_N(x) = \textstyle\int_0^{\mathscr{T}} \int_0^{\mathscr{T}} C_N(\min(t, s), x)\, \mathrm{d}\mu(t)\, \mathrm{d}\mu(s)\,.$$

Let us now remind two propositions from the Appendix of Tsiatis' work [5], in order to prove consistency of the derived estimates. The first proposition follows from the Glivenko-Cantelli lemma, the second one is a special case of Lemma 6.1 from Aalen [1].

Lemma T1. Let $g(x)$ be a continuous function on $\mathscr{X}$, such that $\mathsf{E}(g(x))^2$ is finite. Then

$$\sup_{0 \leqq t \leqq \mathscr{T}} |\mathsf{E}_N(g(x), t) - \mathsf{E}(g(x), t)| \to 0 \quad \text{a.s.}\,,$$

where $\mathsf{E}_N(g(x), t) = \sum g(x_i)\, I[T(x_i) \geqq t]/N$ is the empirical estimate of $\mathsf{E}(g(x), t)$.

Lemma T2. Let $Z_N(t)$, $Q_N(t)$ be random functions on $[0, \mathscr{T}]$ that converge almost surely in sup norm (uniformly as to t) to a continuous function $Z(t)$ and to a continuous (sub)survival function $Q(t)$ respectively. Let $f(z)$ be a continuous function such that $\mathrm{d}f/\mathrm{d}z$ exists and is continuous on the range space of $Z(t)$, $t \in [0, \mathscr{T}]$. Then

$$\sup_{0 \leqq t \leqq \mathscr{T}} \left| \int_0^t f(Z_N(s)) (-\mathrm{d}Q_N(s)) - \int_0^t f(Z(s)) (-\mathrm{d}Q(s)) \right| \to 0 \quad \text{a.s.}$$

Applying Lemma T2, we can prove the strong uniform consistency of $C_N(t, x)$. P corresponds to Q of Lemma P. H stands instead of function Z, $f(z) = 1/z$, x is a fixed point from $\mathscr{X}$.

Immediately the strong consistency of $D_N(x)$ may be stated, as

$$|D_N(x) - D(x)| \leqq \{\int_0^{\mathscr{T}} \mathrm{d}\mu(t)\}^2 \sup_{0 \leqq t \leqq \mathscr{T}} |C_N(t, x) - C(t, x)| .$$

5. ESTIMATION OF FUNCTION $A(t)$

Let us imagine that the function $B(x)$ is known. According to [5]

$$A(t) = \int_0^t - \mathrm{d}R(s)/\mathsf{E}[B(x), s] . \tag{6}$$

both $\mathsf{E}(g(x), t)$ and $R(s)$ are defined at the beginning of our contribution. The empirical estimate of $R(s)$ is $R_N(s) = \sum I[\delta(x_i) = 1]\, I[T(x_i) \geqq s]/N$, under A1 – A3 it is strongly uniformly consistent on $[0, \mathscr{T}]$. Due to Lemmas T1 and T2 the following holds:

Lemma 1. Under assumptions A1 – A3 the function

$$A_N(t) = 1/N \sum_i \delta(x_i)\, I[T(x_i) \leqq t]/\mathsf{E}_N(B(x), T(x_i)) \tag{7}$$

is a strongly consistent estimate of $A(t)$, with respect to the sup norm on $[0, \mathscr{T}]$.

Assumptions A1 – A3 guarantee that all demands of Lemmas T1, T2 are fulfilled, including the accurate definition of the integral (6). The asymptotic distribution of $A_N(t)$ could be derived in the same way as Breslow and Crowley [2] proved the asymptotic normality of the cumulative hazard estimate. However, from Tsiatis [5] the more straightforward proof follows. Tsiatis deals with parametrized model, where $B(x)$ is proportional to $\exp(\beta x)$. Estimate $\hat{\beta}$ of β is obtained by iterative Newton-Raphson procedure. Estimation of $A(t)$ follows, its result resembles our function $A_N(t)$, with $B(x) = \exp(\hat{\beta} x)$. Therefore the asymptotic distribution of $A_N(t)$ can be obtained directly as a special case, namely the case with β known. Properly modified results from [5] yield the following proposition.

Lemma 2. Let the assumptions A1 – A3 hold. Then the random functions $Z_N(t) = \sqrt{(N)}\,(A_N(t) - A(t))$ converge on $[0, \mathscr{T}]$ weakly to a Gaussian random function $Z(t)$, which has zero mean and the covariance function for $0 \leqq s \leqq t \leqq \mathscr{T}$ given by

$$\operatorname{cov}(Z(s), Z(t)) = C(s) = \int_0^s - \mathrm{d}R(u)/\mathsf{E}^2(B(x), u) .$$

Immediately the estimate of the asymptotic variance can be proposed:

$$C_N(s) = \sum \frac{N\,\delta(x_i)}{\mathsf{E}_N^2(B(x), T(x_i))}\, I[T(x_i) \leqq s]\,.$$

Its strong consistency in sup norm on $[0, \mathcal{T}]$ follows again from Lemmas T1 and T2.

Let $B_N^*(x)$ be an estimate of $B(x)$, strongly consistent with respect to the sup norm on $\mathcal{X}$. The example has been presented in Corollary 1. Let us construct the estimate of $A(t)$ by

$$A_N^*(t) = \sum_i \delta(x_i)\, I[T(x_i) \leqq t] / S_N(T(x_i))\,,$$

with $S_N(s) = \sum_i B_N^*(x_j)\, I[T(x_j) \geqq s]$.

Theorem 3. Let the assumptions A1 – A3 hold, then

$$\sup_{0 \leqq t \leqq \mathcal{T}} |A_N^*(t) - A(t)| \to 0 \quad \text{a.s.}$$

Proof. In order to apply Lemma T2, the almost sure, uniform convergence of $S_N(s)/N$ to $\mathsf{E}(B(x), s)$ remains to be demonstrated. However, the comparison with $\mathsf{E}_N(B(x), s)$ suffices. It yields:

$$\sup_{0 \leqq s \leqq \mathcal{T}} |S_N(s)/N - \mathsf{E}_N(B(x), s)| \leqq \sup_{0 \leqq s \leqq \mathcal{T}} 1/N \sum_i |B_N^*(x_i) - B(x_i)|\,.$$

$$.\, I[T(x_i) \geqq s] \leqq \sup_{x \in \mathcal{X}} |B_N^*(x) - B(x)|\,,$$

which tends to zero a.s. Then all assumptions of Lemma T2 hold (with $f(z) = 1/z$) and strong uniform consistency of $A_N(t)$ on $[0, \mathcal{T}]$ is proved. □

5. CONCLUSION

The next step would naturally consist in substitution the estimate $B_N(x)$ into the expression for $A_N(t)$. However, the properties of such an estimate $\hat{A}_N(t)$ are not clear. Certainly this problem is worth of further examination.

As the functions $L(t, x)$ and $B(x)$ are estimated by the kernel method, the asymptotics is rather slow and requires large extent of sample. The actual measure of convergence speed might be a subject of simulation study. In Theorem 2 assumption A5 indicates that the error of $B_N(x)$ decreases more slowly than $M^{-1/3}$. On the other hand, it is well known that $\sup_{0 \leqq t \leqq \mathcal{T}} |L_N(t, x) - L(t, x)| = \mathcal{O}\{\ln M_N(x)/M_N(x)\}^{1/2}$ a.s. When the assumptions A1 – A3 hold only, the proper choice of d_N (or the kernel estimation with another kernel function) makes possible to achieve the asymptotic error not significantly larger than $(\ln N/N)^{1/2}$.

The experience with the presented method is encouraging. The method has been tested both by real and simulated data. The output of the procedure consists of graphs or tables of estimated functions. As it is mostly the goal of nonparametric estimation, resulting estimates could suggest the proper form of parametrized model. When more-dimensional regression is analysed, the graphical projections of $B_N(x)$ are available.

Remark. In the following examples, the measure μ was chosen as the uniform distribution function on $[0, \mathscr{T}_1]$, $\mathscr{T}_1$ near 60% sample quantile from realized T_i's. Such choice μ and $\mathscr{T}_1$ guaranteed that the estimate of $B(x)$ would really reflect the shape of the cumulative hazard function in x. Varied window-width $2d_N$ ensured at least $N^{1/2}$ measurements in the window.

The samples of our examples were not censored.

Example 1. The simulated samples followed one-dimensoonal Cox's model with $B(x) = \exp(\beta x)$, $\beta = -1{\cdot}5$, $A(t) = t^2$. The covariate x was distributed uniformly through $[0, 10]$, the extents of samples differ from 100 to 500. The resulting

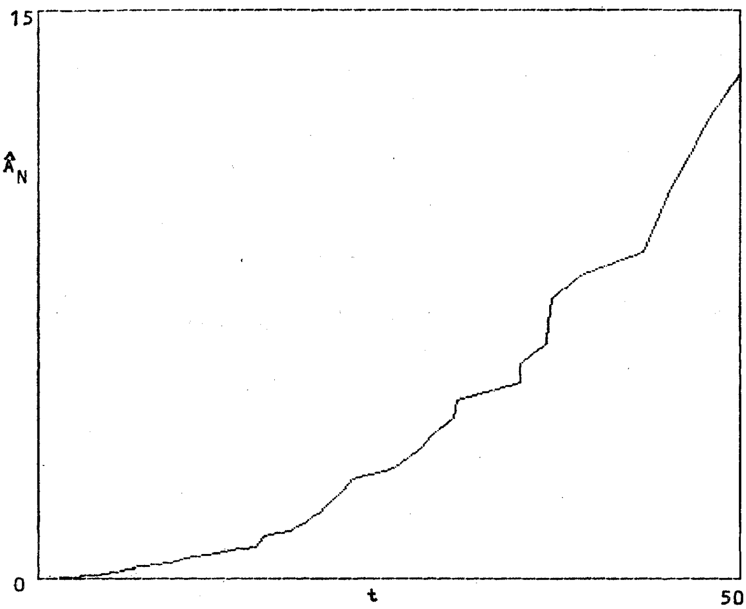

Fig. 1.

graphs of $\hat{A}_N(t)$ clearly showed parabolic trend (cf. Fig. 1 for $N = 200$). When the least squares line $c + dx$ was led through the points $\{x_i, \ln B_N(x_i)\}$, its parameter d could be regarded as an estimate of β. The estimation of the line is summarized in Table 1. $S2$ is the normalized sum of squares, $\hat{\beta}$ denotes the direct estimate of Cox's β.

Naturally, the original functions $A(t)$, $B(x)$ cannot be determined uniquely, we are able to estimate their shape, multiplied by some constant. This constant is connected with the normalizing condition (2). Therefore the graphs of $\hat{A}_N(t)$ have different scales, the estimates $\ln B_N(x)$ are shifted from $\ln B(x)$ (by the constant c in Example 1).

Table 1.

N	estimated c	d	correlation	$S2$	$\hat{\beta}$
100	4·167	−1·278	−0·994	0·163	−1·408
200	4·965	−1·401	−0·992	0·259	−1·462
300	6·376	−1·372	−0·994	0·193	−1·428
400	6·390	−1·411	−0·995	0·182	−1·411
500	6·295	−1·424	−0·997	0·125	−1·430

Example 2. The samples followed two-dimensional model $L(t, x, z) = A(t) \cdot B_1(x) \cdot B_2(z)$, with $A(t) = 2t$, $\ln B_1(x) = x^{-0\cdot5}$, $\ln B_2(z) = z^{0\cdot3}$. The covariates x, z were distributed uniformly in $[0, 10]$, $[0, 20]$ respectively. Figures 2, 3 display

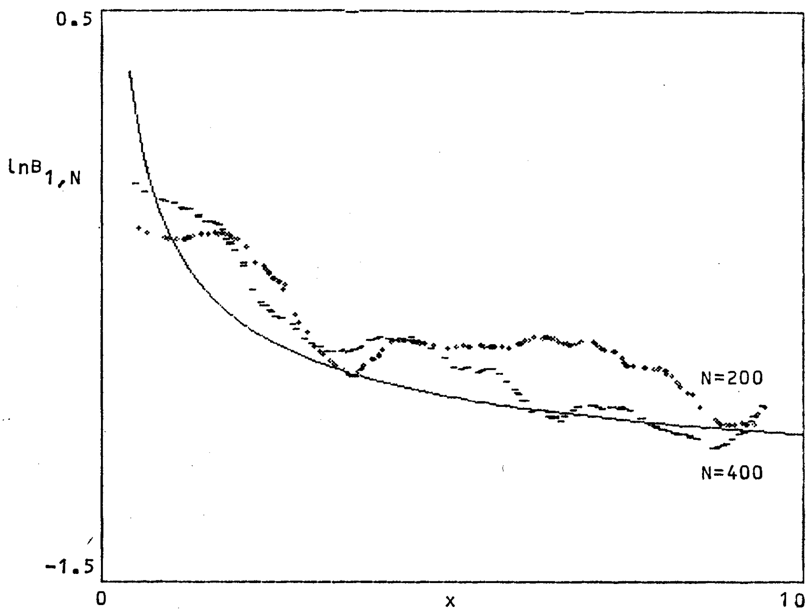

Fig. 2.

logarithm of the estimated $B_{1,N}(x)$, $B_{2,N}(z)$ after secondary smoothing by a fixed bandwidth moving average. They are compared with shifted model functions $x^{-0\cdot5}$, $z^{0\cdot3}$ respectively. Figure 4 shows part of $A_N(t)$ for $N = 200$.

(Received January 5, 1989.)

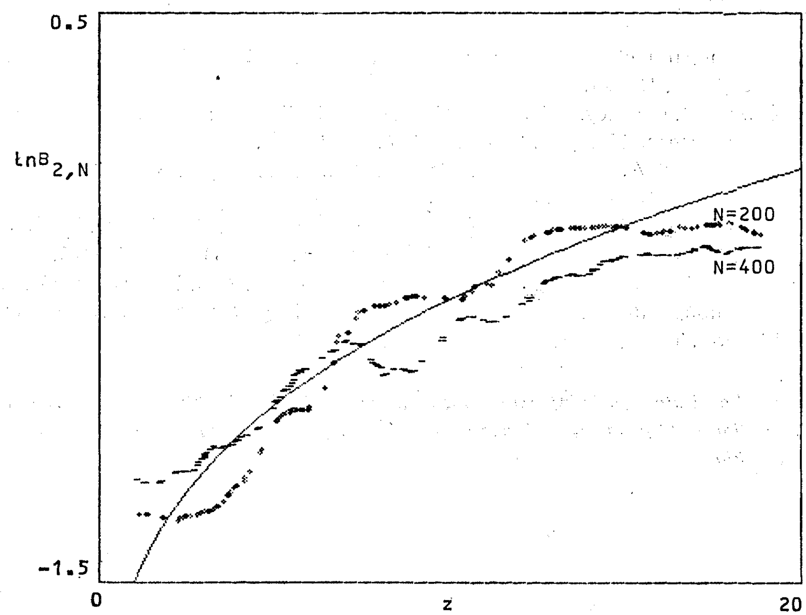

Fig. 3.

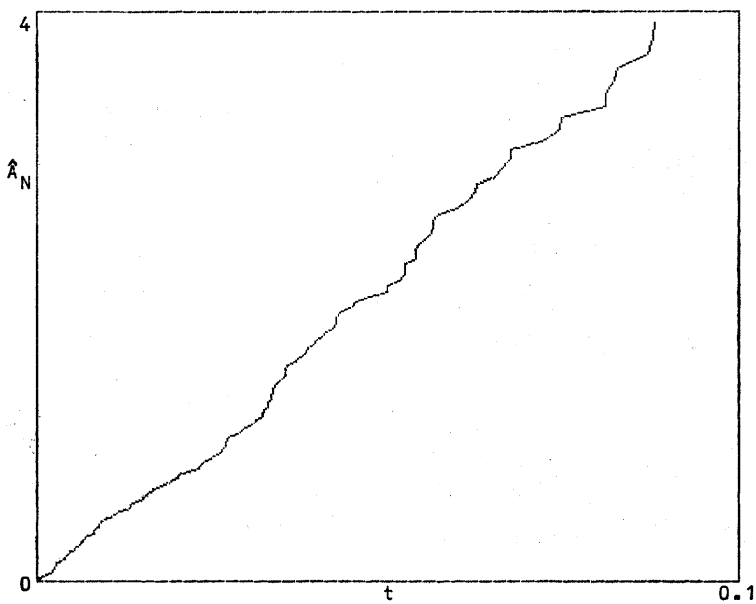

Fig. 4.

REFERENCES

[1] O. Aalen: Nonparametric inference in connection with multiple decrement models. Scand. J. Statist. *3* (1976), 15—27.

[2] N. Breslow and J. Crowley: A large sample study of the life table and product limit estimates under random censorship. Ann. Statist. *2* (1974), 437—453.

[3] I. I. Gikhman and A. V. Skorochod: Introduction to Theory of Random Processes (in Russian). Nauka, Moscow 1977.

[4] L. Rejtö: On the fixed censoring model and the consequences for the stochastic case. In: Trans. Ninth Prague Conference Vol. B, Academia, Prague 1983, pp. 141—147.

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

[6] P. Volf: A nonparametric analysis of proportional hazard regression model. Problems Control Inform. Theory *18* (1989), 5, 311—322.

Petr Volf, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.