The Korean Communications
in Statistics Vol. 4, No. 2, 1997
pp. 435 - 444

Testing Goodness of Fit in Nonparametric Function Estimation Techniques for Proportional Hazards Model[1)]

Jong-Tae Kim[2)]

Abstract

The objective of this study is to investigate the problem of goodness of fit testing based on nonparametric function estimation techniques for the random censorship model. The small and large sample properties of the proposed test, E_{mn}, were investigated and it is shown that under the proportional hazard model E_{mn} has higher power compared to the powers of the Kolmogorov -Smirnov, Kuiper, Cramér-von Mises, and analogue of the Cramér-von Mises type test statistic.

1. Introduction

Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be an independent, identically distributed sequence of nonnegative random variables. Assume that X_j and Y_j are independent with absolutely continuous distribution functions F and G respectively. If X_j is randomly censored by Y_j on the right, then the observations available consist of the pairs (Z_j, δ_j) for $1 \le j \le n$, where $Z_j = \min(X_j, Y_j)$ and, with $I(\cdot)$ standing for the indicator function $\delta_j = I(Z_j = X_j)$. Hence the obsevered Z_j constitute a random sample from the distribution function H. The survival function $S_Z(t)$ of Z_j has the property $S_Z(t) = S_X(t) S_Y(t)$ where $S_X(t)$ and $S_Y(t)$ are the survival functions of X_j and Y_j respectively. The usual censorship model is investigated by Kaplan and Meier (1958), Breslow and Crowley (1974) and many others.

The proportional hazards model is an appealing and potentially useful special nonparametric or parametric model of random censorship in which there exists a positive

1) This paper was supported in part by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1995.

2) Assistant Professor, Department of Statistics, Taegu University, Taegu, Korea

constant β, the censoring parameter, such that

$$S_Y(t)=(S_X(t))^{\beta} \text{ for all } t, \tag{1.1}$$

then this model is called the KG model (Koziol and Green (1976)). In this case the expected proportion of uncensored observations is given by

$$E(\delta_i)=P_r(X_i\le Y_i)=\int_0^{\infty}(1-F(T))^{\beta}dF(t)=(1+\beta)^{-1}=\alpha, \tag{1.2}$$

where $0<\alpha=(1+\beta)^{-1}<1$.The case $\beta=0$, or $\alpha=1$, corresponds to non-censoring and the expected number of the censored observations increases as β increases.

Abduskhurov (1987) and Cheng and Lin (1987) independently studied the large sample properties of the maximum likelihood estimator of S_X which is given by

$$\hat{S}_X(t)=(\hat{S}_Z(t))^{\alpha_n}, \tag{1.3}$$

where $\alpha_n=n^{-1}\sum_{i=1}^{n}\delta_i$ and $\hat{S}_Z(t)=n^{-1}\sum_{i=1}^{n}I(Z_i>t)$. This is called ACL estimator by Csörgo (1988). One clear advantage of the ACL estimator over the product-limit is its much simpler and more informative structure. Csörgo (1988) showed that the ACL estimator has theoretical advantage compare with to Kaplan-Meier (1958) estimator defined as

$$1-\hat{F}_n(t)=\prod_{i=1}^{n}\{1-\frac{\delta_{[i:n]}}{n-i-1}\}^{I[Z_{i:n}\le t]},$$

with $Z_{1:n}\le\cdots\le Z_{n:n}$ denoting the ordered Z-sample and $\delta_{[i:n]}$ being the δ-value associated with the ith Z-order statistic.

In this paper we are concerned with the problem of testing the hypothesis

$$H_0: F(t)=F_0(t) \quad \text{for all } t, \tag{1.4}$$

where F_0 is a completely known continuous distribution function with density f. Thus, by using the survival function $S_X(t)$, we see that testing H_0 is equivalent to testing

$$H_0: S_X(t)=S_0(t) \quad \text{for all } t. \tag{1.5}$$

Let $\hat{Z}_n(t)=\sqrt{n}\{\hat{F}_0(t)-F_0(t)\}\{1-A_n(t)\}/\{1-F_0(t)\}$, where $A_n(t)=a_n(t)/\{1+a_n(t)\}$, and $a_n(t)=n\sum_{\{i:X_i<t\}}(n-i)^{-1}(n-i+1)^{-1}\delta_i$. Koziol and Green (1976), Koziol (1980), Ghorai (1992), and many others considered a traditional statistics with the estimator $\hat{F}_n$ of F to test for H_0 and determined its asymptotic distribution in certain case; namely

the Kolmogorov-Smirnov, Kuiper, Cramér-von Mises, and analogue of the Cramér-von Mises type test statistic are defined by, respectively,

$$\hat{K}_n \equiv \hat{K}_{nLT}(t) = \sup_{0<t<LT} |\hat{Z}_n(t)|, \qquad \hat{V}_n \equiv \hat{V}_{nLT} = \sup_{0<t<LT} \hat{Z}_n(t) - \inf_{0<t<LT} \hat{Z}_n(t),$$

$$\hat{W}_n \equiv \hat{W}_{nt} = \int_0^{LT} \{ \hat{Z}_n(t) \}^2 dA_n(t), \quad \text{and} \quad \Psi_n^2 = n \int \{ \hat{F}_0(t) - F_0(t) \}^2 dF_0(t),$$

where LT is smaller than the largest observed lifetime; in each case, one would reject the H_0 for suitably large values of the test statistic.

2. Comparison Density view of Goodness-of-Fit

In the general setting, the problem discussed in previous section can be specified as follows: Let $T_1, \cdots, T_n$ be independent continuous failure times. The hypothesis to be tested is that the $U_i = S_0(T_i)$ are independent and identically distributed with uniform distribution on (0, 1), with the alternative of interest being that the U_i are independent and identically distributed with other distribution.

Proposition Let D be the distribution function of $U = S_0(t)$, then

$$D(u) = S_X(S_0^{-1}(u)), \quad 0 \le u \le 1. \tag{2.1}$$

Proof Let $S_0(t) = u$, then $t = F_0^{-1}(1-u) = S_0^{-1}(u)$ Therefore,

$$\begin{aligned} D(u) &= \Pr(S_0(T) \le u) = \Pr(T \ge F_0^{-1}(1-u)) \\ &= 1 - F(F_0^{-1}(1-u)). \end{aligned}$$

Thus, by using the ACL estimator in (1.3), the estimator of $D(u)$ can be written by

$$\begin{aligned} \hat{D}_n(u) &= \hat{S}_X(S_0^{-1}(u)) = (\hat{S}_Z(S_0^{-1}(u))^{\alpha_n} \\ &= (n^{-1} \sum_{i=1}^{n} I(Z_i > S_0^{-1}(u))^{n^{-1}\sum\delta_i} \end{aligned} \tag{2.2}$$

The corresponding density function of $U_1, \cdots, U_n$ is obtained by

$$d(u) = \frac{f_X(S_0^{-1}(u))}{f_0(S_0^{-1}(u))}. \tag{2.3}$$

The function d in (2.3) is called the comparison density function (Parzen (1979)). Thus, the H_0 in (1.4) is equivalent to

$$H_0 : D(u)=u , \quad 0\le u\le 1, \tag{2.4}$$

and, by using the comparison density function for the comparing F and F_0 the goodness-of-fit hypothesis become equivalent to

$$H_0 : d(u)=1, \quad 0\le u\le 1. \tag{2.5}$$

To measure the disparity between the comparison density and 1 in (2.5), Eubank, LaRiccia and Rosenstein (1987), and Kim (1994) used a generalized Fourier series expansion of d in the viewpoint of nonparametric density estimator which can be written as

$$d(u)=1+\sum_{k=1}^{\infty}\gamma_k p_k(u), \quad 0\le u\le 1, \tag{2.6}$$

where $\{p_k\}_{k=0}^{\infty}$ be a complete orthornomal sequence (CONS) for $L_2[0,1]$ with $p_0=1$ and γ_k's are generalized Fourier coefficients

$$\gamma_k=\int_0^1 (d(u)-1)\ p_k(u)\,du, \quad k=1,2,\cdots.$$

3. The Proposed Test

In the practical problem of $d(u)$ of (2.6), the comparison density can be estimated by first truncating the series after m terms and then plugging in estimates for the a_k's by letting $p_k(u)=\sqrt{2}\cos(k\pi u)$. Thus, the estimate $\hat{d}_{mn}$ of d in (2.6) is given by

$$\hat{d}_{mn}(u)=1+\sum_{k=1}^{m}\hat{\gamma}_{kn}\sqrt{2}\cos(k\pi u), \tag{3.1}$$

where an unbiased $\sqrt{n}$ - consistent estimator of γ_k is provided

$$\hat{\gamma}_{kn}=n^{-1}\sum_{i=1}^{m}\sqrt{2}\cos(k\pi U_i), \tag{3.2}$$

which is obtained by replacing $D(u)$ by the ACL estimator $\hat{D}_n(u)$ for $U_i, \cdots, U_n$. Using the squared $L_2[0,1]$ norm as a measure of distance then gives the test statistic

$$\begin{aligned} E_{mn} &= n\int_0^1 (\hat{d}_{mn}(u)-1)^2 du \\ &= n\sum_{k=1}^{m}\hat{\gamma}_{kn}^2 \end{aligned} \tag{3.3}$$

with $\hat{\gamma}_{kn}=n^{-1}\sum_{i=1}^{n}\sqrt{2}\cos(k\pi U_i)$. This reveals that choosing m is tantamount to selecting the proper amount of smoothing for the series estimator $\hat{d}_{mn}$ in (3.1) for the comparison density $d(u)$, Thus, a test for H_0 could be based directly on an estimator of $d(u)$ if H_0 is rejected.

Neyman (1936) show that the $n^{-1}\gamma_{kn}$ are independent asymptotically $N(0,1)$ random variable under H_0 and as result E_{mn} is approximately chi-squared distributed with m degree of freedom under H_0. Thus the null hypothesis will be rejected for large values of E_{mn}.

Theorem 3.1 Under the local alternatives comparison density $d_n(u)=1+b(n)\delta(u)$ with $\delta\in L^2[0,1]$ and $b(n)\to 0$ as $n\to\infty$, then, for $b(n)=n^{-1/2}$ and fixed m,

$$E_{mn}\xrightarrow{d}\chi^2_m(\sum_{j=1}^{m}\delta_j^2),$$

where $\delta_j=\int_0^1\delta(u)p_j(u)\,du$ and p_j is a complete orthonormal sequence.

Proof For E_{mn} it is known that if $b(n)=n^{-1/2}$, then, from Shorack and Wellner (1986),

$$E_{mn}\xrightarrow{d}\sum_{j=1}^{m}(Z_j+\delta_j)^2,$$

where, " $\xrightarrow{d}$ " signifies convergence in distribution, the Z_j are *iid* $N(0,1)$ random variables and $\delta_j=\int_0^1\delta(u)p_j(u)du$ with p_j defined in (2.6). Let $\chi^2{}_a(b)$ denoted a chi-squared random variables with a degrees of freedom and noncentral parameter b. Now, for fixed m and $b(n)=n^{-1/2}$, E_{mn} converges in distribution $\chi^2_m(\sum_{j=1}^{m}\delta_j^2)$. (See Eubank and LaRiccia (1992)).

4. Power of The Tests

In this section, we study the finite sample power properties of five tests: namely, the Kolmogorov-Smirnov, $\hat{K}_n$, Kuiper, $\hat{V}_n$, Cramér-von Mises, $\hat{W}_n$, the analogue of the Cramér-von Mises type test statistic, $\Psi_n{}^2$, and the proposed test, E_{mn}, for randomly censored data introduced. From the simulation studies for finding the optimal m, the optimal m is 2 in the cases of exponential and Weibull alternatives cases generally. Therefore, one can use E_{2n} instead of E_{mn}.

We consider the two cases of particular interest in survival studies to compare the power of statistic by using the results of simulation study in Koziol (1980): (Case I) $F_0(t)=1-\exp(-t)$, exponential survival, and alternatives $F_0(t)=1-\exp(-\lambda t)$, scale shifts for $\lambda=1.0, 0.8, 0.6, 0.4$. (Case II) $F_0(t)=1-\exp(-t)$, and alternatives $F_0(t)=1-\exp(-t^{\lambda})$, Weibull alternatives for $\lambda=1.0, 0.6, 0.35, 0.2$ to exponentiality. With each case, the censoring distribution G in the proportional hazard model is assumed to be rated to F_0 by $(1-G)=(1-F_0)^{\beta}$.

The results of the simulation study are presented in Table 1 and Table 2: the top line in each block of the tables, the case of $\lambda=1$, gives the frequencies of the exceeding the nominal asymptotic 0.05 critical values, and other lines give the frequencies of exceeding estimated 0.05 critical values. Each figure is based on 1000 samples for the censoring parameter $\beta=0.5, 1$, and the sample size $n=20, 50$. The critical values for E_{2n} were found by simulation with 20000 replicate experiments for the each sample size n.

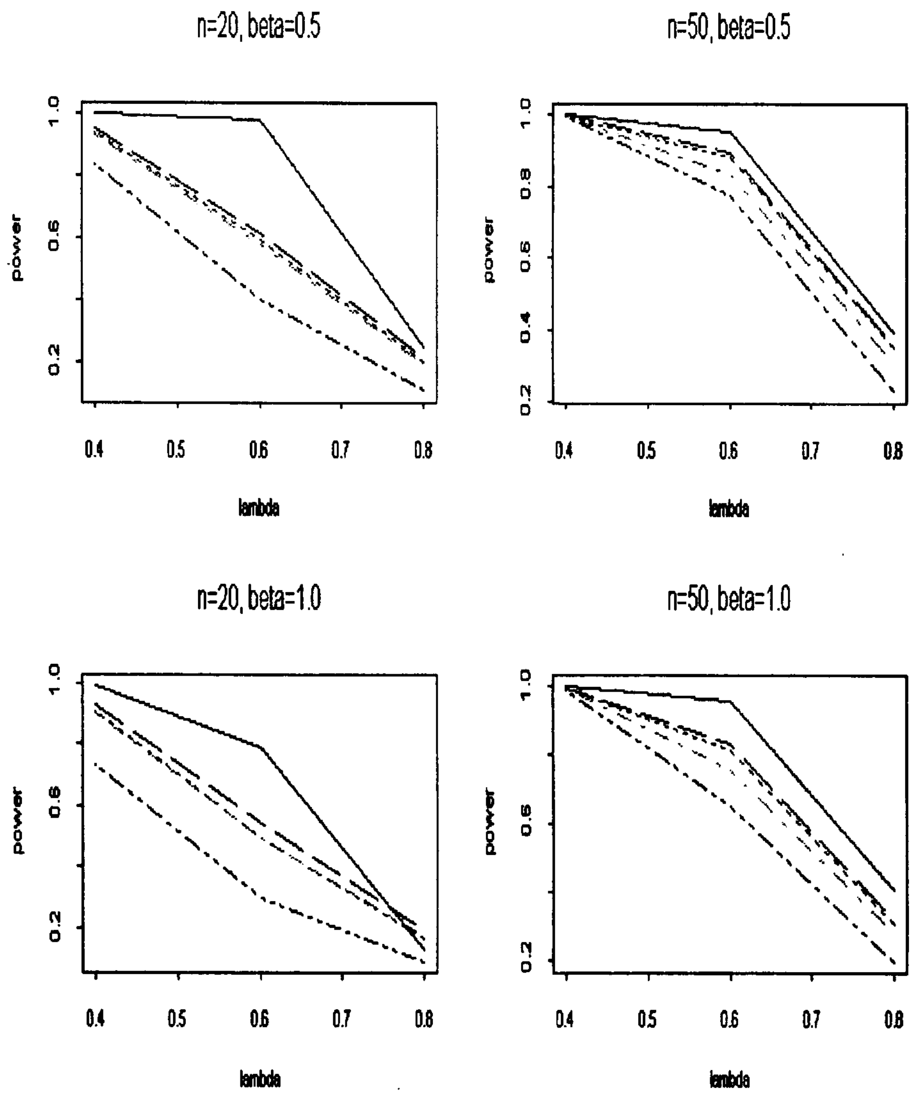
From the Table 1 and Table 2, the power comparisons reveal the expected trends: power increases with sample size, but decreases as the amount of censorship increases. The test using E_{2n} is clearly more powerful than the others. Eubank and LaRiccia (1992), and Kim (1994) showed that the analogue of E_{mn} type is superior to the statistic of Cramer-von Mises type. In the random censoring data, we can see that E_{2n} have significantly better powers than the others from Figure 1 and Figure 2.

< Table 1 > Frequencies of exceeding nominal and estimated 0.05 level critical values for the exponential alternatives.

n	β	λ	$\hat{K}_n$	$\hat{V}_n$	$\hat{W}_n$	Ψ_n^2	E_{2n}
20	0.5	1.0	0.104	0.091	0.077	0.053	0.049
		0.8	0.200	0.194	0.215	0.109	0.253
		0.6	0.592	0.579	0.618	0.399	0.978
		0.4	0.941	0.935	0.952	0.838	1.000
50	0.5	1.0	0.091	0.079	0.059	0.049	0.051
		0.8	0.353	0.313	0.359	0.230	0.391
		0.6	0.880	0.837	0.893	0.776	0.953
		0.4	0.999	0.998	1.000	0.993	1.000
20	1.0	1.0	0.066	0.054	0.049	0.051	0.048
		0.8	0.166	0.172	0.195	0.086	0.130
		0.6	0.497	0.492	0.544	0.299	0.786
		0.4	0.911	0.899	0.932	0.733	0.991
50	1.0	1.0	0.084	0.069	0.054	0.047	0.053
		0.8	0.305	0.280	0.318	0.190	0.406
		0.6	0.814	0.759	0.832	0.652	0.959
		0.4	0.997	0.995	0.997	0.989	1.000

< Table 2 > Frequencies of exceeding nominal and estimated 0.05 level critical values for the Weibull alternatives.

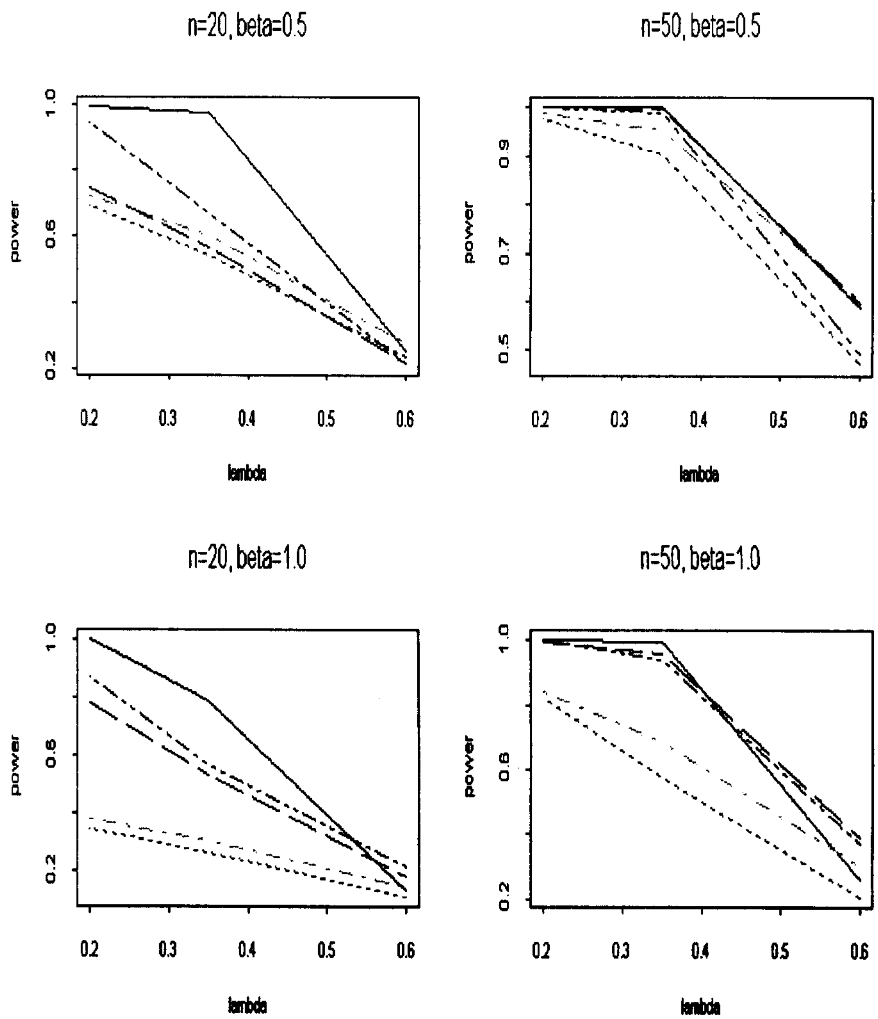
n	β	λ	$\hat{K}_n$	$\hat{V}_n$	$\hat{W}_n$	Ψ_n^2	E_{2n}
20	0.5	1.0	0.107	0.090	0.079	0.058	0.049
		0.6	0.234	0.276	0.216	0.216	0.253
		0.35	0.544	0.606	0.568	0.667	0.978
		0.2	0.695	0.721	0.748	0.945	0.993
50	0.5	1.0	0.104	0.095	0.072	0.056	0.051
		0.6	0.471	0.604	0.596	0.493	0.587
		0.35	0.904	0.952	0.995	0.990	1.000
		0.2	0.977	0.987	1.000	1.000	1.000
20	1.0	1.0	0.067	0.049	0.046	0.048	0.048
		0.6	0.110	0.146	0.184	0.212	0.130
		0.35	0.264	0.308	0.531	0.564	0.786
		0.2	0.344	0.382	0.781	0.866	0.999
50	1.0	1.0	0.092	0.073	0.061	0.059	0.053
		0.6	0.203	0.303	0.389	0.363	0.263
		0.35	0.576	0.687	0.958	0.942	0.997
		0.2	0.821	0.841	0.998	0.999	1.000



** ———— E_{2n} , — — — — — $\Psi_n^{\,2}$, — — — — $\hat{W}_n$

— - — - — $\hat{V}_n$, - - - - - - $\hat{K}_n$.

< Figure 1 > Empirical Power functions for the exponential alternatives with $\alpha=0.05$.



** ———— E_{2n} , – – – – – – – Ψ_n^2 , — — — — $\hat{W}_n$

— – — – — $\hat{V}_n$, - - - - - - - $\hat{K}_n$.

< Figure 2 > Empirical Power functions for the Weibull alternatives with $\alpha = 0.05$

References

[1] Abdushukurov, A. A. (1987). Estimation of Probability Density and the Hazard Rate Function in the Koziol-Green Model of Random Censorship, *Izv. Akud. Nauk-UzzSSR Ser. Fiz. Math Nauk* 1987, No.3, 3-10 (in Russian).

[2] Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit Estimates under Random Censorship, *Annals of Statistics,* 2, 437-453.

[3] Cheng, P. E. and Lin, G. D. (1987). Maximum Likelihood Estimation of a survival Function under the Koziol-Green Proportional Hazard Model, *Statistics & Probability Letters,* 5, 75-80.

[4] Csörgo, S. (1988). Estimation in the proportional hazards model of random censorship, *Statistics*, 19, 437-463.

[5] Eubank, R. LaRiccia, V. N., and Rosenstein, R. (1987). Test Statistical Derived as Components of Pearson's Phi-Squared Distance Measure, *Journal of the American Statistical Association,* 82, 816-825.

[6] Eubank, R. LaRiccia, V. N. (1992). Asymptotic Comparison of Cramér von Mises and Nonparametric Function Estimation Techniques for Testing Goodness-of-Fit, *Annals of Statistics,*

[7] Ghorai, J. K. (1992). Cramér-von Mises statistic for Testing Goodness-of-Fit under Proportional Hazard Model, *Commun Statist-Theory and Method,* 20, 1107-1126.

[8] Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete observations, *Journal of the American Statistical Association,* 53, 457-481.

[9] Kim, J. T. (1994). Goodness-of-Fit Test Based on Smoothing Parameter Selection Criteria, *The Korean Communications in Statistics,* 2, 122-136.

[10] Koziol, J. A. (1980). Goodness-of-fit tests for random censored data, *Biometrica*, 67, 693-696

[11] Koziol, J. A. and Green, S. B. (1976). A Cramér-von Mises statistic for randomly censored data, *Biometrica,* 63, 465-474.

[12] Neyman, J. (1937). Smooth Test for Goodness-of-Fit, *Skandinavisk Aktuarietidskrift,* 20, 149-199.

[13] Parzen, E. (1979). Nonparametric Statistical Data Modeling, *Journal of the American Statistical Association,* 74, 105-131.

[14] Shorack, G. and Wellner, J. (1986). Empirical Processes with Applications to Statistics, Wiley; New York.