PRACE IPI PAN • ICS PAS REPORTS

Jan Mielniczuk

Kernel estimators of a density function in case of censored data

560

February 1985

WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P.O. Box 22, POLAND

Jan Mielniczuk

KERNEL ESTIMATORS OF A DENSITY FUNCTION IN CASE OF CENSORED DATA

Warsaw, February 1985

Pracę zgłosił Antoni Mazurkiewicz

Mailing address: Jan Mielniczuk
Institute of Computer Science
Polish Academy of Sciences
P.O. Box 22
00-901 Warszawa, PKiN

ISSN 0138-0648

P r i n t e d   a s   a   m a n u s c r i p t
N a   p r a w a c h   r ę k o p i s u

Nakład 250 egz., Ark. wyd. 0.45; ark. druk. 1,00. Papier offset.kl. III, 70 g, 70 x 100. Oddano do druku w lutym 1985 r. W.D.N. Zam. Nr 130/85

Abstract . Содержание . Streszczenie

The kernel estimator is a widely used tool for estimation of a density function in the classical case. In the paper its adaptation to censored data using the Kaplan-Meier estimator is being considered. Asymptotic properties of four estimators, arising naturally as a result of considering various types of bandwidths, are investigated. In particular we prove that (i) both proposed estimators steming from the nearest neighbor estimator have censoring-free variances, (ii) one of them is L^1-consistent.

Ядерные статистики плотности правдоподобий для цензурированных выборок

Ядерные статистики плотности правдоподобия часто применяются в классическом случае. В работе рассмотрена их адаптация к случаю случайно цензурированных справа выборок. Изучаются асимптотические свойства четырех статистик возникающих при рассмотрении разных типов окна. В особенности доказано, что (i) асимптотическая варианция статистик, которые являются адаптациями ядерной статистики типа ближайший сосед не зависит от цензурирования (ii) одна из них состоятельна в среднем.

Estymatory jądrowe gęstości dla danych obciętych

Szeroko stosowanym narzędziem do estymacji funkcji gęstości w przypadku klasycznym są estymatory jądrowe. W pracy rozpatrzona jest ich adaptacja dla danych obciętych. Zbadane są własności asymptotyczne czterech estymatorów powstałych wskutek rozpatrzenia różnych typów rozstępu. W szczególności udowadnia się, że (i) estymatory będące modyfikacjami estymatora jądrowego typu najbliższy sąsiad mają wariancje asymptotyczne niezależne od obcinania (ii) jeden z nich jest punktowo L^1-zgodny.

1. INTRODUCTION

Consider the random censorship model with two sequences $X_1,\dots,X_n$ and $Y_1,\dots,Y_n$ of i.i.d. nonnegative random variables such that X_i, Y_i are independent $(i=1,\dots,n)$. Let F and G be unknown rightcontinuous distribution functions of X's and Y's respectively. It is assumed that X_i and Y_i have densities f and g with respect to Lebesque measure on R^1. We want to estimate f using the following data

$$Z_i = \min(X_i, Y_i), \quad \delta_i = [X_i \leq Y_i] \qquad (i=1,\dots,n),$$

where A for any event $[A]$ denotes the indicator function of A. Let H be the distribution function of Z's. The well-known product-limit (K-M) estimator is defined by

$$1 - \hat{F}_n(u) = \prod_{i: Z_i \leq u} \left(\frac{n-i}{n-i+1}\right)^{\delta(i)} \qquad u < Z_{(n)}$$

$$= \quad 0 \qquad\qquad u > Z_{(n)}$$

$\delta_{(i)}$ being the concomitant of $Z_{(i)}$. The K-M empirical survival function $1 - \hat{F}_n$ will be denoted by $\hat{\bar{F}}_n$.

Földes, Rejtő and Winter (1981) introduced a kernel type estimator of f based on the K-M estimator:

$$(1.1) \qquad \hat{f}_n(x) = \frac{1}{h(n)} \int_R K\left(\frac{x-y}{h(n)}\right) d\hat{F}_n(y)$$

AMS 1980 subject classifications. Primary 62G05; Secondary 60F15. Key words and phrases. Censored data, density estimator, k-nearest neighbor estimator, Kaplan-Meier estimator, kernel, random censorship model.

where $h(n)$ is a sequence of positive numbers such that $h(n) \to 0$, $nh(n) \to \infty$ and K is a density function. Analogously, we define a $k(n)$th nearest uncensored neighbor estimator

(1.2) $$f_n(x) = \frac{1}{R(n)} \int_R K\left(\frac{x-y}{R(n)}\right) d\hat{F}_n(y)$$

where $R(n)$ is the distance from x to its $k(n)$th nearest uncensored neighbor and $k(n)$ is a given sequence of integers such that $k(n) \to \infty$ and $k(n)/n \to 0$. Classical nearest neighbor estimators were studied by Moore and Yackel (1976, 1977), Mack and Rosenblatt (1980) and Mack (1980, 1982).

Moreover, we introduce

(1.3) $$\hat{f}^*_n(x) = \frac{1}{h(n_1)} \int_R K\left(\frac{x-y}{h(n_1)}\right) dF_n(y)$$

(1.4) $$f^*_n(x) = \frac{1}{R(n_1)} \int_R K\left(\frac{x-y}{R(n_1)}\right) dF_n(y)$$

where $n_1 = \sum_{i=1}^{n} \delta_i$ is the number of uncensored ($\delta_i = 1$) observations. It seems natural to consider estimators (1.3) and (1.4) since K-M estimator has jumps in uncensored observations and additionally in last observation being censored or not. We shall show that some properties of all these estimators may be deduced from the properties of classic kernel estimators when the observations are not censored. The connections with the classical case are stated in Lemma 1 for $\hat{f}_n$ and f_n and in Lemma 2 for $\hat{f}^*_n$ and f^*_n. The link is evident in the second case since then the uncensored observations may be treated as n_1 random variables distributed as $(Z|\delta=1)$ where $(Z,\delta) \sim (Z_i, \delta_i)$ for any i. In the first case we consider

$$\breve{F}_n(y) = \frac{1}{n} \sum_{i=1}^{n} [Z_i \leq y,\ \delta_i = 1]$$

which, on any compact interval $[0,d]$ may be interpreted as the empirical distribution function for some random variable $\breve{X}_d$. In section 3 some results on consistency and weak convergence of the introduced estimators are proved. In particular it is shown that the asymptotic variances of f_n and f_n^* do not depend on censoring, as opposed to the asymptotic variances of $\hat{f}_n$ and $\hat{f}_n^*$. Finally, we state a L^1-consistency result for f_n^*. The method of proving this result confirms our opinion that the estimators with bandwiths based on the subsample of uncensored observation are worth consideration.

2. CLASSICAL ANALOGUES FOR $\hat{f}_n, f_n, \hat{f}_n^*$ and f_n^*

Put $p=P(\delta=1)$ and $q=1-p$. Observe first that defining $R(n)$ as the distance to $k(n)$th uncensored observation leaves $R(n)$ undefined on the set $A_n = \{k(n) > n_1\}$. However this has no influence on asymptotic properties of $R(n)$ since taking n_o such that $\forall\, n > n_o\ np-k(n) > npq$ we have

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} P(Bin(n,p) < k(n)) < \sum_{n=1}^{\infty} P(Bin(n,p)-np > np-k(n))$$

$$< n_o + \sum_{n=n_o}^{\infty} P(Bin(n,p)-np > npq) < n_o + \sum_{n=n_o}^{\infty} 2\exp(-2/9\ npq) < \infty$$

(In the end of this argument Bernstein's inequality was used).

From now on we denote by x a fixed point of R^+ such that $f(x)\bar{G}(x)>0$ where $\bar{G}=1-G$. Let $\breve{X} = Z[\delta=1] + (Z+x+1)[\delta=0]$ and let W_i for $i=1,\dots,n$ be defined as $\breve{X}$ with Z and δ replaced by Z_i and δ_i respectively. Obviously, $(W_1,\dots,W_n)$ is an i.i.d. sequence with W_i's distributed as X. Let $\bar{R}(n)$ be the distance from x to its $k(n)$th nearest neighbor in the sequence $(W_1,\dots,W_n)$. For any function f denote by $C(f)$ a set of continuity points of f.

Lemma 1. If $k(n)/\log\log n \to \infty$ and $x \in C(f\,\bar{G})$ then

(i) $\bar{R}(n) = R(n)$ for all but finite number of n a.s.

(ii) $k(n)/nR(n) \to 2f(x)\bar{G}(x)$ a.s.

Proof. Using the results of Moore and Yackel (1976) we have

$$\bar{R}(n) \to 0 \quad \text{a.s.}$$

Thus the first $k(n)$ neighbors lie in a small neighborhood of x and by the definition of $\bar{R}(n)$ they must be uncensored. Since on $[0,x+1)$ $\breve{X}$ has the density equal to $f(y)\bar{G}(y)$ and continuous in x then (ii) follows from (i) and Theorem 1 of Moore and Yackel (1976).

Note that $\breve{F}_n(y)$ is the e.d.f. of $\breve{X}$ on $[0,x+1)$ and thus we may estimate $f(x)\bar{G}(x)$ by means of

$$\frac{1}{h(n)} \int_R K\left(\frac{x-y}{h(n)}\right) d\breve{F}_n(y)$$

or

$$\frac{1}{R(n)} \int_R K\left(\frac{x-y}{R(n)}\right) d\breve{F}_n(y)$$

using the kernel K which has a compact support.

Theorem 1. Let K be a bounded density function with support in $[-1,1]$. Assume that $x \in C(f\,\bar{G}) \cap C(g)$.

(i) If $k(n)/\log\log n \to \infty$ then

$$P\left(\left|f_n(x) - \frac{1}{\bar{G}(x)\bar{R}(n)} \int_R K\left(\frac{x-y}{R(n)}\right) d\breve{F}_n(y)\right| = O\left(\frac{\log n}{n^{1/2}}\right) + O\left(\frac{k(n)}{n}\right)\right) = 1$$

(ii) If $nh(n)/\log\log n \to \infty$ then

$$P\left(\left|\hat{f}_n(x) - \frac{1}{\bar{G}(x)h(n)} \int_R K\left(\frac{x-y}{h(n)}\right) d\breve{F}_n(y)\right| = o\left(\frac{\log n}{n^{1/2}}\right) + O(h(n))\right) = 1$$

Proof of (i). Let $S(x,r) = \{ y: |y-x| \le r \}$

$$\left| f_n(x) - \frac{1}{\bar{G}(x)R(n)} \int_R K\left(\frac{x-y}{R(n)}\right) d\tilde{F}_n(y) \right| < \sup K \frac{k(n)}{R(n)} \max_{\substack{z_i \in S(x,R(n)) \\ \delta_i = 1}} \left| 1/nG(x) - a_n(z_i) \right|$$

where $a_n(z_i)$ is the value of the jump of K-M estimator in z_i.
Using (ii) of Lemma 1 it is enough to show that

$$(2.1) \qquad \max_{\substack{z_i \in S(x,R(n)) \\ \delta_i = 1}} \left| 1/\bar{G}(x) - na_n(z_i) \right| = o\left(\frac{\log n}{n^{1/2}}\right) + O\left(\frac{k(n)}{n}\right).$$

Since $na_n(z_i) = F_n(z_i-0)/H_n(z_i-0)$ (Efron (1967)), where $\bar{H}_n$ is the empirical survival function for all observations, then (2.1) is majorized by

$$\sup_{t \le x+R(n)} \left| \frac{\hat{\bar{F}}_n(t)}{\bar{H}_n(t)} - \frac{\hat{\bar{F}}_n(t)}{\bar{H}(t)} \right| + \sup_{t \le x+R(n)} \left| \frac{\hat{\bar{F}}_n(t)}{\bar{H}(t)} - \frac{\bar{F}(t)}{\bar{H}(t)} \right| + \sup_{t \in S(x,R(n))} \left| \frac{1}{\bar{G}(x)} - \frac{1}{\bar{G}(t)} \right|$$

Two first terms are $o(\log n \; n^{-1/2})$ a.s. in view of Glivenko theorem and the result of Földes and Rejtő (1981), respectively. For the last term it is enough to show that

$$P(Y \in S(x,R(n)) = O\left(\frac{k(n)}{n}\right) \text{ a.s.}$$

which is implied by

$$P(Y \in S(x,R(n))/R(n) \to 2g(x) \text{ a.s.}$$

and $n\,R(n)/k(n) = O(1)$ a.s.
Proof of (ii) follows the lines of (i) with the equality

$$(2.2) \quad \lim_{n \to \infty} \sum_{i=1}^{n} [Z_i \in S(x,h(n)), \delta_i = 1]/n\,h(n) = 2f(x)\,G(x)$$

used instead of (ii) of Lemma 1. Formula (2.2) is obvious in view of the strong consistency result for the classical kernel estimator with the kernel uniform on $[-1,1]$ applied to $(W_1,\dots,W_n)$.

Let us turn to $\hat{f}_n^*$ and f_n^*. Consider $(Z|\delta=1)$ instead of $\check{X}$, and the sequence $Z_{i_1},\dots,Z_{i_{n_1}}$ of uncensored observations instead of $W_1,\dots,W_n$. The density of $(Z|\delta=1)$ is equal to $f_1(x) = f(x)\bar{G}(x)/p$.

Lemma 2. Let k be a fixed integer $\leq n$. Conditionally on $n_1=k$, $Z_{i_1},\dots,Z_{i_{n_1}}$ is an i.i.d. sequence with density f_1.

From Lemma 2, for $T_{n_1}(t) = \frac{1}{n_1}\sum_{i=1}^{n}[Z_i<t,\delta_i=1]$ the estimators

$$\frac{1}{h(n_1)}\int_R K\left(\frac{x-y}{h(n_1)}\right)dT_{n_1}(y), \qquad \frac{1}{\bar{R}(n_1)}\int_R K\left(\frac{x-y}{\bar{R}(n_1)}\right)dT_{n_1}(y)$$

can be viewed as the classical kernel estimators for f_1, based on a sample of random size n_1. From Lemma 2 it is also easy to see that asymptotic properties of these estimators such as convergence in probability and weak convergence, are the same as the respective asymptotic properties of analogous estimators based on n observations from the distribution of $(Z|\delta=1)$.
Moreover, an exact analogue of Theorem 1 is true for f_n^* and $\hat{f}_n^*$ and its proof is similar to that of Theorem 1. This allows us to study some asymptotic properties of $\hat{f}_n^*$ and f_n^*.

3. ASYMPTOTIC PROPERTIES OF THE INTRODUCED ESTIMATORS

Below we list some properties of estimators f_n and $\hat{f}_n$. They rely upon analogous properties of their classical counterparts and, as for the consistency results, on the following theorem of

Moore and Yackel (1977). Let K satisfied the assumptions of Theorem 1 and the condition $K(cu) \geq K(u)$ for any $0 < \; < 1$. Then any consistency result which is true for a kernel estimator with $h(n) = k(n)/n$ and with kernel K and the uniform kernel, respectively, is also true for a nearest neighbor estimator with kernel K. We assume that conditions of Theorem 1 are satisfied.

Corollary 1. (i) If $\sum_{n=1} \exp(-ck(n)) < +\infty$ for every $c > 0$, $K(cu) \geq K(u)$ for $0 \leq c \leq 1$ then

$$f_n(x) - f(x) \to 0 \text{ a.s.}$$

(ii) If $\sum_{n=1} \exp(-cnh(n)) < +\infty$ for every $c > 0$ then

$$\hat{f}_n(x) - f(x) \to 0 \text{ a.s.}$$

If K is the uniform kernel on $[-1,1]$ then $f_n(x)$ and $\hat{f}_n(x)$ are strongly consistent under the assumptions of Theorem 1.

Corollary 2. Let g be continuous and $f\,\bar{G}$ continuous and positive on $[0,T]$.

(i) Suppose that K is continuous and $K(cu) \geq K(u)$ for $0 \leq c \leq 1$. If $k(n)$ is a sequence of integers such that $k(n)/\log n \to +\infty$ then

$$(3.1) \qquad P(\lim_n \sup_{0 \leq x \leq T} |f_n(x) - f(x)| = 0) = 1$$

for T such that $\bar{H}(T) > 0$.

(ii) If K is a continuous kernel with the bounded variation then (3.1) holds with f_n replaced by $\hat{f}_n$.

Proof of (i). Observe that in view of Theorem A in Silverman (1978) it is enough to show that under the above assumptions strong convergence in Theorem 1 can be replaced by uniform strong convergence on $[0,T]$. To see this observe that since

$k(n)/\log n \longrightarrow \infty$ and $f \cdot \bar{G}$ is uniformly continuous on [O,T] then in view of Theorem 1 in Devroye and Wagner (1977)

(3.2) $$\sup_{0 \leq x \leq T} |k(n)/nR(n,x) - 2f(x)G(x)| \to 0 \text{ a.s.}$$

It remains to consider the last term of the majorant occurring in the proof of Theorem 1 and show that

$$\sup_{0 \leq x \leq T} |G(x+R(n)) - G(x-R(n))| = O(k(n)/n) \text{ a.s.}$$

We have

$$\sup_{0 \leq x \leq T} |G(x+R(n)) - G(x-R(n))| = \sup_{0 \leq x \leq T} \frac{k(n)}{n} \frac{nR(n)}{k(n)} \frac{P(Y \in S(x,R(n))}{R(n)}$$

Since sup R(n) on [O,T] tends to O a.s. and g is uniformly continuous we have

$$\sup_{0 \leq x \leq T} |P(Y \in S(x,R(n))/R(n) - 2g(x)| \longrightarrow 0 \text{ a.s.}$$

Thus the proof of (i) is completed in view of (3.2) and the fact that $f \cdot \bar{G}$ is positive on O,T .

A proof of (ii) is similar.

Remark. Observe that the uniform strong convergence of $\hat{f}_n$ on [O,T] is obtained, with the stronger condition on h(n) : $\sum \exp(-cnh^2) < +\infty$ for all positive c and with K of bounded variation, using the result of Nadaraya (1965) and the inequality (Földes and Rejtő (1981))

$$P\left(\sup_{0 \leq x \leq T} |F_n(x) - F(x)| > \varepsilon \right) < d_0 \exp(-n^2 \delta^4 d_1)$$

where $\bar{H}(T) > \delta > 0$ and $\varepsilon > 2^7/n\delta^2$, d_0, d_1 being universal positive constants.

Corollary 3. (i) Assume that $f\,G$ has the bounded derivative in the neighborhood of x. If $k(n) = o(n^{2/3})$, then

$$(3.3)\quad (k(n))^{1/2}(f_n(x)-f(x)) \xrightarrow{\mathcal{L}} N(0, 2\,f^2(x)\int_R K^2(y)dy)$$

(ii) Assume that K is an even function, $f\bar{G}$ has the second derivative which is bounded in the neighborhood of x and $h(n) = o(n^{-1/3})$. Then

$$(3.4)\quad (nh(n))^{1/2}(f_n(x)-f(x)) \xrightarrow{\mathcal{L}} N(0, f(x)/G(x)\int_R K^2(y)dy)$$

Proof of (i). Observe that for $w_n(x) = 1/R(n)\int_R K((x-y)/R(n))d\hat{F}_n(y)$ we have

$$(3.5)\quad (k(n))^{1/2}(w_n(x)-f(x)G(x)) \xrightarrow{\mathcal{L}} N(0, 2(f(x)\bar{G}(x))^2\int_R K^2(y)dy)$$

(Moore and Yackel (1976)). Since for $k(n) = o(n^{2/3})$, $k(n)^{1/2}$ is $o(n/k(n))$, then (i) follows from (3.5) and Theorem 1.

Proof of (ii). Rosenblatt (1971) proved that under the conditions imposed on K in (ii) and $h(n) = o(n^{1/5})$ $\hat{w}_n(x) = \frac{1}{h(n)}\int_R K(\frac{x-y}{h(n)})d\hat{F}_n(y)$ is asymptotically normal with mean $f(x)G(x)$ and asymptotic variance $1/(nh(n))\ f(x)\bar{G}(x)\int_R K^2(y)dy$. The result follows from the fact that $(nh(n))^{1/2} = o(1/h(n))$ for $h(n) = o(n^{-1/3})$.

Analogues of Corollaries 1, 2 and 3 for f_n^* and f_n^* are also true. The only change is that in Corollary 3 the scaling sequences $(nh(n))^{1/2}$ and $(k(n))^{1/2}$ are replaced by $(n_1h(n_1)))^{1/2}$ and $(k(n_1))^{1/2}$, respectively.

4. POINTWISE CONVERGENCE IN THE MEAN OF $f_n^*(x)$

Theorem 2. Assume that conditions of Corollary 1 (i) are satisfied. Suppose that $\log n\, k(n)/n \to 0$, f is a bounded density function, and x satisfies $f(x)\bar{G}(x) > 0$. Then

$$\int |f_n^*(x) - f(x)|\, dP \to 0.$$

Let us first state the following simple consequence of Bernstein's inequality given in Bjerve (1977)

Lemma 3. Let $m = \alpha n + O(1)$ for some $\alpha < 1$. Consider a point x such that $F(x) < \alpha$. Then

$$P(X^n_{(m)} \leqslant x) \leqslant 2e^{-cn}$$

where $X^n_{(1)},\dots,X^n_{(n)}$ is a sequence of order statistics for a sample of size n with d.f. F, while c is a positive constant not depending on n.

Proof of Theorem 2. Let F_1 be the d.f. of $(Z|\delta=1)$ and let be a positive constant such that $F_1(x) + 2\varepsilon < 1$. Using Lemma 3 for the i.i.d. sample $X_1,\dots,X_{n-k(n)}$ with $X_i \sim F_1\ (i=1,\dots,n-k(n))$ we obtain

$$P(X^{n-k(n)}_{(i-k(n))} \leqslant x) < e^{-c(n-k(n))}$$

for i such that $i-k(n) \geqslant (F_1(x) + \varepsilon)(n-k(n))$.

Observe that if we choose $i > (F_1(x)+2\varepsilon)n$, this condition is satisfied for sufficiently large n. Since the analogue of Corollary 1(i) is valid for f_n^* then it is enough to prove the uniform integrability of f_n^*:

$$\lim_{a\to\infty} \sup_n \int_{\{f_n^* > a\}} |f_n^*|\, dP = 0 \qquad (4.1)$$

Since the values $a_n(z_{i_j})$ $(j=1,\dots,n_1)$ of jumps of the K-M estimator in uncensored observations form a nondecreasing sequence (Efron (1967)) then

$$(4.2)\quad |f_n^*(x)| \leq \frac{K(0)\,k(n_1)}{R(n_1)} \max_{\substack{z_i \in S(x,R(n_1)) \\ \delta_i = 1}} (a_n(z_i)) \leq \frac{K(0)\,k(n_1)}{R(n_1)} \frac{1}{n_1-i+1}$$

where i is the rank of the last uncensored observation in $S(x,R(n_1))$ in the subsample $Z_{i_1},\dots,Z_{i_{n_1}}$ of uncensored observations.

Therefore, it is enough to prove the uniform integrability of the sequence

$$(4.3)\qquad h_{n_1}(x) = \frac{K(0)k(n_1)}{R(n_1)(n_1-i+1)}$$

Since $\int_{\{h_{n_1} > a\}} |h_{n_1}|\, dP = \sum_{i=1}^{n} \int_{\{h_i > a\}} |h_i| dP\, p_{in}$

where $p_{in} = P(n_1=i)$, we have to show uniform integrability of estimator h_n based on the i.i.d. sample with distribution F_1 of size n.

Put $\alpha = F_1(x)+2\varepsilon$ and let L be an arbitrary positive constant. We have

$$\int_{\{|h_m|>a\}} |h_n| dP = \int_{\{|h_n|>a,\, i\leq[\alpha n]\}} |h_m| dP + \int_{\{|h_n|>a,\, i>[\alpha n]\}} |h_n| dP$$

On the set $\{i \leq [\alpha n]\}$ the sequence h_n is majorized by $k(n)/(nR(n)(1-\alpha))$ which is uniformly integrable (Theorem 4 in Moore, Yackel (1976)). We shall consider the second integral. The density of $R(n)$ for fixed i is given by

$$(4.4)\quad n\binom{n-1}{k-1}\binom{n-k}{i-k}W^{k-1}(r)\,(P((Z|\delta=1)<x-r)^{i-k}P((Z|\delta=1)>x+r)^{n-i}dW(r)$$

$$= n\binom{n-1}{k-1}\,W^{k-1}(r)t_{in}(r)\,dW(r)$$

where $W(r)$ is the probability that $|(Z|\delta=1)-x|\leq r$.

Put $c_{in} = K(0)\cdot k(n)/(a\cdot(n-i+1))$. Then

$$(4.5)\qquad \int_{\{|h_n|>a,\, i>[\alpha n]\}} |h_n|\,dP = n\binom{n-1}{k-1}\sum_{i=[\alpha n]}^{n}\int_0^{c_{in}} \frac{k(n)}{(n-i+1)r}\,t_{in}W^{k-1}dW(r) <$$

$$< \binom{n}{k}k^2\int_0^{c_{nn}}\sum_{i=[\alpha n]}^{n} t_{in}(r)\,\frac{1}{r}W^{k-1}(r)\,dW(r)$$

But

$$(4.6)\qquad \sum_{i=[\alpha n]}^{n} t_{in}(r) \leq \sum_{i=[\alpha n]}^{n}\binom{n-k}{i-k}P((Z|\delta=1)\leq x)^{i-k}P((Z|\delta=1)>x)^{n-i}$$

$$= P(X^{n-k}_{([\alpha n]-k)} \leq x)$$

Then by Lemma 3 the considered integral is less than

$$(4.7)\quad \binom{n}{k}^2 e^{-c(n-k)}\int_0^{c_{nn}}\frac{W^{k-1}(r)}{r}dW(r) < L\binom{n}{k}k^2e^{-c(n-k)}\int_0^{c_{nn}}W^{k-2}dW(r) =$$

$$= L\binom{n}{k}\frac{k^2}{k-1}\,e^{-c(n-k)}(c_{nn})^{k-1} = L\binom{n}{k}\frac{k^k}{n^k}\left(\frac{K(0)}{a}\right)^{k-1}e^{-c(n-k)}n^k.$$

(the last inequality holds since f_1 is bounded).

It is enough to prove that

$$(4.8)\qquad \forall c\ \exists\delta_0\,\forall\delta<\delta_0\ \sup_n\ \delta^{k-1}\binom{n}{k}\left(\frac{k}{n}\right)^k < \varepsilon$$

and

$$(4.9)\qquad n^ke^{-c(n-k)}\to 0.$$

But (4.8) follows easily from Stirling formula

$$\delta^{k-1}\binom{n}{k}\left(\frac{k}{n}\right)^k < \frac{\delta^{k-1}k^k}{k!} \quad \frac{\delta^{k-1}\ e^k}{(2\pi k)^{1/2}} \to 0$$

for $\delta<1/e$ while (4.9) is the consequence of $\log n \cdot k(n)/n \to 0$.

REFERENCES

[1] Bjerve, S. (1977). Error bounds for linear combinations of order statistics. Ann. Statist. 5 357-269.

[2] Devroye, L.P. and Wagner, T.J. (1977). The strong uniform consistency of nearest neighbor density estimates. Ann. Statist. 5 536-540.

[3] Efron, B. (1967). The two-sample problem with consored data. Proc. Fifth Berkeley Symp. Math. Statist. Probability 4 831-853.

[4] Földes, A. and Rejtő, L. (1981). A LIL type result for the product limit estimator. Z. Wahrsch. Verw. Gebiet. 56 75-86

[5] Földes, A., Rejtő, L. and Winter B.B. (1981). Strong consistency properties of nonparametric estimators for randomly censored data, II: Estimation of density and failure rate. Period. Math. Hungar. 12 15-29.

[6] Mack, Y.P. and Rosenblatt, M. (1979). Multivariate k-nearest neighbor density estimates. J. of Mult. Analysis 9 1-15.

[7] Mack, Y.P. (1980). Asymptotic normality of multivariate k-NN density estimates. Sankhya A 42 53-63.

[8] Mack, Y.P. (1982). On a goodness-of-fit problem of some nonparametric density estimates. Scand. J. Statist. 8 179-182.

[9] Moore, D.S. and Yackel, J.W. (1976). Large sample properties of nearest neighbor density function estimators in Statistical Decision Theory and Related Topics (S.S. Gupta and D.S. Moore, eds.) Academic Press, New York.

[10] Moore, D.S. and Yackel, J.W. (1977). Consistency properties of nearest neighbor density estimates. Ann. Statist. 5 143-154.

[11] Nadaraya, E.A. (1965). On nonparametric estimates of density functions and regression curves. Theor. Probability. Appl. 10 186-190.

[12] Rosenblatt, M. (1971). Curve estimates. Ann. Math. Statist. 42 1815-1842

[13] Silverman, B.W. (1978). Weak and strong uniform consistency of the kernel estimates of a density and its derivatives. Ann. Statist. 6 177-184.