Chin. Ann. of Math.
6B (4) 1985

A NECESSARY AND SUFFICIENT CONDITION FOR CONVERGENCE OF ERROR PROBABILITY ESTIMATES IN *K*-NN DISCRIMINATION

SUN ZHIGANG (孙志刚)*

Abstract

Let(X, θ)be $R^d\times\{1, \cdots, s\}$valued random vector, (X_j, θ_j), $j=1, \cdots, n$, be its observed values, $\theta_{nj}^{(k)}$ be the K-nearest neighbor estimate of θ_j, $R^{(k)}$ be the limit of error probability and $\hat{R}_{nk}\triangleq\frac{1}{n}\sum_{j=1}^{n}I_{(\theta_j\neq\theta_{nj}^{(k)})}$ be the error probability estimate. In this paper it is shown that $\forall\varepsilon>0$, $\exists$ constants $a>0$, $c<\infty$ such that

$$P(|\hat{R}_{nk}-R^{(k)}|>\varepsilon)<ce^{-an}$$

if add only if there is no unregular atom of (X, θ) defined below and the various convergences $\hat{R}_{nk}\to R^{(k)}$ are equivalent.

Let (X, θ), $(X_1, \theta_1),\cdots$, (X_n, θ_n) be independent identically distributed random vectors from $R^d\times\{1, \cdots, s\}$, where $d\geqslant1$, $s\geqslant2$ and $Z^n\triangleq\{(X_j, \theta_j), j=1, \cdots, n\}$ are observed values of(X, θ). Let μ be the probability measure of X and

$$P_i(x)\triangleq P(\theta=i\,|\,X=x),\ \text{for}\ x\in R^d,\ i=1, \cdots, s.$$

The k-nearest neighbor estimate $\theta_n^{(k)}$ of θ, introduced by E. Fix and J. L. Hodges[1], is defined as follows: arrange $\|X_j-X\|$, $j=1, \cdots, n$ in increasing order $\|X_{R_1}-X\|\leqslant\cdots\leqslant\|X_{R_n}-X\|$, where $\|X_j-X\|$ is the usual Euclidean distance in R^d between X_j and X; put $i<j$ when $\|X_{R_i}-X\|=\|X_{R_j}-X\|$ and $R_i<R_j$; set $\theta_n^{(k)}$ equal to the integer which has a majority vote among $\theta_{R_1}, \cdots, \theta_{R_k}$; in the case of a voting tie, set $\theta_n^{(k)}$ equal, with same probability, to each integer which has a majority vote. We write the error probability $R_n^{(k)}\triangleq P(\theta_n^{(k)}\neq\theta)$ and the conditional error probability $L_n^{(k)}\triangleq P(\theta_n^{(k)}\neq\theta\,|\,Z^n)$.

It is known that there exists $R^{(k)}\triangleq\lim\limits_{n\to\infty}R_n^{(k)}$ and under some conditions there exists $\lim\limits_{n\to\infty}L_n^{(k)}=R^{(k)}$ a. s. and $R^{(1)}=1-\sum\limits_{i=1}^{s}EP_i^2(X)$ (See, for example, [2, 3] for $k=1$. The posterior error probability $L_n^{(k)}$ of $\theta_n^{(k)}$ for given Z^n is very interesting from a practical point of view, since one can only work with the "training sample" Z^n at his disposal. But it is impossible to get the exact probability distribution of $L_n^{(k)}$

Manuscript received October 21, 1983.
* Department of Mathematics, Hangzhou Normal College, Hongzhou, China.

when that of (X, θ) is unknown. Many mathematicians have studied the convergence of the error probability estimate of $L_n^{(k)}$

$$\hat{R}_{nk} \triangleq \frac{1}{n}\sum_{j=1}^{n} I_{(\theta_j \neq \theta_{nj}^{(k)})},$$

where I_A is the indicator function of set A and $\theta_{nj}^{(k)}$ is the k-nearest neighbor estimate of θ_j based on $\{(X_i, \theta_i), i=1, \cdots, n, i\neq j\}$. Recently, Doctor Bai Zhidong proved

Theorem. *If μ is nonatomic, then $\forall \varepsilon>0$, $\exists$ constants $a>0$, $c<\infty$ independent of n such that*

$$P(|\hat{R}_{nk}-R^{(k)}|\geq\varepsilon)<ce^{-an}.$$

In this paper we introduce the following

Definition. *A point $x\in R^d$ is called a regular atom of (X, θ) if x is an atom of μ and there exists a nonempty subset $\{i_1, \cdots, i_{g(x)}\}$ of set $\{1, \cdots, s\}$ such that*

$$P(\theta=i_m|X=x)=\frac{1}{g(x)} \quad m=1,\cdots, g(x).$$

The goal of the present paper is to prove

Theorem 1. *The following conditions are equivalent each to other:*

(i) *there is no unregular atom in the distribution of (X, θ),*

(ii) $\forall\varepsilon>0$, $\exists$ *constants $a>0$, $c<\infty$ independent of n such that*

$$P(|\hat{R}_{nk}-R^{(k)}|\geq\varepsilon)<ce^{-an},$$

(iii) $\hat{R}_{nk}\to R^{(k)}$ a. s. $(n\to\infty)$,

(iv) $\forall\alpha>0$ $\hat{R}_{nk}\xrightarrow{L_\alpha} R^{(k)}$ $(n\to\infty)$,

(v) $\exists\alpha>0$, $\hat{R}_{nk}\xrightarrow{L_\alpha} R^{(k)}$ $(n\to\infty)$,

(vi) $\hat{R}_{nk}\xrightarrow{P} R^{(k)}$ $(n\to\infty)$,

(vii) $\exists$ *constant r such that* $\hat{R}_{nk}\xrightarrow{F} r$.

In this paper we denote by a a positive constant and by c a finite constant. Both a and c are independent of n and take their own values in each formula. First we shall show the following

Lemma 1. *Let $X_1, \cdots, X_n$ be iid R^d valued random vectors and $\rho>0$. Then $\forall\varepsilon>0$, $\exists$ constants $a>0$, $c<\infty$ independent of n such that*

$$P\left(\frac{1}{n}\#\{j\leq n: \|X_j-X_{nj}^{(k)}\|>\rho\}>\varepsilon\right)<ce^{-an},$$

where $X_{nj}^{(k)}$ is the k-th nearest neighbor of X_j among $\{X_i, i=1, \cdots, n, i\neq j\}$.

Proof of Lemma 1. $\forall\varepsilon>0$, $\exists$ constant $M>0$ such that $P(\|X_1\|>M)<\varepsilon/2$. Since

$$\frac{1}{n}\#\{j\leq n: \|X_j-X_{nj}^{(k)}\|>\rho\}=\frac{1}{n}\#\{j\leq n: \|X_j-X_{nj}^{(k)}\|>\rho, \|X_j\|>M\}$$
$$+\frac{1}{n}\#\{j\leq n: \|X_j-X_{nj}^{(k)}\|>\rho, \|X_j\|\leq M\}\triangleq\Sigma_1+\Sigma_2, \quad (1)$$

where $\#A$ denotes the number of elements of set A, by Hoeffding's inequality, $\exists$ constants $a>0$, $c<\infty$ independent of n such that

$$P(\Sigma_1>\varepsilon/2)\leqslant P\left(\frac{1}{n}\#\{j\leqslant n:\|X_j\|>M\}>\varepsilon/2\right)$$
$$\leqslant P\left(\left|\frac{1}{n}\#\{j\leqslant n:\|X_j\|>M\}-P(\|X_j\|>M)\right|\right.$$
$$\left.>\varepsilon/2-P(\|X_j\|>M)\right)<ce^{-an}. \tag{2}$$

To consider Σ_2 we suppose that $j\leqslant n$, $\|X_j-X_{nj}^{(k)}\|>\rho$ and $\|X_j\|\leqslant M$. Write $B_j\triangleq\{x:\|x-X_j\|<\rho/2\}$. It is not difficult to see that each B_j intersects at most with $k-1$ of balls B_i, otherwise, it contradicts $\|X_j-X_{nj}^{(k)}\|>\rho$. Thus $\forall x:\|x\|\leqslant M$, there are at most k of balls B_j containing x. Clearly, each $B_j\subset\{x:\|x\|\leqslant M+\rho/2\}$. So

$$\#\{B_j:j\leqslant n,\ \|X_j-X_{nj}^{(k)}\|>\rho,\ \|X_j\|\leqslant M\}\leqslant k\left[\left(\frac{M+\rho/2}{\rho/2}\right)^d\right]$$

and for sufficiently large n,

$$\Sigma_2\leqslant\varepsilon/2. \tag{3}$$

From (1)—(3) this lemma follows.

Proof of Theorem 1. Clearly, (ii)$\Rightarrow$(iii), (iv)$\Rightarrow$(v)$\Rightarrow$(vi)$\Rightarrow$(vii). Since $|\hat{R}_{nk}|\leqslant1$, (iii)$\Rightarrow$(iv). Thus to complete the proof it remains to show that (i)$\Rightarrow$(ii) and (vii)$\Rightarrow$(i).

Step 1 (i)$\Rightarrow$(ii). For brevity of the proof, we may suppose $k=1$, without loss of generelity, and this is to discuss [the nearest neighbor discrimination. We denote by A, B respectively the set of regular atoms of $(X,\ \theta)$ with $g(x)=1$ and $g(x)>1$. $H\triangleq A\cup B$. Clearly, the set H is finite or denumerable. Write

$$A\triangleq\{a_1,\ a_2,\cdots\},\ B\triangleq\{b_1,\ b_2,\ \cdots\}\text{ and }H\triangleq\{h_1,\ h_2,\cdots\}.$$

Denote $P(X\in A)$, $P(X\in B)$, $P(X\in H)$ and $P(X=b)$ respectively by $P(A)$, $P(B)$, $P(H)$ and $P(b)$.

It is easy to see that

$$R\triangleq R^{(1)}=1-\sum_{i=1}^{s}\left(\int_A+\int_B+\int_{R^d\setminus H}\right)P_i^2(x)d\mu$$
$$=1-P(A)-\sum_{m=1}^{\infty}\frac{P(b_m)}{g(b_m)}-\sum_{i=1}^{s}\int_{R^d\setminus H}P_i^2(x)d\mu, \tag{4}$$

$$\hat{R}_n\triangleq\hat{R}_{n1}=\frac{1}{n}\Big(\sum_{x_j\in A}+\sum_{x_j\in B}+\sum_{x_j\notin H}\Big)I_{(\theta_j\neq\theta_{nj})}\triangleq\Sigma_3+\Sigma_4+\Sigma_5. \tag{5}$$

From now on we denote $\sum_{j=1}^{n}$ by $\sum_j$.

In the case of $H=\emptyset$ Step 1 follows from Bai's theorem. So we may suppose $H\neq\emptyset$.

In the case of $A=\emptyset$, $\Sigma_3=0$. To consider Σ_3, we may suppose that $A\neq\emptyset$. Then $\forall\varepsilon>0$, $\exists$ constant M such that $\sum_{m=M+1}^{\infty}P(X=a_m)<\varepsilon/2$.

Thus

$$P(\Sigma_3>\varepsilon)\leqslant P\left(\frac{1}{n}\sum_{m=1}^{M}\sum_{x_j=a_m}I_{(\theta_j\neq\theta_{nj})}>\varepsilon/2\right)+P\left(\frac{1}{n}\sum_{m+M=1}^{\infty}\sum_{x_j=a_m}1>\varepsilon/2\right)$$
$$\leqslant P\left(\frac{M}{n}>\varepsilon/2\right)+P\left\{\left|\frac{1}{n}\sum_{m=M+1}^{\infty}\sum_{x_j=a_m}1-\sum_{m=M+1}^{\infty}P(X=a_m)\right|>\varepsilon/2-\sum_{m=M+1}^{\infty}P(X=a_m)\right\}.$$

The last inequality follows from the fact that by the definition of the regular atom of (X, θ) with $g(a_m)=1\ \forall m$

$$\sum_{x_j=a_m}I_{(\theta_j\neq\theta_{nj})}\leqslant 1\quad \text{a. s.}$$

So for sufficiently large n, $P\left(\frac{M}{n}>\frac{\varepsilon}{2}\right)=0$ and by Hoeffding's inequality $\forall\varepsilon>0$, $\exists$ constants $a>0$, $c<\infty$ independent of n such that

$$P(\Sigma_3>\varepsilon)<ce^{-an}. \tag{6}$$

In the case of $B=\emptyset$, $\Sigma_4=0$. To consider Σ_4 we may suppose that $B\neq\emptyset$. Similarly, $\forall\varepsilon>0$, $\exists$ constant M such that $\sum_{m=M+1}^{\infty}P(X=b_m)<\varepsilon/4$. By Borel's strong law of large numbers, $j(m)\triangleq\min\{j\colon X=b_m\}$ is defined with probability one. Now in the case of $X_j=b_m$, $j\neq j(m)$ we have $\theta_{nj}=\theta_{j(m)}$ and

$$P(|\Sigma_4-P'|>\varepsilon)\leqslant P\left(\left|\frac{1}{n}\sum_{m=1}^{M}\sum_{x_j=b_m}I_{(\theta_j\neq\theta_{nj})}-\sum_{m=1}^{M}\frac{g(b_m)-1}{g(b_m)}P(X=b_m)\right|>\varepsilon/2\right)$$
$$+P\left(\frac{1}{n}\sum_{m=M+1}^{\infty}\sum_{x_j=b_m}1>\varepsilon/4\right)$$
$$\leqslant\sum_{m=1}^{M}P\left(\left|\frac{1}{n}\sum_{x_j=b_m}I_{(\theta_j\neq\theta_{j(m)})}-\frac{g(b_m)-1}{g(b_m)}P(X=b_m)\right|>\frac{\varepsilon}{4M}\right)$$
$$+P\left(\frac{M}{n}>\frac{\varepsilon}{4}\right)+P\left\{\left|\frac{1}{n}\sum_{m=M+1}^{\infty}\sum_{x_j=b_m}1-\sum_{m=M+1}^{\infty}P(X=b_m)\right|>\frac{\varepsilon}{4}-\sum_{m=M+1}^{\infty}P(X=b_m)\right\},$$

where $P'\triangleq\sum_{m=1}^{\infty}\frac{g(b_m)-1}{g(b_m)}P(b_m)$. So for sufficiently large n, $P\left(\frac{M}{n}>\frac{\varepsilon}{4}\right)=0$ and by Hoeffding's inequality $\forall\varepsilon>0$, $\exists$ constants $a>0$, $c<\infty$ independent of n such that

$$P(|\Sigma_4-P'|>\varepsilon)<ce^{-an}. \tag{7}$$

To consider Σ_5, we write

$$X'\triangleq\begin{cases}X, & X\bar{\in}H,\\ h_1, & X\in H,\end{cases}\quad\text{and}\quad X'_j\triangleq\begin{cases}X_j, & X_j\bar{\in}H,\\ h_1, & X_j\in H.\end{cases}$$

Clearly, (X', θ), (X'_1, θ_1), $\cdots$, (X'_n, θ_n) are iid random vectors, and there is only one atom h_1 of X' and $P(X'=h_1)=P(H)$.

$$\Sigma_5=\frac{1}{n}\Big(\sum_{x'_j\neq h_1,x_{nj}\neq h_1}+\sum_{x'_j\neq h_1,x'_{nj}=h_1}\Big)I_{(\theta_j\neq\theta_{nj})}\triangleq\Sigma_6+\Sigma_7. \tag{8}$$

By Lemma 1 of[4], there exists constant m independent of n such that

$$\Sigma_7 \leqslant \frac{1}{n} \sum_{x_j' \neq h_1, x_{nj}' = h_1} 1 \leqslant \frac{m}{n}. \tag{9}$$

To consider Σ_6 we introduce random vectors (X'', θ''), (X_1'', θ_1''), $\cdots$, (X_n'', θ_n'') as follows: write $B(\rho) \triangleq \{x: \|x-h_1\| < \rho\}$. By continuity of probability $\forall \varepsilon > 0$, $\exists \rho > 0$ such that

$$P(X' \in B(\rho)) \leqslant P(X' \in B(2\rho)) < P(H) + \varepsilon/2. \tag{10}$$

Write $(X'', \theta'') = (X', \theta)$ when $X' \overline{\in} B(\rho)$, otherwise $\theta'' = 1$ and X'' is uniformly distributed in $B(\rho/2)$ and $P(X'' \in B(\rho/2)) = P(X' \in B(\rho))$. Similarly, we can defined (X_j'', θ_j''), $j=1, \cdots, n$, based on (X_j', θ_j), $j=1, \cdots, n$, such that (X'', θ''), (X_j'', θ_j''), $j=1, \cdots, n$, are iid. Now there is no atom of X'' and $P(X'' \in B(\rho) \backslash B(\rho/2)) = 0$. Thus

$$\begin{aligned}
|\Sigma_6 - R + P'| \leqslant & \left| \Sigma_6 - \frac{1}{n} \sum_{\|x_j'' - x_{nj}''\| < \rho/2;\, x_j'', x_{nj}'' \overline{\in} B(\rho)} I_{(\theta_j'' \neq \theta_{nj}'')} \right| + \left| \frac{1}{n} \sum_j I_{(\theta_j'' \neq \theta_{nj}'')} - R + P' \right| \\
& + \frac{1}{n} \sum_{\|x_j'' - x_{nj}''\| > \rho/2} 1 + \frac{1}{n} \sum_{\|x_j'' - x_{nj}''\| < \rho/2;\, x_j'', x_{nj}'' \in B(\rho)} I_{(\theta_j'' \neq \theta_{nj}'')} \\
= & \Sigma_8 + \Sigma_9 + \Sigma_{10} + \Sigma_{11}.
\end{aligned}$$

Since $\{j \leqslant n: \|X_j'' - X_{nj}''\| < \rho/2;\ X_j'', X_{nj}'' \overline{\in} B(\rho)\} \supset \{j \leqslant n: \|X_j' - X_{nj}'\| < \rho/2;\ X_j', X_{nj}' \overline{\in} B(\rho)\}$ by Lemma 1 (for $k=1$) and Hoeffding's inequality $\forall \varepsilon > 0$, $\exists$ constants $a > 0$, $c < \infty$ independent of n such that

$$\begin{aligned}
P(\Sigma_8 < \varepsilon) \leqslant & P\left(\frac{1}{n} \sum_{\|x_j' - x_{nj}'\| \geqslant \rho/2} 1 > \frac{\varepsilon}{2}\right) + P\left(\frac{1}{n} \sum_{x_j' \in B(2\rho) \backslash \{h_1\}} 1 > \varepsilon/2\right) \\
\leqslant & P\left(\frac{1}{n} \sum_{\|x_j' - x_{nj}'\| > \rho/2} 1 > \varepsilon/2\right) + P\left\{\left|\frac{1}{n} \sum_{x_j' \in B(2\rho) \backslash \{h_1\}} 1 - P(X' \in B(2\rho) \backslash \{h_1\})\right| \right. \\
& \left. > \frac{\varepsilon}{2} - P(X' \in B(2\rho) \backslash \{h_1\}) \right\} \leqslant c e^{-an}.
\end{aligned} \tag{12}$$

Write $X' \sim \mu'$, $X'' \sim \mu''$. $R'' \triangleq 1 - \sum_{i=1}^{s} E P^2(\theta'' = i \mid X'')$. Then

$$\begin{aligned}
|R'' - R + P'| = & \left| 1 - \sum_{i=1}^{s} \left(\int_{B(\rho)} + \int_{R^d \backslash B(\rho)} \right) P^2(\theta'' = i \mid X'' = x) d\mu'' - R + P' \right| \\
= & \left| 1 - P(X'' \in B(\rho)) - \sum_{i=1}^{s} \int_{R^d \backslash B(\rho)} P^2(\theta = i \mid X' = x) d\mu' - R + P' \right| \\
\leqslant & \left| 1 - P(H) - \sum_{i=1}^{s} \int_{R^d \backslash H} P_i^2(x) d\mu - R + P' \right| \\
& + |P(H) - P(X'' \in B(\rho))| + P(X' \in B(\rho) \backslash \{h_1\}) < \varepsilon.
\end{aligned}$$

Thus by Bai's theorem $\forall \varepsilon > 0$, $\exists$ constants $a > 0$, $c < \infty$ independent of n such that

$$P(\Sigma_9 > \varepsilon) \leqslant P\left(\left| \frac{1}{n} \sum_j I_{(\theta_j'' \neq \theta_{nj}'')} - R'' \right| > \varepsilon - |R'' - R + P'| \right) < c e^{-an}. \tag{13}$$

By Lemma 1 $\forall \varepsilon > 0$, $\exists$ constants $a > 0$, $c < \infty$ independent of n such that

$$P(\Sigma_{10} > \varepsilon) < c e^{-an}. \tag{14}$$

Since $P(X'' \in B(\rho) \backslash B(\rho/2)) = 0$ and in the case of X_j'', $X_{nj}'' \in B(\rho/2)$, we have

$$\theta_j'' = \theta_{nj}'' = 1, \ \Sigma_{11} = 0 \ \text{ a. s.} \tag{15}$$

By (5)—(9) and (11)—(15), taking the largest number c and the smallest number a, we have $P(|\hat{R}_n - R| > 10\varepsilon) < ce^{-an}$. This terminates the proof of Step 1.

Step 2 (vii)$\Rightarrow$(i). By contradiction.

Assume that there exists an unregular atom $x_0 \in R^d$ of (X, θ), i_1, i_2 such that $P(\theta \neq i_1 | X = x_0) > P(\theta \neq i_2 | X = x_0) > 0$ and (vii) holds. Then

$$\hat{R}_{nk} = \frac{1}{n}\Big(\sum_{X_j = x_0} + \sum_{X_j \neq x_0}\Big) I_{(\theta_j \neq \theta_{nj}^{(k)})} = \Sigma_{12} + \Sigma_{13}. \tag{16}$$

Let $A_l \triangleq \{(X_j, \theta_j) = (x_0, i_l),\ j = 1, \cdots, k\}$, $l = 1, 2$. Clearly,

$$P(A_1) = \prod_{j=1}^{k} P(X_j = x_0) P(\theta_j \neq i_1 | X_j = x_0)$$

$$> \prod_{j=1}^{k} P(X_j = x_0) P(\theta_j \neq i_2 | X_j = x_0) = P(A_2) > 0.$$

Let $$\delta \triangleq P(X = x_0)\{P(\theta \neq i_1 | X = x_0) - P(\theta \neq i_2 | X = x_0)\} > 0.$$

By (vii), there exists a constant r such that $\hat{R}_{nk} \xrightarrow{P} r$. Thus there exists a subsequence $\hat{R}_{n_m k} \to r$ a. s. Since $\theta_{nj}^{(k)} = i_l$ on A_l, when $X_j = x_0$, $j > k$. So by Borel's strong law of large numbers the corresponding subsequence $\Sigma_{12} \to P(X = x_0) P(\theta \neq i_l | X = x_0)$ a. s.

Thus $$\Sigma_{13} \to r - P(X = x_0) P(\theta \neq i_l | X = x_0) \quad \text{a. s. } (l = 1, 2). \tag{17}$$

Denote $B \triangleq \{x: \Delta > \|x - x_0\| > 0\}$. Clearly, $\exists \Delta > 0$ such that $P(X \in B) < \frac{\delta}{3}$. We have

$$\Sigma_{13} = \frac{1}{n}\Big(\sum_{X_j \in B} + \sum_{X_j \neq x_0, X_j \overline{\in} B}\Big) I_{(\theta_j \neq \theta_{nj}^{(k)})} \triangleq \Sigma_{14} + \Sigma_{15}, \tag{18}$$

$$\Sigma_{14} \leqslant \frac{1}{n} \sum_{X_j \in B} 1 \to P(X \in B) \quad \text{a. s. } (n \to \infty), \tag{19}$$

$$\Sigma_{15} = \frac{1}{n}\Bigg(\sum_{X_j \neq x_0, X_j \overline{\in} B, \|X_j - X_{nj}^{(k)}\| \geqslant \Delta} + \sum_{X_j \neq x_0, X_j \overline{\in} B, \|X_j - X_{nj}^{(k)}\| < \Delta}\Bigg) I_{(\theta_j \neq \theta_{nj}^{(k)})} \triangleq \Sigma_{16} + \Sigma_{17}. \tag{20}$$

By Lemma 1 $\Sigma_{16} \leqslant \frac{1}{n} \sum_{\|X_j - X_{nj}^{(k)}\| \geqslant \Delta} 1 \leqslant \frac{\delta}{3}$ a. s. for n sufficiently large. (21)

For $(X_j, \theta_j) = (x_0, i_l)$, $j = 1, \cdots, k$, we denote $g_{ln}(X_{k+1}, \cdots, X_n;\ \theta_{k+1}, \cdots, \theta_n) \triangleq \Sigma_{17}$ $(l = 1, 2)$. Since $X_{nj}^{(k)} \neq x_0$ for $X_j \neq x_0$, $X_j \overline{\in} B$, $\|X_j - X_{nj}^{(k)}\| < \Delta$, $g_{1n} = g_{2n}$. (22)

By (18)—(22) the difference of values of Σ_{13} on A_1 and A_2 is less than $\frac{2\delta}{3}$ a. s. for n sufficiently large. This contradicts (17). Thus Step 2 is proved and the proof of Theorem 1 is completed.

References

[1] Fix, E. & Hodges, J. L., Discriminatory Analysis, Nonparametric Discrimination, Consistency Properties, Project 21-49-004, Report No. 4, School of Aviation Medicine, Randolph Field, Texas.

[2] Dtvroye, L., On the asymptotic probability of error in nonparametric discrimination, *Ann. Statist.*, **9**(1981), 1320—1327.

[3] Wagner, T. J., Convergence of the nearest neighbor rule, *IEEE Trans. Inform. Theory*, 17, 566—571.

[4] Bai Zhidong(白志东), The strong Consistency of Error Probabilty Estimates in K-NN Discrimination, *Chin. Ann. of Math.*, **6B**: 3(1985), 299—308.