应用概率统计 第二卷
第三期 1986年8月

Chinese Journal of Applied Probability
and Statistics Vol. 2. No. 3. Aug. 1986

下标降序法最近邻判别分析

陈桂景 陈 刚
（安徽大学）

§1 引 言

设$(X_1, \theta_1), \cdots, (X_n, \theta_n), (X, \theta)$为在$R^d\times\{1, \cdots, M\}$上取值的i.i.d.随机向量. 问题是要利用$X$的观察值及历史样本$(X_i, \theta_i), i=1, \cdots, n$对类别变量$\theta$进行判别. 假定在$R^d$上给定了某一距离函数$\rho(\cdot, \cdot)$(比如欧氏距离等)，那么可按照诸$X_i$与$X$的距离由小到大把诸$X_i$重新排列为$X_{R_1}, X_{R_2}, \cdots, X_{R_n}$，相应的$\theta_i$也被排列为$\theta_{R_1}, \theta_{R_2}, \cdots, \theta_{R_n}$. 若采用$\theta_{R_1}$来判别$\theta$，这就是所谓的最近邻判别法. Derroye[1], Wagner[2], Fritz[3], 陈希孺[4]及白志东[5]在X的分布Q无原子的假定下研究了最近邻判别分析，得到了许多很好的结果. 显然，在实际问题中常会碰到Q含有原子的情况. 陈桂景、孔繁超[6], [7]及白志东、陈希孺、陈桂景[8]进一步地考虑了在Q含原子的情况下的最近邻判别. 当Q含有原子时，诸X_i将以正概率与X距离相等，因此如上提出的最近邻判别程式就不确定了，为消除这种不确定性，我们在[8]中考虑了如下三种方法：

(i) 下标升序法. 记

$$k_n=\min\{j: \rho(X_j, X)=\min_{1\leqslant i\leqslant n}\rho(X_i, X), j\leqslant n\},$$

则用θ_{k_n}来判别θ，并记之为$\theta_n^{(1)}$.

(ii) 下标降序法. 记

$$k_n=\max\{j: \rho(X_j, X)=\min_{1\leqslant i\leqslant n}\rho(X_i, X), j\leqslant n\},$$

则用θ_{k_n}来判别θ，并记之为$\theta_n^{(2)}$.

(iii) 等权随机化法. 记

$$J^{(n)}=\{j: \rho(X_j, X)=\min_{1\leqslant i\leqslant n}\rho(X_i, X), j\leqslant n\}, N^{(n)}={}^{\#}J^{(n)},$$

其中$^{\#}(\cdot)$表示集合$(\cdot)$中元素个数. 定义θ的判别值为一随机变量，它以$1/N^{(n)}$的概率取θ_j，对每一$j\in J^{(n)}$. 这种判别记为$\theta_n^{(3)}$.

这三种判别的错判概率、条件错判概率分别记为

$$R_n^{(i)}=P(\theta_n^{(i)}\neq\theta),\ T_n^{(i)}=P(\theta^{(i)}\neq\theta\,|\,X^n) \tag{1}$$

$$L_n^{(i)}=P(\theta_n^{(i)}\neq\theta\,|\,Z^n) \tag{2}$$

对$i=1, 2, 3$. 其中$X^n\triangleq(X_1, \cdots, X_n)$, $Z^n\triangleq((X_i, \theta_i), i=1, \cdots, n)$. 又记

$$R=1-\sum_{j=1}^{M}E[P^2(\theta=j\,|\,X)] \tag{3}$$

本文1985年4月9日收到.

陈、孔[6,7]及白、陈、陈[8]分别研究了在下标升序法与等权随机化法下 $R_n^{(4)}$, $T_n^{(4)}$, $L_n^{(4)}$ 等诸量的收敛性及其与 R 值的关系，本文进一步在下标降序法下研究了这些量的收敛性质. 我们发现，在这三种判别法下，这些量的收敛性虽有类似之处，但也存在着明显的差别.

§2 若干结果及其证明

现来讨论在下标降序法下诸量 $L_n^{(2)}$, $T_n^{(2)}$, $R_n^{(2)}$ 的收敛性质. 先考虑指标 X 为离散变量时的情形.

设 X 的分布 Q 为纯原子的，记其原子全体为$\{a_1, a_2, \cdots\}$. 又记 $q_i=P(X=a_i)$, $q_{ij}=P(\theta=j|X=a_i)$, $p_{ij}=P(X=a_i, \theta=j)$, $j=1, \cdots, M$; $i=1, 2, \cdots$.

定理 1 若 Q 为纯原子的，则在下标降序法判别下，下述三个命题互相等价:

(I) $\lim\limits_{n\to\infty} L_n^{(2)}=R$, a.s.;

(II) 对每一原子 a_i, 存在集 $J\triangleq\{1, \cdots, M\}$ 的一个非空子集 $J_i=\{j\langle i, 1\rangle, \cdots, j\langle i, s(i)\rangle\}$, 使得有

$$0<P_{ij\langle i,1\rangle}=\cdots=P_{ij\langle i,s(i)\rangle},\ P_{ij}=0,\ 对\ j\overline{\in}J_i;$$

(III) 对每一原子 a_i, 有

$$P_{ij}P_{ik}(P_{ij}-P_{ik})^2=0,\ 1\leqslant j<k\leqslant M. \tag{4}$$

当上述命题不成立时，$L_n^{(2)}$ a.s. 发散.

证明 因证明较长，我们仅写出证明的主要步骤.

(a) (II)$\Leftrightarrow$(III)证明容易，略去.

(b) 由 Hoeffding 不等式(见[10])，可以证明

$$P\{X_n=a_i,\ \text{i.o. 关于}\ n;\ i=1, 2, \cdots\}=1.$$

假若记 $N_{n,i}={}^{\#}\{j: X_j=a_i, j\leqslant n\}$,

$$A=\{(x_1, x_2, \cdots): N_{n,i}\to\infty,\ 当\ n\to\infty;\ i=1, 2, \cdots\},$$

则有 $P(A)=1$.

(c) 由 $L_n^{(2)}$ 之定义有

$$\begin{aligned}L_n^{(2)}&=1-\sum_{j=1}^{M}P\{\theta_n^{(2)}=j,\ \theta=j|Z^n\}=1-\sum_{i=1}^{\infty}\sum_{j=1}^{M}P\{\theta_n^{(2)}=j,\ X=a_i,\ \theta=j|Z^n\}\\&=1-\sum_{i=1}^{\infty}\sum_{j=1}^{M}I_{(\theta_{n,i}^{(2)}=j)}\,p_{ij},\end{aligned}\tag{5}$$

其中 $\theta_{n,i}^{(2)}$ 表示当 $X=a_i$, Z^n 给定时最近邻判别 $\theta_n^{(2)}$ 之值，并用到了如下事实:

$$\begin{aligned}&P\{\theta_n^{(2)}=j,\ X=a_i,\ \theta=j|Z^n\}\\&=P\{X=a_i|Z^n\}P\{\theta_n^{(2)}=j,\ X=a_i,\ \theta=j|Z^n,\ X=a_i\}\\&=P\{X=a_i\}I_{(\theta_{n,i}^{(2)}=j)}P\{\theta=j|Z^n,\ X=a_i\}\\&=I_{(\theta_{n,i}^{(2)}=j)}P\{X=a_i\}P(\theta=j|X=a_i)=I_{(\theta_{n,i}^{(2)}=j)}P_{ij}.\end{aligned}$$

再由 R 之定义有

$$R=1-\sum_{i=1}^{\infty}\sum_{j=1}^{M}q_{ij}^2q_i. \tag{6}$$

对每一给定的$(X_1, X_2, \cdots)\in A$, 以及每一 a_i, 存在自然数子列 $1\leqslant j(1, i)<j(2, i)<j(3, i)<\cdots$, 使得有 $X_{j(k,i)}=a_i$, 对 $k=1, 2, \cdots$但 $X_j\neq a_i$ 当 $j\neq j(k, i)$, $k=1, 2, \cdots$. 对任一自然数

N, 记 $n_N=\max\{j(1,i), i=1,2,\cdots,N\}$. 对每一 $n\geqslant n_N$ 以及每一个 $i=1,\cdots,N$, 记 $j[n,i]=\max\{j(k,i), k=1,2,\cdots; j(k,i)\leqslant n\}$. 于是, 当 $X=a_i$ 时, 在 $(X_1,\cdots,X_n)$ 中的下标最大的最近邻点为 $X_{j[n,i]}$, 故而 $\theta_{n,i}^{(2)}=\theta_{j[n,i]}$. 由(5)式便得

$$L_n^{(2)}=1-\sum_{i=1}^{N}\sum_{j=1}^{M}I_{(\theta_{j[n,i]}=j)}P_{ij}+I_{N,n}, \tag{7}$$

其中
$$I_{N,n}=-\sum_{i=N+1}^{\infty}\sum_{j=1}^{M}I_{(\theta_{n,i}^{(2)}=j)}P_{ij}.$$

(d) 证(II)⇒(I). 给定 $(X_1,\cdots,X_n,\cdots)\in A$. 对任意给定的 $\varepsilon>0$, 存在充分大的 N, 使有 $\sum\limits_{i=N+1}^{\infty}q_i<\varepsilon$. 那么在(7)式中, $|I_{N,n}|\leqslant\sum\limits_{i=N+1}^{\infty}\sum\limits_{j=1}^{M}P_{ij}<\varepsilon$. 由(7)式, 当(II)成立时, 对 $n\geqslant n_N$, 有

$$\begin{aligned}L_n^{(2)}&=1-\sum_{i=1}^{N}\sum_{k=1}^{s(i)}I_{(\theta_{j[n,i]}=j(i,k))}P_{ij(i,k)}+I_{N,n}=1-\sum_{i=1}^{N}P_{ij(i,1)}\sum_{k=1}^{s(i)}I_{(\theta_{j[n,i]}=j(i,k))}+I_{N,n}\\&=1-\sum_{i=1}^{N}P_{ij(i,1)}+I_{N,n},\end{aligned} \tag{8}$$

其中用到了事实 $\sum\limits_{k=1}^{s(i)}I_{(\theta_{j[n,i]}=j(i,k))}=1$, 由(6), (II),

$$\begin{aligned}R&=1-\sum_{i=1}^{N}\sum_{j=1}^{M}(P_{ij}^2/q_i)+I_N=1-\sum_{i=1}^{N}(1/q_i)\sum_{k=1}^{s(i)}P_{ij(i,k)}^2+I_N\\&=1-\sum_{i=1}^{N}(1/q_i)s(i)P_{ij(i,1)}^2+I_N=1-\sum_{i=1}^{N}P_{ij(i,1)}+I_N,\end{aligned} \tag{9}$$

其中用到了事实 $s(i)P_{ij(i,1)}=q_i$, 而 $|I_N|=\left|\sum\limits_{i=N+1}^{\infty}\sum\limits_{j=1}^{M}q_{ij}^2q_i\right|\leqslant\sum\limits_{i=N+1}^{\infty}q_i<\varepsilon$. 由(8), (9), 当 $n\geqslant n_N$ 时有

$$|L_n^{(2)}-R|\leqslant|I_{N,n}|+|I_N|<2\varepsilon.$$

这证明了 $\lim\limits_{n\to\infty}L_n^{(2)}=R$, 当 $(X_1,X_2,\cdots)\in A$. 即有 $L_n^{(2)}\to R$, a.s. 当 $n\to\infty$.

(e) 最后证明, 若(II)不成立⇒$L_n^{(2)}$ a.s. 发散. 对每一原子 a_i, 记 $j(i)=\max\{j: P_{ij}=\max\limits_{1\leqslant k\leqslant M}P_{ik}, j\in J\}$, $j[i]=\max\{j: P_{ij}=\min(P_{ik}>0, k\in J), j\in J\}$. 那么当(II)不成立时, 有

$$\sum_{i=1}^{\infty}P_{ij(i)}-\sum_{i=1}^{\infty}P_{ij[i]}\triangleq4\varepsilon>0 \tag{10}$$

于是存在充分大的 N, 使得有

$$\sum_{i=1}^{N}(P_{ij(i)}-P_{ij[i]})\geqslant3\varepsilon, \tag{11}$$

并且仍有 $\sum\limits_{i=N+1}^{\infty}q_i<\varepsilon$. 对给定的 $(X_1,X_2,\cdots)\in A$, 当 $n\geqslant k_N$ 时, (7)式成立, 用归纳法定义一个自然数子列 $n_N\leqslant n^{(1)}<n^{(2)}<n^{(3)}<\cdots$, 使其中 $n^{(1)}\geqslant n_N$ 任意选定, 而 $n^{(2)}>n^{(1)}$, 且有 $j[n^{(2)},i]>j[n^{(1)},i]$, $i=1,\cdots,N$. 一般地, 当 $n^{(k)}$ 选定后, 再选取 $n^{(k+1)}$, 使有 $j[n^{(k+1)},i]>j[n^{(k)},i]$, $i=1,\cdots,N$ 且 $n^{(k+1)}>n^{(k)}$. 记 $I^{(N,k)}=\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{M}(I_{(\theta_{j[n^{(2k)},i]}=j)}-I_{(\theta_{j[n^{(2k-1)},i]}=j)})P_{ij}$, $B_k=\{|I^{(N,k)}|\geqslant3\varepsilon\}$, $k=1,2,\cdots$. 由(11)式知

$$D_k\triangleq\{\theta_{j[n^{(2k)},i]=j(i)},\ \theta_{j[n^{(2k-1)},i]=j[i],i=1,\cdots,N}\}\subset B_k, \tag{12}$$

若记 $\tilde{P}(\cdot)=P(\cdot|X_1,X_2,\cdots)$. 注意 $\theta_{j[n^{(2k)},i]}$, $\theta_{j[n^{(2k-1)},i]}$ $i=1,\cdots,N$ 是条件独立的, 于是有

$$\widetilde{P}(B_k) \geqslant \widetilde{P}(D_k) = \prod_{i=1}^{N} [P(\theta = j(i) \mid X = a_i) P(\theta = j[i] \mid X = a_i)]$$

$$= \prod_{i=1}^{N} (P_{ji(i)} P_{ji[i]} / q_i^2) \triangleq \gamma > 0 \tag{13}$$

故有 $\sum_{k=1}^{\infty} \widetilde{P}(B_k) = \infty$. 因为 $(X_1, X_2, \cdots) \in A$ 给定时, $B_1, B_2, \cdots$ 是条件独立的, 那么由 Borel-Cantelli 引理知, $\widetilde{P}(B_k, \text{ i.o.}) = 1$. 注意, 在 (7) 式中 $|I_{N,n}| < \varepsilon$, 故有 $B_k \subset \{|L_{n(2k)}^{(2)} - L_{n(2k-1)}^{(2)}| \geqslant \varepsilon\}$,

$$\widetilde{P}\{L_n^{(2)} \text{ 发散}\} \geqslant \widetilde{P}\{|L_n^{(2)} - L_m^{(2)}| \geqslant \varepsilon, \text{ i.o. 对于} (n, m)\}$$
$$\geqslant \widetilde{P}\{|L_{n(2k)}^{(2)} - L_{n(2k-1)}^{(2)}| \geqslant \varepsilon, \text{ i.o. 对于 } k\} \geqslant \widetilde{P}\{B_k, \text{ i.o. 对于 } k\} = 1.$$

即有 $\widetilde{P}(L_n^{(2)}$ 发散$) = 1$. 因为 $P(A) = 1$, 由 Fubini 定理, 便证得 $P(L_n^{(2)}$ 发散$) = 1$, 从而证明了 (e). 定理 1 证毕.

注　由第(e)步的证明可以看出, 当定理中的条件(II)或(III)不满足时, 则有

$$P(|L_n^{(2)} - L_m^{(2)}| \geqslant 4\varepsilon - \delta, \text{ i.o. 对于 } n, m) = 1,$$

其中 $\varepsilon > 0$ 由 (10) 式定义, $\delta > 0$ 任意给定. 因此由 (10) 式可看出, $L_n^{(2)}$ a.s. 发散的振幅为 $\sum_{i=1}^{\infty} P_{ij(i)} - \sum_{i=1}^{\infty} P_{ij[i]}$. 事实上, 应有

$$\limsup_{n\to\infty} L_n^{(2)} = \sum_{i=1}^{\infty} P_{ij(i)}, \quad \text{a.s.}$$

$$\liminf_{n\to\infty} L_n^{(2)} = \sum_{i=1}^{\infty} P_{ij[i]}. \quad \text{a.s.}$$

定理 2　当 Q 为纯原子分布即 X 为离散型随机变量, 不管 X 之分布 Q 如何, 在下标降序法下总有

$$\lim_{n\to\infty} T_n^{(2)} = R, \text{ a.s.}; \quad \lim_{n\to\infty} R_n^{(2)} = R. \tag{14}$$

证明　继续采用定理 1 中所使用的记号. 对任意给定的 $(X_1, X_2, \cdots) \in A$, 及对任意给定的 $N > 0$, 当 $n \geqslant n_N$ 时, 由 (7) 式可得

$$\begin{aligned}
T_n^{(2)} &= E(L_n^{(2)} \mid X_1, X_2, \cdots) \\
&= 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} P(\theta_{j[n,i]} = j \mid X_{j[n,i]} = i) p_{ij} + I_{(N,n)} \\
&= 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} P(\theta = j \mid X = i) p_{ij} + I_{(N,n)} = 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} (p_{ij}^2 / q_i^2) q_i + I_{(N,n)} \\
&= 1 - \sum_{i=1}^{N} \sum_{j=1}^{M} p_{ij}^2 / q_i + I_{(N,n)},
\end{aligned}$$

其中

$$\begin{aligned}
|I_{(N,n)}| &= \left| E\left(\sum_{i=N+1}^{\infty} \sum_{j=1}^{M} I_{(\theta_{n,i}^{(2)} = j)} p_{ij} \mid X_1, X_2, \cdots \right) \right| \\
&\leqslant \left| E\left(\sum_{i=N+1}^{\infty} q_i \mid X_1, X_2, \cdots \right) \right| = \left| \sum_{i=N+1}^{\infty} q_i \right| < \varepsilon.
\end{aligned}$$

其中 $\varepsilon > 0$ 是任意预先给定的, 选择 N 充分大使上式成立. 从而知当 $n \geqslant n_N$ 时, 有

$$|T_n^{(2)} - R^{(2)}| \leqslant 2 \sum_{i=N+1}^{\infty} q_i < 2\varepsilon.$$

再由 ε 的任意性及事实 $P(A) = 1$ 知, $T_n^{(2)} \to R$, a.s. 当 $n \to \infty$, 成立. 注意 $R_n^{(2)} = ET_n^{(2)}$, 于是再由有界控制收敛定理, 便证明了 $\lim_{n\to\infty} R_n^{(2)} = R$. 定理 2 证毕.

现来转向考虑 X 为一般变量即 X 的分布 Q 既可能含有原子但又未必为纯原子时的情形．记 Q 的原子的全体为 $\mathscr{X}^{(1)}=\{a_1, a_2, \cdots\}$，并令 $\mathscr{X}^{(2)}=R^d-\mathscr{X}^{(1)}$ 为非原子部分，又记

$$Q_i(\cdot)=Q(\cdot\cap\mathscr{X}^{(i)}),\quad i=1, 2,$$

则 Q_1 Q_2 为 $(R^d, \mathscr{B}^d)$ 上的有限测度，我们称在 R^d 上定义的某一可测函数 $f(x)$ 为 KQ 连续的，如果存在一个在 R^d 上定义的关于 Q_2 a.s. 连续的函数 $\tilde{f}(x)$，使得有 $f(x)=\tilde{f}(x)$，当 $x\in\mathscr{X}^{(2)}$，并且有

$$\lim_{k\to\infty}[f(a_k)-\tilde{f}(a_k)]=0.$$

显然，KQ 连续要比关于 Q a.s. 连续的要求条件要弱些．对于一般的指标变量，我们有

定理 3 假设对每一 $j(=1, \cdots, M)$，$P(\theta=j|X=x)$ 作为 x 的函数是 KQ 连续的，则在 (X, θ) 的任意分布下，总有

$$\lim_{n\to\infty}R_n^{(2)}=R,\quad \lim_{n\to\infty}T_n^{(3)}=R\quad \text{a.s.}$$

并且仍有(I)⇔(II)⇔(III)，其中命题(I)，(II)，(III)与定理 1 中叙述的相同．当这些命题不真时，则 $L_n^{(2)}$ a.s. 发散。

证明　若记

$$L_{n,i}^{(2)}=P(\theta_n^{(2)}\neq\theta,\ X\in\mathscr{X}^{(i)}|Z^n)$$
$$T_{n,i}^{(2)}=P(\theta_n^{(2)}\neq\theta,\ X\in\mathscr{X}^{(i)}|Z^n)$$
$$R_{(i)}=1-\sum_{j=1}^{M}E\{I_{(X\in\mathscr{X}^{(i)})}P^2(\theta=j|X)\}$$

$i=1, 2$，则有 $L_n^{(2)}=L_{n,1}^{(2)}+L_{n,2}^{(2)}$，$R=R_{(1)}+R_{(2)}$ 及 $T_n^{(2)}=T_{n,1}^{(2)}+T_{n,2}^{(2)}$．采用我们在[8]中所使用的方法，类似地可以证明

$$\lim_{n\to\infty}L_{n,2}^{(2)}=R_{(2)}\quad \text{a.s.}$$

从而利用有界控制收敛定理又可证得

$$\lim_{n\to\infty}T_{n,2}^{(2)}=R_{(2)}\quad \text{a.s.}$$

再利用本文上面在定理 1, 2 中所采用的方法，同样可证

$$\lim_{n\to\infty}T_{n,1}^{(2)}=R_{(1)}\quad \text{a.s.},$$

并且 (II)⇔(II)⇒$\lim\limits_{n\to\infty}L_{n,1}^{(2)}=R_{(1)}$ a.s.，而且当(II)或(III)不真时，$L_{n,1}^{(2)}$ a.s. 发散．综合这些结果便完成了本定理的证明．

§3　三种判别法的比较

由文献[6]～[8]，对于 §1 节中考虑的第(i)、(iii)种最近邻判别法，已得到如下结果：

关于下标升序法(i)，当 X 的分布 Q 为纯原子型时，总有 $\lim\limits_{n\to\infty}R_n^{(1)}=R$，$\lim\limits_{n\to\infty}T_n^{(1)}=R$，a.s.，$\lim\limits_{n\to\infty}L_n^{(1)}=L(\varDelta)$，a.s.，其中 $\varDelta=((x_i, \theta_i), i=1, 2, \cdots)$，而

$$L(\varDelta)=1-\sum_{i=1}^{\infty}\sum_{j=1}^{M}I_{(\theta_{j_i}=j)}P(\theta=j,\ X=a_i)\tag{15}$$

$$j_i=\min\{j: x_j=a_i\},\quad i=1, 2, \cdots.$$

$L(\varDelta)$ 一般为一随机变量，而且有：$L(\varDelta)=R$, a.s. ⇔(II)(或(III))．对一般变量 X，假定对

每一$j(=1,\cdots,M)$，$P(\theta=j|x)$是x的KQ连续函数，那么有

$$\lim_{n\to\infty}R_n^{(1)}=R,\ \lim_{n\to\infty}T_n^{(1)}=R,\ \text{a.s.},\ \lim_{n\to\infty}L_n^{(1)}=L'(\varDelta),\ \text{a.s.},$$

其中

$$L'(\varDelta)=1-\sum_{j=1}^{M}E\{I_{(X\in\mathscr{X}^{(2)})}P^2(\theta=j|X)\}-\sum_{i=1}^{\infty}\sum_{j=1}^{M}I_{(\theta_{i1}=j)}P\{\theta=j,\ X=a_i\}\tag{16}$$

而且$L'(\varDelta)=R$, a.s. $\Leftrightarrow$(II)(或(III))。

关于等权随机化法(iii)．在(X,θ)的任一分布下，总有$\lim\limits_{n\to\infty}R_n^{(3)}=R$，$\lim\limits_{n\to\infty}T_n^{(3)}=R$, a.s.，在当$P(\theta=j|x)$为$KQ$连续时，有$\lim\limits_{n\to\infty}L_n^{(3)}=R$, a.s.． 事实上，当$X$为离散变量时，对任意$\varepsilon>0$，总存在$b=b(\varepsilon)>0$，$c=c(\varepsilon)<\infty$，使有

$$P\{|L_n^{(3)}-R|\geqslant\varepsilon\}\leqslant ce^{-nb},$$

对一般变量X，当$P(\theta=j|x)$为KQ连续时，也有

$$P\{|L_n^{(3)}-R|\geqslant\varepsilon\}<ce^{-\sqrt{n}b}.$$

由以上所列结果可以看出，对于(i)，(ii)，(iii)三种判别法，$R_n^{(i)}$，$T_n^{(i)}$的收敛性是一致的．不过在实际问题中，人们更有兴趣的问题是后验错判概率$L_n^{(i)}$的大小及稳定性(当$n\to\infty$)如何．我们的结论是，当(II)或(III)条件成立时，并且$P(\theta=j|x)$作为x的函数为KQ连续时，$L_n^{(1)}$，$L_n^{(2)}$，$L_n^{(3)}$均a.s.收敛于常值R，即上述三种判别法其后验错判概率具有相同的强收敛性质．可惜，(II)，(III)条件是不常见的不足道的情况。但当这些条件不成立时，$L_n^{(2)}$ a.s.发散，$L_n^{(1)}$ a.s.收敛于某一随机变量，其均值为R，但$L_n^{(3)}$却总是a.s.收敛于R的．下标升序法与降序法的收敛性的差异，反映出在最近邻判别大样本理论中，历史样本的排列秩序对后验错判概率有明显影响．这一事实多少有点出人预料． 此因人们从直观上看，优先考虑较新的样本(相应于下标降序法)比优先考虑较老的样本(相应于下标升序法)其效果可能要好些，但是上述结论指出这个直观想法是不对的． 但若采用等权随机化法判别，便总能保证后验错判概率有较好的收敛性质．这从一个侧面说明了采用随机化法的合理性．

参 考 文 献

[1] Devroye, L., *Ann. Statist.*, (1981), 1320—1327.

[2] Wagner, T. J., *IEEE Trans. Inform. Theory.*, (1971), 566—570.

[3] Jozsef, Fritz, *IEEE Trans. Inform. Theory.*, (1975), 592.

[4] 陈希孺, The Exponeatial Bound of Error Probability In K-NN Rule (待发表).

[5] Bai Zhidong, Strong Consistency of Error Probability *Estimates in NN Discrimination* (待发表).

[6] 陈桂景，孔繁超，安徽大学学报(自然科学版)，(1983)，1: 18—25.

[7] 陈桂景，孔繁超，Sufficient and Necessary Condition For Convergence of Conditional Error Probability *In NN-Pattern Discrimination* (待发表).

[8] 白志东，陈希孺，陈桂景，再论最近邻判别分析(待发表).

[9] 孙志刚，The Necessary and Sufficient Condition of Convergence of Error Probability Estimates in K-NN Discrimination (待发表).

[10] Hoeffding, W., *J. Amer. Statist. Assoc.*, 58 (1963), 13—50.

NN-PATTERN DISCRIMINATION IN DECREASING ORDER OF AFFIXES

CHEN GUIJING CHEN GANG
(*Anhui University*)

Let (θ, X) be a random vector with $\theta \in \{1, \cdots, M\}$, $X \in R^d$, and (θ_i, X_i), $i=1,\cdots, n$, i.i.d. random samples of (θ, X). For a distance function $\rho(\cdot, \cdot)$ given on R^d, denote

$$k_n = \max\{j: \rho(x_j, x) = \min_{1<i<n} \rho(x_i, x),\ j \leqslant n\}, \quad \theta_n^{(2)} = \theta_{k_n},$$

Where $\theta_n^{(2)}$ is called NN discrimination in decreasing order of affixes of θ. In this paper we prove that if $P(\theta=j|X=x)$, $j=1, \cdots, M$, are QK-continuous then the following statements are equivalent:

(I) $L_n^{(2)} \to R$, a.s. as $n \to \infty$,

(II) For every atom a_i of X, $1 \leqslant j < k < M$

$$P(\theta=j, X=a_i)P(\theta=k, X=a_i)[P(\theta=j, X=a_i)-P(\theta=k, X=a_i)]^2=0$$

If (II) is not true, then $L_n^{(2)}$ is divergent, a.s., where

$$L_n^{(2)} = P(\theta_n^{(2)} \neq \theta | Z^n),\ Z^n = ((\theta_i, X_i),\ i=1, \cdots, n).$$