第4卷 第2期
1984年4月
数学研究与评论
JPURNAL OF MATHEMATICAL
RRSEARCH AND EXPOSITION
Vol.4 No.2
April 1984

近五年来中国数理统计工作者的部分理论成果(Ⅱ)*

陈希孺

(中国科技大学数学系)

§2 概率密度的估计

设X_1, …, X_n是从具概率密度f的总体中抽出的iid.样本,要由之估计f.这个问题自六十年代以来一直受到不少统计学家的重视。近年来,我国数理统计工作者也在这方面获得了不少成果。以下分几个方面叙述。

(一)核估计的一致收敛速度

在已提出的众多的密度估计方法中,最重要的且得到最深入研究的,要算是由Rosenblatt及Parzen提出和发展的核估计,选定一函数K(称为核函数或简称核)及常数$h_n>0$,则

$$f_n(x)=f_n(x;X_1,\cdots,X_n)=\frac{1}{nh_n^m}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h_n}\right) \tag{2.1}$$

称为f的以K为核的核估计。此处m为总体的维数。

Parzen[2]讨论了核估计的若干初步的大样本性质。往后的研究的一个重要方面,是f_n的强相合性,包括在指定点处的强相合及相对于全空间R^m的一致强相合,即在什么条件下有

$$\lim_{n\to\infty}f_n(x)=f(x),\ a.s.,\ (x\text{ 指定})$$

及

$$\lim_{n\to\infty}\sup_{x\in R^m}|f_n(x)-f(x)|=0,\ a.s. \tag{2.2}$$

关于这个问题的新近进展,历史情况和有关参考文献,可参看Devroye和Wagner的工作[3]。

与此同时产生了收敛速度的问题。为明确计,引入如下的记号。设$\{\xi_n\}$为一串随机变量,而$\{A_n\}$为一串趋于∞的常数。若

$$\lim_{n\to\infty}A_n\xi_n=0,\ a.s.$$

则记为$\xi_n=o(A_n^{-1})$。若

$$\limsup_{n\to\infty}|A_n\xi_n|<\infty,\ a.s.$$

则记为$\xi_n=O(A_n^{-1})$。

*1982年5月12日收到。本文(Ⅰ)刊登在本刊Vol. 3 (1983), No. 4, pp. 109—120.

首先，Schuster 在 $m=1$ 的情况下证明了（见[4]）：在核 K 满足某些条件（如为概率密度函数、在 R^1 连续、有界变差等）且 f' 在 R^1 一致有界时，选择 $h_n=n^{-1/4}$，可以达到

$$\sup_{x\in R^1}|f_n(x)-f(x)|=o(n^{-1/4+\varepsilon}) \tag{2.3}$$

1977 年 Singh 在 f 的 r 阶导数 $f^{(r)}$ $(r\geqslant 0)$ 在 R^1 一致有界的条件下，证明对适当选择的 K 和 h_n，可以达到（见[5]）

$$\sup_{x\in R^1}|f_n(x)-f(x)|=O(n^{-r/(2r+2)}\sqrt{\log\log n}). \tag{2.4}$$

最近 Susarla 等在[6]中对一般 m 的情况，对包括核估计在内的一大类估计 $r_n(x)$ 考察了这个问题。在 f 的一切一阶偏导数都在 R^m 一致有界的条件下，达到了数量级

$$\sup_{x\in R^m}|r_n(x)-f(x)|=O(n^{-1/(2m+2)}\sqrt{\log\log n}). \tag{2.5}$$

若只假定 f 满足 δ 阶的 Lipschitz 条件（$0<\delta\leqslant 1$，$\delta=1$ 就是上述情况），则可达到

$$\sup_{x\in R^m}|r_n(x)-f(x)|=O(n^{-\delta/(2m+2\delta)}\sqrt{\log\log n}). \tag{2.6}$$

这些工作在方法上相似，故得出的数量级也相当。

本文作者在[7]中证明了如下结果：若 f 的一切 k 阶偏导数都在 R^m 上一致有界，则对适当选择的 K 和 h_n，可以达到

$$\sup_{x\in R^m}|f_n(x)-f(x)|=O(n^{-k/(2k+m)}\sqrt{\log n}). \tag{2.7}$$

当 $k=1$ 时，数量级的主要部分为 $n^{-1/(m+2)}$，比（2.5）有显著改进。当 $k=r$ 而 $m=1$ 时，为 $n^{-r/(2r+1)}$，也比（2.4）有显著改进，往下我们将指出：$k/(2k+m)$ 这个值已不能再有改进，但 log 部分能改进到何种程度则尚属未知。

陈桂景在[8]中也研究了这个问题。设 $m=1$，而 $f^{(2)}$ 在 R^1 有界，他证明了与本文作者大体上一样的数量级，但对核函数 K 的要求较宽些，他还研究了当 $m=1$ 而 f 满足 δ 阶 Lipschitz 条件的情况，证明了对适当选择的核估计 f_n，可以达到

$$\sup_{x\in R^1}|f_n(x)-f(x)|=O(n^{-\delta/(2\delta+1)}(\log n)^{\delta/(2\delta+1)}). \tag{2.8}$$

这也比（2.6）当 $m=1$ 的情况有显著改进，有理由猜测，在（2.8）中指数 $\delta/(2\delta+1)$ 已不能再有改进。

与此附带有关的是总体分布的众数（mode）的估计问题。Parzen 在[2]中已提出用 f_n 的众数，即满足条件

$$f_n(\theta_n)=\sup_x f_n(x) \tag{2.9}$$

的 θ_n，去估计总体众数 θ。在一定的条件下，Parzen 证明了 θ_n 的渐近正态性与弱相合性，陈桂景在上引文章中也研究了这个问题，在总体密度有三阶有界导数及其他某些假定之下，他证明了：由（2.9）确定的 θ_n 满足

$$\theta_n-\theta=O(n^{-2/7}).$$

关于众数估计的工作还有本文作者的[9]，其特点是直接估计众数而不通过密度估计。

(二) 最佳收敛速度问题

不论用什么方法估计密度 f，当样本大小 n 愈来愈大时，估计量 $r_n=r_n(x;X_1,\cdots,X_n)=r_n(x)$ 与 f 之间，在某种意义上的差距应愈来愈小，并当 $n\to\infty$ 时趋于 0（这就是基本的相合性要求）。当然，希望这个收敛于 0 的速度愈快愈好，这样就提出了在种种意义下可能的最佳收敛速度问题。

这方面的工作可以说始于 Farrell 的工作[10]，他引进了这样一个密度族（$a>0$ 给定）

$$\mathscr{F}_a=\{f:\ f^{(2)}(x)\ \text{对一切}\ x\ \text{存在连续，且}\ \sup_x f(x)\leqslant a\},$$

设要在 $\mathscr{F}_a$ 中估计 $f(0)$，以 $r=(r_N,N)$ 记任一序贯估计，其中 N 为停止法则，r_N 为最终判决，估计的均方误差记为

$$R(f,r)=E_f[r_N(X_1,\cdots,X_N)-f(0)]^2,$$

Farrell 在[10]中证明了：若 $a\geqslant 3$，则不可能存在这样的序贯估计 $r=(r_N,N)$，致

$$\sup_{f\in\mathscr{F}} E_f(N)<\infty, \tag{2.10}$$

而

$$\sup_{f\in\mathscr{F}} \mathrm{R}(f,r)<\frac{1}{16}.$$

本文作者对这个结果作了较大的改进，在一项工作（已投《科学探索》）中证明了：对任给 $a>0$ 存在只与 a 有关的常数 $b_a>0$，使对任何满足条件

$$P_f(N<\infty)=1,\ \text{对任何}\ f\in\mathscr{F}_a \tag{2.11}$$

的序贯估计 $r=(r_N,N)$，必有

$$\sup_{f\in\mathscr{F}} \mathrm{R}(f,r)\geqslant b_a. \tag{2.12}$$

就是说，不论用什么方法，序贯的还是非序贯的，均方误差在 $\mathscr{F}$ 上的上确界只能降低到一定的限度。这个结果多少有些出人意外，因为它对所允许用的估计没有任何限制。

在对 f 的到某阶为止的导数作出有界的限制的情况下，Farrell[11]得出了一项重要的结果。它可表述如下：给定自然数 k 及常数 $a>0$。以 C_{ka} 记一切满足条件

$$\sup_x\left|\frac{\partial^k f(x_1,\cdots,x_m)}{\partial x_1^{k_1}\cdots\partial x_m^{k_m}}\right|<a,\ \text{对任何}\ k_1\geqslant 0,\cdots,k_m\geqslant 0,\ \sum k_i=k$$

的密度函数 f 的集，以 $\{r_n(0)=r_n(0;X_1,\cdots,X_n)\}$ 记 $f(0)$ 的任一串估计，$\{a_n\}$ 为一串常数，满足

$$\liminf_{n\to\infty}\inf_{f\in C_{ka}} P_f(|r_n(0)-f(0)|\leqslant a_n)=1, \tag{2.13}$$

则必有

$$\lim_{n\to\infty} a_n n^{-k/(2k+m)}=\infty. \tag{2.14}$$

条件 (2.13) 的直接解释是：作为 $f(0)$ 的区间估计 $r_n(0)\pm a_n$，其"渐近置信系数"为 1，而 (2.14) 就表示这时 a_n 趋于 0 不能太快。利用这个结果，Farrell 很容易得出：不可能找到这样一串估计 $\{r_n(0)\}$，使

$$\sup_{f\in C_{ka}} E[r_n(0)-f(0)]^2=o(n^{-2k/(2k+m)}). \tag{2.15}$$

另一方面，可找到适当的核估计 f_n，致

$$\sup_{f\in C_{km}} E[f_n(0)-f(0)]^2 = O(n^{-2k/(2k+m)}). \tag{2.16}$$

这二者结合，就给均方误差一致趋于0的速度定出了一个最佳界限，即（2.16）右边．但 Farrell 的上述结果有其弱点，即条件（2.13）过强．它相当于要求

$$\lim_{n\to\infty} P_f(|r_n(0)-f(0)|\leqslant a_n)=1 \tag{2.17}$$

对 $f\in C_{km}$ 一致地成立．而 Farrell 的结果无法知道：是否存在这样一串估计 $\{r(0)\}$，致

$$E_f[r_n(0)-f(0)]^2 = o(n^{-2k/(2k+m)}),\ \text{对任何}\ f\in C_{km}\ \text{成立}. \tag{2.18}$$

也无法知道：$r_n(0)-f(0)$ 收敛于0的速度能达到多快．

本文作者最近在一项工作［12］中，把 Farrell 上述定理中的条件（2.13）去掉，而代之以较弱的条件：（2.17）对一切 $f\in C_{km}$ 都成立，这时可证明

$$\liminf_{n\to\infty} a_n n^{-k/(2k+m)} > 0. \tag{2.19}$$

（2.19）虽然不如（2.14）强，但已足够解决我们上面指出的两个有关最佳速度界限的问题．实际上，由（2.19）不难得出：

1．不存在任何一串估计 $\{r_n(0)\}$，满足（2.18）；

2．不存在任何一串估计 $\{r_n(0)\}$，满足

$$r_n(0)-f(0)=o(n^{-k/(2k+m)}),\ \text{对任何}\ f\in C_{km}. \tag{2.20}$$

再结合（2.7）即知，在任何情况下 $r_n(0)-f(0)$ 达不到 $o(n^{-k/(2k+m)})$，但 $\sup\limits_x|f_n(x)-f(0)|$ 可达到 $O(n^{-k/(2k+m)}\sqrt{\log n})$．因此指数部分 $n^{-k/(2k+m)}$ 已无可改进．

（三）最近邻估计的收敛速度

这是 Loftsgarden 等于1965年在［13］中提出的一种密度估计方法，指定自然数 $k_n\leqslant n$，找最小的 $a_n(x)=a_n(x;X_1,\cdots,X_n)$，使区间 $x\pm a_n(x)$ 内至少包含 $X_1,\cdots,X_n$ 中的 k_n 个，然后以

$$\hat{f}_n(x)=\frac{k_n}{2na_n(x)} \tag{2.21}$$

去估计 $f(x)$，在多维情况下有类似推广，只须以中心在 x 点的球代替区间．

如在核估计的情况，对 $\hat{f}_n$ 的研究重点在于其强相合性，这一点在1977年由 Devroye 和 Wagner 在［15］中基本上彻底解决了．我国学者在 $\hat{f}_n$ 收敛于 f 的速度问题上作了一些工作，首先，对固定一点 x 的情况（$m=1$），本文作者在［16］中证明：若 $r\geqslant 2$ 为自然数，$f(x)>0$，$f^{(2)}(x)=\cdots=f^{(r-1)}(x)=0$，$f^{(r)}(x)\neq 0$（若 $r=2$，则只要 $f^{(2)}(x)\neq 0$），则 $\hat{f}_n(x)-f(x)$ 达不到 $O(n^{-r/(2r+1)})$，但可适当选择 k_n，使之达到 $O(n^{-r/(2r+1)}\sqrt{\log n})$．对 $\sup\limits_x|\hat{f}_n(x)-f(x)|$（仍设 $m=1$），本文作者在［16］中，在 f 满足Lipschitz条件之下，得到了速度 $O(n^{-1/6}\cdot(\log\log n)^{1/6})$，且证明不可能达到 $O(n^{-1/4}(\log\log n)^{-1/4})$．后来徐业基在一项工作（已投《数学年刊》）中，在 $f^{(2)}$ 于 R^1 有界的条件下，得到了速度 $O(n^{-1/5}(\log\log n)^{1/5})$．接着，本文作者在［16］中，仍沿用［15］的基本思想，但利用了［3］中的一个有力的不等式，在 f

满足 δ 阶 $(0<\delta\leqslant 1)$ Lipschitz 条件之下，达到了速度 $O(n^{-\delta/(1+3\delta)}\sqrt{\log n})$，且证明了不可能达到速度 $o(n^{-\delta/(1+3\delta)})$，这就（在 f 满足 Lipschitz 条件之下）完全解决了 $\sup\limits_x|\hat{f}_n(x)-f(x)|$ 的最佳收敛速度的主要部分。至于其 log 部分，最近杨振海在一项工作[17]中进一步将上述速度中的 $\sqrt{\log n}$ 部分，在 $\delta=1$ 的情况下改进为 $(\log n)^{1/4}$，与此收敛速度有关的另一项工作是柴根象[18]，在 $f^{(2)}$ 于 R^1 有界的条件下，将 $\sup\limits_x|\check{f}_n(x)-f(x)|$ 的收敛速度改进为 $O(n^{-2/7}\sqrt{\log n})$。不论 f 有多少阶有界导数，指数部分 $n^{-2/7}$ 已不能再改进。

最近邻估计从形式上很接近于均匀核估计。从以上介绍的结果可知，在强收敛速度方面看，总的说前者不如后者，然而有趣的是：若从均方误差的角度看，至少从主要部分而言二者分不出高低，事实上，本文作者在[19]中证明了：在 $f^{(3)}(x)$ 存在（x 指定），$f(x)\neq 0\neq f^{(2)}(x)$ 的条件下，适当选择 k_n 和 h_n，$E_f[\hat{f}_n(x)-f(x)]^2$ 和 $E_f[f_n(x)-f(x)]^2$（f_n 是均匀核密度估计）的最优值都是

$$\frac{5}{4}144^{-1/5}(f^4(x)f^{(2)}(x))^{1/5}n^{-4/5}+o(n^{-4/5})\ .$$

（四）核估计一致收敛的必要条件问题

以 f_n 记核估计。在所有关于 $\sup\limits_x|f_n(x)-f(x)|\to 0\ a.s.$ 的条件的结果中，一无例外地都有"f 在全空间一致连续"这个要求，Schuster 是第一个提出和研究这个条件的必要性的人。他在[4]中证明了如下结果：设 f_n 为由 K 及 h_n 产生的核估计。假定

1° $\sum\limits_{n=1}^{\infty}e^{-cnh_n^2}<\infty$ 对任何 $c>0$，$0<h_n\to 0$；

2° K 满足以下条件：a. K 为在 R^1 上有界变差的概率密度函数，b. K 在 R^1 上处处连续，c. $\lim\limits_{|x|\to\infty}xK(x)=0$，d. $\int_{-\infty}^{\infty}|x||dK(x)|<\infty$，e. 存在 $r\in(0,1)$，使 $\lim\limits_{x\to\infty}\{x[\bigvee_{-\infty}^{-x^r}(K)+\bigvee_{x^r}^{\infty}(K)]\}=0$；

3° $\sup\limits_x|f_n(x)-g(x)|\to 0\ a.s.$，对某个函数 g。

则 g 必在 R^1 一致连续，总体密度函数存在且就是 g。

这个结果中关于核 K 的要求太强，特别是排除了某些重要的非连续核，例如均匀核，本文作者在[20]中专门考虑了均匀核的情况，证明了：不论 $h_n>0$ 如何取，只要有上述条件 3°，则 Schuster 定理的结论仍成立，后来，柴根象在一项工作（已设《数学研究与评论》）中把这个结论推广到核为阶梯形概率密度函数的情况。

最近本文作者在一项工作[21]中，对 Schuster 的上述结果作了较显著的改进。证明了：若 $0<h_n\to 0$，$\limsup\limits_{n\to\infty}nh_n^2=\infty$ 而 K 满足上述条件 2°a，则当 3° 成立时，g 必在 R^1 一致连续且总体密度存在并等于 g。这项工作简化了 Schuster 的证明，并使推广到高维空间成为可能。也举例指出了：条件 2° a 中关于 K 在 R^1 上有界变差的部分并非必要，但去掉这一部分后结论是否仍成立则尚未知。

§3　经验 Bayes 估计的大样本理论

经验 Bayes(Empirical Bayes，简称为 EB）方法是 H. Robbins 在1955年[22]中提出的一种统计判决思想，已去世的统计界元老 J. Neyman 曾称它是统计判决的两大突破之

一。虽然 Neyman 的估计与尔后的发展不甚符合，EB 方法仍不失为战后几十年来数理统计学的一项重要成就。

EB 方法的一个方面的工作是对一些常见问题寻找性能良好的 EB 方法，并用包括模拟在内的种种方法去探讨其性质，另一方面是大样本理论，我国统计学者在这方面做了一些工作。

以（$\mathscr{X}, \mathscr{B}_x$）记样本空间，$\{P_\theta, \theta\in\Theta\}$ 记样本 X 的分布族，（$D, \mathscr{B}_D$）为行动空间，$L(\theta, d)$ 为损失函数，G 为先验分布，以 $X_1, \cdots, X_n$ 记历史样本，则任一经验 Bayes 判决 $\delta_n=\delta_n(x)=\delta_n(x_1, \cdots, x_n; x)$ 的"全面 Bayes 风险"定义为

$$R(\delta_n, G)=E\{L(\theta, \delta_n(X_1, \cdots, X_n; X))\}$$
$$=\int_{\mathscr{X}}\cdots\int_{\mathscr{X}} dP^*(x_1)\cdots dP^*(x_n)\int_\Theta\int_{\mathscr{X}} L(\theta, \delta(x_1, \cdots, x_n; x)dP_\theta(x)dG(\theta),$$

此处 P^* 为在先验分布 G 之下，X 的边缘分布。即

$$P^*(X\in A)=\int_\Theta P_\theta(A)dG(\theta).$$

以 $R(G)$ 记在先验分布 G 之下的 Bayes 解的 Bayes 风险。显然，对任何 n 与 δ_n，有 $R(\delta_n, G)\geqslant R(G)$。设 $\mathscr{F}$ 为一先验分布族。若 δ_n 为经验 Bayes 判决，满足条件

$$\lim_{n\to\infty} R(\delta_n, G)=R(G),\text{ 对一切 } G\in\mathscr{F},$$

则称 δ_n 为相对于 $\mathscr{F}$ 的渐近最优（asymptotically optimal，简记为 $a.o.$）EB 解，大样本理论的一个基本问题是寻找有 $a.o.$ 性的 EB 解，或证明某种特定的 EB 解有 $a.o.$ 性。

$a.o.$ EB 解的存在与否在很大的程度上取决于 $\mathscr{F}$ 的性质，也与分布族（$P_\theta, \theta\in\Theta$）的性质有很大的关系。例如，本文作者曾在[23]中证明：即使在平方损失下估计二项分布概率 p 这样简单的问题，如不对先验分布族 $\mathscr{F}$ 施加性质特殊的限制，则 $a.o.$ EB 解将不存在。同时，$a.o.$ EB 解的存在也与损失函数 L 的性质有关。在假设检验问题中往往假定 L 有界。在这种情况下 $a.o$ 性质较易获得。对估计问题，损失函数大多是无界的，问题就困难得多。到目前为止得出的结果很少。

考虑写成下面形状的一维离散型指数分布族

$$P_\theta(X=x)=h(x)\beta(\theta)\theta^x,\quad x\in\mathscr{X}=\{0, 1, 2, \cdots\}, \tag{3.1}$$

此处 $h(x)\geqslant 0$，而 $\Theta=\{\theta: \sum_{x=0}^\infty h(x)\theta^x\in(0, \infty,)\}$。(3.1) 包含了一些常见的离散型分布。如当 $h(x)=1/x!$，$\beta(\theta)=e^{-\theta}$ 的情况，得 Poisson 分布族，在平方损失 $(\theta-d)^2$ 之下并设先验分布为 G，则分布族 (3.1) 中的参数 θ 的 Bayes 估计为

$$\delta_G(x)=\frac{h(x)}{h(x+1)}f_G(x+1)/f_G(x), \tag{3.2}$$

f_G 为 X 的边缘分布，即

$$f_G(x)=\int_\Theta h(x)\beta(\theta)\theta^x dG(\theta). \tag{3.3}$$

若有了历史样本 $X_1, \cdots, X_n$，当前样本 $X=X_{n+1}$，则 $f_G(x)$ 可用 $u_n(X)$ 去估计，此处

$$u_n(j)=\frac{1}{n+1}\{X_1, \cdots, X_{n+1}\text{中，等于 } j \text{ 的个数}\} \tag{3.4}$$

（注意 $u_n(X)>0$）由（3.2）、(3.4)，得到一"自然的" EB 估计（也称为 Robbins 估计）为

$$\delta_n(x_1,\cdots,x_n;x)=\delta_n(x)=\frac{h(x)}{h(x+1)}\frac{u_n(x+1)}{u_n(x)}. \tag{3.5}$$

根据 Robbins 在[24]中的介绍，M.V.Jones 1956 年在其博士论文中研究了在 Poisson 分布特例下，EB 估计（3.5）的 $a.o.$ 性，证明了它相对于先验分布族

$$\mathscr{F}_2=\left\{G:\int_0^\infty\theta^2dG(\theta)<\infty\right\} \tag{3.6}$$

有 $a.o.$ 性。可惜的是，这结果的证明在所有引述它的公开文献中都没有给出过。Jones 在 1957年在[25]中发表了其博士论文的一些结果，但也未提及这个结果及其证明。这触动了本文作者对一般情况下 EB 估计（3.5）的 $a.o.$ 性进行探讨，在[26]中得出了如下的结果：

设 $h(x)>0$ 对 $x=0,1,2,\cdots$，又设存在常数 A，致

$$h^3(x)\leqslant Ah(x-2)h^2(x+1),\quad x=2,3,\cdots \tag{3.7}$$

则由（3.5）确定的分布族（3.1）中参数 θ 的 EB 估计，相对于先验分布族 $\mathscr{F}_2$（见(3.6)）有 $a.o$ 性。

条件（3.7）对一些常见的分布，如 Poisson 分布、负二项分布 $\left(\beta(\theta)=(1-\theta)^r,\ r\right.$ 为已知自然数，$\Theta=(0,1)$，$\left.h(x)=\binom{x+r-1}{x}\right)$ 及对数分布（$\beta(\theta)=-\log(1-\theta)$，$\Theta=(0,1)$，$h(0)=0$，$h(x)=x^{-1}$ 当 $x\geqslant1$）。但在理论上仍可提出问题：可否对（3.5）确定的 δ_n 加以适当修改，使之对任意 h（当然，要求 $h(x)>0$）都有 $a.o.$ 性？本文作者在[26]中肯定地回答了这个问题，结果如下：将 $u_n(x)$ 代之以

$$v_n(x)=\begin{cases}u_n(x+1), & \text{若 } x_1,\cdots,x_n \text{ 中至少有两个为 } x+1\\ 0, & \text{其他情况}\end{cases}$$

而将（3.5）中的 δ_n 修改为

$$\tilde{\delta}_n(x)=\frac{h(x)}{h(x+1)}\frac{v_n(x)}{u_n(x)}$$

则只要 $h(x)>0$（当 $x=1,2,\cdots$），则 $\tilde{\delta}_n$ 相对于先验分布族 $\mathscr{F}_2$ 为 θ 的 $a.o.$ EB 估计。[26]中还讨论了 $h(x)$ 对某些 x 可以为 0 的情况，二项分布族是其一特例。

后来，方兆本考虑了估计 θ 的多项式的问题。这时若限制多项式不高于 k 阶，则先验分布族 $\mathscr{F}_2$ 自应代之以 $\mathscr{F}_{2k}=\left\{G:\int_0^\infty\theta^{2k}dG(\theta)<\infty\right\}$。陶波在[27]中进一步考虑了估计 θ 的一连续函数的问题。

$a.o.$ 问题的进一步发展，是 $R(\delta_n,G)-R(G)$ 收敛于 0 的速度问题。这方面的工作在70年代才开始出现。赵林城[29]对离散指数族的情况作出了较好的结果。这个工作是发现 P.E.Lin（林丕二）的工作[28]中的一个本质错误开始的。

林丕二在[28]中，针对分布族(3.1)，用

$$f_n(j)=\frac{1}{n}\#\{i:\ 1\leqslant i\leqslant n,\ x_i=j\}$$

估计 $f_G(j)$，并作截断：

$$f_n^*(j)=\begin{cases}f_n(j), & \text{当 } f_n(j)\geqslant n^{-1/3},\\ n^{-1/3}, & \text{当 } f_n(j)<n^{-1/3}.\end{cases}$$

然后取

$$\delta_n^*(x)=\frac{h(x)}{h(x+1)}f_n(x+1)/f_n^*(x) \tag{3.8}$$

作为 θ 的 EB 估计。他在[28]中提出的主要结果如下：记 $a_1(x)=\frac{h(x)}{h(x+1)}$，设

1. $\sum\limits_{x=0}^{\infty}a_1^2(x)f_G(x)f_G(x+1)<\infty$,

2. $\sum\limits_{x=0}^{\infty}a_1^2(x)f_G^2(x+1)/f_G(x)<\infty$,

3. 存在 $t\in(0,1]$，致 $P^*(f(X)<\varepsilon)=O(\varepsilon^t)$.

(P^* 的意义见前）则有

$$R(\delta_n^*,G)-R(G)=O(n^{-t/3}) \tag{3.9}$$

赵林城在[29]中指出了林丕二证明中的一处本质错误，并举反例证明上述结果确实不成立。反例是

$$P_\theta(X=x)=\frac{1}{x!}e^{-\theta}\theta^x,\quad x=0,1,2,\cdots,$$

$$dG(\theta)=3\theta^{-4}d\theta,\quad \text{当 } \theta\geqslant1.$$

林丕二的条件对 $t=3/4$ 成立，故依（3.9），应有 $R(\delta_n^*,G)-R(G)=O(n^{-1/4})$。但赵林城经过仔细论证，得出

$$R(\delta_n^*,G)-R(G)\geqslant Cn^{-1/12},\quad C>0 \text{ 与 } n \text{ 无关}.$$

然后赵林城对林丕二的 EB 估计（3.8）作了如下的修改：令

$$\delta_n^{\cdot}(x)=[a_1(x)f_n(x+1)/f_n(x)]_{n^\nu},\quad 0<\nu<1,\quad \nu \text{ 待定}, \tag{3.10}$$

其中 $a_1(x)=h(x)/h(x+1)$，$f_n(x)$ 已定义于前，而 $[b]_c=b$，当 $|b|\leqslant c$，$=0$ 当 $|b|>c$，换言之，赵的估计(3.10)是整个截断而不止是截断分母。赵林城在[29]中证明了如下的结果：

设存在 $\delta>2$ 致 $\int_0^\infty\theta^\delta dG(\theta)<\infty$，又存在 $\lambda\in(0,1)$ 致

$$\sum_{x=0}^{\infty}\left\{1+\left(\frac{a_1^2(x)f_G(x+1)}{f_G(x)}\right)^\lambda\right\}f_G^{\frac{1}{2}-\lambda}(x)<\infty,$$

则当取（3.10）中的 $\nu=\lambda/\delta$ 时，将有

$$R(\delta_n^{\cdot},G)-R(G)=O(n^{-\lambda(\delta-2)/\delta}),$$

这个数量级可取任意接近于 n^{-1} 的形态，而按（3.9），最多只能达到 $O(n^{-1/3})$。实际上，赵林城工作中得出的结果比上面引述的略广，参看[29]。

在此顺便提到，本文作者在[26]中也考虑了林的估计 $\delta_n^*(x)$。按林原来的证明，不仅不能证明 $R(\delta_n^*,G)-R(G)$ 趋于0有一定的速度，就是要证它趋于0也不可能。本文作者证明了如下结果：

设 $h(x)>0$，$x=0,1,2,\cdots$，且满足条件：存在常数 A，

$$h^2(x)\leqslant Ah(x-1)h(x+1),\quad x=1,2,\cdots$$

则相对于先验分布族 $\mathscr{F}_2$（见（3.6）），δ_n^* 为 θ 的 a.o. EB 估计。这个定理对函数 h 的要求不高。特别是，它对 h 和先验分布 G 的要求是分开的。

在指数族以外，关于 EB 估计的结果很少。只Fox[30]考虑了包括均匀分布族$\{R(0,\theta):\theta>0\}$ 和 $\{R(\theta,\theta+1):-\infty<\theta<\infty\}$ 在内的若干情况，给出了 a.o. EB 估计（仍在平方损失$(\theta-d)^2$之下）。方兆本等在［31］中考虑了较一般的均匀分布族 $\{R(\theta,a\theta+b)\}$，其中 $a\geqslant1$ 和 b 为已知常数，而

$$\begin{cases}-\infty<\theta<\infty, & \text{当 } a=1 \text{（且这时 } b>0\text{）}\\ \theta>b/(1-a), & \text{当 } a>1.\end{cases}$$

方兆本等在[31]中，仍在平方损失 $(\theta-d)^2$ 之下，给出了 θ 的一种 EB 估计，它有一定的收敛速度，这速度与先验分布 G 有关。例如，对 $a>1$ 的情况有如下结果：设先验分布族为

$$\mathscr{F}=\left\{G:\int|\theta|^\mu dG(\theta)<\infty \text{ 对某个 } \mu>4,\ G \text{ 具有有界的密度} g\right\}$$

则对适当构造的 EB 估计 δ_n（见[31]），有

$$R(\delta_n,G)-R(G)=O(n^{-(\mu-4)/[4(\mu+1)]+\varepsilon}),$$

对任何 $\varepsilon>0$。对 $a=1$ 的情况[31]中也给出了一个结果。

另外，关于 EB 估计的收敛速度问题，Singh 在[32] 中提出过一个猜测。他考虑绝对连续的一维指数族

$$dP_\theta(x)=f_\theta(x)dx=C(\theta)e^{\theta x}h(x)dx,\quad \theta\in\Theta. \tag{3.11}$$

Θ为 R^1 上的一区间，损失函数为 $(\theta-d)^2$。Singh 提出的猜测是：不论（3.11）中的$C(\theta)$ 和 $h(x)$ 的具体形状如何，即使把先验分布族 $\mathscr{F}$ 取为

$$\mathscr{F}=\{G:\ G \text{ 的负荷集包含在一固定的有界区间 I 内}\},$$

也不可能找到 EB 估计 δ_n，使

$$R(\delta_n,G)-R(G)=O(1/n),\quad \text{对任何 } G\in\mathscr{F}, \tag{3.12}$$

其中 $O(1/n)$ 中的常数可与 G 有关。

Singh 还提出了一个较弱的猜测：不存在 EB 估计 δ_n，使 $R(\delta_n,G)-R(G)=o(1/n)$。

本文作者在[33]中证实了 Singh 的较弱的猜测是正确的。较强的猜测（即(3.12）不成立)是否正确尚属未知。

§4 非参数统计

近年来，我国数理统计工作者在几类重要的非参数统计量方面做了一些工作。

（一）U-统计量

设 $\varphi(x_1,\cdots,x_m)$ 为一个 m 元的对称函数，$X_1,\cdots,X_n$ 为随机变量，$n\geqslant m$。则

$$U_n=\binom{n}{m}^{-1}\sum_{1\leqslant\alpha_1<\cdots<\alpha_m\leqslant n}\varphi(X_{\alpha_1},\cdots,X_{\alpha_m}) \tag{4.1}$$

称为以 φ 为核的 U 统计量。最常见的情况是 $X_1,\cdots,X_n$ 为独立同分布。

这是 Hoeffding 在1948年于[34]中引进的一类很重要的统计量，在非参数统计中有广泛的应用。比此略早，von Mises 在1947年引进了一类统计量

$$V_n=n^{-m}\sum_{a_1,\cdots,a_m=1}^{n}\varphi(X_{a_1},\cdots,X_{a_m})$$

称为 von Mises 统计量，它与 U_n 有密切关系，但应用不如 U 统计量广。

U 统计量之用于估计问题，是通过如下的途径。设 $\mathscr{F}$ 为一个分布族（如，一切其均值存在有限的分布族），$\theta(F)$ 为定义于 $\mathscr{F}$ 上的一个泛函（例如，$\theta(F)=\int_{-\infty}^{\infty}xdF$）。设 $X_1,\cdots,X_n$ 为自 F 中抽出的独立随机样本。$\varphi(x_1,\cdots,x_m)$ 为 m 个变元的对称函数，满足条件

$$E_F\varphi(X_1,\cdots,X_m)=\theta(F)，对任何 F\in\mathscr{F}.$$

则在 $\mathscr{F}$ 满足一定的条件（以使 $X_1,\cdots,X_n$ 的次序统计量 $X_{(1)}\leqslant\cdots\leqslant X_{(n)}$ 为完全统计量，见[35]，第一章§6）时，由（4.1）定义的 U_n 是 $\theta(F)$ 的方差最小无偏估计。在核的二阶矩存在时，不难证明 U_n 是 $\theta(F)$ 的强相合估计，由于 U-统计量是古典的独立和的自然推广（当 $m=1$ 时，U_n 成为独立同分布和），人们自然会预期，U_n 的强相合性在核只有一阶矩时仍成立。这个结果到1966年由 Berk 证明（见[36]）。这样，又发生了 U_n 收敛于 $\theta(F)$ 的速度问题。本文作者在[37]中完全解决了这个问题，得到了以下的结果：

1．若存在 $r\in(1,2)$ 使 $E|\varphi(X_1,\cdots,X_m)|^r<\infty$，则

$$U_n-\theta(F)=o(n^{-(1-1/r)}),\ a.s. \tag{4.2}$$

2．若存在 $r>1$ 致 $E|\varphi(X_1,\cdots,X_m)|^r<\infty$，则对任何 $\varepsilon>0$ 有

$$P(\sup_{k>n}|U_n-\theta(F)|\geqslant\varepsilon)=\begin{cases}o(n^{1-r}), & 当\ 1<r<2,\\ O(n^{1-r}), & 当\ r\geqslant2\end{cases} \tag{4.3}$$

这两个结果中的指数都达到了最佳值，因为在古典的独立同分布和的情况，它们曾分别由 Marcinkiewicz 和 Katz 及 Baum 证明是最佳的，既然 U 统计量是独立同分布和的推广，这种指数当然无法再改进。在此之前，Grams 等曾在1973年[38]中，证明过（4.3）式当 r 为偶自然数的情况。

另外，本文作者在[37]中，在核的 r 阶矩（$r\geqslant1$ 不必为整数）存在的情况下，找到了沟通 U_n 和 V_n 的一个有用的公式：

$$V_n=U_n+\frac{a}{n}+W_n,$$

此处 a 为常数，而 W_n 满足

1．$E|W_n|^r=\begin{cases}o(n^{1-2r}), & 当\ 1\leqslant r<2,\\ O(n^{-3r/2}), & 当\ r\geqslant2.\end{cases}$

2．$W_n=\begin{cases}o(n^{-2+1/r}),\ a.s. & 当\ 1\leqslant r<2,\\ O(n^{-3/2}(\log n)^{1/2+\varepsilon})\ a.s.\ 对任何\ \varepsilon>0, & 当\ r\geqslant2,\end{cases}$

利用这个关系可以把 U_n 的一些大样本性质推到 V_n 上去。例如，前面关于 $U_n-\theta(F)$ 收敛于0的数量级的结果，对于 V_n 也成立。又如，关于 U_n（经过标准化）的分布收敛于 Φ 的 $O(n^{-1/2})$ 速度的结果，对 V_n 也成立，详见[37]。

在[37]中，还解决了 U_n 矩收敛于 $\theta(F)$ 的速度问题，结果如下：若存在 $r\geqslant1$ 使 $E|\varphi(X_1,\cdots,X_m)|^r<\infty$，则

$$E|U_n-\theta(F)|^r=\begin{cases}o(n^{1-r}), & \text{当 } 1\leqslant r<2,\\ O(n^{r/2}), & \text{当 } r\geqslant 2,\end{cases} \tag{4.4}$$

特别，当 $r=1$ 时，知若核 φ 的一阶矩存在，则 U_n 是 $\theta(F)$ 的平均相合估计，即

$$\lim_{n\to\infty}E|U_n-\theta(F)|=0.$$

早在 Hoeffding 于1948年建立 U-统计量的理论时，就证明了其渐近正态性（在核的二阶矩有限时）。以 F_n 记 $(U_n-\theta(F))/\sqrt{VarU_n}$ 的分布函数。$\|F_n-\Phi\|$ 收敛于0的速度问题，经过若干统计学者的研究与逐步改进，才由 Y. K. Chan 等在1978年[39]彻底解决。他们在只假定核的三阶矩有限时证明了

$$\|F_n-\Phi\|=O(n^{-1/2}), \tag{4.5}$$

我国统计学者在这个方面作了一些工作，其中比较显著的是赵林城的工作[40]。他在只假定核的三阶矩有限的条件下，得到了非一致性收敛速度

$$|F_n(x)-\Phi(x)|\leqslant C(n^{-1/2}(1+|x|)^{-3}),$$

其中常数 C 与 n 及 x 都无关，这达到了与古典的独立同分布和情况下同样好的结果。赵林城在 U-统计量收敛速度问题上的另一项工作是[41]，其中在假定核的 $2+\delta$ 阶矩（$0<\delta\leqslant 1$）有限的条件下，证明了相应于（4.5）式的估计

$$\|F_n-\Phi\|=O(n^{-\delta/2}), \tag{4.6}$$

林正炎在 U-统计量的极限理论方面作了一系列的工作，在[42]中，他把赵林城的结果（4.6）推广为如下的形式：设 U_n 是以 $\varphi(x_1,x_2,)$ 为核的 U-统计量，$E\varphi(X_1,X_2)=0$ $(X_1,X_2,\cdots$ 独立同分布$)$，记 $\psi(X_1)=E[\varphi(X_1,X_2)|X_1]$，设 $0<Var\psi(X_1)<\infty$，又 $g(x)$ 为定义在 $(-\infty,\infty)$ 的非负偶函数，在 $[0,\infty)$ 非降，又 $x/g(x)$ 在 $(0,\infty)$ 非降，$x^2g(x)$ 为凸函数，使 $E[\varphi^2(X_1,X_2)g(|\varphi(X_1,X_2)|)]<\infty$，则代替（4.6）有

$$\|F_n-\Phi\|=O(g^{-1}(\sqrt{n})).$$

他还考虑了只假定核的二阶矩存在的情况，以及 $X_1,X_2,\cdots$ 独立但不必同分布的情况（见[43]）。

关于 U-统计量分布的渐近展开，一般都认为是个难题。1980年，Calleant 在[44]中给出了这方面第一个结果，其中要求的条件很强，林正炎在[45]中将 Calleart 等上述工作中最不自然的一个条件去掉了，仍保持其原来的结论。林正炎还在[46]中研究了随机足标的 U-统计量逼近于正态分布的阶。并与陆传荣合作（见[47]），研究了随机指标随机元序列的弱收敛性，作为一个特例，得到了随机足标 U-统计量的弱收敛性的一项结果，改进了 Sproule 1974年（Trans. Amer. Math. Soc. 1974, p. 55）的工作。

本文作者的工作[48]属于 U-统计量的应用的性质，在其中考虑了下述问题：设 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_n$ 分别是从具连续分布 F 和 G 的一维总体中抽出的 iid. 样本，要检验假设

$$H\text{：存在 } a>0 \text{ 和 } b\text{，致 } G(x)=F\left(\frac{x-b}{a}\right).$$

在[48]中使用经过修正的 U-统计量，提出了上述假设 H 的一个检验方法，证明了检验统计量的渐近正态性，以及检验的相合性。

（二）线性置换统计量

设 $A_\nu=\{a_{\nu 1},\cdots,a_{\nu N_\nu}\}$，$B_\nu=\{b_{\nu 1},\cdots,b_{\nu N_\nu}\}$，$\nu=1,2,\cdots$，这里 $a_{\nu i}$，$b_{\nu i}$ 都是实常数，$\lim\limits_{\nu\to\infty}N_\nu=\infty$。以 $(X_{\nu 1},\cdots,X_{\nu N_\nu})$ 记一随机向量，它以 $1/N_\nu!$ 的概率取 $(b_{\nu 1},\cdots,b_{\nu N_\nu})$ 的任一置换为值。则

$$L_\nu=\sum\nolimits_{i=1}^{N_\nu}a_{\nu i}X_{\nu i}$$

称为（由 A_ν 和 B_ν 生成的）线性置换统计量。这种统计量用于非参数统计中的置换检验。易证

$$\lambda_\nu=EL_\nu=N_\nu\bar{a}_\nu\bar{b}_\nu\quad\left(\bar{a}_\nu=\frac{1}{N_\nu}\sum_{i=1}^{N_\nu}a_{\nu i}.\ \ \bar{b}_\nu=\frac{1}{N_\nu}\sum_{i=1}^{N_\nu}b_{\nu i}\right),$$

$$\sigma_\nu^2=\mathrm{Var}L_\nu=\frac{1}{N_\nu-1}\sum_{i=1}^{N_\nu}(a_{\nu i}-\bar{a}_\nu)^2\sum_{i=1}^{N_\nu}(b_{\nu i}-\bar{b}_\nu)^2,$$

记

$$\mu_r(A_\nu)=\frac{1}{N_\nu}\sum_{i=1}^{N_\nu}(a_{\nu i}-\bar{a}_\nu)^r,\quad r=2,3,\cdots,$$

称 A_ν 满足条件 W，若对任何整数 $r\geqslant 3$，有

$$\mu_r(A_\nu)/[\mu_2(A_\nu)]^{r/2}=O(1),\ \text{当}\ \nu\to\infty$$

又若对任何整数 $r\geqslant 3$，有

$$\mu_r(A_\nu)/[\mu_2(A_\nu)]^{r/2}=o(N_\nu^{r/2-1}),\ \text{当}\ \nu\to\infty.$$

则称 A_ν 满足条件 N，对 B_ν 可作类似定义。

关于在什么条件下 $\tilde{L}_\nu=(L_\nu-\lambda_\nu)/\sigma_\nu$ 依分布收敛于标准正态分布 Φ 的问题，经历了一段长时间的研究。1944年Wald 和Wolfowitz证明了当 A_ν 和 B_ν 都满足条件 W 时，$\tilde{L}_\nu\xrightarrow{\mathscr{L}}N(0,1)$。1949 年 Noether 将它改进为只须 A_ν 和 B_ν 中有一个满足条件 W 而另一个满足条件 N。但是，只假定 A_ν 和 B_ν 都满足条件 N 不足以保证 $\tilde{L}_\nu\xrightarrow{\mathscr{L}}N(0,1)$。1961 年 Hajek 在[49]中找到了在 A_ν，B_ν 都满足条件 N 的前提下，使 $\tilde{L}_\nu\xrightarrow{\mathscr{L}}N(0,1)$ 成立的充要条件。这样就产生一个问题：在 A_ν 和 B_ν 都满足条件 N 的前提下，$\tilde{L}_\nu$ 的一切可能的极限分布的族是什么？本文作者在[50]中回答了这个问题。结论是：这个极限分布族就是方差不超过 1 的无穷可分分布的族，即其特征函数有形式

$$\varphi(t)=\exp\left\{\int_{-\infty}^{\infty}(e^{itx}-1-itx)\frac{1}{x^2}dK(x)\right\}$$

的分布的族，此处 $K(-\infty)=0$，$K(\infty)\leqslant 1$，K 单调非降且右连续。

方开泰在[51]中研究了当 A_ν 和 B_ν 不全满足条件 N 的某些情况，建立了其极限定理。例如，均匀分布是可能的极限分布之一，他并且指出了结果的某些应用。

在 $\tilde{L}_\nu\xrightarrow{\mathscr{L}}N(0,1)$ 成立时，这个收敛的速度问题，在一般情况下还没有多少结果，更谈不上彻底解决了：在 A_ν 和 B_ν 都满足条件 W，或更一般地，其中之一满足条件 W，另一个（设为 B_ν）满足条件

$$\frac{1}{N_\nu}\sum_{i=1}^{N_\nu}|b_{\nu i}-\bar{b}_\nu|^3/[\mu_2(B_\nu)]^{3/2}=O(1) \tag{4.7}$$

之下，本文作者在[50]中证明了

$$\|F_\nu-\Phi\|=O(N_\nu^{-1/6+\varepsilon}),\ 对任给\ \varepsilon>0. \tag{4.8}$$

这里 F_ν 为 $\tilde{L}_\nu$ 的分布函数。目前还不知道：在 A_ν 和 B_ν 都满足条件W的情况下，(4.7) 中的指数 1/6 可改进到何种程度。但是，在 $a_{\nu i}$ 或 $b_{\nu i}$ 具有某种特殊构造时，这个速度可大大改善。例如，设 B_ν 满足条件 (4.7)，$a_{\nu i}$ 有形式

$$a_{\nu i}=\phi(i/(N_\nu+1)),\ i=1,\cdots,N_\nu.$$

其中 ϕ 定义在 (0,1) 区间 ϕ 和 ϕ' 在上述区间有界时，Jurěckovǎ 在[52]中证明了

$$\|F_\nu-\Phi\|=O(N_\nu^{-1/2+\varepsilon}),\ 对任何\ \varepsilon>0, \tag{4.9}$$

而当进一步假定 ϕ'' 在 (0,1) 也有界时，Hušckovǎ 在[53]中得到了理想的速度 $O(N_\nu^{-1/2})$，

本文作者在[54]中改进了 Jurěckovǎ 的上述结果：设 B_ν 满足条件 (4.7)，将 $a_{\nu i}$ 按大小次序排列，不妨就假定 $a_{\nu 1}\leqslant a_{\nu 2}\leqslant\cdots\leqslant a_{\nu N_\nu}$。又将 A_ν 规则化，使 $\bar{a}_\nu=0$，$\mu_2(A_\nu)=1$，对一切 ν。令

$$C_\nu=\max_i(a_{\nu,i+1}-a_\nu),$$

若对任给 $\varepsilon>0$ 有 $C_\nu=O(N_\nu^{-1+\varepsilon})$，则 (4.9) 式成立。

在[54]中，本文作者针对区组内的小区效应不可忽略的情况，讨论了两样本置换检验（两个处理，每个处理在每一区组中可重复若干次）的大样本理论。在只假定随机误差有二阶矩的情况下，建立了这问题的置换统计量的极限定理，并证明了在正态对立假设下，所讨论的置换检验与最优无偏检验的 ARE（渐近相对效率）为 1 .

参　考　文　献

[1] 陈桂景等。数学研究与评论，第3卷 (1983) 第4期，109.

[2] Parzen, E., AMS(1962), 1065.

[3] Devroye L. P. & Wagner, T. J., Multivariate Analysis V(文集)(1980), 59.

[4] Schuster, E. P., AMS(1969), 1187.

[5] Singh, R. S., AS(1976), 431.

[6] Susarla, V. & Walter, G., AS(1981), 347.

[7] 陈希孺，密度核估计的一致收敛速度（已投《中国科学》）

[8] 陈桂景 Convergence rates of the kernel estimates of probability density and its derivatives and mode（已投《中国科学》）.

[9] 陈希孺，科学探索，1卷3期 (1981)，23.

[10] Farrell, R. H., AMS(1967), 471.

[11] Farrell, R. H., AMS(1972), 170.

[12] 陈希孺，概率密度的最佳收敛速度问题，（已投《数学年刊》）.

[13] Loftsgarden, D. O. & Quesenberry, C. P., AMS(1965), 1049.

[14] Devroye, L. P. & Wagner, T. J., AS(1977), 536.

[15] 陈希孺，中国科学，1981年12月，1419.

[16] 陈希孺，数学研究与评论，1983年1期，61，
[17] 杨振海，最近邻估计的一致收敛速度（已投《中国科学》）。
[18] 柴根象，一类密度函数 N. N. 估计的一致收敛速度（已投《数学学报》）。
[19] 陈希孺，数学年刊，2卷4期（1981），425。
[20] 陈希孺，数学研究与评论，1982年4期，85。
[21] 陈希孺，Schuster 定理的改进（已投《中国科学》）。
[22] Robbins, H., Proc. 2nd Berkeley Symp. Math. Statist. Prob., (1955), 131.
[23] 陈希孺，中国科技大学学报，12卷1期（1982），1。
[24] Robbins, H. AMS(1964), 1.
[25] Johns, M. V., AMS(1957), 649.
[26] 陈希孺 On asymptotically optimal EB estimates for parameters of one-dimensional exponential distributions，（已投《数学年刊》）。
[27] 陶波，系统科学与数学，1982年第2期。
[28] Lin, P. E.(林丕二), *Ann. Inst. Statist. Math.* (1972), 319.
[29] 赵林城，数学研究与评论，1981年7月，59。
[30] Fox, R. J., AS(1978), 846.
[31] 方兆本等，中国科技大学学报，12卷1期（1982），120。
[32] Singh, R. S., AS(1976), 431.
[33] 陈希孺，On one conjecture of Singh（已投《数学年刊》）。
[34] Hoeffding, W., AMS(1948), 293.
[35] 陈希孺，数理统计引论，科学出版社，1981。
[36] Berk, R. H., AMS(1966), 51.
[37] 陈希孺，中国科学，1980年6月，522。
[38] Grams, M. P. & Serfling, R. J., AS(1973), 153.
[39] Chan, Y. K. & Wierman, J., Ann. Prob. (1978), 186.
[40] 赵林城，陈希孺，U-统计量分布的非一致收敛速度（已投《中国科学》）。
[41] 赵林城，科学探索，1981年1期，89。
[42] 林正炎，关于U-统计量的 Berry-Esseen 不等式的推广（已投《应用数学学报》）。
[43] 林正炎，关于U-统计量和 von-Mises 统计量的 Esseen 不等式（已投《数学学报》）。
[44] Callaery, H. etc., AS(1980), 299.
[45] 林正炎，关于U-统计量的渐近展开的注（已投《数学学报》）。
[46] 林正炎，随机足标U-统计量逼近正态的阶，《数学研究与评论》，Vol. 4, No. 1. 73。
[47] 林正炎，陆传荣，随机指标随机元序列的弱收敛性（已投《数学进展》）。
[48] 陈希孺，数学学报，1979年9月，620。
[49] Hajek, J., AMS(1961), 506.
[50] 陈希孺，应用数学学报，1981年11月，342。
[51] 方开泰，应用数学学报，1981年2月，69。
[52] Jurěckovă, J., Inst. Math. Statist. Univ. of Copenhagen. 1973.
[53] Huskovă, M., AS(1977), 658.
[54] 陈希孺，武汉大学学报1980年第4期，1。