关于正态逼近的最佳非一致界限

苏　淳

（中国科学技术大学，合肥）

如所周知，如果随机变数 $X_1, X_2, \cdots$ iid，服从公共分布 $F(x)$，且

$$EX_1 = 0,\ 0 < Ex_1^2 = \sigma^2 < \infty \tag{1}$$

则当 $n \to \infty$ 时，对一切 $x \in \boldsymbol{R}_1$，有

$$F_n(x) = P\left(\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n} X_i < x\right) \to \Phi(x), \tag{2}$$

其中 $\Phi(x)$ 是正态 $N(0,1)$ 的分布函数.

本世纪的四十年代和六十年代，在条件(1)与 $E|X_1|^3 < \infty$ 之下，人们分别获得关于收敛速度的一致性估计和非一致性估计[1]:

$$\sup_{-\infty<x<\infty} |F_n(x) - \Phi(x)| \leqslant c \cdot n^{-\frac{1}{2}}, \tag{3}$$

$$|F_n(x) - \Phi(x)| \leqslant cn^{-\frac{1}{2}}(1 + |x|)^{-3}. \tag{4}$$

(3)、(4)式中的常数 $c > 0$ 仅与 F 有关，与 n, x 无关.

关于一致性估计(3)式的不可改进性，人们早已弄清(参阅文献[1]第五章§2). 但关于非一致性估计(4)式，近二十年来人们一直怀疑它的不可改进性. 许多人认为，可能存在某个与 F 无关的、定义在 $x > 0$ 上的函数 $0 < \psi(x) \uparrow \infty (x \to \infty)$，使得

$$|F_n(x) - \Phi(x)| \leqslant cn^{-\frac{1}{2}}\psi^{-1}(|x|)(1 + |x|)^{-3}$$

对一切 n 和一切 $x \in \boldsymbol{R}_1$ 成立. 有人并猜测 $\psi(x)$ 可能取 $\ln(1 + |x|)$ 或 $\ln\ln(1 + |x|)$ 的形式. 特别是最近，陈希孺和赵林城[2,3] 相继解决了难度极大的线性模型误差方差估计渐近正态和 U 统计量渐近正态的非一致性估计，他们在与 $E|X_1|^3 < \infty$ 相当的条件下，获得与 i.i.d 场合相同的结果(4)式. 这就再次把结果(4)式是否不可改进这个敏感的课题尖锐地提了出来，引起了更多的兴趣和关注.

最近，作者认真探讨了这个课题，获得了"结果(4)式的确不可改进"的明确答案，从而肯定了结果(4)式以及陈希孺、赵林城的工作都是关于非一致估计的不可改进的最佳上限. 同时，作者的工作还肯定了 Осипов 和 Петров[4] 1967 年获得的如下结果也是不可改进的最佳结果:

设 $X_1, X_2, \cdots$ iid $\sim F(x)$，满足条件(1)式，且对某 $0 < \delta < 1$ 有

$$E|X_1|^{2+\delta} < \infty, \tag{5}$$

则存在与 F 有关的、定义在 $x > 0$ 上的有界函数 $0 < h(x) \downarrow 0 (x \to \infty)$，使得对一切 n 和一切 $x \in \boldsymbol{R}_1$ 有

本文 1982 年 5 月 3 日收到.

$$|F_n(x)-\Phi(x)|\leqslant\frac{h(\sqrt{n}(1+|x|))}{n^{\delta/2}(1+|x|)^{2+\delta}}. \tag{6}$$

作者的工作表明：不存在与 F 无关的、定义在 $x>0$ 上的有界函数 $0<h(x)\downarrow 0(x\to\infty)$，使得

$$|F_n(x)-\Phi(x)|\leqslant\frac{c(F)\cdot h(\sqrt{n}(1+|x|))}{n^{\delta/2}(1+|x|)^{2+\delta}}, \tag{7}$$

对一切 n 和一切 $x\in R_1$ 成立，其中常数 $c=c(F)>0$ 仅与 F 有关.

作者的结果包含在下述定理中.

定理 1 对任何定义在 $x>0$ 上的函数 $\psi(x):0<\psi(x)\uparrow\infty(x\to\infty)$ 和任何 $0<\delta\leqslant 1$，都存在一个分布 $F:\int xdF(x)=0,\int x^2dF(x)=1,\int|x|^{2+\delta}dF(x)<\infty$，使得当随机变数 $X_1,X_2,\cdots\text{iid}\sim F(x)$ 时，无论怎样选择常数 c，都不能对一切充分大的 n 和一切 $x\in R_1$ 成立估计式：

$$|F_n(x)-\Phi(x)|\leqslant cn^{-\frac{\delta}{2}}\psi(|x|)^{-1}(1+|x|)^{-(2+\delta)}. \tag{8}$$

定理 2 对任何定义在 $x>0$ 上的有界函数 $0<h(x)\downarrow 0(x\to\infty)$ 和任何 $0<\delta\leqslant 1$，都存在一个分布 $F:\int xdF(x)=0,\int x^2dF(x)=1,\int|x|^{2+\delta}dF(x)<\infty$，使得当随机变数 $X_1,X_2,\cdots\text{iid}\sim F(x)$ 时，无论怎样选择常数 c，都不能对一切充分大的 n 和一切 $x\in R_1$ 成立(7)式.

定理 1 的证明：设 $0<\delta\leqslant 1$ 及 $\psi(x)$ 给定，不妨 $\psi(x)\leqslant x^{2-\delta}$.

定义一个自然数序列 $\{l_k\}$：

$$l_1=1,\ l_k=\min\{l,\ l>l_{k-1}^2,\ \psi(l)\geqslant k^3\},\ (k\geqslant 2).$$

再定义一个正数序列 $\{a_l\}$：

$$a_l=\begin{cases}l^{-(2+\delta)}2^{-l}, & \text{当 } l\neq\{l_k\};\\ l_k^{-(2+\delta)}k^{-2}, & \text{当 } l=l_k,\ k=1,2,\cdots.\end{cases}$$

则 $c_0^{-1}=\sum_{l=1}^{\infty}a_l<\infty$. 定义对称分布 $Q(x)$ 如下：$Q(x)$ 为离散分布，在 $x=\pm l$ 赋有正测度

$$q_{\pm l}=Q(\pm l+0)-Q(\pm l)=\frac{c_0}{2}a_l,\ l=1,2,\cdots.$$

容易验证 $Q(x)$ 确为一个概率分布，且

$$\int|x|^{2+\delta}dQ(x)=2\sum_{n=1}^{\infty}l^{2+\delta}q_l\leqslant c_0\left(\sum_{l=1}^{\infty}2^{-l}+\sum_{k=1}^{\infty}k^{-2}\right)<\infty. \tag{9}$$

但对固定的自然数 m，只要 $l_k\geqslant m$ 便有

$$\left(\frac{l_{k+1}}{m}\right)^{2+\delta}\psi\left(\frac{l_{k+1}}{m}\right)(1-Q(l_{k+1}))\geqslant\left(\frac{l_{k+1}}{m}\right)^{2+\delta}\psi(l_k)\sum_{l\geqslant l_{k+1}}q_l$$

$$\geqslant\left(\frac{l_{k+1}}{m}\right)^{2+\delta}\cdot k^3\cdot q_{l_{k+1}}=\frac{c_0}{2m^{2+\delta}}\cdot\frac{k^3}{(k+1)^2},$$

这表明对任何自然数 m 和任何常数 $A>0$，有

$$\sup_{A\leqslant x<+\infty}\{x^{2+\delta}\psi(x)(1-Q(mx))\}=\infty,\ \text{更有}$$

$$\sup_{A\leqslant x<\infty}\{(1+|x|)^{2+\delta}\psi(|x|)(1-Q(mx))\}=\infty. \tag{10}$$

记 $g(t)=\int e^{itx}dQ(x)-c_0\sum_{l=1}^{\infty}a_l\cos lt$. 定义一个无穷可分分布 $F(x)$，其特征函数为

$$f(t)=\exp\left\{\lambda\int(e^{itx}-1)dQ(x)\right\}=e^{\lambda(g(t)-1)}, \tag{11}$$

其中 $\lambda>0$ 待定. 我们来证明 $F(x)$ 即为所求.

记 $F(x)$ 的 Levy-Хинчин 谱函数为 $G(x)$，则 $G(x)$ 纯粹离散，跳点集合为 $\{\pm1,\pm2,\cdots\}$，在 $x=\pm l$ 处的跃度为 $G(\pm l+0)-G(\pm l)=\frac{\lambda l^2}{1+l^2}q_l$，所以

$$\int|x|^{2+\delta}dG(x)=2\sum_{l=1}^{\infty}\frac{\lambda l^2}{1+l^2}q_l l^{2+\delta}$$

$$\leqslant 2\lambda\sum_{l=1}^{\infty}q_l l^{2+\delta}=\int|x|^{2+\delta}dQ(x)<\infty.$$

根据周知的结果[5]

$$\int|x|^{2+\delta}dG(x)<\infty\Leftrightarrow\int|x|^{2+\delta}dF(x)<\infty,$$

知有 $\int|x|^{2+\delta}dF(x)<\infty$.

其次，由 $g(t)$ 为实值函数知 $f(t)$ 为实值的特征函数，因而 $F(x)$ 是对称分布，所以

$$\int xdF(x)=0.$$

我们有 $\int x^2dF(x)=-f''(0)$，$\int x^2dQ(x)=-g''(0)$，又

$$g'(t)=-c_0\sum_{l=1}^{\infty}a_l l\sin lt,\ g'(0)=0;$$

$$f'(t)=\lambda g'(t)e^{\lambda(g(t)-1)},\ \text{故}\ f'(0)=0;$$

又
$$f''(t)=\lambda g''(t)e^{\lambda(g(t)-1)}+(\lambda g'(t))^2e^{\lambda(g(t)-1)},$$

故 $-f''(0)=-\lambda g''(0)=\lambda\int x^2dQ(x)$. 所以只要取 $\lambda=(\int x^2dQ(x))^{-1}$，就有

$$\int x^2dF(x)=1.$$

设 $X_1,X_2,\cdots$ i.i.d $\sim F(x)$，则 $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}X_j$ 的特征函数为

$$f_n(t)=\exp\left\{n\lambda\left(g\left(\frac{t}{\sqrt{n}}\right)-1\right)\right\}=e^{-n\lambda}\sum_{k=0}^{\infty}\frac{(n\lambda)^k}{k!}g^k\left(\frac{t}{\sqrt{n}}\right),$$

所以 $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}X_j$ 的分布函数 $F_n(x)$ 可表示为

$$F_n(x)=e^{-n\lambda}\sum_{k=0}^{\infty}\frac{(n\lambda)^k}{k!}Q^{*k}(\sqrt{n}x),$$

其中 $Q^{*0}=I_{(x>0)}$; $Q^{*k}(x)=\underbrace{Q*\cdots*Q}_{k}(x)$, $k\geqslant 1$. 于是

$$1-F_n(x)=1-e^{-n\lambda}\sum_{k=0}^{\infty}\frac{(n\lambda)^k}{k!}Q^{*k}(\sqrt{n}\,x)$$

$$=e^{-n\lambda}\sum_{k=0}^{\infty}\frac{(n\lambda)^k}{k!}(1-Q^{*k}(\sqrt{n}\,x))$$

$$>e^{-n\lambda}n\lambda(1-Q(\sqrt{n}\,x)). \tag{12}$$

为证定理 1，只须再证对一切 $n=m^2$，不论怎样选择常数 $c>0$，都不可能使(8)式对一切 $x\in R_1$ 成立. 亦只须证

$$\sup_{x\in R_1}\{m^{\delta}\psi(|x|)(1+|x|)^{2+\delta}|F_{m^2}(x)-\Phi(x)|\}=\infty. \tag{13}$$

由于 m 为定值，而

$$1-\Phi(x)\sim\frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-\frac{x^2}{2}}(x\to\infty),$$

所以存在 $x(m)>0$ 使当 $x>x(m)$ 后就有

$$1-\Phi(x)\leqslant\frac{1}{m^{\delta}(1+|x|)^4}\leqslant\frac{1}{m^{\delta}\psi(|x|)(1+|x|)^{2+\delta}},$$

因此为证(13)式，只须再证

$$\sup_{x(m)<x<\infty}\{(1+|x|)^{2+\delta}\psi(|x|)(1-F_{m^2}(x))\}=\infty,$$

再由(12)式知亦只要证

$$\sup_{x(m)<x<\infty}\{(1+|x|)^{2+\delta}\psi(|x|)(1-Q(mx))\}=\infty.$$

由(10)式知其为显然. 定理 1 证毕.

定理 2 证明类似.

注：本文所构造的 $F(x)$ 是纯粹离散的[6]，如果将 $Q(x)$ 在某对跳点 $\pm l_0$ 上所赋的测度分别均匀分散到 $\left[l_0, l_0+\frac{1}{2}\right]$ 与 $\left[-l_0-\frac{1}{2}, -l_0\right]$ 中，则可使 $F(x)$ 含有绝对连续成分[6]，从而有 $\overline{\lim_{|t|\to\infty}}|f(t)|<1$，即满足 Cramer C条件.

致谢：感谢陈希孺教授和赵林城博士的指导和帮助.

参 考 文 献

[1] Петров В. В., *Суммы независимых случайных величин*, Наука, Москва, 1972.
[2] 赵林城、陈希孺，中国科学，A辑，1982，5：408—420.
[3] 赵林城、陈希孺，U-统计量的分布的非一致性收敛速度(待发表).
[4] Осипов Л. В. и Петров В. В., *Теория вероят. и её примин.*, **XII** (1967), 2: 322—329.
[5] Ramachandran, B., *On Characteristic Functions and Moments*, Sankhya, **A31** (1969), 1: 1—12.
[6] Лукач Е., *Характеристические функции*, Москва, Наука, 1979.