文章编号:1000-5471(2012)01-0011-05

离散型随机变量序列最大值的收敛速度①

张 耿, 陈守全, 王 超

西南大学 数学与统计学院, 重庆 400715

摘要: 设 $X_1, X_2, \cdots, X_n$ 是独立同分布离散型随机变量序列, $M_n = \max\{X_1, X_2, \cdots, X_n\}$. 当 $n \to \infty$ 时, $(M_n - b_n)/a_n$ 的极限分布已知. 然而, 当离散分布的参数随着 n 而变化时, 有可能得到它的非退化极限分布及其收敛速度. 研究了 3 类离散型随机变量序列最大值的收敛速度.

关 键 词: 均匀分布; 二项分布; 泊松分布; 最大值; 收敛速度

中图分类号: O211.4 **文献标志码:** A

设 X 是非负整值离散型随机变量, 它的概率质量函数(pmf) 为 $Pr(X=k)=p_k$, 分布函数(cdf) 为 F. 设 $X_1, X_2, \cdots, X_n$ 是独立且与 X 同分布的随机变量序列, 最大值 $M_n = \max\{X_1, X_2, \cdots, X_n\}$. 对某些实数 $a_n > 0$ 和 b_n, 当 $n \to \infty$ 时 $(M_n - b_n)/a_n$ 的极限分布已知. 负二项分布的一些性质和应用详见文献[1-2]. 文献[3]给出了 $(M_n - b_n)/a_n$ 的极限分布及其收敛速度. 文献[4]还研究了离散型随机变量序列完备及非完备样本最大值的联合分布.

定理 A 设 M_n 为 $\{X_{n1}, \cdots, X_{nn}\}$ 的最大值, $\{X_{n1}, \cdots, X_{nn}\}$ 是具有参数 $r > 0$, $p = \dfrac{1}{n}$ 的独立负二项分布随机变量序列, 令 $u_n(x) = n(x + \ln n + (r-1)\ln\ln n - \ln\Gamma(r))$. 则对任意 $x \in \mathbb{R}$, 有

$$\lim_{n\to\infty} P(M_n \leqslant u_n(x)) = \exp(-\exp(-x))$$

若定义 $\Delta_n(r, x) = P(M_n \leqslant u_n(x)) - \exp(-\exp(-x))$, 则当 $n \to \infty$ 时, M_n 的收敛速度为:

$$\Delta_n(r, x) \sim \frac{\exp(-x)\exp(-\exp(-x))}{2}\frac{\ln n}{n} \qquad r=1$$

$$\Delta_n(r, x) \sim (r-1)^2\exp(-x)\exp(-\exp(-x))\frac{\ln\ln n}{\ln n} \qquad r \neq 1$$

事实上, 对于正整数 r, 这个定理的结论在很早以前已被证明[5-6]. 文献[3]证明了 r 不是整数的情形. 最大值的极限分布的收敛速度已被许多作者研究(参见文献[7]及其参考文献). 本文主要讨论著名离散分布包括泊松分布、二项分布、均匀分布的最大值的极限分布的收敛速度. 均匀分布最大值的收敛速度容易得到, 而我们将用正态分布分别与泊松、二项分布的关系去研究泊松和二项分布的最大值的收敛速度.

定理 1 设 M_n 为 $\{X_{n1}, \cdots, X_{nn}\}$ 的最大值, $\{X_{n1}, \cdots, X_{nn}\}$ 是具有参数为 $N=n^2$ 的独立均匀分布随机变量序列, 令 $a_n = n$, $b_n = n^2$, $u_n(x) = a_n x + b_n$. 若定义 $\Delta_n(x) = P(M_n \leqslant u_n(x)) - \exp(x)$, 则当 $n \to \infty$ 时, M_n 的收敛速度为:

$$\Delta_n(x) \sim \exp(x)\left(-\frac{1}{n} - \frac{x^2}{2n}\right) \tag{1}$$

---

① 收稿日期: 2010-11-23
基金项目: 重庆市首届高校人才基金资助项目(120060-20600204).
作者简介: 张 耿(1986-), 女, 四川岳池人, 硕士研究生, 主要从事概率论与数理统计的研究.
通信作者: 陈守全, 副教授, 硕士生导师.

定理 2 设 M_n 为 $\{X_{n1}, \cdots, X_{nn}\}$ 的最大值，$\{X_{n1}, \cdots, X_{nn}\}$ 是具有参数 $N=n$，$p>0$ 的独立二项分布随机变量序列. 令 $a_n=\frac{1}{\sqrt{2\log n}}$，$b_n=\sqrt{2\log n}-\frac{\log\log n+\log 4\pi}{2\sqrt{2}\log n}$，$u_n(x)=np+\sqrt{np(1-p)}(a_n x+b_n)$. 若定义 $\Delta_n(x)=P(M_n\leqslant u_n(x))-\exp(-\exp(-x))$，则当 $n\to\infty$ 时，M_n 的收敛速度为：

$$\Delta_n(x)\sim\left(\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}-\frac{1}{n\sqrt{4\pi\log n}}\right)^n+\frac{\exp(-\exp(-x))\exp(-x)}{16}\frac{(\log\log n)^2}{\log n} \tag{2}$$

其中 A 是绝对正常数.

定理 3 设 M_n 为 $\{X_{n1}, \cdots, X_{nn}\}$ 的最大值，$\{X_{n1}, \cdots, X_{nn}\}$ 是具有参数 $\lambda=n$ 的独立泊松分布随机变量序列. 令 $a_n=\frac{1}{\sqrt{2\log n}}$，$b_n=\sqrt{2\log n}-\frac{\log\log n+\log 4\pi}{2\sqrt{2}\log n}$，$u_n(x)=n+\sqrt{n}(a_n x+b_n)$. 若定义 $\Delta_n(x)=P(M_n\leqslant u_n(x))-\exp(-\exp(-x))$，则当 $n\to\infty$ 时，M_n 的收敛速度为：

$$\Delta_n(x)\sim\exp(-\exp(-x))\left\{\frac{\exp(-2x)}{2n}+\left(1+\frac{(\log\log n)^2}{16\log n}\right)\left(1+\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n}\right)-1\right\} \tag{3}$$

为了证明结论，先给出几个引理.

引理 1 若 $F(x)$ 是二项分布函数，$\Phi(x)$ 是标准正态分布函数，定义 $a_n=\frac{1}{\sqrt{2\log n}}$，$b_n=\sqrt{2\log n}-\frac{\log\log n+\log 4\pi}{2\sqrt{2\log n}}$，$u_n(x)=\sqrt{np(1-p)}(a_n x+b_n)+np$，则对于任意的 $x\in\mathbb{R}$，有

$$F(u_n(x))=\Phi(a_n x+b_n)+\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}+o\left(\frac{1}{\sqrt{n}}\right) \tag{4}$$

证 令

$$X_i\sim\begin{pmatrix}-p & 1-p\\ 1-p & p\end{pmatrix}$$

有

$$EX_j=0 \qquad E\,|\,X_j\,|^3=p(1-p)(2p^2-2p+1)<\infty$$

$$\sum_{j=1}^{n}X_j=X'-np \qquad \sigma^2=p(1-p) \qquad B_n=np(1-p)$$

$$L_n=\frac{2p^2-2p+1}{\sqrt{np(1-p)}} \qquad F_n(x)=P\left(\frac{X'-np}{\sqrt{np(1-p)}}<x\right)=F(\sqrt{np(1-p)}\,x+np)$$

其中 X' 是二项分布随机变量.

由文献[9]中的定理 5.4 得

$$\sup_x\,|\,F(\sqrt{np(1-p)}\,x+np)-\Phi(x)\,|\leqslant\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}$$

由于

$$x=a_n y+b_n\sim\sqrt{2\log n}$$

从而得到

$$F(u_n(x))=\Phi(a_n x+b_n)+\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}+o\left(\frac{1}{\sqrt{n}}\right) \tag{5}$$

引理 2 若 $F(x)$ 是泊松分布函数，$\Phi(x)$ 是标准正态分布函数，定义 $a_n=\frac{1}{\sqrt{2\log n}}$，$b_n=\sqrt{2\log n}-\frac{\log\log n+\log 4\pi}{2\sqrt{2\log n}}$，$u_n(x)=\sqrt{n}(a_n x+b_n)+n$. 则对于任意的 $x\in\mathbb{R}$，有

$$\frac{1-F(u_n(x))}{1-\Phi(a_n x+b_n)}=\exp\left(\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n}+o\left(\frac{1}{n}\right)\right) \tag{6}$$

证 由文献[10]知，可定义 $X'_1, X'_2, \cdots, X'_n$ 为中心的标准泊松变量($\lambda_1=\lambda_2=\cdots\lambda_n=1$)，令 $X_i=X'_i$

$-\lambda_i$. 则有

$$E(X_i)=0 \qquad \mathrm{Var}X_i=\sigma^2=1 \qquad S_n=S'_n-n \qquad F_n(x)=P\left(\frac{S'_n-n}{\sqrt{n}}<x\right)$$

且 S'_n 是一个参数为 $\lambda=n$ 的泊松随机变量.

因为

$$E\ \exp(tX)=\exp(-nt-n+n\ \exp(t))<\infty$$

所以满足 Cramer 条件.

从文献[9] 的定理 5.23 的证明中，能够得到

$$R(t)=E\ \exp(tX_1)=(\exp(-nt-n+n\ \exp(t)))^{\frac{1}{n}}$$

因此

$$\log R(t)=\sum_{j=2}^{\infty}\frac{\gamma_j}{\Gamma(j+1)}t^j=-t-1+\exp(t)=\sum_{j=2}^{\infty}\frac{t^j}{\Gamma(j+1)}$$

于是有

$$\gamma_1=0 \qquad \gamma_2=1 \qquad \gamma_3=1 \qquad \gamma_4=1 \qquad \lambda(t)=\frac{1}{6}-\frac{1}{12}t+\cdots$$

所以

$$\frac{1-F(\sqrt{n}x+n)}{1-\Phi(x)}=\exp\left(\frac{x^3}{6\sqrt{n}}+\frac{x^4}{12n}+o(\frac{1}{n})\right) \tag{7}$$

又因为

$$x=a_ny+b_n\sim\sqrt{2\log n} \tag{8}$$

故由(7) 和(8) 式可得(6) 式.

定理 1 的证明　当 $n\to\infty$ 时 $\tau_n=n(1-F(a_nx+b_n))\to\tau=-x$. 从文献[7] 中的定理 2.4.2 知

$$P(M_n\leqslant u_n(x))-\exp(-\tau_n)\sim\frac{-x^2\exp(x)}{2n} \tag{9}$$

从文献[8] 知，因为 $x-1<[x]\leqslant x$，则

$$n\left(1-\frac{a_nx+b_n}{n^2}\right)\leqslant n(1-F(a_nx+b_n))\leqslant n\left(1-\frac{a_nx+b_n-1}{n^2}\right)$$

有

$$n(1-F(a_nx+b_n))=n\left(1-\frac{a_nx+b_n-1}{n^2}+o(\frac{1}{n^2})\right)=\frac{1-nx}{n}+o(\frac{1}{n})=-x+\frac{1}{n}+o(\frac{1}{n})$$

因此

$$\exp(-\tau_n)-\exp(-\tau)=\exp(x)\left(-\frac{1}{n}+o(\frac{1}{n})\right) \qquad n\to\infty \tag{10}$$

最后，由(9) 式和(10) 式得到(1) 式.

定理 2 的证明　从引理 1 知

$$F^n(u_n(x))=\Phi^n(a_nx+b_n)+\left(\Phi(a_nx+b_n)+\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}+o(\frac{1}{n})\right)^n-\Phi^n(a_nx+b_n)$$

有

$$F^n(u_n(x))-\exp(-\exp(-x))=\Phi^n(a_nx+b_n)-\exp(-\exp(-x))+$$

$$\left(\Phi(a_n x+b_n)+\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}+o(\frac{1}{n})\right)^n-\Phi^n(a_n x+b_n) \tag{11}$$

而且，由文献[7]中的定理2.4.2可得

$$\Phi^n(a_n x+b_n)-\exp(-\exp(-x))\sim\frac{\exp(-\exp(-x))\exp(-x)}{16}\frac{(\log\log n)^2}{\log n} \tag{12}$$

又因为

$$\Phi(x)=1-\Phi(-x)=-\frac{\exp(-\frac{x^2}{2})}{x\sqrt{2\pi}}(1+o(x^2)),\ x=a_n y+b_n\sim\sqrt{2\log n}$$

于是有

$$\Phi(a_n x+b_n)=-\frac{1}{n\sqrt{4\pi\log n}}+o(\frac{1}{\log n}) \tag{13}$$

由(11)式、(12)式和(13)式，有

$$F^n(u_n(x))-\exp(-\exp(-x))=P(M_n\leqslant u_n(x))-\exp(-\exp(-x))=$$

$$\frac{\exp(-\exp(-x))\exp(-x)}{16}\frac{(\log\log n)^2}{\log n}+\left(\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}-\frac{1}{n\sqrt{4\pi\log n}}\right)^n-\left(\frac{-1}{n\sqrt{4\pi\log n}}\right)^n+o(\frac{1}{n})\sim$$

$$\frac{\exp(-\exp(-x))\exp(-x)}{16}\frac{(\log\log n)^2}{\log n}+\left(\frac{A(2p^2-2p+1)}{\sqrt{np(1-p)}}-\frac{1}{n\sqrt{4\pi\log n}}\right)^n$$

证明完毕.

定理3的证明 当$n\to\infty$时，$\tau_n=n(1-F(a_n x+b_n))\to\tau=\exp(-x)$. 从文献[7]中的定理2.4.2知

$$P(M_n\leqslant u_n(x))-\exp(-\tau_n)\sim\frac{\exp(-2x)\exp(-\exp(-x))}{2n} \tag{14}$$

我们知道

$$n(1-\Phi(a_n x+b_n))-\exp(-x)\sim\frac{\exp(-x)}{16}\frac{(\log\log n)^2}{\log n} \tag{15}$$

由引理2中的(6)式和(15)式可得

$$n(1-F(u_n(x)))=\exp(-x)(1+(\log\log n)^2(16\log n)^{-1})\exp\left(\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n}\right)+o(\frac{1}{n}) \tag{16}$$

因此

$$\exp(-\tau_n)-\exp(-\tau)=$$

$$\exp(-\exp(-x)(1+\frac{(\log\log n)^2}{16\log n})\exp(\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n}))-\exp(-\exp(-x))+o(\frac{1}{n})\sim$$

$$\exp(-\exp(-x))(\exp((1+\frac{(\log\log n)^2}{16\log n})(1+\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n})-1)-1)\sim$$

$$\exp(-\exp(-x))((1+\frac{(\log\log n)^2}{16\log n})(1+\frac{\sqrt{2}(\log n)^{\frac{3}{2}}}{3\sqrt{n}}+\frac{(\log n)^2}{3n})-1) \tag{17}$$

由(14)式和(17)式得到(3)式.

参考文献:

[1] JOHNSON N L, KOTZ S. Discrete Distributions [M]. Boston: Houghton Mifflin Company, 1969.

[2] FELLER W. An Introduction to Probability Theory and Its Applications [M]. New York: John Wiley and Sons, 1970.

[3] MLADENOVIC P, VUKMIROVIC J. Rates of Convergence in Certain Limit Theorem for Extreme Valus [J]. Journal of Mathematical Analysis and Applications, 2010, 363: 287-295.

[4] TONG Jin-jun, PENG Zuo-xiang. Joint Distribution of Maxima of Complete and Incomplete Samples from Discrete Ran-

dom Variables [J]. 西南大学学报：自然科学版，2010，32(1)：25－28.

[5] MLADENOVIC P. Limit Theorems for the Maximum Terms of a Sequence of Random Variables with Marginal Geometric Distributions [J]. Extreme，1999，2(4)：405－419.

[6] MLADENOVIC P. A Generalization of the Mejzler-de Haan Theorem [J]. Theory Probab App，2006，50(1)：141－153.

[7] LEADBETTER M R，LINDGREN G，ROOTZEN H. Extremes and Related Properties of Random Sequence and Processes [M]. New York：Springer-Verlag，1983.

[8] NADARAJAH S，MITOV K. Asymptotics of Maxima of Discrete Random Variables [J]. Extreme，2002，5(3)：287－294.

[9] PETROV V V. Limit Theorems of Probability Theory：Sequence of Independent Random Varibles [M]. Ventnor：Clarendon Press，1995.

[10] ANDERSON C W，COLES S G，HUSLER J. Maxima of Poisson-like Variables and Related Triangular Arrays [J]. Annals of Applied Probability，1997，7(4)：953－971.

Rate of Convergence of Maxima of Discrete Random Variables

ZHANG Geng，　CHEN Shou-quan，　WANG Chao

School of Mathematics and Statistics，Southwest University，Chongqing 400715，China

Abstract：Let $X_1, X_2, \cdots, X_n$ be a sequence of independent and identically distributed discrete random variables and $M_n=\max\{X_1, X_2, \cdots, X_n\}$. We have known the limiting behavior of $\frac{M_n-b_n}{a_n}$ as $n\to\infty$. However，by tuning the parameter of the discrete distribution to vary as $n\to\infty$，it is possible to obtain the non-degenerate rate of convergence for $\frac{M_n-b_n}{a_n}$. We consider the rate of convergence of the maxima of three families of discrete random variables in this paper.

Key words：uniform distribution；binomial distribution；Poisson distribution；maximum；rate of convergence

责任编辑　张　枸