Serdica
Bulgariacae mathematicae publicationes

Сердика
Българско математическо списание

DER SATZ VON RAO-BAHADUR FÜR GROSSE ABWEICHUNGEN BEI UNABHÄNGIGEN UND NICHT IDENTISCH VERTEILTEN ZUFALLSGRÖSSEN

ELISAVETA I. PANČEVA

Es sei $X_1, X_2, \ldots, X_n, \ldots$ eine beliebige Folge von unabhängigen Zufallsgrößen. Man bezeichne $S_n = X_1 + X_2 + \cdots + X_n$, $F_n(x) = \mathsf{P}(S_n \leq x)$. Die vorliegende Arbeit enthält hinreichende Bedingungen, bei denen $1 - F_n(x) = \varrho_n(x) b_n(x)(1 + o(1))/\sqrt{2\pi n}$, wo $\varrho_n(x)$ die Chernoffsche Funktion der Summe S_n ist. Damit wird der Satz von Rao-Bahadur für den Fall nicht identisch verteilter Zufallsgrößen verallgemeinert.

1. Aufgabenstellung. Bei der Betrachtung des asymptotischen Verhaltens der Verteilung $F_n(x)$ einer Summe von unabhängigen und identisch verteilten Zufallsgrößen $X_1, X_2, \ldots, X_n$ erhielten R. Bahadur und R. Rao [1] das folgende Ergebnis:

$$1 - F_n(x) = \varrho^n(x)(2\pi n)^{-1/2} b_n(x)(1 + o(1)), \tag{1}$$

wo $\varrho(x) = \inf\{e^{-tx} \mathsf{E}\, e^{tX} | t \in R\}$ die sogenannte Chernoffsche Funktion von X_1 ist. Aus (1) folgt unmittelbar (z. B. [5])

$$\lim_{n\to\infty} (1 - F_n(x))^{1/n} = \varrho(x). \tag{2}$$

Ziel der vorliegenden Arbeit ist ähnliche asympthotische Abschätzungen für den Fall nicht identisch verteilter Zufallsgrößen zu erhalten.

2. Hilfsergebnisse und Bezeichnungen. Es seien $X_1, X_2, \ldots, X_n$ unabhängige Zufallsgrößen, deren Verteilungsfunktionen $V_r(x)$, $r = 1, 2, \ldots, n$ durch $V_r(x) = k_r G_{1r}(x) + (1 - k_r) G_{2r}(x)$, $k_r \in [0, 1]$ dargestellt sind. Dabei ist $G_{1r}(x)$ absolut stetig und $G_{2r}(x)$ enthält keine absolut stetige Komponente. Die Summe $S_n = X_1 + \cdots + X_n$ hat die Verteilungsfunktion $F_n(x) = V_1 * \cdots * V_n(x)$ und die momenterzeugende Funktion $R_n(t) = \mathsf{E}\, e^{tS_n}$. Die Verteilungsfunktion $\bar{F}_n(x) = (R_n(t))^{-1} \int_{-\infty}^{x} e^{ty}\, dF_n(y)$ nennt man konjugierte Verteilungsfunktion der Summe S_n. Wir führen folgende Bezeichnungen ein: $\bar{m}_n$ sei der Erwartungswert, $\bar{\sigma}_n^2$ — die Varianz und $\bar{\varkappa}_{\nu,n}$ — die ν-te Semiinvariante bezüglich der Verteilungsfunktion $\bar{F}_n$. Entsprechend sind m_n, σ_n^2 und $\varkappa_{\nu,n}$ die bezüglich $F_n(x)$ gebildeten Charakteristika. Es gelten die Beziehungen:

$$\bar{m}_n = [\log R_n(t)]', \quad \bar{\sigma}_n^2 = [\log R_n(t)]'', \quad \bar{\varkappa}_{3,n} = \int (x - \bar{m}_n)^3\, d\bar{F}_n(x).$$

Die Verteilungsfunktionen $\bar{F}_n$ hängen auch vom freien Parameter t ab. Bekanntlich (z. B. [2, 3, 4]) läßt die normierte konjugierte Verteilungsfunktion $\bar{F}_n^*(x) = \bar{F}_n(\bar{\sigma}_n x + \bar{m}_n)$ die formale asymptotische Entwicklung zu:

$$\bar{F}_n^*(x)=\Phi(x)+\sum_{\nu=1}^{\infty} Q_{\nu,n}(x)n^{-\nu/2},$$

wo Φ die standarte Normalverteilung ist. Die Funktionen $e^{x^2/2}Q_{\nu,n}(x)$ sind Polynome $(3\nu-1)$-ten Grades bezüglich x, und hängen nur von den Semiinvarianten $\bar{\varkappa}_{3,n},\ldots,\bar{\varkappa}_{\nu+2,n}$ ab. Es ist, z. B.:

$$Q_{1,n}(x)=(2\pi)^{-1/2}(1-x^2)e^{-x^2/2}\lambda_{3,n}/3!,\quad \lambda_{3,n}=\sqrt{n}\,\bar{\varkappa}_{3,n}\bar{\sigma}_n^{-3}. \tag{3}$$

Die Funktion $\varrho_n(x)=\inf\{e^{-tx}R_n(t)|t\in R\}$ nennt man Chernoffsche Funktion der Summe S_n. Ihre analytischen Eigenschaften sind in [5] und [7] untersucht worden. Hier werden wir von den folgenden Eigenschaften Gebrauch machen:

a) Das Infimum wird für $t=t_x$ erreicht, wobei t_x die einzige Lösung der. Differentialgleichung $R_n'(t)/R_n(t)=x$ ist. Es sei angenommen, daß $R_n(t)<\infty$ für $t\in[0, T]$. Man definiere $a_{T,n}:=d/dt(\log R_n(t))|_{t=T-0}$. Dann existiert für alle $x<a_{T,n}$ eine Lösung t_x der obigen Differentialgleichung.

b) $t_x=0$ für $x=\bar{m}_n$; $t_x>0$ für $x>m_n$; $t_x<0$ für $x<m_n$.

3. Satz von Rao—Bahadur—Verallgemeinerung. *Die Zufallsgrößen $X_1,\ldots,X_n$ seien unabhängig, nicht unbedingt identisch verteilt und genügen den folgenden Bedingungen:*

i) $R_n(t)<\infty$, $t\in[0, T)$, $T>0$;

ii) $[R_n(t)]^{1/n}\to f(t)$, $n\to\infty$, $t\in[0, T)$;

iii) $[\log n]^{-1}\Sigma_{r=1}^n k_r\to\infty$, $n\to\infty$.

Dann gilt für $m_n/n<x<a_{T,n}/n$

$$\mathsf{P}\left(\frac{X_1+\cdots+X_n}{n}>x\right)=\varrho_n(nx)(2\pi n)^{-1/2}\,b_n(x)(1+o(1)).$$

B e w e i s. Wir haben

$$\mathsf{P}(S_n/n>x)=1-F_n(nx)=R_n(t)\int_{nx}^{\infty}e^{-ty}\,d\bar{F}_n(y)=R_n(t)\int_{(nx-\bar{m}_n)/\bar{\sigma}_n}\exp[-t(\bar{m}_n+\bar{\sigma}_nz)]\,d\bar{F}_n^*(z).$$

Die Differentialgleichung $R_n'(t)/R_n(t)=nx$ hat für x aus dem oben angegebenen Interval eine einzige Lösung $t_{n,x}$. Für $t=t_{n,x}$ ist die untere Grenze des Integrals gleich Null, da $\bar{m}_n=nx$ ist. Deshalb gilt weiter

$$1-F_n(nx)=R_n(t_{n,x})\exp(-t_{n,x}nx)\int_0^{\infty}\exp(-t_{n,x}\bar{\sigma}_n z)\,d\bar{F}_n^*(z)=\varrho_n(nx)\,.\,I_n,$$

wo $I_n=\int_0^{\infty}\exp(-t_{n,x}\bar{\sigma}_nz)\,d\bar{F}_n^*(z)$. Nach partieller Integration erhält man

$$I_n=t_{n,x}\bar{\sigma}_n\int_0^{\infty}\exp(-t_{n,x}\bar{\sigma}_nz)[\bar{F}_n^*(z)-\bar{F}_n^*(0)]\,dz.$$

An dieser Stelle brauchen wir den

L e m m a 1. *Unter den Bedingungen i)—iii) gilt die folgende asympthotiche Entwicklung:*

$$\bar{F}_n^*-\Phi(z)=(2\pi n)^{-1/2}e^{-z^2/2}(1-z^2)\frac{\lambda_{3,n}}{3!}+o\left(\frac{1}{\sqrt{n}}\right). \tag{4}$$

Laut [4, Satz 26] gilt (4), wenn folgende Bedingungen erfüllt sind:

α) es existiert das dritte Moment $\bar{\beta}_{3,n}$ von $\bar{F}_n$,

β) für $n \to \infty$ gilt

$$\frac{1}{\log n} \sum_{r=1}^{n} \frac{k_r}{1+v_r^2} \to \infty$$

und

$$\frac{T_{3,n}}{\bar{\sigma}_n \log n} \sum_{r=1}^{n} \frac{k_r}{1+v_r^2} \to \infty.$$

Dabei ist $T_{3,n} = \bar{\sigma}_n^3 / 4\bar{\beta}_{3,n}$ und v_r die totale Variation von $G'_{1r}(x)$. Gelten die asympthotischen Beziehungen $\bar{\sigma}_n \sim \sqrt{n}\, g_1$ und $\bar{\beta}_{3,n} \sim n g_2$ (g_1 und g_2 sind Konstanten), so nimmt β) die einfachere Form an:

β') $\qquad [\log n]^{-1} \Sigma_{r=1}^{\infty} k_r \to \infty.$

B e w e i s von Lemma 1. Es bleibt zu zeigen, daß aus i) — iii) die Bedingungen des zitierten Satzes folgen. Die Existenz von $\bar{\beta}_{3,n}$ folgt unmittelbar aus i). Weiter zeigen wir, daß falls die Bedingung ii) in Kraft ist, die asympthotischen Beziehungen $\bar{\sigma}_n \sim \sqrt{n}\, g_1$ und $\bar{\beta}_{3,n} \sim n g_2$ gelten.

Auf Grund vom [6, Satz 10.8] ist $\log f(t)$, $t \in [0, T)$ eine konvexe Funktion, da sie Grenzfunktion der Folge $n^{-1} \log R_n(t)$ aus konvexen Funktionen ist. Für $t \in [0, T_1]$, $0 < T_1 < T$ ist diese Konvergenz sogar gleichmäßig. Aus der gleichmäßigen Konvergenz der Folge $n^{-1} \log R_n(t)$ von analytischen Funktionen läßt es sich schließen, daß

1) $\log f(t)$ für $t \in (0, T_1)$ eine analytische Funktion ist, und

2) $n^{-1} (\log R_n(t))'' \to (\log f(t))''$,
$n^{-1} (\log R_n(t))''' \to (\log f(t))'''$.

Jetzt sieht man leicht, daß die gewünschten asymptotischen Beziehungen gelten, da

$$\bar{\sigma}_n^2 = (\log R_n(t))'', \quad \bar{\beta}_{3,n} = [\log R_n(t)]''' + 3[\log R_n(t)]'[\log R_n(t)]'' + [(\log R_n(t))']^3.$$

B e m e r k u n g. Es gilt

$$\bar{F}_n^*(z) - \Phi(z) - \frac{1}{\sqrt{2\pi n}} e^{-z^2/2} \frac{\lambda_{3,n}}{3!} = w(n^{-1/2}, t_{n,x}).$$

Aber wegen ii) strebt $t_{n,x} \to t_x$ für $n \to \infty$, und deshalb ist für das Restglied die asympthotische Abschätzung $w(n^{-1/2}, t_{n,x}) = o(n^{-1/2})$ gültig. Damit ist Lemma 1 bewiesen.

Wir kehren jetzt zur Berechnung von I_n unter Verwendung von (4) zurück:

$$I_n = t_{n,x} \bar{\sigma}_n \int_0^\infty \exp(-t_{n,x} \bar{\sigma}_n z)[\Phi(z) - \Phi(0)]\, dz$$

$$+ \frac{\lambda_{3,n} t_{n,x} \bar{\sigma}_n}{3! \sqrt{2\pi n}} \int_0^\infty \exp(-t_{n,x} \bar{\sigma}_n z)[e^{-z^2/2}(1-z^2) - 1]\, dz + o(n^{-1/2}).$$

Asympthotisch verhält sich $\lambda_{3,n} = \sqrt{n} \varkappa_{3,n} / \bar{\sigma}_n^3 \sim \bar{\beta}_{3,n} / n (\bar{\sigma}_n / \sqrt{n})^3$ wegen ii) wie die Konstante g_2/g_1. Wir spalten I_n in vier Teilintegrale auf, die jetzt einzeln berechnet werden.

$$I_{n,1} = t_{n,x} \bar{\sigma}_n \int_0^\infty \exp(-t_{n,x} \bar{\sigma}_n z)[\Phi(z) - \Phi(0)]\, dz = \int_0^\infty \exp(-t_{n,x} \bar{\sigma}_n z) \Phi'(z)\, dz$$

$$=\frac{1}{\sqrt{2\pi}}\exp(t_{n,x}^2\bar{\sigma}_n^2/2)\int_0^\infty \exp[-(t_{n,x}\bar{\sigma}_n+z)^2/2]\,dz=\exp(t_{n,x}^2\bar{\sigma}_n^2/2)[1-\Phi(t_{n,x}\bar{\sigma}_n)].$$

Für $x\to\infty$ gilt $1-\Phi(x)\sim e^{-x^2/2}/x\sqrt{2\pi}$ und deshalb

$$I_{n,1}\sim 1/\sqrt{2\pi}\,t_{n,x}\bar{\sigma}_n,$$

$$I_{n,2}=\frac{\lambda_{3,n}t_{n,x}\bar{\sigma}_n}{3!\sqrt{2\pi n}}\int_0^\infty \exp(-t_{n,x}\bar{\sigma}_n z)\,e^{-z^2/2}\,dz\sim\lambda_{3,n}/3!\sqrt{2\pi n},$$

$$I_{n,3}=-\frac{\lambda_{3,n}t_{n,x}\bar{\sigma}_n}{3!\sqrt{2\pi n}}\int_0^\infty \exp(-t_{n,x}\bar{\sigma}_n z)\,e^{-z^2/2}\,dz,\quad\sim-\lambda_{3,n}\exp(-t_{n,x}^2\bar{\sigma}_n^2/2)/3!\sqrt{2\pi n},$$

$$I_{n,4}=-\frac{\lambda_{3,n}t_{n,x}\bar{\sigma}_n}{3!\sqrt{2\pi n}}\int_0^\infty \exp(-t_{n,x}\bar{\sigma}_n z)\,dz\sim\lambda_{3,n}/3!\sqrt{2\pi n}.$$

Daraus folgt insgesamt

$$I_n\sim 1/\sqrt{2\pi n}\,t_{n,x}g_1+o(n^{-1/2}).$$

Setzen wir $b_n(x)=1/t_{n,x}g_1$ an, so bekommen wir die Behauptung des Rao-Bahadurschen Theorems in der gewünschten Form

$$1-F_n(nx)=\varrho_n(nx)\,b_n(x)(1+o(1))/\sqrt{2\pi n}.$$

4. Der Chernoffsche Satz—Verallgemeinerung. *Die Zufallsgrößen $X_1, X_2, \ldots, X_n$ seien wie in 3. unabhängig, nicht unbedingt identisch verteilt und es gelten die Bedingungen i) — (iii). Dann ist*

$$\lim_{n\to\infty}(1-F_n(nx))^{1/n}=\inf\{e^{-tx}f(t)\mid t\in(0,T_1)\}. \tag{5}$$

Beweis. Man setzt in (5) das Ergebnis des vorliegenden Absatzes $1-F_n(nx)=\varrho_n(nx)\,I_n$ ein. Offensichtlich gilt $I_n^{1/n}\leq 1$. Andererseits ist aber $I_n\geq\int_0^\varepsilon\exp(-t_{n,x}\bar{\sigma}_n z)\,d\bar{F}_n^*(z)$, $\varepsilon>0$ und damit auch $I_n^{1/n}\geq(t_{n,x}\bar{\sigma}_n)^{1/n}\exp(-t_{n,x}\bar{\sigma}_n\varepsilon/n)\times[\bar{F}_n^*(\varepsilon)-\bar{F}_n^*(0)]^{1/n}\sim(t_x\sqrt{n}\,g_1)^{1/n}\exp(-t_x g_1\varepsilon/\sqrt{n})[\Phi(\varepsilon)-\Phi(0)]^{1/n}\sim 1$. Aus den beiden Ungleichungen für $I_n^{1/n}$ erhält man die asympthotische Beziehung $(1-F_n(nx))^{1/n}\sim\varrho_n^{1/n}(nx)$. Nach Grenzübergang bekommt man ein Analogon zum Satz von Chernoff und zwar

$$\lim\varrho_n^{1/n}(nx)=e^{-t_x x}f(t_x)=\inf\{e^{-tx}f(t)\mid t\in(0,T_1)\}.$$

LITERATUR

1. R. Rao, R. Bahadur. On deviation of the sample mean. *Ann. Math. Statistics*, **31**, 1960, 1015—1027.
2. В. В. Петров. Суммы независимых случайных величин. Москва, 1972.
3. B. W. Gnedenko, A. N. Kolmogoroff. Grenzverteilungen von Summen unabhängiger Zufallsgrößen. Berlin, 1959.
4. H. Cramer. R ndom variables and probability distribution. Cambridge, 1963.
5. H. Chernoff. A measure of asymptotic efficiency for test of a hypothesis based on the sum of observations. *Ann. Math. Statistics*, **23**, 1952, 493—507.
6. R. T. Rockafellar. Convex analysis. Princeton, 1970.
7. А. А. Боровков. Анализ больших уклонений в граничных задачах с произвольными границами, I. *Сиб. мат. ж.*, **5**, 1964, 253—289.

Centre for Research and Education
in Mathematics and Mechanics
1000 Sofia P. O. Box 373

Eingegangen an 2. 7. 1975