Liet. matem. rink, **45**, *spec. nr.*, 2005, 562–565

On the normal approximation for weakly dependent random variables

Jonas SUNKLODAS (MII, VGTU)
e-mail: sunkl@ktl.mii.lt

Abstract. In this report, we present the estimates of the difference $|\mathbb{E}h(Z_n) - \mathbb{E}h(N)|$, where Z_n is a sum of n centered and normalized random variables which satisfy the strong mixing condition (without assuming a stationarity), and N is a standard normal random variable for the function h: $\mathbb{R} \to \mathbb{R}$ which is finite and satisfies the Lipschitz condition. In a particular case, the obtained upper bounds are of order $\mathrm{O}(n^{-1/2})$.

1. Introduction

Consider a sequence of real random variables (r.v.'s) $X_1, X_2, \ldots$.

For a subset $U \subset \{1, 2, \ldots\}$, let $\mathcal{F}_U$ be the minimal σ-algebra such that all r.v.'s X_i with $i \in U$ are measurable. For any subsets $U, V \subset \{1, 2, \ldots\}$, let $d(U, V)$ be the distance between U and V, i.e., $d(U, V) = \inf\{|i - j|\colon i \in U, j \in V\}$.

We say that the sequence of r.v.'s $X_1, X_2, \ldots$ satisfies the strong mixing condition, if for any subsets $U, V \subset \{1, 2, \ldots\}$

$$\sup_{A \in \mathcal{F}_U, B \in \mathcal{F}_V} \bigl|\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)\bigr| \leqslant \alpha\bigl(d(U, V)\bigr) \to 0, \tag{1}$$

as $d(U, V) \to \infty$.

In what follows, N is a standard normal r.v. with the distribution function $\Phi(x)$ and the density $\varphi(x) = \Phi'(x)$.

Consider a space $BL(\mathbb{R})$ of bounded functions h: $\mathbb{R} \to \mathbb{R}$ ($\mathbb{R}$ is the real line) satisfying the Lipschitz condition, that is, such that

$$\|h\|_\infty = \sup_{x \in \mathbb{R}} |h(x)| < \infty \quad \text{and} \quad \|h\|_L = \sup_{x \neq y} \frac{|h(x) - h(y)|}{|x - y|} < \infty.$$

Denote $\|h\|_{BL} = \|h\|_\infty + \|h\|_L$ for the function $h \in BL(\mathbb{R})$.

Write

$$Z_n = \sum_{i=1}^n A_i, \quad A_i = X_i / B_n, \quad B_n^2 = \mathbb{E}\biggl(\sum_{i=1}^n X_i\biggr)^2, \quad L_3 = \sum_{i=1}^n \mathbb{E}|A_i|^3.$$

It is assumed that $B_n > 0$.

We estimate here the difference

$$\bigl|\mathbb{E}h(Z_n) - \mathbb{E}h(N)\bigr|$$

for the sequences of r.v.'s $X_1, X_2, \ldots$, satisfying the strong mixing condition (1), and for the functions $h \in BL(\mathbb{R})$.

Our main results of this report are the following two statements.

THEOREM 1. *Assume that a sequence of real r.v.'s* $X_1, X_2, \ldots$ *satisfies the strong mixing condition* (1) *and the following conditions*: $\mathbb{E}X_i = 0$ *and* $\mathbb{P}\{|X_i| \leqslant M\} = 1$, $i = 1, \ldots, n$, *where* $M > 0$ *is a non-random constant. Then, for any function* $h \in BL(\mathbb{R})$,

$$\left|\mathbb{E}h(Z_n) - \mathbb{E}h(N)\right| \leqslant C\|h\|_{BL}\left(L_3 + \frac{nM^3}{B_n^3}\sum_{r=1}^{n-1} r\alpha(r)\right), \tag{2}$$

where C *is an absolute positive constant.*

THEOREM 2. *Assume that a sequence of real r.v.'s* $X_1, X_2, \ldots$ *satisfies the strong mixing condition* (1) *and the following conditions*: $\mathbb{E}X_i = 0$, $i = 1, \ldots, n$, *and* $d_{2+\delta} = \max_{1\leqslant i\leqslant n}\mathbb{E}|X_i|^{2+\delta} < \infty$, *where* $\delta > 1$. *Then, for any function* $h \in BL(\mathbb{R})$,

$$\left|\mathbb{E}h(Z_n) - \mathbb{E}h(N)\right| \leqslant C\|h\|_{BL}\left(L_3 + \frac{nd_{2+\delta}^{\frac{3}{2+\delta}}}{B_n^3}\sum_{r=1}^{n-1} r\big(\alpha(r)\big)^{\frac{\delta-1}{2+\delta}}\right), \tag{3}$$

where C *is an absolute positive constant.*

Note that the stationarity of the approached sequence of r.v.'s is not requested. In the special cases (i.e., $B_n^2 \geqslant c_0 n$, where c_0 is a positive constant, and $\sum_{r=1}^{\infty} r\alpha(r) < \infty$ in Theorem 1, and $\sum_{r=1}^{\infty} r(\alpha(r))^{\frac{\delta-1}{2+\delta}} < \infty$ in Theorem 2), the obtained bounds of $|\mathbb{E}h(Z_n) - \mathbb{E}h(N)|$ are of optimal order $\mathrm{O}(n^{-1/2})$.

Barbour and Eagleson (1985) obtained the estimate of the bounded Lipschitz metric for dissociated r.v.'s in [1].

Sunklodas (1989, see also (2000)) obtained the estimates of the bounded Lipschitz metric for m-dependent r.v.'s, and for that satisfying ψ-mixing, and β-mixing (absolute regularity) conditions defined between the "past" and the "future".

The proof of Theorems 1 and 2 in this report is based on the method of Stein [6] and Equality (6), using Inequality (5).

2. Auxiliary lemmas

Recall that N is a standard normal r.v. with the distribution function $\Phi(x)$ and the density $\varphi(x) = \Phi'(x)$.

Consider the linear differential equation proposed by Stein [6]

$$f'(y) - yf(y) = h(y) - \mathbb{E}h(N). \tag{4}$$

LEMMA 3 (see [1]). *Let f and f' be given by* (4), *and let the function $h\colon \mathbb{R}\to\mathbb{R}$ satisfy the condition $h\in BL(\mathbb{R})$. Then, for any $y,z\in\mathbb{R}$,*

$$\bigl|f'(y+z)-f'(y)\bigr|\leqslant\bigl[(c_1+c_2)\|h_0\|_\infty+2\|h\|_L\bigr]\cdot|z|=D|z|, \tag{5}$$

where $h_0(y)=h(y)-\mathbb{E}h(N)$, $c_1=\sup_{x\geqslant 0}\Xi(x)$, $c_2=\sup_{x\geqslant 0}x(1-x\,\Xi(x))$, and $\Xi(x)=\frac{1-\Phi(x)}{\varphi(x)}$. Moreover, $c_1\leqslant\sqrt{\pi/2}$ and $c_2\leqslant 1/2$.

To present Equality (6) for the estimation of the difference $\mathbb{E}h(Z_n)-\mathbb{E}h(N)$, we introduce additional notation

$$t_i^{(r)}=\sum_{p\colon |p-i|=r}A_p\quad (t_i^{(0)}=A_i),\qquad T=\sum_{r\geqslant 0}t_i^{(r)}\quad (T=Z_n),$$

$$T_i^{(m)}=T-\sum_{r=0}^{m}t_i^{(r)}\quad (T_i^{(-1)}=T),\qquad u_i^{(r)}=A_it_i^{(r)}-\mathbb{E}A_it_i^{(r)}.$$

Note that in Lemma 4 below no conditions are proposed on the dependence of r.v.'s $X_1,X_2,\dots$ (i.e., the r.v.'s $X_1,X_2,\dots$ may be anyhow dependent, or independent, too).

LEMMA 4. *Let $X_1,X_2,\dots$ be a sequence of real r.v.'s with $\mathbb{E}X_i=0$, $i=1,\dots,n$, and the function $f\colon\mathbb{R}\to\mathbb{C}$ ($\mathbb{C}$ is the complex plane) has a continuous derivative of the first order f'. Then*

$$\sum_{i=1}^{n}\mathbb{E}A_if(Z_n)=\mathbb{E}f'(Z_n)+E_1+\dots+E_7, \tag{6}$$

where

$$E_1=\sum_{i=1}^{n}\sum_{r\geqslant 1}\mathbb{E}\biggl\{A_i\int_0^{t_i^{(r)}}\bigl[f'(T_i^{(r)}+y)-f'(T_i^{(r)})\bigr]\,\mathrm{d}y\biggr\},$$

$$E_2=\sum_{i=1}^{n}\mathbb{E}\biggl\{A_i\int_0^{A_i}\bigl[f'(T_i^{(0)}+y)-f'(T_i^{(0)})\bigr]\,\mathrm{d}y\biggr\},$$

$$E_3=\sum_{i=1}^{n}\sum_{r\geqslant 1}\sum_{q=r+1}^{2r}\mathbb{E}\Bigl\{u_i^{(r)}\bigl[f'(T_i^{(q-1)})-f'(T_i^{(q)})\bigr]\Bigr\},$$

$$E_4=\sum_{i=1}^{n}\sum_{r\geqslant 1}\sum_{q\geqslant 2r+1}\mathbb{E}\Bigl\{u_i^{(r)}\bigl[f'(T_i^{(q-1)})-f'(T_i^{(q)})\bigr]\Bigr\},$$

$$E_5=-\sum_{i=1}^{n}\sum_{r\geqslant 1}\mathbb{E}A_it_i^{(r)}\sum_{q=0}^{r}\mathbb{E}\bigl[f'(T_i^{(q-1)})-f'(T_i^{(q)})\bigr],$$

$$E_6 = \sum_{i=1}^{n} \sum_{q \geqslant 1} \mathbb{E}\Big\{u_i^{(0)} \big[f'(T_i^{(q-1)}) - f'(T_i^{(q)})\big]\Big\},$$

$$E_7 = -\sum_{i=1}^{n} \mathbb{E}A_i^2 \mathbb{E}\big[f'(T) - f'(T_i^{(0)})\big],$$

if the moments of both sides of (6) *are finite.*

References

1. A.D. Barbour, G.K. Eagleson, Multiple comparisons and sums of dissociated random variables, *Adv. Appl. Probab.*, **17**(1), 147–162 (1985).
2. V. Bentkus, A new method for approximation in probability and operator theories, *Liet. matem. rink.*, **43**(4), 444–470 (2003) (in Russian).
3. P. Hall, C.C. Heyde, *Martingale Limit Theory and Its Applications*, Academic Press, New York, London, Toronto, Sydney, San Francisko (1980).
4. Z.Y. Lin, C.R. Lu, *Limit Theory for Mixing Dependent Random Variables*, Kluwer Academic Publishers and Science Press, Dordrecht-Beijing (1996).
5. E. Rio, Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes, *Probab. Theory Related Fields*, **104**(2), 255–282 (1996).
6. Ch. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: *Proc. Math. Statist. and Probab.*, vol. 2, Univ. Calif. Press, Berkeley, CA (1972), pp. 583–602.
7. J. Sunklodas, Estimation of the bounded Lipschitz metric for sums of weakly dependent random variables, *Lith. Math. J.*, **29**(2), 187–193 (1989).
8. A.N. Tikhomirov, On the rate of convergence in the central limit theorem for weakly dependent variables, *Theory Probab. Appl.*, **25**(4), 790–809 (1980).
9. S.A. Utev, On a method of studying sums of weakly dependent random variables, *Sib. Math. J.*, **4**, 675–690 (1991).
10. T.M. Zuparov, On the convergence in the central limit theorem for weakly dependent random variables, *Theory Probab. Appl.*, **36**, 635–643 (1991).

REZIUMĖ

J. Sunklodas. Apie normalųjį aproksimavimą stipriai susimaišiusiems atsitiktiniams dydžiams

Šiame pranešime pateikiame skirtumo $|\mathbb{E}h(Z_n) - \mathbb{E}h(N)|$ įverčius, kur Z_n yra centruotų ir normuotų atsitiktinių dydžių, tenkinančių stipraus susimaišymo sąlygą, suma (nereikalaujant stacionarumo), N yra standartinis normalusis a.d., o funkcija h: $\mathbb{R} \to \mathbb{R}$ yra baigtinė ir tenkina Lipšico sąlygą. Atskiru atveju, gauti viršutiniai įverčiai yra eilės $\mathrm{O}(n^{-1/2})$.