Liet. mat. rink. LMD darbai, **48/49**, 2008, 417–421

On limit uniform distribution

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

Abstract. In the first part of the present paper, we estimate the difference $\Delta_n^{(1)} = \sup_{-\infty<x<\infty} |\mathbb{P}\{Y_n < x\} - \mathbb{P}\{\xi < x\}|$, where $Y_n = X_n/n$, X_n is a discrete r.v. with $\mathbb{P}\{X_n = j\} = \frac{1}{(l-k)n}$, for $j = nk,\ nk+1, \dots, nl-1$, as $n = 1, 2, \dots$, $k < l$, and k, l are any integers; the absolutely continuous r.v. ξ is uniformly distributed in the interval $[k, l]$. The upper bound of $\Delta_n^{(1)}$ is $\frac{1}{(l-k)n}$.

In the second part of the present paper, we estimate the difference $\Delta_n^{(2)} = \sup_{-\infty<x<\infty} |\mathbb{P}\{S_n < x\} - \mathbb{P}\{\xi < x\}|$, where $S_n = \sum_{j=1}^n X_j/2^j$, the r.v.'s $X_1, \dots, X_n$ are independent, X_j is a discrete r.v. with $\mathbb{P}\{X_j = -a\} = \mathbb{P}\{X_j = a\} = 1/2$ for $j = 1, \dots, n$ and any real number $a > 0$; the absolutely continuous r.v. ξ is uniformly distributed in the interval $[-a, a]$. The obtained upper bound of $\Delta_n^{(2)}$ is $C2^{-n}$, where $C < 4$.

Keywords: uniform distribution, independent random variables.

1. A "bridge" between discrete uniform and absolutely continuous uniform distributions

Consider an absolutely continuous random variable (r.v.) ξ, uniformly distributed in the interval $[k, l]$, where $k < l$, and k and l are any integers (we use the notation $\xi \sim \mathcal{U}[k, l]$), i.e., consider the r.v. ξ which has the distribution and characteristic functions, respectively,

$$\mathbb{P}\{\xi < x\} = \frac{x-k}{l-k} 1_{(k,l]}(x) + 1_{(l,\infty)}(x). \tag{1}$$

Here and in what follows 1_A is the indicator of set A. $\mathbb{R}$ is a real line.

We are interested in constructing a uniformly distributed discrete r.v. Y_n, the distribution function $\mathbb{P}\{Y_n < x\}$ of which converges to the distribution function $\mathbb{P}\{\xi < x\}$ of the absolutely continuous r.v. $\xi \sim \mathcal{U}[k, l]$, as well as in determining the rate of convergence.

The case where the r.v. Y_n takes on the values $0, 1/n, 2/n, \dots, (n-1)/n$ with equal probabilities $1/n$ and the convergence fact of Y_n to the absolutely continuous and unformly distributed r.v. ξ in the interval $[0, 1]$ is discussed, for example, in the book [1, p. 37].

The following statement is valid.

*The research was partially supported by the Lithuanian State Science and Studies Foundation, grant No. T-25/08

THEOREM 1. *Assume that for all* $n = 1, 2, \dots$ *and for all* $k < l$*, where* k *and* l *are any integers*,

$$\mathbb{P}\{X_n = j\} = \frac{1}{(l-k)n}, \quad j = nk, nk+1, \dots, nl-1.$$

Denote

$$Y_n = \frac{X_n}{n},$$

$$\Delta_n^{(1)}(x) = \mathbb{P}\{Y_n < x\} - \mathbb{P}\{\xi < x\},$$

$$\mathbb{P}\{\xi < x\} = \frac{x-k}{l-k} 1_{(k,l]}(x) + 1_{(l,\infty)}(x).$$

Then for all $n = 1, 2, \dots$

$$\sup_{x \in \mathbb{R}} \left| \Delta_n^{(1)}(x) \right| \leqslant \frac{1}{(l-k)n}. \tag{2}$$

Proof. The proof of Theorem 1 follows comparing the distribution functions $\mathbb{P}\{Y_n < x\}$ and $\mathbb{P}\{\xi < x\}$ in each separate interval of changes of the argument x: for $x \leqslant k$, afterwords for $k < x \leqslant k + \frac{1}{n}$, next for $k + \frac{1}{n} < x \leqslant k + \frac{2}{n}$, and so on, and finally for $x > l - \frac{1}{n}$.

2. Uniform bound for asymptotic uniformity of independent random variables

In what follows, a symmetric r.v. $\xi \sim \mathcal{U}[-a, a]$, where $a > 0$ is any real number, with the characteristic function $f(t) = \frac{\sin at}{at}$.

The following statement is valid.

THEOREM 2. *Let* $X_1, \dots, X_n$ *be independent r.v.'s with the probabilities*

$$\mathbb{P}\{X_j = -a\} = \mathbb{P}\{X_j = a\} = \frac{1}{2}, \quad j = 1, \dots, n, \tag{3}$$

where $a > 0$ *is any real number. Let*

$$S_n = \sum_{j=1}^{n} Y_j, \quad Y_j = \frac{X_j}{2^j},$$

$$\Delta_n^{(2)}(x) = \mathbb{P}\{S_n < x\} - \mathbb{P}\{\xi < x\},$$

$$\mathbb{P}\{\xi < x\} = \frac{x+a}{2a} 1_{(-a,a]}(x) + 1_{(a,\infty)}(x).$$

Then, for all $n = 1, 2, \dots$

$$\sup_{x \in \mathbb{R}} \left| \Delta_n^{(2)}(x) \right| \leqslant C_2 2^{-n}, \tag{4}$$

where $C_2 = \frac{12 - 10(2^{-1}(\mathrm{e} - \mathrm{e}^{-1}) - 1)}{(1 - (2^{-1}(\mathrm{e} - \mathrm{e}^{-1}) - 1))\pi} = 3.9549\dots$

To prove Theorem 2, we need Lemma 3.

LEMMA 3. *Let the conditions and notation of Theorem* 2 *be satisfied. In addition, denote the characteristic functions of the sum* S_n *and symmetric r.v.* $\xi \sim \mathcal{U}[-a,a]$, *where* $a > 0$ *is any real number, with the characteristic function* $f(t) = \frac{\sin at}{at}$*:*

$$f_n(t) = \mathbb{E}e^{itS_n} \quad \textit{and} \quad f(t) = \mathbb{E}e^{it\xi}.$$

Then for all $|at| \leqslant 2^n$

$$\big| f_n(t) - f(t) \big| \leqslant \frac{C_1}{1-C_1} \frac{|at \sin at|}{2^{2n}}, \tag{5}$$

where $C_1 = 2^{-1}(e - e^{-1}) - 1 = 0.1752\dots$

Proof. To prove Lemma 3, further we write $\sin t$ in a way useful to us. By means of iteration, it is easy to see that, for all $t \in \mathbb{R}$ and all finite $n = 1, 2, \dots,$

$$\sin t = 2^n \sin \frac{t}{2^n} \prod_{j=1}^{n} \cos \frac{t}{2^j}, \tag{6}$$

and therefore

$$f(t) = \frac{2^n \sin \frac{at}{2^n}}{at} \prod_{j=1}^{n} \cos \frac{at}{2^j}. \tag{7}$$

Since r.v.'s $Y_1, \dots, Y_n$ are independent, one has that the characteristic function $f_n(t)$ of the sum S_n is

$$f_n(t) = \prod_{j=1}^{n} \cos \frac{at}{2^j}. \tag{8}$$

It folows from (7) and (8) that

$$\big| f_n(t) - f(t) \big| = \frac{\big| \frac{at}{2^n} - \sin \frac{at}{2^n} \big|}{\big| \sin \frac{at}{2^n} \big|} \frac{|\sin at|}{a|t|}. \tag{9}$$

It only remains to estimate the right-hand side of (9). To this end, we use the following fact: for all $|x| \leqslant 1$

$$\sin x = x + \theta C_1 |x|^3, \tag{10}$$

where θ is a function such that $|\theta| \leqslant 1$. Note that (10) follows from the expansion of $\sin x$ in the power series [3, p. 42]

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \tag{11}$$

and the numerical series [3, p. 13]

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)!} = 2^{-1}\big(\mathrm{e} - \mathrm{e}^{-1}\big). \tag{12}$$

Thus, using (10) we get that, for $|at| \leqslant 2^n$,

$$\left|\frac{at}{2^n} - \sin\frac{at}{2^n}\right| \leqslant C_1 \frac{|at|^3}{2^{3n}}, \tag{13}$$

$$\left|\sin\frac{at}{2^n}\right| \geqslant (1 - C_1)\frac{|at|}{2^n}. \tag{14}$$

Now Lemma 3 follows from (9), (13), and (14).

The end of the proof of Theorem 2. Since

$$\sup_{x\in\mathbb{R}} \left|\left(\frac{x+a}{2a} 1_{(-a,a]}(x) + 1_{(a,\infty)}(x)\right)'_x\right| = \frac{1}{2a},$$

we derive from our Lemma 3 and Lemma 2, for example, from [2, p. 302] that, for the distribution function $\mathbb{P}\{S_n < x\}$ with the characteristic function $f_n(t)$ and the distrubution function $\mathbb{P}\{\xi < x\} = \frac{x+a}{2a} 1_{(-a,a]}(x) + 1_{(a,\infty)}(x)$ with the characteristic function $f(t)$, and $T = 2^n/a$,

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \left|\Delta_n^{(2)}(x)\right| &\leqslant \frac{2}{\pi}\int_0^T \left|\frac{f_n(t) - f(t)}{t}\right| \mathrm{d}t + \frac{12}{\pi a}\frac{1}{T} \\
&\leqslant \frac{C_1 a}{(1 - C_1)\pi 2^{2n-1}} \int_0^{2^n/a} |\sin at| \,\mathrm{d}t + \frac{12}{\pi 2^n} \\
&\leqslant \frac{12 - 10C_1}{(1 - C_1)\pi}\frac{1}{2^n}.
\end{aligned}$$

Theorem 2 is proved.

References

1. N. Balakrishnan, V.B. Nevzorov, *A Primer on Statistical Distributions*, Wiley, New Jersey (2003).
2. Y.S. Chow, H. Teicher, *Probability Theory, Independence, Interchangeability, Martingales*, Springer-Verlag, New York, Berlin, Heidelberg (1988).
3. I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products*, 7th Ed., Academic Press, San Diego, CA, USA (2007).

REZIUMĖ

J. Sunklodas. Apie tolygųjį ribinį skirstinį

Sukonstruotas diskretusis atsitiktinis dydis, tolygiai pasiskirstęs baigtiniame intervale, kurio pasiskirstymo funkcija konverguoja į absoliučiai tolydaus atsitiktinio dydžio, tolygiai pasiskirsčiusio baigtiniame intervale, pasiskirstymo funkciją ir gautas konvergavimo greitis tolygiojoje metrikoje.

Gautas diskrečiųjų atsitiktinių dydžių sumos pasiskirstymo funkcijos konvergavimo greitis tolygiojoje metrikoje į absoliučiai tolydaus atsitiktinio dydžio, tolygiai pasiskirsčiusio baigtiname intervale, pasiskirstymo funkciją.

Raktiniai žodžiai: tolygusis skirstinys, nepriklausomi atsitiktiniai dydžiai .