Liet. matem. rink., T. 40, spec. nr., 464–470

On the closeness of lattice distributions

Vydas ČEKANAVIČIUS* (VU)
e-mail:vydas.cekanavicius@maf.vu.lt

Let ξ and η be two independent lattice random variables concentrated at non-negative integers and having distributions F and G, respectively. Let denote factorial cumulants of ξ and η by

$$\nu_k = \sum_{m=0}^{\infty} m(m-1)...(m-k+1)P(\xi = m),$$
$$\mu_k = \sum_{m=0}^{\infty} m(m-1)...(m-k+1)P(\eta = m).$$

Let γ_k, $k = 1, 2, \ldots$ denote factorial cumulants of ξ.

Let $S_n = \xi_1 + \xi_2 + \ldots + \xi_n$, $Z_n = \eta_1 + \eta_2 + \ldots + \eta_n$. Here $\xi_1, \ldots, \xi_n$ are independent copies of ξ; $\eta_1, \ldots, \eta_n$ are independent copies of η. Note that the distribution of S_n is F^{*n}. Here by F^{*n} we denote the n-fold convolution of F.

The closeness of S_n to Z_n is usually measured in total variation

$$\| F^{*n} - G^{*n} \| = \sum_{m=0}^{n} | P(S_n = m) - P(Z_n = m) | \tag{1}$$

or in weaker uniform Kolmogorov distance

$$| F^{*n} - G^{*n} | = \sup_x \big| P(S_n < x) - P(Z_n < x) \big|. \tag{2}$$

We shall review some approaches to estimating $F^{*n} - G^{*n}$ in distances (1) and (2). Obviously,

$$F^{*n} - G^{*n} = \sum_{m=0}^{n-1} F^{*m} * G^{*(n-m-1)} * (F - G). \tag{*}$$

This form implies that we must take into account the smallness of $F - G$ and possible smoothing effect of convolution $F^{*m} * G^{*(n-m-1)}$. Further on we denote by C the generic constant which can vary from line to line, whereas the constant $C(s)$ depends on s.

*AMS subject classification 60 F 10

1. Le Cam's operator approach

The most general approach for estimating the difference $F^{*n} - G^{*n}$ was introduced by Le Cam in [7]. However, the effect of convolution was ignored. Let, for some fixed $s \geqslant 2$, $\nu_k = \mu_k$, $k = 1, \dots, s-1$ and $\nu_s + \mu_s < \infty$. Then

$$\| F^{*n} - G^{*n} \| \leqslant n(\nu_s + \mu_s)(2^s/s!). \tag{3}$$

Estimate (3) follows from the general results of [5]. Obviously, it is non-trivial for $\nu_s + \mu_s < Cn^{-1}$ only. Such an assumption is very restrictive and leaves out the scheme of sequences (i.e., the case where F and G do not depend on n) completely.

2. Franken's condition

We say that random variable ξ satisfies Franken's condition (F) if

$$\nu_1 - \nu_1^2 - \nu_2 > 0. \tag{F}$$

Note that Franken's condition ensures the smallness of the characteristic function of ξ outside the neighbourhood of zero. Originally introduced by Franken in [6], condition (F) was used in many subsequent papers. Let, for some fixed $s \geqslant 3$, $\nu_k = \mu_k$, $k = 1, \dots, s-1$, $\nu_s + \mu_s < \infty$ and let ξ and η satisfy Franken's condition (F). Then

$$| F^{*n} - G^{*n} | \leqslant C(s)n(\nu_s + \mu_s) \min\Big(1, \big(\nu_1 - \nu_1^2 - \nu_2\big)^{-s/2} n^{-s/2}\Big), \tag{4}$$

see [11]. Similar estimate holds for the total variation norm. Comparing estimate (3) with estimate (4) we can see the benefits of condition (F). Indeed, all ν_k and μ_k might not depend on n (the classical situation when CLT holds), but even in this case the right-hand side of (4) is of the order $n^{-(s-2)/2}$. What are the main drawbacks of the Franken condition? The set of distributions satisfying (F) is not very large. If ξ satisfies Franken's condition, then $\nu_1 = \mathbf{E}\xi < 1$. Moreover, $\nu_2 < \nu_1$. In terms of probabilities, this means that all probabilities, appart from $P(\xi = 0)$ and $P(\xi = 1)$, are small. The simplest example of ξ satisfying (F) is the Bernoulli variable.

3. Statulevičius condition ($\tilde{S}$)

We say that ξ satisfies condition ($\tilde{S}$) if, for some $\Delta > 1$,

$$| \gamma_k | \leqslant \frac{k!\mu}{\Delta^{k-1}}, \quad k = 2, 3, \dots \tag{$\tilde{S}$}$$

Obviously, condition ($\tilde{S}$) is the lattice analogue of Statulevičius condition (S) for cumulants. For approximations under ($\tilde{S}$) see [1, 3, 4, 12] and references therein. Condition ($\tilde{S}$)

usually appears in papers on large deviations. However, it also can be applied to integral estimates, see [12]. If Δ is sufficiently large, condition $(\tilde{S})$ ensures the smallness of the characteristic function outside the neighbourhood of zero, i.e., works just like Franken's condition (F). Let, for some fixed $s \geqslant 3$, $\nu_k = \mu_k$, $k = 1, \dots, s-1$ and let ξ and η satisfy $(\tilde{S})$ with $\Delta \geqslant 5$. Then

$$| F^{*n} - G^{*n} | \leqslant C(s) n(\nu_s + \mu_s)(\nu_1 n)^{-s/2}. \tag{5}$$

Indeed, $\Delta \geqslant 5$ ensures that

$$| \widehat{F}(t) |, | \widehat{G}(t) | \leqslant \exp\{-(2/3)\nu_1 \sin^2(t/2)\}, \tag{6}$$

see [12, formula (2.22)]. Here $\widehat{F}(t)$ and $\widehat{G}(t)$ denote the characteristic functions of F and G, respectively. Moreover,

$$| \widehat{F}(t) - \widehat{G}(t) | \leqslant (\nu_s + \mu_s) 2^s | \sin(t/2) |^{s/2} / s!\,. \tag{7}$$

To get (7) one should use estimates (5) and (6), identity (*) and Tsaregradskii's inequality.

As shown in [12], to some extent Franken's condition can replace $(\tilde{S})$ and vice versa. Of course, these conditions are not equivalent (note that $(\tilde{S})$ requires the existence of all finite moments and just two moments suffice for (F) to hold). In general, both conditions ensure that F^{*n} is close to some (probably centered) Poisson law – see [12].

4. Approximation when conditions (F) and $(\tilde{S})$ are not satisfied

The main result of this note is to show that F^{*n} and G^{*n} can be close even if both conditions (F) and $(\tilde{S})$ fail. To prove this we utilize one result of Barbour and Xia [2] about the total variation distance between $S_n + 1$ and S_n. Unlike Babour and Xia we do not use the Stein equation and need no information about the concrete structure of G. Let E_1 denote degenerate distribution concentrated at 1, E denote degenerate distribution concentrated at zero. Obviously, the distribution of $\xi + 1$ is $F * E_1$. Set

$$u_1 = 1 - \| F * E_1 - F \| / 2, \quad u_2 = 1 - \| G * E_1 - G \| / 2. \tag{8}$$

As follows from Proposition 4.6 [2], for any natural k,

$$\| F^{*k} * (E_1 - E) \| \leqslant 2(k u_1)^{-1/2}, \quad \| G^{*k} * (E_1 - E) \| \leqslant 2(k u_2)^{-1/2}. \tag{9}$$

We assume that the right-hand sides in (9) are infinite, if u_1 or u_2 equals zero. Combining Le Cam's approach with (9) we obtain the following general result.

Theorem. *Let, for some fixed* $s \geqslant 2$, $\nu_k = \mu_k$, $k = 1, \dots, s-1$, $\nu_s + \mu_s < \infty$ *and* $n > 6(s+1)$. *Then*

$$\| F^{*n} - G^{*n} \| \leqslant (\nu_s + \mu_s) n^{-(s-2)/2} (u_1^{-s/2} + u_2^{-s/2}) (12s)^{s/2} / s! \tag{10}$$

Proof. Set $v(n, s) = [\,[n/2]/s]$. Here $[a]$ denotes the integer part of a. From the equality of moments we get the following expansion in distributions

$$F - G = W * (E_1 - E)^{*s}(\nu_s + \mu_s)/s!, \tag{11}$$

where W is a finite measure satisfying $\| W \| \leqslant 1$, see [5]. Taking into account (*), (9), (11) and the properties of total variation norm (see, for example, [5]) we get:

$$\begin{aligned}
\| F^{*n} - G^{*n} \| &= \left\| \sum_{m=0}^{n-1} F^{*m} * G^{*(n-m-1)} * (F - G) \right\| \\
&\leqslant \sum_{m=0}^{n-1} \| F^{*m} * G^{*(n-m-1)} * (F - G) \| \\
&\leqslant n\Big(\| F^{*[n/2]} * (F - G) \| + \| G^{*[n/2]} * (F - G) \|\Big) \\
&\leqslant n(\nu_s/s! + \mu_s/s!)\Big(\| F^{*[n/2]} * (E_1 - E)^{*s} \| + \| G^{*[n/2]} * (E_1 - E)^{*s} \|\Big) \\
&\leqslant n(\nu_s/s! + \mu_s/s!)\Big(\| F^{*v(n,s)]} * (E_1 - E) \|^{*s} + \| G^{*v(n,s)} * (E_1 - E) \|^{*s}\Big) \\
&\leqslant 2^s n(\nu_s/s! + \mu_s/s!)\big((v(n, s)u_1)^{-s/2} + (v(n, s)u_2)^{-s/2}\big) \\
&\leqslant (\nu_s + \mu_s)n^{-(s-2)/2}\big(u_1^{-s/2} + u_2^{-s/2}\big)(12s)^{s/2}/s!
\end{aligned}$$

In the last inequality we used the following estimate

$$v(n, s) \geqslant [n/2]/s - 1 \geqslant n/(2s) - 1/s - 1 \geqslant n/(3s),$$

which follows from $n \geqslant 6(s + 1)$. This completes the proof of (9).

REMARK 1. Note, in the case of unimodal distribution F, we can use the following estimate:

$$u_1 \geqslant 1 - \max_k P(\xi = k).$$

EXAMPLE. We shall exemplify Theorem assuming that η has the geometric distribution (i.e., Z_n has the negative binomial distribution). The negative binomial (and, particularly, the geometric) distribution attracts a lot of attention. Firstly, it naturally appears as one of aggregate claims distributions in risk theory, see [9], and algorithms are developed for its computation, see [8]. The geometric distribution also is used to approximate Polya distribution, see [10]. Therefore, from the practical point of view the negative binomial distribution is certainly acceptible. Secondly, from the theoretical point of view the negative binomial (and geometric) distribution is infinitely divisible and is one of the simplest compound Poisson distributions. Indeed, let η have the geometric distribution, i.e., let it have the characteristic function:

$$\widehat{G}^n(t) = p/(1 - qe^{it}), \quad q \leqslant 1/2, \quad p + q = 1. \tag{12}$$

Then

$$\widehat{G}(t) = \exp\left\{ \sum_{m=1}^{\infty} \frac{q^m}{m}(e^{itm} - 1) \right\}.$$

Expanding $\widehat{G}(t)$ and $\ln \widehat{G}(t)$ in the powers of $(e^{it}-1)$ we establish that factorial moments and factorial cumulants of η equal, respectively,

$$k!\left(\frac{q}{p}\right)^k \quad \text{and} \quad (k-1)!\left(\frac{q}{p}\right)^k, \quad k = 1, 2, \ldots$$

It is easy to check that Franken's condition (F) implies that $q/p < 1/3$. Formally the Statulevičius condition $(\tilde{S})$ requires $q/p < 1$ only. However, in applications, the usual requirement would be at least $q/p \leqslant 1/4$.

Let ξ be concentrated at 4 points and have the following distribution:

$$P(\xi = 0) = 20/45, \quad P(\xi = 1) = 18/45,$$
$$P(\xi = 3) = 5/45, \quad P(\xi = 6) = 2/45.$$

Consequently,

$$\widehat{F}(t) = 1 + (e^{it} - 1) + (e^{it} - 1)^2 + (e^{it} - 1)^3 + \theta| e^{it} - 1 |^4/3. \tag{13}$$

Here $|\theta| \leqslant 1$. It is easy to check that $\nu_1 = 1$. Consequently, Franken's condition (F) is not satisfied. Moreover, $\gamma_2 = 1/2$. Therefore, condition $(\tilde{S})$ is not satisfied too. However, it is easy to check that

$$\begin{aligned} u_1 &= 1 - \big(20/45 + |20/45 - 18/45| + 18/45 + 5/45 + 5/45 + 2/45 + 2/45\big)/2 \\ &= 18/45. \end{aligned} \tag{14}$$

Let η have the geometric distribution defined by (12) with $p = q = 1/2$. Again, we see that conditions (F) and $(\tilde{S})$ are not satisfied. Moreover,

$$\widehat{G}(t) = 1 + (e^{it} - 1) + (e^{it} - 1)^2 + (e^{it} - 1)^3 + \theta_1| e^{it} - 1 |^4. \tag{15}$$

Here $|\theta_1| \leqslant 1$. The quantity u_2 can be easily computed for any geometric distribution, because $\| G * E_1 - G \|$ equals to

$$\begin{aligned} & p + p(1-q) + pq(1-q) + pq^2(1-q) + pq^3(1-q) + \ldots \\ &= p + p^2(1 + q + q^2 + \ldots) = 2p. \end{aligned}$$

Consequently, in our example,

$$u_2 = 1 - p = 1 - 1/2 = 1/2. \tag{16}$$

Combining (13)–(16) with the statement of the Theorem we get the following corollary.

COROLLARY. Let ξ and η be defined as in example above. Then

$$\| F^{*n} - G^{*n} \| \leqslant Cn^{-1}. \tag{17}$$

REMARK 2. In our example, $s = 4$. By conditioning, Theorem can be applied when $n > 6(4+1) = 30$ only. However, if $n \leqslant 30$, estimate (17) follows from the fact that the left hand side of (17) is less then 2.

REMARK 3. In our example F and G do not depend on n. Certainly, in this case, one can apply the central limit theorem. However, the result will not be as good as estimate (17). Firstly, the normal approximation holds in uniform (thus weaker) distance only. Indeed, the total variation norm of the difference of any continuous and any discrete distributions equals 2. Secondly, for obtaining the same accuracy, one should need one member of asymptotics. Thirdly, even then the estimate can not be expressed as absolute constant (which does not depend on F and G) multiplied by n^{-1}.

REMARK 4. In practice, the more important case is when S_n is the sum of independant not identically distributed variables. In principle, some analogue of Theorem also can be obtained for this case. However the estimate then becomes very cumbersome.

References

[1] A. Aleškevičienė, V. Statulevičius, Large deviations in power zones in the approximation by the Poisson law, *Uspekhi mat. nauk.*, **50**(5), 64–82 (1995) (Russian).

[2] A.D. Barbour, A. Xia, Poisson perturbations, *ESAIM: Probability and Statistics*, **3** (1999).

[3] A. Bikelis, A. Žemaitis, Asymptotic expansions for probabilities of large deviations, Normal approximation, *Liet. Matem. Rink.*, **16**(3), 31–50 (1976) (Russian).

[4] A. Bikelis, A. Žemaitis, Asymptotic expansions for probabilities of large deviations, *Application of Probab. Theory and Math. Statist.*, **3**, Vilnius, 9–39 (1980) (Russian).

[5] V. Čekanavičius, Estimates in total variation for convolutions of compound distributions, *J. London Math. Soc.*, **58**(2), 748–760 (1998).

[6] P. Franken, Approximation der Verteilungen von Summen unabhängiger nichtnegativen ganzzahliger Zufalgrösen durch Poissonsche Verteilungen, *Math. Nachr.*, **27**, 303–340 (1964).

[7] L. Le Cam, On the distribution of sums of independent random variables, In: *Bernoulli, Bayes, Laplace (Anniversary volume)*, Springer-Verlag, Berlin, Heidelberg, New York, 179–202 (1965).

[8] H.H. Panjer, S. Wang, Computational aspects of Sundt's generalized class, *ASTIN Bull.*, **25**, 5–17 (1995).

[9] H.H. Panjer, G.E. Willmot, Models for the distribution of aggregate claims in risk theory, *Transactions of Soc. Actuaries*, **36**, 399–452 (1984).

[10] M.J. Philips, G.V. Weinberg, Non-uniform bounds for geometric approximation, *Statist. Probab. Lett.*, **49**, 305–311 (2000).

[11] J. Šiaulys, V. Čekanavičius, Approximation of distributions of integer-valued additive functions by discrete charges I, *Lithuanian Math. J.*, **28**, 392–401 (1988).

[12] P. Vaitkus, V. Čekanavičius, On a centered Poisson approximation, *Lith. Math. J.*, **38**(4), 391–404 (1998).

Apie gardelinių skirstinių artumą

V. Čekanavičius

Gautas bendras pilnosios variacijos sveikaskaičių skirstinių skirtumo įvertis, galiojantis net ir tuo atveju, kai Frankeno sąlyga nepatenkinta.