Liet. matem. rink., T. 41, spec. nr., 615–619

Kornya approximation for dependent indicators

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

In probability theory, sums of dependent indicators play a very important role. Usually such sums are approximated by the prelimiting Poisson law, see [2], [6]. In this note, we shall explore Kornya type approximation. The Stein method is used.

Let $W = \sum_1^n I_i$ be the sum of (possibly dependent) indicators, s be arbitrary chosen fixed positive integer,

$$P(I_i = 1) = p_i = 1 - P(I_i = 0) = 1 - q_i, \quad W_i = \sum_{j \neq i}^{n} I_j. \tag{1}$$

Let

$$\lambda_j = \frac{(-1)^{j+1}}{j} \sum_{k=1}^{n} \left(\frac{p_k}{q_k}\right)^j, \quad \lambda = \sum_{j=1}^{s} \lambda_j. \tag{2}$$

In [5], Kornya proposed to use signed compound Poisson approximation D having the following characteristic function

$$\exp\left\{\sum_{j=1}^{s} \lambda_j (\mathrm{e}^{itj} - 1)\right\}. \tag{3}$$

Note that some of λ_j are negative, which means that D is *the signed* compound Poisson measure. In [4], [5], D and its generalizations were used for approximating of sums of independent random variables. In the proofs, convolution property of measures was crucial. Consequently, for the case of dependent summands we need a different approach. The Stein method (especially its adaptation for discrete variables by Chen) is a relatively simple tool fulfilling our needs. For a comprehensive discussion of the Stein-Chen method see [2]. Originally the method was developed for the Poisson approximation. For a long time its adaptation for the compound Poisson case was far from satisfactory. The breakthrough was achieved in [3], for further developements see [1]. We shall use some ideas and results of [1], [3].

Further on we assume that

$$p_i < 1/3, \quad i = 1, \ldots, n, \tag{4}$$

and

$$\theta = \lambda^{-1} \sum_{k=1}^{n} \sum_{j=2}^{s} (j-1)(p_k/q_k)^j < 1/2. \tag{5}$$

By V_j we denote a random variable, having the same distribution as W_j when $X_j = 1$, i.e., $P(V_j = k) = P(W_j = k | X_j = 1)$.

Theorem 1. *Under conditions* (4) *and* (5), *given any bounded* $f\colon \mathcal{Z} \to \mathcal{R}$, *the following estimate holds:*

$$\begin{aligned} &\left| \mathbf{E}f(W) - \sum_{k=0}^{\infty} f(k) D\{k\} \right| \\ &\leqslant \frac{2\|f\|}{1-2\theta} \left(\lambda^{-1/2} \sum_{k=1}^{n} p_k \left(\frac{p_k}{q_k} \right)^s + 2\lambda^{-1} \sum_{k=1}^{n} \frac{p_k}{q_k} \mathbf{E}|V_k - W_k| \right). \end{aligned} \tag{6}$$

Here $\|f\|$ *denotes the supremum norm.*

REMARK 1. It is evident, that for the estimate (6) to be small, the distributions of V_k and W_k must be quite similar, which means that the dependence of the indicators can not be very strong.

REMARK 2. Taking supremum over all f satisfying $\|f\| \leqslant 1$ we get the estimate for the total variation norm (for it we shall use the same notation $\|\cdot\|$). Hence, the following corollary.

COROLLARY 1. Under conditions (4) and (5)

$$\begin{aligned} \|\mathcal{L}(W) - D\| &= \sum_{k=1}^{\infty} |P(W = k) - D\{k\}| \\ &\leqslant \frac{2}{1-2\theta} \left(\lambda^{-1/2} \sum_{k=1}^{n} p_k \left(\frac{p_k}{q_k} \right)^s + 2\lambda^{-1} \sum_{k=1}^{n} \frac{p_k}{q_k} \mathbf{E}|V_k - W_k| \right). \end{aligned} \tag{7}$$

Here $\mathcal{L}(W)$ denotes the distribution of W.

REMARK 3. For the sum of independent indicators the accuracy of approximation is of the order $O\left(\sum_{i=1}^{n} p_i^{s+1} \lambda^{-1/2}\right)$.

The Stein-Chen method can be adapted for obtaining estimates in other, different from total variation, distances, see [2]. We shall demonstrate this by considering the Wasserstein distance, which is *stronger* than the total variation distance.

Theorem 2. *Under conditions* (4) *and* (5), *the following estimate holds:*

$$\sup_{f\in\mathcal{F}_W}\left|\mathbf{E}f(W)-\sum_{k=0}^{\infty}f(k)D\{k\}\right|=\sum_{k=1}^{\infty}\left|P(W<k)-D\{[0,k)\}\right|$$
$$\leqslant\frac{3}{1-2\theta}\left(\sum_{k=1}^{n}p_k\left(\frac{p_k}{q_k}\right)^s+2\lambda^{-1/2}\sum_{k=1}^{n}\frac{p_k}{q_k}\mathbf{E}|V_k-W_k|\right). \tag{8}$$

Here $\mathcal{F}_W=\{f:\mathcal{Z}\to\mathcal{R}:\sup_k|f(k+1)-f(k)|\leqslant 1\}$.

REMARK 4. Note that the estimate (8) does not follow from (7) directly, because the set $\mathcal{F}_W$ contains unbounded functions as well as the bounded ones.

Proof. The Stein method depends on the properties of the solution of the so-called Stein equation. In our case, the Stein equation is

$$\sum_{i=1}^{s}i\lambda_i g(j+i)-jg(j)=f(j)-\sum_{k=0}^{\infty}D\{k\},\quad j\geqslant 0. \tag{9}$$

Take any bounded $f:\mathcal{Z}\to\mathcal{R}$. If g solves (9) then

$$\|g\|\leqslant\lambda^{-1/2}2\|f\|/(1-2\theta),\quad \|\Delta g\leqslant\lambda^{-1}2\|f\|/(1-2\theta), \tag{10}$$

see [1]. Here $\Delta g(k)=g(k+1)-g(k)$ and $g(k)=0$, $k\leqslant 0$.

Now it remains to estimate the mean of the left-hand side of (9). We have

$$\mathbf{E}\left\{\sum_{i=1}^{s}i\lambda_i g(W+i)-Wg(W)\right\}$$
$$=\sum_{i=1}^{n}\left\{\frac{p_i}{q_i}\mathbf{E}g(W+1)-p_i\mathbf{E}\big(g(W_i+1)|X_i=1\big)+\sum_{k=2}^{s}(-1)^{k+1}\left(\frac{p_i}{q_i}\right)^k\mathbf{E}g(W+k)\right\}$$
$$=\sum_{i=1}^{n}\left\{\frac{p_i^2}{q_i}\mathbf{E}\big(g(W_i+2)|X_i=1\big)+\frac{p_i}{q_i}q_i\mathbf{E}\big(g(W_i+1)|X_i=0\big)\right.$$
$$-p_i\mathbf{E}\big(g(W_i+1)|X_i=1\big)+\sum_{k=2}^{s}(-1)^{k+1}\left(\frac{p_i}{q_i}\right)^k p_i\mathbf{E}\big(g(W_i+k+1)|X_i=1\big)$$
$$\left.+\sum_{k=2}^{s}(-1)^{k+1}\left(\frac{p_i}{q_i}\right)^k q_i\mathbf{E}\big(g(W_i+k)|X_i=0\big)\right\}$$
$$=\sum_{i=1}^{n}\left\{\mathbf{E}\big(g(W_i+k+1)|X_i=1\big)\left((-1)^{k+1}\left(\frac{p_i}{q_i}\right)^k p_i+(-1)^{k+2}\left(\frac{p_i}{q_i}\right)^{k+1}q_i\right)\right.$$
$$+\mathbf{E}\big(g(W_i+s+1)|X_i=1\big)(-1)^{s+1}p_i\left(\frac{p_i}{q_i}\right)^s$$

$$+\sum_{k=1}^{s}(-1)^k\left(p_iq_i\right)^k q_i\big(\mathbf{E}(g(W_i+k)|X_i=0)-\mathbf{E}\big(\mathbf{E}(g(W_i+k)|X_i=1)\big)\Big\}$$

$$=\sum_{i=1}^{n}\left\{(-1)^{s+1}p_i\left(\frac{p_i{}^s}{q_i}\right)\big(\mathbf{E}(g(W_i+s+1)|X_i=1)\big)\right.$$

$$\left.+\sum_{k=1}^{s}(-1)^{k+1}\left(\frac{p_i}{q_i}\right)^k\big(\mathbf{E}g(W_i+k)-\mathbf{E}g(V_i+k)\big)\right\}. \tag{11}$$

It is easy to check that

$$\big|\mathbf{E}g(W_i+k)-\mathbf{E}g(V_i+k)\big|\leqslant\|\Delta g\|\mathbf{E}|W_i-V_i|, \tag{12}$$

$$\Big\|\mathbf{E}\big(g(W_i+s+1)|X_i=1\big)\Big\|\leqslant\|g\|. \tag{13}$$

Combining (11), (12) and (13) we get

$$\Big\|\mathbf{E}\Big\{\sum_{i=1}^{s}i\lambda_i g(W+i)-Wg(W)\Big\}\Big\|$$

$$\leqslant\|g\|\sum_{i=1}^{n}p_i\left(\frac{p_i}{q_i}\right)^s+\|\Delta g\|\sum_{i=1}^{n}\sum_{k=1}^{s}\left(\frac{p_i}{q_i}\right)^k\mathbf{E}|W_i-V_i|. \tag{14}$$

Taking into account condition (4) we can write

$$\sum_{k=1}^{s}\left(\frac{p_i}{q_i}\right)^k\leqslant\sum_{k=1}^{\infty}\left(\frac{p_i}{q_i}\right)^k\leqslant 2\frac{p_i}{q_i}.$$

Therefore, estimate (6) follows from (14) and (10).

For the proof of estimate (8) we need some analogue of (10). If g solves (9) for $f\in\mathcal{F}_W$ then

$$\|g\|\leqslant 3/(1-2\theta),\quad \|\Delta g\|\leqslant\lambda^{-1/2}3/(1-2\theta). \tag{15}$$

The proof of (15) almost coincides with the proof of Theorem 2.1 from [1]. One must estimate (2.6) replace by its analogue from Lemma 1.1.5, [2]. Now (8) follows from (14) and (15).

References

[1] A.D. Barbour, V. Čekanavičius, Total variation asymptotics for sums of independent integer random variables, *Preprint*, No 07–01, Institute of Mathematics, University of Zurich (2001).
[2] A.D. Barbour, L. Holst, S. Janson, *Poisson Approximations*, Oxford University Press, Oxford (1992).
[3] A.D. Barbour, A. Xia, Poisson perturbations, *ESAIM: Probability and Statistics*, 3, 131–150 (1999).

[4] C. Hipp, Improved approximations for the aggregate claims distribution in the individual model, *ASTIN Bull.*, **16**, 89–100 (1986).

[5] P. Kornya, Distribution of aggregate claims in the Individual Risk Theory model, *Society of Actuaries: Transactions*, **35**, 823–858 (1983).

[6] R.J. Serfling, A general Poisson approximation theorem, *Ann. Probab.*, **3**, 726–731 (1975).

Priklausomų indikatorių Kornya aproksimacija

V. Čekanavičius

Priklausomų indikatorių sumos skirstinys aproksimuojamas ženklą keičiančiu puasoninės struktūros matu, 1983 m. pasiūlytu Kornya darbe. Įvertis gautas pilnosios variacijos ir Vaseršteino metrikoms. Rezultatams įrodyti taikomas modifikuotas Steino metodas.