Math-Net.Ru

Общероссийский математический портал

B. Roos, Poisson approximation via the convolution with Korneya–Presman signed measures, *Теория вероятн. и ее примен.*, 2003, том 48, выпуск 3, 628–632

DOI: <https://doi.org/10.4213/tvp278>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.6.47.48

26 апреля 2023 г., 06:43:15

© 2003 г.

ROOS B.*

POISSON APPROXIMATION VIA THE CONVOLUTION WITH KORNIA-PRESMAN SIGNED MEASURES

Получена верхняя граница для расстояния по вариации между обобщенным полиномиальным распределением и конечной мерой со знаком, которая является сверткой двух конечных мер со знаком, одна из которых — мера Корниа-Пресмана. В одномерном пуассоновском случае такая конечная мера со знаком была впервые рассмотрена В. Пфайфером и К. Боровковым (1996 г.), [1].

Мы даем асимптотические соотношения в одномерном случае и в качестве примера рассматриваем модель рекордов для независимых одинаково распределенных величин.

Оказывается, что в этом случае аппроксимация имеет порядок $O(n^{-s}(\ln n)^{-(s+1)/2})$, где s — фиксированное натуральное число, тогда как при аппроксимации простыми мерами Корниа-Пресмана со знаком мы имеем скорость только $O((\ln n)^{-(s+1)/2})$.

Ключевые слова и фразы: асимптотическое соотношение, обобщенное полиномиальное распределение, модель рекордов для независимых одинаково распределенных случайных величин, мера Корниа-Пресмана со знаком, пуассоновская аппроксимация, расстояние по вариации, верхняя граница.

1. Introduction and main results. Let $X_1, X_2, \dots$ be a sequence of independent Bernoulli random vectors in $\mathbf{R}^k$ ($k \in \mathbf{N} = \{1, 2, \dots\}$) with probabilities

$$\mathbf{P}\{X_j = e_r\} = p_{j,r} \in [0, 1], \quad \mathbf{P}\{X_j = 0\} = 1 - p_j \in [0, 1]$$

for $j \in \mathbf{N}$ and $r \in \{1, \dots, k\}$, where $p_j = \sum_{r=1}^k p_{j,r}$ and e_r denotes the vector in $\mathbf{R}^k$ with entry 1 at position r and 0 otherwise. Let $n \in \mathbf{N}$. We assume that $\lambda_r = \sum_{j=1}^n p_{j,r} > 0$ for all r and that $\sum_{j=1}^\infty p_j^2 < \infty$.

In [7], we used Kerstan's method to prove a result concerning the approximation of the distribution P^{S_n} of the sum $S_n = \sum_{j=1}^n X_j$ by the finite signed measures $\mathcal{P}_s$ ($s \in \mathbf{N}$) of Kornya-Presman type (cf. [3], [4]), which are concentrated on $\mathbf{Z}_+^k$ and have the generating functions

$$\Psi_{\mathcal{P}_s}(z) = \sum_{l \in \mathbf{Z}_+^k} \mathcal{P}_s(\{l\}) z^l = \exp \left(\sum_{m=1}^s G_m(z) \right) \quad (z = (z_1, \dots, z_k) \in \mathbf{C}^k),$$

where $\mathbf{Z}_+ = \{0, 1, 2, \dots\}$, $z^l = z_1^{l_1} \dots z_k^{l_k}$ for $l = (l_1, \dots, l_k) \in \mathbf{Z}_+^k$, and, for $m \in \{1, \dots, s\}$ and $z \in \mathbf{C}^k$,

$$G_m(z) = \frac{(-1)^{m+1}}{m} \sum_{j=1}^n [H_j(z)]^m, \quad H_j(z) = \sum_{r=1}^k p_{j,r}(z_r - 1) \quad (j \in \mathbf{N}).$$

For relevant references concerning the Kornya-Presman expansion, see [7].

In the present paper, we give some results for the approximation of P^{S_n} by the convolution $\mu_s = \mathcal{P}_s * \nu_s$ ($s \in \mathbf{N}$), where ν_s is the finite signed measure with generating function $\Psi_{\nu_s}(z) = \prod_{j=1}^\infty W_s(-H_j(z))$ ($z \in \mathbf{C}^k$), where $W_s(y) = (1-y) \exp(\sum_{m=1}^s y^m/m)$ ($y \in \mathbf{C}$) is the Weierstrass prime function. Since $\sum_{j=1}^\infty p_j^{s+1} < \infty$, $\Psi_{\nu_s}(z)$ is an entire function (see,

* Schwerpunkt Math. Statistik und Stoch. Prozesse, Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, D-20146 Hamburg, Germany; e-mail: roos@math.uni-hamburg.de

e.g., [2, p. 169]). Therefore ν_s and, in its turn, μ_s are finite signed measures. The idea behind the above expansion is that the generating function of S_n

$$\Psi_{S_n}(z) = \sum_{l \in \mathbb{Z}_+^k} \mathbf{P}\{S_n = l\} z^l = \exp \left(\sum_{j=1}^n \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} [H_j(z)]^m \right)$$

can be approximated by

$$\exp \left(\sum_{m=1}^s G_m(z) + \sum_{j=1}^{\infty} \sum_{m=s+1}^{\infty} \frac{(-1)^{m+1}}{m} [H_j(z)]^m \right) = \Psi_{\mu_s}(z), \quad (1)$$

if s or n is large. The main difference between the approximants μ_s and $\mathcal{P}_s$ lies in the fact that, in contrast with $\mathcal{P}_s$, μ_s contains the information about the whole sequence of distributions of the X_j ($j \in \mathbb{N}$). This can be a disadvantage, but it turns out that the approximation by μ_s is much more precise than the one by $\mathcal{P}_s$. Note that, in the one-dimensional Poisson case $k = s = 1$, a result concerning the Kolmogorov metric has been given by Borovkov and Pfeifer [1] (see Section 2 below).

In order to present the first result, we need further notation. Let

$$s \in \mathbb{N}, \quad p_0 = \max_{1 \leq j \leq n} p_j, \quad \hat{p}_0 = \max_{j \geq n+1} p_j, \quad \tilde{p}_0 = \sum_{r=1}^k \max_{1 \leq j \leq n} p_{j,r},$$

$$\kappa_{s,j}(x) = \left[\min \left\{ x \sum_{r=1}^k \frac{p_{j,r}^2}{\lambda_r}, p_j^2 \right\} \right]^{(s+1)/2} \quad (x \in [0, \infty), j \in \mathbb{N}),$$

$$\beta_s(x) = \sum_{j=1}^n \kappa_{s,j}(x), \quad \hat{\beta}_s(x) = \sum_{j=n+1}^{\infty} \kappa_{s,j}(x) \quad (x \in [0, \infty)).$$

Observe that, for all $x \in [0, \infty)$, $\hat{\beta}_s(x) < \infty$. For $y \in \mathbb{C}$, let

$$V_s(y) = [1 - W_s(y)] \frac{s+1}{y^{s+1}}.$$

Some properties of the function V_s can be found in [7, Remark (a)]. For $x \in \mathbb{R}$, let $[x]$ denote the smallest integer $\geq x$. Let $A, B \in [0, 1]$ with $A + B = 1$. In this paper, we frequently deal with power series of the type $Q(z) = \sum_{l \in \mathbb{Z}_+^k} q_l z^l$ ($q_l \in \mathbb{R}$), which are absolutely convergent for all $z \in \mathbb{C}^k$. In particular, the order of summation may be chosen arbitrarily. We write $\|Q(z)\| = \sum_{l \in \mathbb{Z}_+^k} |q_l|$ and use the easy fact that $\|Q_1(z)Q_2(z)\| \leq \|Q_1(z)\| \|Q_2(z)\|$ for two power series $Q_1(z)$ and $Q_2(z)$. Let d_s (resp. d'_s) be the total variation distance between P^{S_n} and $\mathcal{P}_s$ (resp. μ_s). For example, we have $d'_s = 2^{-1} \|\delta_s(z)\|$ with $\delta_s(z) = \Psi_{S_n}(z) - \Psi_{\mu_s}(z)$. In [7], we gave an upper bound for d_s . In the following theorem, we present our main result for d'_s .

Theorem 1. *Let*

$$\begin{aligned} c_1(s) &= \begin{cases} (s+1)2^{-5/2}, & \text{for odd } s, \\ (s+1)2^{1/[2(s+1)]-5/2}, & \text{for even } s, \end{cases} \\ c_2(s, p_0) &= \frac{e 2^s [s/2 - 1]!}{\sqrt{2\pi} (s+1)} V_s(2p_0), \quad c_3(s, p_0) = \frac{e 2^{s+1}}{s+1} V_s(2p_0). \end{aligned}$$

If $c_3(s, \hat{p}_0) \hat{p}_0^{s-1} \hat{\beta}_1(2^{-3/2} A^{-1}) < 1$ and $c_3(1, p_0) \beta_1(2^{-3/2} B^{-1}) < 1$, then

$$d'_s \leq \frac{c_2(s, \hat{p}_0) \hat{\beta}_s(c_1(s) B^{-1})}{[1 - c_3(s, \hat{p}_0) \hat{p}_0^{s-1} \hat{\beta}_1(2^{-3/2} B^{-1})]^{[s/2]} [1 - c_3(1, p_0) \beta_1(2^{-3/2} A^{-1})]}. \quad (2)$$

Remark 1. There exist positive constants C_1 and $C_2(s)$, such that, if n is sufficiently large and $\tilde{p}_0 \leq C_1$, then, for all s , we have $d'_s \leq C_2(s) \hat{\beta}_s(1)$. Note that C_1 is indeed

independent of s . Letting $A = 0.9$, we see that we can set $C_1 = 0.3$, since $\lim_{n \rightarrow \infty} \widehat{\beta}_1(1) = 0$ and an $\varepsilon > 0$ exists, such that, for $\widetilde{p}_0 \leq 0.3$ and A as above,

$$c_3(1, p_0) \beta_1(2^{-3/2} A^{-1}) \leq \frac{e V_1(2 \widetilde{p}_0)}{\sqrt{2} A} \widetilde{p}_0 \leq 1 - \varepsilon.$$

In [7, Remark (c)], we have shown a similar assertion concerning the total variation distance d_s between P^{S_n} and $\mathcal{P}_s$: If $\widetilde{p}_0 \leq \frac{1}{4}$, then $d_s \leq C_3(s) \beta_s(1)$, where $C_3(s)$ is a constant depending only on s . In view of $\lim_{n \rightarrow \infty} \widehat{\beta}_s(1) = 0$, we see that, in the case of n being large, the approximation by μ_s is better than the one by $\mathcal{P}_s$.

We now give a recursive formula for the counting density of μ_s . Observe that $\mu_s(\{0\}) = \Psi_{\mu_s}(0)$. For $l = (l_1, \dots, l_k) \in \mathbb{Z}_+^k$, we use the standard multi-index notation $|l| = l_1 + \dots + l_k$ and $l! = l_1! \dots l_k!$. Further, if additionally $t \in \mathbb{Z}_+^k$, we write $t \leq l$ in the case that $t_r \leq l_r$ for all r . For a set M , $1_M(x)$ is set to be one if $x \in M$ and zero otherwise.

Proposition 1. *Let $l \in \mathbb{Z}_+^k$ with $|l| \geq 1$, $I_l = \{t \in \mathbb{Z}_+^k \mid |t| \geq 1; t \leq l\}$ and, for $t \in I_l$,*

$$a_t = \frac{(-1)^{|t|+1} |t|!}{t!} \left[1_{\{1, \dots, s\}}(|t|) \sum_{j=1}^n \sum_{m=|t|-1}^{s-1} f_{j,m,t} + \sum_{j=1}^\infty \sum_{m=\max\{s, |t|-1\}}^\infty f_{j,m,t} \right],$$

where $f_{j,m,t} = \binom{m}{|t|-1} p_j^{m-|t|+1} \prod_{r=1}^k p_{j,r}^{t_r}$. Then

$$\mu_s(\{l\}) = \frac{1}{|l|} \sum_{t \in I_l} \mu_s(\{l-t\}) a_t.$$

The proof of this proposition will be omitted, since it can easily be done by using (1) and following the lines in the proof of Proposition 1 in [7].

2. The one-dimensional case. In this section, let always $k = 1$. For $l \in \mathbb{N}$, let $\alpha_l = \sum_{j=1}^n p_j^l$ and $\widehat{\alpha}_l = \sum_{j=n+1}^\infty p_j^l$. Instead of λ_1 , we simply write λ . In the case $s = 1$, Borovkov and Pfeifer [1, Theorem 2] derived the following asymptotic relation for the Kolmogorov distance $\rho_1 = \sup_{x \in \mathbb{R}} |\mathbf{P}\{S_n \leq x\} - \mu_1((-\infty, x])|$ between P^{S_n} and μ_1 : If $\sum_{j=1}^\infty p_j = \infty$, then

$$\rho_1 = \frac{\widehat{\alpha}_2}{2\sqrt{2\pi e} \lambda} (1 + o(1)) \quad (n \rightarrow \infty). \tag{3}$$

In the following theorem, we show a similar but somewhat more precise asymptotic relation for the total variation distance in the general case $s \in \mathbb{N}$.

Theorem 2. *Let*

$$\tau_s = \int_{\mathbb{R}} |\varphi_{s+1}(x)| dx, \quad \varphi_{s+1}(x) = \frac{1}{\sqrt{2\pi}} \frac{d^{s+1}}{dx^{s+1}} e^{-x^2/2} \quad (x \in \mathbb{R}).$$

If $\sum_{j=1}^\infty p_j = \infty$, then

$$d'_s = \frac{\tau_s \widehat{\alpha}_{s+1}}{2(s+1)\lambda^{(s+1)/2}} \left(1 + O\left(\frac{1}{\sqrt{\lambda}}\right) \right) \quad (n \rightarrow \infty). \tag{4}$$

Similarly, under the same assumptions, the approximation by $\mathcal{P}_s$ yields the expansion

$$d_s = \frac{\tau_s \alpha_{s+1}}{2(s+1)\lambda^{(s+1)/2}} \left(1 + O\left(\frac{1}{\sqrt{\lambda}}\right) \right) \quad (n \rightarrow \infty). \tag{5}$$

Note that simple calculations yield

$$\begin{aligned} \tau_1 &= \frac{4}{\sqrt{2\pi e}} \approx 0.968, & \tau_2 &= \frac{2(4 + e^{3/2})}{\sqrt{2\pi e^{3/2}}} \approx 1.510, \\ \tau_3 &= \frac{4\sqrt{6}[\sqrt{3 - \sqrt{6}} \exp(\sqrt{6}/2) + \sqrt{3 + \sqrt{6}} \exp(-\sqrt{6}/2)]}{\sqrt{2\pi e^{3/2}}} \approx 2.801. \end{aligned}$$

Example 1. In what follows, we consider the i.i.d. record model (cf. [1]). Let $Z_1, Z_2, \dots$ be independent and identically distributed random variables with a continuous cumulative distribution function. We say that Z_j ($j \in \mathbb{N}$) is a record of this sequence when $Z_j > \max\{Z_1, \dots, Z_{j-1}\}$, where we set $\max \emptyset = -\infty$. The corresponding record indicators X'_j ($j \in \mathbb{N}$) are defined by $X'_j = 1$ when Z_j is a record and $X'_j = 0$ otherwise. Rényi [5] has shown that, here, the X'_j form a sequence of independent Bernoulli random variables with $\mathbf{P}\{X'_j = 1\} = 1/j$ ($j \in \mathbb{N}$). Since $X'_1 = 1$ is a constant, we omit it and set $X_j = X'_{j+1}$ ($j \in \mathbb{N}$). Now $p_j = 1/(j+1)$ ($j \in \mathbb{N}$). As $n \rightarrow \infty$, we have $\hat{\alpha}_{s+1} = s^{-1}n^{-s} + O(n^{-(s+1)})$ and $\lambda = \ln n + \gamma + O(n^{-1})$, where $\gamma = 0.5772\dots$ is Euler's constant. In fact, the first asymptotic relation can easily be deduced from the simple equality

$$\frac{1}{(n+1)^s} = s \sum_{j=n+1}^{\infty} \frac{1}{j(j+1)^s} + \sum_{m=0}^{s-2} \binom{s}{m} \sum_{j=n+1}^{\infty} \frac{1}{j^{s-m}(j+1)^s}.$$

Therefore, by Theorem 2,

$$d'_s = \frac{\tau_s [1 + O((\ln n)^{-1/2})]}{2s(s+1)n^s (\ln n)^{(s+1)/2}}, \quad d_s = \frac{\tau_s (\zeta(s+1) - 1) [1 + O((\ln n)^{-1/2})]}{2(s+1)(\ln n)^{(s+1)/2}},$$

where ζ is the Riemann zeta function. Here, we see that the approximation error d'_s is considerably smaller than d_s . The difference in the order is the factor n^{-s} . It should be mentioned, that an asymptotic relation for the Kolmogorov distance between P^{S_n} and $\mathcal{P}_1$ can be found (without proof) in [1, formula (4)]. A proof can easily be obtained from the more general relation (32) in [6]. Observe that Borovkov and Pfeifer [1, Theorem 1] derived another finite signed measure with the help of the gamma function and gave an explicit bound of order $O(n^{-2})$ for the Kolmogorov distance to P^{S_n} .

3. Proofs. For the proof of Theorem 1, we need the following two lemmas.

Lemma 1 (cf. [7]). *Let $x \in (0, \infty)$ and $j \in \mathbb{N}$. Then*

$$\|V_s(-H_j(z))\| = V_s(2p_j),$$

$$\|[H_j(z)]^{s+1} \exp(xG_1(z))\| \leq \left[\min \left\{ \frac{4c_1(s)}{x} \sum_{r=1}^k \frac{p_{j,r}^2}{\lambda_r}, 4p_j^2 \right\} \right]^{(s+1)/2}$$

Lemma 2. *We have*

$$K_1(B) := \|\Psi_{S_n}(z) \exp(-BG_1(z))\| \leq \frac{1}{1 - c_3(1, p_0) \beta_1(2^{-3/2}A^{-1})}.$$

Proof. Using the expansion

$$\begin{aligned} \Psi_{S_n}(z) \exp(-BG_1(z)) &= \prod_{j=1}^n \left[1 - \frac{[H_j(z)]^2}{2} V_1(-H_j(z)) \right] \exp\left(\frac{A}{j} G_1(z)\right) \\ &= \sum_{j=1}^n \sum_{1 \leq i(1) < \dots < i(j) \leq n} \prod_{a=1}^j \left[\frac{-V_1(-H_{i(a)}(z))}{2} [H_{i(a)}(z)]^2 \exp\left(\frac{A}{j} G_1(z)\right) \right] \\ &\quad + \exp\left(\frac{A}{j} G_1(z)\right) \end{aligned}$$

and the polynomial theorem, we are led to

$$K_1(B) \leq 1 + \sum_{j=1}^{\infty} \frac{1}{j!} \left[\frac{1}{2} \sum_{i=1}^n V_1(2p_i) \left\| [H_i(z)]^2 \exp\left(\frac{A}{j} G_1(z)\right) \right\| \right]^j.$$

Application of Lemma 1 and Stirling's formula yields the assertion.

Proof of Theorem 1. We have

$$\delta_s(z) = \lim_{\eta \rightarrow \infty} \left[1 - \prod_{j=1}^{\eta} \left[1 - \frac{[-H_{j+n}(z)]^{s+1}}{s+1} V_s(-H_{j+n}(z)) \right] \right] \Psi_{S_n}(z)$$

$$= \lim_{\eta \rightarrow \infty} \left[- \sum_{j=1}^{\eta} \sum_{1 \leq i(1) < \dots < i(j) \leq \eta} \prod_{a=1}^j \left[\frac{(-1)^s V_s(-H_{i(a)+n}(z))}{s+1} \right. \right. \\ \left. \left. \times [H_{i(a)+n}(z)]^{s+1} \exp\left(\frac{B}{j} G_1(z)\right) \right] \right] \Psi_{S_n}(z) \exp(-B G_1(z)). \quad (6)$$

By using the polynomial theorem, we obtain $2d_s \leq K_1(B) K_2(B)$, where $K_1(B)$ is defined as in Lemma 2 and

$$K_2(B) := \sum_{j=1}^{\infty} \frac{1}{j!} \left[\sum_{i=n+1}^{\infty} \frac{V_s(2p_i)}{s+1} \left\| [H_i(z)]^{s+1} \exp\left(\frac{B}{j} G_1(z)\right) \right\| \right]^j.$$

Proceeding as in the proof of Theorem 1 in [7], the proof is easily completed.

P r o o f o f T h e o r e m 2. We have

$$\left| 2d'_s - \frac{\hat{\alpha}_{s+1}}{s+1} \|(z-1)^{s+1} \Psi_{S_n}(z)\| \right| \leq Y_1 + Y_2,$$

where, using the properties of V_s (see [7, Remark (a)]),

$$Y_1 = \left\| \sum_{i=n+1}^{\infty} \frac{p_i^{s+2}}{s+1} \frac{[V_s(-p_i(z-1)) - 1]}{p_i(z-1)} (z-1)^{s+2} \Psi_{S_n}(z) \right\| \\ = O\left(\hat{\alpha}_{s+2} \|(z-1)^{s+2} \Psi_{S_n}(z)\|\right),$$

and Y_2 denotes the norm of the sum of terms over $j = 2, 3, \dots$ in the expansion (6), i.e., the summand for $j = 1$ has to be left out. As in the proof of Theorem 1, one can now show that $Y_2 = O(\hat{\alpha}_{s+1}^2/\lambda^{s+1})$. In view of formula (32) in [6], we see that, for $l \in \mathbb{N}$,

$$\|(z-1)^l \Psi_{S_n}(z)\| = \|(z-1)^l \exp(\lambda(z-1))\| + O\left(\frac{\alpha_2}{\lambda^{(l+2)/2}}\right),$$

whereas Proposition 4 in [6] gives

$$\|(z-1)^l \exp(\lambda(z-1))\| = \frac{n-1}{\lambda^{l/2}} + O\left(\frac{1}{\lambda^{(l+1)/2}}\right).$$

Combining the above estimates, we obtain (4). The proof of (5) is similar and therefore omitted.

REFERENCES

1. Borovkov K., Pfeifer D. On improvements of the order of approximation in the Poisson limit theorem. — J. Appl. Probab., 1996, v. 33, № 1, p. 146–155.
2. Conway J. B. Functions of One Complex Variable. Berlin: Springer-Verlag, 1978, 317 p.
3. Kornya P. S. Distribution of aggregate claims in the individual risk theory model. — Trans. Soc. Actuaries, 1983, v. 35, p. 823–858.
4. Пресман Э. Л. О сближении биномиальных и безгранично делимых распределений. — Теория вероятн. и ее примен., 1983, т. 28, в. 2, с. 372–382.
5. Rényi A. Théorie des éléments saillants d'une suite d'observations. — Colloquium on Combinatorial Methods in Probability Theory. Aarhus: Matematisk Institut, Aarhus Universitet, 1962, p. 104–115.
6. Roos B. Asymptotics and sharp bounds in the Poisson approximation to the Poisson-binomial distribution. — Bernoulli, 1999, v. 5, № 6, p. 1021–1034.
7. Roos B. Kerstan's method in the multivariate Poisson approximation: an expansion in the exponent. — Теория вероятн. и ее примен., 2002, т. 47, в. 2, с. 397–402.

Поступила в редакцию
18.II.2003