

А. П. К о в а л е в с к и й, Н. С. З а к р е в с к а я (Новосибирск, НГТУ, НГУ). Функциональная центральная предельная теорема для процессов количеств непустых урн в бесконечных урновых моделях.

УДК 519.233

DOI https://doi.org/10.52513/08698325_2022_29_1.1

Резюме: Мы изучаем совместное распределение двух процессов количеств непустых урн в бесконечной урновой схеме. Первый процесс подсчитывает количество непустых урн в прямом направлении времени (от первого шара к последнему), а второй в обратном. Мы доказываем функциональную центральную предельную теорему для разности этих процессов.

Ключевые слова: бесконечная урновая модель, слабая сходимости, гауссовский процесс.

Бесконечная урновая модель состоит в следующем: случайные величины $X_1, \dots, X_n$ независимы и одинаково распределены в соответствии с некоторым распределением на положительных целых числах $\mathbf{P}(X_1 = i) = p_i > 0, i = 1, 2, \dots$.

В этих обозначениях X_j — номер урны при j -м бросании шара. Мы будем предполагать, что вероятности упорядочены по неубыванию: $p_1 \geq p_2 \geq \dots$.

Изучим следующие последовательности случайных величин:

$$R_k = 1 + \sum_{i=2}^k \mathbf{I}\{X_i \neq X_j, \quad j = 1, \dots, i-1\},$$
$$R'_k = 1 + \sum_{i=n-k+1}^{n-1} \mathbf{I}\{X_i \neq X_j, \quad j = i+1, \dots, n\}, \quad k \geq 2, \quad R_0 = R'_0 = 0, \quad R_1 = R'_1 = 1.$$

Это последовательности количеств непустых урн при k первых (R_k) и последних (R'_k) бросаниях шара. Отметим, что $R_n = R'_n$ — общее число непустых урн при бросании n шаров. По-видимому, впервые статистика R_n изучалась в [1]. В работах [2–13] доказаны различные варианты центральной предельной теоремы для статистики R_n , а также для последовательности статистик $R_0, \dots, R_n$ (функциональный вариант). Нашим новым результатом является рассмотрение последовательности статистик в обратном направлении времени $R'_0, \dots, R'_n$ и доказательство (совместно с М. Г. Чебуниным) предельной теоремы для совместного распределения этих последовательностей статистик.

Теорема 1. Пусть функция $\alpha(x) = \max\{k : p_k \geq 1/x\}$ является правильно меняющейся с показателем $\theta, 0 < \theta < 1$, т. е.

$$\alpha(x) = x^\theta L(x),$$

где $L(x)$ — медленно меняющаяся на бесконечности функция. Тогда случайный процесс

$$\left\{ Q_n(t) = \frac{R_{[nt]} - R'_{[nt]}}{\sqrt{R_n}}, \quad 0 \leq t \leq 1 \right\}$$

сходится по распределению в равномерной метрике в $D[0, 1]$ к центрированному гауссовскому случайному процессу $Q = \{Q(t), 0 \leq t \leq 1\}$ с непрерывными п.н. траекториями и ковариационной функцией

$$K_Q(s, t) = \mathbf{E}Q(s)Q(t) = 2(s+t)^\theta \mathbf{1}\{s+t \leq 1\} + 2\mathbf{1}\{s+t > 1\} - 2\max(s^\theta, t^\theta).$$

Допределенный процесс определяется полностью по эмпирическим данным, а распределение предельного процесса зависит только от параметра θ . Мы изучаем асимптотику ряда функционалов от допределенного процесса и их применение к анализу данных.

Исследование поддержано Математическим центром в Академгородке в соответствии с соглашением № 075-15-2019-1675 с Минобрнауки России.

СПИСОК ЛИТЕРАТУРЫ

1. *Bahadur R. R.* On the number of distinct values in a large sample from an infinite discrete distribution. — Proceedings of the National Institute of Sciences of India, 1960, v. 26A, Supp. II, c. 67–75.
2. *Barbour A. D., Gnedin A. V.* Small counts in the infinite occupancy scheme. — Electronic Journal of Probability, 2009, v. 14, Paper no. 13, c. 365–384.
3. *Chakrabarty A., Chebunin M., Kovalevskii A., Pupyshv I., Zakrevskaya N., Zhou Q.* A statistical test for correspondence of texts to the Zipf - Mandelbrot law. — Siberian Electronic Mathematical Reports, 2020, v. 17, c. 1959–1974.
4. *Chebunin M. G.* Functional central limit theorem in an infinite urn scheme for distributions with superheavy tails. — Siberian Electronic Mathematical Reports, 2017, v. 14, c. 1289–1298.
5. *Chebunin M., Kovalevskii A.* Functional central limit theorems for certain statistics in an infinite urn scheme. — Statistics and Probability Letters, 2016, v. 119, c. 344–348.
6. *Chebunin M., Kovalevskii A.* A statistical test for the Zipf's law by deviations from the Heaps' law. — Siberian Electronic Mathematical Reports, 2019, v. 16, c. 1822–1832. DOI 10.33048/semi.2019.16.129
7. *Chebunin M., Zuyev S.* Functional Central Limit Theorems for Occupancies and Missing Mass Process in Infinite Urn Models. — J. Theor. Probab., 2022, v. 35, c. 1–19.
8. *Durieu O., Wang Y.* From infinite urn schemes to decompositions of self-similar Gaussian processes. — Electron. J. Probab., 2016, v. 21, Paper no. 43. MR3530320
9. *Durieu O., Samorodnitsky G., Wang Y.* From infinite urn schemes to self-similar stable processes. — Stochastic Processes and their Applications, 2020, v. 130, № 4, c. 2471–2487.
10. *Dutko M.* Central limit theorems for infinite urn models. — Ann. Probab., 1989, v. 17, c. 1255–1263.
11. *Gnedin A., Hansen B., Pitman J.* Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. — Probability Surveys, 2007, v. 4, c. 146–171.
12. *Karlin S.* Central Limit Theorems for Certain Infinite Urn Schemes. — Journal of Mathematics and Mechanics, 1967, v. 17, No. 4, c. 373–401.
13. *Zakrevskaya N., Kovalevskii A.* An omega-square statistics for analysis of correspondence of small texts to the Zipf-Mandelbrot law. — Applied methods of statistical analysis. Statistical computation and simulation — AMSA'2019, 18–20 September 2019, Novosibirsk: Proceedings of the International Workshop, Novosibirsk: NSTU, 2019, c. 488–494.

UDC 519.233

DOI https://doi.org/10.52513/08698325_2022_29_1_1

Kovalevskii A. P., Zakrevskaya N. S. (Novosibirsk, Novosibirsk State Technical University, Novosibirsk State University). **Functional Central Limit Theorem for processes of numbers of nonempty urns in an infinite urn model.**

Abstract: We study the joint distribution of two processes of numbers of non-empty urns in an infinite urn model. The first process counts numbers of non-empty urns in forward time (from the first ball to the last), the second does it in backward time. We prove the Functional Central Limit Theorem for the difference of these processes.

Keywords: infinite urn model, weak convergence, Gaussian process.