ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
**Том 24** **МАТЕМАТИКИ** **Выпуск 3**
**2017**

**М. Г. Чебунин** (Новосибирск, НГУ). **О точности пуассонизации в бесконечной урновой схеме.**

Изучается следующая вероятностная модель [1], [2]: имеется $n$ шаров и бесконечное число ячеек («урн»), занумерованных числами $1, 2, \ldots$. Каждый шар попадает в $i$-ю ячейку с вероятностью $p_i > 0$ независимо от всех других шаров ( $\sum_{i=1}^{\infty} p_i = 1$ ). Обозначим через $X_i$ номер ячейки, в которую попал $i$-й шар, а через $J_i(n)$ количество шаров, попавших в $i$-ю ячейку. Нас интересуют асимптотические свойства распределения числа различных элементов в выборке ( $X_1, \ldots, X_n$ ) объема $n$ при стремлении $n \to \infty$. А именно, мы рассматриваем статистики числа элементов, встретившихся не менее $k \geqslant 1$ раз $R^*_{n,k} = \sum_{i=1}^{\infty} \mathbf{I}(J_i(n) \geqslant k)$, и статистики числа элементов, встретившихся ровно $k$ раз; $R_{n,k} = R^*_{n,k} - R^*_{n,k+1}$. Наряду с выборкой объема $n$ будем рассматривать выборку случайного объема $P(t)$, где $\{P(t),\ t \geqslant 0\}$ — пуассоновский процесс с интенсивностью 1 и не зависящий от последовательности $X_1, X_2, \ldots$. Обозначим

$$\alpha(x) = \max\{j \ :\ \ p_j \geqslant 1/x\} = x^{\theta} L(x) \text{ при } \theta \in [0, 1] \tag{1}$$

и будем предполагать, что функция $\alpha(x)$ правильно меняется на бесконечности, где $L(x)$ — медленно меняющаяся функция.

В предположении, что $\theta > 0$, Карлин [2] доказал УЗБЧ и ЦПТ для $R_n$. Чебунин и Ковалевский [3] доказали функциональную ЦПТ для $Z_n(\cdot) = (R_{[n\cdot]} - \mathbf{E}R_{[n\cdot]})/\sqrt{\mathbf{E}R_n}$. Предельный процесс $Z_\theta$ — центрированный гауссовский с непрерывными п.н. траекториями и ковариационной функцией $K(s,t) = (s+t)^{\theta} - \max(s^{\theta}, t^{\theta})$. Отметим, что этот же предельный процесс возникает в [4] при изучении последовательности сумм индикаторов непустых урн со случайными знаками. Закревская и Ковалевский [5] получили состоятельную оценку параметра $\theta$ в виде неявной функции от $R_n$ при более жестком по сравнению с (1) условии $p_i = C i^{-1/\theta}$, $i \geqslant 1$. В этой ситуации константа $C$ является функцией от $\theta$. Чебунин [6] доказал состоятельность оценки $\ln R_n / \ln n$ в ситуации более общей, чем (1). Мы будем изучать асимптотическое поведение $R_n - R_{P(n)}$ и $R_n - \mathbf{E}R_n$ со средней нормировкой между законом больших чисел и центральной предельной теоремой. Обозначим при $k \geqslant 1$

$$Y^*_{n,k} = R^*_{n,k} - \mathbb{E}R^*_{n,k};\ B^*_{n,k} = \mathbb{D}R^*_{P(n),k};\ Y_{n,k} = R_{n,k} - \mathbb{E}R_{n,k};\ B_{n,k} = \mathbb{D}R_{P(n),k};$$

и определим при $\theta = 1$ медленно меняющуюся функцию ([1], лемма 4)

$$L^*(t) \stackrel{def}{=} \int_0^{\infty} \frac{e^{-1/y}}{y} L(ty)\,dy \to 0 \ \text{ при } \ t \to \infty.$$

**Теорема 1.** *В условиях* (1), *для любого* $k \geqslant 1$ *и* $\theta \in [0, 1]$,

$$b_n(R^*_{n,k} - R^*_{P(n),k}) \xrightarrow[n\to\infty]{п.н.} 0 \quad и \quad b_n(R_{n,k} - R_{P(n),k}) \xrightarrow[n\to\infty]{п.н.} 0,$$

*где*

$$b_n = \begin{cases} (nL^*(n)\ln\ln n)^{-\frac{1}{2}}, & \theta = 1, \quad k = 1; \\ (nL(n)\ln\ln n)^{-\frac{1}{2}}, & \theta = 1, \quad k \geqslant 2; \\ o(\min\{n^{\frac{1}{2}-\theta}(\ln\ln nL(n))^{-1}, (\ln n)^{-1}\}), & \theta < 1, \quad k \geqslant 1. \end{cases}$$

**Теорема 2.** *В условиях* (1), *если* $\frac{\mathbb{E}R_{P(n),k_0}}{\ln n} \to \infty$ *некоторого* $k_0 \geqslant 1$ *при* $n \to \infty$, *то для любого* $k \leqslant k_0$

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{|Y^*_{n,k}|}{\sqrt{2B^*_{n,k}\ln n}} \leqslant 1\right) = 1, \quad \mathbb{P}\left(\limsup_{n\to\infty} \frac{|Y_{n,k}|}{\sqrt{2B_{n,k}\ln n}} \leqslant 1\right) = 1.$$

Автор благодарит А. П. Ковалевского за многочисленные полезные обсуждения. Исследование поддержано грантом РФФИ 17-01-00683.

## 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, 26A, Supp. II, c. 67–75.
2. *Karlin S.* Central limit theorems for certain infinite urn schemes. — J. Mathematics and Mechanics, 1967, v. 17, №4, c. 373–401.
3. *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.
4. *Durieu O.*, *Wang Y.* From infinite urn schemes to decompositions of self-similar Gaussian processes. — Electron. J. Probab., 2016, v. 21, № 43, 23 c.
5. *Закревская Н. С.*, *Ковалевский А. П.* Однопараметрические вероятностные модели статистик текста. — Сиб. ж. индустр. матем., 2001, т. 4, в. 2, с. 142–153.
6. *Чебунин М. Г.* Оценивание параметров вероятностных моделей по числу различных элементов выборки. — Сиб. ж. индустр. матем., 2014, т. 17, в. 3, с. 135—147.