Оценка постоянной в теореме о сходимости суммы целочисленных разнораспределённых случайных величин к поправке Чебышёва-Эджворта

Мельников Алексей Денисович

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: alekseym822@gmail.com

Научный руководитель — *Ульянов Владимир Васильевич*

Рассмотрим целочисленную случайную величину ξ с характеристической функцией $v(t) = \mathbb{E}e^{it\xi}$. Введём следующую величину

$$V(\xi) := -\sup_{0<t<2\pi} \frac{|v(t)|}{1-\cos t} = -\sup_{0<t<2\pi} \frac{|\mathbb{E}e^{it\xi}|}{1-\cos t}. \tag{1}$$

Далее, нам понадобятся n независимых случайных величин $X_1, ..., X_n$ с конечным четвёртым моментом. Через $S_n := X_1 + ... + X_n$ обозначим их сумму. Введём для них поправку Чебышёва-Эджворта.

Определение 1. *Поправкой Чебышёва-Эджворта называется функция вида*

$$\Phi_3(x) = \Phi(x) - \frac{\gamma_3}{6}(x^2-1)\varphi(x), \tag{2}$$

где $\Phi(x)$ — функция распределения стандартного нормального закона, $\varphi(x)$ — его плотность, $\gamma_3 = \frac{1}{\sigma^3}\mathbb{E}(S_n - \mathbb{E}S_n)^3 = \frac{1}{\sigma^3}\sum_{k=1}^n \mathbb{E}(X_k - \mathbb{E}X_k)^3$.

Ранее был известен следующий результат [1], где фигурировала неизвестная постоянная c.

Теорема 1. *Пусть целочисленные случайные величины $X_1, ..., X_n$ независимы и имеют конечный четвёртый момент. Для суммы $S_n = X_1 + ... + X_n$ положим $\mu = \mathbb{E}S_n, \sigma^2 = \mathbb{D}S_n$ и $V = \sum_{k=1}^n V(X_k)$. Тогда существует абсолютная постоянная $c > 0$, для которой справедливо соотношение*

$$\sup_{k\in\mathbb{Z}} \left| \mathbf{P}\{S_n \leqslant k\} - \Phi_3\left(\frac{k+\frac{1}{2}-\mu}{\sigma}\right)\right| \leqslant \frac{c\sigma^2}{V} L_4, \tag{3}$$

где $L_4 = \frac{1}{\sigma^4}\sum\limits_{k=1}^n \mathbb{E}|X_k - \mathbb{E}X_k|^4$.

Результатом работы стала оценка постоянной c, что позволяет применять данную теорему в практических задачах.

Теорема 2. *Условия теоремы те же, что и выше. Тогда справедливо*

$$\sup_{k\in\mathbb{Z}}\left|\mathbf{P}\{S_n \leqslant k\} - \Phi_3\left(\frac{k+\frac{1}{2}-\mu}{\sigma}\right)\right| \leqslant 34.41\cdot\frac{\sigma^6}{V^3}L_4. \tag{4}$$

С помощью полученного результата можно проводить анализ биномиальных схем разнораспределённых случайных величин, часто встречающихся на практике. Ниже приведена таблица для синтетических биномиальных схем с разнораспределёнными бернуллевскими величинами.

Кол-во сл.в.	Теоретическая погрешность аппроксимации
10	7.265
100	0.913
1000	0.079
10000	0.008

Таблица 1: Теоретическая погрешность для биномиальной схемы

Из неё видно, что при выборке в 10000 значений, теоретическая погрешность составляет не более 1% в абсолютном выражении. Выборки такого размера очень часто встречаются в анализе больших данных, что делает данный результат хорошо применимым в реальной практике.

Литература

1. Бобков С. Г., Ульянов В. В. Поправка Чебышёва–Эджворта в центральной предельной теореме для целочисленных независимых слагаемых // Теория вероятн. и ее примен., 66:4 (2021), 676–692; Theory Probab. Appl., 66:4 (2022), 537–549.
2. Bobkov S. G. Asymptotic Expansions for Products of Characteristic Functions Under Moment Assumptions of Non-integer Orders // In: Carlen, E., Madiman, M., Werner, E. (eds) Convexity and Concentration. The IMA Volumes in Mathematics and its Applications, vol 161. Springer, New York, NY.