ДОКЛАДЫ АКАДЕМИИ НАУК, 2011, том 436, № 1, с. 26–28

МАТЕМАТИКА

УДК 519.21

ОБ ОЦЕНКЕ БЛИЗОСТИ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ К НОРМАЛЬНОМУ

 С. В. Нагаев, В. И. Чеботарев

Представлено академиком Ю.В. Прохоровым 24.05.2010 г.

Поступило 03.08.2010 г.

Пусть $X, X_1, \ldots, X_n$ – последовательность независимых одинаково распределенных случайных величин с нулевым средним и конечным третьим моментом. Обозначим

$$b^2 = \mathrm{E}X^2, \quad \beta_3 = \mathrm{E}|X|^3, \quad S_n = \sum_{i=1}^n X_i.$$

Пусть $\Phi(x)$ – стандартный нормальный закон.

А. Берри [1] и К.-Г. Эссеен [2] доказали, что

$$\Delta_n := \sup_x \left|\mathrm{P}(n^{-1/2}S_n < bx) - \Phi(x)\right| < C_0\frac{\beta_3}{b^3\sqrt{n}},$$

где C_0 – абсолютная постоянная. Поиску оптимального значения C_0 посвящено большое количество работ (см., например, [3–8]). Эссеен [9] показал, что C_0 не может быть меньше

$$C_E = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0.409732\ldots \tag{1}$$

Что касается верхних оценок для C_0, то наилучший результат в этом направлении получен в недавней работе В.Ю. Королева и И.Г. Шевцовой [8], а именно,

$$C_0 \le 0.5129. \tag{2}$$

На самом деле, в [8] получен более сильный результат

$$\Delta_n \le \frac{0.34445\beta_3 + 0.16844b^3}{b^3\sqrt{n}}, \tag{3}$$

из которого очевидным образом следует (2).

В настоящей работе мы дадим оценку для C_0 в частном случае, когда X принимает только два значения.

Итак, пусть $\mathrm{P}(X = a) = q$, $\mathrm{P}(X = d) = p$, где $p + q = 1$, $a < 0 < d$, $\mathrm{E}X = 0$. Будем считать для простоты, что $b^2 = 1$. Тогда

$$\beta_3 = \frac{p^2+q^2}{\sqrt{pq}}, \quad \mathrm{E}X^3 = \frac{q-p}{\sqrt{pq}}, \quad d - a = \frac{1}{\sqrt{pq}}. \tag{4}$$

Определим функцию $\mathscr{E}(p)$ равенством

$$\mathscr{E}(p) = \frac{1}{\beta_3\sqrt{2\pi}}\left(\frac{\mathrm{E}X^3}{6} + \frac{d-a}{2}\right).$$

Нетрудно показать, используя формулы (4), что

$$\mathscr{E}(p) = \frac{2-p}{3\sqrt{2\pi}[p^2 + (1-p)^2]}. \tag{5}$$

Обозначим

$$\sigma = \sigma(p, n) = \sqrt{npq},$$

$$\omega_3(p) = q - p, \quad \omega_4(p) = |q^3 + p^3 - 3pq|,$$

$$\omega_5(p) = q^4 - p^4, \quad \omega_6(p) = q^5 + p^5 + 15(pq)^2.$$

Положим

$$K_1(p, n) = \frac{\omega_3(p)}{4\sigma\sqrt{2\pi}(n-1)}\left(1 + \frac{1}{4(n-1)}\right) + \frac{\omega_4(p)}{12\sigma^2\pi}\left(\frac{n}{n-1}\right)^2 + \frac{\omega_5(p)}{40\sigma^3\sqrt{2\pi}}\left(\frac{n}{n-1}\right)^{5/2} + \frac{\omega_6(p)}{90\sigma^4\pi}\left(\frac{n}{n-1}\right)^3.$$

Далее, обозначим

$$e_5 = 0.0277905, \quad \tilde\omega_5(p) = p^4 + q^4 + 5!e_5(pq)^{3/2},$$

$$V_6(p) = \omega_3^2(p), \quad V_7(p) = \omega_3(p)\omega_4(p),$$

$$V_8(p) = \frac{2\tilde\omega_5(p)\omega_3(p)}{5!3!} + \frac{\omega_4^2(p)}{(4!)^2},$$

$$V_9(p) = \tilde\omega_5(p)\omega_4(p), \quad V_{10}(p) = \frac{2^6\cdot 3}{(5!)^2}\tilde\omega_5^2(p),$$

$$\zeta(p) = \left(\frac{p^2+q^2}{6}\right)^{2/3}, \quad e(n, p) = \exp\left\{\frac{1}{24\sigma^{2/3}\zeta^2(p)}\right\},$$

$$A_k(n) = \left(\frac{n}{n-2}\right)^{k/2}\frac{n-1}{n},$$

Институт математики им. С.Л. Соболева Сибирского отделения Российской Академии наук, Новосибирск

Вычислительный центр Дальневосточного отделения Российской Академии наук, Хабаровск

$$\gamma_6 = \frac{1}{9}, \quad \gamma_7 = \frac{5\sqrt{2\pi}}{96}, \quad \gamma_8 = 24,$$

$$\gamma_9 = \frac{7\sqrt{2\pi}}{4!16}, \quad \gamma_{10} = \frac{2^6 \cdot 3}{(5!)^2},$$

$$\tilde{\gamma}_6 = \frac{2}{3}, \quad \tilde{\gamma}_7 = \frac{7}{8}, \quad \tilde{\gamma}_8 = \frac{10}{9}, \quad \tilde{\gamma}_9 = \frac{11}{8}, \quad \tilde{\gamma}_{10} = \frac{5}{3}.$$

Пусть

$$K_2(p, n) = \frac{1}{\pi\sigma}\sum_{j=1}^{5}\frac{\gamma_{j+5}A_{j+5}(n)V_{j+5}(p)}{\sigma^j} \times$$
$$\times\left[1 + \frac{\tilde{\gamma}_{j+5}e(n, p)n}{\sigma^2(n-2)}\right].$$

Наконец, положим

$$K_3(p, n) = \frac{1}{\pi}\left\{\frac{1}{12\sigma^2} + \left(\frac{1}{36} + \frac{\mu}{8}\right)\frac{1}{\sigma^4} + \left(\frac{e^{A_1/6}}{36} + \frac{\mu}{8}\right)\frac{1}{\sigma^6} +\right.$$
$$+ \frac{5\mu e^{A_2/6}}{24\sigma^8} + \frac{1}{3}e^{-\sigma\sqrt{A_1} + A_1/6} + (\pi - 2)\mu e^{-\sigma\sqrt{A_2} + A_2/6} +$$
$$+ \frac{1}{4}e^{-\sigma\sqrt{A_3} + A_3/6}\ln\left(\frac{\pi^4\sigma^2}{4A_3}\right) + \exp\left\{-\frac{\sigma^{2/3}}{2\zeta(p)}\right\}\times$$
$$\left.\times\left[\frac{2\zeta(p)}{\sigma^{2/3}} + e^{A_3/6}\frac{1 + \chi(p, n)}{24\zeta(p)\sigma^{4/3}}\right]\right\},$$

где

$$A_1 = 5.40466, \quad A_2 = 7.52058, \quad A_3 = 5.2335,$$

$$\mu = \frac{3\pi^2 - 16}{\pi^4},$$

$$\chi(p, n) = \begin{cases} \dfrac{2\zeta(p)}{\sigma^{2/3}} \ \text{при} \ 0 < p < 0.085, \\ 0 \ \text{при} \ 0.085 \le p \le 0.5. \end{cases}$$

Обозначим через $\Delta_n(p)$ значение Δ_n при заданном p.

Теорема 1. *Для любого* $0.02 \le p \le 0.5$

$$\Delta_n(p) \le \frac{\beta_3(p)}{\sqrt{n}}\mathscr{E}(p) + R(p, n),$$

где

$$\beta_3(p) = \frac{p^2 + q^2}{\sqrt{pq}}, \quad R(p, n) = \sum_{j=1}^{3}K_j(p, n),$$

причем для каждого фиксированного $0.2 < p \le 0.5$ *последовательность* $R_0(p, n) \equiv \dfrac{\sqrt{n}}{\beta_3(p)}R(p, n)$, *убывая по* $n \ge 200$, *сходится к* 0.

Теорема 1 уточняет применительно к биномиальному распределению асимптотический результат Эссеена [9].

Приведенные выше формулы для $K_i(p, n)$, $i = 1$, 2, 3, через которые выражается $R_0(p, n)$, очень сложны. Конечно, можно оценить $K_i(p, n)$ сверху более простыми выражениями, но при этом мы много потеряем в точности.

При доказательстве теоремы 1 используется, как и в почти всех работах, посвященных оценке постоянной в неравенстве Берри–Эссеена, метод сглаживания. Однако в отличие от традиционного после работы С. Цаля [4] сглаживания посредством знакопеременных мер, мы применяем с этой целью равномерное распределение, сосредоточенное на отрезке $\left(-\dfrac{1}{2\sqrt{pq}}, \dfrac{1}{2\sqrt{pq}}\right)$.

Функция $\mathscr{E}(p)$ достигает максимума в точке $p = p_0 = \dfrac{4 - \sqrt{10}}{2} = 0.418861...$ Действительно, из (5) следует, что $\mathscr{E}'(p) = \dfrac{1}{\sqrt{2\pi}}\dfrac{2p^2 - 8p + 3}{3(p^2 + q^2)^2}$. Уравнение $2p^2 - 8p + 3$ имеет единственный корень $p_0 = \dfrac{4 - \sqrt{10}}{2}$, не превосходящий 1. Функция $\mathscr{E}(p)$ возрастает при $0 < p < p_0$ и убывает при $p_0 < p \le \dfrac{1}{2}$, т.е. p_0 является точкой максимума $\mathscr{E}(p)$. Нетрудно проверить, что $\mathscr{E}(p_0) = \dfrac{\sqrt{10} + 3}{6\sqrt{2\pi}}$. Таким образом, $\mathscr{E}(p_0)$ совпадает с постоянной C_E, определенной равенством (1).

Вычисления показывают, что

$$\sup_{0.02 \le p < 0.5} R_0(p, 200) < 0.011768.$$

Поэтому справедливо

Следствие. *При* $n \ge 200$ *и* $0.02 \le p \le 0.5$ *имеет место неравенство*

$$\sup_{0.02 \le p < 0.5} \sqrt{n}\beta_3^{-1}(p)\Delta_n(p) < 0.4215. \tag{6}$$

С другой стороны, прямые вычисления приводят к оценке

$$\max_{1 \le n \le 200}\sup_{0.02 \le p \le 0.5} \sqrt{n}\beta_3^{-1}(p)\Delta_n(p) < 0.4096. \tag{7}$$

Пусть теперь $0 < p < 0.02$. В этом случае из оценки (3) следует, что

$$\frac{\sqrt{n}}{\beta_3(p)}\Delta_n < 0.369. \tag{8}$$

Действительно, $\beta_3(p)$ убывает с ростом p. Поэтому для $p < 0.02$

$$\Delta_n(p) < \frac{\beta_3(p)}{\sqrt{n}}(0.34445 + 0.16844\beta_3^{-1}(0.02)) <$$

$$< 0.369\frac{\beta_3(p)}{\sqrt{n}}.$$

Собирая оценки (6)–(8), мы получаем следующее утверждение.

Теорема 2. *Для любых* $0 < p \le 0.5$ *и* $n \ge 1$

$$\Delta_n(p) \le 0.4215\frac{\beta_3(p)}{\sqrt{n}}. \quad (9)$$

Мы видим, что постоянная в правой части неравенства (9) отличается от C_E приблизительно на 0.0118. Это дает основание предполагать с большой долей уверенности, что наименьшая постоянная в неравенстве Берри–Эссеена для биномиального распределения на самом деле равна C_E. В этой связи уместно упомянуть недавнюю работу [11], где получена неулучшаемая оценка

$$\sup_n \sqrt{n}\Delta_n(0.5) \le \frac{1}{\sqrt{2\pi}}.$$

Кстати, в этой работе метод сглаживания не используется.

Заметим, что в [12] оценка для случая $0 < p \le$ ≤ 0.02 получена другим способом, а именно, с помощью оценок пуассоновского приближения как к биномиальному, так и нормальному распределениям.

Авторы благодарят А.С. Кондрика и К.В. Михайлова за помощь в компьютерных вычислениях.

Работа выполнена при частичной финансовой поддержке грантов: РФФИ 09–01–98524-р_восток_а, СО РАН № 30, ДВО РАН 09–II–СО–01–003, ДВО РАН 09–I–ОМН–02.

СПИСОК ЛИТЕРАТУРЫ

1. *Berry A.C.* // Trans. Amer. Math. Soc. 1941. V. 49. P. 122–126.
2. *Esseen C.-G.* // Ark. Mat. Astr. Fys. 1942. V. 28A. P. 1–19.
3. *Золотарев В.М.* // Теория вероятностей и ее применения. 1966. Т. 11. В. 1. С. 108–119.
4. *Zahl S.* // SIAM J. Appl. Math. 1966. V. 14. № 6. P. 1225–1245.
5. *Van Beek P.* // Z. Wahrsch. verw. Geb. 1972. Bd. 23. P. 187–196.
6. *Шиганов И.С.* Проблемы устойчивости стохастических моделей. М.: ВНИИСИ, 1982. С. 109–115.
7. *Prawitz H.* // Scand. Aktuarietidskr. J. 1975. V. 3/4. P. 145–156.
8. *Королев В.Ю., Шевцова И.Г.* // Теория вероятностей и ее применения. 2009. Т. 53. В. 4. С. 671–695.
9. *Esseen C.-G.* // Scand. Aktuarietidskr. J. 1956. V. 3/4. P. 160–170.
10. *Esseen C.-G.* // Acta Math. 1945. V. 77. P. 1–125.
11. *Hipp C., Mattner L.* // Теория вероятностей и ее применения. 2007. Т. 52. В. 3. С. 610–617.
12. *Нагаев С.В., Чеботарев В.И.* Об оценке близости биномиального распределения к нормальному. Препр. 2009/142. Хабаровск: Вычисл. центр ДВО РАН, 2009. 47 с.