УДК 519.2

НЕЗАВИСИМЫЕ ПОВТОРНЫЕ ИСПЫТАНИЯ КАК АСИМПТОТИЧЕСКИ ГАУССОВСКИЙ СЛУЧАЙНЫЙ ПРОЦЕСС

д-р техн. наук С.Г. ЕХИЛЕВСКИЙ, канд. физ.-мат. наук, доц. О.В. ГОЛУБЕВА, Е.П. ПОТАПЕНКО, Т.С. РУДЬКОВА
(Полоцкий государственный университет)

Независимые повторные испытания предложено рассматривать как случайный процесс с дискретным временем. Методом моментов установлена асимптотика процесса и определен вклад асимметрий и эксцессов произвольных порядков в отклонение распределения Бернулли от асимптотической формулы, фигурирующей в локальной теореме Лапласа. Показано, что последнюю нельзя применять для вычисления вероятностей чисел успехов далеких от своих наиболее вероятных значений.

Ключевые слова: *схема Бернулли, случайный процесс, метод моментов.*

Вероятность m успехов в n независимых испытаниях определяется формулой Бернулли [1]

$$P(n,m) = C_n^m p^m (1-p)^{n-m}, \tag{1}$$

в которой C_n^m – числа сочетаний, p – вероятность успеха в одном опыте. Для больших n вычисление C_n^m становится проблемой. В частности, MathCAD не может работать с числами, превышающими 10^{307}. Для того чтобы обойти эту трудность, используют предельные теоремы Лапласа, которые, однако, не всегда обеспечивают требуемую точность и вовсе неприменимы для m далеких от своих наиболее вероятных значений.

Вместе с тем, если n растет, схему Бернулли можно рассматривать как случайный процесс с дискретным временем. При этом эволюцию закона распределения m удобно исследовать методом моментов, реализованном в [2–5] для описания динамики сорбции и диффузии. Таким образом можно не только выяснить асимптотику процесса, но и определить поправки к ней, обусловленные асимметриями и эксцессами закона распределения. Решению этих задач посвящена данная публикация.

Как известно, вся информация о законе распределения m содержится в начальных моментах:

$$\nu_k(n) = \sum_{m=0}^{n} m^k \cdot P(n,m) \quad (k = 0, 1, \ldots). \tag{2}$$

В частности,

$$\nu_1(n) = np, \qquad \sigma(n)^2 = \nu_2(n) - \nu_1(n)^2 = np(1-p), \tag{3}$$

соответственно математическое ожидание и дисперсия m.

Чтобы определить асимптотику процесса, исследуем асимметрии и эксцессы m:

$$A_{2k-1}(n) = \frac{\mu_{2k-1}(n)}{\sigma(n)^{2k-1}}, \qquad E_{2k}(n) = \frac{\mu_{2k}(n)}{\sigma(n)^{2k}} - \omega(k), \ (k = 2, 3, \ldots), \tag{4}$$

где $\omega(k)$ и $\mu_k(n)$ – центральные моменты:

$$\omega(k) = \int_{-\infty}^{\infty} x^{2k} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 \cdot 3 \cdot \ldots \cdot (2k-1), \tag{5}$$

$$\mu_k(n) = \sum_{m=0}^{n} \left(m - \nu_1(n)\right)^k \cdot P(n,m), \tag{6}$$

Их удобно вычислять с помощью характеристической функции:

$$\theta(\tau, n) = \sum_{m=0}^{n} e^{i\tau(m-np)} \cdot P(n,m) = \left(p e^{i\tau q} + q e^{-i\tau p}\right)^n, \tag{7}$$

$$\mu_k(n) = i^{-k} \theta^{(k)}(0, n), \tag{8}$$

где i – мнимая единица, (k) – порядок производной по τ.

Из формул (4)–(8) следует, что

$$A_{2k-1}(n,p)=\sum_{l=1}^{k-1}\frac{f_{2l-1,2k-1}(p)}{\sigma(n)^{2l-1}},\qquad E_{2k}(n,p)=\sum_{l=1}^{k-1}\frac{f_{2l,2k}(p)}{\sigma(n)^{2l}},\tag{9}$$

где $f_{1,3}(p)$, $f_{2,4}(p)$, $f_{1,5}(p)$ – независящие от n коэффициенты разложения:

$$f_{1,3}(p)=1-2p,\qquad f_{2,4}(p)=6p^2-6p+1,\qquad f_{1,5}(p)=10(1-2p).\tag{10}$$

Согласно (3), (9) при $n\to\infty$ асимметрия и эксцесс исчезают, что характерно для гауссовского процесса

$$P_0(n,m)=\frac{1}{\sqrt{2\pi}\sigma(n)}e^{-\frac{x(n,m)^2}{2}},\qquad x(n,m)=\frac{m-np}{\sigma(n)}.\tag{11}$$

Выясним вклад в

$$\Delta(n,m)=P(n,m)-P_0(n,m)\tag{12}$$

асимметрий и эксцессов произвольных порядков. С учетом (9) запишем

$$\Delta(n,m)=P_0(n,m)\sum_{l=1}^{\infty}\frac{\varphi_l(x(n,m))}{\sigma(n)^l}=\Delta_1(n,m)+\Delta_2(n,m)+\dots,\tag{13}$$

где старшая степень полиномов

$$\varphi_{2l-1}(x)=\sum_{i=0}^{l}x^{2i+1}C_{2l-1,\,2i+1},\qquad \varphi_{2l}(x)=\sum_{i=0}^{l+1}x^{2i}C_{2l,\,2i}\qquad (l=1,2,\dots)\tag{14}$$

равна количеству линейно независимых уравнений, получаемых с помощью вытекающего из (6) и (11)–(13) тождества

$$\frac{\mu_k(n)}{\sigma(n)^k}=\sum_{m=0}^{n}x(n,m)^k\cdot P_0(n,m)\left(1+\sum_{l=1}^{\infty}\frac{\varphi_l(x(n,m))}{\sigma(n)^l}\right)\qquad (k=0,1,\dots)\tag{15}$$

для вычисления фигурирующих в (14) коэффициентов $C_{l,i}$. При $n\to\infty$ сумма по m в (15) становится интегральной

$$\Delta x(n,m)=x(n,m+1)-x(n,m)=\sigma(n)^{-1}\xrightarrow[n\to\infty]{}0,$$

$$x(n,0)=-\sqrt{np/q}\xrightarrow[n\to\infty]{}-\infty,\qquad x(n,n)=\sqrt{nq/p}\xrightarrow[n\to\infty]{}\infty,$$

поэтому

$$\frac{\mu_k(n)}{\sigma(n)^k}\xrightarrow[n\to\infty]{}\int_{-\infty}^{\infty}x^k\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\left(1+\sum_{l=1}^{\infty}\frac{\varphi_l(x)}{\sigma(n)^l}\right)dx.\tag{16}$$

Уравнения относительно $C_{l,i}$ получаются приравниванием в (16) выражений при одинаковых степенях $\sigma(n)$. Поскольку интеграл от нечетной функции в симметричных пределах равен нулю, четные и нечетные k будем рассматривать отдельно:

$$\frac{\mu_{2k+1}(n)}{\sigma(n)^{2k+1}}\xrightarrow[n\to\infty]{}\int_{-\infty}^{\infty}x^{2k+1}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\sum_{l=1}^{\infty}\frac{\varphi_{2l-1}(x)}{\sigma(n)^{2l-1}}dx,\tag{17}$$

$$\frac{\mu_{2k}(n)}{\sigma(n)^{2k}}\xrightarrow[n\to\infty]{}\int_{-\infty}^{\infty}x^{2k}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\left(1+\sum_{l=1}^{\infty}\frac{\varphi_{2l}(x)}{\sigma(n)^{2l}}\right)dx.\tag{18}$$

Из (17) с учетом (5), (14) следует

$$\frac{\mu_{2k+1}(n)}{\sigma(n)^{2k+1}} \xrightarrow[n\to\infty]{} \sum_{l=1}^{\infty} \frac{1}{\sigma^{2l-1}} \sum_{i=0}^{l} C_{2l-1,2i+1}\omega(k+i+1)\,, \qquad (19)$$

откуда для $k=0$ с учетом вытекающего из (2), (6) тождества $\mu_1(n)\equiv 0$ получим

$$\sum_{i=0}^{l} C_{2l-1,2i+1}\omega(i+1)=0 \quad (l=1,2,\ldots)\,. \qquad (20)$$

Для остальных k с помощью (4), (9) преобразуем левую часть (19)

$$\sum_{l=1}^{k} \frac{f_{2l-1,2k+1}(p)}{\sigma^{2l-1}} \xrightarrow[n\to\infty]{} \sum_{l=1}^{\infty} \frac{1}{\sigma^{2l-1}} \sum_{i=0}^{l} C_{2l-1,2i+1}\omega(k+i+1) \quad (k=1,2,\ldots)\,. \qquad (21)$$

Пусть в (21) $k=1$, тогда

$$\sum_{i=0}^{1} C_{1,2i+1}\omega(i+2)=f_{13}(p)\,, \qquad \sum_{i=0}^{l} C_{2l-1,2i+1}\omega(i+2)=0 \quad (l=2,3,\ldots). \qquad (22)$$

Аналогично для $k=2$ из (21) следует

$$\sum_{i=0}^{1} C_{1,2i+1}\omega(i+3)=f_{15}(p)\,, \qquad \sum_{i=0}^{2} C_{3,2i+1}\omega(i+3)=f_{35}(p)\,, \qquad (23)$$

$$\sum_{i=0}^{l} C_{2l-1,2i+1}\omega(i+3)=0 \quad (l=3,4,\ldots)\,.$$

Действуя так и далее, можно записать уравнения для определения коэффициентов любого нечетного полинома. В частности, объединив первые уравнения (20), (22) в систему, найдем

$$\left.\begin{array}{r} C_{11}+3C_{13}=0 \\ 3C_{11}+15C_{13}=f_{13}(p) \end{array}\right\} \Rightarrow \begin{pmatrix} C_{11} \\ C_{13} \end{pmatrix} = \frac{f_{13}(p)}{2}\begin{pmatrix} -1 \\ 1/3 \end{pmatrix}. \qquad (24)$$

Согласно (10) первое уравнение в (23) удовлетворяется теми же значениями коэффициентов:

$$15C_{11}+105C_{13}=\frac{f_{13}(p)}{2}\left(15\cdot(-1)+105\cdot\frac{1}{3}\right)=10\,f_{13}(p)\equiv f_{15}(p)\,,$$

т.е. является линейно зависимым с фигурирующими в (24). Так же можно убедиться в линейной зависимости вытекающих из (16) уравнений относительно коэффициентов $\varphi_l(x)$, если количество этих уравнений больше числа коэффициентов в (14). Значит (14) правильно определяет старшие степени полиномов в разложении (13).

Для определения $\varphi_3(x)$ положим в (20), (22) $l=2$ и дополним полученную систему вторым уравнением из (23)

$$\left.\begin{array}{r} C_{31}+3C_{33}+15C_{35}=0 \\ 3C_{31}+15C_{33}+105C_{35}=0 \\ 15C_{31}+105C_{33}+945C_{35}=f_{35}(p) \end{array}\right\} \Rightarrow \begin{pmatrix} C_{31} \\ C_{33} \\ C_{35} \end{pmatrix} = \frac{f_{35}(p)}{8}\begin{pmatrix} 1 \\ -2/3 \\ 1/15 \end{pmatrix}. \qquad (25)$$

Решения (24), (25) позволяют заметить общую закономерность

$$C_{2l-1,2i+1}=\frac{f_{2l-1,2l+1}(p)}{\rho(l)}\cdot\frac{(-1)^{i+l}C_l^i}{\omega(i+1)} \qquad \begin{pmatrix} l=1,2,\ldots \\ i=0,1,\ldots,l. \end{pmatrix}, \qquad (26)$$

где C_l^i – числа сочетаний;

$\rho(l)$ – произведения четных чисел

$$\rho(l) = 2\cdot 4\cdot \ldots \cdot 2l\,. \tag{27}$$

Аналогично рассмотрим в (16) четные k. Из формулы (18) с учетом (5), (14) следует

$$\frac{\mu_{2k}(n)}{\sigma(n)^{2k}} - \omega(k) \xrightarrow[n\to\infty]{} \sum_{l=1}^{\infty}\frac{1}{\sigma^{2l}}\sum_{i=0}^{l+1} C_{2l,2i}\,\omega(k+i)\,. \tag{28}$$

Полагая в (28) $k=0$ и $k=1$, с учетом тождеств $\mu_0(n)\equiv 1$, $\mu_2(n)\equiv\sigma(n)^2$ и равенств $\omega(0)=1=\omega(1)$ (см. формулу (5)) получим

$$\sum_{i=0}^{l+1} C_{2l,2i}\,\omega(i) = 0\,, \qquad \sum_{i=0}^{l+1} C_{2l,2i}\,\omega(i+1) = 0 \qquad (l=1,2,\ldots)\,. \tag{29}$$

Для остальных k с помощью (4), (5), (9) преобразуем левую часть (28):

$$\sum_{l=1}^{k-1}\frac{f_{2l,2k}(p)}{\sigma^{2l}} \xrightarrow[n\to\infty]{} \sum_{l=1}^{\infty}\frac{1}{\sigma^{2l}}\sum_{i=0}^{l+1} C_{2l,2i}\,\omega(k+i) \qquad (k=2,3,\ldots)\,. \tag{30}$$

Пусть в (30) $k=2$, тогда

$$\sum_{i=0}^{2} C_{2,2i}\,\omega(i+2) = f_{2,4}(p)\,, \qquad \sum_{i=0}^{l+1} C_{2l,2i}\,\omega(i+2) = 0 \qquad (l=2,3,\ldots)\,. \tag{31}$$

Аналогично для $k=3$ из (30) следует

$$\sum_{i=0}^{2} C_{2,2i}\,\omega(i+3) = f_{2,6}(p)\,, \qquad \sum_{i=0}^{3} C_{4,2i}\,\omega(i+3) = f_{4,6}(p)\,, \tag{32}$$

$$\sum_{i=0}^{l+1} C_{2l,2i}\,\omega(i+3) = 0 \qquad (l=3,4,\ldots)\,. \tag{33}$$

Полагая в (29) $l=1$, дополним полученную систему первым уравнением (31)

$$\left.\begin{aligned} C_{20}+C_{22}+3C_{24}&=0\\ C_{20}+3C_{22}+15C_{24}&=0\\ 3C_{20}+15C_{22}+105C_{24}&=f_{24}(p)\end{aligned}\right\}\Rightarrow \begin{pmatrix}C_{20}\\ C_{22}\\ C_{24}\end{pmatrix} = \frac{f_{24}(p)}{8}\begin{pmatrix}1\\ -2\\ 1/3\end{pmatrix}. \tag{34}$$

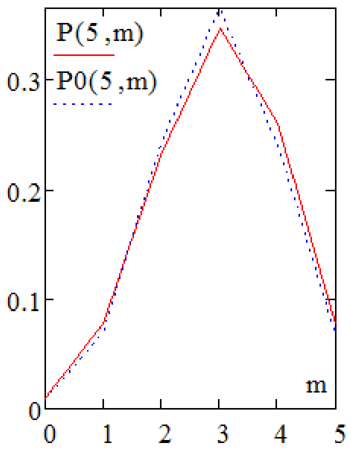
Для получения $\varphi_4(x)$ положим в (29), (31) $l=2$ и дополним полученную систему вторым уравнением (32)

$$\left.\begin{aligned} C_{40}+C_{42}+3C_{44}+15C_{46}&=0\\ C_{40}+3C_{42}+15C_{44}+105C_{46}&=0\\ 3C_{40}+15C_{42}+105C_{44}+945C_{46}&=0\\ 15C_{40}+105C_{42}+945C_{44}+945\cdot 11C_{46}&=f_{46}(p)\end{aligned}\right\}\Rightarrow \begin{pmatrix}C_{40}\\ C_{42}\\ C_{44}\\ C_{46}\end{pmatrix} = \frac{f_{46}(p)}{48}\begin{pmatrix}-1\\ 3\\ -1\\ 1/15\end{pmatrix}. \tag{35}$$

Полученные решения позволяют заметить общую закономерность для четных полиномов

$$C_{2l,2i} = \frac{f_{2l,2(l+1)}(p)}{\rho(l+1)}\cdot\frac{(-1)^{i+l+1}C_{l+1}^{i}}{\omega(i)} \qquad \begin{pmatrix} l=1,2,\ldots\\ i=0,1,\ldots,l+1.\end{pmatrix} \tag{36}$$

Таким образом, формулы (3), (11)–(14), (26), (27), (36) позволяют заменить трудно вычислимое при больших n выражение для вероятности m успехов в схеме Бернулли. Результаты выполненных по ним расчетов в графической форме представлены на рисунке 1. Видно, что даже для $n=5$ ошибка нулевого приближения (рис. 1, *а*) практически полностью исчезает (рис. 1, *б*) при учете асимметрии и эксцесса минимальных порядков ($P_2(n,m) = P_0(n,m) + \Delta_1(n,m) + \Delta_2(n,m)$, см. (12), (13)).



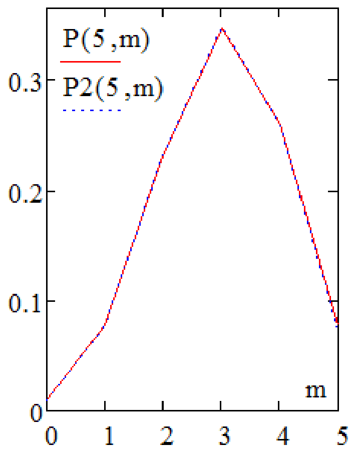


Рисунок 1. – Влияние асимметрии и эксцесса минимальных порядков на вероятность m успехов в 5 опытах схемы Бернулли с *p* = 0,6

Пользоваться локальной теоремой Лапласа (отвечающей нулевому приближению в разложении (12), (13)) нельзя, если $|\Delta(m,n)| \neq o\left(P_0(n,m)\right)$. В частности, для m далеких от наиболее вероятного числа успехов, $|x(n,m)|$ при увеличении числа опытов растет пропорционально $\sqrt{n}$ (см. (3), (11)). При этом, согласно (13), (14), главная часть общего члена ряда для $|\Delta(m,n)| / P_0(n,m)$ пропорциональна n ($x(n,m)^{l+2} / \sigma(n)^l \sim n \to \infty$). То есть относительная погрешность асимптотической формулы (11), фигурирующей в локальной теореме Лапласа, неограниченно возрастает (рис. 2) и пользоваться ею нельзя, что, однако, не всегда оговаривается [1].

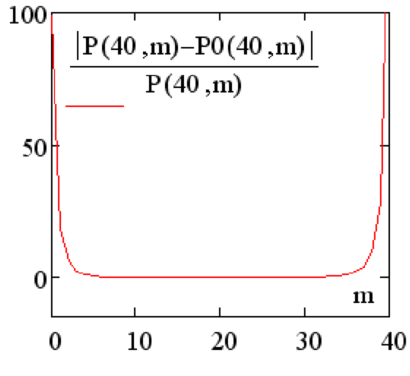


Рисунок 2. – Относительная погрешность нулевого приближения для *p* = 0,6

Выводы. Независимые повторные испытания предложено рассматривать как случайный процесс с дискретным временем. Методом моментов установлена асимптотика процесса и определен вклад асимметрий и эксцессов произвольных порядков в отклонение распределения Бернулли от асимптотической формулы, фигурирующей в локальной теореме Лапласа. Показано, что последнюю нельзя применять для вычисления вероятностей чисел успехов, далеких от своих наиболее вероятных значений.

ЛИТЕРАТУРА

1. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман // М. : Высш. шк., 1975. – 334 с.

2. Теоретико-вероятностный подход к решению уравнения диффузии / С. Г. Ехилевский // Вестн. Полоц. гос. ун-та. Сер. С, Фундаментальные науки. – 2015. – № 4. – С. 94–105.
3. Ехилевский, С. Г. Влияние асимметрии высших порядков на динамику сорбции вредной примеси / С. Г. Ехилевский, О. Н. Мурашкевич // Вестн. Полоц. гос. ун-та. Сер. В, Промышленность. Прикладные науки. – 2014. – № 3. – С. 115–122.
4. Ехилевский, С. Г. Теоретико-вероятностный подход к моделированию динамической сорбционной активности / С. Г. Ехилевский, О. В. Голубева, Д.В. Пяткин // Вестн. Полоц. гос. ун-та. Сер. В, Промышленность. Прикладные науки. – 2013. – № 11. – С. 144–151.
5. Ехилевский, С. Г. Метод моментов и динамика сорбционной активности при малых временах / С. Г. Ехилевский, О. В. Голубева, С. А. Ольшаников // Вестн. Полоц. гос. ун-та. Сер. В, Промышленность. Прикладные науки. – 2013. – № 3. – С. 150–156.

Поступила 20,09,2016

INDEPENDENT RE-TESTING AS THE ASYMPTOTIC GAUSSIAN RANDOM PROCESS

S. EHILEVSKY, O. GOLUBEVA, E. POTAPENKO, T. RUDKOVA

Independent re-tests are considered as a random process with discrete time. The method of moments set the asymptotic behavior of the process and defined contribution asymmetries and excesses of arbitrary order in the deviation of the Bernoulli distribution from the asymptotic formulae appearing in the local theorem of Laplace. It is shown that the latter cannot be applied to calculate the probability of success is far from their most probable values.

Keywords: *Bernoulli scheme, a random process, the method of moments.*