УДК 519.1+519.2 *Н.А. Колокольникова, к.ф.-м.н., доцент, ИГУ*

Одно обобщение урновой схемы Маркова-Пойа

В работе описывается один вариант так называемой Ф-схемы последовательных испытаний, обобщающий известную урновую схему Маркова-Пойа. Приводится явный вид возникающего вероятностного распределения и находятся его числовые характеристики.

Ключевые слова: Ф-схема последовательных испытаний, вероятность, распределение, обобщённые числа Стирлинга, база.

Введение

При решении ряда задач о случайных размещениях удобно использовать специальные комбинаторные схемы, в частности, так называемые А-, В-, Ф-схему последовательных испытаний [1]. Комбинаторные числа, участвующие в записи получаемых при этом вероятностных распределений, строятся из элементов одной или двух последовательностей, называемых базовыми последовательностями (или базами). Конкретизация баз приводит к конкретной задаче случайного размещения. Таким образом, исследование распределения, записанного в А-, В- или Ф-форме – по сути исследование целого класса вероятностных распределений, наделённых общей структурой. Следует отметить, что один частный случай Ф-схемы последовательных испытаний был достаточно подробно исследован [2] и показал свою работоспособность при решении ряда задач, связанных с изучением числа пустых ячеек при размещении комплектов разного объёма, а также при размещении случайного числа частиц (см., например, [3–5]).

В данной статье изучается другой вариант Ф-схемы последовательных испытаний, частным случаем которого является хорошо известная схема Маркова-Пойа (см., например, [6]), впервые введённая в рассмотрение А.А. Марковым в работе [7], а затем часто используемая и обобщаемая разными авторами, о чём подробно написано в статье [8].

1. Ф-схема последовательных испытаний

Пусть проводятся испытания типа «успех-неуспех». Обозначим ξ_n – число успехов в n испытаниях. Предполагаем, что условные вероятности успехов имеют такую структуру:

$$p_{ni} = P\{\xi_n = i + 1 \mid \xi_{n-1} = i\} = \alpha_{n-1}\beta_k,$$

$$q_{ni} = 1 - p_{ni} = P\{\xi_n = i \mid \xi_{n-1} = i\} = \alpha_{n-1}(\delta_{n-1} + \gamma_k),$$

$i = \overline{0, n-1}, \quad n = \overline{1, \infty}.$

Тогда распределение случайной величины ξ_n может быть записано в виде

$$P\{\xi_n = 0\} = \Phi_0^n \cdot \prod_{i=0}^{n-1} \alpha_i,$$

$$P\{\xi_n = k\} = \Phi_k^n \cdot \prod_{i=0}^{n-1} \alpha_i \cdot \prod_{j=0}^{k-1} \beta_j, \qquad k = \overline{1,n},\ n = \overline{1,\infty}. \qquad (1)$$

Распределение (1) называется Ф-распределением.

Комбинаторные числа Φ_k^n , участвующие в распределении (1), могут быть представлены в виде

$$\Phi_k^n = \sum_{i=k}^{n} B_i^n A_k^i,$$

где B_i^n– обобщённые числа Стирлинга 1-го рода, строящиеся на базе $\{\alpha_i\}_{i=0}^{n-1}$, A_k^i– обобщённые числа Стирлинга 2-го рода, строящиеся на базе $\{\gamma_i\}_{i=0}^{k-1}$.

При этом

$$\Phi_n^n = 1, \qquad n = \overline{0,\infty},$$

$$\Phi_k^n = 0, \text{ если } n < k \text{ или } k < 0. \qquad (2)$$

2. **Частный случай**

Ранее детально изучался случай, когда $\beta_k = C - k,$ где C –некоторое бльшое натуральное число (которое в дальнейшем можно устремлять к бесконечности), α_{n-1} $(n = \overline{1,\infty)}$ – вообще говоря, произвольные положительные числа такие, что $0 \le \alpha_{n-1}(C - i) \le 1$ при каждом i. Этот вариант Ф-распределения, как было отмечено выше, показал свою эффективность при изучении распределения числа пустых ячеек, если частицы размещаются комплектами, а также если размещается случайное число частиц.

Рассмотрим теперь другой частный случай Ф-распределения. Пусть $\beta_k = C + k,\ \gamma_k = -k,$ где $C,$ как и в упомянутом выше случае натуральное число. Тогда $\delta_{n-1} = \frac{1}{\alpha_{n-1}} + C$, а распределение (1) запишется в виде:

$$P\{\xi_n = k\} = \Phi_k^n \prod_{i=0}^{n-1} \alpha_i [C]_k, \qquad k = \overline{0,n},\ n = \overline{1,\infty}, \qquad (3)$$

где $[C]_0 = 1,\ \ [C]_k = \prod_{j=0}^{k-1}(C + j),\ k \ge 1;$ Φ_k^n – комбинаторные числа, которые могут быть определены следующим образом:

$$\Phi_k^n = \Phi_{k-1}^{n-1} + (\frac{1}{\alpha_{n-1}} - C + k)\Phi_k^{n-1}, \quad k = \overline{0,n},\ n = \overline{1,\infty}. \qquad (4)$$

При этом, естественно, выполняется условие (2).

3. Пример (урновая схема Маркова-Пойа)

Урна содержит a белых и b чёрных шаров. Произвольно вынимается шар, фиксируется его цвет, а потом шар возвращается в урну вместе с d дополнительными шарами того же цвета. Описанная процедура осуществляется последовательно n раз.

Пусть ξ_n – число вынутых белых шаров. Считая успехом при каждом вынимании появление белого шара, а неуспехом – черного, имеем

$$p_{nk} = \frac{a+kd}{a+b+(n-1)d}, \quad q_{nk} = \frac{b+(n-1-k)d}{a+b+(n-1)d}, \ k = \overline{0, n-1}, n = \overline{1, \infty}.$$

Полагаем $\alpha_{n-1} = \frac{d}{a+b+(n-1)d}$, $\beta_k = \frac{a}{d} + k$, $\delta_{n-1} = \frac{b}{d} + n - 1$, $\gamma_k = -k$.

Видим, что структура элементов α_{n-1}, β_k, δ_{n-1}, γ_k соответствует рассматриваемому варианту Ф-распределения, $C = \frac{a}{d}$. Формула (3) принимает вид

$$P\{\xi_n = k\} = \Phi_k^n \frac{[\frac{a}{d}]_k}{[\frac{a+b}{d}]_n}, \qquad k = \overline{0, n}, \qquad n = \overline{1, \infty}.$$

Итак, распределение Маркова-Пойа является частным случаем изучаемого варианта Ф-распределения.

4. Числовые характеристики

Для нахождения числовых характеристик будем использовать производящую функцию распределения (3):

$$\Phi_n(x) = \prod_{i=0}^{n-1} \alpha_i \cdot \sum_{k=0}^{n} \Phi_k^n \, [C]_k \, x^k. \qquad (5)$$

Рассмотрим факториальные моменты этого распределения. Имеет место утверждение:

Теорема. Для любого натурального n факториальный момент порядка $1 \le r \le n$ имеет вид

$$m_r = \Phi_n^{(r)}(1) = r! \, [C]_r \sum_{i=r}^{n} a_r^i \, B_{n-i}^n, \qquad (6)$$

где $\Phi_n^{(r)}(x)$ – производная r-го порядка от производящей функции (5),

$[C]_k$ – как указывалось выше, k-я факториальная степень,

a_r^i –числа Стирлинга 2-го рода, B_{n-i}^n – обобщённые числа Стирлинга 1-го рода, строящиеся на базе $\{\alpha_i\}_{i=0}^{n-1}$.

Доказательство проводится индукцией по n.

При $n = 1$ имеем очевидное равенство

$$m_1 = \Phi_1'(1) = \alpha_0 C = 1!\,[C]_1\alpha_0.$$

Предположим, что равенство (6) выполняется при некотором натуральном $n \geq 1$ и всех факториальных моментов порядка, не превышающего r. Тогда

$$m_r = m_r(n+1) = \Phi_{n+1}^{(r)}(1) =$$

$$= (1 + r\alpha_n)\Phi_n^{(r)}(1) + r\alpha_n(C + r - 1)\Phi_n^{(r-1)}(1) =$$

$$= (1 + r\alpha_n)r!\,[C]_r \textstyle\sum_{i=r}^{n} a_r^i\, B_{n-i}^n +$$

$$+ r\alpha_n(C + r - 1)(r - 1)!\,[C]_r \sum_{i=r-1}^{n} a_{r-1}^i\, B_{n-i}^n =$$

$$= r!\,[C]_r \left(\sum_{i=r}^{n} a_r^i\, B_{n-i}^n + \alpha_n \sum_{i=r}^{n+1} a_r^i\, B_{n-i+1}^n\right) =$$

$$= r!\,[C]_r \left(\sum_{i=r}^{n+1} a_r^i\, (B_{n-i}^n + \alpha_n B_{n-i+1}^n)\right) =$$

$$= \; r!\,[C]_r \textstyle\sum_{i=r}^{n+1} a_r^i\, B_{n-i+1}^{n+1}.$$

Если $r = n + 1$, то

$$m_{n+1} = m_{n+1}(n+1) = \Phi_{n+1}^{(n+1)}(1) = (n+1)!\, \textstyle\prod_{i=0}^{n} \alpha_i \;\cdot [C]_{n+1}\Phi_{n+1}^{n+1} =$$

$$= \quad (n+1)!\; [C]_{n+1} B_0^{n+1}.$$

Таким образом, теорема доказана.

Используя результат (6), получим формулы для математического ожидания и дисперсии величины ξ_n. Имеем

$$\mathrm{M}\xi_n = \Phi_n'(1) = C\big(\textstyle\prod_{i=0}^{n-1}(1 + \alpha_i) - 1\big). \qquad (7)$$

Дисперсия ξ_n находится по формуле

$$\mathrm{D}\xi_n = \Phi''_n(1) + \Phi'_n(1) - (\Phi'_n(1)^2.$$

На основании формулы (6),

$$\mathrm{D}\xi_n = C(C+1)\left(1 - 2\prod_{i=0}^{n-1}(1+\alpha_i) + \prod_{i=0}^{n-1}(1+2\alpha_i)\right) +$$

$$+C\left(\prod_{i=0}^{n-1}(1+\alpha_i) - 1\right) - C^2\left(\prod_{i=0}^{n-1}(1+\alpha_i) - 1\right)^2 =$$

$$= C(C+1)\prod_{i=0}^{n-1}(1+2\alpha_i) + C\prod_{i=0}^{n-1}(1+\alpha_i) - C^2\prod_{i=0}^{n-1}(1+\alpha_i)^2. \quad (8)$$

5. Асимптотические формулы математического ожидания и дисперсии

Из задания вероятностей p_{ni} и q_{ni} следует, что $\alpha_i \le \frac{1}{n+C}$ для любого i. Тогда при $n \to \infty$ имеют место представления

$$\prod_{i=0}^{n-1}(1+\alpha_i) = \exp\left\{\sum_{i=0}^{n-1}\alpha_i - \frac{1}{2}\sum_{i=0}^{n-1}{\alpha_i}^2 + o\left(\sum_{i=0}^{n-1}{\alpha_i}^2\right)\right\},$$

$$\prod_{i=0}^{n-1}(1+2\,\alpha_i) = \exp\left\{2\sum_{i=0}^{n-1}\alpha_i - 2\sum_{i=0}^{n-1}{\alpha_i}^2 + o\left(\sum_{i=0}^{n-1}{\alpha_i}^2\right)\right\}.$$

Обозначим $a = \sum_{i=0}^{n-1}\alpha_i,\ \ b = \sum_{i=0}^{n-1}{\alpha_i}^2$. Тогда

$$\prod_{i=0}^{n-1}(1+\alpha_i) = e^{a-\frac{b}{2}+o(b)},$$

Подставляя полученные выражения (7) и (8), получим: при $n \to \infty$

$$\mathrm{M}\xi_n = C\left(e^{a-\frac{b}{2}+o(b)} - 1\right), \quad (9)$$

$$\mathrm{D}\xi_n = Ce^a\left(C\left(e^{2a-2b+o(b)} - e^{a-b+o(b)}\right) + e^{a-2b+o(b)} - e^{-\frac{b}{2}+o(b)}\right). \quad (10)$$

Поскольку $\alpha_i \le \frac{1}{n+C}$ для любого i, то a должно удовлетворять неравенству

$a \le \frac{n}{C+n}$. Отсюда следует, что при $n \to \infty, C \to \infty$ возможны случаи:

1. $a \to 0$, если $n = o(C)$,

2. $a \to \gamma,\ \ 0 < \gamma < 1,$ если $n = O(C)$,
3. $a \to 1$, если $C = o(n)$.

Согласно заданию b, должно выполняться неравенство $b \le \frac{a}{n+C}$. Следовательно, $n \to \infty, C \to \infty$ могут возникнуть случаи:

1) $Cb = O(a)$, если $n = o(C)$,
2) $Cb = o(a)$, если $C = o(n)$.

При ситуациях 1) и 2) запишем выражения для $\mathrm{M}\xi_n$ и $\mathrm{D}\xi_n$ и проанализируем их поведение, если $n \to \infty, C \to \infty$.

5.1. Случай $\boldsymbol{n = o(C)}$ при $\boldsymbol{n} \to \infty, \boldsymbol{C} \to \infty$

Очевидно, что при сделанных предположениях $a \to 0$.

1. Пусть $Ca \to \infty$. Тогда $b = o(a^2)$. В этом случае формула (9) примет вид

$$\mathrm{M}\xi_n = Ca + \frac{Ca^2}{2} + o(Ca^2).$$

$\mathrm{M}\xi_n \to \infty$ при $n, C \to \infty$.

Формула (10) может быть записана в виде

$$\mathrm{D}\xi_n = Ce^a\left(a - Cb + \frac{Ca^2}{2}\right)(1 + o(1)).$$

Поведение дисперсии при $n, C \to \infty$ может быть различным (в зависимости от дополнительно накладываемых условий): она может стремиться либо к бесконечности, либо к нулю, либо к константе, отличной от нуля. Это приводит к различным предельным распределениям.

2. Если $Ca \to const$ при $n \to \infty, C \to \infty$, то $b = O(a^2)$. Тогда

$$\mathrm{D}\xi_n = Ca(1 + o(1)).$$

При этом

$$\mathrm{M}\xi_n = Ca(1 + o(1)).$$

Следовательно, если $Ca \to \lambda < \infty$, то $\mathrm{M}\xi_n \to \lambda$, $D\xi_n \to \lambda$.

3. Если $Ca \to 0$ при $n \to \infty, C \to \infty$, то $\mathrm{M}\xi_n \to 0$, $D\xi_n \to 0$.

5.2. Случай $\boldsymbol{n = O(C)}$ при $\boldsymbol{n} \to \infty, \boldsymbol{C} \to \infty$

В данном случае $a \to \gamma,\ \ 0 < \gamma < 1, Ca \to \infty$. При этом $Cb \to a\delta,$ где $0 < \delta < 1$. Таким образом,

$$D\xi_n = Ce^a a(1-\delta+o(1)) \to \infty.$$

Из формулы (9) получим

$$M\xi_n = Ca(1+o(1)) \to \infty.$$

5.3. Случай $C = o(n)$ при $n \to \infty, C \to \infty$

При данном предположении имеем: $a \to 1$, $Ca \to \infty$. Тогда

$$D\xi_n = Ce^a\big(a - Cb + o(a)\big) \to \infty,$$

$$M\xi_n = Ca(1+o(1)) \to \infty.$$

Асимптотическое представление и анализ поведения математического ожидания и дисперсии изучаемой величины важны для нахождения предельных законов распределения. Заметим, что поиск предельных распределений упрощается, когда все корни производящей функции являются действительными и неположительными. Тогда рассматриваемую величину можно представить в виде суммы независимых случайных индикаторов. Но производящая функция $\Phi_n(x)$ случайной величины ξ_n имеет комплексные корни уже при $n = 2$. Поэтому для доказательства предельных теорем используются другие методы, в частности, метод моментов и теорема Бендера [9, С. 57].

Литература

1. Комбинаторные числа и полиномы в моделях дискретных распределений / В. Н. Докин, В. Д. Жуков, Н. А. Колокольникова, О. В. Кузьмин, М. Л. Платонов. Иркутск: Изд-во Иркут. ун-та, 1990. 208 с.
2. Колокольникова Н. А. Предельные теоремы для числа успехов в одной схеме зависимых испытаний // ВИНИТИ. 25.02.1992. №649. В.92.Деп. 25 с.
3. Колокольникова Н. А. О некоторых обобщениях классической задачи о дробинках // Платоновские чтения: материалы Регион. конф. Иркутск, 2019. 7 с. URL: http://math.isu.ru/ru/chairs/tpdm/docs/Platonovskie2019/Kolokolnikova-PR-2019-s.pdf.
4. Имыхелова В. П., Колокольникова Н. А. Одна схема случайного размещения частиц («скользящий комплект») // Асимптотические и перечислительные задачи комбинаторного анализа. Иркутск: Иркут. ун-т, 1997. С. 43-53.
5. Колокольникова Н. А. Вероятностные модели теории страхования, использующие схемы случайного размещения частиц // Информационные технологии и проблемы математического моделирования сложных систем. Вып. 15. Иркутск: ИрГУПС, 2015. С. 66-72.
6. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1. М: Мир, 1984. 528 с.

7. Марков А. А. О некоторых предельных формах исчисления вероятностей. Изв. Акад. Наук, Пг.VI сер.1917, т.11, № 3. С.177-186.
8. Ивченко Г..И., Медведев Ю. И. Об урновой схеме Маркова-Пойа: от 1917 г. до наших дней. // Обозрение прикладной и промышленной математики, сер. Дискретная математика, 1996. Т.3, вып.4. С. 484 -511.
9. Сачков В. Н. Вероятностные методы в комбинаторном анализе. М.: Наука, 1978. 288 с.

One generalization of the Markov-Poya urn scheme

N. A. Kolokolnikova

The paper describes a variant of the so-called Φ-scheme of sequential tests, which generalizes the well-known Markov-Poya urn scheme. An explicit form of the resulting probability distribution is given and its numerical characteristics are found.

Keywords: Φ-scheme of sequential tests, probability, distribution, generalized Stirling numbers, base.