УДК 519.21 *Н.А. Колокольникова*

к.ф.-м.н., доцент, ИГУ

О некоторых обобщениях классической задачи о дробинках

Так называемая «классическая задача о дробинках» [1] служит «отправным пунктом» при решении различных задач, связанных со случайным размещением частиц (дробинок) по ячейкам (ящикам). Многие авторы (например, [2], [3]) изучали различные обобщения «классической схемы», рассматривая размещения в полиномиальной схеме, размещения частиц комплектами и т.д.

Заметим, что при решении некоторых задач удобно использовать специальные схемы последовательных испытаний (В-схема, А-схема, Ф-схема и др.). Применение этих схем позволяет записывать в явном виде распределение изучаемой случайной величины. Покажем это на примере так называемой А-схемы последовательных испытаний. Во-первых, продемонстрируем эффективность этой схемы для представления распределения числа пустых ячеек в классической задаче о дробинках, во-вторых, рассмотрим случай, когда имеются ячейки двух типов.

1. А-схема последовательных испытаний

Предположим, что проводятся испытания типа «успех-неуспех», и после каждого успеха вероятность следующего успеха (а значит, и неуспеха) может меняться. После неуспеха изменения вероятностей не происходит. Такая схема называется А-схемой последовательных испытаний.

Пусть ξ_n – число успехов в n испытаниях, проводимых в условиях А-схемы, P_i- вероятность (i+1)-го успеха, т.е.

$$P_i = P\{\xi_{n+1} = i + 1 \,|\xi_n = i\}, \qquad i = \overline{0, n.}$$

Обозначим

$$Q_i = P\{\xi_{n+1} = i \,|\xi_n = i\} = 1 - P_i.$$

Теорема 1. [4] Если последовательные испытания проводятся в условиях А-схемы, то

$$P\{\xi_n = 0\} = A_{0,}^n$$

$P\{\xi_n = k\} = A_k^n \times \prod_{i=0}^{k-1} P_i$, $k = \overline{1, n}$, $n = \overline{1, \infty}$. (1)

Здесь A_k^n – обобщенные числа Стирлинга 2-го рода, строящиеся на базе $\{Q_i\}_{i=0}^{\infty}$.

2. Классическая задача о дробинках

Пусть имеется N ячеек, в которых случайным образом размещаются n частиц. Предполагается, что каждая частица с вероятностью $\frac{1}{N}$ может попасть в любую ячейку. Обозначим $\mu_0 = \mu_0(N, n)$число ячеек, оставшихся распределения этой величины:

$$P\{\mu_0(N, n) = k\} = \binom{n}{k} (1 - \frac{k}{N})^n P\{\mu_0(N - k, n) = 0\}, \quad (2)$$

где

$$P\{\mu_0(N, n) = 0\} = \sum_{i=0}^{N} \binom{N}{i} (-1)^i (1 - \frac{i}{N})^n. \quad (3)$$

Заметим, что распределение (2), использующее формулу (3), можно записать в более компактной форме. Будем считать, что частицы последовательно «бросаются» в ячейки. Под успехом будем понимать попадание очередной частицы в пустую ячейку, под неуспехом – попадание в ячейку, где уже находится по меньшей мере одна частица. Нетрудно видеть, что выполняются условия А-схемы последовательных испытаний. Вероятность успеха в очередном испытании, если ранее было зафиксировано i успехов, равна

$$P_i = \frac{N - i}{N}, \quad i = \overline{0, \min(n, N)}.$$

Для случайной величины ξ_n – числа непустых ячеек после размещения n частиц – имеем

$$P\{\xi_n = 0\} = 0,$$

$$P\{\xi_n = k\} = A_k^n (N)_k \frac{1}{N^k}, \quad k = \overline{1, \min(n, N)}, \ n = \overline{1, \infty}, \ (4)$$

где числа A_k^n строятся на базе $\left\{\frac{i}{N}\right\}_{i=0}^{\min(n,N)}$,

$$(N)_k = N(N - 1) \dots (N - k + 1).$$

В данном случае, используя свойства чисел A_k^n, получим

$$A_k^n = \frac{1}{N^{n-k}} a_k^n,$$

где a_k^n – числа Стирлинга второго рода. Поэтому

$$P\{\xi_n = k\} = a_k^n (N)_k \frac{1}{N^n}, \quad k = \overline{1, \min(n, N)},\ n = \overline{1, \infty}. \qquad (5)$$

Поскольку

$$\mu_0 = \mu_0(N, n) = N - \xi_n,$$

то

$$P\{\mu_0(N, n) = k\} = a_{N-k}^n (N)_{N-k} \frac{1}{N^n}.$$

Формула (5), безусловно, является более удобной для использования, чем формула (2) с использованием (3).

3. Размещение в две совокупности ячеек

Предположим, что, как и ранее, частиц n размещаются в N ячейках, но при этом каждая частица с вероятностью α попадает в любую и $N_1 < N$ ячеек и с вероятностью β – в любую из остальных $N_2 = N - N_1$ячеек, т.е.

$\alpha N_1 + \beta N_2 = 1$. Не умаляя общности, можно считать, что имеется две совокупности ячеек объемом соответственно N_1 и N_2, причем $N_1 + N_2 = N$. Каждая из размещаемых частиц с вероятностью p_1 попадает в первую совокупность и с вероятностью p_2 – во вторую. Попав в какую-либо совокупность ячеек, частица с одинаковой вероятностью может оказаться в любой ячейке этой совокупности (т.е. попадание в каждую ячейку первой совокупности возможно с вероятностью $\frac{1}{N_1}$, а в любую ячейку второй совокупности – $\frac{1}{N_2}$).

Как и ранее, успехом будем считать попадание частицы в пустую ячейку, ξ_n - общее число непустых ячеек после размещения n частиц.

Пусть P_i – вероятность (i+1)-го успеха, i=0,1,…Очевидно, что $P_0 = 1$. Вероятность 2-го успеха находится следующим образом:

$$P_1 = p_1^2 \frac{N_1 - 1}{N_1} + p_1 p_2 \frac{N_2}{N_2} + p_2 p_1 \frac{N_1}{N_1} + p_2^2 \frac{N_2 - 1}{2} =$$

$$= p_1^2 + 2p_1p_2 + p_2^2 - \frac{p_1^2}{N_1} - \frac{p_2^2}{N_2} = 1\text{-}\left(\frac{p_1^2}{N_1} + \frac{p_2^2}{N_2}\right).$$

Продолжая процесс нахождения вероятностей P_i, предполагая, что $i \le \min(N_1, N_2)$, получим

$$P_i = p_1^{i+1}\frac{N_1 - i}{N_1} + \binom{i}{1} p_1^i p_2 \frac{N_1-(i-1)}{N_1} + \cdots + \binom{i}{i} p_1 p_2^i \frac{N_1}{N_1} +$$

$$+ p_2^{i+1}\frac{N_2 - i}{N_2} + \binom{i}{1} p_2^i p_1 \frac{N_2-(i-1)}{N_2} + \cdots + \binom{i}{i} p_2 p_1^i \frac{N_2}{N_2} =$$

$$= \sum_{j=0}^{i} \binom{i}{j} \left(p_1^{i+1-j} p_2^j \frac{N_1-(i-j)}{N_1} + p_2^{i+1-j} p_1^j \frac{N_2-(i-j)}{N_2}\right) =$$

$$= 1 - i\left(\frac{p_1^2}{N_1} + \frac{p_2^2}{N_2}\right). \qquad (6)$$

Поскольку вероятности успехов меняются после каждого успеха, то находимся в условиях А-схемы последовательных испытаний, и распределение числа непустых ячеек – это распределение числа успехов в n испытаниях, которое определяется формулой (1). В данном случае базовые элементы, из которых строятся числа A_k^n, имеют вид

$$Q_i = i\left(\frac{p_1^2}{N_1} + \frac{p_2^2}{N_2}\right),\ i = \overline{0, \min(N_1, N_2)}.$$

Обозначим

$$S = \frac{p_1^2}{N_1} + \frac{p_2^2}{N_2}.$$

Тогда

$$P_i = 1 - iS,\quad Q_i = iS,\ i = \overline{0, \min(N_1, N_2)}\ .$$

Это значит, что $A_k^n = S^{n-k} a_k^n$, а

$$P\{\xi_n = k\} = S^{n-k} a_k^n \prod_{i=0}^{k-1} (1 - iS)\,,\quad k = \overline{1, n},\ n = \overline{1, \min(N_1, N_2)}. \qquad (7)$$

Таким образом, согласно формуле (7),

$$P\{\mu_0(N, n) = k\} = S^{n-N+k} a_{N-k}^n \prod_{i=0}^{N-k-1} (1 - iS).$$

Данное распределение, как и предыдущее, имеет место, если

$n \le \min(N_1, N_2)$.

Замечание 1. Пусть $N_1 < n \le N_2$. Тогда при $i \le N_1$вероятности P_i имеют вид (6), а при $i > N_1$

$$P_i = \ 1 - i\left(\frac{p_1^2}{N_1} + \frac{p_2^2}{N_2}\right) - \frac{1}{N_1}\sum_{j=0}^{i-N_1}\binom{i}{j} p_1^{i+1-j} p_2^j (N_1 - i + j) -$$

$$-\frac{1}{N_2}\sum_{j=0}^{i-N_1-1}\binom{i}{j} p_1^{i-j} p_2^{j+1} (N_2 - j).$$

Замечание 2. Если $n > \min\,(N_1, N_2)$, то при $i > \min\,(N_1, N_2)$ вероятности P_i удобнее вычислять по общей формуле

$$P_i = \sum_{j=\max(0, i-N_1)}^{\min(i, N_2)}\binom{i}{j} p_1^{i-j+1} p_2^j \frac{N_1 - (i-j)}{N_1} \ +$$

$$+\sum_{j=\max\,(0, i-N_2)}^{\min\,(i, N_1)}\binom{i}{j} p_2^{i-j+1} p_1^j \frac{N_2 - (i-j)}{N_1}.$$

4. Числовые характеристики

Если размещение частиц производится в две совокупности ячеек и

$n \le \min(N_1, N_2)$, то математическое ожидание и дисперсия величины ξ_n

имеют соответственно вид:

$$M\xi_n = \frac{1}{S}(1 - (1-S)^n); \qquad (8)$$

$$D\xi_n = \frac{1}{S}(\frac{1}{S} - 1)(1 - 2S)^n + \frac{1}{S}(1-S)^n - \frac{1}{S^2}(1-S)^{2n}. \qquad (9)$$

При этом

$M\mu_0(N, n) = N - \frac{1}{S}(1 - (1-S)^n);\ \ D\mu_0(N, n) = D\xi_n$.

В частном случае, когда $p_1 = p_2 = \frac{1}{2}$, $N_1 = N_2 = \frac{N}{2}$, получим известные результаты для числа пустых ячеек [1, с. 11, 12]. Формулы (8) и (9) следуют из известных результатов для характеристик распределений вида (1) [4].

5. Предельные теоремы

Сформулируем две теоремы, которые получены на основании утверждений, содержащихся в работах [4] и [5].

Теорема 2. Пусть $n, N_1, N_2 \to \infty$, причем $n = o(\max(N_1, N_2))$ и выполнено условие

$$\frac{Sn^2}{2} \to \lambda < \infty.$$

Тогда

$$P\{\xi_n = n - k\} \to \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \geq 0$$

или

$$P\{\mu_0(N, n) = N - n + k\} \to \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k \geq 0.$$

Теорема 3. Пусть $n, N_1, N_2 \to \infty$, причем $n = o(\max(N_1, N_2))$ и выполнено условие

$$Sn^2 \to \infty.$$

Тогда

$$P\left\{\frac{\xi_n - M\xi_n}{\sqrt{D\xi_n}} < x\right\} \to \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in (-\infty, \infty).$$

Использованные источники

1. Колчин В.Ф. Случайные размещения / В.Ф. Колчин, Б.А. Севастьянов, В.П.Чистяков. – М: Наука, 1976. – 288 с.
2. Чистяков В.П. Дискретные предельные распределения в задаче о дробинках с произвольными вероятностями попадания в ящики / В.П. Чистяков. // Матем. заметки. – 1967 –Т.1, №1.– С. 9-16.
3. Ватутин В.А. Предельные теоремы для числа пустых ячеек в равновероятной схеме размещения частиц комплектами / В.А. Ватутин, В.Г. Михайлов.// Теория вероятностей и ее применения. – 1982. – Т.27, №4. – С.684-692.

4. Докин В.Н. 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.
5. Колокольникова Н.А. Предельные теоремы для числа успехов в одной схеме зависимых испытаний / Н.А. Колокольникова. – ВИНИТИ, 25.02.1992. – №649. – В 92 ДЕП. – 25 с.