Доклады Академии наук СССР
1973. Том 209, № 6

УДК 519.24 *МАТЕМАТИКА*

Я. П. ЛУМЕЛЬСКИЙ

СЛУЧАЙНЫЕ БЛУЖДАНИЯ, ОТВЕЧАЮЩИЕ ОБОБЩЕННЫМ УРНОВЫМ СХЕМАМ

(Представлено академиком Ю. В. Линником 28 VI 1972)

Случайные блуждания из начала координат по точкам с неотрицательными целочисленными координатами k-мерного эвклидова пространства, соответствующие выборкам с возвращением (полиномиальные или мультиномиальные блуждания) и выборкам без возвращения (многомерные гипергеометрические блуждания), обладают рядом общих свойств. Так, в [1] указано, что необходимые и достаточные условия полноты плана одни и те же для обеих схем блужданий. Несмещенные достаточные оценки для переходных вероятностей имеют один и тот же вид [2] в случаях полиномиальных и многомерных гипергеометрических блужданий.

Цель настоящей заметки состоит в рассмотрении более общей схемы блужданий, частными случаями которой являются вышеупомянутые блуждания. В этой общей схеме блужданий могут быть получены известные многомерные дискретные распределения.

Ниже будут рассматриваться лишь замкнутые планы первого вхождения [3] для блужданий из начала координат по точкам с неотрицательными целочисленными координатами k-мерного эвклидова пространства. Используем основные понятия и определения для таких планов, следуя работам [3-6].

Пусть вероятность того, что за первый шаг частица, выходящая из начала координат, продвинется на единицу по оси OT_i, равна p_i, $0 < p_i < 1$; $i = 1, \ldots, k$; $\sum\limits_{i=1}^{k} p_i = 1$. Вероятность перехода за один шаг из произвольной точки $B_1(t_1, t_2, \ldots, t_i, \ldots, t_k)$ в проходную точку $B_2(t_1, t_2, \ldots, t_{i+1}, \ldots \ldots, t_k)$ равна $\dfrac{p_i + \alpha t_i}{1 + \alpha\tau(B_1)}$, где $\tau(B_1)$ обозначает сумму координат точки B_1, а α — некоторый скалярный параметр (для неограниченных планов параметр α может быть лишь неотрицательным).

Тогда вероятность перехода из начала координат в любую достижимую точку $D(t_1, t_2, \ldots, t_{k-1}, t_k)$ (проходимую или граничную) имеет следующий вид:

$$P_{\mathrm{p},\alpha}(0, D) = \mathscr{K}(0, D)\,\frac{\prod\limits_{i=1}^{k}\prod\limits_{h=0}^{t_i-1}(p_i + \alpha h)}{\prod\limits_{l=0}^{\tau(D)-1}(1 + \alpha l)}, \tag{1}$$

где $\mathscr{K}(B, C)$ — число различных траекторий в области блужданий R, соединяющих точки B и C. Положив в формуле (1) $\alpha = 0$, получим полиномиальные блуждания. Значения параметров $p_i = N_i / N$, $i = 1, \ldots, k$; $0 < N_i < N$; $\sum\limits_{i=1}^{k} N_i = N$, и $\alpha = -1/N$ приводят к многомерным гипергеометрическим блужданиям [2].

Рассмотренную общую схему блужданий назовем **блужданиями Пойя**, поскольку из нее могут быть получены известные распределения Пойя.

Замкнутый план Пойя $\Pi(G)$ с границей остановки G определяет при каждом α, $\alpha \geqslant 0$, семейство вероятностных мер $\mathfrak{P}_{\mathbf{p}}{}^{G}$, где вероятности перехода в граничные точки Γ, $\Gamma \in G$, задаются формулой (1) и $\sum\limits_{\Gamma \in G} P_{\mathbf{p},\alpha}(0,\Gamma) = 1$.

Граница остановки $\sum\limits_{i=1}^{k} t_i = U$, U — целое положительное число, приводит к семейству многомерных распределений Пойя, у которого вероятности перехода в граничную точку $\Gamma(t_1, t_2, \ldots, t_k)$, $\Gamma \in G$, определяются формулой

$$P_{\mathbf{p},\alpha}(0,\Gamma) = \frac{U!}{\prod\limits_{i=1}^{k} t_i!} \frac{\prod\limits_{i=1}^{k}\prod\limits_{h=0}^{t_i-1}(p_i + \alpha h)}{\prod\limits_{l=0}^{U-1}(1+\alpha l)}. \tag{2}$$

Распределение (2) подробно рассмотрено с указанием имеющейся литературы в ([7]). Частными случаями распределения (2) при вышеуказанных значениях параметров α и $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ являются полиномиальное и многомерное гипергеометрическое распределения.

В предложенной общей схеме блужданий граница остановки инверсного плана $t_1 = m$, m — целое положительное число, задает семейство отрицательных многомерных распределений Пойя вида

$$P_{\mathbf{p},\alpha}(0,\Gamma) = \frac{[\tau(\Gamma)-1]!}{(m-1)!\prod\limits_{i=2}^{k} t_i!} \frac{\prod\limits_{h=0}^{m-1}(p_1+\alpha h)\prod\limits_{i=2}^{k}\prod\limits_{h=0}^{t_i-1}(p_i+\alpha h)}{\prod\limits_{l=0}^{\tau(\Gamma)-1}(1+\alpha l)}. \tag{3}$$

Распределение (3) является обобщением соответствующих полиномиального и многомерного гипергеометрического распределений с той же границей остановки.

Аналогично из (1) можно получить другие распределения вероятностей, подобрав соответствующий замкнутый план Пойя.

К блужданиям Пойя с естественным сужением пространства параметров приводит следующая урновая схема. Пусть урна содержит N шаров, каждый из которых окрашен в один из k цветов, причем имеется N_i шаров i-го цвета, $0 < N_i < N$; $i = 1, \ldots, k$; $\sum\limits_{i=1}^{k} N_i = N$. Процесс случайного выбора шаров из урны будем интерпретировать как блуждание точки из начала координат вдоль осей OT_i, $i = 1, \ldots, k$, в точки с неотрицательными целочисленными координатами. Точка передвигается на единицу по i-й оси, если последовательно извлечен шар i-го цвета. Пусть состав урны изменяется перед извлечением последующего шара следующим образом: если на предшествующем шаге извлечен шар i-го цвета, то в урну добавляется (убирается) a шаров того же цвета. Случаям $a=0$ и $a=-1$ отвечают соответственно выборки с возвращением и без возвращения. При задании замкнутого плана и целого числа a вероятность перехода из начала координат в достижимую точку D равна

$$P_{N,a}(0,D) = \mathscr{K}(0,D)\frac{\prod\limits_{i=1}^{k}\prod\limits_{h=0}^{t_i-1}(N_i+ah)}{\prod\limits_{l=0}^{\tau(D)-1}(N+al)}. \tag{4}$$

Следует заметить, что при конечных N неограниченные планы могут рассматриваться лишь при положительных a. Формула (4) является частным случаем формулы (1), когда $p_i = N_i / N$ и $\alpha = a / N$.

Другой метод получения блужданий Пойя дается следующей теоремой, обобщающей результаты ([7], [8]).

Теорема 1. *Пусть задано семейство распределений, отвечающее некоторому замкнутому полиномиальному плану, причем параметр* $(\mathbf{p})$ *является случайной величиной, имеющей распределение Дирихле*

$$f_{\beta}(\mathbf{p}) = \mathscr{T}(\beta_0) \prod_{i=1}^{k} \frac{p_i^{\beta_i - 1}}{\mathscr{T}(\beta_i)} \tag{5}$$

с параметром $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$, $\beta_i > 0$; $\beta_0 = \sum_{i=1}^{k} \beta_i$.

Тогда безусловные вероятности попадания из начала координат в граничные точки задаются формулой (1) *с параметрами* $p_i' = \beta_i / \beta_0$, $i = 1, \ldots, k$, *и* $\alpha' = 1 / \beta_0$, а $\mathscr{K}(0, \Gamma)$ *определяется исходным полиномиальным планом.*

Следующие теоремы позволяют строить при блужданиях Пойя для вероятностей перехода в достижимые точки несмещенные оценки, выраженные через достаточную статистику.

Теорема 2. *Пусть замкнутый план Пойя* $\Pi(G)$ *определяет семейство распределений* $\mathfrak{P}_{\mathbf{p}}^{G}$. *Достаточной статистикой для этого семейства распределений являются координаты граничной точки.*

Теорема 3. *Пусть задан план Пойя с границей остановки* G, *тогда несмещенная оценка* $\xi(\Gamma)$ *для* $P_{\mathbf{p}\alpha}(0, D)$ — *вероятности перехода из начала координат в достижимую точку* D — *дается формулой*

$$\xi(\Gamma) = \hat{P}_{\mathbf{p}, \alpha}(0, D) = \frac{\mathscr{K}(0, D) \cdot \mathscr{K}(D, \Gamma)}{\mathscr{K}(0, \Gamma)} \tag{6}$$

при $\forall \Gamma$, $\Gamma \in G$.

План $\Pi(G')$, граница остановки G', $\Gamma' \in G'$, которого образована из достижимых точек исходного плана $\Pi(G)$, назовем **вложенным по отношению к плану** $\Pi(G)$. С помощью теоремы 3, используя результаты наблюдений по исходному плану, можно строить несмещенные оценки для распределений вероятностей, определенных вложенными планами. Так, например, построенная на основе распределения (2) несмещенная оценка для вероятности $P_{\mathbf{p}, \alpha}[0, \Gamma'(t_1', \ldots, t_k')]$ семейства распределений Пойя, определенного вложенным планом с границей G': $\sum_{i=1}^{k} t_i' = U'$, $U' \leqslant U$, имеет вид

$$\hat{P}_{\mathbf{p}, \alpha}[0, \Gamma'(t_1', \ldots, t_k')] = \frac{\prod_{i=1}^{k} \binom{t_i}{t_i'}}{\binom{U}{U'}}. \tag{7}$$

Аналогично несмещенная оценка для вероятности $P_{\mathbf{p}, \alpha}[0, \Gamma(m', t_2', \ldots, t_k')]$ в случае вложенного инверсного плана с границей G': $t_1 = m'$, $m' < m$, на основе исходного плана с соответствующим распределением (3) задается формулой

$$\hat{P}_{\mathbf{p}, \alpha}[0, \Gamma'(m', t_2', \ldots, t_k')] = \frac{(m-1)\binom{m-2}{m'-1}\prod_{i=2}^{k}\binom{t_i}{t_i'}}{\left(m + \sum_{i=2}^{k} t_i - 1\right)\binom{m + \sum_{i=2}^{k} t_i - 2}{m' + \sum_{i=2}^{k} t_i' - 1}}. \tag{8}$$

Пусть точка D является достижимой при планах $\Pi(G)$ и $\Pi(G')$ и α задано. Тогда верна

Теорема 4. *Если на основе ограниченных полных планов Пойя — исходного* $\Pi(G)$ *и вложенного* $\Pi(G')$ *— получены несмещенные оценки вида* (6) *для* $P_{\mathbf{p},\alpha}(0, D)$ *соответственно* $\xi(\Gamma)$, $\Gamma \in G$, $\xi(\Gamma')$, $\Gamma' \in G'$, *то дисперсии этих оценок при всех значениях параметров* $\mathbf{p}$ *связаны соотношением*

$$\bigvee_G \xi(\Gamma) \leqslant \bigvee_{G'} \xi(\Gamma'). \tag{9}$$

Аналогичная теорема может быть сформулирована для полных неограниченных замкнутых планов.

Заметим, что у исходного и вложенного планов при всех значениях параметров средние длины траекторий удовлетворяют неравенству

$$E_G \tau(\Gamma) \geqslant E_{G'} \tau(\Gamma'). \tag{10}$$

В связи с этим, при несмещенном оценивании вероятностей перехода в случае полных планов Пойя, нельзя достигнуть при всех значениях параметров одновременно уменьшения среднего числа наблюдений и дисперсии оценки.

Пермский государственный университет
им. А. М. Горького

Поступило
23 VI 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Ю. В. Линник, И. В. Романовский, ДАН, **194**, № 2, 270 (1970). [2] Я. П. Лумельский, Теория вероятностей и ее применения, **14**, № 4, 751 (1969). [3] Р. А. Зайдман, Ю. В. Линник, И. В. Романовский, ДАН, **185**, № 6, 1222 (1969). [4] M. A. Girshick, F. Mosteller, L. J. Savage, Ann. Math. Statist., **17**, № 1, 13 (1946). [5] M. H. De Groot, Ann. Math. Statist., **30**, № 1, 80 (1959). [6] А. М. Каган, Ю. В. Линник, С. Р. Рао, Характеризационные задачи математической статистики, М., 1971. [7] K. G. Janardan, G. P. Patil, Preprints Pennsylvania Univ., № 9 (1969). [8] M. H. de Groot, Optimal Statistical Decisions, N. Y., 1970.