Доклады Академии наук СССР
1972. Том 207, № 6

УДК 519.214 *МАТЕМАТИКА*

Б. Г. ПИТТЕЛЬ

СЛУЧАЙНОЕ РАЗМЕЩЕНИЕ С ОГРАНИЧЕНИЯМИ И ПРИНЦИП МАКСИМУМА ВЗВЕШЕННОЙ ЭНТРОПИИ

(Представлено академиком Ю. В. Линником 7 IV 1972)

Рассмотрим схему последовательности независимых испытаний $A(\nu)$ с n исходами $A_1, \ldots, A_n$, $P(A(\nu) = A_i) = p_i > 0$, $1 \leqslant i \leqslant n$, $\nu = 1, 2, \ldots$ Пусть $m_i(\nu)$ — число i-х исходов в первых ν испытаниях, а $k_i(\nu) = \min(m_i(\nu), b_i)$, где b_i — заданные целые неотрицательные числа. (Предположим, что исход, скажем i-й, ν-го испытания регистрируется в случае, если общее число i-х исходов в первых ν испытаниях не превзошло b_i. Тогда $k_i(\nu)$ — общее число зарегистрированных i-х исходов в этих испытаниях.) Определим случайный момент ν^* обрыва последовательности испытаний как номер испытания, при котором впервые о б щ е е число зарегистрированных исходов становится равным N, где $N > 0$ и фиксировано, т. е.

$$\nu^* = \min(\nu \mid \|k(\nu)\| = N), \quad k(\nu) = (k_1(\nu), \ldots, k_n(\nu)), \quad \left(\|k\| = \sum_{i=1}^{n} k_i\right).$$

(Естественно, что $N \leqslant \sum_{i=1}^{n} b_i$.)

Эту схему можно содержательно истолковать как процесс последовательного дележа мест между N претендентами в n физически сопоставимых состояниях с ограниченными емкостями $b_1, \ldots, b_n$. При этом отношение каждого претендента к выбору состояния характеризуется вектором априорного вероятностного предпочтения $p = (p_1, \ldots, p_n)$, и очередной претендент «генерирует» в соответствии с распределением $(p_1, \ldots, p_n)$ последовательность независимых выборов до тех пор, пока не «наткнется» на состояние, не исчерпанное полностью его предшественниками.

Пусть $b_i / N \to c_i$ при $N \to +\infty$, $i = 1, \ldots, n$. Нас интересуют асимптотические по N свойства случайной последовательности $\{k(\nu)\}$, $1 \leqslant \nu \leqslant \nu^*$.

Введем в рассмотрение следующую задачу А: максимизировать $H(z) = \sum_i z_i \ln \frac{p_i}{z_i}$ при ограничениях $0 \leqslant z_i \leqslant c_i$, $\sum_{i=1}^{n} z_i = 1$.

Поскольку $H(z)$ (взвешенная энтропия) — строго выпуклая непрерывная функция, то задача А имеет и притом единственное решение $\bar{z} = (\bar{z}_1, \ldots, \bar{z}_n)$. Используя основную теорему нелинейного программирования (теорему Куна — Таккера), ограничениям $z_i \leqslant c_i$, $1 \leqslant i \leqslant n$, можно сопоставить вектор $\bar{\sigma} = (\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$ двойственных оценок таких, что $0 \leqslant \bar{\sigma}_i < +\infty$ при $c_i > 0$, $\bar{\sigma}_i = +\infty$ при $c_i = 0$ и $\bar{z}$ максимизирует функцию Лагранжа $L(z) = H(z) + \sum_{i=1}^{n} \bar{\sigma}_i (c_i - z_i)$ при ограничениях $\sum_{i=1}^{n} z_i = 1$, $z_i \geqslant 0$. При этом $\bar{z}$ и $\bar{\sigma}$ связаны условием дополняющей нежесткости: если $\bar{z}_i < c_i$, то $\bar{\sigma}_i = 0$, $1 \leqslant i \leqslant n$. Оказывается, что оценка $\bar{\sigma}_i$ монотонно зависит от c_i / p_i, точнее, неравенство $c_{i_1} / p_{i_1} \geqslant c_{i_2} / p_{i_2}$ эквивалентно неравенству $\bar{\sigma}_{i_1} \leqslant \bar{\sigma}_{i_2}$ при условии $\max(\bar{\sigma}_{i_1} \bar{\sigma}_{i_2}) > 0$.

Свяжем случайные величины ν_i с последовательностью $\{k(\nu)\}$. $1 \leqslant \nu \leqslant \nu^*$:

$$\nu_i = \begin{cases} \min\{\nu \mid k_i(\nu) = b_i\}, & \text{если } k_i(\nu^*) = b_i, \\ \nu^*, & \text{если } k_i(\nu^*) < b_i \end{cases}$$

(если $\nu_i < \nu^*$, то ν_i — номер испытания, когда регистрируется последний i-й исход, или, пользуясь языком схемы дележа, когда впервые исчерпывается емкость i-го состояния).

Теорема. 1) *Для любого* $\varepsilon > 0$

$$\lim_{N \to +\infty} P\left\{\max_{1 \leqslant i \leqslant n} \left| \frac{k_i(\nu^*)}{N} - \bar{z}_i \right| \leqslant \varepsilon \right\} = 1.$$

2) *Если* $\bar{\sigma}_{i_1} > \bar{\sigma}_{i_2}$ *для некоторой пары индексов* $1 \leqslant i_1, i_2 \leqslant n$, *то*

$$\lim_{N \to +\infty} P\{\nu_{i_1} < \nu_{i_2}\} = 1.$$

Нестрого говоря, теорема утверждает, что при больших N доли заполнения $(k_1(\nu^*)/N, \ldots, k_n(\nu^*)/N)$ «почти» максимизируют взвешенную энтропию $H(z)$ при ограничениях на емкости состояний и для «почти» всех реализаций последовательности дележа $\{k(\nu), 1 \leqslant \nu \leqslant \nu^*$, емкости состояний исчерпываются в порядке убывания двойственных оценок $\bar{\sigma}_i$ (т. е. в порядке возрастания отношения c_i/p_i). Заметим, что при $b_i = +\infty$ утверждение теоремы превращается в обычный закон больших чисел для схемы последовательности независимых испытаний, так как $\bar{z} = p$ при $c = +\infty$.

Изложенное выше позволяет считать значение взвешенной энтропии $H(z)$ оценкой качества распределения $(z_1, \ldots, z_n)$, степени его близости к «идеальному» распределению $(p_1, \ldots, p_n)$. В таком случае двойственные оценки $\bar{\sigma}_i$ становятся оценками дефицитности мест в различных состояниях, учитывающих как их априорную популярность, определяемую распределением $(p_1, \ldots, p_n)$, так и их ограниченную емкость. Последнее делает интуитивно понятным второе утверждение теоремы.

Примечание. Интересно, что распределение $(\bar{z}_1, \ldots, \bar{z}_n)$ является также решением задачи Б: минимизировать $F(z) = \sum\limits_i \frac{(z_i - p_i)^2}{p_i}$ при ограничениях $0 \leqslant z_i \leqslant c_i$, $\sum\limits_i z_i = 1$. Однако функционал $F(z)$ тесно связан со взвешенной энтропией $H(z)$, именно, на симплексе $Z = \{z \mid z_i \geqslant 0, \sum\limits_i z_i = 1\}$

$$H(z) = -{}^1\!/_2 F(z) + O(\|z - p\|^3).$$

(На возможность существования естественно квадратичного функционала, минимизируемого распределением $(\bar{z}_1, \ldots, \bar{z}_n)$, указал автору Л. М. Брэгман.)

По поводу доказательства теоремы: описанный случайный механизм построения распределения $(k_1(\nu^*), \ldots, k_n(\nu^*))$ и обычный закон больших чисел делают естественным следующий конечно-шаговый алгоритм построения точки $\bar{z} \in Z$, относительно которой можно ожидать, что в исчезающе малой окрестности ее сосредоточиваются «почти» все нормированные распределения $(k_1(\nu^*)/N, \ldots, k_n(\nu^*)/N)$. Определим точку $\bar{z}^{(1)}$ как первую (при движении из начала координат по лучу $z = \lambda p$, $\lambda \geqslant 0$) точку пересечения с одной из плоскостей $z_i = c_i$, $1 \leqslant i \leqslant n$, $\|z\| = 1$; если $\|\bar{z}^{(1)}\| = 1$, то полагаем $\bar{z} = \bar{z}^{(1)}$; в противном случае вводим множество $I^{(1)} = \{i \mid \bar{z}_i^{(1)} < c_i\}$, вектор $p^{(1)} = (p_1^{(1)}, \ldots, p_n^{(1)})$, $p_i^{(1)} = 0$ при $i \notin I^{(1)}$, $p_i^{(1)} = \left(\sum\limits_{s \in I^{(1)}} p_s\right)^{-1} p_i$ при $i \in I^{(1)}$, и определяем точку $\bar{z}^{(2)}$ как первую (при движении из точки $\bar{z}^{(1)}$ по лучу $z = \bar{z}^{(1)} + \lambda p^{(1)}$, $\lambda \geqslant 0$) точку пересечения с одной из поверхностей $z_i = c_i$, $i \in I^{(1)}$, $\|z\| = 1$, полагая $\bar{z} = \bar{z}^{(2)}$, если $\|\bar{z}^{(2)}\| = 1$, и переходя к следующему шагу в противном случае. За конечное число шагов (не превосходящее n) придем к точке $\bar{z} \in Z$ такой, что $\bar{z}_i \leqslant c_i$, $1 \leqslant i \leqslant n$. Так построенное распределение является решением задачи А (а также задачи Б). Ключом к доказательству этого служит доказательство равенства $z^* = \bar{z}^{(1)} + z^{**}$, где z^* — решение задачи А, а z^{**} максимизирует $\sum_{i \in I^{(1)}} z_i \ln \frac{p_i}{z_i}$ при ограничениях $0 \leqslant z_i \leqslant c_i - \bar{z}_i^{(1)}$, $1 \leqslant i \leqslant n$, $\sum_i z_i = 1 - \sum_i \bar{z}_i^{(1)}$ (очевидно, что в силу этих ограничений $z_i = 0$ при $i \notin I^{(1)}$). В описанном алгоритме плоскости $z_i = c_i$ исключаются в порядке, соответствующем возрастанию отношения c_i / p_i, что, как оказывается, совпадает с убыванием двойственных оценок в задаче А и в задаче Б.

С учетом сказанного доказательство теоремы в остальном связано лишь с трудностями технического характера.

В заключение отметим, что тематически статья примыкает к работам [1-3]. В частности, в [2, 3] построена марковская цепь, являющаяся весьма идеализированной математической моделью стихийной миграции населения в городе, вызванной парными обменами квартирами. В этой модели стационарное расселение максимизирует (асимптотически по числу жителей) взвешенную энтропию при всех ограничениях на возможные способы расселения.

Ленинградское отделение
Центрального экономико-математического института
Академии наук СССР

Поступило
31 III 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Л. М. Брэгман, Журн. вычислит. матем. и матем. физики, 7, № 1 (1967). [2] Б. Г. Питтель, Пробл. передачи информации, 3, № 3 (1967). [3] Б. Г. Питтель, Статистический подход к задаче математического моделирования массового поведения в системах обслуживания, Тез. докл. Симпозиум по моделированию народного хозяйства, Секция экономико-математических моделей зарубежных стран, Новосибирск, 22—27 июля, 1970 г., М., 1970.