УДК 517.11

А.Х. ШЕНЬ

ПОНЯТИЕ (α, β)-СТОХАСТИЧНОСТИ ПО КОЛМОГОРОВУ И ЕГО СВОЙСТВА

(Представлено академиком А.Н. Колмогоровым 11 XI 1982)

Мы исследуем свойства введенного А.Н. Колмогоровым понятия стохастического конечного объекта. Говоря неформально, стохастические объекты — это "элементы общего положения" простых множеств. Приведем точные определения.

В качестве конечных объектов мы будем рассматривать натуральные числа. Говоря об энтропии (= сложности) числа x, будем иметь в виду его простую колмогоровскую энтропию, введенную в [1]. Энтропия числа x будет обозначаться $K(x)$. Нам потребуется говорить также об энтропии конечных множеств натуральных чисел. Для этого мы фиксируем какую-либо естественную нумерацию конечных множеств (например, описанную в [2]) и под энтропией множества будем понимать энтропию его номера. Энтропия множества A будет обозначаться $K(A)$.

Определение (А.Н. Колмогоров). Пусть α, β — натуральные числа. Число x будем называть (α, β)-стохастическим, если существует такое конечное множество $A \subset \mathbf{N}$, что

$$x \in A, \quad K(A) \leqslant \alpha, \quad K(x) \geqslant \log_2 |A| - \beta;$$

здесь через $|A|$ обозначено число элементов множества A.

Первое неравенство (если α невелико) означает, что множество A достаточно просто. Второе (если β невелико) означает, что элемент x является "элементом общего положения" в множестве A. Действительно, если бы x обладал какими-нибудь особенностями, которые свойственны лишь очень малой части Q множества A, то их можно было бы использовать для простого описания x, указав его порядковый номер в списке всех элементов Q, что потребовало бы $\log_2 |Q|$ бит, т.е. много меньше $\log_2 |A|$.

Установим связь понятия (α, β)-стохастичности с основаниями математической статистики. Пусть мы проводим некоторый вероятностный эксперимент, результатом которого a priori может быть любое натуральное число. Пусть результатом этого эксперимента оказалось число x. Зная x, мы хотим восстановить распределение вероятностей P на множестве $\mathbf{N}$ всех натуральных чисел. Разумно требовать, чтобы, во-первых, P имело простое описание, а во-вторых, x было бы "типичным" исходом опыта с распределением вероятностей P. (На практике специфика задачи часто заранее подсказывает возможный вид P и остается выбрать какие-то его параметры; мы, однако, считаем, что единственное имеющееся у нас сведение о P — это полученное значение x.)

Уточним сказанное. В качестве распределений вероятностей будем рассматривать всюду определенные функции $P\colon \mathbf{N} \to \mathbf{Q}$, все значения которых неотрицательны, $P(x) = 0$ для всех x, кроме конечного числа, и $\sum\limits_x P(x) \leqslant 1$. (Мы допускаем возможность $\sum P(x) < 1$, считая, что наш эксперимент может и не дать результата.) Чтобы говорить об энтропии таких функций, фиксируем какую-либо естественную нумерацию их натуральными числами и, говоря об энтропии функции, будем иметь в виду (простую колмогоровскую) энтропию ее номера. Энтропию распределения P будем обозначать $K(P)$. Теперь требование простоты распределения P превращается в требование малости его энтропии. Требование

"типичности" x для распределения P мы уточним так:

$K(x)$ не должно быть много меньше $-\log_2 P(x)$.

Если, например, P приписывает вероятность $1/2^n$ всем числам от 0 до $2^n - 1$, то "типичными" будут те $x \in \{0, \ldots, 2^n - 1\}$, для которых энтропия $K(x)$ близка к n.

Отметим, что $K(x)$ не может сильно превосходить $-\log_2 P(x)$, если распределение P достаточно просто. Именно, при любом x справедливо неравенство

$$K(x) \leqslant -\log_2 P(x) + K(P) + O(\log_2(-\log_2 P(x) + K(P))).$$

В самом деле, пусть $1/2^{k+1} \leqslant P(x) \leqslant 1/2^k$. Рассмотрим множество всех t, для которых $P(t) \geqslant 1/2^{k+1}$. В нем не более 2^{k+1} элементов, и x — один из них. Чтобы задать x, достаточно указать это множество и порядковый номер элемента x в нем. Для указания множества достаточно указать P и число k; указание порядкового номера требует не более $k + 1$ бит. Отсюда и вытекает написанное неравенство. Требование "типичности" x гарантирует, что это неравенство будет близко к равенству (если энтропия P невелика).

Следующее определение выделяет те x, для которых возможно найти распределение P с описанными свойствами.

Определение. Пусть α, β — натуральные числа. Число назовем (α, β)-квазистохастическим, если существует такое распределение P (из описанного класса), что

$$K(P) \leqslant \alpha, \quad K(x) \geqslant -\log_2 P(x) - \beta.$$

Понятия стохастичности и квазистохастичности оказываются весьма близкими. Именно, имеет место

Теорема 1. *Существуют такие константы C_1 и C_2, что для любого числа x:*

а) *если x является (α, β)-стохастическим, то x является $(\alpha + C_1, \beta)$-квазистохастическим;*

б) *если x является (α, β)-квазистохастическим и $x \in \{0, \ldots, 2^n - 1\}$, то x является $(\alpha + C_1 \log_2 n, \beta + C_2)$-стохастическим.*

Эта теорема показывает, что стохастичность и квазистохастичность "совпадают с точностью до $\log_2 n$".

Доказательство. Утверждение а) доказывается легко. Пусть $x \in A$, (энтропия A) $\leqslant \alpha$, $K(x) \geqslant \log_2 |A| - \beta$. Рассмотрим распределение P, приписывающее всем элементам множества A одинаковые вероятности, равные $1/|A|$, и всем остальным числам — нулевые вероятности. Очевидно, энтропия P превосходит энтропию A не более, чем на константу, а $\log_2 |A| = -\log_2 P(x)$. Отсюда и получаем требуемое.

Чуть более сложно доказывается утверждение б). Пусть число x является (α, β)-квазистохастическим. Тогда существует такое распределение P, энтропия которого не превосходит α, а $K(x) \geqslant -\log_2 P(x) - \beta$. Пусть k — такое число, что $2^{-(k+1)} \leqslant P(x) < 2^{-k}$. Тогда $K(x) \geqslant k - \beta$. Из этого неравенства и из неравенства $K(x) \leqslant n + O(1)$ вытекает, что $k \leqslant n + \beta + O(1)$ (эта оценка потребуется нам в дальнейшем). Рассмотрим теперь множество A, состоящее из всех $y \in \mathbf{N}$, для которых $P(y) \geqslant 2^{-(k+1)}$. Чтобы задать A, достаточно указать P и указать k, поэтому энтропия A не превосходит $\alpha + C\log_2(n + \beta)$. Множество A содержит не более 2^{k+1} элементов, поэтому $K(x) \geqslant \log_2 |A| - (\beta + 1)$. Таким образом, x является $(\alpha + C\log_2(n + \beta), \beta + 1)$-стохастическим. Если $\beta \leqslant n$, то утверждение п. б) доказано. Если же $\beta > n$, то любое число от 0 до $2^n - 1$ является $(C\log_2 n, \beta)$-стохастическим (достаточно в качестве A взять множество всех чисел от 0 до $2^n - 1$). Теорема 1 доказана.

Обратимся теперь к вопросу о том, при каких α и β среди чисел от 0 до $2^n - 1$ существуют не (α,β)-стохастические. Ответ на этот вопрос дает

Теорема 2. а) *Существует такое C, что при любом n и любых α и β, для которых $\alpha \geqslant \log_2 n + C$, $\alpha + \beta \geqslant n + 4\log_2 n + C$, все числа от 0 до $2^n - 1$ являются (α, β)-стохастическими.*

б) *Существует такое C, что при любом n и любых α и β, для которых $2\alpha + \beta < n - 6\log_2 n - C$, не все числа от 0 до $2^n - 1$ являются (α, β)-стохастическими.*

Доказательство. а) Пусть сначала $\beta \leqslant n$. Разобьем числа от 0 до $2^n - 1$ на $2^{n-\beta}$ множеств по 2^β элементов в каждом (например, отнеся к i-му множеству числа от $2^\beta i$ до $2^\beta (i + 1)$). Чтобы задать любое из этих множеств, нужно задать n, β и число от 0 до $2^{n-\beta}$, указывающее, каким по счету оно является в нашем разбиении. Поэтому энтропия любого из множеств разбиения не превосходит $2\log_2 n + 2\log_2 \beta + (n - \beta) + C$, т.е. $\leqslant n - \beta + 4\log_2 n + C$ (здесь C – константа, не зависящая от n, α, β.) Если $\alpha + \beta \geqslant n + 4\log_2 n + C$, то энтропия любого из множеств разбиения не превосходит α, поэтому все числа от 0 до $2^n - 1$ будут (α, β)-стохастическими. (Второе неравенство из определения (α, β)-стохастичности выполнено, так как в правой части его стоит $\log_2 2^\beta - \beta = 0$.) Если же $\beta \geqslant n$, то, взяв в качестве A множество $\{0, 1, \ldots, 2^n - 1\}$, мы убедимся, что его энтропия не превосходит $\log_2 n + C$ (и, следовательно, не превосходит α) и все его элементы являются (α, β)-стохастическими.

б) Пусть α фиксировано. Рассмотрим список $A_1, A_2, \ldots, A_s$ всех конечных множеств, энтропия которых не превосходит α. Очевидно, $s \leqslant 2^{\alpha+1}$. Мы хотим оценить энтропию семейства $A_1, A_2, \ldots, A_s$. Чтобы задать это семейство, достаточно указать (помимо α) то из описаний множеств $A_1, A_2, \ldots, A_s$, на обработку которого выбранному способу описания требуется больше всего шагов. Поэтому энтропия указанного семейства не превосходит $\alpha + 2\log_2 \alpha + C_1$, где C_1 – некоторая константа, не зависящая от α. Рассмотрим те из множеств $A_1, A_2, \ldots, A_s$, которые имеют менее $2^{n-\alpha-1}$ элементов. Рассмотрим наименьшее число x, не содержащееся в их объединении. Это число меньше 2^n, так как s не превосходит $2^{\alpha+1}$, а каждое из множеств $A_1, A_2, \ldots, A_s$ имеет менее $2^{n-\alpha-1}$ элементов.

Чтобы задать число x, нужно указать $A_1, A_2, \ldots, A_s, \alpha$ и n, поэтому его энтропия не превосходит

$$\alpha + 2\log_2 \alpha + 2\log_2 \alpha + 2\log_2 n + C',$$

и тем более $\alpha + 6\log_2 n + C'$ (здесь C' – константа, не зависящая от n и α.) Докажем, что если

$$\beta < n - 6\log_2 n - (C' + 1) - 2\alpha,$$

то построенное нами x не является (α, β)-стохастическим. Из этого будет следовать утверждение теоремы при $C = C' + 1$. В самом деле, если x является (α, β)-стохастическим, то $x \in A_i$ и $K(x) \geqslant \log_2 |A_i| - \beta$ при некотором i. Множество A_i должно содержать не менее $2^{n-\alpha-1}$ чисел (иначе x не принадлежало бы ему), поэтому $K(x) \geqslant n - \alpha - 1 - \beta$. Но $K(x) \leqslant \alpha + 6\log_2 n + C'$, откуда $\alpha + 6\log_2 n + C' \geqslant n - \alpha - 1 - \beta$ и $\beta \geqslant n - 6\log_2 n - 2\alpha - 1 - C'$. Теорема 2 доказана.

Утверждение теоремы 2 указывает границу для α/n и β/n, при переходе через которую исчезают последние нестохастические объекты. Эта граница (для случая $\alpha = \beta$) находится где-то между 1/2 и 1/3.

Следующая теорема отвечает на вопрос о доле (α, β)-стохастических чисел среди всех чисел от 0 до $2^n - 1$.

Теорема 3. *Существует такое C, что для всех n и для всех α и β, для которых $\alpha \geqslant C \log_2 n$, количество чисел от 0 до $2^n - 1$, не являющихся*

(α, β)-стохастическими, заключено между

$$[2^{n-2\alpha-\beta-C\log_2 n}] \text{ и } 2^{n-\alpha-\beta+C\log_2 n};$$

$[a]$ есть целая часть числа a.

Доказательство. Оценим сначала количество нестохастических чисел сверху. Как и при доказательстве теоремы 2, разобьем множество чисел от 0 до $2^n - 1$ на 2^p частей по 2^{n-p} чисел в каждой. Энтропия каждой части не превосходит $p + O(\log_2 n)$, поэтому, выбрав $p = \alpha - C\log_2 n$ при подходящем C, можно добиться, чтобы энтропия любой части не превосходила α. При этом все числа, энтропия которых больше $n - p - \beta$, будут (α, β)-стохастическими. Поэтому количество нестохастических чисел не превосходит $2^{n-p-\beta} = 2^{n-\alpha-\beta+C\log_2 n}$. Верхняя оценка получена.

Чтобы получить нижнюю оценку, рассмотрим все множества, энтропия которых не превосходит α, а число элементов не превосходит $2^{n-\alpha-2}$. Энтропия списка всех таких множеств не превосходит $\alpha + O(\log_2 n)$. Объединение всех миожеств этого списка содержит не более половины всех чисел от 0 до $2^n - 1$. Через a_i обозначим i-е (в порядке возрастания) число, не входящее в это объединение (при $i < 2^{n-1}$). В силу сказанного $a_i < 2^n$ при любом $i < 2^{n-1}$. Энтропия a_i не превосходит $\alpha + O(\log_2 n) + O(\log_2 \alpha) + \log_2 i$; (α, β)-стохастическими среди a_i могут быть лишь те, для которых энтропия превосходит $n - 2 - \alpha - \beta$, т.е. $\alpha + O(\log_2 n) + \log_2 i \geqslant n - \alpha - 2 - \beta$ и $\log_2 i \geqslant n - 2\alpha - \beta - O(\log_2 n)$. Поэтому имеется по крайней мере $[2^{n-2\alpha-\beta-O(\log_2 n)}]$ нестохастических чисел. Теорема 3 доказана.

Она показывает, что (с точностью до $\log_2 n$) доля (α, β)-стохастических чисел среди всех чисел от 0 до $2^n - 1$ заключена между $1 - 1/2^{\alpha+\beta}$ и $1 - 1/2^{2\alpha+\beta}$.

С точки зрения рассмотренной статистической интерпретации представляет интерес вопрос о том, какова вероятность появления нестохастических чисел в вероятностном эксперименте. Более точно, пусть P — некоторое распределение вероятностей на множестве чисел от 0 до $2^n - 1$. Что можно сказать о $P(Q)$, где Q — множество всех нестохастических чисел (при данных α и β)? Естественно желать, чтобы $P(Q)$ было мало. Если не требовать ничего от P, то добиться этого нельзя: P, например, может придавать вероятность 1 некоторому нестохастическому числу. Однако если P имеет малую энтропию, то можно получить желаемую оценку.

Теорема 4. *Существует такое C, что для любого распределения вероятностей P, энтропия которого не превосходит α, величина $P(Q)$, где Q — множество всех чисел от 0 до $2^n - 1$, не являющихся $(\alpha + C\log_2 n, \beta)$-стохастическими, не превосходит $2^{-\beta+C\log_2 n}$.*

Доказательство. Применяя теорему 1, можно доказывать утверждение с заменой $(\alpha + C\log_2 n, \beta)$-стохастичности на (α, β)-квазистохастичность. А это делается так. Для всех x, не являющихся (α, β)-квазистохастическими, $K(x) < -\log_2 P(x) - \beta$ или $P(x) < 2^{-K(x)-\beta}$. Отсюда

$$\sum_{x \text{ — не } (\alpha,\beta)\text{-квазистохастическое}} P(x) < 2^{-\beta} \sum_{x \in \{0,\ldots,2^n-1\}} 2^{-K(x)};$$

правая часть не превосходит $n + O(1)$, так как количество тех x, для которых $K(x) = a$, не превосходит 2^a. Отсюда мы и получаем утверждение теоремы 4.

Институт проблем передачи информации
Академии наук СССР, Москва

Поступило
26 XII 1982

ЛИТЕРАТУРА

1. *Колмогоров А.Н.* — Пробл. передачи информации, 1965, т. 1, вып. 1, с. 3–11. 2. *Роджерс Х.* Теория рекурсивных функций и эффективная вычислимость. М.: Мир, 1972. 624 с.