ДОКЛАДЫ АКАДЕМИИ НАУК, 2003, том 391, № 6, с. 1–3

МАТЕМАТИКА

УДК 519.211+510.52

ОБ УТОЧНЕНИИ ОЦЕНОК КОЛМОГОРОВА, ОТНОСЯЩИХСЯ К ДАТЧИКАМ СЛУЧАЙНЫХ ЧИСЕЛ И СЛОЖНОСТНОМУ ОПРЕДЕЛЕНИЮ СЛУЧАЙНОСТИ

© 2003 г. А. Л. Семенов, Ан. А. Мучник

Представлено академиком Ю.И. Журавлевым 04.01.2003 г.

Поступило 10.01.2003 г.

В 1930-е годы А.Н. Колмогоров построил обоснование теории вероятностей с помощью теории меры. Однако не все связанные с обоснованием вопросы были решены. В работе 1963 г. [1] он начал разрабатывать новый подход к проблеме – теорию сложности описания. В этой работе Колмогоров получил верхнюю и нижнюю оценки максимального количества допустимых правил выбора, для которого гарантированно существует датчик случайных чисел, и поставил проблему получения точной по порядку оценки. В настоящем сообщении проблема Колмогорова решена, показано, что его нижняя оценка является (по порядку) точной.

В работе [1] Колмогоров определил понятия правила выбора из конечной двоичной последовательности **t**. Неформальное описание таково (точное определение можно прочитать в [1]). Представим, что имеется набор карт в количестве, равном длине **t**. Цифры последовательности **t** нарисованы на лицевой стороне карт, оборотная сторона у всех карт одинакова. Вначале карты лежат лицевой стороной вниз в том порядке, в котором идут цифры в **t**. Правило решает, какую карту перевернуть и (еще до того, как эта карта перевернута) включить ли цифру, нарисованную на этой карте, в подпоследовательность. При принятии очередного решения правило может учитывать последовательность цифр, нарисованных на ранее перевернутых картах. Подпоследовательность, выбранная правилом r из последовательности **t**, будет обозначаться $r[\mathbf{t}]$.

Иногда полезно рассматривать более узкие классы правил. Монотонные правила (впервые рассмотренные Черчем в 1940 г.) всегда переворачивают карты в их исходном порядке. Неадаптивные правила сразу указывают некоторое множество карт, и подпоследовательность состоит из цифр, нарисованных на этих картах и расположенных в исходном порядке.

О п р е д е л е н и е 1. Пусть $\mathcal{R}_L$ – некоторое множество правил на последовательностях длины L. Последовательность **t** длины L называется (n, ε)-д а т ч и к о м с л у ч а й н ы х ч и с е л д л я. $\mathcal{R}_L$, если каждое правило из $\mathcal{R}_L$ выбирает в **t** подпоследовательность $r[\mathbf{t}]$ с таким свойством:

при условии, что длина подпоследовательности не меньше n, доля нулей в ней отличается от $\frac{1}{2}$ менее чем на ε.

Абсолютная величина разности между $\frac{1}{2}$ и долей нулей в последовательности называется о т к л о н е н и е м (доли нулей).

З а м е ч а н и е. Приводимые далее результаты обобщаются на случай, когда 0 и 1 появляются с частотами, близкими к p и $1 - p$ во всех длинных подпоследовательностях, выбранных простыми правилами.

Для уточнения слов о не слишком сложных правилах выбора сложностью конечного множества назовем двоичный логарифм его мощности. Введем обозначение: $d(n, \varepsilon) \rightleftharpoons 2n\varepsilon^2\log_2 e$.

Т е о р е м а 1 (Колмогоров, 1963). *Рассмотрим произвольные числа L (длина последовательности), $\varepsilon > 0$ (отклонение частоты от вероятности) и $n \geq \varepsilon^{-4}$ (длина выборки). Для любого множества правил $\mathcal{R}_L$, сложность которого меньше $d(n, \varepsilon)(1 - \varepsilon)$, существует (n, ε)-датчик случайных чисел.*

В статье [1] А.Н. Колмогоров только наметил доказательство этой теоремы. Полное его изложение (связанное с некоторыми тонкостями) будет приведено в работе, которую авторы готовят к публикации в журнале “Проблемы передачи информации”.

Следующая теорема дает алгоритмический аналог для теоремы 1 (понятие условной энтропии было введено Колмогоровым в [2]).

Научный совет по комплексной проблеме “Кибернетика”
Российской Академии наук, Москва

Т е о р е м а 1'. *Рассмотрим произвольные числа* L (*длина последовательности*), $\varepsilon > 0$ (*отклонение частоты от вероятности*) *и* $n \geq \varepsilon^{-4}$ (*длина выборки*). *Тогда для множества* $\mathcal{R}_L$, *состоящего из всех правил, условная энтропия которых при известном* L *меньше* $d(n, \varepsilon)(1 - \varepsilon)$, *существует* (n, ε)-*датчик случайных чисел*.

Т е о р е м а 2 (Колмогоров, 1963). *Рассмотрим произвольные числа* L (*длина последовательности*), $\varepsilon \in \left(0, \frac{1}{2}\right)$ *и* $n \in \left[\varepsilon^{-3}, \frac{L}{2}\right]$ (*длина выборки*). *Существует множество неадаптивных правил* $\mathcal{R}_L$, *сложность которого меньше* $4n\varepsilon(1 + 5\varepsilon)$, *для которого не существует* (n, ε)-*датчика случайных чисел*.

Теорема 1 дает нижнюю оценку, а теорема 2 – верхнюю оценку на максимальное число τ, для которого всякому L и всякому множеству правил, имеющему сложность меньше τ, соответствует хотя бы один (n, ε)-датчик случайных чисел длины L. Поскольку $d(n, \varepsilon) = 2n\varepsilon^2 \log_2 e$ гораздо меньше $4n\varepsilon(1 + 5\varepsilon)$ при малых ε, Колмогоров стремился устранить разрыв между степенями ε в нижней и верхней оценках. Оказалось, что нижняя оценка, полученная Колмогоровым, по порядку точна (даже для неадаптивных правил).

Т е о р е м а 3. *Рассмотрим произвольные числа* L (*длина последовательности*), $\varepsilon \in \left(0, \frac{1}{3}\right)$ (*отклонение частоты от вероятности*) *и* $n \in \left[2\varepsilon^{-3}\log_2 L, \frac{L}{2}\right]$ (*длина выборки*). *Существует множество неадаптивных правил* $\mathcal{R}_L$, *сложность которого меньше* $d(n, \varepsilon)\frac{1 + \varepsilon}{1 - n/(L-1)}$, *для которого не существует* (n, ε)-*датчика случайных чисел*.

Существование такого множества правил доказывается, как и в теореме 1, вероятностно. Однако теперь мы рассматриваем распределение вероятностей на правилах и показываем, что событие, "для множества правил $\mathcal{R}_L$ существует (n, ε)-датчик" имеет вероятность меньше единицы.

Нужное множество правил будем искать среди неадаптивных правил, которые выбирают подпоследовательности длины ровно n, т.е. правило задается n-элементным подмножеством множества 1, 2, …, L. Всего таких правил $\binom{L}{n}$; на них мы введем равномерное распределение вероятностей.

Зафиксируем некоторую последовательность **t** длины L и оценим снизу вероятность того, что она не является (n, ε)-датчиком для случайно взятого правила r, т.е. отклонение в $r[\mathbf{t}]$ не меньше ε. Предположим, что нулей в **t** не меньше, чем единиц (противоположный случай рассматривается симметрично).

Рассмотрим ситуацию, когда числа $\frac{L}{2}$ и $n\left(\frac{1}{2} + \varepsilon\right)$ являются целыми; общий случай несложно сводится к этому. Мы хотим оценить снизу вероятность такого отклонения, когда доля нулей в выборке по крайней мере на ε больше, чем доля единиц; очевидно, что эта вероятность минимальна, если нулей и единиц в **t** поровну. Достаточно оценить вероятность отклонения, точно равного ε. Очевидно, эта вероятность равна

$$\frac{\binom{\frac{L}{2}}{\left(\frac{1}{2} - \varepsilon\right)n}\binom{\frac{L}{2}}{\left(\frac{1}{2} + \varepsilon\right)n}}{\binom{L}{n}}.$$

Используя верхнюю и нижнюю оценки на биномиальные коэффициенты, вытекающие из формулы Стирлинга, и оценку на функцию шеноновской энтропии, получим, что искомая вероятность больше e^{-K}, где

$$K = \frac{2n\varepsilon^2}{1 - \frac{n}{L}}\left(1 + \frac{4\varepsilon^2}{3}\right) + \frac{1}{2}(1 + \ln n).$$

Вероятность того, что из фиксированной последовательности одно правило не выберет подпоследовательность с отклонением не менее ε, оказывается меньше $1 - e^{-K}$. Теперь укажем независимо N случайных правил (некоторые из этих правил могут совпасть друг с другом). Вероятность того, что фиксированная последовательность **t** является (n, ε)-датчиком для множества указанных правил, будет меньше

$$(1 - e^{-K})^N < e^{-Ne^{-K}} \qquad (*)$$

(здесь использовано неравенство $\left(1 - \frac{1}{x}\right)^x < e^{-1}$, верное при всех $x > 1$). Умножив величину из правой части неравенства $(*)$ на количество последовательностей длины L, получим строгую верхнюю оценку вероятности того, что для множества указанных правил есть хотя бы один (n, ε)-датчик, а именно

$$2^L e^{-Ne^{-K}} = e^{L\ln 2 - Ne^{-K}},$$

что не превышает единицы для $N = \lceil e^K L \ln 2 \rceil < e^K L$. При этом сложность множества указанных правил меньше

$$\frac{2n\varepsilon^2}{1-\frac{n}{L}}\left(1+\frac{4\varepsilon^2}{3}\right)\log_2 e + 2\log_2 L.$$

Остается убедиться, что последняя величина меньше $d(n,\varepsilon)\frac{1+\varepsilon}{1-n/L}$ при $\varepsilon < \frac{1}{3}$ и $n \geq 2\varepsilon^{-3}\log_2 L$.

Следующая теорема дает алгоритмический аналог для теоремы 3.

Теорема 3'. *Рассмотрим произвольные числа* L (*длина последовательности*), *рациональное* $\varepsilon \in \left(0, \frac{1}{3}\right)$ (*отклонение частоты от вероятности*) *и* $n \in \left[2\varepsilon^{-3}\log_2 L, \frac{L}{2}\right]$ (*длина выборки*). *Для множества* $\mathcal{R}_L(n, \varepsilon)$ *всех неадаптивных правил, у которых условная энтропия при известных* L, n, ε *меньше*

$$d(n,\varepsilon)\frac{1+\varepsilon}{1-n/(L-1)} + C,$$

не существует (n, ε)*-датчика случайных чисел. Здесь* C *– константа, зависящая только от выбора оптимального языка программирования.*

По теореме 3 для некоторого множества неадаптивных правил, имеющего сложность меньше $d(n,\varepsilon)\frac{1+\varepsilon}{1-n/(L-1)}$, не существует (n, ε)-датчика длины L. Покажем, что, зная L, n, ε, множество правил с таким свойством можно построить алгоритмически.

Действительно, по каждому множеству неадаптивных правил $\mathcal{R}_L$ и последовательности $\mathbf{t}$ длины L можно эффективно узнать, является ли $\mathbf{t}(n, \varepsilon)$-датчиком для $\mathcal{R}_L$ (надо применить к $\mathbf{t}$ каждое правило из $\mathcal{R}_L$ и вычислить отклонение). Перебрав все последовательности длины L, можно проверить, существуют ли (n, ε)-датчики для $\mathcal{R}_L$. Перебирая все множества неадаптивных правил данного размера, найдем нужное (если множеств с нужным свойством несколько, возьмем первое в нашем перечислении).

Условная (относительно L, n, ε) энтропия каждого правила из найденного множества не превосходит его сложности плюс длина программы, действие которой описано в предыдущем абзаце. При добавлении правил (всех остальных с энтропией меньше $d(n,\varepsilon)\frac{1+\varepsilon}{1-n/(L-1)} + C$) отсутствие (n, ε)-датчика сохранится.

Авторы признательны Н.К. Верещагину за интересные обсуждения вопросов колмогоровской теории. Большую помощь нам оказал А.В. Чернов при подготовке текста к публикации, за что авторы ему очень благодарны. Основное содержание сообщения было доложено на Колмогоровском семинаре Московского университета весной 2002 г. Мы признательны его участникам за внимание.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты 01–01–00505 02–01–10904).

Примечание при корректуре. В статье [1] Колмогоров определил понятие таблицы (n, ε, p)-случайных чисел для $\mathcal{R}_L$, которое отличается от (n, ε)-датчика случайных чисел из нашего определения 1 тем, что в "длинных" выборках, полученных правилами из $\mathcal{R}_L$, доля вхождений единицы близка не к 1/2, а к p. Через $l(n, \varepsilon)$ Колмогоров обозначил максимальное число l, для которого всякому p, всякому L и всякому множеству правил, имеющему сложность меньше l, соответствует хотя бы одна таблица (n, ε, p)-случайных чисел длины L. Из [1] вытекает, что при достаточно малых ε $\left(\text{например, } \varepsilon < \frac{1}{20}\right)$ и $n \geq \varepsilon^{-4}$

$$(2\log_2 e)n\varepsilon^2(1-\varepsilon) < l(n,\varepsilon) < 4n\varepsilon(1+5\varepsilon).$$

Колмогоров поставил проблему устранения расхождения в степенях ε в нижней и верхней оценках на $l(n, \varepsilon)$. Из нашей теоремы 3 следует неравенство

$$l(n,\varepsilon) < (2\log_2 e)n\varepsilon^2(1+2\varepsilon),$$

выполненное при достаточно малых ε и $n \geq \varepsilon^{-4}$. (Для доказательства в теореме 3 мы полагаем $L = \lfloor 2^{n\varepsilon^3/2} \rfloor$). Таким образом, верхняя оценка на $l(n, \varepsilon)$, как и нижняя, имеет вид $(2\log_2 e)n(\varepsilon^2 + o(\varepsilon^2))$.

СПИСОК ЛИТЕРАТУРЫ

1. *Kolmogorov A.N.* // Ind. J. Stat. Ser. A. 1963. V. 25. Pt. 4. P. 369–376. (Repr. in Theoret Comp. Sci. 1998. V. 207. P. 387–395.

пер. с добавлением автора: Колмогоров А.Н. // Семиотика и информатика. 1982. В. 18. С. 3–13; перепеч. в сб. Колмогоров А.Н. Теория информации и теория алгоритмов. М.: Наука, 1987. С. 204–213.

2. *Колмогоров А.Н.* // Пробл. передачи информации. 1965. Т. 1. № 1. С. 3–11.