Симметризация игральной кости

И. Р. Высоцкий

С дошкольного детства помню настольную игру, в которой были два кубика, которые папа почему-то называл костями, картонное поле, полное приключений, и разноцветные фишки.

Однажды одна из костей бесследно исчезла, и папа предложил слепить игральный кубик из белого хлеба. Разметку нужно было нанести с помощью спичечных головок — они оставляли во влажном сладковатом мякише коричневые углубления. Хлебная кость получилась отличная, хотя и кривоватая, отчего сразу возник вопрос, равны ли шансы у разных очков. К счастью, идея заняться балансировкой на основе собранной статистики не пришла нам в голову. Мы с папой просто решили, что несимметричность кости не нарушает справедливость игры, поскольку игроки в одинаковых условиях.

* * *

Вопрос о симметризации неправильной кости (неравномерного генератора случайных чисел) вновь возник у меня недавно в связи с составлением задач для олимпиады по теории вероятностей. Теперь задача была уже почти формализована: нужно симметризовать генератор алгебраически, не прибегая к помощи напильника или иным физическим воздействиям.

Орёл-орёл (ОО)	p^2
Орёл-решка (ОР)	pq
Решка-орёл (РО)	qp
Решка-решка (РР)	q^2

Простейший случай — перекошенный бинарный генератор, например, в виде гнутой монеты. Пусть вероятности орла и решки у такой монеты равны p и $q = 1 - p$ соответственно, причём обе положительны. Бросая такую монету дважды, получим четыре элементарных исхода (см. таблицу), и два из них равновероятны, каким бы ни было p: $P(\text{ОР}) = P(\text{РО}) = pq$.

Так получается *точная симметризация*. Если выпала комбинация ОР, то будем считать, что случился орёл, а случай РО отождествим с решкой. Если же при двух бросках выпало одно и то же, то скажем, что опыт не удался и нужно опять бросить монету дважды. Если p и q несильно отличаются от 0,5, то метод вполне работает на практике с тем лишь недостатком, что отменять результаты бросков придётся довольно часто.

Этот способ без труда переносится на несимметричный игральный кубик (генератор шести случайных чисел или, для простоты, 6-генератор).

1. Каждая тройка попарно различных граней может реализоваться шестью упорядоченными равновозможными комбинациями. Сопоставим каждой из этих комбинаций какое-то число от 1 до 6. Поступим так для всех возможных троек (всего 20 штук).
2. Бросим кость трижды.
3. Если какое-то число повторилось, то тройка бросков не засчитывается.
4. Если все три результата различны (120 из 216 вариантов), то смотрим, какое из чисел от 1 до 6 сопоставлено выпавшей тройке. Считаем, что именно это число получилось.

Этот алгоритм для кубической кости требует серий по три броска[1)] и предварительного составления целой таблицы соответствия, но зато он в каком-то смысле эффективнее, чем аналогичный алгоритм для монеты, поскольку можно ждать, что три различных числа при трёх бросках не очень несимметричной кости будут случаться чаще, чем две различные стороны монеты при двух бросках.

Понятно, что таким же способом можно точно симметризовать любой неравномерный генератор, дающий 24, 120, 720 случайных чисел и так далее. Чуть труднее точно симметризовать N-генератор, когда N не является факториалом, но и тут найдутся идеи, часто приводящие к весьма запутанным методам, скорость которых вдобавок сильно зависит от N.

* * *

Займёмся теперь неточной симметризацией, т. е. алгоритмом, делающим из исходного неравномерного генератора другой, менее неравномерный. Пусть неточность компенсируется простотой и скоростью.

[1)] Один из участников олимпиады, где была предложена эта задача, решил, что вместо независимых серий по три броска можно бросать кость до тех пор, пока три разных числа не выпадут подряд. Легко убедиться, что в таком случае разные порядки одной и той же тройки неравновероятны, поэтому такой алгоритм не даст точной симметризации.

Пусть G — генератор случайных чисел $1, \dots, N$, где $N \geqslant 2$. Такой генератор можно понимать как алгоритм или некоторое аппаратное устройство. Будем понимать его как распределение

$$G = \begin{pmatrix} 1 & 2 & 3 & \dots & N \\ p_1 & p_2 & p_3 & \dots & p_N \end{pmatrix},$$

где $p_1, p_2, \dots, p_N$ — вероятности значений $1, 2, \dots, N$. Дополнительно потребуем, чтобы выполнялись неравенства $p_k > 0$ при всех $k = 1, \dots, N$ и чтобы не все вероятности равнялись друг другу. Последнее требование формально излишне, но немного упростит рассуждения[2)], и к тому же оно содержательно: нет смысла симметризовать то, что симметризовано до нас. Главный вопрос — как измерить неравномерность (несимметричность) генератора. Сделать это можно разными способами, но все они будут опираться на какую-то меру рассеивания вероятностей. Используем простейшую — размах.

Неравномерностью генератора G_n назовём величину

$$\max_{1 \leqslant k,\, j \leqslant N} (p_k - p_j) = M - m,$$

где M — наибольшая, а m — наименьшая из вероятностей p_k.

Построим следующий алгоритм. Пусть генератор дал некоторое случайное число в пределах от 1 до N. «Запустим» генератор ещё раз и получим ещё одно случайное число. Формально говоря, мы получаем пару независимых случайных величин ξ и η из распределения G. Если $\xi = \eta$, то сочтём, что процедура неудачна, иначе будем считать, что итогом процесса служит число ξ.

Так получается генератор $G^{(1)}$. С помощью такой же процедуры можно перейти от $G^{(1)}$ к генератору $G^{(2)}$ и так далее. Мы покажем, что каждый следующий генератор лучше предыдущего, т. е. менее неравномерен. Неравномерность не просто уменьшается, а стремится к нулю, и поэтому начиная с какого-то момента можно считать, что генератор стал приемлемо равномерным[3)].

Вероятности генератора $G^{(n)}$ выражаются через распределение $G^{(n-1)}$:

$$p_k^{(n)} = \mathrm{P}(\xi = k,\ \eta \neq k \mid \xi \neq \eta) = \frac{p_k^{(n-1)}(1 - p_k^{(n-1)})}{1 - \sum_{k=1}^{N} (p_k^{(n-1)})^2}, \quad \text{где } \xi, \eta \sim G^{(n-1)}. \quad (1)$$

[2)] В частности, можно будет писать только строгие неравенства, а не сопровождать нестрогие каждый раз фразой «равенство достигается, только если все вероятности равны между собой».

[3)] Например, когда систематическое отклонение изучаемой случайной величины от её математического ожидания, обусловленное неравномерностью исходного генератора, становится пренебрежимо мало по сравнению с отклонением, вызванным естественной случайной изменчивостью.

Здесь и далее верхними индексами в скобках будем помечать все величины, относящиеся к соответствующему генератору. Нулевой индекс — номер начального генератора — для простоты опускаем.

Утверждение. *Промежутки, на которых лежат вероятности генераторов $G^{(n)}$, образуют последовательность вложенных отрезков, длины которых стремятся к нулю. А именно, если $m^{(n)}$ и $M^{(n)}$ — соответственно наименьшая и наибольшая из вероятностей $p_k^{(n)}$, то*

1) *для любого $n \in \mathbb{N}$ верна цепочка неравенств*

$$m^{(0)} < m^{(1)} < \ldots < m^{(n)} < \frac{1}{N} < M^{(n)} < \ldots < M^{(1)} < M^{(0)};$$

2) $\lim\limits_{n\to\infty}(M^{(n)} - m^{(n)}) = 0$, *или, что то же самое,* $\lim\limits_{n\to\infty} m^{(n)} = \lim\limits_{n\to\infty} M^{(n)} = \frac{1}{N}$.

Перед доказательством обсудим некоторые подробности и вспомогательные утверждения. Пусть числа $x_1, x_2, \ldots, x_N$ положительны, не все равны между собой и в сумме дают 1. Для краткости обозначим через σ сумму квадратов:

$$\sigma = \sum_{k=1}^{N} x_k^2.$$

Лемма о крайних. *Наименьшее из чисел $x_1, x_2, \ldots, x_N$ строго меньше, а наибольшее строго больше суммы квадратов всех этих чисел*:

$$\min_{k=1,\ldots,N}(x_k) < \sigma < \max_{k=1,\ldots,N}(x_k).$$

Доказательство.

$$\min(x_k) = \min(x_k) \cdot \sum_k x_k < \sigma < \max(x_k) \cdot \sum_k x_k = \max(x_k). \qquad \square$$

Для дальнейших рассуждений удобно ввести в рассмотрение функцию

$$f(x) = \frac{x(1-x)}{1-\sigma},$$

которая зависит от чисел $x_1, x_2, \ldots, x_N$, поскольку в знаменателе участвует сумма их квадратов σ. Функция f обладает рядом занятных свойств, причём некоторые из них очевидны.

1°. $f(x_k) > 0$ для любого x_k, и $\sum_k f(x_k) = 1$. Свойство очевидно.

2°. Если $N = 2$, то для любых положительных чисел x_1 и $x_2 = 1 - x_1$

$$f(x_1) = f(1-x_1) = f(x_2) = \frac{1}{2}.$$

Это очевидное свойство отражает точную симметризацию гнутой монеты, которую мы обсуждали выше.

3°. Если $N \geqslant 3$, то функция f сохраняет порядок чисел: $f(x_k) < f(x_j)$ тогда и только тогда, когда $x_k < x_j$, для любых $j, k = 1, \ldots, N$.

Доказательство свойства 3°. Рассмотрим разность значений в двух точках

$$f(x_k)-f(x_j)=\frac{x_k(1-x_k)-x_j(1-x_j)}{1-\sigma}=\frac{(x_k-x_j)(1-x_k-x_j)}{1-\sigma}.$$

Поскольку $N\geqslant 3$, разность $1-x_k-x_j$ положительна. Знаменатель положителен, так как $\sigma<1$. Следовательно, выражения $f(x_k)-f(x_j)$ и x_k-x_j имеют один и тот же знак. Отсюда следует утверждение 3°. □

4°. Если $N\geqslant 3$ и $x_k<\sigma<x_j$, то $x_k<f(x_k)<f(x_j)<x_j$.

Если принять во внимание свойство 3°, то свойство 4° очевидно. Оно описывает встречное движение: малые числа функция двигает вверх, большие — вниз. Границей между малыми и большими служит σ.

Доказательство утверждения. Вернёмся к генераторам $G^{(n-1)}$ и $G^{(n)}$ и покажем, что при $N\geqslant 3$ для любого натурального n

$$m^{(n-1)}<m^{(n)}<\frac{1}{N}<M^{(n)}<M^{(n-1)}.$$

Функцию $f(x)$ будем рассматривать теперь не для произвольных чисел, а для вероятностей генератора $G^{(n-1)}$:

$$f(x)=f_n(x)=\frac{x(1-x)}{1-\sigma^{(n)}},\quad \text{где}\quad \sigma^{(n)}=\sum_{k=1}^{N}\left(p_k^{(n)}\right)^2.$$

Формула (1) принимает вид

$$p_k^{(n)}=f\left(p_k^{(n-1)}\right).$$

Будем считать, что в генераторе $G^{(n-1)}$ не все вероятности $p_k^{(n-1)}$ равны между собой; это по условию верно для начального генератора G. Из свойства 3° (сохранение порядка) следует, что числа $p_k^{(n)}=f\left(p_k^{(n-1)}\right)$ расположены в том же порядке, что и $p_k^{(n-1)}$. Поэтому $m^{(n)}=f(m^{(n-1)})$ и $M^{(n)}=f(M^{(n-1)})$. Кроме того, поскольку не все $p_k^{(n-1)}$ равны между собой, из леммы о крайних следует, что $m^{(n-1)}<\sigma^{(n-1)}<M^{(n-1)}$. Из свойства 4° (встречное движение) вытекает, что $m^{(n-1)}<m^{(n)}<M^{(n)}<M^{(n-1)}$.

Наибольшая вероятность больше среднего арифметического всех вероятностей, а наименьшая — меньше, поэтому $m^{(n)}<1/N<M^{(n)}$.

Осталось показать, что $(M^{(n)}-m^{(n)})\to 0$ при $n\to\infty$. Для этого оценим отношение неравномерностей двух последовательных генераторов:

$$\frac{M^{(n)}-m^{(n)}}{M^{(n-1)}-m^{(n-1)}}=\frac{f(M^{(n-1)})-f(m^{(n-1)})}{M^{(n-1)}-m^{(n-1)}}=\frac{(M^{(n-1)}-m^{(n-1)})(1-M^{(n-1)}-m^{(n-1)})}{(M^{(n-1)}-m^{(n-1)})(1-\sigma^{(n-1)})}=$$
$$=\frac{1-M^{(n-1)}-m^{(n-1)}}{1-\sigma^{(n-1)}}<\frac{1-\sigma^{(n-1)}-m^{(n-1)}}{1-\sigma^{(n-1)}}.$$

Последняя оценка вытекает из неравенства $M^{(n-1)} > \sigma^{(n-1)}$. Учтём ещё, что $m^{(n-1)} \geqslant m^{(0)} = m$ и что в силу неравенства о средних

$$\sigma^{(n-1)} = \sum_{k=1}^{N} \left(p_k^{(n-1)}\right)^2 > N\left(\frac{1}{N}\sum_{k=1}^{N} p_k^{(n-1)}\right)^2 = \frac{1}{N}.$$

Теперь цепочку оценок можно продолжить:

$$\frac{M^{(n)} - m^{(n)}}{M^{(n-1)} - m^{(n-1)}} < 1 - \frac{m^{(n-1)}}{1 - \sigma^{(n-1)}} < 1 - \frac{m}{1 - 1/N} = 1 - \frac{Nm}{N-1} = \alpha < 1.$$

Таким образом,

$$0 < M^{(n)} - m^{(n)} < \alpha^n (M^{(0)} - m^{(0)}) < \alpha^n,$$

где число $\alpha = 1 - \dfrac{Nm}{N-1} < 1$ не зависит от n. Следовательно,

$$\lim_{n\to\infty} (M^{(n)} - m^{(n)}) = 0.$$

Утверждение доказано. □

Теперь можно лучше оценить скорость сходимости. Из равенства

$$\lim_{n\to\infty} (M^{(n)} - m^{(n)}) = 0$$

следует, что

$$m^{(n)} \xrightarrow[n\to\infty]{} \frac{1}{N}, \quad M^{(n)} \xrightarrow[n\to\infty]{} \frac{1}{N},$$

поэтому

$$\sigma^{(n)} \xrightarrow[n\to\infty]{} \frac{1}{N},$$

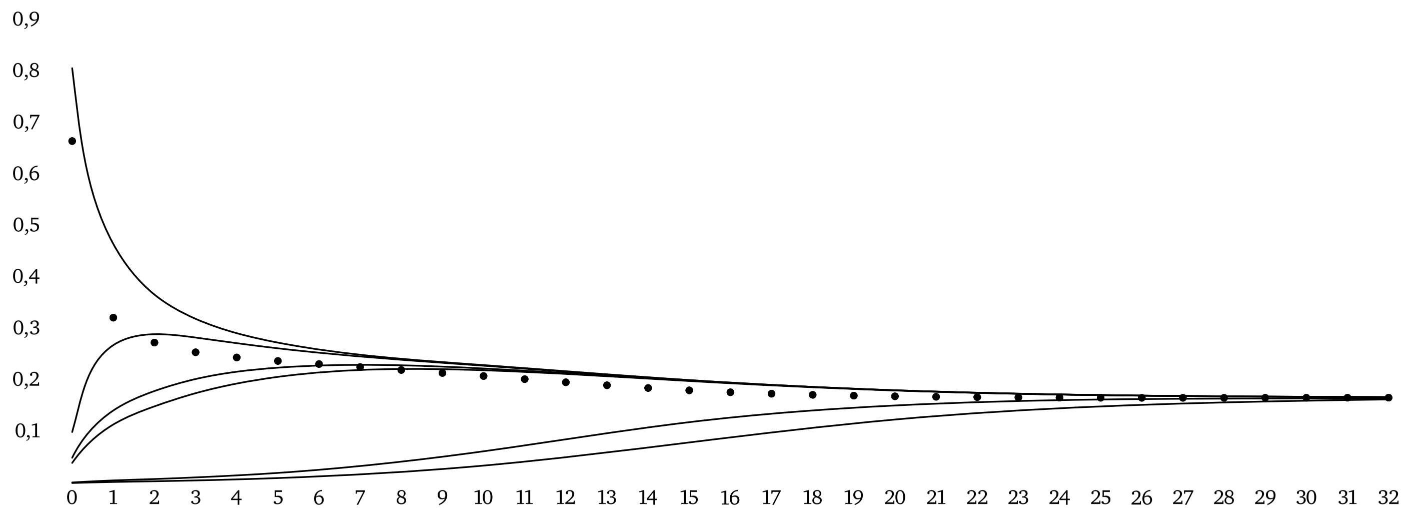
и можно найти предел верхней оценки отношения неравномерностей:

$$\frac{M^{(n)} - m^{(n)}}{M^{(n-1)} - m^{(n-1)}} < 1 - \frac{m^{(n-1)}}{1 - \sigma^{(n-1)}} \xrightarrow[n\to\infty]{} 1 - \frac{1/N}{1 - 1/N} = \frac{N-2}{N-1}.$$

* * *

Описанный алгоритм симметризации на примере шестигранной кости сводится к двум броскам. Первый бросок — определяющий. Он показывает, какое число выпало. Второй бросок — «подтверждающий». Если во второй раз выпало другое число очков, то всё в порядке. Если же выпала та же грань, что и в первый раз, нужно забыть выпавшее число и повторить всё сначала. Алгоритм является аналогом описанного вначале способа симметризации монеты с той лишь разницей, что не даёт точной симметризации. При необходимости алгоритм повторяется. Получается процесс, полностью детерминированный исходным случайным генератором.

	$G^{(0)}$	$G^{(1)}$	$G^{(2)}$	$G^{(3)}$	$G^{(4)}$	$G^{(5)}$	$G^{(10)}$	$G^{(15)}$	$G^{(20)}$	$G^{(30)}$	$G^{(40)}$
p_1	0,1000	0,2689	0,2901	0,2836	0,2727	0,2627	0,2289	0,2008	0,1812	0,1685	0,1669
p_2	0,0020	0,0060	0,0087	0,0119	0,0158	0,0206	0,0626	0,1192	0,1516	0,1653	0,1665
p_3	0,0400	0,1147	0,1499	0,1755	0,1942	0,2072	0,2201	0,1992	0,1808	0,1684	0,1669
p_4	0,0010	0,0030	0,0044	0,0060	0,0080	0,0106	0,0348	0,0798	0,1243	0,1609	0,1660
p_5	0,0500	0,1419	0,1797	0,2030	0,2172	0,2251	0,2235	0,1999	0,1810	0,1685	0,1669
p_6	**0,8070**	0,4654	0,3671	0,3200	0,2921	0,2738	0,2301	0,2010	0,1812	0,1685	0,1669
s	0,6654	0,3223	0,2738	0,2550	0,2449	0,2381	0,2089	0,1810	0,1695	0,1667	0,1667
M	0,8070	0,4654	0,3671	0,3200	0,2921	0,2738	0,2301	0,2010	0,1812	0,1685	0,1669
m	0,0010	0,0030	0,0044	0,0060	0,0080	0,0106	0,0348	0,0798	0,1243	0,1609	0,1660
Нерав-ть $M-m$	0,8060	0,4624	0,3627	0,3139	0,2840	0,2632	0,1953	0,1213	0,0570	0,0076	0,0008
$\frac{M^{(n)}-m^{(n)}}{M^{(n-1)}-m^{(n-1)}}$		0,5737	0,7844	0,8654	0,9047	0,9269	0,9371	0,8884	0,8431	0,8060	0,8007



Как вытекает из последних оценок, процесс сходится тем медленнее, чем больше N. В случае с костью размах вероятностей каждый раз умножается примерно на 0,8. Поэтому процесс может потребовать довольно много шагов, но если неравномерность генератора мала, то уже однократное применение алгоритма существенно улучшает ситуацию.

В заключение — таблица и графики вероятностей (точки соединены непрерывными линиями только для наглядности), которые иллюстрируют процесс последовательной симметризации генератора (см. с. 140)

$$G=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6\\ 0{,}1 & 0{,}002 & 0{,}04 & 0{,}001 & 0{,}05 & 0{,}807\end{pmatrix}.$$

Последовательность значений $\sigma^{(n)}$ показана жирными точками.

Видно, что вероятности на всём протяжении сохраняют свой порядок. При этом малые вероятности (меньшие 1/6) могут сперва расти, стать больше, чем 1/6, а затем убывать, в то время как большие только убывают. Те, что сначала очень малы, всё время остаются меньше σ, а поэтому только растут.

В таблице в строке p_6 первое число больше, чем 0,5 (выделено жирным шрифтом). Однако после первой же итерации вероятность p_6 упала с $p_6^{(0)}=0{,}8070$ до $p_6^{(1)}=0{,}4654$. Это не случайность (см. задачу 1).

Задачи

1. Докажите, что если $N\geqslant 3$, то $M^{(1)}<0{,}5$, т. е. уже после первой симметризации все вероятности оказываются меньше чем 1/2.

2. Докажите, что если не все вероятности p_k равны между собой, то последовательность

$$\sigma^{(n)}=\sum_{k=1}^{N}\left(p_k^{(n)}\right)^2,$$

где $n=0,1,2,\ldots$, монотонно убывает.

Иван Ростиславович Высоцкий, МЦНМО
i_r_vysotsky@hotmail.com