*3.5. ЧТО ТАКОЕ СЛУЧАЙНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ

А. Вводные замечания. Выше в данной главе рассказывалось, как генерировать последовательности

$$\langle U_n \rangle = U_0,\ U_1,\ U_2,\ \ldots \tag{1}$$

действительных чисел в области $0 \le U_n < 1$. Эти последовательности назывались случайными, даже несмотря на то, что они совершенно детерминированные по характеру. Чтобы оправдать использование этой терминологии, мы утверждали, что числа "ведут себя так, как будто они действительно случайны". Такое утверждение может быть удовлетворительным для практических целей (в настоящее время), но это шаг в сторону очень важного методического и теоретического вопроса "Что именно подразумевается под случайным поведением?". Необходимо количественное определение. Нежелательно говорить о понятиях, которых мы на самом деле не понимаем, главным образом потому, что, по-видимому, можно сделать много парадоксальных утверждений, касающихся случайных чисел.

Теория вероятностей и математическая статистика тщательно избегают обсуждения спорных вопросов. Они воздерживаются от безусловных утверждений и вместо этого выражают все в терминах *вероятностей*, связанных с последовательностью случайных событий. Аксиомы теории вероятностей установлены так, что теоретические вероятности можно легко вычислить, но о том, что на самом деле означает "вероятность" или как это понятие можно разумно применить к действительности, ничего не говорится. В книге *Probability, Statistics, and Truth* (New York: Macmillan, 1957), Р. фон Мизес (R. von Mises) подробно обсуждает эту ситуацию и предлагает следующую точку зрения: определение вероятностей зависит от того, какое используется определение случайных последовательностей.

Процитируем здесь несколько утверждений двух авторов, комментировавших эту тему.

> *Д. Г. Лехмер* (D. H. Lehmer) (1951): "Случайная последовательность является смутным понятием, олицетворяющим идею последовательности, в которой каждый член является непредсказуемым для непосвященных и значения которой проходят определенное количество проверок, традиционных у статистиков и отчасти зависящих от пользователей, которым предложена последовательность".
>
> *Д. Н. Франклин* (*J. N. Franklin*) (1962): "Последовательность (1) случайна, если она обладает любыми свойствами, присущими всем бесчисленным последовательностям независимых выборок случайных равномерно распределенных величин".

Утверждение Франклина, по существу, обобщает высказывание Лехмера о том, что последовательность должна удовлетворять *всем* статистическим критериям. Его определение не вполне точное, и позже мы убедимся, что разумное объяснение его утверждения приводит к заключению о том, что не существует такого объекта, как случайная последовательность! Так давайте начнем с менее ограничительного утверждения Лехмера и попытаемся сделать *его* точным. Что нам действительно необходимо — так это сравнительно короткий перечень математических свойств, каждое из которых удовлетворяет нашим интуитивным представлениям о случайной

последовательности. К тому же этого перечня будет вполне достаточно, чтобы согласиться с тем, что *любая* последовательность, удовлетворяющая этим свойствам, является "случайной". В данном разделе рассматривается то, что кажется адекватным точному определению случайности согласно этим критериям, хотя много интересных вопросов остается без ответов.

Пусть u и v — действительные числа, $0 \le u < v \le 1$. Если U — случайная величина, равномерно распределенная между 0 и 1, вероятность того, что $u \le U < v$, равна $v - u$. Например, вероятность, что $\frac{1}{5} \le U < \frac{3}{5}$, равна $\frac{2}{5}$. Как можно выразить это свойство единственного числа U свойством бесконечной последовательности U_0, U_1, U_2, ...? Очевидный ответ — подсчитать, сколько раз U_n находится между u и v, и сказать, что среднее число попаданий в этот интервал равно $v - u$. Наше интуитивное понятие вероятности основано, таким образом, на частоте появления события.

Точнее, пусть $\nu(n)$ — число значений j, $0 \le j < n$, таких, что $u \le U_j < v$, и отношением $\nu(n)/n$ необходимо приблизить значение $v - u$, когда n стремится к бесконечности:

$$\lim_{n\to\infty} \frac{\nu(n)}{n} = v - u. \tag{2}$$

Если это условие выполняется для всех вариантов u и v, говорят, что последовательность *равнораспределена*.

Пусть утверждение $S(n)$ относится и к целому числу n, и к последовательности U_0, U_1, ..., например $S(n)$ может быть приведенным выше утверждением "$u \le U_n < v$". Можно обобщить понятие, используемое в предыдущем разделе, чтобы определить вероятность того, что $S(n)$ справедливо по отношению к некоторой бесконечной последовательности.

Определение А. *Пусть $\nu(n)$ — число значений j, $0 \le j < n$, таких, что $S(j)$ верно. Мы говорим, что $S(n)$ выполняется с вероятностью λ, если предел $\nu(n)/n$, когда n стремится к бесконечности, равен λ.* А именно: $\Pr\big(S(n)\big) = \lambda$, если $\lim_{n\to\infty} \nu(n)/n = \lambda$.

В терминах этой записи последовательность $U_0, U_1, \ldots$ равнораспределена тогда и только тогда, когда $\Pr(u \le U_n < v) = v - u$ для всех действительных чисел u, v при $0 \le u < v \le 1$.

Последовательность может быть равнораспределена, даже если она не случайна. Например, если $U_0, U_1, \ldots$ и $V_0, V_1, \ldots$ — равнораспределенные последовательности, то нетрудно показать, что последовательность

$$W_0, W_1, W_2, W_3, \ldots = \tfrac{1}{2}U_0,\ \tfrac{1}{2}+\tfrac{1}{2}V_0,\ \tfrac{1}{2}U_1,\ \tfrac{1}{2}+\tfrac{1}{2}V_1,\ \ldots \tag{3}$$

также равнораспределена, поскольку подпоследовательность $\frac{1}{2}U_0, \frac{1}{2}U_1, \ldots$ равнораспределена между 0 и $\frac{1}{2}$, в то время как промежуточные члены $\frac{1}{2}+\frac{1}{2}V_0, \frac{1}{2}+\frac{1}{2}V_1, \ldots$ равнораспределены между $\frac{1}{2}$ и 1. Но в последовательности W_j, $j = 0, 1, 2, \ldots$, значения, меньшие $\frac{1}{2}$, всегда следуют за значениями, бóльшими или равными $\frac{1}{2}$ соответственно. Значит, последовательность не случайна согласно любому разумному определению. Необходимы свойства, которые сильнее равнораспределенности.

Естественная возможность обобщить свойство равнораспределенности, которое позволяет развеять сомнения предыдущего раздела, — рассмотреть смежные пары

членов нашей последовательности. Можно потребовать, чтобы последовательность удовлетворяла условиям

$$\Pr(u_1 \le U_n < v_1 \text{ и } u_2 \le U_{n+1} < v_2) = (v_1 - u_1)(v_2 - u_2) \tag{4}$$

для любых членов u_1, v_1, u_2, v_2 при $0 \le u_1 < v_1 \le 1$, $0 \le u_2 < v_2 \le 1$. И вообще, для любого положительного целого k можно потребовать, чтобы наша последовательность была *k-распределенной* в смысле определения B.

Определение B. *Говорят, что последовательность* (1) *будет k-распределенной, если*

$$\Pr(u_1 \le U_n < v_1, \ \ldots, \ u_k \le U_{n+k-1} < v_k) = (v_1 - u_1) \ldots (v_k - u_k) \tag{5}$$

для всех вариантов действительных чисел u_j, v_j при $0 \le u_j < v_j \le 1$ для $1 \le j \le k$.

Равнораспределенная последовательность является 1-распределенной. Заметим, что, если $k > 1$, k-распределенная последовательность всегда является $(k-1)$-распределенной, так как в (5) можно положить $u_k = 0$ и $v_k = 1$. Таким образом, в частности, любая последовательность, о которой известно, что она 4-распределена, должна быть также 3- и 2-распределенной. Можно определить наибольшее k, для которого данная последовательность является k-распределенной, что приведет нас к формулировке более сильного свойства.

Определение C. *Говорят, что последовательность ∞-распределена, если она k-распределена для всех положительных целых k.*

До сих пор мы рассматривали $[0\,..\,1)$-последовательности, т. е. последовательности действительных чисел, лежащих между 0 и 1. Такие же понятия применяются к целочисленным последовательностям. Говорят, что последовательность $\langle X_n \rangle = X_0, X_1, X_2, \ldots$ является *b-ичной последовательностью*, если каждый член последовательности X_n является одним из целых чисел $0, 1, \ldots, b-1$. Таким образом, 2-ичная (бинарная) последовательность — это последовательность нулей и единиц.

Определим также *b-ичное число*, состоящее из k цифр, как строку k целых чисел $x_1 x_2 \ldots x_k$, где $0 \le x_j < b$ для $1 \le j \le k$.

Определение D. *Говорят, что b-ичная последовательность является k-распределенной, если*

$$\Pr(X_n X_{n+1} \ldots X_{n+k-1} = x_1 x_2 \ldots x_k) = 1/b^k \tag{6}$$

для всех b-ичных чисел $x_1 x_2 \ldots x_k$.

Из этого определения ясно, что если $U_0, U_1, \ldots$ является k-распределенной последовательностью $[0\,..\,1)$, то последовательность $\lfloor bU_0 \rfloor, \lfloor bU_1 \rfloor, \ldots$ является k-распределенной b-ичной последовательностью. (Если положить $u_j = x_j/b$, $v_j = (x_j + 1)/b$, $X_n = \lfloor bU_n \rfloor$, то равенство (5) превратится в равенство (6).) Более того, каждая k-распределенная b-ичная последовательность является также $(k-1)$-распределенной, если $k > 1$: мы складываем вероятности для b-ичных чисел $x_1 \ldots x_{k-1}\,0$, $x_1 \ldots x_{k-1}\,1$, $\ldots$, $x_1 \ldots x_{k-1}\,(b-1)$, чтобы получить

$$\Pr(X_n \ldots X_{n+k-2} = x_1 \ldots x_{k-1}) = 1/b^{k-1}$$

(Вероятности для несовместных событий аддитивны; см. упр. 5.) Следовательно, естественно ввести понятие ∞-распределенных b-ичных последовательностей, как в определении C.

Представление положительных действительных чисел в системе с основанием b можно рассматривать как b-ичную последовательность, например π соответствует 10-ичной последовательности 3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, 8, 9, ... Предполагается, что эта последовательность ∞-распределенная, но никто, однако, не в состоянии даже доказать, что она является точно 1-распределенной.

Проанализируем это понятие более подробно для случая, когда k равно миллиону. Бинарная последовательность, являющаяся 1 000 000-распределенной, может содержать серию из миллиона нулей! Аналогично $[0\,..\,1)$-последовательность, являющаяся 1 000 000-распределенной, может содержать миллион последовательных значений, каждое из которых меньше $\frac{1}{2}$. Правда, в среднем такое случается только в одной из $2^{1000000}$ ситуаций, но это действительно *происходит*. Действительно, данный феномен встречается в любой поистине случайной последовательности. Мы используем пока наше интуитивное понятие "истинная случайность". Каждый может легко себе представить, что такая ситуация будет иметь значительные последствия, если такое множество из миллиона "истинно случайных" чисел использовать в эксперименте компьютерного моделирования. Это будет хорошим поводом для того, чтобы вызвать недовольство генератором случайных чисел. Однако при наличии последовательности чисел, которые никогда не пробегают миллион последовательных U_j, меньших $\frac{1}{2}$, последовательность будет не случайной и она не будет подходящим источником чисел для других предполагаемых применений, использующих на входе крайне длинные блоки U_j. В итоге *истинно случайная последовательность будет проявлять локальную "неслучайность"* . Локальная "неслучайность" необходима в одних применениях, но она гибельна в других. Можно сделать вывод, что нет последовательности "случайных" чисел, которую можно было бы использовать в любом случае.

В подобном духе каждый может утверждать, что невозможно решить, будет ли *конечная* последовательность случайной; любая конкретная последовательность ничем не отличается от любой другой. Эти факты действительно представляют собой камни преткновения всякий раз, когда необходимо дать полезное определение случайности, но они не являются истинной причиной смятения. Все еще можно сформулировать такое определение случайности бесконечной последовательности действительных чисел, чтобы соответствующая теория (рассмотренная должным образом) много дала для понимания обычных конечных последовательностей рациональных чисел, которые на самом деле генерируются на компьютере. Более того, ниже в этом разделе будет показано, что существует несколько внушающих доверие определений случайности конечных последовательностей.

B. ∞-распределенные последовательности. Кратко рассмотрим теорию ∞-распределенных последовательностей. Чтобы описать ее адекватно, понадобится немного высшей математики, поэтому в остальной части этого раздела предполагается, что читатель знаком с понятиями, необходимыми для понимания дальнейшего материала.

Во-первых, нужно обобщить определение A, так как фигурирующий в нем предел не существует для всех последовательностей. Определим

$$\overline{\Pr}\big(S(n)\big) = \limsup_{n\to\infty} \frac{\nu(n)}{n}, \qquad \underline{\Pr}\big(S(n)\big) = \liminf_{n\to\infty} \frac{\nu(n)}{n}. \tag{7}$$

Тогда вероятность $\Pr\big(S(n)\big)$ является общим значением $\underline{\Pr}\big(S(n)\big)$ и $\overline{\Pr}\big(S(n)\big)$ (если она существует).

Мы видели, что k-распределенная $[0\,..\,1)$-последовательность приводит к k-распределенной b-ичной последовательности, если U заменить $\lfloor bU \rfloor$. Наша теорема показывает, что обратное утверждение также справедливо.

Теорема A. *Пусть* $\langle U_n \rangle = U_0,\ U_1,\ U_2,\ \ldots$ — $[0\,..\,1)$*-последовательность. Если последовательность*

$$\langle \lfloor b_j U_n \rfloor \rangle = \lfloor b_j U_0 \rfloor,\ \lfloor b_j U_1 \rfloor,\ \lfloor b_j U_2 \rfloor,\ \ldots$$

является k*-распределенной* b_j*-ичной последовательностью для любого* b_j *из бесконечной последовательности целых чисел* $1 < b_1 < b_2 < b_3 < \cdots$, *то исходная последовательность* $\langle U_n \rangle$ k*-распределенная.*

В качестве примера предположим, что $b_j = 2^j$. Последовательность $\lfloor 2^j U_0 \rfloor$, $\lfloor 2^j U_1 \rfloor, \ldots$ является, по существу, последовательностью первых j двоичных разрядов бинарного представления $U_0, U_1, \ldots$. Если все эти последовательности целых чисел k-распределены в смысле определения D, то последовательность действительных чисел $U_0, U_1, \ldots$ также должна быть k-распределенной в смысле определения B.

Доказательство. Если последовательность $\lfloor bU_0 \rfloor, \lfloor bU_1 \rfloor, \ldots$ k-распределена, то с помощью сложения вероятностей получим, что (5) выполняется всякий раз, когда каждое u_j и v_j — рациональные числа со знаменателем b. Предположим, что u_j, v_j — любые действительные числа, и пусть u'_j, v'_j — рациональные числа со знаменателем b, такие, что

$$u'_j \le u_j < u'_j + 1/b, \qquad v'_j \le v_j < v'_j + 1/b.$$

Пусть $S(n)$ — утверждение, что $u_1 \le U_n < v_1, \ldots, u_k \le U_{n+k-1} < v_k$. Получим

$$\begin{aligned}\overline{\Pr}\big(S(n)\big) &\le \Pr\Big(u'_1 \le U_n < v'_1 + \frac{1}{b},\ \ldots,\ u'_k \le U_{n+k-1} < v'_k + \frac{1}{b}\Big) \\ &= \Big(v'_1 - u'_1 + \frac{1}{b}\Big) \ldots \Big(v'_k - u'_k + \frac{1}{b}\Big); \\ \underline{\Pr}\big(S(n)\big) &\ge \Pr\Big(u'_1 + \frac{1}{b} \le U_n < v'_1,\ \ldots,\ u'_k + \frac{1}{b} \le U_{n+k-1} < v'_k\Big) \\ &= \Big(v'_1 - u'_1 - \frac{1}{b}\Big) \ldots \Big(v'_k - u'_k - \frac{1}{b}\Big).\end{aligned}$$

Сейчас $\big|(v'_j - u'_j \pm 1/b) - (v_j - u_j)\big| \le 2/b$. Так как наше неравенство выполняется для всех $b = b_j$ и так как $b_j \to \infty$ при $j \to \infty$, получим

$$(v_1 - u_1) \ldots (v_k - u_k) \le \underline{\Pr}\big(S(n)\big) \le \overline{\Pr}\big(S(n)\big) \le (v_1 - u_1) \ldots (v_k - u_k). \quad \blacksquare$$

Следующая теорема является основным орудием для доказательства различных утверждений о k-распределенных последовательностях.

Теорема В. *Предположим, что* $\langle U_n \rangle$ — *k-распределенная* $[0\,..\,1)$*-последовательность, и пусть* $f(x_1, x_2, \ldots, x_k)$ — *интегрируемая по Риману функция k переменных. Тогда*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0\le j<n} f(U_j, U_{j+1}, \ldots, U_{j+k-1}) = \int_0^1 \cdots \int_0^1 f(x_1, x_2, \ldots, x_k)\, dx_1 \ldots dx_k. \tag{8}$$

Доказательство. Из определения k-распределенной последовательности следует, что этот результат справедлив в частном случае, когда

$$f(x_1, \ldots, x_k) = [u_1 \le x_1 < v_1,\ \ldots,\ u_k \le x_k < v_k], \tag{9}$$

для некоторых постоянных u_1, v_1, $\ldots$, u_k, v_k. Более того, (8) выполняется каждый раз, когда $f = a_1 f_1 + a_2 f_2 + \cdots + a_m f_m$ и каждая функция f_j является функцией вида (9). Другими словами, (8) выполняется каждый раз, когда f является "ступенчатой функцией", полученной путем разбиения единичного k-мерного куба на подкубы, грани которых параллельны оси координат, и f принимает постоянные значения на каждом подкубе.

Пусть f — любая интегрируемая по Риману функция. Если ϵ — любое положительное число, то (по определению интегрируемости по Риману) получим, что существуют ступенчатые функции $\underline{f}$ и $\overline{f}$, такие, что $\underline{f}(x_1, \ldots, x_k) \le f(x_1, \ldots, x_k) \le \overline{f}(x_1, \ldots, x_k)$, и такие, что разность интегралов $\underline{f}$ и $\overline{f}$ меньше ϵ. Поэтому (8) выполняется для $\underline{f}$ и $\overline{f}$. И поскольку

$$\begin{aligned} \frac{1}{n} \sum_{0\le j<n} \underline{f}(U_j, \ldots, U_{j+k-1}) &\le \frac{1}{n} \sum_{0\le j<n} f(U_j, \ldots, U_{j+k-1}) \\ &\le \frac{1}{n} \sum_{0\le j<n} \overline{f}(U_j, \ldots, U_{j+k-1}), \end{aligned}$$

можно заключить, что (8) верно также для f. ∎

Теорема В может применяться, например, в *критерии перестановок* из раздела 3.3.2. Пусть $(p_1, p_2, \ldots, p_k)$ — любая перестановка чисел $\{1, 2, \ldots, k\}$. Покажем, что

$$\Pr(U_{n+p_1-1} < U_{n+p_2-1} < \cdots < U_{n+p_k-1}) = 1/k!. \tag{10}$$

Для доказательства предположим, что последовательность $\langle U_n \rangle$ k-распределена и пусть

$$f(x_1, \ldots, x_k) = [x_{p_1} < x_{p_2} < \cdots < x_{p_k}].$$

Имеем

$$\begin{aligned} &\Pr(U_{n+p_1-1} < U_{n+p_2-1} < \cdots < U_{n+p_k-1}) \\ &\quad = \int_0^1 \cdots \int_0^1 f(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \\ &\quad = \int_0^1 dx_{p_k} \int_0^{x_{p_k}} \cdots \int_0^{x_{p_3}} dx_{p_2} \int_0^{x_{p_2}} dx_{p_1} = \frac{1}{k!}. \end{aligned}$$

Следствие Р. *Если* $[0\,..\,1)$*-последовательность k-распределена, то она удовлетворяет критерию перестановок порядка k в смысле равенства* (10). ∎

Также можно показать, что последовательность удовлетворяет *критерию сериальной корреляции.*

Следствие S. *Если* $[0..1)$*-последовательность* $(k+1)$*-распределена, то коэффициент сериальной корреляции между* U_n *и* U_{n+k} *стремится к нулю:*

$$\lim_{n\to\infty} \frac{\frac{1}{n}\sum U_jU_{j+k} - \left(\frac{1}{n}\sum U_j\right)\left(\frac{1}{n}\sum U_{j+k}\right)}{\sqrt{\left(\frac{1}{n}\sum U_j^2 - \left(\frac{1}{n}\sum U_j\right)^2\right)\left(\frac{1}{n}\sum U_{j+k}^2 - \left(\frac{1}{n}\sum U_{j+k}\right)^2\right)}} = 0.$$

(Все суммирования здесь выполняются по $0 \le j < n$.)

Доказательство. По теореме B значения

$$\frac{1}{n}\sum U_jU_{j+k}, \qquad \frac{1}{n}\sum U_j^2, \qquad \frac{1}{n}\sum U_{j+k}^2, \qquad \frac{1}{n}\sum U_j, \qquad \frac{1}{n}\sum U_{j+k}$$

стремятся соответственно к пределу $\frac{1}{4}, \frac{1}{3}, \frac{1}{3}, \frac{1}{2}, \frac{1}{2}$ при $n \to \infty$. ▮

Рассмотрим несколько более общие свойства распределений последовательностей. Мы определяли понятие k-распределения, рассматривая все смежные строки размерности k, например последовательность является 2-распределенной тогда и только тогда, когда точки

$$(U_0,U_1),\ (U_1,U_2),\ (U_2,U_3),\ (U_3,U_4),\ (U_4,U_5),\ \ldots$$

равнораспределены в единичном квадрате. Это вполне возможно несмотря на то, что пары точек (U_1,U_2), (U_3,U_4), (U_5,U_6), ... могут быть *не* равнораспределенными. Если в некоторой области не хватает точек (U_{2n-1},U_{2n}), их можно компенсировать другими точками: (U_{2n},U_{2n+1}). Например, периодическая бинарная последовательность

$$\langle X_n\rangle = 0,0,0,1,\ 0,0,0,1,\ 1,1,0,1,\ 1,1,0,1,\ 0,0,0,1,\ \ldots \tag{11}$$

с периодом длины 16 будет 3-распределенной; однако последовательность четных элементов $\langle X_{2n}\rangle$ = 0, 0, 0, 0, 1, 0, 1, 0, ... имеет в три раза больше нулей, чем единиц, тогда как подпоследовательность нечетных элементов $\langle X_{2n+1}\rangle$ = 0, 1, 0, 1, 1, 1, 1, 1, ... имеет в три раза больше единиц, чем нулей.

Предположим, что последовательность $\langle U_n\rangle$ является ∞-распределенной. Пример (11) показывает, что подпоследовательность чередующихся членов $\langle U_{2n}\rangle = U_0$, U_2, U_4, U_6, ... не обязана быть ∞-распределенной или даже 1-распределенной. Но $\langle U_{2n}\rangle$ на самом деле является ∞-распределенной и многое еще будет верным.

Определение E. $[0..1)$*-последовательность* $\langle U_n\rangle$ *называют* (m,k)*-распределенной, если*

$$\Pr(u_1 \le U_{mn+j} < v_1,\ u_2 \le U_{mn+j+1} < v_2,\ \ldots,\ u_k \le U_{mn+j+k-1} < v_k) = (v_1-u_1)\ldots(v_k-u_k)$$

для любого выбора действительных чисел u_r, v_r *при* $0 \le u_r < v_r \le 1$ *для* $1 \le r \le k$ *и для всех целых* j *при* $0 \le j < m$.

Поэтому k-распределенная последовательность является частным случаем определения E при $m = 1$; случай, когда $m = 2$, означает, что строки размерности k,

начинающиеся с четного номера, должны иметь такую же плотность, как и строки размерности k, начинающиеся с нечетного номера, и т. д.

Следующие свойства определения E очевидны:

$$(m,k)\text{-распределенная последовательность является } (m,\kappa)\text{-распределенной для } 1 \le \kappa \le k; \tag{12}$$

$$(m,k)\text{-распределенная последовательность является } (d,k)\text{-распределенной для всех делителей } d \text{ числа } m. \tag{13}$$

(См. упр. 8.) Также можно определить понятие (m,k)-распределенности b-ичной последовательности, как в определении D, и доказать теорему A, остающуюся верной для (m,k)-распределенных последовательностей.

Следующая теорема, которая во многих отношениях является, скорее, сюрпризом, показывает, что свойства, присущие ∞-распределению, действительно очень строгие, более строгие, чем мы представляли себе, когда впервые рассматривали определение этого понятия.

Теорема C. (Иван Нивен (Ivan Niven) и Г. С. Цукерман (H. S. Zuckerman).) *∞-распределенная последовательность является (m,k)-распределенной для всех положительных m и k.*

Доказательство. Достаточно доказать теорему для b-ичной последовательности, используя обобщение только что упомянутой теоремы A. Более того, можно предположить, что $m = k$, так как (12) и (13) утверждают, что последовательность (m,k)-распределена, если она (mk,mk)-распределена.

Таким образом, докажем, что *любая ∞-распределенная b-ичная последовательность $X_0, X_1, \ldots$ является (m,m)-распределенной для всех положительных целых m.* Наше доказательство — это упрощенная версия первоначального доказательства, приведенного в статье Niven and Zuckerman, *Pacific J. Math.* **1** (1951), 103–109.

Ключевая используемая идея является важным методом, применяемым во многих ситуациях в математике: "Если и сумма m величин, и сумма их квадратов согласуются с гипотезой, что m величин равны, то гипотеза верна". В строгом виде этот принцип можно сформулировать так.

Лемма E. *Заданы m последовательностей чисел $\langle y_{jn}\rangle = y_{j0}, y_{j1}, \ldots$ для $1 \le j \le m$. Предположим, что*

$$\begin{aligned} \lim_{n\to\infty} (y_{1n} + y_{2n} + \cdots + y_{mn}) &= m\alpha, \\ \limsup_{n\to\infty} (y_{1n}^2 + y_{2n}^2 + \cdots + y_{mn}^2) &\le m\alpha^2. \end{aligned} \tag{14}$$

Тогда для каждого j существует $\lim_{n\to\infty} y_{jn}$ и он равен α.

Невероятно простое доказательство этой леммы приведено в упр. 9. ▮

Продолжим доказательство теоремы C. Пусть $x = x_1x_2\ldots x_m$ — b-ичное число; скажем, что x *встречается* на позиции p, если $X_{p-m+1}X_{p-m+2}\ldots X_p = x$. Пусть $\nu_j(n)$ — число x, находящихся на позиции p, когда $p < n$ и $p \bmod m = j$. Пусть

$y_{jn} = \nu_j(n)/n$ и нужно доказать, что

$$\lim_{n\to\infty} y_{jn} = \frac{1}{mb^m}. \tag{15}$$

Во-первых, известно, что

$$\lim_{n\to\infty} (y_{0n} + y_{1n} + \cdots + y_{(m-1)n}) = \frac{1}{b^m}, \tag{16}$$

так как последовательность m-распределена. Согласно лемме E и равенству (16) теорема доказана, если показать, что

$$\limsup_{n\to\infty} (y_{0n}^2 + y_{1n}^2 + \cdots + y_{(m-1)n}^2) \le \frac{1}{mb^{2m}}. \tag{17}$$

Однако это неравенство не очевидно; необходимо несколько деликатных маневров, прежде чем можно будет его доказать. Пусть q кратно m. Рассмотрим

$$C(n) = \sum_{0\le j<m} \binom{\nu_j(n) - \nu_j(n-q)}{2}. \tag{18}$$

Это число появлений пар x-в на позициях p_1 и p_2, для которых $n - q \le p_1 < p_2 < n$ и $p_2 - p_1$ кратно m. Рассмотрим сумму

$$S_N = \sum_{n=1}^{N+q} C(n). \tag{19}$$

Каждое появление пары x-в, встречающейся на позициях p_1 и p_2 с $p_1 < p_2 < p_1 + q$, где $p_2 - p_1$ кратно m и $p_1 \le N$, учитывается точно $p_1 + q - p_2$ раз в общей сумме S_N (т. е. когда $p_2 < n \le p_1 + q$), а такие пары, которые появляются на позициях p_1 и p_2 с $N < p_1 < p_2 < N + q$, учитываются точно $N + q - p_2$ раз.

Пусть $d_t(n)$ — число пар x, встречающихся на позициях p_1 и p_2 с $p_1 + t = p_2 < n$. Приведенный выше анализ показывает, что

$$\sum_{0<t<q/m} (q - mt) d_{mt}(N + q) \ge S_N \ge \sum_{0<t<q/m} (q - mt) d_{mt}(N). \tag{20}$$

Так как начальная последовательность является q-распределенной, то

$$\lim_{N\to\infty} \frac{1}{N} d_{mt}(N) = \frac{1}{b^{2m}} \tag{21}$$

для всех t, $0 < t < q/m$, и, следовательно, из (20) получаем

$$\lim_{N\to\infty} \frac{S_N}{N} = \sum_{0<t<q/m} \frac{q - mt}{b^{2m}} = \frac{q(q-m)}{2mb^{2m}}. \tag{22}$$

Из этих равенств после нескольких преобразований получим утверждение теоремы.

По определению

$$2S_N = \sum_{n=1}^{N+q} \sum_{0\le j<m} \big((\nu_j(n) - \nu_j(n-q))^2 - (\nu_j(n) - \nu_j(n-q))\big),$$

и можно удалить не возведенные в квадрат члены и, применяя (16), получить

$$\lim_{N\to\infty}\frac{T_N}{N}=\frac{q(q-m)}{mb^{2m}}+\frac{q}{b^m},\tag{23}$$

где

$$T_N=\sum_{n=1}^{N+q}\sum_{0\le j<m}\big(\nu_j(n)-\nu_j(n-q)\big)^2.$$

Используя неравенство

$$\frac{1}{r}\Big(\sum_{j=1}^{r}a_j\Big)^2\le\sum_{j=1}^{r}a_j^2$$

(см. упр. 1.2.3–30), находим, что

$$\limsup_{N\to\infty}\sum_{0\le j<m}\frac{1}{N(N+q)}\Big(\sum_{n=1}^{N+q}\big(\nu_j(n)-\nu_j(n-q)\big)\Big)^2\le\frac{q(q-m)}{mb^{2m}}+\frac{q}{b^m}.\tag{24}$$

Также получим

$$q\,\nu_j(N)\le\sum_{N<n\le N+q}\nu_j(n)=\sum_{n=1}^{N+q}\big(\nu_j(n)-\nu_j(n-q)\big)\le q\nu_j(N+q).$$

Подставив это неравенство в (24), получим

$$\limsup_{N\to\infty}\sum_{0\le j<m}\Big(\frac{\nu_j(N)}{N}\Big)^2\le\frac{q-m}{qmb^{2m}}+\frac{1}{qb^m}.\tag{25}$$

Данная формула справедлива всякий раз, когда q кратно m, и если мы устремим $q\to\infty$, то получим (17), что и завершает доказательство.

Более простое доказательство можно найти у Дж. В. С. Касселя (J. W. S. Cassels, *Pacific J. Math.* **2** (1952), 555–557). ▮

В упр. 29 и 30 иллюстрируется нетривиальность этой теоремы и тот факт, что q-распределенная последовательность имеет вероятности, отклоняющиеся от истинных вероятностей (m,m)-распределения, по существу, не более чем на $1/\sqrt{q}$ (см. (25)). Для доказательства теоремы необходима гипотеза о ∞-распределении последовательности.

Используя теорему С, можно доказать, что ∞-распределенная последовательность проходит критерий серий, критерий "максимум-t", критерий конфликтов, критерий промежутков между днями рождений и критерий подпоследовательностей, о которых упоминалось в разделе 3.3.2. Нетрудно показать, что она также удовлетворяет критерию интервалов, покер-критерию и критерию монотонности (см. упр. 12–14). Критерий собирания купонов является значительно более трудным, но и его последовательность проходит (см. упр. 15 и 16).

Существование ∞-распределенной последовательности достаточно простого вида гарантирует следующая теорема.

Теорема F. (Дж. Н. Франклин (J. N. Franklin).) $[0\,..\,1)$-*последовательность* U_0, U_1, U_2, ... *с*

$$U_n = \theta^n \bmod 1 \tag{26}$$

является ∞-*распределенной для почти всех действительных чисел* $\theta > 1$. *Другими словами, множество*

$$\{\theta \mid \theta > 1 \text{ и (26) не } \infty\text{-распределено}\}$$

имеет меру нуль.

Доказательство этой теоремы и некоторые обобщения приведены в *Math. Comp.* **17** (1963), 28–59. ▮

Франклин показал, что θ должно быть трансцендентным числом для того, чтобы (26) была ∞-распределенной. Раньше, в 60-е годы, степени $\langle \pi^n \bmod 1 \rangle$ получали в результате трудоемких вычислений, использующих многократную точность для $n \leq 10000$, и 35 самых старших двоичных разрядов этих чисел, оставленных в файле на диске, успешно использовались как источник равномерно распределенных случайных чисел. Согласно теореме F вероятность того, что степени $\langle \pi^n \bmod 1 \rangle$ ∞-распределены, равна 1, однако существует несчетное множество действительных чисел, поэтому теорема не дает информации о том, действительно ли последовательность для π имеет ∞-распределение. Можно совершенно спокойно держать пари, что никто при нашей жизни не *докажет*, что именно данная последовательность *не* является ∞-распределенной, хотя это и возможно. Руководствуясь такими соображениями, он может законно удивиться, если окажется, что существует *какая-либо* ∞-распределенная последовательность: *существует ли алгоритм вычисления действительных чисел* U_n *для всех* $n \geq 0$, *такой, что последовательность* $\langle U_n \rangle$ ∞-*распределена?* Ответ — "Да", как показано, например, в работе D. E. Knuth, *BIT* **5** (1965), 246–250. Построенная последовательность полностью состоит из рациональных чисел. На самом деле каждое число U_n имеет ограниченное представление в двоичной системе счисления. Другое построение определенной ∞-распределенной последовательности, отчасти более сложное, чем построение предыдущей последовательности, вытекает из теоремы W, приведенной ниже. (См. также Н. М. Коробов, *Изв. Акад. наук СССР* **20** (1956), 649–660.)

C. Эквивалентно ли понятие ∞-распределенности понятию случайности? Принимая во внимание все теоретические результаты относительно ∞-распределенных последовательностей, можно быть уверенным в одном: понятие "∞-распределенная последовательность" является важным в математике. Кроме того, кажется очевидным, что интуитивное понятие случайности можно формализовать следующим образом.

Определение R1. $[0\,..\,1)$-*последовательность называется случайной, если она является* ∞-*распределенной.*

Мы видели, что последовательности, удовлетворяющие этому определению, удовлетворяют всем статистическим критериям раздела 3.3.2 и многим другим.

Попытаемся объективно критиковать это определение. Прежде всего, всякая ли "поистине случайная" последовательность ∞-распределена? Существует бесконечное число последовательностей U_0, U_1, ... действительных чисел между нулем и единицей. Если генератор истинно случайных чисел выдает значения U_0, U_1, ..., то любую из возможных последовательностей можно рассматривать как в равной степени подходящую и некоторые из последовательностей (на самом деле бессчетное количество) даже не равнораспределены. С другой стороны, используя любое разумное определение вероятности на этом пространстве всех возможных последовательностей, можно прийти к заключению, что случайная последовательность является ∞-распределенной *с вероятностью 1*. Итак, получена следующая формализация определения случайности Франклина, приведенного в начале этого раздела.

Определение R2. $[0\,..\,1)$-*последовательность* $\langle U_n \rangle$ *называется случайной, если всякий раз, когда* P *является таким свойством, что* $P(\langle V_n \rangle)$ *выполняется с вероятностью 1 для последовательности* $\langle V_n \rangle$ *независимых случайных равномерно распределенных величин,* $P(\langle U_n \rangle)$ *также выполняется.*

Можно ли предположить, что определение R1 эквивалентно определению R2? Давайте выдвинем возможные возражения против определения R1 и проанализируем их.

Во-первых, определение R1 распространяется только на предельные свойства последовательностей при $n \to \infty$. Существуют ∞-распределенные последовательности, в которых первый миллион элементов — нули. Можно ли такие последовательности рассматривать как случайные?

Это возражение не очень существенно. Если ϵ — любое положительное число, то нет причины каждому элементу последовательности из первого миллиона не быть меньше ϵ. С вероятностью 1 истинно случайная последовательность содержит бесконечно много рядов в миллион последовательных элементов, меньших ϵ, так почему это не может произойти в начале последовательности?

Во-вторых, рассмотрим определение R2, и пусть P — такое свойство, что все элементы последовательности различны; P справедливо с вероятностью 1, поэтому любая последовательность с миллионом нулей не является случайной по *этому* критерию.

Пусть P — такое свойство, что *нет* элемента последовательности, равного нулю. P снова справедливо с вероятностью 1, поэтому по определению R2 любая последовательность с нулевым элементом не случайна. Вообще говоря, пусть x_0 — любое фиксированное число между нулем и единицей и пусть P — такое свойство: нет элемента последовательности, равного x_0. Из определения R2 сейчас следует, что нет случайной последовательности, которая может содержать элемент x_0! Можно доказать, что *не существует последовательности, удовлетворяющей условиям определения R2.* (Если $U_0, U_1, \ldots$ — такая последовательность, то выберем $x_0 = U_0$.)

Следовательно, если R1 — слишком слабое определение, то R2 является, несомненно, слишком строгим. "Правильное" определение должно быть менее строгим, чем R2. В действительности мы не показали, что определение R1 слишком слабое, поэтому продолжим исследование. Как упоминалось выше, ∞-распределенная последовательность *рациональных* чисел построена. (В самом деле, это не так уди-

вительно; см. упр. 18.) Почти все действительные числа иррациональны. Возможно, следовало бы потребовать, чтобы для случайной последовательности выполнялось

$$\Pr(U_n \text{ рациональное}) = 0.$$

Из определения равнораспределенности (см. равенство (2)) следует, что $\Pr(u \le U_n < v) = v-u$. Существует очевидный способ обобщения этого определения, используя теорию меры: "если $S \subseteq [0\,..\,1)$ — множество меры μ, то

$$\Pr(U_n \in S) = \mu \tag{27}$$

для всех случайных последовательностей $\langle U_n \rangle$". В частности, если S — множество рациональных чисел, то оно имеет меру нуль; значит, нет последовательности рациональных чисел, равнораспределенных в этом обобщенном смысле. Разумно ожидать, что теорема B может распространяться на интегрирование по Лебегу вместо интегрирования по Риману, если оговорено свойство (27). Однако мы снова найдем, что определение (27) слишком строгое, так как *нет* последовательностей, удовлетворяющих этому свойству. Если U_0, U_1, ... — любая последовательность, множество $S = \{U_0, U_1, \ldots\}$ есть множество меры нуль. Кроме того, $\Pr(U_n \in S) = 1$. Поэтому в силу тех же аргументов, из-за которых рациональные числа были исключены из случайных последовательностей, можно исключить все случайные последовательности.

До сих пор определение R1 можно было считать приемлемым. Однако существует несколько совершенно обоснованных возражений по этому поводу. Например, если имеется случайная в интуитивном смысле последовательность, то бесконечная подпоследовательность

$$U_0,\ U_1,\ U_4,\ U_9,\ \ldots,\ U_{n^2},\ \ldots \tag{28}$$

также должна быть случайной. Это не всегда верно для ∞-распределенных последовательностей. В самом деле, если взять любую ∞-распределенную последовательность и присвоить $U_{n^2} \leftarrow 0$ для всех n, количество $\nu_k(n)$, появляющихся в критерии k-распределенности, изменится самое большее на $\sqrt{n}$. Значит, отношение $\nu_k(n)/n$ не изменится. Таким образом, определение R1 не удовлетворяет этому критерию случайности.

Можно было бы усилить R1 следующим образом.

Определение R3. $[0\,..\,1)$-*последовательность называется случайной, если каждая ее бесконечная подпоследовательность является* ∞-*распределенной.*

Однако еще раз определение вышло очень строгим; любая равнораспределенная последовательность $\langle U_n \rangle$ имеет монотонную подпоследовательность с $U_{s_0} < U_{s_1} < U_{s_2} < \cdots$.

Секрет состоит в том, чтобы ограничиться подпоследовательностями, при построении которых можно было бы заранее сказать, принадлежит ли заданное U_n этой подпоследовательности. Предлагаем следующее определение.

Определение R4. $[0\,..\,1)$-*последовательность* $\langle U_n \rangle$ *называется случайной, если для любого эффективного алгоритма, точно определяющего бесконечную последовательность различных неотрицательных целых чисел* s_n *для* $n \ge 0$, *подпоследова-*

тельность $U_{s_0}, U_{s_1}, U_{s_2}, \ldots$, соответствующая этому алгоритму, является ∞-распределенной.

Алгоритм, относящийся к определению R4, — это эффективная процедура вычисления s_n при заданном n (см. обсуждение в разделе 1.1). Так, например, последовательность $\langle \pi^n \bmod 1 \rangle$ *не* удовлетворяет R4, поскольку она или не равнораспределена, или существует эффективный алгоритм, определяющий бесконечную подпоследовательность s_n с $(\pi^{s_0} \bmod 1) < (\pi^{s_1} \bmod 1) < (\pi^{s_2} \bmod 1) < \cdots$. Точно так же никакая *явно определенная последовательность не может удовлетворять определению R4*. Это справедливо, если согласиться с тем, что никакая явно определенная последовательность не является случайной. Судя по ее виду, явная последовательность $\langle \theta^n \bmod 1 \rangle$ на самом деле, однако, удовлетворяет определению R4 для почти всех действительных чисел $\theta > 1$; это не противоречие, так как почти все θ не могут быть вычислены алгоритмом. Ж. Ф. Коксма (J. F. Koksma) доказал, что $\langle \theta^{s_n} \bmod 1 \rangle$ является 1-распределенной для почти всех $\theta > 1$, если $\langle s_n \rangle$ — любая последовательность различных положительных целых чисел [*Compositio Math.* **2** (1935), 250–258]. Г. Нидеррейтер (H. Niederreiter) и Р. Ф. Тичи (R. F. Tichy) усилили теорему Коксма, заменив "1-распределенность" "∞-распределенностью" [*Mathematika* **32** (1985), 26–32]. Только счетное множество последовательностей $\langle s_n \rangle$ эффективно определимо; значит, $\langle \theta^n \bmod 1 \rangle$ почти всегда удовлетворяет R4.

Определение R4 более строгое, чем определение R1, но все еще можно утверждать, что определение R4 слишком слабое. Пусть, например, $\langle U_n \rangle$ — истинно случайная последовательность. Определим подпоследовательность $\langle U_{s_n} \rangle$ следующим образом: $s_0 = 0$ и, если $n > 0$, s_n — наименьшее целое число $\geq n$, для которого все $U_{s_n-1}, U_{s_n-2}, \ldots, U_{s_n-n}$ меньше $\frac{1}{2}$. Таким образом мы определили подпоследовательность значений, следующих за первой серией n значений, меньших $\frac{1}{2}$. Предположим, что $U_n < \frac{1}{2}$ соответствует выпадению "герба" при бросании монеты. Игроки склонны считать, что длинный ряд "гербов" предполагает, что появление "решки" становится более вероятным, если монета не поддельная. В этом случае подпоследовательность $\langle U_{s_n} \rangle$ определяет систему азартной игры для человека, который делает n-ю ставку на первую решку, следующую после серии из n "гербов." Игрок, возможно, думает, что $\Pr(U_{s_n} \geq \frac{1}{2})$ больше $\frac{1}{2}$, но, конечно, в истинной случайной последовательности $\langle U_{s_n} \rangle$ будет совершенно случайным. Нет системы азартных игр, которая всегда приводит к победе! Определение R4 ничего не говорит о подпоследовательности, формируемой в соответствии с такой системой азартных игр; так что, по-видимому, необходимо нечто большее.

Пусть определено "правило подпоследовательности" $\mathcal{R}$ как бесконечной последовательности функций $\langle f_n(x_1, \ldots, x_n) \rangle$, где для $n \geq 0$, f_n — функция n переменных и значение $f_n(x_1, \ldots, x_n)$ равно либо 0, либо 1. Здесь $x_1, \ldots, x_n$ — элементы некоторого множества S. (Таким образом, в данном случае f_0 — это постоянная функция, равная либо 0, либо 1.) Правило подпоследовательности $\mathcal{R}$ определяет подпоследовательность бесконечной последовательности $\langle X_n \rangle$ элементов S следующим образом: *n-й член X_n принадлежит подпоследовательности $\langle X_n \rangle \mathcal{R}$ тогда и только тогда, когда* $f_n(X_0, X_1, \ldots, X_{n-1}) = 1$. Заметим, что подпоследовательность

$\langle X_n\rangle\mathcal{R}$, определенная таким образом, необязательно бесконечна; фактически в ней вообще может не быть элементов.

Например, подпоследовательность азартного игрока, описанная выше, соответствует такому правилу подпоследовательности: $f_0 = 1$ и для $n > 0$, $f_n(x_1, \ldots, x_n) = 1$ тогда и только тогда, когда найдется некоторое k, $0 < k \le n$, такое, что k последовательных чисел x_m, x_{m-1}, $\ldots$, x_{m-k+1} все будут $< \frac{1}{2}$, когда $m = n$, но не когда $k \le m < n$.

Правило подпоследовательностей $\mathcal{R}$ называют *исчислимым*, если существует эффективный алгоритм, определяющий значение $f_n(x_1, \ldots, x_n)$, когда n и $x_1, \ldots, x_n$ заданы на входе. При попытке определить случайность лучше ограничиться исчислимыми правилами подпоследовательностей, когда пытаемся определить случайность, чтобы не получить чрезмерно ограничительное определение, подобное R3. Но эффективный алгоритм нельзя рассматривать с произвольными входными действительными числами. Например, если действительное число x точно определено бесконечным разложением в десятичной системе счисления, то не существует алгоритма для того, чтобы определить, будет ли $x < \frac{1}{3}$, так как для этого нужно исследовать все цифры числа $0.333\ldots$. Поэтому исчислимое правило подпоследовательностей неприменимо ко всем $[0\,..\,1)$-последовательностям и служит основой для следующего определения b-ичных последовательностей.

Определение R5. *b-ичная последовательность называется случайной, если каждая бесконечная подпоследовательность, определенная исчислимым правилом подпоследовательностей, является 1-распределенной.*

$[0\,..\,1)$-последовательность $\langle U_n\rangle$ называется случайной, если b-ичная последовательность $\langle\lfloor bU_n\rfloor\rangle$ является случайной для всех целых чисел $b \ge 2$.

Заметим, что определение R5 говорит только об 1-распределении, а не о ∞-распределении. Интересно проверить, что это может быть сделано без потери общности. Очевидно, можно определить исчислимое правило подпоследовательности $\mathcal{R}(a_1 \ldots a_k)$ следующим образом при любом заданном b-ичном числе $a_1 \ldots a_k$: пусть $f_n(x_1, \ldots, x_n) = 1$ тогда и только тогда, когда $n \ge k - 1$ и $x_{n-k+1} = a_1$, $\ldots$, $x_{n-1} = a_{k-1}$, $x_n = a_k$. Если $\langle X_n\rangle$ является k-распределенной b-ичной последовательностью, то правило $\mathcal{R}(a_1 \ldots a_k)$, с помощью которого выбирается подпоследовательность, содержащая точно следующие за появлением $a_1 \ldots a_k$ члены, определяет бесконечную подпоследовательность, и, если эта подпоследовательность 1-распределена, каждая из строк $a_1 \ldots a_k a_{k+1}$ размерности $(k+1)$ для $0 \le a_{k+1} < b$ появляется с вероятность $1/b^{k+1}$ в $\langle X_n\rangle$. Таким образом, индукцией по k можно доказать, что последовательность, удовлетворяющая определению R5, является k-распределенной для всех k. Подобным образом, рассматривая "композицию" правил подпоследовательностей, получаем: если $\mathcal{R}_1$ определяет бесконечную подпоследовательность $\langle X_n\rangle\mathcal{R}_1$, то можно определить правило подпоследовательности $\mathcal{R}_1\mathcal{R}_2$, для которого $\langle X_n\rangle\mathcal{R}_1\mathcal{R}_2 = (\langle X_n\rangle\mathcal{R}_1)\mathcal{R}_2$, и найти, что все подпоследовательности, рассматриваемые в определении R5, ∞-распределенные (см. упр. 32).

Факт, что ∞-распределение следует из определения R5 как очень частный случай, ободряет, и это хороший признак того, что можно, наконец, найти требуемое определение случайности. Но, увы, это все еще проблема. Не ясно, почему последовательность, удовлетворяющая определению R5, должна удовлетворять определе-

нию R4. Введенные нами исчислимые правила подпоследовательностей всегда дают подпоследовательности $\langle X_{s_n}\rangle$, для которых $s_0 < s_1 < \cdots$, но $\langle s_n\rangle$ не должны быть монотонны в R4; они только должны удовлетворять условию $s_n \neq s_m$ для $n \neq m$.

Итак, определения R4 и R5 можно скомбинировать следующим образом.

Определение R6. *b-ичная последовательность $\langle X_n\rangle$ называется случайной, если для каждого эффективного алгоритма, точно определяющего бесконечную последовательность различных неотрицательных целых чисел $\langle s_n\rangle$ как функцию n и значений $X_{s_0}, \ldots, X_{s_{n-1}}$, подпоследовательность $\langle X_{s_n}\rangle$, которая соответствует этому алгоритму, является случайной в смысле определения R5.*

$[0\,..\,1)$-последовательность $\langle U_n\rangle$ называется случайной, если b-ичная последовательность $\langle \lfloor bU_n\rfloor\rangle$ является случайной для всех целых чисел $b \geq 2$.

Автор утверждает*, что это определение точно удовлетворяет всем разумным философским требованиям случайности, а значит, отвечает на основной вопрос, поставленный в этом разделе.

D. Существование случайных последовательностей. Как было показано, определение R3 слишком строгое в том смысле, что не существует удовлетворяющей ему последовательности, и в приведенных выше определениях R4–R6 предпринимались попытки сохранить существенные элементы определения R3. Для того чтобы показать, что определение R6 не чрезмерно ограничительно, все еще необходимо доказать, что последовательность, удовлетворяющая всем этим условиям, существует. Интуитивно мы совершенно правильно ощущаем, что это не проблема, так как мы верим, что истинно случайная последовательность существует и удовлетворяет определению R6, но доказательство действительно необходимо, чтобы показать, что определение состоятельно.

Интересный метод построения последовательностей, удовлетворяющих определению R5, найден А. Вальдом (A. Wald). Он начинается с построения простой 1-распределенной последовательности.

Лемма T. *Пусть последовательность действительных чисел $\langle V_n\rangle$ определена следующим образом в терминах двоичной системы:*

$$V_0 = 0, \quad V_1 = .1, \quad V_2 = .01, \quad V_3 = .11, \quad V_4 = .001, \quad \ldots$$
$$V_n = .c_r \ldots c_1 1, \qquad \text{если } n = 2^r + c_1 2^{r-1} + \cdots + c_r. \tag{29}$$

Обозначим через $I_{b_1\ldots b_r}$ множество всех действительных чисел в $[0\,..\,1)$, которые имеют двоичное представление, начиная с $0.b_1\ldots b_r$. Таким образом,

$$I_{b_1\ldots b_r} = \big[(0.b_1\ldots b_r)_2\ ..\ (0.b_1\ldots b_r)_2 + 2^{-r}\big). \tag{30}$$

Тогда, если $\nu(n)$ — число V_k в $I_{b_1\ldots b_r}$ при $0 \leq k < n$, получим

$$\big|\nu(n)/n - 2^{-r}\big| \leq 1/n. \tag{31}$$

Доказательство. Так как $\nu(n)$ — число k, для которых $k \bmod 2^r = (b_r\ldots b_1)_2$, получим $\nu(n) = t$ или $t+1$, когда $\lfloor n/2^r\rfloor = t$. Следовательно, $\big|\nu(n) - n/2^r\big| \leq 1$. ∎

* По крайней мере, он сделал такое заявление, когда готовил этот материал в 1966 году.

Из (31) следует, что последовательность $\langle \lfloor 2^r V_n \rfloor \rangle$ — это равнораспределенная 2^r-ичная последовательность. Отсюда согласно теореме A получаем, что $\langle V_n \rangle$ — равнораспределенная $[0\,..\,1)$-последовательность. Действительно, достаточно ясно, что $\langle V_n \rangle$ настолько равнораспределена на $[0\,..\,1)$, насколько это вообще возможно. (Чтобы получить дополнительную информацию о данной и родственных последовательностях, обратитесь к работам J. G. van der Corput, *Proc. Koninklijke Nederl. Akad. Wetenschappen* **38** (1935), 813–821, 1058–1066; J. H. Halton, *Numerische Math.* **2** (1960), 84–90, 196; S. Haber, *J. Research National Bur. Standards* **B70** (1966), 127–136; R. Béjian and H. Faure, *Comptes Rendus Acad. Sci. Paris* **A285** (1977), 313–316; H. Faure, *J. Number Theory* **22** (1986), 4–20; S. Tezuka, *ACM Trans. Modeling and Comp. Simul.* **3** (1993), 99–107.

Л. Г. Рамшоу (L. H. Ramshaw) показал, что последовательность $\langle \phi n \bmod 1 \rangle$ немного более равнораспределена, чем $\langle V_n \rangle$ (см. *J. Number Theory* **13** (1981), 138–175).

Пусть сейчас $\mathcal{R}_1$, $\mathcal{R}_2$, ... — бесконечное множество правил подпоследовательностей, и необходимо найти последовательность $\langle U_n \rangle$, для которой все бесконечные подпоследовательности $\langle U_n \rangle \mathcal{R}_j$ равнораспределены.

Алгоритм W (*Последовательность Вальда*). Задана бесконечная последовательность правил подпоследовательностей $\mathcal{R}_1$, $\mathcal{R}_2$, ..., которая определяет подпоследовательности $[0\,..\,1)$-последовательностей рациональных чисел. Эта процедура определяет $[0\,..\,1)$-последовательность $\langle U_n \rangle$. Для вычисления требуется бесконечно много вспомогательных переменных $C[a_1, \ldots, a_r]$, где $r \geq 1$ и $a_j = 0$ или 1 для $1 \leq j \leq r$. Вначале все эти переменные равны нулю.

W1. [Инициализация n.] Присвоить $n \leftarrow 0$.

W2. [Инициализация r.] Присвоить $r \leftarrow 1$.

W3. [Проверка $\mathcal{R}_r$.] Если элемент U_n должен принадлежать подпоследовательности, определенной $\mathcal{R}_r$, которая основана на значениях U_k для $0 \leq k < n$, присвоить $a_r \leftarrow 1$; иначе — присвоить $a_r \leftarrow 0$.

W4. [Случай $[a_1, \ldots, a_r]$ не окончен?] Если $C[a_1, \ldots, a_r] < 3 \cdot 4^{r-1}$, перейти к шагу W6.

W5. [Увеличить r.] Присвоить $r \leftarrow r + 1$ и возвратиться к шагу W3.

W6. [Присвоить U_n.] Увеличить значение $C[a_1, \ldots, a_r]$ на 1, и пусть k будет его новым значением. Присвоить $U_n \leftarrow V_k$, где V_k определено в приведенной выше лемме T.

W7. [Увеличение n.] Увеличить n на 1 и возвратиться к шагу W2. ▮

Строго говоря, это не алгоритм, так как он не заканчивается, но мы, конечно, слегка преобразуем процедуру, чтобы он останавливал работу, когда n достигает заданного значения. Чтобы понять идею построения, выполните алгоритм вручную, заменяя число $3 \cdot 4^{r-1}$ шага W4 значением 2^r.

Алгоритм W не предназначен для получения случайных чисел на практике. Имеется в виду, что он имеет только теоретическое значение.

Теорема W. *Пусть $\langle U_n \rangle$ — последовательность рациональных чисел, определенная алгоритмом W, и пусть k — положительное целое число. Если подпоследовательность $\langle U_n \rangle \mathcal{R}_k$ бесконечна, то она 1-распределена.*

Доказательство. Пусть $A[a_1, \ldots, a_r]$ означает подпоследовательность $\langle U_n \rangle$ (возможно пустую), включающую те элементы U_n, которые для всех $j \le r$ принадлежат подпоследовательности $\langle U_n \rangle \mathcal{R}_j$, если $a_j = 1$, и не принадлежат подпоследовательности $\langle U_n \rangle \mathcal{R}_j$, если $a_j = 0$.

Достаточно доказать для всех $r \ge 1$ и всех пар двоичных чисел $a_1 \ldots a_r$ и $b_1 \ldots b_r$, что $\Pr(U_n \in I_{b_1 \ldots b_r}) = 2^{-r}$ по отношению к подпоследовательности $A[a_1, \ldots, a_r]$ всякий раз, когда последняя бесконечна (см. равенство (30)). Если $r \ge k$, для бесконечной последовательности $\langle U_n \rangle \mathcal{R}_k$ существует конечное объединение непересекающихся подпоследовательностей $A[a_1, \ldots, a_r]$ для $a_k = 1$ и $a_j = 0$ или 1 для $1 \le j \le r$, $j \ne k$. Из этого следует, что $\Pr(U_n \in I_{b_1 \ldots b_r}) = 2^{-r}$ по отношению к $\langle U_n \rangle \mathcal{R}_k$ (см. упр. 33). Этого достаточно, чтобы показать, что согласно теореме A последовательность 1-распределена.

Пусть $B[a_1, \ldots, a_r]$ означает подпоследовательность $\langle U_n \rangle$ для тех n, в которых $C[a_1, \ldots, a_r]$ увеличивается на единицу на шаге W6 алгоритма. Согласно алгоритму $B[a_1, \ldots, a_r]$ — конечная последовательность с самое большее $3 \cdot 4^{r-1}$ элементами. За исключением конечного числа, все члены $A[a_1, \ldots, a_r]$ являются членами подпоследовательностей $B[a_1, \ldots, a_r, \ldots, a_t]$, где $a_j = 0$ или 1 для $r < j \le t$.

Предположим, что $A[a_1, \ldots, a_r]$ бесконечна, и пусть $A[a_1, \ldots, a_r] = \langle U_{s_n} \rangle$, где $s_0 < s_1 < s_2 < \cdots$. Если N — большое целое число с $4^r \le 4^q < N \le 4^{q+1}$, логически вытекает, что число значений $k < N$, для которых U_{s_k} принадлежит $I_{b_1 \ldots b_r}$, равно (за исключением конечного множества элементов начальных подпоследовательностей)

$$\nu(N) = \nu(N_1) + \cdots + \nu(N_m).$$

Здесь m — число перечисленных выше подпоследовательностей $B[a_1, \ldots, a_t]$, в которых U_{s_k} появляется для некоторых $k < N$, N_j — число значений k с U_{s_k} в соответствующей подпоследовательности и $\nu(N_j)$ — число таких значений, которые также принадлежат $I_{b_1 \ldots b_r}$. Значит, согласно лемме T

$$\begin{aligned} \left|\nu(N) - 2^{-r}N\right| &= \left|\nu(N_1) - 2^{-r}N_1 + \cdots + \nu(N_m) - 2^{-r}N_m\right| \\ &\le \left|\nu(N_1) - 2^{-r}N_1\right| + \cdots + \left|\nu(N_m) - 2^{-r}N_m\right| \\ &\le m \le 1 + 2 + 4 + \cdots + 2^{q-r+1} < 2^{q+1}. \end{aligned}$$

Неравенство для m следует здесь из того, что согласно сделанному выбору N элемент U_{s_N} принадлежит $B[a_1, \ldots, a_t]$ для некоторых $t \le q + 1$.

Мы доказали, что $|\nu(N)/N - 2^{-r}| \le 2^{q+1}/N < 2/\sqrt{N}$. ∎

Чтобы показать окончательно, что существуют последовательности, удовлетворяющие определению R5, сначала заметим, что, если $\langle U_n \rangle$ — $[0\,..\,1)$-последовательность рациональных чисел и если $\mathcal{R}$ — исчисляемое правило подпоследовательностей для b-ичной последовательности, $\mathcal{R}$ можно преобразовать в исчисляемое правило подпоследовательности $\mathcal{R}'$ для $\langle U_n \rangle$, полагая $f'_n(x_1, \ldots, x_n)$ в $\mathcal{R}'$ равным $f_n(\lfloor bx_1 \rfloor, \ldots, \lfloor bx_n \rfloor)$ в $\mathcal{R}$. Если $[0\,..\,1)$-последовательность $\langle U_n \rangle \mathcal{R}'$ равнораспределена, значит, существует b-ичная последовательность $\langle \lfloor bU_n \rfloor \rangle \mathcal{R}$. Сейчас множество всех исчисляемых правил подпоследовательностей для b-ичных последовательностей для всех значений b является счетным (так как возможно только счетное множество эффективных алгоритмов). Значит, они могут образовать

некоторую последовательность $\mathcal{R}_1$, $\mathcal{R}_2$, ...; таким образом, алгоритм W задает $[0..1)$-последовательность, случайную в смысле определения R5.

Это приводит к возникновению в некоторой степени парадоксальной ситуации. Как уже упоминалось, не существует эффективного алгоритма для задания последовательности, удовлетворяющей определению R4. По тем же соображениям неэффективен алгоритм, который задает последовательность, удовлетворяющую определению R5. Доказательство существования такой случайной последовательности неизбежно неконструктивно. Как тогда алгоритм W может построить такую последовательность?

Здесь нет противоречия, есть только заминка, связанная с тем фактом, что множество всех эффективных алгоритмов не может быть пронумеровано эффективным алгоритмом. Другими словами, не существует эффективного алгоритма выбора j-го исчислимого правила подпоследовательности $\mathcal{R}_j$, поскольку нет эффективного алгоритма определения того, заканчивается ли данный вычислительный метод. Но важные большие классы алгоритмов *можно* систематически перенумеровывать. Так, например, в алгоритме W показано, что с помощью эффективного алгоритма можно построить последовательность, удовлетворяющую определению R5, если ограничиться рассмотрением "примитивных рекуррентных" правил подпоследовательностей.

Преобразовав шаг W6 алгоритма W (присвоив $U_n \leftarrow V_{k+t}$ вместо V_k, где t — любое неотрицательное целое число, зависящее от a_1, ..., a_r), можно показать, что существует *несчетное* множество $[0..1)$-последовательностей, удовлетворяющих определению R5.

В приведенной ниже теореме дан другой метод доказательства существования несчетного множества случайных последовательностей, который использует менее прямые доводы, основанные на теории меры, даже если применять строгое определение случайной последовательности R6.

Теорема М. *Пусть действительное число* x, $0 \le x < 1$, *соответствует двоичной последовательности* $\langle X_n \rangle$, *если* $(0.X_0X_1\ldots)_2$ *является двоичным представлением* x. *При этом почти все* x *соответствуют случайным в смысле определения R6 последовательностям.* (Другим словами, множество всех действительных чисел x, соответствующих неслучайным по определению R6 двоичным последовательностям, имеет меру нуль.)

Доказательство. Пусть $\mathcal{S}$ — эффективный алгоритм, определяющий бесконечную последовательность различных неотрицательных чисел $\langle s_n \rangle$, где выбор s_n зависит только от n и X_{s_k} для $0 \le k < n$, и пусть $\mathcal{R}$ — исчислимое правило подпоследовательности. Тогда любая двоичная последовательность $\langle X_n \rangle$ приводит к подпоследовательности $\langle X_{s_n} \rangle \mathcal{R}$ и по определению R6 эта подпоследовательность должна быть либо конечной, либо 1-распределенной. Достаточно доказать, что *для фиксированных* $\mathcal{R}$ *и* $\mathcal{S}$ *множество* $N(\mathcal{R}, \mathcal{S})$ *всех действительных* x, *соответствующих* $\langle X_n \rangle$, *такое, что* $\langle X_{s_n} \rangle \mathcal{R}$ *бесконечна и не* 1*-распределена, имеет меру нуль.* Для x существует неслучайное двоичное представление тогда и только тогда, когда x принадлежит $\bigcup N(\mathcal{R}, \mathcal{S})$, взятому по счетному множеству выборов $\mathcal{R}$ и $\mathcal{S}$.

Следовательно, пусть $\mathcal{R}, \mathcal{S}$ фиксированы. Рассмотрим множество $T(a_1a_2\ldots a_r)$, определенное для всех двоичных чисел $a_1a_2\ldots a_r$ как множество всех x, соответству-

ющих $\langle X_n \rangle$, такое, что $\langle X_{s_n} \rangle \mathcal{R}$ имеет $\geq r$ элементов, из которых первые r элементов равны $a_1, a_2, \ldots, a_r$ соответственно. Первым результатом будет неравенство

$$T(a_1 a_2 \ldots a_r) \text{ имеет меру } \leq 2^{-r}. \tag{32}$$

Доказательство начнем с замечания, что $T(a_1 a_2 \ldots a_r)$ является измеримым множеством: каждый элемент $T(a_1 a_2 \ldots a_r)$ — действительное число $x = (0.X_0 X_1 \ldots)_2$, для которого существует такое целое число m, что алгоритм $\mathcal{S}$ определяет различные значения $s_0, s_1, \ldots, s_m$, и правило $\mathcal{R}$ определяет подпоследовательность $X_{s_0}, X_{s_1}, \ldots, X_{s_m}$, такую, что X_{s_m} является r-м элементом этой подпоследовательности. Множество всех действительных $y = (0.Y_0 Y_1 \ldots)_2$, таких, что $Y_{s_k} = X_{s_k}$ для $0 \leq k \leq m$, также принадлежит $T(a_1 a_2 \ldots a_r)$ и является измеримым множеством, состоящим из конечного объединения двоичных подынтервалов $I_{b_1 \ldots b_t}$. Поскольку существует только счетное множество таких двоичных интервалов, то $T(a_1 a_2 \ldots a_r)$ — счетное объединение двоичных интервалов, и оно, следовательно, измеримо. Более того, данный довод может быть распространен, чтобы показать, что мера $T(a_1 \ldots a_{r-1} 0)$ равна мере $T(a_1 \ldots a_{r-1} 1)$, так как последнее множество является объединением двоичных интервалов, полученных из предшествующей рекуррентной формулы $Y_{s_k} = X_{s_k}$ для $0 \leq k < m$ и $Y_{s_m} \neq X_{s_m}$. Поскольку

$$T(a_1 \ldots a_{r-1} 0) \cup T(a_1 \ldots a_{r-1} 1) \subseteq T(a_1 \ldots a_{r-1}),$$

мера $T(a_1 a_2 \ldots a_r)$ равна самое большее половине меры $T(a_1 \ldots a_{r-1})$. Неравенство (32) теперь легко получить индукцией по r.

Сейчас, когда (32) установлено, осталось, по существу, показать, что двоичное представление почти всех действительных чисел равнораспределено. Пусть для $0 < \epsilon < 1$ $B(r, \epsilon)$ — это $\bigcup T(a_1 \ldots a_r)$, где объединение берется по всем двоичным строкам $a_1 \ldots a_r$, для которых число единиц $\nu(r)$ среди $a_1 \ldots a_r$ удовлетворяет неравенству

$$\left| \nu(r) - \tfrac{1}{2} r \right| \geq \epsilon r.$$

Число таких двоичных строк равно $C(r, \epsilon) = \sum \binom{r}{k}$, и суммирование выполняется по всем значениям k с $|k - \frac{1}{2} r| \geq \epsilon r$. В упр. 1.2.10–21 доказано, что $C(r, \epsilon) \leq 2^{r+1} e^{-\epsilon^2 r}$, отсюда согласно (32)

$$B(r, \epsilon) \text{ имеет меру } \leq 2^{-r} C(r, \epsilon) \leq 2 e^{-\epsilon^2 r}. \tag{33}$$

На следующем шаге определим

$$B^*(r, \epsilon) = B(r, \epsilon) \cup B(r+1, \epsilon) \cup B(r+2, \epsilon) \cup \cdots.$$

Мера $B^*(r, \epsilon)$ равна самое большее $\sum_{k \geq r} 2 e^{-\epsilon^2 k}$ и является остатком сходящегося ряда, так что

$$\lim_{r \to \infty} \big(\text{мера } B^*(r, \epsilon) \big) = 0. \tag{34}$$

Теперь, если x — действительное число, двоичное разложение $(0.X_0 X_1 \ldots)_2$ которого приводит к бесконечной не 1-распределенной последовательности $\langle X_{s_n} \rangle \mathcal{R}$, и если $\nu(r)$ обозначает число 1 в первых r элементах последней последовательности, то

$$\left| \nu(r)/r - \tfrac{1}{2} \right| \geq \epsilon$$

для некоторого $\epsilon > 0$ и бесконечного множества r. Это означает, что x принадлежит $B^*(r,\epsilon)$ для всех r. И наконец, находим, что

$$N(\mathcal{R},\mathcal{S}) = \bigcup_{t\ge 2} \bigcap_{r\ge 1} B^*(r,1/t).$$

Согласно (34) $\bigcap_{r\ge 1} B^*(r,1/t)$ имеет меру нуль для всех t. Следовательно, $N(\mathcal{R},\mathcal{S})$ имеет меру нуль. ▮

Основываясь на факте существования *двоичных* последовательностей, удовлетворяющих определению R6, можно показать существование $[0\,..\,1)$ случайных в этом смысле последовательностей (подробности — в упр. 36). Состоятельность определения R6 в связи с этим установлена.

E. Случайные конечные последовательности. Доводы, приведенные выше, показывают, что невозможно определить понятие случайности конечной последовательности: заданная конечная последовательность так же вероятна, как и любая другая. До сих пор почти каждый согласен, что последовательность 011101001 "более случайна", чем 101010101, и последняя последовательность "более случайна", чем 000000000. Хотя верно, что истинно случайные последовательности проявляют локально неслучайное поведение, мы предполагаем такое поведение только у длинной конечной последовательности, а не у короткой.

Предлагались различные способы определения случайности конечной последовательности, но здесь будут сделаны только наброски нескольких идей. Для простоты ограничимся рассмотрением b-ичных последовательностей.

Пусть задана b-ичная последовательность X_0, X_1, ..., X_{N-1}. Можно сказать, что

$$\Pr\bigl(S(n)\bigr) \approx p, \qquad \text{если } \bigl|\nu(N)/N - p\bigr| \le 1/\sqrt{N}, \tag{35}$$

где $\nu(n)$ — величина, появившаяся в определении A в начале раздела. Приведенную выше последовательность можно назвать k-распределенной, если

$$\Pr(X_n X_{n+1} \ldots X_{n+k-1} = x_1 x_2 \ldots x_k) \approx 1/b^k \tag{36}$$

для всех b-ичных чисел $x_1 x_2 \ldots x_k$. (Ср. с определением D. К сожалению, последовательность может быть k-распределенной согласно этому новому определению, когда она не $(k-1)$-распределена.)

Сейчас можно дать следующее аналогичное определению R1 определение случайности.

Определение Q1. *b-ичная последовательность длины N случайна, если она k-распределена (в вышеприведенном смысле) для всех положительных целых чисел $k \le \log_b N$.*

Например, согласно этому определению существует 178 неслучайных двоичных последовательностей длины 11,

```
00000001111  10000000111  11000000011  11100000001  11110000000
00000001110  10000000110  11000000010  11100000000  11010000000
00000001101  10000000101  11000000001  10100000001  10110000000 ,
00000001011  10000000011  01000000011  01100000001  01110000000
00000000111
```

плюс 01010101010 и все последовательности с девятью или более нулями, плюс все последовательности, полученные из предшествующих последовательностей, если единицы и нули поменять местами.

Подобным образом можно сформулировать определение, аналогичное определению R6, для конечной последовательности. Пусть **A** — множество алгоритмов, каждый из которых является процедурой выделения и выбора, дающей подпоследовательность $\langle X_{s_n}\rangle\mathcal{R}$, как в доказательстве теоремы M.

Определение Q2. *b-ичная последовательность X_0, X_1, ..., X_{N-1} является (n,ϵ)-случайной относительно множества алгоритмов* **A**, *если для каждой подпоследовательности X_{t_1}, X_{t_2}, ..., X_{t_m}, определенной алгоритмом* **A**, *мы имеем либо $m < n$, либо*

$$\left|\frac{1}{m}\nu_a(X_{t_1},\ldots,X_{t_m}) - \frac{1}{b}\right| \le \epsilon \qquad \text{для } 0 \le a < b.$$

Здесь $\nu_a(x_1,\ldots,x_m)$ — количество чисел a в последовательности x_1, ..., x_m.

(Другими словами, каждая достаточно длинная подпоследовательность, определенная алгоритмом из множества **A**, должна быть приближенно равнораспределенной.) Основной идеей в этом случае является предположение, что **A** — множество "простых" алгоритмов и количество (и сложность) алгоритмов в **A** может увеличиваться при росте N.

В качестве примера определения Q2 рассмотрим двоичную последовательность. Пусть **A** состоит из следующих четырех алгоритмов.

a) Взять всю последовательность.
b) Взять нечетные члены последовательности, начиная с первого.
c) Взять члены последовательности, следующие за нулем.
d) Взять члены последовательности, следующие за единицей.

Сейчас последовательность X_0, X_1, ..., X_7 является $(4,\frac{1}{8})$-случайной относительно **A**, если:

по (a) $\left|\frac{1}{8}(X_0+X_1+\cdots+X_7)-\frac{1}{2}\right| \le \frac{1}{8}$, т. е. если последовательность содержит 3, 4 или 5 единиц;

по (b) $\left|\frac{1}{4}(X_0+X_2+X_4+X_6)-\frac{1}{2}\right| \le \frac{1}{8}$, т. е. если последовательность содержит точно 2 единицы на четной позиции;

по (c) существуют три возможности в зависимости от того, как много нулей занимают позиции X_0, ..., X_6: если здесь 2 или 3 нуля, то нет необходимости в проверке (так как $n = 4$); если 4 нуля, то за ними должны следовать 2 нуля и 2 единицы; если 5 нулей, то за ними должны следовать 2 единицы и 3 нуля;

по (d) мы получим условия, подобные условиям в (c).

Оказывается, что только следующие двоичные последовательности длины 8 являются $(4, \frac{1}{8})$-случайными в соответствии с этими правилами,

00001011	00101001	01001110	01101000
00011010	00101100	01011011	01101100
00011011	00110010	01011110	01101101
00100011	00110011	01100010	01110010 '
00100110	00110110	01100011	01110110
00100111	00111001	01100110	

плюс те, которые получены из этих, если единицы и нули поменять местами.

Ясно, что множество алгоритмов можно сделать таким большим, что не будет последовательностей, не удовлетворяющих определению, когда n и ϵ малы. А. Н. Колмогоров доказал, что (n, ϵ)-случайная двоичная последовательность всегда *будет* существовать для любого заданного N, если число алгоритмов в **A** не превышает

$$\tfrac{1}{2} e^{2n\epsilon^2(1-\epsilon)}. \tag{37}$$

Этот результат не достаточно строгий, чтобы показать, что последовательности, удовлетворяющие определению Q1, существуют, но последние могут быть эффективно построены с использованием процедуры Риса (Rees) из упр. 3.2.2–21. Обобщенный спектральный критерий, основанный на дискретном преобразовании Фурье, можно использовать для проверки, насколько последовательность соответствует определению Q1 [см. A. Compagner, *Physical Rev.* **E52** (1995), 5634–5645].

Другие интересные подходы к определению случайности приведены Пером Мартин-Лёфом (Per Martin-Löf, *Information and Control* **9** (1966), 602–619). Пусть задана конечная b-ичная последовательность $X_1, \ldots, X_N$, $l(X_1, \ldots, X_N)$ — длина самой короткой программы Тьюринга, которая генерирует эту последовательность. (Вместо этого можно использовать другие классы эффективных алгоритмов, например такие, которые обсуждались в разделе 1.1.) Тогда $l(X_1, \ldots, X_N)$ — мера "хаотичности" последовательности и это понятие можно отождествить со случайностью. Последовательности длины N, максимизирующие $l(X_1, \ldots, X_N)$, можно называть случайными. (С практической точки зрения генерирование случайного числа компьютером — это, конечно, наихудшее определение "случайности", какое можно себе представить!)

По существу, почти в то же время такое определение случайности независимо предложил Г. Чайтин (G. Chaitin, *JACM* **16** (1969), 145–159.) Интересно отметить, что хотя в данных определениях не упоминается о свойствах равнораспределенности, как в наших определениях, Мартин-Лёф и Чайтин доказали, что случайные последовательности этого вида также имеют ожидаемые свойства равнораспределенности. На самом деле Мартин-Лёф продемонстрировал, что такие последовательности удовлетворяют *всем* вычислимым статистическим критериям случайности в соответствующем смысле.

Дополнительную информацию об определении случайной конечной последовательности можно найти в следующих работах: Звонкин А. К. и Левин Л. А. *Успехи мат. наук* **25**, 6 (Ноябрь, 1970), 85–127; Левин Л. А. *Докл. Акад. наук СССР* **212** (1973), 548–550; Левин Л. А. *Информация и контроль* **61** (1984), 15–37.

F. Псевдослучайные числа. С теоретической точки зрения утешительно знать, что существуют случайные конечные последовательности с разными особенностями, но такие теоремы не дают ответов на вопросы, с которыми сталкиваются в действительности программисты. Недавние исследования привели к более практичной теории, основанной на изучении *множеств* конечных последовательностей. Точнее, рассмотрим *мультимножества*, в которых последовательности могут появляться более одного раза.

Пусть S — мультимножество, содержащее двоичные строки длины N; назовем S *N-источником.* Пусть $\$_N$ означает определенный N-источник, содержащий все 2^N возможных N-двоичных строк. Каждый элемент S представляет последовательность, которую можно использовать в качестве источника псевдослучайных двоичных разрядов, выбор различных "начальных" значений приводит к различным элементам S. Например, возможно такое множество S

$$\{B_1B_2\ldots B_N \mid B_j \text{ старший значащий двоичный разряд } X_j\} \tag{38}$$

в линейной конгруэнтной последовательности, определенной равенством $X_{j+1} = (aX_j + c) \bmod 2^e$, в котором существует одна строка $B_1B_2\ldots B_N$ для каждого из 2^e начальных значений X_0.

Основной идеей псевдослучайных последовательностей, как будет показано в этой главе, является получение N двоичных разрядов, появляющихся случайно, несмотря на то что мы используем лишь несколько "истинно случайных" двоичных разрядов, когда выбираем начальное значение. В только что рассмотренном примере понадобилось e истинно случайных двоичных разрядов для выбора X_0. Вообще, для использования отбирается $\lg|S|$ из S истинно случайных двоичных разрядов, после чего процедура становится детерминированной. Если $N = 10^6$ и $|S| = 2^{32}$, получаем более 30 000 "кажущихся случайными" двоичных разрядов из выбранного истинно случайного двоичного разряда. При $\$_N$ вместо S мы не получим такого большого числа "случайных разрядов", поскольку $\lg|\$_N| = N$.

Что означает "кажущихся случайными"? Э. Ч. Яо (A. C. Yao) в 1982 году предложил хорошее определение: рассмотрим любой алгоритм A, который при применении к строке двоичных разрядов $B = B_1\ldots B_N$ выдает значение $A(B) = 0$ или 1. Можно рассматривать A как критерий случайности, например алгоритм A может вычислить распределение серий последовательных нулей и единиц, выдавая на выходе единицу, если длины серий значительно отличаются для ожидаемого распределения. Что бы ни делал A, вероятность $P(A, S)$ можно рассматривать как вероятность того, что $A(B) = 1$, когда B — случайно выбранный элемент из S, и можно сравнивать с вероятностью $P(A, \$_N)$ того, что $A(B) = 1$, когда B — истинно случайная строка двоичных разрядов длины N. Если $P(A, S)$ будет очень близким к $P(A, \$_N)$ для всех статистических критериев A, то мы ничего не сможем сказать о различии между последовательностями S и истинно случайными двоичными последовательностями.

Определение Р. *Мы говорим, что N-источник S проходит статистический критерий A с допустимым отклонением ϵ, если $\big|P(A, S) - P(A, \$_N)\big| < \epsilon$. Критерий "проваливается", если $\big|P(A, S) - P(A, \$_N)\big| \geq \epsilon$.*

Нет необходимости в том, чтобы алгоритм A задавали статистики. *Любой* алгоритм можно рассматривать как статистический критерий случайности согласно определению P. Мы позволяем A бросать монеты (т. е. использовать истинно случайные двоичные разряды), а также выполнять вычисления. Единственное требование — A должен выдавать на выходе 0 или 1.

Конечно, в действительности существуют другие требования: мы утверждаем, что алгоритм A должен давать результат на выходе за разумное время, по крайней мере в среднем. Нам не интересен алгоритм, который работает много лет, потому что мы никогда не заметим какого-нибудь несоответствия между S и $\$_N$, если наш компьютер не сможет обнаружить их в течение нашей жизни. Последовательности S содержат только $\lg|S|$ двоичных разрядов информации, так что, несомненно, существуют алгоритмы, которые в конечном счете обнаружат избыточность. Но ведь нас интересует, как долго S будет проходить все реально имеющие значение критерии.

Эти качественные идеи можно выразить, как мы сейчас увидим, в количественной форме. Такая теория весьма тонкая, но она достаточно красивая и важная, так что читатель, нашедший время внимательно изучить детали, будет вознагражден.

В последующем обсуждении *время выполнения* $T(A)$ алгоритмом A на N двоичных строк определяется как максимальное ожидаемое число шагов, необходимых для выхода $A(B)$, максимум берется по всем $B \in \$_N$, ожидаемое число является средним по распределению, соответствующему подбрасыванию монеты.

Первый шаг количественного анализа — показать, что можно ограничиться критерием очень специального вида. Пусть A_k — алгоритм, зависящий только от первых k двоичных разрядов во входной строке $B = B_1 \ldots B_N$, где $0 \le k < N$, и пусть $A_k^P(B) = \big(A_k(B) + B_{k+1} + 1\big) \bmod 2$. Тогда A_k^P выводит 1 тогда и только тогда, когда A_k успешно предсказало B_{k+1}; назовем A_k^P критерием *прогноза*.

Лемма P1. *Пусть S — N-источник. Если S не проходит критерий A с допустимым отклонением ϵ, то существует целое число $k \in \{0, 1, \ldots, N-1\}$ и критерий прогноза A_k^P с $T(A_k^P) \le T(A) + O(N)$, такой, что S не проходит A_k^P с допустимым отклонением ϵ/N.*

Доказательство. Дополнительно при необходимости можно предположить, что $P(A,S) - P(A,\$_N) \ge \epsilon$. Рассмотрим алгоритмы F_k, которые начинают подбрасывать $N-k$ монет, и заменим $B_{k+1} \ldots B_N$ случайными двоичными разрядами $B'_{k+1} \ldots B'_N$ до выполнения A. Алгоритм F_N такой же, как и A, в то время как F_0 действует на S, как A действует на $\$_N$. Пусть $p_k = P(F_k, S)$. Поскольку $\sum_{k=0}^{N-1} (p_{k+1} - p_k) = p_N - p_0 = P(A,S) - P(A,\$_N) \ge \epsilon$, существуют k, такие, что $p_{k+1} - p_k \ge \epsilon/N$.

Пусть A_k^P — алгоритм, выполняющий вычисления F_k и предсказывающий значение $(F_k(B) + B'_{k+1} + 1) \bmod 2$. Другими словами, он выводит

$$A_k^P(B) = \big(F_k(B) + B_{k+1} + B'_{k+1}\big) \bmod 2. \tag{39}$$

Внимательный анализ вероятностей показывает, что $P(A_k^P, S) - P(A_k^P, \$_N) = p_{k+1} - p_k$ (см. упр. 40). ∎

Большая часть N-источников S, имеющих практическое преимущество, являются источниками с *симметричным сдвигом* в том смысле, что каждая подстрока

$B_1 \ldots B_k$, $B_2 \ldots B_{k+1}$, $\ldots$, $B_{N-k+1} \ldots B_N$ длины k имеет одно и то же вероятностное распределение. Это выполняется, например, когда S соответствует линейной конгруэнтной последовательности, как в (38). В таких случаях лемму P1 можно улучшить, взяв $k = N - 1$.

Лемма P2. *Если S является N-источником с симметричным сдвигом, который не проходит критерий A с допустимым отклонением ϵ, то существует алгоритм A' с $T(A') \le T(A) + O(N)$, который предсказывает B_N из $B_1 \ldots B_{N-1}$ с вероятностью, равной по крайней мере $\frac{1}{2} + \epsilon/N$.*

Доказательство. Если $P(A, S) > P(A, \$_N)$, пусть A' есть A_k^P в доказательстве леммы P1, только примененное к $B_{N-k} \ldots B_{N-1} 0 \ldots 0$ вместо $B_1 \ldots B_N$. Тогда A' в среднем ведет себя так же из-за симметричного сдвига. Если $P(A, S) < P(A, \$_N)$, пусть A' есть $1 - A_k^P$ в том же виде. Ясно, что $P(A', \$_N) = \frac{1}{2}$. ▮

Введем более жесткие ограничения на S. Предположим, что каждая из последовательностей $B_1 B_2 \ldots B_N$ имеет вид $f\big(g(X_0)\big) f\big(g(g(X_0))\big) \ldots f\big(g^{[N]}(X_0)\big)$, где X_0 — это упорядочение некоторого множества X, g является перестановкой X и $f(x)$ равно 0 или 1 для всех $x \in X$. Наш линейный конгруэнтный пример удовлетворяет этим ограничениям с $X = \{0, 1, \ldots, 2^e - 1\}$, $g(x) = (ax + c) \bmod 2^e$ и $f(x) =$ старший значащий двоичный разряд x. Такие N-источники назовем *итеративными*.

Лемма P3. *Если S — итеративный N-источник, который не удовлетворяет критерию A с допустимым отклонением ϵ, то существует алгоритм A' с $T(A') \le T(A) + O(N)$, предсказывающий B_1 по $B_2 \ldots B_N$ по крайней мере с вероятностью $\frac{1}{2} + \epsilon/N$.*

Доказательство. Итеративный N-источник является источником с симметричным сдвигом. Значит, он является своим отражением $S^R = \{B_N \ldots B_1 \mid B_1 \ldots B_N \in S\}$. Следовательно, лемма P2 применима к S^R. ▮

Перестановка $g(x) = (ax + c) \bmod 2^e$ легко обратима в том смысле, что x можно определить через $g(x)$ всякий раз, когда a нечетное. Однако множества легко вычисляемых функций перестановок являются "односторонними" (трудно обратимыми), и мы увидим, что это делает их вероятно хорошими источниками псевдослучайных чисел.

Лемма P4. *Пусть S — итеративный N-источник, соответствующий f, g и X. Если S не удовлетворяет критерию A с допустимым отклонением ϵ, то существует алгоритм G, который правильно оценивает $f(x)$ по заданной $g(x)$ с вероятностью $\ge \frac{1}{2} + \epsilon/N$, когда x — случайный элемент из X. Время вычисления $T(G)$ будет не более чем $T(A) + O(N)\big(T(f) + T(g)\big)$.*

Доказательство. Зададим $y = g(x)$. Требуемый алгоритм G вычисляет $B_2 = f\big(g(x)\big)$, $B_3 = f\big(g(g(x))\big)$, $\ldots$, $B_N = f\big(g^{[N-1]}(x)\big)$ и применяет алгоритм A' леммы P3. Он приближенно оценивает $f(x) = B_1$ с вероятностью $\ge \frac{1}{2} + \epsilon/N$, так как g является перестановкой X, и $B_1 \ldots B_N$ — элемент S, соответствующий начальному значению X_0, для которого выполняется $g(X_0) = x$. ▮

Для того чтобы использовать лемму P4, нужно усилить способность приближенно оценивать один двоичный разряд $f(x)$ для того, чтобы уметь оценивать x только по значению $g(x)$. Для этого существует только тонкий общий способ —

использовать свойства булевых функций, если расширить S так, что появится потребность приближенно оценивать множество различных функций $f(x)$. (Однако метод является несколько техническим. Так, при первом чтении следует, возможно, перейти к теореме G, прежде чем внимательно рассматривать следующие ниже детали.)

Предположим, $G(z_1 \dots z_R)$ — бинарнозначная функция (принимающая значения 0 или 1) на строках из R двоичных разрядов, которая хорошо приближает функцию вида $f(z_1 \dots z_R) = (x_1 z_1 + \cdots + x_R z_R) \bmod 2$ для некоторого фиксированного $x = x_1 \dots x_R$. Удобно измерять, насколько это приближение успешно, вычисляя среднее значение

$$s = \mathrm{E}\big((-1)^{G(z_1 \dots z_R) + x_1 z_1 + \cdots + x_R z_R}\big), \tag{40}$$

где усреднение берется по всем возможным значениям $z_1 \dots z_R$. Это сумма правильных оценок минус неправильные оценки деленная на 2^R; так, если p — вероятность того, что G правильно, то получим $s = p - (1-p)$ или $p = \frac{1}{2} + \frac{1}{2}s$.

Например, предположим, что $R = 4$ и $G(z_1 z_2 z_3 z_4) = [z_1 \neq z_2][z_3 + z_4 < 2]$. Функция дает хорошую оценку $s = \frac{3}{4}$ (и $p = \frac{7}{8}$), если $x = 1100$, так как она равна $x \cdot z \bmod 2 = (z_1 + z_2) \bmod 2$ для всех строк z, содержащих 4 двоичных разряда, исключая 0111 или 1011. Она также дает хорошую оценку $\frac{1}{4}$, когда $x = 0000$, 0011, 1101 или 1110, так что существует пять вероятных возможностей для x. Остальные одиннадцать x_k дают $s \leq 0$.

Следующий алгоритм магически находит x в большинстве случаев, когда G успешно оценивает в только что описанном смысле. Точнее, алгоритм строит короткий перечень элементов, имеющих хорошую возможность содержать x.

Алгоритм L (*Усиление линейных оценок*). Пусть задана бинарнозначная функция $G(z_1 \dots z_R)$ и положительное целое число k. Этот алгоритм дает таблицу 2^k двоичных последовательностей $x = x_1 \dots x_R$ с таким свойством, что x, вероятно, находится в этой таблице, когда $G(z_1 \dots z_R)$ является хорошим приближением функции $(x_1 z_1 + \cdots + x_R z_R) \bmod 2$.

L1. [Построение случайной матрицы.] Генерировать случайные двоичные разряды B_{ij} для $1 \leq i \leq k$ и $1 \leq j \leq R$.

L2. [Подсчет знаков.] Для $1 \leq i \leq R$ и для всех двоичных разрядов строк $b = b_1 \dots b_k$ вычислить

$$h_i(b) = \sum_{c \neq 0} (-1)^{b \cdot c + G(cB + e_i)}, \tag{41}$$

где e_i — строка, содержащая R двоичных разрядов, $0 \dots 010 \dots 0$ с 1 на позиции i и где cB — строка $d_1 \dots d_R$ с $d_j = (B_{1j} c_1 + \cdots + B_{kj} c_k) \bmod 2$. (Другими словами, двоичный вектор $c_1 \dots c_k$ является кратным двоичной матрице B размера $k \times R$.) Сумма взята по всем $2^k - 1$ строкам двоичных разрядов, $c_1 \dots c_k \neq 0 \dots 0$. Для каждого i можно оценить $h_i(b)$ $k \cdot 2^k$ сложениями и вычитаниями с помощью метода Ятеса (Yates) для преобразования Уолша (см. замечание к равенству 4.6.4–(38)).

L3. [Вывод оценок.] Для всех 2^k выборов $b = b_1 \dots b_k$ вывод строки $x(b) = [h_1(b) < 0] \dots [h_R(b) < 0]$. ∎

Для доказательства того, что алгоритм L хорошо работает, необходимо показать, что заданная строка x, вероятно, выводится всякий раз, когда она этого заслуживает. Сначала заметим, что, если заменить G на G', где $G'(z) = (G(z) + z_j) \bmod 2$, начальное $G(z)$ будет хорошим приближением $x \cdot z \bmod 2$ тогда и только тогда, когда новое $G'(z)$ будет хорошим приближением $(x + e_j) \cdot z \bmod 2$, где e_j — единичный вектор-строка, определенная на шаге L2. Кроме того, если применить алгоритм G' вместо G, можно получить

$$h'_i(b) = \sum_{c\neq 0}(-1)^{b\cdot c+G(cB+e_i)+(cB+e_i)\cdot e_j} = (-1)^{\delta_{ij}} h_i\big((b + B_j) \bmod 2\big),$$

где B_j — j-й столбец B. Следовательно, на шаге L3 выводится вектор $x'(b) = x\big((b+B_j) \bmod 2\big) + e_j$ по модулю 2. Поскольку b пробегает все строки, состоящие из k двоичных разрядов, $(b + B_j) \bmod 2$ также пробегает эти строки, следствием чего является дополнение j-м двоичным разрядом каждого x на выходе.

Следовательно, достаточно доказать, что вектор $x = 0 \ldots 0$ можно вывести, как только $G(z)$ хорошо аппроксимирует постоянную функцию 0. В действительности мы покажем, что $x(0 \ldots 0)$ равняется $0 \ldots 0$ на шаге L3 с большой вероятностью всякий раз, когда $G(z)$ с большей вероятностью принимает значение 0, чем 1, и k является достаточно большим. Точнее говоря, условие

$$\sum_{c\neq 0}(-1)^{G(cB+e_i)} > 0$$

выполняется для $1 \le i \le R$ с вероятностью $> \frac{1}{2}$, если $s = \mathrm{E}\big((-1)^{G(z)}\big)$ положительно, где среднее берется по всем 2^R возможным z и если k достаточно велико.

Ключом исследования является утверждение, что для каждого фиксированного $c = c_1 \ldots c_k \neq 0 \ldots 0$ строка $d = cB$ равномерно распределена: каждое значение d появляется с вероятностью $1/2^R$, так как двоичные разряды B случайны. Более того, когда $c \neq c' = c'_1 \ldots c'_k$, строки $d = cB$ и $d' = c'B$ *независимы*: каждое значение пары (d, d') происходит с вероятностью $1/2^{2R}$. Следовательно, можно рассуждать, как при доказательстве неравенства Чебышева: для любого фиксированного i сумма $\sum_{c\neq 0}(-1)^{G(cB+e_i)}$ отрицательна с вероятностью, не большей, чем $1/((2^k - 1)s^2)$. (Подробности содержатся в упр. 42.) Поэтому $R/((2^k - 1)s^2)$ — верхняя грань вероятности того, что $x(0)$ не является нулем на шаге L3.

Теорема G. *Если* $s = \mathrm{E}\big((-1)^{G(z)+x\cdot z}\big) > 0$ *и* $2^k > \lceil 2R/s^2 \rceil$, *то алгоритм L выводит* x *с вероятностью* $\ge \frac{1}{2}$. *Время счета равно* $O(k2^kR)$ *плюс время получения* 2^kR *оценок* G. ▮

Сейчас мы готовы доказать, что последовательность смешанно-квадратичного метода, заданная в 3.2.2–(17), является хорошим источником (псевдо)случайных чисел. Предположим, что $2^{R-1} < M = PQ < 2^R$, где P и Q — простые числа вида $4k+3$, удовлетворяющие неравенствам $2^{(R-2)/2} < P < 2^{(R-1)/2}$, $2^{R/2} < Q < 2^{(R+1)/2}$. Назовем M, состоящее из R двоичных разрядов, *целым числом Блюма*, поскольку на важность таких чисел для криптографии было впервые указано Мануэлем Блюмом (Manuel Blum, *COMPCON* **24** (Spring, 1982), 133–137). Блюм первоначально предложил выбрать P и Q так, чтобы они оба имели $R/2$ двоичных разрядов, но

алгоритм 4.5.4D показал, что лучше выбрать P и Q, как сформулировано здесь: чтобы $Q - P > .29 \times 2^{R/2}$.

Выбрать X_0 наугад среди чисел $0 < X_0 < M$ с $X_0 \perp M$. Пусть также Z — случайная, состоящая из R двоичных разрядов, маска. Можно построить итеративный N-источник S, полагая X множеством всех (x, z, m), которые возможны для (X_0, Z, M) с дополнительным ограничением $x \equiv a^2$ (по модулю m) для некоторых a. Легко показать, что функция $g(x, z, m) = (x^2 \bmod m, z, m)$ — это перестановка X (см., например, упр. 4.5.4–35). Функция $f(x, z, m)$, извлекающая двоичные разряды в этом итеративном источнике, равна $x \cdot z \bmod 2$. Наше начальное значение (X_0, Z, M) не является необходимым в X, но $g(X_0, Z, M)$ имеет равномерное распределение в X, так как точно четыре значения X_0 имеют данный квадрат $X_0^2 \bmod M$.

Теорема Р. *Пусть S — N-источник, который определен смешанно-квадратичным методом согласно модулю, содержащему R двоичных разрядов, и предположим, что S не удовлетворяет некоторому статистическому критерию A с допустимым отклонением $\epsilon \geq 1/2^N$. Тогда можно построить алгоритм F, который найдет множители состоящего из R двоичных разрядов случайного целого числа Блюма $M = PQ$, имеющего вид, описанный выше, с вероятностью по крайней мере $\epsilon/(4N)$ и временем счета $T(F) = O(N^2R^2\epsilon^{-2}T(A) + N^3R^4\epsilon^{-2})$.*

Доказательство. Умножение по модулю M можно осуществить за $O(R^2)$ шагов; следовательно, $T(f) + T(g) = O(R^2)$. Поэтому лемма Р4 утверждает существование оценочного алгоритма G с успешной оценкой ϵ/N и $T(G) \leq T(A) + O(NR^2)$. Построить G по A можно, используя метод из упр. 41. Этот алгоритм G имеет такое свойство, что $s = \mathrm{E}\big((-1)^{G(y,z,m)+z\cdot x}\big) \geq (\frac{1}{2} + \epsilon/N) - (\frac{1}{2} - \epsilon/N) = 2\epsilon/N$, где среднее значение взято по всем $(x, z, m) \in X$ и где $(y, z, m) = g(x, z, m)$.

Требуемый алгоритм F получается следующим образом. Задано случайное число $M = PQ$ с неизвестными P и Q. Алгоритм вычисляет случайное число X_0 между 0 и M и немедленно останавливается с известным разложением, если $\gcd(X_0, M) \neq 1$. В других случаях применяется алгоритм L с $G(z) = G(X_0^2 \bmod M, z, M)$ и $k = \lceil \lg(1 + 2N^2R/\epsilon^2) \rceil$. Если одно из 2^k значений x на его выходе удовлетворяет $x^2 \equiv X_0^2$ (по модулю M), существует 50:50 шансов, что $x \not\equiv \pm X_0$. Тогда $\gcd(X_0 - x, M)$ и $\gcd(X_0 + x, M)$ являются простыми множителями M (см. "SQRT-ящик" Рабина (Rabin) в разделе 4.5.4).

Ясно, что время счета этого алгоритма равно $O(N^2R^2\epsilon^{-2}T(A) + N^3R^4\epsilon^{-2})$, так как $\epsilon \geq 2^{-N}$. Вероятность, что алгоритм достигнет цели в разложении M, можно оценить следующим образом. Пусть $n = |X|/2^R$ — число выборов (x, m) и пусть $s_{xm} = 2^{-R}\sum(-1)^{G(y,z,m)+z\cdot x}$ — суммирование по всем содержащим R двоичных разрядов числам z. Тогда $s = \sum_{x,m} s_{xm}/n \geq 2\epsilon/N$. Пусть t — число таких (x, m), что $s_{xm} \geq \epsilon/N$. Вероятность, что наш алгоритм оперирует подобными парами (x, m), равна

$$\frac{t}{n} \geq \sum_{x,m} [s_{xm} \geq \epsilon/N] \frac{s_{xm}}{n} = \sum_{x,m} \big(1 - [s_{xm} < \epsilon/N]\big) \frac{s_{xm}}{n}$$
$$\geq \frac{2\epsilon}{N} - \sum_{x,m} [s_{xm} < \epsilon/N] \frac{s_{xm}}{n} \geq \frac{\epsilon}{N}.$$

И в таком случае алгоритм по теореме G найдет x с вероятностью $\geq \frac{1}{2}$, поскольку мы имеем $2^k > \lceil 2R/s_{xm}^2 \rceil$. Значит, он находит множитель с вероятностью $\geq \frac{1}{4}$. ▮

Что дает теорема P с практической точки зрения? Наше доказательство показывает, что константы, включенные в O, малы. Предположим, что время счета для разложения на множители равно самое большее $10(N^2R^2\epsilon^{-2}T(A) + N^3R^4\epsilon^{-2})$. Многие известнейшие математики мира работали над проблемой разложения на множители больших чисел, в особенности после того, как в конце 70-х годов было показано, что разложение на множители в высшей степени связано с криптографией. Так как они не могли найти хорошее решение, мы имеем основание считать, что разложение на множители является трудным делом. Следовательно, теорема P показывает, что $T(A)$ должно быть большим для всех алгоритмов, которые обнаруживают неслучайность двоичных разрядов, полученных смешанно-квадратичным методом.

Длительные вычисления удобно измерять в MIP-годах (это число операций, выполняемых за год машиной, которая совершает миллион операций в секунду, т. е. $31{,}556{,}952{,}000{,}000 \approx 3.16 \times 10^{13}$). В 1995 году время разложения на множители числа из 120 десятичных цифр (400 двоичных разрядов) при использовании в высшей степени совершенных алгоритмов было больше 250 MIP-лет. Наиболее оптимистически настроенные исследователи, работающие над разложением на множители, могут удивиться, если алгоритм обнаружит, что требуется всего $\exp\bigl(R^{1/4}(\ln R)^{3/4}\bigr)$ команд, когда $R \to \infty$. Только допустим, что это количество может быть достигнуто для по крайней мере не слишком малых частей целых чисел Блюма M, состоящих из R двоичных разрядов. Тогда можно будет умножить много чисел, состоящих из приблизительно 50 000 двоичных разрядов (15 000 цифр), за 2×10^{25} MIP-лет. Если генерировать $N = 1000$ случайных двоичных разрядов смешанно-квадратичным методом с $R = 50000$ и если предположить, что все алгоритмы достаточно хороши, то умножение по крайней мере $\frac{1}{400000}$ на состоящие из 50 000 двоичных разрядов числа Блюма должно выполняться минимум 2×10^{25} MIP-лет. Из теоремы P следует, что каждое такое множество из 1 000 двоичных разрядов проходит все статистические критерии на случайность, время счета $T(A)$ которых меньше 70 000 MIP-лет: не существует алгоритма A, который мог бы отличить такие двоичные разряды от истинно случайной последовательности с вероятностью $\geq \epsilon = \frac{1}{100}$.

Впечатляет? Нет. Такой результат вряд ли является сюрпризом, так как необходимо точно определить около 150 000 истинно случайных двоичных разрядов, точно начинающихся в смешанно-квадратичном методе с X_0, Z и M, когда $R = 50000$. *Конечно*, можно получить 1 000 случайных двоичных разрядов из такого вклада!

Но вообще, формула

$$T(A) \geq \frac{1}{100000} N^{-2}R^{-2}\exp\bigl(R^{1/4}(\ln R)^{3/4}\bigr) - NR^2$$

справедлива при наших умеренных предположениях, когда $\epsilon = \frac{1}{100}$, NR^2 членов являются незначительными и когда R велико. Положим, $R = 200000$ и $N = 10^{10}$. Тогда смешанно-квадратичным методом получим десять миллиардов псевдослучайных двоичных разрядов из $3R \approx 600000$ истинно случайных двоичных разрядов, проходящих все статистические критерии, которые требуют меньше 7.486×10^{10} MIP-лет, что равно 74.86 гигаMIP-годам. При $N = 10^{13}$ и $R = 333333$ время

вычисления, необходимое для определения статистического смещения, возрастает до 53.5 тераMIP-лет.

Простой псевдослучайный генератор 3.2.2–(16), который избегает случайной маски Z, что также можно показать, проходит все полиномиальные критерии случайности, если разложение на множители трудно осуществить (см. упр. 4.5.4–43). Но известные преобразования гарантируют для методов, которые отчасти слабее смешанно-квадратичного метода, порядок роста $O\big(N^4R\epsilon^{-4}\log(NR\epsilon^{-1})\big)$ по сравнению с $O(N^2R^2\epsilon^{-2})$ теоремы P.

Каждый думает, что не существует алгоритма разложения на множители для чисел, состоящих из R двоичных разрядов, время счета которых равно полиному в степени R. Если это предположение верно в строгом виде, то нельзя будет получить даже $1/R^k$ для целого числа Блюма, состоящего из R двоичных разрядов, за полиномиальное время для любого фиксированного k. Теорема P доказывает, что смешанно-квадратичный метод генерирует псевдослучайные числа, проходящие все полиномиальные критерии случайности.

Сформулируем это другим способом: если генерировать случайные двоичные разряды смешанно-квадратичным методом для подходящим образом выбранных N и R, можно также получить числа, проходящие все разумные статистические критерии, или открыть новый алгоритм разложения на множители.

G. Выводы, история и библиография. Выше были определены различные степени случайности, которыми может обладать последовательность.

Конечная ∞-распределенная последовательность удовлетворяет великому множеству полезных свойств, которыми могут обладать случайные последовательности, и существует огромная теория, посвященная ∞-распределенным последовательностям. (В упражнениях, которые приводятся ниже, развиваются некоторые важные, не упомянутые в разделе, свойства таких последовательностей.) Определение R1 поэтому является подходящей основой для теоретического изучения случайности.

Понятие "∞-распределение b-ичной последовательности" было введено в 1909 году Эмилем Борелем (Emile Borel). Он, по существу, ввел понятие (m,k)-распределенной последовательности и показал, что b-ичное представление почти всех действительных чисел является (m,k)-распределенным для всех m и k. Борель назвал такие числа *нормальными* по отношению к основанию b. Превосходное обсуждение этой темы появилось в его хорошо известной книге *Leçons sur la Théorie des Fonctions*, 2nd edition (1914), 182–216.

Понятие ∞-распределенной последовательности *действительных* чисел, также носящее название *полностью равнораспределенной последовательности*, впервые появилось в заметке Н. М. Коробова (*Доклады Акад. Наук СССР* **62** (1948), 21–22). Коробов и несколько его коллег достаточно широко развили теорию таких последовательностей в ряде статей в течение 50-х годов. Полностью равнораспределенные последовательности независимо изучались Джоэлем Н. Франклиным (Joel N. Franklin, *Math. Comp.* **17** (1963), 28–59) в статье, которая особенно заслуживает внимания, так как она была вдохновлена проблемой генерирования случайных чисел. Книга L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences* (New York: Wiley, 1974) является чрезвычайно полным источником информации об огромной математической литературе, содержащей k-распределенные последовательности всех видов.

Тем не менее мы увидели, что ∞-распределенные последовательности не обладают достаточным количеством свойств, чтобы их можно было считать совершенно "случайными". Определения R4–R6, приведенные выше, предусматривают дополнительные условия; в частности, определение R6, видимо, было подходящим способом определения понятия бесконечной случайной последовательности. Это точное количественное утверждение, хорошо совпадающее с интуитивным понятием истинной случайности.

Исторически развитие этих определений было стимулировано, главным образом, поисками Р. фон Мизесом (R. von Mises) хорошего определения вероятности. В *Math. Zeitschrift* **5** (1919), 52–99, фон Мизес предложил определение, по духу сходное с определением R5, хотя формулировка была слишком строга (подобно нашему определению R3), так что последовательностей, удовлетворяющих этим условиям, не существует. Многие исследователи отмечали это противоречие, и А. Г. Коуплэнд (A. H. Copeland, *Amer. J. Math.* **50** (1928), 535–552) предлагал ослабить определение фон Мизеса заменой, которую он назвал допустимыми числами (или последовательностями Бернулли). Существует эквивалент ∞-распределенных $[0\,..\,1)$-последовательностей, в которых все входные U_n заменяются 1, если $U_n < p$ или 0 и если $U_n \geq p$ для заданной вероятности p. Так, Коуплэнд, по существу, предложил вернуться к определению R1. Затем Абрахам Вальд (Abraham Wald) показал, что нет необходимости так решительно ослаблять определение фон Мизеса, и предложил заменить счетное множество правилами подпоследовательностей. В важной статье *Ergebnisse eines math. Kolloquiums* **8** (Vienna, 1937), 38–72, Вальд, по существу, доказал теорему W, хотя он сделал ошибочный вывод, что последовательность, построенная алгоритмом W, также удовлетворяет сильным условиям — $\Pr(U_n \in A) =$ мера A для всех измеримых по Лебегу $A \subseteq [0\,..\,1)$. Заметим, что не существует последовательности, которая может удовлетворять этому условию.

Понятие "вычисляемость" играло большую роль на ранней стадии, когда Вальд написал статью и А. Черч (A. Church, *Bull. Amer. Math. Soc.* **46** (1940), 130–135) показал, как точное понятие "эффективный алгоритм" можно присоединить к теории Вальда, делая его определение совершенно строгим. Практически тогда же дополнение к определению R6 было предложено А. Н. Колмогоровым [*Sankhyā* **A25** (1963), 369–376], как и определение Q2 для конечных последовательностей. Другое определение случайности для конечных последовательностей, находящееся где-то между определениями Q1 и Q2, намного раньше сформулировал А. С. Безикович (A. S. Besicovitch, *Math. Zeitschrift* **39** (1934), 146–156).

В публикациях Черча и Колмогорова рассматривались только двоичные последовательности, для которых $\Pr(X_n = 1) = p$ с заданной вероятностью p. В этом разделе анализировалась более общая ситуация, поскольку $[0\,..\,1)$-последовательность, по существу, представляет все p сразу. Определение фон Мизеса-Вальда-Черча еще одним интересным способом усовершенствовал Дж. В. Говард (J. V. Howard, *Zeitschr. für math. Logik und Grundlagen der Math.* **21** (1975), 215–224).

Следующий важный вклад был сделан Дональдом В. Лавлендом (Donald W. Loveland, *Zeitschr. für math. Logik und Grundlagen der Math.* **12** (1966), 279–294), который обсудил определения R4–R6 и несколько других понятий. Лавленд доказал, что существуют R5-случайные последовательности, не удовлетворяющие определению R4. В связи с этим он установил, что необходимо более строгое определение,

такое как R6. На самом деле Лавленд определил простую перестановку $\langle f(n)\rangle$ неотрицательных целых чисел и алгоритм W′, сходный с алгоритмом W, такой, что

$$\overline{\Pr}(U_{f(n)} \ge \tfrac{1}{2}) - \underline{\Pr}(U_{f(n)} \ge \tfrac{1}{2}) \ge \tfrac{1}{2}$$

для каждой R5-случайной последовательности $\langle U_n\rangle$, выдаваемой алгоритмом W′, когда он задан бесконечным множеством правил подпоследовательностей $\mathcal{R}_k$

Хотя определение R6 интуитивно строже определения R4, очевидно, строго доказать это совсем не просто. В течение нескольких лет данный вопрос оставался открытым, поскольку R4 так или иначе включает в себя R6. В конце концов, Томас Герцог (Thomas Herzog) и Джеймс К. Оуингс (мл.) (James C. Owings, Jr.) сумели построить большое семейство последовательностей, удовлетворяющих R4, но не удовлетворяющих R6. [См. *Zeitschr. für math. Logik und Grundlagen der Math.* **22** (1976), 385–389.]

Другую важную статью написал Колмогоров [*Проблемы передачи информации* **1** (1965), 3–11]. В ней он рассмотрел проблему определения "информационного содержимого" последовательности, и эта работа привела к интересному определению Чайтина (Chaitin) и Мартин-Лёфа (Martin-Löf) конечных случайных последовательностей через "хаотичность". [См. *IEEE Trans.* **IT-14** (1968), 662–664.] Их идея может быть также прослежена в работах Р. Дж. Соломонова (R. J. Solomonoff, *Information and Control* **7** (1964), 1–22, 224–254; *IEEE Trans.* **IT-24** (1978), 422–432; *J. Comp. System Sci.* **55** (1997), 73–88).

Обсуждение случайных последовательностей с философской точки зрения можно найти у К. Р. Поппера (K. R. Popper, *The Logic of Scientific Discovery* (London, 1959)); особенно интересно построение на с. 162–163, впервые опубликованное в 1934 году.

Дальнейшие связи между случайными последовательностями и теорией рекурсивных функций исследовались в работе D. W. Loveland, *Trans. Amer. Math. Soc.* **125** (1966), 497–510. См. также работу К.-П. Шнорр (C.-P. Schnorr, *Zeitschr. Wahr. verw. Geb.* **14** (1969), 27–35), нашедшего сильные связи между случайными последовательностями и "категориями меры", которые были определены Л. Э. Я. Брауэром (L. E. J. Brouwer) в 1919 году. В следующей книге Шнорра, *Zufälligkeit und Wahrscheinlichkeit* [*Lecture Notes in Math.* **218** (Berlin: Springer, 1971)], дан подробный обзор всей темы случайности и превосходное вступление к новым публикациям по этой теме. Обзор важнейших усовершенствований в течение следующих двух десятилетий можно найти в книге *An Introduction to Kolmogorov Complexity and Its Applications* (Springer, 1993), Ming Li and Paul M. B. Vitányi.

Основы теории псевдослучайных последовательностей и эффективной информации заложены Мануэлем Блюмом (Manuel Blum), Сильвио Микали (Silvio Micali) и Эндре Яо (Andrew Yao) в работах [*FOCS* **23** (1982), 80–91, 112–117; *SICOMP* **13** (1984), 850–864], в которых построены первые явные последовательности, удовлетворяющие всем возможным статистическим критериям. Блюм и Микали ввели понятие жесткого ядра двоичного разряда, булевой функции f, такой, что $f(x)$ и $g(x)$ легко вычисляются, хотя функция $f(g^{[-1]}(x))$ не вычисляется. В их статье берет начало лемма P4. Леонид Левин развил теорию в работе *Combinatorica* **7** (1987), 357–363. Затем он и Одед Голдрейч (Oded Goldreich) [*STOC* **21** (1989), 25–32] проанализировали такие алгоритмы, как смешанно-квадратичный метод, и

показали, что, используя маску подобным образом, можно получить жесткое ядро во многих случаях. Наконец, Левин в работе *J. Symbolic Logic* **58** (1993), 1102–1103, усовершенствовал методы предыдущей работы, введя алгоритм L и проанализировав его.

Свой вклад в теорию внесли многие авторы — особенно Импаглиаззо (Impagliazzo), Левин, Лаби (Luby) и Хастад (Håstad) [*STOC* **21** (1989), 12–24; **22** (1990), 395–404], которые показали, что псевдослучайные последовательности можно построить из любой однозначной функции. Однако такие результаты здесь не рассматриваются, так как они применяются, главным образом, в сложной абстрактной теории, а не в практическом генерировании случайных чисел. Практическое применение теоретических работ к псевдослучайности впервые эмпирически исследовано в работе P. L'Ecuyer and R. Proulx, *Proc. Winter Simulation Conf.* **22** (1989), 467–476.

Если числа не случайны, то
они по крайней мере в полном беспорядке.

— ДЖОРДЖ МАРСАЛЬЯ (GEORGE MARSAGLIA) (1984)

УПРАЖНЕНИЯ

1. [*10*] Может ли периодическая последовательность быть равнораспределенной?

2. [*10*] Рассмотрите периодическую двоичную последовательность $0, 0, 1, 1, 0, 0, 1, 1, \ldots$. Она 1-, 2- или 3-распределенная?

3. [*M22*] Постройте троичную периодическую 3-распределенную последовательность.

4. [*HM14*] Докажите, что $\Pr(S(n)\text{ и }T(n)) + \Pr(S(n)\text{ или }T(n)) = \Pr(S(n)) + \Pr(T(n))$ для любых двух утверждений $S(n)$ и $T(n)$, предполагая, что по крайней мере три из этих пределов существуют. Например, если последовательность 2-распределена, то можно найти, что

$$\Pr(u_1 \le U_n < v_1 \text{ или } u_2 \le U_{n+1} < v_2) = v_1 - u_1 + v_2 - u_2 - (v_1 - u_1)(v_2 - u_2).$$

▶ **5.** [*HM22*] Пусть $U_n = (2^{\lfloor \lg(n+1) \rfloor}/3) \bmod 1$. Чему равна $\Pr(U_n < \frac{1}{2})$?

6. [*HM23*] Пусть $S_1(n)$, $S_2(n)$, ... — бесконечная последовательность утверждений о совместных непересекающихся событиях, т. е. $S_i(n)$ и $S_j(n)$ не могут выполняться одновременно, если $i \ne j$. Предположим, что $\Pr(S_j(n))$ существует для каждого $j \ge 1$. Покажите, что $\underline{\Pr}(S_j(n)$ выполняется для некоторого $j \ge 1) \ge \sum_{j\ge1} \Pr(S_j(n))$, и приведите пример, показывающий, что равенство может не выполняться.

7. [*HM27*] Пусть $\{S_{ij}(n)\}$ — семейство утверждений, таких, что $\Pr(S_{ij}(n))$ существует для всех $i, j \ge 1$. Предположим, что для всех $n > 0$ $S_{ij}(n)$ выполняется для точно одной пары целых чисел i, j. Если $\sum_{i,j\ge1} \Pr(S_{ij}(n)) = 1$, то следует ли из этого, что "$\Pr(S_{ij}(n)$ выполняется для некоторого $j \ge 1)$" существует для всех $i \ge 1$ и равна $\sum_{j\ge1} \Pr(S_{ij}(n))$?

8. [*M15*] Докажите (13).

9. [*HM20*] Докажите лемму E. [*Указание.* Рассмотрите $\sum_{j=1}^{m}(y_{jn} - \alpha)^2$.]

▶ **10.** [*HM22*] Где в доказательстве теоремы C используется тот факт, что m делит q?

11. [*M10*] Применяя теорему C, докажите, что если последовательность $\langle U_n \rangle$ ∞-распределена, то она является подпоследовательностью $\langle U_{2n} \rangle$.

12. [*HM20*] Покажите, что k-распределенная последовательность удовлетворяет критерию "максимум-k" в следующем смысле: $\Pr(u \le \max(U_n, U_{n+1}, \ldots, U_{n+k-1}) < v) = v^k - u^k$.

▶ **13.** [*HM27*] Покажите, что ∞-распределенная $[0\,..\,1)$-последовательность проходит критерий интервалов в следующем смысле: если $0 \le \alpha < \beta \le 1$ и $p = \beta - \alpha$, пусть $f(0) = 0$ и для $n \ge 1$ пусть $f(n)$ — наименьшее целое число $m > f(n-1)$, такое, что $\alpha \le U_m < \beta$, тогда

$$\Pr(f(n) - f(n-1) = k) \;=\; p(1-p)^{k-1}.$$

14. [*HM25*] Покажите, что ∞-распределенная последовательность проходит критерий монотонности в следующем смысле: если $f(0) = 0$ и если для $n \ge 1$ $f(n)$ — наименьшее целое число $m > f(n-1)$, такое, что $U_{m-1} > U_m$, тогда

$$\Pr(f(n) - f(n-1) = k) \;=\; 2k/(k+1)! - 2(k+1)/(k+2)!.$$

▶ **15.** [*HM30*] Покажите, что ∞-распределенная последовательность проходит критерий собирания купонов, в котором существует только два вида купонов, в следующем смысле: пусть $X_1, X_2, \ldots$ — ∞-распределенная двоичная последовательность. Пусть $f(0) = 0$ и для $n \ge 1$ пусть $f(n)$ — наименьшее целое число $m > f(n-1)$, такое, что $\{X_{f(n-1)+1}, \ldots, X_m\}$ является множеством $\{0, 1\}$. Докажите, что $\Pr(f(n) - f(n-1) = k) = 2^{1-k}$ для $k \ge 2$ (см. упр. 7).

16. [*HM38*] Выполняется ли критерий собирания купонов для ∞-распределенных последовательностей, когда существует больше двух видов купонов? (См. предыдущее упражнение.)

17. [*HM50*] Для любого заданного рационального числа r Франклин (Franklin) доказал, что последовательность $\langle r^n \bmod 1\rangle$ не является 2-распределенной. Но существует ли рациональное число r, для которого эта последовательность равнораспределена? В частности, равнораспределена ли последовательность при $r = \frac{3}{2}$? [См. K. Mahler, *Mathematika* **4** (1957), 122–124.]

▶ **18.** [*HM22*] Докажите, что если $U_0, U_1, \ldots$ k-распределена, то такой же будет последовательность $V_0, V_1, \ldots$, где $V_n = \lfloor nU_n \rfloor / n$.

19. [*HM35*] Рассмотрите модификацию определения R4, требуя, чтобы подпоследовательность была только 1-распределенной, а не ∞-распределенной. Существует ли последовательность, удовлетворяющая этому более слабому определению, которая не будет ∞-распределенной? (Действительно ли это определение более слабое?)

▶ **20.** [*HM36*] (Н. Г. де Брейн (N. G. de Bruijn) и П. Эрдеш (P. Erdős).) Первые n точек любой $[0\,..\,1)$-последовательности $\langle U_n\rangle$ с $U_0 = 0$ делят интервал $[0\,..\,1)$ на n подынтервалов. Пусть эти подынтервалы имеют длины $l_n^{(1)} \ge l_n^{(2)} \ge \cdots \ge l_n^{(n)}$. Очевидно, что $l_n^{(1)} \ge \frac{1}{n} \ge l_n^{(n)}$, поскольку $l_n^{(1)} + \cdots + l_n^{(n)} = 1$. Одним из способов измерения равномерности распределения $\langle U_n\rangle$ является рассмотрение пределов

$$\overline{L} = \limsup_{n\to\infty} n l_n^{(1)} \qquad \text{и} \qquad \underline{L} = \liminf_{n\to\infty} n l_n^{(n)}.$$

a) Чем являются $\overline{L}$ и $\underline{L}$ для последовательности Ван дер Корпута (van der Corput) (29)?

b) Покажите, что $l_{n+k-1}^{(1)} \ge l_n^{(k)}$ для $1 \le k \le n$. Используйте этот результат для доказательства того, что $\overline{L} \ge 1/\ln 2$.

c) Докажите, что $\underline{L} \le 1/\ln 4$. [*Указание.* Для каждого n существуют такие числа a_1, $\ldots$, a_{2n}, что $l_{2n}^{(k)} \ge l_{n+a_k}^{(n+a_k)}$ для $1 \le k \le 2n$. Кроме того, каждое целое число $2, \ldots, n$ появляется самое большее дважды в $\{a_1, \ldots, a_{2n}\}$.]

d) Покажите, что последовательность $\langle W_n\rangle$, определенная равенством $W_n = \lg(2n+1) \bmod 1$, удовлетворяет $1/\ln 2 > nl_n^{(1)} \ge nl_n^{(n)} > 1/\ln 4$ для всех n. Следовательно, она достигает оптимальных значений $\overline{L}$ и $\underline{L}$.

21. [*HM40*] (Л. Г. Рамшоу (L. H. Ramshaw).)

a) Продолжаем предыдущее упражнение. Будет ли последовательность $\langle W_n \rangle$ равнораспределена?

b) Покажите, что $\langle W_n \rangle$ является только $[0\,..\,1)$-последовательностью, для которой выполняется $\sum_{j=1}^{k} l_n^{(j)} \le \lg(1 + k/n)$ всякий раз, как только $1 \le k \le n$.

c) Пусть $\langle f_n(l_1, \ldots, l_n) \rangle$ — любая последовательность непрерывных функций на множествах строк размерности n $\{(l_1, \ldots, l_n) \mid l_1 \ge \cdots \ge l_n$ и $l_1 + \cdots + l_n = 1\}$, удовлетворяющая следующим двум свойствам:

$$f_{mn}(\tfrac{1}{m}l_1, \ldots, \tfrac{1}{m}l_1, \ \ldots\ , \tfrac{1}{m}l_n, \ldots, \tfrac{1}{m}l_n) = f_n(l_1, \ldots, l_n);$$

$$\text{если} \quad \sum_{j=1}^{k} l_j \ge \sum_{j=1}^{k} l'_j \text{ для } 1 \le k \le n, \quad \text{то} \quad f_n(l_1, \ldots, l_n) \ge f_n(l'_1, \ldots, l'_n).$$

[Примеры: $nl_n^{(1)}$; $-nl_n^{(n)}$; $l_n^{(1)}/l_n^{(n)}$; $n(l_n^{(1)2} + \cdots + l_n^{(n)2})$.] Пусть

$$\overline{F} = \limsup_{n\to\infty} f_n(l_n^{(1)}, \ldots, l_n^{(n)})$$

для последовательности $\langle W_n \rangle$. Покажите, что $f_n(l_n^{(1)}, \ldots, l_n^{(n)}) \le \overline{F}$ для всех n относительно $\langle W_n \rangle$, а также $\limsup_{n\to\infty} f_n(l_n^{(1)}, \cdots, l_n^{(n)}) \ge \overline{F}$ относительно каждой другой $[0\,..\,1)$-последовательности.

▸ **22.** [*HM30*] (Герман Вейль (Hermann Weyl).) Покажите, что $[0\,..\,1)$-последовательность $\langle U_n \rangle$ k-распределена тогда и только тогда, когда

$$\lim_{N\to\infty} \frac{1}{N} \sum_{0\le n<N} \exp(2\pi i(c_1U_n + \cdots + c_kU_{n+k-1})) = 0$$

для каждого множества не всех равных нулю целых чисел c_1, c_2, $\ldots$, c_k.

23. [*M32*] (a) Покажите, что $[0\,..\,1)$-последовательность $\langle U_n \rangle$ k-распределена тогда и только тогда, когда все последовательности $\langle (c_1U_n + c_2U_{n+1} + \cdots + c_kU_{n+k-1}) \bmod 1 \rangle$ 1-распределены всякий раз, когда c_1, c_2, $\ldots$, c_k — целые числа, не все равные нулю. (b) Покажите, что b-ичная последовательность $\langle X_n \rangle$ k-распределена тогда и только тогда, когда все последовательности $\langle (c_1X_n + c_2X_{n+1} + \cdots + c_kX_{n+k-1}) \bmod b \rangle$ 1-распределены всякий раз , когда c_1, c_2, $\ldots$, c_k — целые числа с $\gcd(c_1, \ldots, c_k) = 1$.

▸ **24.** [*M35*] (Й. Г. Ван дер Корпут (J. G. van der Corput).) (a) Докажите, что $[0\,..\,1)$-последовательность $\langle U_n \rangle$ равнораспределена всегда, когда последовательности $\langle (U_{n+k} - U_n) \bmod 1 \rangle$ равнораспределены для всех $k > 0$. (b) Следовательно, $\langle (\alpha_d n^d + \cdots + \alpha_1 n + \alpha_0) \bmod 1 \rangle$ равнораспределена, когда $d > 0$ и α_d — иррациональные числа.

25. [*HM20*] Последовательность называется белой, если все сериальные коэффициенты корреляции равны нулю, т. е. если соотношение в следствии S выполняется для *всех* $k \ge 1$. (Согласно следствию S ∞-распределенная последовательность является белой.) Покажите, что если $[0\,..\,1)$-последовательность равнораспределена, то она белая тогда и только тогда, когда

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0\le j<n} (U_j - \tfrac{1}{2})(U_{j+k} - \tfrac{1}{2}) = 0 \qquad \text{для всех } k \ge 1.$$

26. [*HM34*] (Дж. Франклин (J. Franklin).) Белая последовательность, определенная в предыдущем упражнении, может не быть случайной. Пусть U_0, U_1, $\ldots$ — ∞-распределенная последовательность. Определим последовательность V_0, V_1, $\ldots$ следующим образом:

$$(V_{2n-1}, V_{2n}) = (U_{2n-1}, U_{2n}), \qquad \text{если } (U_{2n-1}, U_{2n}) \in G,$$
$$(V_{2n-1}, V_{2n}) = (U_{2n}, U_{2n-1}), \qquad \text{если } (U_{2n-1}, U_{2n}) \notin G,$$

где G — множество

$$\{(x,y) \mid x - \tfrac{1}{2} \le y \le x \text{ или } x + \tfrac{1}{2} \le y\}.$$

Покажите, что (a) последовательность V_0, V_1, ... равнораспределенная и белая, (b) $\Pr(V_n > V_{n+1}) = \frac{5}{8}$. (Это указывает на слабость критерия сериальной корреляции.)

27. [*HM48*] Чему равно наибольшее возможное значение для $\Pr(V_n > V_{n+1})$ в равнораспределенной белой последовательности? (Д. Копперсмит (D. Coppersmith) построил такую последовательность, для которой это значение достигает $\frac{7}{8}$.)

▶ **28.** [*HM21*] Воспользуйтесь последовательностью (11), чтобы построить 3-распределенную $[0\,..\,1)$-последовательность, для которой $\Pr(U_{2n} \ge \frac{1}{2}) = \frac{3}{4}$.

29. [*HM34*] Пусть X_0, X_1, ... — $(2k)$-распределенная двоичная последовательность. Покажите, что

$$\overline{\Pr}(X_{2n} = 0) \le \frac{1}{2} + \binom{2k-1}{k} \Big/ 2^{2k}.$$

▶ **30.** [*M39*] Постройте $(2k)$-распределенную двоичную последовательность, для которой

$$\Pr(X_{2n} = 0) = \frac{1}{2} + \binom{2k-1}{k} \Big/ 2^{2k}.$$

(Таким образом, неравенство в предыдущем упражнении является наилучшим.)

31. [*M30*] Покажите, что существуют $[0\,..\,1)$-последовательности, удовлетворяющие определению R5, однако $\nu_n/n \ge \frac{1}{2}$ для всех $n > 0$, где ν_n — число $j < n$, для которых $U_n < \frac{1}{2}$. (Это можно рассматривать как неслучайное свойство последовательности.)

32. [*M24*] Задана $\langle X_n \rangle$ "случайная" b-ичная последовательность, удовлетворяющая определению R5, и $\mathcal{R}$ — исчисляемое правило подпоследовательности, точно задающее бесконечную подпоследовательность $\langle X_n \rangle \mathcal{R}$. Покажите, что последняя подпоследовательность не только 1-распределена, но и "случайна" согласно определению R5.

33. [*HM22*] Пусть $\langle U_{r_n} \rangle$ и $\langle U_{s_n} \rangle$ — бесконечные непересекающиеся подпоследовательности последовательности $\langle U_n \rangle$. (Иначе говоря, $r_0 < r_1 < r_2 < \cdots$ и $s_0 < s_1 < s_2 < \cdots$ — возрастающие последовательности целых чисел и $r_m \ne s_n$ для любых m, n.) Предположим, что $\langle U_{t_n} \rangle$ — комбинированная подпоследовательность, такая, что $t_0 < t_1 < t_2 < \cdots$, и положим $\{t_n\} = \{r_n\} \cup \{s_n\}$. Покажите, что если $\Pr(U_{r_n} \in A) = \Pr(U_{s_n} \in A) = p$, то $\Pr(U_{t_n} \in A) = p$.

▶ **34.** [*M25*] Определите правила подпоследовательностей $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$, ..., такие, что с этими правилами можно использовать алгоритм W, чтобы задать эффективный алгоритм построения $[0\,..\,1)$-последовательности, удовлетворяющей определению R1.

▶ **35.** [*HM35*] (Д. В. Лавленд (D. W. Loveland).) Покажите, что если двоичная последовательность $\langle X_n \rangle$ R5-случайна и если $\langle s_n \rangle$ — любая исчисляемая последовательность соответственно с определением R4, то $\overline{\Pr}(X_{s_n} = 1) \ge \frac{1}{2}$ и $\underline{\Pr}(X_{s_n} = 1) \le \frac{1}{2}$.

36. [*HM30*] Пусть $\langle X_n \rangle$ — двоичная последовательность, "случайная" согласно определению R6. Покажите, что $[0\,..\,1)$-последовательность $\langle U_n \rangle$, определенная в двоичной системе счисления по схеме

$$U_0 = (0.X_0)_2, \quad U_1 = (0.X_1X_2)_2, \quad U_2 = (0.X_3X_4X_5)_2, \quad U_3 = (0.X_6X_7X_8X_9)_2, \quad \ldots,$$

случайна в смысле определения R6.

37. [*M37*] (Д. Копперсмит (D. Coppersmith).) Постройте последовательность, удовлетворяющую определению R4, но не определению R5. [*Указание.* Рассмотрите преобразование U_0, U_1, U_4, U_9, ... истинно случайной последовательности.]

38. [*M49*] (A. N. Колмогоров.) Пусть заданы N, n и ϵ. Чему равно наименьшее число алгоритмов в множестве **A**, таких, чтобы не существовали (n,ϵ)-случайные двоичные последовательности длины N по отношению к **A**? (Если нельзя задать точные формулы, можно ли найти асимптотические формулы? Суть этой проблемы — обнаружить, как точная грань (37) становится "наилучшей возможной".)

39. [*HM45*] (B. M. Шмидт (W. M. Schmidt).) Пусть U_n — $[0\,..\,1)$-последовательность и пусть $\nu_n(u)$ — число таких неотрицательных целых чисел $j \le n$, что $0 \le U_j < u$. Докажите, что существует такая положительная постоянная c, что для любого N и любой $[0\,..\,1)$-последовательности $\langle U_n \rangle$ мы получим

$$|\nu_n(u) - un| > c \ln N$$

для некоторых n и u при $0 \le n < N$, $0 \le u < 1$. (Другими словами, никакая $[0\,..\,1)$-последовательность не может быть *слишком* равнораспределена.)

40. [*M28*] Завершите доказательство леммы P1.

41. [*M21*] В лемме P2 показано существование критерия прогноза, но при доказательстве предполагается существование подходящего k без объяснения, как конструктивно находить k из A. Покажите, что любой алгоритм A можно обратить в алгоритм A' с $T(A') \le T(A) + O(N)$, который предсказывает B_N по $B_1 \ldots B_{N-1}$ с вероятностью, по крайней мере равной $\frac{1}{2} + (P(A,S) - P(A,\$_N))/N$ на любом симметричном сдвиге N-источника S.

▸ **42.** [*M28*] (*Попарная независимость.*)

a) Пусть $X_1, \ldots, X_n$ — случайные величины со средним, равным $\mu = \mathrm{E}\,X_j$, и дисперсией $\sigma^2 = \mathrm{E}\,X_j^2 - (\mathrm{E}\,X_j)^2$ при $1 \le j \le n$. Докажите неравенство Чебышева

$$\Pr((X_1 + \cdots + X_n - n\mu)^2 \ge tn\sigma^2) \le 1/t$$

при дополнительных предположениях, что $\mathrm{E}(X_i X_j) = (\mathrm{E}\,X_i)(\mathrm{E}\,X_j)$ всякий раз, когда $i \ne j$.

b) Пусть B — случайная двоичная матрица размера $k \times R$. Докажите, что если c и c' — фиксированные не равные нулю двоичные векторы размерности k, то cB и $c'B$ — независимые случайные двоичные векторы размерности R (по модулю 2).

c) Примените (a) и (b) к анализу алгоритма L.

43. [*20*] Кажется, точно так же тяжело найти множители любого *фиксированного* целого числа Блюма M, состоящего из R двоичных разрядов. Как найти множители *случайного* целого числа, состоящего из R двоичных разрядов? Почему тогда теорема P сформулирована для случайного, а не фиксированного M?

▸ **44.** [*16*] (И. Дж. Гуд (I. J. Good).) Может ли правильная таблица случайных чисел содержать точно одну ошибку?