но построение правила, производящего оценку этих характеристик по наблюдениям реализаций самого функционала. Оценочные расчеты и данные моделирования позволяют выбрать целесообразные значения параметров указанного правила.

Построенный таким образом алгоритм начала — остановки, раздельный по всем настраиваемым параметрам c_j, позволяет автоматически исключать из процесса многопараметрической минимизации малоинформативные параметры (те параметры, к вариациям которых функционал (2) становится малочувствительным), тем самым уменьшая размерность настройки.

ЛИТЕРАТУРА

1. Я. З. Цыпкин. Основы теории обучающихся систем. М., «Наука», 1970.
2. Я. З. Цыпкин. Адаптация и обучение в автоматических системах. М., «Наука», 1968.
3. С. А. Понырко, И. В. Семушин. Использование активного принципа при построении самонастраивающихся фильтров.— Изв. АН СССР, Техническая кибернетика, 1971, № 1.
4. С. А. Понырко, И. В. Семушин. Построение обучающихся винеровских фильтров при ограниченном объеме априорной информации.— Изв. АН СССР, «Техническая кибернетика», 1971, № 5.
5. В. С. Пугачев. Теория случайных функций и ее применение к задачам автоматического управления. М., Физматгиз, 1962.
6. И. В. Семушин. Многоканальный адаптивный фильтр активного типа.— ИВУЗ, Приборостроение, 1969, № 10.

Поступила в редакцию 18 февраля 1972 г.

УДК 518.5+519.2

М. В. АНТИПОВ

(Новосибирск)

О МОДЕЛИРОВАНИИ НЕКОТОРЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

При изучении процессов, происходящих в сложных системах, не всегда удается получить законченные аналитические соотношения, которые можно было бы использовать при вычислениях. В таких случаях часто бывает полезен метод статистических испытаний. Это относится также к вопросам планирования эксперимента, когда требуется описать характер потоков заявок на обслуживание (см., например, [1]).

Для решения задач методом статистических испытаний необходимо иметь последовательности случайных величин с заданной функцией распределения. В работе предлагаются стандартные алгоритмы и программы получения псевдослучайных величин с наиболее распространенными функциями распределения для ЭВМ типа «Минск-22». В соответствии с требованиями, предъявленными к датчикам псевдослучайных чисел [2] при хорошем статистическом качестве и наибольшем периоде псевдопоследовательности, минимизировалось время получения одного псевдослучайного числа.

Алгоритм I. Равномерное распределение в интервале [0,1]. Получению и оценке псевдослучайных чисел с равномерным распределением посвящена весьма обширная литература (см. [3]). Обоснование приведенного алгоритма и начальных данных содержится в [2—5]. В [4] была дана стандартная программа получения псевдослучайных чисел на ЭВМ типа «Минск-22» по формуле

$$x_{n+1} \equiv \lambda x_n (\mathrm{mod}\ 2^{36}), \qquad (1)$$

где x_{n+1} — очередное псевдослучайное число, равномерно распределенное в [0,1]. Автором была предпринята еще одна проверка для $\lambda = 261\,047\,521\,715$, найденного в [2, 5], на равномерность девяти и трех старших разрядов, затем вычислены значения χ^2_{511} и χ^2_7 для выборки объемом 2^{17}, получены положительные результаты для некоторых x_0: 000...001, 00...03, 00...05, 00...0011.

Алгоритм II. Равномерное распределение в интервале [а, в]. Поскольку имеется датчик псевдослучайных чисел, равномерно распределенных в интервале [0, 1], равномерное распределение в интервале $[a, в]$ может быть получено следующей несложной процедурой [3]: $y_n = a + (b - a) x_n$. Здесь x_n вычисляется по алгоритму (1), а y_n — искомое число.

Алгоритм III. Равномерная логическая шкала. При решении логических задач бывает необходим набор из нулей и единиц таких, что вероятность $P\{\xi_i = 0\} = P\{\xi_i = 1\} = 1/2$. Псевдослучайное число x_n алгоритма I не может считаться удовлетворительным, так как статистическое качество разрядов заметно ухудшается в сторону младших разрядов. Можно, разумеется, получить шкалу, «склеивая» старшие разряды нескольких псевдослучайных чисел, однако проще воспользоваться алгоритмом работы [5]

$$y_{n+1} \equiv (\lambda x_n + [\lambda x_n / 2^{36}]) (\mathrm{mod}\ 2^{36}), \qquad (2)$$

или

$$y_{n+1} = (\lambda x_n \otimes [\lambda x_n / 2^{36}]), \qquad (3)$$

где $\otimes$ означает поразрядное сложение (сложение по mod 2).

Алгоритмы (2) и (3) равноправны, что следует из [5]; это подтверждает также следующая статистическая проверка.

А. Было подсчитано число единиц в шкале (36-разрядной ячейке) a_i. Вероятность $P\{a_i = n\} = p_n = C^n_{36}/2^{36}$, если N — объем выборки, а s_n — число опытов с $a_i = n$; тогда $\chi^2_{n_2 - n_1} = \sum\limits_{n=n_1}^{n_2+1} \frac{(s_n - Np_n)^2}{Np_n}$. $\chi^2_{24} = 29{,}4$; 31,7 для (2) и (3) соответственно.

Б. Были подсчитаны величины $b^{с}_i$ и $b^{м}_i$ — число единиц в старших и младших шести разрядах; полученные значения равны $\chi^2_6 (b^{с}_i, b^{м}_i) = 1{,}7$; 7,2 и 2,5; 8,9 для (2) и (3) соответственно.

В. У трех следующих друг за другом значений шкалы выбирались по три разряда старших, средних и младших, затем тройки «склеивались» в девятиразрядные слова, набиралась статистика и подсчитывались три величины χ^2_{511}, оценивающие тройные корреляции старших, средних и младших разрядов. Хорошо известно, что величину $\Phi(\chi^2_{511}) = \frac{\chi^2_{511} - 511}{\sqrt{1022}}$ можно считать распределенной нормально с параметрами (0,1). Результаты проверки:

вариант (2) χ^2_{511} (ст.) $= 538{,}9$ ($\Phi = 0{,}87$); χ^2_{511} (ср.) $= 504{,}0$ ($\Phi = -0{,}22$); χ^2_{511} (мл.) $= 489{,}6$ ($\Phi = -0{,}67$);

вариант (3) χ^2_{511} (ст.) $= 531{,}6$ ($\Phi = 0{,}64$); χ^2_{511} (ср.) $= 506{,}8$ ($\Phi = -0{,}13$); χ^2_{511} (мл.) $= 489{,}5$ ($\Phi = -0{,}67$).

Алгоритм IV. Экспоненциальное распределение. Для реализации экспоненциального распределения e^{-x} достаточно было бы вычислить $-\ln x_n = y_n$ [3], где x_n находится из (1), но прямое вычисление логарифма отнимает много машинного времени.

Метод серий моделирования логарифма Неймана может дать выигрыш во времени на некоторых ЭВМ в 1,5 раза [6—8], но Г. А. Михайловым [6] предложен более экономичный метод — моделируются пять величин x_n^i по схеме (1):

$$
\begin{aligned}
y_n^1 &= -\min\left(x_n^4,\ x_n^5\right)\ln\left(x_n^1x_n^2x_n^3\right);\\
y_n^2 &= -\left|x_n^5 - x_n^4\right|\ln\left(x_n^1x_n^2x_n^3\right);\\
y_n^3 &= -\{1-\max\left(x_n^4,\ x_n^5\right)\}\ln\left(x_n^1x_n^2x_n^3\right).
\end{aligned}
\tag{4}
$$

Три величины y_n^1, y_n^2, y_n^3 будут распределены экспоненциально. Так как метод Неймана не применим для вычисления $\ln\left(x_n^1x_n^2x_n^3\right)$, то логарифм вычисляется по приближенной формуле: пусть $x_n^1x_n^2x_n^3 = 2^{-r}\gamma$, где $1/2 \leqslant \gamma < 1$; $1-\gamma=\varepsilon$; тогда $\ln\left(x_n^1x_n^2x_n^3\right) \approx -r\ln 2 - \left(\varepsilon + \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} + \frac{\varepsilon^4}{4} + +0{,}33667\varepsilon^5\right)$. Коэффициент при ε^5 выбран таким образом, чтобы Δ — абсолютная ошибка вычисления $\ln\gamma$ — была минимальной. Так как множитель λ исходного датчика (1) находился по наибольшему корреляционному коэффициенту, меньшему 1/1000 [2, 5], то достаточна достигнутая точность $\Delta_{\max}(\ln\gamma) < 3{,}3\cdot 10^{-4}$. Пять используемых множителей взяты из [5].

Для проверки алгоритма прямая $(0, \infty)$ разбивалась на десять частей, затем набиралась статистика и находилась величина $X_9^2 = 11{,}1$ и 13,2 для несколько измененных начальных данных.

Алгоритм V. Показательное распределение. Так как функцией показательного распределения является e^{-kx}, то $y_n = -\frac{1}{k}\ln x_n$ и от экспоненциального распределения показательное будет отличаться постоянным множителем $1/k$.

Алгоритм VI. Распределение Пуассона. Распределение Пуассона определяется формулой $P\{y_n = m\} = p_m = \frac{s^m}{m!}e^{-s}$, где $s>0$— параметр распределения Пуассона.

В [8] приводится алгоритм и программа получения чисел Пуассона, основанные на соотношении

$$p_m = P\left\{\sum_{i=0}^{m-1} p_i \leqslant x_n < \sum_{i=0}^{m} p_i\right\},$$

где x_n находится из (1). Поскольку известно, что величина $p_{[s]}$ наибольшая из $p_i (i=0,\ 1,\ 2,\ldots)$, то поиск m ведется последовательным сравнением, начиная с $[s]$, причем сравнение идет по убыванию, если $x_n < \sum_{i=0}^{[s]} p_i$, и по возрастанию — в противном случае. Этот алгоритм требует сравнений в $s+1/(\sqrt{2s/\pi}+2)$ раз меньше. К недостаткам алгоритма надо отнести вычисление на каждом шаге величин $\frac{s^m}{m!}e^{-s}$(два сложения, умножение, деление).

Чтобы избежать многократного повторения одинаковых действий, предлагается табличный метод получения чисел Пуассона. Строится таблица значений $e^{-s}\sum_{i=0}^{m}\frac{s^i}{i!}$ ($m=0, 1, 2, \ldots$). Границы для s и таблицы определяются из технических возможностей ЭВМ следующим образом:

а) так как ненулевое положительное нормализованное число не может быть меньше 2^{-64}, то $s>44$ приводит к «занулению» e^{-s} (машинный ноль);

б) при $s>42$ в процессе вычислений образуется величина $s^{[s+1]}/[s]!>2^{64}$ и происходит переполнение;

в) так как для ЭВМ типа «Минск» $\max\limits_{x<1} x=1-2^{-28}$; то таблица строится до тех пор, пока $P_M=e^{-s}\sum\limits_{i=0}^{M}\frac{s^i}{i!}\leqslant 1-2^{-28}$; тогда P_M полагается равной единице и таблица после занесения в начало величины — ε готова для получения чисел Пуассона;

г) для $s=42$ (наибольшего) общая длина таблицы равна 86, для $s=40$—82, $s=20$—52, $s=10$—34, $s=5$—23, $s=1$—12;

д) предусмотрен останов ЭВМ, если $s<0$ и $s>42$;

е) известно, что даже для $s\ll 42$ с высокой точностью величина $\frac{y_n-s}{\sqrt{s}}$ распределена нормально с параметрами (0,1), поэтому предполагается, что околопредельные значения s будут использоваться редко, ибо число сравнений, а вместе с ним и время получения числа Пуассона пропорциональны $\sqrt{2s/\pi}+2$;

ж) предварительное аппроксимирование числа y_n по числу x_n оправдано только для $s>30$, даже для сравнительно простых аппроксимационных функций (отвергнуто в силу причин, изложенных в п. е).

Для проверки распределения Пуассона были взяты параметры $s=2\pi$ (получено $\mathrm{X}^2_{19}=10,1$ и $\mathrm{X}^2_{19}=13,2$ с измененным начальным x_0) и $s=40$ ($\mathrm{X}^2_{48}=36,6$).

Алгоритм VII. Нормальное распределение с параметрами (0, 1). Плотность нормального распределения задается функцией $f(x)=\frac{1}{\sqrt{2\pi}}\times\times e^{-x^2/2}$, а функция распределения $F(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{t}e^{-x^2/2}dx$.

Как показано в [7, 8], случайные величины

$$\xi=\sqrt{-2\ln x_1}\cos 2\pi x_2 \text{ и } \eta=\sqrt{-2\ln x_1}\sin 2\pi x_2 \tag{5}$$

независимы и распределены нормально с параметрами (0, 1) при условии, что x_1 и x_2 — случайные числа, равномерно распределенные в [0, 1]. О. А. Махоткин в [7] сравнил несколько методов получения нормально распределенных величин и нашел, что наиболее экономичен алгоритм (для ЭВМ БЭСМ-6) — модифицированный Г. А. Михайловым метод Дядькина с заменой моделирования $\ln x$ методом нисходящих серий на моделирование одновременно трех экспоненциально распределенных случайных величин (см. Алгоритм IV).

Для ЭВМ типа «Минск-22» наиболее экономичным является моделирование величин по алгоритму V, где $\ln x_1$ взят из (4), а $\sin 2\pi x_2$ и $\cos 2\pi x_2$ моделируются методом исключения [8]: 1) по алгоритму I получим x_n^6 и x_n^7; 2) вычисляем $\beta=1-2x_n^6$ и $\gamma=1-2x_n^7$; $d=\beta^2+\gamma^2$; 3) если $d>1$, то выполняется 1), иначе $\sin 2\pi x_2=\beta/\sqrt{d}$; $\cos 2\pi x_2=\gamma/\sqrt{d}$.

Так как $P\{d<1\}=\pi/4$, то для получения $\sin 2\pi x_2$ и $\cos 2\pi x_2$ необходимо в среднем 2,55 псевдослучайных чисел, равномерно распределенных в [0, 1]. Из (5) получаем:

$$\xi = \left(1 - 2x_n^6\right) \sqrt{\frac{-2\ln x_1}{d}}\,; \quad \eta = \left(1 - 2x_n^7\right) \sqrt{\frac{-2\ln x_1}{d}}\,. \tag{6}$$

Если $\ln x_1$ находить по методу Г. А. Михайлова [6] и (4), то можно показать, что для получения шести нормально распределенных чисел требуется в среднем 12,6 равномерно распределенных в [0, 1] псевдослучайных чисел, одно вычисление $\ln x$ и три вычисления квадратного корня. Отсутствие в ЭВМ типа «Минск-22» операции извлечения квадратного корня заставляет находить корень приближенно. Пусть $a = \left(\frac{1}{2} + \varepsilon_1\right) \cdot 2^r$ — подкоренное выражение $0 \leqslant \varepsilon_1 < 1/2$. Из (6) легко заметить, что $0 < a < 2^{64}$, т. е. r может быть как положительным, так и отрицательным $\left(\varepsilon = 1 - \frac{1}{2} - \varepsilon_1 = \frac{1}{2} - \varepsilon_1;\ \ 0 < \varepsilon < \frac{1}{2}\right)$. Тогда

$$\sqrt{a} \approx 2^{\operatorname{sgn} r\,\{[|r|/2] + (|r|/2 - [|r|/2])\}} \left(1 - \frac{1}{2}\varepsilon - \frac{1}{8}\varepsilon^2 - 0{,}08842\varepsilon^3\right),$$

причем коэффициент при ε^3 выбран таким образом, чтобы абсолютная ошибка $\Delta_{\max}\left(\sqrt{\frac{1}{2} + \varepsilon_1}\right) < 3{,}4 \cdot 10^{-4}$. В силу причин, изложенных выше, этой точности достаточно.

Для проверки нормального распределения прямая $(-\infty, \infty)$ разбивалась на десять частей, набиралась статистика, затем вычислялась величина $\chi_9^2 = 7{,}9$.

ВЫВОДЫ

Рассмотрены алгоритмы семи наиболее часто встречаемых функций распределений псевдослучайных величин для ЭВМ типа «Минск-22». Критерием для выбора той или иной процедуры была скорость генерирования псевдослучайной величины при высоком статистическом качестве. Все проводимые проверки дали согласие с гипотезой о случайности в пределах 95%.

ЛИТЕРАТУРА

1. А. М. Искольдский, Ю. М. Крендель, О. Е. Трофимов. Статистические характеристики информационного потока в эксперименте по исследованию импульсных процессов.— Автометрия, 1972, № 1.
2. М. В. Антипов. Анализ производящих множителей мультипликативного датчика псевдослучайных чисел. Препринт ИЯФ СО АН СССР. Новосибирск, 1971.
3. С. М. Ермаков. Метод Монте-Карло и смежные вопросы. М., «Наука», 1971.
4. М. В. Антипов, Ф. М. Израйлев, Б. В. Чириков. Статистическая проверка датчика псевдослучайных чисел.— В сб. «Вычислительные системы», вып. 30. Новосибирск, 1968.
5. М. В. Антипов. Оптимальный производящий множитель мультипликативного датчика.— В сб. «Вероятностные методы решения задач математической физики». Новосибирск, ВЦ СО АН СССР, 1971.
6. Г. А. Михайлов. Экономичный способ моделирования показательно распределенной случайной величины.— Тезисы докладов III Всесоюзной конференции по методам Монте-Карло. Новосибирск, 1971.
7. О. А. Махоткин. Численное исследование процедур моделирования нормального и показательного распределений.— В сб. «Вероятностные методы решения задач математической физики». Новосибирск, ВЦ СО АН СССР, 1971.
8. З. А. Королева, Г. А. Михайлов. Оптимальные процедуры моделирования некоторых случайных величин. Препринт ВЦ СО АН СССР. Новосибирск, 1969.

Поступила в редакцию 17 октября 1972 г.