4. Статистическое моделирование систем на ЭВМ

4.2 Моделирование случайных воздействий

Формирование случайных величин с заданным законом распределения

Исходный материал – СВ, имеющая равномерное распределение на (0, 1).

Возможные значения СВ *ξ*, равномерно распределенной на (0, 1)

$$x_i \Rightarrow y_i$$

Возможные значения СВ *η*, имеющей заданный закон распределения

Основные пути преобразования.

1) *Прямой* – реализация некоторой операции над x_i , формирующей число y_i , имеющее (точно или приближенно) заданный закон распределения.
2) Отсеивание чисел из первоначальной случайной последовательности.
3) Моделирование условий соответствующей предельной теоремы теории вероятностей.

Прямое преобразование (метод обратных функций)

Идея построения требуемого преобразования вытекает из следующей теоремы (курс теории вероятностей).

Теорема.

Если СВ η имеет плотность распределения вероятности $f(y)$, то распределение СВ

$$\xi = \int_0^{\eta} f(y)\,dy$$

является равномерным на (0, 1).

- ***Правило построения возможных значений непрерывной СВ.***

Чтобы получить одно из возможных значений $\boldsymbol{y_i}$ СВ $\boldsymbol{\eta}$, имеющей плотность распределения $\boldsymbol{f(y)}$, нужно разрешить относительно $\boldsymbol{y_i}$ уравнение

$$\int_{-\infty}^{y_i} f(y)dy = x_i\,, \qquad \textbf{(4.1)}$$

где $\boldsymbol{x_i}$ – одно из возможных значений равномерно распределенной СВ.

<u>Замечание</u>: (4.1) равносильно $\boldsymbol{F_\eta(y_i) = x_i}$.

Это правило называется также ***методом обратной функции*** (используется преобразование $\boldsymbol{\eta = F_\eta^{-1}(\xi)}$).

В ряде случаев это правило может быть использовано непосредственно.

Пример.

Пусть требуется получить случайные числа $\boldsymbol{y_i}$ с показательным законом распределения

$$\boldsymbol{f(y) = \lambda \mathrm{e}^{-\lambda y}} \quad \text{при} \quad \boldsymbol{y} > 0.$$

В силу (4.1)

$$\lambda \int_0^{y_i} e^{-\lambda y}\, dy = x_i ,$$

где $\boldsymbol{x_i}$ – случайное число, имеющее равномерное на (0, 1) распределение.

После вычисления интеграла

$$1 - e^{-\lambda y_i} = x_i\ ,$$

откуда

$$y_i = -\frac{1}{\lambda}\ln(1 - x_i).$$

Окончательно:

$$\boxed{y_i = -\frac{1}{\lambda}\ln x_i\ .}$$

СВ $\xi_1 = 1 - \xi$ также имеет равномерное на (0, 1) распределение

Недостатки рассмотренного метода:

- в большинстве практически важных случаев уравнение (4.1) не решается точно относительно y_i (пример – нормальное распределение);
- даже в случаях, когда (4.1) разрешено, требуется достаточно много машинных операций для определения y_i (вычисление логарифмов, извлечение корней и т. п.).

- ***Формирование возможных значений дискретной СВ.***

Пусть η – дискретная СВ, имеющая конечное (счетное) число возможных значений

$$y_1, y_2, \ldots, y_s \quad (y_1, y_2, \ldots, y_s, \ldots),$$

$$y_1 < y_2 < \ldots < y_s \quad (y_1 < y_2 < \ldots < y_s < \ldots),$$

вероятности которых равны $p_1, p_2, \ldots, p_s$ $(p_1, \ldots, p_s, \ldots)$,

$$\sum_{k=1}^{s} p_k = 1 \quad \left(\sum_{k=1}^{\infty} p_k = 1\right).$$

Тогда

$$x_i = F_\eta(y_i) = \sum_{k=1}^{m} p_k .$$

Обозначим

$$l_r = \sum_{k=1}^{r} p_k \,, \qquad r = 1, 2, \; \ldots$$

Правило построения последовательности случайных чисел y_i :

1) выбрать число x_i из исходной квазиравномерной совокупности;

2) проверять справедливость неравенств

$$l_{r-1} < x_i \le l_r \,, \qquad \textbf{(4.2)}$$

$r = 1, 2, \ldots ,\qquad l_0 = 0;$

3) если неравенство (4.2) выполнено при некотором r, то $y_i = y_r$.

Пример.

Пусть требуется получить случайные числа y_i с распределением Пуассона

$$P(\eta = n) = \frac{a^n}{n!} e^{-a}, \quad n = 0, 1, 2, \ \ldots$$

Выберем очередное x_i – случайное число, имеющее равномерное распределение на (0, 1).

Проверим справедливость

$$\sum_{k=0}^{n-1} p_k < x_i \le \sum_{k=0}^{n} p_k ,$$

или

$$\underbrace{e^{-a} \sum_{k=0}^{n-1} \frac{a^k}{k!}}_{=0 \text{ при } n=0} < x_i \le e^{-a} \sum_{k=0}^{n} \frac{a^k}{k!}, \quad n = 0, 1, 2, \ \ldots .$$

Если неравенство верно, то $\boldsymbol{y_i} = \boldsymbol{n}$.

Полученная таким образом СВ имеет распределение, близкое к распределению Пуассона.

Метод просеивания фон Неймана

Основная идея:

из равномерно распределенной последовательности случайных чисел отбирается подпоследовательность, имеющая заданный закон распределения.

Пусть требуется получить последовательность случайных чисел $\boldsymbol{y_i}$ с функцией плотности $\boldsymbol{f_\eta(y)}$.

Будем считать, что область определения $\boldsymbol{f_\eta(y)}$ ограничена интервалом $(\boldsymbol{a}, \boldsymbol{b})$;

$\boldsymbol{m}$ – максимальное значение $\boldsymbol{f_\eta(y)}$.

Пусть

- последовательность u_i имеет равномерное распределение на интервале (a, b),
- последовательность v_i имеет равномерное распределение на интервале $(0, m)$.

Для каждой пары (u_i, v_i) проверим выполнение условия

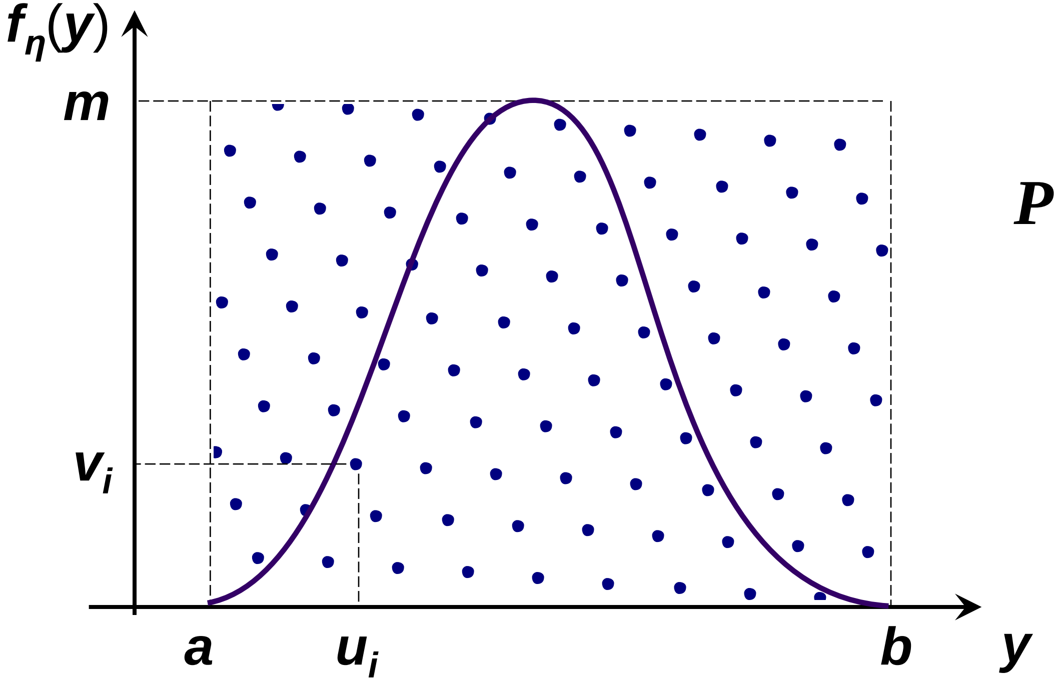
$$v_i \geq f_\eta(u_i).$$

Если это неравенство выполнено, то случайное число u_i должно быть отброшено.

Последовательность случайных чисел u_i, которые не были отвергнуты, имеет плотность распределения $f_\eta(y)$.

Геометрическая иллюстрация.

После просеивания останутся только точки, находящиеся под кривой $v = f_\eta(y)$.



Для таких точек

$$P(U < u_i) = \int_a^{u_i} f(u)du.$$

Геометрическая вероятность попадания в область под кривой левее $U = u_i$

Процедура получения последовательности y_i, имеющей плотность $f_\eta(y)$:

1) выбрать пару чисел (x_{2i-1} , x_{2i}) из исходной квазиравномерной совокупности,
2) для этих чисел проверить справедливость выполнения неравенства

$$x_{2i} \le \frac{1}{m} f_\eta\left(a+(b-a)x_{2i-1}\right), \qquad (*)$$

$v_i \le f_\eta(u_i)$,
$u_i = a + (b - a)x_{2i-1}$,
$v_i = m \cdot x_{2i}$

3) если это неравенство выполнено, то очередное число y_i положить равным

$$y_i = a + (b - a)x_{2i-1} .$$

Замечания.

1. Эффективность данного метода (вероятность принятия числа $\boldsymbol{u_i = a + (b - a)x_{2i-1}}$) может быть повышена, если для формирования СВ $\boldsymbol{v_i}$ использовать не равномерный на (0, $\boldsymbol{m}$) закон распределения, а плотность распределения, близкую к $\boldsymbol{f_\eta(y)}$, но имеющую достаточно простой вид (чтобы можно было использовать метод обратных функций).

2. В случае использования квазиравномерного распределения (вместо равномерного) появляется систематическая погрешность, вызванная дискретностью исходной совокупности. Ошибка вероятности неравенства (*) всегда отрицательна и ограничена величиной $\boldsymbol{m}\cdot 2^{-\boldsymbol{k}}$.

Моделирование условий предельных теорем теории вероятностей

Такие методы ориентированы на получение последовательностей чисел с конкретным законом распределения (не являются универсальными).

- ***Моделирование нормально распределенных случайных чисел***.

Пусть требуется получить последовательность случайных чисел $\boldsymbol{x_i}$, имеющих нормальное распределение с математическим ожиданием $\boldsymbol{m}$ и средним квадратическим отклонением $\boldsymbol{\sigma}$:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}} .$$

На основании ЦПТ случайные числа x_i можно построить в виде сумм последовательностей равномерно распределенных на (0, 1) случайных чисел.

ЦПТ для одинаково распределенных СВ:

если независимые СВ $\xi_1, \xi_2, \ldots$ имеют одно и то же распределение с одним и тем же мат. ожиданием m_1 и средним квадратическим отклонением σ_1, то СВ

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n$$

имеет асимптотически нормальное распределение с параметрами $m = n \cdot m_1$ и $\sigma = \sigma_1 \sqrt{n}$.

Расчеты показывают:

сумма ξ имеет распределение, близкое к нормальному, даже при сравнительно небольших n (практически достаточно $n = 8 \div 12$).

Для СВ ξ_i, имеющих равномерное распределение на (0, 1),

$$m_1 = 0{,}5; \quad \sigma_1 = \frac{1}{2\sqrt{3}}$$

$$\Downarrow$$

сумма ***n*** слагаемых будет иметь мат. ожидание и с.к.о.

$$m_n = \frac{n}{2}; \quad \sigma_n = \frac{1}{2}\sqrt{\frac{n}{3}}.$$

Для квазиравномерного распределения

$$m_1^* = 0{,}5; \quad \sigma_1^* = \frac{1}{2\sqrt{3}}\sqrt{\frac{2^k+1}{2^k-1}},$$

и

$$m_n^* = \frac{n}{2}; \quad \sigma_n = \frac{1}{2}\sqrt{\frac{n}{3}} \cdot \sqrt{\frac{2^k+1}{2^k-1}}.$$

Способы приближения закона распределения СВ $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$ к нормальному:

- увеличение n;
- использование специальных преобразований.

Пример. Если

$$\eta = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i ,$$

где ξ_i равномерно распределены на $(-h, h)$, то СВ

$$\xi = \eta - \frac{1}{20n}\left(3\eta - \eta^3\right)$$

имеет распределение, достаточно близкое к нормальному, уже при $n = 5$.

- ***Моделирование случайных чисел, распределенных по закону Пуассона***.

Пусть требуется получить последовательность случайных чисел y_i , имеющих распределение Пуассона с математическим ожиданием a:

$$P(\xi = k) = \frac{a^k}{k!} \cdot e^{-a} .$$

Можно использовать предельную теорему Пуассона.

Теорема Пуассона:

если *p* – вероятность наступления события *A* при одном испытании, то вероятность наступления *k* событий в *n* независимых испытаниях

при $n \to \infty$, $p \to 0$ и $np = a$ асимптотически равна

$$P(k) = \frac{a^k}{k!} \cdot e^{-a}.$$

Процедура получения последовательности y_i:

1) выбрать достаточно большое n, чтобы

$$p_n = \frac{a}{n} < 1,$$

Для практических целей p_n должно быть не более 0,1–0,2

2) из совокупности равномерно распределенных на (0, 1) случайных чисел x_i выбирать серии по n значений;

3) в серии с номером i подсчитывать число y_i случаев выполнения неравенства $x_i < p_n$.

Количество наступлений события A в n независимых испытаниях

Числа y_i имеют распределение, близкое к распределению Пуассона (тем точнее, чем больше n).

Моделирование случайных событий

Пусть

- имеются случайные числа x_i – возможные значения случайной величины ξ, равномерно распределенной в интервале (0, 1).
- необходимо реализовать случайное событие A, наступающее с заданной вероятностью p.

Определим A как событие, состоящее в том, что выбранное значение x_i случайной величины ξ удовлетворяет неравенству

$$x_i \le p.$$

Обоснование: $$P(A) = \int_0^p f(x)dx = \int_0^p dx = p.$$

Плотность распределения СВ ξ

Противоположное событие $\bar{A}$ состоит в том, что

$$x_i > p,$$

его вероятность равна $1-p$.

Процедура моделирования испытаний:

1) выбор значений x_i и сравнение их с величиной p;
2) если выполняется неравенство $x_i \le p$, то исходом испытания считается наступление события A;

   в противном случае исходом испытания считается наступление события $\bar{A}$.

Обобщение на группу событий.

Пусть $A_1, A_2, \dots, A_s$ – полная группа событий, наступающих с вероятностями $p_1, p_2, \dots, p_s$.

Определим A_m как событие, состоящее в том, что выбранное значение x_i случайной величины ξ удовлетворяет неравенству

$$l_{m-1} < x_i \le l_m ,$$

где

$$l_r = \sum_{i=1}^{r} p_i .$$

Обоснование: $P(A_m) = \int_{l_{m-1}}^{l_m} dx = p_m .$

Процедура моделирования испытаний:

1) выбор значений $\boldsymbol{x_i}$ и сравнение их с величинами $\boldsymbol{l_r}$;
2) исходом испытания считается наступление события $\boldsymbol{A_m}$, если выполняется неравенство

$$l_{m-1} < x_i \le l_m .$$

Эта процедура называется ***определением исхода испытания по жребию*** с вероятностями $p_1, p_2, \ldots, p_s$.

Эта же процедура – при формировании реализации дискретной СВ η, принимающей возможные значения $y_1, y_2, \ldots, y_s$ с вероятностями $p_1, p_2, \ldots, p_s$

Рассмотренные правила моделирования справедливы в предположении, что для испытаний применяются случайные числа x_i , имеющие равномерное распределение в интервале (0, 1).

Можно показать:
при использовании k-разрядных псевдослучайных чисел с квазиравномерным распределением ошибка в определении вероятности события не превосходит величины $\frac{1}{2^k - 1}$.

Моделирование совместных испытаний

В процессе моделирования функционирования систем необходимо бывает осуществить испытания, при которых искомый результат является сложным событием, зависящим от двух или нескольких простых событий.

- Пусть ***A*** и ***B*** – <u>независимые</u> события, вероятности наступления которых равны $\boldsymbol{p_A}$ и $\boldsymbol{p_B}$.

  Возможные исходы совместных испытаний:

$$AB, \quad \overline{A}B, \quad A\overline{B}, \quad \overline{A}\,\overline{B}$$

  с вероятностями

$$p_A \cdot p_B, \quad (1-p_A)p_B, \quad p_A(1-p_B), \quad (1-p_A)(1-p_B).$$

Для моделирования совместных испытаний могут быть использованы две процедуры.

1. Последовательная проверка выполнения неравенств, аналогичных неравенству $x_i \le p$, относительно событий ***A*** и ***B.***

Требует использования 2 случайных чисел и 2 сравнений

2. Определение одного из исходов $\boldsymbol{AB}$, $\boldsymbol{\bar{A}B}$, $\boldsymbol{A\bar{B}}$, $\boldsymbol{\bar{A}\bar{B}}$ по жребию с соответствующими вероятностями.

Достаточно 1 случайного числа, но сравнений может потребоваться больше

В практическом моделировании выбор процедуры определяется соображениями

- удобства построения алгоритма,
- экономией количества операций и оперативной памяти.

В среднем первая процедура более экономна, чем вторая

- Пусть события ***A*** и ***B*** не являются независимыми.

  Пусть условная вероятность $\boldsymbol{P(B|A)}$ известна.

  Описанные выше процедуры могут быть модифицированы следующим образом.

1. Аналог процедуры 1 для независимых событий **A** и **B**.

Процедуры моделирования испытаний:

1) из совокупности $\boldsymbol{x_i}$ выбрать очередное число $\boldsymbol{x_n}$ и сравнить его с величиной $\boldsymbol{p_A}$;

Событие *A* наступило

2) если выполняется неравенство $\boldsymbol{x_n} \le \boldsymbol{p_A}$, то

2.1) для очередного числа $\boldsymbol{x_{n+1}}$ проверить условие $\boldsymbol{x_{n+1}} \le \boldsymbol{P(B|A)}$;

в зависимости от того, выполнено оно или нет, исходом испытания будет $\boldsymbol{AB}$ или $\boldsymbol{A\overline{B}}$.

Наступило событие $\bar{A}$

если выполняется $x_n > p_A$, то

2.2) найти условную вероятность $P(B|\bar{A})$:

т. к. $p_B = p_A \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})$,

то

$$P\left(B \mid \overline{A}\right) = \frac{p_B - p_A \cdot P\left(B \mid A\right)}{1 - p_A};$$

2.3) для очередного числа x_{n+1} проверить выполнение условия $x_{n+1} \le P(B|\bar{A})$;

в зависимости от того выполняется оно или нет, исходом испытания будет $\overline{A}B$ или $\overline{A}\overline{B}$.

2. Определение исхода по жребию (аналог процедуры 2 для независимых событий **A** и **B**).

События AB, $A\bar{B}$, $\bar{A}B$ и $\bar{A}\bar{B}$ образуют полную группу и имеют вероятности, соответственно,

$$p_A \cdot P(B|A), \quad p_A \cdot (1 - P(B|A)),$$

$$(1 - p_A) \cdot P(B|\bar{A}), \quad (1 - p_A) \cdot (1 - P(B|\bar{A})),$$

где

$$P(B\,|\,\bar{A}) = \frac{p_B - p_A \cdot P(B\,|\,A)}{1 - p_A}.$$

<u>Задание</u> (5 баллов): составить блок-схемы моделирующих алгоритмов для процедур 1 и 2 (для независимых и зависимых событий **A** и **B**).

Моделирование простых цепей Маркова

Простая однородная цепь Маркова определяется матрицей переходов

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1K} \\ p_{21} & p_{22} & \cdots & p_{2K} \\ \cdots & \cdots & \cdots & \cdots \\ p_{K1} & p_{K2} & \cdots & p_{KK} \end{pmatrix}$$

p_{ij} – вероятность перехода системы из состояния z_i в состояние z_j

В данном случае исход испытания – переход системы в одно из состояний z_1, z_2, … z_K – наступление одного из событий A_1, A_2, … A_K .

p_{ij} – условная вероятность наступления события A_j в данном испытании при условии, что результатом предыдущего испытания было наступление события A_i .

Моделирование такой цепи – последовательный выбор событий A_j по жребию в соответствии с вероятностями p_{ij}.

Процедура моделирования испытаний.

1) Выбрать начальное состояние, определяемое начальными вероятностями $p_{01}, p_{02}, \dots p_{0K}$:

выбрать случайное число x_n и проверить выполнение неравенств

$$l_{m-1} < x_n \le l_m, \quad \text{где} \quad l_r = \sum_{i=1}^{r} p_{0i};$$

если для некоторого номера m_0 выполняется

$$l_{m_0-1} < x_n \le l_{m_0},$$

то начальное событие данной реализации цепи – событие A_{m_0}.

2) Выбрать следующее случайное число x_{n+1} и выполнить проверку условий

$$l'_{m-1} < x_{n+1} \le l'_m \,, \quad \text{где} \quad l'_r = \sum_{i=1}^{r} p_{m_0 i} \,;$$

если для некоторого номера m_1 выполняется

$$l_{m_1-1} < x_{n+1} \le l_{m_1} \,,$$

то следующее событие данной реализации цепи – событие A_{m_1}.

И т. д.

Каждый номер $\boldsymbol{m_i}$ определяет

- очередное событие $\boldsymbol{A_{m_i}}$ формируемой реализации;
- распределение вероятностей $\boldsymbol{p_{m_i1}}, \boldsymbol{p_{m_i2}}, \ldots \boldsymbol{p_{m_iK}}$ для выбора следующего номера $\boldsymbol{m_{i+1}}$.

Замечание.

Для эргодической цепи Маркова в качестве значений $\boldsymbol{p_{01}}, \boldsymbol{p_{02}}, \ldots, \boldsymbol{p_{0K}}$ могут быть выбраны произвольные значения, например,

$$\boldsymbol{p_{01}} = \boldsymbol{p_{02}} = \ldots = \boldsymbol{p_{0K}} = 1/\boldsymbol{K}.$$