

Моделирование случайных величин

«Теория информационных процессов и систем»

Лекция 4

Способы получения случайных величин

физические генераторы (датчики) случайных величин;

программные генераторы (датчики)
псевдослучайных чисел – позволяют получить
периодические детерминированные числовые
последовательности с большим периодом,
называемые псевдослучайными.

Линейные конгруэнтные генераторы (ЛКГ)

$$\xi_{i+1} = (a\xi_i + c) \pmod{m},$$
$$\xi_i \in (0, m-1), |(0, m-1)| = m$$

Теорема: ЛКГ имеет полный период, когда выполняются следующие условия:

- m и c являются взаимно простыми числами;
- если m делится на простое число q , то $a-1$ тоже делится на q ;
- если m делится на 4, то $a-1$ тоже делится на 4.

Линейные конгруэнтные генераторы (ЛКГ)

Пример 1:

$$\xi_{i+1} = (a\xi_i + c) \pmod{m},$$

$m=5, c=3, a=6, \xi_0=4.$

$$\xi_1 = (6 \cdot 4 + 3) \pmod{5} = 2,$$

$$\xi_2 = (6 \cdot 2 + 3) \pmod{5} = 0,$$

$$\xi_3 = (6 \cdot 0 + 3) \pmod{5} = 3,$$

$$\xi_4 = (6 \cdot 3 + 3) \pmod{5} = 1,$$

$$\xi_5 = (6 \cdot 1 + 3) \pmod{5} = 4,$$

$$\xi_6 = (6 \cdot 4 + 3) \pmod{5} = 2.$$

Линейные конгруэнтные генераторы (ЛКГ)

Пример 2:

$$\xi_{i+1} = (a\xi_i + c) \pmod{m},$$

$$m=5, c=5, a=6, \xi_0=4.$$

$$\xi_1 = (6 \cdot 4 + 5) \pmod{5} = 4,$$

$$\xi_2 = (6 \cdot 4 + 5) \pmod{5} = 4.$$

Мультипликативные генераторы

$$\xi_{i+1} = (a\xi_i) \pmod{m},$$
$$\xi_i \in (1, m-1), |(1, m-1)| = m-1$$

Теорема: Мультипликативный генератор имеет период $m-1$, когда выполняются следующие условия:
 m является простым числом;
 $a-1$ является первообразным элементом по модулю m , т.е. наименьшее целое число l , для которого $a^l - 1$ делится на m , есть $l = m-1$.

Мультипликативные генераторы

Пример 3:

$$\xi_{i+1} = (a\xi_i) \pmod{m},$$
$$m=5, a=2, \xi_0=4.$$

$$\xi_1 = (2 \cdot 4) \pmod{5} = 3,$$

$$\xi_2 = (2 \cdot 3) \pmod{5} = 1,$$

$$\xi_3 = (2 \cdot 1) \pmod{5} = 2,$$

$$\xi_4 = (2 \cdot 2) \pmod{5} = 4,$$

$$\xi_5 = (2 \cdot 4) \pmod{5} = 3.$$

Мультипликативные генераторы

Пример 4:

$$\xi_{i+1} = (a\xi_i) \pmod{m},$$
$$m=5, a=4, \xi_0=4.$$

$$\xi_1 = (4 \cdot 4) \pmod{5} = 1,$$

$$\xi_2 = (4 \cdot 1) \pmod{5} = 4,$$

$$\xi_3 = (4 \cdot 4) \pmod{5} = 1.$$

Моделирование дискретной случайной величины

Необходимо получить последовательность значений x_i случайной величины X с распределением:

x_1	x_2	$\dots$	x_n
p_1	p_2	$\dots$	p_n

Интегральная функция распределения:

$$F_X(x) = P(X \leq x) = \sum_{i=1}^n p_i; x_n \leq x \leq x_{n+1}; n = 1, 2, \dots;$$

$$F_X(x) = 0; x < x_1$$

Моделирование дискретной случайной величины Метод обратной функции

Если γ - равномерно распределенная на интервале $(0,1)$ случайная величина, то искомая случайная величина X получается с помощью преобразования

$$X = F_X^{-1}(\gamma),$$

где F_X^{-1} - функция, обратная F_X .

Общая формула

$$x = \min \{x : F(x) \geq \gamma\}$$

Моделирование дискретной случайной величины

- Интервал $(0,1)$ разбивается на n частей с длинами $p_1, p_2, \dots, p_n$. Полученные интервалы нумеруются цифрами $1, 2, \dots, n$. Координаты точек деления $y_0=0$, $y_1=p_1$, $y_2=p_1+p_2$, $y_n=p_1+p_2+\dots+p_n$.
- Выбирается стандартно равномерно распределенная случайная величина γ и строится точка $y = \gamma$.
- Если эта точка попадает в интервал с номером i , то $X=x_i$.

Пример

x	1	2	3
p	0.2	0.5	0.3

$\gamma = \{0.43, 0.75, 0.11, 0.98, 0.35, 0.64, 0.23\}$

$x = \{2, 3, 1, 3, 2, 2, 2\}$

Распределение Бернулли

Алгоритм эквивалентен методу обратного преобразования, если y и $1 - y$ поменять местами.

1. Генерируем $y \sim U(0,1)$.
2. Если $y \leq p$, возвращаем $X = 1$; в противном случае возвращаем $X = 0$.

p – вероятность успеха в испытании Бернулли

Биномиальное распределение

Биномиальное распределение задает вероятность k удачных исходов в n реализациях некоторого эксперимента:

$$p_k = P(X = k) = C_n^k p^k q^{n-k},$$

где $q = 1-p$, $C_n^k = \frac{n!}{k!(n-k)!}$ - биномиальные коэффициенты.

$$M[X] = np; D[X] = npq.$$

Биномиальное распределение

I. Сумма n независимых и одинаково распределенных величин с распределением Бернулли имеет биномиальное распределение.

1. Генерируем $Y_1, Y_2, \dots, Y_n$ как независимые и одинаково распределенные случайные величины с распределением Бернулли.

2. Возвращаем $X = Y_1 + Y_2 + \dots + Y_n$.

Время выполнения алгоритма пропорционально величине n .

II. Метод обратного преобразования: $p_0 = (1 - p)^n$;

$$\frac{p_k}{p_{k-1}} = \frac{n - k + 1}{k} \frac{p}{1 - p}$$

Распределение Пуассона

Дискретная случайная величина X называется распределенной по закону Пуассона с параметром $\lambda > 0$, если ее возможные значения – все неотрицательные целые числа: $0, 1, 2, \dots$, а вероятность события $\{X = x\}$ выражается формулой:

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

$$M[X] = D_X = \lambda.$$

Распределение Пуассона

I. Метод обратных функций

Ряд пуассоновского распределения имеет вид:

x	0	1	2	...	k	...
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!} e^{-\lambda}$	...	$\frac{\lambda^k}{k!} e^{-\lambda}$	...

Следовательно, $p_0 = e^{-\lambda}$, $p_k / p_{k-1} = \lambda / k$.

Распределение Пуассона

II. Пусть $Y_1, Y_2, \dots, Y_n, \dots$ – последовательность независимых равномерно распределенных на $(0, 1]$ случайных величин. Тогда случайная величина

$$\xi = \max \left\{ n \geq 0 : \prod_{i=1}^n \gamma_i > e^{-\lambda} \right\}, \lambda > 0$$

описывается распределением Пуассона.

Элемент выборки может быть получен путем последовательного увеличения числа членов n в произведении, пока не нарушится условие $\prod_{i=1}^n \gamma_i > e^{-\lambda}$.

Максимальное значение n , удовлетворяющее этому условию, – очередное значение случайной величины.

Распределение Пуассона

II. Алгоритм:

1. Инициализируем $a = e^{-\lambda}$, $b = 1$, $i = 0$.
2. Генерируем $\gamma_{i+1} \sim U(0, 1)$ и заменяем b на $b \cdot \gamma_{i+1}$.
Если $b < a$, возвращаем $X = i$. В противном случае переход к шагу 3.
3. Увеличиваем i на единицу и возвращаемся к шагу 2.

Геометрическое распределение

I. Метод обратных функций, основанный на пересчете

$$p_n = (1 - p) p_{n-1}, p_0 = p.$$

II. Моделирование испытаний Бернулли с вероятностью успеха p до первого успеха с подсчетом количества неудач.

III. Накопленная вероятность $S_{n+1} = p_0 + \dots + p_n$ для геометрического распределения имеет вид:

$$S_{n+1} = \sum_{i=0}^n p(1-p)^i = 1 - (1-p)^{n+1}$$

Геометрическое распределение

Распределение случайной величины X называется геометрическим с параметром p ($p \in (0,1)$), если значения случайной величины X – все натуральные числа, а вероятность события $\{X = x\}$ выражается формулой:

$$p_k = P(X = k) = q^{k-1} \cdot p, \quad k = 1, 2, \dots,$$

где $q = 1 - p$.

Вероятности p_k образуют бесконечную убывающую геометрическую прогрессию со знаменателем q .

$$M[X] = \frac{1}{p}; \quad D_X = \frac{q}{p^2}.$$

Геометрическое распределение

III. Поэтому событие $\{X = n\}$ приобретает вид:

$$\{X = n\} = \{S_n < \gamma < S_{n+1}\} = \left\{n < \frac{\ln(1-\gamma)}{\ln(1-p)} \leq n+1\right\}$$

$$X = \left\lfloor \ln(1-\gamma)/\ln(1-p) \right\rfloor$$

Алгоритм:

1. Генерируем $\gamma \sim U(0, 1)$.
2. Возвращаем $X = \left\lfloor \ln(1-\gamma)/\ln(1-p) \right\rfloor$.

Константа $\ln(1-p)$ вычисляется заранее.

При больших значениях p рекомендуется использовать другие алгоритмы.

Моделирование непрерывных случайных величин

Метод обратной функции

Для нахождения значения непрерывной случайной величины X , распределенной в интервале (a, b) с плотностью $p(x)$ необходимо решить уравнение:

$$\int_a^x p(\tau) d\tau = \gamma$$

где $\gamma \sim U(0, 1)$.