Лекция 6. Моделирование дискретных случайных величин.

Основной метод моделирования дискретных случайных величин

Для моделирования дискретных случайных величин, заданных следующим образом

$$\nu = \begin{pmatrix} x_1 & x_2 \dots & x_m \\ p_1 & p_2 \dots & p_m \end{pmatrix},$$

используется метод моделирования полной группы событий, где события A_k заменены значениями x_k случайной величины ν.

Этот метод моделирования дискретных случайных величин хотя и является универсальным, но требует значительных затрат машинного времени. Поэтому для моделирования ряда дискретных распределений, имеющих важное практическое применение, разработаны более эффективные методы.

Геометрическое распределение

Если некоторое событие происходит с вероятностью p, то число ν независимых испытаний, необходимых, чтобы это событие произошло (или число испытаний между событиями), подчиняется геометрическому распределению. Следовательно, $\nu = 1$ с вероятностью p, $\nu = 2$ с вероятностью $(1 - p) * p$ и в общем случае

$$P\{\nu = k\} = p * (1-p)^{k-1} = p_k. \tag{29}$$

Математическое ожидание геометрического распределения равно

$$M[\nu] = m_x = \frac{(1-p)}{p},$$

а дисперсия

$$D[\nu] = \sigma_x^2 = \frac{(1-p)}{p^2}.$$

Для моделирования случайных величин с геометрическим распределением, когда p достаточно мало, можно использовать формулу

$$\nu = \left[\frac{\ln \xi}{\ln(1-p)} \right], \tag{30}$$

где ν равно наименьшему целому числу, большему или равному значению выражения, заключенного в квадратные скобки.

Алгоритм для моделирования геометрического распределения можно представить в виде:

Шаг 1. Положить $j = 1$.

Шаг 2. Получить реализацию z случайной величины ξ.

Шаг 3. Вычислить реализацию x_j случайной величины ν

$$x_j = Ц\left[\frac{\ln z}{\ln(1-p)} \right] + 1$$

и принять $j = j + 1$.

Шаг 4. Проверить условие завершения счета $j > n$. При нарушении этого условия вернуться на шаг 2.

Шаг 5. Вывод $\{x_j\}$.

Распределение Пуассона

Распределение Пуассона называют законом появления редких событий, и оно характеризует число событий в единицу времени, каждое из которых может произойти в любой момент. Например, число вызовов на станцию скорой медицинской помощи, количество происходящих в стране пожаров, ураганов и т.д.

Вероятность того, что случайная величина ν примет некоторое целочисленное значение k, определяется формулой Пуассона

$$P\{\nu = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \tag{31}$$

где λ – среднее число событий, имеющих место в единицу времени.

Математическое ожидание и дисперсия соответственно равны

$$M[\nu] = m_x = \lambda, \qquad D[\nu] = \sigma_x^2 = \lambda. \tag{32}$$

Для моделирования случайной величины по закону Пуассона воспользуемся предельной теоремой Пуассона. На основе теоремы Пуассона можно выбрать следующую схему моделирования дискретной случайной величины ν.

Необходимо проводить серии по n независимых испытаний, в каждом из которых событие происходит с вероятностью p, и подсчитать число k_i случаев наступления события A в серии с номером i. Числа k_i и будут реализациями распределения Пуассона. Число испытаний n определяют из условия $n = \frac{\lambda}{p}$.

Алгоритм:

Шаг 1. Положить $i = 1$, $k_i = 0$.

Шаг 2. Положить $j = 1$.

Шаг 3. Выработать реализацию z базовой случайной величины ξ.

Шаг 4. Проверить условие $z \le p$. При нарушении этого условия переход на шаг 6.

Шаг 5. Принять $k_i = k_i + 1$.

Шаг 6. Положить $j = j + 1$.

Шаг 7. Проверить условие $j > n$. При нарушении этого условия переход на шаг 3.

Шаг 8. Принять $i = i +1$.

Шаг 9. Проверить условие окончания счета $i > N$. При нарушении этого условия переход на шаг 2.

Шаг 10. Вывод содержимого счетчиков k_i, $x_i = k_i$, $i = \overline{1,N}$.

Биномиальное распределение

Биномиальное распределение, полученное Бернулли, является дискретным распределением и характеризует число благоприятных исходов в n испытаниях по схеме «успех-неудача» с вероятностями p и (1-p).

Вероятность того, что в последовательности из n испытаний "успехи" выпадут k раз, выражается формулой

$$P\{V = k\} = \frac{n!\, p^k (1-p)^{(n-k)}}{k!(n-k)!}.$$

$$M[\nu] = m_x = n * p, \qquad D[\nu] = \sigma_x^2 = n * p * (1-p).$$

Выбор метода моделирования биномиального распределения зависит от значения параметров n и p. При небольших значениях n моделируется n базовых случайных чисел $z[k]$ и величина ν равняется количеству z, не превосходящих по величине p. При больших значениях n и малых p проверяется неравенство

$$z \le \sum_{i=0}^{k} r[i],$$

где $r[i+1] = \frac{r[i]p(n-i)}{(i+1)(1-p)}$, $r[0] = (1-p)^n$.

Дискретная случайная величина ν в этом случае равняется числу итераций k, которое необходимо для выполнения условия $z \le \sum_{i=0}^{k} r[i]$.

Алгоритм моделирования биномиального распределения состоит из следующих шагов.

Шаг 1. Ввод исходных данных N, p и n.

Шаг 2. Задание начальных условий $i = 1$, $r[0] = (1-p)^n$.

Шаг 3. Обнуление счетчика $k = 0$.

Шаг 4. Проверка условия $n \le 10$. При нарушении этого условия переход на шаг 10.

Шаг 5. Запуск счетчика числа испытаний: $j = 1$.

Шаг 6. Моделирование z.

Шаг 7. Проверить условие $z \le p$. При нарушении этого условия переход на шаг 9.

Шаг 8. Увеличение содержимого счетчика $k = k + 1$.

Шаг 9. Увеличение содержимого счетчика j = j + 1 для организации n-кратного повторения шагов 6 и 7 и переход при выполнении условия j > n на шаг 12.

Шаг 10. Моделирование z.

Шаг 11. Проверка условия $z \le \sum_{l=0}^{k} r[l]$ до его выполнения.

Шаг 12. $V[i] = k$.

Шаг 13. Переход к следующему испытанию $i = i + 1$.

Шаг 14. Проверка условия $i > N$. При нарушении этого условия переход на шаг 3.

Шаг 15. Печать результатов моделирования $V[i]$.

Шаг 16. Конец.

Контрольные вопросы:

1. Приведите алгоритм моделирования геометрического распределения.
2. Приведите блок-схему алгоритма моделирования распределения Пуассона.
3. Приведите блок-схему алгоритма основного метода моделирования дискретных случайных величин.