16 МОДЕЛИРОВАНИЕ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

16.1 Моделирование случайного события

Пусть событие A имеет вероятность p. Будем считать, что случайная величина X принимает значение 1, если произошло событие A и 0 – если событие A не произошло. Для моделирования такой случайной величины генерируем стандартное число g. Если $g < p$, то $X = 1$, то есть произошло событие A, в противном случае $X = 0$ и событие A не произошло (рисунок 1).

16.2 Общая схема моделирования

Дискретные случайные величины характеризуются решетчатым распределением (рисунок 2).

Рассмотрим в начале вариант, когда случайная величина задана таблично. Например, величина X принимает значения $x_1, x_2, ..., x_i, ..., x_n$ с вероятностями $P_1, P_2, ..., P_i, ..., P_n$. При этом выполняется условие нормировки вероятностей, то есть

$$P_1 + P_2 + ... + P_i + ... + P_n = 1.$$

При моделировании дискретной случайной величины исходят из стандартной псевдослучайной величины γ, равномерно распределенной в интервале 0...1.0. При этом отрезок (0...1.0) разбивается на интервалы длиной $P_1, P_2, ..., P_n$ и определяют вспомогательные накопленные вероятности (рисунок 3):

$$F_1 = P_1, F_2 = F_1 + P_2, ..., F_i = F_{i-1} + P_i, ..., F_N = F_{N-1} + P_N = 1.$$

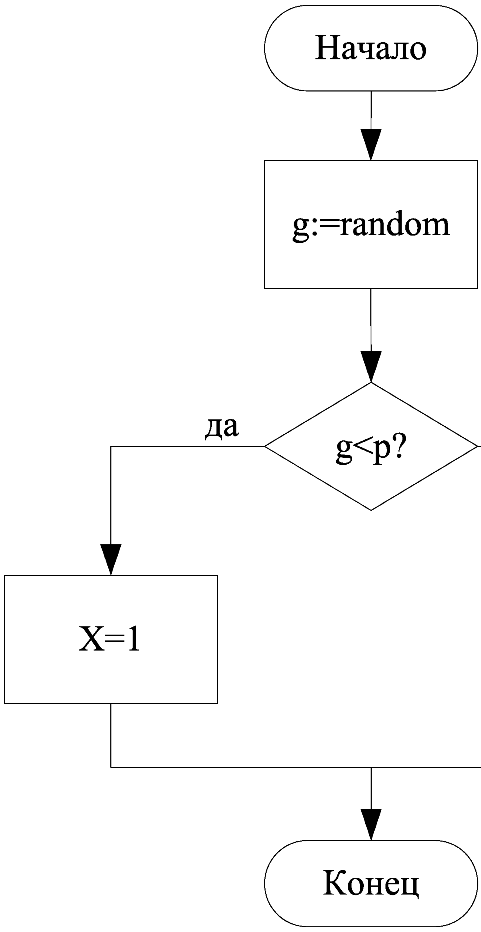

Рисунок 1 – Блок-схема моделирования случайного события

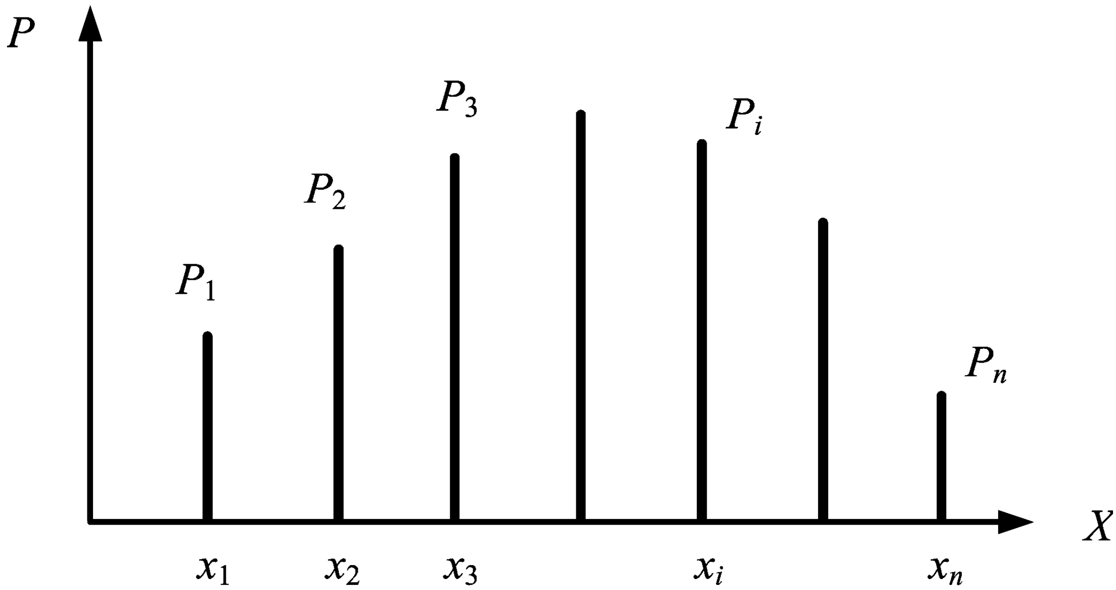

Рисунок 2 – Решетчатое распределение дискретной случайной величины X

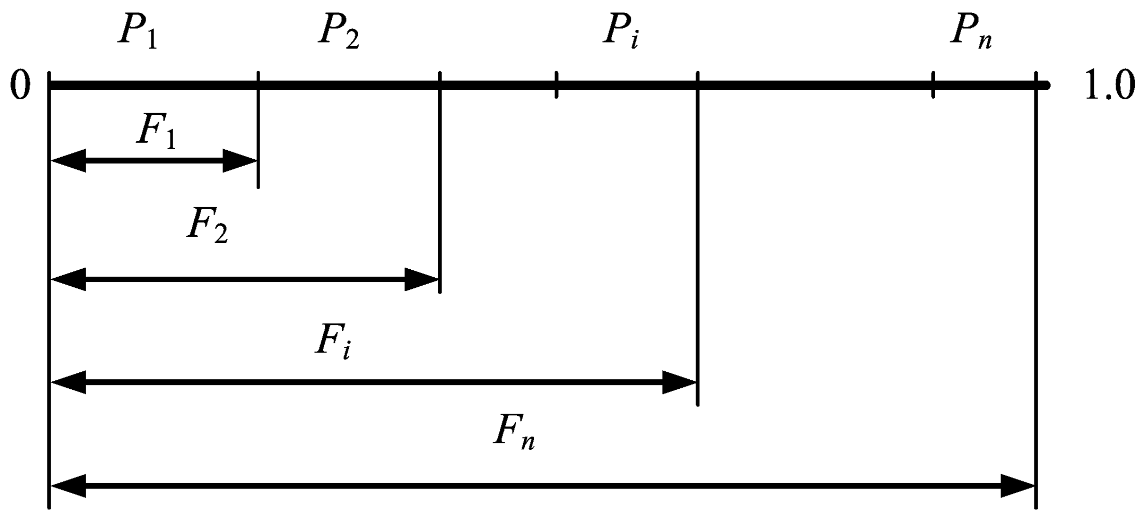

Рисунок 3 – Иллюстрация к алгоритму моделирования дискретной случайной величины

Алгоритм преобразования стандартного непрерывно распределенного псевдослучайного числа в дискретно распределенное число X состоит из следующих шагов:

1) генерируется стандартное число γ;
2) определяется номер поддиапазона i, в который попадает число γ;
3) осуществляется присвоение $X = x_i$.

Номер поддиапазона определяется в цикле. Блок-схемы алгоритмов приведены на рисунках 4 и 5.

Вариант на рисунок 4 имеет ту особенность, что накопленные вероятности F_i рассчитываются в цикле при каждом обращении к подпрограмме генерации числа X и не запоминаются. В подпрограмму передаются только массивы вероятностей P_i и возможных значений X_i. В варианте рисунка 5 F_i рассчитываются заранее один раз и в подпрограмму передается массив F_i вместо P_i. Таким образом, второй вариант предпочтительнее по быстродействию, но сложнее в реализации, так как где-то отдельно надо подсчитать F_i

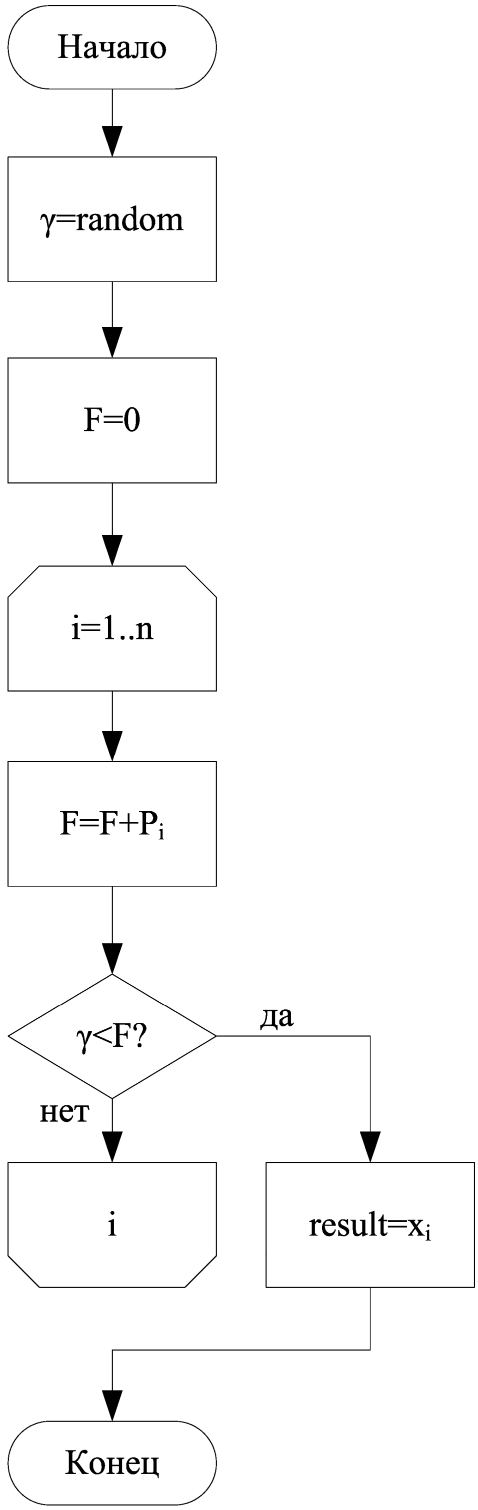

Рисунок 4 – Блок-схема подпрограммы-функции, генерирующей при каждом обращении новое значение дискретной случайной величины X без хранения накопленной вероятности F

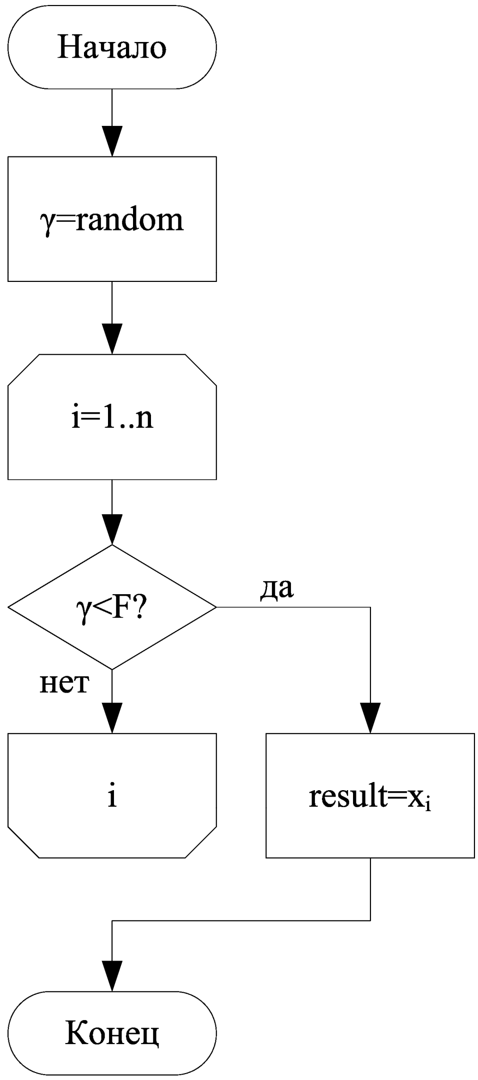

Рисунок 5 – Блок-схема подпрограммы-функции, генерирующей при каждом обращении новое значение дискретной случайной величины X с заранее рассчитанной накопленной вероятности F

16.3 Распределение Пуассона

В случае распределения Пуассона вероятности значений $x = (0,1,...,\infty)$ определяется по следующей формуле:

$$P(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}, \tag{1}$$

где λ – параметр распределения. Таким образом, случайная величина X принимает свои значения $0, 1, 2, ...$ с вероятностями $P(0), P(1), P(2), ...$ соответственно.

Математическое ожидание и дисперсия равны соответственно

$$\overline{X} = \lambda, \quad D_X = \lambda .$$

При моделировании вероятности $P(x)$ удобнее считать по следующей рекуррентной формуле:

$$P(0) = e^{-\lambda}, \; P(x) = P(x-1) \cdot \varphi(x),$$

где

$$\varphi(x) = \frac{P(x)}{P(x-1)} = \frac{\lambda}{x}. \tag{2}$$

Блок-схема алгоритма моделирования в этом случае может выглядеть, как на рисунках 6 и 7.

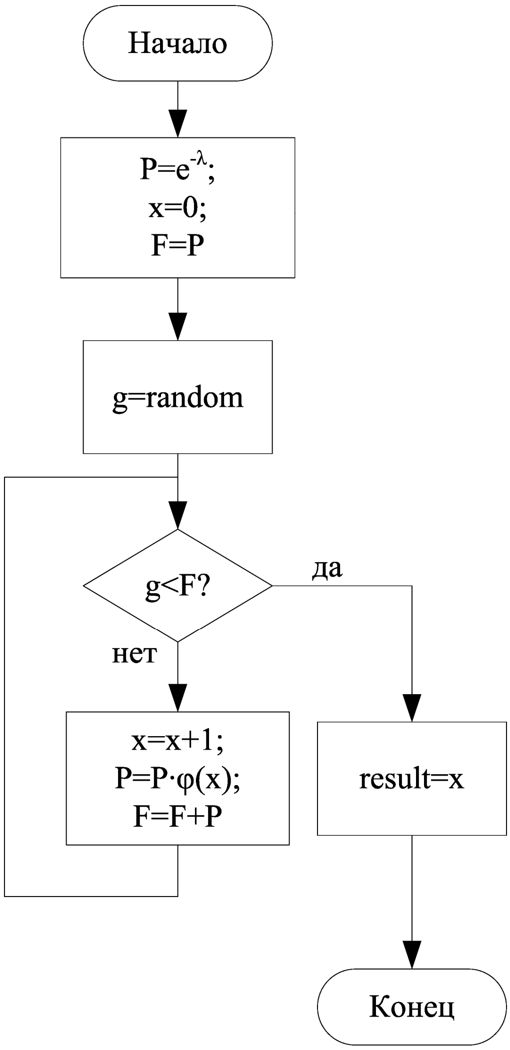

Рисунок 6 – Блок-схема первого варианта подпрограммы-функции моделирования распределения Пуассона

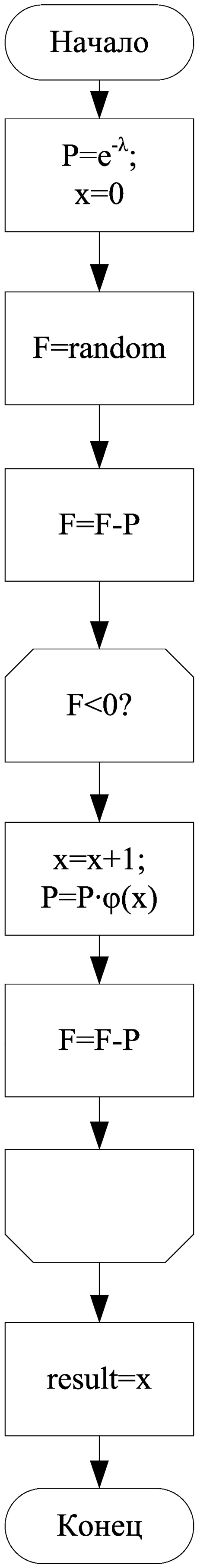

Рисунок 7 – Блок-схема второго варианта подпрограммы-функции моделирования распределения Пуассона

При моделировании распределения Пуассона (рисунки 6, 7) $\varphi(x)$ рассчитывается по формуле (2).

16.4 Биномиальное распределение

Биномиальное распределение связано с повторением опыта. Если p вероятность появления некоторого события в результате опыта, то вероятность того, что после повторения опыта n раз это событие появится x раз,

$$P_n(x) = C_n^x \cdot p^x \cdot (1-p)^{n-x}, x = 0,...,n, \tag{3}$$

при среднем значении $\overline{X} = p \cdot n$ и дисперсии $D_X = n \cdot p \cdot (1-p)$. $C_n^x = \frac{n!}{x! \cdot (n-x)!}$ – это число сочетаний из n по x.

Для вычисления вероятностей $P_n(x)$ удобно пользоваться следующей рекуррентной формулой:

$$P_n(0) = (1-p)^n, \quad P_n(x) = P_n(x-1) \cdot \varphi(x), \tag{4}$$

где

$$\varphi(x) = \frac{n-x+1}{x} \cdot \frac{p}{1-p}. \tag{5}$$

Блок-схема алгоритма моделирования подобна приведенным на рисунках 6 и 7. Особенность только в определении вероятности при $x = 0$ и функции $\varphi(x)$.

При моделировании биноминального распределения (рисунок 8) $\varphi(x)$ вычисляется по формуле (5).

16.5 Геометрическое распределение

Геометрическое распределение то же связано с повторением опыта. Если вероятность p появления некоторого события в результате опыта, то вероятность того, что это событие впервые появится при испытании номер x,

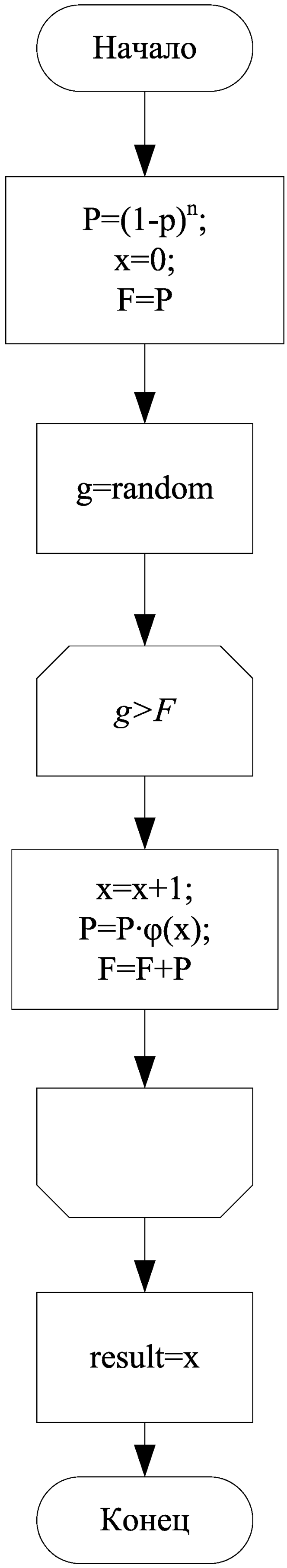

Рисунок 8 – Блок-схема алгоритма моделирования биномиального распределения

$$P(x) = p \cdot (1-p)^{x-1}, x = 1,2,3,... \qquad (6)$$

Математическое ожидание и дисперсия равны соответственно

$$\overline{X} = \frac{1}{p},$$

$$D_X = \frac{1-p}{p^2}.$$

Моделируется это распределение подобно пуассоновскому с той лишь особенностью, что

$$\varphi(x) = \frac{P(x)}{P(x-1)} = 1 - p, \qquad (7)$$

начальное значение $x = 1$ и $P(1) = p$.

При моделировании геометрического распределения $\varphi(x)$ вычисляется по формуле (7).

16.6 Гипергеометрическое распределение

Гипергеометрическое распределение связано с выборочным контролем. Если N размер партии изделий, из которых D дефектны, то вероятность того, что в выборке размера n окажется ровно x дефектных изделий,

$$P(x) = \frac{C_D^x \cdot C_{N-D}^{n-x}}{C_N^n}, \quad x = \max(0, n - N + D), ..., \min(n, D). \qquad (8)$$

Математическое ожидание и дисперсия в этом случае равны соответственно:

$$\overline{X} = n \cdot \frac{D}{N}, \quad D_X = n \cdot \frac{D \cdot (N-D) \cdot (N-n)}{N^2 \cdot (N-1)}$$

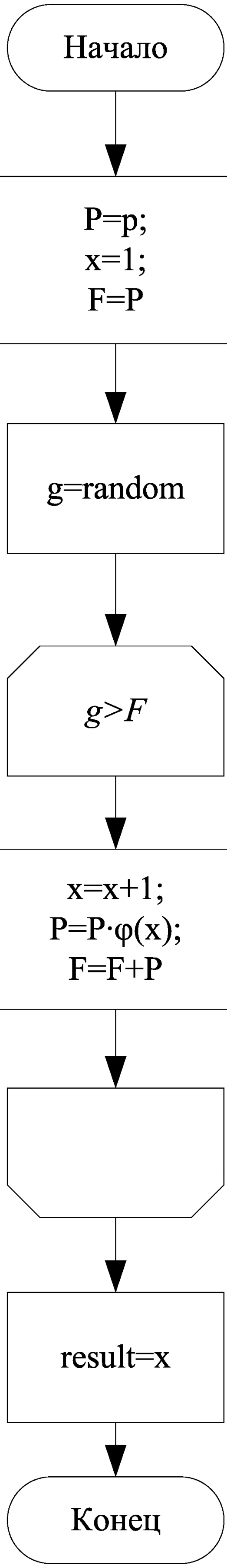

Рисунок 9. Блок-схема алгоритма моделирования геометрического распределения

Рекуррентная формула для расчета $P(x)$ аналогична предыдущим случаям, но при расчете начальной вероятности следует учитывать, что первое значение x с положительной вероятностью равно 0, если количество кондиционных изделий в партии больше или равно размеру выборки, то есть $N - D \geq n$ и $x = n - N + D$, если количество кондиционных изделий в партии меньше размера выборки, то есть $N - D < n$. В первом случае

$$P(0) = \frac{C_{N-D}^{n}}{C_{N}^{n}},$$

а во втором случае

$$P(n - N + D) = \frac{C_{D}^{n-N+d}}{C_{N}^{n}};$$

$$\varphi(x) = \frac{(D - x + 1) \cdot (n - x + 1)}{x \cdot (N - D - n + x)}. \tag{9}$$

Блок-схема алгоритма моделирования гипергеометрического распределения аналогична схемам, приведенным на рисунках 6 и 7, особенность только в расчете $P(0)$ и $\varphi(x)$.

16.7 Специальные алгоритмы моделирования дискретных случайных величин

16.7.1 Равномерное дискретное распределение

Это распределение задается диапазоном возможных целочисленных значений $N_1 ... N_2$ и вероятностью этих значений

$$p = \frac{1}{n}, \tag{10}$$

где $n = N_2 - N_1 + 1$ – число значений в диапазоне.

Математическое ожидание этой случайной величины

$$\overline{X} = \frac{N_1 + N_2}{2},$$

а дисперсия

$$D_X = \frac{n^2 - 1}{12}.$$

Для моделирования такой случайной величины целесообразно воспользоваться функцией Паскаля *random(N)*, выдающей равновероятные целые числа в диапазоне $0...N-1$. В результате получаем следующую формулу для генерации равномерного дискретного распределения:

$$X = N_1 + random(n). \tag{5.11}$$

16.7.2 Биномиальное распределение

Как уже отмечалось в п. 16.4, биномиальное распределение связано с повторением опытов при заданной вероятности p появления некоторого события в каждом опыте. Воспользуемся этим фактом. Пусть n число повторений опыта, тогда, моделируя n раз этот опыт и подсчитывая число реализаций отмеченного события, получаем значение нашей случайной величины X как число реализаций этого события после n испытаний.

Блок-схема этого алгоритма показана на рисунке 10. Этот алгоритм требует n стандартных случайных чисел для генерации одного числа, распределенного по биномиальному закону. В этом смысле проще использовать алгоритм, рассмотренный раньше, который требует только одного стандартного числа.

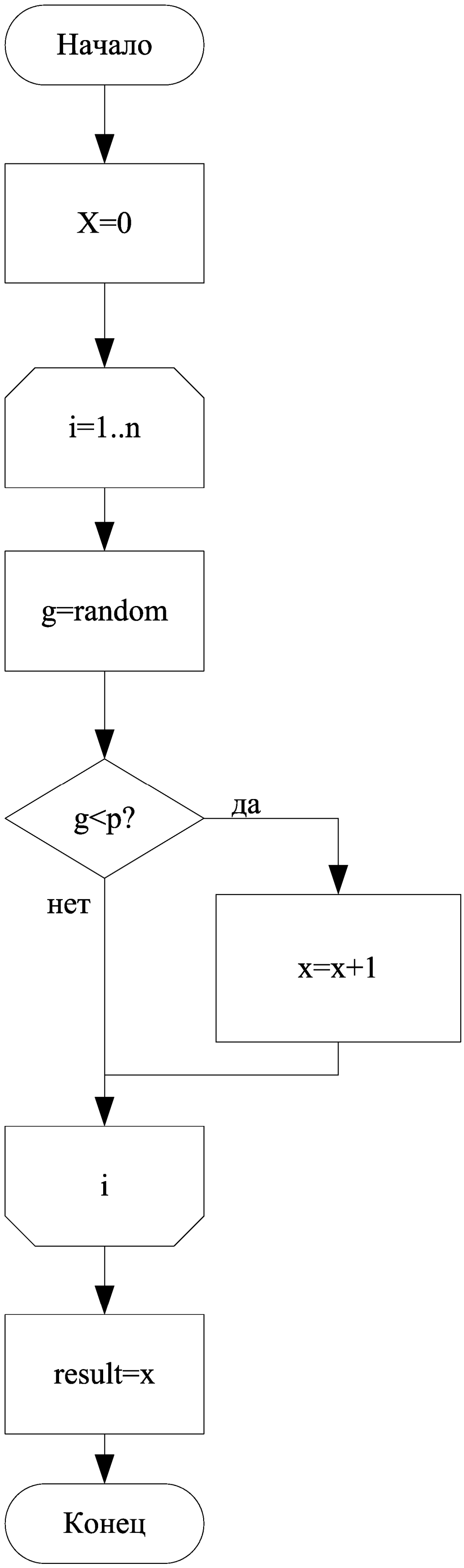

Рисунок 10 – Блок-схема алгоритма моделирования биномиального распределения

16.7.3 Геометрическое распределение

Рассмотрим еще один вариант генерации геометрического распределения, используя тот факт, что это распределение связано с повторением опыта. Как и в предыдущем случае, считаем, что p – вероятность некоторого события в одном опыте. Моделируем опыт до первой реализации нашего события. Номер опыта, при котором это событие реализовалось первый раз, и есть число X с геометрическим распределением (6).

Блок-схема алгоритма, реализующая эту идею, приведена на рисунке 11.

Если сравнивать этот алгоритм с рассмотренным раньше в п. 16.5, то этот алгоритм с точки зрения вычислительных операций проще, но требует переменное количество стандартных чисел на одно геометрически распределенное число.

16.7.4 Гипергеометрическое распределение

Воспользуемся тем фактом, что гипергеометрическое распределение связано с выборочным контролем. Пусть, как и в п. 16.7, N – размер партии изделий, D – число дефектных изделий, n – размер выборки, X – число дефектных изделий в выборке. Чтобы получить число X, моделируем процесс выборки из партии n изделий без возвращения с учетом того, что вероятность извлечения дефектного изделия зависит от того, сколько их осталось в партии до данного шага. Блок-схема соответствующего алгоритма приведена на рисунке 12.

Этот вариант алгоритма моделирования значительно проще ранее рассмотренного в пункте 16.7. Но особенность данного алгоритма состоит в том, что для генерации одного значения X, распределенного по гипергеометрическому закону, требуется n стандартных чисел.

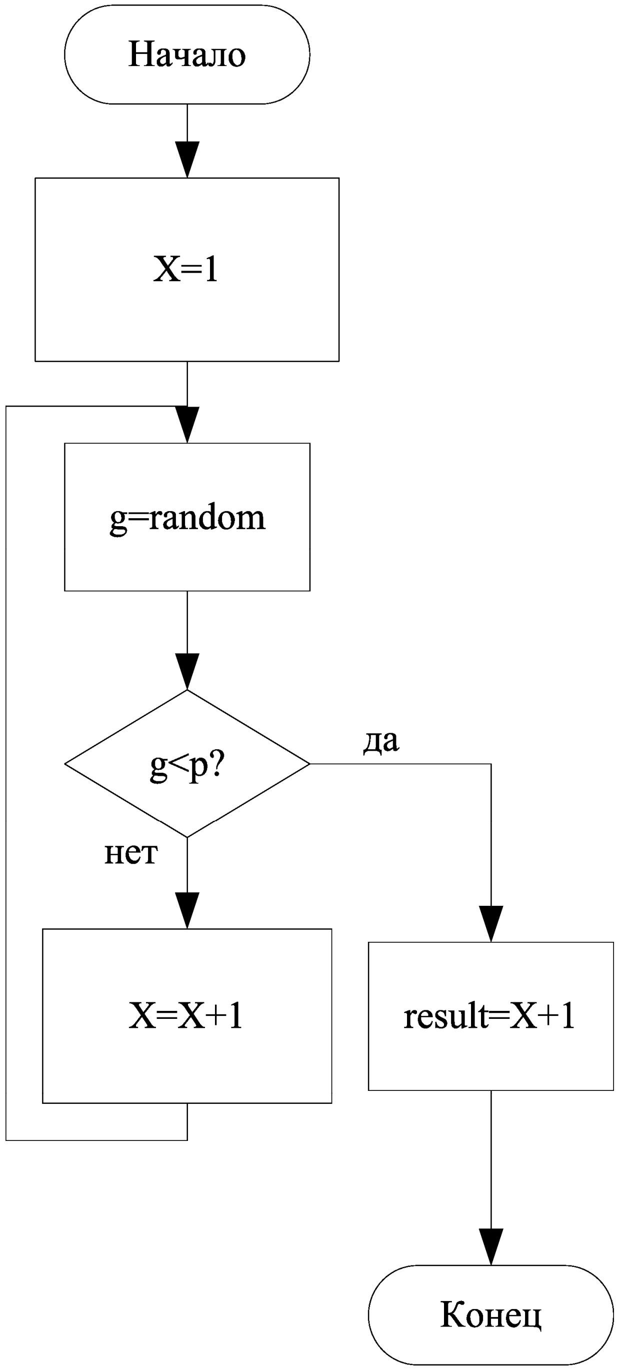

Рисунок 11 – Блок-схема алгоритма моделирования геометрического распределения

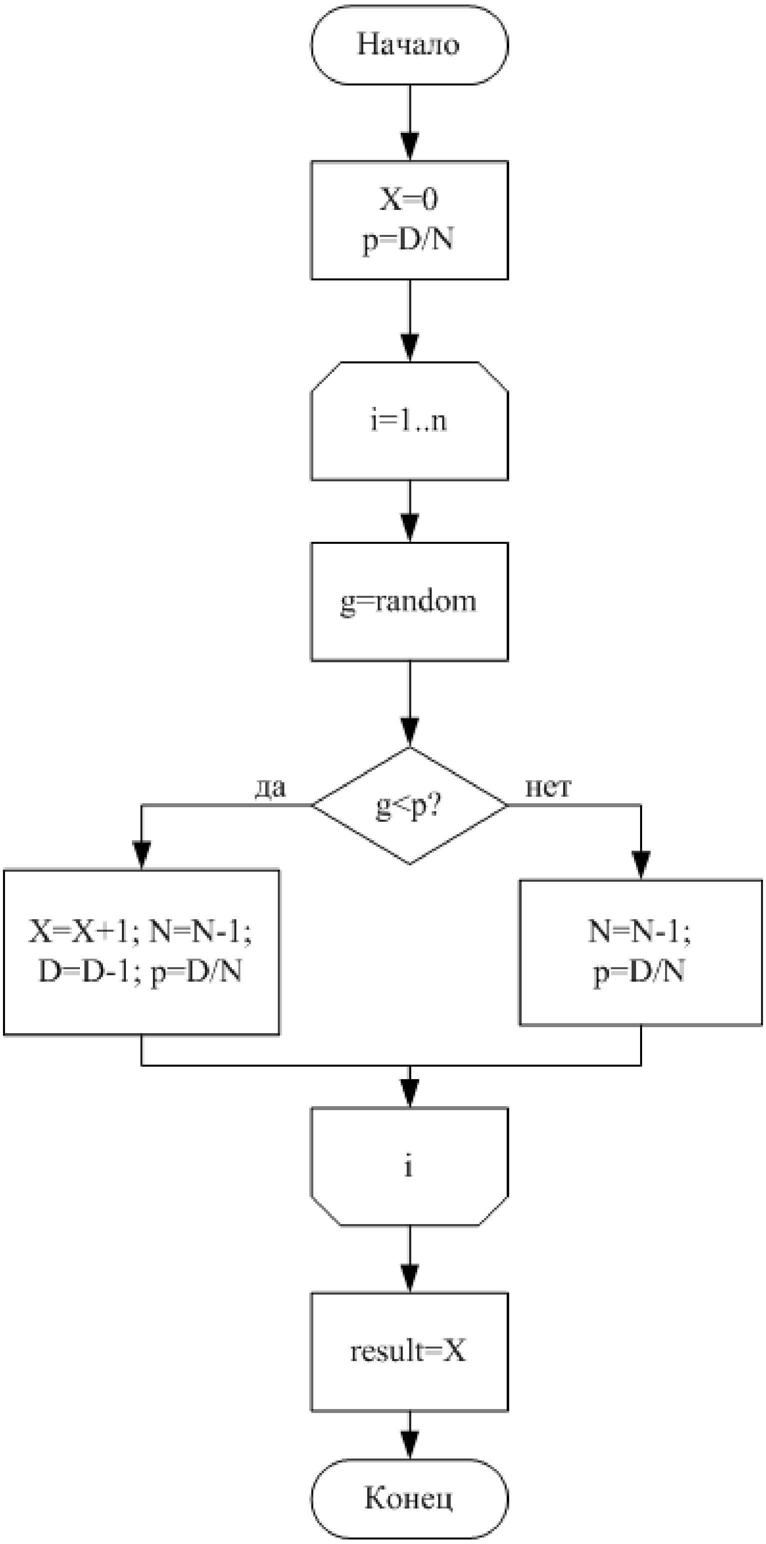

Рис. 5.14. Блок-схема алгоритма моделирования гипергеометрического распределения

16.8 Контрольные вопросы

1. Каким распределением характеризуются дискретные случайные вели-

чины?

2. Опишите алгоритм преобразования стандартного непрерывно распределенного псевдослучайного числа в дискретно распределенное число X .

3. По какой формуле считают вероятность $P(x)$ в случае распределения Пуассона?

4. По какой формуле считают вероятности $P_n(x)$ в случае биноминального распределения?

5. По какой формуле считают вероятность $P(x)$ в случае геометрического распределения?

6. По какой формуле считают вероятность $P(x)$ в случае гипергеометрического распределения?

7. Какая формула используется для генерации равномерного дискретного распределения?