Моделирование случайных факторов

Базовый генератор R(0,1)

Случайные события

Случайные величины:
- Дискретные
- Непрерывные

Случайные вектора

Случайные процессы

Построение генератора R(0,1) – квазиравномерное распределение

$$\alpha = \sum_{j=1}^{k} \alpha_j 2^j$$

, где

$$\alpha_j = \begin{cases} 0, P_j = 1/2 \\ 1, P_j = 1/2 \end{cases}$$

$$x = \alpha / (2^k - 1)$$

$$i/(2^k - 1), \; i = 0, 1, 2, \ldots, 2^k - 1$$

$$M[x] = \sum_{i=0}^{2^k - 1} \frac{i}{2^k - 1} \frac{1}{2^k} = \frac{1}{2},$$

$$D[x] = \sum_{i=0}^{2^k - 1} \left(\frac{i}{2^k - 1} - \frac{1}{2} \right)^2 \frac{1}{2^k} = \frac{1}{12} \frac{2^k + 1}{2^k - 1}.$$

Построение генератора R(0,1)

Основные методы

Методы получения псевдослучайных чисел

Методы «перемешивания»:

1. **Сдвиг предыдущего числа влево на k разрядов**
2. **Специальное сложение рез-та с предыдущим числом**

Аналитические методы:

$$x_i = \varphi(x_{i-1}, x_{i-1}, \ldots, x_{i-p})$$

p **– порядок генератора**

Метод вычетов – один из методов

Построение генератора R(0,1)

Метод вычетов

Методы получения псевдослучайных чисел

$$x_0 = \frac{\sqrt{3}}{3}, \frac{\sqrt{2}}{2}, \frac{\sqrt{5}}{5}, \frac{\sqrt{\pi}}{\pi},$$

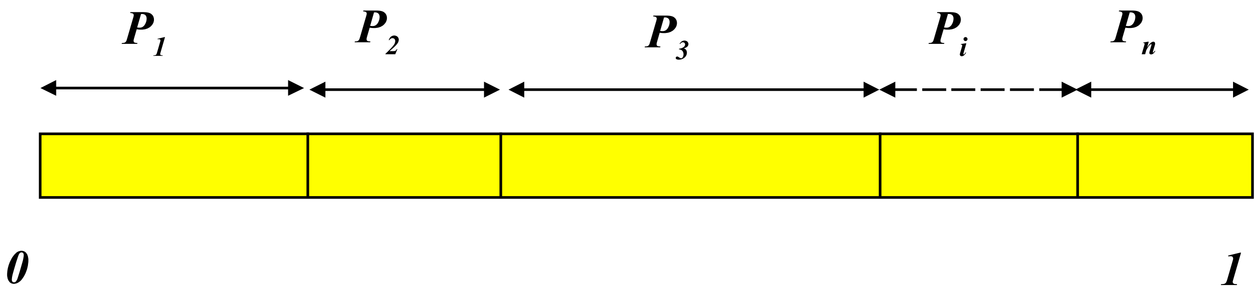
$$x_{i+1} = b \cdot x_i (\,ModM\,),$$

$$x_{i+1\,норм} = \frac{x_{i+1}}{M}, \qquad x_{i\,норм} \in [0,1]$$

$b \cdot x_i (\,ModM\,)$ - **остаток от деления** bx_i **на** M

Моделирование дискретных СВ

$$y = R(0,1)$$



$$P_2 + P_1 + P_3 + \ldots + P_i + \ldots + P_n = 1$$

$$P\left(\sum_{j=1}^{i-1} p_j \le y \le \sum_{j=1}^{i} p_j \right) = P_i$$

Моделирование законов распределения дискретных СВ

$$p_i = P(X = i),\; i = 1, 2, \ldots,$$

$$p_{i+1} = p_i r(i)$$

Биномиальное распределение:

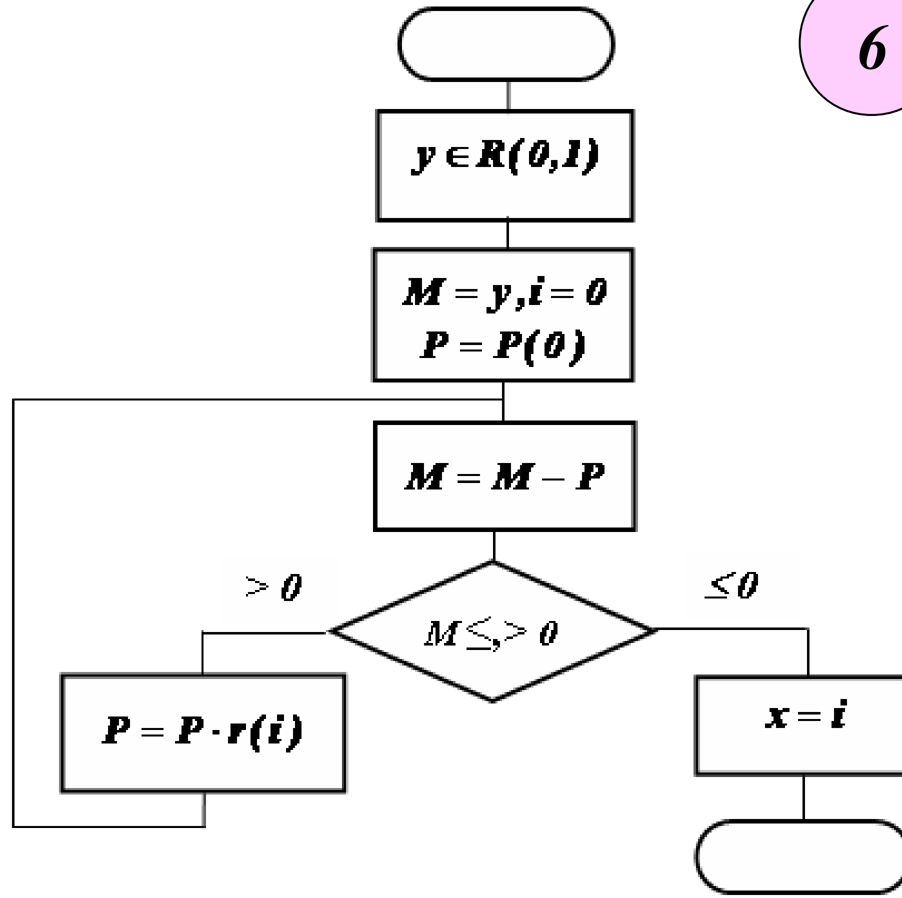
$$p_i = P(x = i) = C_n^i p^i (1 - p)^{n-i}$$

$$r(i) = \frac{p_{i+1}}{p_i} = \frac{(n-i)p}{(i+1)(1-p)}$$

Распределение Пуассона:

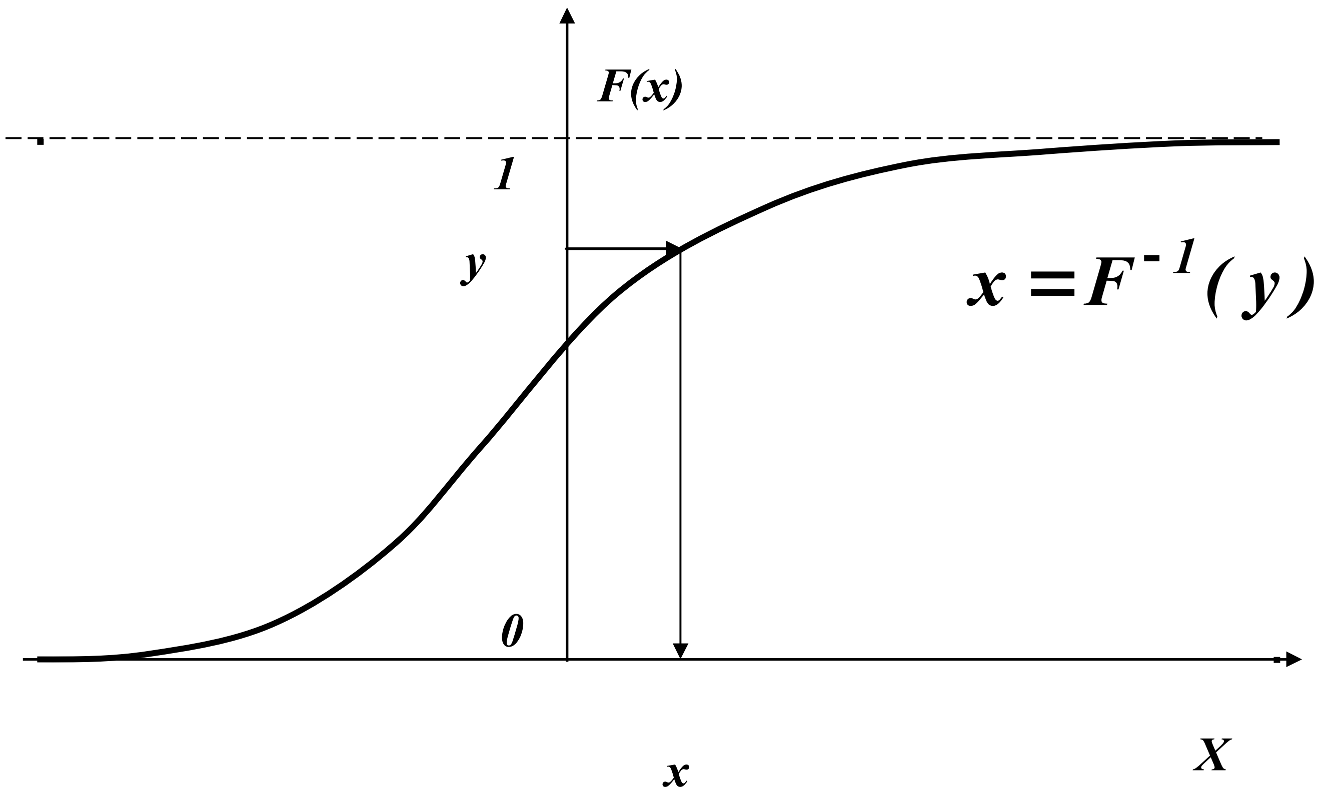
$$p_i = \frac{\lambda^i}{i!} e^{-\lambda}, \quad r(i) = \frac{\lambda}{i+1}$$

$$i = 0, 1, 2, \ldots$$



Моделирование законов распределения непрерывных СВ

Метод обратной теоремы



$$x = F^{-1}(y)$$

$$P[y_1 < y < y_2] = y_2 - y_1$$

Моделирование законов распределения непрерывных СВ

Экспоненциальный закон

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & 0 \le x < \infty \\ 0, & x < 0 \end{cases}.$$

$$y_i = F(x_i) = \int_0^{x_i} \lambda e^{-\lambda z} dz = 1 - e^{-\lambda x_i}$$

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

Моделирование случайных векторов

Пусть $$x = N(\mu, R)$$

$R = COVx$ – *ковариационная матрица*

$$y = N(O, E), \; COVy = E, \quad y_i = N(0,1)$$

$$\Downarrow$$

$$x = Ay + \mu, \; COVx = ACOVyA^T = AA^T$$

$$\Downarrow$$

$$R = AA^T$$

Найти матрицу A

Моделирование случайных векторов

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$a_{11} = \sqrt{r_{11}}, \quad a_{k1} = r_{k1} / a_{11}, \quad k = \overline{2,n}$$

$$a_{kl} = \frac{r_{kl} - \sum_{j=1}^{l-1} a_{kj} a_{lj}}{a_{ll}}, \quad l = 2,...,(k-1)$$

$$a_{kk} = \sqrt{r_{kk} - \sum_{j=1}^{k-1} a_{kj}^2}, \quad k = \overline{1,n}$$

Моделирование марковских процессов

Дискретные марковские цепи

Матрица передач

	1	2	3	4
1	P_{11}	P_{12}	P_{13}	P_{14}
2	P_{21}	P_{22}	P_{23}	P_{24}
3	P_{31}	P_{32}	P_{33}	P_{34}
4	P_{41}	P_{42}	P_{43}	P_{44}

Начальное состояние

1	2	3	4

0 **1**

P_{21}	P_{22}	P_{23}	P_{24}

Методы тестирования программных генераторов

1. Оценка моментов СВ $M[\xi]$ и $D[\xi]$

2. Тест частот (оценка формы распределения по критерию χ^2)

3. Тест пар (проверка независимости соседних значений СВ)

4. Тест на периодичность (определение интервала апериодичности)

Требования, предъявляемые к оценкам параметров распределения СВ

1. Состоятельность

$P\{|t_n - \Theta| \le \varepsilon\} > 1 - \eta$, где ε и η могут быть сколь угодно малыми числами при увеличении n. Иначе говоря, при $n \to \infty$ $t_n \to \Theta$

2. Несмещенность

$M(t_n) = \Theta$ для любого n

3. Эффективность означает, что оценка t_n является наилучшей из возможных в смысле минимальной дисперсии отклонения от значения параметра Θ:

$M\{(t_n - \Theta)^2\} \to min$ для любого n

Вычисление доверительных интервалов

Для каждого малого $\alpha > 0$ можно указать такое ε, что

$$P\left(\left|\Theta - \hat{\Theta}\right| \le \varepsilon\right) = P\left(\hat{\Theta} - \varepsilon \le \Theta \le \hat{\Theta} + \varepsilon\right) = 1 - \alpha$$

Чем меньше для данного α будет ε, тем точнее оценка $\hat{\Theta}$.

Интервал $\left[\hat{\Theta} - \varepsilon, \hat{\Theta} + \varepsilon\right]$ – доверительный интервал,

$1 - \alpha$ – доверительная вероятность,

α – уровень значимости.

Методы тестирования программных генераторов

Оценка моментов СВ

Несмещенные и эффективные точечные оценки

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \Rightarrow \quad \overline{x_N} = \overline{x_{N-1}}\,\frac{N-1}{N} + \frac{x_N}{N}$$

Рекуррентные формулы

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 \quad \Rightarrow \quad S_N^2 = S_{N-1}^2\,\frac{N-2}{N-1} + \frac{1}{N}(x_N - \overline{x_{N-1}})^2$$

Методы тестирования программных генераторов

Оценка моментов СВ

Оценка математического ожидания

$$\frac{\bar{x} - M(x)}{S^2 / \sqrt{N}} \in St(N-1)$$

$$P\left\{\bar{x} - t_{\frac{1+Q}{2}}(N-1)\frac{S}{\sqrt{N}} \le M(x) \le \bar{x} + t_{\frac{1+Q}{2}}(N-1)\frac{S}{\sqrt{N}}\right\} = Q$$

Оценка дисперсии (СКО)

$$\frac{(N-1)S^2}{\sigma^2} \in \chi^2(N-1)$$

$$P\left\{\chi^2_{\alpha}(N-1) \le \frac{(N-1)S^2}{\sigma^2} \le \chi^2_{1-\alpha}(N-1)\right\} = Q$$

$$P\left\{\frac{(N-1)S^2}{\chi^2_{1-\alpha}(N-1)} \le \sigma^2 \le \frac{(N-1)S^2}{\chi^2_{\alpha}(N-1)}\right\} = Q$$

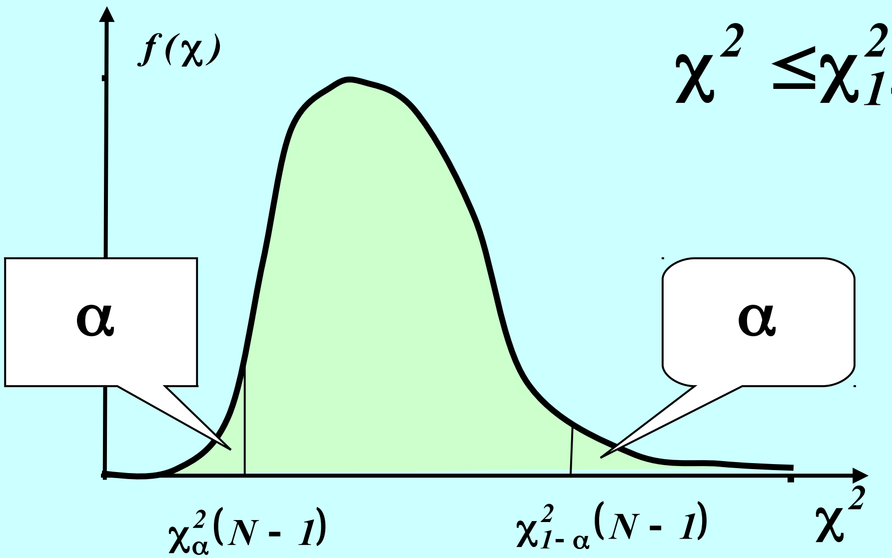
Методы тестирования программных генераторов

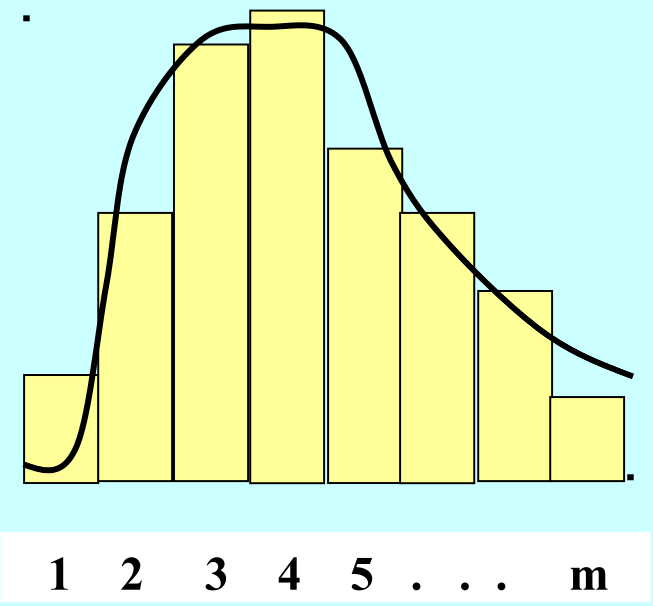
Тест частот

Критерий χ^2

$$\chi^2(m-1) = \frac{1}{NP_i}\sum_{i=1}^{m}(n_i - NP_i)^2$$

$$\chi^2 \le \chi^2_{1-\alpha}$$



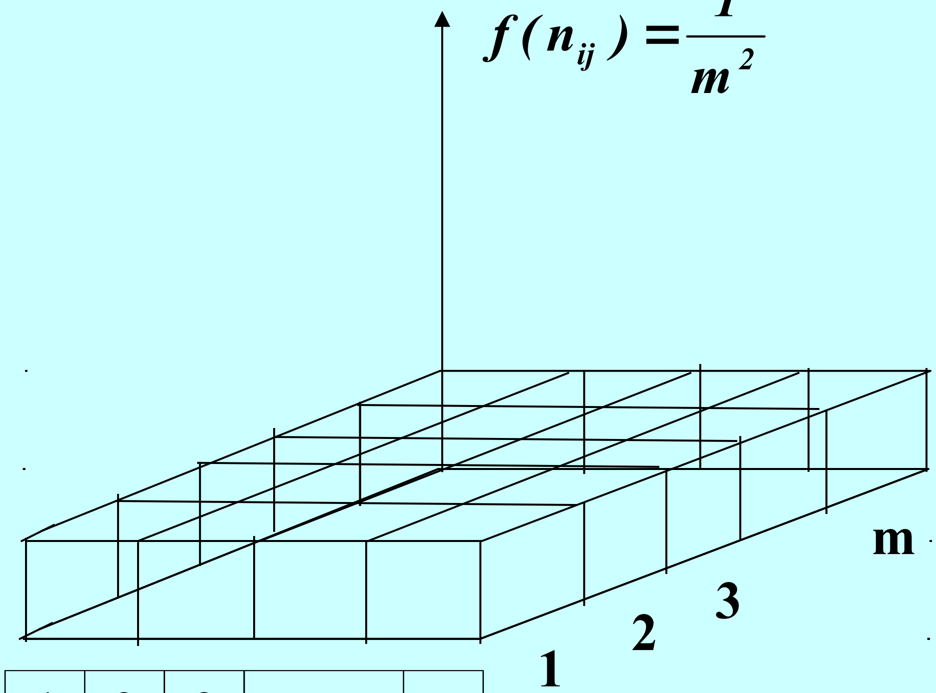


Методы тестирования программных генераторов

Тест пар

$$\chi^2\left(m^2 - m\right) = \left(\frac{2m^2}{N}\right) \sum_{i=1,\ j=1}^{m} \left(n_{ij} - \frac{N}{2m^2}\right)^2$$

$$f(n_{ij}) = \frac{1}{m^2}$$



1	2	3		m

	1	2	.	i	.	m
1						
2						
.						
j				n_{ij}		
.						
m						

Выбор объема экспериментов

Обеспечение заданной статистической точности

3 формы задачи оценки статистической точности:

1. $N = F(\varepsilon, Q)$ $\quad \varepsilon = t_{Q/2} \cdot \frac{p(1-p)}{\sqrt{N}}, \quad \varepsilon = t_{Q/2} \cdot \frac{\sigma_x}{\sqrt{N}}, \quad \varepsilon = t_{Q/2} \cdot \frac{S}{\sqrt{N}}$
2. $\varepsilon = F(N, Q)$
3. $Q = F(\varepsilon, N)$

N – объем выборки =?

$$x_i = \begin{cases} 0, & p \\ 1, & 1-p \end{cases}$$ → **Найти точность оценки вероятности p**

$$\hat{p} = \frac{m}{N}, \text{ где } m = \sum_{i=1}^{N} x_i \Rightarrow D\{\hat{p}\} = D\left\{\frac{\sum_{i=1}^{N} x_i}{N}\right\} = \frac{1}{N^2} D\left\{\sum_{i=1}^{N} x_i\right\} = \frac{D\{x_i\}}{N}$$

$$\overline{x_i} = p, \quad D\{x_i\} = p(1-p)$$

$$D\{\hat{p}\} = \frac{p(1-p)}{N}$$