

Лекция №3

Раздел 1. Модели стохастических факторов

5.2 Приближенные методы

5.2.1 Методы, основанные на воспроизведении условий соответствующей предельной теоремы теории вероятностей

5.2.1.1 Моделирование случайной величины с нормальным законом распределения вероятностей

Пусть требуется получить реализации случайной величины, имеющей нормальное распределение с математическим ожиданием m и среднеквадратическим отклонением σ .

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

На Рисунке 9 представлены плотность и функция распределения случайной величины с нормальным законом распределения вероятностей.

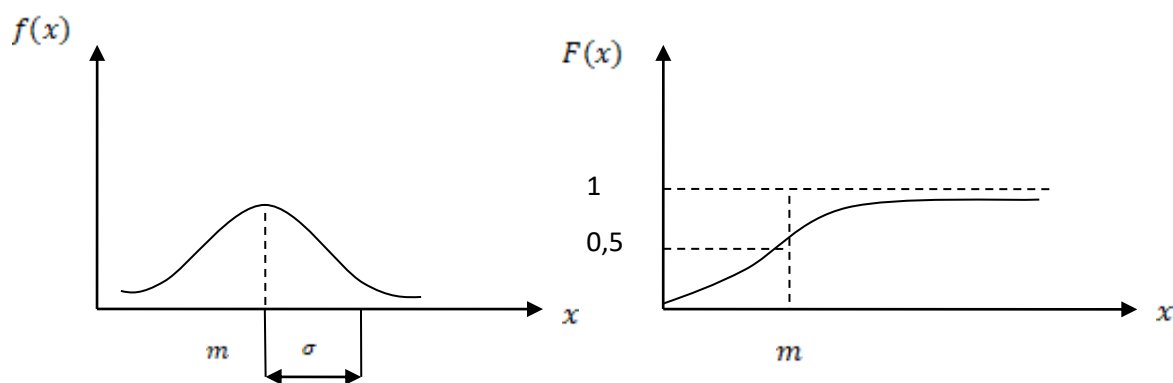


Рисунок 9 – плотность и функция распределения случайной величины с нормальным законом распределения вероятностей

Для моделирования случайной величины с нормальным законом распределения вероятностей можно использовать метод обратной функции, но тогда надо вводить таблицу значений интегральной функции распределения, что неудобно.

На практике для моделирования случайной величины с нормальным законом распределения вероятностей используют метод, основанный на центральной предельной теореме теории вероятностей.

На основании центральной предельной теоремы, реализация случайной величины (СВ) с нормальным законом распределения вероятностей вычисляется как сумма n независимых реализаций равномерно распределенной СВ.

$$norm = \sum_{i=1}^n r_i, \quad r_i \in R \sim [0,1) \quad (23)$$

$$M(norm) = M\left(\sum_{i=1}^n r_i\right) = \sum_{i=1}^n M r_i = \frac{n}{2}$$

$$D(norm) = D\left(\sum_{i=1}^n r_i\right) = \sum_{i=1}^n D r_i = \frac{n}{12}$$

Достаточным считается $n = 12$, следовательно, $M(norm) = 6$, $D(norm) = 1$.

Для получения реализации нормированной нормальной случайной величины $nnorm \sim (0,1)$ величину $norm$ необходимо центрировать

$$nnorm = (norm - 6)$$

В общем случае для получения нормированной нормальной величины при $n \neq 12$, производятся две операции – центрирование и нормирование

$$nnorm = \underbrace{\left(norm - \frac{n}{2}\right)}_{\text{центрирование}} * \underbrace{\frac{\sqrt{12}}{\sqrt{n}}}_{\text{нормирование}} \quad (24)$$

$$M(nnorm) = \frac{\sqrt{12}}{\sqrt{n}} M\left(norm - \frac{n}{2}\right) = 0$$

$$D(nnorm) = \frac{12}{n} D\left(norm - \frac{n}{2}\right) = \frac{12}{n} \left[\frac{n}{12} + 0\right] = 1$$

Для получения реализаций СВ $snorm \sim (m, \sigma^2)$ с произвольным математическим ожиданием и дисперсией используется формула

$$snorm = \sigma * nnorm + m \quad (25)$$

5.2.1.2. Моделирование случайной величины с биномиальным законом распределения вероятностей

Биномиальное распределение является распределением вероятности появления k событий в n независимых испытаниях, в каждом из которых вероятность появления события равна p . Вероятность возможного числа появления событий вычисляется по формуле Бернулли

$$P_n(X = k) = C_n^k p^k (1 - p)^{n-k} \quad (26)$$

$X=k$	0	1	\dots	n
$P_n(x = k)$	$(1 - p)^n$	$np(1 - p)^{n-1}$	\dots	p^n

Моделирующий алгоритм основан на представлении реализации случайной величины X , подчиненной биномиальному закону распределения в виде суммы n независимых реализаций случайной величины $Y \sim y_1, y_2, \dots$, имеющей распределение

y_i	1	0
p_i	p	$1-p$

Алгоритм получения реализаций СВ X , распределенной по биномиальному закону:

- 1) Получить реализацию $r_i \in R \sim [0,1)$;
- 2) Для каждой реализации r_i проверить выполнение неравенства

$$r_i < p, i = 1, 2, \dots, n$$

Если неравенство выполняется, $y_i = 1$, иначе $y_i = 0$.

- 3) Вычислить сумму n реализаций СВ Y и принять эту сумму за реализацию СВ X , имеющий смысл числа успехов в вероятностной схеме испытаний Бернулли

$$X = k = \sum_{i=1}^n y_i \quad (27)$$

При повторении этого алгоритма s раз, получаем последовательность значений $x^{(1)}, x^{(2)}, \dots, x^{(s)}$, которые являются реализациями случайной величины X с биномиальным законом распределения вероятностей.

5.2.1.3 Моделирование случайной величины, распределенной по закону Пуассона

Алгоритм моделирования 2.

В основе алгоритма лежит предельная теорема Пуассона, согласно которой закон Пуассона является предельным для биномиального: если p – вероятность наступления события при одном испытании, то вероятность наступления k событий в n независимых испытаниях при $n \rightarrow \infty, p \rightarrow 0, np = \lambda$ асимптотически равна

$$\lim_{n \rightarrow \infty} C_n^k p^k (1-p)^{n-k} = \lim_{n \rightarrow \infty} C_n^k \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\text{где } p = \frac{\lambda}{n}.$$

Моделирование основано на соотношении между экспоненциальным распределением и распределением Пуассона. Если промежутки времени между появлениями событий на некотором временном интервале распределены по экспоненциальному закону, то число событий на этом интервале распределено по закону Пуассона.

Пусть реализации СВ $Y \sim y_1, y_2, \dots$ независимы и имеют экспоненциальный закон распределения с математическим ожиданием 1, тогда неотрицательное целое число k , для которого выполняется неравенство (28), являются реализацией СВ, распределенной по закону Пуассона с параметром λ :

$$\sum_{i=1}^k y_i \leq \lambda < \sum_{i=1}^{k+1} y_i \quad (28)$$

Так как реализация СВ с экспоненциальным законом распределения и математическим ожиданием, равным 1, определяется как $y_i = -\ln r_i, r_i \in R \sim [0,1)$, то от неравенства (28) переходим к (29)

$$\prod_{i=1}^{k+1} r_i \leq e^{-\lambda} < \prod_{i=1}^k r_i \text{ или } \prod_{i=1}^k r_i > e^{-\lambda} \geq \prod_{i=1}^{k+1} r_i \quad (29)$$

Алгоритм получения СВ, распределенной по закону Пуассона:

- 1) Получить последовательности независимых реализаций $\{r_i\}$ равномерно распределенной СВ, $r_i \in R \sim [0,1)$;
- 2) Вычислять произведения $r_1 \cdot r_2, r_1 \cdot r_2 \cdot r_3, \dots, r_1 \cdot r_2 \cdot \dots \cdot r_k$ до тех пор, пока не выполнится неравенство (29);
- 3) При выполнении неравенства (29) значение СВ $X = k$.

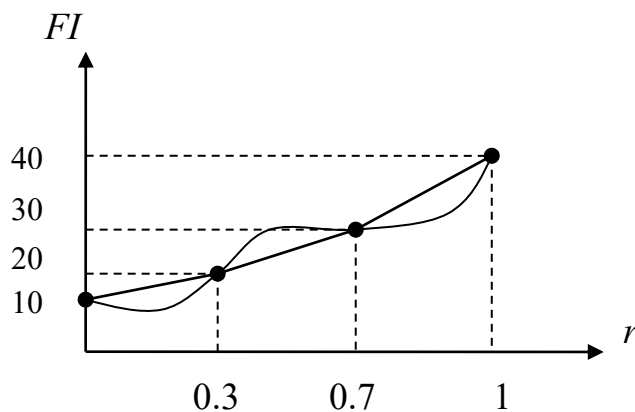
5.2.2 Методы, основанные на аппроксимации функции распределения

На практике часто бывает необходимо получить распределение СВ по эмпирической функции распределения, построенной на основе некоторой статистической информации, собранной на модели или реальной системе. В этом случае функция распределения $F(x)$ аппроксимируется некоторой функцией $G(x)$, обратная функция от которой $G^{-1}(x)$ имеет достаточно простое аналитическое выражение и легко вычислима. Чаще всего вместо определения функции $G(x)$ используется кусочно-линейная аппроксимация, заключающаяся в том, что функция распределения $F(x)$ заменяется функцией $F^*(x)$, составленной из отрезков прямых. Для этого интервал изменения $F(x)$ разбивается на n подинтервалов, число которых зависит от требуемой точности конечных результатов. На каждом из линейных участков функции $F^*(x)$, соответствующих этим подинтервалам, моделируется случайная величина с равномерным законом распределения вероятностей согласно методу обратной функции.

Пример.

Пусть о СВ X собрана следующая статистика: 30% значений находятся в диапазоне $[10,20]$, 40% значений – в диапазоне $[20,30]$, 30% значений – в диапазоне $[30,40]$.

Функция моделирования СВ X , образованная путем инверсии $F(x)$, имеет вид.



Алгоритм моделирования:

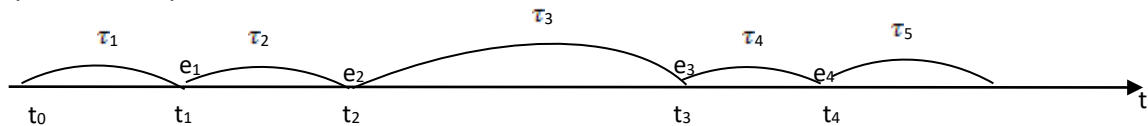
- 1) Получить реализации r_1 и r_2 , $r_i \in R \sim [0,1]$;
 $r_1 = \text{random}(0,1)$;
 $r_2 = \text{random}(0,1)$;
- 2) $\text{if } (r_1 \leq 0.3) \text{ FI} = 10 + (20 - 10) * r_2$
 $\text{else if } (r_1 \leq 0.7) \text{ FI} = 20 + (30 - 20) * r_2$
 $\text{else FI} = 30 + (40 - 30) * r_2$

6. Моделирование информационных потоков

6.1. Понятие информационного потока

Процесс функционирования любой сложной системы во времени можно рассматривать, как последовательность моментов времени, в которые система изменяет свое состояние, т.е. происходит некоторое событие. Последовательность таких событий в системе образует информационный поток.

Информационный поток, в котором события не отличаются одно от другого, является потоком однородных событий. Он определяется моментами $\{t_1, t_2, t_3, \dots\}$ событий $\{e_1, e_2, e_3, \dots\}$.



t_0 – момент начала моделирования

τ_i – интервалы времени между наступлениями двух соседних событий, которые представляют собой реализации СВ T с заданным законом распределения вероятностей.

Информационные потоки, для которых можно рассчитать моменты наступления следующих событий на основании моментов наступления предыдущих, называются рекуррентными.

$$\begin{cases} t_1 = t_0 + \tau_1 \\ t_2 = t_0 + \tau_1 + \tau_2 \\ t_k = t_0 + \sum_{i=1}^k \tau_i \end{cases} \quad (30)$$

Таким образом, для моделирования информационного потока достаточно задать закон распределения $F(\tau)$ случайной величины $T \sim \tau_1, \tau_2, \dots$, реализации которой определяют интервалы времени между наступлениями соседних событий.

6.2 Моделирование стационарных потоков однородных событий

Информационные потоки, в которых все события относятся к одному типу, называются *потоками однородных событий*.

Поток однородных событий называется *стационарным*, если вероятность $P_k(t_0, t)$ появления k событий за промежуток времени (t_0, t_0+t) не зависит от t_0 , а зависит только от t и k .

Поток однородных событий называется *потоком без последствия*, если вероятность $P_k(t_0, t)$ наступления k событий за промежуток времени (t_0, t_0+t) не зависит от наступления событий до момента t_0 .

Поток однородных событий называется *ординарным*, если вероятность $P_k(t_0, t)$ наступления двух и более событий за промежуток времени (t_0, t_0+t) при любом значении t_0 является бесконечно малой величиной.

На практике широко используется **стационарный, ординарный поток однородных событий без последствия**, называемый **простейшим**.

В простейшем потоке интервалы времени между событиями распределены по экспоненциальному закону

$$F(\tau) = \lambda e^{-\lambda\tau}, \lambda > 0, \tau \geq 0, \quad (31)$$

а число событий на интервале времени распределено по закону Пуассона

$$P\{k\} = \frac{a^k}{k!} e^{-a}; \quad k = 0, 1, 2, \dots \quad (32)$$

с параметром $a = \lambda\tau$, λ – интенсивность (или параметр) потока ($\lambda > 0$).

Также широко используется **информационный поток Эрланга**. Поток Эрланга k -го порядка с параметром (интенсивностью) $\lambda_{эрл}$ называется стационарный ординарный поток с ограниченным последствием. Применительно к распределению Эрланга воспользоваться методом нелинейного преобразования, обратной функции распределения не удастся, ввиду сложности вычисления обратной функции. Поэтому при моделировании случайных величин, распределенных по закону Эрланга, используют то обстоятельство, что этому закону распределения подчиняется сумма k независимых реализаций случайной величины, распределенной по экспоненциальному закону

$$T_{эрл} = \sum_{i=1}^k T_{iэксп} \quad (33)$$

с параметром

$$\lambda_{эксп} = k * \lambda_{эрл}. \quad (34)$$

Для формирования на ЭВМ реализаций потоков событий нужно, последовательно обращаясь к датчику случайных чисел, формировать реализации $\tau_1, \tau_2, \tau_3, \dots$ случайной величины T с законом распределения вероятностей $F(\tau)$. Суммирование получаемых реализаций позволит определить моменты наступления событий в системе:

$$t_k = t_{k-1} + \tau_k, k = 1, 2, 3, \dots$$

В качестве законов распределения $F(\tau)$ могут использоваться самые различные:

- 1) экспоненциальный,
- 2) распределение Эрланга,
- 3) равномерное распределение,
- 4) усеченно-нормальное распределение и т.п.

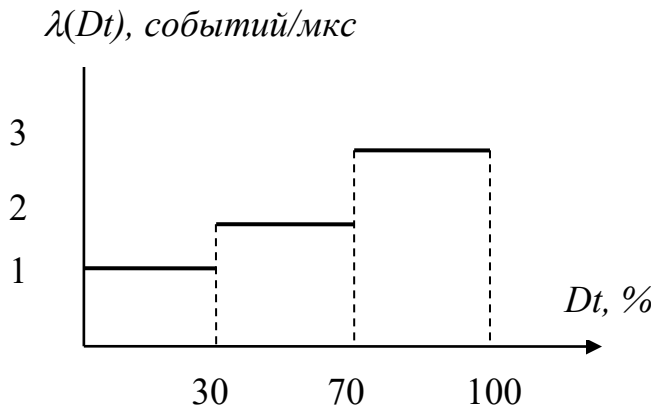
6.3 Моделирование нестационарных потоков однородных событий

Нередко потоки событий в системах отличаются *нестационарностью* – случай, когда параметр λ оказывается зависящим от времени $\lambda(t)$. В качестве примера нестационарного потока рассмотрим поток Пуассона с переменным параметром, для которого интегральная функция распределения величины τ , отделяющей моменты наступления двух соседних событий, зависит от того, где на оси времени расположено первое из событий и определяется выражением

$$P(\tau < u) = F(u, t) = 1 - e^{-\Lambda(t, u)}, \quad \Lambda(t, u) = \int_t^{t+u} \lambda(t) dt \quad (35)$$

Интенсивность нестационарного потока зависит от начала интервала и от ширины интервала на оси времени. Постоянство $\lambda(t)$ на некотором интервале (t_{k-1}, t_k) , $k=1,2,3\dots$ позволяет рассматривать нестационарный поток как локально-стационарный и использовать для его моделирования метод обратной функции.

В качестве модели зависимости интенсивности потока от времени может быть выбрана финитная функция нестационарности (ФН), определенная на интервале времени, называемом периодом нестационарности. В качестве такого периода может быть выбран произвольный временной интервал, для которого известна или эмпирически определена зависимость интенсивности потока от времени (например, 1 с).



В качестве аргумента ФН используется относительное время, представляющее собой процент длительности от периода нестационарности (36)

$$Dt = (t \bmod T) / T * 100, \quad (36)$$

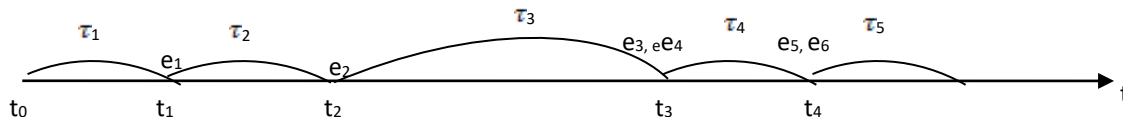
где t – текущее значение времени, T – период нестационарности, \bmod – операция деления по модулю. Таким образом, ФН определяет повторяющиеся изменения интенсивности потока. Функция нестационарности вычисляется с помощью кусочно-постоянной интерполяции, т.е. в течение 300 мкс (30% от 1 с) интенсивность потока составляет в среднем 1 запрос в мкс, в течение следующих 400 мкс (40% от периода нестационарности) – 2 запроса в мкс, в течение оставшихся 300 мкс (30% от периода нестационарности) – 3 запроса в мкс.

Такое определение функции нестационарности позволяет рассматривать поток как локально-стационарный и использовать для его моделирования метод обратной функции, согласно которому интервал времени между событиями в потоке определяется по формуле (37)

$$\Delta t = - (1 / \lambda(Dt)) * \ln r_i, \quad r_i \in R \sim [0,1) \quad (37)$$

6.4 Моделирование неординарных потоков однородных событий

В неординарных потоках в моменты t_i может наступать не одно, а несколько событий (группа событий). Для того чтобы описать неординарный поток, кроме моментов t_i наступления событий необходимо задать распределение количества событий, наступающих в каждый из моментов времени t_i .



Если число наступающих событий является случайной величиной, независимой от t_i , достаточно задать вероятность p_k того, что в произвольный момент t_i наступает k событий. Количество событий моделируется с помощью дискретной случайной величины с ограниченным спектром значений.

$$X = \left\{ \begin{array}{ll} x_0=0 & \dots \quad x_k=k \\ p_0=p(\text{число событий в момент } t_i=0) & \dots \quad p_k=p(\text{число событий в момент } t_i=k) \end{array} \right\}$$

6.5 Моделирование потоков неоднородных событий

В потоке неоднородных событий в моменты t_i ($i = 1, 2, 3, \dots$) может появиться одно из нескольких событий E_k , $k=1..K$, причем, события несовместны и образуют полную группу

$$\sum_{k=1}^K P(E_k) = 1.$$

Подобный поток моделируется как обычный однородный, т.е. последовательно определяются моменты появления t_i очередного события, но в дополнение к этому для каждого t_i методом моделирования определяется тип события, т.е. какое из событий E_k произошло в момент t_i . Это осуществляется путем моделирования дискретной случайной величины X с ограниченным спектром значений:

$$x = \left\{ \begin{array}{l} x^{(0)} := E_1, x^{(1)} := E_2, \dots, x^{(K-1)} := E_K \\ P_0 = P(E_1), P_1 = P(E_2), \dots, P_{K-1} = P(E_K) \end{array} \right\}.$$