14 РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

14.1 Распределение неслучайной величины

Согласно теории вероятностей неслучайная величина является частным случаем случайной и по этому как любая случайная величина она характеризуется соответствующим законом распределения. А именно, если $X = a$, где a константа, то

$$F(x) = \begin{cases} 0 & при\ x < a, \\ 1 & при\ x \geq a \end{cases}.$$

Графическая иллюстрация этого закона приведена на рисунке 14.1.

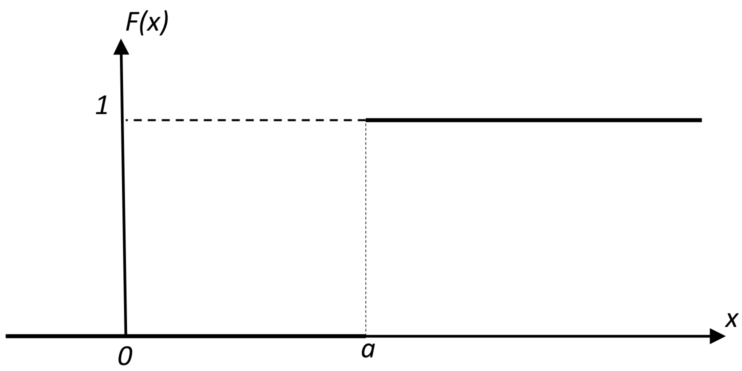

Рисунок 14.1

Плотность распределения неслучайной величины выражается через обобщенную функцию Дирака $\delta(x)$. То есть

$$f(x) = \delta(x-a) = \begin{cases} 0 & при\ x < a, \\ \infty & при\ x = 0, \\ 0 & при\ x > 0. \end{cases}$$

Эта плотность подчиняется условию нормировки вероятностей

$$\int_{-\infty}^{\infty} \delta(x-a)dx = 1.$$

Математическое ожидание

$$\overline{X} = \int_{-\infty}^{\infty} \delta(x-a)x dx = a \ .$$

Дисперсия

$$D_X = \int_{-\infty}^{\infty} \delta(x-a)(x-a)^2 dx = 0 \, .$$

Медиана и мода равны *a* как и математическое ожидание.

С помощью функции Дирака можно аналитически выражать плотности дискретных случайных величин. Например, пусть случайная величина может принимать только два значения *x1, x2* с вероятностями *p1, p2* соответственно, тогда

$$f(x) = p1 \cdot \delta(x - x1) + p2 \cdot \delta(x - x2) \, .$$

14.2 Равномерное распределение

В некоторых задачах встречаются непрерывные случайные величины, все возможные значения которых лежат в пределах определённого интервала и равновероятны. Говорят, что они распределяются по закону равномерной плотности. Примеры: рулетка, угол поворота сбалансированного колеса и т. д.

Рассмотрим случайную величину, подчинённую закону равномерной плотности на участке от α до β.

Плотность распределения *f(x)* постоянна и равна *C* на участке от α до β. Вне этого отрезка она равна 0.

$$f(x) = \begin{cases} c \quad при \quad \alpha < x < \beta, \\ 0 \quad при \quad x < \alpha \quad или \quad x > \beta. \end{cases}$$

Так как площадь под кривой распределения равна 1:

$$c(\beta - \alpha) = 1,$$

$$c = \frac{1}{\beta - \alpha}.$$

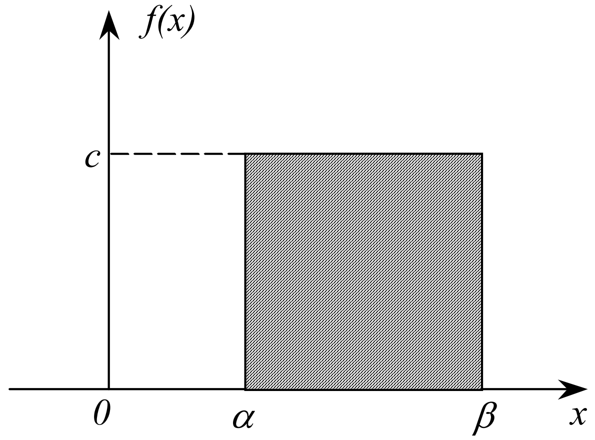

Рисунок 14.2 – График равномерной плотности

Функция распределения:

$$F(x) = \begin{cases} 0 \quad при \quad x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} \quad при \quad \alpha < x < \beta \\ 1 \quad при \quad \beta < x \end{cases}$$

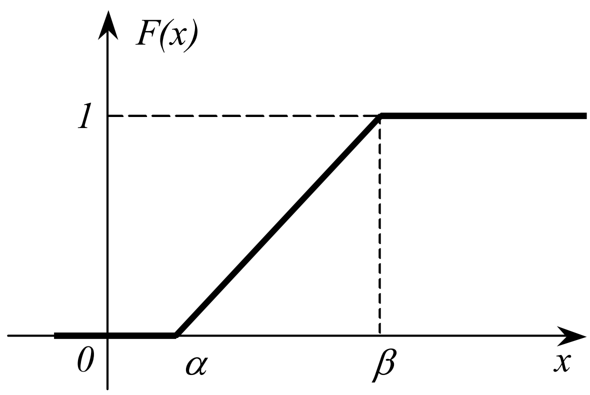

Рисунок 14.3 – График равномерного распределения

Определим основные числовые характеристики случайной величины с равномерным распределением на участке от α до β.

Математическое ожидание:

$$m_x = \int_{\alpha}^{\beta} \frac{x}{\beta-\alpha} dx = \frac{\alpha+\beta}{2}.$$

В силу симметричности распределения медиана также равна $\dfrac{\alpha+\beta}{2}$. Моды у закона равномерной плотности нет.

Дисперсия:

$$D_x = \mu_2 = \frac{1}{\beta - \alpha}\int\limits_{\alpha}^{\beta}\left(x - \frac{\alpha + \beta}{2}\right)dx = \frac{(\beta - \alpha)^2}{12}.$$

Среднее квадратичное отклонение:

$$\sigma_x = \sqrt{D_x} = \frac{\beta - \alpha}{2\sqrt{3}}.$$

Асимметрия у симметричного распределения равна нулю.

$$Sk = 0.$$

Для определения эксцесса находим четвёртый центральный момент:

$$\mu_4 = \frac{1}{\beta - \alpha}\int\limits_{\alpha}^{\beta}\left(x - \frac{\alpha + \beta}{2}\right)^4 dx = \frac{(\beta - \alpha)^4}{80}$$

$$Ex = \frac{\mu_4}{\sigma_4} - 3 = -1{,}2.$$

Найдём вероятность попадания случайной величины X в участок *[a,b]*.

$$P(a < X < b) = \frac{b - a}{\beta - \alpha}.$$

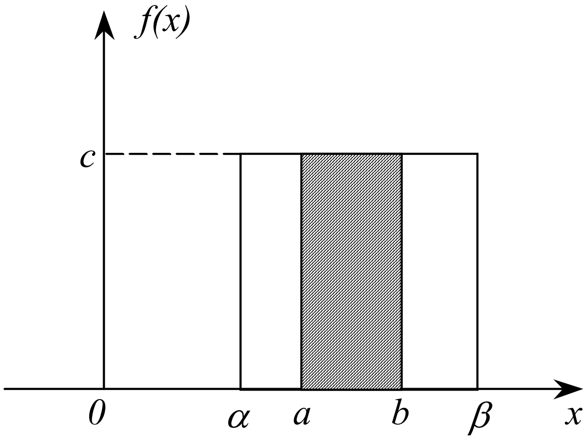

Рисунок 14.4 – К определению вероятности $P(a < X < b)$

14.3 Распределение Симпсона

Плотность этого распределения имеет вид равнобедренного треугольника (см. рисунок 14.5).

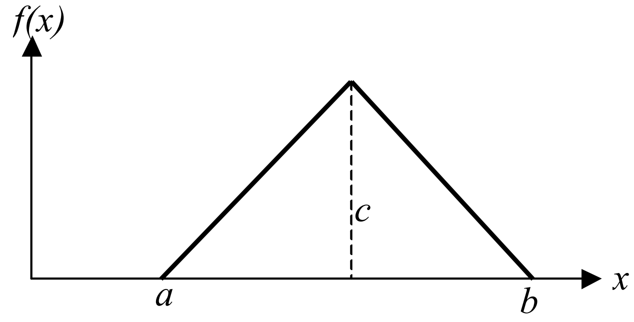

Рисунок 14.5 – График плотности Симсона

Аналитическое выражение для плотности имеет вид:

$$f(x) = \begin{cases} \dfrac{4(x-a)}{(b-a)^2} \text{ } при \text{ } a \le x \le \dfrac{b+a}{2}; \\ \dfrac{4(b-x)}{(b-a)^2} \text{ } при \text{ } \dfrac{b+a}{2} \le x \le b. \end{cases}$$

При $x < a$ и $x > b$ плотность $f(x) = 0$, $c = 2/(b-a)$.

Математическое ожидание

$$\overline{X} = M(X) = \frac{a+b}{2}.$$

Так как распределение симметрично относительно среднего значения, то мода и медиана совпадают с математическим ожиданием.

Дисперсия

$$D(X) = \frac{(b-a)^2}{24}.$$

Квадратичное отклонение

$$\sigma_X = \sqrt{D(X)} = \frac{(b-a)}{2\sqrt{6}}.$$

Ассиметрия $S_k = 0$ из-за симметричности распределения.

Четвертый центральный момент

$$\mu_4 = \frac{(b-a)^4}{240}.$$

Эксцесс $Ex = \dfrac{\mu_4}{\sigma_X^4} - 3 = -0.6$.

14.4 Биномиальное распределение

Дискретная случайная величина *X,* принимающая значения $x_m=m$, где $m=0,1,\dots,n$, имеет биномиальное распределение, если вероятности её значений определяются следующей формулой:

$$P_n(m) = P(X = m) = C_n^m p^m (1-p)^{n-m}, \quad m = 0, 1, \dots, n.$$

Здесь

$$C_n^m = \frac{n!}{m!(n-m)!} -$$

число сочетаний из n по m, а параметр p имеет смысл вероятности, то есть $0 \le p \le 1$.

Это распределение связано с повторением опытов. Если в результате опыта событие A имеет вероятность p и опыт повторяется n раз, то вероятность того, что это событие произойдет m раз, рана $P_n(m)$. Действительно, конкретная реализация n испытаний, в которых событие A произошло m раз, а противоположное событие $\overline{A}$ соответственно n-m раз, имеет вероятность $p^m(1-p)^{n-m}$. Но m событий среди n испытаний могут распределиться C_n^m равновозможными способами. Отсюда и получается формула для биномиального распределения.

Сумма

$$\sum_{m=1}^{n} P_n(m) = \sum_{m=1}^{n} C_n^m p^m q^{n-m} = (p+q)^n = 1,$$

так как $q=1-p$ а $p+q=1$. Выражение $P_n(m)$ является членом разложения бинома Ньютона $(p+q)^n$, поэтому это распределение называется биномиальным.

Математическое ожидание

$$M(X) = \sum_{m=0}^{n} C_n^m p^m q^{n-m} \cdot m = np \sum_{m=1}^{n} C_{n-1}^{m-1} p^{m-1} q^{n-m} = np.$$

Дисперсия

$$D(X)=\sum_{m=1}^{n} C_n^m p^m q^{n-m}\cdot m^2-(np)^2=npq=np(1-p).$$

Квадратичное отклонение

$$\sigma_X=\sqrt{np(1-p)}\,.$$

Если n устремить к бесконечности и одновременно p к нулю так, чтобы выполнялось соотношение

$$p=\frac{a}{n},$$

где a положительная константа, то в пределе

$$P(m)=\frac{a^m}{m!}e^{-a},$$

а это известное распределение Пуассона. То есть в пределе при $n\to\infty$ и $p=a/n$ биномиальное распределение совпадает с распределением Пуассона.

14.5 Распределение Пуассона

Рассмотрим дискретную случайную величину *X*, которая может принимать только целые неотрицательные значения: 0, 1, 2, …

Говорят, что случайная величина *X* распределена по *закону Пуассона*, если вероятность того, что она примет определённое значение *m*, выражается формулой

$$P_m=\frac{a^m}{m!}e^{-a}\quad (m=0,1,\ldots),$$

где a – некоторая положительная величина, называемая *параметром распределения Пуассона*.

Ряд распределения по закону Пуассона:

x_m	0	1	2	…	m
p_m	e^{-a}	$\frac{a}{1!}e^{-a}$	$\frac{a^2}{2!}e^{-a}$	…	$\frac{a^m}{m!}e^{-a}$

Убедимся, что суммарная вероятность равна единице.

$$\sum_{m=0}^{\infty} P_m = \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!}.$$

Но $\sum_{m=0}^{\infty} \frac{a^m}{m!} = e^a$,

следовательно $\sum_{m=0}^{\infty} P_m = e^{-a} e^a = 1.$

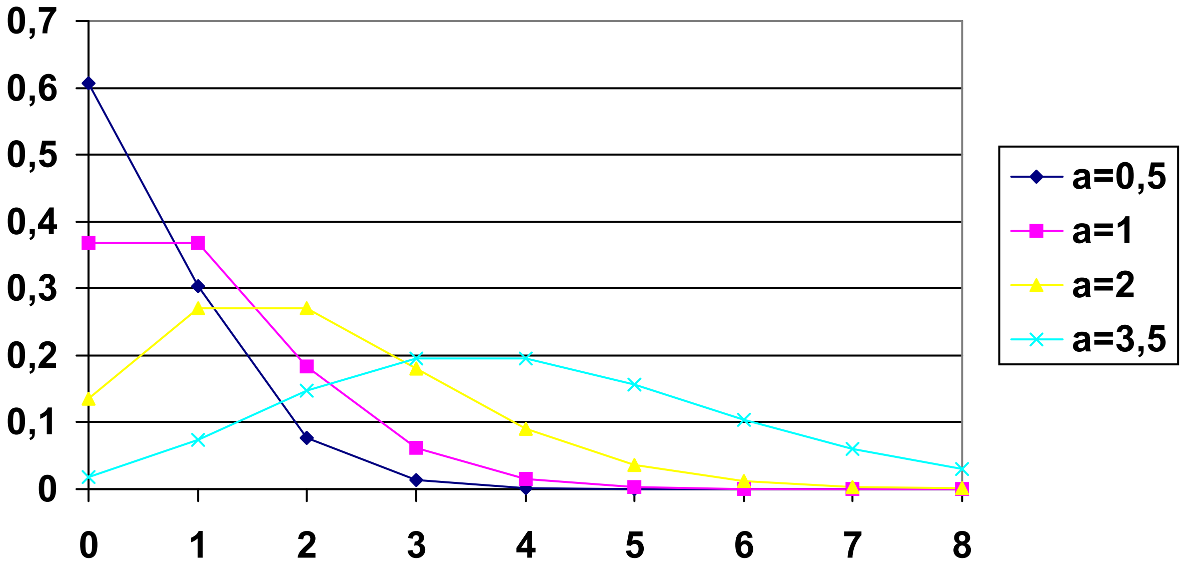

Рисунок 14.6 – Полигон распределения Пуассона

Вычислим вероятность того, что X окажется больше 0:

$$R_1 = 1 - P_0 = 1 - e^a.$$

Математическое ожидание

$$m_x = \sum_{m=0}^{\infty} m P_m = \sum_{m=0}^{\infty} m \frac{a^m}{m!} e^{-a} = \sum_{m=1}^{\infty} \frac{ma}{m} \frac{a^{m-1}}{(m-1)!} e^{-a} = a e^{-a} \sum_{m=1}^{\infty} \frac{a^{m-1}}{(m-1)!} = a e^{-a} e^a = a.$$

Т.е. параметр a – есть математическое ожидание.

Дисперсия

$$D_x = a.$$

Вывод. Дисперсия равна математическому ожиданию. Это свойство часто используют для определения, распределена ли случайная величина X по закону Пуассона.

Закон Пуассона является предельным для биномиального распределения.

$$P_{m,n} = C_n^m p^m (1-p)^{n-m},$$

Причём параметр $a = np$.

Предельное свойство биномиального распределения можно записать в виде:

$$\lim_{n\to\infty} C_n^m p^m (1-p)^{n-m} = \frac{a^m}{m!} e^{-a}$$

Из-за малой вероятности события закон Пуассона носит название «закона редких явлений».

14.6 Геометрическое распределение

Геометрическое распределение выражается следующим образом:

$$P(X=n) = p(1-p)^{n-1}, \quad (n = 1,2,3,...). \tag{14.1}$$

Название распределения связано с тем, что вероятности при различных *n* образуют геометрическую прогрессию со знаменателем, равным $q=1-p$. Действительно

$$\frac{P(X=n+1)}{P(X=n)} = \frac{p(1-p)^n}{p((1-p)^{n-1}} = (1-p).$$

Параметр *p* имеет смысл вероятности. Пусть при повторении опыта событие *A* имеет вероятность *p*, тогда число опытов *X* до первого появления события *A* как раз определяется выражением (14.1). Действительно, вероятность того, что в первых $n-1$ опытах событие *A* не произойдет, равна $(1-p)^{n-1}$.А вероятность появления его при *n*-ом испытании равна *p*. Отсюда получаем вероятность реализации такой серии событий равна $(1-p)^{n-1} \cdot p$.

Математическое ожидание

$$M(X) = \sum_{n=1}^{\infty} (1-p)^n p^2 = \frac{1}{p}.$$

Дисперсия

$$D(X)=\sum_{n=1}^{\infty}(1-p)^n p\cdot n^2-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

14.7 Распределение Паскаля

Распределение Паскаля дискретной случайной величины X выражается так:

$$P(X=x)=C_{x+k-1}^{x}p^k(1-p)^x,\quad (x=0,1,2,...),\tag{14.2}$$

где p и k – параметры распределения. Параметр p имеет смысл вероятности, то есть $0\le p\le 1$. Параметр k – целое положительное число.

Математическое ожидание

$$M(X)=\sum_{x=0}^{\infty}C_{x+k-1}^{x}p^k(1-p)^x\cdot x=\frac{k(1-p)}{p}.$$

Дисперсия

$$D(X)=\sum_{x=0}^{\infty}C_{x+k-1}^{x}p^k(1-p)^x\cdot x^2-(M(X))^2=\frac{k(1-p)}{p^2}.$$

Это распределение связано, как и геометрическое, с повторением опытов.

Если p – вероятность события A в одном опыте, то до появления этого события k раз потребуется всего k+x испытаний, где конкретное значение x имеет вероятность (14.2).

Это распределение обобщает геометрическое распределение. То есть если k=1, то распределение Паскаля совпадает с геометрическим. Действительно, если в (1) подставить k=1, то получим

$$P(X=x)=C_x^x p(1-p)^x,$$

что совпадает с геометрическим распределением, если положить, что

$$x=n-1$$

и учесть, что $C_x^x = 1$.

14.8 Гипергеометрическое распределение

Гипергеометрическое распределение дискретной случайной величины X выражается так:

$$P(X = x) = \frac{C_k^x C_{N-k}^{n-x}}{C_N^n}, \quad (x = 0, 1, 2, ..., \min(k, n)), \qquad (14.3)$$

где N, n, k – целые положительные величины, играющие роль параметров распределения, причем $n \le N$, $k \le N$.

Это выражение уже встречается в связи с выборкой размера n из партии деталей размером N, в которой k – число дефектных деталей. Тогда x – число дефектных деталей в выборке из n деталей, а (14.3) – вероятность этого значения.

Математическое ожидание

$$M(X) = \sum_{x=0}^{\min(n,k)} P(X = x)x = \frac{nk}{N}.$$

Дисперсия

$$D(X) = \sum_{x=0}^{\min(n,k)} P(X = x) \cdot x^2 - (M(X))^2 = \frac{nk(N-k)(N-n)}{N^2(N-1)}.$$

14.9 Контрольные вопросы

1. Что такое неслучайная величина?
2. Как распределены случайные величины в случае равномерного распределения?
3. Какой вид имеет плотность распределенеия Симпсона?
4. Как определяется вероятность значений биноминального распределения?

5. Как определяется вероятность значений распределения Пуассона?

6. Как определяется вероятность значений геометрического распределения?

7. Как выражается распределение Паскаля дискретной случайной величины X?

8. Как выражается гипергеометрическое распределение дискретной случайной величины X?