УНИВЕРСИТЕТ ИТМО

Теория вероятностей

Лекция 2. ***Схема Бернулли.*** *СЛУЧАЙНЫЕ ВЕЛИЧИНЫ*

Независимые испытания. Схема Бернулли.

С понятием «независимых событий» связано понятие «независимых испытаний (опытов)».

Несколько опытов называются *независимыми*, если их исходы представляют собой независимые события (независимые в совокупности).

Последовательность n независимых испытаний, в каждом из которых может произойти некоторое событие A (его называют *успехом*) с вероятностью $P(A) = p$ или противоположное ему событие $\overline{A}$ (его называют *неудачей*) с вероятностью $P(\overline{A}) = q = 1 - p$, называется *схемой Бернулли*.

Например, при стрельбе по мишени: событие A — попадание (успех), событие $\overline{A}$ — промах (неудача); при обследовании n изделий на предмет годности: событие A — деталь годная (успех), событие $\overline{A}$ — деталь бракованная (неудача) и т. д.

Независимые испытания. Схема Бернулли.

В каждом таком опыте ПЭС состоит только из двух элементарных событий, т. е. $\Omega = \{w_0, w_1\}$, где w_0 — неудача, w_1 — успех, при этом $A = \{w_1\}$, $\overline{A} = \{w_0\}$. Вероятности этих событий обозначают через p и q соответственно ($p + q = 1$). Множество элементарных исходов для n опытов состоит из 2^n элементов. Например, при $n = 3$, т. е. опыт повторяется 3 раза, $\Omega = \Big\{ \underbrace{(\overline{A}, \overline{A}, \overline{A})}_{w_0}; \underbrace{(A, A, \overline{A})}_{w_1}; \underbrace{(A, \overline{A}, A)}_{w_2}; \underbrace{(\overline{A}, A, A)}_{w_3}; \underbrace{(A, \overline{A}, \overline{A})}_{w_4}; \underbrace{(\overline{A}, A, \overline{A})}_{w_5}; \underbrace{(\overline{A}, \overline{A}, A)}_{w_6}; \underbrace{(A, A, A)}_{w_7} \Big\}$. Вероятность каждого элементарного события определяется однозначно. По теореме умножения вероятность события, скажем $w_6 = (\overline{A}, \overline{A}, A)$, равна $q \cdot q \cdot p = pq^2$, события w_7 — $p \cdot p \cdot p = p^3 q^0 = p^3$ и т. д.

Часто успеху сопоставляют число 1, неудаче — число 0. Элементарным событием для n опытов будет последовательность из n нулей и единиц. Тройка чисел $(0, 0, 0)$ означает, что во всех трех опытах событие A не наступило; тройка чисел $(0, 1, 0)$ означает, что событие A наступило во 2-м опыте, а в 1-м и 3-м — не наступило.

Формула Бернулли.

Простейшая задача, относящаяся к схеме Бернулли, состоит в определении вероятности того, что в n независимых испытаниях событие A наступит m раз ($0 \leqslant m \leqslant n$). Обозначается искомая вероятность так: $P_n(m)$ или $P_{n,m}$ или $P(\mu_n = m)$, где μ_n — число появления события A в серии из n опытов.

Например, при бросании игральной кости 3 раза $P_3(2)$ означает вероятность того, что в 3-х опытах событие A — выпадение цифры 4 — произойдет 2 раза. Очевидно,

$$P_3(2) = p^2q + p^2q + p^2q =$$

$$= \left[\{(A, A, \overline{A}); (A, \overline{A}, A); (\overline{A}, A, A)\}\right] = 3p^2q = 3 \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} = \frac{5}{72} = 0{,}069.$$

Теорема Если производится n независимых испытаний, в каждом из которых вероятность появления события A равна p, а вероятность его непоявления равна $q = 1 - p$, то вероятность того, что событие A произойдет m раз определяется *формулой Бернулли*

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}, \quad m = 0, 1, 2, \ldots, n. \qquad (1.32)$$

Формула Бернулли.

Доказательство: ❑ Вероятность одного сложного события, состоящего в том, что событие A в n независимых опытах появится m раз в первых m опытах и не появится $(n-m)$ раз в остальных опытах (это событие $\underbrace{A \cdot A \cdot A \cdot \ldots \cdot A}_{m \text{ раз}} \cdot \underbrace{\overline{A} \cdot \overline{A} \cdot \ldots \cdot \overline{A}}_{(n-m) \text{ раз}}$) по теореме умножения вероятностей равна $p^m q^{n-m}$. Вероятность появления события A снова m раз, но в другом порядке (например, $\overline{A} \cdot \underbrace{A \cdot A \ldots \cdot A}_{m \text{ раз}} \overline{A} \cdot \overline{A} \cdot \ldots \cdot \overline{A}$ или $A\overline{A}A\overline{A} \cdot \ldots \cdot A\overline{A}$ и т. д.) будет той же самой, т. е. $p^m q^{n-m}$.

Число таких сложных событий — в n опытах m раз встречается событие A в различном порядке — равно числу сочетаний из n по m, т. е. C_n^m. Так как все эти сложные события несовместны, то по теореме сложения вероятностей искомая вероятность равна сумме вероятностей всех возможных сложных событий, т. е.

$$P_n(m) = \underbrace{p^m q^{n-m} + \ldots + p^m q^{n-m}}_{C_n^m \text{ слагаемых}} = C_n^m p^m q^{n-m}, \quad m = 0, 1, \ldots, n. \quad \blacksquare$$

Биномиальный закон распределения вероятностей

Можно заметить, что вероятности $P_n(m)$, $m = 0, 1, \ldots, n$ являются коэффициентами при x^m в разложении $(q+px)^n$ по формуле бинома Ньютона:

$$(q+px)^n = q^n + C_n^1 q^{n-1} px + C_n^2 q^{n-2} p^2 x^2 + \ldots + C_n^m q^{n-m} p^m x^m + \ldots + p^n x^n.$$

Поэтому совокупность вероятностей $P_n(m)$ называют *биномиальным законом распределения вероятностей* (см. п. 2.7), а функцию $\varphi(x) = (q+px)^n$ — *производящей функцией* для последовательности независимых опытов.

Биномиальный закон распределения вероятностей

Определение. Число k называется наивероятнейшим числом наступления события A в схеме независимых испытаний, если вероятность того, что событие A наступит k раз не меньше вероятности остальных, возможных исходов опыта.

Теорема. Пусть в схеме n независимых испытаний событие A наступает в каждом из них с вероятностью p. Тогда для наивероятнейшего числа k наступлений события A справедливо двойное неравенство: $np - q \le k \le np + p$ $(q = 1 - p)$

Доказательство. Предположим, что k — наивероятнейшее число в схеме Бернулли при общем числе испытаний n. Тогда, согласно определению 4.4, k удовлетворяет следующей системе неравенств:

$$\begin{cases} P_n(k) \geq P_n(k-1), \\ P_n(k) \geq P_n(k+1). \end{cases}$$

Согласно формуле Бернулли (см. refu901) имеем

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k};$$

Биномиальный закон распределения вероятностей

$$P_n(k-1) = \frac{n!}{(k-1)!(n-k+1)!}p^{k-1}q^{n-k+1};$$

$$P_n(k+1) = \frac{n!}{(k+1)!(n-k-1)!}p^{k+1}q^{n-k-1};$$

Подставляя данные представления в систему, получаем

$$\begin{cases} \dfrac{n!}{k!(n-k)!}p^k q^{n-k} \geq \dfrac{n!}{(k-1)!(n-k+1)!}p^{k-1}q^{n-k+1}, \\ \dfrac{n!}{k!(n-k)!}p^k q^{n-k} \geq \dfrac{n!}{(k+1)!(n-k-1)!}p^{k+1}q^{n-k-1}. \end{cases}$$

Сократив обе части первого неравенства системы на

$$\frac{n!}{(k-1)!(n-k)!}p^{k-1}q^{n-k},$$

получим

$$\frac{p}{k} \geq \frac{q}{n-k+1}.$$

Биномиальный закон распределения вероятностей

Далее, подставляя $q = 1 - p$ и домножая обе части неравенства на

$$k(n-k+1),$$

имеем

$$(n-k+1)p \geq k(1-p) \Rightarrow np - pk + p \geq k - pk.$$

В результате, после приведения подобных получаем, что

$$k \leq np + p.$$

Проделаем аналогичные действия со вторым неравенством системы. Для этого сократим обе части второго неравенства системы на

$$\frac{n!}{k!(n-k-1)!} p^k q^{n-k-1},$$

и получим

$$\frac{q}{n-k} \geq \frac{p}{k+1}.$$

Биномиальный закон распределения вероятностей

Далее, домножая обе части неравенства на

$$(n-k)(k+1),$$

имеем

$$q(k+1)p \geq p(n-k) \Rightarrow qk + pk \geq pn - pk.$$

Прибавляя к обеим частям неравенства pk и вычитая q, получим

$$qk + pk \geq pn - q \Rightarrow (p+q)k \geq pn - q.$$

Так как $p + q = 1$, то в итоге имеем

$$k \geq pn - q.$$

Таким образом, действительно имеет место двойное неравенство

$$np - q \leq k \leq np + p.$$

Полиномиальное распределение

Если в каждом из независимых испытаний вероятности наступления события A *разные*, то вероятность того, что событие A наступит m раз в n опытах, равна коэффициенту при m-й степени многочлена $\varphi_n(z) = (q_1 + p_1 z)(q_2 + p_2 z) \cdot \ldots \cdot (q_n + p_n z)$, где $\varphi_n(z)$ — производящая функция.

Если в серии из n независимых опытов, в каждом из которых может произойти одно и только одно из k событий $A_1, A_2, \ldots, A_k$ с соответствующими вероятностями $p_1, p_2, \ldots, p_k$, то вероятность того, что в этих опытах событие A_1 появится m_1 раз, событие A_2 — m_2 раз, ..., событие A_k — m_k раз, равна

$$P_n(m_1, m_2, \ldots, m_k) = \frac{n!}{m_1! m_2! \ldots m_k!} p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k}, \tag{1.33}$$

где $m_1 + m_2 + \ldots + m_k = n$. Вероятности (1.33) называются *полиномиальным распределением*.

Теорема Пуассона

Теорема Если число испытаний неограничено увеличивается ($n \to \infty$) и вероятность p наступления события A в каждом испытании неограничено уменьшается ($p \to 0$), но так, что их произведение np является постоянной величиной ($np = a = \text{const}$), то вероятность $P_n(m)$ удовлетворяет предельному равенству

$$\lim_{n\to\infty} P_n(m) = \frac{a^m e^{-a}}{m!}.$$

Это выражение называется ассимптотической формулой Пуассона.

Теорема Пуассона

Доказательство:

Преобразуем формулу Бернулли с учетом того, что $p = \frac{a}{n}$:

$$P_n(m) = \frac{n!}{m!(n-m)!} \cdot \left(\frac{a}{n}\right)^m \cdot \left(1 - \frac{a}{n}\right)^{n-m} =$$

$$= \frac{n(n-1)(n-2)\ldots(n-(m-1))}{m!} \cdot \frac{a^m}{n^m} \cdot \left(1 - \frac{a}{n}\right)^n \cdot \left(1 - \frac{a}{n}\right)^{-m} =$$

$$= \frac{a^m}{m!} \cdot \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-(m-1)}{n} \cdot \left(1 - \frac{a}{n}\right)^n \cdot \left(1 - \frac{a}{n}\right)^{-m} =$$

$$= \frac{a^m}{m!} \cdot 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \ldots \left(1 - \frac{m-1}{n}\right) \cdot \left(1 - \frac{a}{n}\right)^n \cdot \left(1 - \frac{a}{n}\right)^{-m}.$$

Переходя к пределу при $n \to \infty$, получим $\lim\limits_{n\to\infty} P_n(m) = \frac{a^m e^{-a}}{m!}$

($\lim\limits_{n\to\infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}$ согласно второму замечательному пределу). ■

Формула Пуассона

Из предельного равенства теоремы Пуассона при больших n и малых p вытекает приближенная формула Пуассона:

$$P_n(m) \approx \frac{a^m e^{-a}}{m!}, \quad a = np, \quad m = 0, 1, 2, \ldots .$$

Формула применяется, когда вероятность успеха крайне мала, т. е. сам по себе успех является редким событием. Её обычно используют, когда

$$n \geqslant 50, \text{ а } np \leqslant 10.$$

Формула Пуассона находит применение *в теории массового обслуживания.*

Математическая модель простейшего потока событий

Формулу Пуассона можно считать математической моделью простейшего потока событий.

Потоком событий называют последовательность событий, наступающих в случайные моменты времени (например, поток посетителей в парикмахерской, поток вызовов на телефонной станции, поток отказов элементов, поток обслуженных абонентов и т. п.).

Поток событий, обладающий свойствами стационарности, ординарности и отсутствия последствия называется *простейшим* (*пуассоновским*) *потоком.*

Свойство *стационарности* означает, что вероятность появления k событий на участке времени длины τ зависит только от его длины (т. е. не зависит от начала его отсчета). Следовательно, *среднее число событий, появляющихся в единицу времени*, так называемая *интенсивность* λ потока, есть величина постоянная: $\lambda(t) = \lambda$.

Математическая модель простейшего потока событий

Свойство *ординарности* означает, что событие появляется не группами, а поодиночке. Другими словами, вероятность появления более одного события на малый участок времени Δt пренебрежительно мала по сравнению с вероятностью появления только одного события (например, поток катеров, подходящих к причалу, ординарен).

Свойство *отсутствия последствия* означает, что вероятность появления k событий на любом участке времени длины τ не зависит от того, сколько событий появилось на любом другом не пересекающимся с ним участком (говорят: «будущее» потока не зависит от «прошлого». например, поток людей, входящих в супермаркет).

Можно доказать, что вероятность появления m событий простейшего потока за время продолжительностью t определяется формулой Пуассона

$$P_t(m) = p_m = \frac{(\lambda t)^m \cdot e^{-\lambda t}}{m!}.$$

Локальная и интегральная теоремы Муавра-Лапласа

В тех случаях, когда число испытаний n велико, а вероятность p не близка к нулю ($p \neq 0$, $p \neq 1$), для вычисления биномиальных вероятностей используют теоремы Муавра–Лапласа. Приведем только их формулировки в силу сложности доказательства.

Локальная теорема Муавра-Лапласа

Теорема **Локальная теорема Муавра–Лапласа** Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, а число независимых испытаний достаточно велико, то вероятность $P_n(m)$ может быть вычислена по приближенной формуле

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \text{где} \quad x = \frac{m - np}{\sqrt{npq}}.$$

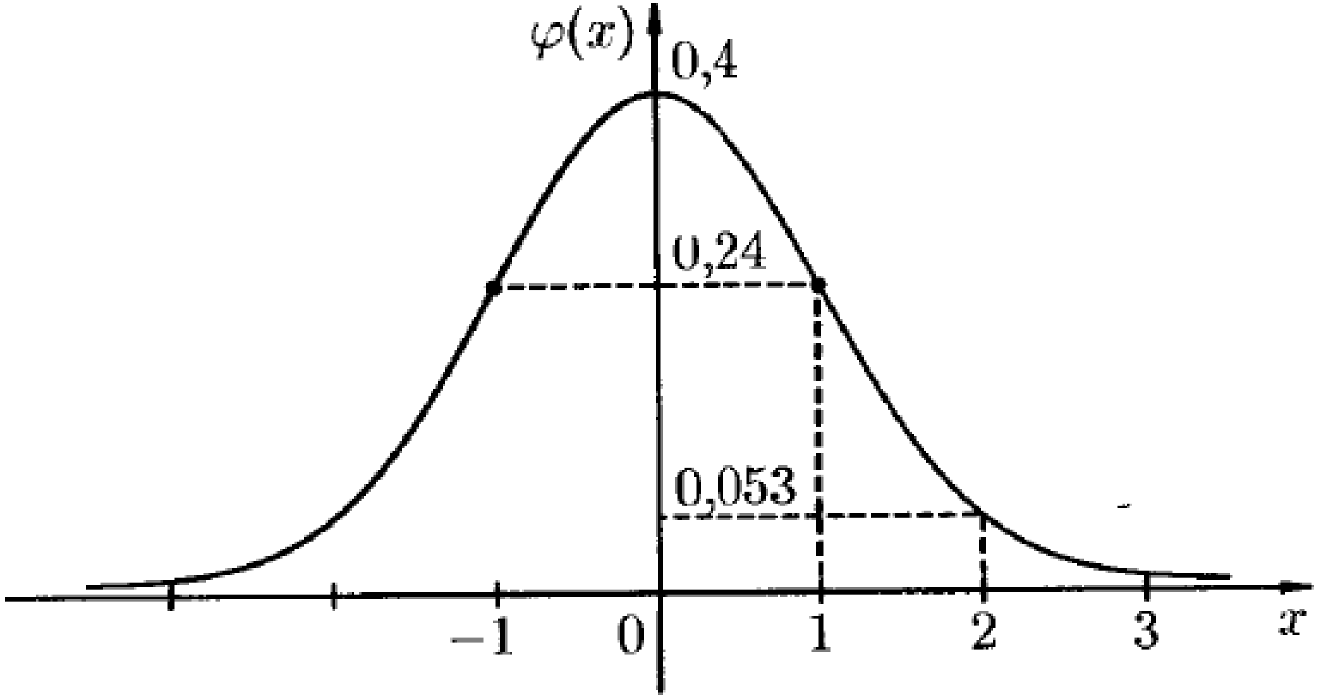
Равенство тем точнее, чем больше n.

Функция Гаусса

Выражение

$$\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} = \varphi(x)$$

называется *функцией Гаусса*, а ее график — *кривой вероятностей*



Интегральная теорема Муавра-Лапласа

В тех случаях, когда требуется вычислить вероятность того, что в n независимых испытаниях событие A появится не менее k_1 раз, но не более k_2 раз, т. е. $P_n(k_1 \leqslant m \leqslant k_2)$ или $P_n(k_1; k_2)$, используют интегральную теорему Муавра–Лапласа (является частным случаем более общей теоремы — центральной предельной теоремы).

Теорема (Интегральная теорема Муавра–Лапласа). Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k_1 \leqslant m \leqslant k_2)$ может быть найдена по приближенной формуле

$$P_n(k_1 \leqslant m \leqslant k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}}\, dx, \quad \text{где} \quad x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

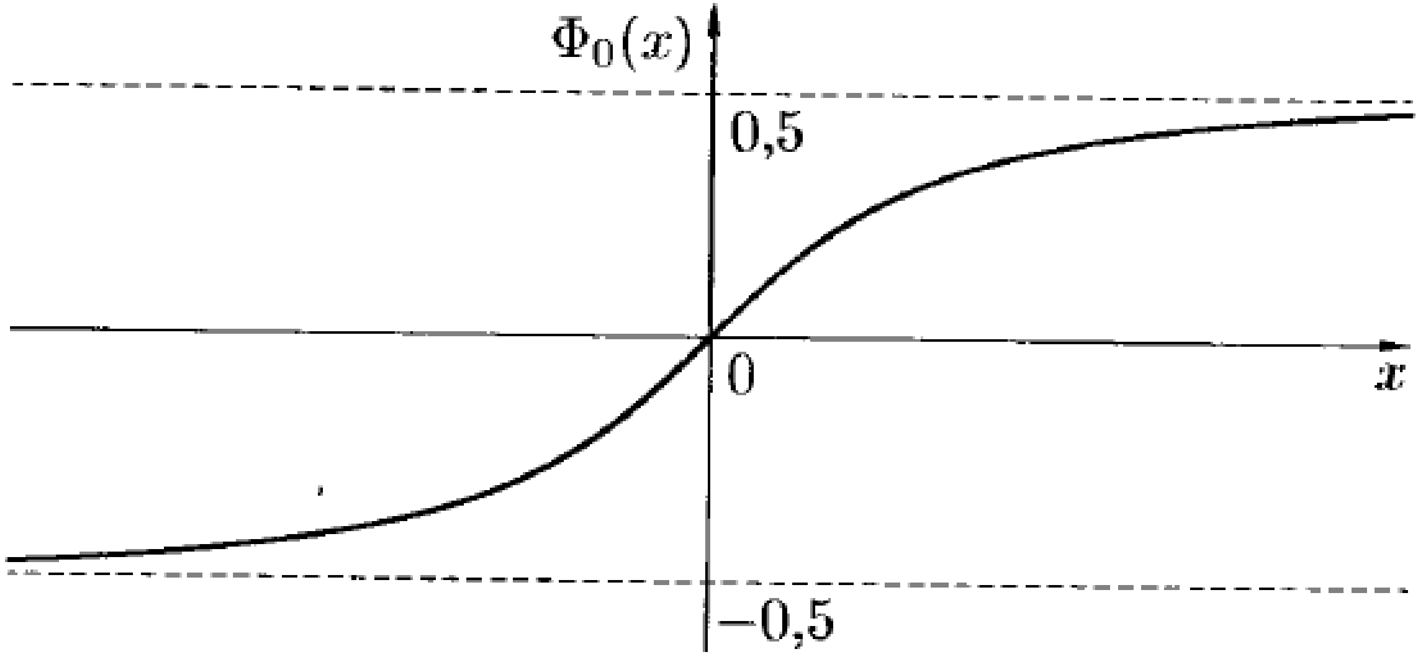
Равенство тем точнее, чем больше n.

Нормированная функция Лапласа

для упрощения вычислений вводят специальную функцию

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}}\, dt,$$

называемую *нормированной функцией Лапласа.*



Приложение 2. Значение функции $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}}\, dt$

x	Сотые доли x: 0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0112	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3079	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3553	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1.6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4700	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4762	4767

x	Десятые доли x: 0	1	2	3	4	5	6	7	8	9
2,	4773	4821	4861	4893	4918	4938	4953	4965	4974	4981
3,	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000[1]

Функция Лапласа

Наряду с нормированной функцией Лапласа используют функцию

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt,$$

называемую также *функцией Лапласа*. Для нее справедливо равенство $\Phi(-x) + \Phi(x) = 1$; она связана с функцией $\Phi_0(x)$ формулой

$$\Phi(x) = 0{,}5 + \Phi_0(x).$$

Приближенную формулу для вычисления вероятности $P_n(k_1 \leqslant m \leqslant k_2)$ можно записать в виде

$$P_n(k_1 \leqslant m \leqslant k_2) = \Phi(x_2) - \Phi(x_1) = \Phi_0(x_2) - \Phi_0(x_1),$$

где $x_1 = \dfrac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \dfrac{k_2 - np}{\sqrt{npq}}.$

Случайная величина

Под *случайной величиной* понимают величину, которая в результате опыта принимает то или иное значение, причем неизвестно заранее, какое именно.

Примерами с. в. могут служить: 1) X число очков, появляющихся при бросании игральной кости; 2) Y — число выстрелов до первого попадания в цель; 3) Z — время безотказной работы прибора и т. п. (рост человека, курс доллара, количество бракованных деталей в партии, температура воздуха, выигрыш игрока, координата точки при случайном выборе ее на $[0; 1]$, прибыль фирмы, …).

Случайная величина, принимающая конечное или счетное множество значений, называется *дискретной* (сокращенно: д. с. в.).

Если же множество возможных значений с. в. несчетно, то такая величина называется *непрерывной* (сокращенно: н. с. в.).

Строгое определение

Случайной величиной X называется числовая функция, определенная на пространстве элементарных событий Ω, которая каждому элементарному событию w ставит в соответствие число $X(w)$, т. е. $X = X(w)$, $w \in \Omega$ (или $X = f(w)$).

Пример. Опыт состоит в бросании монеты 2 раза. На ПЭС $\Omega = \{w_1, w_2, w_3, w_4\}$, где $w_1 =$ ГГ, $w_2 =$ ГР, $w_3 =$ РГ, $w_4 =$ РР, можно рассмотреть с. в. X — число появлений герба. С. в. X является функцией от элементарного события w_i: $X(w_1) = 2$, $X(w_2) = 1$, $X(w_3) = 1$, $X(w_4) = 0$; X — д. с. в. со значениями $x_1 = 0$, $x_2 = 1$, $x_3 = 2$.

Отметим, что если множество Ω конечно или счетно, то случайной величиной является любая функция, определенная на Ω. В общем случае функция $X(w)$ должна быть такова, чтобы для любых $x \in \mathbb{R}$ событие $A = \{w : X(w) < x\}$ принадлежало σ-алгебре множеств S и, значит, для любого такого события была определена вероятность $P(A) = P(X < x)$.

Закон распределения случайной величины

Для полного описания с. в. недостаточно лишь знания ее возможных значений; необходимо еще знать вероятности этих значений.

Любое правило (таблица, функция, график), позволяющее находить вероятности произвольных событий $A \subseteq S$ (S — σ-алгебра событий пространства Ω), в частности, указывающее вероятности отдельных значений случайной величины или множества этих значений, называется *законом распределения случайной величины* (или просто: *распределением*). Про с. в. говорят, что «она подчиняется данному закону распределения»:

Закон распределения случайной величины

Пусть X — д. с. в., которая принимает значения $x_1, x_2, x_3, \ldots, x_n, \ldots$ (множество этих значений конечно или счетно) с некоторой вероятностью p_i, где $i = 1.2.3.\ldots.n.\ldots$. *Закон распределения* д. с. в. удобно задавать с помощью формулы $p_i = P\{X = x_i\}$, $i = 1, 2, 3, \ldots, n, \ldots$, определяющей вероятность того, что в результате опыта с. в. X примет значение x_i. Для д. с. в. X закон распределения может быть задан в виде *таблицы распределения*:

X	x_1	x_2	$\ldots$	x_n	$\ldots$
P	p_1	p_2	$\ldots$	p_n	$\ldots$

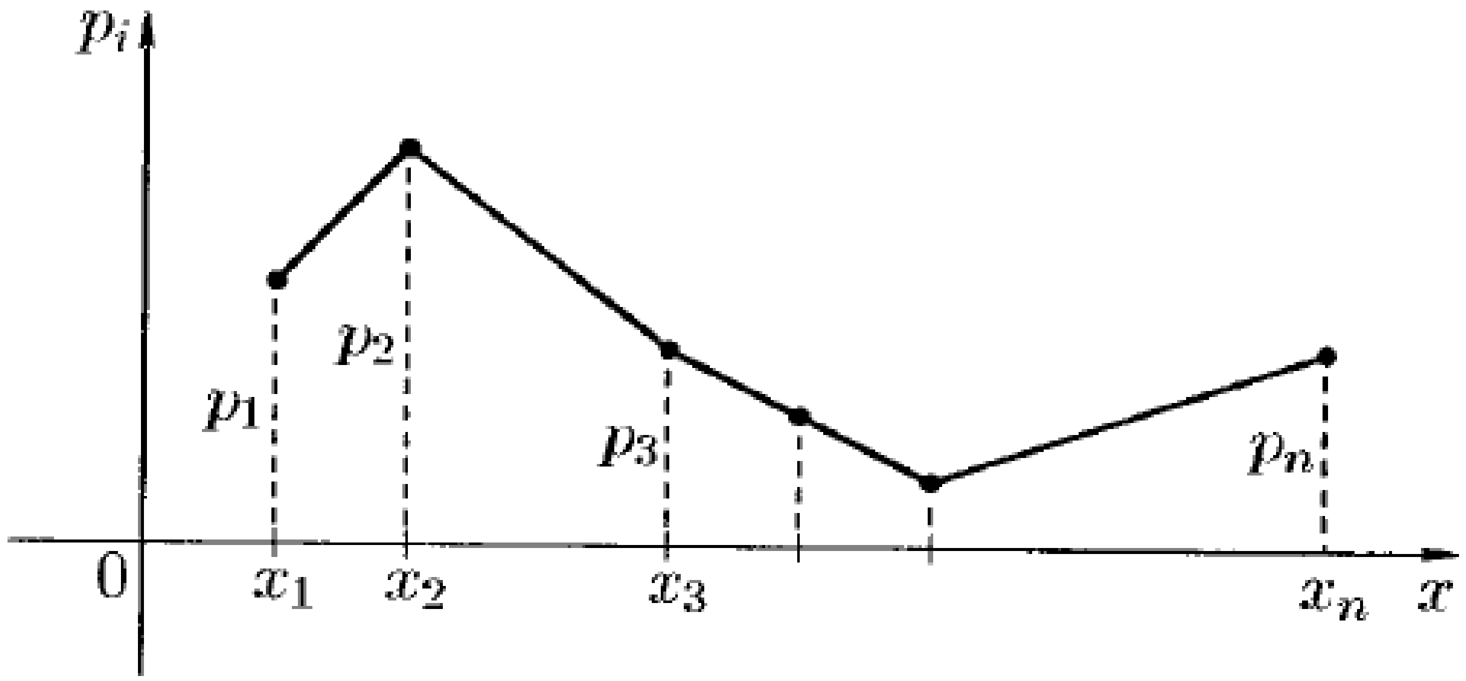
,

где первая строка содержит все возможные значения (обычно в порядке возрастания) с. в., а вторая — их вероятности. Такую таблицу называют *рядом распределения*.

Так как события $\{X = x_1\}$, $\{X = x_2\}$ $\ldots$ несовместны и образуют полную группу, то сумма их вероятностей равна единице т. е. $\sum\limits_i p_i = 1$.

Многоугольник (полигон) распределения

Закон распределения д. с. в. можно задать *графически*, если на оси абсцисс отложить возможные значения с. в., а на оси ординат — вероятности этих значений. Ломаную, соединяющую последовательно точки (x_1, p_1), (x_2, p_2), ... называют *многоугольником* (или *полигоном*) *распределения* (см. рис. 17).



Дискретная случайная величина

Случайная величина X *дискретна*, если существует конечное или счетное множество чисел x_1, x_2, ... таких, что $P\{X = x_i\} = p_i > 0$ $(i = 1, 2, \ldots)$ и $p_1 + p_2 + p_3 + \ldots = 1$.

Определим математические операции над дискретными с. в.

Суммой (*разностью, произведением*) д. с. в. X, принимающей значения x_i с вероятностями $p_i = P\{X = x_i\}$, $i = 1, 2, \ldots, n$ и д. с. в. Y, принимающей значения y_j с вероятностями $p_j = P\{Y = y_j\}$, $j = 1, 2, \ldots, m$, называется д. с. в. $Z = X + Y$ $(Z = X - Y,\ Z = X \cdot Y)$, принимающая значения $z_{ij} = x_i + y_j$ $(z_{ij} = x_i - y_j,\ z_{ij} = x_i \cdot y_j)$ с вероятностями $p_{ij} = P\{X = x_i, Y = y_j\}$ для всех указанных значений i и j. В случае совпадения некоторых сумм $x_i + y_j$ (разностей $x_i - y_j$, произведений $x_i y_j$) соответствующие вероятности складываются.

Дискретная случайная величина

Произведение д. с. в. *на число* c называется д. с. в. cX, принимающая значения cx_i с вероятностями $p_i = P\{X = x_i\}$.

Две д. с. в. X и Y называются *независимыми*, если события $\{X = x_i\} = A_i$ и $\{Y = y_j\} = B_j$ независимы для любых $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$, т. е.

$$P\{X = x_i; Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}.$$

В противном случае с. в. называются *зависимыми*. Несколько с. в. называются взаимно независимыми, если закон распределения любой из них не зависит от того, какие возможные значения приняли остальные величины.

Функция распределения случайной величины

Очевидно, ряд распределения с. в. может быть построен только для д. с. в.; для н. с. в. нельзя даже перечислить все ее возможные значения.

Для характеристики поведения н. с. в. целесообразно использовать вероятность события $\{X < x\}$ (а не $\{X = x\}$), где x — некоторое действительное число. С точки зрения практики нас мало интересует событие, состоящее, например, в том, что лампочка проработает ровно 900 часов, т. е. $X = 900$. Более важным является событие вида $\{X < 900\}$ (или $\{X > 900\}$). Такое событие имеет ненулевую вероятность; при изменении x вероятность события $\{X < x\}$ в общем случае будет меняться. Следовательно, вероятность $P\{X < x\}$ является функцией от x.

Универсальным способом задания закона распределения вероятностей, пригодным как для дискретных, так и для непрерывных случайных величин, является ее функция распределения, обозначаемая $F_X(x)$ (или просто $F(x)$, без индекса, если ясно, о какой с. в. идет речь).

Функция распределения случайной величины

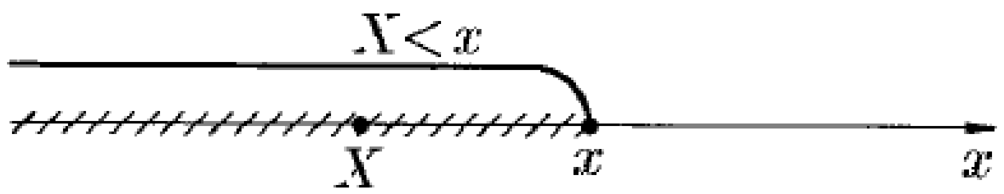
Функцией распределения с. в. X называется функция $F(x)$, которая для любого числа $x \in R$ равна вероятности события $\{X < x\}$.

Таким образом, по определению

$$F(x) = P\{X < x\} \quad \text{т. е.} \quad F(x) = P\{w : X(w) < x\}.$$

Функцию $F(x)$ называют также *интегральной функцией распределения*.

Геометрически равенство можно истолковать так: $F(x)$ есть вероятность того, что с. в. X примет значение, которое изображается на числовой оси точкой, лежащей *левее точки* x, т. е. случайная точка X попадет в интервал $(-\infty, x)$.



Свойства функции распределения

Функция распределения обладает следующими свойствами:

1. $F(x)$ ограничена, т. е.

$$0 \leqslant F(x) \leqslant 1.$$

2. $F(x)$ — неубывающая функция на R, т. е. если $x_2 > x_1$, то

$$F(x_2) \geqslant F(x_1).$$

3. $F(x)$ обращает в ноль на минус бесконечности и равна единице в плюс бесконечности, т. е.

$$F(-\infty) = 0, \qquad F(+\infty) = 1.$$

4. Вероятность попадания с. в. X в промежуток $[a, b)$ равна приращению ее функции распределения на этом промежутке, т. е.

$$P\{a \leqslant X < b\} = F(b) - F(a).$$

5. $F(x)$ непрерывна слева, т. е.

$$\lim_{x \to x_0 - 0} F(x) = F(x_0).$$

Непрерывная случайная величина

С помощью функции распределения можно вычислить вероятность события $\{X \geqslant x\}$:

$$P\{X \geqslant x\} = 1 - F(x).$$

Можно дать *более точное определение н. с. в.*

Случайную величину X называют *непрерывной*, если ее функция распределения непрерывна в любой точке и дифференцируема всюду, кроме, может быть, отдельных точек.

для н. с. в. справедливы равенства

$$P\{a \leqslant x < b\} = P\{a < x < b\} = P\{a \leqslant x \leqslant b\} = P\{X \in (a, b]\}.$$

Функция распределения дискретной случайной величины

Функция распределения д. с. в. имеет вид

$$F(x) = \sum_{x_i < x} p_i.$$

Здесь суммирование ведется по всем i, для которых $x_i < x$.

Плотность распределения вероятностей

Плотностью распределения вероятностей (*плотностью распределения, плотностью вероятностей* или просто *плотностью*) непрерывной случайной величины X называется производная ее функции распределения.

Обозначается плотность распределения н. с. в. X через $f_X(x)$ (или $p_X(x)$) или просто $f(x)$ (или $p(x)$), если ясно о какой с. в. идет речь.

Таким образом, по определению

$$f(x) = F'(x).$$

Функцию $f(x)$ называют также *дифференциальной функцией распределения*; она является одной из форм закона распределения случайной величины, существует только для непрерывных случайных величин.

Плотность распределения вероятностей

Установим вероятностный смысл плотности распределения. Из определения производной следует

$$f(x) = \lim_{\Delta x \to 0} \frac{\Delta F(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Но согласно формуле $P\{a \leqslant X < b\} = F(b) - F(a)$

$$F(x + \Delta x) - F(x) = P\{x \leqslant X < x + \Delta x\}.$$

Отношение $\dfrac{P\{x \leqslant X < x + \Delta x\}}{\Delta x}$ представляет собой среднюю вероятность, которая приходится на единицу длины участка $[x, x + \Delta x)$, т. е. среднюю плотность распределения вероятности. Тогда

$$f(x) = \lim_{\Delta x \to 0} \frac{P\{x \leqslant X < x + \Delta x\}}{\Delta x},$$

т. е. плотность распределения есть предел отношения вероятности попадания с. в. в промежуток $[x; x + \Delta x)$ к длине Δx этого промежутка, когда Δx стремится к нулю.

Плотность распределения вероятностей

Из равенства следует, что

$$P\{x \leqslant X < x + \Delta x\} \approx f(x)\Delta x.$$

То есть *плотность вероятности определяется как функция* $f(x)$, *удовлетворяющая условию* $P\{x \leqslant X < x + dx\} \approx f(x)\,dx$; выражение $f(x)\,dx$ называется *элементом вероятности*.

Отметим, что плотность $f(x)$ аналогична таким понятиям, как плотность распределения масс на оси абсцисс или плотность тока в теории электричества.

Свойства плотности распределения

Плотность распределения обладает следующими свойствами:

1. $f(x)$ неотрицательная, т. е.

$$f(x) \geqslant 0.$$

2. Вероятность попадания н. с. в. в промежуток $[a; b]$ равна определенному интегралу от ее плотности в пределах от a до b, т. е.

$$P\{a \leqslant X \leqslant b\} = \int_a^b f(x)\,dx.$$

3. Функция распределения н. с. в. может быть выражена через ее плотность вероятности по формуле

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

4. Условие нормировки: несобственный интеграл от плотности вероятности н. с. в. в бесконечных пределах равен единице, т. е.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Числовые характеристики случайных величин

Закон распределения полностью характеризует случайную величину. Однако при решении многих практических задач достаточно знать лишь некоторые *числовые параметры, характеризующие отдельные существенные свойства (черты) закона распределения с. в.* Такие числа принято называть *числовыми характеристиками с. в.*

Важнейшими среди них являются характеристики положения: математическое ожидание (центр распределения с. в.), мода, медиана; характеристики рассеяния: дисперсия (отклонение значений с. в. от ее центра), среднее квадратическое отклонение.

Математическое ожидание дискретной случайной величины

Математическим ожиданием (или *средним* значением) д.с.в. X, имеющей закон распределения $p_i = P\{X = x_i\}$, $i = 1, 2, 3, \ldots, n$, называется число, равное сумме произведений всех ее значений на соответствующие им вероятности.

Математическое ожидание (сокращенно: м.о.) обозначается через MX (или: $M[X]$, $M(X)$, EX, m_X, a_X).

Таким образом, по определению

$$MX = \sum_{i=1}^{n} x_i \cdot p_i.$$

Если число возможных значений с. в. X бесконечно (счетно), то

$$MX = \sum_{i=1}^{\infty} x_i \cdot p_i,$$

причем ряд в правой части предполагается сходящимся (в противном случае с. в. X не имеет м.о.).

Вероятностный смысл математического ожидания

Вероятностный смысл математического ожидания состоит в том, что оно является средним значением с. в. Действительно, т. к. $\sum_{i=1}^{n} p_i = 1$, то

$$MX = \sum_{i=1}^{n} x_i p_i = \frac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i} = x_{\text{среднее}}.$$

Математическое ожидание дискретной случайной величины

Значение математического ожидания удовлетворяет следующему неравенству:

$$\boxed{\min\{x_i\} \le M(X) \le \max\{x_i\}}.$$

Для вычисления математического ожидания достаточно знать закон распределения случайной величины. •

Математическое ожидание непрерывной случайной величины

Математическим ожиданием н. с. в. X с плотностью вероятности $f(x)$, называется число

$$MX = \int_{-\infty}^{\infty} x \cdot f(x)\, dx.$$

Интеграл в правой части равенства предполагается абсолютно сходящимся, т. е.

$$\int_{-\infty}^{\infty} |x| \cdot f(x)\, dx < \infty$$

(в противном случае н. с. в. X не имеет м. о.).

Свойства математического ожидания

1. Математическое ожидание постоянной равно самой этой постоянной, т. е.

$$Mc = c.$$

2. Постоянный множитель выносится за знак м. о., т. е.

$$M(cX) = cMX.$$

3. М. о. суммы с. в. равно сумме их м. о., т. е.

$$M(X + Y) = MX + MY.$$

4. М. о. отклонения с. в. от ее м. о. равно нулю, т. е.

$$M(X - MX) = 0.$$

5. М. о. произведения независимых с. в. равно произведению их м. о., т. е. если X и Y независимы, то

$$M(X \cdot Y) = MX \cdot MY.$$

Дисперсия случайной величины

Дисперсией (*рассеянием*) с. в. X называется математическое ожидание квадрата ее отклонения от своего математического ожидания.

Обозначается дисперсия через DX (или $D[X]$, D_X, $D(X)$). Таким образом, по определению

$$DX = M(X - MX)^2,$$

или $DX = M\dot{X}^2$, или $DX = M(X - m_X)^2$. Дисперсия характеризует разброс значений с. в. X относительно ее м. о. Из определения дисперсии следуют формулы для ее вычисления:

$$DX = \sum_i (x_i - MX)^2 \cdot p_i \text{ — для д. с. в. } X,$$

$$DX = \int_{-\infty}^{+\infty} (x - MX)^2 \cdot f(x)\, dx \text{ — для н. с. в. } X.$$

Дисперсия случайной величины

На практике дисперсию с. в. удобно находить по формуле

$$DX = MX^2 - (MX)^2.$$

$$D(X) = M(X^2) - (M(X))^2 = \sum_{i=1}^{n} x_i^2 p_i - (M(X))^2 =$$

$$= x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + \ldots + x_n^2 \cdot p_n - (M(X))^2.$$

Доказательство:

$$DX = M(X^2 - 2X \cdot MX + (MX)^2) =$$

$$= MX^2 - M(2X \cdot MX) + M(MX)^2 = MX^2 - 2MX \cdot MX + (MX)^2 =$$

$$= MX^2 - (MX)^2.$$

Свойства дисперсии

1. Дисперсия постоянной равна нулю, т. е.

$$Dc = 0.$$

2. Постоянный множитель можно выносить за знак дисперсии, возведя его в квадрат, т. е.

$$DcX = c^2 DX.$$

3. Дисперсия суммы независимых с. в. равна сумме их дисперсий, т. е. если X и Y независимы, то

$$D(X+Y) = DX + DY.$$

4. Дисперсия с. в. не изменится, если к этой с. в. прибавить постоянную, т. е.

$$D(X+c) = DX.$$

5. Если с. в. X и Y независимы, то

$$D(XY) = MX^2 \cdot MY^2 - (MX)^2 \cdot (MY)^2.$$

Свойства дисперсии

Пояснение. Случайные величины X и Y называются *независимыми* тогда и только тогда, когда случайные события $\{X = x_i\}$ и $\{Y = y_j\}$ при любых i и j являются независимыми. Из данного определения следует, что при любых i и j вероятность события $\{X = x_i\}$ не зависит от того, произошло событие $\{Y = y_j\}$ или нет. Другими словами, независимые случайные величины X и Y не могут влиять друг на друга, взаимовлияние отсутствует. ●

Среднее квадратическое отклонение

Дисперсия DX имеет размерность квадрата с. в. X, что в сравнительных целях неудобно. Когда желательно, чтобы оценка разброса (рассеяния) имела размерность с. в., используют еще одну числовую характеристику — среднее квадратическое отклонение (сокращенно: с. к. о.).

Средним квадратическим отклонением или *стандартным отклонением* с. в. X называется квадратный корень из ее дисперсии, обозначают через σ_X (или σX, $\sigma[X]$, σ). Таким образом, по определению

$$\sigma_X = \sqrt{DX}.$$

Из свойств дисперсии вытекают соответствующие свойства с. к. о.: $\sigma c = 0$, $\sigma_{cX} = |c|\sigma_X$, $\sigma(c + X) = \sigma_X$.

Стандартная случайная величина

Для изучения свойств случайного явления, независящих от выбора масштаба измерения и положения центра группирования, исходную случайную величину X приводят к некоторому стандартному виду: ее центрируют, т. е. записывают разность $X - MX$ (геометрически означает, что начало координат переносится в точку с абсциссой, равной м. о.), затем делят на с. к. о. σ_X.

Случайную величину $Z = \dfrac{X - MX}{\sigma_X}$ называют *стандартной случайной величиной*. Ее м. о. равно 0, а дисперсия равна 1. Действительно,

$$MZ = M\left(\frac{X - MX}{\sigma_X}\right) = \frac{1}{\sigma_X} M(X - MX) = 0,$$

$$DZ = \frac{1}{\sigma_X^2} D(X - MX) = \frac{DX}{\sigma_X^2} = \frac{DX}{DX} = 1.$$

То есть Z — *центрированная* ($MZ = 0$) и *нормированная* ($DZ = 1$) случайная величина.

Моменты случайной величины

Математическое ожидание и дисперсия являются частными случаями следующих более общих понятий — моментов с. в.

Начальным моментом порядка k с. в. X называется м. о. k-й степени этой величины, обозначается через α_k.

Таким образом, по определению

$$\alpha_k = M(X^k).$$

Для д. с. в. начальный момент выражается суммой:

$$\alpha_k = \sum_i x_i^k \cdot p_i,$$

а для н. с. в. — интегралом: $$\alpha_k = \int_{-\infty}^{\infty} x^k \cdot f(x)\,dx.$$

В частности, $\alpha_1 = MX$, т. е. начальный момент 1-го порядка есть м. о.

Центральный момент

Центральным моментом порядка k с. в. X называется м. о. величины $(X - MX)^k$, обозначается через μ_k.

Таким образом, по определению

$$\mu_k = M(X - MX)^k.$$

В частности, $\mu_2 = DX$, т. е. центральный момент 2-го порядка есть дисперсия; $\mu_1 = M(X - MX) = 0$ (см. свойство 4 м. о.).

Для д. с. в.:

$$\mu_k = \sum_i (x_i - MX)^k \cdot p_i,$$

а для н. с. в.:

$$\mu_k = \int_{-\infty}^{\infty} (x - MX)^k \cdot f(x)\, dx.$$

Производящая функция

Нахождение важнейших числовых характеристик д. с. в. с целыми неотрицательными значениями удобно производить с помощью производящих функций.

Пусть д. с. в. X принимает значения $0, 1, 2, \ldots, k, \ldots$ с вероятностями $p_0, p_1, p_2, \ldots, p_k = P\{X = k\}, \ldots$.

Производящей функцией для д. с. в. X называется функция вида

$$\varphi(z) = \sum_{k=0}^{\infty} p_k \cdot z^k = p_0 + p_1 z + p_2 z^2 + \ldots,$$

где z — произвольный параметр, $0 < z \leqslant 1$.

Отметим, что коэффициентами степенного ряда являются вероятности закона распределения д. с. в. X.

Производящая функция

Дифференцируя по z производящую функцию, получим

$$\varphi'(z) = \sum_{k=0}^{\infty} k \cdot p_k \cdot z^{k-1}.$$

Тогда

$$\varphi'(1) = \sum_{k=0}^{\infty} k \cdot p_k = MX = \alpha_1,$$

т. е.

$$\alpha_1 = MX = \varphi'(1).$$

Производящая функция

Взяв вторую производную функции $\varphi(z)$ и положив в ней $z = 1$, получим:

$$\varphi''(z) = \sum_{k=0}^{\infty} k(k-1) \cdot p_k \cdot z^{k-2}, \quad \varphi''(1) = \sum_{k=0}^{\infty} k^2 \cdot p_k - \sum_{k=0}^{\infty} k \cdot p_k = \alpha_2 - \alpha_1,$$

где α_2 и α_1 — начальные моменты соответственно 2-го и 1-го порядков ($\alpha_2 = MX^2$, $\alpha_1 = MX$). Тогда $DX = MX^2 - (MX)^2 = \alpha_2 - \alpha_1^2 = (\alpha_2 - \alpha_1) + \alpha_1 - \alpha_1^2 = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2$, т. е.

$$DX = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2.$$