УНИВЕРСИТЕТ ИТМО

Теория вероятностей и математическая статистика

Лекция 4. ***Системы случайных величин***

Системы случайных величин

При изучении случайных явлений часто приходится иметь дело с двумя, тремя и большим числом случайных величин. Совместное рассмотрение нескольких случайных величин приводит к системам случайных величин. Так, точка попадания снаряда характеризуется системой (X, Y) двух случайных величин: абсциссой X и ординатой Y;

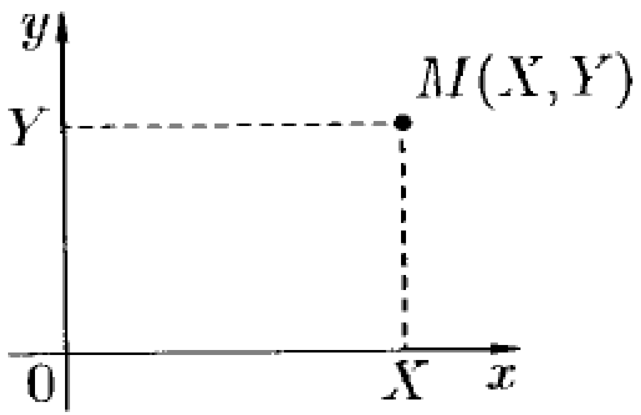
Упорядоченный набор $(X_1, X_2, \ldots, X_n)$ случайных величин X_i $(i = \overline{1,n})$, заданных на одном и том же ПЭС Ω, называется *n-мерной случайной величиной* или *системой n случайных величин*.

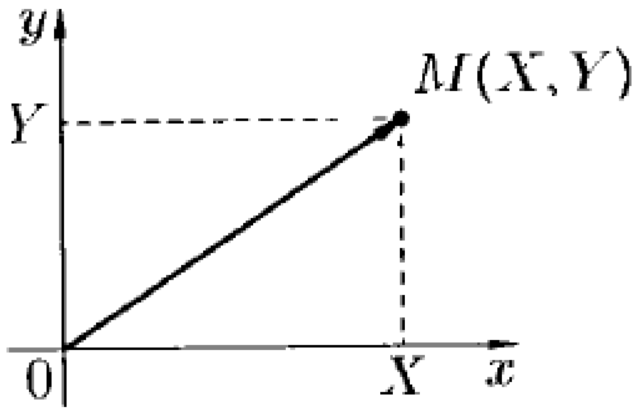
Одномерные с. в. $X_1, X_2, \ldots, X_n$ называются *компонентами* или *составляющими* n-мерной с. в. $(X_1, X_2, \ldots, X_n)$. Их удобно рассматривать как координаты случайной точки или случайного вектора $\bar{X} = (X_1. X_2, \ldots, X_n)$ в пространстве n измерений.

Системы случайных величин

Упорядоченная пара (X, Y) двух случайных величин X и Y называется *двумерной случайной величиной* или *системой двух одномерных случайных величин* X и Y.

Систему (X, Y) можно изобразить *случайной точкой* $M(X, Y)$ или *случайным вектором* OM





Система (X, Y) есть функция элементарного события: $(X, Y) = \varphi(w)$. Каждому элементарному событию w ставится в соответствие два действительных числа x и y (или x_1 и x_2) — значения X и Y (или X_1 и X_2) в данном опыте. В этом случае вектор $\bar{x} = (x_1, x_2)$ называется *реализацией* случайного вектора $\bar{X} = (X_1, X_2)$.

Системы случайных величин

Системы случайных величин могут быть *дискретными*, *непрерывными* и *смешанными* в зависимости от типа случайных величин, образующих систему. В первом случае компоненты этих случайных систем дискретны, во втором — непрерывны. в третьем — разных типов.

Полной характеристикой системы (X, Y) является ее *закон распределения вероятностей*, указывающий область возможных значений системы случайных величин и вероятности этих значений. Как и для отдельных случайных величин закон распределения системы может иметь разные формы (таблица, функция распределения. плотность, ...).

Так, закон распределения дискретной двумерной с. в. $(X. Y)$ можно задать формулой

$$p_{ij} = P\{X = x_i, Y = y_j\}, \qquad i = \overline{1, n}, \quad j = \overline{1, m}$$

Системы случайных величин

или в форме таблицы с двойным входом:

$X \setminus Y$	y_1	y_2	y_3	$\dots$	y_m
x_1	p_{11}	p_{12}	p_{13}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	p_{23}	$\dots$	p_{2m}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	p_{n2}	p_{n3}	$\dots$	p_{nm}

Причем, сумма всех вероятностей p_{ij}, как сумма вероятностей полной группы несовместных событий $\{X = x_i, Y = y_j\}$, равна единице:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1.$$

Системы случайных величин

Зная закон распределения двумерной дискретной случайной величины, *можно найти законы распределения каждой из компонент* (обратное, вообще говоря, неверно). Так, $p_{x_1} = P\{X = x_1\} = p_{11} + p_{12} + \ldots + p_{1m}$, что следует из теоремы сложения несовместных событий $\{X = x_1, Y = y_1\}$, $\{X = x_1, Y = y_2\}$, $\ldots$, $\{X = x_1, Y = y_m\}$. Аналогично можно найти

$$p_{x_i} = P\{X = x_i\} = \sum_{j=1}^{m} p_{ij}, \qquad p_{y_j} = P\{Y = y_j\} = \sum_{i=1}^{n} p_{ij}.$$

Функция распределения двумерной случайной величины

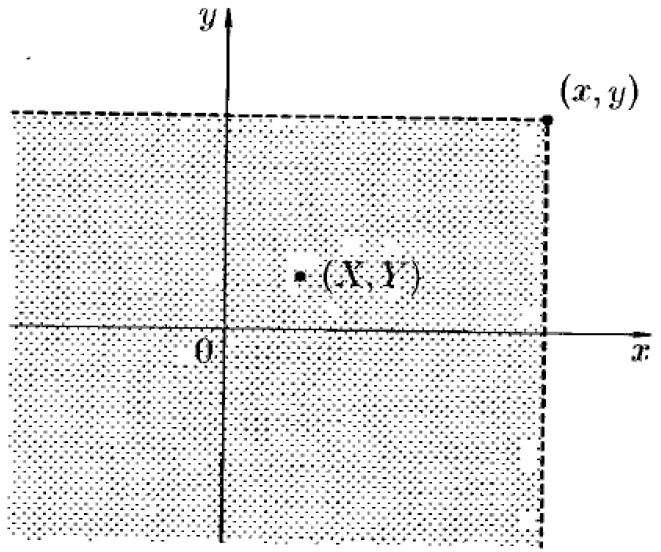
Функцией распределения двумерной случайной величины (X, Y) называется функция $F(x, y)$, которая для любых действительных чисел x и y равна вероятности совместного выполнения двух событий $\{X < x\}$ и $\{Y < y\}$.

Таким образом, по определению

$$F(x, y) = P\{X < x, Y < y\}$$

(событие $\{X < x, Y < y\}$ означает произведение событий $\{X < x\}$ и $\{Y < y\}$).

Геометрически функция $F(x, y)$ интерпретируется как вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной в точке (x, y), лежащий левее и ниже ее



Свойства функции распределения двумерной случайной величины

1. Функция распределения $F(x, y)$ ограничена, т. е.

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

2. $F(x, y)$ не убывает по каждому из своих аргументов при фиксированном другом, т. е.

$$F(x_2, y) \geqslant F(x_1, y) \quad \text{при} \quad x_2 > x_1;$$
$$F(x, y_2) \geqslant F(x, y_1) \quad \text{при} \quad y_2 > y_1.$$

3. Если хотя бы один из аргументов обращается в $-\infty$, то функция распределения $F(x, y)$ равна нулю, т. е.

$$F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0.$$

Свойства функции распределения двумерной случайной величины

4. Если оба аргумента обращаются в $+\infty$, то $F(x, y)$ равна 1, т. е.

$$F(+\infty, +\infty) = 1.$$

5. Если один из аргументов обращается в $+\infty$, то функция распределения системы случайных величин становится функцией распределения с. в., соответствующей другому элементу, т. е.

$$F(x, +\infty) = F_1(x) = F_X(x), \quad F(+\infty, y) = F_2(y) = F_Y(y). \tag{3.4}$$

6. $F(x, y)$ непрерывна слева по каждому из своих аргументов, т. е.

$$\lim_{x \to x_0 - 0} F(x, y) = F(x_0, y), \quad \lim_{y \to y_0 - 0} F(x, y) = F(x, y_0).$$

Плотность распределения вероятностей двумерной случайной величины

Двумерная случайная величина называется *непрерывной*, если ее функция распределения $F(x, y)$ есть непрерывная функция, дифференцируемая по каждому из аргументов, у которой существует вторая смешанная производная $F''_{xy}(x, y)$.

Плотностью распределения вероятностей (или *совместной плотностью*) непрерывной двумерной случайной величины (X, Y) называется вторая смешанная производная ее функции распределения.

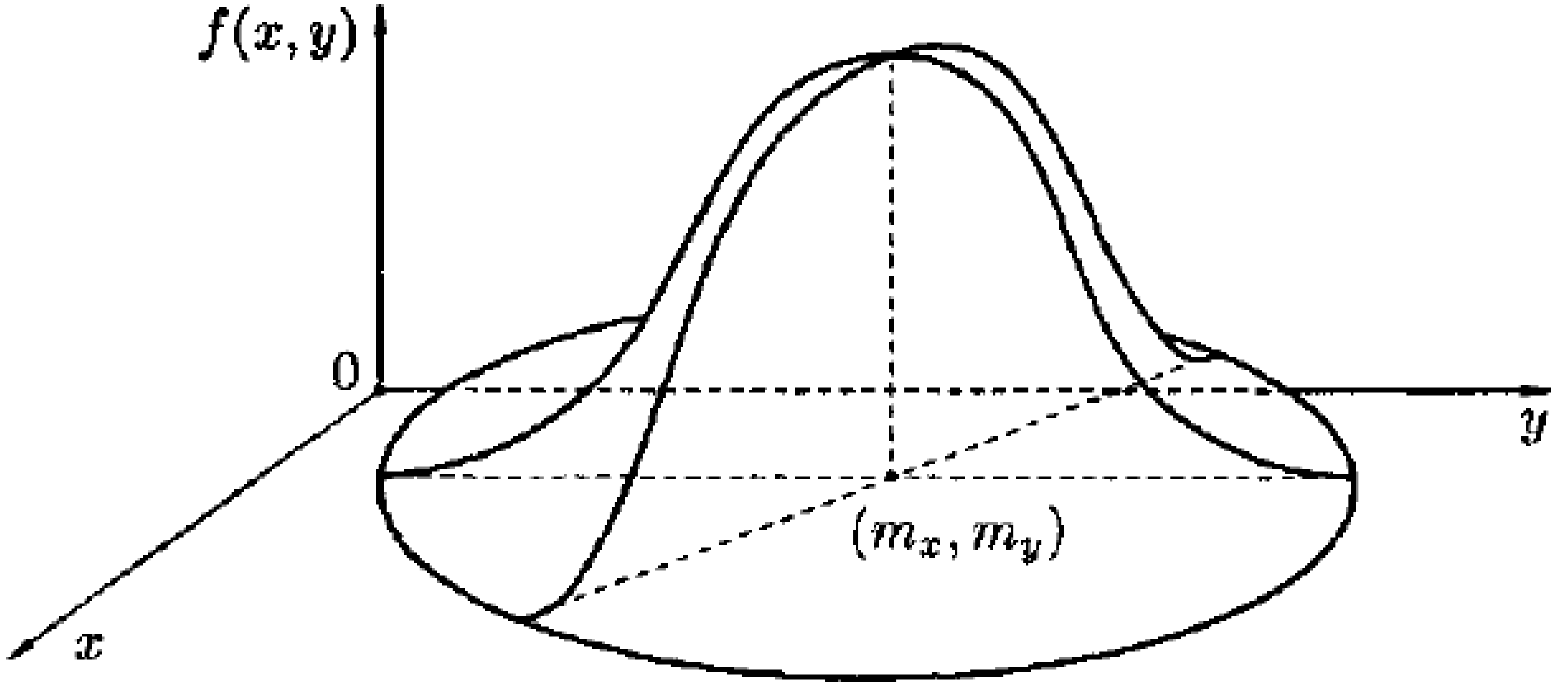
Обозначается совместная плотность системы двух непрерывных случайных величин (X, Y) через $f(x, y)$ (или $p(x, y)$).

Таким образом, по определению

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x\,\partial y} = F''_{xy}(x, y).$$

Плотность распределения вероятностей двумерной случайной величины

Геометрически плотность распределения вероятностей $f(x, y)$ системы двух случайных величин (X, Y) представляет собой некоторую поверхность, называемую *поверхностью распределения*



Свойства плотности распределения вероятностей двумерной случайной величины

1. Плотность распределения двумерной случайной величины неотрицательна, т. е.

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

2. Вероятность попадания случайной точки (X, Y) в область D равна двойному интегралу от плотности по области D, т. е.

$$P\{(X, Y) \in D\} = \iint\limits_D f(x, y)\, dx dy. \tag{3.8}$$

3. Функция распределения двумерной случайной величины может быть выражена через ее плотность распределения по формуле:

$$F(x, y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f_{X,Y}(u, v)\, du dv. \tag{3.9}$$

Свойства плотности распределения вероятностей двумерной случайной величины

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

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

5. Плотности распределения одномерных составляющих X и Y могут быть найдены по формулам:

$$\int_{-\infty}^{+\infty} f(x,y)\,dy = f_1(x) = f_x(x), \quad \int_{-\infty}^{+\infty} f(x,y)\,dx = f_2(y) = f_y(y). \qquad (3.10)$$

Числовые характеристики двумерной случайной величины

Математическим ожиданием двумерной с. в. (X, Y) называется совокупность двух м. о. MX и MY, определяемых равенствами:

$$MX = m_x = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i p_{ij}, \quad MY = m_y = \sum_{i=1}^{n} \sum_{j=1}^{m} y_j p_{ij},$$

если (X, Y) — дискретная система с. в. (здесь $p_{ij} = P\{X = x_i, Y = y_j\}$)

и

$$MX = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f(x, y)\, dx dy, \quad MY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, f(x, y)\, dx dy,$$

если (X, Y) — непрерывная система с. в. (здесь $f(x, y)$ — плотность распределения системы).

Числовые характеристики двумерной случайной величины

Дисперсией системы с. в. (X, Y) называется совокупность двух дисперсий DX и DY, определяемых равенствами:

$$DX = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - m_x)^2 p_{ij}, \quad DY = \sum_{i=1}^{n}\sum_{j=1}^{m}(y_j - m_y)^2 p_{ij},$$

если (X, Y) — дискретная система с. в. и

$$DX = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - m_x)^2 f(x,y)\,dxdy, \quad DY = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y - m_y)^2 f(x,y)\,dxdy,$$

если (X, Y) — непрерывная система с. в.

Ковариация

Корреляционным моментом (или *ковариацией*) двух случайных величин X и Y называется м. о. произведения отклонений этих с. в. от их м. о. и обозначается через K_{XY} или $\operatorname{cov}(X, Y)$.

Таким образом, по определению

$$K_{XY} = \operatorname{cov}(X, Y) = M[(X - m_x)(Y - m_y)]$$

При этом: если (X, Y) — дискретная двумерная с. в., то ковариация вычисляется по формуле

$$K_{XY} = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i - m_x)(y_j - m_y) p_{ij};$$

если (X, Y) — непрерывная двумерная с. в., то

$$K_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y)\, dx dy$$

Ковариация

Ковариацию часто удобно вычислять по формуле

$$K_{XY} = \operatorname{cov}(X, Y) = MXY - MX \cdot MY,$$

если (X, Y) — непрерывная двумерная с. в., то

$$K_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x, y)\, dx dy - m_x m_y.$$

Свойства ковариации

1. Ковариация симметрична, т. е.

$$K_{XY} = K_{YX}.$$

2. Дисперсия с. в. есть ковариация ее с самой собой, т. е.

$$K_{XX} = DX, \qquad K_{YY} = DY.$$

3. Если случайные величины X и Y независимы, то

$$K_{XY} = 0.$$

4. Дисперсия суммы (разности) двух случайных величин равна сумме их дисперсий плюс (мину , удвоенная ковариация этих случайных величин, т. е.

$$D(X \pm Y) = DX + DY \pm 2K_{XY}.$$

Свойства ковариации

5. Постоянный множитель можно вынести за знак ковариаций. т. е.

$$K_{cX,Y} = c\cdot K_{XY} = K_{X,cY} \quad \text{или} \quad \operatorname{cov}(cX,Y) = c\cdot\operatorname{cov}(X,Y) = \operatorname{cov}(X,cY).$$

6. Ковариация не изменится, если к одной из с. в. (или к обоим сразу) прибавить постоянную, т. е.

$$K_{X+c,Y} = K_{XY} = K_{X,Y+c} = K_{X+c,Y+c}$$

или

$$\operatorname{cov}(X+c,Y) = \operatorname{cov}(X,Y) = \operatorname{cov}(X,Y+c) = \operatorname{cov}(X+c,Y+c).$$

7. Ковариация двух случайных величин по абсолютной величине не превосходит их с. к. о., т. е.

$$|K_{XY}| \leqslant \sigma_x\cdot\sigma_y.$$

Коэффициент корреляции

Коэффициентом корреляции r_{XY} двух с. в. X и Y называется отношение их ковариации (корреляционного момента) к произведению их с. к. о.:

$$r_{XY} = \frac{K_{XY}}{\sigma_x \sigma_y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}.$$

Очевидно, коэффициент корреляции равен ковариации стандартных с. в. $Z_1 = \dfrac{X - m_x}{\sigma_x}$ и $Z_2 = \dfrac{Y - m_y}{\sigma_y}$, т. е. $r_{XY} = \operatorname{cov}(Z_1, Z_2)$.

Свойства коэффициента корреляции

1. Коэффициент корреляции по абсолютной величине не превосходит 1, т. е.

$$|r_{XY}| \leqslant 1 \quad \text{или} \quad -1 \leqslant r_{XY} \leqslant 1.$$

2. Если X и Y независимы, то

$$r_{XY} = 0.$$

3. Если с. в. X и Y связаны линейной зависимостью, т. е. $Y = aX + b$, $a \neq 0$, то

$$|r_{XY}| = 1,$$

причем $r_{XY} = 1$ при $a > 0$,
$r_{XY} = -1$ при $a < 0$.

4. Если $|r_{XY}| = 1$, то с. в. X и Y связаны линейной функциональной зависимостью.

Предельные теоремы теории вероятностей

Предельные теоремы теории вероятностей устанавливают связь между теоретическими и экспериментальными характеристиками случайных величин при большом числе испытаний над ними.

Обычно рассматривают два крупных блока предельных теорем: **закон больших чисел** и **центральную предельную теорему**.

Закон больших чисел утверждает, что среднее арифметическое большого числа независимых случайных величин сходится по вероятности к среднему арифметическому из их математических ожиданий.

Согласно **центральной предельной теореме** достаточно большая сумма примерно одинаковых случайных величин ведет себя приближенно как нормально распределенная случайная величина.

Закон больших чисел

Закон больших чисел базируется на понятии сходимости случайных величин по вероятности и неравенствах Маркова и Чебышева.

Определение сходимости по вероятности.

Последовательность случайных величин $X_1, X_2, \ldots, X_n, \ldots$ *сходится по вероятности к случайной величине* X *, что записывают в виде* $X_n \xrightarrow{P} X$ *при* $n \to \infty$ *, если для любого* $\varepsilon > 0$ *выполняется следующее условие:*

$$\lim_{n\to\infty} P\left(|X_n - X| \geq \varepsilon\right) = 0.$$

В дальнейшем часто будет использоваться сходимость по вероятности к каким-либо константам, так что в данном выше определении случайная величина X может быть равна некоторой постоянной величине C .

Неравенство Маркова

Теорема (Неравенство Маркова). *Если неотрицательная случайная величина $X(\omega)$ имеет математическое ожидание m_x, то для любого положительного числа ε справедливо неравенство Маркова*

$$P(X(\omega) \geq \varepsilon) \leq \frac{m_x}{\varepsilon}.$$

Неравенство Маркова позволяет оценить сверху вероятность попадания неотрицательной случайной величины X на промежуток $[\varepsilon, +\infty)$.

Первое неравенство Чебышева

Теорема (Первое неравенство Чебышева). *Если случайная величина X имеет математическое ожидание m_x и дисперсию $D_x = \sigma_x^2$, то для любого положительного числа ε справедливо неравенство Чебышева*

$$P\left(|X - m_x| \geq \varepsilon\right) \leq \frac{\sigma_x^2}{\varepsilon^2}.$$

Второе неравенство Чебышева

Неравенство Чебышева позволяет оценить сверху вероятность события $(-\infty, \varepsilon] \cup [\varepsilon, +\infty)$ для центрированной случайной величины $X - m_x$. Если перейти к противоположному событию

$$\overline{\{|X - m_x| \geq \varepsilon\}} = \{|X - m_x| < \varepsilon\},$$

то его вероятность равна $P\{|X - m_x| < \varepsilon\} = 1 - P\{|X - m_x| \geq \varepsilon\}$.

Отсюда следует, **второе неравенство Чебышева**:

$$P(|X - m_x| < \varepsilon) = P(m_x - \varepsilon < X < m_x + \varepsilon) \geq 1 - \frac{\sigma_x^2}{\varepsilon^2}.$$

Данное неравенство позволяет оценить снизу вероятность попадания случайной величины X на интервал $(m_x - \varepsilon, m_x + \varepsilon)$.

Второе неравенство Чебышева

Пример. Найти нижнюю оценку вероятности попадания случайной величины X с математическим ожиданием m_x и стандартным отклонением σ_x на симметричный относительно m_x интервал длины $6\sigma_x$

$$(m_x - 3\sigma_x; m_x + 3\sigma_x) = \{|X - m_x| < 3\sigma_x\}.$$

Положим во втором неравенстве Чебышева значение ε равным $3\sigma_x$. Тогда вероятность попадания на указанный интервал будет более чем $8/9$. $\square$

Теорема Чебышева

Теорема (Чебышева). *Если случайные величины в последовательности* $X_1, X_2, \ldots, X_n, \ldots$ *попарно независимы, одинаково распределены, имеют одинаковые математические ожидания* m *и одинаковые дисперсии* D, *то последовательность средних арифметических*

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \quad n \in \mathbb{N}$$

сходится по вероятности к математическому ожиданию m:

$$\bar{X}_n \xrightarrow{P} m.$$

Закон больших чисел

Теорема Чебышева сформулирована для случая одинаково распределенных случайных величин. Есть теоремы, в которых рассматриваются случайные величины с различными математическими ожиданиями и дисперсиями. В них, при разнообразных требованиях к существованию дисперсий и их ограниченности, доказывается сходимость по вероятности средних арифметических случайных величин к среднему арифметическому их математических ожиданий. Во всех указанных теоремах принято говорить, что к таким случайным величинам применим **закон больших чисел**.

Закон больших чисел

Теорема Чебышева лежит в основе асимптотических методов в математической статистике. Действительно, выборочные наблюдения случайной величины X рассматриваются в математической статистике как одинаково распределенные и взаимно независимые компоненты n- мерного случайного вектора. Вычисляя их среднее арифметическое, можно за счет увеличения объема выборки n сколь угодно точно оценить математическое ожидание.

Теорема Бернулли

Теорема (Бернулли). *Относительная частота* $\frac{k}{n}$ *события* A *при* n *независимых испытаниях по схеме Бернулли сходится по вероятности к вероятности* $p = P(A)$ *события* A:

$$\frac{k}{n} \xrightarrow{P} p = P(A).$$

Теорема Бернулли является следствием теоремы Чебышева.

Теорема Бернулли теоретически обосновывает возможность сколь угодно точной оценки вероятности события A по формуле $P(A) = \frac{k}{n}$ в достаточно большой серии из n независимых испытаний.

Центральная предельная теорема

Приведем без доказательства центральную предельную теорему для одинаково распределенных случайных величин в формулировке Ляпунова.

Теорема (Ляпунова). *Пусть случайные величины в последовательности $X_1, X_2, \ldots, X_n, \ldots$ взаимно независимы, одинаково распределены, имеют равные математические ожидания m и равные дисперсии $D = \sigma^2$.*

Тогда функция распределения $F_{Z_n}(z)$, $n \in \mathbb{N}$ центрированной и нормированной суммы данных случайных величин

$$Z_n = \frac{\sum_{k=1}^{n} X_k - nm}{\sigma\sqrt{n}}, \quad n \in \mathbb{N}$$

сходится для всех $z \in \mathbb{R}$ к функции распределения Лапласа

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z).$$

Теорема Муавра–Лапласа

Следствием центральной предельной теоремы является теорема Муавра–Лапласа, относящаяся к схеме Бернулли.

Теорема (Муавра–Лапласа). *Пусть случайная величина* B_n *имеет биномиальное распределение и определена как число появлений события* A *при* n *независимых испытаниях по схеме Бернулли с вероятностью* $p = P(A)$ *успеха события* A *в каждом испытании.*

Тогда функции распределения $F_{\bar{B}_n}$, $n \in \mathbb{N}$ *центрированной и нормированной последовательности случайных величин* $\bar{B}_n = \frac{B_n - np}{\sqrt{npq}}$, $n \in \mathbb{N}$ *сходятся для всех* $z \in \mathbb{R}$ *к функции распределения Лапласа* $\lim\limits_{n\to\infty} F_{\bar{B}_n}(z) = \Phi(z)$.

Теорема Муавра–Лапласа

Теорема Муавра–Лапласа теоретически обосновывает возможность приближения дискретного биномиального распределения непрерывным нормальным распределением с параметрами $m = pn$ и $\sigma = \sqrt{npq}$.

В частности, для вычисления вероятности попадания биномиальной случайной величины на дискретный интервал $[k_1, k_2)$ при достаточно больших значениях npq применяется интегральная формула Муавра – Лапласа

$$P(k_1 \le B_n < k_2) = P\left(\frac{k_2 - np}{\sqrt{npq}} \le \frac{B_n - np}{\sqrt{npq}} < \frac{k_1 - np}{\sqrt{npq}}\right) =$$

$$= \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right) + O\left(\frac{1}{\sqrt{npq}}\right),$$

Теорема Муавра–Лапласа

а для вычисления вероятности любого из возможных значений $k = 0, 1, \ldots, n$ биномиальной случайной величины – локальная формула Муавра–Лапласа

$$P(B_n = k) = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} + O\left(\frac{1}{\sqrt{npq}}\right).$$

Высокая точность приближения достигается в тех случаях, когда выполняются условия $npq > 5$ и $0,1 < p < 0,9$. При условии $npq > 25$ приближение можно применять при любых вероятностях $p = P(A)$ события A.