5. Нормальное распределение. Закон больших чисел

Нормальное распределение непрерывной случайной величины. Плотность и функция распределения. Вероятность попадания в интервал. Вероятность заданного отклонения. Стандартная нормальная случайная величина. Законы больших чисел. Центральная предельная теорема.

5.1. Плотность и функция распределения

Рассмотрим ещё одно распределение непрерывной случайной величины, имеющее большое теоретическое и прикладное значение.

Определение 5.1. *Случайная величина ξ имеет **нормальное распределение** с параметрами a и σ, если её плотность распределения имеет вид:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(x-a)^2}{2\sigma^2}}. \tag{5.1}$$

Этот факт будем записывать так: $\xi \sim N(a;\sigma)$.

Нормальное распределение определяется двумя параметрами a и σ.

Докажем, что $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$. Действительно:

$$\int\limits_{-\infty}^{+\infty} f(x)dx = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left[\begin{array}{c} \frac{(x-a)}{\sigma} = t \implies x = \sigma t + a \\ dx = \sigma dt \end{array}\right] =$$

$$= \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} \sigma dt = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt.$$

Полученный интеграл называется интегралом Пуассона и его значение равно $\sqrt{2\pi}$.

$$\int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}. \tag{5.2}$$

Подставив этот результат в последнее выражение, получим

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1.$$

Найдем функцию распределения:

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt.$$

Выразим функцию распределения нормального закона $F(x)$ через функцию Лапласа $\Phi(x)$, введённую в лекции 2 (формула 2.17)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-\frac{t^2}{2}} dt.$$

Сделаем замену переменных:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \left[\begin{array}{c} \frac{(t-a)}{\sigma} = z \implies t = \sigma z + a \\ dt = \sigma dz \end{array} \right] =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} \sigma\, dz = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{0} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

$$= 0{,}5 + \Phi\left(\frac{x-a}{\sigma}\right).$$

Здесь использовался тот факт, что из четности подынтегральной функции в интеграле Пуассона следует:

$$\int\limits_{-\infty}^{0} e^{-\frac{z^2}{2}} dz = \int\limits_{0}^{+\infty} e^{-\frac{z^2}{2}} dz = \frac{\sqrt{2\pi}}{2}. \tag{5.3}$$

Итак, для функции распределения нормального закона получим выражение:

$$F(x) = 0{,}5 + \Phi\left(\frac{x-a}{\sigma}\right). \tag{5.4}$$

Отметим, что при $a = 0$ и $\sigma = 1$

$$F(x) = 0{,}5 + \Phi\big(x\big). \tag{5.5}$$

Найдём математическое ожидание нормально распределённой случайной величины.

$$M(\xi)=\int\limits_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left[\begin{array}{c}\frac{(x-a)}{\sigma}=t \implies x=\sigma t+a\\ dx=\sigma d\,t\end{array}\right]=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}(\sigma t+a)e^{-\frac{t^2}{2}}dt=\frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt+\frac{a}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=$$

$$=0+\frac{a}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=0+a=a.$$

При выводе значения для $M(\xi)$ использовалась формула для интеграла Пуассона (5.3) и нечётность подынтегральной функции $te^{-\frac{t^2}{2}}$, поэтому $\int\limits_{-\infty}^{+\infty} te^{-\frac{t^2}{2}}dt=0$.

Найдём теперь числовое значение для дисперсии

$$D(\xi)=M(\xi-M(\xi))^2=\int\limits_{-\infty}^{+\infty}(x-a)^2\tfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=$$
$$=\left[\begin{array}{c}\frac{(x-a)}{\sigma}=t \implies x=\sigma t+a\\ dx=\sigma d\,t\end{array}\right]=\tfrac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{+\infty}\sigma^2t^2e^{-\frac{t^2}{2}}\sigma dt=$$
$$=-\tfrac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}t\,de^{-t^2/2}=-\tfrac{\sigma^2}{\sqrt{2\pi}}\left(te^{-t^2/2}\Big|_{-\infty}^{+\infty}-\int\limits_{-\infty}^{+\infty}e^{-t^2/2}\,dt\right)=$$
$$=-\tfrac{\sigma^2}{\sqrt{2\pi}}(0-\sqrt{2\pi})=\sigma^2.$$

Итак, для случайной величины ξ, имеющей нормальное распределение, параметры a и σ имеют простой вероятностный смысл:

$$M(\xi)=a;\quad D(\xi)=\sigma^2;\quad \sigma(\xi)=\sigma. \tag{5.6}$$

Графики плотности и функции распределения нормального закона приведены на рис. 22. График плотности нормального распределения иногда называют кривой Гаусса.

На рис. 23 проиллюстрирована зависимость плотности от параметра σ при $a=1$. При увеличении σ значение максима функции, которое равно $\dfrac{1}{\sqrt{2\pi}\sigma}$ линейно уменьшается и график функции становиться более пологим. При уменьшении параметра σ максимум функции линейно возрастает и график

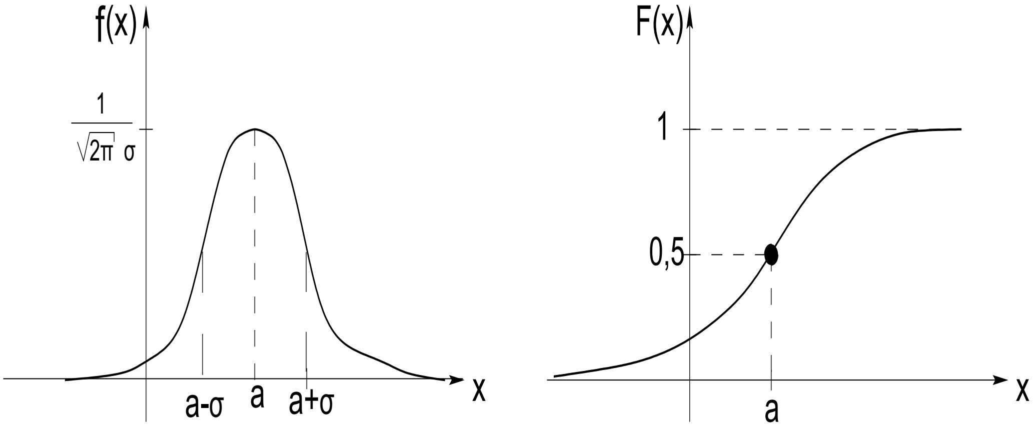

Рисунок 22. *Плотность и функция распределения нормального распределения*

функции растягивается вверх. Отметим, что максимальное значение функции плотности нормального распределения приблизительно равно $0{,}4/\sigma$.

На рис. 24 приведена зависимость плотности от параметра a при $\sigma = 1$. График функции плотности нормального распределения симметричен относительно прямой $x = a$.

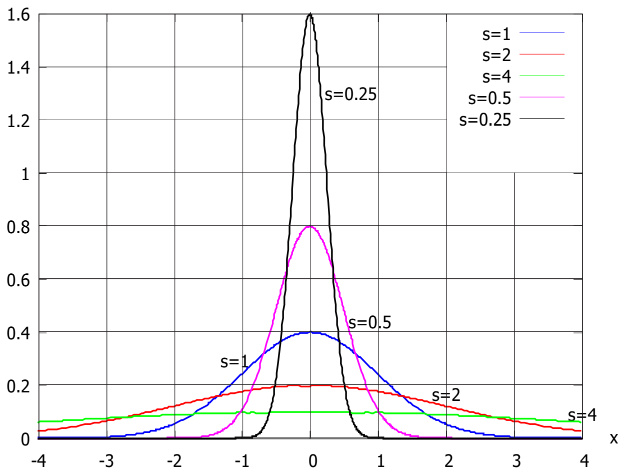

Рисунок 23. *Зависимость от σ*

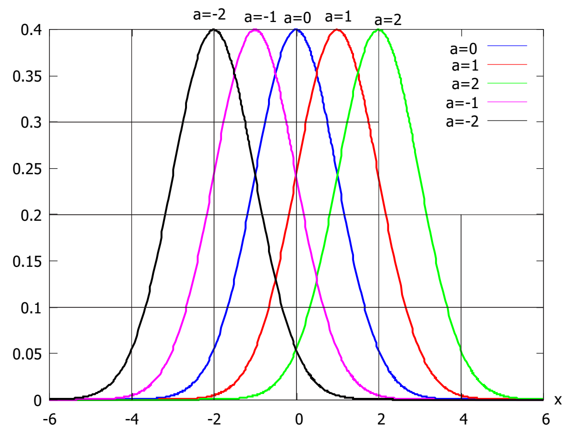

Рисунок 24. *Зависимость от a*

В соответствии со свойством 2 функции распределения получаем формулу для вычисления вероятности попадания нормальной случайной величины в заданный интервал:

$$P(x_1 \leqslant \xi < x_2) = F(x_2) - F(x_1) = \left(0{,}5 + \Phi\left(\frac{x_2 - a}{\sigma}\right)\right) -$$

$$-\left(0{,}5 + \Phi\left(\frac{x_1 - a}{\sigma}\right)\right) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right),$$

$$P(x_1 \leqslant \xi < x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right). \tag{5.7}$$

В частности, если интервал полубесконечный, учитывая тот факт, что $\Phi(+\infty) = 0{,}5,\ \Phi(-\infty) = -0{,}5$, получаем:

$$P(\xi < x_2) = P(-\infty < \xi < x_2) = \Phi\Big(\frac{x_2 - a}{\sigma}\Big) - \Phi(-\infty) = \Phi\Big(\frac{x_2 - a}{\sigma}\Big) + 0{,}5,$$
$$P(x_1 \leqslant \xi) = P(x_1 \leqslant \xi < \infty) = 0{,}5 - \Phi\Big(\frac{x_1 - a}{\sigma}\Big).$$

5.2. Вероятность заданного отклонения для нормального распределения

Пользуясь формулой (5.7), можно получить формулу для вычисления вероятности заданного отклонения нормальной случайной величины от математического ожидания:

$$P(|\xi - a| < \varepsilon) = P(-\varepsilon < \xi - a < \varepsilon) = P(a - \varepsilon < \xi < a + \varepsilon) =$$
$$= \Phi\Big(\frac{a + \varepsilon - a}{\sigma}\Big) - \Phi\Big(\frac{a - \varepsilon - a}{\sigma}\Big) = \Phi\Big(\frac{\varepsilon}{\sigma}\Big) - \Phi\Big(-\frac{\varepsilon}{\sigma}\Big) = 2\Phi\Big(\frac{\varepsilon}{\sigma}\Big).$$

Окончательно имеем:

$$P(|\xi - a| < \varepsilon) = 2\Phi\Big(\frac{\varepsilon}{\sigma}\Big). \tag{5.8}$$

Замечание 5.1. *Познакомившись с нормальным распределением, заметим, что локальная и интегральные теоремы Лапласа дают приближения для вероятностей биномиально распределённой случайной величины через соответствующие вероятности нормально распределённой случайной величины. Аналогично, с помощью формулы (5.7) получается приближённая формула (2.20) для вероятности отклонения частоты от вероятности в испытаниях Бернулли.*

Пример 5.1. $\xi \sim N(20; 10)$. *Найти* $P(|\xi - 20| < 3)$ *и* $P(|\xi - 10| < 3)$.

◀По формуле (5.8) определяем

$$P(|\xi - 20| < 3) = 2\Phi\Big(\frac{3}{10}\Big) \approx 2 \cdot 0{,}1179 = 0{,}2358.$$

Значение $\Phi(0{,}3) = 0{,}1179$ находим по таблице приложения 2.

Для нахождения $P(|\xi - 10| < 3)$ нельзя применить формулу (5.7), т.к. $a = 20 \neq 10$. Эту вероятность найдём по формуле (5.6):

$$P(|\xi - 10| < 3) = P(-3 < \xi - 10 < 3) = P(7 < \xi < 13) =$$
$$= \Phi\Big(\frac{13-20}{10}\Big) - \Phi\Big(\frac{7-20}{10}\Big) = \Phi(1,3) - \Phi(0{,}7) \approx$$
$$\approx 0{,}4032 - 0{,}2580 = 0{,}1452.$$

Ответ: $P(|\xi - 20| < 3) \approx 0{,}236$; $P(|\xi - 10| < 3) \approx 0{,}145$.

Применим формулу (5.8) для вычисления вероятности отклонения при $\varepsilon = 3\sigma$.

$$P(|\xi - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) \approx 2 \cdot 0{,}49865 = 0{,}9973.$$

Мы получили известное в технике «правило трёх сигм»: *для нормально распределённой случайной величины практически невозможно её отклонение от математического ожидания по абсолютной величине более трёх σ.*

На практике в менее ответственных случаях можно также применять аналогичное «правило двух сигм» т.к.

$$P(|\xi - a| < 2\sigma) \approx 0{,}9544.$$

Пример 5.2. *Случайная величина ξ подчиняется нормальному закону распределения вероятностей с параметрами $a = 3$, $\sigma = 2$. найти:*
1) $P(1 < \xi < 4)$; *2)* $P(|\xi - 3| < 1)$; *3)* $P(|\xi - 4| < 4)$.

◀ 1) По формуле (5.7) $P(x_1 \leqslant \xi < x_2) = \Phi\Big(\frac{x_2 - a}{\sigma}\Big) - \Phi\Big(\frac{x_1 - a}{\sigma}\Big)$ имеем:

$P(1 < \xi < 4) = \Phi\left(\frac{4-3}{2}\right) - \Phi\left(\frac{1-3}{2}\right) = \Phi(0{,}5) - \Phi(-1) =$
$= \Phi(0{,}5) + \Phi(1) = 0{,}1915 + 0{,}3413 = 0{,}5328.$

2) По формуле (5.8) $P(|\xi - a| < \varepsilon) = 2\Phi\Big(\frac{\varepsilon}{\sigma}\Big)$, получаем

$P(|\xi - 3| < 1) = 2\Phi\left(\frac{1}{2}\right) = 2\Phi(0{,}5) = 0{,}383.$

3) Для нахождения $P(|\xi - 4| < 4)$ нельзя применить формулу (5.8), т.к. $a = 3 \neq 4$. Эту вероятность найдём по формуле (5.7):

$P(|\xi - 4| < 4) = P(-4 < \xi - 4 < 4) = P(0 < \xi < 8) = \Phi\left(\frac{8-3}{2}\right) -$
$- \Phi\left(\frac{0-3}{2}\right) = \Phi(2{,}5) - \Phi(-1{,}5) = 0{,}4938 + 0{,}4332 = 0{,}927.$ ▶

Пример 5.3. *Ошибка измерения некоего измерительного прибора имеет нормальное распределение. Прибор не имеет систематической ошибки, а средняя квадратическая ошибка равна 5. Найти вероятность того, что ошибка измерения не превзойдет по абсолютной величине 2.*

◀ Применяем формулу (5.8):
$P(|\xi - a| < \varepsilon) = 2\Phi\left(\dfrac{\varepsilon}{\sigma}\right)$.

Здесь ξ — ошибка измерительного прибора,
a — математическое ожидание,
$\sigma = 5$ — средняя квадратическая ошибка измерения,
$\varepsilon = 2$ — ошибка измерения.

Получаем,
$P(|\xi - a| < 2) = 2\Phi\left(\dfrac{2}{5}\right) = 2\Phi(0{,}4) = 2 \cdot 0{,}1554 = 0{,}3108$. ▶

Пример 5.4. *Рассеивание скорости снаряда подчинено нормальному распределению. Найти среднее квадратическое отклонение рассеивания, если с вероятностью 0,996 оно не превосходит по абсолютной величине 5 м/с. Систематическая ошибка отсутствует.*

◀ Применяем формулу (5.8):$P(|\xi - a| < \varepsilon) = 2\Phi\left(\dfrac{\varepsilon}{\sigma}\right)$.
Здесь ξ — рассеивания скорости снаряда,
a — математическое ожидание,
σ — искомая величина обозначающая среднее квадратическое отклонение скорости рассеивания снаряда,
$\varepsilon = 5$ — максимальное отклонение скорости рассеивания снаряда.

Получаем,
$P(|\xi - a| < 5) = 2\Phi\left(\dfrac{5}{\sigma}\right) \;\Rightarrow\; 0{,}996 = 2\Phi\left(\dfrac{5}{\sigma}\right)$.

Получили уравнение относительно аргумента функции

$$\Phi\left(\frac{5}{\sigma}\right) = 0{,}498.$$

Из таблицы приложение 2, находим при каком значении аргумента значение функции равно 0,498. Получаем линейное уравнение

$\dfrac{5}{\sigma} \approx 2{,}88 \;\Rightarrow\; \sigma = 5/2{,}88 \approx 1{,}736$. ▶

Ответ: $\approx 1{,}736$.

Пример 5.5. *Срок службы электрической лампы является случайной величиной, распределённой по нормальному закону. Найти среднее время T срока службы лампы, если с вероятностью 0,9505 лампа работает более 1000ч. Среднее квадратическое отклонение 20ч.*

◀ Применяем формулу (5.7):
$P(x_1 \leqslant \xi < x_2) = \Phi\Big(\dfrac{x_2 - a}{\sigma}\Big) - \Phi\Big(\dfrac{x_1 - a}{\sigma}\Big).$

Здесь ξ — срок службы электрической лампы,
a — искомая величина обозначающая среднее время T срока службы лампы,
$\sigma = 20$ — среднее квадратическое отклонение срока службы лампы,
$x1 = 1000$, $x2 = +\infty$ — границы интервала на котором вероятность равна 0,9505.

Получаем,
$P(\xi > 1000) = \Phi(+\infty) - \Phi\left(\dfrac{1000 - a}{20}\right)$
$\Phi\left(\dfrac{1000 - a}{20}\right) = 0{,}9505 - 0{,}5 = 0{,}4505.$

Из таблицы приложение 2, находим при каком значении аргумента значение функции равно 0,4505.

Задачу свели к линейному уравнению

$\dfrac{1000 - a}{20} = 1{,}65 \ \Rightarrow\ 1000 - a = 33 \ \Rightarrow\ a = 1033.$ ▶
Ответ: ≈ 1033.

5.3. Стандартная нормальная случайная величина

Теорема 5.1. *Если $\xi \sim N(a;\sigma)$, то $\eta = k\xi + b \sim N(ka + b;\ |k|\ \sigma)$.*

◀Найдём функцию распределения $F_\eta(x)$ случайной величины η при $k > 0$:

$$F_\eta(x) = P(\eta < x) = P(k\xi + b < x) = P\Big(\xi < \frac{x - b}{k}\Big) =$$
$$= 0{,}5 + \Phi\left(\frac{\frac{x-b}{k} - a}{\sigma}\right) = 0{,}5 + \Phi\Big(\frac{x - (ka + b)}{k\sigma}\Big).$$

Таким образом, доказано, что $\eta \sim N(ka+\sigma;\ k\sigma)$ при $k > 0$. Проведем аналогичные выкладки при $k < 0$:

$$F_\eta(x) = P(\eta < x) = P(k\xi + b < x) = P\Big(\xi > \frac{x-b}{k}\Big) =$$
$$= 1 - P\Big(\xi < \frac{x-b}{k}\Big) = 1 - 0{,}5 - \Phi\Bigg(\frac{\frac{x-b}{k} - a}{\sigma}\Bigg) =$$
$$= 0{,}5 - \Phi\Big(\frac{x-(ka+b)}{k\sigma}\Big) = 0{,}5 + \Phi\Big(\frac{x-(ka+b)}{-k\sigma}\Big),$$

т.е. при $k < 0$ $\eta \sim N(ka+\sigma;\ -k\sigma)$. Обобщая эти два вывода, получим утверждение теоремы.▶

Теорема 5.2. *Если $\eta \sim N(a;\ \sigma)$, то $\xi_{cm} = \dfrac{\eta - a}{\sigma} \sim N(0; 1)$.*

Действительно, так как $\xi_{\text{ст}} = \dfrac{\eta}{\sigma} - \dfrac{a}{\sigma}$, то по теореме 5.1 для $k = \dfrac{1}{\sigma}$, $b = -\dfrac{a}{\sigma}$, получаем, что $\xi_{\text{ст}}$ имеет нормальное распределение с параметрами

$$\frac{1}{\sigma}\cdot a - \frac{a}{\sigma} = 0 \quad \text{и} \quad \frac{1}{\sigma}\cdot\sigma = 1.$$

Определение 5.2. *Случайная величина, имеющая нормальное распределение с параметрами $a = 0$ и $\sigma = 1$, называется* ***стандартной*** *(нормированной)* ***нормальной случайной величиной*** *, а её распределение — стандартным (нормированным) нормальным.*

Плотность и функция стандартного нормального распределения даются формулами:

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}; \quad F_{\text{ст}}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = 0{,}5 + \Phi(x). \tag{5.9}$$

5.4. Предельные теоремы теории вероятностей

Законы больших чисел

Теперь познакомимся с разделом теории вероятностей, посвящённым получению приближённых формул для вероятностей суммы большого числа случайных величин.

Неравенство Чебышёва даёт оценку вероятности того, что случайная величина примет значение, далёкое от своего среднего.

Теорема 5.3. *(**Неравенство Чебышева**.) Для случайной величины ξ при $\forall\ \varepsilon > 0$ верно неравенство:*

$$P(|\xi - M(\xi)| < \varepsilon) \geqslant 1 - \frac{D(\xi)}{\varepsilon^2}. \tag{5.10}$$

Доказательство проведём для непрерывной случайной величины ξ с плотностью $f(x)$, хотя теорема верна и для дискретных случайных величин.

Оценим вероятность противоположного события:

$$P(|\xi - M(\xi)| \geqslant \varepsilon) = P(\xi \geqslant M(\xi) + \varepsilon \text{ или } \xi \leqslant M(\xi) - \varepsilon) =$$

$$= P(\xi \leqslant M(\xi) - \varepsilon) + P(\xi \geqslant M(\xi) + \varepsilon) = \int\limits_{-\infty}^{M(\xi)-\varepsilon} f(x)dx + \int\limits_{M(\xi)+\varepsilon}^{+\infty} f(x)dx \leqslant$$

$$\leqslant \int\limits_{-\infty}^{M(\xi)-\varepsilon} \frac{\big(x - M(\xi)\big)^2}{\varepsilon^2} f(x)dx + \int\limits_{M(\xi)+\varepsilon}^{+\infty} \frac{\big(x - M(\xi)\big)^2}{\varepsilon^2} f(x)dx \leqslant$$

$$\leqslant \int\limits_{-\infty}^{M(\xi)} \frac{\big(x - M(\xi)\big)^2}{\varepsilon^2} f(x)dx + \int\limits_{M(\xi)}^{+\infty} \frac{\big(x - M(\xi)\big)^2}{\varepsilon^2} f(x)dx =$$

$$= \frac{1}{\varepsilon^2} \int\limits_{-\infty}^{+\infty} \big(x - M(\xi)\big)^2 f(x)dx = \frac{D(\xi)}{\varepsilon^2}.$$

Первое неравенство в этой цепочке объясняется тем, что подынтегральные функции умножили на выражение $\dfrac{\big(x - M(\xi)\big)^2}{\varepsilon^2}$, которое больше или равно 1, т.к. в области интегрирования x удовлетворяет неравенству $|x - M(\xi)| \geqslant \varepsilon$.

Второе неравенство верно, т.к. при увеличении интервала интегрирования интеграл от неотрицательной функции не уменьшается.

Из полученного неравенства:

$$P(|\xi - M(\xi)| \geqslant \varepsilon) \leqslant \frac{D(\xi)}{\varepsilon^2},$$

переходя к вероятности противоположного события, получаем неравенство Чебышева.

Пример 5.6. *В партии 10 лампочек вероятность отказа каждой из которых* 0,05. *Оценить вероятность того, что абсолютная величина отклонения числа отказавших ламп от математического ожидания меньше одного.*

◀Пусть ξ — число отказавших лампочек; эта случайная величина имеет биномиальное распределение с параметрами $n = 10$, $p = 0{,}05$.

$$M(\xi) = np = 0{,}5;\ D(\xi) = npq = 0{,}475.$$

По теореме 5.3 имеем:

$$P(|\xi - 0{,}5| < 1) \geqslant 1 - \frac{0{,}475}{1}.$$

Ответ: $P(|\xi - 0{,}5| < 1) \geqslant 0{,}525$.▶

Пример 5.7. *Средний дневной расход электроэнергии на предприятии составляет 2000 кВт.ч, а среднее квадратическое отклонение расхода электроэнергии не превышает 250 кВт.ч. Оценить вероятность того, что в конкретный день расход электроэнергии не превзойдет 3000кВт.ч.*

◀ Введем случайную величину ξ — дневной расход электроэнергии на предприятии (кВт.ч).
По условию задачи $M(\xi) = 2000$. Дисперсия $D(\xi) = \sigma^2 = 250^2$.

По теореме 5.3 имеем:

$P(|\xi - 2000| < 1000) \geqslant 1 - \frac{250^2}{1000^2} = \dfrac{15}{16} = 0{,}9375$. ▶

Ответ: $\geqslant 0{,}9375$.

Пример 5.8. *Автомат в смену выпускает 4000 деталей. Вероятность выхода бракованной детали равна 0,02. Оценить с помощью неравенства Чебышева вероятность того, что число бракованных деталей находится в диапазоне от 60 до 100.*

◀ Введем случайную величину ξ — число бракованных деталей. По условию задачи данная случайная величина имеет биномиальное распределение с параметрами $M(\xi) = np = 4000 \cdot 0{,}02 = 80$, дисперсия $D(\xi) = np(1-p) = 80 \cdot 0{,}98 = 78{,}4$.

Искомую вероятность можно записать в виде:

$P(60 \leqslant m \leqslant 100) = P(-20 \leqslant m - 80 \leqslant 20) = P(|m-80| \leqslant 20)$.

По теореме 5.3 имеем:

$P(|m-80| \leqslant 20) \geqslant 1 - \dfrac{78{,}4}{20^2} = 0{,}804$.

Данную задачу можно решить при помощи следствия из теоремы Муавра-Лапласа

$$P(|m-80| \leqslant 20) \approx 2 \cdot \Phi\left(\frac{20}{78{,}4}\right) = 2\Phi(2{,}26) = 0{,}976.$$

Полученные результаты показывают, что неравенство Чебышева достаточно грубо оценивает результат. ►

Замечание 5.2. *Неравенство Чебышева используется при доказательстве ряда теорем (иногда его называют лемма Чебышева), однако оно даёт довольно грубую оценку для приведённой вероятности.*

Так, в примере 5.6, раскрывая модуль, мы получили неравенство:

$$P(-0{,}5 < \xi < 1{,}5) \geqslant 0{,}525.$$

Однако приведённый интервал может быть заведомо уменьшен, т.к. $\xi \geqslant 0$*:*

$$P(-0{,}5 < \xi < 1{,}5) = P(0 \leqslant \xi < 1{,}5).$$

П. Л. Чебышёв получил общую формулировку закона больших чисел: если математические ожидания серии случайных величин и квадраты этих математических ожиданий ограничены в совокупности, то среднее арифметическое этих величин с ростом сходится по вероятности к среднему арифметическому для их математических ожиданий.

Теорема 5.4. ***(Закон больших чисел в форме Чебышева.)*** *Если* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *— независимые случайные величины с равномерно ограниченными дисперсиями* $\Big(D(\xi_i) \leqslant C,\ i = 1, 2, \ldots\Big)$, *то для* $\forall\ \varepsilon > 0$ *будет:*

$$\lim_{n\to\infty} P\left(\left|\ \frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M(\xi_i)\ \right| < \varepsilon\right) = 1.$$

◀Обозначим $\eta_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$.

Пользуясь свойствами математического ожидания и дисперсии, получаем:

$$M(\eta_n) = \frac{\sum\limits_{i=1}^{n} M(\xi_i)}{n}, \quad D(\eta_n) = \frac{\sum\limits_{i=1}^{n} D(\xi_i)}{n^2} \leqslant \frac{n \cdot C}{n^2} = \frac{C}{n}.$$

На основании неравенства Чебышева для η_n получаем:

$$P(|\eta_n - M(\eta_n)| < \varepsilon) \geqslant 1 - \frac{D(\eta_n)}{\varepsilon^2} \Longleftrightarrow$$

$$\Longleftrightarrow 1 \geqslant P\left(\left|\frac{\sum\limits_{i=1}^{n}\xi_i}{n} - \frac{\sum\limits_{i=1}^{n} M(\xi_i)}{n}\right| < \varepsilon\right) \geqslant 1 - \frac{C}{n\varepsilon^2}.$$

Переходя к пределу при $n \to \infty$, поскольку пределы левой и правой частей равны 1, получаем утверждение теоремы.▶

Закон больших чисел в форме Чебышева утверждает, что для большого числа независимых случайных величин практически невозможны значительные отклонения их среднего арифметического от среднего арифметического их математических ожиданий.

Следствие 5.1. *Если в условиях теоремы 5.4 $M(\xi_1) = M(\xi_2) = \ldots = a$, то для $\forall\ \varepsilon > 0$ будет:*

$$\lim_{n\to\infty} P\left(\left|\frac{\sum \xi_i}{n} - a\right| < \varepsilon\right) = 1.$$

Действительно, в этом случае $M(\eta_n) = \dfrac{\sum\limits_{i=1}^{n} M(\xi_i)}{n} = \dfrac{na}{n} = a$.

Замечание 5.3. *На практике закон больших чисел в форме Чебышева применяют, например, в теории ошибок. Следует отметить, что результат любого измерения есть случайная величина. При этом различают грубые ошибки измерения, которые можно устранить, основываясь на физической природе измеряемого объекта, Так, если в ряду измерения роста группы людей встретилось значение 17,8 м. — это, очевидно, грубая ошибка измерения. Данный результат следует изъять, если нельзя его уточнить. Далее, бывают систематические ошибки измерения. Эти ошибки,* *как правило, вызываются неисправностью измерительного прибора; они не являются случайными и их можно устранить, проверив прибор и внеся поправку в измерения. Так, например, если часы спешат на 5 минут, то от измеренной величины нужно отнять 5 минут, чтобы получить верное время. Наконец, все остальные ошибки — случайные ошибки измерения, вызываются множеством различных факторов: дрожание стрелки прибора, неточное считывание показаний («косо взглянул» на стрелку), отклонения в условиях измерения и проч. Таким образом, результат измерения можно считать случайной величиной, равной сумме большого числа других случайных величин. В соответствии с теоремой 5.4 для уточнения результата нужно произвести n независимых измерений и усреднить их результат. Следует, однако, заметить, что все равно результат будет получен с точностью, не превышающей точности самого измерительного прибора, которая обычно указывается в технической документации на него.*

Теорема 5.5. *(**Закон больших чисел в форме Бернулли**.) В независимых испытаниях Бернулли с вероятностью p появления события A в каждом для $\forall\ \varepsilon > 0$ будет:*

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1,$$

здесь m — число появлений события A в n испытаниях.

◀Представим относительную частоту $\dfrac{m}{n}$ в виде отношения $\dfrac{\xi_1 + \ldots + \xi_n}{n}$, где случайная величина $\xi_i = 1$, если в i-м испытании появилось событие A.

Таблица 5.1

ξ_i	1	0
p_i	p	$1-p$

Для случайных величин $\xi_1, \xi_2, \ \ldots, \xi_n$ выполняется следствие 5.1, т.к. $M(\xi_1) = M(\xi_2) = \ldots = p,\ D(\xi_1) = D(\xi_2) = \ldots = pq \leqslant 1$.
На основании следствия 5.1 получаем утверждение теоремы 5.5.▶

Теорема 5.5 даёт теоретическое обоснование статистическому определению вероятности, т.к. утверждает, что при большом числе независимых испытаний практически невозможны значительные отклонения относительной частоты события A от вероятности p его появления в каждом испытании.

Из законов больших чисел не следует, что при $n \to \infty$ предел последовательности случайных величин равен какому-то числу (среднему арифметическому математических ожиданий). Обычное понятие предела неприменимо к последовательности случайных величин.

Определение 5.3. *Последовательность случайных величин* $\xi_1, \xi_2, \ \ldots, \xi_n, \ldots$ *сходится по вероятности к числу* a*, если для* $\forall\ \varepsilon > 0$ *будет:*

$$\lim_{n\to\infty} P(|\xi_n - a| < \varepsilon) = 1.$$

Итак, закон больших чисел в форме Чебышева утверждает, что при выполнении определённых условий среднее арифметическое n независимых случайных величин сходится по вероятности к среднему арифметическому их математических ожиданий при $n \to \infty$.

Самостоятельно сформулируйте закон больших чисел в форме Бернулли, используя сходимость по вероятности.

5.5. Центральная предельная теорема

Известно, что нормальные случайные величины широко распространены на практике, что и объясняет их название. В чём причина этого? Ответ на этот вопрос даёт следующая теорема, доказанная русским математиком А.М. Ляпуновым.

Теорема 5.6. ***(Центральная предельная теорема.)*** *Если случайная величина* η_n *является суммой большого числа* n *независимых случайных величин, удовлетворяющих условию Ляпунова, то* η_n *имеет распределение, близкое к нормальному:*

$$\lim_{n\to\infty} P\left(\frac{\eta_n - A_n}{B_n} < x\right) = 0{,}5 + \Phi(x),$$

$$\text{где} \quad \eta_n = \xi_1 + \ldots + \xi_n, \quad A_n = M(\eta_n) = \sum_{i=1}^{n} M(\xi_i) = \sum_{i=1}^{n} a_i,$$

$$B_n^2 = D(\eta_n) = \sum_{i=1}^{n} D(\xi_i) = \sum_{i=1}^{n} b_i^2, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt.$$

Условие Ляпунова заключается в следующем:

(1) Все случайные величины $\xi_1, \xi_2, \ldots$ независимы и имеют одинаковое распределение.

(2) Все дисперсии $D(\xi_1), D(\xi_2), \ldots$ конечны и отличны от нуля.

(3) $$\lim_{n\to\infty}\frac{\sum\limits_{i=1}^{n} M\left|\xi_i - M(\xi_i)\right|^{2+\delta}}{\left(\sum_{i=1}^{n} D(\xi_i)\right)^{\frac{2+\delta}{2}}} = 0 \text{ для некоторого } \delta > 0.$$

Условия Ляпунова приводят к тому, что в сумме $\dfrac{\eta_n - A_n}{B_n}$ каждое слагаемое оказывает на сумму малое влияние. Мы примем эту теорему без доказательства.