# ЧАСТЬ 7

# ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

## Лекция 13

## ЗАКОН БОЛЬШИХ ЧИСЕЛ И ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

*ЦЕЛЬ ЛЕКЦИИ: доказать неравенство Чебышева; сформулировать и доказать закон больших чисел и его следствия; доказать центральную предельную теорему для случая суммы независимых случайных величин.*

Закон больших чисел утверждает, что среднее арифметическое большого числа случайных величин ведет себя как среднее арифметическое их математических ожиданий. А согласно центральной предельной теореме достаточно большая сумма сравнительно малых случайных величин ведет себя как нормальная случайная величина. Различные формы закона больших чисел вместе с различными вариантами центральной предельной теоремы образуют совокупность так называемых предельных теорем теории вероятностей и имеют большой практический смысл, так как составляют теоретическую основу математической статистики.

В качестве леммы, необходимой для доказательства теорем, относящихся к группе "предельных", докажем неравенство Чебышева.

### Неравенство Чебышева

Если у случайной величины $X$ известна дисперсия $D_X = \sigma_X^2$, то она в некотором смысле является мерой "случайности" величины $X$.

Так, для случайной величины, имеющей равномерный закон распределения

$$f(x) = 1/S,\ x \in [0, S],$$

дисперсия равна

$$D_X = S^2/12.$$

При малых $S$ мала и дисперсия, но невелико и отличие любого значения случайной величины от ее математического ожидания.

Аналогично для нормального распределения: чем больше дисперсия, тем больше область вероятных (имеющих отличные от нуля вероятности) значений случайной величины $X$, хотя и с меньшей вероятностью.

Таким образом, чем больше величина дисперсии $D_X = \sigma_X^2$, тем более вероятны значительные отклонения возможных значений случайной величины от центра группирования – математического ожидания $m_X$.

Если у случайной величины $X$ известна плотность распределения $f(x)$, то для любого положительного $\alpha$ можно вычислить вероятность события вида $\{|X - m_X| \geq \alpha\}$.

Однако чаще встречается вариант, когда при неизвестном законе распределения, но по известной дисперсии $\sigma_X^2$ необходимо оценить вероятность события $\{|X - m_X| \geq \alpha\}$. Эту задачу решил Чебышев Пафнутий Львович (1821–1894) посредством неравенства, названного его именем.

**Неравенство Чебышева**. Для любой случайной величины, имеющей конечную дисперсию $D_X$ и математическое ожидание $m_X$, для любого положительного $\alpha$ имеет место неравенство

$$P\{|X - m_X| \geq \alpha\} \leq {D_X}/{\alpha^2}.$$

Доказательство. Для дискретной случайной величины $X$, заданной рядом распределения

$$X\!: \quad \begin{array}{|cccc} x_1 & x_2 & \dots & x_n \\ \hline p_1 & p_2 & \dots & p_n \end{array},$$

изобразим возможные значения этой величины на числовой оси в виде точек (см. рис. 7.1).

Зададимся некоторым значением $\alpha > 0$ и вычислим вероятность того, что случайная величина $X$ отклонится от своего математического ожидания $m_X$ на величину, большую чем $\alpha$:

$$P\{|X - m_X| \geq \alpha\},$$

т. е. вероятность того, что $X$ попадет не внутрь отрезка $AB$, а вне его,

$$P\{|X - m_X| \geq \alpha\} = P\{X \not\subset AB\}.$$

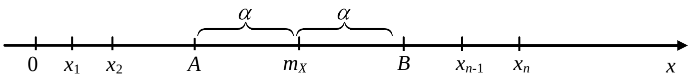


*Рис. 7.1.* Возможные значения дискретной случайной величины

Чтобы вычислить эту вероятность, необходимо просуммировать вероятности всех значений $x_i$, которые лежат вне отрезка $AB$, т. е.

$$P\{|X - m_X| \geq \alpha\} = \sum_{i:|x_i - m_X| \geq \alpha} p_i \ , \tag{7.1}$$

где запись $i:|x_i - m_X| \geq \alpha$ под знаком суммы означает, что суммирование распространяется на все значения $i$, для которых точки $x_i$ лежат вне отрезка $AB$.

По определению дисперсия дискретной случайной величины

$$D_X = \sum_{i=1}^{n} (x_i - m_X)^2 p_i = \sum_{i=1}^{n} |x_i - m_X|^2 p_i .$$

Так как все члены суммы неотрицательны, то эта сумма только уменьшится, если распространить суммирование не на все значения $x_i$, а только на те, что лежат вне отрезка $AB$:

$$D_X \geq \sum_{i:|x_i - m_X| \geq \alpha} |x_i - m_X|^2 p_i \ .$$

Так как $|x_i - m_X| \geq \alpha$, то при замене под знаком суммы величины $|x_i - m_X|$ на $\alpha$, значение этой суммы еще больше уменьшится, и будем иметь неравенство

$$D_X \geq \sum_{i:|x_i - m_X| \geq \alpha} \alpha^2 p_i = \alpha^2 \sum_{i:|x_i - m_X| \geq \alpha} p_i \ .$$

Стоящая в правой части сумма есть не что иное, как вероятность непопадания случайной величины $X$ на отрезок $AB$ (см. выражение 7.1), и поэтому

$$D_X \geq \alpha^2 P\{|X - m_X| \geq \alpha\},$$

откуда окончательно получаем

$$P\{|X - m_X| \geq \alpha\} \leq D_X / \alpha^2;$$

$$P\{|X - m_X| \leq \alpha\} \geq 1 - D_X / \alpha^2.$$

Как и всякий общий результат, не использующий данные о конкретном виде распределения случайной величины $X$, неравенство Чебышева дает лишь грубую оценку сверху для вероятности события $\{|X - m_X| \geq \alpha\}$. Если оценивать вероятность события $\{|X - m_X| \geq 3\sigma_X\}$ для случайной величины $X$ с неизвестным законом распределения, то получим по неравенству Чебышева

$$P\{|X - m_X| \geq 3\sigma_X\} \leq \sigma_X^2 / (3\sigma_X)^2 = 1/9 = 0{,}111.$$

Для нормального распределения эта вероятность равна 0,0027 – разница в 40 раз.

## Закон больших чисел

Одной из основных задач теории вероятностей является установление закономерностей, происходящих с вероятностями, близкими к единице; при этом особую роль играют закономерности, возникающие в результате наложения большого числа независимых или слабо зависимых факторов. Закон больших чисел устанавливает связь между средним арифметическим наблюдаемых значений случайной величины и ее математическим ожиданием и является одним из важнейших приложений теории вероятностей.

Предварительно решим следующую задачу. Есть случайная величина $X$, имеющая математическое ожидание $m_X$ и дисперсию $D_X$. Над этой величиной производится $n$ независимых испытаний и вычисляется среднее арифметическое всех наблюдаемых значений случайной величины $X$. Полученное значение среднего арифметического является случайной величиной. Поэтому необходимо найти числовые характеристики этого среднего арифметического, т. е. вычислить математическое ожидание и дисперсию, а также выяснить, как они изменяются с увеличением $n$.

В результате опытов получена последовательность из $n$ возможных значений: $x_1, x_2, \ldots, x_n$. Удобно посмотреть на эту совокупность чисел как на систему случайных величин $(X_1, X_2, \ldots, X_n)$. Очевидно, что эта система представляет собой $n$ независимых случайных величин, каждая из которых распределена по тому же закону, что и сама исходная величина $X$, т. е. выполняются следующие условия:

$$M[X_1] = M[X_2] = \ldots = M[X_n] = m_X;$$

$$D[X_1] = D[X_2] = \ldots = D[X_n] = D_X.$$

Среднее значение этих случайных величин

$$Y = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{7.2}$$

является случайной величиной с математическим ожиданием

$$m_Y = M[Y] = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}M\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] =$$

$$= \frac{nm_X}{n} = m_X \tag{7.3}$$

и дисперсией

$$D_Y = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}D\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i] =$$

$$= \frac{nD_X}{n^2} = \frac{D_X}{n}. \tag{7.4}$$

Получили, что математическое ожидание случайной величины $Y$ не зависит от числа испытаний $n$ и равно математическому ожиданию исследуемой случайной величины $X$, а дисперсия величины $Y$ неограниченно убывает с увеличением числа опытов и при достаточно большом $n$ может быть сделана сколь угодно малой. Таким образом, *среднее ариф-*

*метическое есть случайная величина с какой угодно малой дисперсией и при большом числе опытов ведет себя почти как неслучайная величина.*

**Теорема Чебышева**. Если $X_1, X_2, \ldots, X_n$ – последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной

$$D[X_1] \le S;\ D[X_2] \le S; \ldots ; D[X_n] \le S,$$

и математические ожидания

$$M[X_1] = M[X_2] = \ldots = M[X_n] = m_X,$$

то, каковы бы ни были постоянные $\varepsilon > 0$ и $\delta > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| < \varepsilon\right\} > 1 - \delta$$

либо

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| < \varepsilon\right\} = 1;\ \frac{1}{n}\sum_{i=1}^{n} X_i \underset{n\to\infty}{\overset{p}{\to}} m_X,$$

т. е. *среднее арифметическое последовательности независимых случайных величин* $X_1, X_2, \ldots, X_n$ *сходится по вероятности к их математическому ожиданию.*

Доказательство. Применим для случайной величины

$$Y = \frac{1}{n}\sum_{i=1}^{n} X_i \ \text{ с } \ m_Y = m_X \ \text{ и } \ D_Y = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i]$$

неравенство Чебышева

$$P\{|Y - m_Y| \ge \varepsilon\} \le \frac{D_Y}{\varepsilon^2} = \frac{\sum_{i=1}^{n} D[X_i]}{n^2 \varepsilon^2}.$$

Но из условия теоремы получаем

$$\frac{\sum_{i=1}^{n} D[X_i]}{n^2\varepsilon^2} < \frac{S}{n\varepsilon^2}.$$

Следовательно, каким бы малым ни было число $\varepsilon$, можно взять $n$ таким большим, чтобы выполнялось неравенство

$$\frac{S}{n\varepsilon^2} < \delta,$$

где $\delta$ – сколь угодно малое число.

И тогда получаем

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| \geq \varepsilon\right\} < \delta,$$

и, переходя к противоположному событию, имеем

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| < \varepsilon\right\} > 1 - \delta.$$

**Обобщенная теорема Чебышева.** Если законы распределения случайной величины $X$ от опыта к опыту изменяются, то приходится иметь дело со средним арифметическим последовательности случайных величин с различными математическими ожиданиями и с различными дисперсиями. Для таких случайных величин существует обобщенная теорема Чебышева.

Теорема (без доказательства). Если $X_1, X_2, \ldots, X_n$ – последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной

$$D[X_1] \leq S;\ D[X_2] \leq S; \ldots; D[X_n] \leq S,$$

и математические ожидания

$$M[X_1] = m_{X_1};\ M[X_2] = m_{X_2}; \ldots; M[X_n] = m_{X_n},$$

то, каковы бы ни были постоянные $\varepsilon > 0$ и $\delta > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}m_{X_i}\right|<\varepsilon\right\}>1-\delta$$

или

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow[n\to\infty]{p}\frac{1}{n}\sum_{i=1}^{n}m_{X_i},$$

т. е. *среднее арифметическое последовательности независимых случайных величин* $X_1, X_2, \ldots, X_n$ *сходится по вероятности к среднему арифметическому их математических ожиданий.*

**<u>Теорема Маркова</u>**. Закон больших чисел может быть распространен и на зависимые случайные величины. Это обобщение принадлежит Маркову.

<u>Теорема</u> (без доказательства). Если имеются зависимые случайные величины $X_1, X_2, \ldots, X_n$ и при $n\to\infty$

$$\frac{D\left[\sum_{i=1}^{n}X_i\right]}{n^2}\to 0,$$

то *среднее арифметическое наблюдаемых значений случайных величин* $X_1, X_2, \ldots, X_n$ *сходится по вероятности к среднему арифметическому их математических ожиданий:*

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow[n\to\infty]{p}\frac{1}{n}\sum_{i=1}^{n}m_{X_i}.$$

## Следствия закона больших чисел

Теорема Я. Бернулли, устанавливающая связь между частотой события и его вероятностью, может быть доказана как прямое следствие закона больших чисел (теоремы Чебышева).

**<u>Теорема Бернулли</u>**. Если производится $n$ независимых испытаний, в каждом из которых может появиться или не появиться некоторое событие $A$, вероятность появления которого в каждом опыте равна $p$, то при неограниченном увеличении числа опытов $n$ частота события $A$ сходится по вероятности к его вероятности $p$.

Обозначив частоту события $A$ через $p^*$, теорему Бернулли можно записать в виде

$$P\{|p^* - p| < \varepsilon\} > 1 - \delta \text{ или } p^* \xrightarrow[n\to\infty]{p} p,$$

где $\varepsilon$ и $\delta$ – сколь угодно малые положительные числа.

Доказательство. Рассмотрим независимые случайные величины: $X_1$ – число появлений события $A$ в первом опыте; $X_2$ – число появлений события $A$ во втором опыте; …; $X_n$ – число появлений события $A$ в $n$-м опыте. Все эти случайные величины дискретны и имеют один и тот же закон распределения в виде индикатора событий. Поэтому математическое ожидание каждой из величин $X_i$ равно $p$, а дисперсия равна $pq$, где $q = 1 - p$.

Частота $p^*$ представляет собой не что иное, как среднее арифметическое случайных величин $X_1, X_2, \ldots, X_n$

$$p^* = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

которая, согласно теореме Чебышева, сходится по вероятности к общему математическому ожиданию этих случайных величин, равному $p$.

Теорема Бернулли утверждает свойство устойчивости частот при постоянных условиях опыта, но и при изменяющихся условиях испытаний аналогичная устойчивость также существует.

**Теорема Пуассона** (следствие обобщенной теоремы Чебышева). Если производится $n$ независимых опытов и вероятность появления события $A$ в $i$-м опыте равна $p_i$, то при неограниченном увеличении числа опытов $n$ частота события $A$ сходится по вероятности к среднему арифметическому вероятностей $p_i$:

$$P\left\{\left|p^* - \frac{1}{n}\sum_{i=1}^{n} p_i\right| < \varepsilon\right\} > 1 - \delta \text{ или } p^* \xrightarrow[n\to\infty]{p} \frac{1}{n}\sum_{i=1}^{n} p_i.$$

## Центральная предельная теорема

Докажем центральную предельную теорему для случая одинаково распределенных случайных величин (в форме Ляпунова).

**<u>Теорема</u>**: *Если* $X_1, X_2, \ldots, X_n$ *– независимые случайные величины, имеющие одно и то же распределение с математическим ожиданием* $m$ *и дисперсией* $\sigma^2$, *то при увеличении* $n$ *закон распределения суммы* $Y_n = \sum_{i=1}^{n} X_i$ *неограниченно приближается к нормальному.*

<u>Доказательство</u>. Докажем теорему для случая непрерывных случайных величин, применив аппарат характеристических функций. Согласно одному из свойств, характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых. Случайные величины $X_1, X_2, \ldots, X_n$ имеют одну и ту же плотность распределения $f(x)$, а значит, и одну и ту же характеристическую функцию $v_X(t)$. Не нарушая общности, можно перенести начало отсчета всех случайных величин $X_1, X_2, \ldots, X_n$ в их общее математическое ожидание $m$, что равнозначно их центрированию и, значит, тому, что математическое ожидание каждой из них будет равно нулю.

Для доказательства теоремы найдем характеристическую функцию гауссовой случайной величины с нулевым математическим ожиданием и единичной дисперсией, плотность распределения которой

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

Характеристическая функция такой случайной величины

$$v_X(t) = \int_{-\infty}^{\infty} f(x) e^{jtx} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{jtx} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2 - 2jtx - t^2)} e^{-\frac{t^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x - jt)^2} dx =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \left| z = \frac{u}{\sqrt{2}},\ \int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi} \right| = e^{-\frac{t^2}{2}}.$$

Получили, что характеристическая функция нормальной случайной величины $X$ с $m_X = 0$ и $D_X = \sigma_X^2 = 1$ имеет вид

$$v_X(t) = e^{-\frac{t^2}{2}}. \tag{7.5}$$

По определению характеристическая функция случайной величины $X$

$$v_X(t) = \int_{-\infty}^{\infty} f(x) e^{jtx} dx. \tag{7.6}$$

Характеристическая функция случайной величины $Y_n = \sum_{i=1}^{n} X_i$ равна произведению $n$ характеристических функций слагаемых, т. е.

$$v_{Y_n} = [v_X(t)]^n. \tag{7.7}$$

Разложим $v_X(t)$ в окрестности точки $t = 0$ в ряд Макларена, ограничившись тремя членами

$$v_X(t) \approx v_X(0) + v_X'(0)t + \left[ {v_X''(0)}\big/{2} + \alpha(t) \right] t^2, \tag{7.8}$$

где $\alpha(t) \to 0$ при $t \to 0$.

Вычислим $v_X(0), v_X'(0), v_X''(0)$.

Так, $v_X(0) = \int_{-\infty}^{\infty} f(x) dx = 1$ по свойству нормировки функции $f(x)$.

Продифференцируем выражение (7.6) по $t$

$$v_X'(t) = \int_{-\infty}^{\infty} jxe^{jtx} f(x) dx = j \int_{-\infty}^{\infty} xe^{jtx} f(x) dx \tag{7.9}$$

и получаем при $t = 0$

$$v_X'(0) = j \int_{-\infty}^{\infty} xf(x) dx = jM[X],$$

а так как все $X_1, X_2, \ldots, X_n$ имеют одну и ту же плотность распределения $f(x)$ и нулевое математическое ожидание, то $v'_X(0) = 0$.

Продифференцируем теперь (7.9): $v''_X(t) = -\int\limits_{-\infty}^{\infty} x^2 e^{jtx} f(x)dx$ и соответственно при $t = 0$ получим

$$v''_X(0) = -\int\limits_{-\infty}^{\infty} x^2 f(x)dx = -\sigma^2 .$$

После подстановки в (7.8) имеем, что

$$v_X(t) = 1 - \left[\frac{\sigma^2}{2} - \alpha(t)\right] t^2. \tag{7.10}$$

Для случайной величины $Y_n$ докажем, что при увеличении $n$ ее закон распределения приближается к нормальному закону распределения. Для этого перейдем к нормированной случайной величине

$$Z_n = Y_n \Big/ (\sigma\sqrt{n}),$$

которая линейно связанной с $Y_n$ и удобна тем, что ее дисперсия равна единице для любого $n$. Если докажем, что случайная величина $Z_n$ имеет нормальное распределение, то это будет означать, что и величина $Y_n$ тоже распределена нормально.

Докажем, что характеристическая функция $v_{Z_n}$, однозначно определяющая плотность распределения случайной величины $Z_n$, приближается к характеристической функции нормального закона с теми же, что и у $Z_n$, параметрами: $m_{Z_n} = 0; \sigma_{Z_n} = 1$.

Найдем характеристическую функцию случайной величины $Z_n$, используя свойства характеристических функций и выражения (7.5) и (7.8):

$$v_{Z_n}(t) = v_{Y_n}\left( t \big/ \sigma\sqrt{n} \right) = \left[ v_{X_n}\left( t \big/ \sigma\sqrt{n} \right) \right]^n =$$

$$= \left\{ 1 - \left[ \sigma^2 \big/ 2 - \alpha\left( t \big/ \sigma\sqrt{n} \right) \right] t^2 \big/ \sigma^2 n \right\}^n .$$

Прологарифмируем это выражение и получим

$$\ln v_{Z_n}(t) = n \ln \left\{ 1 - \left[ \sigma^2 \big/ 2 - \alpha\left( t \big/ \sigma\sqrt{n} \right) \right] t^2 \big/ \sigma^2 n \right\} .$$

Пусть $\chi = \left[ \sigma^2 \big/ 2 - \alpha\left( t \big/ \sigma\sqrt{n} \right) \right] t^2 \big/ \sigma^2 n$, и тогда $\ln v_{Z_n}(t) = n \ln(1-\chi)$. Если неограниченно увеличивать $n$, то величина $\chi$ будет стремиться к нулю. Поэтому разложим $\ln(1-\chi)$ в ряд по степеням $\chi$, ограничившись первым членом разложения, т. е. $\ln(1-\chi) \approx -\chi$. Таким образом, получаем

$$\lim_{n\to\infty} \ln v_{Z_n}(t) = \lim_{n\to\infty} n(-\chi) = \lim_{n\to\infty} \left[ -t^2 \big/ 2 + \alpha\left( t \big/ \sigma\sqrt{n} \right) t^2 \big/ \sigma^2 \right] =$$

$$= -t^2 \big/ 2 + \left[ \lim_{n\to\infty} \alpha\left( t \big/ \sigma\sqrt{n} \right) \right] t^2 \big/ \sigma^2 \approx -t^2 \big/ 2 ,$$

так как функция $\alpha(x) \to 0$, когда аргумент $x = t \big/ \sigma\sqrt{n} \to 0$ при $n \to \infty$.

Получили, что $\lim\limits_{n\to\infty} \ln v_{Z_n} = -t^2 \big/ 2$, следовательно,

$$v_{Z_n} = e^{-\frac{t^2}{2}} ,$$

но это и есть характеристическая функция нормально распределенной случайной величины с нулевым математическим ожиданием и единичной дисперсией (см. выражение (7.5)). Следовательно, и линейно связанная со случайной величиной $Z_n$ случайная величина $Y_n$ имеет нормальное распределение.