

Московский государственный технический университет
имени Н. Э. Баумана



Факультет: Информатика и системы управления

Кафедра: Программное обеспечение ЭВМ и информационные технологии

Математическая статистика

Лекции

Москва, 2015 г.

Содержание

1	Предельные теоремы теории вероятностей	3
1.1	Неравенства Чебышева	3
1.2	Сходимость последовательности случайных величин	5
1.3	Закон больших чисел (ЗБЧ)	5
1.4	Центральная предельная теорема (ЦПТ)	7

1 Пределные теоремы теории вероятностей

1.1 Неравенства Чебышева

Теорема 1.1 (*первое неравенство господина Чебышева*).

$$\left. \begin{array}{l} X \geq 0; \\ \exists MX. \end{array} \right\} \Rightarrow \forall \varepsilon > 0, P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon} \quad (1)$$

- X — случайная величина;
- $P\{X \leq 0\} = 0$ так как $X \geq 0$.

Доказательство. Для непрерывной случайной величины X и зная, что при $X \geq 0 \Rightarrow f(x) = 0, x < 0$

$$MX = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x f(x) dx = \underbrace{\int_0^{\varepsilon} x f(x) dx}_{\geq 0} + \int_{\varepsilon}^{+\infty} x f(x) dx$$

учитывая $x \geq \varepsilon$

$$\underbrace{\int_0^{\varepsilon} x f(x) dx}_{\geq 0} + \int_{\varepsilon}^{+\infty} x f(x) dx \geq \int_{\varepsilon}^{+\infty} x f(x) dx \geq \varepsilon \cdot \int_{\varepsilon}^{+\infty} f(x) dx$$

где

$$\varepsilon \cdot \int_{\varepsilon}^{+\infty} f(x) dx = \varepsilon \cdot P\{X \geq \varepsilon\}$$

таким образом

$$MX \geq \varepsilon \cdot P\{X \geq \varepsilon\} \Rightarrow P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon}$$

□

Теорема 1.2 (*второе неравенство лорда Чебышева*).

$$\exists MX, \exists DX \Rightarrow \forall \varepsilon > 0, P\{|X - MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2} \quad (2)$$

- X — случайная величина.

Доказательство. Выпишем дисперсию

$$DX = M[(X - MX)^2]$$

Рассмотрим случайную величину $Y = (X - MX)^2$, где $Y \geq 0$. Тогда из *первого неравенства Чебышева* следует, что $\forall \delta \geq 0, MY \geq \delta P\{Y \geq \delta\}$, где получается, что $\delta = \varepsilon^2$.

$$\left[DX = M[(X - MX)^2] \right] \geq \left[\varepsilon^2 \cdot P\{(X - MX)^2 \geq \varepsilon^2\} = \varepsilon^2 \cdot P\{|X - MX| \geq \varepsilon\} \right]$$

таким образом

$$DX \geq \varepsilon^2 \cdot P\{|X - MX| \geq \varepsilon\} \Rightarrow P\{|X - MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}$$

□

Пример 1.1. Предельно допустимое давление в пневмосистеме ракеты равна 200 (Па). После проверки большого количества ракет было получено среднее значение давления 150 (Па). Оценить вероятность того, что давление в пневмосистеме очередной ракеты будет больше 200 (Па), если по результатам проверки ракет было получено среднеквадратичное отклонение 5 (Па).

Решение. Имеем следующее:

- случайная величина X — давление в пневмосистеме;
- $X \geq 0$;
- $MX = 150$ (Па);
- $DX = 25$ (Па);

Решим поставленную задачу с помощью *первого неравенства Чебышева*

$$\left[P\{X \geq \varepsilon\} = P\{X \geq 200\} \right] \leq \left[\frac{MX}{\varepsilon} = \frac{150}{200} = \frac{3}{4} = 0.75 \right]$$

$$P\{X \geq 200\} \leq 0.75$$

Поскольку нам известна дисперсия почему бы не воспользоваться *вторым неравенством Чебышева*? Действуем. Для начала рассмотрим вероятность следующего события

$$P\{X \geq \varepsilon\} = P\{X \geq 200\} = P\{X - \underbrace{150}_{MX} \geq \underbrace{50}_{\varepsilon}\}$$

Остаётся построить вероятность, которая будет удовлетворять форме *второго неравенства Чебышева* (т.е. сделать модуль).

$$P\{X - 150 \geq 50\} \leq P\{X - 150 \geq 50\} + P\{X - 150 \leq -50\}$$

Так как *события* $\{X - 150 \geq 50\}$ и $\{X - 150 \leq -50\}$ *несовместные*, то по *формуле сложения вероятностей несовместных событий* получаем

$$P\{X - 150 \geq 50\} + P\{X - 150 \leq -50\} =$$

$$= P\{\{X - 150 \geq 50\} + \{X - 150 \leq -50\}\} = P\{|X - 150| \geq 50\}$$

Таким образом применяем *второе неравенство Чебышева*

$$\left[P\{|X - MX| \geq \varepsilon\} = P\{|X - 150| \geq 50\} \right] \leq \left[\frac{DX}{\varepsilon^2} = \frac{25}{50^2} = \left(\frac{5}{50} \right)^2 = 0.01 \right]$$

$$P\{|X - 150| \geq 50\} \leq 0.01$$

Ответ:

- с использованием *первого неравенства Чебышева* $P \leq 0.75$;
- с использованием *второго неравенства Чебышева* $P \leq 0.01$.

Замечание. *Второе неравенство Чебышева* даёт более точную оценку, так как используется информация о дисперсии случайной величины.

Замечание. Использование *первого неравенства Чебышева* при $\varepsilon < MX$ и *второго неравенства Чебышева* при $\varepsilon < \sqrt{DX}$ даёт тривиальную оценку: $P \leq 1$.

1.2 Сходимость последовательности случайных величин

Будем считать, что X_1, \dots, X_n, \dots — последовательность случайных величин, заданных на одном вероятностном пространстве.

Определение 1.1. Последовательность случайных величин X_1, \dots, X_n, \dots *сходится по вероятности* к случайной величине Z , если $\forall \varepsilon > 0, \mathbf{P}\{|X_n - Z| \geq \varepsilon\} \xrightarrow{n \rightarrow \infty} 0$

Обозначение:

$$X_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} Z$$

Определение 1.2. Последовательность случайных величин X_1, \dots, X_n, \dots *слабо сходится* к случайной величине Z , если $\forall x \in \mathfrak{R}$ где F_Z непрерывна в точке x , числовая последовательность $F_{X_1}(x), \dots, F_{X_n}(x), \dots$ сходится к $F_Z(x)$. Обозначение:

$$F_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} F_Z(x)$$

Замечание. Данные виды сходимости *неэквивалентны*.

1.3 Закон больших чисел (ЗБЧ)

Будем считать, что

- X_1, \dots, X_n, \dots — последовательность случайных величин;
- $\exists MX_i = m_i$, где $i = \overline{1, \infty}$.

Определение 1.3. Последовательность случайных величин X_1, \dots, X_n, \dots удовлетворяет *закону больших чисел*, если

$$\forall \varepsilon > 0, \mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - \frac{1}{n} \cdot \sum_{i=1}^n m_i\right| \geq \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0$$

Замечание. Выполнение *закона больших чисел* для последовательности X_1, \dots, X_n, \dots означает, что при достаточно больших n величина

$$Y_n = \frac{1}{n} \cdot \sum_{i=1}^n X_i - \frac{1}{n} \cdot \sum_{i=1}^n m_i$$

практически теряет случайный характер.

Теорема 1.3 (*Закон больших чисел в форме Чебышева или достаточное условие выполнимости для последовательности случайных величин*). Последовательность случайных величин X_1, \dots, X_n, \dots удовлетворяет *закону больших чисел* тогда и только тогда, когда выполняются следующие условия:

- случайные величины X_1, \dots, X_n, \dots — независимы;
- $\exists MX_i = m_i, \exists DX_i = \sigma_i^2, i = 1, 2, \dots$;
- Дисперсия случайных величин ограничена в совокупности т. е.

$$\exists C > 0 : \sigma_i^2 \leq C, \quad i = 1, 2, \dots$$

Доказательство.

$$\begin{aligned}\bar{X}_n &= \frac{1}{n} \cdot \sum_{i=1}^n X_i \quad n \in N \\ M\bar{X}_n &= \frac{1}{n} \cdot \sum_{i=1}^n m_i \\ D\bar{X}_n &= \frac{1}{n^2} \cdot D\left(\sum_{i=1}^n X_i\right) = \{X_i \text{ независимы}\} = \frac{1}{n^2} \cdot \sum_{i=1}^n DX_i = \frac{1}{n^2} \cdot \sum_{i=1}^n \sigma_i^2\end{aligned}$$

Используем второе неравенство Чебышева

$$\mathbf{P}\{|\bar{X}_n - M\bar{X}_n| \geq \varepsilon\} \leq \frac{D\bar{X}_n}{\varepsilon^2}$$

В нашем случае

$$\mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - \frac{1}{n} \cdot \sum_{i=1}^n m_i\right| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \cdot \sum_{i=1}^n \sigma_i^2 \leq \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \cdot \sum_{i=1}^n C = \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \cdot n C = \frac{C}{\varepsilon^2 n}$$

таким образом

$$0 \leq \mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - \frac{1}{n} \cdot \sum_{i=1}^n m_i\right| \geq \varepsilon\right\} \leq \frac{C}{\varepsilon^2 n}$$

По теореме о двух милиционерах

$$\mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - \frac{1}{n} \cdot \sum_{i=1}^n m_i\right| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0$$

□

Следствие. Пусть

- выполняется теорема о ЗБЧ в форме Чебышева
- X_i одинаково распределены т.е. $MX_i = m$, $DX_i = \sigma^2$, $i \in \mathbb{N}$

Тогда

$$\mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - m\right| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0 \quad (3)$$

Следствие (Теорема Бернулли, ЗБЧ в форме Бернулли). Пусть

- проводится серия испытаний по *схеме Бернулли*
 - с вероятностью успеха p ;
 - с вероятностью неудачи $q = 1 - p$;
- наблюдаемая частота успеха $r_n = \{\text{число успехов в первых } n \text{ испытаниях}\} / n$.

Тогда

$$r_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} p \quad (4)$$

Доказательство. Рассмотрим случайную величину

$$X_i = \begin{cases} 1, & \text{если в } i\text{-ом испытании успех;} \\ 0, & \text{иначе.} \end{cases}$$

тогда $MX_i = p$, $DX_i = pq$. Поскольку $\exists MX_i$, $\exists DX_i$ и X_i — независимы по определению *схемы Бернулли* \Rightarrow выполняются все условия предыдущего следствия

$$P\left\{\underbrace{\left|\frac{1}{n} \cdot \sum_{i=1}^n X_i - p\right|}_{r_n} > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0$$

Таким образом

$$r_n \xrightarrow[n \rightarrow \infty]{P} p$$

□

1.4 Центральная предельная теорема (ЦПТ)

Пусть имеется:

- последовательность независимых случайных величин X_1, \dots, X_n, \dots ;
- X_i одинаково распределены;
- $MX_i = m$, $DX_i = \sigma^2$, $i \in \mathbb{N}$

Составим последовательность:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad n \in \mathbb{N}; \quad M\bar{X}_n = m; \quad D\bar{X}_n = \frac{1}{n^2} \sum_{i=1}^n DX_i = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.$$

Составим случайную величину $Y = \frac{\bar{X}_n - M\bar{X}_n}{\sqrt{D\bar{X}_n}} = \frac{1/n \cdot \sum_{i=1}^n X_i - m}{\sigma/\sqrt{n}}$. Тогда $MY_n = 0$, $DY_n = 1$.

Теорема 1.4. Пусть выполнены условия приведённые выше, тогда последовательность Y_n слабо сходится к случайной величине $Z \sim N(0, 1)$ т. е.

$$\forall x \in \mathbb{R}, \quad F_{Y_n}(x) \xrightarrow[n \rightarrow \infty]{} \Phi(x), \quad \text{где}$$

$$\bullet \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Замечание. В этом случае говорят, что *случайная величина Y_n имеет асимптотически стандартное нормально распределение.*

Замечание. Для последовательности X_i удовлетворяющей *центральной предельной теореме* также выполнены условия *закона больших чисел в форме Чебышева*. Поэтому из ЗБЧ $\Rightarrow \bar{X}_n \xrightarrow[n \rightarrow \infty]{} m$. Центральная предельная теорема уточняет характер этой сходимости.

Замечание. Центральная предельная теорема иллюстрирует особую роль *нормального распределения*. Все естественные процессы, протекание которых обусловлено многочисленными случайными факторами (независимыми) “в среднем” имеют нормальное распределение.