Лекция 7

Предельные теоремы теории вероятностей

Введение. Закон больших чисел и его значение

- Основой для **математической статистики** служит математический аппарат и выводы теории вероятностей, изучающей закономерности, происходящие в массовых, однородных случайных явлениях и процессах.
- Связующим звеном между теорией вероятностей и математической статистикой являются так называемые <u>предельные теоремы</u>, к которым относится <u>закон больших чисел.</u>

- ***Под законом больших чисел*** в теории вероятностей понимается совокупность теорем, в которых устанавливается связь между средним арифметическим достаточно большого числа случайных величин и средним арифметическим их математических ожиданий.

Введение. Закон больших чисел и его значение

- ***Закон больших чисел устанавливает условия, при которых совокупное воздействие множества факторов приводит к результату, не зависящему от случая.***

В самом общем виде закон больших чисел сформулировал П.Л. Чебышев. Большой вклад в изучение закона больших чисел внесли А.Н. Колмогоров, А.Я. Хинчин, Б.В. Гнеденко.

- К ***предельным теоремам*** относится также так называемая центральная ***предельная теорема А. Ляпунова***, устанавливающая условия, при которых сумма случайных величин будет стремиться к случайной величине с нормальным законом распределения.
- Эта теорема позволяет обосновать методы проверки статистических гипотез, корреляционно-регрессионный анализ и другие методы классической статистики.

Введение. Закон больших чисел и его значение

1) Теория вероятностей и классическая математическая статистика трактуют понятие неопределенности только с точки зрения вероятности (вероятностная неопределенность).

- Вероятность имеет место на практике (равно как и законы распределения вероятностей) только при наличии устойчивой частоты появления события, стремящейся к некоторому числу. В других случаях говорить о вероятностной неопределенности нельзя.

2) Практическое применение методов теории вероятностей и математической статистики основано на двух принципах, фактически основывающихся на предельных теоремах:

  - *Принцип невозможности наступления маловероятного события*;
  - *Принцип достаточной уверенности в наступлении события, вероятность которого близка к 1*.

3) Уже в конце XX века была известна ограниченность предельных теорем, в силу того, что выборки, имеющие место на практике - конечны.

1. Неравенство и теорема Чебышева

Рассмотрим закон больших чисел в форме Чебышева.

- ***Лемма Чебышева*** *(Маркова)*. Если случайная величина *X* принимает только неотрицательные значения и имеет математическое ожидание *M(X),* то для любого $\alpha > 0$ имеет место неравенство:

$$P(X \geq \alpha) = \frac{M(X)}{\alpha} \qquad (1)$$

- ***Неравенство Чебышева***. Если случайная величина *X* имеет математическое ожидание *M(X)* и дисперсию *D(X)*, то для любого $\varepsilon > 0$ имеет место неравенство:

$$\mathrm{P}(|\mathrm{x} - \mathrm{M}(\mathrm{X})| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2} \qquad (2)$$

- Неравенство Чебышева является в теории вероятностей общим фактом и позволяет оценить нижнюю границу вероятности.

1. Неравенство и теорема Чебышева

- Если произведено n независимых испытаний по схеме Бернулли, где p - вероятность успеха, q - вероятность неудачи, n - число опытов, k - число успехов, то для случайной величины K имеет место неравенство:

$$\mathrm{P}(|k - np| < \varepsilon) \geq 1 - \frac{npq}{\varepsilon^2} \qquad (3)$$

- Для относительной частоты появления события $\frac{k}{n}$ аналогичное неравенство имеет вид:

$$\mathrm{P}\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2} \qquad (4)$$

1. Неравенство и теорема Чебышева

- **Теорема.** *Закон больших чисел Чебышева.*

Пусть $X_1, X_2, \ldots, X_n$ –последовательность попарно независимых случайных величин, имеющих конечные математические ожидания и дисперсии, ограниченные сверху постоянной

C = const ($D(X_i) \leq$ C (i= 1, 2, ... ,n)). Тогда для любого $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| < \varepsilon\right) = 1 \qquad (5)$$

- Теорема показывает, что среднее арифметическое большого числа случайных величин с вероятностью, сколь угодно близкой к 1, будет мало отклоняться от среднего арифметического математических ожиданий.

1. Неравенство и теорема Чебышева

- **Следствие 1.** Если вероятность наступления события *A* в каждом из n независимых испытаний равна *p, k* – число наступлений события *A* в серии из *n* независимых испытаний, то каково бы ни было число $\varepsilon > 0$, имеет место предел:

$$\lim_{n\to\infty} P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 1. \qquad (6)$$

- Таким образом устанавливается связь между относительной частотой появления события *A* и постоянной вероятностью *p* в серии из *n* независимых испытаний.

- **Следствие 2**. *Теорема Пуассона.* Если в последовательности независимых испытаний вероятность появления события *A* в r-м испытании равна p_r то

$$\lim_{n\to\infty} P\left(\left|\frac{k}{n} - \frac{p_1+p_2+\cdots+p_n}{n}\right| < \varepsilon\right) = 1 \qquad (7)$$

где k – число появлений события А в серии из n испытаний.

1. Неравенство и теорема Чебышева

- **Следствие 3.** *Теорема Бернулли*.

Если $X_1, X_2, \ldots, X_n$ – последовательность независимых случайных величин, таких, что $M(X_1) = M(X_2) = \ldots = M(X_n) = a$, $D(X_1) < C$, $D(X_2) < C$, ..., $D(X_n) < C$, где $C = \text{const}$, то, каково бы ни было постоянное число $\varepsilon > 0$, имеет место, предел:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < \varepsilon\right) = 1. \qquad (8)$$

- Этот частный случай закона больших чисел позволяет обосновать правило средней арифметической.
- Законы больших чисел не позволяют уменьшить неопределенность в каждом конкретном случае, они утверждают лишь о существовании закономерности при достаточно большом числе опытов. Например, если при подбрасывании монеты 10 раз появился герб, то это не означает, что в 11-й раз появится цифра.

- **Пример 1.** Оценить с помощью неравенства Чебышева вероятность того, что при подбрасывании 12 игральных костей сумма очков *(СВ X)* отклонится от математического ожидания меньше, чем на 15. *СВ* X_i- число очков на i-й кости (i = 1, 2, ..., 12).

*Решение***.** *СВ* $X=X_1+\ldots+X_{12}$, где $M(X_1)= M(X_2)=\ldots =M(X_{12})$, $D(X_1)= D(X_2)=\ldots =D(X_{12})$.

$$M(X_1)=\frac{1}{6}(1+2+3+4+5+6)=\frac{21}{6}=\frac{7}{2}=3{,}5,$$

$$M(X_1^2)=\frac{1}{6}(1+2^2+3^2+4^2+5^2+6^2)=\frac{91}{6},$$

$$M(X)=3{,}5\cdot 12=42$$

$$D(X_1)=\frac{91}{6}-\frac{49}{4}=\frac{182-147}{12}=\frac{35}{12},$$

$$D(X)=(35/12)\cdot 12=35.$$

Согласно неравенству Чебышева имеем

$$P(|X-42|<15)\geq 1-\frac{35}{225},$$

$$P(|X-42|<15)\geq 0{,}844.$$

2. Понятие о центральной предельной теореме

- В теории вероятностей и математической статистике большое значение имеет ***центральная предельная теорема Ляпунова***, в которой утверждается, что <u>если сложить большое число случайных величин, имеющих один или различные законы распределения, то случайная величина, являющаяся результатом суммы, при некоторых условиях будет иметь нормальный закон распределения</u>.
- Примером *центральной предельной теоремы* (для последовательности независимых случайных величин) является интегральная теорема Муавра-Лапласа.

2. Понятие о центральной предельной теореме

- **Теорема 1.** Пусть производится n независимых опытов, в каждом из которых вероятность наступления события A равна p (не наступления $q = 1 - p$, $p \neq 0$, $p \neq 1$).

Если K – число появлений события A в серии из n испытаний, то при достаточно больших n СВ K можно считать нормально распределенной ($M(K)=np$, $\sigma(K) = \sqrt{D(K)} = \sqrt{npq}$):

$$\mathrm{P}(K < \text{к}) \to \mathrm{P}(\mathrm{X} < \mathrm{x}_0) = \int_{-\infty}^{x_2} \varphi(x)\, dx = \frac{1}{2} + \Phi(x_0), \quad (9)$$

где $x_0 = \frac{k-np}{\sqrt{npq}}$, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $\Phi(\mathrm{x}_0)$ – функция Лапласа.

2. Понятие о центральной предельной теореме

В более общем случае верна следующая теорема.

- **Теорема 2.** Если случайные величины $X_1, X_2, \ldots, X_n$ независимы, одинаково распределены и имеют конечную дисперсию, то при $n \to \infty$:

$$P\left(\frac{X_1+X_2+\cdots+X_n-na}{\sigma\sqrt{n}}<t\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{u^2}{2}}du, \qquad (10)$$

где $M(X_i)=a$, $\sigma^2=D(X_i)$; U – нормально распределенная случайная величина, $M(U)=0$, $D(U) = 1$.

Пример 2. На отрезке [0; 1] случайным образом выбрано 100 чисел, точнее рассматриваются 100 независимых средних X_1, X_2,... X_n, равномерно распределенных на отрезке [0; 1]. Найти вероятность того, что их сумма заключена между 51 и 60, т.е. $P(51 < \sum X_i < 60)$.

*Решение***.** В силу теоремы 2: $\frac{\sum X_i - na}{\sigma\sqrt{n}} = U,\quad \sum(X_i) - na = \sigma\sqrt{n}U.$

Из условия, в силу равномерности *СВ* X_i, следует, что

$$M(X_i) = \frac{1+0}{2} = \frac{1}{2}, \sigma^2 = \frac{(1-0)^2}{12}.$$

Имеем, $M(\sum X_i) = M(na + \sigma\sqrt{n}U) = na = 100\frac{1}{2} = 50,$

$$D\left(\sum X_i\right) = D(na + \sigma\sqrt{n}U) = na = 100\frac{1}{12} = \frac{100}{12}.$$

Пример 2. На отрезке [0; 1] случайным образом выбрано 100 чисел, точнее рассматриваются 100 независимых средних X_1, X_2,... X_n, равномерно распределенных на отрезке [0; 1]. Найти вероятность того, что их сумма заключена между 51 и 60, т.е. $P(51 < \sum X_i < 60)$.

*Решение***.**

Итак, $\sum X_i \in N(na, n\sigma^2)$ – сумма, нормально распределенная случайная величина с математическим ожиданием na=50 и дисперсией $n\sigma^2$=100/12. Отсюда,

$$P(51 \le \sum X_i \le 60) = \Phi\left(\frac{60-na}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{51-na}{\sqrt{n}\sigma}\right) = \Phi\left(\frac{60-50}{\frac{10}{\sqrt{12}}}\right) - \Phi\left(\frac{51-50}{\frac{10}{\sqrt{12}}}\right) = \Phi(\sqrt{12}) -$$

$$- \Phi\left(\frac{\sqrt{12}}{10}\right) = \Phi(3{,}464) - \Phi(0{,}3464) \approx 0{,}49971 - 0{,}1353 = 0{,}3644.$$

То есть вероятность того, что сумма 100 независимых средних X_1, X_2,..., X_n, равномерно распределенных на отрезке [0; 1], заключена между 51 и 60 и равна 0,3644.