Лекция 7
Предельные теоремы теории вероятностей

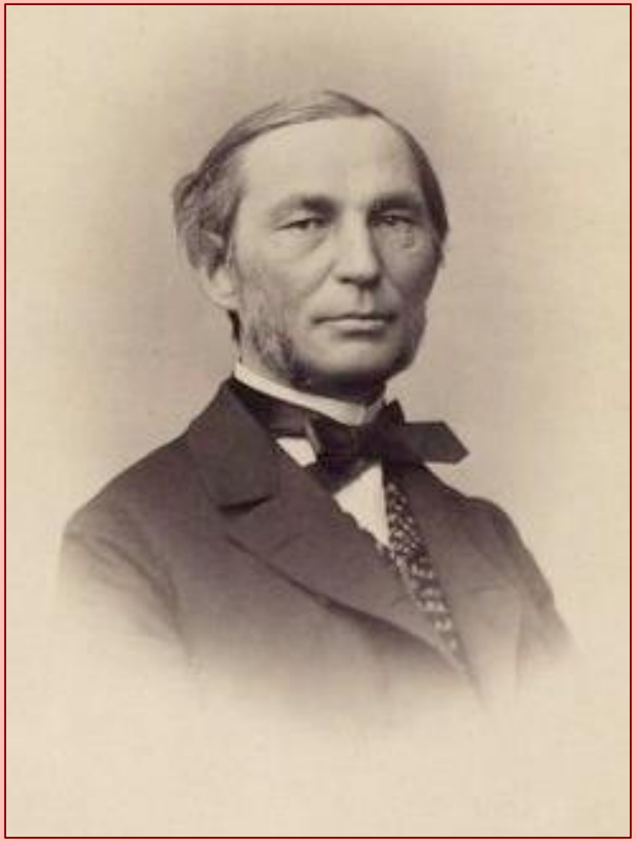
1. Неравенство Чебышёва. Сходимость по вероятности.
2. Закон больших чисел.
3. Центральная предельная теорема.

ITMO University

Предельные теоремы теории вероятностей – группа теорем, которые устанавливают:

- соответствие между теоретическими и экспериментальными характеристиками случайных величин при большом числе экспериментов (закон больших чисел)
- законы распределения случайных величин при большом числе экспериментов (центральная предельная теорема)

1. Неравенство Чебышёва. Сходимость по вероятности



Пафнутий Львович Чебышёв
1821-1894

Неравенство Чебышёва объясняет вероятностный смысл дисперсии.

Лемма

Для любой с.в. X с математическим ожиданием m_x и дисперсией D_x, и $\forall \varepsilon > 0$ имеет место неравенство:

$$P(|X - m_x| \geq \varepsilon) \leq \frac{D_x}{\varepsilon^2} \qquad (1)$$

(1) – “оценка сверху” вероятности больших отклонений с.в. от ее математического ожидания.

$$P(|X - m_x| < \varepsilon) \geq 1 - \frac{D_x}{\varepsilon^2} \qquad (2)$$

(2) – “оценка снизу” вероятности небольших отклонений с.в. от ее математического ожидания.

Пусть (Ω, Σ, P) – вероятностное пространство,

$\{X_n\}_{n=1}^{\infty} = X_1, X_2, \ldots, X_n, \ldots$ – последовательность случайных величин, заданных на (Ω, Σ, P),

X – случайная величина, заданная на (Ω, Σ, P)

Определение 1

Последовательность $\{X_n\}_{n=1}^{\infty}$ сходится по вероятности к случайной величине X, если

$$\forall \varepsilon > 0 \ \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

Обозначение: $X_n \xrightarrow[n\to\infty]{p} X$

Свойства сходимости по вероятности

Пусть имеем две последовательности с.в.:

$$X_n \xrightarrow[n\to\infty]{p} X, \quad Y_n \xrightarrow[n\to\infty]{p} Y$$

1. $X_n + Y_n \xrightarrow[n\to\infty]{p} X + Y$

2. $X_n \cdot Y_n \xrightarrow[n\to\infty]{p} X \cdot Y$

3. $g(x)$ – непрерывная функция $\Rightarrow g(X_n) \xrightarrow[n\to\infty]{p} g(X)$

Сходимость по вероятности к постоянной $\boldsymbol{c}$ последовательности случайных величин $\{X_n\}_{n=1}^{\infty}$ означает, что

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} P(|X - \boldsymbol{c}| < \varepsilon) = 1$$

Обозначение: $X_n \xrightarrow[n\to\infty]{p} \boldsymbol{c}$

2. Закон больших чисел

Закон больших чисел – группа теорем, в каждой из которых для тех или иных условий устанавливается факт приближения средних характеристик случайных величин, при большом числе экспериментов, к определенным *постоянным, неслучайным величинам*.

Теорема (закон больших чисел в форме Чебышёва)

Пусть $\{X_n\}_{n=1}^{\infty}$ – последовательность попарно независимых случайных величин с математическими ожиданиями $M[X_i] = m_i$ и ограниченными в совокупности дисперсиями, т. е. $D[X_i] \le \boldsymbol{c} < +\infty$.

Тогда среднее арифметическое первых n величин последовательности, сходится по вероятности к среднему арифметическому их математических ожиданий:

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\left(\left| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} m_i \right| \ge \varepsilon\right) = 0$$

СР Докажите

Следствие

Если с.в. X_i , $i = 1, 2, \ldots$ также одинаково распределены, т.е. $\forall i \;\; m_i = m$ и $\sigma_i^2 = \sigma^2$, то последовательность $\{X_n\}_{n=1}^{\infty}$ случайных величин удовлетворяет закону больших чисел:

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

Схема Бернулли

- n независимых экспериментов
- в каждом эксперименте возможны только два исхода – появилось A (*успех*) или $\overline{A}$ (*неудача*)
- вероятность A в каждом эксперименте одна и та же и равна p.

Теорема (закон больших чисел в форме Бернулли)

При неограниченном увеличении числа экспериментов относительная частота $\frac{m^*}{n^*}$ события A в схеме Бернулли сходится по вероятности к вероятности p события A в отдельном эксперименте:

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} P\left(\left| \frac{m^*}{n^*} - p \right| \geq \varepsilon\right) = 0$$

X_i – число появлений A в i –том эксперименте: $X_i \sim B(p)$

x_i	0	1
p_i	q	p

$M[X_i] = p$

$D[X_i] = pq = const$

Относительная частота появления A есть среднее арифметическое $\frac{1}{n}\sum_{i=1}^{n} X_i$

Среднее арифметическое математических ожиданий $\frac{np}{n} = p$

Т. Чебышёва $\Rightarrow$

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - p \right| \geq \varepsilon \right) = 0$$

То есть выполняется

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\left(\left| \frac{m^*}{n^*} - p \right| \geq \varepsilon \right) = 0$$

Теорема Бернулли утверждает, что при большом числе испытаний относительная частота события обладает *свойством устойчивости* и обосновывает *статистическое определение вероятности*.

3. Центральная предельная теорема

Центральная предельная теорема – это группа теорем, в которых устанавливаются условия, при которых предельный закон распределения суммы случайных величин является *нормальным*.

Обозначим: $Y_n = \sum_{i=1}^{n} X_i$ – суммарное значение первых n случайных величин.

Говорят, что для последовательности $\{X_n\}_{n=1}^{\infty}$ выполняется *центральная предельная теорема* (ЦПТ), если

$$\frac{Y_n - M[Y_n]}{\sqrt{D[Y_n]}}$$

сходится по распределению к *стандартной нормальной* величине, т.е. сумма Y_n *асимптотически нормальна* с параметрами $M[Y_n]$ и $D[Y_n]$:

$$Y_n \sim N(M[Y_n], D[Y_n])$$

Теорема (центральная предельная теорема)

Пусть $\{X_n\}_{n=1}^{\infty}$ – последовательность

- независимых
- одинаково распределенных случайных величин

с математическими ожиданиями $M[X_i]=m$ и дисперсиями $D[X_i]=\sigma^2 < +\infty$.

Тогда при $n\to\infty$ закон распределения с.в. Y_n неограниченно приближается к нормальному с параметрами nm и $\sigma\sqrt{n}$, т.е.

$$P(\alpha \le Y_n \le \beta) \approx \Phi_0\left(\frac{\beta - nm}{\sigma\sqrt{n}}\right) - \Phi_0\left(\frac{\alpha - nm}{\sigma\sqrt{n}}\right)$$

Другими словами, функция распределения центрированной с.в.

$$Z_n = \frac{Y_n - nm}{\sigma\sqrt{n}}$$

при $n \to \infty$ неограниченно приближается к функции распределения $\Phi(x)$ стандартного нормального закона с математическим ожиданием $a=0$ и дисперсией $\sigma^2=1$:

$$P(Z_n < x) \xrightarrow[n\to\infty]{} \Phi(x)$$

Если случайные величины X_i – число появлений события A в i-том эксперименте, то каждая из них принимает только два значения: 0 и 1.

Тогда получаем частный случай ЦПТ.

Теорема (интегральная Муавра-Лапласа)

Пусть вероятность p события A в каждом из n независимых экспериментов постоянна и $p \neq 0$, $p \neq 1$.

Тогда вероятность того, что событие A появится в n экспериментах от m_1 до m_2 раз при $n \to \infty$ приближенно равна

$$P_n(m_1, m_2) \approx \Phi_0(x_2) - \Phi_0(x_1)$$

где $x_1 = \dfrac{m_1 - np}{\sqrt{npq}}, \quad x_2 = \dfrac{m_2 - np}{\sqrt{npq}}.$

Y – число появлений события A в n независимых экспериментах:

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

Закон распределения с.в. X_i :

x_i	0	1
p_i	q	p

$$M[X_i] = p, \quad D[X_i] = p - p^2 = pq.$$

Тогда при $n \to \infty$ закон распределения с.в. Y неограниченно приближается к нормальному и

$$P(m_1 < Y < m_2) = \Phi_0\left(\frac{m_2 - m_y}{\sigma_y}\right) - \Phi_0\left(\frac{m_1 - m_y}{\sigma_y}\right)$$

Вычислим:

$$m_y = M[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} M[X_i] = \sum_{i=1}^{n} p = np$$

$$D_y = D[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} D[X_i] = \sum_{i=1}^{n} pq = npq$$

$$\sigma_y = \sqrt{npq}$$

Вывод: $Y \sim Bin\,(n,p)$, $p \in (0,\,1)$, $n \in \boldsymbol{N}$.

Тогда

$$P(m_1 < Y < m_2) = \Phi_0\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi_0\left(\frac{m_1 - np}{\sqrt{npq}}\right) \quad (3)$$

$$P_n(m_1, m_2) \approx \Phi_0(x_2) - \Phi_0(x_1)$$

Интегральная Муавра-Лапласа обосновывает факт замены дискретного *биномиального распределения* непрерывным *нормальным распределением* при неограниченном увеличении числа экспериментов.

Если при данных n и p выполняются условия:

$$np - 3\sqrt{npq} > 0$$

$$np + 3\sqrt{npq} < n$$

то вероятности можно вычислять по (3)

Пример *

Станок с ЧПУ выдает за смену 1000 изделий , из которых 2% дефектных.

Найдите вероятность того, что за смену будет изготовлено не менее 970 доброкачественных изделий, если изделия оказываются доброкачественными независимо друг от друга.

Теорема (локальная Муавра-Лапласа)

Пусть вероятность p события A в каждом из n независимых экспериментов постоянна и $p \neq 0, p \neq 1$.

Тогда вероятность того, что событие A появится в n экспериментах ровно m раз при $n \to \infty$ приближенно равна

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi(x)$$

где $x = \dfrac{m - np}{\sqrt{npq}}$.

Пример

*

Найдите вероятность того, что событие A наступит ровно 80 раз в 400 экспериментах, если вероятность появления этого события в одном эксперименте равна 0,2.