Закон больших чисел

Неравенство Чебышева. Для любой случайной величины X с математическим ожиданием m_X и дисперсией D_X выполняют следующее неравенство:

$$p(|X - m_X| \geq \varepsilon) \leq \frac{D_X}{\varepsilon^2} \qquad (12.1)$$

где $\varepsilon > 0$.

Доказательство. Рассмотрим вероятность $p(|X| \geq \varepsilon)$

$$p(|X| \geq \varepsilon) = \int\limits_{|x| \geq \varepsilon} f(x)dx = \int\limits_{|x| \geq \varepsilon} \frac{x^2}{\varepsilon^2} \cdot \frac{\varepsilon^2}{x^2} f(x)dx \leq \int\limits_{|x| \geq \varepsilon} \frac{x^2}{\varepsilon^2} f(x)dx = \frac{1}{\varepsilon^2} \int\limits_{|x| \geq \varepsilon} x^2 f(x)dx \leq$$

$$\leq \frac{1}{\varepsilon^2} \int\limits_{-\infty}^{+\infty} x^2 f(x)dx = \frac{M[X^2]}{\varepsilon^2}.$$

Таким образом, $$p(|X| \geq \varepsilon) \leq \frac{M[X^2]}{\varepsilon^2}$$

Заменив нецентрированную величину *X* на центрированную $\overset{\circ}{X} = X - m_X$

получим

$$p(|X - m_X| \geq \varepsilon) \leq \frac{M[(X - m_X)^2]}{\varepsilon^2} = \frac{D_X}{\varepsilon^2}$$

Пример. Определим вероятность, что случайная величина примет значение за пределами интервала $3\sigma_X$. Полагаем в неравенстве Чебышева $\varepsilon = 3\sigma_X$

имеем:

$$p(|X - m_X| \geq 3\sigma_X) \leq \frac{D_X}{9\sigma_X^2} = \frac{1}{9}$$

Теорема Чебышева.

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{p} m_X \qquad (12.2)$$

Доказательство. Рассмотрим величину $Y = \frac{1}{n}\sum_{i=1}^{n} X_i$

Определим числовые характеристики Y (см. (11.5), (11.7)):

$$m_Y = M[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = \frac{1}{n} n m_X = m_X;$$

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

Запишем неравенство Чебышева для величины Y:

$$p(\left|Y - m_Y\right| \ge \varepsilon) = p(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| \ge \varepsilon) \le \frac{D_Y}{\varepsilon^2} = \frac{D_Y}{\varepsilon^2 n}.$$

$p(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| \ge \varepsilon) < \delta$. Переходя к противоположному событию

$p(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m_X\right| < \varepsilon) > 1 - \delta$ т.е. Y сходится по вероятности к m_X.

Теорема Бернулли.

$$p^*(A) \xrightarrow[n \to \infty]{p} p(A) \qquad (12.3)$$

где $p^*(A) = \dfrac{m}{n}$ - частота события A в n опытах;

m - число опытов в которых произошло событие A;

n - число проведенных опытов.

Пусть случайная величина X – индикатор события A:

$$X = \begin{cases} 1, & A \\ 0, & \overline{A} \end{cases}$$

тогда X_i – индикатор события A в i-м опыте.

Числовые характеристики индикатора X случайного события (см. (6.1)):

$$m_X = p, D_X = qp$$

где $q = 1 - p$ - вероятность осуществления

□A.

Применим теорему Чебышева:

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{m}{n} = p^*(A) \underset{n\to\infty}{\overset{p}{\rightarrow}} m_X = p = p(A)$$

Центральная предельная теорема

$$m_Y = n \cdot m, \;\; \sigma_Y = \sigma\sqrt{n} \qquad (12.4)$$

Доказательство.

Согласно свойству (7.16) характеристическая функция суммы равна произведению характеристических функций слагаемых:

$$\upsilon_Y(t) = \left[\upsilon_X(t)\right]^n \qquad (12.5)$$

Разложим функцию $\upsilon_X(t)$ в окрестности точки $t = 0$ в ряд Маклорена с тремя членами:

$$\upsilon_X(t) = \upsilon_X(0) + \upsilon_X'(0)t + [\upsilon_X''(0)/2 + \alpha(t)]t^2 \tag{12.6}$$

где производные берутся по t; $\alpha(t) \to 0$ при $t \to 0$.

Используя свойство (7.15) характеристических функций определим значения

$$\upsilon_X(0) = \upsilon_X^{(0)}(0) = \alpha_0(x)i^0 = 1$$

$$\upsilon_X'(0) = \upsilon_X^{(1)}(0) = \alpha_1(x)i^1 = m \cdot i = 0$$

$$\upsilon_X''(0) = \upsilon_X^{(2)}(0) = \alpha_2(x)i^2 = \sigma^2 \cdot i^2 = -\sigma^2$$

Подставив их в (12.5) получим

$$\upsilon_X(t) = 1 - [\sigma^2/2 - \alpha(t)]t^2 \tag{12.7}$$

Найдем характеристическую функцию случайной величины *Z*. Из свойства (7.14) характеристической функции имеем:

$$\upsilon_Z(t) = \upsilon_Y\left(\frac{t}{\sigma\sqrt{n}}\right) \tag{12.8}$$

Подставив (12.5) и (12.7) в (12.8) получим

$$\upsilon_Z(t) = \left\{1 - \left[\frac{\sigma^2}{2} - \alpha(t/\sigma\sqrt{n})\right] t^2/(n\sigma^2)\right\}^n \tag{12.9}$$

Пронологарифмируем это выражение:

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

Введем обозначение $\left[\frac{\sigma^2}{2} - \alpha(t/\sigma\sqrt{n})\right] t^2/(n\sigma^2) = \varepsilon$,тогда $\ln\upsilon_Z(t) = n\ln(1-\varepsilon)$

Тогда

$$\lim_{n\to\infty} \ln \upsilon_Z(t) = \lim_{n\to\infty} n \cdot (-\varepsilon) = \lim_{n\to\infty} \left\{ -\frac{t^2}{2} + \alpha(t/\sigma\sqrt{n})t^2/\sigma^2 \right\} =$$

$$= -\frac{t^2}{2} + \lim_{n\to\infty} \alpha(t/\sigma\sqrt{n})t^2/\sigma^2 = -\frac{t^2}{2}.$$

Теорема.

$$m_Y = \sum_{i=1}^{n} m_i, \ \sigma_Y = \sqrt{\sum_{i=1}^{n} D_i} \qquad (12.10)$$

Пример 1.

$$Y = \sigma\left(\sum_{i=1}^{12} x_i - 6\right) + m \qquad (12.11)$$

Пример 2.

$$0 < np - 3\sqrt{npq}, np + 3\sqrt{npq} < n \qquad (12.12)$$

Пусть X_i – индикатор события A в i-м опыте, тогда число появлений событий A в n опытах равно $Y = \sum_{i=1}^{n} X_i$, причем $0 \le Y \le n.$

На основании центральной предельной теоремы величина Y будет иметь практически нормальный закон распределения с параметрами

$$m_Y = \sum_{i=1}^{n} m_i = np, \ \sigma_Y = \sqrt{\sum_{i=1}^{n} D_i} = \sqrt{npq}.$$