21. Центральная предельная теорема: формулировка, обсуждение, примеры применения. Теорема Муавра-Лапласа

Status	Completed

Центральная предельная теорема

Для любой последовательности $X_1, X_2, ...$ независимых и одинаково распределенных случайных величин с конечной и ненулевой дисперсией $0 < DX_1 < \infty$ имеет место сходимость:

S_n - сумма $X_1 + ... + X_n$.

Замечания:

1. К S_n просто применили операцию стандартизации, так как $ES_n = nEX_1; DS_n = nDX_1$.
2. Интервальная формулировка

3. Можно сформулировать ЦПТ в эквивалентной форме: для любых $A \le B$

$$\mathbf{P}\left(A \le \frac{S_n - na}{\sigma\sqrt{n}} \le B\right) \to \frac{1}{\sqrt{2\pi}} \int\limits_A^B e^{-t^2/2}\,dt.$$

Именно такая форма чаще всего используется при решении задач. Делается это следующим образом. Предположим, что нам необходимо найти вероятность $\mathbf{P}(C \le S_n \le D)$ при больших значениях n. Первое, что мы должны сделать, это подогнать наше выражение под формулировку теоремы:

$$\mathbf{P}(C \le S_n \le D) = \mathbf{P}\left(\frac{C - na}{\sigma\sqrt{n}} \le \frac{S_n - na}{\sigma\sqrt{n}} \le \frac{D - na}{\sigma\sqrt{n}}\right),$$

после чего объявляем эту вероятность почти равной

$$\frac{1}{\sqrt{2\pi}} \int\limits_A^B e^{-t^2/2}\,dt = \Phi_{0,1}(B) - \Phi_{0,1}(A),$$

где

$$A = \frac{C - na}{\sigma\sqrt{n}}, \quad B = \frac{D - na}{\sigma\sqrt{n}}.$$

Численные значения функции $\Phi_{0,1}(y)$ обычно находятся из таблиц.

Неравенство Берри-Эссена (вопрос на 5)

Пусть в условиях ЦПТ еще $E|X|^3 < \infty$, тогда

$$\sup_{y \in \mathbb{R}} \left| P\left(\frac{S_n - nEX_1}{\sqrt{nDX_1}} < y\right) - \Phi_{0,1}(y) \right| \le \frac{C}{\sqrt{n}} \frac{E|X_1 - EX_1|^3}{(DX_1)^{3/2}}$$

где C - некая константа, значение которой продолжает уточняться (сейчас $C < 0.4784$).

Примеры применения

Пример 9.2. 1000 раз бросается игральная кость. Найти пределы, в которых с вероятностью 0,95 будет лежать сумма выпавших очков.

Решение. Обозначим через S_n сумму выпавших очков. S_n есть сумма независимых случайных величин, каждая из которых принимает значения от 1 до 6 с равными вероятностями. Нетрудно вычислить: $a = \mathbf{E}X_1 = 3,5$; $\mathbf{E}X_1^2 = 91/6$;

$\sigma^2 = \mathbf{D}X_1 = 35/12$. В силу ЦПТ случайная величина $(S_n - 3500)\,/\sqrt{1000 \cdot 35/12}$ имеет почти стандартное нормальное распределение (число n велико!), поэтому

$$\mathbf{P}\left(-1,96 < \frac{S_n - 3500}{\sqrt{1000 \cdot 35/12}} < 1,96\right) \simeq \frac{1}{\sqrt{2\pi}} \int\limits_{-1,96}^{1,96} e^{-t^2/2}\,dt = 0,95.$$

Последнее мы заранее находим из таблиц. Таким образом,

$$\mathbf{P}(|S_n - 3500)| < 1,96\sqrt{1000 \cdot 35/12}) \simeq 0,95,$$

$$1,96\sqrt{1000 \cdot 35/12} = 105,85\ldots.$$

Считаем так (для тех, у кого в голове интерпретатор питона есть):

```
import scipy.stats as s

n = 1000
ex = 3.5
dx = 35/12
q = s.norm.ppf((0.95+1)/2)

print(f'q = {q:.3}')
right = +q*sqrt(n*dx) + ex*n
left = -q*sqrt(n*dx) + ex*n
print(f'{left:.7}, {right:.7}')

# Вывод:
# q = 1.96
# 3394.15, 3605.85
```

Теорема Муавра-Лапласа

Теорема 42 (предельная теорема Муавра—Лапласа). *Пусть событие A может произойти в любом из n независимых испытаний с одной и той же вероятностью p и пусть $\nu_n(A)$ — число осуществлений события A в n испытаниях. Тогда*

$$\frac{\nu_n(A) - np}{\sqrt{np(1-p)}} \Rightarrow \mathrm{N}_{0,1} \;\; при \;\; n \to \infty,$$

Доказательство: просто частный случай ЦПТ для распределения Бернулли.