Тема 13. Предельные теоремы

Закон больших чисел

Одной из важных проблем теории вероятностей является изучение поведения при $n \to \infty$ распределения суммы $S_n = X_1 + \ldots + X_n$ независимых и одинаково распределённых слагаемых. Сложность проблемы связана с невозможностью явного вычисления n-кратных свёрток для произвольной плотности $f_X(x)$. Если слагаемые положительны, то интуитивно понятно, что $S_n \to \infty$. Хотелось бы как-то сдержать рост суммы S_n, чтобы в некотором смысле наступила стабилизация. С какой скоростью растёт сумма S_n? Очевидно, что она растёт в среднем как $n\mathsf{M}X$. Для обеспечения стабилизации разделим сумму S_n на n. Получим приближённое равенство $\overline{X}_n \equiv S_n/n \approx \mathsf{M}X$. Как сформулировать строгое утверждение? Ответ даёт классическая теорема, которая называется

Закон больших чисел. Пусть $X_1, X_2, \ldots$ — независимые и одинаково распределённые случайные величины с математическим ожиданием $\mu = \mathsf{M}X$. Тогда для любого действительного числа $\varepsilon > 0$

$$\mathbf{P}(|\overline{X}_n - \mu| > \varepsilon) \to 0 \quad \text{при } n \to \infty. \tag{1}$$

Иначе говоря, распределение случайной величины $\overline{X}_n$ с ростом n «стягивается» к константе $\mu = \mathsf{M}X$. Утверждение (1) называется *сходимостью по вероятности* последовательности $\overline{X}_n$ к константе μ.

Доказательство. Ради простоты установим сходимость (1) при избыточном допущении $\sigma^2 = \mathsf{D}X < \infty$. Применим неравенство Чебышёва (см. формулу (5) из темы 7):

$$\mathsf{P}(|\overline{X}_n - \mu| > \varepsilon) \le \frac{\mathsf{D}(\overline{X}_n - \mu)}{\varepsilon^2} = \frac{\mathsf{D}\overline{X}_n}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \to 0 \quad \text{при } n \to \infty.$$

Предположим, что проводятся многократные независимые измерения X_i в одних и тех же условиях некоторого показателя μ со случайными ошибками E_i, т. е. $X_i = \mu + E_i$. Ошибки обычно можно рассматривать как независимые и одинаково распределённые с $\mathsf{M}E_i = 0$. По закону больших чисел при увеличении числа измерений n погрешность среднего арифметического $\Delta_n = |\overline{X}_n - \mu|$ будет по вероятности стремиться к нулю («*семь раз отмерь, один — отрежь*»).

Центральная предельная теорема

Какова типичная величина абсолютной погрешности Δ_n, если число измерений n велико? Оказывается, Δ_n имеет тот же порядок малости, что и стандартное отклонение $\sqrt{\mathsf{D}\overline{X}_n}$, т. е. $const/\sqrt{n}$. Более точный ответ даёт

Центральная предельная теорема. Пусть $X_1, X_2, \ldots$ — независимые и одинаково распределённые случайные величины с математическим ожиданием $\mu = \mathsf{M}X$ и дисперсией $0 < \sigma^2 = \mathsf{D}X < \infty$. Тогда

$$\mathbf{P}\left(\frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu) \le x\right) \to \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \quad \text{при } n \to \infty. \tag{2}$$

Функция $\Phi(x)$ называется функцией распределения *стандартного нормального закона*. Доказано, что эту функцию нельзя выразить через элементарные функции. Её производная

$$\varphi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

имеет форму «колокола». Величины некоторых площадей под графиком $\varphi(x)$ показаны на рис. 1.

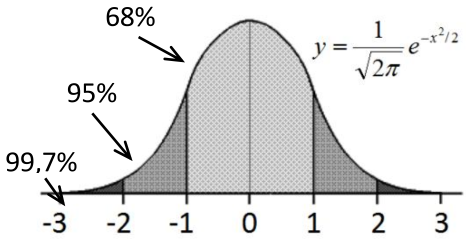

Рис. 1

Видим, что $\Phi(3) - \Phi(-3) \approx 0{,}997$, поэтому *стандартную нормальную случайную величину* Z, имеющую $\Phi(x)$ в качестве функции распределения, можно считать практически ограниченной.

Поскольку $\Phi(2) - \Phi(-2) \approx 0{,}95$, можно утверждать, что для достаточно больших n погрешность $\Delta_n = |\overline{X}_n - \mu|$ не будет превышать $2\sigma/\sqrt{n}$ с вероятностью около 0,95. Действительно, в силу центральной предельной теоремы имеем:

$$\mathbf{P}\left(\Delta_n \le \frac{2\sigma}{\sqrt{n}}\right) = \mathbf{P}\left(|\overline{X}_n - \mu| \le \frac{2\sigma}{\sqrt{n}}\right) = \mathbf{P}\left(-2 \le \frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu) \le 2\right) \to \Phi(2) - \Phi(-2) \approx 0{,}95.$$

Этот результат называют **«*правилом двух сигм*»**.

Пример 1. Рассмотрим эксперимент с бросанием монетки: $X_1, X_2, \ldots$ — это бернуллиевские случайные величины с параметром $p = 0{,}5$. Найдём вероятность, что при n = 100 бросаниях монеты частота выпадений «герба» $\overline{X}_n$ будет отличаться от 0,5 не больше, чем на 0,1. В данном случае $\sigma^2 = \mathrm{D}X = p(1-p) = 0{,}25$. Отсюда имеем равенство $2\sigma/\sqrt{n} = 0{,}1$. Поэтому согласно центральной предельной теореме искомая вероятность приближённо равна 0,95.

Вычислим с помощью Excel точное значение. Заметим, что событие $\{|\overline{X}_{100} - 0{,}5| \le 0{,}1\}$ совпадает с событием $\{40 \le X_1 + \ldots + X_{100} \le 60\}$. Сумма независимых бернуллиевских случайных величин $X_1, \ldots, X_n$ имеет биномиальное распределение с параметрами n и p. Биномиальные вероятности для 0 ≤ i ≤ 100 вычисляются с помощью функции БИНОМ.РАСП (BINOM.DIST) с параметром «Интегральная» («Cumulative») = 0. На рис. 2 приведена столбиковая диаграмма, построенная по этим вероятностям.

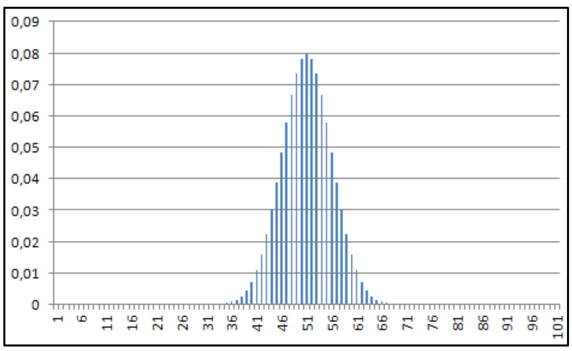

Рис. 2

Выделив блок вероятностей от 40 до 60, видим, что их сумма равна 0,965, т. е. отличается от 0,95 примерно на 0,015.

Из рис. 2 видно, что столбики практически «зануляются» при i < 35 и при i > 65. Аналогично, выделив блок вероятностей от 35 до 65, находим, что их сумма равна 0,998. Центральная предельная теорема даёт для этой вероятности приближение 0,997. Отличие составляет всего 0,001.

Теорема Пуассона (закон редких событий)

Напомним условия примера 1 из темы 5: «среди m экзаменационных билетов l — лёгкие, n студентов по очереди наудачу с возвращением берут билеты». Пусть S_n — общее число лёгких билетов, вынутых всеми n студентами. В примере 1 из темы 5 было установлено, что случайная величина S_n имеет биномиальное распределение: $\mathbf{P}(S_n = i) = C_n^i p^i (1-p)^{n-i}$, где $i = 0, 1, \ldots, n;$ $p = l/m$ — доля лёгких билетов.

Как можно приближённо вычислить биномиальные вероятности $\mathbf{P}(S_n = i)$, если параметр p очень мал (скажем, $p < 0{,}01$), а параметр n, напротив, довольно велик (скажем, $n > 100$)? Проблема заключается в том, что в таком случае в биномиальном коэффициенте C_n^i будут присутствовать факториалы больших чисел, а величины p^i будут крайне малы даже при умеренных значениях i.

Пусть $\lambda = pn$. Предположим, что число λ не слишком велико (скажем, $\lambda < 10$). Тогда при фиксированном значении i для вероятности $\mathbf{P}(S_n = i)$ можно использовать приближение

$$\mathbf{P}(S_n = i) \approx \frac{\lambda^i}{i!} e^{-\lambda},$$

теоретическим обоснованием которого служит

***Теорема Пуассона*.** Если целое $i \geq 0$ фиксировано, $n \to \infty,$ $p \to 0,$ $\lambda = pn < \infty,$ то

$$\mathbf{P}(S_n = i) \to \frac{\lambda^i}{i!} e^{-\lambda}.$$

Доказательство.

$$\mathbf{P}(S_n = i) = \frac{n(n-1)\ldots(n-i+1)}{i!} p^i (1-p)^{n-i} = \frac{(pn)^i}{i!}(1-p)^n \left[\frac{n(n-1)\ldots(n-i+1)}{n^i}(1-p)^{-i}\right],$$

где $(1-p)^n = (1-\lambda/n)^n \to e^{-\lambda}$ по определению экспоненты, а выражение в квадратных скобках равно $\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{i-1}{n}\right)(1-p)^{-i}$ и стремится к 1, поскольку каждый сомножитель стремится к 1.

Следующая таблица показывает точность пуассоновского приближения при n = 100 и p = 0,01:

i	0	1	2	3	4
$\mathbf{P}(S_n = i)$	0,366	0,370	0,185	0,061	0,015
$\lambda^i e^{-\lambda}/i!$	0,368	0,368	0,184	0,061	0,015

Пример 2. Вычисление страхового тарифа. Согласно статистике, накопленной страховой компанией, вероятность угона в течение года автомобиля определенной марки $p = 0{,}0123$. Из $n = 500$ страхуемых машин в следующем году предположительно будет угнано $\lambda = pn = 6{,}15$ автомобилей. Поскольку вероятность отдельного угона мала, а угоны предполагаются независимыми, то можно использовать пуассоновское приближение для распределения *числа ожидаемых угонов* K.

Проблема состоит в определении величины *страхового процента* δ, компенсирующего убытки в случае неожиданно большого числа угонов. Пусть C — *средняя цена автомобиля* данной марки, и в случае угона она выплачивается владельцу полностью. Тогда доход страховой компании равен δCn, а *страховые выплаты* составят сумму CK. Зададим *коэффициент значимости* α — малую вероятность, которой мы готовы пренебречь (скажем, положим $\alpha = 0{,}05$), и вычислим значение δ из условия

$$\mathsf{P}(CK > \delta Cn) = \mathsf{P}(K > \delta n) = \alpha.$$

В частности, для приведённых выше значений p и n, используя встроенную в программу Excel функцию ПУАССОН (POISSON), находим, что $\mathsf{P}(K > 10) \approx 0{,}05$. Отсюда $\delta n = 10$ или $\delta = 10/500 = 2\%$.

В заключение приведём цитату о важности предельных теорем:

«При формальном построении курса теории вероятностей предельные теоремы появляются в виде своего рода надстройки над элементарными главами теории вероятностей, в которых все задачи имеют конечный, чисто арифметический характер. В действительности, однако, познавательная ценность теории вероятностей раскрывается только предельными теоремами. Более того, без предельных теорем не может быть понято реальное содержание самого исходного понятия всей нашей науки — понятия вероятности».

Б. В. Гнеденко, А. Н. Колмогоров «*Предельные распределения для сумм независимых случайных величин*»