Закон больших чисел, центральная предельная теорема

Теорема 1. Неравенство Чебышева. Если случайная величина ξ имеет конечную дисперсию ($D(\xi) < \infty$), то для любого $\varepsilon > 0$ справедливо неравенство

$$P\{|\xi - E\xi| \geq \varepsilon\} \leq \frac{D(\xi)}{\varepsilon^2}.$$

Доказательство. Рассмотрим неотрицательную случайную величину η, имеющую математическое ожидание $E\eta$. Будем считать для определенности, что η непрерывная; для дискретной случайной величины все легко повторить.

$$E\eta = \int_{-\infty}^{\infty} xp_\eta(x)dx = \int_{\{x\geq\varepsilon\}} xp_\eta(x)dx + \int_{\{x<\varepsilon\}} xp_\eta(x)dx \geq$$

$$\geq \int_{\{x\geq\varepsilon\}} xp_\eta(x)dx \geq \varepsilon \int_{\{x\geq\varepsilon\}} p_\eta(x)dx = \varepsilon\, P\{\eta \geq \varepsilon\},$$

откуда следует **неравенство Маркова**

$$P\{\eta \geq \varepsilon\} \leq \frac{E\eta}{\varepsilon}.$$

Если применить его к случайной величине $\eta = (\xi - E\xi)^2$, то получится неравенство Чебышева.■

Теорема 2. ***Закон больших чисел Чебышева***. Пусть $\xi_1, \xi_2, \dots, \xi_n$ – независимые одинаково распределенные случайные величины с математическим ожиданием $E\xi_k = m$ и дисперсией $D\xi_k = \mathcal{D} < \infty$, тогда для любого $\varepsilon > 0$ справедливо неравенство

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - m\right| \geq \varepsilon\right\} \leq \frac{D}{n\varepsilon^2}.$$

Доказательство. Рассмотрим случайную величину

$$\xi = \frac{1}{n}\sum_{k=1}^{n}\xi_k,$$

для которой

$$E\xi = m, \qquad D\xi = D\left(\sum_{k=1}^{n}\frac{\xi_k}{n}\right) = \sum_{k=1}^{n} D\left(\frac{\xi_k}{n}\right) = \sum_{k=1}^{n}\frac{D\xi_k}{n^2} = \frac{\mathcal{D}}{n}.$$

Подставляя ξ в неравенство Чебышева, получаем утверждение Теоремы.■

Применим доказанное неравенство к схеме Бернулли. Пусть p – вероятность успеха в одном испытании, тогда число успехов в n независимых испытаниях можно представить как сумму $\xi_1 + \xi_2 + \cdots + \xi_n$ независимых бинарных случайных величин ξ_k:

$$P\{\xi_k = 1\} = p, \qquad P\{\xi_k = 0\} = 1 - p.$$

Рассмотрим частоту появления успеха в серии испытаний,

$$\hat{p} = \frac{1}{n}\sum_{k=1}^{n}\xi_k = \frac{\{\text{число успехов в } n \text{ испытаниях}\}}{n};$$

поскольку $E\xi_k = p,\ D\xi_k = p(1-p)$, согласно неравенству Чебышева получаем

$$P\{|\hat{p} - p| \geq \varepsilon\} \leq \frac{p(1-p)}{n\varepsilon^2}.$$

Мы получили **закон больших чисел Бернулли**: наблюдаемая частота события в серии испытаний стремится к вероятности этого события при большом числе испытаний, в том смысле, что

$$\lim_{n\to\infty} P\{|\hat{p} - p| \geq \varepsilon\} = 0,$$

то есть, вероятность любого, сколь угодно малого отклонения стремится к 0.

Определение. Последовательность $\{\xi_n\}$ сходится к случайной величине ξ_* **по вероятности**, (обозначается $\xi_n \xrightarrow{P} \xi_*$), если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|\xi_n - \xi_*| \geq \varepsilon\} = 0.$$

С использованием этого понятия ***закон больших чисел в схеме Бернулли*** можно сформулировать так: наблюдаемая частота события в длительной серии экспериментов стремится по вероятности к теоретической вероятности этого события.

Теорема 3. ***Центральная предельная теорема***. Пусть $\xi_1, \xi_2, \dots, \xi_n$ – независимые одинаково распределенные случайные величины с математическим ожиданием $E\xi_k = a$ и дисперсией $D\xi_k = \mathcal{D} = \sigma^2 < \infty$, тогда

$$\lim_{n\to\infty} P\left\{\sqrt{n}\frac{\bar{\xi} - a}{\sigma} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du,$$

здесь обозначено $\bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \xi_k$ - называется выборочное среднее; далее это среднее арифметическое будет очень часто встречаться во многих задачах.

Слева под знаком предела стоит функция распределения случайной величины $\sqrt{n}\frac{\bar{\xi}-a}{\sigma}$, а в правой части равенства – функция распределения стандартного нормального закона $F_0(x)$, поэтому утверждение Теоремы состоит в том, что распределение выборочного среднего $\bar{\xi}$ стремится к нормальному $\mathcal{N}\left(a,\ \frac{\sigma^2}{n}\right)$.

Доказательство. Обозначим $\gamma_n = \sqrt{n}\frac{\bar{\xi}-a}{\sigma}$ и преобразуем ее к виду суммы случайных величин:

$$\gamma_n = \frac{\bar{\xi} - a}{\frac{\sigma}{\sqrt{n}}} = \frac{1}{n}\frac{\sum_{k=1}^{n}(\xi_k - a)}{\frac{\sigma}{\sqrt{n}}} = \sum_{k=1}^{n}\frac{\xi_k - a}{\sigma\sqrt{n}} = \sum_{k=1}^{n}\eta_k.$$

Относительно так определенных $\eta_k = \frac{\xi_k - a}{\sigma\sqrt{n}}$, легко установить следующее: они имеют одинаковые распределения, $E\eta_k = 0,\ D\eta_k = \frac{\sigma^2}{\sigma^2 n} = \frac{1}{n}$, поэтому характеристическая функция

$$\varphi_{\eta_k}(t) = 1 + itE\eta_k + \frac{(it)^2}{2}E\eta_k^2 + o(t) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right).$$

Далее, характеристическую функцию случайной величины γ_n – суммы независимых случайных величин – можно представить в виде

$$\varphi_{\gamma_n}(t) = \prod_{k=1}^{n} \varphi_{\eta_k}(t) = \left[1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right]^n,$$

следовательно, $\lim\limits_{n\to\infty} \varphi_{\gamma_n}(t) = e^{-\frac{t^2}{2}}$ – это и есть характеристическая функция стандартного нормального закона. А тогда, согласно теореме Хелли (Параграф 12), последовательность функций распределения $P\left\{\sqrt{n}\frac{\bar{\xi}-a}{\sigma} \le x\right\}$ сходится к функции распределения стандартного нолрмального. ■

В качестве следствия легко получаем следующий результат для схемы Бернулли.

Теорема 4. Интегральная теорема Муавра – Лапласа. Обозначим в схеме Бернулли с вероятностью успеха в одном испытании p,

$$U_n = \{\text{число успехов в } n \text{ испытаниях}\},$$

тогда

$$\lim_{n\to\infty} P\left\{\frac{U_n - np}{\sqrt{np(1-p)}} \le x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$$

Это прямо следует из предыдущей Теоремы, поскольку в схеме Бернулли $m = np, \sigma^2 = p(1-p)$. При использовании этой формулы для приближенных вычислений вероятность p должна быть не очень близка к 0 или 1.

Пример 1. Оценка неизвестной вероятности и числа π. Применим закон больших чисел и центральную предельную теорему к задаче оценивания неизвестной вероятности p на примере иглы Бюффона (Пример 2 из Лекции 2).

Согласно закону больших чисел Бернулли, для оценки вероятности p в виде частоты $\hat{p} = \frac{U_n}{n}$ справедливо неравенство

$$P\{|\hat{p} - p| \ge \varepsilon\} \le \frac{p(1-p)}{n\varepsilon^2}.$$

Спрашивается: сколько раз надо бросить иглу, чтобы ошибка $|\hat{p} - p|$ была не более 0,003 с вероятностью по крайней мере 0,95?

$$P\{|\hat{p} - p| < 0{,}003\} \ge 0{,}95 = 1 - \frac{p(1-p)}{n(0{,}003)^2}.$$

Из соотношения

$$\frac{p(1-p)}{n(0{,}003)^2} \le 0{,}05$$

можно найти оценку для n,

$$n \ge \frac{p(1-p)}{0{,}05(0{,}003)^2}.$$

Оценкой этой воспользоваться нельзя, так справа находится неизвестная вероятность p, которую мы как раз и оцениваем. Но стоящее в числителе выражение $p(1-p)$ принимает максимальной значение на интервале $[0,1]$, равное $\frac{1}{4}$, в точке $p = \frac{1}{2}$ поэтому можно принять

$$n \ge \frac{1}{4 \cdot 0{,}05 \cdot (0{,}003)^2} = 555\ 556.$$

Как показано в Примере 2 из Лекции 2, число π связано с вероятностью пересечения p линии в задаче Бюффона формулой

$$\pi = \frac{2l}{ap}$$

для $\frac{l}{a} = 1$ погрешность Δp = 0,003 в оценке вероятности p приводит к погрешности примерно в 0,01 в оценке числа π. Не очень хорошая получилась оценка.

Но закон больших чисел не лучший инструмент получения таких оценок, он дает очень грубое приближение; лучший результат можно получить с помощью центральной предельной теоремы. Действительно, применим интегральную теорему Муавра-Лапласа,

$$P\left\{\frac{|U_n - np|}{\sqrt{np(1-p)}} \le z\right\} = \frac{1}{\sqrt{2\pi}} \int_{-z}^{z} e^{-\frac{u^2}{2}} du = 2F_0(z) - 1 = 0{,}95;$$

подберем z, удовлетворяющее последнему равенству,

$$2F_0(z) - 1 = 0{,}95 \ \Rightarrow z = F_0^{-1}\left(\frac{1 + 0{,}95}{2}\right) = 1{,}96.$$

Из соотношения

$$\frac{|U_n - np|}{\sqrt{np(1-p)}} = \frac{|n\hat{p} - np|}{\sqrt{np(1-p)}} = \sqrt{\frac{n}{p(1-p)}}\Delta p = z$$

находим

$$n = \frac{z^2 p(1-p)}{\Delta p^2} = \frac{1{,}96^2}{4(0{,}003)^2} = 106711.$$

Следующая теорема об асимптотической нормальности позволяет получать асимптотические распределения для выборочных моментов и для других функций от выборки.

Теорема об асимптотической нормальности. Если распределение последовательности случайных величин

$$\frac{\xi_n - a}{\sigma/\sqrt{n}}$$

сходится к $\mathcal{N}(0, 1)$ при $n \to \infty$, и $g(x)$ – дифференцируемая функция, причем, $g'(a) \neq 0$, то распределение последовательности

$$\frac{g(\xi_n) - g(a)}{|g'(a)|\sigma/\sqrt{n}}$$

также стремится к $\mathcal{N}(0, 1)$.

Это результат также известен в литературе под названием “**дельта метод**” и применяется для построения приближенных асимптотических распределений выборочных статистик. Он основан на том факте, что даже нелинейная функция от асимптотически нормальной случайной последовательности также имеет в пределе нормальное распределение: если функция g дифференцируема, то в окрестности точки a она ведет себя приближенно как линейная, $g(x) \approx g(a) + g'(a)(x - a)$, а линейное преобразование нормально распределенной случайной величины дает нормально распределенную. Таким образом, если дисперсия ξ_n мала, то ξ_n с вероятностью, близкой к 1, попадает в малую окрестность точки a, где $g(\xi_n)$ похожа на линейную функцию, так что

$$g(\xi_n) \sim \mathcal{N}\left(g(a), \frac{\sigma^2 (g'(a))^2}{n}\right).$$