Теория вероятностей и статистика

Тема 2. Независимые испытания

Белов А.И.

Уральский федеральный университет

Екатеринбург, 2020

Схема Бернулли

Пусть мы проводим эксперимент со случайными исходами, в ходе которого наблюдаем за появлением события A.
Если оно произойдет, будем называть это успехом, если нет — неудачей.

Ограничимся всего двумя элементарными исходами — успехом и неудачей. Такой эксперимент будем называть испытанием.

$P(A) = p$ — вероятность успеха,
$P(\overline{A}) = 1 - p = q$ — вероятность неудачи.

Определение

Пусть мы проводим n одинаковых независимых испытаний с вероятностью успеха p в каждом испытании. Такая серия испытаний называется схемой (испытаний) Бернулли с параметрами n и p.

Примеры

- Монета бросается 100 раз. Успехом считается то, что она упала вверх орлом. Схема Бернулли с параметрами $n = 100$ и $p = \frac{1}{2}$.
- Бросаются три игральные кости. Будем считать успехом, если на кости выпало одно или пять очков.
  Схема Бернулли с параметрами $n = 3$ и $p = \frac{1}{3}$.
- В урне находятся три белых и два чёрных шара. Наудачу вынимаются два шара. Успехом считается то, что вытащенные шары одного цвета. Такой эксперимент проводится четыре раза.
  Схема Бернулли с $n = 4$.
  Найдем p, используя классическое определение вероятности:
  $N = C_5^2 = 10$, $m(A) = C_3^2 + C_2^2 = 3 + 1 = 4$.
  $$p = \frac{m(A)}{N} = \frac{4}{10} = \frac{2}{5}.$$

Формула Бернулли

Теорема (формула Бернулли)

Вероятность $P_n(k)$ того, что в схеме Бернулли с параметрами n и p наступит ровно k $(0 \le k \le n)$ успехов равна

$$P_n(k) = C_n^k p^k q^{n-k}.$$

Доказательство.

Пусть $A_{i_1,\ldots,i_k}$ — произошло k успехов в испытаниях $i_1, \ldots, i_k$.
Тогда очевидно, что $P(A_{i_1,\ldots,i_k}) = p^k q^{n-k}$. Поскольку k из n испытаний, в которых произошёл успех, можно выбрать C_n^k способами и все события $A_{i_1,\ldots,i_k}$ для различных наборов индексов попарно несовместны, мы получаем формулу теоремы. $\square$

Следствия формулы Бернулли

$$\boxed{P_n(k) = C_n^k p^k q^{n-k}}$$

В силу свойств биномиальных коэффициентов

- $P_n(n) = p^n$ (вероятность успеха в каждом испытании);
- $P_n(0) = q^n$ (вероятность неудачи в каждом испытании);
- $P_n(n-1) = np^{n-1}q$ (вероятность ровно одной неудачи);
- $P_n(1) = npq^{n-1}$ (вероятность ровно одного успеха);
- $P_n(0) + P_n(1) + \ldots + P_n(n) = 1.$

Применение формулы Бернулли

Пример

В урне находятся три белых и два чёрных шара. Наудачу вынимаются два шара. Успехом считается то, что вытащенные шары одного цвета. Такой эксперимент проводится четыре раза. Найти вероятность того, что успех наступит а) три раза, б) не менее трёх раз, в) не более трёх раз.

Решение. Имеем схему Бернулли с $n = 4$ и $p = 0{,}4$.

а). $P_4(3) = C_4^3 \cdot 0{,}4^3 \cdot 0{,}6 = 4 \cdot 0{,}064 \cdot 0{,}6 = 0{,}1536.$

б). $P_4(3) + P_4(4) = 0{,}1536 + 0{,}4^4 = 0{,}1536 + 0{,}0256 = 0{,}1792.$

в). $P_4(0) + P_4(1) + P_4(2) + P_4(3) = 1 - P_4(4) = 1 - 0{,}0256 = 0{,}9744.$

Наивероятнейшее число успехов

Определение

Число k_0 такое, что $P_n(k_0) = \max\limits_{k=0,\dots n} P_n(k)$ называется **наивероятнейшим числом успехов**.

Чтобы понять, как изменяется $P_n(k)$ с увеличением k рассмотрим $\dfrac{P_n(k)}{P_n(k-1)}$.

Непосредственными выкладками получаем $\dfrac{P_n(k)}{P_n(k-1)} = \dfrac{np+p-k}{kq} + 1.$

- Если $k < np + p$, то $P_n(k) > P_n(k-1)$.
- Если $k > np + p$, то $P_n(k) < P_n(k-1)$
- Если $k = np + p$, то $P_n(k) = P_n(k-1)$

Из этого следует следующая теорема.

Теорема о наивероятнейшем числе успехов

Пусть $[x]$ — целая часть числа x.

Теорема (о наивероятнейшем числе успехов)

Если $np + p \notin \mathbb{Z}$, то единственное наивероятнейшее число успехов $k_0 = [np + p]$.

Если $np + p \in \mathbb{Z}$, то существует два наивероятнейших числа успехов: $k_0 = np - q$ и $k_0 + 1 = np + p$.

Задача

Испытывается устройство, содержащее 15 независимо работающих одинаковых элементов. Вероятность того, что элемент выдержит проверку, равна 0,9. Найти наивероятнейшее число элементов, которые выдержат проверку.

Решение. Проверка элементов — схема Бернулли с параметрами $n = 15$ и $p = 0.9$. $np + p = 15 \cdot 0{,}9 + 0{,}9 = 14{,}4 \notin \mathbb{Z}$, k_0 единственное и равно $[np + p] = [14{,}4] = 14$.

Локальная теорема Лапласа

Определение

Функция $g(x)$ называется асимптотическим приближением функции $f(x)$ при $x \to \infty$, если существует $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = 1.$

Теорема (Лапласа локальная)

В схеме Бернулли с параметрами n, p вероятность $P_n(k)$ при $n \to \infty$ асимптотически приближается функцией

$$\frac{1}{\sqrt{npq}}\varphi(x),$$

где $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $x = \dfrac{k-np}{\sqrt{npq}}$.

Без доказательства. $\square$

Применение локальной теоремы Лапласа

Для $n \geq 100$ можно считать

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

Для вычислений $\varphi(x)$ используют калькулятор, компьютер или таблицу функции $\varphi(x)$.

Таблица составлена для $0 \leq x \leq 4$:

- $\varphi(x)$ — чётная функция;
- для $x > 4$ можно считать $\varphi(x) \approx 0$.

Пример применения локальной теоремы Лапласа

Задача

Найти с точностью до 4-го знака после запятой вероятность того, что в серии из 400 испытаний успех наступит ровно 104 раза, если вероятность успеха равна 0,2.

Решение. Схема Бернулли с параметрами $n = 400$ и $p = 0{,}2$.

$q = 1 - p = 0{,}8;$

$\sqrt{npq} = \sqrt{400 \cdot 0{,}2 \cdot 0{,}8} = 8;$

$x = \dfrac{k - np}{\sqrt{npq}} = \dfrac{104 - 400 \cdot 0{,}2}{8} = \dfrac{24}{8} = 3;$

$\varphi(3) \approx 0{,}0044;$

$P_{400}(104) \approx \dfrac{1}{\sqrt{npq}}\varphi(x) \approx \dfrac{1}{8} \cdot 0{,}0044 \approx 0{,}0006.$

Интегральная теорема Лапласа

Введем обозначение $P_n(k_1, k_2) = P(k_1 \le k \le k_2)$.

Очевидно, что $P_n(k_1, k_2) = \sum\limits_{k=k_1}^{k_2} P_n(k)$.

Теорема (Лапласа интегральная)

Вероятность $P_n(k_1, k_2)$ *при* $n \to \infty$ *асимптотически приближается функцией*

$$\int\limits_{x_1}^{x_2} \varphi(z)\, dz,$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $x_i = \frac{k_i - np}{\sqrt{npq}}$ *(*$i = 1, 2$*).*

Без доказательства. □

Применение интегральной теоремы Лапласа

Для вычислений используется функция Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-z^2/2}\, dz$$

По формуле Ньютона-Лейбница имеем

$$\boxed{P_n(k_1, k_2) \approx \Phi(x_2) - \Phi(x_1)}$$

Для вычислений используют компьютер или таблицу функции $\Phi(x)$ ($0 \leq x \leq 5$).

- $\Phi(x)$ — нечетная функция;
- для $x > 5$ можно считать $\Phi(x) \approx 0{,}5$;
- $\Phi(x)$ монотонно возрастает на всей числовой оси.

Пример применения интегральной теоремы Лапласа

Задача

Стрелок попадает по мишени с вероятностью 0,75. Найти с точностью до тысячных вероятность того, что стрелок попадёт не менее 70 раз из 100 выстрелов.

Решение. Схема Бернулли с параметрами $n = 100$, $p = 0{,}75$.
Требуется вычислить $P_{100}(70, 100)$ с точностью до тысячных.

$q = 1 - p = 0{,}25.$

$$x_1 = \frac{70 - 100 \cdot 0{,}75}{\sqrt{100 \cdot 0{,}75 \cdot 0{,}25}} \approx -1{,}15;\ x_2 = \frac{100 - 100 \cdot 0{,}75}{\sqrt{100 \cdot 0{,}75 \cdot 0{,}25}} \approx 5{,}77.$$

$\Phi(x_1) \approx \Phi(-1{,}15) \approx -0{,}3749$; $\Phi(x_2) \approx \Phi(5{,}77) \approx 0{,}5$.

$P_{100}(70, 100) \approx \Phi(x_2) - \Phi(x_1) \approx 0{,}5 + 0{,}3749 \approx 0{,}875.$

Отклонение частоты от вероятности

Пусть $\left|\frac{k}{n} - p\right|$ — отклонение частоты от вероятности по абсолютной величине.
Тогда из интегральной теоремы Лапласа следует

Теорема (об отклонении частоты от вероятности)

Пусть мы имеем схему Бернулли с параметрами n, p и k — число успехов. Пусть также $\varepsilon > 0$ — произвольное положительное число. Тогда вероятность $P\left(\left|\frac{k}{n} - p\right| \leq \varepsilon\right)$ при $n \to \infty$ асимптотически приближается функцией

$$2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Применение теоремы об отклонении частоты от вероятности

При достаточно больших n

$$\boxed{P\left(\left|\frac{k}{n}-p\right|\leq\varepsilon\right)\approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)}$$

Задача (о бросании монеты)

Какое минимальное число раз нужно бросить монету, чтобы с вероятностью 0,95 можно было утверждать, что частота появлений орла отклониться по абсолютной величине от 0,5 не более, чем на 0,01?

Решение задачи о бросании монеты

Имеем схему Бернулли. Параметр $p = 0{,}5$, параметр n неизвестен.

Величина максимального отклонения частоты от вероятности $\varepsilon = 0{,}01$.

$$P\left(\left|\frac{k}{n} - 0{,}5\right| \leq 0,01\right) \approx 2\Phi\left(0,01\sqrt{\frac{n}{0{,}5 \cdot 0{,}5}}\right) = 2\Phi\left(0{,}02\sqrt{n}\right) \geq 0,95.$$

$\Phi\left(0{,}02\sqrt{n}\right) \geq 0{,}475.$

По таблице $\Phi(1{,}96) \approx 0{,}475.$

$\Phi(x)$ монотонно возрастает, т. е. $0{,}02\sqrt{n} \geq \Phi^{-1}(0{,}475) \approx 1{,}96.$

$\sqrt{n} \geq 98$ или $n \geq 9604.$

Минимальное значение $n = 9604.$

Теорема Пуассона

В ситуации серии редких событий, т. е. когда вероятность успеха p мала ($p \leq 0{,}1$), гораздо лучший результат даёт асимптотическое приближение Пуассона:

Теорема (Пуассон)

Пусть в схеме Бернулли с параметрами n, p имеет место $n \to \infty$, $p \to 0$ так, что $np \to \lambda$, где λ — постоянная величина. Тогда

$$P_n(k) \to \frac{\lambda^k}{k!} e^{-\lambda}.$$

Без доказательства. $\square$

Применение теоремы Пуассона

При $n \geq 100$ и $npq > 9$ применяют асимптотику Лапласа.

В случае $npq \leq 9$ — теорему Пуассона:

$$\boxed{P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}}$$

Для вычислений берут $\lambda = np$.

Для небольшого числа слагаемых при $npq \leq 9$ тоже лучше использовать теорему Пуассона:

$$\boxed{P_n(k_1, k_2) \approx \sum_{k=k_1}^{k_2} \frac{\lambda^k}{k!} e^{-\lambda}}$$

Пример применения теоремы Пуассона

Задача

Вероятность повреждения аппаратуры спутника при попадании в него высокоэнергетической частицы равна 0,002. За год эксплуатации в спутник попало 1000 частиц. Определить вероятность необходимости ремонта спутника, если необходимость его проведения наступает, когда спутник имеет не менее трёх повреждений.

Решение. Схема Бернулли с параметрами $n = 1000$ и $p = 0{,}002$.
$n = 1000 \geq 100$; $p = 0{,}002 \leq 0{,}1$; $npq = 1{,}996 < 9$.

Используем теорему Пуассона. $\lambda = 1000 \cdot 0{,}002 = 2$.

$P_{1000}(0, 2) \approx \sum_{k=0}^{2} \frac{2^k}{k!} e^{-2} = e^{-2} \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} \right) \approx 0{,}1553 \cdot 5$

$\approx 0{,}766$. $P_{1000}(3, 1000) = 1 - P_{1000}(0, 2) \approx 1 - 0{,}766 = 0{,}234$.

Поток событий

Поток событий — это последовательность событий, которые наступают в случайный момент времени.

Определение

Поток событий называется стационарным, если вероятность появления k событий на любом промежутке времени зависит только от k и от длины промежутка.

Определение

Поток событий называется потоком с отсутствием последствий, если вероятность появления k событий на любом промежутке времени не зависит от того, появлялись или нет эти события до начала этого временного промежутка.

Простейший (пуассоновский) поток событий

Определение

Поток событий называется ординарным, если вероятность появления двух или более событий за малый промежуток времени пренебрежительно мала, т. е. если k — число событий, произошедших за промежуток времени длиной t, то $\lim\limits_{t \to 0} P(k \geq 2) = 0$.

Определение

Стационарный ординарный поток событий с отсутствием последствий называется простейшим (или пуассоновским) потоком.
Число λ — среднее число событий в единицу времени — называют интенсивностью потока.

Теорема о простейшем потоке событий

Теорема

Пусть дан простейший поток событий с интенсивностью λ. Тогда вероятность того, что за время t произойдет k событий равна

$$P_t(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Доказательство

Разобьем промежуток времени длины t на n одинаковых последовательных промежутков (длиною $\frac{t}{n}$).

Если $n \to \infty$, то $\frac{t}{n} \to 0$.

Если n достаточно велико, то эта длина будет достаточно малой.

Доказательство теоремы о простейшем потоке событий

Продолжение доказательства

Ординарность потока $\Rightarrow$ считаем что в каждый из этих промежутков может произойти не более одного события.

Отсутствие последствий $\Rightarrow$ события потока в каждом из промежутков наступают независимо.

Стационарность потока $\Rightarrow$ вероятность наступления события будет одинакова для всех промежутков $(p = \frac{\lambda t}{n})$.

Итак налицо схема Бернулли с параметрами n, $p = \frac{\lambda t}{n}$.

Событие, что в промежуток времени t наступит k событий потока, ни что иное как k успехов в серии из n одинаковых независимых испытаний.

Доказательство теоремы о простейшем потоке событий

Окончание доказательства

Пусть $n \to \infty$.

Тогда $\lim\limits_{n\to\infty} p = \lim\limits_{n\to\infty} \frac{\lambda t}{n} = 0$.

Кроме того $np = n \cdot \frac{\lambda t}{n} = \lambda t$ — константа.

$$P_t(k) = \lim_{n\to\infty} P_n(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

по теореме Пуассона. $\square$

Задача о простейшем потоке событий

Задача

Среднее число вызовов, поступающих на сотовую вышку в минуту, равно 5. Найти вероятность того, что за 2 минуты поступит а) два вызова; б) менее двух вызовов; в) не менее двух вызовов.

Решение. Простейший поток событий с интенсивностью $\lambda = 5$ (в качестве единицы времени выбрана минута).

Длина временного интервала $t = 2$. Тем самым $\lambda t = 10$.

а) $P_2(2) = \frac{10^2}{2!} e^{-10} \approx 50 \cdot 0{,}0000454 \approx 0{,}0023.$

б) $P_2(0) + P_2(1) = \frac{10^0}{0!} e^{-10} + \frac{10^1}{1!} e^{-10} = e^{-10}(1 + 10) \approx 0{,}0000454 \cdot 11 \approx 0{,}0005.$

в) $P_2(2) + P_2(3) + \ldots = 1 - (P_2(0) + P_2(1)) \approx 0{,}9995.$