Лекция 5

Схема испытаний и соответствующее распределение

Введем обозначения:

n - число испытаний

p - вероятность успеха при одном испытании

$q = 1 - p$ - вероятность неудачи

I. Схема Бернулли

$\urcorner \nu_n$ - число успехов в серии из n испытаний

$P_n(\nu_n = k) = C_n^k p^k q^{n-k}, \qquad k = 0, 1, \dots, n$

Def. Соответствие $k \to C_n^k p^k q^{n-k}, \quad k = 0, \dots, n$ называется биномиальным распределением (обозначается $B_{n,p}$ или $B(n, p)$)

II. Схема до первого успешного испытания

Пусть проводится бесконечная серия испытаний, которая заканчивается после первого успешного испытания под номером τ

Th. $P(\tau = k) = q^{k-1}p, \qquad k = 1, 2, \dots$

□

$P(\tau = k) = P(\underbrace{\text{Н} \dots \text{Н}}_{k-1}\text{У}) = q^{k-1}p$

□

Def. Соответствие $k \to q^{k-1}p, k \in \mathbb{N}$ называется геометрическим распределение вероятности (обозначается G_p или $G(p)$)

Nota. Геометрическое распределение обладает свойством нестарения или свойством отсутствия последействия

Th. $\urcorner P(\tau = k) = q^{k-1}p, k \in \mathbb{N}$. Тогда $\forall n, k \geq 0 \quad P(\tau > n + k \mid \tau > n) = P(\tau > k)$

□

Заметим, что $P(\tau > m) = q^m$, первые m - неудачи

$P(\tau > n+k|\tau > n) = \dfrac{P(\tau > n+k, \tau > n)}{P(\tau > n)} = \dfrac{P(\tau > n+k)}{P(\tau > n)} = \dfrac{q^{n+k}}{q^n} = q^k$

□

Nota. $P(\tau = n+k \mid \tau > n) = p(\tau = k)$ - Lab. доказать

III. Схема испытаний с несколькими исходами

Пусть при n независимых испытаний могут произойти m исходов (несовместных)

p_i - вероятность i-ого исхода при одном испытании

Th. Вероятность того, что при n испытаниях первый исход появится n_1 раз, второй - n_2 раз, m-ый - n_m $(\sum\limits_{i=1}^{m} n_i = n)$ равно

$P_n(n_1, n_2, \ldots, n_m) = \dfrac{n!}{n_1! n_2! \ldots n_m!} p_1^{n_1} p_2^{n_2} \ldots p_m^{n_m}$

При $m = 2$ получаем формулу Бернулли

□

Рассмотрим следующий благоприятный исход, обозначим A_1

$A_1 = \underbrace{11\ldots1}_{n_1}\underbrace{22\ldots2}_{n_2}\ldots\underbrace{mm\ldots m}_{n_m}$

$p(A_1) = p_1^{n_1} p_2^{n_2} \ldots p_m^{n_m}$

Все остальные благоприятные исходы имеют ту же вероятность и отличаются лишь расположением i-ых исходов на n позициях, получаем мультиномиальную теорему: $\dfrac{n!}{n_1! n_2! \ldots n_m!}$

В итоге получаем требуемую формулу

□

Ex. Два одинаковых сильных шахматиста играют шесть партий

Вероятность ничьи в партии - 0.5. Какова вероятность того, что второй игрок выиграет две партии, а еще три сведет к ничьей

1-ый исход - выиграл 1 игрок

2-ой исход - выиграл 2 игрок

3-ий исход - ничья

$n = 6; \quad p_3 = 0.5; \quad p_1 = p_2 = \dfrac{1 - p_3}{2} = 0.25$

$$P_6(1;2;3) = \frac{6!}{1!2!3!}\left(\frac{1}{4}\right)^1\left(\frac{1}{4}\right)^2\left(\frac{1}{2}\right)^3 = \frac{4\cdot5\cdot6}{2}\frac{1}{2^9} \approx 0.12$$

IV. Урновая схема

В урне N шаров, из которых K шаров белые, $N-K$ - черные

Из урны вынимаем (без учета порядка) n шаров. Найти вероятность, что из них k белых

а) Схема с возвратом (после каждого раза кладем шар обратно). В этом случае вероятность вынуть белый шар одинакова и равна $\frac{K}{N}$. Получаем схему Бернулли: $P_n(k) = C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$

б) Схема без возврата - вынутый шар мы выбрасываем, тогда $P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}$

Def. Соответствие $k \longrightarrow \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}, k = 0, \dots, n$ называется гипергеометрическим распределением

Nota. Если $K, N \to \infty$ так, что $\frac{K}{N} \approx p$ (не меняется), а n и k зафиксировать, то после выбора n шаров пропорции состава шаров не сильно изменятся, поэтому логично предположить, что гипергеометрическое распределение будет сходиться к биномиальному

Th. Если $K, N \to \infty$ таким образом, что $\frac{K}{N} \to p \in (0;1)$, а n и $0 \le k \le n$ фиксированы, то вероятность при гипергеометрическом распределении будет стремиться к биномиальному:

$$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \longrightarrow C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$$

Воспользуемся леммой: $C_n^k \sim \frac{n^k}{k!}$ при $n \to \infty$ и фиксированном k

Доказательство леммы: $C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{n^k}{k!} = 1\left(1-\frac{1}{n}\right)\dots\left(1-\frac{k-1}{n}\right)\frac{n^k}{k!} \sim \frac{n^k}{k!}$

□

$$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \sim \frac{K^k}{k!}\frac{(N-K)^{n-k}}{(n-k)!}\frac{n!}{N^n} = \frac{n!}{k!}\frac{(N-K)^{n-k}}{(n-k)!}\frac{K^k}{N^n} = C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k} \longrightarrow C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$$

□

V. Схема Пуассона. Теорема Пуассона для схемы Бернулли

Nota. Если вероятность успеха p в схеме Бернулли мала или близка к 1, то предельная формула Лапласа при недостаточно большом числе испытаний дает достаточно большую погрешность. В этой ситуации следует использовать формулу Пуасоона (формула редких событий)
Схема: вероятность числа успеха при одном испытании p_n зависит от числа испытаний n, причем таким образом, что $np_n \approx \lambda = const$
λ - интенсивность появления редких событий в единицу времени в потоке испытаний

Th. 1. (формула Пуассона) Пусть $n \to \infty, p_n \to 0$ таким образом, что $np_n \to \lambda = const > 0$
Тогда вероятность k успехов при n испытаниях: $P_n(k) = C_n^k p_n^k (1-p_n)^{n-k} \underset{n\to\infty}{\to} = \dfrac{\lambda^k}{k!}e^{-\lambda}$

$\square$

Обозначим $\lambda_n = np_n$. Тогда $p_n = \dfrac{\lambda_n}{n}$ и

$$P_n(k) = C_n^k \left(\frac{\lambda_n}{n}\right)^k \left(1-\frac{\lambda_n}{n}\right)^{n-k} \underset{n\to\infty}{\to} \frac{n^k}{k!}\frac{\lambda_n^k}{n^k}\left(1-\frac{\lambda_n}{n}\right)^n \cancelto{1}{\left(1-\frac{\lambda_n}{n}\right)^{-k}} \underset{n\to\infty}{\to} \frac{\lambda_n^k}{k!}\left(1-\frac{\lambda_n}{n}\right)^n =$$
$$\frac{\lambda_n^k}{k!}\left(\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right)^{-\frac{\lambda_n}{n}n} \underset{n\to\infty}{\to} \frac{\lambda_n^k}{k!}e^{-\lambda_n} \underset{n\to\infty}{\to} \frac{\lambda^k}{k!}e^{-\lambda}$$

$\square$

Th. 2. (оценка погрешности в формуле Пуассона) Пусть ν_n - число успехов при n испытаниях в схеме Бернулли
p - вероятность успеха при одном испытании, $\lambda = np$, $A \subset \{0, 1, \dots, n\}$ - произвольное подмножество чисел
Тогда $|P_n(\nu_n \in A) - \sum\limits_{k\in A} \dfrac{\lambda^k}{k!}e^{-\lambda}| \le \min(p, np^2) = \min(p, p\lambda)$
(без доказательства)

Def. Соответствие $k \to \dfrac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, \dots$ называется распределением Пуассона с параметром $\lambda > 0$ (обозначается Π_λ)

Ex. Прибор состоит из 1000 элементов, вероятность отказа каждого элемента равна 0.001. Какова вероятность отказа больше двух элементов
$P_n(k) \approx \dfrac{\lambda^k}{k!}e^{-\lambda}$
$n = 1000, p = 0.001, \lambda = 1$

$P_n(k > 2) = 1 - P_n(k \leq 2) = 1 - P(0) - P(1) - P(2) \approx 1 - \left(\frac{1^0}{0!}e^{-1} + \frac{1^1}{1!}e^{-1} + \frac{1^2}{2!}e^{-1}\right) = 1 - \left(1 + 1 + \frac{1}{2}\right)e^{-1} \approx 0.0803$