

Схема Бернулли

Определение.

Схемой Бернулли называется последовательность независимых испытаний, в каждом из которых возможны лишь **два исхода** — «успех» и «неудача», при этом «успех» в одном испытании происходит с вероятностью p , «неудача» — с вероятностью $q = 1 - p$.

Теорема. (Формула Бернулли).

Обозначим через ν_n число успехов в n испытаниях схемы Бернулли. Тогда для любого $k = 0, 1, \dots, n$

$$P(\nu_n = k) = C_n^k p^k (1 - p)^{n-k} = C_n^k p^k q^{n-k}.$$

Доказательство. Событие

$$A = \{\nu_n = k\}$$

означает, что в серии из n испытаний произошло ровно k успехов.

Рассмотрим один из благоприятствующих событию A исходов.

$$\underbrace{(Y, Y, \dots, Y)}_k, \underbrace{(H, \dots, H)}_{n-k}$$

(y – «успех», n – «неудача»)

Т.к. испытания независимы, то вероятность такого элементарного исхода равна

$$p^k (1 - p)^{n-k}$$

Первые k испытаний завершились успехом, а остальные $(n - k)$ - неудачей.

Другие, благоприятствующие нашему событию исходы, отличаются от данного лишь иным расположением k успехов по n местам.

Число благоприятствующих исходов равно числу сочетаний из n элементов по k , а вероятность события A равна сумме вероятностей всех элементарных событий, составляющих данное.

- Определение. Набор чисел

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

Называется биномиальным распределением вероятностей и обозначается

$$B_{n,p} \text{ или } B(n, p)$$

$$(p + q)^n = C_n^0 p^0 q^n + C_n^1 p^1 q^{n-1} + \square + C_n^n p^n q^0$$

Наиболее вероятное число успехов

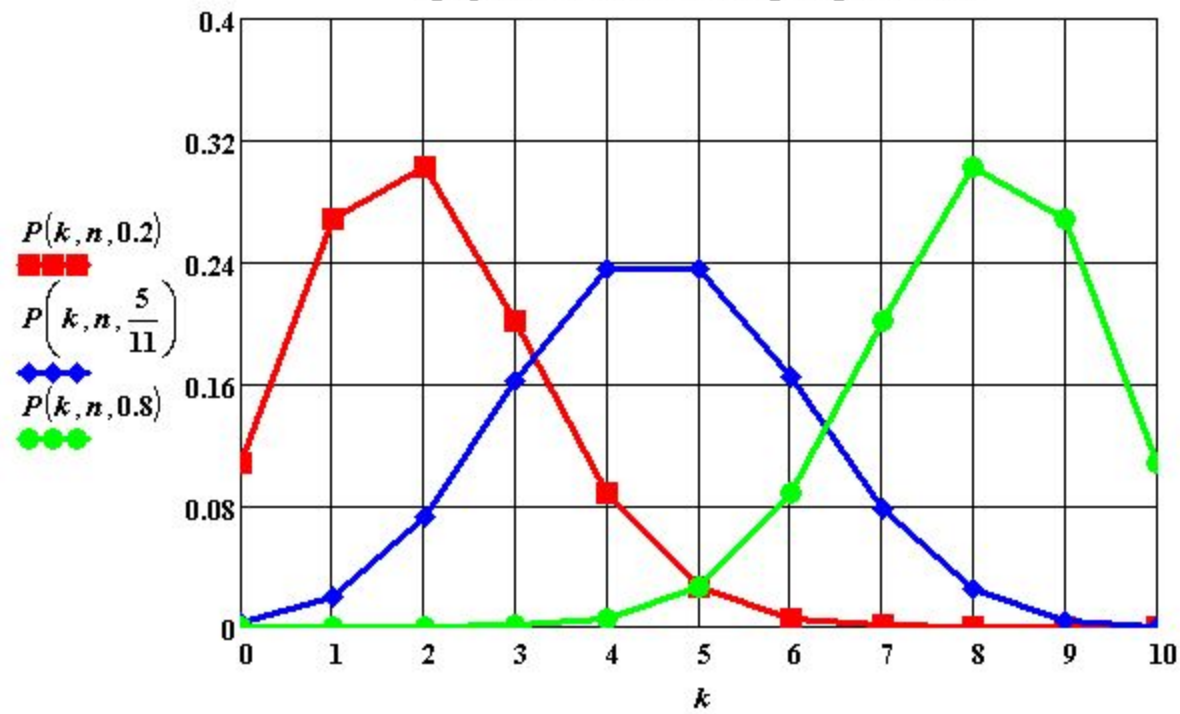
$$\begin{aligned}\frac{P(\nu_n = k)}{P(\nu_n = k - 1)} &= \frac{n!}{k!(n - k)!} \frac{(k - 1)!(n - k + 1)!}{n!} \frac{p^k q^{n - k}}{p^{k - 1} q^{n - k + 1}} = \\ &= \frac{(n - k + 1)p}{kq} = 1 + \frac{(n - k + 1)p}{kq} - 1 = 1 + \frac{np + p - k}{kq}.\end{aligned}$$

Видим, что

- (a) $P(\nu_n = k) > P(\nu_n = k - 1)$ при $np + p - k > 0$, то есть при $k < np + p$;
- (b) $P(\nu_n = k) < P(\nu_n = k - 1)$ при $np + p - k < 0$, то есть при $k > np + p$;
- (c) $P(\nu_n = k) = P(\nu_n = k - 1)$ при $np + p - k = 0$, что возможно лишь если $np + p$ — целое число.

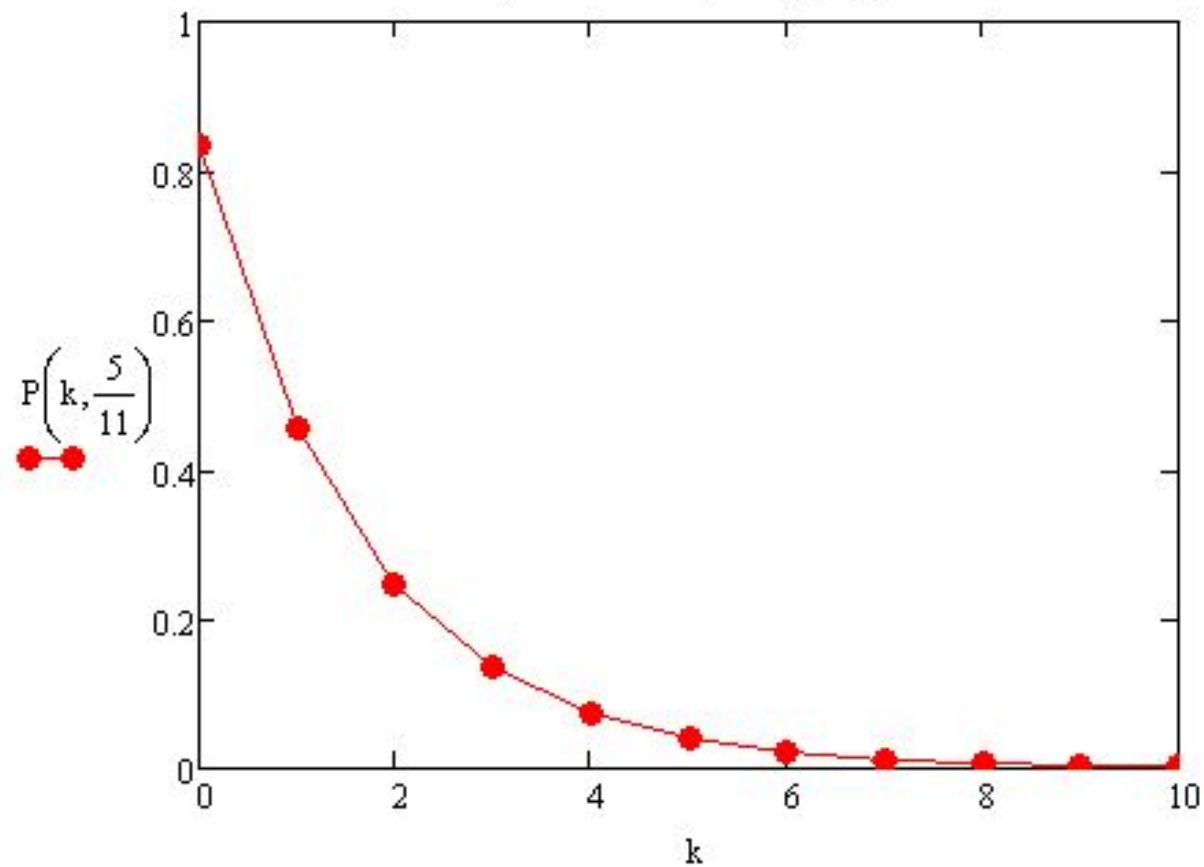
$$np - q < k^* < np + p$$

Графики биномиального распределения



$$n := 10 \quad k := 0..n \quad P(k,p) := p \cdot (1-p)^{k-1}$$

Геометрическое распределение



Номер первого успешного испытания в схеме Бернулли

Введем величину τ , принимающую

значения из $\{1, 2, 3, \dots\}$, равную *номеру первого успешного испытания*.

Теорема Вероятность того, что первый успех произойдет в испытании с номером $k \in \mathbb{N} = \{1, 2, 3, \dots\}$, равна $P(\tau = k) = p q^{k-1}$.

Доказательство. Действительно, $P(\tau = k) = P(\underbrace{n, n, \dots, n}_{k-1}, y) = p q^{k-1}$.

Определение Набор чисел $\{p q^{k-1}, k = 1, 2, 3, \dots\}$ называется *геометрическим распределением вероятностей* и обозначается G_p или $G(p)$.

Выбор без возвращения

Рассмотрим урну, содержащую N шаров, из которых K шаров — белые, $N - K$ шаров — черные.

Из урны наудачу выбирают n шаров

Вероятность $P_{N,K}(n, k)$ того, что будет выбрано ровно k белых и $n - k$ черных находится по формуле

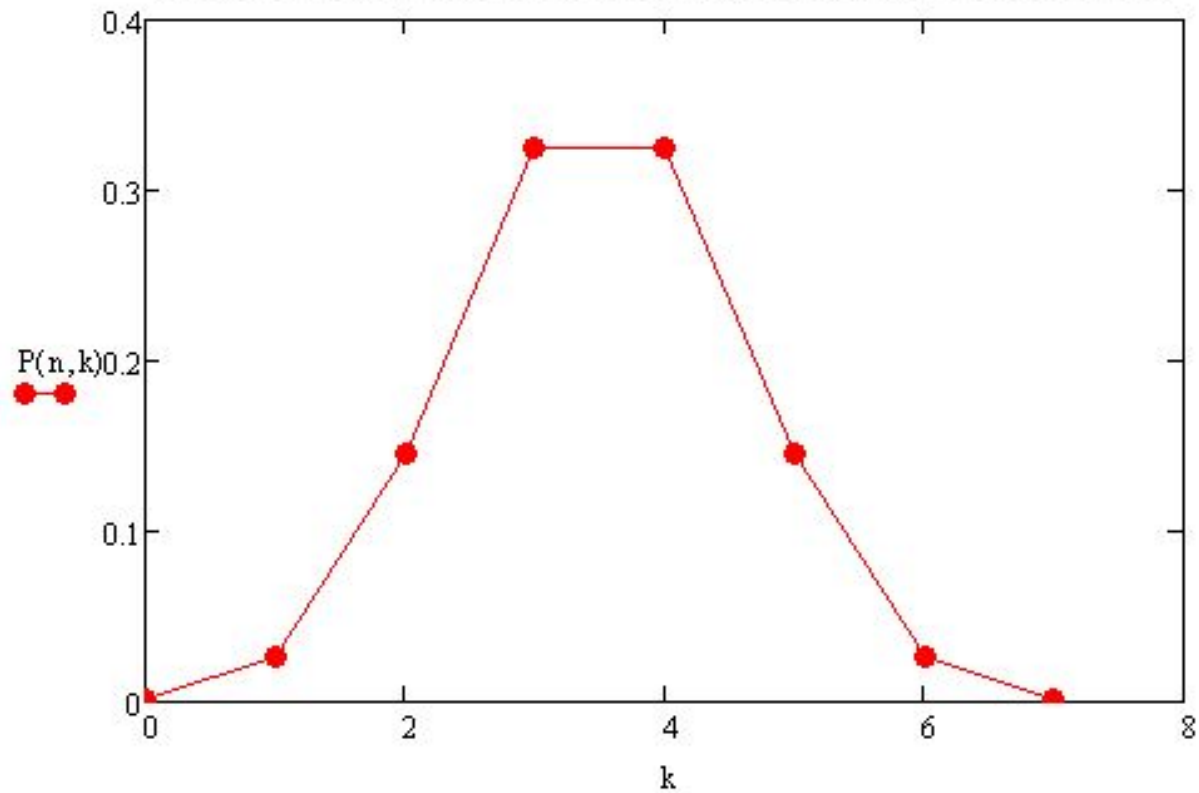
$$P_{N,K}(n, k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}.$$

Такое распределение вероятностей называется *гипергеометрическим*

$$\underline{N} := 20 \quad \underline{K} := 7 \quad \underline{P}(n, k) := C(\underline{K}, k) \cdot \frac{C(\underline{N} - \underline{K}, n - k)}{C(\underline{N}, n)}$$

$$k := 0..n$$

Гипергеометрическое распределение вероятностей



Предельное поведение гипергеометрического распределения

Теорема □ Если $N \rightarrow \infty$ и $K \rightarrow \infty$ так, что $K/N \rightarrow p \in (0, 1)$, то для любых фиксированных n , $0 \leq k \leq n$

$$P_{N,K}(n, k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \rightarrow C_n^k p^k (1-p)^{n-k}.$$

Независимые испытания с несколькими исходами

Пусть в одном испытании возможны m исходов. Обозначим их цифрами $1, 2, \dots, m$. Пусть исход i в одном испытании случается с вероятностью p_i , $1 \leq i \leq m$, и $\sum_{i=1}^m p_i = 1$. Обозначим через $P(n_1, \dots, n_m)$ вероятность того, что в $n = n_1 + \dots + n_m$ независимых испытаниях исход 1 появился n_1 раз, исход 2 — n_2 раз, ..., исход m — n_m раз.

Теорема Для любого n и любых целых $n_1 \geq 0, \dots, n_m \geq 0$ таких, что $n_1 + \dots + n_m = n$, верна формула:

$$P(n_1, \dots, n_m) = \frac{n!}{n_1! \dots n_m!} p_1^{n_1} \cdot \dots \cdot p_m^{n_m}.$$

Полиномиальное распределение

Предельные теоремы с схеме Бернулли

Теорема Пуассона

Предположим, нам нужна вероятность получить не менее десяти успехов в 1000 испытаний схемы Бернулли с вероятностью успеха 0.003. Вероятность этого события равна любому из следующих выражений:

$$\sum_{k=10}^{1000} C_{1000}^k (0.003)^k (0.997)^{1000-k} = 1 - \sum_{k=0}^9 C_{1000}^k (0.003)^k (0.997)^{1000-k},$$

и вычисление даже одного слагаемого в каждом из этих выражений весьма проблематично.

Теорема ■ (Теорема Пуассона).

Пусть $n \rightarrow \infty$, $p_n \rightarrow 0$ так, что $np_n \rightarrow \lambda > 0$. Тогда для любого $k \geq 0$ вероятность получить k успехов в n испытаниях схемы Бернулли с вероятностью успеха p_n стремится к величине $\frac{\lambda^k}{k!} e^{-\lambda}$:

$$P(\nu_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \text{ при } n \rightarrow \infty, p_n \rightarrow 0 \text{ так, что } np_n \rightarrow \lambda > 0.$$

Доказательство. Положим $\lambda_n = n \cdot p_n \rightarrow \lambda > 0$. $C_n^k \sim \frac{n^k}{k!}$ при фиксированном k и при $n \rightarrow \infty$. Тогда

$$C_n^k p_n^k (1 - p_n)^{n-k} = C_n^k \frac{\lambda_n^k}{n^k} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \sim \frac{n^k}{k!} \frac{\lambda_n^k}{n^k} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^n}_{\downarrow e^{-\lambda}} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-k}}_{\downarrow 1} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}.$$

Определение 23. Пусть $\lambda > 0$ — некоторая постоянная. Набор чисел

$$\left\{ \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \dots \right\}$$

называется *распределением Пуассона с параметром λ* .

Теорема 17 (Теорема Пуассона с оценкой погрешности).

Пусть $A \subseteq \{0, 1, 2, \dots, n\}$ — произвольное множество целых неотрицательных чисел, ν_n — число успехов в n испытаниях схемы Бернулли с вероятностью успеха p , $\lambda = n \cdot p$. Тогда

$$\left| \mathbf{P}(\nu_n \in A) - \sum_{k \in A} \frac{\lambda^k}{k!} e^{-\lambda} \right| = \left| \sum_{k \in A} C_n^k p^k (1-p)^{n-k} - \sum_{k \in A} \frac{\lambda^k}{k!} e^{-\lambda} \right| \leq np^2 = \frac{\lambda^2}{n}.$$

Предельная теорема Муавра-Лапласа

Пусть A — событие, которое может произойти в любом из n независимых испытаний с одной и той же вероятностью $p = \mathbf{P}(A)$. Пусть $\nu_n(A)$ — число осуществлений события A в n испытаниях. Тогда $\frac{\nu_n(A) - np}{\sqrt{np(1-p)}} \Rightarrow \mathbf{N}_{0,1}$. Иначе говоря, для любых вещественных $x < y$ при $n \rightarrow \infty$ имеет место сходимость

$$\mathbf{P} \left(x \leq \frac{\nu_n(A) - np}{\sqrt{np(1-p)}} \leq y \right) \rightarrow \Phi_{0,1}(y) - \Phi_{0,1}(x) = \int_x^y \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt;$$