ЛЕКЦИЯ 4

Теория вероятностей и математическая статистика

Схема Бернулли. Предельные теоремы

Схема Бернулли

<u>Определение</u>

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, а «неудача» — с вероятностью $q = 1 - p$.

Теорема (формула Бернулли)

Обозначим через m число успехов в n испытаниях схемы Бернулли. Тогда

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

Доказательство

Событие A = {число успехов равно m} означает, что в n испытаниях схемы Бернулли произошло ровно m успехов. Рассмотрим один из благоприятных исходов:

$$\left(\underbrace{У, У, \ldots, У}_{m}, \underbrace{Н, Н, \ldots, Н}_{n-m} \right)$$

Поскольку испытания независимы, вероятность такого элементарного исхода равна

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

Другие благоприятствующие событию A элементарные исходы отличаются от рассмотренного выше лишь расположением m успехов на n местах. Есть ровно

$$C_n^m$$

способов расположить m успехов на n местах.

Поэтому событие A состоит из

$$C_n^m$$

элементарных исходов, вероятность каждого из которых равна

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

т.е.

$$P_n(m) = C_n^m p^m q^{n-m} \quad . \Diamond$$

Наивероятнейшее число успехов

В испытаниях схемы Бернулли наиболее вероятным числом успехов является

- **а)** единственное число $m_0 = [\boldsymbol{np + p}]$ (целая часть), если число $\boldsymbol{np + p}$ не целое;
- **б)** два числа

$m_0 = \boldsymbol{np + p}$ и $m_0' = \boldsymbol{np + p - 1}$, если число $\boldsymbol{np + p}$ целое.

Пример

Вычислить вероятности всех возможных значений появления «герба» при 5 бросаниях монеты. Построить график распределения этих вероятностей.

Решение

Число независимых испытаний $n = 5$.

Число успехов $m = 0, 1, 2, 3, 4, 5$.

Вероятность успеха в одном испытании $p = 0{,}5$.

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

$n = 5, m = 0, 1, 2, 3, 4, 5$

$$P_5(0) = C_5^0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32} = 0.03125,$$

$$P_5(1) = C_5^1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{32} = 0.15625,$$

$$P_5(2) = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{32} = 0.31250,$$

и т.д.

- **Наивероятнейшее число успехов:**

Вычисляем $np + p$ = **5·1/2 + ½ = 3.**

Это целое число, поэтому

$$m_0 = np + p = 3 \quad \text{и} \quad m_0' = np + p - 1 = 2.$$

Самые большие (и равные между собой) вероятности у двух и трех появлений герба.

m	0	1	2	3	4	5
$P_n(m)$	0.03125	0.15625	0.31250	0.31250	0.15625	0.03125

Еще один пример

- Вероятность сдать экзамен равна 0,8. Найти наивероятнейшее число студентов, сдавших экзамен в группе из 30 человек.

Решение.

Вычисляем $\mathbf{np + p = 30 \cdot 0{,}8 + 0{,}8 = 24{,}8}$.

Это не целое число, поэтому

$\mathbf{m_0 = [24{,}8] = 24}$.

Полиномиальная схема

<u>Определение</u>

Полиномиальной схемой **называется последовательность *n* независимых одинаковых испытаний, в каждом из которых возможны *k* исходов**

$$A_1, A_2, \ldots, A_k, \quad \bigcup_{i=1}^{k} A_i = \Omega,$$

при этом вероятность любого исхода в каждом испытании постоянна,

$$P\left(A_i^{\nu}\right) = p_i, \quad \nu = \overline{1,n}, \quad i = \overline{1,k}, \quad \sum_{i=1}^{k} p_i = 1$$

Полиномиальная формула

$$P_n(m_1, m_2, \ldots m_k) =$$

$$= P\left\{\begin{array}{c}\text{событие } A_1 \text{ произошло ровно } m_1 \text{ раз,} \\ \ldots \\ \text{событие } A_k \text{ произошло ровно } m_k \text{ раз}\end{array}\right\} =$$

$$= \frac{n!}{m_1! \cdot m_2! \cdot \ldots \cdot m_k!} p_1^{m_1} \cdot p_2^{m_2} \cdot \ldots \cdot p_k^{m_k}.$$

Пример

- Человек с вероятностью 0,2 оказывается брюнетом, с вероятностью 0,4 —шатеном, с вероятностью 0,3 — блондином и с вероятностью 0,1—рыжим. Выбирается наугад группа из шести человек. Найти вероятность того, что в составе группы два брюнета, один шатен и три блондина.

$$P_6(2,1,3,0) =$$
$$= \frac{6!}{2!\cdot 1!\cdot 3!\cdot 0!!} 0,2^2 \cdot 0,4^1 \cdot 0,3^3 \cdot 0,1^0.$$

Гипергеометрические испытания

Пусть из совокупности $\mathbf{n}$ предметов, среди которых $\mathbf{n_1}$ предметов первого вида и $\mathbf{n_2}$ предметов второго вида ($\mathbf{n_1} + \mathbf{n_2} = \mathbf{n}$) производится выборка без возвращения $\mathbf{m}$ предметов, $1 \le \mathbf{m} \le \mathbf{n}$.

Вероятность того, что в выборке будет $\mathbf{m_1}$ предметов первого вида и $\mathbf{m_2}$ предметов второго вида ($\mathbf{m_1} + \mathbf{m_2} = \mathbf{m}$), согласно классическому определению вероятности, выражается формулой

Гипергеометрические вероятности

$$P_{n_1,n}(m_1,m)=\frac{C_{n_1}^{m_1}\cdot C_{n_2}^{m_2}}{C_n^m}$$

Данные испытания являются зависимыми.

Пример

- В урне 6 белых и 5 черных шаров.
  Из урны вынимают наугад 5 шаров. Найти вероятность, что два из них будут белыми, а три – черными.
- **Решение:**

$$m(A) = C_6^2 \cdot C_5^3$$

$$n = C_{11}^5$$

$$P(A) = \frac{C_6^2 \cdot C_5^3}{C_{11}^5}.$$

Теорема

Пусть $n \to \infty$ и $n_1 \to \infty$ так, что

$$\frac{n_1}{n} \to p, \quad 0 \le p \le 1$$

Тогда

$$\boxed{P_{n_1,n}(m_1, m) \to P_m(m_1).}$$

Доказательство

$$P_{n_1,n}(m_1,m)=\frac{C_{n_1}^{m_1}\cdot C_{n_2}^{m_2}}{C_n^m}=$$

$$=\frac{m!(n-m)!}{n!}\cdot\frac{n_1!}{m_1!(n_1-m_1)!}\cdot\frac{n_2!}{m_2!(n_2-m_2)!}=$$

$$\frac{m!}{m_1!\cdot m_2!}\cdot\frac{\frac{n_1}{n}\left(\frac{n_1}{n}-\frac{1}{n}\right)\ldots\left(\frac{n_1}{n}-\frac{m_1-1}{n}\right)\cdot\frac{n_2}{n}\left(\frac{n_2}{n}-\frac{1}{n}\right)\ldots\left(\frac{n_2}{n}-\frac{m_2-1}{n}\right)}{\frac{n}{n}\left(\frac{n}{n}-\frac{1}{n}\right)\ldots\left(\frac{n}{n}-\frac{m-1}{n}\right)}\to$$

$$\to\frac{m!}{m_1!\cdot m_2!}\cdot p^{m_1}\cdot(1-p)^{m_2}=P_m(m_1)\quad\Diamond$$

Предельные теоремы для схемы Бернулли

При числе испытаний, превышающем 20, вычисление точного значения $P_n(m)$ затруднительно. В этих случаях применяют приближенные формулы, вытекающие из предельных теорем.

Различают два случая:

- когда p мало, используют приближение Пуассона,
- когда p не мало (и не очень близко к единице), справедливо приближение Муавра –Лапласа.

Существует область, в которой возможно применение обоих приближений.

Теорема Пуассона

Если n → ∞, p → 0 так, что np → λ, 0 < λ < ∞, то для любого фиксированного m∈N справедливо:

$$\boxed{P_n(m) = C_n^m p^m (1-p)^{n-m} \to p_\lambda(m) = \frac{\lambda^m e^{-\lambda}}{m!}}$$

Доказательство

Пусть np $=\lambda_n$. Тогда

$$P_n(m) = C_n^m p^m (1-p)^{n-m} =$$

$$= \frac{n!}{(n-m)!\,m!} \cdot \left(\frac{\lambda_n}{n}\right)^m \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-m} =$$

$$= \frac{n(n-1)\cdots(n-m+1)}{m!} \cdot \left(\frac{\lambda_n}{n}\right)^m \cdot \left(1 - \frac{\lambda_n}{n}\right)^{n-m} =$$

$$= \frac{\lambda_n^m}{m!} \cdot \left(1 - \frac{\lambda_n}{n}\right)^n \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right) \cdot \left(1 - \frac{\lambda_n}{n}\right)^{-m}$$

При n $\to \infty$, $\lambda_n = np \to \lambda$

$$\left(1-\frac{\lambda_n}{n}\right)^n \to e^{-\lambda},$$

$$\left(1-\frac{1}{n}\right)\to 1,\ \left(1-\frac{2}{n}\right)\to 1, \ldots ,\left(1-\frac{m-1}{n}\right)\to 1,$$

$$\left(1-\frac{\lambda_n}{n}\right)^{-m} \to 1.$$

Следовательно,

$$P_n(m) = C_n^m p^m (1-p)^{n-m} \to p_\lambda(m) = \frac{\lambda^m e^{-\lambda}}{m!} \quad \Diamond$$

Приближенная формула Пуассона

$$\boxed{P_n(m) \approx p_\lambda(m) = \frac{\lambda^m e^{-\lambda}}{m!}}$$

где λ = np. Приближенную формулу Пуассона применяют при

$n > 30$,

$p < 0.1$,

$0.1 < \lambda = np < 10$.

Пример (дни рождения)

Какова вероятность, что среди 500 случайно выбранных людей ни один не родился 1 января ?

Решение

По формуле Бернулли

$$P_{500}(0) = C_{500}^{0} p^{0} (1-p)^{500} = \left(\frac{364}{365}\right)^{500} \approx 0.2537$$

- По приближенной формуле Пуассона

$$\lambda = np = 500 \cdot \frac{1}{365} \approx 1.3699$$

$$P_{500}(0) = \frac{\lambda^0 e^{-\lambda}}{0!} \approx e^{-1.3699} \approx 0.2541$$

Предельная теорема Муавра –Лапласа

Если при n→ ∞ и постоянном p, не равном 0 или 1, величина

$$x_m = \frac{m - np}{\sqrt{npq}}$$

ограничена так, что $-\infty < a \leq x_m \leq b < +\infty$, то

$$P_n(m) = \frac{\varphi(x_m)}{\sqrt{npq}}\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right),$$

$$\text{где} \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Доказательство этой теоремы основано на применении формулы Стирлинга

$$n! = n^n e^{-n} \sqrt{2\pi n e^{O\left(\frac{1}{n}\right)}}$$

Локальная приближенная формула Муавра – Лапласа

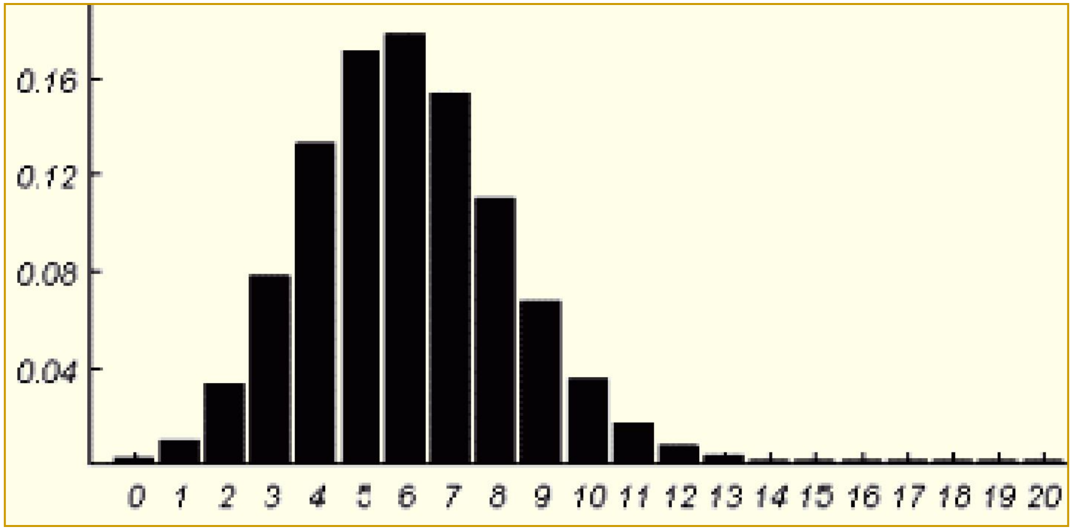
$$\boxed{P_n(m) \approx \frac{\varphi(x_m)}{\sqrt{npq}}}$$

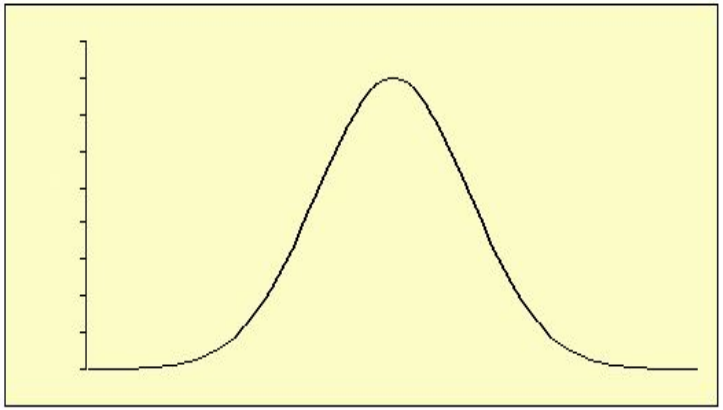
$$\boxed{x_m = \frac{m - np}{\sqrt{npq}}}$$

Локальную приближенную формулу Муавра – Лапласа применяют при

$n > 30$, $0.1 \le p \le 0.9$, $npq > 9$.

График биномиальных вероятностей при n=30, p=0,2 и график φ(X)





Интегральная предельная теорема Муавра –Лапласа

При $n \to \infty$ и постоянном p, не равном 0 или 1,

$$\lim_{n\to\infty} p\left(x_1 \le \frac{m-np}{\sqrt{npq}} \le x_2\right) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx = \Phi(x_2) - \Phi(x_1),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{x} \varphi(t) dt$$

Доказательство

$$p\left(x_1 \le \frac{m-np}{\sqrt{npq}} \le x_2\right) =$$

$$= p\left(x_1 \cdot \sqrt{npq} + np \le m \le x_2 \cdot \sqrt{npq} + np\right) =$$

$$= \sum_{m=x_1\cdot\sqrt{npq}+np}^{x_2\cdot\sqrt{npq}+np} p_n(m) = \sum_{x_m=x_1}^{x_m=x_2} p_n(m)$$

$$x_m = \frac{m-np}{\sqrt{npq}}$$

По локальной предельной теореме

$$\sum_{x_m=x_1}^{x_m=x_2} p_n(m) = \frac{1}{\sqrt{2\pi}} \sum_{x_m=x_1}^{x_m=x_2} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{npq}} (1+\alpha_n) = I_n + A_n$$

$$I_n = \frac{1}{\sqrt{2\pi}} \sum_{x_m=x_1}^{x_m=x_2} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{npq}} = \sum_{x_m=x_1}^{x_m=x_2} \frac{1}{\sqrt{npq}} \varphi(x_m)$$

$$x_m = \frac{m - np}{\sqrt{npq}}, \quad \Delta x_m = \frac{m + 1 - np}{\sqrt{npq}} - \frac{m - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}$$

$$I_n = \sum_{x_m = x_1}^{x_m = x_2} \varphi(x_m)\Delta x_m = \int_{x_1}^{x_2} \varphi(x)dx$$

$$A_n = \sum_{x_m = x_1}^{x_m = x_2} \varphi(x_m)\, \Delta x_m \;\alpha_n$$

$$|A_n| \le \sum_{x_m = x_1}^{x_m = x_2} \varphi(x_m)\, \Delta x_m \;|\alpha_n| \le \frac{C}{\sqrt{n}} I_n$$

при n→ ∞ $A_n \to 0$

$$\lim_{n\to\infty} p\left(x_1 \le \frac{m-np}{\sqrt{npq}} \le x_2 \right) =$$

$$= \lim_{n\to\infty}\left(I_n + A_n\right) =$$

$$= \int_{x_1}^{x_2} \varphi(x)\,dx =$$

$$\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}}\,dx = \Phi(x_2) - \Phi(x_1). \ \Diamond$$

Интегральная приближенная формула Муавра –Лапласа

$$\lim_{n\to\infty} p\left(x_1 \le \frac{m-np}{\sqrt{npq}} \le x_2 \right) \approx \Phi(x_2) - \Phi(x_1),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{x} \varphi(t) dt$$

Интегральную приближенную формулу Муавра – Лапласа применяют при **n > 30, 0.1 ≤ p ≤ 0.9, npq > 9.**

Следствия

$$p(a \le m \le b) \approx \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

$$p\left(\alpha_1 \le \frac{m}{n} \le \alpha_2\right) \approx \Phi\left((\alpha_2 - p)\sqrt{\frac{n}{pq}}\right) - \Phi\left((\alpha_1 - p)\sqrt{\frac{n}{pq}}\right)$$

$$p\left(\beta_1 \le \frac{m}{n} - p \le \beta_2\right) \approx \Phi\left(\beta_2\sqrt{\frac{n}{pq}}\right) - \Phi\left(\beta_1\sqrt{\frac{n}{pq}}\right)$$

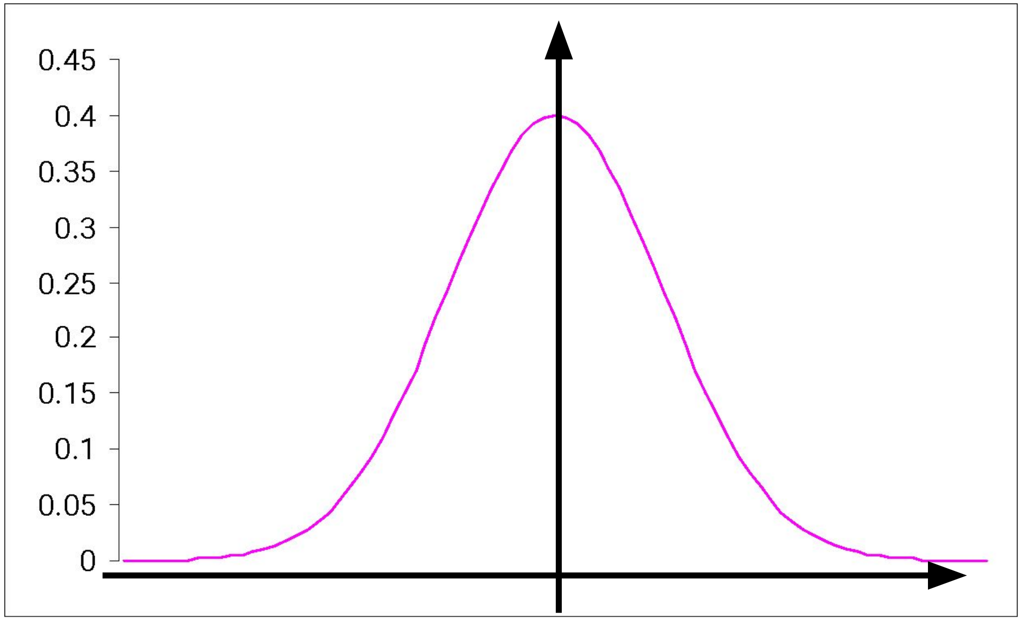
Свойства функции $\phi(x)$

$$\varphi(-x) = \varphi(x)$$

$$\varphi(0) = \frac{1}{\sqrt{2\pi}} \approx 0.3989$$

$$\lim_{x \to \pm\infty} \varphi(x) = 0$$

$$\varphi(\pm 4) < 0.001$$



Свойства функции Ф(x)

$$\Phi(x)+\Phi(-x)=1$$

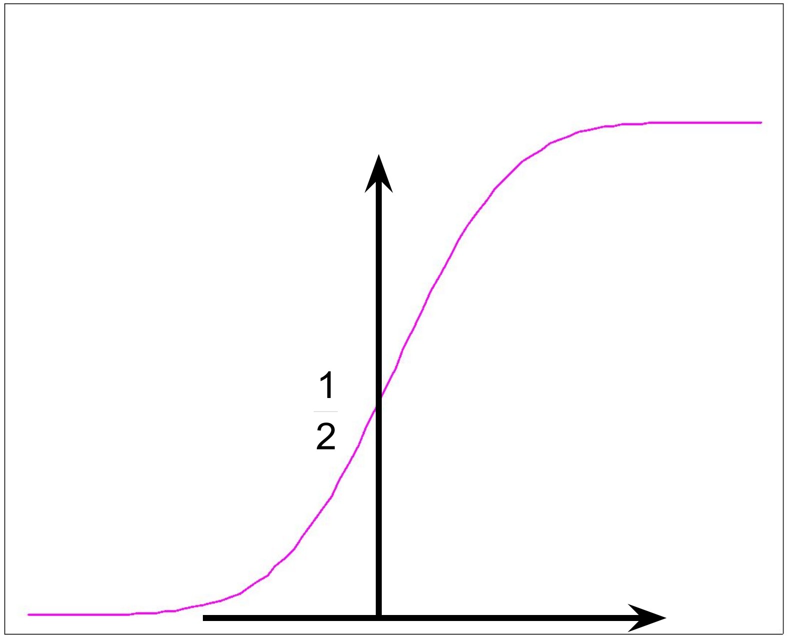
$$\lim_{x\to-\infty}\Phi(x)=0$$

$$\lim_{x\to+\infty}\Phi(x)=1$$

$$\Phi(0)=\frac{1}{2}$$

$$\Phi(3.8)>0.9999$$

$$\Phi(-3.8)<0.0001$$



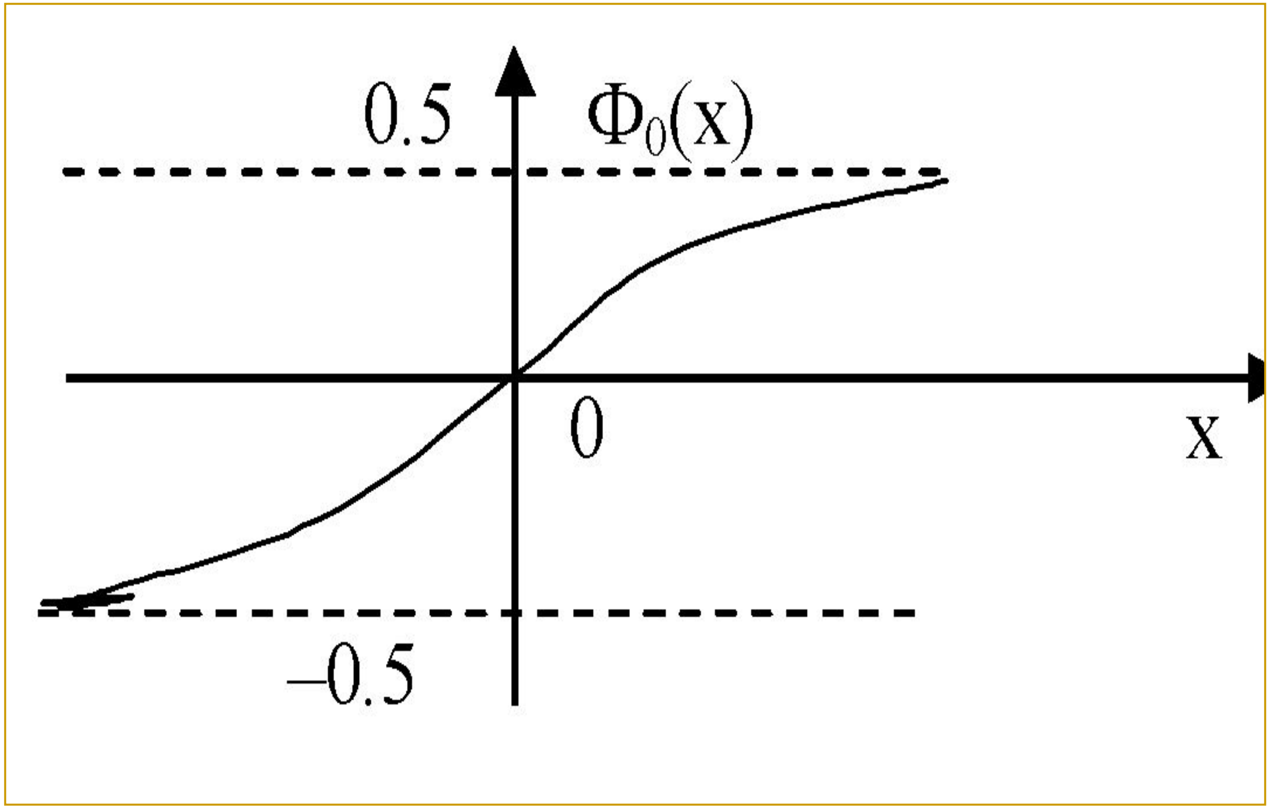
Функция Лапласа $\Phi_0(x)$.

- Вместо $\Phi(x)$ часто используют функцию Лапласа $\Phi_0(x)$.

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

$$\Phi(x) = \Phi_0(x) + \frac{1}{2}$$

График функции $\Phi_0(x)$



Замечания

$$\Phi_0(x_2) - \Phi_0(x_1) = \Phi(x_2) - \Phi(x_1),$$

- поэтому в формулах может использоваться как $\Phi(x)$, так и $\Phi_0(x)$.
- Значения функций находят в таблицах.

Пример

- **Вероятность рождения мальчика p = 0,5. Найти вероятность того, что в группе из 100 новорожденных мальчиков не меньше 60.**

Решение.

$$p\left(60\le m\le 100\right)\approx$$

$$\Phi\left(\frac{100-100\cdot 0,5}{\sqrt{100\cdot 0,5\cdot 0,5}}\right)-\Phi\left(\frac{60-100\cdot 0,5}{\sqrt{100\cdot 0,5\cdot 0,5}}\right)=$$

$$=\Phi\left(\frac{50}{\sqrt{25}}\right)-\Phi\left(\frac{10}{\sqrt{25}}\right)=\Phi(10)-\Phi(2)\approx 1-0,98=0,02.$$