

§ 4. Непрерывные случайные величины. Плотность распределения.

Важный класс случайных величин – непрерывные случайные величины, для которых, в отличие от дискретных, распределение вероятностей нельзя задать в виде таблицы распределения. Введем понятие непрерывной случайной величины.

Определение 8. Говорят, что случайная величина ξ имеет непрерывное распределение, если ее функция распределения $F_{\xi}(x)$ непрерывна по $x \in \mathbb{R}$.

Определение 9. Говорят, что случайная величина ξ имеет абсолютно непрерывное распределение, если существует такая неотрицательная функция $f_{\xi}(x)$, называемая *плотностью распределения вероятностей*, такая, что:

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt .$$

Интеграл понимается в смысле Римана, а в общем случае в смысле Лебега.

Непрерывность случайной величины подразумевает, что функция распределения не имеет точек разрыва. В качестве примеров рассмотрим несколько функций распределения: разрывная функция распределения представлена на рис. 2.1а, непрерывная функция распределения представлена на рис. 2.1б, непрерывная функция распределения представлена на рис. 2.1в.

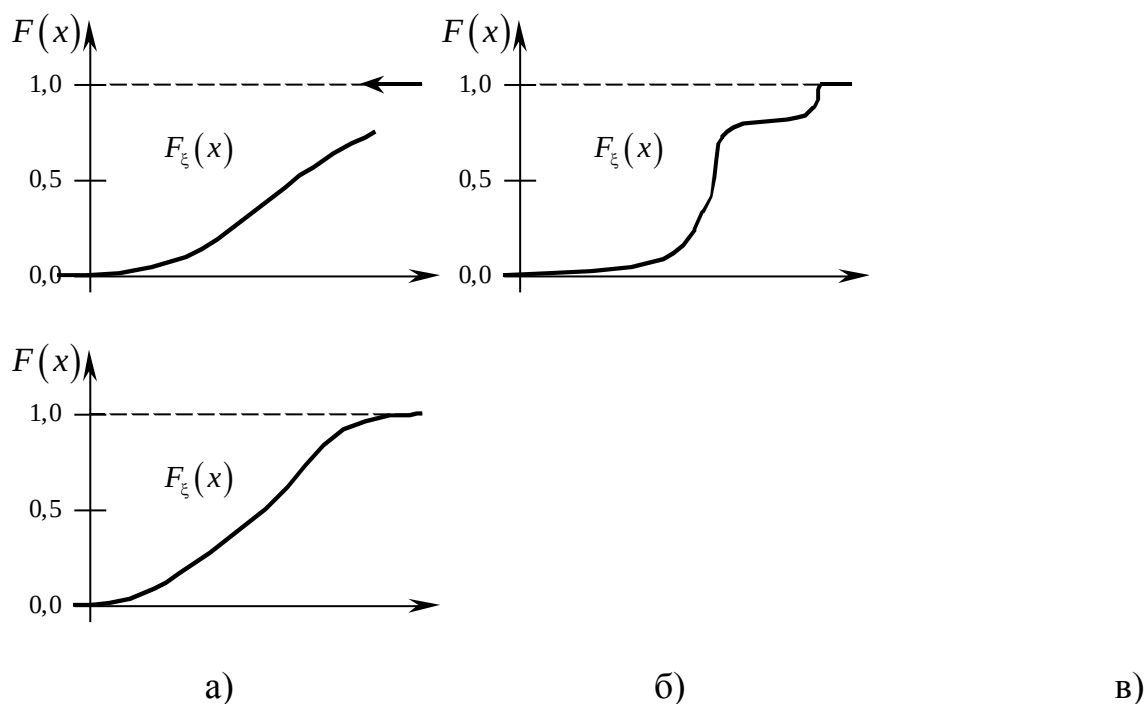


Рис. 2.1. Функция распределения:

а) разрывная; б) непрерывная; в) абсолютно непрерывная

Плотность распределения вероятностей (или плотность) $f_\xi(x)$ характеризует закон распределения случайной величины. Однако, значения плотности распределения вероятностей, которые она принимает в конкретных точках, не являются вероятностями. Если сравнивать с физическими величинами, то функцию распределения можно сравнить с массой, а плотность распределения вероятностей с плотностью веществ. Примеры плотностей распределения вероятностей представлены на рис. 2.2.

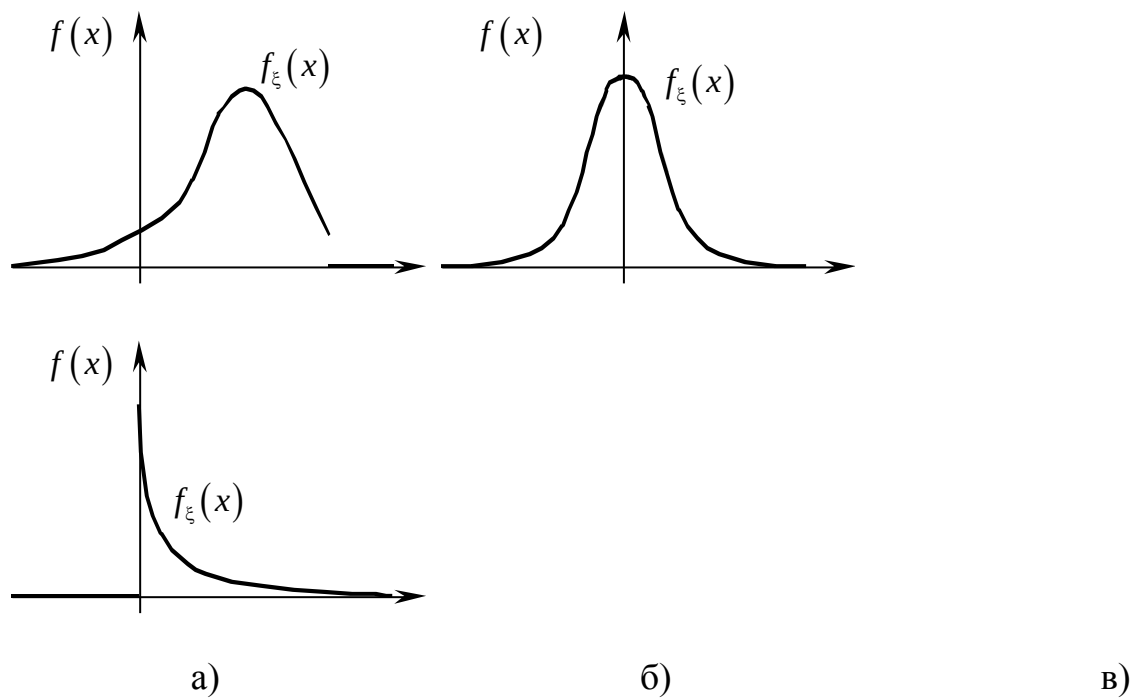


Рис. 2.2. Плотности распределения вероятностей

Значения плотности не ограничены верхним пределом. Рассмотрим некоторые свойства плотности.

Свойство 1. Свойство неотрицательности:

$$f(x) \geq 0 \quad \forall x \in \mathbb{R}.$$

Доказательство: по определению.

Свойство 2. Вероятность принадлежности случайной величины ξ интервалу $[x_1, x_2)$ вычисляется по формуле:

$$P(x_1 \leq \xi < x_2) = \int_{x_1}^{x_2} f_{\xi}(x) dx.$$

Доказательство: см. свойства функции распределения.

Свойство 3. Если случайная величина ξ имеет абсолютно непрерывное распределение, тогда $\forall x \in \mathbb{R}$ выполняется равенство:

$$P(\xi = x_2) = 0.$$

Доказательство: см. свойства функции распределения.

Свойство 4. Если случайная величина ξ имеет абсолютно непрерывное распределение, тогда $\forall x_1, x_2 \in \mathbb{R} : x_1 \leq x_2$ выполняется равенство:

$$P(x_1 \leq \xi < x_2) = P(x_1 < \xi \leq x_2) = P(x_1 < \xi < x_2) = P(x_1 \leq \xi \leq x_2).$$

Доказательство: см. свойства функции распределения.

Свойство 5. Если случайная величина ξ имеет абсолютно непрерывное распределение, тогда почти для всех $x \in \mathbb{R}$ функция распределения дифференцируема:

$$f_{\xi}(x) = F'_{\xi}(x) = \frac{d}{dx} F_{\xi}(x).$$

Доказательство:

Свойство 6.

$$\int_{-\infty}^{\infty} f_{\xi}(x) dx = 1$$

Доказательство: см. свойства функции распределения.

§ 5. Примеры непрерывных случайных величин

5.1. Равномерное распределение

Случайная величина ξ имеет равномерное распределение на отрезке $[a, b]$ если ее функцию распределения можно представить в виде

$$F_{\xi}(x) = P(\xi < x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b \end{cases}$$

плотность распределения в виде

$$f_{\xi}(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{b-a}, & a \leq x \leq b. \\ 0, & x > b \end{cases}$$

Принадлежность случайной величины равномерному распределению с параметрами a и b будем обозначать $\xi \in U:(a,b)$. Плотность и функция равномерного распределения случайной величины с параметрами $a=1$ и $b=5$ представлены на рис. 7а и 7б.

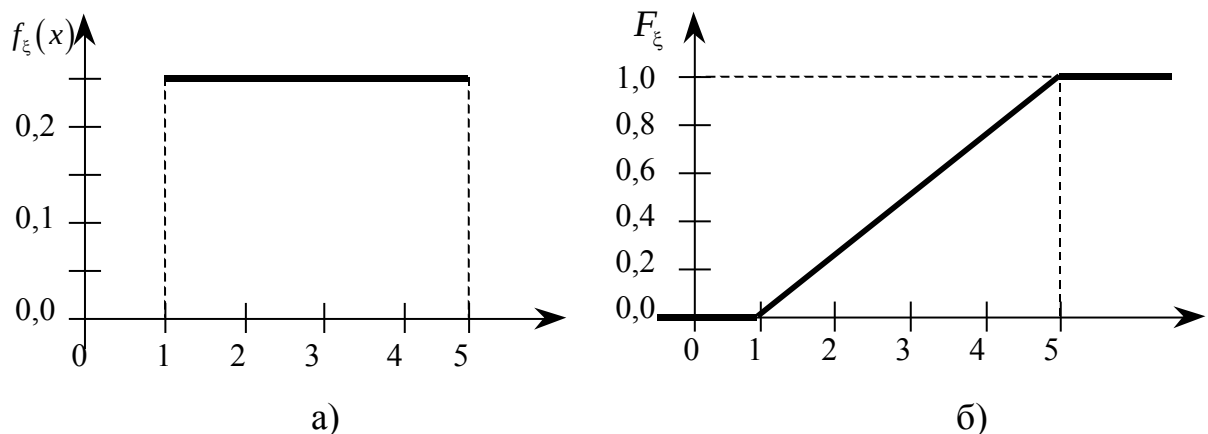


Рис. 7. Плотность (а) и функция (б) равномерного распределения с параметрами $a=1$ и $b=5$

Равномерное распределение с параметрами $a=0$ и $b=1$ имеет важное значение при моделирование случайных величин на компьютере. Их достаточно легко запрограммировать, а случайные величины из других семейств распределения можно получить с помощью функционального преобразования равномерно распределенных случайных величин.

5.2. Показательное распределение

Случайная величина ξ имеет показательное распределение с параметром α : $\alpha > 0$ если функция распределения имеет вид:

$$F_{\xi}(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\alpha x}, & x \geq 0 \end{cases},$$

а плотность распределения в виде

$$f_{\xi}(x) = \begin{cases} 0, & x < 0 \\ \alpha e^{-\alpha x}, & x \geq 0 \end{cases}.$$

Принадлежность случайной величины показательному распределению с параметром α будем обозначать $\xi \in E : \alpha$. Плотность и функция показательного распределения случайной величины с параметром $\alpha = 1$ представлены на рис. 8а и 8б.

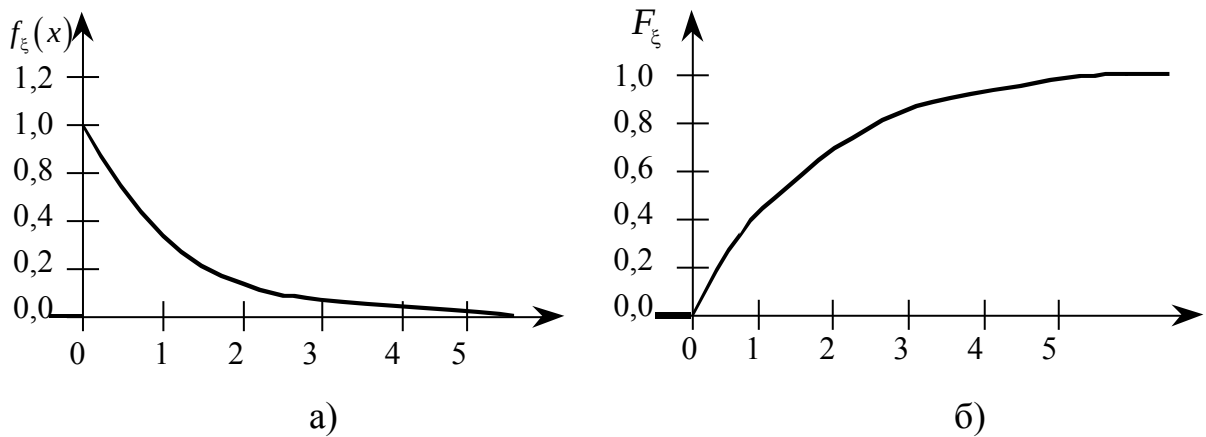


Рис. 8. Плотность (а) и функция (б) показательного распределения с параметром $\alpha = 1$

Показательное распределение нашло широкое применение при изучении случайных величин, связанных со временем, например, в теории надежности. Одним из важных свойств показательного распределения является, так называемое, свойство «нестарения». Сформулируем это свойство в виде теоремы.

Теорема 1. Пусть случайная величина $\xi \in E : \alpha$. Тогда $\forall x, y > 0$

$$P(\xi \geq x + y | \xi \geq x) = P(\xi \geq y).$$

Доказательство: Обозначим события $A = (\xi \geq x + y)$ и $B = (\xi \geq x)$. По определению условная вероятность может быть выражена в виде:

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

Событие $AB = (\xi \geq x + y)(\xi \geq x)$ представляет собой событие $A = (\xi \geq x + y)$. Это следует из того, что $y > 0$. Тогда вероятность события AB равна

$$P[(\xi \geq x+y)(\xi \geq x)] = P(\xi \geq x+y) = 1 - F_{\xi}(x+y) = 1 - [1 - e^{-\alpha(x+y)}] = e^{-\alpha(x+y)},$$

а вероятность события B равна

$$P[(\xi \geq x)] = 1 - F_{\xi}(x) = 1 - [1 - e^{-\alpha x}] = e^{-\alpha x}.$$

Таким образом, условная вероятность

$$P(\xi \geq x+y | \xi \geq x) = \frac{e^{-\alpha(x+y)}}{e^{-\alpha x}} = e^{-\alpha y} = 1 - [1 - e^{-\alpha y}] = 1 - F_{\xi}(y) = 1 - P(\xi < y) = P(\xi \geq y)$$

.

Теорема доказана.

5.3. Нормальное распределение

Случайная величина ξ имеет нормальное распределение с параметрами a и σ^2 : $a, \sigma \in \mathbb{R}$ и $\sigma > 0$ если плотность распределения имеет вид:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Принадлежность случайной величины нормальному распределению будем обозначать $\xi \in N:(a, \sigma^2)$.

Примечание. Для выполнения условия нормировки необходимо, чтобы функция распределения $F_{\xi}(x)$ была равна единице при $x \rightarrow \infty$. Покажем, что

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}. \text{ Рассмотрим интеграл:}$$

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right)^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+t^2}{2}} dx dt$$

Перейдем в цилиндрическую систему координат и заменим переменные x и y на $r \cos \varphi$ и $r \sin \varphi$. Пределы изменения радиуса r : $0 \leq r \leq \infty$ и угла φ : $0 \leq \varphi < 2\pi$. Якобиан $|J|$ преобразования в цилиндрическую систему координат $|J| = r$.

Таким образом, получаем

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+t^2}{2}} dx dt = \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2(\cos\varphi)^2 + r^2(\sin\varphi)^2}{2}} r dr d\varphi = \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} d\left(\frac{r^2}{2}\right) d\varphi = 2\pi,$$

следовательно

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

и функция распределения случайной величины $\xi \in N:(0,1)$ имеет вид:

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

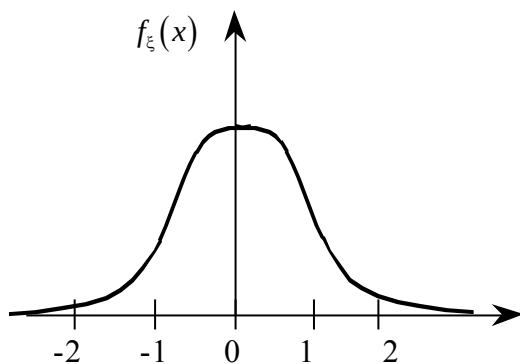
Замечание. Можно было воспользоваться интегралом Дирихле:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}.$$

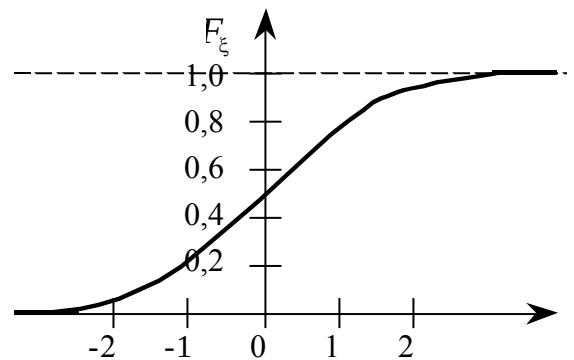
Покажем, что и при любых других параметрах нормального распределения выполняется условие нормировки:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{array}{l} t = \frac{x-a}{\sigma} \\ dx = \sigma dt \end{array} \right| = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \sigma dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1.$$

Отметим, что параметр a является характеристикой положения, а второй параметр σ^2 – характеристикой степени разброса случайной величины. Плотность и функция нормального распределения случайной величины с различными параметрами распределения представлены на рис. 9а и 9б.



а)



б)

Рис. 9. Плотность (а) и функция (б) нормального распределения

Свойства нормального распределения.

Введем определение квантили распределения.

Определение 10. Квантилем уровня q непрерывной случайной величины ξ называется такое значение u_q этой случайной величины для которого выполняется условие:

$$F_{\xi}(u_q) = P(\xi < u_q) = q.$$

Квантиль уровня q случайной величины равна значению, при котором площадь плотности распределения от $-\infty$ до u_q равна q (см. рис. 10).

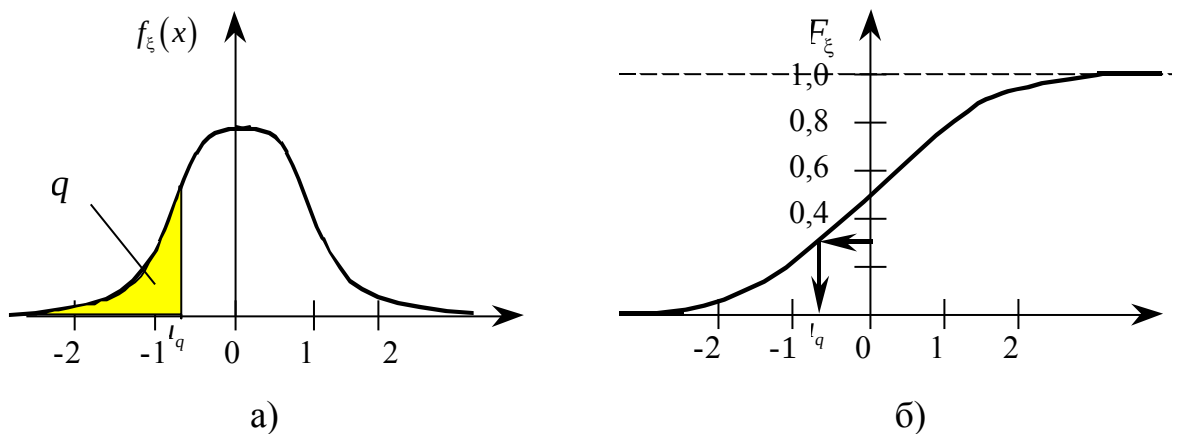


Рис. 10. Вычисление квантили уровня q

Таким образом, квантиль распределения это решение уравнения

$$F_{\xi}(u_q) = q,$$

где $F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t)dt$ и заданном значении q .

Табулированные значения квантилей распределений (нормального, χ^2 распределения, распределения Фишера и т.п.) приводятся в приложениях практически всех учебников по теории вероятностей и математической статистике (например, С.А. Айвазян, В.С. Мхитарян «Теория вероятностей и

прикладная статистика»). В том числе и для стандартного нормального распределения, т.е. распределения случайной величины $\eta \in N:(0,1)$.

В некоторых учебниках (например, Гмурман) вместо квантилей стандартного нормального распределения в таблицах приводятся значения функции Лапласа:

$$\Phi(x) = F_{\eta}(x) - 0,5,$$

т.е. такой функции:

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

Чтобы найти значения квантиля u_q по функции Лапласа необходимо найти значение аргумента функции Лапласа, при котором выполняется следующее равенство:

$$\Phi(u_q) = q - 0,5.$$

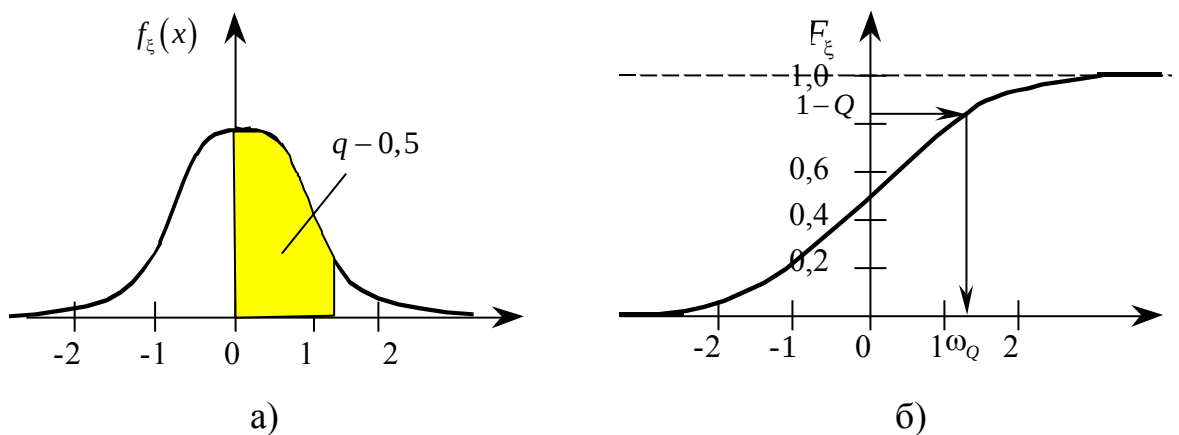


Рис. 11. Вычисление квантилей стандартного нормального распределения с помощью функции Лапласа

Одно из свойств функции Лапласа можно применить для вычисления квантиля уровня $q < 0,5$, а именно:

$$\Phi(-x) = -\Phi(x),$$

т.е. если $q < 0,5$, то для вычисления квантили u_q можно сначала найти

$$\Phi(u_q) = 0,5 - q$$

и взять значение u_q с отрицательным знаком.

Наряду с квантилем уровня q введем понятие $Q \cdot 100\%$ процентной точки, которая равна значению, при котором площадь плотности распределения от ω_Q до ∞ равна Q (см. рис. 11).

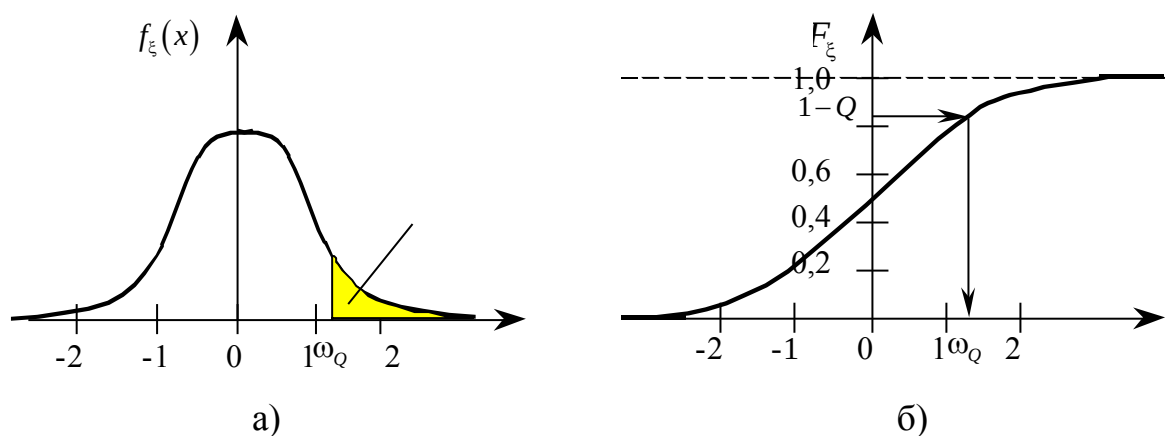


Рис. 11. Вычисление $Q \cdot 100\%$ процентной точки уровня q

Значения квантилей, функции Лапласа и процентных точек приводятся в табл. 1.

Таблица 1

Вычисление квантилей и процентных точек

$q = F_{\eta}(x)$	$q - 0,5 = \Phi(x)$	$Q = 1 - q =$ $= 1 - F_{\eta}(x)$	$u_q = \omega_Q$
0,00135	-0,49865	0,99865	-3,000
0,5	0	0	0
0,95	0,45	0,05	1,645
0,99	0,49	0,01	2,326
0,99865	0,49865	0,00135	3,000
0,999	0,499	0,001	3,090

Свойство 1. Если случайная величина $\xi \in N:(a, \sigma^2)$, а $\eta \in N:(0,1)$, тогда

$$F_{\xi}(x) = F_{\eta}\left(\frac{x-a}{\sigma}\right).$$

Доказательство:

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \left| \begin{matrix} y = \frac{t-a}{\sigma} \\ dt = \sigma dy \end{matrix} \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{y^2}{2}} dy = F_{\eta}\left(\frac{x-a}{\sigma}\right).$$

Свойство 2. Если случайная величина $\xi \in N:(a, \sigma^2)$, тогда случайная величина $\eta = \frac{\xi - a}{\sigma} \in N:(0,1)$ имеет стандартное нормальное распределение.

Доказательство: (доказательство приводится при изучении характеристических функций).

Свойство 3. Если случайная величина $\xi \in N:(a, \sigma^2)$, тогда $F_{\xi}(a) = 0,5$.

Доказательство: из свойства симметрии плотности $f_{\xi}(x)$.

Свойство 4. Если случайная величина $\xi \in N:(a, \sigma^2)$, а $\eta \in N:(0,1)$, тогда

$$P(x_1 < \xi < x_2) = F_{\eta}\left(\frac{x_2 - a}{\sigma}\right) - F_{\eta}\left(\frac{x_1 - a}{\sigma}\right).$$

Доказательство: см. предыдущее свойство и свойство 3 функции распределения.

Свойство 5 «Правило трех сигм». Если случайная величина $\xi \in N:(a, \sigma^2)$, тогда $P(|\xi - a| < 3\sigma) = 0,9973$, т.е. если случайная величина имеет нормальное распределение, то с высокой долей вероятности можно говорить, что все значения случайной величины находятся в пределах 3σ от значения a .

Доказательство: вычислить по таблицам.