Лекція 8. Нормальний розподіл.

План лекції

Питання, що розглядаються:

Нормальний розподіл, стандартний нормальний розподіл, функція розподілу, щільність функції розподілу, функція Гауса, функція Лапласа, ймовірність попадання нормальної випадкової величини в заданий інтервал, відхилення нормальної випадкової величини від її математичного сподівання, правило "трьох сигм".

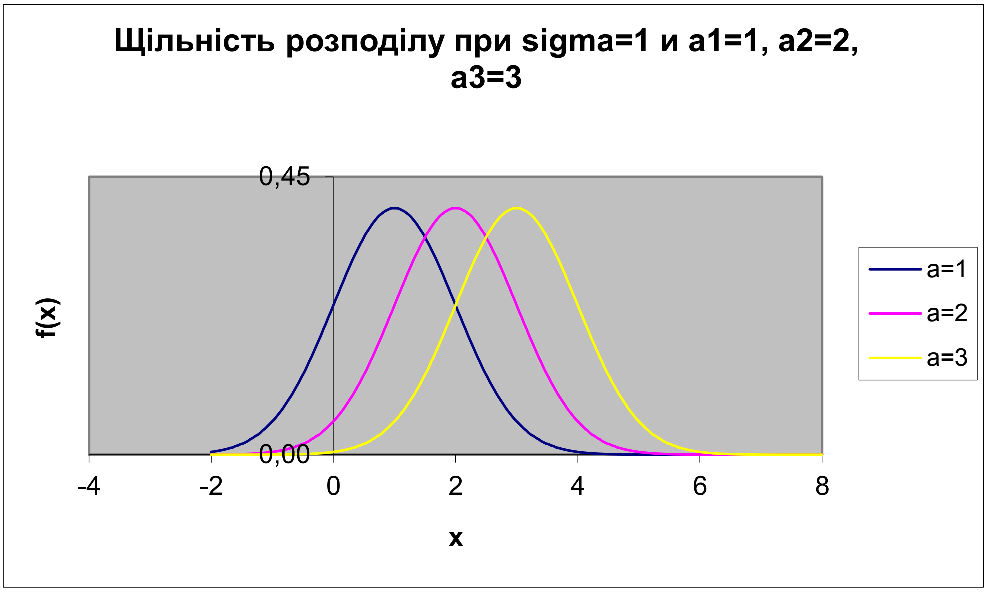
1. Означення нормального розподілу

Неперервна випадкова величина має нормальний закон розподілу з параметрами a і σ, якщо її щільність ймовірності має вигляд функції Гауса

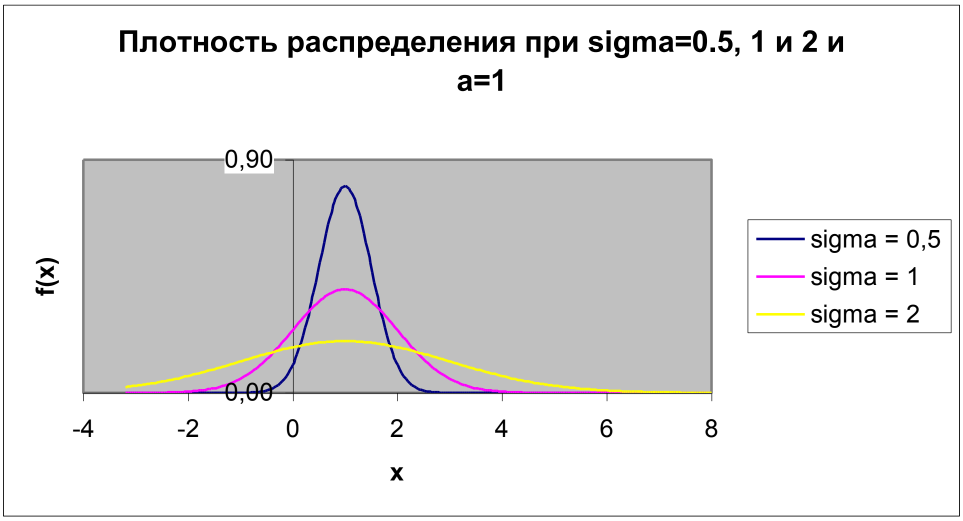
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Де $\sigma > 0$. За допомогою безпосереднього обчислення математичного сподівання і дисперсії нормального розподілу легко з'ясувати імовірнісний зміст його параметрів : a - є математичне сподівання, а σ - середнє квадратичне відхилення нормального розподілу. При $a = 0$, $\sigma = 1$ розподіл називається ***стандартним нормальним розподілом***.

Графіки для ряду конкретних значень математичного сподівання і середнього квадратичного відхилення приведені нижче.



Мал. 1. Зміна вигляду функції $f(x)$ при зміні математичного сподівання

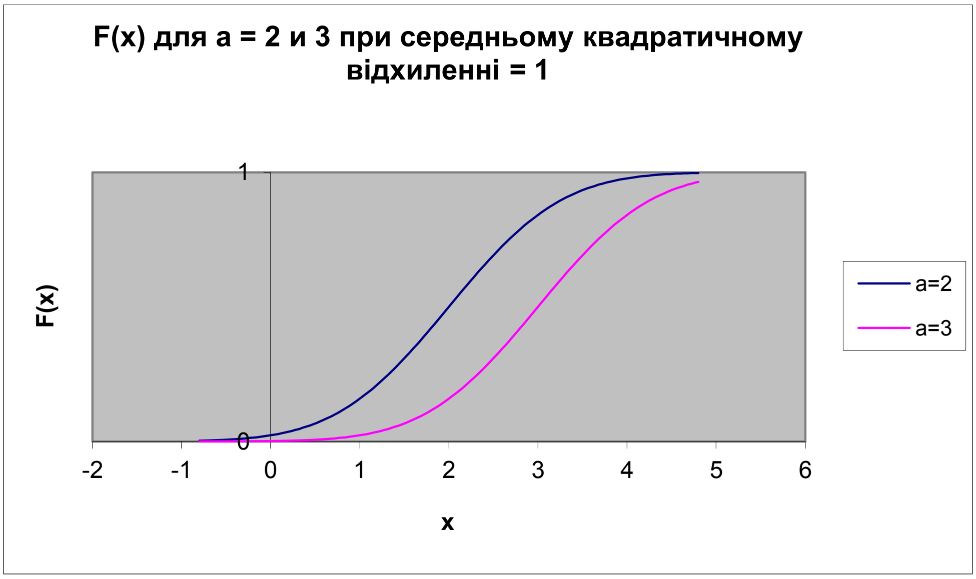


Мал. 2. Зміна вигляду функції $f(x)$ при зміні середнього квадратичного відхилення

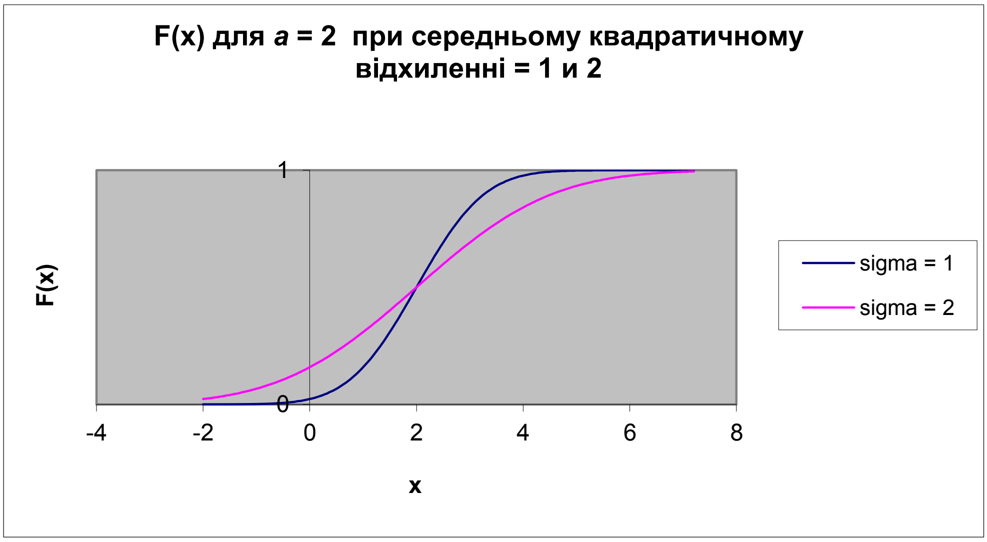
Функція розподілу у випадку нормального розподілу, очевидно, дорівнює

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt .$$

Графіки функції для ряду значень математичного сподівання і середнього квадратичного відхилення зображені на малюнках, що наводяться нижче



Мал. 3. Залежність функції розподілу від величини a



Мал. 4. Залежність функції розподілу від величини σ

Нормальний розподіл має виключно важливе значення для практичних застосувань, оскільки багато неперервних випадкових величин описуються саме цим розподілом. Виявляється, що підсумовування великого числа випадкових величин з різними законами розподілу призводить до нормального розподілу результуючої суми. Ця властивість підтверджується центральною граничною теоремою (***теорема Ляпунова***). Зміст цієї теореми полягає в наступному. Якщо випадкова величина X є сумою дуже великого числа взаємно незалежних випадкових величин, вплив кожної з яких на усю суму дуже малий, то X має розподіл, близький до нормального.

Слід мати на увазі, що при посиленні впливу окремих факторів можуть з'являтися відхилення від нормального розподілу результуючого параметра. Тому велике значення на практиці приділяється експериментальній перевірці висунених гіпотез, у тому числі і гіпотези про нормальний розподіл.

2. Властивості функції Гауса.

Графік щільності нормального розподілу називають ***нормальною кривою Гауса.***

Досліджуємо поведінку функції щільності ймовірності $f(x)$.

1. Очевидно, що функція визначена на усій осі x.

2. Функція набуває лише додатних значень, тобто нормальна крива розташована над віссю Ox.
3. Вісь Ox служить горизонтальною асимптотою графіка. Інших асимптот у графіка немає.
4. При $x = a$ функція має максимум, рівний $\varphi(x)_{\max} = \dfrac{1}{\sigma\sqrt{2\pi}}$.
5. Функція парна : її графік симетричний відносно прямої $x = a$
6. При $x = a \pm \sigma_X$ графік функції має точки перегину.

При будь-яких значеннях параметрів a і σ, площа, обмежена нормальною кривою і віссю x, дорівнює одиниці.

3. Ймовірність попадання нормальної випадкової величини в заданий інтервал.

Часто вимагається визначити ймовірність попадання випадкової величини в заданий інтервал. Ця ймовірність може бути виражена у вигляді різниці функції розподілу ймовірності в граничних точках цього інтервалу :

$$P(x_1 < X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx.$$

У разі нормального розподілу

$$P(x_1 < X < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

Використовуючи заміну змінної : $t = \dfrac{x-a}{\sigma}$, $x = t\sigma + a$, $dx = \sigma dt$,

отримаємо

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-\frac{t^2}{2}} dt,$$

де $t_1 = \dfrac{x_1 - a}{\sigma}$, $t_2 = \dfrac{x_2 - a}{\sigma}$.

Розіб'ємо отриманий інтеграл на два:

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}}\left(\int_{t_1}^{0} e^{-\frac{t^2}{2}}dt + \int_{0}^{t_2} e^{-\frac{t^2}{2}}dt\right) = \frac{1}{\sqrt{2\pi}}\left(\int_{0}^{t_2} e^{-\frac{t^2}{2}}dt - \int_{0}^{t_1} e^{-\frac{t^2}{2}}dt\right)$$

Тоді шукана ймовірність може бути виражена у вигляді:

$$P(x_1 < X < x_2) = \Phi_0(t_2) - \Phi_0(t_1)$$

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

Функція Лапласа протабульована, що істотно спрощує розрахунок попадання нормально розподіленої випадкової величини в будь-який заданий інтервал.

4. Функція Лапласа і її властивості.

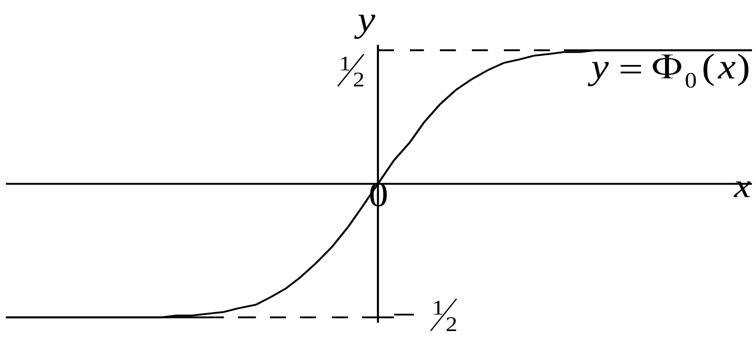
Функція Лапласа не виражається через елементарні функції:.

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

Для її обчислення використовуються спеціальні таблиці або методи наближеного обчислення.

Функція $\Phi_0(x)$ має наступні властивості:

1. $\Phi_0(0) = 0$;
2. $\Phi_0(+\infty) = \frac{1}{2}$;
3. функція $\Phi_0(x)$ - непарна, тобто $\Phi_0(-x) = -\Phi_0(x)$, тому в таблицях зазвичай приводяться значення $\Phi_0(x)$ тільки для додатних ;
4. функція $\Phi_0(x)$ - монотонно зростаюча функція (це випливає з того, що $\Phi_0'(x) = \varphi_0(x) > 0$).



5. Відхилення нормальної випадкової величини від її математичного сподівання. Правило "трьох сигм".

Часто вимагається обчислити ймовірність того, що відхилення нормально розподіленої випадкової величини за абсолютною величиною від математичного сподівання менше заданого додатного числа δ, тобто вимагається знайти ймовірність того, що виконується нерівність .

$|X - M_X| < \delta$

Замінимо цю нерівність рівносильною їй подвійною нерівністю .

$M_X - \delta < X < M_X + \delta$

Скористаємося формулою: $P(x_1 < X < x_2) = \Phi_0(t_2) - \Phi_0(t_1)$

Отримаємо: $P(M_X - \delta < X < M_X + \delta) =$

$$= \Phi_0\left(\frac{(M_X + \delta) - M_X}{\sigma}\right) - \Phi_0\left(\frac{(M_X - \delta) - M_X}{\sigma}\right)$$

$$= \Phi_0\left(\frac{\delta}{\sigma}\right) - \Phi_0\left(-\frac{\delta}{\sigma}\right) = 2\Phi_0\left(\frac{\delta}{\sigma}\right)$$

Якщо в якості δ узяти потрійне значення середнього квадратичного відхилення, то отримаємо:

$$P(|X - M_X| < 3\sigma) = 2\Phi_0(3) = 0{,}9973,$$

Таким чином, ймовірність того, що абсолютна величина відхилення перевищить потрійне середнє квадратичне відхилення, дуже мала (0,0027 або 0,27%). Такі події можна вважати практично неможливими.

Іншими словами, якщо випадкова величина розподілена нормально, то абсолютна величина її відхилення від математичного сподівання не перевищує потрійного середнього квадратичного відхилення. У цьому і полягає ***суть правила "трьох сигм".***

На практиці правило "трьох сигм" застосовують так: якщо розподіл випадкової величини, що вивчається, невідомий, але правило "трьох сигм" виконується, то є підстава вважати, що величина, що вивчається, розподілена нормально, і навпаки.

Приклад 1. Поточна ціна цінного паперу є нормально розподіленою випадковою величиною з середнім 100 у.о. і дисперсією 9. Знайти ймовірність того, що ціна активу знаходитиметься в межах від 91 до 109 у.о.

Розвязання. Оскільки $a=100$, $\sigma=\sqrt{9}=3$, то

$$P(91< X <109)=\Phi_0\left(\frac{109-100}{3}\right)-\Phi_0\left(\frac{91-100}{3}\right)=\Phi_0(3)-\Phi_0(-3)=2\Phi_0(3)=0.9973$$

Приклад 2. Магазин продає чоловічі костюми. За даними статистики відомо, що розподіл за розмірами є нормальним з математичним очікуванням і середнім квадратичним відхиленням, рівними 48 і 2 відповідно. Визначити відсоток попиту на 50-й розмір, за умови розкиду значень цієї величини в інтервалі (49,51).

Розвязання. За умовою, a=48, σ = 2, α = 49, β =51. Використовуючи формулу

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)\,dx$$

отримуємо, що ймовірність попиту на 50-й розмір в заданому інтервалі дорівнює

$$P(49< X <51) = \Phi_0\left(\frac{51-48}{2}\right) - \Phi_0\left(\frac{49-48}{2}\right) =$$
$$= \Phi_0(1{,}5) - \Phi_0(0{,}5) = 0{,}4332 - 0{,}1915 = 0{,}2417$$

Висновок: попит на 50-й розмір складе приблизно 24% і магазину треба передбачити це в загальному об'ємі закупівлі.

Питання для самоперевірки

1. Що таке нормальний розподіл?
2. Записати функцію розподілу для нормальної величини.
3. Записати властивості функції Гауса.
4. Як обчислюється ймовірність попадання нормальної величини в заданий інтервал?
5. Записати функцію Лапласа та її властивості.
6. За якою формулою обчислюється ймовірність відхилення нормальної випадкової величини від її математичного сподівання?
7. Сформулювати правило «трьох сигм».