

# Ймовірнісне моделювання.

## Метод статистичних випробувань.

У тих випадках, коли при моделюванні необхідно враховувати деякий випадковий фактор (елемент або явище), який неможливо описати аналітично, використовують метод моделювання під назвою **метода статистичних випробувань або метода Монте-Карло.**

За допомогою цього методу може бути вирішена будь-яка імовірнісна задача.

Однак використовувати його доцільно в тому випадку, якщо вирішити задачу цим методом простіше, ніж будь-яким іншим.

Суть методу полягає в тому, що замість опису випадкових явищ аналітичними залежностями проводиться розіграш випадкового явища за допомогою деякої процедури, яка дає випадковий результат.

За допомогою розіграшу отримують одну реалізацію випадкового явища.

Здійснюючи багаторазово такий розіграш, накопичують статистичний матеріал (множина реалізацій випадкової величини), який можна обробляти статистичними методами.

Розглянемо на прикладах.

**Приклад 1.** Нехай чотири стрілка одночасно стріляють по рухомій цілі. Ймовірність влучення в ціль кожним стрільцем дорівнює 0,5 (потрапив або не потрапив).

Ціль вважається ураженою, якщо в неї потрапило два або більше стрілка. Знайти ймовірність ураження цілі.

Це завдання можна легко вирішити методами теорії ймовірності:

Ймовірність поразки цілі  $P_{\text{ураж}} = 1 - P_{\text{неураж}}$

Ймовірність неураження  $P_{\text{неураж}}$  визначають як число комбінацій, коли в ціль не потрапив жоден стрілок, плюс потрапив один із стрільців:

$$P_{\text{неураж}} = 0,5^4 + C_4^1 \cdot 0,5^1 \cdot 0,5^3 = 0,5^4 + 4 \cdot 0,5^1 \cdot 0,5^3 = 0,3125$$

$$P_{\text{ураж}} = 1 - P_{\text{неураж}} = 1 - 0,3125 = 0,6875$$

Вирішимо цю задачу методом статистичних випробувань.

Процедуру розіграшу реалізуємо підкиданням одночасно чотирьох монет.

Якщо монета падає лицьовою стороною, то вважаємо, що стрілок влучив у ціль. Позначимо через  $m$  число успішних випробувань.

Зробимо  $N$  випробувань, тоді відповідно до теореми Бернуллі

$$P_{\text{ймов}} = m/N$$

У разі значного збільшення числа випробувань  $N$  частота ураження цілі буде збігатися з ймовірністю

$$p_{\text{ймов}} = 0.6875.$$

**Приклад 2.** Нехай є деяка ціль, на яку бомбардувальники скидають  $n$  бомб.

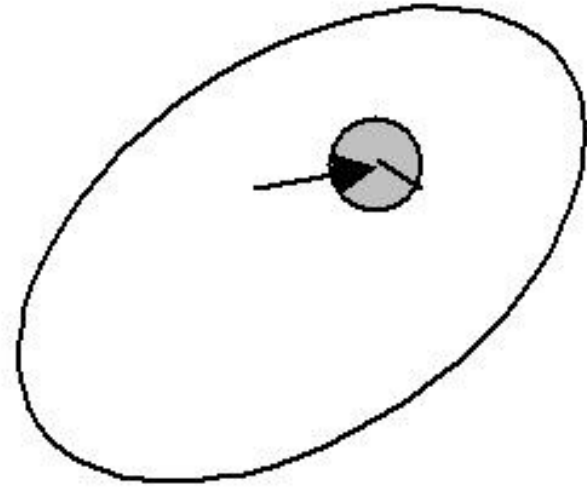
Кожна бомба вражає область у вигляді кола радіусом  $r$  (рис.).

Ціль вважається ураженою, якщо одночасно бомбами накрите  $K$  відсотків площі  $S$ .

Знайти ймовірність ураження цілі.

Аналітично вирішити цю задачу дуже важко.

Покажемо, як її можна вирішити методом статистичних випробувань.



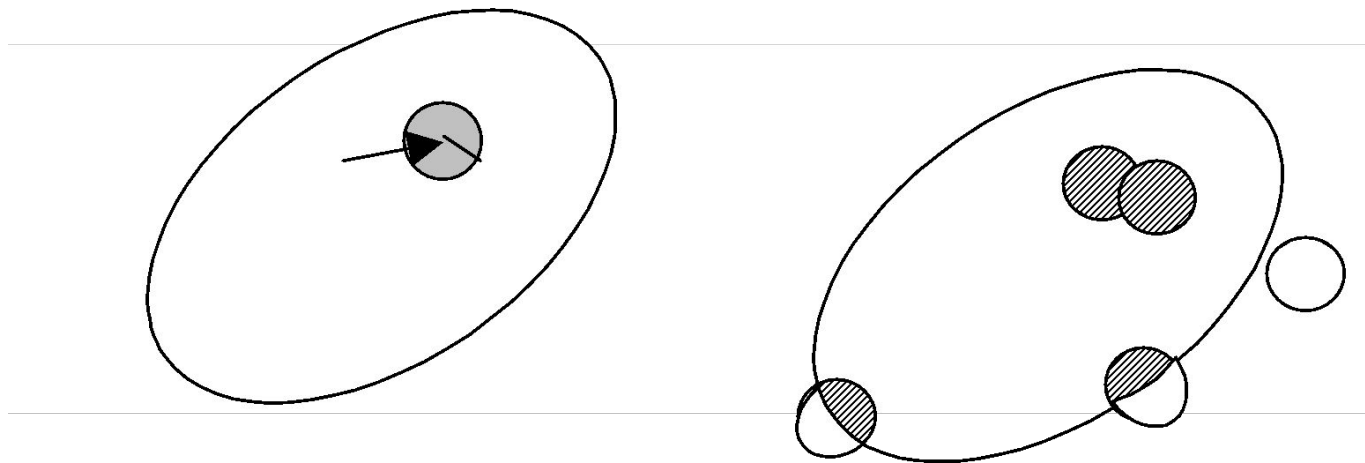
Накладемо координатну сітку на всю можливу область попадання бомб.

Розіграємо  $n$  точок - координат попадання бомб.

Наведемо біля кожної точки коло радіусом  $r$  (рис.) і визначимо заштриховану площу ураження.

Якщо заштрихована площа становитиме  $K$  відсотків і більше всієї площі цілі  $S$ , то ціль вважається ураженою, а випробування успішним.

В іншому випадку ціль не буде вражена і випробування не успішне.



Виконаємо  $N$  випробувань. Тоді ймовірність ураження цілі

$$P_{\text{ураж}} = \frac{m}{N}$$

де  $m$  – кількість випробувань, у яких ціль була вражена.

Методом статистичних випробувань можна оцінити математичне сподівання і інші ймовірні характеристики. Наприклад, оцінку математичного сподівання площі ураження цілі можна визначити як

$$M(S) = \frac{1}{N} \sum_{i=1}^N S_i$$

При  $N \rightarrow \infty$  ця оцінка буде наближатися до математичного сподівання у відповідності до закону великих чисел.

В цьому виразі  $S_i$  площа ураження в  $i$ -му випробуванні.

## **Алгоритм методу статистичних випробувань такий:**

1. Визначити, що собою являтиме випробування або розіграш.
2. Визначити, яке випробування є успішним, а яке - ні.
3. Провести велику кількість випробувань.
4. Обробити отримані результати статистичними методами і розрахувати статистичні оцінки шуканих величин.



До недоліків методу можна віднести необхідність проведення великої кількості випробувань, щоб отримати результат з заданою точністю.

Таким чином, метод статистичних випробувань - це метод математичного моделювання випадкових величин, в якому сама випадковість безпосередньо включена в процес моделювання і є його важливим елементом.

Кожен раз, коли в хід виконання деякої операції втручається випадковий фактор, його вплив моделюється за допомогою розіграшу.

Для ефективного розіграшу випадкових величин використовують генератори випадкових чисел.

Такі генератори будуються апаратними та програмними методами.

**Апаратні методи** генерування випадкових величин базуються на використанні деяких фізичних явищ (шумів електронних пристроїв, радіоактивного випромінювання).

Під час використання апаратних генераторів випадковий електричний сигнал перетворюється в двійковий код, який вводиться в комп'ютер за допомогою аналогово-цифрових перетворювачів.

Один з найбільш поширених способів - використання шумів електронних пристроїв.

Якщо на підсилювач не подавати ніякий сигнал і включити його на повну потужність, то буде чути шипіння (шум).

Цей безперервний сигнал можна перетворити в дискретний.

У більшості випадків його підсилюють і встановлюють граничне значення напруги шумового сигналу, перевищення якого можна вважати значенням двійковій одиниці на деякому малому проміжку часу  $t$ , в протилежному випадку отримуємо двійковий нуль.

Найбільш поширеними на практиці є **програмні генератори**.

Вони повинні відповідати таким вимогам:

- генерувати статистично незалежні випадкові числа, рівномірно розподілені в інтервалі  $[0,1]$ ;
- мати можливість відтворювати задані послідовності випадкових чисел;
- витрати ресурсів процесора на роботу генератора повинні бути мінімальними;
- легко створювати незалежні послідовності випадкових чисел (потоки).

# Лінійні конгруентні генератори.

У більшості сучасних програмних генераторів використовується властивість конгруентності, яке полягає в тому, що два цілих числа  $A$  і  $B$  є конгруентними за модулем  $m$ , якщо їх різниця  $(A-B)$  є число, яке ділиться на  $m$  без залишку.

Записується це так:

$$A = B \pmod{m}$$

Зазвичай використовується лінійний мультиплікативний конгруентний метод, рекурентне співвідношення для якого має вигляд

$$X_i = aX_{i-1} \pmod{m}$$

де  $a$  та  $m$  – деякі константи.

Необхідно взяти останнє випадкове число  $X_{i-1}$ , помножити його на постійний коефіцієнт  $a$  і взяти модуль отриманого числа по  $m$ , тобто розділити на  $m$  і отримати залишок.

Цей залишок і буде наступним псевдовипадковим числом  $X_i$ .

Одержані за формулою значення  $X_i$  належать діапазону  $0 \leq X_i \leq m-1$  і мають рівномірний дискретний розподіл.

$$X_i = aX_i \pmod{m}$$

Щоб отримати випадкове значення з інтервалу  $[0,1]$ , треба  $X_i$  розділити на  $m$ .

Для двоїчного комп'ютера  $m = 2^g - 1$ , де  $g$  – довжина розрядної сітки.

Наприклад, для 32-розрядного комп'ютера  $m = 2^{31} - 1 = 2147483647$ , оскільки один розряд задає знак числа.

У мові GPSS World використовується мультиплікативний конгруентний алгоритм Лемера з максимальним періодом, який генерує 2147483647 унікальних випадкових чисел без повторення.

Ці числа генерують спеціальні генератори, які позначаються RN <№>, де № - номер генератора випадкових чисел (може приймати значення від 1 до 7).

Під час відвідування таких генераторів видаються цілі випадкові числа в діапазоні від 0 до 999 включно.

При використанні генераторів в випадкових функціях розподілів випадкові числа генеруються в діапазоні від 0 до 0,999999 включно.



# Перевірка послідовностей випадкових чисел.

Статистичні властивості всіх послідовностей випадкових чисел потрібно перевіряти. Для цього використовують критерії. Найбільш часто використовують такі критерії (тести):

**Частотний** - з використанням критерію Колмогорова-Смирнова або критерію  $\chi^2$ ;

**Автокореляційний** – з вимірюванням кореляції між  $X_i$  и  $X_{i+k}$ , де  $k$  – зсув по послідовності ( $k = 1, 2, \dots$ );

**Серіальний** – з фіксацією частоти появи всіх можливих комбінацій чисел і виконанням оцінювання по критерію  $\chi^2$ ;

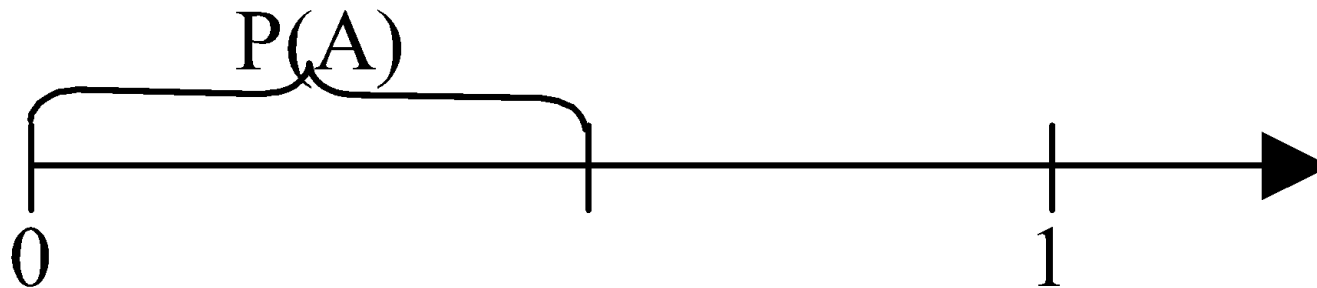
**Циклічний** – з перевіркою кількості циклів більше і менше деякої константи.

## Моделювання дискретних випадкових величин.

**Моделювання подій.** Нехай необхідно змоделювати появу деякої події  $A$ , ймовірність настання якої дорівнює  $P(A)=P$ .

Позначимо звертання до генератора, який розігрує випадкові, рівномірно розподілені на інтервалі  $(0, 1)$  числа  $r_i$ , через  $R$ .

Подія  $A$  при розіграші будет наставати тоді, коли  $r_i \leq P$  (рис.), а якщо  $r_i > P$ , то відбувається подія  $\bar{A}$ .



Дійсно, якщо  $f(r)$  – функція густини рівномірно розподіленої випадкової величини  $r$ ,

то

$$P(r < P) = \int_0^P f(x)dx = \int_0^P f(r)dr = P = P(A)$$

Даний метод використовується в мові GPSS для блока TRANSFER в статистичному режимі роботи.

## Моделювання групи несумісних подій.

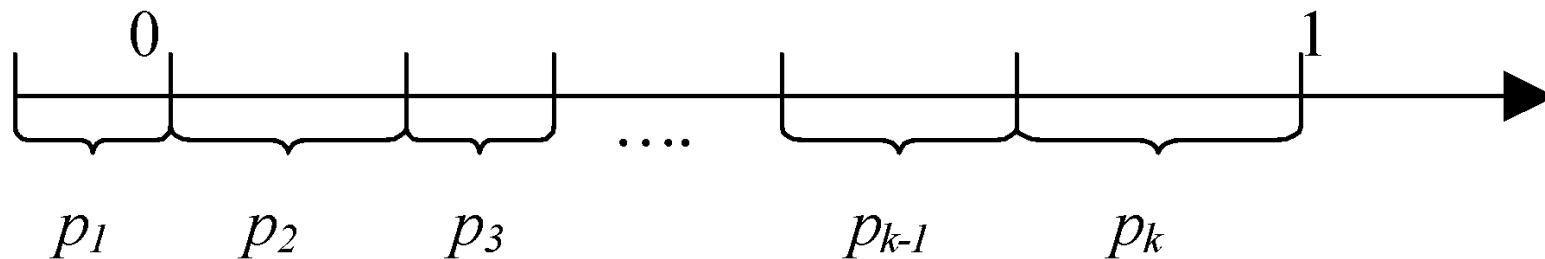
Нехай  $\epsilon$  група несумісних подій,  $A_1, A_2, \dots, A_k$ .

Відомі ймовірності настання подій  $P(A_1), P(A_2), \dots, P(A_k)$ .

Тоді через несумісність випробувань

$$\sum_{i=1}^k P(A_i) = 1$$

Нехай  $p_i = P(A_i), p_0 = 0$ . На відрізку  $(0, 1)$  відкладемо ці ймовірності (рис.).



Якщо одержане число потрапило в інтервал від  $\sum_{k=0}^{i-1} p_k$  до  $\sum_{k=0}^i p_k$ , то відбулася подія

$A_i$ .

Таку процедуру називають визначенням результату випробування по жеребу, і вона ґрунтується на формулі

$$P\left\{\sum_{k=0}^{i-1} p_k < r \leq \sum_{k=0}^i p_k\right\} = p_i = P(A_i)$$

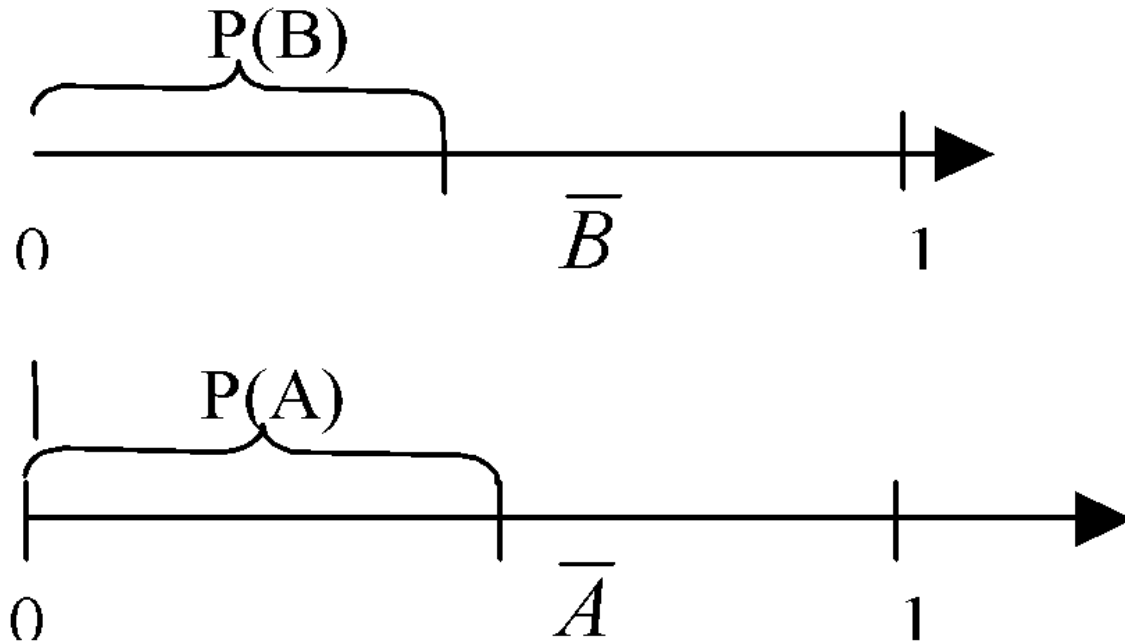
, де  $p_0 = 0$ .

# Моделювання умовної події.

Моделювання умовної події  $A$ , яке відбувається за умови, що настає подія  $B$  з ймовірністю  $P(A / B)$ , показано на рис.

Спочатку моделюємо подію  $B$ .

Якщо подія  $B$  відбувається, то моделюємо настання події  $A$ , якщо ж подія  $B$  не відбувається, то не моделюємо настання події  $A$ .



# Моделювання випадкової дискретної величини.

Моделювання випадкової дискретної величини виконується аналогічно моделюванню групи несумісних подій.

Дискретна випадкова величина  $X$  задається відповідно до таблиці, в якій вказані можливе значення величини і її ймовірність.

Випадкову величину  $X$  можна уявити як повну групу подій:

$$A_1 = (X = x_1), A_2 = (X = x_2), \dots, A_n = (X = x_n).$$

Даний метод використовується в мові GPSS для моделювання дискретних випадкових функцій розподілу.

# Геометричний розподіл.

Для моделювання випадкової величини  $X$  з геометричним розподілом необхідно задати таблицю її значень та їх ймовірність

Значення $X$	0	1	...	$n$
Ймовірності	$p$	$(1-p)p$		$(1-p)^n p$

Прикладом випадкової величини з таким розподілом може бути загальна кількість випробувань, які потрібно провести до першого успішного випробування, наприклад, кількість пострілів, які потрібно виконати до першого попадання в ціль.



# Біноміальний розподіл.

Біноміальний розподіл або розподіл Бернуллі - це розподіл випадкової величини, яка приймає тільки 2 значення: 1 - true (істина) і 0 - false (брехня).

Цей розподіл показує ймовірність настання деякої події за  $n$  незалежних повторних випробувань, в кожному з яких подія настає з імовірністю  $p$ , тобто ймовірність  $S$  успішних наслідків в  $n$  випробуваннях.

$$f(s) = \frac{n!}{s!(n-s)!} p^s (1-p)^{n-s}$$

Функція розподілу ймовірності має такий вигляд:

$$F(k) = \sum_{s=0}^k \frac{n!}{s!(n-s)!} p^s (1-p)^{n-s}, \quad k=0. 1. \dots n$$

## Розподіл Пуассона.

Випадкову величину з розподілом Пуассона можна отримати, якщо допустити, що кількість незалежних випробувань  $n$  в біноміальному розподілі прямує до нескінченності, а ймовірність успішного випробування  $p$  - до нуля, причому добуток  $np$  постійний і дорівнює  $\lambda$ .

Функція густини розподілу Пуассона задається виразом

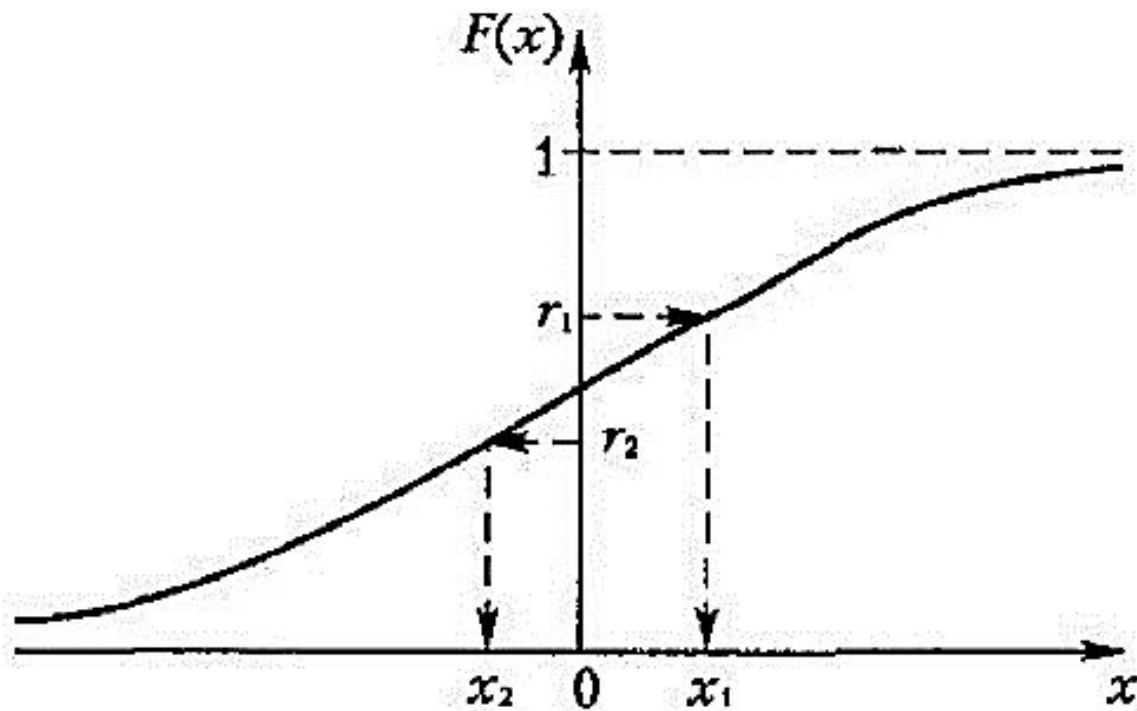
$$f(s) = \frac{\lambda^s}{s!} e^{-\lambda}$$

Таким чином, пуассонівський розподіл є граничним випадком біноміального і описує випадкові події, які мають місце дуже рідко.

На практиці так розподілені кількість дефектів в готовому виробі і кількість аварій на транспорті за деякий тривалий проміжок часу, кількість дзвінків в телефонній мережі за одиницю часу.

# Моделювання неперервних випадкових величин.

В даному випадку використовується метод зворотної функції. Випадкова величина має функцію густини ймовірності  $f(x)$  і монотонно зростаючу функцію розподілу  $F(x)$  (рис.).



**Суть метода.** За допомогою генератора випадкових чисел генеруємо значення випадкової величини  $r_1$ , якій відповідає точка на осі ординат.

Значення випадкової величини  $x_i$  з функцією розподілу  $F(x)$  можемо отримати з рівняння  $F(x_i) = r_i$ .

Дійсно, якщо на осі ординат відкласти значення  $r_i$  випадкової величини, розподіленої рівномірно в інтервалі  $[0,1]$  і на осі абсцис знайти значення  $x_i$  випадкової величини, при якому  $F(x_i) = r_i$ ,

то випадкова величина  $X = F^{-1}(r)$  матиме функцію розподілу  $F(x)$ .

За визначенням функція розподілу  $F(x)$  випадкової величини  $X$  дорівнює ймовірності  $P(X < x)$ :

$$P(X < x) = P(r < F(x)) = \int_0^{F(x)} f(r) dr = F(x)$$

Таким чином, послідовність  $r_1, r_2, r_3 \dots$ , що належить  $R(0, 1)$ , перетворюється в послідовність  $x_1, x_2, x_3 \dots$ , яка має задану функцію густини розподілу  $f(x)$ .

## Рівномірний розподіл.

В загальному випадку випадкова величина  $X$  рівномірно розподілена на відрізку  $[a, b]$ , якщо густина розподілу ймовірності має вигляд:

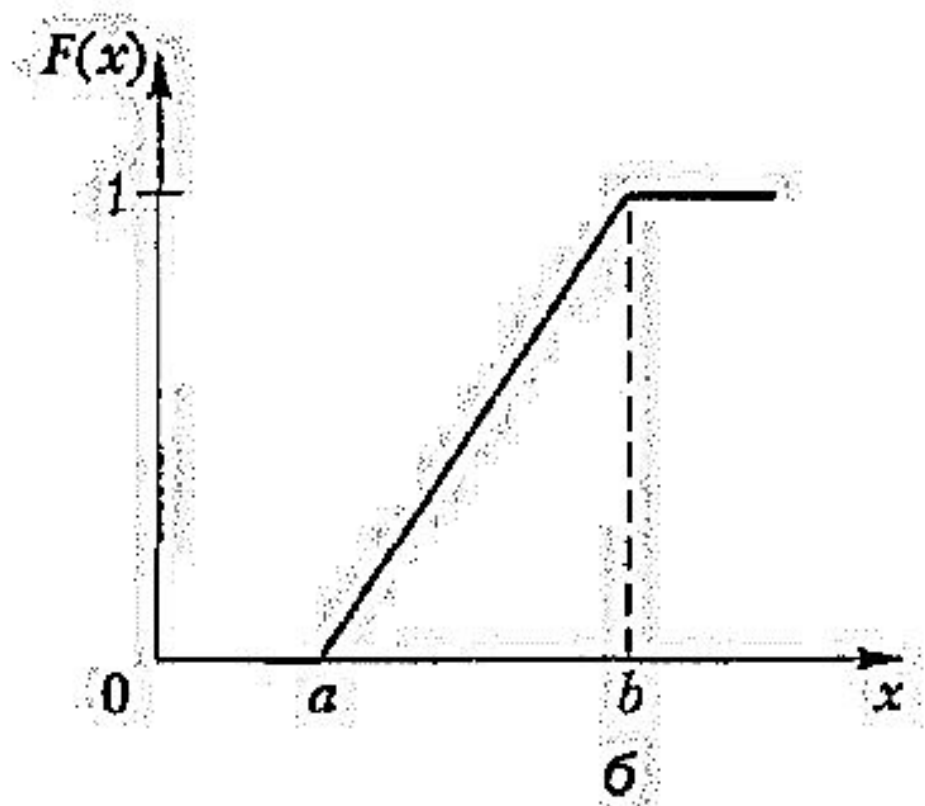
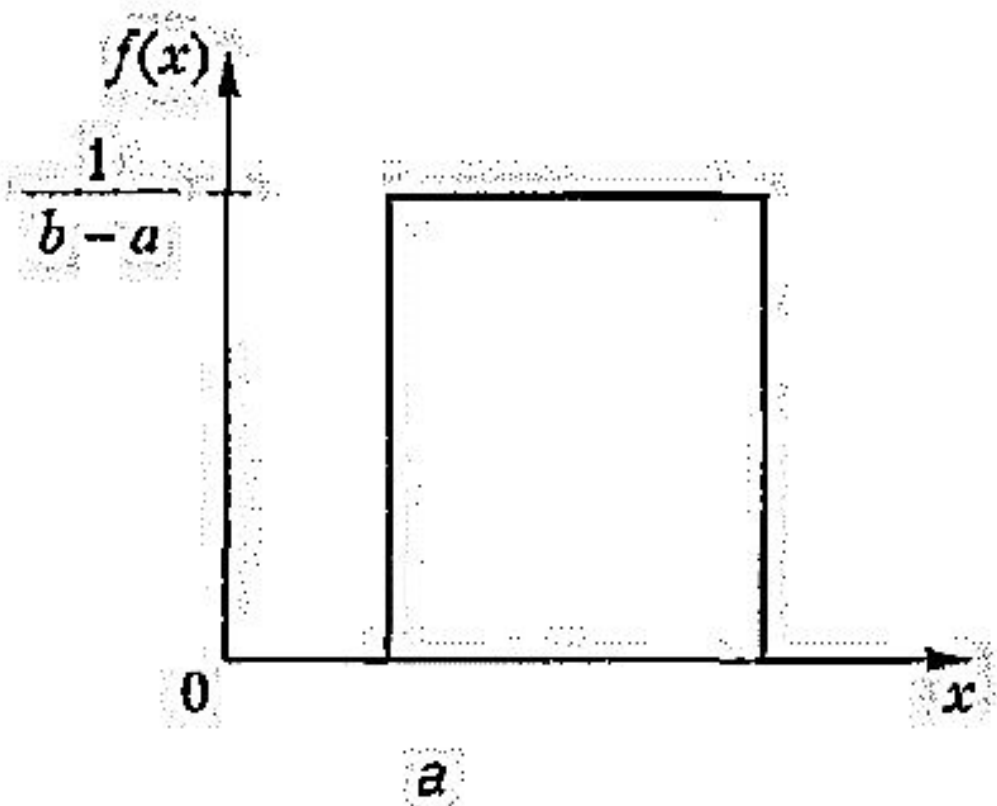
$$f(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x > b. \end{cases}$$

Функцію розподілу ймовірностей можна знайти як

$$F(x) = \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a}$$

тобто

$$F(x) = \begin{cases} 0, & x < a \\ (x-a) / (b-a), & a \leq x \leq b \\ 1, & x > b \end{cases}$$



Математичне сподівання і дисперсія випадкової величини  $X$  визначається як:

$$M(x) = \frac{a+b}{2}, \quad D(x) = \frac{b-a}{12}$$

Для моделювання випадкової рівномірно розподіленої на відрізку  $[a, b]$  величини можна скористатися методом зворотної функції.

Обчислимо функцію розподілу випадкової величини і прирівняємо її значенням  $r_i$ :

$$r_i = \int_a^{x_i} \frac{dx}{b-a} = \frac{x_i - a}{b-a}$$

Звідси знаходимо значення випадкової величини з функцією розподілу  $f(x)$ :

$$x_i = (b - a) r_i + a$$

У мові GPSS такий розподіл часто використовується в блоках Advance для моделювання затримки проходження інформації або під час генерування потоків транзактов в блоках GENERATE.

Наприклад, щоб згенерувати потік транзактов, які надходять в модель кожні  $5 \pm 2$  хвилини, використовують блок GENERATE 5.2.

Приклади реальних задач, в яких виникає необхідність моделювання рівномірно розподілених випадкових величин, - аналіз помилок округлення під час проведення числових розрахунків (точність задається числом десяткових знаків), час переміщення головок в магнітних накопичувачах (мінімальне і максимальне час), відхилення від графіка руху транспортних засобів (наприклад, метро).



## Експоненціальний розподіл.

Експонентний закон набув широкого поширення в теорії надійності складних систем. Функція густини експоненціального розподілу

випадкової величини має вигляд:

$$f(x) = \lambda e^{-\lambda x}$$

Для її моделювання скористуємось методом зворотної функції. Маємо

$$r_i = \int_0^{x_i} f(x) dx = \int_0^{x_i} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x_i}$$

Знаходимо  $x_i$ :

$$x_i = -\frac{1}{\lambda} \ln(1 - r_i)$$

Можна показати, що випадкові величини  $(1 - r_i)$  мають такий самий розподіл, як і  $r_i$ .

Тоді, замінюючи  $(1 - r_i)$  на  $r_i$ , одержимо

$$x_i = -\frac{1}{\lambda} \ln r_i$$

Випадкові величини з експоненціальним розподілом широко використовуються в задачах моделювання і аналізу СМО, наприклад, під час моделювання виходу з ладу і ремонту устаткування, які виникають в складних системах.

## Пуассонівський потік.

Розглянемо моделювання пуассонівського потоку з інтенсивністю  $\lambda$ , основна властивість якого полягає в тому, що ймовірність надходження  $k$  вимог протягом інтервалу тривалістю  $t$  становить:

$$p_k(t) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad k = 1, 2, \dots$$

Для пуассонівського потоку інтервали часу між надходженнями двох сусідніх вимог мають експонентний закон розподілу.

Тому для його моделювання досить отримати ряд чисел з таким розподілом.

Це можна реалізувати за допомогою методу зворотного функції, якщо ряд випадкових чисел  $r_i$  рівномірно розподілених в інтервалі  $[0,1]$  перетворити відповідно до функції, зворотної експоненційної функції розподілу

$$t_j = F^{-1}(x) = -\bar{T} \ln(r_j),$$

де  $t_j$  –  $j$ -й проміжок часу між надходженнями двох сусідніх вимог,

$\bar{T} = 1/\lambda$  – середнє значення проміжку часу між надходженнями двох сусідніх вимог,  $r_j$  –  $j$  – число в послідовності випадкових чисел з рівномірним розподілом в інтервалі  $[0,1]$ .

В мові GPSS для моделювання пуассонівського потоку вимог з  $\bar{T} = 1/\lambda = 2$  години використовується блок GENERATE 120 (одиниця часу в моделі 1 хвилина).

## Нормальний розподіл.

Випадкова величина  $X$  має нормальний розподіл (розподіл Гауса), якщо її густина розподілу ймовірностей описується законом:

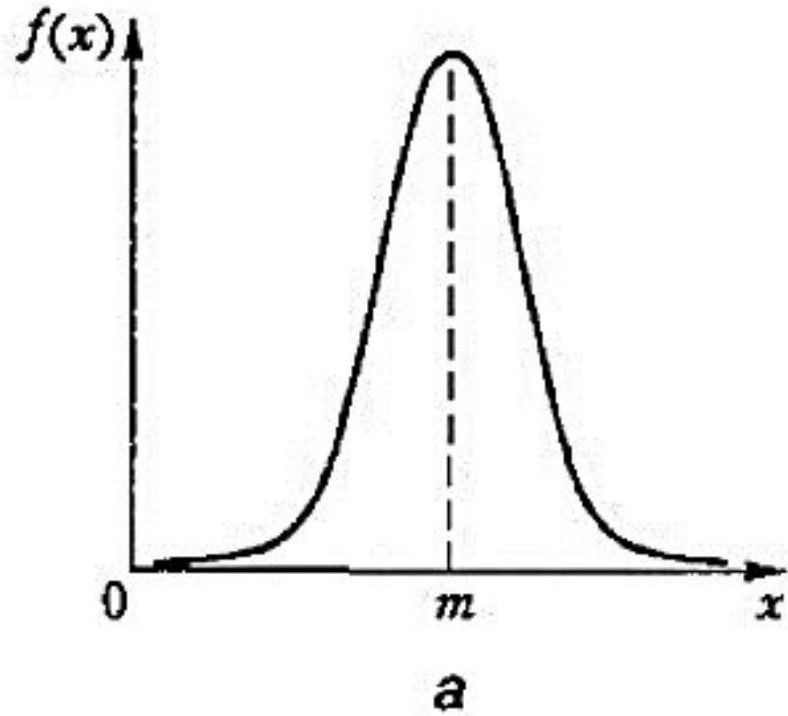
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

де  $m$  – математичне сподівання,  $\sigma$  – середньоквадратичне відхилення.

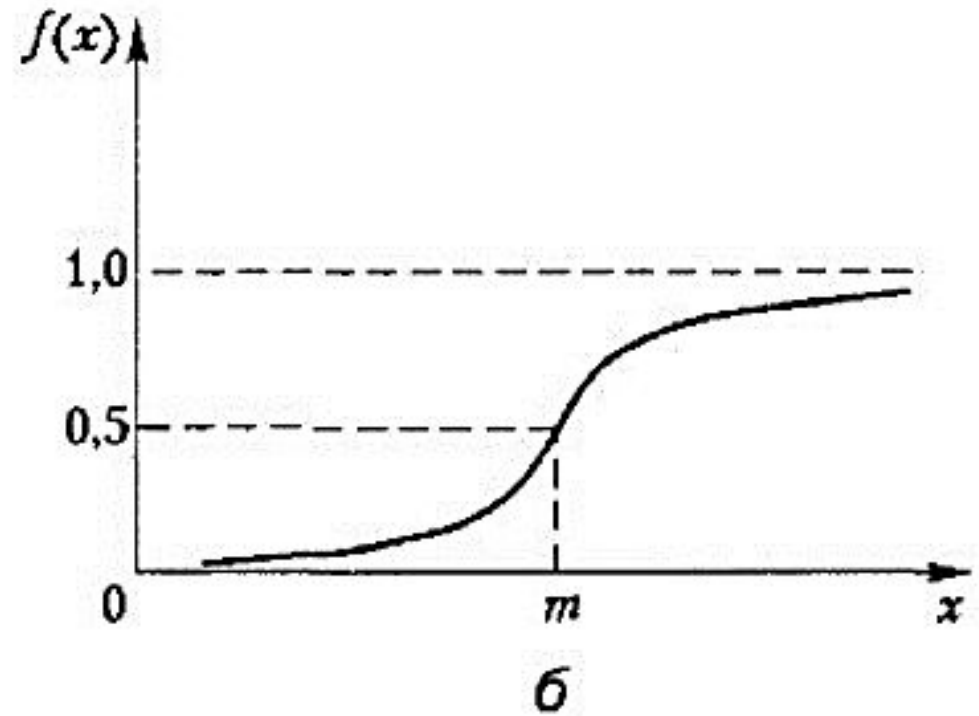
Функція розподілу нормально розподіленої величини  $X$  має вигляд:

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

Графіки функцій густини ймовірностей  $f(x)$  і розподілу  $F(x)$  зображені на рис.



$F(x)$



# **Збір статистичних даних для отримання оцінок характеристик випадкових величин.**

Основними елементами, з сукупності яких складається імовірнісна модель методу статистичних випробувань, є випадкові реалізації.

Очевидно, що при вирішенні деякої задачі визначення характеристик або параметрів вихідного випадкового процесу повинен бути визначений цей випадковий процес.

Шуканими величинами при використанні методу статистичних випробувань є оцінки:



- ймовірності настання деякої події;
- математичного очікування випадкової величини;
- дисперсії випадкової величини;
- коефіцієнтів коваріації або кореляції випадкової величини.

Для оцінки ймовірності  $p$  настання деякої події  $A$  використовується частота настання цієї події:

$$\hat{p} = \frac{m}{N}$$

де  $m$  – частота настання події, а  $N$  – число дослідів.

Для оцінки математичного очікування випадкової величини використовується середнє значення

$$\hat{x} = \hat{E}[x] = \frac{1}{N} \sum_{i=1}^N x_i$$

де  $x_i$  –  $i$ -а реалізація випадкової величини.

Для оцінки дисперсії випадкової величини  $\xi$  використовують формулу

$$S^2 = \frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{x})^2$$

де  $S^2$  – оцінка дисперсії випадкової величини  $\xi$ .

Безпосередньо використовувати ці формули для обчислення дисперсії складно, оскільки середнє значення змінюється в міру накопичення  $x_i$ , тобто потрібно запам'ятовувати всі  $N$  значень  $x_i$ .

Тому для обчислення використовують формулу:

$$S^2 = \hat{\sigma}^2 = \frac{1}{N-1} \left[ \sum_{i=1}^N x_i^2 - \frac{1}{N} \left( \sum_{i=1}^N x_i \right)^2 \right]$$

В цьому випадку достатньо накопичити дві суми значень –  $x_i$  и  $x_i^2$ .

Для випадкових величин  $\xi$  та  $\eta$  з можливими значеннями  $x_k, y_k$  оцінка кореляційного моменту визначається так:

$$\hat{K}_{\xi\eta} = \frac{1}{N-1} \left[ \sum_{k=1}^N (x_k - \bar{x})(y_k - \bar{y}) \right]$$

або в зручній для обчислень формі:

$$\hat{K}_{\xi\eta} = \frac{1}{N-1} \left[ \sum_{k=1}^N x_k y_k - \frac{1}{N} \sum_{k=1}^N x_k \sum_{k=1}^N y_k \right]$$

# Визначення кількості реалізацій

Точність оцінювання параметрів системи, які отримують під час обробки результатів моделювання, залежить від кількості випробувань  $N$ .

Обсяг вибірки  $N$  завжди обмежений, тому раніше розглянуті оцінки матимуть різні похибки і дисперсії.

Якщо потрібно оцінити значення деякого параметра  $a$  за результатами моделювання  $x_i$ , то в якості його оцінки потрібно брати величину  $\bar{x}$ , яка є функцією всіх значень  $x_i$ .

Статистична оцінка  $\bar{x}$  також є випадковою величиною, тому вона буде відрізнятися від  $\mathbf{a}$ , тобто

$$|\mathbf{a} - \bar{x}| < \varepsilon, \quad \varepsilon \text{ — точність або помилка оцінки.}$$

Ймовірність того, що ця нерівність виконується, позначимо  $\alpha$ :

$$P(|\mathbf{a} - \bar{x}| < \varepsilon) \geq \alpha \quad (1), \text{ де}$$

$\varepsilon$  — довірчий інтервал для  $\alpha$ , довжина якого дорівнює  $2\varepsilon$ , а  $\alpha$  - довірчий рівень або надійність оцінки.

Вираз (1) використовують для визначення точності результатів статистичних випробувань.

## Оцінка ймовірності.

Припустимо, що метою моделювання є оцінка ймовірності настання деякої події  $A$ , яка визначає стан системи.

У кожній з  $N$  реалізацій процесу настання події  $A$  є випадковою величиною  $\xi$ , яка приймає значення  $x_1 = 1$  з ймовірністю  $p$  і  $x_2 = 0$  з ймовірністю  $1 - p$ .

Тоді можна визначити математичне очікування і дисперсію відповідно до формулами:

$$M[\xi] = x_1 p + x_2 (1 - p) = p,$$

$$D[\xi] = (x_1 - M[\xi])^2 p + (x_2 - M[\xi])^2 (1 - p) = p(1 - p).$$

Як оцінку  $p$  використовують частоту настання події  $A$ .

За умови, що  $N$  задано, для отримання цієї оцінки досить накопичувати  $m$ :

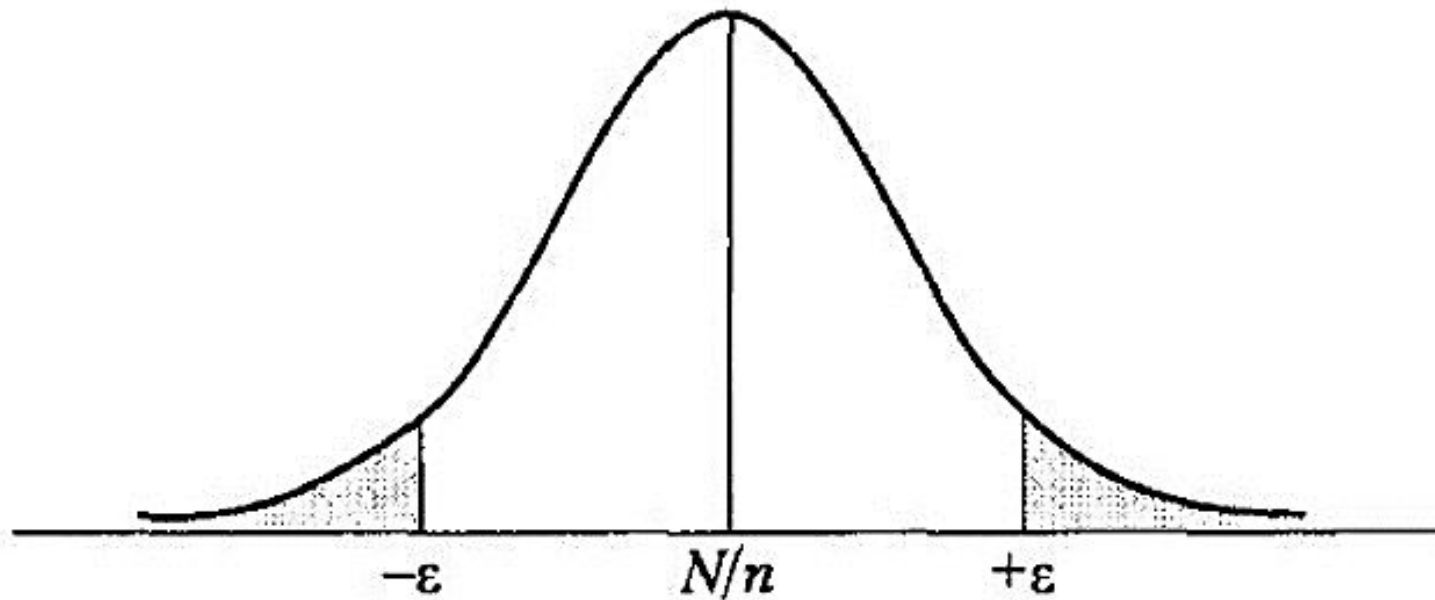
$$\frac{m}{N} = \frac{1}{N} \sum_{i=1}^N x_i,$$

$X_i$  – настання події  $A$  в реалізації  $i$ .

Визначимо вибіркове математичне очікування  $M[m/N] = p$  і дисперсію  $D[m/N] = p(1-p)/(N-1)$ .

Випадкова величина  $m/N$  буде мати розподіл, близький до нормального.

Тому для кожного рівня достовірності  $\alpha$  по таблицям нормального



розподілу можна знайти таку величину  $t_\alpha$ , при якій точність обчислюється за формулою

$$\varepsilon = t_\alpha \sqrt{D [m/N]}.$$



Якщо  $\alpha = 0.95$ , то  $t_{\alpha} = 1.96$ , а якщо  $\alpha = 0.003$ , то  $t_{\alpha} = 3$ .

Підставимо в останню формулу вираз дисперсії:

$$\varepsilon = t_{\alpha} \sqrt{\frac{p(1-p)}{N-1}}.$$

звідки

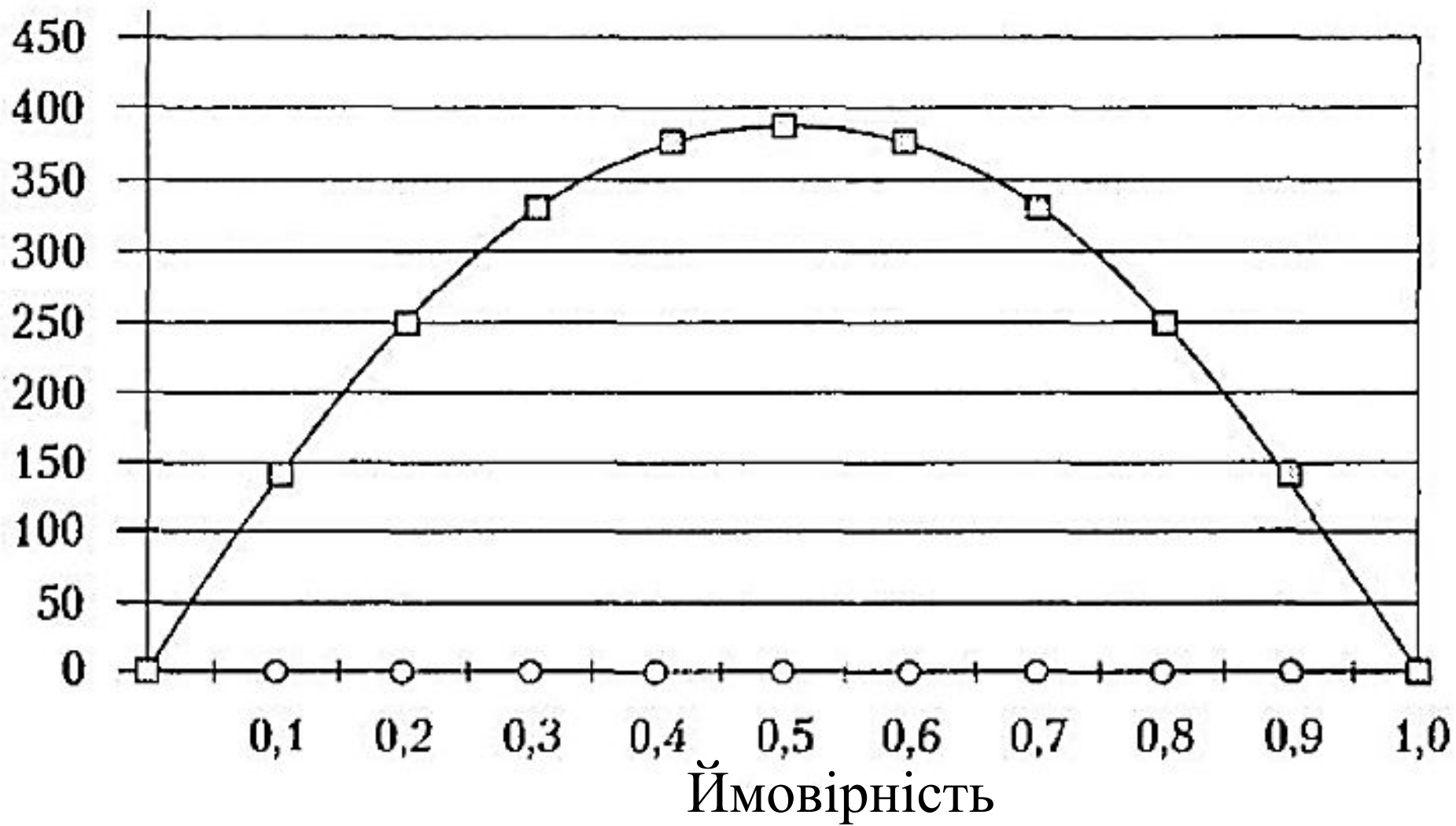
$$N = t_{\alpha}^2 \frac{p(1-p)}{\varepsilon^2} + 1.$$

З останньої формули видно, що при  $p = 1$  або  $p = 0$ , кількість реалізацій, які необхідно зробити для підтвердження того, що подія  $A$  настає (чи ні), дорівнює 1.

Але оскільки ймовірність  $p$  заздалегідь невідома, проводять випробування ( $N = 50 \dots 100$ ), оцінюють частоту  $m / N$  і підставляють її значення в останній вираз замість  $p$ , після чого визначають остаточну кількість реалізацій.

Графік числа реалізацій для  $\alpha = 0.05$  та різних значень  $p$ , якщо  $\varepsilon = 0.05$ , наведено на рис.

Кількість  
реалізацій



## Оцінка середнього значення.

Нехай випадкова величина має математичне сподівання  $a$  і дисперсію  $\sigma^2$ . В  $i$  – й реалізації вона приймає значення  $X_i$ .

Як оцінку математичного очікування  $a$  використовуємо середнє арифметичне:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i.$$

Для кількості реалізації  $N$  одержимо  $N = t_{\alpha}^2 \cdot \sigma^2 / \varepsilon^2 + 1$