

Міністерство освіти і науки України
Запорізький національний університет

Н.В.Матвіїшина, Н.О.Кондрат'єва

ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ

Методичні рекомендації до лабораторних занять для здобувачів ступеня вищої освіти бакалавра спеціальності «Комп'ютерні науки» освітньо-професійної програми «Комп'ютерні науки»

Затверджено
вченою радою ЗНУ
Протокол № від

Запоріжжя
2020

УДК: 311.21:004(076.5)
М338

Матвіїшина Н.В., Кондрат'єва Н.О. Інформаційне забезпечення статистичних досліджень : Методичні рекомендації до лабораторних занять для здобувачів ступеня вищої освіти бакалавра спеціальності «Комп'ютерні науки» освітньо-професійної програми «Комп'ютерні науки». Запоріжжя : ЗНУ, 2020. 50 с.

Методичні рекомендації спрямовані на закріплення студентами теоретичних знань з дисципліни «Інформаційне забезпечення статистичних досліджень» за такими темами: основні інструменти та засоби роботи з математичними пакетами Mathcad та Maple, теорія оцінювання та статистичні гіпотези в Mathcad та Maple, основи кореляційного та регресійного аналізу в Mathcad та Maple, елементи дисперсійного аналізу та реалізація методів Монте-Карло засобами математичного пакету Mathcad.

Відповідно до навчальної програми дисципліни, в методичних рекомендаціях надано необхідну теоретичну інформацію, а також розглянуто питання щодо практичного використання набутих знань і навичок. До кожної теми, що розглядається в методичних рекомендаціях, надано контрольні запитання.

Методичні рекомендації з дисципліни «Інформаційне забезпечення статистичних досліджень» призначені для студентів спеціальності «Комп'ютерні науки», але представлене видання може бути корисним студентам різних спеціальностей.

Рецензент:

С. П. Швидка, кандидат фізико-математичних наук, доцент кафедри прикладної математики і механіки

Відповідальний за випуск:

С.Ю. Борю, кандидат технічних наук, доцент, завідувач кафедри комп'ютерних наук

ВСТУП

Основними завданнями вивчення дисципліни «Інформаційне забезпечення статистичних досліджень» є оволодіння основними методами та засобами технологій, що використовуються для статистичної обробки електронної інформації.

Метою вивчення навчальної дисципліни є засвоєння теоретичних основ і принципів роботи із сучасними програмними математичними додатками та засобами обробки статистичної інформації.

У результаті вивчення дисципліни студенти повинні:

- знати принципи використання операторів, команд та функцій математичних пакетів Mathcad та Maple для статистичних досліджень: обчислення відповідних характеристик, перевірки гіпотез;
- знати принципи проведення кореляційного, регресійного, дисперсійного аналізу;
- знати основні правила роботи з формулами, числами, графіками, текстами в математичних пакетах Mathcad та Maple;
- використовувати відповідні команди та функції математичних пакетів Mathcad та Maple для вирішення задач статистики та побудови різноманітних статистичних графіків.

Вивчення дисципліни дає підґрунтя для подальшого засвоєння можливостей сучасних математичних пакетів для статистичної обробки інформації.

Структура методичних рекомендацій

Методичні рекомендації складаються зі вступу, 4 теоретичних тем, що вивчаються студентами спеціальності «Комп'ютерні науки» в рамках дисципліни «Інформаційне забезпечення статистичних досліджень» та завдань лабораторних робіт. Теоретичний матеріал структурований та проілюстрований скріншотами, та рисунками, що надає наочності представленому матеріалу. Наприкінці кожного теоретичного розділу є контрольні запитання.

Друга частина методичних рекомендацій містить завдання до 8 лабораторних робіт за такими темами:

- 1.Числові характеристики вибірки. Побудова гістограм і полігонів в MathCAD.
- 2.Побудова графіків щільності вірогідності і функції розподілу. Обчислення довірчого інтервалу.
- 3.Числові характеристики вибірки. Побудова графіків в Maple.
- 4.Перевірка статистичних гіпотез в MathCAD.
- 5.Кореляційний та регресійний аналіз в MathCAD.
- 6.Кореляційний та регресійний аналіз в Maple.
- 7.Однофакторний дисперсійний аналіз в MathCAD.
- 8.Метод Монте-Карло в MathCAD.

Після виконання лабораторної роботи студентам рекомендується відповісти на запитання для самоперевірки.

1 ОСНОВНІ ІНСТРУМЕНТИ ТА ЗАСОБИ РОБОТИ З МАТЕМАТИЧНИМ ПАКЕТОМ MATHCAD

1.1 Інтерфейс користувача математичного пакету MathCAD

Додаток MathCAD [5, 8] – один з найбільш популярних комп'ютерних математичних пакетів. З його допомогою можна вирішувати найрізноманітніші математичні задачі і оформляти результати розрахунків на високому професійному рівні. За допомогою цього пакета здійснюються не тільки прості і допоміжні обчислення але і доволі складні розрахунки і наукові дослідження, що використовують комбінації різних чисельних алгоритмів і аналітичних перетворень.

Основні елементи вікна MathCAD представлено на Рис. 1.

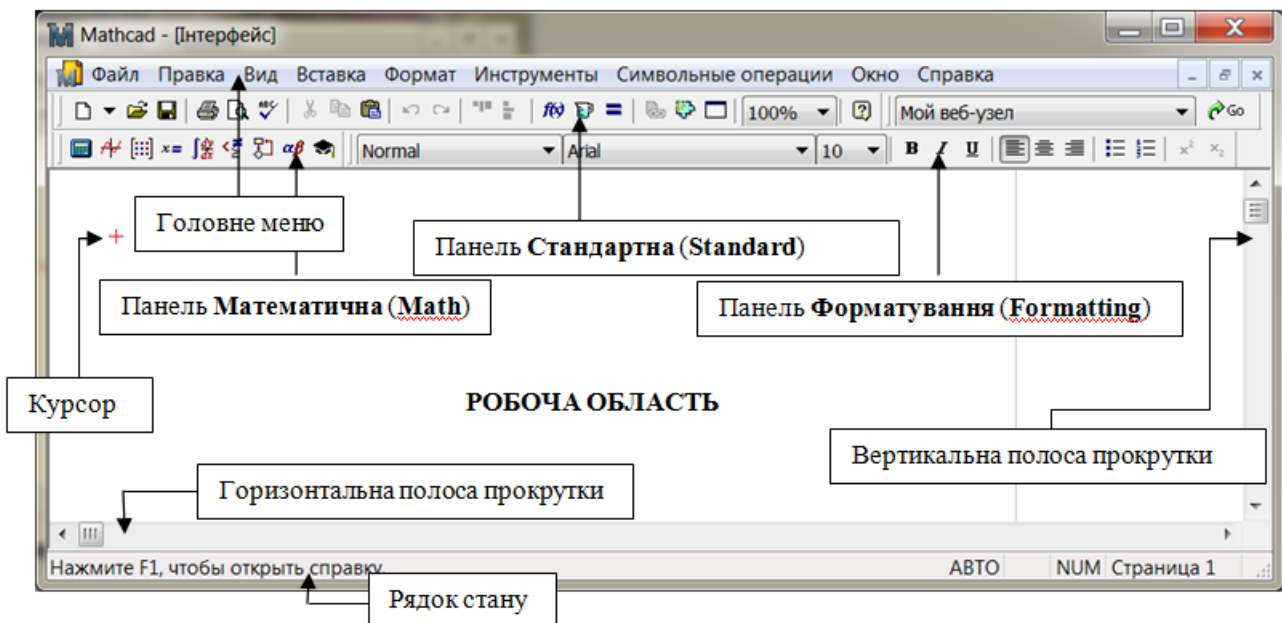


Рис. 1 Елементи вікна MathCAD

Головне меню MathCAD містить команди, які об'єднані у відповідності з функціями, які виконуються:

Файл (File) – загальні команди роботи з документом (створення, відкриття, зберігання, друк і т.д.);

Правка (Edit) – команди редагування документу (копіювання, видалення, вставки і т.д.);

Вид (View) – команди управління виглядом інтерфейсу (показ, приховування панелей інструментів, зміна масштабу і т.д.)

Вставка (Insert) – команди вставки в документ різних об'єктів (графіки, таблиці, функції і т.д.);

Формат (Format) – команди управління форматом відображення всіх об'єктів MathCAD;

Инструменты (Tools) – допоміжні команди (перевірка орфографії, зміна налаштувань документу і т.д.);

Символьные операции (Symbolics) – команди, які управляють символьними обчисленнями;

Окно (Window) – стандартні команди управління вікнами відкритих документів;

Помощь (Help) – довідкова система MathCAD.

Робоча область і курсор документу MathCAD. Робоча область займає більшу частину вікна MathCAD і саме тут можна вводити формули, графіки та інші об'єкти MathCAD. Курсор показує те місце на сторінці документу, де буде починатися наступне введення інформації. В процесі введення формули курсор MathCAD набуває вигляд двох ліній синього кольору (Рис. 2).

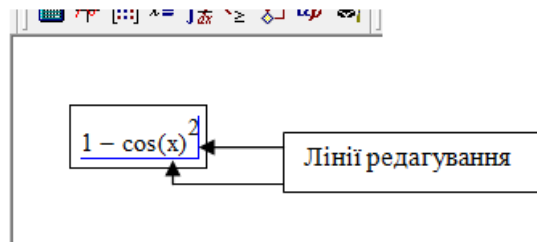


Рис. 2 Лінії редагування формули MathCAD

Панелі інструментів MathCAD. Панелі інструментів призначені для швидкого виконання різних команд. Показати або приховати панелі можна за допомогою команди «Вид»/«Панели инструментов» («View» / «Toolbars») (Рис. 3).

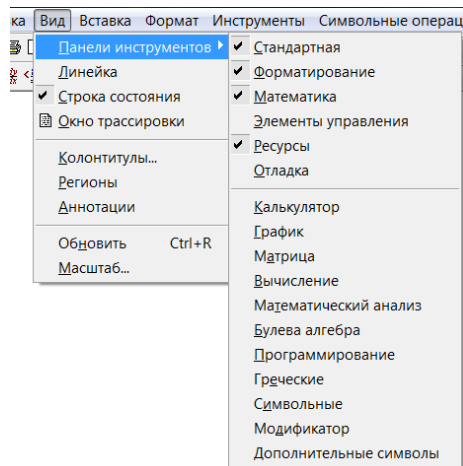


Рис. 3 Меню панелей інструментів

За замовчуванням в MathCAD показано декілька панелей інструментів: «Форматирование» («*Formatting*»), «Стандартная» («*Standard*»), «Ресурсы» («*Resources*»), а також панель «Математика» («*Math*»), яка містить посилання на панелі зі специфічними елементами.

Панель інструментів «Математика» («Math»). Панель Математика (Math) призначена для відтворення панелей інструментів, які необхідні для роботи з формулами, графіками і т.д. (Рис. 4, Рис. 5, Рис. 6).

Панель «*Калькулятор*» («*Calculator*») – введення в документ основних математичних символів та операцій.

Панель «*График*» («*Graph*») – вставка в документ шаблонів для графіків різного вигляду.

Панель «*Матрица*» («*Matrix*») – операції з векторами і матрицями.

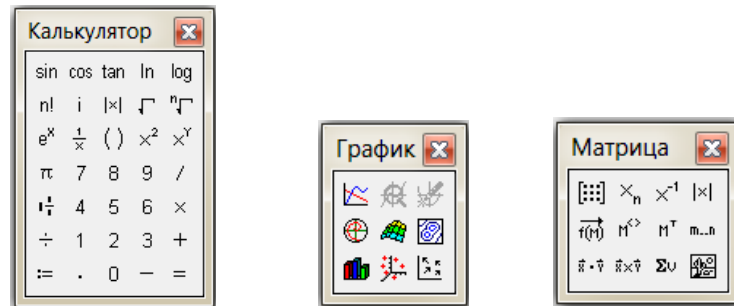


Рис. 4 Вигляд панелей інструментів «*Калькулятор*», «*График*», «*Матрица*»

Панель «*Вычисление*» («*Evaluation*») – доступ до команд, які управляють процесом обчислення (присвоювання значень, виведення результатів розрахунків).

Панель «*Математический анализ*» («*Calculus*») – вставка в документ шаблонів диференціювання, інтегрування і т.д.

Панель «*Булева алгебра*» («*Boolean*») – доступ до команд, які задають рівняння, нерівності, а також логічні операції.

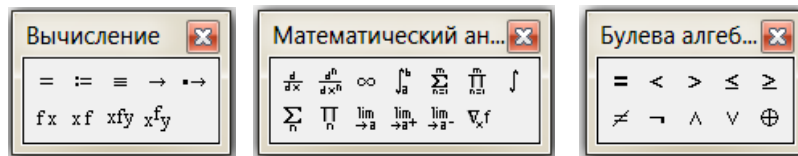


Рис. 5 Вигляд панелей інструментів «*Вычисление*», «*Математический анализ*», «*Булева алгебра*»

Панель «*Программирование*» («*Programming*») – доступ до операторів, необхідних для створення програмних модулів.

Панель «*Греческий*» («*Greek*») – вставка в документ грецьких символів.

Панель «*Символьные*» («*Symbolic*») – вставка в документ ключових слів для управління символічними операціями.

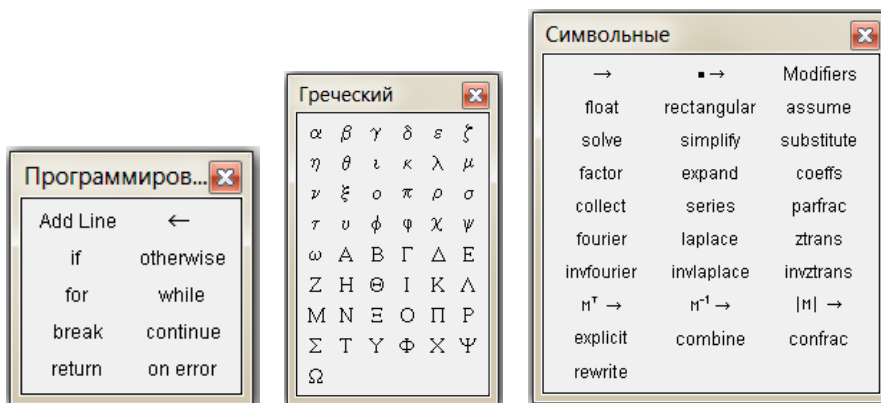


Рис. 6 Вигляд панелей інструментів «*Программирование*», «*Греческий*», «*Символьные*»

1.2 Введення та виведення інформації

1.2.1 Створення формул

Виконання обчислень – це основне призначення пакету MathCAD. Кожному математичному оператору відповідає та або інша кнопка на панелях математичних інструментів або клавіатурне поєднання для швидкого вводу з клавіатури.

Прості арифметичні вирази можуть бути введені безпосередньо з клавіатури. Наприклад, необхідно ввести вираз $\frac{1}{3} + 5 \cdot \left(\frac{2}{5} - 7\right)$ та отримати результат обчислень.

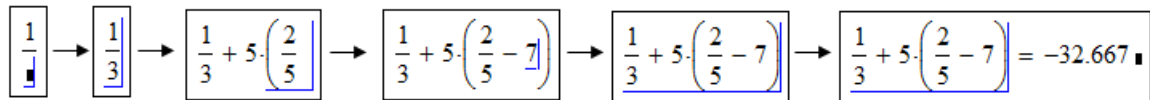


Рис. 7 Введення простих арифметичних виразів

Для створення складних математичних конструкцій використовуються шаблони відповідного оператора, які обираються на панелі інструментів «*Математический анализ*» («*Calculus*») (Рис. 8).

Рис. 8 Шаблони математичних операторів

Після вставки шаблону в документ необхідно заповнити місцезаповнювачі відповідними значеннями (Рис. 9).

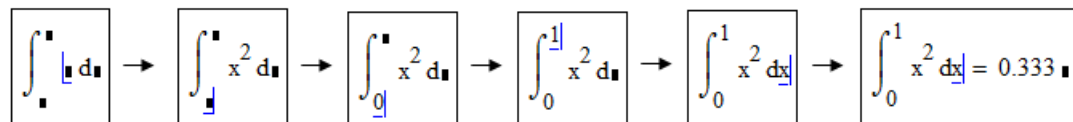


Рис. 9 Приклад заповнення шаблону

1.2.2 Вставка функцій

Крім операторів, вирази, що обчислюються можуть містити функції. В MathCAD є дві групи функцій:

- 1) функції користувача;
- 2) вбудовані функції.

Функція характеризується своїм іменем та аргументами (змінні, що вказуються в дужках):

ім'я_функції (змінна_1, змінна_2, ..., змінна_n)

Загальний вигляд функції користувача:

ім'я_функції (змінна_1, змінна_2, ..., змінна_n) := вираз

Функцію користувача можна, наприклад, задати таким чином (Рис. 10):

$$a := 1 \quad b := 3$$

$$f(x, y) := a \cdot \sin(x) + b \cdot \cos(y)$$

$$f(0, 0) = 3 \quad \text{значення функції при } x=0 \text{ та } y=0$$

Рис. 10 Приклад використання функції користувача

Вбудовані функції – це функції, що задані в MathCAD за замовчуванням. Для того, щоб їх використати, достатньо коректно ввести ім'я функції з клавіатури, скористатися панеллю «*Калькулятор*» («*Calculator*») для вставки найбільш поширених функцій, або скористатися бібліотекою функцій MathCAD (команда «Вставка» / «Функція» («*Insert*» / «*Function*»)) (Рис. 11).

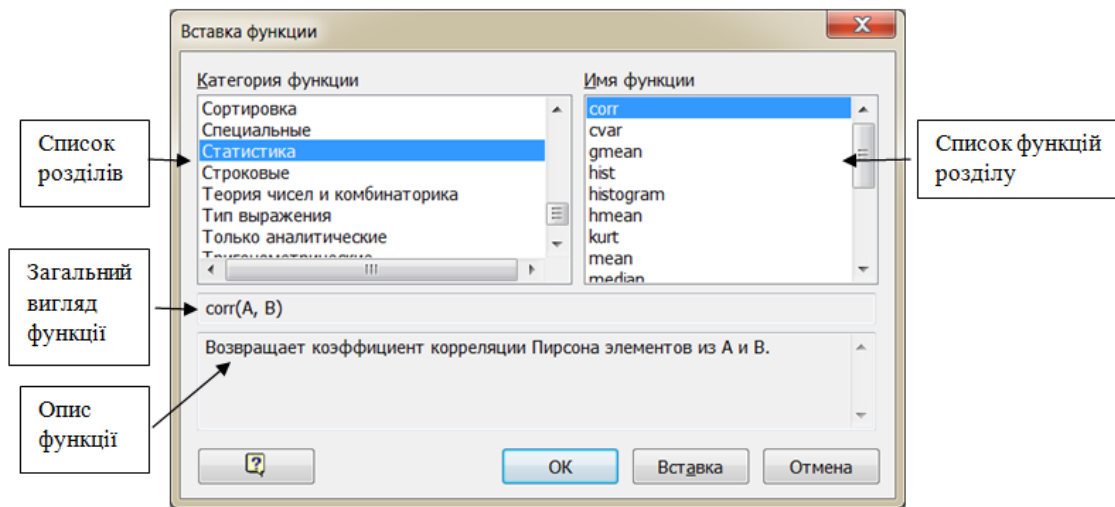


Рис. 11 Діалогове вікно для вставки вбудованої функції

Слід зазначити, що імена функцій користувача не повинні співпадати з іменами вбудованих функцій.

1.2.3 Створення тексту

Для того, щоб створювати будь-які коментарі, а також якісно оформлювати задачі, що розв'язуються, корисно використовувати текстовий режим MathCAD. Щоб ввести текстову область, можна використати комбінацію клавіш Shift+ «'», або команду «Вставка» / «Регион тексту» («*Insert*» / «*Text Region*»). Текст розміщується в області, що спеціально створено (Рис. 12).

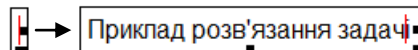


Рис. 12 Вставка тексту

1.2.4 Введення/виведення даних в текстові файли

Часто аналіз даних пов'язаний з імпортом даних із зовнішніх джерел та експортом на зовнішні носії. Для роботи із зовнішніми файлами даних в MathCAD передбачено декілька функцій (Рис. 13, Рис. 14):

READPRN(file) – читання даних в матрицю з текстового файлу;

WRITEPRN(file) – запис даних в текстовий файл;

APPENDPRN (file) – дозапис даних в текстовий файл, що вже існує;
 READFILE (file, type, [colwidths, rows, cols, emptyfill]) – читання даних з файлу (більш універсальна функція):
 file – назва файлу;
 type – тип файлу (“delimited”, “fixed”, “Excel”);
 colwidths – ширина стовпця даних, якщо зчитування з файлу проводиться з параметром “fixed” (з фіксованою шириною);
 rows – початковий рядок імпорту даних;
 cols – початковий стовпець імпорту даних;
 emptyfill – значення, яке використовується для заміни даних, що відсутні.

```
x := READPRN("d:\mnv\z.txt")
xT = (4.2 4.4 5.5 4.7 5.09 6.2)
p := READFILE("d:\mnv\z.txt", "delimited")
pT = (4.2 4.4 5.5 4.7 5.09 6.2)
```

```
k := (90) → значення повинно бути вектором або матрицею!!!
APPENDPRN("d:\mnv\z.txt") := k
y := READPRN("d:\mnv\z.txt")
yT = (4.2 4.4 5.5 4.7 5.09 6.2 90)
```

Рис. 13 Використання функцій READPRN, READFILE, APPENDPRN

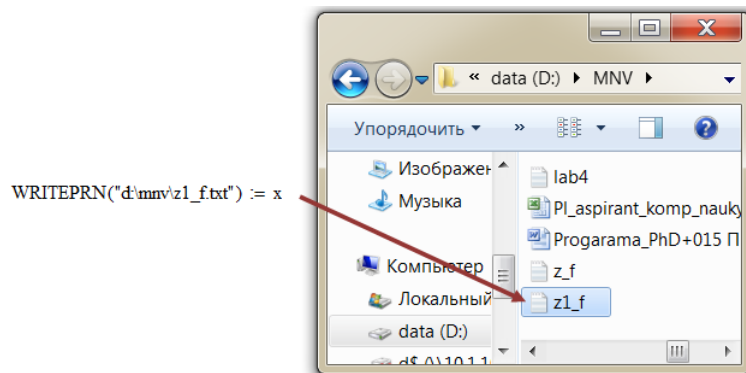


Рис. 14 Використання функції WRITEPRN

1.3 Типи даних. Обчислення в MathCAD

В MathCAD існує три типи даних:

- 1) числа (дійсні та комплексні);
- 2) рядки (будь-який текст в лапках);
- 3) масиви (матриці, вектори, таблиці, тобто будь-яка впорядкована послідовність елементів)

1.3.1 Змінні

Змінна – це іменованій об’єкт, має певне значення, яке може змінюватися в процесі виконання програми.

Для того, щоб визначити деяку змінну, необхідно використати символ \equiv , або символ \equiv з панелі інструментів «Вычисление» («Evaluation») (Рис. 15). Якщо

присвоювання здійснюється за допомогою $:=$, то змінна – **локальна**, якщо за допомогою \equiv , то змінна – **глобальна**.

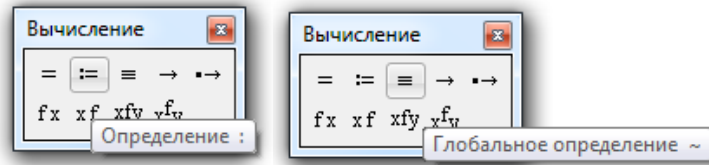


Рис. 15 Локальне та глобальне присвоювання

MathCAD прочитує документ **двічі**: зліва направо і зверху вниз. Спочатку виконуються всі дії, що приписані глобальним оператором присвоювання (в цьому випадку визначення змінної можна здійснювати в будь-якій точці документу), а потім дії, пов'язані з локальним присвоюванням, а також виводяться всі необхідні результати обчислень за допомогою оператора чисельного виводу ($=$) (Рис. 16).

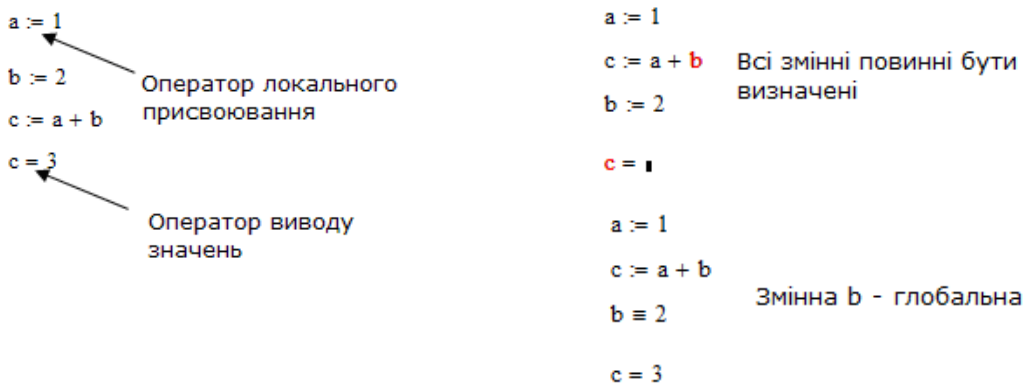


Рис. 16 Приклади локального, глобального присвоювання, виводу значень

Змінні можуть набувати значення із заданого інтервалу (Рис. 17). Локальне присвоювання для такої змінної виглядає так:

- 1) значення змінної змінюються з шагом 1
Name:=Nbegin .. Nend
- 2) значення змінної змінюються з шагом Step
Name:=Nbegin, Nbegin+Step.. Nend

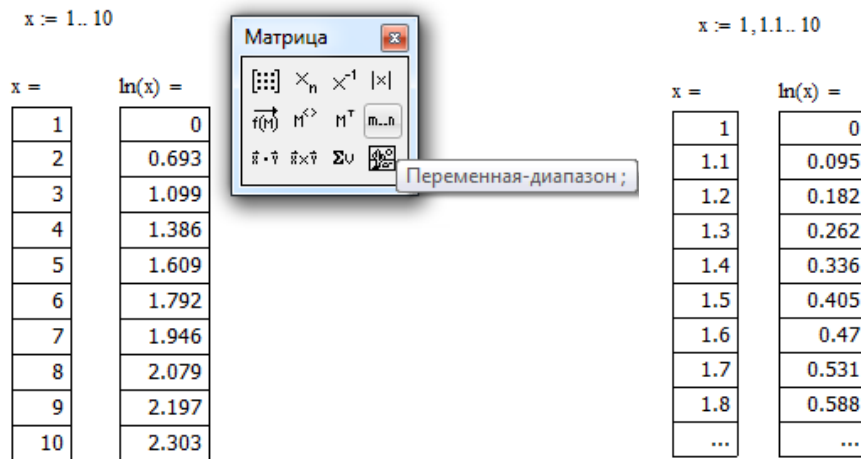


Рис. 17 Приклад набуття змінною x значень з інтервалу

1.3.2 Символьні обчислення

Значення багатьох функцій або виразів може бути обчислено як символічно такі чисельно. В системі MathCAD існує два оператори виводу: чисельний та символічний. В якості оператора чисельного виводу виступає \equiv (див. п. 1.3.1).

Символьне обчислення пов'язано з отриманням результату аналітичним методом.

Для того, щоб задіяти символічний процесор MathCAD, необхідно використати спеціальний оператор виводу \rightarrow з панелі інструментів «*Вычисление*» («*Evaluation*») (Рис. 18, Рис. 19).

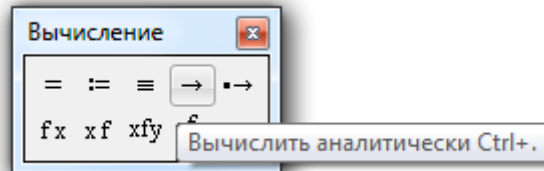


Рис. 18 Символьний оператор виводу

$$\frac{1}{3} + \frac{5}{9} \rightarrow \frac{8}{9} \quad \frac{1}{3} + \frac{5}{9} = 0.889$$








$$\cos\left(\frac{\pi}{6}\right) \rightarrow \frac{\sqrt{3}}{2} \quad \cos\left(\frac{\pi}{6}\right) = 0.866$$

$$\int x^2 dx \rightarrow \frac{x^3}{3} \quad \int_0^1 x^2 dx \rightarrow \frac{1}{3} \quad \int_0^1 x^2 dx = 0.333$$

Рис. 19 Приклад символічного та чисельного розрахунку

1.4 Побудова та форматування графіків

Графік – це зручний спосіб візуалізації інформації. MathCAD має значну кількість засобів для побудови графіків різних типів, і всі основні типи графіків та інструменти роботи з ними розташовані на панелі «*График*» («*Graph*»), що включено до панелі інструментів «*Математика*» («*Math*»):

- графік кривої у двовимірній системі координат  (X-Y Plot);
- графік кривої у полярній системі координат  (Polar Plot);
- поверхня  (Surface);
- контурний графік  (Contour Plot);
- стовпчаста тривимірна діаграма  (3D Bar Plot);
- точковий тривимірний графік  (3D Scatter Plot);
- векторне поле  (Vector Field).

Список всіх типів графіків MathCAD можна отримати за допомогою команди «*Вставка*» / «*График*» («*Insert*» / «*Graph*») (Рис. 20).

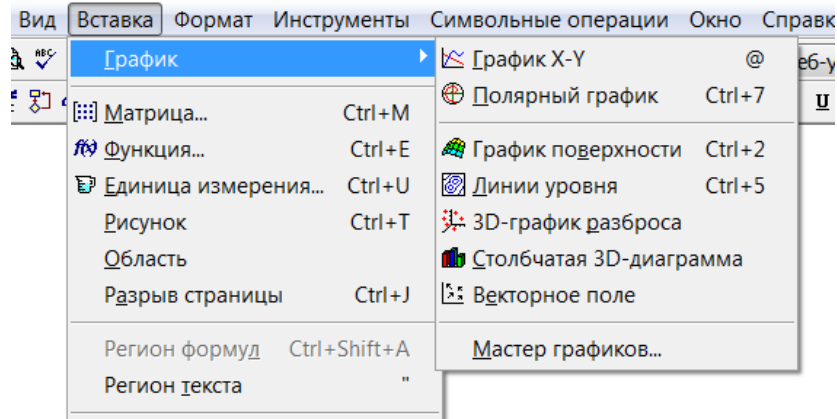


Рис. 20 Типи графіків MathCAD

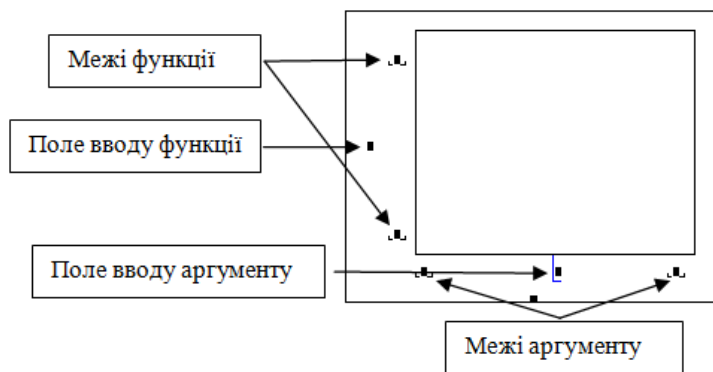
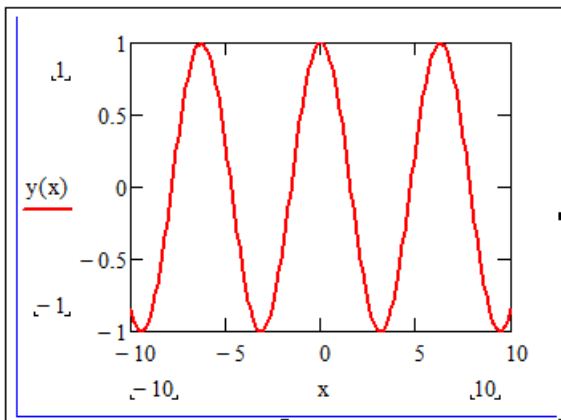


Рис. 21 Шаблон графіка MathCAD в декартовій системі координат

Приклади графіків функцій, що задано у декартових та полярних координатах наведено на Рис. 22.

$$x := -10, -9.9 .. 10$$

$$y(x) := \cos(x)$$



$$z(\varphi) := \cos(\varphi)^3 \cdot \sin(\varphi)^3$$

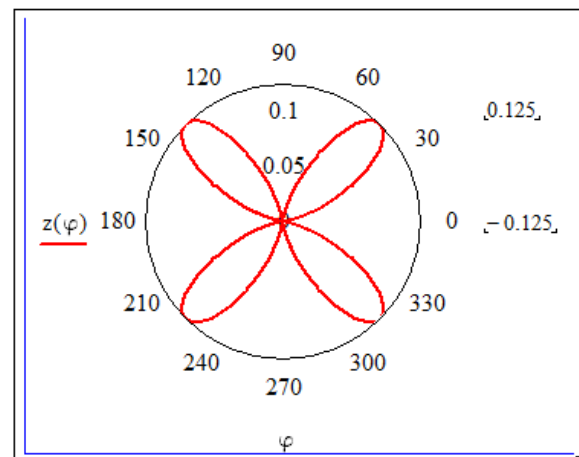


Рис. 22 Приклади побудови графіків в MathCAD

В MathCAD на одній площині можна зобразити декілька графіків функцій. Для цього необхідно ввести їх в поле вводу функції через **КОМУ**. Кожна функція може залежати від свого аргументу, тому, за необхідністю, можна ввести через **КОМУ** аргументи в поле вводу аргументів (Рис. 23).

```

x := -10, -9.9.. 10
y(x) := cos(x)
z(x) := sin(x)

```

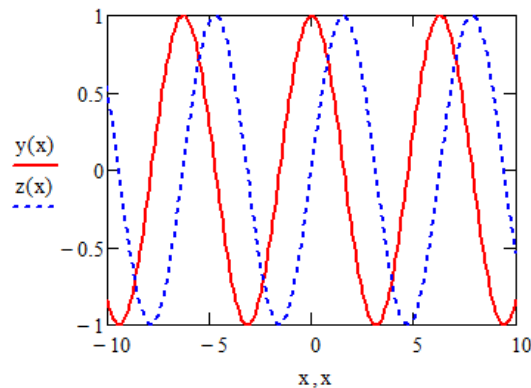


Рис. 23 Побудови декількох графіків в MathCAD

Після того, як графік побудовано, можна налаштувати його зовнішній вигляд. Для цього використовується команда «**Формат**» / «**Графік**» («**Format**» / «**Graph**»). Вікно форматування будь-якого графіку можна визвати з контекстного меню командою «**Формат...**» («**Format...**»).

Контрольні запитання

1. Які команди містить головне меню MathCAD?
2. Назвіть основні елементи вікна MathCAD.
3. Для чого призначена панель інструментів «Математика»?
4. Які функції використовуються для читання/запису текстових файлів?
5. Поняття локального та глобального присвоювання. Як визначити локальну та глобальну змінну в MathCAD?
6. Яким чином в MathCAD здійснюються символні обчислення?
7. За допомогою якого інструменту та які типи графіків можна побудувати в MathCAD?
8. Який загальний вигляд має функція користувача?
9. Яким чином побудувати в MathCAD декілька графіків функцій?

2 ВИКОРИСТАННЯ ІНСТРУМЕНТІВ МАТНСАД ДЛЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ

2.1 Вибіркові статистичні характеристики

У більшості статистичних розрахунків використовують вибіркові сукупності даних (вибірки): випадкові дані, які отримані в ході експерименту, або дані, які є результатом генерації випадкових чисел [2, 9]. Наближене значення шуканого статистичного параметра генеральної сукупності, встановлене за даними вибірки, називають **вибірковою оцінкою параметра**.

MathCAD має ряд вбудованих функцій для обчислення числових статистичних характеристик рядів випадкових даних (Рис. 24) [4, 5]:

- mean (x) – вибіркове середнє значення;
- median (x) – вибіркова медіана (середнє значення впорядкованого ряду значень випадкової величини);
- var (x) – вибіркова дисперсія;
- Var (x) – незміщена оцінка дисперсії;
- stdev (x) – середнє квадратичне відхилення;
- Stdev (x) – незміщена оцінка середнього квадратичного відхилення;
- max (x) – максимальнє значення вибірки;
- min (x) – мінімальнє значення вибірки;
- mode (x) – вибіркова мода (значення вибірки, яке зустрічається частіше за інші);
- cvar (x1, x2) – коваріація двох вибірок;
- corr (x1, x2) – коефіцієнт кореляції двох вибірок;
- x – вектор (матриця) з вибіркою випадкових даних;
- x1, x2 – вектори (матриці) одного розміру з вибірками випадкових даних.

```
x := (0.6 2 2.8 3.5 4.2 4.6 4 3.2 2 4 3.7 4.1 2.8 3.1 4 3.4 2.5 2 3.9 2)T

mean(x) = 3.12          max(x) = 4.6
median(x) = 3.3        min(x) = 0.6
var(x) = 0.989         mode(x) = 2
stdev(x) = 0.994      sqrt(var(x)) = 0.994
```

Рис. 24 Використання вбудованих функцій MathCAD для розрахунку вибіркових статистичних характеристик

2.2 Побудова гістограми та полігона

Одним з етапів первинної обробки є групування та графічне представлення даних. Найбільш наочна форма графічного представлення групування – **гістограма**. Гістограма – це стовпчаста діаграма, яка відображає частотні дані.

Під час побудови гістограми область значень випадкової величини розбивається на відповідну кількість інтервалів k , а потім підраховується частота потрапляння даних в кожній інтервал. Для побудови гістограм в MathCAD є декілька вбудованих функцій.

1) Використання функції `hist`. Гістограми з довільними інтервалами (Рис. 25)

`hist(interval, x)` – вектор частоти потрапляння даних в інтервали гістограми:

- `interval` – вектор, елементи якого задають інтервали побудови гістограми в порядку збільшення;
- `x` – вектор випадкових даних.

Інша форма графічного представлення групованих даних – **полігон** (ламана лінія, яка з'єднує середини верхніх основ прямокутників гістограми).

```

N := 20
k := 5
x := (0.6 2 2.8 3.5 4.2 4.6 4 3.2 2 4 3.7 4.1 2.8 3.1 4 3.4 2.5 2 3.9 2)T

h :=  $\frac{\max(x) - \min(x)}{k} = 0.8$ 

i := 0..k

intervali := min(x) + h·i      gi :=  $\frac{1}{N \cdot h} \cdot \text{hist}(\text{interval}, x)$ 

```

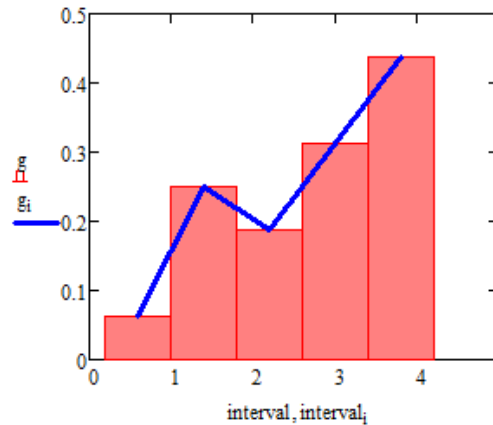


Рис. 25 Гістограми з довільними інтервалами та полігон

2) Використання функції `histogram` (спрощений варіант побудови) (Рис. 26)

`histogram(k, x)` – матриця гістограми розміру $k \times 2$, яка містить стовпчик інтервалів розбиття та стовпчик частоти потрапляння даних в інтервали:

k – кількість інтервалів гістограми;

x – вектор випадкових даних.

Для того, щоб назначити двомірному графіку тип **гістограма**, у діалоговому вікні «**Формат...**» («**Format...**») встановити на вкладці «**Трассировка**» («**Trace**») тип списку «**Столбцы**» («**bar**») або «**Гистограмма**» («**solidbar**»).

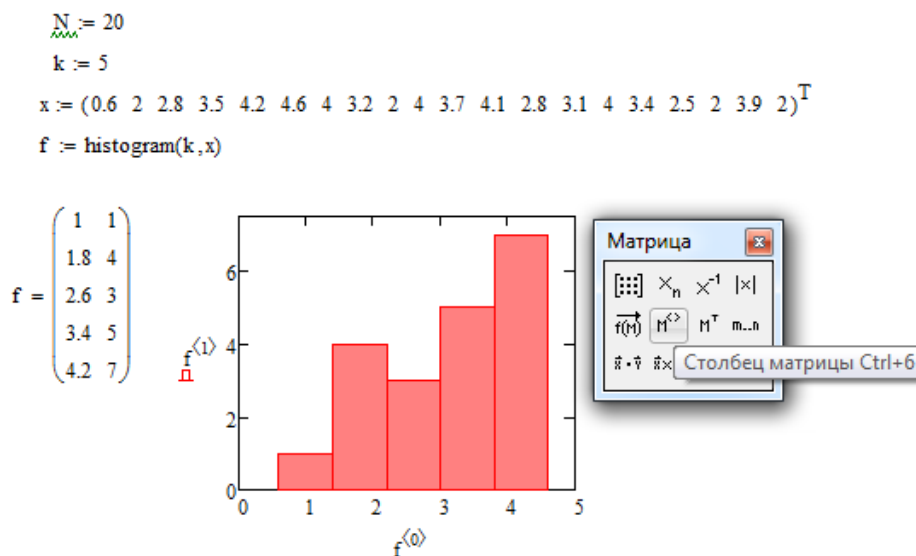


Рис. 26 Побудова гістограми за допомогою функції `histogram`

2.3 Функції розподілу ймовірностей

В MathCAD вбудована інформація про велику кількість різних статистичних розподілів, що використовуються в математичній статистиці: 1) табульовані функції ймовірності; 2) можливість генерації послідовності випадкових чисел з відповідним законом розподілу. Для реалізації цих можливостей в MathCAD є чотири основних категорії вбудованих функцій. Їх назви є складовими: перша літера ідентифікує відповідний закон розподілу, а частина, що залишилась (умовно позначена *) задає змістовну частину вбудованої функції:

$d^*(x, par)$ – щільність ймовірності;

$p^*(x, par)$ – функція розподілу;

$q^*(P, par)$ – квантиль розподілу (функція, зворотна функції розподілу);

$r^*(M, par)$ – вектор M незалежних випадкових чисел, кожне з яких має відповідний розподіл:

x – значення випадкової величини (аргумент функції);

P – значення ймовірності;

Par – список параметрів розподілу.

Наприклад, щоб отримати функції, які задають рівномірний розподіл, замість * необхідно записати `unif`, та ввести список параметрів: `runif(a, b)`.

Деякі типи розподілів, які реалізовані в MathCAD (* – це перша літера вбудованих функцій):

*`binom(k, n, p)` – біноміальний розподіл (n – цілий параметр, $0 \leq k \leq n$ та $0 \leq p \leq 1$ – параметр, якій дорівнює ймовірності успіху одиничного випробування);

*`chisq(x, d)` – χ^2 розподіл ($d > 0$ – число ступенів вільності);

*`exp(x, r)` – експоненціальний розподіл ($r > 0$ – показник експоненти);

*`F(x, d1, d2)` – розподіл Фішера ($d1, d2 > 0$ – числа ступенів вільності);

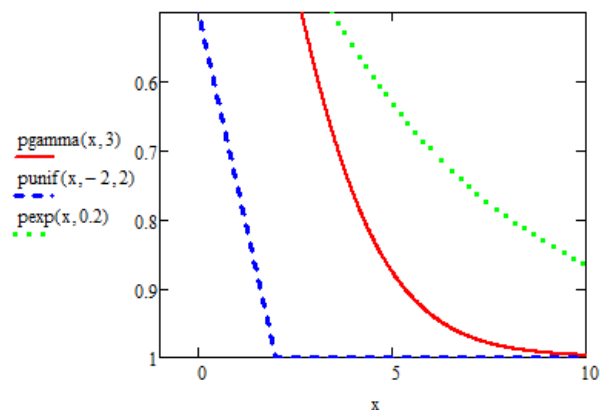
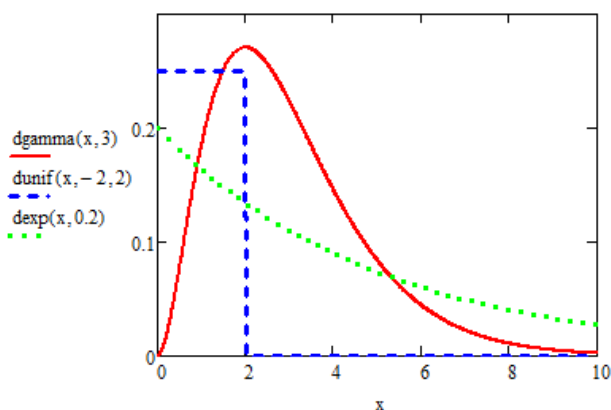
*`gamma(x, s)` – гамма-розподіл ($s > 0$ – параметр форми);

*`norm(x, μ , σ)` – нормальний розподіл (μ – математичне очікування, $\sigma > 0$ – середньоквадратичне відхилення);

*`t(x, d)` – розподіл Стюдента ($d > 0$ – число ступенів вільності);

*`unif(x, a, b)` – рівномірний розподіл ($a < b$ – границі інтервалу).

Приклади графіків щільності ймовірності, функцій розподілів, квантилей ймовірності показано на Рис. 27.



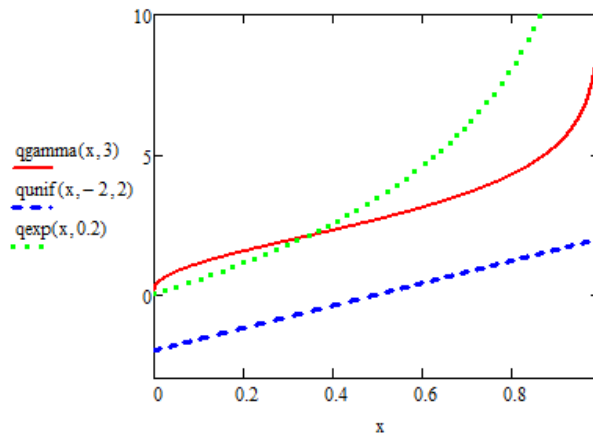


Рис. 27 Щільності ймовірності, функції розподілів, квантілі ймовірності

В теорії ймовірностей доведено, що сума різних незалежних випадкових величин (незалежно від їх закону розподілу) є випадковою величиною, яка розподілена відповідно до **нормального** закону. Тому нормальний розподіл добре моделює широкий спектр явищ, для яких відомо, що на них впливає декілька незалежних випадкових параметрів.

Система MathCAD має вбудовані функції для опису нормального розподілу ймовірностей:

$dnorm(x, \mu, \sigma)$ – щільність ймовірності нормального розподілу;

$pnorm(x, \mu, \sigma)$ – функція нормального розподілу;

$qnorm(P, \mu, \sigma)$ – зворотна функція нормального розподілу (квантиль);

$rnorm(M, \mu, \sigma)$ – вектор M незалежних випадкових чисел, кожне з яких має нормальний розподіл;

x – значення випадкової величини;

P – значення ймовірності;

μ – математичне очікування;

σ – середньоквадратичне відхилення.

Математичне очікування та дисперсія є параметрами нормального розподілу. Щільність розподілу для трьох пар значень параметрів показана на Рис. 28. Щільність розподілу $dnorm$ задає ймовірність потрапляння випадкової величини x в інтервал від x до $x+\Delta x$.

Функція розподілу $F(x)$ – це ймовірність того, що випадкова величина набуває значення, менше, або таке, що дорівнює x . Функції розподілу для трьох пар значень параметрів показана на Рис. 29.

Функція, зворотна до $F(x)$ – квантиль розподілу, дозволяє за заданим аргументом P визначити значення x , та випадкова величина буде менше або дорівнювати x з ймовірністю P .

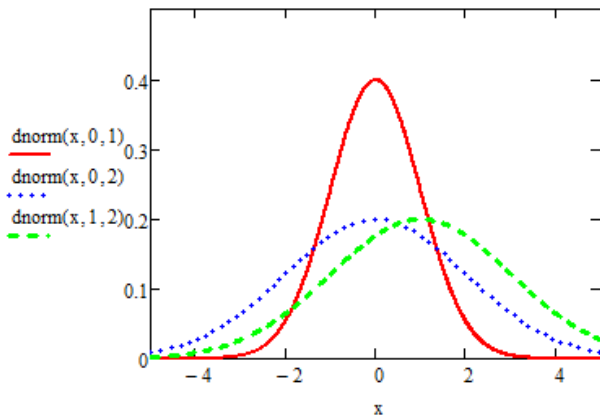


Рис. 28 Щільності ймовірності нормальних розподілів

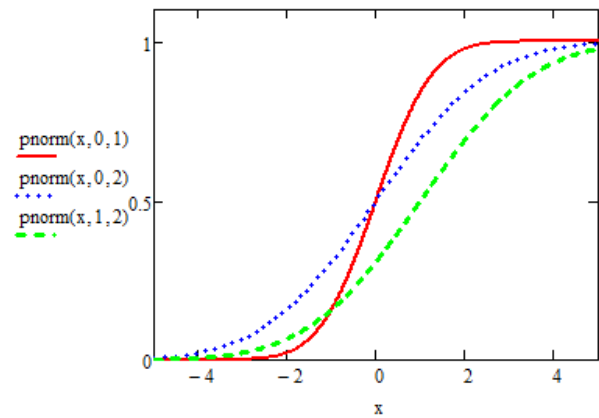


Рис. 29 Нормальні функції розподілу

Приклади використання розглянутих функцій (випадкова величина розподілена за нормальним законом з $\mu = 0$, $\sigma = 1$):

- 1) Генерування ряду даних з нормальним законом розподілу

$$N := 250 \quad \mu := 0 \quad \sigma := 1$$

$$x := \text{norm}(N, \mu, \sigma)$$

$x^T =$		23	24	25	26	27
	0	-0.245	0.09	1.262	-0.706	...

- 2) Ймовірність того, що величина x буде менше 2.5

$$\text{pnorm}(2.5, \mu, \sigma) = 0.994$$

- 3) Квантиль (95%) нормального розподілу

$$\text{qnorm}(0.95, \mu, \sigma) = 1.645$$

- 4) Ймовірність того, що величина x буде більше 2.5

$$1 - \text{pnorm}(2.5, \mu, \sigma) = 0.00621$$

- 5) Ймовірність того, що величина x буде належати інтервалу (2,3)

$$\text{pnorm}(3, \mu, \sigma) - \text{pnorm}(2, \mu, \sigma) = 0.021$$

2.4 Інтервальне оцінювання. Обчислення довірчого інтервалу для дисперсії

На практиці параметри генеральної сукупності, такі як математичне очікування і дисперсія, невідомі. Точкові оцінки цих параметрів (вибіркове середнє і вибіркова дисперсія) є випадковими величинами. При малому числі

спостережень точкові оцінки є малонадійними. Тому в статистиці використовують інтервальні оцінки. Для цього на числовій прямій будують інтервал, в який потрапляє параметр, що оцінюється, із заздалегідь обраною ймовірністю $1-\alpha$, близькою до одиниці.

Довірчий інтервал (θ_1, θ_2) [2, 9] – це інтервал, який із заданою ймовірністю $P=1-\alpha$, містить невідоме значення параметру θ , тобто

$$P(\theta_1 < \theta < \theta_2) = 1-\alpha.$$

Число $1-\alpha$ – довірна ймовірність, а значення α – рівень значущості.

У разі нормального розподілу, межі довірного інтервалу, а також довірна ймовірність не залежать від шуканого параметру.

Побудуємо, наприклад, довірчий інтервал для дисперсії. Завдання полягає в тому, що необхідно визначити числовий (довірчий) інтервал, якому буде належати з певною ймовірністю дисперсія нормальної випадкової величини, виходячи з деякого об'єму вибірки. Ця задача розв'язується в статистиці за допомогою χ^2 -розподілу.

Складемо для σ^2 оцінку $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ та розглянемо відношення:

$$\frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2$$

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – оцінка невідомого математичного очікування.

Можна довести, що випадкова величина $\frac{(n-1)s^2}{\sigma^2}$ має χ^2 -розподіл з $n-1$ ступенями вільності.

Для побудови довірного інтервалу з рівнем значущості α необхідно визначити квантілі χ_0^2 і χ_1^2 та виконати умову:

$$P\left(\chi_0^2 \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi_1^2\right) = 1-\alpha$$

$$\frac{(n-1)s^2}{\chi_1^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_0^2},$$

Таким чином, $\left(\frac{(n-1)s^2}{\chi_1^2}, \frac{(n-1)s^2}{\chi_0^2}\right)$ – довірчий інтервал для невідомої дисперсії з довірчою ймовірністю $1-\alpha$.

Квантілі χ_0^2 та χ_1^2 в MathCAD обчислюються за допомогою функції `qchisq(x, d)`.

Приклад обчислення довірного інтервалу для дисперсії нормального розподілу при відомому математичному очікуванні за допомогою інструментів MathCAD надано на Рис. 30.

$$\underline{\underline{N}} := 250 \quad \underline{\underline{\mu}} := 10 \quad \underline{\underline{\sigma}} := 2 \quad \alpha := 0.05 \quad P := 1 - \alpha = 0.95$$

$$\underline{\underline{x}} := \text{mom}(N, \mu, \sigma)$$

$x^T =$	0	1	2	3	4	5	6	7	
	0	11.269	10.584	10.283	10.695	9.233	9.86	8.072	...

$$\chi^2_{-0} := \text{qchisq}\left(\frac{\alpha}{2}, N - 1\right) \quad \chi^2_{-0} = 207.186$$

$$\chi^2_{-1} := \text{qchisq}\left(1 - \frac{\alpha}{2}, N - 1\right) \quad \chi^2_{-1} = 294.601$$

$$\underline{\underline{L}} := (N - 1) \cdot \frac{\text{Var}(x)}{\chi^2_{-1}} \quad L = 3.303$$

$$\underline{\underline{R}} := (N - 1) \cdot \frac{\text{Var}(x)}{\chi^2_{-0}} \quad R = 4.697$$

Рис. 30 Обчислення довірчого інтервалу для дисперсії

Таким чином, довірчий інтервал для заданої задачі – (3.303, 4.697).

2.5 Перевірка статистичних гіпотез

Статистичною гіпотезою H називають будь-яке припущення відносно параметрів або форми розподілу генеральної сукупності, яке перевіряється на основі даних вибірки [2, 9]. Висунуту гіпотезу називають **основною** або **нульовою** і позначають H_0 . Гіпотезу H_1 , яка суперечить основній, називають **альтернативною**.

Для перевірки статистичної гіпотези використовують спеціально підібрану випадкову величину ϕ , точний або наближений розподіл якої відомий. Цю величину називають **статистичним критерієм** або просто критерієм чи статистикою.

Сукупність значень критерію, при яких основна гіпотеза відхиляється, називають **критичною областю**, а сукупність значень, при яких гіпотезу приймають, – **областю прийняття гіпотези** або **областю допустимих значень**.

Основний принцип статистичної перевірки статистичних гіпотез – якщо спостережуване значення критерію належить критичній області, то основну гіпотезу **відхиляють**; якщо спостережуване значення критерію належить області допустимих значень – основну гіпотезу **приймають**.

Оскільки критерій ϕ – випадкова величина, то всі її можливі значення належать деякому інтервалу числової прямої (або всій числовій прямій). Отже, існують точки, які відокремлюють критичну область від області допустимих значень. Ці точки називають критичними точками критерію ϕ . Розрізняють односторонню (правосторонню і лівосторонню) і двосторонню критичні області.

Розглянемо алгоритм перевірки гіпотези про числове значення математичного очікування при невідомій дисперсії.

Сформулюємо нульову гіпотезу H_0 про те, що невідомий параметр μ дорівнює заданому числу μ_0 , тобто $H_0: \mu = \mu_0$. Альтернативну гіпотезу H_1 можна сформулювати трьома способами:

- 1) $H_1: \mu \neq \mu_0$;
- 2) $H_1: \mu > \mu_0$;
- 3) $H_1: \mu < \mu_0$.

Якщо $x=(x_1, x_2, \dots, x_n)$ – вибірка з нормального розподілу,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{– оцінка математичного очікування,}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{– оцінка дисперсії,}$$

то при перевірці нульової гіпотези $H_0: \mu = \mu_0$ використовується критерій:

$$\varphi = \frac{\bar{x} - \mu_0}{\sqrt{\frac{s^2}{n}}}$$

який при виконанні гіпотези H_0 має розподіл Стюдента з числом ступенів вільності $n-1$.

Розглянемо кожен з трьох випадків.

- 1) Нульова гіпотеза $H_0: \mu = \mu_0$ та альтернативна $H_1: \mu \neq \mu_0$.

Критична область двостороння (Рис. 31) і її границі визначаються умовами:

$$P(\varphi < x_{1-\alpha}) = 0.5\alpha \quad \text{та} \quad P(\varphi > x_{r\alpha}) = 0.5\alpha.$$



Рис. 31 Критична область для альтернативної гіпотези $H_1: \mu \neq \mu_0$

Розподіл Стюдента є симетричним, тому достатньо обчислити $x_{r\alpha}$, тому що $x_{1\alpha} = -x_{r\alpha}$.

Задамо рівень значущості α та обчислимо значення $x_{r\alpha}$ як розв'язання рівняння

$$F_{n-1}(x_{r\alpha}) = 1 - 0.5\alpha$$

де $F_{n-1}(x)$ – функція розподілу Стюдента з числом ступенів вільності $n-1$.

Якщо критичну область знайдено, можна обчислити за вибіркою значення критерію φ та перевірити, чи потрапляє воно в критичну область. Якщо $\varphi < x_{1\alpha}$,

або $\varphi > x_{r\alpha}$, то гіпотеза H_0 відхиляється і приймається гіпотеза H_1 . Якщо $x_{1\alpha} < \varphi < x_{r\alpha}$, то приймається гіпотеза H_0 .

Нижче наведено документ MathCAD, який містить обчислення для вибірки з файлу (Рис. 32).

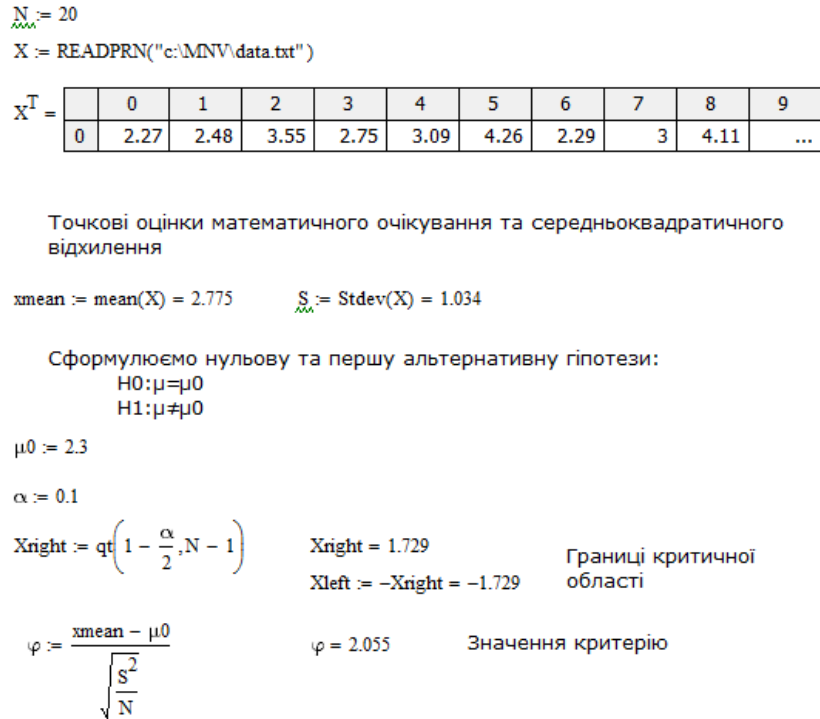


Рис. 32 Перевірка гіпотези H_0 проти першої альтернативної

Таким чином, критерій $\varphi = 2.055$ належить інтервалу $(-1.729, 1.729)$ і можна зробити висновок, що гіпотеза $H_0: \mu = \mu_0$ приймається, а гіпотеза $H_1: \mu \neq \mu_0$ відхиляється.

2) Нульова гіпотеза $H_0: \mu = \mu_0$ та альтернативна $H_1: \mu > \mu_0$.

Критична область значень критерію φ , при яких гіпотеза H_0 відхиляється, є правосторонньою. Критична область – інтервал $(x_{r\alpha}, +\infty)$ (Рис. 33)

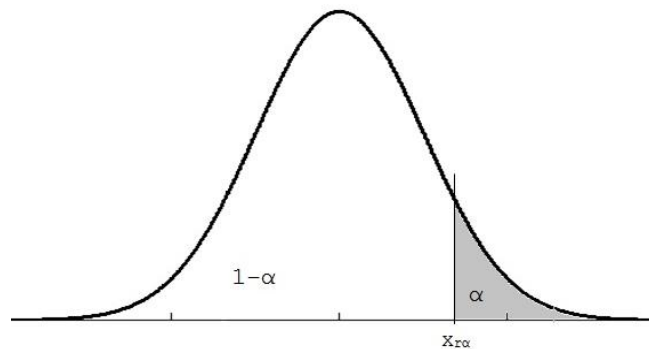


Рис. 33 Критична область для альтернативної гіпотези $H_1: \mu > \mu_0$

Критична точка задовольняє вимозі $P(\varphi > x_{r\alpha}) = \alpha$ і знаходиться (з відповідним рівнем значущості α) як розв'язання рівняння:

$$F_{n-1}(x_{r\alpha}) = 1 - \alpha$$

Якщо $\varphi > x_{r\alpha}$, то гіпотеза H_0 відхиляється, а гіпотеза H_1 приймається.

Приклад розв'язання в системі MathCAD:

$$\begin{aligned} x_{\text{mean}} &:= \text{mean}(X) = 2.775 & S &:= \text{Stdev}(X) = 1.034 \\ \text{Сформулюємо нульову та другу альтернативну гіпотези:} \\ H_0: \mu &= \mu_0 \\ H_1: \mu &> \mu_0 \\ \mu_0 &:= 2.3 \\ \alpha &:= 0.1 \\ X_{\text{right}} &:= \text{qt}(1 - \alpha, N - 1) & X_{\text{right}} &= 1.328 & \text{Границі критичної} \\ X_{\text{left}} &:= -X_{\text{right}} = -1.328 & & & \text{області} \\ \varphi &:= \frac{x_{\text{mean}} - \mu_0}{\sqrt{\frac{S^2}{N}}} & \varphi &= 2.055 & \text{Значення критерію} \end{aligned}$$

Рис. 34 Перевірка гіпотези H_0 проти другої альтернативної

Для даних попереднього прикладу, але для $\mu_0 = 2.3$, отримуємо, що $\varphi = 2.055$ і $\varphi > x_{r\alpha}$ ($2.055 > 1.328$), тому гіпотеза H_0 відхиляється, а гіпотеза H_1 приймається.

3) Нульова гіпотеза $H_0: \mu = \mu_0$ та альтернативна $H_1: \mu < \mu_0$.

Критична область значень критерію φ , при яких гіпотеза H_0 відхиляється, є лівосторонньою. Критична область – інтервал $(-\infty, x_{1\alpha})$ (Рис. 35)

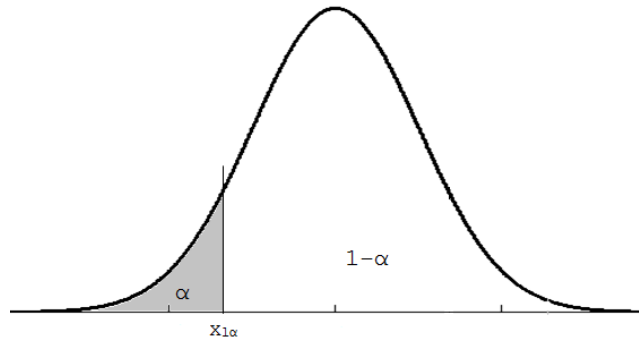


Рис. 35 Критична область для альтернативної гіпотези $H_1: \mu < \mu_0$

Критична точка задовольняє вимозі $P(\varphi < x_{1\alpha}) = \alpha$ і знаходиться (з відповідним рівнем значущості α) як розв'язання рівняння:

$$F_{n-1}(x_{1\alpha}) = 1 - \alpha$$

Якщо $\varphi < x_{1\alpha}$, то гіпотеза H_0 відхиляється, а гіпотеза H_1 приймається.

Приклад розв'язання в системі MathCAD:

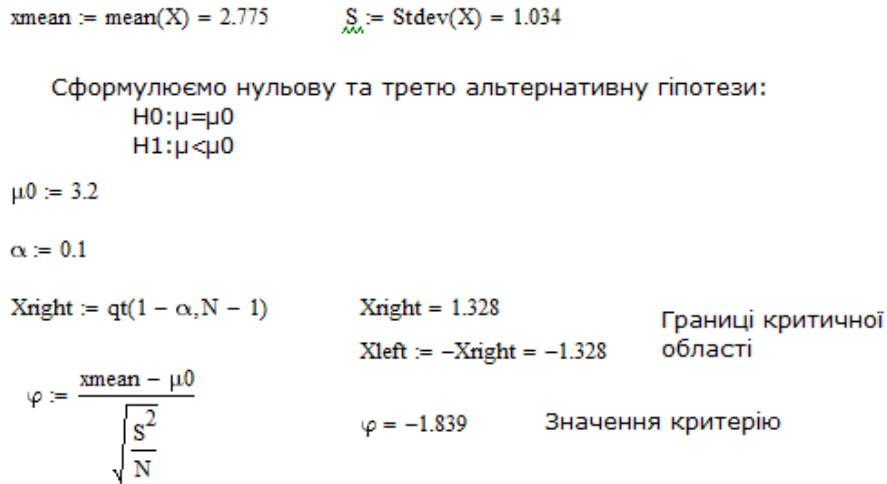


Рис. 36 Перевірка гіпотези H_0 проти третьої альтернативної

для $\mu_0=3.2$, отримуємо, що $\varphi=-1.839$ і $\varphi < x_{1\alpha}$ ($-1.839 < -1.328$), тому гіпотеза H_0 відхиляється, а гіпотеза H_1 приймається.

2.6 Елементи кореляційного та регресійного аналізу

В багатьох задачах потрібно встановити та оцінити залежність випадкової величини Y від однієї або декількох інших величин.

Дві випадкові величини можуть бути пов'язані **функціональною залежністю** (залежність виду $Y=f(X)$, де кожному значенню незалежної змінної X відповідає єдине значення змінної Y), або статистичною залежністю, або бути незалежними.

Функціональна залежність виникає рідко, тому що на обидві величини або на одну з них можуть впливати випадкові фактори. В цьому випадку виникає статистична залежність [2, 9].

Статистичною називають залежність, якщо кожному значенню X відповідає ряд значень Y . Зокрема, статистична залежність виявляється у тому, що при зміні однієї величини змінюється середнє значення іншої; така залежність називається **кореляційною** (Рис. 37). **Кореляційним полем** називають нанесені на графік точки, які відповідають одночасно значенням двох величин.

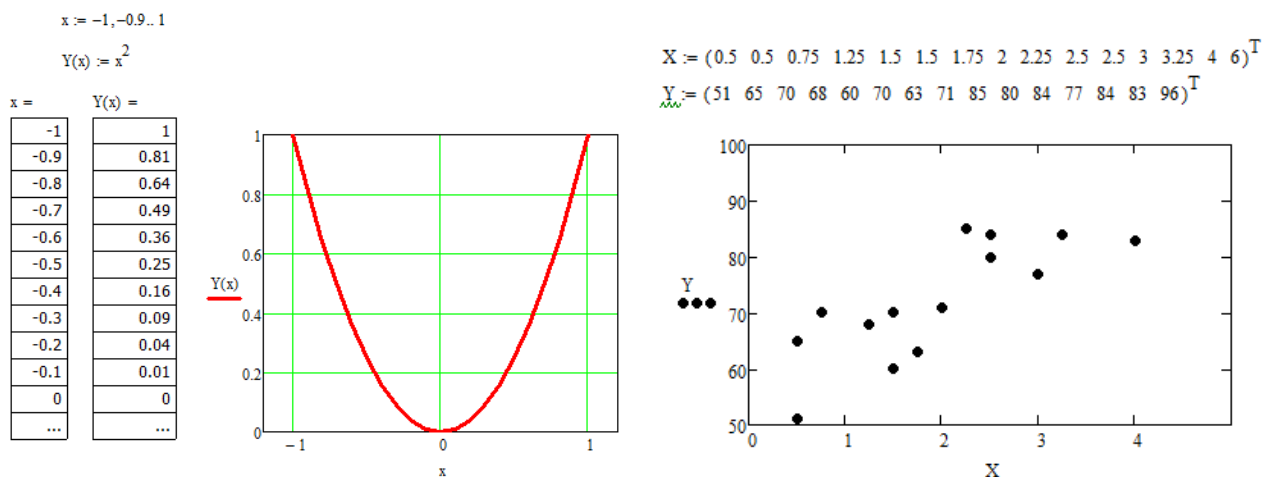


Рис. 37 Приклад функціональної та кореляційної залежності

Статистичні зв'язки між змінними можна вивчати методами кореляційного й регресійного аналізу. Основним завданням кореляційного аналізу є виявлення зв'язку між випадковими змінними та оцінка її тісноти.

Кількісна оцінка тісноти зв'язку між досліджуваними параметрами встановлюється за допомогою **коефіцієнту кореляції**.

Для обчислення лінійного коефіцієнту кореляції Пірсона r_{xy} використовується формула [2]:

$$r_{xy} = \frac{n \sum x \cdot y - \sum x \cdot \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}}$$

Тісноту зв'язку між двома величинами можна визначити візуально по полю кореляції (Рис. 38).

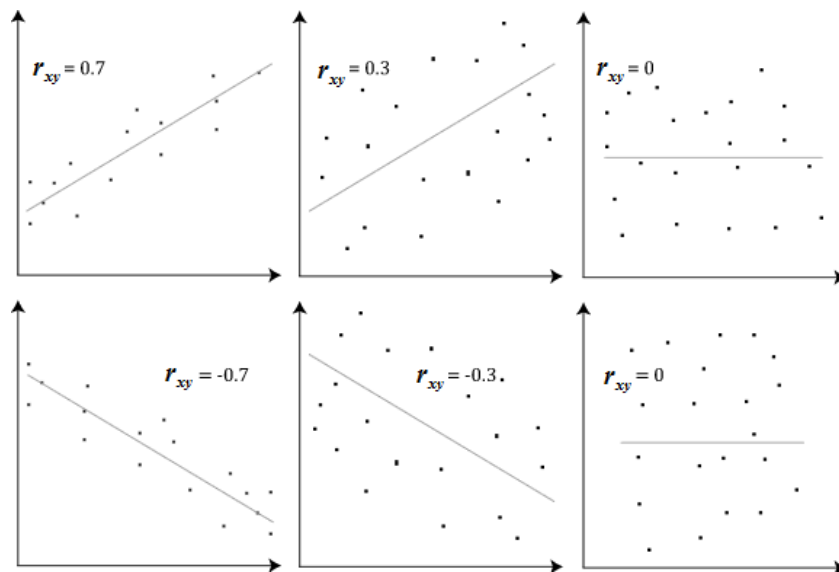


Рис. 38 Типи зв'язку між ознаками

Таким чином, коефіцієнт кореляції має такі властивості:

- 1) r_{xy} змінюється від -1 до 1 ;
- 2) якщо величини X та Y незалежні, $r_{xy}=0$;
- 3) якщо $|r_{xy}|=1$, то величини X та Y зв'язані лінійною функціональною залежністю;
- 4) якщо $|r_{xy}| \rightarrow 1$, то зв'язок між величини X та Y сильний;
- 5) якщо $0.3 < |r_{xy}| < 0.5$, то зв'язок між величини X та Y слабкий;
- 6) якщо $r_{xy} > 0$, то зв'язок прямий; якщо $r_{xy} < 0$, то зв'язок зворотний.

В MathCAD обчислення коефіцієнту кореляції виглядатиме таким чином (Рис. 39):

$$X := (0.5 \ 0.5 \ 0.75 \ 1.25 \ 1.5 \ 1.5 \ 1.75 \ 2 \ 2.25 \ 2.5 \ 2.5 \ 3 \ 3.25 \ 4 \ 6)^T$$

$$Y := (51 \ 65 \ 70 \ 68 \ 60 \ 70 \ 63 \ 71 \ 85 \ 80 \ 84 \ 77 \ 84 \ 83 \ 96)^T$$

$$\text{corr}(X, Y) = 0.845$$

$$\text{ccorel} := \frac{\text{cvar}(X, Y)}{\text{stdev}(X) \cdot \text{stdev}(Y)} = 0.845$$

Рис. 39 Обчислення коефіцієнту кореляції

Значення коефіцієнту кореляції $r_{xy}=0.845$ свідчить про суттєвий прямий зв'язок між величинами X та Y .

Основним завданням регресійного аналізу є встановлення форми залежності між заданими параметрами, визначення функції регресії та вивчення залежності між змінними.

Розглянемо лінійну регресію. Математичне рівняння, яке оцінює лінію парної лінійної регресії має вид:

$$y=b+kx,$$

b, k – коефіцієнти регресії.

В MathCAD для обчислення коефіцієнтів регресії призначені відповідні функції:

- для обчислення коефіцієнту k – функція $\text{slope}(X, Y)$;
- для обчислення коефіцієнту b – функція $\text{intercept}(X, Y)$.

Приклад побудови кореляційного поля та лінії регресії в MathCAD наведено на Рис. 40.

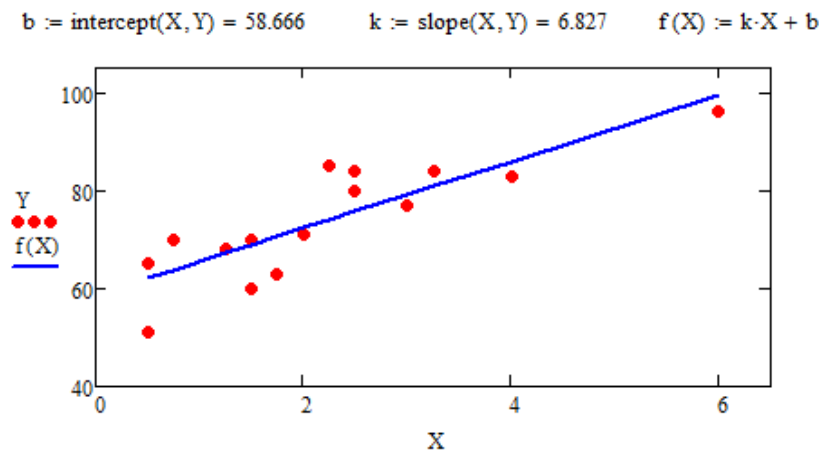


Рис. 40 Обчислення коефіцієнтів лінійної регресії. Побудова лінії регресії

2.7 Однофакторний дисперсійний аналіз

Дисперсійний аналіз застосовується для дослідження впливу однієї або декількох якісних змінних (факторів) на одну залежну кількісну змінну [2].

Суть дисперсійного аналізу полягає в поділі загальної дисперсії досліджуваної ознаки на окремі компоненти, які зумовлені впливом конкретних факторів, і перевірці гіпотез про значущість впливу цих факторів на досліджувану ознаку. Порівнюючи компоненти дисперсії за допомогою F-критерію Фішера, можна визначити, наскільки результативна ознака зумовлена дією регульованих факторів.

На практиці часто зустрічається ситуація, коли можна вказати один фактор, що впливає на кінцевий результат, і цей фактор приймає кінцеве число значень. Така ситуація може бути проаналізована за допомогою **однофакторного дисперсійного аналізу**.

Розглянемо математичну модель, в якій передбачається, що кожна випадкова величина x_{ij} може бути представлена у вигляді $x_{ij} = a_i + \varepsilon_{ij}$, де a_i – досліджувана ознака, характерна для i -го рівня фактору, ε_{ij} – незалежні випадкові величини

випадкові величини, які описують сумарний внесок всіх випадкових факторів, що впливають на кінцевий результат. Найчастіше вважають, що всі ε_{ij} мають нормальний розподіл з нульовим математичним очікуванням і з однаковою дисперсією σ^2 . Необхідно з'ясувати, чи впливає вибір рівня фактору на досліджувану ознаку.

Таким чином, за результатами експерименту необхідно перевірити нульову гіпотезу про рівність математичних очікувань (вибір фактору **не впливає** на ознаку)

$$H_0: a_1 = a_2 = \dots = a_n.$$

Аналіз результатів буде заснований на порівнянні двох оцінок невідомої дисперсії σ^2 . Якщо обидві оцінки близькі, то гіпотеза H_0 приймається. Якщо оцінки суттєво відрізняються, гіпотеза H_0 відхиляється.

Приклад: Нехай на кількісну нормально розподілену ознаку X впливає фактор, який має чотири рівні. В таблиці наведено дані спостережень для кожного рівня. Реалізація в MathCAD – на Рис. 41

Рівні фактора	Ознака X		
	Значення X під впливом фактору рівня 1	Значення X під впливом фактору рівня 2	Значення X під впливом фактору рівня 3
1	40	41	40
2	50	49	50
3	47	47	
4	44	47	42

Побудуємо середні та оцінки дисперсії:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad i = 1, 2, \dots, m \quad s_1^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} \quad s^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

де m – кількість рівнів $n = n_1 + n_2 + \dots + n_m$ – кількість спостережень.

Порівняємо оцінки s_1^2 та s^2 . Якщо гіпотеза H_0 вірна, то величина

$$F_H = \frac{\frac{s^2}{n-1}}{\frac{s_1^2}{n-m}} = \frac{s^2(n-m)}{s_1^2(n-1)}$$

має розподіл Фішера з $n-1$ та $n-m$ ступенями вільності.

Задамо достатньо малий рівень значущості α і розв'яжемо рівняння:

$$F_{n-1, n-m}(\alpha) = 1 - \alpha$$

Порівняємо корінь цього рівняння x_α із значенням F_H . Якщо $F_H > x_\alpha$, гіпотеза H_0 відхиляється, в протилежному випадку – приймається.

$n_1 := 3$	$x_{1,1} := 40$	$x_{1,2} := 41$	$x_{1,3} := 40$	
$n_2 := 3$	$x_{2,1} := 50$	$x_{2,2} := 49$	$x_{2,3} := 50$	
$n_3 := 2$	$x_{3,1} := 47$	$x_{3,2} := 47$		
$n_4 := 3$	$x_{4,1} := 44$	$x_{4,2} := 47$	$x_{4,3} := 42$	
$m := 4$			Кількість факторів	
$i := 1..m$				
$N := \sum_{i=1}^m n_i = 11$			Кількість вимірів	
$X_i := \frac{\sum_{j=1}^{n_i} x_{i,j}}{n_i}$	$s_{1_2} := \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{i,j} - X_i)^2 = 14$		Оцінка дисперсії s_1^2	
$XN := \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} x_{i,j}}{N}$	$s_{_2} := \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{i,j} - XN)^2 = 153.636$		Оцінка дисперсії s^2	
$\alpha := 0.05$	$x_{\alpha} := qF(1 - \alpha, N - 1, N - m)$	$x_{\alpha} = 3.637$	Квантиль розподілу Фішера з рівнем значущості 0.05	
	$FH := \frac{s_{_2} \cdot (N - m)}{s_{1_2} \cdot (N - 1)}$	$FH = 7.682$	Значення критерія	
	$r := \frac{s_{_2} - s_{1_2}}{s_{_2}}$	$r = 0.909$	Коефіцієнт детермінації	

Рис. 41 Реалізація однофакторного дисперсійного аналізу

В результаті обчислень отримано $F_H = 7.682$ та $x_{\alpha} = 3.637$, тобто $F_H > x_{\alpha}$. Таким чином, гіпотеза H_0 відхиляється і можна зробити висновок, що вибір рівню фактора **впливає** на ознаку X . В розглянутому прикладі коефіцієнт детермінації $r = 0.909$, тобто можна сказати, що на 90.9% результат ознаки X обумовлений фактором, який на неї впливає.

2.8 Реалізація метода Монте-Карло в системі MathCAD

Для моделювання різних фізичних, економічних та інших явищ широко розповсюджені статистичні методи, які називають методами Монте-Карло, основна ідея яких полягає в генерації випадкових послідовностей, що моделюють відповідний ефект, наприклад шум в фізичному експерименті. [2]

Приклад моделювання сигналу та шуму наведено на Рис. 42. Для генерації одного псевдовипадкового числа в MathCAD використовується функція:

$\text{rnd}(x)$ – випадкове число, яке має рівномірну щільність розподілу на інтервалі $(0, x)$.

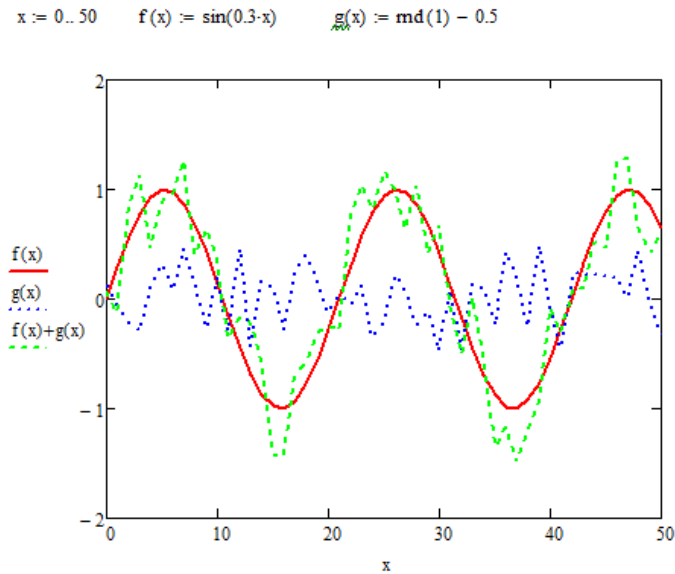


Рис. 42 Приклад візуалізації сигналу та шуму у фізичному експерименті

У XVIII столітті Жорж Луї Леклерк де Бюффон сформулював задачу про знаходження ймовірності того, що голка, яку кидають на лист розграфленого паперу, перетинає одну з ліній. Виявилось, що ця ймовірність пов'язана з числом π , і це зробило можливим пошук цього числа ймовірнісними методами, а саме, методом Монте-Карло.

Реалізацію метода Монте-Карло для пошуку числа π в MathCAD представлено на Рис. 43.

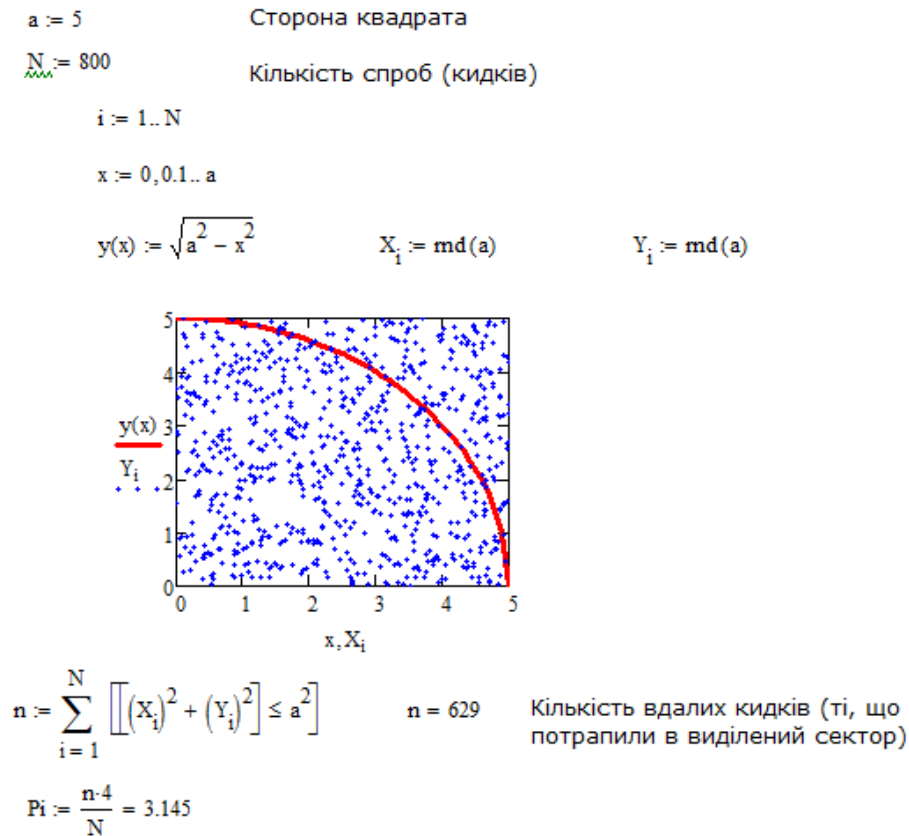


Рис. 43 Обчислення числа π

Контрольні запитання

1. Які вбудовані функції MathCAD застосовуються для обчислення числових статистичних характеристик рядів даних?
2. Поняття гістограми і полігону.
3. За допомогою яких функцій MathCAD можна згенерувати послідовність випадкових чисел з певним законом розподілу?
4. Поняття довірчого інтервалу.
5. Яку статистичну гіпотезу називають основною (нульовою)?
6. Яка залежність між параметрами називається кореляційною?
7. Яке основне завдання регресійного аналізу?
8. Для яких задач статистики застосовується метод факторного аналізу?
9. Яка основна ідея методів Монте-Карло?

3 ОСНОВНІ ІНСТРУМЕНТИ ТА ЗАСОБИ РОБОТИ З МАТЕМАТИЧНИМ ПАКЕТОМ MAPLE

3.1 Інтерфейс користувача математичного пакету Maple

Система символічної (аналітичної) математики Maple – визнаний світовий лідер в області аналітичних обчислень [1, 3, 6]. Інтерфейс користувача Maple (Рис. 44) містить одночасно текстові коментарі, команди вхідної мови (з можливістю перетворення їх в звичну математичну форму), результати обчислень та графічні дані.

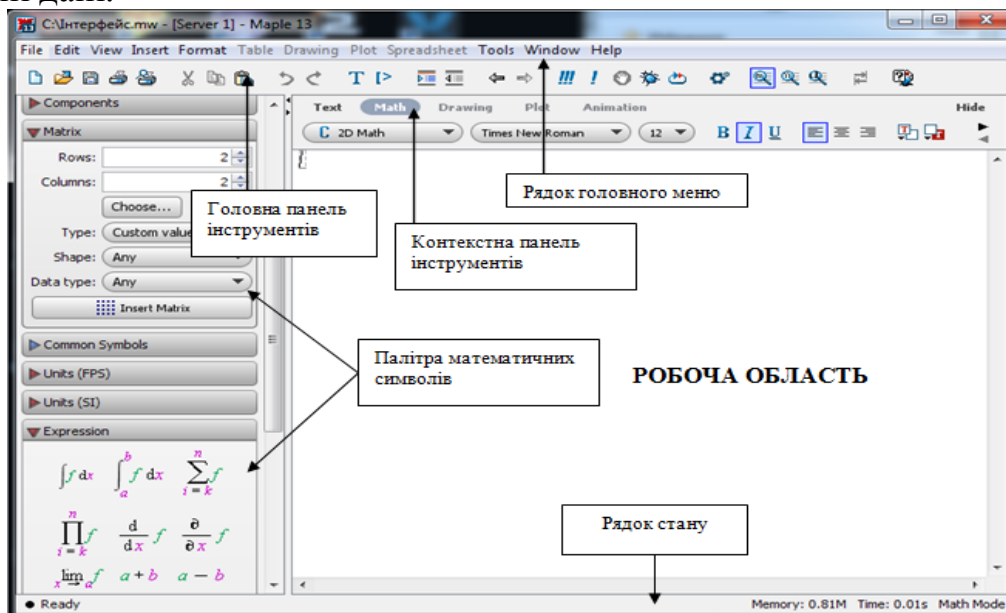


Рис. 44 Елементи вікна Maple

Головне меню Maple є контекстно-залежним, і його вигляд може змінюватися в залежності від поточного стану системи і містить такі команди:

- File – робота з файлами та друк документів;
- Edit – команди редагування документу;
- View – команди управління виглядом інтерфейсу;
- Insert – команди вставки;

Format – команди управління форматом відображення об'єктів Maple;

Table – активується, якщо вставлена таблиця («Insert» / «Table»)

Drawing – активується, якщо вставлено, наприклад, зображення («Insert» / «Image»);

Plot – активується, якщо користувач працює з графіками;

Spreadsheet – активується, якщо вставлена електронна таблиця («Insert» / «Spreadsheet»)

Tools – допоміжні команди;

Window – стандартні команди управління вікнами відкритих документів;

Help – довідкова система Maple.

В системі Maple підтримуються два режими функціонування:

– режим Worksheet («File» / «New» / «Worksheet Mode») (Рис. 45).

Створений шаблон орієнтовано на традиційну для Maple форму взаємодії. Переключення в режим Document здійснюється при створенні блоку документа: «Format» / «CreateDocumentBlock»;

– режим Document («File» / «New» / «Document») (Рис. 46) – це вільний від синтаксису режим, в якому знак позиціонування рядка вводу \triangleright не відображається. Передбачено можливості редагування типових властивостей документів (вибір шрифту, створення структури документу і т.д.).

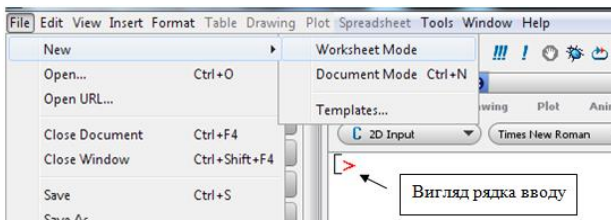


Рис. 45 Режим функціонування Worksheet

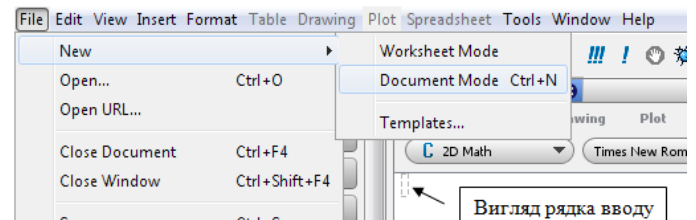


Рис. 46 Режим функціонування Document

Maple дозволяє обрати режим представлення виразів в рядку введення (Рис 47):

– введення виразів в Maple-нотації;

– введення виразів в звичайному математичному вигляді.

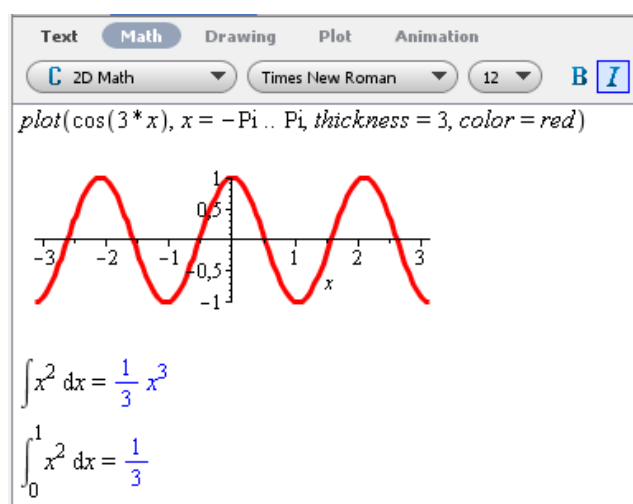
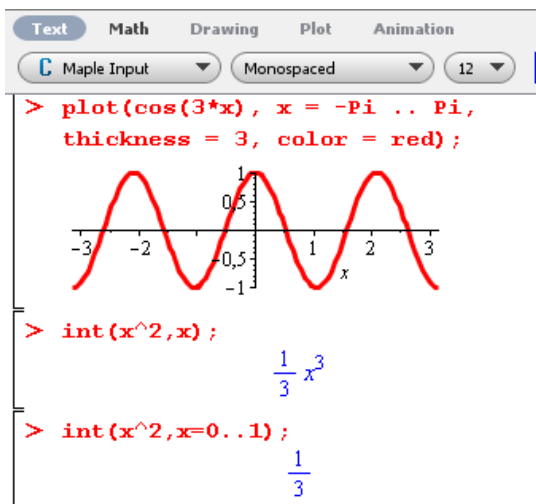


Рис. 47 Режими представлення виразів

Стандартна команда Maple складається з імені команди та її параметрів, які розділяються комами. Наприкінці кожної команди в попередніх версіях повинен бути знак (;) або (:). У поточних версіях Maple для обчислень можна не використовувати знак (;) в режимі Math; в режимі Text залишається класичний синтаксис команд. Якщо команда Maple завершується символом (:), команда виконується системою, але результат на екран не виводиться.

Система Maple містить велику кількість бібліотек (пакетів) (Рис. 48), підключення яких здійснюється за допомогою команди

```
with(ім'я_пакету)
```

```
with(linalg) :
```

```
with(plots)
```

```
[animate, animate3d, animatecurve, arrow, changecoords, complexplot, complexplot3d, conformal, conformal3d, contourplot, contourplot3d, coordplot, coordplot3d, densityplot, display, dualaxisplot, fieldplot, fieldplot3d, gradplot, gradplot3d, graphplot3d, implicitplot, implicitplot3d, inequal, interactive, interactiveparams, intersectplot, listcontplot, listcontplot3d, listdensityplot, listplot, listplot3d, loglogplot, logplot, matrixplot, multiple, odeplot, pareto, plotcompare, pointplot, pointplot3d, polarplot, polygonplot, polygonplot3d, polyhedra_supported, polyhedraplot, rootlocus, semilogplot, setcolors, setoptions, setoptions3d, spacecurve, sparsematrixplot, surfdata, textplot, textplot3d, tubeplot]
```

```
with(geometry) :
```

Рис. 48 Підключення пакетів Maple

3.2 Спеціальні символи та основні об'єкти Maple

Для скасування всіх зроблених призначень і початку нового сеансу без виходу з системи, використовується команда `restart`.

В Maple використовуються спеціальні символи (Рис. 49):

- види скобок: круглі, квадратні (для роботи з індексними величинами, списками), фігурні (для роботи з множинами)
- дві послідовні точки (..) в параметрах команд використовуються для змінювання значення змінної в заданому інтервалі;
- (‘) використовується як обмежувач рядка;
- (“) задає рядок;
- \$ задає впорядкованість послідовностей;
- % служить для використання результату попередньої команди в якості аргументу нової операції;
- знак # – покажчик коментаря;
- знак := – оператор присвоювання;
- знак = – використовується при формуванні рівнянь;
- D – для роботи з диференціальним оператором.

Найпростішими об'єктами є числа, константи, змінні, рядки.

Числа можуть бути дійсними або комплексними (Рис. 50б). Дійсні числа – цілі та раціональні. Раціональні числа можна представити у трьох видах:

- раціональна дріб, наприклад $7/31$;
- з плаваючою комою, наприклад 6.728 ;
- в експоненціальній формі, наприклад $1.709e-9$.

Десяткова точка має особливий статус: якщо вказати її у будь-якому місці числа, або наприкінці числа, то число стає дійсним з плаваючою комою.

Кількість цифр, які виводяться після десяткової точки, можна задавати за допомогою системної змінної `Digits`.

Основні математичні константи:

`Pi` – число π , `I` – уявна одиниця, `infinity` – нескінченність.

Змінні ідентифікуються іменем – набором символів, які починаються з літери (великі та малі літери розрізняються).

Рядок – будь-який набір символів в подвійних лапках.

<code>x := [1, 2, 5, 7, 8,]</code>	<code>Digits := 5</code>	
	<code>[1, 2, 5, 7, 8]</code>	<code>5</code>
<code>z := {a, b, c}</code>	<code>{a, b, c}</code>	<code>$\frac{31}{58}$</code>
<code>k := 1..5</code>	<code>1..5</code>	<code>$\frac{31}{58}$</code>
<code>restart</code> <code>x := \$1..5</code>	<code>1, 2, 3, 4, 5</code>	<code>0.53448</code>
<code>z\$6</code>	<code>z, z, z, z, z, z</code>	<code>evalf($\frac{31}{58}$)</code>
<code>y² - 1 = 0</code> #рівняння	<code>y² - 1 = 0</code>	<code>k := 7001.8</code>
		<code>7001.8</code>
<code>s1 := "String"</code>	<code>"String"</code>	<code>k1 := 7.0018e3</code>
		<code>7001.8</code>
<code>s2 := 'Symbol'</code>	<code>Symbol</code>	<code>p := x + I·y</code> # комплексне число
		<code>x + I·y</code>
<code>'Map'</code>	<code>Map</code>	<code>Pi</code>
		<code>π</code>
<code>cat(%, 'le')</code>	<code>Maple</code>	<code>infinity</code>
		<code>∞</code>

Рис. 49 Використання спеціальних символів Maple

Рис. 50 Числа та константи Maple

3.3 Функції

Система Maple має повний набір математичних функцій. Звернення до вбудованої функції – елементарна операція, якщо знати ім'я потрібної функції.

Приклади вбудованих функцій Maple:

<code>exp(x)</code>	Експоненціальна функція
<code>ln(x)</code>	Натуральний логарифм
<code>log10(x)</code>	Логарифм за основою 10
<code>sqrt(x)</code>	Квадратний корінь
<code>abs(x)</code>	Абсолютне значення
<code>sin(x), cos(x)</code>	<code>sinx, cosx</code>
<code>cot(x)</code>	<code>ctgx</code>
<code>ceil(x)</code>	Найменше ціле, яке більше або дорівнює x

<code>floor(x)</code>	Найбільше ціле, яке менше або дорівнює x
<code>round(x)</code>	Округлене значення числа
<code>signum(x)</code>	Функція знака (-1 при $x < 0$, 0 при $x = 0$, $+1$ при $x > 0$)

Розглянемо основні способи представлення функції в Maple.

1. Задання функції алгебраїчним виразом за допомогою оператора присвоювання (`:=`) (Рис. 51). Якщо задати конкретне значення x , то результат дорівнюватиме значенню функції при поточному x .

$$\begin{array}{l}
 f := \sin(x) - \cos(2 \cdot x) \\
 x := \frac{\pi}{4} \\
 f \\
 f\left(\frac{\pi}{4}\right)
 \end{array}
 \qquad
 \begin{array}{l}
 \sin(x) - \cos(2x) \\
 \frac{1}{4} \pi \\
 \frac{1}{2} \sqrt{2} \\
 \frac{1}{2} \sqrt{2}
 \end{array}$$

Рис. 51 Задання функції за допомогою оператора присвоювання

2. Використання **підстановки** (Рис. 52), синтаксис якої:

$$\text{subs}(\{x_1=a_1, x_2=a_2, \dots\}, f),$$

в фігурних скобках вказані змінні x_1, x_2 та їх нові значення a_1, a_2 , які потрібно підставити у функцію f .

$$\begin{array}{l}
 f := x \cdot \exp(-\lambda) \\
 \text{subs}(\{x=5, \lambda=3\}, f) \\
 \text{evalf}(\%)
 \end{array}
 \qquad
 \begin{array}{l}
 x e^{-\lambda} \\
 5 e^{-3} \\
 0.2489353418
 \end{array}$$

Рис. 52 Задання функції за допомогою підстановки

3. Визначення функції за допомогою **функціонального оператора** ставить у відповідність набору змінних (x_1, x_2, \dots) одно або декілька виразів (f_1, f_2, \dots) (Рис. 53).

$$\begin{array}{l}
 f := (x, y) \rightarrow \cos(x - y) \\
 f\left(\frac{\pi}{2}, 0\right) \\
 g := x \rightarrow \ln(x + 0.5) \\
 g(1)
 \end{array}
 \qquad
 \begin{array}{l}
 (x, y) \rightarrow \cos(x - y) \\
 0 \\
 x \rightarrow \ln(x + 0.5) \\
 0.4054651081
 \end{array}$$

Рис. 53 Задання функції за допомогою функціонального оператора

3.4 Побудова графіків функцій в системі Maple

Для побудови графіків функції $y=f(x)$ однієї змінної, $a \leq x \leq b$, $c \leq y \leq d$ використовується команда (Рис. 54)

```
plot(f(x), x=a..b, y=c..d, parameters)
```

parameters – параметри управління зображенням.

Основні параметри команди plot:

- title=text – заголовок графіка;
- color – колір кривих (red, green, blue і т.д.);
- coords=polar – встановлення полярних координат (за замовчуванням – декартові);
- style – стиль побудови графіка (point – точка, line – лінія);
- linestyle – стиль лінії (1 – суцільна лінія, 2 – точки, 3 – пунктир, 4 – штрихпунктир)
- axes – встановлення виду координатних осей;
- thickness=n – визначення товщини лінії;
- symbol – вид символу для точок графіка (circle – коло, point – точка) і т.д.

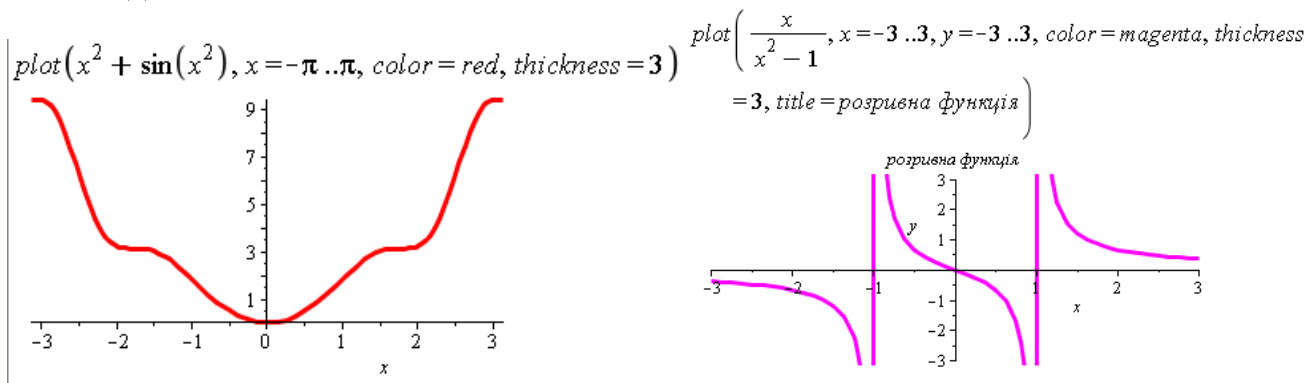


Рис. 54 Побудова графіків функцій, заданих явно

За допомогою команди plot, крім графіків функцій, заданих явно, можна будувати графіки функцій, які задано **параметрично** $y=y(t)$, $x=x(t)$ (Рис. 55):

```
plot([y=y(t), x=x(t), t=a..b], parameters)
```

```
plot([sin(3*t), cos(5*t), t=0..2*pi], color=blue, thickness=2)
```

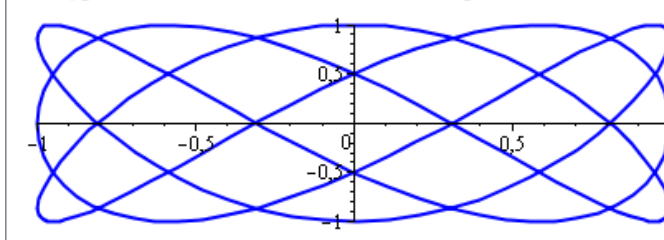


Рис. 55 Побудова графіків функцій, заданих параметрично

Для побудови графіка в **полярних координатах** необхідно вказати вид координатної осі (Рис. 56):

```
plot(cos(3·φ), φ=0..2·π, coords=polar, color=coral, thickness=2)
```

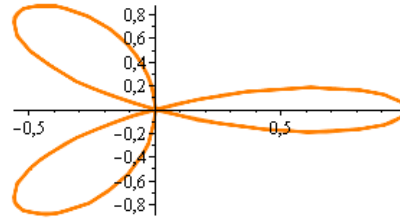


Рис. 56 Побудова графіків функцій, заданих в полярних координатах

Якщо потрібно побудувати декілька графіків функцій, в якості першого аргументу команди необхідно задати **список**, який містить алгебраїчні вирази залежностей (Рис. 57).

```
plot([sin(x), cos(x)], x=-π..π, color=[red, green], linestyle=[1, 3], thickness=2)
```

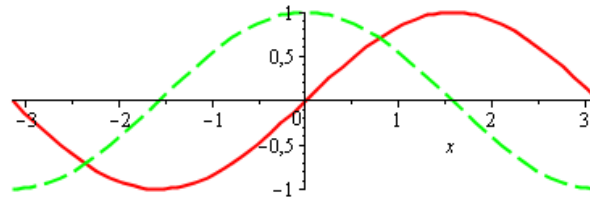


Рис. 57 Побудова декількох графіків функцій

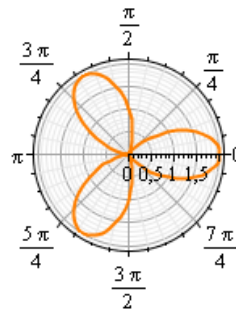
Використання графіки пакету plots

Пакет `plots` містить графічні функції, які значно розширюють можливості побудови різних видів графіків (Рис. 58). Пакет підключається командою

```
with(plots)
```

```
with(plots) :
```

```
polarplot(1 + cos(3·φ), φ=0..2·π, color=coral, thickness=2)
```



```
display({plot(sin(x), x=-π..π, color=black, thickness=2), plot(x^2 - 1, x=-2..2, y=-2..2, color=blue, thickness=2), plot(x^3, x=-2..2, thickness=2)})
```

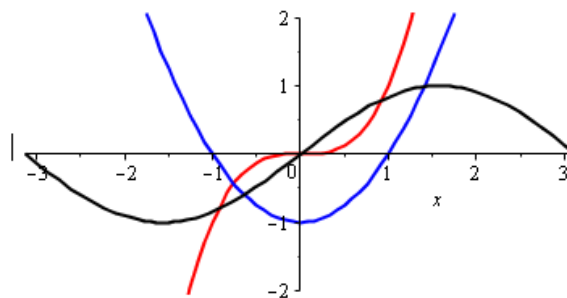


Рис. 58 Приклади використання пакету plots

3.5 Робота з файлами в системі Maple

Обмін інформацією між Maple та зовнішньою середою здійснюється, частіше всього, через файли текстового формату (Рис. 59).

Команда `writedata()` записує в текстовий файл дані Maple, і має такий синтаксис:

```
writedata(fileID, data, format, procedure)
```

де `fileID` – ім'я файлу даних, `data` – список, вектор або матриця даних, `format` – специфікація формату даних (`integer`, `float` або `string`), `procedure` – задає процедуру, яка використовується для перетворення даних, якщо вони не відповідають типу, який вказано в параметрі.

Команда `readdata()` читає табульовані дані з текстового файлу, і має синтаксис:

```
readdata(fileID, format, n)
```

`n` – ціле позитивне число, яке задає кількість стовпців, що читаються.

```
primer := readdata("c:\\MNV\\data.txt")
[2.27, 2.48, 3.55, 2.75, 3.09, 4.26, 2.29, 3.00, 4.11, 3.89, 0.13, 0.84, 3.21, 2.38, 1.82,
 3.09, 3.77, 3.33, 3.01, 2.23 ]
writedata("c:\\data1.txt") := primer
[2.27, 2.48, 3.55, 2.75, 3.09, 4.26, 2.29, 3.00, 4.11, 3.89, 0.13, 0.84, 3.21, 2.38, 1.82,
 3.09, 3.77, 3.33, 3.01, 2.23 ]
```

Рис. 59 Команди `writedata()` та `readdata()`

Контрольні запитання

1. Назвіть основні елементи вікна Maple.
2. За допомогою якої команди здійснюється підключення пакетів системи Maple?
3. Назвіть декілька спеціальних символів Maple та їх призначення.
4. Назвіть основні об'єкти Maple.
5. Яка команда використовується для побудови графіка функції однієї змінної?
6. Назвіть способи представлення функції в Maple.
7. Який синтаксис має команда для побудови графіка функції, яка задана в полярних координатах?
8. Для чого призначений пакет `plots`?
9. Які команди Maple використовуються для читання/запису текстових файлів?

4 ВИКОРИСТАННЯ ІНСТРУМЕНТІВ MAPLE ДЛЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ

Пакет статистичних обчислень `stats` складається з підпакетів (Рис. 60), в яких згруповані команди, що мають відношення до відповідних розділів статистики [6, 7]. Підключення пакету виконується командою `with(stats)`.

```
with(stats)
_ [anova, describe, fit, importdata, random, statevalf, statplots, transform ]
```

Рис. 60 Підключення пакету stats

4.1 Підпакети describe, transform, random, statevalf

Підпакет **describe** містить команди для обчислення різноманітних статистичних характеристик (Рис. 61).

Формат виклику:

```
stats[describe, function](arg)
describe[function](arg)
```

де *arg* – масив даних, а замість *function* може бути використано, наприклад, одне з таких ключових слів:

- *count* – кількість елементів в масиві даних;
- *linearcorrelation* – коефіцієнт лінійної кореляції;
- *mean* – середнє арифметичне;
- *mode* – мода розподілу;
- *median* – медіана розподілу;
- *standarddeviation* – стандартне відхилення;
- *variance* – дисперсія.

Фрагмент документу Maple:

```
with(stats) :
data1 := [1, 2, 2.4, 3, 4, 5, 6, 2, 7, 3.5, 2, 4.7] :
data2 := [4, 5, 7, 8, 8.4, 9.7, 7, 9, 6.8, 8.9, 7, 8] :
describe[count](data1)
12
describe[mean](data1)
3.550000000
describe[mode](data2)
7
describe[standarddeviation](data1)
1.752854814
describe[linearcorrelation](data1, data2)
0.3612356669
```

Рис. 61 Використання підпакету describe

Підпакет **transform** містить функції перетворення даних (Рис. 62).

Формат виклику:

```
stats[transform, function](arg)
transform [function](arg)
```

де *arg* – масив даних, а замість *function* може бути використано, наприклад, одне з таких ключових слів:

- *frequency* – частотність кожного елементу даних;
- *statsort* – сортування статистичних даних;
- *tally* – підрахунок повторюваності кожного елементу;
- *tallyinto* – підрахунок повторюваності кожного елементу в певному класі.

Фрагмент документу Maple:

Вага елементів ряду

```
V := transform[tally](data2)
      [4, 5, Weight(7, 3), Weight(8, 2), 9, 6.8, 8.4, 8.9, 9.7]
```

Скільки елементів потрапляє у вказані інтервали

```
K := transform[tallyinto](data2, [0..2, 2..4, 4..6, 6..8, 8..10])
      [Weight(0..2, 0), Weight(2..4, 0), Weight(4..6, 2), Weight(6..8, 4), Weight(8..10, 6)]
```

Визначення частотності однакових елементів

```
transform[frequency](K)
      [0, 0, 2, 4, 6]
```

Рис. 62 Використання підпакету transform

Підпакет **random** призначений для генерації випадкових чисел з різними законами розподілу (Рис. 63).

Формат виклику:

```
stats[random, distribution](k, uniform, method)
random[distribution](k, uniform, method)
```

де *distribution* – закон розподілу; *k* – позитивне число, яке вказує на кількість випадкових чисел; *uniform* – процедура генерації чисел з рівномірним розподілом; *method* – один з трьох методів, за допомогою якого буде проведена генерація чисел ('auto', 'inverse', 'builtin').

Підпакет містить дискретні та безперервні розподіли, наприклад:

binomiald[n, p] – біноміальний (*n* – кількість спроб, *p* – ймовірність успіху);

discreteuniform[a, b] – дискретний рівномірний (*a, b* – нижня та верхня границі параметру);

chisquare[nu] – χ^2 - розподіл (*nu* – ступінь вільності);

fratio[nu1, nu2] – розподіл Фішера (*nu1, nu2* – ступені вільності);

normald[mu, sigma] – нормальний (*mu* – математичне очікування, *sigma* – середньоквадратичне відхилення);

students[nu] – розподіл Стьюдента (*nu* – ступінь вільності);

uniform[a, b] – рівномірний (*a, b* – нижня та верхня границі параметру).

Фрагмент документу Maple:

```
x := stats[random, normald[0, 1]](5)
      0.3151206337, -0.6711662977, -0.3957955372, 0.3570600824, 0.004561604546
y := random[normald[0, 1]](5)
      0.8186971245, -0.7093245232, -0.1522233390, -0.01405914095, -0.6157994795
μ := 1
      1
σ := 2
      2
z := random[normald[μ, σ]](5)
      2.568511373, 5.909999802, 2.667725771, -0.898466859, -0.548551503
random[uniform[3, 4]](6)
      3.352263657, 3.819571011, 3.849634042, 3.704117256, 3.169276114, 3.371184668
```

Рис. 63 Генерація випадкових чисел

Підпакет `statevalf` призначений для обчислення статистичних функцій та отримання оцінок для масивів даних (Рис 64).

Формат виклику:

```
stats[statevalf, function, distribution](arg)
statevalf[function, distribution](arg)
```

де `distribution` – закон розподілу; `function` – команди для безперервного розподілу:

- `cdf` – функція ймовірності;
- `icdf` – квантиль;
- `pdf` – функція щільності ймовірності;

для дискретного розподілу:

- `dcd` – дискретна функція ймовірності;
- `icd` – квантиль;
- `pf` – функція щільності ймовірності.

Фрагмент документу Maple:

Значення функції щільності ймовірності для розподілу Пуассона в точці 5

```
statevalf[pf, poisson[3]](5)
0.1008188134
```

Значення функції щільності ймовірності нормального розподілу в точці 0.9

```
stats[statevalf, pdf, normald](0.9)
0.2660852499
```

Значення квантиля нормального розподілу в точці 0.7

```
statevalf[icdf, normald](0.7)
0.5244005127
```

Графік функції щільності ймовірності нормального розподілу

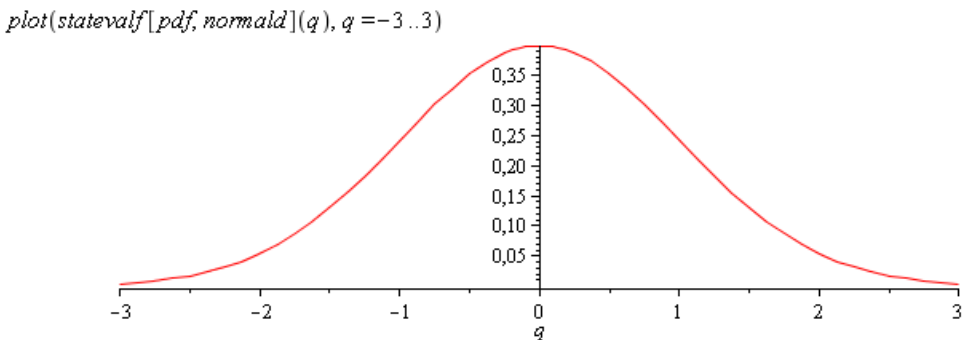


Рис. 64 Використання підпаketу `statevalf`

4.2 Підпакет візуалізації статистичних даних `statplots`

Статистичний пакет `stats` має свою невелику бібліотеку `statplots` для побудови графіків.

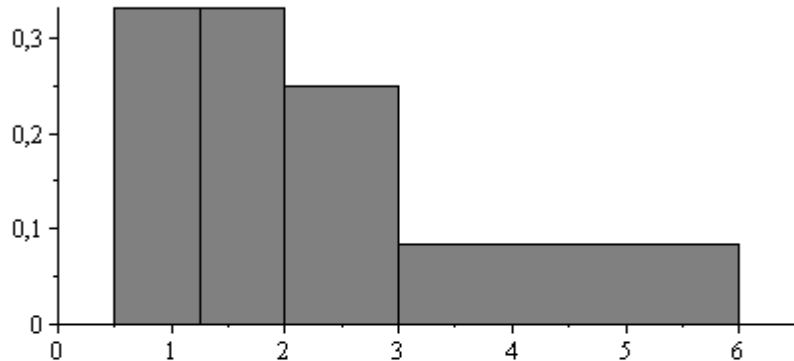
Формат виклику:

```
stats[statplots, function](arg)
statplots [function](arg)
```

де `arg` – масив даних, а вид графіку задається описом `function`: `boxplot`, `scatterplot`, `histogram` та інші (Рис 65).

Фрагмент документу Maple:

```
Xdata := readdata("c:\MN\X.txt")
        [0.5, 0.5, 0.75, 1.25, 1.5, 1.5, 1.75, 2., 2.25, 2.5, 2.5, 3., 3.25, 4., 6.]
Ydata := readdata("c:\MN\Y.txt")
        [51., 65., 70., 68., 60., 70., 63., 71., 85., 80., 84., 77., 84., 83., 96.]
stats[statplots, histogram](Xdata, color = gray)
```



```
statplots[scatterplot](Xdata, Ydata, color = blue, symbol = solidcircle)
```

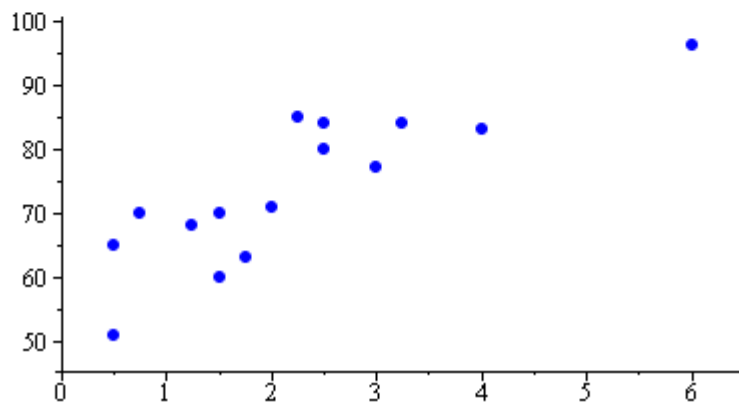


Рис. 65 Використання підпакету statplots

4.3 Елементи кореляційного та регресійного аналізу в Maple

Для обчислення коефіцієнту лінійної кореляції використовується функція `linearcorrelation` з підпакету `describe` (п. 4.1). Для побудови кореляційного поля можна використати підпакет `statplots` з пакету `stats` (п. 4.2), або пакет `plots` (п. 3.4).

Для визначення кореляційних відносин, а також апроксимації статистичних даних обраними залежностями призначений підпакет `fit` (Рис. 660).

Формат виклику:

```
stats[fit, leastsquare[vars, eqn, parms]](data)
fit[leastsquare[vars, eqn, parms]](data)
```

де `data` – список даних, `vars` – список змінних для представлення даних, `eqn` – рівняння, яке задає апроксимуючу залежність (за замовчуванням залежність – лінійна), `parms` – множина параметрів, які замінюються значеннями, що обчислені.

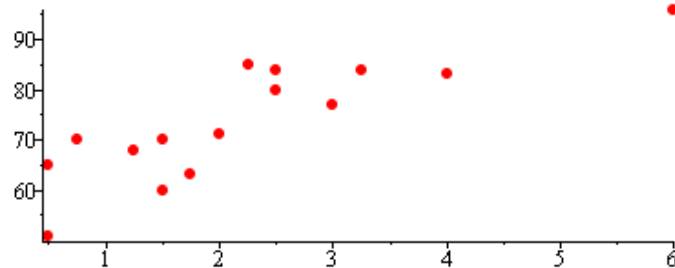
Фрагмент документу Maple:

```
with(stats) : with(plots) :
Xdata := readdata("c:\MNV\X.txt") :
Ydata := readdata("c:\MNV\Y.txt") :
```

Обчислення коефіцієнту кореляції, побудова кореляційного поля та регресійної залежності

```
rx := describe[linearcorrelation](Xdata, Ydata)
0.8446827462
```

```
pointplot(Xdata, Ydata)
```



```
y := fit[leastsquare[[x1, y1]]](Xdata, Ydata)
y1 = 58.66577540 + 6.827469744 x1
```

```
z := fit[leastsquare[[x1, y1], y1 = a · x12 + b · x1 + c, {a, b, c}]](Xdata, Ydata)
y1 = -0.6809985403 x12 + 10.91446725 x1 + 54.29664172
```

```
display({plot(-0.6809985403 x12 + 10.91446725 x1 + 54.29664172, x1 = -2..7, color = red, thickness = 2),
plot(58.66577540 + 6.827469744 x1, x1 = -2..7, color = blue, thickness = 2), pointplot(Xdata, Ydata, symbol
= solidcircle, color = black)})
```

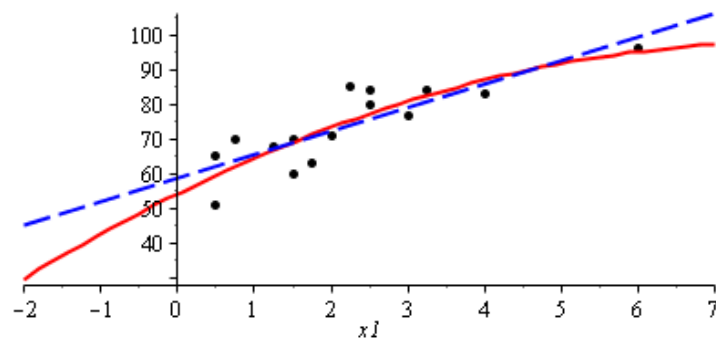


Рис. 66 Використання підпакету fit

Контрольні запитання

1. Як підключити пакет для статистичних обчислень?
2. Які підпакекти містить пакет stats?
3. Яке призначення підпакету fit?
4. Назвіть декілька прикладів дискретних розподілів пакету stats.
5. Як згенерувати ряд випадкових чисел з певним розподілом засобами Maple?
6. Який підпакет призначений для візуалізації статистичних даних?
7. Який підпакет містить функції перетворення даних?

5 ЛАБОРАТОРНІ РОБОТИ

Лабораторна робота №1. Числові характеристики вибірки. Побудова гістограм і полігонів в MathCAD

Завдання:

1. Змоделювати вибірку 250 вибірових значень з нормальним розподілом з математичним очікуванням і середньоквадратичним відхиленням.
2. Обчислити статистичні характеристики для вибіркової сукупності.
3. Виконати угруповання для значень $k=10, 20$. Побудувати відповідні гістограми, полігони частот.

Запитання для самоперевірки:

1. Назвіть основні статистичні характеристики ряду даних.
2. За допомогою яких функцій MathCAD можна обчислити моду та коваріацію?
3. За допомогою якої функції MathCAD можна обчислити незміщену оцінку дисперсії?
4. Назвіть функції для побудови гістограм.
5. Назвіть функції для опису нормального розподілу ймовірностей.

Теоретичні відомості: див. 1.2, 1.4, 2.1, 2.2.

Література: основна [2, 3, 6, 8]; додаткова [9, 11, 14].

Лабораторна робота №2. Побудова графіків щільності ймовірності і функції розподілу. Обчислення довірчого інтервалу

Завдання:

1. Побудувати графіки щільності ймовірності нормальних розподілів, якщо математичне сподівання $\mu=0, 1, 3$ і середньоквадратичне відхилення $\sigma=1, 2, 3$. Використати функцію $dnorm(x, \mu, \sigma)$.
2. Побудувати графіки нормальних функцій розподілу, якщо математичне сподівання $\mu=0, 1, 3$ і середньоквадратичне відхилення $\sigma=1, 2, 3$. Використати функцію $pnorm(x, \mu, \sigma)$.
3. Знайти 90%-вий довірчий інтервал для дисперсії при невідомому математичному сподіванні.

Запитання для самоперевірки:

1. Назвіть вбудовані функції MathCAD для опису нормального розподілу.
2. Визначення функції розподілу, щільності ймовірності, квантилі.
3. Записати 90%-у квантиль нормального розподілу за допомогою відповідної функції MathCAD.
4. Назвіть параметри нормального та експоненціального розподілів.
5. Який інтервал називається довірчим?

Теоретичні відомості: див. 1.2, 1.4, 2.3, 2.4.

Література: основна [2, 3, 6]; додаткова [9, 11].

Лабораторна робота №3. Числові характеристики вибірки. Побудова графіків в Maple

Завдання:

1. Змодельовати вибірку 250 вибіркових значень (нормальний закон розподілу), якщо математичне сподівання $\mu = 150$ та дисперсія $\sigma = 10$.
2. Обчислити вибіркові характеристики для вибірки, що сформована. Побудувати гістограму.
3. Відобразити на площині сукупність з 1000 випадкових точок (з нормальним розподілом) в обмеженому прямокутному просторі. Колір точок – синій.

Запитання для самоперевірки:

1. Який формат виклику має підпакек `random`?
2. Як визначити частотність однакових елементів?
3. Як визначити значення квантиля нормального розподілу в точці 0.9?
4. Який формат виклику має підпакек `statplots`?
5. Як при візуалізації статистичних даних в Maple можна задати вид графіку – гістограма?

Теоретичні відомості: див. 3.2., 3.4, 4.1, 4.2.

Література: основна [1, 3, 5, 7]; додаткова [9, 13].

Лабораторна робота №4. Перевірка статистичних гіпотез в MathCAD

Завдання:

1. Скопіювати в особисту папку файл **lab4.txt** з папки «Вихідні дані до лабораторних робіт».
2. Прочитати з файлу вибірку об'єму 50.
3. Сформулювати нульову гіпотезу про величину математичного очікування (при невідомій дисперсії) і перевірити для заданих рівнів значущості три альтернативні гіпотези.

Запитання для самоперевірки:

1. Назвіть основний принцип статистичної перевірки гіпотез.
2. Визначення критичної області та області прийняття гіпотези.
3. Сформулювати першу альтернативну гіпотезу.
4. Який критерій використовується при перевірці гіпотези про математичне очікування з невідомою дисперсією?
5. Що можна обчислити за допомогою функції MathCAD `qt(x, d)`?

Теоретичні відомості: див. 1.2, 2.7.

Література: основна [1, 2, 3; 6]; додаткова [9, 11].

Лабораторна робота №5. Кореляційний та регресійний аналіз в MathCAD

Завдання:

1. Скопіювати в особисту папку файли *lab5_x.txt* та *lab5_y.txt* з папки «Вихідні дані до лабораторних робіт».
2. Прочитати з файлу вибірку заданого об'єму.
3. Побудувати кореляційне поле, обчислити коефіцієнт кореляції, надати інтерпретацію.
4. Обчислити коефіцієнти лінійної регресії.
5. Побудувати графік лінійної регресії.

Запитання для самоперевірки:

1. Поняття кореляційної залежності.
2. Які функції MathCAD використовуються для обчислення коефіцієнту кореляції?
3. Як в MathCAD побудувати кореляційне поле ?
4. Які властивості має коефіцієнт кореляції?
5. Які функції MathCAD використовуються для обчислення коефіцієнтів регресії?

Теоретичні відомості: див. 1.2, 1.4, 2.6.

Література: основна [3, 8, 6]; додаткова [9, 11].

Лабораторна робота №6. Кореляційний та регресійний аналіз в Maple

Завдання:

1. Скопіювати в особисту папку файли *lab6_x.txt* та *lab6_y.txt*.
2. Прочитати з файлу вибірку заданого об'єму.
3. Побудувати кореляційне поле, використовуючи функції **scatterplot** та **pointplot**.
4. Обчислити коефіцієнт кореляції, надати інтерпретацію.
5. Отримати рівняння лінійної та поліноміальної регресії.
6. Побудувати комбінований графік: точковий, лінійної регресії, поліноміальної регресії.

Запитання для самоперевірки:

1. Який формат виклику має підпакект `fit`?
2. Який вигляд має графік, якщо використовується опис `scatterplot`?
3. До якого пакету відноситься функція `pointplot`?
4. Назвіть функцію Maple для обчислення коефіцієнта кореляції.
5. Назвіть способи побудови комбінованого графіку.

Теоретичні відомості: див. 3.2, 3.3, 3.4, 3.5, 4.3,4.4.

Література: основна [1, 3, 7]; додаткова [9, 13].

Лабораторна робота №7. Однофакторний дисперсійний аналіз в MathCAD

Завдання:

Встановити вплив типу реклами, що використовується, на об'єм продажу товару для заданого значення α . Визначити параметри і побудувати графіки щільності вірогідності розподілу числа продажу для кожного типу реклами.

Тип реклами	Рік					
	2014	2015	2016	2017	2018	2019
A	2.391	3.001	4.473	3.228	3.387	3.276
B	4.324	4.106	2.56	2.82	2.82	4.127
C	2.336	1.92	1.849	2.035	3.002	
D	2.28	2.581	3.636	1.865		

Запитання для самоперевірки:

1. Яка основна ідея однофакторного дисперсійного аналізу?
2. Сформулювати нульову гіпотезу однофакторного дисперсійного аналізу.
3. Яким чином обчислити оцінки середнього та дисперсії?
4. Назвіть функцію Maple для обчислення квантилю розподілу Фішера.
5. У якому випадку можна зробити висновок, що вибір рівню фактора не впливає на ознаку?

Теоретичні відомості: див. 1.2, 1.3, 1.4, 2.7.

Література: основна [1, 3, 6]; додаткова [9, 11, 12].

Лабораторна робота №8. Розв'язання задач методом Монте-Карло в MathCAD

Завдання:

1. Змодельовати фізичний експеримент: сигнал і шум з рівномірним законом розподілу. Побудувати відповідні графіки.
2. Обчислити число π методом Монте-Карло (задати різну кількість спроб, проаналізувати результат).

Запитання для самоперевірки:

1. Для розв'язання яких задач використовуються методи Монте-Карло?
2. Поняття рівномірного розподілу.
3. Назвіть функцію MathCAD для генерації одного псевдовипадкового числа.
4. Алгоритм обчислення числа π методом Монте-Карло?

Теоретичні відомості: див. 1.2, 1.3, 1.4, 2.8.

Література: основна [1, 3, 6]; додаткова [9, 11].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Алексеев Е. Р., Чеснокова О. В. Решение задач вычислительной математики в пакетах Mathcad 12, MATLAB 7, Maple 9. Москва : НТ Пресс, 2006. 496с.
2. Гмурман В. Е. Теория вероятностей и математическая статистика : учебное пособие для вузов. 9-е изд., стереотип. Москва : Высшая школа, 2003. 479 с.
3. Дьяконов В. П. Maple 9.5/10 в математике, физике и образовании. Москва : СОЛОН-Пресс, 2006. 720 с.
4. Кирьянов Д. В. Самоучитель Mathcad 11. Санкт-Петербург: БХВ-Петербург, 2003. 560 с.
5. Макаров Е. Г. Mathcad : учебный курс. Санкт-Петербург : Питер, 2009. 384 с.
6. Матросов А. В. Maple 6. Решение задач высшей математики и механики. Санкт-Петербург : БХВ-Петербург, 2001. 528 с.
7. Сдвижков О. А. Математика на компьютере : Maple 8. Москва : СОЛОН-Пресс, 2003. 175 с.
8. Херхагер М., Партолль Х. Mathcad 2000 : Полное руководство. Киев : Ирина: ВНУ, 2000. 416с.
9. Швець В.Т. Теорія ймовірностей і математична статистика : Видавництво ВМВ, 2018. 218 с.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна:

1. Алексеев Е. Р., Чеснокова О. В. Решение задач вычислительной математики в пакетах Mathcad 12, MATLAB 7, Maple 9. Москва : НТ Пресс, 2006. 496с.
2. Волощенко Ф. Б., Джалладова І. А. Теорія ймовірностей та математична статистика: навчально-методичний посібник для самостійного вивчення дисципліни. Київ : КНЕУ, 2003. 256 с.
3. Гмурман В. Е. Теория вероятностей и математическая статистика : учебное пособие для вузов. 9-е изд., стереотип. Москва : Высшая школа, 2003. 479 с.
4. Махней О. В., Гой Т. П. Математичне забезпечення автоматизації прикладних досліджень : навч. посіб. Івано-Франківськ : Сімик, 2013. 304 с.
5. Дьяконов В. П. Maple 9.5/10 в математике, физике и образовании. Москва : СОЛОН-Пресс, 2006. 720 с.
6. Макаров Е. Г. Mathcad : учебный курс. Санкт-Петербург : Питер, 2009. 384 с.
7. Матросов А. В. Maple 6. Решение задач высшей математики и механики. Санкт-Петербург : БХВ-Петербург, 2001. 528 с.
8. Тюрин О. В., Ахмеров О. Ю. Теорія ймовірностей і математична статистика : навч. посіб. Одеса : Одес. нац. ун-т ім. І. І.Мечникова, 2018. 170 с.

Додаткова:

9. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. Санкт-Петербург : Изд-во "Лань", 1999. 224 с.
10. Боровиков В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. Санкт-Петербург : Питер, 2003. 688 с.
11. Кирьянов Д. В. Самоучитель Mathcad 11. Санкт-Петербург: БХВ-Петербург,

2003. 560 с.

12. Костів Б. В. Сучасні засоби обчислювальної практики та програмування : лабораторний практикум. Івано-Франківськ : ІФНТУНГ, 2010. 40 с.

13. Сдвижков О. А. Математика на комп'ютері : Maple 8. Москва : СОЛОН-Пресс, 2003. 175 с.

14. Херхагер М., Партолль Х. Mathcad 2000 : Полное руководство. Киев : Ирина: ВНУ, 2000. 416с.

ЗМІСТ

ВСТУП.....	3
1 ОСНОВНІ ІНСТРУМЕНТИ ТА ЗАСОБИ РОБОТИ З МАТЕМАТИЧНИМ ПАКЕТОМ MATHCAD	4
1.1 Інтерфейс користувача математичного пакету MathCAD	4
1.2 Введення та виведення інформації.....	7
1.2.1 Створення формул.....	7
1.2.2 Вставка функцій	7
1.2.3 Створення тексту.....	8
1.2.4 Введення/виведення даних в текстові файли	8
1.3 Типи даних. Обчислення в MathCAD	9
1.3.1 Змінні	9
1.3.2 Символьні обчислення	11
1.4 Побудова та форматування графіків	11
Контрольні запитання	13
2 ВИКОРИСТАННЯ ІНСТРУМЕНТІВ MATHCAD ДЛЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ.....	13
2.1 Вибіркові статистичні характеристики	13
2.2 Побудова гістограми та полігона	14
2.3 Функції розподілу ймовірностей.....	16
2.4 Інтервальне оцінювання. Обчислення довірчого інтервалу для дисперсії	18
2.5 Перевірка статистичних гіпотез	20
2.6 Елементи кореляційного та регресійного аналізу	24
2.7 Однофакторний дисперсійний аналіз.....	26
2.8 Реалізація методу Монте-Карло в системі MathCAD	28
Контрольні запитання	30
3 ОСНОВНІ ІНСТРУМЕНТИ ТА ЗАСОБИ РОБОТИ З МАТЕМАТИЧНИМ ПАКЕТОМ MAPLE	30
3.1 Інтерфейс користувача математичного пакету Maple	30
3.2 Спеціальні символи та основні об'єкти Maple.....	32
3.3 Функції	33
3.4 Побудова графіків функцій в системі Maple.....	35
3.5 Робота з файлами в системі Maple	37
Контрольні запитання	37
4 ВИКОРИСТАННЯ ІНСТРУМЕНТІВ MAPLE ДЛЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ.....	37
4.1 Підпакети <i>describe, transform, random, statevalf</i>	38
4.2 Підпакет візуалізації статистичних даних <i>statplots</i>	40
4.3 Елементи кореляційного та регресійного аналізу в Maple	41
Контрольні запитання	42
5 ЛАБОРАТОРНІ РОБОТИ	43
Лабораторна робота №1. Числові характеристики вибірки. Побудова гістограм і полігонів в MathCAD	43
Лабораторна робота №2. Побудова графіків щільності ймовірності і функції розподілу. Обчислення довірчого інтервалу	43
Лабораторна робота №3. Числові характеристики вибірки. Побудова графіків в Maple	44
Лабораторна робота №4. Перевірка статистичних гіпотез в MathCAD	44
Лабораторна робота №5. Кореляційний та регресійний аналіз в MathCAD.....	45
Лабораторна робота №6. Кореляційний та регресійний аналіз в Maple	45

Лабораторна робота №7. Однофакторний дисперсійний аналіз в MathCAD	46
Лабораторна робота №8. Розв'язання задач методом Монте-Карло в MathCAD.....	46
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	47
РЕКОМЕНДОВАНА ЛІТЕРАТУРА	47

Навчальне видання
(українською мовою)

Надія Вікторівна Матвіїшина
Наталія Олександрівна Кондрат'єва

ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ СТАТИСТИЧНИХ ДОСЛІДЖЕНЬ

Методичні рекомендації до лабораторних занять для здобувачів ступеня вищої освіти бакалавра спеціальності «Комп'ютерні науки» освітньо-професійної програми «Комп'ютерні науки»

Рецензент *С.П.Швидка*
Відповідальний за випуск *С. Ю. Борю*
Коректор *Н.В.Матвіїшина*

Підп. до друку 00.00.2015. Формат 60×90/16. Папір офсетний.
Друк цифровий. Гарнітура Таймс. Умовн. друк. арк. 7,5.
Замовлення № 000. Тираж 20 прим.

Запорізький національний університет

69600, м. Запоріжжя, МСП-41
вул. Жуковського, 66

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців, виготівників
і розповсюджувачів видавничої продукції
ДК № 2952 від 30.08.2007