Лабораторна робота № 5

**ВИКОРИСТАННЯ СТАТИСТИЧНИХ ФУНКЦІЙ**

**програми EXCEL**

**Мета роботи:**

* освоєння принципів роботи зі статистичними функціями програми **Microsoft Excel**;
* • вивчення принципів роботи електронних таблиць **Microsoft Excel**зекспериментальними даними;
* • освоєння методики прийняття рішень на основі отриманих статистичних показників.

**Короткі теоретичні відомості**

Програма **Excel**має багатий набор статистичних функцій, які використовуються для швидкої і елегантною оцінки різного роду імовірнісних та статистичних параметрів. Для їх застосування треба розбиратися в математичній статистиці.

Пакет аналізу є надбудовою**Excel**. Нижче указані основні види (показники) аналізу з короткою характеристикою, підтримувані пакетом **Excel**.

• Розрахунок набору статистичних показників (мода, медіана, дисперсія і т.д.) одновимірного набору даних.

• t-тест (t-тестСтьюдента) з однаковими і різними дисперсіями для перевірки гіпотези про рівність (розходженні) середніх двох вибірок.

• Однофакторний і двофакторний дисперсійний аналіз двох або більше вибірок, що належать одній і тій же генеральної сукупності.

• F-тест для порівняння дисперсій двох генеральних сукупностей.

• Кореляція для кількісної оцінки взаємозв'язку двох наборів даних, представлених у безрозмірному вигляді.

• Коваріація для обчислення середнього твору відхилень точок даних від відносних середніх.

• Експоненційне згладжування для передбачення значення на основі прогнозу для попереднього періоду, скоригованого з урахуванням похибок у цьому прогнозі.

• Аналіз Фур'є для вирішення завдань в лінійних системах та аналізу періодичних даних з використанням методу швидкого перетворення Фур'є.

• Гістограма для обчислення вибіркових та інтегральних частот попадання даних у вказані інтервали значень, при цьому генеруються числа влучень для заданого діапазону комірок.

• Ковзаюче середнє для розрахунку значень у прогнозованому періоді на основі середнього значення змінної для вказаного числа попередніх періодів.

• Генерація випадкових чисел, витягнутих з одного або декількох розподілів, для моделювання об'єктів, що мають випадкову природу.

• Регресія для аналізу впливу на окрему залежну змінну значень декількох незалежних змінних.

Усі вони мають дуже простий інтерфейс,що полегшує їх застосування. Після виклику функції необхідно вказати комірки з даними щоб отримати відповідь. В деяких випадках необхідно додатково вказати допоміжні параметри (показники), що характеризують функцію чи її представлення.

**Виконання роботи**

* 1. Ввести укомірки таблиці**Excel**два набори вихідних даних з таблиці 4.1 відповідно до варіанта завдання, зазначеному у таблиці 4.2.

Таблиця 4.1

|  |
| --- |
| **Номер наборів вихідних даних** |
| **№ 1** | **№ 2** | **№ 3** | **№ 4** | **№ 5** | **№ 6** | **№ 7** | **№ 8** |
| 96 | 276 | 366 | 322 | 239 | 216 | 256 | 229 |
| 181 | 265 | 315 | 177 | 325 | 153 | 123 | 506 |
| 352 | 419 | 439 | 309 | 224 | 312 | 312 | 271 |
| 296 | 378 | 548 | 138 | 441 | 283 | 328 | 362 |
| 322 | 353 | 353 | 282 | 219 | 185 | 185 | 130 |
| 130 | 315 | 315 | 289 | 352 | 419 | 419 | 285 |
| 491 | 456 | 115 | 273 | 321 | 418 | 418 | 325 |
| 121 | 366 | 466 | 257 | 617 | 296 | 296 | 61 |
| 258 | 314 | 314 | 124 | 183 | 332 | 332 | 95 |
| 423 | 279 | 279 | 267 | 294 | 317 | 317 | 257 |
| 236 | 388 | 388 | 256 | 298 | 281 | 281 | 160 |
| 287 | 398 | 390 | 184 | 361 | 372 | 372 | 269 |
| 276 | 393 | 393 | 187 | 350 | 241 | 241 | 281 |
| 535 | 172 | 448 | 170 | 49 | 501 | 518 | 271 |
| 288 | 340 | 340 | 63 | 359 | 286 | 286 | 329 |
| 337 | 383 | 383 | 176 | 235 | 311 | 311 | 266 |
| 394 | 256 | 356 | 481 | 338 | 406 | 649 | 319 |
| 287 | 392 | 392 | 231 | 299 | 535 | 305 | 99 |
| 235 | 529 | 553 | 361 | 298 | 290 | 290 | 198 |
| 229 | 339 | 339 | 284 | 298 | 216 | 216 | 189 |

Таблиця 4.2

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| варіант | Номера наборів | варіант | Номера наборів | варіант | Номера наборів | варіант | Номера наборів | варіант | Номера наборів |
| 1 | 1,2 | 6 | 1,7 | 11 | 2,6 | 16 | 3,6 | 21 | 4,7 |
| 2 | 1,3 | 7 | 1,8 | 12 | 2,7 | 17 | 3,7 | 22 | 4,8 |
| 3 | 1,4 | 8 | 2,3 | 13 | 2,8 | 18 | 3,8 | 23 | 5,6 |
| 4 | 1,5 | 9 | 2,4 | 14 | 3,4 | 19 | 4,5 | 24 | 5,7 |
| 5 | 1,6 | 10 | 2,5 | 15 | 3,5 | 20 | 4,6 | 25 | 5,8 |

2. Використовуючи статистичні функції програми розрахувати: середнє значення, дисперсію і стандартне відхилення для кожного набору.

3. Визначити відносне відхилення мінімального і максимального значень вихідних даних в кожному наборі за формулою:

, (4.1)

де Хmin / max - мінімальне або максимальне значення в оброблюваному наборі,

S - експериментальне значення середньоквадратичного відхилення.

4. Визначити можливість виключення розглянутого показання шляхом порівняння отриманої величини зі значеннями t-критерію одностороннього розподілу Стьюдента. Аналізований показник може бути виключений з подальшої обробки, якщо ймовірність помилкової оцінки менше 0,025. Для решти даних провести розрахунки за п. 3.

5. Використовуючи статистичні функції програми визначити медіану, асиметрію (скіс) і ексцес залишилися даних. Порівняти отримані значення з допустимими:



Якщо вони в 2 ... 3 рази перевищують розрахункові, то припущення про нормальність розподілу необхідно піддати сумніву. Зробити на основі отриманих результатів висновки.

6. Визначити значимість відмінностей розбіжностей середніх значень двох решти наборів даних. Для чого:

* оцінити можливу дисперсію узагальненого набору даних за формулою:

; (4.2)

* розрахуватикоефициентСтьюдента:

, (4.3)

* порівняти отриману величину з табличній при обраному рівні значущості і числі ступенів свободи ().

Якщо отримані значення перевершують табличні значення t-критерію двостороннього розподілу Стьюдента (ймовірність помилки не більше 0,025), то розбіжності середніх величин двох розподілів можна визнати значущими.

7. Порівняти дисперсії двох наборів і перевірити значимість їх розбіжностей по F-розподілу.

8. Якщо розбіжності по п. 6 і 7 виявляться незначними, то звести набори в один, визначити середнє значення, дисперсію, а також довірчий інтервал знаходження істинного значення середньої величини за формулою:

, (4.4)

*Хсробъед* – середнє значення об'єднаного набору;

*k* – число ступенів свободи, рівне числу величин, включених в об'єднаний набір.

Якщо набори (при зазначеному рівні ймовірності помилки) неприпустимо об'єднувати, то за формулою (4.4), обліком відповідної зміни входять до неї величин *Хср*, *S*та*k*, підрахувати довірчий інтервал знаходження дійсної величини середнього значення для кожного окремого набору.

9. Розбити інтервал зміни величини *Х* на піддіапазони. Визначити кількість даних потрапили в поддиапазон і середнє значення *Х*ср, що характеризують поддиапазон. Побудувати графік розподілу числа вимірів в кожному піддіапазоні від середньої величини *Х*ср.