Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

Звіт з практичної роботи №1

Дисципліна

Графові ймовірнісні моделі

Варіант 8

Студентка: Пороскун О. О.

Група: ПМ.м-21

Викладач: Хоменко О. В.

Суми, Сумська область

2023

**Порядок виконання роботи**

1. *Згенеруємо вихідні дані.*

Вихідні дані є вибіркою {x1, x2, ..., xn}, згенерованою за одним із законів розподілу в залежності від варіанту завдання (Додаток А, Табл. 1.2). Для генерації вихідних даних використовується функція [Генерація випадкових чисел] статистичної надбудови Microsoft Excel.

Згенеровані випадкові числа необхідно округлити до цілих і розмістити в другій стовпчик таблиці. Для цього можна скористатися математичної функцією ОКРУГЛ, що має 2 аргументи: число, що округлюється, і значення десяткового розряду, до якого його потрібно округлити. Число розрядів дорівнює 0 в разі округлення до цілого.

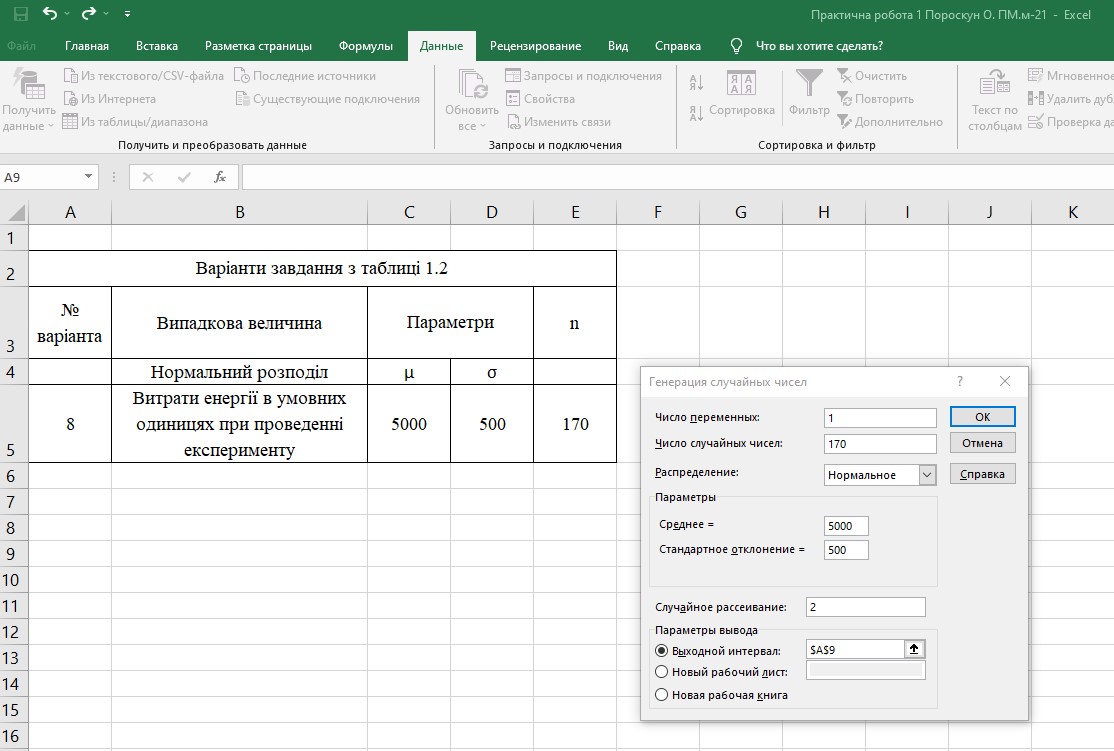


Рис. 1.1. Вибір параметрів генератора нормального розподілу

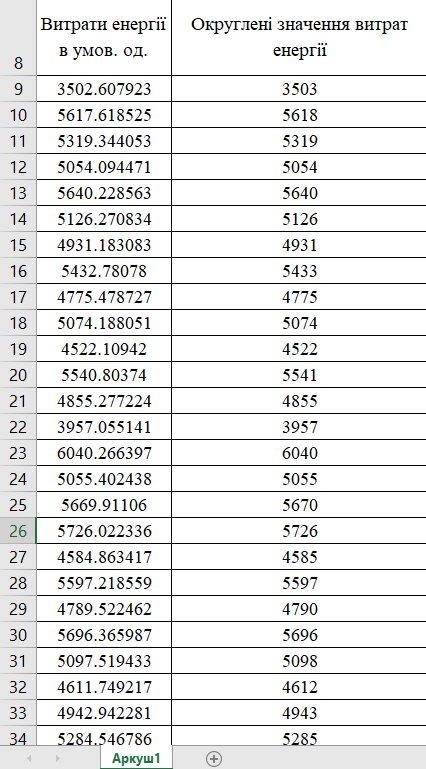
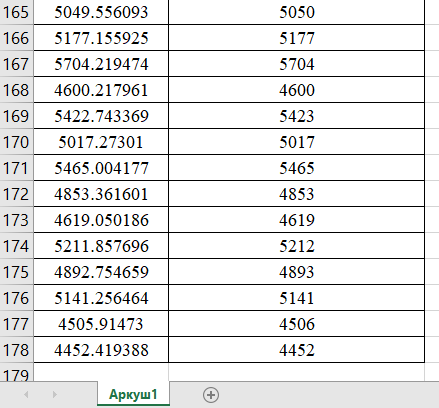
 

Рис. 1.2. Результати генератора розподілу та його округлені значення

1. *Обчислимо основні статистичні показники трьома способами:  
   - За допомогою формул;  
   - За допомогою статистичних функцій;  
   - За допомогою статистичної надбудови.*

Скористаємося спочатку статистичною надбудовою. Для обчислення показників варіації застосовується функція Описова статистика статистичної надбудови Microsoft Excel. У діалоговому вікні потрібно вибрати*Вхідний інтервал, Мітки в першому рядку, Вихідний інтервал і Підсумкова статистика.*

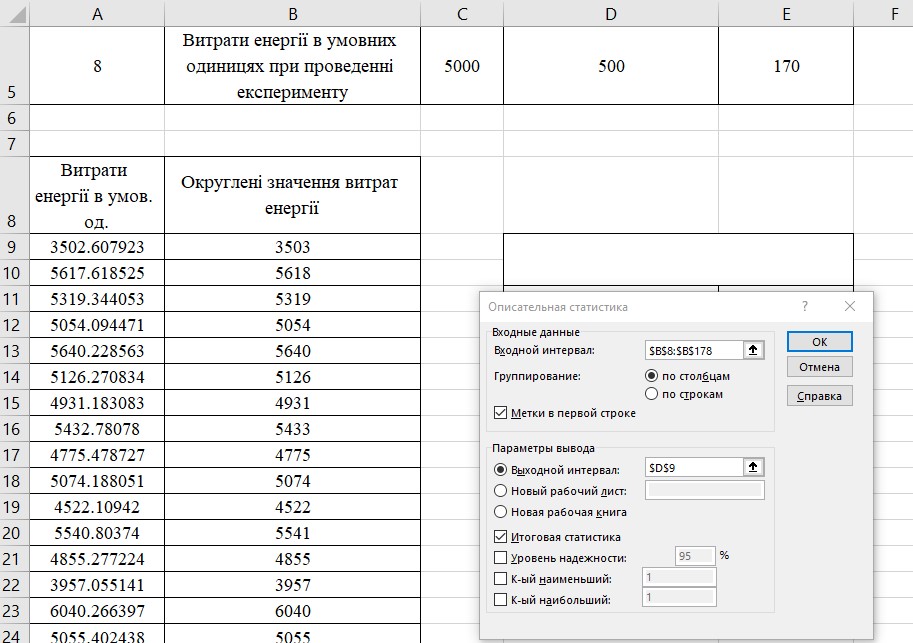


Рис. 1.3. Вибір параметрів описової статистики

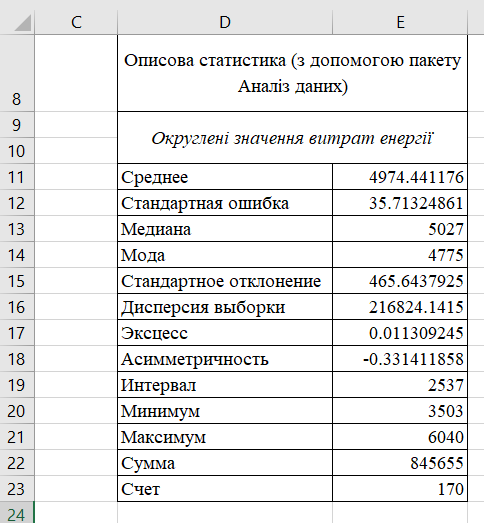
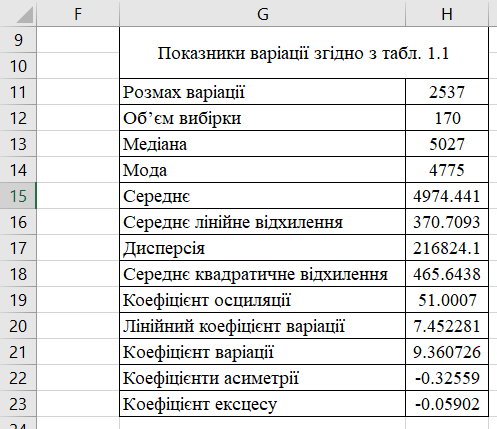


Рис. 1.4. Результати описової статистики

Далі розрахуємо показники варіації, вказані в Табл 1.1 (Додаток А) за допомогою вбудованих в Microsoft Excel формул.



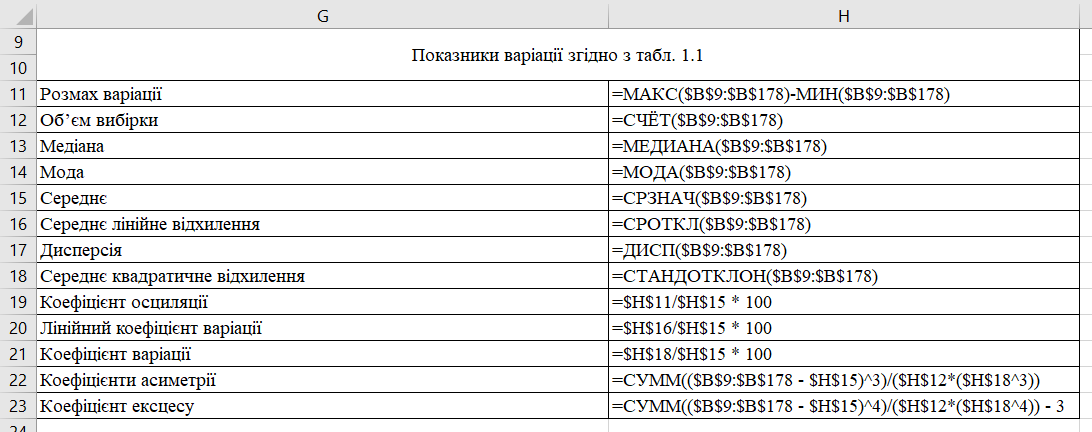


Рис. 1.5. Результати показників варіації за допомогою вбудованих формул

Далі розрахуємо деякі показники варіації без застосування вбудованих формул, використаємо формули з Рис. 1.6.

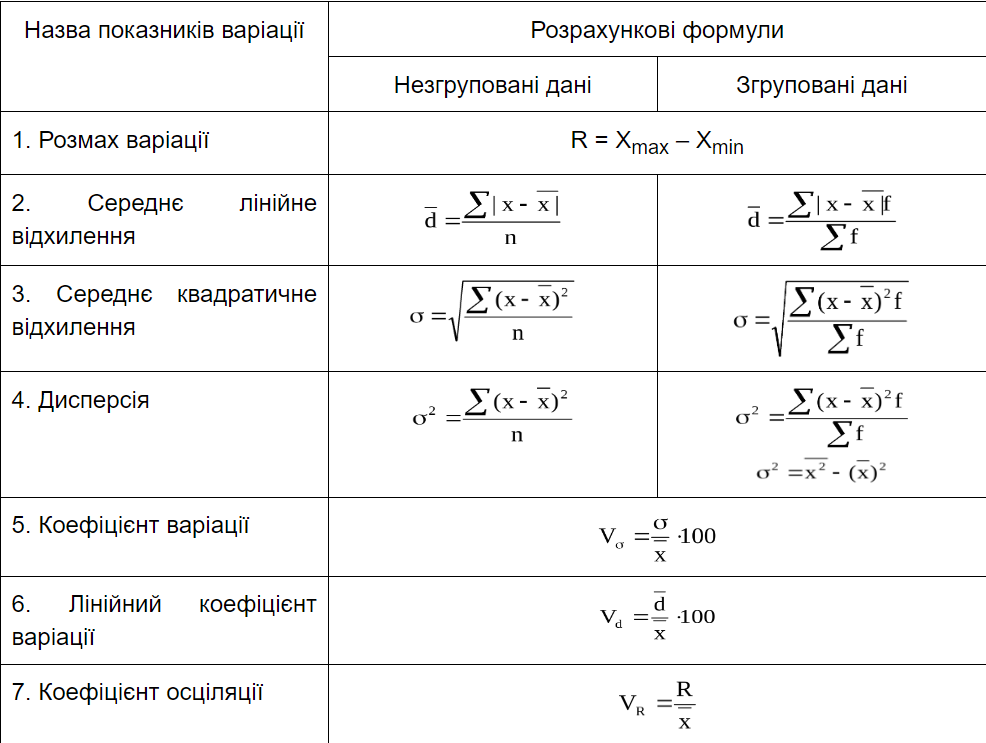
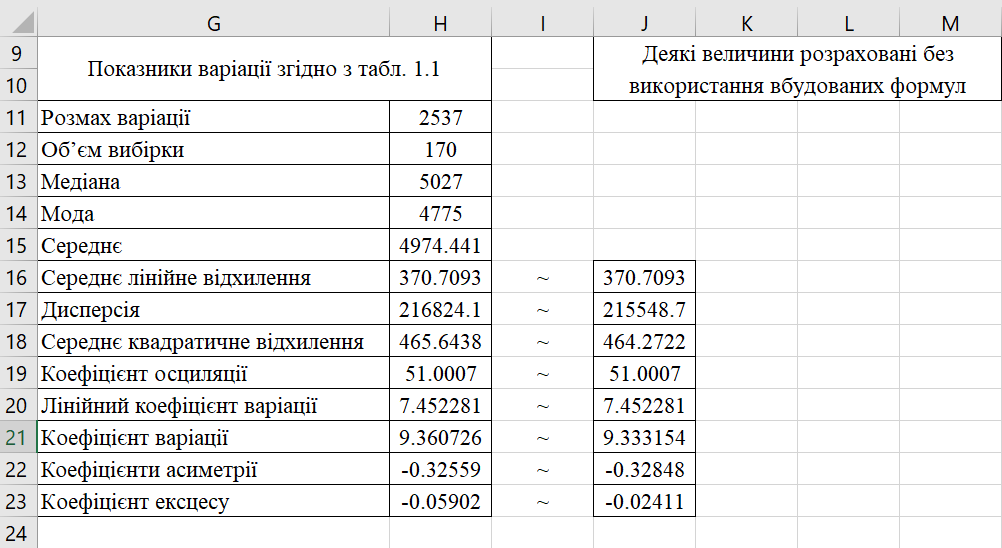


Рис. 1.6. Показники варіації та формули для їх обчислення



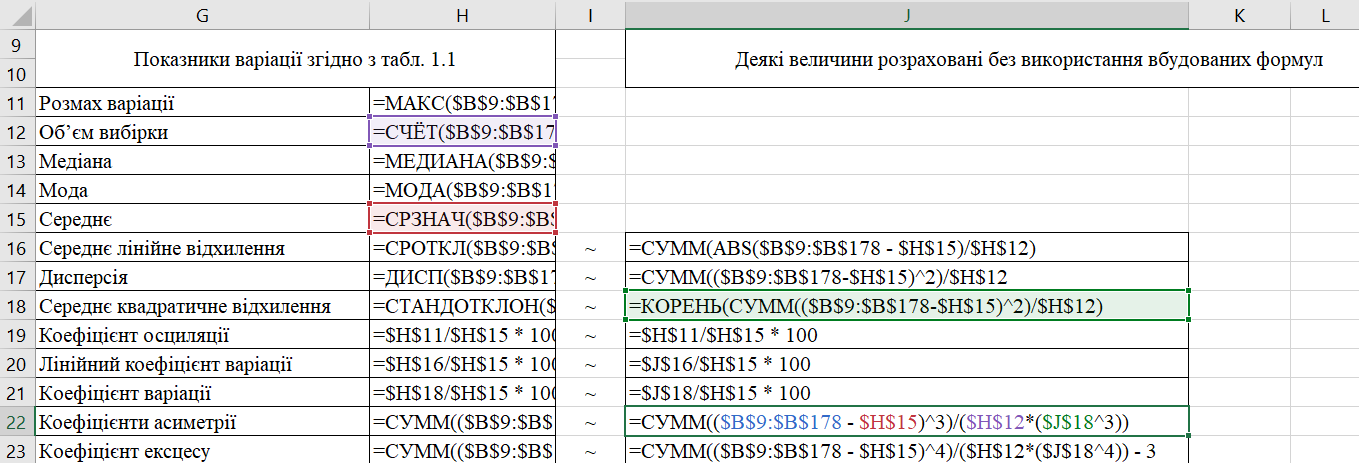


Рис. 1.7. Результати показників варіації без допомоги вбудованих формул (комірки J16 : J23)

1. *Порівняймо результати розрахунків.*

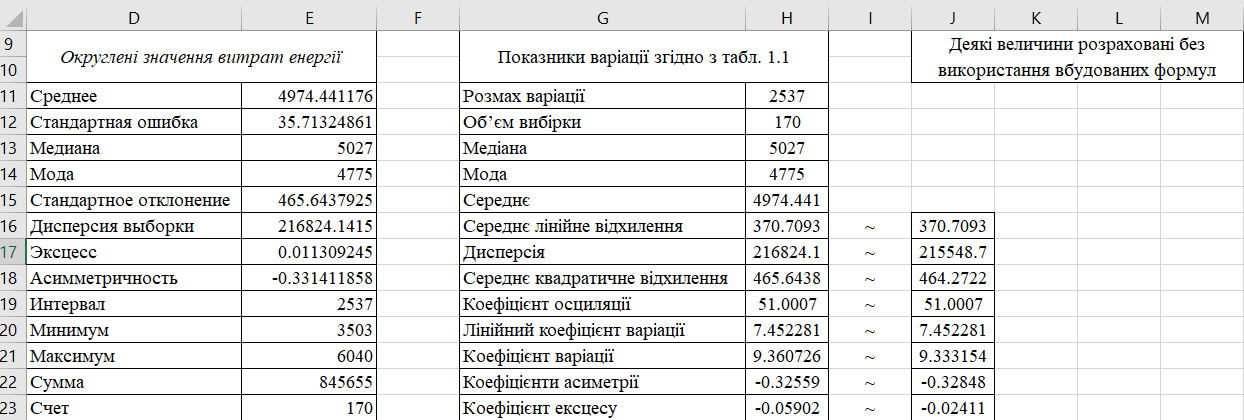


Рис. 1.8. Результати показників варіації для 3 способів

Бачимо що всі значення або однакові або наближені, крім значення коефіцієнту ексцесу.

1. *Зробіть висновок про однорідність вибірки.*
2. *Зробіть висновок про близькість до нормального розподілу.*

При аналізі показників варіації можна використовувати такі правила:

- Вибірка вважається однорідною, якщо коефіцієнт варіації Vσ ≤ 30%;

- Якщо коефіцієнти асиметрії та ексцесу близькі до нуля, то форму розподілу можна вважати близькою до нормальної. Критичні значення А і Е обчислюють за формулами:

де n = 170.

Критерій згоди:

Отже, так як коефіцієнт варіації Vσ 9,3% (≤ 30%), то вибірка вважається однорідною. А коефіцієнти асиметрії та ексцесу близькі до нуля, то форму розподілу можна вважати близькою до нормальної.

1. *Проведіть групування даних двома способами:*

*- За допомогою стандартних функцій Excel;*

*- За допомогою статистичної надбудови.*

1. *Побудуйте зведені таблиці і порівняйте результати групування, отримані двома способами.*

*Групування за допомогою статистичної надбудови*

Спочатку створимо таблицю нижніх меж інтервалів групування.

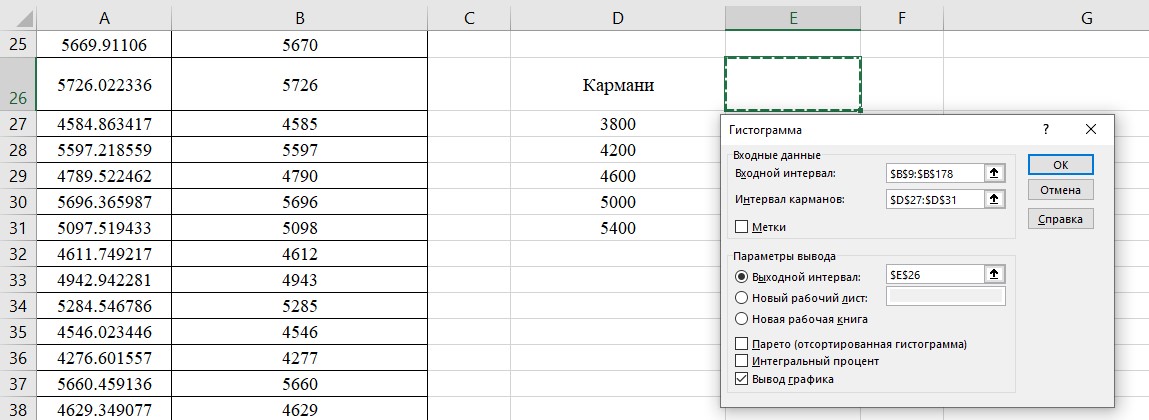


Рис.1.9 Встановлення функції Гістограма

Для групування даних за допомогою статистичної надбудови вибираємо меню *[Сервіс → Аналіз даних → Гістограма].* Вказуємо наступні параметри:

***Вхідні дані***:

-Вхідний інтервал – вибірка вихідних даних;

-Інтервал кишень – нижні межі інтервалів групування.

***Параметри виходу:***

-Вхідний інтервал – розташування результатів групування на аркуші;

-Виведення графіка – побудова гістограми.

-Інтегральний відсоток – обчислення накопичених частот.

Результат роботи функції Гістограма представлений на Рис. 1.10.

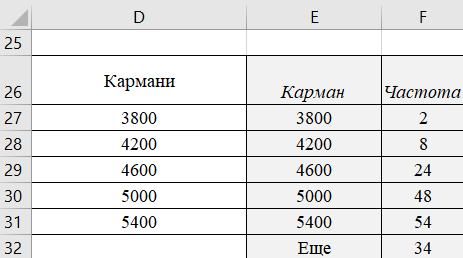


Рис. 1.10 Результат виклику функції «Гістограма»

Згенеровану таблицю необхідно доповнити відсутніми стовпчиками. Графік необхідно настроїти для коректного відображення.

Таблицю необхідно доповнити наступними стовпцями:

* Інтервал - підписи для стовпців гістограми у вигляді текстових міток, що описують межі інтервалів групування;
* Частота ;
* Відносна частота ;
* Накопичена частота

У графі **Всього** виконують підрахунок суми частот.

Для обчислення відсутніх характеристик інтервалів слід використовувати формули. Наприклад, для обчислення накопиченої частості для інтервалу 3800-4200 (комірка J28) використовується формула =$J27+$I28.

Обчислення відносної частоти для інтервалу <3800 (комірка I27) виконується за допомогою формули: =100\*$H27/$H$33.

Після обчислень слід переконатися у відсутності грубих помилок. Наприклад, накопичена частота повинна дорівнювати 100%.

Розташування стовпчиків гістограми має відповідати межам інтервалів групування даних. Для настройки графіка клацніть по стовпчику гістограми курсором і натисніть праву кнопку миші. Виберіть [*Формат рядів даних → Параметри*] і встановіть нульове значення параметрів Перекриття та Ширина зазору.

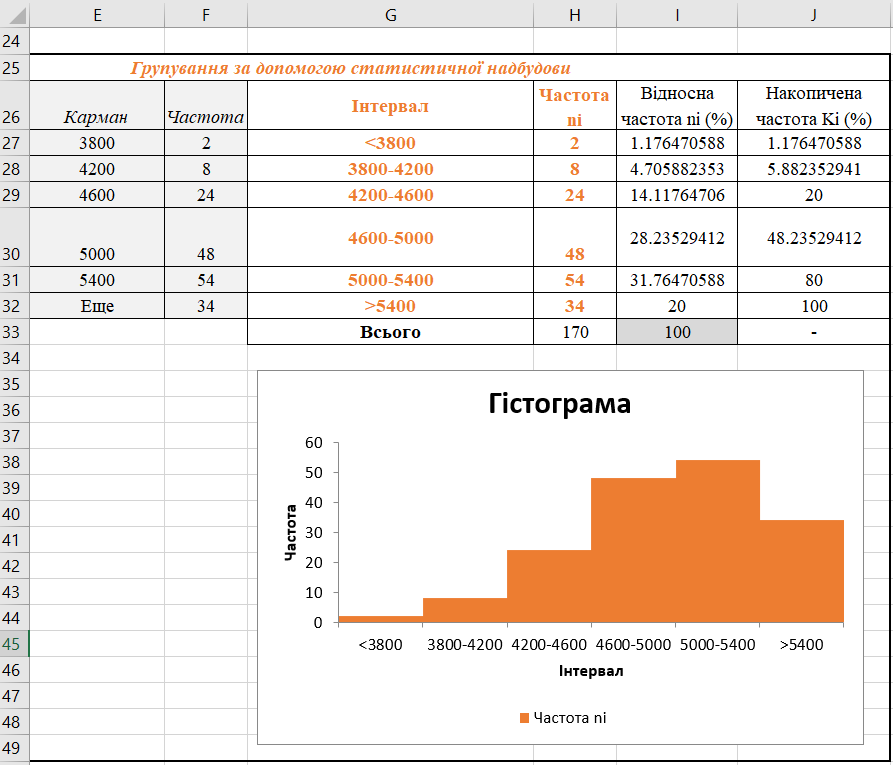


Рис. 1.11 Результат після доповнення таблиці та настройки графіка

*Групування за допомогою формул*

Приклад групування за допомогою формул наводиться на рис. 1.13. Підрахунок частоти потрапляння в інтервал значень визначається як різниця кількості значень менше верхньої межі і менше нижньої межі інтервалу. Наприклад, частота для першого інтервалу (комірка Q29) розрахована за допомогою функції *СЧЁТЕСЛИ*:

*СЧЁТЕСЛИ*(B:B;"<="&O29)-*СЧЁТЕСЛИ*(B:B;"<="&N29)

Відносна часта та накопичена частота розраховуються як у попередній частині про групування за допомогою статистичної надбудови.

За даною таблицею будується гістограма для цього вибираємо в меню *[Вставка → Діаграма → Гістограма → Звичайна гістограма].* Переходимо на закладку *Ряд* і натискаємо кнопку *Додати.* Натискаємо кнопку *Значення* і вказуємо діапазон значень частот. Натискаємо кнопку *Підпису осі Х* і вказуємо діапазон міток для осі Х. натискаємо кнопки *Далі → Готово.* Як міток можна вказати середини інтервалів групування.

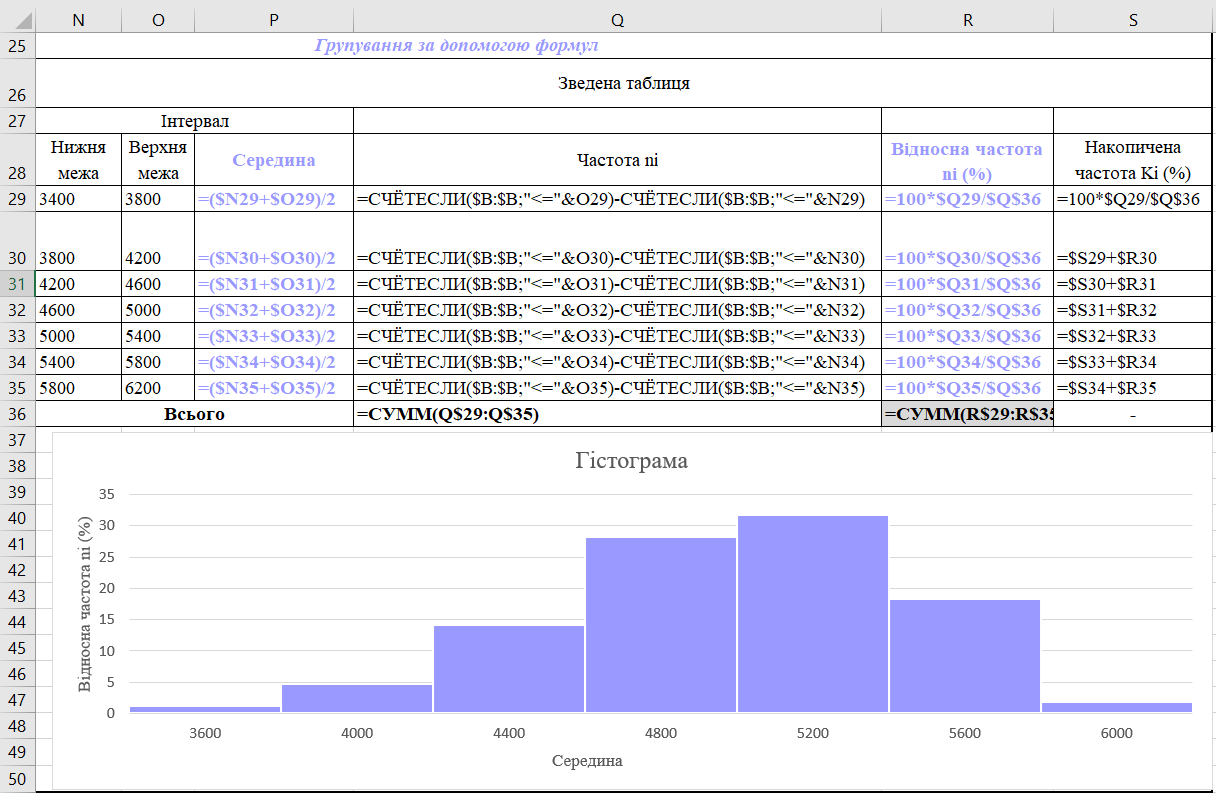


Рис. 1.12 Формули для зведеної таблиці та гістограма

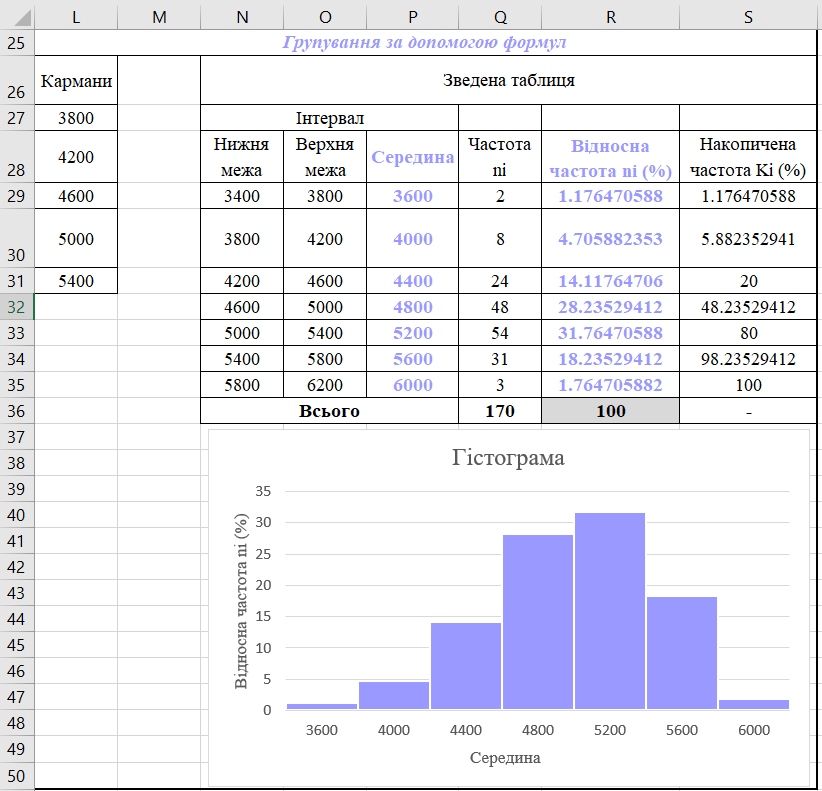
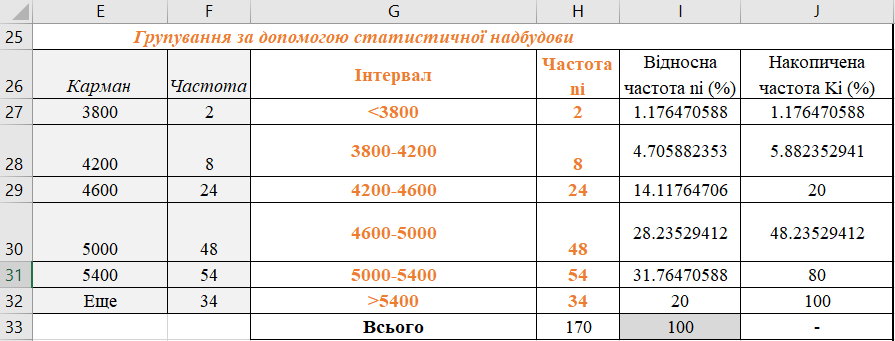


Рис. 1.13 Гістограма за результатом групування

Побудували зведені таблиці і далі порівняємо результати групування, отримані двома способами.

Поглянемо на таблиці отримані 2 способами (рис. 1.14).



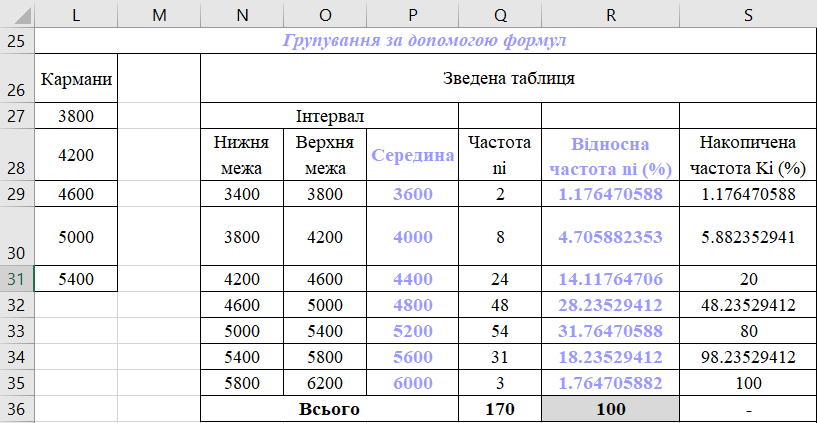


Рис. 1.14 Гістограми за результатом групування 2 способами

Можемо побачити що відносні та накопичені частоти відповідних інтервалів співпадають.

1. *Побудуйте гістограму, полігон і кумуляту двома способами:*

*- За допомогою статистичної надбудови*

*- «Вручну».*

Використаємо відповідні дані отримані за допомогою статистичної надбудови та за допомогою формул відповідно до попереднього кроку.

В першому випадку кумулята – Інтегральний відсоток.

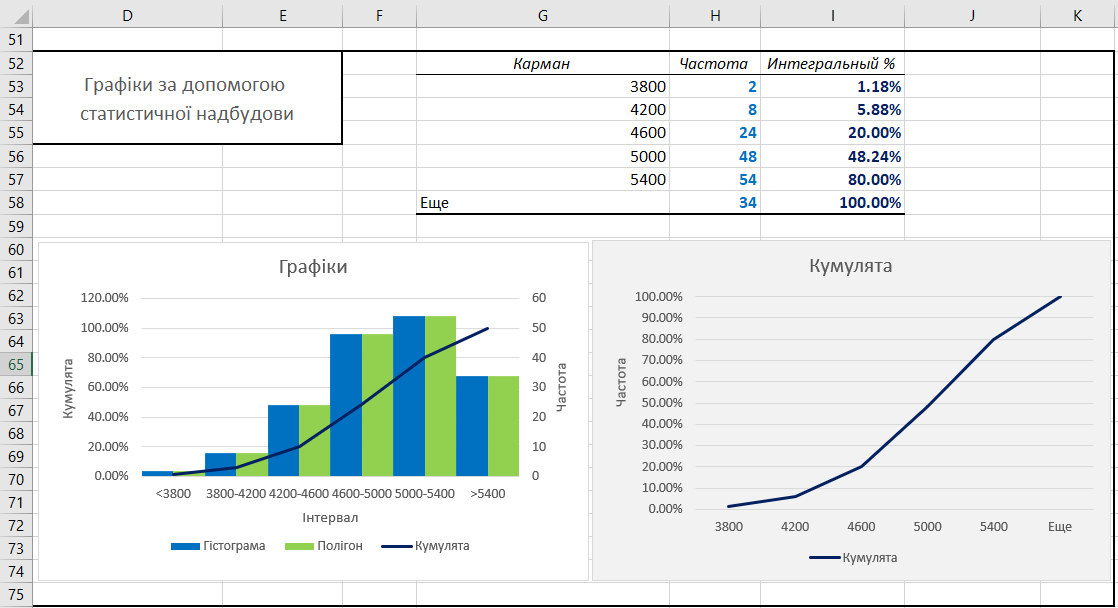
**

Рис. 1.15 Гістограма, полігон та кумулята за допомогою статистичної надбудови

У другому випадку кумулята - Накопичена частота, Ki (%).

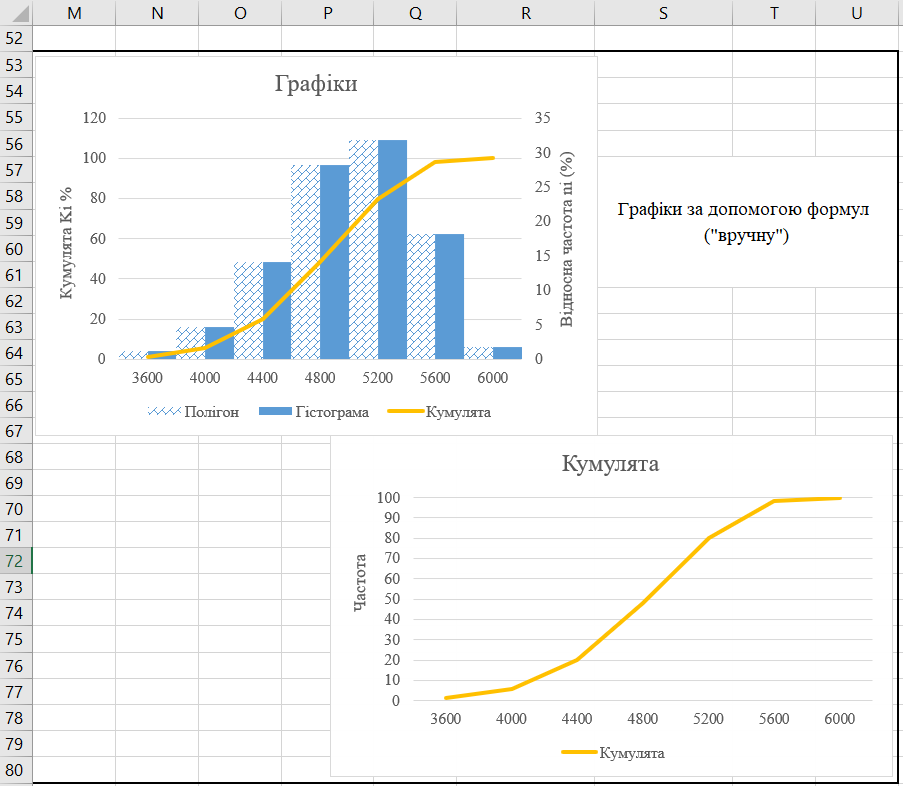
**

Рис. 1.16 Гістограма, полігон та кумулята за допомогою формул (вручну)

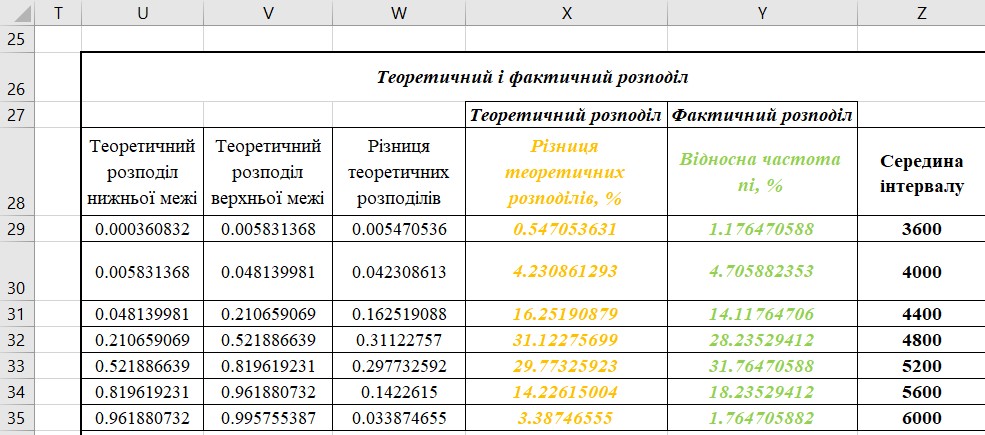
1. *Побудуйте гістограми теоретичного та емпіричного розподілів на одному графіку.*

Вихідна вибірка генерується за стандартним законом розподілу з параметрами відповідно до варіанта завдання. Щоб порівняти фактичний і теоретичний розподіл, необхідно побудувати їх графіки. Для роботи з теоретичними розподілами використовуються готові статистичні функції, наприклад, *НОРМРАСП і НОРМОБР*.

Щоб обчислити значення теоретичної ймовірності попадання випадкової величини в інтервал [x1, x2), необхідно знайти різницю ймовірності попадання в інтервали [0, x2) і [0, x1). Наприклад, ймовірність попадання випадкової величини в інтервал [3400, 3800) для нормального закону розподілу із середнім 5000 і стандартним відхиленням 500, дорівнює:

*= НОРМРАСП (3800; 5000; 500; ІСТИНА) -НОРМРАСП (3400; 5000; 500; ІСТИНА).*

Побудуємо необхідні таблиці з даними для теоретичного та фактичного розподілу та графік розподілів.



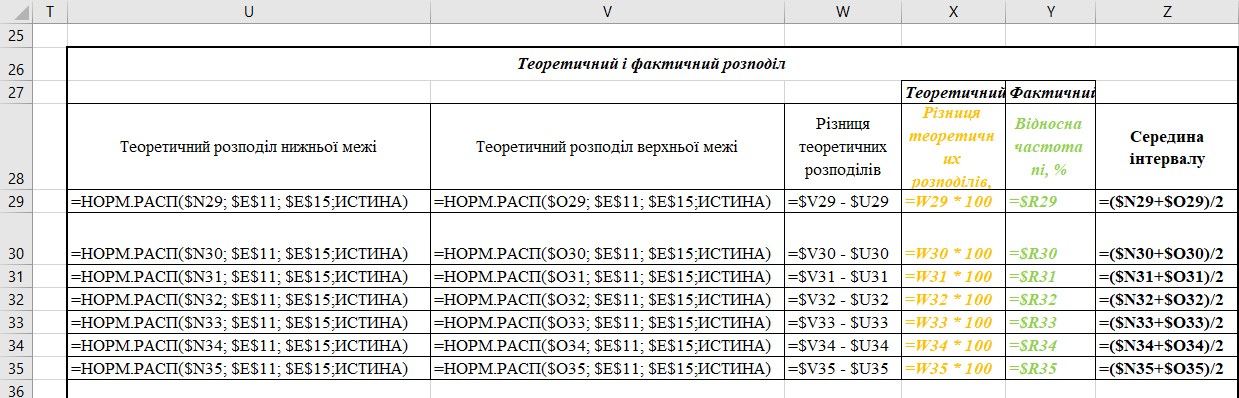


Рис. 1.17 Таблиці значень для теоретичного та фактичного розподілів

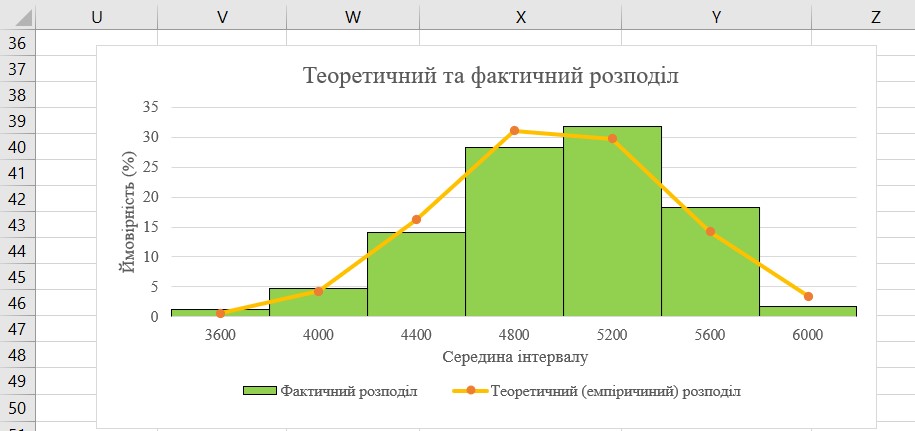


Рис. 1.18 Фактичний (емпіричний) та теоретичний розподіл

1. Зробіть висновок про близькість емпіричного розподілу до теоретичного.

Емпіричний розподіл достатньо наближений до теоретичного.

**Висновки**

В ході виконання практичної роботи було проаналізовано статистичні дані нормального розподілу по витратам енергії в умовних одиницях при проведенні експерименту. Були розраховані основні показники варіації. Під час аналізу отриманих результатів було встановлено, що вибірка однорідна та форма розподілу близька до нормальної. Це підтверджується показниками коефіцієнта варіації, асиметрії та ексцесу. Також були побудовані гістограми теоретичного та емпіричного розподілів, які показують близькість даного за умовою завдання розподілу до нормального.

Контрольна задача до захисту практик 1 - 3

***8*** *В потомстві від схрещування шести пар золотистих хом’ячків отримано 39 золотистих і 14 білих особин. Визначте частку тих і інших та розрахуйте середнє квадратичне відхилення із абсолютних і відносних частот отриманих в потомстві особин.*

Зробимо розрахунки на новому аркуші з назвою «Контрольна задача».

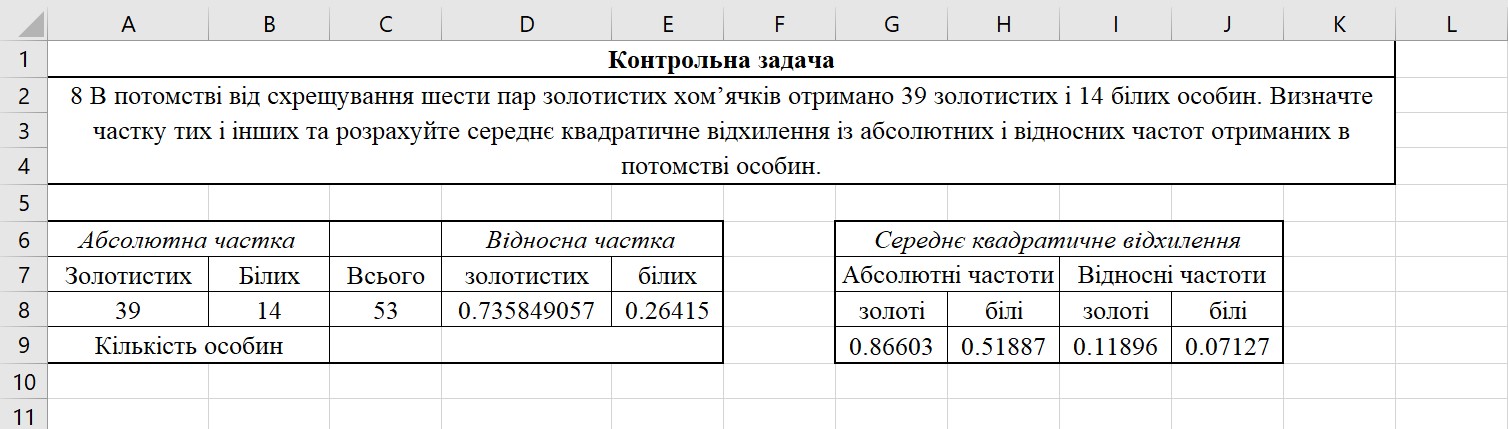


Рис. 1.19 Розраховані величини для контрольної задачі на аркуші «Контрольна задача».

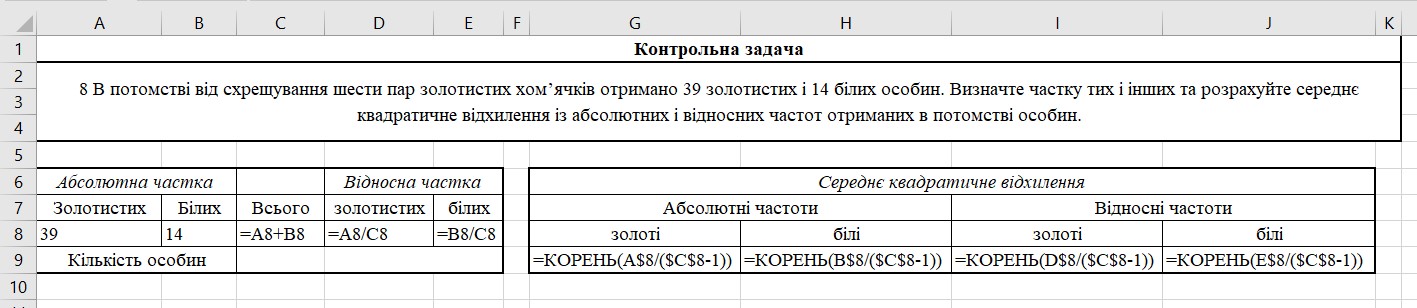


Рис. 1.20 Формули для розрахованих величин контрольної задачі

ДОДАТОК А

Таблиця 1.1

Показники варіації

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Назва** | **Позначення** | **Назва у зводній таблиці** | **Метод обчислення** | **Формула Excel** |
| 1 | 2 | 3 | 4 | 5 |
| Розмах варіації | R | Інтервал | Різниця максимального і мінімального значень | МАКС(інтервал)-МІН(інтервал) |
| Об’єм вибірки | n | Рахунок | Кількість статистичних одиниць | РАХУНОК(інтервал) |
| Медіана | Me | Медіана | Центральне значення відсортованої вибірки | МЕДІАНА(інтервал) |
| Мода | Mo | Мода | Найбільш часто зустрічається значення | МОДА(інтервал) |
| Середнє |  | Середнє | Середнє арифметичне | СРЗНАЧ(інтервал) |
| Середнє лінійне відхилення | , *d* | - | Середній модуль відхилення від середнього значення | СРВІДХ(інтервал) |
| Дисперсія |  | Дисперсія | Середній квадрат відхилення від середнього значення | ДИСП(інтервал) |
| Середнє квадратичне відхилення |  | - | середнє квадратичне відхилення від середнього значення | СТАНДВІДХ(інтервал) |
| Середнє квадратичне відхилення(незміщена оцінка) |  | Стандартне відхилення | середнє квадратичне відхилення від середнього значення з поправкою на обсяг вибірки | СТАНДВІДХ(інтервал)- незміщена оцінка |

Закінчення таблиці 1.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 |
| Коефіцієнт осциляції |  | - |  | - |
| Лінійний коефіцієнт варіації |  | - |  | - |
| Коефіцієнт варіації |  | - |  | - |
| Коефіцієнти асиметрії | A | асиметричність |  | - |
| Коефіцієнт ексцесу | E | ексцес |  | - |

**Варіанти завдання**

Таблиця 1.2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № варіанта | Випадкова величина | Параметри | | n |
|  | **Рівномірний розподіл** | **min** | **max** |  |
| 1 | Зріст,см | 155 | 190 | 150 |
| 2 | Вага,кг | 40 | 90 | 100 |
| 3 | Витрати на споживання | 2000 | 3500 | 170 |
| 4 | Місячна зарплата, грн. | 3000 | 7000 | 180 |
| 5 | ВВП.трлн. дол. | 2 | 5 | 180 |
|  | **Нормальний розподіл** |  |  |  |
| 6 | Ціна автомобіля,тис.дол. | 20 | 5 | 140 |
| 7 | Число студентів у групі, чол. | 31 | 10 | 190 |
| 8 | Витрати енергії в умовних одиницях при проведенні експерименту | 5000 | 500 | 170 |
| 9 | Чисельність складових | 20 | 4 | 150 |
| 10 | Тривалість експерименту, хв. | 240 | 60 | 140 |