Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

**ЗВІТ**

Самостійні роботи № 2,3,4

**Дисципліна**

Теорія ймовірностей та математична статистика

*Варіант 8*

Виконавець: студент

Пороскун Олена Олегівна

Група ПМ-81

Викладач: Гончаров Олександр Андрійович

Суми, Сумська область

2020

***САМОСТІЙНА РОБОТА №2***

СТАТИСТИЧНА ОБРОБКА ОДНОВИМІРНОГО ВИПАДКОВОГО МАСИВУ

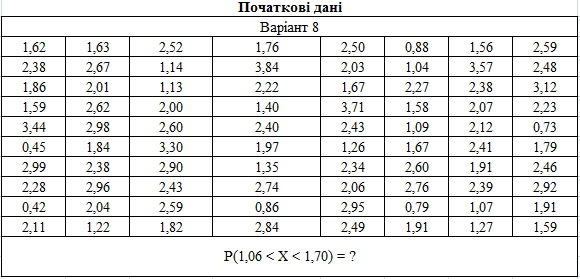
**Мета самостійної роботи**

Повинні бути набуті такі *вміння*: будувати гістограми; обчислювати ймовірності потрапляння випадкової величини в заданий проміжок; обчислювати ймовірність відхилення випадкової величини від математичного сподівання не більше ніж на задану величину.

Мають бути засвоєні такі *поняття*: обсяг вибірки, математичне сподівання, розмах вибірки, середньоквадратичне відхилення, асиметрія, ексцес, теоретична й емпірична щільність нормального розподілу.

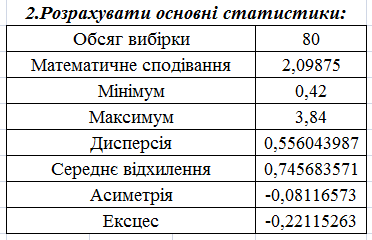
***Завдання до самостійної роботи***

1. Створити файл початкових даних.



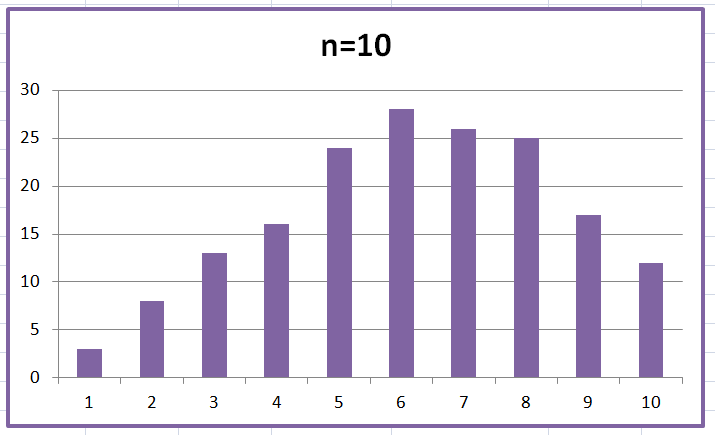
2. Розрахувати основні статистики: обсяг вибірки (функція СЧЕТЗ () або COUNTA()), математичне сподівання (функція СРЗНАЧ () або AVERAGE()), розмах вибірки (функції МИН (), МАКС () або MIN (), MAX()), середнє квадратичне відхилення (функція СТАНДОТКЛОН () або STDEV()), асиметрію (функція СКОС () або SKEW()), ексцес (функція ЭКСЦЕСС () або KURT()).

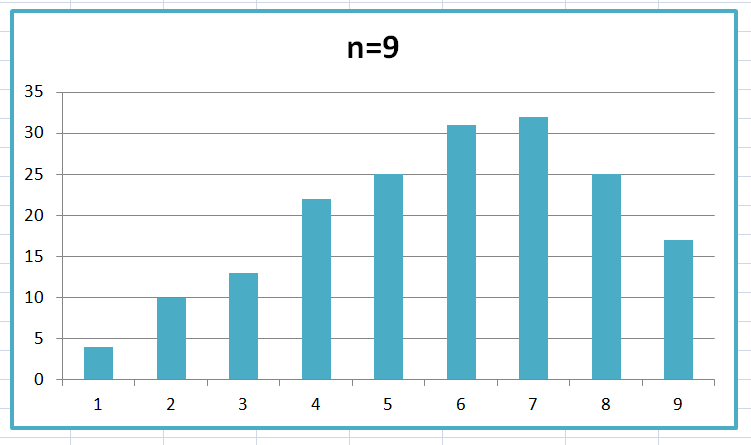
Одержимо (табл. 2.1).

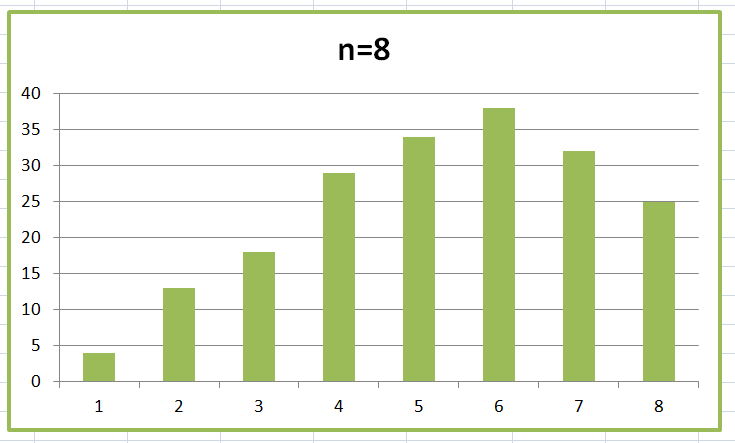


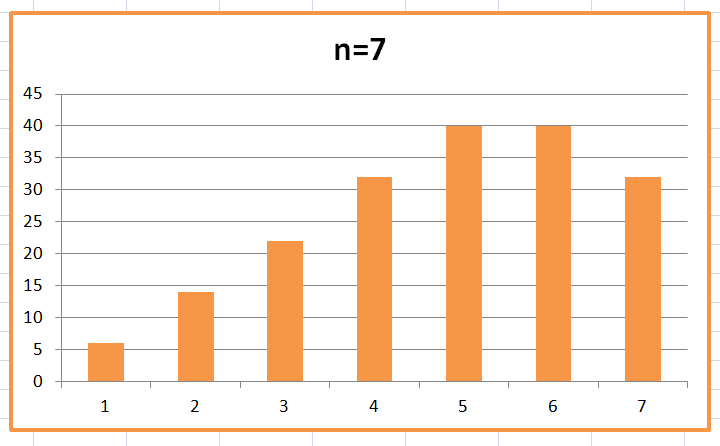
*Таблиця 2.1*

3. Побудувати гістограми. Взяти число часткових інтервалів, що дорівнює 10 (потім 9, 8, 7). Одержимо гістограми (рис. 2.2-2.6).



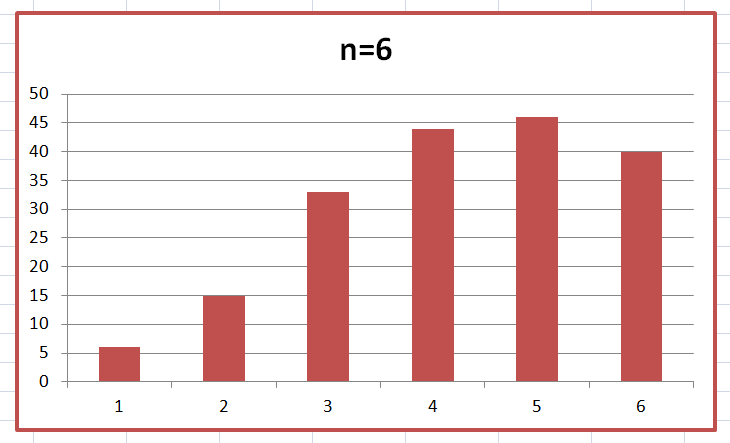




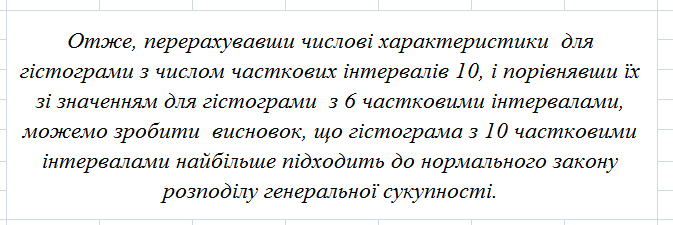


*Рисунки 2.2-2.5*

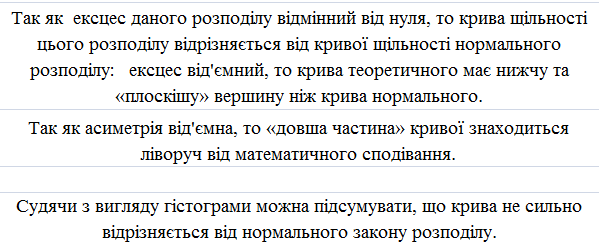
*Визначити, яка з одержаних гістограм задовольняє вимоги, поставлені до гістограм, пояснити чому. Зменшити число часткових інтервалів до 6 або 5 і побудувати нову гістограму. Дайте відповідь на запитання: чи можна ще зменшити число часткових інтервалів?*

**

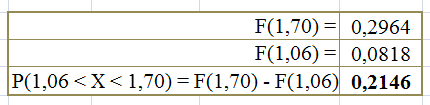
*Рисунок 2.6*

**

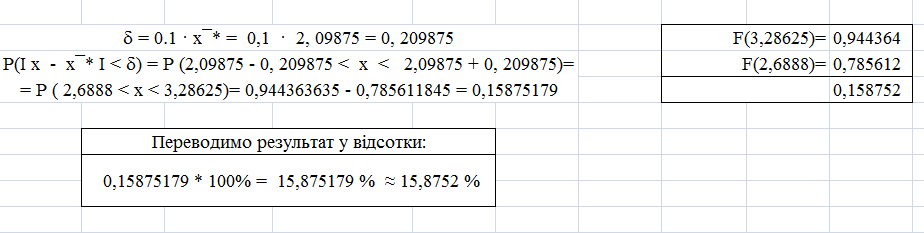
*3. За значеннями асиметрії та ексцесу й виглядом гістограми зробити висновок, чи значно відрізняється розподіл випадкової величини від нормального.*

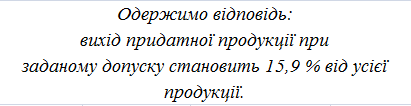
**

*4. Знайти ймовірність потрапляння випадкової величини у проміжок, заданий у варіанті завдання. Імовірність потрапляння випадкової величини в проміжок (a, b) розраховуємо за формулою P (a < X < b) = F (b) – F (a), де F (x) – функція розподілу випадкової величини. Значення функції розподілу знаходимо за допомогою стандартної функції НОРМРАСП () (чи NORMDIST()). Одержимо (наприклад, для а = 1,06; b = 1,70) :*

**

*5. Вважаючи, що технологічний процес відрегульований правильно, а допуск становить 10 % від значення контрольованого параметра, знайти випуск придатної продукції у відсотках.*

**

**

***САМОСТІЙНА РОБОТА №3***

ПРОГНОЗ НА ОСНОВІ ЛІНІЙНОЇ РЕГРЕСІЇ. ТОЧНІСТЬ ПРОГНОЗУ. ЩІЛЬНІСТЬ ЛІНІЙНОГО ЗВ'ЯЗКУ

**Мета самостійної роботи**

Повинні бути набуті такі *вміння*:

1) знаходження графіка та рівняння лінійної регресії в довірчій області при заданому рівні довіри;

2) знаходження за графіком прогнозу відгуку при заданому значенні фактора;

3) знаходження за графіком максимальної абсолютної похибки прогнозу;

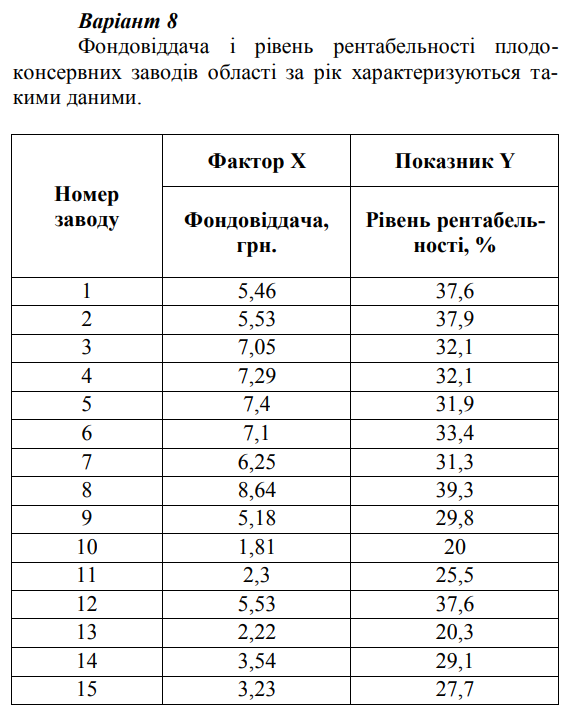
4) розрахунок максимальної відносної похибки прогнозу у відсотках;

5) оцінювання тісноти лінійного зв'язку за значенням коефіцієнта кореляції.

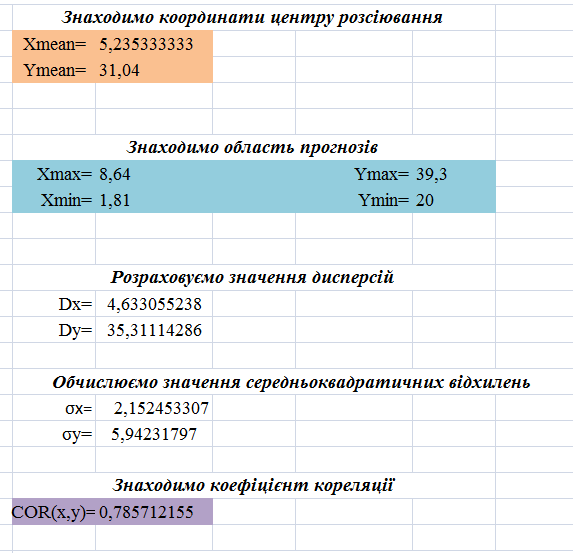
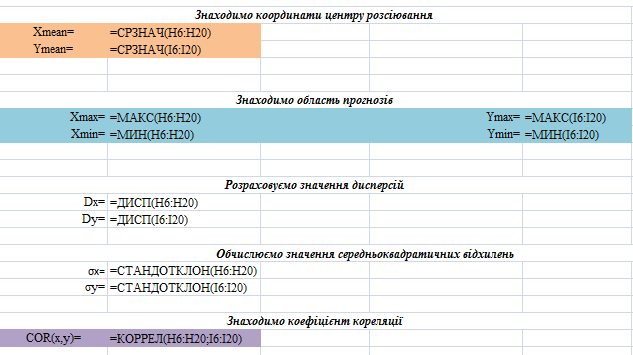
Мають бути засвоєні такі *поняття*: кореляційне поле, область прогнозів, прогноз, довірча область, довірчий інтервал, рівень довіри, напівширина довірчого інтервалу, абсолютна і відносна похибки прогнозу, коефіцієнт кореляції, тіснота лінійного зв'язку.

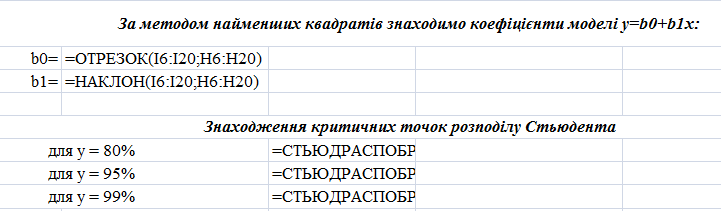
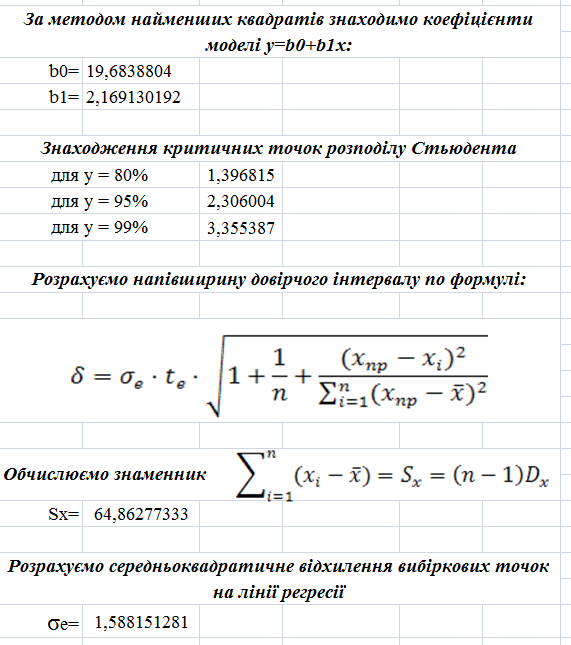
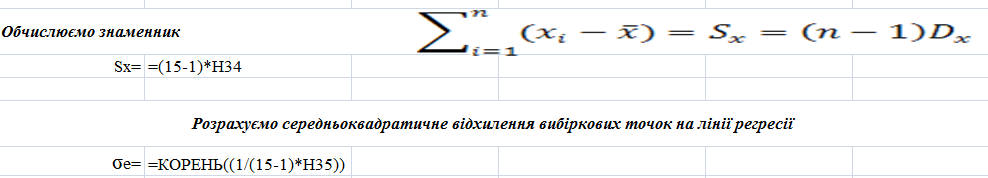
***Завдання до самостійної роботи***

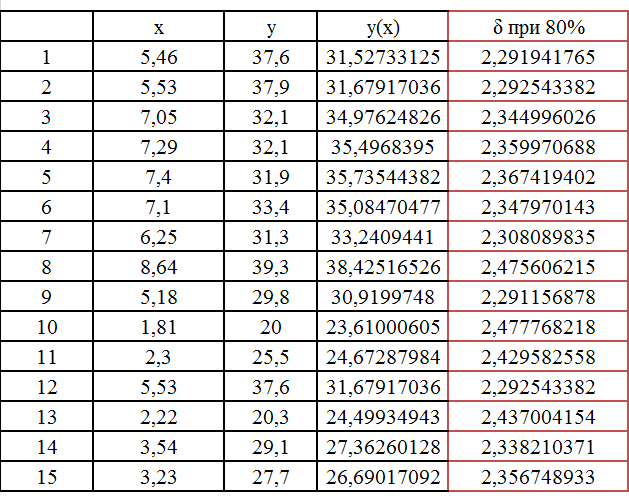
Використовуючи дані зі свого індивідуального завдання, виконати таке:

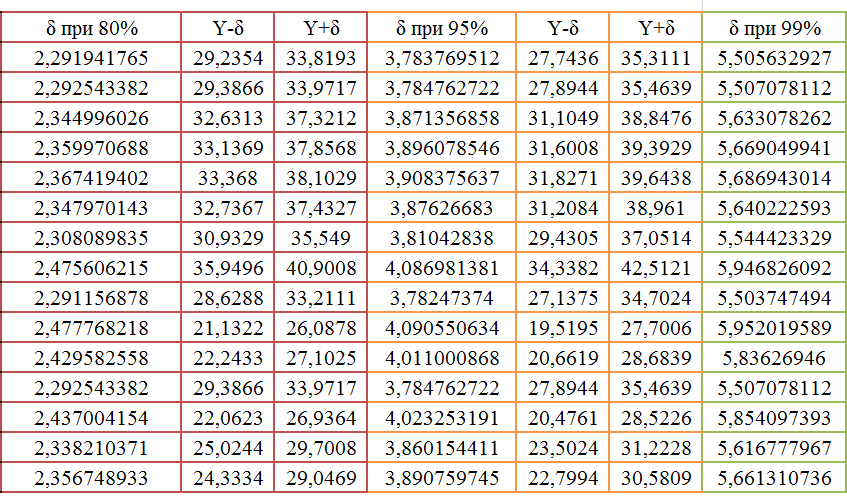


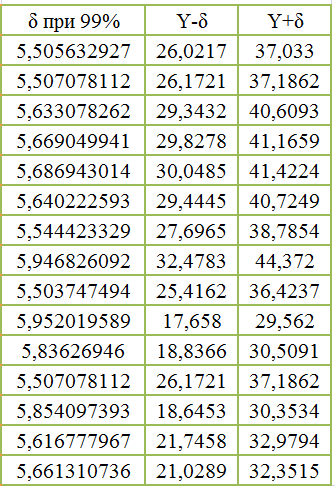
1. Побудувати графіки лінії регресії з 80, 95 і 99 % довірчими областями.



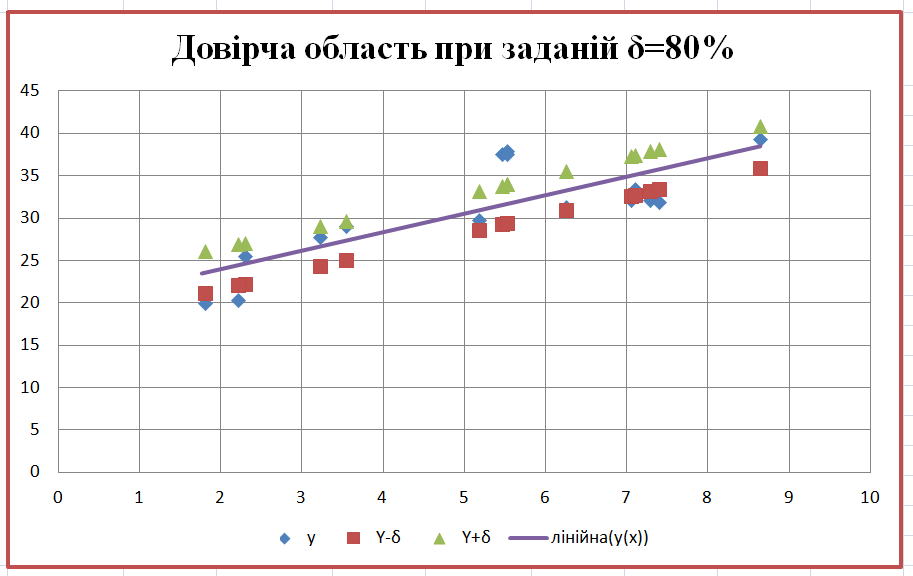


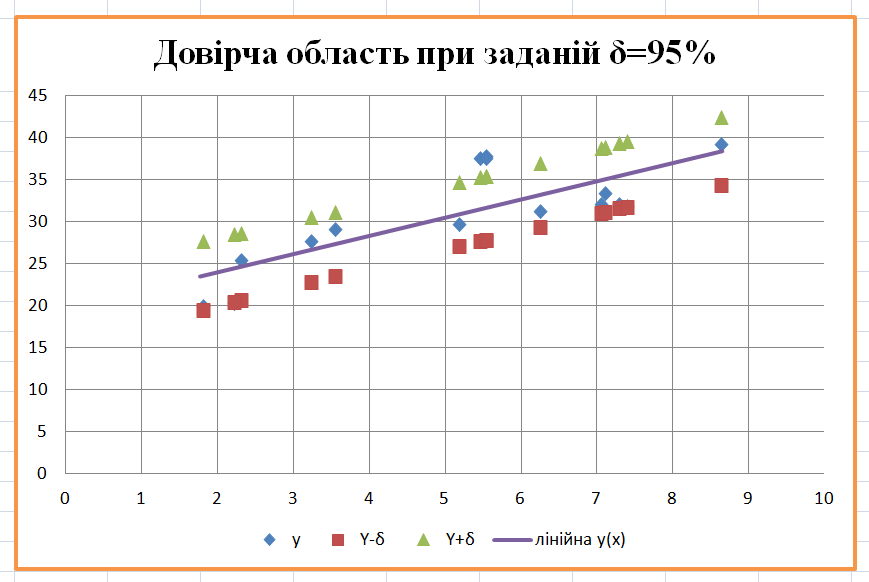


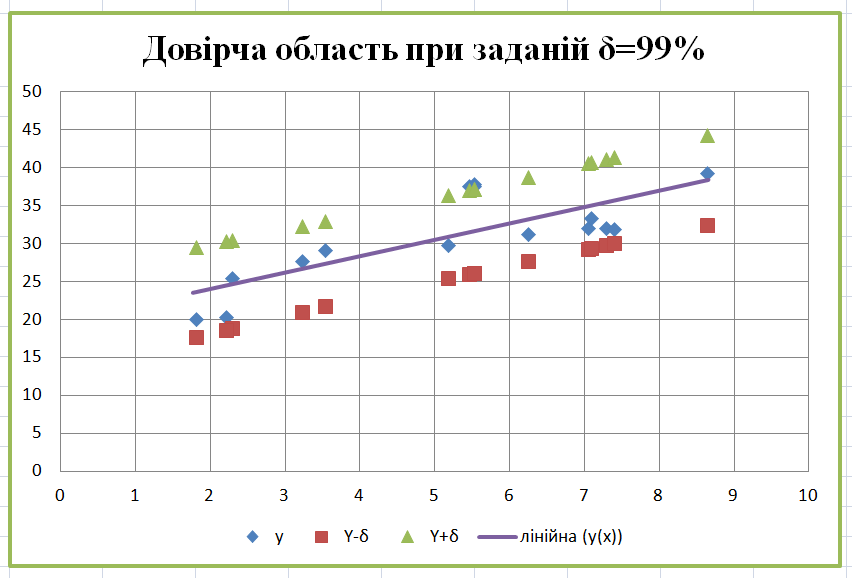




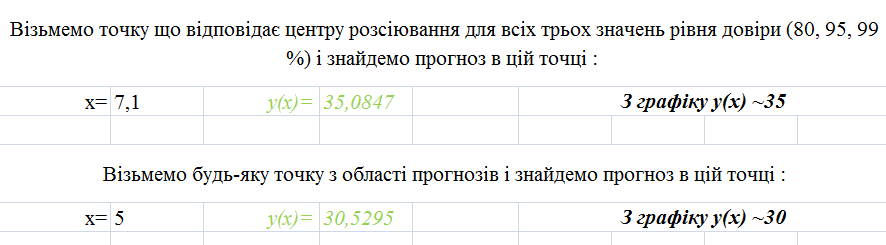
2. Нанести вручну на лінію регресії центр розсіювання.



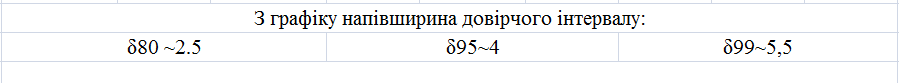




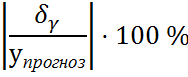
3. Знайти за графіком прогноз в точці, що відповідає центру розсіювання для всіх трьох значень рівня довіри (80, 95, 99 %), а також прогноз в будь-якій довільній точці з області прогнозів.

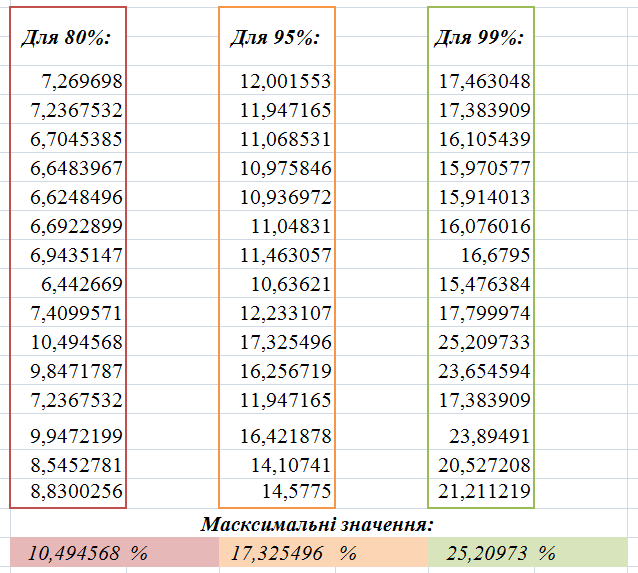
******

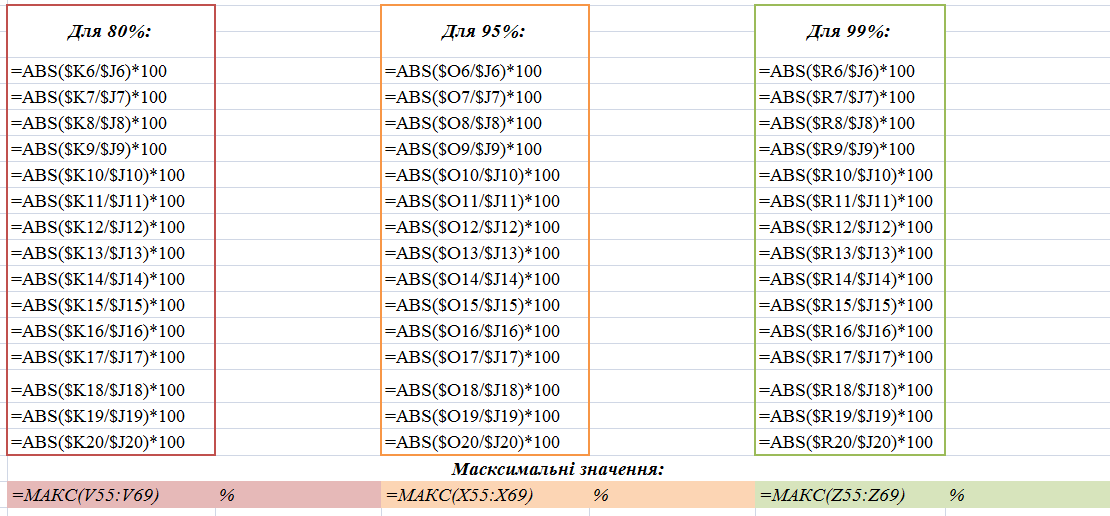
4. Знайти за графіком напівширину довірчого інтервалу в точці, що відповідає центру розсіювання для всіхтрьох значень рівня довіри (80 , 95 , 99 %): δ80, δ95, δ99.



5.Оцінити максимальну відносну помилку прогнозу (у відсотках) для всіх трьох значень рівня довіри (80, 95, 99 %) за формулою

 (δγ і у прогноз знаходимо за кресленням).





***САМОСТІЙНА РОБОТА №4***

КРИТЕРІЙ УЗГОДЖЕНОСТІ ПІРСОНА, КОЛМОГОРОВА – СМІРНОВА

**Мета самостійної роботи**

Повинні бути набуті такі *вміння*: побудова гістограм у парі з кривими розподілу; розбиття розмаху вибірки на часткові інтервали довільної довжини; знаходження критичного значення для заданого рівня значущості; розрахунок теоретичних частот розподілу.

Повинні бути освоєні такі *поняття*: критерій згоди, обрані (емпіричні) частоти, теоретичні частоти, статистична гіпотеза, рівень значущості гіпотези, число ступенів свободи.

*Індивідуальні завдання до самостійної роботи взяти з самостійної роботи 2.*

Для вибірки знаходимо основні її характеристики:

1. Вибіркове (емпіричне) середнє. Формула Excel=СРЗНАЧ (< діапазон вибірки >)

2. Дисперсія. Формула Excel = ДИСП.В (< діапазон вибірки >)

3. Середньоквадратичне відхилення. Формула Excel = КОРЕНЬ (дисперсії)

4. Обсяг вибірки - Формула Exсel = СЧЁТ (< діапазон вибірки >)



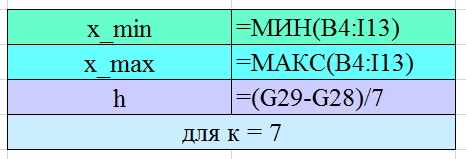
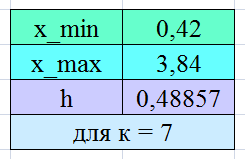
 Одержуємо значення:



Використання **критерію Пірсона** для визначення вигляду розподілу вимагає розділення усього розмаху вибірки на n часткових інтервалів. *Розмахом вибірки* є відстанню між максимальним і мінімальним членами варіаційного ряду (масиву даних).

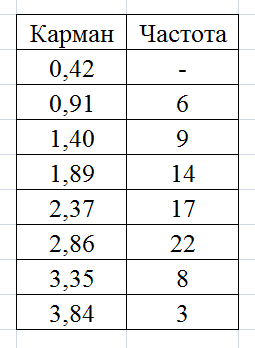
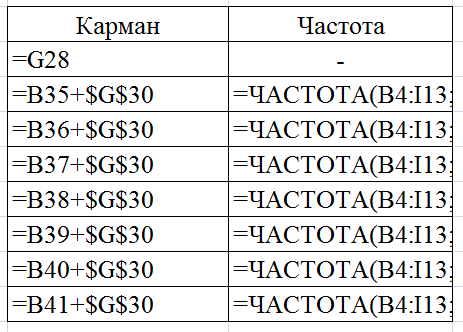
Для знаходження мінімального і максимально елементів вибірки використовуємо формули Excel : = МИН (< діапазон вибірки >) = МАКС (< діапазон вибірки >). Кількість часткових інтервалів беруть 7 – 10.

Для визначення кроку для розбиття розмаху вибірки використовуємо формулу: =(МАКС () -МИН ())/k, де k – обрана кількість інтервалів.



Для одержання початкової гіпотези H0 побудуємо гістограму, відклавши по горизонталі кармани вибірки, по вертикалі – частоти.

Розрахунки для ручної побудови гістограми:

Розраховані значення використовуємо при побудові гістограми.

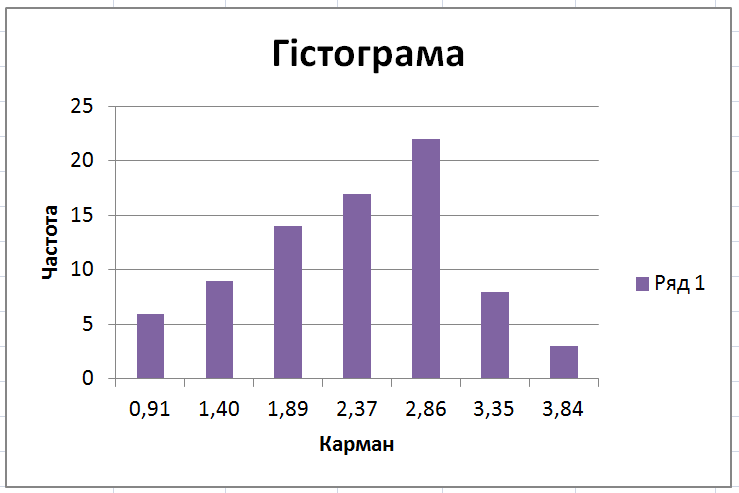


Рисунок 4.1 – Побудована гістограма

Використання критерію Пірсона. Для розрахунку побудуємо таку таблицю:



Кармани, які ми знайшли раніше, переносимо в перший стовпець. Емпіричні частоти – кількість потраплянь у і-й інтервал. Ці частоти ми отримали за допомогою функції = ЧАСТОТА (...).

Теоретичні частоти – це ймовірність попадання в і-й інтервал, помножена на обсяг вибірки.

За припущення, що вибірка має нормальний розподіл, теоретичні частоти обчислюватимемо таким чином: = (НОРМРАСП (X; середнє; стандарт\_відхилення; 1) -НОРМРАСП (X - 1; середнє; стандарт\_відхилення; 1)) \*СЧЁТ(< діапазон вибірки >), де Х- 1 і Х – кінці інтервалу [Х- 1; Х].

За припущення, що вибірка має експоненціальний розподіл, теоретичні частоти обчислюємо таким чином: = (EXP (-Х/середнє) – EXP (- (X - 1) / середнє)) \*СЧЁТ (< діапазон вибірки >), де Х-1 і Х – кінці інтервалу [Х- 1; Х].

За припущення, що вибірка має рівномірний розподіл, теоретичні частоти обчислюємо таким чином: = СЧЁТ (< діапазон вибірки >) \* (X - (X - 1)) / ( МАКС (< діапазон вибірки >) -МИН (< діапазон вибірки >)), де Х- 1 і Х – кінці інтервалу [Х- 1; Х]. Після знаходження теоретичних та емпіричних частот знаходимо X^2 набл = ((Емп.част – Теор.част.) ^ 2 / Теор. част.) на і-му інтервалі.

Після того як заповнена вся таблиця, будуємо гістограму теоретичних і емпіричних частот, щоб оцінити вибірку.

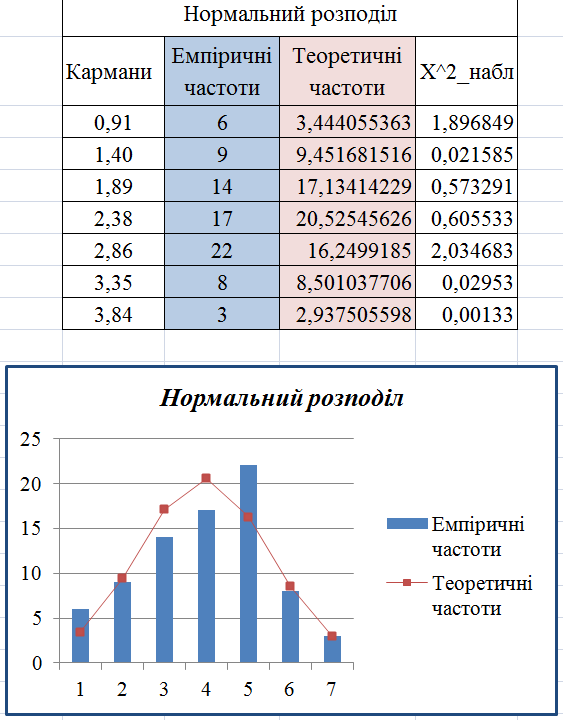


Рис 4.2 – Гістограма теоретичних та емпіричних частот (нормальний розподіл)

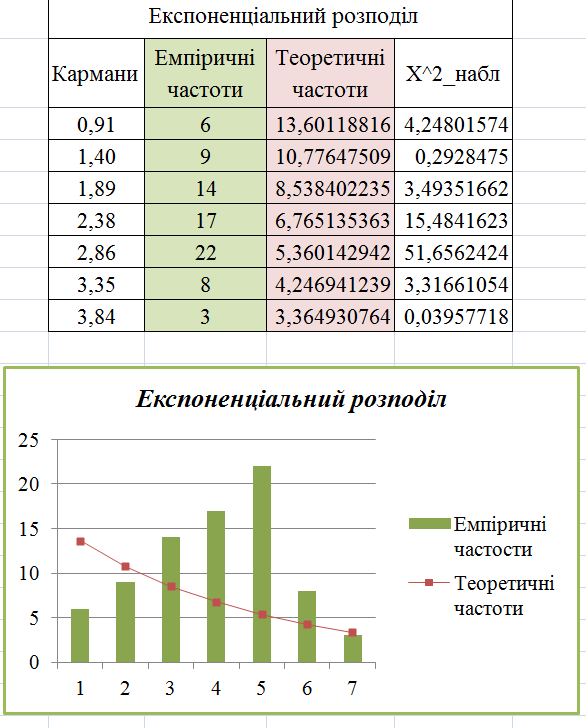


Рисунок 4.3 – Гістограма теоретичних та емпіричних частот (експоненціальний розподіл)

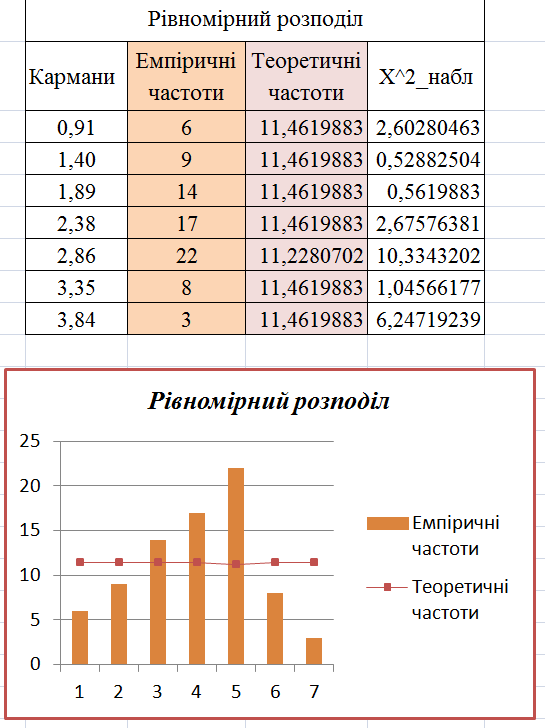
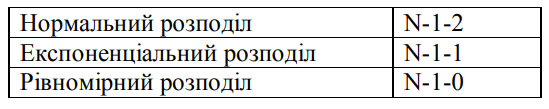


Рисунок 4.4 – Гістограма теоретичних та емпіричних частот (рівномірний розподіл)

Виконуємо оцінювання.

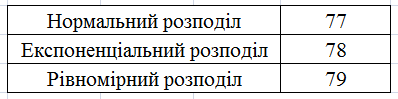
Знайдемо X^2 як суму X^2\_набл на усіх і-х інтервалах. Критичне значення X^2\_кр є табличним. Його значення одержуємо за допомогою функції Excel =ХИ2.ОБР (рівень значущості; кількість ступенів свободи). Рівнем значущості є ймовірність відхилити правильну гіпотезу. Як правило, рівень значущості набуває значень 0,05 і 0,1.

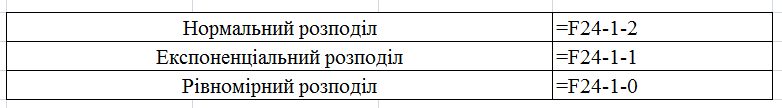
Кількість ступенів свободи визначають таким чином:

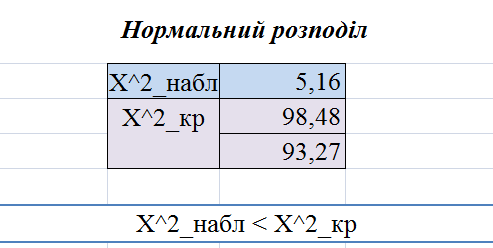
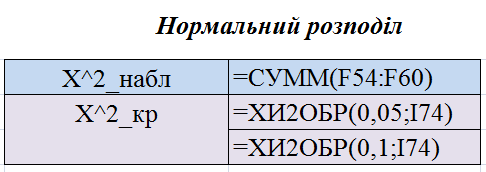


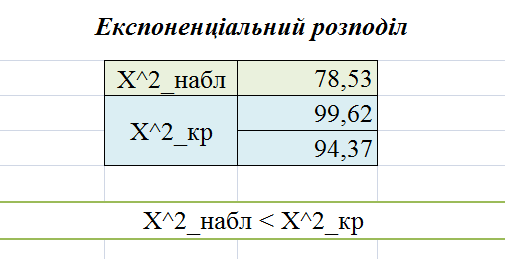
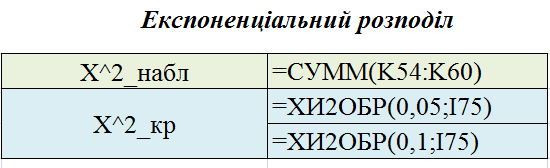
де N – обсяг вибірки, параметр 0; 1 або 2 показує кількість характеристик, від яких залежить розподіл.

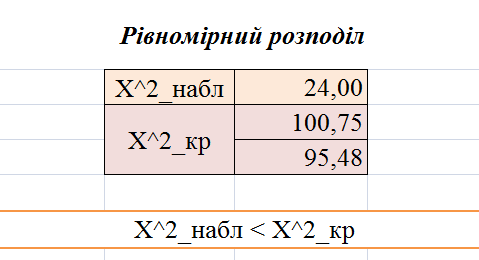
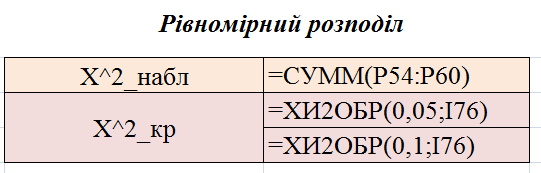
Маємо:











Якщо X^2\_набл < X^2\_кр – гіпотезу приймаємо, в іншому разі гіпотезу відхиляємо.

Якщо гіпотеза виявилася 97 неправильною, перевіряємо вибірку для іншого вигляду розподілу.

Якщо жодна гіпотеза не підтвердилася при використанні критерію Пірсона, то можна збільшити рівень значущості або ж виконати розрахунок за критерієм Колмогорова – Смирнова.

***Отже***, *так як X^2\_набл < X^2\_кр – гіпотезу приймаємо(за 3 розподілами).*