

# ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

методичні рекомендації до виконання  
лабораторних робіт

Запоріжжя 2024



УДК 519.6:311.2(072)  
Т30

Рекомендовано Науково-методичною радою  
ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
«МЕТІНВЕСТ ПОЛІТЕХНІКА»  
(протокол № 8 від 05.07.2024 р.)

**Укладач**

Грудкіна Н.С., д-р техн. наук, доцент

- Т30 Теорія ймовірностей та математична статистика : методичні рекомендації до виконання лабораторних робіт (для студентів економічних спеціальностей усіх форм навчання першого (бакалаврського) рівня вищої освіти) / уклад. Н. С. Грудкіна. Запоріжжя : ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА». 2024. 22 с.

Методичні вказівки містять відомості щодо виконання лабораторних робіт в рамках вивчення дисципліни «Теорія ймовірностей та математична статистика» та критерії оцінювання. Матеріал навчального посібника має на меті підвищити рівень знань студентів із елементів статистичного аналізу та спрямований на вдосконалення практичних навичок з візуалізації даних, в тому числі із використанням можливостей MS Excel, аналізу отриманих результатів та формулювання висновків. Крім того, посібник покликаний покращити якість навчального процесу шляхом поєднання теоретичних знань із практичними завданнями, стимулювати критичне мислення студентів через самостійний пошук рішень і аналіз помилок, а також підвищити зацікавленість у предметі завдяки інтерактивним і прикладним методам навчання. Рекомендовано для студентів економічних спеціальностей усіх форм навчання першого (бакалаврського) рівня вищої освіти.

УДК 519.6:311.2(072)

© ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ МЕТІНВЕСТ ПОЛІТЕХНІКА», 2024



## ЗМІСТ

ВСТУП .....	4
1. Вимоги до виконання лабораторних робіт .....	5
2. Лабораторні роботи .....	6
2.1 Лабораторна робота №1 «Використання діаграми розмаху – «коробки з вусами»» .....	6
2.2 Лабораторна робота №2 «Побудова регресійної моделі»..	12
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	20
ДОДАТОК .....	21



## ВСТУП

Теорія ймовірностей та математична статистика – базовий курс, який належить до циклу математичної підготовки та присвячений формуванню у студентів здатності застосовувати математичні методи та алгоритмічні принципи для моделювання, аналізу й обробки даних технічних, природничих і соціально-економічних об'єктів і процесів, знань з методології та інструментарію побудови і використання різних типів стохастичних математичних моделей, що забезпечує фундамент для опанування професійно-орієнтованих дисциплін спеціальності.

Курс містить відомості з елементів комбінаторики, теорії ймовірностей та математичної статистики. Особливістю курсу є фокус на прикладну направленість математичної підготовки із використанням комп'ютерно-інформаційних технологій та пакетів математичних прикладних програм для глибокого розуміння та критичного осмислення теорій, принципів, методів і понять у сфері професійної діяльності. Сформовані навички сприятимуть розв'язуванню складних спеціалізованих задач аналізу та управління даними та аналітичному обґрунтування бізнес-рішень за комплексності та невизначеності умов. При навчанні за освітніми програмами «Вартісне управління бізнесом» та «Бізнес-аналітика» цей освітній компонент є обов'язковим, оскільки дозволяє набути переваг конкурентоспроможного на ринку праці фахівця, який вільно володіє професією і орієнтується в суміжних галузях діяльності, засвідчує готовність до постійного професійного зростання, соціальної й професійної мобільності.

В рамках освоєння дисципліни студентам пропонується виконати лабораторні роботи в рамках опанування знань із змістовного модуля «Основні поняття математичної статистики», які спрямовані на формування практичних навичок і поглиблення теоретичних знань. Особливу увагу приділено не лише традиційним методам статистичного аналізу та візуалізації даних, а й застосуванню сучасних інструментів для обробки даних та можливостям віртуальних навчальних середовищ. Це дозволить студентам набути навичок застосування математичного апарату до аналізу і прогнозування параметрів управлінських і бізнес-процесів для обґрунтування управлінських рішень, що в подальшому сприятиме підвищенню їхньої конкурентоспроможності на ринку праці.



## 1 ВИМОГИ ДО ВИКОНАННЯ ЛАБОРАТОРНИХ РОБІТ

В рамках курсу лабораторна робота №1 «Використання діаграми розмаху – «коробки з вусами»» виконується із використанням можливостей MS Excel, а лабораторна робота №2 «Побудова регресійної моделі» - у віртуальному навчальному середовищі, створеному на платформі Phet.colorado.edu. Оцінка за виконання кожної з лабораторних робіт виставляється за результатами оцінювання звіту, який долучається до відповідної активності на платформі Moodle. Максимальна оцінка у 5 балів формується з огляду на повноту наведення основних етапів розв'язання, правильність розрахунків та аналізу отриманого розв'язку, геометричної ілюстрації за потреби. Підготовлена лабораторна робота завантажується у вигляді файлу звіту з розширенням .docx або .pdf (за наявності розробленого розрахункового модуля у MS Excel у форматах .xls, .xlsx завантажується додатково) у відповідному розділі на платформі Moodle.

Допускається виправлення незначних ваг оформлення або розрахунку із завантаженням виправленої роботи наприкінці тижня складання роботи, встановленого у розділі «Розподіл балів за контрольними точками та графік їх виконання», що не знижує максимальну оцінку.

Для успішного виконання лабораторних робіт у рамках курсу рекомендується дотримуватися наступних вимог: ознайомлення з теоретичним матеріалом, дотримання інструкцій, викладених у методичних рекомендаціях, формування чіткого та логічного звіту, що включає постановку задачі, наведення основних етапів її виконання, правильність розрахунків та аналізу результатів із зазначенням відповідних висновків.

## 2 ЛАБОРАТОРНІ РОБОТИ

### 2.1 Лабораторна робота №1 «Використання діаграми розмаху – «коробки з вусами»

**Мета:** Лабораторна робота №1 спрямована на ознайомлення студентів із можливостями використання діаграми «коробка з вусами» для візуалізації даних. Основна мета роботи полягає у набутті навичок з наочного представлення порядкової статистики закону одновимірного розподілу, що дозволить виявити та графічно представити результати аналізу для ефективного прийняття управлінських рішень. Досягнення цієї мети через візуалізацію залежить від вибору та правильності застосування відповідного математичного та програмного інструментарію.

#### Хід і результати роботи:

1. **Ознайомитись з наведеним теоретичним матеріалом за темою.**

Серед ефективних інструментів табличного процесора MS Excel доступним та інформативним є діаграма розмаху, яку найчастіше називають «коробкою з вусами». Вона дозволяє отримати інформацію щодо мінімального значення, першого квартилю (Q1), медіани, третього квартилю (Q3), максимального значення та додатково середнього.

Діаграма «коробка з вусами» був запропонований Джоном Тьюкі ще 1969 році, і не зважаючи на меншу її інформативність у порівнянні з гістограмою, до її безперечних переваг можна віднести компактність та простоту побудови, що призвело до її широкого застосування фахівцями-аналітиками. Зазначимо, що при використанні діаграми розмаху не робиться ніяких припущень про закон розподілу вибірок, а отже її розглядають як інструмент непараметричної статистики.

Наведемо пояснення:

- **мінімальне значення:** мінімальне (найменше) значення в наборі даних (min);
- **перший квартиль:** середнє значення між мінімальним та медіанним - 25-й процентиль (Q1);
- **медіана:** таке число  $Me$  дискретного статистичного розподілу вибірки, яке ділить варіаційний ряд, що «породжує» цей розподіл, на дві рівні за кількістю варіант частини (якщо число варіант є непарним, тобто для обсягу вибірки  $n = 2m + 1$ ,  $Me = x_{m+1}$ ; для парного варіанту, тобто для  $n = 2m$ ,  $Me$  дорівнює середньому арифметичному «середньої» (медіанної) пари варіант);
- **третій квартиль:** середнє значення медіани і максимального - 75-й процентиль (Q3);



- **максимальне значення:** максимальне (найбільше значення) у наборі даних (max).

Додатково хрестиком на діаграмі розмаху позначають середнє вибіркоче (середня арифметична варіант) статистичного розподілу  $\bar{x}$ . Висота коробки визначається нижньою межею, що відповідає 25-му процентилю (Q1), тобто 25% спостережень мають значення нижче, та верхньою межею, що відповідає 75-му процентилю (Q3), де також 25% спостережень мають значення вищі за нього. Це означає, що 50% спостережень лежать в межах коробки.

Графічно представимо приклад діаграми «коробка з вусами» з розшифруванням відповідних позначень (рис. 1).

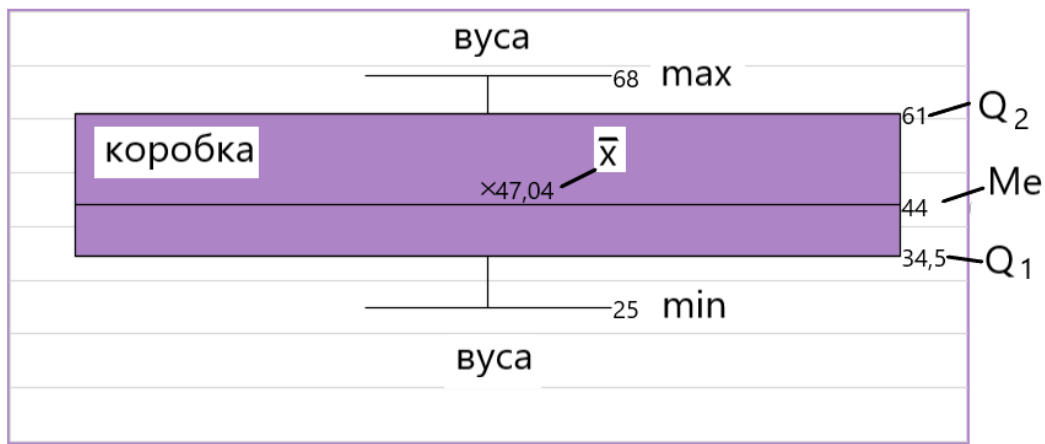


Рисунок 1 – Інтерпретація діаграми «коробка з вусами»

Зазначимо окремо, що вусами називають Т-подібні фігури, що виступають за межі коробки та у 1,5 рази довші за висоту коробки. Переважно їх довжина буде відповідати мінімальному і максимальному значенням. При нормальному розподілі даних близько 95% даних має лежати в діапазоні вусів. Мають місце також викиди – екстремальні значення, при чому зірочками позначають крайні викиди, які мають значення, що перевищують висоту коробки у три рази (рис. 2).

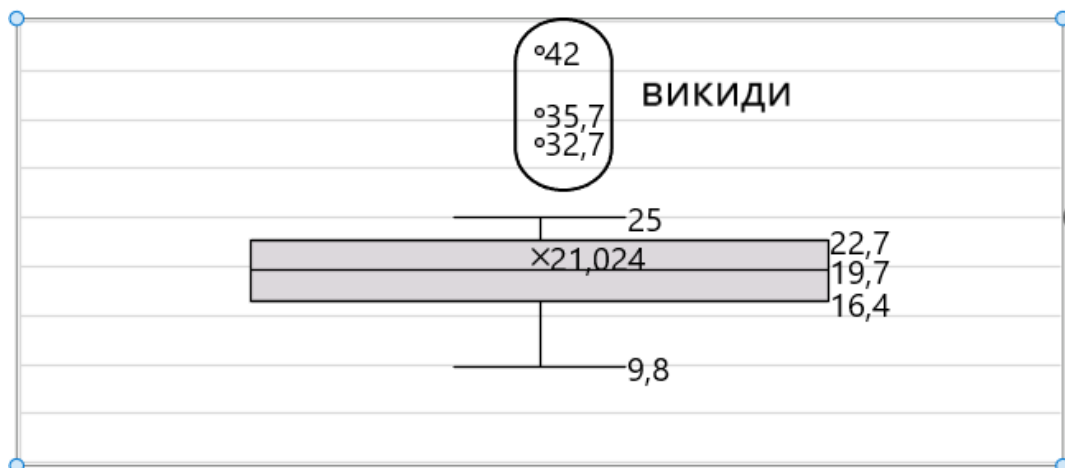


Рисунок 2 – Діаграми «коробка з вусами» з викидами

## 2. Виконати Завдання 1

**Завдання 1.** Виконати побудову діаграм «коробка з вусами» для стовпці ЗП (заробітної платні, тис. грн.), Вік (вік робітників, повних років), Стаж (стаж робітників, повних років) та Непрацездатність (кількість днів лікарняного за останні 5 років) та проаналізувати отримані дані.

Рекомендації щодо виконання завдання 1 містять наступні кроки:

- 1) завантажити дані з файлу ЗП.xlsx;
- 2) у додатково створених стовпцях Вік, Стаж та Непрацездатність за допомогою функції RANDBETWEEN(нижн\_межа;верх\_межа) рандомно у заданих згідно варіанту завдання межах (табл.1, рис. 3);

Таблиця 1 – Дані для завдання 1

№	Вік, повних років	Стаж, повних років	Непрацездатність, днів	№	Вік, повних років	Стаж, повних років	Непрацездатність, днів
1	(27;78)	(1;32)	(3;32)	11	(19;63)	(0;29)	(3;36)
2	(23;66)	(3;27)	(0;27)	12	(26;69)	(3;27)	(0;40)
3	(21;72)	(4;33)	(5;38)	13	(20;73)	(1;48)	(2;32)
4	(22;69)	(2;29)	(0;22)	14	(24;69)	(2;36)	(1;27)
5	(19;65)	(0;40)	(3;32)	15	(27;71)	(3;32)	(1;39)
6	(26;69)	(3;27)	(0;29)	16	(23;68)	(5;37)	(0;28)
7	(28;73)	(1;36)	(5;38)	17	(21;70)	(4;30)	(5;38)
8	(24;65)	(2;39)	(0;27)	18	(22;69)	(2;29)	(0;22)
9	(21;71)	(4;38)	(3;39)	19	(22;73)	(1;48)	(2;39)
10	(28;69)	(2;39)	(2;29)	20	(20;59)	(2;34)	(1;40)

Бібліотека функцій				Python	
2	№	ЗП, тис. грн.	Вік, повних років	Стаж, повних років	днів
3	1	19,8	=RANDBETWEEN(25;68)	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
4	2	22,7	=RANDBETWEEN(25;68)	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
5	3	32,7	=RANDBETWEEN(25;68)	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
6	4	19,7	=RANDBETWEEN(25;68)	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
7	5	12,7	=RANDBETWEEN(25;68)	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)

Рисунок 3 – Заповнення стовпців Вік, Стаж та Непрацездатність

3) для кожного з стовпців ЗП, Вік, Стаж та Непрацездатність для перевірки та наступного аналізу визначаємо:

- середнє: AVERAGE(число1;[число2];...);
- найменше: MIN(число1;[число2];...);
- найбільше: MAX(число1;[число2];...);
- медіана: MEDIAN(число1;[число2];...).

Отримаємо наступні дані (рис.4).

23	21	17,5	=RANDBETWEEN(25;6	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
24	22	13	=RANDBETWEEN(25;6	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
25	23	16,7	=RANDBETWEEN(25;6	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
26	24	25	=RANDBETWEEN(25;6	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
27	25	32,7	=RANDBETWEEN(25;6	=RANDBETWEEN(0;37)	=RANDBETWEEN(0;150)
28	Перевірка				
29	середнє	=AVERAGE(B3:B28)	=AVERAGE(C3:C27)	=AVERAGE(D3:D27)	=AVERAGE(E3:E27)
30	найменше	=MIN(Дані_зп!\$B\$3:\$B\$28)	=MIN(Дані_зп!\$C\$3:\$C\$27)	=MIN(Дані_зп!\$D\$3:\$D\$28)	=MIN(Дані_зп!\$E\$3:\$E\$28)
31	найбільше	=MAX(Дані_зп!\$B\$3:\$B\$28)	=MAX(Дані_зп!\$C\$3:\$C\$27)	=MAX(Дані_зп!\$D\$3:\$D\$28)	=MAX(Дані_зп!\$E\$3:\$E\$28)

Рисунок 4 – Визначення допоміжних даних для перевірки та порівняння

4) будуємо діаграму «коробка з вусами» для даних ЗП, Вік, Стаж та Непрацевдатність:

Вставлення – Усі діаграми – Розмах та обираємо відповідні властивості (рис. 5).

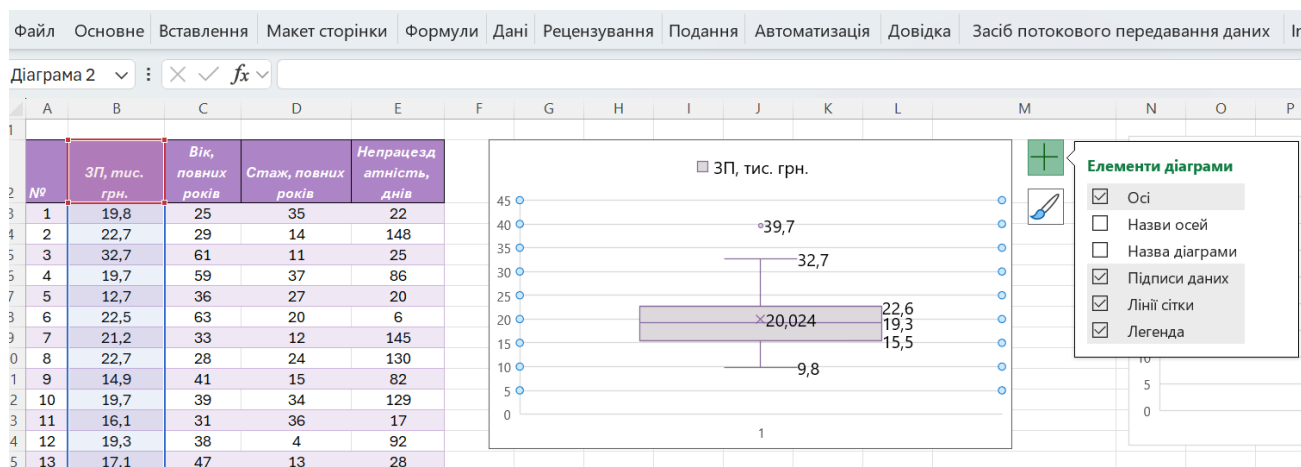


Рисунок 5 – Побудова діаграми «коробка з вусами» для даних ЗП

5) проводимо аналіз даних та робимо відповідні висновки.

**Зауваження.** При аналізі зробити порівняння середнього значення з медіаною, враховуючи що середнє значення та медіана є мірами розташування центру розподілу. Однак на медіану (на відміну від середнього значення) наявність екстремальних спостережень має менший вплив. У прикладі для ЗП (заробітної платні робітників фірми) (рис. 5) медіана 19,3 тис. грн. є меншою за середнє 20,024 тис. грн. значення, а різниця між ними вказує на наявність спостережень з екстремальними значеннями, які збільшують середнє значення. Маємо кілька спостережень з екстремальними значеннями, які збільшують середнє значення (викид у 39,7 тис. грн.) і є одиночним випадком, та не входить навіть у область від 9,8 та 32,7 тис. грн., що покриває 95 % від розкиду заробітної платні робітників). Також можна зазначити, що половина усіх робітників мають заробітну платню від 15,5 до 22,6 тис. грн. (сторони прямокутника) і є зміщеним у сторону нижньої межі (саме це) заробітної платні, що є недоліком.

### 3. Виконати Завдання 2

**Завдання 2.** Запропонувати зміни у даних ЗП та заповнити новий стовпчик ЗП\_зміни, за яких побудована діаграма «коробка з вусами», суттєво відрізнятиметься для стовпці ЗП\_базова (див. завдання 1).

Рекомендації щодо виконання завдання 2 містять наступні кроки:

- 1) дублюємо дані з файлу ЗП.xlsx на другий аркуш та розглядаємо як ЗП\_базова;
- 2) дублюємо дані з ЗП\_базова у стовпчику справа ЗП\_зміни та проводимо модифікацію даних (рис. 6);
- 3) будуємо діаграму «коробка з вусами» для даних ЗП\_зміни;

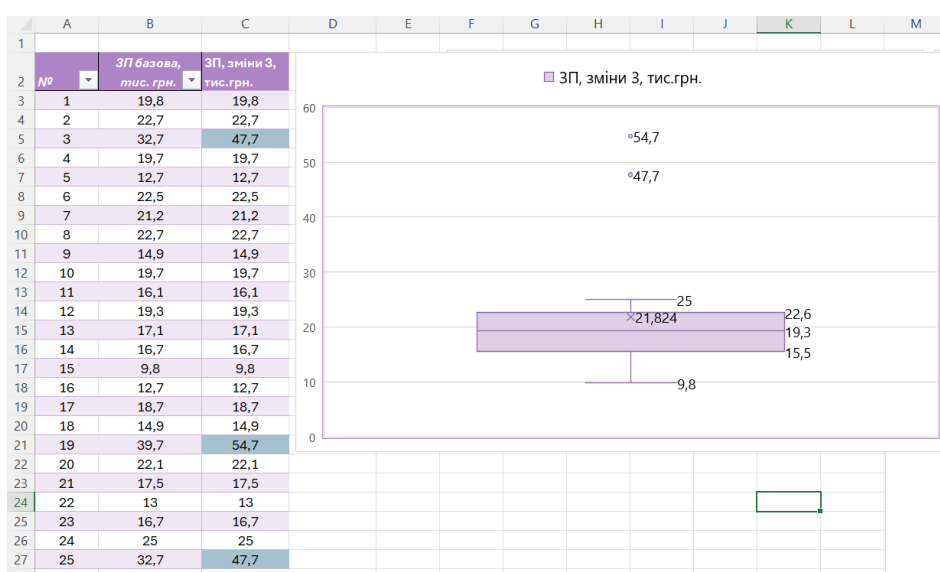


Рисунок 6 – Діаграма «коробка з вусами» для модифікованих даних ЗП\_зміни

- 4) проводимо порівняльний аналіз даних ЗП\_базова та зміни і робимо відповідні висновки.

**Зауваження.** Наведений приклад модифікації даних, а саме: у ЗП\_базова для отримання ЗП\_зміни (рис. 6) збільшено на 15 тис. грн. трьом робітникам (керівникам) з найвищої заробітною платнею.

Порівняльний аналіз базових та модифікованих даних дозволяє зробити наступні висновки: середнє значення заробітної платні фірми становить 24,824 тис. грн. (зросла на 1,8 тис. грн. завдяки збільшенню на 15 тис. грн. заробітної платні трьом керівникам), а медіана при цьому залишилася на тому ж рівні 19,3 тис. грн. як і той факт, що половина усіх робітників мають заробітну платню від 16,7 до 22,6 тис. грн. (сторони прямокутника) і є зміщеним у сторону нижньої межі 9,8 тис. грн. заробітної платні, що залишається недоліком.

Таким чином, середнє значення є більш чуттєвим до наявності екстремальних значень в даних, на медіану ж ці зміни не вплинули, що



надає можливість рекомендувати цю величину до розгляду на противагу з середнім значенням і діаграми «коробка з вусами» для компактного і наочного представлення порядкової статистики.

#### **4. Підготувати звіт за виконаними Завданнями 1, 2 та завантажити на платформу Moodle**

**Рекомендації.** При формуванні звіту дотримуватися наступного: звіт повинен бути повним, чітким і логічно завершеним та має включати титульний аркуш (додаток 1), умову завдання, наведення основних етапів його виконання (використані інструменти та скриншоти з аркуша MS Excel), правильні розрахунки та аналіз результатів із зазначенням відповідних висновків. На платформу Moodle у відповідну секцію звіт завантажується у форматі .docx або .pdf, а розроблений розрахунковий модуль у MS Excel у форматах .xls, .xlsx.

#### **ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ**

1. Яким чином в MS Excel обчислити медіану?
2. Яким чином в MS Excel обчислити середнє (середнє арифметичне)?
3. Яким чином в MS Excel провести рандомне заповнення даних в заданому діапазоні ?
4. Які дані дозволяє отримати діаграма розмаху - «коробка з вусами» ?
5. Чи завжди значення вусів відповідають найменшому та найбільшому значенню даних?
6. Чи позначається на діаграмі розмаху середнє значення?
7. Яким чином можна інтерпретувати висоту «коробки з вусами»?
8. Що представляють собою викиди?

## 2.2 Лабораторна робота №2 «Побудова регресійної моделі»

**Мета:** Лабораторна робота №2 має на меті ознайомлення студентів із основними регресійними моделями, особливостями їх побудови та аналізу у віртуальному навчальному середовищі, створеному на платформі Phet.colorado.edu. Особлива увага приділяється побудові кореляційного поля із відображенням відповідної ліній регресії та дослідженню змін коефіцієнта регресії від змін вибірових даних, що досліджуються. Ця лабораторна робота спрямована на розвиток практичних умінь аналізувати дані, обирати вид відповідної регресійної залежності та перевіряти узгодженість побудованої моделі.

### Хід і результати роботи:

1. **Ознайомитись з наведеним теоретичним матеріалом за темою.**

Статистичний аналіз є основою для обробки даних, яка включає побудову та аналіз невідомих показників математичних моделей, що відображають можливості встановлення певних статистичних закономірностей функціонування систем різних типів. До основних інструментів статистичного аналізу відносять кореляційний та регресійний аналізи. Під **кореляційним зв'язком** або **кореляцією** між випадковими величинами розуміють наявність статистичного або ймовірнісного зв'язку між ними, а під **кореляційним аналізом** - сукупність методів його виявлення. Кореляційний аналіз зазвичай застосовують для формалізованого подання математичних моделей зв'язків між окремими компонентами системи (економічної, соціальної, виробничої, тощо) або окремими процесами, що відбуваються в ній.

В багатьох прикладних задачах ставиться задача виявити залежність між двома ознаками  $X$  та  $Y$  одного і того ж об'єкту (явища, процесу), або між певними ознаками різних об'єктів (явищ, процесів). Якщо ці ознаки мають кількісне вимірювання, і, виходячи з характеристик об'єкту (системи), теоретично можна вказати, що ознака  $Y$  залежить від ознаки  $X$ , тоді незалежну змінну  $X$  вважають факторною ознакою (фактором), а залежну змінну  $Y$  – результативною ознакою. Якщо кожному значенню фактору  $X$  відповідає єдине значення результативної ознаки  $Y$ , то говорять, що між ними існує функціональний зв'язок  $Y=f(X)$ .

При дослідженнями економічних чи соціальних об'єктів головною задачею є встановлення виду цієї функціональної залежності, що за дотримання певних умов дозволяє здійснити прогнозування досліджуваних процесів. **Регресійним аналізом** називають математико-статистичний апарат, що дозволяє встановити вид кореляційної залежності, а відповідну функцію, яка описує цю залежність, **рівнянням регресії**.

Регресійний аналіз включає наступні етапи:

- встановлення виду кореляційної залежності результативної ознаки  $Y$  від фактору  $X$ ;
- побудову регресійної моделі;
- перевірка статистичної значущості запропонованої моделі.

Перший етап регресійного аналізу є визначальним і забезпечується правильним вибором виду запропонованої залежності, що відповідає (найкраще моделює) емпіричним даним та може в подальшому використовуватися для прогнозування. Здійснено це може на основі графічного зображення вибірових даних у вигляді множини точок  $(x_i; y_i)$  координатної площини, що називається **кореляційним полем (полем кореляції)**.

Якщо висунуто гіпотезу про наявність залежності результативної ознаки  $Y$  від фактору  $X$  певного виду (лінійної або нелінійної), то задача полягає у визначенні параметрів обраного рівняння. Одним з найпоширеніших методів визначення параметрів рівняння регресії є **метод найменших квадратів**.

Для рівняння лінійної регресії  $y = ax + b$  необхідним є визначення двох параметрів  $a$  та  $b$ , які для незгрупованих даних визначають з системи лінійних алгебраїчних рівнянь:

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i + bn = \sum_{i=1}^n y_i. \end{cases}$$

Для рівняння квадратичної регресії  $y = ax^2 + bx + c$  необхідним є визначення трьох параметрів  $a$ ,  $b$  та  $c$ , які для незгрупованих даних визначають з системи лінійних алгебраїчних рівнянь:

$$\begin{cases} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i, \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + cn = \sum_{i=1}^n y_i. \end{cases}$$

Аналогічно можна знайти параметри кубічної, показникової, гіперболічної та інших видів регресії.

Адекватність моделі вибіровим даним можна оцінити за коефіцієнтом детермінації  $R^2$ , що показує частину варіації значень результативної ознаки, що пояснюється рівнянням регресії.

Значення коефіцієнта детермінації знаходяться в проміжку  $[0;1]$ , і

чим ближче його значення до 1, тим краще побудоване рівняння регресії пояснює поведінку результативної ознаки.

## 2. Виконати регресійний аналіз у віртуальному навчальному середовищі, створеному на платформі Phet.colorado.edu

1. Перейти за посиланням [Curve Fitting - Polynomials | Error Analysis | Data - PhET Interactive Simulations](https://phet.colorado.edu/sims/html/curve-fitting/latest/curve-fitting_all.html)
2. Запустити симуляцію та виконати наступні завдання (рис. 7).

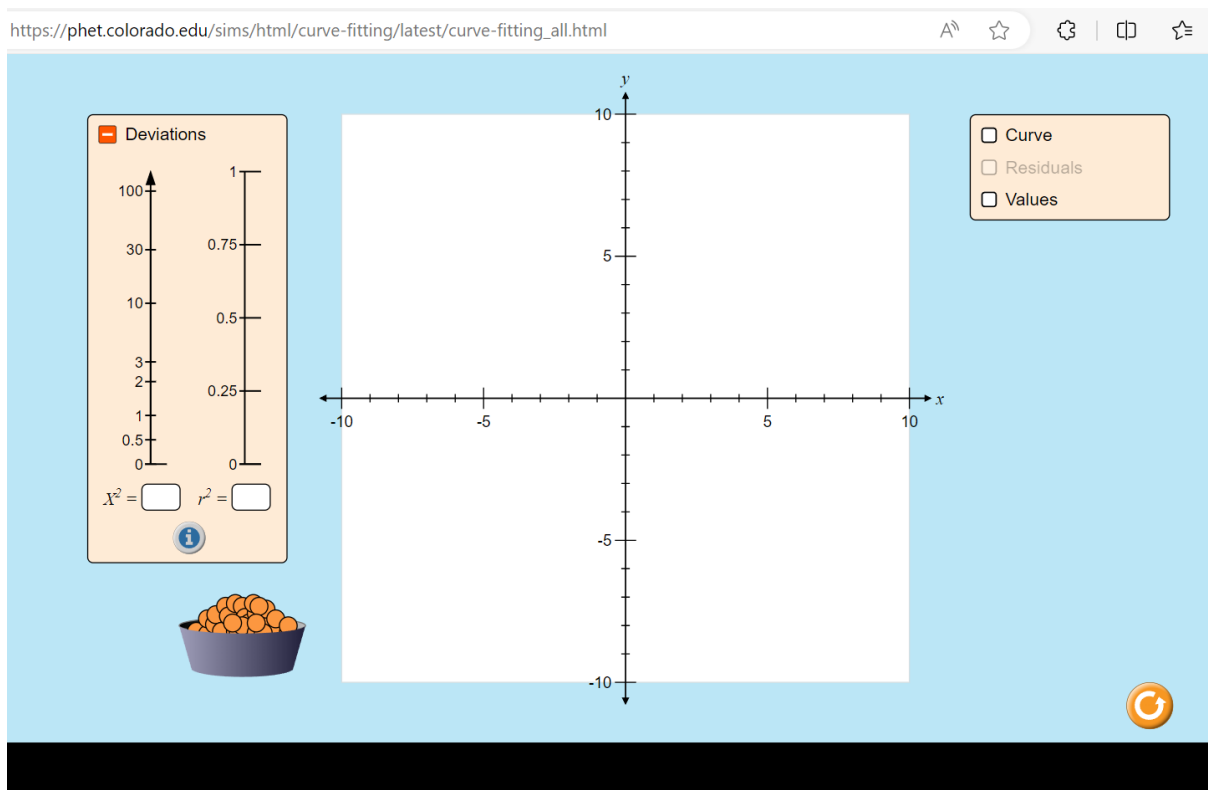


Рисунок 7 - Приклад скриншоту екрану початку роботи з симуляцією

### Завдання. Порівняльний регресійний аналіз.

1) Побудувати кореляційне поле з 20 кульок таким чином, щоб вид регресійної моделі був лінійним з високою узгодженістю ( $0,8 < r^2 < 1$ ). Отримати дані щодо квадратичної та кубічної регресійних моделей. Заповнити таблицю отриманих даних та зробити прогноз для значення  $x_{i+1} = x_{\max} + 0,1$  за усіма видами регресійних моделей, зробити висновки.

Рекомендації щодо виконання пункту 1) містять наступні кроки:

- утворити кореляційне поле з 20 кульок з  $\Delta y = 0$  та вказанням даних  $(x_i; y_i)$  для отримання інформації щодо  $x_{\max}$  та зробити скриншот екрану (рис. 8);

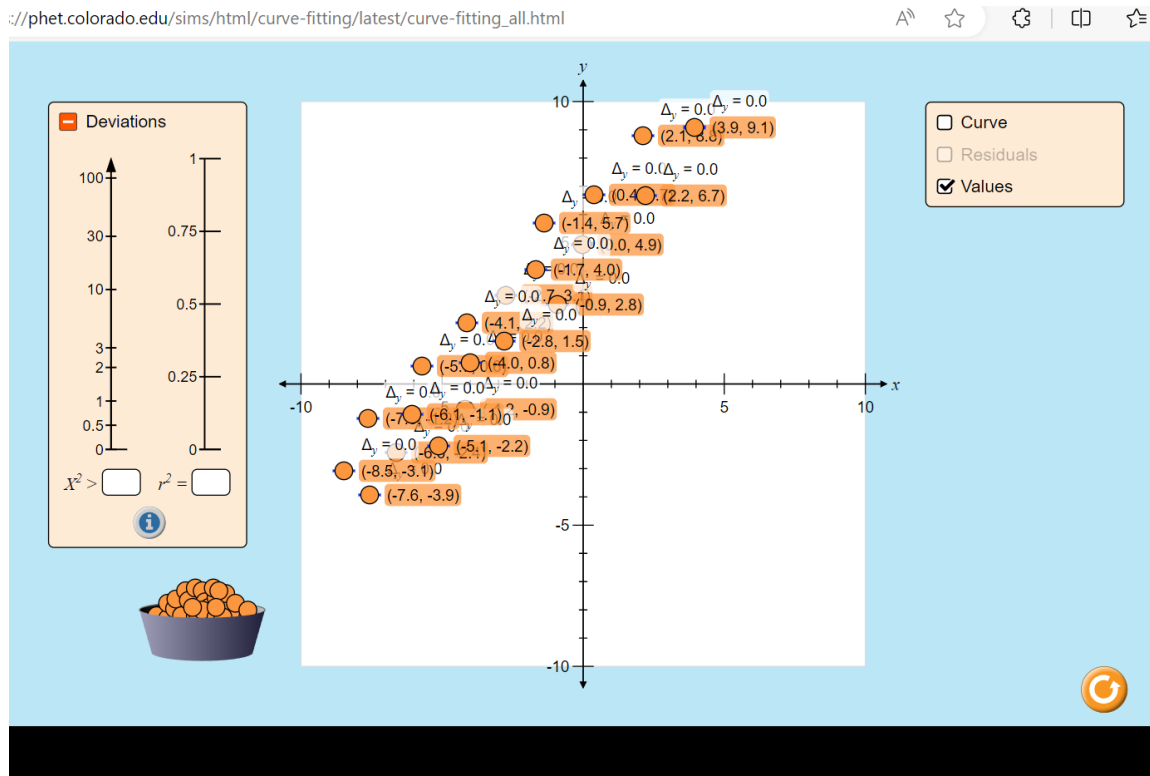


Рисунок 8 - Приклад скриншоту екрану побудови кореляційного поля

- побудувати лінійну регресійну модель та зробити скриншот екрану (рис. 9);

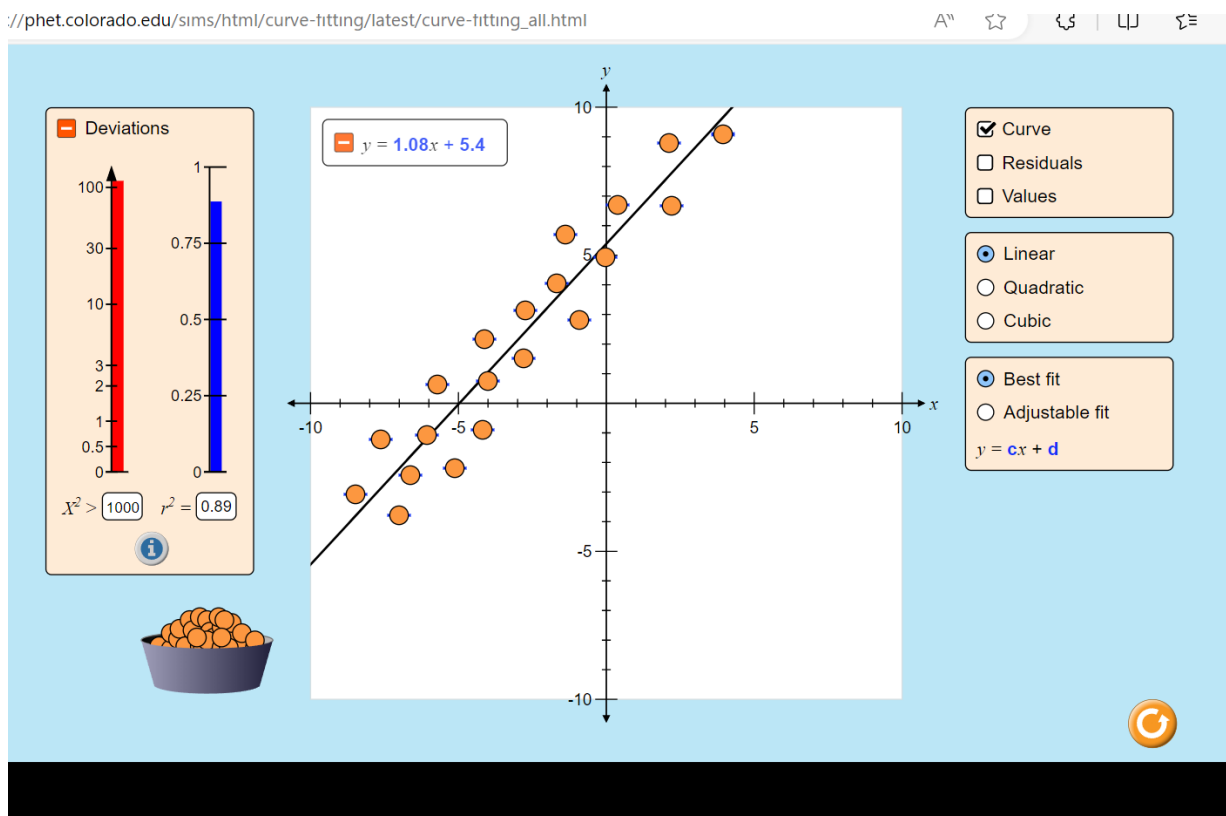


Рисунок 9 - Приклад скриншоту екрану побудови лінійної регресії

- побудувати квадратичну регресійну модель та зробити скриншот екрану (рис. 10);

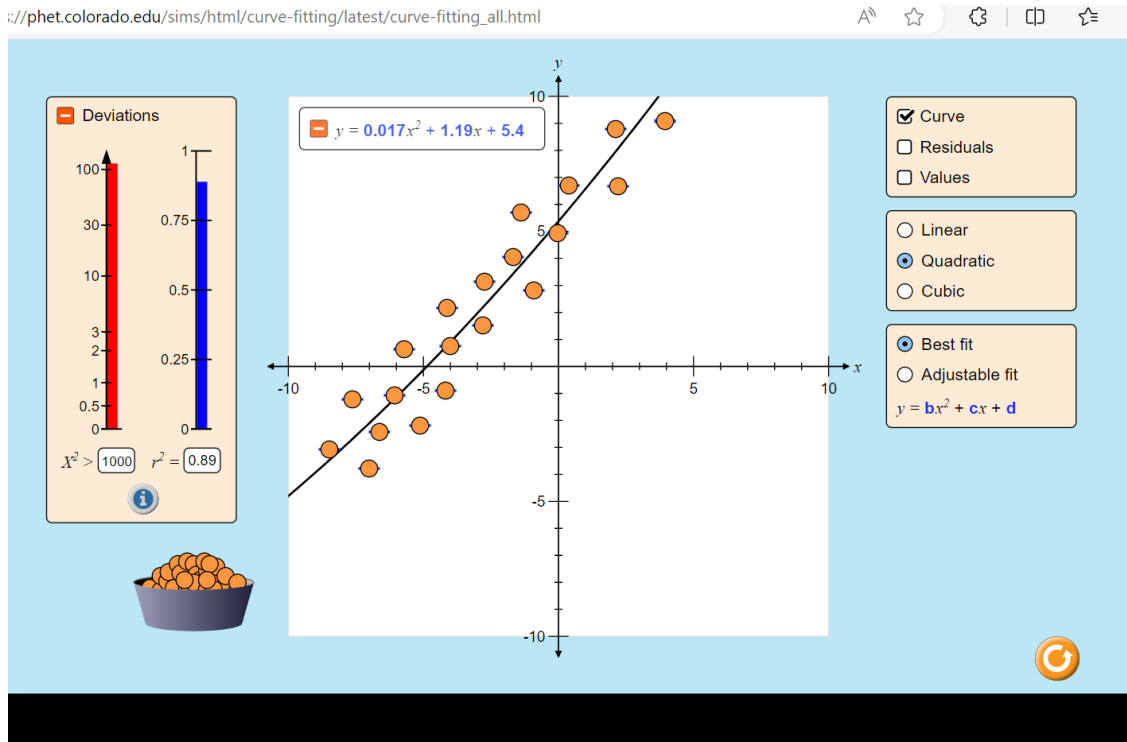


Рисунок 10 - Приклад скриншоту екрану побудови квадратичної регресії

- побудувати кубічну регресійну модель та зробити скриншот екрану (рис. 11);

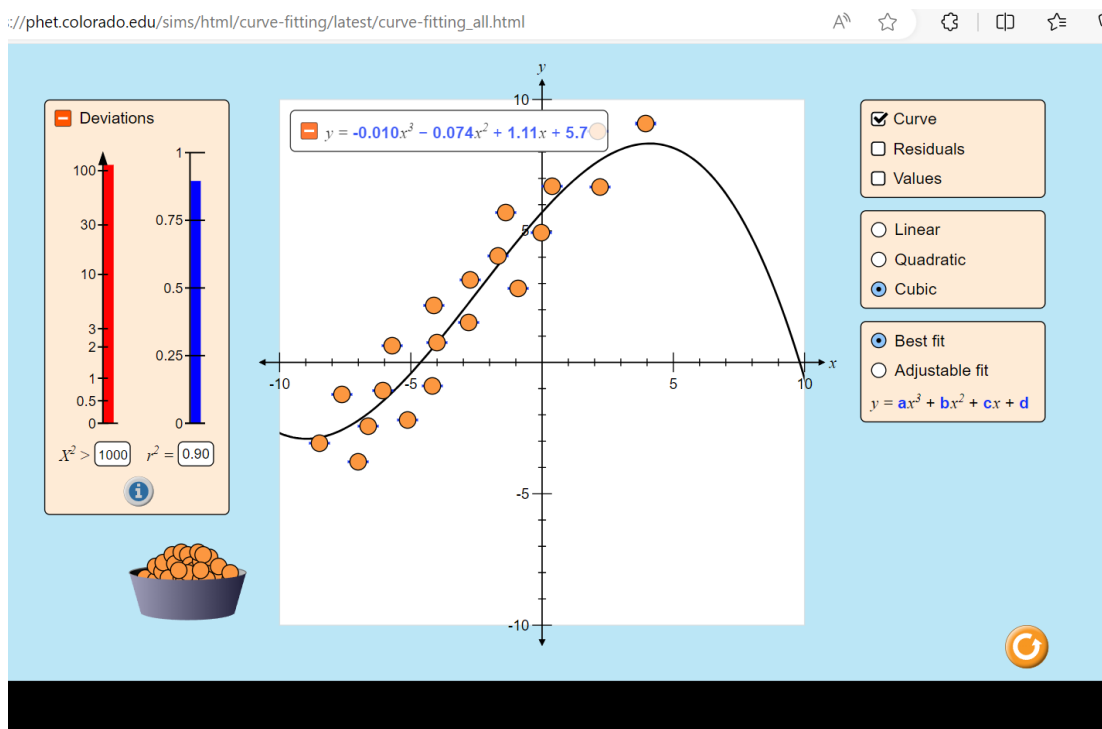


Рисунок 11 - Приклад скриншоту екрану побудови кубічної регресії

- заповнити таблицю отриманих даних та зробити прогноз для значення  $x_{i+1}=x_{\max}+0.1$  за усіма видами регресійних моделей та зробити висновки.

Таблиця 2 – Розрахункові дані

№	Вид	Рівняння	R <sup>2</sup>	Прогноз y(4)
1	лінійна	$y = 1,08x + 5,4$	0,89	9,72
2	квадратична	$y = 0,017x^2 + 1,19x + 5,4$	0,89	10,432
3	кубічна	$y = -0,01x^3 - 0,074x^2 + 1,11x + 5,7$	0,9	8,316

**Зауваження.** Маємо високу узгодженість усіх побудованих моделей, а саме  $r^2 = 0,89$  та  $r^2 = 0,9$  вказують на той факт, що 89% варіації результативної ознаки пояснюється рівняннями лінійної та квадратичної регресії та 90% - кубічної регресії. Прогноз у точці  $x_{21}=x_{\max}+0,1=3,9+0,1=4$  має суттєві відмінності за різними регресійними моделями.

2) Модифікувати кореляційне поле з 20 кульок таким чином, щоб при додаванні трьох додаткових кульок від регресійної моделі описувався квадратичною або кубічною (на Ваш розсуд) з високою узгодженістю ( $0,8 < r^2 < 1$ ) та недостатньою для лінійної ( $r^2 < 0,6$ ). Заповнити таблицю отриманих нових даних та зробити прогноз для значення  $x_{i+1}=x_{\max}+0,1$  з попереднього пункту за обраною нелінійною та лінійними регресіями, зробити висновки.

Рекомендації щодо виконання пункту 2) містять наступні кроки:

- модифікувати кореляційне поле з 20 кульок з  $\Delta y=0$  шляхом додавання трьох додаткових кульок, побудувати кубічну регресію та зробити скриншот екрану (рис. 12);

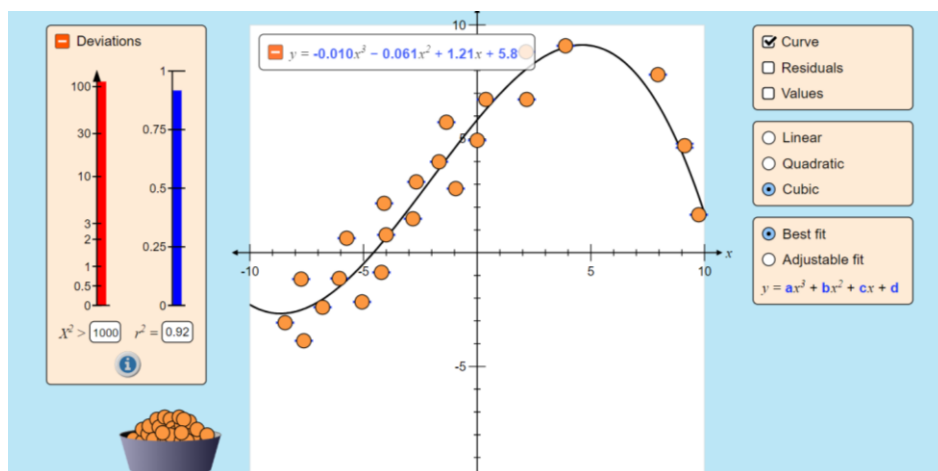


Рисунок 12 - Приклад скриншоту екрану побудови нової кубічної регресії

- побудувати лінійну регресію для модифікованих даних та зробити скриншот екрану (рис. 13);

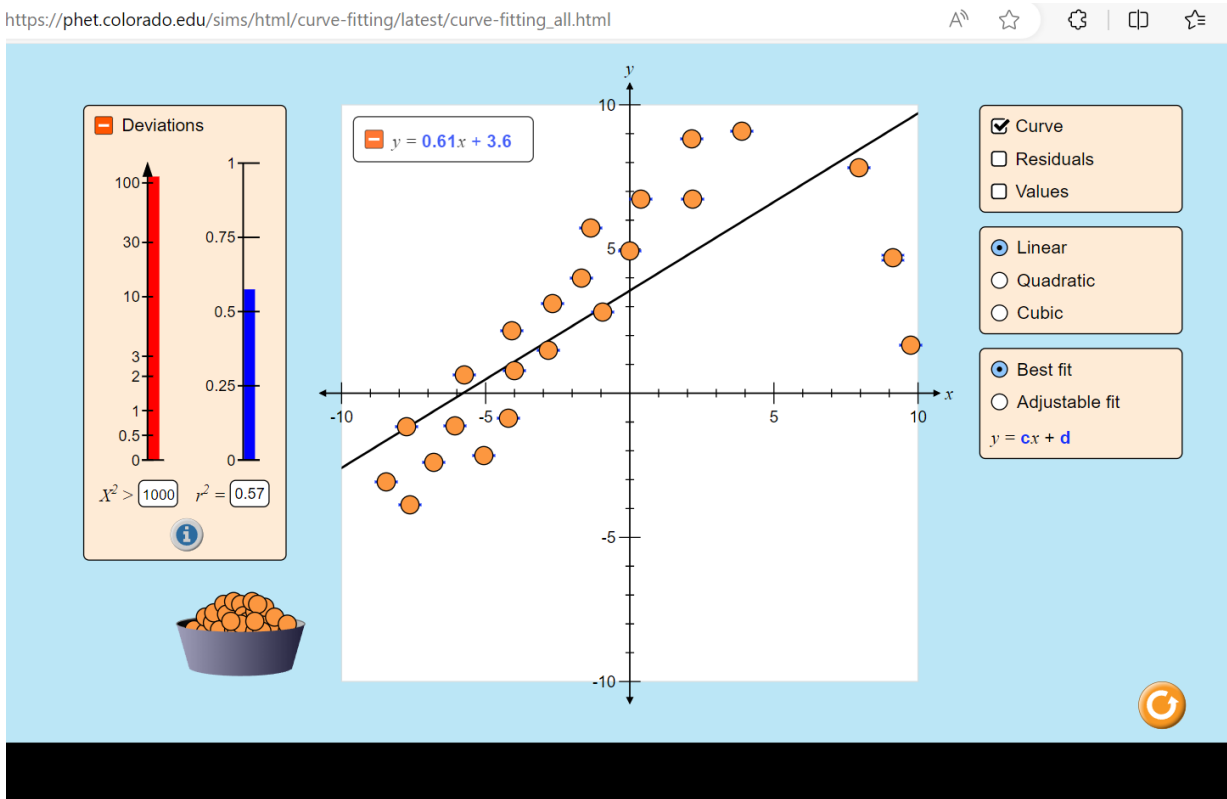


Рисунок 13 - Приклад скриншоту екрану побудови нової лінійної регресії

- заповнити таблицю отриманих даних та зробити прогноз для значення  $x_{i+1} = x_{\max} + 0.1$  за двома видами регресійних моделей та зробити висновки.

Таблиця 3 – Розрахункові дані

№	Вид	Рівняння	R <sup>2</sup>	Прогноз y(4)
1	кубічна	$y = -0,01x^3 - 0,061x^2 + 1,21x + 5,8$	0,92	9,024
2	лінійна	$y = 0,61x + 3,6$	0,57	6,04

**Зауваження.** Маємо високу узгодженість побудованої кубічної моделі, а саме  $r^2 = 0,92$  та низьку  $r^2 = 0,57$  для лінійної. Для модифікованих даних 92% варіації результативної ознаки пояснюється рівняннями кубічної регресії та тільки 57% - для лінійної регресії. Прогноз у точці  $x_{21} = x_{\max} + 0,1 = 3,9 + 0,1 = 4$  має ще більш суттєві відмінності за різними регресійними моделями. Додавання або зміни даних (навіть незначні) можуть суттєво змінювати вид моделі, яка більш інформативно описує залежність результативної ознаки від факторної, що також суттєво впливає на отримання достовірності прогнозу.



### **3. Підготувати звіт за виконаним завданням та завантажити на платформу Moodle**

**Рекомендації.** При формуванні звіту дотримуватися наступного: звіт повинен бути повним, чітким і логічно завершеним та має включати титульний аркуш (додаток 1), умову завдання, наведення основних етапів його виконання (використані інструменти та скриншоти з віртуального навчального середовища, створеного на платформі Phet.colorado.edu), правильні розрахунки та аналіз результатів із зазначенням відповідних висновків. На платформу Moodle у відповідну секцію звіт завантажується у форматі .docx або .pdf.

#### **ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ**

1. Що розуміють під кореляцією (кореляційним зв'язком) між випадковими величинами?
2. В чому полягає регресійний аналіз
3. Які основні етапи включає регресійний аналіз?
4. На що вказує високий коефіцієнт детермінації та як його можна інтерпретувати?
5. Скількома параметрами визначається квадратична регресія?
6. Який вигляд має рівняння лінійної регресії?
7. В яких межах знаходиться значення коефіцієнта детермінації?

## ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бишевец Н. Г., Омецинська Н. В., Юсипів Т. В. Теорія ймовірностей та математична статистика з використанням табличного процесора MS EXCEL : навч. посіб. Одеса : Гельветика, 2021. 233 с.
2. Веригіна І. В., Островська О. В., Проскурін Д. П. Теорія ймовірностей та математична статистика. Збірник задач : навчальний посібник для здобувачів ступеня бакалавра за технічними спеціальностями. Київ : КПІ ім. Ігоря Сікорського, 2019. 48 с. URL: <https://ela.kpi.ua/handle/123456789/27822>.
3. Горват А. А., Молнар О. О., Мінькович В. В. Методи обробки експериментальних даних з використанням MS EXCEL : навчальний посібник. Ужгород : "Видавництво УжНУ Говерла". 2020. 182 с. URL: <https://www.researchgate.net/publication/339145835> .
4. Інтерактивні симуляції для природничих наук і математики : веб-сайт. URL: <https://phet.colorado.edu/uk/> (дата звернення: 20.05.2024).
5. Теорія ймовірностей і математична статистика = Theory of Probability and Mathematical Statistics : навчальний посібник для студентів спеціальностей 121 «Інженерія програмного забезпечення», 126 «Інформаційні системи та технології» / уклад.: Є. О. Покровський, С. Є. Покровський, О. В. Савчук. Київ : КПІ ім. Ігоря Сікорського, 2021. 231 с. URL: <https://ela.kpi.ua/handle/123456789/41905>.
6. Алілуйко А. М., Дзюбановська Н. В., Єрмоменко В. О., Мартинюк О. М., Шинкарик М. І. Практикум з теорії ймовірностей та математичної статистики. Навчальний посібник для студентів економічних спеціальностей. Тернопіль : Підручники і посібники, 2018. 352 с.
7. Kaidan V., Velychko V., Fedorenko E., Kaidan N. The use of computer modeling in the educational process based on the example of studying Coulomb's law. *Journal of Physics: Conference Series, Volume 2871, XVI International Conference on Mathematics, Science and Technology Education (ICon-MaSTEd 2024)*. Kryvyi Rih, Ukraine. 2024. Vol. 2871. № 01201. DOI: 10.1088/1742-6596/2871/1/012014.



## ДОДАТОК

### Приклад титульного листа

ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА»

Кафедра природничо-наукових та загальноінженерних дисциплін

Лабораторна робота №...

з навчальної дисципліни  
«Теорія ймовірностей та математична статистика»

Виконав:

здобувач академічної групи

(шифр групи)

(прізвище та ініціали)

Керівник:

д-р техн.наук, доцент Грудкіна Н.С.

Запоріжжя, 2024



*Навчально-методичне видання*

**Грудкіна Наталія Сергіївна**

# **ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА**

**методичні рекомендації до виконання  
лабораторних робіт**

Самостійне електронне мережеве видання

Публікується в авторській редакції