УДК 378.65.011.56 doi: 10.20998/2220-4784.2021.02.05

О. П. ПРІЩЕНКО, Н. В. ЧЕРЕМСЬКА, Т. Т. ЧЕРНОГОР

ПОБУДОВА МАТЕМАТИЧНИХ МОДЕЛЕЙ ЗА ДОПОМОГОЮ МЕТОДІВ КОРЕЛЯЦІЙНОГО І РЕГРЕСІЙНОГО АНАЛІЗУ

У статті розглядається побудова математичний моделей за допомогою методів кореляційного і регресійного аналізу при визначенні функціональної залежності між величинами. При проведенні експерименту часто доводиться стикатися з необхідністю встановлення взаємозалежності між двома або кількома величинами з метою отримання емпіричної формули. У деяких випадках це виявляється простим завданням, тому що ці зв'язки практично наочні або заздалегідь відомі. Як правило, встановити взаємозв'язок між різними показниками, чинниками і ознаками, далеко не тривіальна задача. Виникає необхідність використання деякої гіпотези в вигляді функціональної залежності. Іншими словами, необхідно замінити цю функціональну залежність досить простим математичним виразом. Таким математичним виразом може бути лінійне рівняння або многочлен. Для того щоб, використовуючи дані експерименту, визначити таку математичну або функціональну залежність між змінними, застосовують методи кореляційного і регресійного аналізів. Кореляційний аналіз дає відповідь на статистичну гіпотезу про відсутність або наявність зв'язку між змінними з деякою наперед заданою довірчою ймовірністю. Визначення функціональної залежності між різними величинами по їх експериментальним значенням здійснюється за допомогою регресійного аналізу. В його основі лежить широко відомий метод найменших квадратів. Пропонуючи ту чи іншу рівняння регресії, дослідник визначає як саме існування залежності між змінними, так і математичний її вид. Регресійний аналіз розглядає зв'язок між залежною величиною і незалежними змінними. Цей зв'язок є за допомогою математичної моделі, тобто рівняння, яке пов'язує залежну і незалежні змінні. Обробка експериментальних даних при використанні кореляційного і регресійного аналізу дає нам можливість побудувати статистичну математичну модель у вигляді рівняння регресії. Таким чином, методи кореляційного і регресійного аналізів тісно пов'язані між собою.

Ключові слова: кореляційний аналіз, регресійний аналіз, функціональна залежність, апроксимація, математичний вираз, лінійне рівняння, коефіцієнт кореляції.

Вступ.

При вивченні різних об'єктів дослідження в лабораторних або виробничих умовах виникає необхідність встановлення найбільш ймовірних взаємозв'язків і взаємозалежностей між двома або більше змінними. Іноді це буває просто, оскільки зв'язок легко виявляється або заздалегідь відома з яких-небудь теоретичних передумов. Однак набагато частіше виявлення таких зв'язків є надзвичайно складним завданням [1–3].

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями.

Дослідники стикаються з необхідністю введення деякої гіпотези про характер зв'язку у вигляді функціональної залежності, тобто Апроксимації її деяким відносно простим математичним виразом, наприклад, лінійним рівнянням або многочленом.

Для пошуку таких математичних функціональних або структурних залежностей між двома або більше змінними (по накопиченим експериментальним даним) дуже корисні методи кореляційного і регресійного аналізів.

Кореляційний аналіз дає відповідь на статистичну гіпотезу про відсутність або наявність зв'язку між змінними з деякою наперед заданою довірчою ймовірністю.

Визначення функціональної залежності між різними величинами (в найпростішому y випадку від x) по їх експериментальним значенням здійснюється за допомогою регресійного аналізу. В його основі лежить широко відомий метод найменших квадратів.

Пропонуючи те чи інше рівняння регресії, дослідник визначає як саме існування залежності між змінними, так і математичний її вид [4–11].

Викладання основного матеріалу досліджень.

1. Коефіцієнт кореляції.

Поняття «кореляція» походить від латинського слова correlatio – співвідношення.

У математичній статистиці під кореляцією розуміється будь-який зв'язок між двома або кількома змінними випадковими величинами.

Про наявність чи відсутність зв'язку між двома випадковими величинами в першому наближенні судять по кореляційному полю (рис. 1).

Позитивна кореляція між випадковими величинами характеризує таку імовірнісну залежність, при якій зі зростанням однієї з них інша в середньому також буде зростати (рис. 1, а). При негативній кореляції зі зростанням однієї випадкової величини інша в середньому убуває (рис. 1, в).

Кореляційний аналіз дозволяє оцінювати тісноту зв'язку різних параметрів або факторів, що впливають на процес. У загальному вигляді задача виявлення і оцінки сили зв'язку в математичній статистиці не вирішена. Існують лише показники, що дозволяють оцінювати ті чи інші сторони випадкового зв'язку. З них найважливішим є коефіцієнт кореляції.

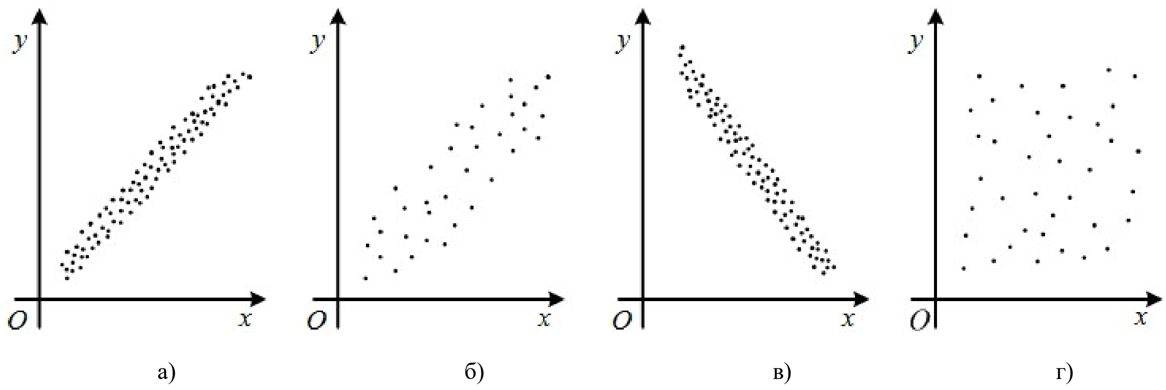

Рис. 1 – Кореляційне поле двох випадкових величин: при різному характері кореляції:
а – сильна позитивна кореляція; б – слабка позитивна кореляція;
в – сильна негативна кореляція; г – нульова кореляція

Якщо допустити, що проведено m випробувань і при кожному відзначалися значення двох випадкових величин x_i і y_i, то вибірковий коефіцієнт кореляції буде дорівнювати:

$$r_{xy} = \frac{\sum_{i=1}^{m}(x_i - \overline{x})(y_i - \overline{y})}{(m-1)S_x S_y},$$

де S_x, S_y – стандартні відхилення.

Коефіцієнт кореляції є безрозмірну величину, модуль якої не перевищує одиниці $r_{xy} \le 1$. Для незалежних величин x і y $r_{xy} = 0$ і означає відсутність лінійної залежності. Рівність $\left|r_{xy}\right| = 1$ свідчить про наявність лінійної залежності між величинами, при якій кожному x значенню відповідає тільки одне значення y.

Коефіцієнт кореляції є досить грубою оцінкою тісноти зв'язку, що має сенс лише при лінійній залежності між параметрами. Навіть при високому коефіцієнті кореляції не можна зробити достовірних висновків про наявність статистичного зв'язку, оскільки одночасне регулювання параметрів призводить до їх штучної (помилковою) кореляції.

Точно також малий коефіцієнт кореляції не завжди є наслідком відсутності зв'язку між параметрами, а може бути результатом нелінійного характеру зв'язку.

2. Побудова математичної моделі за результатами експерименту.

Математична модель представляє собою залежність між параметрами і факторами процесу, отриману теоретично або експериментально.

Математична модель компактна і зручна для дослідження і управління реальним процесом.

Застосування математичної моделі дозволяє:

- вибрати оптимальний технологічний режим процесу;
- скоротити план дослідних робіт при розробці технології виробництва;
- створити систему управління процесом.

При експериментальному вивченні функціональної залежності y від x виробляють ряд вимірювань величини y при різних значеннях величини x.

Результати можуть бути представлені у вигляді таблиць або графіків. Завдання полягає в аналітичному представленні шуканої функціональної залежності, тобто у підборі формули, яка описує результати екс-перімента.

Особливість завдання полягає в тому, що наявність випадкових помилок вимірювання (або, як кажуть, наявність «шуму» в експерименті) унеможливлює підбір такої формули, яка точно описала б все досвідчені значення.

Так, наприклад, послідовне з'єднання всіх експериментальних точок відрізками прямої дасть залежність зображену на рисунку 2.

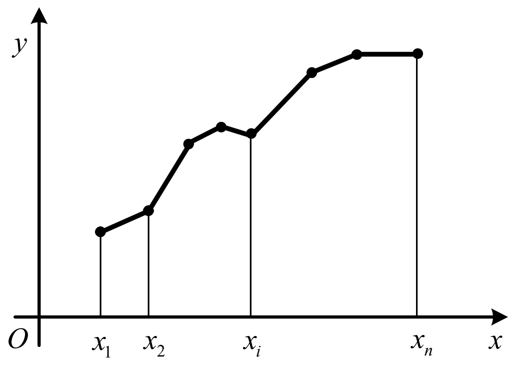

Рис. 2 – Приклад розкиду експериментальних точок

Тому графік шуканої функції не повинен проходити через всі точки, а повинен по можливості згладжувати «шум».

Згладжування «шуму» буде тим точніше і надійніше, чим більше буде вироблено експериментів.

Перш за все, дослідник повинен вибрати вид кривої, для якої він буде шукати апроксимуюче рівняння. Нижче наводиться для довідок кілька найбільш поширених видів апроксимуючих кривих і відповідних їм рівнянь:

пряма лінія $y = b_0 + b_1 x$;

квадратна парабола $y = b_0 + b_1 x + b_2 x^2$;

парабола n-й степені $y = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n$;

гіпербола $y = 1/(b_0 + b_1 x)$ або $1/y = b_0 + b_1 x$;

логарифмічна крива $y = b_0 + b_1 \lg x$.

Звичайно, можуть знайти застосування і криві багатьох інших видів. Щоб вирішити, яку апроксимацію використовувати, слід вивчити кореляційне поле і порівняти розташування експериментальних точок з формою кривих, що відповідають різним рівнянням.

Форма деяких з них показана на рисунку 3.

Форму зв'язку математично не вибирають. Можна лише перевірити, наскільки адекватна інтуїтивно обрана дослідником форма зв'язку експериментальних точок.

Таким чином, завдання зводиться до визначення параметрів рівняння b_o, b_1, і так далі. Крім того, вид формули може бути відомий заздалегідь з теоретичного опису об'єкта, що моделюється.

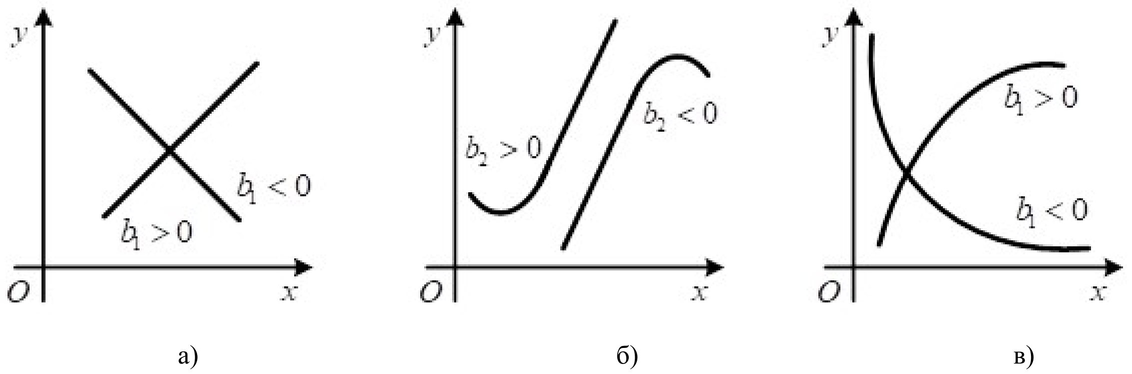

Рис. 3. Форми різних регресійних кривих:
а) $y = b_0 + b_1 x$ за умови $b_1 > 0$, $b_1 < 0$;
б) $y = b_0 + b_1 x + b_2 x^2$, за умови $b_2 > 0$, $b_2 < 0$;
в) $y = b_0 + b_1 \lg x$, за умови $b_1 > 0$, $b_1 < 0$.

Позначимо обрану функціональну залежність через рівняння:

$$y = f(x_i; \tilde{\beta}_i),$$

де $i = 1, 2, \ldots, n$; $\tilde{\beta}_i$ – експериментальні оцінки параметрів рівняння. Це рівняння називається рівнянням регресії, а математичне очікування цієї функціональної залежності

$$M(y_i) = M\left[f(x_i; \tilde{\beta}_i)\right] = \sum_{i=1}^{n} p_i \cdot f(x_i; \tilde{\beta}_i)$$

називається регресією.

Слово «регресія» ввів в статистику Френсіс Гальтон (1822-1911) – англійський математик. Знаходження оцінок параметрів і дослідження одержуваних моделей отримали назву регресійного аналізу.

Отримане рівняння прийнято називати емпіричної формулою, а оцінку параметрів функції *f*-оцінкою параметрів емпіричної формули.

3. Визначення рівняння регресії методом найменших квадратів.

Одним з найбільш поширених методів регресійного аналізу є метод найменших квадратів, перший виклад якого було дано французьким математиком Андре Марі Олександром (1752–1833) і далі розроблено німецьким вченим Карлом Фрідріхом Гаусом (1777–1855).

Для найпростішого, однофакторного випадку, коли маємо дві змінні випадкові величини y і x_1, функція відгуку або рівняння регресії має вигляд:

$$y = b_0 + b_1 x_1.$$

Це рівняння прямої лінії. Мета визначення параметрів емпіричної формули – обчислення невідомих коефіцієнтів b_0 і b_1. Якби всі експериментальні точки лежали строго на прямій лінії, то для кожної з них була б справедлива рівність:

$$y_i - b_0 - b_1 x_{1i} = 0,$$

де $i = 1, 2, \ldots, n$ – номер досліду.

На практиці це рівність порушується, замість нього доводиться писати:

$$y_i - b_0 - b_1 x_{1i} = \Delta y_i ,$$

де Δy_i – різницю між експериментальним і розрахунковим за рівнянням регресії значеннями y в i-й експериментальній точці. Цю величину іноді називають нев'язкої.

Позначаючи через $\hat{y}$ розрахункове значення функції (за рівнянням регресії), отримаємо:

$$\hat{y} = b_0 + b_1 x_{1i} ,$$
$$y_i - \hat{y}_i = \Delta y_i .$$

Нев'язка Δy_i виникає з двох причин:

1) помилка експерименту;
2) непридатність моделі.

Причому ці причини змішані і без додаткової інформації не можна вирішити, яка з причин переважає. Для цього використовують методи оцінки помилок досвіду і придатності моделі (адекватності моделі).

Завжди прагнуть знайти такі коефіцієнти регресії, при яких нев'язка будуть мінімальні. Ось одна з можливих записів:

$$U = \sum_{i=1}^{n} \Delta y_i^2 = \min . \qquad (1)$$

вона призводить до методу найменших квадратів

Можливий і метод найменших кубів:

$$\sum_{i=1}^{n} \left| \Delta y_i^3 \right| = \min .$$

Можливий і метод, в якому мінімізується сума модулів (абсолютних величин) нев'язок:

$$\sum_{i=1}^{n} \left| \Delta y_i \right| = \min .$$

Умова (1), покладена в основу методу найменших квадратів вважають найбільш вдалим компромісом.

При постановці експерименту зазвичай проводиться більше дослідів, ніж число невідомих коефіцієнтів. Тому система лінійних рівнянь

$$\begin{cases} \Delta y_1 = y_1 - b_0 - b_1 x_{11}, \\ \Delta y_2 = y_2 - b_0 - b_1 x_{12}, \\ ---------- \\ \Delta y_i = y_i - b_0 - b_1 x_{1i} \end{cases}$$

часто виявляється суперечливою.

Якщо все експериментальні точки лежать на прямій, то тільки тоді система стає визначеною і має єдине розв'язання щодо b_0 та b_1.

Метод найменших квадратів (МНК) породжує властивість, яка робить визначеною будь-яку систему рівнянь. Рівняння регресії має вигляд:

$$y = b_0 + b_1 x_1 .$$

У ньому два невідомих b_0 і b_1 значення. Застосовуючи МНК, перепишемо рівняння (1) інакше:

$$U = \sum_{i=1}^{n} \Delta y_i^2 = \sum_{i=1}^{n} \left(y_i - b_0 - b_1 x_{1i} \right)^2 = \min .$$

Відомо, що мінімум деякої функції, якщо він існує, досягається при одночасній рівності нулю частинних похідних по всіх невідомих, тобто

$$\frac{\partial U}{\partial b_0} = 0, \ \frac{\partial U}{\partial b_1} = 0 .$$

Обчислимо частинні похідні і розв'яжемо систему рівнянь:

$$\begin{cases} -2\sum_{i=1}^{n} \left(y_i - b_0 - b_1 x_{1i} \right) = 0, \\ -2\sum_{i=1}^{n} \left(y_i - b_0 - b_1 x_{1i} \right) x_{1i} = 0. \end{cases}$$

Розкриємо дужки, проведемо перетворення:

$$\begin{cases} n b_0 + \sum_{i=1}^{n} x_i b_1 = \sum_{i=1}^{n} y_i , \\ \sum_{i=1}^{n} x_{1i} b_0 = \sum_{i=1}^{n} x_{1i}^2 b_1 = \sum_{i=1}^{n} y_i x_{1i} . \end{cases}$$

Ця система називається системою нормальних рівнянь. Формули для обчислення b_0 і b_1 зручно знаходити за допомогою метода Крамера.

Остаточні формули мають вигляд:

$$b_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_{1i}^2 - \sum_{i=1}^{n} y_i x_{1i} \sum_{i=1}^{n} x_{1i}}{n\sum_{i=1}^{n} x_{1i}^2 - \left(\sum_{i=1}^{n} x_{1i} \right)^2},$$

$$b_1 = \frac{n\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_{1i} - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_{1i}}{n\sum_{i=1}^{n} x_{1i}^2 - \left(\sum_{i=1}^{n} x_{1i} \right)^2}.$$

Подивимося тепер, як обчислюються суми, що входять в ці формули. Для виконання розрахунків складають матрицю результатів експерименту, як показано в таблиці 1.

Таблиця 1. Результати експеримента

Номер експерименту	x_1	y	${x_{1i}}^2$	yx_{1i}	$*y^2$	$*x_1+y$	$*\left(x_1+y\right)^2$
1	x_{11}	y_1	${x_{11}}^2$	y_1x_{11}	${y_1}^2$	$x_{11}+y_1$	$\left(x_{11}+y_1\right)^2$
2	x_{12}	y_2	${x_{12}}^2$	y_2x_{12}	${y_2}^2$	$x_{12}+y_2$	$\left(x_{12}+y_2\right)^2$
…	…	…	…	…	…	…	…
n	x_{1n}	y_n	${x_{1n}}^2$	y_nx_{1n}	${y_n}^2$	$x_{1n}+y_n$	$\left(x_{1n}+y_n\right)^2$
$\sum\limits_{i=1}^{n}$	$\sum\limits_{i=1}^{n}x_{1i}$	$\sum\limits_{i=1}^{n}y_i$	$\sum\limits_{i=1}^{n}{x_{1i}}^2$	$\sum\limits_{i=1}^{n}y_ix_{1i}$	$\sum\limits_{i=1}^{n}{y_i}^2$	–	$\sum\limits_{i=1}^{n}\left(x_{1i}+y_i\right)^2$

Видно, що обчислень зроблено більше, ніж потрібно для розрахунку b_0 і b_1. Вони позначені зірочкою. Ці «зайві» дані потрібні для перевірки правильності розрахунків. Можливі два способи перевірки:

1) перевіряється умова

$$\sum_{i=1}^{n}\left(x_{1i}+y_i\right)^2=\sum_{i=1}^{n}{x_{1i}}^2+2\sum_{i=1}^{n}y_ix_{1i}+\sum_{i=1}^{n}{y_i}^2;$$

2) перевірку можна провести і по наступному рівнянню

$$\overline{y}=b_0+b_1\overline{x}_{1i}.$$

Другий спосіб є найбільш повним. З його допомогою перевіряється не тільки обчислення суми, але і коефіцієнтів. На практиці використовуються обидві перевірки.

Завдання. В ході тестування хромель-алюмелеві термопари за постійними (реперних) точок температур кристалізації металів Pb (327,5 °C), Zn (419,6 °C), Al (660,0 °C). Отримані відповідно наступні дані термоерс: 12,1 мВ; 16,0 мВ; 26,1 мВ. Необхідно розрахувати рівняння регресії лінійної залежності температури від величин термоерс термопари.

Розв'язання. Складемо таблицю 2 для розрахунку коефіцієнтів регресії:

Таблиця 2. Розрахунок коефіцієнтів регресії завдання

Номер експерименту	x	y	x^2	$x\cdot y$
1	12,1	327,5	146,4	3962,75
2	16,0	419,6	256,0	6713,6
3	26,1	660,0	681,2	17226,0
$\sum$	54,2	1407,1	1083,6	27902,3

Після виконання розрахунків за таблицею 2 можна визначити:

$$b_0=\frac{1407,1\cdot 1093,6-27902,3\cdot 54,2}{3\cdot 1083,6-54,2^2}=39,75.$$

$$b_1=\frac{3\cdot 27902,3-1407,1\cdot 54,2}{3\cdot 1083,6-54,2^2}=23,76.$$

Таким чином, можна отримати рівняння регресії:

$$y=39,75+23,76x,$$

яке дозволяє розрахувати величину температури (°C) за показаннями термопари.

Коефіцієнти регресії можна знайти за допомогою програми **Excel**.

Порядок дій для обчислення обох коефіцієнтів регресії однаковий.

Невелика відмінність полягає в тому, що в діалозі **Майстер функцій** в категорії **Статистичні** для знаходження лінійного коефіцієнта b_1 вибираємо функцію **НАКЛОН**, а для вільного члена рівняння b_0 – **ОТРЕЗОК** (рис. 4 та рис. 5).

На екрані з'являться відповідно діалоги:

Аргумент функцій – **НАКЛОН**
і
Аргумент функцій – **ОТРЕЗОК**.

У відповідні поля вводимо з робочого листа діапазон значень y і x.

У діалозі **Аргумент функцій** з'явиться значення визначеного коефіцієнта, яке після натискання кнопки **ОК** переноситься в задану комірку робочого аркуша. Програма **Excel** видає більш точні значення коефіцієнтів регресії, ніж розраховані вручну.

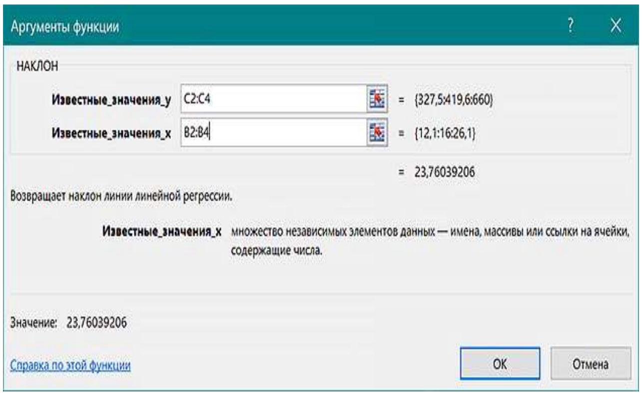

Рис. 4. Діалог «Аргумент функцій» – НАКЛОН

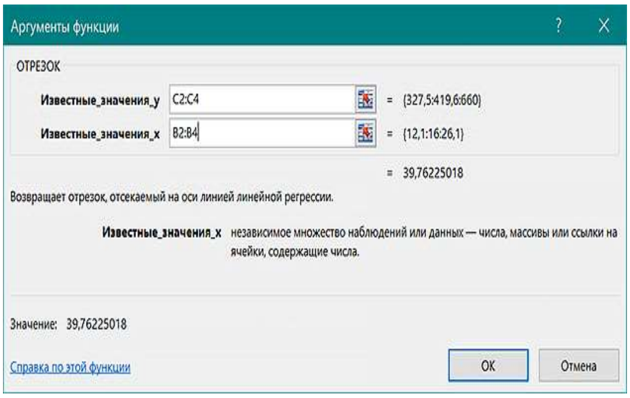

Рис. 5. Діалог «Аргумент функцій» – ОТРЕЗОК

На екрані з'являться відповідно діалоги **Аргумент функцій – НАКЛОН** і **Аргумент функцій – ОТРЕЗОК**. У відповідні поля вводимо з робочого листа діапазон значень y і x.

У діалозі **Аргумент функцій** з'явиться значення визначеного коефіцієнта, яке після натискання кнопки **OK** переноситься в задану комірку робочого аркуша. Програма **Excel** видає більш точні значення коефіцієнтів регресії, ніж розраховані вручну.

Програма Microsoft Office Excel 2016 дозволяє одночасно провести дисперсійний, кореляційний и регресійний аналіз з оцінкою значущості коефіцієнтів отриманого рівняння регресії [12–18].

Висновки та перспективи подальшого розвитку даного напрямку.

У статті розглядається побудова математичний моделей за допомогою методів кореляційного і регресійного аналізу при визначенні функціональної залежності між величинами. Як приклад наводиться завдання з хімічним змістом. Найчастіше, при вивченні курсу вищої математики доводиться розв'язувати задачі загального характеру. Але для студентів Навчально-наукового інституту хімічних технологій та інженерії більший інтерес представляють завдання, які безпосередньо пов'язані з їх спеціальністю. Таким чином, розглядаючи завдання, подібні наведеної в даній статті, ми підвищимо інтерес і мотивацію майбутніх фахівців до вивчення даного матеріалу.

І хоча математики та інженери хіміки-технологи мислять зовсім по різному, ті випадки, коли їм вдається досягти взаємодії, призводять до появи нетривіальних результатів і сприяють збагаченню обох цих наук. Завдяки таким діям можливо досягнути більш конкретизованих результатів за деякими питаннями з теми інноваційної діяльності.

Таким чином, заняття зі студентами та їх самостійна робота формують вміння при формулюванні висновків з проведеної роботи. При цьому у студентів виробляються необхідні навички: користування комп'ютерною технікою з метою виявлення закономірностей процесів та методів дослідження; проведення патентного пошуку та реалізації отриманих результатів; публічний захист наукової розробки, аналітичний компетентнісний аналіз наукової та прикладної частини і т.і. [19–23].

Список літератури

1. Высшая математика в примерах и задачах : уч. пособ. : Т. 2 / Ю.Л. Геворкян, Л.А. Балака, С.С. Габриелян и др. ; под ред. Ю.Л. Геворкяна. – Х.: Підручник НТУ «ХПІ», 2011. – 376 с.
2. Вища математика в прикладах і задачах : у 2 т. Т. 2 : Диференціальне та інтегральне числення функцій багатьох змінних. Диференціальні рівняння та ряди : навч. посіб. / Л.В. Курпа, Н.О. Кириллова, Г.Б. Лінник та ін. ; за ред. Л.В. Курпи. Х: НТУ «ХПІ», 2009. – 432 с.
3. Геворкян Ю.Л. Краткий курс высшей математики : учеб. пособ. : в 2 ч. Ч. 2 / Ю.Л. Геворкян, А.Л. Григорьев, Н.А. Чикина. – Х. : Підручник НТУ «ХПІ», 2011. 476 с.
4. Диференціальні рівняння та їх застосування : н.-мет. посіб. / Пріщенко О.П., Черногор Т.Т. – Х. : НТУ «ХПІ», 2017. – 88 с.
5. Ерёмин В. В. Математика в химии. – 2-е изд., испр. / В.В. Ерёмин. – М. : МЦНМО, 2016. – 64 с.
6. Збірник розрахунково-графічних завдань з вищої математики : у 2 ч. Ч. 1 / Н.О. Чікіна, І.В Антонова, Л.О. Балака та ін. ; за ред. Н.О. Чікіної. – Х. : Підручник НТУ «ХПІ», 2012. – 224 с.
7. Пріщенко О. П., Черногор Т. Т. Аналіз прикладів застосування диференціальних рівнянь в хімічній та харчовій технології // Вісник НТУ «ХПІ». – Харків : НТУ «ХПІ», 2018. – № 40 (1316). – с. 39 – 45.
8. Пріщенко О.П., Черногор Т.Т., Бухкало С.І. Деякі особливості проведення кореляціійного аналізу Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: тези доповідей XXVII міжнародної науково-практичної конференції MicroCAD-2019, 15-17 травня 2019 р.: у 4 ч. Ч. II. / за ред. проф. Сокола Є.І. – Харків: НТУ «ХПІ». – с.320.
9. Пріщенко О.П., Черногор Т.Т. Деякі особливості проведення регресійного аналізу Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: тези доповідей XXVII міжнародної науково-практичної конференції MicroCAD-2019, 15-17 травня 2019 р.: у 4 ч. Ч. II./за ред. проф. Сокола Є.І. – Х: НТУ «ХПІ». – с. 319.
10. Скатецкий В.Г. Математические методы в химии::уч. пос. для студентов ВНЗ/ В.Г. Скатецкий, Д.В. Свиридов, В.И. Яшкин. Минск :ТетраСистемс, 2006. 368 с.
11. Тевяшев А.Д. Вища математика у прикладах та задачах : у 3 ч. Ч. 1 : Лінійна алгебра і аналітична геометрія. Диференціалтне числення функції однієї змінної : навч. пос./А.Д. Тевяшев, О.Г. Литвин. Х.: ХНУРЕ, 2002. 552 с.
12. Бухкало С.І. Деякі властивості полімерних відходів у якості сировини для енерго- і ресурсозберігаючих процесів // Інтегровані технології та енергозбереження.

– Х.: НТУ «ХПІ». 2014. – № 4. – с. 29–33.
13.Бухкало С.І. Моделі енергетичного міксу для утилізації полімерної частки ТПВ // Вісник НТУ «ХПІ». – Х.: НТУ «ХПІ». 2016. – № 19 (1191). – с. 23–32.
14.Prishchenko O. P., Chernogor T. T. Using of methods of cross-correlation and regressive analysis for determination of functional dependence between sizes // Вісник НТУ «ХПІ». Х.: НТУ «ХПІ», 2019. – №15 (1340), – с. 36 – 41.
15.S. Bukhkalo, A. Ageicheva, O. Komarova. Distance learning main trends. Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: тези доповідей XXVI міжн. н-пр. конф. MicroCAD-2018, 16-18.05 2018. Ч. II / за ред. проф. Сокола Є.І. Х:.НТУ «ХПІ», с 205.
16.Бухкало С.І., Іглін С.П. Деякі моделі дослідження структурно-хімічних змін при експлуатації полімерних виробів. Інтегровані технології та енергозбереження. Х.: НТУ «ХПІ», 2016. № 3. – С. 52–57.
17.Бухкало С.І., Білоус О.В., Демидов І.М. Розробка комплексного антиоксиданту із екстрактів листя горіху волоського та календули. Восточно-Европейский журнал передовых технологий. 2015. № 1/6(73), – с. 22–26. – Х. : Технол. центр.
18.Bukhkalo S.I., Klemeš J.J., Tovazhnyanskyy L.L., Arsenyeva O.P., Kapustenko P.O., Perevertaylenko O.Y. Eco-friendly synergetic processes of municipal solid waste polymer utilization. Chemical Engineering Transactions, Vol.70, (2018), – pp.2047–2052.
19.Prishchenko O.P., Chernogor T.T. Analysis of opportunities of analytical method of optimization in chemical technology // Вісник НТУ «ХПІ». Х.: 2020. – №5 (1359). – с. 71 – 77.
20. Пріщенко О. П., Черногор Т. Т. Використання тензорів при аналізі особливостей фізичних властивостей твердих тіл // Вісник НТУ «ХПІ». – Харків : НТУ «ХПІ», 2020. – №6 (1360). – с. 42 – 48.
21. Prishchenko O.P., Chernogor T.T. Application of elements of studying the function of one variable when solving chemical problems // Вісник НТУ «ХПІ». Х : НТУ «ХПІ», 2021. – №1 (1361). – с. 30 – 35.
22. Бухкало С.І. Удосконалювання методів оцінки знань студентів вищих навчальних закладів. Вісник НТУ «ХПІ». Х.: НТУ «ХПІ». 2014, № 16, с. 3–11.
23. Бухкало С.І. Визначення загальної технології комплексних курсових проектів. Інформаційні технології: наука, техніка, технології, освіта, здоровья: тези доповідей XXVII Міжн. н-практ. конференції (MicroCAD-2019), 15–17 мая 2019 р.: у 4 ч. Ч. II. / за ред. проф. Сокола Є.І. – Харків: НТУ «ХПІ». С. 217.

References (transliterated)

1. Vysshaja matematika v primerah i zadachah : ucheb. posob. : T. 2 / Ju.L. Gevorkjan, L.A. Balaka, S.S. Gabrieljan i dr. ; pod red. Ju.L. Gevorkjana. – Khar'kov : Pidruchnik NTU «KhPI», 2011. – 376 p.
2. Vishha matematika v prikladah i zadachah : u 2 t. T. 2 : Diferencial'ne ta integral'ne chislennja funkcij bagat'oh zminnih. Diferencial'ni rivnjannja ta rjadi : navch. posib. / L.V. Kurpa, N.O. Kirillova, G.B. Linnik ta in. ; za red. L.V. Kurpi. – Kharkiv : NTU «KhPI», 2009. – 432 p.
3. Gevorkjan Ju.L. Kratkij kurs vysshej matematiki : ucheb. posob. : Ch. 2 / Ju.L. Gevorkjan, A.L. Grigor'ev, N.A. Chikina. Khar'kiv Pidruchnik NTU «HPI», 2011. – 476 p.
4. Diferencial'ni rivnyannya ta ïh zastosuvannya : n.-met. posib. / Prishchenko O.P., CHernogor T.T. – Kh. : NTU «KhPI», 2017. – 88 p.
5. Eryomin V. V. Matematika v himii. – 2-e izd., ispr. / V.V. Eryomin. – M. : MCNMO, 2016. – 64 p.
6. Zbirnik rozrahunkovo-grafichnih zavdan' z vishchoï matematiki : CH. 1/N.O. CHikina, I.V Antonova, L.O. Balaka ta in. ; za red. N.O. CHikinoï. – Kh. : Pidruchnik NTU «KhPI», 2012. – 224 p.
7. Prishchenko O.P., Chernogor T.T. Analiz prikladiv zastosuvannya diferencial'nih rivnyan' v himichnij ta harchovij tekhnologiï/Visnik NTU «KhPI». 2018. № 40, – pp. 39–45.
8. Prishchenko O.P., Chernogor T.T., Buhkalo S.I. Deyaki osoblivosti provedennya korelyaciijnogo analizu Informacijni tekhnologiï: nauka, tekhnika, tekhnologiya, osvita, zdorov'ya: tezi dopovidej XXVII mizhnarodnoï naukovo-praktichnoï konferenciï MicroCAD-2019, CH. II. / za red. prof. Sokola Є.I. – Kharkiv: NTU «KhPI». – p.320.
9. Prishchenko O.P., CHernogor T.T. Deyaki osoblivosti provedennya regresijnogo analizu Informacijni tekhnologiï: nauka, tekhnika, tekhnologiya, osvita, zdorov'ya: tezi dopovidej HXVII mizhnarodnoï naukovo-praktichnoï konferenciï MicroCAD-2019, 15-17 travnya 2019 r.: u 4 ch. CH. II/za red. prof. Sokola Є.I. – Kh: NTU «KhPI», p. 319.
10.Skateckij V.G. Matematicheskie metody v himii : ucheb. posob. dlya studentov vuzov/V.G. Skateckij, D.V. Sviridov, V.I. YAshkin. – Minsk : TetraSistems, 2006. – 368 p.
11.Tevyashev A.D. Vishcha matematika u prikladah ta zadachah : u 3 ch. CH. 1 : Linijna algebra i analitichna geometriya. Diferencialtne chislennya funkciï odnieї zminnoï : navch. posib. / A.D. Tevyashev, O.G. Litvin. – Kharkiv : HNURE, 2002. – 552 p.
12.Bukhkalo S.I. Deyaki vlastivosti polimernih vidhodiv u yakosti sirovini dlya energo- i resursozberigayuchih procesiv // Integrovani tekhnologiï ta energozberezhennya. – Kh.: NTU «KhPI». 2014. – No. 4, – pp. 29–33.
13.Buhkalo S.I. Modeli energetichnogo miksu dlya utilizaciï polimernoï chastki TPV // Visnik NTU «KhPI». – Kh.: NTU «KhPI». 2016. – № 19 (1191), – pp. 23–32.
14.Prishchenko O.P., Chernogor T.T. Using of methods of cross-correlation and regressive analysis for determination of functional dependence between sizes/Visnik NTU «KhPI». Kh : 2019. No. 15 (1340), – pp. 36 – 41.
15.S. Bukhkalo, A. Ageicheva, O. Komarova. Distance learning main trends. Informacijni tekhnologiï: nauka, tekhnika, tekhnologiya, osvita, zdorov'ya: tezi dopovidej XXVI mizhn. n-pr. konf. MicroCAD-2018, 16-18 travnya 2018r. CH. II.Kh :.NTU «KhPI», p. 205.
16.Bukhkalo S.I., Iglin S.P. Deyaki modeli doslidzhennya strukturno-himichnih zmin pri ekspluataciï polimernih virobiv. Integrovani tekhnologiï ta energozberezhennya. H.: NTU «KhPI», 2016. No. 3, – pp. 52–57.
17.Buhkalo S.I., Bilous O.V., Demidov I.M. Rozrobka kompleksnogo antioksidantu iz ekstraktiv listya gorihu volos'kogo ta kalenduli. Vostochno-Evropejskij zhurnal peredovyh tekhnologij. 2015. No. 1/6(73), – pp. 22–26.
18.Bukhkalo S.I., Klemeš J.J., Tovazhnyanskyy L.L., Arsenyeva O.P., Kapustenko P.O., Perevertaylenko O.Y. Eco-friendly synergetic processes of municipal solid waste polymer utilization. Chemical Engineering Transactions, Vol.70, (2018), – pp.2047–2052.
19.Prishchenko O.P., Chernogor T.T. Analysis of opportunities of analytical method of optimization in chemical technology // Visnik NTU «KhPI». – Kh :2020. №5(1359), pp. 71–77.
20.Prishchenko O. P., Chernogor T. T. Vikoristannya tenzoriv pri analizi osoblivostej fizichnih vlastivostej tverdih til/ Visnik NTU «KhPI», Kh : 2020. No.6 (1360), pp. 42 – 48.
21.Prishchenko O.P., Chernogor T.T. Application of elements of studying the function of one variable when solving chemical problems // VIsnik NTU «KhPI». – Kharkiv : NTU «KhPI», 2021. – No. 1 (1361), – pp. 30 – 35.
22. Bukhkalo S.I. Udoskonaljuvannja metodiv ocinki znan' studentiv vishhih navchal'nih zakladiv. Visnik NTU «KhPI». Kh.: NTU «KhPI». 2014, No. 16, pp. 3–11.
23. Buhkalo S.I. Viznachennja zagal'noï tehnologiï kompleksnih kursovih proektiv. Informacijni tehnologiï: nauka, tehnika, tehnologiï, osvita, zdorov'ja: tezi dopovidej XXVII Mizhn. n-prakt. konferenciï (MicroCAD-2019), 15–17 maja 2019 r.: u 4 ch. Ch. II. / za red. prof. Sokola Є.I. – Kh: NTU «KhPI», pp. 217.

Надійшла (received) 30.09.2021

Відомості про авторів / Сведения об авторах / About the Authors

Пріщенко Ольга Петрівна (Прищенко Ольга Петровна, Prishchenko Olga Petrivna) – старший викладач кафедри вищої математики, Національний технічний університет «Харківський політехнічний інститут», м. Харків, Україна; ORCID: http://orcid.org/0000-0003-0530-2131; e-mail: priolga2305@gmail.com

Черемська Надія Валентинівна (Черемская Надежда Валентиновна, Cheremskaya Nadezhda Valentinovna) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», м. Харків, Україна; тел.: (050) 225-15-44; e-mail: cheremskaya66@gmail.com

Черногор Тетяна Тимофіївна (Черногор Татьяна Тимофеевна, Chernogor Tetyana Timofiyivna) – старший викладач кафедри вищої математики, Національний технічний університет «Харківський політехнічний інститут», м. Харків, Україна; ORCID: http://orcid.org/0000-0002-7823-7628; e-mail: tatyanachernogor54@gmail.com

О. П. ПРИЩЕНКО, Н. В. ЧЕРЕМСКАЯ, Т. Т. ЧЕРНОГОР

ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРИ ПОМОЩИ МЕТОДОВ КОРРЕЛЯЦИОННОГО И РЕГРЕССИОННОГО АНАЛИЗА

В статье рассматривается построение математической модели с помощью методов корреляционного и регрессионного анализа при определении функциональной зависимости между величинами. При изучении различных объектов исследования в лабораторных или производственных условиях возникает необходимость установления наиболее вероятных взаимосвязей и взаимозависимостей между двумя или более переменными. Иногда это бывает просто, поскольку связь легко обнаруживается или заранее известна из каких-либо теоретических предпосылок. Однако гораздо чаще выявление таких связей между различными показателями, факторами, признаками является чрезвычайно сложной задачей. Исследователи сталкиваются с необходимостью введения некоторой гипотезы о характере связи в виде функциональной зависимости, т.е. аппроксимации ее некоторым относительно простым математическим выражением, например, линейным уравнением или многочленом. Для поиска таких математических функциональных или структурных зависимостей между двумя или более переменными весьма полезны методы корреляционного и регрессионного анализов. Корреляционный анализ дает ответ на статистическую гипотезу об отсутствии или наличии связи между переменными с некоторой наперед заданной доверительной вероятностью. Определение функциональной зависимости между различными величинами по их экспериментальным значениям осуществляется с помощью регрессионного анализа. В его основе лежит широко известный метод наименьших квадратов. Предлагая то или иное уравнение регрессии, исследователь определяет как само существование зависимости между переменными, так и математический ее вид. Регрессионный анализ рассматривает связь между зависимой величиной и независимыми переменными. Эта связь представляется с помощью математической модели, т. е. уравнения, которое связывает зависимую и независимые переменные. Обработка экспериментальных данных при использовании корреляционного и регрессионного анализа дает нам возможность построить статистическую математическую модель в виде уравнения регрессии. Таким образом, методы корреляционного и регрессионного анализов тесно связаны между собой.

Ключевые слова: корреляционный анализ, регрессионный анализ, функциональная зависимость, аппроксимация, математическое выражение, линейное уравнение, коэффициент корреляции.

O. P. PRISHCHENKO, N. V. CHEREMSKAYA, T. T. CHERNOGOR

CONSTRUCTION OF MATHEMATICAL MODELS USING THE METHODS OF CORRELATION AND REGRESSION ANALYSIS

The article discusses the construction of a mathematical model using the methods of correlation and regression analysis in determining the functional relationship between the quantities. When conducting an experiment, it is often necessary to establish the interdependence between two or more quantities in order to obtain an empirical formula. In some cases, this is a simple task, because these connections are almost obvious or known in advance. As a rule, to establish the relationship between different indicators, factors and characteristics is not a trivial task. There is a need to use some hypothesis in the form of functional dependence. In other words, it is necessary to replace this functional dependence with a fairly simple mathematical expression. Such a mathematical expression can be a linear equation or a polynomial. In order to use such experimental data to determine such a mathematical or functional relationship between variables, the methods of correlation and regression analysis are used. Correlation analysis provides an answer to the statistical hypothesis of the absence or presence of a relationship between variables with some predetermined confidence probability. Determination of the functional dependence between different values on their experimental values is carried out using regression analysis. It is based on the well-known method of least squares. Proposing one or another regression equation, the researcher determines both the very existence of the relationship between variables and its mathematical form. Regression analysis considers the relationship between the dependent quantity and non-dependent variables. This relationship is represented using a mathematical model, that is, an equation that connects the dependent and independent variables. Processing of experimental data using correlation and regression analysis allows us to build a statistical mathematical model in the form of a regression equation. Thus, the methods of correlation and regression analysis are closely related.

Key words: correlation analysis, regression analysis, functional dependence, approximation, mathematical expression, linear equation, correlation coefficient.