

**ОСОБЛИВОСТІ ПОБУДОВИ ЕМПІРИЧНИХ РЕГРЕСИВНИХ МОДЕЛЕЙ  
ПРИ МОДЕЛЮВАННІ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ ПІДПРИЄМСТВ  
АВТОМОБІЛЬНОГО ТРАНСПОРТУ**

*Г. В. Заяць, к. т. н., доц.*

**Ключові слова:** технологічний процес, модель, регресія, кореляція, імовірність, адекватність, точність

**Постановка проблеми.** Технологічні процеси на підприємствах автомобільного транспорту, як і машинобудівного виробництва, особливо процеси відновлення, дуже складні. Досі відсутні адаптовані аналітичні моделі закономірності процесів щодо систем масового обслуговування рухомого складу автотранспортного підприємства, оптимізації процесу заміни обладнання виробничих зон, зношування і навантаження інструменту при різних видах ремонту тощо. Тому дуже часто використовують моделі, які раніше позначали як емпіричні. Емпіричні моделі об'єктів і процесів є результатом обробки експериментальних даних про поведінку об'єкта або процесу методами математичного статистичного аналізу.

**Аналіз досліджень та публікацій.** Аналіз публікацій [1; 3; 4; 7 – 9] дозволив зробити висновок про необхідність загальної систематизації напрацювань для можливості аналізу їх результатів із використанням сучасного програмного забезпечення (MathLab&Simulink, MathCad, Statistica та ін.).

**Метою** досліджень є визначення найбільш оптимального методу моделювання технологічних процесів за критерієм адекватності моделі, похибки результатів, економічності розрахунків.

**Виклад основного матеріалу.** Для побудування моделей об'єктів за результатами експериментальних досліджень використовують математичний апарат регресійного і кореляційного аналізу. Терміни «регресія» і «кореляція» були вперше запропоновані наприкінці XIX сторіччя до вживання Ф. Гальтоном і К. Пірсоном, які вивчали взаємозалежності зросту і ваги людей різного віку і виявили необхідність уведення показників вказаної залежності, які б надавали зв'язок між досліджуваними характеристиками людини, але не визначали б одне одного однозначно. В даний час «регресія» і «кореляція» – основні поняття статистики.

Основне завдання кореляційного аналізу – виявлення значущості зв'язку між значеннями різних випадкових величин. Залежність між випадковими величинами, при яких одному значенню однієї величини (аргументу) відповідає одне або декілька цілком певних значень іншої величини, називається, відповідно, однозначною або багатозначною функціональною залежністю [6]. Залежність між величинами, при якій кожному значенню однієї величини відповідає з відповідною ймовірністю безліч можливих значень інших, називають імовірнісною, яка, в свою чергу, може бути стохастичною чи статистичною.

Прикладами кореляційного зв'язку (для технологічних процесів у машинобудуванні) є залежності між межами міцності і текучості сталі певної марки, між погрішностями розміру і погрішністю форми поверхні деталі, між температурою випробування і міцністю матеріалу тощо.

Математичний апарат регресійного аналізу дозволяє:

- оцінити невідомі параметри запропонованої до дослідження регресійної моделі;
- перевірити статистичну значущість параметрів моделі;
- перевірити адекватність моделі;
- оцінити точність моделі.

Вид регресійної моделі пропонує виконавач, який повинний враховувати:

- фізичну суть досліджуваного об'єкта;
- характер експериментального матеріалу;
- наявність і можливість аналізу апріорної інформації.

Найпростішим для моделювання є об'єкт, у якого один вхідний і один вихідний фактор (рис. 1). Вхідний фактор характеризується дією на досліджуваний об'єкт. Наприклад, у технологічних процесах машинобудування це може бути температура, сила, час, геометричні параметри інструменту, характеристики оброблюваного й інструментального матеріалів. Вихідний параметр характеризує реакцію об'єкта на дію вхідного чинника. Вихідні фактори в технологічних процесах машинобудування – довжина пройденого інструментом шляху,

величина зносу, напруга, якість обробленої поверхні і т. п.

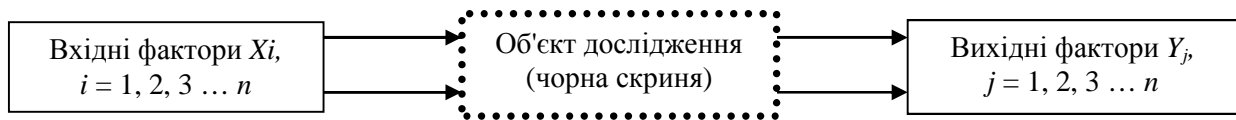


Рис. 1. Об'єкт дослідження в загальному вигляді

Для побудування емпіричної моделі необхідна наявність даних експериментальних досліджень об'єкта (наприклад, у вигляді таблиць), в яких кожному значенню вхідного фактора (X) відповідає значення вихідного фактора (Y). Пари випадкових змінних ( $X_i, Y_j$ ) підкоряються імовірнісному розподілу, графічне відображення якого наведено на діаграмі розсіяння (рис. 2), [7].

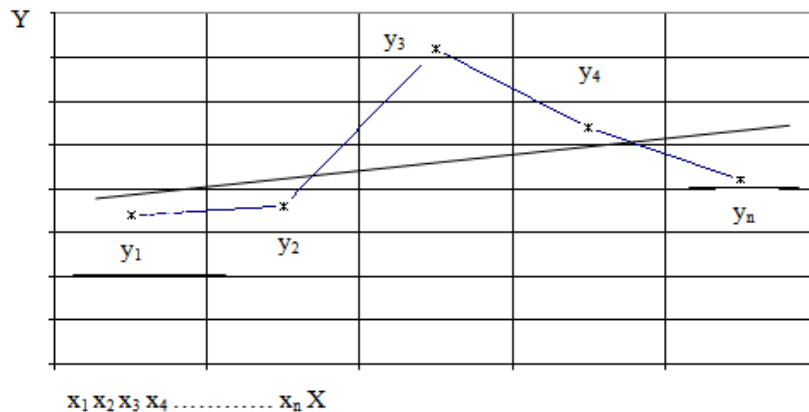


Рис. 2. Графічна інтерпретація дослідження з лінією регресії

Необхідно знайти таку криву, яка б найближче апроксимувала експериментальні точки. Для зручності подальшого дослідження об'єкта ця крива для свого опису повинна мати одну єдину формулу. Якщо ми з'єднаємо точки на графіку, то отримаємо ламану лінію, складену з декількох прямих і описувану відповідною кількістю лінійних моделей. Це вкрай незручно для дослідження. Необхідно знайти криву, що найкращим чином описує всі експериментальні точки (рис. 2). Таку криву називають регресійною кривою Y по X.

У загальному випадку крива регресії може мати будь-який вигляд, але вона не повинна мати розривів. У найпростішому випадку крива регресії має вигляд прямої лінії.

Побудування моделей і дослідження об'єкта починають із найпростіших моделей, лінійних. Лінійній моделі відповідає крива регресії у вигляді простої лінії.

Як правило, є відхилення експериментальних точок від кривої регресії, що викликано впливом неврахованих у моделі зовнішніх факторів на досліджуваний об'єкт. У моделюванні вихідний фактор називають залежною вихідною змінною, а вхідний – незалежною вхідною змінною. У процесі дослідження об'єкта вхідний фактор завжди має детермінований характер, а вихідний – випадковий.

Рівняння, яке встановлює зв'язок між випадковою залежною і детермінованою незалежною змінними, називають рівнянням регресії. Відповідно до [7], термін «рівняння регресії» є загальноприйнятим, але не коректним. Модель, побудована на основі рівняння регресії, називають регресійною моделлю. Для отримання регресійних моделей (рівнянь регресії) використовується математичний апарат регресійного аналізу.

Отже, підбір кривої регресії і регресійної моделі зазвичай починають із побудування простої прямої лінії і, відповідно, з лінійної моделі.

Відповідно до [7], при необмежено великій кількості експериментальних точок лінійна регресійна модель може бути представлена у вигляді рівнянь:

$$y = \beta_0 + \beta_1 \cdot x + \varepsilon, \quad (1)$$

$$\hat{y} = \beta_0 + \beta_1 \cdot x, \quad (2)$$

де  $\hat{y}$  – передбачені значення вихідної змінної для лінійної моделі;

$x$  – значення вхідної змінної;

$\beta_0, \beta_1$  – коефіцієнти регресії;

$\varepsilon$  – нехарактерний залишок.

Визначення коефіцієнтів регресії здійснюється на основі методу найменших квадратів, який застосовують у тих випадках, коли випадкова варіація вхідного параметра досить мала порівняно зі спостережуваним діапазоном його вимірювання [7], тобто значення вхідної змінної вважаються фіксованими. Суть методу полягає в тому, що розраховуються такі значення  $\beta_0$  і  $\beta_1$ , при яких сума квадратів відхилень зміряних величин  $y$  від передбачених  $\hat{y}$  була б мінімальною.

Для пар спостережень доцільно записати:

$$y = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i. \quad (3)$$

Відхилення зміряної величини  $y$  від передбаченої  $\hat{y}$

$$\varepsilon_i = y_i - \hat{y} = y_i - (\beta_0 + \beta_1)x_i. \quad (4)$$

Сума квадратів відхилень записується у вигляді:

$$\Delta = \sum_{i=1}^n \varepsilon_i^2 = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 \cdot x_i), \quad (5)$$

де  $\Delta$  – функція суми квадратів відстаней від експериментальних точок до прямої.

Значення  $\beta_0$  і  $\beta_1$  перебирають і підбирають такими, щоб

$$\Delta \rightarrow \min. \quad (6)$$

Похідні

$$\frac{d\Delta}{d\beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 \cdot x_i), \quad (7)$$

$$\frac{d\Delta}{d\beta_1} = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 \cdot x_i). \quad (8)$$

Найменше значення суми квадратів відхилень досягається у тому випадку, коли коефіцієнти  $\beta_0$  і  $\beta_1$  задовольняють умові [6]:

$$\frac{d\Delta}{d\beta_0} = \frac{d\Delta}{d\beta_1} = 0. \quad (9)$$

Значення  $x_i$ ,  $y_i$  є обмеженою вибіркою із загального числа станів досліджуваного об'єкта. Тому визначити можливо тільки коефіцієнти  $\beta_0$  і  $\beta_1$ , які для зручності позначимо  $b_0$  і  $b_1$  відповідно. Тоді:

$$\hat{y} = b_0 + b_1 x. \quad (10)$$

Такі моделі називають однофакторними регресійними моделями. Коефіцієнт регресії  $b_1$  визначається за формулою [7]:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (11)$$

де  $x_i$  – значення вхідного фактора під час проведення експерименту;

$y_i$  – значення вихідного чинника, відповідного до  $x_i$ ;

$\bar{x}$  – середнє значення вхідного фактора, яке може бути визначене за формулою:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad (12)$$

$\bar{y}$  – середнє значення вихідного фактора, яке може бути визначене за формулою:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}. \quad (13)$$

Коефіцієнт регресії  $b_0$  може бути визначений за залежністю:

$$b_0 = \bar{y} - b_1 \bar{x}. \quad (14)$$

В результаті перетворень отримаємо:

$$\hat{y} = \bar{y} + b_1(x_i - \bar{x}). \quad (15)$$

Адекватність отриманої регресійної моделі експериментальним даним слід оцінювати за критерієм Фішера ( $F_{розр}$ ) [7]:

$$F_{розр} = \frac{\sum_{i=1}^n \left( \hat{y}_i - \bar{y} \right)^2}{\sum_{i=1}^n \left( y_i - \hat{y}_i \right)^2}, \quad (16)$$

який є відношенням суми квадратів відхилень, зумовлених регресією, до суми квадратів відхилень щодо регресії.

При формулюванні нульової гіпотези ( $H_0: b_1 = 0$ ) розрахункове значення коефіцієнта  $F_{розр}$  порівнюється з табличним значенням  $F_{табл}(n$  і  $\alpha$ ), де  $n$  – загальна кількість експериментальних спостережень ( $x_i, y_i$ ), яка впливає на кількість ступенів свободи при визначенні критерію Фішера;  $\alpha$  – ймовірність невірної гіпотези.

Зазвичай в моделюванні використовують значення  $\alpha = 0,05; 0,01$ . При  $F_{розр} > F_{табл}$  нульова гіпотеза відкидається, модель вважається адекватною, а регресія – значущою. При  $F_{розр} < F_{табл}$  регресійна модель неадекватна і використання її для аналізу і дослідження неприпустиме. У цьому випадку необхідно знову проаналізувати апіорну інформацію, знову спланувати і провести експеримент.

Значення коефіцієнта Фішера зазвичай дані в довідниках із математичної статистики і теорії ймовірності.

Для оцінки точності регресійних моделей з однією вхідною використовується вибірковий коефіцієнт кореляції Пірсона  $P(x, y)$ , який визначається за формулою:

$$P(xy) = \frac{\sum_{i=1}^n \left( x_i - \bar{x} \right) \cdot \left( y_i - \bar{y} \right)}{\sqrt{\sum_{i=1}^n \left( x_i - \bar{x} \right)^2 \cdot \sum_{i=1}^n \left( y_i - \bar{y} \right)^2}}. \quad (17)$$

Коефіцієнт кореляції Пірсона  $P(xy)$  характеризує тісноту зв'язку між вихідною змінною  $y$  і вхідною змінною  $x$ . Значення коефіцієнта кореляції Пірсона  $P(xy)$  знаходиться в межах від -1 до +1 включно.

Можна виділити декілька окремих випадків значення коефіцієнта кореляції.

Чим вище значення  $P(xy)$ , тим тісніший зв'язок між вихідною  $y$  і вхідною  $x$  змінними, тим точніша математична модель.

Якщо модель має низьке значення  $P(xy)$ , то вона має низьку точність оцінки і прогнозу поведінки або властивостей об'єкта. Таку модель використовувати для дослідження, опису і прогнозу об'єкта не рекомендується. З декількох моделей, проаналізованих під час моделювання, для дослідження об'єкта вибирається та модель, у якій коефіцієнт кореляції  $P(xy)$  має найбільше значення.

Після розрахунку коефіцієнта кореляції проводять перевірку його значущості за допомогою критерію Стюдента. Коефіцієнт кореляції, розрахований для моделі ( $P_{розр}$ ), порівнюється з граничним значенням ( $P_{табл}$ ).

Якщо  $P_{розр} > P_{табл}$ , то  $P_{розр}$  приймається як показник щільності зв'язку і навпаки [8]. Табличні значення  $P_{табл}$  можна знайти в довідниках з теорії ймовірності і математичної статистики.

При оцінюванні стандартної помилки для передбачених значень вихідного фактора треба розуміти, що якнайкращі прогнози будуть у «центрі тяжіння» експерименту, і чим далі від «центру тяжіння», тим найменш точними будуть прогнози значень вихідного фактора [7].

Якщо в результаті розрахунку коефіцієнта кореляції  $P(xy)$  лінійна модель визнана недостатньо точною, переходять до дослідження складніших моделей: статичної ( $y = b_0 x^{b_1}$ ), експоненційної ( $y = \exp(b_0 + b_1 x)$ ), полінома ( $y = b_0 + b_1 x + b_2 x^2$ ) чи зворотної ( $y = b_0 + b_1 / x$  або

$$y = \frac{1}{b_0 + b_1 \cdot x}).$$

Поліном і зворотні моделі є лінійними за параметрами, тому для оцінки їх коефіцієнтів

регресії, кореляції і критерію адекватності можна використовувати формули (11), (14), (16), (17).

Статична й експоненціальна моделі вимагають додаткових перетворень у вигляді логарифмування.

Для побудови моделі необхідно мати дані експериментальних досліджень процесу, наведені у вигляді таблиці, де кожній комбінації значень вхідних чинників відповідає значення вихідного фактора (табл.).

Таблиця

Результати експериментальних досліджень

№ експерименту	Вхідний фактор					Вихідний фактор
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	.....	X <sub>m</sub>	
1	X <sub>11</sub>	X <sub>21</sub>	X <sub>31</sub>	.....	X <sub>m1</sub>	Y <sub>1</sub>
2	X <sub>12</sub>	X <sub>22</sub>	X <sub>32</sub>	.....	X <sub>m2</sub>	Y <sub>2</sub>
3	X <sub>13</sub>	X <sub>23</sub>	X <sub>33</sub>	.....	X <sub>m3</sub>	Y <sub>3</sub>
4	X <sub>14</sub>	X <sub>24</sub>	X <sub>34</sub>	.....	X <sub>m4</sub>	Y <sub>4</sub>
.....				.....		
n	X <sub>1n</sub>	X <sub>2n</sub>	X <sub>3n</sub>	.....	X <sub>mn</sub>	Y <sub>n</sub>

Моделювання процесу зі складною зовнішньою дією у вигляді декількох вхідних факторів, так само, як і для процесу з одним вхідним фактором, починається з лінійної моделі.

Якщо мати необмежено велику кількість експериментальних точок, то лінійна регресійна модель з декількома вхідними змінними має вигляд:

$$\bar{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_m x_m, \quad (18)$$

де  $x_1, x_2, x_3, \dots, x_m$  – значення вхідної змінної;

$\beta_0, \beta_1, \beta_2, \dots, \beta_m$  – коефіцієнти регресії.

Експериментальні дані у вигляді комбінацій ( $x_{1i}, x_{2i}, x_{3i}, \dots, x_{mi}, y_i$ ) є лише обмеженою вибіркою із загального числа станів досліджуваного процесу. Тому можливо визначити тільки коефіцієнти  $\beta_0, \beta_1, \beta_2, \dots$ , які позначають, відповідно,  $b_0, b_1, b_2, b_3, \dots, b_m$ .

$$\bar{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_m x_m. \quad (19)$$

Такі моделі називаються багатофакторними, в яких визначити коефіцієнти регресії  $b_i$  так, як це робиться для однофакторної моделі, не уявляється можливим. Необхідно використовувати основи алгебри матриць і матричного числення.

У відповідності з даними таблиці, матриці  $X, Y, B$  мають вигляд:

$$X = \begin{bmatrix} 1x_{11} & x_{21} & x_{31} & \dots & x_{m1} \\ 1x_{12} & x_{22} & x_{32} & \dots & x_{m2} \\ 1x_{13} & x_{23} & x_{33} & \dots & x_{m3} \\ \dots & \dots & \dots & \dots & \dots \\ 1x_{1m} & x_{2m} & x_{3m} & \dots & x_{mm} \end{bmatrix}, \quad (20)$$

У матриці  $X$  всі елементи першого стовпчика дорівнюють одиниці. Вважатимемо це фіктивною вхідною змінною  $X_{03}$  постійним значенням.

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_m \end{bmatrix}, \quad (21)$$

$$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ b_k \end{bmatrix}, \quad (22)$$

Розмірність матриць така:

$Y$  – вектор спостережень  $(m \cdot 1)$ ;

$X$  – матриця незалежних змінних  $[m(k + 1)]$ ;

$B$  – вектор коефіцієнтів регресії  $[(k + 1) \cdot 1]$ .

Правила множення матриць і векторів вимагають, щоб вони були узгодженими і мали відповідну розмірність [5; 7].

Вважаємо, що матриця  $A$  має розмірність  $(n \cdot p)$ :

– на матрицю  $A$  тільки зліва може бути помножена матриця  $B$  розмірністю  $(m \cdot n)$ :

$$B \cdot A = (m \cdot n \cdot n \cdot p) = C(m \cdot p); \quad (23)$$

– на матрицю  $A$  тільки справа може бути помножена матриця  $D$  розмірністю  $(p \cdot q)$ :

$$A \cdot C = (n \cdot p \cdot p \cdot q) = F(n \cdot q). \quad (24)$$

Отже, додаток  $B \cdot X$  не існує і лінійна регресія може бути записана у вигляді

$$Y = X \cdot B. \quad (25)$$

Використання апарату лінійної алгебри дозволяє отримати загальну залежність для визначення вектора, що містить коефіцієнти регресії [5]:

$$B = (X' \cdot X)^{-1} \cdot X' \cdot Y, \quad (26)$$

де  $(X' \cdot X)^{-1}$  – зворотна матриця;

$X'$  – транспонована матриця.

Необхідно визначити адекватність і точність запропонованої багатофакторної моделі.

Адекватність моделі характеризує відповідність моделі експериментальним даним і статистичну значущість рівняння регресії.

Адекватність регресійної моделі оцінюється коефіцієнтом Фішера:

$$F_{розр} = \frac{\sum_{i=1}^n \left( \hat{y}_i - \bar{y} \right)^2}{\sum_{i=1}^n \left( y_i - \hat{y}_i \right)^2}. \quad (27)$$

Розрахункове значення коефіцієнта  $F_{розр}$  необхідно порівняти з табличним  $F_{табл}(m, \alpha)$ , де  $m$  – загальна кількість експериментальних спостережень,  $\alpha$  – рівень значущості.

Для багатофакторних моделей табличне значення критерію Фішера залежить ще і від числа вхідних змінних [5].

При  $F_{розр} > F_{табл}$  модель вважається адекватною, а регресія статистично значущою.

При  $F_{розр} < F_{табл}$  регресійна модель неадекватна, а регресія статистично незначуща.

Для оцінки точності регресійних моделей з декількома вхідними змінними використовується множинний коефіцієнт кореляції ( $R^2$ ) [4], який визначається за формулою:

$$R^2 = \frac{\sum_{i=1}^n \left( \hat{y}_i - \bar{y} \right)^2}{\sum_{i=1}^n \left( y_i - \bar{y} \right)^2} \quad (28)$$

і характеризує щільність зв'язку між вихідними і вхідними змінними. При цьому  $0 \leq R^2 \leq 1$ .

При  $R^2 = 0$  вихідний фактор у лінійно не залежить від вхідних факторів  $x_1, x_2, \dots, x_k$ , тобто кореляційний зв'язок між вихідним чинником і вхідними факторами відсутній.

При  $R^2 = 1$  вихідний фактор у лінійно залежить від вхідних  $x_1, x_2, \dots, x_k$  – тобто в наявності сильний кореляційний зв'язок. При вищому значенні  $R^2$ , присутній істотний зв'язок у моделі між змінними вихідними і вхідними змінними факторами. При низькому значенні  $R^2$  будемо мати низьку точність оцінки і прогнозу поведінки або властивостей об'єкта, а використання такої моделі для дослідження, описання і прогнозу об'єкта недоцільне.

З декількох моделей для дослідження вибирається та, у якої відношення  $R^2$  має найбільше

значення. Якщо в результаті розрахунку відношення  $R^2$  лінійна регресія визнається недостатньо точною, то переходять до дослідження складніших моделей, таких як:

$$y = b_0 + b_1x + b_2x^2 + b_3x^3; \quad (29)$$

– *полінома з декількома незалежними змінними*

$$y = b_0b_1x_1 + b_2x_2 + b_3x_1^2x_2 + b_4x_2^2x_1 + b_5x_1x_2; \quad (30)$$

– *звотної моделі*

$$y = b_0 + \frac{b_1}{x_1} + \frac{b_2}{x_2} + \frac{b_3}{x_3} + \frac{b_4}{x_4}, \quad y = b_0 + \frac{b_1}{x_1} + \frac{b_2}{x_2} + \frac{b_3}{x_1^2} + \frac{b_4}{x_2^2} + \frac{b_5}{x_1 \cdot x_2}; \quad (31)$$

– *комбінованої моделі*

$$y = b_0 + b_1x_1 + b_2 \cdot \sqrt{x} + b_3 \cdot \ln x + b_4 \cdot \exp(x). \quad (32)$$

Дослідник сам може запропонувати вид багатofакторної на основі аналізу апriorної інформації. Оцінка коефіцієнтів регресії, критерію адекватності моделі і множинного коефіцієнта кореляції здійснюється за формулами (27), (28), (29). Якщо в результаті побудови лінійних моделей жодна з них не була визнана достатньо точною, переходять до дослідження більш складних моделей. Будь-яку модель, записувану в результаті перетворення у вигляді (18), доцільно аналізувати методами лінійного регресійного аналізу.

На практиці використовують такі види перетворень для нелінійних моделей [7]:

- зворотне перетворення;
- логарифмічне перетворення;
- перетворення типу квадратного кореня.

Якщо за допомогою будь-яких перетворень нелінійна модель може бути приведена до вигляду множинної лінійної регресії, то вона називається нелінійною моделлю з «внутрішньою лінійністю» [6]. До таких моделей належать:

– *статична (мультиплікативна) модель*

$$y = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot x_3^{b_3} \cdot \dots \cdot x_k^{b_k}. \quad (33)$$

Після перетворення:  $\ln y = \ln b_0 + b_1 \ln x_1 + b_2 \ln x_2 + \dots + b_k \ln x_k$ .

Якщо ввести наступні перетворення:  $\ln x_i = z_i$ ,  $w = b_0 + b_1z_1 + b_2z_2 + \dots + b_kz_k$ , то модель матиме вигляд багатofакторної лінійної регресії, аналогічний (18).

Дана модель має вигляд багатofакторної лінійної регресії (див. формулу (18)), отже, *статична* модель може бути перетворена логарифмуванням до вигляду багатofакторної лінійної регресії з відповідним розрахунком коефіцієнтів регресії, критерію адекватності і множинного коефіцієнта кореляції. Але після розрахунків необхідно виконати потенціювання [6] і повернутися до початкового статичного виду моделі;

– *експоненціальна модель*

$$y = \exp(b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k). \quad (34)$$

Логарифмування дозволяє перетворити експоненціальну модель до виду багатofакторної лінійної моделі і використовувати для її дослідження апарат лінійного регресійного аналізу;

– *звотна модель*

$$y = \frac{1}{b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_k \cdot x_k}. \quad (35)$$

Після виконання перетворень ( $\frac{1}{y} = b_0 + b_1 \cdot x_2 + b_3 \cdot x_3 + \dots + b_k \cdot x_k$ ) та введення позначень

( $\frac{1}{y} = w$ ,  $w = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots + b_k \cdot x_k$ ) будемо мати багатofакторну лінійну модель,

що дозволяє використовувати для її дослідження апарат лінійного регресійного аналізу.

У випадку неможливості перетворення нелінійної моделі до вигляду множинної лінійної регресії (так звана нелінійна модель з «внутрішньою нелінійністю») для їх дослідження, слід використовувати апарат нелінійного регресійного аналізу [7].

На практиці при дослідженні технологічних процесів на підприємствах автомобільного транспорту і побудованні регресійних моделей з декількома вхідними змінними використовують крокові (ітераційні) методи, що дозволяють обґрунтовано вводити в модель тільки найбільш впливові і найбільш значущі вхідні фактори. Частіше всього використовується метод виключення змінних і метод включення змінних.

Метод виключення змінних, згідно з [7], складається із семи етапів.

Етап перший: пропонується регресійна модель, що включає всі досліджувані вхідні змінні.

Етап другий: розраховується значення критерію Фішера для кожної вхідної змінної  $F_{x_i}$ ; оцінюється статистична значущість кожної вхідної змінної аналогічно тому, як це виконується для моделі в цілому.

Етап третій: встановлюється вхідна змінна з мінімальним значенням  $F_{x_i}$ .

Етап четвертий: мінімальне значення критерію Фішера порівнюється з граничним значенням. При  $F_{min} > F_{табл}(m, \alpha)$  відповідна вхідна змінна вважається статистично значущою і залишається в моделі. Отже, запропонована на першому етапі модель статистично значуща, адекватна і може бути використана для дослідження об'єкта. Аналіз інших вхідних змінних, що залишилися, не проводиться. Далі розраховуються коефіцієнти регресії і множинний коефіцієнт кореляції. При  $F_{min} < F_{табл}$ , вхідна змінна вважається статистично незначущою і повинна бути видалена з моделі як неадекватна. Слід пам'ятати, що для багатofакторних моделей табличне значення критерію Фішера залежить ще і від числа вхідних змінних [5; 7].

Етап п'ятий: після видалення статистично незначущих вхідних змінних для вхідних змінних, що залишилися, знову перераховуються їх частинні критерії Фішера.

Етап шостий: знову встановлюється вхідна змінна (з тих, що залишилися з мінімальним значенням  $F_{x_i}$ ).

Етап сьомий: четвертий, п'ятий і шостий етапи повторюються доти, поки в моделі не залишаться тільки статистично значущі змінні.

Якщо після виконання вказаних процедур в моделі не залишиться жодної статистично значущої змінної, модель побудувати неможливо. Причиною є прорахунки при плануванні експерименту на основі аналізу апріорної інформації. В цьому випадку слід знову уважно проаналізувати апріорну інформацію, наново спланувати і провести експеримент і обробити його результати.

Якщо в моделі залишаються вхідні змінні, то переходять до розрахунку коефіцієнтів регресії і оцінки точності моделі.

Метод включення змінних складається з восьми етапів:

Етап нульовий: у «моделі» визначається відсутність вхідних змінних.

Етап перший: розраховується кореляційна матриця (з часткових коефіцієнтів кореляції вхідних змінних – одна з іншою і з вихідною змінною):

$$\begin{bmatrix} R_{1y} & R_{11} & R_{21} & R_{31} & R_{41} & \dots & R_{k1} \\ R_{2y} & R_{12} & R_{22} & R_{32} & R_{42} & \dots & R_{k2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ R_{ky} & R_{1k} & R_{2k} & R_{3k} & R_{4k} & \dots & R_{kk} \end{bmatrix}, \quad \dots \dots \dots \quad (36)$$

де  $R_{ky}$  – частинний коефіцієнт кореляції  $k$ -ї вхідної і вихідної змінних;

$R_{ij}$  – частинний коефіцієнт кореляції  $i$ -ї і  $j$ -ї вхідних змінних ( $i \neq j$ ).

Етап другий: підбирається вхідна змінна з максимальним коефіцієнтом кореляції  $R_{iy}$ , яка першою вводиться в модель.

Етап третій: визначається частинний критерій Фішера введеної вхідної змінної, який одночасно є критерієм адекватності всієї моделі. При  $F_{розр} < F_{табл}$  відповідна вхідна змінна статистично незначуща, тобто повинна бути видалена з моделі, а сама модель визнається неадекватною. Вхідні змінні, що залишилися, мають з вихідною змінною ще менш тісний кореляційний зв'язок. Отже, в даній ситуації побудування моделі неможливе.

При  $F_{розр} > F_{табл}(m, \alpha)$  відповідна вхідна змінна враховується статистично значущою і залишається в моделі.

Етап четвертий: кореляційна матриця перераховується без урахування впливу вибраної вхідної змінної.

Етап п'ятий: з вхідних змінних, що залишилися, вибирається змінна з максимальним коефіцієнтом кореляції  $R_{iy}$ . Ця змінна вводиться в модель наступною.

Етап шостий: визначається критерій Фішеранової моделі. При  $F_{розр} < F_{табл}$  введена в модель вхідна змінна статистично незначуща і повинна бути видалена з моделі як неадекватна, а в моделі залишається одна вхідна змінна.

Етап сьомий: етапи чотири, п'ять і шість повторюються до тих пір, поки не буде сформований остаточний вигляд моделі, після чого обов'язкові розрахунки коефіцієнтів регресії і оцінки точності моделі.



**Висновки.** 1. Основною вимогою до побудовання емпіричних регресивних моделей є адекватність моделі, точність і економічність розрахунків.

2. При моделюванні технологічних процесів підприємств автомобільного транспорту найбільш оптимальними є крокові методи побудовання регресійних моделей.

#### ВИКОРИСТАНА ЛІТЕРАТУРА

1. **Цирлин А. М.** Оптимальное управление технологическими процессами / А. М. Цирлин. – М. : Энергопромиздат, 1986. – 400 с.

2. **Тихонов А. Н., Уфимцев М. В.** Статистическая обработка результатов экспериментов: учебн. пособ. / А. Н. Тихонов, М. В. Уфимцев. – М. : Изд-во Моск. ун-та, 1988. – 174 с.

3. **Евсеев Д. Г., Тарасевич О. М., Корноухов А. П.** Автоматизация производственных процессов в машиностроении / Д. Г. Евсеев, О. М. Тарасевич, А. П. Карнаухов. – М. : МИИТ, 2005. – 93 с.

4. **Советов Б. Я., Яковлев С. А.** Моделирование систем : учебн. издан. / Б. Я. Советов, С. А. Яковлев. – М. : Высшая школа, 2001. – 343 с.

5. **Корн Г.** Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М. : Наука, 1972. – 830 с.

6. **Рогов В. А.** Методика и практика технических экспериментов: учебн. пособ. / В. А. Рогов. – М. : Академия, 2005. – 288 с.

7. **Дрейпер Н.** Прикладной регрессионный анализ: пер. с англ. / Н. Дрейпер, Г. Смит. 3-е изд. – М. : Вильямс, 2007. – 912 с.

8. **Спирин Н. А.** Методы планирования и обработки результатов инженерного эксперимента: учебн. пособ. / Н. А. Спирин [и др.]; под ред. Н. А. Спирина; ГОУ ВПО УГТУ – УПИ – Екатеринбург, 2003. – 260 с.

9. **Говорущенко Н. Я.** Основы управления автомобильным транспортом / Н. Я. Говорущенко. – Харків : Вища школа, 1976. – 224 с.

УДК 656.13-049.7

**Особливості побудови емпіричних регресивних моделей при моделюванні технологічних процесів підприємств автомобільного транспорту / Г. В. Заяць // Вісник Придніпровської державної академії будівництва та архітектури. – Д. : ПДАБА, 2013. – № 4. – С. 17 – 25. – рис. 2. – Бібліограф. : (9 назв.).**

Визначено найоптимальніші методи побудови емпіричних регресійних моделей при моделюванні технологічних процесів підприємств автомобільного транспорту.

*Ключові слова:* технологічний процес, модель, регресія, кореляція, імовірність, адекватність, точність.

**Особенности построения эмпирических регрессивных моделей при моделировании технологических процессов предприятий автомобильного транспорта / Г. В. Заяц // Вісник Придніпровської державної академії будівництва та архітектури. – Д. : ПДАБА, 2013. – № 4. – С. 17 – 25. – рис. 2. – Бібліогр.: (9 назв.).**

Определены наиболее оптимальные методы построения эмпирических регрессионных моделей при моделировании технологических процессов предприятий автомобильного транспорта.

*Ключевые слова:* технологический процесс, модель, регрессия, корреляция, вероятность, адекватность, точность.

**Features of development of empiric regressive models at the design of technological processes of enterprises of motor transport / G. Zayats // Visnyk of Pridneprovsk State Academy of Civil Engineering and Architecture. – D. : PSACEA, 2013. – № 4. – P. 17 – 25. – pic. 2. – Bibliogr.: (9 names).**

Optimum methods of construction regressive models at the design of technological processes of enterprises of motor transport certain

*Key words:* technological process, model, regression, correlation, probability, adequacy, exactness.