

ПРАКТИЧНА РОБОТА №3

МОДЕЛІ НЕЛІНІЙНОЇ ПАРНОЇ РЕГРЕСІЇ

Завдання. На основі вхідних статистичних даних за 15 років про динаміку виробничих фондів K та трудових ресурсів L промислового підприємства, табл. 1, необхідно:

1. Спрогнозувати динаміку використання основних виробничих фондів K на наступний період, використовуючи: степеневу та параболічну залежність другого порядку;
2. Спрогнозувати динаміку використання трудових ресурсів L на наступний період, використовуючи: експоненційну, логарифмічну та гіперболічну залежності;
3. З надійністю $P = 0,95$, оцінити достовірність побудованих нелінійних регресій за допомогою критерію Фішера;
4. Побудувати графіки вхідних статистичних даних та відповідних ліній регресії на них.
5. На основі одержаних результатів зробити висновки щодо найбільш вірогідних прогнозів динаміки виробничих фондів K та трудових ресурсів L на наступний період.

Таблиця 1 – Вхідні дані про динаміку виробничих фондів K та трудових ресурсів L промислового підприємства

Рік	K	L
1	$= 311 + 2N$	$= 92 + 2N$
2	$= 305 + 2N$	$= 100 + 2N$
3	$= 348 + 2N$	$= 122 + 2N$
4	$= 398 + 2N$	$= 118 + 2N$
5	$= 400 + 2N$	$= 143 + 2N$
6	$= 471 + 2N$	$= 159 + 2N$
7	$= 465 + 2N$	$= 148 + 2N$
8	$= 534 + 2N$	$= 177 + 2N$
9	$= 527 + 2N$	$= 153 + 2N$
10	$= 544 + 2N$	$= 200 + 2N$
11	$= 586 + 2N$	$= 238 + 2N$

12	= 592 + 2N	= 266 + 2N
13	= 584 + 2N	= 271 + 2N
14	= 610 + 2N	= 290 + 2N
15	= 597 + 2N	= 340 + 2N

Де N – порядковий номер студента по журналу

Рекомендації до виконання.

1. Прогнозування динаміки використання основних виробничих фондів К на наступний період, згідно поставленого завдання, відбувається на основі степеневі та параболічної залежностей другого порядку. Для вирішення даного завдання використовується метод найменших квадратів. Розглянемо специфіку його застосування для вказаних трьох випадків.

а) Степенева залежність $y_{\text{розр}} = a_0 x^{a_1}$. Шляхом послідовного логарифмування та заміни змінних, спочатку перетворюємо степеневу залежність до лінійного вигляду:

$$\ln(y_{\text{розр}}) = \ln(a_0 x^{a_1}), \text{ звідки отримуємо:}$$

$$\ln(y_{\text{розр}}) = \ln(a_0) + \ln(x^{a_1}) = \ln(a_0) + a_1 \times \ln(x).$$

Далі проводимо заміну змінних:

$$\ln(y_{\text{розр}}) = \hat{y};$$

$$\ln(a_0) = \hat{a}_0;$$

$$\ln(x) = \hat{x}.$$

Тоді остаточне рівняння набуває лінійного вигляду:

$$\hat{y} = \hat{a}_0 + a_1 \hat{x}.$$

Після знаходження методом найменших квадратів невідомих параметрів $\hat{a}_0$ та a_1 , необхідно розрахувати початкове a_0 , згідно попередньої заміни:

$$a_0 = e^{\hat{a}_0}.$$

б) Парабола другого порядку $y_{\text{розр}} = a_0 + a_1 x + a_2 x^2$. Введемо заміну змінних: $x_1 = x^2$. Тоді, рівняння параболи прийме вигляд:

$$y_{\text{розр}} = a_0 + a_1 x + a_2 x_1.$$

Як бачимо, початкова нелінійна залежність з однією змінною x перетворилась на лінійну залежність з двома змінними x та x_1 . Тобто, маємо випадок множинної регресії, який буде розглядатись нижче.

2. Прогнозування динаміки використання трудових ресурсів L на наступні 3 періоди, згідно поставленого завдання, слід виконувати за допомогою експоненційної, логарифмічної та гіперболічної залежностей. Для застосування з цією метою методу найменших квадратів, спочатку слід перетворити вказані залежності до лінійного вигляду.

а) Експоненційна залежність $y_{\text{розн}} = a_0 e^{a_1 x}$. Перш ніж провести заміну змінних, виконаємо логарифмування лівої й правої частин нашого рівняння:

$$\ln(y_{\text{розн}}) = \ln(a_0 e^{a_1 x}), \text{ звідки отримуємо:}$$

$$\ln(y_{\text{розн}}) = \ln(a_0) + \ln(e^{a_1 x}) = \ln(a_0) + a_1 x \times \ln(e) = \ln(a_0) + a_1 x.$$

Далі проводимо заміну змінних:

$$\ln(y_{\text{розн}}) = \hat{y};$$

$$\ln(a_0) = \hat{a}_0.$$

Тоді остаточне рівняння набуває лінійного вигляду:

$$\hat{y} = \hat{a}_0 + a_1 x.$$

Після знаходження методом невідомих параметрів $\hat{a}_0$ та a_1 , необхідно розрахувати початкове a_0 :

$$a_0 = e^{\hat{a}_0}.$$

б) Логарифмічна залежність $y_{\text{розн}} = a_0 + a_1 \ln(x)$. Введемо заміну змінних:

$$\ln(x) = \hat{x}.$$

Тоді, початкове рівняння знову приймає лінійну форму:

$$y_{\text{розн}} = a_0 + a_1 \hat{x}.$$

в) Гіперболічна залежність $y_{\text{розн}} = a_0 + \frac{a_1}{x}$. Введемо заміну змінних:

$$\frac{1}{x} = \hat{x}.$$

Тоді, початкове рівняння прийме лінійну форму:

$$y_{\text{розн}} = a_0 + a_1 \hat{x}.$$

3. Кожну з побудованих регресій слід перевірити на достовірність вхідним даним за допомогою критерію Фішера й обрати ту, де $F_{\text{розр}}$ приймає найбільше значення. Для цього спочатку слід обчислити вибірковий коефіцієнт детермінації R^2 :

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

Тоді, критерій Фішера про значущість характеристик парної лінійної регресії буде мати вигляд:

$$F \geq F_{pk},$$

Розрахункове значення критерію Фішера обчислюється за формулою:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}$$

p – заданий рівень ймовірності ($p = 0,95$);

n – кількість спостережень;

m – кількість факторів x , що впливають на y ($m = 1$);

$k = (n - m - 1)$ – кількість ступенів вільності.

Табличне значення F_{pk} можна розрахувати в Microsoft Excel за допомогою функції:

$= FINV(1-p; m; n-m-1)$, або

$= FPACПОБР(1-p; m; n-m-1)$.