УДК 004.94 **doi: 10.26906/SUNZ.2022.4.051**

А. Н. Аль-Амморі, М. М. Дехтяр, Х. І. С. Абдусалам

Національний транспортний університет, Київ, Україна

КОМП'ЮТЕРНІ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ПЛАНУВАННЯ ЕКСПЕРИМЕНТІВ

Анотація. Вирішення завдань аналізу та дослідження економічних, виробничих та управлінських процесів на транспорті базується на використанні новітніх комп'ютерних технологій (математичних та статистичних пакетів, електронних таблиць тощо). Останнім часом підвищений інтерес спостерігається до методів статистичного аналізу. У цій роботі як інструмент імовірнісно-статистичних розрахунків обрано математичний пакет Mathcad. Вибір Mathcad обумовлений тим, що він є потужною та універсальною системою комп'ютерної математики, призначеної для математичних розрахунків. Простота інтерфейсу Mathcad, широкий набір інструментів для комп'ютерної реалізації числових методів вирішення математичних завдань, якісна комп'ютерна графіка, можливість запису формул у їхньому природному вигляді зробили його одним із популярних інструментів математичних розрахунків. Важливим компонентом Mathcad є система програмування, яка має мову, наближену до традиційних професійних мов програмування, що значно розширює обчислювальні можливості Mathcad. Перевагою Mathcad є його відкритість для поповнення новими алгоритмами, функціями користувача та програмами. Робота об'єднує як теоретичну, так і комп'ютерну технології розрахунків із чисельної реалізації математичних моделей запропонованих завдань. Це дає можливість поглибити знання теорії статистичного аналізу, розвинути вміння у вирішенні завдань за наведеними комп'ютерними алгоритмами, розширити уявлення про сфери застосування методів статистичного аналізу для вирішення практичних завдань у галузі економіки, виробництва та управління.

Ключові слова: статистичний метод, повний експеримент, інформаційні технології, математична модель планування експерименту.

Вступ

Планування експериментів передбачає визначення найбільш ефективного способу його проведення з метою одержання статистичного матеріалу, який має заздалегідь задані властивості. Планування експериментів застосовується при розв'язуванні задач оцінки параметрів розподілів, перевірки статистичних гіпотез при заданій потужності критерію, знаходження математичної моделі процесів із заданими статистичними властивостями, пошуку оптимальних значень за заданими критеріями умов протікання процесів. У даній роботі ми розглянемо математичні методи планування експериментів для вивчення механізму процесу, що спостерігається і побудови його статистичної моделі.

Планування регресійних експериментів (статистичне моделювання)

Задача планування експериментів формулюється наступним чином: потрібно одержати деяке уявлення про поверхню відгуку, яка описується моделлю $y = f(x_1, x_2, \ldots, x_n)$, де y –залежна змінна – відгук (наприклад, надійність приладу); x_i –незалежні змінні – фактори, що впливають на відгук, які можна варіювати у ході експерименту (наприклад, параметри навантаження приладу). Невідома функція відгуку частіш представляється поліномом k-го степеня

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{\substack{i,j=1 \\ i<j}}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} a_{ii} x_i^2 + \ldots, \quad (1)$$

де a_0, a_i, a_{ij}, a_{ii} –коефіцієнти поліному.

Планування експерименту полягає у виборі на кожному етапі дослідження оптимального у прийнятих критеріях розташування експериментальних точок у просторі факторів. За критерії оптимальності планів використовуються:

- мінімізація числа випробувань;
- незалежність оцінок коефіцієнтів функції відгуку (ортогональність плану);
- однорідність дисперсій відгуку відносно центру плану (рототабельність плану);
- мінімізація об'єму еліпсоїда розсіяння оцінок коефіцієнтів моделі (*D*–оптимальність плану).

Найбільше застосування знайшли ортогональні плани у сполученні з критеріями *D*–оптимальності.

Як правило, поверхня відгуку відшукується у деякій визначеній області зміни факторів. Найбільш широко застосовується планування на двох рівнях (екстремальний експеримент), коли у експерименті використовуються значення факторів, які відповідають верхньому і нижньому рівню і позначаються +1 і відповідно (або просто + і –). Експериментальні плани, у яких усі фактори варіюються тільки на двох рівнях, називаються планами 2^k, де k – кількість факторів, що варіюються.

При побудові плану експерименту потрібно виходити із деякого апріорного уявлення про можливий вигляд функції відгуку (лінійність, монотонність і т. інше). Спочатку область варіювання факторів визначається, виходячи із припущення про лінійність поверхні відгуку у середині цієї області. Якщо лінійна поверхня відгуку описує експериментальний матеріал неадекватно, то проводяться подальші експерименти по уточненню її вигляду за допомогою поліномів більш високого порядку.

Лінійні ортогональні плани

Повний факторний експеримент. Планом, який дозволяє одержати незалежні оцінки коефіцієнтів рівняння поверхні відгуку, є повний факторний експеримент (ПФЕ). ПФЕ реалізує усі можливі неповторювані комбінації рівнів незалежних факторів.

У даному розділі ПФЕ розглядається стосовно до умов активного екстремального експерименту. Розглянемо ПФЕ для трьох факторів (k = 3). Поверхня відгуку у цьому випадку має вигляд

$$y = a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{\substack{i,j=1 \\ i<j}}^{k} a_{ij} x_i x_j + a_{123} x_1 x_2 x_3. \quad (2)$$

Нехай кожний фактор x_i варіюється від основного рівня x_{i0} на величину $\pm \Delta x_i$. Тоді за допомогою перетворення $\tilde{x}_i = \frac{x_i - x_{i0}}{\Delta x_i}$ можна перейти до кодованих змінних $\tilde{x}_i$, що приймають на границях інтервалу варіювання x_i значення ± 1. План екстремального експерименту прийнято записувати у вигляді матриці, яка визначає у кодованих змінних $\tilde{x}_i$ умови проведення експерименту. У подальшому під x_i будемо розуміти кодовані змінні $\tilde{x}_i$. У літературі k-факторний ПФЕ зі зміною факторів на двох рівнях прийнято називати планом типу 2^k. Наведемо приклад матриці ПФЕ для трьох факторів (табл. 1).

Таблиця 1 – **Матриця повного факторного експерименту (ПФЕ) 2^k**

Номер точки плану, j	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	Відгук y_j
1	+	–	–	–	+	+	+	–	y_1
2	+	+	–	–	–	–	+	+	y_2
3	+	–	+	–	–	+	–	+	y_3
4	+	+	+	–	+	–	–	–	y_4
5	+	–	–	+	+	–	–	+	y_5
6	+	+	–	+	–	+	–	–	y_6
7	+	–	+	+	–	–	+	–	y_7
8	+	+	+	+	+	+	+	+	y_8

Для ПФЕ мають місце співвідношення ($n = 2^k$ –кількість точок плану):

$$\sum_{j=1}^{n} x_{ij} = 0 (i = 1,2,\ldots,2^{k-1}) ; \sum_{j=1}^{n} x_{ij}^2 = n; \quad (3)$$

$$\sum_{j=1}^{n} x_{ij} x_{rj} = 0 (i \neq r; i,r = 0,1,\ldots,2k-1) \ ), \quad (4)$$

що відповідає властивості ортогональності стовпців матриці плану.

Будь-який план 2^k може бути побудований за наступним правилом: у стовпці, який відповідає фактору x_i, знаки + і – чергуються через 2^{i-1}. План 2^k дозволяє оцінити 2^k коефіцієнтів регресії a_i. Однак використовувати ПФЕ для оцінки коефіцієнтів при членах із кратністю не можна, оскільки оцінки для $a_0 і a_{ii}$ змішуються (наприклад, стовпці $x_0 і x_i x_i$ нерозрізними).

Основною перевагою ПФЕ є ортогональність матриці плану, що дозволяє суттєво спростити обчислення коефіцієнтів рівняння відгуку.

Для будь-якого числа факторів k вибіркові коефіцієнти a_i обчислюються за формулами:

$$a_j = \frac{1}{n} \sum_{i=1}^{n} a_{ij} \bar{y}_j , де \bar{y}_j = \frac{1}{m} \sum_{i=1}^{m} y_{ij},$$

де m –число паралельних випробувань у i-й точці плану, n –загальне число точок плану.

Дисперсія, яка характеризує розкид значень y_{jr}, при постійних умовах експерименту (тобто в одній точці плану), знаходиться за формулою

$$S_j^2 = \frac{1}{m-1} \sum_{i=1}^{m} (y_{ij} - \bar{y}_j)^2. \quad (6)$$

Загальна дисперсія, яка характеризує розкид відгуку безвідносно до умов експерименту, дорівнює

$$S_y^2 = \frac{1}{n} \sum_{j=1}^{n} S_j^2 = \frac{1}{n(m-1)} \sum_{j=1}^{n} \sum_{i=1}^{m} (y_{ij} - \bar{y}_j)^2. \quad (7)$$

Якщо кількість паралельних випробувань у точках плану різна, то

$$S_y^2 = \frac{\sum_{j=1}^{n} (m_j - 1) S_j^2}{\sum_{j=1}^{n} (m_j - 1)}, \quad (8)$$

де m_j –число випробувань у j-й точці плану.

Попередньо однорідність дисперсій S_j^2 повинна бути перевірена одним із методів.

Дисперсія коефіцієнта регресії a_i визначаються формулою

$$S_{a_i}^2 \frac{1}{n(m-1)} S_y^2. \quad (9)$$

Коефіцієнт a_i рівняння відгуку з коефіцієнтом значущості α визнається значущим, якщо $|a_i| > t_{\frac{1+\alpha}{2}}(f) S_a$, де $t_{\frac{1+\alpha}{2}}$ – квантіль розподілу Стьюдента при $f = n(m-1)$ ступенях волі.

Для перевірки адекватності математичної моделі відгуку використовується дисперсія

$$S^2 = \frac{m}{n-d} \sum_{j=1}^{n} (\bar{y}_j - \hat{y}_j)^2, \quad (10)$$

де d – число коефіцієнтів апроксимуючого поліному; $\hat{y}_j$– значення відгуку, яке передбачається регресійною моделлю.

Адекватність моделі встановлюється порівнянням дисперсій $S^2 i S_y^2$ за допомогою критерію Фішера $F = \frac{S^2}{S_y^2}$ при $f_1 = n - d i f_2 = n(m-1)$ ступенях свободи.

Якщо усі коефіцієнти лінійної регресії (у тому числі і усі коефіцієнти при взаємодіях) є значущими, то $d = n$ і не залишається ступенів волі для перевірки гіпотези адекватності. У цьому випадку рекомендується поставити випробування у центрі плану (тобто при значенні фактору $x_i = 0$). Тоді, якщо $|a_0 - \bar{y}_0| =< S_y$, де $\bar{y}_0$– середнє значення відгуку у центрі експерименту, лінійна модель визнається адекватною.

Приклад 1. Для ПФЕ 2^3, наведеного у табл. 2, визначити рівняння відгуку і провести його статистичний аналіз.

Таблиця 2 – **Матриця ПФЕ 2^3 і результати експерименту**

	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	Відгук		
	1	2	3	4	5	6	7	8	y_1	y_2	y_3
1	+	–	–	–	+	+	+	–	25	23	27
2	+	+	–	–	–	–	+	+	15	12	21
3	+	–	+	–	–	+	–	+	16	16	19
4	+	+	+	–	+	–	–	–	25	24	23
5	+	–	–	+	+	–	–	+	-20	-24	-22
6	+	+	–	+	–	+	–	–	24	21	24
7	+	–	+	+	–	–	+	–	-12	-14	-16
8	+	+	+	+	+	+	+	+	8	11	14

Алгоритм розв'язання задачі

1. Задання вхідних даних задачі: матриці планування експерименту і вектора значень відгуків.
2. Обчислення середніх значень відгуків.
3. Обчислення коефіцієнтів регресії рівняння відгуку.
4. Обчислення дисперсії значень відгуків, загальної дисперсії відгуків та дисперсії коефіцієнтів регресії.
5. Перевірка значущості коефіцієнтів регресії.
6. Знаходження вигляду рівняння регресії.
7. Перевірка адекватності математичної моделі відгуку за критерієм Фішера.
8. Побудова графіку точок відгуку і рівняння регресії.

Алгоритм у Mathcad

Вхідні дані:

Число випробувань і рівнів варіювання факторів

$$n := 8 \qquad m := 3 \qquad i := 0 \ldots n-1.$$

Матриця планування, значення відгуків і рівні факторів

$$x := \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix}$$

$$y := \begin{pmatrix} 25 & 23 & 27 \\ 15 & 12 & 21 \\ 16 & 16 & 19 \\ 25 & 24 & 23 \\ -20 & -24 & -22 \\ 24 & 21 & 24 \\ -12 & -14 & -16 \\ 8 & 11 & 14 \end{pmatrix} \qquad z := \begin{pmatrix} 5 & 2 & 4 \\ 4 & 2 & 3 \\ 9 & 4 & 7 \\ 1 & 0 & 1 \end{pmatrix}$$

Середні значення відгуків

$$ys_i := \frac{1}{m} \cdot \sum_{j=0}^{m-1} y_{i,j}, \tag{11}$$

$$ys^T = (25 \quad 16 \quad 17 \quad 24 \quad -22 \quad 23 \quad -14 \quad 11)$$

Коефіцієнти регресії рівняння відгуку

$$a_0 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,0} \cdot ys_i, \tag{12}$$

$$a_1 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,1} \cdot ys_i, \tag{13}$$

$$a_2 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,2} \cdot ys_i, \tag{14}$$

$$a_3 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,3} \cdot ys_i, \tag{15}$$

$$a_4 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,4} \cdot ys_i, \tag{16}$$

$$a_5 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,5} \cdot ys_i, \tag{17}$$

$$a_6 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,6} \cdot ys_i, \tag{18}$$

$$a_7 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} x_{i,7} \cdot ys_i, \tag{19}$$

$$a^T = (10 \quad -8.5 \quad 0.5 \quad 10.5 \quad -0.5 \quad 9 \quad -0.5 \quad 4.5).$$

Дисперсія значень відгуків $y_{i,j}$

$$S2_i := \frac{1}{2} \cdot \sum_{j=0}^{m-1} \left(y_{i,j} - ys_i\right)^2, \tag{20}$$

$$S2^T = (4.0 \quad 21.0 \quad 3.0 \quad 1.0 \quad 4.0 \quad 3.0 \quad 4.0 \quad 9.0).$$

Загальна дисперсія

$$Sy2 := \frac{1}{n} \cdot \sum_{i=0}^{n-1} S2_i, \;\; Sy2 = 6.13. \tag{21}$$

Дисперсія коефіцієнтів регресії

$$Sa := \sqrt{S2a}, \;\; Sa = 0.62. \tag{22}$$

Перевірка значущості коефіцієнтів регресії із рівнем значущості α і визначення ступеня волі *f*:

$$\alpha := 0.05 \;\; f := n \cdot (m-1), \;\; f = 16. \tag{23}$$

Квантіль розподілу Стьюдента

$$t_\alpha := qt\left(1 - \frac{\alpha}{2}, f\right), \;\; t_\alpha = 2.12. \tag{24}$$

Статистики Стьюдента для коефіцієнтів регресії

$$t_i := \frac{|a_i|}{Sa}, \tag{25}$$

$$t^T = (16.2 \quad 13.7 \quad 0.8 \quad 17 \quad 0.8 \quad 14.5 \quad 0.8 \quad 7.3).$$

Перевірка гіпотези про значущість коефіцієнтів регресії

$$Q_i := if(t_i < t_\alpha, 0{,}1), \tag{26}$$

$$Q^T = (1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1)\,,$$

$$a_i := a_i \cdot Q_i, \tag{27}$$

$$a^T = (10 \quad -8.5 \quad 0 \quad 10.5 \quad 0 \quad 9 \quad 0 \quad 4.5).$$

Значення $Q_i = 0$ означає, що відповідні коефіцієнти a_i є незначущими. Загальний вигляд рівняння регресії ($Y1r_i$ i $Y2r_i$ – складові загального рівняння) є таким:

$$Y1_i := a_0 + a_1 \cdot x_{i,1} + a_2 \cdot x_{i,2} + a_3 \cdot x_{i,3}\,, \tag{28}$$

$$Y2_i := a_4 \cdot x_{i,4} + a_5 \cdot x_{i,5} + a_6 \cdot x_{i,6} + a_7 \cdot x_{i,7}, \tag{29}$$

$$Y_i := Y1_i + Y2_i\,. \tag{30}$$

Значення відгуку за рівнянням регресії

$$Yr^T = (25.5 \quad 15.5 \quad 16.5 \quad 24.5 \quad -22.5 \quad 21.5 \quad -13.5 \quad 12.5).$$

Перевірка адекватності математичної моделі відгуку за критерієм Фішера:

- кількість значущих коефіцієнтів регресії

$$d := \sum_{i=0}^{n-1} if(Q_i = 1{,}1{,}0)\,, \; d = 5, \tag{31}$$

- залишкова дисперсія

$$S2 = 6.0,$$

- значення статистики критерію

$$F := \frac{S2}{Sy2}, \;\; F = 0.98. \tag{32}$$

Квантиль F –розподілу із ступенями волі

$$f_1 := n - d, \;\; f_1 = 3, \tag{33}$$

$$f_2 := n \cdot (m-1)\,, \;\; f_2 = 16. \tag{34}$$

$$t_\alpha := qF(1-\alpha, f_1, f_2)\ ,\ t_\alpha = 3.24. \qquad (35)$$

Умова прийняття гіпотези про адекватність математичної моделі

$$Q := if(F \le t_\alpha, 1{,}0),\ Q = 1. \qquad (36)$$

Умова прийняття гіпотези про адекватність моделі виконується (рис. 1).

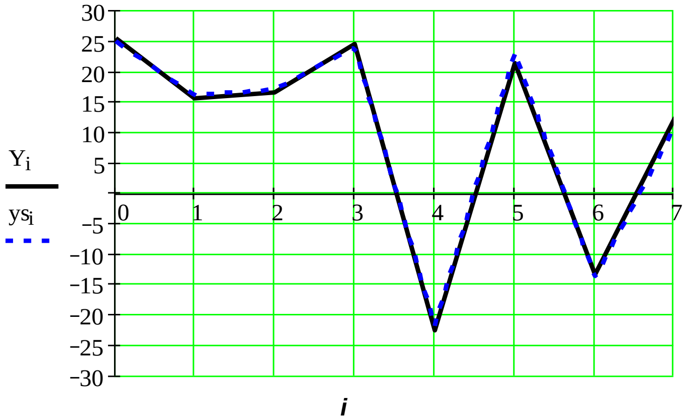

Рис. 1. Графік точок відгуку і рівняння регресії

Одержану модель можна записати у натуральних змінних, переходячи від змінних x_j до $z_j = x_j \Delta z_j + z_{0j}$.

Дробовий факторний експеримент

Якщо деякими взаємодіями можна знехтувати, то регресійна модель, яка пов'язує відгук з основними факторами, може бути одержана при меншій, ніж у ПФЕ, кількості експериментів за допомогою дробового факторного експерименту (ДФЕ). Наприклад, у випадку ПФЕ 2^3, якщо взаємодіями $x_1x_2, x_1x_3,\ x_2x_3,\ x_1x_2x_3$ можна знехтувати, то можна використати 4 стовпця матриці плану, які залишились, для визначення коефіцієнтів регресійної моделі, або знайти моделі для трьох факторів за допомогою чотирьох, а не восьми експериментів. Частина матриці ПФЕ 2^k, у якій к лінійних ефектів (факторів) прирівняні до коефіцієнтів взаємодії, називається реплікою вигляду 2^{k-r}.

Співвідношення, які визначають правила побудови дробових реплік ПФЕ, і які вказують, які фактори прирівняні до взаємодіям, називаються *генеруючими*. Наприклад, дробові репліки ПФЕ 2^{3-1} можуть бути одержані за допомогою генеруючих співвідношень $x_3 = x_1x_2 i x_3 = -x_1x_2$. Матриці планів, які відповідають цим дробовим реплікам, наведені у табл. 3.

Таблиця 3 – **Дробові репліки** 2^{3-1}

$x_3 = x_1x_2$					$x_3 = -x_1x_2$				
j	x_1	x_2	x_3	$x_1x_2x_3$	j	x_1	x_2	x_3	x_1
1	1	1	1	1	1	1	1	1	1
2	–1	–1	1	1	2	–1	–1	–1	1
3	1	–1	–1	1	3	1	1	1	1
4	–1	1	–1	1	4	–1	–1	–1	1

Для розрізнення змішаних ефектів по матриці дробової репліки використовується поняття *визначального контрасту*, який характеризує комбінацію тих факторів, стовпець добутку яких складається тільки із плюсів або тільки із мінусів.

Визначальний контраст J може бути рівним +1 або –1. Наприклад, для дробових реплік

$$J = x_1x_2x_3 = +1 \text{ або } J = x_1x_2x_3 = -1.$$

Визначальний контраст дозволяє встановити систему змішування основних факторів з ефектами взаємодії. Якщо $J = x_1x_2x_3 = +1$, то $x_1 = x_1^2x_2x_3 = x_2x_3$ (оскільки завжди $x_i^2 = +1$), тобто оцінка a_1 змішана з оцінкою a_{23}.

Якщо до ефектів взаємодії прирівнюється не один, а декілька основних факторів, для повного опису розв'язувальної здатності дробової репліки ПФЕ використовується узагальнюючий контраст. Він включає у себе частинні визначальні контрасти та їх добутки. Наприклад, при дослідженні 5 факторів можна поставити не $2^5 = 32$ випробування, а тільки 8, якщо реалізувати дробову репліку 2^{5-2}, тобто прирівняти два фактори до ефектів взаємодії. Припустимо, що вибрані варіанти змішування $x_4 = x_1x_3 i x_5 = x_1x_2x_3$ з визначальними контрастами

$$J = x_1x_3x_4 \quad i \quad J = x_1x_2x_3x_5$$

відповідно. Тоді узагальнюючий визначальний контраст може бути записаний наступним чином:

$$J = x_1x_3x_4 = x_1x_2x_3x_5 = x_2x_4x_5.$$

Для того, щоб виявити з чим змішана та чи інша оцінка коефіцієнтів моделі, необхідно помножити комбінацію факторів, які відповідають їй, на узагальнюючий визначальний контраст. Наприклад, визначимо систему змішування для оцінки коефіцієнта a_{12} ефекту взаємодії $x\ _1x_2$. Маємо кількість варіантів змінних захваток

$$x_1x_2 = x_2x_3x_4 = x_3x_5 = x_1x_4x_5,$$

тобто оцінка a_{12} буде змішаною з оцінкою

$$n := cols(PA) = 6$$

коефіцієнтів a_{234}, a_{35}, a_{145}.

У загальному випадку виходить складна система змішування ефектів взаємодії. Дисперсійний аналіз був застосований для наукового дослідження факторів, що найбільше впливають на витрати палива автомобілями-самоскидами та дорожніми машинами в зоні впливу (тобто – ділянки дороги, в якій змінюється режим руху транспортних засобів під час проведення дорожньо-ремонтних робіт) [10] Аналіз витрат палива методом двофакторного дискретного аналізу для ДМ і АС. Визначення впливу робочої швидкості та довжини смуги ущільнення на витрати палива дорожніми машинами відбувався за наступною методикою:

В програмі MathCad 14 задано нумерацію елементів масиву. Далі вводиться матриця кількісних показників сумарних витрат палива дорожніми машинами і присвоюються змінним m, n, N відповідно функції значення кількості смуг ущільнення, варіантів змінних захваток та загальної кількості операторів:

$$PA := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 102.22 & 0 & 0 & 0 & 0 & 0 \\ 70.94 & 0 & 0 & 0 & 0 & 0 \\ 56.42 & 0 & 0 & 0 & 0 & 0 \\ 48.13 & 171.80 & 0 & 0 & 0 & 0 \\ 42.81 & 140.63 & 232.42 & 0 & 0 & 0 \\ 39.11 & 119.66 & 194.78 & 0 & 0 & 0 \\ 36.39 & 104.66 & 168.01 & 249.31 & 0 & 0 \\ 34.32 & 93.44 & 148.08 & 217.91 & 0 & 0 \\ 32.68 & 84.76 & 132.71 & 193.78 & 268.91 & 0 \\ 31.35 & 77.85 & 120.53 & 174.70 & 241.18 & 0 \\ 30.26 & 72.23 & 110.64 & 159.27 & 218.79 & 289.87 \\ 29.35 & 67.57 & 102.47 & 146.55 & 200.38 & 264.54 \end{pmatrix}$$

Кількість довжин смуг укладання: $n := cols(PA) = 6$, $m := rows(PA) = 13$.

Загальна кількість операторів: $N := m \cdot n$

Задаються цикли для визначення номерів рядків і стовпців матриці: $i := 1 \dots m$, $j := 1 \dots n$.

Розраховується середнє значення витрат палива в залежності факторів А та В

$$XA_i := \frac{1}{n} \cdot \sum_{j=1}^{n} PA_{i,j} \quad (37)$$

$$XB_i := \frac{1}{m} \cdot \sum_{j=1}^{m} PA_{i,j}. \quad (38)$$

Результати обчислення:

$$XA^T = (0 \quad 17.037 \quad 11.823 \qquad 9.403 \quad 36.655)$$

$$XB^T = (42.614 \quad 71.738 \quad 93.049 \quad 87.809 \quad 71.482 \quad 42.647)$$

Розрахунок загального середнього:

$$m_x := \frac{1}{N} \cdot \sum_{i=1}^{m} \sum_{j=1}^{n} PA_{i,j}, m_x = 68.223 \quad (39)$$

Визначення кількості ступенів свободи

$$f_1 := m - 1 = 12, f_2 := n - 1 = 5\ ,$$

$$f_3 := (m-1) \cdot (n-1) = 60, \quad s2_B := \frac{S_B}{f_1}.$$

Суми квадратів відхилення по факторам А та В і відповідні дисперсії:

$$S_A := n + \sum_{i=1}^{m} (XA_i - m_x)^2 = 3.059 \times 10^4,$$

$$s2_A := \frac{S_A}{f_1} = 2.549 \times 10^3,$$

$$S_B := m + \sum_{i=1}^{n} (XB_i - m_x)^2 = 2.346 \times 10^3,$$

$$s2_B := \frac{S_B}{f_1}, \quad s2 := \frac{S}{N-1} = 7.377 \times 10^3\ .$$

$$S_\varepsilon := \sum_{i=1}^{m} \sum_{j=1}^{n} (PA_{i,j} - XA_i - XB_j + m_x)^2 = 3.542 \cdot 10^5.$$

Перевірка гіпотези НА: спостережувальне значення критеріальної статистики по фактору А:

$$F_A := \frac{{s2_A}^2}{{s2_\varepsilon}^2} = 0.186.$$

Критичне значення критерію (квантіль порядку 1-α F-розподілу Фішера з f1 та f2 ступенями свободи, $\alpha := 0.05$): $t_A := qF(1-\alpha, f_1, f_3) = 1.917$.

Перевірка гіпотези НА про рівність математичних очікувань за фактором А: $Q_A := if(F_A \le t_A)$.

Коефіцієнт детермінації $R2_A := \frac{s2_A}{s2} = 0.346$.

Перевірка гіпотези НВ: спостережене значення критеріальної статистики по фактору В:

$$F_B := \frac{s2_B}{s2_\varepsilon} = 0.033.$$

Критичне значення критерію (квантіль порядку 1-α F-розподілу Фішера з f2, f3 ступенями свободи

$$t_B := qF(1-\alpha, f_2, f_3) = 2.368.$$

Перевірка гіпотези НВ про рівність математичних очікувань за фактором В: $Q_B := if(F_B \le t_B)$.

Коефіцієнт детермінації: $R2_B := \frac{s2_B}{s2} = 0.027$.

Оцінка параметрів а та σ^2 нормального розподілу випадкових залишків:

$$a := m_x = 68.223, \quad \sigma^2 = \frac{s2_\varepsilon}{f_1 \cdot f_3} = 8.199.$$

Визначення впливу щільності асфальтобетону та довжини змінної захватки на витрати палива АС:

$$PC := \begin{pmatrix} 11.43 & 28.75 & 28.75 & 43.13 & 43.13 & 57.14 \\ 11.43 & 28.75 & 28.75 & 43.13 & 43.13 & 57.14 \\ 11.43 & 28.75 & 28.75 & 43.13 & 43.13 & 57.14 \\ 11.43 & 28.75 & 43.13 & 43.13 & 43.13 & 57.14 \\ 11.43 & 28.75 & 43.13 & 43.13 & 43.13 & 57.50 \\ 11.43 & 28.75 & 43.13 & 43.13 & 57.14 & 57.50 \end{pmatrix}$$

Кількість варіантів щільності асфальтобетону:

$$m := rows(PC) = 6.$$

Кількість варіантів довжини змінної захватки

$$n := cols(PC) = 6, \quad N := m \cdot n\ .$$

Середнє значення витрат палива в залежності факторів А та В:

$$XA_j := \frac{1}{n} \cdot \sum_{i=1}^{n} PC_{i,j}\ , \quad XB_j := \frac{1}{m} \cdot \sum_{i=1}^{m} PC_{i,j}.$$

$$XA^T = \begin{pmatrix} 35.388 & 35.388 & 35.388 \\ 37.785 & 37.845 & 840.1 \end{pmatrix},$$

$$XB^T = (11.43 \quad 28.75 \quad 35.94 \quad 43.13 \quad 45.465 \quad 57.26),$$

$$m_x := \frac{1}{N} \cdot \sum_{i=1}^{m} \sum_{j=1}^{n} PC_{i,j} = 36.996.$$

Кількість ступенів свободи

$$f_1 := m - 1 = 5, \quad f_2 := n - 1 = 5).$$

$$f_3 := (m-1) \cdot (n-1) = 25.$$

Суми квадратів відхилення по факторам А та В і відповідні дисперсії:

$$S_A := n + \sum_{i=1}^{m} (XA_i - m_x)^2 = 25.235,$$

$$s2_A := \frac{S_A}{f_1} = 5{,}047,\ s2_B := \frac{S_B}{f_1} = 249.742,$$

$$S_B := m + \sum_{i=1}^{n} (XB_i - m_x)^2 = 1.249 \times 10^3,$$

$$S_\varepsilon := 358.506),\quad s2_\varepsilon := \frac{S_\varepsilon}{f_3} = 14.34),$$

$$S := \sum_{i=1}^{m} \sum_{j=1}^{n} \left(PC_{i,j} - m_x\right)^2 = 7.93 \times 10^3,$$

$$s2 := \frac{S}{N-1} = 226.577.$$

Перевірка гіпотези НА: спостережувальне значення критеріальної статистики по фактору А:

$$F_A := \frac{s2_A}{s2_\varepsilon} = 0.352.$$

Критичне значення критерію

$$t_A := qF(1-\alpha, f_1, f_3) = 2.603.$$

Перевірка гіпотези НА про рівність математичних сподівань за фактором А:

$$Q_A := if(F_A \le t_A).$$

Коефіцієнт детермінації

$$R2_A := \frac{s2_A}{s2} = 0.022$$

Перевірка гіпотези НВ: спостережене значення критеріальної статистики по фактору В:

$$F_B := \frac{s2_B}{s2_\varepsilon} = 17.415.$$

Критичне значення критерію

$$t_B := qF(1-\alpha, f_2, f_3) = 2.603.$$

Перевірка гіпотези НВ про рівність математичних очікувань за фактором В:

$$Q_B := if(F_B \le t_B).$$

Коефіцієнт детермінації

$$R2_B := \frac{s2_B}{s2} = 1.102.$$

Оцінка параметрів а та σ^2 нормального розподілу випадкових залишків:

$$m_x = 68.223;\quad \sigma^2 := \frac{s2_\varepsilon}{f_1 \cdot f_3} = 0.115.$$

Проведений дисперсійний аналіз показав, що гіпотеза А та гіпотеза В про рівність математичних очікувань витрат палива АС в залежності від щільності асфальтобетону (фактор А) - приймається та довжини ЗЗ (фактор В) – приймається. Це означає, що обидва фактори мають вплив на витрати палива автомобілями самоскидами. Щодо ДМ – обидва фактори – швидкість руху та довжина ЗЗ (змінної захватки) впливають на витрати палива.

Висновки

Найбільш ефективними дробовими репліками від ПФЕ є репліки, у яких лінійні ефекти змішані із взаємодіями найвищого порядку.

Розрізняють регулярні і нерегулярні дробові репліки. *Регулярні репліки* утворюються із ПФЕ діленням на число частин, кратне двом. Репліки типу 4/4, 5/8 і т.д. називаються *нерегулярними*. Дробові репліки дозволяють суттєво скоротити число факторів і експериментів для моделювання процесу. Особливо ефективне їх застосування при плануванні експериментів для знаходження оптимуму відгуку.

Список літератури

1. Організація наукових досліджень: навчальний посібник / В.М. Кислий. - Суми: Університетська книга, 2011. - 224 с.
2. Планування і обробка даних наукового експерименту / В.В. Полтавець. - Донецьк: ДВНЗ ДонНТУ, 2008 — 52 с.
3. Методологія наукових досліджень технологічних процесів. /П.Білей, М.Адамовський, Я. Ханик, Н. Довга, Л. Сорока/ - Львів: Видав. НУ "Львівська політехніка«, 2003. - 352 с.
4. Засименко В.М. Основи теорії планування експерименту. Навч. посібник. - Львів: Видав. ДУ «ЛП», - 2000. - 205 с.
5. Стеченко Д.М., Чмир О.С. Методологія наукових досліджень. Підручник. - К.: Знання, 2005. - 309 с.
6. Аністратенко В.О., Федоров В.Г. Математичне планування експерименту в АПК. - К.: Вища школа, 1993.- 375 с.
7. Основи наукових досліджень / В.С. Марцин, Н.Г. Міценко, О.А. Даниленко та ін. - Львів: Ромус-Поліграф, 2002.- 128 с.
8. Шенк Х. Теорія інженерного експерименту. М., 1972- 382 с.
9. Теорія планування експериментів: – Київ : КПІ ім. Ігоря Сікорського, 2020. – 86 с.Дехтяр М. М. Інформаційні моделі та метод управління енергоефективністю дорожньо-ремонтних робіт : дис. канд. техн. наук : 05.13.06 / Дехтяр Марина Михайлівна – м. Київ, 2021. – 208 с.

Received (Надійшла) 22.08.2022
Accepted for publication (Прийнята до друку) 26.10.2022

Computer and information technologies for planning experiments

A. N. Al-Ammori, M. Dekhtyar, H. I. S. Abdusalam

Abstract. Solving tasks of analysis and research of economic, production and management processes in transport is based on the use of the latest computer technologies (mathematical and statistical packages, electronic spreadsheets, etc.). Recently, there has been an increased interest in statistical analysis methods. In this work, the mathematical package Mathcad was chosen as a tool for probabilistic and statistical calculations. The choice of Mathcad is due to the fact that it is a powerful and universal computer mathematics system designed for mathematical calculations. The simplicity of the Mathcad interface, a wide set of tools for computer implementation of numerical methods for solving mathematical problems, high-quality computer graphics, the ability to record formulas in their natural form have made it one of the popular tools for mathematical calculations. An important component of Mathcad is the programming system, which has a language close to traditional professional programming languages, which significantly expands Mathcad's computing capabilities. The advantage of Mathcad is its openness to adding new algorithms, user functions and programs. The work combines both theoretical and computer technologies of calculations from the numerical implementation of mathematical models of the proposed tasks. This gives an opportunity to deepen the knowledge of the theory of statistical analysis, to develop skills in solving problems according to the given computer algorithms, to expand the idea of the areas of application of statistical analysis methods for solving practical problems in the field of economics, production and management.

Keywords: statistical method, complete experiment, information technologies, mathematical model of experiment planning.