

Приходько Н.В.

Національний університет кораблебудування імені адмірала Макарова

Макарова Л.М.

Національний університет кораблебудування імені адмірала Макарова

Кудін О.О.

Національний університет кораблебудування імені адмірала Макарова

НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ ТРУДОМІСТКОСТІ РОБІТ У СУДНОБУДІВНИХ ПРОЕКТАХ НА ОСНОВІ ДВОВИМІРНОГО НОРМАЛІЗУЮЧОГО ПЕРЕТВОРЕННЯ ДЖОНСОНА

У статті вперше побудовано нелінійну регресійну модель для оцінювання трудомісткості робіт з розробки конструкторської документації суден на основі двовимірного нормалізуючого перетворення Джонсона сімейства SB. Двовимірне нормалізуюче перетворення Джонсона враховує кореляцію між випадковими змінними двовимірного негаусівського вектору трудомісткості робіт. Використання двовимірного нормалізуючого перетворення Джонсона сімейства SB дозволяє зменшити довірчі інтервали і інтервали передбачення нелінійної регресії у порівнянні з одновимірним нормалізуючим перетворенням Джонсона сімейства SB та одновимірним нормалізуючим перетворенням на основі десятичного логарифму. Це дає змогу підвищити достовірність оцінювання трудомісткості робіт в проектах розроблення конструкторської документації суден.

Ключові слова: нелінійна регресійна модель, інтервал передбачення нелінійної регресії, двовимірні негаусівські дані, двовимірне нормалізуюче перетворення Джонсона, конструкторська документація судна.

Постановка проблеми. Під час управління часом проектів розроблення конструкторської документації (далі – КД) суден трапляються випадки низької достовірності оцінювання трудомісткості проектно-конструкторських робіт і проектів загалом. Низька достовірність оцінювання трудомісткості робіт обумовлена впливом випадкових зовнішніх факторів та використанням для оцінювання експертного методу, оснований на власному досвіді менеджера проектів. Наприклад, при розробці КД для однієї з секцій судна планова трудомісткість дорівнювала 32 годинам, а фактична трудомісткість склала 98 годин. Враховуючи наявність випадкових впливів та помилок в оцінюванні трудомісткості, виникає потреба у розробці математичних моделей, які б дозволили підвищити достовірність оцінювання трудомісткості робіт з розробки конструкторської документації (КД) суден. Однієї з таких математичних моделей, яка дозволить це зробити, є нелінійна регресійна модель, що містить випадкову складову.

Отже підвищення достовірності оцінювання трудомісткості робіт з розробки КД суден на основі побудови нелінійних регресійних моделей є актуальною проблемою.

Аналіз останніх досліджень і публікацій.

Під час побудови нелінійних регресійних моделей використовуються два підходи. Або методом перебору знаходять нелінійну функцію регресії та на її основі будують нелінійну регресійну модель, або знаходять такі попередні перетворення випадкових змінних регресії (нормалізуючі перетворення), які дозволяють отримати близький до лінійного зв'язок між нормалізованими змінними [1–3].

Під час створення нелінійних регресійних моделей для двовимірних негаусівських даних (наприклад, СОСОМО, ISBSG) використовуються відомі нормалізуючі перетворення на основі десятичного логарифму [5–8] або нормалізуючі перетворення Джонсона [9]. Але ці методи базуються на одновимірних нормалізуючих перетвореннях, які не враховують кореляції між двовимірними випадковими величинами [1].

Використання багатовимірних нормалізуючих перетворень з урахуванням кореляції між негаусівськими випадковими величинами для побудови нелінійних регресійних моделей, як правило, приводить до зменшення ширин довірчих інтервалів і інтервалів передбачення нелінійної регресії [4].

Враховуючи те, що дані по трудомісткості робіт з розробки КД суден є негаусівськими, виникає необхідність у побудові відповідної нелінійної регресійної моделі для оцінювання трудомісткості робіт на основі багатовимірних нормалізуючих перетворень.

Нині розроблено метод [4, с. 101] для побудови нелінійних регресійних моделей на основі багатовимірних нормалізуючих перетворень. Цей метод в загальному випадку вимагає використання обернених матриць. Для двовимірних негаусівських даних побудову нелінійних регресійних моделей можна спростити так же само, як це було зроблено при побудові нелінійних регресійних рівнянь [2].

Постановка завдання. Метою роботи є побудова нелінійної регресійної моделі для оцінювання трудомісткості робіт з розробки КД секцій корпусів суден на основі двовимірного перетворення Джонсона сімейства SB з урахуванням кореляції між трудомісткістю робіт і масою секцій корпусів суден, що дозволить зменшити ширину довірчих інтервалів та інтервалів передбачення нелінійної регресії у порівнянні з існуючими моделями на основі одно-вимірних нормалізуючих перетворень і підвищити достовірність оцінювання трудомісткості робіт з розробки КД секцій корпусів суден при управлінні часом в проектах розробки КД суден.

Виклад основного матеріалу дослідження. Побудова нелінійної регресійної моделі для оцінювання трудомісткості робіт з розробки КД секцій корпусів суден складається з трьох етапів. На першому етапі виконується нормалізація вектора $\mathbf{P} = \{X, Y\}^T$ до гаусівського випадкового вектора $\mathbf{T} = \{Z_X, Z_Y\}^T$ на основі нормалізуючого перетворення

$$\mathbf{T} = \Psi(\mathbf{P}), \quad (1)$$

яке має зворотне перетворення

$$\mathbf{P} = \Psi^{-1}(\mathbf{T}). \quad (2)$$

На другому етапі виконується побудову лінійної регресійної моделі, яку можна представити у вигляді:

$$Z_Y = \hat{Z}_Y + \varepsilon = \hat{b}_0 + \hat{b}_1 Z_X + \varepsilon, \quad (3)$$

де \hat{Z}_Y – результат передбачення лінійної регресії для Z_X ; \hat{b}_0, \hat{b}_1 є оцінками для параметрів лінійної регресії b_0, b_1 ; ε – випадкова складова або похибка з нормальним законом розподілу.

На цьому етапі будують також інтервали передбачення лінійної регресії на основі рівняння [6]:

$$Z_{Y_{in}} = \hat{Z}_Y \pm t_{\alpha/2, N-2} S_{Z_Y} \sqrt{1 + \frac{1}{N} + \frac{(Z_{X_i} - \bar{Z}_X)^2}{S_{Z_X Z_X}}}, \quad (4)$$

де $t_{\alpha/2, N-2}$ – квантиль t-розподілу Стюдента з $\frac{\alpha}{2}$ рівнем значимості і $N-2$ ступенями сво-

$$\text{боди; } S_{Z_Y}^2 = \frac{1}{N-1} \sum_{i=1}^N (\hat{Z}_{Y_i} - \bar{Z}_Y)^2; \quad \bar{Z}_Y = \frac{1}{N} \sum_{i=1}^N \hat{Z}_{Y_i}; \\ S_{Z_X Z_X} = \sum_{i=1}^N (Z_{X_i} - \bar{Z}_X)^2.$$

На третьому етапі виконується побудова нелінійної регресійної моделі трудо+місткості робіт з розробки КД секцій корпусів суден та її інтервалів передбачення Y_{in} на основі зворотного перетворення (2):

$$Y = \psi^{-1} \left[\hat{b}_0 + \hat{b}_1 Z_X + \varepsilon \right]; \quad (5)$$

$$Y_{in} = \psi^{-1} \left[\hat{Z}_Y \pm t_{\alpha/2, N-2} S_{Z_Y} \sqrt{1 + \frac{1}{N} + \frac{(Z_{X_i} - \bar{Z}_X)^2}{S_{Z_X Z_X}}} \right]. \quad (6)$$

У якості двовимірного нормалізуючого перетворення (1) використано перетворення Джонсона, яке має вигляд:

$$\mathbf{T} = \gamma + \eta \mathbf{h}[\lambda^{-1}(\mathbf{P} - \phi)] \sim N_m(0_m, \mathbf{S}), \quad (7)$$

де $\gamma, \eta, \phi, \lambda$ – параметри перетворення Джонсона; $\gamma = (\gamma_X, \gamma_Y)^T$; $\eta = \text{diag}(\eta_X, \eta_Y)$; $\phi = (\phi_X, \phi_Y)^T$; $\lambda = \text{diag}(\lambda_X, \lambda_Y)$; $\mathbf{h}[(g_X, g_Y)] = \{h_1(g_X), h_2(g_Y)\}^T$, а $h_i(\cdot)$ – одна з функцій перетворення:

$$h(g) = \begin{cases} \ln(g), & \text{для сімейства } S_L; \\ \ln[g/(1-g)], & \text{для сімейства } S_B; \\ \text{Arsh}(g), & \text{для сімейства } S_U; \\ g, & \text{для сімейства } S_N; \end{cases}$$

\mathbf{S} – коваріаційна матриця

$$\mathbf{S} = \begin{pmatrix} S_{Z_X}^2 & S_{Z_X Z_Y} \\ S_{Z_X Z_Y} & S_{Z_Y}^2 \end{pmatrix}.$$

$$\text{В (7) } g = (x - \phi)/\lambda; \quad \text{Arsh}(g) = \ln(g + \sqrt{g^2 + 1}).$$

Побудова нелінійної регресійної моделі. При побудові нелінійної регресійної моделі для оцінювання трудомісткості робіт з розробки КД секцій корпусів суден на основі двовимірного нормалізуючого перетворення Джонсона сімейства SB використано дані з табл. 1.

Оцінки параметрів перетворення Джонсона здійснено методом максимальної правдоподібності з використанням логарифмічної функції максимальної правдоподібності, яка отримана на основі щільності двовимірного розподілу Джонсона сімейства SB.

Щільність двовимірного розподілу Джонсона сімейства SB з урахуванням кореляції має вигляд:

$$F_B(Y, X) = \frac{\eta_X \eta_Y \lambda_X \lambda_Y}{2\pi (X - \phi_X)(Y - \phi_Y)(\lambda_X + \phi_X - X)(\lambda_Y + \phi_Y - Y) \sqrt{1 - S_{Z_Y}^2}} \times \\ \times \exp \left\{ -\frac{1}{2} \left[\frac{1}{1 - S_{Z_X Z_Y}^2} [Z_X^2 + Z_Y^2 - 2S_{Z_X Z_Y} Z_X Z_Y] \right] \right\}. \quad (8)$$

Таблиця 1

Вхідні дані для побудови нелінійної регресії

Маса секцій (т)	Тривалість робіт (години)	Маса секцій (т)	Тривалість робіт (години)	Маса секцій (т)	Тривалість робіт (години)
6,15	20	64,58	49	144,78	124
13,08	34	67,24	65	145,69	112
24,83	31	86,64	98	148,06	84
33,97	16	90,32	21	159,06	103
34,15	24	90,43	50	163,5	36
35,07	69	91,25	64	164,27	71
37,07	17	95,31	44	166,01	121
37,14	17	96,54	48	168,53	52
37,15	12	98,35	55	168,68	82
37,22	18	101,06	44	174,14	53
37,25	13	102,35	53	179,03	17
37,65	12	103,14	58	180,14	95
39,81	31	105,19	42	180,94	111
44,15	26	107,37	93	187,88	79
44,36	34	110,69	107	196,15	62
52,39	74	112,36	43	206,16	79
53,91	48	125,43	78	216,81	30
61,13	54	140,51	70	217,82	38

На основі (8) отримано логарифмічну функцію правдоподібності у вигляді:

$$L_B(\theta) = n \ln \eta_X + n \ln \eta_Y + n \ln \lambda_X + n \ln \lambda_Y - n \ln 2\pi - \sum_{i=1}^n \ln(x_i - \phi_X) - \sum_{i=1}^n \ln(y_i - \phi_Y) - \sum_{i=1}^n \ln(\lambda_X + \phi_X - x_i) - \sum_{i=1}^n \ln(\lambda_Y + \phi_Y - y_i) - \frac{1}{2} \sum_{i=1}^n \ln(1 - S_{Z_X Z_Y}^2) - \frac{1}{2} \sum_{i=1}^n \left\{ \frac{1}{1 - S_{Z_X Z_Y}^2} [Z_{X_i}^2 + Z_{Y_i}^2 - 2S_{Z_X Z_Y} Z_{X_i} Z_{Y_i}] \right\}. \quad (9)$$

Після нормалізації двовимірного негаусівського вектора на основі двовимірного нормалізуючого перетворення Джонсона сімейства SB і побудови лінійної регресійної моделі виконано побудову нелінійної регресійної моделі для оцінювання трудомісткості робіт з розробки КД секцій корпусів суден, яка має вигляд:

$$Y = \hat{\phi}_Y + \hat{\lambda}_Y \left[1 + e^{-\frac{(\hat{h}_0 + \hat{h}_X Z_X + \varepsilon - \hat{\gamma}_Y)}{\hat{h}_Y}} \right]^{-1}, \quad (10)$$

де ε – гаусівська випадкова величина з математичним сподіванням 0 і дисперсією 1.

За методом максимальної правдоподібності з використанням функції (9) для двовимірного нормалізуючого перетворення Джонсона сімейства SB отримано наступні оцінки параметрів: $\hat{\gamma}_X = 0,2651$; $\hat{\eta}_X = 0,7792$; $\hat{\phi}_X = 6,209$; $\hat{\lambda}_X = 230,728$; $\hat{\gamma}_Y = 0,0446$; $\hat{\eta}_Y = 0,3808$; $\hat{\phi}_Y = 30,9$; $\hat{\lambda}_E = 48,2$. Оцінки

параметрів лінійної регресії для нормалізованих даних за цим перетворенням знайдено за методом найменших квадратів: $\hat{b}_0 = 0$; $\hat{b}_1 = 0,5055$.

Результати побудови нелінійної регресії за моделлю (10) наведено на рис. 1.

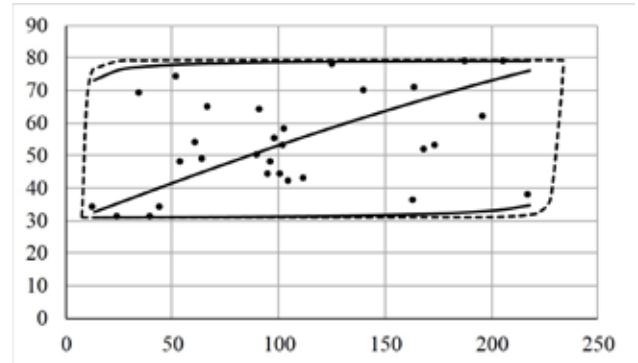


Рис. 1. Нелінійна регресія, побудована з використанням двовимірного нормалізуючого перетворення Джонсона сімейства SB

На рис. 1 по осі абсцис відкладено масу секцій корпусів суден (тони), а по осі ординат – трудомісткість робіт з розробки КД цих секцій (години). Колами позначено вхідні данні. Суцільними лініями зображено відповідну нелінійну регресію і інтервали передбачення цієї регресії, пунктиром – трансформований еліпс передбачення.

Для порівняння отриманих результатів були побудовані відповідні нелінійні регресійні моделі на основі одновимірних нормалізуючих перетворень: у вигляді десяткового логарифму та Джонсона сімейства SB та нормалізуючого перетворення з використанням десяткового логарифму. Результати побудови нелінійних регресій за моделями, що побудовані на основі зазначених одновимірних нормалізуючих перетворень, наведено на рис. 2 і рис. 3.

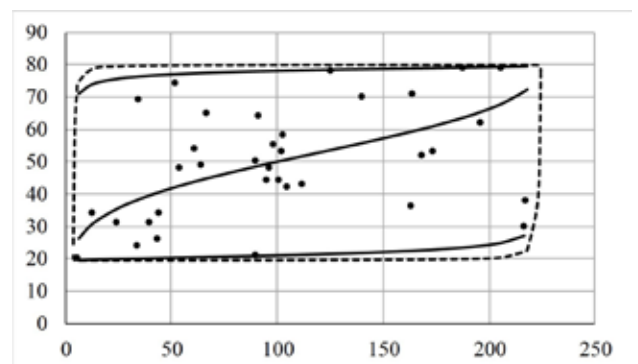


Рис. 2. Нелінійна регресія, побудована з використанням одновимірного нормалізуючого перетворення Джонсона сімейства SB

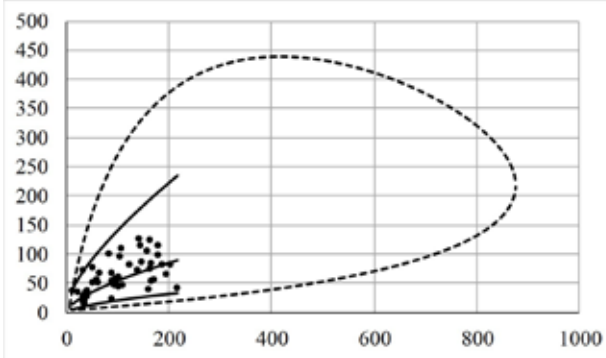


Рис. 3. Нелінійна регресія, побудована з використанням нормалізуючого перетворення на основі десяткового логарифму

На рис. 2, 3, як і на рис. 1, по осі абсцис відкладено масу секцій корпусів суден (тони), а по осі ординат – трудомісткість робіт з розробки КД цих секцій (години). Колами позначено вхідні данні. Суцільними лініями зображено відповідну нелінійну регресію і інтервали передбачення цієї регресії, пунктиром – трансформований еліпс передбачення.

За методом максимальної правдоподібності для одновимірного нормалізуючого перетворення Джонсона сімейства SB отримано такі оцінки параметрів: $\hat{\gamma}_X = 0,1554$; $\hat{\eta}_X = 0,626$; $\hat{\phi}_X = 3,0683$; $\hat{\lambda}_X = 221,9643$; $\hat{\gamma}_Y = -0,0105$; $\hat{\eta}_Y = 0,5302$; $\hat{\phi}_Y = 0,19,4724$; $\hat{\lambda}_Y = 60,3176$. Оцінки параметрів лінійної регресії для нормалізованих даних за цим перетворенням знайдено за методом найменших квадратів: $\hat{b}_0 = 0$; $\hat{b}_1 = 0,4467$.

Як можна побачити з рис. 1-3 Найменша ширина інтервалу передбачення нелінійної регресії є у випадку використання двовимірного нормалізуючого перетворення Джонсона сімейства S_B (рис. 1). На початку нелінійної регресії ширина інтервалів передбачення, побудованих на основі двовимірного перетворення Джонсона сімейства S_B , менше на 17,79% відносно ширини інтервалів передбачення, побудованих на основі одно-

вимірного перетворення Джонсона сімейства S_B , в середині нелінійної регресії – менше на 16,82%, в кінці нелінійної регресії – менше на 14,72%. Приблизно такі ж результати отримано для ширин довірчих інтервалів регресії. Нелінійна регресійна модель на основі двовимірного перетворення Джонсона сімейства S_B виявилась кращою у порівнянні з моделями на основі одновимірних перетворень також за середньою відносною похибкою MMRE і процентом передбачених значень $\text{Pred}(0,25)$. Для двовимірного перетворення Джонсона $\text{MMRE} = 0,2266$ і $\text{Pred}(0,25) = 0,6667$. Для одновимірного перетворення Джонсона $\text{MMRE} = 0,3028$ і $\text{Pred}(0,25) = 0,6571$. Для перетворення на основі десяткового логарифму $\text{MMRE} = 0,4016$; $\text{Pred}(0,25) = 0,4118$.

У випадку двовимірного нормалізуючого перетворення розмір трансформованого еліпсу передбачення менше трансформованих еліпсів передбачення, побудованих на основі одновимірних нормалізуючих перетворень. Найбільший трансформований еліпс передбачення отримано для перетворення на основі десяткового логарифму. Отримані результати пояснюються кращою нормалізацією і врахуванням кореляції між випадковими величинами при використанні двовимірного нормалізуючого перетворення.

Висновки. *Вперше* побудовано нелінійну регресійну модель для оцінювання трудомісткості робіт з розробки КД секцій корпусів суден на основі двовимірного нормалізуючого перетворення Джонсона сімейства SB, що дозволяє зменшити ширину інтервалів передбачення нелінійної регресії і підвищити достовірність оцінювання трудомісткості робіт при управлінні часом проєктів розробки КД суден у порівнянні з моделями, що отримані за одновимірними нормалізуючими. В подальшому планується побудова відповідних нелінійних регресійних моделей за іншими наборами даних.

Список літератури:

1. Гур'янова Л.С., Клебанова Т.С., Сергієнко О.А., Прокопович С.В. Економетрика : навчальний посібник. Харків : ХНЕУ ім. С. Кузнеця, 2015. 384 с.
2. Prykhodko S.B., Prykhodko N.V., Makarova L.M., Kudin O.O., Smykodub T.G. Constructing non-linear regression equations on the basis of bivariate normalizing transformations. *Вісник Херсонського національного технічного університету*. Херсон: ХНТУ, 2017. Вип. 3(62). Т. 1. С. 333–337.
3. Pardoe Iain. *Applied regression modelling*. Wiley, 2012. 325 p.
4. Prykhodko N.V., Prykhodko S.B. Constructing the non-linear regression models on the basis of multivariate normalizing transformations. *Electronic modeling*. 2018. Т. 40. № 6. С. 99–108.
5. Oligny S, Bourque P., Abran A., Fournier B. Exploring the relation between effort and duration in software engineering projects. *In proc. of the World Computer Congress*, Aug. 2000. P. 175–178.
6. Bates D.M., Watts D.G. *Nonlinear regression analysis and its applications*. Wiley, 1988. 384 p.
7. Ryan T.P. *Modern regression methods*. Wiley, 1997. 529 p.

8. Seber G.A.F., Wild C.J. Nonlinear regression. John Wiley & Sons, Inc., 2003. 792 p.
9. Johnson R.A., Wichern D.W. Applied multivariate statistical analysis. Pearson Prentice Hall, 2007. 800 p.
10. Приходько С.Б., Пухалевич А.В. Confidence interval estimation of PC software project duration regression based on Johnson transformation. *Радіоелектронні і комп'ютерні системи*. Харків : Харківський авіаційний інститут, 2014. № 2 (66). С. 104–107.

НЕЛИНЕЙНАЯ РЕГРЕССИОННАЯ МОДЕЛЬ ДЛЯ ОЦЕНИВАНИЯ ТРУДОЕМКОСТИ РАБОТ В СУДОСТРОИТЕЛЬНЫХ ПРОЕКТАХ НА ОСНОВЕ ДВУМЕРНОГО НОРМАЛИЗИРУЮЩЕГО ПРЕОБРАЗОВАНИЯ ДЖОНСОНА

В статье впервые построена нелинейная регрессионная модель для оценивания трудоемкости работ по разработке конструкторской документации судов на основе двумерного нормализующего преобразования Джонсона семейства SB. Двумерное нормализующее преобразование Джонсона семейства SB учитывает корреляцию между случайными переменными двумерного негауссовского вектора трудоемкости работ. Использование двумерного нормализующего преобразования Джонсона семьи SB позволяет уменьшить доверительные интервалы и интервалы предсказания нелинейной регрессии по сравнению с одномерным нормализующим преобразованием Джонсона семейства SB и одномерным нормализующим преобразованием на основе десятичного логарифма. Это дает возможность увеличить достоверность оценивания трудоемкости работ в проектах разработки конструкторской документации судов.

Ключевые слова: нелинейная регрессионная модель, интервал предсказания нелинейной регрессии, двумерные негауссовские данные, двумерное нормализующее преобразование Джонсона, конструкторская документация судна.

NON-LINEAR REGRESSION MODEL FOR ESTIMATING THE EFFORT OF WORKS IN SHIPBUILDING PROJECTS BASED ON THE JONSON BIVARIATE NORMALIZING TRANSFORMATION

Non-linear regression model for estimating the effort of works of the ship design documentation development based on the Johnson bivariate normalizing transformation of SB family is firstly built. The Johnson bivariate normalizing transformation of SB family takes into account the correlation between the random variables of the bivariate non-Gaussian vector of the effort of works. Using the Johnson bivariate normalizing transformation of the SB family allows to reduce the confidence and prediction intervals of the non-linear regression compared to the Johnson univariate normalizing transformation of the SB family and the univariate normalizing transformations based on the decimal logarithm. This makes it possible to increase the reliability of estimating the effort of works in projects of the ship design documentation development.

Key words: non-linear regressions model, non-linear regression prediction interval, bivariate non-Gaussian data, Johnson's bivariate normalizing transformation, ship design documentation.