

УДОСКОНАЛЕННЯ НЕЛІНІЙНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ ДЛЯ МОДЕЛЮВАННЯ ТРУДОМІСТКОСТІ ВИКОНАННЯ ПРОГРАМНИХ ПРОЕКТІВ

Куриленко О.С.¹, Приходько С.Б. д.т.н., професор²
Національний університет кораблебудування імені адмірала Макарова
Україна, Миколаїв

¹os.kurylenko@gmail.com, ²sergiy.prykhodko@nuos.edu

Анотація. Удосконалено нелінійну регресійну модель для моделювання трудомісткості виконання програмних проектів на основі двовимірного перетворення Джонсона сім'ї S_B , що дозволе підвищити достовірність оцінювання залежної змінної регресії у порівнянні з існуючими регресійними моделями.

Ключові слова: нелінійна регресійна модель; моделювання; трудомісткість; програмний проект.

Вступна частина. Розробка програмного забезпечення (ПЗ) сьогодні це складний процес, який не завжди є успішним. Так, за результатами дослідження Standish Group у 2017 р. з понад 50000 програмних проектів лише 36% були успішними, 45% завершилися лише з частковим успіхом (оскаржені), 19% не були завершені взагалі (невдалі).

Однією із причин цього є не достовірне оцінювання трудомісткості виконання програмних проектів. Тому достовірне оцінювання трудомісткості виконання програмних проектів залишається важливою задачею.

На сьогоднішній день при оцінюванні трудомісткості програмних проектів часто використовуються нелінійні регресійні моделі трудомісткості виконання таких проектів в залежності від певних метрик програмного забезпечення, наприклад, таких як кількість строк коду, функціональних точок, сторінок проектної документації, дефектів, тощо. Побудова зазначених моделей (зокрема СОСОМО II) здійснюється на основі нормалізації емпіричних даних за допомогою десятичного логарифму. Але логарифмічне перетворення не завжди дозволяє добре нормалізувати емпіричні дані, що приводить до зменшення достовірності оцінювання трудомісткості програмних проектів. Крім того зазначені моделі, як правило, не враховують негаусівський розподіл випадкової складової нелінійної регресійної моделі, що також веде до зменшення достовірності оцінювання трудомісткості програмних проектів.

Підвищити достовірність моделювання трудомісткості виконання програмних проектів можна за рахунок застосування інших перетворень, наприклад, багатовимірних нормалізуючих перетворень Джонсона [1]. Крім того при побудові існуючих математичних моделей не враховують той факт, що дані, як правило, є двовимірними випадковими величинами з негаусівським розподілом. Ігнорування цього може приводити до суттєвих помилок при оцінюванні трудомісткості виконання програмних проектів [2].

Метою роботи є удосконалення нелінійної регресійної моделі для моделювання трудомісткості виконання програмних проектів на основі двовимірного перетворення Джонсона сім'ї S_B , що дозволе підвищити достовірність оцінювання залежної змінної регресії у порівнянні з існуючими регресійними моделями.

Основна частина. Як і в роботі [2] для моделювання трудомісткості виконання програмних проектів ми побудували нелінійну регресійну модель в залежності від кількості

функціональних точок на основі двовимірного перетворення Джонсона сім'ї S_B . Але для побудови цієї моделі ми додатково застосували ще і ітераційний метод заснований на багатовимірних нормалізуючих перетвореннях та інтервалах передбачення [3].

Нелінійна регресійна модель для моделювання трудомісткості виконання програмних проектів Y у годинах в залежності від кількості функціональних точок X_1 має вигляд

$$Y = \hat{\phi}_Y + \hat{\lambda}_Y \left[1 + e^{-(\hat{Z}_Y + \varepsilon - \hat{\gamma}_Y) / \hat{\eta}_Y} \right]^{-1}, \quad (1)$$

де ε – гаусівська випадкова величина, яка визначає залишки, $\varepsilon \sim N(0,1)$; $\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1$,

$$\hat{b}_0 = 0, \quad \hat{b}_1 = 0,923975; \quad Z_1 = \gamma_1 + \eta_1 \ln \frac{X_1 - \phi_1}{\phi_1 + \lambda_1 - X_1}, \quad \phi_1 < X_1 < \phi_1 + \lambda_1; \quad \hat{\gamma}_Y = 1,97416, \quad \hat{\gamma}_1 = 3,61583,$$

$$\hat{\eta}_Y = 0,860386, \quad \hat{\eta}_1 = 0,925973, \quad \hat{\phi}_Y = 17,4111, \quad \hat{\phi}_1 = 118,627, \quad \hat{\lambda}_Y = 2812,481, \quad \hat{\lambda}_1 = 75516,537.$$

У порівнянні з іншими регресійними моделями ця модель (1) має більші значення коефіцієнту множинної детермінації R^2 та відсотка прогнозованих результатів, менші середні величини відносної похибки та ширини інтервалу передбачення нелінійної регресії. Так, якщо значення R^2 для моделі, що наведено у [2], дорівнює всього 0,566, то для моделі (1) воно значно краще: 0,809.

Висновки. Удосконалено нелінійну регресійну модель в залежності від кількості функціональних точок для моделювання трудомісткості виконання програмних проектів на основі двовимірного перетворення Джонсона сім'ї S_B . Це дозволяє підвищити достовірність оцінювання залежної змінної регресії у порівнянні з існуючими регресійними моделями. Модель, що побудовано на основі двовимірного нормалізуючого перетворення Джонсона сім'ї S_B , в порівнянні з іншими регресійними моделями має більші значення коефіцієнту множинної детермінації та відсотка прогнозованих результатів, менші середні величини відносної похибки та ширини інтервалу передбачення нелінійної регресії.

REFERENCES

- [1] Prykhodko, N.V., & Prykhodko, S.B. (2018). Constructing the non-linear regression models on the basis of multivariate normalizing transformations. *Electronic modeling*, 6(40), 101-110. DOI: <https://doi.org/10.15407/emodel.40.06.101>
- [2] Prykhodko, S.B., Prykhodko, N.V., Makarova, L.M., Kudin, O.O., & Smykodub, T.G. (2017). Constructing the non-linear regression equations on the basis of bivariate normalizing transformations. *Visnyk of Kherson National Technical University*, 3(62), 333-337.
- [3] Prykhodko, S., & Prykhodko, N. (2021). Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations. *Advances in Intelligent Systems and Computing*, 1265, 166-174. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-58124-4_16

Kurylenko O.S., Prykhodko S.B.

Improving the nonlinear regression model for modeling of efforts of developing software projects

Abstract. The nonlinear regression model for modeling of efforts of developing software projects based on the bivariate Johnson normalizing transformations for S_B family is improved, which improves the confidence of the dependent variable regression estimate compared to the existing regression models.

Keywords: nonlinear regression model; modeling; efforts; software project.