

На основі нормалізуючого перетворення Джонсона сімей SB та SU побудовано нелінійну регресійну модель тривалості виконання проектів підготовки жінок-боксерів високої кваліфікації до змагань в залежності від тривалості тренувань. Виконано порівняння побудованої моделі з моделлю на основі логарифмічного перетворення. Рекомендовано використання регресійної моделі на основі перетворення Джонсона для оцінювання тривалості проектів підготовки жінок-боксерів до змагань

Ключові слова: управління часом, перетворення Джонсона, регресійна модель, логарифмічне перетворення

На основе нормализующего преобразования Джонсона семейств SB и SU построена регрессионная модель длительности проектов подготовки женщин-боксеров высокой квалификации к соревнованиям в зависимости от длительности тренировки. Выполнено сравнение построенной модели с моделями на основе логарифмического преобразования. Рекомендовалось использование модели на основе преобразования Джонсона для оценивания длительности проектов подготовки женщин-боксеров к соревнованиям

Ключевые слова: управление временем, преобразование Джонсона, регрессионная модель, логарифмическое преобразование

УДК 005.8
DOI: 10.15587/1729-4061.2015.38674

РОЗРОБКА РЕГРЕСІЙНОЇ МОДЕЛІ ТРИВАЛОСТІ ВИКОНАННЯ ПРОЕКТІВ ПІДГОТОВКИ ЖІНОК- БОКСЕРІВ ДО ЗМАГАНЬ

С. Б. Приходько

Доктор технічних наук, професор, завідувач кафедри*

E-mail: sergiy.prykhodko@nuos.edu.ua

Н. В. Князь

Аспірант*

E-mail: natalia.knyaz@nuos.edu.ua

*Кафедра програмного забезпечення
автоматизованих систем
Національний університет
кораблебудування ім. адм. Макарова
пр. Героїв Сталінграда, 9, м. Миколаїв, Україна, 54025

1. Вступ

Бокс – один з найбільш масових олімпійських видів спорту. Велика популярність боксу пояснюється його видовищністю, високим емоціональним напруженням спортивної боротьби і різнобічним впливом на рухові, психічні та вольові якості людини. З кожним роком жіночий бокс набуває все більшого поширення в нашій країні. Щороку ряди боксерів високої кваліфікації поповнюють чимало майстерних молодих спортсменів, які домагаються видатних успіхів на рингу. Збільшений рівень конкурентної боротьби на найбільших міжнародних турнірах викликає необхідність вдосконалення засобів і методів підготовки боксерів високого класу. На сучасному етапі є актуальним проведення досліджень, які підвищать ефективність підготовки жінок-боксерів високої кваліфікації до змагань, тому що це дозволить покращити спортивну майстерність боксерів і досягти кращих результатів на змаганнях. Одним з напрямів досліджень є визначення тривалості підготовки спортсменів до змагань.

2. Аналіз літературних даних та постановка проблеми

При підготовці жінок-боксерів високої кваліфікації до змагань існує проблема визначення часу для

виводу спортсменів на пік форми перед важливими боями. На сьогодні не враховується, по-перше, те, що підготовка до змагань є певним проектом і потребує управління низкою ресурсів, зокрема часом, а, по-друге, що час підготовки є випадковою величиною, яка залежить від ряду факторів, в тому числі і від тривалості тренувань. Подібних залежностей у вигляді математичних моделей на сьогодні не розроблено. А існують лише загальні рекомендації наближених строків підготовки жінок-боксерів до змагань, а самі строки, як правило, є точковими оцінками, що отримані експертним шляхом [1]. Експертні оцінки є досить суб'єктивними.

Згідно [2] замість експертного оцінювання тривалості проекту можна також застосувати параметричне оцінювання, наприклад, використовуючи регресійні моделі. Проте для оцінювання тривалості підготовки жінок-боксерів до змагань неможливо побудувати адекватну лінійну регресійну модель, тому що, як показують дослідження, розподіл тривалості тренувань або тривалості проектів підготовки жінок-боксерів до змагань не є нормальним [3]. Через це виникає необхідність побудови нелінійної регресійної моделі. На сьогоднішній час для побудови нелінійної регресійної моделі є такі методи: метод простого перебору, методи лінеаризуючих та нормалізуючих перетворень. Метод простого перебору вимагає завдання різних видів

рівняння регресії [4, 5] і вибору найкращого наближення із заданих за певним критерієм [6], тому краще використовувати перетворення [7, 8], в першу чергу нормалізуючі. У якості таких перетворень зазвичай застосовують логарифмічні перетворення, але вони не завжди дозволяють зробити добру нормалізацію. З цієї причини краще використовувати інші перетворення, наприклад, нормалізуюче перетворення Джонсона. В [3] для нормалізації тривалості проектів підготовки жінок-боксерів до змагань було обрано саме таке перетворення Джонсона.

Тому і виникає необхідність побудови нелінійних моделей тривалості підготовки жінок-боксерів до змагань на основі перетворення Джонсона.

3. Мета та задачі дослідження

Мета роботи полягає у побудові нелінійної регресійної моделі тривалості виконання проектів підготовки жінок-боксерів високої кваліфікації до змагань в залежності від тривалості тренувань з використанням перетворення Джонсона.

Для досягнення поставленої мети вирішувалися наступні задачі:

- вибір сім'ї перетворення Джонсона;
- знаходження параметрів перетворення Джонсона;
- нормалізація емпіричних даних;
- побудова лінійної регресійної моделі тривалості проектів підготовки жінок-боксерів високої кваліфікації до змагань за нормалізованими даними;
- побудова нелінійної регресійної моделі тривалості проектів підготовки жінок-боксерів високої кваліфікації до змагань.

4. Матеріали та методи досліджень розробки регресійної моделі тривалості виконання проектів підготовки жінок-боксерів до змагань

Побудова нелінійної регресійної моделі тривалості проектів підготовки жінок-боксерів високої кваліфікації до змагань в залежності від його тривалості тренувань виконується в 3 етапи, так як було зроблено в [9]. Спочатку емпіричні дані нормалізуються з використанням нормалізуючого перетворення; далі за нормалізованими даними будується лінійна регресійна модель; на останньому етапі з лінійної моделі, застосовуючи обране нормалізуюче перетворення, отримується нелінійна регресійна модель.

Для побудови регресійної моделі було використано емпіричні дані тривалості тренувань та емпіричні дані тривалості підготовки до 35 змагань з боксу серед жінок майстра спорту міжнародного класу Н. В. Князь, яка є членом національної збірної України. Значення випадкової величини x – тривалості тренувань (у хвиликах) такі: 558, 572, 580, 606, 629, 667, 687, 701, 710, 726, 797, 827,

837, 920, 960, 972, 1029, 1104, 1127, 1130, 1139, 1158, 1175, 1212, 1229, 1233, 1305, 1362, 1421, 1495, 1517, 1551, 1600, 1790, 1958. Відповідні їм значення випадкової величини y – тривалості підготовки до змагань (у днях) були такими: 6, 8, 6, 10, 8, 11, 11, 8, 8, 14, 10, 11, 10, 16, 15, 13, 12, 21,15, 14, 14, 12, 13, 16, 14, 16, 14, 20, 13, 14, 14, 15, 13, 24, 17. За наведеними значеннями x були обчислені наступні оцінки: $\bar{x} = 1065,26$; $S_x = 369,94$; $\hat{\epsilon}_x = 2,56$; $\hat{A}_x = 0,49$. За відповідними значеннями y були обчислені такі оцінки: $\bar{y} = 13,03$; $S_y = 3,97$; $\hat{\epsilon}_y = 3,85$; $\hat{A}_y = 0,51$. Значення оцінок ексцесу і асиметрії вказують на те, що величини x і y є негаусівськими.

Отримане значення критерію Пірсона при перевірці гіпотези про нормальність закону розподілу випадкової величини x складає $\chi^2 = 9,52$ (критичне значення $\chi_{кр}^2 = 5,99$, $\nu = 5 - 2 - 1 = 2$, $\alpha = 0,05$). Гістограма випадкової величини x та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 1.

Для випадкової величини y значення критерію Пірсона складає $\chi^2 = 3,18$ (критичне значення $\chi_{кр}^2 = 5,99$, $\nu = 5 - 2 - 1 = 2$, $\alpha = 0,05$). Гістограма випадкової величини y та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 2.

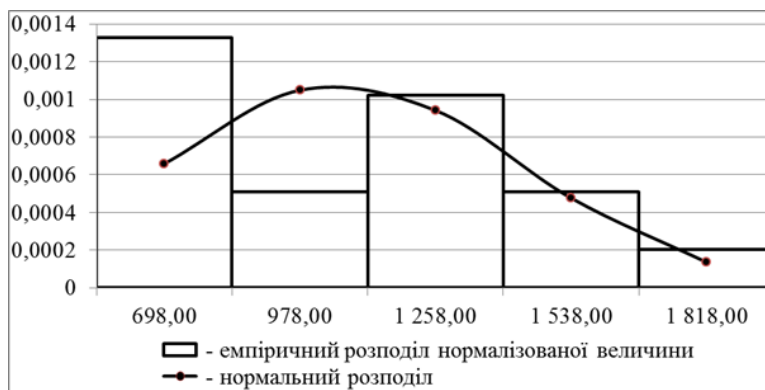


Рис. 1. Розподіл випадкової величини x - тривалості тренувань в проектах підготовки жінок-боксерів до змагань

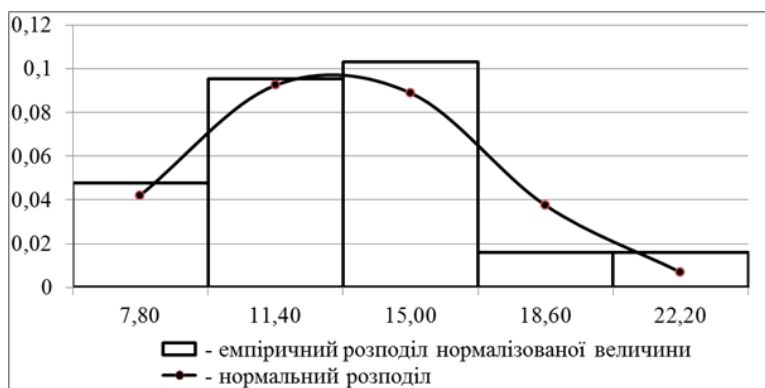


Рис. 2. Розподіл випадкової величини y – тривалості підготовки до змагань

Для нормалізації використаємо перетворення Джонсона, яке в загальному випадку має вигляд [10]

$$z = \gamma + \eta h(x, \phi, \lambda); \eta > 0; -\infty < \gamma < \infty; \lambda > 0; -\infty < \phi < \infty, \quad (1)$$

де z – нормально розподілена випадкова величина з математичним сподіванням нуль і дисперсією одиниця; x – випадкова величина з розподілом Джонсона;

$\gamma, \eta, \phi, \lambda$ – параметри перетворення або розподілу Джонсона; h – функція з певної сім'ї:

$$h = \begin{cases} \ln(\tilde{x}), & x > \phi & \text{для сім'ї } S_U; \\ \ln[\tilde{x}/(1-\tilde{x})], & \phi < x < \phi + \lambda & \text{для сім'ї } S_B; \\ \text{Arsh}(\tilde{x}), & -\infty \leq x \leq +\infty & \text{для сім'ї } S_U, \end{cases}$$

де $\tilde{x} = (x - \phi) / \lambda$; $\text{Arsh}(\tilde{x}) = \ln(\tilde{x} + \sqrt{\tilde{x}^2 + 1})$.

Сім'я розподілу Джонсона обирається за значеннями асиметрії та ексцесу [11, 12]. Оцінки параметрів для обраної сім'ї перетворення (1) можна знайти шляхом рішення наступної задачі математичного програмування [10]:

$$\hat{\theta} = \arg \min_{\theta} \left\{ A_z^2 + (\epsilon_z - 3)^2 + \bar{z}^2 + (S_z^2 - 1)^2 \right\}, \quad (2)$$

де $\theta = \{\gamma, \eta, \phi, \lambda\}$;

$$A_z = \frac{1}{n S_z^3} \sum_{i=1}^n (z_i - \bar{z})^3; \quad \epsilon_z = \frac{1}{n S_z^4} \sum_{i=1}^n (z_i - \bar{z})^4;$$

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i; \quad S_z^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2;$$

z_i – i -значення величини z_u вибірці довжиною $n, i \in [1, n]$.

За значеннями асиметрії та ексцесу для нормалізації випадкової величини x було вибрано перетворення Джонсона (1) сім'ї S_B . Згідно з [3] для нормалізації випадкової величини y було вибрано перетворення Джонсона (1) сім'ї S_U . За рішенням задачі (2) було знайдено параметри Джонсона (1) γ, η, ϕ і λ для нормалізації випадкових величин x та y . Значення параметрів перетворення Джонсона (1) із сім'ї S_B для нормалізації величини x є такими: $\gamma_x = 0,59, \eta_x = 0,71, \phi_x = 517,04, \lambda_x = 1546,29$. Значення параметрів перетворення Джонсона (1) із сім'ї S_U для нормалізації величини y є такими: $\gamma_y = -0,96, \eta_y = 2,50, \phi_y = 9,42, \lambda_y = 8,54$. Було отримано нормалізовані випадкові величини z_x та z_y . Статистичні характеристики величин z_x були: $\bar{z}_x = 0; S_{z_x} = 1; \hat{\epsilon}_x = 3; \hat{A}_{z_x} = 0$; та z_y : $\bar{z}_y = 0; S_{z_y} = 1; \hat{\epsilon}_y = 3; \hat{A}_{z_y} = 0$.

Для випадкової величини z_x значення критерію Пірсона складає $\chi^2 = 1,16$ (критичне значення $\chi_{кр}^2 = 5,99, \nu = 5 - 2 - 1 = 2, \alpha = 0,05$). Гістограма отриманої випадкової величини z_x та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 3.

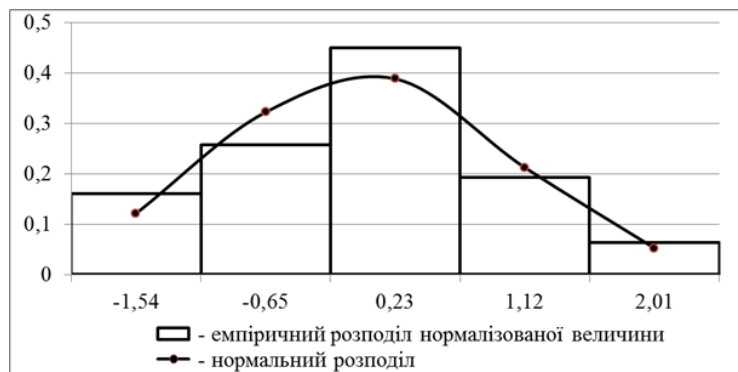


Рис. 3. Розподіл значень випадкової величини z_x – нормалізованих за перетворенням Джонсона

Для випадкової величини z_y значення критерію Пірсона складає $\chi^2 = 5,81$ (критичне значення $\chi_{кр}^2 = 5,99, \nu = 5 - 2 - 1 = 2, \alpha = 0,05$). Гістограма отриманої випадкової величини z_y та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 4.

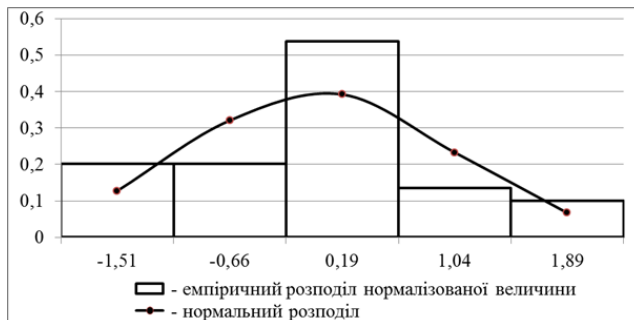


Рис. 4. Розподіл значень випадкової величини z_y – нормалізованих за перетворенням Джонсона

Так як розподіл величин z_x та z_y є нормальним, то це дозволяє побудувати лінійну регресійну модель для нормалізованих значень:

$$z_y = b_0 + b_1 z_x, \quad (3)$$

де b_0 і b_1 – коефіцієнти лінійної регресійної моделі.

Значення коефіцієнтів b_0 і b_1 лінійної регресійної моделі (3) знаходяться за методом найменших квадратів:

$$b_1 = \frac{\overline{z_x z_y} - \bar{z}_x \cdot \bar{z}_y}{\overline{z_x^2} - \bar{z}_x^2}; \quad b_0 = \bar{z}_y - b_1 \cdot \bar{z}_x, \quad (4)$$

$$\text{де } \bar{z}_x = \frac{1}{n} \sum_{i=1}^n z_{xi}; \quad \bar{z}_y = \frac{1}{n} \sum_{i=1}^n z_{yi}; \quad \overline{z_x z_y} = \frac{1}{n} \sum_{i=1}^n z_{xi} z_{yi}.$$

Отримані за (4) значення коефіцієнтів лінійної регресійної моделі були такими: $b_0 = 0$ і $b_1 = 0,79$. Лінійна регресійна модель (3) з отриманими значеннями коефіцієнтів показана на рис. 5.

Враховуючи (3) і те, що $z_y = \gamma_y + \eta_y \text{Arsh}\left(\frac{y - \phi_y}{\lambda_y}\right)$, то нелінійна негаусівська регресійна модель тривалості проектів підготовки жінок-боксерів до змагань в залежності від тривалості тренувань може бути представлена як

$$y = \lambda_y + \phi_y \text{Sh}\left(\frac{b_0 + b_1 z_x - \gamma_y}{\eta_y}\right), \quad (5)$$

$$\text{де } z_x = \gamma_x + \eta_x \ln\left(\frac{x - \phi_x}{\lambda_x + \phi_x - x}\right).$$

Для побудови моделі на основі логарифмічного перетворення виконано нормалізацію величини x та z з використанням цього перетворення.

Для випадкової величини z_x отриманої значення критерію Пірсона складає $\chi^2 = 6,84$ (критичне значення $\chi_{кр}^2 = 5,99, \nu = 5 - 2 - 1 = 2,$

$\alpha = 0,05$). Гістограма отриманої випадкової величини z_x отриманої з використанням логарифмічного перетворення та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 6.

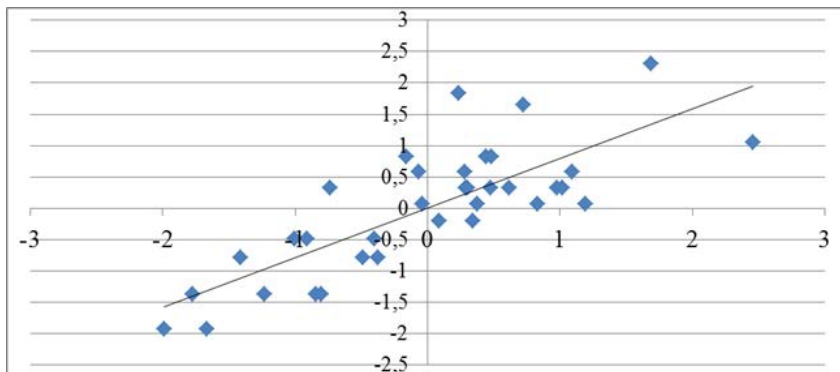


Рис. 5. Лінійна регресійна модель для нормалізованих значень тривалості проектів підготовки жінок-боксерів до змагань в залежності від нормалізованих значень тривалості тренувань

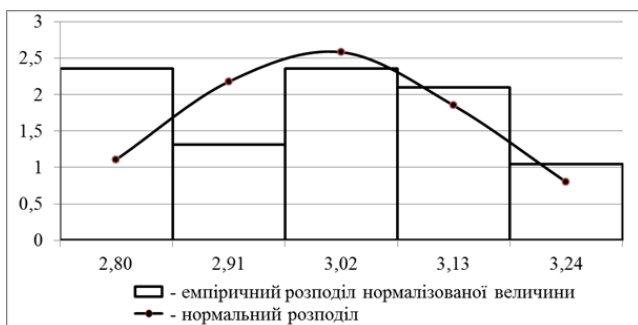


Рис. 6. Розподіл значень випадкової величини z_x – нормалізованих за логарифмічним перетворенням

Для випадкової величини z_y отриманої з використанням логарифмічного перетворення значення критерію Пірсона складає $\chi^2 = 3,57$ (критичне значення $\chi_{кр}^2 = 5,99$, $\nu = 5 - 2 - 1 = 2$, $\alpha = 0,05$). Гістограма отриманої випадкової величини z_y отриманої з використанням логарифмічного перетворення та її теоретичний розподіл у вигляді щільності ймовірності наведені на рис. 7.

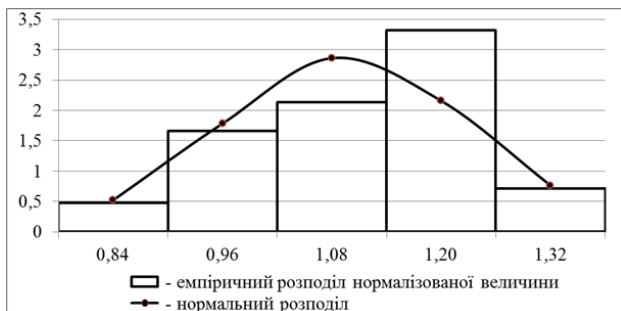


Рис. 7. Розподіл значень випадкової величини z_y – нормалізованих за логарифмічним перетворенням

Якщо виконувати нормалізацію за логарифмічним перетворенням (тобто $z_x = \log_{10} x$, $z_y = \log_{10} y$), то, враховуючи (3), нелінійна негаусівська регресійна модель

тривалості проектів підготовки жінок-боксерів до змагань в залежності від тривалості тренувань може бути представлена як

$$y = 10^{b_0 x^{b_1}}, \quad (6)$$

де b_0 і b_1 – коефіцієнти лінійної регресійної моделі (3) для нормалізованих значень за логарифмічним перетворенням.

Значення коефіцієнтів цієї регресійної моделі (6) були такими: $b_0 = -1,23$ і $b_1 = 0,77$.

5. Результати досліджень розробки регресійної моделі тривалості виконання проектів підготовки жінок-боксерів до змагань

Побудована нелінійна негаусівська регресійна модель тривалості проектів підготовки жінок-боксерів до змагань в залежності від тривалості тренувань на основі перетворення Джонсона та логарифмічного перетворення показані на рис. 8.

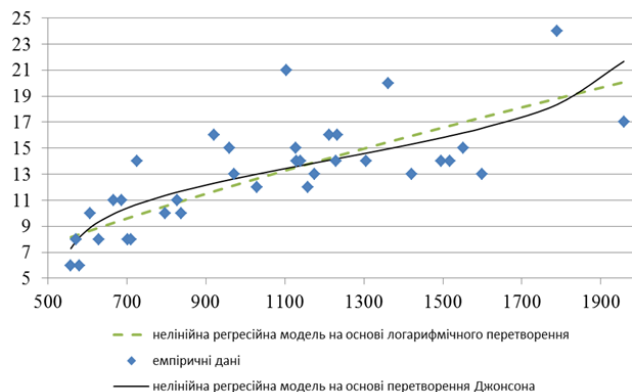


Рис. 8. Нелінійна негаусівська регресійна модель тривалості підготовки жінок-боксерів до змагань в залежності від тривалості тренувань

Модель на основі перетворення Джонсона і на основі логарифмічного перетворення практично співпадають, тому для вибору кращої регресійної моделі потрібно використати метод математичної статистики.

6. Обговорення результатів розробки регресійної моделі тривалості виконання проектів підготовки жінок-боксерів до змагань

Для того щоб порівнювати нелінійні регресійні моделі можна використовувати такий критерій як сума квадратів відхилень між тривалістю, прогнозованою за моделлю та емпіричними даними:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}(x))^2, \quad (7)$$

де n – це кількість емпіричних даних; y_i – фактичне значення випадкової величини y ; $\hat{y}(x)$ – прогнозоване значення випадкової величини y .

Кращою є та модель, для якої значення SSE є меншим.

Також для порівняння регресійних моделей можна використовувати значення коефіцієнта детермінації R^2 :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}(x))^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (8)$$

де n – це кількість емпіричних даних; y_i – фактичне значення випадкової величини y ; $\hat{y}(x)$ – прогнозоване значення випадкової величини y ; \bar{y} – середнє значення

випадкової величини y , $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Кращою є та модель, для якої значення R^2 є більшим.

Значення (7) і (8) для отриманих нелінійних негаусівських регресійних моделей на основі логарифмічного перетворення (6) і перетворення Джонсона (7) були наступні: сума квадратів відхилень між тривалістю, прогнозованою за моделлю та емпіричними даними (7) була $SSE = 235,72$ для моделі (5) і $SSE = 245,59$ для

моделі (6); коефіцієнт детермінації (8) був $R^2 = 0,56$ для моделі (5) і $R^2 = 0,54$ для моделі (6).

7. Висновки

В роботі вибрано перетворення Джонсона сім'ї SB для нормалізації емпіричних даних тривалості тренувань та перетворення Джонсона сім'ї Su для нормалізації емпіричних даних тривалості підготовки жінок-боксерів до змагань, та знайдено параметри цих перетворень. Було виконано нормалізацію емпіричних даних за обраним перетворенням Джонсона зі знайденими параметрами. За нормалізованими даними було побудовано лінійну регресійну модель. Далі за лінійною моделлю було побудовано нелінійну регресійну модель тривалості проектів підготовки жінок-боксерів високої кваліфікації до змагань в залежності від їх тривалості тренувань з використанням перетворення Джонсона. Побудована модель на основі перетворення Джонсона (5) краща за модель на основі логарифмічного перетворення (6) як за коефіцієнтом детермінації, так і за сумою квадратів відхилень. Тому при оцінюванні тривалості проектів підготовки жінок-боксерів високої кваліфікації до змагань краще використовувати нелінійну регресійну модель побудовану за перетворенням Джонсона. В подальших дослідженнях планується побудова довірчого інтервалу нелінійної регресії за алгоритмом, який наведений в роботі [10].

Література

1. Ділення, М. О. Бокс: Жінки. Навчальна програма для дитячо-юнацьких спортивних шкіл, спеціалізованих дитячо-юнацьких шкіл олімпійського резерву, шкіл вищої спортивної майстерності [Текст] / М. О. Ділення, В. Н. Ост'янов, Г. І. Комісаренко. – К.: Національний Олімпійський комітет України, 2009. – 112 с.
2. A guide to the project management body of knowledge. Fifth edition [Text] / Project Management Institute, 2013. – 589 p.
3. Приходько, С. Б. Вибір нормалізуючого перетворення для інтервального оцінювання часу виконання проектів підготовки боксерів-жінок до змагань [Текст] / С. Б. Приходько, Н. В. Князь // 36. наук. праць НУК. – 2014. – № 5 (455). – С. 87–91.
4. Демиденко, Е. З. Линейная и нелинейная регрессии [Текст] / Е. З. Демиденко. – М.: Финансы и статистика, 1981. – 302 с.
5. Bates, D. M. Nonlinear Regression Analysis and Its Applications [Text] / D. M. Bates, D. G. Watts. – Wiley, 1988. – 384 p.
6. Pardoe, I. Applied regression modeling [Text] / I. Pardoe. – Wiley, 2012. – 325 p. doi: 10.1002/9781118345054
7. George, A. F. Nonlinear Regression [Text] / A. F. George, C. J. Seber, Wild. – John Wiley & Sons, Inc., 2003. – 792 p.
8. Ryan, T. P. Modern Regression Methods [Text] / T. P. Ryan – Wiley, 2008. – 672 p.
9. Приходько, С. Б. Розробка нелінійних регресійних моделей тривалості програмних проектів на основі перетворення Джонсона [Текст] / С. Б. Приходько, А. В. Пухалевич // Збірник наукових праць НУК. – 2014. – № 2 (2014). – С. 76–80.
10. Приходько, С. Б. Інтервальне оцінювання статистичних моментів негаусівських випадкових величин на основі нормалізуючих перетворень [Текст] / С. Б. Приходько // Науковий журнал “Математичне моделювання”. – 2011. – № 1 (24). – С. 9–13.
11. Johnson, N. L. System of Frequency Curves Generated by Methods of Translation [Text] / N. L. Johnson // Biometrika. – 1949. – Vol. 36, Issue 1/2. – P. 149–176. doi: 10.2307/2332539
12. Приходько, С. Б. Аналитическая зависимость для выбора распределения Джонсона семейства SL [Текст] / С. Б. Приходько, Л. Н. Макарова // Вестник ХНТУ. – Херсон : ХНТУ, 2012. – № 2 (45). – С. 101–104.