

## ЗАСТОСУВАННЯ МОВИ R ДЛЯ АНАЛІЗУ СТАТИСТИЧНИХ ПОКАЗНИКІВ НАВЧАЛЬНОЇ УСПІШНОСТІ СТУДЕНТІВ

**Павленко Лілія Василівна**

Кандидат педагогічних наук, доцент  
(Бердянський державний педагогічний університет)

**Павленко Максим Петрович**

Кандидат педагогічних наук, доцент  
(Бердянський державний педагогічний університет)

**Венецький Дмитро Сергійович**

Студент факультету фізико-математичної, комп'ютерної та технологічної освіти  
(Бердянський державний педагогічний університет)

**Актуальність.** Пошук шляхів і засобів підвищення якості освіти – предмет багатьох педагогічних досліджень. Застосування статистичних методів для аналізу якості навчання дозволяє отримати відомості про основні закономірності процесу навчання у вищому навчальному закладі, виявити зв'язок різних компонентів цього процесу між собою, визначити чинники, що вимагають першочергової уваги. Проте, комерційні продукти світового рівня типу Statistica вимагають істотних матеріальних витрат. Тому актуальним є вибір і використання вільно розповсюджуваних програмних продуктів, що мають ліцензією GNU General Public License.

**Мета і методи дослідження.** Визначити основні можливості мови R для аналізу статистичних показників навчальної успішності студентів

**Сутність дослідження.** Одним з продуктів, що відповідають високим сучасним вимогам, є RStudio – вільне середовище розробки програмного забезпечення з відкритим вихідним кодом для мови програмування R, яке призначене для статистичної обробки даних і роботи з графікою.

До його переваг можна віднести:

- обробка великих масивів даних;
- графічні можливості надання результатів досліджень;
- можливість самостійного написання необхідних функцій;

Недоліки, що відзначаються рядом користувачів:

– на відміну від більшості комерційних програм, R не має графічного інтерфейсу, тому потрібно знати необхідні для роботи функції і синтаксис мови програмування.

Для оцінки ступеня застосовності середовища моделювання R можна проаналізувати показники навчальної успішності студентів.

Успішність студентів – це багатовимірне явище, тому його дослідження необхідно проводити із застосуванням багатовимірних методів, зокрема, прийомів кореляційно-регресійного аналізу. Як правило, незалежні змінні в регресійних моделях мають безперервні області зміни. Однак в дослідженні

успішності потрібно брати до уваги якісні ознаки, які можуть приймати всього два значення або дискретну множину значень.

В якості незалежних змінних були відібрані:

- шкільний середній бал;
- наявність роботи (так – 1, ні – 0);
- кількість вільного часу (мало – 1, багато – 5)
- місце проживання (у будинку з батьками – 1, інше 0);
- фінансове становище (погано – 1, добре – 5);
- мета навчання (диплом – 1, інше 0);
- здоров'я студента (погано – 1, добре – 5);

Для отримання регресійної моделі успішності студентів був обраний метод покрокової регресії, який дозволяє з множини незалежних змінних відібрати тільки ті, які найбільш значимі для адекватного опису багатопараметричної регресії.

Побудувати модель лінійної регресії можна за допомогою функції `lm`, яка приймає у якості параметрів список назв змінних, які будуть враховуватися при побудові моделі.

Отримані коефіцієнти дозволяють зробити висновок, що найбільший вплив на успішність надає фактор «шкільний середній бал», оскільки основи успішності студента у вищому навчальному закладі закладаються ще в школі. Другим за значимістю стимулом до навчання є можливість отримання стипендії (фінансове становище).

**Висновок:** Визначені основні основні можливості мови R для аналізу статистичних даних шляхом побудови регресійної моделі навчальної успішності студентів.

#### Література

1. Храмов Д. А. Сбор данных в интернете на языке R. – М.: ДМК Пресс, 2017. – 283 с.
2. А. Б. Шипунов, Е. М. Балдин. Наглядная статистика. Используем R! [Електроний ресурс] // <http://ashipunov.info/shipunov/school/books/rbook.pdf>