

УДК 378:004.451.5

DOI: 10.31866/2617-796X.7.2.2024.317729

**Андрій Ясінський,**

кандидат педагогічних наук,  
доцент кафедри математичного моделювання,  
Приватний вищий навчальний заклад  
«Міжнародний економіко-гуманітарний університет  
імені академіка Степана Дем'янчука»,  
Рівне, Україна  
yasinskiy@meta.ua  
<https://orcid.org/0000-0002-1894-1314>

**Людмила Соловей,**

старший викладач кафедри інформаційних систем  
та обчислювальних методів,  
Приватний вищий навчальний заклад  
«Міжнародний економіко-гуманітарний університет  
імені академіка Степана Дем'янчука»,  
Рівне, Україна  
lyuda\_solovej@ukr.net  
<https://orcid.org/0009-0001-2832-1741>

## ЗАСТОСУВАННЯ СЕРЕДОВИЩА RSTUDIO ПРИ ДОСЛІДЖЕННІ ЗАДАЧ З МАТЕМАТИКИ, ФІЗИКИ ТА ЕКОНОМІКИ

**Мета статті** – дослідження можливостей застосування середовища RStudio для вирішення завдань з математики, фізики та економіки. Зокрема, акцент зроблено на інтеграції інструментів статистичного аналізу, використання набору математичних функцій та операцій, що забезпечують ефективність та точність у проведенні розрахунків.

**Методи дослідження:** проведено аналіз наукових праць, методичних рекомендацій та прикладів застосування RStudio у різних галузях знань, що стосуються математики, фізики та економіки, випробувано функціонал середовища RStudio на практичних завданнях, здійснено порівняння отриманих результатів із застосуванням RStudio з результатами, отриманими іншими методами, для оцінки ефективності середовища.

**Наукова новизна** полягає у використанні функціональних можливостей R, що дозволяє розробляти та впроваджувати нові методи та підходи до аналізу даних та розв'язання наукових проблем.

**Висновки.** Середовище RStudio демонструє свою ефективність для вирішення задач з математики, фізики та економіки завдяки широкому набору функцій для чисельного моделювання, статистичного аналізу та візуалізації даних. Використання RStudio дозволяє суттєво спростити процес розв'язання складних математичних задач, зокрема сприяє інтеграції математичних методів у дослідження фізичних явищ і економічних процесів, що дозволяє формувати нові підходи до аналізу даних і моделювання.

RStudio може бути використане як ефективний інструмент у навчанні, сприяючи формуванню навичок аналітичного мислення та програмування у студентів, які займаються математичними, фізичними та економічними задачами. Зокрема застосування критерію Тьюкі та функцій середовища RStudio для порівняння даних у великих масивах, таких як результати експериментів або вимірювань. Результати дослідження підтверджують доцільність впровадження RStudio в практику навчання та наукових досліджень, а також вказують на необхідність подальшого дослідження можливостей середовища у вирішенні задач наукових галузей.

**Ключові слова:** мова програмування R; RStudio; моделювання; статистичні дані; дисперсійний аналіз; критерій Тьюкі.

**Вступ.** Сучасний розвиток науки і техніки вимагає застосування потужних інструментів для аналізу даних, моделювання та візуалізації результатів досліджень. Середовище RStudio, яке базується на мові програмування R, стало одним із провідних інструментів у багатьох галузях, зокрема у математиці, фізиці та економіці. Його популярність пояснюється широким спектром функцій для обробки даних, підтримкою чисельних методів і розвинутими засобами графічного відображення результатів.

На сьогодні важливим завданням є пошук ефективних підходів до інтеграції RStudio в освітній процес та наукові дослідження. Зокрема, використання цього середовища у міждисциплінарних задачах дозволяє оптимізувати дослідницьку діяльність, об'єднуючи методи аналізу з різних галузей знань.

У дослідженні розглянуто можливості використання RStudio для вирішення прикладних задач із математики, фізики та економіки. Особливу увагу приділено аналізу його функціональних можливостей, зокрема статистичного аналізу та опрацюванню даних з використанням великого набору математичних функцій та операцій. Результати дослідження мають на меті сприяти поширенню практичного застосування RStudio у науці, навчанні та професійній діяльності.

**Результати дослідження.** Сучасний етап формування інформаційного суспільства характеризується високим ступенем інтеграції інформаційних технологій у систему управління економічними процесами. Постійно зростають вимоги щодо збору, обробки, аналізу, систематизації і накопиченню інформації. Фактор розширення інформаційного простору, забезпечення вільного інформаційного обміну стає умовою сталого розвитку суспільства.

У роботі (Горбачук, 2018) досліджується питання розробки комп'ютерно-орієнтованої методики викладання математичної статистики з метою формування у майбутніх фахівців умінь застосовувати сучасні інструменти математичної статистики у професійній діяльності для вирішення прикладних економічних задач із використанням інформаційних технологій та спеціалізованого програмного забезпечення.

Дослідження авторів (Грод, Андреева, 2019) підкреслює роль RStudio як сучасного інструменту, що забезпечує інтеграцію навчання теоретичної математики з її прикладним використанням.

Сучасний розвиток науки та освіти вимагає від фахівців не лише глибоких теоретичних знань, але й володіння інструментами аналізу, моделювання та візуалізації даних. Це особливо актуально у галузях, які зосереджені на точних розрахунках та аналізі даних, таких як математика, фізика та економіка. RStudio, як потужне інтегроване середовище для аналізу даних, пропонує широкі можливості для вирішення задач цих дисциплін. Його інструменти дозволяють реалізовувати чисельні методи, будувати моделі, виконувати статистичний аналіз і створювати якісну візуалізацію результатів. Завдяки цьому RStudio є ефективним засобом підготовки здобувачів вищої освіти до вирішення прикладних завдань у професійній діяльності.

Особливу увагу привертає міждисциплінарний підхід, який об'єднує методи математичного моделювання, фізичних процесів та економічного аналізу. Впровадження RStudio у навчальний процес сприяє розвитку критичного мислення, вміння працювати з великими обсягами даних та застосовувати сучасні інформаційні технології.

Теоретичний зміст освітніх компонентів, що формують програмні результати фахівця, передбачають огляд задач класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил та вивчення ряду технологій таких до прикладу, як DataMining, Big Data, WebMining та інші.

О. Верес та Р. Оливко аналізуючи методи інтелектуального аналізу даних (Data Mining) приходять до висновку, що дана технологія зорієнтована на розв'язок таких задач: «класифікація (Classification); кластеризація (Clustering); асоціація (Associations); послідовність (Sequence), або послідовна асоціація (sequential association); прогнозування (Forecasting); визначення відхилень (Deviation Detection), аналіз відхилень або викидів; оцінювання (Estimation); аналіз зв'язків (Link Analysis); візуалізація (Visualization, Graph Mining); підбивання підсумків (Summarization) – опис конкретних груп об'єктів за допомогою аналізованого набору даних» (Верес, 2017).

Дослідження, представлені у навчальному посібнику (Кобець, 2021), зосереджені на вивченні можливостей програмного середовища RStudio для економічного аналізу. Посібник орієнтований на здобувачів вищої освіти бакалаврського та магістерського рівнів за спеціальностями інформаційних технологій, а також корисний для аналітиків, економістів, Data Scientist і маркетологів, які займаються обробкою та аналізом даних.

Працюючи із здобувачами вищої освіти, орієнтованими на роботу в IT галузі, зосереджуємось на декількох аспектах при використанні пакету RStudio:

- опрацювання даних типу: вектор, матриця, масив;
- опрацювання статистичних даних;
- графічний аналіз даних;
- лінійна, множинна та нелінійна регресія;
- дисперсійний аналіз;
- факторний аналіз.

Найбільш зручною є така послідовність у засвоєнні технологій опрацювання економічної інформації із використанням RStudio:

- якісна оцінка та попередня підготовка даних;

- опрацювання, очистка, і підготовка даних до обстеження;
- проведення статистичних тестів;
- проведення оціночного аналізу даних;
- опрацювання даних що розміщені в таблицях;
- побудова графіків;
- проведення регресійного аналізу.

Застосування однофакторного дисперсійного аналізу дозволяє перевірити нульову гіпотезу, яка стверджує, що порівнювані групи є вибірками з однієї генеральної сукупності, а їх середні значення не мають статистично значущих відмінностей.

Якщо нульову гіпотезу не вдається відкинути із заданого рівня значущості (наприклад,  $\alpha=0.05$ ) то подальший аналіз є недоцільним, оскільки дані не вказують на наявність відмінностей, які потребують детальнішого дослідження.

Якщо нульова гіпотеза відкидається ми робимо висновок про те, що середні значення порівнюваних груп значно різняться (іншими словами, фактор, що вивчається, істотно впливає на цікаву для нас змінну). Це єдиний висновок, який можна зробити за допомогою дисперсійного аналізу.

Проте зазвичай цікаво піти далі і з'ясувати, в чому полягають відмінності, тобто, які саме групи відрізняються одна від одної. Тому далі необхідно виконати попарні порівняння середніх значень наявних груп.

Для виконання великої кількості попарних порівнянь групових середніх без втрати статистичної потужності було розроблено декілька методів. Одним із найбільш популярних методів для аналізу результатів є критерій достовірно значимої різниці Тьюкі (Tukey's HSD test), оскільки він не тільки контролює помилки, але й дає чіткі інтерпретації для порівняння кожної пари груп.

У RStudio множинні порівняння групових середніх за допомогою тесту Тьюкі можна виконати кількома способами.

Як приклад використаємо дані щодо вмісту стронцію (мг/мл) у п'яти водоймах США (приклад запозичений з книги Zar1999):

```
> waterbodies <-data.frame (water =rep(c("Grayson", "Beaver",  
+                               "Angler", "Appletree",  
+                               "Rock"),each =6),  
+                               Sr =c(28.2, 33.2, 36.4, 34.6, 29.1, 31.0,  
+                               39.6, 40.8, 37.9, 37.1, 43.6, 42.4,  
+                               46.3, 42.1, 43.5, 48.8, 43.7, 40.1,  
+                               41.0, 44.1, 46.4, 40.2, 38.6, 36.3,  
+                               56.3, 54.1, 59.4, 62.7, 60.0, 57.3)  
+ )  
>
```

Рис. 1. Вміст стронцію у п'яти водоймах США (Zar, 1999)

З'ясуємо: 1) наявність суттєвих відмінностей у водоймах за вмістом стронцію в цілому, 2) які саме водойми відрізняються між собою. Для відповіді на перше запитання виконаємо дисперсійний аналіз з допомогою функції aov:

Як очевидно з отриманих результатів, обстежені водойми статистично значно різняться за вмістом стронцію.

```

> M <-aov(Sr ~ water, data= waterbodies)
> summary(M)
          Df Sum Sq Mean Sq F value    Pr(>F)
water      4 2193.4    548.4   56.16 3.95e-12 ***
Residuals 25  244.1     9.8
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

Рис. 2. Дисперсійний аналіз з допомогою функції aov

Щоб з'ясувати, де саме лежать відмінності, достатньо подати об'єкт M на функцію TukeyHSD():

```

> TukeyHSD(M)
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = Sr ~ water, data = waterbodies)

$water
      diff      lwr      upr    p adj
Appletree-Angler -2.983333 -8.281979  2.315312 0.4791100
Beaver-Angler    -3.850000 -9.148645  1.448645 0.2376217
Grayson-Angler  -12.000000 -17.298645 -6.701355 0.0000053
Rock-Angler      14.216667  8.918021 19.515312 0.0000003
Beaver-Appletree -0.8666667 -6.165312  4.431979 0.9884803
Grayson-Appletree -9.016667 -14.315312 -3.718021 0.0003339
Rock-Appletree   17.200000 11.901355 22.498645 0.0000000
Grayson-Beaver  -8.1500000 -13.448645 -2.851355 0.0011293
Rock-Beaver     18.066667 12.768021 23.365312 0.0000000
Rock-Grayson    26.216667 20.918021 31.515312 0.0000000

```

Рис. 3. Результати порівнянь за допомогою функції TukeyHSD

У першому стовпці отриманої таблиці перераховані пари порівнюваних водойм. У другому стовпці містяться різниці між відповідними груповими середніми. Третій і четвертий стовпці містять значення нижньої (lwr) і верхньої (upr) меж різниць середнього значення. Нарешті, у п'ятому стовпці представлені P-значення кожної з порівнюваних пар водойм. Добре видно, що істотної різниці в парах "Appletree-Angler", "Beaver-Angler" та "Beaver-Appletree" немає ( $P > 0.05$ ), тоді як у всіх інших випадках різниця статистично значуща.

Хоча теорія того не вимагає, критерій Тьюкі та інші подібні до нього методи рекомендується застосовувати після того, як дисперсійний аналіз встановив наявність суттєвої різниці між групами в цілому. У зв'язку з цим критерій Тьюкі відноситься до методів апостеріорного аналізу (post-hoc analysis).

Основна ідея критерію Тьюкі полягає в тому, щоб виявити, які пари груп відрізняються значимо, не підвищуючи ймовірність помилок першого роду (відкидання правильної нульової гіпотези). Для цього використовується корекція рівня значущості за допомогою специфічних критеріїв, які враховують кількість порівнюваних груп та кількість всіх можливих пар.

Критерій Тьюкі має такі ж самі умови застосовності, як і дисперсійний аналіз. Дані мають відповідати нормальному розподілу та груповій дисперсії – однорідні. Відхилення від цих умов, так само як і статистична потужність критерію Тьюкі, зростають при однаковій кількості спостережень у всіх порівнюваних групах.

RStudio допомагає вирішувати задачі лінійної алгебри, підтримує чисельне розв'язання диференціальних рівнянь, побудову регресійних моделей і використання методів машинного навчання, дозволяє моделювати динамічні системи у фізиці, аналізувати експериментальні дані, проводити обчислення за допомогою пакетів для візуалізації тривимірних графіків, а також використовувати методи для імовірнісного моделювання.

**Висновки.** Мова програмування R це не стільки засіб розробки комп'ютерних програм, в класичному розумінні цих термінів, а і спеціалізоване середовище роботи з даними. Середовище RStudio об'єднує інструменти, методи статистичного дослідження та цілий набір засобів візуалізації. Загалом, RStudio робить процес аналізу даних на мові R більш зручним та ефективним завдяки своєму зручному інтерфейсу та широкому спектру інструментів. ІТ фахівці повинні бути готові до використання різних інструментів та методів, включаючи статистичні методи та інструменти для обробки великих масивів даних, щоб успішно виконувати обчислення в сферах математики, фізики та економіки.

## СПИСОК ПОСИЛАНЬ

Biostatistical analysis / Jerrold H. Zar. Publisher Upper Saddle River, N.J. : Prentice-Hall, 1999, 663р.

Верес О. М., Оливко Р. М., 2017. Класифікація методів аналізу Великих даних // Вісник Національного університету "Львівська політехніка". Серія : Інформаційні системи та мережі. № 872. – С.84-92. <https://science.lpnu.ua/sites/default/files/journal-paper/2018/jun/13005/ilovepdfcom-84-92.pdf>

Горбачук В.О., 2021. Комп'ютерно-орієнтована методика навчання математичної статистики студентів економічних спеціальностей педагогічних університетів: дис. канд. пед. наук: 13.00.02 / Нац. пед. ун-т ім. М. П. Драгоманова. К., 325с. URL: [https://npu.edu.ua/images/file/vidil\\_aspirant/dicer/%D0%94\\_26.05 3.03/Horbachuk.pdf](https://npu.edu.ua/images/file/vidil_aspirant/dicer/%D0%94_26.05%203.03/Horbachuk.pdf)

Грод, І. М., Андреева, Ю. Ю., 2019. RStudio як один із інструментів дослідження задач з курсу «Комп'ютерна математика». Сучасні інформаційні технології та інноваційні методи навчання: досвід, тенденції, перспективи. Матеріали III Міжнародної науковопрактичної інтернет-конференції (м. Тернопіль, 5 квітня, 2019) – 174 с.

Економетрика в RStudio [Текст] : навч. посіб. / В. М. Кобець ; Херсон. держ. ун-т. – Одеса : Гельветика, 2021. – 131 с.

---

**REFERENCES**

---

- Biostatistical analysis / Jerrold H. Zar. Publisher Upper Saddle River, N.J. : Prentice-Hall, 1999, 663p.
- Veres O. M., Olyvko R. M., 2017. Klasyfikatsiia metodiv analizu Velykykh danykh / // Visnyk Natsionalnoho universytetu "Lvivska politekhnika". Seriia : Informatsiini systemy ta merezhi. № 872. – S.84-92. <https://science.lpnu.ua/sites/default/files/journal-paper/2018/jun/13005/ilovepdfcom-84-92.pdf>
- Horbachuk V.O., 2021. Kompiuterno-oriientovana metodyka navchannia matematychnoi statystyky studentiv ekonomichnykh spetsialnostei pedahohichnykh universytetiv: dys. kand. ped. nauk: 13.00.02 / Nats. ped. un-t im. M. P. Drahomanova. K., 325s. URL: [https://npu.edu.ua/images/file/vidil\\_aspirant/dicer/%D0%94\\_26.05.3.03/Horbachuk.pdf](https://npu.edu.ua/images/file/vidil_aspirant/dicer/%D0%94_26.05.3.03/Horbachuk.pdf)
- Hrod, I. M., Andreieva, Yu. Yu., 2019. RStudio yak odyin iz instrumentiv doslidzhennia zadach z kursu «Kompiuterna matematika». Suchasni informatsiini tekhnolohii ta innovatsiini metodyky navchannia: dosvid, tendentsii, perspektyvy. Materialy III Mizhnarodnoi naukovopraktychnoi internet-konferentsii (m. Ternopil, 5 kvitnia, 2019) – 174 s.
- Ekonometryka v RStudio [Tekst] : navch. posib. / V. M. Kobets ; Kherson. derzh. un-t. – Odesa : Helvetyka, 2021. – 131 s.

**UDC 378:004.451.5**

**Andriy Yasinsky,**

*PhD in Pedagogy of the Department Mathematical Modeling,  
Private Higher Educational Institution*

*"International University of Economics and Humanities  
named after Academician Stepan Demyanchuk",  
Rivne, Ukraine*

*yasinsky@meta.ua*

*<https://orcid.org/0000-0002-1894-1314>*

**Liudmyla Solovej,**

*Senior Lecturer of the Department Information Systems  
and Computing Methods,*

*Private Higher Educational Institution*

*"International University of Economics and Humanities  
named after Academician Stepan Demyanchuk",*

*Rivne, Ukraine*

*lyuda\_solovej@ukr.net*

*<https://orcid.org/0009-0001-2832-1741>*

**USING THE RSTUDIO ENVIRONMENT IN RESEARCHING PROBLEMS  
IN MATHEMATICS, PHYSICS AND ECONOMICS**

**The purpose of the article** is to study the possibilities of using the RStudio environment to solve problems in mathematics, physics and economics. In particular, the emphasis is on



the integration of statistical analysis tools, the use of a set of mathematical functions and operations that ensure efficiency and accuracy in calculations.

**Research methods.** An analysis of scientific papers, methodological recommendations and examples of RStudio application in various fields of knowledge related to mathematics, physics and economics was conducted, the functionality of the RStudio environment was tested on practical tasks, and the results obtained using RStudio were compared with the results obtained using other methods to assess the effectiveness of the environment.

**The scientific novelty** lies in the use of R functionality, which allows developing and implementing new methods and approaches to data analysis and solving scientific problems.

**Conclusions.** The RStudio environment demonstrates its effectiveness for solving problems in mathematics, physics and economics due to a wide range of functions for numerical modeling, statistical analysis and data visualization. Using RStudio allows you to significantly simplify the process of solving complex mathematical problems, in particular, it contributes to the integration of mathematical methods in the study of physical phenomena and economic processes, which allows you to form new approaches to data analysis and modeling.

RStudio can be used as an effective tool in education, contributing to the formation of analytical thinking and programming skills in students engaged in mathematical, physical and economic problems. In particular, the use of the Tukey criterion and the functions of the RStudio environment for comparing data in large arrays, such as the results of experiments or measurements. The results of the study confirm the feasibility of implementing RStudio in the practice of teaching and scientific research, and also indicate the need for further research into the capabilities of the environment in solving problems in scientific fields.

**Keywords:** R programming language; RStudio; modeling; statistical data; analysis of variance; Tukey test.

16.09.2024