

УДК 519.2, 004.4

DOI 10.36910/775.24153966.2024.78.3

Г.М. Губаль

Луцький національний технічний університет

ПОБУДОВА ВИБІРКОВОГО РІВНЯННЯ ПРЯМОЇ РЕГРЕСІЇ В СИСТЕМІ LATEX

У статті створено кореляційну таблицю у системі LaTeX. Автоматизовано обчислення вибірових числових характеристик і побудову вибірового рівняння прямої регресії. Для цього створено нові команди (макриси), які заокруглюють десяткові дроби до двох знаків після точки (коми), обчислюють вибіровий коефіцієнт кореляції, генерують вибірове рівняння прямої регресії. Створені нові команди спрощують обчислення, написання і читання кодів, покращують структурованість LaTeX документів та підвищують продуктивність.

Ключові слова: регресія, LaTeX, макриси, рівняння, таблиця.

H.M. Hubal

A CONSTRUCTION OF A SAMPLE LINEAR REGRESSION EQUATION IN THE LATEX SYSTEM

A correlation table is created in the LaTeX system in this article. It is performed the automation of calculations of numerical characteristics of a sample and a construction of a sample linear regression equation. For this, new commands (macros) round decimal numbers to two decimal places, calculate a sample correlation coefficient, generate the sample linear regression equation. The created new commands simplify calculations, writing and reading codes, improve the structure of LaTeX documents and increase productivity.

Keywords: regression, LaTeX, macros, equation, table.

Постановка проблеми. Для програмування в системі LaTeX створено ефективні пакети з підтримкою обчислень з плаваючою або фіксованою точкою (комою), наприклад, пакети `\calculator`, `\xfr`, `\fr`, `\pst-fr`, `\minifr`, `\pgf`, `\xint`, `\Lua`, `\fltpoint`, `\arpm` [1-9]. Для швидкого програмування ряд пакетів вимагають, щоб кожен числовий аргумент передавався безпосередньо до реєстру розмірностей LaTeX. Оскільки ці пакети забезпечують різну точність (не менше 8 знаків), вибір найефективнішого пакета для виконання відповідного завдання нелегкий. Якщо не задана точність, то найкраще балансувати між продуктивністю та точністю. Для цього застосовується пакет `\xfr` [5, 10, 11]. Це найшвидший пакет для виконання операцій.

Створимо нові команди (макриси) у LaTeX документі, які:

- заокруглюють десяткові дроби до двох знаків після точки (коми);
- обчислюють вибіровий коефіцієнт кореляції;
- генерують вибірове рівняння прямої регресії.

Ці команди спрощують обчислення, написання і читання кодів у LaTeX документах, особливо у документах зі складними і повторюваними завданнями. Ці команди покращують структурованість і автоматизують створення LaTeX документів та підвищують продуктивність. Внесені зміни в командах автоматично застосовуються до їхніх використань у LaTeX документах.

Ці команди можна використовувати повторно в різних частинах LaTeX документа і в різних LaTeX документах. Можна створити пакети команд для визначених типів LaTeX документів.

У статті створимо кореляційну таблицю у системі LaTeX. Автоматизуємо обчислення вибірових числових характеристик і побудову вибірового рівняння прямої регресії, використовуючи створені нові команди.

Аналіз останніх досліджень і публікацій. У системі LaTeX пакети `\calculator`, `\xfr`, `\fr`, `\pst-fr`, `\minifr`, `\pgf`, `\xint`, `\Lua`, `\fltpoint`, `\arpm` використовуються для обчислення і друку значень математичних виразів без використання зовнішнього програмного забезпечення. Створення нових команд дає можливість автоматизувати обчислення [12, 13] і створювати тестові питання для перевірки знань студентів [14-22].

Виклад основного матеріалу. Знайдемо вибірове рівняння прямої регресії η на ξ за даними кореляційної таблиці:

$\xi \backslash \eta$	10	15	20	25	30	35
15	3	3	0	0	0	0
18	1	5	4	0	0	0
21	0	2	8	35	4	0
24	0	0	5	10	5	1
27	0	0	0	4	5	5

Для цього у LaTeX документі `\article` підключимо пакети `\bnumexpr` і `\xfr`. Використаємо команду `\sm`, створену в роботі [9], що приймає шість аргументів, які необхідно ввести. Ця команда обчислює і друкує суму введених аргументів.

Якщо необхідно ввести іншу кількість рядків і стовпців у таблиці, то необхідно змінити число аргументів команди `\sm`.

Створимо нову команду `\rnd`, яка заокруглює десяткові дробі до двох знаків після точки (коми). Ця команда приймає два аргументи, які необхідно ввести. Перший аргумент `\#1` очікує число, а другий аргумент `\#2` очікує кількість знаків після точки (коми).

Створимо нову команду `\corelcf` для обчислення коефіцієнта кореляції. Ця команда приймає п'ять аргументів, які необхідно ввести. Перший аргумент `\#1` очікує значення вибіркового середнього добутку \overline{xy} , другий аргумент `\#2` очікує значення вибіркового середнього \overline{x} , третій аргумент `\#3` очікує значення вибіркового середнього \overline{y} , четвертий аргумент `\#4` очікує значення вибіркового середнього квадратичного відхилення S_x , четвертий аргумент `\#4` очікує значення вибіркового середнього квадратичного відхилення S_y .

У новій команді `\corelcf` створимо нову глобальну команду `\crlcf` за допомогою команди `\xdef`, тобто в цьому LaTeX документі команда `\crlcf` буде глобально представлятись як число, отримане в результаті виконання команди `\fpeval`.

Створимо нову команду `\regreqn`, яка генерує вибіркове рівняння прямої регресії η на ξ за отриманими вибірковими числовими характеристиками заданого статистичного розподілу за допомогою створених команд `\sm`, `\rnd`, `\corelcf`, `\crlcf`. Ця команда приймає чотири аргументи, які необхідно ввести. Перший аргумент `\#1` очікує значення вибіркового середнього \overline{y} , другий аргумент `\#2` очікує значення вибіркового коефіцієнта кореляції R_{xy} , третій аргумент `\#3` очікує

значення відношення $\frac{S_y}{S_x}$, четвертий аргумент `\#4` очікує значення вибіркового середнього \overline{x} . Ці аргументи заокруглені до двох знаків після точки (коми) за допомогою нової команди `\rnd`.

Наведемо програмний код для знаходження вибіркового рівняння прямої регресії η на ξ за допомогою створених команд `\sm`, `\rnd`, `\corelcf`, `\crlcf` `\regreqn` у преамбулі LaTeX документа (рис. 1).

```
\documentclass{article}

\usepackage{bnumexpr,xfr}

\newcommand{\sm}[6]{%
\the\numexpr\#1+\#2+\#3+\#4+\#5+\#6\relax
}%

\newcommand{\rnd}[2]{%
\fpeval{round(\#1,\#2)}
}%

\newcommand{\corelcf}[5]{%
\fpeval{(\#1-(\#2*\#3))/(\#4*\#5)}
\xdef\crlcf{\fpeval{(\#1-\#2*\#3)/(\#4*\#5)}}
}%

\newcommand{\regreqn}[4]{%
\langle y-\rnd{\#1}{2}=\rnd{\#2*\#3}{2}(x-\rnd{\#4}{2})\rangle
}%
```

Рис. 1. Команди для знаходження вибіркового рівняння прямої регресії η на ξ у преамбулі LaTeX документа

У тілі LaTeX документа створимо кореляційну таблицю, яка містить 7 рядків і 8 стовпців. У кожній комірці останнього рядка і останнього стовпця таблиці викличемо команду `\sm`, в якості

аргументів якої введемо відповідні значення елементів (частот) таблиці. Ця команда обчислить суми елементів (частот) рядків і стовпців таблиці. У кожній комірці останнього рядка і останнього стовпця таблиці визначимо глобальні команди `\x`, `\y`, `\z`, `\u`, `\v`, `\w`, `\t`, `\a`, `\b`, `\c`, `\d`, `\e`, як числа, отримані в результаті виконання команди `\sm`.

Наведемо програмний код для створення кореляційної таблиці у тілі LaTeX документа (рис. 2).

```
\begin{document}
\begin{center}
The correlation table

\vspace{2mm}

\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
\(\xi\backslash\eta\) & 10 & 15 & 20 & 25 & 30 & 35 & sum \\
\hline
15 & 3 & 3 & 0 & 0 & 0 & 0 & \sm{3}{3}{0}{0}{0}{0} \\
\hline
\hline
\hline
18 & 1 & 5 & 4 & 0 & 0 & 0 & \sm{1}{5}{4}{0}{0}{0} \\
\hline
\hline
\hline
21 & 0 & 2 & 8 & 35 & 4 & 0 & \sm{0}{2}{8}{35}{4}{0} \\
\hline
\hline
\hline
24 & 0 & 0 & 5 & 10 & 5 & 1 & \sm{0}{0}{5}{10}{5}{1} \\
\hline
\hline
\hline
27 & 0 & 0 & 0 & 4 & 5 & 5 & \sm{0}{0}{0}{4}{5}{5} \\
\hline
\hline
\hline
\hline
sum & \sm{3}{1}{0}{0}{0}{0} & \sm{3}{5}{2}{0}{0}{0} & \sm{0}{4}{8}{5}{0}{0} & \sm{0}{0}{35}{10}{4}{0} & \sm{0}{0}{4}{5}{5}{0} & \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{3}{5}{2}{0}{0}{0} & \sm{0}{4}{8}{5}{0}{0} & \sm{0}{0}{35}{10}{4}{0} & \sm{0}{0}{4}{5}{5}{0} & \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{0}{4}{8}{5}{0}{0} & \sm{0}{0}{35}{10}{4}{0} & \sm{0}{0}{4}{5}{5}{0} & \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{0}{0}{35}{10}{4}{0} & \sm{0}{0}{4}{5}{5}{0} & \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{0}{0}{4}{5}{5}{0} & \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{0}{0}{0}{1}{5}{0} & \sm{4}{10}{17}{49}{14}{6} \\
& \sm{4}{10}{17}{49}{14}{6} \\
\hline
\hline
\end{tabular}
\end{center}

\vspace{5mm}
```

Рис. 2. Створення кореляційної таблиці у тілі LaTeX документа

Після створення кореляційної таблиці обчислимо вибірконе середнє \bar{x} , вибірконе середнє \bar{y} , вибірконе середнє $\overline{x^2}$, вибірконе середнє $\overline{y^2}$, вибіркону дисперсію S_x^2 , вибіркону дисперсію S_y^2 , вибірконе середнє квадратичне відхилення S_x , вибірконе середнє квадратичне відхилення S_y , відношення $\frac{S_y}{S_x}$, вибірконе середнє добутку \overline{xy} за допомогою команди `\fpeval`.

За допомогою команди `\xdf` визначимо глобальні команди `\xavg`, `\yavg`, `\xsqragv`, `\ysqragv`, `\varix`, `\vari`, `\devx`, `\devy`, `\ratiodev`, `\xyavg` для вибіркового середнього \bar{x} , вибіркового середнього \bar{y} , вибіркового середнього $\overline{x^2}$, вибіркового середнього $\overline{y^2}$, вибіркової дисперсії S_x^2 , вибіркової дисперсії S_y^2 , вибіркового середнього квадратичного відхилення S_x ,

вибіркового середнього квадратичного відхилення S_y , відношення $\frac{S_y}{S_x}$, вибіркового середнього добутку \overline{xy} , як числа згенеровані командою `\fpeval`.

Наведемо програмний код для обчислення вибірових числових характеристик заданого статистичного розподілу у тілі LaTeX документа (рис. 3).

```
The sample mean  $\overline{x}=\backslash$ 
\fpeval{(15*\a+18*\b+21*\c+24*\d+27*\e)/\t}
\undef\xavg{\fpeval{(15*\a+18*\b+21*\c+24*\d+27*\e)/\t}}
```

```
The sample mean  $\overline{y}=\backslash$ 
\fpeval{(10*\x+15*\y+20*\z+25*\u+30*\v+35*\w)/\t}
\undef\yavg{\fpeval{(10*\x+15*\y+20*\z+25*\u+30*\v+35*\w)/\t}}
```

```
The sample mean  $\overline{x^2}=\backslash$ 
\fpeval{(15^2*\a+18^2*\b+21^2*\c+24^2*\d+27^2*\e)/\t}
\undef\xsqravg{\fpeval{(15^2*\a+18^2*\b+21^2*\c+24^2*\d+27^2*\e)/\t}}
```

```
The sample mean  $\overline{y^2}=\backslash$ 
\fpeval{(10^2*\x+15^2*\y+20^2*\z+25^2*\u+30^2*\v+35^2*\w)/\t}
\undef\ysqravg{\fpeval{(10^2*\x+15^2*\y+20^2*\z+25^2*\u+30^2*\v+35^2*\w)/\t}}
```

```
The sample variance  $(S_x^2=\backslash)$ 
\fpeval{\xsqravg-\xavg^2}
\undef\varix{\fpeval{\xsqravg-\xavg^2}}
```

```
The sample variance  $(S_y^2=\backslash)$ 
\fpeval{\ysqravg-\yavg^2}
\undef\vari y{\fpeval{\ysqravg-\yavg^2}}
```

```
The sample standard deviation  $(S_x=\backslash)$ 
\fpeval{\sqrt{\varix}}
\undef\devx{\fpeval{\sqrt{\varix}}}
```

```
The sample standard deviation  $(S_y=\backslash)$ 
\fpeval{\sqrt{\vari y}}
\undef\devy{\fpeval{\sqrt{\vari y}}}
```

```
The ratio  $\frac{S_y}{S_x}=\backslash$ 
\fpeval{\devy/\devx}
\undef\ratiodev{\fpeval{\devy/\devx}}
```

```
The sample mean of the product  $\overline{xy}=\backslash$ 
\fpeval{(15*(10^3+15^3+20^3+25^3+30^3+35^3)
+18*(10^1+15^5+20^4+25^0+30^0+35^0)
+21*(10^0+15^2+20^8+25^35+30^4+35^0)
+24*(10^0+15^0+20^5+25^10+30^5+35^1)
+27*(10^0+15^0+20^0+25^4+30^5+35^5))/\t}
\undef\xyavg{\fpeval{(15*(10^3+15^3+20^3+25^3+30^3+35^3)
+18*(10^1+15^5+20^4+25^0+30^0+35^0)
+21*(10^0+15^2+20^8+25^35+30^4+35^0)
+24*(10^0+15^0+20^5+25^10+30^5+35^1)
+27*(10^0+15^0+20^0+25^4+30^5+35^5))/\t}}
```

Рис. 3. Обчислення вибірових числових характеристик у тілі LaTeX документа

Для обчислення вибірового коефіцієнта кореляції R_{xy} виклинемо команду `\corelcf`, створену у преамбулі LaTeX документа. В якості її аргументів введемо глобальні команди `\xyavg`, `\xavg`, `\yavg`, `\devx`, `\devy`.

Знайдемо вибіркоче рівняння прямої регресії η на ξ , викликавши команду `\regreqn`, створену у преамбулі LaTeX документа. В якості її аргументів введемо глобальні команди `\yavg`, `\crlcf`, `\ratiodev` і `\xavg`.

Наведемо програмний код для знаходження вибіркового рівняння прямої регресії η на ξ за отриманими вибірковими числовими характеристиками заданого статистичного розподілу у тілі LaTeX документа (рис. 4).

```
The correlation coefficient \(\R_{\{xy\}}=\)
\corelcf{\xyavg}{\xavg}{\yavg}{\devx}{\devy}

The linear regression equation is
\regreqn{\yavg}{\crlcf}{\ratiodev}{\xavg}
\end{document}
```

Рис. 4. Звертання до створених команд для знаходження вибіркового рівняння прямої регресії η на ξ у тілі LaTeX документа

Наведений програмний код на рис. 1 – рис. 4 генерує результат, представлений на рис. 5.

The correlation table

$\xi \backslash \eta$	10	15	20	25	30	35	sum
15	3	3	0	0	0	0	6
18	1	5	4	0	0	0	10
21	0	2	8	35	4	0	49
24	0	0	5	10	5	1	21
27	0	0	0	4	5	5	14
sum	4	10	17	49	14	6	100

The sample mean $\bar{x} = 21.81$

The sample mean $\bar{y} = 23.85$

The sample mean $\overline{x^2} = 485.01$

The sample mean $\overline{y^2} = 600.25$

The sample variance $S_x^2 = 9.3339$

The sample variance $S_y^2 = 31.4275$

The sample standard deviation $S_x = 3.055143204499586$

The sample standard deviation $S_y = 5.606023546151051$

The ratio $\frac{S_y}{S_x} = 1.834946243401799$

The sample mean of the product $\overline{xy} = 532.95$

The correlation coefficient $R_{xy} = 0.7462688858496572$

The linear regression equation is $y - 23.85 = 1.37(x - 21.81)$

Рис. 5. Результат, згенерований програмним кодом

Висновки. У статті побудовано вибіркоче рівняння прямої регресії у системі LaTeX. При цьому створено нові команди (макриси), які:

- заокруглюють десяткові дробу до двох знаків після точки (коми);
- обчислюють вибірковий коефіцієнт кореляції;
- генерують вибіркоче рівняння прямої регресії.

Створено кореляційну таблицю у системі LaTeX. Автоматизовано обчислення вибіркових числових характеристик і побудову вибіркового рівняння прямої регресії з використанням створених нових команд. Команди, створені у цій статті, покращують структурованість і автоматизують створення LaTeX документів та підвищують продуктивність.

Список використаних джерел:

1. Губаль, Г. М. (2013). LATEX як видавнича система для створення математичних текстів і

для програмування. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*, (12), 23-26.

2. Губаль, Г. М. (2013). Стратегії для створення математичної статті у видавничій системі LATEX. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*, (13), 10-13.

3. Gryshchenko, T.V., Deineko, Zh.V. & Nikitenko, O.M. (2019). Using the LaTeX system during the preparation of scientific publications. *IV International Scientific and Technical Conference "Print, Multimedia & Web"*, 96–98.

4. Hubal, H.M. (2023) Improvement of references and footnotes in mathematical and other texts by creating macros in the LaTeX programming language. *International Journal on Information Technologies & Security*, (15(3)), 15–22.

5. Lammport, L. (1994). *LaTeX: a document preparation system*. 2nd ed. Boston: Addison-Wesley. (272 p.).

6. Lode, C. (2019). *Better books with LaTeX the agile way*. Clements Lode Verlag E.K. (254 p.).

7. Goossens, M., Rahtz, S. P., & Mittelbach, F. (1997). *The LATEX graphics companion: illustrating documents with TEX and PostScript* (Vol. 3). Addison-Wesley Professional.

8. Goossens, M., Rahtz, S. P., & Rahtz, S. (1999). *The Latex Web Companion: Integrating TEX, HTML, and XML* (Vol. 4). Addison-Wesley Professional.

9. Griffiths, D. F., & Higham, D. J. (2016). *Learning LaTeX*. Society for Industrial and Applied Mathematics.

10. Kottwitz, S. (2021). *LaTeX beginner's guide*. Packt Publishing, USA, (354 p.).

11. Kottwitz, S. (2023). *LaTeX graphics with TikZ*. Packt Publishing, USA, (304 p.).

12. Губаль, Г. М. (2024). Вдосконалення створення таблиць і деяких обчислень на мові програмування LATEX для математичних досліджень. *Наукові нотатки*, (77), 78-84.

13. Hubal, H. (2024). Analysis of Approaches to the Study of Limits of Sequences and Functions and the Use of Information Technologies. *Grail of Science*, (35), 226-231.

14. Губаль, Г. М. (2014). Особливості створення інтерактивних математичних тестів у видавничій системі LATEX. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*, (15), 9-13.

15. Hubal, H. (2023). LATEX system in distance learning of the probability theory. *Grail of Science*, (33), 244-248.

16. Hubal, H. (2024). Optimization of the Use of Some Modules in the Moodle System. *Grail of Science*, (38), 226-231.

17. Romansky, R. (2023). Empirical evaluation of the transfer of information resources in active learning. *International Journal on Information Technologies & Security*, 15(1), 39–48.

18. Hubal, H.M. (2019) Mathematical description of the non-equilibrium state of symmetric particle systems. *International Journal of Applied Mathematics*, 32(5), 767–774.

19. Hubal, H.M. (2021) Mathematical modeling of biochemical processes rates in biological systems. *Computer-Integrated Technologies: Education, Science, Production*, (42), 43–49.

20. Губаль, Г. М. (2022). Математичне дослідження стійкості особливих точок систем диференціальних рівнянь, які описують швидкості біохімічних процесів. *Наукові нотатки*, (73), 29-39.

21. Van Dongen, M.R.C. (2012). *LaTeX and friends*. Springer-Verlag Berlin and Heidelberg GmbH & Co. K. (300 p.).

22. Verdaguer-Codina, J. (2022). Linux, LaTeX, and Python in secondary and baccalaureate. *Proceedings of IEEE Global Engineering Education Conference (EDUCON)*, Tunisia, 763–768.