


**ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«МЕТІНВЕСТ ПОЛІТЕХНІКА»**

ЕКОНОМЕТРІЯ

**методичні рекомендації
до виконання практичних та індивідуальних робіт**

Запоріжжя 2024



УДК 330.43(072)
Е45

Рекомендовано Науково-методичною
радою
ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«МЕТІНВЕСТ ПОЛІТЕХНІКА»
(протокол № 3 від 22.11.2024 р.)

Укладачі

Жерліцин Д.М., доктор економічних наук, професор

Подскребко О.С., кандидат економічних наук, доцент

Е45 **Економетрія** : методичні рекомендації до виконання практичних та індивідуальних робіт / уклад. Д. М. Жерліцин, О. С. Подскребко. Запоріжжя : ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА», 2025. 43 с.

Методичні рекомендації включають особливості побудови однофакторних та багатофакторних економетричних моделей, розглядають сукупність статистичних тестів, які використовуються в процесі економетричного дослідження, метрики оцінки якості отриманих моделей. Основний акцент робиться на дотриманні передумов застосування методу найменших квадратів (теорема Гауса-Маркова), наслідках та методах визначення мультіколінеарності, гетероскедастичності. Рекомендовано для здобувачів вищої освіти за освітньо-професійною програмою першого (бакалаврського) рівня спеціальності 051 «Економіка».

УДК 330.43(072)

© ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ «МЕТІНВЕСТ ПОЛІТЕХНІКА», 2024



ЗМІСТ

ВСТУП.....	4
ПРАКТИЧНА РОБОТА №1. ПОБУДОВА ЕКОНОМЕТРИЧНОЇ МОДЕЛІ. ПАРНА ЛІНІЙНА РЕГРЕСІЯ НА ОСНОВІ ПОКРОКОВОЇ РЕГРЕСІЇ.....	7
Питання для самоаналізу.....	16
ПРАКТИЧНА РОБОТА №2 ПОБУДОВА ЕКОНОМЕТРИЧНОЇ МОДЕЛІ НА ОСНОВІ ПОКРОКОВОЇ РЕГРЕСІЇ	17
Питання для самоаналізу.....	22
ПРАКТИЧНА РОБОТА №3 ПЕРЕВІРКА ПЕРЕДУМОВИ ВІДСУТНОСТІ МУЛЬТИКОЛІНЕАРНОСТІ МІЖ ФАКТОРАМИ- АРГУМЕНТАМИ. МЕТОДИ УСУНЕННЯ МУЛЬТИКОЛІНЕАРНОСТІ	23
Питання для самоаналізу.....	27
ПРАКТИЧНА РОБОТА № 4 ПЕРЕВІРКА ПЕРЕДУМОВИ ГОМОСКЕДАСТИЧНОСТІ ЗБУРЕНЬ. УЗАГАЛЬНЕНИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ ОЦІНКИ ПАРАМЕТРІВ МОДЕЛІ З ГЕТЕРОСКЕДАСТИЧНИМИ ЗАЛИШКАМИ.....	28
Питання для самоаналізу.....	34
ІНДИВІДУАЛЬНА РОБОТА 1. МОДЕЛЬ ПАРНОЇ РЕГРЕСІЇ	35
ІНДИВІДУАЛЬНА РОБОТА 2. ОЦІНКА ПАРАМЕТРІВ МОДЕЛІ ЗА ДОПОМОГОЮ АЛГОРИТМУ ПОКРОКОВОГО ВІДБОРУ ФАКТОРІВ ...	39
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	42



Вступ

Методичні рекомендації з виконання практичних та індивідуальних робіт із дисципліни «Економетрія» створені для здобувачів вищої освіти першого (бакалаврського) рівня за спеціальністю 051 «Економіка». Метою методичних рекомендацій є допомога студентам у засвоєнні основних особливостей побудови економетричних моделей та їх формування висновків про закономірності, кількісні зв'язки та динаміку економічних процесів.


Економетрія є потужним інструментом, який спрямован на набуття майбутніми фахівцями систематизованих практичних навичок щодо вивчення та аналізу соціально-економічних процесів засобами економетричного моделювання, розвиток навичок використання даних методів та моделей, відповідного програмного забезпечення для обробки, аналізу даних, інтерпретації отриманих результатів та використання набутих знань у плануванні, прийнятті обґрунтованих управлінських рішень.

Методичні рекомендації охоплюють тематику від особливості побудови однофакторних та багатфакторних економетричних моделей, розглядають сукупність статистичних тестів, які використовуються в процесі економетричного дослідження, метрики оцінки якості отриманих моделей. Основний акцент робиться на дотриманні передумов застосування методу найменших квадратів (теорема Гауса-Маркова), наслідках та методах визначення мультиколінеарності, гетероскедастичності та автокореляції залишків.

Завданнями курсу є:

- ознайомлення з теоретичними основами економетричного моделювання, формування у студентів базових знань щодо побудови, аналізу та інтерпретації економетричних моделей;
- розвиток практичних навичок застосування економетричних методів, навчання методам оцінки параметрів моделей, перевірки їх адекватності та діагностики порушення основних передумов;
- засвоєння інструментів обробки й аналізу економічних даних
- використання сучасних програмних засобів для розв'язання економетричних задач.
- розвиток аналітичного мислення – навчання аналізу реальних економічних процесів і розробки моделей, які допомагають у прогнозуванні та прийнятті управлінських рішень.
- підготовка до наукових досліджень і професійної діяльності, набуття компетенцій, необхідних для виконання наукових і практичних завдань у сфері економіки.

Після опанування дисципліни «Економетрія» здобувачі освіти набудуть таких програмних результатів навчання:



1. Застосовувати відповідні економіко-математичні методи та моделі для вирішення економічних задач.

2. Ідентифікувати джерела та розуміти методологію визначення і методи отримання соціально-економічних даних, збирати та аналізувати необхідну інформацію, розраховувати економічні та соціальні показники.

3. Вміти використовувати дані, надавати аргументацію, критично оцінювати логіку та формувати висновки з наукових та аналітичних текстів з економіки.

4. Використовувати інформаційні та комунікаційні технології для вирішення соціально-економічних завдань, підготовки та представлення аналітичних звітів.

5. Застосовувати моделі та методи економічної статистики для оцінки та прогнозування соціально-економічних процесів.

6. Використовувати інструменти MsExcel та надбудови Аналіз даних для оцінки економетричних моделей.

7. Вміти абстрактно мислити, застосовувати аналіз та синтез для виявлення ключових характеристик економічних систем різного рівня, а також особливостей поведінки їх суб'єктів.

8. Демонструвати розуміння взаємозв'язку між перебігом технологічних, організаційних та інших процесів та економічними показниками під час аналітичного супроводу розробки і реалізації проєктів розвитку бізнес-діяльності.

9. Аналізувати та моделювати бізнес-процеси компанії на основі дослідження закономірностей у великих обсягах даних з використанням передових методологій та цифрових інструментів.

Очікується, що після завершення курсу студенти матимуть змогу ефективно використовувати економетричне моделювання для аналізу економічних процесів, пошуку взаємозв'язків та інтєпретації отриманих результатів.


Критерії оцінювання

Максимальні 5 балів за практичною роботою передбачають, що здобувач освіти:

- підготував і завантажив звіт з виконання практичної (лабораторної) роботи, що виконано у повній відповідності до поставлених завдань, у т.ч. індивідуального характеру (2 бали);

- дав пряму і релевантну відповідь на поставлене питання щодо виконаного завдання, у т.ч. у вигляді додаткових запитань / зміг стисло формалізувати вербально сутність проблеми за ситуацією, ідентифікувати ключові складові і пріоритети вирішення, запропонував логічне розв'язання (3 бали).

Виконання та захист індивідуального завдання: підготовлене аналітичне завдання у вигляді файлу *.docx, *.xlsx, *.ipynb, *.pdf розміщується у відповідному розділі дисципліни в Moodle і



перевіряється протягом тижня після завершення терміну подачі. Оскарження оцінки може бути здійснене на останньому практичному занятті модуля.

Максимальна оцінка за індивідуальну роботу складає 15 балів, що виставляється студенту:

— який підготував завдання (звіт) за ситуаційним завданням, в якому: правильно визначив проблеми, комплекс факторів, які могли вплинути на їх виникнення, обґрунтував своє бачення теоретичними концепціями або моделями, виконав необхідні розрахунки в разі потреби, представив висновок або власне бачення виходу з проблеми і окреслив можливі перспективи і обмеженість такого рішення; завдання структуровані, викладені діловим, науковим або публіцистичним стилем української (або часткового, англійської) мови з використанням вивчених методів підготовки аналітичних звітів (5 балів);

— оформив аналітичний звіт містить комплексну, логічну і оригінальну розв'язку поставлених завдань аж до міждисциплінарного підходу; використання штучного інтелекту (ШІ) не забороняється, оскільки пропозиції відомих застосунків ШІ суттєво залежать від обміркованої постановки питання і уточнюючих питань; однак в разі, якщо відповідь, отримана з використанням ШІ, не є комплексною або не відповідає за стилем і викладеними позиціями іншим частинам завдання, містить очевидно неправдиву інформацію, то оцінка за цим критерієм знижується (5 балів)

— під час презентації / захисту аналітичного звіту демонструє володіння термінологічним апаратом, відповідає на запитання, здатний швидко адаптувати позицію під зміни у вихідному ситуаційному завданні (5 бали).



Практична робота №1. Побудова економетричної моделі. Парна лінійна регресія на основі покрокової регресії

Мета практичної роботи

Навчитися будувати однофакторні регресійні моделі, оцінювати їх якість і статистичну значущість змінних з використанням MS Excel, а також проводити діагностику та інтерпретацію отриманих результатів.

Терміни виконання практичної роботи

Практична робота виконується протягом двох занять:

Практичне заняття 1 — виконання завдань 1–3.

Практичне заняття 2 — виконання завдань 4–6, завантаження звіту та захист роботи.

Завдання

На основі статистичних даних доходу підприємства (у млн. грн.) Y та кількості працюючих (у тис. чол.) X

y	x
10,8	2,53
11,9	3,54
12,4	3,84
13,2	3,84
14,1	4,22
15,2	4,81
16,0	6,53
17,4	5,82
18,6	6,43
19,4	7,73
20,5	8,19
21,3	7,65
22,5	9,31
23,7	9,26
25,0	9,86
Прогнозне значення	9,98

1. Знайти оцінки параметрів лінійної регресії $y = a_0 + a_1 x + \varepsilon$.
2. Перевірити модель на адекватність з рівнем надійності 95%.
3. Визначити значимість коефіцієнту нахилу регресії з рівнем надійності 95%.
4. Оцінити значущість коефіцієнта кореляції з рівнем надійності 95%.

5. Визначити надійні інтервали для коефіцієнтів регресії з рівнем надійності 95%.

6. Обчислити середній коефіцієнт еластичності.

7. Перевірити точність економетричної моделі за допомогою середньої відносної похибки апроксимації (коефіцієнту апроксимації).

8. Визначити точковий та інтервальний прогнози для заданого останнього значення незалежної змінної.

9. Побудувати графіки фактичних даних, лінію регресії.

10. Побудувати економетричну модель за допомогою функції ЛИНЕЙН та інструменту ДАННЫЕ – АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

11. Побудувати та проаналізувати графіки і коефіцієнти детермінації рівнянь нелінійної регресії.

12. Виконати економіко-математичний аналіз характеристик економетричної моделі

Завдання 1. Знайти оцінки параметрів лінійної регресії $y = a_0 + a_1x + \varepsilon$.

$$a_1 = \frac{\overline{yx} - \bar{y} \cdot \bar{x}}{x^2 - (\bar{x})^2},$$

$$a_0 = \bar{y} - a_1 \bar{x}.$$

Таким чином вибіркова регресійна функція записується у такому вигляді:

$$\hat{y} = 5,7486 + 1,8787x$$

При збільшенні кількості працюючих на 1000 чоловік доход підприємства зростає на 1,8787 млн. грн.

Формула розкладу дисперсії має вигляд

■ $TSS = ESS + RSS$, (Total Sum of Squares, Residual Sum of Squares and Explained Sum of Squares)

$$TSS = \sum_{i=1}^n (y_i - \bar{y})^2 - \text{загальна сума квадратів}$$

$$ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 - \text{пояснена сума квадратів}$$

$$RSS = \sum_{i=1}^n e_i^2 - \text{сума квадратів залишків}$$

Таким чином, $RSS = 10,175$, $ESS = 277,618$, $TSS = 287,793$.

Коефіцієнт детермінації моделі становить

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS},$$

$$R^2 = 0,965$$

2. Перевірити модель на адекватність з рівнем надійності 95%.

Перевіримо модель на адекватність. Практичне значення статистики Фішера дорівнює

$$F_{\text{факт}} = \frac{\frac{\sum(\hat{y}_i - \bar{y})^2}{m}}{\frac{\sum(y_i - \bar{y})^2}{n - m - 1}} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2) = \frac{R^2}{1 - R^2} \cdot \frac{1}{n - 2}$$

$$F_{\text{факт}} = 354,71$$

Якщо $F_{\text{розр}} > F_{\text{табл}} = F_{\alpha, m, n - m - 1}$, то рівняння статистично значимо (з надійністю $1 - \alpha$), тобто модель адекватна.

$F_{\text{табл}}$ знаходиться за допомогою функції: $F_{\text{РАСПОБР}}(\text{вероятность}, \text{степени_свободы1}, \text{степени_свободы2}) = 4,67$ (или функція $F_{\text{ОБР.ПХ}}$ для поздних версий).

$F_{\text{розр}} > F_{\text{табл}}$ таким чином модель є адекватною.

3. Визначити значимість коефіцієнту нахилу регресії з рівнем надійності 95%.

$$t_{pr} = \left| \frac{a_1}{s.e.(a_1)} \right|$$

$$s.e.(a_1) = \sqrt{\frac{\frac{RSS}{n - 2}}{\sum(x - \bar{x})^2}},$$

$$\sigma^2 = \frac{RSS}{n - 2}$$

$$s.e.(a_1) = 0,0998$$

$$t_{pr} = 18,8338$$

Якщо $|t_{a_j}| > t_{tabl} = t_{\alpha, n-m-1}$, то коефіцієнт a_j статистично значимий (з надійністю $1 - \alpha$).

$$t_{tabl} = 2,16. \text{ В Excel функція } \text{СТЪЮДЕНТ.ОБР.2X}(\alpha; n-2)$$

Таким чином, гіпотеза повинна бути відхилена, а значить коефіцієнт a_1 є значимим.

4. Оцінити значущість коефіцієнта кореляції з рівнем надійності 95%.

Розрахуємо коефіцієнт кореляції:

$$r_{xy} = a_1 \frac{\sigma_x}{\sigma_y} = a_1 \frac{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = 1,8787 \frac{\sqrt{78,6563}}{\sqrt{287,7933}} = 0,9822$$

Для перевірки гіпотези про те, що коефіцієнт кореляції рівний 0, використовують статистики

$$t_{pr} = |r_{xy}| \sqrt{\frac{n-2}{1-r_{xy}^2}} = 18,8338, \quad t_{tabl} = t(\alpha, n-2) = 2,16$$

Оскільки практичне значення більше за теоретичне, то коефіцієнт кореляції не можна прийняти рівним 0, тобто він є значущим.

5. Визначити надійні інтервали для коефіцієнтів регресії з рівнем надійності 95%.

$$s.e.(a_0) = \sqrt{\frac{RSS}{n-2} \cdot \frac{\sum x^2}{n \cdot \sum (x_i - \bar{x})^2}}$$

$$a_0 - \Delta_{a_0} \cdot t_{tabl} < a_0 < a_0 + \Delta_{a_0} \cdot t_{tabl}$$

$$a_1 - \Delta_{a_1} \cdot t_{tabl} < a_1 < a_1 + \Delta_{a_1} \cdot t_{tabl}$$

$$4,316715 < a_0 < 7,180448;$$

$$1,663201 < a_1 < 2,094201.$$

6. Обчислити середній коефіцієнт еластичності.

Середній коефіцієнт еластичності

$$\varepsilon = a_1 \frac{\bar{x}}{\bar{y}}$$

$$\varepsilon = 0,67.$$

При збільшенні кількості працюючих на 1% доход підприємства зросте на 0,67%.

7. Визначити середню відносну помилку апроксимації (коефіцієнт апроксимації).

$$E_{\text{ср.отн.}} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \cdot 100\% = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{y_i} \cdot 100\%$$

$$E_{\text{ср.отн.}} = 4,1\%.$$

8. Визначити точковий та інтервальний прогнози для заданого останнього значення незалежної змінної.

Для знаходження прогнозного значення результативної ознаки y_{pr} необхідно прогнозне значення факторної ознаки x_{pr} підставити в рівняння регресії

$$y_{pr} = a_0 + a_1 x_{pr}$$

$$y_{pr} = 24,5.$$

Оцінимо точність прогнозу, розрахувавши помилку прогнозу і його довірчий інтервал. Оцінка помилки прогнозу (при довірчій ймовірності $1-\alpha$) розраховується за формулою

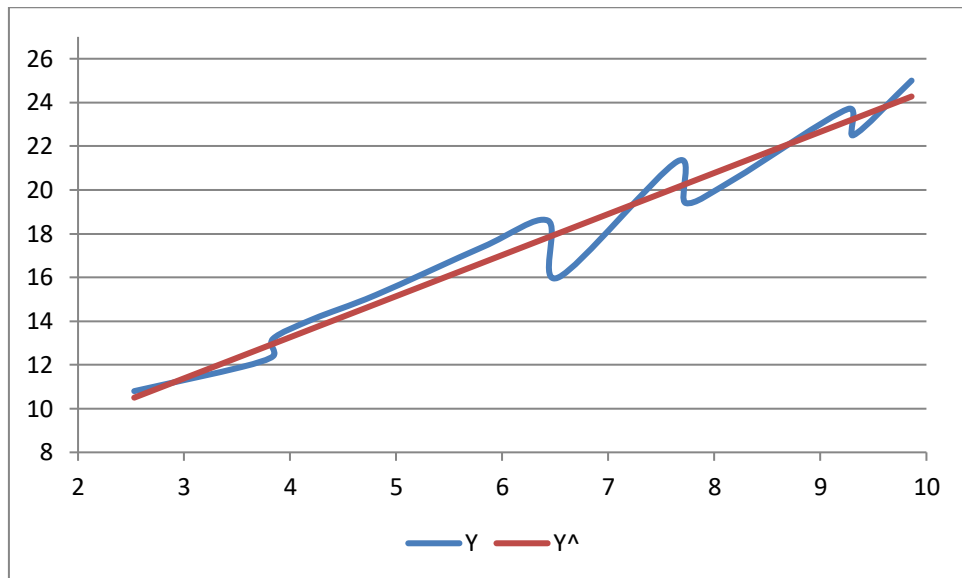
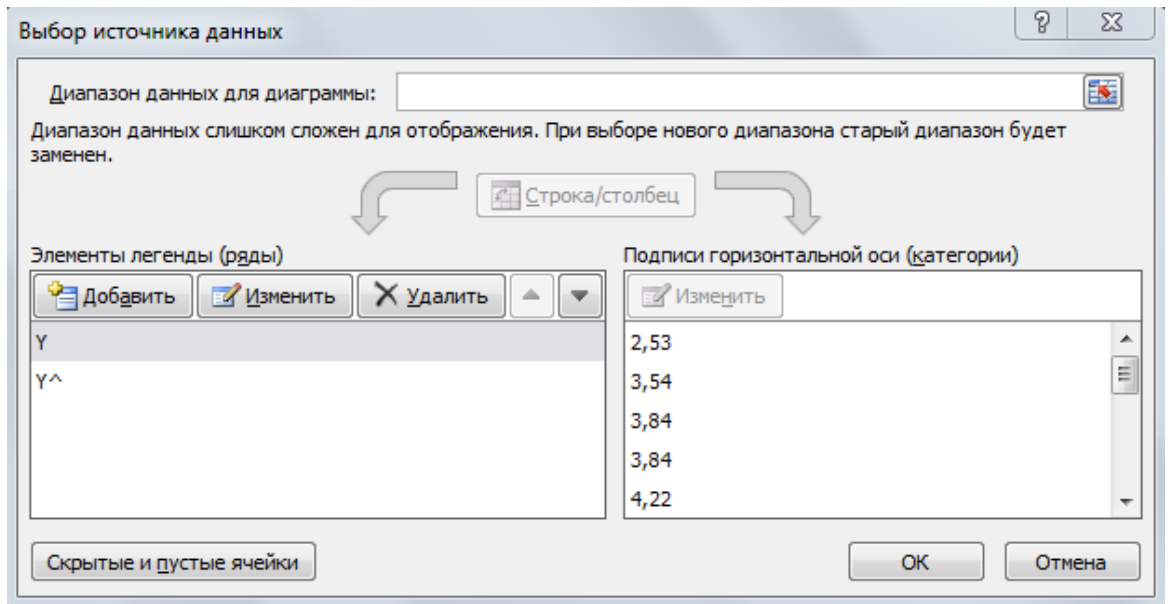
$$\Delta y_{pr} = \frac{RSS}{n-2} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_{pr} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

$$\Delta y_{pr} = 0,8732.$$

$$y_{pr} - \Delta y_{pr} \cdot t_{\text{tabl}} < y_{pr} < y_{pr} + \Delta y_{pr} \cdot t_{\text{tabl}}$$

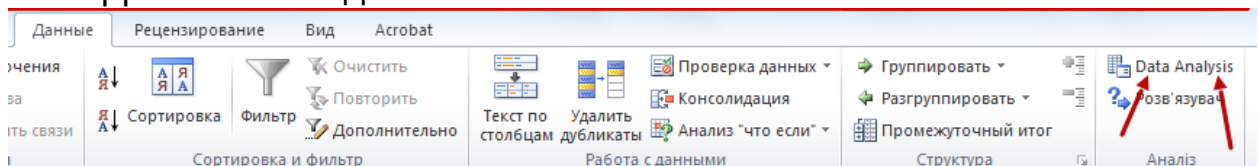
$$22,6116 < y_{pr} < 26,3844$$

9. Побудувати графіки фактичних даних, лінію регресії.

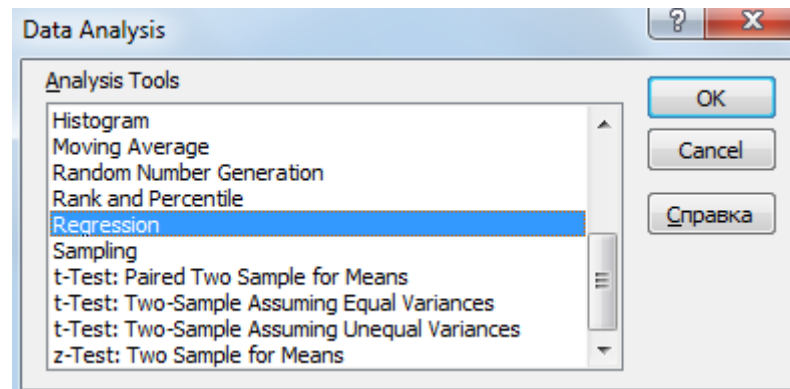


10. Побудувати економетричну модель за допомогою функції ЛИНЕЙН та інструменту ДАННЫЕ – АНАЛИЗ ДАННЫХ – РЕГРЕССИЯ.

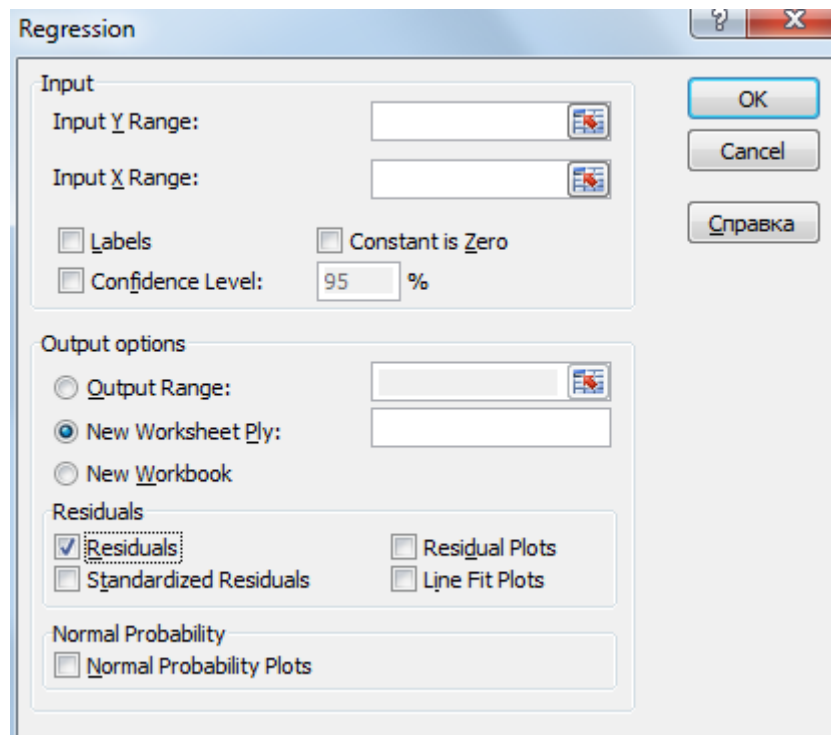
Дані – Аналіз даних



У діалоговому віконці обрати Регресія.



Та ввести свої дані.

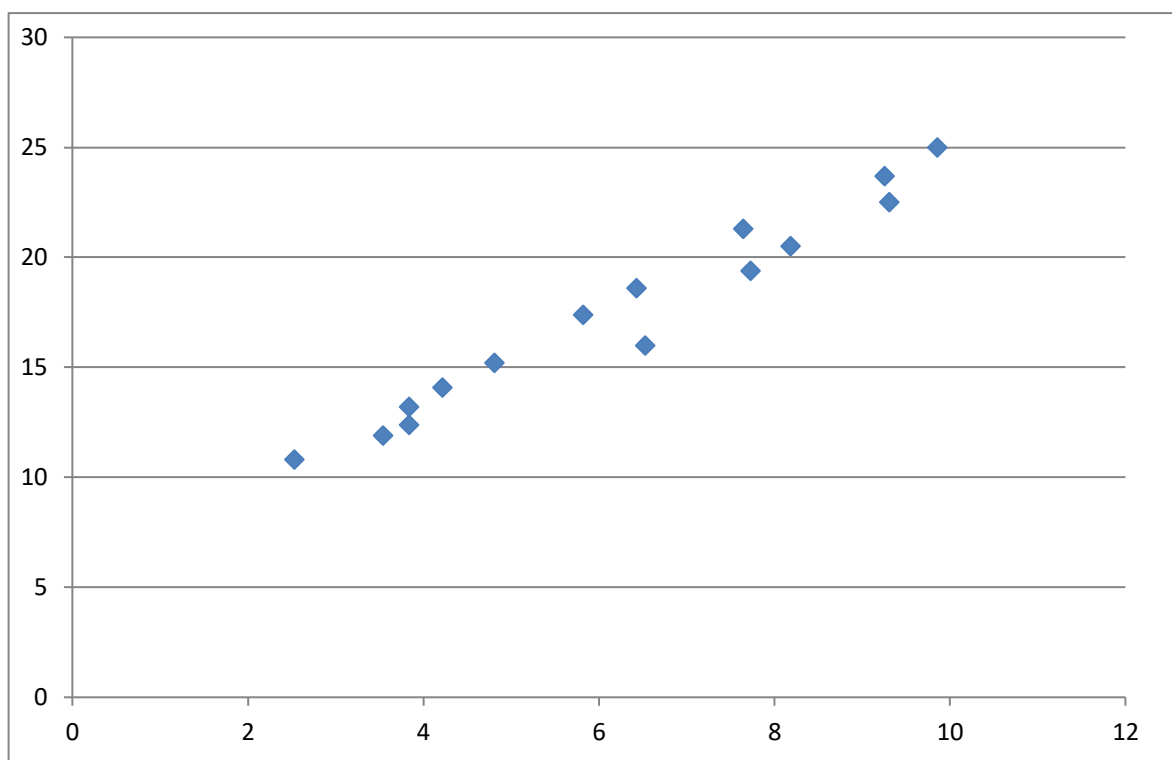


Отримуємо наступну таблицю зображену на рисунку.

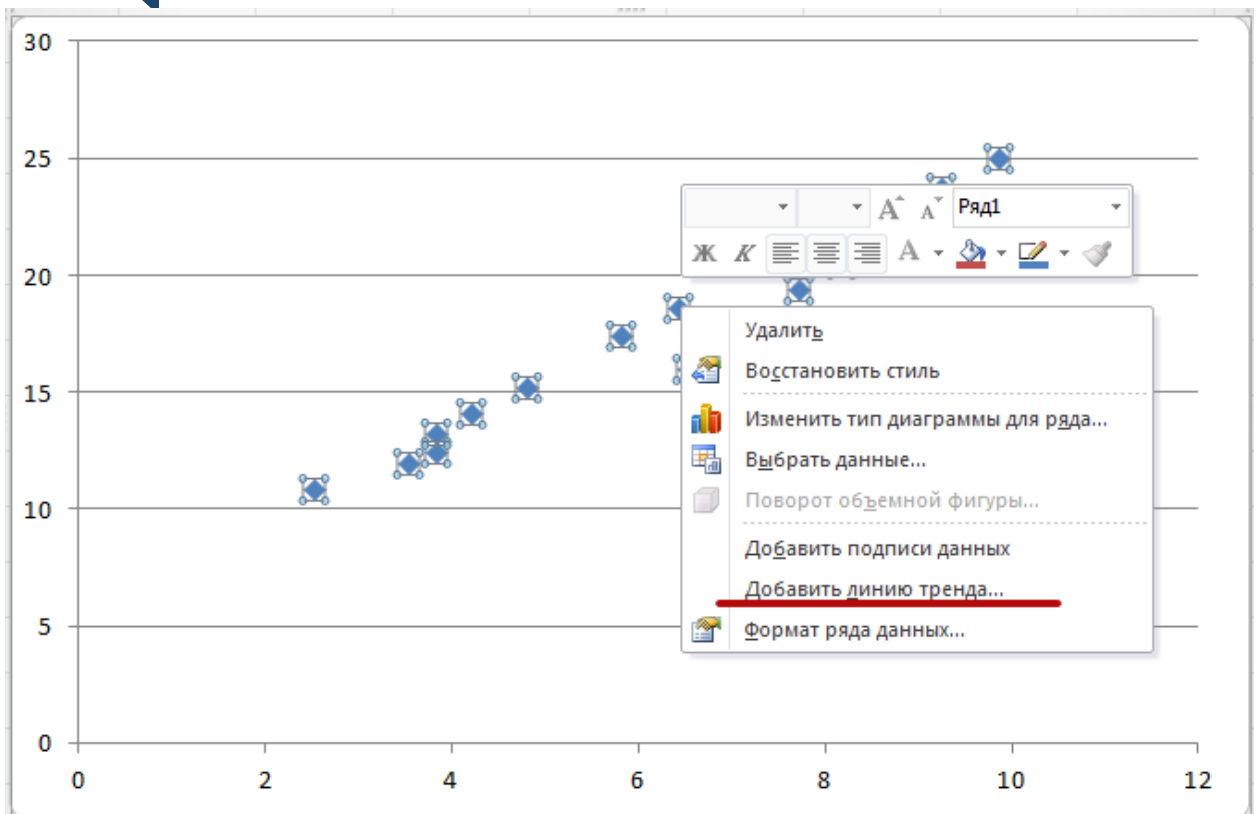
	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	<i>Regression Statistics</i>								
4	Multiple R	0,9821641							
5	R Square	0,9646464							
6	Adjusted R Square	0,9619269							
7	Standard Error	0,8846789							
8	Observations	15							
9									
10	ANOVA								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>gnificance F</i>			
12	Regression	1	277,6187962	277,619	354,713	8,1E-11			
13	Residual	13	10,17453712	0,78266					
14	Total	14	287,7933333						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
17	Intercept	5,7485815	0,662788289	8,67333	9,1E-07	4,31671	7,18045	4,31671	7,18045
18	X Variable 1	1,8787011	0,099751381	18,8338	8,1E-11	1,6632	2,0942	1,6632	2,0942

11. Побудувати та проаналізувати графіки і коефіцієнти детермінації рівнянь нелінійної регресії.

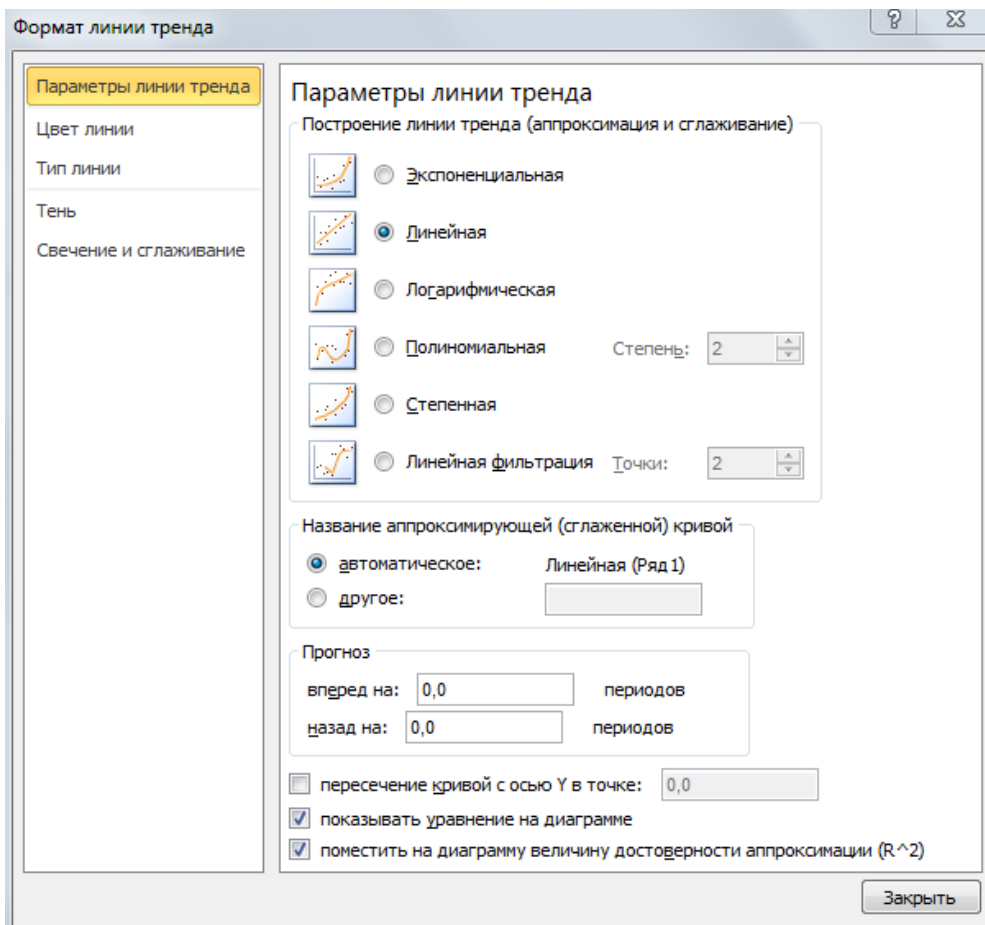
Побудувати точковий графік залежності Y від X.

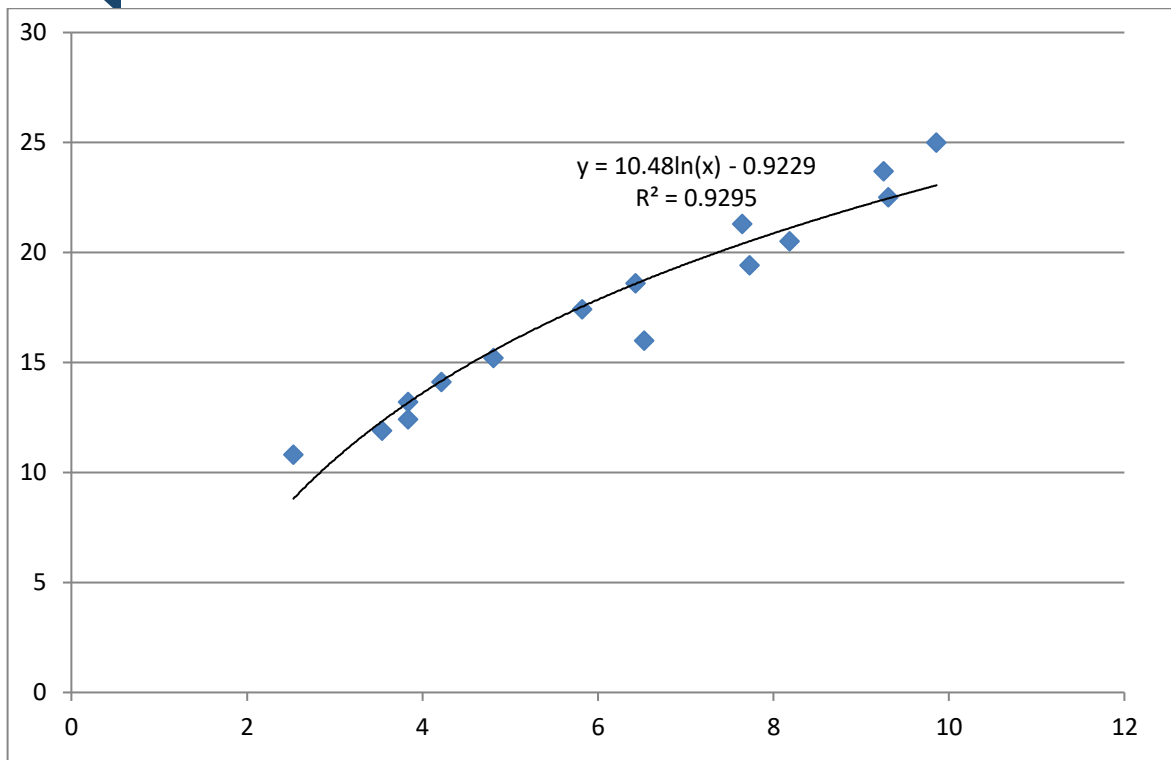


ПКМ на побудованому графіку (на точках графіку) та вибрати «Додати лінію тренду».



В діалоговому вікні вибираємо тип лінії тренду і ставимо необхідні галочки, як показано на рисунку.





Питання для самоаналізу

1. Які основні етапи побудови парної лінійної регресії? Що включає специфікація моделі?
2. Як інтерпретуються параметри регресійної моделі (β_0 та β_1) у контексті реальних даних?
3. Яким чином тест Фішера використовується для оцінки загальної значущості моделі? Як інтерпретувати його результати?
4. Для чого застосовується тест Стюдента в контексті парної лінійної регресії? Як він допомагає перевірити значущість окремих коефіцієнтів?
5. Що таке коефіцієнт детермінації (R^2), і як він відображає якість побудованої моделі? Які його межі та як їх інтерпретувати?
6. Як розрахувати коефіцієнт еластичності на основі параметрів регресійної моделі? Як цей показник дозволяє оцінити чутливість залежної змінної до змін незалежної змінної?
7. Які обмеження існують для використання парної лінійної регресії, і як ці обмеження впливають на точність прогнозування?



Практична робота №2

Побудова економетричної моделі на основі покрокової регресії

Мета практичної роботи

Ознайомитися з методом покрокової регресії для побудови економетричних моделей, навчитися обирати значущі предиктори для регресійної моделі, оцінювати її якість та інтерпретувати результати.

Терміни виконання практичної роботи

Практична робота виконується протягом двох занять:

Практичне заняття 1 — виконання завдань 1–3.

Практичне заняття 2 — виконання завдань 4–6, завантаження звіту та захист роботи.

Завдання

На основі статистичних даних значення показників продуктивності праці - Y (млн.грн./р.), фондомісткості X_1 (тис. грн), стажу роботи X_2 (років), текучості кадрів X_3 (в частках), середньої заробітної плати X_4 (тис. грн.) і втрат робочого часу X_5 (в частках). Необхідно застосувати алгоритм побудови покрокової регресії, відібрати значущі змінні та оцінити якість побудованих моделей.

Y	X_1	X_2	X_3	X_4	X_5
3,41	28,27	9	0,16	1,61	0,063
9,04	37,67	16	0,11	2,43	0,063
17,45	51,64	15	0,09	2,43	0,049
21,28	44,22	29	0,09	3,23	0,034
16,21	45,33	22	0,08	2,89	0,06
17,66	44,02	22	0,1	2,68	0,038
17,26	39,72	22	0,08	2,89	0,036
14,83	45,62	20	0,11	2,72	0,058
15,5	51,54	19	0,13	2,54	0,058
15,82	57,79	20	0,12	2,47	0,069
18,23	52,27	24	0,13	2,84	0,047
20,85	43,3	22	0,07	2,89	0,019
17,03	44,89	20	0,08	2,82	0,037
14,18	45,53	16	0,14	2,48	0,05
15,22	39,25	26	0,1	3,1	0,058
10,95	42,66	10	0,1	2,2	0,053
16,88	41,98	27	0,06	3,28	0,054
17,7	49,79	24	0,13	2,87	0,052
10,44	30,52	15	0,12	2,33	0,047
19,43	50,01	24	0,1	2,86	0,043
12,27	44,34	22	0,09	2,8	0,07

20,32	56,14	25	0,1	2,54	0,046
18,99	49,69	28	0,08	3,11	0,051
9,587	35,8	20	0,12	2,6	0,068

Для побудови моделі застосуємо алгоритм покрокової регресії. На першому кроці побудуємо п'ятифакторну лінійну модель

$$y = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + a_4 X_4 + a_5 X_5 + \varepsilon$$

Для цього у вкладці «Дані» виберемо пункт «Аналіз даних», в якому в свою чергу виберемо функцію «Регресія». Після цього з'явиться діалогове вікно (рис. 1), в якому потрібно ввести посилання на вхідні інтервали, виділяючи їх послідовно курсором. Слід зазначити, що посилання на значення факторів-аргументів має бути нерозривним. Також потрібно встановити інші необхідні параметри завдання (наприклад, «Виведення залишків») і натиснути «ОК». Після цього з'явиться звіт, як представлено на рис. 2.

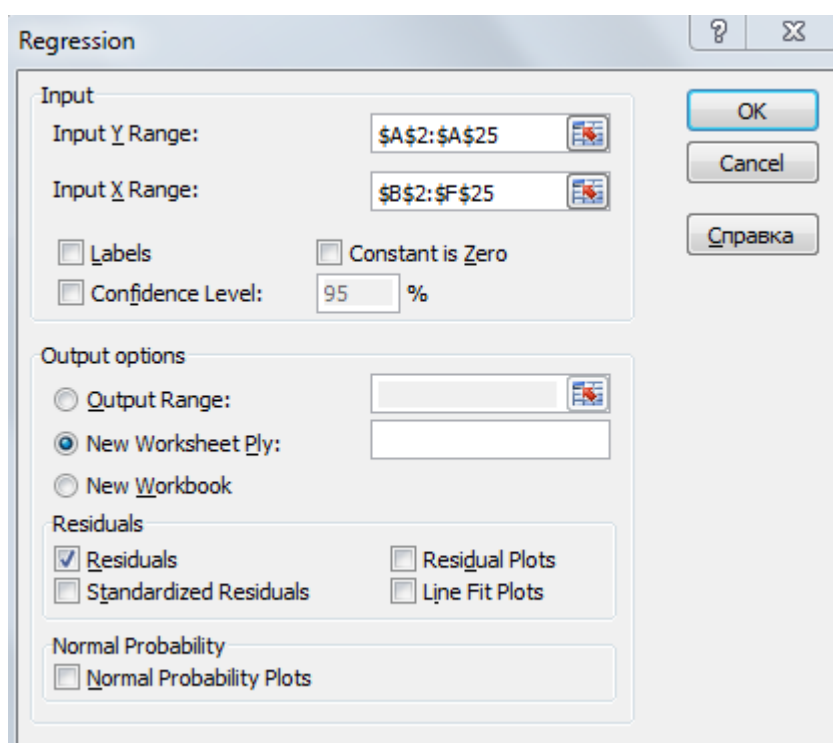


Рис. 1 - Діалогове вікно функції «Регресія»

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	<i>Regression Statistics</i>								
4	Multiple R	0,99153							
5	R Square	0,983132							
6	Adjusted R Square	0,978446							
7	Standard Error	0,629892							
8	Observations	24							
9									
10	ANOVA								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>gnificance F</i>			
12	Regression	5	416,239739	83,24795	209,8174	2,67E-15			
13	Residual	18	7,141747941	0,396764					
14	Total	23	423,381487						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
17	Intercept	1,465366	2,682084832	0,546353	0,591532	-4,16949	7,100217	-4,16949	7,100217
18	X Variable 1	0,299344	0,020256321	14,7778	1,65E-11	0,256787	0,341901	0,256787	0,341901
19	X Variable 2	0,200979	0,069423573	2,894965	0,00965	0,055125	0,346832	0,055125	0,346832
20	X Variable 3	-10,1363	8,221288555	-1,23294	0,233459	-27,4086	7,135942	-27,4086	7,135942
21	X Variable 4	2,043331	1,108255996	1,843736	0,081749	-0,28503	4,371691	-0,28503	4,371691
22	X Variable 5	-157,148	11,68338682	-13,4506	7,87E-11	-181,694	-132,602	-181,694	-132,602

Рис. 2 - Звіт функції «Регресія» (крок 1)

Побудова моделі множинної регресії методом виключення (включення) факторів в модель.

З отриманого звіту (рис. 2) випишемо рівняння регресії

$$y = 1,4654 + 0,2993x_1 + 0,201x_2 - 10,1363x_3 + 2,0433x_4 - 157,1481x_5$$

($R^2 = 0,9831$, $R^2_{ск} = 0,9784$).

Значимість рівняння множинної регресії в цілому визначається за допомогою статистичного F-критерію Фішера. У звіті функції «Регресія» (розділ «Дисперсійний аналіз») в комірці E12 зберігається значення Fфакт., яке можна використовувати для проведення F-тесту.

Можна використовувати другий підхід і перевірити виконання нерівності $P(F < F_{факт.}) \leq \alpha$ (де $\alpha = 0,05$ заданий рівень значимості). Для нашого прикладу $P(F < F_{факт.}) = 2,67 \cdot 10^{-15}$ - (рис. 2, значення в осередку F12). Так як виконується умова $P(F < F_{факт.}) \leq \alpha$ ($2,67 \cdot 10^{-15} < 0,05$), то п'ятифакторне рівняння регресії значимо з надійністю не менше 95%.

Перевірку значущості коефіцієнтів отриманого рівняння можна виконати двома способами. Перший - за критерієм Стьюдента. Другий - перевірити виконання нерівності $P(t < t_{j,факт.}) \leq \alpha$, $j = 1, 5$ (де $\alpha = 0,05$ заданий рівень значимості). Значення $P(t < t_{j,факт.})$ зберігаються в стовпці «P-значення» звіту функції «Регресія» (комірки E43: E48). Зауважимо, що вільний член a_0 зазвичай не перевіряється на статистичну значущість. У стовпці «P-значення» (рис. 2) тільки одне значення (комірка E46) для змінної X_3 більше заданого рівня значущості $\alpha = 0,05$. Отже оцінка коефіцієнта $a_3 = -10,1363$ не значима, оцінки

решти коефіцієнтів регресії $a_1 = 0,2993$, $a_2 = 0,201$, $a_4 = 2,0433$, $a_5 = -157,1481$ статистично значущі з достовірністю 95%.

На **другому** кроці побудуємо чотирьох-факторну лінійну модель

$$y = a_0 + a_1x_1 + a_2x_2 + a_4x_4 + a_5x_5.$$

З отриманого звіту (рис. 3) випишемо рівняння регресії

$$y = -0,9238 + 0,2999x_1 + 0,1720x_2 + 2,8268x_4 - 161,0162x_5$$

($R^2 = 0,982$, $R^2_{\text{скор.}} = 0,978$).

Аналізуючи значення P ($t < t_{j, \text{факт.}}$) В стовпці «P-значення» звіту функції «Регресія» (комірки E17: E21, рис. 3) робимо висновок, що всі вони задовольняють умові P ($t < t_{j, \text{факт.}}$) $\leq \alpha$, $j = 1, 2, 4, 5$.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	<i>Regression Statistics</i>								
4	Multiple R	0,9908							
5	R Square	0,9817							
6	Adjusted R Square	0,9779							
7	Standard Error	0,6385							
8	Observations	24							
9									
10	ANOVA								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>gnificance F</i>			
12	Regression	4	415,6366034	103,909	254,9133	3,2E-16			
13	Residual	19	7,744883524	0,40763					
14	Total	23	423,381487						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
17	Intercept	-0,9238	1,879519822	-0,4915	0,628683	-4,8577	3,01004	-4,8577	3,01004
18	X Variable 1	0,2999	0,02052645	14,6109	8,73E-12	0,25695	0,34287	0,25695	0,34287
19	X Variable 2	0,1720	0,06621572	2,59776	0,017669	0,03342	0,3106	0,03342	0,3106
20	X Variable 3	2,8268	0,920311333	3,0716	0,00628	0,9006	4,75307	0,9006	4,75307
21	X Variable 4	-161,0162	11,40728407	-14,115	1,6E-11	-184,89	-137,14	-184,89	-137,14

Рис. 3 - Звіт функції «Регресія» (крок 2)

Отже всі коефіцієнти моделі статистично значущі. Використовуючи дані звіту функції «Регресія» (розділ «Висновок залишків») розраховуємо середню відносну помилку апроксимації чотирьох-факторної моделі. Для цього в комірки D28:D51 (рис. 4) скопіюємо значення результативної ознаки Y , в комірки E28:E51 введемо формули

$$E_{\text{ср.відн.}}(i) = \frac{|e_i|}{y_i}$$

а в комірці E52

$$E_{\text{ср.відн.}} = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{y_i} \cdot 100\%$$


	A	B	C	D	E
25	RESIDUAL OUTPUT				
26					
27	<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>	<i>Y</i>	<i>Еср. Відн(i)</i>
28	1	3,5098972	-0,09989723	3,41	0,029295
29	2	9,8511377	-0,81113774	9,04	0,089728
30	3	16,123086	1,32691371	17,45	0,076041
31	4	20,982646	0,29735374	21,28	0,013973
32	5	14,963912	1,246087773	16,21	0,076872
33	6	17,519752	0,140247689	17,66	0,007942
34	7	17,145809	0,114191326	17,26	0,006616
35	8	14,548331	0,281668738	14,83	0,018993
36	9	15,642953	-0,14295263	15,5	0,009223
37	10	15,720343	0,099656723	15,82	0,006299
38	11	19,341178	-1,11117828	18,23	0,060953
39	12	20,95676	-0,10675958	20,85	0,00512
40	13	17,99342	-0,96342035	17,03	0,056572
41	14	14,442978	-0,26297788	14,18	0,018546
42	15	14,744181	0,47581873	15,22	0,031263
43	16	11,275599	-0,32559924	10,95	0,029735
44	17	16,887841	-0,00784128	16,88	0,000465
45	18	17,877127	-0,1771269	17,7	0,010007
46	19	9,8283482	0,611651816	10,44	0,058587
47	20	19,363984	0,066015687	19,43	0,003398
48	21	12,802425	-0,53242523	12,27	0,043392
49	22	19,986807	0,333192539	20,32	0,016397
50	23	19,374643	-0,38464314	18,99	0,020255
51	24	9,653839	-0,06683898	9,587	0,006972
52				Еср.відн	2,9%

Рис. 4 - Звіт «Висновок залишку» функції «Регресія» (крок 2) і розрахунок середньої відносної помилки апроксимації моделі

В результаті отримаємо значення середньої відносної помилки апроксимації чотирьох-факторної моделі $E_{\text{ср.відн.}} = 2,9\%$. Так як виконується умова $2,9\% < 7\%$, то якість моделі хороша. За значенням коефіцієнта детермінації $R^2 = 0,982$ робимо висновок, що отримане рівняння регресії пояснює коливання результативної ознаки Y на 98,2%, решта 1,8% припадають на фактори, невраховані в моделі. Отже, побудовану модель можна використовувати для подальшого економічного аналізу і прогнозу.

На підставі отриманого рівняння регресії

$$y = -0,92 + 0,3x_1 + 0,17x_2 + 2,83x_4 - 161,02x_5$$



зробимо наступні висновки. При збільшенні фондомісткості на 1 тис.грн (за умови незмінності інших чинників) продуктивність праці збільшується в середньому на 0,3 млн.грн./р. При збільшенні стажу роботи на 1 рік (за умови незмінності інших чинників) продуктивність праці збільшується в середньому на 0,17 млн.грн./р. При збільшенні середньої заробітної плати на 1 тис.грн. (за умови незмінності інших чинників) продуктивність праці збільшується в середньому на 2,83 млн.грн/р. При збільшенні втрат робочого часу на 1% (за умови, незмінності інших факторів) продуктивність праці зменшується в середньому на 1,61 млн.грн./р.

Для того, щоб зробити остаточний висновок про можливість використання моделі в економічному аналізі необхідно перевірити для побудованої моделі виконання умов Гаусса-Маркова (наявність мультиколінеарності між факторами-аргументами, гетероскедастичності та автокореляції збурень).

Питання для самоаналізу

1. Що таке покрокова регресія, і які основні етапи її застосування? У чому полягає різниця між Forward Selection, Backward Elimination та Stepwise Regression?

2. Які критерії використовуються для відбору змінних під час покрокової регресії? Як AIC (критерій інформації Акаїке) впливає на вибір моделі?

3. Як тест Фішера застосовується для оцінки загальної значущості моделі? Що означає значення F-статистики та як її інтерпретувати?

4. Як за допомогою тесту Стюдента перевіряється значущість окремих коефіцієнтів у моделі? Які висновки можна зробити, якщо р-значення для певної змінної більше за рівень значущості?

5. Що таке коефіцієнт детермінації (R^2) та скоригований коефіцієнт детермінації (R_{adj}^2)? Як вони використовуються для оцінки якості побудованої моделі?



Практична робота №3

Перевірка передумови відсутності мультиколінеарності між факторами-аргументами. Методи усунення мультиколінеарності

Мета практичної роботи

Ознайомитися з передумовою відсутності мультиколінеарності в економетричних моделях, навчитися діагностувати мультиколінеарність за допомогою статистичних показників (коефіцієнт кореляції, VIF, алгоритм Фаррара-Глобера), а також засвоїти основні методи її усунення.

Терміни виконання практичної роботи

Практична робота виконується протягом двох занять:

Практичне заняття 1 — виконання завдань 1–3.

Практичне заняття 2 — виконання завдань 4–6, завантаження звіту та захист роботи.

Завдання

Необхідно перевірити наявність мультиколінеарності між чотирма факторами-аргументами, вплив, яких на результативний фактор вивчалось в попередній роботі.

Для перевірки мультиколінеарності в масиві даних виконаємо розрахунки на аркуші MS Excel.

1. Побудуємо кореляційну матрицю за допомогою функціоналу «Аналіз даних» - «Кореляція» MS Excel (рис 4.1).

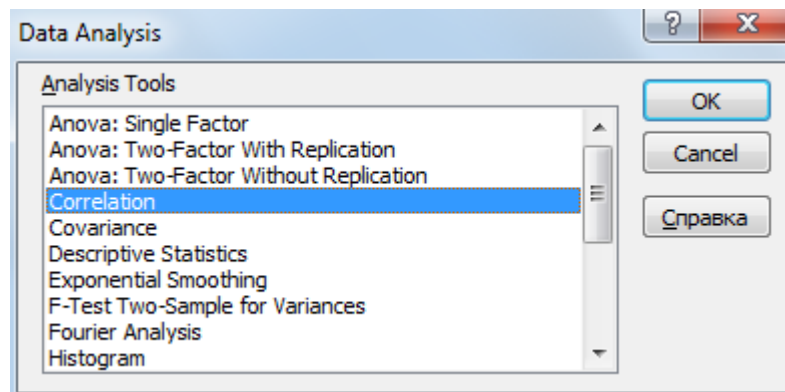
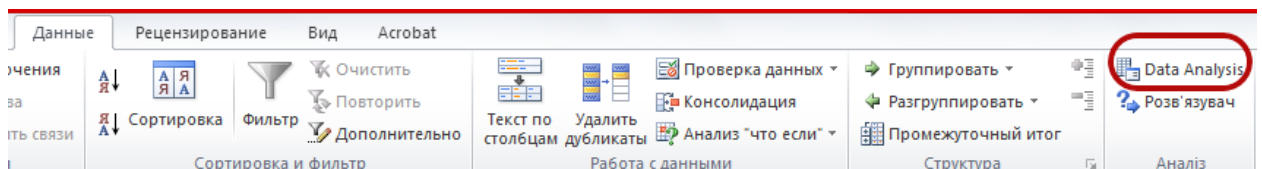


Рис 4.1. Побудова кореляційної матриці в MS Excel, лист 1

	A	B	C	D	E	F	G	H	I	J	K	L
1	Y	X1	X2	X4	X5							
2	3,41	28,27	9	1,61	0,063							
3	9,04	37,67	16	2,43	0,063							
4	17,45	51,64	15	2,43	0,049							
5	21,28	44,22	29	3,23	0,034							
6	16,21	45,33	22	2,89	0,06							
7	17,66	44,02	22	2,68	0,038							
8	17,26	39,72	22	2,89	0,036							
9	14,83	45,62	20	2,72	0,058							
10	15,5	51,54	19	2,54	0,058							
11	15,82	57,79	20	2,47	0,069							
12	18,23	52,27	24	2,84	0,047							
13	20,85	43,3	22	2,89	0,019							
14	17,03	44,89	20	2,82	0,037							
15	14,18	45,53	16	2,48	0,05							
16	15,22	39,25	26	3,1	0,058							
17	10,95	42,66	10	2,2	0,053							
18	16,88	41,98	27	3,28	0,054							
19	17,7	49,79	24	2,87	0,052							
20	10,44	30,52	15	2,33	0,047							
21	19,43	50,01	24	2,86	0,043							
22	12,27	44,34	22	2,8	0,07							
23	20,32	56,14	25	2,54	0,046							
24	18,99	49,69	28	3,11	0,051							
25	9,587	35,8	20	2,6	0,068							

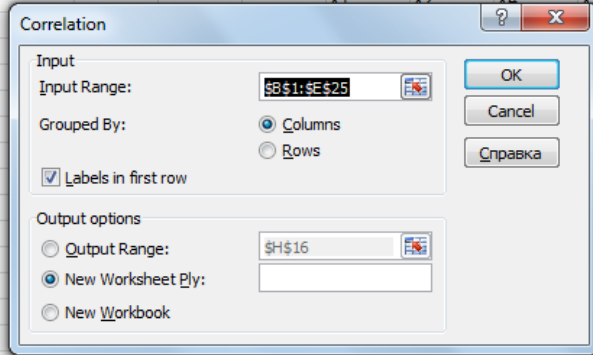


Рис 4.1. Побудова кореляційної матриці в MS Excel, лист 2

- Слідуючи алгоритму Фаррара-Глобера треба обчислити визначник кореляційної матриці $|r_{xx}|$ використовуючи функцію МОПРЕД.
- Наступним кроком треба обчислити критерій χ^2 :

$$\chi^2 = - \left(n - 1 - \frac{1}{6} (2m + 5) \right) \ln (r_{xx})$$

$\chi^2_{\text{табл}}$ знаходиться за допомогою функції ХИ2.ОБР.ПХ(вероятность;степени_свободы)

Крок 2.

2.1. Обчислимо матрицю похибок за допомогою функції МОБР.

2.2. Розрахуємо F-критерії:

$$F_k = \frac{(c_{kk} - 1)(n - m)}{m - 1}$$

$$C = r_{xx}^{-1} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \end{pmatrix}$$

За допомогою функції знайдемо табличне значення F-критерію.

Порівняємо значення F_k з табличним при $n-m$ і $m-1$ ступенями свободи і рівні значущості α (якщо $F_k > F_{\text{табл}}$, то відповідна k -та незалежна змінна мультиколінеарна з іншими).



	G	H	I	J	K	L	M
1		n=	24	m=	4		
2			X1	X2	X4	X5	
3		X1	1	0,41284	0,30937	-0,06958	
4	r=	X2	0,41284	1	0,910515	-0,26818	
5		X4	0,30937	0,910515	1	-0,32824	
6		X5	-0,06958	-0,26818	-0,32824	1	
7							
8	det r	0,121826					
9							
10	Chi^2розр	43,85746					
11	Chi^2табл	12,59159					
12		1,244406	-0,95252	0,478415	-0,01183		
13		-0,95252	6,614646	-5,79152	-0,19337		
14	C=	0,478415	-5,79152	6,305764	0,549919		
15		-0,01183	-0,19337	0,549919	1,127825		
16							
17	F1	1,629372		r12	0,332	t12	1,574028
18	F2	37,43097		r14	-0,17079	t14	0,775171
19	F4	35,37176		r15	0,009985	t15	0,044655
20	F5	0,852167		r24	0,896748	t24	9,062172
21				r25	0,070796	t25	0,317404
22	Ftabl=	8,66019		r45	-0,20621	t45	0,942452
23						ttabl=	2,085963

Рис. 1 - Організація даних та розрахунків за алгоритмом Фаррара-Глобера

3. Розрахуємо коефіцієнти часткової кореляції:

$$r_{kj} = \frac{-c_{kj}}{\sqrt{c_{kk}c_{jj}}}$$

Розраховуємо t- критерії Ст'юдента:

$$t_{kj} = \frac{|r_{kj}| \sqrt{n-m}}{\sqrt{1-r_{kj}^2}}$$

Порівняємо значення t_{kj} з табличним при $n-m$ ступенями свободи і рівні значущості α (якщо $t_{kj} > t_{табл}$ то між незалежними змінними x_k та x_j існує мультиколінеарність).

На підставі проведених розрахунків можна зробити такі висновки.

1. Так як $\chi_{\text{факт.}} > \chi_{\text{табл}}$ ($43,857 > 12,592$) то в масиві змінних X_1, X_2, X_4, X_5 існує мультиколінеарність.

2. Так як умова $F_{\text{факт.}} > F_{\text{табл.}}$ ($F_{\text{табл.}} = 8,66019$) виконується для статистик F_2 і F_4 то робимо висновок (з вірогідністю 95%) про статистичної значущості коефіцієнтів множинної кореляції показників X_2, X_4 і тісної лінійної залежності кожного з факторів з іншими.

3. Так як умова $t_{\text{факт.}} > t_{\text{табл.}}$ ($t_{\text{табл.}} = 2,086$) виконується тільки для статистики $t_{24} = 9,062$, то коефіцієнт приватної кореляції $R_{24} = 0,895$ статистично значимий. Отже, між показником стажу роботи X_2 (роки) і середньою заробітною платою X_4 (тис. грн.) існує тісна лінійна залежність (виключаючи вплив інших факторів).


Для того, щоб зменшити вплив мультиколінеарності на оцінки параметрів моделі виключимо змінну стаж роботи X_2 з моделі (так як $F_2 > F_4$). Використовуючи функцію «Регресія» оцінимо параметри трехфакторної моделі і перевіримо, чи вдалося позбутися мультиколінеарності (рис. 2).

На підставі значень параметрів моделі (рис. 2) запишемо рівняння трьохфакторної моделі $y = -4,03 + 0,32x_1 + 4,97x_4 - 158,92x_5$. Порівнюючи з моделлю, яка включала другий фактор $y = -0,92 + 0,3x_1 + 0,17x_2 + 2,83x_4 - 161,02x_5$ ($R_2^2 = 0,9908$) бачимо, що параметри a_1 і a_5 змінилися незначно, параметр a_4 - більш істотно. Коефіцієнт детермінації на третьому кроці зменшився незначно: $R_3^2 = 0,9875$.

	n=	24		m=	3
		X1	X4	X5	
	X1	1	0,30937	-0,06958	
r=	X4	0,30937	1	-0,32824	
	X5	-0,06958	-0,32824	1	
det r	0,805839				
Chi^2розр	4,569283				
Chi^2табл	7,814728				

Рис. 2. Перевірка мультиколінеарності пояснюючих чинників трьохфакторної моделі

Так як $\chi_{\text{факт.}} < \chi_{\text{табл}}$ ($4,569 < 7,815$), то в масиві змінних трьохфакторної моделі X_1, X_4, X_5 мультиколінеарності відсутня. Внаслідок того, що для побудови моделі використовувалися



«просторові» дані для остаточного висновку про якість моделі необхідно перевірити гіпотезу про відсутність гетероскедастичності збурень.

Питання для самоаналізу

1. Що таке мультиколінеарність, і чому її наявність може впливати на якість економетричної моделі?
2. Які статистичні інструменти використовуються для діагностики мультиколінеарності? Як інтерпретувати коефіцієнт кореляції, значення VIF (Variance Inflation Factor) і результати тесту Фаррара-Глобера?
3. Які значення VIF свідчать про наявність сильної мультиколінеарності? Як це впливає на оцінку регресійних коефіцієнтів?
4. Які методи усунення мультиколінеарності існують? Як практично застосувати методи зменшення кількості факторів, регуляризації або трансформації даних?
5. Як перевірити, що після усунення мультиколінеарності модель стала більш стабільною та її результати стали більш інтерпретованими?



Практична робота № 4

Перевірка передумови гомоскедастичності збурень. Узагальнений метод найменших квадратів оцінки параметрів моделі з гетероскедастичними залишками

Мета практичної роботи

Ознайомитися з передумовою гомоскедастичності залишків у регресійних моделях, навчитися діагностувати гетероскедастичність за допомогою статистичних тестів (тест Голдфельда-Квандта, Глейсера), а також засвоїти використання узагальненого методу найменших квадратів (GLS) для оцінювання параметрів моделі в умовах гетероскедастичності.

Терміни виконання практичної роботи

Практична робота виконується протягом двох занять:

Практичне заняття 1 — виконання завдань 1–3.

Практичне заняття 2 — виконання завдань 4–6, завантаження звіту та захист роботи.

Завдання

Для перевірки виконання умови гомоскедастичності для трьохфакторної моделі

$$y = -4,03 + 0,32x_1 + 4,97x_4 - 158,92x_5$$

Спочатку проведемо графічний аналіз залишків. Для цього при введенні параметрів функції «Регресія» потрібно відзначити поле «Графік залишків». В результаті при виведенні звіту функції «Регресія» будуть виведені графіки залишків (рис. 5.1 – 5.3).

Аналізуючи дані графіки, складно зробити висновок про гетероскедастичності залишків і змінної, яка є причиною цього явища. Виконаємо перевірку гетероскедастичності моделі за критерієм μ (рис. 5.4-5.5). Ці графіки лише є візуальним інструментом для оцінки розподілу залишків щодо передбачуваних значень моделі.

Аналізуючи отримані графіки залишків, можна побачити характер їх розподілу. Однак на основі лише графічного представлення часто складно зробити остаточний висновок щодо наявності гетероскедастичності. Це також ускладнює ідентифікацію конкретної незалежної змінної, яка може бути причиною цього явища.

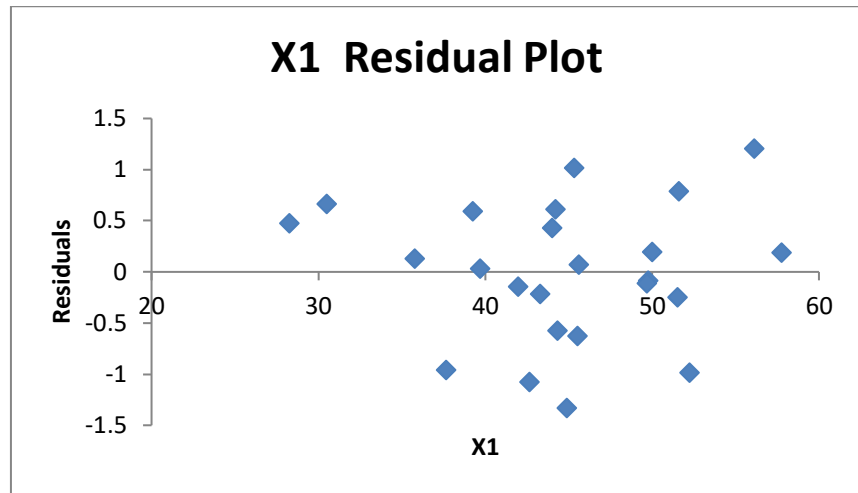


Рис. 5.1 Графік залишків по змінній X_1

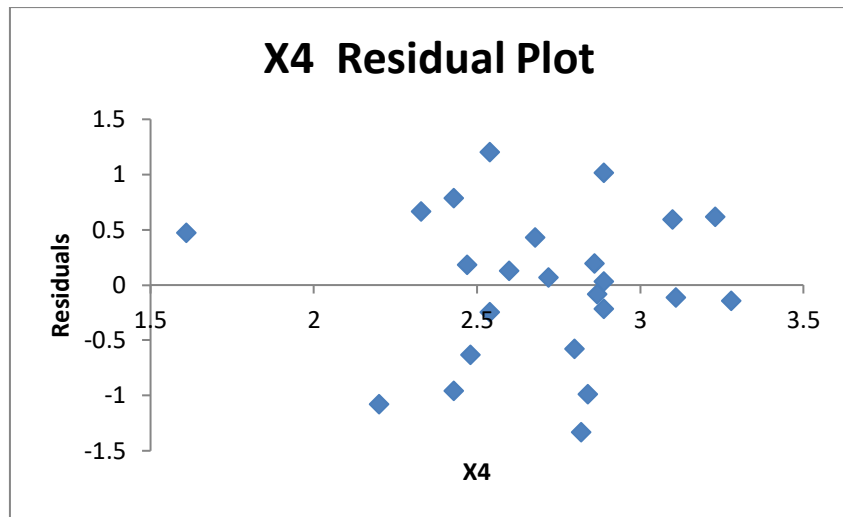


Рис. 5.2 Графік залишків по змінній X_4

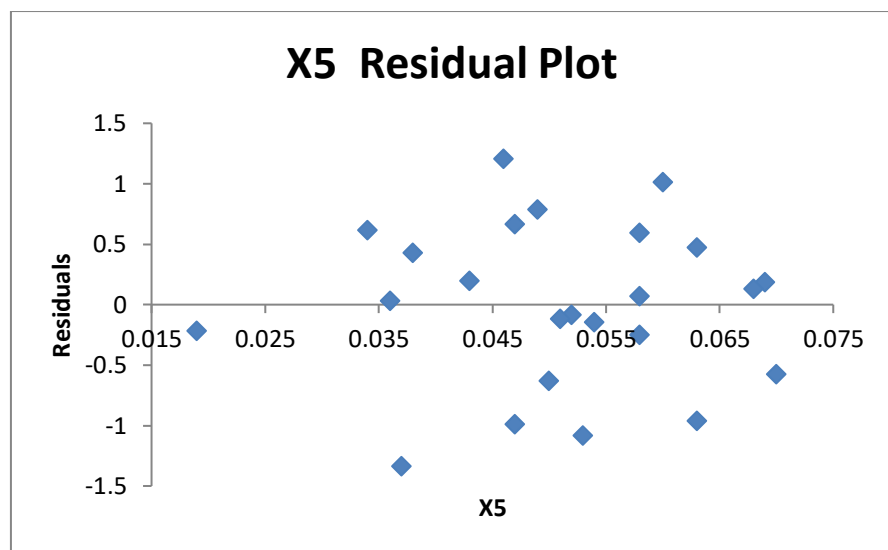


Рис. 5.3 Графік залишків по змінній X_5



	A	B	C	D	E	F
1	Y		Група 1	Група 2	Група 3	
2	3,41	1	3,41	15,22	17,66	
3	9,04	2	9,04	15,5	17,7	
4	9,587	3	9,587	15,82	18,23	
5	10,44	4	10,44	16,21	18,99	
6	10,95	5	10,95	16,88	19,43	
7	12,27	6	12,27	17,03	20,32	
8	14,18	7	14,18	17,26	20,85	
9	14,83	8	14,83	17,45	21,28	
10	15,22	Дисперсія	11,10014	0,630086	1,733794	
11	15,5	Сума квадрат	88,80109	5,040688	13,87035	107,7121
12	15,82		15181,45	0,157616	9,036281	66779051
13	16,21	w=	0,000324			24
14	16,88	m=	16,07084			
15	17,03					
16	17,26	ХИ^2tabl	5,991465			
17	17,45					
18	17,66					
19	17,7					
20	18,23					
21	18,99					
22	19,43					
23	20,32					
24	20,85					
25	21,28					

Рис. 5.4 Перевірка гетероскедастичності моделі за критерієм μ

	A	B	C	D	E	F
1	Y		Група 1	Група 2	Група 3	
2	3,41	1	3,41	15,22	17,66	
3	9,04	2	9,04	15,5	17,7	
4	9,587	3	9,587	15,82	18,23	
5	10,44	4	10,44	16,21	18,99	
6	10,95	5	10,95	16,88	19,43	
7	12,27	6	12,27	17,03	20,32	
8	14,18	7	14,18	17,26	20,85	
9	14,83	8	14,83	17,45	21,28	
10	15,22	Дисперсія	=ДИСПР(C2:C9)	=ДИСПР(D2:D9)	=ДИСПР(E2:E9)	
11	15,5	Сума квадрат	=C10*\$B\$9	=D10*\$B\$9	=E10*\$B\$9	=СУММ(C11:E11)
12	15,82		=C10^(\$B\$9/2)	=D10^(\$B\$9/2)	=E10^(\$B\$9/2)	=(F11/F13)^(F13/2)
13	16,21	w=	=ПРОИЗВЕД(C12:E12)/F12			24
14	16,88	m=	=-2*LN(C13)			
15	17,03					
16	17,26	ХИ^2tabl	=ХИ2.ОБР.ПХ(0,05;3-1)			

Рис. 5.5 Перевірка гетероскедастичності моделі за критерієм μ .

На підставі проведених розрахунків робимо висновок, що спостерігається гетероскедастичність залишків моделі так як $\mu > \chi^2_{\text{табл.}}$ ($16,07 > 5,99$).

Так як априорі важко визначити яка з пояснюють змінних є причиною гетероскедастичності, зробимо перевірку по кожній змінній за допомогою тесту Гольдфельда-Квандта.

Для цього впорядкуємо масив значень результативного і факторних ознак в комірках A2: D25 по змінній X_4 (рис. 5.6).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N						
1	Y	X1	X4	X5	SUMMARY OUTPUT				n= 24	Rfact=S1/S2=	3,301031902	SUMMARY OUTPUT								
2	3,41	28,27	1,61	0,063					c= 6	F(0,05;k1=c;k2=c)=	4,283865714									
3	10,95	42,66	2,2	0,053	Regression Statistics				m= 3	F(0,1;k1=c;k2=c)=	3,054550682	Regression Statistics								
4	10,44	30,52	2,33	0,047	Multiple F				0,990707	Multiple F				0,982944						
5	9,04	37,67	2,43	0,063	R Square				0,981501	R Square				0,966178						
6	17,45	51,64	2,43	0,049	Adjusted R Square				0,970401	Adjusted R Square				0,945885						
7	15,82	57,79	2,47	0,069	Standard Error				0,875948	Standard Error				0,482118						
8	14,18	45,53	2,48	0,05	Observations				9	Observations				9						
9	15,5	51,54	2,54	0,058																
10	20,32	56,14	2,54	0,046	ANOVA								ANOVA							
11	9,587	35,8	2,6	0,068					df	SS	MS	F	Significance F							
12	17,66	44,02	2,68	0,038	Regression				3	203,5477	67,84924	88,4277	9,42E-05	Regression				3	33,20017	
13	14,83	45,62	2,72	0,058	Residual				5	3,836425	0,767285	Residual				5	1,16219			
14	12,27	44,34	2,8	0,07	Total				8	207,3842	Total				8	34,36236				
15	17,03	44,89	2,82	0,037																
16	18,23	52,27	2,84	0,047	Coefficients				Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Coefficients						
17	19,43	50,01	2,86	0,043	Intercept				0,513876	4,088307	0,125694	0,904872	-9,99545	11,02320245	Intercept				-2,16611	4,206828
18	17,7	49,79	2,87	0,052	X Variable 1				0,360359	0,042529	8,473235	0,000376	0,251034	0,469682973	X Variable 1				0,29024	0,042871
19	16,21	45,33	2,89	0,06	X Variable 2				3,355629	1,624417	2,065743	0,093746	-0,82007	7,531326409	X Variable 2				4,610688	1,096001
20	17,26	39,72	2,89	0,036	X Variable 3				-206,549	40,59942	-5,08749	0,00381	-310,913	-102,1849064	X Variable 3				-144,418	13,18462
21	20,85	43,3	2,89	0,019																
22	15,22	39,25	3,1	0,058																
23	18,99	49,69	3,11	0,051																
24	21,28	44,22	3,23	0,034																
25	16,88	41,98	3,28	0,054																

Рис. 5.6 - Перевірка гетероскедастичності моделі за критерієм Гольдфельда-Квандта (по змінній X_4)

Визначимо кількість спостережень, яке необхідно відкинути

$$c = \frac{4}{15} \cdot n = \frac{4}{15} \cdot 24 = 4,8 \approx 5. \text{ Покладемо } c=6, \text{ щоб кількість}$$

залишених спостережень було парним. На наступному кроці побудуємо моделі лінійної регресії за першими і останніми дев'ятьма спостереженнями за допомогою функції «Регресія».

Для обчислення значення статистики $R^* = \frac{S_1}{S_2}$ розділимо значення суми квадратів відхилень по першій моделі S_1 (комірка G13 звіту функції «Регресія») на суму квадратів відхилень по другій моделі S_2 (комірка N13 функції «Регресія»). В результаті отримаємо $R_{\text{факт.}} = 3,301$. Потім визначимо табличне значення статистики при рівні значущості $\alpha=0,05$ і ступенях свободи $k_1 = k_2 = (n - c - 2m) / 2 = (24 - 6 - 2 \cdot 3) / 2 = 6$:

$$F_{\text{табл.}} (\alpha = 0,05, k_1 = 6, k_2 = 6) = 4,28.$$

Висновки за результатами перевірки передумови гомоскедастичності остатків.

На підставі проведених розрахунків робимо висновок (з надійністю 95%), що змінна X_4 не є причиною гетероскедастичності залишків моделі, так як

$$F_{\text{факт}}^4 < F_{\text{табл}} (3,301 < 4,284).$$

Аналогічно проводяться тести Гольфельд-Квандта по змінним X_1 і X_5 . В результаті отримуємо, що $F_{\text{факт}}^1 < F_{\text{табл}}$ ($2,758 < 4,284$) і $F_{\text{факт}}^5 < F_{\text{табл}}$ ($1,832 < 4,284$). Отже, змінні X_1 і X_5 не є причиною гетероскедастичності залишків моделі.

Так як за критерієм μ масив даних Y (24 спостереження) визнаний гетероскедастичних (при рівні значущості $\alpha = 0,05$), а жодна з змінних X_1, X_4, X_5 не є джерелом гетероскедастичності, то можна припустити, що причиною гетероскедастичності є фактори, невраховані в моделі. Також слід зазначити, що якщо в тесті Гольфельд-Квандта по змінній X_4 , задати рівень значущості $\alpha = 0,1$ то змінна X_4 буде визнана причиною гетероскедастичності залишків моделі, так як в цьому випадку табличне значення F-статистики при рівні значущості $\alpha = 0,1$ і ступенях свободи $k_1 = k_2 = 6$ дорівнює $F_{\text{табл}}$. ($\alpha = 0,1; k_1 = 6; k_2 = 6$) = 3,055 і $F_{\text{факт}}^4 > F_{\text{табл}}$ ($3,301 > 3,055$). Отже, для оцінки параметрів слід використовувати узагальнений МНК. Припустимо, що дисперсія залишків пропорційна квадрату значень змінної X_4 (умова $\sigma_{\varepsilon_i}^2 = \sigma^2 x_{ij}^2$, то $\lambda_i = \frac{1}{x_{ij}^2}, i = \overline{1, n}$). Тоді в матриці перетворень Ω , $\lambda_i = \frac{1}{x_{i4}^2}$. Використовуємо оператор оцінювання (умова $\sigma_{\varepsilon_i}^2 = \sigma^2 x_{ij}^2$, то $\lambda_i = \frac{1}{x_{ij}^2}, i = \overline{1, n}$) для оцінки параметрів УМНК (рис. 5.7, 5.8).

Де,

$$W = \begin{pmatrix} \frac{1}{l_1} & 0 & \dots & 0 \\ 0 & \frac{1}{l_2} & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \frac{1}{l_n} \end{pmatrix}$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	Y	X1	X4	X5																									
2		3,41	28,27	1,61	0,063		2,6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3		10,95	42,66	2,2	0,053		0	4,8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4		10,44	30,52	2,33	0,047		0	0	5,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5		9,04	37,67	2,43	0,063		0	0	0	5,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6		17,45	51,64	2,43	0,049		0	0	0	0	5,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7		15,82	57,79	2,47	0,069		0	0	0	0	0	6,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8		14,18	45,53	2,48	0,05		0	0	0	0	0	0	6,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9		15,5	51,54	2,54	0,058		0	0	0	0	0	0	0	6,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10		20,32	56,14	2,54	0,046		0	0	0	0	0	0	0	6,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11		9,587	35,8	2,6	0,068	W=	0	0	0	0	0	0	0	0	6,8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12		17,66	44,02	2,68	0,038		0	0	0	0	0	0	0	0	0	7,2	0	0	0	0	0	0	0	0	0	0	0	0	0
13		14,83	45,62	2,72	0,058		0	0	0	0	0	0	0	0	0	0	7,4	0	0	0	0	0	0	0	0	0	0	0	0
14		12,27	44,34	2,8	0,07		0	0	0	0	0	0	0	0	0	0	0	7,8	0	0	0	0	0	0	0	0	0	0	0
15		17,03	44,89	2,82	0,037		0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0
16		18,23	52,27	2,84	0,047		0	0	0	0	0	0	0	0	0	0	0	0	0	8,1	0	0	0	0	0	0	0	0	0
17		19,43	50,01	2,86	0,043		0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,2	0	0	0	0	0	0	0	0
18		17,7	49,79	2,87	0,052		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,2	0	0	0	0	0	0	0
19		16,21	45,33	2,89	0,06		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,4	0	0	0	0	0	0
20		17,26	39,72	2,89	0,036		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,3521	0	0	0	0	0
21		20,85	43,3	2,89	0,019		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8,4	0	0	0	0
22		15,22	39,25	3,1	0,058		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9,6	0	0	0
23		18,99	49,69	3,11	0,051		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9,7	0	0
24		21,28	44,22	3,23	0,034		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0
25		16,88	41,98	3,28	0,054		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
27			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	Xtr	28,27	42,66	30,52	38	52	58	46	52	56	36	44	46	44	44,9	52	50	50	45	39,7	43	39	49,69	44	42				
29		1,61	2,2	2,33	2,4	2,4	2,5	2,5	2,5	2,5	2,6	2,7	2,7	2,8	2,82	2,8	2,9	2,9	2,9	2,89	2,9	3,1	3,11	3,2	3,3				
30		0,063	0,053	0,047	0,1	0	0,1	0,1	0,1	0	0,1	0	0,1	0,1	0,04	0	0	0,1	0,1	0,04	0	0,1	0,051	0	0,1				
31																													
32		0,385788	0,206612	0,184199	0,2	0,2	0,2	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,13	0,1	0,1	0,1	0,1	0,12	0,1	0,1	0,10339	0,1	0,1				
33	Xtr*(W^1)	10,90622	8,81405	5,621765	6,4	8,7	9,5	7,4	8	8,7	5,3	6,1	6,2	5,7	5,64	6,5	6,1	6	5,4	4,76	5,2	4,1	5,13746	4,2	3,9				
34		0,621118	0,454545	0,429185	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,35	0,4	0,3	0,3	0,3	0,35	0,3	0,3	0,32154	0,3	0,3				
35		0,024305	0,01095	0,008657	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,00527	0	0				
36																													
37		3,550508	154,2928	9,107386	0,2								28	-0,1	-5,4	-195						50,5			-3,436				
38	Xtr*W^1*X	154,2928	6949,08	402,451	8			(Xtr*W^1*X)^1					-0,1	0	-0,1	-0,2				Xtr*W^1*Y	2312	A=		0,317					
39		9,107386	402,451	24	0,5								-5,4	-0,1	2,5	29,6					136			4,793					
40		0,185963	8,022534	0,469895	0								-195	-0,2	30	2450					2,51			-160,564					

Рис. 5.7 - Розрахунок параметрів моделі УМНК

	0,4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0,2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W^1=	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0,14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,12	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,11973	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,1	0	0

Рис. 5.8 - Розрахунок матриці Ω^{-1} в УМНК

Модель, побудована УМНК, матиме вигляд $y' = -3,4355 + 0,3168x_1 + 4,793x_4 - 160,564x_5$ з коефіцієнтом детермінації $R^2 = 97,5\%$ і середньої відносної помилкою апроксимації $E_{\text{ср.відн.}} = 3,85\%$.

Так як в масиві даних спостерігалася незначна гетероскедастичність, то якість моделей розрізняються незначно (менше 1%). Для прогнозу і економічного аналізу можна використовувати обидві моделі. Виберемо, наприклад, другу модель. Розрахуємо її параметри матричних способом і використовуємо для прогнозу і аналізу.

Порівняння економетричних моделей продуктивності праці

Модель	Метод	R^2	R^2 скорег.	$E_{\text{ср.відн.}}$
$y = -4,03 + 0,32x_1 + 4,97x_4 - 158,92x_5$	МНК	97,52%	97,15	4,07
$y' = -3,436 + 0,317x_1 + 4,793x_4 - 160,564x_5$	УМНК	97,5%	97,12	3,85%

Питання для самоаналізу

1. Що таке гомоскедастичність, і чому ця передумова є важливою для класичної регресійної моделі?
2. Які тести використовуються для перевірки гетероскедастичності залишків? Як інтерпретувати результати тестів Голдфельда-Квандта, Бреуша-Пагана та Уайта?
3. Як впливає гетероскедастичність на оцінки параметрів регресійної моделі та їх статистичну значущість?
4. У чому полягає принцип узагальненого методу найменших квадратів (GLS)? Як він допомагає враховувати гетероскедастичність залишків?
5. Які практичні кроки потрібно виконати для побудови моделі з використанням GLS?

Індивідуальна робота 1. Модель парної регресії

Мета індивідуальної роботи: навчитися будувати однофакторні регресійні моделі, оцінювати їх якість і статистичну значущість змінних з використанням MS Excel, а також проводити діагностику та інтерпретацію отриманих результатів.

Дані для виконання практичної роботи

1. Y- Витрати на харчування на одну особу за тиждень (грн),
X - середня зарплата за тиждень (грн).

Дані спостережень показників Y та X

Рік	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Y	376,4	301,2	329,2	378,2	371,7	387,4	416,1	448,6	489,9	414,1
X	604,8	627,1	663,3	691,1	686,5	701,4	740,7	777,8	812,8	834,8

2. Y-дохід підприємства (млн грн), X-кількість працюючих (тис. чол.).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8	9	10	11	12	13
Y	10,8	11,9	12,4	13,2	14,1	15,2	16,0	17,4	18,6	19,4	20,5	21,3	22,5
X	2,53	3,54	3,84	3,84	4,22	4,81	6,53	5,82	6,43	7,73	8,19	7,65	9,31

3. Y- Випуск продукції на місяць на підприємстві (тис. грн),
X - основний капітал (тис. грн).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8	9	10	11	12	13
Y	49	58	65	65	78	79	80	87	88	89	89	89	89
X	340	352	431	446	470	487	488	491	494	504	521	571	574

4. Y- Витрати на дозвілля (грн/день), X - дохід (грн/тиждень).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8	9	10	11
Y	13,5	18,0	31,5	49,5	85,5	126,0	144,0	187,2	243,4	316,4	411,3
X	486	567	666	810	1008	1260	1710	2223	2890	3757	4884

5. Y- Прибуток, млрд грн, X - обсяг виробленої продукції, млрд грн.

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8	9	10	11
Y	1,12	1,50	1,62	1,37	1,30	1,53	1,55	1,54	1,71	1,80	1,88
X	14,19	14,39	15,94	21,32	28,23	36,44	43,78	51,35	58,90	66,46	73,92

6. Y- Рівень рентабельності торгового підприємства (%),
X - частина продовольчих товарів у товарообігу (%).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7
Y	3,61	3,80	2,67	2,14	4,55	3,99	1,99
X	75,3	73,8	78,8	85,76	68,7	71,7	84,1

7. Y - витрати матеріалу на 1 одиницю продукції (грн), X – випуск продукції (тис. шт.).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8
Y	1,4	1,1	8,8	5,1	4,5	2,1	6,1	7,6
X	1,1	0,7	4,7	2,23	1,87	2,35	2,88	3,08

8. Y- рівень витрат у товарообігу (%), X - товарообіг (тис. грн).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8
Y	9	8	7,7	6,1	6,5	5,9	5,3	5,2
X	8	11	14	20	30	45	70	110

9. Y- товарні запаси магазину (тис. грн.), X – товарообіг (тис. грн.).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8
Y	2,9	4,2	4,5	4,9	5,7	6,3	5,8	4,1
X	40	60	64	77	83	95	90	80

10. Y- собівартість 1 одиниці продукції (грн), X-випуск продукції (тис. шт.).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6
Y	18	16	18	16	19	22
X	3,9	3,4	4,2	4,3	5,6	6,1

11. Y- експорт послуг України (млн дол.), X - інвестиції у сектор послуг України (млн дол.).

Дані спостережень показників Y та X

Рік	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Y	3 544,4	4 055,3	4 248	53 07,9	6 134,74	7 505,5	9 038,87	11 741,3	9 598,3	11 759,4	13 697
X	16 000	18 083	26 902	38 782	45 408	61 431	96 811	11 4817	70 624	8 8407	11 6378

12. Y- Експорт послуг України (млн дол.), X - індекс стану інфраструктури в цілому.

Дані спостережень показників Y та X

Рік	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Y	4 055,3	4 248,0	5 307,9	6 134,7	7 505,5	9 038,8	11 741,3	9 598,3	11 759,4	13 697,6
X	3,14	3,19	3,20	3,26	3,30	3,09	3,10	3,50	4,10	4,20

13. Y- експорт послуг України (млн дол.), X - чисельність зайнятого населення у сфері послуг та обсягу експорту послуг (тис. чол.).

Дані спостережень показників Y та X

Рік	2003	2004	2005	2006	2007	2008	2009	2010	2011
Y	4 248,0	5 307,9	6 134,7	7 505,5	9 038,8	11 741,3	9 598,3	11 759,4	13 697,6
X	8 116,7	8 234,2	8 426,9	8 637,0	8 882,8	9 034,4	8 763,3	8 784,1	8 851,4

14. Y- чисельність населення України зайнятого у сфері послуг (тис. чол.), X – інвестиції у сектор послуг.

Дані спостережень показників Y та X

Рік	2004	2005	2006	2007	2008	2009	2010	2011
Y	8234,2	8426,9	8637	8882,8	9034,4	8763,3	8784,1	8851,4
X	38782	45408,2	61431	96811,3	114817,2	70624	88407,3	116378

15. Y- чисельність населення зайнятого у сфері послуг та обсягу експорту послуг (тис. чол.), X – експорт послуг України (млн дол.).

Дані спостережень показників Y та X

Рік	2001	2002	2003	2004	2005	2006	2007	2008	2009
Y	7765,9	7981,4	8116,7	8234,2	8426,9	8637	8882,8	9034,4	8763,3
X	3544,4	4055,3	4248	5307,9	6134,74	7505,52	9038,87	11741,3	9598,33

16. Y- ціна акцій ВАТ «Астра» на Варшавській фондовій біржі (євро), X - біржовий індекс WIG-SPOZYW (індекс WIG-SPOZYW - є індексом суб-сектору, який включає компанії, що належать до сектору «Харчова промисловість» на Варшавській Фондовій Індекс розраховується з 1998 р., його початкове значення становило 1279,56 пунктів).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8
Y	8,93	9,07	4,36	2,64	2,64	4,42	5,76	9,74
X	2 923,1	2 657,8	1 953,4	1 555,5	1 431,1	1 700,3	2 821,6	3 008,4
Номер	9	10	11	12	13	14	15	16
Y	9,74	13,99	18,91	23,21	23,21	21	11,8	27,71
X	3 256,7	3 872,4	4 040,4	4 199,5	4 410,4	4 106,6	3 344,5	4 388,6

17. Y – ціна акції ВАТ «Астра» (виробника цукру) на Варшавській фондовій біржі (євро), X – ціна ф'ючерсу на цукор RSUGAOPEN на Варшавській товарній біржі (євро за фунт).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7
Y	8,93	9,07	4,36	2,64	2,64	4,42	5,76
X	0,5868	0,6259	0,7243	0,7757	0,8059	0,8313	1,0919
Номер	8	9	10	11	12	13	14
Y	9,74	13,99	18,91	23,21	23,21	21,00	11,80
X	1,2707	0,9413	1,0493	1,5233	1,8064	1,6256	1,7665

18. Y- товарообіг супермаркету (млн. грн.), X – площа (тис. м²).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8	9
Y	480	510	530	540	555	564	570	575	580
X	3	2,5	3,1	3,2	3,8	4,1	4,0	4,6	4,9

19. Y- врожайність цукрових буряків (ц/га), X - кількість внесених добрив (кг/га).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7	8
Y	247,3	280,2	270,5	295,6	320,1	312,8	336,2	320,4
X	83	92	112	102	124	134	132	149

20. Y- собівартість (грн/кг), X – кількості вирощеного врожаю соняшника (т/га).

Дані спостережень показників Y та X

Номер	1	2	3	4	5	6	7
Y	10,36	11,56	13,29	14,51	15,6	14,25	17,36
X	1,23	1,33	1,43	1,53	1,63	1,73	1,83



Індивідуальна робота 2.

Оцінка параметрів моделі за допомогою алгоритму покрокового відбору факторів

Мета індивідуальної роботи – ознайомитися з принципами покрокового відбору факторів для побудови економетричної моделі, навчитися оцінювати параметри моделі та визначати найбільш значущі змінні.

Завдання

Завдання 1. Оцінка параметрів моделі за допомогою алгоритму покрокового відбору факторів в модель

Вивчається взаємозв'язок наступних показників виробничо-господарської діяльності підприємств машинобудування: Y_1 - продуктивність праці; Y_2 - індекс зниження собівартості продукції; Y_3 - рентабельність; Y_4 - темп зростання прибутку; X_1 - трудомісткість одиниці продукції; X_2 - питома вага робітників у складі промислово-виробничого персоналу; X_3 - питома вага покупних виробів; X_4 - коефіцієнт змінності обладнання; X_5 - премії і винагороди на одного робітника; X_6 - питома вага втрат від браку; X_7 - фондвіддача; X_8 - середньорічна вартість основних виробничих фондів; X_9 - фондоозброєність праці; X_{10} - оборотність нормованих оборотних коштів; X_{11} - оборотність ненормованих оборотних коштів; X_{12} - невиробничі витрати.

Припускаючи, що між розглянутими показниками існує лінійна залежність, виконайте пункти 1); 2) (значення показників перебувають в табл. 1-2).

1. За допомогою «Пакета аналізу», застосувавши алгоритм покрокового регресійного аналізу, побудуйте рівняння регресії з максимальним числом значущих коефіцієнтів регресії (зробити висновки за коефіцієнтами детермінації на кожному кроці). Оцініть якість побудованої моделі, використовуючи коефіцієнт середньої відносної помилку апроксимації (для розрахунків використовуйте дані звіту «Висновок залишку»).

2. Використовуючи матричний підхід, побудуйте рівняння регресії з максимальним числом значущих коефіцієнтів регресії. Прийнятий рівень значущості $\alpha = 0,05$.

Завдання 2. Перевірка гіпотези відсутності мультиколінеарності між факторами-аргументами

Перевірте наявність мультиколінеарності між факторами-аргументами в моделі, побудованої в завданні 1. При виявленні мультиколінеарності запропонуйте заходи щодо її усунення.

Відбір показників для індивідуальної роботи

Номер варіанту	Номер показника а Y	Номер показника X	кількість спостереже нь <i>n</i>	номер варіанту	Номер показника Y	Номер показника X	кількість спостере жень <i>n</i>
1	1	1;2;3;11	35	11	2	7;3;5;12	30
2	1	1;2;11;4	34	12	2	6;7;9;12	32
3	1	6;3;4;5	27	13	3	3;4;12;6	29
4	1	1;3;4;12	32	14	3	4;10;6;7	33
5	1	1;8;5;12	33	15	3	3;4;12;11	31
6	1	7;12;4;5	31	16	3	3;8;10;9	33
7	2	6;7;10;11	34	17	3	3;8;10;12	35
8	2	6;7;4;11	26	18	4	2;3;4;12	29
9	2	7;4;5;12	25	19	4	2;9;4;5	30
10	2	6;7;3;5	26	20	4	2;3;8;5	35

Статистичні дані показників індивідуальної роботи

Но- мер	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
1	16,82	0,55	1,43	0,71	0,25	0,86	0,44	1,51	1,35	0,25	1,60	184,46	7,04	182,95	11,09	19,49
2	19,21	0,42	0,94	0,54	0,26	0,83	0,29	1,64	1,14	0,43	1,43	204,71	8,58	102,17	16,24	20,23
3	17,95	0,81	1,36	0,51	0,21	0,75	0,44	1,58	1,98	0,47	1,51	242,50	0,84	173,84	7,13	29,11
4	18,22	0,20	1,36	0,61	0,19	0,77	0,55	1,56	0,47	0,20	1,82	186,23	8,69	103,36	24,16	24,61
5	17,19	0,36	1,43	0,64	0,25	0,68	0,44	1,49	0,97	0,17	2,10	43,48	5,89	191,27	13,07	30,94
6	17,67	0,27	0,98	0,76	0,47	0,84	0,21	1,53	0,63	0,37	1,85	44,45	10,89	178,53	13,86	19,31
7	13,58	0,64	1,70	1,06	0,34	0,80	0,28	1,28	1,89	0,42	2,13	113,26	4,95	97,42	12,67	24,11
8	15,02	0,72	1,78	0,76	0,29	0,78	0,48	1,40	1,87	0,10	2,08	40,78	5,37	111,28	9,11	21,47
9	14,04	0,64	1,47	1,06	0,54	0,76	0,19	1,28	0,92	0,15	2,13	50,31	3,81	127,95	12,67	26,39
10	15,22	0,25	1,57	0,94	0,40	0,80	0,43	1,38	0,66	0,23	2,27	44,08	3,96	154,84	35,64	23,94
11	13,49	0,58	1,74	1,16	0,41	0,75	0,36	1,24	0,90	0,46	2,16	49,98	3,92	141,37	12,67	28,25
12	12,89	0,44	1,70	1,17	0,47	0,81	0,28	1,21	0,92	0,06	1,12	45,19	6,22	195,62	19,01	19,94
13	14,02	0,46	1,67	1,09	0,39	0,73	0,35	1,27	0,74	0,32	2,04	149,75	4,71	125,93	17,82	28,31
14	15,52	0,32	1,10	1,03	0,42	0,79	0,02	1,35	1,14	0,53	0,97	46,63	9,74	102,56	14,65	23,33
15	18,18	0,19	0,78	0,73	0,46	0,75	0,07	1,53	0,73	0,45	0,68	41,13	9,37	139,39	19,01	25,27
16	17,68	0,60	0,96	0,84	0,33	0,85	0,17	1,52	0,95	0,68	1,20	111,96	7,91	100,98	10,69	18,02
17	17,35	0,73	0,92	0,89	0,35	0,86	0,09	1,49	0,87	0,62	1,76	52,31	5,30	76,03	17,82	14,53
18	18,58	0,45	0,79	1,08	0,28	0,86	0,22	1,56	0,37	1,94	1,68	35,87	6,01	72,86	27,32	15,93
19	16,81	0,63	1,12	0,98	0,34	0,89	0,22	1,51	1,76	1,44	1,54	113,58	6,82	74,45	16,24	14,72
20	17,45	0,95	1,24	0,69	0,29	0,87	0,33	1,55	1,61	0,50	2,44	42,85	4,68	55,44	8,32	15,06
21	16,82	0,55	1,43	0,71	0,25	0,86	0,44	1,51	1,35	0,25	1,60	184,46	7,04	182,95	11,09	19,49
22	19,21	0,42	0,94	0,54	0,26	0,83	0,29	1,64	1,14	0,43	1,43	204,71	8,58	102,17	16,24	20,23
23	17,95	0,81	1,36	0,51	0,21	0,75	0,44	1,58	1,98	0,47	1,51	242,50	0,84	173,84	7,13	29,11
24	18,22	0,20	1,36	0,61	0,19	0,77	0,55	1,56	0,47	0,20	1,82	186,23	8,69	103,36	24,16	24,61
25	17,19	0,36	1,43	0,64	0,25	0,68	0,44	1,49	0,97	0,17	2,10	43,48	5,89	191,27	13,07	30,94
26	17,67	0,27	0,98	0,76	0,47	0,84	0,21	1,53	0,63	0,37	1,85	44,45	10,89	178,53	13,86	19,31
27	13,58	0,64	1,70	1,06	0,34	0,80	0,28	1,28	1,89	0,42	2,13	113,26	4,95	97,42	12,67	24,11
28	15,02	0,72	1,78	0,76	0,29	0,78	0,48	1,40	1,87	0,10	2,08	40,78	5,37	111,28	9,11	21,47
29	14,04	0,64	1,47	1,06	0,54	0,76	0,19	1,28	0,92	0,15	2,13	50,31	3,81	127,95	12,67	26,39
30	15,22	0,25	1,57	0,94	0,40	0,80	0,43	1,38	0,66	0,23	2,27	44,08	3,96	154,84	35,64	23,94
31	14,94	0,53	1,54	0,90	0,38	0,88	0,30	1,40	2,08	0,46	2,35	124,58	5,45	107,16	13,94	26,52
32	16,53	0,62	1,62	0,57	0,31	0,86	0,53	1,54	2,06	0,11	2,29	44,85	5,90	122,40	10,02	23,62
33	15,45	0,53	1,28	0,89	0,59	0,83	0,21	1,40	1,02	0,17	2,35	55,35	4,19	140,75	13,94	29,03
34	16,75	0,10	1,38	0,76	0,44	0,88	0,47	1,51	0,73	0,25	2,49	48,48	4,36	170,32	39,20	26,33
35	14,84	0,46	1,57	1,00	0,45	0,82	0,40	1,37	0,99	0,51	2,37	54,98	4,31	155,51	13,94	31,07



ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Економетрика з R : навчальний посібник / А. В. Скрипник, Д. М. Жерліцин, Ю. О. Нам'ясенко. Київ : ФОР Ямчинський О. В., 2020. 251 с.
2. Прикладна економетрика : навч. посіб. у 2-х ч. Ч. 1 / Л. С. Гур'янова та ін. Харків : ХНЕУ ім. С. Кузнеця, 2016. 248 с. URL: <http://repository.hneu.edu.ua/handle/123456789/19846> (дата звернення: 02.09.2024).
3. Cameron A. C. Analysis of Economics Data: An Introduction to Econometrics. Independently published, 2022. 533 p.
4. Chung-ki Min. Applied Econometrics: A Practical Guide. Routledge, 2019. 312 p.
5. Asteriou D., Hall S. G. Applied Econometrics. 4th edition. Bloomsbury Academic, 2021. 568 p.
6. Wooldridge J. M. Introductory Econometrics: A Modern Approach. 7th edition. Cengage Learning, 2019. 848 p.
7. Hill R. C., Griffiths W. E., Lim G. C. Principles of econometrics. 5th edition. New York : Wiley, 2018. 912 p.



Навчально-методичне видання

Жерліцин Дмитро Михайлович
Подскребко Олександр Сергійович

ЕКОНОМЕТРІЯ

**методичні рекомендації
до виконання практичних та індивідуальних робіт**

Самостійне електронне мережеве видання

Публікується в авторській редакції