

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

ЕКОНОМЕТРИКА

**Методичні рекомендації
до виконання практичних завдань
для здобувачів вищої освіти всіх спеціальностей
першого (бакалаврського) рівня**

**Харків
ХНЕУ ім. С. Кузнеця
2025**

УДК 330.43(072.034)

E40

Укладачі: Л. М. Малярець

О. В. Мінєнкова

Р. П. Шкребень

Затверджено на засіданні кафедри економіко-математичного моделювання.

Протокол № 9 від 20.10.2024 р.

Самостійне електронне текстове мережеве видання

Економетрика [Електронний ресурс] : методичні рекомендації до виконання практичних завдань для здобувачів вищої освіти всіх спеціальностей першого (бакалаврського) рівня / уклад. Л. М. Малярець, О. В. Мінєнкова, Р. П. Шкребень. – 2-ге вид., випр. та допов. – Харків : ХНЕУ ім. С. Кузнеця, 2025. – 82 с.

Подано вправи та наведено приклади й основні теоретичні відомості, необхідні для їх виконання.

Рекомендовано для здобувачів вищої освіти всіх спеціальностей першого (бакалаврського) рівня.

УДК 330.43(072.034)

© Харківський національний економічний
університет імені Семена Кузнеця, 2025

Вступ

На сьогодні характерною особливістю розвитку інструментів дослідження соціально-економічних процесів і явищ на різних рівнях їхнього управління є інтенсивне поширення математичних методів та моделей. Вони закладають фундаментальну основу вирішення реальних аналітичних завдань у всіх сферах діяльності суб'єктів господарювання.

Економетрика займає одне з провідних місць у системі підготовки економістів нового покоління, разом з іншими математичними й економічними навчальними дисциплінами формує нове економічне мислення в майбутніх фахівців.

Основними темами методичних рекомендацій є парна регресія й кореляція в економетричних дослідженнях; перевірка якості рівняння регресії; лінійні моделі множинної регресії; оцінювання надійності загальної багатофакторної лінійної моделі; мультиколінеарність, її наслідки та методи усунення; гетероскедастичність та методи її визначення; автокореляція залишків моделі та методи її усунення; проблеми інтерпретації параметрів багатофакторної моделі; системи одночасних рівнянь; динамічні економетричні моделі; моделювання одновимірних часових рядів; моделювання тенденції часового ряду.

Наведені приклади складають методичні рекомендації та положення щодо розв'язання завдань з економетрики. Вони не тільки роз'яснюють загальнотеоретичні положення, але й наочно показують можливі сфери застосування в економічному аналізі. Самостійне вирішення запропонованих завдань сприяє закріпленню теоретичної бази знань і формуванню практичних навичок, які необхідні для вирішення питань, пов'язаних із дослідженням економічних процесів і явищ, а також дає можливість виявити якість засвоєння матеріалу.

Набуті здобувачами вищої освіти знання, навички, вміння щодо формування економетричних задач та моделей в економіці, обґрунтування методів їхнього розв'язання, аналіз визначених результатів закладають фундамент компетентностей сучасного економіста.

Практичне заняття 1

Особливості економетричних моделей та принципи їхньої побудови. Парна лінійна модель

1.1. Приклади розв'язання завдань

Предметом економетрики є економічні та соціально-економічні системи, а саме модельний опис конкретних взаємозв'язків, які аналізують. Результатом застосування інструментів економетрики є побудова моделі. Економетрична модель містить набір регресійних рівнянь, що описують стохастичні зв'язки між досліджуваними економічними показниками, а також окремі тотожності, що характеризують співвідношення між економічними показниками.

Вважають, що найбільш поширеним математичним видом досліджуваних зв'язків є лінійний. Економетричне моделювання спрямоване на отримання двох типів кінцевих прикладних результатів, а саме: опис різних станів соціально-економічних систем, імітація різних можливих сценаріїв їхнього розвитку, а також прогнозування економічних показників, що характеризують стан і розвиток системи.

Парна регресія – рівняння зв'язку двох змінних y та x , де y – залежна змінна (результативна ознака); x – незалежна, пояснювальна змінна (ознака-фактор).

Розрізняють лінійні й нелінійні регресії.

Рівняння лінійної регресії має вигляд: $\hat{y}_x = a + bx$, де b – коефіцієнт регресії, a – вільний член рівняння регресії. Параметри рівняння лінійної регресії розраховують у результаті розв'язання системи нормальних рівнянь:

$$\begin{cases} n \cdot a + b \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \\ a \cdot \sum_{i=1}^n x_i + b \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i. \end{cases}$$

Варто зазначити, що нелінійну модель можна привести до виду лінійної.

Розгляньмо приклади побудови математичних моделей різних економічних задач.

Приклад 1.1. Побудувати лінійне рівняння парної регресії залежності прибутку y від вироблення продукції x за даними табл. 1.1 та провести його аналіз.

Таблиця 1.1

Дані для побудови та перевірки моделі парної лінійної регресії

y	x
133	78
148	82
134	87
154	79
162	89
195	106
139	67
158	88
152	73
162	87
159	76
173	115

Розв'язання

Рівняння лінійної парної регресії має такий вигляд: $\hat{y}_x = a + bx$, де b – коефіцієнт регресії; a – вільний член рівняння регресії.

Для визначення коефіцієнтів рівняння регресії слід побудувати табл. 1.2, для знаходження середніх значень ознак, дисперсії та коваріації.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ – середнє значення пояснювальної змінної } x,$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \text{ – середнє значення результативної ознаки } y,$$

$$\mu_{xy} = \overline{xy} - \bar{x}\bar{y} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x}\bar{y} \text{ – коефіцієнт коваріації,}$$

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 7\,492,25 - 85,583^2 = 167,743 \text{ – дисперсія фактора } x.$$

Знаходження середніх значень

№ з/п	y	x	xy	y^2	x^2
1	133	78	10 374	17 689	6 084
2	148	82	12 136	21 904	6 724
3	134	87	11 658	17 956	7 569
4	154	79	1 266	23 716	6 241
5	162	89	14 418	26 244	7 921
6	195	106	20 670	38 025	11 236
7	139	67	9 313	19 321	4 489
8	158	88	13 904	24 964	7 744
9	152	73	11 096	23 104	5 329
10	162	87	14 094	26 244	7 569
11	159	76	12 084	25 281	5 776
12	173	115	19 895	29 929	13 225
Σ	1 869	1 027	150 908	294 377	89 907
Сер. знач.	155,750	85,583	13 484,000	24 531,417	7 492,250

Коефіцієнти регресії визначають за співвідношеннями:

$$b = \frac{\mu_{xy}}{\sigma_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{13\,484 - 85,583 \cdot 155,75}{7\,492,25 - 85,583^2} = 0,92;$$

$$a = \bar{y} - b\bar{x} = 155,75 - 0,92 \cdot 85,583 = 76,976.$$

Визначено рівняння регресії $\hat{y}_x = 76,976 + 0,92x$.

Коефіцієнт регресії $b = 0,92$ вказує на те, що з підвищенням вироблення продукції x на одну одиницю результативна ознака прибутку y зросте на 0,92 одиниці в середньому.

Приклад 1.2. За умови прикладу 1.1 знайти лінійний коефіцієнт парної кореляції, коефіцієнт детермінації, середню помилку апроксимації та надати їхню інтерпретацію.

Розв'язання

Використовуючи табл. 1.2, знайти дисперсію результативної ознаки y : $\sigma_y^2 = \overline{y^2} - \bar{y}^2 = 24\,531,417 - 155,75^2 = 273,354$.

Для оцінювання щільності лінійного кореляційного зв'язку використовують коефіцієнт кореляції:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} = b \frac{\sigma_x}{\sigma_y} = 0,92 \frac{12,952}{16,533} = 0,721.$$

Величина коефіцієнта кореляції $r_{xy} = 0,721$ означає достатньо тісний зв'язок між ознаками.

Коефіцієнт детермінації $R^2 = r_{xy}^2 = 0,52$ вказує, що 52 % варіації ознаки y пов'язано з варіацією фактора x , а інші 48 % залежать від факторів, не врахованих у моделі.

Для перевірки якості моделі слід знайти середню помилку апроксимації. Для цього треба знайти теоретичні значення y за рівнянням $\hat{y}_x = 76,976 + 0,92x$ табл. 1.3.

Таблиця 1.3

Знаходження середньої помилки апроксимації

№ з/п	y	x	$\hat{y}_x$	$A = \left \frac{y_i - \hat{y}_i}{y_i} \right \cdot 100$
1	133	78	149	12,0
2	148	82	152	2,7
3	134	87	157	17,2
4	154	79	150	2,6
5	162	89	159	1,9
6	195	106	174	10,8
7	139	67	139	0
8	158	88	158	0
9	152	73	144	5,3
10	162	87	157	3,1
11	159	76	147	7,5
12	173	115	183	5,8

Середню помилку апроксимації знаходять за формулою:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \% = 5,758 \ \%.$$

Тобто в середньому теоретичні значення ознаки y відхиляються від своїх фактичних значень на 5,758 %, що є допустимим, бо не перевищує 8 – 10 %.

Приклад 1.3. За умови прикладу 1.1 – 2 перевірити статистичну значущість рівняння регресії загалом, коефіцієнтів регресії та коефіцієнта кореляції, побудувати довірчі інтервали для коефіцієнтів регресії з ймовірністю 95 %.

Розв'язання

Значущість моделі загалом слід перевірити за допомогою критерію Фішера. Для цього знайти його фактичне значення і порівняти з табличним:

$$F_{\text{факт.}} = \frac{r_{xy}^2}{1 - r_{xy}^2} \cdot \frac{(n - 2)}{1} = \frac{0,52}{1 - 0,52} \cdot \frac{(12 - 2)}{1} = 10,828.$$

Табличні значення знаходять за таблицями за різних рівнів значущості, так $F_{0,01}(1; 10) = 10,04$, $F_{0,05}(1; 10) = 4,96$ (додатки А і Б).

Отже, тому що $F_{\text{факт.}} > F_{0,01}$, то модель лінійної регресії значуща та кореляційний зв'язок між факторами також статистично значущий.

Статистичну значущість кожного з параметрів рівняння регресії окремо, а також коефіцієнта кореляції перевіряють за критерієм Стюдента. Фактичне значення критерію знаходять із співвідношень:

$$t_r = \frac{r}{m_r}; \quad t_b = \frac{b}{m_b}; \quad t_a = \frac{a}{m_a},$$

де m_r – середня помилка коефіцієнта кореляції; m_a , m_b – середні помилки параметрів лінійної регресії.

Використовуючи табл. 1.3, буде отримано $\sum (y_i - \hat{y}_i)^2 = 1574,939$.

$S = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{1574,939}{12-2}} = 12,55$ – стандартна помилка регресії.

Середні помилки відповідних величин визначають за формулами:

$$m_a = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2}} = S \frac{\sqrt{\sum x_i^2}}{n \sigma_x} = 12,55 \frac{\sqrt{89\,907}}{12 \cdot 12,952} = 24,212;$$

$$m_b = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2} \cdot \frac{1}{\sum (x_i - \bar{x})^2}} = \frac{S}{\sigma_x \cdot \sqrt{n}} = \frac{12,55}{12,952 \cdot \sqrt{12}} = 0,279;$$

$$m_r = \sqrt{\frac{1-r_{xy}^2}{n-2}} = \sqrt{\frac{1-0,52}{12-2}} = 0,219.$$

$$\text{Отже } t_r = \frac{0,721}{0,219} = 3,291; \quad t_b = \frac{0,92}{0,279} = 3,291; \quad t_a = \frac{76,976}{24,212} = 3,179.$$

Слід порівняти фактичні значення критерію Стьюдента та його табличне значення на 5 % рівні значущості $t_{0,05}(12-2) = 2,23$ (додаток В). Коефіцієнти регресії та коефіцієнт кореляції є статистично значущими, тому що для кожного з них $t_{\text{факт.}} > t_{\text{табл.}}$.

Для знаходження довірчих інтервалів треба скористатися формулами:

$$b \in (b \pm t_{\alpha, n-2} m_b); \quad a \in (a \pm t_{\alpha, n-2} m_a).$$

$$\text{Отже, } b \in (0,92 \pm 2,23 \cdot 0,279) = (0,297; 1,544);$$

$$a \in (76,976 \pm 2,23 \cdot 24,212) = (22,984; 130,969).$$

Приклад 1.4. Надати точковий та інтервальний прогноз прибутку залежно від вироблення продукції з ймовірністю 95 % за умови прикладу 1.1 – 2 та $x = 92$.

Розв'язання

Точковий прогноз знаходять за рівнянням регресії $\hat{y}_x = 76,976 + 0,92x$:

$$y_p = 76,976 + 0,92 \cdot 92 = 161,656.$$

Формула обчислення стандартної похибки прогнозованого значення:

$$m_{\hat{y}_p} = S \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{n\sigma_x^2}} = 12,55 \sqrt{1 + \frac{1}{12} + \frac{(92 - 85,583)^2}{12 \cdot 12,952^2}} = 13,185.$$

Обчислення прогнозного значення $\hat{y}_x$ з 95 % довірчим інтервалом виконують за формулою: $y_p \in (\hat{y}_p \pm t_\alpha \cdot m_{\hat{y}_p})$.

Отже, інтервальний прогноз:

$$y_p \in (161,656 \pm 2,23 \cdot 13,185) = (132,254; 191,058).$$

Приклад 1.5. За заданого рівняння парної регресії $\hat{y}_x = 3,2 \cdot x^{1,5}$ знайти середній коефіцієнт еластичності та здійснити його аналіз, якщо $\bar{x} = 1,8$.

Розв'язання

Щоб знайти коефіцієнт еластичності, знадобиться похідна від функції регресії: $\hat{y}_x' = 3,2 \cdot 1,5x^{0,5}$.

Коефіцієнт еластичності знаходять за формулою:

$$E = f'(x) \frac{\bar{x}}{y} = 3,2 \cdot 1,5x^{0,5} \frac{\bar{x}}{3,2 \cdot \bar{x}^{1,5}} = 1,5 \ %.$$

Він указує на те, що із збільшенням фактора x на 1 % від свого середнього значення, результативна ознака y збільшиться на 1,5 % від свого середнього рівня. Коефіцієнт еластичності відображає силу впливу фактора на результативну ознаку.

Приклад 1.6. Залежність вартості однокімнатної квартири від її площі за даними 40 квартир має такі характеристики: $\hat{y}_x = 5 \cdot x^{0,7}$, $R_{xy} = 0,8$, $\sigma_{зал}^2 = 0,12$. Провести дисперсійний аналіз отриманих результатів.

Розв'язання

Для дисперсійного аналізу варто скористатися формулами:

$$SS_{зал} = \sigma_{зал}^2 \cdot n = 0,12 \cdot 40 = 4,8;$$

$$SS_{загал} = SS_{зал} : (1 - R_{xy}^2) = 1 : (1 - 0,64) = 13,3;$$

$$SS_{фак} = SS_{загал} - SS_{зал} = 13,3 - 4,8 = 8,53;$$

$$F_{факт.} = \frac{R_{xy}^2}{1 - R_{xy}^2} \cdot \frac{(n - 2)}{1} = \frac{0,64}{1 - 0,64} \cdot \frac{(40 - 2)}{1} = 67,5.$$

Результати дисперсійного аналізу наведені в табл. 1.4.

Таблиця 1.4

Дисперсійний аналіз

Компоненти дисперсії	Число ступенів свободи df	Сума квадратів відхилень, SS	Дисперсія на одну ступінь свободи, MS	$F_{факт.}$	$F_{\alpha}(k_1, k_2)$
Загальна	$df = n - 1 = 39$	13,3	–	–	–
Факторна	$k_1 = m = 1$	8,53	8,53	67,5	4,08
Залишкова	$k_2 = n - m - 1 = 38$	4,8	0,126	–	–

Отже, через те що $F_{факт.} = 67,5 > F_{0,05} = 4,08$, гіпотеза про випадковість відхилень факторної та залишкової дисперсій відхиляється. Ця різниця значуща, рівняння регресії значуще, показник тісноти зв'язку надійний та відображає залежність вартості квартири від її площі.

1.2. Завдання для самостійної роботи

1.2.1. Проведено 15 спостережень та отримано такі результати:

$$\sum x_i = 23,19; \quad \sum y_i = 149,95; \quad \sum x_i y_i = 227,23;$$

$$\sum x_i^2 = 36,829; \quad \sum y_i^2 = 1523,986.$$

Оцінити коефіцієнти рівнянь регресії Y на X . Оцінити коефіцієнт кореляції r_{xy} , зробити висновки.

1.2.2. Залежність продуктивності праці (Y) від тривалості робочого дня (x), годин 20-ти заводів концерну характеризується таким чином:

Рівняння регресії: $Y = 200 + 0,2x - 0,06x^2$; частка остаточної дисперсії в загальній складає 25 %.

Визначити значущість рівняння регресії, використовуючи критерій Фішера ($\alpha = 0,1$).

Визначити коефіцієнт еластичності за умовою середньої тривалості робочого дня 7 годин.

1.2.3. За даними, поданими в табл. 1.5 про щотижневе споживання сировини на підприємстві, обчислити:

- а) коефіцієнт детермінації R^2 у регресії Y на X ;
- б) визначити значущість коефіцієнта детермінації.

Таблиця 1.5

Вихідні дані

$X, \$$	50	70	75	90	100	150	200	300	350	400
$Y, \text{кг}$	0.7	0.8	0.8	1	1.2	1.3	2	2.5	2.5	3

1.2.4. Проаналізувати залежність собівартості продукції від працёмісткості одиниці продукції: $y = 9,3 + 9,83 \cdot x$. Зробити висновок щодо коефіцієнта рівняння регресії. За допомогою середнього коефіцієнта еластичності визначити силу впливу фактора x ($\bar{x} = 1.38$).

1.2.5. За результатами 15-ти спостережень випадкових величин X та Y отримано такі дані:

$$\sum x_i = 26,61; \quad \sum y_i = 131,95; \quad \sum x_i y_i = 228,272; \quad \sum x_i^2 = 48,319; \\ \sum y_i^2 = 1194,959.$$

Припускаючи, що виконані передумови застосування МНК, оцінити:

- а) коефіцієнти a та b лінійного рівняння регресії та стандартні помилки коефіцієнтів;
- б) 90 % та 99 % довірчі інтервали для коефіцієнтів;
- в) коефіцієнт детермінації R^2 .

1.2.6. Для прогнозування можливого обсягу експорту ВВП (табл. 1.6) запропоновано використати лінійну регресійну модель.

Таблиця 1.6

Вихідні дані

Роки	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
ВНП	1 000	1 090	1 150	1 230	1 300	1 360	1 400	1 470	1 500	1 580
Експорт	190	220	240	240	260	250	280	290	310	350

Необхідно: побудувати відповідну регресійну модель, оцінити параметри моделі, подати їхню інтерпретацію;

обчислити стандартну помилку регресії; обчислити стандартні помилки коефіцієнтів;

визначити 90 та 95 %-і довірчі інтервали для теоретичних коефіцієнтів регресії;

проаналізувати статистичну значущість коефіцієнтів за умови рівнів значущості α та $\alpha = 0,05$;

оцінити коефіцієнт кореляції між ВНП та експортом;

обчислити коефіцієнт детермінації і порівняти його з коефіцієнтом кореляції;

навести прогнозне значення обсягу експорту на 2025 рік;

зробити висновки за побудованою моделлю.

1.2.7. Проаналізувати залежність собівартості продукції від державної частки прибутку: $y = 14,87 \cdot 1,016^x$. Зробити висновок щодо коефіцієнта рівняння регресії. За допомогою середнього коефіцієнта еластичності визначити силу впливу фактора x ($\bar{x} = 26,3$).

1.2.8. За 20-ма заводами отримано інформацію, наведену в табл. 1.7.

Таблиця 1.7

Вихідні дані

Показники	Середнє значення	Коефіцієнт варіації
Врожайність, ц/га	27	20
Кількість добрив, кг	5	15

Визначити лінійний коефіцієнт детермінації.

Побудувати рівняння лінійної регресії.

Множинні регресійні моделі: визначення МНК-оцінок параметрів моделі

Розв'язання

Результати обчислень середніх квадратичних відхилень та парних коефіцієнтів кореляції наведено в табл. 2.1.

Таблиця 2.1

Основні числові характеристики розподілу

Середнє вибірки	Середнє квадратичне відхилення	Парний коефіцієнт кореляції
$\bar{y} = 6,689$	$\sigma_y = 0,345$	$r_{yx_1} = 0,931$
$\bar{x}_1 = 13,167$	$\sigma_{x_1} = 1,672$	$r_{yx_2} = 0,808$
$\bar{x}_2 = 5,689$	$\sigma_{x_2} = 1,750$	$r_{x_1x_2} = 0,719$

Лінійне рівняння множинної регресії має такий вигляд:

$$\hat{y} = a + b_1x_1 + b_2x_2.$$

Для знаходження його параметрів слід використати метод стандартизації змінних та побудувати рівняння у стандартизованому масштабі:

$$\hat{t}_y = \beta_1 \cdot t_{x_1} + \beta_2 \cdot t_{x_2}.$$

Значення β -коефіцієнтів знаходять за формулами:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2} = \frac{0,931 - 0,808 \cdot 0,719}{1 - 0,719^2} = 0,724;$$

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2} = \frac{0,808 - 0,931 \cdot 0,719}{1 - 0,719^2} = 0,288.$$

Отже, є економетрична модель у стандартизованих змінних:

$$\hat{t}_y = 0,724 \cdot t_{x_1} + 0,288 \cdot t_{x_2}.$$

Оскільки $|\beta_1| > |\beta_2|$, можна стверджувати, що сила впливу x_1 на y більша ніж x_2 .

Для побудови рівняння регресії у натуральних змінних за допомогою β -коефіцієнтів знаходять $b_j = \beta_j \cdot \frac{\sigma_y}{\sigma_{x_j}}$, де $j = \overline{1,2}$:

$$b_1 = 0,724 \cdot \frac{0,345}{1,672} = 0,149; \quad b_2 = 0,288 \cdot \frac{0,345}{1,75} = 0,057;$$

$$a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 = 6,689 - 0,149 \cdot 13,167 - 0,057 \cdot 5,689 = 4,403.$$

Отже, економетрична модель у натуральній формі має такий вигляд:

$$\hat{y} = 4,403 + 0,149 \cdot x_1 + 0,057 \cdot x_2.$$

Коефіцієнт регресії $b_1 = 0,149$ вказує на те, що у разі збільшення кваліфікації праці на одну одиницю, то продуктивність праці зростає на 0,149 одиниць в середньому, а $b_2 = 0,057$ – збільшення автоматизації праці приведе до зростання продуктивності на 0,057 одиниць в середньому.

Приклад 2.2. За умовами прикладу 2.1 знайти та проаналізувати коефіцієнти множинної кореляції та детермінації; знайти коефіцієнти частинної кореляції, порівняти їх з коефіцієнтами парної кореляції та зробити висновки.

Розв'язання

Знайти коефіцієнт множинної кореляції:

$$R_{yx_1x_2} = \sqrt{r_{yx_1}\beta_1 + r_{yx_2}\beta_2} = \sqrt{0,931 \cdot 0,149 + 0,808 \cdot 0,288} = \sqrt{0,906} \approx 0,952.$$

Визначити коефіцієнт детермінації: $R_{yx_1x_2}^2 = 0,952^2 \approx 0,906$.

Отже, залежність y від x_1 та x_2 можна характеризувати як тісну, у якій мінливість результативної ознаки y на 90,6 % визначено мінливістю факторів x_1 та x_2 . Інші 9,4 % обумовлено впливом факторів, які не враховано в цій моделі.

Слід обчислити частинні коефіцієнти кореляції:

$$r_{yx_1|x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}} = \frac{0,931 - 0,808 \cdot 0,719}{\sqrt{(1 - 0,808^2)(1 - 0,719^2)}} \approx 0,854;$$

$$r_{yx_2|x_1} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}} = \frac{0,808 - 0,931 \cdot 0,719}{\sqrt{(1 - 0,931^2)(1 - 0,719^2)}} \approx 0,547.$$

Знайдено частинний коефіцієнт кореляції $r_{x_1x_2|y}$, що характеризує зв'язок між факторами x_1 та x_2 за фіксованого значення результативної ознаки y :

$$r_{x_1x_2|y} = \frac{r_{x_1x_2} - r_{yx_1} \cdot r_{yx_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{yx_2}^2)}} = \frac{0,719 - 0,931 \cdot 0,808}{\sqrt{(1 - 0,931^2)(1 - 0,808^2)}} \approx -0,154.$$

Через достатньо тісний зв'язок між факторами x_1 та x_2 коефіцієнти парної та частинної кореляції відрізняються ($r_{x_1x_2} = 0,719$). Висновки про тісний зв'язок між фактором x_1 та результативною ознакою збігаються, а для x_2 – відрізняються, бо за парним коефіцієнтом зв'язок більш тісний, ніж за частинним.

Приклад 2.3. Проведено дослідження на 51-му підприємстві (табл. 2.2), де y – фондоемність (грн/тис. грн), x_1 – витрати виробництва, що пов'язані з усуненням усіх видів простою (тис. грн), x_2 – витрати виробництва, що пов'язані з модернізацією встаткування й удосконалюванням техніки й технології виробництва (тис. грн).

Таблиця 2.2

Вихідні дані

Ознаки	Середнє значення	Середнє квадратичне відхилення	Коефіцієнт кореляції
y	25,12	3,72	$r_{yx_1} = -0,6375$
x_1	145,14	50,09	$r_{yx_2} = -0,6835$
x_2	122,69	36,75	$r_{x_1x_2} = 0,7941$

Побудувати економетричну модель у стандартизованих змінних й натуральних змінних.

Обчислити основні множинні характеристики, зробити висновки.

Розв'язання

Загальний вид моделі в стандартизованих змінних: $t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2}$.

Для оцінювання параметрів множинної регресії застосовують метод найменших квадратів (МНК). Наявна така система рівнянь:

$$\begin{cases} r_{yx_1} = \beta_1 + r_{x_1x_2} \beta_2, \\ r_{yx_2} = \beta_1 r_{x_1x_2} + \beta_2. \end{cases} \quad \begin{cases} -0,6375 = \beta_1 + 0,7941 \beta_2, \\ -0,6835 = 0,7941 \beta_1 + \beta_2; \end{cases} \quad \begin{cases} \beta_2 \approx -0,4800, \\ \beta_1 \approx -0,2563. \end{cases}$$

Таким чином, модель у стандартизованих змінних має вигляд:

$$t_y = -0,2563 \cdot t_{x_1} - 0,1141 \cdot t_{x_2}.$$

Завдяки тому, що $|\beta_1| > |\beta_2|$, можна стверджувати, що витрати x_1 , які пов'язані із модернізацією устаткування й удосконалюванням техніки та технології виробництва, найбільш сильно впливають на фондомісткість y .

Для побудови моделі в натуральних змінних $\hat{y} = a + b_1 x_1 + b_2 x_2$ знайти значення b_i за формулою: $b_i = \beta_i \cdot \frac{\sigma_y}{\sigma_{x_i}}$, $i = \overline{1,2}$, тоді $a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2$.

$$b_1 = \beta_1 \cdot \frac{\sigma_y}{\sigma_{x_1}} = -0,2563 \cdot \frac{3,72}{50,09} \approx -0,0191;$$

$$b_2 = \beta_2 \cdot \frac{\sigma_y}{\sigma_{x_2}} = -0,48 \cdot \frac{3,72}{36,75} \approx -0,0486;$$

$$a = 25,12 - (-0,0191) \cdot 145,14 - (-0,0486) \cdot 122,69 \approx 33,8501.$$

Отже, модель у натуральних змінних має вигляд:

$$y = 33,8501 - 0,0191 \cdot x_1 - 0,0486 \cdot x_2.$$

Знайти часткові коефіцієнти еластичності за формулою:

$$\overline{E_{x_i}} = b_i \cdot \frac{\overline{x_i}}{\overline{y}}, \quad i = \overline{1,2}.$$

$$\overline{E_{x_1}} = b_1 \cdot \frac{\overline{x_1}}{\overline{y}} = -0,0191 \cdot \frac{145,14}{25,12} \approx -0,11;$$

$$\overline{E_{x_2}} = b_2 \cdot \frac{\overline{x_2}}{\overline{y}} = -0,0486 \cdot \frac{122,69}{25,12} \approx -0,24.$$

Таким чином, можна зробити висновок, що зі збільшенням середньої величини витрат виробництва, пов'язаних з усуненням всіх видів простою (x_1), на 1 % середнє значення фондомісткості (y) зменшиться на 0,11 %; та зі збільшенням середньої величини витрат виробництва, пов'язаних з модернізацією устаткування й удосконалюванням техніки й технології виробництва (x_2), на 1 % середнє значення фондомісткості (y) зменшиться на 0,24 %.

Обчислити основні множинні характеристики моделі.

Знайти часткові коефіцієнти кореляції:

$$\begin{aligned} r_{yx_1x_2} &= \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}} = \frac{-0,6375 - (-0,6835) \cdot 0,7941}{\sqrt{(1 - (-0,6835)^2)(1 - 0,7941^2)}} = \\ &= \frac{-0,09473265}{\sqrt{0,196829336}} \approx -0,21356; \end{aligned}$$

$$\begin{aligned} r_{yx_2x_1} &= \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}} = \frac{-0,6835 - (-0,6375) \cdot 0,7941}{\sqrt{(1 - (-0,6375)^2)(1 - 0,7941^2)}} = \\ &= \frac{-0,17726125}{\sqrt{0,219276612}} \approx -0,3786; \end{aligned}$$

$$r_{x_1 x_2 y} = \frac{r_{x_1 x_2} - r_{yx_1} \cdot r_{yx_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{yx_2}^2)}} = \frac{0,7941 - (-0,6375) \cdot (-0,6835)}{\sqrt{(1 - (-0,6375)^2)(1 - (-0,6835)^2)}} =$$

$$= \frac{0,35836875}{\sqrt{0,316283222}} \approx 0,6372.$$

Оскільки $r_{x_1 x_2 y}$ досить близько до +1, то це означає, що міжфакторний зв'язок сильний, тому часткові коефіцієнти кореляції відрізняються від коефіцієнтів парної кореляції.

Визначити коефіцієнт множинної кореляції декількома способами:

$$R_{yx_1 x_2} = \sqrt{r_{yx_1} \beta_1 + r_{yx_2} \beta_2} = \sqrt{-0,6375 \cdot (-0,2563) + (-0,6835) \cdot (-0,48)} =$$

$$= \sqrt{0,49147125} \approx 0,7010.$$

У випадку, коли ми маємо два пояснювальних фактори, то визначник Δ_r має вигляд:

$$\Delta_r = \begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} \\ r_{yx_1} & 1 & r_{x_1 x_2} \\ r_{yx_2} & r_{x_1 x_2} & 1 \end{vmatrix}, \Delta_{11} = \begin{vmatrix} 1 & r_{x_1 x_2} \\ r_{x_1 x_2} & 1 \end{vmatrix}.$$

Таким чином, коефіцієнт множинної кореляції дорівнює:

$$R_{yx_1 x_2} = \sqrt{1 - \frac{0,1879}{0,3694}} = \sqrt{0,4913} \approx 0,7010.$$

Визначити коефіцієнт детермінації: $R_{yx_1 x_2}^2 = 0,7010^2 \approx 0,4915$.

Отриманий результат свідчить про помірну залежність y від факторів x_1 і x_2 , оскільки 49,15 % мінливості фондомісткості (y) пояснюють мінливістю величини витрат виробництва, пов'язаних з усуненням усіх видів простою (x_1), і величини витрат виробництва, пов'язаних із модернізацією устаткування й удосконалюванням техніки та технології виробництва (x_2). Залишок 50,85 % – вплив неврахованих факторів.

2.2. Завдання для самостійної роботи

2.2.1. Скласти рівняння множинної регресії у натуральних змінних, якщо $\beta_1 = 0,838$; $\beta_2 = 0,557$; $n = 20$; $\sum x_1 = 62,80$; $\sum x_2 = 54,40$; $\sum y = 383,2$; $S^2(x_1) = 0,98$; $S^2(x_2) = 4,02$; $S^2(y) = 10,8$. Обчислити середні часткові коефіцієнти еластичності, зробити висновки.

2.2.2. Побудувати економетричну модель за даними, що наведені в табл. 2.3. Оцінити параметри моделі. Зробити економічні висновки.

Таблиця 2.3

Вихідні дані

Номер підприємства	1	2	3	4	5
Витрати на перевезення, грн	25,2	32,3	28,1	26,1	32,5
Збір овочів з 1 га, ц	52,8	72,6	50,4	55,2	31,5
Рівень прибутковості, %	31,4	40,9	37,1	45,7	24,5

2.2.3. На 30 заводах вивчають залежність споживання електроенергії (y) від обсягу виробництва на одного робітника (x_1) та рівня автоматизації праці (x_2). Дані наведено в табл. 2.4.

Таблиця 2.4

Вихідні дані

Ознаки	Середнє значення	Середнє квадратичне відхилення	Парний коефіцієнт кореляції
y	600	17	$r_{yx_1} = 0.77$
x_1	210	22	$r_{yx_2} = 0.43$
x_2	21,2	8	$r_{x_1x_2} = 0.38$

Знайти часткові коефіцієнти еластичності.

Побудувати множинну регресію у натуральному масштабі.

2.2.4. На 40-ка заводах вивчають залежність продуктивності праці (y) від рівня кваліфікації робітника (x_1) та рівня автоматизації праці (x_2). Дані наведено в табл. 2.5.

Таблиця 2.5

Вихідні дані

Ознаки	Середнє значення	Середнє квадратичне відхилення	Парний коефіцієнт кореляції
y	200	5,2	$r_{yx_1} = 0.77$
x_1	4	1,2	$r_{yx_2} = 0.54$
x_2	6,3	0,9	$r_{x_1x_2} = 0.65$

Побудувати множинну регресію у стандартизованому та натуральному масштабах, надати економічну інтерпретацію коефіцієнтів регресії. Знайти часткові коефіцієнти еластичності та зробити висновки.

2.2.5. За даними діяльності 20-ти підприємств легкої промисловості отримано інформацію, що характеризує залежність обсягів виробництва продукції (млн грн) y від кількості опрацьованих за рік людино-годин (тис. люд.-годин) x_1 і середньорічної вартості виробничого обладнання (млн грн) x_2 (табл. 2.6).

Таблиця 2.6

Вихідні дані

Рівняння регресії	$y = 35 + 0,06x_1 + 2,5x_2$
Множинний коефіцієнт кореляції	0,9
Сума квадратів відхилень обчислених значень результатів від фактичних	3 000

Знайти коефіцієнт детермінації у цій моделі. Скласти таблицю результатів дисперсійного аналізу. Проаналізувати отримані результати регресійного аналізу.

Практичне заняття 3

Множинні регресійні моделі: перевірка значущості економетричної моделі в цілому і кожного з її параметрів окремо

3.1. Приклади розв'язання завдань

Якість побудованої моделі в цілому можна оцінити за допомогою коефіцієнта (індексу) детермінації: $R^2 = R^2_{yx_1 \dots x_n}$.

Значущість рівняння множинної регресії у цілому оцінюють за допомогою F -критерію Фішера, розрахункове значення якого одержане за формулою: $F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}$, де n – кількість спостережень; m – кількість факторів; R^2 – коефіцієнт (індекс) детермінації.

Доцільність присутності кожного з факторів у множинній моделі встановлюють за допомогою часткових F -критеріїв Фішера, які розраховують для кожного фактора за формулою:

$$F_{частx_i} = \frac{R^2 - R^2_{yx_1 \dots x_{i-1} x_{i+1} \dots x_n}}{1 - R^2} \cdot \frac{n-m-1}{1}.$$

Оцінка значущості коефіцієнтів регресії на основі t -критерію Стюдента приводить до обчислення: $t_{b_i} = \frac{b_i}{m_{b_i}} = \sqrt{F_{x_i}}$, де m_{b_i} – середня квадратична помилка коефіцієнта регресії:

$$m_{b_i} = \frac{\sigma_y \sqrt{1-R^2}}{\sigma_{x_i} \sqrt{1-R^2_{x_1 x_1 \dots x_n}}} \cdot \frac{1}{\sqrt{n-m-1}}.$$

Приклад 3.1. За умови прикладу 2.1 перевірити значущість побудованої моделі в цілому, окремих її параметрів та факторів.

Розв'язання

Перевірити значущість рівняння множинної регресії загалом слід за критерієм Фішера. За формулою визначають емпіричне значення критерію:

$$F_{emn.} = \frac{R_{yx_1x_2}^2}{1 - R_{yx_1x_2}^2} \cdot \frac{n - m - 1}{m} = \frac{0,906}{1 - 0,906} \cdot \frac{9 - 2 - 1}{2} \approx 29,01.$$

Потрібно визначити критичні значення розподілу Фішера для $\alpha = 0,05$ і $\alpha = 0,01$: $F_{0,05}(2;6) = 5,14$ і $F_{0,01}(2;6) = 10,93$ (див. додатки А і Б). Оскільки $F_{emn.} > F_{0,01}$, то з надійністю 99 % можна стверджувати, що регресія загалом є значущою.

Перевіряють доцільність наявності кожного з факторів в економетричній моделі за допомогою частинних F -критеріїв. Обчислюють їхні емпіричні значення за формулами:

$$F_{x_1} = \frac{R^2 - r_{yx_2}^2}{1 - R^2} \cdot \frac{n - m - 1}{1} = \frac{0,906 - 0,808^2}{1 - 0,906} \cdot \frac{9 - 2 - 1}{1} \approx 16,17;$$
$$F_{x_2} = \frac{R^2 - r_{yx_1}^2}{1 - R^2} \cdot \frac{n - m - 1}{1} = \frac{0,906 - 0,931^2}{1 - 0,906} \cdot \frac{9 - 2 - 1}{1} \approx 2,56.$$

Слід порівняти ці емпіричні значення із критичними $F_{0,05}(1; 6) = 5,99$ і $F_{0,01}(1; 6) = 13,75$. Для фактора x_1 буде $F_{emn.} > F_{0,01}$, тому з надійністю 99 % вплив фактора x_1 на ознаку y не можна вважати несуттєвим, тобто доцільно включити в модель фактор x_1 після фактора x_2 . Для фактора x_2 $F_{emn.} < F_{0,05}$, тобто гіпотезу про відсутність впливу цього фактора на ознаку y немає підстав відхилити, отже, його наявність в економетричній моделі є недоцільною.

У процесі аналізу y доцільно використовувати модель парної регресії, яка враховує тільки вплив фактора x_1 . У стандартизованих змінних ця модель має такий вигляд: $\hat{t}_y = 0,931 \cdot t_{x_1}$.

Після перетворення визначено модель у натуральних змінних:

$$\hat{y} = 4,163 + 0,192x_1.$$

Перевіряють статистичну значущість коефіцієнтів регресії за критерієм Стюдента:

$$t_{b_1} = \sqrt{F_{x_1}} = \sqrt{16,17} = 4,02;$$

$$t_{b_2} = \sqrt{F_{x_2}} = \sqrt{2,56} = 1,6.$$

Знаходять табличне значення критерію $t_{0,05}(6) = 2,45$ (див. додаток В). Порівнюючи табличне та фактичне значення, можна бачити, що коефіцієнт b_1 значущий, оскільки $t_{b_1} > t_{0,05}$, а b_2 , навпаки, статистично не значущий, тому що $t_{b_2} < t_{0,05}$. Це збігається з висновками щодо доцільності включення факторів x_1 та x_2 у модель.

Приклад 3.2. За умови прикладу 2.3 оцінити значущість регресії за допомогою F -критерію Фішера. Установити доцільність присутності факторів у рівнянні регресії за допомогою часткових F -критеріїв Фішера.

Розв'язання

Оцінити значущість рівняння регресії за допомогою F -критерію Фішера.

Визначити розрахункове значення загального F -критерію Фішера:

$$F = \frac{0,4915}{1 - 0,4915} \cdot \frac{51 - 2 - 1}{2} = 23,194.$$

Визначити табличні значення F -критерію Фішера з параметрами $k_1 = m = 2$ та $k_2 = n - m - 1 = 51 - 2 - 1 = 48$ за умови рівня значущості $\alpha = 0,05$ й $\alpha = 0,01$: $F_{табл}(\alpha = 0,05) \approx 3,185$; $F_{табл}(\alpha = 0,01) \approx 5,1$.

Оскільки розрахунковий F -критерій Фішера перевищує табличні значення F -критерію Фішера, можна із упевненістю 99 % стверджувати, що рівняння регресії значуще.

Установимо доцільність присутності кожного з факторів у моделі за допомогою часткових F -критеріїв, які розраховують для кожного фактора:

$$F_{x_1} = \frac{R^2 - r_{yx_2}^2}{1 - R^2} \cdot \frac{n - m - 1}{1}; \quad F_{x_2} = \frac{R^2 - r_{yx_1}^2}{1 - R^2} \cdot \frac{n - m - 1}{1}.$$

Звідси, наявні такі значення розрахункових часткових F -критеріїв:

$$F_{x_1} = \frac{0,4915 - (-0,6835)^2}{1 - 0,4915} \cdot \frac{51 - 2 - 1}{1} \approx 2,291;$$

$$F_{x_2} = \frac{0,4915 - (-0,6375)^2}{1 - 0,4915} \cdot \frac{51 - 2 - 1}{1} \approx 8,033.$$

Визначити табличні значення F -критерію за умови $\alpha = 0,05$ й $\alpha = 0,01$ з параметрами $k_1 = 1$ та $k_2 = n - m - 1 = 48$:

$$F_{\text{табл}}(1; 49) = \begin{cases} \alpha = 0,05 & 4,035 \\ \alpha = 0,01 & 7,2 \end{cases}.$$

Оскільки розрахунковий F -критерій для фактора x_1 не перевищує табличні значення, то присутність цього фактора (витрат виробництва, пов'язаних з усуненням усіх видів простою) у моделі недоцільно. Оскільки розрахунковий F -критерій для фактора x_2 перевищує табличні значення, то із упевненістю 99 % можна затверджувати, що присутність фактора витрат виробництва, пов'язаних із модернізацією устаткування й удосконалюванням техніки й технології виробництва доцільно.

3.2. Завдання для самостійної роботи

3.2.1. На 15 підприємствах вивчають залежність прибутку (y) від обсягу виробництва на 1 робітника (x_1) та індексу ціни на продукцію (x_2). Дані наведено в табл. 3.1.

Вихідні дані

Ознаки	Середнє значення	Середнє квадратичне відхилення	Парний коефіцієнт кореляції
y	250	38	$r_{yx_1} = 0.68$
x_1	47	12	$r_{yx_2} = 0.63$
x_2	112	21	$r_{x_1x_2} = 0.42$

Побудувати лінійні моделі парної регресії $y(x_1)$ та $y(x_2)$, зробити висновок про якість кожної з моделей.

Визначити доцільність присутності факторів x_2 та x_1 в економетричній множинній моделі за допомогою часткових критеріїв Фішера.

3.2.2. Проведено аналіз залежності рівня рентабельності підприємств від рівня інвестицій – x_1 та обсягів виробництва – x_2 на 50 підприємствах. Отримано $R^2 = 0.86$, $r_{yx_1} = 0.65$, $r_{yx_2} = 0.52$.

Зробити висновок щодо коефіцієнта детермінації.

Оцінити значущість рівняння множинної регресії за критерієм Фішера, а також установити доцільність присутності факторів у рівнянні регресії.

3.2.3. За 20-ма актами купівлі-продажу однокімнатної квартири отримано дані про вартість квартири (тис. дол. США) y , її загальну площу (m^2) x_1 та відстань від неї до метро (хв) x_2 :

$$\sum x_1 = 739; \quad \sum x_2 = 180; \quad \sum y = 734; \quad \sum x_1^2 = 27551; \quad \sum x_1x_2 = 6615; \\ \sum x_2^2 = 1806; \quad \sum y^2 = 28020; \quad \sum yx_1 = 27513; \quad \sum yx_2 = 6357.$$

Побудувати лінійне рівняння множинної регресії. Знайти коефіцієнт множинної детермінації, у тому числі скоригований; зробити висновки.

Оцінити значущість рівняння регресії через F -критерій Фішера.

Оцінити значущість коефіцієнтів регресії через t -критерій Стьюдента.

Дати інтервальну оцінку коефіцієнтів регресії (з імовірністю 0,95).

Дати інтервали прогнозу вартості квартири з імовірністю 0,95, припускаючи прогнозні значення $x_1 = 40$, $x_2 = 5$.

3.2.4. У процесі вивчення залежності обсягів реалізації y від розміру торгової площі x_1 і товарних запасів x_2 за 19-ма підприємствами оптової торгівлі було отримано такі варіанти рівнянь регресії:

1) $y = 25 + 15x_1$, $r^2 = 0,90$;

2) $y = 42 + 27x_2$, $r^2 = 0,84$;

3) $y = 30 + 10x_1 + 8x_2$, $R^2 = 0,92$;
(2,5) (4,0)

4) $y = 21 + 14x_1 + 20x_2 + 0,6x_2^2$, $R^2 = 0,95$.
(5,0) (12,0) (0,2)

У дужках вказано значення стандартних помилок для коефіцієнтів регресії.

Провести аналіз тісноти зв'язку результату з кожним із факторів. Обрати найкраще рівняння регресії, обґрунтувати висновок.

3.2.5. Проведено дослідження впливу x_1 – чисельності працівників, x_2 – середньорічної вартості основних засобів на обсяг виробництва продукції на 20-ти підприємствах. Отримано матрицю парних коефіцієнтів кореляції (табл. 3.2).

Таблиця 3.2

Матриця парних коефіцієнтів

Ознаки	y	x_1	x_2
y	1	0,3	0,6
x_1	0,3	1	0,8
x_2	0,6	0,8	1

Оцінити значущість рівняння регресії за допомогою F -критерію Фішера.

Побудувати економетричну модель у стандартизованих змінних. Зробити висновок щодо впливу факторів x_1 , x_2 на результативну ознаку та перевірити доцільність їхньої присутності в моделі.

Обчислити часткові коефіцієнти кореляції з кожним із факторів. Порівняти їх із парними коефіцієнтами кореляції, зробити висновок.

3.2.6. Під час вивчення рівня споживання м'яса (кг на душу населення) y залежно від доходу (грн на одного члена родини) x_1 і y відношенні з рівнем споживання риби (кг на душу населення) x_2 отримано такі результати (з 50-ти родин) (табл. 3.3). Використовуючи t -критерій Стюдента, оцінити значущість параметрів рівняння.

Таблиця 3.3

Основні результати

Рівняння регресії	$y = -180 + 0,2x_1 - 0,4x_2$		
Стандартні помилки параметрів	20	0,01	0,25
Множинний коефіцієнт кореляції	0,85		

Обчислити F -критерій Фішера.

Оцінити за частковими F -критеріями Фішера доцільність включення у модель: фактора x_1 після фактора x_2 ; фактора x_2 після фактора x_1 .

3.2.7. На 30-ти підприємствах галузі отримані результати табл. 3.4.

Таблиця 3.4

Основні результати

Коефіцієнт детермінації	?		
Множинний коефіцієнт кореляції	0,85		
Рівняння регресії	$y = ? + 0,48x_1 + 20x_2$		
Стандартні помилки параметрів	2	0,06	?
t -критерій параметрів	1,5	?	4

Заповнити пропущенні характеристики та проаналізувати результати регресійного аналізу.

Практичне заняття 4

Проблеми в побудові лінійних множинних регресійних моделей

Однією з умов використання методу найменших квадратів (МНК) для знаходження параметрів економетричної моделі є те, що пояснювальні змінні у матриці X мають бути незалежними між собою. Проте на практиці можуть мати місце випадки, коли пояснювальні змінні пов'язані між собою, що стає перешкодою до використання МНК.

Явище існування тісної лінійної залежності, або сильної кореляції, між двома або більше пояснювальними змінними називають **мультиколінеарністю**. Вона негативно впливає на кількісні характеристики економетричної моделі або взагалі робить неможливою її побудову.

Основні наслідки мультиколінеарності такі:

- падає точність оцінювання параметрів моделі;
- оцінки деяких параметрів моделі можуть показати порушення гіпотези про значущість зв'язку через наявність мультиколінеарності пояснювальних змінних;
- оцінки параметрів моделі стають дуже чутливими до розмірів сукупності спостережень і навіть збільшення цієї сукупності іноді може призвести до значних змін в оцінках параметрів.

Головними ознаками мультиколінеарності є такі:

- наявність високих значень парних коефіцієнтів кореляції $r_{x_i x_j} \geq 0,8$

(це означає, що пояснювальні змінні x_i та x_j пов'язані між собою лінійною залежністю та $r_{x_i x_j} \rightarrow 1, i \neq j$);

- значне наближення коефіцієнта кореляції до одиниці;
- наявність малих значень оцінки параметрів моделі за умови високого рівня коефіцієнта детермінації R^2 і F -критерію;
- істотна зміна оцінок параметрів моделі за умови додаткового введення до неї пояснювальної змінної.

Однією з чотирьох необхідних умов для застосування МНК у процесі оцінювання параметрів економетричної моделі є вимоги постійної

дисперсії залишків для кожного спостереження. Цю властивість незмінної дисперсії в спостереженнях називають **гомоскедастичністю**.

Явище, коли дисперсія залишків змінюється для кожного спостереження або групи спостережень, називається **гетероскедастичністю**.

За наявності гетероскедастичності залишків оцінки параметрів моделі МНК будуть незміщеними, обґрунтованими, але неефективними. При цьому формулу для обчислення стандартної помилки спостережень застосовувати не можна.

Використання того чи іншого методу перевірки економетричної моделі на наявність гетероскедастичності залежить від вхідних даних. Найбільш поширеними методами перевірки наявності чи відсутності гетероскедастичності є тест Спірмета та тест Гольфельда – Квандта.

Автокореляція – це кореляція між значеннями результуючої змінної, яка виникає унаслідок залежності значень випадкової величини в різних спостереженнях.

Основні наслідки автокореляції

- оцінки параметрів перестають бути ефективними, проте вони залишаються лінійними;
- дисперсії оцінок параметрів є зміщеними;
- висновки із t -статистики Стюдента та F -статистики Фішера про значущість коефіцієнтів регресії і коефіцієнта детермінації можуть бути неправильними.

Перевірку моделі на присутність автокореляції здійснюють за тестом Дарбіна – Уотсона.

Для n – кількість спостережень, m – кількість факторів моделі, рівень значущості $\alpha = 0,05$ за таблицями розподілу Дарбіна – Уотсона (додаток Г) знаходимо числа d_l і d_u .

Статистику Дарбіна – Уотсона обчислюють за формулою:

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}.$$

Якщо $DW < d_l$, то наявна додатна автокореляція залишків.

Якщо $DW > 4 - d_l$, то наявна від'ємна автокореляція залишків.

Якщо $d_u < DW < 4 - d_u$, то гіпотезу про відсутність автокореляції залишків приймають.

Якщо $d_l < DW < d_u$ або $4 - d_u < DW < 4 - d_l$, то значення DW потрапляє в зону невизначеності й не можна зробити висновок ні про наявність, ні про відсутність автокореляції.

4.1. Приклади розв'язання завдань

Приклад 4.1. Перевірити наявність мультиколінеарності у двофакторній моделі $\hat{y}_x = 4,403 + 0,149x_1 + 0,057x_2$ (табл. 4.1).

Таблиця 4.1

Дані для перевірки множинної регресійної моделі

y	x_1	x_2
6,0	10,0	1,4
6,3	11,2	5,7
6,5	12,0	6,5
6,8	14,5	5,1
6,8	13,3	6,0
6,8	13,2	5,2
6,8	14,5	6,0
7,0	15,3	7,5
7,2	14,5	7,8

$$\sigma_y = 0,345, \sigma_{x_1} = 1,672, \sigma_{x_2} = 1,75,$$

$$\mu_{yx_1} = 0,536, \mu_{yx_2} = 0,488, \mu_{x_1x_2} = 2,105.$$

Розв'язання

Для перевірки наявності мультиколінеарності знайдемо парні коефіцієнти кореляції за формулами:

$$r_{x_1y} = \frac{\mu_{x_1y}}{\sigma_{x_1} \sigma_y} = b_1 \frac{\sigma_{x_1}}{\sigma_y};$$

$$r_{x_2y} = \frac{\mu_{x_2y}}{\sigma_{x_2} \sigma_y} = b_2 \frac{\sigma_{x_2}}{\sigma_y};$$

$$r_{x_1x_2} = \frac{\mu_{x_1x_2}}{\sigma_{x_1} \sigma_{x_2}}.$$

Результати наведено в кореляційній табл. 4.2.

Таблиця 4.2

Лінійні коефіцієнти парної кореляції

Ознаки	y	x_1	x_2
y	1,000		
x_1	0,931	1,000	
x_2	0,808	0,719	1,000

Варто перевірити значущість міжфакторного коефіцієнта кореляції за критерієм Стюдента:

$$m_r = \sqrt{\frac{1 - r_{x_1x_2}^2}{n - 2}} = \sqrt{\frac{1 - 0,719^2}{7}} = 0,263;$$

$$t_r = \frac{0,719}{0,263} = 2,733.$$

Треба порівняти фактичні значення критерію Стюдента та його табличне значення на 5 % рівні значущості $t_{0,05}(7) = 2,45$. Міжфакторний коефіцієнт кореляції є статистично значущим, оскільки $t_{табл.} < t_{факт.}$.

Ураховуючи, що міжфакторний коефіцієнт кореляції не перевищує 0,8, r_{yx_1} , r_{yx_2} , а також є статистично значущим, то можна зробити висновки про відсутність мультиколінеарності.

Приклад 4.2. Знайти критерій Дарбіна – Уотсона та перевірити наявність автокореляції для моделі прикладу 4.1. Чи можна використовувати регресійну модель для прогнозу та чому?

Розв'язання

Знайти теоретичне значення результативної ознаки, використовуючи рівняння регресії, і перевірити за критерієм Дарбіна – Уотсона, чи існує автокореляція залишків цієї моделі.

За результатами обчислень (табл. 4.3) знаходять емпіричне значення критерію Дарбіна – Уотсона.

Таблиця 4.3

Результати обчислень залишків моделі

y	$\hat{y}$	e_i	e_{i-1}	e_i^2	$(e_i - e_{i-1})^2$
6,000	5,973	0,027	–	0,001	–
6,300	6,396	–0,096	0,027	0,009	0,015
6,500	6,561	–0,061	–0,096	0,004	0,001
6,800	6,854	–0,054	–0,061	0,003	0,000
6,800	6,726	0,074	–0,054	0,005	0,016
6,800	6,666	0,134	0,074	0,018	0,004
6,800	6,905	–0,105	0,134	0,011	0,057
7,000	7,110	–0,110	–0,105	0,012	0,000
7,200	7,007	0,193	–0,110	0,037	0,091
Σ	–	–	–	0,100	0,185

$$DW = \frac{\sum (e_i - e_{i-1})^2}{\sum e_i^2} = \frac{0,185}{0,100} = 1,85.$$

Табличні значення нижньої та верхньої меж статистичного критерію Дарбіна – Уотсона для рівня значущості $\alpha = 0,05$, кількості спостережень $n = 9$ та кількості пояснювальних змінних $m = 2$: $d_L = 0,63$, $d_U = 1,70$ (див. додаток Г). Щодо емпіричного значення критерію Дарбіна – Уотсона виконується подвійна нерівність: $d_U \leq DW \leq 4 - d_U$.

Отже, можна зробити висновок про відсутність автокореляції залишків цієї регресійної моделі та її можна використовувати для знаходження прогнозного значення результативної ознаки.

Коефіцієнт автокореляції дорівнює $r_a = 1 - \frac{DW}{2} = 0,075$, що підтверджує відсутність автокореляції через слабкий зв'язок залишків регресійної моделі.

Приклад 4.3. За даними за 18 місяців побудовано рівняння регресії залежності прибутку підприємства (млн грн) y від цін на ресурси (тис. грн за 1 т) x_1 та продуктивність праці (од. продукції на 1 робітника) x_2 . Аналізуючи залишки, було отримано такі дані:

$$\sum e_t^2 = 500, \quad \sum (e_t - e_{t-1})^2 = 1900.$$

Розрахувати критерій Дарбіна – Уотсона, оцінити отриманий результат за 5 %-м рівнем значущості. Чи можна використовувати регресійну модель для прогнозу?

Розв'язання

Критерій Дарбіна – Уотсона знайти за формулою:

$$DW = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2} = \frac{1900}{500} = 3,8.$$

Фактичне значення порівняти з табличними значеннями на 5 %-му рівні значущості. За $n = 18$ і $m = 2$ (число факторів) нижня границя дорівнює 1,05, а верхня – 1,54. Оскільки $DW > 4 - d_l$, то наявна від'ємна автокореляція залишків. Такий самий результат буде отримано, знайшовши коефіцієнт автокореляції: $r_a = 1 - \frac{DW}{2} = -0,9$.

Рівняння регресії не може бути використано для прогнозу, оскільки в ньому не усунуто автокореляцію у залишках, яка може мати різні причини. Автокореляція у залишках може означати, що в рівняння не включений який-небудь істотний фактор. Можливо також, що форма зв'язку не точна.

Приклад 4.4. Застосувати тест рангової кореляції Спірмена для оцінювання гетероскедастичності, ранжуючи за фактором x_1 для моделі прикладу 4.1.

Розв'язання

Для перевірки наявності в моделі гетероскедастичності знаходять емпіричне значення коефіцієнта рангової кореляції Спірмена. Для цього слід проранжувати значення фактора x_1 , значення залишків та різницю рангів. Обчислення наведено в табл. 4.4.

Таблиця 4.4

Результати обчислень коефіцієнта рангової кореляції

y	x_1	x_2	$\hat{y}$	e_i	R_{x_1}	$R_{ e_i }$	$d = R_x - R_{ e_i }$	d^2
6,00	10,00	1,40	5,97	0,03	1	1	0	0
6,30	11,20	5,70	6,40	-0,10	2	5	-3	9
6,50	12,00	6,50	6,56	-0,06	3	3	0	0
6,80	14,50	5,10	6,85	-0,05	7	2	5	25
6,80	13,30	6,00	6,73	0,07	5	4	1	1
6,80	13,20	5,20	6,67	0,13	4	8	-4	16
6,80	14,50	6,00	6,91	-0,11	7	6,5	0,5	0,25
7,00	15,30	7,50	7,11	-0,11	9	6,5	2,5	6,25
7,20	14,50	7,80	7,01	0,19	7	9	-2	4
Σ	-	-	-	-	-	-	-	61,5

За результатами обчислень (див. табл. 4.4) знаходять емпіричне значення коефіцієнта рангової кореляції Спірмена:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 61,5}{9(9^2 - 1)} = 0,488.$$

Знаходять t -критерій для рангової кореляції:

$$t_\rho = \frac{\rho \sqrt{n-2}}{\sqrt{1-\rho^2}} = \frac{0,488 \sqrt{7}}{\sqrt{1-0,488^2}} = 1,48.$$

Порівняти це значення t_ρ з табличним на 5 %-му рівні значущості $t_{0,05}(7) = 2,36$ (див. додаток В). Оскільки $t_\rho < t_{табл.}$, варто прийняти гіпотезу про відсутність гетероскедастичності.

Приклад 4.5. За 20 спостереженнями було побудовано модель залежності витрат на харчування y від прибутку x : $\hat{y} = 20,84 + 0,44x$; $r_{yx}^2 = 0,916$; $e_i = y_i - \hat{y}_i$ (залишки). Розміри залишків за кожного значення x виявилися такими (табл. 4.5).

Таблиця 4.5

Вихідні дані

№ з/п	x	e_i
1	30	-12
2	36	-11,7
3	40	-5,4
4	45	-5,6
5	50	-2,8
6	60	0,8
7	70	-1,6
8	80	-4
9	85	-6,2
10	90	6,6
11	92	13,7
12	100	12,2
13	120	4,4
14	130	4
15	145	3,4
16	150	23,2
17	200	16,2
18	250	-16,8
19	300	-27,8
20	360	9,8

Перевірити наявність гетероскедастичності за критерієм Спірмена та критерієм Гольдфельда – Квандта. Порівняти отримані результати.

Розв'язання

Проранжувати значення x_i та абсолютні величини e_i у порядку зростання і знайти коефіцієнт рангової кореляції Спірмена за формулою:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)},$$

де d – різниця рангів ($R_x - R_{|e|}$).

Розрахунки $\sum d^2$ у табл. 4.6.

Таблиця 4.6

Розрахунок рангової кореляції

x	e_i	Розрахунок рангової кореляції			
		R_x	$R_{ e }$	$d = R_x - R_{ e }$	d^2
30	-12	1	14	-13	169
36	-11,7	2	13	-11	121
40	-5,4	3	8	-5	25
45	-5,6	4	9	-5	25
50	-2,8	5	3	2	4
60	0,8	6	1	5	25
70	-1,6	7	2	5	25
80	-4	8	5,5	2,5	6,25
85	-6,2	9	10	-1	1
90	6,6	10	11	-1	1
92	13,7	11	16	-5	25
100	12,2	12	15	-3	9
120	4,4	13	7	6	36
130	4	14	5,5	8,5	72,25
145	3,4	15	4	11	121
150	23,2	16	19	-3	9
200	16,2	17	17	0	0
250	-16,8	18	18	0	0
300	-27,8	19	20	-1	1
360	9,8	20	12	-10	100

$$\rho = 1 - \frac{6 \cdot 739,5}{20(20^2 - 1)} = 0,444.$$

Знайти t -критерій для рангової кореляції:

$$t_\rho = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}} = \frac{0,444\sqrt{18}}{\sqrt{1-0,444^2}} = 2,102.$$

Порівняємо це значення t_ρ з табличним на 5 %-му рівні значущості та числа ступенів свободи 18: $_{0,05}t_\rho = 2,101$. Оскільки $t_\rho < t_{табл}$, на рівні 5 % приймають гіпотезу про відсутність гетероскедастичності. У цьому випадку такий висновок зробити не можна. Можна відхилити цю гіпотезу тільки за умови $\alpha = 0,10$.

Для застосування тесту Гольдфеля – Квандта потрібна інформація про y . Хоча ці дані в умові не наведено, але їх можна знайти. На підставі рівняння регресії $\hat{y} = 20,84 + 0,44x$ знайти розрахункові значення $\hat{y}$. Далі знайти фактичні значення y : $y = \hat{y} + e$ (табл. 4.7).

Таблиця 4.7

Розрахунок залишків за групами

№ з/п	x	$\hat{y}$	e	y	$\hat{y}_{по\ гр}$	$e_{по\ гр}$	$e^2_{по\ гр}$
1	2	3	4	5	6	7	8
1	30	34,04	-12,00	22	24,69	-2,65	7,04
2	36	36,68	-11,70	25	28,43	-3,45	11,90
3	40	38,44	-5,40	33	30,95	2,10	4,39
4	45	40,64	-5,60	35	34,09	0,95	0,90
5	50	42,84	-2,80	40	37,24	2,81	7,87
6	60	47,24	0,80	48	43,53	4,51	20,39
7	70	51,64	-1,60	50	49,82	0,23	0,05
8	80	56,04	-4,00	52	56,11	-4,07	16,52
							$\Sigma 69,05$
9	85	58,24	-6,20	52	—	—	—
10	90	60,44	6,60	67	—	—	—
11	92	61,32	13,70	75	—	—	—
12	100	64,84	12,20	77	—	—	—

1	2	3	4	5	6	7	8
13	120	73,64	4,40	78	82,25	-4,21	17,69
14	130	78,04	4,00	82	85,89	-3,85	14,79
15	145	84,64	3,40	88	91,35	-3,31	10,93
16	150	86,84	23,20	110	93,17	16,87	284,73
17	200	108,84	16,20	125	111,37	13,67	186,98
18	250	130,84	-16,80	114	129,57	-15,53	241,06
19	300	152,84	-27,80	125	147,77	-22,73	516,47
20	360	179,24	9,80	189	169,61	19,43	377,68
							Σ 1650,33

Виключивши чотири центральних спостереження (9 – 12), сукупність буде розділена на дві частини по 8 спостережень у кожній. За кожною частиною знайти рівняння регресії, для першої групи воно складе:

$$\hat{y} = 5,785 + 0,629x; \quad R^2 = 0,924; \quad F = 73,2.$$

Для другої групи:

$$\hat{y} = 38,566 + 0,364x; \quad R^2 = 0,814; \quad F = 26,3.$$

Для кожної групи визначити теоретичні значення $\hat{y}$, залишки e та їхні квадрати e^2 (див. табл. 4.7). Знайти співвідношення:

$$R^* = \frac{\sum e^2_{1\text{групи}}}{\sum e^2_{2\text{групи}}} = \frac{1650,33}{69,05} = 23,9.$$

Порівнявши цю величину з табличним значенням F -критерію за 5 %-м рівнем значущості і числом ступенів свободи 6 для кожної залишкової суми квадратів: $F(0,05; 6; 6) = 4,28$. Оскільки $R^* > F_{\text{табл}}$, то можна зробити висновок про наявність гетероскедастичності залишків.

Порівнявши результати двох критеріїв, можна бачити, що вони однакові. Отже, в моделі присутня гетероскедастичність.

4.2. Завдання для самостійної роботи

4.2.1. Проведено дослідження впливу обсягів виробництва (x_1) та витрат на сировину (x_2) на прибуток 30 виробничих підприємств. Отримано матрицю парних коефіцієнтів кореляції (табл. 4.8).

Таблиця 4.8

Матриця парних коефіцієнтів кореляції

Ознаки	y	x_1	x_2
y	1	0,75	-0,55
x_1	0,75	1	0,8
x_2	-0,55	0,8	1

За допомогою t -критерію Стюдента визначити наявність чи відсутність мультиколінеарності.

4.2.2. Проводять аналіз середньої заробітної плати робітників (з/п) та стажу. Статистичні дані наведено в табл. 4.9.

Таблиця 4.9

Вихідні дані

Стаж	5	10	15	20	25	30	35	40
З/п	10 000	12 500	14 300	18 700	25 400	29 000	32 000	34 300

Побудувати рівняння регресії, в якій заробітна плата є залежною змінною, а стаж праці – пояснювальною змінною; оцінити якість побудованої регресії. Чи є підстави вважати, що для заданої регресійної моделі імовірна гетероскедастичність? Якщо так, то чому?

4.2.3. Для вивчення змін *ВНП* за 10 років необхідно побудувати регресійну модель з пояснювальними змінними – споживанням (C) та інвестиціями (I) (табл. 4.10).

Оцінити коефіцієнти рівняння лінійної регресійної моделі $BHP = \beta_0 + \beta_1 I + \beta_2 C + \varepsilon$. Визначити значення статистики DW – Дарбіна – Уотсона. Чи наявна автокореляція залишків?

Таблиця 4.10

Вихідні дані

C , млрд	8	9,5	11	12	13	14	15	16,5	17	18
I , млрд	1,65	1,8	2,0	2,1	2,2	2,4	2,65	2,85	3,2	3,55
BHP , млрд	14	16	18	20	23	23,5	25	26,5	28,5	30,5

4.2.4. Проводять аналіз витрат за період 35-ти років, побудовано таке рівняння регресії:

$$y_t = 43,5 + 0,251x_t + 0,545y_{t-1}, \quad DW = 1,9,$$

$$(S) \quad (0,105) \quad (0,135)$$

де y_t – витрати на розваги молоді в момент часу t ;

x_t – прибуток у момент часу t .

Перевірити гіпотезу про відсутність автокореляції залишків.

4.2.5. Відбувається аналіз середніх річних витрат (Y) здобувачів вищої освіти на розваги. За статистичними даними за 32 роки побудовано таке рівняння регресії:

$$y_t = 41,2 + 0,254x_t + 0,539y_{t-1}; \quad R^2 = 0,783, \quad DW = 1,86,$$

$$(S) \quad (0,107) \quad (0,133)$$

де X – прибуток здобувача вищої освіти після виплати за навчання і гуртожиток.

Надати оцінку якості побудованої моделі. Перевірити гіпотезу про відсутність автокореляції залишків за умови альтернативної гіпотези про додатну автокореляцію з рівнем значущості $\alpha = 0,01$.

4.2.6. Проводять аналіз залежності між інфляцією (INF) та безробіттям (U). Використовують статистичні дані за 12 років (табл. 4.11).

Таблиця 4.11

Вихідні дані

INF	3,07	0,70	4,08	2,20	2,38	0,90	1,10	5,12	0,93	2,54	1,55	3,45
U	3,69	9,10	3,92	6,50	4,63	8,50	9,55	3,71	5,80	3,60	6,53	4,32

Побудувати лінійну парну регресію, обчислити статистику DW Дарбіна – Уотсона та на її основі перевірити наявність автокореляції; перевірити наявність автокореляції за допомогою методу рядів.

4.2.7. Обчислення залишків моделі залежності (y) – прибутку підприємства від (x_1) – вартості на сировину й (x_2) – продуктивності праці $y = f(x_1, x_2)$ наведено в табл. 4.12.

Таблиця 4.12

Вихідні дані

t	1	2	3	4	5	6
e_t	12,0	21,2	-6,7	10,1	1,5	12,6

Обчислити критерій Дарбіна – Уотсона, оцінити отриманий результат за 5 %-м рівнем значущості. Зробити висновки щодо адекватності моделі.

4.2.8. За даними 10-річних спостережень (табл. 4.13) побудовано функцію інвестицій $i_t = \beta_0 + \beta_1 y_t + \beta_2 r_t + \varepsilon_t$, де i_t – обсяг інвестувань за рік; y_t – ВВП за рік t ; r_t – відсоткова ставка за рік t .

Таблиця 4.13

Вихідні дані

Y	8,58	10,45	8,35	10,65	9,7	12,0	13,45	14,2	14,45	13,85
R	18,12	11,05	9,0	17,0	16,25	13,8	19,95	18,74	13,8	9,55
I	11,55	13,25	10,9	10,45	15,1	17,5	17,77	16,1	10,59	10,65

Оцінити за МНК коефіцієнти рівняння регресії. Використовуючи статистику DW Дарбіна – Уотсона, оцінити наявність автокореляції залишків моделі. Спрогнозувати обсяг інвестицій на наступний рік, якщо $y_{t+1} = 37$ та $r_{t+1} = 15$.

4.2.9. У табл. 4.14 наведено дані з 10 спостережень моделі вибірки.

Таблиця 4.14

Вихідні дані

i	x_i	e_i^2
1	21,3	2,3
2	22,6	5,6
3	32,7	12,8
4	41,9	10,1
5	43,8	14,6
6	49,7	13,9
7	56,9	24
8	59,7	21,9
9	67,8	19,7
10	71,5	23,8

Перевірити гіпотезу про гомоскедастичність, використовуючи тест Голдфеля – Квандта.

Практичне заняття 5

Узагальнені схеми регресійного аналізу

Серед економічних чинників, що розглядають як пояснювальні змінні моделі, можуть бути й такі, які не можна виміряти кількісно, але вони істотно впливають на рівень зв'язку між залежною і пояснювальними змінними.

Такі зміни отримали назву **фіктивних змінних (dummy-змінних)**. Ці змінні можуть істотно розширити сферу застосування лінійних моделей, включивши їх до матриці пояснювальних змінних X . Фіктивні зміни є певним чином сконструйованими змінними, що описують якісні показники, ознаки, а також відображають змінні в таких чинниках, як ефект зрушення в часі (сезонність) або змінюваність у просторі, або ж навіть включаються як змінна, що змінює інші пояснювальні змінні, яких раніше не було в моделі.

5.1. Приклади розв'язання завдань

Приклад 5.1. Досліджується ефективність ліків y залежно від віку пацієнта x (табл. 5.1). Необхідно порівняти ефективність ліків a та b .

Таблиця 5.1

Вихідні дані та додаткові розрахунки

Ліки	y	x	D	Dx		Ліки	y	x	D	Dx
a	54	69	0	0		b	30	40	1	40
b	30	48	1	48		b	23	41	1	41
a	58	73	0	0		a	21	55	0	0
b	66	64	1	64		b	43	45	1	45
b	67	60	1	60		a	38	58	0	0
a	64	62	0	0		b	43	58	1	58
a	67	70	0	0		a	43	64	0	0
a	33	52	0	0		b	45	55	1	55
a	33	63	0	0		b	48	57	1	57
b	42	48	1	48		a	48	63	0	0
b	33	64	1	64		a	53	60	0	0
a	28	55	0	0		b	58	62	1	62

Розв'язання

Вводять фіктивну змінну D : $D = \begin{cases} 0, & \text{якщо ліки } a, \\ 1, & \text{якщо ліки } b. \end{cases}$

Може бути один з трьох варіантів:

$$\hat{y}_1 = a + bx;$$

$$\hat{y}_2 = a + b_1x + c_1D;$$

$$\hat{y}_3 = a + b_1x + c_1D + c_2Dx.$$

Знайти коефіцієнти парної та множинних регресій, а також відповідні коефіцієнти детермінації для аналізу отриманих моделей:

$$\hat{y}_1 = -14,9 + 1,03x, \quad R^2 = 0,4;$$

$$\hat{y}_2 = -37,62 + 1,33x + 10,33D, \quad R^2 = 0,5;$$

$$\hat{y}_3 = -61,67 + 1,72x + 45,68D - 0,6Dx, \quad R^2 = 0,53.$$

Аналізуючи отримані результати, можна зазначити, що найкраще описує ефективність ліків залежно від віку пацієнтів модель, яка використовує фіктивну змінну як в адитивному, так і в мультиплікативному вигляді.

Приклад 5.2. Приймаючи господарські рішення, фірма, визначаючи темп випуску продукції, часто користується "згладженими" значеннями середнього продажу. Тому до відповідної економетричної моделі необхідно включати реальне виробництво як функцію скоригованого на сезонність продажу.

Варто зауважити, що в багатьох випадках, особливо на макрорівні, не можна сформулювати достатньо чіткі вказівки щодо того, які з даних, скориговані чи не скориговані, необхідно включати в економетричну модель, але на практиці проблема специфікації саме з цього боку не є не вирішуваною.

Під час побудови економетричної моделі можна брати змінні із сезонним компонентом і для урахування цього компонента застосовувати фіктивні змінні, а можна спочатку скорегувати вихідні дані, вилучивши сезонний компонент, наприклад за допомогою ковзної середньої чи регресії на одну із матриць D або $(PD)'$.

Слід зауважити, що коли сезонні коливання дуже сильні, потрібно зосередитися на специфікації моделі і внести необхідні корективи.

Приклад 5.3. Робітник щомісяця отримує заробітну плату, яка не містить сезонних коливань, але його споживчі витрати можуть мати ці коливання, і вони досягають свого максимуму у третьому кварталі через відпускні витрати. Щоб знайти залежність між споживчими витратами та заробітною платою, необхідно скоригувати квартальні витрати.

У цьому разі скориговані квартальні витрати можна отримати як функцію від доходу (зарплати) і подати реальні витрати як суму скоригованих витрат і сезонної складової.

Економетрична модель тоді буде моделлю реальних витрат від доходу та необхідної кількості сезонних фіктивних змінних. Але оскільки і сам дохід часто містить сезонні коливання, які не завжди можуть збігатися із сезонними коливаннями в споживчих витратах, то в цьому випадку модель реального споживання від доходу має містити споживчі витрати з вилученою сезонністю. До цієї моделі будуть включені необхідні фіктивні зміни, що відображають сезонні коливання залежної змінної.

5.2. Завдання для самостійної роботи

5.2.1. Нехай наявна модель:

$$y_i = a_1 + a_2 D_{2i} + a_3 D_{3i} + a_4 (D_{2i} D_{3i}) + \beta x_i + \varepsilon_i,$$

де y_i – щорічна зарплата викладача ЗВО;

$D_{2i} = 1$ – якщо це чоловік, $i = 0$ у цьому випадку;

$D_{3i} = 1$ – якщо це жінка, $i = 0$ у цьому випадку;

x_i – роки викладацького досвіду.

Що означає вираз $(D_{2i} D_{3i})$?

Знайти $E(y_i / D_2 = 1, D_3 = 1, x_i)$ і проінтерпретувати його.

5.2.2. Американські вчені Шіско і Росткер зацікавилися факторами, які визначають зарплати заробітчани. Базуючись на вибірці щодо 318 заробітчани, вони отримали таку регресію:

$$\begin{matrix} 90,06D_1 + 75,51D_2 + 47,33D_3 + 113,64D_4 + 2,26x_1 ; \\ (24,47) \quad (21,60) \quad (23,42) \quad (27,62) \quad (0,94) \end{matrix}$$

$$\varpi_m = 37,07 + 0,403\varpi_0; \quad R^2 = 0,34, \quad df = 311,$$

де D_1 – раса = 0, якщо білий, = 1, якщо не білий;

D_2 – місто = 0, якщо не міський, = 1, якщо міський;

D_3 – освіта = 0 – без освіти, = 1 має освіту;

D_4 – область = 0, якщо не західний район, = 1, якщо західний район;

x_1 – вік людини;

ϖ_m – зарплата заробітчанина (центів на годину);

ϖ_0 – початкова зарплата (центів на годину).

З цього рівняння отримати рівняння погодинної оплати праці для таких типів заробітчанин:

- а) білий, неміський, із західних районів, із вищою освітою;
- б) темношкірий, міський, не із західних районів, без вищої освіти;
- в) білий, неміський, не із західних районів, із вищою освітою.

5.2.3. Досліджують надійність верстатів трьох виробників a , b та c . Ураховують вік верстатів M (місяці) та час H (години) безаварійної праці до останньої несправності. Вибірка з 40 верстатів дала такі результати (табл. 5.2).

Таблиця 5.2

Вихідні дані

Фірма	H	M	Фірма	H	M
1	2	3	4	5	6
a	280	23	a	200	52
b	230	30	b	265	20
c	112	65	c	148	70
a	176	69	c	150	62
c	90	75	b	176	40
a	176	63	a	123	66
b	216	25	a	245	20
c	110	75	c	176	39
b	45	75	b	260	25
a	236	48	b	220	22
a	205	59	c	194	33

1	2	3		4	5	6
<i>a</i>	240	25		<i>a</i>	156	48
<i>b</i>	65	69		<i>a</i>	100	75
<i>a</i>	115	71		<i>b</i>	240	21
<i>c</i>	200	26		<i>a</i>	170	56
<i>b</i>	126	45		<i>c</i>	116	58
<i>a</i>	225	40		<i>b</i>	120	40
<i>c</i>	210	30		<i>a</i>	240	37
<i>b</i>	45	69		<i>b</i>	88	56
<i>a</i>	260	30		<i>a</i>	120	67

Побудувати рівняння регресії, не враховуючи різниці верстатів різних фірм виробників $H = a + b_1M$.

Ввести фіктивні змінні та побудувати множинне рівняння регресії. Урахувати фіктивну змінну виробника верстатів в адитивному вигляді $H = a + b_1M + b_2F + b_3R$, якщо:

$$F = \begin{cases} 0, & \text{якщо виробник } a, \\ 1, & \text{якщо виробник } b \text{ або } c. \end{cases} \quad R = \begin{cases} 0, & \text{якщо виробник } a \text{ або } b, \\ 1, & \text{якщо виробник } c. \end{cases}$$

Побудувати множинне рівняння регресії, ураховуючи фіктивну змінну F в мультиплікативному вигляді $H = a + b_1M + b_2F + b_3FM$.

5.2.4. Якщо є щомісячні дані за кілька років, скільки дитму-змінних необхідно ввести для тестування гіпотез:

- усі 12 місяців року мають вплив сезонних коливань;
- тільки лютий, квітень, червень, серпень, жовтень і грудень мають вплив сезонних коливань.

5.2.5. Logit-модель була застосована до вибірки, у якій $y = 1$, якщо кількість зайнятих у фірмі виросло ($y = 0$ – в іншому випадку), x_1 – дохід фірми, у млн дол. США; $x_2 = 1$, Якщо фірма належить до сфери високих технологій ($x_2 = 0$ – в іншому випадку). Отримано таку модель: $\hat{z} = 0,40 + 0,20 \cdot x_1 + 0,10 \cdot x_2$. Потрібно визначити оцінку ймовірності

зростання зайнятості для фірми високотехнологічної фірми *A* з доходом 5 млн дол. США і для фірми *B*, яка не належить до сфери високих технологій і має дохід 7 млн дол. США.

Практичне заняття 6

Системи економетричних рівнянь

6.1. Приклади розв'язання завдань

Досить часто складні економічні процеси моделюються декількома пов'язаними між собою рівняннями, що містять як повторювані, так і власні змінні. Найбільшого поширення в економетричних дослідженнях отримала система взаємозалежних або одночасних рівнянь. У ній одні і ті ж залежні змінні в одних рівняннях входять в ліву частину, а в інших рівняннях – у праву частину системи:

[illegible]

Таку систему взаємозалежних рівнянь називають також структурною формою моделі. Структурна форма моделі дозволяє побачити вплив змін будь-якої екзогенної змінної на значення ендогенної змінної. У такій системі кожне рівняння не може розглядатися самостійно, і для знаходження його параметрів використовують спеціальні прийоми оцінювання.

Найбільшого поширення набули такі методи оцінювання коефіцієнтів структурної моделі:

- 1) непрямий метод найменших квадратів;
- 2) двокроковий метод найменших квадратів;
- 3) трикроковий метод найменших квадратів;
- 4) метод максимальної правдоподібності з повною інформацією;
- 5) метод максимальної правдоподібності з обмеженою інформацією.

Непрямий метод найменших квадратів застосовують для ідентифікованої структурної моделі.

Систему рівнянь називають **ідентифікованою**, якщо за оцінками коефіцієнтів наведених рівнянь можна однозначно визначити коефіцієнти структурних рівнянь (зазвичай це вдається зробити, коли кількість рівнянь для визначення коефіцієнтів структурних рівнянь у точності дорівнює кількості цих коефіцієнтів).

Приклад 6.1.

1. Перевірити структурну модель на ідентифікацію:

$$\begin{cases} y_1 = b_{13}y_3 + a_{11}x_1 + a_{13}x_3, \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2, \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

2. Знайти структурні коефіцієнти моделі, якщо зведена форма моделі рівнянь така:

$$\begin{cases} y_1 = 2x_1 + 4x_2 + 10x_3, \\ y_2 = 3x_1 - 6x_2 + 2x_3, \\ y_3 = -5x_1 + 8x_2 + 5x_3. \end{cases}$$

Розв'язання

Модель має три ендогенних (y_1, y_2, y_3) і три екзогенних (x_1, x_2, x_3) змінних. Перевірити для кожного рівняння необхідну та достатню умови ідентифікації.

Перше рівняння.

Необхідна умова. Ендогенних змінних – дві (y_1, y_3) , відсутніх екзогенних – одна (x_2) . Виконується необхідне рівняння: $2 = 1 + 1$, отже, рівняння ідентифікується.

Достатня умова. У першому рівнянні відсутні y_2, x_2 . Побудувати матрицю з коефіцієнтів цих змінних в інших рівняннях системи:

$$A = \begin{pmatrix} -1 & a_{22} \\ b_{32} & 0 \end{pmatrix} \quad \text{Det}A = -1 \cdot 0 - b_{32} \cdot a_{22} \neq 0.$$

Визначник матриці не нульовий, ранг матриці дорівнює двом, отже, виконується достатня умова ідентифікації. Перше рівняння ідентифіковано.

Друге рівняння.

Необхідна умова. Ендогенних змінних – три (y_1, y_2, y_3) , відсутніх екзогенних – дві (x_1, x_3) . Виконується необхідне рівняння: $3 = 2 + 1$, отже, рівняння ідентифікується.

Достатня умова. У другому рівнянні відсутні x_1, x_3 . Побудувати матрицю з коефіцієнтів цих змінних в інших рівняннях системи:

$$A = \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \quad \text{Det}A = a_{11} \cdot a_{33} - a_{31} \cdot a_{13} \neq 0.$$

Визначник матриці не нульовий, ранг матриці дорівнює двом, отже, виконується достатня умова ідентифікації. Друге рівняння ідентифіковано.

Третє рівняння:

Необхідна умова. Ендогенних змінних – дві (y_2, y_3) , відсутніх екзогенних – одна (x_2) . Виконується необхідне рівняння: $2 = 1 + 1$, отже, рівняння ідентифікується.

Достатня умова. У другому рівнянні відсутні y_1, x_2 . Побудувати матрицю з коефіцієнтів цих змінних в інших рівняннях системи:

$$A = \begin{pmatrix} -1 & 0 \\ b_{21} & a_{22} \end{pmatrix} \quad \text{Det}A = -1 \cdot a_{22} - b_{21} \cdot 0 \neq 0.$$

Визначник матриці не нульовий, ранг матриці дорівнює двом, отже, виконується достатня умова ідентифікації. Третє рівняння ідентифіковано.

Таким чином, уся система ідентифікується.

Визначити структурні коефіцієнти моделі.

Виключивши з першого та третього зведених рівнянь змінну x_2 , визначено структурні коефіцієнти першого рівняння, а саме:

$$\begin{cases} y_1 = 2x_1 + 4x_2 + 10x_3, & | \cdot (-2) \\ y_3 = -5x_1 + 8x_2 + 5x_3, & \oplus \end{cases}$$

$$y_3 - 2y_1 = -9x_1 - 15x_3.$$

Отже, $y_1 = 0,5y_3 + 4,5x_1 + 7,5x_3$ – перше рівняння структурної моделі.

Щоб знайти друге рівняння, спочатку слід додати всі три рівняння, а потім по черзі виключити з рівнянь змінні x_1, x_3 :

$$\begin{cases} y_1 = 2x_1 + 4x_2 + 10x_3, \\ y_2 = 3x_1 - 6x_2 + 2x_3, \\ y_3 = -5x_1 + 8x_2 + 5x_3, \oplus \end{cases}$$

$$y_1 + y_2 + y_3 = 6x_2 + 17x_3;$$

$$\begin{cases} y_1 = 2x_1 + 4x_2 + 10x_3, \mid \cdot (-3) \\ y_2 = 3x_1 - 6x_2 + 2x_3, \mid \cdot 2 \oplus \end{cases}$$

$$-3y_1 + 2y_2 = -24x_2 - 26x_3;$$

$$\begin{cases} -3y_1 + 2y_2 = -24x_2 - 26x_3, \mid \cdot 17 \\ y_1 + y_2 + y_3 = 6x_2 + 17x_3, \mid \cdot 26 \oplus \end{cases}$$

$$-25y_1 + 60y_2 + 26y_3 = -252x_2.$$

Відокремивши y_2 , знаходять друге рівняння структурної моделі –

$$y_2 = \frac{5}{12}y_1 - \frac{13}{30}y_3 - 4\frac{3}{15}x_2.$$

Виключивши із другого та третього зведених рівнянь змінну x_2 , визначено структурні коефіцієнти третього рівняння, а саме:

$$\begin{cases} y_2 = 3x_1 - 6x_2 + 2x_3, \mid \cdot 4 \\ y_3 = -5x_1 + 8x_2 + 5x_3, \mid \cdot 3 \oplus \end{cases}$$

$$3y_3 + 4y_2 = -3x_1 + 23x_3.$$

Отже, $y_3 = -1\frac{1}{3}y_2 - x_1 + 7\frac{2}{3}x_3$ – третє рівняння структурної моделі.

Таким чином, структурна модель за заданої зведеної форми має такий вигляд:

$$\begin{cases} y_1 = 0,5y_3 + 4,5x_1 + 7,5x_3, \\ y_2 = \frac{5}{12}y_1 - \frac{13}{30}y_3 - 4\frac{3}{15}x_2, \\ y_3 = -1\frac{1}{3}y_2 - x_1 + 7\frac{2}{3}x_3. \end{cases}$$

Приклад 6.2. Побудувати економетричну модель, яка характеризує залежність між обсягом валового національного продукту від виробничих ресурсів, а саме: основних виробничих фондів, робочої сили, матеріальних ресурсів.

Розв'язання

Для побудови такої моделі доцільно будувати економетричну модель на основі системи одночасних структурних рівнянь:

$$X_t = f(F_t, L_t, u_t),$$

$$M_t = f(X_t, v_t),$$

$$Y_t = X_t - M_t,$$

де X_t – випуск продукції;

t – період часу;

F_t – основні виробничі фонди;

L_t – робоча сила;

M_t – матеріальні ресурси;

Y_t – валовий національний продукт.

Запишемо два перших рівняння у вигляді:

$$X_t = a_0 F_t^{a_1} L_t^{a_2} u_t,$$

$$M_t = b_0 + b_1 X_t + v_t,$$

$$Y_t = X_t - M_t,$$

де a_0, a_1, a_2, b_0, b_1 – параметри моделі, u_t, v_t – залишки.

Таким чином, побудована економетрична модель складається з трьох одночасних рівнянь: два перших є регресійними, а третє – тотожність. Усі ці рівняння описують економічні процеси, які відбуваються одночасно і значить вони повинні мати спільний розв'язок.

Приклад 6.3. Кейнсіанська модель формування доходів. Розглядається замкнута економіка без державних витрат.

$$\begin{cases} c_t = \beta_0 + \beta_1 y_t + \varepsilon_t, \\ y_t = c_t + i_t, \end{cases}$$

де перше рівняння – функція споживання;

друге рівняння – макроекономічне тотожність;

y_t і i_t – значення сукупного випуску та інвестицій в момент часу t ;

ε_t – випадкова складова.

6.2. Завдання для самостійної роботи

6.2.1. Перевірити структурну модель на ідентифікацію:

$$\begin{cases} y_1 = b_{13}y_3 + a_{11}x_1 + a_{13}x_3, \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2, \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

Знайти структурні коефіцієнти моделі, якщо зведена форма моделі рівнянь така:

$$\begin{cases} y_1 = 3x_1 + x_2 + 2x_3, \\ y_2 = x_1 - 2x_2 + 6x_3, \\ y_3 = -4x_1 + x_2 - 3x_3. \end{cases}$$

6.2.2. Перевірити структурну модель на ідентифікацію:

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + a_{12}x_2, \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2, \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

Знайти структурні коефіцієнти моделі, якщо зведена форма моделі рівнянь така:

$$\begin{cases} y_1 = 3x_1 - 6x_2 + 2x_3, \\ y_2 = 2x_1 + 4x_2 + 10x_3, \\ y_3 = -5x_1 + 6x_2 + 5x_3. \end{cases}$$

6.2.3. Перевірити структурну модель на ідентифікацію:

$$\begin{cases} y_1 = b_{13}y_3 + a_{11}x_1 + a_{13}x_3, \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2, \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

Знайти структурні коефіцієнти моделі, якщо зведена форма моделі рівнянь така:

$$\begin{cases} y_1 = 2x_1 - 4x_2 + 4x_3, \\ y_2 = 3x_1 + 2x_2 - 5x_3, \\ y_3 = -5x_1 + 2x_2 + 3x_3. \end{cases}$$

6.2.4. Перевірити структурну модель на ідентифікацію:

$$\begin{cases} y_1 = b_{12}y_2 + a_{11}x_1 + a_{12}x_2, \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2, \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

Знайти структурні коефіцієнти моделі, якщо зведена форма моделі рівнянь така:

$$\begin{cases} y_1 = 2x_1 - 4x_2 + 10x_3, \\ y_2 = x_1 + 3x_2 + 2x_3, \\ y_3 = -6x_1 + 7x_2 + 6x_3. \end{cases}$$

6.2.5. Наведені результати розрахунків деякої економетричної моделі. Структурна форма моделі:

$$\begin{cases} y_1 = -4 + ? y_2 - 9,4x_2 + e_1, \\ y_2 = 12,83 - 2,67 y_1 + ? x_1 + e_2, \\ y_3 = 1,36 - 1,76 y_1 + 0,828 y_2 + e_3. \end{cases}$$

Зведена форма моделі така:

$$\begin{cases} y_1 = 2 + 4x_1 - 3x_2 + v_1, \\ y_2 = 7,5 + 5x_1 + 8x_2 + v_2, \\ y_3 = 4 + ? x_1 + ? x_2 + v_3. \end{cases}$$

Знайдіть невідомі характеристики. Якими методами отримані параметри структурної та зведеної форм моделі? Поясніть можливість використання МНК для розрахунку структурних параметрів моделі.

6.2.6. Є дані за п'ять років (табл. 6.1).

Таблиця 6.1

Вихідні дані

Номер року	Річне споживання свинини на душу населення, фунтів, y_1	Оптова ціна за фунт, дол. США, y_2	Дохід на душу населення, дол. США, x_1	Затрати на обробку м'яса, % до ціни, x_2
1	2	3	4	5
1	60	5,0	1 300	60
2	62	4,0	1 300	56

1	2	3	4	5
3	65	4,2	1 500	56
4	62	5,0	1 600	63
5	66	3,8	1 800	50

Побудувати модель виду:

$$\begin{cases} y_1 = f(y_2, x_1), \\ y_2 = f(y_1, x_2). \end{cases}$$

Розрахувати відповідні структурні коефіцієнти.

Практичне заняття 7

Вивчення взаємозв'язків за часовими рядами

7.1. Приклади розв'язання завдань

Часовим рядом називають упорядковану в часі послідовність показників, що характеризує рівні розвитку досліджуваного явища в послідовні моменти або періоди часу. Окремі спостереження часового ряду називають рівнями цього ряду.

Часові ряди бувають моментні, інтервальні і похідні.

Моментні ряди характеризують значення показника на певні моменти часу.

Інтервальні ряди характеризують значення показника за деякі інтервали часу.

Похідні ряди виходять з середніх або відносних величин показника.

Важливою умовою правильного відображення часового ряду реального процесу розвитку є порівнянність рівнів часового ряду. Несумісність найчастіше зустрічають у вартісних характеристиках, змінах цін, територіальних змінах, під час укрупнення підприємств. Для несумісних величин показника неправомірно проводити прогнозування.

Прийнято вважати, що значення рівнів часових рядів економічних показників складаються з таких компонент: тренда (T), сезонної (S), циклічної (C) і випадкової (E).

Кореляційний залежність між послідовними рівнями тимчасового ряду називають автокореляцією рівнів ряду. Кількісно її можна виміряти за допомогою лінійного коефіцієнта кореляції між рівнями вихідного часового ряду і рівнями цього ряду, зсунутими на кілька кроків у часі.

Приклад 7.1.

1. Визначити абсолютні прирости та зробити висновки щодо тенденції розвитку кожного ряду. Указати основні шляхи усунення тенденції (табл. 7.1).

Таблиця 7.1

Початкові дані динамічних рядів

Роки \ Показники	y	x
2010	30	100
2011	35	103
2012	39	105
2013	44	109
2014	50	115
2015	53	118

2. Побудувати лінійну модель, використовуючи перші різниці динамічних рядів. Пояснити економічний зміст коефіцієнта регресії.

3. Побудувати лінійну модель, включивши до неї фактор часу. Зробити висновки.

Розв'язання

1. Щорічні прирости визначають за формулами:

$$\Delta y_t = y_t - y_{t-1}, \quad \Delta x_t = x_t - x_{t-1}.$$

Обчислення наведено в табл. 7.2.

Таблиця 7.2

Щорічні прирости

y_t	Δy_t	x_t	Δx_t
30	–	100	–
35	5	103	3
39	4	105	2
44	5	109	4
50	6	115	6
53	3	118	3

Значення Δy не мають чітко вираженої тенденції, вони знаходяться біля середнього рівня, що означає наявність лінійного тренда (лінійної тенденції). Такий саме висновок можна зробити й з ряду x : абсолютні прирости не мають систематичної спрямованості, вони приблизно стабільні, тобто характеризуються лінійною тенденцією.

2. Для побудови регресійної моделі необхідно усунути тенденцію, оскільки ряди динаміки мають спільну тенденцію до зростання. Для цього можна будувати модель за першими різницями, тобто $\Delta y = f(\Delta x)$, якщо ряди динаміки характеризуються лінійною тенденцією. Інший спосіб урахувати тенденцію під час побудови моделі – це знайти з кожного ряду рівняння тренда:

$$\Delta y = f(\Delta x) \text{ та } \Delta x = f(\Delta x),$$

а також відхилення від нього:

$$\Delta y_t = y_t - y_{t-1}, \quad \Delta x_t = x_t - x_{t-1}.$$

Далі модель будують за відхиленнями від тренда: $dy = f(dx)$.

Під час побудови економетричних моделей частіше використовують інший шлях урахування тенденції – включення до моделі фактора часу. Інакше кажучи, модель будують за початковими даними, але до неї долучається як самостійний фактор час, тобто $\hat{y}_t = f(x, t)$.

Модель з використанням перших різниць має такий вигляд:

$$\Delta \hat{y} = a + b \cdot \Delta x.$$

Для визначення параметрів моделі слід використати метод найменших квадратів.

Система нормальних рівнянь така:

$$\begin{cases} \sum \Delta y = n \cdot a + b \cdot \sum \Delta x, \\ \sum \Delta y \Delta x = a \cdot \sum \Delta x + b \cdot \sum \Delta^2 x. \end{cases}$$

Необхідні обчислення наведено в табл. 7.3.

Таблиця 7.3

Вихідні дані для визначення параметрів моделі

y	x	xy	x^2
5	3	15	9
4	2	8	4
5	4	20	16
6	6	36	36
3	3	9	9
23	18	88	74

Наявні такі дані:

$$\begin{cases} 23 = 5 \cdot a + 18 \cdot b, \\ 88 = 18 \cdot a + 74 \cdot b. \end{cases}$$

Розв'язавши цю систему, знаходять: $a = 2,57$ та $b = 0,57$. Модель має такий вигляд:

$$\Delta \hat{y} = 2,57 + 0,57 \cdot \Delta x.$$

Коефіцієнт регресії $b = 0,57$ означає, що зі зростанням факторної ознаки x , результативна ознака y зростає із середнім прискоренням 0,57 одиниць.

3. Якщо до моделі включити фактор часу, то вона буде мати такий вигляд:

$$\hat{y} = a + b \cdot x + c \cdot t.$$

Використавши МНК, визначають систему нормальних рівнянь:

$$\begin{cases} \sum y = n \cdot a + b \cdot \sum x + c \cdot \sum t, \\ \sum yx = a \cdot \sum x + b \cdot \sum x^2 + c \cdot \sum xt, \\ \sum yt = a \cdot \sum t + b \cdot \sum xt + c \cdot \sum t^2. \end{cases}$$

За результатами обчислень, наведених у табл. 7.4, система рівнянь має такий вигляд:

$$\begin{cases} 251 = 6 \cdot a + 650 \cdot b + 21 \cdot c, \\ 27\,500 = 650 \cdot a + 70\,664 \cdot b + 2\,340 \cdot c, \\ 961 = 21 \cdot a + 2\,340 \cdot b + 91 \cdot c. \end{cases}$$

Таблиця 7.4

Вихідні дані для визначення параметрів моделі

t	y	x	xy	yt	xt	x^2	t^2
1	30	100	3 000	30	100	10 000	1
2	35	103	3 605	70	206	10 609	4
3	39	105	4 095	117	315	11 025	9
4	44	109	4 796	176	436	11 881	16
5	50	115	5 750	250	575	13 225	25
6	53	118	6 254	318	708	13 924	36
21	251	650	27 500	961	2 340	70 664	91

Після її обчислення знаходять параметри моделі: $a = -5,42$, $b = 0,32$ та $c = 3,52$. Отже, рівняння регресії має такий вигляд:

$$\hat{y} = -5,42 + 0,32 \cdot x + 3,52 \cdot t.$$

Параметр $b = 0,32$ фіксує силу зв'язку y та x . Його величина означає, що зі зростанням x на одну одиницю за умови незмінної тенденції результативна ознака y зростає у середньому на 0,32 одиниці. Параметр $c = 3,52$ характеризує середньорічний абсолютний приріст ознаки y під впливом інших факторів за умови незмінної x .

7.2. Завдання для самостійної роботи

7.2.1. Визначити абсолютні прирости та зробити висновки щодо тенденції розвитку кожного ряду. Указати основні шляхи усунення тенденції (табл. 7.5).

Таблиця 7.5

Початкові дані динамічних рядів

y	x
90	120
95	117
103	132
115	136
134	155
132	163

Побудувати лінійну модель, використовуючи перші різниці динамічних рядів. Побудувати лінійну модель, включивши до неї фактор часу. Зробити висновки.

7.2.2. Визначити абсолютні прирости та зробити висновки щодо тенденції розвитку кожного ряду (табл. 7.6).

Таблиця 7.6

Початкові дані динамічних рядів

y	x
46	53
50	57
54	64
59	70
62	73
67	82

Побудувати лінійну модель, використовуючи перші різниці динамічних рядів.

Побудувати лінійну модель, включивши до неї фактор часу. Зробити висновки.

7.2.3. Будують модель залежності інвестицій (млн грн) y від прибутку компанії (млн грн) x за кварталними даними за 7 років у табл. 7.7.

Таблиця 7.7

Початкові дані

Номер року	Ознаки	Квартали			
		I	II	III	IV
1	x	9	10	9	11
	y	10	12	11	10
2	x	10	12	13	14
	y	11	13	15	16
3	x	14	17	18	18
	y	11	11	10	11
4	x	20	21	24	26
	y	15	15	22	21
5	x	24	25	26	28
	y	21	23	24	24
6	x	31	31	26	27
	y	25	26	25	23
7	x	32	33	34	33
	y	27	28	29	31

Необхідно побудувати рівняння лінійної регресії та перевірити його значущість за допомогою критерію Фішера, та значущість його параметрів за критерієм Стюдента. За допомогою критерію Дарбіна – Уотсона оцінити автокореляцію залишків.

7.2.4. У табл. 7.8 наведені дані про прибутки господарств (X) та затрати господарств на роздрібні закупки (Y) за 18 років.

Таблиця 7.8

Початкові дані

X	9,098	9,137	9,095	9,280	9,230	9,348	9,525	9,755	10,28
Y	5,49	5,54	5,305	5,505	5,42	5,32	5,54	5,69	5,87
X	11,02	12,055	12,088	12,215	12,495	11,305	11,43	11,45	11,697
Y	6,342	6,48	6,395	6,555	6,755	5,905	6,125	6,185	6,225

Оцінити параметри рівнянь регресії $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$ та $y_t = \beta_0 + \beta_1 x_t + \gamma y_{t-1} + \varepsilon_t$. Оцінити якість побудованих моделей та порівняти результати. Яка з моделей має переваги?

7.2.5. У табл. 7.9 наведено статистичні дані відсотків змін заробітної плати (Y), зростання продуктивності праці (x_1) та рівня інфляції (x_2) за 20 років.

Таблиця 7.9

Початкові дані

Y	6,0	8,9	9,0	7,1	3,2	6,5	9,1	14,6	11,9	9,4
x_1	2,8	6,3	4,5	3,1	1,5	7,6	6,7	4,2	2,7	3,5
x_2	3,0	3,1	3,8	3,8	1,1	2,3	3,6	7,5	8,0	6,3
Y	12,0	12,5	8,5	5,9	6,8	5,6	4,8	6,7	5,5	4,0
x_1	5,0	2,3	1,5	6,0	2,9	2,8	2,6	0,9	0,6	0,7
x_2	6,1	6,9	7,1	3,1	3,7	3,9	3,9	4,8	4,3	4,8

За МНК побудувати рівняння регресії $y_t = b_0 + b_1 x_{1t} + b_2 x_{2t} + e_t$ та $y_t = c_0 + c_1 x_{1t-1} + c_2 x_{2t-1} + v_t$ ($x_{10} = 3,5$; $x_{20} = 4,5$). Оцінити якість рівнянь. Порівняти побудовані моделі. Яка з них має переваги і чому?

7.2.6. Вивчають залежність обсягів продажу товару А (y_t) від динаміки споживчих цін (x_t). Отримано дані за 12 кварталів (табл. 7.10).

Відомо, що $\sum_{t=1}^{12} x_t = 1\,360$; $\sum_{t=1}^{12} y_t = 470$; $\sum_{t=1}^{12} x_t y_t = 54\,945$; $\sum_{t=1}^{12} x_t^2 = 20\,320$.

Початкові дані

Квартали	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Індекс споживчих цін, у % до I кварталу	100	102	104	110	112	112	114	117	120	124	124	130
Середній за день обсяг продажу товару А, кг	20	24	22	30	35	37	45	47	46	50	54	60

Побудувати модель залежності обсягів продажу товару А від індексу споживчих цін з включенням фактору часу ($Y = a_1x_t + a_2t$) та навести інтерпретацію параметрів моделі.

7.2.7. З метою прогнозування обсягів експорту країни на майбутні періоди було зібрано дані за 30 років таких показників: Y_t – обсяги експорту (млрд грн); x_t – індекс обсягів промислового виробництва (у % до попереднього року).

Отримано такі результати оброблення вихідних даних:
рівняння лінійних трендів:

$$Y_t = 3,1 + 1,35t + e_t; \quad R^2 = 0,91; \quad DW = 2,31;$$

$$X_t = -8,4 + 4,8t + e_t; \quad R^2 = 0,89; \quad DW = 2,08;$$

рівняння регресії за рівнями часових рядів:

$$Y_t = -10,5 + 0,5x_t + e_t; \quad R^2 = 0,95; \quad DW = 2,21;$$

рівняння регресії за першими різницями рівнів часових рядів:

$$\Delta Y_t = 1,4 + 0,03\Delta x_t + e_t; \quad R^2 = 0,86; \quad DW = 2,25;$$

рівняння регресії за другими різницями рівнів часових рядів:

$$\Delta^2 Y_t = 0,7 + 0,012\Delta^2 x_t + e_t; \quad R^2 = 0,47; \quad DW = 2,69;$$

рівняння регресії за рівнями часових рядів з включенням фактору часу:

$$Y_t = 4,23 + 0,24x_t + 0,78t + e_t; \quad R^2 = 0,97; \quad DW = 0,9.$$

Перевірити гіпотези щодо величини коефіцієнта автокореляції. Обрати найкраще рівняння регресії, яке можна використати для прогнозування обсягів експорту, дати інтерпретацію його параметрів.

Практичне заняття 8

Моделювання одновимірних часових рядів

Для моделювання тенденції розвитку процесу або явища в реальних економічних задачах найчастіше використовують моделі кривих росту. Це є функції часу $y = f(x)$, при цьому вважають, що вплив інших факторів несуттєво чи непрямо враховують через фактор часу.

Прогнозування на основі моделей кривих росту ґрунтується на екстраполяції, тобто на продовженні на наступні періоди тенденції, яка встановлена за попередніми періодами. Оскільки даний метод прогнозування передбачає побудову аналітичної функції, що характеризує залежність рівнів ряду від часу або тренда, то називають аналітичним вирівнюванням часового ряду.

У процедурі прогнозування на основі кривих росту виокремлюють такі етапи:

- 1) вибір однієї або декількох кривих, форма яких відповідає характеру зміни часового ряду;
- 2) оцінювання параметрів вибраних кривих;
- 3) перевірка адекватності вибраних кривих прогнозованого процесу чи явища, оцінювання точності моделей і остаточний вибір кривої росту;
- 4) розрахунок точкового та інтервального прогнозів.

У прогнозуванні економічних показників за допомогою кривих росту найчастіше застосовують такі функції: лінійний тренд: $y = a + bt$; гі-

пербола: $y = a + \frac{b}{t}$; експоненціальний тренд: $y = e^{a+bt}$; тренд у формі

степеневої функції: $y = a \cdot t^b$; парабола другого та високих порядків:

$y = a + b_1t + b_2t^2 + \dots + b_kt^k$. У переважній більшості випадків розрахунок оцінок параметрів моделей здійснюється за допомогою методу найменших квадратів у формі регресійних моделей, в яких як залежна змінна слугують значення показників, а фактором є час. Для нелінійних трендових моделей застосовують процедуру лінеаризації.

Для визначення типу тенденції використовують кілька способів: якісний аналіз досліджуваного процесу, графічний метод, а також методи, що використовують коефіцієнти автокореляції рівнів ряду.

Найбільш простим методом вибору кривих зростання вважають візуальний, основний на графічному зображенні часового ряду. Якщо на графічному зображенні чітко не простежується тенденція, то рекомендують провести деякі стандартні перетворення ряду (наприклад, згладжування) і підібрати відповідну функцію. За допомогою сучасних програмних засобів, наприклад, статистичних пакетів процедуру вибору кривої зростання здійснити дуже легко.

Вибір найкращого рівняння для побудови тренда можна здійснити шляхом перебору основних форм тренда, розрахунку за кожним рівнянням скоригованого коефіцієнта детермінації і вибору рівняння тренда з максимальним значенням скоригованого коефіцієнта детермінації.

Статистичну якість побудованих моделей кривих росту для прогнозу перевіряють за критеріями перевірки якості побудованих регресійних моделей: критерієм Стюдента, критерієм Фішера, критерієм Дарбіна – Уотсона. Існування автокореляції залишків може істотно спотворювати прогнозні значення.

Показники точності моделі прогнозування описують значення випадкових похибок, отриманих у процесі використання моделі. Відомо, що похибка прогнозу – це величина, що характеризує відхилення між фактичним і прогнозним значенням показника.

Абсолютну похибку прогнозу визначають за формулою: $\Delta_t = \hat{y}_t - y_t$, де $\hat{y}_t$ – прогнозне значення показника, y_t – фактичне значення показника.

Абсолютну похибку прогнозу доповнюють обчисленням відносної похибки, яку виражають у процентах відносно фактичного значення показника:

$$\delta_t = \frac{\hat{y}_t - y_t}{y_t} \cdot 100 \%$$

Прийнято вважати, що коли абсолютна і відносна похибки більші 0, то це свідчить про "завищену" прогнозну оцінку, а коли менші 0, то прогнозне значення було занижено.

Також рекомендують використовувати середні похибки за модулем:

$$|\bar{\Delta}_t| = \frac{\sum_{t=1}^n |\hat{y}_t - y_t|}{n}, \quad |\bar{\delta}_t| = \frac{1}{n} \sum_{t=1}^n \frac{\hat{y}_t - y_t}{y_t} \cdot 100 \%,$$

де n – число рівнів часового ряду, для яких визначалось прогнозне значення.

Характеристику $|\bar{\delta}_t|$ широко використовують для порівняння точності прогнозів різнорідних об'єктів прогнозування. Фахівці вважають, якщо значення $|\bar{\delta}| < 10\%$, то це свідчить про високу точність моделі, якщо $10\% < |\bar{\delta}| < 20\%$, то прогноз визнають добрим, а за умови $20\% < |\bar{\delta}| < 50\%$ – задовільним. Проте такий підхід не враховує специфіку часових рядів і характеру динаміки, призначення прогнозів та інших характеристик, що обумовлює "механічність" підходу.

Також під час порівняння моделей прогнозування використовують

середню квадратичну похибку: $s = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}$, іноді в знаменнику

замість довжини часового ряду n використовують величину $(n - k)$, де k – число коефіцієнтів, які оцінюють в моделі. У ході вибору кращого методу прогнозування використовують суми квадратів похибок, середні квадрати похибок. І чим менші значення всіх розглянутих характеристик, тим вища точність моделі прогнозу.

Фахівці за статистичними методами прогнозування вважають, що не може бути чисто формальних підходів до вибору методів і моделей прогнозування. Успішне застосування статистичних методів прогнозування на практиці можливо лише під час поєднання знань в області самих методів з глибоким знанням об'єкта дослідження, із змістовним аналізом досліджуваного явища.

8.1. Приклади розв'язання завдань

Приклад 8.1. Для економічного аналізу зовнішньоекономічної діяльності промислового підприємства важливо прогнозувати значення показника ефективності реалізації продукції на внутрішньому ринку (y). У табл. 8.1 наведені квартальні значення показника протягом п'яти років.

Квартальні значення показника

Y	1,225	1,242	1,264	1,205	1,277	1,289	1,241	1,235	1,236	1,234
t	1	2	3	4	5	6	7	8	9	10
Y	1,193	1,182	1,236	1,201	1,199	1,236	1,182	1,184	1,168	1,201
t	11	12	13	14	15	16	17	18	19	20

Побудувати модель кривої росту показника ефективності реалізації продукції промислового підприємства на внутрішньому ринку і прогнозувати його значення на два наступні квартали.

Розв'язання

Спочатку була обчислена модель кривої зростання в лінійній формі: $y = 1,2605 - 0,0037t$, при цьому статичну якість моделі було перевірено за допомогою критеріїв Стюдента, Фішера, скоригованого коефіцієнта детермінації і критерію Дарбіна – Уотсона: $t_a = 105,9$; $t_b = -3,74$; $F = 14,0$; $R^2 = 0,4062$; $DW = 1,999$.

Модель відносно якісна. На наступні квартали прогноз значення показника ефективності реалізації продукції на внутрішньому ринку такий: 1,18245; 1,17873. Можна спостерігати, що ефективність реалізації продукції на внутрішньому ринку у промислового підприємства буде трохи знижуватися в наступних двох кварталах.

У табл. 8.2 подано розраховані альтернативні моделі за основними кривими зростання за рейтингом значень коефіцієнтів детермінації.

З табл. 8.2 видно, що виходячи із значення скоригованого коефіцієнта детермінації найкращою моделлю є $y = \frac{1}{0,7931 + 0,0025 t}$. Прогноз за цією моделлю ефективності реалізації продукції на внутрішньому ринку такий: 1,18282; 1, 17935. Розбіжність прогнозних значень, розрахованих за двома моделями, досягає тисячних одиниць.

Таблиця 8.2

Альтернативні моделі

Альтернативні моделі	Скоригований коефіцієнт детермінації
$y = \frac{1}{0,7931 + 0,0025 t}$	0,4124
$y = e^{0,2317 - 0,003t}$	0,4094
$y = (1,1228 - 0,0017t)^2$	0,4078
$y = 1,2605 - 0,0037 t$	0,4062
$y = 1,248 - 0,00029 t - 0,000163 t^2$	0,3961
$y = 1,2829 - 0,0199\sqrt{t}$	0,3407
$y = 1,2672 t^{-0,0175}$	0,2348
$y = 1,2667 - 0,0213 \ln t$	0,2320
$y = \frac{1}{0,8242 - \frac{0,0275}{t}}$	0,0256
$y = e^{0,1937 + \frac{0,033}{7}}$	0,0238
$y = 1,2142 + \frac{0,0403}{t}$	0,0219

8.2. Завдання для самостійної роботи

8.2.1. Дані про рівень безробіття за 8 місяців подано в табл. 8.3. Знайти коефіцієнти автокореляції рівней цього ряду першого та другого порядку. Обґрунтуйте вибір кривої росту та інтерпретуйте отримані результати.

Таблиця 8.3

Початкові дані для аналізу

Місяці	1	2	3	4	5	6	7	8
y_t	8,8	8,6	8,4	8,1	7,9	7,6	7,4	7,0

8.2.2. Отримані дані про врожайність зернових у господарствах деякої області (табл. 8.4).

Таблиця 8.4

Початкові дані

Роки	Урожайність зернових, ц/га
1	10,2
2	10,7
3	11,7
4	13,1
5	14,9
6	17,2
7	20,0
8	23,2

Обґрунтувати вибір типу кривої росту. Знайти параметри кривої росту та спрогнозувати врожайність на наступні два роки.

8.2.3. Нехай є такий часовий ряд:

t	1	2	3	4	5	6	7	8
x_t	20	...	...	...	...	...	...	10

Відомо також $\sum x_t = 150$, $\sum x_t^2 = 8\,100$, $\sum x_t x_{t-1} = 7\,350$. Знайти коефіцієнт автокореляції рівней цього ряду першого порядку. Установити, включає чи ні досліджуваний часовий ряд тенденцію.

8.2.4. Динаміка продажів легкових авто в країні характеризується такими даними (табл. 8.5).

Таблиця 8.5

Початкові дані

Роки	2010	2011	2012	2013	2014	2015	2017	2018
Тис. шт.	8,8	8,6	8,4	8,1	7,9	7,6	7,4	7,0

Визначити коефіцієнт автокореляції першого порядку та надати його інтерпретацію. Побудувати криву росту у формі параболи другого порядку. За допомогою критерію Дарбіна – Уотсона зробіть висновки відносно автокореляції у залишках.

8.2.5. Побутові послуги населенню (млрд грн) на території України за 2008 – 2018 рр. характеризують такими кривими росту (табл. 8.6).

Таблиця 8.6

Таблиця кривих росту

Рівняння кривих росту	Коефіцієнт детермінації
$y = -5\,573,5 + 17\,300\,t$	0,963
$y = 130\,209 - \frac{138\,733}{t}$	0,519
$y = -8\,126 + 64\,684 \ln t$	0,789
$y = 19\,194,7t^{0,89847}$	0,962
$y = 22\,368,4e^{0,2189t}$	0,974
$y = 22\,368,4 \cdot 1,245^t$	0,974
$y = 22\,119 + 3\,453\,t + 1\,259\,t^2$	0,995

t приймає значення: 1, 2, ..., 10.

Вибрати найкраще рівняння. Надати інтерпретацію показникової та експоненціальної кривих. Ураховуючи, що $\sum (y_t - \hat{y}_t)^2 = 8\,977\,241\,773$ для параболи другого порядку, надати з ймовірністю 0,95 довірчий інтервал для прогнозу на 2019 рік.

8.2.6. Отримано дані, які відображають динаміку змін обсягів продажу промислової продукції за період 15-ти місяців (табл. 8.7).

Таблиця 8.7

Початкові дані

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_i	187	331	408	408	395	426	455	506	505	528	544	590	605	629	753

За умовою лінійності тренда знайти його рівняння та оцінити його значущість на рівні $\alpha = 0,05$. Установити за допомогою критерію Дарбіна – Уотсона наявність чи відсутність автокореляції залишків на 5 %-му рівні значущості.

8.2.7. Наявні поквартальні дані за роздрібним товарообігом України за 5 років (табл. 8.8).

Таблиця 8.8

Початкові дані

Номер кварталу	Товарообіг, % до минулого періоду	Номер кварталу	Товарообіг, % до минулого періоду
1	100,0	11	98,8
2	93,9	12	101,9
3	96,5	13	113,1
4	101,8	14	98,4
5	107,8	15	97,3
6	96,3	16	102,1
7	95,7	17	97,6
8	98,2	18	83,7
9	104,0	19	84,3
10	99,0	20	88,4

Побудувати графік часового ряду.

Побудувати мультиплікативну модель часового ряду.

Оцінити якість моделі через показник середньої абсолютної помилки та середнього відносного відхилення.

8.2.8. Дані з експорту та імпорту Німеччини (млн дол. США) за 12 років подано в табл. 8.9.

Статистичні дані

Роки	Експорт	Імпорт	Роки	Експорт	Імпорт
1	184	158	7	403	390
2	243	191	8	422	402
3	294	228	9	382	346
4	323	280	10	430	385
5	341	270	11	524	464
6	410	346	12	521	456

За кожним рядом побудувати тренди і вибрати кращий із них.

Побудувати рівняння регресії та оцінити тісноту і силу зв'язку двох рядів (за відхиленнями від тренда і за множинною регресійною моделлю з включенням у неї фактора часу).

Виконати прогноз рівнів одного ряду, виходячи з його зв'язку з рівнями іншого ряду.

8.2.9. В економетричній моделі вивчають залежність заробітної платні W_t (тис. дол. США) від прибутку X_t (тис. дол. США) за даними за 40 років. Оцінка параметрів моделі дала такі результати:

$$W_t = -4,79 + 0,237 X_t + 0,201 X_{t-1} + 0,229 X_{t-2} - 0,512t + e_t;$$

$$(-3,87)(2,10) \quad (1,49) \quad (2,32) \quad (-4,79)$$

$$R^2 = 0,9689; \quad DW = 0,4919.$$

У дужках вказано фактичні значення критерію Стюдента для коефіцієнтів регресії, t – змінна фактору часу.

Проаналізувати отримані результати регресійного аналізу: визначити короткостроковий та довгостроковий мультиплікатори, охарактеризувати структуру лагу. Перерахувати основні економетричні проблеми, які виникають під час побудови моделей з розподіленим лагом.

Рекомендована література

1. Економетрика. Методичні рекомендації до практичних завдань для студентів усіх спеціальностей першого (бакалаврського) рівня [Електронний ресурс] / уклад. Л. М. Малярець, О. В. Мартинова ; Харківський національний економічний університет ім. С. Кузнеця. – Харків : ХНЕУ ім. С. Кузнеця, 2020. – 81 с. – Режим доступу : <http://repository.hneu.edu.ua/handle/123456789/2342>.
2. Єгоршин О. О. Лабораторний практикум з економетрики в Excel : навч.-практ. посіб. / О. О. Єгоршин, Л. М. Малярець. – Харків : Вид. ХНЕУ, 2011. – 140 с. – Режим доступу : <http://repository.hneu.edu.ua/handle/123456789/5866>.
3. Малярець Л. М. Економіко-математичне моделювання : навч. посіб. / Л. М. Малярець. – Харків : Вид. ХНЕУ ім. С. Кузнеця, 2014. – 412 с. – Режим доступу : <http://repository.hneu.edu.ua/handle/123456789/11891>.
4. Малярець Л. М. Економіко-математичні методи і моделі : навч.-практ. посіб. / Л. М. Малярець. – Харків : Вид. ХНЕУ, 2010. – 312 с.
5. Малярець Л. М. Збірник вправ з навчальної дисципліни "Економіко-математичне моделювання" для студентів всіх галузей знань усіх форм навчання / Л. М. Малярець, Е. Ю. Железнякова, Є. Ю. Місюра. – Харків : Вид. ХНЕУ, 2011. – 320 с.
6. Малярець Л. М. Лабораторний практикум з навчальної дисципліни "Економетрика" в середовищі MATLAB : навч. посіб. / Л. М. Малярець, К. О. Ковальова. – Харків : ХНЕУ ім. С. Кузнеця, 2015. – 192 с. – Режим доступу : <http://repository.hneu.edu.ua/handle/123456789/13407>.
7. Малярець Л. М. Методичні рекомендації і завдання до виконання контрольних робіт з навчальної дисципліни "Економетрика" для студентів галузей знань 0305 "Економіка і підприємництво", 0306 "Менеджмент і адміністрування" денної форми навчання / уклад. Л. М. Малярець, О. В. Мінєнкова. – Харків : ХНЕУ ім. С. Кузнеця, 2016. – 56 с. – Режим доступу : <http://repository.hneu.edu.ua/handle/123456789/16028>.
8. Наконечний С. І. Економетрія : підручник / С. І. Наконечний, Т. О. Терещенко, Т. П. Романюк. – 3-тє вид., доп. та перероб. – Київ : КНЕУ, 2005. – 520 с.
9. Econometrics: multimedia guidelines to independent work for Bachelor's (first) degree students of speciality 051 "Economics" (content module 1) : [Electronic multimedia interactive edition of the combined use] / compiled by L. Malyarets, O. Martynova, Ie. Misiura. – Kharkiv : S. Kuznets KhNUE, 2023. – Access mode : <https://pns.hneu.edu.ua/course/view.php?id=9921>.

Додатки

Додаток А

Критичні точки розподілу F Фішера – Снедекора

(k_1 – кількість ступенів свободи більшої дисперсії,

k_2 – кількість ступенів свободи меншої дисперсії)

за рівня значущості $\alpha = 0,05$

k_2	k_1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4
3	10,1	9,55	9,28	9,13	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80
400	3,86	3,02	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,78
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75

Критичні точки розподілу F Фішера – Снедекора

(k_1 – кількість ступенів свободи більшої дисперсії, k_2 – кількість ступенів свободи меншої дисперсії) за рівня значущості $\alpha = 0,01$

k_2	k_1								
	1	2	3	4	5	6	7	8	9
1	4052,2	4999,5	5403,3	5624,6	5763,7	5859,0	5928,3	5981,6	6022,5
2	98,503	99,000	99,166	99,249	99,299	99,332	99,356	99,374	99,388
3	34,116	30,817	29,457	28,710	28,237	27,911	27,672	27,489	27,345
4	21,198	18,000	16,694	15,977	15,522	15,207	14,976	14,799	14,659
5	16,258	13,274	12,060	11,392	10,967	10,672	10,456	10,289	10,158
6	13,745	10,925	9,780	9,148	8,746	8,466	8,260	8,102	7,976
7	12,246	9,547	8,451	7,847	7,160	7,191	6,993	6,840	6,719
8	11,259	8,649	7,591	7,006	6,632	6,371	6,178	6,029	5,911
9	10,561	8,022	6,992	6,422	6,057	5,802	5,613	5,467	5,351
10	10,044	7,559	6,552	5,994	5,636	5,386	5,200	5,057	4,942
11	9,646	7,206	6,217	5,668	5,316	5,069	4,865	4,745	4,632
12	9,330	6,927	5,953	5,412	5,064	4,812	4,640	4,499	4,388
13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,191
14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	4,030
15	8,683	6,359	5,417	4,893	4,556	4,318	4,142	4,005	3,895
16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,780
17	8,400	6,112	5,185	4,669	4,336	4,102	3,927	3,791	3,682
18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,597
19	8,185	5,926	5,010	4,500	4,171	3,939	3,765	3,631	3,523
20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,457
21	8,017	5,780	4,874	4,369	4,042	3,812	3,640	3,506	3,398
22	7,945	5,719	4,817	4,313	3,988	3,758	3,587	3,453	3,346
23	7,881	5,664	4,765	4,264	3,939	3,710	3,539	3,406	3,299
24	7,823	5,614	4,718	4,218	3,895	3,667	3,496	3,363	3,256
25	7,770	5,568	4,676	4,177	3,865	3,627	3,457	3,324	3,217
26	7,721	5,526	4,637	4,140	3,818	3,591	3,421	3,288	3,182
27	7,677	5,488	4,601	4,106	3,785	3,558	3,388	3,256	3,149
29	7,598	5,421	4,538	4,045	3,725	3,500	3,330	3,198	3,092
30	7,568	5,390	4,510	4,018	3,699	3,474	3,305	3,173	3,067
40	7,314	5,179	4,313	3,828	3,514	3,291	3,124	2,993	2,888
60	7,077	4,977	4,126	3,649	3,339	3,119	2,953	2,823	2,719
120	6,851	4,787	3,949	3,480	3,174	2,956	2,792	2,663	2,559
∞	6,635	4,605	3,782	3,319	3,017	2,802	2,639	2,511	2,407

Критичні точки розподілу Стьюдента

Рівень значущості α (двостороння критична область)				
Кількість ступенів свободи k	$\alpha = 0,01$	$\alpha = 0,02$	$\alpha = 0,05$	$\alpha = 0,1$
1	63,70	31,82	12,70	6,31
2	9,92	6,97	4,30	2,92
3	5,84	4,54	3,18	2,35
4	4,60	3,75	2,78	2,13
5	4,03	3,37	2,57	2,01
6	3,71	3,14	2,45	1,94
7	3,50	3,00	2,36	1,89
8	3,36	2,90	2,31	1,86
9	3,25	2,82	2,26	1,83
10	3,17	2,76	2,23	1,81
11	3,11	2,72	2,20	1,80
12	3,05	2,68	2,18	1,78
13	3,01	2,65	2,16	1,77
14	2,98	2,62	2,14	1,76
15	2,95	2,60	2,13	1,75
16	2,92	2,58	2,12	1,75
17	2,90	2,57	2,11	1,74
18	2,88	2,55	2,10	1,73
19	2,86	2,54	2,09	1,73
20	2,85	2,53	2,09	1,73
21	2,83	2,52	2,08	1,72
22	2,82	2,51	2,07	1,72
23	2,81	2,50	2,07	1,71
24	2,80	2,49	2,06	1,71
25	2,79	2,49	2,06	1,71
26	2,78	2,48	2,06	1,71
27	2,77	2,47	2,05	1,71
28	2,76	2,46	2,05	1,70
30	2,75	2,46	2,04	1,70
40	2,70	2,42	2,02	1,68
60	2,66	2,39	2,00	1,67
120	2,62	2,36	1,98	1,66
∞	2,58	2,33	1,96	1,64
k	$\alpha = 0,005$	$\alpha = 0,01$	$\alpha = 0,025$	$\alpha = 0,05$
Рівень значущості α (одностороння критична область)				

Критичне значення критерія Дарбіна – Уотсона DW за $\alpha = 0,05$

n	Кількість незалежних змінних											
	$m = 1$		$m = 2$		$m = 3$		$m = 4$		$m = 5$		$m = 6$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,61	1,40										
7	0,70	1,37	0,47	1,90								
8	0,76	1,33	0,56	1,78	0,37	2,29						
9	0,82	1,32	0,63	1,70	0,46	2,13	0,30	2,59				
10	0,88	1,32	0,70	1,64	0,53	2,02	0,38	2,41	0,24	2,82		
11	0,93	1,32	0,76	1,60	0,60	1,93	0,44	2,28	0,32	2,65	0,12	2,89
12	0,97	1,33	0,81	1,58	0,66	1,86	0,51	2,18	0,38	2,51	0,16	2,67
13	1,01	1,34	0,86	1,56	0,72	1,82	0,57	2,09	0,45	2,39	0,21	2,49
14	1,05	1,35	0,91	1,55	0,77	1,78	0,63	2,03	0,51	2,30	0,26	2,35
15	1,08	1,36	0,95	1,54	0,81	1,75	0,69	1,98	0,56	2,22	0,30	2,24
16	1,11	1,37	0,98	1,54	0,86	1,73	0,73	1,94	0,62	2,16	0,35	2,15
18	1,16	1,39	1,05	1,54	0,93	1,70	0,82	1,87	0,71	2,06	0,44	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,89	1,83	0,79	1,99	0,52	1,92
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94	1,57	1,85
25	1,29	1,45	1,21	1,55	1,12	1,65	1,04	1,77	0,95	1,89	1,68	1,78
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85	0,76	1,73
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82	0,86	1,69
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80	0,91	1,67
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79	1,00	1,64
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78	1,07	1,64
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77	1,12	1,64
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77	1,21	1,64
70	1,58	1,64	1,55	1,67	1,53	1,70	1,49	1,74	1,46	1,77	1,28	1,65
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77	1,34	1,65
90	1,64	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78	1,38	1,66
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78	1,42	1,67
150	1,72	1,75	1,71	1,76	1,69	1,77	1,68	1,79	1,67	1,80	1,54	1,71
200	1,76	1,78	1,75	1,79	1,74	1,80	1,73	1,81	1,72	1,82	1,61	1,74

Зміст

Вступ	3
Практичне заняття 1. Особливості економетричних моделей та принципи їхньої побудови. Парна лінійна модель	4
1.1. Приклади розв'язання завдань	4
1.2. Завдання для самостійної роботи	11
Практичне заняття 2. Множинні регресійні моделі: визначення МНК-оцінок параметрів моделі	14
2.1. Приклади розв'язання завдань	14
2.2. Завдання для самостійної роботи	21
Практичне заняття 3. Множинні регресійні моделі: перевірка значущості економетричної моделі в цілому і кожного з її параметрів окремо	23
3.1. Приклади розв'язання завдань	23
3.2. Завдання для самостійної роботи	26
Практичне заняття 4. Проблеми в побудові лінійних множинних регресійних моделей	30
4.1. Приклади розв'язання завдань	32
4.2. Завдання для самостійної роботи	41
Практичне заняття 5. Узагальнені схеми регресійного аналізу	45
5.1. Приклади розв'язання завдань	45
5.2. Завдання для самостійної роботи	47
Практичне заняття 6. Системи економетричних рівнянь	50
6.1. Приклади розв'язання завдань	50
6.2. Завдання для самостійної роботи	55
Практичне заняття 7. Вивчення взаємозв'язків за часовими рядами	58
7.1. Приклади розв'язання завдань	58
7.2. Завдання для самостійної роботи	63
Практичне заняття 8. Моделювання одновимірних часових рядів	67
8.1. Приклади розв'язання завдань	69
8.2. Завдання для самостійної роботи	71
Рекомендована література	76
Додатки	77

НАВЧАЛЬНЕ ВИДАННЯ

ЕКОНОМЕТРИКА

**Методичні рекомендації
до виконання практичних завдань
для здобувачів вищої освіти усіх спеціальностей
першого (бакалаврського) рівня**

Самостійне електронне текстове мережеве видання

Укладачі: **Малярець** Людмила Михайлівна
Мінєнкова Олена Вадимівна
Шкребень Роман Петрович

Відповідальний за видання *Л. М. Малярець*

Редактор *В. О. Дмитрієва*

Коректор *В. О. Дмитрієва*

План 2025 р. Поз. № 133 ЕВ. Обсяг 82 с.

Видавець і виготовлювач – ХНЕУ ім. С. Кузнеця, 61165, м. Харків, просп. Науки, 9-А

*Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру
ДК № 4853 від 20.02.2015 р.*