МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Запорізький національний технічний університет

МЕТОДИЧНІ ВКАЗІВКИ ТА

РОЗРАХУНКОВО-ГРАФІЧНІ ЗАВДАННЯ

для самостійної роботи студентів економічних спеціальностей
усіх форм навчання

з курсу

"*Е К О Н О М Е Т Р И К А*"

2017

Методичні вказівки та розрахунково-графічні завдання для самостійної роботи студентів економічних спеціальностей усіх форм навчання з курсу “Економетрика” / Укл. Г.А.Шишканова – Запоріжжя: ЗНТУ, 2017. – 62 с.

Укладач:

Шишканова Ганна Анатоліївна, доцент, к.ф.-м.н.

Рецензент:

Мастиновський Юрій Вікторович, професор, к.т.н.

Відповідальний за випуск:

Шишканова Ганна Анатоліївна, доцент, к.ф.-м.н.

Затверджено
на засіданні кафедри прикладної математики ЗНТУ
Протокол № *9* від *06.04.2017*

Рекомендовано
до видання НМК факультету економіки та управління ЗНТУ
Протокол № *8* від *18.04.2017*

ЗМІСТ

ВСТУП

Основні задачі економетрії пов'язані з проблемами побудови типових економіко-математичних моделей, розробкою методів визначення їх параметрів та аналізом їх властивостей, а також прогнозуванням на майбутнє та оцінкою точності і надійності зробленого прогнозу. Економетрія дає основні напрямки використання цих моделей в економічних дослідженнях. Економетричні методи принесуть більше користі студентам, якщо в процесі їх вивчення розглядаються конкретні приклади їх застосування. Тому завдання по економетрії побудовані, виходячи з постановки економічних задач.

Оскільки економетрія базується на математичній статистиці та матричній алгебрі, то для вивчення її необхідно мати знання по цих двох дисциплінах. Економетрія має тісний зв'язок з іншими областями економічної науки. Знання, що набуті студентами при вивченні економетрії, широко використовуються в спеціальних економічних дисциплінах: макроекономіці, мікроекономіці, основах менежменту, маркетінгу, управлінні трудовими ресурсами і т.п.

Крім того, знайомство студентів з економетричними методами дає додаткові можливості в використанні обчислювальної техніки, розвиває аналітичні навички та є основою для проведення економічних досліджень.

Дані методичні вказівки містять розрахунково-графічні завдання, теоретичні відомості та приклади з курсу „Економетрія" для студентів усіх форм навчання економічних спеціальностей.

Контрольні завдання включають 10 варіантів. Номер варіанту виконується згідно останньої цифри залікової книжки, якщо остання цифра нуль, то потрібно виконувати 10 варіант.

Перед виконанням завдань необхідно створити певну теоретичну базу, використовуючи конспект лекцій та рекомендовану літературу. Передбачається, що всі завдання використовуються за допомогою ЕОМ пакета *EXCEL*, але без використання спеціальних економіко-статистичних пакетів прикладних програм.

Робота виконується в окремому зошиті, подається викладачеві на перевірку та захищається студентом.

1 ПАРНИЙ ЛІНІЙНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

1.1 Парна лінійна регресія

Для вивчення зв'язку між показником та факторами на основі статистичних даних використовується регресійний аналіз. Серед парних регресій найбільш поширеною, найкраще вивченою й простою в практиці моделювання, є парна регресія. Для двох випадкових величин *X* та *Y* ***регресія Y на X*** називається умовне математичне сподівання:

$$\hat{Y} = M(Y / X = x) = \varphi(x) . \qquad (1.1)$$

Графік такої функції називається *кривою регресії*. Припустимо гіпотезу, що між показником *Y* та фактором *X* існує статистична *лінійна залежність (* $\varphi(x) = \alpha + \beta x$ *)*, тоді зв'язок між показником *Y* та фактором *X*, з урахуванням випадкових відхилень, запишемо у вигляді:

$$Y = \alpha + \beta X + \varepsilon . \qquad (1.2)$$

Позначимо оцінки параметрів α та β через a та b. Тоді рівняння парної лінійної регресії

$$\hat{Y} = a + bX \qquad (1.3)$$

є оцінкою моделі $Y = \alpha + \beta X + \varepsilon$ ($M(\varepsilon) = 0$, $D(\varepsilon) = \sigma^2$ - невідома).

Параметри α та β оцінюються, за відомими статистичними даними вибірки (x_i , y_i), $i = \overline{1,n}$, величинами a і b за допомогою **методу найменших квадратів (*МНК*)**.

Розглянемо різницю y_i та $\hat{y}_i$

$$e_i = y_i - \hat{y}_i , i = \overline{1,n}$$

-де y_i - фактичні, а $\hat{y}_i$, розрахункові значення показника, e_i - відхилення спостережуваної точки (x_i, y_i) від точки $(x_i, \hat{y}_i)$ загладжуваної прямої. МНК полягає в підборі таких оцінок параметрів регресії, для яких сума квадратів відхилень є мінімальною, тобто:

$$Q(a,b) = \sum_{i=1}^{n} e_i^2 \Rightarrow \min,$$

де функція

$$Q(a,b) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2 .$$

Необхідною умовою існування мінімуму функції *Q(a, b)* є рівність нулю її частинних похідних:

$$\frac{\partial Q}{\partial a} = -2\sum_{i=1}^{n} (y_i - a - bx_i) = 0;$$

$$\frac{\partial Q}{\partial b} = -2\sum_{i=1}^{n} (y_i - a - bx_i)x_i = 0,$$

звідки отримаємо систему лінійних рівнянь

$$\begin{cases} \sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i, \\ \sum_{i=1}^{n} y_i x_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2, \end{cases} \tag{1.4}$$

яка має єдинйй розв'язок

$$b = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{1}{2}\left(\sum_{i=1}^{n} x_i\right)^2}; \tag{1.5}$$

$$a = \frac{\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i}{n}. \qquad (1.6)$$

Після спрощення вираз (1.6) набуде вигляду

$$a = \overline{y} - b\overline{x} \qquad (1.7)$$

де середні значення фактору та показника

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \ , \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ .$$

З метою спрощення виразу для b чисельник та знаменник у виразі (1.5) помножимо на $1/n$. Тоді отримаємо, що

$$b = \frac{\frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i\right) - \overline{x}\cdot\overline{y}}{\frac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) - \overline{x}^2}. \qquad (1.8)$$

Вираз (1.8) можна записати ще так:

$$b = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Або

$$b = \frac{Cov(x;y)}{Var(x)}; \qquad (1.9)$$

де

$$Cov(x;y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y}; \ Var(x) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2.$$

1.2 Коефіцієнт кореляції

Коефіцієнт кореляції характеризує щільність лінійної залежності між показником *Y* і фактором *X*. Його обчислюють за формулою:

$$r = \frac{Cov(x;y)}{\sqrt{Var(x)\cdot Var(y)}}. \qquad (1.10)$$

Він змінюється у межах від **-1<** r **<1**.

Додатне значення коефіцієнта кореляції свідчить про прямий зв’язок між показниками, а від’ємне – про зворотний.

Коли коефіцієнт кореляції за абсолютною величиною є близьким до 1, це свідчить про наявність сильного лінійного зв’язку $(r \approx \pm 1)$ - *щільність зв’язку велика*; у протилежному випадку, коли коефіцієнт кореляції є близьким до нуля $(r \approx 0)$, *зв’язку немає*.

1.3 Коефіцієнт детермінації

За допомогою ***коефіцієнта детермінації*** також вимірюється щільність лінійного зв’язку між двома або більше показниками та перевіряється адекватність побудованої регресійної моделі реальній дійсності. Коефіцієнт детермінації обчислюють за формулою:

$$R^2 = \frac{Var(\hat{y})}{Var(y)}. \qquad (1.11)$$

Коефіцієнт детермінації завжди невід’ємний і перебуває у межах від нуля до одиниці $\left(0 \le R^2 \le 1\right)$. Коли R^2 близький до нуля зв’язку немає. Коли R^2 близький до 1 зв’язок хороший.

Коефіцієнт кореляції та коефіцієнт детермінації зв'язані такою рівністю $R^2 = r^2$.

1.4 Коефіцієнт еластичності

В економічних задачах для оцінки впливу на показник будь-якого фактора використовують ***коефіцієнт еластичності*** фактора $\widehat{Y}$ за показником *X*, який обчислюють за формулою:

$$K_X = \frac{\widehat{Y}' \cdot X}{\widehat{Y}},. \qquad (1.12)$$

де Y' похідна по X .

Коефіцієнт еластичності показує, на скільки відсотків зміниться показник, якщо фактор зміниться на один відсоток.

Для парної лінійної регресії (1.3) коефіцієнт еластичності обчислюється за формулою:

$$K_{x_i} = \frac{(a + bx_i)' x_i}{\widehat{y}_i} = \frac{bx_i}{\widehat{y}_i}, \quad i = \overline{1,n}. \qquad (1.13)$$

1.5 Перевірка гіпотези про значущість коефіцієнта кореляції

Коефіцієнт кореляції є випадковою величиною, тому доцільно зробити перевірку гіпотези про відсутність кореляційного зв'язку. Для цього перевіряється *нульова гіпотеза* $H_0 : r = 0$.

Тоді *альтернативна гіпотеза* $H_1 : r \neq 0$.

Припускаючи, що двовимірна випадкова величина розподілена за нормальним законом, для вибірки обчисляємо *t-статистику,* яка розподілена за законом ***Стьюдента*** з $k = n - k_2 - 1$ ступенями вільності

$$t = r\sqrt{\frac{n - k_2 - 1}{1 - r^2}},$$

де k_2 число факторів включених у регресію. В даному випадку парної лінійної регресії $k_2 = 1$. Якщо $-t_{крит} < t < t_{крит}$, де $t_{крит}(\gamma, k)$ для заданої надійності $\gamma = 1 - \alpha$ (α - рівень значущості) береться із таблиць, то гіпотеза H_0 вірна з заданою надійністю, а якщо ця умова не виконується то слід відкинути про відсутність кореляції і прийняти альтернативну гіпотезу H_1 про наявність залежності між випадковими величинами Y та X .

1.6 Деякі припущення у парному регресивному аналізі

Основні ймовірнісні припущення щодо лінійної моделі (1.2) $y_i = \alpha + \beta x_i + \varepsilon_i$:

1. залишки ε_i, $i = \overline{1,n}$ є випадкові величини з однаковим математичним сподіванням $M(\varepsilon_i) = 0$ і однаковою дисперсією $D(\varepsilon_i) = \sigma^2$, σ^2 - невідома;

2. залишки ε_i та ε_j некорельовані при $i \neq j$, тому $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$, $i = \overline{1,n}$, $j = \overline{1,n}$, $i \neq j$;

3. залишки ε_i, $i = \overline{1,n}$ є нормально розподілені випадкові величини з математичним сподіванням, що дорівнює нулю, і дисперсією σ^2, тобто $e_i \approx N\left(0, \sigma^2\right), i = \overline{1,n}.$

1.7 Стандартні відхилення параметрів регресії та їх надійні інтервали

Якщо припущення 1-3 п. 1.6 щодо розкиду спостережень відносно лінії регресії вірні, то надійні інтервали для параметрів *α, β* лінійної регресії є такими:

$$b - \Delta b < \beta < b + \Delta b, \tag{1.14}$$

$$a - \Delta a < \alpha < a + \Delta a, \tag{1.15}$$

де $\Delta b = c.n.(b) \cdot t_{крит}$, $\Delta a = c.n.(a) \cdot t_{крит}$. Стандартні похибки (с.п.) параметрів *b*, *a* лінійної регресії обчислюються так:

$$c.n.(b) = \sqrt{\frac{S_u^2}{nVar(x)}}\,,\ c.n.(a) = \sqrt{\frac{S_u^2}{n}\left(1 + \frac{\bar{x}^2}{Var(x)}\right)}\,, \text{ де } S_u^2 = \frac{n}{n-2} Var(e)\,.$$

1.8 Надійна зона регресії

Надійна зона значень $\widehat{y}_i$ визначається нерівностями:

$$\bar{y}_i - \Delta\widehat{y}_i < y_i < \widehat{y}_i + \Delta\widehat{y}_i \qquad (1.16)$$

де $\Delta\widehat{y}_i = c.n.(\widehat{y}) \cdot t_{kpuт}$, $c.n.(\widehat{y}_i) = \sqrt{\dfrac{S_u^2}{n}\left(1 + \dfrac{\left(x_i - \bar{x}^2\right)}{Var(x)}\right)}$.

1.9 Прогноз і його надійний інтервал

Прогнозуванням називається наукове передбачення ймовірносних шляхів розвитку явищ і процесів для більш-менш віддаленого майбутнього. Основою регресійного аналізу є можливість спрогнозувати (з певною надійністю) значення показника $\widehat{y}_p$ при новому значенні фактора x_p .

Значення прогнозу показника при заданому значенні фактора x_p згідно з побудованою лінійною регресією визначається так

$$\widehat{y}_p = a + bx_p \, .$$

Довірчий інтервал для прогнозу:

$$\widehat{y}_p - \Delta\widehat{y}_p < y_p < \widehat{y}_p + \Delta\widehat{y}_p \, , \qquad (1.17)$$

де $\Delta\widehat{y}_p = c.n.(\widehat{y}_p) \cdot t_{kpum}$, $c.n.(\widehat{y}_p) = \sqrt{\dfrac{S_u^2}{n}\left(n + 1 + \dfrac{(x_p - \bar{x})^2}{Var(x)}\right)}$.

1.10 *F*-тест для перевірки адекватності регресійної моделі

Для оцінки адекватності парної регресії спостережуваним даним можна використовувати ***критерій Фішера*** (***F-тест***). Розрахункове значення *F*-критерію знаходиться за формулою

$$F = \frac{R^2 / k_2}{(1 - R^2)/(n - k_2 - 1)}, \qquad (1.18)$$

де R^2 - коефіцієнт детермінації, k_2 - число включених у регресію факторів, (для лінійної регресії k_2=1), n - число дослідів.

Для заданої надійності $\gamma = 1 - \alpha$ (α - рівень значущості) та числі ступенів вільності $k_1 = k$, $k = n - k_2 - 1$, знаходимо табличне значення $F_{крит}$ з таблиць розподілу Фішера.

Якщо $F > F_{крит}$, то з надійністю γ можна вважати, що розглянута математична модель адекватна експериментальним даним і на підставі прийнятої моделі проводити економічний аналіз, у протилежному випадку з надійністю γ розглянуту парну регресію не можна вважати адекватною експериментальним даним.

1.11 Приклад побудови економічної моделі парної лінійної регресії

Завдання.

Побудувати регресійну залежність $\hat{y} = a + bx$ між *y* (витрати на бензин) та *x* (час) відповідно до другого та третього стовпчиків таблиці 1.1, де величини задані в мільярдах умовних грошових одиниць за період з 1959 по 1983 роки, і зробити повний економічний аналіз отриманої моделі.

Результати розв'язку.

Для розрахунків використовуємо пакет *Excel*. Результати обчислень наведено у таблиці 1.1.

Таблиця 1.1

№	*x*	*y*	*y*^	*e*	*y*^-	*y*^+	*Кел*
1	1959	13,7	14,18	- 0,48	12,58	15,78	87,11
2	1960	14,2	14,81	- 0,61	13,31	16,31	83,44
3	1961	14,3	15,44	- 1,14	14,03	16,85	80,08
4	1962	14,9	16,07	- 1,17	14,75	17,39	76,97
5	1963	15,3	16,70	- 1,40	15,47	17,93	74,11
6	1964	16	17,33	- 1,33	16,18	18,48	71,45
7	1965	16,8	17,96	- 1,16	16,89	19,03	68,97
8	1966	17,8	18,59	- 0,79	17,59	19,60	66,67
9	1967	18,4	19,22	- 0,82	18,28	20,17	64,52
10	1968	19,9	19,85	0,05	18,96	20,75	62,50
11	1969	21,4	20,48	0,92	19,63	21,34	60,61
12	1970	22,9	21,11	1,79	20,28	21,95	58,83
13	1971	24,2	21,74	2,46	20,92	22,57	57,15
14	1972	25,4	22,37	3,03	21,54	23,21	55,57
15	1973	26,2	23,00	3,20	22,15	23,86	54,07
16	1974	24,8	23,64	1,16	22,74	24,53	52,66
17	1975	25,6	24,27	1,33	23,32	25,21	51,31
18	1976	26,8	24,9	1,90	23,89	25,90	50,04
19	1977	27,7	25,53	2,17	24,45	26,60	48,83
20	1978	28,3	26,16	2,14	25,01	27,31	47,68
21	1979	27,4	26,79	0,61	25,56	28,02	46,58
22	1980	25,1	27,42	-2,32	26,10	28,74	45,53
23	1981	25,1	28,05	-2,95	26,64	29,46	44,53
24	1982	25,3	28,68	-3,38	27,18	30,18	43,57
25	1983	26,1	29,31	-3,21	27,71	30,91	42,66
прогноз	***1984***		***29,94***	***-29,94***	***25,48***	***34,40***	***41,78***
∑	49275	543,6	543,6				
Ср. арифм.	1971	21,74	21,74				

n=25 b=0,63 $r(x,y)$=0,92 var(y^)=20,67

k_1=23 a=-1220,896 R^2=0,85 var(y)=24,32

с.п.(a)=108,92 $a- = -1446{,}26$ $a+ = -995{,}53$ var(x)=52,0

$t_{крит}$=2,069 $F_{крит}$=4,28 cov(x,y)=32,78 var(e)=3,65

t=11,41 F=130,15 с.п.(b)=0,55 $b- = 0{,}516$ $b+ = 0{,}745$ S_u^2=3,97

Для побудови графіків у „*Майстері діаграм*” вибрано тип „*Точкова*”.

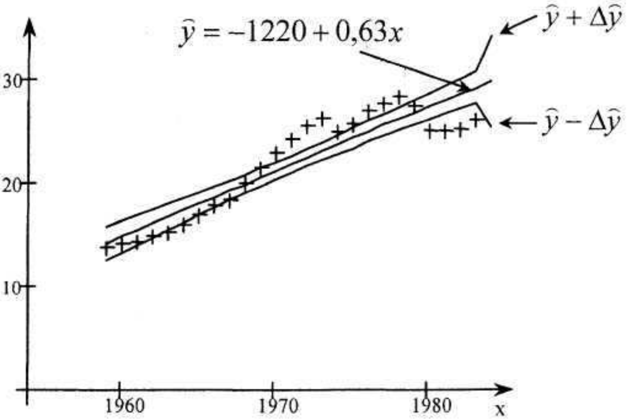

Рисунок 1.1

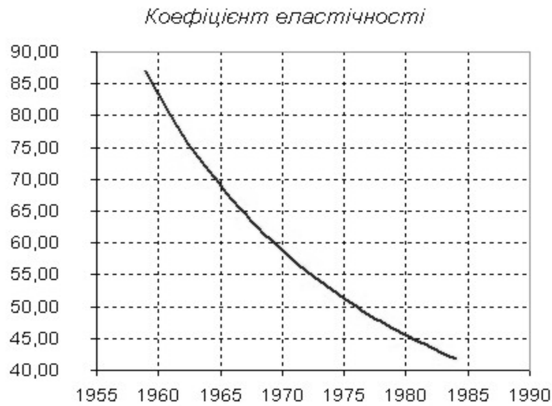

Рисунок 1.2

Висновок

1. Модель має вигляд $\widehat{y} = -1220{,}896 + 0{,}63x$.
2. Якщо x збільшити на 1 одиницю то y збільшиться на 0,63 одиниць у своїх вимірах, тобто на 630 млн. умовних грошових одиниць.

3. Коефіцієнт кореляції r=0,92. Оскільки він близький до 1 то можна сказати, що між x та y існує пряма лінійна залежність (точки знаходяться на зростаючій прямій). Отже із збільшенням часу x (тобто років), витрати на бензин збільшуються.
4. Коефіцієнт детермінації R^2=0,85. Він близький до 1 і говорить про те, що між x та y є тісний лінійний зв'язок.
5. Коефіцієнт еластичності змінюється в межах від 41,78 до 87,11.
6. Перевірка за допомогою t-тесту показала, що t=11,41 є більшим за $t_{крит}$=2,069, тобто коефіцієнт кореляції знайдено правильно.
7. Перевірка за допомогою критерію Фішера показала, що F=130,15 є більшим за $F_{крит}$=4,28 тобто модель адекватна статистичним даним.
8. Параметри α та β знаходяться в таких межах:

$$-1446{,}26<\alpha<-995{,}53;$$
$$0{,}516<\beta<0{,}745.$$

9. Значення прогнозу в точці x_p=1984 дорівнює y_p=29,94.

Довірчий інтервал для прогнозу є 34,4< y_p <41,78.

2 НЕЛІНІЙНИЙ ПАРНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

2.1 Нелінійна парна регресія

Разом з лінійною регресією в економічній практиці використовуються нелінійні моделі, які характеризуються нелінійною залежністю між показником і фактором. За параметрами, парні нелінійні моделі поділяються на два типи:

a) ***нелінійна за факторами***, але лінійна відносно невідомих параметрів (квазілінійна);

b) ***нелінійна за факторами та параметрами***.

Загальний випадок квазілінійних парних регресії можна записати у вигляді:

$$Y = a + b\varphi(X), \tag{2.1}$$

де $\varphi(X)$ - відома нелінійна функція, a та b - невідомі параметри, які потрібно визначити за допомогою МНК, використовуючи статистичні дані (x_i , y_i), де $i = \overline{1,n}$. Заміною незалежної змінної на нову $Z = \varphi(X)$ нелінійна парна регресія приводиться до лінійної. Формули для оцінок параметрів набудуть вигляду:

$$b = \frac{Cov(Z,Y)}{Var(Z)}; \tag{2.2}$$

$$a = \bar{y} - b\bar{z}; \tag{2.3}$$

де $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$; $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

Коефіцієнт еластичності для парної квазілінійної регресії оцінюється за формулою:

$$K_{\varphi(X)} = \frac{b\varphi'(x)X}{\widehat{Y}} = \frac{b\varphi'(x)X}{a + b\varphi(X)}. \tag{2.4}$$

2.2 Приклади деяких парних квазілінійних регресій

В загальному випадку процедура визначення параметрів здійснюється наступним чином. За заданим статистичним рядом треба побудувати новий, виконуючи заміну $Z = \varphi(X)$, треба знайти оцінки параметрів, їх надійні інтервали, визначити надійні зони базисних даних та необхідних коефіцієнтів, аналогічно як і в розділі 1 для лінійної парної регресії.

Розглянемо деякі види квазілінійних моделей.

1. $\hat{y} = a + \frac{b}{x}$ - модель парної нелінійної регресії, яка після заміни $z = \frac{1}{x}$ зводиться до лінійної регресії $\hat{y} = a + bz$, $K_{ел_i} = -\frac{b}{x_i \hat{y}_i}$ - коефіцієнт еластичності;

2. $\hat{y} = a + b \ln x$, заміна $z = \ln x$, $K_{ел_i} = \frac{b}{\hat{y}_i}$;

3. $\hat{y} = a + be^x$, заміна $z = e^x$, $K_{ел_i} = \frac{be^{x_i} x_i}{\hat{y}_i}$;

4. $\hat{y} = a + b\sqrt{x}$, заміна $z = \sqrt{x}$, $K_{ел_i} = \frac{b\sqrt{x_i}}{2\hat{y}_i}$;

5. $\hat{y} = a + bx^2$, заміна $z = x^2$, $K_{ел_i} = \frac{2bx_i^2}{\hat{y}_i}$;

6. $\hat{y} = a + bx^3$, заміна $z = x^3$, $K_{ел_i} = \frac{3bx_i^3}{\hat{y}_i}$.

2.3 Приклади регресій нелінійних за факторами і параметрами

Для таких регресій не існує єдиного способу зведення їх до лінійних.

1. Регресія вигляду $\hat{y} = \dfrac{1}{a+bx}$, заміною $y_1 = \dfrac{1}{y}$ призводиться до лінійної $\hat{y}_1 = a + bx$, Коефіцієнт еластичності матиме вигляд $K_{ел_i} = \dfrac{bx_i}{a+bx_i}$.

2. Регресія $\hat{y} = \dfrac{1}{a+be^{-x}}$, після двох замін $y_1 = \dfrac{1}{y}$ та $z = e^{-x}$, зводиться до лінійної регресії $\hat{y}_1 = a + bz$, $K_{ел_i} = \dfrac{bx_i}{ae^{-x_i}+b}$.

3. Регресію $\hat{y} = ab^x$, за допомогою логарифмування та послідовної зміни величин та параметрів приводимо до лінійної регресії. Логарифмуємо $\ln\hat{y} = \ln a + x\ln b$, вводимо заміни $y_1 = \ln y$, $a_1 = \ln a$, $b_1 = \ln b$, лінійна регресія набуде вигляду $\hat{y}_1 = a_1 + b_1 x$. Для такої регресії спочатку знаходять оцінки параметрів a_1 та b_1, а потім оцінки параметрів a та b як $a = e^{a_1}$, $b = e^{b_1}$. Оцінка коефіцієнта еластичності знаходиться за формулою $K_{ел_i} = x_i \ln b$.

4. Регресія $\hat{y} = ae^{bx}$, після логарифмування та замін $y_1 = \ln y$, $a_1 = \ln a$ зводиться до лінійної регресії з коефіцієнтом еластичності $K_{ел_i} = bx_i$;

5. Регресія $\hat{y} = ax^b$, після логарифмування та замін $y_1 = \ln y$, $a_1 = \ln a$, $z = \ln x$ зводиться до лінійної регресії. Коефіцієнт еластичності обчислюється за формулою $K_{ел_i} = b$.

2.4 Оцінка адекватності нелінійної парної регресії

Для перевірки нелінійної парної регресії на адекватність, використовують F-критерій Фішера, аналогічно як і в розділі 1.10 для парної лінійної регресії.

2.5 Довірча зона базисних даних нелінійної регресії

Довірча зона базисних даних нелінійної за факторами регресії знаходиться за тими ж формулами що і для лінійної, тільки замість X використовується $Z=\varphi(X)$:

$$\widehat{y}_i-\Delta\widehat{y}_i<y_i<\widehat{y}_i+\Delta\widehat{y}_i,$$

де $\Delta\widehat{y}_i=c.n.(\widehat{y}_i)\cdot t_{крит}$, $\quad c.n.(\widehat{y}_i)=\sqrt{\frac{S_u^2}{n}\left(1+\frac{(z_i-\bar{z})^2}{Var(z)}\right)}$.

2.6 Прогноз і його надійний інтервал

Для квазілінійних регресій прогноз знаходять значення прогнозу $\widehat{y}_p$ при заданому значенні фактора x_p ($z_p=\varphi(x_p)$) обчислюється за формулою:

$$\widehat{y}_p=a+bz_p.$$

Довірчий інтервал для прогнозу визначається нерівностями

$$\widehat{y}_p-\Delta\widehat{y}_p<y_p<\widehat{y}_p+\Delta\widehat{y}_p,$$

де $\Delta\widehat{y}_p=c.n.(\widehat{y}_p)\cdot t_{крит}$, $\quad c.n.(\widehat{y}_p)=\sqrt{S_u^2\left(1+\frac{1}{n}+\frac{(z_p-\bar{z})^2}{nVar(z)}\right)}$.

2.7 Довірчі інтервали нелінійної за параметрами і факторами регресії

У тих випадках, коли нелінійна регресія зводиться до лінійної за допомогою логарифмування та заміни змінних, довірча зона спочатку шукається для лінійної регресії, потім використовуючи зворотні перетворення знаходимо надійні інтервали нелінійної регресії. Показникова регресія $\widehat{y}=ab^x$, за допомогою логарифмування $\ln\widehat{y}=\ln a+x\ln b$ та заміни $y_1=\ln y$, $a_1=\ln a$, $b_1=\ln b$ зводиться до лінійної регресії вигляду $\widehat{y}_1=a_1+b_1x$.

За формулами (1.7), (1.9), (1.14), (1.15), (1.16), (1.17) лінійної парної регресії обчислюємо коефіцієнти a_1 та b_1, Δa_1, Δb_1, Δy_1 та Δy_{1p} і після потенціювання цих значень знаходимо межі надійних інтервалів для заданої нелінійної регресії:

$$\hat{y}_i \pm \Delta\hat{y}_i = \exp(\hat{y}_{1i} \pm \Delta\hat{y}_{1i}).$$

Зауваження. Логарифмічне перетворення призводить до того, що оцінки параметрів базуються тепер не на мінімізації суми квадратів відхилень, а на мінімізації суми квадратів відхилень логарифмів, тобто:

$$Q = \sum_{i=1}^{n}(\ln y_i - \ln \hat{y}_i)^2 = \sum_{i=1}^{n}\ln^2\frac{y_i}{\hat{y}_i} \neq \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}e_i^2 .$$

Звідси отримується, деяке зміщення оцінок параметрів знайдених звичайним МНК.

2.8 Приклад побудови економічної моделі парної нелінійної регресії

Завдання.

Побудувати нелінійну регресійну залежність, яка б найкраще описувала зв'язок між y - дохід який знаходиться у власному розпорядженні та x - витрати на бензин, і зробити повний економічний аналіз отриманої моделі. Розглянути нелінійні моделі

1. $\hat{y} = a + \frac{b}{x}$; 2. $\hat{y} = a + b\ln x$, 3. $\hat{y} = \frac{1}{a + bx}$; 4. $\hat{y} = ax^b$, 5. $\hat{y} = ab^x$,

побудувати усі вказані моделі парних нелінійних регресій.

Результати розв'язку.

Для розрахунків використовуємо пакет *Excel*. Для побудови графіків у „*Майстері діаграм*” вибрано тип „*Точкова*”.

Точність моделі. Із декількох моделей найбільш ***точною*** є та, в якій найменша сума квадратів відхилень: $\sum_{i=1}^{n}e^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$.

де y та $\hat{y}$ змінні, які описують нелінійну модель.

Із п'яти моделей найкраще описує статистичні дані перша модель (сума квадратів відхилень в якої є найменша), тому, для скорочення, тут наведено розрахункові дані тільки для цієї регресії – таблиця 2.1.

На рис. 2.1-2.5 зображено графіки, які відповідають математичним моделям нелінійних парних регресій 1-5, відповідно.

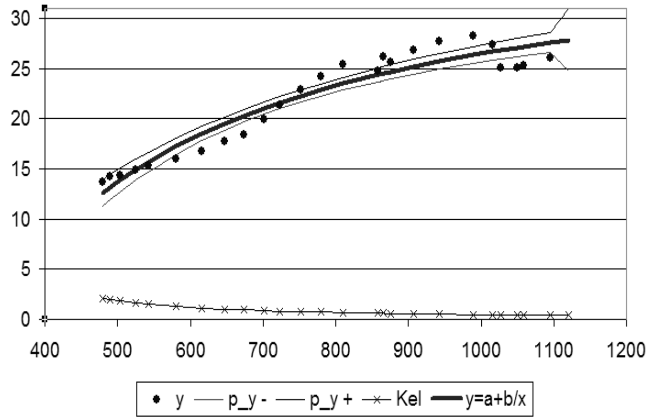

Рисунок 2.1

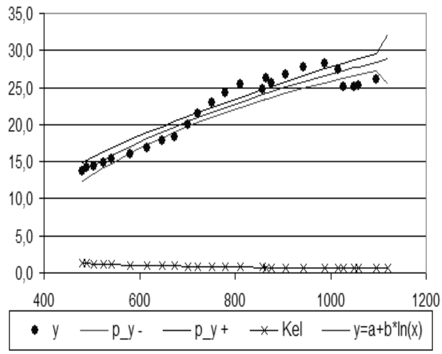

Рисунок 2.2

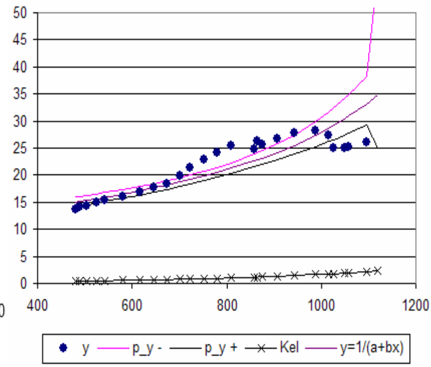

Рисунок 2.3

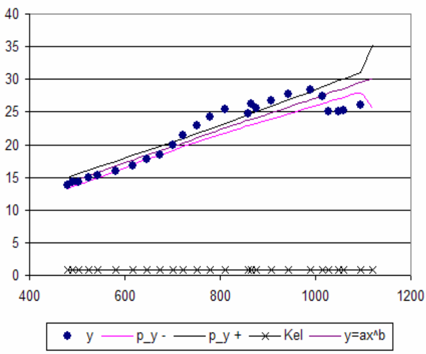

Рисунок 2.4

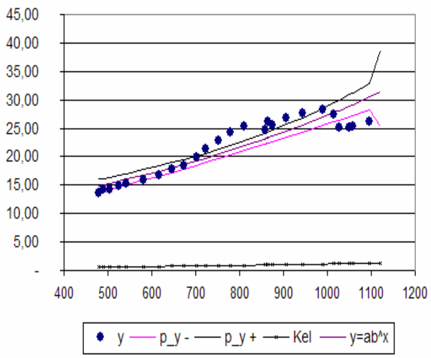

Рисунок 2.5

Висновки

1. Із п’яти моделей найкраще описує статистичні дані перша модель, тому що для цієї регресії сума квадратів відхилень є найменша, яка дорівнює 48,37.

Модель набуває вигляд $\widehat{y} = 39{,}3 - \dfrac{12795{,}12}{x}$.

2. Коефіцієнт детермінації R^2 =0,92; він близький до 1 і свідчить про те, що між x та y є тісний зв'язок;

3. Коефіцієнт кореляції у лінеарізованій моделі r=-0,96; він близький до 1, тому можна стверджувати, що між x та y існує відповідна залежність.

4. Коефіцієнт еластичності змінюється в межах від 0,42 до 2,11.

5. Перевірка за допомогою критерію Фішера показала, що F=266,09 є більшим за $F_{крит}$ =4,28, тобто модель є адекватною статистичним даним.

6. Параметри α та β задовольняють такі нерівності:

$$36{,}99 < \alpha < 41{,}61;$$
$$-11172{,}24 < \beta < -14418{,}01.$$

7. Значення прогнозу в точці x_p=1120 дорівнює y_p=27,88. Довірчий інтервал для прогнозу є 24,72 $< y_p <$ 31,03.

Таблиця 2.1

№	***x***	***y***	***x1=1/x***	***y^***	***y^-***	***y^+***	***Кел***	***e^2***
1	479,7	13,7	0,00208	12,63	11,32	13,93	2,11	1,153292939
2	489,7	14,2	0,00204	13,17	11,93	14,41	1,98	1,059316025
3	503,8	14,3	0,00198	13,90	12,74	15,06	1,83	0,158376125
4	524,9	14,9	0,00191	14,92	13,87	15,98	1,63	0,00052698
5	542,3	15,3	0,00184	15,71	14,73	16,68	1,50	0,164092583
6	580,8	16,0	0,00172	17,27	16,44	18,10	1,28	1,610592795
7	616,3	16,8	0,00162	18,54	17,81	19,26	1,12	3,020884064
8	646,8	17,8	0,00155	19,52	18,85	20,18	1,01	2,948321467
9	673,5	18,4	0,00148	20,30	19,67	20,93	0,94	3,614964667
10	701,3	19,9	0,00143	21,05	20,45	21,66	0,87	1,332633829
11	722,5	21,4	0,00138	21,59	20,99	22,19	0,82	0,036004522
12	751,6	22,9	0,00133	22,28	21,67	22,88	0,76	0,390106604
13	779,2	24,2	0,00128	22,88	22,26	23,50	0,72	1,746585538
14	810,3	25,4	0,00123	23,51	22,87	24,15	0,67	3,577161614
15	865,3	26,2	0,00116	24,51	23,82	25,21	0,60	2,848196923
16	858,4	24,8	0,00116	24,39	23,71	25,08	0,61	0,165258973
17	875,8	25,6	0,00114	24,69	23,98	25,40	0,59	0,828790357
18	906,8	26,8	0,00110	25,19	24,45	25,93	0,56	2,595101722
19	942,9	27,7	0,00106	25,73	24,94	26,51	0,53	3,883686826
20	988,8	28,3	0,00101	26,36	25,52	27,20	0,49	3,766668317
21	1 015,5	27,4	0,00098	26,70	25,83	27,57	0,47	0,490791036
22	1 026,6	25,1	0,00097	26,84	25,95	27,72	0,46	3,01254835
23	1 049,3	25,1	0,00095	27,11	26,20	28,01	0,45	4,021228613
24	1 058,3	25,3	0,00094	27,21	26,29	28,13	0,44	3,644281232
25	1 095,4	26,1	0,00091	27,62	26,66	28,58	0,42	2,305793455
рг	1120,0		0,00089	27,88	24,72	31,03	0,41	
∑	19505,8	543	0,03	543,60				48,37

K=23 b=-12795,12 r(x,y)=-0,96 Var(y^)=22,387

$t_{крит}$ =2,069 a=39,30 R^2 =0,92 Var(y)=24,322

t=-16,31 с.о.(a)=1,11 с.о.(b)=784,38 Var(x)=0,00000014

$F_{крит}$ =4,28 a+=41,61 b+=-11172,24 Var(e)=1,935

F=266,09 a-=36,99 b-=-14418,01 cov(x,y)=-0,00175

3 МНОЖИННИЙ ЛІНІЙНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

3.1 Множинна регресія. Основні поняття

Явище, яке залежить від багатьох факторів, можна описати за допомогою множинної регресії. Дослідивши взаємозв'язок процесів у минулому можна побудувати функціональний зв'язок між ними і з деякою вірогідністю планувати майбутнє.

Нехай Y - деяка випадкова величина, яка є лінійною функцією деяких змінних (факторів). Наприклад, якщо Y - річний попит на товари першої необхідності на душу населення, то факторами можуть бути x_1 - хліб та хлібобулочні вироби, x_2 — молоко та молочні продукти, x_3 - овочі та фрукти, x_4 - одяг та швейні вироби і т.д. Якщо величини x_i, $i=\overline{1,m}$ виражають річну реалізацію товару в у.о. на душу населення, то лінійна модель попиту набуває вигляду

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \ldots + \alpha_m x_{mi} + l_i\,, \tag{3.1}$$

де $x_{1i}, x_{2i}, \ldots, x_{mi}$ - відомі незалежні змінні (наприклад, кількість хліба та хлібобулочних виробів проданих в i-му обліковому році на душу населення), l_i - випадкова величина, яка є відхиленням від передбачуваної функціональної залежності в i-му році (флуктуація або помилка), α_i - невідомі параметри. Систему рівнянь (3.1) можна записати у матричній формі

$$\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{21} & \ldots & x_{m1} \\ 1 & x_{12} & x_{22} & \ldots & x_{m2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} & \ldots & x_{mn} \end{pmatrix} \times \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \ldots \\ \alpha_m \end{pmatrix} + \begin{pmatrix} l_1 \\ l_2 \\ \ldots \\ l_n \end{pmatrix}, \tag{3.2}$$

або

$$Y = X\alpha + l\,. \tag{3.3}$$

Матриця X розмірності $n\times(m+1)$ називається регресійною, матриця стовпчик Y розмірності $n\times1$, α розмірності $(m+1)\times1$ і вектор l розмірності $n\times1$.

Модель (3.1) є досить загальною. Наприклад, розглянемо квазілінійну модель

$$y_i=\alpha_0+\alpha_1x_{1i}+\alpha_2x_{2i}^2+\alpha_3\sqrt{x_{3i}}+\alpha_4\frac{1}{x_{4i}}+l_i,\ \ i=\overline{1,n}.$$

після заміни $z_{1i}=x_{1i}, z_{1i}=x_{2i}^2, z_{1i}=\sqrt{x_{3i}}, z_{1i}=\dfrac{1}{x_{4i}},\ \ i=\overline{1,n},$ вона зводиться до моделі (3.1).

Для вектора похибок l введемо такі припущення:

- математичне сподіванням i-го значення випадкової величини дорівнює нулю $M(l_i)=0$, $i=\overline{1,n}$, ;

- l_i для кожного спостереження має однакову теоретичну дисперсію $D(l_i)=\sigma^2$ (σ^2 - невідома), $i=\overline{1,n}$, ;

- залишки l_i та l_j некорельовані при $i\neq j$, тому $\operatorname{cov}(l_i,l_j)=0$, $i=\overline{1,n}$, $j=\overline{1,n}$, $i\neq j$;

- розподіл l не залежить від розподілу пояснюючих змінних x_k, $k=\overline{1,n}$;

- змінні x_{ki} та x_{pi} - незалежні, $\operatorname{cov}(x_{ki},x_{pi})=0$, $k\neq p$, $i=\overline{1,n}$;

- випадковий доданок l має нормальний розподіл.

3.2 Оцінки параметрів множинної лінійної регресії МНК

Нехай Y результати спостережень над об'єктом за n періодів років, кварталів, місяців, декад) або спостереження за один період над n об'єктами, в яких показник Y залежить від m факторів x_i, $i=\overline{1,m}$, отримані такі дані (нехай m = 2).

Y	y_1	y_2	...	y_j	...	y_n
x_1	x_{11}	x_{12}	...	x_{1j}	...	x_{1n}
x_2	x_{21}	x_{22}	...	x_{2j}	...	x_{2n}

Приймається гіпотеза, що між показником Y та факторами x_i $i \in \{1,2\}$ існує лінійна залежність

$$Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + l\,,$$

де $\alpha_0, \alpha_1, \alpha_2$ - параметри лінійної множинної регресії, l - випадкові відхилення. Для n дослідів маємо n рівнянь:

$$\begin{aligned} y_1 &= \alpha_0 + \alpha_1 x_{11} + \alpha_2 x_{21} + l_1; \\ y_2 &= \alpha_0 + \alpha_1 x_{12} + \alpha_2 x_{22} + l_2; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_j &= \alpha_0 + \alpha_1 x_{1j} + \alpha_2 x_{2j} + l_j; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= \alpha_0 + \alpha_1 x_{1n} + \alpha_2 x_{2n} + l_n \end{aligned} \qquad (3.4)$$

Якщо позначити

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix};\ X = \begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{pmatrix};\quad a = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix};\ l = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix},$$

де Y - матриця-стовпець спостережуваних значень показника, X - матриця спостережуваних значень факторів x_i, $i \in \{1,2\}$, x_0 - фіктивний фактор, всі значення якого дорівнюють одиниці, a – матриця стовпець оцінюваних параметрів, l - матриця-стовпець відхилень фактичних даних від розрахункових.

Рівняння (3.4) можна записати в матричній формі (3.3)

$$Y = Xa + l\,. \qquad (3.5)$$

Оскільки $l = Y - Xa$, то функція Q набуває вигляду

$$Q(a) = \sum_{i=1}^{n} l_i^2 = l^T l = Y^T Y - 2a^T X^T Y + a^T X^T Xa.$$

Функція Q в точці екстремуму задовольняє умову

$$\frac{\partial Q}{\partial a^T} = -2X^T Y + 2X^T Xa = 0.$$

Із попереднього рівняння отримаємо, що

$$X^T Y = X^T X a\,. \qquad (3.6)$$

Матриці $X^T X$ та $X^T Y$ обчисляються за формулами:

$$X^T X = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \end{pmatrix} \times \begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{pmatrix} =$$

$$= \begin{pmatrix} n & \sum\limits_{i=1}^{n} x_{1i} & \sum\limits_{i=1}^{n} x_{2i} \\ \sum\limits_{i=1}^{n} x_{1i} & \sum\limits_{i=1}^{n} x^2{}_{1i} & \sum\limits_{i=1}^{n} x_{1i} x_{2i} \\ \sum\limits_{i=1}^{n} x_{2i} & \sum\limits_{i=1}^{n} x_{2i} x_{1i} & \sum\limits_{i=1}^{n} x_{2i}{}^2 \end{pmatrix};$$

$$X^T Y = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} y_i \\ \sum\limits_{i=1}^{n} x_{1i} y_i \\ \sum\limits_{i=1}^{n} x_{2i} y_i \end{pmatrix}.$$

Із рівності (3.6) можемо знайти матрицю-стовпець коефіцієнтів множинної лінійної регресії; вона набуває вигляду:

$$a = \left(X^T X\right)^{-1} \cdot X^T Y\,. \qquad (3.7)$$

3.3 Коефіцієнт детермінації

Коефіцієнт множинної детермінації є оцінкою близькості математичної моделі до вибіркових даних і обчисляється за формулою

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} l_i}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{a^T X^T Y - n\bar{y}^2}{Y^T Y - n\bar{y}^2}.$$

3.4 Критерій Фішера для оцінки адекватності множинної регресії

Розрахункове значення критерію Фішера набуває вигляду

$$F = \frac{R^2(n-m-1)}{(1-R^2)m},$$

де $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, m - число факторів включених в регресію, n – число дослідів. Для заданої надійності γ та числа ступенів вільності $k_1 = m$, $k_2 = n - m - 1$ знаходиться табличне значення з таблиці розподілу Фішера F_{γ,k_1,k_2}. Якщо $F > F_{\gamma,k_1,k_2}$, то з надійністю γ можна вважати, що математична модель адекватна статистичним даним. В протилежному випадку з ймовірністю γ вона не адекватна статистичним даним тому її потрібно змінити на іншу.

3.5 Надійні інтервали базисних даних прогнозу та параметрів

Оскільки x_i, $i \in \{1,2\}$ - випадкова величина, то Y також є випадковою величиною. Для визначення надійного інтервалу необхідно знайти оцінку дисперсії випадкової величини Y. Оцінка дисперсії y_i для лінійної множинної регресії у матричній формі набуває вигляду

$$c.n.(\Delta\hat{y}_i) = \sqrt{S^2\left(x_i^T\left(X^T X\right)^{-1} x_i + 1\right)},\ i = \overline{1,n},$$

де $S^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-m-1} = \frac{l^T l}{n-m-1} = \frac{Y^T Y - a^T X^T Y}{n-m-1}$,

m - число факторів у регресії. Для регресії

$$Y = a_0 + a_1 x_1 + a_2 x_2$$

з надійністю γ можна вважати, що істинне значення задовольняє нерівність

$$\hat{y}_i - \Delta\hat{y}_i < y_i < \hat{y}_i + \Delta\hat{y}_i,$$

$$\Delta\hat{y}_i = c.п.(\hat{y}_i) \cdot t_{крит}, \qquad (3.8)$$

де $t_{крит}$ - значення, яке визначається із таблиць для заданої ймовірності γ та k ступенів вільності.

Якщо прогноз міститься у точці $x_p = (1, x_{1p}, x_{2p})$, то надійний інтервал для цієї точки набуває вигляду:

$$c.п.(\Delta\hat{y}_p) = \sqrt{S^2\left(x_p^T\left(X^T X\right)^{-1} x_p + 1\right)}, \qquad (3.9)$$

$$\hat{y}_p - \Delta\hat{y}_p < y_p < \hat{y}_p + \Delta\hat{y}_p$$

$$\Delta\hat{y}_p = c.п.(\hat{y}_p) \cdot t_{крит};$$

де $\hat{y}_p$ - прогнозне значення.

Надійні інтервали для параметрів набуває вигляду

$$a_i - \Delta a_i < \alpha_i < a_i + \Delta a_i, \qquad (3.10)$$

де $\Delta a_i = \sqrt{S^2 z_{ii}} \cdot t_{крит}$, z_{ii} - діагональний елемент матриці $Z = \left(X^T X\right)^{-1}$

3.6 Мультиколінеарність

Обговорення мультиколінеарності. При побудові структури регресії, з одного боку, потрібно включити в регресію всі фактори, які мають суттєвий статистичний вплив на показник, а з іншого боку, потрібно, щоб була виконана ***умова лінійної незалежності*** між факторами. Якщо існує лінійна залежність хоча б між двома факторами, то говорять, що між цими факторами існує ***мультиколінеарність***. Якщо між факторами x_i і x_j існує лінійна залежність $x_i = ax_j$, то говорять, що між цими факторами існує ***строга мультиколінеарність***.

Враховуючи, що x_i і x_j - випадкові величини то, як правило, між ними існує приблизна лінійна залежність $x_i = ax_j + \varsigma$, де ς - відхилення. В таких випадках ***мультиколінеарність*** між факторами ***нестрога***. Оскільки одним із основних припущень МНК є відсутність лінійної залежність між факторами, то постає питання: як позбутися мультиколінеарності. Якщо в регресії присутнє явище мультиколінеарності, то не виконується умова $\det\left(X^T X\right) \neq 0$, і при строгій мультиколінеарності неможливо отримати оцінки параметрів МНК. Якщо мультиколінеарність не строга, то отримані оцінки параметрів мало надійні. В цьому випадку незначні зміни вибіркових даних призводять до значних змін оцінок параметрів. Очевидно, що у випадку, коли $\det\left(X^T X\right)$близький до нуля, то в регресії присутнє явище мультиколінеарності. Взагалі при визначенні лінійної структури корисно будувати кореляційну матрицю, в яку включені фактори і показник, і в першу чергу у регресію потрібно включати фактори, які корегують з показником і не корегують між собою.

Якщо зв'язок між незалежними змінними не функціональний, проте статистично істотний, то незважаючи на те, що теоретично допустима оцінка параметрів методом найменших квадратів, вона може привести до таких помилок параметрів, що сама модель стає беззмістовною.

Тому, будуючи економетричну модель, яка повинна кількісно описати залежність економічних процесів та явищ, потрібно мати інформацію про те, що між незалежними змінними не існує колінеарності.

Ознаки мультиколінеарності. Якщо серед парних коефіцієнтів кореляції незалежних змінних є такі, рівень яких наближається або дорівнює множинному коефіцієнту кореляції, то це свідчить про можливість існування мультиколінеарності.

Інформацію про парну залежність може дати симетрична матриця коефіцієнтів парної кореляції, або кореляції нульового порядку:

$$r = \begin{matrix} r_{x1x1} & r_{x1x2} & r_{x1x3} & \dots & r_{x1xm} \\ r_{x2x1} & r_{x2x2} & r_{x2x3} & \dots & r_{x2xm} \\ r_{x3x1} & r_{x3x2} & r_{x3x3} & \dots & r_{x3xm} \\ \dots & \dots & \dots & \dots & \dots \\ r_{xmx1} & r_{xmx2} & r_{xmx3} & \dots & r_{xmxm} \end{matrix}$$

Але якщо в моделі фігурує більше двох незалежних змінних, то вивчення питання про мультиколінеарність не може обмежуватись інформацією, що дає ця матриця. Явище мультиколінеарності ні в якому разі не зводиться тільки до існування парної кореляції між незалежними змінними.

Більш загальна перевірка передбачає визначення визначники (детермінанта) матриці r, який називається детермінанте кореляції і позначається $|r|$. Числові значення детермінанта кореляції знаходяться на множині $|r| \in [0;1]$.

Якщо $|r|=0$, то існує повна мультиколінеарність, якщо $|r|=1$ - мультиколінеарність відсутня, чим ближче $|r|$ до нуля, тим повніше можна стверджувати, що між незалежними змінними існує мультиколінеарність.

Незважаючи на те, що на числове значення $|r||$ впливає дисперсія незалежних змінних, цей показник можна вважати точковою мірою тісноти мультиколінеарності.

Якщо в економетричній моделі одержано мале значення параметра як при високому рівні коефіцієнта детермінації і при цьому F-критерій суттєво відрізняється від нуля, то це також свідчить про наявність мультиколінеарності.

Якщо коефіцієнт детермінації R^2, що розрахований для регресійних залежностей між однією незалежною змінною і іншими, має значення, яке близьке до одиниці, то можна говорити про наявність мультиколінеарності.

Якщо при побудові економетричної моделі на основі покрокової регресії включення нової незалежної змінної, суттєво змінює оцінку параметрів моделі при незначно» підвищенні (або зниженні) коефіцієнта кореляції чи детермінації, то ця змінна, очевидно знаходиться в лінійній залежності від інших, які введені в модель раніше.

Всі ці методи виявлення мультиколінеарності мають один загальний недолік: ні один із них не проводить чіткої межі між тим, що треба вважати "суттєвою" мультиколінеарністю, яку треба враховувати, і тим, коли мультиколінеарністю можна знехтувати.

Існує багато методів дослідження мультиколінеарності. Основою цих методів є дослідження кореляційної матрці. Найбільш повне дослідження мультиколінеарності можні здійснити на основі застосування алгоритму Феррара-Глобера. В економетричних задачах для дослідження мультиколінеарності широко застосовується цей метод.

Він включає три види статистичних критеріїв, на основі яких перевіряється мультиколінеарність всього масиву незалежних змінних (χ^2, хі-квадрат), кожної незалежної змінної з всіма іншими (F-критерій) і мультиколінеарність кожної пари незалежних змінних (t-критерій).

Всі ці критерії при порівнянні з їх критичними значеннями дають можливість зробити конкретні висновки відносно наявності чи відсутності мультиколінеарності незалежних змінних.

Метод Феррара-Глобера. Для дослідження загальної мультиколінеарності і мултиколінеарності між окремими факторами використовується кореляційна матриця r і обернена до неї матриця K

$$r = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, \qquad K = r^{-1},$$

$$r_{ij} = \frac{Cov(x_i, x_j)}{\sqrt{Var(x_i)Var(x_j)}}, \quad i = \overline{1,n}, \quad j = \overline{1,n}.$$

Частинний коефіцієнт кореляції обчислюється за формулою

$$r_{ij1\ldots m} = -\frac{k_{ij}}{\sqrt{k_{ii}k_{jj}}}, \quad i = \overline{1,n}, \quad j = \overline{1,n},$$

де k_{ij} - елементи матриці K.

Для дослідження загальної мультиколінеарності використовується тест χ^2. Для цього треба обчислити визначник кореляційної матриці $|r|$ і знайти розрахункове значення

$$\chi^2 = -\left(n - 1 - \frac{2m+5}{6}\right)\ln\det|r|.$$

Для заданої довірчої ймовірності γ та $k = m(m-1)/2$ - що є числом ступенів вільності, треба знайти табличне значення ${\chi_{\gamma k}}^2$. Якщо $\chi^2 \le {\chi_{\gamma k}}^2$ то із заданою надійністю γ можна вважати, що загальна мультиколінеарність відсутня і на цьому закінчується дослідження мультиколінеарності. Якщо $\chi^2 > {\chi_{\gamma k}}^2$ то з прийнятою ймовірністю γ можна вважати, що між факторами існує мультиколінеарність. Для з'ясування питання, між якими факторами існує мультиколінеарність, використовується F- та t-статистика. t-статистика обчислюється

$$t_{ij} = \frac{r_{ij1...m}\sqrt{n-m-1}}{\sqrt{1-r^2_{ij1...m}}},\ i = \overline{1,n},\quad j = \overline{1,n}.$$

Для заданої ймовірності γ та числа ступешв вільності $k = n - m - 1$ знаходимо критичне значення статистики Стьюдента $t_{\gamma k}$ із таблиць. Якщо $t_{ij} > t_{\gamma k}$, то з надійністю γ можна стверджувати, що між факторами x_i і x_j існує мультиколінеарність.

Усунення мультиколінеарності. Для усунення мультиколінеарності існує декілька способів. Розглянемо найпростіші.

Перший спосіб. Якщо між двома факторами x_i і x_j існує мультиколінеарність, то один із факторів виключається з розгляду.

Другий спосіб полягає в заміні фактора x_i на різницю $x_i^* = x_j - x_i$, після чого слід перевірити на наявність мультиколінеарності між факторами x_i^* і x_j.

При наявності мультиколінеарності і значної кількості факторів використовують перший спосіб. При наявності мультиколінеарності і малої кількості факторів замість фактора x_i розглядається фактор x_i^*.

3.7 Коефіцієнт еластичності

Для множинної регресії коефіцієнти частинної еластичності визначаються за формулами

$$K_{xij} = \frac{\partial \widehat{Y}}{\partial x_i} \frac{x_{ij}}{\widehat{y}_j} \tag{3.11}$$

де i номер фактора регресії, x_{ij} - точка в якій визначається коефіцієнт еластичності.

Для лінійної множинної регресії коефіцієнт еластичності набуває вигляду

$$K_{xij} = a_j \frac{x_{ij}}{\widehat{y}_j} \tag{3.12}$$

Коефіцієнт еластичності показує на скільки відсотків зміниться показник $\widehat{y}_j$, якщо фактор x_{ij} зміниться на 1 % при незмінних значеннях інших факторів x_i ($i = \overline{1,m}$) при $i \neq j$.

4 МНОЖИННИЙ НЕЛІНІЙНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

4.1 Множинна нелінійна регресія. Основні поняття

Лише деякі природні та економічні процеси можна моделювати за допомогою лінійної моделі. Тому для більш точного й повного опису процесу, необхідно використовувати нелінійну регресійну залежність.

Як для парного так і для багатофакторного регресійного аналізу можна розглядати два типи моделей:

- нелінійна за факторами, але лінійна за невідомими параметрами;
- нелінійна за факторами та параметрами.

Багатофакторні регресійні моделі першого типу

$$Y = \alpha_0 + \alpha_1 F_1(x_1) + \alpha_2 F_2(x_2) + \ldots + \alpha_m F_m(x_m) + l, \tag{4.1}$$

де $F_i(x_i)$, $i = \overline{1,m}$ - будь-які функції.

Заміною змінних $z_i = F_i(x_i)$, $i = \overline{1,m}$, такі регресії зводяться до лінійної моделі

$$Y = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \ldots + \alpha_m z_m + l\ . \tag{4.2}$$

Нелінійність за параметрами є більш серйозною проблемою. Але коли права частина моделі складається з членів вигляду x^β, $e^{\beta x}$ помножених між собою, а випадковий член мультиплікативний до добутку цих функцій, то таку модель можна за допомогою логарифмування обох її частин звести до лінійної. Таким прикладом є степенева модель

$$Y = \alpha_0 x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \ldots \cdot x_m^{\alpha_m} \cdot l\ , \tag{4.3}$$

яка може описувати деякий виробничий процес, або степенево-показникова регресія (де m = 4).

$$Y = \alpha_0 \cdot \alpha_1^{x_1} \cdot \alpha_2^{x_2} \cdot \ldots \cdot \alpha_m^{x_m} \cdot l\ . \tag{4.4}$$

Логарифмуванням обох частин (4.3) та (4.4) можна звести до лінійної (4.2).

Досліджуючи регресії (4.3) та (4.4), спочатку треба знайти оцінки та надійні інтервали для лінійної регресії, а потім за допомогою зворотних перетворень, знайти оцінки та надійні інтервали і для регресій (4.3) та (4.4).

4.2 Приклади застосування нелінійних регресій в економіці

Аналіз індивідуального ринку.

Регресійне рівняння попиту залежно від ціни, на деякий вид товару, дає можливість знайти оцінку оптимальної ціни та прийняти оптимальне рішення. Нехай відома таблиця попиту

Pi	P_1	P_2	...	P_n
Di	D_1	D_2	...	D_n

Якщо між ***ціною*** P та величиною ***попиту*** D існує залежність

$$D = \alpha_0 + \alpha_1 P + \alpha_2 P^2 , \tag{4.5}$$

то оцінки параметрів та їх надійні інтервали знаходимо аналогічно лінійній множинній регресії.

Вплив еластичності попиту на ринкові оборот.

Нехай на ринку продається деякий однорідний товар. Введемо позначення: $P > 0$ - ціна одиниці цього товару;

$\widehat{D} = f(P) > 0$ - ***величина попиту на товар***, тобто кількість проданого товару, або регресія попиту; якщо P зростає то $f(P)$ спадає і $f'(P) < 0$.

$\widehat{Z} = P \cdot f(P)$ - ***товарообіг у грошовому виразі***, тобто кількість грошей, виручених за проданий товар;

$$\widehat{Z}'_p = \frac{dZ}{dP} \tag{4.6}$$

- ***граничний товарообіг*** (швидкість зміни товарообігу при фіксованій ціні P); якщо $\widehat{Z}'_p > 0$, то товарообіг зростає, якщо $\widehat{Z}'_p < 0$ – то спадає.

$K_D = \dfrac{Pf'(P)}{f(P)}$ - ***коефіцієнт еластичності попиту*** за фіксованою ціною P. Ясно, що $K_D < 0$, бо $P > 0$, $f(P) > 0$, а $f'(P) < 0$.

Треба дослідити граничний товарообіг залежно від коефіцієнта еластичності. За означенням

$$\widehat{Z} = Pf(P), \tag{4.7}$$

тому

$$\widehat{Z}'_p = f(P) + Pf'(P) = f(P)\,(1 + K_D\,). \tag{4.8}$$

Можливі три випадки: $\widehat{Z}'_p > 0$, $\widehat{Z}'_p < 0$, $\widehat{Z}'_p(P) = 0$, які розглянемо окремо.

1. Нехай $\widehat{Z}'_p >$ 0. тоді добуток $f(P)\,(1 + K_D\,) >$ 0, де $f(P) > 0$.

Тому $1 + K_D >$ 0 і, отже, $K_D >$ -1. Відомо що $K_D <$ 0, тому вірними є нерівності: $-1 < K_D <$ 0. в економіці попит з таким коефіцієнтом еластичності називається ***нееластичним***.

Економічне тлумачення. Зміна ціна на 1% викликає зміну попиту у зворотному напрямку на $|K_D|$ %, де $-1 < K_D <$ 0, при цьому товарообіг у грошовому виразі зростає. Таке явище не є природним, тому попит не є еластичний.

2. Якщо $\widehat{Z}'_p <$ 0, то з підвищенням ціни на товар товарообіг спадає у грошовому виразі. Оскільки ($f(P)\,(1 + K_D\,)$)< 0, а $f(P) > 0$, то $1 + K_D < 0$, звідси випливає, що $K_D < -1$. Попит при цьому є ***еластичний***.

Економічне тлумачення. Зміна ціни на 1% викликає зміну попиту в зворотному напрямку на $|K_D|$ %, в цьому випадку $K_D <$ -1. Тобто із зростанням ціни, спадає. Це явище в економіці є природним і тому попит є еластичним.

3. Якщо $\widehat{Z}'_p(P) =$ 0, то товарообіг залишається ***сталим*** для деякого проміжку цін. Тоді $K_D = 1$. Якщо $\widehat{Z}'_p(P) = 0$ лише в одній точці P_0, то ці точка називається ***критичною***. Якщо при переході через цю точку похідна $\widehat{Z}'_p$ змінює знак з додатного на від'ємний то при ціні P_0 товарообіг у грошовому виразі є ***максимальним***.

Визначення максимального прибутку.

Прибуток підприємства визначається за формулою

$$\widehat{F} = \widehat{D}P - (C + V\widehat{D}), \tag{4.9}$$

де C – *сталі затрати*, $V\widehat{D}$ – *змінні витрати*.

Знайдемо точку P_0 в якій прибуток $\widehat{F}$ досягає екстремуму. З необхідної умови екстремуму випливає, що

$$\frac{d\widehat{F}}{dP} = \frac{d}{dP}\,[a_2P^3 + (a_1 - Va_2)P^2 + (a_0 + Va_1)P - C - Va_0] = 0$$

або

$$3a_2P^2 + 2(a_1 - Va_2)P + (a_0 + Va_1) = 0,$$

а звідси отримуємо, що

$$P_{3,4} = Va_2 - a_1 \pm \frac{\frac{1}{2}\sqrt{D_0}}{3a_2},$$

де $D_0 = 4((a_1 - Va_2)^2 + 3a_2(Va_1 - a_0))$.

Припустимо, що в точці P_4 функція $\frac{dF}{dP}$ змінює знак з (+) на (-), тоді в цій точці прибуток є максимальним.

Оптимальну кількість продукції можна визначити за такою формулою

$$\widehat{D}_1 = a_0 + a_1P_3 + a_2P_4{}^2,$$

а максимальний прибуток

$$F(P_4) = Z(P_4) - Vf(P_4) =$$
$$= a_2P^3{}_4 - (a_1 - Va_2)P^2{}_4 + (a_0 + Va_1)p_4 - C - Va_0. \tag{4.10}$$

Оцінка параметрів кривої Лаффера.

Залежність податкових надходжень від податкових ставок є функцією, графік якої називається кривою Лаффера.

Нехай задано динамічний ряд з даними за певний період про розміри податкових ставок X і відповідними їм податкові надходження Y. Будемо вважати, що регресія є показниковою з невідомими параметрами a, b, c:

$$\widehat{Y} = ae^{b(x-c)^2}, \tag{4.11}$$

де Y – відомі податкові надходження, X – відома податкова ставка. Зведемо (4.11) до лінійної множинної регресії. Після логарифмування

$$\ln\widehat{Y} = \ln a + bx^2 - 2bcx + bc^2.$$

і заміни $\widehat{Y_1} = \ln \widehat{Y}$, $b_0 = \ln a + bc^2$, $b_1 = -2bc$, $b_2 = b$, $z_1 = x$, $z_2 = x^2$.

Регресія набуде вигляду:

$$\widehat{Y_1} = b_0 + b_1 z_1 + b_2 z_2 . \qquad (4.12)$$

Отримавши оцінку параметрів b_0, b_1, b_2, Залежності (4.12) знаходимо оцінки параметрів рівняння (4.11)

Якщо $b < 0$, то $a = e^{b_0 - \frac{b_1}{4b_2}}$ – ***максимальне надходження податків*** при податковій ставці $x = c$.

Виробнича регресія.

Обсяг виробництва Y Залежить від двох типових факторів – затрат на працю x_1 і від основних засобів (капіталу) даної галузі x_2

$$Y = F(x_1, x_2) . \qquad (4.13)$$

Для такої залежності використовують регресію типу ***Кобба-Дугласа***

$$\widehat{Y} = \alpha_0 x_1^{\alpha 1} x_2^{\alpha 2}. \qquad (4.14)$$

За допомогою логарифмування (4.14) можна звести до такої множинної регресії

$$\ln \widehat{Y} = \ln \alpha_0 + \alpha_1 \ln x_1 + \alpha_2 \ln x_2 ,$$

і заміни $\widehat{Y_1} = \ln \widehat{Y}$, $a_0 = \ln \alpha_0$, $z_1 = \ln x_1$, $z_2 = \ln x_2$, отримуємо лінійну модель

$$\widehat{Y_1} = a_0 + \alpha_1 z_1 + \alpha_2 z_2 . \qquad (4.15)$$

Для деяких підприємств виконується приблизна рівність $\alpha_1 + \alpha_2 \approx 1$. В цьому випадку регресія Кобба-Дугласа набуватиме вигляд:

$$\widehat{Y} = \alpha_0 \left(\frac{x_1}{x_2} \right)^{\alpha 1} x_2 \qquad (4.16)$$

Після заміни $\widehat{Y_1} = \dfrac{\widehat{Y}}{x_2}, z = \dfrac{x_1}{x_2}$ отримаємо

$$\widehat{Y_1} = \alpha_0 z^{\alpha 1} \qquad (4.17)$$

Оцінивши параметри (4.15) знайдемо частинні коефіцієнти еластичності для виробничої регресії (4.14):

$$K_{xi} = \frac{\partial f}{\partial x_1}\frac{x_1}{f} = \frac{\partial(\alpha_0 x_1^{\alpha 1} x_2^{\alpha 2}}{\partial x_1}\frac{x_1}{\alpha_0 x_1^{\alpha 1} x_2^{\alpha 2}} \cong$$
$$\cong \frac{\alpha_0 \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} x_1}{\alpha_0 x_1^{\alpha 1} x_2^{\alpha 2}}, \qquad (4.18)$$

Отже K_{xi} (4.18) є частинним коефіцієнтом еластичності фактора x_i. Він показує, що показник Y Змінюється на K_{xi} % при зміні фактора x_i на 1% при незмінних значеннях інших факторів.

4.3 Сумарний коефіцієнт еластичності

Збільшимо обсяг факторів у будь-яке число разів λ і прослідкуємо зміну обсягу випуску продукції.

Нехай у деякий момент часу фактори й показники мали значення Y_0, x_{10} , x_{20}, тобто $Y_0 = \alpha_0 x_{10}{}^{\alpha_1} x_{20}{}^{\alpha_2}$. Після збільшення факторів у λ разів отримаємо

$$Y = \lambda^{\alpha 1 + \alpha 2} \alpha_0 x_{10}^{\alpha 1} x_{20}^{\alpha 2} = \lambda^{\alpha 1 + \alpha 2} Y_0 .$$

У даному випадку бачимо, що обсяг продукції збільшився в $\alpha_1 + \alpha_2$ разів, якщо $\lambda > 1$. Тому цей показник $\alpha_1 + \alpha_2$ можна назвати ***сумарним коефіцієнтом еластичності***. На основі цих елементарних розрахунків можна зробити висновки:

1. якщо $\alpha_1 + \alpha_2 = 1$, то при збільшені факторів в λ разів ($\lambda > 1$), *обсяг виробництва збільшиться* в стільки ж разів.

2. якщо $\alpha_1 + \alpha_2 > 1$ і $\lambda > 1$, то в даному випадку *маємо економію ресурсів на масштабах виробництва*.

3. якщо $\alpha_1 + \alpha_2 < 1$, то збільшення факторів в $\lambda > 1$ разів викличе *зменшення обсягу виробництва* в число разів менше за λ тобто в $\lambda^{\alpha 1 + \alpha 2}$, де $\alpha_1 + \alpha_2 < 1$. У цьому випадку при зростанні обсягу виробництва *зростають витрати на одиницю продукції*.

4.4 Темп приросту показника виробничої регресії для двох факторів

Якщо x_1, x_2 та Y залежить неявно від часу $(x_1(t), x_2(t))$, то приріст величин у момент часу t можна охарактеризувати у вигляді $\frac{\partial Y}{\partial t}, \frac{\partial x_1}{\partial t}, \frac{\partial x_2}{\partial t}$, а темпи приросту, відповідно:

$$\varepsilon_y = \frac{\partial Y}{\partial t}\frac{1}{Y}, \;\; \varepsilon_{x1} = \frac{\partial x_1}{\partial t}\frac{1}{x_1}, \;\; \varepsilon_{x2} = \frac{\partial x_2}{\partial t}\frac{1}{x_2}.$$

Якщо враховувати, що $\widehat{Y} = f(x_1(t), x_2(t))$ складна функція від t, то

$$\frac{\partial Y}{\partial t} = \frac{\partial Y}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial Y}{\partial x_2}\frac{\partial x_2}{\partial t}. \tag{4.19}$$

Для виробничої регресії частині похідні по x_1, x_2 мають значення:

$$\frac{\partial Y}{\partial x_1} = \alpha_0 \alpha_1 x_1^{\alpha 1-1} x_2^{\alpha 2}; \quad \frac{\partial Y}{\partial x_2} = \alpha_0 \alpha_2 x_{x1}^{\alpha 1} x_2^{\alpha 2-1}. \tag{4.20}$$

Після підстановки (4.20) в (4.19) і ділення Y отримуємо

$$\frac{\partial Y}{\partial t}\frac{1}{Y} = \alpha_1 \frac{\partial x_1}{\partial t}\frac{1}{x_1} + \alpha_2 \frac{\partial x_2}{\partial t}\frac{1}{x_2}.$$

Одержуємо

$$\varepsilon_y = \alpha_1 \varepsilon_{x1} + \alpha_2 \varepsilon_{x2}.$$

Коли в виробничу регресію включають залежність від науково-технологічного прогресу, який проявляється в удосконаленні техніки й технології, підвищення кваліфікації робітників та ін. Технічний прогрес у виробничих регресіях залежить від часу. Таку залежність можна задати у формі множника $e^{\lambda t}$, де λ – показник темпу розвитку виробництва, пов’язаного з технічним прогресом. Отже виробнича регресія набуде вигляду

$$\widehat{Y} = \alpha_0 x_1^{\alpha 1} x_2^{\alpha 2} e^{\lambda t}.$$

Темп приросту виробництва в цьому разі визначається за формулою $\varepsilon_y = \alpha_1 \varepsilon_{x1} + \alpha_2 \varepsilon_{x2} + \lambda$.

5 АВТОКОРЕЛЯЦІЯ

5.1 Основні положення теми

Розглянемо економетричну модель виду:

$$Y = AX + u,$$

де Y - вектор залежної змінної; X - матриця незалежних змінних; A - вектор параметрів моделі; u - вектор помилок або залишків.

Однією з основних передумов, на якій базується оцінка параметрів моделі МНК є відсутність коваріації залишків. Але дуже часто приходиться мати справу з ситуаціями, коли ця передумова не виконується і залишки u_t автокорельовані з залишками u_{t-1}.

Автокореляція залишків виникає частіше всього тоді, коли економетрична модель будується на основі часових рядів. Вона також може бути наслідком помилкової специфікації форми залежності між змінними, тобто вибраний вид стохастичного рівняння не адекватний до статистичних даних. Крім того, наявність автокореляції залишків може вказувати на необхідність введення в модель нової незалежної змінної, яка має важливе значення при дослідженні економічного явища, або на значні помилки при отриманні статистичних даних і т.п.

5.2 Критерії перевірки автокореляції

Для перевірки наявності автокореляції залишків частіше всього застосовується критерій ***Дарбіна-Уотсона DW***:

$$DW = \frac{\sum_{t=2}^{n}(u_t - u_{t-1})^2}{\sum_{t=1}^{n}(u_t)^2}.$$

Він може приймати значення: $DW \in [0, 4]$.

Якщо залишки u_t є випадковими, тобто не автокорельовані величини, то значення DW=2, або знаходяться поблизу 2.

При додатній автокореляції DW<2, при від'ємній DW>2. Фактичні значення критерію DW порівнюються з критичними (табличними) для різного числа спостережень n і числа незалежних змінних m при вибраному рівні значущості а. Табличні значення мають нижню межу DW_1 і верхню - DW_2.

Якщо $DW_{факт}$.< DW_1, то залишки мають автокореляцію. Якщо $DW_{факт}$.> DW_2, то приймається гіпотеза про відсутність автокореляції. Якщо DW_1< $DW_{факт}$ < DW_2, то конкретних висновків зробити не можна: необхідно проводити подальші дослідження, беручи більшу сукупність спостережень.

Коли розрахункове значення DW_1< $DW_{факт}$ < DW_2.>2, то з DW_1 та з DW_2 порівнюється не сам коефіцієнт $DW_{факт}$, а вираз (4- $DW_{факт}$).

Для виявлення автокореляції залишків використовуються також інші критерії, наприклад, критерій фон-Неймана.

Оцінка параметрів моделі з автокорельованими залишками.

Автокореляція залишків є важливою проблемою при виборі методу оцінки і верифікації економетричної моделі.

Вона приводить до неефективності оцінок. А - це означає, що стандартні помилки параметрів можуть бути дуже великими Якщо матриця X містить запізнюючі значення залежної змінної, то оцінки, що одержані МНК, можуть бути ще й зміщеними і необгрунтованими.

Для оцінки параметрів економетричної моделі, що має автокореляцію залишків можна застосувати узагальнений метод найменших квадратів (метод Ейткена), який базується на скоригованій вихідній інформації з урахуванням коваріації залишків.

6 ГЕТЕРОСКЕДАСТИЧНІСТЬ

6.1 Основні положення теми

Гетероскедастичність – це явище, коли випадкові величини мають різну дисперсію (неоднаковий розкид статистичних даних). Термін походить від грецького слова «гетеро» (інший) та «skedasis» (дисперсії).

Відповідно, послідовність випадкових величин називається гомоскедастичною, якщо вона має постійну дисперсію.

Наявність гетероскедастичності порушує одну з умов Гауса-Маркова про те, що залишковий член має незмінну дисперсію. Тому оцінки, отримані МНК можуть бути неефективними.

Вплив гетероскедастичності на оцінку інтервалу прогнозування і перевірку гіпотези полягає в тому, що хоча коефіцієнти і не зміщені, дисперсії і, відповідно, стандартні похибки цих оцінок будуть зміщеними. Гетероскедастичність може виявлятися при аналізі часових рядівб та якщо спостереження x та y збільшуються з часом, то може статися, що і дисперсія випадкового члена буде зростати з часом.

Гетероскедастичність у деяких випадках може бути усунена за допомогою переосмислення рівняння регресії.

У 2003 р. Роберт Енгл отримав Нобелівську премію з економіки за дослідження з регресійного аналізу в присутності гетероскедастичності, що призвело до розробки ним техніки моделювання авторегресійної умовної гетероскедастичності (ARCH).

6.2 Перевірка наявності гетероскедастичності

Існує кілька способів, щоб перевірити на наявність гетероскедастичності, серед яких: тест Бройша-Паґана, тест Уайта, тест Гольфельда-Квандта, тест Глейзера, тест Спірмена та інші.

Тест Бройша-Паґана використовується для перевірки гетероскедастичності в лінійній регресійній моделі. Він перевіряє, чи оціночна дисперсія залишків з регресії залежить від значень незалежних змінних.

Припустимо, що ми оцінюємо рівняння $y = \alpha + \beta x + u$.

Створимо регресію квадратів залишків на незалежні змінні, що і є Бройш-Паган тестом: $\hat{u}^2 = \alpha_1 + \beta_1 x + v$.

Якщо *F*-тест підтверджує, що незалежні змінні спільно значнимі, то ми можемо відхилити нульову гіпотезу про наявність гомоскедастичності.

Це χ^2-квадрат тест з *k* ступенями свободи. Якщо Бройш-Паган тест показує, що існує умовна гетероскедастичність, це може бути виправлено за допомогою методу Хансен.

Тест рангової кореляції Спірмена оцінює наскільки добре можна описати відношення між двома змінними за допомогою монотонної функції. Ранг, у даному випадку, - це порядковий номер значення деякої величини, впорядкованої у зростаючому порядку. Коефіцієнт кореляції рангу Спірмена обчислюється за формулою

$$r_{x,u} = 1 - \frac{6\sum_{t=1}^{n} D_t^2}{n(n^2-1)}, ,$$

де $D_t = N_{xt} - N_{ut}$; N_{xt} - ранг x_t; N_{ut} - ранг u_t.

Якщо $r_{x,u} \geq r_{kr}$, то можна відхилити гіпотезу про гомоскедастічность.

Якщо в моделі регресії є більше однієї пояснювальної змінної, то перевірка гіпотези може виконуватися з використанням будь-якої з них. Можна використати *t*-статистику, тоді:

$$t = \frac{r_{x,u}\sqrt{n-k-1}}{\sqrt{1-r^2}}.$$

Якщо $t > t_{kr}$, то це підтверджує гіпотезу про наявність гетероскедастичності. Якщо t

Тест Гольфельда-Квандта застосовується у випадку, коли є підстави вважати, що стандартне відхилення помилок може бути пропорційно деякої змінної. Тест також ґрунтується на припущенні нормальності розподілу випадкових помилок регресійної моделі. Фактично це *F*-тест, оскільки статистика тесту має розподіл Фішера.

В першу чергу, дані упорядковуються за спаданням незалежної змінної, відносно якої є підозри на гетероскедастичність.

Далі звичайним МНК оцінюється регресійна модель для двох різних вибірок – перших і останніх m спостережень в даному впорядкування, де $m<n/2$. Середні $(n\text{-}2m)$ спостережень виключаються з розгляду. Якщо n=30, то Доугерті рекомендує брати m=11, а для n=60 – m=22. Також рекомендують обсяг виключаються середніх спостережень – близько треті або чверті загального обсягу вибірки. Цей тест працює і без виключення середніх спостережень, але в цьому випадку потужність тесту менше.

Для отриманих двох оцінок регресійної моделі знаходять суми квадратів залишків і розраховують GK-статистику, рівну відношенню більшої суми квадратів залишків до меншої, яку порівнюють з табличним значенням $F_{kr}(m-k-1, m-k-1)$. Якщо $GK \geq F_{kr}$, то можна відхилити гіпотезу про гомоскедастічность.

Тест Глейзера дозволяє дещо більш ретельно розглянути характер гетероскедастичності. Знімається припущення про те, що залишки пропорційні одній з незалежних змінних, і перевіряється, чи може бути більш підходяща інша функціональна залежність.

Тест Уайта регресує квадрат залишку від регресійної моделі на регресори, крос-добуток регресорів і квадрат регресорів. Потім слід перевірити R^2. Якщо гомоскедастічность відхилена, можна використовувати моделі GARCH.

Цей тест, і оцінка для гетероскедастичності-узгоджених стандартних помилок, були запропоновані Халбертом Уайтом в 1980 році. Щоб порахувати LM статистику для тесту, слід перемножити значення R^2 і розмір вибірки: $LM = nR^2$. Ця статистика розподілена згідно з χ^2 розподілом з k ступенями свободи, що дорівнюють числу оцінюваних параметрів в допоміжній регресії мінус один.

Узагальнений МНК застосовується до перетворених даних і дозволяє отримати оцінки, які мають не тільки властивість незміщеності, але й мають найменші вибіркові дисперсії. На відміну від звичайного МНК, узагальнений метод враховує інформацію про неоднаковість дисперсії.

7 ЗАВДАННЯ

7.1 Завдання №1

На основі статистичних даних *Y* та *X* (таблиця 6.1):

1. Побудувати усі 5 вказаних нижче моделей парних нелінійних регресій, знайти для всіх оцінки параметрів.

2. Зробити аналіз нелінійних регресій за:
   - коефіцієнтом кореляції;
   - коефіцієнтом детермінації; за.
   - коефіцієнтом еластичності.
   -критерієм Фішера з надійністю p=0.95, оцінити адекватність моделі статистичним даним

3. Знайти довірчі інтервали нелінійних регресій.

4. Знайти прогнози і його надійні інтервали.

5. Побудувати графіки для усіх моделей:
   - за фактичними даними і прогнозом;
   - лінії регресії та її довірчий інтервал;
   - лінії коефіцієнта еластичності.

6. Вияснити яка з моделей найкраще відповідає вхідним даним. Зробити висновок.

1 $y = b \cdot 10^{\frac{\alpha}{x}}$; 2 $y = \dfrac{x}{\alpha x + b}$; 3 $y = \dfrac{1}{\alpha e^{-x} + b}$; 4 $y = b\alpha^{x}$; 5 $y = bx^{\alpha}$.

Примітка. Розрахунки рекомендовано виконувати, застосовуючи табличний редактор *Excel*.

Таблиця 7.1

1.	X	1.03	1.59	2.12	2.61	3.05	3.56	4.00	4.50	5.03	5.56
	Y	3.02	2.94	2.71	3.32	2.69	2.99	3.22	3.42	4.00	3.43
2.	X	1.10	1.33	1.58	1.81	2.09	2.32	2.59	2.85	3.14	3.43
	Y	2.27	2.91	3.18	3.50	3.71	3.88	4.06	4.18	4.39	4.44
3.	X	8.68	9.86	10.65	11.2	11.87	12.32	12.59	12.89	13.18	13.40
	Y	0.21	0.69	1.03	1.38	1.71	2.07	2.47	2.92	3.38	3.87
4.	X	0.38	0.93	1.61	2.35	30.2	3.36	4.35	5.02	5.72	2.65
	Y	48.33	7.97	4.60	4.02	3.75	3.72	3.33	3.27	2.67	3.28
5.	X	0.11	0.39	0.66	0.89	1.15	1.43	1.67	1.95	2.33	2.46
	Y	1.78	2.22	4.30	5.49	6.57	7.15	10.48	12.52	17.53	24.23
6.	X	0.21	0.69	1.03	1.38	1.71	2.07	2.47	2.92	3.38	3.87
	Y	0.19	0.47	0.67	0.66	0.92	1.00	1.15	1.39	1.48	1.72
7.	X	1.03	1.63	2.16	2.71	3.26	3.77	4.35	4.91	5.50	6.01
	Y	1.27	1.42	1.93	2.36	2.73	3.93	5.12	6.55	9.05	12.24
8.	X	0.32	0.93	1.59	2.29	2.92	3.62	4.30	4.99	5.63	6.25
	Y	2.44	3.62	4.98	5.43	5.59	5.94	6.67	9.62	11.39	12.60
9.	X	1.08	1.53	2.05	2.58	3.02	3.58	4.06	4.56	5.01	5.51
	Y	4.37	4.01	3.29	3.10	3.22	2.99	2.90	2.37	1.87	1.82
10	X	1.10	1.55	2.09	2.52	3.07	3.57	4.06	4.56	5.01	5.53
	Y	1.84	2.22	3.54	3.50	5.30	5.42	7.12	8.27	8.77	10.09

7.2 Завдання № 2

На основі статистичних даних, які наведені в таблиці 5.2, де P – ціна в у.г.о. за одиницю товару, D – кількість товару, оцінити параметри регресії, якщо математична модель

$$D = \alpha_0 + \alpha_1 P + \alpha_2 P^2 .$$

За критерієм Фішера з надійністю p=0,95 оцінити адекватність моделі статистичним даним. Якщо модель адекватна статистичним даним , знайти :

1. коефіцієнт еластичності для всіх цін;
2. проміжки зростання та спадання товарообігу а грошових одиницях;
3. ціну на товар, при якій товарообіг буде максимальним, а також оцінку цін на товар, де прибуток буде максимальним та значення його;
4. побудувати графіки:
- статистичних даних;
- лінії регресії;
- лінії еластичності.

Таблиця 7.2

№	1	2	3	4	5	6	7	8	9	10
P	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
1	8.15	8.71	7.76	10.00	8.33	7.60	9.04	9.18	8.39	8.05
2	7.24	7.64	6.83	8.12	8.96	6.70	7.24	7.79	7.37	6.66
3	6.31	6.90	6.02	6.68	8.06	5.79	6.55	7.20	6.60	6.14
4	6.24	6.28	6.06	7.70	6.34	6.08	6.88	6.52	6.44	5.21
5	5.47	6.28	5.02	5.64	7.09	5.37	6.15	5.74	5.50	4.85
6	4.53	4.55	4.18	6.33	5.64	3.71	4.55	4.61	4.75	3.91
7	3.67	3.94	3.67	3.93	4.85	2.68	3.97	3.83	3.90	3.21
8	3.08	3.30	2.85	4.46	4.06	2.36	3.58	3.38	3.26	2.17
9	2.44	3.23	2.12	2.64	2.94	2.13	3.08	3.23	2.51	1.92
10	1.81	2.15	1.77	2.26	2.90	1.48	1.93	2.33	2.03	1.35
11	1.45	1.83	1.19	1.57	1.65	1.18	2.14	1.55	1.53	0.45

7.3 Завдання №3

На основі статистичних даних із таблиці 6.3 показника Y та факторів x_1 та x_2 знайти оцінку параметрів регресії, якщо припустити, що залежність між факторами x_1, x_2 та показником Y набуває вигляду:

$$Y = F(a_0, a_1, a_2, x_1, x_2)\ .$$

З надійністю p=0,95, перевірити на адекватність модель, використовуючи критерій Фішера. Якщо модель адекватна статичним даним, то знайти:

1. коефіцієнт детермінації,
2. оцінки прогнозу та з надійністю p=0,95 його надійний інтервал;
3. Оцінки статичних коефіцієнтів еластичності і для прогнозу;
4. перевірити наявність автокореляції залишків;
5. перевірити наявність мультиколінеарності,
6. перевірити наявність гетероскедастичності за допомогою двох тестів, порівняти результати,
6. зробити висновок.

7.4 Завдання №4

Для виробничої регресії Кобба-Дугласа

$$\widehat{Y} = \alpha_0 x_1^{\alpha 1} x_2^{\alpha 2}.$$

На основі статистичних даних з таблиці 6.4 оцінити параметри регресії за МНК. З надійністю p=0,95 встановити адекватність моделі статистичним даним (критерії Фішера). Якщо модель адекватна своїм даним, то знайти:

1. коефіцієнт детермінації,
2. частині коефіцієнти еластичності;
3. побудувати ізокванту Y=Y_z (Y_z – вибрати самостійно);
4. прогноз в точці (x_{1p},x_{2p}),
5. перевірити наявність автокореляції залишків;
6. зробити висновок.

Таблиця 7.3

1	$Y = \exp(a_0 + a_1 x_1 + a_2 x_2)$	x_1	0.43	0.78	0.91	1.03	1.29	1.30	1.62	1.70	1.71
		x_2	5.12	5.09	4.70	4.23	3.87	3.43	3.26	2.85	2.59
		Y	21.22	23.91	21.78	17.83	17.36	12.70	14.63	11.06	9.82
2	$Y = \exp(a_0 + a_1 \ln x_1 + {}$ $+ a_2 x_2)$	x_1	5.12	5.36	7.15	8.73	9.96	10.66	11.79	12.08	12.17
		x_2	3.02	2.89	2.86	2.76	2.61	2.41	2.36	2.32	2.15
		Y	31.39	31.68	43.69	54.43	59.50	59.16	66.65	67.91	62.26
3	$Y = \exp(a_0 + a_1 / x_1 + {}$ $+ a_2 \ln x_2)$	x_1	0.51	0.63	0.88	1.11	1.25	1.36	1.50	1.53	1.78
		x_2	24.15	22.91	22.77	20.78	19.93	19.50	19.38	18.72	18.23
		Y	972.26	406.1	150.32	87.03	65.67	56.43	48.17	45.75	38.23
4	$Y = \exp(a_0 + a_1 \ln x_1 + {}$ $+ a_2 / x_2)$	x_1	15.78	15.26	13.43	13.07	12.13	11.52	11.35	10.22	8.26
		x_2	0.49	0.81	1.14	1.30	1.55	1.65	1.73	2.11	2.31
		Y	41.11	42.82	40.16	39.65	37.51	37.45	36.38	36.96	31.18
5	$Y = \exp(a_0 + a_1 / x_1 + {}$ $+ a_2 x_2)$	x_1	5.12	5.70	6.50	7.24	7.78	8.15	8.41	9.26	9.66
		x_2	12.11	11.86	11.80	11.71	11.23	10.88	10.30	9.57	9.56
		Y	11.78	9.41	8.74	9.06	10.22	10.31	8.62	8.67	9.01

6	$Y = a_0 + a_1 / x_1 + a_2 x_2$	x_1	1.12	1.39	1.53	1.96	2.20	2.32	2.76	2.97	3.19
		x_2	13.13	15.48	15.58	17.26	17.44	19.72	21.51	23.50	24.93
		Y	49.48	54.28	54.24	58.46	59.08	64.84	70.09	76.45	80.95
7	$Y = a_0 + a_1 \ln x_1 + a_2 \ln x_2$	x_1	50.76	58.53	97.01	126.17	143.64	152.62	189.46	198.89	231.0
		x_2	52.07	57.44	73.24	74.97	83.16	88.94	132.44	155.39	193.8
		Y	16.90	18.12	18.98	19.57	19.81	20.61	21.19	21.82	22.3
8	$Y = a_0 + a_1 \ln x_1 + a_2 x_2$	x_1	52.06	73.80	73.85	84.28	98.84	103.30	105.39	108.43	114.5
		x_2	25.17	27.33	26.17	27.39	27.17	27.20	28.17	30.47	29.17
		Y	68.26	73.07	70.62	73.88	74.16	74.75	75.59	80.15	79.34
9	$Y = a_0 + a_1 / x_1 + a_2 \ln x_2$	x_1	3.52	4.676	5.507	7.729	7.995	10.07	12.98	14.84	17.83
		x_2	52.06	53.10	53.62	55.07	57.63	58.86	59.28	62.22	65.95
		Y	20.99	19.23	19.53	18.10	20.16	19.92	19.27	18.28	18.53
10	$Y=\sqrt{\alpha_0 + \alpha_1 \ln(x_1)}$	x_1	52.06	52.63	54.10	57.86	58.47	61.71	66.08	68.23	72.11
		x_2	21.06	22.85	24.73	29.00	31.07	35.48	35.90	39.93	41.57
		Y	5.625	5.509	5.645	5.812	5.997	6.092	6.11	6.463	6.356

Продовження таблиці 7.3

1	x_1	21.1	23.6	24.4	24.8	27.0	28.6	31.0	33.1	33.7	34.9	35.2	36.4	37.2	39.3
	x_2	52.1	57.0	60.7	65.7	69.9	74.6	77.8	78.0	83.0	86.9	89.2	94.6	97.0	100.4
	Y	69.5	75.9	79.9	84.6	89.0	95.6	99.8	103.1	107.5	111.9	114.5	120.9	122.8	?
2	x_1	30.1	32.6	33.7	35.1	36.4	39.4	41.8	43.3	44.2	46.0	47.8	49.5	49.7	49.9
	x_2	52.1	53.5	53.1	56.5	54.1	58.2	55.1	57.2	56.1	56.1	57.1	58.7	58.1	58.1
	Y	78.2	82.4	83.8	86.7	87.0	92.8	91.5	95.3	94.7	96.7	99.5	102.9	102.6	?
3	x_1	26.9	28.7	29.8	31.5	34.4	35.5	35.9	37.3	40.1	42.7	44.6	44.8	46.5	47.2
	x_2	52.1	56.7	60.9	62.9	67.2	67.5	72.7	76.9	81.0	84.4	85.8	89.8	93.7	95.9
	Y	75.2	81.4	84.9	90	96.3	97.2	102.2	107.1	114.0	118.5	123.1	126.2	131.7	?
4	x_1	23.8	23.8	24.2	24.8	24.9	25.9	26.7	27.8	30.3	31.5	33.4	35.6	37.7	39.8
	x_2	42.7	46.6	49.3	51.8	53.3	55.8	58.8	59.0	64.4	65.7	68.8	69.2	74.4	74.7
	Y	63.4	67.2	70.9	72.2	74.2	76.2	81.2	81.4	88.8	92.7	96.5	99.1	108.5	?
5	x_1	43.1	45.2	45.5	47.8	49.5	50.3	53.3	53.7	53.7	55.5	57.4	59.2	61.2	61.8
	x_2	32.4	34.8	38.0	40.3	45.7	48.2	53.3	55.4	55.5	56.7	62.2	65.9	68.4	73.0
	Y	68.7	73.3	77.6	80.1	88.1	91.0	100.4	102.2	101.1	103.9	112.1	118.3	121.9	?

6	x_1	46.0	46.6	48.7	49.2	49.3	51.7	54.3	54.7	55.7	58.6	61.5	64.1	67.1	67.3
	x_2	32.8	34.3	38.9	42.8	43.3	48.4	48.7	52.5	56.9	59.7	62.1	66.3	66.9	69.2
	Y	69.9	72.4	80.3	85.4	85.1	92.3	94.9	99.4	104.1	110.2	115.8	121.6	127.7	?
7	x_1	32.1	34.3	36.3	38.4	40.8	42.1	43.5	45.8	48.7	51.8	53.7	54.4	55.7	58.5
	x_2	44.1	45.4	46.4	50.9	55.5	57.9	63.4	63.8	66.5	66.9	68.3	71.5	76.5	80.5
	Y	74.2	76.2	78.5	85.9	91.0	95.8	102.4	104.4	108.3	111.3	114.6	119.4	124.6	?
8	x_1	33.1	34.2	36.3	38.4	41.0	43.7	44.3	46.0	47.0	49.3	50.1	51.8	53.9	56.8
	x_2	25.1	25.4	30.2	32.7	35.7	41.1	43.5	47.9	53.3	58.0	62.2	64.2	69.7	69.9
	Y	54.1	54.9	62.0	66.3	70.5	79.3	81.3	88.0	94.8	99.7	106.2	110.4	117.7	?
9	x_1	54.2	56.8	59.7	61.4	63.5	64.7	64.8	67.4	69.0	70.7	71.3	73.7	75.9	77.5
	x_2	33.6	39.1	40.4	42.9	44.0	46.8	51.9	56.3	56.6	58.7	59.6	62.4	63.9	67.2
	Y	75.4	85.4	88.5	92.7	95.2	99.5	106.2	113.2	114.5	118.1	118.7	123.0	127.4	?
10	x_1	43.6	44.0	44.8	46.3	47.6	47.9	50.1	50.5	51.1	52.9	53.1	54.0	56.8	57.6
	x_2	34.7	37.7	40.4	42.8	44.7	49.5	51.8	53.8	57.6	60.4	63.3	63.6	66.0	66.7
	Y	71.9	75.1	78.1	83.1	86.7	92.3	94.9	97.5	102.0	106.4	110.0	110.0	114.6	?

ЛІТЕРАТУРА

1. Толбатов, Ю.А. Економетрика [Текст]/ Ю.А. Толбатов Ю.А.— К.: Четверта хвиля, 1997.— 320 с
2. Доугерти, К. Введение в эконометрику [Текст]: пер. с англ. / К. Доугерти - М.: ИНФРА-2001.— С. 402.
3. Черняк, О.І. Економетрика [Текст]: підручник/ О.І. Черняк, О.В. Комашко, А.В. Ставицький, О.В.Баженова – Київ: Видавничо-поліграфічний центр "Київський університет", 2010. - 359 с.
4. Доля, В.Т. Економетрія [Текст]: навч. посібник / В.Т. Доля; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2010. – 171 с.
5. Корольов, О. А. Практикум з економетрії: завдання з практичними рекомендаціями, алгоритмами та прикладом їх наскрізного виконання [Текст]/ О. А. Корольов, В. В. Рязанцева. – Київ: Вид-во Європ. Ун-ту, 2002. – 250 с.
6. Руська, Р. В. Економетрика [Текст]: навч. посібник / Р. В. Руська. – Тернопіль : Тайп, 2012. – 224с.

Додаток А

Таблиця А.1 - Критичні значення *t*-статистики Стюдента

Число ступенів вільності	*2х стор.*	*10%*	*5%*	*1%*	*0,10%*
	1но стор.	*5%*	*2,50%*	*0,50%*	*0,05%*
1		6,314	12,706	63,657	636,62
2		2,92	4,303	9,925	31,598
3		2,353	3,182	5,841	12,924
4		2,132	2,776	4,604	8,61
5		2,015	2,571	4,032	6,869
6		1,943	2,447	3,707	5,959
7		1,895	2,365	3,499	5,408
8		1,86	2,306	3,355	5,041
9		1,833	2,262	3,25	4,781
10		1,812	2,228	3,169	4,587
11		1,796	2,201	3,106	4,437
12		1,782	2,179	3,055	4,318
13		1,771	2,16	3,012	4,221
14		1,761	2,145	2,977	4,14
15		1,753	2,131	2,947	4,073
16		1,746	2,12	2,921	4,015
17		1,74	2,11	2,898	3,965
18		1,734	2,101	2,878	3,922
19		1,729	2,093	2,861	3,883
20		1,725	2,086	2,845	3,85
21		1,721	2,08	2,831	3,819
22		1,717	2,074	2,819	3,792
23		1,714	2,069	2,807	3,767
24		1,711	2,064	2,797	3,745
25		1,708	2,06	2,787	3,725
26		1,706	2,056	2,779	3,707
27		1,703	2,052	2,771	3,69
28		1,701	2,048	2,763	3,674
29		1,699	2,045	2,756	3,659
30		1,697	2,042	2,75	3,646
40		1,684	2,021	2,704	3,551

Таблиця А.2 - Критичні значення F-статистики Фішера, для 5%

$k1\backslash k2$	1	2	3	4	5
1	161,4	199,5	215,7	224,6	230,2
2	18,51	19	19,16	19,25	19,3
3	10,13	9,55	9,28	9,12	9,01
4	7,71	6,94	6,59	6,39	6,26
5	6,61	5,79	5,41	5,19	5,05
6	5,99	5,14	4,76	4,53	4,39
7	5,59	4,74	4,35	4,12	3,97
8	5,32	4,46	4,07	3,84	3,69
9	5,12	4,26	3,86	3,63	3,48
10	4,96	4,1	3,71	3,48	3,33
11	4,84	3,98	3,59	3,36	3,2
12	4,75	3,89	3,49	3,26	3,11
13	4,67	3,81	3,41	3,18	3,03
14	4,6	3,74	3,34	3,11	2,96
15	4,54	3,68	3,29	3,06	2,9
16	4,49	3,63	3,24	3,01	2,85
17	4,45	3,59	3,2	2,96	2,81
18	4,41	3,55	3,16	2,93	2,77
19	4,38	3,52	3,13	2,9	2,74
20	4,35	3,49	3,1	2,87	2,71
21	4,32	3,47	3,07	2,84	2,68
22	4,3	3,44	3,05	2,82	2,66
23	4,48	3,42	3,03	2,8	2,64
24	4,26	3,4	3,01	2,78	2,62
25	4,24	3,39	2,99	2,76	2,6
26	4,23	3,37	2,98	2,74	2,59
27	4,21	3,35	2,96	2,73	2,57
28	4,2	3,34	2,95	2,71	2,56
29	4,18	3,33	2,93	2,7	2,55
30	4,17	3,32	2,92	2,69	2,53
40	4,08	3,23	2,84	2,61	2,45

Таблиця А.3 - Критичні значення F-статистики Фішера, для 1%

*k*1*k*2	1	2	3	4	5
1	4052	4999	5404	5624	5764
2	98,50	99,00	99,16	99,25	99,30
3	34,12	30,82	29,46	28,71,	28,24
4	21,20	18,00	16,69	15,98	15,52
5	16,26	13,27	12,06	11,39	10,97
6	13,75	10,92	9,78	9,15	8,75
7	12,25	9,55	8,45	7,85	7,46
8	11,26	8,65	7,59	7,01	6,63
9	10,56	8,02	6,99	6,42	6,06
10	10,04	7,56	6,55	5,99	5,64
11	9,65	7,21	6,22	5,67	5,32
12	9,33	6,93	5,95	5,41	5,06
13	9,07	6,70	5,74	5,21	4,86
14	8,86	6,51	5,56	5,04	4,69
15	8,68	6,36	5,42	4,89	4,56
16	8,53	6,23	5,29	4,77	4,44
17	8,40	6,11	5,19	4,67	4,34
18	8,29	6,01	5,09	4,58	4,25
19	8,18	5,93	5,01	4,50	4,17
20	8,10	5,85	4,94	4,43	4,10
21	8,02	5,78	4,87	4,37	4,04
22	7,95	5,72	4,82	4,31	3,99
23	7,88	5,66	4,76	4,26	3,94
24	7,82	5,61	4,72	4,22	3,90
25	7,77	5,57	4,68	4,18	3,85
26	7,72	5,53	4,64	4,14	3,82
27	7,68	5,49	4,60	4,11	3,78
28	7,64	5,45	4,57	4,07	3,75
29	7,60	5,42	4,54	4,04	3,73
30	7,56	5,39	4,51	4,02	3,70
40	7,31	5,18	4,31	3,83	3,51

Таблиця А.4 - Критичні значення $\chi_{\gamma k}^{2}$

k	γ=5 %	γ=1%	k	γ=5 %	γ=1%
1	3,8415	6,6349	**31**	45,0	52,2
2	5,9915	9,2104	**32**	46,2	53,5
3	7,8147	11,3449	**33**	47,4	54,8
4	9,4877	13,2767	**34**	48,6	56,1
5	11,0705	15,0863	**35**	49,8	57,3
6	12,5916	16,8119	**36**	51,0	58,6
7	14,0671	18,4753	**37**	52,2	59,9
8	15,5073	20,0902	**38**	53,4	61,2
9	16,9190	21,6660	**39**	54,6	62,4
10	18,3070	23,2093	**40**	55,8	63,7
11	19,6752	24,7250	**41**	56,9	65,0
12	21,0261	26,2170	**42**	58,1	66,2
13	22,3620	27,6882	**43**	59,3	67,5
14	23,6848	29,1412	**44**	60,5	68,7
15	24,9958	30,5780	**45**	61,7	70,0
16	26,2962	31,9999	**46**	62,8	71,2
17	27,5871	33,4087	**47**	64,0	72,4
18	28,8693	34,8052	**48**	65,2	73,7
19	30,1435	36,1908	**49**	66,3	74,9
20	31,4104	37,5663	**50**	67,5	76,2
21	32,6706	38,9322	**52**	69,8	78,6
22	33,9245	40,2894	**54**	72,2	81,1
23	35,1725	41,6383	**56**	74,5	83,5
24	36,4150	42,9798	**58**	76,8	86,0
25	37,6525	44,3140	**60**	79,1	88,4
26	38,8851	45,6416	**62**	81,4	90,8
27	40,1133	46,9628	**64**	83,7	93,2
28	41,3372	48,2782	**66**	86,0	95,6
29	42,5569	49,5878	**68**	88,2	98,0
30	43,7730	50,8922	**70**	90,5	100

Таблиця А.5 - Статистика Дарбіна-Уотсона, для 5% рівня значущості

	***k* = 1**		***k* = 2**		***k* = 3**		***k* = 4**		***k* = 5**	
n	DW_1.	DW_2.	DW_1.	DW_2.	DW_1.	DW_2.	DW_1.	DW_2.	DW_1.	DW_2.
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Таблиця А.5 - Критичні значення коефіцієнта кореляції рангів Спірмена

	γ			γ			γ	
n	**5 %**	**1%**	n	**5 %**	**1%**	n	**5 %**	**1%**
5	0,91	-	17	0,48	0,62	29	0,37	0,48
6	0,85	-	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40