МІЖРЕГІОНАЛЬНА
АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ

МАУП

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ ДЛЯ ВИКОНАННЯ КОНТРОЛЬНОЇ РОБОТИ

з дисципліни

"ЕКОНОМЕТРІЯ"

(для бакалаврів)

Київ 2008

Підготовлено доцентом кафедри математики *О. О. Юньковою* і професором кафедри математики *І. І. Юртиним*

Затверджено на засіданні кафедри математики (протокол № 9 від 28.05.07 р.)

Схвалено Вченою радою Міжрегіональної Академії управління персоналом

Юнькова О. О., Юртин І. І. Методичні рекомендації для виконання контрольної роботи з дисципліни “Економетрія” (для бакалаврів). — К.: МАУП, 2008. — 30 с.

Методичні рекомендації містять пояснювальну записку, варіанти завдань контрольної роботи з економетрії, вказівки до її виконання та зразки розв’язування завдань контрольної роботи, а також список літератури.

ПОЯСНЮВАЛЬНА ЗАПИСКА

Контрольна робота складається з чотирьох задач, кожна з яких має 10 варіантів початкових значень. Номер варіанта *N* визначається останньою цифрою у номері залікової книжки студента. Початкові значення для завдання визначаються додаванням виразу *N/10*, після підстановки замість *N* визначену цифру. Наприклад, якщо останньою цифрою у номері залікової книжки студента є цифра 2, то вираз *N*/10 набуває значення *N*/10=2/10=0,2.

КОНТРОЛЬНІ ЗАВДАННЯ

Задача № 1

На базі статистичних даних (економічного показника X за 12 місяців):

1) побудувати графік парної лінійної регресії $x(t) = \hat{a}_0 + \hat{a}_1 t$
2) оцінити всі її параметри;
3) визначити довірчі інтервали для параметрів регресії a_0, a_1 при рівні значущості α=0,05;
4) знайти коефіцієнти детермінації R^2 та кореляції R;
5) обчислити прогнозні значення показника X для наступних трьох місяців ($x(13)$, $x(14)$, $x(15)$):

T	X
1	5,93+N/10
2	6,17+N/10
3	7,15+N/10
4	6,87+N/10
5	7,71+N/10
6	8,20+N/10
7	7,77+N/10
8	7,36+N/10
9	9,45+N/10
10	9,57+N/10
11	10,24+N/10
12	10,70+N/10

Задача № 2

На базі **11** статистичних даних певного регіону

1) визначити параметри лінійної моделі залежності витрат на споживання ***(C)*** від рівня доходів ***(D)***, збережень ***(S)*** та заробітної плати ***(L)***;
2) оцінити коефіцієнт детермінації;
3) перевірити наявність автокореляції залишків;
4) дослідити мультиколінеарність між факторами.

i	*C(i)*	*D(i)*	*S(I)*	*L(i)*
1	*5,25+N/10*	*9,11+N/10*	*7,05+N/10*	*16,05+N/10*
2	*11,24+N/10*	*13,57+N/10*	*8,68+N/10*	*18,68+N/10*
3	*16,27+N/10*	*14,01+N/10*	*9,57+N/10*	*20,06+N/10*
4	*18,75+N/10*	*17,29+N/10*	*10,11+N/10*	*29,67+N/10*
5	*21,78+N/10*	*19,58+N/10*	*11,55+N/10*	*31,55+N/10*
6	*24,58+N/10*	*21,07+N/10*	*13,31+N/10*	*34,01+N/10*
7	*27,09+N/10*	*22,47+N/10*	*15,37+N/10*	*35,34+N/10*
8	*31,76+N/10*	*24,68+N/10*	*17,01+N/10*	*36,01+N/10*
9	*35,94+N/10*	*25,75+N/10*	*19,67+N/10*	*38,54+N/10*
10	*38,57+N/10*	*27,05+N/10*	*21,92+N/10*	*41,92+N/10*
11	*41,47+N/10*	*30,87+N/10*	*25,08+N/10*	*43,27+N/10*

Задача № 3

За статистичними показниками *Y*, *K* та *L* за *9* років проаналізувати класичну модель виробничої функції Кобба-Дугласа, що описує залежність між продуктивністю праці $y=Y/L$ та фондоозброєністю $x=K/L$ з урахуванням впливу технічного прогресу у виробництві регіону:

1) оцінити параметри нелінійної моделі;
2) оцінити коефіцієнт детермінації;
3) перевірити наявність автокореляції залишків;

t	*Y(i)*	*K(i)*	*L(i)*
1	*65,04+N/10*	*4,03+N/10*	*7,45+N/10*
2	*54,27+N/10*	*5,25+N/10*	*8,68+N/10*
3	*78,22+N/10*	*7,57+N/10*	*9,55+N/10*
4	*82,06+N/10*	*7,99+N/10*	*10,67+N/10*
5	*79,14+N/10*	*8,91+N/10*	*11,68+N/10*
6	*90,48+N/10*	*10,67+N/10*	*13,31+N/10*
7	*85,69+N/10*	*11,51+N/10*	*14,27+N/10*

8	*76,26+N/10*	*10,23+N/10*	*13,01+N/10*
9	*82,05+N/10*	*10,84+N/10*	*15,05+N/10*

Задача № 4

На основі статистики за *n* років визначити параметри найпростішої мультиплікативної моделі споживання Кейнса для певного регіону:

$C(t) = a_0 + a_1 Y(t) + u(t),$

$Y(t) = C(t) + I(t),$

де $C(t)$ — споживання, $Y(t)$ — національний дохід, $I(t)$ — інвестиції, $u(t)$ — стохастичне відхилення, похибка.

t	*C(t)*	*I(t)*	*Y(t)*
1	*15,25+N/10*	*11,11+N/10*	*22,05+N/10*
2	*15,84+N/10*	*13,25+N/10*	*24,68+N/10*
3	*16,27+N/10*	*14,57+N/10*	*27,28+N/10*
4	*16,75+N/10*	*15,29+N/10*	*29,67+N/10*
5	*17,14+N/10*	*16,21+N/10*	*31,55+N/10*
6	*17,68+N/10*	*17,67+N/10*	*33,31+N/10*
7	*18,58+N/10*	*18,47+N/10*	*35,74+N/10*
8	*19,26+N/10*	*19,23+N/10*	*37,01+N/10*
9	*20,24+N/10*	*20,61+N/10*	*40,25+N/10*
10	*21,57+N/10*	*21,05+N/10*	*41,92+N/10*

Зразки розв'язування контрольного завдання

Задача 1.

Нехай задано зміни обсягу споживання X (у. о.) домогосподарства протягом року на підставі вибірки *n*=12 спостережень (щомісячно впродовж року), яка наведена в таблиці 1.

Таблиця 1

I	1	2	3	4	5	6	7	8	9	10	11	12
x_I	107	109	110	113	120	122	123	128	136	140	145	150
y_I	102	105	108	110	115	117	119	125	132	130	141	144

На базі цих статистичних даних:

1) побудувати графік парної лінійної регресії $x(t) = \hat{a}_0 + \hat{a}_1 t$;
2) оцінити всі її параметри;

3) визначити довірчі інтервали для параметрів регресії a_0, a_1 при рівні значущості a=0,05;
4) знайти коефіцієнти детермінації R^2 та кореляції R;
5) обчислити прогнозні значення показника X для наступних трьох місяців (*x(13), x(14), x(15)*):

Розв'язання.

Виконувати завдання будемо засобами Excel. Час t у даній задачі є незалежною змінною, а витрати на споживання x — залежною змінною.

Початкову інформацію запишемо у таблиці Excel: заголовки робочої таблиці будемо писати у першому рядку, числові дані — починаючи з другого рядка. У перших двох колонках запишемо початкові дані: значення незалежної змінної — в стовпці А, значення залежної — у стовпці В.

Для знаходження параметрів лінійної моделі $x(t)=\hat{a}_0+\hat{a}_1 t$ використаємо формули:

$$\hat{a}_1=\frac{\overline{tx}-\bar{t}\cdot\bar{x}}{\overline{t^2}-\left(\bar{t}\right)^2},\ \hat{a}_0=\bar{x}-\hat{a}_1\bar{t}.$$

Для цього виконаємо додаткові розрахунки.

Обчислимо для всіх спостережень квадрати незалежної змінної $t_i^2, i=1,2,...,n,$ і попарні добутки значень залежної і незалежної змінної $t_i\,x_i, i=1,2,...,n$. Для цього у комірках C3 і D3 набрати відповідно до формули "=A2*A2" і "=A2*B2". Після отримання результату, продовжити формулу у наступні рядки (курсор у правому нижньому кутку у вигляді знака "+").

Середні значення $\bar{t},\bar{x},\overline{t^2},\overline{tx}$ можна визначити за допомогою функції СРЗНАЧ програми EXEL. Таким чином, маючи значення $\bar{t},\,\bar{x},\,\overline{tx},\,\overline{t^2}$, обчислюємо за вищезазначеною формулою параметри $\hat{a}_0=3{,}78,\,\hat{a}_1=96{,}08$.

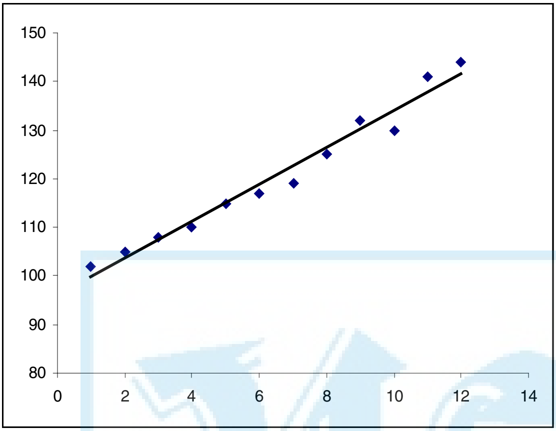
Отже, рівняння парної лінійної регресії має вигляд:

$$x(t)=96{,}08+3{,}78t.$$

В стовпці E запишемо наближені значення x_i^, які обчислюються за формулою

$$x_i\hat{}=96{,}08+3{,}78t_i.$$

Застосовуючи майстер діаграм EXEL, графічно зобразимо графіки статистичної залежності x від t та їх лінійну модель:



Залишки моделі $u_i = x_i - \hat{x}_i, i = 1,...,n,$ і їх квадрати обчислимо у стовпцях F і G, набираючи формули в F2 «=B2-E2», а в G2 «=F2*F2» і продовжуючи їх для всіх спостережень. Суму квадратів усіх залишків $\sum_{i=1}^{n} u_i^2$ обчислимо за допомогою функції СУММ, а середнє цих же значень дає вибіркову дисперсію залишків $S_u^2 = \frac{1}{n}\sum_{i=1}^{n} u_i^2$. Далі обчислюємо коефіцієнти детермінації R^2 та кореляції R

$$R^2 = 1 - \frac{S_u^2}{S_x^2}, \text{ де } S_x^2 = \overline{x^2} - \left(\overline{x}\right)^2; \; R = \sqrt{R^2}.$$

Маємо: $R^2 = 0{,}97, \; R = 0{,}99$.

Обчислене значення коефіцієнта кореляції дозволяє зробити висновок про сильну (пряму) лінійну залежність між змінними T та X. Це також підтверджується розташуванням точок і графіка тренду на кореляційному полі.

Прогнозоване споживання для наступних трьох місяців також визначається за отриманою формулою регресії, тобто X (t) = 96,08 + + 3,78*t, де t = 13,14,15. Значення легко отримати, задавши прогнозні значення незалежної змінної у першому стовпці і продовживши формулу у стовпці модельних значень.

Побудоване рівняння регресії в будь-якому випадку потребує певної інтерпретації та аналізу.

В нашому прикладі коефіцієнт $\hat{a}_1$ може розглядатися як гранична схильність до споживання. Фактично він показує, на яку величину зміниться обсяг споживання у наступному місяці, якщо тенденції минулого періоду залишаться незмінними.

Вільний член $\hat{a}_0$ рівняння регресії визначає прогнозне значення X при змінній T, що дорівнює нулю (тобто автономне споживання). В нашому випадку значення $\hat{a}_0$ =96,08 у. о. Цей параметр може визначати накопичені або позичені кошти.

Необхідно пам'ятати, що емпіричні коефіцієнти регресії $\hat{a}_0$ та $\hat{a}_1$ є лише оцінками теоретичних коефіцієнтів a_0 та a_1, а саме рівняння відображає лише загальну тенденцію в поведінці розглянутих змінних.

При змінюванні статистичної бази (початкових даних) результати оцінювання, очевидно, відрізнятимуться від попередніх, але з високою ймовірністю можуть опинитися в певних межах — в межах довірчого інтервалу параметрів регресії. Цей інтервал визначається для кожного параметра за відповідною формулою:

$$(\hat{a}_0-\Delta_0,\hat{a}_0+\Delta_0),\ (\hat{a}_1-\Delta_1,\hat{a}_1+\Delta_1),$$

де $\Delta_0 = t_{табл}\sqrt{\sigma_u^2 c_{00}}$, $\Delta_1 = t_{табл}\sqrt{\sigma_u^2 c_{11}}$.

У загальному випадку $\sqrt{\sigma_u^2 c_{jj}} = S_{a_j}$ — стандартна похибка параметра. Для парної регресії $\sigma_u^2 = \dfrac{1}{n-2}\sum_{i=1}^{n} u_i^2$ — незміщена дисперсія залишків,

$$c_{00} = \frac{\overline{t^2}}{n\sigma_t^2}, c_{11} = \frac{1}{n\sigma_t^2}, \sigma_t^2 = \overline{t^2} - \left(\overline{t}\right)^2.$$

$t_{табл} = t(\dfrac{\alpha}{2}, n-2)$ — табличне значення розподілу Стьюдента. Його визначають із стандартної таблиці розподілу Стьюдента (Додаток 2 у посібнику “Економетрія”) або за допомогою статистичної функції СТЬЮДРАСПОБР, параметри якої набираються вручну: у запиті “Вероятность” — 0,05/2; у запиті “Степени свободы” — 12–2.

Для аналізу щільності лінійної залежності обчислимо коефіцієнт кореляції:

$$r_{tx} = \frac{\overline{tx} - \overline{x}\cdot\overline{t}}{\sqrt{\overline{t^2} - \overline{t}^2}\cdot\sqrt{\overline{x^2} - \overline{x}^2}}.$$

Остаточна таблиця розрахунків набуде вигляду:

A	B	C	D	E	F	G	H
T	x	t^2	tx	x^	u	u^2	x^2
1	102	1	102	99,85	2,14	4,58	10404
2	105	4	210	103,64	1,35	1,84	11025
3	108	9	324	107,42	0,57	0,33	11664
4	110	16	440	111,20	-1,20	1,460	12100
5	115	25	575	114,99	0,008	0	13225
6	117	36	702	118,77	-1,778	3,150	13689
7	119	49	833	122,55	-3,55	12,66	14161
8	125	64	1000	126,34	-1,34	1,80	15625
9	132	81	1188	130,12	1,87	3,51	17424
10	130	100	1300	133,90	-3,91	15,27	16900
11	141	121	1551	137,69	3,31	10,95	19881
12	144	144	1728	141,47	2,52	6,38	20736
$\overline{t}$ =	$\overline{x}$ =	$\overline{t^2}$ =	$\overline{tx}$ =			$\sum u^2$ =	$\overline{x^2}$ =
6,5	120,66	54,166	829,41			61,94	14736,16
a_1=	3,78	Sa1=	0,21		S_u^2=	5,16	
a_0=	96,07	Sa0=	1,53		Sigma u^2=	6,19	
$a_0 \in$	(93,30	98,85)			T_{tabl}=	1,81	
$a_1 \in$	(3,41	4,16)			C_{00}=	0,38	
			R_{tx}=	0,99	C_{11}=	0,007	
			R^2=	0,97			

Задача № 2.

На базі ***n=15*** статистичних даних певного регіону

1) визначити параметри лінійної моделі залежності прибутку підприємства ***(Y)*** від рівня інвестицій ***(I)***, витрат на рекламу ***(Cr)*** та заробітної плати ***(L)***;
2) оцінити коефіцієнт детермінації R^2;
3) перевірити наявність автокореляції залишків;
4) дослідити мультиколінеарність між факторами.

№	Y	X1 (I)	X2 (Cr)	X3 (L)
1	15,70	17,37	5,28	1,42
2	17,34	18,24	6,47	1,58
3	21,57	22,47	6,98	1,98
4	33,50	18,47	7,05	2,04
5	32,30	16,82	7,94	2,38
6	37,90	17,60	8,12	3,48
7	40,78	17,12	8,69	3,07
8	48,02	19,81	9,31	3,84
9	43,30	18,67	10,45	4,28
10	49,57	20,83	10,47	4,67
11	52,14	22,84	13,48	5,98
12	55,17	28,85	15,78	6,51
13	59,18	29,61	17,65	7,82
14	62,22	35,67	18,47	8,58
15	77,58	47,87	19,64	9,47

Розв'язання.

В даному випадку загальна лінійна модель має вигляд:

$$y = \hat{a}_0 + \hat{a}_1 x_1 + \hat{a}_2 x_2 + +\hat{a}_3 x_3 + u,$$

де y — досліджувана (залежна) змінна Y — прибуток підприємства;

x_1, x_2, x_3 — незалежні, пояснюючі змінні або регресори (I, Cr і L відповідно);

$\widehat{a}_0, \widehat{a}_1, ..., \widehat{a}_m$ — п аметри моделі, u — випадкова складова регресійного рівняння.

1) Оцінки параметрів моделі $\widehat{a}_0, \widehat{a}_1, ..., \widehat{a}_m$ оцінимо методом найменших квадратів, матричний запис якого має вигляд:

$$A = (X^T X)^{-1}(X^T Y),$$

складемо вектор-стовпець і матрицю спостережень у вигляді:

$$
Y = \begin{pmatrix} 15{,}7 \\ 17{,}34 \\ 21{,}57 \\ 33{,}5 \\ 32{,}3 \\ 37{,}9 \\ 40{,}78 \\ 48{,}02 \\ 43{,}3 \\ 49{,}57 \\ 52{,}14 \\ 55{,}17 \\ 59{,}18 \\ 62{,}22 \\ 77{,}58 \end{pmatrix}
\qquad
X = \begin{pmatrix}
1 & 17{,}37 & 5{,}28 & 1{,}42 \\
1 & 18{,}24 & 6{,}47 & 1{,}58 \\
1 & 22{,}47 & 6{,}98 & 1{,}98 \\
1 & 18{,}47 & 7{,}05 & 2{,}04 \\
1 & 16{,}82 & 7{,}94 & 2{,}38 \\
1 & 17{,}6 & 8{,}12 & 3{,}48 \\
1 & 17{,}12 & 8{,}69 & 3{,}07 \\
1 & 19{,}81 & 9{,}31 & 3{,}84 \\
1 & 18{,}67 & 10{,}45 & 4{,}28 \\
1 & 20{,}83 & 10{,}47 & 4{,}67 \\
1 & 22{,}84 & 13{,}48 & 5{,}98 \\
1 & 28{,}85 & 15{,}78 & 6{,}51 \\
1 & 29{,}61 & 17{,}65 & 7{,}82 \\
1 & 35{,}67 & 18{,}47 & 8{,}58 \\
1 & 47{,}87 & 19{,}64 & 9{,}47
\end{pmatrix}
$$

Стовпчик одиниць у матриці ***X*** відповідає коефіцієнту 1 при параметрі a_0.

Виконувати розрахунки будемо за допомогою вбудованих функцій Excel поетапно:

1) ***X'*** – функція ТРАНСП(массив) (ТРАНСПонована матриця) із категорії "ссылки и массивы";
2) $X'X$, $X'Y$, A – функція МУМНОЖ(массив1, массив2) (Матричное УМНОЖение) із категорії "математические";
3) $(X'X)^{-1}$ – функція МОБР(массив) (Матрица ОБРатная) також із категорії "математические".

Для роботи з вказаними функціями потрібно:

1) виділити місце під результат
2) викликати функцію (натиснути кнопку $\boxed{f_x}$ на панелі інструментів, вказати категорію, вибрати функцію);
3) вказати аргументи функції (у тому порядку, як вони записані у формулі!) (зауваження: аргументами усіх перелічених функцій є масиви (стовпці чи таблиці даних), вказувати їх зручніше мишкою);
4) після виходу з діалогового вікна функції у рядку формул натиснути ліву клавішу мишки (аргументи виділяться рамками), а потім натиснути три клавіші на клавіатурі: Ctrl+Shift+Enter ***одночасно***.

Врешті отримаємо наступні результати

$$X'=\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 17{,}37 & 18{,}24 & 22{,}47 & 18{,}47 & 16{,}82 & 17{,}6 & 17{,}12 & 19{,}81 & 18{,}67 & 20{,}83 & 22{,}84 & 28{,}85 & 29{,}61 & 35{,}67 & 47{,}87 \\ 5{,}28 & 6{,}47 & 6{,}98 & 7{,}05 & 7{,}94 & 8{,}12 & 8{,}69 & 9{,}31 & 10{,}45 & 10{,}47 & 13{,}48 & 15{,}78 & 17{,}65 & 18{,}47 & 19{,}64 \\ 1{,}42 & 1{,}58 & 1{,}98 & 2{,}04 & 2{,}38 & 3{,}48 & 3{,}07 & 3{,}84 & 4{,}28 & 4{,}67 & 5{,}98 & 6{,}51 & 7{,}82 & 8{,}58 & 9{,}47 \end{pmatrix}$$

$$\boldsymbol{X'X}=\begin{pmatrix} 15 & 352{,}24 & 165{,}78 & 67{,}1 \\ 352{,}24 & 9335{,}74 & 4404{,}383 & 1858{,}071 \\ 165{,}78 & 4404{,}38 & 2147{,}268 & 914{,}9516 \\ 67{,}1 & 1858{,}07 & 914{,}9516 & 397{,}2576 \end{pmatrix}$$

$$\boldsymbol{(X'*X)^{-1}}\begin{pmatrix} 2{,}14866 & -0{,}0276 & -0{,}51745 & 0{,}958316 \\ -0{,}0276 & 0{,}00428 & -0{,}0056 & -0{,}00245 \\ -0{,}5174 & -0{,}0056 & 0{,}18797 & -0{,}31932 \\ 0{,}95831 & -0{,}0024 & -0{,}31932 & 0{,}58755 \end{pmatrix}$$

$$\boldsymbol{X'*Y}=\begin{pmatrix} 646{,}27 \\ 16861{,}1 \\ 8209{,}78 \\ 3498{,}18 \end{pmatrix} \qquad \boldsymbol{A}=\begin{pmatrix} 26{,}10789 \\ -0{,}2518 \\ -2{,}72767 \\ 11{,}85602 \end{pmatrix}$$

2. Запишемо функцію регресії з урахуванням знайдених оцінок коефіцієнтів моделі

$\hat{y} = 26{,}10789 - 0{,}2518x_1 - 2{,}72767x_2 + 11{,}85602x_3;$

1.4) модельні значення $\hat{y}_i$ зручно розташувати у тих же рядках, де записано початкові дані. Обчислити їх можна двома способами:

- за формулою $y = \hat{a}_0 + \hat{a}_1x_1 + \hat{a}_2x_2 + \hat{a}_3x_3$, яка набирається у першому рядку і продовжується у всіх інших. Значення параметрів $\widehat{a}_0,\ \widehat{a}_1, ..., \widehat{a}_m$ у кожному рядку є незмінними, тому їх адреса фіксується (клавіша F4).
- за формулою $\hat{Y} = XA$, яка реалізується вбудованою функцією МУМНОЖ(X, A) (місце під її результат вказується перед викликом функції).

Отже, ми побудували загальну лінійну модель (2.5) залежності прибутку від інвестицій, витрат на рекламу та заробітної плати. Наступним кроком наших досліджень є проведення дисперсійно-кореляційного аналізу та аналізу залишків.

2-й крок:

2.1) обчислимо залишки моделі $u_i = y_i - \hat{y}_i$, $i = 1,2,...,15$; і їх квадрати аналогічно тому, як це виконувалось у першому завданні;

2.2) обчислимо середньоквадратичну похибку дисперсії залишків

$$\sigma_u = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-m-1}} = \sqrt{\frac{\sum_{i=1}^{n}u_i^2}{n-m-1}},$$

Маємо $\sigma_u = 5{,}7357$;

2.3) перевіримо тісноту загального впливу незалежних змінних на залежну змінну: обчислити коефіцієнт детермінації за формулою

$$R^2 = 1 - \frac{\sum_{i=1}^{n}u_i^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}.$$

Отримаємо $R^2 = 0{,}91436$.

Висновок: *оскільки коефіцієнт детермінації наближається до одиниці, то варіація залежної змінної Y значною мірою визначається варіацією незалежних змінних.*

Для знаходження параметрів регресії $\widehat{a}_0, \widehat{a}_1, ..., \widehat{a}_m$ можна використати функцію ***ЛИНЕЙН***. Після використання цієї функції з'являється регресійна статистика у вигляді таблиці:

$\widehat{a}_m$	$\widehat{a}_{m-1}$	...	$\widehat{a}_2$	$\widehat{a}_1$	$\widehat{a}_0$
$S_{\widehat{a}_m}$	$S_{\widehat{a}_{m-1}}$	...	$S_{\widehat{a}_2}$	$S_{\widehat{a}_1}$	$S_{\widehat{a}_0}$
R^2	S_u				
$F_{експ}$	k				
$S_{регр}$	$\sum u_i^2$				

$S_{a_i}\left(i = \overline{0; m}\right)$ — стандартні значення похибок для параметрів моделі $a_0, a_1, ..., a_m$;

F — спостережуване значення F-статистики;

k — кількість ступенів вільності;

$S_{регр}$ — регресійна сума ($S_{регр} = \sum(y_i - \overline{y})^2 - \sum u_i^2$).

Для заданих статистичних даних застосування цієї функції приводить до наступної таблиці:

11,85602	-2,72767	-0,2518	26,10789
4,396525	2,486744	0,375368	8,407584
0,91436	5,735708	#Н/Д	#Н/Д
39,14827	11	#Н/Д	#Н/Д
3863,741	361,8818	#Н/Д	#Н/Д

Як бачимо, результати, одержані за допомогою функції **ЛИНЕЙН**, повністю відповідають результатам, одержаним за допомогою формул обчислення параметрів моделі з кроку 2.

3-й крок:

Перевірка статистичних гіпотез

3.1. Перевіримо значущість вибіркового коефіцієнта кореляції: обчислимо R= $\sqrt{R^2}$ — коефіцієнт кореляції (характеризує тісноту лінійного зв'язку усіх незалежних факторів x_i з залежною змінною у)

R=0,956222

Коефіцієнт кореляції, який близький до одиниці, свідчить про те, що *існує тісний лінійний зв'язок усіх незалежних факторів* x_1, x_2, x_3 *із залежною змінною у*.

Однак потрібна ще перевірка його значущості.

3.2. **Гіпотеза 1.** (H_0: $R = 0$) обчислимо t-статистику за формулою

$$t = \frac{R\sqrt{n-m-1}}{\sqrt{1-R^2}}; \qquad t=37,03215;$$

знайдемо $t_{табл.(a/2,\, n-m-1)}$ — табличне значення t-розподілу з рівнем значущості δ***=0,05*** і ***(n-m-1)=11*** ступенями свободи. Його можна визначити за таблицею розподілу Стьюдента (дод. 4) або за допомогою статистичної функції СТЬЮДРАСПОБР(вероятность, степени свободы), де "вероятность" — 0,025 ***(δ/2)***, "степени свободы" — 11. Отримаємо:

$$t_{табл.(0,025,11)}=2,201.$$

Оскільки $|t|>t_{табл.(0,025,11)}$, то можна зробити висновок про достовірність коефіцієнта кореляції, який характеризує тісноту зв'язку між залежною і незалежними змінними моделі.

3.3) Для вибраного рівня значущості α=0,05 і степеня свободи ***k=n-m-1=11*** запишемо межі надійності для множинного коефіцієнта кореляції R

$$(R\text{-}\Delta R;\ R+\Delta R),$$

де $\Delta \boldsymbol{R} = \boldsymbol{t_{0,025,11}} \cdot \dfrac{\boldsymbol{1-R}}{\sqrt{\boldsymbol{15}}} = 2{,}201*(1-0{,}956222)/\sqrt{15} = 0{,}029311,$

довірчий інтервал для множинного коефіцієнта кореляції R

$$(R\text{-}\Delta R;\ R+\Delta R)=(0{,}926911;\ 0{,}985533)$$

3.4) **Гіпотеза 2.** (H_0:$R^2 = 0$, що рівносильно a_1=a_2=a_3=0). Обчислимо F-статистику за формулою

$$F_{експ} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m};$$

F = 39,14827

знайдемо табличне значення *F*-статистики $F(m, n-m-1, \alpha)$ (дод. 5 або статистична функція FРАСПОБР(вероятность, степени свободы1, степени свободы2):

FРАСПОБР(0,05; 3; 11)=3,59 і порівняємо з обчисленою $F_{експ}$-статистикою: оскільки ***F> FРАСПОБР(0,05;3;11),*** то *нульова гіпотеза відхиляється, тобто коефіцієнт детермінації є значущим*. Це значить, що гіпотеза a_1=a_2=a_3=0 відхиляється, тобто показник *y* істотно залежить хоча б від одного із факторів x_1, x_2, x_3.

3.5) **Гіпотеза 3.** (H_0: a_0=0, a_1=0, a_2=0, a_3=0). Обчислити експериментальне значення *t*-статистики для кожного коефіцієнта a_i:

$$t_i = \frac{a_i}{S_{a_i}},\ i = \overline{0;m},$$

де Sa_0, Sa_1, Sa_2, Sa_3 – стандартні значення похибки параметрів a_0, a_1, a_2, a_3, які обчислені на другому кроці. Їх можна взяти з другого рядка таблиці, яка є результатом роботи функції ЛИНЕЙН.

Таким чином, маємо:

$$t_0= 3{,}11,\ t_1= -0{,}67,\ t_2= -1{,}10,\ t_3= 2{,}70.$$

Порівняємо абсолютні величини цих *t*-статистик з табличним значенням $\boldsymbol{t_{табл.(0,025,11)}}$ ***=2,201***. Будемо мати:

$|t_0| > t_{табл}$. Гіпотеза $a_0 = 0$ відхиляється. Цей параметр істотно відмінний від нуля, він є вагомим.

$|t_1| < t_{табл}$. Гіпотеза $a_1 = 0$ приймається. Цей параметр випадково відмінний від нуля, він є невагомим.

$|t_2| < t_{табл}$. Гіпотеза $a_2 = 0$ приймається. Цей параметр випадково відмінний від нуля, він є невагомим.

$|t_3| > t_{табл}$. Гіпотеза $a_3 = 0$ відхиляється. Цей параметр істотно відмінний від нуля, він є вагомим.

Внаслідок перевірки цієї гіпотези зроблено такі висновки: фактори x_1, x_2 слабо впливають на показник y; фактор x_3 сильно впливає на показник y. Також є істотним постійний фактор, який визначається параметром моделі $a_0 = 0$.

4-й крок:

Надійні зони регресії (довірчі інтервали для значень y_i) обчислимо за формулою

$$(\hat{y}_i - t_{\alpha/2,k}\sigma\sqrt{X_i\,(X^T X)^{-1} X_i^T}\,; \hat{y}_i + t_{\alpha/2,k}\sigma\sqrt{X_i\,(X^T X)^{-1} X_i^T}\,),$$

де $t_{\alpha/2,k} = t_{табл.(\alpha/2,\, n-m-1)} = t_{табл.(0,025,11)}$ — табличне значення t-розподілу з ***(n-m-1)=11*** степенями свободи і рівнем значущості α = 0,05 (дод. 4): $\boldsymbol{t_{табл.(0,025,11)} = 2,201}$ (обчислене у попередньому пункті). σ — незміщена дисперсія залишків рівняння ($\sigma = 5,7357$ з п. 2.5), X_i — i-тий рядок матриці спостережень X, X_i^T — i-тий стовпець транспонованої матриці X^T.

Розрахунки потрібно виконати поетапно.

Обчислення під знаком кореня виконуються за допомогою математичної функції МУМНОЖ, яку потрібно застосувати двічі:

1) помножити матрицю X на обернену $(X^T X)^{-1}$ — отримати матрицю розмірності (15×4);

2) помножити її на транспоновану X^T — отримати матрицю розмірності (15×15) і записати окремим стовпцем її діагональні елементи.

Обчислити корені квадратні від кожного елемента останнього стовпця (математична функція КОРЕНЬ) і за обчисленими раніше $t_{табл.}(0,025,11)$ і σ визначити межі довірчих інтервалів для кожного модельного значення $\hat{y}_i$ — два нових вектори-стовпці по 15 елементів.

5-й крок:

Довірчі інтервали для окремого параметра моделі b_j, $j = 0,1,2,...,m$, обчислюються за формулою

$$(a_j - t_{\alpha/2,k}\sqrt{\sigma^2 c_{jj}};\ a_j + t_{\alpha/2,k}\sqrt{\sigma^2 c_{jj}}),$$

$$\text{або}\left(a_j - t_{\alpha/2,k}\sigma\sqrt{c_{jj}}; a_j + t_{\alpha/2,k}\sigma\sqrt{c_{jj}}\right)$$

де $t_{\alpha/2,k}$, σ — визначені в попередніх пунктах, c_{jj} — діагональні елементи матриці $C=(X^T X)^{-1}$ (для зручності можна переписати їх у вигляді окремого стовпця і розрахунки виконати за спільною формулою, зафіксувавши адреси $t_{\alpha/2,k}$ і σ). Як і в попередньому пункті буде утворено ще два нових стовпця, тепер по чотири елементи.

6-й крок:

Прогнозування за моделлю складається з двох варіантів: точкового та інтервального прогнозу.

6.1) Точковий прогноз:

за моделлю (3.5) обчислюють значення залежної змінної Y_p для заданих прогнозних значень $X1_p$, $X2_p$, $X3_p$.

6.2) Інтервальний прогноз:

межі надійних інтервалів індивідуальних прогнозованих значень обчислюють за формулою

$$(y_{pi} - \Delta y_{pi};\ y_{pi} + \Delta y_{pi}),$$

де $\Delta y_{pi} = t_{\alpha/2,k}\sigma\sqrt{1 + X_p(X^T X)^{-1} X_p^T}$, $t_{\alpha/2,k}$, σ — визначені в попередніх пунктах, X_p — вектор-рядок незалежних змінних, що лежить за межами базового періоду, X_p^T — транспонований у стовпець рядок X_p, $\hat{Y}_p = X_p B$ — точкова оцінка математичного сподівання прогнозного значення Y_p, яка може розглядатися також як індивідуальне значення залежної змінної для відповідного вектора незалежних змінних.

6.3) Межі надійних інтервалів для математичного сподівання значення y_{pi} знаходять за формулою

$$(\hat{Y}_p - t_{\alpha/2,k}\sigma\sqrt{X_p(X^T X)^{-1} X_p^T} \le M(Y_p(X_p)) \le$$

$$\le \hat{Y}_p + t_{\alpha/2,k}\sigma\sqrt{X_p(X^T X)^{-1} X_p^T}).$$

Розрахунки під знаком кореня виконуються, як описано раніше, із застосуванням функції МУМНОЖ, значення $t_{\alpha/2,k}$ і σ ті ж самі, що і раніше.

II) **Перевірка наявності автокореляції залишків.** Ця перевірка виконується за критерієм Дарбіна-Уотсона

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

Щоб визначити експериментальне значення DW потрібно:

- визначити різниці між сусідніми значеннями залишків (оскільки таких різниць 14, то формулу для їх розрахунку потрібно набирати з другого рядка відносно елементів вектора залишків моделі);
- піднести їх до квадрату і додати (кнопка Σ);
- отриману суму поділити на суму квадратів залишків, яка обчислювалася раніше.

Порівняти отриманий результат з табличними значеннями розподілу Дарбіна-Уотсона (Додаток 4) на рівні α=0,05.

За таблицею Дарбіна–Уотсона при заданому рівні значущості α, кількості факторів ***m*** та кількості спостережень ***n***, знаходимо два значення $\boldsymbol{DW_1}$ та $\boldsymbol{DW_2}$.

- Якщо $0 < DW < DW_1$, то наявна додатна автокореляція.
- Якщо $\boldsymbol{DW_1 \le DW \le DW_2}$ або $\boldsymbol{4 - DW_2 \le DW \le 4 - DW_1}$, то ми не можемо зробити висновки ні про наявність, ні про відсутність автокореляції (DW потрапляє в зону невизначеності).
- Якщо $\boldsymbol{4 - DW_1 < DW < 4}$, то ми маємо від'ємну автокореляцію.
- Якщо $\boldsymbol{DW_2 < DW < 4 - DW_2}$, то автокореляція відсутня.

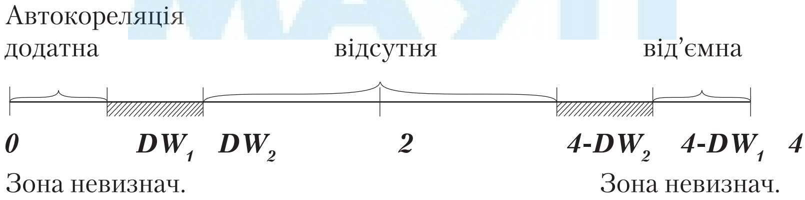


Рис. 1. Зони автокореляційного зв'язку за критерієм Дарбіна-Уотсона.

III) дослідження мультиколінеарності виконуємо за алгоритмом Феррара-Глобера.

III. Перевірка на наявність або відсутність мультиколінеарності. Алгоритм Фаррара-Глобера.

Наявність мультиколінеарності в моделі свідчить про те, що фактори x_1, x_2, x_3 є залежними між собою. Це приводить до того, що не можна вказати, який же вплив кожного фактора на показник y. Перевірка здійснюється за алгоритмом Фаррара-Глобера. Наведемо покроковий алгоритм.

1-й крок:
нормалізувати змінні $\boldsymbol{x_1, x_2, \ldots x_m}$ моделі, для чого обчислити

$$x_{ij}^{*} = \frac{(x_{ij} - \overline{x}_j)}{\sqrt{n\sigma_{x_j}^2}}, \text{ або } x_{ij}^{*} = \frac{(x_{ij} - \overline{x}_j)}{\sqrt{\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2}}$$

де $\boldsymbol{n}$ — число спостережень, $\boldsymbol{(i=1,2,\ldots, n)}$;

$\boldsymbol{m}$ — число незалежних змінних, $\boldsymbol{(j=1, m)}$;

$\overline{x}_j$ — середня арифметична j-ї незалежної змінної;

$\sigma^2 \boldsymbol{xj}$ — дисперсія $\boldsymbol{j}$-ї незалежної змінної.

2-й крок:
побудувати нову матрицю $\boldsymbol{X^*}$ — матрицю, елементами якої є нормалізовані незалежні змінні x_{ij}^{*}, і обчислити кореляційну матрицю (матрицю моментів нормалізованої системи нормальних рівнянь)

$$R = X^{*tr} X^{*} = \begin{pmatrix} 1 & r_{12} & \ldots & r_{1m} \\ r_{21} & 1 & \ldots & r_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ r_{m1} & r_{m2} & \ldots & 1 \end{pmatrix},$$

де X^{*tr} — транспонована матриця до матриці $\boldsymbol{X^*}$ (елементи матриці $\boldsymbol{R}$ характеризують тісноту зв'язку однієї незалежної змінної з іншою $(r_{ij} = r_{x_i x_j})$ — парні коефіцієнти кореляції.

3-й крок:

3.1) визначити $|R|$ — визначник кореляційної матриці $\boldsymbol{R}$;

3.2) визначити критерій χ^2, як

$$\chi^2 = -[n-1-\frac{1}{6}(2m+5)]\cdot\ln|R|;$$

3.3) порівняти значення χ^2 з табличним при $\frac{1}{2}m(m-1)$ ступенях свободи і рівні значущості α (якщо $\chi^2 > \chi^2_{табл}$, то в масиві незалежних змінних існує мультиколінеарність).

4-й крок:

визначити матрицю $\boldsymbol{C}$ — помилок

$$C = R^{-1} = (X^{*tr}X^*)^{-1} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \end{pmatrix}.$$

5-й крок:

5.1) розрахувати $\boldsymbol{F}$-критерії $\boldsymbol{F_k = \frac{(c_{kk}-1)\cdot(n-m)}{(m-1)}}$,

де c_{kk} — діагональні елементи матриці C;

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

5.3) розрахувати коефіцієнти детермінації для кожної змінної:

$$R_k^2 = 1 - \frac{1}{c_{kk}}.$$

6-й крок:

знайти часткові коефіцієнти кореляції, які характеризують тісноту зв'язку між двома змінними за умови, що всі інші змінні $\boldsymbol{x_{l1}}$, $\boldsymbol{x_{l2}}$,..., $\boldsymbol{x_{lm}}$ не впливають на цей зв'язок (існування парної мультиколінеарності)

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

де c_{kj} — елементи матриці C, що знаходиться в $\boldsymbol{k}$-му рядку та $\boldsymbol{j}$-му стовпці, $\boldsymbol{k=1,2,\dots, m; j=1,2,\dots, m, c_{kk}}$ і $\boldsymbol{c_{jj}}$ — діагональні елементи матриці $\boldsymbol{C}$.

Однак, якщо порівняти конкретні числові значення часткових та парних коефіцієнтів, то можна побачити, що перші значно менші за останні. Тому на основі знання парних коефіцієнтів кореляції висновок про мультиколінеарність робити неможливо. Для цього необхідно ще виконати сьомий крок.

7-й крок:

7.1) розрахувати t-критерії $t_{kj} = |r_{kj}| \dfrac{\sqrt{n-m}}{\sqrt{1-r_{kj}^2}}$;

7.2) значення критеріїв t_{kj} порівняти з табличними при $(m\text{-}n)$ ступенях свободи і рівні значущості α; якщо $t_{kj} > t_{табл.}$, то між незалежними змінними x_k і x_j існує мультиколінеарність.

Висновок:

1. Між незалежними змінними може існувати лінійна залежність, але вона може не бути явищем мультиколінеарності змінних і тому не буде негативно впливати на кількісні параметри моделі, які розраховані за допомогою звичайного МНК.
2. Якщо $F_k > F_{табл}$, то x_k залежить від усіх інших незалежних змінних і треба вирішити питання про її виключення з переліку змінних.
3. Якщо $t_{kj} > t_{табл}$, то x_k і x_j тісно пов'язані між собою.
4. Аналізуючи F і t критерії, робиться висновок, яку із змінних треба виключити з розгляду в будованій моделі (зрозуміло, що треба при цьому виходити з економіко-логіко-теоретичних міркувань).
5. Якщо після вилучення певної змінної ми ще не позбавились мультиколінеарності, то оцінку параметрів моделі слід отримувати за допомогою іншого методу, наприклад, методу головних компонентів (або одного з його модифікацій).

Приклад. (Дослідження наявності мультиколінеарності на основі алгоритму Фаррара-Глобера).

Розглянемо задачу дослідження впливу на економічний показник y –реальне споживання країни (в млрд грн) трьох факторів x_1-купівля та оплата товарів та послуг (в млрд грн), x_2-всього заощаджень від загального грошового доходу, в % від загальної суми доходу, x_3-рівень ставки ПДВ (в %). Необхідно перевірити фактори на мультиколінеарність.

№	$y(i)$	$x_1(i)$	$x_2(i)$	$x_3(i)$
1	25,74	4,69	11,97	29,23
2	25,34	5,64	13,43	29,35
3	31,26	6,26	12,92	33,40
4	33,50	6,99	14,74	30,97
5	32,30	6,36	14,64	32,92

6	38,90	7,60	17,10	37,27
7	41,58	7,12	15,63	30,97
8	48,02	6,81	15,35	33,58
9	43,30	8,67	15,85	35,62
10	51,78	7,83	18,05	34,99
11	52,14	7,84	17,24	39,34
12	54,94	8,85	20,52	41,50
13	59,18	9,61	19,18	45,58
14	62,22	10,67	19,03	41,08
15	63,62	11,04	21,45	40,54
16	65,01	11,85	22,25	42,75
17	67,78	12,94	24,75	43,89
18	71,45	14,24	25,03	41,95
19	75,24	15,67	27,87	44,06
20	77,38	16,33	30,48	46,77

Розв'язання:

1-й крок:

нормалізуємо змінні $\boldsymbol{x_1}$, $\boldsymbol{x_2}$, $\boldsymbol{x_3}$ економетричної моделі, для чого обчислимо

$$x_{ij}^* = \frac{(x_{ij} - \bar{x}_j)}{\sqrt{n\sigma_{x_j}^2}},$$

де $\boldsymbol{n}$=15 — число спостережень, ($\boldsymbol{i=1,2,\dots, n}$);

$\boldsymbol{m}$ =3 — число незалежних змінних, ($\boldsymbol{j=1, m}$);

$\bar{x}_j$ — середня арифметична j-ї незалежної змінної;

σ_{xj}^2 — дисперсія j-ї незалежної змінної: $\sigma_{xj}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2$.

розрахунки за цією формулою проводяться поетапно:

— під кожним стовпцем незалежних змінних обчислити середнє значення:

$\overline{x_1}$ =9,3505; $\overline{x_2}$ =18,874; $\bar{x}_3$ =37,788;

— від кожного елемента у стовпцях незалежних змінних відняти відповідне середнє значення, тобто обчислити $(x_{ij} - \bar{x}_j)$ (в результаті утвориться робоча матриця Р1 розмірності (15×3));

— отримані різниці піднести до квадрата, тобто обчислити $(x_{ij} - \bar{x}_j)^2$ (ще одна матриця Р2 такої ж розмірності, як і Р1);

— додати елементи у кожному стовпці матриці Р2, тобто отримати

$$n\sigma_{xj}^2 = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2;$$

— обчислити їх корені квадратні, тобто $\sqrt{n\sigma_{xj}^2}$ для j=1,2,3;

— розділити елементи відповідних стовпців матриці Р1 на значення $\sqrt{n\sigma_{xj}^2}(j = 1,2,3)$.

В результаті отримаємо нову матрицю ***X****- матрицю, елементами якої є нормалізовані незалежні змінні x_{ij}^*

$$X^* = \begin{pmatrix} -0,309287 & -0,302374 & -0,339018 \\ -0,246242 & -0,238430 & -0,334264 \\ -0,205096 & -0,260766 & -0,173826 \\ -0,156651 & -0,181056 & -0,270089 \\ -0,198460 & -0,185436 & -0,192841 \\ -0,116169 & -0,077695 & -0,020520 \\ -0,148024 & -0,142077 & -0,270089 \\ -0,168596 & -0,154340 & -0,166696 \\ -0,045160 & -0,132441 & -0,085883 \\ -0,100905 & -0,036088 & -0,110840 \\ -0,100242 & -0,071564 & 0,061481 \\ -0,033215 & 0,072089 & 0,147047 \\ 0,017221 & 0,013401 & 0,308673 \\ 0,087566 & 0,006832 & 0,130409 \\ 0,112121 & 0,112820 & 0,109018 \\ 0,165875 & 0,147858 & 0,196565 \\ 0,238212 & 0,257350 & 0,241725 \\ 0,324485 & 0,269613 & 0,164874 \\ 0,419385 & 0,393997 & 0,248460 \\ 0,463185 & 0,508307 & 0,355814 \end{pmatrix}$$

2-й крок:

Обчислимо кореляційну матрицю (матрицю моментів нормалізованої системи нормальних рівнянь)

$$R = X^{*tr} X^{*} = \begin{pmatrix} 1 & r_{12} & \dots & r_{1m} \\ r_{21} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & 1 \end{pmatrix},$$

де X^{*tr} — транспонована матриця X* (елементи матриці R характеризують тісноту зв'язку однієї незалежної змінної з іншою, тобто $r_{ij} = r_{x_i x_j}$ — парні коефіцієнти кореляції. Застосувавши функцію МУМНОЖ, отримаємо

$$R = \begin{pmatrix} 0{,}95 & 0{,}92883 & 0{,}819534 \\ 0{,}92883 & 0{,}95 & 0{,}831277 \\ 0{,}81953 & 0{,}83127 & 0{,}95 \end{pmatrix}.$$

3-й крок:

3.1) обчислимо |R| — визначник кореляційної матриці R за допомогою математичної функції МОПРЕД:

$$|R|=0{,}00881;$$

3.2) визначимо значення критерію χ^2, як

$$\chi^2 = -(n-1-\frac{1}{6}(2m+5))\ \mathbf{ln|R|} \quad \chi^2= 81{,}23005;$$

3.3) порівняємо значення χ^2 з табличним при $\frac{1}{2}m(m-1)$ = 3 степенях свободи і рівні значущості α=0,05 (додаток 3 або статистична функція ХИ2ОБР):

$$\chi^2_{табл}=7{,}814724$$

Оскільки $\chi^2>\chi^2_{табл.}$, то в масиві незалежних змінних існує мультиколінеарність в сукупності.

4-й крок:

визначимо матрицю С — помилок,

$$C = R^{-1} = (X^{*tr} X^{*})^{-1} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \end{pmatrix}.$$

Скориставшись математичною функцією МОБР, отримаємо:

$$c = \begin{matrix} 24,00377 & -22,82912 & -0,7311322 \\ -22,82912 & 26,20416 & -3,235457 \\ -0,731132 & -3,235457 & 4,5144749 \end{matrix}.$$

5-й крок:

5.1) розрахуємо F-критерії

$$F_k = \frac{(c_{kk}-1)\cdot(n-m)}{(m-1)}, k = 1, 2, 3.$$

Винесемо в окремий стовпець c_{kk} — діагональні елементи матриці С, помножимо кожен з них на (15–3)/(3–1)=6, отримаємо:

$F_1=195,5320; F_2=214,2354; F_3=29,87303;$

5.2) значення критеріїв F_k порівняємо з табличним при (n-m)=12 і (m-1)=2 степенях свободи і рівні значущості $\alpha=0,05$ (додаток 5 або статистична функція FРАСПОБР):

$$F_{табл.} = 19,43703$$

Так як $F_1 > F_{табл.}$, $F_2 > F_{табл.}$, $F_3 > F_{табл.}$, то робимо висновок, що *перша, друга і третя незалежні змінні мультиколінеарні з іншими*;

5.3) розрахуємо коефіцієнти детермінації для кожної змінної:

$$R_k^2 = 1 - \frac{1}{c_{kk}}$$

(при розрахунках можна використати виписані раніше значення c_{kk}).

$R_1^2 = 0,958339;$ $R_2^2 = 0,961838;$ $R_3^2 = 0,778490.$

6-й крок:

знайдемо часткові коефіцієнти кореляції, які характеризують тісноту зв'язку між двома змінними за умови, що всі інші змінні $x_{l1}, x_{l2}, \ldots, x_{lm}$ не впливають на цей зв'язок (існування парної мультиколінеарності)

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

де c_{kj} — елементи матриці С, що знаходиться в k-му рядку та j-му стовпці, коли $k \neq j$, тобто елементи c_{12}, c_{13}, c_{23}, c_{kk}, і c_{jj} — діагональні елементи матриці С (c_{11}, c_{22}, c_{33} у відповідних комбінаціях).

r_{12}=0,910257 r_{13}=0,070234 r_{23}=0,297472.

Однак, якщо порівняти абсолютні значення часткових та парних коефіцієнтів, то можна побачити, що перші значно менші за останні. Тому на основі знання парних коефіцієнтів кореляції висновок про мультиколінеарність робити неможливо. Для цього необхідно ще виконати сьомий крок.

7-й крок:

7.1) розрахуємо t-критерії $t_{kj} = |r_{kj}| \dfrac{\sqrt{n-m}}{\sqrt{1-r^2{kj}}}$

t_{12}= 9,064506 t_{13}=0,290302 t_{23}=1,284666;

7.2) значення критеріїв t_{kj} порівняємо з табличними при $(n\text{-}m)=12$ ступенях свободи і рівні значущості $\alpha=0{,}05$ (додаток 4 або статистична функція СТЬЮДРАСПОБР) $t_{табл}$ = 2,109818.

Так як $t_{12}>t_{табл.}$, $t_{13}<t_{табл.}$, $t_{23}<t_{табл.}$, то *між першою і другою незалежними змінними існує мультиколінеарність.*

Якщо *F*- критерій більше табличного значення, а це значить, що *k*- та змінна залежить від всіх інших в масиві, то необхідно вирішувати питання про її виключення з переліку незалежних змінних моделі.

Якщо t_{kj}-критерій більше табличного, то ця пара змінних (x_k і x_j) тісно взаємопов'язана. Звідси, аналізуючи рівень обох критеріїв *F* і *t*, можна зробити обґрунтований висновок про те, яку із змінних необхідно виключити із дослідження чи замінити іншою. Але заміна масиву незалежних змінних завжди повинна узгоджуватись з економічною доцільністю, що випливає з мети дослідження.

Задача № 3.

Класична виробнича функція Кобба-Дугласа має вигляд

$$Y(i) = aK(i)^{b_1} L(i)^{b_2}.$$

Для оцінки параметрів моделі необхідно прологарифмувати цю рівність:

$$\ln Y(i) = \ln a + b_1 \ln K(i) + b_2 \ln L(i) + \xi(t).$$

Після заміни змінних

$$y_i = \ln Y(i);\ b_0 = \ln a;\ x_{i_1} = \ln K(i);\ x_{i_2} = \ln L(i);$$

отримаємо лінійну форму багатофакторної моделі,

$$y = b_0 + b_1 x_1 + b_2 x_2 \quad ,$$

параметри якої оцінимо за МНК (далі див. співвідношення, що застосовуються при розв'язанні задачі № 2). За знайденим b_0 параметр задачі a визначиться за формулою:

$$a = e^{b_0}.$$

Задача № 4

На основі статистики за n років визначити параметри найпростішої мультиплікативної моделі споживання Кейнса для певного регіону:

C(t) = a0 + a1 Y(t) + u(t),

Y(t) = C(t) +I(t),

де *C(t)* — споживання, *Y(t)* — національний дохід, *I(t)* — інвестиції, *u(t)* — стохастичне відхилення, похибка.

Підставивши значення Y(t) з другого рівняння моделі в перше, дістанемо:

$$C(t) = b_1 + b_2 I(t) + \xi(t),$$

де $b_1 = \dfrac{a_0}{1 - a_1}$, $b_2 = \dfrac{a_1}{1 - a_1}$, $\xi(t) = \dfrac{u(t)}{1 - a_1}$.

Знайдемо оцінки параметрів отриманої моделі: $\hat{b}_1, \hat{b}_2$ та повернемось до параметрів початкової моделі $\hat{a}_1 = \dfrac{b_2}{1 + b_2}$, $\hat{a}_0 = b_1(1 - a_1)$.

СПИСОК ЛІТЕРАТУРИ

Основна

1. *Грубер Й.* Економетрія: Вступ до множинної регресії та економетрії: В 2 т. — К.: Нічлава, 1998. — Т. 1: Вступ до економетрії. — 384 с.; 1999. — Т. 2. — 308 с.
2. *Доугерти К.* Введение в эконометрику. — М.: ИНФРА-М, 1997. — 402 с.
3. *Корольов О. А.* Економетрія. Навч. посіб. — К.: КНТЕУ, 2000. — 660 с.
4. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика: Підручник. — К.: Знання; КОО, 1998. — 494 с.
5. *Магнус Я. Р., Катышев П. К., Пересецкий А. А.* Эконометрика. Нач. курс. — М.: Дело, 1998. — 248 с.
6. *Наконечний С. І., Терещенко Т. О., Романюк Т. П..* Економетрія. — Вид. 2-ге, допов. та переробл.: Підручник. — К.: КНЕУ, 2000. — 296 с.

Додаткова

1. *Айвазян С. А., Мхитарян В. С.* Прикладная статистика и основы эконометрики: Учеб. для вузов. — М.: ЮНИТИ, 1998. — 1022 с.
2. *Джонстон Дж.* Эконометрические методы. — М.: Статистика, 1980. — 444 с.
3. *Дрейпер С.* Прикладной регрессионный анализ: В 2 т. — М.: Мир, 1988.
4. *Катышев П. К., Пересецкий А. А.* Сборник задач к начальному курсу эконометрики. — М.: Дело, 1999. — 72 с.
5. *Маленво Э.* Статистические методы в эконометрии. — Вып. 1. — М.: Статистика, 1975. — 423 с.; Вып. 2. — 1976. — 325 с.
6. *Винн Р., Холден К.* Введение в прикладной эконометрический анализ. — М.: Финансы и статистика, 1981. — 294 с.
7. *Клас А., Гергели К., Колек Ю., Шуян И.* Введение в эконометрическое моделирование. — М.: Статистика, 1978. — 152 с.
8. *Тинтнер Г.* Введение в эконометрию. — М.: Статистика, 1965. — 361 с.
9. *Клейнер Б. Г.* Производственные функции. — М.: Финансы и статистика, 1995.

10. *Алмон Д.* Система функций потребления и ее оценка для Бельгии // Экономика и матем. методы. — 1978, XIV. — Вып. 3. — С. 480–502.
11. Рабочая книга по прогнозированию / Отв. ред. И. В. Бестужев-Лада. — М.: Мысль, 1982. — 430 с.
12. *Хазанова Л. Э.* Математическое моделирование в экономике: Учеб. пособие. — М.: БЕК, 1998. — 141 с.
13. *Дубров А. М., Лагоша Б. А., Хрусталев Е. Ю.* Моделирование рисковых ситуаций в экономике и бизнесе: Учеб. пособие. — М.: Финансы и статистика, 1999. — 176 с.
14. *Жданов С. А.* Экономические модели и методы в управлении. — М.: Дело и сервис, 1998. — 176 с.
15. *Єлейко В. М.* Основи економетрії. — Львів: Марка Лтд, 1995. — 191 с.
16. *Емельянов А. С.* Эконометрия и прогнозирование — М.: Экономика,1985. — 117 с.

ЗМІСТ

Пояснювальна записка 3
Контрольні завдання 3
Список літератури 28

Відповідальний за випуск *А. Д. Вегеренко*
Редактор *О. М. Коваленко*
Комп'ютерне верстання *О. М. Каденко*

Зам. № ВКЦ-3304

Міжрегіональна Академія управління персоналом (МАУП)
03039 Київ-39, вул. Фрометівська, 2, МАУП