

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ МОЛОДІ
ТА СПОРТУ УКРАЇНИ**



**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
АВТОМОБІЛЬНО - ДОРОЖНІЙ УНІВЕРСИТЕТ**

***МЕТОДИЧНІ ВКАЗІВКИ
ДО ВИКОНАННЯ ПРАКТИЧНИХ РОБІТ
З ДИСЦИПЛІНИ “СТАТИСТИКА”
ЧАСТИНА 1***

**для студентів спеціальності
7.030504 – Економіка підприємства
(денна форма навчання)**

Харків 2012

Методичні вказівки до виконання практичних робіт з дисципліни „Статистика” (для студентів спеціальності 7.030504 - „Економіка підприємства” (денна форма навчання)) / Укладачі: Бредіхін В.М., Вербицька В.І, Першина Г.О. - Харків: РІО ХНАДУ, 2012. - ____ с.

Укладачі: В.М. Бредіхін
В.І. Вербицька
Г.О. Першина

Кафедра економіки підприємства

Методичні вказівки складено згідно з вимогами державної освітньо-професійної програми бакалаврів з напрямку «Економіка підприємства».

ПРАКТИЧНЕ ЗАНЯТТЯ 1

АНАЛІЗ СТАТИСТИЧНИХ ІНДЕКСІВ

Мета – здобути навички знаходження статистичних індексів, індивідуальних та агрегатних.

Завдання 1

З метою визначення зміни рівня цін на деякі товари було проведено статистичне дослідження. Потрібно розрахувати:

- 1) індивідуальні індекси цін на окремі товари;
- 2) індивідуальні індекси фізичного об'єму окремих товарів;
- 3) агрегатний індекс цін на товари;
- 4) агрегатний індекс фізичного об'єму товарообігу;
- 5) загальний індекс товарообігу.

Дані для розрахунку статистичних індексів приведені в табл.1. Розрахунок індексів виробити згідно індивідуальному варіанту завдання по заліковій книжці (i - остання цифра залікової книжки; j - передостання цифра). Результати розрахунків оформити в табличному вигляді (табл.2).

За результатами роботи зробити загальний висновок.

Таблиця 1 – Результати статистичного дослідження

Вид товару	Ціна од. продукції, грн.			Кількість проданого товару, шт.		
	Базисний період	Звітний період	$(i;j)$	Базисний період	Звітний період	$(i;j)$
	p_0	p_1		q_0	q_1	
Молоко, л	1,8	2,2	$+0,3i$	800	700	$+9i$
М'ясо, кг	15	17	$+0,5(i+j)$	120	100	$+5(i+j)$
Пальто, шт.	260	300	$+5j$	50	40	$+3j$
Костюми, шт	410	450	$+7(i+j)$	60	70	$+4(i+j)$

Таблиця 2 – Форма для здійснення розрахунків

Вид товару	Ціна од. продукції, грн.		Кількість проданого товару, шт.		Вартість проданих товарів, грн.		
	Базисний період	Звітний період	Базисний період	Звітний період	Базисний період	Звітний період	У звітному періоді в цінах базисного періоду
	p_0	p_1	q_0	q_1	$q_0 p_0$	$q_1 p_1$	$q_1 p_0$
Молоко, л							
М'ясо, кг							
Продовольча група	-	-	-	-			
Пальто, шт.							
Костюми, шт							
Непродовольча група	-	-	-	-			
ЗАГАЛОМ	-	-	-	-			

Вказівки до виконання завдання

1. Для визначення зміни рівня цін не окремі товари використовуються індивідуальні індекси цін і відповідно для характеристики зміни фізичного об'єму - індивідуальні індекси фізичного об'єму.

Індивідуальні індекси цін на окремі товари розраховується за наступною формулою:

$$i_p = \frac{p_1}{p_0} \quad (1.1)$$

Індивідуальні індекси фізичного об'єму окремих товарів визначаються так:

$$i_q = \frac{q_1}{q_0} \quad (1.2)$$

2. Для характеристики зміни рівня цін і фізичного об'єму товарообігу використовуються загальні індекси, розрахунок яких виробляється по агрегатних формулах.

Агрегатний індекс цін розраховується за формулою 1.3

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} \quad (1.3)$$

Агрегатний індекс фізичного об'єму товарообігу знаходиться так:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} \quad (1.4)$$

3. Загальний індекс товарообігу розраховується за формулою 1.5

$$I_{qp} = \frac{\sum q_1 p_1}{\sum q_0 p_0} \quad (1.5)$$

4. Зробити висновок про виконану роботу.

Завдання 2

Є дані про кількість продукції і витрати на 1 ц за планом і фактично (табл. 3).

Таблиця 3 – Відомості про кількість продукції і витрати на 1 ц за планом і фактично

Продукція	Кіл-сть продукції (q), тис. грн.	Кіл-сть праці (l) на 1 ц, чол.-год.	
		план	факт
Картопля	15,6	3,7	3,0
Цукровий буряк	64,3	1,2	1,0
Молоко	34,6	10,0	10,0

Необхідно визначити: індивідуальні індекси продуктивності праці; загальний індекс продуктивності праці; абсолютну економію (перевитрату) витрат праці.

Вказівки до виконання завдання

1) Визначаємо індивідуальні індекси продуктивності праці: по картоплі; по цукровому буряку; по молоку за формулою 1.6

$$i_{li} = \frac{l_{1i}}{l_{0i}} \quad (1.6)$$

2) Далі визначаємо загальний індекс продуктивності праці

$$i_l = \frac{1}{3}(i_{l1} + i_{l2} + i_{l3}) \quad (1.7)$$

Прим.: такий спосіб розрахунку - найбільш грубий і застосовується, коли немає можливості зв'язати вироблення за окремими показниками з відповідною кількістю продукції, оскільки не вказані ціни на центнер продукції.

3) Розрахуємо абсолютну економію (або перевитрату) витрат праці за кожним видом продукції, тобто пов'яжемо зростання продуктивності праці з грошовим показником:

$$e = q_i \left(1 - \frac{1}{i_{li}}\right) \quad (1.8)$$

- по картоплі: економія в (3,7 / 3,0) - в 1,23 разу, в рублях:

4) На підставі проведених розрахунків формулюються висновки.

ПРАКТИЧНЕ ЗАНЯТТЯ 2

ДОСЛІДЖЕННЯ ЗАЛЕЖНОСТІ РІЧНОГО ВАЛОВОГО ДОХОДУ ПІДПРИЄМСТВА ВІД ТЕХНІКО- ЕКОНОМІЧНИХ ПОКАЗНИКІВ І РОЗРОБКА ЙОГО СТАТИСТИЧНОЇ МОДЕЛІ

Мета – провести дослідження залежності річного валового доходу підприємства від техніко-економічних показників і розробити його статистичної моделі

Завдання

Підприємству в процесі розробки бізнес-плану необхідно розрахувати валовий дохід від різних варіантів здійснення виробничої діяльності. Метод прямого розрахунку цього показника в силу низки причин використовувати неможливо. Пропонується провести дослідження залежності річного валового доходу підприємства від техніко-економічних показників і розробити його статистичну модель, що дозволить швидко і якісно прогнозувати значення валового доходу на підставі факторів, які визначають його величину.

Передбачається, що на річний валовий дохід підприємства (Y) можуть здійснювати значний вплив наступні фактори: середньорічна вартість виробничих фондів (X_1), середньорічна чисельність працюючих (X_2); фондівіддача (X_3); фондоозброєність (X_4); продуктивність праці (X_5).

Для аналізу залежності і побудови моделі необхідно:

1) одержати випадкову 10-відсоткову вибірку з генеральної сукупності, представленої 360 підприємствами;

2) зробити розрахунок по кожному відібраному для дослідження підприємств значень показники фондівіддачі (X_3), фондоозброєності (X_4) та продуктивності праці (X_5);

3) розрахувати для всіх показників описову статистику, побудувати гістограму та визначити закон розподілу результативної перемінної, перевірити вибірку на присутність аномальних спостережень (при необхідності виключити відповідні підприємства з подальшого дослідження);

4) перевірити достатність обсягу вибірки для одержання достовірних результатів;

5) провести парний кореляційно-регресійний аналіз залежностей $Y = \varphi(X_1)$, $Y = \varphi(X_2)$, $Y = \varphi(X_3)$, $Y = \varphi(X_4)$, $Y = \varphi(X_5)$, відзначити можливість прогнозування Y на підставі парних залежностей, заповнити матрицю парних коефіцієнтів кореляції і вибрати з її використанням два-три фактора, що мають найменшу кореляцію між собою, але найвищу кореляцію з результативним показником Y ;

б) провести багатофакторний кореляційно-регресійний аналіз залежності Y від двох-трьох факторів, відібраних з комплексу досліджуваних (X_1, X_2, X_3, X_4, X_5);

7) розробити рекомендації з використання багатофакторної регресійної моделі на практиці (розробити приклад використання моделі для планування).

Вказівки до виконання завдання

За умовами завдання вихідні дані для рішення поставленої задачі варто одержати з генеральної сукупності, представленої 360 підприємствами (Додаток А). За розміром досліджуваного результативного показника (Y) вона розділена на 10 типових груп.

Для формування репрезентативної вибірки необхідний механізм випадкового відбору об'єктів з генеральної сукупності. В завданні в основу даного механізму було закладено передостанню й останню цифри номера залікової книжки студента.

Добір вихідних даних виробляється в такий спосіб:

за передостанньою цифрою залікової книжки встановлюється номер першого досліджуваного об'єкта в кожній типовій групі;

за останньою цифрою - періодичність добору об'єктів у цій групі.

Добір здійснюється пропорційно розмірам типових груп у генеральній сукупності (табл. 4). Вибіркова сукупність повинна складатися з 36 підприємств ($n = 36$).

Таблиця 4 - Структура генеральної та вибіркової сукупностей

Типичні групи	Генеральна сукупність		Вибіркова сукупність	
	Обсяг, N_j	Частка, %	Обсяг, n_j	Частка, %
1	2	3	4	5
I	10	2,78	1	2,78
II	20	5,56	2	5,56
III	40	11,11	4	11,11
IV	50	13,89	5	13,89
V	60	16,67	6	16,67
VI	60	16,67	6	16,67

1	2	3	4	5
VII	50	13,89	5	13,89
VIII	40	11,11	4	11,11
IX	20	5,56	2	5,56
X	10	2,78	1	2,78
Всього	360	100,0	36	100,0

Приклад. Останні цифри залікової книжки 08. Звідси в кожній типовій групі добір об'єктів варто починати з об'єкта під номером 10 і далі відбирати кожен наступний восьмий об'єкт. При цих умовах буде отримана вибірка, в яку з кожної групи будуть відібрані підприємства з наступними порядковими номерами в групах (табл.5).

Таблиця 5 - Вибіркова сукупність (приклад для варіанта 08)

Група	Номера об'єктів в групі	Кількість об'єктів в групі
I	10	1
II	10,18	2
III	10, 18, 26, 34	4
IV	10, 18, 26, 34, 42	5
V	10, 16, 26, 34, 42, 50	6
VI	10, 18, 26, 34, 42, 50	6
VII	10, 18, 26, 34, 42	5
VIII	10, 18, 26, 34	4
IX	10, 18	2
X	10	1
Всього		36

По кожному обраному підприємству в Додатку А приведено три показники, значення яких слід вписати в таблицю 3. Далі на підставі цих первинних показників потрібно розрахувати похідні показники. До них, відповідно до завдання, відносяться показники фондоддачі (x_3), фондоозброєності (x_4) та продуктивності праці (x_5). Їх розраховують для кожного підприємства, яке потрапило до вибірки, наступним чином

Фондовіддача	$X_3 = Y / X_1$,	грн./грн.
Фондоозброєність	$X_4 = X_1 / X_2$	тис.грн./чол.
Продуктивність праці	$X_5 = Y / X_2$	тис.грн./чол.

Розрахунки можуть бути виконані на ЕОМ з використанням табличного процесора Excel. Значення перемінних X_{i3} , X_{i4} , X_{i5} повинні мати 4 цифри після десяткової коми.

Результати розрахунків потрібно привести в табл.6.

Таблиця 6 - Вихідні данні до статистичного дослідження (приклад для варіанта 08)

Номер за порядком			Значення перемінних					
Групи	Підприємства в групі	Спостереження, i	Y_i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}
I	10	1	1075	1155	456	0,9307	2,5329	2,3575
II	10	2	1650	1673	6801	0,9863	2,4603	2,4265
III	18	3	1964	1689	701	1,1628	2,4094	2,8017
...
X	10	36	9994	8012	2099	1,2474	3,8171	4,7613
Сума								
Середнє значення								

Таблиця з вихідними даними не повинна містити розрахункових помилок і описок, оскільки вони можуть позначитися на остаточних результатах дослідження. В зв'язку з цим необхідно стежити за правильністю всіх обчислень і записів.

Далі необхідно виконати групування даних, яке виробляється з метою аналізу структури і закономірностей розподілу досліджуваних показників.

1. Групування з використанням рівних інтервалів

В випадку групування з використанням рівних інтервалів можуть бути розглянуті два способи.

Перший спосіб. Базується на попередньому розрахунку кількості групувальних інтервалів по формулі Стреджеса.

Другий спосіб. Бере за основу попередній розрахунок величини середнього квадратичного відхилення досліджуваного показника.

Перший спосіб використовують студенти, у яких остання цифра залікової книжки парна (0, 2, 4, 6, 8). Другий спосіб використовують студенти, у яких остання цифра залікової книжки непарна (1, 3, 5, 7, 9).

Спосіб 1. Групування значень досліджуваного показника по методу Стреджеса

Визначається кількість інтервалів групування за формулою Стреджеса

$$K = 1 + 3,322 \cdot Lg(n),$$

де n - кількість спостережень (обсяг вибірки);

$Lg(n)$ - десятковий логарифм числа n .

Отримане значення K потрібно округлити до цілого в більшу сторону числа. Після цього розрахувати ширину групувального інтервалу

$$\Delta Y = (Y_{\max} - Y_{\min}) / K,$$

де Y_{\max} - максимальне значення показника, що вивчається, у вибірці;

Y_{\min} - мінімальне значення показника, що вивчається, у вибірці.

Значення ΔY округляють відповідно до правил, зазначених в [10].

Після чого встановлюють межі інтервалів групування:

- нижня межа першого групувального інтервалу

$$a_1 = Y_{\min} - 0,5 \cdot \Delta Y$$

- верхня межа першого групувального інтервалу

$$b_1 = a_1 + \Delta Y$$

- межі наступних інтервалів встановлюють згідно з наступним правилом: нижня межа чергового інтервалу дорівнює верхній межі попереднього інтервалу, а верхня межа дорівнює

нижня плюс ширина групувального інтервалу. Наприклад, для другого інтервалу межі будуть наступними: $a_2 = b_1$, $b_2 = a_2 + \Delta Y$. В результаті весь діапазон зміни значень перемінної розбивається на K рівних за розміром інтервалів.

Одночасно з установами меж групувальних інтервалів задають умови віднесення спостережень на інтервал. Їх задають у вигляді подвійної нерівності

$$a_k < Y \leq b_k, \quad k = 1, 2, 3, \dots, K$$

Відповідно до цієї умови на інтервал з номером k відносять ті значення досліджуваної перемінної, які перевершують нижню межу і більше або дорівнюють верхній межі.

Далі варто розподілити одиниці вибіркової сукупності (підприємства) по інтервалах в залежності від величини результативної ознаки. Для цього рекомендується скласти таблицю наступного виду (табл.4).

Якщо в процесі віднесення даних спостережень на інтервал максимальне значення ознаки не попадає на останній інтервал, то можна створити ще один додатковий інтервал або зробити межу останнього інтервалу відкритою ($Y > a_k$).

Після рознесення даних по інтервалах, у табл.7 підраховується частота попадання спостережень на інтервал (f_k), розраховуються частоти ($\omega_k = f_k / n$), визначаються кумулятивні частоти ($\sum f_k = f_1 + f_2 + \dots + f_k$) і частоти ($\sum \omega_k = \omega_1 + \omega_2 + \dots + \omega_k$).

Таблиця 7 – Групування підприємств за розміром валового доходу, тис.грн.

Номер інтервала, k	Межі інтервалі в	Частота, f_j	Накопичена частота, $\sum f_j$	Частість, ω_o	Накопичена частість, $\sum \omega_o$
1	2	3	4	5	6
1	$a_1 < Y \leq b_1$	f_1	f_1	ω_1	ω_1
2	$a_2 < Y \leq b_2$	f_2	$f_1 + f_2$	ω_2	$\omega_1 + \omega_2$
K	$a_k < Y \leq b_k$	f_k	$f_1 + f_2 + \dots + f_k$	ω_k	$\omega_1 + \omega_2 + \dots + \omega_k$
Всього	-	$f_1 + f_2 + \dots + f_k = n$	-	$\omega_1 + \omega_2 + \dots + \omega_k = 1$	-

Спосіб 2. Групування з використанням середнього квадратичного відхилення

Спочатку по незгрупованим даним обчислюють середню (\bar{Y}) і середнє квадратичне відхилення σ_y . Розрахункові формули приведені в табл. 8.

Потім, починаючи від середнього значення, рухаються спочатку вліво, а потім вправо та утворюють групувальні інтервали. При цьому, якщо потрібно розподілити одиниці сукупності по 6 групам використовують значення σ ; якщо потрібно розподілити одиниці сукупності по 9 групам, використовують значення $\frac{2}{3}\sigma$; якщо потрібно розподілити одиниці сукупності по 12 групам використовують значення $0,5\sigma$.

Наприклад, границі інтервалів для розподілу об'єктів сукупності по 6 групам будуть наступними:

Таблиця 8 – Розрахунок інтервалів

$$\begin{array}{lll} 1) \bar{Y} - 3\sigma_y < Y \leq \bar{Y} - 2\sigma_y & 2) \bar{Y} - 2\sigma_y < Y \leq \bar{Y} - \sigma_y & 3) \bar{Y} - \sigma_y < Y \leq \bar{Y} \\ 4) \bar{Y} < Y \leq \bar{Y} + \sigma_y & 5) \bar{Y} + \sigma_y < Y \leq \bar{Y} + 2\sigma_y & 6) \bar{Y} + 2\sigma_y < Y \leq \bar{Y} + 3\sigma_y \end{array}$$

Далі варто скласти таблицю аналогічну таблиці 5 і зробити в ній підрахунок частоти влучення спостережень на інтервал, розрахувати частоти, кумулятивні частоти і частоти

2. Групування з використанням нерівних інтервалів

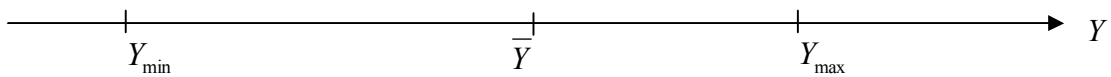
Групування на нерівні інтервали застосовуються для опису статистичних даних, що мають явну асиметрію розподілу частот і частостей. Ширину і межі цих інтервалів встановлюють на основі логічного аналізу попередніх зведень про якісні і кількісні характеристики досліджуваного явища.

В економіці широко застосовуються групування економічних об'єктів з поділом їх на великі, дрібні та середні. Однак такий розподіл залежить від галузевих особливостей і неминуче пов'язаний з низкою умовностей. Наприклад, хлібозавод з чисельністю працівників більш 300 чоловік вважається великим, в той час, як машинобудівний завод з

такою ж чисельністю працівників може бути віднесений до групи дрібних або середніх підприємств.

В практичній роботі в якості одного з можливих рішень задачі групування підприємств по розміру валового доходу рекомендується використовувати досить просту формалізовану процедуру розподілу підприємств на групи.

Процедура виділення груп об'єктів з нерівними інтервалами досліджуваної ознаки наступна. Необхідно проранжувати значення ознаки. Потім весь інтервал його можливих значень $[Y_{\min}; Y_{\max}]$ слід розділити на два інтервали, що відокремлені друг від друга середнім значенням показника \bar{Y} .



На першому інтервалі $[Y_{\min}; \bar{Y}]$ будуть розташовані значення досліджуваної ознаки менше середнього значення \bar{Y} . На другому інтервалі $[\bar{Y}; Y_{\max}]$ будуть розташовані значення досліджуваної ознаки більше, ніж середнє значення \bar{Y} .

У випадку асиметричного розподілу точка, що відповідає середньому значенню ознаки \bar{Y} , не буде поділяти інтервал $[Y_{\min}; Y_{\max}]$ на рівні частини, а буде зміщена до якого-небудь з кінців інтервалу.

Вибираємо з двох інтервалів, розділених значенням середньої величини \bar{Y} , інтервал найменшої довжини, для чого порівнюємо по модулю величини $|Y_{\min} - \bar{Y}|$ і $|\bar{Y} - Y_{\max}|$. Довжину найменшого з двох порівнюваних інтервалів поділяємо навпіл і отримане значення ΔY додаємо до середнього \bar{Y} і віднімаємо від нього. Одержуємо координати двох точок $\bar{Y} - \Delta Y$ і $\bar{Y} + \Delta Y$, які відзначаємо на числовій осі варіаційного ряду вліво і вправо від середнього значення \bar{Y} .



В результаті числова вісь, що відповідає проранжованому варіаційному ряду досліджуваної ознаки, розділяється на три

інтервали $[Y_{\min}; \bar{Y} - \Delta Y]$, $[\bar{Y} - \Delta Y; \bar{Y} + \Delta Y]$ та $[\bar{Y} + \Delta Y; Y_{\max}]$, довжини яких можуть бути інтерпретовані як величини що відмежовують дрібні, середні і великі одиниці сукупності.

При досить великій величині розмаху варіації досліджуваної ознаки процедура дроблення всієї числової осі може бути повторена, у результаті чого будуть визначені границі груп самих дрібних, середніх, більше середнього і самих великих об'єктів.

Після встановлення границь інтервалів варто розробити таблицю частот і частостей, побудувати гістограму розподілу підприємств.

Результати групувань з використанням рівних і нерівних інтервалів варто проаналізувати і зробити попередній висновок щодо закономірності розподілу підприємств за розміром досліджуваного показника. Крім того, необхідно вказати спосіб групування, що дав більш чітке представлення щодо закономірностей розподілу досліджуваного показника.

3. Розрахунок узагальнюючих характеристик вибіркової сукупності

Наступним етапом аналізу сукупності спостережень є розрахунок узагальнюючих характеристик (описової статистики) досліджуваної статистичної сукупності. Розрахунок описової статистики можна виконати по незгрупованим або згрупованим даним (на підставі частотної таблиці). Більш точними є результати, отримані з використанням незгрупованих даних.

В практичній роботі розрахунок зазначених показників описової статистики слід виконати для кожної перемінної по незгрупованим даним. Для розрахунку рекомендується використовувати наступні формули (табл.9).

Примітка: i - номер спостереження, $i = 1, 2, 3, \dots, n$; j - номер фактора, $j = 1, 2, 3, 4, 5$

Якщо розрахунок виконується без застосування ЕОМ і обсяг вибірки $n > 30$, то для прискорення значення σ_y , σ_{x_j} можна обчислити по формулах $\sigma_y = \sqrt{Y^2 - (\bar{Y})^2}$ та $\sigma_{x_j} = \sqrt{X_j^2 - (\bar{X}_j)^2}$.

Розрахунок виконується на підставі вихідних даних, приведених у табл.3. Результати розрахунків рекомендується

представити в окремій таблиці. Також слід провести їхній аналіз і зробити висновок про однорідність досліджуваної сукупності об'єктів. Чим менше значення коефіцієнта варіації, тим однорідніше об'єкти досліджуваної сукупності і надійніше рішення, прийняті з використанням описової статистики.

Таблиця 9 - Формули для розрахунку узагальнюючих показників вибірових сукупностей

Узагальнюючі показники	Результативна перемінна, Y	Факторні перемінні, X_j
Середнє	$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$	$\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij}$
Середнє квадратичне відхилення	$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}}$	$\sigma_{X_j} = \sqrt{\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n-1}}$
Коефіцієнт варіації	$K_v = \frac{\sigma_y}{\bar{Y}} \cdot 100\%$	$K_{vj} = \frac{\sigma_{X_j}}{\bar{X}_j} \cdot 100\%$

Неоднорідні сукупності характеризуються великими значеннями коефіцієнтів варіації, що може бути наслідком присутності у вибірках аномальних спостережень. Значення перемінних, що різко виділяються, прийнято вважати аномальними. Об'єкти з такими значеннями перемінних потрібно виключати з вибірки, оскільки, як правило, вони мають іншу структуру або/і зовнішні умови існування (відмінні від інших об'єктів, що потрапили у вибірку), і далі ці об'єкти вивчаються окремо. Присутність аномальних спостережень у вибірках робить помилковими висновки про досліджуване явище. Перевірка значень перемінних на присутність аномальних спостережень може бути виконана по різних методиках. В курсовій роботі можна обмежитися правилом „три-сигма”. Відповідно до цього правила значення перемінної вважається аномальним, якщо воно виходить за межі припустимого інтервалу:

$$\bar{Y} - 3\sigma_y \leq Y \leq \bar{Y} + 3\sigma_y \quad \text{і} \quad \bar{X}_j - 3\sigma_{X_j} \leq X_j \leq \bar{X}_j + 3\sigma_{X_j}$$

Об'єкти, що мають значення однієї і більш перемінних, які різко виділяються, варто виключити з вибірки і перерахувати значення середніх, дисперсії і середніх квадратичних відхилень. Подальший аналіз проводиться по зменшеній вибірці.

4. Оцінка достатності обсягу вибірки

Обсяг вибірки повинен бути достатнім для одержання достовірних висновків про досліджуване явище. В зв'язку з цим у практичній роботі варто визначити, яким повинен бути обсяг вибірки для проведення дослідження. З цією метою варто розрахувати мінімальний обсяг виборки, потрібний для оцінки генеральної середньої результативного показника з 5-відсотковою помилкою на рівні довірчої імовірності 0,95. Розрахунок обсягу виборки виконується по формулі для безперервної власне випадкової вибірки.

В результаті розрахунків може бути виявлено, що фактичний обсяг вибірки більше або менше мінімального. Якщо фактичний обсяг вибірки більше або менше мінімального, то, використовуючи формулу для розрахунку обсягу вибірки, варто перерахувати величину помилки оцінки генеральної середньої. Для цього у формулу, по якій вироблявся розрахунок обсягу вибірки варто підставити фактичне значення обсягу вибірки, значення дисперсії ознаки і коефіцієнта довіри, і знайти, таким чином, відповідне їм значення помилки оцінки генеральної середньої.

5. Парний кореляційно-регресійний аналіз залежностей

5.1 Кореляційний аналіз парних зв'язків $Y = \varphi(X_j)$

Кореляційний аналіз проводять з метою оцінки сили й істотності залежності результативної перемінної від факторних перемінних. Якщо факторних перемінних небагато (у загальному випадку m), то на першому етапі дослідження проводять аналіз залежності результативної перемінної Y від кожної факторної перемінної

$$Y = \varphi(X_j), \quad j = 1, 2, \dots, m.$$

Таким чином, у практичній роботі необхідно провести кореляційний аналіз залежності валового доходу підприємства від середньорічної вартості основних фондів $Y = \varphi(X_1)$, від

середньоспискової чисельності працюючих $Y = \varphi(X_2)$, від фондівіддачі $Y = \varphi(X_3)$, від фондоозброєності $Y = \varphi(X_4)$, і від продуктивності праці $Y = \varphi(X_5)$.

Для виявлення наявності залежності однієї перемінної від іншої необхідно побудувати кореляційне поле і розрахувати коефіцієнт парної кореляції. Рекомендації з виконання цієї частини роботи наводяться в літературі по дисципліні [1-3, 7-12]. Оскільки в літературі для розрахунку коефіцієнта парної кореляції приводяться різні формули, то в курсовій роботі рекомендується використовувати наступну

$$r_{YX_j} = \frac{\sigma_{YX_j}}{\sigma_Y \sigma_{X_j}} = \frac{\overline{YX_j} - \bar{Y} \bar{X_j}}{\sqrt{Y^2 - (\bar{Y})^2} * \sqrt{X_j^2 - (\bar{X_j})^2}}, \quad (j = 1,2,3,4,5)$$

$$\text{де } \overline{YX_j} = \frac{1}{n} \cdot \sum_{i=1}^{i=n} Y_i X_{ij}; \quad \bar{Y}^2 = \frac{1}{n} \sum_{i=1}^{i=n} Y_i^2; \quad \bar{X_j}^2 = \frac{1}{n} \sum_{i=1}^{i=n} X_{ij}^2$$

Розрахунок середніх для добутоків і середніх для квадратів значень досліджуваних перемінних приводиться в таблицях 6 і 7.

Результати кореляційного аналізу представляються в пояснювальній записці. Для кожної пари зв'язків в записці потрібно привести графік кореляційного поля, розрахунок коефіцієнта парної кореляції і висновки щодо істотності зв'язку перемінних.

5.2 Регресійний аналіз парних зв'язків $Y = \varphi(X_j)$

Методика регресійного аналізу викладена в літературі [1-3, 7-12]. В практичній роботі потрібно представити найбільш важливі результати регресійного аналізу. Зокрема для кожної пари зв'язків $Y = \varphi(X_1)$, $Y = \varphi(X_2)$, $Y = \varphi(X_3)$, $Y = \varphi(X_4)$, $Y = \varphi(X_5)$ повинні бути дані відповіді на наступні запитання:

1) яка форма зв'язку (пряма або зворотня, лінійна або нелінійна) має місце між Y і кожною з досліджуваних факторних перемінних X_j ;

2) яке рівняння регресії найкраще описує залежність між Y та X_j ;

3) чи є це рівняння статистично значимим.

Таблиця 10 – Розрахунок середньої суми квадратів значень перемінних та суми добутків

i	Y_i^2	X_{i1}^2	X_{i2}^2	X_{i3}^2	X_{i4}^2	X_{i5}^2	$Y_i X_{i1}$	$Y_i X_{i2}$	$Y_i X_{i3}$	$Y_i X_{i4}$	$Y_i X_{i5}$
1											
2											
...											
i											
...											
n											
Сума	$\sum Y_i^2$	$\sum X_{i1}^2$	$\sum X_{i2}^2$	$\sum X_{i3}^2$	$\sum X_{i4}^2$	$\sum X_{i5}^2$	$\sum Y_i X_{i1}$	$\sum Y_i X_{i2}$	$\sum Y_i X_{i3}$	$\sum Y_i X_{i4}$	$\sum Y_i X_{i5}$
Середнє	$\overline{Y_i^2}$	$\overline{X_{i1}^2}$	$\overline{X_{i2}^2}$	$\overline{X_{i3}^2}$	$\overline{X_{i4}^2}$	$\overline{X_{i5}^2}$	$\overline{Y_i X_{i1}}$	$\overline{Y_i X_{i2}}$	$\overline{Y_i X_{i3}}$	$\overline{Y_i X_{i4}}$	$\overline{Y_i X_{i5}}$

Таблиця 11 – Розрахунок середньої суми квадратів добутків факторних перемінних

i	Для X_1				Для X_2			Для X_3		Для X_4
	$X_{i1} X_{i2}$	$X_{i1} X_{i3}$	$X_{i1} X_{i4}$	$X_{i1} X_{i5}$	$X_{i2} X_{i3}$	$X_{i2} X_{i4}$	$X_{i2} X_{i5}$	$X_{i3} X_{i4}$	$X_{i3} X_{i5}$	$X_{i4} X_{i5}$
1										
2										
...										
i										
...										
n										
Сума	$\sum X_{i1} X_{i2}$	$\sum X_{i1} X_{i3}$	$\sum X_{i1} X_{i4}$	$\sum X_{i1} X_{i5}$	$\sum X_{i2} X_{i3}$	$\sum X_{i2} X_{i4}$	$\sum X_{i2} X_{i5}$	$\sum X_{i3} X_{i4}$	$\sum X_{i3} X_{i5}$	$\sum X_{i4} X_{i5}$
Середнє	$\overline{X_{i1} X_{i2}}$	$\overline{X_{i1} X_{i3}}$	$\overline{X_{i1} X_{i4}}$	$\overline{X_{i1} X_{i5}}$	$\overline{X_{i2} X_{i3}}$	$\overline{X_{i2} X_{i4}}$	$\overline{X_{i2} X_{i5}}$	$\overline{X_{i3} X_{i4}}$	$\overline{X_{i3} X_{i5}}$	$\overline{X_{i4} X_{i5}}$

Для вибору форми зв'язку потрібно використовувати раніше побудований графік із зображенням кореляційного поля; побудувати емпіричну лінію регресії; по її вигляду визначити, чи є залежність $Y = \varphi(X_j)$ прямою або зворотною; яке рівняння парної регресії доцільно використовувати для апроксимації емпіричної лінії регресії.

Визначення форми зв'язку. Якщо зі збільшенням значень перемінної X_j , значення перемінної Y також збільшуються, то має місце пряма лінійна (рис. 1а) або пряма нелінійна (рис. 1б) залежність. Якщо зі збільшенням значень перемінної X_j , значення перемінної Y зменшуються, то має місце зворотній лінійний (рис.2а), або нелінійний (рис.2б) зв'язок.

При виконанні дослідження варто враховувати, що у всіх випадках, коли має місце нелінійний зв'язок, а сила зв'язку неправомірно намагаються оцінити за допомогою коефіцієнта кореляції, його значення отримуються незначними, а іноді близькими до нуля.

Якщо кореляційне поле „розмите” і не дозволяє виявити закономірності між зміною значень перемінних X_j і Y , то є підстави вважати, що зв'язок між ними відсутній (рис.3).

При виявленні форми зв'язку між перемінними важливо переконатися в тому, що спостереження мають приблизно однакову варіацію щодо передбачуваної теоретичної лінії регресії. У випадку, коли дані розподіляються вздовж лінії регресії нерівномірно (рис.4), вважається що метод найменших квадратів непридатний для підбора лінії регресії.

Вибір рівняння регресії. У випадку, коли між перемінними Y і X_j передбачається прямий або лінійний зв'язок, обирають рівняння вигляду

$$\hat{Y}_i = a + b \cdot X_{ij} + \varepsilon_i,$$

де \hat{Y}_i - теоретичне значення результативної перемінної, розраховане за рівнянням регресії за умови, що об'єкт з номером i має значення факторної перемінної, яке дорівнює X_{ij} ;

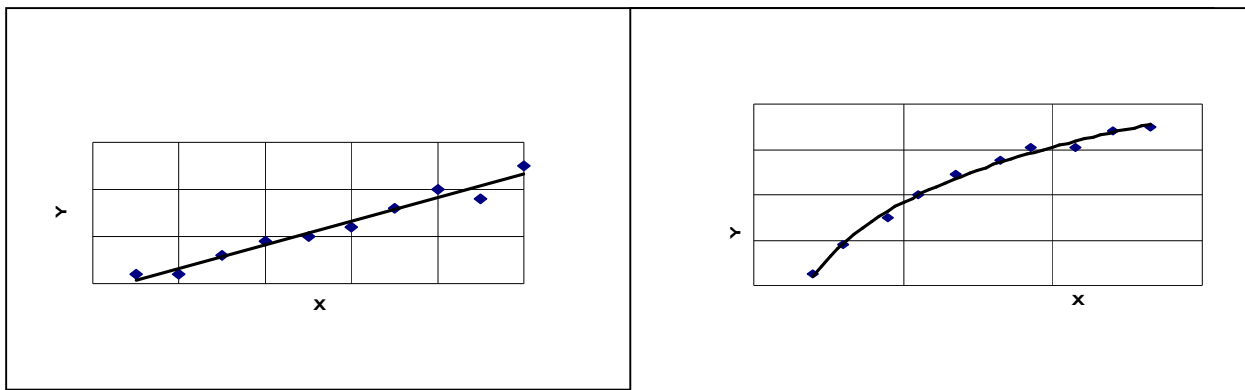
a, b - параметри рівняння;

X_{ij} - значення j -ї факторної перемінної на i -му спостереженні;

ε_i - випадкова похибка результативної перемінної.

Далі розраховують невідомі значення параметрів a і b за даними вибірки.

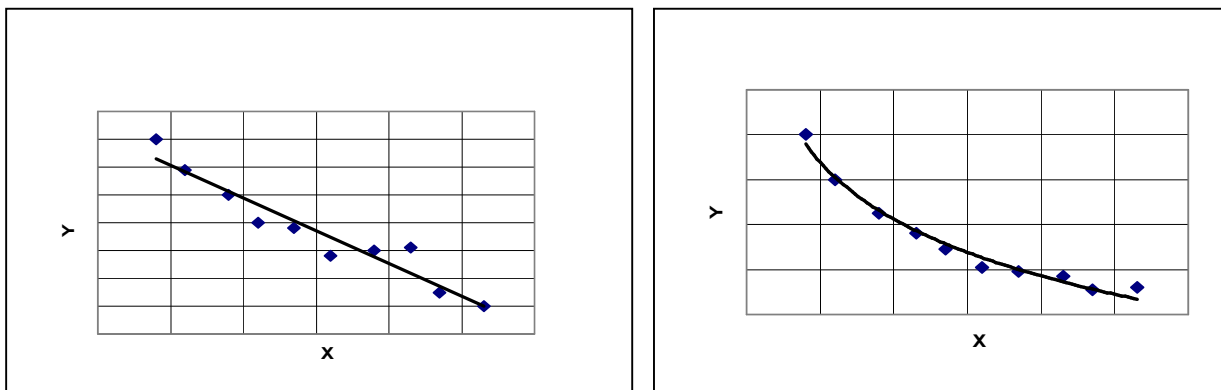
$$b = \frac{n \cdot \sum X_j Y_j - \sum X_j \cdot \sum Y_j}{n \cdot \sum X_j^2 - (\sum X_j)^2}$$



а

б

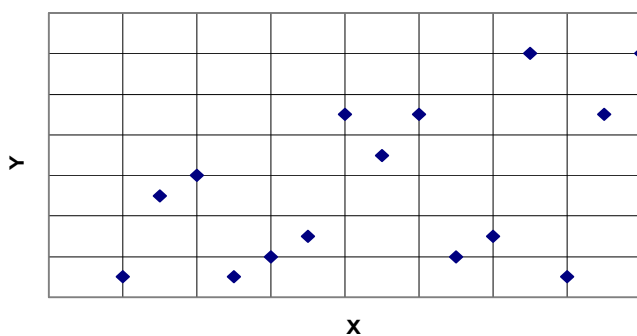
Рисунок 1 - Прямая лінійна (а) і пряма нелінійна (б) кореляційна залежність



а

б

Рисунок 2 - Зворотня лінійна (а) і зворотня нелінійна (б) кореляційна залежність



х

Рисунок 3 – Відсутність кореляційного зв'язку між перемінними

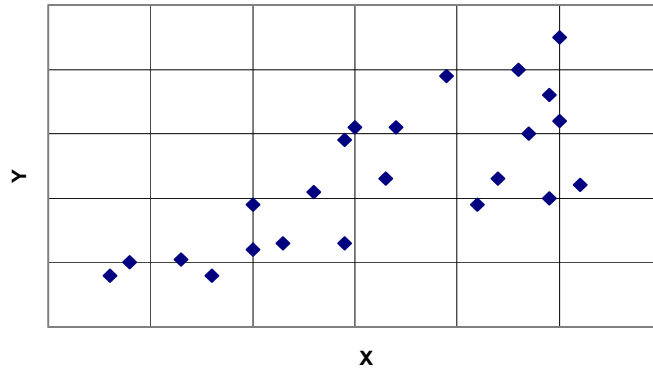


Рисунок 4 – Регресія з нерівномірною варіацією, яка призводить до викривлення оцінок Y

$$a = \frac{\sum X_j^2 \cdot \sum Y_j - \sum X_j \cdot \sum X_j Y_j}{n \cdot \sum X_j^2 - (\sum X_j)^2}$$

або

$$a = \bar{Y} - b \cdot \bar{X}$$

У випадку, коли між перемінними передбачається пряма або зворотня нелінійна залежність, рівняння регресії рекомендується підібрати по Додатку Д. Далі використовуючи прийом лінеаризації вихідних значень перемінних і параметрів (Додаток Д), варто перевести обране рівняння до лінійного виду і по відомих формулах обчислити значення параметрів a і b .

Слід враховувати, що іноді тенденція в розкиді точок в поле кореляції може бути не видна через недолік обраного масштабу зображення. В зв'язку з цим рекомендується використовувати техніку побудови багаторівневих графіків. Вона припускає, крім основного графіка, побудову декількох додаткових графіків, які являють собою вертикально і горизонтально стиснуті зображення основного графіка. Цей стиск графіків полегшує ідентифікацію затіненої (що чітко не виявляється) тенденції розкиду точок поля кореляції двох перемінних.

Приклад. Вивчається залежність валового доходу підприємства (Y) від середньорічної вартості основних виробничих фондів (X_1), середньорічної чисельності працюючих (X_2); фондівіддачі (X_3); фондоозброєності (X_4) та продуктивності

праці (X_5). Для перевірки гіпотези отримані вибіркоким методом (з Додатка А) дані по 19 підприємствам (табл. 12).

Потрібно побудувати кореляційні поля і вибрати рівняння регресії для опису залежностей між досліджуваними показниками.

Таблиця 12 – Вихідні дані до прикладу

i	Y	X_1	X_2	X_3	X_4	X_5
1	1075	1155	456	0,93074	2,35746	2,53289
2	1650	1673	680	0,98625	2,42647	2,46029
3	2718	2205	879	1,23265	3,09215	2,50853
4	3166	1980	862	5,59899	3,67285	2,29698
5	3852	2554	965	1,50822	3,99171	2,64663
6	4490	2982	1064	1,50570	4,21992	2,80263
7	4822	3165	1162	1,52354	4,14974	2,72375
8	4938	3253	1191	1,51798	4,14610	2,73132
9	5484	3704	1337	1,48056	4,10172	2,77038
10	5854	3446	1383	1,69878	4,23283	2,49168
11	5954	3890	1364	1,53059	4,36510	2,85191
12	6262	4111	1381	1,52323	4,53440	2,07683
13	6823	4369	1502	1,56168	4,54261	2,90879
14	6964	4376	1618	1,59141	4,30142	2,70290
15	7464	4890	1720	1,52638	4,33953	2,84302
16	7888	5164	1457	1,52750	5,41386	3,54427
17	8325	5684	1901	1,46464	4,37927	2,99001
18	9167	5993	2077	1,52962	4,41358	2,83541
19	9994	8012	2099	1,24738	4,76131	3,81706
Всього	106890	72606	25099	27,48585	77,44204	53,48530

На рис. 5, 6, 7, 8 і 9 представлені багаторівневі графіки кореляційних полів, результати апроксимації тенденцій у розкіді точок лінійними і нелінійними формами зв'язку. Ці рисунки добре ілюструють вплив масштабу зображення полів кореляції на результати вибору форми зв'язку двох перемінних.

Якщо для опису регресії можуть бути використані кілька рівнянь, то кращим з них звичайно вважається те, що має найбільше значення критерію Фішера (однак можна використовувати й інші критерії, наприклад, найбільше значення

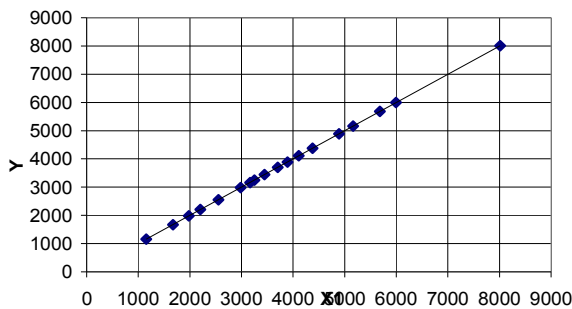
коефіцієнта кореляції, найменшу відносну помилку прогнозування тощо).

Критерій Фішера використовується для перевірки різних гіпотез. У випадку перевірки гіпотези про значущість обраної моделі регресії спочатку розраховують його емпіричне значення.

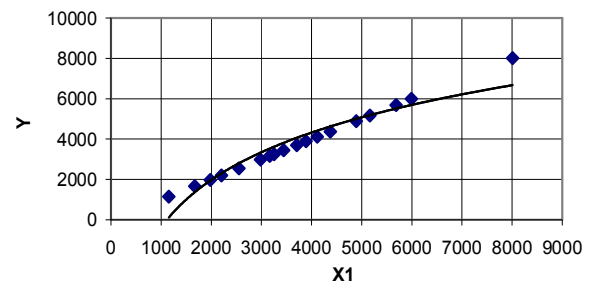
$$F = \frac{\frac{1}{n} \cdot \left(\sum_{i=1}^n (Y_i - \bar{Y})^2 \right)}{\frac{1}{n - m - 1} \cdot \left(\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right)},$$

де m - кількість факторних ознак моделі (для моделі парної регресії $m = 1$).

Це значення порівнюють із критичним значенням, що знаходять по таблицях (Додаток Е) для обраного рівня довірчої імовірності (прийняти рівним 0,95) на перетинанні стовпця, що відповідає числу ступенів волі $df = n - 1$, і рядка, що відповідає числу ступенів волі $df = n - m - 1$.

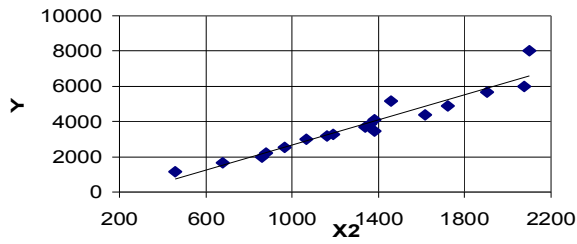


а

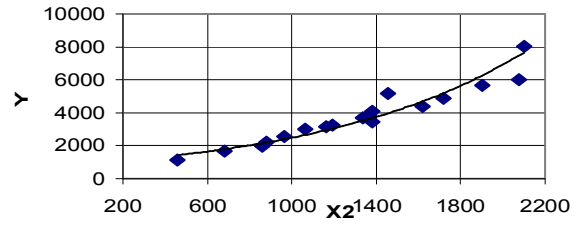


б

Рисунок 5 – Поля кореляції для перемінної X_1 та лінії регресії, які їх апроксимують: а) лінійний варіант; б) нелінійний варіант (логарифмічна крива)

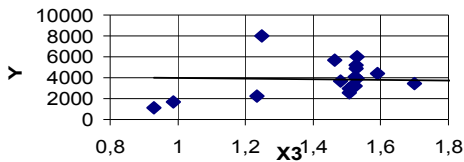


а)

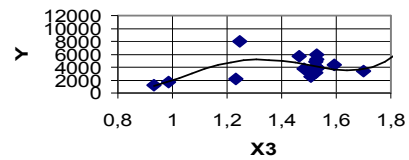


б)

Рисунок 6 - Поля кореляції для перемінної X_2 та лінії регресії, які їх апроксимують: а) лінійний варіант; б) нелінійний варіант (експонента)

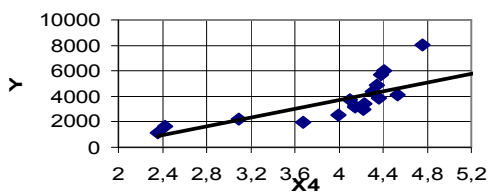


а)

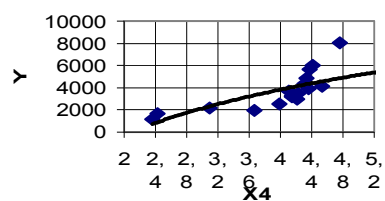


б)

Рисунок 7 - Поля кореляції для перемінної X_3 та лінії регресії, які їх апроксимують: а) лінійний варіант; б) нелінійний варіант (полінома)

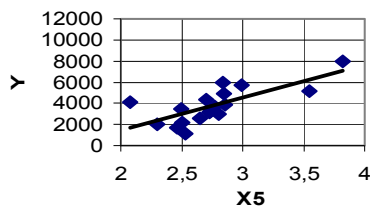


а)

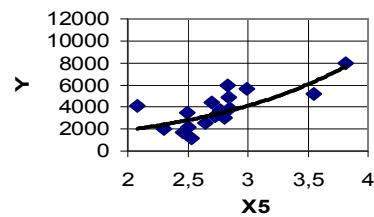


б)

Рисунок 8 - Поля кореляції для перемінної X_4 та лінії регресії, які їх апроксимують: а) лінійний варіант; б) нелінійний варіант (логарифмічна крива)



а)



б)

Рисунок 9 - Поля кореляції для перемінної X_5 та лінії регресії, які їх апроксимують: а) лінійний варіант; б) нелінійний варіант (експонента)

Якщо розрахункове значення більше критичного, то модель вважається значущою на обраному рівні довірчої імовірності. В протилежному випадку модель вважається незначущою. Однак при зниженні рівня довірчої імовірності вона може виявитися статистично значущою.

Розрахунок F для обраних форм зв'язку перемінних варто виконати в таблицях, а хід перевірки гіпотез і зроблені висновки описати в пояснювальній записці.

6. Добір факторів для багатофакторної моделі

Модель множинної регресії повинна включати, фактори, які мають сильний зв'язок з результативною перемінною і не мають сильного зв'язку між собою. Взаємозв'язок факторних перемінних прийнято називати мультиколінеарністю. Мультиколінеарність розглядається як небажане явище.

Добір факторів, що підлягають включенню в модель регресії, рекомендується виконати за спрощеною методикою, не прибігаючи до методів багатокрокового кореляційно-регресійного аналізу.

Для добору факторів, що підлягають включенню в модель регресії, рекомендується розрахувати всі елементи матриці парних коефіцієнтів кореляції (табл. 13).

Матриця парних коефіцієнтів кореляції є симетричною: значення коефіцієнтів кореляції над головною діагоналлю і під нею рівні (тобто $r_{12} = r_{21}$ тощо). Значення елементів, розташованих на головній діагоналі матриці завжди дорівнюють одиниці.

Розташування перемінних у цій матриці може бути й іншим.

Таблиця 13 – Загальний вигляд матриці парних коефіцієнтів кореляції

Перемінні	X_1	X_2	X_3	X_4	X_5	Y
X_1	1,0	r_{12}	r_{13}	r_{14}	r_{15}	r_{1Y}
X_2	r_{21}	1,0	r_{23}	r_{24}	r_{25}	r_{2Y}
X_3	r_{31}	r_{32}	1,0	r_{34}	r_{35}	r_{3Y}
X_4	r_{41}	r_{42}	r_{43}	1,0	r_{45}	r_{4Y}
X_5	r_{51}	r_{52}	r_{53}	r_{54}	1,0	r_{5Y}
Y	r_{Y1}	r_{Y2}	r_{Y3}	r_{Y4}	r_{Y5}	1,0

Заповнення матриці значеннями парних коефіцієнтів кореляції роблять у такий спосіб. Коефіцієнти парної кореляції між результативною перемінною і факторними перемінними обчислені раніше (п.3.1), тому їхні значення записують в останній стовпець і останній рядок матриці. Далі, починаючи з першого рядка, роблять розрахунок відсутніх у матриці коефіцієнтів парної кореляції. Для розрахунку використовується формула вже відома з курсу „Статистика” коефіцієнта парної кореляції між перемінними з порядковими номерами k і l .

$$r_{kl} = \frac{\overline{X_k X_l} - \overline{X_k} \overline{X_l}}{\sqrt{\overline{X_k^2} - (\overline{X_k})^2} * \sqrt{\overline{X_l^2} - (\overline{X_l})^2}}$$

Для розрахунку використовуються суми і середні, обчислені в табл.7 і табл.8. Матрицю парних коефіцієнтів кореляції потрібно проаналізувати і виділити пари факторних перемінних, які мають між собою високий кореляційний зв'язок ($r > 0,8$). У кожній з таких пар виключити з подальшого розгляду ту факторну перемінну, яка має менший коефіцієнт кореляції з результативною перемінною. Для подальшого аналізу доцільно залишити дві-три факторні перемінні.

Приклад. Вивчається залежність валового доходу підприємства (Y) від середньорічної вартості основних виробничих фондів (X_1), середньорічної чисельності працюючих (X_2); фондівіддачі (X_3); фондоозброєності (X_4) та продуктивності

праці (X_5). На підставі раніше приведених вихідних даних (табл.8) розраховані коефіцієнти парної кореляції між усіма перемінними в припущенні, що усі вони мають лінійну форму зв'язку.

Результати розрахунків коефіцієнтів приведені в табл.14.

Таблиця 14 – Вихідна матриця парних коефіцієнтів кореляції (приклад)

Перемінні	Y	X_1	X_2	X_3	X_4	X_5
Y	1,00000					
X_1	0,97185	1,00000				
X_2	0,98190	0,95827	1,00000			
X_3	0,49370	0,30479	0,46313	1,00000		
X_4	0,76283	0,84296	0,67057	0,05252	1,00000	
X_5	0,85150	0,76241	0,76382	0,74584	0,69968	1,00000

Аналіз матриці показує, що між факторними перемінними X_1 і X_2 існує сильний внутрішній кореляційний зв'язок (коефіцієнт кореляції дорівнює 0,95827). З них трохи сильніше впливає на результативний показник факторна перемінна X_2 (коефіцієнт кореляції 0,9819). Тому виключаємо з подальшого розгляду факторну перемінну X_1 .

Матриця парних коефіцієнтів кореляції для змінних, що залишилися, має такий вигляд (табл.15).

Таблиця 15 – Зменшена матриця парних коефіцієнтів кореляції (приклад)

Перемінні	Y	X_2	X_3	X_4	X_5
Y	1,00000				
X_2	0,98190	1,00000			
X_3	0,49370	0,46313	1,00000		
X_4	0,76283	0,67057	0,05252	1,00000	
X_5	0,85150	0,76382	0,74584	0,69968	1,00000

У цій матриці немає факторних перемінних, між якими є тісний кореляційний зв'язок з коефіцієнтом кореляції більш 0,8.

Таким чином, для подальшого дослідження впливу на річний валовий доход підприємства (Y) слід залишити чотири факторні перемінні - середньоспискову чисельність працюючих (X_2), фондівіддачу (X_3), фондоозброєність (X_4) і продуктивність праці (X_5).

7. Багатофакторний кореляційно-регресійний аналіз

7.1 Сутність і загальна процедура побудови моделі

У процесі багатофакторного кореляційного аналізу оцінюється сила зв'язку досліджуваних перемінних, а за допомогою множинного регресійного аналізу для опису форми зв'язку факторів обирається багатофакторна статистична модель.

В практичній роботі потрібно побудувати модель лінійної регресії з m незалежними (факторними) перемінними.

$$Y = a_0 \cdot X_{i0} + a_1 \cdot X_{i1} + a_2 \cdot X_{i2} + \dots + a_m \cdot X_{im} + \varepsilon_i$$

де X_{ij} - факторні перемінні, що спостерігаються на i -ому об'єкті;

i - номер за порядком об'єкта в сукупності, що вивчається,
 $i = 1, 2, \dots, n$;

ε_i - випадкова помилка, яка має математичне очікування 0,
та дисперсію σ^2 ;

X_{i0} - фіктивна перемінна, що дорівнює 1 в усіх спостереженнях.

У цій моделі невідомими є параметри a_{ij} , які підлягають оцінці.

7.2 Розробка моделі й оцінка її статистичної значимості

7.2.1 Обґрунтування форми зв'язку перемінних

Розробка багатофакторної моделі починається з вивчення форми зв'язку, між перемінними. Якщо в результаті буде виявлена можливість використання лінійної моделі, то на наступному етапі виробляється оцінка її параметрів за даними виборки. В зв'язку з цим виклад даного розділу слід розпочати з обґрунтування можливості використання моделі лінійної регресії. При цьому можна використовувати результати, отримані в попередньому розділі й інші приклади з навчальної літератури.

7.2.2 Оцінка параметрів багатofакторної лінійної моделі регресії

При вирішенні задачі як від руки, так і на ЕОМ, параметри лінійної багатofакторної моделі варто розрахувати за методом найменших квадратів (МНК) з використанням матриць.

Оцінка параметрів багатofакторної лінійної моделі регресії методом МНК із використанням матриць. У матричному вигляді лінійна моделі регресії записується в такий спосіб

$$\hat{Y} = Xa + e,$$

де \hat{Y} - вектор теоретичних значень залежної перемінної,

$$\hat{Y} = (Y_i), i = 1, \dots, n;$$

$X = (x_{ij})$ - матриця значень незалежних перемінних розміром $n \times (m + 1)$;

n, m - кількість спостережень та кількість невідомих параметрів;

a - вектор оцінок параметрів моделі, $a = (a_j), j = 0, 1, \dots, m$;

$e = (\varepsilon_i)$ - вектор випадкових помилок, які не залежать від моделі.

Оцінку параметрів цієї моделі виконують методом найменших квадратів. Він дозволяє вибрати параметри моделі таким чином, щоб сума квадратів відхилень значень результативної перемінної (Y_i), що спостерігаються, від значень, обчислених по моделі (\hat{Y}_i), була мінімальною

$$Q = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n e_i^2 \rightarrow \min$$

або в матричному вигляді

$$Q = e^T e = Y^T Y - 2a^T X^T Y + a^T X^T X a \rightarrow \min$$

Якщо продиференціювати Q по a , прирівняти похідну до нуля та зробити інші перетворення, то буде отримано вираз для визначення параметрів моделі

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

Стосовно до лінійної моделі регресії матриці, що використовуються для оцінки невідомих параметрів, мають наступний вигляд

$$X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1m} \\ 1 & X_{21} & X_{22} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nm} \end{bmatrix} \quad X^T X = \begin{bmatrix} n & \sum X_{i1} & \dots & \sum X_{im} \\ \sum X_{i1} & \sum X_{i1}^2 & \dots & \sum X_{i1} X_{im} \\ \dots & \dots & \dots & \dots \\ \sum X_{im} & \sum X_{i1} X_{im} & \dots & \sum X_{im}^2 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} \sum Y_i \\ \sum Y_i X_{i1} \\ \dots \\ \sum Y_i X_{im} \end{bmatrix}$$

Підсумовування виробляється за кількістю спостережень n . Після визначення значень параметрів, слід записати числову модель регресії і дати їм економічну характеристику.

Приклад. На підставі попереднього парного кореляційно-регресійного аналізу встановлена залежність валового доходу підприємства (Y) від чотирьох показників: середньорічної чисельності працюючих (X_2); фондівдачі (X_3); фондоозброєності (X_4) та продуктивності праці (X_5). Усі факторні перемінні не мають мультиколінеарних зв'язків. Передбачається, що багатфакторна модель регресії повинна бути лінійною

$$Y = a_0 \cdot X_{i0} + a_1 \cdot X_{i1} + a_2 \cdot X_{i2} + \dots + a_m \cdot X_{im} + \varepsilon_i$$

Потрібно оцінити параметри цієї моделі методом найменших квадратів.

Для оцінки параметрів використовуємо дані, що раніше розглядалися по 19 підприємствам без фактора X_1 . Вихідні дані для оцінки параметрів моделі регресії приведено в табл.16, доповненої фіктивною перемінною X_0 .

Таблиця 16 - Вихідні дані до приклада оцінки параметрів моделі регресії

i	Y	X_0	X_2	X_3	X_4	X_5
1	2	3	4	5	6	7
1	1075	1	456	0,93074	2,53289	2,35746
2	1650	1	680	0,98625	2,46029	2,42647
3	2718	1	879	1,23265	2,50853	3,09215
4	3166	1	862	5,59899	2,29698	3,67285
5	3852	1	965	1,50822	2,64663	3,99171
6	4490	1	1064	1,50570	2,80263	4,21992
7	4822	1	1162	1,52354	2,72375	4,14974
8	4938	1	1191	1,51798	2,73132	4,14610
9	5484	1	1337	1,48056	2,77038	4,10172
10	5854	1	1383	1,69878	2,49168	4,23283
11	5954	1	1364	1,53059	2,85191	4,36510
12	6262	1	1381	1,52323	2,07683	4,53440
13	6823	1	1502	1,56168	2,90879	4,54261
14	6964	1	1618	1,59141	2,70290	4,30142
15	7464	1	1720	1,52638	2,84302	4,33953
16	7888	1	1457	1,52750	3,54427	5,41386
17	8325	1	1901	1,46464	2,99001	4,37927
18	9167	1	2077	1,52962	2,83541	4,41358
19	9994	1	2099	1,24738	3,81706	4,76131

Оцінимо параметри $a_0, a_1, a_2, a_3, a_4, a_5$ цієї моделі за методом найменших квадратів з використанням техніки матричного рахування [4].

Сформуємо матриці необхідні для оцінки параметрів.

$$X = \begin{vmatrix} 1 & 456 & 0,93074 & 2,53289 & 2,35746 \\ 1 & 680 & 0,98625 & 2,46029 & 2,42647 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 2077 & 1,52962 & 2,88541 & 4,41358 \\ 1 & 2099 & 1,24738 & 3,81706 & 4,76131 \end{vmatrix} \quad Y = \begin{vmatrix} 1075 \\ 1650 \\ \dots \\ 9167 \\ 9994 \end{vmatrix}$$

$$X^T = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 456 & 680 & 879 & \dots & 2077 & 2099 \\ 0,9307 & 0,9863 & 1,2327 & \dots & 1,5296 & 1,2474 \\ 2,5329 & 2,4603 & 2,5085 & \dots & 2,8854 & 3,8171 \\ 2,3575 & 2,4265 & 3,0922 & \dots & 4,4136 & 4,7613 \end{vmatrix} \quad X^T Y = \begin{vmatrix} 106890,0 \\ 160603392,0 \\ 159030,0 \\ 313010,1 \\ 463580,1 \end{vmatrix}$$

$$X^T X = \begin{vmatrix} 19,00 & 25099,00 & 27,49 & 53,49 & 77,44 \\ 25099,00 & 36777927,00 & 37065,52 & 72606,00 & 106890,00 \\ 27,49 & 37065,52 & 40,50 & 77,44 & 114,05 \\ 53,49 & 72606,00 & 77,44 & 152,90 & 221,38 \\ 77,44 & 106890,00 & 114,05 & 221,38 & 325,61 \end{vmatrix}$$

$$\frac{1}{X^T X} = (X^T X)^{-1} = \begin{vmatrix} 233664552 & 0,004414 & -165980835 & -88326984 & 61167152 \\ 0,004414 & 0,000001 & -0,002651 & -0,001629 & 0,000729 \\ -165980835 & -0,002651 & 120067263 & 62845875 & -44437166 \\ -88326984 & -0,001629 & 62845875 & 33812359 & -23459449 \\ 61167152 & 0,000729 & -44437166 & -23459449 & 16730803 \end{vmatrix}$$

Знайдемо вектор параметрів моделі регресії

$$a = (X^T X)^{-1} \cdot X^T Y = \begin{vmatrix} 233664552 & 0,004414 & \dots & 61167152 \\ 0,004414 & 0,000001 & \dots & 0,000729 \\ -165980835 & -0,002651 & \dots & -44437166 \\ -88326984 & -0,00169 & \dots & -23459449 \\ 61167152 & 0,000729 & \dots & 16730803 \end{vmatrix} \begin{vmatrix} 106890 \\ 160603392 \\ 159030 \\ 313010 \\ 463580 \end{vmatrix} = \begin{vmatrix} -20142994 \\ 4,1598 \\ -19615899 \\ -1713482 \\ 13408085 \end{vmatrix}$$

Записуємо модель регресії з використанням числових значень параметрів регресії

$$Y_i = -2014,3 + 4,1598 \cdot X_{i2} - 1961,5899 \cdot X_{i3} - 171,3482 \cdot X_{i4} + 1340,8085 \cdot X_{i5}$$

та визначаємо область змінення факторних перемінних

$$\begin{aligned} 456 &\leq X_{i2} \leq 2099 \\ 0,93074 &\leq X_{i3} \leq 1,69878 \\ 2,29698 &\leq X_{i4} \leq 3,54427 \\ 2,35746 &\leq X_{i5} \leq 5,41386 \end{aligned}$$

7.2.3 Аналіз багатфакторної лінійної моделі регресії

Позитивні знаки параметрів моделі вказують на те, що збільшення відповідних показників-факторів веде до збільшення результативного показника, а негативні знаки параметрів моделі вказують на те, що збільшення відповідних показників-факторів веде до зменшення результативного показника.

Так, у розглянутому прикладі, збільшення середньорічної чисельності працюючих (X_{i2}) і продуктивності праці одного працюючого (X_{i5}) викликають збільшення результативної перемінної, тоді як ріст фондівдачі (X_{i3}) і фондоозброєності (X_{i4}) викликають зниження результативної перемінної - річного валового доходу. Характер впливу на річний валовий дохід двох останніх перемінних суперечить економічному змісту. Однак це викликано тією обставиною, що ми в межах лінійної регресійної моделі нелінійний зв'язок фондівдачі (X_{i3}) і фондоозброєності (X_{i4}) з валовим доходом грубо замінили на лінійну, що і є однією з причин появи знака мінус у параметрів регресії при цих перемінних.

Величина кожного параметра моделі показує, на яку величину зміниться значення результативної перемінної при збільшенні або зменшенні відповідного фактора на одну одиницю.

Так, в розглянутому прикладі, зі збільшенням факторної перемінної X_{i2} на 1 чоловіка результативна перемінна (доход) збільшиться на 4,1598 тис.грн, а зі збільшенням факторної перемінної X_{i5} на 1 тис.грн результативна перемінна збільшиться на 1340,8 тис.грн. Збільшення двох інших факторних перемінних

веде до зниження результативної перемінної, а зниження - до збільшення.

7.2.4 Оцінка ступеня впливу факторних перемінних

Ступінь впливу факторних перемінних на результативну перемінну може бути виявлена двома шляхами: 1) шляхом аналізу параметрів моделі регресії; 2) шляхом аналізу приватних коефіцієнтів кореляції. У курсовій роботі використовується перший спосіб, що є найменш трудомістким і простим.

Оцінка ступеня впливу факторів за допомогою аналізу параметрів (коефіцієнтів) моделі регресії може бути виконана двома способами:

- з використанням бета-коефіцієнтів;
- з використанням коефіцієнтів еластичності.

Спосіб 1. Оцінка впливу факторних перемінних на результативну перемінну з використанням бета-коефіцієнтів.

Пряме порівняння коефіцієнтів регресії в рівнянні множинної регресії дає представлення про ступінь впливу факторних ознак на результативну ознаку тільки тоді, коли вони виражаються в однакових одиницях і мають приблизно однакові коливання. Щоб зробити коефіцієнти регресії порівнянними, застосовують нормовані коефіцієнти регресії β_j . Коефіцієнт β_j показує величину зміни результативного фактора в значеннях середньої квадратичної помилки при зміні факторної ознаки X_j на одну середньоквадратичну помилку.

$$\beta_j = a_j \left(\frac{\sigma_{X_j}}{\sigma_y} \right),$$

де a_j - коефіцієнт регресії при X_j факторі, $j = 1, 2, \dots, m$.

Приклад. Розрахуємо коефіцієнти β_j , що характеризують вплив чисельності (β_2), фондівіддачі (β_3), фондоозброєності (β_4) і продуктивності праці (β_5) на валовий доход у моделі

$$Y_i = -2014,3 + 4,1598 \cdot X_{i2} - 1961,5899 \cdot X_{i3} - 171,3482 \cdot X_{i4} + 1340,8085 \cdot X_{i5}$$

Середні квадратичні відхилення перемінних:

$$\sigma_Y = -2,447,1; \quad \sigma_{X_2} = 448,6; \quad \sigma_{X_3} = 0,2023531; \\ \sigma_{X_4} = -0,360497; \quad \sigma_{X_5} = 0,7440662.$$

Для розрахунку бета-коефіцієнтів використовуємо формулу

$$\beta_j = a_j \left(\frac{\sigma_{X_i}}{\sigma_Y} \right). \text{ Отримуємо наступні значення}$$

$$\beta_2 = 4,15898 \cdot (448,6 / 2447,1) = 0,76242 \\ \beta_3 = 1961,5899 \cdot (0,2023531 / 2447,1) = 0,16221 \\ \beta_4 = 171,3482 \cdot (0,360497 / 2447,1) = 0,02524 \\ \beta_5 = 1340,8085 \cdot (0,7440662 / 2447,1) = 0,407687$$

Звідси бачимо, що найбільш сильний вплив на річний валовий доход здійснюють чисельність працюючих $\beta_2 = 0,76242$, продуктивність їхньої праці $\beta_5 = 0,407687$.

При цьому чисельність працюючих впливає на валовий доход у 1,9 рази $(0,76242/0,407687)$ сильніше, ніж продуктивність праці.

Спосіб 2. Оцінка впливу факторних перемінних на результативну перемінну з використанням коефіцієнтів еластичності.

Для оцінки ступеня впливу факторних ознак використовуються часні коефіцієнти еластичності ε_j , відносно X_j .

$$\varepsilon_j = \frac{\partial \mathcal{F}}{\partial X_j} \cdot \frac{X_j}{\mathcal{F}} \quad \text{або} \quad \varepsilon_j \approx a_j \cdot \frac{\bar{X}_j}{\bar{Y}},$$

де $\frac{\partial \mathcal{F}}{\partial X_j}$ - часна похідна від регресії по перемінній X_j .

Коефіцієнт ε_j показує, на скільки відсотків зміниться результативна ознака при зміні факторної ознаки на один відсоток при фіксуванні значень інших факторів на якому-небудь рівні.

Приклад. Розрахуємо коефіцієнти еластичності ε_j , що характеризують вплив чисельності (ε_2), фондівдачі (ε_3),

фондоозброєності (Θ_4) і продуктивності праці (Θ_5) на валовий дохід у моделі

$$Y = -2014,3 + 4,1598 \cdot X_{i2} - 1961,5899 \cdot X_{i3} - 171,3482 \cdot X_{i4} + 1340,8085 \cdot X_{i5}$$

Середні значення факторів, що входять в модель: $\bar{Y} = 5625,8$; $\bar{X}_2 = 1321,0$; $\bar{X}_3 = 1,4466$; $\bar{X}_4 = 2,815$; $\bar{X}_5 = 4,076$. За цих умов коефіцієнти еластичності наступні

$$\begin{aligned}\Theta_2 &= 4,1598 \cdot (1321/5625,8) = 0,9768 \\ \Theta_3 &= 1961,5899 \cdot (1,4466/5625,8) = 0,5044 \\ \Theta_4 &= 171,3482 \cdot (2,815/5625,8) = 0,0857 \\ \Theta_5 &= 1340,8085 \cdot (4,076/5625,8) = 0,9714\end{aligned}$$

Порівняння коефіцієнтів показує, що в середньому чисельність працюючих впливає на річний дохід з такою же силою, що і продуктивність праці ($\Theta_2 = 0,9768$; $\Theta_5 = 0,9714$), але майже вдвічі сильніше, ніж фондівіддача ($\Theta_3 = 0,5044$).

Цей висновок не збігається з висновком, отриманим з використанням бета-коефіцієнтів.

Прийнято вважати, що більш точний висновок дає використання бета-коефіцієнтів. Однак кращу економічну інтерпретацію мають коефіцієнти еластичності.

ПРАКТИЧНЕ ЗАНЯТТЯ 3

КЛАСТЕРНИЙ АНАЛІЗ

Мета – отримати навички зі здійснення кластеризації об'єктів.

Теоритична довідка

Кластерний аналіз все частіше знаходить вживання в маркетингових дослідженнях. **Кластерним аналізом** користуються і університетські фахівці, і практикуючі маркетологи, вирішуючи, в першу чергу, проблему угруповання. Мова може йти про продукти, клієнтів, співробітників - при грамотному вживанні кластерний аналіз може працювати з

вельми широким спектром об'єктів. Проте, не дивлячись на високу популярність цього апарату, частенько дослідник не знайомий з його специфікою і особливостями вживання. Це, у свою чергу, вабить розчарування в аналітичних можливостях методу і необгрунтований скептицизм.

Як правило, при вживанні кластерного аналізу аналітик стикається з двома групами завдань:

1. На основі теоретичних передумов вибрати адекватний алгоритм

2. Грамотно провести аналіз і проінтерпретувати результати

Виділяють дві групи методів **кластерного аналізу**: ієрархічні і неієрархічні.

Основними методами ієрархічного кластерного аналізу є метод ближнього сусіда, метод повного зв'язку, метод середнього зв'язку і метод Варда. Найбільш універсальним є останній. Існують також центроїдні методи і методи, що використовують медіану, але Сніат і Сокал 1973 акуратно показали, що їх вживання може привести до деяких вельми небажаних наслідків.

Неієрархічні методів більше, хоча працюють вони на одних і тих же принципах. По суті, вони є ітеративними методами дроблення вихідної сукупності. В процесі ділення формуються нові кластери, і так до тих пір, поки не буде виконано правило зупинки. Між собою методи розрізняються вибором початкової точки, правило формування нових кластерів і правилом зупинки. Найчастіше використовується алгоритм *до-середніх*. Він має на увазі, що аналітик заздалегідь фіксує кількість кластерів в результуючому розбитті.

Вибираючи між ієрархічними і неієрархічними методами, слід звернути увагу на наступні моменти:

Неієрархічні методи виявляють вищу стійкість по відношенню до викидів, невірному вибору метрики, включенню незначимих змінних в базу для кластеризації і ін. Але платою за це є слово «апріорі». Дослідник повинен заздалегідь фіксувати результуючу кількість кластерів, правило зупинки і, якщо на те є підстави, початковий центр кластера. Останній момент істотно відбивається на ефективності роботи алгоритму. Якщо немає

підстав штучно задати цю умову, взагалі кажучи, рекомендується використовувати ієрархічні методи.

Завдання

Використовуючи дані 2 практичної роботи вибрати 15 значень. Результати вибірки внести до таблиці 1. З використанням даних вибірки необхідно здійснити кластерний аналіз та побудувати дендрограму. За результатами розрахунків зробити висновки, які підприємства об'єднані на вищому рівні ієрархії, а які на найнижчому.

Вказівки до виконання завдання

Вибірка виробляється таким чином: значення всіх чинників (В і X) з 1 по 5; з 10 по 15; з 30 по 35.

Таблиця 17 - Вихідні дані для кластерного аналізу

№	Yi	Xi1	Xi2	Xi3	Xi4	Xi5
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
Сума						
Середнє						

Далі необхідно графічно представити отриманні дані (приклад рис. 10).

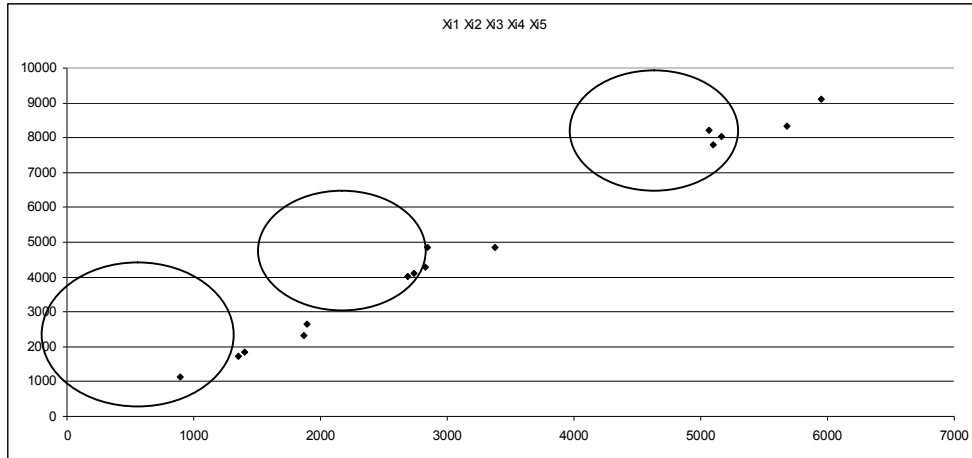


Рисунок 10 - Графічне відображення результатів кластерного аналізу

Об'єкти, обведені кружечками - кластери.

На наступному етапі кластерного аналізу здійснюється нормалізація значень X по відношенню до Y (r) по формулі 3.1:

$$r_{ij} = \frac{X_{ij} - \bar{X}_i}{\delta_i} \quad (3.1)$$

где δ_i – середнє квадратичне відхилення (3.2).

$$\delta_i = \sqrt{\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_i)^2}{n-1}} \quad (3.2)$$

Далі розраховується евклідова відстань (d) (табл. 18).

Таблиця 18 - Евклідова відстань

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0														
2		0													
3			0												
4				0											
5					0										
6						0									
7							0								
8								0							

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
9									0						
10										0					
11											0				
12												0			
13													0		
14														0	
15															0

Таблиця 19 - Приклад розрахунку

	1
1	0
2	$=\sqrt{(X1_1 - X1_2)^2 + (X2_1 - X2_2)^2 + (X3_1 - X3_2)^2 + (X4_1 - X4_2)^2 + (X5_1 - X5_2)^2}$
3	$=\sqrt{(X1_1 - X1_3)^2 + (X2_1 - X2_3)^2 + (X3_1 - X3_3)^2 + (X4_1 - X4_3)^2 + (X5_1 - X5_3)^2}$
...	
...	
...	
...	
13	
14	
15	

Результати розрахунків відображаються на дендрограмі.
На рис. 11 показаний скорочений приклад дендрограми

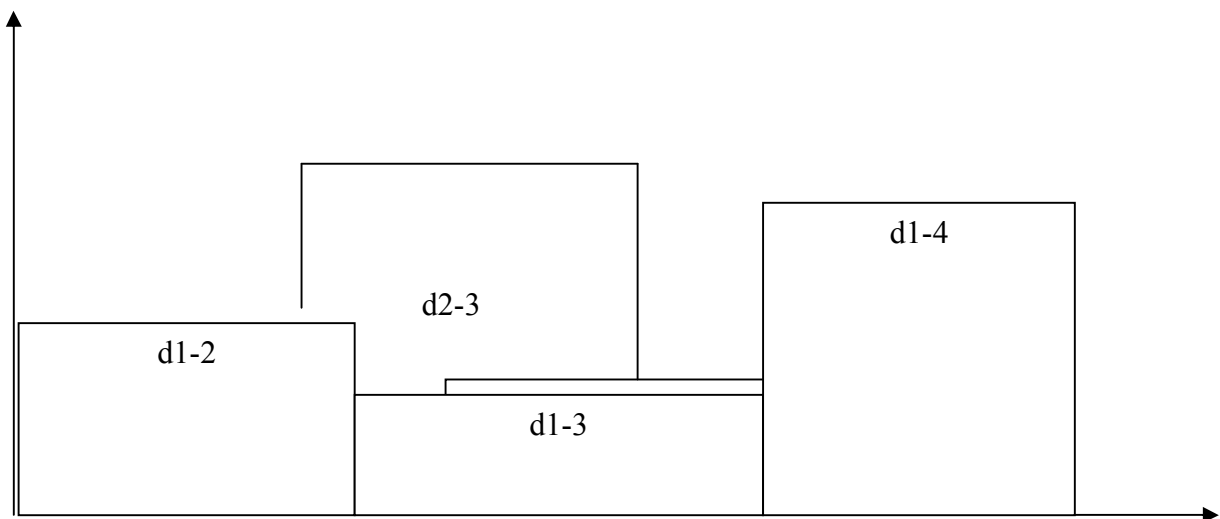


Рисунок 11 – Приклад дендрограми

ПРАКТИЧНЕ ЗАНЯТТЯ 4

КОЕФІЦІЄНТ КОНКОРДАЦІЇ

Мета – отримати навички з оцінки узгодженості думок експертів при проведенні досліджень.

Теоретична довідка

При проведенні різних вибірових досліджень (маркетингових, клінічних, епідеміологічних, соціологічних) дуже часто виникає необхідність оцінити узгодженість думок експертів. Це надалі дозволяє виявити різні підходи експертів до оцінки різних явищ (ознак, критеріїв), що дозволяє потім провести поглиблений аналіз ситуації і прийняти обґрунтоване рішення.

У зв'язку з природою даних, які є результатами експертних оцінок (зазвичай бали), для їх аналізу зазвичай використовуються рангові (непараметричні) методи.

Для того, щоб оцінити узгодженість думок двох експертів при їх оцінці ряду ознак (або об'єктів, що має значення при рішенні завдання ранжирування), можна скористатися коефіцієнтом кореляції Спірмена або коефіцієнтом кореляції Кендалла.

В разі, якщо експертів не два, а більш (наприклад, при з'ясуванні узгодженості думок групи експертів), використовується коефіцієнт конкордації, запропонований Кендаллом:

$$W = \frac{12}{m^2(n^2 - n)} S \quad (4.1)$$

де $S = \sum_{j=1}^n (\sum_{i=1}^m R_{ij} - \frac{m(n+1)}{2})^2$, n - кількість аналізованих об'єктів, m - кількість експертів, R_{ij} - ранг j -го об'єкту, який привласнений йому i -им експертом.

Слід звернути увагу на відмінність значень коефіцієнта конкордації від коефіцієнта кореляції, оскільки він існує в межах від 0 до 1. Якщо думки експертів повністю протилежні,

коефіцієнт конкордації дорівнює нулю ($W = 0$), а коефіцієнт кореляції в цьому випадку буде рівний -1 .

За наявності в'язок (однакових значень) формула (4.1) набуває наступного вигляду:

$$W = \frac{1}{\frac{1}{12}m^2(n^3 - n) - m \sum_{j=1}^m T_j} \sum_{j=1}^n \left(\sum_{i=1}^m R_{ij} - \frac{m(n+1)}{2} \right)^2 \quad (4.2)$$

де $T = \frac{\sum_{i=1}^{L_j} (n_i^2 - n_i)}{12}$, при цьому L_i - число зв'язок, n_i - кількість елементів в i -й в'язці для j -го експерта.

Перевірка значущості коефіцієнта конкордації W

З метою перевірки значущості коефіцієнта конкордації W формулюється дві статистичні гіпотези:

H_0 : думки (оцінки) експертів не узгоджуються;

H_a : думки (оцінки) експертів узгоджуються.

Нульова гіпотеза відкидається якщо $W > W_{кр}$.

Значущість коефіцієнта конкордації при малій кількості об'єктів перевірити складно. Для малих значень існують неповні таблиці, наприклад таблиця 20.

Таблиця 20 - Критичне значення коефіцієнта конкордації

n=3; m=10		n=5; m=3	
S	Рівень значущості Q	S	Рівень значущості Q
50	0,092	56	0,096
62	0,046	62	0,056
104	0,0034	78	0,053
126	0,0008	86	0,0009

Для набуття критичного значення коефіцієнта конкордації необхідне узятє з таблиці значення підставити у формулу: $\frac{12}{m^2(n^3 - n)} S$. Проте можна поступити і іншим способом. Якщо ми задамося прийнятним для нас рівнем значущості (наприклад, $\alpha=0,05$), то для з'ясування, чи значимий коефіцієнт конкордації,

необхідно узяти мінімальний рівень значущості Q , відповідний приведену в таблиці значенню S , і порівняти його із заданим нами рівнем значущості α . Якщо $Q < \alpha$, то нульова гіпотеза про неузгодженість думок експертів відкидається і приймається альтернативна (думки експертів погоджені).

Слідуює, також відзначити, що критичні значення коефіцієнта конкордації $W_{кр}$ задовільно апроксимуються (незалежно від значення n) за допомогою наступного вираження:

$$W_{кр.} = W_{\alpha, m, n} = 1 - X_{\alpha, A, B} + \frac{36}{m^2(n^3 - n)} \quad (4.3)$$

де $A = (m - 1) \times B$; $B = \frac{n - 1}{2} - \frac{1}{m}$; $X_{\alpha, A, B}$ - зворотня функція стандартного бета розподілу.

Якщо ж кількість об'єктів (n) більше 7, то нульова гіпотеза відхиляється на наближеному рівні значущості α якщо $W \geq W_{кр.} = \chi_{\alpha, \nu}^2 / (m(n - 1))$, де $\nu = n - 1$ - міри свободи для розподілу χ^2 .

Завдання

Є 7 об'єктів, кожен з яких оцінюється незалежно трьома експертами за десятибальною шкалою. Необхідно визначити міру узгодженості думок експертів - коефіцієнт конкордації.

Вказівки до виконання завдання

1. Формуємо таблицю (матрицю) рангів для вихідних даних. Дані ранжуємо для експертів (у нашому випадку по стовпцях). Для цього у вічко E3 поміщаємо формулу $=\text{Rank1}(B3:B9;0)$ і погоджено копіюємо її у всі вічка діапазону E3:G9.

2. У вічка B10 і B11 вводимо число експертів $m = \text{ЧИСЛОСТОЛБ}(C3:E3)$ і число об'єктів $n = \text{ЧСТРОК}(C3:C9)$. В принципі, у вказані вічка можна прямо ввести відповідні значення: у B10 - 3 (кількість експертів), а в B11 - 3 (кількість об'єктів).

3. Формуємо стовпець допоміжних результатів для вираження $\sum_{i=1}^m R_{ij} - \frac{m(n+1)}{2}$. Для цього підсумовуємо відрядковий (у нашому випадку) ранги для і-го об'єкту, з суми віднімаємо $m(n+1)/2$ і результат зводимо в квадрат. З цією метою у вічко Н3 поміщаємо формулу: $=(\text{СУМ}(E3:G3)-\$H\$11*(\$H\$12+1)/2)^2$. Після цього погоджено копіюємо її у всі вічка діапазону Н3:Н9.

4. Знаходимо значення виразу $\sum_{j=1}^n (\sum_{i=1}^m R_{ij} - \frac{m(n+1)}{2})^2$, підсумувавши значення по стовпцю, для чого у вічко Н10 вводим формулу $=\text{СУММ}(Н3:Н9)$.

5. Розраховуємо значення коефіцієнта конкордації, помістивши у вічко Н11 формулу $=12*\$H\$10/(\$B\$10*\$B\$10*(\$B\$11*\$B\$11*\$B\$11-\$B\$11))$

Розрахований коефіцієнт конкордації вимагає коректування, оскільки в нім не враховане те, що в отриманих ранжуваннях є в'язки. Для того, щоб врахувати в'язки, необхідно виконати наступні дії:

- Візуально визначимо в для кожного ранжування кількість в'язок і помістимо їх у відповідні вічка діапазону E12:G12.

- Обчислимо вираження, яке використовується для коректування коефіцієнта конкордації в разі наявності в'язок. Формули, які необхідно ввести у відповідні вічка, приведені в таблицю. 21.

Таблиця 21 - Формули робочого аркуша для розрахунку вираження, що коректує

	E	F	G
15	$=E13^3-E13$	$=F13^3-F13$	$=G13^3-G13$
16			$=G14^3-G14$
17	$=\text{СУММ}(E15:E16)$	$=\text{СУММ}(F15:F16)$	$=\text{СУММ}(G15:G16)$
18			$=\text{СУММ}(E17:G17)$

- Обчислимо значення коефіцієнта конкордації з врахуванням в'язок.

Для цього у вічко Н19 помістимо: $=12*\$H\$10/(\$B\$10*\$B\$10*(\$B\$11*\$B\$11*\$B\$11-\$B\$11)-B10*G18)$. В результаті буде розрахований коефіцієнт конкордації.

- Перевіримо значущість даного коефіцієнта конкордації. Згідно таблиці додатка М, для даного випадку, критичне значення коефіцієнта конкордації (при $m=3$, $n=7$ а рівень значущості: 0,05) рівне: $W_{кр.} = 0,6242$. Оскільки $W < W_{кр.}$, то нульова гіпотеза, що думки (оцінки) експертів не узгоджуються, приймається.

Вихідні дані, проміжні і кінцеві результати розрахунків приведені на мал. 1. Враховуючи невелику кількість даних і експертів, неузгодженість думок експертів можна показати графічно (див. рис.12).

	A	B	C	D	E	F	G	H
1		Експерты						
2	Объекты	Эксперт1	Эксперт2	Эксперт3	Матрица рангов			
3	1	5	2	5	5	2	3,5	2,25
4	2	6	3	7	6	4	7	25
5	3	9	7	5	7	7	3,5	30,25
6	4	3	1	6	2	1	5,5	12,25
7	5	4	5	4	3,5	6	2	0,25
8	6	4	3	6	3,5	4	5,5	1
9	7	1	3	3	1	4	1	36
10	m=	3						107
11	n=	7					W=	0,4246
12		Количество связей			1	1	2	
13		Размеры связей			2	3	2	
14							2	
15		сумма			6	24	6	
16							6	
17					6	24	12	
18	Корректирующая сумма						42	
19	Козэффициент конкордации скорректированный						W=	0,4431
20							S=	157,3
21							Wkr=	0,6242
22								

Рис. 12 - Вихідні дані і результати розрахунку коефіцієнту конкордації

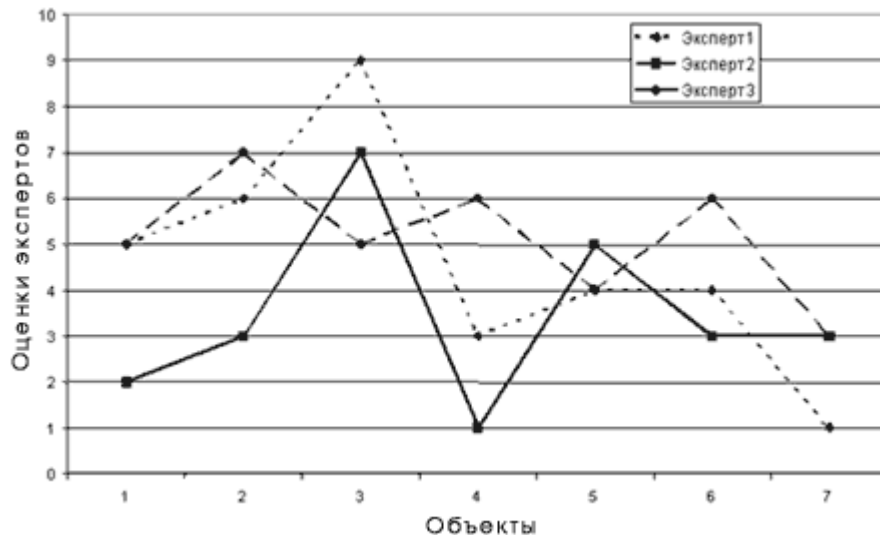


Рис. 13 - Ілюстрація неузгодженості думок експертів

ПРАКТИЧНЕ ЗАНЯТТЯ 5

ФАКТОРНИЙ АНАЛІЗ

Мета заняття – отримати навички зі здійснення детермінованого, стохастичного, дискримінантного та багатофакторного аналізів.

Теоретична довідка

Факторний аналіз - визначення впливу чинників на результат - є одним з сильних методичних рішень в аналізі господарської діяльності компаній для ухвалення рішень.

Види факторного аналізу

Розрізняють два основні види факторного аналізу - детермінований і стохастичний.

Детермінований факторний аналіз

Сутність детермінованого факторного аналізу полягає в знаходженні оцінок впливу зміни кожного з чинників на абсолютне або відносне відхилення результативного показника.

Детермінований факторний аналіз має досить жорстку послідовність виконуваних процедур:

- побудова економічно обґрунтованої детермінованої факторної моделі;

- вибір прийому факторного аналізу і підготовка умов для його виконання;
- реалізація рахункових процедур аналізу моделі;
- формулювання висновків і рекомендацій за результатами аналізу.

Перший етап особливо важливий, оскільки неправильно побудована модель може привести до логічно невиправданих результатів. Сенс цього етапу полягає в наступному: будь-яке розширення жорстко детермінованої факторної моделі не повинне перечити логіці зв'язку “причина - слідство”.

На другому етапі вибирається один з прийомів факторного аналізу: інтегральний, ланцюгових підстановок, логарифмічний і ін.

Існують наступні моделі детермінованого аналізу:

- аддитивна модель, тобто модель, в яку чинники входять у вигляді суми алгебри;
- мультиплікативна модель, тобто модель, в яку чинники входять у вигляді твору;
- кратна модель, тобто модель, що є відношенням чинників;
- змішана модель, тобто модель, в яку чинники входять в

різних комбінаціях, наприклад:
$$P_T = \frac{P}{OC + OB}$$

Жорстко детермінована модель, що має більше двох чинників, називається багатофакторною.

Типові завдання детермінованого факторного аналізу

У детермінованому факторному аналізі можна виділити чотири типові завдання:

- Оцінка впливу відносної зміни чинників на відносну зміну результативного показника.
- Оцінка впливу абсолютної зміни і-го чинника на абсолютну зміну результативного показника.
- Визначення відношення величини зміни результативного показника, викликаного зміною і-го чинника, до базової величини результативного показника.
- Визначення долі абсолютної зміни результативного показника, викликаного зміною і-го чинника, в загальній зміні результативного показника.

Основні методи детермінованого факторного аналізу

Одним з найважливіших методологічних в АХД є визначення величини впливу окремих чинників на приріст результативних показників. У детермінованому факторному аналізі (ДФА) для цього використовуються наступні способи: виявлення ізольованого впливу чинників, ланцюгової підстановки, абсолютних різниць, відносних різниць, пропорційного ділення, інтегральний, логарифмування і ін.

Перші три способи ґрунтуються на методі елімінювання. Елімінувати - означає усунути, відхилювати, виключити дію всіх чинників на величину результативного показника, окрім одного. Цей метод виходить з того, що всі чинники змінюються незалежно один від одного: спочатку змінюється один, а всі інші залишаються без зміни, потім змінюються два, потім три і т. д., при незмінності останніх. Це дозволяє визначити вплив кожного чинника на величину досліджуваного показника окремо.

Дамо коротку характеристику найбільш поширеним способам.

1. Прийом виявлення ізольованого впливу чинників.

Зміна результативного показника під впливом якого-небудь чинника обчислюється за формулою:

$$\Delta x_i y = f(x_1^0, \dots, x_{i-1}^0, x_i^1, x_{i+1}^0, \dots, x_n^0) - f(x_1^0, \dots, x_{i-1}^0, x_i^0, x_{i+1}^0, \dots, x_n^0)$$

2. Прийом цепних підстановок:

$$\Delta x_i y = f(x_1^1, \dots, x_{i-1}^1, x_i^1, x_{i+1}^0, \dots, x_n^0) - f(x_1^1, \dots, x_{i-1}^1, x_i^0, x_{i+1}^0, \dots, x_n^0)$$

Спосіб ланцюгової підстановки є вельми простим і наочним методом, найбільш універсальним зі всіх. Він використовується для розрахунку впливу чинників у всіх типах детермінованих факторних моделей: аддитивних, мультиплікативних, кратних і змішаних. Цей спосіб дозволяє визначити вплив окремих чинників на зміну величини результативного показника шляхом поступової заміни базисної величини кожного факторного показника в об'ємі результативного показника на фактичну в звітному періоді. З цією метою визначають ряд умовних величин результативного показника, які враховують зміну одну, потім два, потім три і так далі чинників, допускаючи, що останні не міняються. Порівняння величини результативного показника до і після зміни рівня того або іншого чинника дозволяє визначити дію конкретного чинника на приріст результативного показника,

виключивши вплив останніх чинників. При використанні цього методу досягається повне розкладання.

Ознака, що безпосередньо відноситься до явища, що вивчається, і характеризує його кількісну сторону, називається первинною або кількісною. Ці ознаки: а) абсолютні (об'ємні); б) їх можна підсумовувати у просторі та часі. Як приклад можна привести об'єм реалізації, чисельність, вартість оборотних коштів і так далі

Ознаки, що відносяться до явища, що вивчається, не безпосередньо, а через один або декілька інших ознак і що характеризують якісну сторону явища, що вивчається, називаються вторинними або якісними. Ці ознаки: а) відносні; б) їх не можна підсумовувати у просторі та часі. Прикладами можуть служити фондоозброєність, рентабельність і ін. У аналізі виділяють вторинні чинники 1-го, 2-го і так далі порядків, отримувані шляхом послідовної деталізації.

Жорстко детермінована факторна модель називається повною, якщо результативний показник кількісний, і неповною, якщо результативний показник якісний. У повній двофакторній моделі один чинник завжди кількісний, другий - якісний. В цьому випадку заміну чинників рекомендують починати з кількісного показника. Якщо ж є декілька кількісних і декілька якісних показників, то спочатку слід змінити величину чинників першого рівня підпорядкування, а потім нижчого. Таким чином, вживання способу ланцюгової підстановки вимагає знання взаємозв'язку чинників, їх соподчиненности, уміння правильно їх класифікувати і систематизувати.

Стохастичний факторний аналіз

Суть стохастичного методу - вимір впливу стохастичних залежностей з невизначеними і приблизними чинниками. Стохастичний метод доцільно застосовувати для економічних досліджень з неповною (імовірнісною) кореляцією: наприклад, для завдань маркетингу.

Стохастичне моделювання факторних систем взаємозв'язків окремих сторін господарської діяльності спирається на узагальнення закономірностей варіювання значень економічних показників - кількісних характеристик чинників і результатів господарської діяльності. Кількісні параметри зв'язку

виявляються на основі зіставлення значень показників, що вивчаються, в сукупності господарських об'єктів або періодів. Таким чином, першою передумовою стохастичного моделювання є можливість скласти сукупність спостережень, тобто можливість повторно виміряти параметри одного і того ж явища в різних умовах. Другий - достатня чисельність сукупності спостереження, що дозволяє з точністю і надійністю виявити наявні закономірності (у теорії статистики вважається, що кількість спостережень повинна в 6-8 разів перевищувати кількість чинників). А третьою передумовою є - наявність методів, тобто спеціального математичного апарату, що дозволяє виявити тісноту зв'язку між показниками, що вивчаються, і оцінити величину впливу чинників на зміну результативного показника.

В цілому стохастичне моделювання призначене для вирішення трьох завдань:

- 1) встановлення факту наявності або відсутності зв'язку між ознаками, що вивчаються;
- 2) виявлення причинних зв'язків між показниками, що вивчаються, і кількісний вимір дії чинників на результативний показник;
- 3) прогнозування невідомих значень результативних показників.

Методи стохастичного факторного аналізу

1. Спосіб парної кореляції - метод кореляційного і регресійного (стохастичного) аналізу широко використовується для визначення тісноти зв'язку між показниками, що не знаходяться у функціональній залежності, тобто зв'язок виявляється не у кожному окремому випадку, а в певній залежності.

За допомогою кореляції вирішуються два головні завдання:

- 1) складається модель чинників, що діють (рівняння регресії);
- 2) дається кількісна оцінка тісноти зв'язку (коефіцієнт кореляції).

2. Матричних моделей - є схематичним віддзеркаленням економічного явища або процесу за допомогою наукової абстракції. Найбільшого поширення тут

набув метод аналізу «витрати-випуск», що будується за шаховою схемою і дозволяє в найбільш компактній формі представити взаємозв'язок витрат і результатів виробництва.

3. **Математичне програмування** - це основний засіб вирішення завдань по оптимізації виробничо-господарської діяльності.

4. **Метод дослідження операцій** - направлений на вивчення економічних систем, у тому числі виробничо-господарській діяльності підприємств, з метою визначення такого поєднання структурних взаємозв'язаних елементів систем, яке найбільшою мірою дозволить визначити найкращий економічний показник з ряду можливих.

5. **Теорія ігор** як розділ дослідження операцій - це теорія математичних моделей ухвалення оптимальних рішень в умовах невизначеності або конфлікту декількох сторін, що мають різні інтереси.

Дискримінантний аналіз

Дискримінантний аналіз — різновид багатовимірною аналізу, призначеного для попередньої класифікації даних. Ґрунтується на дискримінантній функції і дає можливість визначити відмінність між двома сукупностями об'єктів. Застосовується в маркетингових дослідженнях при вирішенні питань сегментації ринку, при об'єктивній оцінці ступеня новизни товарів тощо.

Аналіз дискримінанта використовується для ухвалення рішення про те, які змінні розрізняють (дискримінують) дві або що більш виникають сукупності (групи). Наприклад, деякий дослідник в області освіти може захотіти досліджувати, які змінні відносять випускника середньої школи до однієї з трьох категорій: (1) що поступає в коледж, (2) що вступає до професійної школи або (3) відмовляється від подальшої освіти або професійної підготовки. Для цієї мети дослідник може зібрати дані про різні змінні, пов'язані з учнями школи. Після випуску більшість учнів природно повинно попасти в одну з названих категорій. Потім можна використовувати аналіз дискримінанта для визначення того, які змінні дають найкраще передбачення вибору такими, що вчаться подальшої дороги.

Найбільш загальним вживанням аналізу дискримінанта є включення в дослідження багатьох змінних з метою визначення тих з них, які щонайкраще розділяють сукупності між собою. Наприклад, дослідник в області освіти, що цікавиться передбаченням вибору, який зроблять випускники середньої школи відносно своєї подальшої освіти, виробить з метою здобуття найбільш точних прогнозів реєстрацію можливо більшої кількості параметрів що виучуються, наприклад, мотивацію, академічну успішність і так далі

Покроковий аналіз з включенням. У покроковому аналізі функцій дискримінантів модель дискримінації будується по кроках. Точніше, на кожному кроці є видимими всі змінні і знаходиться та з них, яка вносить найбільший вклад до відмінності між сукупностями. Ця змінна має бути включена в модель на даному кроці, і відбувається перехід до наступного кроку.

Покроковий аналіз з виключенням. Можна також рухатися у зворотному напрямі, в цьому випадку всі змінні будуть спочатку включені в модель, а потім на кожному кроці усуватимуться змінні, що вносять малий вклад до передбачень. Тоді як результат успішного аналізу можна зберегти лише "важливі" змінні в моделі, тобто ті змінні, чий вклад в дискримінацію більше останніх.

F для включення, F для виключення. Ця покрокова процедура "керується" відповідним значенням F для включення і відповідним значенням F для виключення. Значення F статистики для змінної вказує на її статистичну значущість при дискримінації між сукупностями, тобто, вона є мірою вкладу змінної в передбачення членства в сукупності.

Розрахунок на випадок. Покроковий аналіз дискримінанта заснований на використанні статистичного рівня значущості. Тому за своєю природою покрокові процедури розраховують на випадок, оскільки вони "ретельно перебирають" змінні, які мають бути включені в модель для здобуття максимальної дискримінації. При використанні покрокового методу дослідник повинен усвідомлювати, що використовуваний при цьому рівень значущості не відображає дійсного значення альфа, тобто, вірогідність помилкового відхилення гіпотези H_0 (нульової

гіпотези, що полягає в тому, що між сукупностями немає відмінності).

Інтерпретація функції дискримінації для двох груп. Для двох груп аналіз дискримінанта може розглядатися також як процедура множинної регресії (і аналогічна їй). (З обчислювальної точки зору всі ці підходи аналогічні). Якщо ви кодуєте дві групи як 1 і 2, і потім використовуєте ці змінні як залежні змінні в множинній регресії, то отримаєте результати, аналогічні тим, які отримали б за допомогою аналізу дискримінанта. Загалом, в разі двох сукупностей ви підганяєте лінійне рівняння наступного типу:

$$\text{Група} = a + b_1 * x_1 + b_2 * x_2 + \dots + b_m * x_m$$

де a є константою, і $b_1 \dots b_m$ є коефіцієнтами регресії. Інтерпретація результатів завдання з двома сукупностями тісно слідує логіці вживання множинної регресії: змінні з найбільшими регресійними коефіцієнтами вносять найбільший вклад до дискримінації.

Функції дискримінантів для декількох груп

Якщо є більше двох груп, то можна оцінити більш, ніж одну функцію дискримінанта подібно до того, як це було зроблено раніше. Наприклад, коли є три сукупності, можна оцінити: (1) - функцію для дискримінації між сукупністю 1 і сукупностями 2 і 3, узятими разом, і (2) - іншу функцію для дискримінації між сукупністю 2 і сукупності 3. Наприклад, є одна функція, дискримінуюча між тими випускниками середньої школи, які йдуть в коледж, проти тих, хто цього не робить (але хоче отримати роботу або піти в училищі), і другу функцію для дискримінації між тими випускниками, які хочуть отримати роботу проти тих, хто хоче піти в училищі. Коефіцієнти b в цих дискримінуючих функціях можуть бути проінтерпретовані тим же способом, що і раніше.

Канонічний аналіз. Коли проводиться аналіз дискримінанта декількох груп, не вказують, яким чином слід комбінувати групи для формування різних дискримінуючих функцій. Замість цього, автоматично визначають деякі оптимальні комбінації змінних, так що перша функція проведе найкращу дискримінацію між всіма групами, друга функція буде другою найкращою і так далі. Більш того, функції будуть

незалежними або ортогональними, тобто їх вклади в розділення сукупностей не перекриватимуться. З обчислювальної точки зору система ви проводить аналіз канонічних, які визначатимуть послідовне канонічне коріння і функції. Максимальне число функцій дорівнюватиме числу сукупностей мінус один або числу змінних в аналізі залежно від того, яке з цих чисел менше.

Завдання 1

Об'єм валової продукції (ВП) залежить від двох основних чинників першого рівня: чисельності працівників (ЧР) і середньорічного вироблення (ГВ). Маємо двофакторну мультиплікативну модель: $ВП = ЧР \cdot ГВ$. Розглянемо ситуацію, коли і вироблення, і чисельність робітників в звітному періоді відхилялися від запланованих значень. Завдання виконується з використанням прийому виявлення ізольованого впливу чинників і прийому ланцюгових підстановок.

Дані для розрахунків приведені в таблиці 22 (і – остання цифра залікової книжки; j – передостання цифра залікової книжки) .

Таблиця 22 - Дані для факторного аналізу об'єму валової продукції.

Показник	Умовні позначки	План	Факт	Відхилення	Індивідуальний варіант
1	2	3	4	5	6
Валова продукція, тис. грн.	ВП	160 000	240 000	+80 000	+30*i

Продовження табл. 22

1	2	3	4	5	6
Середньорічна чисельність робочих, чол.	ЧР	1000	1200	+200	+15*(i+J)
Середньорічна виробка одного робочого, тис. грн.	ГВ	160	200	+40	+5j

Вказівки до виконання завдання

1. Виконання завдання з використанням прийому виявлення ізольованого впливу чинників побудоване на основі простої двофакторної моделі $p = a \cdot b$. При аналізі динаміки цих показників виконуватиметься наступне співвідношення між індексами: $I_p = I_a \cdot I_b$, де значення індексу знаходиться відношенням значення показника в звітному періоді до базисного.

Розрахуємо індекси чисельності працівників і середньорічного вироблення по формулах 5.1 і 5.2:

$$I_{\text{чр}} = \frac{\text{ЧР}_{\phi}}{\text{ЧР}_{\text{нл}}} \quad (5.1)$$

$$I_{\text{в}} = \frac{\text{ГВ}_{\phi}}{\text{ГВ}_{\text{нл}}} \quad (5.2)$$

2. Далі виробляється розрахунок індексу валової продукції (5.3):

$$I_{\text{вп}} = \frac{\text{ЧР}_{\phi} \cdot \text{ГВ}_{\phi}}{\text{ЧР}_{\text{нл}} \cdot \text{ГВ}_{\text{нл}}} \quad (5.3)$$

3. На підставі проведених розрахунків робиться висновок про вплив чисельності працівників і середньорічного вироблення на валову продукцію

4. Алгоритм розрахунку способом ланцюгової підстановки для даної моделі виглядає таким чином (формули 5.4 – 5.6):

$$\text{ВП}_{\text{нл}} = \text{ЧР}_{\text{нл}} \cdot \text{ГВ}_{\text{нл}} \quad (5.4)$$

$$\text{ВП}_{\text{усл}} = \text{ЧР}_{\phi} \cdot \text{ГВ}_{\text{нл}} \quad (5.5)$$

$$\text{ВП}_{\phi} = \text{ЧР}_{\phi} \cdot \text{ГВ}_{\phi} \quad (5.6)$$

5. На підставі розрахунків необхідно зробити висновок про те, які чинники надали дії на перевиконання або недовиконання плану.

Сума алгебри чинників при використанні даного методу обов'язково має дорівнювати загальному приросту результативного показника (формула 5.7). Відсутність такої рівності говорить про допущені помилки у розрахунках.

$$\Delta ЧР + \Delta ГВ = \Delta ВП \quad (5.7)$$

Завдання 2

Використовуючи інтегральний спосіб факторного аналізу необхідно оцінити резерви росту обсягу товарної продукції. Вихідні дані наведені у таблиці 23

Таблиця 23 – Вихідні дані для аналізу

Показники	План	Факт	Зміна	Індивідуальний варіант
1. Товарна продукція, м ³	35480	32560	-2920	+2(i+j)
2. Чисельність робітників, чел.	135	115	-20	+i
3. Кількість днів, відпрацьованих одним робітником, дні	242	240	-2	-2j
4. Фонд робочого часу, год.	261360	220800	-40560	-3i
5. Виробка в м ³ /год	0,14	0,15	+0,01	-

Вказівки до виконання завдання

1. Визначимо плановий і фактичний фонд робочого часу Т (5.8):

$$T = ЧР \cdot Д \cdot П \quad (5.8)$$

де ЧР - чисельність, чіл.;

Д - кількість відпрацьованих днів одним робітником, дні;

П - середня тривалість робочого дня, година Середню тривалість робочого дня, приймемо рівною 8 годин.

2. Розрахуємо планову і фактичну продуктивність (вироблення) в годину, V:

$$V = \frac{N}{T} \quad (5.9)$$

де N - товарно-годинна продукція, мЗ;

T - фонд робочого часу, година

3. Виробимо розрахунок показника інтенсивності роботи $K_{\text{інт}}$:

$$K_{\text{інт}} = \frac{V_{\Phi}}{V_{\pi}} \quad (5.10)$$

4. Визначимо інтегральним способом вплив чинників на зміну об'єму товарній продукції.

- Вплив зміни величини фонду робочого часу (5.11):

$$\Delta N_{\text{т}} = \frac{1}{2} \Delta \text{ЧР} (\text{Д}_{\pi} \cdot V_{\Phi} + \text{Д}_{\Phi} \cdot V_{\pi}) + \frac{1}{3} \cdot \Delta \text{ЧР} \cdot \Delta \text{Д} \cdot \Delta V \quad (5.11)$$

- Вплив зміни кількості відпрацьованих одним робітником днів (5.12):

$$\Delta N_{\text{д}} = \frac{1}{2} \Delta V (\text{ЧР}_{\pi} \cdot V_{\Phi} + \text{ЧР}_{\Phi} \cdot V_{\pi}) + \frac{1}{3} \cdot \Delta \text{ЧР} \cdot \Delta \text{Д} \cdot \Delta V \quad (5.12)$$

- Вплив зміни величини вироблення (5.13):

$$\Delta N_{\text{в}} = \frac{1}{2} \Delta V (\text{ЧР}_{\pi} \cdot \text{Д}_{\Phi} + \text{ЧР}_{\Phi} \cdot \text{Д}_{\pi}) + \frac{1}{3} \cdot \Delta \text{ЧР} \cdot \Delta \text{Д} \cdot \Delta V \quad (5.13)$$

- Сума впливу чинників (спад (приріст) об'єму товарної продукції) (5.14):

$$\Delta N = \Delta N_{\text{в}} + \Delta N_{\text{т}} + \Delta N_{\text{д}} \quad (5.14)$$

5. За результатами розрахунків зробити висновки.

Завдання 3

На підставі даних таблиць. 1 необхідно проаналізувати залежність між витратами на оплату праці (Y) і виручкою від продажу товарів (X).

Таблиця 24 - Дані про виручку від продажу товарів і суму витрат на оплату праці в розрізі торговельних організацій тис. грн.

№ кранн иць А	Виручка від продажу товарів	Сума розходів на оплату праці	№ кранн иць Б	Виручка від продажу товарів	Сума розходів на оплату праці	Индиві дуальни й варіант
1.	3 200	190	15.	1 690	177	+2(i+j)
2.	500	45	16.	7 450	230	
3.	12 000	670	17.	12 900	587	
4.	8 560	345	18.	2 010	166	
5.	14 100	713	19.	1 650	105	
6.	11 300	470	20.	5 115	241	
7.	4 300	194	21.	8 945	400	
8.	1 010	98	22.	11 900	523	
9.	8 230	244	23.	14 200	780	
10.	12 560	510	24.	10 300	576	
11.	6 201	215	25.	11 450	425	
12.	11 500	603	26.	13 000	606	
13.	13 300	575	27.	6 100	210	
14.	1 000	95	28.	7 500	249	

Вказівки до виконання

1. На підставі даних таблиць. 1 побудуємо графік залежності зміни витрат на оплату праці від зміни товарообігу.

2. Далі виміряємо тісноту зв'язку між показниками, що вивчаються, на основі коефіцієнта кореляції, для чого згрупуємо магазини по сумі виручки від продажу товарів і складемо наступну таблицю (таблиця. 25).

Коефіцієнт кореляції розраховується на підставі наступної формули:

$$r_{YX_j} = \frac{\overline{YX_j} - \bar{Y} \cdot \bar{X_j}}{\sqrt{Y^2 - (\bar{Y})^2} \cdot \sqrt{X_j^2 - (\bar{X_j})^2}} \quad (5.15)$$

Коефіцієнт детермінації: r^2 (показує на скільки відсотків зміну витрат на оплату праці пояснюється зміною виручки від продажу).

3. Побудуємо рівняння регресії наступного вигляду: $Y(x) = a + b \cdot x$,

де Y - витрати на оплату праці; x - виручка від продажу товарів.

Для визначення параметрів a і b в слід вирішити систему нормальних рівнянь методом найменших квадратів:

$$\begin{cases} \sum y = na + b\sum x \\ \sum xy = a\sum x + b\sum x^2. \end{cases}$$

Звідси значення коефіцієнта b визначається по формулі

$$b = \frac{n\sum xy - \sum y\sum x}{n\sum x^2 - \sum x\sum x} \quad (5.16)$$

Підставимо значення даного параметра в перше рівняння системи і визначимо значення параметра a :

$$\sum y = na + b\sum x \quad (5.17)$$

Значення параметра a показує для торгівельних організацій з яким розміром виручки від продажу за рік може бути застосовний коефіцієнт регресії.

Таблиця 25 - Таблиця для визначення показників, використовуваних при розрахунку коефіцієнта кореляції

Групи магазинів по сумі виручки від продажу товарів	Кількість магазинів	Виручка від продажу товарів (X_i), тис. грн.	Сума розходів на оплату праці (Y_i), тис. грн.	$Y_i X_i$	X_i^2	Y_i^2
От 500 до 3 220 включ.						
От 3 221 до 5 440 включ.						
От 5 441 до 8 160 включ.						
От 8 161 до 10 880 включ.						
Понад 10 881						
Сума						

Завдання 4

Поліграфічне видавництво "Синя птиця" закінчило рік з наступними фінансовими результатами: Валовий прибуток підприємства за поточний рік склав 236,54 тис. гривень, торік було 189,48 тис. гривень. Рівень оподаткування 37,1% в поточному році і 45,3% в прошедшем. Виручка від всіх видів продажів була 742,8 тис. гривень, а торік 685,6 тис. гривень. Розмір середньорічного сукупного капіталу рівний 475,1 тис. гривень, торік був 482,12 тис. гривень. Розмір власного капіталу склав 253,7 тис. гривень, торік 289,45 тис. гривень. Необхідно провести 4-х факторний аналіз прибутковості власного капіталу та дати відповіді на запитання: Що вплинуло на зміну прибутковості власного капіталу? Визначите величину впливу співвідношення цих чинників на зміну рентабельності власного капіталу.

Порядок виконання завдання

1. Розрахуємо долю чистого прибутку в об'ємі валового прибутку за минулий рік і за поточний рік:

$$Дчп_{\text{пр}} = ВП_{\text{пр}} \cdot \frac{1 - Пб_{\text{нпр}}}{ВП} \quad (5.18)$$

$$Дчп_{\text{тр}} = ВП_{\text{тр}} \cdot \frac{1 - Пб_{\text{нтр}}}{ВП} \quad (5.19)$$

де $ВП_{\text{пр}}$, $ВП_{\text{тр}}$ - валовий прибуток за минулий і поточний рік;

$Пб_{\text{нпр}}$, $Пб_{\text{нтр}}$ - податковий прибуток за минулий і поточний рік.

2. Далі проведемо розрахунок зміни за рік долі чистого прибутку в об'ємі валового прибутку (5.20):

$$\Delta Дчп = Дчп_{\text{тр}} - Дчп_{\text{пр}} \quad (5.20)$$

3. Визначимо значення мультиплікатора капіталу за минулий (5.21) і поточний рік (5.21) і зміну (5.23):

$$МК_{\text{пр}} = \frac{\overline{\Sigma K_{\text{пр}}}}{СК_{\text{пр}}} \quad (5.21)$$

$$MK_{тг} = \frac{\overline{\Sigma K_{шт}}}{СК_{тг}} \quad (5.22)$$

де $\overline{\Sigma K_{шт}}$ - розмір середньорічного сукупного капіталу у минулому і поточному році;

$СК_{шт,тг}$ - розмір власного капіталу у минулому і поточному році.

$$\Delta MK = MK_{тг} - MK_{пг} \quad (5.23)$$

4. Визначимо значення і зміну коефіцієнта оборотності капіталу:

$$Коб_{пг} = \frac{В_{продпг}}{\Sigma A_{пг}} \quad (5.24)$$

$$Коб_{тг} = \frac{В_{продтг}}{\Sigma A_{тг}}, \quad (5.25)$$

де $В_{продпг,тг}$ - валова продукція у минулому і поточному році;

$\Sigma A_{пг,тг}$ - сукупні активи у минулому і поточному році.

$$\Delta Коб = Коб_{тг} - Коб_{пг} \quad (5.26)$$

5. Визначимо значення і зміну рентабельності обігу капіталу:

$$R_{обпг} = \frac{ВП_{пг}}{В_{продпг}} \quad (5.27)$$

$$R_{обтг} = \frac{ВП_{тг}}{В_{продтг}} \quad (5.28)$$

$$\Delta R_{об} = R_{обтг} - R_{обпг} \quad (5.29)$$

6. Обчислимо значення (5.30) і зміну рентабельності власного капіталу (5.31):

$$ROE_{пг,тг} = \frac{ВП_{пг,тг} \cdot (1 - Пб_{пг,тг})}{СК_{пг,тг}} \quad (5.30)$$

$$\Delta ROE = ROE_{ТГ} - ROE_{ПГ} \quad (5.31)$$

7. Визначимо міру впливу чинників на зміну ROE:

а) за рахунок зміни питомої ваги чистому прибутку в загальній сумі балансового прибутку (5.32):

$$ROE = Дчп_{ПГ} \cdot МК_{ПГ} \cdot Коб_{ПГ} \cdot Rob_{ПГ} \quad (5.32)$$

б) мультиплікатора капіталу:

$$ROE = Дчп_{ТГ} \cdot \Delta МК \cdot Коб_{ПГ} \cdot Rob_{ПГ} \quad (5.33)$$

в) оборотності капіталу:

$$ROE = Дчп_{ТГ} \cdot МК_{ТГ} \cdot Коб \cdot Rob_{ПГ} \quad (5.34)$$

г) рентабельності продажів:

$$ROE = Дчп_{ТГ} \cdot МК_{ТГ} \cdot Коб_{ТГ} \cdot Rob \quad (5.35)$$

Завдання 5

Маємо дані по двом групам промислових підприємств машинобудівного комплексу (табл. 26):

X1 – фондвіддача основних виробничих фондів, грн.;

X2- витрати на грн. виробленої продукції, коп.;

X3 – витрати сировини та матеріалів на одну гривню продукції, коп.

Таблиця 26 – Вихідні дані для виконання завдання

Групи	Номер підприємства	X ₁	X ₂	X ₃
1-я група	1	0,50	94,0	8,50
	2	0,67	75,4	8,79
	3	0,68	85,2	9,10
	4	0,55	98,8	8,47
2-я група	5	1,52	81,5	4,95
	6	1,20	93,8	6,95
	7	1,46	86,5	4,70

Необхідно провести класифікацію чотирьох нових підприємств, що мають наступні значення вихідних змінних, застосовуючи дискримінантний факторний аналіз:

1-е підприємство: $x_1 = 1,07$; $x_2 = 93,5$; $x_3 = 5,30$,

2-е підприємство: $x_1 = 0,99$; $x_2 = 84,0$; $x_3 = 4,85$,

3-е підприємство: $x_1 = 0,70$; $x_2 = 76,8$; $x_3 = 3,50$,

4-е підприємство: $x_1 = 1,24$; $x_2 = 88,0$; $x_3 = 4,95$.

Порядок виконання завдання

1. Необхідно представити дані у вигляді матриць M1 и M2:

$$M_{1-я} = \begin{pmatrix} X_{11} & X_{21} & X_{31} \\ X_{12} & X_{22} & X_{32} \\ X_{13} & X_{23} & X_{33} \\ X_{14} & X_{24} & X_{34} \end{pmatrix}, M_{2-я} = \begin{pmatrix} X_{15} & X_{25} & X_{35} \\ X_{16} & X_{26} & X_{36} \\ X_{17} & X_{27} & X_{37} \end{pmatrix}.$$

2. Далі слід розрахувати середнє значення кожної змінної в окремих групах для визначення положення центрів цих груп.

3. На наступному етапі розраховуються коефіцієнти для дискримінантної функції:

$$f(x) = a_1 x_1 + a_2 x_2 + a_3 x_3 \quad (5.36)$$

Коефіцієнти a_1 , a_2 и a_3 розраховуються за формулою 6.1:

$$A = S_*^{-1}(\bar{X}_1 - \bar{X}_2) \quad (5.37)$$

де \bar{X}_1, \bar{X}_2 - вектори середніх в першій і другій групах;

A - вектор коефіцієнтів;

S_*^{-1} - матриця, зворотна спільній ковариационной матриці.

Для визначення спільної ковариационной матриці S^* потрібно розрахувати матриці S_1 і S_2 . Кожен елемент цих матриць є різницею між відповідним значенням вихідної змінної x_{ij} і середнім значенням цієї змінної в даній групі \bar{x}_{ik} (k - номер групи):

$$S_1 = \begin{pmatrix} X_{11} \cdot X'_{11} & X_{21} \cdot X'_{21} & X_{31} \cdot X'_{31} \\ X_{12} \cdot X'_{12} & X_{22} \cdot X'_{22} & X_{32} \cdot X'_{32} \\ X_{13} \cdot X'_{13} & X_{23} \cdot X'_{23} & X_{33} \cdot X'_{33} \end{pmatrix}; S_2 = \begin{pmatrix} X_{15} \cdot X'_{15} & X_{25} \cdot X'_{25} & X_{35} \cdot X'_{35} \\ X_{16} \cdot X'_{16} & X_{26} \cdot X'_{26} & X_{36} \cdot X'_{36} \\ X_{17} \cdot X'_{17} & X_{27} \cdot X'_{27} & X_{37} \cdot X'_{37} \end{pmatrix}$$

де X'_{ij} - транспонована матриця відхилень спостережуваних значень вихідних змінних від їх середніх величин.

$$S^* = \frac{1}{n_1 + n_2 - 2} (S_1 + S_2), \quad (5.38)$$

где n_1, n_2 – число об'єктів 1-ої и 2-ої груп.

4. Далі по формулі 6.2 розраховуємо вектор коефіцієнтів функції дискримінанта.
5. Підставляємо набутих значень коефіцієнтів у формулу 5.36 і розраховуємо значення функції дискримінанта для кожного об'єкту і знаходимо середні значення.
6. На наступному етапі виробляється розрахунок константи дискримінації C по формулі 5.39:

$$C = 1/2(\bar{f}_1 - \bar{f}_2) \quad (5.39)$$

Після здобуття константи дискримінації можна перевірити правильність розподілу об'єктів у вже існуючих двох класах, а також провести класифікацію нових об'єктів. Для того, щоб віднести об'єкти до одного з двох безлічі, необхідно розрахувати для них значення функцій дискримінантів (по трьох змінним).

ПРАКТИЧНЕ ЗАНЯТТЯ 6 ДИСПЕРСІЙНИЙ АНАЛІЗ

Мета – набути навички зі здійснення одно- та багатофакторного дисперсійного аналізу для незв'язаних та зв'язаних виборок.

Теоретична довідка

Дисперсійний аналіз, запропонований Р. Фішером, є статистичним методом, призначеним для виявлення впливу ряду окремих чинників на результати експериментів.

У основі дисперсійного аналізу лежить припущення про те, що одні змінні можуть розглядатися як причини (чинники, незалежні змінні), а інші як следствия (залежні змінні). Незалежні змінні називають інколи регульованими чинниками саме тому, що в експерименті дослідник має можливість варіювати ними і аналізувати результат, що виходить.

Сутність дисперсійного аналізу полягає в розчленуванні загальної дисперсії ознаки, що вивчається, на окремі компоненти, обумовлені впливом конкретних чинників, і перевірці гіпотез про значущість впливу цих чинників на досліджувану ознаку. Порівнюючи компоненти дисперсії один з одним за допомогою F - критерію Фішера, можна визначити, яка доля загальної варіативної результативної ознаки обумовлена дією регульованих чинників.

Вихідним матеріалом для дисперсійного аналізу служать дані дослідження три і більш за вибірки, які можуть бути як рівними, так і нерівними за чисельністю, як зв'язковими, так і незв'язковими. По кількості регульованих чинників, що виявляються, дисперсійний аналіз може бути однофакторним (при цьому вивчається вплив одного чинника на результати експерименту), двофакторним (при вивченні впливу двох чинників) і багатофакторним (дозволяє оцінити не лише вплив кожного з чинників окремо, але і їх взаємодію).

Дисперсійний аналіз відноситься до групи параметричних методів і тому його слід застосовувати лише тоді, коли доведено, що розподіл є нормальним.

6.1 Однофакторний дисперсійний аналіз для незв'язаних вибірок

Вивчається дія лише одній змінної (чинника) на досліджувану ознаку. Дослідника цікавить питання, як змінюється певна ознака в різних умовах дії змінної (чинника). Наприклад, як змінюється час рішення задачі за різних умов мотивації випробовуваних (низькій, середній, високій мотивації) або при різних способах пред'явлення завдання (усно, письмово або у вигляді тексту з графіками і ілюстраціями), в різних умовах роботи із завданням (наодинці, в кімнаті з викладачем, в класі). У першому випадку чинником є мотивація, в другому - міра наочності, в третьому - чинник публічності.

У даному варіанті методу впливу кожною з градацій піддаються різні вибірки випробовуваних. Градацій чинника повинно бути не менше три.

Дисперсійний однофакторний аналіз дозволяє перевірити гіпотези:

H₀: відмінності в об'ємі відтворення слів між групами є не більш вираженими, ніж випадкові відмінності усередині кожної групи

H₁: Відмінності в об'ємі відтворення слів між групами є більш вираженими, ніж випадкові відмінності усередині кожної групи.

6.2 Дисперсійний аналіз для зв'язаних вибірок

Метод дисперсійного аналізу для зв'язаних вибірок застосовується в тих випадках, коли досліджується вплив різних градацій чинника або різних умов на одну і ту ж вибірку випробовуваних. Градацій чинника повинно бути не менше три.

В даному випадку відмінності між випробовуваними - можливе самостійне джерело відмінностей. Одинфакторний дисперсійний аналіз для зв'язаних вибірок дозволить визначити, що переважає - тенденція, виражена кривою зміни чинника, або індивідуальні відмінності між випробовуваними. Чинник індивідуальних відмінностей може виявитися значимішим, ніж чинник зміни експериментальних умов.

Завдання 1

Три різні групи з шести випробовуваних отримали списки з десяти слів. Першій групі слова пред'являлися з низькою швидкістю - 1 слово в 5 секунд, другій групі з середньою швидкістю - 1 слово в 2 секунди, і третій групі з великою швидкістю - 1 слово в секунду. Було передбачене, що показники відтворення залежатимуть від швидкості пред'явлення слів. Результати представлені в таблиці. 27.

Таблиця 27 - Кількість відтворених слів

№ іспитуємого	Група 1 низька швидкіст ь	Група 2 середня швидкіст ь	Група 3 висока швидкість	Індивідуальний варіант
1	8	7	4	+і
2	7	8	5	+і
3	9	5	3	+і
4	5	4	6	+і
5	6	6	2	+і
6	8	7	4	+і
Сума	43	37	24	
Середнє	7,17	6,17	4	
Загальна сума	104			

Вказівки до виконання завдання

1. Підраховується $SS_{\text{факт}}$ - варіативну ознаки, обумовлену дією досліджуваного чинника. Позначення SS , що часто зустрічається, - скорочення від "суми квадратів" (sum of squares).

$$SS_{\text{факт}} = \frac{\sum T_c^2}{n} - \frac{(\sum x_i)^2}{N} \quad (6.1)$$

де T_c - сума індивідуальних значень по кожній з умов. Для нашого прикладу 43, 37, 24 (див. таблиці. 1);

c - кількість умов (градацій) чинника (=3);

n - кількість випробовуваних в кожній групі (=6);

N - загальна кількість індивідуальних значень (=18);

$(\sum x_i)^2$ - квадрат загальної суми індивідуальних значень (=104²=10816)

Відзначимо різницю між $(\sum x_i)^2$, в якій всі індивідуальні значення спочатку зводяться в квадрат, а потім підсумовуються і $(\sum x_i)^2$, де індивідуальні значення спочатку підсумовуються для здобуття загальної суми, а потім вже ця сума зводиться в квадрат.

2. Далі необхідно розрахувати $SS_{\text{общ}}$ - загальну варіативну ознаки:

$$SS_{\text{общ}} = \frac{\sum x_i^2 - (\sum x_i)^2}{N} \quad (6.2)$$

3. Наступним етапом є розрахунок випадкової (залишковою) величини $SS_{\text{сл}}$, обумовленої неврахованими чинниками:

$$SS_{\text{сл}} = SS_{\text{общ}} - SS_{\text{факт}} \quad (6.3)$$

4. число ступенів свободи дорівнює:

$$\begin{aligned} k_{\text{факт}} &= k_1 = c - 1 \\ k_{\text{общ}} &= N - 1 \\ k_{\text{сл}} &= k_2 = k_{\text{общ}} - k_{\text{факт}} \end{aligned}$$

(6.4)

5. На наступному етапі розраховується «середній квадрат» або математичне чекання суми квадратів, усереднена величина відповідних сум квадратів SS рівна:

$$\begin{aligned} MS_{\text{факт}} &= \frac{SS_{\text{факт}}}{k_{\text{факт}}} \\ MS_{\text{сл}} &= \frac{SS_{\text{сл}}}{k_{\text{сл}}} \end{aligned} \quad (6.5)$$

6. Далі знаходимо значення статистики критерію $F_{\text{емп}}$ по формулі 6.6:

$$F_{\text{емп}} = \frac{MS_{\text{факт}}}{MS_{\text{случ}}} \quad (6.6)$$

7. Визначаємо $F_{\text{крит}}$ по статистичних таблицях. Для $df1=k1=2$ і $df2=k2=15$ табличне значення статистики дорівнює 3,68

8. Якщо $F_{\text{емп}} < F_{\text{крит}}$, то приймається нульова гіпотеза, інакше приймається альтернативна гіпотеза.

По закінченню розрахунків необхідно зробити висновок чи є відмінності в об'ємі відтворення слів між групами більш вираженими, ніж випадкові відмінності усередині кожної групи

чи ні, і чи впливає швидкість пред'явлення слів на об'єм їх відтворення чи ні.

Завдання 2

1. Визначите чи впливає чинник освіти на рівень зарплати співробітників фірми на підставі наступних даних (див. таблиці. 28).

Таблиця 28 - Освіта і зарплата співробітників

Освіта	Зарплата співробітників	Індивідуальний варіант
Вища	3200 3000 2600 2000 1900 1900	+2ij
Середня спец.	2600 2000 2000 1900 1800 1700	+3i
Середня	2000 2000 1900 1800 1700 1700	+(i+j)

2. Дослідник порівнює ефективність чотирьох різних методик навчання виробничим навикам. Для цієї мети зі всіх випускників ПТУ вибрано чотири групи учнів, що виучувалися, відповідно чотирма різними методами. Ефективність методик оцінювалася по сумі оброблених такими, що вчать деталей протягом одного дня (див. таблицю 29). Необхідно перевірити гіпотезу про відсутність впливу регульованого чинника (методик навчання) на продуктивність діяльності учня

Таблиця 29 - Ефективність чотирьох різних методик навчання виробничим навикам

№ особи, що навчається	1 група	2 група	3 група	4 група
1	60	75	60	95
2	80	66	80	85
3	75	85	65	100
4	80	80	60	80
5	85	70	86	
6	70	80	75	
7		90		

Вказівки до виконання завдання:

1. Ввести дані в таблицю, так аби в кожному стовпці виявилися дані, відповідні одному значенню досліджуваного чинника, а стовпці розташовувалися в порядку зростання (убування) величини досліджуваного чинника;

2. Виконати команду **Сервис > Анализ данных**;

3. У діалоговому вікні, що з'явилося, **Анализ данных** в списку **Инструменты анализа** вибрати процедуру **Однофакторный дисперсионный анализ**, потім натискувати кнопку **ОК**;

4. У діалоговому вікні, що з'явилося, задати **Входной интервал**, тобто ввести зазначення на діапазон аналізованих даних, що містить всі стовпці даних.

5. У розділі **Группировка** перемикач встановити в положення *по столбцам*;

6. Вказати **Выходной интервал**, тобто ввести зазначення на вічко, з яким будуть показані результати аналізу. Розмір вихідного діапазону буде визначений автоматично, і на екран буде виведено повідомлення в разі можливого накладення вихідного діапазону на вихідні дані. Натискувати кнопку **ОК**.

Вихідний діапазон включатиме результати дисперсійного аналізу: середні, дисперсії, критерій Фішера і інші показники. Вплив досліджуваного чинника визначається по величині значущості критерію Фішера, яка знаходиться в таблиці **Дисперсионный анализ** на пересіченні рядка **Между группами** і стовпця **P-значение**. У випадках, коли $P\text{-значение} < 0,05$, критерій Фішера значимий і вплив досліджуваного чинника можна визнавати доведеним.

Окрім розглянутої процедури однофакторного дисперсійного аналізу, для проведення двофакторного дисперсійного аналізу в пакеті аналізу реалізовані процедури **Двухфакторный дисперсионный анализ с повторениями** і **Двухфакторный дисперсионный анализ без повторений**.

Завдання 3

Група з 5 випробовуваних була обстежена за допомогою трьох експериментальних завдань, направлених на вивчення інтелектуальної, наполегливості. Кожному випробовуваному індивідуально пред'являлися послідовно три однакові анаграми: чотирьохбуквена, п'ятибуквена і шестибуквена. Чи можна вважати, що чинник довжини анаграми впливає на тривалість спроб її рішення?

Таблиця 30 - Тривалість вирішення анаграм (сек)

Код іспитуємого	Умова 1. Анаграма з чотирма літерами	Умова 2. Анаграма з п'ятьма літерами	Умова 3. Анаграма з шістьма літерами	Суми по іспитуємым	Індивідуальний варіант
1	5	235	7	247	+2i
2	7	604	20	631	+(i+j)
3	2	93	5	100	+i
4	2	171	8	181	+(i+j)
5	35	141	7	183	-0,5j
Суми	51	1244	47	1342	

Вказівки до виконання завдання

1. Сформулюємо гіпотези. Наборів гіпотез в даному випадку два.

Набір А.

H₀(А): Відмінності в тривалості спроб вирішення анаграм різної довжини є не більш вираженими, ніж відмінності, обумовлені випадковими причинами.

H₁(А): Відмінності в тривалості спроб вирішення анаграм різної довжини є більш вираженими, ніж відмінності, обумовлені випадковими причинами.

Набір Б.

H₀(Б): Індивідуальні відмінності між випробовуваними є не більш вираженими, ніж відмінності, обумовлені випадковими причинами.

H1(Б): Індивідуальні відмінності між випробовуваними є більш вираженими, ніж відмінності, обумовлені випадковими причинами.

2. Далі виробимо розрахунки по аналогії із завданням 2.

3. Після проведення розрахунків порівнюємо $F_{\text{емп_факт}}$ і $F_{\text{крит_факт}}$, $F_{\text{емп_исп}}$ та $F_{\text{крит_исп}}$ за результатами приймаємо гіпотези.

По завершенні формується висновок, є чи ні відмінності в об'ємі відтворення слів в різних умовах більш вираженими, ніж відмінності, обумовлені випадковими причинами. Також необхідно вказати чи є чи ні індивідуальні відмінності між випробовуваними більш вираженими, ніж відмінності, обумовлені випадковими причинами.

ПРАКТИЧНЕ ЗАНЯТТЯ 7

КОВАРІАЦІЙНИЙ АНАЛІЗ (ANCOVA)

Мета – набути навички в побудові статистичних оцінок та статистичних критеріїв для перевірки різних гіпотез з використанням коваріаційного аналізу.

Теоретична довідка

Коваріаційний аналіз - сукупність методів математичної статистики, моделей залежності середнього значення деякої випадкової величини Y , що відносяться до аналізу, від набору декількох чинників F і одночасно від набору кількісних чинників X . У коваріаційному аналізі по відношенню до Y змінні X називають супутніми; чинники F задають поєднання умов якісної природи, при яких отримані спостереження Y і X , і описуються за допомогою так званих індикаторних змінних; серед супутніх і індикаторних змінних можуть бути як випадкові, так і не випадкові (контрольовані в експерименті); якщо випадкова величина Y є вектором, то говорять про багатовимірний коваріаційний аналіз.

Основні теоретичні і прикладні проблеми ковариационного аналізу відносяться до лінійних моделей. Зокрема, якщо аналізуються n спостережень Y_1, \dots, Y_n з p супутніми змінними

$(x = (x^{(1)}, \dots, x^{(p)}))$, k можливими типами умов експерименту $(F=(f_1, \dots, f_k))$, то лінійна модель відповідного коваріаційного аналізу задається рівнянням:

$$Y_i = \sum_{j=1}^k f_{ij}\theta_j + \sum_{s=1}^p \beta_s(f_i)x_i^{(s)} + \varepsilon_i(f_i) \quad (7.1)$$

де $i = 1, \dots, n$, індикаторні змінні дорівнюють 1, якщо j -а умова експерименту мала місце при спостереженні, і рівні 0 в іншому випадку.

(f_{ij}) можуть відповідати результатам дихотомізації номінальної ознаки P з градаціями f_1, \dots, f_k ; номінальна ж ознака може бути складною: кожній його градації може відповідати поєднання значень деяких первинних, наприклад, узятих з анкети, ознак; коефіцієнти θ_j визначають ефект впливу j -го умови; $x_i^{(s)}$ - значення супутньої змінної, при якому отримано спостереження, $i = 1, \dots, n$; $s = 1, \dots, P$; $\beta_s(f_i)$ - значення відповідних коефіцієнтів регресії Y по $x^{(s)}$, взагалі кажучи, залежні від конкретного поєднання умов експерименту, тобто від вектора $f_i = (f_{i1}, \dots, f_{ik})$; $\varepsilon_i(f_i)$ - випадкові помилки, що мають нульові середні значення.

Основне призначення коваріаційного аналізу - використання в побудові статистичних оцінок $\theta_1, \dots, \theta_k$; β_1, \dots, β_p і статистичних критеріїв для перевірки різних гіпотез відносно значень цих параметрів. Якщо в моделі (1) постулювати апріорі $\beta_1 = \dots = \beta_p = 0$, то вийде модель аналізу дисперсійного; якщо з (1) виключити вплив некілкісних чинників (покласти $\theta_1 = \dots = \theta_k = 0$), то вийде модель аналізу регресійного.

Свою назвою коваріаційний аналіз зобов'язаний тій обставині, що в його обчисленнях використовується розбиття коваріації величин Y і X так само як в дисперсійному аналізі використовується розбиття суми квадратів відхилень.

Коваріаційний аналіз необхідний, якщо дослідник хоче вивчити переваги користувачів в групах з різним рівнем вжитку і рівнем лояльності, взявши до уваги відношення респондентів до складу продуктів харчування і до значення сніданку як способу їди. Дві останні змінні вимірюються за дев'ятибальною шкалою Лайкерта. В цьому випадку категоріальні незалежні змінні (вжиток продукту і лояльність до торгівельної марки) як і раніше

називаються чинниками, тоді як метричні незалежні змінні (відношення до складу продуктів харчування і значення, що додається сніданку) - ковариатами (covariates).

Ковариата (covariates) - метрична незалежна змінна, використовувана в ковариационному аналізі.

Завдання

Дані в таблиці 1 отримані з експериментального свинарника, пристосованого для індивідуального вигодовування 6 свиней в кожній з 5 загород. Від кожного з 5-ти опоросів було відібрано 6 поросят (3 самці і 3 самки) і розподілено по загородах. Було використано три склади корму. Позначимо їх через А, В, З в порядку, відповідному збільшенню вмісту ($\rho_A < \rho_B < \rho_C$) білка. У кожній загороді корм кожного складу видавався одному самцеві і одній самці. Поросята індивідуально зважували кожного тижня в течії 16 тижнів. Для кожного поросяти була обчислена швидкість приросту у фунтах на тиждень як кутовий коефіцієнт прямої, підбраної методом найменших квадратів. У таблиці ця швидкість позначена через v , а вага на початку тижня - через x . Необхідно провести ковариационний аналіз з лінійною регресією v на x , вважаючи, що коефіцієнти регресії однакові для кожного поросяти і що можливе взаємодія «кормпол», але немає інших взаємодій. Які пари з трьох досліджуваних складів кормів значимо розрізняються по S-методу з $b=10$?

Таблиця 31 – Корми для тварин

Загін	Змінна	А		В		С	
		Самець	Самка	Самець	Самка	Самець	Самка
1	у	9,52	9,94	8,51	10,00	9,11	9,75
	х	38	48	39	48	48	48
2	у	8,21	9,48	9,95	9,24	8,50	8,66
	х	35	32	38	32	37	28
3	у	9,32	9,32	8,43	9,34	8,90	7,63
	х	41	35	46	41	42	33
4	у	10,56	10,90	8,86	9,68	9,51	10,37
	х	48	46	40	46	42	50
5	у	10,42	8,82	9,20	9,67	8,76	8,57
	х	43	32	40	37	40	30

Вказівки до виконання завдання

1) На першому етапі необхідно провести регресійний аналіз. Для цього вибираємо меню Сервис - Анализ данных - Регрессия і у вікні, що відкрилося, вводимо значення x і y .

2) Далі знов вибираємо Сервис - Анализ данных - Ковариация і робимо висновок про те, який корм здійснює більший вплив на поросят залежно від статті.

3) Для того, щоб підтвердити або спростувати зроблені висновки вибираємо Сервис - Анализ данных - Двухфакторный дисперсионный анализ без повторений

4) Для виявлення помилки виконаємо наступні дії: Сервис - Анализ данных - Описательная статистика

5) На наступному етапі необхідно розробити лінійну модель коваріаційного аналізу. Зазначена модель буде мати наступний вигляд:

$$Y_{ij} = \mu + \alpha_i + \gamma z_{ij} + \delta z_{ij}^2 + e_{ij}$$

де μ – адитивна постійна (дорівнює 1);

α_i – головний ефект i -го виду корму;

γ – коефіцієнт регресії залежності ваги свиней від корму;

z_{ij} – стать свиней;

e – похибка.

Дані, необхідні для розробки моделі беруться за результатами проведених розрахунків з використанням пакету аналізу.

6) Висновки

ПРАКТИЧНЕ ЗАНЯТТЯ 8

СТАНДАРТИЗАЦІЯ ВИХІДНИХ ДАНИХ

Мета – здобуття навичок з проведення стандартизації вихідних даних.

Теоретична довідка

Стандартизація вихідних даних - в статистиці, цей термін визначає трансформацію даних. Віднімаючи кожне значення з деякого базисного значення (наприклад - вибіркоче середнє) і

ділячи на стандартне відхилення (зазвичай - вибіркоче стандартне відхилення). Це важливе перетворення всі вихідні значення (не звертає уваги на їх початкові розподіли і одиниці виміру) наводить до набору порівнянних елементів з розподілу з нульовим середнім і стандартним відхиленням рівним 1. Даний вигляд перетворення дуже широко застосовний, оскільки він спрощує механізм порівняння змінних. Крім того, застосування стандартизації до вихідних даних робить результати статистичних методів абсолютно незалежними їх розмахів і одиниць виміру.

Завдання

Необхідно провести стандартизацію даних про доходи на одного члена сім'ї і середньодушовий вжиток м'яса з коректировою на середньоквадратичне відхилення. За результатами розрахунків зробити висновки

Таблиця 1 - Співвідношення доходу на одного члена сім'ї з середнедушевым вжитком м'яса

Порядок виконання завдання

1. Необхідно розрахувати величини середньоквадратичних відхилень для X і Y.

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}$$

$$\sigma_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{N}} \quad (8.1)$$

2. На підставі отриманих величин середньоквадратического відхилення слід виконати стандартизацію значень X і Y за формулою, наведеною нижче:

$$t_{xi} = \frac{x_i - \bar{x}}{\sigma_x}$$

де t_x і t_y - це стандартизовані значення варіантів ряду X і Y відповідно.

Література

1. Закон України "Про державну статистику" Політвидавництво України, К. 2003.
2. Статистика: Підручник / за ред. Герасименка С.С., К.: КНЕУ, 2007.–467 с.
3. Статистика /за ред. Головача А.В. Київ. Вища школа. 2003.
4. Єріна А.М., Кальян З.О. Теорія статистики: Практикум. –К.: ТОВ «Знання», КОО, 2007 — с. 325
5. Елисеєва И.И., Юзбашев М.М. Общая теория статистики. – М.: Финансы и статистика, 2009. –480с.
6. Практикум по статистике: Учеб. Пособие / Под ред. Симчеры В.М. –М.:ЗАО «Финстатинформ», 2009. –259 с.
7. Статистика. Збірник задач. Навч. посібник / А.В. Головач и др. - К.: Вища школа, 2004.
8. Курс социально–экономической статистики / Под ред. Проф. М.Г. Назарова. – М.: «Финстатинформ», 2006. – 771 с.
9. Экономическая статистика: Учебник / Под ред. Ю.Н. Иванова. – М.: ИНФРВМ, 2007. – 487 с.
10. Фещур Р.В., Барвінський А.Ф., Кігор В.П., Статистика. Теоретичні засади і прикладні аспекти, Львів, 2003

Додатки
Додаток А
Вихідні дані (генеральна сукупність)

Номер	y	X_1	X_2		Номер	y	X_1	X_2
Група 1								
1	700	507	303		6	1010	766	387
2	754	571	324		7	1108	960	400
3	856	636	345		8	1126	895	428
4	964	701	366		9	1175	1090	491
5	994	831	407		10	1075	1155	456
Група 2								
1	1500	1025	470		11	1794	1414	596
2	1512	1087	523		12	1800	1381	586
3	1588	1252	544,		13	1777	1543	638
4	1552	1219	508		14	1853	1398	607
5	1688	1187	554		15	1825	1576	649
6	1674	1317	565		16	1888	1705	690
7	1736	1479	528		17	891	1511	628
8	1722	1349	496		18	1954	1689	701
9	1782	1548	669		19	1920	1608	659
10	1650	1673	680		20	1905	1770	598
Група 3								
1	2100	1610	727		21	2854	1950	860
2	2101	1806	738		22	2844	1895	800
3	2141	1835	684		23	2936	2025	850
4	2150	1803	722		24	3020	2143	832
5	2248	1848	753		25	3010	2175	842
6	2326	1884	748		26	3166	1980	862
7	2198	2013	790		27	3235	2256	869
8	2325	1867	705		28	3248	2402	916
9	2440	19S7	785		29	3204	2208	853
10	1650	1673	680		30	3216	2321	890
11	2514	1916	698		31	3362	2224	835
12	2623	1965	774		32	3414	2289	879
13	2584	2094	816		32	3427	2246	674
14	2651	1890	764		34	3469	2370	905
15	2620	1981	780		35	3505	2226	888
16	2665	1830	806		36	3S26	2528	957
17	2685	2058	844		37	3508	2451	931
18	2718	2205	879		38	354S	2337	895
19	2785	1994	798		39	3575	2434	926

20	2816	2029	795		40	3547	2645	980
----	------	------	-----	--	----	------	------	-----

Продовження Додатка А

Номер	Y	X_1	X_2		Номер	Y	X_1	X_2
Група 4								
1	3720	2364	915		26	4122	2736	1024
2	3714	2476	940		27	4172	2762	1032
3	3746	2606	972		28	4215	2839	1174
4	3760	2489	924		29	4240	2801	1045
5	3780	2502	948		30	4236	2686	947
6	3754	2418	984		31	4265	2671	1003
7	3800	2301	922		32	4280	2826	1203
8	3840	2541	951		33	4320	2782	1061
9	3848	2306	945		34	4299	2865	1066
10	3852	2554	965		35	4360	3001	995
11	3878	2507	969		36	4380	2791	1074
12	3900	2580	973		37	4420	2917	856
13	3879	2593	968		38	4440	2876	1087
14	3905	2559	999		39	4460	2943	1091
15	3894	2305	884		40	4471	2765	1095
16	3960	2619	386		41	4500	2969	1099
17	3880	2532	990		42	4490	2982	1064
18	3951	2697	983		43	4518	3110	1145
19	4011	2723	1020		44	4540	2995	1108
20	4021	2688	955		45	4530	2900	1032
21	4055	2536	1009		46	4545	3021	1116
22	4045	2708	1007		47	4564	3034	1120
23	3990	2794	1015		48	4521	3047	1124
24	4125	2788	1040		49	4560	3187	1170
25	4084	2904	1078		50	4587	3060	1129
Група 5								
1	4805	3073	1138		31	5165	3528	1280
2	4824	3086	1107		32	5176	3261	1202
3	4861	3099	1141		33	5210	3429	1248
4	4815	3121	1148		34	5227	3340	1235
5	4835	3231	1184		35	5244	3448	126
6	4819	3132	1152		36	5261	3462	1259
7	4835	3143	1155		37	5255	3473	1262
8	4865	2847	1159		38	5295	3484	1266
9	4884	3418	1244		39	5329	3506	1273
10	4822	3165	1162		40	5346	3517	1171
11	4819	3176	966		41	5339	3693	1333
12	4824	3286	1202		42	5380	3429	1284
13	4853	3198	1173		43	5378	3550	1287
14	4864	3374	1230		44	5414	3261	1291

Продовження Додатка А

Номер	Y	X_1	X_2		Номер	Y	X_1	X_2
16	4887	3220	1180		46	5420	3670	1326
17	4911	3142	1187		47	5415	3583	1298
18	4938	3253	1191		48	5465	3594	1201
19	4955	3264	1195		49	5465	3594	1201
20	4962	3175	1193		50	5484	3704	1337
21	5006	3297	1205		51	5518	3627	1312
22	5023	3208	1209		52	5533	3428	1242
23	5040	3319	1112		53	5542	3660	1323
24	5057	3330	1216		54	5601	3682	1330
25	5062	3241	1189		55	5580	3564	1289
26	5091	3352	1223		56	5652	3675	1311
27	5108	3305	1227		57	5648	3727	1344
28	5126	3451	1255		58	5670	3649	1319
29	5112	3385	1206		59	5656	3738	1348
30	5159	3396	1237		60	5635	3771	1358
Група 6								
1	5703	3507	1351		31	6175	4155	1483
2	5717	3758	1314		32	6211	4078	1458
3	5336	3791	1299		33	6198	4100	1465
4	5783	3924	1408		34	6262	4111	1381
5	5765	3802	1369		35	6258	4276	1482
6	5803	3784	1372		36	6296	3699	1409
7	5824	3824	1312		37	6313	4144	1479
8	5837	3835	1379		38	6347	4166	1486
9	5840	3688	1361		39	6325	4321	1425
10	5854	3446	1383		40	6364	4177	1356
11	5849	3890	1407		41	6281	4188	1493
12	5865	3998	1458		42	6387	3572	1294
13	5871	3857	1387		43	6398	4098	1497
14	5858	3946	1415		44	6415	4210	1501
15	5888	3879	1394		45	6449	4232	1508
16	5939	3847	1275		46	6426	4287	1525
17	5954	4045	1447		47	6466	4243	1511
18	5990	3890	1364		48	6483	4152	1483
19	5999	3935	1411		49	6498	4320	1536
20	5984	3968	1422		50	6551	4195	1529
21	6008	4012	1436		51	6535	4434	1573
22	6024	4067	1454		52	6558	4265	1620
24	6015	3913	1404		54	6555	4459	1524
25	6058	3979	1426		55	6602	4331	1540
26	6042	4001	1433		56	6594	4302	1543

Продовження Додатка А

Номер	Y	X_1	X_2		Номер	Y	X_1	X_2
27	6126	3999	1440		57	6630	4362	1550
28	6152	4122	1501		58	6650	4375	1554
29	6077	3892	1451		59	6681	4386	1558
30	6185	4089	1461		60	6674	4395	1560
Група 7								
1	6700	4395	1560		26	7140	4680	1653
2	6704	4459	1501		27	7131	4591	1657
3	6720	4108	1564		28	7186	4796	1690
4	6742	4485	1590		29	7167	4719	1665
5	6736	4702	1568		30	7220	4732	1669
6	6780	4446	1577		31	7222	4835	1703
7	6758	4511	1598		32	7235	4874	1725
8	6820	4372	1585		33	7280	4770	1712
9	6798	4421	1569		34	7266	4783	1686
10	6723	4369	1502		35	7362	4809	1694
11	6860	4498	1594		36	7355	4688	1724
12	6857	4267	1648		37	7360	4822	1699
13	6880	4337	1886		38	7358	4661	1711
14	6894	4602	1627		39	7480	4900	1724
15	6871	4524	1602		40	7385	4839	1704
16	6900	4563	1615		41	7426	4865	1545
17	6940	4550	1611		42	7464	4890	1720
18	6964	4376	1619		43	7512	4921	1645
19	6995	4628	1606		44	7581	4766	1745
20	7001	4606	1566		45	7562	4953	1741
21	6994	4289	1603		46	7620	4991	1621
22	7040	4512	1632		47	7624	4993	1754
23	7024	4641	1640		48	7660	5017	1654
24	7114	4654	1444		49	7658	5015	1761
25	7102	4758	1678		50	7676	5027	1487
Група 8								
1	7725	5059	1568		21	8225	5383	1780
2	7758	5172	1782		22	8215	5376	1878
3	7775	5091	1679		23	8176	5351	1870
4	7812	5115	1793		24	8246	5331	1884
5	7794	5104	1790		25	8261	5406	1807
6	7850	5140	1801		26	8325	5684	1901
7	7836	5131	1798		27	8305	5435	1897
8	7784	5097	1787		28	8354	5704	1825
9	7925	5188	1817		29	8365	5474	1909
10	7888	5164	1457		30	8406	5500	1918

Продовження Додатка А

Номер	Y	X_1	X_2		Номер	Y	X_1	X_2
11	7975	5015	1828		31	8366	5047	1909
12	8004	5240	1834		32	8425	5512	1922
13	7997	5235	1832		33	8462	5536	1930
14	8050	5168	1814		34	8503	5563	1938
15	8075	5286	1848		35	8499	5560	1937
16	8100	5302	1854		36	8518	5418	2001
17	8094	5294	1655		37	8535	5584	1924
18	8150	5334	1864		38	8584	5558	1955
19	8175	5350	1869		39	8590	5619	1904
20	8204	5069	1876		40	8606	5586	1960
Група 9								
1	8801	5658	1963		11	9150	5982	1974
2	8829	5774	2007		12	9206	9206	2086
3	8850	5788	2011		13	9256	9256	2096
4	8900	5820	2021		14	9288	9288	2026
5	8895	5953	2054		15	9360	9360	2118
6	8953	6065	2033		16	9457	9457	2138
7	9051	5918	2053		17	9488	9488	2145
8	9100	5950	2063		18	9506	9506	2088
9	9126	5967	2069		19	9562	9562	2160
10	9167	5993	2077		20	9558	9558	2005
Група 10								
1	9884	7215	2545		6	9979	7315	2418
2	9867	6447	2531		7	9956	6505	2836
3	9905	6438	3087		8	9989	6526	3120
4	9964	6324	2482		9	9986	8714	2249
5	9948	8554	2507		10	9994	8012	2099

Додаток Б

Таблиці критичних значень для перевірки значимості коефіцієнту конкордації при малій кількості оцінюваних об'єктів (< 7)¹

Критичні значення S для рівня значимості (альфа) = 0,05						Додаткові значення S для n=3 (альфа=0,05)	
Експерти	n					m	S
m	3	4	5	6	7	m	S
3			64,4	103,9	157,3	9	54
4		49,5	88,4	143,3	217	12	71,9
5		62,6	112,3	182,4	276,2	14	93,8
6		75,7	136,1	221,4	335,2	16	95,8
8	48,1	101,7	183,7	299	453,1	18	107,7
10	60	127,8	231,2	376,7	571		
15	89,8	192,9	349,8	570,5	864,9		
20	119,7	258	468,5	764,4	1158,7		

Критичні значення W для рівня значимості (альфа) = 0,05						Додаткові значення S для n=3 (альфа=0,05)	
Експерти	n					m	W
m	3	4	5	6	7	m	W
3	--	--	0,7156	0,6597	0,6242	9	0,3333
4	--	0,6188	0,5525	0,5118	0,4844	12	0,2497
5	--	0,5008	0,4492	0,4169	0,3946	14	0,2393
6	--	0,4206	0,3781	0,3514	0,3325	16	0,1871
8	0,3758	0,3178	0,2870	0,2670	0,2528	18	0,1662
10	0,3000	0,2556	0,2312	0,2153	0,2039		
15	0,1996	0,1715	0,1555	0,1449	0,1373		
20	0,1496	0,1290	0,1171	0,1092	0,1035		

Критичні значення S для рівня значимості (альфа) = 0,01						Додаткові значення S для n=3 (альфа=0,01)	
Експерти	n					m	S
m	3	4	5	6	7	m	S
3			75,6	122,8	185,6	9	75,9
4		61,4	109,3	176,2	265	12	103,5
5		80,5	142,8	229,4	343,8	14	121,9
6		99,5	176,1	282,4	422,6	16	140,2
8	66,8	137,4	242,7	388,3	579,9	18	158,6
10	85,1	175,3	309,1	494	737		
15	131	269,8	475,2	758,2	1129,5		
20	177	364,2	641,2	1022,2	1521,9		

Критичні значення W для рівня значимості (альфа) = 0,01						Додаткові значення S для n=3 (альфа=0,01)	
Експерти	n					m	W
m	3	4	5	6	7		
3	--	--	0,8400	0,7797	0,7365	9	0,4685
4	--	0,7675	0,6831	0,6293	0,5915	12	0,3594
5	--	0,6440	0,5712	0,5243	0,4911	14	0,3110
6	--	0,5528	0,4892	0,4483	0,4192	16	0,2738
8	0,5219	0,4294	0,3792	0,3467	0,3236	18	0,2448
10	0,4255	0,3506	0,3091	0,2823	0,2632		
15	0,2911	0,2398	0,2112	0,1926	0,1793		
20	0,2213	0,1821	0,1603	0,1460	0,1359		

Методичні вказівки до виконання практичних занять з
дисципліни «Статистика» для студентів спеціальності 07.050107
«Економіка підприємства»

Укладачі
Бредихін Володимир Михайлович
Першина Ганна Олександрівна
Вербицька Вікторія Іванівна

Відповідальний за випуск
І.А.Дмитрієв

Підп. до друку	Формат 60*80	1/16 Папір
тип. Друк офсетний.	Розум. печ. л. 1,4	Навч. вид.л.
1,5 Зак. №	Тираж 200 экз.	Ціна договірна

ХНАДУ, 61002, Харків, вул. Петровського, 25

Підготовлено до друку РІО Харківського національного
автомобільно-дорожнього університету