

## РАЗДЕЛ «ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ»

УДК 519.24: 330.43

АВРАМЕНКО В.І., к.т.н., доцент

Дніпродзержинський державний технічний університет

### ДОСЛІДЖЕННЯ ПОХИБОК ВИЗНАЧЕННЯ ПАРАМЕТРІВ РЕГРЕСІЙ ПОПИТУ І ПРОПОЗИЦІЇ

**Вступ.** Багато залежностей в економіці мають суттєво нелінійний характер. До таких відносяться зміна суми надходжень до бюджету в залежності від податкової ставки (крива Лаффера), залежності попиту на товари різних груп (криві Торнквіста), виробничі функції різного виду і т.п. За відсутності детермінованих теоретичних формул таких залежностей їх вираз для конкретних умов звичайно шукається за результатами обробки статистичних даних. Зрозуміло, що результати такої обробки містять як систематичні, так і випадкові похибки, які суттєво можуть вплинути на прогнозовані значення.

**Постановка задачі.** Метою даної роботи є оцінка впливу похибок вихідних даних на статистичну обґрунтованість результатів статистичної обробки. Наведені в літературі [1, 2, 3] оцінки відносяться до лінійних економетричних функцій. Нами були вибрані суттєво нелінійні залежності попиту і пропозиції від ціни за одиницю товару, кожна з залежностей містить три невідомих параметри.

Дослідження впливу похибок виконувалось методом статистичного моделювання. Задаючись видом залежності, для деяких конкретних значень параметрів можна отримати теоретичні еталонні криві. Далі до еталонних кривих додавались випадкові збурення, розподілені за нормальним законом з параметрами  $a=0$  і  $\sigma \neq 0$ . Отримані дані оброблялись методом найменших квадратів з метою оцінки параметрів залежності заданого виду, значення яких порівнювались з заданими параметрами. Оцінювались похибки у визначенні параметрів і залежності похибок від дисперсії збурення  $\sigma^2$  і довжини вибірки.

В роботі використані функції і інструменти табличного процесора MS Excel і програмні засоби Delphi.

**Результати роботи.** Аналіз функції попиту. Залежність обсягів попиту  $Y=\{y_i\}$  від ціни  $X=\{x_i\}$  слід розглядати [1] як спадну функцію, обмежену знизу. Будемо шукати її у вигляді обмеженої експоненти, а саме

$$\hat{y}_i = y_0 + a \cdot \exp(b \cdot x_i) \quad b < 0, \quad (1)$$

де  $a$ ,  $b$  та  $y_0$  – невідомі параметри залежності; величину  $y_0$  можна розглядати як мінімальні рентабельні обсяги виробництва. На рис.1 наведено типовий графік залежності попиту від ціни для модельного прикладу, що буде розглядатись в якості еталонного:

$$\hat{y}_i = 35 + 50 \cdot \exp(-0,1 \cdot x_i) \quad x_i \in (10, 20). \quad (2)$$

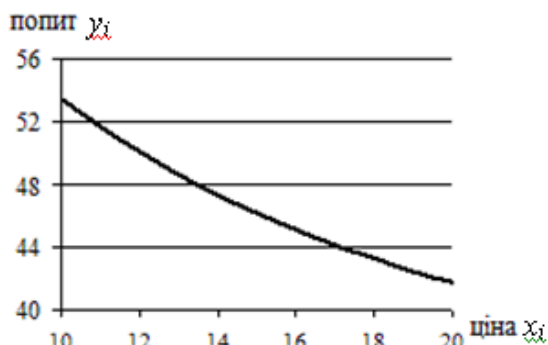


Рисунок 1 – Графік залежності попиту у від ціни  $x$

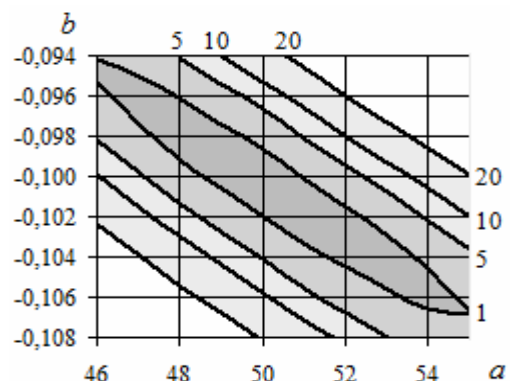


Рисунок 2 – Ізолінії рельєфу функції  $F(a, b, 35)$

При обробці результатів статистичного експерименту, коли є пари спостережень  $\{x_i, y_i\}, i=1 \dots n$ , де  $n$  – кількість спостережень, невідомі параметри визначаються звичайно методом найменших квадратів, коли шукається мінімум функціоналу

$$F(a, b, y_0) = \sum_{i=1}^n (y_i - (y_0 + a \cdot \exp(b \cdot x_i)))^2 \Rightarrow \min. \quad (3)$$

На рис.2 наведені ізолінії рельєфу функції  $F(a, b, 35)$ . Розрахунки показали, що зміна кожного параметра регресії суттєво змінює координати мінімуму  $a^*, b^*, y_0^*$  навіть за відсутності випадкових збурень. Координати мінімуму найбільш чутливі до зміни вільного члена залежності  $y_0$ . Збільшення вільного члена  $y_0$  на 0,35 (1%) одночасно збільшує величину  $a^*$  на 2,1% і зменшує  $b^*$  на 3,7%.

Для знаходження конкретних значень параметрів регресії для різних вибірок розв'язується система нормальних рівнянь [2], яка для рівняння регресії (1) має вигляд

$$\begin{cases} \sum_{i=1}^n y_i - n \cdot y_0 - a \cdot \sum_{i=1}^n \exp(b \cdot x_i) = 0 \\ \sum_{i=1}^n y_i \cdot \exp(b \cdot x_i) - y_0 \cdot \sum_{i=1}^n \exp(b \cdot x_i) - a \cdot \sum_{i=1}^n (\exp(b \cdot x_i))^2 = 0 \\ \sum_{i=1}^n y_i \cdot x_i \cdot \exp(b \cdot x_i) - y_0 \cdot \sum_{i=1}^n x_i \cdot \exp(b \cdot x_i) - a \cdot \sum_{i=1}^n x_i \cdot (\exp(b \cdot x_i))^2 = 0 \end{cases} \quad (4)$$

Позначивши  $u_i(b) = \exp(b \cdot x_i)$  і послідовно розв'язуючи систему, отримуємо нелінійне рівняння, яке містить тільки невідомий параметр  $b$  під знаком експоненти

$$\begin{aligned} & \left( n \cdot \sum_{i=1}^n u_i^2 - \left( \sum_{i=1}^n u_i \right)^2 \right) \cdot \left( n \cdot \sum_{i=1}^n x_i y_i u_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i u_i \right) - \\ & - \left( n \cdot \sum_{i=1}^n y_i u_i - \sum_{i=1}^n y_i \sum_{i=1}^n u_i \right) \cdot \left( n \cdot \sum_{i=1}^n x_i u_i^2 - \sum_{i=1}^n u_i \sum_{i=1}^n x_i u_i \right) = 0 \end{aligned} \quad (5)$$

Розв'язуючи рівняння (5) числовими методами з заданою похибкою, отримуємо значення  $b^*$ , після чого можна обчислити величини  $a^*$  і  $y_0^*$  за формулами

$$a^* = \frac{n \cdot \sum_{i=1}^n y_i u_i^* - \sum_{i=1}^n y_i \sum_{i=1}^n u_i^*}{n \cdot \sum_{i=1}^n u_i^{*2} - \left( \sum_{i=1}^n u_i^* \right)^2}; \quad y_0^* = \frac{\sum_{i=1}^n y_i - a^* \cdot \sum_{i=1}^n u_i^*}{n}. \quad (6)$$

Оцінка похибок визначення параметрів регресії за результатами моделювання виконувалась методами статистичних випробувань. Випадковим чином задавались координати  $x_i$  для  $n$  точок, рівномірно розподілених на інтервалі  $x_i \in (10, 20)$ . У вибраних точках обчислювались еталонні значення функції попиту (2), до яких додавались випадкові збурення  $\delta_i$ , розподілені за нормальним законом з параметрами  $M(\delta)=0, D(\delta)=\sigma^2, \sigma^2 \in (0,1 \div 1,5)$ . Отримана таким чином конкретна реалізація оброблялась за формулами (5), (6) і результати порівнювались з використаними еталонними значеннями параметрів залежності. Для досягнення статистично обґрунтованих висновків кількість реалізацій для кожного варіанту складала не менше 1000. Значення дисперсії збурень  $\sigma^2$  визначають величину коефіцієнта детермінації  $R^2$  отриманої регресії.

В табл.1 наведені деякі статистичні характеристики відхилень параметрів регресії від еталонних для вибірки з 1000 послідовностей, кожна довжиною  $n=15$  при  $\sigma^2=0,25$  ( $R^2=0,96$ ); на рис.3 показані полігони відносних частот  $w_i$  відхилень параметрів  $b^*$ ,  $a^*$  і  $y_0^*$  від еталонних.

Таблиця 1 – Статистичні характеристики відхилень параметрів регресії попиту

Статистичні характеристики	Відхилення $\Delta b$	Відхилення $\Delta a$	Відхилення $\Delta y_0$
Середнє значення	-0,0027	4,542	-1,31
Медіана	-0,0003	0,70	0,01
Асиметрія	-0,491	5,28	-2,92
Сер. квадрат. відхилення	0,0348	12,54	5,71

Наведені результати свідчать, що розподіл відхилів  $\Delta b$  близький до зсуеного симетричного, розподіли інших відхилень  $\Delta a$  і  $\Delta y_0$  є суттєво асиметричними, причому асиметрія відхилів коефіцієнта  $a$  – додатна ( $K_{ac}=5,28$ ), асиметрія відхилів вільного члена  $y_0$  є від’ємною ( $K_{ac}=-2,92$ ). Про асиметричність розподілу цих відхилень свідчить також суттєва різниця між вибірковим середнім і медіаною цих відхилів в табл.1.

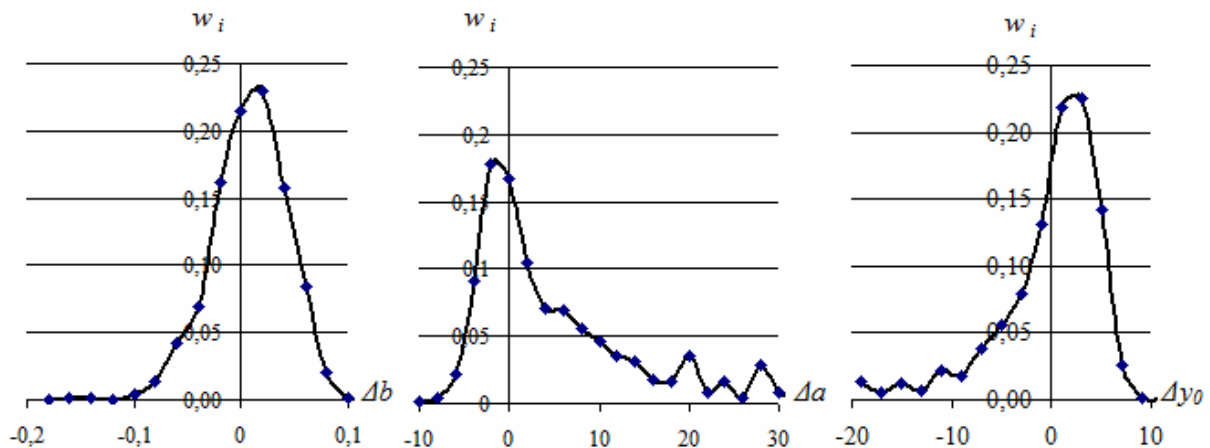


Рисунок 3 – Полігони відхилів параметрів функції попиту

Крім того, значення відхилів різних параметрів суттєво пов’язані між собою, про що свідчить кореляційна матриця, наведена в табл.2. Невелике значення коефіцієнта кореляції між відхиленнями  $\Delta a$  і  $\Delta y_0$  пояснюється нелінійним характером цієї залежності, про що свідчить кореляційне поле, наведене на рис.4.

Таблиця 2 – Кореляційна матриця відхилень параметрів регресії

	$\Delta b$	$\Delta a$	$\Delta y_0$
$\Delta b$	1,00	-0,76	-0,83
$\Delta a$	-0,76	1,00	0,35
$\Delta y_0$	-0,83	0,35	1,00

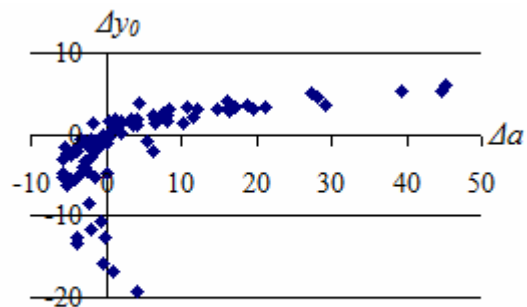


Рисунок 4 – Кореляційне поле відхилів

Досліджувалась залежність систематичних і випадкових відхилень параметрів функції попиту від довжини вибірових реалізацій та коефіцієнту детермінації. На рис.5 наведені графіки, які відображають зміну систематичних похибок.

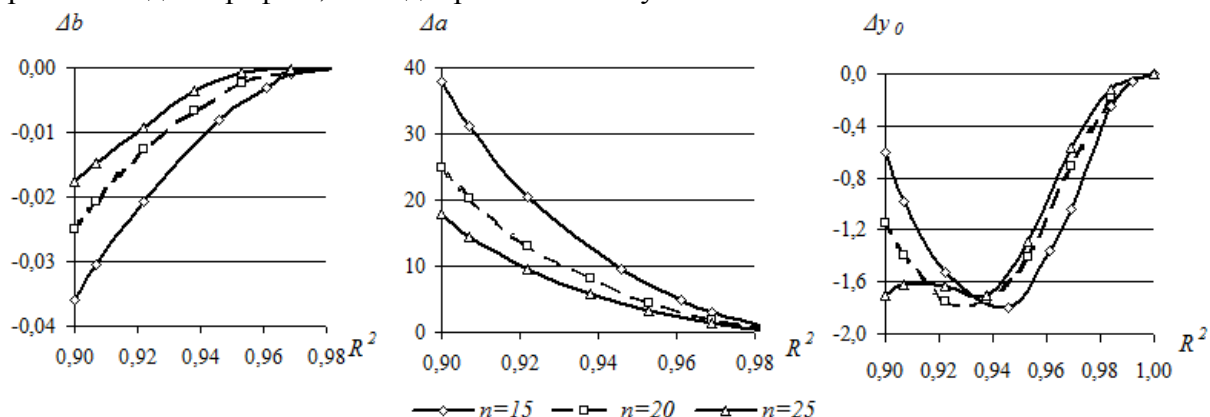


Рисунок 5 – Графіки зміни систематичних похибок визначення параметрів попиту

Як і слід очікувати, систематичні похибки швидко зростають при зменшенні довжини вибірки  $n$  і зменшенні коефіцієнту детермінації  $R^2$  (збільшенні дисперсії збурень  $D(\delta)=\sigma^2$ ). Відхилення параметра  $b$  і коефіцієнта  $a$  від еталонних є монотонними функціями, в той час як залежність зміни вільного члена  $y_0$  від коефіцієнта детермінації є більш складною, що можна пояснити нелінійним взаємозв'язком між цими параметрами регресії.

Випадкові відхилення оцінок параметрів регресії оцінюються середніми квадратичними відхиленнями. В усіх випадках похибки обчислення параметрів регресії зростають при зменшенні коефіцієнту детермінації і зменшенні об'єму вибірки. Для малих вибірок ( $n=15$ ) при  $R^2=0,90$  середні квадратичні відхилення досягають 70% величини параметра  $b$ , 80% величини коефіцієнта  $a$  і 40% величини вільного члена  $y_0$ .

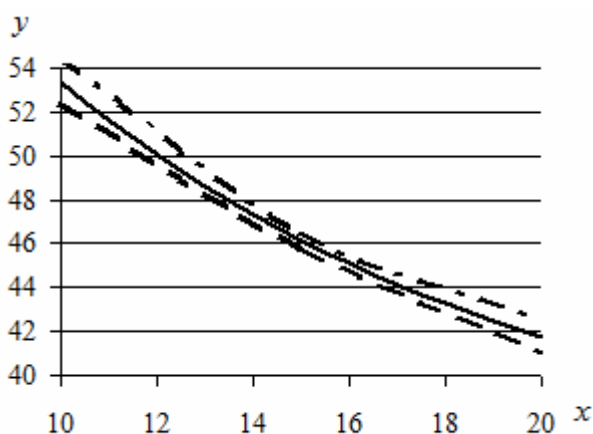


Рисунок 6 – Довірчий інтервал регресії попиту

Для ілюстрації сумісного впливу корельованих факторів на рис.6 побудовано довірчий інтервал функції регресії попиту для випадку  $m=15$ , довірна ймовірність  $\gamma=1-\alpha=0,95$ . Як видно на графіку, зона надійності суттєво несиметрична відносно еталонної кривої. В табл.3 наведені значення границь зони надійності на кінцях і в середині інтервалу зміни значень  $x$  ціни за одиницю товару. Як слідує з табл.3, з імовірністю 0,95 навіть на кінцях інтервалу відносна похибка регресійного значення не відхиляється від еталонного більш ніж на 2,2%.

Таблиця 3 – Відхили границь довірчого інтервалу функції попиту

Ціна $x$	Верхня границя надійності				Нижня границя надійності			
	$\tilde{y}_{верх}$	$y_{ет}$	$\tilde{y}-y_{ет}$	%	$\tilde{y}_{нижн}$	$y_{ет}$	$\tilde{y}-y_{ет}$	%
10	54,57	53,39	1,18	2,2	52,44	53,39	-0,95	1,8
15	46,54	46,16	0,38	0,8	45,77	46,16	-0,39	0,8
20	42,56	41,77	0,75	1,9	41,05	41,77	-0,72	1,7

Такий результат можна вважати досить несподіваним: при суттєвих відхиленнях від еталонного кожного параметру залежності окремо їх сумісний вплив дозволяє отримати несуттєві відхилення значень функції попиту від еталонної.

Аналіз функції пропозиції. Із загальних уявлень про природу залежності обсягів пропозиції на ринку від ціни  $x$  за одиницю продукції [3] функція пропозиції  $Z$  є зростаючою функцією із спадним темпом зростання. Розглянемо степеневу функцію

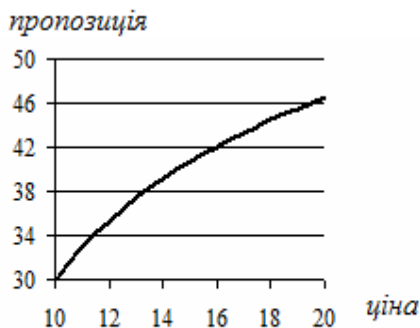


Рисунок 7 – Графік залежності пропозиції  $z$  від ціни  $x$

$$z_i = c \cdot (x_i - x_0)^d \quad 0 < d < 1, \quad (7)$$

де  $c, d, x_0$  – деякі параметри регресії, причому  $x_0$  має зміст мінімальної рентабельної ціни за одиницю продукції. На рис.7 наведено графік функції пропозиції при  $c=25, d=0,25, x_0=8,0, x_i \in (10, 20)$ .

За результатами статистичних спостережень невідомі параметри регресії шукаються методом найменших квадратів, коли мінімізується функція

$$F(c, d, x_0) = \sum_{i=1}^n (y_i - c \cdot (x_i - x_0)^d)^2 \Rightarrow \min. \quad (8)$$

Система нормальних рівнянь має вигляд:

$$\begin{aligned} \sum_{i=1}^n z_i \cdot (x_i - x_0)^d - c \cdot \sum_{i=1}^n (x_i - x_0)^{2d} &= 0 \\ \sum_{i=1}^n z_i c (x_i - x_0)^d \ln(x_i - x_0) &= c^2 \sum_{i=1}^n (x_i - x_0)^{2d} \ln(x_i - x_0). \\ \sum_{i=1}^n z_i \cdot c \cdot (x_i - x_0)^{d-1} - c^2 \sum_{i=1}^n (x_i - x_0)^{2d-1} &= 0 \end{aligned} \quad (9)$$

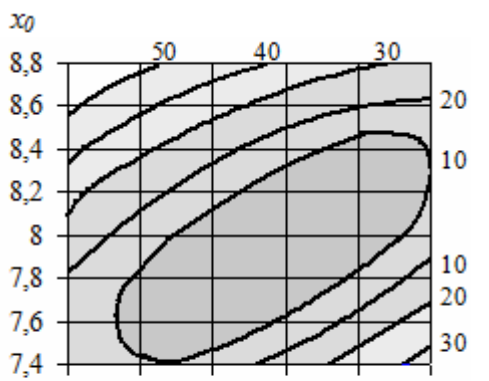


Рисунок 8 – Ізолінії рельєфу функції  $F(25, d, x_0)$

На рис.8 наведено ізолінії рельєфу цільової функції  $F$ . На відміну від функції попиту (рис.2) мінімум виражений більш чітко, в той же час чутливість значень параметрів ( $c^*, d^*, x_0^*$ ) мінімуму до зміни інших параметрів достатньо суттєва. Так, при збільшенні множника  $c$  на 1% значення показника степеню  $d^*$  зменшується на 2,56%, одночасно величина  $x_0^*$  збільшується на 0,89%.

Розв'язок системи рівнянь (9) є досить складним, тому продуктивним виявився пошук числовими методами безпосередньо мінімуму функції  $F(c, d, x_0)$ , визначеної рівнянням (8).

Послідовність обчислень для оцінки похибок параметрів не відрізняється від вище наведеної для функції попиту.

В табл.4 наведені деякі статистичні характеристики відхилень обчислених значень параметрів залежності від еталонних для варіанту  $R^2 = 0,95$ , довжини вибірки  $n=15$ , кількості реалізацій  $m=1000$ . Невелика різниця між середніми вибірковими значеннями, статистичними медіаною і модою свідчить про симетричність розподілів відхилень емпіричних оцінок від еталонних. В той же час суттєві значення коефіцієнту ексцесу є ознакою гостровершинності розподілів. Отже, для всіх параметрів залежності великі відхилення від еталонних, більші за „три сигми”, мають більшу ймовірність для емпіричних розподілів у порівнянні з нормальним.

Таблиця 4 – Статистичні характеристики відхилень параметрів регресії пропозиції

Статистичні характеристики	Відхилення параметрів		
	$\Delta c$	$\Delta d$	$\Delta x_0$
Середнє значення	0,051	-0,000	-0,015
Медіана	-0,020	0,000	-0,001
Мода	-0,041	0,00	-0,021
Коефіцієнт ексцесу	1,53	1,72	1,49
Сер. квадрат. відхилення	2,20	0,0322	0,693

Обчислення кореляційної матриці відхилень параметрів показало майже функціональну лінійну залежність між значеннями відхилів різних параметрів функції пропозиції (парні коефіцієнти кореляції близькі по модулю до 1,00), причому між відхиленнями параметра  $x_0$  і множника  $c$  залежність позитивна, інші співвідношення – негативні.

Виконані обчислення по оцінці похибок відхилень емпіричних характеристик від еталонних підтвердили дані табл.4 про відсутність систематичних похибок для різних довжин вибірки  $n$  і значень коефіцієнтів детермінації  $R^2$ . В той же час випадкові

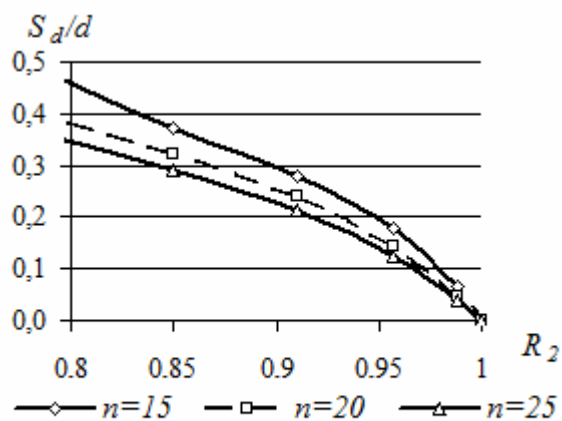


Рисунок 9 – Графік відносного стандартного відхилення показника  $d$

похибки визначення параметрів пропозиції, які оцінюються середніми квадратичними відхиленнями  $S_i$ , суттєво зростають при зменшенні об'єму вибірки  $n$  і зменшенні коефіцієнта детермінації  $R^2$ . На рис.9 наведені графіки залежності відносної випадкової похибки обчислення параметра  $d$  від довжини вибірки  $n$  і коефіцієнта детермінації  $R^2$  для показника степеню. Як свідчать графіки, при  $R^2 \approx 0,8$  відносні похибки оцінки показника степеню  $d$  досягають 50%. Для інших параметрів похибки міняються за такою ж закономірністю, досягаючи 30%.

Сумісний вплив корельованих факторів відображений в табл.5 та на рис.10, де побудовано довірчий інтервал функції регресії попиту для випадку  $n=15$ , довірна ймовірність  $\gamma = 1 - \alpha = 0,95$ . Як видно на графіку, зона надійності суттєво несиметрична відносно еталонної кривої. В табл.5 наведені значення границь зони надійності на кінцях

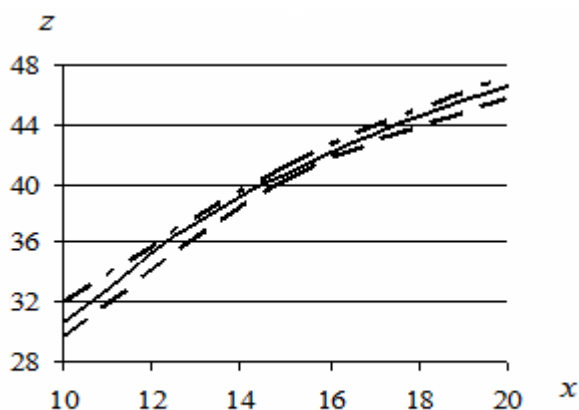


Рисунок 10 – Довірчі зони функції пропозиції

( $x=10$  і  $x=20$ ) і в середині інтервалу ( $x=15$ ) зміни значень ціни за одиницю товару. Як слідує з таблиці, з імовірністю 0,95 на кінцях інтервалу відносна похибка регресійного значення відхиляється від еталонного і сягає 4,5%, причому відхилили на початку інтервалу, при  $x=10$ , більші, ніж в кінці. Зрозуміло, що при зменшенні коефіцієнта детермінації інтервал надійності зростає, при збільшенні об'єму вибірки інтервал надійності зменшуватиметься.

Таблиця 5 – Абсолютні і відносні відхили границь довірчого інтервалу функції пропозиції

Ціна $x$	Верхня границя надійності				Нижня границя надійності			
	$\tilde{z}_{\text{аадо}}$	$z_{em}$	$\tilde{z}_{\text{аадо}} - z_{em}$	%	$\tilde{z}_{\text{іеаі}}$	$z_{em}$	$\tilde{z}_{\text{іеаі}} - z_{em}$	%
10	32,07	30,70	1,37	4,5	29,6	30,70	-1,10	3,6
15	41,16	40,66	0,50	1,2	40,23	40,66	-0,43	1,1
20	47,27	46,53	0,74	1,6	45,72	46,53	-0,81	1,7

**Висновки.** Виконані методом статистичних випробувань обчислення систематичних і випадкових похибок визначення параметрів функцій попиту і пропозиції дозволили кількісно оцінити їх залежність як від довжини вибірки, так і від величини коефіцієнта детермінації. Для досягнення статистично обґрунтованих результатів використання вибірок довжиною  $n \leq 20$  і коефіцієнта детермінації  $R^2 \leq 0,9$  недоцільне. І навіть при виконанні цих умов інтерпретація кожного параметра залежності окремо, в тому числі тих, що мають економічний зміст, не є виправданим внаслідок великих систематичних і/або випадкових похибок. В той же час взаємно компенсуючий вплив різних параметрів дозволяє оцінити функції регресії з достатньою для використання точністю.

## ЛІТЕРАТУРА

1. Назаренко О.М. Основи економетрики: підручник / О.М. Назаренко. – К.: „Центр навчальної літератури”, 2004. – 392с.
2. Айвазян С.А. Прикладная статистика и основы эконометрики / С.А. Айвазян, В.С. Мхитарян. – М.: Издательское объединение „ЮНИТИ”, 1998. – 1022с.
3. Красс М.С. Математика для экономистов / М.С. Красс, Б.П. Чупринов. – СПб.: Питер, 2007. – 464с.

УДК 65.011.56:681.3

ТИТЮК В.К., к.т.н.

МИХАЙЛЕНКО О.Ю., аспірант

Криворізький технічний університет

### АНАЛІЗ ПРОДУКТИВНОСТІ БАГАТОСТАДІЙНИХ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ З УРАХУВАННЯМ СТАТИСТИЧНИХ ВІДХИЛЕНЬ ПРОДУКТИВНОСТІ ОКРЕМИХ СТАДІЙ

**Вступ.** Періодичним є технологічний процес, в якому за порівняно невеликий проміжок часу (години або дні) виробляється певна, обмежена кількість кінцевого продукту. При цьому протягом відведеного проміжку часу періодичний процес є безперервним.

В [1] дається наступне визначення для періодичного режиму роботи технологічного процесу: «Режими, в которых характеризующие процесс переменные периодичны, а состояния процесса испытывают разрывы, называют периодическими. Обычно такие разрывы связаны с загрузкой и выгрузкой продуктов из аппарата, что сопровождается прекращением процесса и остановкой аппарата на некоторое время».

До широкого класу періодичних технологічних процесів (ТП) відносяться процеси плавки в металургії, нагрів зливків в нагрівальних колодязях і камерних печах перед гарячою обробкою тиском (плющення, пресування, кування), транспортування вантажів, процеси металообробки, виробництво полімерів і багато інших. Часто виробни-