Лабораторна робота № 1

Мета роботи: навчитися будувати лінійну двофакторну економетричну модель класичним методом МНК за допомогою табличного редактору Microsoft Excel та виконувати економічний аналіз характеристик взаємозв'язку.

Приклад виконання лабораторної роботи

Завдання. За допомогою табличного редактора Microsoft Excel на основі даних про роздрібний товарообіг і доходи населення у грошових одиницях (табл. 1.1) побудувати та дослідити економетричну модель парної лінійної регресії. Необхідно:

1. Виконати ідентифікацію змінних та специфікацію моделі.
2. Оцінити параметри моделі методом МНК.
3. Побудувати базову таблицю дисперсійного аналізу (ANOVA − таблицю) та визначити дисперсії.
4. Визначити коефіцієнти детермінації, кореляції та еластичності.
5. Перевірити статистичну значущість коефіцієнта детермінації, коефіцієнта кореляції та визначити його надійний інтервал.
6. Визначити стандартні помилки оцінок параметрів моделі та оцінити їх.
7. Перевірити статистичну значущість оцінок параметрів моделі та визначити їх надійні інтервали.
8. Перевірити адекватність економетричної моделі.
9. Визначити надійні інтервали базисних значень.
10. Визначити точковий та інтервальний прогнози для заданого останнього значення незалежної змінної.
11. Побудувати графіки фактичних даних, лінію регресії, її надійні інтервали та лінію тренда.
12. Перевірити точність економетричної моделі за допомогою середньої відносної похибки апроксимації.
13. Побудувати економетричну модель за допомогою функції ЛИНЕЙН.
14. Виконати економіко-математичний аналіз характеристик економетричної моделі.

Таблиця 1.1 – Вихідні дані

№	Доходи населення, ум.од.	Роздрібний товарообіг, ум.од.
1	31	27
2	18	17
3	29	26
4	20	18
5	21	19
6	25	23
7	24	21
8	28	25

9	22	20
10	27	24
Прогнозне значення	$x_{np} = 34{,}1$	

Розв'язання.

1. Ідентифікація змінних та специфікація моделі.

Ідентифікуємо змінні економетричної моделі:

X – вектор доходів населення (факторна, незалежна, екзогенна змінна);

Y – вектор роздрібного товарообігу (результативна, залежна, ендогенна змінна);

u – вектор залишків (стохастична складова).

Загальний вигляд моделі:

$$Y = f\left(X, u\right).$$

За допомогою функції СЧЕТ (значення 1; значення2; …), яка повертає кількість чисел у списку аргументів) визначимо обсяг вихідної сукупності. Для цього у комірку А14 записуємо ***n***, а у комірку С14 функцію СЧЕТ () (рис. 1.1), отримаємо $n = 10$.

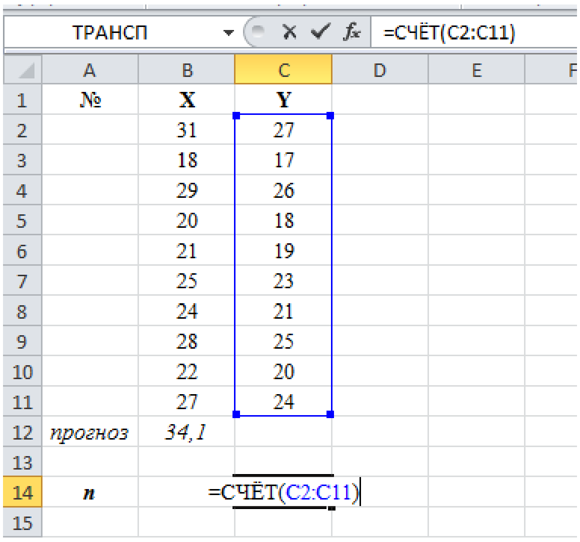


Рисунок 1.1 – Обчислення обсяг вихідної сукупності за допомогою функції СЧЕТ ()

Застосувавши команду ***Сортировка*** з меню ***Данные***, або піктограму 

до виділених стовпчиків X та Y обсягом 10 одиниць (рис. 1.2), розташуємо вихідні дані у порядку зростання значень незалежної змінної X (рис. 1.3, табл. 1.2).

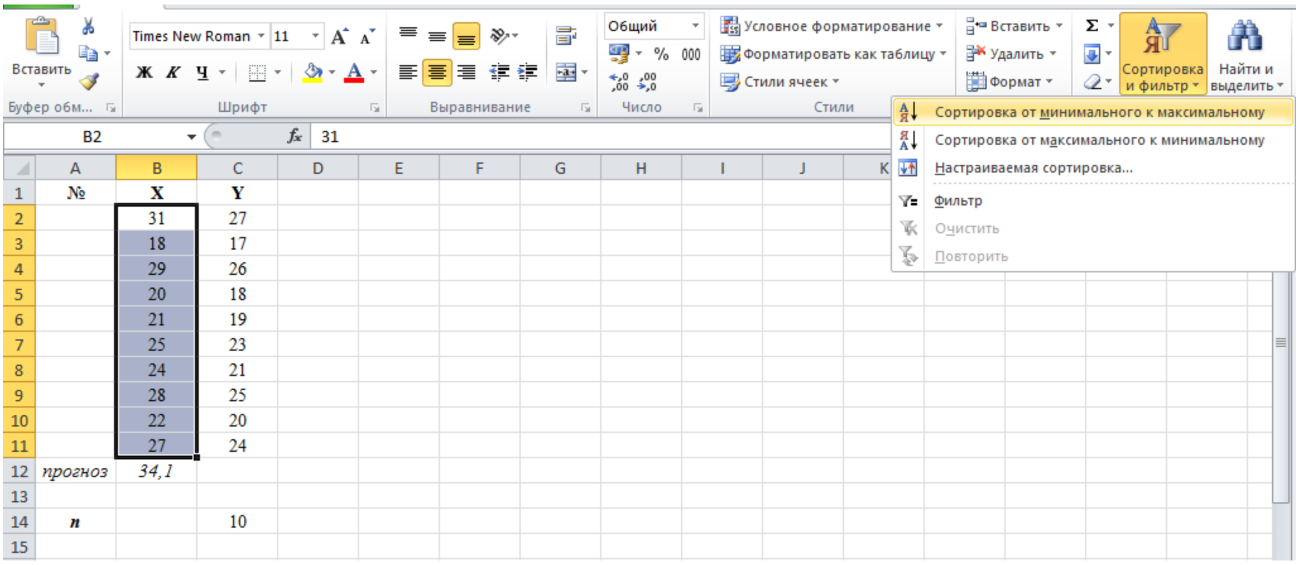
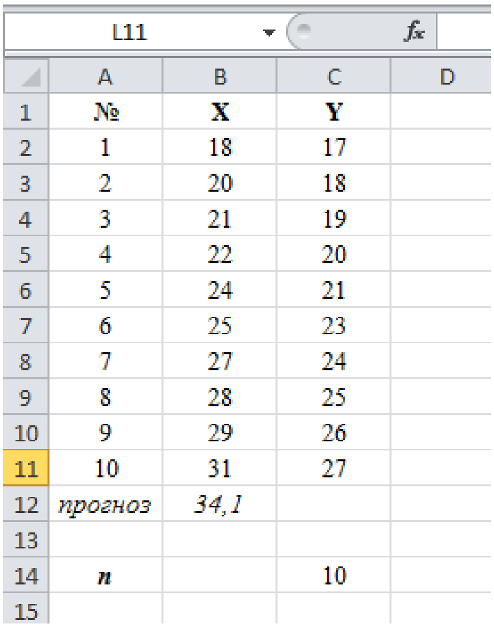


Рисунок 1.2 – Сортування вихідних даних у порядку зростання значень незалежної змінної *X*



	A	B	C	D
1	№	X	Y	
2	1	18	17	
3	2	20	18	
4	3	21	19	
5	4	22	20	
6	5	24	21	
7	6	25	23	
8	7	27	24	
9	8	28	25	
10	9	29	26	
11	10	31	27	
12	прогноз	34,1		
13				
14	n		10	
15				

Рисунок 1.3 – Відсортовані вихідні дані

Таблиця 1.2 – Відсортовані дані

№	Доходи населення, ум.од.	Роздрібний товарообіг, ум.од.
	X	*Y*
1	18	17
2	20	18
3	21	19
4	22	20
5	24	21
6	25	23
7	27	24
8	28	25

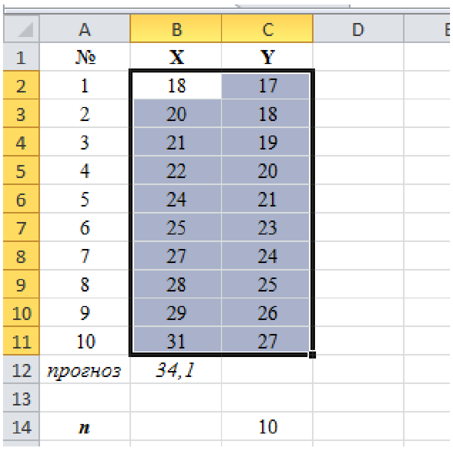
9	29	26
10	31	27
Прогнозне значення	$x_{пр} = 34{,}1$	

Виберемо специфікацію моделі, попередньо побудувавши для відсортованих даних діаграму розсіювання (кореляційне поле) залежності *Y* від *X* за допомогою ***Мастер диаграмм*** табличного редактора Microsoft Excel.

Розглянемо порядок побудови графіка:

1) відмічаємо необхідний для побудови графіка діапазон числових даних залежної змінної *Y*: B2: C11 (рис. 1.4).

Якщо потрібно виділити декілька несуміжних діапазонів даних, то тримаючи натиснутою клавішу Ctrl , відмічаємо кожний із них. При переході до нового діапазону ліву клавішу миші відпускаємо.



	A	B	C	D
1	№	X	Y	
2	1	18	17	
3	2	20	18	
4	3	21	19	
5	4	22	20	
6	5	24	21	
7	6	25	23	
8	7	27	24	
9	8	28	25	
10	9	29	26	
11	10	31	27	
12	*прогноз*	*34,1*		
13				
14	*n*		10	

Рисунок 1.4 – Вікно виділених даних залежної змінної *Y*

2) на панелі інструментів наводимо курсор на вкладку ***Вставка*** ⇒ ***Точечная*** (або ***График***) (рис. 1.5) і натискуємо ліву клавішу мишки. З’являється графік (рис. 1.6).

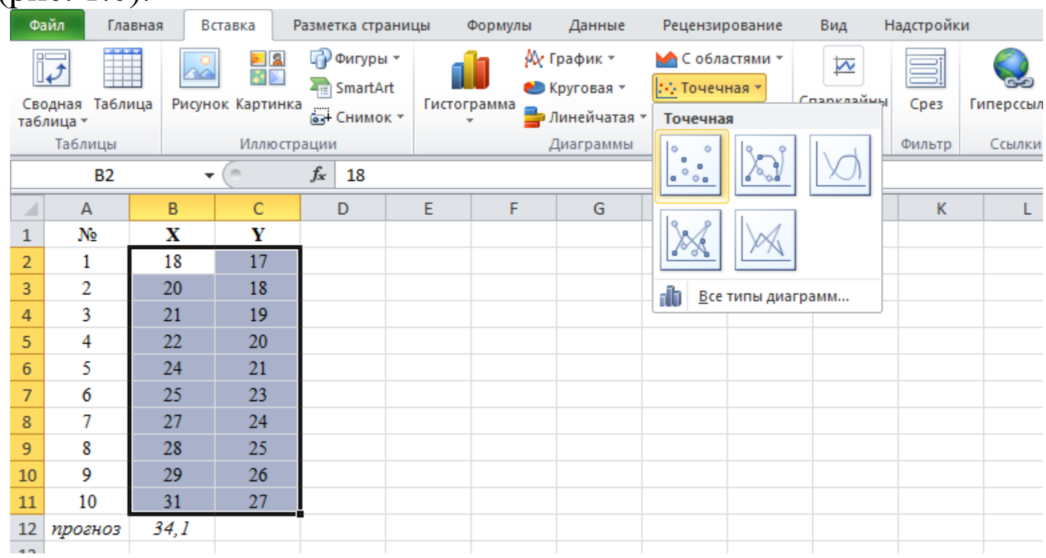


Рисунок 1.5 – Вікно типів діаграм та вихідних даних для діаграми ***Точечная***

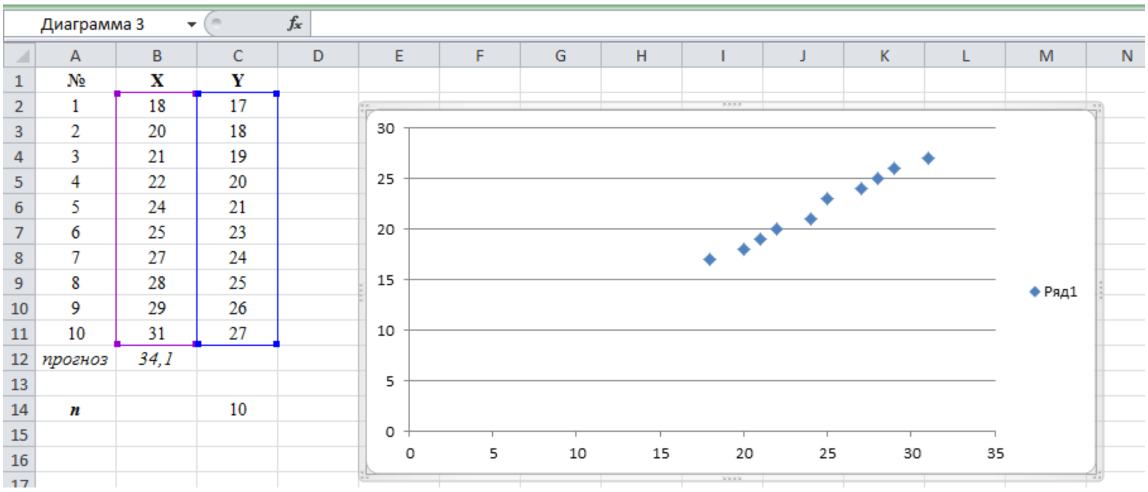


Рисунок 1.6 – Перший графік

3) виділяємо графік і переходимо на вкладку ***Работа с диаграммами*** ⇒ ***Макет*** ⇒ ***Названия осей***, для горизонтальної осі пишемо назву *Значення фактора X* і для вертикальної – *Значення показника Y*, а назву діаграми *Діаграма розсіювання* за допомогою ***Работа с диаграммами*** ⇒ ***Макет*** ⇒ ***Название диаграммы*** (рис. 1.7). Для зміни назви ряду обрати ***Выбрать данные*** ⇒ ***Изменить***, потім уведемо ім’я ряду *Yf* в поле ***Имя*** (рис. 1.8). Діаграму розміщуємо на окремому листі через натискання правої кнопки миші і обрання ***Переместить диаграмму*** ⇒ ***Лист 2***, який можна назвати інакше, наприклад, діаграма.

Для додавання одного чи декількох рядів даних натиснемо праву кнопку миші на точки графіка, обрати ***Выбрать данные*** ⇒ ***Добавить***, потім уведемо ім’я ряду в поле **Имя** натиснемо у полі ***Значения Y*** і введемо діапазон даних.

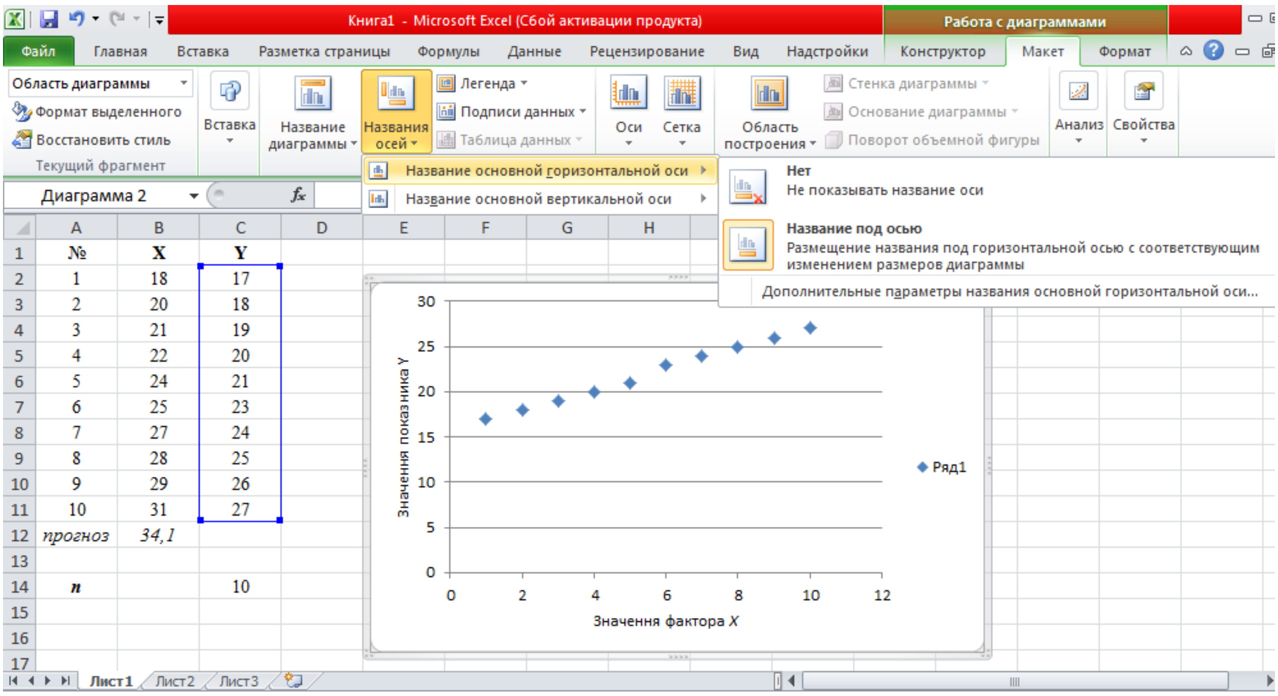


Рисунок 1.7 – Позначення назв осей

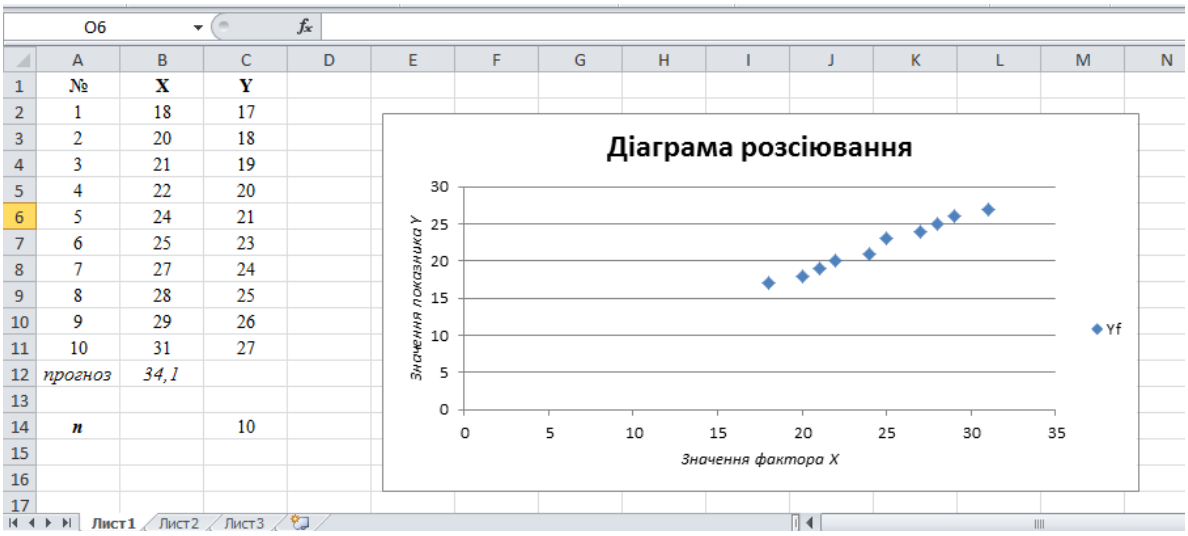


Рисунок 1.8 – Графік на робочому аркуші

З рис. 1.8 бачимо, що зі збільшенням значень незалежної ознаки X залежна змінна Y має тенденцію до збільшення. Візуальний аналіз графіка дозволяє зробити припущення, що між ознаками X та Y існує лінійна залежність.

Розглянемо специфікацію економетричної моделі $Y = f(X, u)$ у вигляді лінійної функції:

$$Y = a_0 + a_1 X + u .$$

Рівняння регресії:

$$\hat{Y} = \hat{a}_0 + \hat{a}_1 X ,$$

де a_0, a_1 – невідомі параметри моделі;

$\hat{a}_0, \hat{a}_1$ – їх оцінки;

$\hat{Y}$ – теоретичне (регресійне) значення результативної змінної;

$u = Y - \hat{Y}$ – вектор залишків (стохастична складова).

2. Оцінювання параметрів моделі методом МНК.

Для визначення параметрів рівняння регресії найчастіше застосовується класичний *метод найменших квадратів* (МНК) основоположниками якого є К. Гаусс та П. Лаплас. МНК дозволяє підібрати певну неперервну аналітичну функцію для апроксимації (наближення) дискретного набору вихідних даних. Задача зводиться до знаходження такої лінії (прямої, кривої, ламаної тощо), яка мінімізує суму квадратів відхилень між фактичними y та отриманими розрахунковим шляхом за рівнянням регресії теоретичними значеннями $\hat{y}$:

$$F = \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \to \min .$$

Оцінимо параметри $\hat{a}_0, \hat{a}_1$ лінійної моделі $\hat{Y} = \hat{a}_0 + \hat{a}_1 X$ класичним методом найменших квадратів (МНК), попередньо висунувши гіпотезу, що всі передумови (умови Гаусса-Маркова) для його застосування дотримані.

Для мінімізації функції

$$F\left(a_0, a_1\right)=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2=\sum_{i=1}^{n}\left(y_i-\left(a_0+a_1 x_i\right)\right)^2$$

визначимо перші частинні похідні функції F по a_0 та a_1, прирівняємо їх до нуля та отримаємо систему нормальних рівнянь:

$$\begin{cases} n\hat{a}_0+\hat{a}_1\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i, \\ \hat{a}_0\sum_{i=1}^{n}x_i+\hat{a}_1\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}x_i y_i. \end{cases}$$

Використовуючи вихідні дані з таблиці 1.2, зробимо необхідні обчислення. Для цього в Excel треба записати необхідні формули у вказані комірки (табл. 1.3).

Таблиця 1.3 – Розрахункові формули

комірка	*формула*	*коментар*
D2	=B2^2	аналогічно для D3:D11
E2	=B2*C2	аналогічно для E3:E11
B13	=СУММ(B2:B11)	аналогічно для C13:E13

Зробити дії, вказані в третьому стовпчику табл. 1.3, для інших комірок таблиці. Отже, будемо мати розрахункову таблицю (рис. 1.9, табл. 1.4).

Таблиця 1.4 – Розрахункова таблиця для МНК

№	X	Y	X^2	XY
1	18	17	324	306
2	20	18	400	360
3	21	19	441	399
4	22	20	484	440
5	24	21	576	504
6	25	23	625	575
7	27	24	729	648
8	28	25	784	700
9	29	26	841	754
10	31	27	961	837
Прогнозне значення	$x_{пр}=34{,}1$			
сума	**245**	**220**	**6165**	**5523**
n		**10**		

	A	B	C	D	E
1	№	X	Y	X^2	XY
2	1	18	17	324	306
3	2	20	18	400	360
4	3	21	19	441	399
5	4	22	20	484	440
6	5	24	21	576	504
7	6	25	23	625	575
8	7	27	24	729	648
9	8	28	25	784	700
10	9	29	26	841	754
11	10	31	27	961	837
12	*прогноз*	*34,1*			
13	***сума***	245	220	6165	5523
14	***n***		10		

Рисунок 1.9 – Розрахункова таблиця для МНК

Потім підставимо в систему знайдені значення сум і отримаємо систему нормальних рівнянь:

$$\begin{cases} 10\hat{a}_0 + 245\hat{a}_1 = 220, \\ 245\hat{a}_0 + 6165\hat{a}_1 = 5523. \end{cases}$$

Розв’язати цю систему рівнянь відносно невідомих оцінок параметрів $\hat{a}_0, \hat{a}_1$ можна будь-яким з відомих математичних методів або за допомогою ***Поиск решения*** табличного редактора Microsoft Excel.

Ми застосуємо ***Поиск решения***. Для цього комірки в G15:J15 запишемо відповідно:

G15 ^a_0
H15 ^a_1
I15 лев часть
J15 пр часть

Для подальшого застосування оператора ***Поиск решения*** запишемо необхідні функції згідно таблиці 1.5.

Таблиця 1.5 – Підготовка комірок до застосування оператора ***Поиск решения***

комірка	*формула*
I16	=C14*G16+H16*B13
J16	=C13
I17	=B13*G16+D13*H16
J17	=E13

Далі ставимо курсор в комірку I16 і обираємо ***Данные⇒ Поиск решения***. Заповнюємо вікно аналогічно рис. 1.10, натискаємо ***Найти решение*** і отримаємо розв’язок (рис. 1.11).

Таким чином, будемо мати:

$$\hat{a}_0 = 1{,}94770, \qquad \hat{a}_1 = 0{,}81846,$$

а економетрична модель роздрібного товарообігу матиме вигляд:

$$\hat{Y} = 1{,}94770 + 0{,}81846X.$$

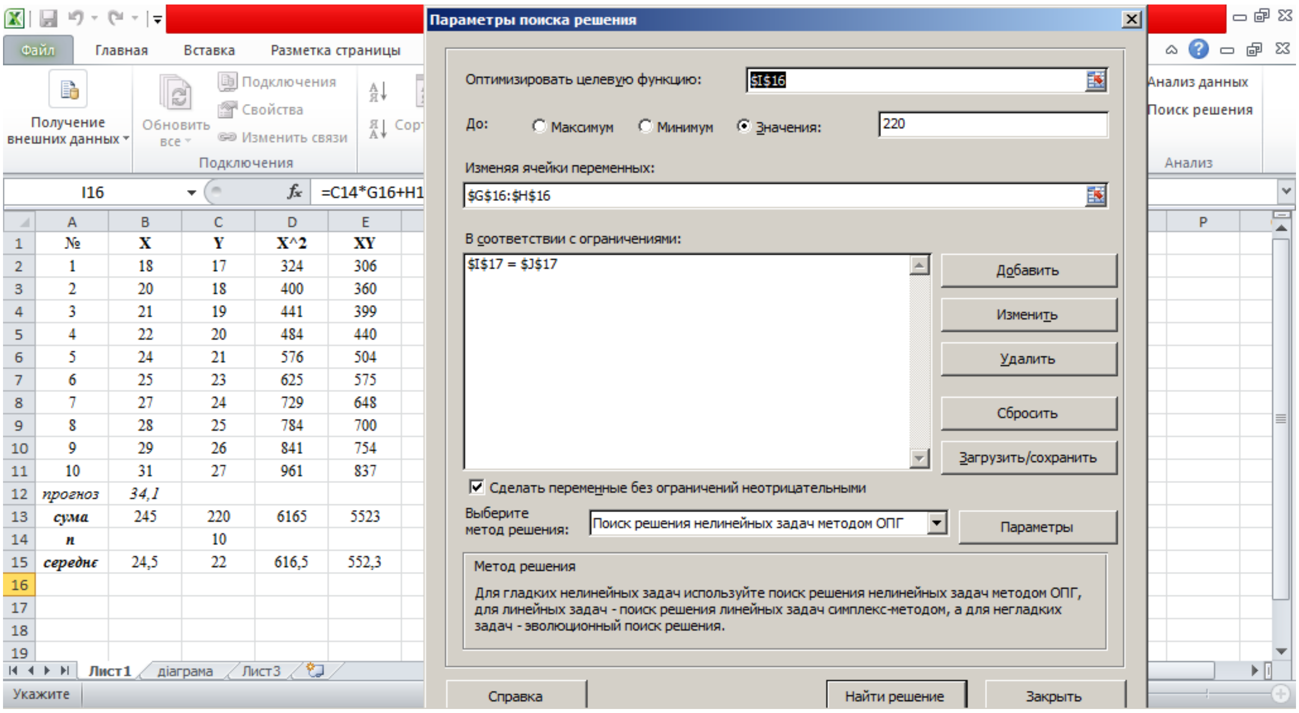
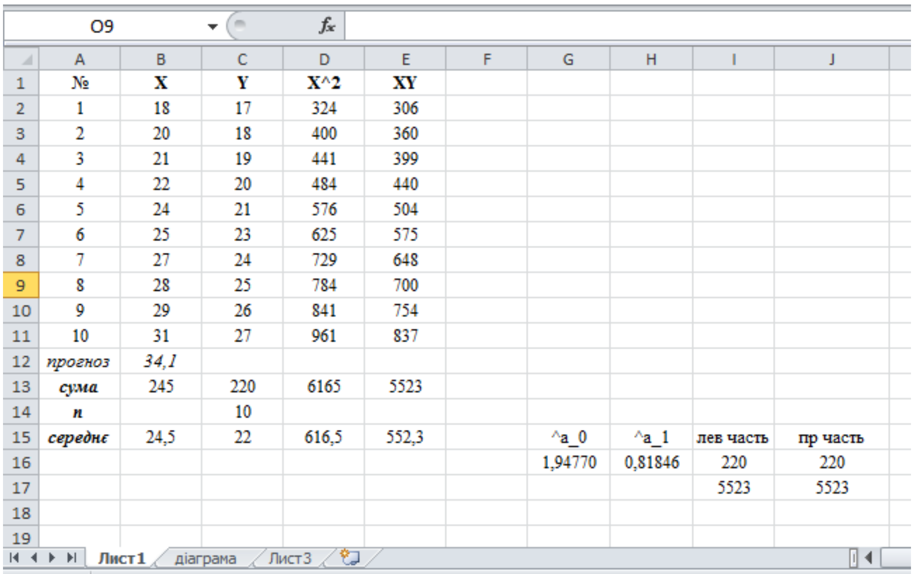


Рисунок 1.10 – Вікно оператора ***Поиск решения***



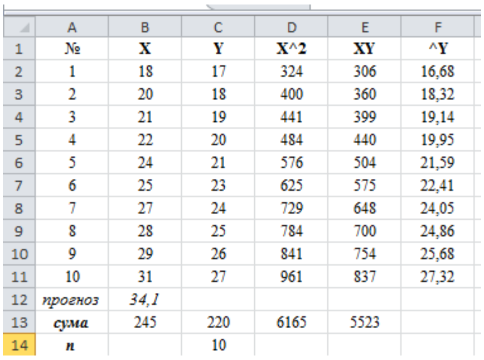
	A	B	C	D	E	F	G	H	I	J
1	№	X	Y	X^2	XY					
2	1	18	17	324	306					
3	2	20	18	400	360					
4	3	21	19	441	399					
5	4	22	20	484	440					
6	5	24	21	576	504					
7	6	25	23	625	575					
8	7	27	24	729	648					
9	8	28	25	784	700					
10	9	29	26	841	754					
11	10	31	27	961	837					
12	*прогноз*	*34,1*								
13	*сума*	245	220	6165	5523					
14	*n*		10							
15	*середнє*	24,5	22	616,5	552,3		^a_0	^a_1	лев часть	пр часть
16							1,94770	0,81846	220	220
17									5523	5523
18										
19										

Рисунок 1.11 – Результат застосування оператора ***Поиск решения***

Оскільки вільний член моделі $\hat{a}_0 = 1{,}94770 \neq 0$, то рівень роздрібного товарообігу не є точно пропорційним до рівня доходів населення. Заповнимо відповідний стовпчик таблиці теоретичними регресійними значеннями $\hat{Y}$, підставивши в одержане рівняння моделі послідовно всі фактичні значення X (рис. 1.12, табл. 1.6):

$$\hat{Y} = 1{,}94770 + 0{,}81846 \cdot 18 \text{ (=\$G\$16+\$H\$16*B2)},$$

$$\hat{Y} = 1{,}94770 + 0{,}81846 \cdot 20 \text{ (=\$G\$16+\$H\$16*B3) і т.д.}$$



	A	B	C	D	E	F
1	№	X	Y	X^2	XY	^Y
2	1	18	17	324	306	16,68
3	2	20	18	400	360	18,32
4	3	21	19	441	399	19,14
5	4	22	20	484	440	19,95
6	5	24	21	576	504	21,59
7	6	25	23	625	575	22,41
8	7	27	24	729	648	24,05
9	8	28	25	784	700	24,86
10	9	29	26	841	754	25,68
11	10	31	27	961	837	27,32
12	*прогноз*	*34,1*				
13	*сума*	245	220	6165	5523	
14	*n*		10			

Рисунок 1.12 – Теоретичні регресійні значення $\hat{Y}$

Таблиця 1.6 – Розрахункова таблиця для МНК

№	X	Y	$\hat{Y}$
1	18	17	16,68
2	20	18	18,32
3	21	19	19,14
4	22	20	19,95
5	24	21	21,59
6	25	23	22,41
7	27	24	24,05
8	28	25	24,86
9	29	26	25,68
10	31	27	27,32

Крім цього, теоретичні значення автоматично можна отримати в матричному виді, застосувавши функцію ПРЕДСКАЗ(X (число); известные значения Y (массив); известные значения X (массив)).

3. Побудова базової таблиці дисперсійного аналізу (ANOVA − таблиці) та визначення дисперсій.

Для подальших розрахунків необхідно визначити середні значення, дисперсії, середньоквадратичні відхилення змінних *X* і *Y*, використовуючи статистичні функції, та виконаємо послідовно необхідні розрахунки (табл. 1.7) та заповнимо інші стовпчики таблиці (рис. 1.13, табл. 1.8).

Таблиця 1.7 – Розрахункові формули

позначення	*комірка*	*формула*	*коментар*
середні значення			
$\overline{X}$	B15	=СРЗНАЧ(B2:B11)	аналогічно для C15 ($\overline{Y}$)
дисперсії (варіації)			
σ_X^2	B16	=ДИСП.Г(B2:B11)	аналогічно для C16 (σ_Y^2)
середньоквадратичні відхилення			

σ_X	B17	=СТАНДОТКЛОНП(B2:B11)	аналогічно для C17 (σ_Y)
проміжні обчислення			
$X-\overline{X}$	G2	=B2-B15	аналогічно для G3:G11
$Y-\overline{Y}$	H2	=C2-C15	аналогічно для H3:H11
$\left(X-\overline{X}\right)^2$	I2	=G2^2	аналогічно для I3:I11
$\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)$	J2	=G2*H2	аналогічно для J3:J11
u^2	K2	=(C2-F2)^2	аналогічно для K3:K11
$\left(Y-\overline{Y}\right)^2$	L2	=H2^2	аналогічно для L3:L11
$\left(\hat{Y}-\overline{Y}\right)^2$	M2	=(F2-C15)^2	аналогічно для M3:M11
сума	I13	=СУММ(I2:I11)	аналогічно для J13:M13

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	№	X	Y	X^2	XY	^Y	X-_X	Y-_Y	(X-_X)^2	(X-_X)(Y-_Y)	u^2	(Y-_Y)^2	(^Y-_Y)^2
2	1	18	17	324	306	16,68	-6,5	-5	42,25	32,5	0,1024	25	28,302389
3	2	20	18	400	360	18,32	-4,5	-4	20,25	18	0,1004	16	13,565050
4	3	21	19	441	399	19,14	-3,5	-3	12,25	10,5	0,0183	9	8,206018
5	4	22	20	484	440	19,95	-2,5	-2	6,25	5	0,0021	4	4,186744
6	5	24	21	576	504	21,59	-0,5	-1	0,25	0,5	0,3490	1	0,167470
7	6	25	23	625	575	22,41	0,5	1	0,25	0,5	0,3490	1	0,167470
8	7	27	24	729	648	24,05	2,5	2	6,25	5	0,0021	4	4,186744
9	8	28	25	784	700	24,86	3,5	3	12,25	10,5	0,0183	9	8,206019
10	9	29	26	841	754	25,68	4,5	4	20,25	18	0,1004	16	13,565051
11	10	31	27	961	837	27,32	6,5	5	42,25	32,5	0,1024	25	28,302391
12	*прогноз*	*34,1*											
13	*сума*	245	220	6165	5523	220			162,5	133	1,144615	110	108,855345
14	*n*		10										
15	*середнє*	24,5	22				^a_0	^a_1	лев часть	пр часть			
16	*sig^2*	16,25	11				1,94770	0,81846	220	220			
17	*sig*	4,031129	3,316625						5523	5523			
18													
19													

Лист1 | діаграма | Лист3

Рисунок 1.13 – Розрахункова таблиця в Excel

Таблиця 1.8 – Базова таблиця

№	X	Y	$\hat{Y}$	$X-\overline{X}$	$Y-\overline{Y}$	$\left(X-\overline{X}\right)^2$	$\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)$	u^2	$\left(Y-\overline{Y}\right)^2$	$\left(\hat{Y}-\overline{Y}\right)^2$
1	18	17	16,68	-6,5	-5	42,25	32,5	0,1024	25	28,302389
2	20	18	18,32	-4,5	-4	20,25	18	0,1004	16	13,565050
3	21	19	19,14	-3,5	-3	12,25	10,5	0,0183	9	8,206018
4	22	20	19,95	-2,5	-2	6,25	5	0,0021	4	4,186744
5	24	21	21,59	-0,5	-1	0,25	0,5	0,3490	1	0,167470
6	25	23	22,41	0,5	1	0,25	0,5	0,3490	1	0,167470

7	27	24	24,05	2,5	2	6,25	5	0,0021	4	4,186744
8	28	25	24,86	3,5	3	12,25	10,5	0,0183	9	8,206019
9	29	26	25,68	4,5	4	20,25	18	0,1004	16	13,565051
10	31	27	27,32	6,5	5	42,25	32,5	0,1024	25	28,302391
прогн	*34,1*									
сума	**245**	**220**	**220**			**162,5**	**133**	**1,144615**	**110**	**108,855345**
n		**10**								
середнє	**24,5**	**22**								
sig^2	**16,25**	**11**								
sig	**4,031129**	**3,316625**								

Використовуючи результати розрахунків таблиці 1.8 або за допомогою статистичних функцій редактора Excel, визначимо основні елементи базової таблиці дисперсійного аналізу (ANOVA – таблиці, табл. 1.9).

Таблиця 1.9 – Основні елементи базової таблиці дисперсійного аналізу

позначення	*формула*	*статистична функція*
SSR	$\sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2$	=КВАДРОТКЛ($\hat{Y}$)
SSE	$\sum_{i=1}^{n}u_i^2 = \sum_{i=1}^{n}\left(\hat{y}_i - y\right)^2$	=СУММКВРАЗН($Y;\hat{Y}$)
MSE	$\frac{SSE}{n-2}$	$\frac{SSE}{n-2}$
SST	$\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$	=КВАДРОТКЛ(Y)
MST	$\frac{SST}{n-1}$	$\frac{SST}{n-1}$

Перевіримо розрахунки, відповідно до теореми додавання дисперсій:

$$SSR + SSE = SST\,,$$

$$108{,}855345 + 1{,}14461538 = 110\,.$$

Отже, отримаємо ANOVA − таблицю (табл. 1.10). У даній таблиці $n = 10$ − обсяг сукупності, $m = 2$ − кількість параметрів моделі. Величини $k_1 = m - 1 = 1$, $k_2 = n - m = 8$, $n - 1 = 9$ називаються *ступенями свободи* (*вільності*) – це числа, що показують, скільки незалежних елементів інформації із змінних $y_i,\ i = \overline{1,\,n}$ потрібно для розрахунку відповідної суми квадратів.

Використовуючи таблицю 1.9, визначимо та занесемо до табл. 1.10:

– дисперсію регресії

$$\hat{\sigma}_r^2 = MSR = SSR = 108{,}855345\,;$$

– незміщену оцінку дисперсії залишків

$$s^2 = \hat{\sigma}_u^2 = MSE = \frac{SSE}{n-2} = \frac{1{,}14461538}{8} = 0{,}143077;$$

– стандартизовану похибку залишків

$$s = \hat{\sigma}_u = \sqrt{\sigma_u^2} = \sqrt{0{,}143077} = 0{,}378255;$$

– дисперсію залежної змінної (загальну дисперсія)

$$\hat{\sigma}_y^2 = MST = \frac{SST}{n-1} = \frac{110}{9} = 12{,}222222.$$

Чим менша стандартна помилка залишків $\hat{\sigma}_u$, тим краще підібрана функція.

Таблиця 1.10 – ANOVA – таблиця

Джерело варіації	Ступені свободи	Сума квадратів	Дисперсії (середні квадрати)
Регресії	$k_1 = m - 1 = 1$	$SSR = 108{,}855345$	$\hat{\sigma}_r^2 = MSR = 108{,}855345$
Залишків	$k_2 = n - m = 8$	$SSE = 1{,}14461538$	$\hat{\sigma}_u^2 = MSE = 0{,}143077$
Загальної змінної	$n - 1 = 9$	$SST = 110$	$\hat{\sigma}_y^2 = MST = 12{,}222222$

4. *Визначення коефіцієнтів детермінації, кореляції та еластичності.*

Визначимо потрібні коефіцієнти, використовуючи результати розрахунків таблиць 1.9 і 1.10 або за допомогою статистичних функцій редактора Excel коефіцієнт детермінації. Заповнимо таблицю 1.11.

Таблиця 1.11 – Визначення коефіцієнтів детермінації, кореляції та еластичності

назва та позначення	*формула*	*статистична функція*	*значення*
коефіцієнт детермінації R^2	$\frac{SSR}{SST}$	=КВПИРСОН(Y;X)	0,9896
коефіцієнт кореляції Пірсона r_{xy}	$\frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 (x_i - \bar{x})^2}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}$	=КОРРЕЛ(X; Y) або =PEARSON(Y;X)	0,9948
коефіцієнт еластичності $E_{Y/X}$	$\hat{a}_1 \cdot \frac{\bar{X}}{\bar{Y}}$		0,9115

Коефіцієнт детермінації є певною мірою універсальною характеристикою ступеня щільності статистичного зв'язку. Його значення 0,9896 свідчить, що варіація обсягу роздрібного товарообігу на 98,96% визначається варіацією доходів населення.

Коефіцієнт кореляції характеризує щільність зв'язку, його значення 0,9948 свідчить про прямий достатньо сильний зв'язок між доходами населення та роздрібним товарообігом. Оскільки значення коефіцієнта кореляції близьке до одиниці, то це означає, що залежність між *Y* та *X* – практично лінійна.

Для парної лінійної регресії повинна справджуватися рівність:

$$R^2 = r_{xy}^2.$$

Дійсно, $0{,}9896 = 0{,}9948^2$, $0{,}9896 = 0{,}9896$.

Коефіцієнт еластичності характеризує відносний ефект впливу фактора *X* на результативний показник *Y*, тобто при збільшенні рівня доходів населення на 1% роздрібний товарообіг гранично зростає на 0,91%.

5. Перевірка статистичної значущості коефіцієнта детермінації, коефіцієнта кореляції та визначення надійного інтервалу коефіцієнта кореляції.

Перевірка статистичної значущості коефіцієнта детермінації. Розглянемо нульову гіпотезу $H_0 : d = R^2 = 0$ проти альтернативної гіпотези $H_A : d = R^2 > 0$.

Порівняємо фактичне значення коефіцієнта детермінації $d = R^2 = 0{,}9896$ із табличним для рівня значущості $\alpha = 0{,}05$ та чисел ступенів свободи $k_1 = m - 1 = 1$, $k_2 = n - m = 8$: $d_{\text{табл}}(0{,}05; 1; 8) = 0{,}399$ (додаток Б). Оскільки $d > d_{\text{табл}}$, то гіпотеза H_0 відхиляється і коефіцієнт детермінації є значущим.

Рівень значущості $\alpha = 0{,}05$ означає, що в 5% випадків ми можемо помилитися, а в 95% випадків наші висновки правильні.

Перевірка статистичної значущості коефіцієнта кореляції. Для цього перевіримо нульову гіпотезу $H_0 : r_{xy} = 0$ проти альтернативної гіпотези $H_A : r_{xy} \neq 0$. Порівняємо отримане значення $r_{xy} = 0{,}9948$ із табличним. Для $n = 10$: $r_{xy\ \text{табл}} = 0{,}63$ (додаток В). Оскільки умова $r_{xy} > r_{xy\ \text{табл}}$ виконується, то гіпотеза H_0 відхиляється на користь альтернативної, що свідчить про значущість коефіцієнта кореляції.

Оцінка значущості коефіцієнта кореляції за $t-$ критерієм Стьюдента. Для цього перевіримо нульову гіпотезу $H_0 : r_{xy} = 0$ проти альтернативної гіпотези $H_A : r_{xy} \neq 0$.

Фактичне значення $t-$ критерію визначається за формулою:

$$t_{\text{факт}} = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}} = 27{,}5829.$$

Для заданого рівня значущості $\alpha = 0,05$ та числа ступенів свободи $k = k_2 = n - m = 8$ визначимо табличне значення t − статистики: $t_{\text{табл}} = t(0,05;\, 8) = 2,306$ (додаток Д). Табличне значення можна також визначити за допомогою функції =СТЬЮДЕНТ.ОБР.2Х(0,05;8)=2,306004.

Оскільки $|t_{\text{факт}}| > t_{\text{табл}}$, то гіпотеза H_0 відхиляється на користь альтернативної та коефіцієнт кореляції вважається статично значущим.

Визначення надійних інтервалів для коефіцієнта кореляції за $t-$розподілом Стьюдента при рівні значущості $\alpha = 0,05$. Надійні інтервали коефіцієнта кореляції визначаються за формулою:

$$r - \Delta r \le r \le r + \Delta r\,,$$

$$r - tS_r \le r \le r + tS_r\,.$$

Для значення $r_{xy} = 0,9948$ при $t_{\text{табл}} = t(0,05;\, 8) = 2,306$, маємо:

$$S_r = \frac{1-r}{\sqrt{n}} = 0,00165\,.$$

Тоді,

$$0,9948 - 2,306 \cdot 0,00165 \le r \le 0,9948 + 2,306 \cdot 0,00165\,,$$

$$0,990980 \le r \le 0,998588\,.$$

6. Визначення стандартних помилок оцінок параметрів моделі та їх оцінка.

Визначимо стандартні помилки оцінок параметрів, використовуючи дисперсію залишків. Використовуючи результати розрахунків із таблиці 1.8, дістанемо:

$$S_{\hat{a}_0} = \hat{\sigma}_u \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}} = 0,736758, \qquad S_{\hat{a}_1} = \hat{\sigma}_u \frac{1}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = 0,029673.$$

Визначимо матрицю коваріацій оцінок параметрів моделі, використовуючи дисперсію залишків і матрицю помилок. Для цього розглянемо матрицю незалежних змінних *X*:

$$X \quad \begin{Vmatrix} 1 & 18 \\ 1 & 20 \\ 1 & 21 \\ 1 & 22 \\ 1 & 24 \\ 1 & 25 \\ 1 & 27 \\ 1 & 28 \\ 1 & 29 \\ 1 & 31 \end{Vmatrix}$$

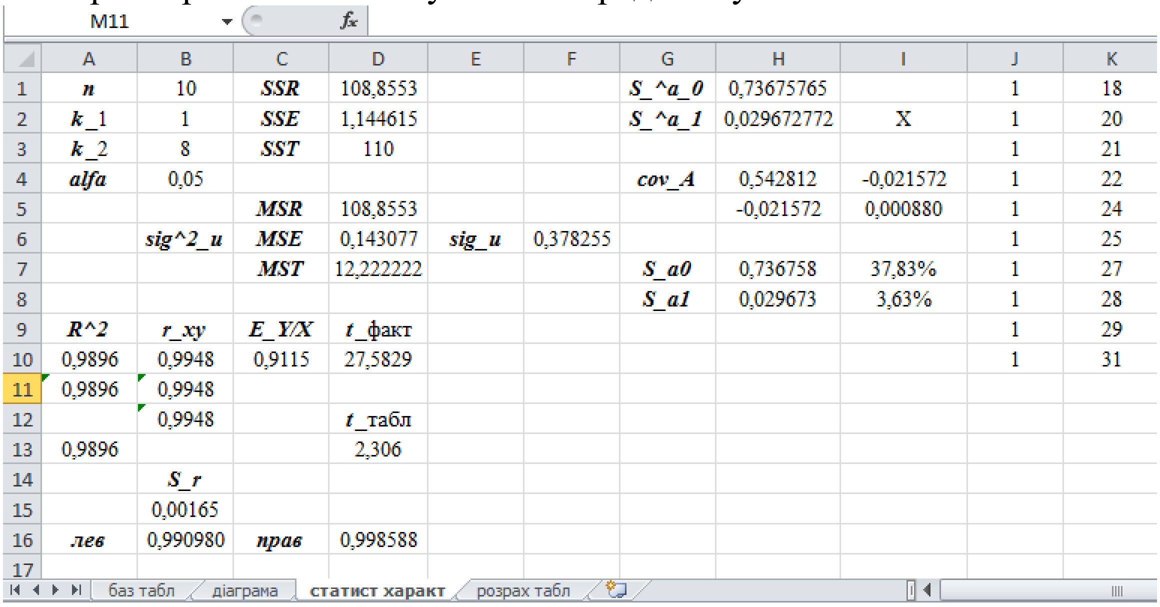
Тоді, ввівши в комірку H4 функцію

$$=D6*МОБР(МУМНОЖ(L1:U2;ТРАНСП(L1:U2))),$$

виділивши H4:I5, натиснувши F2, а потім Ctrl+Shift+Enter, отримаємо матрицю (рис. 1.14):

$$\operatorname{cov}\left(\hat{A}\right)=\hat{\sigma}_u^2\left(X^T X\right)^{-1}=\begin{Vmatrix} 0{,}542812 & -0{,}021572 \\ -0{,}021572 & 0{,}000880 \end{Vmatrix}$$

Діагональні елементи цієї матриці характеризують дисперсії оцінок параметрів моделі: $\hat{\sigma}_{\hat{a}_0}^2=0{,}542812$, $\hat{\sigma}_{\hat{a}_1}^2=0{,}000880$. Інші елементи даної матриці визначають рівень коваріації між оцінками параметрів моделі. Знак «мінус» перед оцінками коваріацій $\hat{\sigma}_{\hat{a}_j\hat{a}_k}$ вказує на те, що зі збільшенням однієї оцінки параметрів інша зменшується в середньому та навпаки.



	A	B	C	D	E	F	G	H	I	J	K
1	n	10	SSR	108,8553			S_^a_0	0,73675765		1	18
2	k_1	1	SSE	1,144615			S_^a_1	0,029672772	X	1	20
3	k_2	8	SST	110						1	21
4	alfa	0,05					cov_A	0,542812	-0,021572	1	22
5			MSR	108,8553				-0,021572	0,000880	1	24
6		sig^2_u	MSE	0,143077	sig_u	0,378255				1	25
7			MST	12,222222			S_a0	0,736758	37,83%	1	27
8							S_a1	0,029673	3,63%	1	28
9	R^2	r_xy	E_Y/X	t_факт						1	29
10	0,9896	0,9948	0,9115	27,5829						1	31
11	0,9896	0,9948									
12		0,9948		t_табл							
13	0,9896			2,306							
14		S_r									
15		0,00165									
16	лев	0,990980	прав	0,998588							
17											

Рисунок 1.14 – Матриця коваріації

Визначимо стандартні помилки оцінок параметрів, використовуючи матрицю коваріацій. Оскільки

$$\operatorname{cov}\left(\hat{A}\right)=\begin{Vmatrix} 0{,}542812 & -0{,}021572 \\ -0{,}021572 & 0{,}000880 \end{Vmatrix}$$

то відповідно дістанемо (рис. 1.14):

$$S_{\hat{a}_0}=\sqrt{\hat{\sigma}_{\hat{a}_0}^2}=0{,}736758, \qquad S_{\hat{a}_1}=\sqrt{\hat{\sigma}_{\hat{a}_1}^2}=0{,}029673.$$

Стандартні похибки характеризують середні лінійні коливання оцінок параметрів моделі навколо свого математичного сподівання. Чим менші ці похибки, тим стійкіші оцінки параметрів моделі. Остаточні висновки стосовно стійкості оцінок можна зробити лише тоді, коли порівняти їх з абсолютними значеннями оцінок параметрів моделі.

Вибіркова оцінка параметрів $\hat{A}$ називається *незміщеною*, якщо вона задовольняє рівність $M\left(\hat{A}\right)=A$. *Незміщеність* – це мінімальна вимога, яка ставиться до оцінок параметра $\hat{A}$. Про наявність зміщеності оцінки можна стверджувати тоді, коли її стандартна похибка перевищує 10% від абсолютного значення оцінки.

Порівняємо стандартні помилки оцінок параметрів моделі зі значеннями цих оцінок:

$$\frac{S_{\hat{a}_0}}{|\hat{a}_0|}\cdot 100\% = 37{,}83\%, \qquad \frac{S_{\hat{a}_1}}{|\hat{a}_1|}\cdot 100\% = 3{,}63\%.$$

У результаті визначимо (рис. 1.14), що стандартна помилка оцінки параметра $\hat{a}_0$ становить 37,83% абсолютного значення цієї оцінки (1,94770), а це означає, що даний параметр може мати зміщення, що зумовлюється невеликою сукупністю спостережень ($n=10$). Стандартна помилка оцінки параметра $\hat{a}_1$ становить 3,63% абсолютного значення цієї оцінки (0,81846), що свідчить про незміщеність даної оцінки параметра моделі.

7. *Перевірка статистичної значущості оцінок параметрів моделі та визначення їх надійних інтервалів.*

Перевіримо статистичну значущість оцінок параметрів моделі за t–статистикою Стьюдента.

Розглянемо гіпотези:

$H_0 : \hat{a}_j = 0\ (j=0,1)$ – оцінка параметра незначуща;

$H_A : \hat{a}_j \neq 0\ (j=0,1)$ – оцінка параметра значуща.

Визначимо фактичні значення t–статистики (рис. 1.15):

$$t_0 = \frac{|\hat{a}_0|}{S_{\hat{a}_0}} = 2{,}644, \qquad t_1 = \frac{|\hat{a}_1|}{S_{\hat{a}_1}} = 27{,}583.$$

G10 | f_x t_0

	A	B	C	D	E	F	G	H	I	J	K
1	***n***	10	***SSR***	108,8553			***S_^a_0***	0,73675765		1	18
2	***k_1***	1	***SSE***	1,144615			***S_^a_1***	0,029672772	X	1	20
3	***k_2***	8	***SST***	110						1	21
4	***alfa***	0,05					***cov_A***	0,542812	-0,021572	1	22
5			***MSR***	108,8553				-0,021572	0,000880	1	24
6		***sig^2_u***	***MSE***	0,143077	***sig_u***	0,378255				1	25
7			***MST***	12,222222			***S_a0***	0,736758	37,83%	1	27
8							***S_a1***	0,029673	3,63%	1	28
9	***R^2***	***r_xy***	***E_Y/X***	***t*_факт**						1	29
10	0,9896	0,9948	0,9115	27,5829			***t_0***	2,644		1	31
11	0,9896	0,9948					***t_1***	27,583			
12		0,9948		***t*_табл**							
13	0,9896			2,306							
14		***S_r***									
15		0,00165									
16	***лев***	0,990980	***прав***	0,998588							
17											

баз табл / діаграма / **статист характ** / розрах табл

Рисунок 1.15 – Фактичні значення t –статистики

Порівняємо знайдені значення з табличним значенням t– статистики:

$$t_{табл} = \text{СТЬЮДЕНТ.ОБР.2Х}(0{,}05;8) = 2{,}306.$$

Оскільки $t_0 > t_{табл}$ і $t_1 > t_{табл}$, то нульову гіпотезу для обох параметрів відхиляємо на користь альтернативної і параметри моделі статично значущі, тобто мають значний вплив на залежну змінну Y.

Визначимо інтервали надійності для оцінок параметрів моделі за t–розподілом при рівні значущості $\alpha = 0{,}05$. Оскільки $t_{\text{табл}} = 2{,}306$, то відповідно:

– для параметра $\hat{a}_0 = 1{,}94770$ маємо:

$$\hat{a}_0 - \Delta\hat{a}_0 \le \hat{a}_0 \le \hat{a}_0 + \Delta\hat{a}_0,$$

$$\hat{a}_0 - tS_{\hat{a}_0} \le \hat{a}_0 \le \hat{a}_0 + tS_{\hat{a}_0},$$

$$1{,}94770 - 2{,}306 \cdot 0{,}736758 \le \hat{a}_0 \le 1{,}94770 + 2{,}306 \cdot 0{,}736758,$$

$$0{,}2487299 \le \hat{a}_0 \le 3{,}6466623;$$

– для параметра $\hat{a}_1 = 0{,}81846$ маємо:

$$\hat{a}_1 - \Delta\hat{a}_1 \le \hat{a}_1 \le \hat{a}_1 + \Delta\hat{a}_1,$$

$$\hat{a}_1 - tS_{\hat{a}_1} \le \hat{a}_1 \le \hat{a}_1 + tS_{\hat{a}_1},$$

$$0{,}81846 - 2{,}306 \cdot 0{,}029673 \le \hat{a}_1 \le 0{,}81846 + 2{,}306 \cdot 0{,}029673,$$

$$0{,}750036 \le \hat{a}_1 \le 0{,}886887.$$

8. Перевірка адекватності економетричної моделі.

Перевіримо за допомогою F−критерію Фішера адекватність економетричної моделі фактичним даним.

Розглянемо гіпотезу $H_0 : R^2 = 0$ проти альтернативної $H_A : R^2 > 0$. Це рівнозначне тому, що перевіряється значущість водночас усіх параметрів моделі (гіпотеза $H_0 : \hat{a}_0 = \hat{a}_1 = 0$ проти альтернативної $H_A : \hat{a}_0 \ne 0, \hat{a}_1 \ne 0$).

Використовуючи ANOVA − таблицю (табл. 1.10) і значення коефіцієнта детермінації (табл. 1.11), визначимо фактичне значення F−критерію:

$$F_{\text{факт}} = \frac{MSR}{MSE} = 760{,}8169.$$

Табличне значення для заданого рівня значущості $\alpha = 0{,}05$ (рівня помилки) або $p = 1 - \alpha = 0{,}95$ (довірчої ймовірності) та числа ступенів свободи $k_1 = m - 1 = 1$ і $k_2 = n - 2 = 8$ можна взяти з таблиці (додаток Е): $F_{\text{табл}} = F(0{,}05; 1; 8) = 5{,}32$, або, використавши функцію редактора Excel:

$$F_{\text{табл}} = \text{F.ОБР.ПХ}(\alpha; k_1; k_2) = 5{,}31766.$$

Оскільки $F_{\text{факт}} > F_{\text{табл}}$, то нульову гіпотезу відхиляємо і з заданою ймовірністю $p = 0{,}95$ економетричну модель вважаємо адекватною фактичним даним, тобто гіпотеза про значущість зв'язку між незалежною та залежною змінними підтверджується.

9. Визначення надійних інтервалів базисних значень.

Використовуючи раніше знайдені значення $\hat{\sigma}_u = 0{,}378255$ і $t_{\text{табл}} = 2{,}306$, визначимо надійні інтервали базисних значень $\hat{Y}$ за формулою:

$$\hat{y}_i \pm \Delta\hat{y}_i,$$

де $\hat{y}_i$ – теоретичні значення, обчислені за допомогою побудованої економетричної моделі;

$$\Delta\hat{y}_i = t\hat{\sigma}_{\hat{y}_i} = t\hat{\sigma}_u \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

і внесемо їх до таблиці 1.12.

Виконавши відповідні розрахунки, визначимо надійні інтервали базисних значень за формулами:

$$\hat{y}_{i,\ \min} = \hat{y}_i - \Delta\hat{y}_i,$$

$$\hat{y}_{i,\ \max} = \hat{y}_i + \Delta\hat{y}_i.$$

Таблиця 1.12 – Розрахункова таблиця

№	X	Y	^Y	delta^Y	^Y_min	^Y_max	abs((Y-^Y)/Y)*100
1	18	17	16,68	0,52335	16,15665	17,20336	1,88235
2	20	18	18,32	0,41339	17,90353	18,73032	1,76069
3	21	19	19,14	0,36529	18,77009	19,50068	0,71255
4	22	20	19,95	0,32457	19,62928	20,27842	0,23077
5	24	21	21,59	0,27795	21,31282	21,86872	2,81319
6	25	23	22,41	0,27795	22,13128	22,68718	2,56856
7	27	24	24,05	0,32457	23,72158	24,37072	0,19231
8	28	25	24,86	0,36529	24,49932	25,22991	0,54154
9	29	26	25,68	0,41339	25,26968	26,09647	1,21894
10	31	27	27,32	0,52335	26,79664	27,84335	1,18518
сума	**245**	**220**	**220**				**13,10607**

10. Визначення точкового та інтервальних прогнозів.

Визначимо значення точкового прогнозу роздрібного товарообігу $\hat{Y}_{\text{прогн}}$. Для визначення точкової оцінки прогнозу підставимо задане прогнозне значення $X_{\text{прогн}} = 34{,}1$ (табл. 1.1) у рівняння регресії $\hat{Y} = 1{,}94770 + 0{,}81846X$:

$$\hat{Y} = 1{,}94770 + 0{,}81846 \cdot 34{,}1 = 29{,}857229.$$

Значення $\hat{Y}_{\text{прогн}}$ можна інтерпретувати як точкову оцінку прогнозного значення математичного сподівання та індивідуального значення роздрібного товарообігу, коли відомі доходи населення $X_{\text{прогн}} = 34{,}1$.

Перевіримо знайдене прогнозне значення за допомогою відповідної статистичної функції:

$Y_{\text{прогн}}$=ПРЕДСКАЗ(34,1;*Y*;*X*)= 29,857231.

Визначимо прогнозний інтервал математичного сподівання $M\left(\hat{Y}_{\text{прогн}}\right)$. Обчислимо

– стандартну помилку прогнозного значення або стандартну похибку прогнозу математичного сподівання $M\left(\widehat{Y}_{\text{прогн}}\right)$, враховуючи, що стандартна помилка залишків $\widehat{\sigma}_u = 0,378255$:

$$\widehat{\sigma}_{\text{прогн}} = \widehat{\sigma}_u \sqrt{\frac{1}{n} + \frac{\left(x_{\text{прогн}} - \overline{x}\right)^2}{\sum\limits_{i=1}^{n}\left(x_i - \overline{x}\right)^2}} = 0,30895;$$

– прогнозний інтервал математичного сподівання $M\left(\widehat{Y}_{\text{прогн}}\right)$ при $t_{\text{табл}} = 2,306$:

$$\widehat{Y}_{\text{прогн}} - \Delta\widehat{Y}_{\text{прогн}} \le M\left(\widehat{Y}_{\text{прогн}}\right) \le \widehat{Y}_{\text{прогн}} + \Delta\widehat{Y}_{\text{прогн}};$$

$$\widehat{Y}_{\text{прогн}} - t\widehat{\sigma}_{\text{прогн}} \le M\left(\widehat{Y}_{\text{прогн}}\right) \le \widehat{Y}_{\text{прогн}} + t\widehat{\sigma}_{\text{прогн}};$$

$$29,857231 - 2,306 \cdot 0,30895 \le M\left(\widehat{Y}_{\text{прогн}}\right) \le 29,857231 + 2,306 \cdot 0,30895;$$

$$29,14478 \le M\left(\widehat{Y}_{\text{прогн}}\right) \le 30,56968.$$

Визначимо прогнозний інтервал індивідуального значення $\widehat{Y}_{\text{прогн}}$:

– дисперсія прогнозу індивідуального значення $\widehat{Y}_{\text{прогн}}$:

$$\widehat{\sigma}^2_{\text{прогн}(i)} = \widehat{\sigma}^2_u + \widehat{\sigma}^2_{\text{прогн}} = 0,143077 + 0,30895^2 = 0,238529;$$

– стандартна похибка прогнозу індивідуального значення $\widehat{Y}_{\text{прогн}}$:

$$\sigma_{\text{прогн}(i)} = \widehat{\sigma}^2_{\text{прогн}(i)} = \sqrt{0,238529} = 0,488394.$$

Тоді ***прогнозний інтервал індивідуального значення*** $\widehat{Y}_{\text{прогн}}$ при $t_{\text{табл}} = 2,306$ матиме вигляд:

$$\widehat{Y}_{\text{прогн}} - \Delta\widehat{Y}_{\text{прогн}(i)} \le \widehat{Y}_{\text{прогн}} \le \widehat{Y}_{\text{прогн}} + \Delta\widehat{Y}_{\text{прогн}(i)};$$

$$\widehat{Y}_{\text{прогн}} - t\widehat{\sigma}_{\text{прогн}(i)} \le \widehat{Y}_{\text{прогн}} \le \widehat{Y}_{\text{прогн}} + t\widehat{\sigma}_{\text{прогн}(i)};$$

$$29,857231 - 2,306 \cdot 0,488394 \le \widehat{Y}_{\text{прогн}} \le 29,857231 + 2,306 \cdot 0,488394;$$

$$28,7309914 \le \widehat{Y}_{\text{прогн}} \le 30,9834702.$$

Отже, з ймовірністю $p = 0,95$ $(\alpha = 0,05)$ прогноз математичного сподівання $M\left(\widehat{Y}_{\text{прогн}}\right)$ потрапляє в інтервал $\left[29,14478; 30,56968\right]$, а прогноз індивідуального значення – в інтервал $\left[28,7309914; 30,9834702\right]$.

Економічна інтерпретація: якщо у прогнозному періоді доходи населення мають рівень 34,1, то середній роздрібний товарообіг потрапляє в інтервал:

$$29,14478 \le M\left(\widehat{Y}_{\text{прогн}}\right) \le 30,56968.$$

Водночас окреме (індивідуальне) значення роздрібного товарообігу міститься в ширшому інтервалі:

$$28,7309914 \le \widehat{Y}_{\text{прогн}} \le 30,9834702 .$$

11. Побудова графіків лінії тренда, фактичних даних, лінії регресії та її надійних інтервалів.

Використовуючи вкладку ***Работа с диаграммами*** ⇒ ***Макет***, побудуємо лінію тренда для фактичних даних, що задані діаграмою розсіювання на рис. 1.6.

Для цього натискуємо ліву клавішу миші на діаграмі. Потім подамо команду ***Добавить линию тренда***. У результаті з'являється діалогове вікно, в якому визначаємо тип діаграми та задаємо її параметри (рис. 1.16).

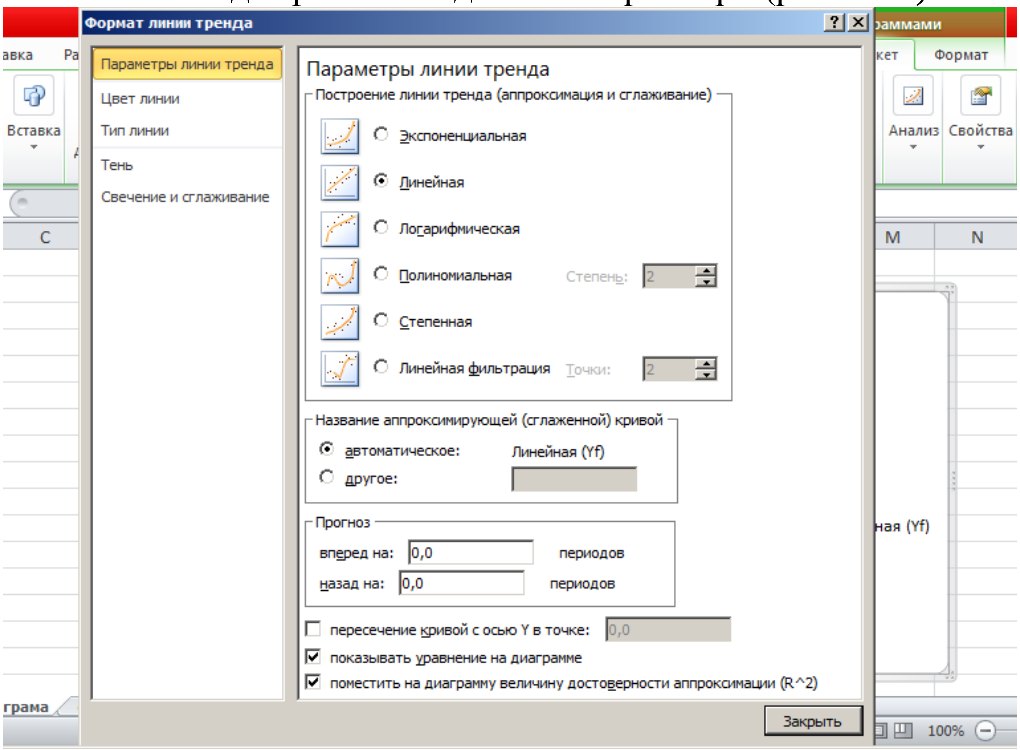


Рисунок 1.16 – Вікно задання лінії тренда

Натискуємо клавішу ***Закрыть*** і отримаємо на діаграмі лінію тренда, її рівняння та коефіцієнт детермінації (рис. 1.17).

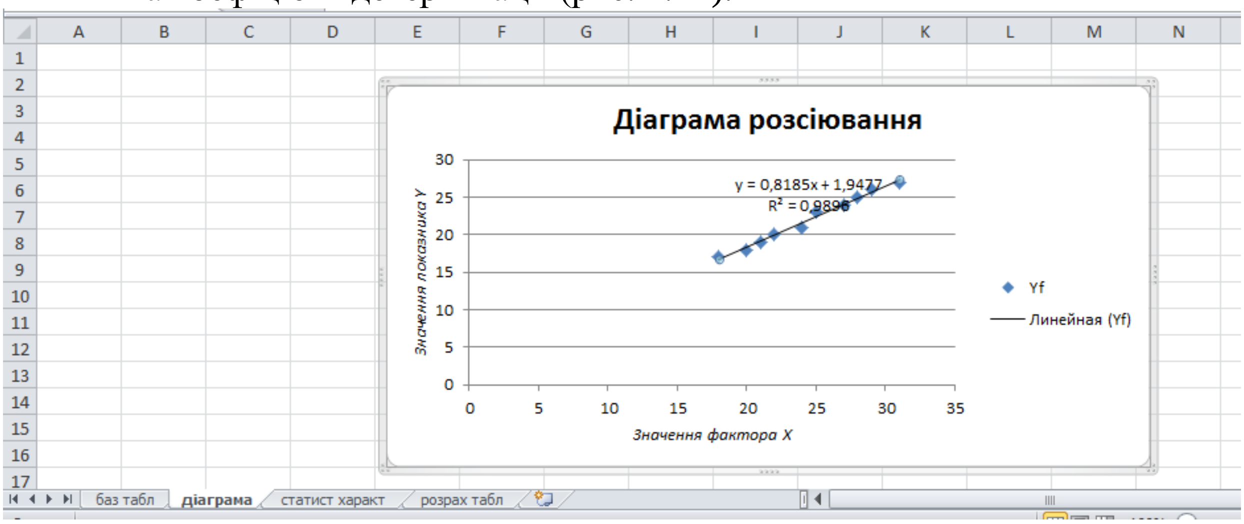


Рисунок 1.17 – Діаграма лінії тренда

Результати, що були отримані за допомогою ***Работа с диаграммами***, а саме рівняння тренда та коефіцієнт детермінації, співпадають з раніше одержаними.

Для візуального порівняння одержаних результатів, за допомогою ***Работа с диаграммами***, побудуємо графіки фактичних даних *y*, лінію регресії $\hat{y}$ та її надійні інтервали $\hat{y}_{\min}$, $\hat{y}_{\max}$ (рис. 1.18).

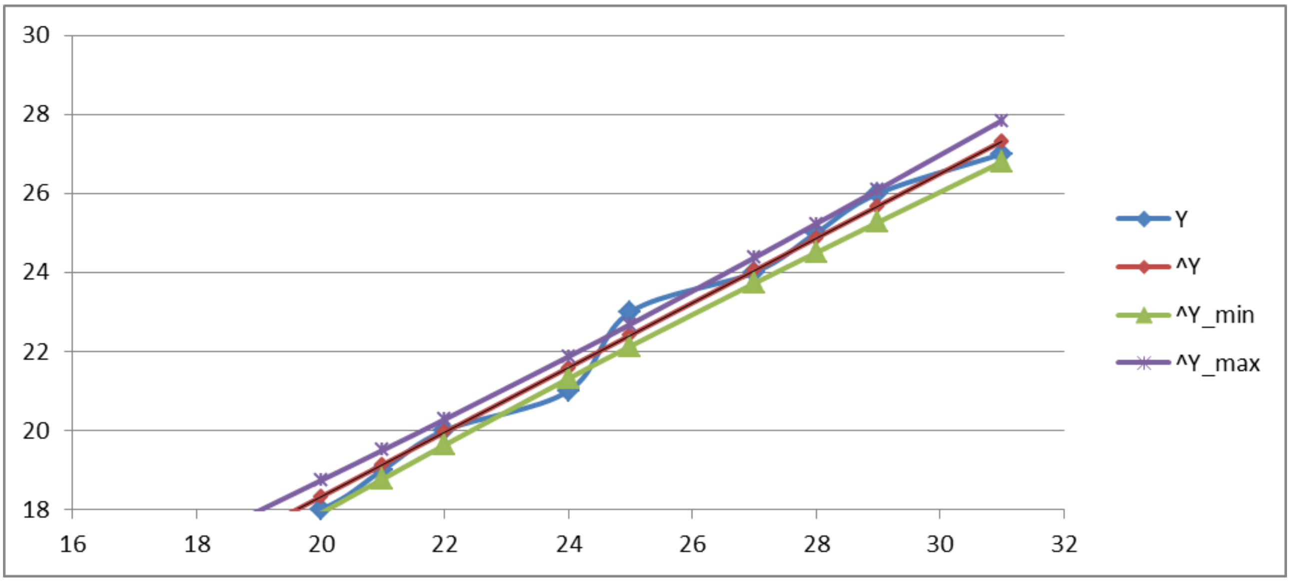


Рисунок 1.18 – Графіки фактичних даних *y*, лінію регресії $\hat{y}$ та її надійні інтервали $\hat{y}_{\min}$, $\hat{y}_{\max}$

12. Перевірка точності економетричної моделі за допомогою середньої відносної похибки апроксимації.

Визначимо ***абсолютну середню відносну похибку апроксимації*** $\bar{\varepsilon}$. Використовуючи розрахунки з таблиці 1.12, маємо

$$\bar{\varepsilon} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \cdot 100\% = \frac{1}{10} \cdot 13{,}10607 = 1{,}31\% < 10\%,$$

що свідчить про високу якість моделі (додаток Г).

Це також ілюструють рис. 1.17 та рис. 1.18, на якихспостерігаються незначні відхилення теоретичних значень $\hat{y}_i$ від відповідних фактичних значень y_i. При порівнянні різних моделей, побудованих на одній статистичній базі, перевагу надають моделям з меншою похибкою апроксимації $\bar{\varepsilon}$.

13. Побудова економетричної моделі за допомогою функції ЛИНЕЙН.

Виділивши область розміром 5 рядків і 2 стовпці ($m = 2$ – кількість параметрів моделі), застосуємо функцію ЛИНЕЙН(Y;X;1;1) (рис. 1.19) і заповнимо таблицю 1.13.

Таблиця 1.13 – Застосування функції ЛИНЕЙН

$\widehat{a}_1$	$\widehat{a}_0$	0,818462	1,947692
$S_{\widehat{a}_1}$	$S_{\widehat{a}_0}$	0,029673	0,736758
R^2	σ_u	0,989594	0,378255
F	$k=n-m$	760,8172	8
$SSR=\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\widehat{Y}_i-Y_i\right)^2$	108,8554	1,144615

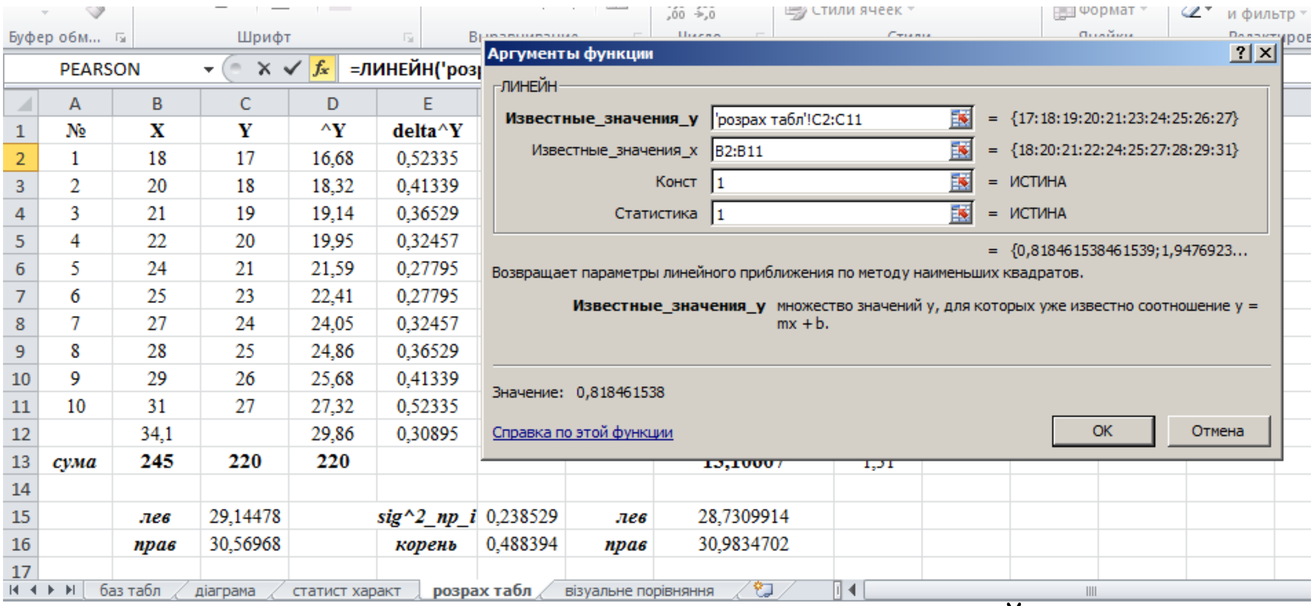


Рисунок 1.19 – Вікно аргументів функції ЛИНЕЙН

Перший рядок результатів розрахунку містить оцінки параметрів моделі: $\widehat{a}_1=0{,}818462$; $\widehat{a}_0=1{,}947692$.

Другий рядок містить стандартні похибки оцінок параметрів моделі: $S_{\widehat{a}_1}=0{,}029673$; $S_{\widehat{a}_0}=0{,}736758$.

Третій рядок містить коефіцієнт детермінації та стандартну помилку залишків: $R^2=0{,}989594$; $\sigma_u=0{,}378255$.

Четвертий рядок містить F–критерій та число ступенів свободи $k=n-m$: $F=760{,}8172$; $k=n-m=8$.

П'ятий рядок таблиці результатів містить суми квадратів: $SSR=\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2=108{,}8554$; $SSE=\sum_{i=1}^{n}\left(\widehat{Y}_i-Y_i\right)^2=1{,}144615$.

Як бачимо, результати побудови економетричної моделі, отримані раніше, повністю збіглися з результатами розрахунку функції ЛИНЕЙН, що і підтверджує їх правильність.

14. Економіко-математичний аналіз характеристик економетричної моделі $\hat{Y}=1{,}94770+0{,}81846X$.

1) коефіцієнт детермінації $R^2 = 0,9896$ показує, що на 98,96% варіація роздрібного товарообігу залежить від рівня доходів населення;

2) коефіцієнт кореляції $r = 0,9948$ (додаток В) свідчить про досить сильний прямий зв'язок між роздрібним товарообігом та рівнем доходів населення;

3) коефіцієнт регресії $\hat{a}_1 = 0,81846$ характеризує граничний розмір витрат на купівлю товарів у роздрібній торгівлі, коли дохід збільшується на 1. Тобто збільшення рівня доходів населення на одиницю сприятиме граничному зростанню роздрібного товарообігу в середньому на 0,81846 одиниць;

4) коефіцієнт еластичності $E_{Y/X} = 0,91$ свідчить, що за збільшенні рівня доходів населення на 1% роздрібний товарообіг гранично зросте на 0,91%;

5) оскільки значення залишкової дисперсії $\hat{\sigma}_u^2 = 0,143077$ наближається до нуля, то це свідчить, що розрахункові значення роздрібного товарообігу дуже близькі до фактичних. Чим менше $\hat{\sigma}_u^2$, тим краще підібрана модель відповідає фактичним даним;

6) згідно з $F -$ критерієм, з надійністю $p = 0,95$ економетричну модель можна вважати адекватною фактичним експериментальним даним і на підставі прийнятої моделі проводити економічний аналіз та знаходити значення прогнозу;

7) оскільки $\bar{\varepsilon} = 1,31\% < 10\%$, це свідчить про досить високу якість прогнозу (додаток Г).