

Звіт з лабораторних робіт № 1-3.

Виконав:
студент групи 201 м
Іванов А.Л.
Перевірила:
доц. Бесклінська О.П.

Тема: Побудова та аналіз економетричної моделі з двома змінними.

Завдання: зробити аналіз залежності обсягу споживання y (у.о.) домогосподарства від наявного прибутку x (у.о.) за вибіркою обсягом $n=12$, результати якої наведено у таблиці. Визначити вид залежності, оцінити параметри рівняння регресії, оцінити силу лінійної залежності між x та y , а також спрогнозувати споживання прибутку при $x=146$ і $x=160$.

№	1	2	3	4	5	6	7	8	9	10	11	12
Наявний прибуток (x у.о.)	107	109	110	113	120	122	123	128	136	140	145	150
Обсяг споживання (y у.о.)	102	105	108	110	115	117	119	125	132	130	141	144

1. Числові характеристики змінних:

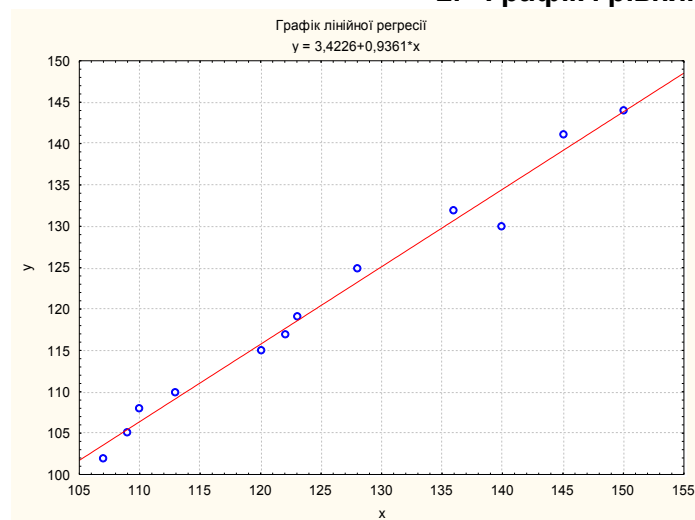
Таблиця 1: Числові характеристики для ознаки x :

Variable	Descriptive Statistics (labor1.sta)			
	Mean	Minimum	Maximum	Std.Dev.
x	125,2500	107,0000	150,0000	14,66675

Таблиця 2: Числові характеристики для ознаки y :

Variable	Descriptive Statistics (labor1.sta)			
	Mean	Minimum	Maximum	Std.Dev.
y	120,6667	102,0000	144,0000	13,84547

2. Графік і рівняння лінійної регресії:



Висновки: Отже рівняння парної лінійної регресії:

$y = 3,423 + 0,936x$. Коефіцієнт **0,936** показує на яку величину зміниться обсяг споживання, якщо доступний дохід збільшиться на одиницю. Він також визначає тангенс кута нахилу прямої регресії відносно додатного напрямку осі абсцис. Вільний член **3,423** визначає прогнозоване значення y при величині наявного прибутку x , що дорівнює нулю.

Середнє значення $x_{сер} = 125,25$ задає центр області прогнозів (Таблиця 1). Середнє квадратичне відхилення $\sigma = 14,666$ характеризує середнє значення розсіювання значень наявного прибутку відносно $x_{сер}$.

Прогнозоване споживання при доступному доході $x=146$ і $x=160$ за допомогою моделі становить $y(146) = 3,423 + 0,936 \cdot 146 = 140,08$ у.о..

$y(160) = 3,423 + 0,936 \cdot 160 = 153,18$ у.о..

3. Перевірка однофакторної лінійної регресії на адекватність.

Знайдемо коефіцієнти *кореляції*, *детермінації*, значення критерію *Фішера*, що спостерігається, число степенів вільності критеріїв *Фішера* і *Стьюдента*.

Таблиця 3: Підсумки регресійного аналізу

N=12	Regression Summary for Dependent Variable: y (labor2.sta)					
	R = ,99160670 R_ = ,98328384 Adjusted R_ = ,98161222 F(1,10)=588,22 p<,00000 Std.Error of estimate: 1,8775					
	Beta	Std.Err. of Beta	B	Std.Err. of B	t(10)	p-level
Intercept			3,422610	4,864432	0,70360	0,497738
x	0,991607	0,040885	0,936080	0,038596	24,25332	0,000000

У цій таблиці знаходять всі необхідні відомості: бачимо, що **коефіцієнт кореляції $R=0,9916067$** . Оскільки він дуже близький до одиниці, то робимо висновок, що лінійний зв'язок досить сильний.

Коефіцієнт детермінації $R^2 = 0,98328384$. Коефіцієнт детермінації досить близький до 1, отже варіація наявного прибутку суттєво впливає на варіацію обсягу споживання.

Значення **критерію Фішера**, що спостерігається, $F(1,10) = 588,22$. Число степенів вільності критерію Фішера $\kappa_1=1$, $\kappa_2=10$. Число степенів вільності критерію Стюдента **t(10):** $k=10$.

Для перевірки значимості коефіцієнтів a_1 і a_0 використовуються значення **критерію Стюдента**, що спостерігаються, для кожного коефіцієнта. Ці значення знаходяться в стовпці **t (10)** таблиці 3:

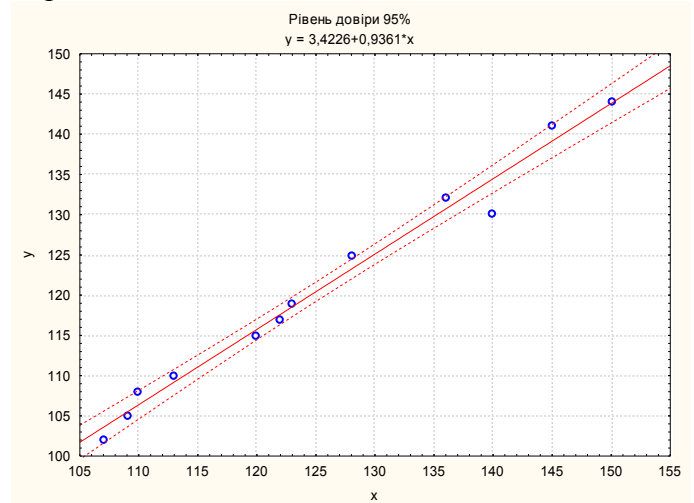
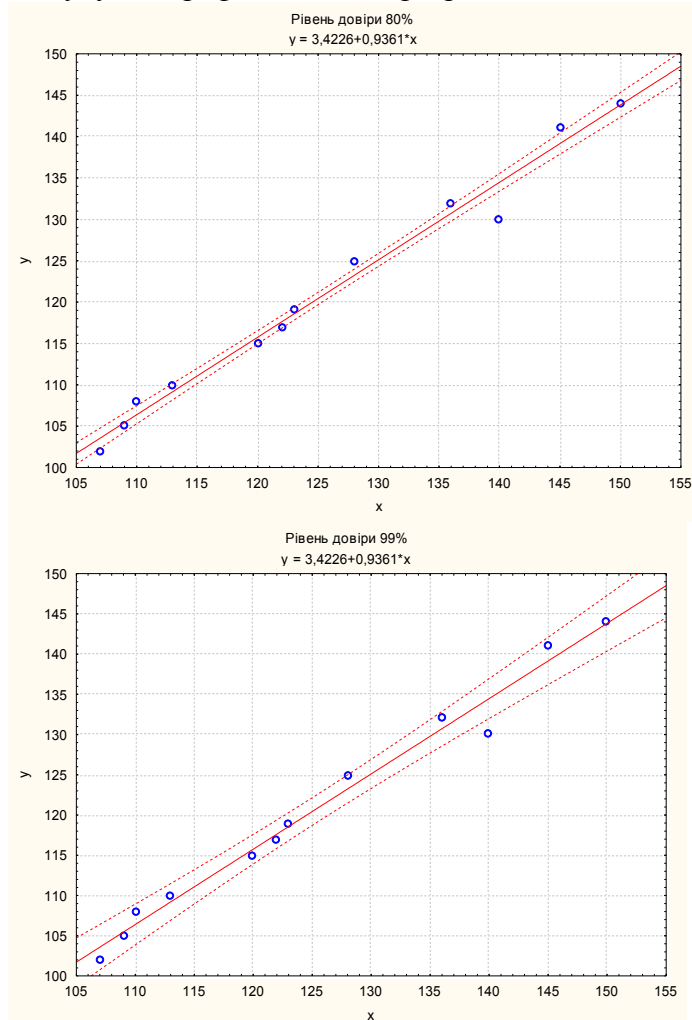
$t_{\text{спост}}(a_0) = 0,7036$, $t_{\text{спост}}(a_1) = 24,25332$. Знайдемо критичне значення **критерію Стюдента** з рівнем значущості $\alpha=0,05$. Маємо $t(10) = 2,22814$ при $p = 0,025$.

Перевіримо статистичну значимість коефіцієнтів a_0 і a_1 . Так як значення критерію Стюдента, що спостерігаються, для коефіцієнта $t_{\text{спост}}(a_0) = 0,7036 < 2,22814$, то цей коефіцієнт *незначущий*. Для коефіцієнта $t_{\text{спост}}(a_1) = 24,25332 > 2,22814$ це означає, що він є *значущий*.

Знайдемо критичне значення критерію Фішера з рівнем значущості $\alpha=0,05$. Маємо, $F(1, 10) = 4,9646$ при $p = 0,95$. Оскільки $F > F_{\text{табл}}(1;10;0,05)$, то гіпотезу $H_0: a_1=0$, про неістотність зв'язку між залежною і незалежною змінною моделі *відкидаємо* з ризиком $\alpha=0,05$, тобто лінійна модель **адекватна**.

4. Прогноз на підставі лінійної регресії. Точність прогнозу.

Побудуємо графіки лінійної регресії з 80, 95 і 99% довірчими областями.



Знайдемо t_γ — критичну точку розподілу Стюдента, що відповідає рівню довіри $\gamma=80$ (95,99)%.

Маємо: $t(10) = 1,37218$ при $p = 0,8$,

$t(10) = 2,22814$ при $p = 0,95$,

$t(10) = 3,16927$ при $p = 0,99$.

Знаходимо прогноз у точках вибірки і двох додаткових точках. У таблицю даних додамо дві точки з області прогнозів, наприклад 146 і 160. Розрахуємо **delta** — напівширина довірчого інтервалу і **pogr** — максимальна помилка прогнозу.

Для розрахунку **delta** при $\gamma=0,80$ у вікно Long Name вводим формулу
 $=1,37218 * 1,8775 * 1 / (12)^{0,5} * (1 + 12 * q / 2366,25)^{0,5}$.

Висновки: Побудовано довірчі інтервали для лінійної регресії $y = 3,4226 + 0,9361 \cdot x$ для трьох рівнів довіри 80%, 95%, 99%. Розраховано прогноз за лінійною регресією у всіх точках вибірки і в двох додаткових точках з області прогнозів $x=146$ і $x=160$, а також розраховані помилки прогнозів.

З порівняння відносних похибок прогнозів видно, що підвищення рівня довіри з 80% до 99% знижує точність прогнозу. Для оцінки в скільки разів знижується точність прогнозу порівняємо відносні похибки для значення $x=146$. Маємо **pogr** (80%)=0,9472, **pogr**(95%)=1,5380, **pogr**(99%)=2,1877.

З приведених помилок видно, що підвищення рівня довіри з 80 до 99% знижує точність прогнозу в

$$\frac{2,1877}{0,9472} = 2,31 \text{ рази.}$$

Для кожного рівня довіри ($\gamma=0,80$, $\gamma=0,95$, $\gamma=0,99$) розрахуємо окрему таблицю. Оцінюємо максимальну відносну помилку прогнозу (у відсотках) для всіх трьох значень коефіцієнта довіри (80%, 95%, 99%).

Таблиця 4: Рівень довіри 80%

	1 x	2 y	3 yregr	4 q	5 delta	6 pogr
1	107	102	103,59	333,0625	1,2196	1,1773
2	109	105	105,46	264,0625	1,1374	1,0786
3	110	108	106,39	232,5625	1,0979	1,0319
4	113	110	109,20	150,0625	0,9869	0,9038
5	120	115	115,75	27,5625	0,7940	0,6859
6	122	117	117,63	10,5625	0,7634	0,6490
7	123	119	118,56	5,0625	0,7532	0,6353
8	128	125	123,24	7,5625	0,7578	0,6149
9	136	132	130,73	115,5625	0,9366	0,7164
10	140	130	134,48	217,5625	1,0786	0,8021
11	145	141	139,16	390,0625	1,2834	0,9223
12	150	144	143,84	612,5625	1,5071	1,0478
13	146		140,09	430,5625	1,3269	0,9472
14	160		153,20	1207,563	1,9850	1,2957
	125,25			2366,25		

Таблиця 5: Рівень довіри 95%

5 delta	6 pogr
1,9803	1,9118
1,8470	1,7514
1,7828	1,6757
1,6026	1,4675
1,2893	1,1138
1,2395	1,0538
1,2230	1,0315
1,2306	0,9985
1,5209	1,1633
1,7514	1,3024
2,0840	1,4976
2,4472	1,7014
2,1547	1,5380
3,2232	2,1040

Таблиця 6: Рівень довіри 99%

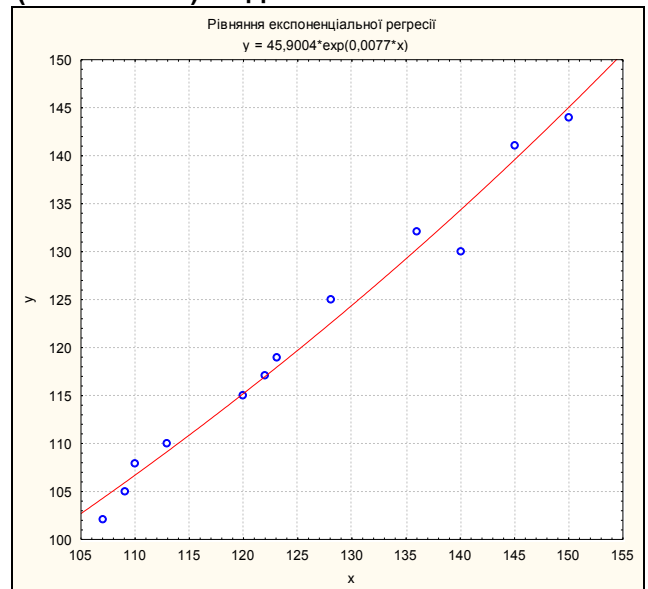
5 delta	6 pogr
2,8168	2,7193
2,6271	2,4911
2,5358	2,3834
2,2794	2,0874
1,8338	1,5842
1,7631	1,4989
1,7396	1,4673
1,7503	1,4202
2,1633	1,6547
2,4912	1,8525
2,9643	2,1302
3,4808	2,4200
3,0648	2,1877
4,5847	2,9926

Висновки.

Побудована модель досить точно описує залежність між обсягом споживання у (у.о.) домогосподарства від наявного прибутку x(у.о.). Значення споживання для кожного рівня довіри знаходяться між $y - \text{delta}$; $y + \text{delta}$. Наприклад для рівня довіри 99% при $x=160$ маємо: (153,2-4,5847; 153,2+4,5847) або (148,62; 157,78).

5. Прогноз за експоненціальною (нелінійною) моделлю.

	1 x	2 y	3 yregr	4 yexp	5 u	6 u ²	7 √ ²
1	107	102	103,58	104,62	-1,01	2,481	6,886
2	109	105	105,45	106,25	-0,76	0,200	1,557
3	110	108	106,38	107,07	-0,64	2,615	0,867
4	113	110	109,19	109,57	-0,36	0,654	0,184
5	120	115	115,74	115,64	0,09	0,552	0,408
6	122	117	117,62	117,43	0,15	0,378	0,188
7	123	119	118,55	118,34	0,18	0,202	0,434
8	128	125	123,23	122,99	0,20	3,129	4,055
9	136	132	130,72	130,80	-0,06	1,641	1,439
10	140	130	134,46	134,89	-0,32	19,918	23,931
11	145	141	139,14	140,19	-0,75	3,448	0,662
12	150	144	143,82	145,69	-1,30	0,031	2,853
13	146		140,08	141,27	-0,85		
14	160		153,18	157,35	-2,72		
1-12						35,25	
1-12							43,462



Висновки.

Сума квадратів залишків для лінійної моделі дорівнює 35,25, а для нелінійної моделі дорівнює 43,462. Так як сума квадратів залишків для лінійної моделі менше, ніж сума квадратів залишків для експоненціальної моделі, лінійна модель є оптимальною.