



ECONOMETRIE

Planul cursului

1. Introducere
2. Modelul de regresie liniară simplă
3. **Modelul de regresie liniară multiplă**
4. Modele de regresie neliniară
5. Ipoteze statistice: normalitatea erorilor,
homoscedasticitatea, necorelarea erorilor,
multicoliniaritatea.

3. Modelul de regresie liniară multiplă

3.1. *Identificarea modelului*

- presupune reprezentarea punctelor (y_i, x_i) .
- dacă toate legăturile sunt liniare, atunci regresia multiplă este liniară.

3.2. *Prezentarea modelului*

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

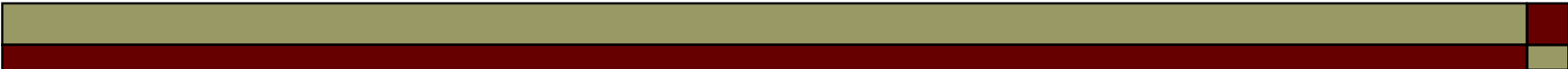
□ *Interpretare:*

- β_0 este valoarea medie a lui Y , atunci când valorile X_i sunt egale cu zero.
- β_i reprezintă *variația medie absolută* a variabilei Y la o creștere cu o unitate a variabilei independente X_i , în condițiile în care influența celorlalte variabile independente este constantă. Măsoară influența parțială a fiecărei variabile independente asupra variabilei dependente.

Example:

1. În studiul legăturii dintre salariul obținut (Y , *euro/oră*), numărul de ani de școală (X_1), numărul de ani de experiență profesională în domeniu (X_2) și vechimea în muncă (X_3), s-a obținut următorul model de regresie estimat:

$$Y_X = 0,284 + 0,092 \cdot X_1 + 0,041 \cdot X_2 + 0,022 \cdot X_3$$

- 
-
2. În studiul legăturii dintre consumul anual pentru un produs (kg/pers.), prețul produsului (cenți/kg) și venitul anual disponibil (mii euro), s-a obținut următorul model de regresie estimat:

$$Y_X = 37,54 - 0,88 \cdot X_1 + 11,9 \cdot X_2$$



3.3. Ipoteze

- normalitatea erorilor;
- homoscedasticitate;
- autocorelarea erorilor;
- lipsa coliniarității.

3.4. Estimarea parametrilor modelului

- Ecuația estimată a modelului de regresie este:

$$y_x = b_0 + b_1 x_1 + b_2 x_2 + \dots b_p x_p$$

Estimarea parametrilor modelului de regresie liniară multiplă se realizează prin MCMMP.

Dacă punem condiția de minim:

$$S = \sum (y_i - y_{x_i})^2 = \sum (y_i - b_0 - b_1 x_{1i} - b_2 x_{2i} - \dots - b_p x_{pi})^2 = \text{minim}$$

De exemplu, în cazul unui model cu 2 variabile independente se obține sistemul de ecuații:

$$\begin{cases} nb_0 + b_1 \sum_i x_{1i} + b_2 \sum_i x_{2i} = \sum_i y_i \\ b_0 \sum_i x_{1i} + b_1 \sum_i x_{1i}^2 + b_2 \sum_i x_{1i} x_{2i} = \sum_i y_i x_{1i} \\ b_0 \sum_i x_{2i} + b_1 \sum_i x_{1i} x_{2i} + b_2 \sum_i x_{2i}^2 = \sum_i y_i x_{2i} \end{cases}$$

3.4. Estimarea parametrilor modelului

- a) **Estimarea punctuală:** b_i sunt estimații punctuale ale parametrilor modelului.
- b) **Estimarea prin IC:** $\left[b_i \pm t_{\alpha/2; n-k} \cdot s_{\hat{\beta}_i} \right]$

Exemplu

În studiul legăturii dintre valoarea vânzărilor unei firme (Y , *mii euro*) și cheltuielile de publicitate (X_1 , *sute euro*), cheltuielile ocazionate de diferite promoții (X_2 , *sute euro*) și vânzările anuale realizate de principalul concurent (X_3 , *mii euro*), s-au obținut următoarele rezultate:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	65,705	27,731		2,369	,037
	X1	48,979	10,658	,581	4,596	,001
	X2	59,654	23,625	,359	2,525	,028
	X3	-1,838	,814	-,324	-2,258	,045

a. Dependent Variable: Y

Exemplu

Știind că $n=15$, se cere:

1. Să se scrie modelul legăturii dintre variabila Y și variabilele independente X_i .
2. Să se interpreteze valoarea parametrului care arată legătura dintre valoarea vânzărilor firmei și cheltuielile de publicitate.
3. Să se calculeze limitele intervalului de încredere pentru parametrul β_1 , pentru un risc de 0,05.
4. Să se estimeze cu cât crește valoarea vânzărilor la o creștere a cheltuielilor de publicitate cu 15 sute euro.
5. Să se estimeze cu cât variază valoarea vânzărilor la o creștere a cheltuielilor de publicitate cu 15 sute euro și a cheltuielilor ocazionate de diferite promoții cu 10 sute euro.

3.5. Măsurarea intensității legăturii dintre variabile

- Se poate efectua cu ajutorul :
 - coeficienților de corelație bivariată și parțială;
 - raportului de determinație multiplă;
 - raportului de corelație multiplă;
 - raportului de determinație multiplă ajustat.



a). Coeficienții de corelație bivariată și parțială

- Coeficienții de corelație bivariată măsoară dependența dintre variabile, fără a lua în considerare influența celorlalte variabile.
- Notatii.

a). Coeficienții de corelație bivariată și parțială

Formula de calcul:

$$r_{yl} = \frac{n \sum_i x_{li} y_i - \sum_i x_{li} \sum_i y_i}{\sqrt{\left[n \sum_i x_{li}^2 - \left(\sum_i x_{li} \right)^2 \right] \left[n \sum_i y_i^2 - \left(\sum_i y_i \right)^2 \right]}}$$

a). Coeficienții de corelație bivariată și parțială

- Coeficienții de corelație parțială măsoară dependența dintre variabile, considerând influența celorlalte variabile constantă.

Exemplu:

- $r_{y1.2}$ – coeficientul de corelație dintre Y și X_1 , considerând constantă influența lui X_2
- $r_{y2.1}$ – coeficientul de corelație dintre Y și X_2 , considerând constantă influența lui X_1
- $r_{12.y}$ - coeficientul de corelație dintre X_1 și X_2 , considerând constantă influența lui Y .

a). Coeficienții de corelație bivariată și parțială

$$r_{y1.2} = \frac{r_{y1} - r_{y2}r_{12}}{\sqrt{(1 - r_{y2}^2)(1 - r_{12}^2)}}$$

$$r_{y2.1} = \frac{r_{y2} - r_{y1}r_{12}}{\sqrt{(1 - r_{y1}^2)(1 - r_{12}^2)}}$$

Coeficienții de corelație bivariată și parțială în SPSS

Correlations

		Y	X1	X2	X3
Y	Pearson Correlation	1	,708**	,612*	-,625*
	Sig. (2-tailed)		,003	,015	,013
	N	15	15	15	15
X1	Pearson Correlation	,708**	1	,161	-,213
	Sig. (2-tailed)	,003		,566	,446
	N	15	15	15	15
X2	Pearson Correlation	,612*	,161	1	-,494
	Sig. (2-tailed)	,015	,566		,061
	N	15	15	15	15
X3	Pearson Correlation	-,625*	-,213	-,494	1
	Sig. (2-tailed)	,013	,446	,061	
	N	15	15	15	15

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Coeficienții de corelație bivariată și parțială în SPSS

Correlations

Control Variables			Y	X1
X2	Y	Correlation	1,000	,780
		Significance (2-tailed)	.	,001
		df	0	12
	X1	Correlation	,780	1,000
		Significance (2-tailed)	,001	.
		df	12	0

b). Raportul de determinație multiplă

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \bar{y})^2}$$

Interpretare: arată cât la sută din variația lui Y depinde de variația simultană a variabilelor factoriale considerate.

c). Raportul de corelație multiplă

$$R = \sqrt{\frac{ESS}{TSS}} = \sqrt{1 - \frac{RSS}{TSS}} = \sqrt{1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \bar{y})^2}}$$

Măsoară influența simultană a variabilelor factoriale asupra variabilei rezultative.

d). Raportul de determinatie multiplă ajustat

O estimatie a raportului de determinatie multiplă ajustat este:

$$\bar{R}^2 = 1 - (1 - R^2) \cdot \frac{n - 1}{n - k}$$

- Raportul de determinatie multiplă ajustat este întotdeauna mai mic sau egal cu raportul de determinatie multiplă.
- Raportul de determinatie multiplă ajustat are o importanță deosebită atunci când se compară mai multe modele de regresie cu un număr diferit de parametri.

Exemplu. Pe baza rezultatelor prezentate în tabelul de mai jos, se cere:

a) să se calculeze și să se interpreteze raportul de determinatie

a) să se calculeze raportul de determinatie multiplă ajustat

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16997,537	3	5665,846	18,290	,000 ^a
	Residual	3407,473	11	309,770		
	Total	20405,009	14			

a. Predictors: (Constant), X3, X1, X2

b. Dependent Variable: Y

3.6. Testarea parametrilor modelului

- Ipoteze: $H_0 : \beta_i = 0$
 $H_1 : \beta_i \neq 0$

- Calculul statisticii test: $t_{calc} = \frac{b_i}{s_{\hat{\beta}_i}}$

- Regula de decizie.

Exemplu:

- Pentru datele prezentate în output-ul de mai jos, se cere să se testeze valoarea parametrului β_1 , considerând un risc de 0,05.

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	65,705	27,731		2,369	,037
X1	48,979	10,658	,581	4,596	,001
X2	59,654	23,625	,359	2,525	,028
X3	-1,838	,814	-,324	-2,258	,045

a. Dependent Variable: Y



3.7. Testarea modelului de regresie

- Ipoteze
- Statistica test
- Regula de decizie

3.8. Testarea indicatorilor de corelație

a) Coeficientul de corelație bivariată

□ Ipoteze statistice $H_0: \rho_{y1}=0$

$H_1: \rho_{y1} \neq 0$

□ Statistica test - t Student

□ Pragul de semnificație

□ Calculul statisticii test:

$$t_{\text{calculat}} = \frac{r_{y1} \sqrt{n-2}}{\sqrt{1-r_{y1}^2}}$$

□ Decizie și interpretare: $\nu=n-2$ grade de libertate.

3.8. Testarea indicatorilor de corelație

b) Coeficientul de corelație parțială

□ Ipoteze statistice $H_0: \rho_{y1.2}=0$

$H_1: \rho_{y1.2} \neq 0$

□ Alegerea statisticii test - Student

□ Alegerea pragului de semnificație

□ Calculul statisticii test:
$$t_{calculat} = \frac{r_{y1.2} \sqrt{n-k}}{\sqrt{1-r_{y1.2}^2}}$$

□ Decizie și interpretare: $\nu=n-k$ grade de libertate.

3.8. Testarea indicatorilor de corelație

c). Raportul de corelație și Raportul de determinație

- se realizează în mod identic ca la modelul de regresie liniară simplă

(Statistica test Fisher)

3.9. Testarea influenței marginale a unei variabile independente

- Atunci când se dorește testarea influenței marginale a unei variabile nou introduse în model sau excluse din model, se folosește un test Fisher.

a) Cazul unei variabile excluse din model

- Prin procedura *Backward* din SPSS, modelarea se realizează într-o primă etapă cu toate variabilele independente. Variabilele care au cea mai mică influență sunt eliminate pas cu pas din modelul de regresie.

3.9. Testarea influenței marginale a unei variabile independente

Ipoteze statistice

H_0 : variabila exclusă din model nu are o influență semnificativă asupra variației variabilei dependente

H_1 : variabila exclusă din model are o influență semnificativă asupra variației variabilei dependente.

3.9. Testarea influenței marginale a unei variabile independente

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.910 ^a	.828	.818	5.2961	.828	83.271	4	69	.000
2	.907 ^b	.822	.814	5.3542	-.006	2.544	1	69	.115

a. Predictors: (Constant), Average female life expectancy, People living in cities (%), Daily calorie intake, People who read (%)

b. Predictors: (Constant), Average female life expectancy, Daily calorie intake, People who read (%)

3.9. Testarea influenței marginale a unei variabile independente

b) Cazul unei variabile nou introduse în model

- Prin procedura *Stepwise* din SPSS, modelarea presupune adăugarea, pas cu pas, a variabilelor care au cea mai mare influență asupra variabilei dependente.



Ipoteze

H_0 : variabila nou adăugată în model nu are o influență semnificativă asupra variației variabilei dependente

H_1 : variabila nou adăugată în model are o influență semnificativă asupra variației variabilei dependente.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.885 ^a	.784	.778	5.398	.784	128.666	2	71	.000
2	.905 ^b	.819	.811	4.973	.035	13.656	1	70	.000

a. Predictors: (Constant), Gross domestic product / capita, People who read (%)

b. Predictors: (Constant), Gross domestic product / capita, People who read (%), Daily calorie intake

3.10. Modele cu variabile standardizate

- Un obiectiv al regresiei liniare multiple este acela de a identifica *ordinea* importanței factorilor care au o influență asupra variabilei dependente.
- Pentru aceasta, se construiește un model cu variabile standardizate, în care variația se măsoară în unități de abatere standard pentru fiecare variabilă.
- Valoarea estimată a coeficienților în modul sau în valoare absolută arată impactul parțial al variației cu o unitate a variabilei independente standardizate asupra variabilei dependente standardizate.



3.10. Modele cu variabile standardizate

- Cel mai mare coeficient în valoare absolută indică cea mai mare influență asupra variabilei dependente, iar semnul coeficientului arată sensul acestei influențe.

Exemplu

În studiul legăturii dintre valoarea vânzărilor unei firme (Y , *mii euro*) și cheltuielile de publicitate (X_1 , *sute euro*), cheltuielile ocazionate de diferite promoții (X_2 , *sute euro*) și vânzările anuale realizate de principalul concurent (X_3 , *mii euro*), s-au obținut următoarele rezultate:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	65,705	27,731		2,369	,037
	X1	48,979	10,658	,581	4,596	,001
	X2	59,654	23,625	,359	2,525	,028
	X3	-1,838	,814	-,324	-2,258	,045

a. Dependent Variable: Y