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.

5. Verificarea ipotezelor modelului de regresie

Ipotezele statistice care se formulează în modelarea econometrică sunt ipoteze asupra componentei aleatoare (erorilor) şi ipoteze asupra componentei deterministe (variabilelor independente).

5.1. Ipotezele asupra componentei aleatoare (erorilor) sunt următoarele:

a. Media erorilor este nulă: $M(\varepsilon_i)=0$.

b. Ipoteza de homoscedasticitate: $V(\varepsilon_i)=\sigma^2$

c. Ipoteza de normalitate: $\varepsilon_i \sim N(0,\sigma^2)$

5. Verificarea ipotezelor modelului de regresie

d. Ipoteza de necorelare sau de independență a erorilor: cov$(\varepsilon_i, \varepsilon_j)=0$.

5.2. Ipotezele asupra variabilelor independente

a. Variabilele independente sunt nestochastice sau deterministe.
b. Variabilele independente şi variabila eroare sunt necorelate, cov $(X_i, \varepsilon_i)=0$.
c. Ipoteza de necoliniaritate a variabilelor independente.

5.1. Ipotezele asupra componentei aleatoare (erorilor)

a. Media erorilor este nulă

1. Definire: $M(\varepsilon_i)=0$ care este echivalentă cu $M(Y/X)=\beta_0+\beta_1 X$.

2. Efectele încălcării acestei ipoteze:

- dacă această ipoteză este încălcată, atunci se modifică proprietățile estimatorilor parametrilor modelului de regresie (parametrii sunt estimați deplasat sau cu o eroare sistematică).

a. Media erorilor este nulă

3. Testarea ipotezei cu privire la media erorilor

- Ipoteze:

H_0: $M(\varepsilon)=0$

H_1: $M(\varepsilon)\#0$

- Alegerea testului:

$$t = \frac{\hat{M}(\varepsilon) - M(\varepsilon)}{\sqrt{\hat{V}(\hat{M}(\varepsilon))}}$$

- Calculul statisticii test:

$$t_{calc} = \frac{M(e_i)}{s_{\hat{M}(\varepsilon)}}$$

- Decizie

a. Media erorilor este nulă

4. Exemplu

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	58508,12	6084511	1582428	1957596,554	15
Residual	-98125,2	131202,7	,00000	73271,63549	15
Std. Predicted Value	-,778	2,300	,000	1,000	15
Std. Residual	-1,290	1,725	,000	,964	15

a. Dependent Variable: salariu

a. Media erorilor este nulă

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Unstandardized Residual	15	,0000000	73271,63549	18918,65

a. Media erorilor este nulă

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Unstandardized Residual	,000	14	1,000	,00000000	-40576,5	40576,48

5. Corectarea modelului

- Modelul inițial se corectează cu ajutorul estimației erorilor calculate la nivelul eşantionului.
- Modelul corectat este de forma:

$$y_i^* = \beta_0 + \beta_i x_i + u_i$$, unde:

$$y_i^* = y_i - M(\varepsilon_i)$$

b. Ipoteza de homoscedasticitate a erorilor

1. ***Definire***

- ipoteza de homoscedasticitate presupune ca varianța erorilor să fie constantă:

$$V(\varepsilon_i) = \sigma^2$$

- această ipoteză presupune o varianță constantă şi egală a erorilor la nivelul distribuțiilor condiționate de forma

$$Y \mid X = x_i$$

2. Efectele încălcării acestei ipoteze

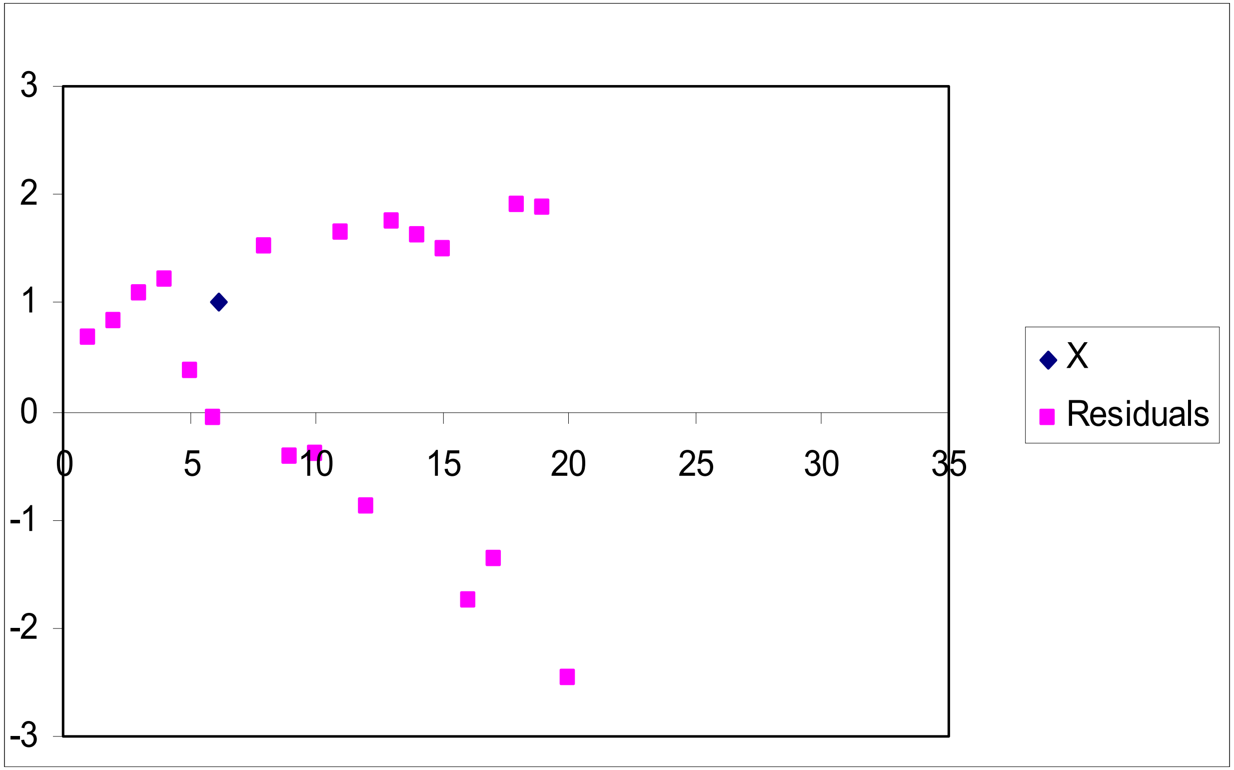
- pierderea eficienței estimatorilor parametrilor modelului de regresie (estimează parametrul cu o varianță mai mare).

3. Identificarea heteroscedasticității

3.1. Procedee grafice

- presupun reprezentarea distribuției erorilor și aprecierea varianței acesteia.

- cazul existenței heteroscedasticității:



3.2. Procedee numerice
a. Testul Glejser (pentru modele simple)

- are la bază un model de regresie între variabila reziduală estimată şi variabila independentă.

Etapele testării:

1. Se estimează modelul de regresie de forma:

$$Y = \beta_0 + \beta_1 \cdot X + \varepsilon$$

2. Se calculează erorile estimate e_i.
3. Se construieşte un model de regresie pe baza erorilor estimate în valoare absolută şi variabila

a. Testul Glejser

independentă. Exemplu:

$$|\varepsilon_i| = \alpha_0 + \alpha_1 \cdot x_i + u_i$$

4. Se testează parametrii acestui model: dacă parametrul α_1 este semnificativ, atunci modelul inițial este heteroscedastic.

Exemplu:

a. Testul Glejser

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	50921,663	12000,771		4,243	,001
	PIB	,016	,012	,348	1,337	,204

a. Dependent Variable: erori

b. Testul Breusch-Pagan-Godfrey (pentru modele multiple)

- este similar testului Glejser, cu excepția faptului că se consideră ca variabilă dependentă pătratul erorilor.

b. Testul Breusch-Pagan-Godfrey (pentru modele multiple)

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.991388	Prob. F(2,7)	0.4177
Obs*R-squared	2.207309	Prob. Chi-Square(2)	0.3317
Scaled explained SS	1.612479	Prob. Chi-Square(2)	0.4465

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.516496	658.6977	-0.005339	0.9959
X1	0.855525	2.856270	0.299525	0.7732
X2	-0.480219	3.606007	-0.133172	0.8978

R-squared	0.220731	Mean dependent var	44.96155
Adjusted R-squared	-0.001917	S.D. dependent var	81.83752
S.E. of regression	81.91594	Akaike info criterion	11.89259
Sum squared resid	46971.55	Schwarz criterion	11.98336
Log likelihood	-56.46295	Hannan-Quinn criter.	11.79301
F-statistic	0.991388	Durbin-Watson stat	1.709325
Prob(F-statistic)	0.417740		

4. Corectarea heteroscedasticității

4.1. Dacă se cunosc parametrii σ_i^2

Corecția heteroscedasticității este aplicată modelului de regresie liniară simplă:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Corectarea heteroscedasticității presupune ponderarea modelului inițial cu variabila $\frac{1}{\sigma_i}$.

Noul model de regresie (corectat) se obține astfel:

$$\frac{y_i}{\sigma_i} = \frac{\beta_0}{\sigma_i} + \beta_1 \frac{x_i}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i}$$

Estimarea parametrilor acestui model se realizează pe baza MCMMP ponderată (*method of weighted least squares*).

4.2. Dacă nu se cunosc parametrii σ_i^2

Corecția heteroscedasticității se realizează pe baza relației:

$$\sigma_i^2 = \sigma^2 x_i^2$$

Corectarea heteroscedasticității presupune ponderarea modelului inițial cu variabila $1/x_i$.

Noul model de regresie (corectat) este de forma:

$$\frac{y_i}{x_i} = \frac{\beta_0}{x_i} + \beta_1 + \frac{\varepsilon_i}{x_i}$$

c. Ipoteza de normalitate a erorilor

1. Formularea problemei

- erorile ε urmează o lege normală de medie 0 şi varianță σ^2:

$$\varepsilon_i \sim N(0, \sigma^2)$$

2. Efectele încălcării acestei ipoteze

- ipoteza de normalitate a erorilor este importantă pentru stabilirea proprietăților estimatorilor parametrilor modelului de regresie.

- dacă $\varepsilon_i \sim N(0,\sigma^2)$, atunci estimatorii parametrilor modelului de regresie urmează, de asemenea, o lege normală:

$$\hat{\beta}_i \sim N(\beta_i, \sigma^2_{\hat{\beta}_i})$$

- dacă ipoteza de normalitate este încălcată, atunci estimatorii parametrilor modelului de regresie nu urmează o lege normală (pentru eşantioane de volum mare, proprietatea de normalitate este atinsă asimptotic).

3. Verificarea normalității erorilor

3.1. Procedee grafice

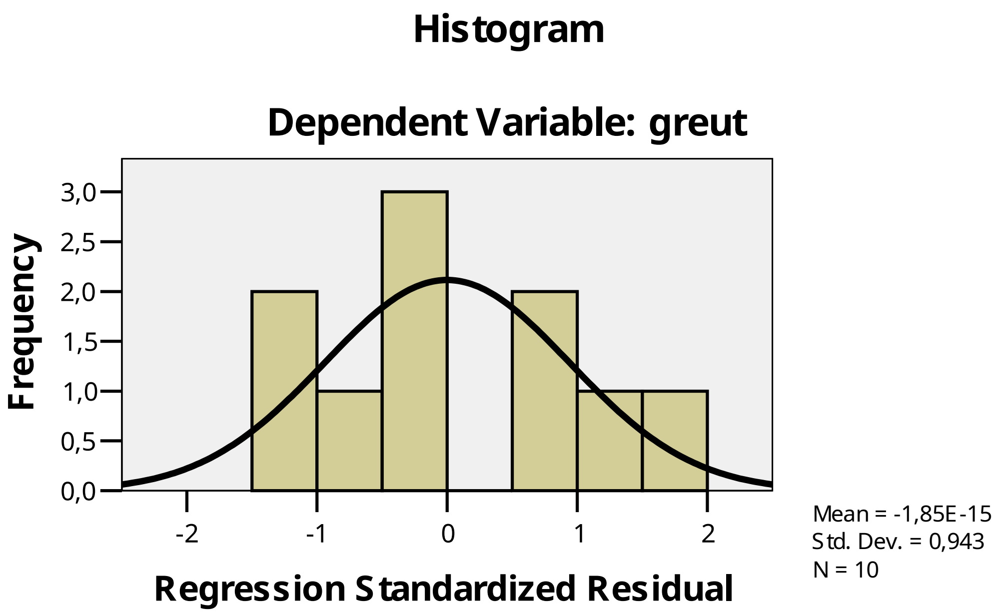
- Histograma (curba frecvenței):
- P-P Plot sau Q-Q Plot;
- Box-Plot.

a. *Reprezentarea histogramei şi a curbei frecvențelor*

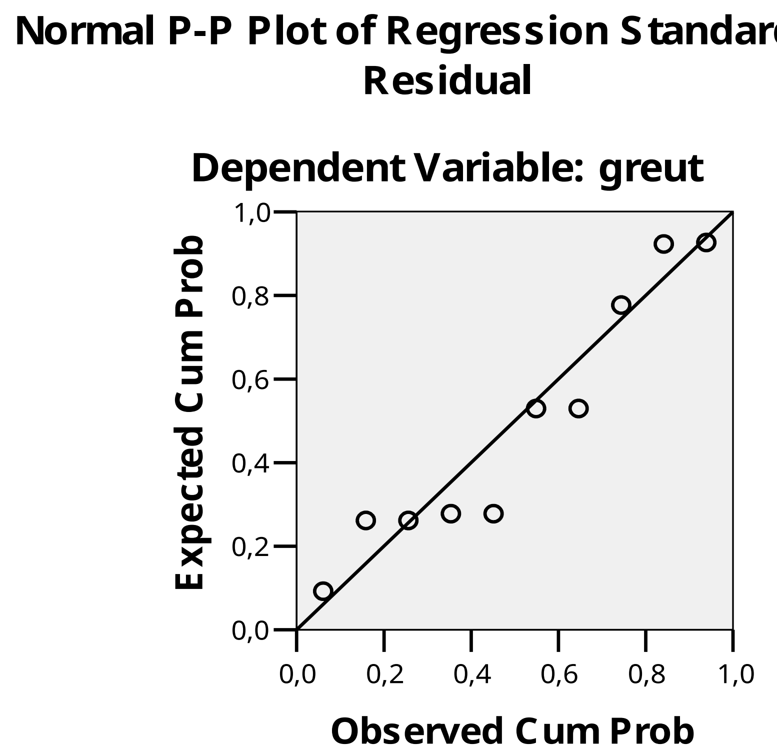
- se reprezintă curba frecvenței sau histograma reziduurilor şi se observă dacă forma distribuției acestora are alură de clopot.

Exemplu:

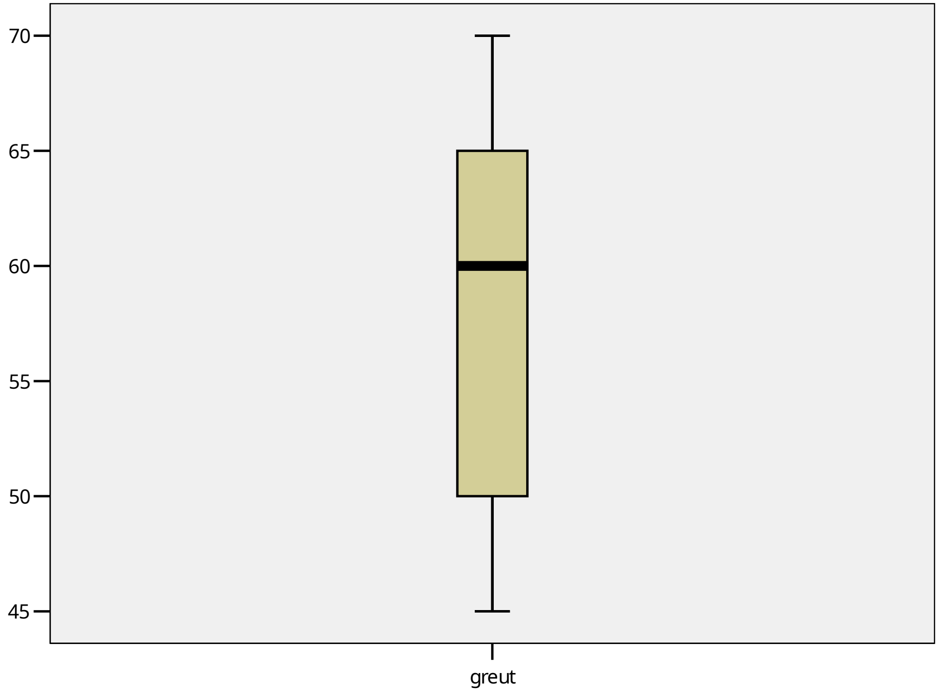
Histograma şi curba frecvențelor



b. P-P Plot sau *Q-Q Plot*



c. Box-plot



3.2. Procedee numerice

a. Testul Kolmogorov-Smirnov

- presupune compararea frecvențelor cumulate (calculate) cu frecvențele teoretice cumulate extrase din tabelul Gauss.

Ipoteze statistice:

H_0: ipoteza de normalitate

H_1: distribuția erorilor nu urmează o lege normală

Regula de decizie:

- valoarea probabilității asociate statisticii test calculate (Sig.) se compară cu α (0,05): dacă Sig.<0,05, atunci se respinge ipoteza de normalitate a erorilor.

Exemplu:

One-Sample Kolmogorov-Smirnov Test

		greut
N		10
Normal Parameters [a,b]	Mean	58,5000
	Std. Deviation	7,83511
Most Extreme Differences	Absolute	,276
	Positive	,161
	Negative	-,276
Kolmogorov-Smirnov Z		,873
Asymp. Sig. (2-tailed)		,432

a. Test distribution is Normal.

b. Calculated from data.

b. Testul Jarque-Bera

- se bazează pe verificarea simultană a proprietăților de asimetrie şi boltire ale seriei reziduurilor. Pentru o distribuție normală, valoarea coeficientului de asimetrie Fisher (*sw*) este zero, iar valoarea coeficientului de boltire Fisher (*k*) este zero.

Ipoteze statistice

H_0: ipoteza de normalitate

H_1: distribuția erorilor nu urmează o lege normală

Calculul statisticii test:

- statistica test JB se calculează după relația:

$$JB = \frac{n}{6} \cdot \left(sw^2 + \frac{k^2}{4} \right)$$

unde: *sw* este coeficientul de asimetrie (*Skewness*):

$$sw = \frac{\mu_3}{\sigma^3}$$

k este coeficientul de boltire (*Kurtosis*):

$$k = \frac{\mu_4}{\mu_2^2} - 3$$

- În cazul unui model de regresie, estimațiile celor doi parametri ai formei unei distribuții au următoarele relații:

$$sw = \sqrt{\frac{\left(\sum_i \frac{e_i^3}{n-2}\right)^2}{\left(\sum_i \frac{e_i^2}{n-2}\right)^3}}$$

$$k = \frac{\sum_i \frac{e_i^4}{n-2}}{\left(\sum_i \frac{e_i^2}{n-2}\right)2} - 3$$

unde: $e_i = y_i - y_{x_i}$

Regula de decizie:
Statistica JB urmează o lege $\chi^2_{\alpha,2}$.

- dacă valoarea calculată a statisticii test JB > $\chi^2_{\alpha;2}$ sau Sig.<0,05, atunci se respinge ipoteza Ho.

Exemplu.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Unstandardized Residual	7	-2.61905	2.14286	.0000000	1.889822	-.529	.794	-1.375	1.587
Valid N (listwise)	7								