



Βιομετρία

Ενότητα 3^η:

Πολλαπλή Γραμμική Παλινδρόμηση

Γεώργιος Μενεξές

Τμήμα Γεωπονίας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

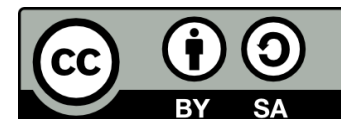


ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

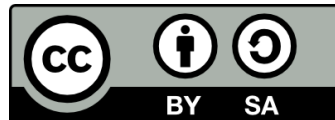


ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.





Πολλαπλή Γραμμική Παλινδρόμηση

Ενότητα 3^η

Το μαθηματικό μοντέλο της ΠΓΠ (1)

Μελέτη της γραμμικής σχέσης μεταξύ 1 εξαρτημένης (Y) και 2 ή περισσότερων ανεξάρτητων μεταβλητών (X_i)

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \dots + \beta_k X_{k-1,i} + \varepsilon_i$$

Σταθερός όρος

Συντελεστές Παλινδρόμησης (slopes)

Τυχαίο Σφάλμα



Το μαθηματικό μοντέλο της ΠΓΠ (2)

Οι συντελεστές παλινδρόμησης εκτιμώνται με τη βοήθεια των δεδομένων

Εκτιμώμενη τιμή
της Y

Εκτιμώμενος
Σταθερός
Όρος

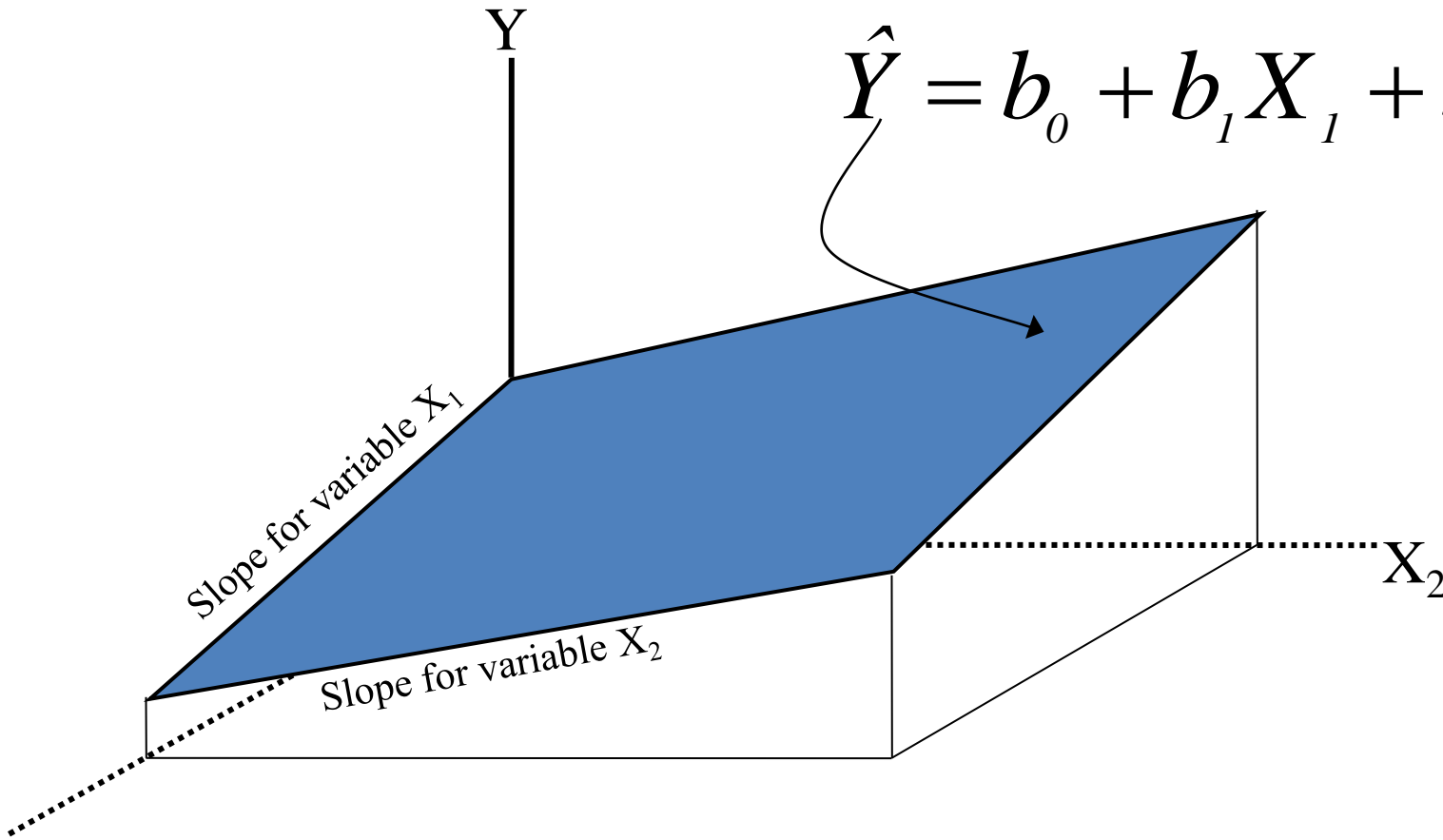
Εκτιμώμενοι συντελεστές

$$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{k,i}$$



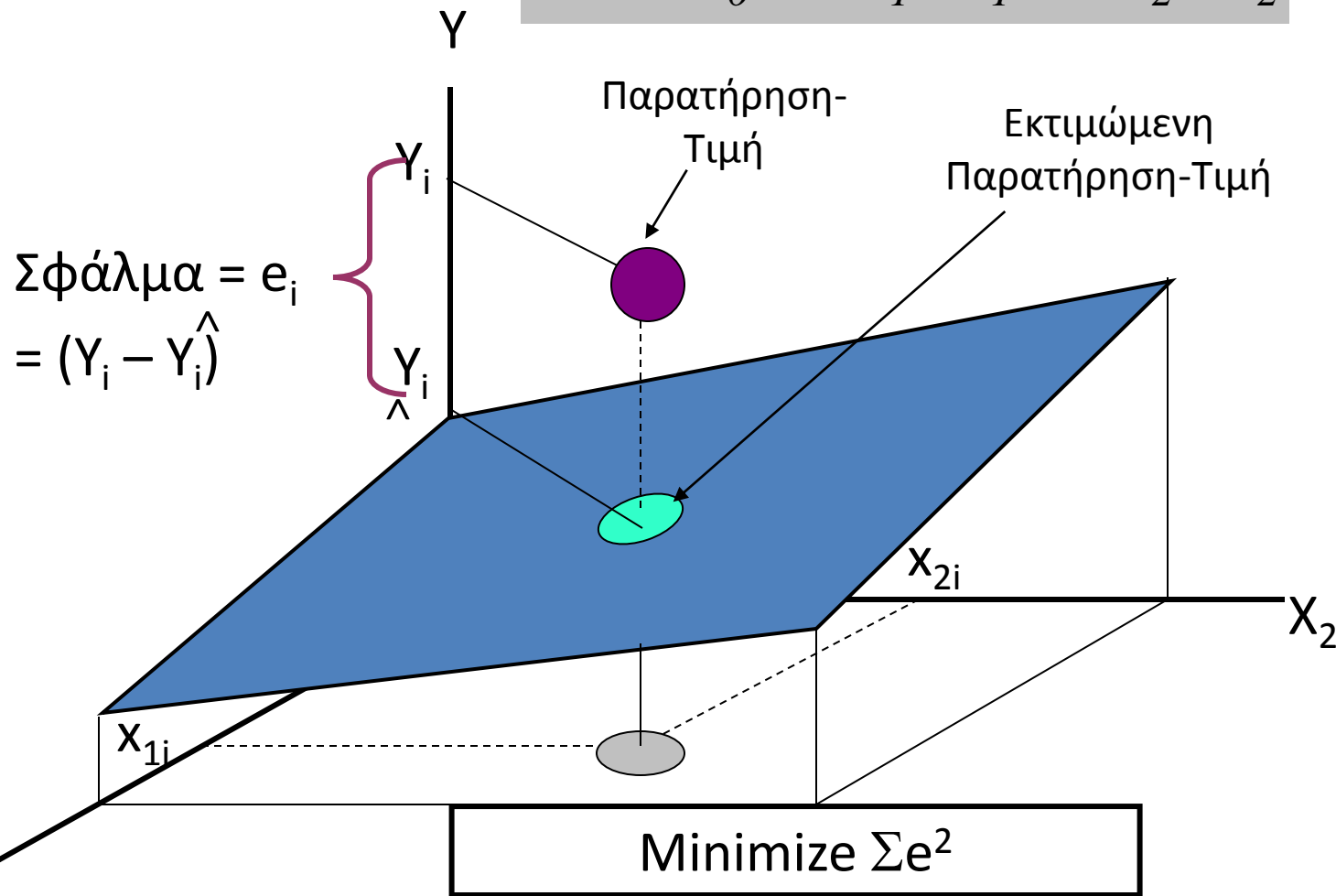
Παράδειγμα

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$$



Μέθοδος Ελαχίστων Τετραγώνων

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$$



Πρόβλημα

- Από μια ποικιλία σιταριού εκτιμήθηκαν τρία χαρακτηριστικά του κόκκου, το βάρος Y , το μήκος X και το πλάτος Z . Μελετήθηκε η σχέση που συνδέει το βάρος Y με το μήκος X και το πλάτος Z μέσω του υποδείγματος της Πολλαπλής Ευθύγραμμης Συμμεταβολής.
- Τα διαθέσιμα δεδομένα αναλύθηκαν μέσω στατιστικού πακέτου και μέρος των αποτελεσμάτων δίνεται στους πίνακες που ακολουθούν. Με βάση τα αποτελέσματα να απαντήσετε στα παρακάτω ερωτήματα:



Εισαγωγή δεδομένων

*Data Exams Biometry Regression.sav [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 :

	Y	X	Z	var	var	var	var
1	49,00	71,00	32,00				
2	39,00	68,00	30,00				
3	54,00	73,00	34,00				
4	42,00	67,00	33,00				
5	45,00	68,00	34,00				
6	37,00	65,00	30,00				
7	52,00	74,00	32,00				
8	30,00	64,00	28,00				
9	35,00	64,00	29,00				
10	47,00	70,00	32,00				
11							
12							
13							
14							
15							
16							

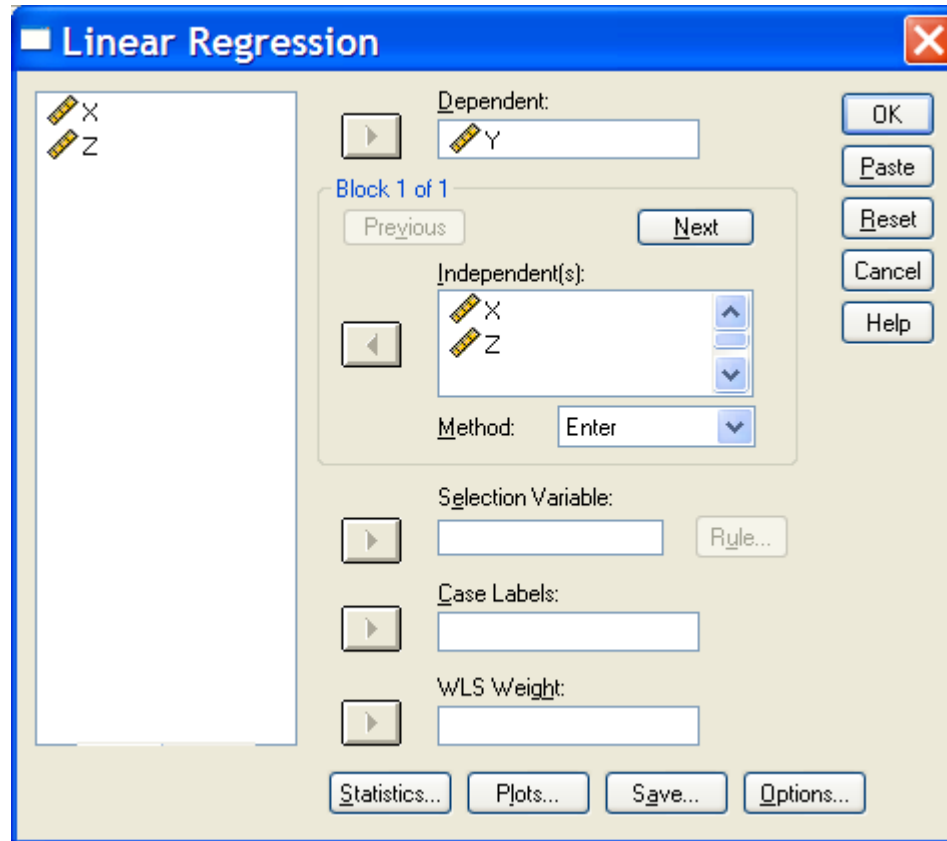


Εντολές (1)

The screenshot shows the SPSS Data Editor interface. The title bar reads '*Data Exams Biometry Regression.sav [DataSet1] - SPSS Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, and Help. The 'Analyze' menu is open, showing a list of options: Reports, Descriptive Statistics, Tables, Compare Means, General Linear Model, Generalized Linear Models, Mixed Models, Correlate, Regression (highlighted), Loglinear, Classify, Data Reduction, Scale, Nonparametric Tests, Time Series, Survival, Multiple Response, Missing Value Analysis..., Complex Samples, Quality Control, and ROC Curve... The 'Regression' submenu is also open, listing: Linear... (highlighted), Curve Estimation..., Binary Logistic..., Multinomial Logistic..., Ordinal..., Probit..., Nonlinear..., Weight Estimation..., 2-Stage Least Squares..., and Optimal Scaling... The data grid shows 19 rows with columns Y and X. The data points are: (1, 49.00, 71.00), (2, 39.00, 68.00), (3, 54.00, 73.00), (4, 42.00, 67.00), (5, 45.00, 68.00), (6, 37.00, 65.00), (7, 52.00, 74.00), (8, 30.00, 64.00), (9, 35.00, 64.00), (10, 47.00, 70.00), and rows 11-19 are empty.

	Y	X
1	49,00	71,00
2	39,00	68,00
3	54,00	73,00
4	42,00	67,00
5	45,00	68,00
6	37,00	65,00
7	52,00	74,00
8	30,00	64,00
9	35,00	64,00
10	47,00	70,00
11		
12		
13		
14		
15		
16		
17		
18		
19		

Εντολές (2)



Επιλογές (1)

Linear Regression: Statistics

Regression Coefficients

- Estimates
- Confidence intervals
- Covariance matrix
- Model fit
- R squared change
- Descriptives
- Part and partial correlations
- Collinearity diagnostics

Residuals

- Durbin-Watson
- Casewise diagnostics
 - Outliers outside: 3 standard deviations
 - All cases

Continue
Cancel
Help

Επιλογές (2)

Linear Regression: Plots

Scatter 1 of 1

DEPENDNT
*ZPRED
*ZRESID
*DRESID
*ADJPRED
*SRESID
*SDRESID

Previous Next

Y: *ZRESID

X: *ZPRED

Continue
Cancel
Help

Standardized Residual Plots

Produce all partial plots

Histogram

Normal probability plot

Επιλογές (3)

Linear Regression: Save [X]

Predicted Values

- Unstandardized
- Standardized
- Adjusted
- S.E. of mean predictions

Distances

- Mahalanobis
- Cook's
- Leverage values

Prediction Intervals

- Mean Individual
- Confidence Interval: %

Residuals

- Unstandardized
- Standardized
- Studentized
- Deleted
- Studentized deleted

Influence Statistics

- DiBeta(s)
- Standardized DiBeta(s)
- DiFit
- Standardized DiFit
- Covariance ratio

Coefficient statistics

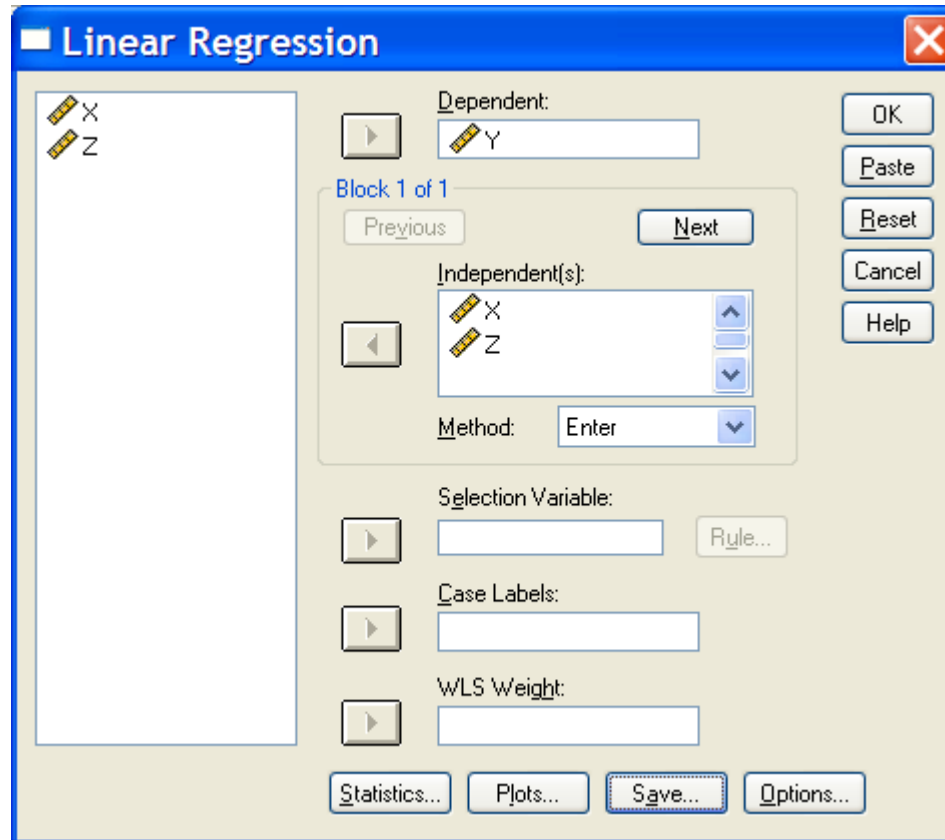
- Create coefficient statistics
- Create a new dataset
Dataset name:
- Write a new data file

Export model information to XML file

- Include the covariance matrix



Τέλος διαδικασίας



Αποτελέσματα (1)

Descriptive Statistics

	Mean	Std. Deviation	N
Y	43,0000	7,77460	10
X	68,4000	3,56526	10
Z	31,4000	2,06559	10

Correlations

		Y	X	Z
Pearson Correlation	Y	1,000	,954	,837
	X	,954	1,000	,685
	Z	,837	,685	1,000
Sig. (1-tailed)	Y	.	,000	,001
	X	,000	.	,014
	Z	,001	,014	.
N	Y	10	10	10
	X	10	10	10
	Z	10	10	10



Αποτελέσματα (2)

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,987 ^a	,974	,966	1,42827

a. Predictors: (Constant), Z, X

b. Dependent Variable: Y

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	529,720	2	264,860	129,835	,000 ^a
	Residual	14,280	7	2,040		
	Total	544,000	9			

a. Predictors: (Constant), Z, X

b. Dependent Variable: Y



Αποτελέσματα (3)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-104,843	9,244		-11,342	,000	-126,701	-82,986					
	X	1,564	,183	,717	8,530	,000	1,130	1,997	,954	,955	,522	,531	1,884
	Z	1,303	,316	,346	4,117	,004	,554	2,051	,837	,841	,252	,531	1,884

a. Dependent Variable: Y

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions		
				(Constant)	X	Z
1	1	2,997	1,000	,00	,00	,00
	2	,002	39,201	,67	,00	,47
	3	,001	59,059	,33	1,00	,53

a. Dependent Variable: Y



Αποτελέσματα (4)

Residuals Statistics^a

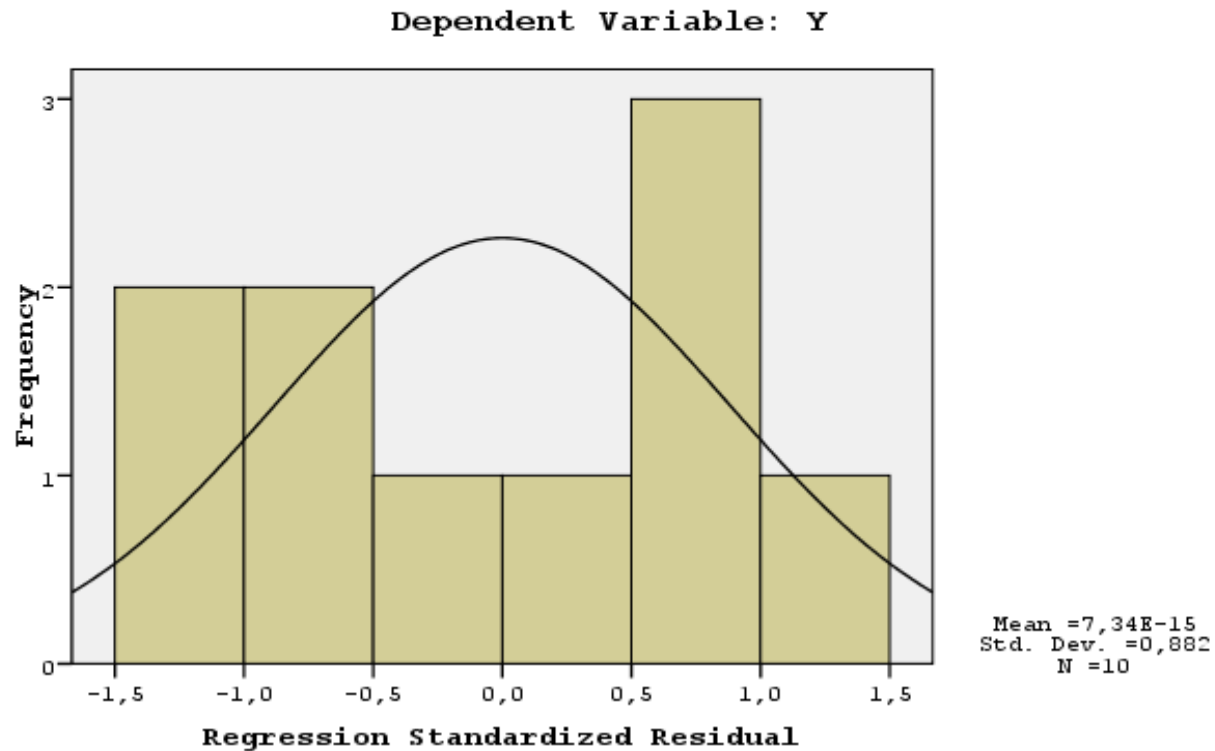
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	31,6920	53,5787	43,0000	7,67188	10
Std. Predicted Value	-1,474	1,379	,000	1,000	10
Standard Error of Predicted Value	,500	1,013	,764	,177	10
Adjusted Predicted Value	32,1746	53,3855	43,1548	7,69430	10
Residual	-1,69200	2,00549	,00000	1,25962	10
Std. Residual	-1,185	1,404	,000	,882	10
Stud. Residual	-1,534	1,667	-,044	1,052	10
Deleted Residual	-2,83617	2,82545	-,15483	1,80982	10
Stud. Deleted Residual	-1,743	1,987	-,037	1,142	10
Mahal. Distance	,202	3,629	1,800	1,195	10
Cook's Distance	,013	,530	,153	,169	10
Centered Leverage Value	,022	,403	,200	,133	10

a. Dependent Variable: Y



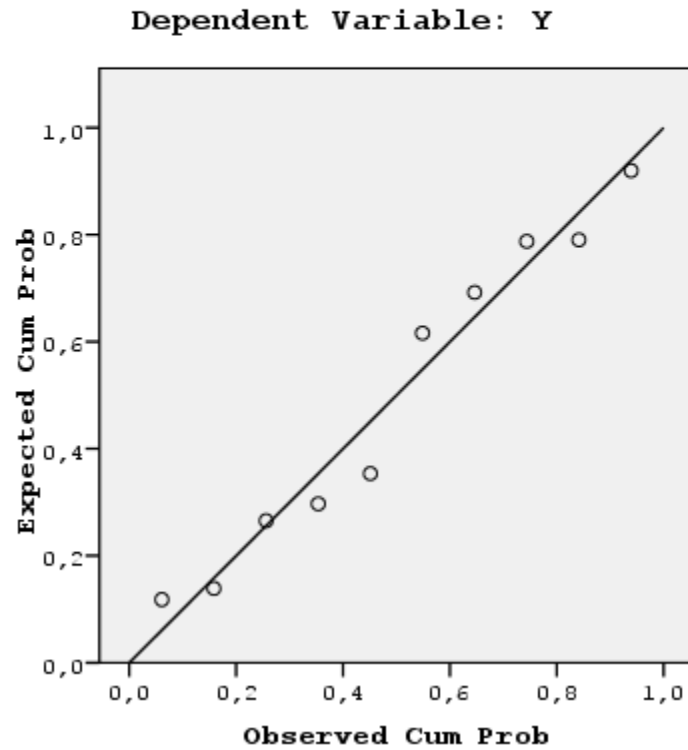
Αποτελέσματα (5)

Histogram



Αποτελέσματα (6)

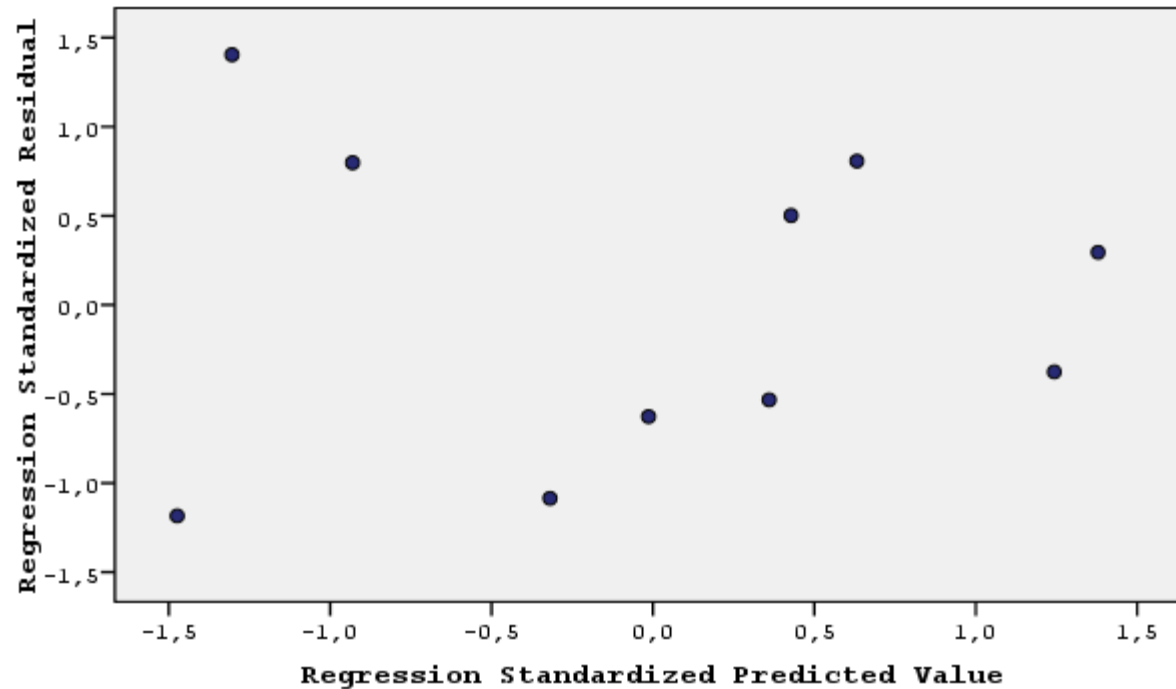
Normal P-P Plot of Regression Standardized Residual



Αποτελέσματα (7)

Scatterplot

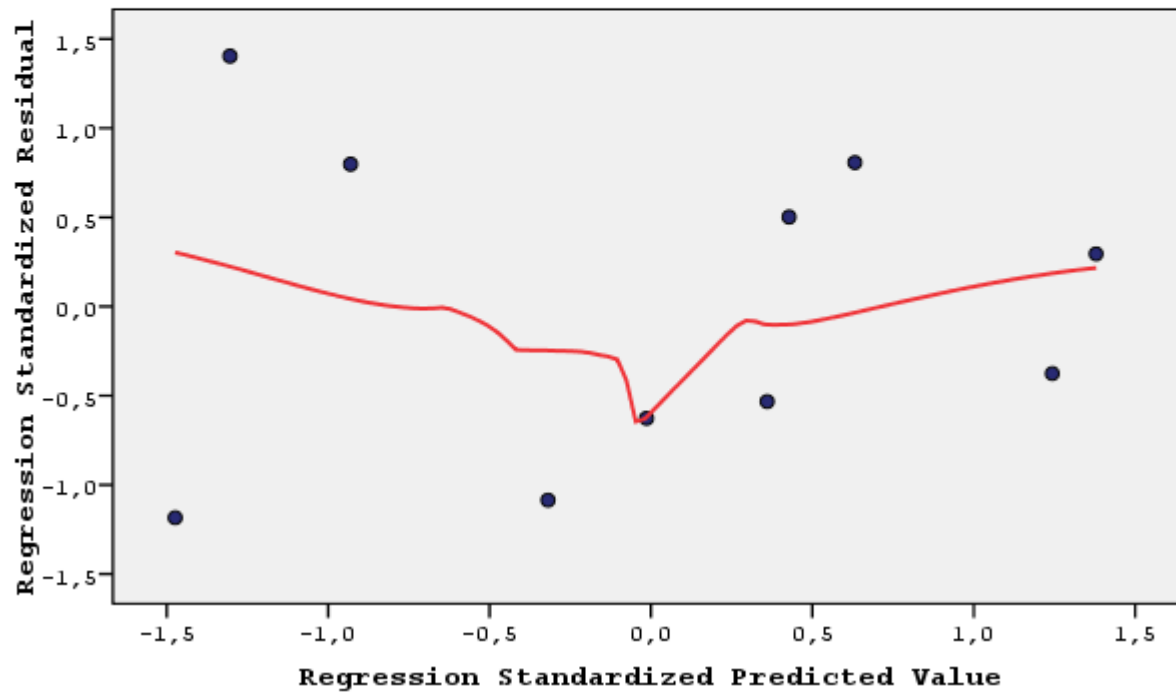
Dependent Variable: Y



Αποτελέσματα (8)

Scatterplot

Dependent Variable: Y



Αποτελέσματα (9)

ometry Regression.sav [DataSet1] - SPSS Data Editor

Transform Analyze Graphs Utilities Window Help

-0,201087672602483

	Z	PRE_1	RES_1	MAH_1	SDF_1	SDB0_1	SDB1_1	SDB2_1	var
,00	32,00	47,84664	1,15336	,61415	,39123	-20109	,23142	-,09123	
,00	30,00	40,55108	-1,55108	,69289	-,57630	-,12664	-,22066	,37662	
,00	34,00	53,57869	,42131	1,92953	,22540	-,17466	,07872	,06896	
,00	33,00	42,89510	-,89510	2,20588	-,54429	-,06973	,39138	-,44238	
,00	34,00	45,76113	-,76113	3,37305	-,67381	,03578	,43595	-,59754	
,00	30,00	35,86054	1,13946	,91058	,44054	,32530	-,21992	-,01103	
,00	32,00	52,53718	-,53718	3,62933	-,50772	,32175	-,44917	,25718	
,00	28,00	31,69200	-1,69200	2,73080	-1,43305	-1,07237	,11008	,82650	
,00	29,00	32,99451	2,00549	1,71184	1,27037	1,04476	-,47284	-,34152	
,00	32,00	46,28313	,71687	,20194	,18920	-,07096	,06180	-,00419	



Βιβλιογραφία

- **Hair, J., Anderson, R., Tatham, R. & Black, W. (1995).** *Multivariate Data Analysis With Readings*. New Jersey: Prentice-Hall International, Inc.
- **Kirk, R. (1995).** *Experimental Design: Procedures for the Behavioral Sciences*. Pacific Grove: Brooks/Cole Publishing Company.
- **Ζαχαροπούλου, Χ. (1994).** *Παλινδρόμηση - Συσχέτιση: Θεωρία και Πράξη*. Θεσσαλονίκη: Προσωπική Έκδοση.





Τέλος Ενότητας

Επεξεργασία: Μαρία Αλεμπάκη
Θεσσαλονίκη, Φεβρουάριος 2014



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

