

Estimation of Linear and Log linear Models in SPSS

M.A. / M.Sc. in Economics
Semester IV

Paper: ECO 401 (Computer Application)

Dr. Pinaki Das

Associate Professor and Head

Department of Economics

Vidyasagar University



Data

Demand for roses (q_r)	price of rose(P_x)	price of carnation(P_y)	income(M)
11484	2.3	3.5	158.1
9348	2.5	2.9	173.4
8429	3.1	4.1	165.3
10079	2.9	3.6	172.9
9240	2.7	3.2	178.5
8862	2.8	3.7	198.6
6216	3.6	3.8	186.3
8253	3.2	3.5	189.0
8038	2.6	3.1	180.5
7476	2.9	3.2	183.3
5911	3.8	3.7	181.9
7950	3.6	3.6	185.0
6134	2.8	2.9	184.0
5868	3.0	3.1	188.2
3160	4.2	3.6	175.7
5872	3.7	3.5	188.0



Step 1: Open data editor → transform data from excel → name the variables

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	cr	prose	pcar	y	var	var	var	var
1	11484.00	2.30	3.50	158.10				
2	9348.00	2.50	2.90	173.40				
3	8429.00	3.10	4.10	165.30				
4	10079.00	2.90	3.60	172.90				
5	9240.00	2.70	3.20	178.50				
6	8862.00	2.80	3.70	198.60				
7	6216.00	3.60	3.80	186.30				
8	8253.00	3.20	3.50	189.00				
9	8038.00	2.60	3.10	180.50				
10	7476.00	2.90	3.20	183.30				
11	5911.00	3.80	3.70	181.90				
12	7950.00	3.60	3.60	185.00				
13	6134.00	2.80	2.90	184.00				
14	5868.00	3.00	3.10	188.20				
15	3160.00	4.20	3.60	175.70				
16	5872.00	3.70	3.50	188.00				
17								
18								
19								
20								
21								
22								
23								

Data View Variable View

Windows taskbar: Type here to search, Internet Explorer, File Explorer, Camera, Google Chrome



Step 2: Taking Logarithm of the variables *Transform* → Compute Variables

*Untitled1 [DataSet0] - PASW Statistics Data Editor

The screenshot shows the PASW Statistics Data Editor interface. The 'Transform' menu is open, and 'Compute Variable...' is selected. The data table has the following content:

	cr				
1	11484.00				
2	9348.00				
3	8429.00				
4	10079.00				
5	9240.00				
6	8862.00				
7	6216.00				
8	8253.00				
9	8038.00				
10	7476.00				
11	5911.00				
12	7950.00				
13	6134.00				
14	5868.00	3.00	3.10	188.20	
15	3160.00	4.20	3.60	175.70	
16	5872.00	3.70	3.50	188.00	
17					
18					
19					
20					
21					
22					
23					



Step 2: Name the target variable (i. e., new variable)

Taking natural log \rightarrow LN (original variable name)

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	cr	prose	pcar
1	11484.00	2.30	3.50
2	9348.00	2.50	2.90
3	8429.00	3.10	4.10
4	10079.00	2.90	3.60
5	9240.00	2.70	3.20
6	8862.00	2.80	3.70
7	6216.00	3.60	3.80
8	8253.00	3.20	3.50
9	8038.00	2.60	3.10
10	7476.00	2.90	3.20
11	5911.00	3.80	3.70
12	7950.00	3.60	3.60
13	6134.00	2.80	2.90
14	5868.00	3.00	3.10
15	3160.00	4.20	3.60
16	5872.00	3.70	3.50
17			
18			
19			
20			
21			
22			
23			

Compute Variable

Target Variable: = Numeric Expression:

Type & Label...

cr
prose
pcar
y

Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation

Functions and Special Variables:

If... (optional case selection condition)

OK Paste Reset Cancel Help

Data View Variable View

PASW Stat

Type here to search



Step 2: New variables with original the variables

*Untitled1 [DataSet0] - PASW Statistics Data Editor

	cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	
1	11484.00	2.30	3.50	158.10	9.35	.83	1.25	5.06	
2	9348.00	2.50	2.90	173.40	9.14	.92	1.06	5.16	
3	8429.00	3.10	4.10	165.30	9.04	1.13	1.41	5.11	
4	10079.00	2.90	3.60	172.90	9.22	1.06	1.28	5.15	
5	9240.00	2.70	3.20	178.50	9.13	.99	1.16	5.18	
6	8862.00	2.80	3.70	198.60	9.09	1.03	1.31	5.29	
7	6216.00	3.60	3.80	186.30	8.73	1.28	1.34	5.23	
8	8253.00	3.20	3.50	189.00	9.02	1.16	1.25	5.24	
9	8038.00	2.60	3.10	180.50	8.99	.96	1.13	5.20	
10	7476.00	2.90	3.20	183.30	8.92	1.06	1.16	5.21	
11	5911.00	3.80	3.70	181.90	8.68	1.34	1.31	5.20	
12	7950.00	3.60	3.60	185.00	8.98	1.28	1.28	5.22	
13	6134.00	2.80	2.90	184.00	8.72	1.03	1.06	5.21	
14	5868.00	3.00	3.10	188.20	8.68	1.10	1.13	5.24	
15	3160.00	4.20	3.60	175.70	8.06	1.44	1.28	5.17	
16	5872.00	3.70	3.50	188.00	8.68	1.31	1.25	5.24	
17									
18									
19									
20									
21									
22									
23									

Data View

Variable View



Step 3: Estimation linear regression (Analyze → Regression → Linear)

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform **Analyze** Direct Marketing Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...

Incr Inprose Inpcar Iny var

	Incr	Inprose	Inpcar	Iny	var
1	9.35	.83	1.25	5.06	
2	9.14	.92	1.06	5.16	
3	9.04	1.13	1.41	5.11	
4	9.22	1.06	1.28	5.15	
5			1.16	5.18	
6			1.31	5.29	
7			1.34	5.23	
8			1.25	5.24	
9			1.13	5.20	
10			1.16	5.21	
11			1.31	5.20	
12			1.28	5.22	
13			1.06	5.21	
14			1.13	5.24	
15			1.28	5.17	
16			1.25	5.24	
17					
18					
19					
20					
21					
22					
23					

Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...
Multinomial Logistic...
Ordinal...
Probit...
Nonlinear...
Weight Estimation...
2-Stage Least Squares...
Optimal Scaling (CATREG)...

Data View Variable View



Step 3: Estimation linear regression (Selection of Dependent and Independent Variables)

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help



	cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	var	var
1	11484.00	2.30	3.50	158.10						
2	9348.00	2.50	2.90	173.40						
3	8429.00	3.10	4.10	165.30						
4	10079.00	2.90	3.60	172.90						
5	9240.00	2.70	3.20	178.50						
6	8862.00	2.80	3.70	198.60						
7	6216.00	3.60	3.80	186.30						
8	8253.00	3.20	3.50	189.00						
9	8038.00	2.60	3.10	180.50						
10	7476.00	2.90	3.20	183.30						
11	5911.00	3.80	3.70	181.90						
12	7950.00	3.60	3.60	185.00						
13	6134.00	2.80	2.90	184.00						
14	5868.00	3.00	3.10	188.20						
15	3160.00	4.20	3.60	175.70						
16	5872.00	3.70	3.50	188.00						
17										
18										
19										
20										
21										
22										
23										

Linear Regression

Dependent: cr

Block 1 of 1

Independent(s): prose, pcar, y

Method: Enter

Selection Variable: Rule...

Case Labels:

WLS Weight:

OK Paste Reset Cancel Help

Statistics... Plots... Save... Options... Bootstrap...

Data View Variable View



Step 3: Estimation linear regression (Selection of Statistics)

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	var	var
1	11484.00	2.30	3.50	158.10	9.35	.83	1.25	5.06		
2	9						1.06	5.16		
3	8						1.41	5.11		
4	10						1.28	5.15		
5	9						1.16	5.18		
6	8						1.31	5.29		
7	6						1.34	5.23		
8	8						1.25	5.24		
9	8						1.13	5.20		
10	7						1.16	5.21		
11	5									
12	7									
13	6									
14	5									
15	3									
16	5									
17										
18										
19										
20										
21										
22										
23										

Linear Regression

Dependent: cr

Block 1 of 1

Independent(s): prose, pcar, y

Method: Enter

Selection Variable: [] Rule...

Case Labels: []

WLS Weight: []

OK Paste Reset Cancel Help

Linear Regression: Statistics

Regression Coefficients

- Estimates
- Confidence intervals
- Level(%): 95
- 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

Data View Variable View



Step 3: Estimation linear regression (Save Residuals)

SPSS Statistics Data Editor - PASW Statistics Data Editor

	cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	var	var
1	11484.00	2.30	3.50	158.10	9.35	.83	1.25	5.06		

Linear Regression

Dependent: cr

Block 1 of 1

Independent(s): prose, pcar, y

Method: Enter

Selection Variable: [] Rule...

Case Labels: []

WLS Weight: []

Buttons: OK, Paste, Reset, Cancel, Help

Linear Regression: Save

Predicted Values

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

Residuals

- Unstandardized
- Standardized
- Studentized
- Deleted
- Studentized deleted

Distances

- Mahalanobis
- Cook's
- Leverage values

Influence Statistics

- DfBeta(s)
- Standardized DfBeta(s)
- DfFit
- Standardized DfFit
- Covariance ratio

Prediction Intervals

- Mean Individual
- Confidence Interval: 95 %

Coefficient statistics

- Create coefficient statistics
- Create a new dataset
 - Dataset name: []
- Write a new data file
 - File... []

Export model information to XML file []

Step 4: Estimation linear regression (Output for interpretation)

*Output1 [Document1] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help



- Output
 - Log
 - Regression
 - Title
 - Notes
 - Active D
 - Variable
 - Model S
 - ANOVA
 - Coeffici
 - Residu:
 - Log
 - Regression
 - Title
 - Notes
 - Active D
 - Variable
 - Model S
 - ANOVA
 - Coeffici
 - Collinez

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.895 ^a	.802	.752	1016.64051

a. Predictors: (Constant), y, pcar, prose

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.019E7	3	1.673E7	16.188	.000 ^a
	Residual	1.240E7	12	1033557.917		
	Total	6.260E7	15			

a. Predictors: (Constant), y, pcar, prose

b. Dependent Variable: cr

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	12741.856	6101.129		2.088	.059		
	prose	-3809.243	618.537	-.992	-6.158	.000	.636	1.573
	pcar	2783.078	936.369	.461	2.972	.012	.686	1.457
	y	-15.682	29.178	-.076	-.537	.601	.832	1.202

a. Dependent Variable: cr

Collinearity Diagnostics^a

Model	Dimension	Condition Index	Condition Number	Maximum Variance Inflation Factor
1	4	1.000	1.000	1.000



Step 5: Test of Normality of Residuals

(Analyze → Descriptive Statistics → Explore → Plots)

The screenshot displays the SPSS interface with the 'Explore' dialog box open. The 'Dependent List' contains 'Unstandardized Residuals'. The 'Display' section has 'Both' selected. The 'Explore: Plots' sub-dialog box is also open, with 'Normality plots with tests' checked. The 'Descriptive' section has 'Stem-and-leaf' and 'Histogram' checked. The 'Spread vs Level with Levene Test' section has 'None' selected.

cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	RES_1
11484.00							1.25	241.90021
9348.00							1.06	777.52478
8429.00							1.41	-1322.64394
10079.00							1.28	1076.22682
9240.00							1.16	676.42651
8862.00							1.31	-396.98694
6216.00							1.34	-466.78432
8253.00								923.78213
8038.00								-596.82675
7476.00								-250.45292
5911.00								199.37288
7950.00								1803.44510
6134.00								-1127.47676
5868.00								-1122.38073
3160.00								-846.84823
5872.00	3.70	3.50	188.00					431.72216



Step 5: Test of Normality of Residuals : Output for interpretation

*Output1 [Document1] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help

[Dataset1] C:\Users\Dr.P.Das\Desktop\Estimation demand func of rose.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Unstandardized Residual	16	100.0%	0	.0%	16	100.0%

Descriptives

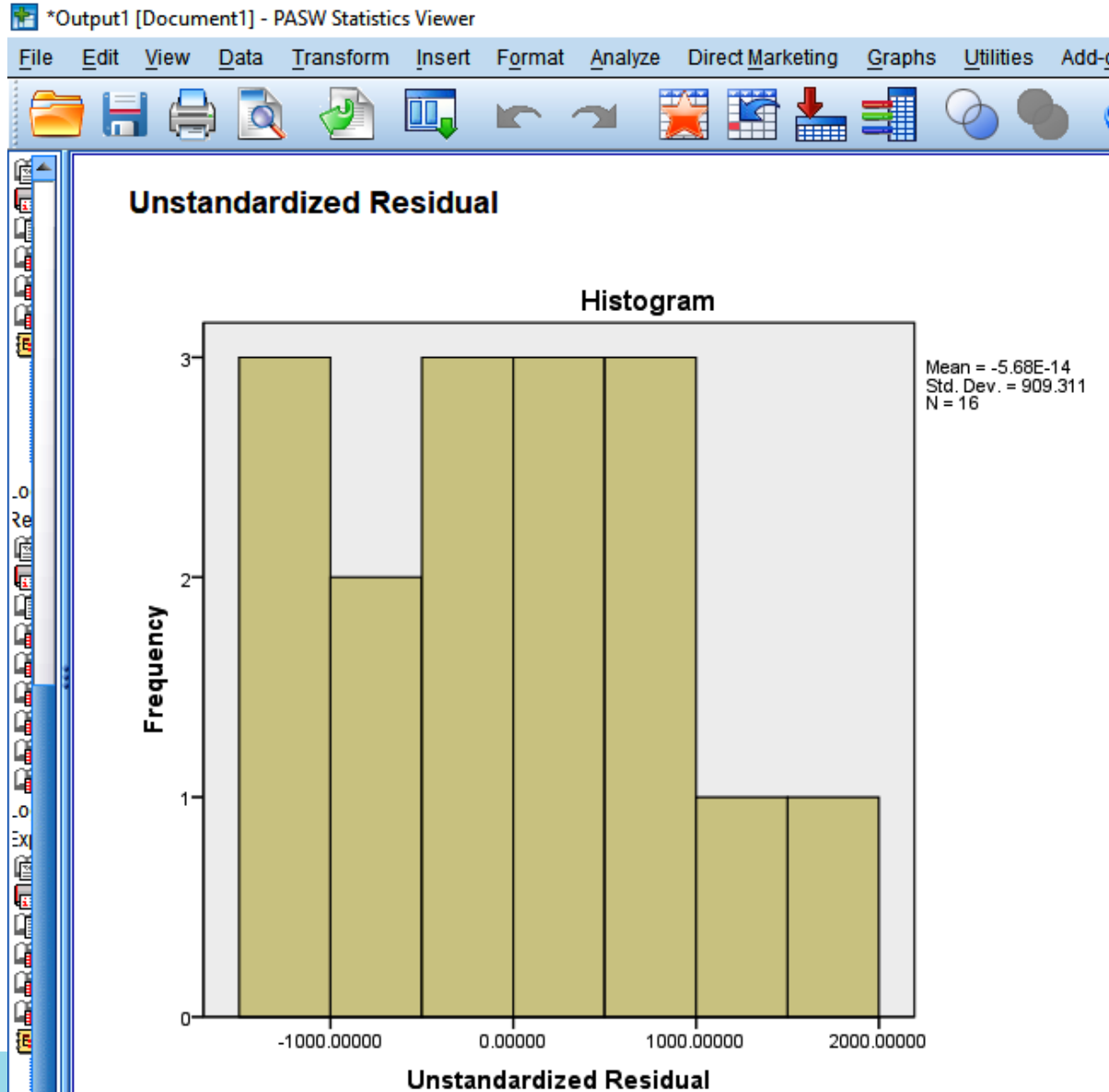
				Statistic	Std. Error
Unstandardized Residual	Mean			.0000000	227.3277278
	95% Confidence Interval for Mean	Lower Bound		-484.5375822	
		Upper Bound		484.5375822	
	5% Trimmed Mean			-26.7111755	
	Median			-25.5400206	
	Variance			826846.333	
	Std. Deviation			909.3109113	
	Minimum			-1322.64394	
	Maximum			1803.44510	
	Range			3126.08904	
	Interquartile Range			1536.59307	
	Skewness			.266	.564
	Kurtosis			-.759	1.091

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.109	16	.200*	.964	16	.733



Step 5: Test of Normality of Residuals : Output for interpretation



Step 6: Estimation Log linear regression (Analyze → Regression → Linear)

*Estimaiton demand func of rose.sav [DataSet1] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	cr	prose	pcar	y	Incr	Inprose	Inpcar	Iny	RES_1	v
1	11484.00	2.30	3.50	158.10	9.35	.83	1.25	5.06	241.90021	
2							1.06	5.16	777.52478	
3							1.41	5.11	-1322.64394	
4							1.28	5.15	1076.22682	
5							1.16	5.18	676.42651	
6							1.31	5.29	-396.98694	
7							1.34	5.23	-466.78432	
8							1.25	5.24	923.78213	
9							1.13	5.20	-596.82675	
10							1.16	5.21	-250.45292	

Linear Regression

Dependent: **Incr**

Block 1 of 1

Independent(s): **Inprose, Inpcar, Iny**

Method: **Enter**

Selection Variable:

Case Labels:

WLS Weight:

OK Paste Reset Cancel Help

Linear Regression: Save

Predicted Values

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

Residuals

- Unstandardized
- Standardized
- Studentized
- Deleted
- Studentized deleted

Distances

- Mahalanobis
- Cook's
- Leverage values

Influence Statistics

- DfBeta(s)
- Standardized DfBeta(s)
- DfFit
- Standardized DfFit
- Covariance ratio

Prediction Intervals

- Mean Individual

Confidence Interval: %

Step 7: Estimation Log linear regression (Output for interpretation)

*Output1 [Document1] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help



Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.875 ^a	.766	.707	.16599

a. Predictors: (Constant), Iny, Inpcar, Inprose

b. Dependent Variable: Incr

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.082	3	.361	13.090	.000 ^a
	Residual	.331	12	.028		
	Total	1.413	15			

a. Predictors: (Constant), Iny, Inpcar, Inprose

b. Dependent Variable: Incr

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5.635	4.637		1.215	.248		
	Inprose	-1.960	.335	-.1075	-5.847	.000	.578	1.732
	Inpcar	1.573	.535	.510	2.941	.012	.649	1.540
	Iny	.679	.876	.123	.775	.453	.772	1.296

a. Dependent Variable: Incr

Collinearity Diagnostics^a

Model	Dimension	Condition	Variance Proportions			



Step 8: Test of Normality for Residuals of Log Linear Model (Analyze → Descriptive Statistics → Explore → Plots)

*Estimaition demand func of rose.sav [DataSet1] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1 : RES_2 -.06251230386604

	Incr	Inprose	Inpcar	Iny	RES_1	RES_2
1					41.90021	-.06251
2					77.52478	.12816
3					22.64394	-.06597
4					76.22682	.15615
5					76.42651	.09283
6					96.98694	-.17849
7					66.78432	-.03919
8					22.78212	.13303
9						.7812
10						.0300
11						.7463
12						.9666
13						.1137
14						.4073
15						.8884
16	8.68	1.31	1.25			.8076
17						
18						
19						
20						
21						
22						
23						

Explore

Dependent List: Unstandardized Re...

Factor List:

Label Cases by:

Display: Both Statistics Plots

OK Paste Reset Cancel

Explore: Plots

Boxplots: Factor levels together Dependents together None

Descriptive: Stem-and-leaf Histogram

Normality plots with tests

Spread vs Level with Levene Test: None Power estimation Transformed Power: Natural log Untransformed

Continue Cancel Help

Data View Variable View



Step 9: Test of Normality of Residuals Log linear Model: Output for interpretation

*Output1 [Document1] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons W

[Dataset1] C:\Users\Dr.P.Das\Desktop\Estimation demand func of Iose.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Unstandardized Residual	16	100.0%	0	.0%	16	100.0%

Descriptives

		Statistic	Std. Error	
Unstandardized Residual	Mean	.0000000	.03711600	
	95% Confidence Interval for Mean	Lower Bound	-.0791109	
		Upper Bound	.0791109	
	5% Trimmed Mean	-.0004342		
	Median	-.0180952		
	Variance	.022		
	Std. Deviation	.14846402		
	Minimum	-.28884		
	Maximum	.29666		
	Range	.58550		
	Interquartile Range	.22238		
	Skewness	.026	.564	
Kurtosis	-.086	1.091		

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.130	16	.200 [*]	.983	16	.982



Step 9: Test of Normality of Residuals Log linear Model: Output for interpretation

*Output1 [Document1] - PASW Statistics Viewer

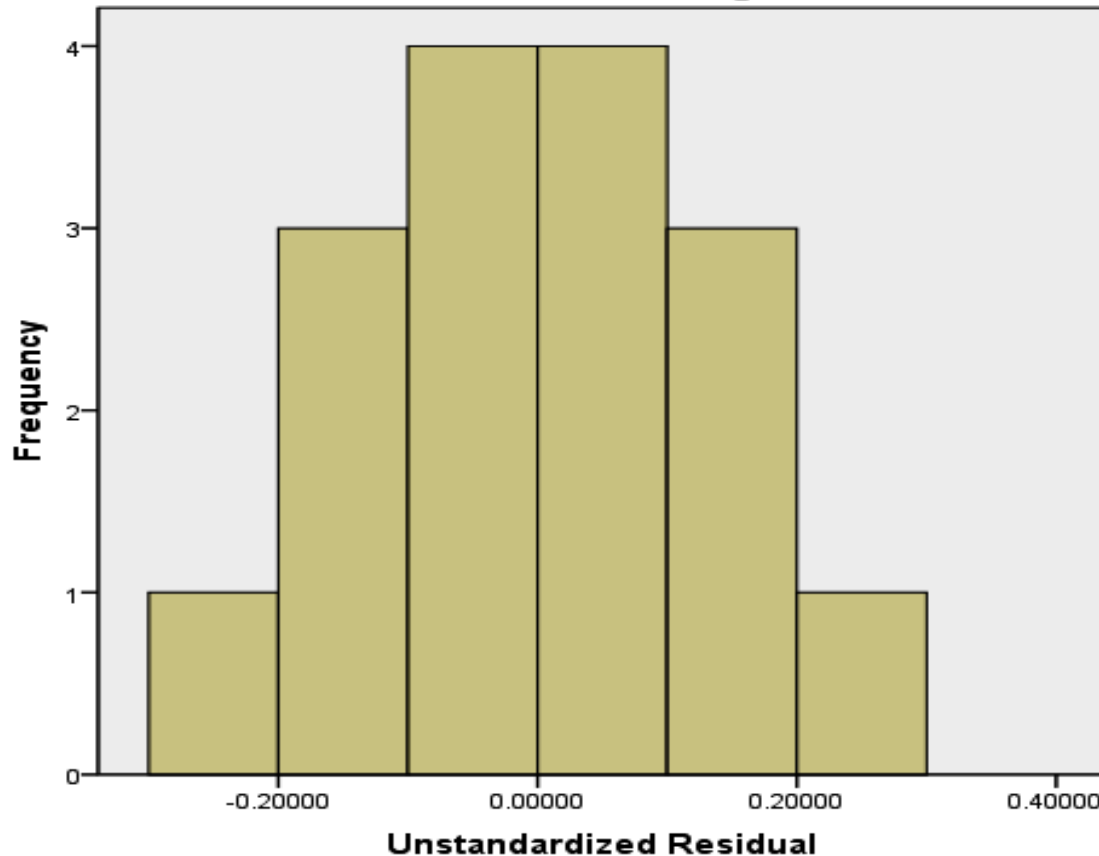
File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-c



*. This is a lower bound of the true significance.

Unstandardized Residual

Histogram



Mean = -9.65E-16
Std. Dev. = 0.148
N = 16



Assignment

1. Specify Linear Model and Log Linear Model for the following function:

$$D = f(P_1, P_2, Y)$$

2. Write down the steps of estimation procedure of linear regression in SPSS. Interpret the estimated results (as given in Slide 11).
3. How you can check the normality of the residuals of a linear model? Interpret the results of normality test as given in Slides 13 and 14).
4. Interpret the estimated results of log linear model (as given in Slide 16) and distinguish with the linear model.
5. Interpret the results of normality test of log linear model (as given in Slides 18 and 19).



Thanks

