NPTEL Project on

Econometric Modelling

Prof. Rudra P. Pradhan

VINOD GUPTA SCHOOL OF MANAGEMENT

INDIAN INSTITUTE OF TECHNOLOGY, KHARAGPUR

==

Q. 1

Samples	Mileage	Octane rating
1	32	90
2	30	90
3	35	90
4	33	90
5	35	90
6	34	90
7	29	90
8	32	90
9	36	90
10	34	90
11	35	95
12	38	95
13	37	95
14	40	95
15	41	95
16	35	95
17	37	95
18	41	95
19	36	95
20	40	95
21	44	100
22	46	100
23	47	100
24	47	100
25	46	100
26	43	100
27	47	100
28	45	100
29	48	100
30	47	100

1. Examine the various descriptive statistics.
2. Fit a bivariate regression modelling and point out the significance of parameters.
3. Find out the correlation coefficient? Is it different to coefficient of determination?
4. Prepare the ANOVA table. What is its significance here?

Answers:

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
VAR00001	30	19.00	29.00	48.00	39.0000	5.80725	.128	.427	-1.283	.833
VAR00002	30	10.00	90.00	100.00	95.0000	4.15227	.000	.427	-1.554	.833

Regression Estimates (Coefficients)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-84.500	9.268		-9.117	.000
	VAR00002	1.300	.097	.930	13.338	.000

a. Dependent Variable: VAR00001

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	845.000	1	845.000	177.895	.000[a]
	Residual	133.000	28	4.750		
	Total	978.000	29			

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.930[a]	.864	.859	2.17945

Correlations

		VAR00001	VAR00002
VAR00001	Pearson Correlation	1	.930**
	Sig. (2-tailed)		.000
	N	30	30
VAR00002	Pearson Correlation	.930**	1
	Sig. (2-tailed)	.000	
	N	30	30

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

Q. 2

Sample Units	Sales	Price	Net Income
1994	80	5.00	2620
1995	86	4.87	2733
1996	93	4.86	2898
1997	99	4.79	3056
1998	106	4.79	3271
1999	107	4.87	3479
2000	109	5.01	3736
2001	110	5.31	3868
2002	111	5.55	4016
2003	113	5.72	4152
2004	110	5.74	4336
2005	112	5.59	4477
2006	131	5.50	4619
2007	136	5.48	4764
2008	137	5.47	4802
2009	139	5.49	4916

A company is the sole owner of a product says **'XYZ'** and they are in the process of expansion plan for the year **2015**. The MD asked his accountant to provide some information about the **'XYZ'**. That is with respect to sales, price and net income (see in the above table) and engaged an analyst to suggest few strategic decisions about their expansion plan. The analyst is now in the process. Please answer the followings:

5. Prepare a correlation matrix between sales, price and income. And examine their statistical significance level.
6. Report the values of partial correlation coefficients and multiple correlation coefficients. Are they also statistically significant?
7. Fit a bivariate regression model with respect to net income and sales. Give the model interpretation.
8. Fit a multivariate regression model with respect to sales, price and net income. Give the model interpretation.
9. Is there any difference between these two models? If so, then what grounds and which conditions?
10. Between the two models, which one is the best and how do you justify the same.

11. Using multivariate model, forecast the net income of **'XYZ'** for the year **2015**.
12. Let us assume multivariate model is better fitted than bivariate. Now for multivariate model, what are the basic problems you can address further? This should be elaborated logically.
13. What are the basic assumptions you need to use for getting the model better fitted?
14. Which particular variable is more consistent than others? Give suitable logical reasoning's for the same.
15. On the basis of above statistical analysis, please suggest various strategic decisions about the product **'XYZ'** for its expansion plan in **2015**.

Answers:

Correlations

		Sales	Price	Income
Sales	Pearson Correlation	1	.625**	.939**
	Sig. (2-tailed)		.010	.000
	N	16	16	16
Price	Pearson Correlation	.625**	1	.831**
	Sig. (2-tailed)	.010		.000
	N	16	16	16
Income	Pearson Correlation	.939**	.831**	1
	Sig. (2-tailed)	.000	.000	
	N	16	16	16

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

Partial Correlations

Control Variables			Sales	Income
Price	Sales	Correlation	1.000	.966
		Significance (2-tailed)	.	.000
		df	0	13
	Income	Correlation	.966	1.000
		Significance (2-tailed)	.000	.
		df	13	0

Bivariate Model:

Regression Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	29.352	8.161		3.597	.003
	Income	.021	.002	.939	10.216	.000

a. Dependent Variable: Sales

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4047.453	1	4047.453	104.357	.000[a]
	Residual	542.985	14	38.785		
	Total	4590.438	15			

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.939[a]	.882	.873	6.22773

Multivariate Model

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	123.472	19.399		6.365	.000
	Income	.031	.002	1.356	13.551	.000
	Price	-24.845	4.952	-.502	-5.017	.000

a. Dependent Variable: Sales

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4405.521	2	2202.760	154.858	.000[a]
	Residual	184.917	13	14.224		
	Total	4590.438	15			

a. Predictors: (Constant), Price, Income

b. Dependent Variable: Sales

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.980[a]	.960	.954	3.77152

a. Predictors: (Constant), Price, Income

Q. 3

Year	X1	X2	Y
1982	-0.4	6.7	53
1983	0.4	2.1	31.2
1984	2.9	1.8	3.7
1985	3.0	-0.4	-13.8
1986	1.7	6.0	41.7
1987	1.5	2.1	10.5
1988	1.8	2.6	-1.3
1989	0.8	5.8	26.1
1990	1.8	4.1	-10.5
1991	1.6	5.3	21.2
1992	1.1	6.0	15.5
1993	2.3	6.1	10.2
1994	3.2	2.7	-13.3
1995	2.7	4.6	21.3
1996	4.3	2.8	6.8
1997	5.0	-0.2	-13.5
1998	4.4	3.4	-0.4
1999	3.8	5.7	10.5
2000	3.6	5.8	15.4
2001	7.9	-0.6	-22.6
2002	10.8	-1.2	-37.3
2003	6.0	5.4	31.2
2004	4.7	5.5	19.1
2005	5.9	5.1	-13.1
2006	7.9	2.8	-1.3
2007	9.8	-0.3	8.6
2008	10.2	2.6	-22.2
2009	7.3	-1.9	-12.2

Where, Y is stock return, X1 is rate of inflation and X2 is GDP. Answer the followings:

1.1 Fit a bivariate regression model (Y on X1 and Y on X2) and compare the same with multivariate model.
1.2 Justify the best fitted model with respect to the presence of autocorrelation.
1.3 Forecast the stock price for the year 2015. You can assume that rate of inflation and GDP will grow at 10% from the base year 2009.
1.4 Examine, whether there is any multicolinearity? If so, justify some solution.
1.5 Find out RMSE for all the three above models.

Answers:

BVM (Y on X1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	24.031	5.241		4.585	.000
	VAR00006	-4.382	1.023	-.643	-4.283	.000

a. Dependent Variable: VAR00008

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4878.139	1	4878.139	18.344	.000[a]
	Residual	6913.893	26	265.919		
	Total	11792.033	27			

a. Predictors: (Constant), VAR00006

b. Dependent Variable: VAR00008

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.643[a]	.414	.391	16.30702

a. Predictors: (Constant), VAR00006

BVM (Y on X2)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-12.175	4.726		-2.576	.016
	VAR00007	5.591	1.152	.689	4.852	.000

a. Dependent Variable: VAR00008

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5603.913	1	5603.913	23.545	.000[a]
	Residual	6188.119	26	238.005		
	Total	11792.033	27			

a. Predictors: (Constant), VAR00007

b. Dependent Variable: VAR00008

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.689[a]	.475	.455	15.42740

a. Predictors: (Constant), VAR00007

Multicolinearity:

Correlations

		VAR00006	VAR00007	VAR00008
VAR00006	Pearson Correlation	1	-.567**	-.643**
	Sig. (2-tailed)		.002	.000
	N	28	28	28
VAR00007	Pearson Correlation	-.567**	1	.689**
	Sig. (2-tailed)	.002		.000
	N	28	28	28
VAR00008	Pearson Correlation	-.643**	.689**	1
	Sig. (2-tailed)	.000	.000	
	N	28	28	28

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

Multivariate:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.839	8.143		.471	.641
	VAR00007	3.880	1.293	.478	3.000	.006

VAR00006	-2.532	1.087	-.372	-2.330	.028

a. Dependent Variable: VAR00008 (Y)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6708.174	2	3354.087	16.494	.000[a]
	Residual	5083.859	25	203.354		
	Total	11792.033	27			

a. Predictors: (Constant), VAR00006, VAR00007

b. Dependent Variable: VAR00008

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.754[a]	.569	.534	14.26024

a. Predictors: (Constant), VAR00006, VAR00007

Q. 4

Set A		Set B		Set C		Set D	
X	**Y**	**X**	**Y**	**X**	**Y**	**X**	**Y**
10	**8.04**	**10**	**9.14**	**10**	**7.46**	**8**	**6.58**
8	**6.95**	**8**	**8.14**	**8**	**6.77**	**8**	**5.76**
13	**7.58**	**13**	**8.74**	**13**	**12.74**	**8**	**7.71**
9	**8.81**	**9**	**8.77**	**9**	**7.11**	**8**	**8.84**
11	**8.33**	**11**	**9.26**	**11**	**7.81**	**8**	**8.47**
14	**9.96**	**14**	**8.10**	**14**	**8.84**	**8**	**7.04**
6	**7.24**	**6**	**6.13**	**6**	**6.08**	**8**	**5.25**
4	**4.26**	**4**	**3.10**	**4**	**5.39**	**8**	**12.5**
12	**10.84**	**12**	**9.13**	**12**	**8.15**	**8**	**5.56**
7	**4.82**	**7**	**7.26**	**7**	**6.42**	**8**	**7.91**
5	**5.68**	**5**	**4.74**	**5**	**5.73**	**8**	**6.89**

1. Determine the best fitting equation.
2. Test it for statistical significance.
3. Calculate the residuals.
4. Prepare the ANOVA table.

Answers:

Set A:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.000	1.125		2.667	.026
	VAR00010	.500	.118	.816	4.241	.002

a. Dependent Variable: VAR00009

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27.510	1	27.510	17.990	.002[a]
	Residual	13.763	9	1.529		
	Total	41.273	10			

a. Predictors: (Constant), VAR00010

b. Dependent Variable: VAR00009

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.816[a]	.667	.629	1.23660

a. Predictors: (Constant), VAR00010

Set B:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.995	2.435		-.408	.692
	VAR00012	1.332	.314	.816	4.239	.002

a. Dependent Variable: VAR00011

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	73.287	1	73.287	17.966	.002[a]
	Residual	36.713	9	4.079		
	Total	110.000	10			

a. Predictors: (Constant), VAR00012

b. Dependent Variable: VAR00011

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.816[a]	.666	.629	2.01972

a. Predictors: (Constant), VAR00012

Set C:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.000	2.436		-.411	.691
	VAR00014	1.333	.315	.816	4.239	.002

a. Dependent Variable: VAR00013

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	73.296	1	73.296	17.972	.002[a]
	Residual	36.704	9	4.078		
	Total	110.000	10			

a. Predictors: (Constant), VAR00014

b. Dependent Variable: VAR00013

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.816[a]	.666	.629	2.01947

a. Predictors: (Constant), VAR00014

Set D:
No results, since dependent is same for all X.

Q. 5

X_0	X1	X2	Y
1	1	8	6
1	4	2	8
1	9	-8	1
1	11	-10	0
1	3	6	5
1	8	-6	3
1	5	0	2
1	10	-12	-4
1	2	4	10
1	7	-2	-3
1	6	-4	5

1. Write out the analysis of variance
2. Determine the overall fitness of the model
3. Examine, if any multicolinearity is present in the model.

Answers:

Correlations

		VAR00003	VAR00004
VAR00003	Pearson Correlation	1	-.973**
	Sig. (2-tailed)		.000
	N	11	11
VAR00004	Pearson Correlation	-.973**	1
	Sig. (2-tailed)	.000	
	N	11	11

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

Regression Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	14.000	6.095		2.297	.051
	VAR00003	-2.000	1.198	-1.522	-1.669	.134
	VAR00004	-.500	.599	-.761	-.834	.428

a. Dependent Variable: VAR00005

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	122.000	2	61.000	7.176	.016[a]
	Residual	68.000	8	8.500		
	Total	190.000	10			

a. Predictors: (Constant), VAR00004, VAR00003

b. Dependent Variable: VAR00005

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.801[a]	.642	.553	2.91548

a. Predictors: (Constant), VAR00004, VAR00003

Q. 6

Items	Y	X1	X2
1	66	38	47.5
2	43	41	21.3
3	36	34	36.5
4	23	35	18
5	22	31	29.5
6	14	34	14.2
7	12	29	21.0
8	7.6	32	10.0

1. Fit the Trivariate econometric model.
2. Do you find any difference with bivariate model?
3. Is the overall regression model significant?
4. Examine the nature of heteroskedasticity in the Trivariate modelling.

Answers:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-94.552	9.963		-9.490	.000
	VAR00007	2.802	.301	.552	9.308	.000
	VAR00008	1.073	.093	.683	11.505	.000

a. Dependent Variable: VAR00006

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2618.979	2	1309.490	151.699	.000[a]
	Residual	43.161	5	8.632		
	Total	2662.140	7			

a. Predictors: (Constant), VAR00008, VAR00007

b. Dependent Variable: VAR00006

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.992[a]	.984	.977	2.93805

a. Predictors: (Constant), VAR00008, VAR00007

Bivariate:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-101.406	47.588		-2.131	.077
	VAR00007	3.777	1.382	.745	2.733	.034

a. Dependent Variable: VAR00006

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1476.356	1	1476.356	7.470	.034[a]
	Residual	1185.784	6	197.631		
	Total	2662.140	7			

a. Predictors: (Constant), VAR00007

b. Dependent Variable: VAR00006

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.745[a]	.555	.480	14.05812

a. Predictors: (Constant), VAR00007

Q. 7

Y	X1	X2	X_1X_2
1	610	0	0
1	950	0	0
1	720	0	0
1	840	0	0
1	980	0	0
1	530	0	0
1	680	0	0
1	540	0	0
1	890	0	0
1	730	0	0
1	670	1	670
1	770	1	770
1	880	1	880
1	1000	1	1000
1	760	1	760
1	590	1	590
1	910	1	910
1	650	1	650
1	810	1	810
1	500	1	500

1. Fit a model "$Y = a + b_1X_1+b_2X_2+b_3X_1X_2$
2. Examine the significance of interactive effect.
3. Is the overall regression model significant?
4. Prepare the ANOVA table.

Answer:

The problem is inconsistent, as all the items of DV are equal.

Q. 8

Year	X_1	X_2	X_3	X_4	X_5	X_6	Y
1	0.447	4.68	4.87	-1	8.4	8.4	0.829
2	0.146	5.19	4.50	-1	6.5	6.5	1.11
3	0.146	4.82	4.70	-1	7.9	7.9	1.17
4	0.518	4.85	4.76	-1	8.3	8.3	1.1
5	0.230	4.86	4.95	-1	8.4	8.4	0.92
6	0.462	5.16	4.45	-1	7.4	7.4	1.02
7	0.568	4.82	5.05	-1	6.8	6.8	0.862
8	0.230	4.86	4.70	-1	8.6	8.6	1.103
9	-0.036	4.78	4.84	-1	6.7	6.7	1.1
10	-0.167	5.16	4.76	-1	7.7	7.7	1.31
11	0.778	4.57	4.82	-1	7.4	7.4	0.55
12	0.633	4.61	4.65	-1	6.7	6.7	0.85
13	-0.22	5.07	5.10	-1	7.5	7.5	1.42
14	0.253	4.66	5.09	-1	8.2	8.2	1.07
15	0.778	5.42	4.41	-1	5.8	5.8	0.88
16	0.643	5.01	4.74	-1	7.1	7.1	1.09
17	1.945	4.97	4.66	1	6.5	6.5	-0.119
18	1.792	4.01	4.72	1	8.0	8.0	0.13
19	1.699	4.96	4.90	1	6.8	6.8	0.15
20	1.763	5.2	4.7	1	8.2	8.2	0.204
21	1.954	4.8	4.6	1	6.6	6.6	0.041
22	1.819	4.98	4.69	1	6.4	6.4	0.07
23	2.146	5.35	4.76	1	7.3	7.3	0.079
24	2.380	5.04	4.8	1	7.8	7.8	0.251
25	2.623	4.8	4.8	1	7.4	7.4	0.142
26	2.699	4.83	4.36	1	6.7	6.7	0.327
27	2.255	4.66	4.72	1	7.2	7.2	0.481
28	2.431	4.67	4.5	1	6.3	6.3	0.585
29	2.230	4.72	4.7	1	6.8	6.8	0.119
30	1.99	5.00	5.07	1	7.2	7.2	0.097
31	1.544	4.70	4.80	1	7.7	7.7	0.301

1. Fit an econometric model by using all these information.

2. Examine the presence of multicolinearity, if any?

3. Perform stepwise regression to avoid multicolinearity, if any.

Answers:

Excluded Variables[b]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	VAR00005	.[a]	.	.	.	.000

a. Predictors in the Model: (Constant), VAR00006, VAR00002, VAR00004, VAR00003, VAR00001

b. Dependent Variable: VAR00007

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.253	1.409		.889	.382
	VAR00001	-.169	.121	-.344	-1.397	.175
	VAR00002	.055	.141	.032	.391	.699
	VAR00003	-.211	.224	-.086	-.942	.355
	VAR00004	-.267	.108	-.584	-2.464	.021
	VAR00006	.040	.058	.063	.696	.493

a. Dependent Variable: VAR00007

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.496	5	1.099	28.443	.000[a]
	Residual	.966	25	.039		
	Total	6.462	30			

a. Predictors: (Constant), VAR00006, VAR00002, VAR00004, VAR00003, VAR00001

b. Dependent Variable: VAR00007

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.922[a]	.850	.821	.19658

a. Predictors: (Constant), VAR00006, VAR00002, VAR00004, VAR00003, VAR00001

Q. 9

Sample Units	Y	X
1990	3748	21777
1991	4010	22418
1992	3711	22308
1993	4004	23319
1994	4151	24180
1995	4569	24893
1996	4582	25310
1997	4697	25799
1998	4753	25886
1999	5062	26868
2000	5669	28314
2001	5628	29091
2002	5736	29450
2003	5946	30705
2004	6501	32372
2005	6549	33152
2006	6705	33764
2007	7104	34411
2008	7609	35429
2009	8100	36200

1. Fit the regression model and examine its significance level
2. Investigate the overall fitness of the model
3. Explore the existence o autocorrelation in the model

Answers:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-2467.027	244.710		-10.081	.000
	VAR00002	.280	.009	.992	32.738	.000

a. Dependent Variable: VAR00001

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.267E7	1	3.267E7	1071.780	.000[a]
	Residual	548677.886	18	30482.105		
	Total	3.322E7	19			

a. Predictors: (Constant), VAR00002

b. Dependent Variable: VAR00001

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.992[a]	.983	.983	174.59125	.885

a. Predictors: (Constant), VAR00002

b. Dependent Variable: VAR00001

Q. 10

$\mathbf{Y_1 = 4Y_2 - 3X_1 + u_1}$

$\mathbf{Y_2 = 2Y_3 + 2X_3 + u_2}$

$\mathbf{Y_3 = 2Y_1 - 3Y_2 - X_2 - X_3 + u_3}$

1. Establish for each equation whether it is exactly identified, over identified or under identified
2. Prepare the reduced form of the model

Answers:

First equation: Over

Second equation: Exact

Third equation: Over

Q. 11

$$\mathbf{D = a + b_1X_1 + b_2X_2 + U_1}$$

$$\mathbf{S = c + d\ X_1 + u_2}$$

D = S

1. Can we identify this model?
2. Use rank method to examine the third equation.

Answers:

First equation: Over

Second equation: Over

Q. 12

students	GRE Non verbal score	GRE Verbal score	Admitted to the programme 1 for yes; 0 for no
1	760	550	1
2	600	350	0
3	720	320	0
4	710	630	1
5	530	430	0
6	650	570	0
7	800	500	1
8	650	680	1
9	520	660	0
10	800	250	0
11	670	480	0
12	670	520	1
13	780	710	1

1. Use BCM to predict the probability of admission to the program based on quantitative and verbal score in the GRE.
2. Is this a reliable mode? If not, why

Answers:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-2.867	.838		-3.422	.007
	VAR00001	.003	.001	.561	2.976	.014
	VAR00002	.002	.001	.649	3.441	.006

a. Dependent Variable: VAR00003

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.098	2	1.049	9.258	.005[a]
	Residual	1.133	10	.113		
	Total	3.231	12			

a. Predictors: (Constant), VAR00002, VAR00001

b. Dependent Variable: VAR00003

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.806[a]	.649	.579	.33659

a. Predictors: (Constant), VAR00002, VAR00001

Q. 13

Sample Units	Y	X
1990	38596	59822
1991	43356	70242
1992	44840	72377
1993	47987	76122
1994	46443	73175
1995	51694	79516
1996	54063	87304
1997	55879	89052
1998	54201	87055
1999	59729	92097
2000	60827	94719
2001	61159	95580
2002	65662	101049
2003	68995	105463
2004	73682	111504
2005	80283	120929
2006	87187	136824
2007	90918	145681
2008	98794	156611
2009	105812	170400

Where, Y = Sales of Company I; X = Sales of Company 2

Answer the followings:

2.1 Fit an ARIMA Model of order (2, 2). Justify, whether model can be used for forecasting.

2.2 Fit a MAM of order 2 and justify its model feasibility.

2.3 Prepare an ARDL model of order 2. Is it better than that of previous two models? Justify.

2.4 Choose the best model (ARIMA, MAM and ARDL) on the basis of AIC, SIC and Rice Criterion.

2.5 Forecasting the sales of Y and X for the year 2015.

Answers:

```
    Exact AR(1) Inverse Interpolation Method Converged after 6
iterations
*********************************************************************
***********
 Dependent variable is SC1
 20 observations used for estimation from 1 to 20
*********************************************************************
***********
 Regressor                Coefficient       Standard Error         T-
Ratio[Prob]
 SC2                          .63329           .0060307
105.0111[.000]
*********************************************************************
***********
```

```
 R-Squared                       .99546   R-Bar-Squared
.99521
 S.E. of Regression             1333.4   F-Stat.    F(1,18)
3949.7[.000]
 Mean of Dependent Variable   64505.3   S.D. of Dependent Variable
19268.7
 Residual Sum of Squares     3.19E+07   Equation Log-likelihood
-171.4355
 Akaike Info. Criterion     -173.4355   Schwarz Bayesian Criterion
-174.4313
 DW-statistic                   1.7899
*******************************************************************
***********

             Parameters of the Autoregressive Error Specification
*******************************************************************
***********
 U=     .57481*U(-1)+E
   (    3.1414)[.006]
 T-ratio(s) based on asymptotic standard errors in brackets
 Log-likelihood ratio test of AR(1) versus OLS   CHI-SQ(1)=
6.7189[.010]
*******************************************************************
***********
```

Part 2:

```
Autoregressive Distributed Lag Estimates
            ARDL(1,1) selected based on Schwarz Bayesian Criterion
*******************************************************************
***********
 Dependent variable is SC1
 19 observations used for estimation from 2 to 20
*******************************************************************
***********
 Regressor              Coefficient       Standard Error         T-
Ratio[Prob]
 SC1(-1)                   .74086             .23194
3.1942[.006]
 SC2                       .51247            .078952
6.4909[.000]
 SC2(-1)                  -.34161             .13889             -
2.4595[.026]
*******************************************************************
 R-Squared                       .99565   R-Bar-Squared
.99511
 S.E. of Regression             1313.3   F-Stat.    F(2,16)
1832.2[.000]
 Mean of Dependent Variable   65869.0   S.D. of Dependent Variable
18779.1
 Residual Sum of Squares     2.77E+07   Equation Log-likelihood
-161.7530
 Akaike Info. Criterion     -164.7530   Schwarz Bayesian Criterion
-166.1696
 DW-statistic                   2.3855   Durbin's h-statistic
*NONE*
*******************************************************************
```

```
Testing for existence of a level relationship among the variables in
the ARDL model
*******************************************************************
 F-statistic  95% Lower Bound  95% Upper Bound  90% Lower Bound  90%
Upper Bound
    1.5773          3.6266          4.6806          2.6689
3.5746

 W-statistic  95% Lower Bound  95% Upper Bound  90% Lower Bound  90%
Upper Bound
    3.1547          7.2531          9.3612          5.3379
7.1491
*******************************************************************
 If the statistic lies between the bounds, the test is inconclusive.
If it is
 above the upper bound, the null hypothesis of no level effect is
rejected. If
 it is below the lower bound, the null hypothesis of no level effect
can't be
 rejected. The critical value bounds are computed by stochastic
simulations
 using 20000 replications.


                               Diagnostic Tests
*******************************************************************
*    Test Statistics  *        LM Version       *          F Version
*
*******************************************************************
*
* A:Serial Correlation*CHSQ(1)  =   2.6327[.105]*F(1,15)      =
2.4128[.141]*
*                     *                         *
*
* B:Functional Form   *CHSQ(1)  =   .59792[.439]*F(1,15)      =
.48738[.496]*
*                     *                         *
*
* C:Normality         *CHSQ(2)  =   1.3887[.499]*       Not
applicable        *
*                     *                         *
*
* D:Heteroscedasticity*CHSQ(1)  =   .44513[.505]*F(1,17)      =
.40783[.532]*
*******************************************************************
***********
   A:Lagrange multiplier test of residual serial correlation
   B:Ramsey's RESET test using the square of the fitted values
   C:Based on a test of skewness and kurtosis of residuals
```

Q. 14

Trend	Sales	Quarter
1	38596	1
2	43356	2
3	44840	3
4	47987	4
5	46443	1
6	51694	2
7	54063	3
8	55879	4
9	54201	1
10	59729	2
11	60827	3
12	61159	4
13	65662	1
14	68995	2
15	73682	3
16	80283	4
17	87187	1
18	90918	2
19	98794	3
20	105812	4
21	111504	1
22	120929	2
23	136824	3
24	145681	4
25	156611	1
26	170400	2
27	181150	3
28	183500	4

Using above information, answer the followings:

3.1 Prepare a trend analysis and examine, which quarter is very effective for sales forecasting.

3.2 Examine the stationarity of time series sales.

Answers:

```
ADF tests for variable SALES
       The Dickey-Fuller regressions include no intercept and no
trend
********************************************************************
 26 observations used in the estimation of all ADF regressions.
 Sample period from 3 to 28
********************************************************************
        Test Statistic          LL          AIC          SBC
HQC
 DF          9.3652         -247.1431   -248.1431   -248.7722   -
248.3243
 ADF(1)      3.1392         -247.1319   -249.1319   -250.3899   -
249.4941
********************************************************************
 LL  = Maximized log-likelihood      AIC = Akaike Information
Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

```
                         ADF tests for variable SALES
      The Dickey-Fuller regressions include an intercept but not a
trend
*******************************************************************
 26 observations used in the estimation of all ADF regressions.
 Sample period from 3 to 28
*******************************************************************
        Test Statistic          LL         AIC         SBC
HQC
 DF         4.3104         -247.0279   -249.0279   -250.2860   -
249.3902
 ADF(1)     2.2904         -247.0273   -250.0273   -251.9144   -
250.5707
*******************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =
-2.9798

 LL  = Maximized log-likelihood      AIC = Akaike Information
Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion


                         ADF tests for variable SALES
    The Dickey-Fuller regressions include an intercept and a linear
trend
*******************************************************************
 26 observations used in the estimation of all ADF regressions.
 Sample period from 3 to 28
*******************************************************************
        Test Statistic          LL         AIC         SBC
HQC
 DF        -.36052         -244.9435   -247.9435   -249.8307   -
248.4869
 ADF(1)    -.34796         -244.9391   -248.9391   -251.4553   -
249.6637
*******************************************************************
 95% published asymptotic critical value corresponding to ADF(0) =
-3.5943

 LL  = Maximized log-likelihood      AIC = Akaike Information
Criterion
 SBC = Schwarz Bayesian Criterion    HQC = Hannan-Quinn Criterion
```

Q. 15

$\Delta X_t = 31.03 - 0.19 X_{t-1}$

se	(12.50)	(0.08)
t		[-2.35]

Using above information, answer the followings:

1. Is the model attains stationary? How do you know?

2. What statistics do you use to know the unit root?

Answer:

It is not stationary; since the calculated statistics is lower than the tabulated value.

Statistics Use: DF

Q. 16

$\Delta^2 X_t = 4.76 - 13\ \Delta X_{t-1} + 0.313\ \Delta^2 X_{t-1}$

se	(5.06)	(0.236)	(0.163)
t		[-5.89]	

Using above information, answer the followings:

1. Is the model attains stationary? Is there any difference to question number 14?

2. Could you add an extra lag in this model? How do you decide for the same?

Answer:

It is not stationary; since the calculated statistics is lower than the tabulated value.

Yes, it is different to question 14. That is with respect to the use of second order difference equation.

If data is available, we can.

Q. 17

$u^2_t = 0.0000008 + 0.373\ U^2_{t-1}$

t (7.84) (10.235)

R^2 0.1397 d = 1.989

Using above information, answer the followings:

1. Identify this model

2. Is the model significant?

Answer:

Moving average model

Yes

Q. 18

$$X = \begin{vmatrix} 1 & -0.75 & -0.95 & -1.13 \\ 1 & 0.9 & -1 & -1 \\ 1 & -0.95 & 1.1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1.1 & -1.05 & 1.4 \\ 1 & 1.15 & -1 & 1 \\ 1 & -0.9 & 1 & 1 \\ 1 & 1.25 & 1.15 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}$$

$$Y = \begin{vmatrix} 32 \\ 46 \\ 57 \\ 65 \\ 36 \\ 48 \\ 57 \\ 68 \\ 50 \\ 44 \\ 53 \\ 56 \end{vmatrix}$$

1. Fit the econometric model by using matrix approach.
2. Forecast Y, if $X_1 = 20$, $x_2 = 30$; and $X_3 = 40$

Answer:

Y (acp) = 50.36 +5.419X1 +10.17X2 + 1.08X3

Forecasting value is about Y = 499

Q. 19

$C_{it} = 2.315 + 10.11X_{1t} + 2.385X_{2t} + 16.17X_{3t} + 1.119Q_{it}$

Using above information, answer the followings:

1. Find out the cost of company if Q = 1000

2. Find out the variation of cost if Q = 2000

Answer:

Cost for company 1 is: 1121.32

Cost for company 2 is: 1131.43

Cost for company 3 is: 1123.72

Cost for company 4 is: 1137.49

For Q = 2000, 2 will be multiplied to Q; but other steps are same.

Q. 20

Sample Units	Y	X
1990	3748	21777
1991	4010	22418
1992	3711	22308
1993	4004	23319
1994	4151	24180
1995	4569	24893
1996	4582	25310
1997	4697	25799
1998	4753	25886
1999	5062	26868
2000	5669	28314
2001	5628	29091
2002	5736	29450
2003	5946	30705
2004	6501	32372
2005	6549	33152
2006	6705	33764
2007	7104	34411
2008	7609	35429
2009	8100	36200

1. Examine unit root test for both X and Y
2. Examine the cointegration by EG and JJ method

Answers:

```
Cointegration with no intercepts or trends in the VAR
   Cointegration LR Test Based on Maximal Eigenvalue of the
Stochastic Matrix
********************************************************************
 19 observations from 2 to 20. Order of VAR = 1.
 List of variables included in the cointegrating vector:
 SC1             SC2
 List of eigenvalues in descending order:
.75521     .13552
********************************************************************
 Null    Alternative    Statistic     95% Critical Value  90%
Critical Value
 r = 0      r = 1        26.7398           11.0300
9.2800
 r<= 1      r = 2         2.7669            4.1600
3.0400
********************************************************************
 Use the above table to determine r (the number of cointegrating
vectors).
```

```
Cointegration with no intercepts or trends in the VAR
          Cointegration LR Test Based on Trace of the Stochastic
Matrix
*******************************************************************
 19 observations from 2 to 20. Order of VAR = 1.
 List of variables included in the cointegrating vector:
 SC1            SC2
 List of eigenvalues in descending order:
.75521     .13552
*******************************************************************
***********
 Null    Alternative    Statistic     95% Critical Value  90%
Critical Value
 r = 0      r>= 1        29.5067           12.3600
10.2500
 r<= 1      r = 2         2.7669            4.1600
3.0400
*******************************************************************
***********
 Use the above table to determine r (the number of cointegrating
vectors).

            Cointegration with no intercepts or trends in the VAR
Choice of the Number of Cointegrating Relations Using Model
Selection Criteria
*******************************************************************
***********
 19 observations from 2 to 20. Order of VAR = 1.
 List of variables included in the cointegrating vector:
 SC1            SC2
 List of eigenvalues in descending order:
.75521     .13552
*******************************************************************
***********
 Rank      Maximized LL       AIC            SBC            HQC
 r = 0      -360.1591      -360.1591      -360.1591       -
360.1591
 r = 1      -346.7892      -349.7892      -351.2059       -
350.0290
 r = 2      -345.4058      -349.4058      -351.2947       -
349.7255
*******************************************************************
***********
 AIC = Akaike Information Criterion    SBC = Schwarz Bayesian
Criterion
 HQC = Hannan-Quinn Criterion
```

```
OLS estimation of a single equation in the Unrestricted VAR
*******************************************************************
 Regressor              Coefficient       Standard Error         T-
Ratio[Prob]
 SALES(-1)                  1.0635           .0066492
159.9361[.000]
*******************************************************************
 R-Squared                     .99462   R-Bar-Squared
.99462
 S.E. of Regression            3282.1   F-Stat.
*NONE*
 Mean of Dependent Variable   91041.1   S.D. of Dependent Variable
44741.1
 Residual Sum of Squares     2.80E+08   Equation Log-likelihood
-256.4007
 Akaike Info. Criterion     -257.4007   Schwarz Bayesian Criterion
-258.0486
 DW-statistic                  1.6407   System Log-likelihood
-256.4007
*******************************************************************


                               Diagnostic Tests
*******************************************************************
*    Test Statistics  *        LM Version       *          F Version
*
*******************************************************************
*
* A:Serial Correlation*CHSQ(1)  =  .018192[.893]*F(1,25)      =
.016855[.898]*
*                     *                         *
*
* B:Functional Form   *CHSQ(1)  =   .35244[.553]*F(1,25)      =
.33065[.570]*
*                     *                         *
*
* C:Normality         *CHSQ(2)  =   3.4962[.174]*       Not
applicable        *
*                     *                         *
*
* D:Heteroscedasticity*CHSQ(1)  =   5.4329[.020]*F(1,25)      =
6.2977[.019]*
*******************************************************************
```

Plot of the Leverage Measures of the Regression