

POKHARA UNIVERSITY

Level: Bachelor

Semester: Fall

Programme: BBA/BI/BCIS/TT

Course: Data Analysis and Modeling

Year: 2021

Full Marks: 100

Pass Marks: 45

Time: 3 hrs.

Candidates are required to answer in their own words as far as practicable. The figures in the margin indicate full marks.

Section "A"

Very Short Answer Questions

Attempt all the questions. [10×2]

1. Interpret the slope of the regression model $\hat{Y} = 35 - 10X$. What is the estimated value of Y when X=1.5.
2. Detect active and inactive constraint of following LPP.
 Max $Z = 5X + 7Y$
 Subjected to constraint
 $2X + 3Y \leq 30$
 $3X + 5Y \leq 42$
 Where max $Z = 62$ at $X = 4$ and $Y = 6$.
3. Given the following information from a multiple regression analysis:
 $n = 22$, $b_1 = 4.5$, $b_2 = 4$, $S_{b_1} = 2.2$, $S_{b_2} = 1.1$
 Set up a 90% confidence interval estimate of the population slope β_1 .
4. Complete the following ANOVA summary table which was obtained from a multiple regression model with four independent variables:

Sources	Degree of freedom (d.f.)	Sum of square (SS)	Mean square (MS)	F
Regression		58		
Error				
Total	22	160		

5. The Coefficient of correlation between variables y and x is -0.86. Calculate and interpret the value of the coefficient of determination.
6. Convert the following profit matrix into opportunity loss matrix.

Job	Machine			
	P	Q	R	S
A	23	11	16	15
B	18	12	12	13
C	19	15	17	14
D	16	13	11	17

7. If $r = 0.95$ then find VIF. And interpret the result.
8. For $\sum p_1 q_0 = 1900$, $\sum p_0 q_0 = 1360$ then find the Laspyres Index.
9. Draw the network diagram

Activity	A	B	C	D	E	F
Predecessor	-	-	A	B	D	C

10. Find MAD of

D-F	3.4	0.36	1.62	1.16
-----	-----	------	------	------

Section "B"
Descriptive Answer Questions

11. Attempt any six questions. [6×10]

A farm administers a test to sales trainees before they go into the field. The management of the farm is interested in determining the relationship between the test scores and the sales made by the trainees at the end of one year in the field. The following data were collected for 10 sales personnel who have been in the field:

Sales Person Number	Test Score(X)	Number of units sold(Y)
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150

Calculation shows that

$$\sum X = 34 \quad \sum Y = 1371 \quad \sum X^2 = 122.62 \quad \sum Y^2 = 201121 \quad \sum XY = 4954$$

- Find the correlation coefficient between the test score and number of units sold, examine if this linear relationship is significant at the 5% level of significance.
- Compute the 90% prediction interval for the number of units sold of salesperson 9.
- Find coefficient of determination and interpret its meaning.

12. The department store has been expanding market share during the past 7 years, posting the following gross sales in millions of dollars:

Year	1995	1996	1997	1998	1999	2000	2001
Sales	14.8	20.7	24.6	32.9	37.8	47.6	51.7

- Find the linear estimating equation that best describes the data.
- Calculate the percent of trend for these data.
- Calculate the relative cyclical residual for these data.
- In which years does the largest fluctuation from trend occur, and is it the same for both methods.

13. a) Reconstruct the following indices using 2005 as the base

Year	2001	2002	2003	2004	2005	2006	2007	2008
I.N.	110	130	150	175	180	200	220	250

b) The following data represent the annual number of employee (in thousands) in an oil supply company for the years 1990 to 1997.

Year	1990	1991	1992	1993	1994	1995	1996	1997
No. of employee	1.45	1.55	1.61	1.60	1.74	1.92	1.95	2.04

Apply exponential smoothing technique with exponential constant 0.50 to compute the forecast value. What is the forecast value for year 1998.

14. a) Write the dual of the following LP model

Maximize $Z = 2x_1 + 4x_2 - 6x_3 + 12x_4$
 Subject to constraints
 $x_1 - 3x_2 - 7x_3 + x_4 \leq 80$
 $3x_1 + 2x_2 + 8x_3 \geq 15$
 $2x_1 + 5x_2 + 9x_4 \geq 125$
 $x_1, x_2, x_3, x_4 \geq 0$

b) Atv company products two types of TV set the Astro an the Cosmo. There are two production lines, one for each set, and there are two department, both of which are used in the production of each set. The capacity of the Astro production line 70 set per day. The capacity of the Cosmo line is 50 set per day. In department A picture tubes are produced. In this department the Astro set required 5 labor hours and Cosmo set required 4 labor hour. Presently in department A, a maximum of 120 labor hours can be assigned to production of the two types of sets. In department B the chassis is constructed. In this department the Astro set requires 4 labour hours and the Cosmo also require 2 labor hours. Presently in department B a maximum of 90 labor hour per day can be assigned to production of the two types of sets. The profit contributions are 20 and 10 rupees respectively, for each Astro and Comso set.

- i. Formulate the LP model
- ii. Solve graphically.

15. a) Four different ships are to be assigned to three cargo consignors with a view to maximizing the profit. From the following profit matrix representing the problem, work out the optimal assignment plan and determine the maximum possible profit.

Ships	Cargo consignors		
	P	Q	R
A	1	4	5
B	2	3	3
C	3	3	3
D	5	1	2

b) Bindhabasini digital has developed a substantial market share in the PC computer industry. The prices and number of units sold their top four computer products from 2019 to 2021 were:

Model	Selling price (RS)		Number of sold	
	2019	2021	2019	2021
ED 107	1985	1995	45	107
Electra	2700	2780	44	56
Optima	1700	1900	109	124
4K	15000	16000	24	34

16. Construct Fisher's price index.
 Solve the following transportation problem for minimum cost by taking initial feasible solution by Vogel's Approximation Method. The entries in the matrix indicate the cost in rupees of transporting a unit from a particular source to a particular destination.

Origin	D ₁	D ₂	D ₃	D ₄	
O ₁	10	8	11	7	20
O ₂	9	12	14	6	40
O ₃	8	9	12	10	35
Requirement	16	18	31	30	95

17. Following table lists the activity of a project along with their time estimates.

Activity	Predecessor	Most likely (t _m)	Optimistic (t _o)	Pessimistic (t _p)
A	-	5	4	6
B	-	12	8	16
C	A	5	4	12
D	B	3	1	5
E	D,A	2	2	2
F	B	6	4	8
G	C,E,F	14	10	18
H	G	20	18	34

The scheduled completion date for this project is 62 days. Draw the network diagram and find.

- The probability that the project will be finished within the scheduled date.
- The probability that the project will be completed at least 6 days prior to the expected time.
- What should be the scheduled completion time for the probability of completion to be 90%.

Section "C"

Case Analysis

18. The following data given by SPSS software.

ANOVA

	d.f.	SS	MSS	F
Regression	2	809	?	?
Residual	?	?	?	
Total	14	1000		

Coefficient table

	Coefficient	Standred error	t-value
Intercept	52.88	5.51	
Life expt	-0.78	0.122	
Literature	-036	0.067	

Complete the table and calculate:

[20]

- At 5% level of significance, which of the given independent variable is better explanatory variable for birth rate per thousands of population?
- What proportion of the variation in the estimation of birth rate is explained by fitted regression model and what proportion remains to be accounted by other factors?
- Established 95% confidence interval or birth rate for such a nation whose life expectancy at birth is 79.9 and literacy rate 99 percentage.
- Is there any significance relationship between the dependent and two independent variables in the fitted regression model? At 5%.