

Indian Statistical Institute, Bangalore

MS (QMS) First Year

Second Semester - Multivariate Data Analysis

Final Exam
Maximum marks: 50

Date: May 03, 2019
Duration: 3 hours

Answer as many questions as you can. The maximum you can score is 50 marks

1. Explain ridge regression? Give the step by step details of carrying out ridge regression? Briefly, explain how ridge regression can tackle multicollinearity?

[10]

2. A hotels and resorts chain is going to open a new hotel at an emerging city of India. The company has conducted a conjoint analysis to obtain customer preference on different room packages. The details of the conjoint analysis with the aggregate ranking are given below.

SL No	Room Rent (Rs)	Complimentary Breakfast	Airport Pick up & Drop	Wi-Fi	Laundry service	Aggregate Rating
1	< 3000	No	Not Free	Not Free	Not Free	1
2	3000 - 5000	No	Not Free	Free	Free	3
3	3000 - 5000	Yes	Free	Not Free	Not Free	7
4	< 3000	Yes	Free	Free	Free	8
5	> 5000	No	Free	Not Free	Free	1
6	3000 - 5000	No	Free	Free	Not Free	5
7	3000 - 5000	Yes	Not Free	Not Free	Free	6
8	> 5000	Yes	Not Free	Free	Not Free	4

- a. Analyze the data, compute part worth utilities and importance scores.
b. Identify the most preferred room package?
c. Estimate the expected rating for the best package?

[18]

3. A model is developed using regression to estimate the resolution time of tickets (y) in a technical support process using knowledge repository usage (x_1) and programmer skill (x_2) as explanatory variables. The data collected is given below

SL No	x_1	x_2	y	SL No	x_1	x_2	y
1	Low	Average	575	7	Medium	Average	540
2	Low	Excellent	540	8	Medium	Excellent	500
3	Low	Good	560	9	Medium	Good	525
4	Low	Excellent	530	10	Medium	Average	535
5	Low	Good	550	11	High	Average	520
6	Medium	Good	530	12	High	Excellent	490

The coefficient table obtained is given below

	Estimate	Standard Error	t	p-value
(Intercept)	522.763	4.157	125.745	5.31e-13
x_1 Low	47.719	4.596	10.383	1.67e-05
x_1 Medium	15.614	4.596	3.397	0.0115
x_2 Excellent	-35.526	3.795	-9.361	3.30e-05
x_2 Good	-13.18	3.873	-3.403	0.0114

- Check whether the model is significant?
- Is the model accurate?
- Validate the model on the following observations

x_1	Low	Low	Medium	Medium	High	High	High
x_2	Excellent	Good	Average	Excellent	Average	Good	Excellent
y	540	550	545	510	525	525	495

[18]

4. Explain polynomial regression? Give the step by step details of carrying out polynomial regression?

[10]