INDIAN STATISTICAL INSTITUTE
SQC & OR Unit, Hyderabad

MS in Quality Management Science : 2015 -17
III SEMESTER: MID-TERM EXAMINATION

Subject: Applied Regression Analysis (ARA)

Date : 6 September 2016 **Duration : 2 Hours** **Max. Marks : 50**

INSTRUCTIONS

This paper contains questions for 60 marks. Answer as many as you can but the maximum you can score is 50 marks. You will also be given soft copy of this question paper for using data for analysis. You are free to use appropriate statistical software for analysis but the answers will have to be given in the answer sheet given to you. You need also to submit the soft copy of your analysis if any, only in a ***word file format*** *opened in your name before the end of the examination without fail.*

Q1. In a high precision electronic industry, an optic lens is used to produce laser for brazing components. Currently the industry is using one type of optic lens (Type A) which is an imported one. The lenses are to be periodically changed after producing certain number of components when the power (W) drops by a fixed unit from the initial setting for set voltage (V) and frequency (hertz). As the number of components being produced is very high, the industry is incurring high cost of manufacture by way of this consumable. In an attempt to reduce the cost of consumable (optic lens), the industry has found an alternative supplier who can supply indigenously and the validation trials are being conducted before the actual substitution is made in regular production. During the trials though the output quality is found to be at par with that of imported but the number of components that can be produced observed to be less as compared to imported. The imported lens is 3 times costlier than the indigenous and newly developed. In order to study and compare the performance of the two lenses, a study has been conducted for a shift with observations taken at every 15 minutes on the cumulative no. of components (N) produced (counter reading) and the corresponding power drop (W). The data is presented in the table below.

TYPE A (Imported)			**TYPE B (Indigenous)**		
Time	Cum no. produced	Cumu Power Drop	Time	Cum no. produced	Cumu Power Drop
8:15	0	0	10:15	0	0
8:30	97727	0.2	10:30	90461	0.28
8:45	186535	0.6	10:45	180073	0.58
9:00	267585	0.9	11:00	272606	1.03
9:15	357517	1.1	11:15	360317	1.38
9:30	446619	1.3	11:30	450739	1.7
9:45	538991	1.5	11:45	550982	2.06
10:00	624452	1.7	12:00	625211	2.28
10:15	723305	1.8	12:15	731177	2.68
10:30	810596	2.1	12:30	807555	2.88
10:45	907489	2.3	12:45	895347	3.18

11:00	1001031	2.6	1:00	994080	3.48
11:15	1114096	2.7	1:15	1078741	3.68
11:30	1173194	2.8	1:30	1180783	3.98
11:45	1264026	3.1	1:45	1273027	4.28
12:00	1352028	3.4	2:00	1370339	4.58
12:15	1451161	3.6	2:15	1447910	4.68
12:30	1554794	3.8	2:30	1530691	4.98
12:45	1629534	4.2	2:45	1632584	5.23
1:00	1691572	4.3	3:00	1698603	5.48
1:15	1810888	4.4	3:15	1795786	5.68
1:30	1882007	4.5	3:30	1887536	5.83
1:45	1999343	4.9	3:45	1976127	6.08
2:00	2077133	5	4:00	2086482	6.38
2:15	2172125	5.1	4:15	2163752	6.68
2:30	2247125	5.25	4:30	2262795	6.88
2:45	2349989	5.5	4:45	2353327	7.18
3:00	2421638	5.75	5:00	2447640	7.38
3:15	2528523	5.9	5:15	2524740	7.68
3:30	2618214	6.24	5:30	2618922	7.78
3:45	2700765	6.3	5:45	2709228	8
4:00	2798798	6.4	6:00	2799116	8.16

a) Obtain the least square estimates for the linear regression relationship for Type-A optic lens by writing explicitly the normal equations for fitting regression equation between cumulative power drop Vs cumulative no. of components produced.

b) Build separately appropriate linear regression equations for each laser type and examine the adequacy of the fitted model. What would be your approach to improve the adequacy in each case?

c) If the objective is to propose to the plant, a prediction model for the cumulative no. of components produced for a specified power drop then what would be the change in your approach. For the same, fit a single regression model by considering the Type of the optic as a predictor and comment on the adequacy of the final fitted model. State the hypotheses involved clearly for the model parameters.

d) Write down the expressions for the confidence and prediction limits to predict the number of components that can be produced for given power drop and obtain the 95% confidence and prediction limits to estimate for the minimum no. of components that can be produced for a power drop of 5 units. Finally comment on the comparative performance of the lenses by giving a relative efficiency index and whether would you like to conclusively recommend going for import substitution with indigenous one based on economics?

(6+5+8+6=25 Marks)

Q2. Dairy Herd Improvement Cooperative (DHI) in upstate New York collects and analyses data on Milk production. One question of interest here is how to develop a suitable model to predict current milk production from a set of measured variables. The response variable (current milk production in pounds) and the predictor variables are given in the data table. Samples are taken once a month during milking. The period that a cow gives milk is called lactation. Number of lactations is the number of times a cow has calved or given milk. The recommended management practice is to have the cow produce milk for about 305 days and then allow a 60 day rest period before beginning of the next lactation. The data set consists of 199 observations, was compiled from the DHI milk production records and given in the excel file Milk Production Data.xlsx

The variables given in data are:

Variables in Milk Production Data	
Variable	**Definition**
Current	Current Month milk production in pounds
Previous	Previous month production in pounds
Fat	Percent of Fat in the Milk
Protein	Percent of Protein in Milk
Days	Number of days since present Lactation
Lactation	Number of Lactations
I 79	Indicator Variable (0 if days<= 79 and 1 if Days >79)

a) For the above data, use different variable selection approaches and find out the most suitable model that adequately explains the maximum variation in milk production. Comment on the most suitable approach for building the model for this problem and the basis for your final model selection. List the appropriate measures with criteria you will look in your final selection of the model and the merits/demerits in those measures.

b) Perform the residual analysis and comment on the adequacy of the model. State the assumptions involved in your estimation method about residuals and comment on whether your model meets those assumptions.

c) What are the commonly adopted diagnostic measures in model building using OLS method? Do you need to adopt any of the diagnostic measures and if so, explain the basis you have adopted to overcome the problems in your modelling.

d) Explain what approach you would adopt in validating your model when you particularly have large amount of data and the rationale in such approach?

(15+5+10+5=35 Marks)

//*****//

Variables in Milk Production Data

Variable	Definition
Current	Current Month milk production in pounds
Previous	Previous month production in pounds
Fat	Percent of Fat in the Milk
Protein	Percent of Protein in Milk
Days	Number of since present Lactation
Lactation	Number of Lactations
I 79	Indicator Variable (0 days<= 79 and 1 if Days >79)

Dairy Herd Improvement Cooperative (DHI) in upstate Newyork collects and analyzes data on Milk production. One question of interest here is how to develop a suitable model to predict current milk production from a set of measured variables. The response variable (current milk production in pounds) and the predictor variables are given in the data table. Samples are taken once a month during milking. The period that a cow gives milk is called lactation. Number of lactations is the number of times a cow has calved or given milk. The recommended management practice is to have the cow produce milk for about 305 days and then allow a 60 day rest period before begining of the next lactation. The data set consists of 199 observations, was compiled from the DHI milk production records.

CurrentMilk	Previous	Fat	Protein	Days	Lactation	I79
45	45	5.5	8.9	21	5	0
86	86	4.4	4.1	25	4	0
50	50	6.5	4	25	7	0
42	42	7.4	4.1	25	2	0
61	61	3.8	3.8	33	2	0
93	93	4.2	3	45	3	0
91	91	2.9	2.6	46	2	0
90	90	4.7	2.9	46	5	0
53	53	2.5	3.5	46	2	0
84	84	4.3	3.3	50	7	0
70	20	3.6	3.2	53	2	0
76	50	4.3	3.3	56	2	0
95	60	3	2.97	57	1	0
69	55	3.2	2.6	57	1	0
99	81	2.9	2.5	59	6	0
97	24	5.9	3	59	5	0
49	51	3.8	2.9	59	1	0
108	58	3.2	2.7	60	5	0
98	98	3.1	3.1	60	2	0
71	57	3.2	3.2	60	1	0
58	47	3.1	2.9	62	1	0
76	82	3.2	2.7	69	2	0
70	57	3.9	3.4	70	2	0
69	57	3.4	2.6	70	1	0
92	63	3.3	3.1	71	2	0
86	46	3	3.2	73	8	0
95	81	3.1	2.6	77	3	0
84	65	3.3	2.7	77	1	0
56	54	4.6	3.5	78	2	0
92	95	2.9	3	79	3	0
90	85	3.7	3	217	4	1
58	54	2.7	2.8	217	1	1
57	62	4.6	3.5	217	1	1
21	26	2.6	3.2	217	3	1
86	73	3.1	3.1	219	2	1
52	60	2.4	3	219	1	1
51	52	3.8	3.3	219	2	1
42	45	4.6	3.4	223	1	1
59	37	3.8	3.2	224	2	1
54	60	4.1	3.6	226	1	1
56	57	3.4	3.2	227	3	1
70	70	4.2	3.5	228	4	1
37	44	3.5	3.3	228	2	1
74	73	3.3	3.3	232	2	1
20	42	2.5	3	232	4	1
60	63	3.3	3	233	1	1
44	42	2.5	3.2	236	2	1
54	49	2.9	3.1	239	1	1
76	52	4.1	3	240	5	1

62	57	3.7	3.2	243	4	1
20	24	3.4	3.6	245	3	1
63	56	3.7	3.3	246	1	1
83	72	2.8	3.2	250	1	1
45	55	4.1	3.3	252	3	1
58	59	4	3.1	253	6	1
54	51	4.1	3.4	260	1	1
44	51	4.3	3.8	261	4	1
46	49	3.8	3.4	264	1	1
53	58	2.9	3	267	2	1
29	29	3.1	3.2	273	4	1
58	63	3.9	3.2	278	5	1
69	71	4.2	3.7	297	2	1
45	62	3.7	2.9	280	1	1
68	72	2.9	3.2	283	1	1
56	48	4.5	3.3	283	1	1
52	48	3.5	3.5	284	1	1
49	50	4.5	3.9	287	2	1
29	30	3.8	3.6	292	2	1
54	59	3.7	3.3	297	2	1
46	35	3.9	3.4	297	8	1
60	60	3.8	3.6	299	1	1
68	67	3.6	3.1	301	1	1
52	50	4.7	3.7	301	1	1
57	62	3.8	2.9	151	1	1
64	62	3.8	3.3	154	1	1
58	53	4.3	3.2	154	1	1
77	72	3.8	2.9	155	2	1
77	76	2.8	2.8	155	2	1
69	87	2.7	2.7	155	2	1
33	39	5.7	3.5	155	2	1
69	75	2.9	2.8	159	1	1
54	59	4.9	3.4	160	3	1
70	56	2.8	3.1	163	2	1
58	61	2.8	3.2	163	2	1
59	62	2.9	2.7	164	2	1
57	59	2.9	3.4	164	1	1
69	64	3.2	3.1	165	1	1
69	62	4.5	3.3	165	2	1
68	69	2.4	3.1	165	1	1
45	45	4.5	3.4	165	1	1
83	81	5.2	3	169	2	1
60	78	3.4	3.2	169	3	1
53	53	4	3.3	169	4	1
61	62	4.1	3	172	2	1
83	74	2.8	2.9	174	1	1
73	45	2.8	2.7	174	1	1
61	31	4.4	3.5	174	2	1
78	80	3.1	2.9	176	4	1
64	57	3.5	3.3	176	1	1

67	70	6.2	2.9	177	1	1
59	60	4	3.1	178	1	1
50	59	4.6	3	178	2	1
64	66	4.4	3.3	180	1	1
48	42	2.7	3.2	181	2	1
80	75	2.5	3.2	182	2	1
64	63	3.4	3.4	183	1	1
53	60	4.2	3.7	183	2	1
60	62	6.4	4.8	185	3	1
48	71	2.6	3.1	188	2	1
46	57	5.9	3.8	190	2	1
74	67	3.6	2.8	195	1	1
72	70	3.9	3.5	196	5	1
28	36	4.1	3.8	197	2	1
42	47	3.6	3.8	200	2	1
27	31	3.9	3.4	202	2	1
26	28	4.1	3.3	202	2	1
23	17	4.7	4	207	2	1
61	56	3.5	3.2	208	5	1
47	45	2.9	3.1	209	2	1
27	25	5.3	3.6	209	2	1
65	59	3.2	3.1	210	2	1
77	79	3.6	2.8	211	1	1
73	57	3.3	2.8	213	1	1
61	62	3.7	3.2	213	1	1
58	59	6.4	3.4	215	3	1
103	99	2.9	2.8	81	4	1
74	64	4	2.9	81	1	1
70	62	4.3	2.9	83	1	1
78	62	4.1	3.1	87	2	1
54	52	3.7	3.2	87	1	1
85	75	4.2	2.5	88	1	1
79	75	3.5	3.2	91	2	1
75	62	4	2.9	91	1	1
100	93	3.7	3.2	92	2	1
113	87	5.8	2.9	93	3	1
92	92	3.6	3.4	93	3	1
59	53	3.5	3.2	95	1	1
49	75	3.3	3.1	95	2	1
77	67	2.5	2.9	96	1	1
78	78	4.1	3.2	97	2	1
56	51	3.2	2.7	103	3	1
93	85	4.1	2.9	105	4	1
79	86	3.8	3	105	6	1
71	73	3.5	3.4	105	3	1
67	72	5.5	2.6	105	1	1
62	53	3	3	105	1	1
69	56	4	3.3	107	1	1
68	68	3.8	2.8	108	1	1
60	61	4.2	3.1	109	1	1

88	100	3.8	2.9	111	2	1
80	66	4	3.1	111	1	1
79	84	5.1	2.6	111	1	1
61	63	3.3	2.9	111	1	1
75	66	3.9	3.4	114	2	1
69	66	3.8	3.1	115	1	1
70	86	3.3	3.3	118	5	1
56	45	3.4	2.6	118	1	1
70	81	4.8	3	119	1	1
80	93	3.5	2.9	120	4	1
74	89	4.7	3.2	120	4	1
82	63	3.2	2.9	123	4	1
77	72	3.6	2.9	123	1	1
72	65	2.8	2.6	123	1	1
103	104	2.8	2.8	125	3	1
82	85	2.3	2.8	125	2	1
69	58	3.5	3.3	127	1	1
87	96	4.1	2.5	128	3	1
82	93	3.7	3.1	128	2	1
64	64	4.4	3.4	129	2	1
63	63	3.3	3.4	130	3	1
86	72	3.6	2.8	131	2	1
79	91	2.9	2.9	131	2	1
75	78	3.9	2.9	132	1	1
42	60	2.9	3.1	132	2	1
82	75	2.3	2.9	135	2	1
61	52	3.2	3.4	135	2	1
85	75	3	2.9	137	1	1
67	74	3.8	3.7	137	5	1
80	73	3.7	3.1	138	4	1
63	64	2.9	3.4	138	2	1
62	73	3.2	3.5	142	6	1
62	66	4.3	3.1	142	1	1
45	42	3.7	3.2	142	2	1
38	34	3.6	3.5	142	2	1
80	73	3.3	2.8	143	2	1
57	72	2.5	3.3	143	4	1
67	62	3.6	3.2	145	1	1
54	57	3.6	3	145	2	1
46	47	4.3	3.4	145	1	1
14	53	3	3.4	145	2	1
55	64	3.5	3	146	1	1
68	76	1.8	1.3	147	2	1
54	63	2.5	3.3	147	2	1
53	64	3.8	3.4	147	4	1
89	89	3.3	3	149	2	1
68	69	3.5	3.4	149	2	1
55	54	5.9	3.4	149	1	1
47	57	3.5	2.9	149	2	1
73	61	3.6	2.8	151	1	1