Reg. No									

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

II SEMESTER M.TECH. EXTERNAL EXAMINATIONS MAY 2019

SUBJECT: STATISTICAL DESIGN AND ANALYSIS OF EXPERIMENTS IN BIOTECHNOLOGY [BIO 5253]

Date of Exam: **02/05/2019** Time of Exam: **9.00 AM – 12.00 PM** Max. Marks: **50**

Instructions to Candidates:

- Answer ALL the questions & missing data may be suitable assumed

1A. Enzyme activity (mM/min) of lipase is given below, 12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25. Find out the IQR and construct the box-whisker plot. **3**

1B. For cells growing in liquid medium, the overall biomass yield coefficient Yx/s can be estimated as follows: Yx/s=(Xf-Xo)/(So-Sf), where Xf is the final cell concentration, Xo is the initial cell concentration, So is the initial substrate concentration and Sf is the final substrate concentration. For a particular cell culture, So = 30 g/L, Sf = 0.85 g/L, Xf = 12.7 g/L and Xo = 0.66 g/L. If the uncertainty associated with the substrate concertation is ± 0.1 g/L, how accurate must the measurement for cell concentration be if the maximum possible error in Yx/s must be less than 5%. **3**

1C. If your constructing a regression model, how will you check the following, **4**

i. Errors are normally distributed,
ii. Errors have the same variance
iii. Errors are independent
iv. Errors have a mean of zero

2A. Fit the following data into Y = mX and and Y = mX + C models and determine the values of R^2 by least square regression analysis. Which one of these models do you choose? Why? Comment on the results: **6**

X	25	50	75	100
Y	0.3174	0.7327	0.8232	1.155

2B. The following table is the result of 8-experiment PB design, so there are three dummy factors, labelled d1, d2, and d3, Determine the main effect, sum of square (SS), mean square (MS) and F-ratio of each variable. **4**

Run Number	A	D1	B	D2	C	D3	D	Yield
1	+	-	-	+	-	+	+	10
2	+	+	-	-	+	-	+	9
3	+	+	+	-	-	+	-	10
4	-	+	+	+	-	-	+	9
5	+	-	+	+	+	-	-	8
6	-	+	-	+	+	+	-	7
7	-	-	+	-	+	+	+	7
8	-	-	-	-	-	-	-	7

3A. The study on biosorption of chromium by *Sargassum sp* used the 2^3-factorial design. The selected design is given below, **8**

T	1	1	1	1	-1	-1	-1	-1
C	1	1	-1	-1	1	1	-1	-1
pH	1	-1	1	-1	1	-1	1	-1
Y (%)	75	8.4	59.6	83.2	99.5	9.9	65.6	73.8

a. Fit the equation $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_1X_2 + b_5X_1X_3 + b_6X_2X_3 + b_7X_1X_2X_3$.

b. If the value of MSE is 21.22 & if the experiments are done in duplicates, calculate the standard errors of Y and each parameter.

c. Diagrammatically represent the main effect of T, C and pH and interaction effect of variables.

3B. List and describe the methods used to process the research data by descriptive and inferential analysis **2**

4A. Discuss the different types of experimental hypothesis testing methods used in data analysis. **5**

4B. A group of seven chickens reared on a high protein diet of 12, 15, 11, 16, 14, 14, and 16 ounces; a second group of five chickens of same age, similarly treated except that they receive a low protein diet, of 8, 10, 14, 10 and 13 ounces. assumed mean (or A1) = 10 for the sample of 7 chickens and assumed mean (or A2) = 8 for the sample of 5 chickens Test at 5 per cent level whether there is significant evidence that additional protein has increased the weight of the chickens by applying one sample t-test. Note: The tabulated value of t-statistic is 1.812 **5**

5A. A chemical engineer is interested in determining the operating conditions that maximize the yield of a process. The engineer suspects that two factors namely reaction time and reaction temperature influence the yield. Currently the process is set at a reaction time of 35 minutes and reaction temperature of 155° and the yield is only 40 %. Because it is unlikely that the current operating region contains the optimum, the engineer decides to fit a lower order polynomial model to the yield and then move to the optimum using steepest ascent method as shown below. **5**

EXP .NO.	STEPs	X1	X2	Y
	Δ	5	2	
1	origin	35	155	40.3
2	origin + Δ	?	?	41
3	origin + 2Δ	?	?	42.9
4	origin +3 Δ	?	?	47.1
5	origin + 4Δ	?	?	49.7
6	origin + 5Δ	?	?	53.8
7	origin + 6Δ	?	?	59.9
8	origin + 7Δ	?	?	65
9	origin + 8Δ	?	?	70.4
10	origin +9 Δ	?	?	77.6
11	origin + 10Δ	?	?	80.3
12	origin + 11Δ	?	?	76.2

$$y = 40.44 + 0.775x_1 + 0.325x_2$$

i. Is there any curvature in the model?

ii. Why the higher order model terms are missing in the above regression model. Explain

iii. How do you proceed to the optimization using RSM?

5B. The ANOVA results for the quadratic equation of Laccase production (using Minitab software) is presented in the table. Identify the missing information and complete the analysis. **5**

Source	SS	DF	MS	F
Model	59416.222	8	-	-
Intercept	-	1	400900.028	-
A	10683.722	2	-	-
B	-	2	19559.361	-
C	9613.778	-	-	
Error	18230.750	27	-	
Total	478547.00	36	R^2 =?	