Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

A Constituent Institution of Manipal University

II SEMESTER M.TECH (INDUSTRIAL BIOTECHNOLOGY)
END SEMESTER EXAMINATIONS, MAY 2018
SUBJECT: STATISTICAL DESIGN AND ANALYSIS OF EXPERIMENTS IN BIOTECHNOLOGY [BIO 5253]
REVISED CREDIT SYSTEM (25/04/2018)

Time: 3 Hours MAX. MARKS: 50

Instructions to Candidates:

- Answer **ALL** the questions.
- Missing data may be suitable assumed.

1A (5)

Fit the following binodal data to the polynomial equation $W_p = A + B\,W_s^{0.5} + C\,W_s$ and determine the coefficients by multiple regression analysis.

W_P (%)	39	31	13	10	6
W_s (%)	4	6	12	14	18

1B (5)

The calibration data of a biomolecule are shown below:

Concentration (mM)	1	1.2	1.5	1.7	2
Absorbance @ 600 nm	0.99	1.13	1.52	1.73	1.96

a) Let $A = B_o + B_1C$ be the equation of the least-squares line for predicting absorbance and concentration. Compute the values of B_o and B_1.
b) Let $A = B_1C$ be the equation of the least-squares line for predicting absorbance and concentration. Compute the value of B_1.

If the standard error of the B_o and B_1 for the first model are 0.09 and 0.06 and the t_{crit} value is 3.18 and if the standard error of the B_o for the second model is 0.012 and the t_{crit} value is 2.78, which model do you select?

2 (10)

The average yield (Y, %) of a process for a 2^3 factorial design is given below (standard order)
Y (%): 86,78,85,79,82,83,85,82

a) By considering the intercept, linear effects, two-way interactions and three-way interactions calculate the main effects and interaction effects.
b) Diagrammatically represent the main effect of T, C and pH.
c) If the value of MSE is 3.34 & if the experiments are done in duplicates, calculate the standard error of Y and each parameter.
d) Check the significance (95% confidence interval) of each parameter in the model and suggest a best model for this system:

3A (5)

Explain the following concepts in DOE with an example:

a. Hidden replication
b. Randomization

3B (5)

Following is the regression output from Excel spread sheet. Fill up the table by considering the model $Y = A+BX_1+CX_2$

SUMMARY OUTPUT

Regression Statistics	
R Square	-----
Standard Error	159.206
Observations	5
ANOVA	

	df	*SS*	*MS*	*F*	*F crit*
Regression	-----	-----	-----	-----	-----
Residual	-----	-----	-----		
Total	-----	25346.55			

4A. (5)

CCD was suggested for the optimization of pH and temperature for the production of carboxymethyl cellulase synthesized by a fungal system. List all the experimental runs that are needed in CCD, in coded and uncoded forms.

Variable	**pH**	**Temp (°C)**
Low	5	20
High	7	30

4B. (5)

You are interested in optimizing the yield of a process by checking different combinations of two factors namely, time ($30 \leq t \leq 40$min) and temperature ($150 \leq T \leq 160$°C). A first order model in coded variables has been fit to yield data from 2^2 designs. The design and the resulting yields are as follows: Yield at the factorial points (in standard order): 49.3, 50, 50.9, 51.5, and Yield at the center points: 50.3, 50.5, 50.7, 50.2, 50.6. The model is $Y= 44.81 + 0.775X_1+0.325 X_2$.

a) Is there any curvature in the model?
b) After checking curvature, how do you proceed to the optimization?

5A (5)

Assume that you have got a second order model equation at the end of response surface methodology. How do you check the following? Support your answers with specific plots as applicable:

i. significance of the model
ii. significance of each terms in the model
iii. normal distribution of errors
iv. independency of errors

5B (5)

Body weights of 18 diabetics are shown below.

107	119	99	114	120	104
114	124	116	101	121	152
125	114	95	117	100	88

Construct the stem-and-leaf-plot. Comment on the shape, location and spread.