Reg. No.									

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

(A constituent unit of MAHE, Manipal)

VII SEMESTER B. TECH (BIOTECHNOLOGY)
END SEMESTER EXAMINATIONS (REGULAR), MARCH 2021
SUBJECT: BIOSTATISTICS & ANALYTICAL TECHNIQUES [BIO 4103]
REVISED CREDIT SYSTEM

Time: 3 Hours MAX. MARKS: 50

Instructions to Candidates:

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

1A.	List and describe the components of research design with an example	**3**
1B.	What sort of study is it when researchers make use of naturally occurring variation in their independent variable(s)? What is the alternative?	**2**
1C.	A researcher studying bird vocalization cannot tell if she has already recorded a bird or not. Why might this cause a problem with her data analysis?	**3**
1D.	If your constructing a regression model, how will you check the following, i. Errors are normally distributed, ii. Errors have the same variance iii. Errors are independent iv. Errors have a mean of zero	**2**
2A.	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.	**2**
2B.	Obtain least squares estimates of the parameters in the model, $y= \beta_0+\beta_1 x+ \beta_1 x^2+\epsilon$ for given the data. Draw the curve for predicted response. Explain your answer. Substrate (g): 10, 10, 15, 20, 20, 25, 25, 25, 30, 35 Growth rate: 73, 78, 85, 90, 91, 87, 86, 91, 75, 65	**4**
2C.	The luciferase enzyme in fireflies catalyzes the modification of luciferin, consuming both luciferin and ATP, and producing light. A series of experiments are performed in which 5 µM luciferase enzyme is mixed with various concentrations of substrate S_0, and the relative reaction rates are measured in terms of light emission rates (RLU), measured using a photomultiplier tube: **S_0 (µM)**: 5, 10, 20, 40, 80, 200, 500, 1000 **RR (RLU)**: 3554, 6262, 10115, 14611, 18786, 22672, 24718, 25484 From the data, estimate the V_{max} (in RLU) and K_M (in µM) by least square analysis. Determine ARD and R^2 also.	**4**
3A.	A 2^{5-1} fractional factorial design was used to determine the significant variables that affect the enzyme activity. The experimental results are 6.2, 8.4, 9.5, 11.9, 7.5, 9.9, 10.1, 11.9, 10.7, 12.0, 11.6, 13.0, 9.4, 11.4, 10.4 and 13.0. Estimate the main effects of each parameters, six two-factor interactions, three three-factor interactions, one four-factor	**5**

	interaction and fit into the model equation with its co-efficient. If the value of the residual value is 4.01. Calculate the F value for linear variable.	
3B.	Construct the following plots for question number, **3A;** main effect plots, all possible two-factor interaction plots, cube plots for the response and plot a graph of response versus run order.	**5**
4A.	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.	**3**

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

4B.	An IQ test is standardized so that 50% of male students in the general population obtain scores of over 100. Among 43 male applicants at an army recruiting office only 10 obtained scores higher than 100. Is there significant evidence that the applicants are not a random sample from the population used to standardize the test?				**2**
4C.	The Y data below were fit using a linear model in X using a spreadsheet and the output is shown. Fill up the blanks.				**3**

Regression Statistics				**Coefficients**	
R Square	----- (1)		**Intercept**	***βo***	
Standard Error	18.071		**X variable 1 to X variable 7**	***β1-β7***	
ANOVA					
	df	***SS***	***MS***	***F***	***F crit***
Regression	-----(2)	-----(5)	-----(7)	-----(9)	-----(10)
Residual	-----(3)	-----(6)	-----(8)		
Total	-----(4)	11455.94			

4D.	Construct a 2^{6-2} fractional factorial design with high resolution possible. What are the generators of your design? What us defining relation of your design? What is confounded with the main effect 3 of your design? What is confounded with 1x2 interaction?	**2**

5A. (3)

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 $y = 40.44 + 0.775x_1 + 0.325x_2$ and then move to the optimum using steepest ascent method as shown below.

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

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 (2)

Genetic theory states that children having one parent of blood type *A* and the other of blood type *B* will always be of one of three types, *A*, *AB*, *B* and that the proportion of three types will on an average be as 1 : 2 : 1. A report states that out of 300 children having one *A* parent and *B* parent, 30 per cent were found to be types *A*, 45 per cent per cent type *AB* and remainder type *B*. Test the hypothesis by chi-square test and degree of freedom (Table value of □2 at 5 per cent level of significance is 5.991.).

5C. (2)

The β-carotene has λ_{max} 450 nm, and ε = 15,000 $m^2\ mol^{-1}$. Calculate the absorption expected for a solution in which 0.1 mg has been dissolved in 10 ml of water with a path length of 1 cm.

5D. (3)

Sucrose concentration in a fermentation broth is measured using HPLC. Chromatogram peak areas are measured for five standard sucrose solutions to calibrate the instrument. Measurements are performed in triplicate with results as follows:

i. Find an equation for sucrose concentration as a function of peak area using simple linear regression.

ii. A sample containing sucrose gives a peak area of 209.86, what is the sucrose concentration?

Sucrose conc. (g/L)	**Peak area**
6	55.55, 57.01,57.95
12	110.66,114.76,113.05
18	168.9,169.44,173.55
24	233.66,233.89,230.67
30	300.45,304.56,301.11