Reg. No.		
----------	--	--



# MANIPAL INSTITUTE OF TECHNOLOGY

Manipal University Manipal – 576 104



**MAX. MARKS: 100** 

## **II SEMESTER M.Tech.(BME) DEGREE END SEMESTER EXAMINATIONS, MAY 2016**

SUBJECT: Statistical Data Analysis (Program Elective II - BME 567) (REVISED CREDIT SYSTEM)

Thursday, May 12<sup>th</sup>, 2016: 09.00 am- 12.00 pm

### **TIME: 3 HOURS**

### **Instructions to Candidates:**

- 1. Answer any FIVE full questions. Each question is worth 20 marks.
- 2. ALL STEPS MUST BE SHOWN.
- 3. You are allowed to use statistical tables but formula sheet is not allowed. You can use any calculator.
- 4. All answers should be accurate to three decimals or more.

### Q1)

### (Total=20 Marks)

A) With an example, briefly explain the concept of sampling distribution. How does it help to relate sample statistics to the population parameter? (5)

B) What is the basic difference between a Z distribution and a t distribution? When do you use t distribution? (2)

C) Poisson distribution is a form of Binomial distribution. Then why do we need to use Poisson distribution? (3)

D) In two group comparison, give examples as to when you will use (i) independent samples test and (ii) paired samples test. (4)

E) An independent samples test of two groups had the following data. Calculate the Pooled Variance.
(4)

Group	Sample size	Mean height (m)	Std deviation of height (m)
Male	140	1.723	0.15
Female	160	1.654	0.18

F) In comparing the two groups in E, write the null and alternate hypotheses. (2)

The standards for tall building structures require that concrete used must exceed a mean breaking strength of 160 kg /  $cm^2$ . The standard deviation is known to be 20 kg /  $cm^2$ . A sample of 80 pieces of concrete blocks reveals a sample mean breaking strength of 156.75 kg /  $cm^2$ . For this question,

- A) State the Null and the Alternate hypotheses. (2)
- B) Is there evidence that the concrete is not meeting the standards for average breaking strength? (Use  $\alpha = 0.05$ .) (10)
- C) Compute the p-value and interpret its meaning. (4)
- D) What are the Type I and Type II errors in this question? In the real business situation, what are the effects of committing these types of errors? (4)

### Q3)

(Total=20 Marks)

In grown up adults, body weight appears to be dependent on height of respondents. Data collected from twenty adults yielded the following information.

Synod	Body	Height	S.No	Body	Height
	weight Kg	m		weight Kg	m
1	68.8	1.86	11	56.7	1.535
2	64.7	1.72	12	54.2	1.43
3	65.2	1.605	13	55.1	1.455
4	60.8	1.645	14	57.3	1.395
5	68	1.955	15	53.6	1.5
6	61.7	1.64	16	63.9	1.775
7	55.7	1.51	17	61.9	1.735
8	59.9	1.565	18	69.4	1.8
9	61.6	1.725	19	71.8	1.94
10	59.4	1.665	20	75.9	1.88

Summary statistics for the above data is given below:

	Body	Height
	weight Kg	m
Sum	1245.6	33.335
Sum of Squares	78303.54	56.10188
Sum of Cross Products	2093.821	

### Q2)

(2)

- A) Determine the independent and dependent variables in this question. (1)
- B) Assuming a linear relationship, use the least square method to find the regression coefficients b0 and b1 and write the regression equation. Interpret the meaning of b0 and b1
   (10)
- C) Using the regression equation, predict the weight of an adult with a height of 1.8 m.

(1)

- D) Determine the Correlation coefficient and Coefficient of determination. (4)
- E) At the 0.01 level of significance, is there evidence of a linear relationship between the height and weight? (4)

### (Total=20 Marks)

- A) Blood pressure in a population is said to follow a normal distribution with a standard deviation of 30 mm Hg. Calculate the number samples required if you want to be 95% confident of estimating the population mean blood pressure to within a sampling error of 5 mm Hg.
   (3)
- B) Heart attack is treated in 3 ways: Coronary bypass, Intra-arterial balloon and localised enzyme treatment. In a hospital 40% of the patients had coronary bypass, 40% were treated with a balloon technique and the rest with the enzyme technique. The survival rates for the three techniques are 0.8, 0.85 and 0.70 respectively. If a heart attack survivor is randomly chosen, what is the probability that the patient underwent (a) Enzyme treatment (b) Balloon treatment (c) Coronary bypass? Also calculate the total chance of survival of a heart attack patient admitted in the hospital. Draw the tree diagram and show all calculations. (7)
- C) A nutritionist is studying the effect of iron deficiency of pregnant mothers on the severity of child birth related problems. Classifying the iron as low, medium and high and severity of birth problems as Insignificant, average and severe, the following contingency table was obtained.

Severity of Birth problems	Severe	Average	Insignificant	Total
Iron deficiency				
Low	120	5	10	135
Medium	45	110	40	195
High	35	35	100	170
Total	200	150	150	500

At the 0.01 level of significance, is there evidence of a significant relationship between degree of iron deficiency and severity of birth problems? (10)

**Q4**)

IBM Technical Support uses 'problem resolution time' as a measure of its quality of service. Generally, the average 'resolution time' is known to follow a normal distribution. A random sample of 25 phone calls yielded the following 'resolution times' (in minutes).

Sample Statistics: 
$$\overline{x} = 48.56$$
;  $s = 36.74$ 

- A) Set up a 90% confidence interval estimate of the mean 'resolution time'. (4)
- B) At the 0.01 level of significance, test the proposition that the average 'resolution time' is greater than 45 minutes? (10)
- C) In response to complaints that the 'resolution time' is too long, the operations director states that "Almost certainly not longer than 45 minutes". On the basis of the results in (b), evaluate this statement.
- D) Suppose the p value for the hypothesis test in (b) is 0.3. Show how this supports or opposes your conclusion in (b). (3)

## Q6)

## (Total=20 Marks)

**A**) Bacteria are distributed independently of each other in a solution and it is known that the number of bacteria per ml follows a Poisson distribution with mean 2.9.

Find the probability that a sample of 1 ml of solution contains

(i) 0, (ii) 1, (iii) 2, (iv) more than 3 bacteria. (v) Justify the use of Poisson distribution in this problem. (5)

B) In a prestigious US university, a professor decides that the top 5% of the class will get High Distinction irrespective of the class performance. Under this scheme, which is better?

(i) Scoring 88 in a class with an average of 76 and a standard deviation of 8 OR

(ii) Scoring a 73 in a class with an average of 60 and a standard deviation of 9.

Assume that the marks follow a normal distribution.

C) It is known that, in underdeveloped countries, 40% of the women are anaemic. If a sample of 15 women are taken from this population, calculate the probability that

- a) No one will be anaemic.
- b) All will be anaemic.
- c) Exactly 10 will be anaemic.
- d) Is this a binomial distribution? Justify your answer.
- e) What will happen if our sample size is 800?

(5)

(5)

D) The CEO of a private medical Centre is trying to estimate the average number patients admitted per day in his facility. Using the past records for the last 400 days, the CEO arrived at the following figures.

No of patients admitted	No of days
0	4
1	8
5	20
10	80
15	100
20	100
30	48
40	40

- (i) On average, how many are expected to be admitted each day in the facility.(3 marks)
- (ii) On what proportion of days can the CEO expect to admit between 10 to 20 patients (both inclusive)(2)

### END OF THE EXAMINATION