S.Y. BSc. IT - SEM IV - Reg/ATKT Exam - April '18

Q.P. Code: 36161

(Time: 2½ Hours) **[Total Marks: 75]**

N. B.: (1) **All** questions are **compulsory**.
(2) Make **suitable assumptions** wherever necessary and **state the assumptions** made.
(3) Answers to the **same question** must be **written together**.
(4) Numbers to the **right** indicate **marks**.
(5) Draw **neat labelled diagrams** wherever **necessary**.
(6) Use of **Non-programmable** calculators is **allowed**.

1. Attempt *any three* of the following: **15**

a. The following table gives the heights of 100 students at XYZ College. Find the mean height of the students.

Height (in)	No. of Students
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
	100

b. During one year the ratio of milk prices per quart to bread prices per loaf was 3.00, whereas during the next year the ratio was 2.00.
   i. Find the arithmetic mean of these ratios for the 2-year period.
   ii. Find the arithmetic mean of the ratios of bread prices to milk prices for the 2-year period.
   iii. Discuss the advisability of using the arithmetic mean for averaging ratios.
   iv. Discuss the suitability of the geometric mean for averaging ratios.

c. Two variables, X and Y, assume the values $X_1 = 2$, $X_2 = -5$, $X_3 = 4$, $X_4 = -8$ and $Y_1 = -3$, $Y_2 = -8$, $Y_3 = 10$, $Y_4 = 6$, respectively. Calculate:

$$i. \sum XY, \quad ii. \sum X \sum Y, \quad iii. \sum XY^2, \quad iv. \sum X^2, \quad v. \sum (X - Y)(X + Y)$$

d. On a final examination in statistics, the mean grade of a group of 150 students was 78 and the standard deviation was 8.0. In algebra, however, the mean final grade of the group was 73 and the standard deviation was 7.6. In which subject was there the greater (i) absolute dispersion and (ii) relative dispersion?

e. State and explain the properties of standard deviation.

f. For a group of 200 candidates, the mean arid standard deviation of scores were found to be 40 and 15 respectively. Later on, it was discovered that the scores 43 and 35 were misread as 34 and 53 respectively. Find the corrected mean and standard deviation corresponding to the corrected figures.

[TURN OVER]

2. Attempt *any three* of the following: **15**

a. Find the (i) first, (ii) second, (iii) third and (iv) fourth moments about the mean of the set 2, 3, 7, 8, 10.

b. In a frequency distribution1 the co-efficient of skewness based upon the quartiles is 0·6. If the sum of the upper and lower quartiles is 100 and median is 38, find the value of the upper and the lower quartiles.

c. In a survey of 500 adults were asked the three-part question (1) Do you own a cell phone, (2) Do you own an ipod, and (3) Do you have an internet connection? The results of the survey were as follows (no one answered no to all three parts):

cell phone 329	cell phone and ipod 83
ipod 186	cell phone and internet connection 217
internet connection 295	ipod and internet connection 63

(i) answered yes to all three parts, (ii) had a cell phone but not an internet connection, (iii) had an ipod but not a cell phone, (iv) had an internet connection but not an ipod, (v) had a cell phone or an internet connection but not an ipod and, (vi) had a cell phone but not an ipod or an internet connection.

d. One bag contains 4 white balls and 2 black balls; another contains 3 white balls and 5 black balls. If one ball is drawn from each bag, find the probability that (i) both are white, (ii) both are black, and (iii) one is white and one is black.

e. Assume that the heights of 3000 male students at a university are normally distributed with mean 68.0 inches (in) and standard deviation 3.0 in. If 80 samples consisting of 25 students each are obtained, what would be the expected mean and standard deviation of the resulting sampling distribution of means if the sampling were done (i) with replacement and (ii) without replacement? Give the interpretation of the result.

f. Five hundred ball bearings have a mean weight of 5.02 grams (g) and a standard deviation of 0.30 g. Find the probability that a random sample of 100 ball bearings chosen from this group will have a combined weight of (i) between 496 and 500 g and (ii) more than 510 g. (Use the table of area under normal curve from 0 to z).

3. Attempt *any three* of the following: **15**

a. In measuring reaction time, a psychologist estimates that the standard deviation is 0.05 seconds (s). How large a sample of measurements must he take in order to be (i) 95% and (ii) 99% confident that the error of his estimate will not exceed 0.01 s?

b. A measurement was recorded as 216.480 grams (g) with a probable error of 0.272 g. What are the 95% confidence limits for the measurement?

c. A sample poll of 100 voters chosen at random from all voters in a given district indicated that 55% of them were in favor of a particular candidate. Find the (a) 95%, (b) 99%, and (c) 99.73% confidence limits for the proportion of all the voters in favor of this candidate.

d. Explain Type I and Type II errors and Level of Significance.

e. The breaking strengths of cables produced by a manufacturer have a mean of 1800 pounds (lb) and a standard deviation of 100 lb. By a new technique in the manufacturing process, it is claimed that the breaking strength can be increased. To test this claim, a sample of 50 cables is tested and it is found that the mean breaking strength is 1850 lb. Can we support the claim at the 0.01 significance level?

[TURN OVER]

f. Two groups, A and B, consist of 100 people each who have a disease. A serum is given to group A but not to group B (which is called the control); otherwise, the two groups are treated identically. It is found that in groups A and B, 75 and 65 people, respectively, recover from the disease. At significance levels of (a) 0.01, (b) 0.05, and (c) 0.10, test the hypothesis that the serum helps cure the disease. Compute the p-value and show that p-value>0.01, p-value4\>0.05, but p-value<0.10.

4. Attempt *any three* of the following: **15**

a. A random sample of 10 boys had the following I.Q.s :

70, 120, 110, 101, 88, 83, 95, 98, 107, 100.

Do these data support the assumption of a population mean I.Q. of 100? Find a reasonable range in which most of the mean I.Q. values of samples of 10 boys lie.

b. Pumpkins were grown under two experimental conditions. Two random samples of 11 and 9 pumpkins show the sample standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distributions are normal, test the hypothesis that the true variances are equal. against the alternative that they are not. at the 10% level. [Assume that $P(F_{10,8} \geq 3.35) = 0.05$ and $P(F_{8,10} \geq 3.07) = 0.05$

c. The standard deviation of the heights of 16 male students chosen at random in a school of 1000
male students is 2.40 in. Find the (i) 95% and (ii) 99% confidence limits of the standard deviation for all male students at the school.

d. Calculate the chi-square value for the following data:

Colour	Red	Green	Yellow
Observed Frequency	12	16	20
Expected Frequency	16	8	15

e. Acme Toy Company prints baseball cards. The company claims that 30% of the cards are rookies, 60% veterans but not All-Stars, and 10% are veteran All-Stars.
Suppose a random sample of 100 cards has 50 rookies, 45 veterans, and 5 All-Stars. Is this consistent with Acme's claim? Use a 0.05 level of significance. (Use chi-square goodness of fit). Given $P(\chi^2 > 19.58) = 0.0001$

f. A survey of 320 families with 5 children each revealed the following distribution:

Boys	5	4	3	2	1	0
Girls	0	1	2	3	4	5
No of families	14	56	110	88	40	12

Is this result consistent with the hypothesis that male and female births are equally probable?

[TURN OVER]

5. Attempt *any three* of the following: **15**

a. Fit an exponential curve of the form $Y = ab^x$ to the following data:

x	1	2	3	4	5	6	7	8
Y	1.0	1.2	1.8	2.5	3.6	4.7	6.6	9.1

b. The weights of a calf taken at weekly intervals are given below. Fit a straight line using the method of least squares and calculate the average rate of growth, per week.

Age (X)	1	2	3	4	5	6	7	8	9	10
Weight (Y)	52.5	58.7	65.0	70.2	75.4	81.1	87.2	95.5	102.2	108.4

c. Fit a second-degree parabola to the following data taking X as the independent variable:

X	1	2	3	4	5	6	7	8	9
Y	2	6	7	8	10	11	11	10	9

d. Find the coefficient of linear correlation between the variables X and Y presented in Table below:

X	1	3	4	6	8	9	11	14
Y	1	2	4	4	5	7	8	9

e. In a partially destroyed laboratory record of an analysis of correlation data, the following results only are legible:
Variance of $X = 9$, Regression equations: $8X - 10Y + 66 = 0$, $40X - 18Y = 214$
Find (i) Mean values of X and Y. (ii) the correlation coefficient between X and Y. (iii) the standard deviation of Y.

f. Find the equations of lines of regression for the following data:

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

Obtain the estimate of X for $Y = 70$.