CHRIST UNIVERSITY, BENGALURU - 560029

End Semester Examination March - 2016 Bachelor of Computer Applications II SEMESTER

Code: BCA233 Max.Marks: 100
Subject: STATISTICS II FOR BCA Duration: 3Hrs

SECTION A

Answer any 10 questions

10X2=20

- 1 Define regression. Give any two uses of regression.
- 2 Find the value of b_{xy} when r = 0.75, SD(X) = 5.4 and SD(Y) = 3.2.
- 3 In a bivariate data 48 is repeated thrice and 85 is repeated thrice. What is the value of the correction factor?
- 4 If n=10 and p=.4, what is the mean and variance of a binomial distribution?
- 5 Define Poisson distribution.
- 6 Mention any two situations where Poisson distribution can be applied.
- 7 Distinguish between statistic and parameter.
- **8** Write the confidence interval for single mean when variance is unknown.
- **9** Define confidence interval.
- 10 Define level of significance.
- Write the test statistic for testing equality of two means when variances are unknown.
- Write the test statistic for testing independence of attributes.

SECTION B

Answer any 5 questions

5X6=30

- 13 Given the lines of regression 4X-5Y+33=0 and 20X -9Y=107 find the means of X and Y and obtain the correlation coefficient between X and Y.
- A student is taking a multiple-choice exam in which each question has four choices. Assume that the student has no knowledge of the correct answers to any of the questions. She has decided on a strategy in which she will place four balls (marked and D) into a box. She randomly selects one ball for each question and replaces the ball in the box. The marking on the ball will determine her answer to the question. There are five multiple-choice questions on the exam. What is the probability that she will get
 - a. five questions correct?
 - b. at least four questions correct?
 - c. no questions correct?
- 15 The average number of annual trips per family to amusement parks in the United States is Poisson distributed, with a mean of 0.6 trips per year. What is the probability of randomly selecting an American family and finding the following?
 - a. The family did not make a trip to an amusement park last year.
 - b. The family took exactly one trip to an amusement park last year.
 - c. The family took two or more trips to amusement parks last year.
- The life in hours of a 75-watt light bulb is known to be normally distributed with $\sigma = 25$ hours. A random sample of 20 bulbs has a mean life of 1014 hours. Construct a 95% two-sided confidence interval on the mean life.
- 17 Ten engineering schools in the United States were surveyed. The sample contained 250 electrical engineers, 80 being women; 175 chemical engineers, 40 being women. Compute a 90% confidence interval for the difference between the proportion of women in these two fields of engineering.
- Past experience at the Crowder Travel Agency indicated that 44 percent of those persons who wanted the agency to plan a vacation for them wanted to go to Europe. During the most recent busy season, a sampling of 1,000 plans was selected at random from the files. It was found that 480 persons wanted to go to Europe on vacation. Has there been a significant shift upward in the percentage of persons who want to go to Europe? Test at .05 significance level.

Compute Karl Pearson's coefficient of correlation between per capita National Income and per capita Consumer Expenditure from the data given below:

Years	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Per capita National Income	249	251	248	252	258	269	271	272	280	275
Per capita Consumer Expenditure	237	238	236	240	245	255	254	252	258	251

It appears that over the past 45 years, the number of farms in the United States declined while the average size of farms increased. The following data provided by the U.S. Department of Agriculture show five-year interval data for U.S. farms. Use these data to develop the equation of a regression line to predict the average size of a farm by the number of farms. Discuss the slope and y-intercept of the model.

Year	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995
Number of Farms (millions)	5.65	4.65	3.96	3.36	2.95	2.52	2.44	2.29	2.15	2.07
Average Size (acres)	213	258	297	340	374	420	426	441	460	469

- The speed in which the home page of a website is downloaded is an important quality characteristic of that website. Suppose that the mean time to download the home page for the Internal Revenue Service is 1.2 seconds. Suppose that the download time is normally distributed, with a standard deviation of 0.2 second. What is the probability that a download time is
 - a. less than 2 seconds?
 - b. between 1.5 and 2.5 seconds?
 - c. above 1.8 seconds?
 - d. 99% of the download times are slower (higher) than how many seconds?
- The following data represent the number of days absent per year in a population of 5 employees of a small company: 1 3 6 7 9. Assuming that you sample with replacement, select all possible samples of size 2 and construct the sampling distribution of the mean. Compute the population mean. Also obtain the mean and variance of the sampling distribution of the mean.
- 23 Some psychologists believe that there is a statistical correlation between smoking and absenteeism. The management of a leather goods factory has under consideration 'stop smoking' incentive plans and would therefore be interested in knowing whether the employees who have stopped smoking have better absenteeism records than they had while they were smoking. 9 such employees are selected and their absenteeism records before and after they had stopped smoking is compared. The data follows.

Employee	A	В	С	D	Е	F	G	Н	I
Days absent / year (While smoking)	20	30	14	6	42	19	18	12	24
Days absent / year	10	20	16	5	40	15	22	10	20
(After stopping smoking)		120			40			10	

At level of significance 0.05, does the comparison tend to support the theory that the employees have less absenteeism after they stopped smoking?

24 Two hundred men selected at random from various levels of management were interviewed regarding their concern about environmental issues. The response of each person was tallied into one of three categories: no concern, some concern, and great concern. The results were:

		1	
Level of Management	No concern	Some concern	Great concern
Top management	15	13	12
Middle Management	20	19	21
Supervisor	7	7	6
Group leader	28	21	31

Use the .01 significance level to determine whether there is a relationship between management level and environmental concern.