ශ්‍රී ලංකා විභාග දෙපාර්තමේන්තුව
இலங்கைப் பரீட்சைத் திணைக்களம்
Department of Examinations, Sri Lanka

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2015 අගෝස්තු
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2015 ஓகஸ்ற்
General Certificate of Education (Adv. Level) Examination, August 2015

ව්‍යාපාර සංඛ්‍යානය I வணிகப் புள்ளிவிவரவியல் I **Business Statistics I**	**31 E I**	පැය දෙකයි இரண்டு மணித்தியாலம் ***Two hours***

Instructions:

- ✻ *Answer **all** questions.*
- ✻ *Write your **Index Number** in the space provided in the answer sheet.*
- ✻ *Statistical tables will be provided. Calculators are **not allowed**.*
- ✻ *Instructions are given on the back of the answer sheet. Follow those carefully.*
- ✻ *In each of the questions **1** to **50**, pick one of the alternatives from (1), (2), (3), (4), (5) which is **correct** or **most appropriate** and **mark your response on the answer sheet with a cross** (×) on the number of the correct option in accordance with the instructions given at the back of the answer sheet.*

1. Which of the following statements is true?
   (1) Data that are collected for a specific purpose are called secondary data.
   (2) In general the self enumeration method is expensive than the personal interview method.
   (3) A questionnaire is a form containing a set of questions filled out by an interviewer.
   (4) The main advantage of self enumeration method is that it ensures a high response rate.
   (5) A "more than" ogive falls downwards from left to the right.

2. Which of the following statements is/are true?
   A - Ogives for more than type and less than type intersect each other at mode.
   B - If any value of a data set is zero the value of the geometric mean of that data set is zero.
   C - The width of a class interval is measured by the difference between the upper and lower boundaries of the class interval.
   (1) C only (2) A and B only (3) A and C only (4) B and C only (5) A, B and C

3. The following is a stem and leaf diagram of the ages of employees in a small company.

stem	Leaf
2	3, 4, 6, 7, 7, 8, 8, 8, 9
3	2, 2, 3, 4, 6, 7, 8, 9
4	1, 2, 2, 3, 4
5	3, 5, 5, 6

   The first quartile age of the distribution is
   (1) 26.5 (2) 28 (3) 35 (4) 42.25 (5) 53

4. A sample of 100 observations has a mean value of 35 and a median of 35.8. It has later been discovered that an observation which was erroneously recorded as 50 actually had a value of 65. If this correction is made to the data, then mean and median are respectively
   (1) 35.00 and 35.95 (2) 35.0 and 36.15 (3) 35.15 and 35.80
   (4) 35.15 and 35.95 (5) 35.30 and 35.80

5. In a small company two typists are employed. Typist A types one page in ten minutes while typist B takes twenty minutes for the same. Both are asked to type for one hour. What is the average time taken by them for typing one page?
   (1) 6 minutes and 40 seconds. (2) 13 minutes and 20 seconds.
   (3) 14 minutes and 10 seconds. (4) 15 minutes.
   (5) 18 minutes.

[see page two

6. A consumer affairs agency wants to check the weight of a new product. A random sample of 25 items of the product was taken and the weights in grams were recorded as follows.

Weight	Number
74 - 77	3
77 - 80	6
80 - 83	9
83 - 86	3
86 - 89	4

The third quartile of the weight of the product is
(1) 18.75 (2) 75.00 (3) 83.00 (4) 83.75 (5) 84.50

7. Suppose a frequency distribution is skewed with a median 75 and mode 80. Which of the following is a possible value for the mean of the distribution?
(1) 70 (2) 75 (3) 78 (4) 80 (5) 91

8. Which of the following statements is **not** true?
(1) In a symmetric distribution, the mean and the mode are equal.
(2) The first quartile is equal to the twenty fifth percentile.
(3) In a symmetric distribution, the median is halfway between the first and the third quartiles.
(4) The median is greater than the mean in a positively skewed distribution.
(5) The range is the difference between the largest and the smallest observations in the data set.

9. Which of the following statements are/is true?
A - If 9 is subtracted from each observation of a data set, then the variance of that data set is reduced by 9.
B - If the mean salary of sewing machine operators of a garment factory is Rs. 600 per day and the standard deviation of salaries is Rs. 120, then the coefficient of variation is 20%.
C - If the coefficient of variation of a data set A is larger than that of data set B then the data set A is more consistent.
(1) A only (2) B only (3) C only (4) A and B only (5) B and C only

10. The following summary measures have been obtained for a certain set of data.
$Q_1 = 70$, $Q_3 = 96$, $M_d = 82$, $\bar{X} = 83$, $P_{90} = 107$, $P_{10} = 60$, $S^2 = 8.2$
The coefficient of kurtosis is
(1) 0.276 (2) 0.316 (3) 0.541 (4) 0.552 (5) 0.831

11. You have been given the following calculations relevant to a small population.
$\sum x_i^2 = 800$ $\mu = 9$ $\sigma^2 = 19$
What is the size of the population?
(1) 7 (2) 8 (3) 9 (4) 12 (5) 100

12. Which of the following statements are/is true about the skewness?
A - In a positively skewed distribution, mean < median < mode.
B - The median does not always lie between the mean and the mode in a skewed distribution.
C - Bowley's measure of skewness is based on quartiles.
(1) A only (2) B only (3) C only (4) A and B only (5) B and C only

13. Which of the following statements are/is true about the coefficient of determination in analysis of simple linear regression?
A - The square of the coefficient of determination is equal to the coefficient of correlation.
B - The coefficient of determination measures the proportion of the total variation in independent variable explained by the regression model.
C - The coefficient of determination measures the proportion of the total variation of dependent variable explained by the regression model.
(1) A only (2) B only (3) C only (4) A and B only (5) A and C only

[see page three

14. In regression analysis, the method of least squares
    (1) maximizes the value of the coefficient of determination.
    (2) minimizes the error sum of squares.
    (3) maximizes the error sum of squares.
    (4) minimizes the total variation in dependent variable.
    (5) minimizes the sum of errors.

15. Which of the following statements is **not** true about the least square regression line?
    (1) The point $(\bar{X}, \bar{Y})$ always lies on the regression line.
    (2) The sum of the residuals is always zero.
    (3) The sum of the observed values of Y is equal to the sum of the fitted values of Y.
    (4) The intercept in a regression model represents the predicted value of Y when $X = 0$.
    (5) There are always as many points above the regression line as there are below it.

16. Which of the following statements is/are true about the approaches of probability?
    A - The classical approach cannot be applied if the possible outcomes of the random experiment are not equally likely.
    B - One of the limitations of the relative frequency approach is that the experimental conditions may change when the experiment is repeated.
    C - The axiometic approach is not concerned with calculation of the probability of events.
    (1) A only (2) A and B only (3) A and C only (4) B and C only (5) All A, B and C

17. The probability that both events A and B occur, the probability that event A occurs and event B does not occur, and the probability that event B occurs and event A does not occur are all equal to p. The probability that at least one of the events A or B occurs is
    (1) $2p$ (2) p (3) $3p^2$ (4) $3p$ (5) p^3

18. In A and B are independent events with $P(A) < P(B)$, $P(A \cap B) = \frac{6}{25}$ and $P(A|B) + P(B|A) = 1$, then $P(A)$ is
    (1) $\frac{1}{25}$ (2) $\frac{1}{5}$ (3) $\frac{6}{25}$ (4) $\frac{2}{5}$ (5) $\frac{3}{5}$

19. Which of the following statements is true?
    (1) The conditional probability of A given B is always greater than $P(A)$.
    (2) If $P(A|B) \geq P(A)$, then $P(B|A) \leq P(B)$.
    (3) If A and B are two mutually exclusive events and none of which has probability zero, then A and B are independent events.
    (4) If A and B mutually exclusive events, then $P(A|B') = \frac{P(A)}{1 - P(B)}$.
    (5) If the events A and B are independent A' and B' cannot be independent events.

20. The random variable X has the following probability distribution.

x :	0	1	2	3	4	5	6	7
$P(x)$:	k	$2k$	$4k$	$6k$	$9k$	$10k$	$15k$	$17k$

    If k is a constant, the smallest value of x for which $P(X \leq x) > 0.5$
    (1) 2 (2) 3 (3) 4 (4) 5 (5) 6

21. Which of the following statements is true?
    (1) When X is a continuous random variable $P(a \leq X \leq b) \neq P(a < X < b)$.
    (2) A discrete random variable can take decimal values.
    (3) The expectation of a random variable cannot be a negative value.
    (4) If X is a random variable with variance σ^2, then $\text{Var}(2X + 3) = 2\sigma^2 + 3$
    (5) If X is a random variable, $Y = -X^2$ cannot be a random variable.

[see page four

22. The demand (X) for a certain item for a week in a firm has the following probability distribution.

x :	30	40	50	60	70	80
$P(x)$:	0.03	0.15	0.42	0.35	0.04	0.01

If 50 units are ordered for a week what is the probability that they will all be sold?
(1) 0.18 (2) 0.40 (3) 0.42 (4) 0.60 (5) 0.82

23. If X has a binomial distribution with mean 2 and variance 1.6, then $P(X > 2)$ is
(1) 0.3020 (2) 0.3222 (3) 0.3758 (4) 0.6242 (5) 0.6778

24. Customers enter a large business firm randomly at an average of 180 per hour. If a poisson distribution is assumed, the probability that 2 or more customers will arrive in 2 minute interval is
(1) 0.0174 (2) 0.0446 (3) 0.0620 (4) 0.9380 (5) 0.9826

25. If the random variable X is normally distributed with mean 10 and if a $P(X > 12) = 0.1587$ then $P(8 < X < 12)$ is
(1) 0.2772 (2) 0.3413 (3) 0.3830 (4) 0.6826 (5) 0.7228

26. Which of the following statements is true?
(1) Simple random sampling is a method of selecting a sample giving each unit in the population a known probability of being included in the sample.
(2) Without a complete sampling frame cluster sampling cannot be used.
(3) The standard error cannot be calculated in quota sampling because the selection of units is not based on a sampling frame.
(4) Circular systematic sampling is used when we want to select more than one systematic sample from a population.
(5) Cluster sampling is more effective if variation within clusters is larger.

27. Which of the following statements is true?
(1) The effect of the ignoring the finite population correction is to over estimate the standard error of the estimate.
(2) The difference between the estimate and the population parameter is called the precision of the estimate.
(3) In systematic sampling the term $\frac{N}{n}$ is called the sampling fraction.
(4) The standard error of the sample proportion p is given by the formula $\frac{\pi(1-\pi)}{\sqrt{n}}\sqrt{\left(\frac{N-n}{N-1}\right)}$.
(5) Failure to obtain responses from some of the units in the selected sample is an example for sampling error.

28. Which of the following statements is **not** true?
(1) Central limit theorem implies that the sampling distribution of the sample mean is approximately normal if the sample size is large.
(2) An estimator is said to be unbiased if its expected value is equal the parameter being estimated.
(3) An estimator is said to be sufficient if it contains all the information in the data about the parameter it estimates.
(4) An estimator is said to be a consistent estimator of a population parameter if it has the smallest variance among all possible estimators of the parameter.
(5) The standard error of mean of a sample taken from a given population decreases as the sample size increases.

29. Which of the following statements are/is true about the t-distribution?
A - The t-distribution is symmetric about zero.
B - The t-distribution has a larger variance than the standard normal distribution.
C - The t-distribution with K degrees of freedom has a smaller variance than the t-distribution with $K+1$ degrees of freedom.
(1) A only (2) B only (3) A and B only (4) A and C only (5) All A, B and C

[see page five

30. A random sample of size 81 is drawn from a normal population of size 501 with mean 100 and standard deviation 36. The sampling distribution of the sample mean is
(1) approximately normal with mean 100 and variance 16
(2) normal with mean 100 and variance 16
(3) approximately normal with mean 100 and variance 13.44
(4) normal with mean 100 and variance 13.44
(5) approximately normal with mean 100 and variance 3.36

31. A quality control inspector needs to test whether a machine that packages potato chips is working properly. The inspector selects a random sample of packages and weigh the content of potato chips in each. If an estimate for the mean content must be given with 98% confidence and a margin of error no more than 20 grams, what would be the minimum sample size of packages the inspector must select? Assume that the weights of potato chips in packages have a normal distribution with a standard deviation of 50 grams.
(1) 6 (2) 17 (3) 25 (4) 27 (5) 34

32. The analysis of a random sample of 300 house holds in a small town indicates that a 98% confidence interval for the mean family income is (Rs. 42,520, Rs. 49,860). Could this information be used to conduct a test of the null hypothesis $H_0 : \mu = 40{,}000$ agains the alternative hypothesis $H_1 : \mu \neq 40{,}000$ at a 0.02 level of significance?
(1) No, because it is not known whether the data are normally distributed.
(2) No, because the sample standard deviation is not known.
(3) Yes, since the sample mean Rs. 46,190 greater than Rs. 40,000 H_0 would be rejected.
(4) Yes, since Rs. 40,000 is not contained in 98% confidence interval, H_0 would be rejected.
(5) Yes, since Rs. 40000 is not contained in 98% confidence interval, H_0 would not be rejected.

33. According to past records, the mean lifetime for a certain type of battery has been 196 hours. The manufacturer has introduced a change in the production method and wants to perform a hypothesis test to determine whether the mean lifetime has increased as a result. The null hypothesis H_0 and a Alternative hypothesis H_1 for this test are
(1) H_0: $\mu \geq 196$ hours; H_1: $\mu < 196$ hours
(2) H_0: $\mu > 196$ hours; H_1: $\mu \leq 196$ hours
(3) H_0: $\mu = 196$ hours; H_1: $\mu \neq 196$ hours
(4) H_0: $\mu < 196$ hours; H_1: $\mu \geq 196$ hours
(5) H_0: $\mu = 196$ hours; H_1: $\mu > 196$ hours

34. Which of the following statements about hypothesis testing concerning the population mean is true?
(1) When the researcher does not reject a false null hypothesis, a type 1 error occurs.
(2) If there is sufficient evidence to reject a null hypothesis at the 5% level, then there is sufficient evidence to reject it at the 1% level as well.
(3) The maximum probability of a type II error that the decision maker will tolerate is called the power of the test.
(4) When the population standard deviation is unknown and if it is reasonable to assume that the population is normal, the distribution of the test statistic is also normal.
(5) If the test result yields a *p*-value of 0.034, the test is statistically significant at a level of significance of 0.05.

35. Suppose we have obtained the *p*-value after performing a *t*-test for comparison of means. Which of the following statements are/is true at a 5% level of significance?
A - If *p*-value < 0.05, we accept H_0 and reject H_1.
B - If *p*-value < 0.05, we reject H_0 and accept H_1.
C - If *p*-value < 0.05, we reject H_0 but do not accept H_1.
(1) A only (2) B only (3) C only (4) A and B only (5) A and C only

36. It has been found that 20% of the patients die during the first year from a particular disease. When treated with an experimental drug 64 out of 400 patients died during the first year. Is this an evidence to claim that the new medication reduces the mortality rate?
(1) Yes, because the *p*-value of the test is 0.0288.
(2) No, because the *p*-value of the test is only 0.0288.
(3) No, because the *p*-value of the test is 0.0912.
(4) No, because the *p*-value of the test is 0.16.
(5) No, because the sample proportion is close to 20%.

[see page six

37. Two students A and B collected data for a significant test. A found that the results were significant at 5% level while B found that the results were not significant at the same level. When checking their results, they found that they have calculated the same z-test statistic but the only difference in their analysis was that A has used a two-sided test while B has used a one-sided test. Which of the following could have been their test statistic?
(1) −2.156 (2) −1.732 (3) 1.065 (4) 1.732 (5) 1.906

38. The chi-square test for goodness of fit is used to
(1) test whether there is significant differences between two variables.
(2) test the independence of two categorical variables.
(3) test whether several categorical variables are related.
(4) test whether an observed distribution is significantly different from a relevant theoretical probability distribution.
(5) test whether the sampling distribution of the sample variance is normal.

39. Which of the following statements is/are true about the assumptions made in the analysis of variance?
A - Populations, from which samples are drawn, are normally distributed.
B - The means of the populations are equal.
C - The variances of the populations are equal.
(1) A only (2) B only (3) A and C only (4) B and C only (5) A, B and C

40. Which of the following components are/is **not** relevant if the data for a time series analysis is collected on annual basis?
(1) Trend (2) Seasonal (3) Cyclical
(4) Irregular (5) Trend and Cyclical

41. The time series trend equation for annual data is given by $y = 5.2 + 3.1x$. What is the forecast value for 2015 if the origin is at the mid of 2011? (the unit of x is six months)?
(1) 17.6 (2) 21.7 (3) 26.9 (4) 45.5 (5) 50.7

42. When the time series values are divided by the centred moving averages in the process of estimating seasonal component, which time series components are eliminated?
(1) S and I (2) T and C (3) S, C and I (4) T, C and I (5) T, S and I

43. The sum of commodity prices for the current year as percentage of the corresponding sum for base year is called;
(1) weighted average of price relatives. (2) weighted aggregate price index.
(3) simple average price index. (4) simple aggregate price index.
(5) simple aggregate quantity index.

44. If the fisher's and Paasche's price indices are 225 and 250 respectively, then the Laspeyre's price index is
(1) 90 (2) 111.1 (3) 202.5 (4) 237.1 (5) 277.7

45. Suppose your annual salary in 2010 and 2014 were Rs. 480,000 and Rs. 624,000 respectively. If the consumer price index rose from 120 to 200 during this period, then your real income and purchasing power of money in 2014 are
(1) Rs. 312,000 and 0.5 (2) Rs. 312,000 and 0.83
(3) Rs. 520,000 and 0.5 (4) Rs. 520,000 and 0.83
(5) Rs. 780,000 and 1.25

46. Which of the following conditions should be fulfilled by an index which satisfies the circular test?
(1) $P_{1/2} \times P_{2/3} \times P_{3/4} \times P_{4/3} = 1$ (2) $P_{1/3} \times P_{2/4} \times P_{3/2} \times P_{4/3} = 1$
(3) $P_{1/2} \times P_{2/3} \times P_{3/4} \times P_{4/1} = 1$ (4) $P_{1/2} + P_{2/3} + P_{3/4} + P_{4/1} = 1$
(5) $P_{1/2} + P_{2/3} + P_{3/4} = 1$

[see page seven

47. Which of the following statements about statistical process control is **not** true?
    (1) In general finding and correcting an assignable variation represent an improvement in the system.
    (2) When using a statistical control chart, a point outside the control limits is attributed to assignable causes.
    (3) If a company changed the basis for the upper and lower limits on a control chart from three standard deviations to two standard deviations then the chance of making type 1 error would increase.
    (4) An inherent range of variation in the output of a worker that occur on a routine basis represents random cause variation and is uncontrollable.
    (5) Control charts are built so that new data can be quickly compared with past performance data.

48. The mean of the measurements taken from a sample of items is within the control limits, but some items measure too low and other items measure too high, compared to specifications.
    (1) the process is in control, but with only random causes of variation and no further action need be taken.
    (2) the process is in control, but with some assignable causes of variation.
    (3) the process is in control, but not capable of producing within the specification limits.
    (4) the process is out of control but capable of producing within the specification limits.
    (5) the process is out of control and not capable of producing within the specification limits.

49. A manufacturer of light bulbs selects a random sample of 36 bulbs at the end of each shift to test for defective bulbs. The number of defectives in 25 shifts is as follows:
    3, 5, 2, 3, 4, 5, 2, 2, 4, 2, 3, 6, 3
    5, 5, 4, 3, 7, 4, 3, 5, 2, 2, 3, 3
    The upper control limit of the *P*-chart is
    (1) 0.1075 (2) 0.125 (3) 0.25 (4) 0.28 (5) 0.305

50. Which of the following statements is/are true regarding single-sampling plans?
    A - All items of a random sample of size n must be defined and tested.
    B - The lot under test will be rejected if the total defective items in the sample exceeds acceptance number "C".
    C - A rejected lot is subjected to 100% inspection or return of the lot to the producer.
    (1) A only (2) B only (3) A and B only (4) A and C only (5) A, B and C

* * *

ශ්‍රී ලංකා විභාග දෙපාර්තමේන්තුව
இலங்கைப் பரீட்சைத் திணைக்களம்
Department of Examinations, Sri Lanka

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2015 අගෝස්තු
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2015 ஓகஸ்ட்
General Certificate of Education (Adv. Level) Examination, August 2015

வ්‍යාපාර සංඛ්‍යානය II வணிகப் புள்ளிவிவரவியல் II **Business Statistics II**	**31 E II**	පැය තුනයි மூன்று மணித்தியாலம் ***Three hours***

Instructions:

* *Answer **five** questions selecting at least **two** questions from each part.*
* *Statistical tables and graph papers will be provided. Calculators are **not allowed**.*

Part I

1. (a) Describe the major limitations of the subject, statistics. (03 *marks*)

   (b) What are the shortcomings that can be identified by pretesting a questionnaire? (03 *marks*)

   (c) What is editing of data? What kind of deficiencies of data are corrected through editing? (04 *marks*)

   (d) What are the advantages and limitations of diagrammatic presentation of data? (04 *marks*)

   (e) Following data shows the monthly sales in million rupees of a business firm for 2014 and moving annual totals.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sales Moving	07	09	11	13	16	20	35	44	16	08	08	07
Moving Annual Totals	165	166	166	167	169	174	184	188	190	191	192	194

   Construct a Z-chart and comment on the sales of the firm. (06 *marks*)

2. (a) Explain the role of following measures in data analysis.
   (i) Measures of Central tendency.
   (ii) Measures of dispersion
   (iii) Measures of skewness
   (iv) Measures of Kurtosis (04 *marks*)

   (b) What do you mean by weights in a weighted average? Describe the uses of a weighted average through an example. (04 *marks*)

   (c) The following table shows the number of male and female workers in two factories *A* and *B* together with their salaries.

Sex	Factory *A*		Factory *B*	
	Number	Salary per month	Number	Salary per month
Male	250	Rs. 18 000	650	Rs. 17 000
Female	750	Rs. 15 000	350	Rs. 14 000

   (i) Determine the average monthly salary for each factory.

   (ii) Give reasons to have higher average monthly salary in factory B, even though in factory B the monthly salary of both categories of workers is lower than factory A. (04 *marks*)

   (d) The following table shows the lifetime of two models of computers.

Lifetime (No. of years)	Number of computers	
	Model A	Model B
0 - 2	06	02
2 - 4	15	07
4 - 6	12	10
6 - 8	08	19
8 - 10	05	11
10 - 12	04	01

   Calculating Bowley's coefficient of skewness for each model, decide which model has the less skewed distribution? (08 *marks*)

[see page two

3. (a) Define the Laspeyer's and Paasche's price indices and discuss their relative advantages and disadvantages. (05 *marks*)

(b) The prices of some consumer goods for the years 2010 and 2014 are given in the following table.

Item	Unit	Price (Rs)	
		2010	2014
Wheat	kilograms	150	180
Milk	litre	60	72
Egg	dozen	125	200
Sugar	kilograms	85	108

(i) Calculate the simple aggregate price index and the simple average index of price relatives for the year 2014 using 2010 as base year.

(ii) What are the advantages and disadvantages of the two price indices? (05 *marks*)

(c) Describe briefly the following methods in estimating trend stating advantages and disadvantages of each method.

(i) freehand method

(ii) moving average method

(iii) least square method (06 *marks*)

(d) The time series data on production (in '000 units) of a certain company are given in the following table.

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Production ('000 units)	45	48	55	63	65	72	84	90	87	82

(i) obtain the trend line using semi-average method.

(ii) forecast the trend value for year 2016. (04 *marks*)

4. (a) Distinguish between Product Moment Correlation Coefficient and Rank Correlation Coefficient, stating the formula for each coefficient. (03 *marks*)

(b) Fifteen observations were obtained to estimate a simple regression model and in the analysis of data, the following summations were calculated.

$\sum X = 45, \quad \sum Y = 105, \quad \sum XY = 500, \quad \sum X^2 = 250, \quad \sum Y^2 = 1100$

(i) Find the least squares regression line.

(ii) Interpret the regression coefficient.

(iii) Compute the coefficient of determination and interpret it.

(iv) Predict the average value of the dependent variable if the value of the independent variable is 40. (06 *marks*)

(c) A company produces dials for a machine. These dials are supposed to have a constant diameter. To check on the production process, the first 4 dials were selected every half an hour for 12 hours giving a total of 96 observations. It was found that $\sum X = 4896$ mm and $\sum R = 144$ mm. Find the upper and lower control limits for $\bar{X}$-chart and R-chart. If both the $\bar{X}$-chart and the R-chart are out of control, which chart should be dealt with first? Explain. (05 *marks*)

(d) (i) Describe how you would construct an operating characteristic curve of an acceptance sampling plan. Explain the importance of the operating characteristic curve in quality control.

(ii) Consider a single sampling plan with a lot size of 2500, sample size of 100, and acceptance number of 2. If acceptance quality level is 0.01 and lot tolerance percent defective is 0.06, find the producer's risk and the consumer's risk using poisson approximation. (06 *marks*)

[see page three

Part II

5. (a) (i) What do you mean by collectively exhaustive events? Give an example for **two** events, which are collectively exhaustive but not mutually exclusive.

(ii) A and B are two collectively exhaustive events and it is known that $P(A|B) = 0.4$ and $P(B) = 0.7$.

Find $P(A \cap B)$ and $P(A)$. (06 *marks*)

(b) The employees of a company have been classified according to age and salary as shown in the following table.

Age (years)	Salary			Total
	Under Rs. 30000	Rs. 30000 - Rs. 50000	Over Rs. 50000	
Under 30	28	12	5	45
30 - 45	16	26	18	60
Over 45	5	18	12	35
Total	49	56	35	140

One employee is selected at random. If two events A and B are defined as follows, verbally describe the events $A \cup B$, $A \cap B'$ and $A'|B$

A - The employee is under 30 years of age.

B - The employee's salary is under Rs. 30000

Find the following probabilities.

(i) $P(A)$ (ii) $P(A \cap B')$ (iii) $P(A \cup B)$

(iv) $P(B|A)$ (v) $P(A|B')$ (vi) $P(A'|B)$ (08 *marks*)

(c) A doctor recommends a test for a particular disease for a patient having a certain symptom. Before the result of the test, the only evidence the doctor has to go on is that 10 percent having this symptom has the disease. Past experience has shown that in 99 percent of the cases in which disease is present the test reveals the presence of the disease. In 95 percent of the cases in which the disease is not present the test reveals the absence of the disease.

(i) What is the probability that the test reveals the presence of the disease?

(ii) If the test reveals that the disease is present, what is the probability that the patient is actually having the disease? (06 *marks*)

6. (a) (i) Define the binominal distribution. State the conditions which should be satisfied to use the binominal distribution as a probability model.

(ii) A producer claims that in a lot at most 10 percent of the items is defective. To test this claim 15 items are selected at random and the claim is accepted if at most 2 items are defective. If the actual probability that an item is defective is 0.10, find the probability that the producer's claim will be accepted. (06 *marks*)

(b) (i) State the conditions under which poisson distribution may be approximated by normal distribution.

(ii) The weekly demand for a certain item in a particular shop follows a poisson distribution with mean 25. How many of this item should be kept at the beginning of the week so that 95% certain of meeting the demand for that week? (05 *marks*)

(c) (i) Define the poisson distribution and write down its mean and variance.

(ii) In a journey of X km the number of times a certain component of a vehicle needs repair has a poisson distribution with a mean $\frac{X}{1000}$. If the driver is setting out on a journey which involves 3000 km, what is the probability that he has to repair the component at least 3 times? (05 *marks*)

(d) The life time (in hours) of a certain electrical equipment has a normal distribution with mean 30 and variance σ^2. A purchaser of this item requires that at least 90% of the equipments have life times exceeding 150 hours. If the distribution satisfies the purchaser's requirement what should be the largest value of σ? (04 *marks*)

[see page four

7. (a) What is quota sampling? How is quota sampling different from stratified random sampling. Describe **three** arguments for using quota sampling and **three** arguments against quota sampling. (06 *marks*)

(b) Describe the method of systematic sampling. Explain **two** advantages and **two** disadvantages of systematic sampling. Discuss the efficiency of systematic sampling for the following structures of populations.

(i) Populations in random order

(ii) Populations with linear trend

(iii) Populations with periodic variations (07 *marks*)

(c) (i) State the Central Limit Theorem. Explain how is it useful in statistical inference.

(ii) The length of nails produced by a machine has an unknown distribution with mean 2.03 cm and standard deviation 1.5 cm. If a random sample of 100 nails is taken, approximate the probability that the mean length of the sample is less than 3 cm. Is this probability different if the distribution is normal? Explain. (07 *marks*)

8. (a) Explain the properties of a good estimator. (04 *marks*)

(b) A manufacturer of ball point pens claims that a certain type of pen he manufactures has at least a mean writing life 400 pages. A purchasing agent selects a sample of 100 pens and puts them for test. The mean writing life for the sample was 390 pages. If the standard deviation of writing life is 20 pages, should the purchasing agent rejects the manufacturer's claim at 5% level of significance. (04 *marks*)

(c) Two brands of refrigerators A and B each has warranty period of one year. In a random sample of 50 refrigerators of brand A, 12 were observed to fail before the warranty period ended. A random sample of 60 brand B refrigerators also revealed 12 failures during the warranty period.

(i) Estimate the true difference between proportions of failures during the warranty period at 98% confidence level.

(ii) Does the confidence interval reveal that there is no difference between the failure proportions of two types refrigerators? Give reasons for your answer. (05 *marks*)

(d) The following table shows the data concerning the number of societies a student is affiliated with at the school from a sample of 94 students of a certain school.

Number of Societies	Frequency
0	43
1	24
2	16
3	07
4	03
5	01
6 or more	00
Total	94

(i) Fit a poisson distribution for this frequency distribution.

(ii) Test the goodness of fit of the poisson distribution at 5% level of significance. (07 *marks*)

* * *