නව නිර්දේශය/புதிய பாடத்திட்டம்/New Syllabus

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2020
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2020
General Certificate of Education (Adv. Level) Examination, 2020

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. Use of calculator is **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) Making a statistical conclusion with probable error is an example for misusing of statistics.
(2) The difference between the population parameter and the estimate for the parameter is called the sampling error.
(3) Non-response error is an example for non-sampling errors.
(4) If in a study answers are needed to many complicated questions, self-enumeration method is the most suitable method.
(5) In focus group interviews, participants have to answer questions included in a pre-prepared document.

2. Consider the following statements.

A - Pie diagram is a circular representation of the data which can also be represented by a single bar of a percentage component bar-chart.
B - A Lorenz curve can not be constructed when the income is given in the form of a frequency distribution.
C - If the Gini coefficient is zero, it indicates that everyone in a population receives the same amount of income.

The true statement/s from the above is/are,

(1) only A (2) only C
(3) only A and B (4) only A and C
(5) all A, B and C

3. Consider the following statements regarding scales of measurements.

A - Categorising of social classes as upper class, middle class and lower class is an example for nominal scale.
B - The main difference between ordinal scale and interval scale is that interval scale uses unit of measurements.
C - The class interval in a frequency distribution is an example for interval scale.

The true statement/s from the above is/are,

(1) only B (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

[see page two

4. Which of the following statements is true?

(1) Stem-and-leaf diagram cannot be constructed for observation with decimal values.
(2) If the boxes of box-and-whiskers plot are equal, the distribution is exactly symmetric.
(3) In a box-and-whiskers plot, the right whisker contains the largest 25% of values.
(4) Multiple bar diagram is an alternative way of representing data which can also be represented by component bar chart.
(5) If class intervals are unequal, the area under the frequency polygon is not equal to the sum of the areas of the rectangles in a histogram.

5. Milk is sold at the prices of 60, 100, 120 and 150 rupees per liter in four consecutive different months respectively. If a family spent Rs. 600 on milk per month in the four months, the average price in rupees per liter of milk for the family per month was,

(1) Rs. 96.00 (2) Rs. 102.00 (3) Rs. 107.50 (4) Rs. 110.00 (5) Rs. 150.00

6. The geometric mean of 10 observations on a certain variable was calculated as 14.2. It was later discovered that the observed value 21 had been taken as 12 in calculation. The corrected geometric mean is,

(1) $14.2(1.75)^{1/10}$ (2) $14.2(0.57)^{1/10}$ (3) $(24.85)^{1/10}$
(4) 8.11 (5) 24.85

7. Consider the following statements about dispersion.

A - Mean deviation is less affected by the extreme values as compared to standard deviation.
B - If a constant is added to all values of a data set, the coefficient of variation of the resulting values will not change.
C - Both variance and standard deviation have same unit of measurements.

The true statement/s from the above is/are,

(1) only A (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

8. Consider the following stem-and-leaf diagram.

0	8							
1	3	4	5					
2	0	1	2	3	3	5		
3	0	0	1	4	6	7	8	8
4	0	0	2	4	7	7	8	8
5	0	2	3					

For this distribution Kelly's coefficient of skewness is,

(1) – 0.29 (2) – 0.23 (3) – 0.22 (4) 0.22 (5) 0.23

9. The value 5.1 is subtracted from each of 100 observations. The sum of the deviations and the sum of square of deviations are –10 and 401 respectively. The coefficient of variation of the distribution is

(1) 13% (2) 38% (3) 40% (4) 78% (5) 80%

10. Which of the following statements is **false**?

(1) The quartile deviation is not affected by the extreme values.
(2) The sum of deviation of a data set from its mean is always zero.
(3) Pearson's coefficient of skewness cannot be calculated when a distribution has open-ended classes.
(4) For a symmetric distribution the median is the mean of the first quartile and the third quartile.
(5) For a normal distribution the percentage kurtosis coefficient is zero.

[see page three

11. In a distribution Pearson's coefficient of skewness is 0.5, the coefficient of variation is 40% and the mode is 80. The mean of the distribution is,

(1) 40 (2) 100 (3) 160 (4) 200 (5) 320

12. Bowley's coefficient of skewness for a certain distribution is -0.8. If the sum of the lower and upper quartiles is 100 and the median is 58, find the lower and upper quartiles.

(1) $Q_1 = 20$, $Q_3 = 80$ (2) $Q_1 = 25$, $Q_3 = 75$
(3) $Q_1 = 30$, $Q_3 = 70$ (4) $Q_1 = 35$, $Q_3 = 65$
(5) $Q_1 = 40$, $Q_3 = 60$

13. Which of the following statements is true about regression and correlation?

(1) High correlation between X and Y implies that X causes change in Y and Y causes change in X.
(2) If an additional independent variable is included in a simple linear regression equation, the error term will increase.
(3) If a constant is subtracted from all the values of X, the regression coefficient of Y on X will change.
(4) Spearman's rank correlation coefficient cannot be calculated for quantitative data.
(5) Spearman's rank correlation coefficient is equal to the Karl Pearson's product moment correlation coefficient between ranks.

14. Consider the following statements regarding the regression.

A - If the regression of Y on X is simple linear, the expected value of Y vary exactly on a straight line when X vary.

B - If the regression equation of Y on X is $\hat{Y} = -10 + 5x$, then the regression equation of X on Y is $\hat{X} = 0.2y - 2$.

C - If the estimated regression equation is $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$, the coefficient $\hat{\beta}_2$ represent the change of $\hat{Y}$ for a unit change of x_2 if x_1 is kept constant.

The true statement/s from the above is/are,

(1) only A (2) only B
(3) only A and B (4) only A and C
(5) all A, B and C

15. The fitted regression equation for paddy yield (Y) on fertilizer (X) is given below.

$\hat{Y} = 36.4 + 0.05x$

The true statement with respect to the above equation is,

(1) $\hat{Y}$ is the paddy yield when fertilizer level is x.
(2) $\hat{Y}$ is the expected value of paddy yield when fertilizer level is x.
(3) $\hat{Y}$ is the estimate of the expected value of paddy yield when fertilizer level is x.
(4) If the fertilizer level is increased by one unit, the paddy yield will increase by 36.45 units.
(5) If the fertilizer level is increased by one unit, the paddy yield will increase by 36.4 units.

16. Consider the following statements about the approaches to probability.

A - Subjective probability approach is more appropriate when the experiment cannot be repeated.

B - If the random experiment is to select an item at random from a lot, the probability of an event can be obtained without performing the experiment.

C - The probability of an event obtained under the relative frequency approach may be different from the true probability of the event.

The true statement/s from the above is/are,

(1) only B (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

[see page four

17. In a group there are 3 boys and 2 girls. If 3 are selected at random from this group, find the probability that there are 2 boys and 1 girl or 1 boy and 2 girls.

(1) $\frac{1}{5}$ (2) $\frac{3}{10}$ (3) $\frac{1}{2}$ (4) $\frac{3}{5}$ (5) $\frac{9}{10}$

18. Suppose A and B are two independent events. The probability that both A and B occur is $\frac{1}{8}$ and the probability that neither of them occur is $\frac{3}{8}$. If $P(A) > P(B)$, the probability of the occurrence of A is,

(1) $\frac{1}{5}$ (2) $\frac{1}{4}$ (3) $\frac{1}{3}$ (4) $\frac{1}{2}$ (5) $\frac{3}{4}$

19. Suppose that A and B are any two events. The probability that both A and B occur, the probability that A occurs and B does not occur, and the probability that B occurs and A does not occur are all equal to k. The probability that at least one event occurs from the events A and B is,

(1) k (2) 2k (3) 3k (4) $3k^2$ (5) k^3

20. If A and B are any two events with $P(A) = p_1$, $P(B) = p_2$ and $P(A \cap B) = p_3$ then $P(A'|B')$ is,

(1) $\frac{p_1 + p_2 - p_3}{1 - p_1}$ (2) $\frac{p_1 + p_2 - p_3}{1 - p_2}$

(3) $\frac{1 - p_1 - p_2 + p_3}{1 - p_2}$ (4) $\frac{1 - p_1 - p_2 + p_3}{1 - p_1}$

(5) $\frac{1 - p_1 - p_2 - p_3}{1 - p_2}$

21. Consider the following statements.

A - The expected value of a random variable X is a weighted average with weights being the probabilities of possible values of X.

B - The expected value of a random variable is the value which occurs with maximum probability.

C - If X is a random variable and c and d are constants $Var(cX \pm d) = cVar(X) \pm d$.

The true statement/s from the above is/are,

(1) only A (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

22. The manufacturer of nails says that on the average 2.5% of his product is defective. A buyer accepts a box of 100 nails if it contains not more than 4 defective nails. Find the approximate probability that a box of nails will be accepted by the buyer.

(1) 0.1088 (2) 0.2424 (3) 0.5438 (4) 0.7576 (5) 0.8912

23. A student sit for a test consists of 10 multiple choice questions with 5 answers to each question. The student answers each question by selecting one answer randomly as the correct answer. To pass the exam he must get at least 60% correct answers. What is the probability that the student passes the exam?

(1) 0.0064 (2) 0.0328 (3) 0.9672 (4) 0.9936 (5) 0.9991

24. The yield of Potato in 10000 plots of land has a normal distribution with mean 650kg and standard deviation 30kg. The lowest yield of the best 1000 plots is,

(1) 578kg (2) 612kg (3) 688kg (4) 719kg (5) 962kg

25. In an area 50% of the persons favour a certain development proposal. If random sample of 100 persons selected from this area, what is the approximate probability that at least 55 persons favour the proposal?

(1) 0.1587 (2) 0.1841 (3) 0.3159 (4) 0.3413 (5) 0.3682

[see page five

26. Consider the following statements about the cluster sampling.

A - If the variation among clusters is small cluster sampling is more suitable.
B - Cluster sampling can be used even when a complete sampling frame is not available.
C - If the intra-cluster correlation coefficient is close to 1 cluster sampling is as efficient as simple random sampling.

The true statement/s from the above is/are,

(1) only A (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

27. Consider the following statements about sampling.

A - The results of a sample survey may be more reliable than the results of a whole population survey.
B - Method of systematic sampling can be used only when $\frac{N}{n}$ is an integer.
C - In systematic sampling the standard error cannot be calculated using a single sample.

The true statement/s from the above is/are,

(1) only A (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

28. In simple random sampling with replacement from a population with the population proportion π the standard error of the sample proportion of sample size n is,

(1) $\sqrt{\left(\frac{N-n}{N-1}\right)\frac{\pi(1-\pi)}{n}}$ (2) $\frac{\pi(1-\pi)}{\sqrt{n}}$ (3) $\frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}}$

(4) $\sqrt{\left(\frac{N-n}{N}\right)\frac{\pi(1-\pi)}{n}}$ (5) $\frac{\sqrt{\pi(1-\pi)}}{n}$

29. Which of the following statements is true?

(1) If the sample size n is small $t-$distribution depends on the mean of the normal distribution.
(2) The standard deviation of a sampling distribution of an estimator is called the standard error of the estimator.
(3) When degrees of freedom increases χ^2-distribution becomes more skewed.
(4) If the sample size is small the sampling distribution of the sample proportion is unknown.
(5) The central limit theorem can be used only for deciding the sampling distribution of the sample mean.

30. Which of the following statements is true?

(1) If $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators for the parameter θ, the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is defined as $\frac{Var(\hat{\theta}_1)}{Var(\hat{\theta}_2)}$
(2) A biased estimator cannot be a consistent estimator.
(3) If the population mean μ is known, $\frac{1}{n}\sum(x_i-\mu)^2$ is an unbiased estimator for population variance σ^2.
(4) The difference between the population parameter and the estimate for the parameter is called bias of the estimator.
(5) Any function of a random sample is called a statistic.

31. It is required to estimate population proportion π from the sample proportion p with maximum probable error within the range $\pi \pm 0.02$ with probability 0.9544. What is the sample size required to satisfy this requirement?

(1) 900 (2) 1681 (3) 1785 (4) 2401 (5) 2500

[see page six

32. In a random sample of size 25 from a normal population with mean μ and unknown variance σ^2, the sample mean was $\bar{x} = 60$ and the sample variance s^2=16. If the calculated confidence interval for μ is (57.76, 62.24), what is the confidence level?

(1) 80% (2) 90% (3) 95% (4) 98% (5) 99%

33. Which of the following statements is **false**?

(1) The variable which is used to construct confidence intervals for a population parameter contains the parameter and a point estimator for the parameter.

(2) The confidence interval with 99% confidence level is better than confidence interval with 95% confidence level.

(3) A confidence interval for the difference of population means $\mu_1 - \mu_2$ is obtained by adjusting the point estimator $\bar{X}_1 - \bar{X}_2$ for its probable error.

(4) For the mean of a normal population there can be more than one 95% confidence intervals.

(5) The 95% confidence interval for the mean of a normal population with unknown variance is wider than 95% confidence interval with known variance.

34. Consider the following statements about testing hypotheses.

A - The hypothesis test with the minimum probability of type II error for a given value of probability of type I error is called the most powerful test.

B - The probability of accepting H_1 hypothesis when H_0 hypothesis is false is called the power of the test.

C - The sampling distribution of a test statistic depends on the population parameter being tested.

The true statement/s from the above is/are,

(1) only A (2) only A and B

(3) only A and C (4) only B and C

(5) all A, B and C

35. For testing the population proportion H_0: $\pi = 0.1$ against H_1: $\pi \neq 0.1$ at 5% significance level a random sample of size 100 was obtained. If the sample proportion is $p = 0.16$, the conclusion is,

(1) reject H_0 since $p-$value = $0.0228 < 0.05$

(2) reject H_0 since $p-$value = $0.0456 < 0.05$

(3) do not reject H_0 since $p-$value = $0.0526 > 0.05$

(4) do not reject H_0 since $p-$value = $0.2104 > 0.05$

(5) reject H_0 since $Z = 1.62 < 1.96$

36. The mean and variance of the life time of type *A* bulbs of a random sample of size 120 were $\bar{x} = 945$ and $s_1^2 = 240$ and the mean and variance of the life time of type *B* bulbs of a random sample of 100 were $\bar{y} = 940$ and $s_2^2 = 200$. If the critical region for testing the equality of population means H_0: $\mu_1 = \mu_2$ against H_1: $\mu_1 > \mu_2$ is given by $\bar{X}$-$\bar{Y} > 4$, the probability of type I error is

(1) 0.0228 (2) 0.0250 (3) 0.1103 (4) 0.3897 (5) 0.4772

37. The critical region for testing mean of a normal population H_0: μ=120 against H_1: μ=122 by taking a random sample of size 60 is given by $\bar{X} > 121.4$. If the population variance is $\sigma^2 = 240$ the power of the test is

(1) 0.1179 (2) 0.2420 (3) 0.3821 (4) 0.6179 (5) 0.8821

38. The number of deaths occurred in 50 days in a certain hospital are given in the following table.

No. of deaths	0	1	2	3	4	5	6	7
No. of days	2	8	12	13	8	4	2	1
Expected frequency	3	8	11	11	8	5	3	1

The critical value for testing at 5% significance level whether the poission distribution with mean 3 is the relevant distribution for this data is,

(1) 7.82 (2) 9.50 (3) 11.10 (4) 12.60 (5) 14.10

[see page seven

39. Three samples of size 5 each were randomly taken from three varieties of paddy to test whether the mean yields are equal. The calculated sum of squares are as follows;

Total sum of squares SST=224

Between varieties of paddy sum of squares SSB=128

The critical region for testing the hypothesis that population means are equal at 5% level is given by,

(1) $F=5.34>3.49$ (2) $F=8>3.89$

(3) $F=8<19.4$ (4) $F=8>3.34$

(5) $F=8.67>3.81$

40. The mean of the first five years of a time series from 2010 to 2019 is 32.6 and the mean of the second five years is 42.6. The trend equation obtained by the method of semi-average is,

(1) $\hat{Y}=26.6+t$ (2) $\hat{Y}=26.6+2t$

(3) $\hat{Y}=29.24+1.67t$ (4) $\hat{Y}=32.6+2t$

(5) $\hat{Y}=42.6+t$

41. Consider the following statements about moving averages.

A - Moving averages may generate movements which were not present in the original data.

B - Moving averages are not useful in forecasting the values of a variable on the basis of time.

C - Moving averages cannot be used to estimate trend if the shape of the appropriate trend line or curve is unknown.

The true statement/s from the above is/are,

(1) only A (2) only A and B

(3) only A and C (4) only B and C

(5) all A, B and C

42. In a time series with no cyclical variation from 2010 to 2014, the five year totals of trend eliminated percentage for each quarter are given below,

Q_1	Q_2	Q_3	Q_4
450	550	525	500

The seasonal indices for the first quarter and the third quarter are approximately given by,

(1) 89 , 104 (2) 90 , 105 (3) 91 , 106 (4) 92 , 107 (5) 101 , 106

43. Trend line fitted for the yearly data of the time series 2010 - 2014 is given by $\hat{Y}=50+16t$ (t=0 for 2012). If the actual value for the second quarter of the year 2013 is 72, the trend eliminated percentage value for this quarter is approximately

(1) 103 (2) 104 (3) 109 (4) 113 (5) 116

44. If all the sample points lie within control limits in $\overline{X}$ – chart but exhibit some trend by these points it indicates that,

(1) process is in control.

(2) presence of chance causes.

(3) variability has increased.

(4) presence of assignable causes.

(5) presence of errors in selecting random samples.

45. If the inspection unit varies in size, the statistical chart which should be used to control the number of defects per unit of a product is,

(1) p - chart (2) C - chart (3) U - chart (4) $\overline{X}$ - chart (5) np - chart

[see page eight

46. Consider the following statements about *OC*-curve.

A - *OC*-curve shows the ability of an acceptance sampling plan to distinguish between good lots and bad lots.

B - If the *OC*-curve for a product is not satisfactory it can be improved by changing the sample size and acceptance number.

C - *OC*-curve represents the probability of rejecting the lot for variations in the fraction defective in the lot.

The true statement/s from the above is/are,

(1) only A (2) only A and B
(3) only A and C (4) only B and C
(5) all A, B and C

47. Consider an accepting sampling plan with N=1 000, n=100 and acceptance number c=1. If AQL = 0.01 and LTPD = 0.07, the respective consumer's risk and producer's risk are,

(1) 26.42%, 0.73% (2) 26.42%, 99.27%
(3) 36.79%, 0.09% (4) 63.21%, 0.09%
(5) 73.58%, 0.73%

48. A certain firm expects its sale for an item to increase by 50% in next year. If the aim of the firm is to double the gross income, by what percentage should it increase the selling price?

(1) 30% (2) $33\frac{1}{3}\%$ (3) 50% (4) 100% (5) 150%

49. Consider the following statements about index numbers.

A - Marshall-Edgeworth index satisfies the time reversal test but it does not satisfy the factor reversal test.

B - In a situation where prices are increasing Laspeyers's price index tends to be smaller than Paasche's price index.

C - The simple aggregate price index does not take into account the relative importance of various commodities.

The true statement/s from the above is/are,

(1) only A (2) only C
(3) only A and B (4) only A and C
(5) all A, B and C

50. Two raw materials (I and II) are used in different proportions for the production of goods A and B, but the price of each raw material is equal for both products.

	Product A	Product B
Weight for raw material I (w_1)	60	70
Weight for raw material II (w_2)	40	30
Index of production Cost	170	165

The price indices for raw material I and II are respectively,

(1) 15 , 20 (2) 50 , 45 (3) 64.5 , 187.5 (4) 150 , 200 (5) 285 , 235

* * *

නව නිර්දේශය/புதிய பாடத்திட்டம்/New Syllabus

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2020
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2020
General Certificate of Education (Adv. Level) Examination, 2020

ව්‍යාපාර සංඛ්‍යානය II வணிகப் புள்ளிவிவரவியல் II **Business Statistics II**	**31**	**E** **II**

පැය තුනයි
மூன்று மணித்தியாலம்
Three hours

අමතර කියවීම් කාලය - මිනිත්තු 10 යි
மேலதிக வாசிப்பு நேரம் - 10 நிமிடங்கள்
Additional Reading Time - 10 minutes

Use additional reading time to go through the question paper, select the questions you will answer and decide which of them you will prioritise.

Instructions:

* *Answer **five** questions selecting at least **two** questions from each part.*
* *Statistical tables and graph papers will be provided. Use of calculators is **not allowed**.*

Part I

1. (a) Explain the difference between pre-test and editing of a completed questionnaire. Describe **three** weaknesses that could be identified through pre-test and through editing of a completed questionnaire. (04 *marks*)

(b) State whether the following statements are true or false and explain your answers.

(i) Interval scale has units of measurements which vary from scale to scale between arbitrary starting point and terminating point but no mathematical operations can be done except taking the sum and difference.

(ii) Classification of students by their reading ability as Below Average, Average and Above Average is an example for interval scale.

(iii) Variables are classified as qualitative and quantitative considering the scales of measurements. (03 *marks*)

(c) State the most appropriate method of data collection for each of the following studies. Describe how the each method is applied giving one advantage and one disadvantage.

(i) Explore various aspects of the issue of domestic violence by conducting discussions with experts in the field and victims.

(ii) Study the efficiency of nurses in carrying out of their functions. (03 *marks*)

(d) The following stem-and-leaf diagrams represent the marks of Mathematics of students in two classes A and B of grade 10 in a school.

	Class A						
3	2	3	4	5			
4	1	3	4	4	5	6	7
5	0	2	3	4	5	7	8
6	2	3	4	5	6		
7	4	5	5				
8	6						

	Class B					
4	2	3				
5	3	4	5	6	8	8
6	1	4	6	7	8	9
7	0	3	3	7	7	8
8	0	2	6	7	9	
9	6	7				

Draw Box-and-Whiskers plots on a same graph and compare the performance of students of two classes in the subject of Mathematics. (06 *marks*)

[see page two

(e) The hourly payments of 70 workers are given in the following table.

Wage	No. of workers
60-69	8
70-79	10
80-89	15
90-99	16
100-109	10
110-129	8
130-189	3

Construct the histogram for the above frequency distribution and highlight the area corresponding to the payment 90–119. (04 *marks*)

2. (a) What are the properties of a good average? Describe these properties with reference to mean, median and mode. (04 *marks*)

(b) Define the geometric mean of a data set.
If the sales of a certain firm doubles in a period of 10 years, what is the average percentage growth rate per year? (05 *marks*)

(c) In a national examination the mean of the marks scored for a certain subject was 50 and the standard deviation was 10. In the following year for the same subject the mean was 60 while standard deviation increased to 15. Calculate an appropriate measure and compare the performance of the students of two years. (03 *marks*)

(d) The marks obtained by 100 students in an examination are given in the following frequency distribution.

Marks	No. of students
0-9	6
10-19	8
20-29	10
30-39	12
40-49	20
50-59	25
60-69	10
70-79	9

Compute Pearson's first coefficient of skewness and second coefficient of skewness. Comment on the shape of the distribution using your results. (08 *marks*)

3. (a)(i) "It is sometimes stated that Laspeyers's price index tends to overestimate price changes while Paasche's index tends to underestimate price changes." Explain this statement giving reasons. (02 *marks*)

(ii) Explain what is meant by time reversal test and factor reversal test. Show that the Marshall-Edgeworth Price Index satisfies the time reversal test. (03 *marks*)

[see page three

(iii) The prices and quantities of commodities A, B, C and D for years 2016 and 2018 are given in the following table.

Commodity	2016		2018	
	Price	**Quantity**	**Price**	**Quantity**
A	10	8	20	6
B	25	10	30	5
C	20	15	25	15
D	10	20	10	25

Verify that Marshall-Edgeworth Price Index is a good approximation for the Fisher's Ideal Price Index by calculating these indices for the year 2018 taking 2016 as the base year. Explain the reasons for this in your own words. (05 *marks*)

(b)(i) Explain what is meant by trend of a time series. Describe the method of semi-average and the method of moving average for estimating trend. (03 *marks*)

(ii) The quarterly sales (in thousand Rupees) values of a certain item for the years 2015, 2016 and 2017 are given in the following table. The values given in brackets are trend values.

Year	Quarter			
	Q_1	Q_2	Q_3	Q_4
2015	6 (12)	15 (15)	15 (15)	20 (18)
2016	15 (18)	20 (20)	25 (20)	30 (25)
2017	25 (25)	30 (25)	27 (30)	25 (35)

Estimate the seasonal indices by the method of Ratio to Trend. If the actual sale for the first quarter of 2018 is Rs.100 000 what is the expected sales for the fourth quarter? (07 *marks*)

4. (a) An office clerk recorded the time in minutes **Y** taken to travel to the office when he leaves home **X** minutes after 6 a.m. on 8 randomly selected days. The results are as follows,

X	0	5	10	15	20	25	30	35
Y	20	25	39	35	40	45	46	50

$\sum X = 140$ $\sum Y = 300$ $\sum X^2 = 3500$ $\sum Y^2 = 12012$ $\sum XY = 6095$

(i) Fit the regression line of **Y** on **X** by using the method of least squares and explain the meaning of the regression coefficient.

(ii) Calculate the coefficient of determination and comment on the goodness of fit. (05 *marks*)

(b) Ten competitors in a musical contest were ranked by two judges in the following order.

Judge A	4	8	7	6	5	9	10	3	2	1
Judge B	6	7	8	1	5	10	9	2	3	4

Calculate the Spearman's rank correlation coefficient and Karl Pearson's product moment correlation coefficient between ranks and verify that both answers are equal. Explain whether the two judges agree in their judgement. (05 *marks*)

[see page four

(c) A producer receives parts in large lots and it has been decided to use an acceptance sampling plan. The following acceptance sampling plans are considered.

Plan I – Examine a random sample of size 50 and accept the lot if the acceptance number $c \leq 1$.

Plan II– Examine a random sample of size 100 and accept the lot if acceptance number $c \leq 2$.

(i) Calculate the acceptance probabilities of lots for each plan at defective percentage 1%, 2%, 5% and 7%.

(ii) Plot the values obtained in (i) for each plan on the same graph.

(iii) If it is required to have a sampling plan with 95% acceptance at 2% defective percentage and 5% acceptance at 7% defective percentage which plan is close to this requirement? (07 *marks*)

(d) The number of defectives in 10 samples of 100 item each are given below.

Sample No.	1	2	3	4	5	6	7	8	9	10
No. of defectives	8	4	12	3	12	8	8	15	12	8

Find the control limits for constructing *np* - chart and explain whether the process is under control. (03 *marks*)

Part II

5. (a) Explain the difference between the terms of the following pairs.

(i) Sample space and Events.

(ii) Mutually exclusive events and Collectively exhaustive events. (03 *marks*)

(b) 3 students are selected at random from a class which has 10 boys and 5 girls. Find the following probabilities,

(i) exactly **one** girl is selected.

(ii) at least **one** girl is selected. (04 *marks*)

(c) The following table classifies 1000 persons by their sex and by whether or not they favour a certain development proposal.

	Male	**Female**	**Total**
Favour	250	450	700
Oppose	170	130	300
Total	420	580	1000

If a person is selected at random from 1000 persons, find the probabilities that,

(i) the selected person favours the development proposal.

(ii) the selected person favours the development proposal given that the person is male.

(iii) the selected person opposes the development proposal given that the person is female. (03 *marks*)

(d) Bean seeds from supplier A have 80% germination percentage and those from supplier B have 70% germination percentage. A seed packing company purchases 70% of bean seeds from supplier A and 30% from supplier B and mixes these seeds together.

(i) Find the probability that a seed selected at random from the mixed seeds will germinate.

(ii) Given that the selected seed germinates find the probability that the seed was purchased from supplier B. (06 *marks*)

[see page five

(e) An electronic system has three components K_1, K_2 and K_3. If K_1 fails K_2 is used and if K_2 fails K_3 is used. If K_3 fails the system will fail. The probability of failure of any one of these component is 0.2 and failure of these components are mutually independent. What is the probability that this system **does not** fail?

To increase the reliability of the system fourth component with the same probability of failure is added. What is the probability that the new system **does not** fail? (04 *marks*)

6. (a) (i) In a certain town 20% of the households is known to purchase a particular brand of soap. In a survey 100 investigators take random samples of 10 households to see whether the households purchase this soap. How many investigators are expected to report that at most three households purchase this soap in the samples?

(ii) A manufacturer claims that at most 10% of his product is defective. To test this claim 15 units are selected at random and his claim is accepted if among 15 units at most 2 are defective. Find the probability that the manufacturer's claim is accepted if the actual probability that a unit is defective is 0.2. (06 *marks*)

(b) (i) The average number of telephone calls arrived at a certain switchboard is 420 calls per hour. The board can make a maximum of 15 connections per minute. Assuming a poisson distribution find the probability of that some calls **cannot** be connected in a given minute.

(ii) The daily demand for a certain commodity in a shop follows a poisson distribution with mean 2. If shop keeper stocks at the beginning of each three day period how many item should be held at the beginning of the period so that he can be 95% certain of meeting the demand for the period. (06 *marks*)

(c) (i) The life of an electric component is normally distributed with mean of 800 hours and standard deviation of 60 hours. What is the probability that the component will fail before 680 hours.

If the standard deviation remains 60 hours what would have been the mean to ensure that **not** more than 10% of components fail before 800 hours.

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

In a large factory there are 16 breakdowns of machine per month on the average. Assuming breakdowns occur at constant rate, at random and independent of one another, find the probability that there will **not** be more than 22 breakdown in a given month. (08 *marks*)

7. (a) Describe the following methods of sampling giving **two** advantages and **two** disadvantages of each method.

(i) Stratified random sampling

(ii) Cluster sampling

(iii) Quota sampling (06 *marks*)

(b) The electric light bulbs of manufacturer **A** have mean life time of 1600 hours with standard deviation 200 hours, while those of manufacturer **B** has mean life time of 1400 hours with standard deviation 100 hours. If a random samples of 125 bulbs of each type are tested, what is the probability that the sample mean life time of type **A** will exceed the sample mean life time of type **B** by 240 hours? (06 *marks*)

[see page six

(c) (i) For a population of size N=6 the values of the variable Y are 8, 4, 2, 10, 5 and 7. Calculate the sample mean $\bar{y}$ for all possible simple random samples of size 2 from the population.
Verify that the sample mean $\bar{y}$ is an unbiased estimator for the population mean $\bar{Y}$ using the sampling distribution of $\bar{y}$.
Calculate the variance of $\bar{y}$ using only the formula.

(ii) Calculate sample mean $\bar{y}$ for all possible systematic samples taken from the population given in (i).
Verify that the sample mean $\bar{y}$ is an unbiased estimator for the population mean $\bar{Y}$ using the sampling distribution of $\bar{y}$.
Calculate the variance of the sample mean using the sampling distribution of $\bar{y}$ and find the efficiency of the systematic sampling compared with respect to simple random sampling. (08 *marks*)

8. (a) Explain the difference between each of the following pair of terms.
(i) Simple hypothesis and Composite hypothesis.
(ii) Power of the hypothesis test and Most powerful critical region.
(iii) Significance level and p-value. (03 *marks*)

(b) The number of accidents in a certain city in 300 days are given below.

No. of accidents	0	1	2	3	4	5	6	7	8	9	10
No. of days	28	32	70	60	50	30	20	5	3	1	1

(i) Fit a poisson distribution for these data.
(ii) Test the goodness of fit at 5% significance level and state your conclusion. (05 *marks*)

(c) In an epidemic 500 persons contacted the disease and 300 persons received no treatment. From those who received no treatment, 80 persons did not recover and of those who received treatment 70 persons recovered. Test the hypothesis that the treatment was not effective in curing the disease at 5% significance level. What is the p-value of the test? (05 *marks*)

(d) The following table gives the outputs of three machines observed for a random samples of 5 different hours for each machine.

Machine I	**Machine II**	**Machine III**
6	5	10
8	3	7
5	8	11
12	7	10
9	7	12
40	**30**	**50**

$\Sigma\Sigma x_{ij}^2 = 1060$

(i) Write down the analysis of variance model for analysing this data.
(ii) Construct the analysis of variance table and test the hypothesis that mean output of three machines are equal at 5% significance level.
(iii) Construct a 95% confidence interval for the difference of mean output of machine II and the mean output of machine III. (Use the mean square error (MSE) in the analysis of variance table as the estimate for common variance σ^2). (07 *marks*)

* * *