Rehearsal Test for G.C.E. Advance Level Examination

Business Statistics - I   Time: Two hours

NOTE:
- *Answer **all** questions.*
- *In each of the questions **1** to **50**, pick one of the alternatives 1) ,2), 3), 4) ,5) which you consider **is correct** or **most appropriate.***

(1) Which of the following statements is true regarding population and sample?

(1) If the population is finite, there is no difference between population and sample.

(2) A sample should be drawn out, only if the population is infinite.

(3) When the size of population is N and sample size is n, then $\frac{n}{N} < 1$.

(4) When the size of population is N and sample size is n, then $\frac{n}{N} \leq 1$

(5) If the population is finite, the population size N and sample size is 'n' number of samples without replacement that can be drawn from the population is $\frac{N}{n}$.

(2) Which of the following statements regarding classification of data on measuring scale is **false?**

(1) Write the number relevant to the ethnic group you belong to in the given cage.

    1. Sinhala   2. Tamil   3. Muslim   4. Other

    This is an example for norminal scale data.

(2) Classification of house holds as higher, medium and lower under three income levels is an example for ratio scale data.

(3) There is not always an equal difference among mental conditions represented by like very much, like, dislike and dislike at all in a multiple choice question associated with each four choices.

(4) Raw mark received by a candidate in an examination and corresponding Z-score are examples for interval scale data.

(5) Discrete data as well as continuous data can be considered as ratio scale data.

(3) Given below are three statements regarding business data presentation techniques

    A - Using diagrams is more appropriate, in order to meet the need of efficient communication data related to the economy of a country.

    B - In comparing data among categories, a greater accuracy can be assured by using tables.

    C - A rough idea regarding the shape of a distribution can not be get through a stem & leaf diagram.

    Which of these statement/s is/are true?

    (1) Only A   (3) Only C   (5) Only B and C

    (2) Only B   (4) Only A and B

(4) Number of king cocunuts sold by a businessman for 12 days has been recorded as follows,

    24   26   29   33   33   38

    42   44   49   52   55   60

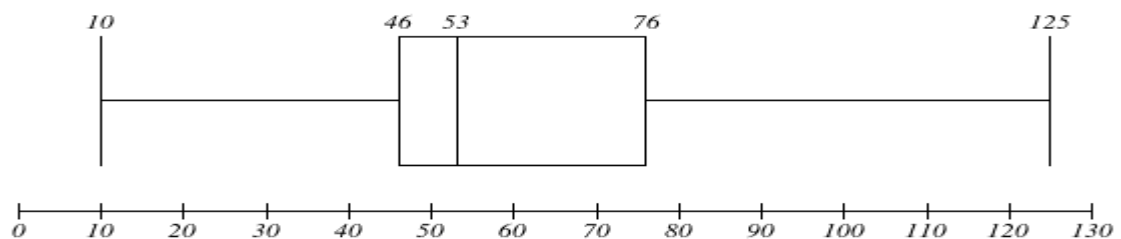   Choose the answer containing the quartile deviation of this data set accurately.

    (1) 9.75   (2) 10.00   (3) 10.62   (4) 19.50   (5) 21.25

(5) Following box & whisker diagram has been presented using a data set related to daily sales of a certain product.

    LIF = 1 $Q_1 = 46$ $Q_2 = 53$ $Q_3 = 76$ UIF = 121

    Min = 10   Max = 125

    In addition to this information the mean and standard deviation of this data set have been computed as 64.2 and 18 respectively.

    Choose the answer containing the most suitable representative measure of dispersion for this data set.

    (1) Mean and standard deviation   (4) Median and standard deviation

    (2) Mean and coefficient of variation   (5) Mean and quartile deviation

    (3) Median and quartile deviation

(6) Three statements related to organizing of data are as follows,

    A- A stem & leaf diagram is not expressed more than in an array of data.

    B- An ungrouped frequency distribution is a method of organizing data without harming the original data.

    C- The difference between lower class boundaries of two consecutive class intervals in frequency distribution with equal size class intervals is equal to the class width.

Which of the above statements is/are **false**?

(1) Only A   (3) Only C   (5) Only B and C

(2) Only B   (4) Only A and B

(7) Choose the answer containing the correct relationship that exists among the measures of central tendency in following relationships.

(1) If Mean $\leq$ Median $\leq$ Mode, it's an asymmetrical distribution.

(2) The relationship such that, Mean - Mode = 3(Mean - Median(is reasonable for any distribution.

(3) Harmonic mean $\leq$ Geometric mean $\leq$ Arithmetic mean is always true related to a same data set.

(4) Geometric mean of any data set is greater than the Harmonic mean as well as the Arithmetic mean.

(5) Any distribution such that, Mean = Median = Mode is a normal distribution.

(8) Choose the answer containing the correct order of requirements in group **A** for measure in group **B**.

Group A	Group B
a. To compute the average monthly sales growth in last 4 months. b. To obtain a measure of changing the general price level of 10 different commodities between two periods in time. c. To compute the average monthly sales of last six months with the purpose of estimating the average monthly income of a certain product. d. To compute the average time taken to supply water for four places in a factory pemisses from the main water tank through four different angled tubes.	-Simple arithmetic mean - Weighted mean -Geometric mean - Harmonic mean

(1) a b c d   (3) c b a d   (5) b a d c

(2) a c b d   (4) c b d a

(9) The mean and variance of mass of some tomato fruits were computed as 20 and 4 respectively.Later it revealed that the scale used to weight the mass always reported 5g less than the correct value.Accordingly choose the answer stating how the mean and coefficient of variation should actually be changed.

   (1) Mean increases by 5g,while coefficient of variation decreases by 20%

   (2) Mean increases by 5g, but coefficient of variation does not change.

   (3) Mean does not change,while coefficient of variation decreases by 2%

   (4) Mean increases by 5g,while coefficient of variation increases by 8%

   (5) Mean increases by 5g,while coefficient of variation decreases by 2%

(10) If the quartile deviation and the median of a symmetrical distribution are 8 and 30 respectively, choose the answer containing the correct values of first quartile and third quartile respectively.

   (1) 26 and 34   (3) 14 and 38   (5) 22 and 46

   (2) 22 and 38   (4) 14 and 46

(11) Which of the following statements regarding skewness and kurtosis of a distribution is **false?**

   (1) If Pearson's coefficient of skewness falls in the range -3 to +3, it's a moderately skewed distribution.

   (2) If Pearson's coefficient of skewness falls in the range -1 to +1, it's a moderately skewed distribution.

   (3) Most of the observations in a Lepto kurtic distribution trend to be flocked around the centre of the distribution.

   (4) Kurtosis coefficient of a Platty kurtic distribution is closer to zero.

   (5) More extreme values of distribution are covered by Kelly's coefficient of skewness than Bowley's coefficient of skewness.

(12) Total number of over time hours completed by a sample of 16 employees in a week is 432 and the sum of squres of those over time hours is 11 782.Choose the correct answer containing the average number of over time hours completed by an employee in this week and its' variance respectively.

   (1) 2.7 and 66.3475   (3) 27 and 7.87   (5) 43.2 and 688.04

   (2) 27 and 7.375   (4) 27and 499

(13) Which of the following statements regarding the movement of a pair of correlated variable is **false?**

(1) If the deviations of both variables are moving ahead on same direction, those are positively correlated variables.

(2) If corresponding values of two variables lie on a straight line or closer to a straight line, a linear relationship can be expected between those two variables.

(3) A perfect linear relationship exists between the two variables, if the variation in one variable is propotional to the corresponding variation in the other.

(4) If the regression line fitted for a pair of variables having a perfect linear correlation is moving up(inclining) from left to right, it represent a strong negative correlation.

(5) If the regression line fitted for a pair of variables having perfect linear correlation is moving down (declining) from left to right, there is a perfect negative correlation between those two variables.

(14) Manager of a particular bakery has constructed following regression lines.

Variables	Regression line
Price of bread and demand D-1 unit = 1 loaf of bread P-1 unit = Rs:10	$\hat{D}$ = 4800 – 12 P
Consumer income and the demand for bread (Y-1 unit = Rs:1000)	$\hat{D}$ = 4800 + 8Y

Comments of three students made with reference to these two models are as follows,

A- If the price of a loaf of bread is reduce by Rs 10/=, the estimate of average demand for bread will increase by 12 loaves of bread.

B- When the consumer income increases by Rs: 1 000/= the estimate of average demand will decrease by 8 loaves of bread.

C -When the value of any independent variable is zero, although the estimate of average demand for bread is 4 800, the actual demand may be more significant.

Which of these three statements is/are true?

(1) Only A   (3) Only C   (5) Only A and C

(2) Only B   (4) Only A and B

(15) Which of the following statements is true, when the regression coefficient of Y on X is - 0.4 and the regression coefficient of X on Y is -1.6?

(1) There is a stong positive correlation between these two variables.

(2) Correlation coefficient is 0.64

(3) 80% of total variation of the dependant variable is explained by independent variable.

(4) There is a strong negative correlation between these variables and 36% of the dependent variable is not explained through this regression model.

(5) Even though coefficient of determination is 0.64, but the direction of correlation may be either positive or negative.

(16) Which of the following statements is/are true regarding probability approaches?

A. Probability of a particular event takes an identical value under subjective approach

B. Probability value can not be assigned under relative frequency approach before the experiment is conducted.

C. Probability is assigned under classical approach based on experiances.

  (1) Only A   (3)   Only C   (5)   Only B and C

  (2) Only B   (4)   Only A and B

(17) Four students should be selected from a group of seven girls and five boys.The probability that the selected group containing at least three girls,

(1) $\dfrac{(7_{C_3} \times 5) + 7_{C_4}}{12_{C_4}}$

(2) $\dfrac{(7_{C_3} \times 5_{C_1})}{12_{C_4}}$

(3) $1 - \dfrac{(7_{C_3} \times 5_{C_1}) + 7_{C_4}}{12_{C_4}}$

(4) $\dfrac{(7_{C_3} + 5_{C_1}) \times 7_{C_4}}{12_{C_4}}$

(5) $\dfrac{(7_{C_3} + 5_{C_1}) + 7_{C_4}}{12_{C_4}}$

(18) **A** and **B** are two mutually exclusive events such that P(A) = 2P(B) . If $P(A' \cap B') = 0.7$, probability that the event **A** will not occur is,

  (1) 0.1   (2) 0.2   (3) 0.7   (4) 0.8   (5) 0.9

(19) If $P(A / B) = \dfrac{1}{4}$, $P(B / A) = \dfrac{1}{3}$ and $P(A \cup B) = \dfrac{3}{4}$, $P(A' \cap B)$ is ,

  (1) $\dfrac{1}{8}$ (2) $\dfrac{1}{3}$ (3) $\dfrac{3}{8}$ (4) $\dfrac{1}{2}$ (5) $\dfrac{7}{8}$

(20) If A and B are collectively exhaustive and independent events such that P(A) = 0.8 ,what is the P(B/A) ?

(1) 0   (2) 0.2   (3) 0.6   (4) 0.8   (5) 1

(21) What is the largest value of X such that $P(X > x) > 3/10$ by considering the following probability distribution.

x	0	1	2	3	4	5
P(x)	y	4y	$9y^2$	y^2	y	3y

(1) 0   (2) 1   (3) 3   (4) 4   (5) 5

(22) Which of the following statements is **true** regarding the probability distribution of a random variable?

(1) If X is a random variable E $(3X) = 3^2 \times E(X)$

(2) If X is a random variable Var $(3X+4) = 3^2 \times Var(X)$

(3) Expected value of a random variable can not be a negative value.

(4) Expected value of a probability distribution is computed based on the random value associated with the highest probability value.

(5) If X is a random variable E $(X+6) = 6 \times E(X)$

(23) Three statements regarding the function of binomial distribution

$$P(X = x) = {}^{10}C_x (0.1)^x (0.9)^{10-x} \text{ are given below.}$$

     A -   n = 10 , p = 0.1

     B -   x = 0 ,1, 2, ...

     C -   This is a negatively skewed distribution

Which of these statements is/are true?

(1) Only A   (3) Only C   (5) Only A and C

(2) Only B   (4) Only A and B

(24) If $3P(x = 3) = 2P(x = 4)$ of the distribution $X \sim Po(\lambda)$, What is the variation of that distribution,

(1) 2   (2) 3^2 (3) 6   (4) 6^2 (5) 3

(25) If $P(x > 45) = 0.8944$ of the distribution $X \sim N(50, \sigma^2)$, What is $P(x > 60)$?

(1) 0.0062   (2) 0.1056   (3) 0.3944   (4) 0.4938   (5) 0.9938

(26) Openions of three students regarding the selection of a sample from a population are as follows,

A- For an efficient study, it is important to selecting a sample at easy access to the invigilator.

B- An efficient study is facilitated with a random sample selected considering the population structure.

C- Based on the requirement and capability of the invigilator a representative sample can be selected even by using a non probability sampling technique.

Which of these three staements is/are **false**?

(1) Only A   (3) Only C   (5) Only B and C

(2) Only B   (4) Only A and B

(27) Which of the following statements is **true** regarding the Sampling?

(1) In systematic sampling the size of a sampling interval is determined as $\frac{n}{N} * 100$

(2) If a sample of 50 elements is drawn out from a population with 400 elements that population is considered as infinite.

(3) If the variation among population elements is greater, simple random sampling is more suitable.

(4) It is easy to assess the precision of an estimator in quota sampling.

(5) It is easy to assess the precision of each estimator built for each stratum separately in stratified sampling

(28) Which of the following statement is **false** regarding sampling distribution?

(1) It's a probability distribution.

(2) It's a distribution in connection with a sample statistic.

(3) It provides with a good basis for statistical estimation.

(4) It is built using all the possible samples in same size drawn out from the population.

(5) It minimizes the sampling errors.

(29) Variance of a particular characteristic of a population with 401 elements is 930.The variance of the sampling distribution of means associated with samples in size 31 is,

(1) 0.925   (2) 2.3192   (3) 2.775   (4) 27.75   (5) 30.00

(30) A student after advanced level studies who is involved in a business of preparing meals for diabetic patients using indegeneous medicinal prescriptions, has found that the mean and variance of the quantity of caleries contained in one portion of meal as 1 200 and 3 600.A sample of 36 portions (doses) of meal is tested.If the mean quantity of caleries is not exceeding 1 212, it will be recommonded to be appropriate.What is the probability that his meal will be recommended?

(1) 0.1151   (2) 0.1587   (3) 0.3413   (4) 0.3849   (5) 0.8849

(31) Self employed housewives in Galle district and Colombo district are 40% and 24 % respectively.If the sampling distribution of the difference between two sample propotions is built using two random samples of 900 housewives drawn from each district, choose the answer containing the mean and variance of that sampling distribution.

(1) 0.16 and $\sqrt{\dfrac{0.24 \times 0.76}{900}}$ (3) 0.16 and $\sqrt{\dfrac{0.8(1+0.76)}{300}}$ (5) 0.16 and $\dfrac{\sqrt{0.24 \times 1.76}}{30}$

(2) 0.16 and $\dfrac{\sqrt{0.24 \times 1.76}}{900}$ (4) 0.16 and $\sqrt{\dfrac{0.24 \times 0.76}{30}}$

(32) Several decisions made regarding point estimation are contained in column A and properties expected in a good point estimator are contained in column B

A	B
i) Since the variance of sampling distribution of sample means is gradually decreasing,then the sample mean $\bar{x}$ is used as an estimator for population mean μ	a. Unbiasedness
ii) $T_1 = \dfrac{x_1 + x_2 + x_3}{3}$ and $T_2 = \dfrac{4x_1 - x_2}{3}$ are two estimators built for population mean μ and T_1 is determined to be the best estimator.	b. efficiency
iii) Using the sample propotion 'P' to estimate the population propotion π assuring the relationship that E(P) =π	c. sufficiency d. consistency

Choose the answer contain the properties of a good point estimator, matching with the decisions mentioned in column A.

(1) d,a,b   (2) d,b,a   (3) d,b,c   (4) a,b,c   (5) a,b,d

(33) Which of the followings should not be done in making an interval estimator for a population parameter?

    (1) Understanding the population structure.

    (2) Knowing the population variance or sample variance.

    (3) Selecting an appropriate confidence level.

    (4) Identifying the sampling distribution of the relevant sample statistic.

    (5) Selecting a desired sample.

(34) Ideas expressed by three students regarding the width of a confidence interval are given below.

    A. If a lower width confidence interval should be estimated, a large size sample should be selected keeping other factors at constant.

    B. If a lower width interval should be estimated, it would be better to select 96% confidence level rather than 99% confidence level.

    C. Width of a confidence interval is independent from sample size.

Choose the answer containing the **false** statement/s from the above statements.

    (1) Only A   (3) Only C   (5) Only B and C

    (2) Only B   (4) Only A and B

(35) The confidence interval $146 \leq \mu \leq 342$ has been estimated using a random sample of size 64 drawn from the population $X \sim N(\mu, 160\,000)$. The confidence level on which this estimation has been made is,

    (1) 90%   (2) 95%   (3) 96%   (4) 98%   (5) 99%

(36) If the sample size is 400, suppose that $P \leq 0.22$ has been determined as the critical region of the test $H_0 : \pi = 0.25$ Vs $H_1 : \pi = 0.20$. What is the probability of occurring Type II error?

    (1) 0.0668   (2) 0.1587   (3) 0.4332   (4) 0.8413   (5) 0.9392

(37) Which of the followings is/are composite hypothesis given that the null hypothesis is true?

    A- $H_0 : \mu = 100$, $H_1 : \mu > 100$, $\sigma^2 = 15$

    B- $H_0 : \mu \geq 100$, $H_1 : \mu < 100$, $\sigma^2 = 15$

    C- $H_0 : \mu = 100$, $H_1 : \mu < 100$, $S^2 = 15$

    (1) Only A   (3) Only C   (5) Only B and C

    (2) Only B   (4) Only A and B

(38) What is the most appropriate distribution to test whether preference of the students for on-line learning is independent from the distance to school from the residence of each student?

(1) Chi squre distribution with (k-1) degrees of freedom.

(2) t distribution with (k-1) degrees of freedom.

(3) Chi squre distribution with)c-1)(r-1) degrees of freedom.

(4) F distribution with (k-1)/k(n-1) degrees of freedom.

(5) Standard normal distribution with mean (μ_1 - $\mu_{2)}$ and variance ($\dfrac{S1^2}{n1} + \dfrac{S2^2}{n2}$)

(39) Which of the following assumptions is not used in analysis of vaiance?

(1) Population variables distribute normally.

(2) Population means are equal.

(3) Population variances are equal.

(4) Errors in variance analysis model distribute normally.

(5) Data are collected using independent random samples.

(40) Examples given by three students for time series components are as follows,

A- Declining in sales of fuel in a filling station during April 2022 is an example for seasonal movements.

B –The prevailing demand for face masks and sanitisers in the market is an example for cyclical movement.

C-The prevailing progress in online marketing is an example for long term trend (secular trend)

Which of the above statements is/are **true**?

(1) Only A (4) Only A and C

(2) Only B (5) Only B and C

(3) Only A and B

(41) The monthly trend equation estimated consideing February 2020 as the base is

$\widehat{Y}$ = 3 977.5 + 5x.Choose the correct answer containing the annual trend line built considering the year 2020 as origin and the trend value in the year 2022.

(1) $\widehat{Y}$ = 15 910 + 80x and 16 070 (4) $\widehat{Y}$ = 48 000 + 72x and 4 944

(2) $\widehat{Y}$ = 48 000 + 60x and 48 120 (5) $\widehat{Y}$ = 47 730 + 720x and 49 170

(3) $\widehat{Y}$ = 48 000 + 720x and 49 440

(42) A statement regarding calculation of seasonal indices of a time seriesis given below.

In calculating seasonal indices on ratio to moving average method.when original time series data are devided by centered moving averages (1) _____ and cyclical movements are removed and (2) _____and irregular movements are remained by calculating the average of those remainders,irregular movments are also removed and only (3) _____ movements are expected to get remained.

Choose the answer containg the three correct terms respectively to be filled the three blanks in this statement.

(1) Trend,seasonal,seasonal   (4) Trend,irregular,seasonal

(2) Irregular,seasonal,cyclical   (5) Seasonal,irregular,seasonal

(3) Seasonal,cyclical,seasonal

(43) The quarterly trend line estimated considering II nd quarter 2020 as origin is given as $\widehat{Y} = 40 + 7.5x$ .Four quarterly indices estimatd for each quarter are 95, 115,105 and 85 respectively.The forecast value of this time series variable in last quarter 2023 is,

(1) 46.75   (2)72.25   (3)102   (4)123.25   (5)170.59

(44) Which of the following satements is **false** regarding the price indices used in Sri lanka and the influence on employee wages in the year 2020?

(1) National Consumer Price Index (NCPI) increased from 137.0 index points in January to 141.2 index points in December.

(2) Colombo Consumer Price Index (CCPI) remained unchanged in February and August due to the movement of food and non-food categories setting off each other.

(3) The real wages of public sector employees were considerably influenced by the Covid-19 pandemic situation.

(4) The real wages of the employees in formal private sector decreased by 4.2% in 2020 compared to 2019.

(5) The real wages of the employees in informal private sector were badly affected by the influence of Covid-19 pandemic.

(45) If the money salary in 2020 increased by 60% in 2022 and the price index increased by 70% in the same period, how should be the change in real salary?

(1) Decreased by 5.89%.　　　　　　　　　(4) Decreased by 6.25%.

(2) Increased by 5.89%.　　　　　　　　　(5) Decreased by 10%.

(3) Increased by 6.25%.

(46) Price indices computed for 2018 and 2021 considering 2017 as base year respectively are 120 and 210.Once the base period is shifted to year 2019, the price index for 2018 has been computed as 80.

Choose the answer containing the price index in 2021 based on 2019 respctively,

(1) 150,140　　　　(2) 170,123　　　　(3) 175,120　　　　(4) 38,552　　　　(5) 262.5,80

(47) Which of the following indices is **not used** to measure the inflation in Sri Lanka?

(1) Colombo Consumer Price Index　　　　(4) All share Price Index

(2) National Consumer Price Index　　　　(5) Producer's Price Index

(3) GDP deflator

(48) If you are an independent Quality controller employeed in a manufacturing factory, which of the followings should not be committed by you?

(1) Examining random samples.

(2) Rectifying the identified assignable causes that effect for variation.

(3) Not having personal relations with officers and other employees in the firm.

(4) Even practicing complete enumeration, if necessary.

(5) Disconnecting the calls received by personal mobile phone.

(49) Using which type of control chart can you conclude "The process is under control"

(1) Laying only a single point above the upper control limit.

(2) A regular pattern in all the points.

(3) Laying all the points between the central line and only one control limit.

(4) Laying all the points between the control limits showing a random variation.

(5) Majority of points lying closer to a single control limit.

(50) Several possible aternative decisions that can be made on the guidance of a quality control officer involved in an acceptance sampling plan are given below.(sample size – n and acceptance level – c are remain unchange)

A. When AQL = 0.01, rejecting a lot with 0.001 propotion of defectives.

B. When AQL = 0.01, rejecting a lot with 0.025 proportion of defectives.

C. When LTPD = 0.07, not rejecting a lot at 0.05 proportion of defectives.

D. When LTPD= 0.07, not rejecting a lot with the same defective proportion.

Choose the answer containg two decisions that minimizing the producers' risk and maximizing the consumers risk.

(1)   Only A and B   (3)   Only A and D   (5)   Only B and D

(2)   Only A and C   (4)   Only B and C

Rehearsal Test for G.C.E. Advance Level Examination

Business Statistics – II   Duration: 3 Hours

NOTE:
- *Answer **five** questions, choosing **at least two** questions from each part.*
- *Calculators are not allowed.*

Part I

1) (a) Explain with reasons whether usage of primary data or usage of secondary data is more appropriate for each study mentioned below.

   (i) In an investigation in to recent sectorial contribution in Gross Domestic Product in Sri Lanka.

   (ii) In a study launched to explore the issues faced by farmers in cultivation and short term and mid term actions to be taken to minimize these issues.

   (iii)In a study launched by a company to evaluate the progress of online marketing, during a period through which a certain pandemic is being rapidly dispersed. **(6 marks)**

  (b) Explain in brief the leading questions and ambiguious questions that should not be contained in a good questionnaire giving examples for each. **(2 marks)**

  (c) Explain in brief whether each of following statements is true or false.

   (i)   Recorrecting the errors identified in a questionnaire through a pretest is known as editing.

   (ii)   Considering the fact that four persons out of five who consumed a recently introduced food item prefer much for it, advertising of that food item as the one which is preferred by 80% of consumers is an example for misuse of statistical data.

   (iii) Data that can be lined up systematically, but not having a true zero are categorized as ratio scale data. **(3 marks)**

(d) Quantity of milk sold by a group of 50 dairy farmers to a milk collecting centre during a week has been recorded as follows (liters),

36	48	33	54	60	55	65	84	73	62
24	39	28	30	53	74	60	66	82	74
33	23	28	64	36	48	44	50	52	63
54	60	66	73	62	62	59	58	54	66
36	29	45	50	55	60	62	58	54	66

(i) Organize these data in a grouped frequency distribution with seven class intervals in the range 19.5 - 89.5

(ii) Construct the relative frequency histogram and comment on the nature of this distribution.

(iii) Construct the "Less than"ogive and "or more" ogive on the same coordinate plane and find median, first quartile and second quartile.   (9 marks)

(Total marks 20)

2) (a)

A- States that the algebraic property and identical property that should be satisfied by a good measure of central tendancy are found in simple arithmetic mean but not in mode and median.

B- states that these two properties are found in mode and median as well in addition to mean, because algebraic formulae are used.

How do you explain these two comments?   (5 marks)

(b)The distribution of salaries in a private firm where 200 workers have been employed is given below.

Basic Salary August 2022)Rs '000(	No: of Emplyees
36-40	08
41-45	18
46-50	30
51-55	40
56-60	32
61-65	26
66-70	20
71-75	16
76-80	10

Considering the inflationary situation prevailing in the country,it was first proposed that the basic salary of every employee of this firm should be increased by 20%.Because implementing of that proposal is not reasonable for lower salary earners,following two alternations are under consideration.

A – Adding ¼ of the average salary of an employee in this month to the basic salary of each and every worker.

B – Adding 20% of the median salary of an employee in this month to the basic salary of each and every employee.

(i) Explain with relevant calculations which of these two proposals is appropriate to control the salary expenses of the firm.

(ii) If you are an independent trade union activist, point out which proposal do you agree with?

(iii) What is the influence on the variation of the salaries in this institute through implementation of each proposal?

(iv) Construct the box and whisker's plot and comment on the shape of this salary distribution. (10 marks)

(c)

(i) Explain the importance of coefficient of variation in comparing the disparities of few data sets (2 marks)

(ii) Using the details given below, comment on the variation of those data sets.

Data set	Mean	Standard deviation
a).Salary of employees	Rs.54 000	Rs.5 400
b).Distance from home to the work place	18.5 km	3.5 km
c).Daily fuel consumption in the institute (l)	40.8 l	8 l

(3 marks)
(Total marks 20)

3) (a)

(i) Distinguish between money salary and real salary of an employee. (1 mark)

(ii) A company added a monthly amount of Rs.5 000 to the money salary of every employee.Even after that, the trade union activists are accusing that their living condition is still at a poor level. Point out how such a situation can arise. (2 marks)

(iii) Explain how rational to raise the salary level in an inflationary economy. (2 marks)

(b) Distribution of average monthly expenses of a middle class family in January 2022 for various purposes has been summarized in following table using a random sample of 1000 families.

Categary of Expenses	Amount (Rs)
Food & Beverage	8 000
Clothing	2 000
Travelling	10 000
Healthh care	3 000
Educational affairs	8 000
Social affairs	5 000
Entertainment	4 000
Total	40 000

The unit price of each expense category in January and June 2022 has been identified as follows.

Expenditure Unit	Jan 2022 (Rs(	Jun 2022 (Rs(
A lunch packet	160	280
A cake of washing soap	80	120
Minimum bus fare	14	32
Prescription of a single meeting with a doctor	400	600
A pack of stationary items including excersice books, pens and pencils	500	900
Individual cost for a wedding feast	8 000	12 000
Admission ticket in to a cinema	150	280

(i) Compute the cost of living index in June considering January 2022 as the base period using Laspyer's method.

(ii) How much should be added to the salary of an employee in June, who received Rs.40 000 as the salary in January to enjoy the same living condition. (5 marks)

(c)

    (i) What do you mean by estimation of seasonal index of a time series variable?

       How it is important for a businessman estimating the seasonal component?　　(2 marks)

    (ii) Below mentioned table contains the original observations of a time series variable in each quarter in the year 2021 and corresponding deseasonalized values.

Quarter	I	II	III	IV
Original observation	27.00	38.40	26.35	29.40
Deseasonalized data	30	32	31	28

   Compute the seasonal indices for each quarter.　　　　　　　　　　　(2 marks)

    (iii) Use a linear graph to compare the original data and deseasonalized data.　　(1 mark)

(d) Annual sales income of a company is given in following table.

Year	Sales income (Rs.Millions)
2012	14
2013	16
2014	20
2015	19
2016	22
2017	25
2018	24
2019	29
2020	32
2021	30

    (i) Estimate the annual trend line using least square method.　　　　　　(3 marks)

    (ii) Forecast the trend value in the year 2014　　　　　　　　　　　　(1 mark)

    (iii)Comment on that trend value and true sales in the year 2014.　　　　(1 mark)

                           (Total marks 20)

4) (a) Reseach and development experiences incurred by 8 independent companies that involve in manufacturing a particular product (Rs. Million) and its annual sales income of each company (Rs.Billions) are given in following table.

Research & Development Expenditure- X (Rs. Million)	2	4	4	5	8	10	12	15
Annual Sales income- Y (Rs.Billion)	16	18	20	22	23	28	30	33

$\sum X = 60$　　　$\sum Y = 190$　　　$\sum XY = 1\,613$　　　$\sum X^2 = 594$　　　$\sum Y^2 = 4\,766$

(i) Comment on the nature of correlation existing between these two variables using the scatter diagram. (1 mark)

(ii) If the regression line of annual sales income (Y) on Reasearch & Development expenses (X) has been estimated as $\hat{Y}$ = 13.925 + 1.31 X, interpret the intercept and the gradient of this regression line. (2 marks)

(iii) If the correlation coefficient between these two variables is 0.98, evaluate the goodness of fit of this regression line. (2 marks)

(iv) Suppose that Reseach & Development expenditure should be explain based on annual sales income. When the annual sales income increases by Rs. 1 Billion, find the estimate of the mean expenses for Reasearch & Development and interpret that value. (2 marks)

(v) When the annual Sales income is Rs.45 Billions, how suitable is this model to be applied to estimate the Reseach & Development expenditure. (1 mark)

(b) Two independent judges assigned ranks for 11 bowlers in a cricket team and their coefficient of agreement has been computed as -0.56. Later another judge C was called and let him assign ranks for the same bowlers and computed the rank correlation coefficient between A and C as 0.92. Comment on this event. (2 marks)

(c)

(i) Mention four advantages of maintaining a proper quality control michanism for a manufacturing firm. (2 marks)

(ii) A school leaver after A/L examination began to produce pepper powder and sell, instead of selling them as seeds. He selected 10 random sample of **5** packets in each and observed the net weight of each packet of pepper powder. Total weight of the packets in each sample and the sample range have been reported as follows.

Sample No	1	2	3	4	5	6	7	8	9	10
Total Net weight($\sum X$)	246	251	257	264	243	252	238	262	258	271
Sample Range(R)	6	4	5	8	7	5	8	6	6	10

Construct the mean control chart and examine whether the process is under control.(4 marks)

(d)

(i) An acceptance sampling plan has been constructed as n = 80 and C = 2
Find the,
- Producers' risk at AQL=0.10 and
- Consumers risk at LTPD = 0.0625 (3 marks)

(ii) State two possible suggestions to decrease the consumers risk you received in (i) above. (1 mark)

(Total marks 20)

Part II

5) (a) "Mutually exclusive events can not be independent, but independent events some times may be mutually exclusive"

Explain whether you agree with this statement. (2 marks)

(b) **A, B** and **C** are three events defined in the sample space **S** such that **A** and **B** are independent, while **B** and **C** are mutually exclusive.

If **P (A) = 2/5**, **P(C) = P (A∩B') =1/5**, find the probability of occurring following events,

(i) Occuring both the events **A** and **B**

(ii) Occuring the event **B**

(iii) Occuring both the events **B** and **C**

(iv) Occuring only **B**

(v) Not occurring **A** or **B** or **C** (6 marks)

(c) A survey was conducted using 500 people to inquire their preference for the proposed constitutional reforms in a country collected data are categorized in following table.

Preference	Rural		Urban	
	Employed	Unemployed	Employed	Unemployed
Like	120	30	140	40
Dislike	30	70	40	30

If an individual is selected at random from this sample undergone to the survey, find the probability that he/she will be,

(i) a person who likes to the reforms

(ii) a person living in an urban area

(iii) an unemployed person

(iv) a person who likes to the reforms and living in an urban area

(v) an employee living in a rural area ,given that he/she does not like to the reforms

(vi) Comment on whether the preference for the proposed reforms is independent from the living area of the people? (6 marks)

(d)

    (i) State the Bayes' theorem.

    (ii) 800 insurance policies issued by Janahitha insurance company during the first quarter in 2022 have been categorized as follows. A percentage of each type of policy that is expected to claim for compensation during the first three years also is mentioned respectively.

Type of insurance	No.of policies issued	Assumed % for claims
Marine insurance	200	5%
Motor insurance	300	15%
Life insurance	250	4%
Intelectual property insurance	50	0.8%

If an insurance policy holder selected at random claims for compensation within the first three years, find the probability that he/she has obtained a marine insurance policy.

(6 marks)

(Total marks 20)

6) (a) Suppose that an investor has invested his money in four investment sources having five equaly likely benefits in each including one disadvantagable benefit.

    (i) If the number of disadvantagable benefits he may experience in these four investment sources is the random variable X, explain the probability distribution of X highliting the relevant conditions that should be satisfied by X in connection with this situation.

    (ii) Find the probability that he may have to experience at most one disadvantageble benefit.

(5 marks)

(b)

    (i) Mean number of accidents expected to occur in a city during five days is one. If the number of accident occurred daily in this city is X, explain the probability distribution of X with the relevant assumptions.

    (ii) Under which conditions is the Poisson distribution used as an approximation to the binomial distribution?

    (iii)A gas filling company accepts that one cylinder out of 100 000 cylinders filled by them is at a risk of being blasted.When a random sample of 20 000 cylinders filled by this company is selected, the probability of at most "K" number of cylinders being blasted is 0.98. Find the value of" K" (6 marks)

(c) Number of electric ovens sold by a company during 50 days is given by following frequency distribution.

No of Ovens sold	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No. of days	1	2	5	10	14	10	5	2	1

(i) Construct the smooth frequency curve for this frequency distribution.

(ii) Suppose that this distribution falls normally.Compute the relevant parameters.

(iii) Show that the area under this frequency curve is very closer to the division of area under normal distribution curve.

(iv) In a random selected day find the probability that this trader has sold,

- less than 39 ovens
- 39 – 60 ovens   (9 marks)

(Total marks 20)

7) (a) With sufficient reasons for your choice mention the most appropriate sampling method to draw out a representative sample for each study given below.

(i) To study the nessasary actions to be taken to manage a rapidly spreading pandemic

(ii) To study the difficulties faced by railway passengers and to propose solutions to promote that service, a sample of residents who use the railway from Colombo to Badulla railway station should have been selected.

(iii)To study regarding financial fraudlences suspected to have been committed in a group of business firms with reference to the source documents of transactions.   (6 marks)

(b) Explain the diference between two terms in each pair of terms given below with examples.

(i) Parameter and statistic

(ii) Estimator and estimate   (4 marks)

(c) Name the sampling distribution with its characteristic that can be applied for the requirement of each study mentioned below.

(i) To test whether the average life times of two electric bulbs A and B produced in two factories are equal.

(ii) To estimate a 95% confidence interval regarding the families who use a particular water filter that was recently introduced in North central province.   (4 marks)

(d) Association of innovators has recently introduced two efficient cookers to the market as a solution for prevailing fuel crisis; one powered by charcoal and the other by bio gas.Sales of each kind of cooker is supposed to fall normally with mean 24.5, satandard deviation 4.2 and mean 23.2, satndard deviation 5.6 respectively.

Find the probability of occurring following events related to a study launched taking two random samples of 25 days from each kind of cookers.

(i) Daily mean sales of charcoal cooker being graeter than that of bio gas cooker.

(ii) Daily mean sales of charcoal cooker being 2.5 more than that of bio gas coooker.

(6 marks)

(Total Marks 20)

8) (a) Name the distinguish propotises of a good point estimator and explain the sample mean satisfies all those propoties for population mean when compared to the other estimators.

(4 marks)

(b)

(i)"A narrower confidence interval can be estimated with a less confidence level, while a narrower confidence interval with a greater confidence level can be estimated only by increasing the sample size"

Explain this statement using relavent evidence. (2 marks)

(ii)Suppose that the time taken by a trained bank officer to provide the service to the customers coming to the bank counter and the time taken by a trainee for the same purpose fall normally .Two random samples of 10 customers who came to each officer were selected. It has been computed the mean time taken by a trained officer to provide his services to a customer and standard deviation as 1.8 minutes and 0.1 minutes respectively ,while same measures related to a trainee are 2.95 minutes and 0.2 minutes respectively.

Construct a 98% confidence interval for the difference between the time taken by trained officer and trainee to provide services to the customers. (4 marks)

(c) An Agro Research Centre is involved in studying whether there is a significant difference in paddy harvest yielded using organic fertilizer in place of chemical fertilizer. Two random samples with 50 plots of equal extent paddy cultivated land under each situation were selected for this study and following measures have been computed using their practical observations.

Situation	Mean harvest (kg)	Standard deviation (kg)
When chemical fertilizer used	1920	80
When organic fertilizer used	1840	120

Test whether the average paddy harvest under each situation is significant at 5% level of significance. (5 marks)

(d) Cupboards manufactured by a carpenter are produced to be sold in four cities. Data collected considering monthly sales in each city as X_1, X_2, X_3 and X_4 are summerised in following table.

City	A	B	C	D
No.of weeks observed	5	4	6	3
Total No.of units sold	36	52	56	42
Sum of the squares of the No.of units sold	280	696	554	596

Test whether the average cupboard sales in these four cities are significant at 2.5% level of significance.

(5 marks)

(Total marks 20)