AL/2024/31/E-I

සියලු ම හිමිකම් ඇවිරිණි /முழுப் பதிப்புரிமையுடையது /All Rights Reserved]

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

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2024
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2024
General Certificate of Education (Adv. Level) Examination, 2024

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**.*
- *Carefully follow the Instructions given on the back of the answer sheet.*
- *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 about statistics?

(1) In statistics only business and socio-economic phenomena can be measured.
(2) Statistical conclusions are always free from uncertainty.
(3) Giving recommendations without a sufficient and representative sample is a misuse of statistics.
(4) Identifying the causal relationships between quantitative variables is the only role of statistics.
(5) Even though many statistical techniques are available only statistical hypotheses testing can be used to make decisions.

2. Consider the following statements.

A - Reliability and accuracy of results only depend on knowledge and experience of the enumerator.
B - The main purpose of data editing is to make necessary adjustments to ensure accuracy, completeness, relevancy and consistency.
C - When assigning the values for ordinal scale variables, there should be a specific scale among the orders.

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

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

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

(1) Pilot-survey is used not only to test the accuracy of a questionnaire but also to test the field activity plan.
(2) It is important to get ethical approval before collecting the data in clinical studies with human participation.
(3) The data generated within an institute can be considered as primary or secondary data mainly based on the user and the purpose of the user.
(4) Focus - group discussion method can be used to collect data from people with expertise even if they are illiterate.
(5) Under the direct observation method, data can be collected accurately using high quality technical instruments despite of involving less experienced observers.

[see page two

0303000018811O534

4. Consider the following statements about data presentation methods.

A - When the Z-chart is given for a monthly sales data of a particular company, the monthly original sales can be retrieved.

B - When the Lorenz curve was drawn using market share and sales data of each company in the insurance industry, the original market share of each company can directly be retrieved.

C - When Ogive is given, the original data can always be retrieved.

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.

5. When the number of votes taken by five leading political parties for the last five presidential elections of a particular country are given, the most appropriate chart to present these data is,

(1) pictogram (2) pie-chart (3) profile chart
(4) simple bar chart (5) multiple bar chart

6. The areas of rectangular bars are considered to get adjusted frequency which is used to construct the most accurate histogram, when the size of class intervals are,

(1) equal. (2) unequal. (3) less than five.
(4) more than five. (5) unlimited.

7. Consider the following statements about the measures of central tendency.

A - Although the median and mode can be obtained graphically, only the mean can be used for further algebraic operations meaningfully.

B - The weighted mean is the most appropriate measure to obtain the average value when the relative importance of each observed data is equal.

C - Among the key measures of central tendency, both mean and median can be considered as measures of relative locations too.

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

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

8. World market crude oil demand (in billion barrels) from 2015 to 2021 are given in the following table.

Year	2015	2016	2017	2018	2019	2020	2021
Demand	94.2	95.7	97.7	99.1	100.3	91.2	97.1

If the calculated mode and mean of crude oil demand for the period from 2015 to 2023 are 99.1 and 97.4 respectively, the demand for the years 2022 and 2023 are,

(1) 99.1, 99.1 (2) 99.1, 100.6 (3) 99.1, 102.2
(4) 100.6, 100.6 (5) 102.2, 100.6

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

(1) Only Bowley's coefficient of skewness can be used to measure the skewness for distributions with open-ended classes.

(2) Box-and-whiskers plot can be used to identify not only extreme values but also the shape of the distribution.

(3) Although the distribution is multi-modal, it has only one median.

(4) Coefficient of variation can be used to make a correct investment decision when there are several investment options with equal average returns.

(5) Kelley's coefficient of skewness is one of the options to measure the skewness of a distribution with extreme values.

[see page three

10. Consider the following stem-and-leaf diagram regarding the length of 37 bolts recorded to the nearest millimeter.

10	4										
11	9	9									
12	5	6	8								
13	2	4	6	7	8	9					
14	0	0	2	2	4	4	5	6	6	7	8
15	0	0	2	3	4	6	7				
16	1	3	4								
17	3	5									
18	8	9									

All possible extreme values of the distribution are,

(1) 104, 119, 119 (2) 104, 119, 125 (3) 104, 119, 189
(4) 104, 188, 189 (5) 175, 188, 189

11. If the D_1 and D_9 of a particular distribution which can be assumed to be approximately normal are respectively 47.25 and 104.75, the inter-quartile range of the distribution is,

(1) 13.570 (2) 15.123 (3) 27.140 (4) 30.245 (5) 39.976

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

(1) A scatter diagram can be used to quantify the relationship between variables accurately.

(2) When the two variables measured by ordinal scale, product moment correlation coefficient is most appropriate to measure its correlation.

(3) When the correlation coefficient is equal to 0.5 the exactly half of the total variation of the dependent variable is explained by the independent variable.

(4) If coefficient of determination is close to 1, always there is a strong positive correlation between the two variables.

(5) The rank correlation coefficient is not an integer, if tied ranks are available in the ranks given by one of the two persons.

13. Consider the following statements about uses of a simple regression line.

A - If a student assumes that the gas price (Y) of a particular country mainly depends on the global gas price (X), the estimated regression equation is $\hat{y}_i = 20 + 0.5x_i + u_i$.

B - When the R^2 of sales on advertising expenses is 0.24, it is more appropriate to investigate other variables which affect sales.

C - Predicting the sales price of a land based on the distance from the town using a simple regression model, the interpretation of the intercept is meaningful.

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.

14. Two popular websites, P and Q rank top 5 world tourist destinations (A, B, C, D and E). P ranks them as 1, 2, 3, 4, and 5 respectively, while Q ranks differently. The rank correlation coefficient between P and Q is -0.95. However, website Q incorrectly ranked destination C as 6 instead of 5. The corrected rank correlation coefficient is,

(1) -0.70 (2) -0.30 (3) 0.30 (4) 0.70 (5) 0.72

[see page four

000534

0203000188110534

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

A - The classical approach is applicable to experiments with any number of outcomes.
B - If outcomes of the experiment are not equally likely occurred, the probability can be calculated under the subjective approach only.
C - Mathematical approach can be applied for infinite sample spaces.

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

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

16. If a dice be weighted so that the probability of getting a number when the dice is rolled is proportional to the appearing face number, the probability that getting an even number is,

(1) $\frac{1}{21}$ (2) $\frac{1}{2}$ (3) $\frac{3}{7}$ (4) $\frac{4}{7}$ (5) $\frac{5}{7}$

17. Let A and B are two events in the sample space S. The probability that both A and B occur is $\frac{1}{12}$ and the probability that neither A nor B occur is $\frac{1}{3}$. Then,

(1) $P(A)=\frac{1}{4}$ and $P(B)=\frac{1}{3}$ (2) $P(A)=\frac{1}{3}$ and $P(B)=\frac{1}{5}$
(3) $P(A)=\frac{1}{3}$ and $P(B)=\frac{5}{12}$ (4) $P(A)=\frac{1}{3}$ and $P(B)=\frac{7}{12}$
(5) $P(A)=\frac{5}{6}$ and $P(B)=\frac{1}{3}$

18. X and Y are two events of a sample space such that $P(X)=\frac{1}{3}$, $P(Y|X)=\frac{1}{2}$ and $P(X|Y)=\frac{1}{4}$. Then $P(X'|Y')$ is,

(1) $\frac{1}{6}$ (2) $\frac{1}{4}$ (3) $\frac{1}{3}$ (4) $\frac{1}{2}$ (5) $\frac{3}{4}$

19. A certain company's shares were bought at Rs. 50,000 and the cost per share was Rs. 100. The probabilities that one share can be sold at Rs. 130, Rs. 120 and Rs. 90 are 0.3, 0.5 and 0.2 respectively. The expected profit of this investment is,

(1) Rs. 1 700 (2) Rs. 8 500 (3) Rs. 10 500 (4) Rs. 40 500 (5) Rs. 58 500

20. Long-term statistical data show that Sri Lankan women's cricket team wins 60% of cricket matches. The probability that the team will win at least 80% of next 10 matches is,

(1) 0.0106 (2) 0.0123 (3) 0.1672 (4) 0.8328 (5) 0.9877

21. If, on average 4 customers arrive at a certain shop in a 6 minutes time interval, then the probability that at most 5 customers arriving in a 9 minutes time interval is,

(1) 0.1563 (2) 0.1606 (3) 0.4457 (4) 0.5543 (5) 0.7852

22. Marks in a particular examination conducted by a certain institute are normal with mean 56 and standard deviation 16. If the management of this institute expect to pass only 16% of the total candidates, the minimum mark that the candidate should get to pass the examination is approximately,

(1) 57 (2) 60 (3) 61 (4) 72 (5) 81

[see page five

23. Consider the following statements about random sampling methods.

A - When the sampling frame is not available, cluster sampling provides the best option for obtaining a representative random sample.

B - When the variation among clusters is minimum, selecting primary clusters using the simple random sampling method is more appropriate.

C - Although cluster sampling is a cost-effective method for a wide range of investigations, it is less accurate than other random sampling methods.

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

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

24. Which of the following statements is true about the sampling?

(1) The multi-stage sampling method is more appropriate to select the sample for household, income and expenditure (HIE) surveys in a country.

(2) The population items can be studied more widely and deeply only by removing technical barriers.

(3) A representative sample can be obtained only by using well-updated and accurate sampling frame.

(4) Both sampling and non-sampling errors can be totally removed by involving the well-trained enumerators.

(5) Even though all the items in the population are included in the updated sampling frame, there is a difference between the sampling population and the target population.

25. A research team has been assigned to a survey with the purpose of estimating the average cost of living of vulnerable people (N known) due to the economic crisis in a country. If the research team has used the formula of $\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ to calculate the variance of the sampling distribution of the estimator, the most appropriate reason for using the formula is,

(1) $E(\overline{X})=\mu$. (2) sampling fraction < 0.05.
(3) infinite of population. (4) sampling with replacement.
(5) sampling without replacement.

26. Which of the following statements is true?

(1) When S^2 is an unbiased estimator for σ^2, the $\frac{S^2}{n}\left(\frac{N-n}{N-1}\right)$ is an unbiased estimator for $\frac{\sigma^2}{n}$.

(2) The variance of an estimator with replacement sampling is always greater than without replacement sampling.

(3) When $N=6$ and $n=3$, the number of samples can be obtained using sampling with replacement and without replacement are equal.

(4) Variance of any estimator can be used to estimate the variance of the population distribution.

(5) When the sampling fraction is close to 1, the variance of the sampling distribution is also close to 1.

000534

0103000018811053 4

[see page six

27. Consider the following statements about the point estimation.

A - The variance of the sampling distribution of the sample mean is always less than the corresponding population variance.

B - The value obtained by substituting values of sample units with size 35 to the equation $\bar{X} = \frac{\sum_{i=1}^{35} X_i}{35}$ is called an estimator.

C - If T is an estimator for the unknown population parameter θ and all the sample data are used to calculate it, T is known as a sufficient estimator for θ.

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.

28. If the sample proportion p is an unbiased estimator for the population proportion π, the standard deviation of the population distribution for π and the standard error of the sampling distribution for p are respectively,

(1) $\sqrt{\pi(1-\pi)}$ and $\sqrt{\frac{p(1-p)}{n}}$ (2) $\sqrt{\pi(1-\pi)}$ and $\sqrt{\frac{\pi(1-\pi)}{n}}$

(3) $\sqrt{\frac{\pi(1-\pi)}{N-1}}$ and $\sqrt{\frac{p(1-p)}{n}}$ (4) $\sqrt{\frac{\pi(1-\pi)}{n}}$ and $\sqrt{\frac{p(1-p)}{n-1}}$

(5) $\sqrt{\frac{\pi(1-\pi)}{n}}$ and $\sqrt{\frac{p(1-p)}{n-1}}$

29. If a 95% confidence interval is constructed for the population proportion π, the correct statement about the confidence interval is,

(1) $\left(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le \pi \le p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \approx 0.95\%$

(2) $\left(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le \pi \le p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \approx 5\%$

(3) $\left(p - Z_{\alpha/2}\sqrt{p(1-p)} \le \pi \le p + Z_{\alpha/2}\sqrt{p(1-p)}\right) \approx 5\%$

(4) $\left(p - Z_{\alpha/2}\sqrt{p(1-p)} \le \pi \le p + Z_{\alpha/2}\sqrt{p(1-p)}\right) \approx 95\%$

(5) $\left(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le \pi \le p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \approx 95\%$

30. Consider the following statements.

A - If two samples were drawn from two populations with sample size $n_1 = 55$ and $n_2 = 105$, the sampling distribution of $p_1 - p_2$ is normal.

B - $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$ and $U = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$. If the Z and U are independent, then $t = \frac{Z}{\sqrt{\frac{U}{(n-1)}}} = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{(n-1)}$

C - Testing the goodness of fit by using a chi-square test, the expected frequency of each interval should be at most 5.

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

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

[see page seven

31. In 10 m air rifle mixed team international event, the results of the target and shooting points are normally distributed and were reported as, $n=6$, $\overline{x}=65.7$mm and $m=12$, $\overline{y}=68.2$mm. When $s_p=3.7$mm, the 98% confidence interval for the $\mu_x-\mu_y$ is approximately,

(1) (-12.05, 7.05) (2) (-11.12, 6.12) (3) (-7.27, 2.27)
(4) (-6.81, 1.81) (5) (-4.89, -0.11)

32. The results of an environmental sustainability survey revealed that 64 out of 80 (P_a) and 36 out of 60 (P_i) answered major causes for carbon emission as 'automobile' and 'industries' respectively. When a 95% confidence interval for ($P_a - P_i$) is to construct, the standard error is,

(1) $\sqrt{0.0003}$ (2) $\sqrt{0.002}$ (3) $\sqrt{0.003}$ (4) $\sqrt{0.005}$ (5) $\sqrt{0.006}$

33. The length of feathers of peacocks in birds sanctuary were measured to identify their healthiness and fitness level. Results were reported as $\overline{x}=16.82$ cm and $\sigma^2=12.5$ cm^2. The sample size such that $\overline{x}\pm0.5$ is an approximate 95% confidence interval for μ is,

(1) 96 (2) 136 (3) 192 (4) 600 (5) 613

34. Consider the following statements about a level of significance in a hypotheses test.

A - The level of significance and p-value are directly related.
B - The level of significance is a predetermined level to determine whether the observed data significantly deviates from the null hypothesis.
C - Reject null hypothesis when the null hypothesis is true, is denoted by α.

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

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

35. A marketing manager is considering whether to introduce a new product into the market. The product will only be introduced if the consumer acceptance rate exceeds 20%. A sample survey of 144 consumers enables the manager to introduce the new product into the market at 5% level of significance. If the sample consumer acceptance rate is 24%, then p-value of the test is,

(1) 0.1151 (2) 0.1357 (3) 0.2302 (4) 0.3643 (5) 0.3849

36. The data related to two packing machines A and B in a company are given in the table below.

Machine	A	B
n	225	200
$\overline{X}$	57	55
s^2	22.5	12

The quality control department of company wants to determine whether the mean number of units packed per hour by the two machines A and B are the same. The test statistic is approximately,

(1) 1.16 (2) 1.30 (3) 1.61 (4) 2.51 (5) 5.00

37. Which of the following statements is true about statistical hypotheses tests?

(1) Critical region is a set of values for test statistic where the null hypothesis is not rejected.
(2) The confidence coefficient is the probability of rejecting the null hypothesis when it is true.
(3) The test statistic measures how far the sample data deviates from the null hypothesis.
(4) A result is statistically significant for a left tail test, if the test statistic value is greater than the critical value.
(5) The p-value is the probability of accepting the null hypothesis when it is true.

[see page eight

38. Consider the following statements about contingency coefficient.

A - The contingency coefficient is a measure used to quantify size of the relationship between two categorical variables.

B - Even though correlation coefficient ranges from -1 to +1, the range of contingency coefficient is between 0 to 1.

C - The contingency coefficient is calculated using the formula $C=\sqrt{\frac{\chi^2}{\chi^2-n}}$.

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.

39. In an ANOVA table, the mean square error is equal to 10. If random samples of six were selected from each of four populations where the total sum of square was 250, the variance between the groups is,

(1) 1.67 (2) 2.50 (3) 10.86 (4) 16.67 (5) 50.00

40. Consider the following statements about the time series analysis.

A - When a set of mortality rates data were collected in equal and successive time intervals, time series models can be used to forecast mortality rates.

B - When the monthly kWh output in a five year period of a solar plant is represented in a time series plot, it can not be identified irregular movements of monthly output.

C - If the monthly tourist arrivals were recorded as zero in a particular month of the last five years, it can be identified as an impact of the cyclical movement.

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.

41. Which of the following statements is true?

(1) Deflating monthly re-exports income by the corresponding seasonal indices, the effects of calendar changes can be removed.
(2) The ticket income variations of a particular cricket stadium during the Asia T-20 tournament is an example for the effects of population changes.
(3) Time series plot is used to determine the most accurate model in modelling the time series data.
(4) The trend line obtained by both free-hand and semi-average methods are subjective when there are odd number of data points in the time series.
(5) The trend line obtained by the moving average method is mathematically more precise than the trend line obtained by the least squares method.

42. Which of the following statements is true about the seasonal indices estimation and forecasting?

(1) The moving average method is more simple and convenient than the average percentage method.
(2) When the order of moving averages is increased the lost of end values of the time series is increased proportionately to the order.
(3) For an additive time series model, summation of all seasonal indices is equal to zero.
(4) The average percentage method is based entirely on theoretical basement of decomposing components of a time series.
(5) If the time series can be assumed to have no cyclical and irregular movements, forecasting can be done using only the trend and seasonal components.

[see page nine

43. $\hat{Y}=9600+0.576t$ is the annual trend equation estimating the demand for skin fairness cream (in thousand units) of a particular brand (origin - 2022). If the seasonal index for May 2025 is 140 and the components are assumed to be independent, the equation can be used to estimate the demand of May 2025 is,

(1) $\hat{Y}=(66.805+0.004t)*1.4$ (2) $\hat{Y}=(66.805+0.004t)*140$
(3) $\hat{Y}=(800.138+0.004t)+1.4$ (4) $\hat{Y}=(800.138+0.048t)+1.4$
(5) $\hat{Y}=(800.138+0.048t)*140$

44. Consider the following statements about statistical quality control.

A - If the manufacturing process is maintained under the control measures, it can be ensured that the consumer receives a product with 100% quality.
B - Statistical process control continuously monitors the production process to identify variations affecting the quality.
C - When increasing the sample size, estimators are more precise and it causes to reduce random variations.

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

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

45. Which of the following statements is true about statistical control charts?

(1) The *c*-chart can be used to monitor the number of defects per unit in the process of inspecting production units.
(2) The range chart inspects how the variations in observations of the relevant variable are deviated from the population mean.
(3) The grand mean is considered as the central line of the range chart.
(4) The proportion of defectives in each sample is variated about the mean number of defectives in a sample is depicted in the *np*-chart.
(5) The *u*-chart is a control chart that monitors the proportion of defects per unit when the size of each unit varies.

46. In a mobile phone manufacturing plant, a *u*-chart is used to monitor defects in the main circuit board of each phone. The average number of defects per board ($\bar{u}$) is 16, with control limits set at ± 3 standard deviations. If the distance between the mean and the lower control limit (LCL) is 1, the sample size of the circuit boards to be tested is,

(1) 3 (2) 24 (3) 36 (4) 48 (5) 144

47. When the defective proportion is 0.05, the probability of accepting a good lot is 0.9138. At a 0.15 defective proportion, the probability of rejecting a bad lot is 0.4560. The producer's and consumer's risks are respectively,

(1) 0.0862, 0.4560 (2) 0.0862, 0.5440 (3) 0.5440, 0.0862
(4) 0.9138, 0.0862 (5) 0.9138, 0.5440

48. Consider the following statements about index numbers.

A - In determining the educational expenses of Advanced Level students in a particular area, weighted aggregated indices are more practical and effective than relative indices.
B - During an economic crisis with import restrictions, it is more appropriate to determine the bucket of goods to construct the price indices by experts in board of investment.
C - It is essential to consider a base year for calculating propaganda expenses of a candidate, when candidates from various parties are represented in a presidential election.

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.

[see page ten

49. If the new base year is considered as 2019 for the price indices data given in table below, the price indices approximately for years 2014 and 2023 are respectively,

2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
100	110.8	111.5	113.0	117.4	121.6	125.3	129.2	137.5	155.8	175.2

(1) (71.52, 113.07) (2) (88.43, 139.82) (3) (110.80, 175.20)
(4) (113.07, 71.52) (5) (158.12, 113.07)

50. Which of the following statements is true about the limitations of the index numbers?

(1) Index numbers are only an approximate calculation, even though an appropriate bucket of items is selected.
(2) Non-sampling errors do not affect for computing the index numbers.
(3) The representation of the population by a sample is only depend on the selected bucket of items.
(4) Accuracy and reliability of the indices are always in a higher level when only the base year is free from unusual events.
(5) Neither the process of calculating appropriate indices nor the process of selecting relevant bucket of items is complicated.

* * *

AL/2024/31/E-II

සියලු ම හිමිකම් ඇවිරිණි /முழுப் பதிப்புரிமையுடையது/All Rights Reserved]

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

අධ්‍යයන පොදු සහතික පත්‍ර (උසස් පෙළ) විභාගය, 2024
கல்விப் பொதுத் தராதரப் பத்திர (உயர் தர)ப் பரீட்சை, 2024
General Certificate of Education (Adv. Level) Examination, 2024

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) Discuss **two** relative advantages of selecting the most appropriate method from 'Personal Interviews', 'Telephone Interviews', and 'Focus-group Interviews' as primary data collection methods for each of the following situations.

(i) The executive officer of a television channel wants to estimate the proportion of viewers who are watching the television channel at a particular hour in a certain country. (02 *marks*)

(ii) The Public Health Inspector (PHI) of a particular area wants to know public opinions regarding problems arisen due to street dogs in the area. (02 *marks*)

(iii) The chairman of the particular financial institute is planning to assess the potential of opening a new financial branch at special economic center. (02 *marks*)

(b) (i) Explain what is meant by 'secondary data' with an example. (02 *marks*)

(ii) State **four** disadvantages of using secondary data for a study. (02 *marks*)

(c) (i) Explain a situation where a questionnaire is needed for collecting primary data. (02 *marks*)

(ii) State **four** properties of a good questionnaire. (02 *marks*)

(d) (i) Discuss the importance of profile chart, compared to Lorenz curve in business field. (02 *marks*)

(ii) A supermarket network operates with at least one supermarket in most of the cities in a certain country. For the eight categories of goods with the top sales in weekends, average sales of the supermarket network and sales of a middle level supermarket in a city approximately are given in the table below. (Sales in hundred thousand rupees)

Category of goods	Average sales of the supermarket network	Sales of a middle level supermarket
A - Fish	78	85
B - Meat	187	235
C - Beverages	33	45
D - Rice	55	50
E - Fruits	92	126
F - Bakery Items	48	66
G - Cooked Foods	61	84
H - Fresh Juice	26	24

Draw the profile chart for the data and state **four** important observations to keep customers with this network. (04 *marks*)

[see page two

0202000018811O534

2. (a) State **three** properties of a good measure of central tendency and compare those properties with mean, median and mode. (04 *marks*)

(b) (i) Explain a situation where the weighted mean is applicable. (02 *marks*)

(ii) The number of shares sold and their market price (in rupees) in a certain month of a company are shown in the following table.

Selling price of a share	15	44	88	50	65	35
Number of shares sold	140	120	50	90	70	130

Calculate the average price of a share of this company, and comment on it. (02 *marks*)

(c) During a specific period when import restrictions were in place in a particular country, the number of imported vehicles and their corresponding tax percentages are shown in the table below.

Tax percentage (%)	**No. of vehicles (f)**
120 - 150	30
150 - 180	44
180 - 210	52
210 - 240	34
240 - 270	20
270 - 300	12
300 - 330	08
	200

(i) Calculate the mean, median, mode, variance and Karl Pearson's coefficient of skewness for the tax percentage distribution of imported vehicles. (07 *marks*)

(ii) Comment on the shape of the tax percentage distribution of imported vehicles. (01 *mark*)

(iii) Calculate the coefficient of variation of tax percentages of imported vehicles and compare with the coefficient of variation which was 12.5 in a general situation. (01 *mark*)

(d) Considering the above (c) situation,

(i) If the 8% of the highest taxed vehicles are in luxury vehicles category, what is the minimum tax percentage imposed for this category? (02 *marks*)

(ii) If the minimum CIF (Taxed on CIF) value of luxury vehicles category is 12 million rupees, what is the minimum tax revenue that can be earned from this vehicles category? (01 *mark*)

3. (a) (i) State **two** practical uses of time series analysis in business field. (01 *mark*)

(ii) The following table shows the number of AI (Artificial Intelligence) users in the education field in a particular country for the period from 2011 to 2023.

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
No. of Users	340	364	450	470	600	650	690	715	872	898	920	1023	1050

Draw the time series plot and forecast the number of AI users for the year 2030 by using the Semi-average method. (04 *marks*)

(b) The estimated least squares trend equation for quarterly chlorine usage (in metric tons) for water refinery in a particular country from 2017 to 2023 is given as $\hat{Y} = 200 + 0.4t$ (Consider Q_1 of year 2021 as the origin). Chlorine usage during the quarters of the rainy seasons is observed as higher compared to other seasons. The quarterly actual chlorine usage in 2023 and the estimated indices of quarterly chlorine usage are mentioned in table below.

[see page three

Quarter	Q_1	Q_2	Q_3	Q_4
Actual chlorine usage in 2023	231	170	180	230
Index	110	85	90	115

(i) Using a multiplicative model with no cyclical movements, estimate the irregular movements for the corresponding quarters in 2023. (04 *marks*)

(ii) Comment on the results of the irregular movements. (01 *mark*)

(c) (i) State the meaning of factor reversal test and show that the Fisher price index satisfies the factor reversal test. (02 *marks*)

(ii) The price relatives and weights of commodities A, B, C and D for a particular year are given in the following table.

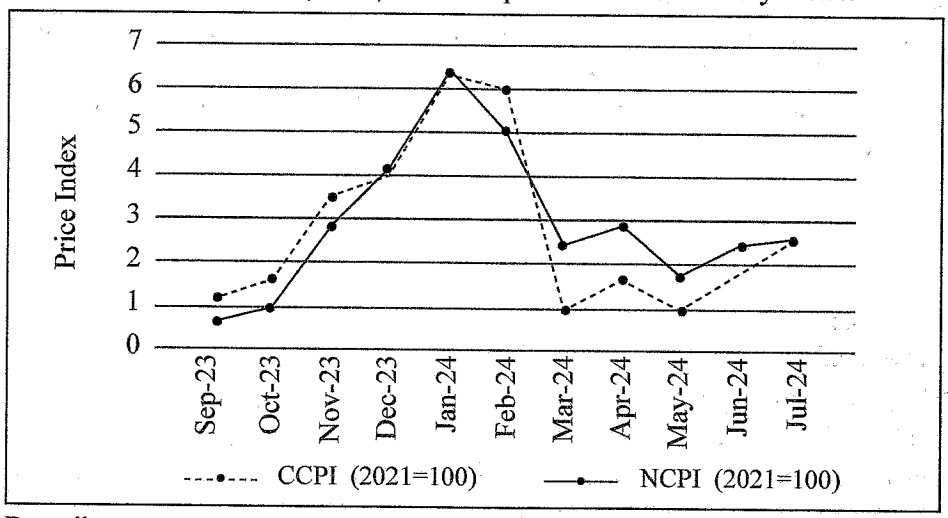
Commodity	A	B	C	D
Price Relative	113	108	115	107
Weight	W_1	$W_1 + 4$	W_2	$2W_2$

If the sum of the weights is 40 and the index of the commodities for the year is 110, find the values of W_1 and W_2. (03 *marks*)

(d) (i) What is meant by a 'Consumer Price Index'? (01 *mark*)

(ii) Explain **two** situations of using consumer price indices. (02 *marks*)

(iii) Following graph depicts Colombo Consumers Price Index (CCPI) and National Consumer Price Index (NCPI) from September 2023 to July 2024.



Describe **two** reasons for reporting the higher values of NCPI than the values of CCPI from March 2024 to June 2024 comparing their basis of constructing. (02 *marks*)

4. (a) What is meant by a 'deterministic relationship' in regression analysis? (01 *mark*)

(b) A survey was conducted to measure the relationship between weekly jogging duration in hours (X) and BMI (Y) among 8 males aged 40-50 years engaging in regular jogging. The relevant data is given in the table below.

Person	Weekly Jogging Duration (X)	BMI (Y)
1	2 hours 30 minutes	36
2	3 hours	30
3	3 hours 30 minutes	24
4	4 hours 30 minutes	26
5	5 hours	22
6	5 hours 30 minutes	20
7	6 hours	18
8	7 hours	20

[see page four

000534
0102000188110534

$$\sum X^2 = 188,\ \sum Y^2 = 5056,\ \sum XY = 849$$

(i) Draw the scatter plot and comment on the relationship between weekly jogging duration and BMI. (02 *marks*)

(ii) Estimate the least squares regression line of BMI on weekly jogging duration and interpret the coefficients. (04 *marks*)

(iii) If the coefficient of determination is 0.77, calculate the correlation coefficient and comment on both coefficients. (02 *marks*)

(iv) Forecast BMI value of a person with 1 hour and 15 minutes weekly jogging duration and comment on it. (01 *mark*)

(c) (i) What is the purpose of using a *p*-chart in statistical quality control? (01 *mark*)

(ii) A company imports 200 hair colour tubes each week. The numbers of defective tubes recorded in a recent 10 weeks period is given in the following table.

Week	1	2	3	4	5	6	7	8	9	10
No. of Defective tubes	8	11	14	10	15	13	14	12	15	20

Draw the *np*-chart and determine whether the process is statistically controlled or not with reasons. (03 *marks*)

(d) (i) State **two** uses of an accepting sampling plan in statistical quality control. (01 *mark*)

(ii) A government is planning to import 2 000 new meters for taxi service. A sample of 50 taxi meters from each lot is inspected and the lot is rejected if at least two taxi meters are faulty. Draw an OC-curve to show probabilities of accepting a lot based on faulty taxi meter proportions of 1%, 3% and 8% and explain how it can be used to assess the sampling plan. (04 *marks*)

(iii) Explain the impact to the OC-curve when only the sample size is increased with respect to the above sampling plan. (01 *mark*)

Part II

5. (a) (i) Explain what 'probability' is and state **two** examples of uses of probability in business field. (03 *marks*)

(ii) A fair dice is rolling one time which takes the sample space (S) as follows.
$S = \{1, 2, 3, 4, 5, 6\}$
Define **two** events with mutually exclusive and **two** events with not mutually exclusive from the above sample space. (02 *marks*)

(b) A and B are two events such that $P(A|B) = 0.4$, $P(A) = 0.8$ and $P(A \cup B) = 0.98$. Find,

(i) $P(B|A)$ (01 *mark*)

(ii) $P(B)$ (01 *mark*)

(iii) whether the two events A and B are independent? (01 *mark*)

(c) There are three tests to select a person for a particular job. A person is successful in getting the job if the person passes the test 1 and test 2 or test 1 and test 3. The probability of a person passing test 1, test 2 and test 3 are respectively p_1, p_2 and $\frac{1}{2}$ (Assume that every person attend for three tests). If the probability of the event that a person getting the job is $\frac{1}{2}$,

(i) Show that $p_1(1+p_2) = 1$ (03 *marks*)

(ii) If $p_2 = \frac{1}{3}$, find the probability of the person passing the test 1. (01 *mark*)

(iii) What is the probability of getting the job for a person who passed all the three test 1, test 2 and test 3? (01 *mark*)

[see page five

(d) (i) State the 'Bayes' Theorem'. (01 *mark*)

(ii) A certain company produces nails from three machines M_1, M_2 and M_3. Machine M_1 produces 64% of the total production. M_2 produces nails three times M_3. If it is known that respectively 2%, 2% and 1% of the nails produced by M_1, M_2 and M_3 are defectives,

(1) find the percentages of nails produce by machines M_2 and M_3 separately. (03 *marks*)

(2) if it is given that a randomly selected nail is defective, then find the probability that it was produced by the machine M_1. (03 *marks*)

6. (a) If X is a discrete random variable that has a probability function $f(x)$ as follows,

$$f(x)=P(X=x)=\frac{1}{3kx};\ x=1,\ 2,\ 3,\ 4$$

(i) What are the **two** conditions that should be satisfied by probability distribution of X? (01 *mark*)

(ii) Find the value of 'k'. (02 *marks*)

(iii) Construct the probability distribution of X? (02 *marks*)

(b) State the following statements are 'True' or 'False' by giving the reasons.

(i) The number of defectives out of 120 total items follows a binomial distribution with mean 5 and variance 6. (01 *mark*)

(ii) Mean of poisson distribution is equal to the square of standard deviation of the distribution. (01 *mark*)

(iii) If X follows a normal distribution with mean 63 and variance 15, then $P(X=50)=P(X<51)-P(X<49)$. (01 *mark*)

(c) A vehicle insurance salesman expects that there is a good chance of finding customers in the age group 27 years to 30 years, as people often buy their first vehicle at this age. Twenty percent of all people in certain two cities in this age group are looking to purchase a vehicle. If the salesmen inquires 15 people in this age group about vehicle insurance covers, find the probability that he will get,

(i) at least 4 insurance covers (02 *marks*)

(ii) no more than 5 insurance covers (02 *marks*)

(d) (i) State the conditions under which the poisson distribution can be applied as approximation to Binomial distribution. (01 *mark*)

(ii) Three children out of 10 000 who were vaccinated against a particular disease have been allergic. If 5 000 children are to be vaccinated in a large city, find the probability that at most 4 children will get allergic. (03 *marks*)

(iii) State **two** importance of normal distribution in statistics. (01 *mark*)

(iv) It is known that 20% of the employees working in a large production company are machine operators. If a random sample of 225 is drawn from these employees, what is the probability that it consists of less than 53 machine operators. (03 *marks*)

7. (a) Describe the following sampling methods, stating **two** advantages and **two** disadvantages for each.

(i) Simple random sampling (03 *marks*)

(ii) Systematic sampling (03 *marks*)

(b) (i) Explain **two** basic reasons considered for each of stratification and clustering in sampling. (02 *marks*)

(ii) Describe the convenience sampling method stating **two** advantages and **two** disadvantages. (02 *marks*)

(c) For a population with size $N=5$, the values of the population characteristic, Y are given as 12, 5, 10, 8 and 3. The population mean $\bar{Y}$ and median M_d are 7.6 and 8 respectively.

(i) Write down all possible simple random samples of size 3 without replacement from the population. (02 *marks*)

[see page six

(ii) Calculate sample mean $\bar{y}$ (nearest to one decimal) and sample median m_d for each sample. (02 *marks*)

(iii) Draw plots of distributions of $\bar{y}$ and m_d, and comment on them. (02 *marks*)

(iv) Calculate and comment on expected value of $\bar{y}$ and m_d comparing with the population mean $\bar{Y}$ and median M_d. (02 *marks*)

(d) Before releasing to the market, the arsenic content (in milligrams) was examined for a random sample of 50 fish tins packed by a new machine. The mean and standard error of this sample were 110 mg and 7.1mg respectively. Construct a 98% confidence interval for the mean arsenic content (μ) of the fish tins packed by the machine. (02 *marks*)

8. (a) (i) Describe how does 'critical value approach to hypotheses tests', differ from '*p*-value approach to hypotheses tests'. (02 *marks*)

(ii) State the difference between 'independent samples t-test' and 'ANOVA' with examples. (02 *marks*)

(b) (i) The owner of a publishing company believes that the daily revenue of the company has increased compared to the last year's daily average revenue of Rs.5 000. A random sample of 256 days revealed a mean daily revenue of Rs.5 125 with a standard deviation of Rs.800. Test the owner's belief at a 1% level of significance. (02 *marks*)

(ii) The quality control department of a soft drink bottling plant wants to ensure that the amount of soft drink automatically dispensed in its 300ml bottles should not vary by pre-determined level per bottle. A random sample of 12 bottles collected by an inspector shows the following amounts of contents in ml.

301 304 297 288 299 302 307 301 291 298 289 296

Find whether this data show any automatic dispensing problem at a 5% level of significance (Consider $s = 5.93$). (03 *marks*)

(c) A particular manufacturing company has experienced an increase in defective units in manufacturing process. The production supervisor believes that the night shift has more defective units than the day shift. The results from random samples taken during the day and night shifts are as follows.

Production shift	**Day (D)**	**Night (N)**
Number of units checked	1800	1200
Number of defective units recorded	63	54

If the production supervisor decides to provide a training program for the workers on the night shift to improve their skills, check whether the program should be implemented at the 5% level of significance using *p*-value approach (Consider $\text{Var}(P_N - P_D) = 0.000049$). (03 *marks*)

(d) The management of a large hotel chain wants to find out whether the consumption of beverage (tea, coffee and milk) during breakfast is same for men as same as for women. A survey was conducted on 500 hotel guests about their gender and the choice of beverage with breakfast. The results are tabulated as follows.

Gender	**Choice of beverages**		
	Tea	**Coffee**	**Milk**
Men	120	90	40
Women	30	160	60

(i) Formulate relevant hypotheses for the above situation. (02 *marks*)

(ii) Calculate the appropriate test statistic. (04 *marks*)

(iii) Draw your conclusion at 5% level of significance. (02 *marks*)

* * *