

Question Paper

Quantitative Methods – II (132) : January 2004

Section A : Basic Concepts (40 Marks)

- • This section consists of questions with serial number 1 - 40.
- • Answer all questions.
- • Each question carries one mark.

1. Which of the following is **not true** for random variables?
 - a. A discrete random variable can assume only a limited number of values
 - b. A continuous random variable can assume any value within a given range
 - c. A random variable follows a predictable sequence
 - d. A random variable is a variable which assumes different numerical values as a result of random experiments
 - e. The value of the random variable changes from occurrence to occurrence in no predictable sequence.
2. Which of the following is **not true** for probability distributions?
 - a. The past frequency distribution may be used as a guide for predicting the future values of the random variable
 - b. The probability distribution gives the chances of occurrence of the various values of the random variable
 - c. A subjective probability distribution of the likely future values can be formed on the basis of expert opinion
 - d. We can never assign probabilities to all likely values of the random variable on the basis of the past data
 - e. The sum of the probability of occurrence of all possible values must be equal to one.
3. Which of the following is **false** for expected value of a random variable?
 - a. The concept of expected value of the random variable can be used for making decisions under conditions of uncertainty
 - b. The expected value is the mean of the probability distribution of the random variable
 - c. The mean is computed as the weighted average of the values that the random variable can assume
 - d. The probability assigned to each possible value of the random variable is used as weight
 - e. The expected value of the random variable predicts the value that the random variable will assume in future.
4. If the covariance of two random variables X and Y is equal to zero then it can be said that
 - a. The random variables are dependent on each other
 - b. The random variables have a positive correlation
 - c. The random variables have a negative correlation
 - d. The random variables are independent of each other
 - e. The variances of the two random variables are always zero.
5. Consider a batch of N cricket balls. Each cricket ball may be defective or non-defective. The experiment involves selecting a cricket ball and checking whether it is defective or non-defective. If this experiment is repeated without replacing the balls then the probability distribution function obtained from such an experiment can be appropriately described by
 - a. The binomial distribution
 - b. The hypergeometric distribution
 - c. The normal distribution
 - d. The lognormal distribution
 - e. The t distribution.
6. In many situations managers resort to sampling to draw some conclusions about a population. Which of

the following is **not** an advantage of sampling over a census?

- a. The population may be too large to be studied in full, hence sampling is the only feasible means
 - b. A study of sample is usually cheaper than a census
 - c. Sampling usually provides information quicker than a census
 - d. In destructive testing sampling is the only available course
 - e. The conclusions obtained from sampling are more accurate than census.
7. Rahul automobiles, wants to study the consumer preference in terms of the features that vehicle owners of the city want to have in their vehicle. It divided the city in fifty sectors, each representing the characteristics of the whole city. Now they randomly selected five sectors and interviewed each household of these sectors for their preferences. They plan to forward the result of this survey to the vehicle producers to make them aware about the consumer preferences. This method of sampling is termed as
- a. Cluster sampling
 - b. Stratified sampling
 - c. Systematic sampling
 - d. Judgmental sampling
 - e. Simple random sampling.
8. The standard error of sample mean is
- a. The mean of the sampling distribution of the mean
 - b. The standard deviation of the sampling distribution of the mean
 - c. The variance of the sampling distribution of the mean
 - d. Equal to the standard deviation of the sampling distribution of the proportion
 - e. The mean of the sampling distribution of the standard deviation.
9. Which of the following is **not true** for estimation?
- a. The probability that we associate with an interval estimate is called the confidence interval
 - b. If the population standard deviation is high, we can go for large confidence interval
 - c. The number of degrees of freedom used in a t-distribution is equal to the sample size
 - d. The t distribution need not be used in estimating if you know the standard deviation of the population
 - e. The confidence interval is the range of the estimate we are making.
10. If in a test of statistical hypothesis, α be the level of significance and β be the probability of committing Type II error then which of the following is **not true** for the hypothesis test?
- a. The probability that the null hypothesis is true but rejected is α
 - b. The probability that the null hypothesis is false but accepted is β
 - c. $(1-\beta)$ is known as the power of test
 - d. A high $(1-\beta)$ value implies that the test is poor
 - e. There is a trade-off between the Type I and Type II error.
11. Which of the following is **false** with regard to simple regression and correlation analysis?
- a. If the coefficient of correlation between two variables is close to one, then there is a strong correlation between the two variables
 - b. If the slope of a regression equation is positive then the coefficient of correlation between the variables involved is positive
 - c. A high correlation always indicates a cause-effect relationship between the variables involved
 - d. If the slope of a regression equation is negative then the coefficient of correlation between the variables involved is negative
 - e. If the coefficient of correlation between two variables is equal to zero, then there is no correlation between the two variables.
12. According to the 'method of least squares' criterion, the regression line should be drawn on the scatter diagram in such a way that
- a. The sum of the squared values of the vertical distances from each plotted point to the line is maximum

- b. The sum of the squared values of the vertical distances from each plotted point to the line is minimum
 - c. The sum of the squared values of the horizontal distances from each plotted point to the line is maximum
 - d. The sum of the squared values of the horizontal distances from each plotted point to the line is minimum
 - e. The sum of the squared values of the vertical distances from each plotted point to the line is equal to zero.
13. The standard error of estimate of a simple regression line is
- a. Equal to the slope of the regression line
 - b. Equal to the square root of the coefficient of determination
 - c. Equal to the Y intercept of the regression line
 - d. The measure of the variability of the observed values around the regression line
 - e. Equal to the coefficient of correlation between two variables.
14. If r is the coefficient of correlation between two variables X and Y , then the variation in the observed Y values which is not explained by the regression line is equal to
- a. r^2
 - b. $1 - r^2$
 - c. r
 - d. $1 - r$
 - e. $1 + r^2$.
15. Which of the following is **false** for the multiple correlation coefficient between dependent variable Y and two independent variables X_1 and X_2 ?
- a. It depends on the correlation coefficient between Y and X_1
 - b. It depends on the correlation coefficient between X_1 and X_2
 - c. It depends on the correlation coefficient between Y and X_2
 - d. It will take values between 0 and 1
 - e. It will take values between -1 and $+1$.
16. In a multiple regression analysis involving two independent variables, X and Y , we find that there is no correlation between dependent variable Z and independent variable Y . We also observe that there is no correlation between two independent variables X and Y . The partial coefficient of correlation between Z and X when Y is kept constant is equal to
- a. Simple coefficient of correlation between Z and X
 - b. Simple coefficient of correlation between Z and Y
 - c. Simple coefficient of correlation between X and Y
 - d. Partial coefficient of correlation between Z and Y when X is kept constant
 - e. Partial coefficient of determination between Z and X when Y is kept constant.
17. If RSS be the regression sum of squares and ESS be the error sum of the squares in a linear regression analysis between two variables with n observations, then the ratio $\frac{RSS/1}{ESS/(n-2)}$ follows
- a. A normal distribution
 - b. A t distribution
 - c. An F distribution
 - d. A chi square distribution
 - e. A binomial distribution.
18. In a multiple regression with n observations and k independent variables, the standard error of estimate is calculated using the number of degrees of freedom as
- a. $n - 2$
 - b. $n - k$
 - c. $n - k - 1$
 - d. $n - k + 1$
 - e. $n + k - 1$.
19. In the regression analysis for a set of observations less than 30, the estimated values of the dependent variable is likely to follow
- a. A normal distribution
 - b. A t distribution
 - c. An F distribution
 - d. A chi square distribution
 - e. A lognormal distribution.
20. In multiple-regression analysis, the regression coefficients often become less reliable as the degree of

correlation between the independent variables increases. This problem is referred by statisticians as

- a. Random variations b. Non-random variations c. Irregular variation
- d. Multicollinearity e. Cyclical variation.

21. Which of the following is **false** about index numbers?

- a. Index numbers are expressed as a percentage
- b. Index number for the base year is always 100
- c. Index number has a unit of the variable that is being compared
- d. Index number is calculated as a ratio of the current value to a base value
- e. None of the above.

22. Which of the following is **not true** about the use of index numbers?

- a. Index numbers when analyzed reveal a general trend of the phenomenon under study
- b. Index numbers help in taking policy decisions, like dearness allowance paid to employees
- c. Index numbers help in determining the purchasing power of money
- d. Index numbers can be used to deflate time series data to reflect real values
- e. The value index is the most popular index as it can distinguish the effects of price and quantity separately.

23. Which of the following price indices has a downward bias?

- a. Marshall-Edgeworth price index b. Fisher's price index
- c. Laspeyres price index d. Paasche's price index
- e. None of the above.

24. The weighted average of price relatives using base values as weights is same as the

- a. Unweighted aggregates price index b. Unweighted aggregates quantity index
- c. Laspeyres price index d. Laspeyres quantity index
- e. Paasche's price index.

25. According to the factor reversal test if we interchange the price and quantity terms in an index number then the product of the original index number and the index number obtained by reversing the factors should be equal to

- a. One b. One hundred c. The value index
- d. The Fisher's ideal index e. The chain index number.

26. Which of the following methods is used to isolate the cyclical variation component from the time series?

- a. Residual method b. Linear regression analysis
- c. Non-linear regression analysis d. Ratio to Moving average method
- e. Multiple regression analysis.

27. An ice-cream parlour has been witnessing similar fluctuations in the demand of ice creams in the same quarters each year over the last five years. The owner wants to study the movement of the demand to streamline his procurement plans. Which of the following components of the time series should be studied to get an accurate behaviour of the demand?

- a. Secular trend b. Cyclical variation c. Seasonal variation
- d. Irregular variation e. Regular variation.

28. The movement of the dependent variable above and below the secular trend line over periods longer than one year is known as

- a. Seasonal variation b. Secular variation c. Irregular variation
- d. Cyclical variation e. Regular variation.

29. If Y be the actual value and $\hat{Y}$ is the estimated value using an estimating equation for a time series,

then the expression $\left(\frac{Y - \hat{Y}}{\hat{Y}} \right) \times 100$ is termed as

- a. Percent of trend measure b. Relative cyclical residual measure
- c. Standard error d. Standard deviation e. Seasonal Index.

30. Which of the following is **not true** about the ratio to moving average method?
- a. This method is used to identify the seasonal variations in a time series
 - b. An index is constructed with a base of 100
 - c. The magnitude of seasonal variations is measured by the individual deviations from the base of 100
 - d. The index is used to eliminate the seasonal effects in the time series
 - e. We multiply the original data points with the relevant seasonal index to deseasonalize the time series.
31. Which of the following control charts is/are used to monitor the variability in a process?
- a. $\bar{X}$ charts b. R chart c. p chart
 - d. Scatter diagram e. Both (b) and (c) above.
32. Which of the following is **not true** about the statistical process control?
- a. Quality standards in manufacturing process and the service industry can be achieved through the application of statistical methods
 - b. The variability present in the manufacturing process can be minimized to the extent possible
 - c. The variation in the specifications of the product due to wearing of the machine is a systematic or non-random variation
 - d. The variation arising out of measurement errors is a random variation
 - e. If the process is out of control, then random variation will be observed.
33. The quality manager of the bottling plant of a cola major wants to ensure that the quantity of soft drinks in the bottles is as printed on the label of the bottle. He takes the samples of ten bottles every day and plots a control chart for the mean quantity of beverage in the bottle. Towards the end of the week he observes that the observations are within the control limits but there are jumps in the level around which the observations vary. Which of the following inferences is most appropriate?
- a. The variations have been reduced to a great extent
 - b. The means of two different populations are being observed
 - c. The process mean has been shifting
 - d. The process is out of control
 - e. None of the above.
34. In a control chart there are some observations that lie outside the upper and lower control limits. These points are referred to as the
- a. Outliers b. Attribute c. Inherent variation
 - d. Assignable variation e. Common variation.
35. A company manufacturing a brand of detergent powder, *Sunshine*, has found the proportion of people using their brand of detergent powder in three different geographic regions by choosing suitable samples from each region. The company wishes to test whether the regions have significantly different proportions. Assuming p_1 , p_2 and p_3 to be the true proportions for the three regions, which of the following represents the alternative hypothesis for this test?
- a. $H_0 : p_1 = p_2 = p_3$ b. $H_1 : p_1 = p_2 = p_3$ c. $H_0 : p_1 \neq p_2 \neq p_3$
 - d. $H_1 : p_1 \neq p_2 \neq p_3$ e. $H_1 : p_1, p_2$ and p_3 are not all equal.
36. Which of the following is **false** about Chi square distribution?

- a. There will be a different chi square distribution for each degrees of freedom
 - b. If the number of degrees of freedom is small, the curve will be skewed to the right
 - c. As the number of degrees of freedom increases the curve tends to become symmetrical
 - d. The area under the probability distribution curve increases as the number of degrees of freedom increase
 - e. None of the above.
37. The psychology department of an institute has developed an IQ test. It has compiled the result of the test conducted on a number of students and faculty members of the institute. After analyzing the test results they have arrived at this conclusion that the distribution of results can be said to follow a binomial distribution with the probability of scoring above average IQ level to be 60%. They want to test whether this particular result really follows a binomial distribution. Which of the following statistical techniques should be used to verify these findings?
- a. Analysis of variance
 - b. Regression analysis
 - c. Correlation analysis
 - d. Chi square test
 - e. Marginal analysis.
38. Which of the following is **not** a step used in Analysis of variance (ANOVA)?
- a. We determine an estimate of the population variance from the variance that exists among the sample means
 - b. We determine an estimate of the population variance from the variance that exists within the samples
 - c. We prepare a table of observed and estimated frequencies
 - d. We accept the null hypothesis if the two estimates of the population variance are not significantly different
 - e. We reject the null hypothesis if the **F** ratio is too high and falls in the critical region.
39. When the chi-square test is used as a test of independence, the number of degrees of freedom is determined by
- a. The sample size
 - b. The ratio of sample size and the population
 - c. Number of rows in the contingency table, only
 - d. Number of columns in the contingency table, only
 - e. Both the number of rows and the number of columns in the contingency table.
40. If $F(n, d, \alpha)$ is the right tail value in a **F** distribution, then the corresponding left tail value will be
- a. $F(n, d, 1-\alpha)$
 - b. $F(d, n, 1-\alpha)$
 - c. $1/F(n, d, 1-\alpha)$
 - d. $1/F(d, n, 1-\alpha)$
 - e. $F(d, n, \alpha)$.

END OF SECTION A

Section B : Problems (60 Marks)

- This section consists of questions with serial number 41 - 70.
- Answer all questions.
- Marks are indicated against each question.

41. The owner of a movie theatre wants to determine the average number of moviegoers in the coming month. From the past experience he assigns probabilities to the number of moviegoers per day for the coming month. [< Answer >](#)

No. of moviegoers per day	Probability of the number of moviegoers per day
500	0.15
550	0.20
600	0.30
650	0.35

Which of the following is the expected number of moviegoers per day for the coming month?

- a. 553 b. 563 c. 573 d. 583 e. 593.

(1 mark)

42. A biased coin has the probability of giving a head when tossed equal to 0.60. What is the probability of getting exactly three heads in 4 tosses? [< Answer >](#)

- a. 0.2546 b. 0.3456 c. 0.4654 d. 0.5346
e. 0.6354.

(1 mark)

43. A simple random sample of size 49 is drawn from a finite population consisting of 205 items. If the population standard deviation is 12.5, then what is the standard error of sample mean when the sample is drawn without replacement? [< Answer >](#)

- a. 1.56 b. 1.65 c. 1.78 d. 1.87 e. 2.01.

(1 mark)

44. For a given data set of 10 observations of two variables X and Y, the covariance between X and Y is -121.5 and the standard deviation of the variable X is 9.0. The sums of the observations of X and Y in the given data set is 55 and 33 respectively. If a regression line is plotted for estimating the value of Y for a given value of X, then what is the estimated value of Y for X equal to 5? [< Answer >](#)

- a. 2.15 b. 3.25 c. 4.05 d. 5.15 e. 6.25.

(1 mark)

45. The following is the matrix of correlation coefficients between three variables Y, X_1 and X_2 . [< Answer >](#)

	Y	X_1	X_2
Y	1	0.7	0.5
X_1	0.7	1	0.6
X_2	0.5	0.6	1

How much variation in Y is explained by X_1 if X_2 is kept constant?

- a. 33 % b. 58 % c. 61 % d. 72 %
e. 83 %.

(1 mark)

[< Answer >](#)

46. If the prices and the quantities consumed for a basket of commodities in two time periods are such that $\sum P_1 Q_0 = 1280$, $\sum P_0 Q_0 = 1140$, $\sum P_1 Q_1 = 1680$ and $\sum P_0 Q_1 = 1500$, then what is the Fisher's ideal price index?

- a. 105.12 b. 108.14 c. 110.26 d. 112.14 e. 116.16.

(1 mark)

[< Answer >](#)

47. The manager of a food chain outlet wants to ensure that the variability in the service time by their bearers is in control. He does a sampling of the average time taken by the bearers in serving the customers in the outlet. He finds that the mean of the sample range to be 1.5 minutes. If he has collected five samples each of ten observations to conduct the test then what is the upper control limit of the R chart? (Given that $d_2 = 3.078$ and $d_3 = 0.797$ for sample size of 10)

- a. 0.335 b. 1.005 c. 1.500 d. 2.005 e. 2.665.

(1 mark)

[< Answer >](#)

48. In a test of ANOVA, the estimate of the population variance based on the variance among the sample means is 3.362 and the estimate of the population variance based on the variance within samples is 1.124. What is the F ratio for this test?

- a. 0.334 b. 1.944 c. 2.243 d. 2.991 e. 3.779.

(1 mark)

[< Answer >](#)

49. The estimation equation for a time series is given by $\hat{Y} = 60 + 1.5x$, where x is the coded time such that $x = X - \bar{X}$. The mean year is given as 1998. What would be the relative cyclical residual for the year 2003, if the observed value for that year were 70?

- a. -3.70 b. 3.57 c. 3.70 d. 96.4 e. 103.7.

(1 mark)

[< Answer >](#)

50. A normal variable has the mean of 12 and a standard deviation of 5. What is the probability that the normal variable will take a value in the interval of 5.6 and 21.8?

- a. 80.54 % b. 85.56 % c. 86.64 % d. 87.47 %
e. 95.52 %.

(1 mark)

[< Answer >](#)

51. The conditional returns on the securities of Magnum Steel and Paragon Pharma for year 2003 with their respective probabilities are as follows:

Market Condition	Probability	Return (%)	
		Magnum Steel	Paragon Pharma
Bullish	40 %	40	35
Stable	50 %	25	25
Bearish	10 %	-15	20

What is the covariance of returns?

- a. 67.0 % squared b. 73.0 % squared c. 77.5 % squared
d. 87.5 % squared e. 93.0 % squared.

(2 marks)

[< Answer >](#)

52. The latest nationwide political poll indicates that for persons who are randomly selected, the probability

that they are pro NDA is 0.55, the probability that they are pro Congress is 0.30, and the probability that they are pro Left Front is 0.15. Assuming that these probabilities are accurate, what is the probability that out of randomly chosen group of 10 persons atleast eight are pro NDA?

- a. 0.0995 b. 0.1980 c. 0.4400 d. 0.5500
e. 0.9950.

(2 marks)

[< Answer >](#)

53. The academic council of Ohio University has decided to bring some changes to its existing grading system. The university wants to determine what proportion of the students is in favour of new grading system. The total number of students enrolled in the university is 60,000 and the university feels that atleast 36,000 of them support the changes. What should be the sample size that will enable the university to be 89.9 percent certain of estimating the true proportion in the favour of new system within plus and minus 0.02?

- a. 1,075 b. 1,411 c. 1,613 d. 1,680
e. 1,711.

(2 marks)

[< Answer >](#)

54. An investment analyst must make a decision about investing in one of the five investment portfolio packages. Each investment portfolio contains different proportions of common stock, industrial bonds and real estate. The table below gives the gain of investing in these portfolios under different states of nature S_1 , S_2 and S_3 based on the projection of changes in stock prices, yields on bonds and appreciation in real estate values.

Portfolio	Gains from investment (in Rs.)		
	State S_1	State S_2	State S_3
A	30,000	4,000	2,000
B	40,000	25,000	-15,000
C	50,000	20,000	-30,000
D	40,000	10,000	5,000
E	20,000	5,000	-5,000

If the analyst specifies his degree of optimism as 0.7, then according to Hurwicz criterion which portfolio will he invest in?

- a. Portfolio A b. Portfolio B c. Portfolio C d. Portfolio D e. Portfolio E.

(2 marks)

[< Answer >](#)

55. John Wright, an overzealous graduate student, has just completed a first draft of his 700-page dissertation. John has typed this paper himself and is interested in knowing the average number of typographical errors per page, but does not want to read the whole paper. Knowing a little bit about business statistics, John selected 50 pages at random to read and found that the average number of typographical errors per page was 4.5 and the sample standard deviation was 1.2 typographical errors per page. What will be the interval estimate for the average number of typographical errors per page in his paper at the 90 % confidence level?

- a. 4.5 ± 0.2683 b. 4.5 ± 0.2783 c. 4.5 ± 0.3206 d. 4.5 ± 0.3326
e. 4.5 ± 0.3665 .

(2 marks)

[< Answer >](#)

56. Atlas Sporting Goods has implemented a special trade promotion for its tennis balls and feels that the promotion should result in a price benefit for the consumer. Atlas knows that before the promotion began, the average retail price of the ball was Rs. 44.95, with a standard deviation of Rs.5.75. Atlas samples 25 of its retailers after the promotion was launched and finds the mean price for the ball is now Rs.42.95. Atlas wants to examine whether the price of balls has significantly reduced after the trade promotion. Which of the following conclusions can be drawn on the basis of the above information at a

significance level of 2%?

- a. The average retail price to the consumers has increased significantly
- b. The average retail price to the consumers has not decreased significantly
- c. The average retail price to the consumer has decreased significantly
- d. We cannot draw any conclusion, as the sample size is too small
- e. We cannot draw any conclusion, as information is insufficient.

(2 marks)

[< Answer >](#)

57. The city council of Goa has gathered data on the number of minor traffic accidents and number of youth soccer games that occur in town over a weekend. He feels that there is a close relationship between the number of accidents and the number of games during weekends.

No. of Soccer games	20	30	10	12	15	25	34
No. of Minor accidents	6	9	4	5	7	8	9

Assuming that there is a linear relation between these two, what would be your prediction about the number of minor accidents that will occur on a weekend during which 33 soccer games take place in Goa?

- a. 6.25
- b. 7.35
- c. 8.65
- d. 9.25
- e. 9.85.

(2 marks)

[< Answer >](#)

58. Rola Cola is studying the effect of its latest advertising campaign. People chosen at random were called and asked how many cans of Rola Cola they had bought in the past week and how many Rola Cola advertisements they had seen on television in the past week.

No. of advertisements seen	3	7	4	2	0	4	1
No. of cans of Rola Cola purchased	10	17	9	4	7	6	3

What is the coefficient of correlation for the above sample of the number of advertisement seen and the number of cans purchased in the past week?

- a. 0.6548
- b. 0.7458
- c. 0.8001
- d. 0.9134
- e. 0.9868.

(2 marks)

[< Answer >](#)

59. A national shopping survey was conducted to study the average weekly buying habits of a typical Indian family in 1990 and 1994. The data collected are as follows:

Items	1990		1994	
	Unit Price (in Rs)	Quantity	Unit price (in Rs)	Quantity
Cheese (8 grams)	11.9	2	20.9	1
Bread (1 loaf)	7.9	3	10.9	3
Eggs (1 dozen)	8.4	2	13.5	1
Milk (1 liter)	13.6	2	23.9	2

What is the Marshall-Edgeworth price index for the year 1994 using 1990 as the base period?

- a. 162.38
- b. 164.17
- c. 166.64
- d. 168.55
- e. 171.34.

(2 marks)

[< Answer >](#)

60. Raj Aryan, the auditor of a state public school system, has reviewed the inventory records to determine if the current inventory holdings of textbooks are typical. The following inventory amounts are from the previous five years:

Year	1999	2000	2001	2002	2003
Inventory (in thousand rupees)	4,620	4,910	5,490	5,730	5,990

If the time variable has been coded as $x = X - \bar{X}$, then which of the following is the linear trend equation

for estimating inventory level for a given year in thousand rupees?

- a. $\hat{Y} = 4,332 + 356 X$ b. $\hat{Y} = 5,348 + 356 x$
 c. $\hat{Y} = 53,48,000 + 356 x$ d. $\hat{Y} = -7,07,008 + 356 x$
 e. $\hat{Y} = -7,07,008 + 2001 X$.

(2 marks)

61. Mithra Motors has recently started its vehicle showroom in the city. Its chief asset is a franchise to sell automobiles of a major Indian manufacturer. Mithra's general manager is planning the staffing of the dealership's garage facilities. From information provided by the manufacturer and from other nearby dealerships, he has estimated the number of annual mechanic hours that the garage will be likely to need.

Hours	10,000	12,000	14,000	16,000
Probability	0.2	0.3	0.4	0.1

The manager plans to pay each mechanic Rs 9.00 per hour and to charge his customer Rs.16.00. Mechanics will work a 40-hour week and get an annual 2-week vacation. Assuming that the mechanics do not get paid holidays, how many mechanics Mithra Motors should hire and what is the maximum amount it should pay to get perfect information about the number of mechanics it needs? (Assume 1 year = 52 weeks)

- a. Mithra Motors should hire 4 mechanics and should pay maximum Rs.20,800 to get perfect information
 b. Mithra Motors should hire 5 mechanics and should pay maximum Rs.19,600 to get perfect information
 c. Mithra Motors should hire 6 mechanics and should pay maximum Rs.12,000 to get perfect information
 d. Mithra Motors should hire 7 mechanics and should pay maximum Rs.14,000 to get perfect information
 e. Mithra Motors should hire 8 mechanics and should pay maximum Rs.28,800 to get perfect information.

(3 marks)

62. Anand Electronics runs a chain of stores selling stereo systems and components. It has been very successful in many university towns, but it has had some failures. Analysis of its failures has led it to adapt a policy of not opening a store unless it can be reasonably certain that more than 15 percent of the students in town own stereo systems costing Rs.5,000 or more. A survey of 300 students at BHU, Varanasi has revealed that 57 of them own stereo systems costing atleast Rs. 5,000. Anand Electronics is willing to run a 5 percent risk of failure, and wants to make a decision on the basis of the sample result. Which of the following are **correct** for this test of hypothesis?

- I. A two-tailed test of hypothesis should be done.
 II. A left tail test of hypothesis should be done.
 III. A right tail test of hypothesis should be done.
 IV. The alternative hypothesis is $p \neq 0.15$.
 V. The alternative hypothesis is $p < 0.15$.
 VI. The alternative hypothesis is $p > 0.15$.
 VII. The standard error of proportion is 0.19.
 VIII. The standard error of proportion is 0.02375.
 IX. The standard error of proportion is 0.02265.
 X. The critical values for the test are -1.96 and $+1.96$.
 XI. The critical value for the test is -1.64 .
 XII. The critical value for the test is $+1.64$.
 XIII. The standardized value of sample proportion is -0.21 .
 XIV. The standardized value of sample proportion is 1.684.
 XV. The standardized value of sample proportion is 1.766.
 XVI. At a 5 percent risk of failure Anand Electronics should open a store in Varanasi.

XVII. At a 5 percent risk of failure Anand Electronics should not open a store in Varanasi.

- a. (I), (IV), (VII), (X), (XIII) and (XVII)
- b. (II), (V), (VIII), (XI), (XIV) and (XVI)
- c. (II), (V), (IX), (XI), (XV) and (XVI)
- d. (III), (VI), (IX), (XII), (XV) and (XVI)
- e. (III), (VI), (VIII), (XII), (XV) and (XVII).

(3 marks)

[< Answer >](#)

63. The Securities Exchange Board of India (SEBI) was considering a proposal to require companies to report the potential effect of employees' stock options on earnings per share (EPS). A random sample of 41 high-technology firms revealed that the new proposal would reduce EPS by an average of 13.8 percent, with a standard deviation of 18.9 percent. A random sample of 35 producers of consumer goods showed that the proposal would reduce EPS by 9.1 percent on average, with a standard deviation of 8.7 percent. On the basis of these samples, is it reasonable to conclude at significance level of 10 percent that the SEBI proposal will cause a greater reduction in EPS for high technology firms than the producers of consumer goods? Which of the following are **correct** for this test of hypothesis?

- I. A two-tailed test of hypothesis on difference between means is appropriate.
- II. A one tailed test of hypothesis on difference between means is appropriate.
- III. The estimated standard error of difference between the means is 2.559.
- IV. The estimated standard error of difference between the means is 3.768.
- V. The estimated standard error of difference between the means is 3.298.
- VI. The standardized difference between the sample means is 1.836.
- VII. The standardized difference between the sample means is 1.247.
- VIII. The standardized difference between the sample means is 1.425.
- IX. At a significance level of 10 percent, we can conclude that the SEBI proposal will not cause a greater reduction in EPS for high technology firms than the producers of consumer goods.
- X. At a significance level of 10 percent, we can conclude that the SEBI proposal will cause a greater reduction in EPS for high technology firms than the producers of consumer goods.

- a. (I), (III), (VI) and (IX)
- b. (II), (V), (VIII) and (X)
- c. (II), (IV), (VII) and (IX)
- d. (I), (IV), (VII) and (X)
- e. (II), (V), (VI) and (IX).

(3 marks)

[< Answer >](#)

64. Goofy Toys is trying to establish a relationship between the number of toys that can be sold in a given month and the price that can be realized for one of its new toys, *Scooby Doo*. The marketing head of the company has provided the following information:

Price (in Rs.)	28	26	25	24	21	19
Quantity that can be sold	135	160	170	180	215	235

The dealer establishes a linear regression equation based on the above information to predict the quantity of toys he can sell for a given price. The dealer also wishes to test the accuracy of the slope of the regression line. Which of the following statements are **true** about the regression equation?

- I. The estimated number of toys that can be sold by the dealer at a price of Rs 30 is approximately 114.
- II. The estimated number of toys that can be sold by the dealer at a price of Rs 30 is approximately 125.
- III. The estimated number of toys that can be sold by the dealer at a price of Rs 30 is approximately 130.
- IV. The standard error of estimate for the regression line is approximately 4.256.
- V. The standard error of estimate for the regression line is approximately 10.568.
- VI. The standard error of estimate for the regression line is approximately 24.346.
- VII. The standard error of b coefficient is 0.5698.
- VIII. The standard error of b coefficient is 1.6785.

IX. The standard error of **b** coefficient is 4.3568.

- a. (I), (IV), and (VII) b. (II), (V) and (VIII)
 c. (III), (VI), and (IX) d. (II), (IV) and (VII)
 e. (I), (V) and (IX).

(3 marks)

65. The Gulmarg Winter Sports resort has just recently tabulated its data on the number of customers (in thousands) it has had during each season of the last 5 years. [< Answer >](#)

	Spring	Summer	Fall	Winter
1999	200	300	125	325
2000	175	250	150	375
2001	225	300	200	450
2002	200	350	225	375
2003	175	300	200	350

Which of the following is the **correct** combination of seasonal index for the quarters Spring, Summer, Fall and Winter respectively? (Use four quarter moving average)

- a. 45.36, 65.58, 51.36, 95.64 b. 48.75, 75.00, 45.00, 93.75
 c. 64.56, 105.45, 71.25, 126.56 d. 73.37, 113.34, 66.02, 147.27
 e. 84.26, 125.64, 75.84, 156.56.

(3 marks)

66. Sportsline Bicycle Parts manufactures precision ball bearings for wheel hubs, bottom brackets, headsets, and pedals. Ashok is responsible for quality control at Sportsline. He has been checking the output of the 5 mm bearings used in front wheel hubs. For each of last 6 hours, he has sampled 5 bearings, with the following results: [< Answer >](#)

Hour	Bearing Diameter (mm)				
1	5.03	5.06	4.86	4.90	4.95
2	4.97	4.94	5.09	4.78	4.88
3	5.02	4.98	4.94	4.95	4.80
4	4.92	4.93	4.90	4.92	4.96
5	5.01	4.99	4.93	5.06	5.01
6	5.00	4.95	5.10	4.85	4.91

Ashok wishes to control the variation in the bearing diameter by plotting a $\bar{X}$ chart. Which of the following is **true** for the given sample?

- I. The grand mean of the bearing diameter is 4.953 mm.
 II. The grand mean of the bearing diameter is 5.001 mm.
 III. The mean of sample range is 0.195 mm.
 IV. The mean of sample range is 0.201 mm.
 V. The Lower Control Limit of the $\bar{X}$ chart will be 4.841 mm.
 VI. The Lower control Limit of the $\bar{X}$ chart will be 4.891 mm.
 VII. The Upper Control Limit of the $\bar{X}$ chart will be 5.066 mm.
 VIII. The Upper Control Limit of the $\bar{X}$ chart will be 5.112 mm.
 IX. On the basis of the control chart it can be concluded that the process is out of control.
 X. On the basis of the control chart it can be concluded that the process is in control.
- a. (I), (III), (V), (VII) and (X) b. (II), (IV), (VI), (VIII) and (IX)
 c. (I), (IV), (VI), (VIII) and (X) d. (II), (III), (V), (VII) and (IX)
 e. (I), (III), (VI), (VIII) and (X).

(3 marks)

67. The post office is interested in modeling the “mangled-letter” problem. It has been suggested that any letter sent to a certain area has a 0.15 chance of being mangled. Because the post office is so big, it can [< Answer >](#)

be assumed that two letters' chances of being mangled are independent. A sample of 310 people was selected, and two test letters were mailed to each of them. The number of people receiving zero, one, or two mangled letters was 260, 40, and 10, respectively. The post office wants to test whether it is reasonable to conclude that the number of mangled letters received by people follows a binomial distribution with p equal to 0.15, at the significance level of 10 percent. Which of the following is **true** for the above test for the goodness of fit?

- I. I. The chi square statistic for the above test for the goodness of fit is 5.325.
- II. II. The chi square statistic for the above test for the goodness of fit is 16.325.
- III. III. The chi square statistic for the above test for the goodness of fit is 26.325.
- IV. IV. The degrees of freedom for the distribution is 2.
- V. V. The degrees of freedom the distribution is 3.
- VI. VI. The degrees of freedom for the distribution is 9.
- VII. VII. The critical value for the test for the given degrees of freedom is 4.605.
- VIII. VIII. The critical value for the test for the given degrees of freedom is 6.251.
- IX. IX. The critical value for the test for the given degrees of freedom is 14.684.
- X. X. At 10 percent level of significance, it can be concluded that the distribution of the mangled letters received by people follows a binomial distribution with $p = 0.15$.
- XI. XI. At 10 percent level of significance, it can be concluded that the distribution of the mangled letters received by people does **not** follow a binomial distribution with $p = 0.15$.

- a. (I), (IV), (VII) and (X) b. (II), (V), (VIII) and (XI)
- c. (III), (IV), (VII) and (XI) d. (III), (VI), (IX) and (X)
- e. (II), (V) (VIII) and (X).

(3 marks)

68. A research company has designed three different systems to clean up oil spills. The following table contains the results, measured by how much surface area (in square meters) is cleared in one hour. The data were found by testing each method in several trials.

[< Answer >](#)

System A	55	60	63	56	59	55
System B	57	53	64	49	62	
System C	66	52	61	57		

The research team wants to test the results whether the three systems are equally effective at a significance level of 5%. Which of the following are **true** for the above test?

- I. I. The estimate of the population variance based on the variance among the sample means is 4.467.
- II. II. The estimate of the population variance based on the variance among the sample means is 4.501.
- III. III. The estimate of the population variance based on the variances within the samples is 26.00.
- IV. IV. The estimate of the population variance based on the variances within the samples is 22.28.
- V. V. The critical value for the given test is 3.89.
- VI. VI. The critical value for the given test is 3.29.
- VII. VII. The critical value for the given test is 3.49.
- VIII. VIII. At a significance level of 5 percent, the research team should conclude that there is **no** significant difference in the surface area cleaned in one hour by the three systems.
- IX. IX. At a significance level of 5 percent, the research team should conclude that there is a significant difference in the surface area cleaned in one hour by the three systems.

- a. (II), (IV), (VI) and (IX) b. (I), (III), (V) and (VIII)
- c. (I), (III), (VII) and (IX) d. (II), (IV), (V) and (VIII)
- e. (I), (III), (VI) and (VIII).

(3 marks)

[< Answer >](#)

69. A management professor has given careful thought to the design of examinations. In order for him to be reasonably certain that an exam does a good job of distinguishing the differences in achievement shown by the students, the standard deviation in the scores on the examination cannot be too small. On the other hand, if the standard deviation is too large, there will tend to be a lot of very low scores, which is bad for the student morale. Past experience has led the professor to believe that a standard deviation of about 15 points on a 100-point exam indicates that the exam does a good job of balancing these two objectives. The professor just gave an examination to his class of 26 students. The mean score was 75 and the sample standard deviation was 16. The professor wants to test whether the test results meet his goodness criterion at a significance level of 10%. Which of the following is **true** for the above test?

- I. I. The chi square statistic for the given situation is 28.44.
- II. II. The chi square statistic for the given situation is 29.58.
- III. III. The chi square statistic for the given situation is 41.64.
- IV. IV. The critical region is in the right tail with the critical value as 35.563.
- V. V. The critical region is in both tails of the distribution with the critical values as 14.611 and 37.652.
- VI. VI. The critical region is in both tails of the distribution with the critical values as 15.379 and 38.885.
- VII. VII. At a significance level of 10 percent, the management professor should conclude that the standard deviation of the scores obtained by students is equal to 15 percent.
- VIII. VIII. At a significance level of 10 percent, the management professor should conclude that the standard deviation of the scores obtained by students is **not** equal to 15 percent.
- IX. IX. At a significance level of 10 percent, the management professor should conclude that the standard deviation of the scores obtained by students is more than 15 percent.

- a. (I), (IV) and (VIII)
- b. (II) (IV) and (VII)
- c. (III), (IV) and (IX)
- d. (III), (VI) and (VIII)
- e. (I), (V) and (VII).

(3 marks)

[< Answer >](#)

70. Darjeeling Tea has developed a new blend of tea, *Good Day* having superior taste and aroma than the existing brands. It wants to test market the new product before launching it nationwide. It chooses four cities and launches the product in these cities. After a month of the launch, they surveyed the individuals in these cities to find whether they liked this brand or not. The result of the survey is as follows:

	Delhi	Mumbai	Kolkata	Chennai	Total
No. of people who liked <i>Good Day</i>	71	75	58	79	283
No. of people who didn't like <i>Good Day</i>	29	45	32	31	137
Total household sampled in each city	100	120	90	110	420

The manager of Darjeeling Tea wants to know whether the proportion of people who like the new brand of tea, *Good Day* is same in all the four cities at a significance level of 10%. Which of the following is **true** for the above test?

- I. I. The chi square statistic is 3.2356.
 - II. II. The chi square statistic is 5.3468.
 - III. III. The chi square statistic is 7.3652.
 - IV. IV. The number of degrees of freedom for the contingency table is 3.
 - V. V. The number of degrees of freedom for the contingency table is 6.
 - VI. VI. The number of degrees of freedom for the contingency table is 8.
 - VII. VII. At a significance level of 10 percent the manager should conclude that there is **no** significant difference in the proportion of people who like their tea in the four cities.
 - VIII. VIII. At a significance level of 10 percent the manager should conclude that there is a significant difference in the proportion of people who like their tea in the four cities.
- a. (I), (IV) and (VII)
 - b. (II), (V) and (VII)

- c. (III), (VI) and (VIII) d. (I), (V), and (VIII) e. (II), (VI) and (VII).

(3 marks)

END OF SECTION B

Suggested Answers

Quantitative Methods – II (132) : January 2004

1. Answer :

Reason : (a) This is a true statement. A discrete random variable can assume only a limited number of values. If the number of students attending a particular class ranges from 50 to 60. It is an example of discrete random variable, as this variable can assume numerical values which are positive integers lying between 50 and 60. It cannot assume any fractional value.

(b) This is a true statement. A continuous random variable can assume any value within a given range. The maximum temperature of a city for a given day is an example of continuous random variable as it can assume any value between a given range i.e. it can have values like 29.35 °C, 29.48 °C, 29.57 °C etc. between 29 °C to 30 °C.

(c) This is **not a true** statement for random variables. A random variable is defined as a value or magnitude that changes from occurrence to occurrence in no predictable sequence.

(d) This is a true statement for random variables. A random variable is a variable, which assumes different numerical values as a result of random experiments.

(e) This is a true statement for random variables. The value of the random variable changes from occurrence to occurrence in no predictable sequence.

2. Answer :

Reason : (a) This statement is true for probability distribution. Usually, the past behaviour of the variable is studied and the frequency distribution of the data is formed. If the past behaviour can be taken as a representative pattern for the future also, then the past frequency distribution can be used as a guide for predicting the future values of the random variable.

(b) This is a true statement for probability distribution. The probability distribution gives an idea of the likely values of a random variable and the chances of occurrence of the various values.

(c) This is a true statement for probability distribution. If the past behaviour does not help in forming a probability distribution then, a subjective probability distribution of the likely future values is formed on the basis of expert opinion.

(d) This is **not true** for probability distribution. We can assign probabilities to all likely values of the random variable on the basis of the past data.

(e) This is a true statement for probability distribution. The probability distribution for a random variable provides a probability for each possible value and that these probabilities must sum to 1.

3. Answer :

Reason : (a) This is true for expected value of the random variable. The concept of expected value of the random variable is used for making decisions under conditions of uncertainty.

(b) This is true for expected value of the random variable. The expected value is the mean of the probability distribution.

(c) This is true for expected value of the random variable. The mean is computed as the weighted average of the values that the random variable can assume.

(d) This is true for expected value of the random variable. The probabilities assigned to each random variables are used as weights. Thus expected value of the random variable is computed by summing up the random variables multiplied by their respective probabilities of occurrence. $E [X] = \sum X P (X)$.

[TOP](#)

[TOP](#)

[TOP](#)

(e) This is **not true** for expected value of the random variable. The expected value of the random variable does not predict the value that the random variable will assume in future. Although there are many different possible values that the random variable can take, the expected value is a single figure. We say that the random variable is most likely to assume this value, it does not predict the value as such.

4. Answer :

Reason : (a) This is a wrong answer. If the covariance of two random variables X and Y is equal to zero then it can be said that the random variables are independent of each other. If the variables are dependent on each other then the covariance will assume any non-zero value.

(b) This is a wrong answer. If the coefficient of correlation is positive then we say that there is a positive correlation between two variables.

(c) This is a wrong answer. If the coefficient of correlation is negative then we say that there is a negative correlation between two variables.

(d) This is the **correct** answer. If the covariance of two random variables X and Y is equal to zero then it can be said that the random variables are independent of each other.

(e) This is a wrong answer. If the covariance of the two random variables is zero it does not mean that the variances of the individual variables will be zero.

5. Answer :

Reason : (a) This is a wrong answer. If the balls are replaced after inspection then it will follow the Bernoulli process and can be described by the binomial distribution.

(b) This is the **right** answer. If the Bernoulli experiment is repeated without replacement, it follows hypergeometric distribution. Therefore the random variables in the given case are best described by hypergeometric distribution.

(c) This is the wrong answer. Any variable studied for a large population is assumed to follow a normal distribution. If we draw samples of size more than 30 from any population the sample means are approximately normally distributed. But in our case the distribution of the random variable can be best described by hypergeometric distribution.

(d) This is a wrong answer. If $\ln(X)$ is a normally distributed random variable, the X is said to be a lognormal variable. The random variables in given case are not described by lognormal distribution.

(e) This is a wrong answer. t-distribution is used for testing of hypothesis when the sample size is 30 or less than 30 and the population standard deviation is not known. The random variables in the given case are not described well with the help of t-distribution.

6. Answer :

Reason : (a) This is an advantage of sampling. The population may be too large to be studied in full. In such situations it is always better to use sampling compared to the study of whole population.

(b) This is an advantage of sampling. A study of sample is usually cheaper than a study of the population. As the number of observations required is less the cost incurred in obtaining and processing these information is also less for sampling.

(c) This is an advantage of sampling. Sampling usually provides information

[TOP](#)

[TOP](#)

[TOP](#)

quicker than a census. Since the time required for studying the whole population is large, it is advisable to go for sampling to reduce the time spent in the studying the whole population.

(d) This is an advantage of sampling. In destructive testing, where the product becomes unusable after the performance of test, we cannot test the whole population. In such cases sampling is the only method of testing.

(e) This is **not** an advantage of sampling. In a sampling exercise we use a small number of samples to study the whole population the degree of accuracy is not very high. For accurate information census is always the preferable method compared to sampling.

7. Answer :

Reason : (a) This is the **right** answer. In cluster sampling, we divide the population into groups, or *clusters*, and then select a random sample of these clusters. We assume that these individual clusters are representative of the population as a whole. Then we choose a certain number of clusters randomly. Every household of these clusters is interviewed.

(b) This is a wrong answer. In stratified sampling we divide the population into relatively homogeneous groups, called strata. Then we use one of two approaches. Either we select at random from each stratum a specified number of elements corresponding to the proportion of that stratum in the population as a whole or we draw an equal number of elements from each stratum and give weight to the results according to the stratum's proportion of total population.

(c) This is a wrong answer. In systematic sampling, elements are selected from the population at a uniform interval that is measured in time, order, or space. In this method of sampling each element has an equal chance of being selected but each sample does not have an equal chance of being selected.

(d) This is a wrong answer. In the judgmental sampling the sample is selected according to the judgment of the investigators or experts.

(e) This is a wrong answer. In simple random sampling we do not divide the population in groups. We select a certain number of elements randomly. Here each possible sample has an equal chance of being selected and each item in the entire population also has an equal chance of being selected.

8. Answer :

Reason : (a) This is the wrong answer. The standard error of the sample mean is not the mean of the sampling distribution of the sample means.

(b) This is the **right** answer. The standard error of sample mean is the standard deviation of the sampling distribution of the sample means. Standard error measures the variability arising from sampling error due to chance. Standard error is used as a tool in tests of hypotheses or tests of significance.

(c) This is the wrong answer. The standard error of sample mean is not the variance of the sampling distribution of the sample means.

(d) This is the wrong answer. The standard error of sample mean is not the standard deviation of the sampling distribution of the sample proportion.

(e) This is the wrong answer. The standard error of sample mean is not the mean of the sampling distribution of the sample standard deviation.

9. Answer :

[<](#)
[TOP](#)
[>](#)

[<](#)
[TOP](#)
[>](#)

[<](#)
[TOP](#)
[>](#)

Reason : (a) This is a true statement. The probability that we associate with an interval estimate is called the confidence interval. This probability, then, indicates how confident we are that the interval estimate will include the population parameter. A higher probability means more confidence.

(b) This is a true statement. Using high confidence level is not always desirable, because high confidence levels produce large confidence interval, and such large intervals are not precise; they give very fuzzy estimates.

(c) This is a **false** statement. The number of degrees of freedom used in t-distribution estimation are equal to $(n-1)$, where n is the sample size. We lose one degree of freedom by estimating the standard deviation of the population with the help of the sample standard deviation.

(d) This is a true statement. The t distribution need not be used in estimating if you know the standard deviation of the population. We can use the normal distribution in such case.

(e) This is a true statement. The confidence interval is the range of the estimate we are making. An interval estimate describes a range of values within which a population parameter is likely to lie.

10. Answer :

Reason : (a) This is true for hypothesis testing. The probability that the null hypothesis is true but rejected is α . In other words probability of committing type I error is α , the level of significance.

(b) This is true for hypothesis testing. The probability that the null hypothesis is false but is accepted is β . In other words, the probability of committing type II error is β .

(c) This is a true statement. $(1-\beta)$ is known as the power of test. A high $(1-\beta)$ implies that the test is doing exactly what it should be doing i.e. Rejecting the null hypothesis when it is false. A low $(1-\beta)$ indicates poor performance.

(d) This is a **false** statement. A high $(1-\beta)$ implies that the test is doing exactly what it should be doing i.e. Rejecting the null hypothesis when it is false. A low $(1-\beta)$ indicates poor performance.

(e) This is a true statement. We can make two types of errors in hypothesis testing namely Type I error and Type II error. There is a trade-off between two types of errors: The probability of making one type of error can be reduced only if we are willing to increase the probability of making the other type of error.

11. Answer :

Reason : (a) This is true for correlation analysis. If the coefficient of correlation between two variables is close to one, then there is a strong correlation between the two variables.

(b) This is true for correlation analysis. If the slope of a regression equation is positive then the coefficient of correlation between the variables involved is positive.

(c) This is **false** for correlation analysis. Even a highly significant correlation does not necessarily mean that a cause and effect relationship exists between the two variables. Correlation does not necessarily imply causation or functional relationship though the existence of causation always implies correlation or association between two variables.

(d) This is true for correlation analysis. If the slope of a regression equation is negative then the coefficient of correlation between the variables involved is negative.

(e) This is a true statement. If the coefficient of correlation between two variables is equal to zero, then there is a no correlation between the two variables.

>

TOP
<

TOP
<

12. Answer :

Reason : (a) This is a wrong answer. The sum of the squared values of the vertical distances from each plotted point to the line should be minimum (**not** maximum).

(b) This is the **right** answer. According to the 'method of least squares' criterion, the regression line should be drawn on the scatter diagram in such a way that the sum of the squared values of the vertical distances from each plotted point to the line are minimum.

(c) This is the wrong answer. The sum of the squared values of the vertical (**not** the horizontal) distances from each plotted point to the line should be minimum (**not** maximum).

(d) This is the wrong answer. The sum of the squared values of the vertical (**not** the horizontal) distances from each plotted point to the line should be minimum.

(e) This is the wrong answer. The sum of the squared values of the vertical distances from each plotted point to the line should be minimum (not necessarily zero).

13. Answer :

Reason : (a) This is a wrong answer. The standard error of the estimate of a regression line is not the slope of the regression line.

(b) This is a wrong answer. The standard error of the estimate of a regression line is not The square root of the coefficient of determination.

(c) This is a wrong answer. The standard error of the estimate of a regression line is not the Y intercept of the regression line.

(d) This is the **right** answer. The standard error of the estimate of a regression line is the measure of the variability of the observed values around the regression line.

(e) This is a wrong answer. The standard error of the estimate of a regression line is not the coefficient of correlation between two variables.

14. Answer :

Reason : (a) This is the wrong answer. The coefficient of determination r^2 measures how good the fit is and gives the variation explained by the model.

(b) This is the right answer. If r be the coefficient of correlation between two variables X and Y , then the variation that is not explained by the regression line is equal to $1 - r^2$.

(c) This is the wrong answer. ' r ' is the coefficient of correlation and it signifies the degree of association between two variables.

(d) This is the wrong answer. ' $1 - r$ ' does not give the variation that is not explained by the model.

(e) This is the wrong answer. ' $1 + r^2$ ' does not have any meaning associated with it.

15. Answer :

Reason : (a) This is true for multiple correlation coefficient. It depends on the correlation coefficient between Y and X_1 .

(b) This is true for multiple correlation coefficient. It depends on the correlation coefficient between X_1 and X_2 .

(c) This is true for multiple correlation coefficient. It depends on the correlation coefficient between Y and X_2 .

(d) This is true for multiple correlation coefficient. It will take values between 0 and 1.

(e) This is **false** for multiple correlation coefficient. It will take values between 0 and 1. It does not take negative values.

16. Answer :

Reason : (a) This is the right answer. The partial coefficient of correlation between Z and X when Y is kept constant is equal to the simple coefficient of correlation between Z and X, if the coefficient of correlation between Z and Y and coefficient of correlation between X and Y are zero.

(b) This is the wrong answer. The partial coefficient of correlation between Z and X when Y is kept constant is not equal to simple coefficient of correlation between Z and Y for the given condition.

(c) This is the wrong answer. The partial coefficient of correlation between Z and X when Y is kept constant is not equal to simple coefficient of correlation between X and Y for the given condition.

(d) This is the wrong answer. The partial coefficient of correlation between Z and X when Y is kept constant is not equal to partial coefficient of correlation between Z and Y when X is kept constant for the given condition.

(e) This is the wrong answer. The partial coefficient of correlation between Z and X when Y is kept constant is not equal to partial coefficient of determination between Z and X when Y is kept constant for the given condition.

17. Answer :

Reason : (a) This is a wrong answer The given ratio does not follow a normal distribution.

(b) This is a wrong answer The given ratio does not follow a t distribution.

(c) This is the **right** answer. The ratio $F = \frac{RSS/1}{ESS/(n-2)}$ follows an F distribution with 1 degree of freedom in the numerator and (n-2) degrees of freedom in the denominator. A high F ratio signifies that the regression model is working fine and a low value of F means the model should be rejected.

(d) This is a wrong answer The given ratio does not follow a chi square distribution.

(e) This is a wrong answer The given ratio does not follow a binomial distribution.

18. Answer :

Reason : (a) This is a wrong answer. The standard error of the estimate is not calculated using the number of degrees of freedom as $n - 2$.

(b) This is a wrong answer. The standard error of the estimate is not calculated using the number of degrees of freedom as $n - k$.

(c) This is the **right** answer. In a multiple regression with **n** number of observations and **k** independent variables, the standard error of the estimate is calculated using the number of degrees of freedom as $n - k - 1$.

(d) This is a wrong answer. The standard error of the estimate is not calculated using the number of degrees of freedom as $n - k + 1$.

(e) This is a wrong answer. The standard error of the estimate is not calculated using the number of degrees of freedom as $n + k - 1$.

19. Answer :

Reason : (a) This is the wrong answer. The estimated values of the dependent variable follows a normal distribution if the number of observations is large. If the number of observations is more than 30 then the estimated value is likely to follow a normal distribution.

(b) This is the **right** answer. In the regression analysis for a set of observations less than 30, the estimated values of the dependent variable is likely to follow a t distribution with $(n - 2)$ degrees of freedom for linear regression and $(n - k - 1)$ for multiple regression.

(c) This is the wrong answer. The estimated values of the dependent variable do not follow a F distribution.

(d) This is the wrong answer. The estimated values of the dependent variable do not follow a chi square distribution.

(e) This is the wrong answer. The estimated values of the dependent variable do not follow a lognormal distribution.

20. Answer :

Reason : (a) This is a wrong answer. This problem is not referred as Random variations.

(b) This is a wrong answer. This problem is not referred as Non-random variations.

(c) This is a wrong answer. This problem is not referred as irregular variations.

(d) This is the right answer. In multiple-regression analysis, the regression coefficients often become less reliable as the degree of correlation between the independent variable increases. This problem is referred by statisticians as Multicollinearity.

(e) This is a wrong answer. This problem is not referred as cyclical variations.

21. Answer :

Reason : (a) This is true for index numbers. Index numbers are expressed as a percentage.

(b) This is true for index numbers. Index number for the base year is always 100.

(c) This is **not true** for index numbers. Index number is a ratio of the price, quantity or value of the commodities under study and therefore does not have any unit.

(d) This is true for index numbers. Index number is calculated as a ratio of the current value to a base value.

(e) This is not the correct answer.

22. Answer :

Reason : (a) This is true for index numbers. Index numbers help in establishing trends. Index numbers when analyzed reveal a general trend of the phenomenon under study.

(b) This is true for index numbers. Index numbers help in policy making. The dearness allowance paid to the employees is linked to the cost of living index, generally the consumer price index.

(c) This is true for index numbers. Usually index numbers are used to determine the purchasing power of the money. Suppose the consumer price index for the urban non-manual employees increased from 100 in 1984 to 202 in 1992, the real purchasing power of the rupee can be found out as $100/202=0.495$. It indicates that if rupee was worth 100 paise in 1984 its purchasing power is 49.5 paise in 1992.

(d) This is true for index numbers. Index numbers help in deflating time series data. It plays an important role in adjusting the original data to reflect reality. For example, nominal income can be transformed into real income by using income deflators.

<
TOP
>

<
TOP
>

<
TOP
>

(e) This is **not true** about index numbers. The value index is a combination index. It combines price and quantity changes to present a more spatial comparison. The value index as such measures changes in net monetary worth. Though the value index enables comparison of value of a commodity in a year to the value of that commodity in a base year, it has limited use. Its limited use is owing to the inability of the value index to distinguish the effects of price and quantity separately.

23. Answer :

Reason : (a) This is a wrong answer. Marshall-Edgeworth method uses both the current year as well as the base year quantities as weights. Therefore it is not biased towards the base year or the current year.

(b) This is a wrong answer. Fisher's ideal index is free from any bias. Both the current year and base year prices and quantities are taken into account by this index.

(c) This is a wrong answer. Laspeyres index tends to overestimate the rise in prices or has an upward bias.

(d) This is the **right** answer. The Paasche index generally tends to underestimate the prices and has a downward bias. Because people tend to spend less on goods when their prices are rising, the use of the Paasche index which bases on current weighting, produces an index which does not estimate the rise in prices rightly showing a downward bias. Since all prices or all quantities do not move in the same order, the goods which have risen in price more than others at a time when prices in general are rising, will tend to have current quantities relatively small than the corresponding base quantities and they will thus have less weight in the Paasche index.

(e) This is a wrong answer.

24. Answer :

Reason : (a) This is the wrong answer. The weighted average of price relatives using base values as weights is not same as the unweighted aggregates price index.

(b) This is the wrong answer. The weighted average of price relatives using base values as weights is not same as the unweighted aggregates quantity index.

(c) This is the **right** answer. The weighted average of price relatives using base values as weights is same as the Laspeyres price index.

(d) This is the wrong answer. The weighted average of price relatives using base values as weights is not same as the Laspeyres quantity index.

(e) This is the wrong answer. The weighted average of price relatives using base values as weights is not same as the Paasche's price index.

25. Answer :

Reason : (a) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to one.

(b) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to one hundred.

(c) This is the **right** answer. According to the factor reversal test if we interchange the price and quantity terms in an index number then the product of the original index number and the index number obtained by reversing the factors should be equal to the value index.

[TOP](#)

[TOP](#)

[TOP](#)

(d) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to the Fisher's ideal index.

(e) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to the chain index number.

26. Answer :

Reason : (a) This is the **right** answer. A time series consisting of annual data for longer periods is depicted by trend lines. This facilitates us to isolate the component of secular trend variation from the series and examine it for cyclical, seasonal and irregular component. We use "Residual Method" to isolate the cyclical variation component. This method uses one of the two measures namely Percent of trend and relative cyclical residual measure. Both these measures are expressed in terms of percentage.

(b) This is the wrong answer. Cyclical component of the time series is not isolated by Linear regression analysis.

(c) This is the wrong answer. Cyclical component of the time series is not isolated by non-linear regression analysis.

(d) This is the wrong answer. Seasonal component of the time series is isolated by ratio to moving average method.

(e) This is the wrong answer. Cyclical component of the time series is not isolated by multiple regression analysis.

27. Answer :

Reason : (a) This is the wrong answer. The variation in the demand of ice creams cannot be explained by secular trend.

(b) This is the wrong answer. The variation in the demand of ice creams cannot be explained by cyclical variation as the variation is observed in different seasons in a year.

(c) This is the **right** answer. The variation in the demand of ice creams can be well explained by seasonal variation. The variations observed in different seasons of a year can be best explained with the help of seasonal variation component. To find a long term forecasting first the time series has to be deseasonalized to get an accurate estimation.

(d) This is the wrong answer. The variation in the demand of ice creams cannot be explained by Irregular variation.

(e) This is the wrong answer.

28. Answer :

Reason : (a) This is a wrong answer. The movement of dependent variable above and below the secular trend line over periods longer than one year is not known as Seasonal variation.

(b) This is a wrong answer. The movement of dependent variable above and below the secular trend line over periods longer than one year is not known as Secular variation.

(c) This is a wrong answer. The movement of dependent variable above and below the secular trend line over periods longer than one year is not known as Irregular

[TOP](#)

[TOP](#)

[TOP](#)

variation.

(d) This is the right answer. The movement of dependent variable above and below the secular trend line over periods longer than one year is known as Cyclical variation.

(e) This is a wrong answer. The movement of dependent variable above and below the secular trend line over periods longer than one year is not known as Regular variation.

29. Answer :

Reason : (a) This is a wrong answer. The expression is not termed as Percent of trend measure.

(b) This is the right answer. If Y be the actual value and $\hat{Y}$ be the estimated value

using an estimating equation for a time series, then the expression $\left(\frac{Y - \hat{Y}}{\hat{Y}} \right) \times 100$ is termed as Relative cyclical residual measure.

(c) This is a wrong answer. The expression is not termed as standard error.

(d) This is a wrong answer. The expression is not termed as standard deviation.

(e) This is a wrong answer. The expression is not termed as Seasonal Index.

30. Answer :

Reason : (a) This is true for the ratio to moving average method. This method is used to identify the seasonal trend in a time series.

(b) This is true for the ratio to moving average method. An index is constructed with a base of 100.

(c) This is true for the ratio to moving average method. The magnitude of seasonal variations is measured by the individual deviations from the base of 100.

(d) This is true for the ratio to moving average method. The index is used to nullify the seasonal effects in the time series.

(e) This is not true for the ratio to moving average method. We divide the original data points with the relevant seasonal index to deseasonalize the time series.

31. Answer :

Reason : (a) This is a wrong answer. The $\bar{x}$ charts are constructed using the mean value as the central limit and three standard deviations from the mean as upper and lower control limits.

(b) This is the right answer. **R** charts are used to monitor the variability in the individual characteristics of the product. The control limit is the mean range and the upper and lower control limits are constructed with the standard deviation in the range values.

(c) This is the wrong answer. **p** charts deal with the attributes that are qualitative in nature. The measure that is applied here is the proportion of the data points satisfying the given criteria and the proportion of the data points that does not satisfy the criteria.

(d) This is the wrong answer.

(e) This is the wrong answer.

32. Answer :

Reason : (a) This is true about statistical process control. Quality standards in manufacturing process and the service industry can be achieved through the application of statistical methods.

- (b) This is true about statistical process control. The variability present in the manufacturing process can either be eliminated completely or minimized to the extent possible.
- (c) This is true about statistical process control. The variation in the specifications of the product due to wearing of the machine is a systematic or non random variation.
- (d) This is true about statistical process control. The variation due to an error made by the personnel in the measurement of the product specifications is a random variation.
- (e) This is **not true** about statistical process control. If the process is 'out of control', it indicates the presence of non-random pattern in the system. The management should first identify the cause of that variation and eliminate it. This elimination or the reduction of the systematic variation results in the process being brought 'in-control'. Once this is done, the whole process can be redesigned to improve or reduce the incidence of random or inherent variability.

33. Answer :

- Reason : (a) This is a wrong answer. If we observe that the sample means are hugging the center line then the deviations are minor and uniform in nature. This indicates that the variation has been reduced to great extent.
- (b) This is a wrong answer. If the observations are such that it can be described to be 'hugging the control limits', then this refers to the deviations being uniform in nature but having large magnitude. This indicates that the means of two different populations are being observed.
- (c) This is the right answer. If the observations can be described as jumps in the level around which the observations vary, then this pattern indicates that the process mean has been shifting.
- (d) This is a wrong answer. If some of the points lie outside the upper or lower control limits or if there is an increasing or decreasing trend in the observations, we conclude that the process is out of control.
- (e) This is a wrong answer.

34. Answer :

- Reason : (a) This is the **right** answer. In a control chart there are some observations that lie outside the upper and lower control limits. These points are referred to as the Outliers.
- (b) This is the wrong answer. Qualitative variables with only two categories are termed as attributes. This is monitored with the help of p chart.
- (c) This is the wrong answer. This is another name of random variation.
- (d) This is the wrong answer. this is the nonrandom, systematic variability in a process. It usually can be corrected without redesigning the entire process.
- (e) This is the wrong answer. This is the random variability inherent in the system. It usually cannot be reduced without redesigning the entire process.

35. Answer :

- Reason : (a) This is the wrong answer. This will be the null hypothesis for the test.
- (b) This is the wrong answer.
- (c) This is the wrong answer.
- (d) This is the wrong answer. This is a common error made while making the alternative hypothesis. We do not make a hypothesis that none of the proportions are

equal to each other. We only say that all the proportions are not all equal, although some of them might be equal to each other it does not mean that null hypothesis is true.

(e) This is the right answer. The alternative hypothesis is that all the three proportions are not all equal.

36. Answer :

Reason : (a) This is true for chi square distribution. There will be different chi square distribution for each degrees of freedom.

(b) This is true for chi square distribution. If the number of degrees of freedom is small, the curve is skewed to the right

(c) This is true for chi square distribution. As the number of degrees of freedom increases the curve also becomes symmetrical.

(d) This is not true for chi square distribution. The area under the chi square distribution remains equal to one irrespective of the number of degrees of freedom.

(e) This is the wrong answer.

37. Answer :

Reason : (a) This is the wrong answer. Analysis of variance is used to test the equality of means for samples drawn from more than two populations.

(b) This is the wrong answer. Regression analysis is used to estimate the future value of the dependent variable for a given independent variable with the help of a set of data.

(c) This is the wrong answer. Correlation analysis is used to find the degree of association between two variables.

(d) This is the correct answer. We use chi square test as a distributional goodness of fit. The managers can employ the chi square distribution to verify the appropriateness of the distribution employed by them and conclude whether any significant differences exist between the observed and theoretical distribution they have employed.

(e) This is the wrong answer. We use marginal analysis for finding the level at which the expected marginal loss is equal to expected marginal profit.

38. Answer :

Reason : (a) This is a step followed in the ANOVA. We determine an estimate of the population variance from the variance that exists among the sample means.

(b) This is a step followed in the ANOVA. We determine an estimate of the population variance from the variance that exists within the samples.

(c) This is **not** a step followed in ANOVA. The contingency table consisting observed and estimated frequency is prepared for testing equality of proportion from more than two samples.

(d) This is a step followed in the ANOVA. We accept the null hypothesis if the two estimates of the population variance are not significantly different.

(e) This is a step followed in the ANOVA. We reject the null hypothesis if the F ratio is too high and falls in the critical region.

39. Answer :

Reason : (a) This is a wrong answer. The number of degrees of freedom does not depend on the sample size.

(b) This is the wrong answer. The number of degrees of freedom does not depend on

[TOP](#)

[TOP](#)

[TOP](#)

[TOP](#)

the ratio of sample size and the population.

(c) This is the wrong answer. The number of degrees of freedom does not depend on the number of rows in the contingency table only.

(d) This is the wrong answer. The number of degrees of freedom does not depend on the number of columns in the contingency table only.

(e) This is the right answer. When the chi-square distribution is used as a test of independence, the number of degrees of freedom is related to both the number of rows and the number of columns in the contingency table.

40. Answer :

Reason : (a) This the wrong answer.

(b) This is the wrong answer.

(c) This is the wrong answer.

(d) This is the right answer. In two-tailed test of the two variances, we find the left tail of the F distribution using $F(n, d, \alpha)$ is equal to $1/F(d, n, 1-\alpha)$.

(e) This is the wrong answer.

[TOP](#)

41. Answer :

Reason : The expected number of moviegoers
 $= 500 \times 0.15 + 550 \times 0.20 + 600 \times 0.30 + 650 \times 0.35$
 $= 592.5 \approx \mathbf{593}$ (approx)

[TOP](#)

42. Answer :

Reason : Given that $p = 0.60$, $q = 0.40$, $n = 4$,

$$P(r) = {}^nC_r p^r q^{n-r}$$

$$\therefore P(3) = {}^4C_3 \times (0.60)^3 \times (0.40)^{4-3}$$

$$= \mathbf{0.3456}$$

[TOP](#)

43. Answer :

Reason : Since the sampling is done without replacement, so we will use

population correction factor, The standard error of mean $\sigma_{\bar{x}} = \left(\frac{\sigma}{\sqrt{n}} \right) \sqrt{\frac{N-n}{N-1}}$

$$= \left(\frac{12.5}{\sqrt{49}} \right) \sqrt{\frac{205-49}{205-1}}$$

$$= 1.56156 \approx \mathbf{1.56}$$

[TOP](#)

44. Answer :

Reason : We know that the regression coefficient $b = \frac{\text{Cov}(X, Y)}{V(X)}$

$$\therefore b = \frac{-121.5}{81} = -1.5$$

Given that, $\sum X = 55$ and $\sum Y = 33$ and the number of observations $n = 10$

Therefore $\bar{X} = \frac{55}{10} = 5.5$ and $\bar{Y} = \frac{33}{10} = 3.3$

Now, The Y intercept, $a = \bar{Y} - b\bar{X}$

$$= 3.3 - (-1.5 \times 5.5) = 11.55$$

The estimating equation is $\hat{Y} = 11.55 - 1.5 X$

[TOP](#)

The estimate of Y for X = 5, $\hat{Y} = 11.55 - 1.5 \times 5 = 4.05$

45. Answer :

Reason : We know that The variation in Y explained by X_1 if X_2 is kept constant is given by the partial coefficient of determination denoted by

$$R^2_{12.3} = \frac{(r_{12} - r_{13} \cdot r_{23})^2}{(1 - r_{13}^2)(1 - r_{23}^2)}$$

$$\begin{aligned} \text{Therefore } R^2_{12.3} &= \frac{(0.7 - 0.5 \times 0.6)^2}{(1 - 0.5^2)(1 - 0.6^2)} \\ &= \frac{(0.7 - 0.3)^2}{(0.75)(0.64)} \\ &= \end{aligned}$$

$$\frac{0.4^2}{0.48} = 0.3333$$

$$\approx 0.33$$

i.e 33%

46. Answer :

$$\begin{aligned} \text{Reason : Fisher's Ideal Price Index} &= \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}} \times 100 \\ &= \sqrt{\frac{1280}{1140} \times \frac{1680}{1500}} \times 100 \\ &= \mathbf{112.14} \end{aligned}$$

47. Answer :

$$\text{Reason : The Upper Control Limit for the R chart is given as, } UCL = \bar{R} + \frac{3d_3}{d_2} \bar{R}$$

$$\begin{aligned} &= 1.5 + \frac{3 \times 0.797}{3.078} \times 1.5 \\ &= 1.5 + 1.165 = \mathbf{2.665} \end{aligned}$$

48. Answer :

Reason : The F ratio is given as

$$\begin{aligned} \text{F ratio} &= \frac{\text{Estimate of population variance based on variance among the sample means}}{\text{Estimate of the population variance based on variance within the samples}} \\ &= \frac{3.362}{1.124} = \mathbf{2.991} \end{aligned}$$

49. Answer :

Reason : The estimating equation is $\hat{Y} = 60 + 1.5x$

$X = 2003$ and $\bar{X} = 1998$, therefore $x = 2003 - 1998 = 5$

The estimated value $\hat{Y} = 60 + 1.5 \times 5 = 67.5$

$$\text{The relative cyclical residual} = \frac{Y - \hat{Y}}{\hat{Y}} \times 100$$

$$= \frac{70 - 67.5}{67.5} \times 100$$

$$= \mathbf{3.7037}$$

50. Answer :

Reason : Probability that the normal variable with mean 12 and standard deviation 5 will take a value in the interval 5.6 and 21.8 is the area under the standard normal curve between the standard normal values $\frac{5.6 - 12}{5}$ and $\frac{21.8 - 12}{5}$.
The area between $z = -1.28$ and $z = 1.96$ is $0.3997 + 0.4750 = 0.8747$.
Therefore the required probability is **87.47 %**

[TOP](#)

51. Answer :

Reason : Given a probability distribution of paired data {X,Y}, we can compute the

X	Y	P(X,Y)	X×p	Y×p	X–E(X)	Y–E(Y)	
		(p)	(a)	(b)	(c)	(d)	(c).(d).(p)
40	35	0.40	16.0	14.0	13	6.5	33.8
25	25	0.50	12.5	12.5	-2	-3.5	3.5
-15	20	0.10	-1.5	2.0	-42	-8.5	35.7
Total			27	28.5			73.00

covariance using the formula $\text{Cov}(X,Y) = \sum [X - E(X)] [Y - E(Y)] P(X,Y)$

The calculations for the covariance

are shown in the table below:

Therefore the covariance of the returns is equal to 73.0 % squared.

52. Answer :

Reason : The probability that a randomly selected person is pro NDA is 0.55. Then the probability that he/she does not support NDA is $1 - 0.55 = 0.45$. Therefore We can assume that it follows a binomial distribution with probability of success $p = 0.55$ and probability of failure $q = 0.45$. The probability that atleast eight out of ten persons chosen are pro NDA is $P(8) + P(9) + P(10) =$

$${}^{10}C_8 \times (0.55)^8 \times (0.45)^2 + {}^{10}C_9 \times (0.55)^9 \times (0.45)^1 + {}^{10}C_{10} \times (0.55)^{10} \times (0.45)^0$$

$$= 0.0763 + 0.0207 + 0.0025$$

$$= \mathbf{0.0995}$$

[TOP](#)

53. Answer :

Reason : The z value in the normal table for 89.9 percent confidence level is 1.64. We want our estimate to be within 0.02, so,

$$Z \sigma_{\bar{p}} = 0.02$$

$$\text{And } z = 1.64$$

$$\text{Then } 1.64 \sigma_{\bar{p}} = 0.02$$

[TOP](#)

Substituting the value of $\sigma_{\bar{p}}$ in the equation,

$$1.64 \sqrt{\frac{pq}{n}} = 0.02$$

$$\text{or, } \sqrt{\frac{pq}{n}} = 0.0122$$

As the university feels that the proportion of students favouring the new system is $\frac{36000}{60000} = 0.60$. We can take this value as the estimate of the population parameter.

$$\text{Therefore, } \sqrt{\frac{0.6 \times 0.4}{n}} = 0.0122$$

$$\text{or, } n = \frac{0.6 \times 0.4}{(0.0122)^2} = 1612.4697$$

Therefore a sample size of **1613** would be appropriate to estimate the proportion of the students favouring the new system in an interval of ± 0.02 at a confidence level of 89.9%.

54. Answer :

Reason : According to Hurwicz criterion, the decision is made based on the weighted profit, which is calculated as,

Weighted profit = α (Maximum profit for alternative) + $(1-\alpha)$ (Minimum profit for alternative)

Hence Weighted profit for Portfolio A = $0.7 \times 30,000 + 0.3 \times 2000 = \text{Rs. } 21,600$

Weighted profit for Portfolio B = $0.7 \times 40,000 + 0.3 \times (-15,000) = \text{Rs. } 23,500$

Weighted profit for Portfolio C = $0.7 \times 50,000 + 0.3 \times (-30,000) = \text{Rs. } 26,000$

Weighted profit for Portfolio D = $0.7 \times 40,000 + 0.3 \times 5,000 = \text{Rs. } 29,500$

Weighted profit for Portfolio E = $0.7 \times 20,000 + 0.3 \times (-5,000) = \text{Rs. } 12,500$

Therefore according to Hurwicz criterion, the analyst should invest in **portfolio D**.

As portfolio D has the highest weighted profit.

55. Answer :

Reason : The sampling proportion is large as $n/N = 50/700 = 0.07$ (This is > 0.05)

Therefore we will use finite population correction factor to find standard error of mean,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{1.2}{\sqrt{50}} \sqrt{\frac{700-50}{700-1}} = 0.1636$$

Since the sample size is more than 30, we use normal distribution to find the interval estimate.

The confidence interval for the estimate is given as $[\bar{X} - z \sigma_{\bar{x}}, \bar{X} + z \sigma_{\bar{x}}]$ Therefore the interval estimate is $4.5 + 1.64 \times 0.1636$ and $4.5 - 1.64 \times 0.1636$ The interval estimate is **4.5 ± 0.2683** .

56. Answer :

Reason : We want to test the hypothesis

$H_0: \mu = 44.95$ (Null Hypothesis: Average retail price has not changed)

$H_1: \mu < 44.95$ (Alternative Hypothesis: Average retail price has reduced after sales promotion)

The standard error of mean = $\frac{\sigma}{\sqrt{n}} = \frac{5.75}{\sqrt{25}} = 1.15$

The standard normal value of the mean retail price of the sample = $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$
 $= \frac{42.95 - 44.95}{1.15} = \frac{-2}{1.15} = -1.739$

Looking at the table for the critical value corresponding the area under the normal curve as 2% in the left tail i.e. -2.05. We find that the sample statistic falls in the acceptance region therefore we accept null hypothesis and at significance level of 2 % we can conclude that the price of the tennis balls has **not** reduced significantly after the trade promotions.

57. Answer :

Reason : Let X be the number of soccer games played and Y be the number of accidents during weekends.

$\sum X = 146$, $\sum Y = 48$, $n = 7$

$\sum XY = 1101$,

$\sum X^2 = 3550$,

$\bar{X} = 20.86$, $\bar{Y} = 6.86$

The coefficient of regression $b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2} = \frac{1101 - 7 \times 20.86 \times 6.86}{3550 - 7 \times (20.86)^2} = 0.197$

The Y intercept $a = \bar{Y} - b \bar{X} = 6.86 - 0.197 \times 20.86 = 2.751$

Therefore the estimating equation is $\hat{Y} = 2.751 + 0.197 X$

The estimate of Y for $X = 33$, is $\hat{Y} = 2.751 + 0.197 \times 33 = 9.251 \approx \mathbf{9.25}$

58. Answer :

Reason : Let the number of advertisement shown be X and the number of cans of cola purchased be Y.

Then $\sum X = 21$, $\sum Y = 56$, $n = 7$

$\bar{X} = 3$, $\bar{Y} = 8$,

$\sum (X - \bar{X})(Y - \bar{Y}) = 52$,

$\sum (X - \bar{X})^2 = 32$ and $\sum (Y - \bar{Y})^2 = 132$

The coefficient of correlation $r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}$
 $= \frac{52}{\sqrt{32 \times 132}} = \mathbf{0.8001}$

59. Answer :

Reason : The Marshall-Edgeworth price index = $\frac{\sum (Q_0 + Q_1)P_1}{\sum (Q_0 + Q_1)P_0} \times 100$

The price and the quantity of the commodities consumed in the base year is described with the subscript '0' and the price and the quantity consumed in the current year is denoted with the subscript '1'.

$$\therefore \text{The Marshall-Edgeworth price index} = \frac{264.2}{162.7} \times 100 = \mathbf{162.38}$$

60. Answer :

Reason : We observe that the number of data points is odd. The working is shown below:

	Years (X)	Inventory (Y)	$x = X - \bar{X}$	$x.Y$	x^2
	1999	4620	-2	-9240	4
	2000	4910	-1	-4910	1
	2001	5490	0	0	0
	2002	5730	1	5730	1
	2003	5990	2	11980	4
Total	10005	26740	0	3560	10

$$\text{The mean of "X"} = \frac{10005}{5} = 2001$$

$$\text{We have } b = \frac{\sum xY}{\sum x^2} = \frac{3560}{10} = 356$$

$$a = \bar{Y} = 5348$$

Therefore the estimating equation is $\hat{Y} = 5,348 + 356x$ (in thousands rupees)

61. Answer :

Reason : The marginal profit per hour = Rs 16 – Rs 9 = Rs 7

The marginal loss per hour = Rs 9

The conditional profit (in rupees) for different levels of demand is tabulated below:

Available (hours)	Demand (hours)				Expected Profit (Rs.)
	10,000 (0.2)	12,000 (0.3)	14,000 (0.4)	16,000 (0.1)	
10,000	70,000	70,000	70,000	70,000	70,000
12,000	52,000	84,000	84,000	84,000	77,600
14,000	34,000	66,000	98,000	98,000	75,600
16,000	16,000	48,000	80,000	112,000	60,800

The manager expects to make maximum profit of Rs 77,600 if he has 12,000 mechanic hours available.

The number of hours a mechanic will work in a year = (52 weeks – 2 weeks) × 40 hours = 2,000 hours.

Therefore the number of mechanics required for 12,000 mechanic hours = $12,000 \div 2,000 = \mathbf{6 \text{ mechanics}}$.

[TOP](#)

[TOP](#)

Expected profit under certainty = $70,000 \times 0.2 + 84,000 \times 0.3 + 98,000 \times 0.4 + 112,000 \times 0.1 = \text{Rs } 89,600$

Therefore Expected value of perfect information (EVPI) = $\text{Rs } 89,600 - \text{Rs } 77,600 = \text{Rs } 12,000$.

62. Answer :

Reason : Anand Electronics wants to open a store if more than 15 percent of the students own a stereo system costing Rs 5,000 or more. They conduct a survey and wants to test whether they should open a store in Varanasi or not.

$H_0 : p = 0.15$ (Null hypothesis: The proportion of students owning a stereo costing Rs 5,000 or more is 15 percent)

$H_1 : p > 0.15$ (Alternative hypothesis: The proportion of students owning a stereo costing Rs 5,000 or more is more than 15 percent)

Therefore we will conduct a **right tailed** test of hypothesis.

The sample size is, $n = 300$

The sample proportion owning a stereo worth Rs 5,000 and more is, $p =$

$\frac{57}{300} = 0.19$ and $q = 1 - 0.19 = 0.81$

The standard error of proportion $\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.19 \times 0.81}{300}} = 0.02265$

The standardized value of the sample proportion is $z = \frac{\bar{p} - p_{H_0}}{\sigma_{\bar{p}}} = \frac{0.19 - 0.15}{0.02265} = 1.766$

At a significance level of 5 percent the critical value is **1.64**.

Since the standardized value of the sample proportion exceeds 1.64, it falls in the rejection region, we can conclude at a significance level of 5 percent that the number of students owning stereos worth Rs 5,000 and more is more than 15 percent.

Therefore they should open a store at Varanasi.

63. Answer :

Reason : We want to test on the basis of two samples, whether it is reasonable to conclude at significance level of 10 percent that the SEBI proposal will cause a greater reduction in EPS for high technology firms than the producers of consumer goods.

Therefore the Hypothesis are,

$H_0: \mu_1 = \mu_2$ [Null Hypothesis: The reduction in the EPS of the high technology companies, with mean μ_1 and the consumer goods companies, with mean μ_2 are equal]

$H_1: \mu_1 > \mu_2$ [Alternative Hypothesis: The reduction in the EPS in the high technology companies, with mean μ_1 is more than the reduction in consumer goods companies with mean μ_2]

Since we are testing whether the mean reduction in the EPS is more in the high technology companies than the consumer goods companies, **we will conduct a right tailed test.**

The sample statistics are as follows:

$n_1 = 41$, $\bar{x}_1 = 13.8$, $s_1 = 18.9$ and $n_2 = 35$, $\bar{x}_2 = 9.1$, $s_2 = 8.7$

Since both the samples have more than 30 observations we will use normal

distribution to test the hypothesis.

Estimated Standard error of the difference between two means,

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{18.9^2}{41} + \frac{8.7^2}{35}} = 3.2977 \approx \mathbf{3.298}$$

The standardized value of the difference between sample means,

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}} = \frac{(13.8 - 9.1) - 0}{3.298} = \mathbf{1.425}$$

At a significance level of 10 percent the critical value is **1.28**.

We see that the standardized value of the difference between the sample means falls in the rejection region as $1.425 > 1.28$. We can conclude at a significance level of 10 percent that the reduction in the EPS in the high technology company is **more than** the reduction in the EPS of the consumer goods company.

64. Answer :

Reason : Let X denote the price in Rs and Y denote the quantity (in thousands) of toys that can be sold for the given price X.

Then, $\sum X = 143$, $\sum Y = 1095$, $\bar{X} = 23.83$, $\bar{Y} = 182.5$, $n = 6$

$$\sum XY = 25490, \sum X^2 = 3463, \sum Y^2 = 206575.$$

$$\text{The regression coefficient, } \mathbf{b} = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} = \frac{6 \times 25490 - 143 \times 1095}{6 \times 3463 - 143^2} = -11.08$$

$$\text{The Y intercept } a = \bar{Y} - b \bar{X} = 182.5 - (-11.08) \times 23.83 = 446.51257 \approx 446.513$$

$$\text{Therefore the estimating equation is } \hat{Y} = 446.513 - 11.08 X$$

Estimated number of toys that can be sold at X= Rs 30,

$$\hat{Y} = 446.513 - (11.08 \times 30) = 114.113 \approx \mathbf{114 \text{ toys (approximately)}}$$

$$\text{Now the standard error of estimate is } S_e = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n - 2}}$$

$$= \sqrt{\frac{206575 - 446.513 \times 1095 - (-11.08)(25490)}{6 - 2}} = \mathbf{4.256}$$

$$\text{The standard error of the regression coefficient, } s_b = \frac{s_e}{\sqrt{\sum X^2 - n \bar{X}^2}} =$$

$$\frac{4.256}{\sqrt{3463 - 6 \times 23.83^2}} = \mathbf{0.5698}$$

65. Answer :

Reason : We compute the Percentage of the actual data to the centered moving average as follows:

[TOP](#)

Year	Season	Number of customers	Four Quarter Moving Total	Four Quarter Moving Average	Centered Moving Average	Percentage of actual to CMA
(1)	(2)	(3)	(4)	(5) = (4) ÷ 4	(6)	(7) = $\frac{(3)}{(6)} \times 100$
1999	Spring	200				
	Summer	300				
			950	237.50		
	Fall	125			234.375	53.33
			925	231.25		
	Winter	325			225	144.44
			875	218.75		
2000	Spring	175			221.875	78.87
			900	225.00		
	Summer	250			231.250	108.11
			950	237.50		
	Fall	150			243.750	61.54
			1,000	250.00		
	Winter	375			256.250	146.34
			1,050	262.50		
2001	Spring	225			268.750	83.72

1.100 275.00

	Summer	300			284.375	105.49
			1.175	293.75		
	Fall	200			290.625	68.82
			1.150	287.50		
	Winter	450			293.750	153.19
			1.200	300.00		
2002	Spring	200			303.250	65.98
			1.225	306.25		
	Summer	350			303.125	65.98
			1.150	287.50		
	Fall	225			284.375	79.12
			1.125	281.25		
	Winter	375			275.000	136.36
			1.075	268.75		
2003	Spring	175			265.625	65.88
			1.050	262.50		
	Summer	300			259.375	115.66
			1.025	256.25		
	Fall	200				
	Winter	350				

Now we calculate the modified mean as follows:

Year	Spring	Summer	Fall	Winter
1999	–	–	53.33	144.44
2000	78.87	108.11	61.54	146.34
2001	83.72	105.49	68.82	153.19
2002	65.98	117.89	79.12	136.36
2003	65.88	115.66	–	–

We cancel the two extreme values for each quarter and take the mean of the remaining values.

We get the modified means of the four quarters as follows:

Spring = 72.43, Summer = 111.89, Fall = 65.18 and Winter = 145.39.

Sum = 394.89

Therefore Adjusting constant = $\frac{400}{394.89} = 1.01294$

Now finding the seasonal Index as follows:

Quarter	Unadjusted means	× Adjusting constant	= Seasonal Index
Spring	72.43	1.01294	73.3672
Summer	111.89	1.01294	113.3379
Fall	65.18	1.01294	66.0234
Winter	145.39	1.01294	147.2713
Total			400

66. Answer :

Reason :

[TOP](#)

The $\bar{x}$ chart for the ball bearing diameter can be plotted as blow:

Hour	Bearing Diameter (mm)					Sample Mean	Sample Range
1	5.03	5.06	4.86	4.90	4.95	4.960	0.20
2	4.97	4.94	5.09	4.78	4.88	4.932	0.31
3	5.02	4.98	4.94	4.95	4.80	4.938	0.22
4	4.92	4.93	4.90	4.92	4.96	4.926	0.06
5	5.01	4.99	4.93	5.06	5.01	5.000	0.13
6	5.00	4.95	5.10	4.85	4.91	4.962	0.25
Sum						29.718	1.17

The Grand Mean, $\bar{\bar{x}} = 4.953$ and the Mean Range $\bar{R} = 0.195$

The d_2 value for the sample size 5 is **2.326**.

$$\text{The Upper Control Limit} = \bar{\bar{x}} + \frac{3 \bar{R}}{d_2 \sqrt{n}} = 4.953 + \frac{3 \times 0.195}{2.326 \times \sqrt{5}} = 5.0655 \approx \mathbf{5.066}$$

$$\text{The Lower control limits} = \bar{\bar{x}} - \frac{3 \bar{R}}{d_2 \sqrt{n}} = 4.953 - \frac{3 \times 0.195}{2.326 \times \sqrt{5}} = 4.8405 \approx \mathbf{4.841}$$

We observe that none of the sample means are out of the upper as well as lower control limits. Therefore, we can say that **the process is in control**.

67. Answer :

Reason : We have to test if the probability of receiving 'mangled letters' follows a binomial distribution.

[TOP](#)

H_0 : A binomial distribution with $p = 0.15$ is a good description of ‘mangled letters’.

H_1 : A binomial distribution with $p = 0.15$ is a **not** good description of ‘mangled letters’.

If the case of mangled letters follows a binomial probability distribution with $p = 0.15$, then

For $n = 2$ (two letters sent to a number of people)

$p = 0.15$ and $q = 1 - p = 0.85$

The probability of receiving zero mangled letter, exactly one mangled letter and exactly two mangled letters is given as :

$$P(0) = {}^2C_0(0.15)^0(0.85)^2 = 0.7225$$

$$P(1) = {}^2C_1(0.15)^1(0.85)^1 = 0.255$$

$$P(2) = {}^2C_2(0.15)^2(0.85)^0 = 0.0225$$

Therefore of the total number of people to whom the two letters are sent is 310, then the number of people who are expected to receive zero, one and two mangled letters will be,

$$n(0) = 0.7225 \times 310 = 223.975 \approx 224$$

$$n(1) = 0.255 \times 310 = 79.05 \approx 79$$

$$n(2) = 0.0225 \times 310 = 6.975 \approx 7$$

Now tabulating the observed and expected frequency as follows:

Number of people receiving mangled letters	Observed Frequency (f_o)	Expected Frequency (f_e)	$\frac{(f_o - f_e)^2}{f_e}$
0	260	224	5.786
1	40	79	19.253
2	10	7	1.286
Total	310	310	26.325

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 26.325$$

The chi square statistic

The degrees of freedom for the goodness of fit test is $(k - 1)$, where k is the number of classes observed. Therefore the number of degrees of freedom is $(3 - 1) = 2$.

The critical value for the 2 degrees of freedom at 10 percent significance level is = **4.605**.

We see that the chi square statistic falls in the rejection region. Therefore **we reject the null hypothesis** at 10 percent significance level. We conclude that the number of mangled letters received by people **do not** follow a binomial distribution with a

probability p equal to 0.15.

68. Answer :

Reason : We first calculate the between column variance as follows:

[TOP](#)

n	$\bar{x}$	$\bar{\bar{x}}$	$n (\bar{x} - \bar{\bar{x}})^2$
6	58	57.93	0.0294
5	57	57.93	4.3245
4	59	57.93	4.5796
		Total	8.9335

The estimate of the population variance based on the variance among the sample means

$$\hat{\sigma}^2 = \frac{\sum n_j (x_j - \bar{\bar{x}})^2}{k-1} = \frac{8.9335}{3-1} = 4.46675 \approx 4.467$$

Now we calculate variance within the column as follows:

	n	$\bar{x}$	$\sum (x - \bar{x})^2$	s^2	$\left(\frac{n-1}{n_T - k} \right) s^2$
System A	6	58	52	10.4	4.3333
System B	5	57	154	38.5	12.8333
System C	4	59	106	35.33	8.8325
				Total	25.9991

The estimate of the population variance based on the variance within the samples is

$$\hat{\sigma}^2 = \sum \left(\frac{n_j - 1}{n_T - k} \right) \times s_j^2 = 25.9991 \approx \mathbf{26.0}$$

Now calculating the F ratio as

$$F \text{ ratio} = \frac{\text{Estimate of the population variance based on the variance among the sample means}}{\text{Estimate of the population variance based on the variance within the samples}}$$

$$F \text{ ratio} = \frac{4.467}{26} = \mathbf{0.172}$$

Now looking at the critical value in the F distribution for a 5 percent significance level.

The degrees of freedom in numerator is, (number of samples – 1) = 3 – 1 = **2**

The degrees of freedom in denominator is, (Total sample size – number of samples) = 15 – 3 = **12**

The F value from the table is $F_{(2, 12, 5\%)} = \mathbf{3.89}$ (Critical value)

We find that the F statistic falls in the acceptance region; therefore we can conclude at 5 percent level of significance that **there is no significant difference** in the time taken by the three systems of cleaning the oil spills.

69. Answer :

Reason : The professor wants to test that whether the marks obtained by the students have standard deviation equal to 15 point in a test of 100 marks question paper.

[<](#)
[TOP](#)
[>](#)

$\sigma_{H_0} = 15$ (Hypothesized value of the population standard deviation)

$s = 16$ (sample standard deviation)

$n = 26$ (sample size)

$H_0: \sigma = 15$ (Null Hypothesis: The true standard deviation is 15 marks)

$H_1: \sigma \neq 15$ (Alternative Hypothesis: The true standard deviation is not 15 marks)

$\alpha = 10$ percent (Level of significance for testing the hypothesis)

The chi square statistic for the sample is

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(26-1) \times 16^2}{15^2} = \mathbf{28.44}$$

The number of degrees of freedom for the chi square distribution is $n - 1 = (26 - 1) = 25$.

This is a two-tailed test, therefore the values of the critical region in the both tails of the chi square distribution at 10 percent level of significance is **14.611** and **37.652**.

We observe that the test statistic falls in the acceptance region of the distribution. Hence we can conclude at a significance level of 10 percent that the standard deviation of the marks obtained by the students are 15 marks and meets the goodness criterion of the professor.

70. Answer :

Reason : We can put the null and alternative hypothesis for the test as follows:

$H_0: p_D = p_M = p_K = p_C$

[<](#)
[TOP](#)
[>](#)

(The proportion of people liking the *Good day* tea are same in all the four cities)

H_1 : p_D , p_M , p_K and p_C are not all equal.

(The proportion of people liking the *Good day* tea are not same in all the four cities)

$\alpha = 0.10$ level of significance for testing the hypothesis.

Calculating the expected frequency by using, $f_e = \frac{RT \times CT}{n}$;

Where, f_e is the expected frequency in a given cell

RT is row total for the row containing that cell

CT is the column total for the column containing that cell

N is the total number of observations

The expected frequency table:

	Delhi	Mumbai	Kolkata	Chennai
Like Good Day	67.38	80.86	60.64	74.12
Do not like Good Day	32.62	39.14	29.36	35.88

Now calculating the Chi square statistic as follows:

Row	Column	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
1	1	71	67.38	0.1945
1	2	75	80.86	0.4247
1	3	58	60.64	0.1149
1	4	79	74.12	0.3213
2	1	29	32.62	0.4017
2	2	45	39.14	0.8774
2	3	32	29.36	0.2374
2	4	31	35.88	0.6637
Total		420	420	3.2356

The Chi square statistic $\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 3.2356$

The number of degrees of freedom for the chi square distribution for the above situation is

(Number of rows – 1) $\times$ (Number of columns – 1) = (2 – 1) $\times$ (4 – 1) = 3

The critical value in the chi square distribution for 3 degrees of freedom and 10 percent level of significance is **6.251**.

We observe that the chi square statistic falls in the acceptance region. Therefore we conclude that at a significance level of 10 percent, the proportions of people who like the taste of Good day tea in the four cities are equal.

[< TOP OF THE DOCUMENT >](#)