

Question Paper

Quantitative Methods-II (132): January 2005

- Answer all questions.
- Marks are indicated against each question.

1. If a fair die is thrown twenty times, what is the probability of obtaining a minimum of 3 in a maximum of 17 throws? [Answer](#)

(a) 0.01759 (b) 0.01428 (c) 0.98572 (d) 0.98241 (e) 0.9997.

(2 marks)

2. The probability that a continuous random variable will assume a particular value is [Answer](#)

(a) Zero (b) Between zero and 0.50 (c) Between 0.50 and 1.00
(d) Equal to 0.50 (e) Unknown.

(1 mark)

3. A random variable, **Y**, has the following probability distribution: [Answer](#)

Y	20	40	60	80	100
Probability	0.15	0.20	0.30	0.25	0.10

15 values of **Y** have been observed. What is the probability that less than 12 observations are such that **40** ≤ **Y** ≤ **80**?

(a) 0.8441 (b) 0.1559 (c) 0.0668 (d) 0.4613 (e) 0.5387.

(2 marks)

4. **Z** is a binomial random variable. For 25 trials the standard deviation of **Z** is $\sqrt{6}$. [Answer](#)

What will be the standard deviation of **Z** for 75 trials?

(a) 6 (b) $\sqrt{6}$ (c) $3\sqrt{2}$ (d) 18 (e) $9\sqrt{2}$.

(1 mark)

5. If two fair dice are thrown simultaneously ten times, what is the likelihood that a minimum of 3 can be simultaneously observed in both the dice, on at least nine throws? [Answer](#)

(a) 0.00376 (b) 0.00030 (c) 0.00406 (d) 0.99594 (e) 0.44444.

(2 marks)

6. Given a probability distribution for a random variable, the covariance of the variable with itself is equal to [Answer](#)

(a) 0
(b) 1
(c) The variance of the variable
(d) The standard deviation of the variable
(e) The expected value of the variable.

(1 mark)

7. Which of the following is **false** with regard to a hypergeometric distribution? [Answer](#)

(a) The trials are not independent
(b) The probability of success is variable

- (c) The outcomes can be labeled as success or failure
- (d) The composition of the population remains unchanged
- (e) The population is finite.

(1 mark)

8. A continuous uniform probability distribution is described by

[Answer](#)

- (a) The expected value of the distribution
- (b) The variance of the distribution
- (c) The upper and lower limits of range of values of the distribution
- (d) The mid-point of the range of values of the distribution
- (e) Only the lower limit of the range of values of the distribution.

(1 mark)

9. **X** and **Y** are two dependent random variables. The following details are available:

[Answer](#)

Standard deviation of **X** = 6

Standard deviation of **Y** = 8

Let **Z** = 5**X** – 10**Y**

It is known that the standard deviation of **Z** is 70.

What is the value of covariance between the variables **X** and **Y**?

- (a) 36
- (b) 24
- (c) 64
- (d) 240
- (e) 2400.

(1 mark)

10. There are 25 students in a class, which consists of 14 boys and 11 girls. 5 students of the class were absent on a particular day. What is the probability that maximum two girls were absent?

[Answer](#)

- (a) 0.2072
- (b) 0.3768
- (c) 0.416
- (d) 0.584
- (e) 0.6232.

(2 marks)

11. A random variable, **Z**, has the following probability distribution:

[Answer](#)

Z	40	60	80	100	120
Probability	0.15	0.25	0.30	0.20	0.10

If twenty values of **Z** are observed, how many observations are expected to be less than 80?

- (a) 2
- (b) 4
- (c) 6
- (d) 8
- (e) 10.

(1 mark)

12. If each of the values of a random variable is transformed into a **z**-value, the distribution of the resulting values will have a standard deviation equal to

[Answer](#)

- (a) Zero
- (b) One
- (c) The mean of the original distribution
- (d) The standard deviation of the original distribution
- (e) A variable, depending upon the shape and spread of the original distribution.

(1 mark)

13. Which of the following is a **required** condition for the probability distribution of a discrete random variable?

[Answer](#)

- (a) Sum of all the probabilities = 0
- (b) Sum of all the probabilities = 1
- (c) Sum of all the probabilities < 0
- (d) Sum of all the probabilities < 1
- (e) Sum of all the probabilities > 1.

(1 mark)

14. If a **z**-value is to the left of the mean in the standard normal distribution, then its value

[Answer](#)

- (a) Is negative (b) Is positive (c) Lies between 0 and ∞
(d) Is zero (e) Is 1.

(1 mark)

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

[Answer](#)

- (a) The expected value of the sum of two random variables is equal to the sum of the expected values of the two random variables
(b) The variance of the sum of two random variables is equal to the sum of the variances of the two random variables if the random variables are independent
(c) The expected value of the sum of two random variables is equal to zero if the expected value of each of the random variables is zero
(d) The variance of the sum of two random variables is equal to zero if the random variables are independent
(e) The covariance between two dependent random variables is not zero.

(1 mark)

16. The probability that a continuous random variable will assume a value within a specific interval is provided by

[Answer](#)

- (a) The height of the graph of the probability distribution at any point within the interval
(b) The height of the graph of the probability distribution at any point within the interval plus a constant
(c) The height of the graph of the probability distribution at any point within the interval minus a constant
(d) The area under the graph of the probability distribution corresponding to the specified interval
(e) The area under the graph of the probability distribution corresponding to the specified interval subtracted from 1.

(1 mark)

17. Which criterion of decision-making is **not** applicable when the decision maker has insufficient information to assign any probabilities of occurrence to the various states of nature?

[Answer](#)

- (a) Maximax criterion (b) Maximin criterion (c) Hurwicz criterion
(d) Regret criterion (e) Expected value criterion.

(1 mark)

18. A discrete uniform random variable **N**, assumes all the integers starting from 1 to 25. The characteristic of the discrete uniform random variable is that all possible values of the variable have equal probability.

[Answer](#)

What is the variance of the random variable **N**?

- (a) 12.5 (b) 25 (c) 52 (d) 624 (e) 625.

(1 mark)

19. A medical shop sells a variety of drugs and one of the drugs sold by the shop is **Drug X**. This drug has a short shelf life and it is quite expensive. Each packet of **Drug X** costs Rs.300 and can be sold for Rs.400, The manager of the shop has observed that the demand for this drug before its expiry has the following probability distribution:

[Answer](#)

Demand level before expiry (number of packets)	300	400	500	600	700
Probability	0.15	0.25	0.40	0.15	0.05

Any packet of this drug not sold before expiry has no value and has to be rejected. It is assumed that the demand for this drug will be one of the aforementioned values only.

How many packets of **Drug X** should be stocked by the shop?

- (a) 300 (b) 400 (c) 500 (d) 600 (e) 700.

(1 mark)

20. A businessman has to decide on the next month's stock level of a particular product that he deals in. He has observed that the monthly demand for the product has the following probability distribution:

Monthly demand (number of units)	4,000	5,000	6,000	7,000	8,000
Probability	0.10	0.25	0.40	0.20	0.05

The following conditional profits (in Rs.) have been assessed by the businessman for various stock and demand levels:

Conditional Profits (in Rs.)

Monthly stock level (in units)	Monthly demand level (in units)				
	4,000	5,000	6,000	7,000	8,000
4,000	4,00,000	4,00,000	4,00,000	4,00,000	4,00,000
5,000	1,00,000	5,00,000	5,00,000	5,00,000	5,00,000
6,000	-2,00,000	2,00,000	6,00,000	6,00,000	6,00,000
7,000	-5,00,000	-1,00,000	3,00,000	7,00,000	7,00,000
8,000	-8,00,000	-4,00,000	0	4,00,000	8,00,000

What is the expected value of perfect information for the decision on the appropriate stock level?

- (a) Rs. 4,60,000 (b) Rs. 5,85,000 (c) Rs. 1,25,000 (d) Rs. 4,00,000
(e) Rs. 8,00,000.

(1 mark)

21. A bicyclist's one-way commute time is uniformly distributed between 40 and 50 minutes. What is the probability that the commute time takes a value between 44 and 46 minutes?

- (a) 0.2000 (b) 0.2736 (c) 0.7264 (d) 0.8000 (e) 0.7264.

(1 mark)

22. When decisions are made under uncertain conditions, which of the following criteria for decision making is based on the opportunity costs of making particular decisions?

- (a) Maximin criterion (b) Maximax criterion (c) Hurwicz criterion
(d) Regret criterion (e) Expected value.

(1 mark)

23. The variance of a random variable, P , is 9, and its expected value is 4. Let $Q = 6P$.

What is the standard deviation of the random variable Q ?

- (a) 3 (b) 9 (c) 81 (d) 18 (e) 324.

(1 mark)

24. If we reject $H_0: \mu = 10$ in favor of $H_1: \mu \neq 10$ at a given level of significance with a positive value of the test statistic, then a test with $H_0: \mu = 10$ versus $H_1: \mu > 10$ using the same sample and the same level of significance will

- (a) Always accept H_0
(b) Always reject H_0
(c) Sometimes accept and sometimes reject H_0
(d) Certainly lead to a Type I error
(e) Certainly lead to a Type II error.

(1 mark)

25. The DEF Electronics Company claims that the batteries it produces have an average useful life of 100 hours. It is known that the population standard deviation is 20 hours. A test is undertaken on a random sample of 49 batteries to check the validity of this claim. Assuming that the company's claim is valid, what is the

probability of observing a sample mean of 96 hours or less?

- (a) 0.0420 (b) 0.0808 (c) 0.4192 (d) 0.96 (e) 0.92.

(1 mark)

26. The following details are available with regard to a hypothesis test on a population mean:

[Answer](#)

$$H_0 : \mu = 20$$

$$H_1 : \mu \neq 20$$

$$\sigma^2 = 81$$

$$n = 36$$

The null hypothesis is rejected if $\bar{x} \leq 17.06$ or $\bar{x} \geq 22.94$. What is the probability of committing a type I error?

- (a) 0.5 (b) 0.05 (c) 0.025 (d) 0.005 (e) 1.00

(2 marks)

27. The variance of the sampling distribution of mean is

[Answer](#)

- (a) Always equal to the variance of the underlying population
(b) Less than the variance of the underlying population if the sample size is more than one
(c) More than the variance of the underlying population if the sample size is more than one
(d) Equal to the standard deviation of the underlying population if the sample size is more than one
(e) Not dependent on the sample size.

(1 mark)

28. The arithmetic mean of the upper and lower limits of the confidence interval for population proportion is equal to the

[Answer](#)

- (a) Population proportion (b) Sample proportion (c) Population variance
(d) Sample variance (e) Standard error of proportion.

(1 mark)

29. A sample has been taken from a population and the sample mean has been found to be 100. The upper limit of a 90 percent confidence interval is 112. What is the lower limit of this confidence interval?

[Answer](#)

- (a) 100 (b) 92 (c) 68.26 (d) 88 (e) 144.76.

(1 mark)

30. A small sample has been taken from a population and the sample mean has been found to be 55. The lower limit of a 95 percent confidence interval for population mean is 35.40. The population variance is known to be 1,600. What is the sample size?

[Answer](#)

- (a) 16 (b) 32 (c) 64 (d) 25 (e) 10.

(1 mark)

31. A sample of size 96 has been taken from a population and the estimated standard error of proportion is found to be 0.05. What is the sample proportion?

[Answer](#)

- (a) 0.05 or 0.95 (b) 0.24 or 0.76 (c) 0.80 or 0.20 (d) 0.40 or 0.60 (e) 0.50.

(1 mark)

32. The following information is available with regard to a sampling distribution of mean:

[Answer](#)

Population mean = 60

Sample size = 64

Probability that the sample mean is less than 55 = 0.0228

It is assumed that the Central Limit Theorem will be applicable.

What is the population variance?

- (a) 6.25 (b) 64 (c) 20 (d) 400 (e) 3600.

(2 marks)

33. A polling agency found that 387 people who planned to vote in the upcoming election said that they would vote for party A, 360 people planned to vote for party B, and 253 are undecided. What is the 90% confidence interval to estimate the proportion of voters who plan to vote for party A in the next election? Assume that the voter population is large vis-à-vis the sample. [Answer](#)

- (a) (0.341, 0.432) (b) (0.376, 0.398) (c) (0.357, 0.417) (d) (0.372, 0.402)
(e) (0.362, 0.412).

(1 mark)

34. The mean and the variance of a simple random sample are found to be 11 and 9.6 respectively. The population variance is known to be 7.84. The sample size is 6 and the population size is 40. What is the 95 percent confidence interval for the population mean, on the basis of the given information? It is assumed that the population is normally distributed. [Answer](#)

- (a) (8.257, 13.743) (b) (8.91, 13.09) (c) (8.76, 13.24) (d) (8.06, 13.94)
(e) (8.24, 13.76).

(1 mark)

35. Other things remaining the same, if the sample mean increases then the width of the confidence interval for the population mean will [Answer](#)

- (a) Increase (b) Decrease (c) Not change (d) Change unpredictably
(e) Increase by twice the increase in sample mean.

(1 mark)

36. The confidence level being same and the population variance being known, a confidence interval for the population mean which is wider than another for the same population, is made from a sample [Answer](#)

- (a) Of larger size (b) Of smaller size (c) Of equal size (d) With equal mean
(e) With higher mean

(1 mark)

37. A 95% confidence interval for the proportion of male drunk drivers in a given state is to be constructed and must be accurate to within plus and minus 0.08 of the point estimate. A preliminary sample provides an initial estimate of the proportion as 0.09. The smallest sample size that will provide the desired accuracy is [Answer](#)

- (a) 77 (b) 26 (c) 150 (d) 50 (e) 40.

(1 mark)

38. The details for a 98% confidence interval for a population proportion are given below: [Answer](#)

Sample size = 200

Lower confidence limit = 0.393

Upper confidence limit = 0.557

If you test at the 2% significance level whether the population proportion is less than 0.60, then what will be the most logical conclusion?

- (a) The population proportion is equal to 0.60
(b) The population proportion is less than 0.60
(c) The population proportion is more than 0.60

- (d) The confidence interval is incorrect
- (e) The sample proportion is incorrect.

(2 marks)

39. The details for a 97% confidence interval for a population mean are given below:

[Answer](#)

Population standard deviation = 4

Sample size = 36

Lower confidence limit = 18.553

Upper confidence limit = 21.447

If you test at the 3% significance level whether the population mean is greater than 22, then what will be the most logical conclusion?

- (a) The population mean is equal to 22
- (b) The population mean is less than 22
- (c) The population mean is more than 22
- (d) The confidence interval is incorrect
- (e) The sample mean is incorrect.

(2 marks)

40. In testing the equality of two population means by using independent random samples, if the null hypothesis of equal population means is rejected at a significance level of 0.025, then it

[Answer](#)

- (a) Will certainly be accepted at a significance level of 0.05
- (b) Will certainly be rejected at a significance level of 0.05
- (c) May or may not be rejected at a significance level of 0.05
- (d) Will certainly be accepted at a significance level of 0.01
- (e) Will certainly be rejected at a significance level of 0.01.

(1 mark)

41. The following details are available with regard to a hypothesis test on means of two populations:

[Answer](#)

$$H_0: \mu_1 = \mu_2 + 1$$

$$H_1: \mu_1 \neq \mu_2 + 1$$

$$n_1 = 64 \quad \sum x_1^2 = 1852 \quad \sum x_1 = 320$$

$$n_2 = 36 \quad \sum x_2^2 = 891 \quad \sum x_2 = 144$$

The samples collected from the two populations are independent.

What is the value of the test statistic?

- (a) 0.707
- (b) -1.789
- (c) 0
- (d) -0.354
- (e) 1.789

(2 marks)

The following details are available with regard to a hypothesis test on difference between means of two populations:

[Answer](#)

$$n_1 = 16 \quad s_1 = 16$$

$$n_2 = 9 \quad s_2 = 12$$

The samples are independently collected and it is assumed that the two populations have the same variance.

What is the estimated standard error of difference between means?

- (a) 37.68
- (b) 6.138
- (c) 13.565
- (d) 3.683
- (e) 4.911.

(1 mark)

43. The following details are available with regard to a hypothesis test on difference between means of two populations:

[Answer](#)

$$H_0: \mu_1 = \mu_2 + 1$$

$$H_1: \mu_1 \neq \mu_2 + 1$$

$$n_1 = 16$$

$$n_2 = 9$$

$$\bar{x}_1 + \bar{x}_2 = 38$$

The samples are independently collected and it is assumed that the two populations have the same variance. The pooled estimate of the underlying common population variance is 217.032. The value of the test statistic is 0.1629. Which of the following is correct with regard to $\bar{x}_1$ and $\bar{x}_2$? Round off your answer to the nearest integer.

- (a) $\bar{x}_1 = 18$ and $\bar{x}_2 = 20$ (b) $\bar{x}_1 = 20$ and $\bar{x}_2 = 18$
 (c) $\bar{x}_1 = 18$ and $\bar{x}_2 = 18$ (d) $\bar{x}_1 = 20$ and $\bar{x}_2 = 20$ (e) $\bar{x}_1 = 38$ and $\bar{x}_2 = 2$

(2 marks)

44. The following details are available with regard to a hypothesis test on means of two populations:

[Answer](#)

$$H_0: \mu_1 = \mu_2 + 1$$

$$H_1: \mu_1 > \mu_2 + 1$$

$$n_1 = 64 \quad \sigma_1^2 = 256 \quad \sum x_1 = 640$$

$$n_2 = 36 \quad \sigma_2^2 = 144 \quad \sum x_2 = 144$$

The samples collected from the two populations are independent. If the significance level is 0.05 then

- (a) The critical value is 1.645 and the null hypothesis is rejected
 (b) The critical value is 1.96 and the null hypothesis is accepted
 (c) The critical value is 1.645 and the null hypothesis is accepted
 (d) The critical values are ± 1.96 and the null hypothesis is rejected
 (e) The critical values are ± 1.645 and the null hypothesis is accepted.

(1 mark)

45. A magazine claims that 25% of its readers are college students. Of a random sample of 200 readers, 42 are college students. It is to be tested at a 0.10 level of significance whether the proportion of college students among all the readers of the magazine is not equal to 0.25. What is the conclusion?

[Answer](#)

- (a) The proportion of college students among the readers of the magazine is 0.25
 (b) The sample data are incorrect
 (c) The proportion of college students among the readers of the magazine is less than 0.25
 (d) The proportion of college students among the readers of the magazine is more than 0.25
 (e) No conclusion can be drawn on the basis of the given information.

(1 mark)

46. A survey firm conducts door-to-door surveys on a variety of issues. Some individuals cooperate with the interviewer and complete the interview questionnaire and others do not. The following data are available.

[Answer](#)

Respondents	Sample Size	Number Not Cooperating
Men	200	110
Women	300	210

It is to be tested at a significance level of 0.05, whether the proportions of men and women who cooperate with the interviewer are equal. Which of the following can be inferred from the test?

- (a) Proportions of men and women who cooperate with the interviewer are equal
 (b) Proportions of men and women who cooperate with the interviewer are not equal
 (c) Proportion of men who cooperate with the interviewer is less than the proportions of women who cooperate
 (d) All men cooperate with the interviewer
 (e) All women cooperate with the interviewer.

(2 marks)

47. The following details are available with regard to a hypothesis test on population mean:

[Answer](#)

$$H_0: \mu = 9$$

$$H_1: \mu \neq 9$$

$$n = 25$$

$$s^2 = 256$$

$$\bar{X} = 15.75$$

$$\text{Significance level} = 0.05$$

The population is normally distributed. It is later known that the true population mean is 9.

Which of the following can be said with regard to the test?

- (a) There is insufficient information for doing the test
- (b) The normal distribution should be used
- (c) The test does not lead to either type I or type II error
- (d) The test leads to a type I error
- (e) The test leads to a type II error.

(1 mark)

48. The following linear trend equations for estimation of national income and population have been constructed for a hypothetical economy for the period 1999 to 2003:

[Answer](#)

$$\text{National income, } \hat{Y} = 4.86 + 1.5x$$

$$\text{Population, } \hat{Z} = 5.25 + 1.25x$$

x is the coded value of the time variable (year).

Assuming that the year 2001 is coded as 0 and time interval of 1 year is considered for developing the

equations, in which year is the estimated per capita income highest? $\left(\text{Per capita income} = \frac{\text{National income}}{\text{Population}} \right)$

- (a) 1999 (b) 2000 (c) 2001 (d) 2002 (e) 2003.

(1 mark)

49. Cosmic Systems Ltd. produces different types of products in the field of information technology. The research and development department of the company has been witnessing a growth in the research activities over the past few years. The data on research expenditure in the recent past are given below:

[Answer](#)

Year	1998	1999	2000	2001	2002	2003
Research expenditure (Rs. in lakhs)	50	57	61	65	72	79

The estimate of research expenditure for the year 2004 is

- (a) Rs.74.25 lakhs (b) Rs.83.39 lakhs (c) Rs.75.25 lakhs (d) Rs.85.25 lakhs
(e) Rs.72.15 lakhs

(2 marks)

50. In a time series analysis, if relative cyclical residual measure is -0.346 and actual value of the dependent variable is 4.516 then the forecasted value of the dependent variable is

[Answer](#)

- (a) 3.4517 (b) 4.1547 (c) 4.3571 (d) 4.5317 (e) 5.1537.

(1 mark)

51. A financial analyst has collected the following data on sales and net profits of some companies manufacturing TV: [Answer](#)

Sales (Rs. in crores)	42	45	52	34	60
Net profit (Rs. in crores)	2.5	2.45	3.25	1.75	4.00

What is the percentage of variations in net profits that is explained by the variations in sales? (Round off your answer to the nearest integer)

- (a) 98% (b) 95% (c) 92% (d) 85% (e) 81%

(2 marks)

52. The **Y** intercept in the simple linear regression equation (**X** is the independent variable and **Y** is the dependent variable) represents the [Answer](#)

- (a) True value of **Y** when **X** = 0 (b) Change in average value of **Y** per unit change in **X**
(c) Expected value of **Y** when **X** = 0 (d) Standard deviation of the values of **X**
(e) Mean of the values of **X**.

(1 mark)

53. A simple regression relationship was developed between two variables **X** and **Y**, with **X** as the independent variable. [Answer](#)

$$\Sigma Y = 792 \quad \Sigma Y^2 = 79144 \quad \Sigma (Y - \hat{Y})^2 = 200 \quad n = 8$$

The coefficient of determination is

- (a) 0.2717 (b) 0.7283 (c) 0.3731 (d) 0.2137 (e) 0.7863.

(1 mark)

54. The following regression relationship between two variables, **X** and **Y**, has been obtained: [Answer](#)

$$\hat{Y} = 1,236 - 104X$$

where **X** is the independent variable and **Y** is the dependent variable.

The following details are also available:

$$\Sigma Y^2 = 19,00,400$$

$$\bar{Y} = 612$$

$$\Sigma XY = 18,100$$

Number of observations = 5

What is the approximate 90 percent prediction interval for **Y** if **X** = 7.50?

- (a) $1,236 \pm 456$ (b) 456 ± 34.38 (c) $1,236 \pm 104$ (d) 456 ± 104 (e) $1,236 \pm 34.38$

(1 mark)

55. In the linear regression model, the residuals are assumed to [Answer](#)

- (a) Have a binomial distribution with a probability of success of 0.50
(b) Have a normal distribution with a mean of 0
(c) Have a chi-square distribution with a mean of 0
(d) Have a **t** distribution with a mean of 1.00
(e) Have a normal distribution with unknown mean.

(1 mark)

56. For a simple regression equation the following results were obtained:

[Answer](#)

If $X = 7$, then $\hat{Y} = 39$

If $X = 12$, then $\hat{Y} = 64$

where X is the independent variable and $\hat{Y}$ is the estimated value of the dependent variable Y .

$\sum X = 96$

Number of observations = 8

What is the mean of the observed values of the dependent variable, Y ?

- (a) 25 (b) 35 (c) 64 (d) 32 (e) 39

(1 mark)

57. Which of the following is **true** when the slope of a regression line is positive?

[Answer](#)

- (a) The correlation coefficient between the dependent and independent variables is zero
(b) There is a positive correlation between the dependent and independent variables
(c) There is a negative correlation between the dependent and independent variables
(d) The regression line is perpendicular to the horizontal axis
(e) The regression line is parallel to the horizontal axis.

(1 mark)

58. The following table shows the respective heights, X (independent variable) and Y (dependent variable), of a sample of 12 fathers and their eldest sons.

Height (X) of father (inches)	65	63	67	64	68	62	70	66	68	67	69	71
Height (Y) of eldest sons (inches)	68	66	68	65	69	66	68	65	71	67	68	70

The least-squares regression equation of Y on X is

- (a) $\hat{Y} = 3.38 + 0.476X$ (b) $\hat{Y} = 35.85 + 0.476X$ (c) $\hat{Y} = -3.38 + 1.036X$
(d) $\hat{Y} = 35.85 + 1.036X$ (e) $\hat{Y} = 35.85 + 2X$.

(2 marks)

59. Two variables, X and Y , are related by a regression equation such that X is the independent variable and Y is the dependent variable. The following details are available:

Coefficient of correlation between X and Y = 0.75

Standard deviation of X = 6

Slope of regression equation = 2.5

The standard deviation of variable Y is

- (a) 20 (b) 12 (c) 6.4 (d) 32 (e) 64.

(1 mark)

60. Variable Y increases as variable X increases and vice versa. Variable X cannot explain 36 percent of the variations in Y . What is the coefficient of correlation between X and Y ?

- (a) 0.80 (b) 0.20 (c) 0.64 (d) -0.64 (e) -0.80.

(1 mark)

61. The standard error of estimate

[Answer](#)

- (a) Is a measure of central tendency for the distribution of the values estimated by the regression equation
- (b) Measures the variability of the observed values of the dependent variable around their mean
- (c) Measures the variability of the observed values of the dependent variable around the regression line
- (d) Measures the variability of the observed values of the independent variable around their mean
- (e) Measures the variability of the observed values of the independent variable around the regression line.

(1 mark)

62. The following details are available with regard to a basket of goods:

[Answer](#)

Ratio of weighted average price of the goods in current year to the weighted average price of the goods in base year, where both the weighted averages have been calculated with base year consumption quantities as weights = 6:5

Ratio of weighted average price of the goods in current year to the weighted average price of the goods in base year, where both the weighted averages have been calculated with current year consumption quantities as weights = 5:4

What is Fisher's ideal price index?

- (a) 109.54
- (b) 122.47
- (c) 111.80
- (d) 120
- (e) 125.

(1 mark)

63. If two variables **X** and **Y** are perfectly positively correlated then their covariance will be equal to

[Answer](#)

- (a) 0
- (b) 1
- (c) The product of the standard deviations of **X** and **Y**
- (d) The sum of the standard deviations of **X** and **Y**
- (e) The difference between the standard deviations of **X** and **Y**.

(1 mark)

64. Which of the following is **true** with regard to a given coefficient of correlation and its corresponding coefficient of determination?

[Answer](#)

- (a) The coefficient of determination is always greater than or equal to zero, and less than or equal to 1
- (b) The coefficient of determination is always less than zero
- (c) The coefficient of determination is always equal to 1
- (d) The coefficient of determination always has the same sign as the coefficient of correlation
- (e) The magnitude of coefficient of determination is always higher than the magnitude of coefficient of correlation.

(1 mark)

65. A simple regression relationship was developed between the variables **X** and **Y**, with **X** as the independent variable. The estimated value of **Y** when **X** = 2, is 7.0; and the estimated value of **Y** when **X** = 8, is 11.8.

[Answer](#)

$$\Sigma Y^2 = 791.31 \quad \Sigma XY = 452.7 \quad \Sigma Y = 78.1 \quad n = 8$$

The standard error of estimate for the regression line is

- (a) 1.1113
- (b) 8.4572
- (c) 7.8482
- (d) 7.5123
- (e) 6.5893.

(1 mark)

66. Two variables, **X** and **Y**, are related by a regression equation such that **X** is the independent variable and **Y** is the dependent variable. The following details are available:

[Answer](#)

Coefficient of correlation between **X** and **Y** = 0.80

Standard deviation of **X** = 5

Standard deviation of **Y** = 8

The slope of the regression equation is

- (a) 0.80 (b) 1.28 (c) 6.4 (d) 32 (e) 64.

(1 mark)

67. In a simple regression relationship, $\hat{Y} = A + BX$, if the null hypothesis $H_0: B = 0$ is rejected then which of the following can be inferred? [Answer](#)

- (a) There is a significant relationship between **X** and **Y**
(b) There is not a significant relationship between **X** and **Y**
(c) Nothing can be inferred
(d) **X** is the cause and **Y** is the effect
(e) There is no correlation between the variables **X** and **Y**.

(1 mark)

68. Which of the following is **not true** with regard to the regression relationship, $\hat{Y} = a + bX$? [Answer](#)

- (a) The point $(\bar{X}, \bar{Y})$ always lies on the regression line
(b) The expected value of the residuals is zero
(c) The mean of the fitted values of **Y** is the same as the mean of the observed values of **Y**
(d) There are always as many points above the fitted line as there are below it
(e) The regression line minimizes the sum of the squared residuals

(1 mark)

69. A fitted regression equation is given by $\hat{Y} = 20 + 0.75X$. What is the value of the residual at the point $X = 100, Y = 90$? [Answer](#)

- (a) 5 (b) 10 (c) 0 (d) 15 (e) -5

(1 mark)

70. Larger values of coefficient of determination in simple regression analysis imply that the observations are more closely grouped about the [Answer](#)

- (a) Average value of the independent variable
(b) Average value of the dependent variable
(c) Least squares regression line
(d) Origin
(e) Highest value of the dependent variable.

(1 mark)

71. Two variables, **A** and **B**, are related by a simple regression equation such that **A** is the independent variable and **B** is the dependent variable. The following details are available: [Answer](#)

Slope of the regression equation = 0.64
Standard deviation of **A** = 5
Standard deviation of **B** = 4

The coefficient of correlation between **A** and **B** is

- (a) 0.64 (b) 0.80 (c) 0.40 (d) 0.32 (e) 0.56.

(1 mark)

72. The following data provides the values of sample range, **R** for samples of size 5 each. [Answer](#)

Sample No.	1	2	3	4	5	6	7	8	9	10
------------	---	---	---	---	---	---	---	---	---	----

Range	7	4	8	5	7	4	8	4	7	9
-------	---	---	---	---	---	---	---	---	---	---

Calculate the value of lower control limit for the **R** chart. (Use the relevant **D₃** or **D₄** factor, whichever is appropriate)

- (a) Zero (b) 0.0101 (c) 0.1024 (d) 1.0126 (e) 1.1020.

(1 mark)

73. Two random samples are drawn from two normal populations, 1 and 2, and their data are:

Sample from population 1	60	65	71	74	76	82	85	87		
Sample from population 2	61	66	67	85	78	63	85	86	88	91

We want to test whether the population 2 has larger variance than the population 1. What is the value of the appropriate test statistic for the above test?

- (a) 1.346 (b) 1.468 (c) 1.582 (d) 1.624 (e) 1.718.

(2 marks)

74. To assess the significance of possible variation in performance in a certain test between the grammar schools of a city, a common test was given to a number of students taken at random from the senior fifth class of each of the four schools concerned. The results are given below.

Schools			
A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15

It is to be tested whether the mean marks obtained is same for the four schools. A test of ANOVA has to be performed.

What is the variance between the sample means?

- (a) 14.5 (b) 15.6 (c) 16.7 (d) 17.8 (e) 18.9.

(1 mark)

75. In response to a number of complaints about slow mail delivery, the Postmaster General initiated a preliminary investigation. An investigation follows nine letters from Mumbai to Pune, to estimate the variance of delivery time. The staff collects a sample of 9 mail deliveries and finds the sample standard deviation to be 23 hours. The Postmaster General wants a 95% confidence interval for the variance in the mail delivery.

Find the lower confidence limit for the estimated population variance.

- (a) 231.25 (b) 241.35 (c) 251.45 (d) 261.55 (e) 271.65.

(1 mark)

76. A beverage company is interested in the preference of the gender on the types of beverages consumed and has collected the following data from a random sample of 150 persons:

Beverages	Male	Female	Total
A	25	30	55

B	40	15	55
C	30	10	40
Total	95	55	150

The managers of the company want to test whether there is any association between the types of beverage consumed and the gender of the consumer. What is the value of the appropriate test statistic for the above test?

- (a) 9.754 (b) 10.875 (c) 11.996 (d) 12.017 (e) 13.138.

(1 mark)

77. A production supervisor is interested in knowing if the number of breakdowns on four machines is independent of the shift using the machines. He collects the following data: [Answer](#)

Shift	Machines				
	A	B	C	D	Total
Morning	15	10	18	12	55
Evening	12	8	15	10	45
Total	27	18	33	22	100

He wants to conduct a test of independence between the number of breakdowns and the shift. What is the critical value for the appropriate statistical test at a 5% significance level?

- (a) 3.841 (b) 5.991 (c) 7.815 (d) 9.488 (e) 11.070.

(1 mark)

78. The movement of the dependent variable above and below the secular trend line over periods longer than one year is known as [Answer](#)

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

(1 mark)

79. A graphical method of representing states of nature and courses of action may be referred to as a [Answer](#)

- (a) Decision tree (b) Histogram (c) Scatter diagram
(d) Frequency distribution (e) Probability distribution.

(1 mark)

80. The sample mean is an unbiased point estimator of [Answer](#)

- (a) The population variance (b) The population mean
(c) The population proportion (d) The standard error of mean
(e) The standard error of proportion.

(1 mark)

81. Which type of sampling is appropriate when the population consists of well-defined groups such that the elements within each group are heterogeneous and each group is a representative of the population as a whole? [Answer](#)

- (a) Simple random sampling (b) Cluster sampling (c) Stratified sampling
(d) Systematic sampling (e) Judgmental sampling.

(1 mark)

82. The property of efficiency with regard to an estimator refers to the

[Answer](#)
<
>

- (a) Sample size
- (b) Size of the standard error of the sample statistic
- (c) Size of the standard deviation of the sample
- (d) Size of the sample variance
- (e) Size of the sample mean.

(1 mark)

83. Which of the following is the graphical plot of the values of the dependent and independent variables, in the context of simple regression analysis?

[Answer](#)
<
>

- (a) Scatter diagram
- (b) Frequency polygon
- (c) Histogram
- (d) $\bar{x}$ chart
- (e) Ogive.

(1 mark)

84. Which of the following price indices uses only current year quantities as weights?

[Answer](#)
<
>

- (a) Fisher's ideal price index
- (b) Laspeyre's price index
- (c) Paasche's price index
- (d) Marshall-Edgeworth price index
- (e) Fixed weight aggregates price index.

(1 mark)

85. Which of the following patterns of $\bar{x}$ chart indicates that the variations have been reduced significantly?

[Answer](#)
<
>

- (a) Hugging the control limits
- (b) Hugging the center line
- (c) Cycles
- (d) Decreasing trend
- (e) Increasing trend.

(1 mark)

Suggested Answers

Quantitative Methods-II (132): January 2005

1. Answer :

Reason : Let X be the number obtained on any throw. The desired event is $X = 3$ or 4 or 5 or 6 i.e, $X \geq 3$.

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

This can be considered as success. $\therefore P(\text{Success}) = P(X \geq 3) = \frac{4}{6} = \frac{2}{3}$

Let Y denote the number of throws in which success ($X \geq 3$) happens. Y follows a binomial

distribution with $p = \frac{2}{3}$ and number of trials = 20

$$\begin{aligned} P(Y \leq 17) &= 1 - [P(Y = 18) + P(Y = 19) + P(Y = 20)] \\ &= 1 - [{}^{20}C_{18}p^{18}(1-p)^2 + {}^{20}C_{19}p^{19}(1-p) + {}^{20}C_{20}p^{20}] \\ &= 1 - [190 \times \left(\frac{2}{3}\right)^{18} \left(\frac{1}{3}\right)^2 + 20 \times \left(\frac{2}{3}\right)^{19} \left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)^{20}] \\ &= 1 - [0.01428 + 0.00301 + 0.0003] \\ &= 0.98241 \end{aligned}$$

2. Answer :

Reason : The probability that a continuous random variable will assume a particular value is zero.

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

3. Answer :

Reason : The event $40 \leq Y \leq 80$ occurs when

$Y = 40$ or 60 or 80

$$\therefore P(40 \leq Y \leq 80) = 0.20 + 0.30 + 0.25 = 0.75$$

Let Z denote the number of observations of Y which satisfy the condition $40 \leq Y \leq 80$.

$\therefore$ Z follows a binomial distribution with number of trials = 15 and probability of success

= 0.75.

$$\begin{aligned} \therefore P(Z < 12) &= P(Z \leq 11) = 1 - P(Z \geq 12) \\ &= 1 - [P(Z=12) + P(Z=13) + P(Z=14) + P(Z=15)] \end{aligned}$$

$$\begin{aligned} &= 1 - [{}^{15}C_{12} (0.75)^{12} (1-0.75)^{15-12} + {}^{15}C_{13} (0.75)^{13} (1-0.75)^{15-13} + {}^{15}C_{14} (0.75)^{14} (1-0.75)^{15-14} + {}^{15}C_{15} (0.75)^{15} (1-0.75)^{15-15}] \end{aligned}$$

$$= 1 - \left[455(0.75)^{12} (0.25)^3 + 105(0.75)^{13} (0.25)^2 + 15(0.75)^{14} (0.25) + 1(0.75)^{15} (0.25)^0 \right]$$

$$= 1 - [0.2252 + 0.1559 + 0.0668 + 0.0134]$$

$$= 1 - 0.4613$$

$$= 0.5387.$$

4. Answer :

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

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

For 25 trials the standard deviation of Z, $\sigma = \sqrt{n.p.q} = \sqrt{6}$

$\therefore$ For 25 trials $V(Z) = \sigma^2 = n.p.q = 6$

or $25 \times p \times q = 6$

or $p \times q = \frac{6}{25} = 0.24$

$\therefore$ For 75 trials $V(Z) = n.p.q = 75 \times p \times q$

$= 75 \times 0.24$

$= 18$

$\therefore$ For 75 trials standard deviation of Z $= \sqrt{Z} = \sqrt{18}$

$= 3\sqrt{2}$

Reason :

5. Answer :

Let success be defined as obtaining at least three from both the dice simultaneously.

On any die $P(\text{At least 3}) = P(3 \text{ or } 4 \text{ or } 5 \text{ or } 6) = \frac{4}{6} = \frac{2}{3}$

$\therefore P(\text{Success}) = P(\text{At least 3 on both the dice}) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$

Let X denote the number of successes.

$\therefore$ X follows a binomial distribution with $n = 10$ and $p = \frac{4}{9}$

$P(X \geq 9) = P(X = 9) + P(X = 10)$

$$= {}^{10}C_9 \left(\frac{4}{9}\right)^9 \left(1 - \frac{4}{9}\right)^{10-9} + {}^{10}C_{10} \left(\frac{4}{9}\right)^{10} \left(1 - \frac{4}{9}\right)^{10-10}$$

$$= 0.00376 + 0.00030$$

Reason : $= 0.00406$

6. Answer :

Reason : The covariance of any random variable with itself is equal to the variance of the variable.

7. Answer :

Reason : In a hypergeometric distribution the composition of the population changes from trial to trial.

8. Answer :

Reason : A continuous uniform probability distribution is described by the upper and lower limits of the interval.

9. Answer :

Reason : $V(Z) = V(5X - 10Y) = 5^2 V(X) + 10^2 V(Y) - 2 \times 5 \times 10 \times \text{Cov}(X, Y)$

Or $70^2 = 25 \times 6^2 + 100 \times 8^2 - 100 \text{Cov}(X, Y)$

Or $4900 = 900 + 6400 - 100 \text{Cov}(X, Y)$

Or $100 \text{Cov}(X, Y) = 900 + 6400 - 4900$

Or $100 \text{Cov}(X, Y) = 2400$

Or $\text{Cov}(X, Y) = \frac{2400}{100} = 24$

10. Answer :

Reason : The number of absentees who are girls follows a hypergeometric

[TOP](#)

[TOP](#)

[TOP](#)

[TOP](#)

[TOP](#)

[TOP](#)

distribution. The following details are available:

N = No. of elements in the population = 25

r = No. of elements in the population labelled success = No. of girls = 11

n = No. of trials = No. of absent students = 5

x = Desired number of successes

$$\frac{{}^r C_x \cdot {}^{(N-r)} C_{(n-x)}}{{}^N C_n}$$

In a hypergeometric distribution $P(x) =$

$$P(x \leq 2) = P(x = 1) + P(x = 2) = \frac{{}^{11} C_1 \cdot {}^{(25-11)} C_{(5-1)}}{{}^{25} C_5} + \frac{{}^{11} C_2 \cdot {}^{(25-11)} C_{(5-2)}}{{}^{25} C_5}$$

$$= \frac{11 \times {}^{14} C_4}{{}^{25} C_5} + \frac{55 \times {}^{14} C_3}{{}^{25} C_5}$$

$$= 0.2072 + 0.3768$$

$$= 0.584$$

11. Answer :

Reason : Our desired event is $Z < 80$. $P(Z < 80) = 0.15 + 0.25 = 0.40$

Let X be the number of times Z assumes values less than 80. Then X is binomially distributed with

number of trials = 20 and probability of success = 0.40

The expected number of observations less than 80 = $E(Z) = n.p = 20 \times 0.40 = 8$

[TOP](#)

12. Answer :

Reason : $Z = (x - \mu)/\sigma$

When each value is reduced by a constant the mean of the resulting values = previous mean - constant. When the mean itself is subtracted from each value the mean of the resulting values = mean - mean = 0. When each value is reduced by a constant the standard deviation remains unchanged. However when each value is divided by a constant the standard deviation of the resulting values = previous standard deviation/constant. When each value is divided by the standard deviation the standard deviation of the resulting values = standard deviation/standard deviation = 1

[TOP](#)

13. Answer :

Reason : A **required** condition for probability distribution of a discrete random variable is: Sum of all the probabilities = 1.

[TOP](#)

14. Answer :

Reason : If a z value is to the left of the mean in the standard normal distribution, then its value is negative.

[TOP](#)

15. Answer :

Reason : The covariance of the sum of two random variables is equal to zero if the random variables are independent. In such a case the variance of the sum will equal to the sum of the individual variances and hence not equal to zero.

[TOP](#)

16. Answer :

Reason : The area under the graph of the probability distribution corresponding to the specified interval.

[TOP](#)

17. Answer :

Reason : When the decision maker has insufficient information to assign any

[TOP](#)

probabilities of occurrence to the various states of nature expected value criterion will not be applicable.

18. Answer :

Reason : The variance of the discrete uniform random variable is obtained by the formula:

Variance = $\frac{k^2 - 1}{12}$ Where k is the last value in the range of the random variable

Hence k = 25.

$$\therefore V(N) = \frac{25^2 - 1}{12} = 52$$

19. Answer :

Reason : Conditional profit table for Drug X (in Rs.):

Stock	Demand					Expected
	300 (0.15)	400 (0.25)	500 (0.40)	600 (0.15)	700 (0.05)	Profit
300	30,000	30,000	30,000	30,000	30,000	30,000
400	0	40,000	40,000	40,000	40,000	34,000
500	(30,000)	10,000	50,000	50,000	50,000	28,000
600	(60,000)	(20,000)	20,000	60,000	60,000	6,000
700	(90,000)	(50,000)	(10,000)	30,000	70,000	(22,000)

Since the expected conditional profit is maximum at a stock level of 400 packets the shop should stock 400 packets.

20. Answer :

Reason : **Expected profit without perfect information:**

Monthly stock level (units)		Expected profit
4,000	4,00,000 (0.10 + 0.25 + 0.40 + 0.20 + 0.05)	4,00,000
5,000	1,00,000 (0.10) + 5,00,000 (0.25 + 0.40 + 0.20 + 0.05)	4,60,000
6,000	(-2,00,000 × 0.10) + (2,00,000 × 0.25) + 6,00,000 (0.40 + 0.20 + 0.05)	4,20,000
7,000	(-5,00,000 × 0.10) + (-1,00,000 × 0.25) + (3,00,000 × 0.40) + 7,00,000 (0.20 + 0.05)	2,20,000
8,000	(-8,00,000 × 0.10) + (-4,00,000 × 0.25) + (0 × 0.40) + (4,00,000 × 0.20) + (8,00,000 × 0.05)	-60,000

Highest expected profit without perfect information = Rs.4,60,000 (A)

Conditional profit with perfect information:

Monthly stock level (units)	Monthly demand level in units				
	4,000 (0.10)	5,000 (0.25)	6,000 (0.40)	7,000 (0.20)	8,000 (0.05)
4,000	4,00,000				
5,000		5,00,000			
6,000			6,00,000		
7,000				7,00,000	

8,000				8,00,000
-------	--	--	--	----------

Expected profit with perfect information

$$= (4,00,000 \times 0.10) + (5,00,000 \times 0.25) + (6,00,000 \times 0.40) + (7,00,000 \times 0.20) + (8,00,000 \times 0.05) = \text{Rs.} 5,85,000 \text{ (B)}$$

$$\text{EVPI} = B - A = \text{Rs.} 1,25,000.$$

21. Answer :

$$\text{Reason : } P(44 < X < 46) = \frac{46 - 44}{50 - 40} = 0.2$$

22. Answer :

Reason : In decision making under uncertainty the regret criterion is based on opportunity costs of decisions.

23. Answer :

$$\begin{aligned} \text{Reason : } \text{Var}[Q] &= \text{Var}[6P] \\ &= 36 \text{ Var}[P] \\ &= 36 \times 9 = 324. \end{aligned}$$

$$\therefore \text{Standard deviation of Q, } \sigma_Q = \sqrt{324} = 18.$$

24. Answer :

Reason : If we reject $H_0: \mu = 10$ in favor of $H_1: \mu \neq 10$ at a given level of significance with a positive value of the test statistic, then a test with $H_0: \mu = 10$ versus $H_1: \mu > 10$ using the same sample and the same level of significance will always reject H_0 because in a one tailed test the critical region on the relevant tail is larger than a two tailed test for the same level of significance.

25. Answer :

$$\text{Reason : } \sigma_{\bar{x}} = \frac{20}{\sqrt{49}} = 2.86$$

$$\text{test statistic: } z = \frac{96 - 100}{2.86} = -1.399$$

According to the C.L.T $\bar{x}$ is normally distributed with mean equal to the population mean and standard deviation equal to $\sigma_{\bar{x}}$.

$$\begin{aligned} \text{Probability of observing a sample mean of 96 hours or less} \\ = P(z \leq -1.399) &= 0.0808 \end{aligned}$$

26. Answer :

Reason : Probability of committing a type I error = Probability of rejecting the null hypothesis when it is true

By the central limit theorem $\bar{x}$ is normally distributed with mean $= \mu$ and

$$\text{variance} = \frac{\sigma^2}{n}$$

Hence Probability of rejecting the null hypothesis when it is true =

$$\begin{aligned}
P(\bar{x} \leq 17.06 \text{ or } \bar{x} \geq 22.94) &= P(\bar{x} \leq 17.06) + P(\bar{x} \geq 22.94) \\
&= P\left(\frac{17.06 - 20}{\sqrt{\frac{\sigma^2}{n}}} \leq z\right) + P\left(\frac{22.94 - 20}{\sqrt{\frac{\sigma^2}{n}}} \geq z\right) \\
&= P\left(\frac{-2.94}{\sqrt{\frac{81}{36}}} \leq z\right) + P\left(\frac{2.94}{\sqrt{\frac{81}{36}}} \geq z\right) \\
&= P(z \leq -1.96) + P(z \geq 1.96) \\
&= 0.025 + 0.025 = 0.05
\end{aligned}$$

27. Answer :

Reason : The variance of the sampling distribution of mean is less than the variance of the underlying population if the sample size is more than one.

28. Answer :

Reason : The arithmetic mean of the upper and lower limits of the confidence interval for population proportion is equal to the sample proportion.

$$UCL = \hat{p} + z\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$LCL = \hat{p} - z\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

29. Answer :

$$\text{Reason : } UCL = \bar{x} + z\sigma_{\bar{x}}$$

$$\text{Or } z\sigma_{\bar{x}} = UCL - \bar{x} = 112 - 100 = 12$$

$$\therefore LCL = \bar{x} - z\sigma_{\bar{x}} = 100 - 12 = 88.$$

30. Answer :

Reason : Since the population variance is known the distribution to be used is the $N(0, 1)$ distribution though the sample size is small.

$$LCL = \bar{x} - 1.96\sigma_{\bar{x}}$$

$$\text{Or } \frac{1.96\sigma_{\bar{x}}}{19.60} = \bar{x} - LCL = 55 - 35.40 =$$

$$\text{Or } \sigma_{\bar{x}} = \frac{19.60}{1.96} = 10$$

$$\text{Or } \frac{\sigma}{\sqrt{n}} = 10 \quad \text{or} \quad n = \left(\frac{\sigma}{10}\right)^2 = \frac{\sigma^2}{100} = \frac{1600}{100} = 16$$

31. Answer :

$$\text{Reason : } \hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.05$$

$$\text{Or } \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.05$$

$$\therefore \hat{p}(1-\hat{p}) = 96(0.05)^2$$

$$\text{or } \hat{p} - \hat{p}^2 = 0.24$$

$$\text{or } \hat{p}^2 - \hat{p} + 0.24 = 0$$

$$\therefore \hat{p} = \frac{-(-1) \pm \sqrt{(-1)^2 - (4 \times 1 \times 0.24)}}{2 \times 1} = \frac{1 \pm 0.20}{2} = 0.40 \text{ or } 0.60.$$

32. Answer :

Reason : By Central Limit Theorem for large samples the sample mean is approximately normally distributed with mean = population mean and

$$\text{standard deviation} = \frac{\sigma}{\sqrt{n}}$$

$$E(\bar{x}) = \mu = 60$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\text{Now, } P(\bar{x} < 55) = 0.0228$$

$$\therefore P(55 < \bar{x} < 60) = 0.50 - 0.0228 = 0.4772 = P(k < z < 0)$$

From the tables we find that $k = -2.0$

$$\therefore z = -2 = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \text{ or } -2 = \frac{55 - 60}{\sigma_{\bar{x}}} \text{ or } \sigma_{\bar{x}} = \frac{5}{-2} = -2.50$$

$$\therefore \frac{\sigma}{\sqrt{n}} = -2.50$$

$$\text{or } \frac{\sigma}{\sqrt{64}} = -2.50$$

$$\text{or } \sigma^2 = (-2.50)^2 \times 64 = 400.$$

33. Answer :

Reason : Assuming the population of the voters to be large vis-à-vis the sample.

Point estimate of population proportion of voters willing to vote for party A

$$= \frac{387}{387 + 360 + 253} = 0.387.$$

$\therefore$ 90% confidence interval for proportion in favour of party A

$$= \hat{p} \pm 1.645 \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= 0.387 \pm 1.645 \sqrt{\frac{0.387(1-0.387)}{1000}} = 0.387 \pm 0.025$$

$$= (0.362, 0.412)$$

34. Answer :

Reason : Since the population variance is known the appropriate distribution to use is the normal distribution. Sample is more than 5% of population.

$\therefore$ This is a case of finite population.

[TOP](#)

[TOP](#)

[TOP](#)

Population standard deviation, $\sigma = \sqrt{7.84}$

$$\text{Estimated standard error of mean} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{\sqrt{7.84}}{\sqrt{6}} \sqrt{\frac{40-6}{40-1}} = 1.067$$

$\therefore$ 95% confidence interval for population mean =

$$\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 11 \pm 1.96 \times 1.067 = (8.91, 13.09)$$

35. Answer :

Reason : Other things remaining the same, if the sample mean increases then the width of the confidence interval for the population mean will not change because the width of the confidence interval is determined by the size of the standard error and confidence level

[TOP](#) [^](#)
[v](#)

36. Answer :

Reason : Other things remaining the same and the population variance being known, a confidence interval for the population mean which is wider than another for the same population, is made from a sample of smaller size.

[TOP](#) [^](#)
[v](#)

37. Answer :

$$1.96\sigma_p \leq 0.08$$

$$\therefore 1.96 \sqrt{\frac{p(1-p)}{n}} \leq 0.08$$

$$\Rightarrow 1.96 \sqrt{\frac{0.09(1-0.09)}{n}} \leq 0.08$$

Squaring both sides we get

$$\frac{0.314627}{n} \leq 0.0064 \Rightarrow \frac{0.314627}{0.0064} \leq n$$

Reason : $\Rightarrow n \geq 49.16$

$\therefore$ The sample size should be at least 50 (we cannot round off 49.16 to the lower value because that will reduce the accuracy).

[TOP](#) [^](#)
[v](#)

38. Answer :

Reason : The 98% confidence interval is (0.393, 0.557).

$$\therefore \text{Sample proportion} = \frac{0.393 + 0.557}{2} = 0.475$$

$$H_0 : p = 0.60$$

$$H_1 : p < 0.60$$

Under the null hypothesis the non-standardised critical value of the sampling distribution of proportion in the lower tail, for the 2% significance level, will be =

$$0.60 - 2.054 \sqrt{\frac{0.60(1-0.60)}{200}} = 0.529$$

The sample proportion is less than the critical value (on non-standardised scale).

$\therefore$ We reject H_0 .

Hence we can conclude at 2% significance level that the population proportion is less than 0.60.

[TOP](#) [^](#)
[v](#)

39. Answer :

[TOP](#) [^](#)
[v](#)

Reason : The 97% confidence interval for the population mean is (18.553, 21.447).

$$\therefore \text{Sample mean} = \frac{(18.553 + 21.447)}{2} = 20$$

$$H_0 : \mu = 22$$

$$H_1 : \mu < 22$$

Under the null hypothesis the non-standardised critical value of the sampling distribution of mean in the lower tail, for the 3% significance level, will be =

$$22 - 1.881 \frac{4}{\sqrt{36}} = 20.746$$

The sample mean is less than this critical value (on non-standardised scale).

$\therefore$ We reject H_0 .

Hence we can conclude at 3% significance level that the population mean is less than 22.

40. Answer :

Reason : In testing the equality of two population means by using independent random samples, if the null hypothesis of equal population means is rejected at a significance level of 0.025, then it will always be rejected at a significance level of 0.05 because the rejection region increases.

41. Answer :

Reason : Estimated standard error of difference between means:

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}^2 = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} = \frac{4}{64} + \frac{9}{36} = 0.3125$$

$$s_1^2 = \frac{1}{n_1 - 1} \left[\sum x_1^2 - n_1 \bar{x}_1^2 \right] = \frac{1}{63} \left[1852 - 64 \left(\frac{320}{64} \right)^2 \right] = 4$$

$$s_2^2 = \frac{1}{n_2 - 1} \left[\sum x_2^2 - n_2 \bar{x}_2^2 \right] = \frac{1}{35} \left[891 - 36 \left(\frac{144}{36} \right)^2 \right] = 9$$

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{0.3125} = 0.559$$

Value of the test statistic =

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}} = \frac{\left(\frac{320}{64} - \frac{144}{36} \right) - 1}{0.559} = \frac{0}{0.559} = 0$$

42. Answer :

Reason : Pooled estimate of the underlying common population variance, $s_p^2 =$

$$\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(16 - 1)16^2 + (9 - 1)12^2}{16 + 9 - 2} = 217.04$$

Estimated standard error of difference between means =

$$s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = \sqrt{217.04 \left(\frac{1}{16} + \frac{1}{9} \right)} = 6.138$$

43. Answer :

$$H_0 : \mu_1 = \mu_2 + 1 \Rightarrow \mu_1 - \mu_2 = 1$$

$$\text{Test statistic} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{(\bar{x}_1 - \bar{x}_2) - 1}{\sqrt{217.032 \left(\frac{1}{16} + \frac{1}{9} \right)}} = 0.1629$$

$$\text{or } (\bar{x}_1 - \bar{x}_2) = 1 + 0.1629 \sqrt{217.032 \left(\frac{1}{16} + \frac{1}{9} \right)} \approx 2$$

$$\therefore \bar{x}_1 - \bar{x}_2 = 2 \quad \dots \dots \dots (A)$$

$$\text{Given: } \bar{x}_1 + \bar{x}_2 = 38 \quad \dots \dots \dots (B)$$

$$(A) + (B): 2\bar{x}_1 = 40 \text{ or } \bar{x}_1 = \frac{40}{2} = 20$$

$$\text{Putting the value of } \bar{x}_1 \text{ in (B): } 20 + \bar{x}_2 = 38 \text{ or } \bar{x}_2 = 38 - 20 = 18$$

Reason : $\therefore \bar{x}_1 = 20$ and $\bar{x}_2 = 18$

44. Answer :

$$\text{Reason : } H_0 : \mu_1 = \mu_2 + 1 \Rightarrow \mu_1 - \mu_2 = 1$$

$$H_1 : \mu_1 > \mu_2 + 1 \Rightarrow \mu_1 - \mu_2 > 1$$

It is a right-tailed test at a significance level of 0.05. Both the samples are large. Hence the normal distribution should be used to approximate the sampling distribution of the difference between sample means.

Value of the test statistic =

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\left(\frac{640}{64} - \frac{144}{36} \right) - 1}{\sqrt{\frac{256}{64} + \frac{144}{36}}} = \frac{6 - 1}{2.828} = 1.768$$

The critical value under the right tail obtained from the standard normal distribution is 1.645. The observed value of the test statistic exceeds the critical value; so it falls in the rejection region. Hence we reject the null hypothesis.

45. Answer : (a)

$$\text{Reason : } H_0 : p = 0.25$$

$$H_1 : p \neq 0.25$$

$$\text{Standard error of proportion} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.25(1-0.25)}{200}} = 0.03062$$

$$\text{Sample proportion} = \frac{42}{200} = 0.21$$

$$\text{Test statistic: } z = \frac{0.21 - 0.25}{0.03062} = -1.31$$

This is a large sample test of proportion. The standard normal distribution will be used to approximate the sampling distribution of proportion. The test will be a two-tailed test. From the standard normal distribution table and using interpolation the critical values are -1.645 and 1.645 .

We can see that the test statistic does not fall below the lower tail critical

[TOP](#)

[TOP](#)

value. Hence the test statistic falls in the acceptance region. We accept H_0 and conclude that proportion of college students among the readers is 0.25.

46. Answer :

Reason : p_1 : proportion of men cooperating

p_2 : proportion of women cooperating

Null hypothesis : $p_1 = p_2$

Alternative hypothesis : $p_1 \neq p_2$

$$\bar{p}_1 = \frac{90}{200} = 0.45$$

$$\bar{p}_2 = \frac{90}{300} = 0.30$$

Estimated value of the population proportion , $p =$

$$\frac{(n_1 \bar{p}_1) + (n_2 \bar{p}_2)}{n_1 + n_2} = \frac{(200 \times 0.45) + (300 \times 0.30)}{200 + 300} = \frac{180}{500} = 0.36$$

$$\begin{aligned} \sigma_{\bar{p}_1 - \bar{p}_2} (\text{estimated}) &= \sqrt{\left(\frac{p(1-p)}{n_1}\right) + \left(\frac{p(1-p)}{n_2}\right)} \\ &= \sqrt{\left(\frac{0.36(1-0.36)}{200}\right) + \left(\frac{0.36(1-0.36)}{300}\right)} \\ &= 0.04382 \end{aligned}$$

$$\text{Test statistic : } \frac{(0.45 - 0.30) - 0}{0.04382} = 3.42$$

Critical values: ± 1.96

The test statistic goes beyond the critical value in the right tail. Hence we reject null hypothesis. We can conclude that the proportions of men and women who cooperate with the interviewer are not equal.

47. Answer :

Reason : H_0 : $\mu = 9$

H_1 : $\mu \neq 9$

The sample is small and the population variance is not known. The sample variance is specified.

The population is normally distributed. Hence we should use the t distribution with $25 - 1 = 24$ d.o.f.

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{\sqrt{256}}{\sqrt{25}} = 3.20$$

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{15.75 - 9}{3.20} = 2.109$$

At $\alpha = 0.05$, the critical values are ± 2.064 .

The test statistic is more than the right tail critical value. So it falls in the rejection region.

$\therefore$ We reject H_0 . But the true mean is 9. So H_0 is true. Hence the test leads to a type I error.

48. Answer :

Reason : By careful observation of equations, it can be interpreted that as the time

(i.e., the value of x increases by one unit) passes on, national income is increasing by 1.5 units, where as population is increasing only by 1.25 units. In other words, growth rate in national income is more than that in population. Hence, as the time passes on, per capita income goes on increasing. Hence, the estimated per capita income is highest in the year 2003. This can be proved through calculations.

Year	1999	2000	2001	2002	2003
National income (A)	1.86	3.36	4.86	6.36	7.86
Population (B)	2.75	4	5.25	6.5	7.75
Per capita income (A/B)	0.676	0.84	0.926	0.978	1.014

49. Answer :

Reason : Let the following notations be used:

X : Year

x : Coded value of year = $2(X - \bar{X})$

Y : Research expenditure (in Rs.)

Year X	$x = 2(X - \bar{X})$	Y	XY	x^2
1998	-5	50	-250	25
1999	-3	57	-171	9
2000	-1	61	-61	1
2001	1	65	65	1
2002	3	72	216	9
2003	5	79	395	25
$\Sigma X = 12003$	0	$\Sigma Y = 384$	$\Sigma XY = 194$	$\Sigma x^2 = 70$

$$\bar{X} = \frac{12003}{6} = 2000.5 \quad \hat{Y} = a + bx$$

$$b = \frac{\Sigma xY}{\Sigma x^2} = \frac{194}{70} = 2.77 \quad a = \bar{Y} = \frac{384}{6} = 64$$

$\therefore$ Three linear trend equation is $\hat{Y} = 64 + 2.77x$

Year (X) = 2003

$$x = 2(2004 - 2000.5) = 7$$

$$Y = 64 + 2.77(7) = 83.39$$

$\therefore$ The research and development expenditure for the year 2004 is Rs.83.39 lakh

50. Answer :

$$\text{Reason : Relative cyclical residual measure} = \frac{Y - \hat{Y}}{\hat{Y}} \times 100$$

$$-0.346 = \frac{4.516 - \hat{Y}}{\hat{Y}} \times 100 \quad \text{or} \quad -0.346\hat{Y} + 100\hat{Y} = 4.516 \times 100$$

[TOP](#)

[TOP](#)

$$\hat{y} = \frac{4.516 \times 100}{(100 - 0.346)} = 4.5317$$

or

51. Answer :

Reason : Let the following notations be used.

X : Sales (Rs. in crores)

Y : Net profit (Rs. in crores)

$$\text{Coefficient of correlation, } r = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\sqrt{\Sigma(X - \bar{X})^2 \Sigma(Y - \bar{Y})^2}}$$

A. X	42	45	52	34	60	$\Sigma X = 233$
B. Y	2.5	2.45	3.25	1.75	4.00	$\Sigma Y = 13.95$
C. X -	-4.6	-1.6	5.4	-12.6	13.4	
D. Y -	-0.29	-0.34	0.46	-1.04	1.21	
E. C × D	1.334	0.544	2.484	13.104	16.214	$\Sigma(X - \bar{X})(Y - \bar{Y}) = 33.68$
F. C ²	21.16	2.56	29.16	158.76	179.56	$\Sigma(X - \bar{X})^2 = 391.2$
G. D ²	0.0841	0.1156	0.2116	1.0816	1.4641	$\Sigma(Y - \bar{Y})^2 = 2.957$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{233}{5} = 46.6 \quad \bar{Y} = \frac{\Sigma Y}{n} = \frac{13.95}{5} = 2.79$$

$$\therefore r = \frac{33.68}{\sqrt{391.2 \times 2.957}} = 0.99 \text{ (approx.)}$$

Percentage of variations in net profit that is explained by the variations in sales

$$= \text{coefficient of determination} = r^2 = (0.99)^2 = 0.9801$$

$\therefore$ 98.01% of the variations in net profits is explained by the variations in sales.

52. Answer :

Reason : The Y intercept of the regression line represents the mean value of Y when X = 0.

53. Answer :

$$\text{Reason : Coefficient of determination, } r^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \bar{Y})^2}$$

Y denotes the observed values and $\hat{Y}$ denotes the estimated values.

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{792}{8} = 99$$

$$\Sigma(Y - \bar{Y})^2 = \Sigma Y^2 - n\bar{Y}^2 = 79144 - 8 \times 99^2 = 736$$

[TOP](#)

[TOP](#)

[TOP](#)

$$\therefore \text{Coefficient of determination, } r^2 = 1 - \frac{200}{736} = 0.7283$$

54. Answer :

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

$$\text{Reason : Standard error of estimate, } s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n-2}}$$

$$\Sigma Y^2 = 19,00,400 \text{ (given)}$$

$$\Sigma Y = \bar{Y} \times n = 612 \times 5 = 3060$$

$$\Sigma XY = 18,100 \text{ (given)}$$

$$\hat{Y} = 1,236 - 104X$$

$$\therefore a = 1,236 \text{ and } b = -104$$

$$\therefore s_e = \sqrt{\frac{19,00,400 - (1236)(3060) - (-104)(18,100)}{5-2}} = 14.61.$$

The sample size is less than 30 and the standard deviation of the population is not known. Hence the appropriate distribution to be used for the prediction interval is the t-distribution.

$$\text{Degrees of freedom} = n - 2 = 5 - 2 = 3.$$

The appropriate t-value is 2.353 (from the t-table)

The limits of the prediction interval are given below:

$$\text{For, } x = 7.50, \quad \hat{Y} = 1236 - 104(7.50) = 456$$

$$\begin{aligned} \text{The prediction interval is: } \hat{Y} - t(s_e) &= 456 \pm 2.353(14.61) \\ &= 456 \pm 34.38 \end{aligned}$$

55. Answer :

Reason : In the linear model, the residuals are assumed to have a normal distribution with a mean of zero.

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

56. Answer :

Reason : A simple regression equation is of the form:

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

$$\hat{Y} = a + bX$$

Given:

$$\text{If } X = 7, \hat{Y} = 39 \therefore 39 = a + 7b \dots\dots\dots (A)$$

$$\text{If } X = 12, \hat{Y} = 64 \therefore 64 = a + 12b \dots\dots\dots (B)$$

Subtracting (A) from (B) we get:

$$64 - 39 = (12 - 7)b$$

$$25 = 5b$$

$$\text{or } b = \frac{25}{5} = 5$$

Putting the values of b in (B):

$$64 = a + 12 \times 5 = a + 60$$

$$\text{or } a = 64 - 60 = 4$$

$$\therefore \hat{Y} = 4 + 5X$$

$$\bar{Y} - b\bar{X} = \bar{Y} - b\left(\frac{\sum X}{n}\right)$$

Now, a =

$$\bar{Y} = a + b\left(\frac{\sum X}{n}\right) = 4 + 5\left(\frac{96}{8}\right) = 64$$

∴

∴ Mean of the observed values of Y, i.e. $\bar{Y} = 64$

57. Answer :

Reason : When the slope of a regression line is positive there is a positive correlation between the dependent and independent variables

[TOP](#)
↕

58. Answer :

Reason : Regression equation of Y on X is derived as shown below:

[TOP](#)
↕

Height X of the Father (in)	65	63	67	64	68	62	70	66	68	67	69	71	$\Sigma X = 800$
Height Y of the son (in)	68	66	68	65	69	66	68	65	71	67	68	70	$\Sigma Y = 811$
XY	4420	4158	4556	4160	4692	4092	4760	4290	4828	4489	4692	4970	54107
X^2	4225	3969	4489	4096	4624	3844	4900	4356	4624	4489	4761	5041	53418

$$n = 12$$

$$b = \frac{n \sum XY - \sum X \cdot \sum Y}{n \sum X^2 - (\sum X)^2} = \frac{(12 \times 54107) - (800 \times 811)}{(12 \times 53418) - (800)^2} = \frac{484}{1016} = 0.476$$

$$a = \bar{Y} - b\bar{X} = \left(\frac{811}{12}\right) - 0.476 \times \left(\frac{800}{12}\right) = 35.85$$

∴ The regression equation is: $\hat{Y} = 35.85 + 0.476X$

59. Answer :

[TOP](#)
↕

Reason : Slope of a regression equation, $b = \frac{\text{Cov}(X, Y)}{\sigma_x^2}$

$$= \frac{r \sigma_x \sigma_y}{\sigma_x^2}$$

$$= \frac{0.75 \times 6 \times \sigma_y}{36} = 2.5$$

$$\text{Or } \sigma_y = \frac{2.5 \times 36}{0.75 \times 6} = 20$$

60. Answer :

[TOP](#)
↕

$$\text{Coefficient of correlation} = + \sqrt{\text{Coefficient of determination}}$$

$$= + \sqrt{1 - \text{Proportion of unexplained variations}}$$

$$= \sqrt{1 - 0.36}$$

Reason : $= 0.80$

(Positive square root because y increases as x increases and vice versa)

61. Answer :

Reason : The standard error of estimate is a measure of the variability of the observed values of the dependent variable around the regression line.

[TOP](#)
[↑](#)

62. Answer :

Reason : Fisher's ideal price index = $\sqrt{\text{Laspeyres index} \times \text{Paasches index}}$

[TOP](#)
[↑](#)

$$\begin{aligned}
 &= \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100} \\
 &= \sqrt{\left(\frac{\sum P_1 Q_0}{\sum Q_0} \right) \times \left(\frac{\sum P_1 Q_1}{\sum Q_1} \right) \times 100} \\
 &= \sqrt{\left(\frac{\sum P_0 Q_0}{\sum Q_0} \right) \times \left(\frac{\sum P_0 Q_1}{\sum Q_1} \right) \times 100} \\
 &= \sqrt{\frac{6}{5} \times \frac{5}{4} \times 100} = 122.47
 \end{aligned}$$

63. Answer :

[TOP](#)
[↑](#)

$$\text{Reason : } r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

$$r = +1 \Rightarrow \text{Cov}(x, y) = \sigma_x \cdot \sigma_y.$$

64. Answer :

[TOP](#)
[↑](#)

Reason : The coefficient of determination is the square of coefficient of correlation (r). Hence it will always be ≥ 0 and ≤ 1 .

From above we can see that the coefficient of determination cannot be less than zero.

The coefficient of determination is will be equal to 1, only if the coefficient of correlation is equal to -1 or 1 ; in other cases it will be > 0 and < 1 .

The coefficient of determination is always positive; The coefficient of correlation may be negative.

Since $-1 \leq r \leq 1$, the coefficient of determination which is the square of coefficient of correlation (r), will always have magnitude less than or equal to r

65. Answer :

[TOP](#)
[↑](#)

$$\text{Reason : } \hat{Y} = a + bX$$

$$\text{Given: } 7 = a + b \quad (2) \quad (\text{For } X=2)$$

$$\text{Or } 7 = a + 2b \dots\dots\dots (A)$$

$$11.8 = a + b \quad (8) \quad (\text{For } X=8)$$

$$\text{Or } 11.8 = a + 8b \dots\dots\dots (B)$$

$$\therefore b = \frac{(11.8-7)}{(8-2)} = 0.80$$

Putting $b = 0.80$ in (A)

$$7 = a + 2(0.80)$$

$$\text{or } a = 7 - 1.6 = 5.4$$

$\therefore$ The regression equation is

$$\hat{Y} = 5.4 + 0.80X$$

$$\text{Standard error of estimate; } S_e = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n-2}}$$

$$\sum Y^2 = 791.31 \quad \sum XY = 452.7$$

$$\sum Y = 78.1 \quad n = 8$$

$$\begin{aligned} \therefore S_e &= \sqrt{\frac{791.31 - (5.4 \times 78.1) - (0.8 \times 452.7)}{8-2}} \\ &= 1.1113 \end{aligned}$$

66. Answer :

$$\begin{aligned} \text{Reason : Slope of a regression equation, } b &= \frac{\text{Cov}(X, Y)}{\sigma_x^2} \\ &= \frac{r \sigma_x \sigma_y}{\sigma_x^2} \\ &= \frac{0.80 \times 5 \times 8}{25} = 1.28 \end{aligned}$$

[TOP](#)

67. Answer :

Reason : In a regression analysis, if a null hypothesis $H_0: B = 0$ is rejected. This indicates that there is a significant relationship between X and Y

[TOP](#)

68. Answer :

Reason : The statement that there are always as many points above the fitted line as there are below it is wrong.

[TOP](#)

69. Answer :

$$\text{Reason : } \hat{Y} = 20 + 0.75X.$$

$$\text{At the point } X=100, Y=90, \hat{Y} = 20 + 0.75 \times 100 = 75 + 20 = 95 \text{ but } Y = 90$$

$$\therefore \text{Residual } Y - \hat{Y} = 90 - 95 = -5$$

[TOP](#)

70. Answer :

Reason : Larger values of r^2 imply that the observations are more closely grouped about the least squares line

[TOP](#)

71. Answer :

[TOP](#)

Reason : The slope (b) of a regression equation is given by

$$b = \frac{\text{Cov}(A, B)}{\sigma_A^2}$$

$$\text{or } \text{Cov}(A, B) = b \sigma_A^2 = 0.64 \times 25 = 16$$

$$\therefore r = \frac{\text{Cov}(A, B)}{\sigma_A \sigma_B} = \frac{16}{5 \times 4} = 0.80$$

72. Answer :

[TOP](#)

Reason : The central line of for the **R** chart = Mean of the sample range ($\bar{R}$)

$$\bar{R} = \sum R / n = 63 / 10 = 6.3.$$

$$\begin{aligned} \text{The value of the lower control limit for the R chart} &= \bar{R} D_3 \\ &= 6.3 \times 0 = 0 \end{aligned}$$

73. Answer :

[TOP](#)

Reason : We calculate the variances of the two samples as follows:

$$H_0 : \sigma_1^2 = \sigma_2^2 \Rightarrow \frac{\sigma_2^2}{\sigma_1^2} = 1$$

$$H_1 : \sigma_1^2 < \sigma_2^2 \Rightarrow \sigma_2^2 > \sigma_1^2 = \frac{\sigma_2^2}{\sigma_1^2} > 1$$

Number	X_1	$(X_1 - \bar{X}_1)$	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)$	$(X_2 - \bar{X}_2)^2$
1	60	-15	225	61	-16	256
2	65	-10	100	66	-11	121
3	71	-4	16	67	-10	100
4	74	-1	1	85	8	64
5	76	1	1	78	1	1
6	82	7	49	63	-14	196
7	85	10	100	85	8	64
8	87	12	144	86	9	81
9				88	11	121
10				91	14	196
Sum	600		636	770		1200
Means	75			77		
Sample Variance			90.857			133.333

Since the values in the F distribution table are appropriate for only upper-tailed tests, we calculate the test statistic by keeping the higher sample variance in the numerator. H_1 is also formulated accordingly, as shown above.

$$\text{The test statistic } F = \frac{S_2^2}{S_1^2} = \frac{133.333}{90.857} = 1.468$$

74. Answer :

[TOP](#)

Reason : The value of the grand mean

$$\bar{\bar{X}} = \frac{\sum X}{N} = \frac{45+50+60+65}{20} = \frac{220}{20} = 11$$

Variance between the sample means can be computed as below:

n_j	$\bar{X}_j$	$(\bar{X}_j - \bar{\bar{X}})^2$	$n_j (\bar{X}_j - \bar{\bar{X}})^2$
5	9	4	20
5	10	1	5
5	12	1	5
5	13	4	20
Total			50

$$\text{The variance within sample means} = s_x^2 = \frac{\sum n_j (\bar{X}_j - \bar{\bar{X}})^2}{k-1} = \frac{50}{4-1} = 16.7$$

75. Answer :

Reason : The lower confidence limit for the estimated population variance is given by

$$\sigma_L^2 = \frac{(n-1)s^2}{\chi_U^2}$$

The value of Chi square statistic from the table with 0.025 area in the right tail and $(9-1) = 8$ degrees of freedom is **17.535**.

∴ The lower confidence limit for the estimated population variance is

$$\sigma_L^2 = \frac{(9-1)(23)^2}{17.535} = 241.35$$

76. Answer :

Reason : The expected frequency can be tabulated as below:

Beverages	Male	Female	Total
A	34.83	20.17	55
B	34.83	20.17	55
C	25.34	14.66	40
Total	95	55	150

The Chi square statistic can be found out as below:

Observed Frequency (O)	Expected Frequency (E)	$(O - E)^2 / E$
25	34.83	2.774
40	34.83	0.767
30	25.34	0.857
30	20.17	4.791
15	20.17	1.325

10	14.66	1.481
----	-------	-------

The value of Chi square statistic, $\chi^2 = \sum \frac{(O-E)^2}{E} = \mathbf{11.996}$

77. Answer :

[TOP](#)
[↑](#)
[↓](#)

Reason : The number of degrees of freedom can be calculated as below:

$$N = (r - 1)(c - 1) = (2 - 1)(4 - 1) = 3.$$

For 3 degrees of freedom and 5% level of significance, the critical value of chi square distribution is **7.815** (from F table).

78. Answer :

[TOP](#)
[↑](#)
[↓](#)

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

79. Answer :

[TOP](#)
[↑](#)
[↓](#)

Reason : A decision tree is a graphical method which represents states of nature and courses of action.

A histogram is a graphical representation of a frequency distribution.

A scatter diagram shows the distribution of data points in regression analysis.

A frequency distribution is a distribution of data along with their frequencies.

A probability distribution is a distribution of values of a random variable along with their respective probabilities

80. Answer :

[TOP](#)
[↑](#)
[↓](#)

Reason : The sample mean is not an estimator of the population variance.

The sample mean is an unbiased estimator of the population mean.

The sample mean is not an estimator of the population proportion.

The sample mean is not an estimator of the standard error of mean.

The sample mean is not an estimator of the standard error of proportion

81. Answer :

[TOP](#)
[↑](#)
[↓](#)

Reason : Simple random sampling may not be appropriate when the population is known to consist of well defined groups such that the elements within each group are heterogeneous and each group is a representative of the population as a whole, because even if the sample is random it may not reflect the nature of the population.

When the population is known to consist of well-defined groups such that the elements within each group are heterogeneous and each group is a representative of the population as a whole, the cluster sampling is appropriate.

When the population is known to consist of well-defined groups such that the elements within each group are homogeneous and the groups vary from

each other significantly, the stratified sampling is appropriate.

When the population is known to consist of well defined groups such that the elements within each group are heterogeneous and each group is a representative of the population as a whole, because even if the sample is random it may not reflect the true nature of the population.

When the population is known to consist of well defined groups such that the elements within each group are heterogeneous and each group is a representative of the population as a whole, judgmental sampling may not be appropriate because the representativeness of the sample depends upon the knowledge and judgment of the decision maker

82. Answer :

Reason :The property of efficiency with regard to an estimator does not refer to the sample size.

The property of efficiency with regard to an estimator refers to the size of the standard error.

The property of efficiency with regard to an estimator does not refer to the size of the standard deviation of the sample.

The property of efficiency with regard to an estimator does not refer to the size of the sample variance.

The property of efficiency with regard to an estimator does not refer to the size of the sample mean

83. Answer :

Reason :The graphical plot of the values of the dependent and independent variables, in the context of regression analysis, is called scatter diagram.

A frequency polygon is a graphical representation of a frequency distribution which uses straight lines to join the top mid points of the rectangles in a histogram.

A histogram is a graphical representation of a frequency distribution.

A p chart is a quality control chart.

An ogive is a graphical plot of a cumulative frequency distribution

84. Answer :

Reason :Fisher's ideal price index uses both current year and base year quantities.

Laspeyres price index uses only base year quantities.

Paasche's price index uses only current year quantities.

Marshall Edgeworth price index uses both current year and base year quantities.

Fixed weight aggregates price index uses the weights from a representative period

85. Answer :

Reason :‘Hugging the control limits’ indicates that two different populations are being observed.

[TOP](#) [↑](#)
[↓](#)

[TOP](#) [↑](#)
[↓](#)

[TOP](#) [↑](#)
[↓](#)

[TOP](#) [↑](#)
[↓](#)

‘Hugging the center line’ indicates that the variations have been reduced significantly.

Cycles indicate the possibility of presence of random variations in the process.

Increasing trend indicates that the process mean is increasing.

Decreasing trend indicates that the process mean is decreasing

[< TOP OF THE DOCUMENT >](#)