

**Question Paper**  
**Quantitative Methods – II (MB152): July 2008**

- Answer all 72 questions.
- Marks are indicated against each question.

Total Marks : 100

1. A box contains 3 black, 5 white and 2 red balls. If 3 balls are drawn at random without replacement, the probability that all the 3 balls are black is <Answer>
- (a)  $\frac{1}{130}$   
(b)  $\frac{1}{120}$   
(c)  $\frac{1}{150}$   
(d)  $\frac{1}{140}$   
(e)  $\frac{1}{125}$  (1 mark) <Answer>
2. In a simultaneous tossing of two unbiased coins, the probability of obtaining at least one tail is <Answer>
- (a) 1.00  
(b) 0.75  
(c) 0.5  
(d) 0.375  
(e) 0.25. (1 mark) <Answer>
3. If two dice are thrown simultaneously, what is the probability of obtaining a total score of eight? <Answer>
- (a)  $\frac{1}{9}$   
(b)  $\frac{1}{12}$   
(c)  $\frac{5}{36}$   
(d)  $\frac{1}{6}$   
(e)  $\frac{2}{9}$  (1 mark) <Answer>
4. A fair coin is thrown in the air four times. If the coin lands with the head up on the first three tosses, what is the probability that the coin will land with the head up on the fourth toss? <Answer>
- (a) 0  
(b) 1  
(c)  $\frac{1}{16}$   
(d)  $\frac{1}{8}$   
(e)  $\frac{1}{2}$  (1 mark) <Answer>
5. If event B is dependent on event A.  $P(A \text{ and } B) = 0.30$  and  $P(B) = 0.60$ , then  $P(A/B)$  is <Answer>
- (a) 0.18  
(b) 0.30  
(c) 0.40 (1 mark)

- (d) 0.50
- (e) 0.90.

6. There are 30 boys and 20 girls in a class. From the class list, a name is picked up at random. What is the probability that it is a girl's name? <Answer>

- (a)  $\frac{1}{5}$
- (b)  $\frac{2}{5}$
- (c)  $\frac{4}{5}$
- (d)  $\frac{6}{5}$
- (e)  $\frac{7}{5}$ .

(1 mark)

7. The probability that A can solve a problem is  $\frac{2}{3}$  and the probability that B can solve the problem is  $\frac{3}{5}$ . What is the probability that at least one of A and B will be able to solve the problem? <Answer>

- (a)  $\frac{2}{15}$
- (b)  $\frac{1}{5}$
- (c)  $\frac{4}{15}$
- (d)  $\frac{7}{15}$
- (e)  $\frac{13}{15}$ .

(2marks)

8. There are 20 good bulbs and 2 defective bulbs in a packet. From the packet two bulbs are drawn randomly without replacement. What is the probability of getting a good bulb in the first trial and a defective bulb in the second trial? <Answer>

- (a)  $\frac{20}{22}$
- (b)  $\frac{2}{21}$
- (c)  $\frac{40}{441}$
- (d)  $\frac{20}{231}$
- (e)  $\frac{21}{130}$ .

(2marks)

9. If A and B are two independent events and the marginal probability of A is  $P_a$  and that of B is  $P_b$ , then the joint probability of A and B will be <Answer>

- (a)  $P_a + P_b$
- (b)  $P_a \times P_b$
- (c)  $P_a + P_b - P_a \times P_b$
- (d)  $P_a + P_b + P_a \times P_b$
- (e)  $(1 - P_a)(1 - P_b)$ .

(1 mark)

10. A random variable X has the following probability distribution: <Answer>

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

Let  $Y = X - 25$

What is the expected value of Y?

- (a) 20
- (b) 25
- (c) 38
- (d) 45
- (e) 63.

(2marks)

[<Answer>](#)

11. Which of the following is **not** a characteristic of Bernoulli's process?

- (a) Each trial has only two outcomes
- (b) The probability of outcome of any trial remains fixed over time
- (c) The trials are statistically independent
- (d) The probability of occurrence of one outcome is equal to the probability of occurrence of the other
- (e) The occurrence of one trial cannot influence the outcome of any other trial.

(1 mark)

[<Answer>](#)

12. The number of mistakes in a page, consisting of large number of words follows a

- (a) Normal distribution
- (b) Poisson distribution
- (c) Hypergeometric distribution
- (d) Uniform distribution
- (e) Binomial distribution.

(1 mark)

[<Answer>](#)

13. A normal probability distribution

- (a) Is a discrete probability distribution
- (b) Is a continuous probability distribution
- (c) Must always have a variance of 1
- (d) Must always have a mean of zero
- (e) Must always have a mean of 1.

(1 mark)

[<Answer>](#)

14. Using Poisson distribution, the probability that the ace of spades will be drawn from a pack of well shuffled cards at least once in 104 consecutive trials is

- (a) 0.8647
- (b) 0.5897
- (c) 0.2356
- (d) 0.0254
- (e) 0.9885.

(2 marks)

[<Answer>](#)

15. A t-statistic calculated for a sample mean is known to have 15 degrees of freedom. What is the sample size?

- (a) 15
- (b) 16
- (c) 30
- (d) 14
- (e) The sample size cannot be determined.

(1 mark)

[<Answer>](#)

16. Y follows binomial distribution with mean of 4 for 20 trials. What will be the standard deviation of Y for 25 trials?

- (a) 2
- (b) 4
- (c) 5
- (d) 20
- (e) 25.

(2 marks)

[<Answer>](#)

17. Which of the following is the **most** appropriate example of a continuous random variable?

- (a) Weight of a randomly selected human being
- (b) Number of patients visiting in a randomly selected doctor
- (c) Number of seats vacant in a randomly selected theatre
- (d) Number of passengers in a randomly selected bus
- (e) Number of students absent in a class.

(1 mark)

[<Answer>](#)

18. Assume that on an average one telephone number out of fifteen is busy. What is the probability that if six randomly selected telephone numbers are called, at most three will be busy?

- (a) 0.5979
- (b) 0.6979
- (c) 0.7979
- (d) 0.8979
- (e) 0.9997.

(2 marks)

[<Answer>](#)

19. A taxi cab company has 6 Ambassadors and 4 Fiats. If 3 of these taxi cabs are in the shop for repairs and Ambassador is as likely to be in for repairs as a Fiat. Then the probability that two of the cabs

(2 marks)

and Ambassador is as likely to be in for repairs as a Fiat. Then the probability that two of the cabs which are in the shop for repair are Ambassadors is

- (a)  $\frac{1}{8}$
- (b)  $\frac{1}{6}$
- (c)  $\frac{1}{4}$
- (d)  $\frac{1}{2}$
- (e)  $\frac{1}{10}$

20. What is the expected value of the number of heads when two fair coins are tossed simultaneously?

<Answer>

- (a)  $\frac{1}{2}$
- (b)  $\frac{1}{4}$
- (c) 1
- (d)  $\frac{3}{4}$
- (e)  $\frac{5}{4}$

(2 marks)

21. From an association consisting of 540 individuals, a sample of 60 individuals is taken. From this sample, the average age of the individuals is found to be 31 years and the standard deviation is found to be 6.84 years.

<Answer>

A 95 percent confidence interval for the mean age of the individuals in the association has to be constructed. The lower and upper confidence limits of the confidence interval are

- (a) 24.16 years and 37.84 years respectively
- (b) 28.25 years and 33.25 years respectively
- (c) 29.37 years and 32.63 years respectively
- (d) 30.17 years and 31.83 years respectively
- (e) 17.6 years and 44.41 years respectively.

(3 marks)

<Answer>

22. Which of the following statements is **false** regarding business research?

- (a) It can be defined as the plan and structure of enquiry, formulated in order to obtain answers to research questions on business aspects
- (b) The research plan constitutes the overall programme of the business research process
- (c) It can be understood as that which gives the blueprint for collection, measurement and analysis of business data
- (d) It helps researchers to utilize available resources efficiently to achieve research objectives
- (e) It is not an important tool for decision-making.

(1 mark)

<Answer>

23. Which of the following statements is **false** regarding variables?

- (a) Continuous variables can take on different values even in decimal places
- (b) Non-continuous variables are those that cannot be measured in decimals
- (c) If a variable is dependent on the result of some other variable, it is then called a dependent variable
- (d) An independent variable is one that is not dependent on any other variable with reference to that particular study
- (e) Extraneous variables are dependent variables.

(1 mark)

<Answer>

24. Which of the following statements is **true** regarding secondary data?

- (a) Secondary data is already available as a result of research carried out by others
- (b) Secondary data is that the data which is collected first time on subject
- (c) Using secondary data every time we can get correct conclusions to current situation
- (d) Secondary data cannot be obtained from journals, magazines and periodicals

(1 mark)

- (d) Secondary data cannot be obtained from journals, magazines and periodicals
- (e) To collect secondary data it takes much time and more money.

<Answer>

25. Which of the following are steps in a sampling process?

- I. Defining the Target Population.
- II. Specifying the Sampling Frame.
- III. Selection of the Sampling Method.
- IV. Specifying the Sampling Plan.
- V. Selecting the Sample.

- (a) (I), (II) and (III) above
- (b) (II), (III) and (IV) above
- (c) (III), (IV) and (V) above
- (d) (I), (III), (IV) and (V) above
- (e) All (I), (II), (III), (IV) and (V) above.

(1 mark)

<Answer>

26. The important characteristic of a good research design, which is essentially devised to minimize the effect of extraneous variables?

- (a) Control
- (b) Time
- (c) Cost
- (d) Data analysis
- (e) Survey.

(1 mark)

<Answer>

27. Which of the following is the key objective of any business research?

- (a) To develop the capital of the company
- (b) To get all the information collected to the management
- (c) To provide accurate, relevant and timely information to the top management
- (d) To remain unchanged in the cost capital
- (e) To have all the records of the data.

(1 mark)

<Answer>

28. Which of the following statement(s) is/are **true** regarding evaluation of the data?

- I. Several mathematical and statistical models are used to evaluate the data.
- II. Evaluation of the data normally starts with editing and coding of the data.
- III. Evaluation of the data is done after the Analysis.

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (II) above
- (e) Both (II) and (III) above.

(1 mark)

<Answer>

29. Which of the following is/are form(s) of business research?

- I. Market Research.
- II. Operations Research.
- III. Motivational Research.

- (a) Only (I) above
- (b) Only (III) above
- (c) Both (I) and (II) above
- (d) Both (II) and (III) above
- (e) All (I), (II) and (III) above.

(1 mark)

<Answer>

30. Which of the following are the parts of a typical research report?

- I. Prefatory information.
- II. Methodology.
- III. Introduction.
- IV. Findings.
- V. Conclusions and Recommendations.

- (a) Both (I) and (II) above
- (b) Both (III) and (IV) above
- (c) Both (IV) and (V) above
- (d) (I), (II) and (III) above
- (e) All (I), (II), (III), (IV) and (V) above.

(1 mark)

- (e) All (I), (II), (III), (IV) and (V) above.

[<Answer>](#)

31. Which of the following statements can be termed as a good research design?

- (a) It can be obtained with maximum utilization of time, money and effort
- (b) It can be obtained with minimum utilization of money, time and effort
- (c) It can be obtained with maximum utilization of time and minimum money and effort
- (d) It can be obtained with minimum utilization of money and time and maximum effort
- (e) It can be obtained with minimum utilization of time and maximum money and effort.

(1 mark)

[<Answer>](#)

32. Which of the following is/are the strengths of qualitative research?

- I. Ideal exploratory tool.
- II. Increase the value of subsequent quantitative research.
- III. Group dynamics enhance results.

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (II) above
- (e) All (I), (II) and (III) above.

(1 mark)

[<Answer>](#)

33. Which of the following is/are components of attitude?

- I. Cognitive Component (Beliefs).
- II. Affective Component (Feelings).
- III. Behavioural Component (Actions).

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (II) above
- (e) All (I), (II) and (III) above.

(1 mark)

[<Answer>](#)

34. Which of the following types of sampling, on proper application can significantly increase the statistical efficiency of sampling?

- (a) Stratified sampling
- (b) Simple random sampling
- (c) Systematic sampling
- (d) Convenience sampling
- (e) Judgment sampling.

(1 mark)

[<Answer>](#)

35. Which of the following is/are the single item scales which measure the attitudes of respondents?

- I. Itemized category scales.
- II. Rank order scales.
- III. Q-sort scales.
- IV. Comparative scales.

- (a) Only (I) above
- (b) Only (IV) above
- (c) Both (I) and (II) above
- (d) Both (III) and (IV) above
- (e) All (I), (II), (III) and (IV) above.

(1 mark)

[<Answer>](#)

36. Which of the following statements is/are **true** regarding sequential sampling?

- I. In this method the size of the sample is not fixed in advance.
- II. This sampling method is used for statistical quality control.
- III. A number of sample lots are drawn in sequence one after another from the population depending on the results of the earlier sample.

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (II) above
- (e) All (I), (II) and (III) above.

(1 mark)

[<Answer>](#)

37. The sampling frame error is said to occur

(1 mark)

- (a) When the list of members corresponds exactly with the target population
- (b) When the list of members does not correspond exactly with the target population
- (c) When the sample size does not correspond exactly with the target sample
- (d) When the sample size correspond exactly with the target population
- (e) When the list of members does not correspond exactly with the target sample.

<Answer>

38. In which of the following methods of sampling the population is divided into groups so that the elements in each group are homogeneous and the groups vary from each other?

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

(1 mark)

<Answer>

39. According to the Central Limit Theorem for large samples the sample mean approaches

- (a) The normal distribution
- (b) The binomial distribution
- (c) The hypergeometric distribution
- (d) The chi-square distribution
- (e) An unknown distribution.

(1 mark)

<Answer>

40. Which of the following statements is **false** with regard to sampling?

- (a) It is useful when the population is too large to be studied in full
- (b) It costs significantly less than a census
- (c) It is not suitable in case of destructive testing
- (d) It involves less work on statistical data processing than census
- (e) It involves less work on collection of statistical data.

(1 mark)

<Answer>

41. Which of the following are the causes for error in the sampling?

- I. Error in selection of the sample.
- II. Bias in the reporting of data.
- III. Diversity of the population.
- IV. Substitution of sampling units for convenience.
- V. Faulty demarcation of sampling universe.

- (a) (I), (II) and (III) above
- (b) (II), (III) and (IV) above
- (c) (III), (IV) and (V) above
- (d) (I), (II), (IV) and (V) above
- (e) All (I), (II), (III), (IV) and (V) above.

(1 mark)

<Answer>

42. The statistical measures like mean, median, mode and standard deviation which are used to describe the characteristics of an universe, are known as

- (a) Statistics
- (b) Measurements
- (c) Parameters
- (d) Constants
- (e) Estimates.

(1 mark)

<Answer>

43. Which of the following is **not** a type of random sampling?

- (a) Multistage sampling
- (b) Stratified sampling
- (c) Systematic sampling
- (d) Judgment sampling
- (e) Cluster sampling.

(1 mark)

<Answer>

44. The mean lifetime of a sample of 100 light bulbs produced by a company is found to be 1,580 hours with a standard deviation of 90 hours. We want to test the hypothesis that the mean lifetime of the bulbs produced by the company is 1,600 hours.

What is the value of the appropriate test statistic for the above test?

- (a) 1.8888
- (b) - 1.2222
- (c) 2.8888

(2marks)

- (c) 2.8888
- (d) - 2.2222
- (e) 3.2222.

[<Answer>](#)

45. The Mumbai Cricket Club, a professional club for the cricketers, has the player who led the league in batting average for many years. Over the past ten years, Amod Kambli has achieved a batting average of 54.50 runs with a standard deviation of 5.5 runs. This year Amod played 25 matches and achieved an average of 48.80 runs only. Amod is negotiating his contract with the club for the next year, and the salary he will be able to obtain is highly dependent upon his ability to convince the team's owner that his batting average this year was not significantly worse than in the previous years.

We want to find out whether Amod's salary will be reduced next year. What would be the appropriate test statistic for this test?

- (a) 2.8498
- (b) 3.9608
- (c) 4.0718
- (d) - 5.1818
- (e) - 6.2938.

(2marks)

[<Answer>](#)

46. The demand for a particular product in a factory was found to vary for every two months in a year. In a sample study the following information was obtained:

Month	Jan	March	May	July	Sep	Nov
Demand (in units)	1032	1026	1020	1019	1023	1012

What is the value of the test statistic for the above test?

- (a) 0.3159
- (b) 0.2250
- (c) 0.2079
- (d) 0.2159
- (e) 0.3079.

(3marks)

[<Answer>](#)

47. The weekly number of automobile accidents in 12 weeks in a certain city were as follows:

8	3	15	7	9	21	19	5	9	11	6	16
---	---	----	---	---	----	----	---	---	----	---	----

It is believed that the traffic conditions were same during the 12 weeks. We want to test whether the numbers obtained are in agreement with the belief. What is the critical value for the appropriate statistical test at a significance level of 5%?

- (a) 13.362
- (b) 19.675
- (c) 15.987
- (d) 16.919
- (e) 18.307.

(2marks)

[<Answer>](#)

48. Which of the following statement(s) is/are **true** regarding the Lognormal Distribution?

- I. If  $Y = \ln(X)$  is normally distributed random variable, then  $X$  is said to be a lognormal variable.
- II. If  $Y = \ln(X)$  is normally distributed random variable, then  $Y$  is said to be a lognormal variable.
- III. In finance "Continuous compounded returns are required to be normal". This property is described as "Stock Prices are Lognormal".

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (III) above
- (e) All (I), (II) and (III) above.

(1 mark)

[<Answer>](#)

49. The time that it takes to find a taxi when leaving a restaurant follows a left skewed distribution with a mean of 20 minutes and a variance of 100 minutes. If 64 restaurant patrons are randomly sampled and the average time that it takes for them to find a taxi is calculated, then what is the probability that the sample mean will be between 18 and 23 minutes?

- (a) 0.4918
- (b) 0.4452
- (c) 0.9918
- (d) 0.9452
- (e) 0.9370.

(3marks)



(e) 0.9370.

<Answer>

50. You are working as a purchase manager of a company. The following information has been supplied to you by two manufacturers of electric bulbs on the basis of sample data:

Particulars	Company A	Company B
Mean life time (in hours)	1,300	1,288
Standard deviation of life time (in hours)	82	93
Sample Size	100	100

Your quality manager hypothesizes that there is no significant difference between the mean lives of the electric bulbs of the two companies. You wish to test this hypothesis at 5% level of significance.

What is the estimated standard error of difference between means for this test?

- (a) 8.8558
- (b) 9.9668
- (c) 10.1778
- (d) 11.2888
- (e) 12.3988.

(2 marks)

<Answer>

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

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2$$

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

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

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 values are  $\pm 1.96$  and the null hypothesis is rejected
- (d) The critical values are  $\pm 1.96$  and the null hypothesis is accepted
- (e) The critical values are  $\pm 1.645$  and the null hypothesis is accepted.

(2 marks)

<Answer>

52. Which of the following statement(s) is/are **true** regarding advantages of sampling?

- I. A study of a sample is usually cheaper than a study of the population.
  - II. Sampling usually provides information quicker than a census so that timely decisions can be taken.
  - III. In sampling, the chances of errors while processing the data are more.
- (a) Only (I) above
  - (b) Only (II) above
  - (c) Only (III) above
  - (d) Both (I) and (II) above
  - (e) All (I), (II) and (III) above.

(2 marks)

<Answer>

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

$$H_0: \mu = 9$$

$$H_1: \mu < 9$$

$$n = 25$$

$$s^2 = 256$$

$$\bar{x} = 3.60$$

Significance level = 0.05

The population is normally distributed. It is later known that the true population mean is 7.

Which of the following is **true** regards 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.

(3 marks)

(e) The test leads to a type II error.

<Answer>

54. A banker claims that the life of a regular savings account opened with his bank averages 15 months with an S.D. of 6.45 months. Assuming that life of a regular savings account is normally distributed, what is the probability that there is balance in the savings account of a depositor at the end of 19 months?

- (a) 0.2754
- (b) 0.3607
- (c) 0.4603
- (d) 0.2676
- (e) 0.5245.

(2 marks)

<Answer>

55. Which of the following statement(s) is/are **true** regarding estimation?

- I. If an estimate of a population parameter is given by a single value, then the estimate is called an interval estimate.
- II. If an estimate of a population parameter is given by two distinct numbers between which the parameter may be considered to lie, then the estimate is called a point estimate.
- III. The interval estimate indicates the accuracy (or precision) of an estimate.

- (a) Only (I) above
- (b) Only (II) above
- (c) Only (III) above
- (d) Both (I) and (II) above
- (e) All (I), (II) and (III) above.

(1 mark)

<Answer>

56. If  $\bar{X} = 51$ ,  $n = 16$ ,  $\mu = 50$  and  $s = 1$ , then the value of t-statistic is

- (a) 1
- (b) 2
- (c) 3
- (d) 4
- (e) 5.

(2 marks)

<Answer>

57. Which of the following statements are **true** regarding Pearson's coefficient of correlation ( $\rho$ )?

- I. It lies between  $-\infty \leq \rho \leq \infty$ .
- II.  $\rho = 0 \Leftrightarrow$  No relationship between X and Y.
- III.  $\rho = 1 \Leftrightarrow$  Perfect positive linear relationship between X and Y.
- IV.  $\rho = -1 \Leftrightarrow$  Perfect negative linear relationship between X and Y.

- (a) Both (I) and (II) above
- (b) Both (II) and (III) above
- (c) Both (III) and (IV) above
- (d) (II), (III) and (IV) above
- (e) All (I), (II), (III) and (IV) above.

(1 mark)

<Answer>

58. The ranks of 5 students are given according to their marks in Statistics and Mathematics.

Student No.	1	2	3	4	5
Statistics	2	3	5	4	1
Mathematics	3	4	5	2	1

What is the Spearman's rank correlation coefficient?

- (a) 0.5
- (b) 0.6
- (c) 0.7
- (d) 0.8
- (e) 0.9.

(1 mark)

<Answer>

59. Students of the EAMCET batch are required to give a model examination before they sit for the main exam. The model exam is expected to give the students a chance of assessing their preparation under examination conditions. For a set of five students, the scores in the model exam and the corresponding main exam are as follows.

(1 mark)

Student Number	Model Exam Score (X)	Main Exam Score (Y)
101	85	89
203	76	80
357	68	78
465	80	70
509	90	92

Construct a regression line for the plot of main exam score vs. model exam score. What is the predicted main exam score, when model exam score is 95? (Round off your answer to the nearest integer)

- (a) 90
- (b) 91
- (c) 92
- (d) 93
- (e) 94.

[<Answer>](#)

60. A simple regression relationship was developed between two variables X and Y, with X as the independent variable.

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

The coefficient of determination is

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

(2 marks)

[<Answer>](#)

61. A regression model **cannot** explain any variations in the values of the dependent variable if

- (a) ESS = 0
- (b)  $R^2 = 1$
- (c) RSS = 0
- (d) RSS > 0
- (e) TSS > 0.

(1 mark)

[<Answer>](#)

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

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

(1 mark)

[<Answer>](#)

63. Which of the following is **true** with regard to simple regression and correlation analysis?

- (a) Correlation analysis provides a measure, which indicates how well an equation for estimating one variable from another, explains the changes in the estimated variable
- (b) If the coefficient of correlation between two variables is close to 1, then there is no correlation between the two variables
- (c) If the slope of a regression equation is positive, then the coefficient of correlation between the variables involved is negative
- (d) If the slope of a regression equation is negative, then the dependent variable decreases as the independent variable decreases
- (e) Correlation analysis necessarily indicates the presence of cause-effect relationship between the variables involved.

(1 mark)

[<Answer>](#)

64. The variable Y in the simple regression model  $Y = A + BX + e$  (where 'e' is the random error), is the

- (a) Intercept on the horizontal axis
- (b) Intercept on the vertical axis
- (c) Slope of the regression line
- (d) Dependent variable
- (e) Independent variable.

(1 mark)

(e) Independent variable.

[<Answer>](#)

65. Pradeep Khanna, a second year MBA student, is doing a study of companies going public for the first time. He is curious to see whether there is a significant relationship between the sizes of the offering (in crores of rupees) and the price per share after the issue. The data are given below:

Size (in crores of rupees)	108.00	39.00	68.40	51.00	10.40	4.40
Price (in Rs.)	12.00	13.00	19.00	12.00	6.50	4.00

Calculate the coefficient of correlation between the issue size and the post-issue price

- (a) 0.6511
- (b) 0.6725
- (c) 0.7533
- (d) 0.8044
- (e) 0.8555.

(3 marks)

[<Answer>](#)

66. From the following data obtain the regression equation of Y on X (independent variable)

X	6	2	10	4	8
Y	9	11	5	8	7

- (a)  $\hat{Y} = 11.9 - 0.65X$
- (b)  $\hat{Y} = 0.65 + 11.9X$
- (c)  $\hat{Y} = -11.9 + 0.65X$
- (d)  $\hat{Y} = 11.9 + 0.65X$
- (e)  $\hat{Y} = 0.65 - 11.9X$ .

(2 marks)

[<Answer>](#)

67. In which of the following conditions will marginal analysis be inappropriate for decision-making?

- (a) When the nature of the decision maker is not known
- (b) When there are a large number of decision alternatives
- (c) When the decision maker cannot assign probabilities of occurrence to the various states of nature
- (d) When the nature of the decision maker is known
- (e) When the decision maker can assign probabilities of occurrence to the various states of nature. (1 mark)

[<Answer>](#)

68. In a decision making problem the expected profit of the optimal decision without perfect information is Rs.40,000 and the expected profit with perfect information is Rs.45,000. What is the expected value of perfect information?

- (a) - Rs. 5,000
- (b) Rs. 5,000
- (c) Rs. 85,000
- (d) Rs. 42,500
- (e) Rs. 2,500.

(1 mark)

[<Answer>](#)

69. A graphical method of representing states of nature and courses of action is referred to as a

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

(1 mark)

[<Answer>](#)

70. The SPSS package enables

- I. Summarization of data.
  - II. Display of data with tables.
  - III. Display of data with graphs.
- (a) Only (I) above
  - (b) Only (II) above
  - (c) Only (III) above
  - (d) Both (II) and (III) above
  - (e) All (I), (II) and (III) above.

(1 mark)

71. In SPSS, the output view displays

- (a) The actual data that you have entered and any new variables created
- (b) The results of the queries like frequency distributions, cross-tabs, statistical tests and charts
- (c) The output as it is generated for printing
- (d) The actual computer code that provides a specific output
- (e) The definition of each variable in the data set.

(1 mark)

72. In SPSS, the draft view displays

- (a) The actual data that you have entered and any new variables created
- (b) The results of the queries like say frequency distributions, cross-tabs and charts
- (c) The output as it is generated for printing
- (d) The actual computer code that provides a specific output
- (e) The definition of each variable in the data set.

(1 mark)

**END OF QUESTION PAPER**

**Suggested Answers**  
**Quantitative Methods – II (MB152): July 2008**

	<b>Answer</b>	<b>Reason</b>	
1.	B	Let E be the event that the total number of ways of drawing 3 black balls out of 3 is ${}^3C_3$ . The total number of ways of drawing 3 balls out of 10 being ${}^{10}C_3$ the required probability is $P(E) = \frac{{}^3C_3}{{}^{10}C_3} = \frac{1}{120}$	≤
2.	B	Here, the sample space is (HH, HT, TH and TT). Let E be the event that getting at least one tail. The number of cases that are favorable is given by 3 while the total number of possible outcomes is 4. $\therefore P(E) = \frac{\text{Favorable Cases}}{\text{Number of possible outcomes}} = \frac{3}{4} = 0.75.$	≤
3.	C	There are six numbers in each of the dice. So total number of ways the dice can roll is $6 \times 6 = 36$ . Following could be the combinations of outcomes of rolling to get a sum of eight: (2, 6), (3, 5), (4, 4), (5, 3), (6, 2) i.e. 5 ways. So, the corresponding probability is $\frac{5}{36}$ .	≤
4.	E	Since, If a fair coin is tossed, coin may lands either with the head up or with the tail up are mutually exclusive events and trials are independent. Therefore, probability of getting either head or tail is remains same trial to trial. $\therefore \text{The probability that the coin will land with the head up on the fourth toss is } \frac{1}{2}$	≤
5.	D	$P(A \text{ and } B) = P(B) \times P(A/B)$ Or $0.30 = 0.60 \times P(A/B)$ $\text{Or } P(A/B) = \frac{0.30}{0.60} = 0.50.$	≤
6.	B	Number boys in a class = 30 Number of girl in a class = 20	≤

Total number of students in a class = 50.

One name is picked up random ,

$$\text{Probability that the name is girl's name} = \frac{{}^{20}C_1}{{}^{50}C_1} = \frac{2}{5}.$$

7. E Let  $E_1$  be the event that A can solve a problem and  $E_2$  the event that B can solve the problem  $\leq$

$$\text{Then } P(E_1) = \frac{2}{3}, P(E_2) = \frac{3}{5}$$

Now, we want the probability of  $P(E_1 \text{ or } E_2)$

$$\begin{aligned} \therefore P(E_1 \text{ or } E_2) &= P(E_1) + P(E_2) - P(E_1 \text{ and } E_2) \\ &= P(E_1) + P(E_2) - P(E_1) \times P(E_2) \end{aligned}$$

[ $\square P(E_1 \text{ and } E_2) = P(E_1) P(E_2)$  since the event  $E_1$  &  $E_2$  are independent]

$$= \frac{2}{3} + \frac{3}{5} - \frac{2}{3} \times \frac{3}{5} = \frac{2}{3} + \frac{3}{5} - \frac{2}{5} = \frac{13}{15}.$$

8. D Let G is the event that the first bulb is good and D is the event that the second bulb if defective.  $\leq$   
As the two events are dependent,

$$P(G \cap D) = P(G) \cdot P(D/G)$$

$$P(G) = \frac{20}{22}, P(D/G) = \frac{2}{21}$$

$$P(G \cap D) = \frac{20}{22} \times \frac{2}{21} = \frac{20}{231}.$$

9. B The joint probability of any two events is estimated by multiplying the probabilities of the  $\leq$   
respective events.

That is  $P_a \times P_b$ .

10. C  $E(Y) = E(X - 25) = E(X) - 25$   $\leq$

$$E(X) = (20 \times 0.10) + (40 \times 0.20) + (60 \times 0.30) + (80 \times 0.25) + (100 \times 0.15) = 63$$

$$\therefore E(Y) = 63 - 25 = 38.$$

11. D Alternatives (a), (b), (c) and (e) are the characteristics of the Bernoulli process. The probability of  $\leq$   
two outcomes may not necessarily be the same.

12. B Since n is said to be very large (in case of words) then the number of mistakes in a page follows  $\leq$   
Poisson distributions.

13. B A normal probability distribution is a continuous probability distribution.  $\leq$

14. A 
$$P(r) = \frac{e^{-\lambda} \lambda^r}{r!}, r = 0, 1, 2, \dots$$
 ≤
- $$\lambda = np = 104 \times \frac{1}{52} = 2$$
- Now, probability of drawing an ace of spades at least once
- $$1 - P(0) = 1 - \frac{e^{-2} 2^0}{0!}$$
- $$1 - e^{-2} = 1 - 0.1353 = 0.8647.$$
15. B The degrees of freedom = sample size - 1. ≤  
Hence sample size = degrees of freedom + 1 = 15 + 1 = 16.
16. A For 20 trials,  $E(Y) = n \times p$  ≤  
 $\Rightarrow 20 \times p = 4$   
 $\Rightarrow p = 0.2.$   
 For 25 trials  $V(Y) = n \times p \times q$   
 $= 25 \times 0.2 \times 0.8$   
 $= 4.$   
 Therefore for 25 trials standard deviation of Y  
 $= \sqrt{V(Y)}$   
 $= \sqrt{4}$   
 $= 2.$
17. A The most appropriate example of a continuous random variable is the weight of a randomly selected human being. ≤
18. E Probability that the phone is busy is  $p = 1/15$  ≤  
 Probability that the phone is not busy is  $q = 14/15$   
 Probability that out of six randomly selected telephone numbers, at most three numbers are busy  
 $P[X \leq 3] = P(0) + P(1) + P(2) + P(3)$   

$$= {}^6C_0 \left(\frac{14}{15}\right)^6 + {}^6C_1 \left(\frac{1}{15}\right) \left(\frac{14}{15}\right)^5 + {}^6C_2 \left(\frac{1}{15}\right)^2 \left(\frac{14}{15}\right)^4 + {}^6C_3 \left(\frac{1}{15}\right)^3 \left(\frac{14}{15}\right)^3$$
  

$$= \left(\frac{1}{15}\right)^6 [(14)^6 + 6(14)^5 + 15(14)^4 + 20(14)^3]$$
  

$$= \frac{11387600}{11390625} = 0.9997.$$
19. D Given that N (total number of taxi cabs) = 10, ≤  
 M (number of Ambassador cars) = 6,  
 n (sample size) = 3,  
 x (number of Ambassadors) = 2

$$P(X = x) = \frac{{}^M C_x \cdot {}^{(N-M)} C_{(n-x)}}{{}^N C_n}$$

$$= \frac{{}^6 C_2 \cdot {}^{(10-6)} C_{(3-2)}}{{}^{10} C_3} = \frac{15 \times 4}{120} = \frac{1}{2}$$

20. C When two fair coins are tossed the probability distribution of the number of heads will be as follows: ≤

Number of heads	Probability
0	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
1	$\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
2	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$$E(\text{Number of heads}) = \left(0 \times \frac{1}{4}\right) + \left(1 \times \frac{1}{2}\right) + \left(2 \times \frac{1}{4}\right) = 1.$$

21. C  $\bar{X} = 31$  ≤  
 $s = 6.84$   
 $n = 60$      $N = 540$

$$\text{Sampling fraction, } \frac{n}{N} = \frac{60}{540} = 0.11 > 0.05$$

$$\therefore \text{Estimated standard error of mean, } \hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$= \frac{6.84}{\sqrt{60}} \sqrt{\frac{540-60}{540-1}}$$

$$= \frac{6.84}{\sqrt{60}} \times \sqrt{\frac{480}{539}}$$

$$= 0.8333$$

Since the sample size is greater than 30, normal distribution will be used.

Z-values for the upper and lower confidence limit are  $\pm 1.96$

$$\therefore \text{Upper confidence limit} = \bar{X} + 1.96 \hat{\sigma}_{\bar{x}} = 31 + (1.96 \times 0.8333) = 32.63 \text{ years}$$

$$\text{Lower confidence limit} = \bar{X} - 1.96 \hat{\sigma}_{\bar{x}} = 31 - (1.96 \times 0.8333) = 29.37 \text{ years}$$

22. E Given option (e) is **false** regarding business research. ≤

Since, the following true regarding to the business research.

- It can be defined as the plan and structure of enquiry, formulated in order to obtain answers to research questions on business aspects.
- The research plan constitutes the overall programme of the business research process.
- It can be understood as that which gives the blueprint for collection, measurement and analysis of business data.
- It helps researchers to utilize available resources efficiently to achieve research objectives.

Hence from above discussion option (e) is correct.

23. E Given option (e) is **false** regarding variables. ≤

Since, Extraneous variables are **Independent** variables.

The following statements are **true** regarding variables.

- Continuous variables can take on different values even in decimal places



- (b) Non-continuous variables are those that cannot be measured in decimals
- (c) If a variable is dependent on the result of some other variable, it is then called a dependent variable
- (d) An independent variable is one that is not dependent on any other variable with reference to that particular study
- Hence option (e) is correct.
24. A Secondary data is already available as a result of research carried out by others ≤  
Options,(b), (c), (d) and (e) are false. Since correct statements is  
**Primary data** is that the data which is collected first time on subject  
Using **Primary data** we can give correct conclusions to current situation  
Secondary data **can** be obtained from journals, magazines and periodicals  
And to collect secondary data it takes **less** time and **less** money.  
Hence option(a) is correct.
25. E The following are steps in a sampling process. ≤  
I. Defining the Target Population.  
II. Specifying the Sampling Frame.  
III. Selection of the Sampling Method.  
IV. Specifying the Sampling Plan.  
V. Selecting the Sample.  
Hence option (e) is correct.
26. A Control is the important characteristic of a good research design, which is essentially devised to ≤  
minimize the effects of extraneous variables.
27. C The key objective of any business research is to provide accurate, relevant and timely ≤  
information to the top management.
28. D The following statements are true regarding Evaluating the Data. ≤  
I. Several mathematical and statistical models are used to evaluate the data.  
II. Evaluation of the data normally starts with editing and coding of the data.  
Hence (d) is the correct option.
29. E The three major forms of business research are: ≤  
Market Research, Operations Research and Motivational Research.  
Hence option (e) is correct.
30. E A typical research report has the following sections: ≤  
I. Prefatory information.  
II. Methodology.  
III. Introduction.  
IV. Findings.  
V. Conclusions and Recommendations.  
Hence (e) is the correct option.

31. B It means that good research results can be obtained with minimum utilization of money, time and effort. ≤
32. E The following are the strengths of qualitative research. ≤  
 I. Ideal exploratory tool.  
 II. Increase the value of subsequent quantitative research.  
 III. Group dynamics enhance results.  
 Hence (e) is the correct option.
33. E The following are components of attitude. ≤  
 I. Cognitive Component (Beliefs).  
 II. Affective Component (Feelings).  
 III. Behavioural Component (Actions).  
 Hence (e) is the correct option.
34. A Stratified sampling on proper application can significantly increase the statistical efficiency of sampling. ≤
35. E The following are the single item scales which measure the attitudes of respondents. ≤  
 I. Itemized category scales.  
 II. Rank order scales.  
 III. Q-sort scales.  
 IV. Comparative scales.  
 Hence (e) is the correct option.
36. E The following statements are **true** regarding Sequential sampling. ≤  
 I. In this method the size of the sample is not fixed in advance.  
 II. This sampling method is used for statistical quality control.  
 III. A number of sample lots are drawn in sequence one after another from the population depending on the results of the earlier sample.  
 Hence (e) is the correct option.
37. B The sampling frame error is said to occur when the list of members does not correspond exactly with the target population. ≤
38. B In the stratified sampling the population is divided into groups such that elements in each group are homogeneous and the groups vary from each other. ≤
39. A According to the Central Limit Theorem for large samples the sample mean approaches the normal distribution. ≤

40. C Sampling is the appropriate technique in case of destructive testing. All other statements are true with regard to sampling. ≤

41. E The following are the causes for error in the sampling. ≤

- I. Error in selection of the sample.
- II. Bias in the reporting of data.
- III. Diversity of the population.
- IV. Substitution of sampling units for convenience.
- V. Faulty demarcation of sampling universe.

Hence (e) is the correct option.

42. C The statistical measures like mean, median, mode and standard deviation which are used to describe the characteristics of a sample are known as parameters. ≤

43. D In judgment sampling the sample is selected according to the judgment of the investigator. Hence some subjectivity creeps into the sampling process and the truly random nature of the sample is lost to some extent. ≤

44. D Given that  $\mu = 1,600$ ,  $\bar{X} = 1,580$ ,  $s = 90$ ,  $n = 100$ . ≤

$$\text{The estimated standard error of mean} = \frac{s}{\sqrt{n}} = \frac{90}{\sqrt{100}} = 9.$$

$$\text{The test statistic for this test is, } z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{1580 - 1600}{9} = -2.2222.$$

45. D Given that  $\mu_0 = 54.50$  runs  $\sigma = 5.5$  runs ≤

The average runs scored this season can be taken as the mean of the sample,

Therefore  $\bar{X} = 48.80$  runs,  $n = 25$

$$\text{Standard error of mean, } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5.5}{\sqrt{25}} = 1.10$$

The standardized value of the test statistic

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{x}}} = \frac{48.80 - 54.50}{1.10} = -5.1818.$$

46. B Here we set up the null hypothesis,  $H_0$  that the number of products demanded does not depend on the months of the year. ≤

Under the null hypothesis, the expected frequencies of the products demanded on each of the six days would be:

$$\frac{1}{6}(1032 + 1026 + 1020 + 1019 + 1023 + 1012) = \frac{6132}{6} = 1022$$

This is a chi-square test of independence and the contingency table is given as:

month	Frequency Observed ( $f_i$ )	Frequency Expected ( $e_i$ )	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
Jan	1032	1022	100	0.0978
March	1026	1022	16	0.0157
May	1020	1022	4	0.0039
July.	1019	1022	9	0.0088
Sep	1023	1022	1	0.0010
Nov.	1012	1022	100	0.0978
Total	6132	6132		0.2250

Then the value of chi-square statistic is given by:

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = 0.2250.$$

i.e.

47. B The test is chi-square goodness of fit. ≤  
 Since the number of elements in the table is 12, the number of degrees of freedom for the chi square test is  $12 - 1 = 11$ . The value of chi-square for 11 degrees of freedom and 5% level of significance from the table is 19.675.

48. D The following statements are **true** regarding to Lognormal distribution. ≤  
 I. If  $\ln(X)$  is a normally distributed random variable, then  $X$  is said to be a lognormal variable.  
 III. In finance “Continuous compounded returns are required to be normal”. This property is described as “Stock Prices are Lognormal”.  
 So option (d) is correct.

49. E By Central Limit Theorem for large samples the sample mean is approximately normally ≤  
 distributed with mean = population mean and standard deviation  $= \frac{\sigma}{\sqrt{n}}$

$$\therefore E(\bar{x}) = \mu = 20$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{100}{64}} = \frac{10}{8} = 1.25$$

$$\begin{aligned} \therefore P(18 < \bar{x} < 23) &= P\left(\frac{18-20}{1.25} < z < \frac{23-20}{1.25}\right) = P(-1.60 < z < 2.40) \\ &= P(z < 2.40) - P(z < -1.60) = P(z < 2.40) - P(z > 1.60) \\ &= P(z < 2.40) - [1 - P(z < 1.60)] \\ &= (0.50 + 0.4918) - 1 + (0.50 + 0.4452) = 0.937. \end{aligned}$$

50. E Given that  $\bar{X}_1 = 1300, s_1 = 82, n_1 = 100$ . ≤

$$\bar{X}_2 = 1288, s_2 = 93, n_2 = 100$$

The estimated standard error of the difference between means is,

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{82^2}{100} + \frac{93^2}{100}} = 12.3988.$$

51. A  $H_0: \mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$  ≤

$$H_1: \mu_1 > \mu_2 \Rightarrow \mu_1 - \mu_2 > 0$$

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 =

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\left(\frac{640}{64} - \frac{144}{36}\right) - 0}{\sqrt{\frac{256}{64} + \frac{144}{36}}} = \frac{6}{2.8284} = 2.1213$$

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

52. D The following statements are **true** regarding advantages of sampling. ≤

- I. A study of a sample is usually cheaper than a study of the population.
- II. Sampling usually provides information quicker than a census so that timely decisions can be taken.

And in sampling the chances of errors while processing the data are **less**.

So option (d) is correct.

53. E  $H_0 : \mu = 9$  ≤

$$H_1 : \mu < 9$$

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  $25 - 1 = 24$  degrees of freedom.

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

$$t = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{3.60 - 9}{3.20} = -1.6875$$

At  $\alpha = 0.05$ , the critical value is  $-1.711$ .

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

∴ We accept  $H_0$ . The true mean is 7. So  $H_0$  is false. Hence the test leads to a type II error.

54. D Let  $X =$  life of a regular savings account. Then we are given that  $X$  follows Normal distribution with mean  $\mu = 15$  and Standard deviation  $\sigma = 6.45$ . ≤

$$\begin{aligned} P(X > 19) &= P\left(\frac{X - \mu}{\sigma} > \frac{19 - 15}{6.45}\right) \\ &= P(Z > 0.62), \text{ Where } Z \text{ is a standard normal variable.} \\ &= 0.5 - \phi(0.62) = 0.5 - 0.2324 = 0.2676. \end{aligned}$$

55. C Given statement (III) is true. ≤

III. The interval estimate indicates the accuracy (or precision) of an estimate.

If an estimate of a population parameter is given by a single value, then the estimate is called a **Point** estimate. And if an estimate of a population parameter is given by two distinct numbers between which the parameter may be considered to lie, then the estimate is called an **interval** estimate.

So option (c) is correct.

56. D **t**-statistic is given by ≤

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ t &= \frac{51 - 50}{1/\sqrt{16}} = \frac{1}{0.25} = 4. \end{aligned}$$

57. D The following statements are **true** regarding Pearson's coefficient of correlation. ≤

- II.  $\rho = 0 \Leftrightarrow$  No relationship between  $X$  and  $Y$ .

III.  $\rho = 1 \Leftrightarrow$  Perfect positive linear relationship between X and Y.

IV.  $\rho = -1 \Leftrightarrow$  Perfect negative linear relationship between X and Y.

And

Pearson's coefficient of correlation lies between  $-1 \leq \rho \leq 1$

So option (d) is correct.

58. C

Let rank of student in Statistics is  $X_i$

Let rank of student in Mathematics is  $Y_i$

Then  $D_i = X_i - Y_i$

n	Statistics	Mathematics	Difference between ranks ( $D_i$ )	$D_i^2$
1	2	3	-1	1
2	3	4	-1	1
3	5	5	0	0
4	4	2	2	4
5	1	1	0	0
Total			0	6

Here  $\sum D_i^2 = 6$  and  $n=5$ .

$$\begin{aligned}
 R &= 1 - \frac{6 \sum D_i^2}{n^3 - n} \\
 &= 1 - \frac{6 \times 6}{5^3 - 5} \\
 &= 1 - \frac{36}{120} \\
 &= 0.7.
 \end{aligned}$$

59. C

The equation of simple regression line is  $\hat{Y} = a + bX$ .

Student	Y	$Y^2$	X	$X \times Y$	$X^2$
101	89	7921	85	7565	7225
203	80	6400	76	6080	5776
357	78	6084	68	5304	4624
465	70	4900	80	5600	6400
509	92	8464	90	8280	8100
Total	409	33769	399	32829	32125

The coefficient b is

$$\begin{aligned}
 b &= \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} \\
 &= \frac{5(32829) - (399)(409)}{5(32125) - (399)^2} \\
 &= \frac{954}{1424} = 0.6699.
 \end{aligned}$$

The coefficient a is

$$\begin{aligned}
 a &= \bar{Y} - b\bar{X} \\
 &= \left(\frac{409}{5}\right) - (0.6699)\left(\frac{399}{5}\right) \\
 &= 28.3420.
 \end{aligned}$$

Therefore, the equation for the regression line is  $\hat{Y} = 28.3420 + 0.6699X$ .

When model exam score (X) is 95,

$$\hat{Y} = 28.3420 + 0.6699(95) = 91.9825 \approx 92.$$

60. B ≤

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

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

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

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

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

61. C If  $RSS = 0$ , then  $R^2 = 0$ . ≤

62. A Coefficient of determination =  $r^2$  (where 'r' is the coefficient of correlation and  $-1 \leq r \leq 1$ ) ≤  
 $\therefore$  Coefficient of determination =  $r^2 \geq 0$

Because square of a positive or, negative number is always positive.

Further if  $r = 0$  then  $r^2 = 0$ , and if  $r \neq 0$  then  $r^2 \neq 0$

Also if  $-1 \leq r \leq 1$  then  $0 \leq r^2 \leq 1$

Hence all other alternatives except (a) are false.

63. A
- a. Correlation analysis provides a measure (coefficient of determination) which indicates, how well an estimating equation explains the changes in the estimated (i.e. dependent variable). ≤
  - b. If the coefficient of correlation between two variables is close to 1 then there is a very high correlation between the two variables.
  - c. If the slope of a regression line is positive then, the coefficient of correlation between the variables involved is positive.
  - d. If the slope of a regression line is negative then the dependent variable decreases as the independent variable increases.
  - e. Correlation analysis may indicate a possibility of a cause-effect relationship. It does not necessarily indicate that there is a cause-effect relationship between the variables involved.

64. D Y is the dependent variable. ≤

65. B To compute the coefficient of correlation between the two variables we tabulate them as below: ≤

Size	Price	$x = X - \bar{X}$	$y = Y - \bar{Y}$	$x^2$	$y^2$	xy
108	12	61.13	0.92	3736.8769	0.8464	56.2396
39	13	-7.87	1.92	61.9369	3.6864	-15.1104

39	13	-7.87	1.92	61.9369	3.6864	-15.1104
68.4	19	21.53	7.92	463.5409	62.7264	170.5176
51.0	12	4.13	0.92	17.0569	0.8464	3.7996
10.4	6.5	-36.47	-4.58	1330.0609	20.9764	167.0326
4.40	4.0	-42.47	-7.08	1803.7009	50.1264	300.6876
Total				7413.1734	139.2084	683.1666

The mean values

$$\bar{X} = \frac{\sum X}{n} = \frac{281.2}{6} = 46.87$$

$$\bar{Y} = \frac{\sum Y}{n} = \frac{66.5}{6} = 11.08$$

We know that the coefficient of correlation

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

$$\therefore r = \frac{683.1666}{\sqrt{7413.1734 \times 139.2084}} = 0.6725.$$

66. A We obtain the regression equation using the following table: ≤

X	Y	XY	X <sup>2</sup>
6	9	54	36
2	11	22	4
10	5	50	100
4	8	32	16
8	7	56	64
$\sum X = 30$	$\sum Y = 40$	$\sum XY = 214$	$\sum X^2 = 220$

The regression equation of Y on X is  $\hat{Y} = a + bX$

To determine the values of a and b the following two normal equations are to be solved:

$$\sum Y = na + b \sum X$$

$$\sum XY = a \sum X + b \sum X^2$$

Substituting the values, we get :

$$40 = 5a + 30b$$

$$214 = 30a + 220b$$

The values of a = 11.9 and b = -0.65

$$\hat{Y} = 11.9 - 0.65X.$$

67. C Marginal analysis will be inappropriate for decision-making when the decision maker cannot assign probabilities of occurrence to the various states of nature. ≤

68. B  $EVPI = \text{Expected profit with perfect information} - \text{Expected profit without perfect information}$  ≤  
 $= 45,000 - 40,000$   
 $= \text{Rs. } 5,000.$

69. A
- 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.

- 70.** E The SPSS package enables: ≤  
(I) Summarization of data.  
(II) Display of data with tables.  
(III) Display of data with graphs.  
Hence option (E) is correct.
- 71.** B In SPSS, the Output View displays: ≤  
The results of the queries like frequency distributions, cross-tabs, statistical tests and charts.
- 72.** C In SPSS, the Draft View displays: ≤  
The output as it is generated for printing.

[< TOP OF THE DOCUMENT >](#)