

Question Paper

Quantitative Methods – II (132) – July 2004

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

1. When the chi-square test is used as a test of independence, the number of degrees of freedom is determined by [< Answer >](#)
- (a) The sample size
 - (b) The ratio of sample size and the population
 - (c) Number of rows in the contingency table, only
 - (d) Number of columns in the contingency table, only
 - (e) Both the number of rows and the number of columns in the contingency table.
- (1 mark)
2. In a control chart there are some observations that lie outside the upper and lower control limits. These points are referred to as the [< Answer >](#)
- (a) Outliers
 - (b) Attribute
 - (c) Inherent variation
 - (d) Assignable variation
 - (e) Common variation.
- (1 mark)
3. The weighted average of price relatives using base values as weights is same as the [< Answer >](#)
- (a) Unweighted aggregates price index
 - (b) Unweighted aggregates quantity index
 - (c) Laspeyres price index
 - (d) Laspeyres quantity index
 - (e) Paasche's price index.
- (1 mark)
4. The standard error of estimate of a simple regression line is [< Answer >](#)
- (a) Equal to the slope of the regression line
 - (b) Equal to the square root of the coefficient of determination
 - (c) Equal to the Y intercept of the regression line
 - (d) The measure of the variability of the observed values around the regression line
 - (e) Equal to the coefficient of correlation between two variables.
- (1 mark)
5. Which of the following is **not true** for estimation? [< Answer >](#)
- (a) The probability that we associate with an interval estimate is called the confidence interval
 - (b) If the population standard deviation is high, we can go for large confidence interval
 - (c) The number of degrees of freedom used in a t-distribution is equal to the sample size
 - (d) The t distribution need not be used in estimating if you know the standard deviation of the population
 - (e) The confidence interval is the range of the estimate we are making.
- (1 mark)
6. In many situations managers resort to sampling to draw some conclusions about a population. Which of the following is **not** an advantage of sampling over a census? [< Answer >](#)
- (a) The population may be too large to be studied in full, hence sampling is the only feasible means
 - (b) A study of sample is usually cheaper than a census

- (c) Sampling usually provides information quicker than a census
- (d) In destructive testing sampling is the only available course
- (e) The conclusions obtained from sampling are more accurate than census.

(1 mark)

7. If the covariance of two random variables X and Y is equal to zero then it can be said that

[< Answer >](#)

- (a) The random variables are dependent on each other
- (b) The random variables have a positive correlation
- (c) The random variables have a negative correlation
- (d) The random variables are independent of each other
- (e) The variances of the two random variables are always zero.

(1 mark)

8. Which of the following is **not true** for probability distributions?

[< Answer >](#)

- (a) The past frequency distribution may be used as a guide for predicting the future values of the random variable
- (b) The probability distribution gives the chances of occurrence of the various values of the random variable
- (c) A subjective probability distribution of the likely future values can be formed on the basis of expert opinion
- (d) We can never assign probabilities to all likely values of the random variable on the basis of the past data
- (e) The sum of the probability of occurrence of all possible values must be equal to one.

(1 mark)

9. Which of the following is **true** with regard to Fisher's ideal price index?

[< Answer >](#)

- (a) It does not consider the base year prices
- (b) It does not consider the base year quantities
- (c) It does not consider the current year prices
- (d) It does not consider the current quantities
- (e) It is the geometric mean of the Laspeyre's and Paasche's price indices.

(1 mark)

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

[< Answer >](#)

- (a) Scatter diagram
- (b) Frequency polygon
- (c) Histogram
- (d) p chart
- (e) Ogive.

(1 mark)

11. In a test involving ANOVA the **F** statistic is calculated on the basis of

[< Answer >](#)

- (a) The mean of the largest sample
- (b) The mean of the smallest sample
- (c) The standard deviation of the smallest sample
- (d) The estimated population variance based on the variance among the sample means and the estimated population variance based on the variance within the samples
- (e) The variance of the largest sample.

(1 mark)

12. 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)

13. According to the factor reversal test if we interchange the price and quantity terms in an index number then the product of the original index number and the index number obtained by reversing the factors should be equal to [< Answer >](#)

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

(1 mark)

14. Which of the following is **false** for the multiple correlation coefficient between dependent variable Y and two independent variables X_1 and X_2 ? [< Answer >](#)

(a) It depends on the correlation coefficient between Y and X_1
(b) It depends on the correlation coefficient between X_1 and X_2
(c) It depends on the correlation coefficient between Y and X_2
(d) It will take values between 0 and 1
(e) It will take values between -1 and $+1$.

(1 mark)

15. If in a test of statistical hypothesis, α is the level of significance and β is the probability of committing Type II error then which of the following is **not true** for the hypothesis test? [< Answer >](#)

(a) The probability that the null hypothesis is true but rejected is α
(b) The probability that the null hypothesis is false but accepted is β
(c) $(1-\beta)$ is known as the power of test
(d) A high $(1-\beta)$ value implies that the test is poor
(e) There is a trade-off between the Type I and Type II error.

(1 mark)

16. 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)

17. Which of the following refers to the random (chance) occurrences that can affect the outcome of an individual's decision? [< Answer >](#)

(a) Payoffs (b) States of nature
(c) Decision alternatives (d) Probabilities (e) None of the above.

(1 mark)

18. 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)

19. Which of the following is **false** with regard to hypothesis testing? [< Answer >](#)
- (a) The purpose of hypothesis testing is to find out whether there is a significant difference between the sample statistic and the hypothesized population parameter
 - (b) Rejection of null hypothesis implies the acceptance of alternative hypothesis
 - (c) The acceptance of null hypothesis proves that it is true
 - (d) The rejection of null hypothesis does not prove that it is false
 - (e) None of the above.
- (1 mark)**
20. For samples of size greater than 1, the finite population correction factor [< Answer >](#)
- (a) Is always equal to 1
 - (b) Is always more than 1
 - (c) Is always constant irrespective of the sample size
 - (d) Is always equal to 0
 - (e) Is always less than 1.
- (1 mark)**
21. In which of the following conditions the finite population correction factor **need not** be used for calculating the standard error of a statistic? [< Answer >](#)
- I. The population is infinite.
 - II. The population is finite and sampling is done without replacement.
 - III. The population is finite and sampling is done with replacement.
- (a) (I) only
 - (b) (II) only
 - (c) (III) only
 - (d) Both (I) and (II)
 - (e) Both (I) and (III).
- (1 mark)**
22. 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)**
23. In a binomial distribution the probability of getting zero or more number of successes is equal to [< Answer >](#)
- (a) 0
 - (b) The probability of getting zero success
 - (c) The probability of getting successes in all the trials
 - (d) 1 minus the probability of getting successes in all the trials
 - (e) 1.
- (1 mark)**
24. The sampling distribution of mean is [< Answer >](#)
- (a) A distribution of all possible sample means for samples of a given size taken from the same population
 - (b) A distribution of all possible sample means for samples of a given size taken from different populations
 - (c) A distribution of all possible sample means for samples of different sizes taken from the same population
 - (d) A distribution of all possible sample means for samples of different sizes taken from different populations
 - (e) A distribution of the observations in a sample.
- (1 mark)**
25. In a right tailed test of hypothesis the [< Answer >](#)

- (a) Acceptance region lies under the right tail
- (b) Rejection region lies under the right tail
- (c) Acceptance region lies under the right and left tails
- (d) Rejection region lies at the center of the distribution
- (e) Entire region lying on the right of the mean of the distribution is the rejection region.

(1 mark)

26. The standard error of estimate is developed from the variations of the

[< Answer >](#)

- (a) Observed values of the dependent variable around a fixed quantity
- (b) Observed values of the dependent variable around the mean of the observed values of the independent variable
- (c) Observed values of the dependent variable around the fitted regression line
- (d) Observed values of the independent variable around their mean
- (e) Observed values of the independent variable around the mean of the observed values of the dependent variable.

(1 mark)

27. If the dependent variable decreases as the independent variable increases in an estimating equation then, the coefficient of correlation will be in the range of

[< Answer >](#)

- (a) 0 to -2.00
- (b) 0 to -1.00
- (c) 0 to 0.50
- (d) 0 to 0.25
- (e) 0 to 1.00.

(1 mark)

28. Which of the following represents the proportion of variation in the dependent variable that is explained by the regression line?

[< Answer >](#)

- (a) Coefficient of determination
- (b) Coefficient of correlation
- (c) Coefficient of variation
- (d) Standard error of estimate
- (e) None of the above.

(1 mark)

29. Weighted average of relatives price index for various years can be readily compared if

[< Answer >](#)

- (a) Base year values are used as weights
- (b) Current year values are used as weights
- (c) Base year prices are used as weights
- (d) Current year prices are used as weights
- (e) Sum of base year and current year quantities are used as weights.

(1 mark)

30. If the regression equation is a perfect estimator of the dependent variable then which of the following is **false**?

[< Answer >](#)

- (a) The standard error of estimate is zero
- (b) The coefficient of correlation is zero
- (c) The coefficient of determination is 1.00
- (d) All the data points fall on the regression line
- (e) None of the above.

(1 mark)

31. Which of the following techniques enables us to test whether means of more than two populations are equal?

[< Answer >](#)

- (a) Analysis of variance
- (b) Chi-square test
- (c) Decision tree analysis
- (d) Marginal analysis
- (e) Correlation analysis.

(1 mark)

32. If a random variable **X** has a lognormal distribution then which of the following statements is **true**? [< Answer >](#)
- (a) **X** is normally distributed
 - (b) $\ln(\mathbf{X})$ is normally distributed
 - (c) The expected value of **X** is zero
 - (d) The expected value of **X** is 1
 - (e) None of the above.

(1 mark)

33. 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)

34. Which of the following is an example of a continuous random variable? [< Answer >](#)
- (a) The volume of liquid present in a container
 - (b) The number of patients admitted to a hospital in a day
 - (c) The number of defective units of a product in a batch of production
 - (d) The number of spectators in a cinema hall
 - (e) The number of professors in a college.

(1 mark)

35. Which of the following best describes the expected value of a discrete random variable? [< Answer >](#)
- (a) It is the geometric average of all possible outcomes of the variable
 - (b) It is the simple average of all possible outcomes of the variable
 - (c) It is the weighted average of all possible outcomes of the variable
 - (d) It is the outcome of the variable, which has the highest probability of occurrence
 - (e) It is the highest probability of occurrence in probability distribution of the random variable.

(1 mark)

36. According to the Central Limit Theorem, the sampling distribution of mean can be approximated by the normal distribution [< Answer >](#)
- (a) As the number of samples decreases
 - (b) As the sample size increases
 - (c) As the population standard deviation increases
 - (d) As the sample standard deviation decreases
 - (e) Only if the population distribution is normal.

(1 mark)

37. As the sample size increases [< Answer >](#)
- (a) The variation of the sample mean from the population mean becomes larger
 - (b) The variation of the sample mean from the population mean becomes smaller
 - (c) The variance of the sample becomes less than the variance of the population
 - (d) The standard deviation of the sample becomes less than the standard deviation of the population
 - (e) The standard deviation of the sample comes close to zero.

(1 mark)

38. 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)

39. Which of the following is **true** with regard to the **p** chart in quality control?

[< Answer >](#)

- (a) The center line is drawn on the basis of overall sample mean
- (b) The lower control limit can only be greater than or equal to zero
- (c) The upper control limit is equal to the overall sample proportion
- (d) The lower control limit is equal to the overall sample proportion
- (e) The upper and the lower control limits are based on the overall sample mean.

(1 mark)

40. The t distribution

[< Answer >](#)

- (a) Is not a symmetrical distribution
- (b) Is a discrete probability distribution
- (c) Is a multimodal distribution
- (d) Is different for different number of degrees of freedom
- (e) Is positively skewed.

(1 mark)

41. The following details are available with regard to a simple regression relationship:

[< Answer >](#)

Total sum of squares = 16730

Error sum of squares = 1630

What percentage of the variations in the dependent variable is explained by the regression relationship?

- (a) 0 (b) 9.75% (c) 5.00% (d) 90.25% (e) 95.00%.

(1 mark)

42. The following details are available with regard to a regression relationship:

[< Answer >](#)

Sum of the observations of the independent variable = 150

Sum of the observations of the dependent variable = 455

Number of data points = 10

Slope of the regression line = 2.5

What is the estimated value of the dependent variable if the independent variable is equal to 8?

- (a) 2.5 (b) 8.0 (c) 25.0 (d) 28.0 (e) 36.0.

(1 mark)

43. If the sample size is 50 and the proportion of defective pieces in a batch of output is 0.08 then, the upper control limit of **p** chart will be

[< Answer >](#)

- (a) 0.042 (b) 0.038 (c) 0.08 (d) 0.115 (e) 0.195.

(1 mark)

44. In a test of ANOVA, the following are found:

[< Answer >](#)

Estimated population variance based on the variance among sample means = 24.60

F ratio = 1.20

What is the estimated population variance based on the variance within the samples?

- (a) 17.08 (b) 20.50 (c) 25.80 (d) 23.40 (e) 29.52.

(1 mark)

45. A multiple regression relationship contains two independent variables. The standard error of estimate is 5.70. $\sum(Y - \hat{Y})^2 = 519.84$. The number of observations is

(a) 18 (b) 14 (c) 16 (d) 19 (e) 20.

(1 mark)

46. A variable X, which has a normal distribution, has a mean of 26 and a standard deviation of 2.5. The probability that the value of X lies in the interval 22 to 30 is approximately

(a) 28.81% (b) 31.59% (c) 57.60% (d) 44.52% (e) 89.04%.

(1 mark)

47. The following index values pertain to a year:

Fisher's ideal price index = 125.50
Laspeyres price index = 122.50

Paasche's price index for the year is approximately

(a) 102.45 (b) 115.50 (c) 120.50 (d) 128.57 (e) 153.74.

(1 mark)

48. The probability of failure in a Bernoulli experiment is 0.40. The experiment is repeated 20 times. The mean of the binomial distribution of the number of successes is

(a) 12 (b) 8 (c) 4.8 (d) 6 (e) 2.4.

(1 mark)

49. The linear trend estimating equation for a time series is given below:

$\hat{Y} = 120 + 2.5x$
where $x = \text{Year} - 2000$

$\hat{Y}$ is the estimated value of the dependent variable Y.
The observed value of Y for the year 2001 is 130.

What is the relative cyclical residual for the year 2001?

(a) -4.25 (b) 93.88 (c) 6.12 (d) 95.75 (e) 106.12.

(1 mark)

50. The following details are available for a test involving ANOVA:

Number of samples = 4
Size of the first sample = 6
Size of the second sample = 6
Size of the third sample = 7
Size of the fourth sample = 5

The number of degrees of freedom in the numerator and denominator of the F ratio are

(a) 4 and 24 respectively (b) 4 and 4 respectively
(c) 20 and 3 respectively (d) 3 and 20 respectively
(e) 4 and 20 respectively.

(1 mark)

51. The probability distributions of independent random variables, X and Y, are given below:

[< Answer >](#)

[< Answer >](#)

[< Answer >](#)

[< Answer >](#)

[< Answer >](#)

[< Answer >](#)

[< Answer >](#)

X	12	17	22	27
Probability	0.20	0.40	0.30	0.10

Y	11	15	19
Probability	0.25	0.55	0.20

If a random variable, **Z**, is defined as: $Z = 4X + 5Y$
What is the expected value of the random variable **Z**?

- (a) 14.5 (b) 18.5 (c) 33 (d) 66 (e) 148.

(2 marks)

52. The latest nationwide political poll indicates that for persons, who are randomly selected, the probability that they support **Party A** is 0.55, the probability that they support **Party B** is 0.30, and the probability that they support **Party C** is 0.15. It is assumed that these probabilities are accurate and that the randomly selected persons have independent opinions towards the different parties.

[< Answer >](#)

What is the probability that out of a randomly chosen group of ten persons, less than eight persons do not support **Party A**?

- (a) 0.02739 (b) 0.02289 (c) 0.0045 (d) 0.99618 (e) 0.97261.

(2 marks)

53. The Traffic Police Department of Kolkata has gathered data on the number of traffic accidents and number of cricket matches that occur in the city over a weekend. According to them, there is a close relationship between the number of accidents and the number of cricket matches played during weekends.

[< Answer >](#)

Number of cricket matches	18	28	8	10	13	23	32
Number of accidents	5	8	3	4	6	7	8

Assuming that there is a linear relation between these two variables, what would be your prediction about the number of accidents that will occur on a weekend during which 25 cricket matches take place in Kolkata?

- (a) 6.13 (b) 7.08 (c) 8.07 (d) 9.45 (e) 10.52.

(2 marks)

54. The projected returns on the equity shares of ABC Ltd. and XYZ Ltd. during the next year along with their respective probabilities, are given below:

[< Answer >](#)

Market Condition	Probability	Return (%)	
		ABC Ltd.	XYZ Ltd.
Bullish	10 %	40	30
Stable	50 %	30	20
Bearish	40 %	10	5

What is the covariance of returns?

- (a) 23 %² (b) 95 %² (c) 15 %² (d) 529 %² (e) 225 %².

(2 marks)

55. The auditor of a publishing company, has reviewed the inventory records to determine whether the current inventory holdings of books are typical. The following inventory amounts are from the previous five years:

[< Answer >](#)

Year	1999	2000	2001	2002	2003
Inventory (in thousand rupees)	4,720	5,010	5,590	5,830	6,090

If the time variable(X) has been coded as $x = X - \bar{X}$, then which of the following is the linear trend equation for estimating the inventory level for a given year?

- (a) $\hat{Y} = 4,432 + 356x$ (b) $\hat{Y} = 5,448 + 356x$ (c)
 $\hat{Y} = 54,48,000 + 356x$
 (d) $\hat{Y} = 356 + 5448X$ (e) $\hat{Y} = 3,56,000 + 5448X$

(2 marks)

56. 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 95 percent prediction interval for the value of **Y** if **X** = 8?

- (a) $1,236 \pm 404$ (b) 404 ± 46.49 (c) $1,236 \pm 104$ (d) 404 ± 104 (e) $1,236 \pm 46.49$.

(2 marks)

57. The following details are available with regard to a regression analysis between two variables, **X** and **Y**, with **X** as the independent variable:

[< Answer >](#)

Percentage of variations in **Y** that is explained by the variations in **X** = 84%.

$$\Sigma(Y - \bar{Y})^2 = 1369$$

Number of observations = 6

What is the standard error of estimate for the given regression relationship?

- (a) 43.81 (b) 6.62 (c) 14.80 (d) 54.76 (e) 7.40.

(2 marks)

58. For a sample randomly collected from a population, the following details are available:

[< Answer >](#)

Sum of the squares of the observations = 4710.40

Sum of the observations = 256

Number of observations = 16

What is the estimated standard error of mean?

- (a) 1.60 (b) 0.40 (c) 2.56 (d) 614.40 (e) 40.96.

(1 mark)

59. The standard deviation of the number of successes in a binomial distribution is 2 and the probability of success on any trial is 1/3.

[< Answer >](#)

What is the probability of obtaining **exactly** ten successes?

- (a) 0.0289 (b) 0.1157 (c) 0.3300 (d) 0.8843 (e) 0.9711.

(1 mark)

60. 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** = 6

Standard deviation of **Y** = 12

The slope of the regression equation is

- (a) 0.80 (b) 1.60 (c) 4.80 (d) 9.60 (e) 2.00.

(1 mark)

61. A retailer has received a consignment of 160 numbers of 500 gram packets of turmeric powder. He has randomly checked 10 packets and has found out the weights (in grams) to be as follows:

[< Answer >](#)

491	490	493	492	494
496	498	491	495	494

It is assumed that the distribution of weights of the packets is normal. The retailer wants to know whether he should reject the entire consignment keeping the customers' interests in mind.

At a significance level of 2.5 percent which of the following can be concluded?

- (a) The value of the test statistic is -8.58 and the mean weight of the consignment is significantly less than 500 grams
(b) The value of the test statistic is -8.58 and the mean weight of the consignment is **not** significantly less than 500 grams
(c) The value of the test statistic is -8.34 and the mean weight of the consignment is significantly less than 500 grams
(d) The value of the test statistic is -8.34 and the mean weight of the consignment is **not** significantly less than 500 grams
(e) The value of the test statistic is 8.58 and the mean weight of the consignment is significantly less than 500 grams.

(3 marks)

62. Popular Garments Company is in the business of trading men's garments. It purchases readymade garments from different manufacturers and sells them in the retail market. The company sells a particular brand of trousers named "Boss" every year during the summer season. It has to decide on the number of these trousers to stock for this season. Each trouser costs Rs.600 and sells for Rs.800. The salesman at the counter is given a commission of Rs.40 on every trouser sold. The stock of this brand of trousers is maintained for a period of 4 months. In the past the sales of this brand of trousers during the same period of time have averaged 300 with a standard deviation of 40. The trousers, which remain unsold at the end of the 4-month period are sold at 45% discount from the selling price. Further, on each trouser sold at discount the salesman at the counter is given a commission of Rs.80. The demand for the trousers is assumed to be normally distributed.

[< Answer >](#)

What is the optimal number of the "Boss" brand trousers which should be stocked by the company? (Round off your answer to the nearest integer).

- (a) 200 (b) 266 (c) 290 (d) 310 (e) 334.

(3 marks)

63. Orient Electricals Company manufactures water pumps. The general manager of the company wants to compare the performance of the water pumps manufactured by the company with the industry standards. He knows that 18% of all the water pumps of similar type sold in the industry require repairs during the first year of sale. The company sampled 200 customers and found that in case of 44 customers the water pumps required repairs in the first year of sale. It is to be tested whether the performance of the water pumps manufactured by the company is different from the industry standards.

[< Answer >](#)

At a significance level of 5% which of the following can be concluded?

- (a) The value of the test statistic is -1.471 and the performance of the water pumps manufactured by the company is **not** significantly different from the industry standards
(b) The value of the test statistic is -54.201 and the performance of the water pumps manufactured by the company is significantly different from the industry standards
(c) The value of the test statistic is 54.201 and the performance of the water pumps manufactured by

- the company is **not** significantly different from the industry standards
- (d) The value of the test statistic is 1.471 and the performance of the water pumps manufactured by the company is **not** significantly different from the industry standards
- (e) The value of the test statistic is 1.471 and the performance of the water pumps manufactured by the company is significantly different from the industry standards.

(3 marks)

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

[< Answer >](#)

If $X = 10$, $\hat{Y} = 55$

If $X = 15$, $\hat{Y} = 75$

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

$$\sum X = 60 \quad \sum X^2 = 645$$

Number of observations = 6

What is the value of $\sum XY$?

- (a) 900 (b) 225 (c) 3480 (d) 2025 (e) 3025.

(3 marks)

65. A survey was conducted on the retail prices of a specific type of rice in a city. At a confidence level of 95 percent it was estimated that the mean retail price of the commodity fell between Rs.23.02 per kg and Rs.24.98 per kg. The sample size was 64 and less than 5 per cent of the population of the grain retailers in the city were sampled.

[< Answer >](#)

What are the limits within which the true mean price has a 90.5 percent probability of falling?

- (a) Rs.21.495 and Rs.26.505 (b) Rs.23.165 and Rs.24.835
(c) Rs.23.50 and Rs.24.50 (d) Rs.20.665 and Rs.27.345
(e) Rs.22.33 and Rs.25.67.

(3 marks)

66. The manager of Ace Computer Parts is deliberating on the existing policy of ordering CDs. Presently 1200 CDs are ordered per month, but the firm runs out of stock one out of every five months. On average, the firm sells 900 CDs per month. It is assumed that the monthly demand for the CDs is normally distributed. The manager wants to order enough CDs so that the probability of meeting the demand in any month is not less than 90 percent.

[< Answer >](#)

What is the number of CDs that should be ordered per month by the firm? (Round off your answer to the nearest integer).

- (a) 457 (b) 1154 (c) 1357 (d) 360 (e) 1800.

(3 marks)

67. Capital Finance Company Ltd. is a non-banking finance company. In the recent years there has been variations in the demand for loans as well as in the rate of interest. The general manager of the company wants to know whether the variations in the amount of advances explain the variations in the net profit of the company.

[< Answer >](#)

The following data pertain to the operations of the company in the recent years:

Year	2003	2002	2001	2000	1999	1998
Amount of advances (Rs. in crores)	6.0	4.5	4.0	3.5	3.0	3.0
Profit after tax (Rs. in lakhs)	50	44	38	23	15	10

What percentage of the variations in net profit is explained by the variations in the amount of advances?

- (a) 84.31% (b) 15.69% (c) 91.82% (d) 28.92% (e) 71.08%.

(3 marks)

68. Pioneer Book Mart sells a variety of books, magazines and newspapers. The manager of the shop is concerned about one particular newspaper, The Chronicle. This newspaper costs Rs.1.80 each to the shop and can be sold for Rs.2.20 each. Any copy of the newspaper not sold by the end of the day can be disposed off for Re.0.40 each to the paper packet manufacturers. The manager of the shop has recorded the daily sales of the newspaper over the past 250 days and his observations are given below:

[< Answer >](#)

Number of copies sold in a day	Number of days sold
250	50
375	75
500	100
625	25
	250

It is assumed that the number of copies of the newspaper sold in a day is a discrete random variable which will assume only the numbers listed above.

What is the optimal number of copies of the newspaper that should be stocked by the shop?

- (a) 250 (b) 375 (c) 500 (d) 625 (e) 400.

(3 marks)

69. A simple regression equation is developed which relates the variables **X** and **Y**; **X** is the independent variable and **Y** is the dependent variable. The following details are available:

[< Answer >](#)

Y	62	80	60	68	50	40
---	----	----	----	----	----	----

$$\sum XY = 4000 \quad \sum X = 60 \quad \sum X^2 = 800$$

Number of data points = 6

What percentage of the variations in Y does the regression line, **not** explain?

- (a) 9.1% (b) 17.4% (c) 34.8% (d) 82.6% (e) 91.0%.

(3 marks)

70. Modern Finance Ltd. offers various types of financing to its clients. Its clientele includes individuals as well as business firms. It offers leasing and hire purchase financing as well as direct loans to the business firms for a variety of asset purchases. Presently, there are 240 business firms among its clients; the remaining are individuals. It has been observed that the borrowers belonging to the business firm category tend to pay their repayment dues regularly. However, of late some irregularities have been observed in the payment habits of its individual borrowers. The individuals constitute 80 percent of all its borrowers. The managing director of the company ordered a random survey of all the accounts in the individual borrowers category. 120 accounts in the individual borrowers category were randomly surveyed and it was found that 90 of them were in excellent standing.

[< Answer >](#)

Which of the following represents the interval estimate, at a confidence level of 95 percent for the absolute number of accounts in the individual borrowers category, which are in excellent standing?

- (a) 90 ± 74.38 (b) 90 ± 69.62 (c) 120 ± 69.62 (d) 720 ± 74.38 (e) 720 ± 69.62 .

(3 marks)

71. A large courier company is interested in modeling the problem of damaged letters. It has been suggested that any letter sent to a certain area has a 0.16 chance of being damaged. Because the courier company is so big, it can be assumed that the possibilities of any two letters being damaged are independent. A sample of 372 clients of the company was selected, and two dummy letters were

[< Answer >](#)

mailed to each of them. The number of clients receiving zero, one, or two damaged letters was 312, 48, and 12, respectively. The company wants to test whether it is reasonable to conclude that the number of damaged letters received by clients follows a binomial distribution with probability of success equal to 0.16.

At the significance level of 5 percent, which of the following can be concluded?

- (a) The value of the test statistic is 5.991 and the number of damaged letters follows a binomial distribution with probability of success equal to 0.16
- (b) The value of the test statistic is 5.991 and the number of damaged letters does **not** follow a binomial distribution with probability of success equal to 0.16
- (c) The value of the test statistic is 37.029 and the number of damaged letters follows a binomial distribution with probability of success equal to 0.16
- (d) The value of the test statistic is 37.029 and the number of damaged letters does **not** follow a binomial distribution with probability of success equal to 0.16
- (e) The value of the test statistic is 9.343 and the number of damaged letters does **not** follow a binomial distribution with probability of success equal to 0.16.

(3 marks)

END OF QUESTION PAPER

Suggested Answers

Quantitative Methods – II (132) – July 2004

1. Answer : (e)

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

(b) This is the wrong answer. The number of degrees of freedom does not depend on the ratio of sample size and the population.

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

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

(e) This is the right answer. When the chi-square distribution is used as a test of independence

2. Answer : (a)

Reason : (a) This is the **right** answer. In a control chart there are some observations that lie outside the upper and lower control limits. These points are referred to as the Outliers.

(b) This is the wrong answer. Qualitative variables with only two categories are termed as attributes. This is monitored with the help of p chart.

(c) This is the wrong answer. This is another name of random variation.

(d) This is the wrong answer. This is the nonrandom, systematic variability in a process. It usually can be corrected without redesigning the entire process.

(e) This is the wrong answer. This is the random variability inherent in the system. It usually cannot be reduced without redesigning the entire process.

3. Answer : (c)

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

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

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

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

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

4. Answer : (d)

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

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

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

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

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

5. Answer : (c)

Reason : (a) This is a true statement. The probability that we associate with an interval estimate is

called the confidence interval. This probability, then, indicates how confident we are that the interval estimate will include the population parameter. A higher probability means more confidence.

- (b) This is a true statement. Using high confidence level is not always desirable, because high confidence levels produce large confidence interval, and such large intervals are not precise; they give very fuzzy estimates.
- (c) This is a **false** statement. The number of degrees of freedom used in t-distribution estimation are equal to $(n-1)$, where n is the sample size. We lose one degree of freedom by estimating the standard deviation of the population with the help of the sample standard deviation.
- (d) This is a true statement. The t distribution need not be used in estimating if you know the standard deviation of the population. We can use the normal distribution in such case.
- (e) This is a true statement. The confidence interval is the range of the estimate we are making. An interval estimate describes a range of values within which a population parameter is likely to lie.

6. Answer : (e)

- Reason :
- (a) This is an advantage of sampling. The population may be too large to be studied in full. In such situations it is always better to use sampling compared to the study of whole population.
 - (b) This is an advantage of sampling. A study of sample is usually cheaper than a study of the population. As the number of observations required is less the cost incurred in obtaining and processing these information is also less for sampling.
 - (c) This is an advantage of sampling. Sampling usually provides information quicker than a census. Since the time required for studying the whole population is large, it is advisable to go for sampling to reduce the time spent in the studying the whole population.
 - (d) This is an advantage of sampling. In destructive testing, where the product becomes unusable after the performance of test, we cannot test the whole population. In such cases sampling is the only method of testing.
 - (e) This is **not** an advantage of sampling. In a sampling exercise we use a small number of samples to study the whole population the degree of accuracy is not very high. For accurate information census is always the preferable method compared to sampling.

7. Answer : (d)

- Reason :
- (a) This is a wrong answer. If the covariance of two random variables X and Y is equal to zero then it can be said that the random variables are independent of each other. If the variables are dependent on each other then the covariance will assume any non-zero value.
 - (b) This is a wrong answer. If the coefficient of correlation is positive then we say that there is a positive correlation between two variables.
 - (c) This is a wrong answer. If the coefficient of correlation is negative then we say that there is a negative correlation between two variables.
 - (d) This is the **correct** answer. If the covariance of two random variables X and Y is equal to zero then it can be said that the random variables are independent of each other.
 - (e) This is a wrong answer. If the covariance of the two random variables is zero it does not mean that the variances of the individual variables will be zero.

8. Answer : (d)

- Reason :
- (a) This statement is true for probability distribution. Usually, the past behaviour of the variable is studied and the frequency distribution of the data is formed. If the past

behaviour can be taken as a representative pattern for the future also, then the past frequency distribution can be used as a guide for predicting the future values of the random variable.

- (b) This is a true statement for probability distribution. The probability distribution gives an idea of the likely values of a random variable and the chances of occurrence of the various values.
- (c) This is a true statement for probability distribution. If the past behaviour does not help in forming a probability distribution then, a subjective probability distribution of the likely future values is formed on the basis of expert opinion.
- (d) This is **not true** for probability distribution. We can assign probabilities to all likely values of the random variable on the basis of the past data.
- (e) This is a true statement for probability distribution. The probability distribution for a random variable provides a probability for each possible value and that these probabilities must sum to 1.

9. Answer : (e)

- Reason :
- a. Fisher's ideal price index considers base year prices.
 - b. Fisher's ideal price index considers base year quantities.
 - c. Fisher's ideal price index considers current year prices.
 - d. Fisher's ideal price index considers current year prices.
 - e. Fisher's ideal price index is the geometric mean of the Laspeyres and Paasche's price indices

10. Answer : (a)

- Reason :
- a. The graphical plot of the values of the dependent and independent variables, in the context of regression analysis, is called scatter diagram.
 - b. 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.
 - c. A histogram is a graphical representation of a frequency distribution.
 - d. A p chart is a quality control chart.
 - e. An ogive is a graphical plot of a cumulative frequency distribution

11. Answer : (d)

- Reason :
- a. The F statistic is not calculated on the basis of the mean of the largest sample.
 - b. The F statistic is not calculated on the basis of the mean of the smallest sample.
 - c. The F statistic is not calculated on the basis of the standard deviation of the smallest sample.
 - d. The F statistic is calculated on the basis of two estimates of the population variance viz., the estimated population variance based on the variance among the sample means and the estimated population variance based on the variance within the samples.
 - e. The F statistic is not calculated on the basis of the variance of the largest sample

12. Answer : (b)

- Reason :
- a. 'Hugging the control limits' indicates that two different populations are being observed.
 - b. 'Hugging the center line' indicates that the variations have been reduced significantly.
 - c. Cycles indicate the possibility of presence of random variations in the process.
 - d. Increasing trend indicates that the process mean is increasing.
 - e. Decreasing trend indicates that the process mean is decreasing

13. Answer : (c)

- Reason : (a) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to one. [TOP](#)
- (b) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to one hundred.
- (c) This is the **right** answer. According to the factor reversal test if we interchange the price and quantity terms in an index number then the product of the original index number and the index number obtained by reversing the factors should be equal to the value index.
- (d) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to the Fisher's ideal index.
- (e) This is the wrong answer. According to the factor reversal test the product of the original index and the index obtained through reversing the factors is not equal to the chain index number.

14. Answer : (e)

- Reason : (a) This is true for multiple correlation coefficient. It depends on the correlation coefficient between Y and X_1 . [TOP](#)
- (b) This is true for multiple correlation coefficient. It depends on the correlation coefficient between X_1 and X_2 .
- (c) This is true for multiple correlation coefficient. It depends on the correlation coefficient between Y and X_2 .
- (d) This is true for multiple correlation coefficient. It will take values between 0 and 1.
- (e) This is **false** for multiple correlation coefficient. It will take values between 0 and 1. It does not take negative values.

15. Answer : (d)

- Reason : (a) This is true for hypothesis testing. The probability that the null hypothesis is true but rejected is α . In other words probability of committing type I error is α , the level of significance. [TOP](#)
- (b) This is true for hypothesis testing. The probability that the null hypothesis is false but is accepted is β . In other words, the probability of committing type II error is β .
- (c) This is a true statement. $(1-\beta)$ is known as the power of test. A high $(1-\beta)$ implies that the test is doing exactly what it should be doing i.e. Rejecting the null hypothesis when it is false. A low $(1-\beta)$ indicates poor performance.
- (d) This is a **false** statement. A high $(1-\beta)$ implies that the test is doing exactly what it should be doing i.e. Rejecting the null hypothesis when it is false. A low $(1-\beta)$ indicates poor performance.
- (e) This is a true statement. We can make two types of errors in hypothesis testing namely Type I error and Type II error. There is a trade-off between two types of errors: The probability of making one type of error can be reduced only if we are willing to increase the probability of making the other type of error.

16. Answer : (c)

- Reason : a. The Y intercept of the regression line does not represent the true value of Y when $X=0$. [TOP](#)
- b. The Y intercept of the regression line does not represent the change in average value of Y per unit change in X.
- c. The Y intercept of the regression line represents the mean value of Y when $X=0$.
- d. The Y intercept of the regression line does not represent the standard deviation of the values of X.

e. The Y intercept of the regression line does not represent mean of the values of X

17. Answer : (b)

- Reason :
- a. The payoffs represent the outcomes of certain decisions.
 - b. The states of nature refer to the chance occurrences that can affect the outcome of an individual's decision
 - c. The decision alternative courses of action that are available to the decision maker.
 - d. The probabilities refer to the likelihood of the chance occurrences or the states of nature.

18. Answer : (b)

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

19. Answer : (c)

Reason : The acceptance of null hypothesis does not prove that it is true because there are situations when the sample statistic falls within the acceptance region, though null hypothesis is false. That is why type II error occurs.

20. Answer : (e)

Reason : Finite population correction factor = $\sqrt{\frac{N-n}{N-1}}$
($N - n < (N - 1)$)

$\therefore$ Finite population correction factor is always less than 1 for $n > 1$.

21. Answer : (e)

Reason : The finite population correction factor need not be used for calculating the standard error of a statistic if the population is finite and sampling is done with replacement.

22. Answer : (c)

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

23. Answer : (e)

Reason : In a binomial distribution the probability of getting zero or more number of successes = $P(x = 0) + P(x = 1) + \dots P(x = n) = 1.00$.

24. Answer : (a)

Reason : The sampling distribution of mean is a distribution of all possible means for samples of a given size taken from the same population. The conditions mentioned in the other alternatives are not true in this context. So, the alternative (a) is correct.

25. Answer : (b)

Reason : In any right tailed test of hypothesis, the rejection region lies under the right tail. So the alternative (b) is correct.

26. Answer : (c)

Reason : The standard error of estimate for any regression equation is given by:

$$S_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}} \text{ where}$$

y = Observed values of the dependent variable

$\hat{y}$ = Estimated values from the estimating equation that correspond to each Y value and

n = Number of data points.

27. Answer : (b)

Reason : The dependent variable decreases as the independent variable increases in an estimating equation signifies that there is a negative nature of association between the two variables. Hence the correlation co-efficient will be negative and will lie in the range of 0 to -1.00 (because correlation coefficient cannot be less than -1).

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

28. Answer : (a)

Reason : Coefficient of determination represents the proportion of variation in the dependent variable that is explained by the regression line. Coefficient of correlation signifies the nature of association between two variables. Coefficient of variation states the extent of variability in a set of data points. Standard error of estimate measures the variability of the observed values around the regression line. So the alternative (a) is correct.

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

29. Answer : (a)

Reason : The weighted average of relatives price index with the base year values used as weights, can be readily compared because the weights in the base year and current year remain the same (and for every year the same weights are used).

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

30. Answer : (b)

Reason : If the regression equation is a perfect estimator of the dependent variable, then the following are possible:

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

- i. The standard error of estimate is zero as $\sum (y - \hat{y})^2 = 0$.
- ii. The coefficient of determination is 1.00 as the coefficient of correlation is either $+1.00$ or -1.00 .
- iii. Naturally, all the data points must lie on the regression line.

Hence, the alternative (b) is correct.

31. Answer : (a)

Reason : Analysis of variance can be used to test whether means of more than two populations are equal. No other techniques, mentioned in the other alternatives, can be applied in this respect. So the alternative (a) is correct.

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

32. Answer : (b)

Reason : If a random variable x has a log normal distribution, then $\ln(x)$ will be normally distributed. The expected value of the variable need not necessarily be 0 or 1.

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

33. Answer : (c)

$$\frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

Reason : $r =$

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

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

34. Answer : (a)

Reason : The volume of liquid present in a container is a continuous random variable.

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

35. Answer : (c)

- Reason :
- a. The expected value of a discrete random variable is not a geometric average of the outcomes of the variable.
 - b. The expected value of a discrete random variable is not a simple average of the outcomes of the variable.
 - c. The expected value of a discrete random variable is a weighted average of the outcomes of the variable.
 - d. The expected value of a discrete random variable is not the outcome, which has the highest frequency.
 - e. The expected value of a discrete random variable is not the highest probability of occurrence in the distribution of the random variable

[TOP](#)

36. Answer : (b)

- Reason :
- b. According to the Central Limit Theorem the sampling distribution of the mean can be approximated by the normal distribution as the sample size increases.
- a, c, d & e. These alternatives are not the correct interpretations of the Central Limit Theorem

[TOP](#)

37. Answer : (b)

- Reason :
- a. As the sample size increases the variation of the sample mean from the population mean becomes smaller.
 - b. From above we can see that (b) is incorrect.
 - c. It can not be said with certainty that as the sample size increases the variance of the sample becomes less than the variance of the population.
 - d. It can not be said with certainty that as the sample size increases the standard deviation of the sample becomes less than the standard deviation of the population.
 - e. It can not be said with certainty that as the sample size increases the standard deviation of the sample comes close to zero.

[TOP](#)

38. Answer : (a)

- Reason :
- a. The coefficient of determination is the square of coefficient of correlation (r). Hence it will always be ≥ 0 and ≤ 1 .
 - b. From above we can see that the coefficient of determination cannot be less than zero.
 - c. 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 .
 - d. The coefficient of determination is always positive; The coefficient of correlation may be negative.
 - e. 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

[TOP](#)

39. Answer : (b)

- Reason :
- a. In a p chart, the center line is drawn on the basis of overall sample proportion.
 - b. In a p chart, the lower control limit can only be greater than or equal to zero.
 - c d & e. The upper and lower control limits are fixed by adding and subtracting the estimated standard error of proportion from the overall sample proportion

[TOP](#)

40. Answer : (d)

- Reason :
- There is a different t distribution for different degrees of freedom. The degree of freedom is dependent on the sample size. Hence, there is a different t distribution for different sample sizes.

[TOP](#)

41. Answer : (d)

- Reason :
- Proportion of variations in the dependent variable explained by the regression relationship

[TOP](#)

$$= 1 - \frac{ESS}{TSS} = 1 - \frac{1630}{16730} = 0.9025 \text{ i.e. 90.25 percent}$$

42. Answer : (d)

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

$$a = \bar{Y} - b\bar{X}$$

$$\text{Given : } b = 2.5 \quad \sum X = 150 \quad \sum Y = 455 \quad n = 10$$

$$\therefore a = \left(\frac{455}{10} \right) - 2.5 \left(\frac{150}{10} \right) = 8.0$$

$\therefore$ Regression equation is $\hat{Y} = 8.0 + 2.5X$

$$\text{If } X = 10, \text{ then } \hat{Y} = 8.0 + (2.5 \times 8) = 28.0$$

43. Answer : (e)

$$\begin{aligned} \text{Reason : The lower control limit} &= \bar{p} + 3 \sqrt{\frac{pq}{n}} = 0.08 + 3 \times \sqrt{\frac{0.08 \times 0.92}{50}} \\ &= 0.08 + 3 \times 0.038 = 0.08 + 0.115 = 0.195 \end{aligned}$$

44. Answer : (b)

Reason : The required estimate of population variance based on the variance within the sample will

$$\text{be } \frac{24.60}{1.20} = 20.50$$

45. Answer : (d)

$$\text{Reason : The number of data points can be obtained as } \frac{\sum (y - \hat{y})^2}{s_e^2} + (k + 1)$$

$$= \frac{519.84}{5.70^2} + 2 + 1 = 19$$

46. Answer : (e)

Reason : X lies within ± 1.6 times standard deviation $\left(\frac{30-26}{2.5} = 1.6 \text{ and } \frac{22-26}{2.5} = -1.6 \right)$ of mean.
So the probability is $2 \times 0.4452 = 0.8904$ i.e. 89.04%

47. Answer : (d)

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

$$\begin{aligned} \text{Paasches price index} &= \frac{(\text{Fisher's ideal price index})^2}{\text{Laspeyres price index}} \\ &= \frac{(125.5)^2}{122.5} = 128.57 \end{aligned}$$

48. Answer : (a)

Reason : Mean of a binomial distribution = $n \cdot p$

>

$\therefore$ The mean of the binomial distribution of the number of successes is = $20(1 - 0.40)$
= 12

49. Answer : (c)

Reason : Year = 2001

<
TOP
>

$$\therefore x = 2001 - 2000 = 1$$

$$\hat{Y} = 120 + (2.5 \times 1) = 122.5$$

$$\text{Relative cyclical residual} = \left(\frac{Y - \hat{Y}}{\hat{Y}} \right) \times 100$$

$$Y = 130$$

$$\therefore \text{Relative cyclical residual for the year 2000} = \left(\frac{130 - 122.5}{125} \right) \times 100 = 6.12.$$

50. Answer : (d)

Reason : In ANOVA:

<
TOP
>

Number of degrees of freedom in the numerator of F ratio = Number of samples – 1

$$= 4 - 1 = 3$$

Number of degree of freedom in the denominator of F ratio

= Total sample size – Number of samples

$$= (6 + 6 + 7 + 5) - 4$$

$$= 20$$

51. Answer : (e)

$$\text{Reason : } E(Z) = E(4X + 5Y)$$

$$= 4E(X) + 5E(Y)$$

$$= 4[(12 \times 0.20) + (17 \times 0.40) + (22 \times 0.30) + (27 \times 0.10)]$$

$$+ 5[(11 \times 0.25) + (15 \times 0.55) + (19 \times 0.20)]$$

$$= 74 + 74$$

$$= 148$$

<
TOP
>

52. Answer : (e)

Reason : The probability that a randomly selected person supports party A is 0.55. Then the probability that he/she does not support party A is $1 - 0.55 = 0.45$

<
TOP
>

$\therefore$ The event that a randomly selected person does not support party A, follows a binomial distribution with probability of success = 0.45.

The probability that less than 8 out of 10 person do not support party A

$$= 1 - P(\text{At least 8 persons do not support party A})$$

$$= 1 - [P(8) + P(9) + P(10)]$$

$$= 1 - \left[{}^{10}C_8 (0.45)^8 (1 - 0.45)^2 + {}^{10}C_9 (0.45)^9 (1 - 0.45) + {}^{10}C_{10} (0.45)^{10} (1 - 0.45)^0 \right]$$

$$= 1 - [0.02289 + 0.00416 + 0.00034]$$

$$= 0.97261$$

53. Answer : (b)

Reason : Let X be the number of cricket matches played and Y be the number of accidents during weekends.

$$\sum X = 132, \sum Y = 41, n = 7$$

$$\sum XY = 873,$$

$$\sum X^2 = 2994,$$

$$\bar{X} = 18.86, \bar{Y} = 5.86$$

$$\text{The coefficient of regression } b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} = \frac{(7 \times 873) - (132 \times 41)}{(7 \times 2994) - (132^2)} = \frac{699}{3534} = 0.198$$

$$\text{The Y intercept } a = \bar{Y} - b \bar{X} = 5.86 - (0.198 \times 18.86) = 2.126$$

$$\text{Therefore the estimating equation is } \hat{Y} = 2.126 + 0.198 X$$

$$\text{The estimate of Y for } X = 25, \text{ is } \hat{Y} = 2.126 + 0.198 \times 25 = 7.08 \text{ (Approx)}$$

54. Answer : (b)

Reason : Let the returns of ABC be denoted by A_i and XYZ be denoted by X_i .

$$\therefore E(A) = \sum A_i p_i = (0.10 \times 40) + (0.50 \times 30) + (0.40 \times 10) = 23\%$$

$$E(X) = \sum X_i p_i = (0.10 \times 30) + (0.50 \times 20) + (0.40 \times 5) = 15\%$$

$$\begin{aligned} \text{Cov}(A, X) &= \sum [A_i - E(A)][X_i - E(X)] p_i \\ &= (40 - 23)(30 - 15)0.10 + (30 - 23)(20 - 15)0.50 + (10 - 23)(5 - 15)0.40 \\ &= (17 \times 15 \times 0.10) + (7 \times 5 \times 0.50) + (-13 \times -10 \times 0.40) \\ &= 95 (\%)^2 \end{aligned}$$

55. Answer : (b)

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

	Years (X)	Inventory (Y)	$x = X - \bar{X}$	$x.Y$	x^2
	1999	4720	-2	-9440	4
	2000	5010	-1	-5010	1
	2001	5590	0	0	0
	2002	5830	1	5830	1
	2003	6090	2	12180	4
Total	10005	27240	0	3560	10

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

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

$$a = \bar{Y} = \frac{27240}{5} = 5448$$

$$\text{Therefore the estimating equation is } \hat{Y} = 5,448 + 356 x$$

[TOP](#)

[TOP](#)

[TOP](#)

56. Answer: (b)

Reason:

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

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

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

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

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

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

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

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.

Degrees of freedom = $n - 2 = 5 - 2 = 3$.

The appropriate t-value is 3.182 (from the t-table).

The limits of the prediction interval are given below:

$$\text{For, } x = 7.50, \quad \hat{Y} = 1236 - 104(8.00) = 404$$

The prediction interval is: $\hat{Y} - t(s_e) = 404 \pm (3.182 \times 14.61) = 404 \pm 46.49$ passengers

57. Answer: (e)

Reason:

Proportion of variations in Y that is explained by variations in X

$$= r^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \bar{Y})^2}$$

From above, we have the following:

$$r^2 = 0.84$$

$$\Sigma(Y - \bar{Y})^2 = 1369 \text{ (given)}$$

$$\therefore 0.84 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{1369}$$

$$\text{or } \frac{\Sigma(Y - \hat{Y})^2}{1369} = 1 - 0.84 = 0.16$$

$$\therefore \Sigma(Y - \hat{Y})^2 = 1369 \times 0.16 = 219.04$$

$$\text{Standard error of estimate} = \sqrt{\frac{\Sigma(Y - \hat{Y})^2}{n-2}} = \sqrt{\frac{219.04}{6-2}} = 7.40$$

58. Answer : (a)

$$\text{Reason : Estimated standard error of mean} = \frac{\text{Sample standard deviation}}{\sqrt{n}}$$

[TOP](#)

[TOP](#)

[TOP](#)

Sample standard deviation, $s = \sqrt{\frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}}$

$$\bar{x} = \frac{\sum x}{n} = \frac{256}{16} = 16$$

$$\therefore s = \sqrt{\frac{4710.4}{15} - \frac{16(16)^2}{15}} = 6.40$$

$$\therefore \text{Estimated standard error of mean} = \frac{6.4}{\sqrt{16}} = \frac{6.4}{4} = 1.6$$

59. Answer : (a)

Reason : $p = \frac{1}{3}$

$$\therefore q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$$

Variance = npq

$$\therefore (2)^2 = n \left(\frac{1}{3}\right) \left(\frac{2}{3}\right)$$

$$\text{or } 4 = \frac{2}{9}n$$

$$\text{or } n = \frac{4 \times 9}{2} = 18$$

$$\therefore P(r = 10) = {}^{18}C_{10} \left(\frac{1}{3}\right)^{10} \left(\frac{2}{3}\right)^8 = 0.0289$$

60. Answer : (b)

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

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

$$= \frac{0.80 \times 6 \times 12}{36} = 1.60$$

61. Answer: (a)

Reason : Mean weight of the sample ($\bar{x}$) = $\frac{\sum x}{n} = \frac{493.4}{10} = 49.34$ gms

$$\text{Sample standard deviation (s)} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{56.4}{10-1}} = 2.503 \text{ gms}$$

$$\text{Sampling fraction} = \frac{10}{160} = 0.0625 > 0.05$$

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

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

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

Hence the standard error applicable to finite populations has to be calculated.

$$\therefore \text{Estimated standard error of mean, } \sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{2.503}{\sqrt{10}} \sqrt{\frac{160-10}{160-1}} = 0.769 \text{ gms.}$$

The consignment should be rejected if it can be inferred at a significance level of 2.5% that the entire consignment is underweight i.e., the mean weight of the packets in the consignment is less than 500 gms.

So a left tailed test of hypothesis has to be conducted. The null and the alternative hypotheses are given below:

$$H_0: \mu = 500$$

$$H_1: \mu < 500$$

Since the sample size is less than 30 and the standard deviation of population is unknown, the 't' distribution has to be used.

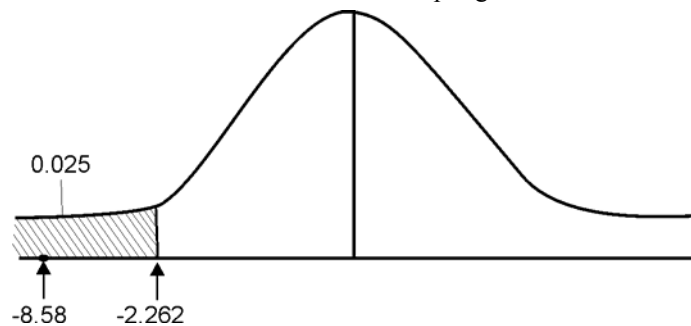
$$\therefore \text{Standardised 't' statistic for the sample} = \frac{\bar{x} - 500}{\sigma_{\bar{x}}}$$

$$\therefore t = \frac{493.4 - 500}{0.769} = -8.58$$

$$\text{Degrees of freedom} = n - 1 = 10 - 1 = 9$$

The t-value from the t-distribution table for 9 degrees of freedom and 5% of the area in both tails combined i.e., 2.5% of the area in each tail = 2.262.

So the cut-off t-value on the left side of the sampling distribution will be -2.262.



Since $-8.58 < -2.262$ (and it falls in the rejection region i.e., the shaded region in left tail) we conclude that the sample mean is significantly less than 500. Hence we reject the null hypothesis and accept the alternative hypothesis. So we infer that the mean weight of the consignment is significantly less than 500 gms.

62. Answer: (c)

Reason:

$$\begin{aligned} \text{Marginal profit (MP)} &= \text{Selling price} - \text{Cost} - \text{Commission} \\ &= 800 - 600 - 40 = \text{Rs. } 160 \end{aligned}$$

Marginal loss (ML):

$$\text{Discounted price of the Trousers} = 800 (1 - 0.45) = \text{Rs. } 440$$

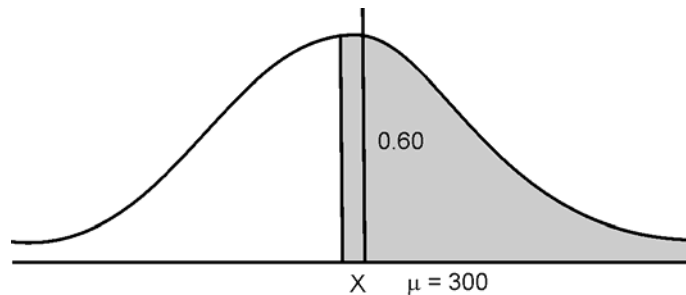
$$\text{Less: Commission payable} = \underline{\text{Rs. } 80}$$

$$\text{Amount receivable from discount sale} = \underline{\text{Rs. } 360}$$

$$\begin{aligned} \text{Marginal loss (ML)} &= \text{Cost of the T-shirt} - \text{Amount receivable from the discount sale} \\ &= 600 - 360 = \text{Rs. } 240 \end{aligned}$$

Minimum required probability of selling an additional T-shirt

$$(p^*) = \frac{ML}{ML + MP} = \frac{240}{240 + 160} = 0.60 \text{ i.e. } 60\%.$$



Since 0.50 of the area under the normal distribution is located between the mean and the right tail 0.10 of the shaded area in the diagram must be to the left of the mean, $(0.60 - 0.50 = 0.10)$. From the standard normal distribution table we find by interpolation the Z-value which

$$\text{corresponds to an area of 0.10 from the mean i.e.} = 0.25 + \frac{(0.26 - 0.25)}{(0.1026 - 0.0987)} (0.10 - 0.0987) = 0.253.$$

Let the stock level be denoted X.

$$\therefore \frac{X - \mu}{\sigma} = -0.253 \text{ (-ve because it is on the left side of the mean).}$$

Given, $\mu = 300$, $\sigma = 40$

$$\therefore \frac{X - 300}{40} = -0.253$$

$$\text{or } X = 300 + (-0.253 \times 40) = 289.88 \cong 290$$

$\therefore$ The company should stock 290 Trousers of the brand.

63. Answer (d)

Reason: Sample size, $n = 200$

$$\text{Sample proportion, } \bar{p} = \frac{44}{200} = 0.22$$

$$H_0 : p = 0.18 \text{ (= } p_0)$$

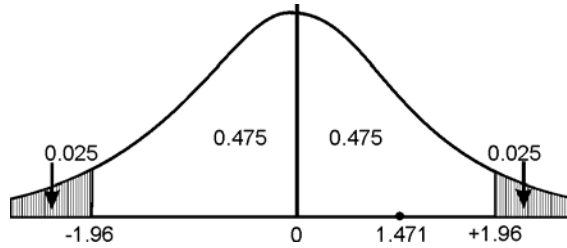
$$H_1 : p \neq 0.18$$

$$\alpha = 0.05$$

$$\begin{aligned} \text{Standard error of proportion, } \sigma_{\bar{p}} &= \sqrt{\frac{p_0 \cdot q_0}{n}} \\ &= \sqrt{\frac{(0.18)(1 - 0.18)}{200}} = 0.0272 \end{aligned}$$

Standardizing the sample proportion:

$$Z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0.22 - 0.18}{0.0272} = 1.471$$



From the normal distribution curve we can find out that the test statistic falls in the acceptance region. So we accept the null hypothesis. Hence, we infer that the performance of the water pumps manufactured by the company does not significantly differ from the industry standards.

64. Answer : (c)

Reason : A simple regression equation is of the form:

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

Given:

$$\text{If } X = 15, \quad \hat{Y} = 75 \therefore 75 = a + 15b \dots\dots\dots (A)$$

$$\text{If } X = 10, \quad \hat{Y} = 55 \therefore 55 = a + 10b \dots\dots\dots (B)$$

Subtracting (A) from (B) we get:

$$(75 - 55) = (15 - 10)b$$

$$\text{or } 20 = 5b$$

$$\text{or } b = \frac{20}{5} = 4$$

Putting the values of b in (B):

$$55 = a + 10 \times 4 = a + 40$$

$$\text{or } a = 55 - 40 = 15$$

$$\therefore \hat{Y} = 15 + 4X$$

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

Now, a =

$$\bar{Y} = a + b\left(\frac{\sum X}{n}\right) = 15 + 4\left(\frac{60}{6}\right) = 55$$

$\therefore$

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$\text{or } 4 = \frac{\sum XY - 6\left(\frac{60}{6}\right)(55)}{645 - 6\left(\frac{60}{6}\right)^2}$$

$$\text{or } 4 = \frac{\sum XY - 3300}{45}$$

$$\text{or } \sum XY = (4 \times 45) + 3300 = 3480$$

65. Answer: (b)

Reason: The sample size is greater than 30 and the standard deviation of population is not known. Hence the appropriate distribution applicable is the normal distribution. A confidence interval of 95 per cent implies that 0.475 of the area of the normal curve is covered on either side of the sample mean. From the normal distribution table it is found that the appropriate z-value is 1.96.

$$\bar{x} = ? \quad s = ? \quad \hat{\sigma}_{\bar{x}} = ?$$

For the given confidence interval :

$$\text{Lower limit} = \text{Rs.23.02 per kg}$$

$$\text{Upper limit} = \text{Rs.24.98 per kg.}$$

$$\therefore \bar{x} - 1.96 \hat{\sigma}_{\bar{x}} = 23.02 \quad (\text{A})$$

$$\bar{x} + 1.96 \hat{\sigma}_{\bar{x}} = 24.98 \quad (\text{B})$$

Adding equations (A) and (B) we get :

$$2\bar{x} = 48$$

$$\text{or } \bar{x} = 24 \text{ (Rs. per kg.)}$$

Subtracting equation (A) from equation (B) we get :

$$3.92 \hat{\sigma}_{\bar{x}} = 1.96$$

$$\therefore \hat{\sigma}_{\bar{x}} = \frac{1.96}{3.92} = 0.50 \text{ (Rs. per kg)}$$

$$\text{Mean price of the sample} = \text{Rs.24.00 per kg}$$

$$\text{Estimated standard error of mean } (\hat{\sigma}_{\bar{x}}) = \text{Rs.0.50 per kg.}$$

For a confidence interval of 90.5%, z-value(s) are : ± 1.67

$$\therefore \text{Lower limit} = \bar{x} - 1.67 \hat{\sigma}_{\bar{x}} = 24 - 1.67 \times 0.50 = \text{Rs.23.165}$$

$$\text{Upper limit} = \bar{x} + 1.67 \hat{\sigma}_{\bar{x}} = 24 + 1.67 \times 0.50 = \text{Rs.24.835}$$

66. Answer: (c)

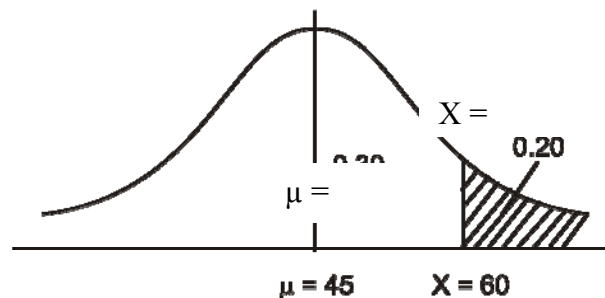
Reason:

Probability that the monthly demand for CDs will exceed 1200 (i.e., existing number of CDs

$$\text{ordered per month}) = \frac{1}{5} = 0.20$$

Average number of CDs sold per month (μ) = 900

The monthly demand for the CDs is normally distributed.



We can see that, $P(900 \leq x \leq 1200) = 0.30$

From the standard normal table, we can find that the z value for a point lying to the right of the mean, so that 0.30 of the area is covered between the mean and the point, is

$$(by\ interpolation) = 0.84 + \frac{(0.85 - 0.84)}{(0.3023 - 0.2995)} \times (0.30 - 0.2995) = 0.842$$

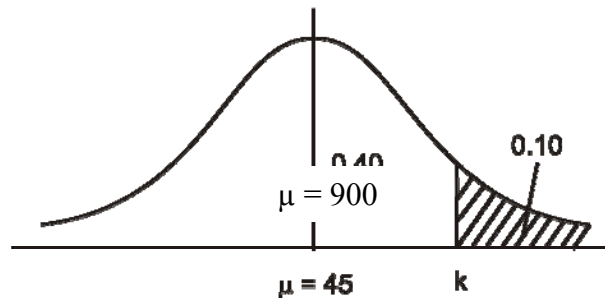
$$z = \frac{x - \mu}{\sigma}$$

$$\therefore 0.842 = \frac{1200 - 900}{\sigma} \quad \text{or} \quad \sigma = \frac{300}{0.842} = 356.29$$

Required: Probability that the demand will be met with the available stock = 0.90

$\therefore$ Probability that the firm will run out of stock = $1 - 0.90 = 0.10$

Let the number of CDs to be ordered be k.



We can see that, $P(900 \leq x \leq k) = 0.40$.

From the standard normal table, we can find that the z value corresponding to k (by interpolation).

$$= 1.28 + \frac{(1.29 - 1.28)}{(0.4015 - 0.3997)} (0.40 - 0.3997) = 1.282$$

$$z = \frac{x - \mu}{\sigma}$$

$$\therefore 1.282 = \frac{k - 900}{356.29}$$

$$\text{or } k = 1.282 \times 356.29 + 900 = 1356.76 \approx 1357$$

$\therefore$ 1357 CDs should be ordered per month.

67. Answer: (a)

Reason: Let the following notations be used:

X : Amount of advances (Rs. in crores)

Y : Net profit (Rs. in lakhs)

Coefficient of correlation is given by:

$$r = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\sqrt{\Sigma(X - \bar{X})^2 \cdot \Sigma(Y - \bar{Y})^2}}$$

A.	X	6.0	4.5	4.0	3.5	3.0	3.0	$\Sigma X = 24$
----	---	-----	-----	-----	-----	-----	-----	-----------------

B. Y	50	44	38	23	15	10	$\Sigma Y = 180$
C. $(X - \bar{X})$	2	0.5	0	-0.5	-1	-1	
D. $(Y - \bar{Y})$	20	14	8	-7	-15	-20	
E. $(X - \bar{X})^2$	4	0.25	0	0.25	1	1	$\Sigma(X - \bar{X})^2 = 6.5$
F. $(Y - \bar{Y})^2$	400	196	64	49	225	400	$\Sigma(Y - \bar{Y})^2 = 1334$
G. $(X - \bar{X})(Y - \bar{Y})$	40	7	0	3.5	15	20	$\Sigma(X - \bar{X})(Y - \bar{Y}) = 85.5$

$$\bar{X} = \frac{\Sigma X}{n} = \frac{24}{6} = 4$$

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{180}{6} = 30$$

$$\therefore r = \frac{85.5}{\sqrt{6.5 \times 1334}} = 0.9182$$

Proportion of variations in net profit that is explained by the variations in the amount of advances = $r^2 = (0.9182)^2 = 0.8431$

$\therefore$ Percentage of variations in profit after tax that is explained by the variations in the amount of advances = $0.8431 \times 100 = 84.31\%$.

68. Answer: (b)

Reason:

Profit per newspaper sold = $2.20 - 1.80 = \text{Re.}0.40$

Loss per unsold newspaper = $1.80 - 0.40 = \text{Re.}1.40$

Deriving the probabilities:

Daily sales	250	375	500	625	
Number of days sold (f)	50	75	100	25	$\Sigma f = 250$
Probability $\left(p = \frac{f}{\Sigma f}\right)$	$\frac{50}{250} = 0.20$	$\frac{75}{250} = 0.30$	$\frac{100}{250} = 0.40$	$\frac{25}{250} = 0.10$	$\Sigma p = 1.00$

Conditional profit table

Stock level	Daily demand (Probability)				Expected daily profit (Rs.)
	250 (0.20)	375 (0.30)	500 (0.40)	625 (0.10)	
250	100	100	100	100	100
375	-75	150	150	150	105
500	-250	-25	200	200	42.5
625	-425	-200	25	250	-110

From the above table we can see that the maximum expected daily profit (Rs.105) is associated with the stock level of 375 newspapers. Hence the optimal number of the newspaper, The Chronicle, that should be stocked is 375.

Working notes:

Conditional profit for any stock level = Number of newspapers demanded and sold $\times$ 0.40 –
Number of unsold newspapers $\times$ 1.40.

69. Answer : (b)

Reason : Proportion of variation in Y, that is not explained by the regression line = 1 – Coefficient of determination (r^2)

$$r^2 = \frac{a \sum y + b \sum XY - n \bar{Y}^2}{\sum Y^2 - n \bar{Y}^2}$$

$$\sum Y = 360 \quad \sum Y^2 = 22568$$

$$b = \frac{n \sum XY - \sum X \cdot \sum Y}{n \sum X^2 - (\sum X)^2} = \frac{(6 \times 4000) - (60 \times 360)}{(6 \times 800) - (60 \times 60)} = \frac{2400}{1200} = 2$$

$$a = \bar{Y} - b \bar{X} = \left(\frac{360}{6} \right) - 2 \times \left(\frac{60}{6} \right) = 40$$

$$r^2 = \frac{(40 \times 360) + (2 \times 4000) - (6 \times 60^2)}{22568 - (6 \times 60^2)} = \frac{800}{968} = 0.826$$

$\therefore$

$\therefore$ Proportion of variation in Y that is not explained by the regression line = $1 - r^2$

$$= 1 - 0.826 = 0.174$$

$\Rightarrow$ Percentage of variations in Y that is not explained by the regression line = 17.4%

70. Answer : (e)

Reason : Total number of individual borrower accounts = $\frac{240}{(1 - 0.80)} \times 0.80 = 960$

Number of individual borrower accounts surveyed = 144

$$\text{Sampling fraction} = \frac{n}{N} = \frac{120}{960} = 0.125 > 0.05$$

Hence this can be treated as sampling done from a finite population.

$$\therefore \text{Proportion of excellent borrowers in the sample, } \bar{P} = \frac{90}{120} = 0.75$$

$$\bar{Q} = 1 - \bar{P} = 1 - 0.75 = 0.25$$

$$\begin{aligned} \text{Estimated standard error of proportion, } \hat{\sigma}_{\bar{P}} &= \sqrt{\frac{\bar{P}\bar{Q}}{n}} \sqrt{\frac{N-n}{N-1}} \\ &= \sqrt{\frac{0.75 \times 0.25}{120}} \sqrt{\frac{960-120}{960-1}} \\ &= 0.03953 \times 0.9359 \\ &= 0.037 \end{aligned}$$

Since both $n\bar{P}$ and $n\bar{Q}$ are greater than 5, the sample can be treated as a large sample and

normal distribution can be used. For a confidence level of 95%, $Z = \pm 1.96$

$$1.96 \hat{\sigma}_{\bar{p}} = 1.96 \times 0.037 = 0.07252$$

Confidence interval for the proportion of accounts in the individual borrower category who are in the excellent standing = $\bar{p} \pm 1.96 \hat{\sigma}_{\bar{p}} = 0.75 \pm 0.07252$

$$\begin{aligned} \therefore \text{Confidence interval for the absolute number of accounts in that category} \\ &= (0.75 \times 960) \pm (0.07252 \times 960) \\ &= 720 \pm 69.62 \end{aligned}$$

71. Answer : (d)

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

H_0 : A binomial distribution with $p = 0.16$ is a good fit with the distribution of damaged letters.

H_1 : A binomial distribution with $p = 0.16$ is **not** a good fit with distribution of damaged letters.

If the case of damaged letters follows a binomial probability distribution with $p = 0.16$, then

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

$$p = 0.16 \text{ and } q = 1 - p = 0.84$$

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

$$P(0) = {}^2C_0 (0.16)^0 (0.84)^2 = 0.7056$$

$$P(1) = {}^2C_1 (0.16)^1 (0.84)^1 = 0.2688$$

$$P(2) = {}^2C_2 (0.16)^2 (0.84)^0 = 0.0256$$

Therefore if the total number of people to whom the two letters are sent is 372, then the number of people who are expected to receive zero, one or two damaged letters will be,

$$n(0) = 0.7056 \times 372 = 262.48$$

$$n(1) = 0.2688 \times 372 = 100$$

$$n(2) = 0.0256 \times 372 = 9.52$$

Now tabulating the observed and expected frequency as follows:

Number of people receiving mangled letters	Observed Frequency (f_o)	Expected Frequency (f_e)	$\frac{(f_o - f_e)^2}{f_e}$
0	312	262.48	9.343
1	48	100	27.04
2	12	9.52	0.646
Total	372	372	37.029

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 37.029$$

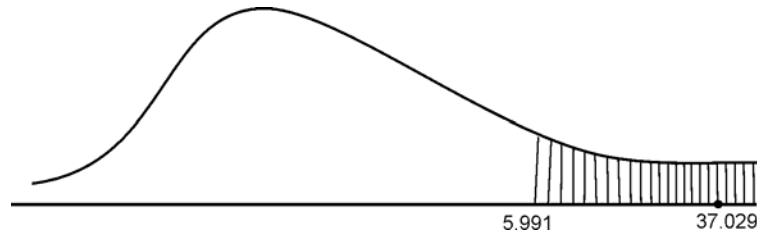
The test statistic,

The degrees of freedom for the goodness of fit test is $(k - 1)$, where k is the number of

classes observed. Therefore the number of degrees of freedom is $(3 - 1) = 2$.

The critical value for the 2 degrees of freedom at 5 percent significance level is = **5.991**.

We see that the test statistic exceeds the critical value. So it clearly falls in the rejection region. Therefore **we reject the null hypothesis** at 5 percent significance level. We conclude that the number of damaged letters received by people **does not** follow a binomial distribution with probability of success equal to 0.16.



[< TOP OF THE DOCUMENT >](#)