PART - III

STATISTICS

(தமிழ் மற்றும் ஆங்கில் பாடு / Tamil & English Versions)

Instructions: (1) Check the question paper for fairness of printing. If there is any lack of fairness, inform the Hall Supervisor immediately.
(2) Use Black or Blue ink to write and underline and pencil to draw diagrams.

பாதி - I/PART - I

தீப்பாட்: (i) தொடர்பு விளக்கங்களுக்கு விளக்கம் பயன்படுத்துநர்.
(ii) நூற்றாண்டு விளக்கங்கள் தக்க மெய்ப்பாட்டு.
Note: (i) Answer all the questions.
(ii) Each question carries one mark.

தீப்பாட்: சின்னங்கள் விளக்கங்கள் சுருக்கியது சடத்தம்.
Note: Choose the correct answer.

1. A, B நிலையில் சார்பு விளக்கங்கள் என்பதாக P(A∩B) = ________.

   (a) $P(A) \cdot P(B)$  (b) $P(A) P(B/A)$  (c) $\frac{P(A \cap B)}{P(A)}$  (d) $P(A) + P(B)$

   If A and B are independent events then P(A∩B) = ________.

   (a) $P(A) \cdot P(B)$  (b) $P(A) P(B/A)$  (c) $\frac{P(A \cap B)}{P(A)}$  (d) $P(A) + P(B)$

[திறுப்பட்ட / Turn over]
2. \[ P(X) = 0.15 \quad P(Y) = 0.25 \quad P(X \cap Y) = 0.10 \] 
   \[ \text{or} \quad P(X \cup Y) \text{ is:} \]
   
   \[ (a) \quad 0.10 \quad (b) \quad 0.20 \quad (c) \quad 0.30 \quad (d) \quad 0.40 \]
   
   If \( P(X) = 0.15 \) \( P(Y) = 0.25 \) \( P(X \cap Y) = 0.10 \) then \( P(X \cup Y) \) is:
   
   (a) 0.10  (b) 0.20  (c) 0.30  (d) 0.40

3. \[ \text{Probability is expressed as:} \]
   
   (a) ratio  (b) percentage  (c) proportion  (d) all the above

4. \[ \text{Probability of drawing a spade queen from a well shuffled pack of cards is:} \]
   
   \[ (a) \quad \frac{1}{13} \quad (b) \quad \frac{1}{52} \quad (c) \quad \frac{4}{13} \quad (d) \quad 1 \]

5. \[ \text{Classical probability is also known as:} \]
   
   (a) Statistical probability  (b) A priori probability  (c) Empirical probability  (d) None of the above
6. The probability of an impossible event is:
(a) 0  (b) 1  (c) -1  (d) 2

7. When three dices are thrown the probability of sum being 17 is:
(a) \(\frac{3}{216}\)  (b) \(\frac{2}{216}\)  (c) \(\frac{1}{216}\)  (d) 0

8. From the given random variable table, the value of \(a\) is:
(a) 1  (b) \(\frac{1}{2}\)  (c) 4  (d) \(\frac{1}{4}\)

9. Variance of the random variable \(x\) is:
(a) \(E(x^2) - [E(x)]^2\)  (b) \([E(x)]^2 - E(x^2)\)  (c) \(E(x^2)\)  (d) \([E(x)]^2\)
10. \( \text{Var}(5x + 2) \) is:

(a) 25 \( x \) \hspace{1cm} (b) 5 \( x \) \hspace{1cm} (c) 2 \( x \) \hspace{1cm} (d) 25

11. \( E(2x + 3) \) is:

(a) \( E(2x) \) \hspace{1cm} (b) \( 2E(x) + 3 \) \hspace{1cm} (c) \( E(3) \) \hspace{1cm} (d) \( 2x + 3 \)

12. \( f(x) \) is the p.d.f. of the continuous random variable \( x \), then \( E(x^2) \) is:

(a) \( \int -\infty^{\infty} f(x) \, dx \) \hspace{1cm} (b) \( \int -\infty^{\infty} xf(x) \, dx \) \hspace{1cm} (c) \( \int -\infty^{\infty} x^2 f(x) \, dx \) \hspace{1cm} (d) \( \int -\infty^{\infty} f(x^2) \, dx \)

13. \( f(x) \) is the p.d.f. of a continuous random variable \( X \) with p.d.f. \( f(x) \) then \( F(X) = \int -\infty^{x} f(x) \, dx \)
14. The random variable \( X \) is defined by \( X = 1 \) with probability \( \frac{1}{2} \) and \( X = 2 \) with probability \( \frac{1}{4} \). Which statement is true?

(a) The mean of \( X \) is 1.5
(b) The mode of \( X \) is 2
(c) The variance of \( X \) is 0.75
(d) The standard deviation of \( X \) is 0.866

Mathematical expectation of random variable \( X \) is also known as ________.

(a) mean
(b) standard deviation
(c) variance
(d) mode

15. \( \left( \frac{2}{3} + \frac{1}{3} \right)^9 \) is the probability of getting 3 successes in a binomial distribution where the probability of success is 0.5. The mean of \( X \) is ________.

(a) \( \sqrt{2} \)
(b) 2
(c) \( \sqrt{3} \)
(d) 3

If \( \left( \frac{2}{3} + \frac{1}{3} \right)^9 \) refers to the binomial distribution then its standard deviation is ________.

(a) \( \sqrt{2} \)
(b) 2
(c) \( \sqrt{3} \)
(d) 3

16. The variance of the binomial distribution is:

(a) \( npq \)
(b) \( np \)
(c) \( \sqrt{npq} \)
(d) 0

The variance of the binomial distribution is:

(a) \( npq \)
(b) \( np \)
(c) \( \sqrt{npq} \)
(d) 0

17. The random variable \( X \) is defined by \( X = 1 \) with probability \( \frac{1}{2} \), \( X = 2 \) with probability \( \frac{1}{4} \), \( X = 3 \) with probability \( \frac{1}{8} \), and \( X = 4 \) with probability \( \frac{1}{8} \). Then the probability \( P(X=1) \) is ________.

(a) \( \frac{1}{12} \)
(b) \( \frac{1}{4} \)
(c) \( \frac{1}{6} \)
(d) \( \frac{1}{8} \)

The mean and variance of a binomial distribution are 8 and 4 respectively. Then \( P(X=1) \) is equal to:

(a) \( \frac{1}{12} \)
(b) \( \frac{1}{4} \)
(c) \( \frac{1}{6} \)
(d) \( \frac{1}{8} \)
18. For a Poisson distribution:
(a) mean > variance  (b) mean = variance
(c) mean < variance  (d) mean ≠ variance

19. Poisson distribution corresponds to:
(a) rare events  (b) certain event
(c) impossible event  (d) sure events

20. The area $P(-\infty < z < 0)$ is equal to:
(a) 1  (b) 0.1  (c) 0.5  (d) 0

21. The mean of a normal distribution is 60 its mode will be:
(a) 60  (b) 40  (c) 50  (d) 30

22. The standard normal distribution has:
(a) $\mu = 1, \sigma = 0$  (b) $\mu = 0, \sigma = 1$
(c) $\mu = 0, \sigma = 0$  (d) $\mu = 1, \sigma = 1$
23. \[ \frac{\sum x^2}{N} - \left( \overline{x} \right)^2 \] is used to find ________.

(a) variance  (b) standard deviation  
(c) mean  (d) mode

24. The formula \[ \frac{\sum x^2}{N} - \left( \overline{x} \right)^2 \] is used to find ________.

(a) variance  (b) standard deviation  
(c) mean  (d) mode

25. The standard error of observed sample proportion \( P \) is:

(a) \( \sqrt{\frac{P(1-P)}{n}} \)  (b) \( \sqrt{\frac{PQ}{n}} \)  
(c) \( \sqrt{\frac{(1-P)Q}{n}} \)  (d) \( \frac{PQ}{n} \)

26. A wrong decision about \( H_0 \) leads to:

(a) One kind of error  (b) Two kinds of error  
(c) Three kinds of error  (d) Four kinds of error
27. \( H_0 : \mu = 100 \) vs. \( H_1 : \mu \neq 100 \) (Choose Correct)

- (a) Right tailed test
- (b) Left tailed test
- (c) Two tailed test
- (d) None of the above

Testing \( H_0 : \mu = 100 \) vs. \( H_1 : \mu \neq 100 \) lead to:

(a) One sided right tailed test
(b) One sided left tailed test
(c) Two tailed test
(d) None of the above

28. Find the variance of the sample mean.

- (a) \( \frac{\sigma^2}{n} \)
- (b) \( \frac{\sigma}{\sqrt{n}} \)
- (c) \( \frac{\sigma^2}{n^2} \)
- (d) \( \frac{\sqrt{\sigma}}{n} \)

The variance of the sample mean is:

(a) \( \frac{\sigma^2}{n} \)
(b) \( \frac{\sigma}{\sqrt{n}} \)
(c) \( \frac{\sigma^2}{n^2} \)
(d) \( \frac{\sqrt{\sigma}}{n} \)

29. Find the statistics for difference between two means.

- (a) \( \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \)
- (b) \( \frac{p - P}{\sqrt{pq/n}} \)
- (c) \( \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \)
- (d) \( \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \)

The statistics for difference between two means is:

(a) \( \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \)
(b) \( \frac{p - P}{\sqrt{pq/n}} \)
(c) \( \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \)
(d) \( \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \)
30. Standard error of the difference of proportions \((p_1 - p_2)\) in two classes under the hypothesis \(H_0 : p_1 = p_2\) with usual notation is:

\[
\begin{align*}
\text{(a)} & \quad \sqrt{pq \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \\
\text{(b)} & \quad \sqrt{p \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \\
\text{(c)} & \quad pq \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\
\text{(d)} & \quad \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}
\end{align*}
\]

31. If \(Z_0 < Z_e\) then the null hypothesis is:

\[
\begin{align*}
\text{(a)} & \quad \text{accepted} \\
\text{(b)} & \quad \text{rejected} \\
\text{(c)} & \quad \text{cannot be formed} \\
\text{(d)} & \quad \text{nothing can be said}
\end{align*}
\]

32. Student's t-distribution was pioneered by:

\[
\begin{align*}
\text{(a)} & \quad \text{Karl Pearson} \\
\text{(b)} & \quad \text{Laplace} \\
\text{(c)} & \quad \text{R.A. Fisher} \\
\text{(d)} & \quad \text{William S. Gosset}
\end{align*}
\]

[ Turn over ]
33. While testing significance of the difference between two sample means in case of small samples the degrees of freedom is:

(a) \( n_1 + n_2 \)  
(b) \( n_1 + n_2 - 1 \)  
(c) \( n_1 + n_2 - 2 \)  
(d) \( n_1 + n_2 + 2 \)

34. \( v = 2 \) gives \( \chi^2_{0.05} \) as.

(a) 5.9  
(b) 5.99  
(c) 5.55  
(d) 5.95

For \( v = 2 \), \( \chi^2_{0.05} \) equals:

(a) 5.9  
(b) 5.99  
(c) 5.55  
(d) 5.95

35. The contingency table (4x3) has an expected value equal to:

(a) 12  
(b) 9  
(c) 8  
(d) 6

Degrees of freedom for chi-square in case of contingency table of order (4x3) are:

(a) 12  
(b) 9  
(c) 8  
(d) 6

36. F distribution has higher variance when it is in the denominator.

(a) variance  
(b) numerator  
(c) denominator  
(d) both variance

Customarily the larger variance in the variance ratio for F statistic is taken:

(a) in the numerator  
(b) in the denominator  
(c) either way  
(d) none of the above

37. The assumption in t-distribution is that the population standard deviation is ________.

(a) not known  
(b) known  
(c) 1  
(d) 0
38. \[ F = \frac{S_1^2}{S_2^2} \] is the test statistic for testing:

(a) \( H_0: \mu_1 = \mu_2 \) (b) \( H_0: \sigma_1^2 = \sigma_2^2 \) (c) \( H_0: \sigma_1 = \sigma \) (d) \( H_0: \sigma^2 = \sigma_0^2 \)

39. Analysis of variance technique originated in the field of:

(a) Agriculture (b) Industry (c) Biology (d) Genetics

40. One of the assumption of analysis of variance is that the population from which the samples are drawn is:

(a) Binomial (b) Poisson (c) \( \chi^2 \) (d) Normal

41. In case of one way classification with \( N \) observations and \( t \) treatments the error degrees of freedom is:

(a) \( N-1 \) (b) \( t-1 \) (c) \( N-t \) (d) \( Nt \)
42. Which is the best measure for trend analysis? (a) Mean (b) Median (c) PM Mean (d) GM Mean

Salient features responsible for seasonal variations are:
(a) Weather (b) Social Customers (c) Festivals (d) All the above

43. Which is the best measure of central tendency?
(a) Mean (b) Median (c) Mode (d) GM Mean

The economic rhythm theory comes under the category of:
(a) Analytical methods (b) Naive method (c) Barometric methods (d) None of the above

44. The component of time-series attached to long term variation is termed as:
(a) Secular Trend (b) Seasonal Variation (c) Irregular Variation (d) All the above

45. If A and B are independent, Yule’s coefficient is equal to _______.
(a) 0 (b) +1 (c) −1 (d) ∞
46. \( N = 500 \text{ (A)} = 300 \text{ (B)} = 250 \text{ (AB)} = 40 \) then data are

(a) Inconsistent
(b) Consistent
(c) Negatively associated
(d) Independent

47. With the two attributes the total number of class frequencies is:

(a) Two
(b) Four
(c) Eight
(d) Nine

48. Which of the following criteria does not apply decision making under uncertainty?

(a) Maximin return
(b) Maximax return
(c) Minimax return
(d) Maximize expected return
49. In which of the following methods we use loss-tables?

(a) Minimax return  (b) Maximin method  
(c) Maximax return  (d) EMV - method

50. A chance node in a decision tree diagram represented as ________.

(a) square  (b) circle  (c) arrow mark  (d) EMV

51. Define statistical probability.

52. For two independent events A, B  \( P(A) = \frac{1}{2} \) and  \( P(B) = \frac{1}{3} \) , find the probability that one of them occur.
53. The function of a continuous variable \( x \) is \( f(x) = A x^2 \), \( 0 < x < 1 \). Find the constant \( A \) so that the given function is a p.d.f.

If \( f(x) = A x^2 \), \( 0 < x < 1 \) is a p.d.f. of a continuous random variable, find the value of \( A \).

54. Define the mathematical expectation for a discrete random variable.

55. Write a short note about 'moment generating function'.

56. \((0.68 + 0.32)^{10}\) is the cumulative distribution of a binomial distribution with \( n = 10 \) and \( p = \frac{3}{4} \). Find the probability of 2 successes.

For the binomial distribution \((0.68 + 0.32)^{10}\) find the probability of 2 success.

57. If the variance of a Poisson Distribution is 0.5, find \( P(x = 3) \) \((e^{-0.5} = 0.6065)\).

If the variance of a Poisson Distribution is 0.5, find \( P(x = 3) \) \((e^{-0.5} = 0.6065)\).

58. Give any two properties of a normal distribution.

59. Define null hypothesis and alternative hypothesis.

60. Explain clearly type I and type II errors.
61. Give the test statistics for proportions.

62. State the assumptions of student's t-test.

63. Write short note on Yate's correction.

64. Define chi-square test.

65. What are the components of time series?

66. Give the names of different methods of measuring trend.

67. Verify whether the given data \( N = 100 \) \( A = 75 \) \( B = 60 \) \( AB = 15 \) are consistent.

68. Give Yule's coefficient of association.
69. Explain the meaning of ‘statistical decision theory’.

70. Write a note on decision tree.

**PART - III**

**6x5 = 30**

Note: Answer any six questions.

71. Two persons A, B appeared for an interview for a job. The probability of selection of A is $\frac{1}{3}$ and that of B is $\frac{1}{2}$. Find the probability that (i) both of them will be selected (ii) only one of them will be selected (iii) none of them will be selected.

72. Let $X$ be a continuous random variable with p.d.f. $f(x) = 3x^2$, $0 < x < 1$. Find mean and variance.
73. Three coins are tossed simultaneously. Find the probability of getting (i) exactly 7 heads (ii) 5 heads (iii) 7 tails.

Ten coins are tossed simultaneously. Find the probability of getting (i) at least seven heads (ii) exactly seven heads (iii) at most seven heads.

74. A factory produces a large number of items. The probability that an item is defective is 1/1000. A sample of 1000 items is tested. The probability that exactly 5 items are defective in the sample is: \( e^{-2} = 0.13534 \)

Suppose on an average one house in 1000 in a certain district gets fire during a year find the probability that exactly 5 houses will have fire out of 2000 houses of that district. \( e^{-2} = 0.13534 \).

75. 10 students have the following heights (in cm):

38, 40, 45, 53, 47, 43, 55, 48, 52, 49

It is known that on average, 10-year-old boys in a certain town have a height of 140 cm. Is it correct?

Weights in kgs of 10 students are given below:

38, 40, 45, 53, 47, 43, 55, 48, 52, 49

Can we say that the variance of distribution of weights of all students from which the above sample of 10 students was drawn is equal to 20 kgs?

76. A large sample of 1000, 2000 items is taken from the population. The mean height of the population is 67.5 cm and 68 cm respectively. The mean of the sample is 67.5 cm. Can we say that the sample is drawn from the population with mean 5% and 2.5% less than the population mean at 5% level of significance?

The means of two large samples of 1000 and 2000 items are 67.5 cms and 68 cms respectively. Can the samples be regarded as drawn from the population with standard deviation 2.5 cms? Test at 5% level of significance.
77. பெண்களின் சீரான விளைந்தத் 3 மாத முறை நீரும் கோளத்துடன்.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>புத்தாண்டு (தாதை)</td>
<td>50</td>
<td>36</td>
<td>43</td>
<td>45</td>
<td>39</td>
<td>38</td>
<td>33</td>
<td>42</td>
<td>41</td>
<td>34</td>
</tr>
</tbody>
</table>

Calculate the three yearly average for the following data:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production (Tons)</td>
<td>50</td>
<td>36</td>
<td>43</td>
<td>45</td>
<td>39</td>
<td>38</td>
<td>33</td>
<td>42</td>
<td>41</td>
<td>34</td>
</tr>
</tbody>
</table>

78. முதல் ச்ராண்டு குழுக்களுக்கு அல்லது பழுப்பு குழுக்களுக்கு அடையாளப்படுத்தும் நூற்றாண்டுகள் காலத்தில் ,

<table>
<thead>
<tr>
<th>பாதுகாப்பு வினைப்படங்கள்</th>
<th>12</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>கார்பொடி உயர்ப்படங்கள்</td>
<td>16</td>
<td>6</td>
</tr>
</tbody>
</table>

In an experiment of immunization of cattle from tuberculosis the following results were obtained:

<table>
<thead>
<tr>
<th>Inculcated</th>
<th>Affected</th>
<th>Unaffected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>Not inculcated</td>
<td>16</td>
<td>6</td>
</tr>
</tbody>
</table>

By calculating Yule’s coefficient of association examine the effect of vaccine in controlling the disease.
79. Consider the pay off table: Using EMV criterion decide the best act.

<table>
<thead>
<tr>
<th>Act</th>
<th>EMV</th>
<th>Probability</th>
<th>EMV</th>
<th>Probability</th>
<th>EMV</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1</td>
<td>2500</td>
<td>0.4</td>
<td>3500</td>
<td>0.4</td>
<td>5000</td>
<td>0.2</td>
</tr>
<tr>
<td>A_2</td>
<td>2500</td>
<td>0.4</td>
<td>3500</td>
<td>0.4</td>
<td>2500</td>
<td>0.2</td>
</tr>
<tr>
<td>A_3</td>
<td>2500</td>
<td>0.2</td>
<td>1500</td>
<td>0.4</td>
<td>1000</td>
<td>0.4</td>
</tr>
</tbody>
</table>

80. A company has two plants to manufacture motorbikes. Plant I manufactures 80 percent of motorbikes and Plant II manufactures 20 percent. At Plant I 85 out of 100 motorbikes are rated standard quality or better. At Plant II only 65 out of 100 motorbikes are rated standard quality or better.

(i) What is the probability that the motorbike selected at random came from Plant I, if it is known that the motorbike is of standard quality?

(ii) What is the probability that the motorbike came from Plant II, if it is known that the motorbike is of standard quality?
81. In a normal distribution 31% of the items are under 45 and 8% are over 64. Find the mean and variance of the distribution.

82. In a certain city 125 men in a sample of 500 are found to be self-employed. In another city the number of self-employed are 375 in a random sample of 1000. Does this indicate that there is a greater population of self-employed in the second city than in the first?

83. The sales data of an item in six shops before and after a special promotional campaign are as under:

<table>
<thead>
<tr>
<th>Shops</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Campaign</td>
<td>53</td>
<td>28</td>
<td>31</td>
<td>48</td>
<td>50</td>
<td>42</td>
</tr>
<tr>
<td>After Campaign</td>
<td>58</td>
<td>29</td>
<td>30</td>
<td>55</td>
<td>56</td>
<td>45</td>
</tr>
</tbody>
</table>

Test whether the campaign is a success at 5% level of significance.
Three varieties of coal were analysed by four chemists and the ash-content in the varieties was found to be as under:

<table>
<thead>
<tr>
<th>Varieties</th>
<th>Chemists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>A</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
</tr>
</tbody>
</table>

Carry out the analysis of variance.

Find the seasonal variations by simple average method for data given below:

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Quarters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>1989</td>
<td>30</td>
</tr>
<tr>
<td>1990</td>
<td>34</td>
</tr>
<tr>
<td>1991</td>
<td>40</td>
</tr>
<tr>
<td>1992</td>
<td>54</td>
</tr>
<tr>
<td>1993</td>
<td>80</td>
</tr>
</tbody>
</table>
A farm owner is considering drilling a farm well. In past only 70% of wells drilled were successful at 20 meters of depth in that area. Moreover on finding no water some persons drilled further 25 meters were successful only for 20%. The prevailing cost for drilling ₹ 500 per meter. In case of not drilling a well he estimates that he has to buy water at ₹ 15,000 over the next ten years from his neighbour. Draw an appropriate decision tree and determine the farm owner's strategy under EMV approach.

The following decision can be optimal:

(i) do not drill any well.
(ii) drill upto 20 meters.
(iii) No water is found at 20 meters drill further 25 meters.

- o O o -