

Indian Statistical Institute, Bangalore Centre

Documentation Research and Training Centre
M.S. (Library and Information Science)
2nd Semester Final Examination (2023-2025)

Paper 09: Elements of Mathematics and Statistics

Time: 10.00 AM - 1:00 PM

Max. Marks: 60

Date: 29-04-2024

INSTRUCTION: Read the question before you attempt.

Section A

[1x10 = 10]

1. A medical treatment has a success rate of 0.8. Two patients will be treated with this treatment. Assuming the results are independent for the two patients, what is the probability that neither one of them will be successfully cured?
 - a. 0.5
 - b. 0.36
 - c. 0.2
 - d. 0.04

$P = 0.8$
 $= (1 - 0.8)(1 - 0.8)$
 $= (0.2)(0.2)$
2. Which of the following statements is correct about a parameter and a statistic associated with repeated random samples of the same size from the same population?
 - a. Values of a parameter will vary from sample to sample but values of a statistic will not.
 - b. Values of both a parameter and a statistic may vary from sample to sample.
 - c. Values of a parameter will vary according to the sampling distribution for that parameter.
 - d. Values of a statistic will vary according to the sampling distribution for that statistic.
3. The probability is $p = 0.80$ that a patient with a certain disease will be successfully treated with a new medical treatment. Suppose that the treatment is used on 40 patients. What is the "expected value" of the number of patients who are successfully treated?
 - a. 40
 - b. 20
 - c. 8
 - d. 32
4. Suppose that the probability of event A is 0.2 and the probability of event B is 0.4. Also, suppose that the two events are independent. Then $P(A|B)$ is:
 - a. $P(A) = 0.2$
 - b. $P(A)/P(B) = 0.2/0.4 = \frac{1}{2}$
 - c. $P(A) \times P(B) = (0.2)(0.4) = 0.08$
 - d. None of the above.

5. The expected value of a random variable is the
- value that has the highest probability of occurring.
 - mean value over an infinite number of observations of the variable.
 - largest value that will ever occur.
 - most common value over an infinite number of observations of the variable.
6. Pulse rates of adult men are approximately normal with a mean of 70 and a standard deviation of 8. Which choice correctly describes how to find the proportion of men that have a pulse rate greater than 78?
- Find the area to the left of $z = 1$ under a standard normal curve.
 - Find the area between $z = -1$ and $z = 1$ under a standard normal curve.
 - Find the area to the right of $z = 1$ under a standard normal curve.
 - Find the area to the right of $z = -1$ under a standard normal curve.
7. The test statistic used in large sample tests is
- Z-statistic
 - t-statistic
 - F-statistic
 - χ^2 -statistic
8. Karl Pearson's coefficient of correlation between x and y is 0.52, their covariance is + 7.8. If the variance of x is 16, find the standard deviation of y .
- 3.25
 - 2.25
 - 3.75
 - 2.75
- $$r = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$
9. Two uncorrelated random variables x and y have standard deviations 3 and 4 respectively. Find the s.d. of $(x + y)$.
- 25
 - 5
 - 9
 - 7
- $$\text{If } r = 0, \text{var}(x + y) = \text{var}(x) + \text{var}(y) = 25$$
- $$\sigma_{x+y} = 5$$
10. If now the prices of all commodities in a place have been decreased by 35% over the base period prices, then the index number of prices for the place is
- 100
 - 135
 - 65
 - 35

Section B

Answer any six of the following questions

[6 * 5 = 30]

- The time, in hours, that each student spent sleeping on a school night was recorded for 550 secondary school students. The distribution of these times was found to be approximately normal with a mean of 7.6 hours and a standard deviation of 0.5 hours.
 - Draw the normal distribution curve to show this data.
 - What amount of time did 95% of these students spend sleeping?
 - What percentage of students spent between 7.1 and 8.1 hours sleeping?
- 10% of the bulbs produced in a factory are of red colour and 2% are red and defective. If one bulb is picked up at random, determine the probability of its being defective if it is red.
- Two dice are thrown together. Let A be the event 'getting 6 on the first die' and B be the event 'getting 2 on the second die'. Are events A and B independent?
- Three machines E_1 , E_2 , E_3 in a certain factory produce 50%, 25% and 25%, respectively, of the total daily output of electric tubes. It is known that 4% of the tubes produced by each of machines E_1 and E_2 are defective, and that 5% of those produced by E_3 are defective. If one tube is picked up at random from a day's production, calculate the probability that it is defective.
- Bag I contains 3 red and 4 black balls while another Bag II contains 5 red and 6 black balls. One ball is drawn at random from one of the bags and it is found to be red. Find the probability that it was drawn from Bag II.
- Find the regression of x on y from the following data:
 $\Sigma x = 24$, $\Sigma y = 44$, $\Sigma xy = 306$, $\Sigma x^2 = 164$, $\Sigma y^2 = 574$, $n = 4$
Find the estimated value of x, when $y = 6$.
- Let the lines of regression concerning two variables x and y be given by $y = 32 - x$ and $x = 13 - 0.25y$. Obtain the values of the means and the correlation coefficient.

Section C

[1 * 10 = 10]

- The amount of a certain trace element in blood is known to vary with a standard deviation of 14.1 ppm (parts per million) for male blood donors and 9.5 ppm for female donors. Random samples of 75 male and 50 female donors yield concentration means of 28 and 33 ppm, respectively. What is the likelihood that the population means of concentrations of the element are the same for men and women? Test the hypothesis at both 5% and 1% level of significance. Clearly specify the hypotheses, the calculated test statistics, and the final decision should be made for both the level of significance.

Answer any one of the following questions (9. OR 10.)

[1 * 10 = 10]

9.

a. What do you mean by a time series? Explain the different components of such a series. With what characteristic component of time series should each of the following be associated?

(3+2+5)

- i. An upturn in business activities;
- ii. Withdrawal of bank deposits by 15th March for tax payment;
- iii. Fire loss in a factory;
- iv. General increase in sale of T.V. sets.

b. Fit a second degree polynomial to the following data:

Year	1882	1883	1884	1885	1886	1887	1888	1889	1890
Price Index	84	82	76	72	69	68	70	72	73

10.

a. Fit a least square line to the data below, using (i) x as the independent variable, (ii) y as the independent variable. Find y when $x = 5$. Also find x when $y = 7$.

(5+3+2)

x	3	5	6	8	9	11
y	2	3	4	6	5	8

b. Prove that the coefficient of correlation is the geometric mean of the coefficients of regression.

c. The coefficient of correlation r of the following data is 0.8

X:	1	2	3	4	5
Y:	3	2	5	4	6

If the values of X are increased by 4 and those of Y decreased by 3, what will be the value of r after the changes? Give reasons for your answer. (unaffected)