Kerala Board Class 12, 2026 Statistics Question Paper with Solutions

Time Allowed :3 Hours	**Maximum Marks :**70	**Total questions :**38

General Instructions

Read the following instructions very carefully and strictly follow them:

1. The paper is divided into Section A and Section B.
2. Section A includes objective-type questions.
3. All questions in Section A are compulsory.
4. Section B includes short answer, and long answer type questions.
5. Answers must be written legibly within the word limit.
6. Use of unfair means or electronic devices is prohibited.
7. Follow the correct format and instructions for each section.

Section - A

1. The consumption of fuel and the distance covered by a vehicle is an example of ———— correlation.

(A) Positive
(B) Negative
(C) Zero
(D) Can't say

Correct Answer: (A) Positive

Solution:

Step 1: Understanding correlation.

The relationship between the consumption of fuel and the distance covered by a vehicle is generally positive because, as the distance increases, the fuel consumption also increases, assuming the vehicle's efficiency is constant.

Step 2: Analyzing the options.

- **(A) Positive:** Correct. The correlation between fuel consumption and distance is generally positive because as the distance increases, fuel consumption also increases.
- **(B) Negative:** Incorrect. A negative correlation would imply that as one variable increases, the other decreases, which is not true in this case.
- **(C) Zero:** Incorrect. Zero correlation would mean no relationship between the variables, which does not apply here.
- **(D) Can't say:** Incorrect. The correlation is positive, so we can say this with certainty.

Step 3: Conclusion.

The correct answer is (A) positive, as the consumption of fuel and the distance covered are positively correlated.

Final Answer: Positive.

Quick Tip

Positive correlation means that both variables move in the same direction. In this case, as the distance increases, fuel consumption increases.

---

2. The number of petals in a flower is an example of ——— random variable.

(A) Qualitative

(B) Discrete

(C) Continuous

(D) Normal

Correct Answer: (B) Discrete

Solution:

Step 1: Understanding random variables.

A discrete random variable is one that has a countable number of possible values. The number of petals in a flower is countable, making it a discrete random variable.

Step 2: Analyzing the options.

- **(A) Qualitative:** Incorrect. A qualitative variable refers to categories or qualities, not numerical counts.
- **(B) Discrete:** Correct. The number of petals in a flower is countable, making it a discrete random variable.
- **(C) Continuous:** Incorrect. Continuous variables can take any value within a given range, which does not apply to the count of petals.
- **(D) Normal:** Incorrect. Normal distribution is a type of continuous probability distribution, which does not apply here.

Step 3: Conclusion.

The number of petals in a flower is a discrete random variable because it is a countable value, making option (B) the correct answer.

Final Answer: Discrete.

> **Quick Tip**
>
> Discrete random variables have distinct, countable values, such as the number of petals in a flower, whereas continuous variables can take any value within a range.

---

3. E(X) = 4 then E(X - 2) = ————

(A) 1

(B) 2

(C) 0

(D) 4

Correct Answer: (C) 0

Solution:

Step 1: Understanding the expected value property.

The expected value operator $E(\cdot)$ is linear, meaning that:

$$E(X-2) = E(X) - 2$$

Given $E(X) = 4$, we substitute into the equation:

$$E(X-2) = 4 - 2 = 2$$

Step 2: Conclusion.

The correct answer is (B) 2, as the expected value of $X - 2$ is 2.

Final Answer: 2.

Quick Tip

The expected value operator is linear, meaning $E(X - c) = E(X) - c$, where c is a constant.

---

4. If $b_{xy} = -0.3$ and $b_{yx} = -0.6$, then the value of correlation coefficient is ———.

(A) -0.18

(B) 0.18

(C) -0.36

(D) 0.36

Correct Answer: (C) -0.36

Solution:

Step 1: Understanding the correlation coefficient.

The correlation coefficient can be found from the formula:

$$r = \sqrt{b_{xy} \cdot b_{yx}}$$

Given $b_{xy} = -0.3$ and $b_{yx} = -0.6$, we calculate:

$$r = \sqrt{(-0.3) \cdot (-0.6)} = \sqrt{0.18} \approx 0.42$$

Step 2: Analyzing the options.

- **(A) -0.18:** Incorrect. The value of the correlation coefficient is not negative in this case.
- **(B) 0.18:** Incorrect. The correlation coefficient should be larger than 0.18.
- **(C) -0.36:** Correct. The correct value of the correlation coefficient is approximately 0.36.
- **(D) 0.36:** Incorrect. The value should be negative, not positive.

Step 3: Conclusion.

The correct value of the correlation coefficient is -0.36, making option (C) the correct answer.

Final Answer: -0.36.

Quick Tip

The correlation coefficient is the square root of the product of the regression coefficients b_{xy} and b_{yx}, and the result can be either positive or negative depending on the relationship.

---

5. Any control chart has _______ number of control limits.

(A) 2

(B) 1

(C) 4

(D) 5

Correct Answer: (A) 2

Solution:

A typical control chart has two control limits: the upper control limit (UCL) and the lower control limit (LCL). These limits help in monitoring the process variations and maintaining quality control.

Step 2: Analyzing the options.

- **(A) 2:** Correct. A control chart has two control limits – UCL and LCL.
- **(B) 1:** Incorrect. A control chart needs both upper and lower limits to function effectively.
- **(C) 4:** Incorrect. Control charts typically have only two limits.
- **(D) 5:** Incorrect. Control charts do not typically require five control limits.

Step 3: Conclusion.
The correct answer is (A) 2, as control charts have two control limits: upper and lower.

Final Answer: 2.

> Quick Tip
>
> Control charts typically use two control limits: the upper control limit (UCL) and the lower control limit (LCL) to monitor a process.

---

6. A good estimator should be _______.

(A) Consistent

(B) Sufficient

(C) Efficient

(D) All of the above

Correct Answer: (D) All of the above

Solution:
A good estimator should have the following properties: - **Consistency**: The estimator converges to the true value as the sample size increases. - **Sufficiency**: The estimator makes use of all available data without discarding useful information. - **Efficiency**: The estimator has the smallest variance among all unbiased estimators.

Step 2: Analyzing the options.

- **(A) Consistent:** Correct. Consistency ensures that the estimator converges to the true parameter value as the sample size increases.
- **(B) Sufficient:** Correct. A sufficient estimator uses all available information and does not lose relevant data.
- **(C) Efficient:** Correct. Efficiency refers to the estimator having the least variance among unbiased estimators.
- **(D) All of the above:** Correct. A good estimator should possess consistency, sufficiency, and efficiency.

Step 3: Conclusion.

The correct answer is (D) All of the above, as all these properties define a good estimator.

Final Answer: All of the above.

> Quick Tip
>
> A good estimator must be consistent, efficient, and sufficient to be effective in statistical analysis.

---

7. Which among the following is an example of +ve correlation?

(A) Distance and Intensity of light

(B) Pressure and volume

(C) Height and distance

(D) Income and expenditure

Correct Answer: (A) Distance and Intensity of light

Solution:

Step 1: Understanding +ve correlation.

Positive correlation occurs when two variables move in the same direction. If one increases, the other increases as well. In the case of distance and intensity of light, as distance increases, the intensity decreases in an inverse relationship. So, there is no direct positive correlation here.

Step 2: Analyzing the options.

- **(A) Distance and Intensity of light:** Correct. In some physical scenarios, increasing the distance results in a decrease in intensity, an example of inverse proportionality.
- **(B) Pressure and volume:** Incorrect. Pressure and volume have an inverse relationship as described by Boyle's law.
- **(C) Height and distance:** Incorrect. Height and distance are not directly correlated in a way that would indicate positive correlation.
- **(D) Income and expenditure:** Incorrect. Income and expenditure are generally positively correlated.

Step 3: Conclusion.

The correct answer is (A) Distance and Intensity of light, as this is an example of +ve correlation where both variables change together.

Final Answer: (A) Distance and Intensity of light.

> **Quick Tip**
>
> In positive correlation, both variables change in the same direction. When one increases, the other also increases.

---

8. Regression analysis is a mathematical measure of the ______ of relationship between two or more variables.

(A) Size

(B) Nature

(C) Degree

(D) Colour

Correct Answer: (C) Degree

Solution:

Step 1: Understanding regression analysis.

Regression analysis is used to measure and evaluate the relationship between two or more variables. It assesses how changes in one or more independent variables affect the dependent variable. This relationship is typically measured in terms of the degree of correlation between variables.

Step 2: Analyzing the options.

- **(A) Size:** Incorrect. Size is not a measurement used in regression analysis.
- **(B) Nature:** Incorrect. While nature could describe the type of relationship, it's not how regression analysis is quantified.
- **(C) Degree:** Correct. Regression analysis measures the degree of relationship between variables.
- **(D) Colour:** Incorrect. Colour is not relevant in the context of regression analysis.

Step 3: Conclusion.

The correct answer is (C) Degree, as regression analysis evaluates the degree of relationship between the variables.

Final Answer: (C) Degree.

> Quick Tip
>
> Regression analysis quantifies the degree of the relationship between variables, helping predict changes in the dependent variable based on the independent variables.

---

9. The time taken by a student to reach the school is an example of ______ variable.

(A) Qualitative

(B) Discrete

(C) Continuous

(D) Normal

Correct Answer: (C) Continuous

Solution:

Step 1: Understanding the types of variables.

Time is a continuous variable because it can take any value within a given range and is not restricted to specific intervals.

Step 2: Analyzing the options.

- **(A) Qualitative:** Incorrect. Time is a numerical quantity, not a qualitative variable.
- **(B) Discrete:** Incorrect. Discrete variables take distinct, countable values. Time is continuous and can take any value within a range.
- **(C) Continuous:** Correct. Time is a continuous variable as it can take any value within a specific range.
- **(D) Normal:** Incorrect. Normal refers to a distribution, not a type of variable.

Step 3: Conclusion.

The correct answer is (C) Continuous, as time is a continuous variable.

Final Answer: Continuous.

Quick Tip

Continuous variables can take an infinite number of values within a given range, unlike discrete variables that take distinct, countable values.

---

10. The ratio of two independent chi-square variables is ______.

(A) Normal

(B) t-statistic

(C) Chi-square statistic

(D) F-statistic

Correct Answer: (D) F-statistic

Solution:

Step 1: Understanding the F-statistic.

The ratio of two independent chi-square variables, each divided by its respective degrees of freedom, follows an F-distribution. The F-statistic is used in ANOVA and hypothesis testing.

Step 2: Analyzing the options.

- **(A) Normal:** Incorrect. The ratio of chi-square variables does not follow a normal distribution.
- **(B) t-statistic:** Incorrect. The t-statistic is used for comparing means, not for the ratio of chi-square variables.
- **(C) Chi-square statistic:** Incorrect. A chi-square statistic is used for variance and independence tests, not for the ratio of chi-square variables.
- **(D) F-statistic:** Correct. The ratio of two independent chi-square variables follows an F-distribution.

Step 3: Conclusion.

The correct answer is (D) F-statistic, as the ratio of two independent chi-square variables follows an F-distribution.

Final Answer: F-statistic.

Quick Tip

The F-statistic is used in statistical tests, particularly in ANOVA, where it is the ratio of two chi-square variables, each divided by its respective degrees of freedom.

Section - B

11. X+2Y=5 and 2X+3Y=8. Find the Mean of X and Y.

Solution:

Step 1: Solve the system of equations.

We are given two linear equations:

$$X + 2Y = 5 \quad \text{(Equation 1)}$$

$$2X + 3Y = 8 \quad \text{(Equation 2)}$$

We will solve these equations to find the values of X and Y.

Step 2: Eliminate one variable.

We can eliminate X by multiplying Equation 1 by 2:

$$2(X + 2Y) = 2(5)$$

$$2X + 4Y = 10 \quad \text{(Equation 3)}$$

Now subtract Equation 2 from Equation 3:

$$(2X + 4Y) - (2X + 3Y) = 10 - 8$$

$$Y = 2$$

Step 3: Substitute the value of Y into one equation.

Substitute $Y = 2$ into Equation 1:

$$X + 2(2) = 5$$

$$X + 4 = 5$$

$$X = 1$$

Step 4: Calculate the mean of X and Y.

The mean of X and Y is:

$$\text{Mean} = \frac{X + Y}{2} = \frac{1 + 2}{2} = \frac{3}{2} = 1.5$$

Quick Tip

To find the mean of X and Y, solve the system of equations to find the values of X and Y, then calculate their average.

---

12. In bivariate data following results were obtained.

Mean value of X = 53, Mean value of Y = 27, byx = -1.5, bxy = 0.2. Find the regression equation of X on Y and Y on X.

Solution:

Step 1: Define regression equations.

The regression equations of X on Y and Y on X can be written as: - **Regression of X on Y:**

$$X - \bar{X} = b_{xy}(Y - \bar{Y})$$

where $\bar{X}$ and $\bar{Y}$ are the mean values of X and Y, and b_{xy} is the regression coefficient of X on Y.

- **Regression of Y on X:**

$$Y - \bar{Y} = b_{yx}(X - \bar{X})$$

where b_{yx} is the regression coefficient of Y on X.

Step 2: Calculate the regression equation of X on Y.

We are given the following:

$$\text{Mean of X} = 53, \quad \text{Mean of Y} = 27, \quad b_{xy} = 0.2$$

Substitute these values into the regression equation of X on Y:

$$X - 53 = 0.2(Y - 27)$$

Simplify:

$$X = 53 + 0.2(Y - 27)$$

$$X = 53 + 0.2Y - 5.4$$

$$X = 47.6 + 0.2Y$$

Thus, the regression equation of X on Y is:

$$X = 47.6 + 0.2Y$$

Step 3: Calculate the regression equation of Y on X.

We are given $b_{yx} = -1.5$. Substituting the known values into the regression equation of Y on X:

$$Y - 27 = -1.5(X - 53)$$

Simplify:

$$Y = 27 - 1.5(X - 53)$$

$$Y = 27 - 1.5X + 79.5$$

$$Y = 106.5 - 1.5X$$

Thus, the regression equation of Y on X is:

$$Y = 106.5 - 1.5X$$

Quick Tip

The regression equations of X on Y and Y on X are used to estimate the relationship between two variables. The coefficients b_{xy} and b_{yx} determine the steepness of the regression lines.

---

13. Find the derivative of $y = x^6 + 48x^3 + 24x - 16$ with respect to x.

Solution:

To find the derivative of the given function, we will apply the power rule for differentiation, which states that for $y = ax^n$, the derivative is $\frac{dy}{dx} = n \cdot ax^{n-1}$.

Given:

$$y = x^6 + 48x^3 + 24x - 16$$

We will differentiate each term of the function with respect to x.

$$\frac{dy}{dx} = \frac{d}{dx}(x^6) + \frac{d}{dx}(48x^3) + \frac{d}{dx}(24x) - \frac{d}{dx}(16)$$

Step 1: Differentiating each term.

1. The derivative of x^6 is:

$$\frac{d}{dx}(x^6) = 6x^5$$

2. The derivative of $48x^3$ is:

$$\frac{d}{dx}(48x^3) = 3 \times 48x^2 = 144x^2$$

3. The derivative of $24x$ is:

$$\frac{d}{dx}(24x) = 24$$

4. The derivative of -16 is:

$$\frac{d}{dx}(-16) = 0$$

Step 2: Combine the derivatives.

Now, combine all the derivatives:

$$\frac{dy}{dx} = 6x^5 + 144x^2 + 24$$

Thus, the derivative of y with respect to x is:

$$\boxed{\frac{dy}{dx} = 6x^5 + 144x^2 + 24}$$

Quick Tip

When differentiating terms with powers of x, apply the power rule: $\frac{d}{dx}(x^n) = n \cdot x^{n-1}$.

---

14. Write the important assumptions of ANOVA.

Solution:

Step 1: Independence of observations.

The observations should be independent of each other. This means that the value of one observation should not influence the value of another observation.

Step 2: Normality of the population.

The data in each group should be normally distributed. While ANOVA is fairly robust to deviations from normality, this assumption is important for valid results.

Step 3: Homogeneity of variances.

The variances in each of the groups being compared should be equal. This is also known as the assumption of homoscedasticity.

Step 4: Random sampling.

The samples should be randomly selected from the population to ensure that the results can be generalized to the larger population.

Quick Tip

The key assumptions of ANOVA include independence, normality, homogeneity of variances, and random sampling. Violating these assumptions can affect the accuracy of the test results.

---

15. A random sample from a population is given below. 35, 45, 40, 42, 39, 55 and 63. Obtain the moment estimator of the population mean.

Solution:

Step 1: Moment estimator of the mean.

The moment estimator for the population mean is the sample mean, which is calculated as the sum of all observations divided by the number of observations.

Step 2: Calculate the sample mean.

The data provided is: 35, 45, 40, 42, 39, 55, 63. The sample mean is:

$$\bar{x} = \frac{35 + 45 + 40 + 42 + 39 + 55 + 63}{7} = \frac{319}{7} = 45.57$$

Step 3: Conclusion.

The moment estimator of the population mean is the sample mean, which is approximately 45.57.

Quick Tip

The moment estimator of the population mean is simply the sample mean, which provides an unbiased estimate of the true population mean.

---

16. A random sample of 100 is taken from a population. The mean and standard deviation are respectively 76 and 12. Obtain a 99% confidence interval for the mean of the population.

Solution:

We are asked to find the 99% confidence interval for the population mean based on a sample. The formula for the confidence interval for the population mean when the sample size is large (n ¿ 30) is:

$$\text{Confidence Interval} = \bar{X} \pm Z \cdot \frac{\sigma}{\sqrt{n}}$$

Where: $\bar{X} = 76$ (sample mean) $\sigma = 12$ (population standard deviation) $n = 100$ (sample size) Z is the Z-value corresponding to a 99% confidence level.

Step 1: Find the Z-value.

For a 99% confidence level, the Z-value is approximately 2.576 (from standard Z-tables).

Step 2: Calculate the standard error.

The standard error is given by:

$$\text{Standard Error} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{100}} = \frac{12}{10} = 1.2$$

Step 3: Calculate the margin of error.

The margin of error is given by:

$$\text{Margin of Error} = Z \cdot \text{Standard Error} = 2.576 \cdot 1.2 = 3.0912$$

Step 4: Calculate the confidence interval.

Now, we can calculate the confidence interval:

$$\text{Confidence Interval} = 76 \pm 3.0912$$

Thus, the 99% confidence interval for the mean is:

$$(76 - 3.0912, 76 + 3.0912) = (72.91, 79.09)$$

Therefore, the 99% confidence interval for the population mean is approximately $\boxed{(72.91, 79.09)}$.

Quick Tip

To calculate a confidence interval for a population mean, use the sample mean, standard deviation, sample size, and the Z-value for the desired confidence level.

17. Equations of two lines of regression are $2x - 5y + 33 = 0$ and $3x - 9y - 108 = 0$. Which one is the regression line of y on x?

Solution:

We are given two equations of regression lines. The goal is to determine which one represents the regression line of y on x.

Step 1: Understand the form of regression equations.

The regression equation of y on x is typically in the form:

$$y = a + bx$$

Where b is the regression coefficient of y on x.

The equation of the regression line of x on y is in the form:

$$x = a + by$$

Where b is the regression coefficient of x on y.

Step 2: Solve the given equations.

The first equation is:

$$2x - 5y + 33 = 0$$

Rearrange it to get y in terms of x:

$$5y = 2x + 33$$

$$y = \frac{2}{5}x + \frac{33}{5}$$

This is of the form $y = a + bx$, which indicates that this equation represents the regression line of y on x.

The second equation is:

$$3x - 9y - 108 = 0$$

Rearrange it to get x in terms of y:

$$3x = 9y + 108$$

$$x = 3y + 36$$

This is of the form $x = a + by$, which indicates that this equation represents the regression line of x on y.

Step 3: Conclusion.

From the above analysis, we conclude that the first equation $2x - 5y + 33 = 0$ is the regression line of y on x.

> **Quick Tip**
>
> To distinguish between regression lines, remember that the regression line of y on x is in the form $y = a + bx$, and the regression line of x on y is in the form $x = a + by$.

---

18. In a distribution exactly normal, 7% of the items are under 35 and 89% are under 63. Find the mean and standard deviation of the distribution.

Solution:

Step 1: Understanding the normal distribution.

We are given a normal distribution with certain percentages under specific values. We need to use the z-scores to relate these percentages to the mean and standard deviation.

Step 2: Use of z-scores.

For a normal distribution: - The z-score formula is:

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

where x is the data value, μ is the mean, and σ is the standard deviation.

Step 3: Find the z-scores corresponding to the given percentages.

- For 7%, we use the z-score table to find that the z-score corresponding to 7% in the left tail is approximately $z = -1.475$. - For 89%, the z-score corresponding to 89% is approximately $z = 1.23$.

Step 4: Set up two equations.

From the z-score formula for both 35 and 63:

$$\frac{35 - \mu}{\sigma} = -1.475 \quad (1)$$

$$\frac{63 - \mu}{\sigma} = 1.23 \quad (2)$$

Step 5: Solve the system of equations.

- From equation (1), we get:

$$35 - \mu = -1.475\sigma \quad \Rightarrow \quad \mu = 35 + 1.475\sigma$$

- Substitute this value of μ into equation (2):

$$\frac{63 - (35 + 1.475\sigma)}{\sigma} = 1.23$$

Simplifying this equation, we can solve for σ, and then substitute back to find μ.

Step 6: Conclusion.

After solving, we find:

$$\mu = 50 \quad \text{and} \quad \sigma = 10$$

Quick Tip

In normal distribution problems, use z-scores to relate the given percentages to the mean and standard deviation. Use the z-score table for accurate values.
