Answer 1

The core difference between logistic and linear regression lies in the type of problem they address. Linear regression is used for \*\*regression problems\*\*, where the goal is to predict a continuous numerical outcome (e.g., house price, temperature). Logistic regression is used for \*\*classification problems\*\*, where the goal is to predict the probability of an instance belonging to a specific category (e.g., spam/not spam, disease present/absent). Logistic regression achieves this by modeling the probability using a sigmoid function, while linear regression directly models the relationship between input and output variables.

Answer 2

The sigmoid function is crucial in logistic regression because it maps the output of a linear model to a probability. The linear model can produce any real number, but probabilities must be between 0 and 1. The sigmoid function, defined as σ(x) = 1 / (1 + e^-x), achieves this.

Key mathematical properties:

\* \*\*Bounded Output:\*\* The exponential term in the denominator ensures that the output is always between 0 and 1.

\* \*\*Differentiability:\*\* The sigmoid function is differentiable, which is essential for training the model using gradient descent.

\* \*\*S-Shape:\*\* The S-shape allows for a smooth transition between the two extremes (0 and 1), making it suitable for modeling probabilities.

\* \*\*Symmetry:\*\* It's symmetric around 0.5, which is useful for binary classification.

Answer 3

Response: "The log-likelihood is a measure of how well the model's predicted probabilities fit the observed data. Maximizing it means finding the model parameters that make the observed data most likely. In logistic regression, we use log-likelihood because it's easier to work with mathematically than just the likelihood, and it avoids numerical issues."

ExplanationThis answer introduces the concept of log-likelihood as a measure of fit and explains that maximizing it is about finding parameters that make the observed data most likely. It also touches on the mathematical convenience and numerical stability of using log-likelihood. It's a good, solid answer.