Readiness Quiz: MXN 441- Advanced Statistical Machine Learning, S1, 2026

Q1. Which of the following best describes a vector dot product **a.b**?

A) The product of the lengths of two vectors
B) The matrix obtained from multiplying **a** and **b**
C) A scalar value representing the projection of one vector onto another
D) The determinant of the two vectors

Answer: C
The dot product returns a scalar and measures how much one vector goes in the direction of another.

Q2. Which of the following expressions gives the least-squares solution $\hat{\beta}$in linear regression, assuming Xis the design matrix and yis the response vector?

A) $\hat{\beta} = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}^T\mathbf{y}$
B) $\hat{\beta} = \mathbf{X}^{-1}\mathbf{y}$
C) $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$
D) $\hat{\beta} = \mathbf{X}\mathbf{y}$

Answer: C
This is the standard normal equation for ordinary least squares (OLS):

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

Q3. Which of the following Python code snippets correctly calculates the mean of a list of numbers?

A) `mean(numbers)`
B) `sum(numbers) / len(numbers)`
C) `np.average(numbers)`
D) Both B and C

Answer: D
Option B is the manual way, and option C uses NumPy.

Q4. You are given a dataset with 10,000 rows and 50 columns. Which of the following is NOT a typical preprocessing step?

A) Handling missing values
B) Normalizing numeric features
C) Sorting the dataset alphabetically
D) Encoding categorical variables

Answer: C
Sorting alphabetically is not usually meaningful for modeling. The others are common preprocessing steps.

Q5. What does the p-value represent in hypothesis testing?

A) The probability that the null hypothesis is true
B) The probability of observing your data if the null hypothesis were true
C) The effect size of your experiment
D) The significance level of the test

Answer: B
The p-value measures how likely your data is under the assumption that the null hypothesis is true it does **not** tell you the probability that the null itself is true.

Q6. Which of the following best describes the bias-variance tradeoff in machine learning?

A) Increasing model complexity reduces both bias and variance.
B) High bias models usually overfit the data.
C) High variance models are likely to underfit the data.
D) Reducing bias usually increases variance and vice versa.

Answer: D

There is a fundamental trade-off: simpler models tend to have high bias and low variance (underfitting), while complex models tend to have low bias and high variance (overfitting). Good generalization requires balancing these two.