

PRACTICE QUESTIONS

SAS

A00-480

Applied Statistics for Machine Learning

SHARED BY [u/nifty_widget_81](#) on aiexamcerts

EDITION Free Set 2

QUESTIONS 10

SHARED October 15, 2026

EXAM SAS A00-480, Applied Statistics for Machine Learning

Shared by [u/nifty_widget_81](#) on aiexamcerts

10 practice questions for the SAS A00-480 exam, Applied Statistics for Machine Learning, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/sas/a00-480. Spotted a mistake? Email support@aiexamcerts.com with the exam code A00-480 and the question number.

1. A bootstrap distribution of a sample statistic is built to estimate its standard error when the theoretical distribution is unknown. Which step describes the bootstrap correctly?
 - A Add a fixed constant to each observation until the distribution becomes normal
 - B Remove the largest observations until the mean is stable
 - C Resample the observed data with replacement many times and recompute the statistic each time
 - D Divide the sample into two halves and report only the first half

Answer: C

The bootstrap draws resamples with replacement from the observed data, recomputes the statistic on each one, and uses the spread of those values as an estimate of its standard error. Adding a constant to make the data normal changes the problem and is not a bootstrap step. Trimming extreme observations until the mean settles biases the estimate, which is the tempting distractor.

2. On a dataset where only 1% of cases are positive, a classifier that always predicts negative can reach 99% accuracy while being useless for finding the positive cases.
 - A True
 - B False

Answer: A (True)

Accuracy rewards the majority class, so a trivial model that never flags a positive case still scores 99 percent. This is why precision, recall, F1 or the area under the precision-recall curve should be reported for imbalanced problems. The statement is true.

3. A regression analyst wants to test whether a categorical predictor with four levels contributes to the model as a whole rather than checking each level separately. Which test is appropriate?
 - A A paired t test on the first two levels only
 - B A partial F test comparing the reduced and full models
 - C A chi-square goodness-of-fit test on the response
 - D A Kolmogorov-Smirnov test on the residuals

Answer: B

A partial F test compares the model with and without all the indicator variables for the factor, which tests the factor as one term. A paired t test is designed for matched observations and does not address a multi-level predictor. The residual test checks distributional shape, which is the most tempting distractor because it is a model check, not a term test.

-
4. A variance inflation factor above about 10 for a predictor signals strong _____ among the predictors in a regression model.

Answer: multicollinearity (also accepted: collinearity)

The variance inflation factor measures how much the variance of a coefficient is inflated by correlation with the other predictors, so large values indicate multicollinearity. Heteroscedasticity refers to unequal residual variance and is the most tempting alternative, but VIF does not measure it. Autocorrelation concerns residuals ordered in time.

5. Which two statements about the normal distribution are correct? (Choose two.)

- A About 68% of values lie within one standard deviation of the mean
- B The mean and the median are always different
- C Its skewness is zero
- D Its excess kurtosis is always negative
- E It is defined only for positive values

Answer: A, C

For a normal distribution about 68 percent of values fall within one standard deviation of the mean, and the distribution is symmetric, so its skewness is zero. The mean and the median coincide for a normal distribution, which makes the claim that they are always different false. The normal distribution covers the whole real line, and its excess kurtosis is zero rather than negative, so the remaining options are also false.

6. Which technique reduces overfitting in a decision tree by pruning branches whose contribution to validation accuracy is negligible?

- A Principal component rotation
- B Bootstrap resampling of predictors
- C Cost-complexity pruning
- D Target encoding of the response

Answer: C

Cost-complexity pruning removes subtrees whose complexity is not justified by their error reduction, which controls overfitting in a tree. Principal component rotation is a dimension reduction method for predictors, not a pruning method. Resampling predictors and target encoding are separate preprocessing ideas, and target encoding can itself leak information if done carelessly.

7. Which criterion balances goodness of fit against the number of parameters, and is used in SAS model selection to prefer simpler models among candidates?
- A Coefficient of determination
 - B Pearson correlation
 - C Akaike information criterion
 - D Mean of the residuals

Answer: C

The Akaike information criterion rewards fit while penalizing the number of parameters, so it favors simpler models among candidates. The R-squared always increases when predictors are added, which makes it the most tempting distractor for model comparison. Pearson correlation measures association between two variables, and the mean of the residuals is zero for a least squares fit with an intercept.

8. Choosing a model by the lowest error on the data used to fit it is a reliable way to estimate how the model will perform on new data.
- A True
 - B False

Answer: B (False)

Training error is optimistic because the model was tuned to those same observations, so it usually underestimates the error on new data. Held-out data, cross-validation or a test set are needed for an honest estimate. The statement is false.

9. Two variables have a Pearson correlation of zero. Which statement is always correct?
- A The two variables are statistically independent
 - B There is no linear association between them
 - C Their covariance is one
 - D Their means are equal

Answer: B

A zero Pearson correlation means there is no linear association, but a nonlinear relationship such as a quadratic can still exist. Statistical independence is a stronger condition, which makes the independence claim the most tempting but incorrect one. Covariance of one and equal means have no connection with a zero correlation.

10. Which technique reduces many correlated numeric predictors to a smaller set of uncorrelated components that capture most of their variance?
- A Stepwise logistic regression
 - B Principal component analysis
 - C Kaplan-Meier estimation
 - D Hierarchical clustering with complete linkage

Answer: B

Principal component analysis finds orthogonal directions ordered by the variance they explain, so a few components can summarize many correlated predictors. Stepwise logistic regression selects original predictors rather than building new components. Kaplan-Meier estimation is a survival method, and hierarchical clustering groups observations rather than reducing predictors.