

01	☒ A	☐ B	☐ C	☐ D	☐ E
02	☐ A	☐ B	☒ C	☐ D	☐ E
03	☐ A	☐ B	☐ C	☐ D	☒ E
04	☐ A	☐ B	☒ C	☐ D	☐ E
05	☐ A	☐ B	☐ C	☐ D	☒ E
06	☐ A	☐ B	☒ C	☐ D	☐ E
07	☐ A	☐ B	☒ C	☐ D	☐ E
08	☐ A	☐ B	☒ C	☐ D	☐ E
09	☐ A	☐ B	☒ C	☐ D	☐ E
10	☐ A	☒ B	☐ C	☐ D	☐ E
11	☐ A	☐ B	☐ C	☐ D	☒ E
12	☐ A	☐ B	☐ C	☐ D	☒ E
13	☐ A	☐ B	☒ C	☐ D	☐ E

14	☒ A	☐ B	☐ C	☐ D	☐ E
15	☐ A	☐ B	☐ C	☒ D	☐ E
16	☒ A	☐ B	☐ C	☐ D	☐ E
17	☐ A	☐ B	☒ C	☐ D	☐ E
18	☐ A	☒ B	☐ C	☐ D	☐ E
19	☐ A	☐ B	☐ C	☐ D	☒ E
20	☐ A	☐ B	☐ C	☐ D	☒ E
21	☐ A	☐ B	☐ C	☐ D	☒ E
22	☐ A	☐ B	☐ C	☐ D	☒ E
23	☐ A	☐ B	☐ C	☐ D	☒ E
24	☒ A	☐ B	☐ C	☐ D	☐ E
25	☒ A	☐ B	☐ C	☐ D	☐ E
26	☐ A	☒ B	☐ C	☐ D	☐ E

27	☐ A	☐ B	☐ C	☐ D	☒ E
28	☐ A	☐ B	☐ C	☐ D	☒ E
29	☐ A	☐ B	☐ C	☒ D	☐ E
30	☐ A	☐ B	☐ C	☐ D	☒ E
31	☐ A	☐ B	☐ C	☐ D	☒ E
32	☐ A	☐ B	☒ C	☐ D	☐ E
33	☐ A	☐ B	☐ C	☒ D	☐ E
34	☐ A	☐ B	☐ C	☐ D	☒ E
35	☒ A	☐ B	☐ C	☐ D	☐ E
36	☐ A	☐ B	☐ C	☐ D	☒ E
37	☐ A	☐ B	☒ C	☐ D	☐ E
38	☒ A	☐ B	☐ C	☐ D	☐ E
39	☒ A	☐ B	☐ C	☐ D	☐ E

PRACTICE QUESTIONS

Society of Actuaries (SOA)

PADEM

Predictive Analytics - Demo

EDITION

Demo

QUESTIONS

9

FORMAT

Limited Edition PDF

Before you begin

What this is

9 practice questions for Society of Actuaries (SOA) **PADEM**, each with its answer key and an explanation. Answer a question before you read its key — the explanation is worth more once you have committed.

If something looks wrong

Send us three things and we will check it:

- the exam code, **PADEM**
- the question number
- the email address on your order

Write to **support@mometrix.net**. We reply within one business day.

QUESTION 1

SINGLE CHOICE

In a Poisson regression model used to model insurance claim counts, which component represents the natural logarithm of the expected claim frequency?

- ☒ **A The link function**
- ☐ B The dispersion parameter
- ☐ C The deviance residual
- ☐ D The offset variable

ANSWER A

WHY In a Poisson GLM with a log link, $g(\mu) = \log(\mu)$, so $\log(\mu) = X\beta$. The link function is what transforms the expected value of the response into the linear predictor; for a log-linked Poisson model, that transformation is the natural logarithm of the expected claim frequency. The offset is an additive term on the linear-predictor scale (also log-scale), not the transformation itself.

QUESTION 2

SINGLE CHOICE

Which measure is most appropriate for comparing the predictive accuracy of two competing regression models fit on the same dataset when the response variable is continuous?

- ☐ A The Akaike information criterion (AIC)
- ☒ **B The coefficient of determination, R-squared**
- ☐ C The Gini coefficient
- ☐ D The likelihood ratio chi-square statistic

ANSWER B

WHY For continuous response variables, the coefficient of determination (R-squared) directly measures the proportion of variance in the response explained by the model and is the standard metric for comparing predictive accuracy across competing regressions on the same data. AIC is useful for model selection but penalizes complexity rather than measuring predictive fit directly; the Gini coefficient is used for ranking/classification problems; the likelihood ratio chi-square applies to nested model comparisons.

QUESTION 3 SINGLE CHOICE

In a classification tree built using the CART algorithm, what criterion is used at each split to determine the best partitioning variable and split point?

- ☐ A Likelihood ratio chi-square
- ☐ B Akaike information criterion
- ☒ C Gini impurity decrease
- ☐ D Hosmer-Lemeshow statistic

ANSWER C

WHY The CART algorithm uses Gini impurity (or entropy, in some variants) as the splitting criterion: at each node it selects the variable and split point that produce the largest decrease in Gini impurity. Likelihood ratio chi-square is used in model-based trees and some inferential procedures, AIC is a model-comparison criterion for parametric models, and Hosmer-Lemeshow is a goodness-of-fit test for logistic regression.

QUESTION 4 SINGLE CHOICE

In logistic regression, the odds ratio for a one-unit increase in a predictor x_k , holding all other predictors constant, is computed as:

- ☒ A $\exp(\beta_k)$
- ☐ B β_k
- ☐ C $1 / (1 + \exp(-\beta_k))$
- ☐ D $\beta_k \times \sigma(x)$

ANSWER A

WHY For a logistic regression model $\text{logit}(p) = X\beta$, the odds ratio comparing two observations that differ by one unit in x_k (with other predictors equal) is $\exp(\beta_k)$. The link inverse $1/(1 + \exp(-\beta_k))$ is the predicted probability, not an odds ratio, and $\sigma(x)$ is not part of standard logistic regression.

QUESTION 5

SINGLE CHOICE

When applying k-fold cross-validation to assess a predictive model, what is the principal purpose of this procedure?

- ☐ A To reduce the variance of the estimated regression coefficients
- ☒ B To estimate the out-of-sample predictive performance of the model
- ☐ C To test the null hypothesis that all regression coefficients are zero
- ☐ D To satisfy the normality assumption of the residuals

ANSWER B

WHY K-fold cross-validation partitions the data into k folds, fits the model on k-1 folds, and evaluates it on the held-out fold, repeating this to obtain an estimate of out-of-sample error. This gives a less biased estimate of predictive performance than training error alone. It does not reduce coefficient variance, test a global null, or enforce residual normality.

QUESTION 6

SINGLE CHOICE

In an ARIMA(p,d,q) model, what role does the differencing order d play?

- ☐ A It specifies the number of lagged forecast errors in the model
- ☐ B It specifies the number of lagged values of the response in the model
- ☒ C It specifies the number of times the series is differenced to achieve stationarity
- ☐ D It specifies the order of the seasonal component of the model

ANSWER C

WHY The d in ARIMA(p,d,q) is the order of differencing applied to the series before fitting the AR and MA parts, used to remove trends and induce stationarity. p is the autoregressive order (lagged response values) and q is the moving-average order (lagged forecast errors). Seasonal effects are handled by seasonal ARIMA components, not by d.

QUESTION 7

SINGLE CHOICE

In principal component analysis (PCA) applied to a set of standardized variables, the first principal component is defined as:

- ☒ **A** The linear combination of the variables with maximum variance
- ☐ **B** The variable with the highest correlation with the response
- ☐ **C** The variable that explains the most of the total variance by itself
- ☐ **D** The linear combination that minimizes the residual sum of squares

ANSWER A

WHY PCA finds orthogonal linear combinations of the variables such that the first principal component accounts for the largest possible variance in the data, each subsequent component accounts for the largest remaining variance subject to being orthogonal to prior components. It is an unsupervised technique, so it does not involve a response variable, and the components are linear combinations rather than individual variables.

QUESTION 8

SINGLE CHOICE

In k-means clustering, what is the effect of increasing the number of clusters k on the within-cluster sum of squares (WCSS)?

- ☐ **A** WCSS always increases as k increases
- ☒ **B** WCSS always decreases as k increases and equals zero only when k equals the number of observations
- ☐ **C** WCSS is unchanged because k-means is invariant to k
- ☐ **D** WCSS is unaffected because k-means minimizes between-cluster sum of squares

ANSWER B

WHY In k-means, the within-cluster sum of squares is a non-increasing function of k . Adding more clusters can only partition the points into smaller groups, reducing WCSS, and reaches zero when each point is its own cluster ($k = n$). Methods such as the elbow rule exploit the rate of decrease to choose an appropriate k .

QUESTION 9

MULTIPLE CHOICE

When preparing a dataset for predictive modeling, which of the following are common data preprocessing steps? (Choose two.)

- ☒ **A** Standardizing numeric variables to have mean zero and standard deviation one
- ☒ **B** Recoding missing values using an indicator variable plus an imputation method
- ☐ **C** Maximizing the within-cluster sum of squares before fitting k-means
- ☐ **D** Adding the response variable to the set of predictors to improve fit

ANSWER **A** • **B**

WHY Standardizing numeric predictors (option A) is a standard preprocessing step that puts variables on comparable scales and is essential for regularized and distance-based models. Treating missing values explicitly through an indicator plus imputation (option B) is also a recognized preprocessing practice. Maximizing within-cluster sum of squares is not a goal of k-means (it is minimized), and adding the response variable to the predictors causes data leakage and overfitting.