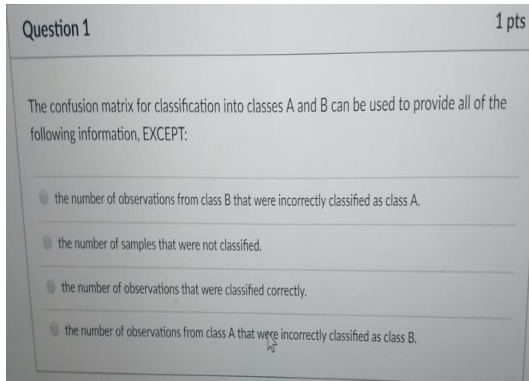
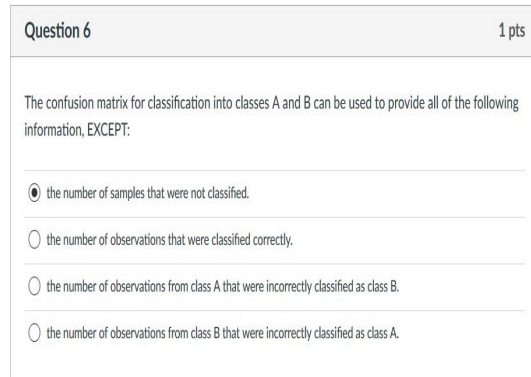


Quiz Review

Question 1 - Confusion Matrix



(a) Current Quiz

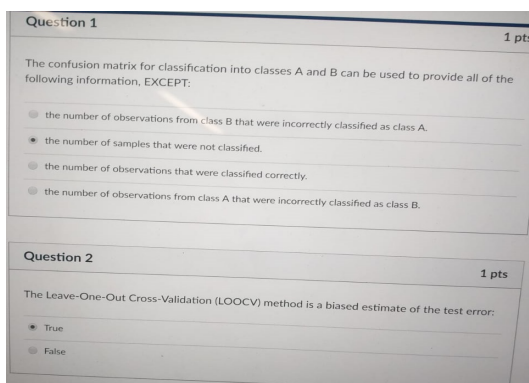


(b) Prior Quiz

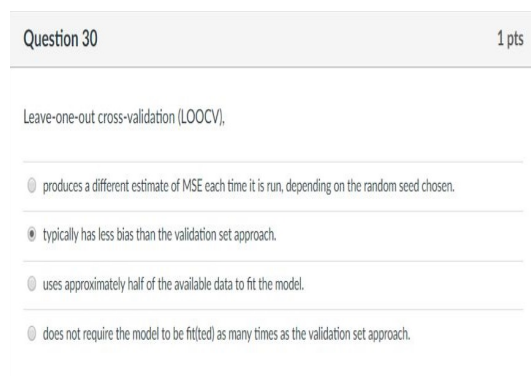
Solution: (2) = 'the number of samples...'

FAITH: 9/10 NOTES: Confusion Matrices show how well classification occurred (correct + incorrect)

Question 2 - LOOCV (Leave-One-Out Cross-Validation)



(a) Current Quiz

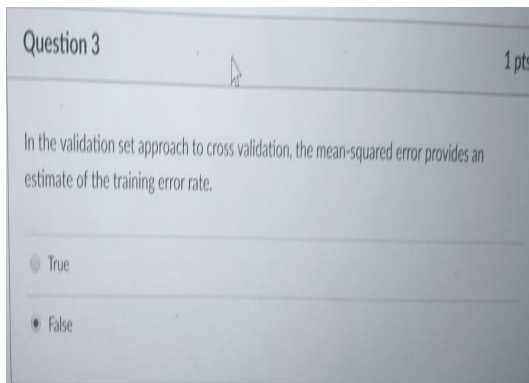


(b) Prior Quiz

Solution: TRUE

NOTES: Note the older quiz answer of 'less bias' but not '0 bias'.

Question 3 - MSE (Mean Squared Error)



(a) Current Quiz

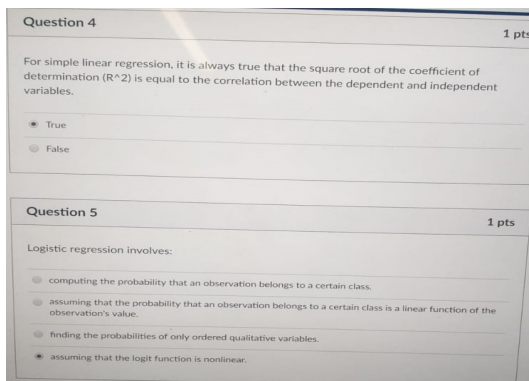


(b) Prior Quiz

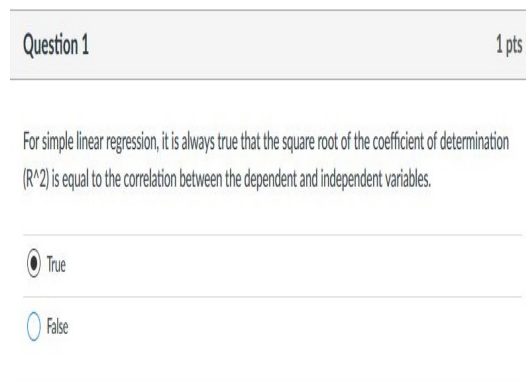
Solution: FALSE

NOTES:

Question 4 - R-Squared (Linear Regression)



(a) Current Quiz

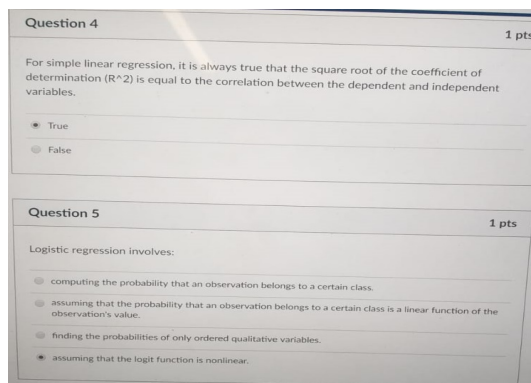


(b) Prior Quiz

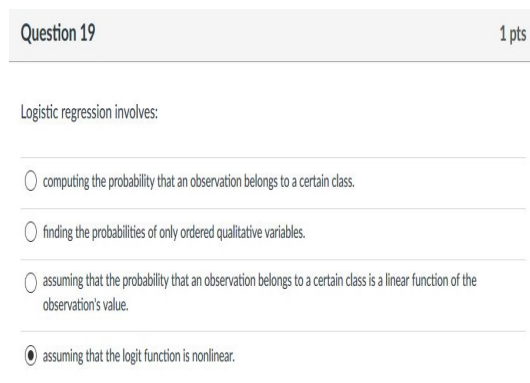
Solution: TRUE

NOTES: N/A

Question 5 - Logistic Regression



(a) Current Quiz

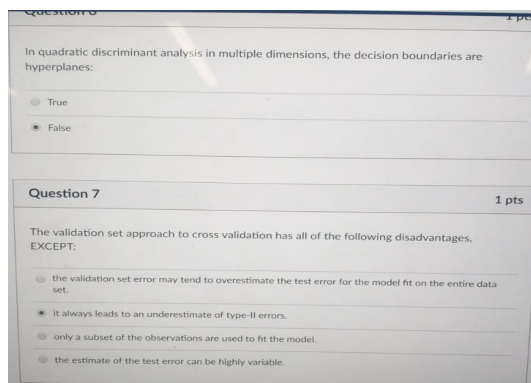


(b) Prior Quiz

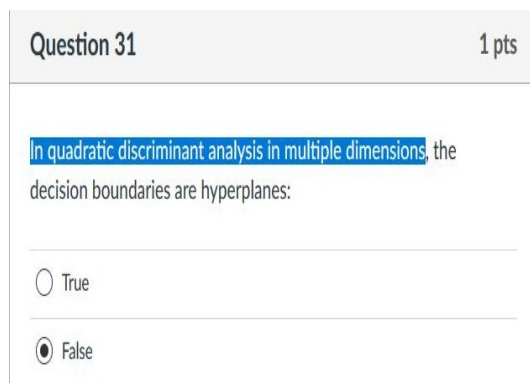
Solution: (4) = 'assuming non-linear...'

NOTES: N/A

Question 6 - QDA (Quadratic Discriminant Analysis)



(a) Current Quiz

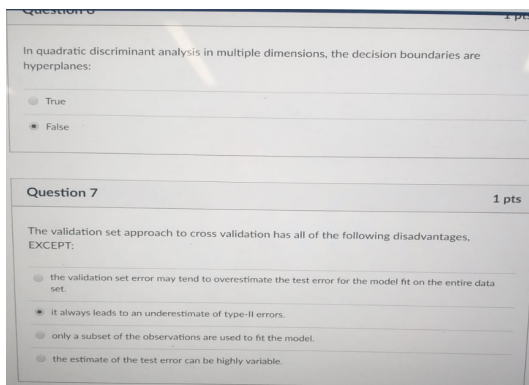


(b) Prior Quiz

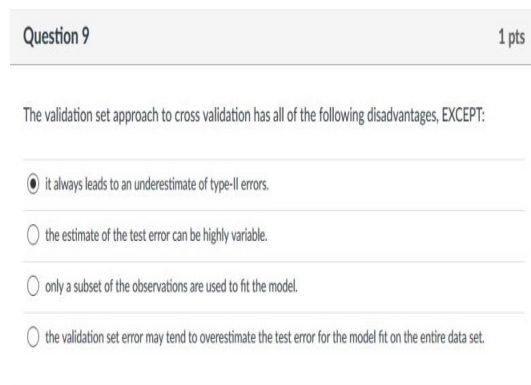
Solution: FALSE

NOTES: N/A

Question 7 - CV (Validation set approach)



(a) Current Quiz

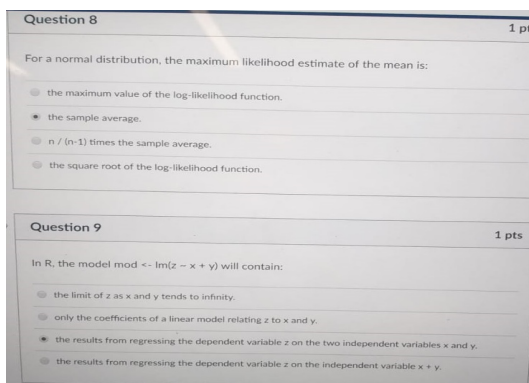


(b) Prior Quiz

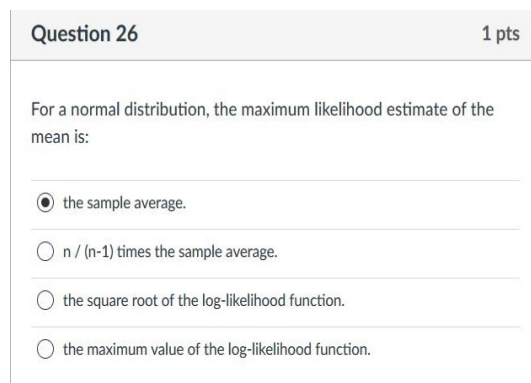
Solution: (2) = '...underestimate type-II errors'

NOTES: N/A

Question 8 - MLE (Normal Distribution)



(a) Current Quiz

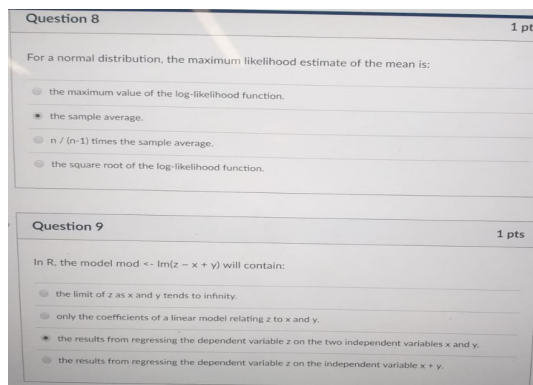


(b) Prior Quiz

Solution: (2) = 'sample average'

NOTES: N/A

Question 9 - R Code ($\text{mod} = \text{lm}(z \sim x + y)$)



(a) Current Quiz

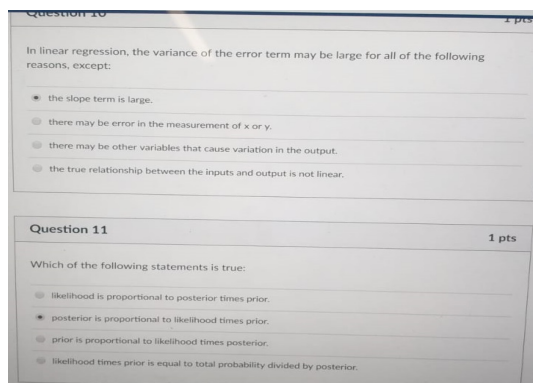


(b) Prior Quiz

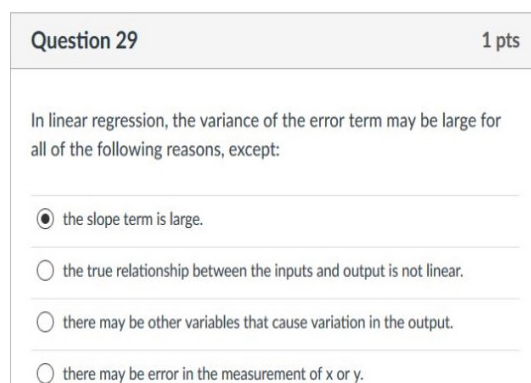
Solution: (3) = '... z on the two independent variables'

NOTES: N/A

Question 10 - Variance (Linear Regression)



(a) Current Quiz

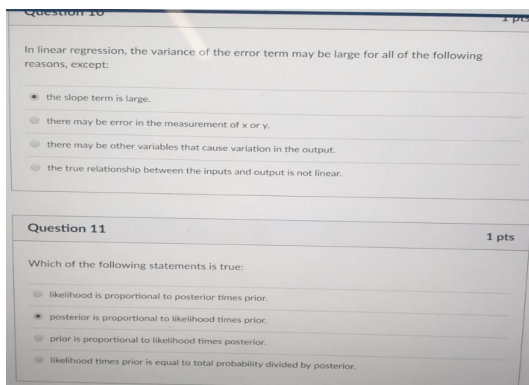


(b) Prior Quiz

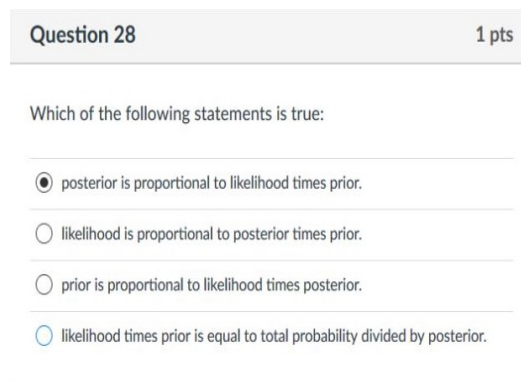
Solution: (1) = 'slope term is large'

NOTES: N/A

Question 11 - Posteriors & likely-hood



(a) Current Quiz

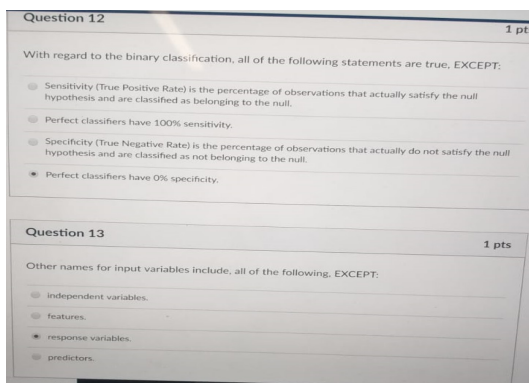


(b) Prior Quiz

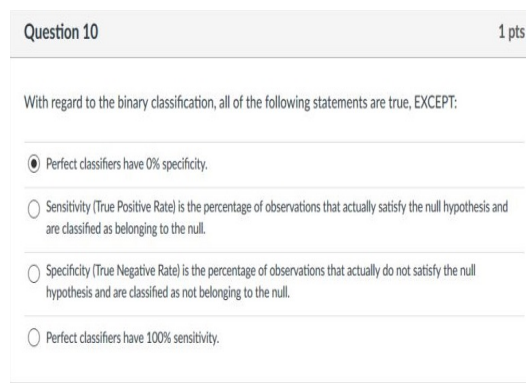
Solution: (2) = 'posterior is proportional to likely-hood times prior'

NOTES: N/A

Question 12 - Binary Classification



(a) Current Quiz

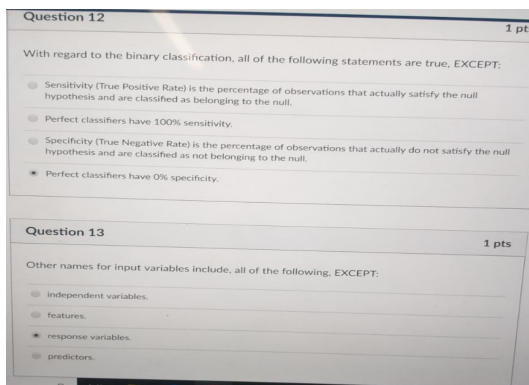


(b) Prior Quiz

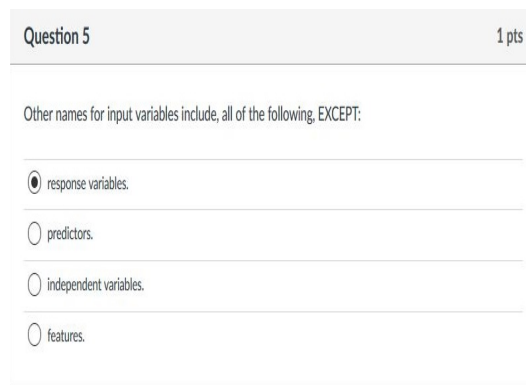
Solution: (4) = '... 0% specificity'

NOTES: N/A

Question 13 - Names of Input Variables



(a) Current Quiz

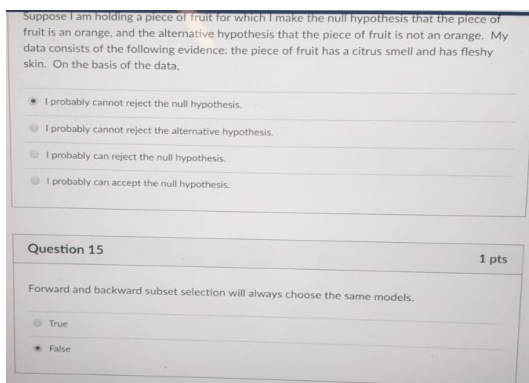


(b) Prior Quiz

Solution: (3) = 'response'

NOTES: N/A

Question 14 - Hypotheses (with fruit)



(a) Current Quiz

?

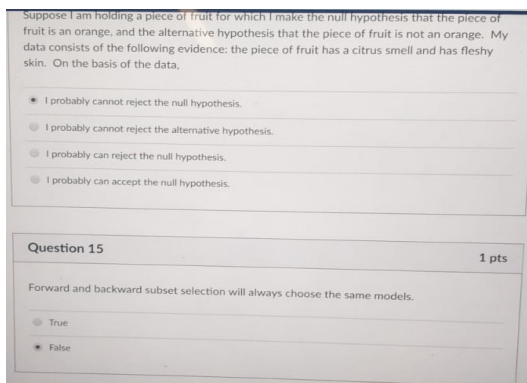
NO OLD REFERENCE

(b) Prior Quiz

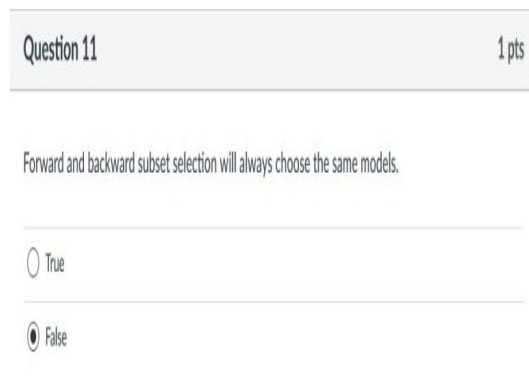
Solution: (1) = '... cannot reject NULL'

NOTES:

Question 15 - Subset Selection (Models & Data)



(a) Current Quiz

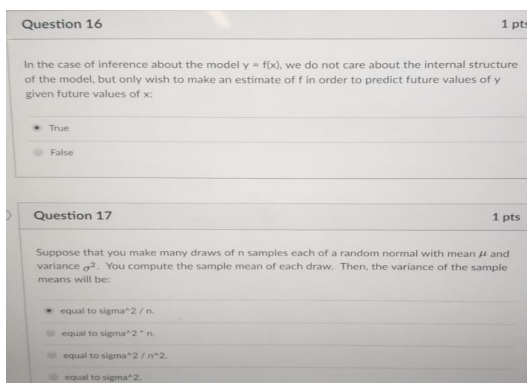


(b) Prior Quiz

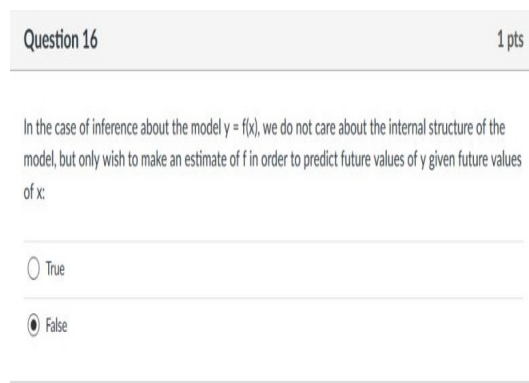
Solution: FALSE

NOTES: N/A

Question 16 - Inference and model structure



(a) Current Quiz



(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 17 - Sample Mean Variance

Question 16 1 pts

In the case of inference about the model $y = f(x)$, we do not care about the internal structure of the model, but only wish to make an estimate of f in order to predict future values of y given future values of x :

☒ True
☐ False

Question 17 1 pts

Suppose that you make many draws of n samples each of a random normal with mean μ and variance σ^2 . You compute the sample mean of each draw. Then, the variance of the sample means will be:

☒ equal to σ^2 / n .
☐ equal to $\sigma^2 \cdot n$.
☐ equal to σ^2 / n^2 .
☐ equal to σ^2 .

(a) Current Quiz

Question 14 1 pts

Suppose that you make many draws of n samples each of a random normal with mean μ and variance σ^2 . You compute the sample mean of each draw. Then, the variance of the sample means will be:

☒ equal to σ^2 / n .
☐ equal to σ^2 / n^2 .
☐ equal to $\sigma^2 \cdot n$.
☐ equal to σ^2 .

(b) Prior Quiz

Solution: (1) = ' σ^2 / n '

NOTES: N/A

Question 18 - R-Squared as determinant

Question 18 1 pts

The coefficient of determination (R^2) alone is a sufficient statistic to use for model order selection.

☐ True
☒ False

Question 19 1 pts

In linear discriminant analysis, the covariance matrices of the data in all classes are assumed to be equal to each other:

☒ True
☐ False

(a) Current Quiz

Question 8 1 pts

The coefficient of determination (R^2) alone is a sufficient statistic to use for model order selection.

☐ True
☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 19 - LDA (Linear Discriminant Analysis)

Question 18 1 pts

The coefficient of determination (R^2) alone is a sufficient statistic to use for model order selection.

☐ True

☒ False

Question 19 1 pts

In linear discriminant analysis, the covariance matrices of the data in all classes are assumed to be equal to each other:

☒ True

☐ False

(a) Current Quiz

Question 12 1 pts

In linear discriminant analysis, the covariance matrices of the data in all classes are assumed to be equal to each other:

☒ True

☐ False

(b) Prior Quiz

Solution: TRUE

NOTES: N/A

Question 20 - Parametric models

Question 20 1 pts

In parametric statistical learning methods, we do not make explicit assumptions about the functional form of f . Instead, we seek some description of f that fits the data points well.

☐ True

☒ False

Question 21 1 pts

Non-parametric statistical learning methods have a major disadvantage in that they need a very large number of observations to obtain an accurate estimate for f because they do not reduce the problem of estimating f to a small number of parameters.

☒ True

☐ False

(a) Current Quiz

Question 3 1 pts

In parametric statistical learning methods, we do not make explicit assumptions about the functional form of f . Instead, we seek some description of f that fits the data points well.

☐ True

☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 21 - Non-parametric models

Question 20 1 pts

In parametric statistical learning methods, we do not make explicit assumptions about the functional form of f . Instead, we seek some description of f that fits the data points well.

☐ True

☒ False

Question 21 1 pts

Non-parametric statistical learning methods have a major disadvantage in that they need a very large number of observations to obtain an accurate estimate for f because they do not reduce the problem of estimating f to a small number of parameters.

☒ True

☐ False

(a) Current Quiz

?

NO OLD REFERENCE

(b) Prior Quiz

Solution: ? change to FALSE

NOTES:

<https://www.investopedia.com/terms/n/nonparametric-statistics.asp>

<https://www.analyticsvidhya.com/blog/2017/11/a-guide-to-conduct-analysis-using-non-parametric-tests/>

<https://stats.stackexchange.com/questions/163915/why-would-parametric-statistics-ever-be-preferred-over-nonpar>

Question 22 - K-fold CV

Question 22 1 p

In k-fold Cross-Validation, the observations are separated into multiple validation sets each with k observations.

☐ True

☒ False

Question 23 1 pts

The R^2 statistic has all of the following advantages, EXCEPT:

☒ it is measured in units of x_i .

☐ it (almost) always has a value between 0 and 1.

☐ it is independent of the scale of y .

☐ it is dimensionless.

(a) Current Quiz

Question 27 1 pts

In k-fold Cross-Validation, the observations are separated into multiple validation sets each with k observations.

☐ True

☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 23 - R-Squared advantages

Question 22 1 pt

In k-fold Cross-Validation, the observations are separated into multiple validation sets each with k observations.

☐ True

☒ False

Question 23 1 pts

The R^2 statistic has all of the following advantages, EXCEPT:

☒ it is measured in units of x.

☐ it (almost) always has a value between 0 and 1.

☐ it is independent of the scale of y.

☐ it is dimensionless.

(a) Current Quiz

Question 21 1 pts

The R^2 statistic has all of the following advantages, EXCEPT:

☒ it is measured in units of x.

☐ it is dimensionless.

☐ it is independent of the scale of y.

☐ it (almost) always has a value between 0 and 1.

(b) Prior Quiz

Solution: (1) = 'measured in units of x'

NOTES: N/A

Question 24 - Bayes Classifier

Question 24 1 pts

The Bayes classifier is known to have the lowest total error rate of all classifiers. Therefore, it is not possible to achieve a lower type-I error at the expense of a higher type-II error.

☐ True

☒ False

Question 25 1 pts

The population regression line will always lie on the least squares line.

☐ True

☒ False

(a) Current Quiz

Question 2 1 pts

The Bayes classifier is known to have the lowest total error rate of all classifiers. Therefore, it is not possible to achieve a lower type-I error at the expense of a higher type-II error.

☐ True

☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 25 - Population Regression Line

Question 24 1 pts

The Bayes classifier is known to have the lowest total error rate of all classifiers. Therefore, it is not possible to achieve a lower type-I error at the expense of a higher type-II error.

☐ True

☒ False

Question 25 1 pts

The population regression line will always lie on the least squares line.

☐ True

☒ False

(a) Current Quiz

Question 4 1 pts

The population regression line will always lie on the least squares line.

☐ True

☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A

Question 26 - MSE (equation checking)

Question 26 1 pts

Under suitable assumptions, mean squared error (MSE) is equal to bias squared plus variance plus the variance of the noise. In estimating f , we attempt to do all of the following, EXCEPT:

☒ minimize the noise variance.

☐ minimize the bias squared.

☐ minimize the mean squared error.

Question 27 1 pts

All of the following statements about Leave-One-Out Cross-Validation (LOOCV) are true, EXCEPT:

☐ In each application of the training method, a single observation is used for the validation set.

☐ LOOCV can be computationally expensive to implement because one must fit the model n times.

☐ the statistical learning method is fit on $n-1$ training observations.

(a) Current Quiz

Question 25 1 pts

Under suitable assumptions, mean squared error (MSE) is equal to bias squared plus variance plus the variance of the noise. In estimating f , we attempt to do all of the following, EXCEPT:

☒ minimize the noise variance.

☐ minimize the bias squared.

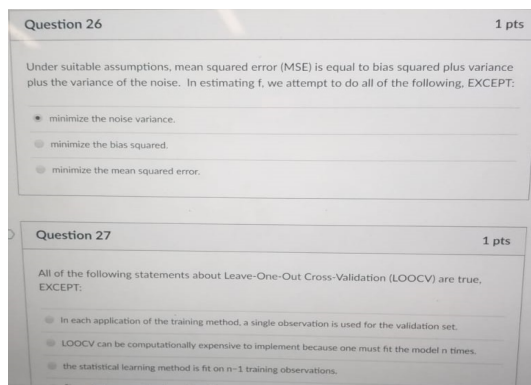
☐ minimize the mean squared error.

(b) Prior Quiz

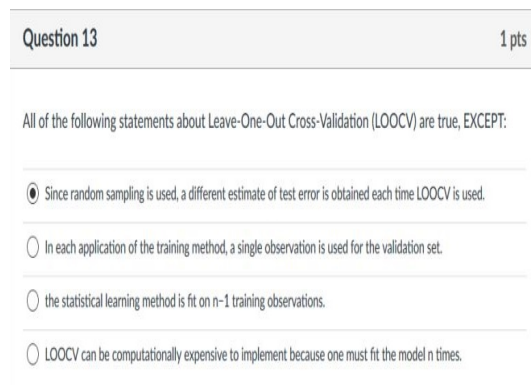
Solution: (1) = 'minimize noise variance'

NOTES: N/A

Question 27 - LOOCV (all true EXCEPT...)



(a) Current Quiz

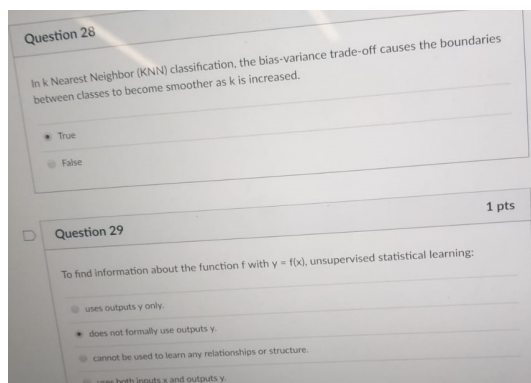


(b) Prior Quiz

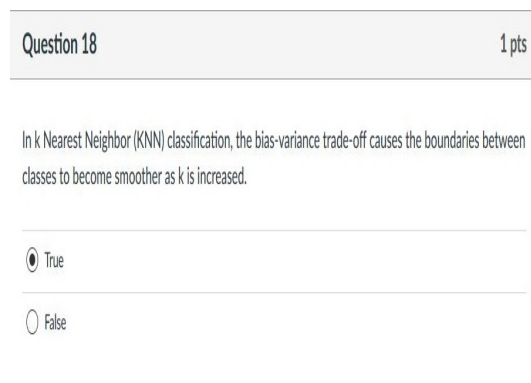
Solution: ??? - Possibly 1 from prior exam

NOTES: CHECK!! User did not submit question with answer. *Possible fix: 6/4

Question 28 - KNN (K Nearest Neighbor)



(a) Current Quiz

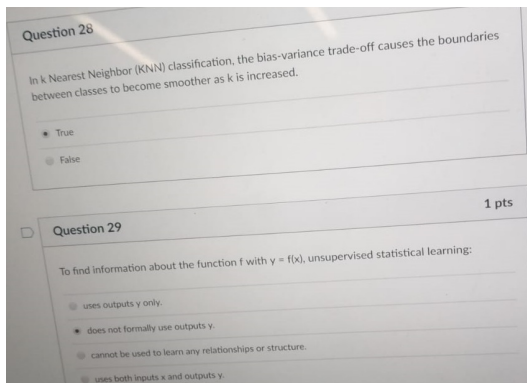


(b) Prior Quiz

Solution: TRUE

NOTES: N/A

Question 29 - Info on $y = f(x)$ (unsupervised learning)



(a) Current Quiz

?

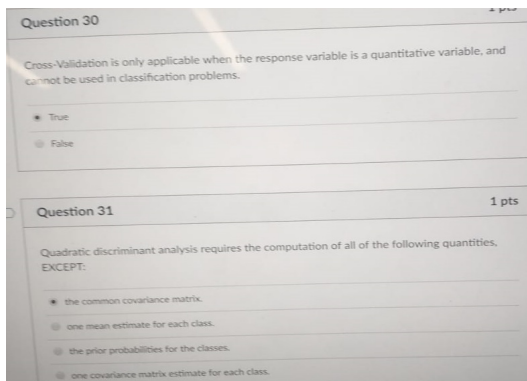
NO OLD REFERENCE

(b) Prior Quiz

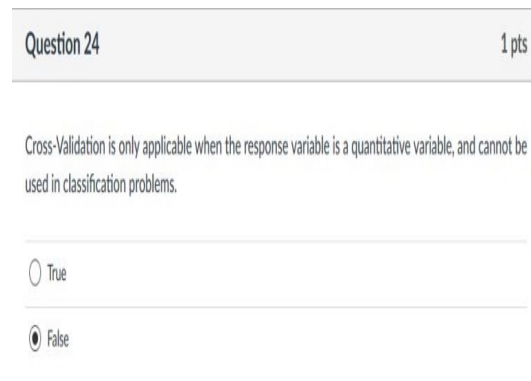
option 2 - use x not y (online research)

NOTES:

Question 30 - CV use for numeric/classification



(a) Current Quiz

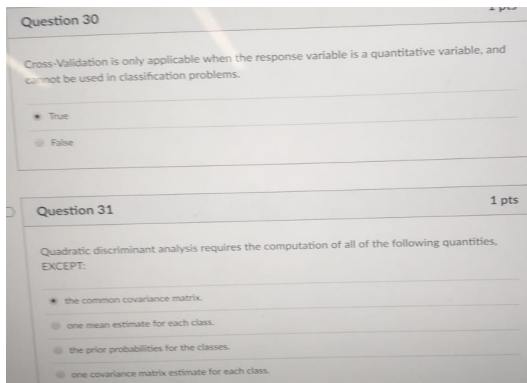


(b) Prior Quiz

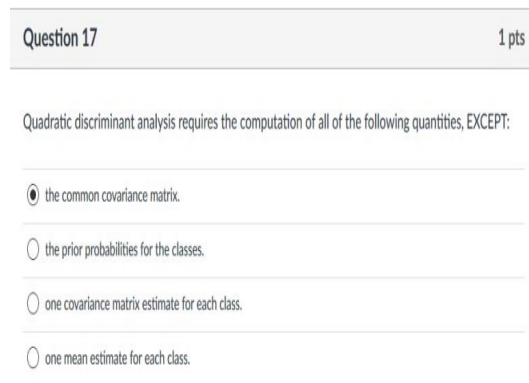
Solution: FALSE

NOTES:

Question 31 - QDA (Quadratic Discriminant Analysis)



(a) Current Quiz

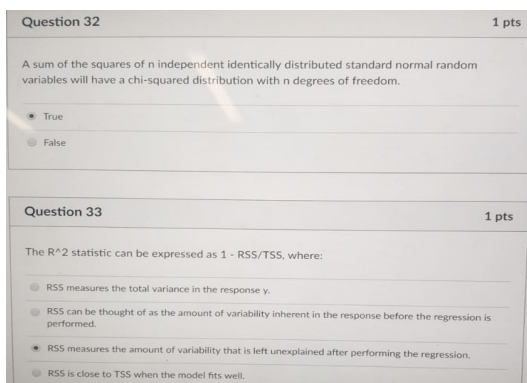


(b) Prior Quiz

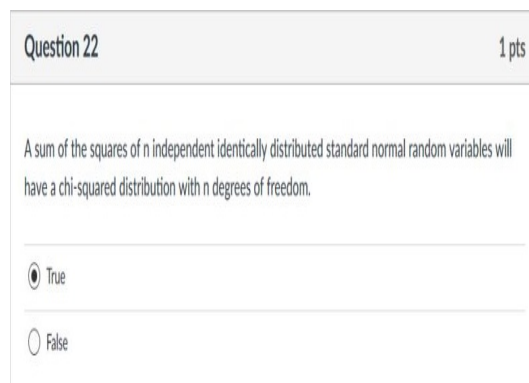
Solution: (1) = 'common covariance matrix'

NOTES: N/A

Question 32 - Sum of Squares (Chi-squared distribution)



(a) Current Quiz



(b) Prior Quiz

Solution: TRUE

NOTES: N/A

Question 33 - R-Squared equation expression

Question 32 1 pts

A sum of the squares of n independent identically distributed standard normal random variables will have a chi-squared distribution with n degrees of freedom.

☒ True
☐ False

Question 33 1 pts

The R^2 statistic can be expressed as $1 - \text{RSS}/\text{TSS}$, where:

☐ RSS measures the total variance in the response y .
☐ RSS can be thought of as the amount of variability inherent in the response before the regression is performed.
☒ RSS measures the amount of variability that is left unexplained after performing the regression.
☐ RSS is close to TSS when the model fits well.

(a) Current Quiz

Question 20 1 pts

The R^2 statistic can be expressed as $1 - \text{RSS}/\text{TSS}$, where:

☒ RSS measures the amount of variability that is left unexplained after performing the regression.
☐ RSS measures the total variance in the response y .
☐ RSS can be thought of as the amount of variability inherent in the response before the regression is performed.
☐ RSS is close to TSS when the model fits well.

(b) Prior Quiz

Solution: (3) = 'RSS ... explains ... variability ... regression'

NOTES: N/A

Question 34 - LOOCV : special case

Question 34 1 pts

LOOCV is a special case of k -fold CV in which $k=n$.

☒ True
☐ False

Question 35 1 pts

With regard to discriminant functions $\delta_{k(x)}$ and conditional densities $f_{k(x)}$ for classification, all of the following statements are true, EXCEPT:

☐ There exists a discriminant function for each class.
☐ An observation x is assigned to class k if $\delta_{k(x)} \leq \delta_{j(x)}$ for all j .
☐ A boundary between classes j and k exists for all x wherever $f_j(x) = f_k(x)$.
☒ In linear discriminant analysis, the input space is divided into regions of constant classification, with piecewise hyperplanar decision boundaries.

(a) Current Quiz

Question 15 1 pts

LOOCV is a special case of k -fold CV in which $k=n$.

☒ True
☐ False

(b) Prior Quiz

Solution: TRUE

NOTES: N/A

Question 35 - Discriminant Functions / Conditional densities ($\delta(x)$, etc...)

Question 34 1 pts

LOOCV is a special case of k-fold CV in which $k=n$.

☒ True

☐ False

Question 35 1 pts

With regard to discriminant functions $\delta_k(x)$ and conditional densities $f_k(x)$ for classification, all of the following statements are true, EXCEPT:

☐ There exists a discriminant function for each class.

☐ An observation x is assigned to class k if $\delta_k(x) \leq \delta_j(x)$ for all j .

☐ A boundary between classes j and k exists for all x wherever $f_j(x) = f_k(x)$.

☒ In linear discriminant analysis, the input space is divided into regions of constant classification, with piecewise hyperplanar decision boundaries.

(a) Current Quiz

Question 30 1 pts

With regard to discriminant functions $\delta_k(x)$ and conditional densities $f_k(x)$ for classification, all of the following statements are true, EXCEPT:

☐ An observation x is assigned to class k if $\delta_k(x) \leq \delta_j(x)$ for all j .

☐ There exists a discriminant function for each class.

☐ A boundary between classes j and k exists for all x wherever $f_j(x) = f_k(x)$.

☐ In linear discriminant analysis, the input space is divided into regions of constant classification, with piecewise hyperplanar decision boundaries.

(b) Prior Quiz

Solution: (2) = 'j' is False

NOTES:

Question 36 - Mallows C_p and AIC (model selection criteria)

Question 36 1 pts

When using subset selection techniques, Mallows C_p will always choose the same number of predictors as the Akaike Information Criterion.

☐ True

☒ False

Question 37 1 pts

Linear regression is a non-parametric learning method:

☐ True

☒ False

(a) Current Quiz

?

NO OLD REFERENCE

(b) Prior Quiz

Solution: FALSE (d vs. $d(\alpha)$), slides of week 4

NOTES:

Question 37 - Linear regression is non-parametric

Question 36 1 pts

When using subset selection techniques, Mallows C_p will always choose the same number of predictors as the Akaike Information Criterion.

☐ True

☒ False

Question 37 1 pts

Linear regression is a non-parametric learning method:

☐ True

☒ False

(a) Current Quiz

Question 7 1 pts

Linear regression is a non-parametric learning method:

☐ True

☒ False

(b) Prior Quiz

Solution: FALSE

NOTES: N/A