Quiz 3

Due No due date **Points** 20 **Questions** 10 **Time Limit** None

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	31 minutes	16 out of 20

Score for this quiz: **16** out of 20
Submitted Oct 22 at 4:56pm
This attempt took 31 minutes.

Question 1

2 / 2 pts

(1) True/False: Models selected by automated variable selection techniques do not need to be validated since they are 'optimal' models.

- True
- False

Correct!

False – All models need to be validated. Models selected by automated variable selection are only 'optimal' with respect to the metric used by the algorithm. It does not mean that the models are correctly specified.

Question 2

2 / 2 pts

(2) Compute the Akaike Information Criterion (AIC) value for the linear regression model

$Y = b0 + b1*X1 + b2*X2 + b3*X3.$

The regression model was fitted on a sample of 250 observations and yielded a likelihood value of 0.18.

(a) 9.49

Correct!

(b) 11.43

(c) 25.52

(d) 15.55

Question 3

2 / 2 pts

(3) Compute the Bayesian Information Criterion (BIC) value for the linear regression model

Y = b0 + b1*X1 + b2*X2 + b3*X3.

The regression model was fitted on a sample of 250 observations and yielded a likelihood value of 0.18.

(a) 9.49

(b) 11.43

Correct!

(c) 25.52

(d) 15.55

Question 4

2 / 2 pts

(4) True/False: Consider a categorical predictor variable that has three levels denoted by 1, 2, and 3. We can include this categorical predictor variable in a regression model using this specification, where X1 is a dummy variable for level 1, X2 is a dummy variable for level 2, and X3 is a dummy variable for level 3.

Y = b0 + b1*X1 + b2*X2 + b3*X3

Correct!

True

False

False – We cannot include all three dummy variables. Any categorical variable with k levels can be included in a regression model with at most (k-1) dummy variables when an intercept is included in the model. One possible specification would be Y = b0 + b2*X2 + b3*X3, where level 1 is taken to be the base category.

Question 5

0 / 2 pts

(5) True/False: The model Y = b0 + exp(b1*X1) + e can be transformed to a linear model.

'ou Answered

True

False – Taking the natural logarithm of both sides does not yield a log-linear model.

orrect Answer

False

Question 6

2 / 2 pts

(6) True/False: A variable transformation can be used as a remedial measure for heteroscedasticity.

Correct!

True

True – Variable transformations are used for two primary reasons: (1) to stabilize the variance, i.e. a remedial measure for heteroscedasticity, and (2) to linearize the model.

False

Question 7

2 / 2 pts

(7) When comparing models of different sizes (i.e. a different number of predictor variables), we can use which metrics?

a. R-Squared and Adjusted R-Squared

b. R-Squared and Mallow's Cp

c. AIC and R-Squared

Correct!

d. AIC and BIC

Question 8

2 / 2 pts

(8) True/False: When using Mallow's Cp for model selection, we should choose the model with the largest Cp value.

True

Correct!

False

False – When using Mallow's Cp, you should select the model with the smallest Cp value that is 'close' to the diagonal line Cp = p. In practice this can be difficult to interpret so one should consider using metrics such as Adjusted R-Squared, AIC, or BIC instead.

Question 9

0 / 2 pts

(9) True/False: Consider the case where the response variable Y is constrained to the interval [0,1]. In this case one can fit a linear regression model to Y without any transformation to Y.

'ou Answered

- (•) True

False – Linear regression assumes a continuous response variable over the range of the real line. In the situation where Y is constrained to the interval [0,1], one should transform Y using the arcsin transformation.

orrect Answer

- ( ) False

Question 10

2 / 2 pts

(10) True/False: Consider the case where the response variable Y takes only two values: 0 and 1. A linear regression model can be fit to this data.

- ( ) True

Correct!

- (•) False

False – In the case of a binary response variable there is no transformation that will produce a valid linear regression model. In this case one should fit a logistic regression model.

Quiz Score: **16** out of 20