

Quiz 5 - Week #5

Due No due date

Points 20

Questions 10

Time Limit None

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	20 minutes	18 out of 20

Score for this quiz: **18** out of 20

Submitted Feb 10 at 12:12pm

This attempt took 20 minutes.

Question 1

0 / 2 pts

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

You Answered

☒ True

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.

Correct Answer

☐ False

Question 2

2 / 2 pts

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

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3.$$

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

Correct!

- ☐ (a) 9.49
- ☒ (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

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3.$$

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

Correct!

- ☐ (a) 9.49
- ☐ (b) 11.43
- ☒ (c) 25.52
- ☐ (d) 15.55

Question 4

2 / 2 pts

(4) Consider the following three models:

(a) Model 1: $Y \sim X_1 + X_2 + X_3 + X_4$ with AIC = -1234.3

(b) Model 2: $Y \sim X_1 + X_3 + X_5 + X_7$ with AIC = -1279.4

(c) Model 3: $Y \sim X_2 + X_4 + X_6 + X_7$ with AIC = -1189.2

Using the AIC values which model should we select as the best model?

☐ (a) Model 1

Correct!

☒ (b) Model 2

Using the AIC values select the value with the smallest AIC value. Here we select Model 2.

☐ (c) Model 3

Question 5

2 / 2 pts

(5) Consider the following three models:

(a) Model 1: $Y \sim X_1 + X_2 + X_3 + X_4$ with BIC = -934.3

(b) Model 2: $Y \sim X_1 + X_3 + X_5 + X_7$ with BIC = -779.4

(c) Model 3: $Y \sim X_2 + X_4 + X_6 + X_7$ with BIC = -1589.2

Using the BIC values which model should we select as the best model?

☐ (a) Model 1

☐ (b) Model 2

Correct!

☒ (c) Model 3

Using the BIC values select the value with the smallest BIC value. Here we select Model 3.

Question 6

2 / 2 pts

(6) True/False: Consider a modeling exercise that has 20 possible predictor variables. You apply both forward and backward variable selection based on AIC. Both variable selection methods will always return exactly the same model, i.e. both variable selection algorithms will select the same set of predictor variables.

☐ True

Correct!

☒ False

False – Applying different variable selection methods to the same set of possible predictor variables may or may not return the same set of optimal predictors. The predictor variables that are selected depend both on metric used by the algorithm and the relationships (correlation) between the possible predictor variables.

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.

Correct!☐ True☒ False

False – When using Mallows' C_p , you should select the model with the smallest C_p value that is 'close' to the diagonal line $C_p = p$. 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**2 / 2 pts**

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

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3.$$

The regression model was fitted on a sample of 750 observations and $SSE = 1256$. Enter your answer with three decimal places of precision, e.g. 0.001.

Correct!

Answer: $750 \cdot \log(1256/750) + 2 \cdot 4 = 394.7106$

Correct Answers

394.7106 (with margin: 0.01)

0.0 (with margin: 0.0)

Question 10**2 / 2 pts**

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

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3.$$

The regression model was fitted on a sample of 750 observations and SSE = 1256. Enter your answer with three decimal places of precision, e.g. 0.001.

Correct!

413.191

Answer: $750 \cdot \log(1256/750) + 4 \cdot \log(750) = 413.1909$

Correct Answers

413.1909 (with margin: 0.01)

0.0 (with margin: 0.0)

Quiz Score: **18** out of 20