Print

Quiz: Model Building and Effect Selection

Your Score: 100%

Congratulations! Your score of 100% indicates that you've mastered the topics in this lesson. If you'd like, you can review the feedback.

When you're finished, exit the lesson.

1. In forward selection, after a variable is added to the model, it can be removed if it becomes non-significant later.

   a. true
   b. false

Your answer: b
Correct answer: b

In forward selection, after a variable is added to the model, it stays in, even if it becomes non-significant later.

---

2. Given the information in this summary of variable selection, which stepwise selection method was specified in the PROC REG step?

Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	RunTime		1	0.7434	0.7434	3.3432	84.00	<.0001
1	Age		2	0.0213	0.7647	2.8192	2.54	<.1222

   a. FORWARD
   b. BACKWARD
   c. STEPWISE
   d. Not enough information is given.

Your answer: c
Correct answer: c

The summary table contains both **Variable Entered** and **Variable Removed** columns. Of the three types of stepwise selection (forward, backward, and stepwise), only stepwise selection can both enter and remove variables.

Therefore, STEPWISE must have been specified in the PROC REG step.

---

3. When you interpret *p*-values from models that are chosen using any automated variable selection technique, which of the following should you be cautious about?

   a. incorrect calculation of degrees of freedom
   b. biases in parameter estimates, predictions, and standard errors
   c. *p*-values that tend to err on the side of overestimating significance
   d. all of the above

   Your answer: d
   Correct answer: d

   You should be cautious when you're interpreting and reporting statistical quantities that are produced by these methods. Using automated model selection results in biases in parameter estimates, predictions, and standard errors, incorrect calculation of degrees of freedom, and p-values that tend to err on the side of overestimating significance.

---

4. When you add predictor variables to a model, which of the following values tend to increase or stay the same (and can never decrease)?

   a. R-square
   b. adjusted R-square
   c. *F* statistic
   d. none of the above
   e. all of the above

   Your answer: a
   Correct answer: a

   The R-square always increases or stays the same when you include more terms in the model. Therefore, choosing the "best" model is not as simple as making the R-square as large as possible.

5. Which of the following statements is true about information criteria such as AIC, AICC, BIC, and SBC?

   a. Formulas for all information criteria begin with the same calculation.
   b. The penalty term to assess the complexity of the model enables information criteria to be a useful means of comparing models with a different number of parameters.
   c. The best model is the one with the smallest information criteria value.
   d. All of the above.

Your answer: d
Correct answer: d

All of the statements about information criteria are true.

Close