Flag for Review
Record Notes
Email the Professor

Question Difficulty: Core

Monitor Difficulty Level

You are given the following table for model selection:

Model	Scaled Deviance	Number of parameters (k+1)	AIC
Intercept+ Age	A	4	225
Intercept+ Risk Class	261	8	
Intercept+ Age+ Vehicle Value	264	X	
Intercept+ Age+ Risk Class + Vehicle Value	B	Y	307

Calculate Y.

A 16

B 23

C 24

D 17

Helpful Strategies To Get You Started

Comprehensive Solutions with Alternative Solutions When Available

Help Me Start

Recall that the AIC is $\text{AIC} = -2\ell + 2p$, where $p = k + 1$ is the number of parameters. What can you infer from the difference between two AICs?

The scaled deviance is $D^* = 2(\ell_{sat} - \ell)$. Can you do the same with scaled deviances?

Solution

The AIC is $\text{AIC} = -2\ell + 2p$, where $p = k + 1$ is the number of parameters.

The difference in AICs is

$$\text{AIC}_2 - \text{AIC}_1 = 2(\ell_1 - \ell_2) + 2(p_2 - p_1) = (D_2^* - D_1^*) + 2(p_2 - p_1).$$

The scaled deviance is $D^* = 2(\ell_{sat} - \ell)$.

Since the saturated model is the same for all four models, the difference in the scaled deviance is,

$$D_2^* - D_1^* = 2(\ell_1 - \ell_2).$$

We can use this relationship to find X, then Y.

Comparing the second and the third model,

$$\begin{aligned} \text{AIC}_3 - \text{AIC}_2 &= (D_3^* - D_2^*) + 2(p_3 - p_2) \\ 233 - 214 &= (264 - 261) + 2(X - 8). \end{aligned}$$

This tells us X is 16.

From the second model alone, we see that Risk Class has 7 parameters.

The fourth model, must have 7 parameter more than the third model.

Therefore, Y is 23.

Leave Feedback By Rating Problems

Rate this problem | Excellent | Needs Improvement | Inadequate

Flag for Review

Record Notes

Email the Professor

actuarial e-learning portal

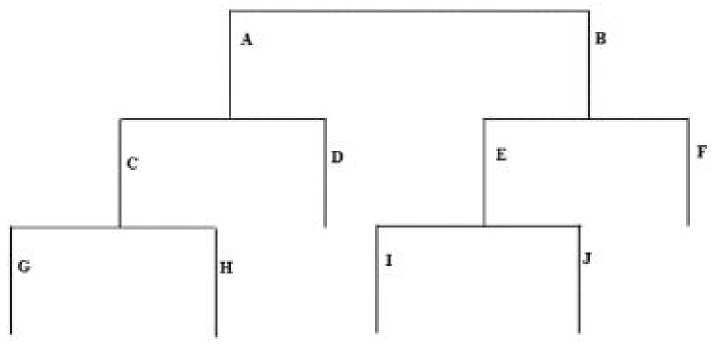
Home Support Log Off

GOAL | Problem # | Go! | Problem 1 of 11 | Prev | Next | Leave

Question | Difficulty: Mastery

Monitor Difficulty Level

The following diagram illustrates an unpruned decision tree:



The table below gives the residual sum of squares (RSS) at each node:

Pruned decision tree		Unpruned decision tree	
Node	RSS	Node	RSS
A	75	D	32
B	50	F	28
C	55	G and H	19
E	40	I and J	22

You are to use the cost-complexity pruning method to prune the tree, and the value of α of the cost-complexity function is 3. Which of the following is preferred?

A Prune the tree at node A only.

B Prune the tree at node B only.

C Prune the tree at node C only.

D Prune the tree at node E only.

E Do not prune the tree.

Helpful Strategies To Get You Started

Comprehensive Solutions with Alternative Solutions When Available

Help Me Start

The cost-complexity function is $f = \sum_{m=1}^{|T|} \sum_{i: y_i \in R_m} (y_i - \hat{y}_{R_m})^2 + \alpha |T|$ where y_i is an observation of node R_m and $|T|$ is the number of nodes.

Solution

<u>If the tree is pruned at node A only:</u>

The number of nodes is $|T| = 4$.

The value of the function is $f = 75 + 22 + 28 + 3\,(4) = 137$.

Graphs and Other Solution Techniques Demonstrated When Available

<u>If the tree is pruned at node B only:</u>

The number of nodes is $|T| = 4$.

The value of the function is $f = 50 + 19 + 32 + 3\,(4) = 113$.

<u>If the tree is pruned at node C only:</u>

The number of nodes is $|T| = 5$.

The value of the function is $f = 55 + 32 + 22 + 28 + 3\,(5) = 152$

<u>If the tree is pruned at node E only:</u>

The number of nodes is $|T| = 5$.

The value of the function is $f = 40 + 19 + 32 + 28 + 3\,(5) = 134$.

<u>If the tree is not pruned:</u>

The number of nodes is $|T| = 6$.

The value of the function is $f = 19 + 22 + 28 + 32 + 3\,(6) = 119$.

We can see that the value of the function is the smallest when the tree is pruned at node B only.

Leave Feedback By Rating Problems

Rate this problem | Excellent | Needs Improvement | Inadequate