Lab 7

Kristen Fitzgerald

2024-07-07

Question 1

Part (a)

Build a regression model predicting the `FE` variable using all the remaining variables. Some of these predictor variables are coded as numeric, but should be treated as categorical. The only numeric variables in your dataset should be `EngDispl`. All remaining variables are categorical. Use **forward** selection with a **0.1 p-value selection criterion**.

a. What is the final model?

Solution:

`FE ~ EngDispl + CarlineClassDesc + NumCyl + DriveDesc + Transmission + IntakeValvePerCyl + VarValveLift + AirAspirationMethod + NumGears + TransLockup + TransCreeperGear + ExhaustValvesPerCyl`

```
# starting model, most start with zero variables - intercept only model
empty.model <- lm(FE ~ 1, data = cars2010)


# Upper model, starts with all variables
full.model <- lm(FE ~ ., data = cars2010)


for.model <- step(empty.model,
                  scope = list(lower = empty.model,
                               upper = full.model),
```

```
                  direction = "forward",
                  k = qchisq(0.1, 1, lower.tail = FALSE))
```

b. What was the first variable added?

Solution: The first variable added was `EngDispl`.

c. What was the last variable added?

Solution: The last added variable was `ExhaustValvesPerCyl`.

Part(b)

Instead of the previous approach, now use **stepwise selection** with a **BIC criterion**.

a. How many variables (not parameter estimates) are left?

Solution: There are 8 variables left.

`FE ~ EngDispl + DriveDesc + CarlineClassDesc + NumCyl + ExhaustValvesPerCyl + NumGears + AirAspirationMethod + TransCreeperGear`

```
empty.model <- lm(FE ~ 1, data = cars2010)

full.model <- lm(FE ~ ., data = cars2010)

for.model <- step(empty.model,
                  scope = list(lower = empty.model,
                               upper = full.model),
                  direction = "both", k = log(nrow(cars2010)))
```

b. Are these the same variables as with the forward selection?

Solution: The forward selection included 12 variables and stepwise included 8. The similar variables include: `EngDispl`, `DriveDesc`, `CarlineClassDesc`, `NumCyl`, `ExhaustValvesPerCyl`, `NumGears`, `AirAspirationMethod`, and `TransCreeperGear`. All variables in the stepwise were in the forward model, however, `Transmission`, `IntakeValvePerCyl`, `VarValveLift`, and `TransLockup`.