Alyssa Semerdjian

Stat 510

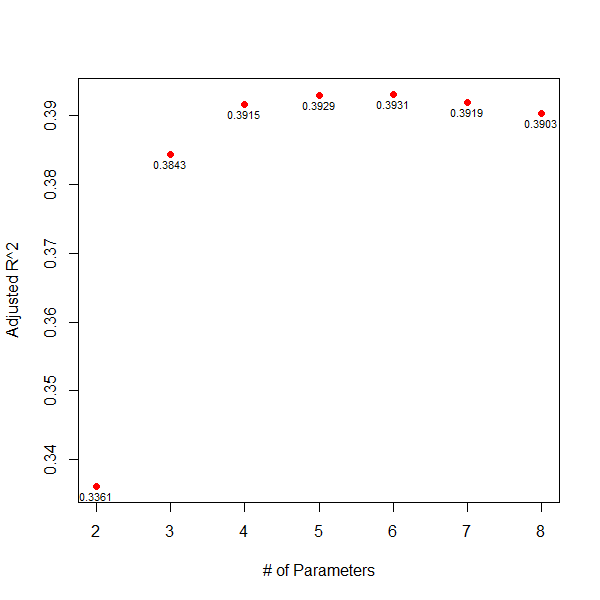
29 Jan 2018

**Homework 2**

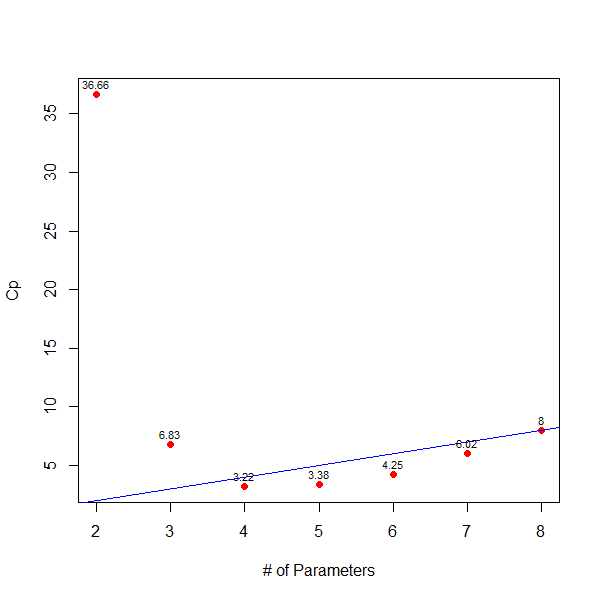
a) The equation for the ‘best’ estimated model is:

glucose = 65.5 + 0.24(diastolic) +0.14(insulin) + 0.60(age).

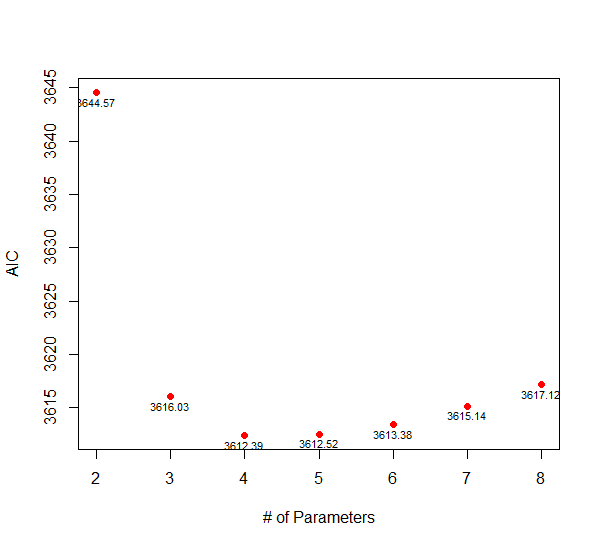
b) The models with 4, 5, 6, 7 and 8 variables all have very similar adjusted R2 values, however model 6 is the ‘best’ because it has the highest adjusted R2.

b

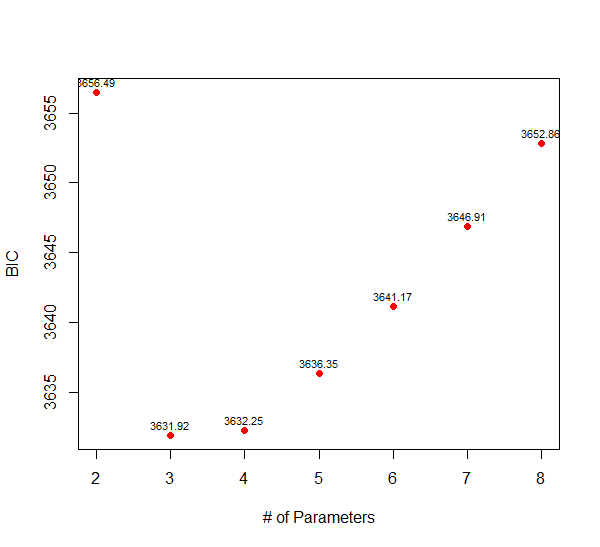
c) graph for Mallows’ Cp model 7 is closest to the target (7-6.02)/7?– not counting model 8? – double check this



d) n/(p+1) is not less than 40 so you’d want to use AIC for this data. However, since the question asks which model is best using AICc, the model with the lowest AICc has \_\_ \_\_ as variables. Below is a graph with AIC scores for each model.



e) The model with the lowest BIC (and therefore the ‘best’ model) has the variables \_\_ \_\_.



f) The equation for AIC is AIC = and the equation for AICc is

AICc = AIC + . Loglikelihood for model 5 as calculated in R is -1799.69, p is 6 and n is 392. So AIC is = = . So AICc =+ =+ = – 239 = **3376.375.**

g) Normal qq is not idea, and the residuals are not perfectly scattered around 0. Could be better