STAT 3400 - Homework #X

Alex Ojemann

Due February 8, 2023

Problem 8.6.2

Adrian is right. Encoding a binary categorical variable as a 0/1 numeric variable allow it to be used in linar regression.

Problem 8.6.4

a. False. Two predictors being collinear means that they have a significant correlation between them, so removing one will affect the coefficient of the other in whichever direction those two predictor variables were correlated.

b. True

c. False. Predictive models output the average outcome value given its predictors, not the exact outcome value.

Problem 8.6.8

These figures do not support the comment in the article because the while the predicted ROI is much higher for horror movies than other genres, the actual ROI is fairly similarly distributed for all genres.

Problem 8.6.10

a. y = -1461 + 18.2b_1 + 67.2b_2 + 16b_3 + 389.9b_4 - 251.5b_5 + 1014.6b_6

b. b_1: For each 1mm increase in bill length, body weight increases by 18g on average. b_2: For each 1mm increase in bill depth, body weight increases by 67.2g on average. b_3: For each 1mm increase in bill length, body weight increases by 16g on average. b_4: The average body weight is 389g higher for males than for females. b_5: The average body weight is 251.5g lower for Chinstraps than for Adelies. b_6: The average body weight is 1014.6g higher for Gentoos than for Adelies.

c. The predicted weight is 3,793.16g. Thus, the model overpredicted its actual weight of 3,750g.

d. 87.5% of the variation in body weight can be accounted for by bill length, bill depth, flipper length, sex, and species.

Problem 8.6.12

None of the variables should be dropped because the adjusted R^2 is lower if any are dropped than it is with all of them included.

Problem 8.6.14

Flipper length should be added to the model first because it results in the highest adjusted R^2.