Quiz 5
MSAN 601

September 30, 2016

Question 1 **(2 pts)**

How is centering of the predictor variable useful in case of a second order polynomial model with one predictor variable?

Answer

Centering helps reduce multicollinearity.

Question 2 **(2 pts)**

While using the Hierarchical Approach to fit polynomial models, what rule must be followed about the polynomial terms of a predictor?

Answer

If a higher order polynomial term of a predictor is used in the model, then all the lower order terms of that predictor must also be included in the model.

Question 3 **(6 pts)**

Explain what an interaction term is. Give a simple numerical example and explain it.

Answer

An interaction term is used in a regression model when there exists a relation between the predictors such that the influence of two predictors on the response is not additive.

A model with two predictors and an interaction term takes a form of

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

The change in the mean response with a unit increase in X_1 when X_2 is held constant is $\beta_1 + \beta_3 X_2$. The increase in X_1 depends on the level of X_2.

Question 4 **(6 pts)**

Discuss two types of interaction effects. Write an example equation for each.

Answer

1. Reinforcement: $E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$
2. Interference: $E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 - \beta_3 X_1 X_2$

Question 5 **(4 pts)**

Consider the predictors `age` (numeric) and `gender` (M and F), and the response `height` (numeric). While constructing a model to predict height using age and gender, how many regressors will a model that is linear in the predictors have? How will the values F and M of gender be represented in the model? Write out the *theoretical* regression equation.

Answer

The model will have 2 predictors, one for age and one for gender.

The gender predictor can be coded as 0 for F and 1 for M.

The model equation is: $Y = \beta_0 + \beta_1\ age + \beta_2\ gender$, where gender = 0 when F and 1 when M

Question 6 **(6 pts)**

What is the purpose of the Ramsay RESET test? What are the advantages and disadvantages of using said test? Write out the null and alternative hypotheses.

Answer

Ramsey RESET is a top-down approach to determine if the predictors raised to their second or higher powers and the interaction terms are significant in the model. The test is very generalized and can only be used as a guidance. Further steps will be needed to determine the predictors to be included in the model. However, it helps save time and efforts while modeling.

H_0: $\alpha_1 = ... = \alpha_k = 0$, H_a: *not all* $\alpha_k = 0$

where $\alpha_1, ..., \alpha_k$

Question 7 **(6 pts)**

Consider the following fitted regression equation:

$$\widehat{height} = 3 + 0.33 age + 0.266 age^2$$

Interpret a positive, one-unit change in the value of the predictor `age`.

Answer

Taking the derivative, $\frac{\partial\ \widehat{height}}{\partial\ age} = 0.33 + (2)0.266\ age$

$\therefore \frac{\partial\ \widehat{height}}{\partial\ age} = 0.33 + 0.532\ age$

Thus, in going from the age of 0 to 1 units, height increases by 0.33, in going from the age of 1 to 2 units, height increases by $0.33 + 0.532\ (1) = 0.862$, in going from the age of 2 to 3 units, height increases by $0.33 + 0.532\ (2) = 1.394$ and so on.