Quiz 4
MSAN 601

September 21, 2016

Question 1 (3 pts)

What does 'the effect of predictor variables on the mean response is additive' mean in the context of first-order regression model?

Answer

In a first-order regression model, the effect of a predictor variable on the mean response does not depend on the level of other predictors in the model. In other words, the predictors do not interact and thus, are said to have an additive effect.

Question 2 (4 pts)

The GLRM in the matrix form can be written as $Y = X\beta + \epsilon$. Write the dimensions of Y, X, β and ϵ matrices. (Assume n data points and $p-1$ predictors.)

Answer

Y: n × 1
X: n × p
β: p × 1
ϵ: n × 1

Question 3 (3 pts)

How does adjusted R^2 help in determining if a predictor should be included in the model?

$$R_a{}^2 = 1 - \frac{n-1}{n-p}\frac{SSE}{SSTO}$$

When a predictor is added to an existing model, SSE decreases and p increases. Thus, $R_a{}^2$ can either increase or decrease. A good predictor will help decrease the SSE enough to account for increased value of p. As a result, the second term of the above equation will decrease and thus, $R_a{}^2$ will increase. These predictors should be included in the model.

Question 4 (6 pts)

For an MLR model $Y = 0.5 + 0.85X_1 - 0.32X_2 + 0.33X_3$, interpret how a unit increase in each of the parameters X_1, X_2, X_3 affect the response Y.

Answer

Holding all other predictors at a constant level,

- For a unit increase in X_1, we expect the response Y to increase by 0.85.

- For a unit increase in X_2, we expect the response Y to decrease by 0.32.

- For a unit increase in X_3, we expect the response Y to increase by 0.33.

Question 5 (3 pts)

Explain the statement 'You cannot unexplain error'.

Answer

The statement is used in the context of linear regression. As more variables are added to the model, SSE decreases and R^2 increases. Additional variables always help to explain more error and not less. Thus, we cannot unexplain error.

Question 6 (6 pts)

For an MLR, $SSTO = 500$, $SSE(X_1, X_2) = 110$, $SSR(X_1|X_2) = 5$
Find: $SSR(X_2)$, $SSR(X_1, X_2)$, $SSE(X_2)$

Answer

multiple answers here due to typo in question

Question 7 (2 pts)

What is coefficient of partial determination?

Answer

A coefficient of partial determination is used to measure the marginal contribution of a predictor when other predictors are already included in the model.

Question 8 (2 pts)

What happens to the MLR model when perfectly correlated variables are included as predictors?

Answer

If perfectly correlated variables are included as predictors, the design matrix X isn't full rank and thus cannot be inverted.

Question 9 (4 pts)

What is multicollinearity? How does it affect an MLR model? How can it be detected?

Answer

Multicollinearity exists when two or more predictors in an MLR are highly correlated.
It can lead to an increase in the standard error of regression coefficients affecting their statistical significance. It also affects their interpretation.

Observing the relation between the predictors using a scatter plot is the simplest method to detect multicollinearity. One can also observe the marginal contribution of additional predictor in reducing SSE. If the predictors are correlated, the contribution will be very small.