

Table 1.11 Various Classifications of Regression Analysis

Type of Regression	Conditions
Univariate	Only one quantitative response variable
Multivariate	Two or more quantitative response variables
Simple	Only one predictor variable
Multiple	Two or more predictor variables
Linear	All parameters enter the equation linearly, possibly after transformation of the data
Nonlinear	The relationship between the response and some of the predictors is nonlinear or some of the parameters appear nonlinearly, but no transformation is possible to make the parameters appear linearly
Analysis of Variance	All predictors are qualitative variables
Analysis of Covariance	Some predictors are quantitative variables and others are qualitative variables
Logistic	The response variable is qualitative

$\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$. The estimated regression equation then becomes

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p. \quad (1.5)$$

A *hat* on top of a parameter denotes an estimate of the parameter. The value \hat{Y} (pronounced as *Y-hat*) is called the *fitted* value. Using (1.5), we can compute n fitted values, one for each of the n observations in our data. For example, the i th fitted value \hat{y}_i is

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_p x_{ip}, \quad i = 1, 2, \dots, n, \quad (1.6)$$

where x_{i1}, \dots, x_{ip} are the values of the p predictor variables for the i th observation.

Note that (1.5) can be used to predict the response variable for any values of the predictor variables not observed in our data. In this case, the obtained \hat{Y} is called the *predicted* value. The difference between fitted and predicted values is that the fitted value refers to the case where the values used for the predictor variables correspond to one of the n observations in our data, but the predicted values are obtained for any set of values of the predictor variables. It is generally not recommended to predict the response variable for a set of values of the predictor variables far outside the range of our data. In cases where the values of the predictor variables represent