

## **Multiple Linear Regression**

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

**The expression above represents a multiple linear regression model relating a dependent variable  $y$  to  $k$  independent variables referred to as predictors. The error term  $\varepsilon$  accounts for variability in  $y$  not explained by the predictors.**

**Basic assumptions-**

- 1) The regression model is linear in the parameters.**
- 2) The predictors are measured without error.**
- 3) The error term is an independent normal random variable, with zero mean and constant variance.**

**Some facts-**

- 1) Any regression model linear in the parameters is a linear regression model regardless of the shape of the surface it generates. A polynomial is an example.**
- 2) Various functional forms are possible. Interactions can be included, i.e.,  $x_1 x_2$ , or  $\exp(-x_1)$  etc. Variables can be ratio, interval, ordinal or nominal.**
- 3) The goal of developing a multiple linear regression model is to include the right variables, exclude unnecessary variables and have the variables expressed in proper functional form.**

## **Model adequacy checking-**

**The validity of a fitted regression model needs to be checked. This involves looking for indications that the fitted model does not meet regression assumptions. Much of this entails examining residuals. There are a variety of methods which can be employed.**

**This video illustrates some of the steps involved in model development, and exploratory methods which can be used to evaluate the adequacy of a fitted regression model. In particular comments will be made regarding:**

- 1) Interpretation of overall F-tests and the multiple correlation coefficient,**
- 2) Use of p-values for determining which variables to retain in a regression model,**
- 3) Interpretation of regression coefficients,**
- 4) Evaluation of residuals, and**
- 5) Consideration of proper functional form.**

**The methods demonstrated in this video are indicative of what will be studied in future courses. This video is not a comprehensive treatment of model adequacy checking. It does point to the considerations that need to be made as part of the second abalone data analysis.**