LCF 280 – Métodos Quantitativos para a Gestão Ambiental
Aula 6

Regressão Linear Múltipla - Modelo Linear Generalizado

Considere agora que *p* representa o número de parâmetros a serem estimados num modelo de regressão linear com múltiplas variáveis. Podemos definir esse modelo linear generalizado, com erros normalmente distribuídos, da seguinte forma:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \ldots \beta_{p-1} X_{i,\,p-1} + \varepsilon_i \qquad \text{onde:}$$

$\beta_0, \beta_1, \beta_2, \ldots \beta_{p-1}$ são parâmetros; $X_{i,1}, \ldots X_{i,\,p-1}$ são constantes conhecidas;
ε_i representa um erro independente c/ distribuição $N(0, \sigma^2)$; e $i = 1, 2, \ldots, n$

Se considerarmos que $E\{\varepsilon_i\} = 0$, podemos assumir que a função de resposta é:

$$E\{Y_i\} = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \ldots \beta_{p-1} X_{i,\,p-1}$$

Esse modelo, aparentemente linear, pode também ser usado em situações não lineares. Por exemplo, em um modelo de regressão polinomial, quando, se adotam variáveis quadráticas de uma única variável independente ($X_{i1} = X_i$ e $X_{i2} = X_i^2$):

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$$

Outro exemplo em que a “linearidade” poder ser contornada é quando transformamos as variáveis. Por exemplo, a variável dependente pode ser transformada em $\log Y_i$ ou $1/Y_i$. Nesses casos, se $Y'_i = \log Y_i$ ou $Y'_i = 1/Y_i$, e o modelo depende de três variáveis independentes, poderíamos ter:

$$Y'_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$$

É possível usar também o mesmo modelo geral linearizado em situações em que se pretende estudar o efeito interativo entre variáveis:. Por exemplo, podemos considerar um modelo de regressão com duas variáveis do tipo:

$$Y'_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + \varepsilon_i$$

Nesse caso, o que se espera é que a mudança média na variável Y, como resultado do aumento de uma unidade na variável X_1 sem alterar X_2 seja igual a:

$$\beta_1 + \beta_3 X_2$$

de onde se conclui que o efeito de variar X_1 ou X_2 depende do nível da outra variável.