

# Regression and Clinical prediction models

Seção 16  
Especificação e Estimação – Parte 4

Marcel de Souza Borges Quintana  
marcel.quintana@ini.fiocruz.br  
2019







# Aditividade

Lembrando:  $f(E(Y)) = \alpha + X_1\beta_1 + X_2\beta_2 + \dots + X_q\beta_q$

Parte direita:  $LP = \alpha + X_1\beta_1 + X_2\beta_2 + \dots + X_q\beta_q$

- $\beta_1, \beta_2, \dots, \beta_q$  - Efeitos principais
- Implicitamente, pressupomos que os efeitos ( $\beta_1, \beta_2, \dots$ ) dependem apenas de suas respectivas variáveis ( $X_1, X_2, \dots$ ).
  - Se esta hipótese não for verdadeira -> Interações
- Exemplo:  $X_1$ : Tratamento ( $T_1$  e  $T_2$ ) e  $X_2$ : Hospital ( $H_1$  e  $H_2$ ).



# Aditividade

- Tipos de interação
  1. Interação Quantitativa
    - Quando uma variável potencializa o efeito da outra.
  2. Interação Qualitativa
    - Quando uma variável inverte o efeito da outra.
- Forma de representar a interação:

$$LP = \alpha + X_1\beta_1 + X_2\beta_2 + (X_1 \times X_2)\beta_{12}$$

- Se  $X_1$  for *dummy* (0 ou 1), é preferível representar a interação por:

$$LP = \alpha + X_1\beta_1 + (1 - X_1)X_2\beta_2 + (X_1 \times X_2)\beta_{12}$$









# Não-linearidade

- Mesmos métodos comentados na análises univariadas:
  1. Splines
    - Restricted Cubic splines
      - Graus definidos *a priori* (depende do tamanho da amostra e da *expertise* no tema)
    - Splines em GAM
      - Graus de liberdade estimados por validação cruzada generalizada (GCV)
  2. Polinômios fracionais
    - Busca iterativamente uma transformação ótima com base no nível de significância.
      - É possível testar:
        1. Se existe efeito da variável;
        2. Se existe efeito não-linear;
        3. Se existe necessidade de transformações mais complexas (F2 vs F1)

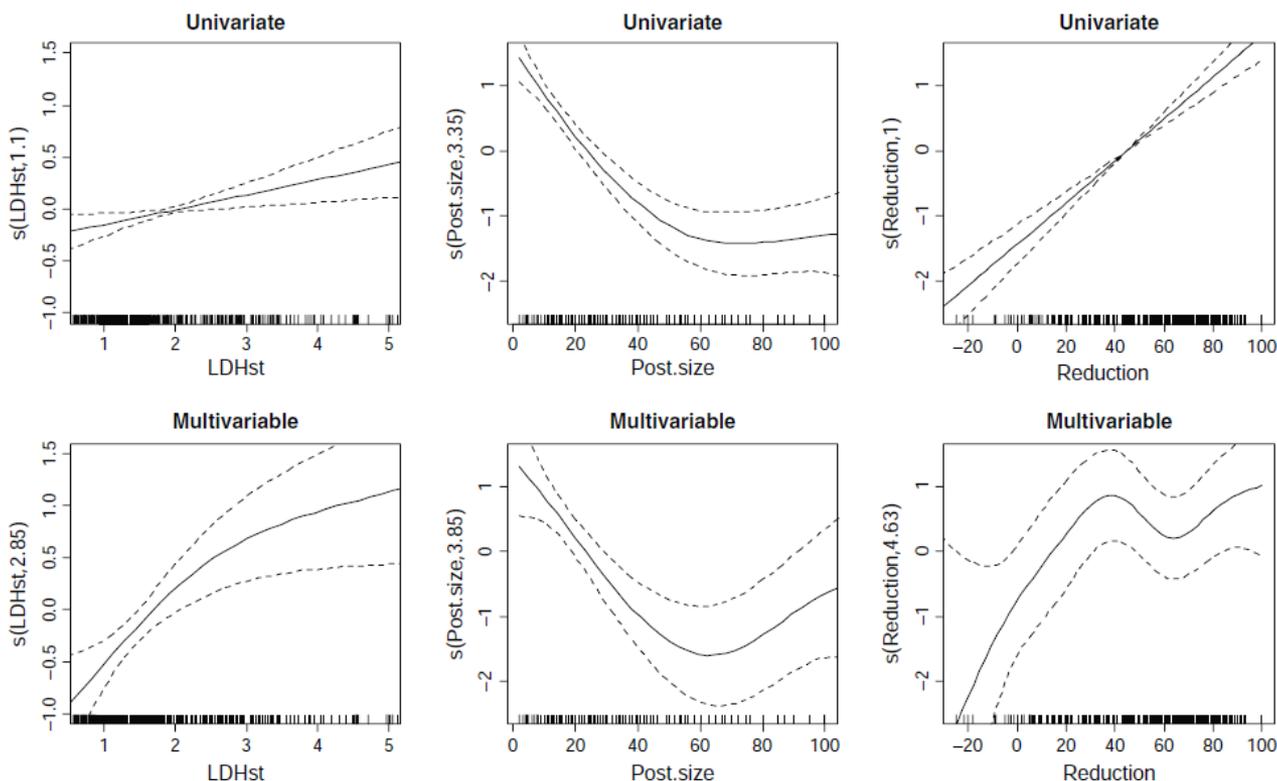


## Não-linearidade

### 2. Ver como se comportam as curvas no modelo multivariado

- Exemplo: GAM com GCV

Fig. 12.5 Generalized additive models with optimal smoothing spline transformations according to a generalized cross-validation procedure in the testicular cancer example ( $n=544$ ). Top row: optimal transformation in univariate logistic regression analysis; bottom row: multivariable logistic regression analysis with six predictors. The degrees of freedom of the optimal smoothing spline transformation are shown in each y-axis label. The distribution of values if shown at the bottom of the graphs





# fim

Seção 16  
Especificação e Estimação – Parte 4

Marcel de Souza Borges Quintana  
marcel.quintana@ini.fiocruz.br  
2019

