

# Introdução a Análise de resíduos e diagnósticos em GLM

# Introdução

Análise de Diagnósticos

Análise de  
Resíduos

- Analisar algumas suposições.
- Verificar se o ajuste dos dados está correto (adequado)

Métodos de  
diagnóstico

avaliar pontos  
que possam ser  
influentes

# Diagnóstico do Ajuste

A análise de Diagnóstico é uma etapa fundamental no ajuste dos modelos de regressão.

Como garantimos que o modelo está ajustado aos dados?

Seu principal objetivo é a avaliação do modelo ajustado. No caso dos GLMs, iremos verificar os seguintes itens:

Distribuição  
proposta

Parte  
sistemática  
proposta

Função de  
ligação

Efeito de  
observações  
mal  
ajustadas

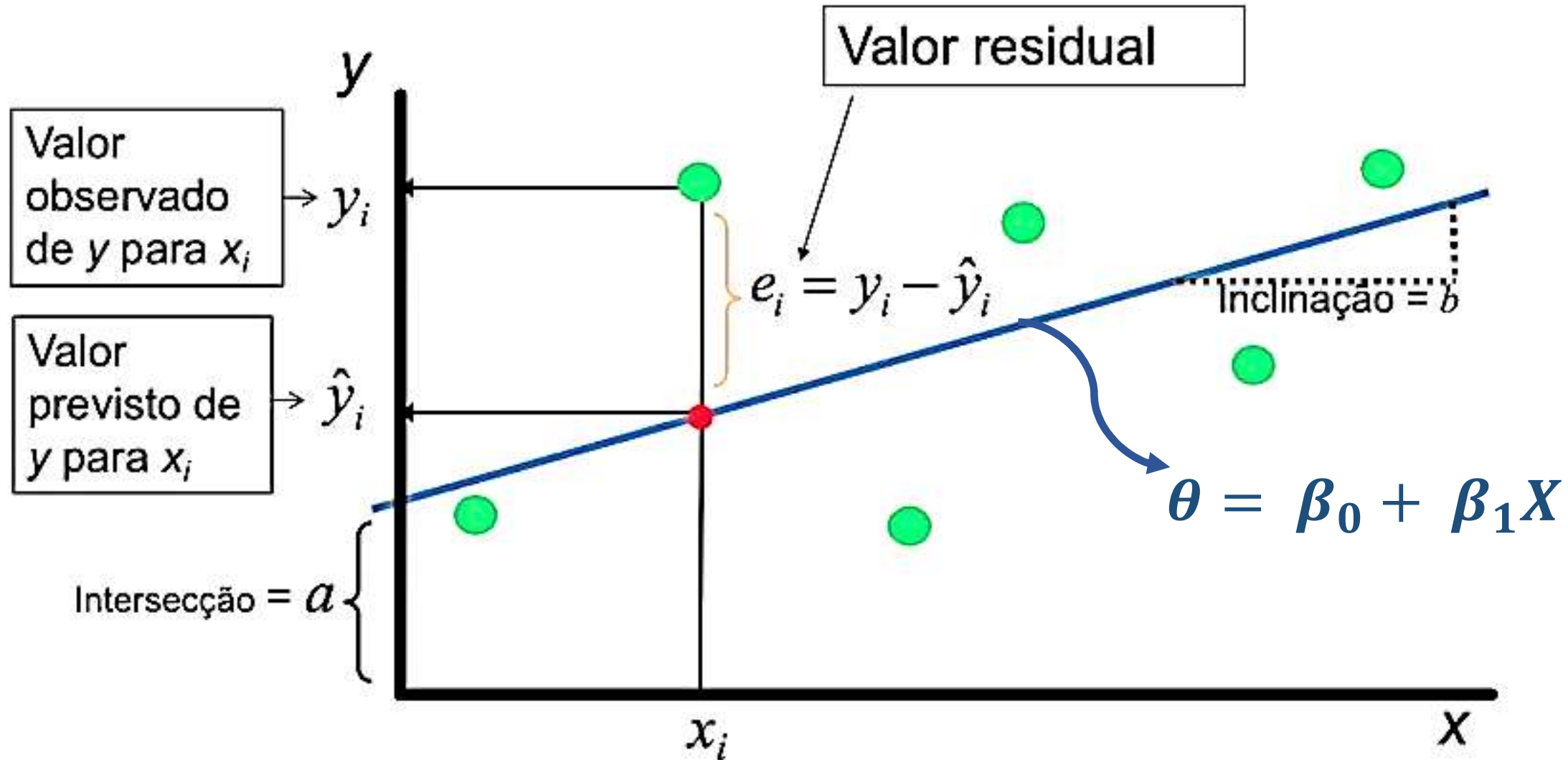
Identificação  
de pontos  
influentes

# Diagnóstico do ajuste

A maioria dos métodos de Diagnóstico em GLMs são extensões de análises utilizadas em regressão linear

Principal componente do diagnóstico é a Análise de Resíduos

# Resíduos no contexto de Regressão clássica



# Análise de Resíduos no Modelo Clássico

Dado o modelo clássico de regressão:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Sabe-se que  $\varepsilon = y_i - \hat{y}_i$

Temos que :  
 $\varepsilon \sim N(0, \sigma^2)$

Variação natural dos dados ou fatores que não foram considerados no modelo

Se as **pressuposição não são atendidas**, origina-se **falhas sistemáticas**(não-linearidade, heterocedasticidade, não-independência) e **presença de pontos atípicos**, que podem ou não influenciar o ajuste do modelos

# Análise de resíduos no modelo clássico

Considerando um modelo ajustado com  $p$  parâmetros, para verificarmos as pressuposições devemos considerar:

- Valores ajustados (estimados)  $\hat{\mu}_i$
- Resíduos -  $\varepsilon = y_i - \hat{\mu}_i$
- Variância residual :  $\hat{\mu}_i$
- Variância residual estimada :  $\hat{\sigma}^2 = s^2 = \sum \frac{(y_i - \hat{\mu}_i)^2}{n-p}$
- Os elementos da diagonal (*leverage*) da matriz de projeção  
 $H = X(X^T X)^{-1} X^T$ , isto é,  $h_{ii} = x_i (X^T X)^{-1} x_i^T$ , em que,  $x_i^T = (x_{i1}, \dots, x_{ip})$ .

Modelo ajustado com todos os pontos X modelo ajustado com os pontos atípicos

# Análise de resíduos nos GLMs

As técnicas utilizadas na análise de resíduo e diagnósticos para os GLMs são semelhantes aquelas usada para o modelo clássico, com algumas adaptações.

- Verificação da pressuposição de linearidade -  $\hat{z}$ (variável ajustada) e  $\hat{\eta}$ (componente linear)
- Variância residual estimada :  $s^2$  é substituída por uma estimativa consistente do parâmetro de dispersão  $\phi$ .
- A matriz de projeção  $H$  é definida por :  $H = W^{1/2}X(X^T W X)^{-1}X^T W^{1/2}$  (substituindo  $X$  por  $W^{1/2}X$ ).

**$H$  depende das variáveis explicativas, função de ligação e variância.**

Difícil interpretação da medida de *leverage*





**PUC Minas**  
**Virtual**