

# **I Workshop em Regressão Não Linear.**

**DES/UFLA**

**Redução de viés utilizando outliers em  
procedimentos bootstrap em modelos de  
regressão não linear.**

**Marcelo Ângelo Cirillo**

# 1- Modelos não-lineares aplicados nas Ciência dos Alimentos - (Curvas de isotermas)

**Isoterma** : relação entre a atividade de água ( $a_w$ ) e umidade do produto em uma determinada temperatura. **Cada produto tem uma isoterma única.**

Atividade de água ( $a_w$ ) está relacionada com a umidade relativa do ambiente (UR%) por meio da equação

$$a_w = \frac{P_w}{P_v}$$

→ Pressão parcial da água contida no alimento.  
→ Pressão de vapor da água pura a mesma temperatura.

**Isoterma** : Descreve a relação de equilíbrio entre o teor total de umidade de um alimento e a pressão de vapor ou umidade relativa, a uma dada temperatura.

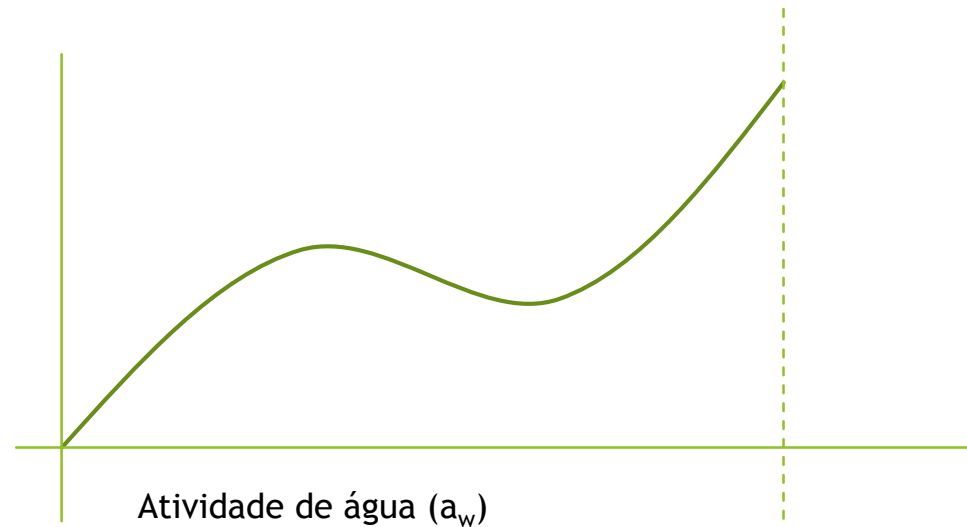
**Isoterma (Sorção)**: Submete um material mais seco, em condições ambientais aumentando a umidade relativa e medindo o aumento do peso devido ao ganho de água.

**Isoterma (Dessertação)**: Submete um material inicialmente úmido nas mesmas condições ambientais, sendo medida a perda de peso devido a saída de água.

**Isoterma (Dessertação) - Isoterma (Sorção) → Histerese**

## Exemplos

1

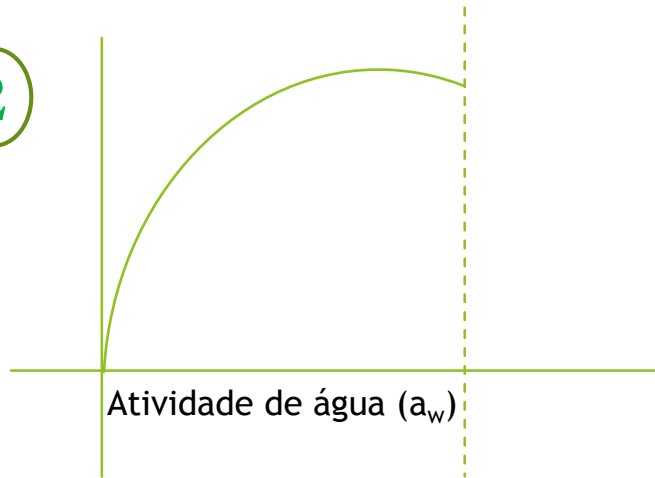


Geração: Obtida para produtos solúveis e mostra a tendência assintótica conforme a atividade de água aumenta.

O formato é  
causado por  
efeitos de aditivos

{ Efeito da capilaridade  
Interação da superfície de água.

2

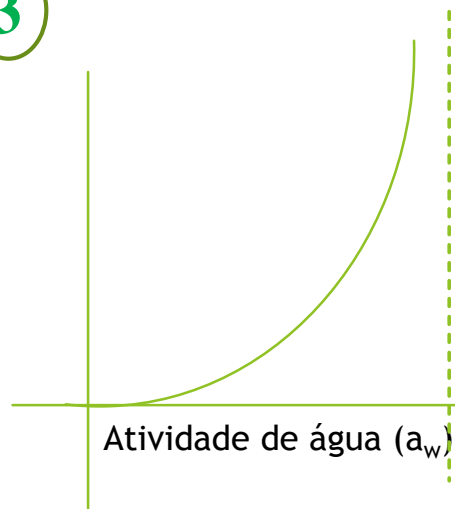


## Isoterma Langmuir

**Geração:** Obtida pela adsorção monomolecular de gás por sólidos porosos.

Típico de agentes anti-aglomerante.

3



## Isoterma formato “J”.

**Geração:** Representa a isoterma de material cristalino como açúcares e sais.

**Caracterização:** Ganho de umidade é muito baixa até o ponto onde os cristais começam a se dissolverem na água absorvida na superfície do cristal (ponto de deliquescência)

Tabela 1 - Modelos biparamétricos para modelagem de curvas de isotermas

Modelo	Equação
Kuhn (linear)	$y = -\frac{a}{\ln(a_w)} + b + \xi$
Halsey	$y = \left[ -\frac{a}{\ln(a_w)} \right]^b + \xi$
Oswin	$y = a \left[ \frac{a_w}{1-a_w} \right]^b + \xi$
Smith (Linear)	$y = a - (b \ln(1 - a_w)) + \xi$

$y$  representa o conteúdo de umidade (g H<sub>2</sub>O/g de sólidos secos);  $a_w$  indica a atividade de água;  $a$   $b$  = constantes características das equações dois parâmetros e  $\xi$  = vetor dos resíduos supostamente distribuídos por  $\xi \sim N(0, \sigma^2)$ . (CHIRIFE e IGLESIAS, 1978)

## 2- Objetivos

**Objetivo Principal :** Apresentar procedimentos com o propósito de reduzir o viés das estimativas.

**Objetivos Secundários :**

1 ° - **Procedimento bootstrap** que utilize outliers, mais especificamente, medidas de influência que incorpore informações aos pesos de reamostragem com o propósito de gerar estimativas de uma curva de isoterma com menor viés.

2 ° - **Teoria assintótica** para determinar melhor faixa de atividade de água que resulte em estimativas de uma curva de isoterma com menor viés.

### 3- Principais aspectos dos modelos de regressão não lineares

Os modelos não lineares são deduzidos a partir de **suposições teóricas** e os parâmetros resultantes são interpretáveis.

A forma clássica do modelo normal não-linear é dada por

$$y_i = f(\mathbf{B}; x) + \xi_i \quad (i = 1, \dots, n)$$

Obs.      Comp. Sist.      Erros i.i.d Normal



Seja a função  $S(\beta) = [Y - f(X, \beta)]^t [Y - f(X, \beta)]$

O estimador não linear de mínimos quadrados é a função

$$S(\hat{\beta}) \leq S(\beta) \quad \forall \beta \in B$$

Se o mínimo de  $S(\beta)$  ocorrer em um ponto interior de  $B$  o estimador NLS verifica a condição em  $\beta$  que

$$Z^t [Y - f(X, \beta)] = 0 \quad Z = \begin{bmatrix} \frac{\partial f(X_1, \beta)}{\partial \beta_1} & \frac{\partial f(X_1, \beta)}{\partial \beta_2} & \dots & \frac{\partial f(X_1, \beta)}{\partial \beta_k} \\ \frac{\partial f(X_2, \beta)}{\partial \beta_1} & \frac{\partial f(X_2, \beta)}{\partial \beta_2} & \dots & \frac{\partial f(X_2, \beta)}{\partial \beta_k} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f(X_n, \beta)}{\partial \beta_1} & \frac{\partial f(X_n, \beta)}{\partial \beta_2} & \dots & \frac{\partial f(X_n, \beta)}{\partial \beta_k} \end{bmatrix}$$

OBS: O SEN de um modelo de regressão linear  $\mathbf{X}^t \mathbf{X} \boldsymbol{\beta} - \mathbf{X}^t \mathbf{Y} = 0$

é um caso particular de  $\mathbf{Z}^t [\mathbf{Y} - f(\mathbf{X}, \boldsymbol{\beta})] = 0$

No modelo linear  $f(\mathbf{X}, \boldsymbol{\beta}) = \mathbf{X}\boldsymbol{\beta}$  e  $\mathbf{Z} = \frac{\partial f(\mathbf{X}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \mathbf{X}$

$$\mathbf{Z}^t [\mathbf{Y} - f(\mathbf{X}, \boldsymbol{\beta})] = 0 \quad \longrightarrow \quad \mathbf{X}^t [\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}] = 0$$

Assim, como um modelo linear temos  $\hat{\boldsymbol{\beta}} = (\mathbf{Z}_0^t \mathbf{Z}_0)^t \mathbf{Z}_0^t \mathbf{Y}_0$

No algoritmo de Gauss-Newton, para a iteração seguinte o modelo seria linearizado numa vizinhança de  $\hat{\boldsymbol{\beta}}_1 = (\mathbf{Z}_1^t \mathbf{Z}_1)^t \mathbf{Z}_1^t \mathbf{Y}_1$

## Método de Gauss-Newton

Fundamentado em aproximações lineares de um modelo não linear

Definição: Aproximação de 1 ordem do modelo não linear

$$f(\mathbf{X}; \underline{\beta}) \approx f(\mathbf{X}; \underline{\beta}_0) + \mathbf{Z}_0(\underline{\beta} - \underline{\beta}_0),$$

$$f(\mathbf{X}; \underline{\beta}) \approx \left[ f(\mathbf{X}; \underline{\beta}_0) - \mathbf{Z}_0 \underline{\beta}_0 \right] + \mathbf{Z}_0 \underline{\beta}_0$$

Temos que  $\underline{Y} = f(\mathbf{X}; \underline{\beta}) + \underline{\xi}$  e substituindo a expressão acima

$$\underline{Y} \approx \left[ f(\mathbf{X}; \underline{\beta}_0) - \mathbf{Z}_0 \underline{\beta}_0 \right] + \mathbf{Z}_0 \underline{\beta}_0 + \underline{\xi} \quad \text{ou}$$

$$\underbrace{\underline{Y} - f(\mathbf{X}; \underline{\beta}_0) + \mathbf{Z}_0 \underline{\beta}_0}_{\underline{Y}_0} \approx \mathbf{Z}_0 \underline{\beta}_0 + \underline{\xi}$$

## 4- Identificação de observações influentes e discrepantes

Projektor ortogonal para o modelo linear

$$\hat{\underset{\sim}{Y}} = \underset{\sim}{X} \hat{\underset{\sim}{\beta}} = \underset{\sim}{X} (\underset{\sim}{X}^t \underset{\sim}{X})^{-1} \underset{\sim}{X}^t \underset{\sim}{Y} = H \underset{\sim}{Y}$$

$$H = \underset{\sim}{X} (\underset{\sim}{X}^t \underset{\sim}{X})^{-1} \underset{\sim}{X}^t$$

Vetor residual escrito em função do projetor

$$\underset{\sim}{\xi} = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}} = \underset{\sim}{Y} - \underset{\sim}{X} \hat{\underset{\sim}{\beta}}$$

$$= \underset{\sim}{Y} - \underset{\sim}{X} (\underset{\sim}{X}^t \underset{\sim}{X})^{-1} \underset{\sim}{X}^t \underset{\sim}{Y}$$

$$= \underset{\sim}{Y} - H \underset{\sim}{Y} = (\mathbf{I}_n - H) \underset{\sim}{Y} \Rightarrow \underset{\sim}{\xi} = M \underset{\sim}{Y}$$

OBS: Valor predito e o resíduo são independentes, pois estão em planos ortogonais.

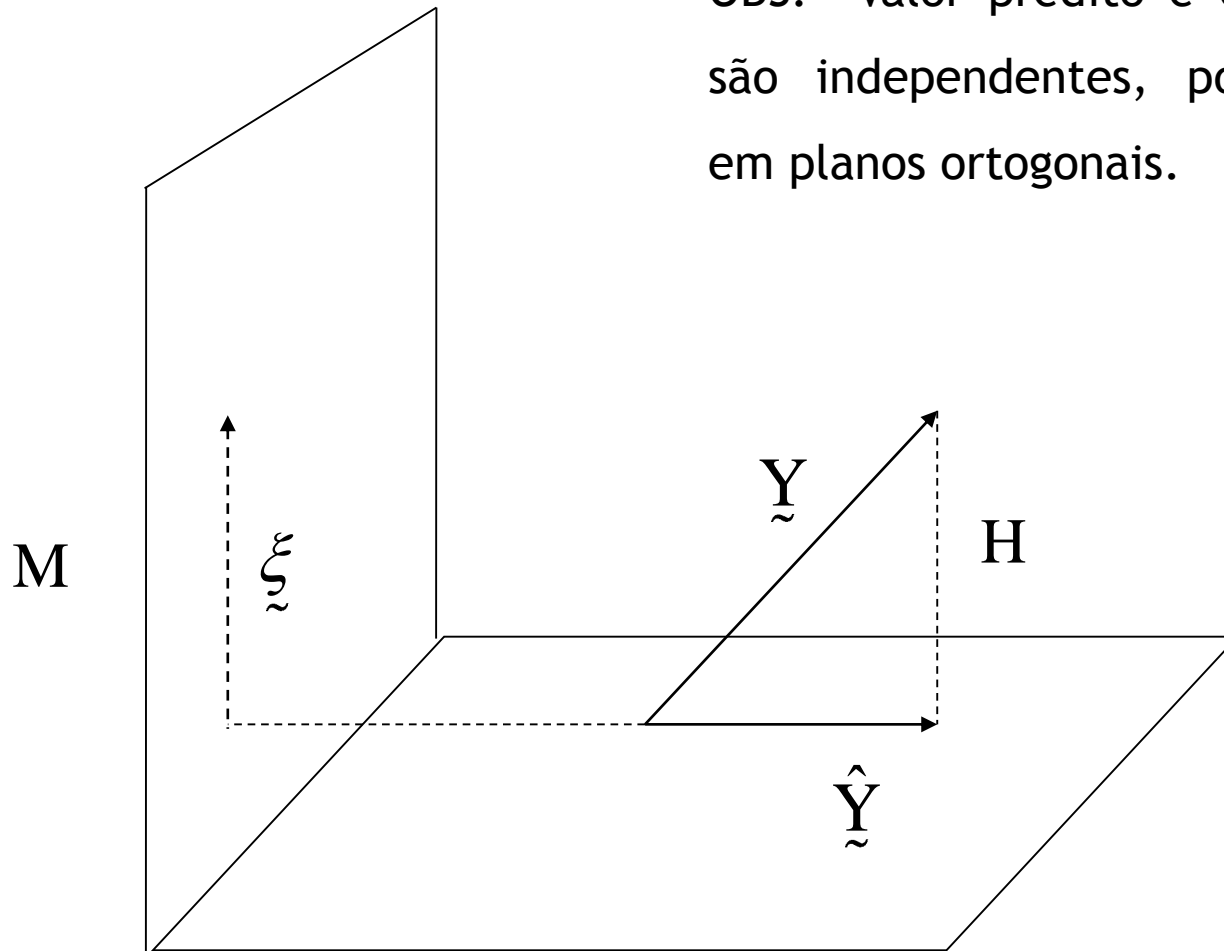


Figura 1 - Interpretação geométrica do projetor  $H$

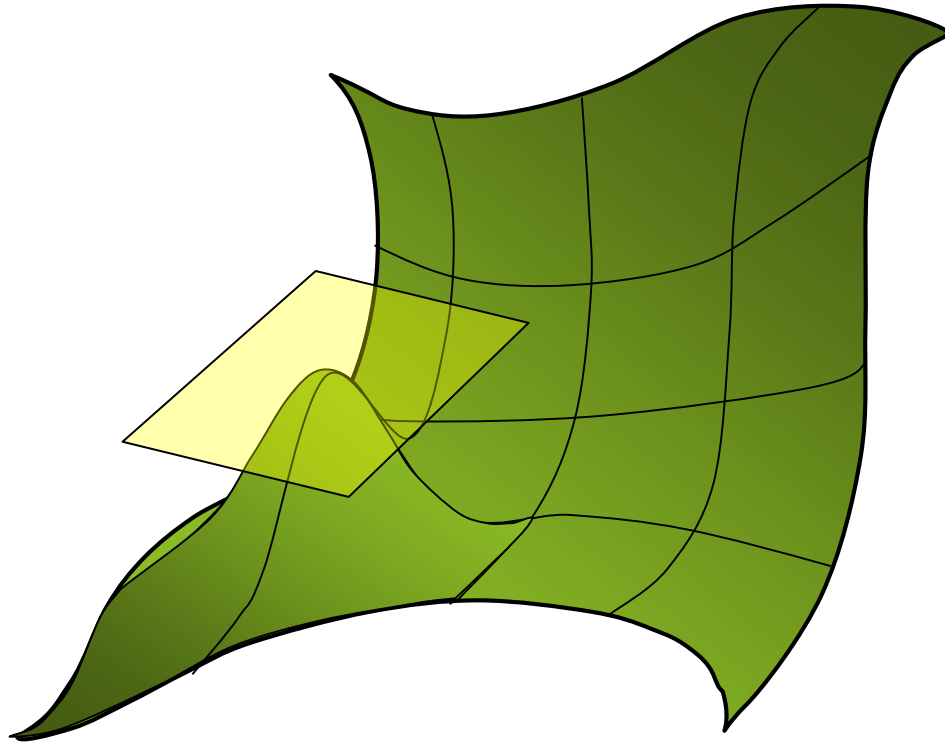


Figura 2 - Projeção não linear é feita no plano tangente a um ponto na curva

$$H = Z(\beta) \left( Z(\beta)^t Z(\beta) \right)^{-1} Z(\beta)^t$$



## 5 – Algumas medidas de influência

### Medida de Leverage ( $h_{ii}$ )

A medida de leverage ( $h_{ii}$ ) mede a importância da  $i$ -ésima observação na determinação do ajuste do modelo.

**COVRATIO**



$$COVRATIO_i = \frac{\det \left[ \text{QMRes}(-i) \left( \mathbf{X}_{(-i)}^t \mathbf{X}_{(-i)} \right)^{-1} \right]}{\det \left[ \text{QMRes} \left( \mathbf{X}^t \mathbf{X} \right)^{-1} \right]}$$

**Distância de Cook**



$$D_i = \frac{\left( \hat{\beta}_{(i)} - \hat{\beta}_{(-i)} \right)^t \mathbf{X}^t \mathbf{X} \left( \hat{\beta}_{(i)} - \hat{\beta}_{(-i)} \right)}{p \text{QMRes}}$$

**Dffits**



$$Dffits_i = \frac{\left( \hat{y}_{(i)} - \hat{y}_{(-i)} \right)}{\sqrt{h_{ii} \text{QMRes}_{(-i)}}}$$

## 6- Método *bootstrap* em modelos de regressão

Dada uma amostra de tamanho  $n$  representada pelas ternas

$(x_1, y_1, \xi_1)$   $(x_2, y_2, \xi_2)$   $(x_3, y_3, \xi_3)$   $(x_4, y_4, \xi_4)$   $\dots$   $(x_n, y_n, \xi_n)$

(1) (2) (3) (4)  $\dots$  (n)

Sub\_amostras de tamanho  $n$

(1), (3), (3),  $\dots$ , (7)

(5), (5), (2),  $\dots$ , (10)

(6), (4), (2),  $\dots$ , (3)

↓  $\beta_{1b}$

↓  $\beta_{2b}$

↓  $\beta_{kb}$

**Distribuição empírica dos parâmetros**

Causas e “solução”

Dados incompletos



Introdução de censura

Dependência



Reamostragem por blocos



## 6 - Algoritmo *bootstrap* incorporando medidas de influência

### Staide & Shater (1990)

**Passo 1:** Determinar a função de influência (FI) do estimador  $\beta$ ,

**Passo 2:** Calcular a função de influência empírica de cada uma

das observações  $\hat{\mathbf{FI}} = \left( \hat{\mathbf{FI}}_1, \dots, \hat{\mathbf{FI}}_n \right)$


**Passo 3:** Atribuir a cada observação  $x_i$ , o peso  $w_i$

$$w_i = I_{[0,c]} \left( \left| \hat{\mathbf{FI}}_i \right| \right) + \left( c / \left| \hat{\mathbf{FI}}_i \right| \right) \times I_{]c,+\infty[} \left( \left| \hat{\mathbf{FI}}_i \right| \right)$$

**Passo 4:** Obter  $P = (p_1, p_2, \dots, p_n)$  em que,

$$p_i = \frac{w_i}{\left( \sum_{j=1}^n w_j \right)}$$

## 7- Metodologia

**Modelo de Oswin**   $y_i = a \left[ \frac{a_w}{1 - a_w} \right]^b + \xi_i$

$y$  = conteúdo de umidade (Exemplo: g H<sub>2</sub>O/g de sólidos secos)

$a_w$  = atividade de água ;  $a, b$  = parâmetros característicos

Para “forçar” a presença de observações atípicas no modelo os valores de  $a_w$  foram gerados da seguinte forma:

$$a_w = (1-\gamma)U_{(0,1)} + \gamma\text{Beta}(0.2,0.8)$$

## Prática no R

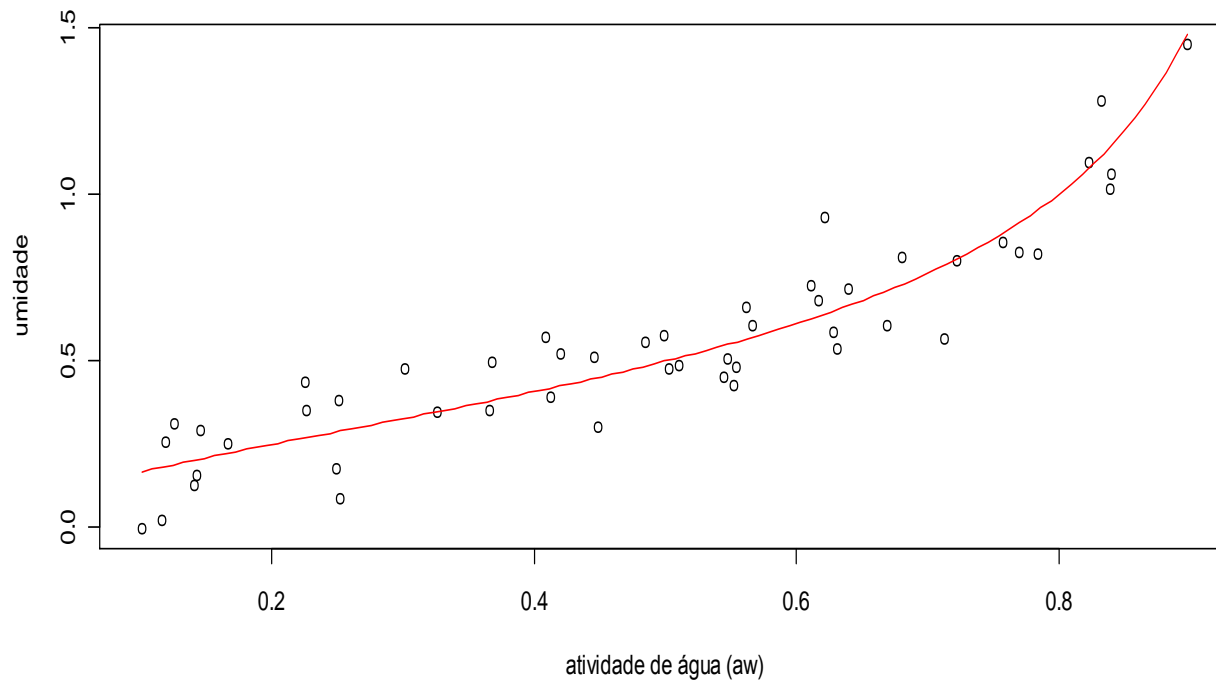


Figura 3 - Diagrama de dispersão de observações geradas pelo modelo de Oswin.

# Teoria Assintótica

---

Fundamentos para entendimento da  
redução de viés em simulação Monte Carlo

---

Notação de ordens de magnitude  $o(\cdot)$  e  $O(\cdot)$

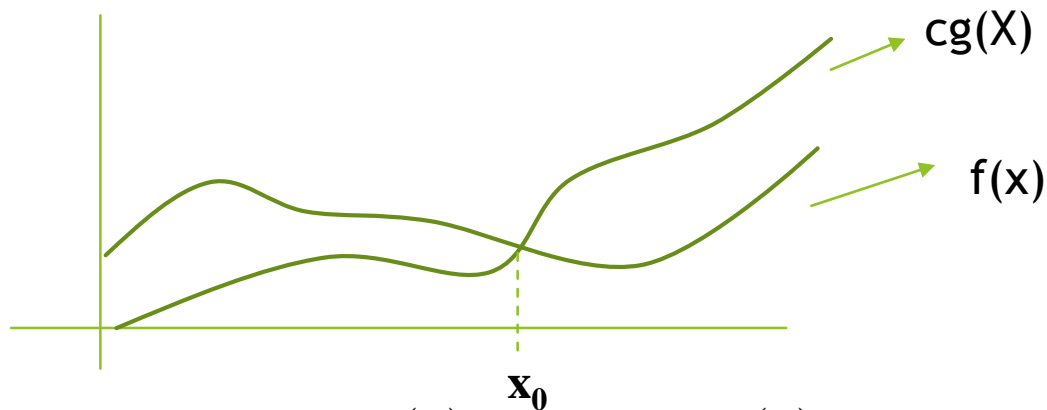
Expansão assintótica da série de Taylor

Equivalência assintótica entre os modelos

$f(x) \sim g(x) \Rightarrow \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1 \implies f(x)$  é assintoticamente igual a  $g(x)$

$f(x) = O(g(x)) \Rightarrow \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} < \infty$  ou  $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} < \infty$

$f(x)$  : Existe constantes positivas  $c$  e  $x_0$  de modo que, para qualquer  $x \geq x_0$ ,  $f(x) \leq cg(x)$



$f(x) = o(g(x)) \Rightarrow \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$  ou  $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$


$f(x)$  : Existe constantes positivas  $c$  e  $x_0$  de modo que, para qualquer  $x \geq x_0$ ,  $f(x) < cg(x)$

## Comportamento assintótico do erro de truncamento na série de Taylor.

$$f(x) = f(k) + f'(k)(x-k) + \frac{f''(k)(x-k)^2}{2!} + \dots + \frac{f^{(n)}(k)(x-k)^n}{n!} + R_{n+1}(k)$$




sendo 
$$R_{n+1}(k) = \frac{f^{(n+1)}(x-k)^{n+1}}{(n+1)!}$$

A questão é como representar  $R_{n+1}(k)$  em  $o(\cdot)$  ou  $O(\cdot)$

$f: \mathbb{R} \rightarrow \mathbb{R}$    $\lim_{h \rightarrow 0} \left( \frac{f(k+h) - f(k)}{h} \right) = f'(k)$

A função  $f(x)$  é derivável em  $k$  quando é possível escrever

$$f(k+h) = f(k) + T \times h + r(h) \quad ; \quad r(h) \rightarrow \lim_{h \rightarrow 0} \left( \frac{r(h)}{h} \right) = 0$$

    
Transf. erro resto  
Linear

$$\lim_{h \rightarrow 0} \left( \frac{r(h)}{h} \right) = \lim_{h \rightarrow 0} \left( \frac{f(k+h) - f(k)}{h} - \frac{Th}{h} \right) = f'(k) = T$$

Exemplo : Expansão de Taylor de primeira ordem e notação  $o(\ )$

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad \left\{ \begin{array}{l} k+h = x \\ k = 0 \\ h = x \end{array} \right. \quad \begin{array}{l} f(k+h) = f(k) + f'(k) \times h + r(h) \\ e^k \approx 1 + x + r(x) \\ e^k \approx 1 + x + o(k) \end{array}$$

Exemplo : Expansão de Taylor de segunda ordem e notação  $o(\ )$

$$f(k+h) = f(k) + f'(k) \times h + \frac{f''(k)h^2}{2!} + r(h)$$

$$r(h) = \lim_{h \rightarrow 0} \frac{r(h)}{h^2} = 0 \Leftrightarrow r(h) = o(h^2)$$

$$e^x \approx 1 + x + x^2 + o(x^2)$$

## 7- Expansão de Taylor para o modelo de Oswin em torno de $k = a_w$

$$y = \ln(a) + b \ln\left(\frac{a_w}{1 - a_w}\right)$$

### Expansão de Taylor de segunda ordem

$$h(k) = y + \frac{b\left(\frac{1}{(1-k)} + \frac{k}{(1-k)^2}\right)(1-k)(y-k)}{k} + \left[ \frac{b\left(\frac{2}{(1-k)^2} + \frac{2k}{(1-k)^3}\right)(1-k) - b\left(\frac{1}{(1-k)} + \frac{k}{(1-k)^2}\right)(1-k) - b\left(\frac{1}{(1-k)} + \frac{k}{(1-k)^2}\right)}{k^2} \right] \frac{(y-k)^2}{2!}$$

### Reorganizando os termos

$$h(k) = y + \frac{b\left(\frac{1}{(1-k)}\right)(y-k)}{k} + \left[ \frac{b\left(\frac{2}{(1-k)^2}\right) - b\left(\frac{1}{(1-k)}\right) - b\left(\frac{1}{(1-k)^2}\right)}{k^2} \right] \frac{(y-k)^2}{2!}$$



$$h(k) = y + \underbrace{\left(\frac{b}{(1-k)}\right) \frac{(y-k)}{k}}_{C_1} + \underbrace{\left[ \left(\frac{2b}{(1-k)^2 k}\right) - \left(\frac{b}{(1-k)k^2}\right) - \left(\frac{b}{(1-k)^2 k}\right) \right] \frac{(y-k)^2}{2!}}_{C_2}$$

Calculando o limite temos

$$\lim_{y \rightarrow k} \frac{C_1(y-k)}{C_2|y-k|} = \frac{C_1}{C_2} \lim_{y \rightarrow k} \frac{y-k}{(y-k)^2} = 0$$

Logo, temos

$$o(|y-k|) = \left(\frac{2b}{(1-k)^2 k}\right) - \left(\frac{b}{(1-k)k^2}\right) - \left(\frac{b}{(1-k)^2 k}\right) \frac{(y-k)^2}{2!}$$

$$h(k) = y - \left(\frac{b}{(1-k)k}\right)(y-k) + o(|y-k|)$$

## 8- Expansão de Taylor para o modelo de Smith em torno de um valor $k = (a_w)$

$$y = a - (b \ln(1 - k))$$

$$g(k) = y + b \left( \frac{1}{1-k} \right) (y - k) + b \left( \frac{1}{(1-k)^2} \right) \frac{(y - k)^2}{2!} + 2b \left( \frac{1}{(1-k)^3} \right) \frac{(y - k)^3}{3!}$$

Separando os termos

$$C_3 = y + b \left( \frac{1}{1-k} \right); \quad C_4 = b \left( \frac{1}{(1-k)^2} \right) \frac{(y - k)^2}{2!} + 2b \left( \frac{1}{(1-k)^3} \right)$$

Calculando os limites

$$\lim_{y \rightarrow k} \frac{C_3(y - k)}{C_4 |(y - k)|} = \frac{C_3}{C_4} \lim_{y \rightarrow k} \frac{y - k}{(y - k)^2} = 0$$

$$o(|y-k|) = b \left( \frac{1}{(1-k)^2} \right) \frac{(y-k)^2}{2!} + 2b \left( \frac{1}{(1-k)^3} \right) \frac{(y-k)^3}{3!}$$

$$g(k) = y + b \left( \frac{1}{1-k} \right) (y-k) + o(|y-k|)$$

### Relação das aproximações das séries do modelo de Oswin e Smith

$$m_1 = \left[ \left( \frac{2b}{(1-k)^2 k} \right) - \left( \frac{b}{(1-k) k^2} \right) - \left( \frac{b}{(1-k)^2 k} \right) \right] \frac{(y-k)^2}{2!}$$

$o(|y-k|)$  - Oswin

$o(|y-k|)$  - Smith

$$m_2 = b \left( \frac{1}{(1-k)^2} \right) \frac{(y-k)^2}{2!} ; m_3 = 2b \left( \frac{1}{(1-k)^3} \right) \frac{(y-k)^3}{3!}$$

### Forçando a igualdade entre os modelos

$$\frac{o(|y-k|)}{o(|y-k|)} = \frac{m_2(y-k)^2 + m_3(y-k)^3}{m_1(y-k)^2} = \frac{m_2}{m_1} + \frac{m_3(y-k)}{m_1}$$

## Calculando o limite

$$\lim_{y \rightarrow k} \left( \frac{m_2}{m_1} + \frac{m_3}{m_1} \frac{(y-k)}{(y-k)} \right) = \frac{m_2}{m_1} + 1 \quad \longrightarrow \quad \frac{m_2}{m_1} + 1 = 0 \Rightarrow m_2 = -m_1$$

## Substituindo $m_2$ e $m_1$

$$b \left( \frac{1}{(1-k)^2} \right) \frac{(y-k)^2}{2!} = - \left[ \left( \frac{2b}{(1-k)^2 k} \right) - \left( \frac{b}{(1-k)k^2} \right) - \left( \frac{b}{(1-k)^2 k} \right) \right] \frac{(y-k)^2}{2!}$$

$$2k^3 - 2k^2 - k + 1 = 0$$

Uma das raízes encontra-se no intervalo **(0,7 ; 0,8)**, considerou-se este intervalo como uma faixa apropriada para os níveis de  $a_w$

Tabela 2 - Vieses relativos das estimativas dos parâmetros do modelo de Oswin.

Faixa de atividade de água ( $a_w$ )	Tamanho amostral	Viés de a	Viés de b	Numero de simulações válidas
(0,7 ; 0,8)	20	0,7348	0,1817	5619
	50	0,7771	0,1765	7566
	100	0,7869	0,1739	8896
(0,1 ; 1,0)	20	0,8872	0,1883	5138
	50	0,9035	0,2031	6684
	100	0,9087	0,2053	7731

Tabela 3 -Vieses relativos das estimativas dos parâmetros do modelo de Smith

Faixa de atividade de água ( $a_w$ )	Tamanho amostral	Viés de a	Viés de b
(0.7-0.8)	20	0,0053	0,0036
	50	0,0389	0,0273
	100	0,0184	0,0112
(0,1 ; 1,0)	20	0,0048	0,0025
	50	0,0008	0,0016
	100	0,0004	0,0007

## Referências Bibliográficas

Andrade, L.R.; Cirillo, M.A; Beijo, L.A. (2013) Proposal of bootstrap regression using measures of influence in non-linear regression models with outliers. *Acta Scientiarum - Technology*. v.36,p.93-97.

Aguiar, O.I.; Cirillo, M.A. Redução do Vício Nas Estimativas Dos Parâmetros Em Modelos Biparamétricos Aplicados Na Ciência Dos Alimentos. Relatório Final PIBIC/CNPq, 2010.

Chirife, J.; Iglesias,H. A. (1978) *Equations for fitting water sorption isotherms of foods: Part 1 - a review*. *J. Food technol.* , v.13, n.3, p.159-174.

Cook, R.D.; Weisberg, S. (1982). *Residuals and Influence in Regression*, New York: Chapman and Hall.

Davison, A. C. ; Hinkley, D.V. (1997). *Bootstrap Methods and their Application*. New York: Cambridge University Press.

Hampel, F. R.; Ronchetti, E. M.; Rousseeuw, P. J.; Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley.

Rahman, M.S. (1995). *Food Properties Handbook*. 1st ed. CRC Press, Boca Raton.

Development Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Staudte, R. G. e Sheather, S. J. (1990). *Robust Estimation and Testing*. New York: John Wiley.

Treybal, R.E. (1981). *Mass Transfer Operations*. 3rd ed. Auckland: McGraw-Hill



---

**OBRIGADO**

---