



TÉCNICAS EXPERIMENTAIS APLICADAS EM CIÊNCIA DO SOLO

Mario de Andrade Lira Junior

[lira.pro.br\wordpress](http://lira.pro.br/wordpress)



REGRESSÃO X CORRELAÇÃO

○ Diferença

- Regressão - equação ligando duas ou mais variáveis
- Correlação – medida do grau de ligação entre duas variáveis

○ Usos

- Regressão – estimar valores intermediários aos realmente estudados durante o experimento
- Correlação – indicar variáveis com comportamento semelhante



COEFICIENTE DE CORRELAÇÃO

- Mede o grau de relação linear entre variáveis
- Interpretação
- Testes de significância
 - Análise de variância da regressão completa
 - No computador, cada componente é testado pelo teste de t



ANÁLISE DE REGRESSÃO

- ▶ É importante diferenciar entre testes de “significância” e “importância”
 - ▶ Significância – chance de acontecer devido ao acaso
 - ▶ Importância – quanto a regressão realmente explica
- ▶ Em modelos de regressão avaliar a importância científica costuma ser mais importante do que a significância
- ▶ Em modelos lineares a importância é principalmente definida por:
 - Proporção da variância atribuída ao modelo
 - O tamanho de um ou mais coeficientes de correlação
 - Intervalos de confiança de interesse



TIPOS DE REGRESSÃO

- ▶ Linear

$$Y_{ij} = \alpha + \beta X_i + \varepsilon_i$$

- ▶ Polinomial

$$Y_{ij} = \alpha + \beta_1 X_i + \beta_2 X_i^2 + \dots + \beta_l X_i^l + \varepsilon_i$$

- ▶ Múltiplo

$$Y_{ij} = \alpha + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$$

- ▶ Modelos não-lineares
 - Exponencial
 - Logarítimico



LIMITAÇÕES DOS TIPOS MAIS COMUNS

- A linear raramente representa bem toda uma série de dados
 - Mas pode e frequentemente representa bem faixas de valores
- Regressões polinomiais não têm interpretação biológica para os parâmetros
 - Úteis como simplificação de situação real
 - Podem ser usadas para estimar valores com significado biológico/prático
 - Polinomiais cúbicas ou mais complexas raramente são boas descritoras de fatos biológicos



REQUISITOS DA REGRESSÃO LINEAR

- ▶ Variável independente medida sem erro
- ▶ O valor esperado de Y é descrito pela função linear de X
- ▶ Para cada X_i os Y 's têm resíduos
 - Independentes
 - Normalmente distribuídos com média zero
 - Homocedástico – variância aproximadamente constante



FONTES DE VARIAÇÃO

- ▶ Médias são função da variável independente
- ▶ Acaso observado pela diferença entre o ponto real e o previsto

Valores ajustados e resíduos

- ▶ Valor ajustado - obtido pela equação estimativa da população
- ▶ Resíduo- diferença entre ajustado e real=acaso
- ▶ Comparação entre resíduo e variável independente é útil para visualizar ajuste do modelo



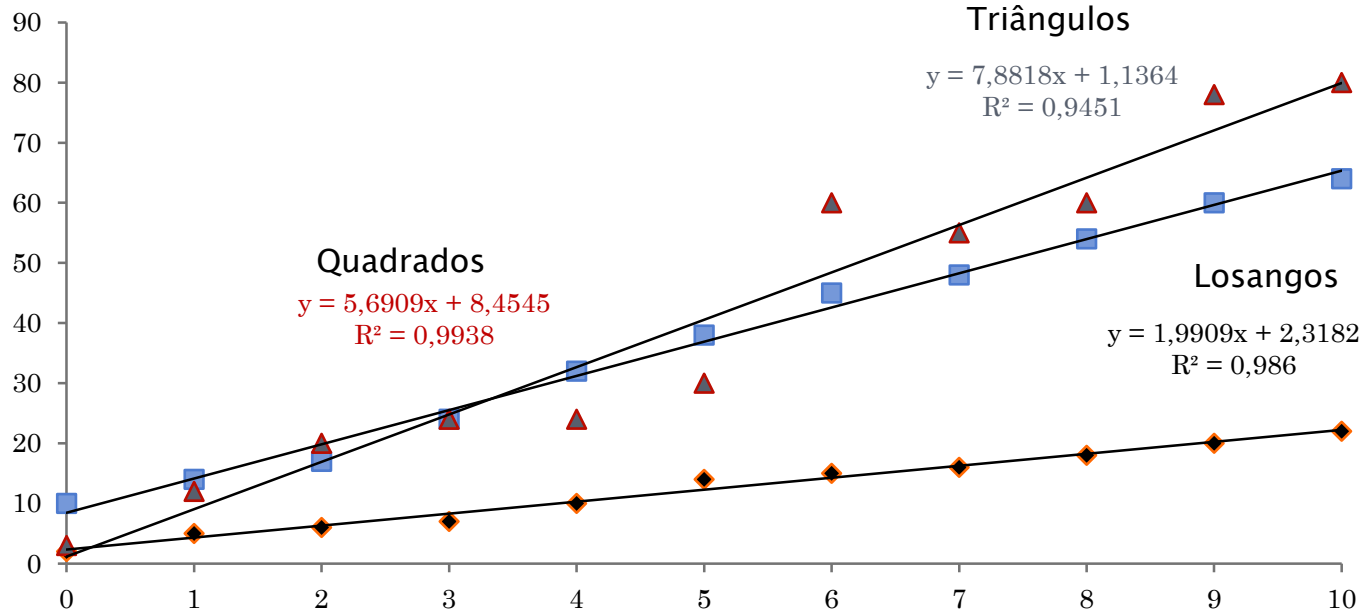
INTERPRETAÇÃO REGRESSÃO LINEAR

$$Y = a + bX$$

- Y – variável dependente
- a – y para x = 0
- b – quanto y varia para cada x
- r – coeficiente de correlação
 - Varia de -1 a 1
 - Quanto da variação de y é explicada por x
- r² – coeficiente de determinação
 - Quanto de y é explicado pela regressão
 - Varia de 0 a 1



INTERPRETAÇÃO



INTERPRETAÇÃO DE ANÁLISE DO OUTPUT DO SAS EM UMA REGRESSÃO LINEAR

Model: MODEL1 - Dependent Variable: _800125888

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	58120	58120	53,45	<,0001
Error	18	19572	1087,32861		
Corrected Total	19	77692			

ANAVA
ou
ANOVA

Root MSE	32,97467	R-Square	0,7481
Dependent Mean	159,31150	Adj R-Sq	0,7341
Coeff Var	20,69823		

Síntese

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr> t	Standardized Estimate
Intercept	1	83,07500	12,77103	6,50	<,0001	0
N	1	0,76237	0,10428	7,31	<,0001	0,86492

Estimativa e
significância

Parameter Estimates

Variable	DF	95% Confidence Limits	
Intercept	1	56,24405	109,90595
N	1	0,54329	0,98144

Intervalos
confiança



MODELOS NÃO LINEARES

- Processos interativos
 - Não tem como separar os coeficientes
- Bons descritores de fenômenos biológicos
- Uso bem mais complexo
- Frequentemente derivadas de modelagem mecanicista
- Parâmetros têm significado biológico
- Cálculo
 - Os cálculos variam de equação para equação
 - Sugiro programas especializados (SigmaPlot)
 - Bons programas gerais como SAS também dão opção



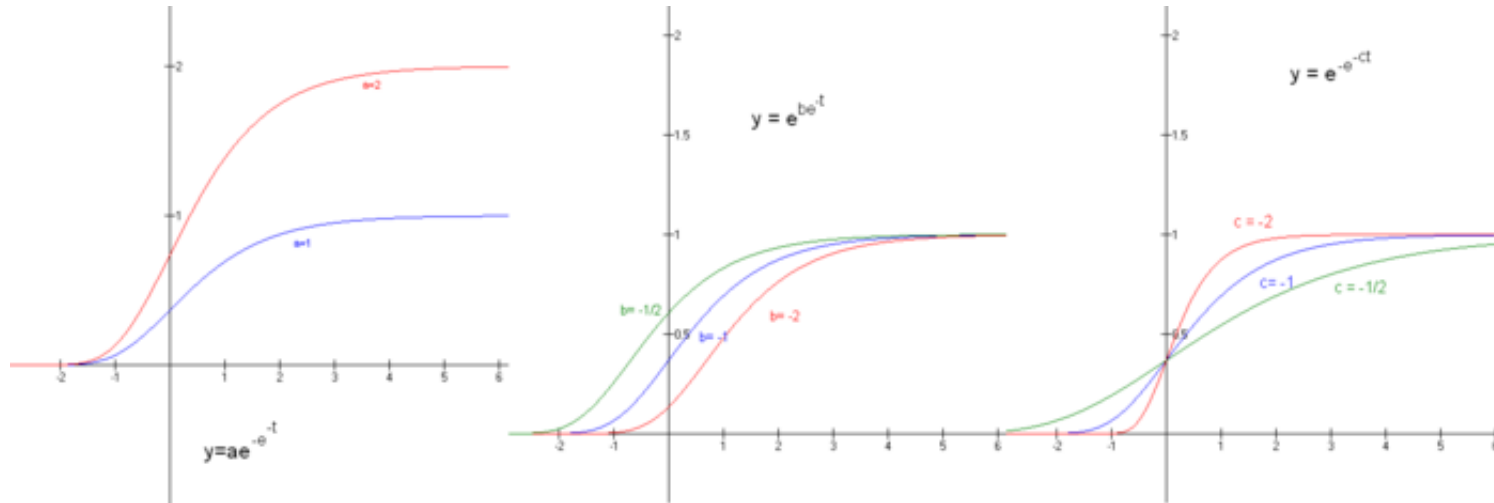
GOMPERTZ

- Curva de crescimento com fases inicial e final lentas
- a é a assíntota do crescimento
- c é a taxa de crescimento
- b e c são constantes negativas
 - Corrigem para redução do crescimento no tempo
- e é a constante neperiana

$$y(t) = ae^{be^{ct}}$$



EXEMPLO DE GOMPERTZ



- http://en.wikipedia.org/wiki/Gompertz_curve



REGRESSÃO LOGÍSTICA

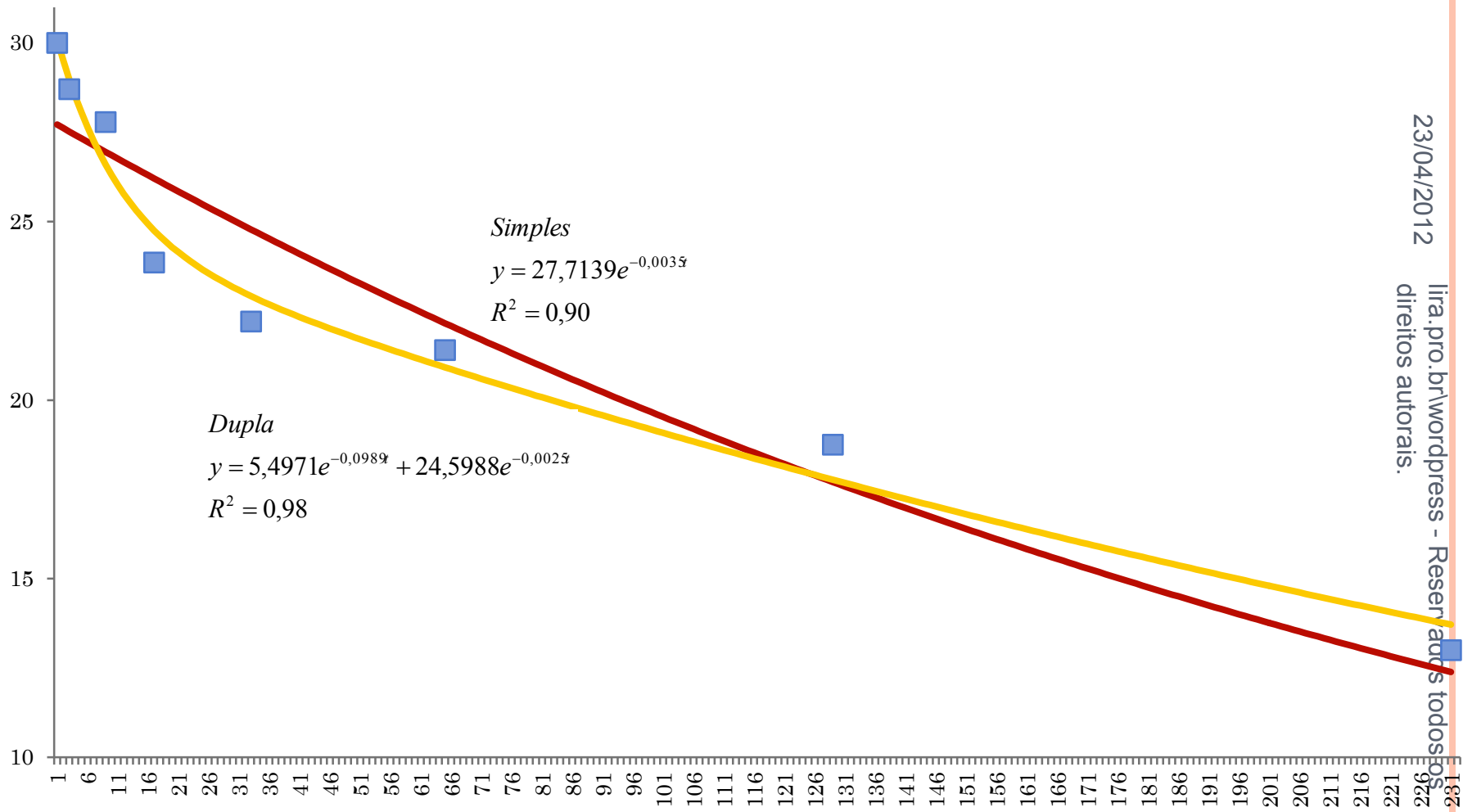
- Também modelo de crescimento
- Crescimento inicial aproximadamente exponencial seguido por redução do crescimento pela competição até estabilização
- Também pode ajudar no estudo de reações autocatalíticas
- Alguns modelos específicos são
 - Verhultz – crescimento populacional
 - Sigmoidal



QUEDA EXPONENCIAL

- ▶ Modelo básico para decomposição de matéria orgânica e liberação de nutrientes
- ▶ Casos típicos queda exponencial simples ou dupla
- ▶ $y = ae^{-bx}$
 - a – pool de elementos
 - b – taxa de decomposição
 - e – constante neperiana
- ▶ $y = ae^{-bx} + ce^{-dx}$
 - c – pool de elementos de decomposição lenta
 - d – taxa de decomposição deste segundo pool





23/04/2012
 lira.pro.br/wordpress - Reserva de todos os direitos autorais.



HIPERBÓLICA

- Adequada para casos em que tende a uma constante
- As constantes também apresentam interpretação biológica pré-definida
- Dividem-se em dois tipos básicos

- Crescimento

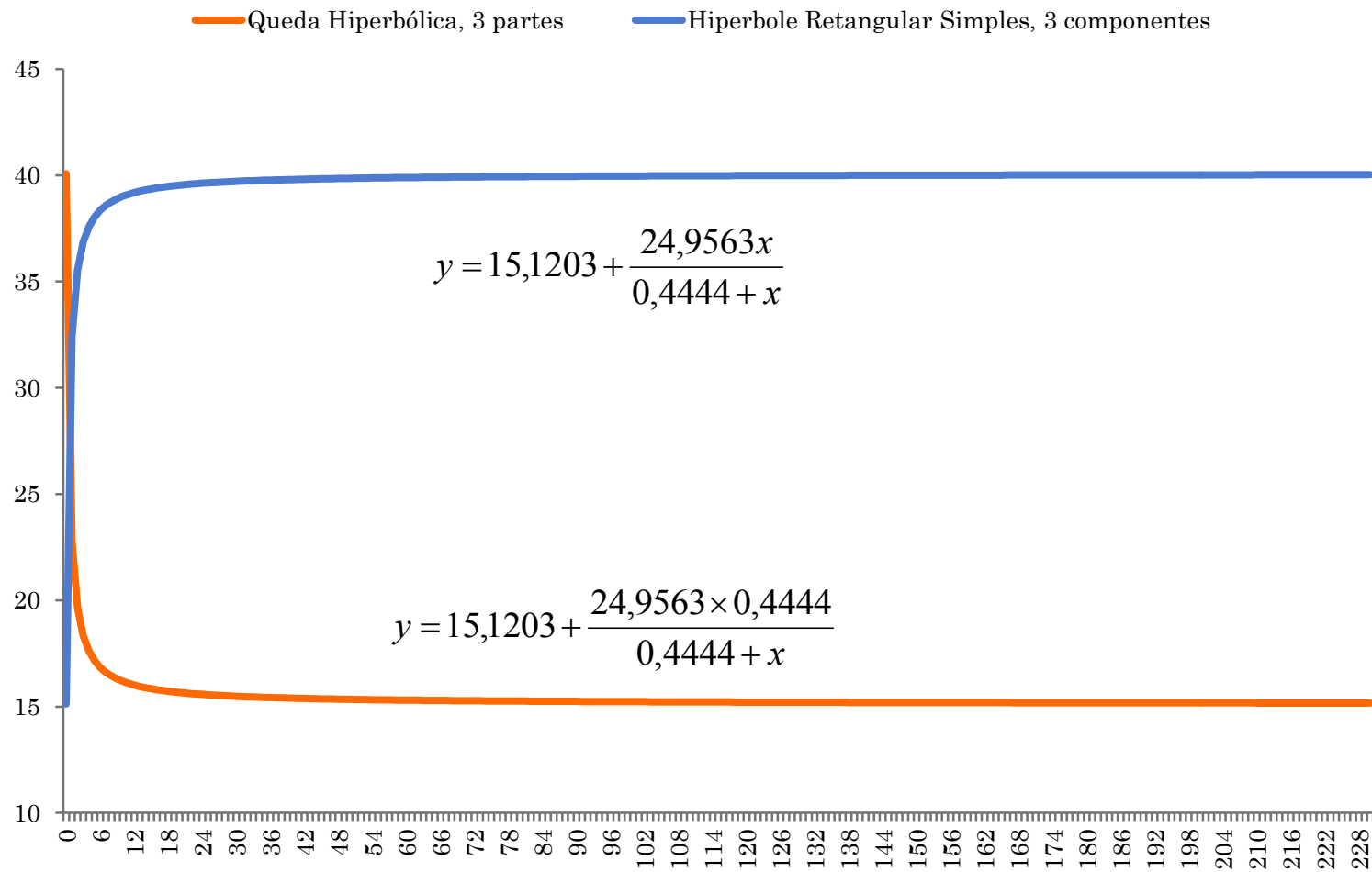
$$y = y_0 + \frac{ax}{b+x}$$

- Queda

$$y = y_0 + \frac{ab}{b+x}$$



CURVAS HIPERBÓLICAS



LITERATURA RECOMENDADA

- Muller e Fetterman Regression and ANOVA. An integrated approach using SAS software
 - Capítulo 2
 - Capítulo 4
 - Capítulo 5
 - Capítulo 11
- Mills, J.L. How to torture your data- Artigo no site

