# MODELO DE GOMPERTZ COM SAZONALIDADE E AUTOCORRELAÇÃO NOS ERROS PARA AJUSTE DO CRESCIMENTO PONDERAL EM VACA LEITEIRA[1]

LUCIO BENEDICTO KROLL[2] e MARIA TERESINHA TROVARELLI TORNERO[3]

RESUMO – Em trabalho recente sobre ajuste de peso de vacas leiteiras em função do tempo, Kroll (1990) ajustou várias funções, entre elas a de Gompertz, com erros auto-regressivos. No entanto, o ajuste sem considerar a autocorrelação explicou a mais até 25% da variação total do que o modelo de erros autocorrelacionados. Neste trabalho, o modelo de Gompertz com erros auto-regressivos é retomado com acréscimo do efeito de estação do ano. A sugestão dessa inclusão foi baseada na função de autocorrelação da série original, depois de descontada a tendência. Para estimação dos seis parâmetros, no processo iterativo foi utilizada uma matriz $\underset{\sim}{X} = [\underset{\sim}{G} \mid \underset{\sim}{S}]$, onde $\underset{\sim}{G}$ é a mesma matriz do método de Stevens para ajuste de Gompertz, e $\underset{\sim}{S}$, a matriz de 0s, 1s e -1s, para sazonalidade. Como resultado, houve elevação do coeficiente de determinação em relação ao maior valor obtido no trabalho acima citado.

Termos para indexação: regressão não linear, resíduos autocorrelacionados, função de Gompertz.

## SEASONALITY AND ERROR AUTOCORRELATION IN A GOMPERTZ MODEL ADJUSTMENT FOR DAIRY COWS WEIGHT GROWTH

ABSTRACT - In recent work (Kroll,1990) on dairy cattle growth using its monthly weights, a Gompertz function, first considering the autocorrelated error structure, and second, the independent error structure, was fitted. The last explained 25% more of the total variance than the former. In this work, a Gompertz function was fitted with autocorrelated error structure adding seasonality to the model. This was suggested taking into account the time series autocorrelation function. In order to estimate the six parameters, an iterative procedure was used, employing a matrix $\underset{\sim}{X} = [\underset{\sim}{G} \mid \underset{\sim}{S}]$, where $\underset{\sim}{G}$ is the usual matrix for fitting the Gompertz function, and $\underset{\sim}{S}$ is a 0s, 1s, -1s matrix necessary to take into account seasonality. With this model, the adjustment was better in terms of the coefficient determination.

Index terms: nonlinear regression, error autocorrelation, Gompertz function. ,

## INTRODUÇÃO

A atividade agropecuária, em nosso meio, visa a atingir dois objetivos principais: a carne e o leite. Para qualquer um deles, o crescimento ponderal do animal é um fato de relevância, que auxiliando no melhoramneto genético, serve de instrumento na avaliação de manejo animal, pode ser usado no acompanhamento do desempenho do animal no tempo, bem como pode servir para correlacioná-lo com outras variáveis de interesse, como a produção leiteira, o número de crias, etc. Justifica-se, assim, a procura de modelos que se ajustem aos dados de crescimento. A finalidade, entre outras, da busca de modelo cada vez mais adequado aos dados, reside na contribuição de maneira mais efetiva e direta nos modelos de melhoramento genético, na análise comparativa de manejo animal, através de parâmetros específicos, ou diretamente, como curva esperada do seu desempenho ponderal.

O termo "curva de crescimento" tem um significado muito especial em Estatística. Nela a análise

[1] Aceito para publicação em 30 de novembro de 1993.

[2] Eng. Agr., Dr., Prof.-Ass., Dep. de Bioestatística, Instituto Biociências, UNESP, Botucatu, Caixa Postal 502, CEP 18618-000, Rubião Júnior, Botucatu, SP.

[3] Bióloga, M.Sc., Profª.-Assist., Dep. de Bioestatística - IB, UNESP, Botucatu.



5

de curvas de crescimento aplica-se a dados longitudinais, que consistem geralmente, de medidas repetidas ao longo do tempo.

Vieira & Mischan (1976) estudaram funções assintótitico-sigmóides em modelos de crescimento. Uma delas foi a função de Gompertz, expressa no modelo:

$$Y_i = exp(\alpha + \beta\, \theta^{x_i})\, e_i,$$

onde $Y_i$ é a medida ponderal no tempo $x_i$; $e_i$ é o respectivo erro aleatório; e $\beta > 0$, $0 < \theta < 1$ e $\alpha > 0$ são parâmetros. Este modelo pode ser ajustado da equação de Spillman, como fez Stevens (1951), considerando os erros independentes.

Quando se trabalha com medidas longitudinais repetidas, fica difícil assegurar a independência dos erros em dois períodos consecutivos. Em recente trabalho, Kroll (1990) ajustou vários modelos, dentre eles o de Gompertz, ao crescimento ponderal em função do tempo para vacas leiteiras, considerando: i) modelo com estrutura de independência dos erros; ii) modelo com estrutura de erros autocorrelacionados. Verificou que o modelo Gompertz em i) explicava até 25% a mais da variação total dos dados do que ii). Este fato sugere alguma modificação no modelo ii) pois a autocorrelação nos resíduos foi significativa.

Quando se trabalha com organismo vivo em ambiente com condições não controladas, e em longo período de observação, é prudente atentar para os efeitos da sazonalidade. No manejo de vacas leiteiras, as estações do ano têm grande influência no meio ambiente, principalmente com respeito à temperatura, disponibilidade de água e de alimentos.

No presente trabalho retoma-se o ajuste do modelo de Gompertz com autocorrelação nos erros e nele se acrescenta o efeito da estação do ano.

## MATERIAL E MÉTODOS

Para estimação dos três parâmetros do modelo Gompertz, considera-se o logaritmo natural da variável dependente, e se obtém :

$$Y_i = \alpha + \beta\, \theta^{x_i} + u_i,$$

em que $u_i$ corresponde a erros independentes.

Como esta regressão é não-linear, as estimativas de mínimos quadrados para $\alpha$, $\beta$, Y e $\theta$ são obtidas por processos iterativos, no caso pelo método de Gauss-Newton. Por este, faz-se a expansão de Taylor da expressão acima, em torno de um valor fixado $\theta_o$, obtendo-se uma regressão linear múltipla de $Y_i$ contra $\theta_o^{x_i}$ e $X_i\, \theta_o^{x_i-1}$. Em termos de estimativas, pode-se escrever:

$$\hat{Y}_i = a + b r_o^{x_i} + c x_i r_o^{x_i - 1} \qquad (1)$$

onde $r_o$, a, b e c são estimativas de $\theta_o$, $\alpha$, $\beta$ e $\beta(\theta - \theta_o)$ respectivamente.

A estimavativa $r_o$ pode ser obtida por qualquer processo. Este valor é utilizado para proceder ao método de mínimos quadrados na obtenção de a, b e c.

No presente caso, $r_o$ poderia ser obtido graficamente ou pela fórmula de Patterson (Hoffmann & Vieira, 1983). Utilizou-se esta última para sete pontos, ou seja:

$$r_o = \frac{Y_7 + Y_6 + Y_5 - Y_3 - 2Y_2}{Y_6 + Y_5 + Y_4 - Y_2 - 2Y_1}.$$

Um vez realizada a primeira estimativa de a, b e c, calcula-se um novo $r_o$ ($r_o'$) através de c e b, ou seja

$$r_o' = r_o + \frac{c}{b}.$$

Repete-se o método de mínimos quadrados para regressão múltipla, agora empregando $r_o'$ no lugar de $r_o$.

A parada do processo iterativo é decidida quando $\Delta_r = \frac{c}{b}$ for considerado desprezível.

Em termos matriciais (1), pode ser escrito como segue:

$$\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{u}$$

onde,

$$\underset{\sim}{X} = \begin{bmatrix} 1 & r_o^{x_1} & x_1 & r_o^{x_1 - 1} \\ 1 & r_o^{x_2} & x_2 & r_o^{x_2 - 1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & r_o^{x_n} & x_n & r_o^{x_n - 1} \end{bmatrix}, \underset{\sim}{\hat{\beta}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \underset{\sim}{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\underset{\sim}{u} = [u_1 ; u_2 ; \ldots ; u_n],$$

Empregando-se programas computacionais, abre-se mão da tabela de Stevens (1951), que fornece os elementos de $(X'X)^{-1}$ para diversos $r_0$. Assim, o processamento de dados foi feito através de um sistema de programas computacionais que realiza operações matriciais, ALGEMA (álgebra de Matrizes) desenvolvido por Dias (1988).

A variância dos erros é estimada por:

$$S^2 = \frac{1}{n-3} \Sigma \left( Y_i - \hat{\alpha} - \hat{\beta} \hat{\theta}^{x_i} \right)^2$$

Considerando o modelo como auto-regressivo de primeira ordem, onde $\underset{\sim}{u}$ tem estrutura:

$$u_i = \rho \, u_{i-1} + \xi_i ,$$

com E $[\xi_i] = 0$; E $[\xi_i \, \xi_{i-\beta}] = 0$, se $s \neq 0$ e $-1 \leq \rho \leq 1$, para $i = 1, \ldots, n$, pode-se escrever $[\underset{\sim}{u}'\underset{\sim}{u}] = \underset{\sim}{V} \delta$.

A matriz $\underset{\sim}{V}$ é tal, que $\underset{\sim}{V}^{-1} = \underset{\sim}{\Lambda}' \underset{\sim}{\Lambda}$, onde $\underset{\sim}{\Lambda}$ segundo Theil (1971) é definido por (2) a seguir,

$$\lambda_{ij} = \begin{cases} \sqrt{1 - \hat{\rho}^2} & , \text{se } i = j \text{ e} i = 1 \\ 1 & , \text{se } i = j \text{ e} i \rangle 1 \\ -\hat{\rho} & , \text{se } i = j + 1 \text{ e } i = 2, 3, \ldots, n \\ O & , \text{nos demais casos.} \end{cases} \qquad (02)$$

Para estimar $\underset{\sim}{\Lambda}$ é nescessária a estimação de $\rho$. Para isto, ajusta-se o modelo (1) pelo método dos mínimos quadrados ordinários, e através dos desvios em tempos consecutivos obtém-se $\hat{\rho}$, (Hoffmann & Vieira, 1983), conforme estes autores, $\underset{\sim}{\beta}$ passa a ser estimado por:

$$\hat{\underset{\sim}{\beta}} = \left( \underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{X} \right)^{-1} \underset{\sim}{X}' \underset{\sim}{V}^{-1} \underset{\sim}{Y},$$

o que equivale ao método dos mínimos quadrados ordinários para Y modificado, ou seja $\underset{\sim}{Z}$, sendo que $\underset{\sim}{Z}$ é estimado por:

$$\hat{Z}_1 = \sqrt{1 - \hat{\rho}^2} \; \hat{Y}_1,$$

para o primeiro valor e

$$\hat{Z}_i = \hat{Y}_i - \rho \hat{Y}_{i-1},$$

para $l = 2, 3, \ldots, n$.

Levando em conta a sazonalidade, a matriz $\underset{\sim}{X}$ passa a ser: $\underset{\sim}{X} = [\underset{\sim}{G} \mid \underset{\sim}{S}]$, e $\underset{\sim}{S}$ é uma matriz de $0_s$, $1_s$, $-1_s$ indicando a estação do ano.

Usou-se o coeficiente de determinação corrigido ($R_c^2$) como medida para comparação dos ajustes dos modelos,

onde

$$R_c^2 = R^2 - \left( \frac{p-1}{n-p} \right) \left( 1 - R^2 \right),$$

sendo p o número de parâmetros estimados e n o número de observações.

Os dados utilizados são referentes a medidas de pesos mensais em quilogramas e em sessenta meses consecutivos, extraídos do registro da Seção Técnica da Zootecnia da ESALQ/USP, Piracicaba. O animal acima citado foi mantido em regime de criação semi-extensiva.

Nesse estudo foram consideradas as medidas ponderais da vaca leiteira identificada como Zanga, nascida em 8 de abril de 1961, oriunda do cruzamento Guernsey x Gir.

## RESULTADOS E DISCUSSÃO

A primeira iteração é feita para o modelo Gompertz, com sazonalidade, independência dos erros considerando os 60 meses, e $r_o = 0.89330486$. Ou seja a matriz $\underset{\sim}{X} = [\underset{\sim}{G} \mid \underset{\sim}{S}]$, onde: ,

$$\underset{\sim}{S} = \begin{bmatrix} -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & \ldots & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & \ldots & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}_{(3 \times 60)}$$

$$\underset{\sim}{G} = \begin{bmatrix} 1 & 0.89 & 1 \\ 1 & 0.89^2 & 2(0.89)^1 \\ \vdots & \vdots & \vdots \\ 1 & 0.89^i & i(0.89)^{i-1} \\ \vdots & \vdots & \vdots \\ 1 & 0.89^{60} & 60(0.89)^{59} \end{bmatrix}_{(60 \times 3)}$$

e o vetor $\underset{\sim}{Y}_{(1x60)}$, dos logaritmos naturais das medidas ponderais, serviram para a estimação dos desvios do modelo, e conseqüente estimativa de $\rho$, que resultou em $\hat{\rho} = 0,605$.

A cada etapa, o vetor $\underset{\sim}{Y}$ foi modificado, utilizando as estimativas dos parâmetros obtidos no processo anterior, considerando $\hat{\rho} = 0,605$ e mantendo a mesma matriz $\underset{\sim}{X}$. Desse modo, o acréscimo da matriz $\underset{\sim}{S}$ à $\underset{\sim}{G}$, ou seja, o fato de se acrescentar ao modelo o fator sazonalidade não alterou o algoritmo. Isto traz vantagem quando se usa o modelo de Gompertz.

A regra de parada do processo iterativo foi estabelecida quando $\Delta < 10^{-6}$. As estimativas dos parâmetros estão na Tabela 1. Ela mostra que, dentre os três parâmetros estimados, $\beta$ foi o mais sensível à alteração do modelo. Enquanto o modelo auto-regressivo com sazonalidade explica 96% da variação no logaritmo das medidas ponderais, o modelo auto-regressivo, sem sazonalidade, explica, apenas 71% dessa variação.

Kroll (1990) observou que o modelo Gompertz sem autocorrelação e sem sazonalidade já explicava 96% da variação. Este resultado sugere que a estrutura de autocorrelação foi perturbada pelo fator sazonalidade, quando este não está contemplado no modelo com autocorrelação; então, o modelo completo, adequado à natureza de medidas repetidas dos dados, novamente eleva o nível de explicação.

A assíntota estimada pelo modelo de Gompertz auto-regressivo sem sazonalidade, em termos de medida ponderal, foi de 333,3 kg, inferior ao modelo com sazonalidade 340,7 kg. A menor medida ponderal observada nos doze últimos meses da curva foi a 342 kg e a maior foi 405 kg. O menor e o maior valor estimado pelo modelo de Gompertz com autocorrelação e sazonalidade foram 342,2 kg e 376,1 kg, respectivamente. Já para o modelo de Gompertz com autocorrelação sem sazonalidade, essas estimativas foram 342,2 kg e 370,3 kg respectivamente. Analisando os desvios entre os valores observados e os estimados em ambos os modelos, observa-se que nos últimos doze meses apenas dois deles forneceram módulos maiores no modelo de Gompertz auto-regressivo com sazonalidade em relação ao modelo auto-regressivo sem sazonalidade.

**TABELA 1 - Estimativas dos parâmetros $\alpha$, $\beta$, $\theta$ e coeficiente de determinação ajustado ($R_c^2$), conforme ajuste do modelo de Gompertz auto-regressivo considerando ou não sazonalidade (S) para medida ponderal em 60 meses de uma vaca leiteira.**

| Parâmetro | Modelo | |
|---|---|---|
| | $\exp(\alpha + \beta\theta^i + \mu_i)$ | $\exp(\alpha + \beta\theta^i + S_i + \mu_i)$ |
| $\alpha$ | 5,809 | 5,831 |
| $\beta$ | -2,771 | -3,083 |
| $\theta$ | 0,858 | 0,872 |
| $R_c^2$ | 71% | 96% , |

As estimativas da medida ponderal no primeiro mês de observação, pelo modelo sem sazonalidade, ou com ela, foram, respectivamente, 30,9 kg e 23,5 kg, enquanto o valor observado foi de 20 kg. Portanto, a estimativa para o primeiro mês no modelo proposto com sazonalidade está mais próxima do valor observado, do que no modelo sem sazonalidade.

## CONCLUSÕES

1. Foi possível incluir componente sazonal em modelo de Gompertz auto-regressivo.

2. O modelo de Gompertz auto-regressivo com sazonalidade explicou mais em termos de variância total do que o modelo de Gompertz auto-regressivo sem sazonalidade.

3. O comportamento dos valores estimados no primeiro mês e nos últimos doze meses pelo modelo de Gompertz auto-regressivo com sazonalidade foi melhor que o do modelo de Gompertz auto-regressivo sem sazonalidade.

4. Atenção é necessária com a presença de fatores que venham perturbar a estrutura de autocorrelação quando se constroem modelos com medidas repetidas.

## REFERÊNCIAS

DIAS, C.T.S. **Algema**: sistema computacional para álgebra de matrizes. Piracicaba: ESALQ-USP, 1988. Tese de Mestrado.

HOFFMAN, R., VIEIRA, S. **Análise de regressão**: uma introdução a econometria. 2.ed. São Paulo:. Hucitec, EDUSP, 1983. 379p.

KROLL, L.B. **Estudo do crescimento de vacas leiteiras através de modelo com autocorrelação nos erros**. Botucatu: UNESP, Fac. Cien. Agron., 1990. 97p. Tese de Doutorado.

STEVENS, W.L. Assymptotic regression. **Biometrics**, n. 7, p.247-267, 1951.

THEIL, H. **Principles of econometrics**. New York: John Willey & Sons, 1971. 736p.

VIEIRA, S.; MISCHAN, M.M. A Logística e a Gompertz: duas funções alternativas no estudo de dados de crescimentos. **Ciência e Cultura**, São Paulo, v. 28, p. 950-952, 1976.