

SONIA VIEIRA

ESTUDO DE FUNÇÕES ASSINTÓTICO-SIGMÓIDES

Tese apresentada a Faculdade de  
Odontologia de Piracicaba, da Uni-  
versidade Estadual de Campinas, para  
obtenção do título de livre-docente  
em Bioestatística.

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ODONTOLOGIA DE PIRACICABA  
BIBLIOTECA

T 245

PIRACICABA  
Estado de São Paulo  
1975

## AGRADECIMENTOS

Devo meus maiores agradecimentos ao Prof. Rubens Murillo Marques, pelas valiosas sugestões apresentadas durante todo o decorrer deste trabalho.

Ao meu marido, Rodolfo Hoffmann, agradeço a revisão do texto original, bem como a ajuda na programação para computador.

Agradeço à C.D. Ana Helena Marcondes o levantamento dos dados que constam na seção 5.3. e ao Dr. Luiz Octavio Coelho Guimarães a discussão dos conceitos de odontologia ali empregados.

Sou grata ao Dr. Décio Teixeira pelo fornecimento de dados não publicados que constam na seção 5.2..

Devo à Sra. Elisa S. Peron a cuidadosa impressão deste trabalho.

Finalmente, nas pessoas do Prof. José Merzel e do Prof. Andrés José Tumang, os meus agradecimentos a todos os membros da Faculdade de Odontologia de Piracicaba, cujo apoio e incentivo foram essenciais para o bom andamento deste trabalho.

## ÍNDICE

	Pág.
1. INTRODUÇÃO .....	1
2. A FUNÇÃO LOGÍSTICA .....	4
2.1. A Função Logística para dados de crescimento .....	4
2.1.1. Características .....	4
2.1.2. Estimativas dos parâmetros .....	14
2.1.3. Variâncias das estimativas dos parâmetros .....	29
2.2. A Função Logística para dados de resposta quântica .....	32
2.2.1. Características .....	32
2.2.2. Estimativas dos Parâmetros .....	35
3. A FUNÇÃO DE GOMPERTZ .....	39
3.1. A Função de Gompertz para dados de crescimento .....	39
3.1.1. Características .....	39
3.1.2. Estimativas dos parâmetros .....	45
3.1.3. Variâncias das estimativas dos parâmetros .....	55
3.2. A Função de Gompertz para dados de resposta quântica ....	57
3.2.1. Características .....	57
3.2.2. Estimativas dos parâmetros .....	59
4. A FUNÇÃO LOG-INVERSA GENERALIZADA .....	65
4.1. A Função log-inversa generalizada para dados de cres- cimento .....	65
4.1.1. Características .....	65
4.1.2. Estimativas dos parâmetros .....	67
4.1.3. Variâncias das estimativas dos parâmetros .....	69

	Pág.
4.2. A Função log-inversa generalizada para dados de res- posta quântica .....	69
4.2.1. Características .....	69
5. EXEMPLOS DE APLICAÇÃO .....	71
5.1. O Problema da Especificação .....	71
5.2. Estudo descritivo do crescimento ponderal de animais ....	73
5.3. Estudo descritivo do número de dentes permanentes ir- rompidos, em crianças brancas, da cidade de Piracica- ba, SP.....	83
6. RESUMO E CONCLUSÕES .....	91
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	95
APÊNDICE .....	100

## 1. INTRODUÇÃO

As funções assintótico-sigmóides desempenham importante papel em estudos de crescimento, sejam eles de natureza biológica, econômica ou demográfica, onde os modelos lineares geralmente não são apropriados.

O ajustamento das funções assintótico-sigmóides aos dados de crescimento se fundamenta na hipótese de que o aspecto gráfico característico do processo é em forma de S e de que existe tendência à estabilização.

Estas condições podem ser satisfeitas por diferentes funções matemáticas. Apenas para certos casos de dados de crescimento, existe uma informação definida a respeito da relação entre as variáveis, caracterizando a forma da função, ou existe uma equação diferencial que a função deve, obrigatoriamente, satisfazer.

Tais restrições praticamente não aparecem em biologia ou, nas palavras de BERKSON (1951), "em muitos, talvez na maioria dos casos

biológicos, é realmente gratuito falar numa teoria dos fenômenos observados em qualquer sentido literal e sério. Os fatores envolvidos são tantos e o conjunto é tão variado e complexo que o modelo usado é melhor considerado como uma curva meramente empírica, empregada para a apresentação estatística sucinta dos dados, com grande utilidade descritiva, mas sem qualquer significado teórico específico".

Portanto, no campo biológico, o tipo de função assintótico-sigmóide que melhor descreve os dados de crescimento deve ser pesquisado.

Entretanto, o ajustamento destas funções é relativamente difícil. Esta afirmativa é comprovada pela quantidade de métodos propostos para o ajustamento das funções assintótico-sigmóides, que procuram diminuir a quantidade de cálculo mesmo em detrimento do rigor estatístico, e pela quantidade de trabalhos que apresentam dados experimentais, mas que mostram o simples traçado da "curva de crescimento", por entre os pontos colocados em um gráfico.

Esta dificuldade do ajustamento das funções assintótico-sigmóides constitui séria restrição à especificação do modelo da função, quando se pretende usar os próprios dados experimentais como ponto de partida.

É claro que, com a atual facilidade no uso de computadores, a quantidade de cálculo exigida pelos processos iterativos de ajustamento das funções assintótico-sigmóides não deve constituir maior problema.

Entretanto, é importante salientar que existem, no ajustamento das funções assintótico-sigmóides, diversos aspectos de ordem teórica que ainda não estão definitivamente resolvidos.

Neste trabalho propõe-se estudar três funções assintótico-sigmóides, ou seja, a função logística, a função de Gompertz e a função log-inversa generalizada. Serão considerados tanto os casos em que a ordenada da assíntota superior deve ser estimada como casos em que essa ordenada é conhecida, e a variável dependente pode ser vista como uma proporção da assíntota, que assume, então, valor igual a um.

Depois, pretende-se ajustar essas funções a dados experimentais de crescimento, obtidos no campo biológico, procurando sempre discutir a validade dos modelos em cada caso.

## 2. A FUNÇÃO LOGÍSTICA

### 2.1. A Função Logística para dados de crescimento

#### 2.1.1. Características

A função logística, definida por

$$Z = \frac{a}{1+e^{-(b+ct)}} \quad (1)$$

onde  $\underline{a}$ ,  $\underline{b}$  e  $\underline{c}$  são parâmetros,  $a > 0$  e  $c > 0$ , foi indicada para o estudo descritivo do crescimento de populações humanas por VERHULST (1845), que a denominou de "curva logística".

Muitos anos mais tarde, PEARL e REED (1920), sem conhecerem a contribuição de VERHULST (1845), obtiveram empiricamente a mesma curva, que utilizaram para descrever o crescimento da população dos EEUU, de 1790 a 1910, com base em dados de censo.

A partir daí, a curva logística tem sido bastante estudada quanto às suas características matemáticas e quanto ao método de estimar seus parâmetros. Ela também tem sido largamente empregada para a

representação de dados empíricos de crescimento de animais e vegetais, de crescimento de populações humanas e de adoção de novos bens econômicos.

A função (1) é monotonicamente crescente e fica entre as duas assíntotas  $Z = 0$  e  $Z = a$ . O parâmetro  $a$ , que é a distância entre as duas assíntotas, é denominado "nível de saturação". O parâmetro  $c$  está relacionado com a taxa de crescimento de  $Z$ , desde que

$$\frac{d(Z/a)}{dt} = c \frac{Z}{a} \left(1 - \frac{Z}{a}\right) \quad (2)$$

Finalmente,  $b$  é um parâmetro de posição, isto é, mudando o valor de  $b$  enquanto os outros parâmetros se mantêm fixos, a função apenas se movimenta horizontalmente.

A função (1) tem ponto de inflexão para a abscissa  $t = -b/c$ , onde  $Z$  vale  $a/2$ . A curva logística é radialmente simétrica em torno do seu ponto de inflexão.

O uso da função logística para o estudo descritivo de crescimento pode ser baseado na equação diferencial:

$$\frac{dZ}{dt} = k Z (a-Z) \quad (3)$$

onde  $k = c/a$  é constante.

A equação (3) mostra que a taxa de crescimento da função logística é proporcional ao valor alcançado pela função e à diferença entre esse valor e o "nível de saturação".

Portanto, reconhece-se na logística, de acordo com a denominação apresentada por LANGE (1967), um "fator de momento" igual a  $Z$  e um "fator de contenção" igual a  $(a-Z)$ .

De (3) segue-se que:

$$Z^{-1} \frac{dZ}{dt} = k (a-Z) \quad (4)$$

isto é, a taxa de crescimento relativo de  $Z$  decresce linearmente com o aumento de  $Z$ .

A função logística é imediatamente recomendável para a descrição de certos tipos de fenômenos, devido às suas características matemáticas peculiares. Entretanto, são essas mesmas características que podem tornar o modelo logístico inadequado para descrever outros tipos de fenômenos.

Assim, a equação diferencial (3) mostra que a função logística corresponde ao processo de autocatálise. De acordo com JOHNSON e KOTZ (1970), numa reação química em que o catalizador  $M$  transforma o composto  $G$  em dois compostos  $J$  e  $K$  e o próprio  $J$  atua como um catalizador para a mesma reação, se  $M_0$  e  $G_0$  são as concentrações originais de  $M$  e  $G$ , respectivamente, e  $y$  é o valor comum das concentrações de  $J$  e  $K$  no tempo  $t$ , então a lei de ação das massas é:

$$\frac{dy}{dt} = c_1 M_0 (G_0 - y) + c_2 y (G_0 - y) \quad (5)$$

onde  $c_1$  e  $c_2$  são constantes catalíticas para a ação de  $M$  e  $J$  respectivamente.

O lado direito da equação (5) pode ser rearranjado como segue:

$$c_2 \left( y + \frac{c_1}{c_2} M_o \right) \left[ \left( G_o + \frac{c_1}{c_2} M_o \right) - \left( y + \frac{c_1}{c_2} M_o \right) \right] \quad (6)$$

$$\text{Fazendo } c_2 = k, \quad y + \frac{c_1}{c_2} M_o = Z \quad \text{e}$$

$$G_o + \frac{c_1}{c_2} M_o = a, \quad \text{tem-se o segundo membro da equação (3).}$$

Por esta razão, a função logística é também conhecida como "curva de autocatálise". BERKSON (1951) apresentou, entre outros exemplos, o ajustamento da função logística para dados experimentais de hidrólise de acetato de etila para formar álcool etílico e ácido acético, pela ação catalítica do ácido acético. Neste caso, como bem lembra o autor, a informação dada pela equação diferencial (3) é totalmente válida. Isto porque a taxa de aumento do ácido, que é igual à taxa de diminuição do acetato do qual o ácido é formado, é proporcional ao produto das massas do acetato de etila e do ácido acético.

Foi também a equação (3), num problema totalmente diverso, que chamou a atenção de ROOS e VON SZELISKI (1939), que ajustaram a função logística para o crescimento da demanda de automóveis nos EE.UU.. Os autores argumentaram que, num determinado período de tempo, o aumento na população de automóveis é dado pelo número potencial de compradores, multiplicado pela probabilidade de que um indivíduo, selecionado ao acaso nesse grupo, venha a adquirir um automóvel. Os autores julgaram lícito supor que tal probabilidade é proporcional ao número de automóveis já

comprados, uma vez que este número está diretamente relacionado com o conhecimento sobre o produto, por parte de possíveis compradores e com o desenvolvimento do sistema de distribuição e de atividades complementares. Então, para determinado intervalo de tempo, tem-se:

$$\Delta C = A_1 C (M - C) \quad (7)$$

onde  $C$  é a população de automóveis,  $M$  é o número máximo de proprietários de automóveis e  $A_1$  e  $M$  dependem de fatores econômicos. Se as condições forem mantidas constantes, tem-se aqui um modelo logístico.

Num estudo sobre o aumento do número de proprietários de televisão no Reino Unido, BAIN (1962) ajustou um modelo logístico, usando argumentos semelhantes aos de ROOS e VON SZELISKI (1939). Entretanto, o próprio autor considerou que os resultados obtidos não foram satisfatórios.

Mais tarde, discutindo o aumento da demanda de novos bens de consumo, BAIN (1963) considerou que o modelo logístico é inadequado para descrever o crescimento da demanda.

De acordo com o autor, a demanda de todo novo bem de consumo depende de um processo de aprendizagem, que se passa em três estágios: primeiro, o consumidor deve se aperceber do novo bem, depois deve se familiarizar com ele e, em terceiro lugar, deve incluí-lo em seu consumo. Ora, a explicação da equação diferencial (3) se concentra nos consumidores que se encontram no segundo estágio do processo de aprendizagem.

BAIN (1963) considerou também que a função logística é simplista na explicação do processo de crescimento da demanda porque, de acordo com o modelo, o número de pessoas adquirindo o novo bem é proporcional ao número de pessoas que já tem o bem de consumo, multiplicado pelo número de pessoas que podem vir a ter esse bem. Entretanto, as pessoas vivem em grupos sociais, onde tais proporções podem ser bastante distintas.

Uma terceira característica do modelo logístico, que BAIN (1963) considera inadequada em relação ao processo de crescimento da demanda, é a de pressupor que a influência que qualquer proprietário possa exercer sobre um comprador em potencial é a mesma durante todo o processo e igual entre todos os indivíduos.

Por outro lado, os novos bens de consumo são mais rapidamente adquiridos pelas camadas da população de mais alta renda o que, evidentemente, torna assimétrica a curva de crescimento. Baseado nesta argumentação, BAIN (1963) considerou que o ajustamento de um modelo simétrico ao crescimento de demanda de novos bens de consumo é injustificável e recomendou, nestes casos, a pesquisa de curvas com assimetria positiva. Para o caso do crescimento do número de televisões na Inglaterra, BAIN (1963) utilizou um modelo baseado na curva lognormal acumulada. O autor considerou que o modelo é intuitivamente plausível e dá explicação razoável aos dados.

Em muitos casos, entretanto, o uso da função logística se baseia mais no aspecto gráfico da função do que na equação diferencial.

Assim, as curvas de crescimento de animais e vegetais tem uma forma característica em S, que está de acordo com o aspecto sigmóide do gráfico da função logística. Ainda, a função logística atende ao fato, já expresso por HAMMOND (1960), de que a taxa de crescimento em peso de todos os animais é pequena nos estágios iniciais, aumenta até um máximo e depois decresce à medida que a maturidade se aproxima.

Os processos de crescimento, desde o início até a maturidade, não podem ser descritos por polinômios, pois os valores se aproximam de uma assíntota. TOOTILL (1963) considerou que, nesses casos, os modelos contendo apenas uma não linearidade mostram-se geralmente satisfatórios. DAY (1966) salientou que famílias de curvas com três parâmetros frequentemente descrevem dados de crescimento de maneira adequada. Para a escolha da família de curvas, partiu da equação diferencial de segunda ordem:

$$f(t) \frac{d^2 f(t)}{dt^2} - (a + 1) \left( \frac{df(t)}{dt} \right)^2 + bf(t) \frac{df(t)}{dt} = 0 \quad (8)$$

onde  $a$  é uma constante para o grupo de indivíduos e  $b$  pode variar entre indivíduos. Segundo o autor, a equação acima surge naturalmente no processo de crescimento  $f(t)$ , onde um processo de envelhecimento  $g(t)$  tende a inibir o crescimento. Se  $G$  é o valor do qual  $g(t)$  se aproxima assintoticamente, então as equações diferenciais de crescimento devem ser da forma:

$$\left. \begin{aligned} \frac{df(t)}{dt} &= \left[ f(t) \right]^{a+1} \left[ G - g(t) \right]^{b_1} \\ \frac{dg(t)}{dt} &= b_2 \left[ G - g(t) \right] \end{aligned} \right\} \quad (9)$$

às quais se reduzem a equação (8), eliminando-se  $g(t)$ , com  $b = b_1 b_2$ .

A solução da equação (8), quando  $a \neq 0$  e  $b \neq 0$ , é:

$$f(t) = K \left[ 1 + ae^{-(bt+c)} \right]^{-1/a} \quad (10)$$

que é uma potência da função logística. A função (10) foi designada por NELDER (1961) como logística generalizada.

Entretanto, HOFFMANN e THAME (1970), quando ajustaram a função logística a dados de crescimento em altura de Pinus caribaea var. caribaea e MISCHAN (1972), quando ajustou a função logística a dados de crescimento ponderal de gado bovino, não obtiveram bons resultados. Os autores de ambos os trabalhos notaram que as estimativas que obtiveram para a assíntota superior eram muito inferiores aos valores mais altos alcançados pelas observações da variável dependente.

O método de ajustamento de STEVENS (1951), usado pelos autores dos dois trabalhos, pode ser parcialmente responsabilizado pelos resultados.

Entretanto, deve-se lembrar que na função logística, a ordenada do ponto de inflexão é  $a/2$ , ou seja, é a metade do valor da distância entre suas duas assíntotas. Isto caracteriza o crescimento logístico como aquele que tem taxa de crescimento crescente até o ponto em que a variável dependente atinge a metade do valor teórico máximo que pode alcançar a taxa de crescimento decrescente após esse ponto. Esta característica da função logística já foi assinalada por BAIN (1963) como altamente restritiva, o que levou OLIVER (1964) a comentar que tal

restrição sã pode ser evitada com o uso de outras funções, de propriedades diversas e eventualmente mais complicadas.

Por outro lado, o fato da função logística ser radialmente simétrica em torno do ponto de inflexão, caracteriza o crescimento logístico como aquele que tem taxas de crescimento simétrico em torno desse ponto. Ora, esta restrição também pode tornar a função logística inadequada para a descrição de certos processos de crescimento.

Assim, por exemplo, sabe-se do estudo de crescimento ponderal de gado bovino, que o peso atingido pelo animal no ponto de inflexão é menor que a metade do peso alcançado na idade adulta, o que implica em afirmar que as taxas de crescimento não são simétricas em torno do ponto de inflexão. É de se esperar, portanto, que a função logística, ajustada a pesagens mensais de uma cabeça de gado bovino, deve subestimar o valor da assíntota e superestimar o ponto de inflexão. Isto realmente ocorreu nos dados analisados por MISCHAN (1972).

No estudo descritivo de populações humanas, a função logística encontra a sua mais antiga aplicação.

Além disso, alguns autores observaram que as populações de organismos vivos, em condições específicas, crescem de acordo com a lei logística. LANGE (1967) cita o exemplo de que uma população de moscas, colocada num ambiente fechado, aumenta de acordo com a lei logística.

OLIVER (1964) verificou que a função logística se ajusta bastante bem aos dados de crescimento de populações de leveduras.

Entretanto, não se pode estender esta idéia ao crescimento de populações humanas, que depende de complicadas condições sócio-econômicas que variam no curso do desenvolvimento histórico.

O ajustamento da função logística aos dados de crescimento de populações humanas tem apenas valor descritivo, no período em análise. As previsões de populações baseadas nesses ajustamentos são falhas, como o comprova o trabalho de DAVIS (1941). Este autor ajustou a função logística aos dados de população dos EE.UU., no período de 1780-1940, e estabeleceu previsões para a população no período de 1940-2100. Segundo LANGE (1967), os desvios consideráveis entre a população efetiva dos EE.UU. e a previsão estabelecida por DAVIS (1941), são justificados por diversos fatores, entre os quais ressalta as restrições à imigração, a crise econômica de 1929-1936 e o pleno emprego depois da 2a. Guerra Mundial.

A argumentação relativa às restrições do uso da função logística com dados de populações humanas também se estende ao uso da função para os índices econômicos. É claro que, nestes casos, não existe uma teoria econômica que leve ao ajustamento da função logística, mas os resultados desse ajustamento podem ter, segundo OLIVER (1964), um valor descritivo. Este é o caso do trabalho de LEAL (1972), que ajustou a função logística aos dados de índice de custo de vida em Curitiba.

Um caso particular da função logística encontra aplicação num problema diverso dos discutidos aqui, por não se tratar de curva de crescimento. É o caso discutido por COLQUHOUN (1971), da função de Michaelis e Menten, usada em bioquímica para descrever a velocidade de

reação enzimática como função da concentração do substrato.

Então, na função logística definida em (1), indicando a variável dependente por  $v$  e fazendo  $a = V$ ,  $b = -\log K$  e  $t = \log x$ , onde  $x$  é a concentração do substrato, obtém-se:

$$v = \frac{V x^c}{K + x^c} \quad (11)$$

Para  $c = 1$ , tem-se a função de Michaelis e Menten, onde  $V$  é o valor assintótico da velocidade e  $K$  é a constante de dissociação.

A função (11), quando  $c = 1$ , também pode ser vista como um caso particular de hipérbole. Neste caso, o ajustamento é mais fácil e deve ser feito na forma indicada por BLISS e JAMES (1966).

#### 2.1.2. Estimativas dos parâmetros

É possível que a dificuldade de estimação dos parâmetros da função logística, definida por

$$Z = \frac{a}{1 + e^{-(b+ct)}} \quad (1)$$

pelo método clássico dos mínimos quadrados, tenha reduzido o seu emprego.

Entretanto, não se pode esquecer que foram propostos diversos outros métodos para a obtenção das estimativas dos parâmetros da função (1).

Entretanto, os métodos que tem apenas o mérito de simplificar os cálculos, como o dos "totais parciais", ou o proposto por DAVIS

(1941), ou o proposto por ERKELENS (1968), perdem em rigor estatístico. Os métodos que usam uma transformada da variável dependente, como os propostos por STEVENS (1951), PATTERSON (1956) e NELDER (1961) pressupõem que os erros de  $Z$  são heterocedásticos e portanto, só devem ser aplicados nesses casos. Os métodos gráficos, como o método da estimação dos parâmetros da função logística por nomograma, proposto por RASOR (1949), também perdem em rigor estatístico.

Parece lícito supor que diferentes métodos de ajustamento, aplicados aos mesmos dados experimentais, devem levar a diferenças sensíveis, tanto nas estimativas dos parâmetros como nos quadrados médios residuais. Este fato é comprovado pelos trabalhos de OLIVER (1964) e de LEAL (1972), que comparam diversos métodos para o ajustamento da função logística.

Um dos métodos mais conhecidos para o ajustamento da função logística é o método dos "totais parciais", que consta em diversos livros textos, principalmente nos mais antigos, como o de CROXTON e COWDEN (1955). Tal método consiste em dividir as observações em três grupos e calcular os subtotais de cada grupo. Se o número de dados não for divisível por 3, abandonam-se 1 ou 2 dados. Obtêm-se, assim, três equações em  $a$ ,  $b$  e  $c$ , que são resolvidas algebricamente. As estimativas dos parâmetros obtidas por este método são consistentes, mas não são eficientes, conforme demonstrou STEVENS (1951).

Outro método para o ajustamento da função logística é o desenvolvido por DAVIS (1941), que estimou os parâmetros da função baseada num procedimento anteriormente proposto por Hotelling. De (1), segue-se

que:

$$\frac{1}{Z} \frac{dZ}{dt} = c - \left(\frac{c}{a}\right) Z \quad (2)$$

Aproximando  $\frac{dZ}{dt}$  por  $\frac{\Delta Z}{\Delta t}$ , pode-se ajustar uma regressão linear simples de  $\frac{\Delta Z}{Z \Delta t}$  contra Z. Obtêm-se, assim, as estimativas dos parâmetros  $\underline{a}$  e  $\underline{c}$ . De (1) segue-se também que

$$e^{-b} = \frac{a-Z}{Z} e^{ct} \quad (3)$$

A estimativa de  $\underline{b}$  é dada pelo valor negativo do logaritmo da média aritmética dos valores obtidos de (3), para cada par de valores t, Z. Se alguns valores de Z excederem a estimativa da assíntota, o que comumente acontece, o autor recomenda que eles sejam excluídos, para o cálculo de  $\underline{b}$ .

É claro que, se o número de pares de valores observados é grande, a estimação de  $\underline{b}$  dependerá de bastante cálculo. TINTNER (1952) sugeriu que se determinasse  $\underline{b}$  apenas com os valores médios de t e Z, isto é, a função passaria obrigatoriamente pelo ponto  $(\bar{t}, \bar{Z})$ . A sugestão de TINTNER (1952) tem apenas o mérito de simplificar os cálculos, uma vez que, obviamente, os resultados só serão viáveis se as observações se distribuírem igualmente em torno do ponto de inflexão.

Também ERKELENS (1968) desenvolveu um método para o ajustamento da função logística, usando a propriedade de que o logaritmo das diferenças dos inversos de valores consecutivos de Z são função linear de t. Se  $t_{i+1} = 1+t_i$ , de (1) pode-se obter:

$$\log \left( \frac{1}{Z_i} - \frac{1}{Z_{i+1}} \right) = -ct_i + \log \left( e^{-b} \frac{1-e^{-c}}{a} \right) \quad (4)$$

que é uma regressão linear simples de  $\log \left( \frac{1}{Z_i} - \frac{1}{Z_{i+1}} \right)$  contra  $t_i$ . A partir de (4), obtêm-se as estimativas de  $c$  e de  $e^{-b}/a$ . A estimativa de  $a$ , e conseqüentemente, a estimativa de  $b$ , pode ser obtida através de uma transformação da própria função logística, pois de (1) vem:

$$\frac{1}{a} = \frac{1}{Z} - \frac{e^{-b}}{a} e^{-ct} \quad (5)$$

Usando-se as médias aritméticas de  $1/Z$  e de  $e^{-ct}$  em (5), facilmente se obtém a estimativa de  $a$ .

Note-se que o método de ajustamento da função logística, proposto por ERKELENS (1968), exige que os intervalos entre valores consecutivos de  $t$  sejam iguais, de forma que cada intervalo possa ser visto como uma unidade de  $t$ .

As estimativas dos parâmetros da função logística, quando se pressupõe que os erros são homocedásticos, podem ser obtidas pelo método dos mínimos quadrados que tem a vantagem de não exigir pressuposição a respeito da forma específica da distribuição dos erros. Então, dada a função (1), com erro aditivo de média zero e variância  $\sigma^2$ , obtém-se, imediatamente, o sistema de equações normais:

$$\left. \begin{aligned} \Sigma \frac{1}{1+e^{-(\hat{b}+\hat{c}t)}} \left( z - \frac{\hat{a}}{1+e^{-(\hat{b}+\hat{c}t)}} \right) &= 0 \\ \hat{a} \Sigma \frac{e^{-(\hat{b}+\hat{c}t)}}{[1+e^{-(\hat{b}+\hat{c}t)}]^2} \left( z - \frac{\hat{a}}{1+e^{-(\hat{b}+\hat{c}t)}} \right) &= 0 \\ \hat{a} \Sigma \frac{te^{-(\hat{b}+\hat{c}t)}}{[1+e^{-(\hat{b}+\hat{c}t)}]^2} \left( z - \frac{\hat{a}}{1+e^{-(\hat{b}+\hat{c}t)}} \right) &= 0 \end{aligned} \right\} \quad (6)$$

onde  $z$  são as observações e  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$  são as estimativas de  $a$ ,  $b$  e  $c$ , respectivamente.

A solução do sistema de equações normais (6) deve ser obtida por processo iterativo.

Partindo de estimativas preliminares  $\hat{a}_0$ ,  $\hat{b}_0$  e  $\hat{c}_0$  de  $a$ ,  $b$ , e  $c$ , respectivamente, define-se  $\Delta a$ ,  $\Delta b$  e  $\Delta c$  como correções, obtidas em determinada etapa, para as estimativas dos mesmos parâmetros obtidas na etapa anterior. Aplicando a série de Taylor a cada uma das equações normais e desprezando os termos onde  $\Delta a$ ,  $\Delta b$  e  $\Delta c$  aparecem com expoente superior a um, obtém-se:

$$\left[ F + Q \right] \begin{bmatrix} \Delta a \\ \Delta b \\ \Delta c \end{bmatrix} = \begin{bmatrix} \Sigma (z - \hat{z}_0) \phi_0 \\ \hat{a}_0 \Sigma (z - \hat{z}_0) e^{-(\hat{b}_0 + \hat{c}_0 t)} \phi_0^2 \\ \hat{a}_0 \Sigma (z - \hat{z}_0) e^{-(\hat{b}_0 + \hat{c}_0 t)} t \phi_0^2 \end{bmatrix} \quad (7)$$

$$\text{onde } \phi_0 = \frac{1}{1+e^{-(\hat{b}_0+\hat{c}_0 t)}}, \quad \hat{z}_0 = \hat{a}_0 \phi_0, \quad e \quad (8)$$

F e Q são matrizes simétricas cujos elementos são dados por:

$$\left. \begin{aligned} F_{11} &= \Sigma \phi_0^2 \\ F_{12} &= \hat{a}_0 \Sigma e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^3 \\ F_{13} &= \hat{a}_0 \Sigma e^{-(\hat{b}_0+\hat{c}_0 t)} t \phi_0^3 \\ F_{22} &= \hat{a}_0^2 \Sigma e^{-2(\hat{b}_0+\hat{c}_0 t)} \phi_0^4 \\ F_{23} &= \hat{a}_0^2 \Sigma e^{-2(\hat{b}_0+\hat{c}_0 t)} t \phi_0^4 \\ F_{33} &= \hat{a}_0^2 \Sigma e^{-2(\hat{b}_0+\hat{c}_0 t)} t^2 \phi_0^4 \end{aligned} \right\} \quad (9)$$

$$\left. \begin{aligned} Q_{11} &= 0 \\ Q_{12} &= - \Sigma (z-\hat{z}_0) e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^2 \\ Q_{13} &= - \Sigma (z-\hat{z}_0) e^{-(\hat{b}_0+\hat{c}_0 t)} t \phi_0^2 \\ Q_{22} &= -\hat{a}_0 \Sigma (z-\hat{z}_0) e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^2 \left[ 2e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^{-1} \right] \\ Q_{23} &= -\hat{a}_0 \Sigma (z-\hat{z}_0) e^{-(\hat{b}_0+\hat{c}_0 t)} t \phi_0^2 \left[ 2e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^{-1} \right] \\ Q_{33} &= -\hat{a}_0 \Sigma (z-\hat{z}_0) e^{-(\hat{b}_0+\hat{c}_0 t)} t^2 \phi_0^2 \left[ 2e^{-(\hat{b}_0+\hat{c}_0 t)} \phi_0^{-1} \right] \end{aligned} \right\} \quad (10)$$

É importante notar que com exceção do elemento nulo, todos os elementos da matriz Q, definida em (10), são somas ponderadas de desvios. Portanto, deve-se esperar que os valores desses elementos

sejam pequenos em comparação com os valores dos elementos da matriz  $F$  definida em (9), desde que a variância residual não seja muito grande. Então, desprezando a matriz  $Q$ , o sistema (7) fica como segue:

$$F \begin{bmatrix} \Delta a \\ \Delta b \\ \Delta c \end{bmatrix} = \begin{bmatrix} \Sigma (z - \hat{z}_0) \phi_0 \\ \hat{a}_0 \Sigma (z - \hat{z}_0) e^{-(\hat{b}_0 + \hat{c}_0 t)} \phi_0^2 \\ \hat{a}_0 \Sigma (z - \hat{z}_0) e^{-(\hat{b}_0 + \hat{c}_0 t)} t \phi_0^2 \end{bmatrix} \quad (11)$$

O sistema de equações (11) pode ser obtido expandindo a estimativa da função pela série de Taylor, desprezando os termos de grau superior a um. Na realidade, minimiza-se a função:

$$S = \begin{pmatrix} z - f(\hat{a}_0, \hat{b}_0, \hat{c}_0) - f'_a(\hat{a}_0, \hat{b}_0, \hat{c}_0) \Delta a \\ -f'_b(\hat{a}_0, \hat{b}_0, \hat{c}_0) b - f'_c(\hat{a}_0, \hat{b}_0, \hat{c}_0) \Delta c \end{pmatrix}^2$$

onde  $f(\hat{a}_0, \hat{b}_0, \hat{c}_0)$  é o valor da função para as estimativas  $\hat{a}_0$ ,  $\hat{b}_0$  e  $\hat{c}_0$  e  $f'_a(\hat{a}_0, \hat{b}_0, \hat{c}_0)$ ,  $f'_b(\hat{a}_0, \hat{b}_0, \hat{c}_0)$  e  $f'_c(\hat{a}_0, \hat{b}_0, \hat{c}_0)$  são, respectivamente, os valores das derivadas parciais da função em relação a  $\underline{a}$ ,  $\underline{b}$  e  $\underline{c}$  no ponto  $(\hat{a}_0, \hat{b}_0, \hat{c}_0)$ .

Note-se que o processo para a solução do sistema de equações normais, dadas em (6), leva ao sistema (7). Uma aproximação razoável de (7) é dada pelo sistema definido em (11), que tem a vantagem de ser mais simples. Repetidos um número suficiente de vezes, tanto o sistema (7) como o sistema (11) devem levar à mesma solução, que independe

das estimativas preliminares. Entretanto a velocidade de convergência para a solução final, obtida pelo sistema (7), pode ser substancialmente maior que a velocidade obtida pelo sistema (11), tornando o primeiro preferível a este.

As estimativas preliminares  $\hat{a}_0$ ,  $\hat{b}_0$  e  $\hat{c}_0$ , necessárias para se iniciar o processo iterativo, podem ser obtidas por qualquer um dos métodos simplificados anteriormente descritos. Neste trabalho será usado o método proposto por DAVIS (1941).

SCHULTZ (1930) foi o primeiro autor que usou o método dos mínimos quadrados para a obtenção das estimativas dos parâmetros da função logística com erro aditivo. Entretanto, provavelmente devido à quantidade de cálculo que o método exige, SCHULTZ (1930) trabalhou com apenas uma iteração. Neste caso, o resultado final deve ser bastante impreciso e totalmente dependente das estimativas preliminares.

OLIVER (1964) ajustou a função logística pelo método dos mínimos quadrados, expandindo a função pela série de Taylor e usando número suficiente de iterações. Concluiu que não existe substituto para este método, uma vez que todos os métodos que simplificam os cálculos levam a estimativas ineficientes dos parâmetros. Entretanto, ressaltou o autor, o método dos mínimos quadrados exige o uso de computador, uma vez que o número de iterações necessárias pode ser relativamente grande.

Em trabalho posterior, OLIVER (1966) obteve as estimativas dos parâmetros da função logística através do método da máxima verossimilhança e estudou as variâncias dessas estimativas através da matriz

de informação. O autor observou que as estimativas obtidas por este método, que são assintoticamente eficientes, são idênticas às obtidas pelo método dos mínimos quadrados.

Os resultados obtidos pela aplicação direta do método dos mínimos quadrados não serão satisfatórios se for válida a pressuposição de que os erros são heterocedásticos.

Nestes casos, o ajustamento da função logística deve ser feito ou pelo método proposto por STEVENS (1951), desde que seja possível admitir que o inverso da variável dependente tem variância constante, ou pelo método proposto por NELDER (1961), desde que seja possível admitir que o logaritmo da variável dependente tem variância constante.

STEVENS (1951) ajustou a função de Spillman definida por:

$$Y = \alpha + \beta \rho^t \quad (13)$$

onde  $\alpha$ ,  $\beta$  e  $\rho$  são parâmetros.

Aplicando o método dos mínimos quadrados, o autor obteve o sistema de equações normais:

$$\left. \begin{aligned} \Sigma (y - \hat{Y}) &= 0 \\ \Sigma (y - \hat{Y}) R^t &= 0 \\ \Sigma (y - \hat{Y}) Bt R^{t-1} &= 0 \end{aligned} \right\} \quad (14)$$

onde  $y$  são as observações,  $\hat{Y}$  são as estimativas dadas pela função e  $A$ ,  $B$  e  $R$  são as estimativas de  $\alpha$ ,  $\beta$  e  $\rho$  respectivamente.

Como o sistema de equações normais (14) não é linear, sua solução é obtida por processo iterativo. Sejam  $\Delta A$ ,  $\Delta B$  e  $\Delta R$  as correções das estimativas de  $\alpha$ ,  $\beta$  e  $\rho$  e sejam  $A_0$ ,  $B_0$  e  $R_0$  as estimativas preliminares desses parâmetros.

Aplicando a série de Taylor a cada uma das equações normais e desprezando os termos onde  $\Delta A$ ,  $\Delta B$  e  $\Delta R$  aparecem com expoente superior a um, obtêm-se:

$$\begin{bmatrix} F + Q \end{bmatrix} \begin{bmatrix} \Delta A \\ \Delta B \\ \Delta R \end{bmatrix} = \begin{bmatrix} \Sigma (y - \hat{Y}_0) \\ \Sigma (y - \hat{Y}_0) R_0^t \\ B_0 \Sigma (y - \hat{Y}_0) t R_0^{t-1} \end{bmatrix} \quad (15)$$

onde  $\hat{Y}_0$  é o valor da função para as estimativas  $A_0$ ,  $B_0$  e  $R_0$ ,

$$F = \begin{bmatrix} n & \Sigma R_0^t & B_0 \Sigma t R_0^{t-1} \\ \Sigma R_0^t & \Sigma R_0^{2t} & B_0 \Sigma t R_0^{2t-1} \\ B_0 \Sigma t R_0^{t-1} & B_0 \Sigma t R_0^{2t-1} & B_0^2 \Sigma t^2 R_0^{2t-2} \end{bmatrix} \quad (16)$$

e

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & - \Sigma (y - \hat{Y}_0) t R_0^{t-1} \\ 0 & - \Sigma (y - \hat{Y}_0) t R_0^{t-1} & - B_0 \Sigma (y - \hat{Y}_0) t (t-1) R_0^{t-2} \end{bmatrix} \quad (17)$$

sendo  $n$  o número de pares de observações.

Note-se que os elementos da matriz Q, definida em (17), ou são nulos ou são somas ponderadas de desvios. Desde que a variância residual não seja muito grande, os valores dos elementos desta matriz são pequenos em comparação com os valores dos elementos da matriz F, definida em (16). Então, desprezando a matriz (17), o sistema de equações definido em (15) fica como segue:

$$F \begin{bmatrix} \Delta A \\ \Delta B \\ \Delta R \end{bmatrix} = \begin{bmatrix} \Sigma (y - \hat{Y}_0) \\ \Sigma (y - \hat{Y}_0) R_0^t \\ B_0 \Sigma (y - \hat{Y}_0) t R_0^{t-1} \end{bmatrix}$$

Lembrando que  $\Delta A$ ,  $\Delta B$  e  $\Delta R$  são correções, dadas pela diferença entre a estimativa dos parâmetros obtidas em duas etapas consecutivas, da relação anterior se obtêm:

$$F \begin{bmatrix} A \\ B \\ \Delta R \end{bmatrix} - F \begin{bmatrix} A_0 \\ B_0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma y R_0^t \\ B_0 \Sigma y t R_0^{t-1} \end{bmatrix} - \begin{bmatrix} \Sigma \hat{Y}_0 \\ \Sigma \hat{Y}_0 R_0^t \\ B_0 \Sigma \hat{Y}_0 t R_0^{t-1} \end{bmatrix}$$

Partindo de  $\hat{Y}_0 = A_0 + B_0 R_0^t$ , pode-se verificar que:

$$F \begin{bmatrix} A_0 \\ B_0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Sigma \hat{Y}_0 \\ \Sigma \hat{Y}_0 R_0^t \\ B_0 \Sigma \hat{Y}_0 t R_0^{t-1} \end{bmatrix}$$

Considerando esta igualdade, o sistema anterior pode ser escrito como segue:

$$F \begin{bmatrix} A \\ B \\ \Delta R \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma y R_0^t \\ B_0 \Sigma y t R_0^{t-1} \end{bmatrix}$$

ou

$$\begin{bmatrix} n & \Sigma R_0^t & \Sigma t R_0^{t-1} \\ \Sigma R_0^t & \Sigma R_0^{2t} & \Sigma t R_0^{2t-1} \\ \Sigma t R_0^{t-1} & \Sigma t R_0^{2t-1} & \Sigma t^2 R_0^{2t-2} \end{bmatrix} \begin{bmatrix} A \\ B \\ B_0 \Delta R \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma y R_0^t \\ \Sigma y t R_0^{t-1} \end{bmatrix} \quad (18)$$

A expressão (18), obtida por STEVENS (1951), permite determinar as estimativas de  $\alpha$  e de  $\beta$  e uma correção para  $R_0$ , desde que se ja dada a estimativa preliminar de  $\rho$ .

Inicia-se assim um processo iterativo de cálculo que deve chegar às estimativas finais A, B e R dos parâmetros  $\alpha$ ,  $\beta$  e  $\rho$ , que STEVENS (1951) mostrou serem eficientes e consistentes. É importante notar que usando a relação (15), é possível obter convergência mais rápida para a solução final, do que usando a relação (18). Entretanto, a relação (18) é mais simples e exige a estimativa preliminar de um único parâmetro.

O método de ajustamento que STEVENS (1951) desenvolveu se aplica às diversas funções que, por anamorfose, se transformam na função (13). Assim, invertendo os membros da função (1), obtêm-se:

$$\frac{1}{Z} = \frac{1}{a} + \frac{e^{-(b+ct)}}{a} \quad (19)$$

Em (19) fazendo  $\frac{1}{Z} = Y$ ;  $\frac{1}{a} = \alpha$ ;  $\frac{e^{-b}}{a} = \beta$ ;  $e^{-c} = \rho$ , obtêm-se a função (13).

Note-se que quando se ajusta a função logística na forma da função (19), minimiza-se a soma dos quadrados dos desvios dos inversos dos valores observados em relação aos valores estimados por essa função. Portanto, o modelo estatístico da função logística em análise é:

$$Z = \frac{a}{1 + e^{-(b+ct)} + a \varepsilon} \quad (20)$$

onde  $\varepsilon$  são variáveis aleatórias independentes, de média zero e variância constante.

NELDER (1961) estudou a família de curvas definida pela equação diferencial

$$\frac{dZ}{dt} = cZ \left[ 1 - \left( \frac{Z}{a} \right)^{1/\theta} \right] \quad (21)$$

onde  $c$ ,  $a$  e  $\theta$  são parâmetros. Fazendo  $\theta > 0$ , o autor obteve a solução da equação diferencial (21):

$$Z = a / \left[ 1 + e^{-(b+ct)/\theta} \right]^\theta \quad (22)$$

onde  $b$  é a constante de integração. NELDER (1961) denominou a função (22) de logística generalizada. Em (22), fazendo  $\theta = 1$ , obtém-se a função logística definida em (1).

Aplicando logaritmo à função (22), o autor obteve:

$$Y = \alpha - \theta \log \left[ 1 + e^{-(b+ct)/\theta} \right] \quad (23)$$

onde  $Y = \log Z$  e  $\alpha = \log a$

O ajustamento da função (23) através do método dos mínimos quadrados resultou em um sistema de equações normais que não tem solução explícita.

Para a solução desse sistema quando  $\theta = 1$ , NELDER (1961) propôs que fossem obtidas, a partir de uma estimativa preliminar de  $\alpha$ , estimativas preliminares para b e para c, fazendo uma regressão linear simples de  $\log (Z/a-Z)$  contra t, pois de (1) pode-se obter:

$$\log \left( \frac{Z}{a-Z} \right) = b + ct \quad (24)$$

Depois, por correções sucessivas, são obtidas as estimativas finais. Para a obtenção da estimativa preliminar  $\alpha_0$ , NELDER (1961) sugere um processo semigráfico que nem sempre é necessário.

Note-se que quando se ajusta a função logística pelo método proposto por NELDER (1961), o modelo estatístico da função em análise é:

$$z = \frac{a}{1+e^{-(b+ct)}} + \varepsilon, \quad (25)$$

sendo que  $\log \varepsilon$  são variáveis aleatórias com média zero e variância constante.

Portanto, minimiza-se a soma dos quadrados dos logaritmos dos valores observados em relação aos valores estimados pela função (22), com  $\theta$  igual a 1.

NELDER (1961), comparando o método de ajustamento da função logística que desenvolveu com o método desenvolvido por STEVENS

(1951), mostrou que ambos implicam em admitir que as variâncias das observações são funções dos valores assumidos por essas observações, devido às anamorfoses utilizadas.

Assim, feita a anamorfose  $y = 1/z$ , exigida pelo processo de STEVENS (1951), admite-se que o inverso da variável dependente tem variância constante  $\sigma^2$ . Uma aproximação da variância de  $z$  pode ser obtida a partir da equação diferencial

$$dy = \frac{1}{z^2} dz \quad (26)$$

donde

$$V(z) = z^4 \sigma^2 \quad (27)$$

Portanto, quando se ajusta a função logística pelo método de STEVENS (1951), admite-se que a variância das observações da variável dependente cresce com a quarta potência dessa variável, o que significa que, na estimação dos parâmetros, será dado maior peso às observações de menor valor (ou seja, próximas do eixo das abcissas), uma vez que as observações próximas da assíntota tem variância relativamente alta.

Quando os dados observados não atendem à pressuposição de heterocedasticia de acordo com (27), o método de ajustamento proposto por STEVENS(1951) não pode levar a uma boa estimativa da assíntota. HOFFMANN e THAME (1970), MISCHAN (1972) e LEAL (1972) observaram que obtiveram estimativas das assíntotas muito inferiores aos valores mais altos alcançados pela variável dependente.

Da mesma forma, feita a anamorfose  $y = \log z$ , exigida pelo método de NELDER (1961), admite-se que a variância do logaritmo da variável dependente é constante, pressuposição comum quando se trata de dados econômicos.

Obtêm-se uma aproximação da variância de  $z$  partindo da equação diferencial

$$dy = \frac{1}{z} dz \quad (28)$$

donde

$$V(z) = z^2 \sigma^2 \quad (29)$$

Se a pressuposição de heterocedasticia não for válida para os dados em análise, o método proposto por NELDER (1961) pode levar a uma subestimativa da assíntota.

### 2.1.3. Variâncias das estimativas dos parâmetros

As estimativas dos parâmetros da função logística, obtidas pelo método dos mínimos quadrados, são idênticas as estimativas obtidas pelo método de máxima verossimilhança, quer se use o método de ajustamento proposto por STEVENS (1951) (STEVENS (1951)), ou o método proposto por NELDER (1961) (NELDER (1961)), ou o método que pressupõe um erro aditivo no modelo (OLIVER (1966)).

Ora, as variâncias das estimativas dos parâmetros obtidas pelo método de máxima verossimilhança, são dadas assintoticamente pelo limite inferior da desigualdade de Cramer-Rao, ou seja, são dadas pelo

inverso da matriz de informação.

Verifica-se que a matriz de informação  $M^{-1}$ , relativa às variâncias de  $s^{*2}$  e de  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$ , é dada por:

$$M^{-1} = \frac{1}{\sigma^2} \begin{bmatrix} \frac{n}{2\sigma^2} & \phi' \\ \phi & F \end{bmatrix} \quad (1)$$

onde  $\phi$  é um vetor de zeros.

Para a obtenção das variâncias das estimativas dos parâmetros pelo método desenvolvido por STEVENS (1951), usa-se a matriz  $F$  definida em (16) na seção 2.1.2.

No método proposto por NELDER (1961), a matriz  $F$  é dada por (NELDER (1961)):

$$F = \begin{bmatrix} n & \Sigma \phi & \Sigma \phi t \\ \Sigma \phi & \Sigma \phi^2 & \Sigma \phi^2 t \\ \Sigma \phi t & \Sigma \phi^2 t & \Sigma \phi^2 t^2 \end{bmatrix} \quad (2)$$

$$\text{onde } \phi = \frac{e^{-(b+ct)}}{1+e^{-(b+ct)}} = \frac{1}{1+e^{(b+ct)}}$$

Para a obtenção das variâncias das estimativas dos parâmetros pelo método que pressupõe um erro aditivo no modelo usa-se a matriz  $F$  definida em (9) na seção 2.1.2.

A estimativa de  $\sigma^2$ , necessária para a obtenção das estimativas das variâncias das estimativas dos parâmetros, é obtida por:

$$s^2 = \text{SQR}/(n-3) \quad (3)$$

onde SQR é a soma de quadrados de resíduo de regressão e n é o número de pares de observações.

Note-se que pode ser obtida a estimativa de máxima verossimilhança de  $\sigma^2$ , dada por:

$$s^{*2} = \text{SQR}/n \quad (4)$$

Entretanto, optou-se pela estimativa dada em (3), por similaridade com os modelos lineares, onde se obtém a estimativa de  $\sigma^2$ , dividindo-se a soma dos quadrados de resíduo de regressão pelos respectivos graus de liberdade. Este divisor assegura, no caso dos modelos lineares, a propriedade de não tendenciosidade.

Cumpra notar que STEVENS (1951), e DRAPER e SMITH (1966) indicam que a estimativa de  $\sigma^2$  seja obtida por (3), também no caso dos modelos de regressão assintótica.

Finalmente deve-se observar que para maior exatidão, a soma de quadrados de resíduo de regressão, usada para a obtenção de  $s^2$ , deve ser calculada da maneira sugerida por STEVENS (1951), isto é, fazendo a soma dos quadrados das diferenças entre cada valor e a respectiva estimativa, que não é o procedimento de cálculo comumente utilizado no caso de modelos lineares.

## 2.2. A Função Logística para dados de resposta quântica

### 2.2.1. Características

Um problema mais simples, mas relacionado com o ajustamento da função logística, é o estudo da função

$$P = \left[ 1 + e^{-(\alpha + \beta x)} \right]^{-1} \quad (1)$$

que pode ser vista como proveniente da função logística definida por

$$Z = \frac{Y}{1 + e^{-(\alpha + \beta x)}} \quad (2)$$

quando o nível de saturação  $\gamma$  é conhecido. Portanto,  $P$  expressa a proporção do máximo valor possível de  $z$ .

A função (1) encontra importante campo de aplicação em dados de resposta quântica\*, largamente estudados por BERKSON (1944, 1951, 1953, 1955).

Na conceituação de REIERSOL (1961), tem-se dados de resposta quântica quando se considera uma sequência de experimentos binomiais independentes, divididos em  $s$  grupos, e se relaciona ao  $j$ ésimo grupo, um número  $n_j$  de ensaios e uma variável não aleatória  $x_j$ .

---

\* O "Dicionário Brasileiro de Estatística" de Milton da Silva Rodrigues, registra reação quântica, em vez de resposta quântica.

Nos exemplos clássicos de resposta quântica,  $x_j$  se refere à dose de uma droga ou ao estímulo dado a um animal experimental e se observa se houve ou não morte desse animal. Entretanto, existem problemas totalmente diversos que se enquadram nesta conceituação. É o caso do exemplo citado por AITCHISON e BENNET (1970), onde o estímulo é a renda pessoal e a resposta é a compra ou não de um automóvel.

Segundo BERKSON (1944), o primeiro autor que estudou dados de resposta quântica foi BLISS (1935). Partindo do pressuposto de que dose letal mede susceptibilidade e de que susceptibilidade tem distribuição normal de média  $\mu$  e variância  $\sigma^2$ , BLISS (1935) argumentou que a proporção de indivíduos mortos, quando expostos à dose  $x_j$ , deve ser dada pela integral, até  $Y$ , da distribuição normal de probabilidades.

Então, define-se próbite como segue:

$$p(x_j) = \int_{-\infty}^Y f(y) dy \quad (3)$$

onde  $p(x_j)$  é a proporção de indivíduos mortos na dose  $x_j$ ,  $f(y)$  é a função normal reduzida, e o próbite é dado por  $Y + 5$ , onde 5 é um valor arbitrário usado apenas para evitar o sinal negativo.

Esta definição consta em diversos livros textos, mas merece especial destaque o livro de FINNEY (1952).

De (3) segue-se também a definição de nórmita (BERKSON (1944)):

$$\text{N\u00f3rmite de } \hat{P} = Y = \frac{(x-\mu)}{\sigma} = \alpha + \beta x \quad (4)$$

com  $\alpha = -\mu/\sigma$  e  $\beta = 1/\sigma$ .

Mais recentemente, foram estudados os m\u00e9todos de estimac\u00e3o dos n\u00f3rmites por BERKSON (1957), CRAMER (1964), LITTLE (1968) e SOWDEN (1972).

BERKSON (1951) argumentou, com muita propriedade, que o pr\u00f3bite mede susceptibilidade ou toler\u00e2ncia mas que, em resposta qu\u00e2ntica, n\u00e3o se observa nem se mede susceptibilidade ou toler\u00e2ncia. Tem-se, apenas, a propor\u00e7\u00e3o  $p_j$  de indiv\u00edduos mortos quando se administra a dose  $x_j$ . Ent\u00e3o, quando se calcula o desvio padr\u00e3o para toler\u00e2ncia, calcula-se uma medida de variabilidade para algo que n\u00e3o foi observado.

Baseado nesta argumenta\u00e7\u00e3o, BERKSON (1944) prop\u00f4s que se ajustasse aos dados de resposta qu\u00e2ntica a fun\u00e7\u00e3o:

$$Y = \log \left( \frac{P}{1-P} \right) = \alpha + \beta x \quad (5)$$

obtida da fun\u00e7\u00e3o (1).

A vari\u00e1vel transformada  $Y = \log \left( \frac{P}{1-P} \right)$  foi denominada l\u00f3gite, por analogia com a palavra pr\u00f3bite.

Diversos autores verificaram que os resultados das an\u00e1lises de pr\u00f3bites e l\u00f3gites, feitas sobre os mesmos dados experimentais, s\u00e3o bastante semelhantes. FINNEY (1952) observou que tais resultados s\u00e3o particularmente concordantes com rela\u00e7\u00e3o \u00e0s doses letais medianas.

A semelhança de resultados se deve às formas similares das distribuições normal e logística, conforme consta em JOHNSON e KOTZ (1970).

Entretanto, a função (5) é bastante mais simples que a integral da curva normal.

Além de ser usada para a descrição de dados de resposta quântica, a função (1) também pode ser usada para descrever a difusão de determinada característica na população. Neste caso, conforme observa OLIVER (1964), a implicação é a de que, quando a variável independente, que aqui é o tempo, tende para infinito, todos os indivíduos da população apresentarão a característica em estudo.

Entretanto, deve-se ter presente a idéia de que quando se ajusta a função (1) aos dados de proporções que crescem com o aumento da dose de uma droga, ou aos dados de proporções que crescem ao longo do tempo, admite-se implicitamente as características do crescimento logístico. Tais características, podem não ser válidas para todos os tipos de fenômenos.

### 2.2.2. Estimativas dos Parâmetros

O ajustamento da função logística aos dados de resposta quântica pode ser feito pelo método da máxima verossimilhança ou pelo método dos mínimos quadrados.

A função é:

$$P = \left[ 1 + e^{-(\alpha + \beta x)} \right]^{-1} \quad (1)$$

Aplicando o método da máxima verossimilhança, BERKSON (1944) obteve o sistema de equações:

$$\left. \begin{aligned} \sum_{i=1}^k n_i (p_i - \hat{P}_i) &= 0 \\ \sum_{i=1}^k n_i x_i (p_i - \hat{P}_i) &= 0 \end{aligned} \right\} \quad (2)$$

onde  $n_i$  é o número de ensaios relativos a cada dose  $x_i$ ,  $p_i$  é a frequência relativa e  $\hat{P}_i$  é a estimativa de  $P_i$ .

A solução do sistema de equações (2) é obtida por processo iterativo.

Para a obtenção das estimativas de  $\alpha$  e de  $\beta$ , pelo método dos mínimos quadrados, BERKSON (1944) mostrou que deve ser minimizada a soma dos quadrados dos desvios ponderados pelo inverso da variância de  $p_i$ , que é dada por  $P_i Q_i / n_i$ .

Obtém-se, assim, um sistema de equações que não tem solução, porque os pesos são valores desconhecidos.

Então, usando como fator de ponderação as estimativas  $\hat{P}_i$  e  $\hat{Q}_i$ , obtém-se o sistema de equações normais:

$$\left. \begin{aligned} \sum_i \frac{n_i (p_i - \hat{P}_i) (p_i - 2p_i \hat{P}_i + \hat{P}_i^2)}{\hat{P}_i \hat{Q}_i} &= 0 \\ \sum_i \frac{n_i x_i (p_i - \hat{P}_i) (p_i - 2p_i \hat{P}_i + \hat{P}_i^2)}{\hat{P}_i \hat{Q}_i} &= 0 \end{aligned} \right\} \quad (3)$$

É claro que se pode chegar à solução do sistema (3) através da série de Taylor, por aproximações sucessivas.

Considerando este procedimento muito trabalhoso, BERKSON (1944) propôs uma simplificação. Em vez de trabalhar diretamente com as frequências relativas, o autor obteve, partindo do lógite e usando algumas aproximações, uma soma de quadrados de desvios ponderados, que corresponde a uma regressão linear de  $y_i$ , definido como  $y_i = \log p_i/q_i$  contra  $x_i$ . Os fatores de ponderação dos quadrados dos desvios são  $n_i p_i q_i$ . BERKSON (1955), denominou a expressão que dá a soma dos quadrados dos desvios ponderados de " $\chi^2$  de lógite", que indicou como segue:

$$\chi^2 (\text{lógite}) = \sum n_i p_i q_i (y_i - \hat{y}_i)^2 \quad (4)$$

A partir de (4) obtém-se o sistema de equações:

$$\left. \begin{aligned} \hat{\alpha} \sum n_i p_i q_i + \hat{\beta} \sum n_i p_i q_i x_i &= \sum n_i p_i q_i y_i \\ \hat{\alpha} \sum n_i p_i q_i x_i + \hat{\beta} \sum n_i p_i q_i x_i^2 &= \sum n_i p_i q_i x_i y_i \end{aligned} \right\} \quad (5)$$

Baseado na facilidade da solução do sistema de equações (5), BERKSON (1944, 1951, 1953, 1955) recomendou enfaticamente este método.

SILVERSTONE (1957) verificou que as estimativas obtidas pela aproximação do método dos mínimos quadrados são consistentes, no sentido de convergência em probabilidade, mas argumentou que tais estimativas não são suficientes. Isto porque se um dos valores  $p$  é 0 ou 1, o sistema de equações (5) fica indeterminado. É claro que se pode usar a

"regra do  $2n$ ", proposta por BERKSON (1955), que consiste em substituir  $p=1$  por  $p=1 - 1/2 n_j$  e  $p=0$  por  $p_j=1/2 n_j$ . REIERSOL (1961) afirmou que esta modificação não muda as propriedades assintóticas dos estimadores. Entretanto, SILVERSTONE (1957) argumentou que, usando esta modificação, os estimadores obtidos não tem a propriedade de suficiência, no sentido de que nem toda informação necessária para a estimação dos parâmetros está contida na amostra.

Lembrando que as dificuldades computacionais para a solução do sistema de equações (2) praticamente desapareceram, desde que se disponha de um computador, SILVERSTONE (1957), recomendou que as estimativas de  $\alpha$  e de  $\beta$  fossem obtidas através do método de máxima verossimilhança, que é o único indicado por JOHNSON e KOTZ (1970).

HODGES (1958) desenvolveu uma técnica, chamada método de transferência, que dá boas estimativas preliminares para se iniciar o processo iterativo de solução das equações obtidas pelo método de máxima verossimilhança. A técnica é, em última análise, um processo semigráfico, cujo uso é pouco defensável com a atual disponibilidade de computadores.

É claro que se obtêm boas estimativas preliminares para o processo iterativo de solução do sistema de equações obtido pelo método da máxima verossimilhança pela solução do sistema de equações (5).

### 3. A FUNÇÃO DE GOMPERTZ

#### 3.1. A Função de Gompertz para dados de crescimento

##### 3.1.1. Características

A função de Gompertz, definida por

$$Z = e^{a-bc^t} \quad (1)$$

onde  $a$ ,  $b$  e  $c$  são parâmetros,  $b > 0$  e  $0 < c < 1$ , foi pela primeira vez indicada para o estudo descritivo de crescimento por PRESCOTT (1922), num estudo sobre as leis de crescimento em previsão de demanda.

Não obstante a data da primeira recomendação, o número de autores que ajustaram a função de Gompertz a dados experimentais é pequeno e as referências a esta função, comparativamente à função logística, são poucas.

A função (1) é monotonicamente crescente e fica entre as duas assíntotas  $Z = 0$  e  $Z = e^a$ . Como no caso da função logística, o parâmetro  $e^a$ , que é a distância entre as duas assíntotas, pode ser denominado "nível de saturação".

A função (1) tem ponto de inflexão para a abcissa  $t = (-\log b)/\log c$ , onde a ordenada vale  $Z = e^{a-1}$ .

O estudo descritivo de crescimento através da função de Gompertz pode ser baseado na equação diferencial:

$$\frac{dZ}{dt} = k Z (a - \log Z) \quad (2)$$

onde  $k = -\log c$  é uma constante.

A equação (2) mostra que a taxa de crescimento da função de Gompertz é proporcional ao valor alcançado pela função e à diferença entre o logaritmo desse valor e o logaritmo da assíntota.

Portanto, na função de Gompertz também se reconhece um "fator de momento" igual a  $Z$  e um "fator de contenção", igual a  $(a - \log Z)$ .

De (2) segue-se que:

$$Z^{-1} \frac{dZ}{dt} = k (a - \log Z) \quad (3)$$

isto é, a taxa de crescimento relativo de  $Z$  decresce linearmente com o logaritmo de  $Z$ .

Estudando as leis de crescimento de demanda, PRESCOTT (1922) argumentou que o crescimento de uma indústria pode ser dividido em quatro estágios: período de experimentação, período de crescimento com taxas crescentes, período de crescimento com taxas decrescentes e período de estabilização. Para este autor, a previsão do futuro de uma indústria pode ser feito através da função de Gompertz, porque as

características desta função atendem às pressuposições do crescimento industrial, mostrando inclusive a tendência à estabilização.

GIRÃO (1964) ajustou a função de Gompertz, por dois métodos alternativos, à produção de leite de ovinos em Portugal, no período 1947-1962. Os métodos simplificados de ajustamento usados pelo autor, levaram apenas às estimativas dos parâmetros. Portanto, não se pode avaliar, com base nos resultados que constam no trabalho, o grau de ajustamento da função aos dados apresentados.

LEAL (1972) utilizou dados de índice de custo de vida em Curitiba, no período compreendido entre janeiro de 1965 e dezembro de 1971, para comparar diversos métodos de ajustamento da função logística. O autor ajustou a função de Gompertz pelo método de STEVENS (1951) aos mesmos dados, e verificou que eram mais plausíveis os resultados obtidos pelo ajustamento desta função.

Os dados de crescimento de animais e vegetais, de acordo com DAY (1966), são adequadamente descritos por famílias de curvas com 3 parâmetros. Para a escolha da família de curvas, o autor partiu da equação diferencial de 2a. ordem:

$$f(t) \frac{d^2 f(t)}{dt^2} - (a+1) \left( \frac{df(t)}{dt} \right)^2 + bf(t) \frac{df(t)}{dt} = 0 \quad (4)$$

discutida na seção 2.1.1.. DAY (1966) mostrou que fazendo  $a = 0$  em (4), obtêm-se como solução a função de Gompertz.

A função de Gompertz foi ajustada pelo método dos totais parciais por AMER e WILLIAMS (1957), aos dados de crescimento em área de

folhas de Pelargonium zonale, submetidos a diferentes condições de umidade. Os autores puderam estabelecer conclusões com base nas estimativas dos parâmetros, mas não verificaram o grau de ajustamento da função.

NELDER (1962) afirmou que a função logística generalizada proposta por NELDER (1961) deu, para dados de crescimento em peso de cenoura, bons ajustamentos em vários casos, mas em alguns as estimativas de  $1/\theta$  ou eram zero ou eram pequenas e negativas.

Desde que o valor de  $\theta$  não pode ser negativo, NELDER (1962) julgou conveniente propor uma "reparametrização" da função, que tornasse possível a obtenção de valores positivos e negativos para  $\theta$ , e gerasse a função de Gompertz como um caso limite. A função "reparametrizada" é:

$$Z = \frac{a}{[1+\phi e^{-(b+ct)}]^{1/\phi}} \quad (5)$$

Em (5), fazendo  $\phi = 1$  obtém-se a função logística, fazendo  $\phi = -1$ , obtém-se a função de Mitscherlich e fazendo  $\phi$  tender para zero, obtém-se a função de Gompertz.

HOFFMANN e THAME (1970) ajustaram a função de Gompertz e a função logística a dados de crescimento de Pinus caribaea var. caribaea e verificaram que os resultados obtidos com o ajustamento da função de Gompertz eram melhores que os resultados obtidos com o ajustamento da função logística.

Estudando dados de crescimento ponderal de gado bovino, MISCHAN (1972) ajustou a função de Gompertz e a função logística. A autora verificou que os resultados obtidos com o ajustamento da função de Gompertz eram melhores que os resultados obtidos com o ajustamento da função logística.

A função de Gompertz foi ajustada a dados de crescimento de quatro espécies de Eucaliptus, em dois espaçamentos, por HOFFMANN e BERGER (1973). Os autores verificaram que a função de Gompertz se ajustou bem aos dados, pois se obteve, nos oito casos analisados, coeficientes de determinação elevados. O valor de F para regressão é sempre alto e o valor de F para falta de ajustamento, obtido de análises de variância em parcelas subdivididas, é significativo ao nível de 5%, em apenas três casos. Entretanto os autores julgaram de interesse lembrar que, além da função de Gompertz, outras funções podem ser ajustadas aos mesmos dados, levando, eventualmente, a melhores resultados.

Note-se que a conclusão de que os resultados obtidos com o ajustamento da função de Gompertz são melhores que os resultados obtidos com o ajustamento da função logística, deve ser vista com cuidado.

Isto porque os autores mencionados, que chegaram a esta mesma conclusão, embora usando dados essencialmente diferentes, utilizaram o mesmo método de STEVENS (1951), tanto para o ajustamento da função logística como para o ajustamento da função de Gompertz.

Ora, o método de STEVENS (1951) exige anamorfose dos dados. Entretanto, a transformação de variáveis exigida para o ajustamento da função de Gompertz é a aplicação de logaritmos à variável dependente, enquanto que a transformação de variáveis exigida para o ajustamento da função logística é a inversão da variável dependente.

A aplicação de logaritmos pressupõe erro multiplicativo, justificável nos modelos para crescimento animal ou vegetal, enquanto que a inversão da variável coloca, no modelo, o erro em posição difícil de explicação, como foi discutido na seção 2.1.2..

Por outro lado, na função logística, a ordenada do ponto de inflexão é a metade do valor da assíntota, o que caracteriza o crescimento logístico como aquele que tem taxas simétricas em torno desse ponto. Ora, pode-se argumentar que esta restrição, embora não seja incompatível com o processo de crescimento animal ou vegetal, pode estar em desacordo com o mesmo.

É interessante lembrar que PEARL e REED (1920) afirmaram textualmente que a função logística, devido às suas características restritivas, representa apenas uma primeira aproximação para a verdadeira lei de crescimento. Estes autores recomendaram que fosse feita uma modificação na função, de modo a desaparecerem as restrições.

Na função de Gompertz, a ordenada do ponto de inflexão é  $e^a/e$ . Este valor, embora relacionado com o valor da assíntota, é menor do que a ordenada do ponto de inflexão da função logística, o que atende melhor às características do processo de crescimento animal ou vegetal.

Portanto, é de se supor que os resultados obtidos por HOFFMANN e THAME (1970) e MISCHAN (1972), não são casos particulares nem são consequências apenas do método de ajustamento. É possível que a função de Gompertz seja realmente melhor do que a função logística para a descrição de crescimento animal ou vegetal, em certos casos.

### 3.1.2. Estimativas dos parâmetros

As estimativas dos parâmetros da função de Gompertz, definida por

$$Z = e^{a-bc^t} \quad (1)$$

devem ser obtidas, fundamentalmente, por métodos iterativos.

Entretanto, a quantidade de cálculo que estes métodos exigem tornam quase imprescindível o uso de computador. Existem métodos para a obtenção das estimativas dos parâmetros da função de Gompertz que tem a vantagem de simplificar e diminuir a quantidade de cálculo.

O método mais antigo, conhecido como dos "totais parciais", consta em livros textos como o de CROXTON e COWDEN (1955). Este método consiste em dividir as observações em três grupos e calcular os subtotais de cada grupo. Se o número de dados não for divisível por 3, abandonam-se 1 ou 2 dados. Obtêm-se, assim, três equações em a, b e c, que são resolvidas algebricamente. As estimativas dos parâmetros, obtidas por este método, são consistentes, mas não são eficientes, conforme demonstrou STEVENS (1951).

DAY (1963) apresentou dois métodos alternativos para a obtenção das estimativas dos parâmetros da função de Gompertz.

O primeiro método apresentado por DAY (1963) consiste em aplicar logaritmo a função (1), para obter:

$$Y = a - bc^t \quad (2)$$

onde  $Y = \log Z$ .

De (2), aplicando logaritmo, obtém-se:

$$\log (a-Y) = \log b + t \log c \quad (3)$$

que é uma regressão linear simples de uma função de  $Z$  contra  $t$ . Desde que se tenha uma estimativa de  $\underline{a}$ , obtida por processo gráfico conforme recomenda o autor, obtém-se facilmente as estimativas de  $\underline{b}$  e de  $\underline{c}$ .

É claro que as estimativas dos parâmetros, obtidas por este método, não apresentam as propriedades estatísticas desejáveis. Isto porque a estimativa da assíntota, da qual dependem as estimativas dos outros dois parâmetros, é obtida por simples observação de pontos colocados em um gráfico.

Outro método para a obtenção das estimativas dos parâmetros da função de Gompertz, também apresentado por DAY (1963), quando  $t_i = m i (i = 0, 1, \dots, n-1)$ , consiste em aplicar logaritmo à função (1), para obter:

$$c^{mi} = (a - Y_i) / b \quad (4)$$

onde  $Y_i = \log Z_i$ .

Substituindo o valor  $c^{mi}$  obtido em (4) na função (2), quando a variável independente assume o valor  $t_i = mi+m$ , obtêm-se:

$$Y_{i+1} = a(1-c^m) + c^m Y_i \quad (5)$$

onde  $Y_{i+1} = \log Z_{i+1}$

Então, fazendo uma regressão linear simples de  $\log Z_{i+1}$  contra  $\log Z_i$ , obtêm-se as estimativas de  $a$  e de  $c$ . A estimativa de  $b$  é obtida aplicando logarítmo a função (1). Desde que  $a$  e  $c$  já foram estimados, obtêm-se:<sup>1/</sup>

$$b = \text{antilog} \left[ \frac{\sum \log(a-y_i)}{n} - \frac{m(n-1)\log c}{Z} \right] \quad (6)$$

PATTERSON (1956) sugeriu um método simplificado para o ajustamento da função de Spillman, apresentada na forma:

$$Y = a - \beta \rho^t \quad (7)$$

nos casos em que existem 4, 5, 6 ou 7 pares de observações e os valores de  $t$  são igualmente espaçados.

Note-se que, por anamorfose, a função de Gompertz se transforma na função (7), pois aplicando logarítmo à função (1), obtêm-se:

$$\log Z = a - bc^t \quad (8)$$

<sup>1/</sup> Note-se que na expressão apresentada por DAY (1963) falta a indicação de logarítmo em  $\frac{\sum \log(a-y_i)}{n}$  e o autor indica  $i=j, j+1, \dots, j+n-1$ .

Ora a expressão (6) só é válida se  $i=0, 1, 2, \dots, n-1$ .

De acordo com PATTERSON (1956), a estimativa de  $\rho$  pode ser obtida pela razão entre dois contrastes de  $y$ . Tais contrastes são escolhidos de maneira que as eficiências das estimativas propostas sejam as mais altas possíveis. Evidentemente, os contrastes são diferentes quando diferem o número de observações. Obtida a estimativa de  $\rho$ , obtêm-se as estimativas de  $\alpha$  e de  $\beta$ , através de uma regressão linear simples.

As estimativas de  $\alpha$ ,  $\beta$  e  $\rho$ , obtidas por este método não são eficientes, conforme discutiu o autor. Entretanto, elas podem ser usadas como estimativas preliminares para iniciar os processos iterativos de ajustamento da função de Gompertz.

Outros dois métodos simplificados para a obtenção das estimativas dos parâmetros da função de Gompertz podem ser propostos.

O primeiro consiste em estabelecer uma regressão linear simples dos logarítmos das diferenças dos logarítmos de cada duas observações consecutivas de  $Z$  contra  $t$ .

Desde que as observações da variável independente sejam equiespaçadas, isto é, desde que  $t_i = mi$ , aplicando logarítmo em (1) e estabelecendo a diferença entre os logarítmos de duas observações consecutivas, obtêm-se:

$$Y_{i+1} - Y_i = bc^{mi} (1-c^m) \quad (9)$$

onde  $Y_{i+1} = \log Z_{i+1}$  e  $Y_i = \log Z_i$

Aplicando logarítmo à função (9), obtêm-se:

$$\log (Y_{i+1} - Y_i) = mi \log c + \log [b(1-c^m)] \quad (10)$$

A regressão linear simples de  $\log (y_{i+1} - y_i)$  contra  $i$ , obtida em (10) dá as estimativas de  $b$  e de  $c$ . De acordo com (2), a estimativa de  $a$  pode ser obtida como segue:

$$a = \frac{\Sigma Y + b \Sigma c^{mi}}{n} \quad (11)$$

O segundo método consiste em aplicar logaritmo a função (1) e diferenciar, para obter:

$$\frac{1}{Z} \frac{dZ}{dt} = -a \log c + Y \log c \quad (12)$$

onde  $Y = \log Z$ .

Aproximando  $\frac{dZ}{dt}$  por  $\frac{\Delta Z}{\Delta t}$ , pode-se ajustar uma regressão linear simples de  $\Delta Z/Z\Delta t$  contra  $Y$ . Obtêm-se, então, as estimativas dos parâmetros  $a$  e  $c$ . De (2) segue-se que:

$$b = \frac{a-Y}{c^t} \quad (13)$$

A estimativa de  $b$  pode ser obtida pela média aritmética dos valores obtidos em (13), para cada par de observações  $t_i, z_i$ .

STEVENS (1951) desenvolveu um método iterativo de ajustamento para a função de Spillman e mostrou que o método se aplica às funções que, por anamorfose, se transformam naquela função. Inclui-se neste caso a função logística, cujo ajustamento é discutido na seção 2.1.2., e a função de Gompertz.

A estimativa preliminar de  $\rho$ , necessária para iniciar o processo iterativo de cálculo pelo método proposto por STEVENS (1951), pode ser obtida por qualquer um dos métodos de ajustamento anteriormente descritos. Neste trabalho é usado o método proposto por PATTERSON (1956).

É importante notar que quando se ajusta a função de Gompertz pelo método proposto por STEVENS (1951), o modelo estatístico da função em análise é:

$$Z = e^{a-bc^t} \epsilon \quad (14)$$

sendo  $\log \epsilon$  variáveis aleatórias independentes de média zero e variância constante.

Portanto, minimiza-se a soma dos quadrados dos desvios dos logaritmos dos valores observados em relação aos valores estimados pela função de Spillman.

Feita a anamorfose  $Y = \log Z$ , exigida pelo método de STEVENS (1951) para o ajustamento da função de Gompertz, admite-se que a variância do logaritmo da variável dependente é constante.

Esta pressuposição, que é comumente feita quando se trata dados econômicos, é a mesma feita para o ajustamento da função logística pelo método proposto por NELDER (1961). Então, no estudo descritivo de dados de crescimento, quando se pode admitir que o logaritmo da variável dependente tem variância constante e o modelo da função não é estabelecido, devem ser ajustadas ambas as funções, isto é, a função logística pelo método proposto por NELDER (1961) e a função de Gompertz pelo método

proposto por STEVENS (1951).

Quando se admite que os erros são aditivos e homocedásticos, deve-se aplicar diretamente o método dos mínimos quadrados. É importante ressaltar que o método dos mínimos quadrados tem a vantagem de não exigir pressuposição a respeito da forma específica da distribuição dos erros.

Então, dado o modelo:

$$z = e^{a-bc^t} + \epsilon \quad (15)$$

onde  $\epsilon$  são variáveis aleatórias com média zero e variância  $\sigma^2$ , obtêm-se imediatamente o sistema de equações normais:

$$\left. \begin{aligned} \Sigma(z-\hat{z}) \hat{z} &= 0 \\ \Sigma(z-\hat{z}) c^t \hat{z} &= 0 \\ \hat{b} \Sigma(z-\hat{z}) t \hat{c}^{t-1} \hat{z} &= 0 \end{aligned} \right\} \quad (16)$$

onde  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$  são as estimativas de  $a$ ,  $b$  e  $c$  respectivamente,  $z$  são os valores observados e  $\hat{z}$  são as estimativas da função.

Para resolver um sistema de equações que não tem solução explícita, como é o caso do sistema (16), utiliza-se processo iterativo.

Sejam  $\Delta a$ ,  $\Delta b$  e  $\Delta c$  as correções obtidas em determinada etapa para as estimativas da etapa anterior, isto é, sejam

$$\left. \begin{aligned} \Delta a &= \hat{a} - \hat{a}_0 \\ \Delta b &= \hat{b} - \hat{b}_0 \\ \Delta c &= \hat{c} - \hat{c}_0 \end{aligned} \right\} \quad (17)$$

onde  $\hat{a}_0$ ,  $\hat{b}_0$  e  $\hat{c}_0$  são estimativas preliminares.

Expandindo cada uma das equações normais pela série de Taylor e desprezando os termos onde  $\Delta a$ ,  $\Delta b$  e  $\Delta c$  tem expoente superior a um, obtém-se:

$$\left[ \begin{array}{c} F + Q \\ \\ \\ \end{array} \right] \left[ \begin{array}{c} \Delta a \\ \Delta b \\ \Delta c \end{array} \right] = \left[ \begin{array}{c} \Sigma (z - \bar{z}_0) \bar{z}_0 \\ - \Sigma (z - \bar{z}_0) \hat{c}_0^t \bar{z}_0 \\ - \hat{b}_0 \Sigma (z - \bar{z}_0) t \hat{c}_0^{t-1} \bar{z}_0 \end{array} \right] \quad (18)$$

onde

$$F = \left[ \begin{array}{ccc} \Sigma \bar{z}_0^2 & - \Sigma \hat{c}_0^t \bar{z}_0^2 & - \hat{b}_0 \Sigma t \hat{c}_0^{t-1} \bar{z}_0^2 \\ - \Sigma \hat{c}_0^t \bar{z}_0^2 & \Sigma \hat{c}_0^{2t} \bar{z}_0^2 & \hat{b}_0 \Sigma t \hat{c}_0^{2t-1} \bar{z}_0^2 \\ - \hat{b}_0 \Sigma t \hat{c}_0^{t-1} \bar{z}_0^2 & \hat{b}_0 \Sigma t \hat{c}_0^{2t-1} \bar{z}_0^2 & \hat{b}_0^2 \Sigma t^2 \hat{c}_0^{2t-2} \bar{z}_0^2 \end{array} \right] \quad (19)$$

e  $Q$  é uma matriz simétrica cujos elementos são dados por:

$$\left. \begin{aligned}
 Q_{11} &= \Sigma (z - \bar{z}_o) \bar{z}_o \\
 Q_{12} &= \Sigma (z - \bar{z}_o) \hat{c}_o^t \bar{z}_o \\
 Q_{13} &= \bar{b}_o \Sigma (z - \bar{z}_o) t \hat{c}_o^{t-1} \bar{z}_o \\
 Q_{22} &= - \Sigma (z - \bar{z}_o) \hat{c}_o^{2t} \bar{z}_o \\
 Q_{23} &= \Sigma (z - \bar{z}_o) t \hat{c}_o^{t-1} \bar{z}_o (1 - \bar{b}_o \hat{c}_o^t) \\
 Q_{33} &= \Sigma (z - \bar{z}_o) t \hat{c}_o^{t-2} \bar{z}_o (t-1 - \bar{b}_o t \hat{c}_o^t)
 \end{aligned} \right\} \quad (20)$$

É importante notar que todos os elementos da matriz  $Q$  definida em (20), são somas ponderadas de desvios, o que torna os valores de seus elementos praticamente desprezíveis em relação aos valores dos elementos da matriz  $F$ , definida em (19).

Então, obtém-se uma aproximação do sistema de equações (18), como segue:

$$F \begin{bmatrix} \Delta a \\ \Delta b \\ \Delta c \end{bmatrix} = \begin{bmatrix} \Sigma (z - \bar{z}_o) \bar{z}_o \\ - \Sigma (z - \bar{z}_o) \hat{c}_o^t \bar{z}_o \\ - \bar{b}_o \Sigma (z - \bar{z}_o) t \hat{c}_o^{t-1} \bar{z}_o \end{bmatrix} \quad (21)$$

O sistema (21) permite obter correções para estimativas sucessivas de  $a$ ,  $b$  e  $c$ , desde que se disponha das estimativas preliminares  $\hat{a}_o$ ,  $\hat{b}_o$  e  $\hat{c}_o$ .

Tais estimativas, necessárias para se iniciar o processo iterativo, podem ser obtidas por quaisquer dos métodos simplificados, discutidos aqui. Neste trabalho é utilizado o método definido por (12) e (13).

Observe-se que o procedimento clássico para a solução do sistema de equações normais (16) leva ao sistema de equações (18), que deve convergir para a solução final mais rapidamente que o sistema de equações (21), desde que este é apenas uma aproximação daquele.

A recomendação de que as estimativas dos parâmetros da função de Gompertz, quando se considera um erro aditivo no modelo, sejam obtidas pelo sistema de equações (21), se baseia tanto no fato de que o sistema é mais simples, como também no fato de que a perda em velocidade de convergência pode não ser suficientemente importante.

Entretanto, CORNFIELD e MANTEL (1950), em situação análoga na análise de próbites, observaram que o aumento na velocidade de convergência para a solução final era considerável, quando usaram as equações exatas.

Este último método para obtenção das estimativas dos parâmetros da função de Gompertz pressupõe que a variável dependente tem variância constante. Esta pressuposição também é feita para o ajustamento da função logística pelo método proposto por SCHULTZ (1930), mas usado com número suficiente de iterações apenas por OLIVER (1964).

## 3.1.3. Variâncias das estimativas dos parâmetros

Obtêm-se as variâncias das estimativas dos parâmetros da função de Gompertz, ajustada pelo método proposto por STEVENS (1951), como foi discutido na seção 2.1.3.

Quando se admite o modelo:

$$z = e^{a-bc^t} + \epsilon \quad (1)$$

onde  $\epsilon$  são variáveis aleatórias independentes com distribuição normal de média zero e variância  $\sigma^2$ , a obtenção das variâncias das estimativas dos parâmetros da função de Gompertz, dadas pelo sistema de equações (21) da seção 3.1.2., podem ser obtidas, à semelhança dos outros métodos, através da matriz de informação, desde que se verifique que as estimativas de máxima verossimilhança coincidem com as estimativas de quadrados mínimos.

Então, diferenciando o logaritmo da função de verossimilhança  $L$  obtida para  $n$  pares de observações, vem:

$$\left. \begin{aligned} \frac{\partial \log L}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (z-Z)^2 \\ \frac{\partial \log L}{\partial a} &= \frac{1}{\sigma^2} \sum [Z (z-Z)] \\ \frac{\partial \log L}{\partial b} &= -c^t \frac{1}{\sigma^2} \sum [-c^t Z (z-Z)] \\ \frac{\partial \log L}{\partial c} &= -b \frac{1}{\sigma^2} \sum [tc^{t-1} Z (z-Z)] \end{aligned} \right\} \quad (2)$$

onde  $Z$  são valores da função.

Sejam  $s^{*2}$ ,  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$  as estimativas de máxima verossimilhança de  $\sigma^2$ ,  $a$ ,  $b$  e  $c$ . Evidentemente:

$$s^{*2} = \frac{1}{n} \sum (z - \bar{z})^2 \quad (3)$$

As estimativas  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$ , obtidas das três últimas equações de (2), independentemente de  $s^{*2}$ , coincidem com as estimativas de quadrados mínimos obtidas em (16), na seção 3.1.2.

De acordo com as propriedades do método de máxima verossimilhança,  $s^{*2}$ ,  $\hat{a}$ ,  $\hat{b}$  e  $\hat{c}$  tem, para um número suficientemente grande de pares de observações, distribuição aproximadamente normal de média ( $\sigma^2$ ,  $a$ ,  $b$ ,  $c$ ) e matriz de variâncias e covariâncias dada pela inversa da matriz de informação.

Esta matriz de informação é obtida de (2) através de nova diferenciação, troca de sinal e aplicação de esperança, isto é:

$$M^{-1} = \frac{1}{\sigma^2} \begin{bmatrix} n/2\sigma^2 & 0 & 0 & 0 \\ 0 & \sum z^2 & -\sum c^t z^2 & -b\sum tc^{t-1} z^2 \\ 0 & -\sum c^t z^2 & \sum c^{2t} z^2 & b\sum tc^{2t-1} z^2 \\ 0 & -b\sum tc^{t-1} z^2 & b\sum tc^{2t-1} z^2 & b^2 \sum t^2 c^{2(t-1)} z^2 \end{bmatrix} \quad (4)$$

A matriz (4) não pode ser invertida na forma algébrica, mas mostra as seguintes propriedades:

- i) A covariância entre  $s^{*2}$  e cada uma das estimativas dos parâmetros é zero.
- ii) A variância de  $s^{*2}$  é dada por  $2\sigma^4/n$ , independentemente dos valores dos outros três parâmetros e dos valores assumidos pela variável in dependente.
- iii) As variâncias e as covariâncias das estimativas dos parâmetros  $a$ ,  $b$  e  $c$  são dadas por  $F^{-1} \sigma^2$ , onde  $F$  é a matriz definida em (19) na seção 3.1.2..

### 3.2. A Função de Gompertz para dados de resposta quântica

#### 3.2.1. Características

A função de Gompertz, definida por

$$Z = e^{\gamma - \alpha\beta^x} \quad (1)$$

com  $a > 0$  e  $0 < \beta < 1$ , desde que o nível de saturação  $e^\gamma$  seja conhecido, pode gerar a função

$$P = e^{-\alpha\beta^x}, \quad (2)$$

onde  $P$  expressa a proporção do valor máximo de  $Z$ .

No estudo de dados de resposta quântica, o ajustamento da função (2) constitui uma alternativa para a análise de próbites e de lôgites.

Então, como nestas análises, a variável independente  $x_i$  é a quantidade da dose ou do estímulo e a variável dependente é a frequência relativa, dada por  $p_i = m_i/n_i$ , sendo  $n_i$  o número de indivíduos submetidos à dose  $x_i$  e  $m_i$  o número de indivíduos afetados.

Para ser ajustada aos dados de resposta quântica, a função (2) deve ser linearizada. Pela aplicação de logaritmo duas vezes consecutivas em (2), obtêm-se:<sup>1/</sup>

$$\log (-\log P) = \log \alpha + x \log \beta = a + bx \quad (3)$$

Conforme observaram JOHNSON e KOTZ (1970), a semelhança dos resultados das análises de próbites e lógites, feitas sobre os mesmos dados experimentais, é devida à semelhança das formas das distribuições normal e logística.

A distribuição de Gompertz, entretanto, difere destas distribuições, principalmente por ser positivamente assimétrica.

Portanto, é de se esperar que o resultado do ajustamento da função (2) aos dados de resposta quântica, sejam diferentes dos resultados obtidos das análises de próbites e de lógites.

Evidentemente, a validade do ajustamento da função (2) deve ser discutida, principalmente em termos de resultados obtidos com dados experimentais.

---

<sup>1/</sup> A variável transformada  $Y = \log (-\log P)$  poderia ser denominada Gômpite, por analogia com a palavra lógite.

Note-se que, em resposta quântica, tem-se a variável  $p$ , definida no intervalo  $[0,1]$ . É interessante assinalar que tanto o lógite como a transformada definida em (3) são definidos no intervalo  $(-\infty, \infty)$ .

### 3.2.2. Estimativas dos parâmetros

As estimativas dos parâmetros da função definida por:

$$p = e^{-\alpha\beta^x} \quad (1)$$

podem ser obtidas pelo método dos mínimos quadrados ou da máxima verossimilhança.

De acordo com o método dos mínimos quadrados, deve-se minimizar a função

$$f = \sum \frac{n_i}{P_i Q_i} (p_i - \hat{P}_i)^2, \quad (2)$$

onde  $n_i$  é o número de ensaios relativos a  $x_i$ ,  $p_i$  é a frequência relativa,  $\hat{P}_i$  é a estimativa de  $P_i$ , e  $Q_i = 1 - P_i$ .

Os quadrados dos desvios são ponderados por  $n_i/P_i Q_i$ , porque  $p_i$  tem variância igual a  $P_i Q_i/n_i$ . Entretanto, como  $P_i$  e  $Q_i$  são valores desconhecidos, serão utilizados, como fatores de ponderação, as estimativas  $\hat{P}_i$  e  $\hat{Q}_i$ . Portanto, minimiza-se a função:

$$f = \sum \frac{n_i}{\hat{P}_i \hat{Q}_i} (p_i - \hat{P}_i)^2 \quad (3)$$

De (3), obtêm-se o sistema de equações normais, onde os índices foram omitidos por simplificação:

$$\left. \begin{aligned} \Sigma \frac{n}{\hat{P}\hat{Q}^2} \hat{\beta}^x (p-\hat{P}) (p-2p \hat{P}+\hat{P}) &= 0 \\ \Sigma \frac{n}{\hat{P}\hat{Q}^2} x \hat{\beta}^{x-1} (p-\hat{P}) (p-2p \hat{P}+\hat{P}) &= 0 \end{aligned} \right\} \quad (4)$$

O sistema (4) não tem solução explícita, devendo ser resolvido por processo iterativo.

Sejam  $\hat{\alpha}_0$  e  $\hat{\beta}_0$  as estimativas preliminares de  $\alpha$  e de  $\beta$  e sejam  $\Delta\alpha = \hat{\alpha} - \hat{\alpha}_0$  e  $\Delta\beta = \hat{\beta} - \hat{\beta}_0$  as correções que devem ser estimadas.

Então, desenvolvendo cada uma das equações do sistema (4), de acordo com a série de Taylor obtêm-se, desprezando os termos em que  $\Delta\alpha$  e  $\Delta\beta$  aparecem com expoente superior a um, o sistema de equações:

$$F \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} - \Sigma \hat{\beta}_0^x \delta \\ - \Sigma x \hat{\beta}_0^{x-1} \delta \end{bmatrix} \quad (5)$$

onde

$$\delta = \frac{n}{\hat{P}_0 \hat{Q}_0} (p-\hat{P}_0) (p-2p \hat{P}_0 + \hat{P}_0) ,$$

com  $\hat{P}_0 = e^{-\hat{\alpha}_0} \hat{\beta}_0^x$  e  $\hat{Q}_0 = 1 - \hat{P}_0$ ;  $F$  é uma matriz cujos elementos são dados por:

$$\begin{aligned}
 F_{11} &= \Sigma \hat{\beta}_0^{2x} \phi \\
 F_{12} &= \hat{\alpha} \Sigma x \hat{\beta}_0^{2x-1} \phi + \Sigma x \hat{\beta}_0^{x-1} \delta \\
 F_{21} &= \Sigma x \hat{\beta}_0^{2x-1} \phi \\
 F_{22} &= \hat{\alpha} \Sigma x^2 \hat{\beta}_0^{2x-2} \phi + \Sigma (x-1) \hat{\beta}_0^{x-2} \phi
 \end{aligned}
 \tag{7}$$

com

$$\phi = \frac{n}{\hat{p}_0 \hat{q}_0^3} \left[ (1-3 \hat{p}_0 + 4 \hat{p}_0^2) p^2 + (1+\hat{p}_0 - 2p-2p \hat{p}_0) p_0^2 \right]
 \tag{8}$$

Se as correções  $\Delta\alpha$  e  $\Delta\beta$  obtidas em (5) não forem desprezíveis, reiniciam-se os cálculos, utilizando os valores corrigidos  $\hat{\alpha}_0 + \Delta\alpha$  e  $\hat{\beta}_0 + \Delta\beta$ , em lugar das estimativas preliminares.

Evidentemente, o processo deve ser repetido até que as correções sejam consideradas desprezíveis.

Para a obtenção das estimativas preliminares de  $\alpha$  e de  $\beta$ , necessárias para se iniciar o processo iterativo de cálculo, pode-se partir da transformada:

$$Y = \log (-\log P)
 \tag{9}$$

De (9) obtêm-se, por diferenciação:

$$dP = P \log P dY
 \tag{10}$$

Aproximando  $dP$  e  $dY$  por  $\Delta P = p - \hat{P}$  e  $\Delta Y = y - \hat{Y}$ , respectivamente, pode-se escrever:

$$p - \hat{P} = P \log P (y - \hat{Y}) \quad (11)$$

Portanto, o método dos mínimos quadrados implica, aproximadamente, em minimizar:

$$f = \sum_n \frac{\hat{P}}{Q} (\log \hat{P})^2 (y - \hat{Y})^2 \quad (12)$$

ou, ainda aproximadamente, em minimizar:

$$f = \sum \frac{np}{1-p} (\log p)^2 (y - \hat{Y})^2 \quad (13)$$

onde

$$y = \log (- \log p) \quad (14)$$

e

$$\hat{Y} = \log (- \log \hat{P}) = \log \hat{\alpha} + x \log \hat{\beta} = \hat{a} + \hat{b}x \quad (15)$$

Portanto, de acordo com (13) e (15), as estimativas preliminares de  $a$  e de  $b$  e, conseqüentemente, as estimativas preliminares de  $\alpha$  e de  $\beta$ , podem ser obtidas fazendo uma regressão linear simples ponderada de  $y$  contra  $x$ , usando como fatores de ponderação  $[np/(1-p)](\log p_i)^2$ .

Note-se que as estimativas assim obtidas podem ser tomadas como finais, evidentemente com prejuízo da exatidão.

Entretanto, a facilidade da obtenção destas estimativas, em situação análoga no estudo de lógites, fez com que BERKSON (1944, 1951, 1953, 1955) recomendasse, enfaticamente, o método.

Para estimar os parâmetros da função (1) através do método da máxima verossimilhança, deve-se admitir que a variável aleatória que representa o número de indivíduos afetados ( $m_i$ ), do total de indivíduos submetidos a dose  $x_i$  ( $n_i$ ), tem distribuição binomial cuja função de probabilidade é:

$$P(m_i) = \binom{n_i}{m_i} P_i^{m_i} Q_i^{n_i - m_i} \quad (16)$$

De acordo com o método da máxima verossimilhança, deve-se maximizar a função, onde o índice  $i$  é omitido por simplificação:

$$\log L = \sum \log \binom{n}{m} + \sum m \log P + \sum (n-m) \log Q \quad (17)$$

Obtém-se, então, o sistema de equações:

$$\left. \begin{aligned} \sum \frac{n}{Q} (p - \hat{P}) \hat{\beta}^x &= 0 \\ \sum \frac{n}{Q} (p - \hat{P}) x \hat{\beta}^{x-1} &= 0 \end{aligned} \right\} \quad (18)$$

O sistema (18), que não tem solução explícita, é resolvido por processo iterativo, baseado na linearização das equações, utilizando apenas os dois primeiros termos do desenvolvimento pela série de Taylor.

Obtém-se, assim, o sistema de equações lineares:

$$H \begin{bmatrix} \Delta \alpha \\ \Delta \beta \end{bmatrix} = \begin{bmatrix} - \sum \frac{n}{Q_0} (p - \hat{P}_0) \hat{\beta}_0^x \\ - \sum \frac{n}{Q_0} (p - \hat{P}_0) x \hat{\beta}_0^{x-1} \end{bmatrix} \quad (19)$$

onde  $H$  é uma matriz cujos elementos são:

$$\begin{aligned}
 H_{11} &= \sum \frac{n\hat{p}_o}{\hat{Q}_o^2} \hat{\beta}_o^{2x} (1-p) \\
 H_{12} &= \hat{\alpha}_o \sum \frac{n\hat{p}_o}{\hat{Q}_o^2} \times \hat{\beta}_o^{2x-1} (1-p) + \sum \frac{n}{\hat{Q}_o} \times \hat{\beta}_o^{x-1} (p-\hat{P}_o) \\
 H_{21} &= \sum \frac{n\hat{p}_o}{\hat{Q}_o^2} \times \hat{\beta}_o^{2x-1} (1-p) \\
 H_{22} &= \hat{\alpha}_o \sum \frac{n\hat{p}_o}{\hat{Q}_o^2} \times x^2 \hat{\beta}_o^{2x-2} (1-p) + \sum \frac{n}{\hat{Q}_o} \times (x-1) (p-\hat{P}_o) \hat{\beta}_o^{x-2}
 \end{aligned} \quad (20)$$

É claro que as estimativas preliminares  $\hat{\alpha}_o$  e  $\hat{\beta}_o$ , necessárias para se iniciar o processo iterativo de solução do sistema (19), podem ser obtidas da forma já proposta, sugerida pelas relações (13)e(15).

Note-se que os estimadores de máxima verossimilhança são diferentes dos estimadores de mínimos quadrados. No caso da função logística, os autores que se preocuparam com as propriedades dos estimadores recomendaram que fosse usado o método da máxima verossimilhança para a estimação dos parâmetros.

Neste caso, não foram estudadas as propriedades dos estimadores, mas o método da máxima verossimilhança conduz a estimadores consistentes e assintoticamente eficientes.

Seria interessante, usando simulação, comparar o comportamento dos estimadores de mínimos quadrados e de máxima verossimilhança, no caso da função de Gompertz, como foi feito por BERKSON (1955) e por SILVERSTONE (1957), no caso da logística.

#### 4. A FUNÇÃO LOG-INVERSA GENERALIZADA

##### 4.1. A Função log-inversa generalizada para dados de crescimento

###### 4.1.1. Características

A função log-inversa, definida por

$$Z = BR^{\frac{1}{t}} \quad (1)$$

onde B e R são parâmetros,  $0 < R < 1$ , foi acrescida de um parâmetro e indicada para o estudo descritivo de crescimento por ARRUDA, VIEIRA e HOFFMANN (1973).

Esta função, aqui denominada de log-inversa generalizada é definida por:

$$Z = A + BR^{\frac{1}{t}} \quad (2)$$

onde A, B e R são parâmetros,  $B > 0$  e  $0 < R < 1$ .

A função (2) é monotonicamente crescente e tem como as síntota a reta  $Z = A + B$ . Tem ponto de inflexão para a abcissa  $t = -\log R/2$ , com ordenada igual a  $A + \frac{B}{e^2}$ .

Note-se que  $\lim_{t \rightarrow 0} Z = A$ . Então, se  $Z$  é o valor da variável dependente em certo instante  $t$ ,  $Z-A$  é o crescimento observado até aquele instante.

O estudo descritivo de crescimento através da função log-inversa generalizada pode ser baseado na equação diferencial:

$$\frac{dZ}{dt} = k (Z-A) \frac{1}{t^2}, \quad (3)$$

onde  $k = -\log R$  é uma constante.

A equação (3) mostra que a taxa de crescimento da função log-inversa generalizada é diretamente proporcional ao acréscimo  $(Z-A)$  e inversamente proporcional ao quadrado da abcissa naquele instante.

Portanto, reconhece-se na função log-inversa generalizada, um "fator de momento" igual a  $(Z-A)$  e um "fator de contenção" igual a  $1/t^2$ .

De (3) segue-se que

$$(Z-A)^{-1} \frac{dZ}{dt} = k \frac{1}{t^2}$$

isto é, a taxa de crescimento de  $Z$ , relativa ao crescimento já observado, é inversamente proporcional ao quadrado de  $t$ .

É interessante notar, também que

$$\lim_{t \rightarrow 0^+} \frac{dZ}{dt} = 0 \quad (5)$$

isto é, a curva apresenta declividade inicial nula. Esta característica pode corresponder a um fenômeno real, se a função for usada para descrever crescimento animal ou vegetal. Isto porque o crescimento ponderal de animais fica praticamente paralizado logo após o nascimento e o crescimento das plantas fica praticamente paralizado logo após o transplante.

ARRUDA, VIEIRA e HOFFMANN (1973) ajustaram a função log-inversa generalizada e a função de Gompertz a dados de crescimento de Pinus caribaea var. caribaea, a dados de crescimento de Pinus caribaea var. hondurensis e a dados de crescimento de uma rês. Comparando os resultados do ajustamento das duas funções, os autores verificaram que o ajustamento da função log-inversa generalizada revelou-se melhor que o ajustamento da função de Gompertz, em um dos três casos estudados. Portanto, no estudo descritivo de crescimento, a função log-inversa generalizada deve ser pesquisada, ao lado da função logística e da função de Gompertz, estas mais conhecidas.

#### 4.1.2. Estimativas dos parâmetros

As estimativas dos parâmetros da função log-inversa generalizada, definida por

$$Z = A + BR^{\frac{1}{t}} \quad (1)$$

podem ser obtidas pelo método proposto por STEVENS (1951) para o ajustamento da função de Spillman, desde que se faça a inversão da variável independente.

A estimativa preliminar de R, necessária para se iniciar o processo iterativo de cálculo é facilmente obtida se for possível estabelecer com precisão razoável a abcissa do ponto de inflexão, através de processo gráfico ou devido ao conhecimento do fenômeno. Isto porque a relação entre o parâmetro R e a abcissa do ponto de inflexão é dada por

$$t = (-\log R)/2 \quad (2)$$

Também se pode obter uma estimativa preliminar de R, desde que se conheçam os valores aproximados das ordenadas da curva para valores da abcissa t iguais a zero, m e nm. Então, facilmente se obtém:

$$R = \left( \frac{Z_m - Z_0}{Z_{nm} - Z_0} \right)^{\frac{nm}{n-1}} \quad (3)$$

Note-se que quando se ajusta a função log-inversa generalizada pelo método proposto por STEVENS (1951), o modelo estatístico da função em análise é:

$$Z_i = A + BR^{\frac{1}{t}} + \epsilon \quad (4)$$

onde  $\epsilon$  são variáveis aleatórias independentes de média zero e variância constante, isto é, admite-se a existência de um erro aditivo e admite-se que a variável dependente tem variância constante.

Portanto, no estudo descritivo de crescimento, quando o modelo da função não é determinado a priori e é razoável admitir que a variável dependente tem variância constante, deve-se ajustar a função logística pelo método proposto por SCHULTZ (1930), a função de Gompertz pelo método dos mínimos quadrados e a função log-inversa generalizada pelo método proposto por STEVENS (1951), para o ajustamento da função de Spillman. Os resultados obtidos com esses ajustamentos são imediatamente comparáveis.

#### 4.1.3. Variâncias das estimativas dos parâmetros

As variâncias das estimativas dos parâmetros da função log-inversa generalizada, ajustada pelo método proposto por STEVENS (1951), são obtidas da maneira discutida na seção 2.1.3.

### 4.2. A Função log-inversa generalizada para dados de resposta quântica

#### 4.2.1. Características

A função log-inversa generalizada, definida por

$$Z = A + B\rho^{\frac{1}{x}} \quad (1)$$

desde que o nível de saturação  $A + B$  seja conhecido, pode gerar a função

$$Q = \beta \left( 1 - \rho^{\frac{1}{x}} \right) \quad (2)$$

onde  $\beta = \frac{B}{A+B}$  e  $P = 1 - Q$  expressa a proporção do valor máximo de  $Z$ .

No estudo de dados de resposta quântica, o ajustamento da função (2) poderia constituir uma alternativa para a análise de pró-bites, ou para o ajustamento das funções logística e de Gompertz.

Entretanto, o ajustamento das funções assintótico-sigmóides aos dados de resposta quântica se baseia na facilidade de, conhecido o nível de saturação, estabelecer uma regressão linear ponderada, de uma função da variável dependente, que é uma proporção, contra o valor da dose ou do estímulo.

Ora, não é possível colocar a função (2) sob forma linear, razão pela qual a função log-inversa generalizada não é ajustada, neste trabalho, aos dados de resposta quântica.

## 5. EXEMPLOS DE APLICAÇÃO

### 5.1. O Problema da Especificação

As funções assintótico-sigmóides tem grande aplicação no estudo descritivo de dados biológicos de crescimento.

Entretanto, na maioria dos casos, a biologia não dispõe de teorias que levem à escolha da forma matemática da função.

Uma pressuposição a respeito da variância da variável dependente, para um dado tempo, que pode ser feita ou com base em conhecimentos anteriores, ou com base em testes clássicos para homocedasticidade, facilita a escolha entre possíveis especificações.

Assim, quando se pode pressupor que o logaritmo da variável dependente tem variância constante, devem ser comparados os resultados dos ajustamentos da função logística pelo método proposto por NELDER (1961) e da função de Gompertz pelo método proposto por STEVENS (1951). Quando se pode pressupor que a variável dependente tem variância constante, devem ser comparados os resultados dos ajustamentos da

função logística pelo método usado por OLIVER (1964), da função de Gompertz como se propõe aqui e da função log-inversa generalizada, pelo método proposto por STEVENS (1951).

Nestes casos, a especificação do modelo pode ser feita através da comparação das variâncias residuais, critério que embora extensivamente usado, não é perfeito (THEIL (1971)).

Entretanto, muitas vezes não existem fundamentos suficientes mesmo para uma pressuposição a respeito da variância dos dados.

A sugestão é obter o ajustamento das três funções pelos diferentes métodos e especificar o modelo através do procedimento empírico, sugerido por THEIL (1971).

Assim, nos casos em que se ajusta uma função usando as próprias observações, além das somas dos quadrados dos desvios dos valores observados em relação aos valores estimados pela função, calcula-se a soma dos quadrados dos desvios dos logarítmos dos valores observados em relação ao logarítmo dos valores estimados pela função. Nos casos em que se ajusta a função usando os logarítmos das observações, além das somas dos quadrados dos desvios dos logarítmos dos valores observados em relação aos valores estimados pela função, calcula-se a soma dos quadrados dos desvios dos valores observados em relação ao antilogarítmo dos valores estimados pela função.

Na comparação dos resultados de funções ajustadas com as próprias observações e funções ajustadas com o logarítmo dessas

observações utiliza-se os valores das duas somas de quadrados de desvios, escolhendo-se a função que levou às menores somas.

É importante frisar aqui que se a teoria biológica define a forma matemática da função, não importam os valores das somas dos quadrados dos desvios, pois o modelo especificado pela teoria deve ser o único utilizado.

## 5.2. Estudo descritivo do crescimento ponderal de animais

O estudo de dados biológicos de crescimento tem diversas aplicações de suma importância.

Assim, a avaliação dos inibidores e dos estimuladores que atuam sobre o crescimento pode ser feita através da comparação de curvas que descrevem o fenômeno.

O uso de processos gráficos é o procedimento mais comum nestes casos, seguido da alternativa de comparação de médias, obtidas sob condições diferentes, para o mesmo tempo, através de testes de significância.

Entretanto, a solução mais completa é a comparação de funções, uma ajustada aos dados coletados sob a ação do inibidor ou do estimulador, e a outra estabelecida para dados coletados em condições normais.

Por outro lado, o estudo de dados de crescimento de animais e de vegetais pode apresentar, em certos casos, a par do interesse

puramente biológico, uma importância econômica.

Assim, as funções de crescimento, ajustadas aos dados de peso de animais de corte, podem ser utilizadas para a determinação da idade econômica de abate, como mostraram MISCHAN (1972) e CRÓCOMO (1973).

As funções de crescimento, ajustadas aos dados de volume de madeira de povoamentos florestais, podem ser utilizadas para a determinação da idade econômica de corte, de acordo com HOFFMANN e THAME (1970) e HOFFMANN e BERGER (1973).

Com a finalidade de ilustrar aspectos metodológicos do ajustamento das funções assintótico-sigmóides, foram utilizados dados de crescimento ponderal de ratos e dados de crescimento ponderal de vacas holandesas.

Os dados relativos ao crescimento ponderal de ratos foram obtidos pelo Dr. Décio Teixeira, com a finalidade de verificar a influência da sialoadenectomia, realizada em idades diferentes, no crescimento ponderal de ratos (TEIXEIRA, NEGREIROS DE PAIVA e ALMEIDA COSTA (1970)). No experimento em questão foram obtidos os pesos individuais diários, em gramas, de 73 ratos normais desde o primeiro até o 60º dia de vida, que conduziram aos pesos médios que constam na Tabela I.

Tabela I. Pesos médios, em g, de 73 ratos normais, desde o primeiro até o 60º dia de vida, obtidos pelo Dr. Décio Teixeira. (Os valores estão separados por ponto e vírgula e ordenados em linha, de acordo com o dia em que foram obtidos, isto é, o primeiro valor corresponde ao peso médio do primeiro dia, e assim por diante).

---

5,56; 6,17; 7,01; 8,37; 8,98; 9,27; 10,46; 11,50; 14,36; 15,15;  
 14,74; 16,03; 16,99; 18,79; 19,84; 20,90; 22,34; 23,94; 27,03; 28,25;  
 31,00; 32,91; 35,49; 37,25; 38,12; 40,97; 42,37; 44,80; 46,34; 48,12;  
 44,64; 46,26; 48,46; 50,45; 53,56; 56,11; 58,10; 60,54; 63,93; 65,48;  
 67,16; 69,60; 70,41; 73,79; 76,25; 79,08; 81,75; 81,85; 85,89; 88,47;  
 88,56; 92,77; 94,07; 95,33; 97,28; 99,00; 100,88; 102,06; 103,00; 105,17.

---

Com estas médias foram feitos os ajustamentos da função logística pelo método de NELDER (1961), da função logística pelo método recomendado por OLIVER (1964), da função de Gompertz pelo método de STEVENS (1951), da função de Gompertz pelo método que pressupõe um erro aditivo no modelo e da função log-inversa generalizada pelo método de STEVENS (1951).

Na Tabela II constam as estimativas dos parâmetros ( $\hat{a}, \hat{b}, \hat{c}$ ), das abscissas dos pontos de inflexão, das ordenadas das assíntotas e das somas de quadrados de resíduo, tanto em termos da variável dependente original como em termos do logaritmo da variável dependente.

A simples observação da Tabela II mostra que as menores somas de quadrados de resíduo referem-se à função de Gompertz.

Entretanto, quando se comparam as somas de quadrados de desvios relativas a função de Gompertz, ajustada pelo método de STEVENS (1951) e pelo método que pressupõe um erro aditivo, verifica-se que o procedimento proposto por THEIL (1971) não pode levar a uma especificação completa do modelo. Isto porque a soma de quadrados de desvios em termos da variável original, obtida pelo ajustamento da função pelo método que pressupõe um erro aditivo, é necessariamente menor que a soma de quadrados de desvios em termos da variável original, obtida pelo ajustamento da função pelo método de STEVENS (1951). Por sua vez, em termos do logaritmo da variável dependente, a soma de quadrados de desvios relativa ao método de STEVENS (1951) é necessariamente menor que a soma de quadrados de desvios relativa ao método que pressupõe um erro aditivo.

Portanto, o procedimento sugerido por THEIL (1971) mostra que a função de Gompertz é a que melhor se ajusta aos dados apresentados de crescimento ponderal de ratos; entretanto, não se pode decidir, com base nesse procedimento, entre o modelo estatístico com erro aditivo e o modelo estatístico com erro multiplicativo.

Tabela II. Estimativas dos parâmetros ( $\hat{a}, \hat{b}, \hat{c}$ ), das abcissas dos pontos de inflexão, das ordenadas das assíntotas e das somas de quadrados de resíduo, tanto em termos da variável dependente original como em termos do logaritmo da variável dependente.

Modelos	Estimativas dos parâmetros	Abcissa do ponto de inflexão	Ordenada da assíntota	Soma de quadrados	
				(original)	(logaritmo)
logística (NELDER)	$\hat{a} = 4,662$ $\hat{b} = 2,801$ $\hat{c} = 0,08627$	32,5	105,88	698,40	0,2818
logística (OLIVER)	$\hat{a} = 129,926$ $\hat{b} = 14,979$ $\hat{c} = 0,068667$	39,4	129,93	235,36	0,8246
Gompertz (STEVENS)	$\hat{a} = 5,1098$ $\hat{b} = 3,4697$ $\hat{c} = 0,96709$	37,2	165,64	181,18	0,1039
Gompertz (aditivo)	$\hat{a} = 5,2558$ $\hat{b} = 3,4865$ $\hat{c} = 0,97083$	42,2	191,68	143,77	0,1711
Log-inversa generalizada	$\hat{a} = 12,014$ $\hat{b} = 238,951$ $\hat{c} = 0,1446$	29,0	250,96	548,44	2,0183

Tendo em vista o conhecimento dos pesquisadores sobre o assunto, o ponto de inflexão, que deve coincidir com a puberdade, e o peso do animal adulto são melhor estimados, pela função de Gompertz com erro aditivo, como mostram os resultados da Tabela II.

Portanto, este modelo parece descrever melhor o crescimento ponderal de ratos normais.

Deve ser lembrado, entretanto, que a especificação do modelo deveria ser feita com base no estudo das variâncias dos pesos e das variâncias dos logarítmos dos pesos, entre animais dentro de idades, o que não pôde ser feito porque não se dispunha dos pesos individuais.

Outro procedimento que poderia ser utilizado para a especificação do modelo, é o que consta em DRAPER e SMITH (1966), e se refere ao exame dos sinais dos desvios. Verificou-se que no ajustamento da função de Gompertz com erro aditivo, existem 13 grupos de sinais, contra 10 grupos existentes para o ajustamento da função de Gompertz pelo método de STEVENS (1951). Este resultado reforça a conclusão estabelecida acima, uma vez que, se os erros fossem aleatórios, existiriam ao redor de 30 grupos de sinais.

Os dados de crescimento ponderal de vacas holandesas foram obtidos pelo Posto Zootécnico "Luiz de Queiroz" de Piracicaba. Tais dados constam na Tabela III do apêndice e referem-se a pesagens mensais dos animais, em quilogramas, desde o nascimento até o 56º mês de idade.

Antes de se proceder ao ajustamento das funções assintótico-sigmóides a estes dados, fêz-se uma análise preliminar da variância dos pesos e da variância dos logarítmos dos pesos, entre animais dentro de idades, obtendo os resultados que constam na Tabela IV.

Observe-se que a variância tendeu à estabilização após a aplicação de logarítmo à variável dependente, o que é mostrado pelo teste de Bartlett para homogeneidade de variâncias, que resultou em  $\chi^2 = 176,23$ , com 56 graus de liberdade, quando se usa a variável

Tabela IV. Estimativas de médias relativas a idades e de variância entre animais dentro de idades para pesos e logarítmos de pesos de 10 vacas holandesas do Posto Zootécnico "Luiz de Queiroz", SP.

Idade (meses)	Peso		Logarítmo do peso	
	Média	Variância	Média	Variância
0	39.1	37.21	3.65	0.02317
1	52.8	59.51	3.95	0.02302
2	70.7	130.45	4.24	0.02997
3	93.5	171.38	4.52	0.02303
4	118.0	155.11	4.76	0.01144
5	134.7	281.12	4.89	0.01772
6	153.5	384.27	5.02	0.01993
7	172.5	392.50	5.14	0.01669
8	176.9	615.21	5.16	0.02476
9	191.8	666.84	5.24	0.01916
10	204.3	536.45	5.31	0.01319
11	219.4	626.93	5.38	0.01295
12	223.6	462.71	5.40	0.00976
13	235.3	397.34	5.45	0.00770
14	230.7	585.34	5.43	0.01116
15	229.2	699.28	5.42	0.01356
16	232.3	2459.12	5.41	0.07452
17	247.0	1163.77	5.50	0.01904
18	261.6	1244.48	5.55	0.01851
19	276.4	1374.26	5.61	0.02029
20	288.4	1425.37	5.65	0.01955
21	302.7	1908.01	5.70	0.02431
22	315.4	1678.04	5.74	0.01909
23	325.9	2300.32	5.77	0.02346
24	342.8	2838.62	5.82	0.02594
25	355.9	2330.76	5.86	0.01945
26	366.9	2737.43	5.89	0.02177
27	375.7	2006.23	5.92	0.01558
28	381.5	1848.94	5.93	0.01440
29	384.4	1373.82	5.94	0.01003
30	389.9	1586.76	5.96	0.01182
31	402.9	1149.21	5.99	0.00730
32	420.3	1337.78	6.03	0.00736
33	415.8	1793.51	6.02	0.01075
34	426.3	1802.45	6.05	0.01048
35	427.5	2006.50	6.05	0.01182
36	439.6	2587.37	6.07	0.01497
37	436.7	3696.23	6.07	0.02115
38	445.7	3713.34	6.09	0.01927
39	449.5	4469.16	6.09	0.02329
40	442.8	4238.62	6.08	0.02101

Tabela IV. Continuação

Idade (meses)	Peso		Logarítmo do peso	
	Média	Variância	Média	Variância
41	451.5	4139.38	6.10	0.01932
42	453.2	3781.73	6.10	0.01768
43	436.0	3151.11	6.07	0.01533
44	438.6	1688.93	6.07	0.00856
45	446.6	1257.37	6.09	0.00630
46	453.6	1480.93	6.11	0.00743
47	450.9	2425.43	6.10	0.01314
48	475.0	2996.66	6.15	0.01266
49	483.8	2421.95	6.17	0.00939
50	487.3	1690.01	6.18	0.00644
51	485.0	1974.44	6.18	0.00822
52	495.4	2121.82	6.20	0.00926
53	487.6	2012.26	6.18	0.00896
54	481.5	1873.16	6.17	0.00836
55	491.9	1441.87	6.19	0.00633
56	491.4	1735.82	6.19	0.00733

dependente na forma original e em  $\chi^2 = 61,60$ , com o mesmo número de graus de liberdade quando se usa o logarítmo da variável dependente.

O primeiro valor de  $\chi^2$  é significativo ao nível de 5% o que não acontece com o segundo.

Portanto, foram ajustadas a função logística pelo método de NELDER (1961) e a função de Gompertz pelo método de STEVENS (1951), obtendo-se as estimativas dos parâmetros que constam na Tabela V.

Tabela V. Estimativas dos parâmetros relativas ao ajustamento das funções logística e de Gompertz, aos dados da Tabela III.

Função logística	Função de Gompertz
$\hat{a} = 6,11018$	$\hat{a} = 6,15289$
$\hat{b} = 1,83436$	$\hat{b} = 2,21665$
$\hat{c} = 0,13798$	$\hat{c} = 0,91668$

Com base nos resultados da Tabela V, foram estimados os valores das duas funções para cada  $t_i$ , que foram utilizados para construir o Gráfico I.

Também foram obtidas as somas dos quadrados de resíduo de regressão, que valem 16,8536 no caso da função logística e 13,0466 no caso da função de Gompertz. É importante lembrar que estas somas de quadrados tem validade para comparação, mas conduzem a subestimativas da variância do erro uma vez que as observações não são independentes. Então, usando o critério de variância mínima, pode-se afirmar que a função de Gompertz ajustou-se ligeiramente melhor aos dados do que a função logística.

O ponto de inflexão, estimado pela função logística, tem abcissa igual a 13,3 e a estimativa da ordenada da assíntota, em termos da variável original, é igual a 450.

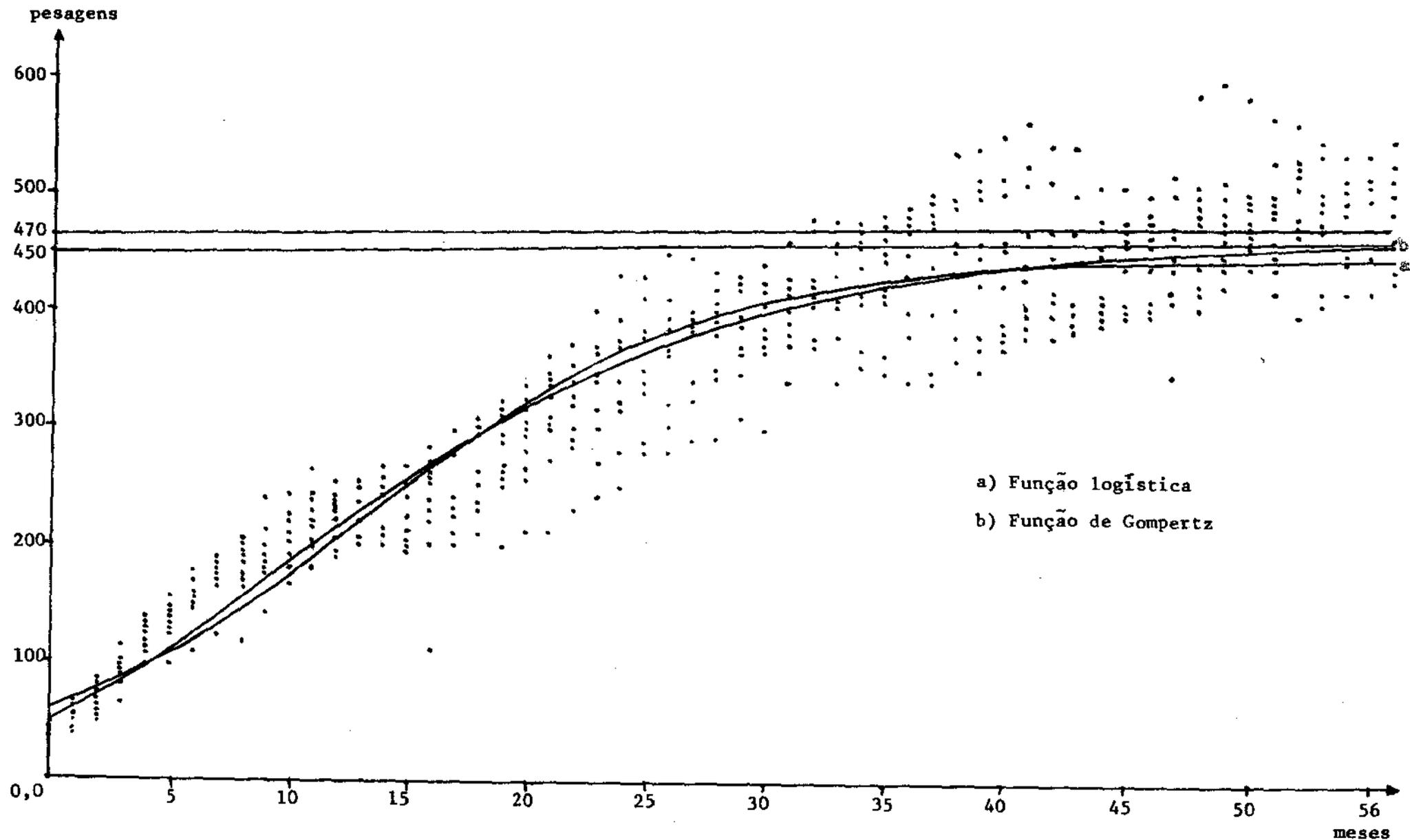


Gráfico I: Função logística e função de Gompertz ajustadas às pesagens mensais de 10 vacas holandesas do Posto Zootécnico "Luiz de Queiroz", Piracicaba, SP.

No caso da função de Gompertz, a estimativa da abcissa do ponto de inflexão é igual a 9,2 e a estimativa da ordenada da assíntota, em termos da variável original, é igual a 470.

Note-se que parâmetros importantes, como época em que ocorre a inflexão da curva e peso do animal adulto, são melhor estimados, do ponto de vista zootécnico, pela função de Gompertz.

Entretanto, tanto no caso da função logística como no caso da função de Gompertz, existem observações, e médias de observações para a mesma idade, com valores mais altos que a ordenada da assíntota.

Apesar dos problemas levantados aqui, parece que a curva de crescimento que melhor descreve dados de crescimento ponderal de gado bovino é a função de Gompertz, ajustada pelo método de STEVENS(1951). Aliás, esta função foi utilizada, para este tipo de dados, por MISCHAN (1972) e por CRÓCOMO (1973).

### 5.3. Estudo descritivo do número de dentes permanentes irrompidos, em crianças brancas, da cidade de Piracicaba, SP

O conhecimento do número de dentes permanentes irrompidos, nas crianças em idade da dentição mista, tem várias aplicações práticas em odontologia.

O estudo desta característica também tem grande importância em odontologia legal, porque o número de dentes permanentes irrompidos pode ser usado como variável independente para a estimação da idade de uma criança.

Com esta finalidade, ARBENS (1961) estabeleceu a regressão linear simples de uma transformada da proporção de dentes permanentes irrompidos, em função da idade em meses, para escolares brancos, de 7 a 13 anos, residentes na cidade de São Paulo, SP.

Baseado na pressuposição de que a transformada da variável dependente, definida como  $y = \arcsen \sqrt{p}$ , onde  $p$  é a proporção de dentes permanentes irrompidos, tem distribuição aproximadamente normal e variância estável, o autor obteve a análise de variância de regressão. O valor de  $F$  calculado para desvio de linearidade com 82 graus de liberdade foi igual a 1,27, que coincide com o  $F$  crítico ao nível de 5% de significância. Baseado neste resultado, o autor não rejeitou a hipótese de linearidade, o que é discutível.

O número de dentes permanentes irrompidos pode ser visto como dado de resposta quântica.

Isto porque os indivíduos estão divididos em  $s$  grupos etários; relaciona-se ao  $j^{\text{ésimo}}$  grupo uma variável não aleatória  $x_j$ , que é a idade em meses, e um número  $n_j = 32 r_j$  de ensaios, onde 32 é o número máximo de dentes permanentes em indivíduos normais e  $r_j$  é o número de indivíduos no grupo. Ora, observa-se uma variável binomial, que é dente permanente irrompido ou não irrompido, sendo  $32 r_j$  o número de ensaios binomiais para aquele grupo.

Como dados de resposta quântica, o número de dentes permanentes irrompidos podem ser estudados através da função logística e da função de Gompertz.

O fenômeno pode ser visto como aproximadamente assintótico, ou seja, definida a proporção de dentes permanentes irrompidos como a razão entre o número de dentes permanentes irrompidos, para cada grupo etário e 32  $r_j$ , à medida que a idade adulta se aproxima, a proporção tende para 1.

O fenômeno também pode ser visto como sigmóide, ou seja, é possível pressupor que a taxa de crescimento da proporção do número de dentes permanentes irrompidos é pequena no início do processo, aumenta depois até um máximo e decresce no final do processo.

Para o estudo descritivo do número de dentes permanentes irrompidos, em função da idade, foram utilizados os dados que constam na Tabela VI do apêndice.

Tais dados referem-se a crianças brancas, com idades entre 60 e 155 meses completos, que frequentavam creches, parques infantis, jardins de infância, grupos escolares e ginásios de Piracicaba, em 1973. Os dados foram levantados pela C.D. Ana Helena Marcondes, que usou os critérios adotados pela disciplina de Odontologia Preventiva e Saúde Pública da FOP, UNICAMP.

Os dados relativos a cada sexo foram analisados separadamente, uma vez que sexo e grupo etário são fatores que afetam o número de dentes permanentes irrompidos (MOREIRA, VIEIRA e TUMANG (1973)).

Os resultados dos ajustamentos das funções logísticas e de Gompertz aos dados coletados, transformados em resposta quântica, constam na Tabela VII. Ambas as funções foram ajustadas pelo método dos mínimos quadrados, tanto através de uma regressão linear simples ponderada, que é um método aproximado, como por processo iterativo, e pelo método da máxima verossimilhança.

Também foram obtidos os valores das funções para cada  $x_i$ , que poderiam ser utilizados para estimar o número de dentes permanentes irrompidos em cada idade. Com estes valores, para as funções ajustadas pelo método da máxima verossimilhança, foram construídos o gráfico II, relativo ao sexo feminino e o gráfico III, relativo ao sexo masculino.

Verifica-se, pela simples comparação das somas de quadrados de resíduo que constam na Tabela VII, que a função de Gompertz dá maior explicação da variação do número de dentes permanentes irrompidos em função da idade que a função logística.

Evidentemente, esta observação é confirmada pelo exame dos gráficos, que mostram a curva de Gompertz mais próxima das observações do que a curva logística.

Os resultados que constam na Tabela VII também mostram, como era de se esperar, que a abcissa do ponto de inflexão, estimada pela função de Gompertz, é sempre menor que a abcissa do ponto de inflexão estimada pela função logística.

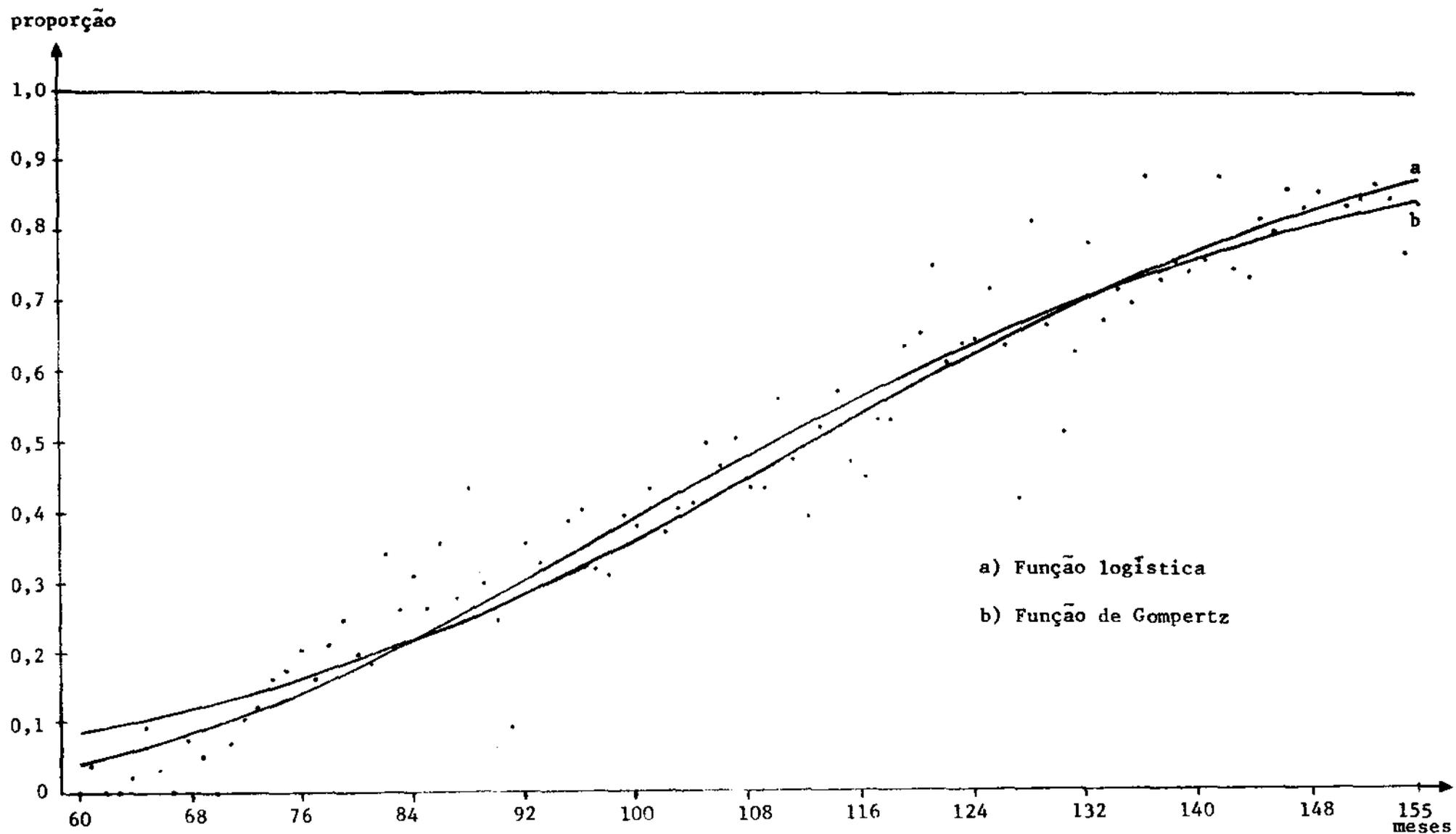


Gráfico II: Função logística e função de Gompertz ajustadas às proporções de dentes permanentes irrompidos em meninas, com idades entre 60 e 155 meses.

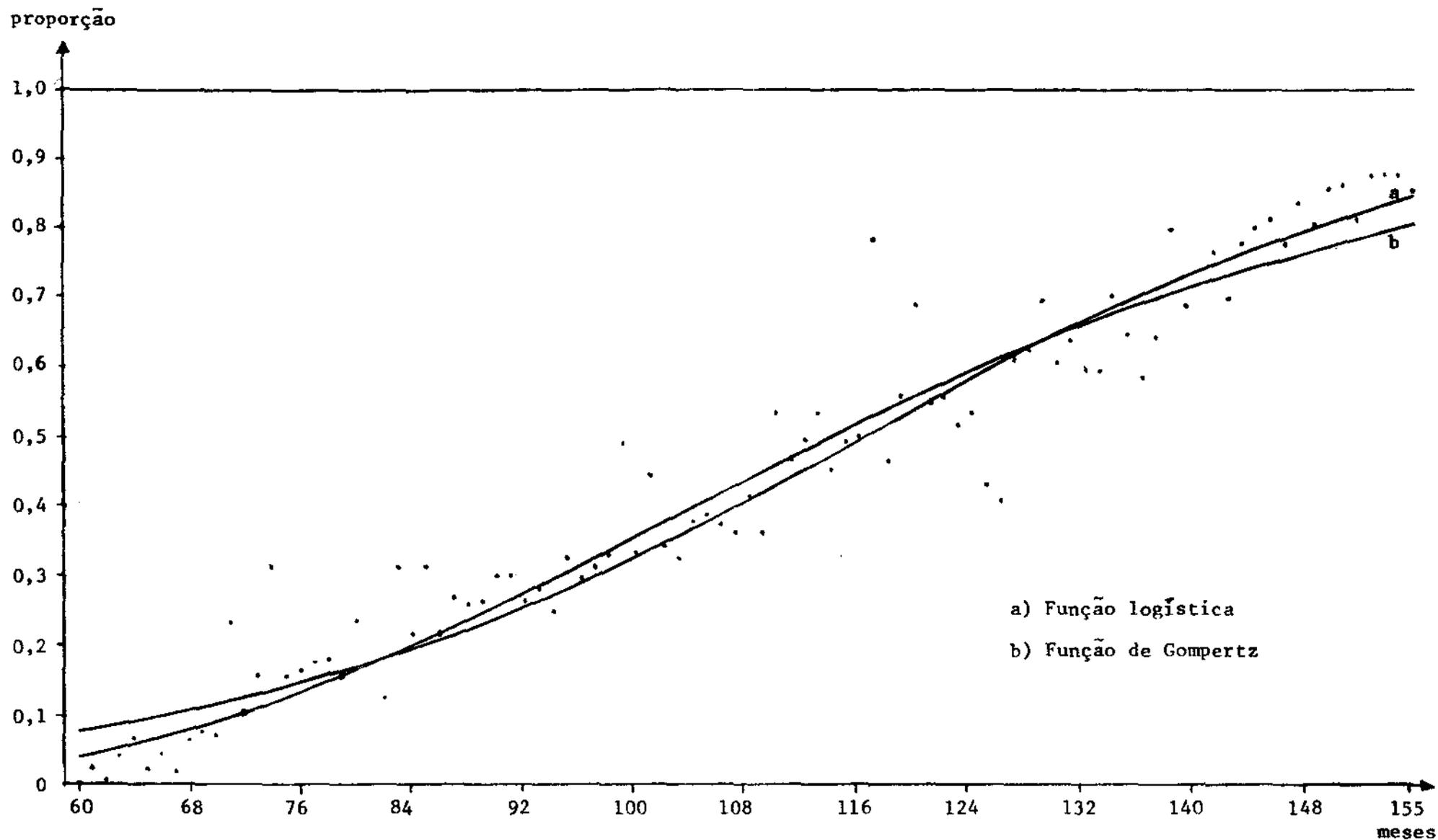


Gráfico III: Função logística e função de Gompertz ajustadas às proporções de dentes permanentes irrompidos em meninos, com idades entre 60 e 155 meses.

Também se verifica, pela simples observação dos resultados que constam na Tabela VII, que em todos os casos, o método dos mínimos quadrados, por processo iterativo, leva a pequenas correções das estimativas preliminares, dadas pelo método que transforma as funções em regressões lineares. Portanto, este procedimento se presta inclusive para a obtenção das estimativas definitivas, dependendo da aproximação admitida.

Os resultados que constam na Tabela VII também mostram que o método da máxima verossimilhança conduz a estimativas ligeiramente maiores que o método dos mínimos quadrados por processo iterativo, o que foi notado por BERKSON (1955) e SILVERSTONE (1957), quando trabalharam com a função logística.

Ainda, uma última observação de interesse para a odontologia, que pode ser feita com base nos resultados que constam na Tabela VII, é a de que, nas meninas, o número estimado de dentes permanentes irrompidos é maior do que nos meninos, quando se considera mesma idade.

As estimativas dos parâmetros, relativas ao sexo feminino, são maiores que as estimativas dos parâmetros relativos ao sexo masculino e a inflexão da curva ocorre, para as meninas, cerca de 6 meses antes que para os meninos. Estes resultados não foram comparados através de testes de significância, porque não foram obtidas as estimativas das variâncias das estimativas dos parâmetros.

Tabela VII. Estimativas dos parâmetros ( $\hat{\alpha}$  e  $\hat{\beta}$ ), das abcissas do ponto de inflexão ( $x_I$ ) e das somas dos quadrados do resíduo, relativas ao ajustamento das funções logísticas e Gompertz, através do método dos mínimos quadrados, tanto por aproximação como por processo iterativo e do método da máxima verossimilhança, para número de dentes permanentes irrompidos em crianças brancas de Piracicaba, SP.

Métodos	Sexo Feminino		Sexo Masculino	
	Logística	Gompertz	Logística	Gompertz
MMQ (aproximado)	$\hat{\alpha} = 4,57$	$\hat{\alpha} = 2,83$	$\hat{\alpha} = 4,82$	$\hat{\alpha} = 2,79$
	$\hat{\beta} = 0,0410$	$\hat{\beta} = 0,0292$	$\hat{\beta} = 0,0415$	$\hat{\beta} = 0,0276$
	$x_I = 111,32$	$x_I = 109,53$	$x_I = 116,09$	$x_I = 114,51$
	SQR=276,44	SQR=209,53	SQR=235,04	SQR=224,54
MMQ (iterativo)	$\hat{\alpha} = 4,78$	$\hat{\alpha} = 2,87$	$\hat{\alpha} = 4,89$	$\hat{\alpha} = 2,74$
	$\hat{\beta} = 0,0427$	$\hat{\beta} = 0,0297$	$\hat{\beta} = 0,0421$	$\hat{\beta} = 0,0273$
	$x_I = 111,76$	$x_I = 109,19$	$x_I = 116,15$	$x_I = 113,90$
	SQR=271,33	SQR=208,64	SQR=234,54	SQR=221,38
MMV	$\hat{\alpha} = 5,03$	$\hat{\alpha} = 3,00$	$\hat{\alpha} = 5,09$	$\hat{\alpha} = 2,84$
	$\hat{\beta} = 0,0449$	$\hat{\beta} = 0,0308$	$\hat{\beta} = 0,0437$	$\hat{\beta} = 0,0282$
	$x_I = 112,04$	$x_I = 109,26$	$x_I = 116,29$	$x_I = 113,82$
	SQR=277,86	SQR=213,47	SQR=238,54	SQR=225,13

MMQ - Método dos mínimos quadrados

MMV - Método da máxima verossimilhança

## 6. RESUMO E CONCLUSÕES

Neste trabalho são estudadas, comparativamente, três funções assintótico-sigmóides, ou seja, a função logística, a função de Gompertz e a função log-inversa generalizada.

São discutidos tanto os casos de dados de crescimento, em que a ordenada da assíntota superior deve ser estimada, como os casos de dados de resposta quântica, em que a ordenada da assíntota superior é conhecida e a variável dependente se transforma numa proporção dessa assíntota.

Do ponto de vista metodológico, dois novos métodos são desenvolvidos: o ajustamento da função de Gompertz a dados de crescimento considerando um erro aditivo e homocedástico e o ajustamento da mesma função a dados de resposta quântica.

Para exemplificação da metodologia estudada, são apresentados os ajustamentos das funções assintótico-sigmóides e dados de

crescimento ponderal de animais e a dados de resposta quântica de número de dentes permanentes irrompidos.

Com base neste estudo, parece lícito estabelecer as seguintes conclusões:

1) Os métodos mais rigorosos para o ajustamento das funções assintótico-sigmóides dependem de processos iterativos, que exigem grande quantidade de cálculo. É claro que não se deve recorrer aos métodos que simplificam os cálculos, porque tais métodos não tem rigor estatístico. Portanto, no ajustamento de funções assintótico-sigmóides a dados experimentais é quase imprescindível o uso de computador.

2) O ajustamento das funções assintótico-sigmóides aos dados experimentais de crescimento tem grande validade descritiva. Entretanto, nem sempre a teoria especifica a forma matemática da função, razão porque devem ser ajustadas, para posterior comparação dos resultados, a função logística, a função de Gompertz e a função log-inversa generalizada. Quando é razoável pressupor que a variável dependente tem variância constante, devem ser impostos modelos onde o erro aleatório é aditivo, ou seja, devem ser ajustadas a função logística pelo método recomendado por OLIVER (1964), a função de Gompertz na forma proposta neste trabalho e a função log-inversa generalizada pelo método de STEVENS (1951). Quando é razoável pressupor que o logarítmo da variável dependente tem variância constante, devem ser impostos modelos onde o erro aleatório aparece como fator, ou seja, devem ser ajustadas a função logística pelo método proposto por NELDER (1961) e a função de Gompertz

pelo método proposto por STEVENS (1951). Nestes casos, a escolha da melhor especificação pode ser feita através do critério de variância residual mínima. Entretanto, quando não se tem informação a respeito da homocedasticia, todos os métodos de ajustamento das três funções podem ser pesquisados e a especificação do melhor modelo só pode ser feito por procedimento empírico. Note-se que não é recomendado o ajustamento da função logística pelo método proposto por STEVENS (1951), porque a pressuposição exigida, de que o inverso da variável dependente tenha variância constante, não parece justificável.

3) Os dados empíricos de resposta quântica são comumente descritos através das análises de próbites ou de lógites. Os resultados obtidos por estes dois métodos de análise são, na opinião de diversos autores, bastante concordantes. Neste trabalho, verificou-se que a função de Gompertz também pode ser utilizada na descrição de dados de resposta quântica. Quando a assíntota superior é conhecida, a função de Gompertz pode ser ajustada na forma de uma regressão linear simples ponderada, com uma transformada da proporção como variável dependente. Esta forma de ajustamento é relativamente simples e deve ser considerada como uma alternativa para a análise de lógites, na forma recomendada por BERKSON (1945). Por outro lado, deve-se assinalar que o resultado do ajustamento da função de Gompertz a dados de resposta quântica deve ser mais exato quando se obtém estimativas de mínimos quadrados ou de máxima verossimilhança, que exigem processo iterativo de cálculo.

4) No estudo descritivo de dados experimentais de crescimento de ratos, foram ajustadas as três funções assintótico-sigmóides,

admitindo-se tanto modelos com erro multiplicativo como modelos com erro aditivo. A comparação dos resultados foi feita pelo procedimento empírico sugerido por THEIL (1971). Pôde-se concluir que a função de Gompertz se ajustou melhor aos dados. Para o ajustamento das funções assintótico-sigmóides aos dados de crescimento ponderal de gado bovino procedeu-se a uma análise preliminar das variâncias dos pesos e das variâncias dos logarítmos dos pesos entre animais, dentro de idades. Como se verificou que os logarítmos dos pesos tendem à homocedasticidade, ajustou-se a função logística pelo método de NELDER (1961) e a função de Gompertz pelo método de STEVENS (1951). O ajustamento da função de Gompertz revelou-se melhor que o ajustamento da função logística, confirmando resultados de outros trabalhos. Portanto, é possível que, no estudo descritivo de dados biológicos de crescimento, a função de Gompertz dê resultados melhores que a função logística, embora esta seja mais conhecida.

5) O ajustamento da função logística e da função de Gompertz aos dados de dentes permanentes irrompidos mostraram que, em ambos os casos, a regressão linear simples ponderada dá resultado bastante próximo do resultado obtido pelo método dos mínimos quadrados. Por outro lado, verificou-se que o ajustamento da função de Gompertz foi melhor que o ajustamento da função logística, mostrando que, no estudo descritivo de dados de resposta quântica, a alternativa proposta aqui deve ser verificada.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

1. AITCHISON, J. & BENNET, J.A. Polychotomus quantal response by maximum indicant. Biometrika, London, 57: 253-62, 1970.
2. AMER, F.A. & WILLIAMS, W.T. Leaf area growth in Pelargonium zonale. Ann. Bot., London, 21: 339-42, 1957.
3. ARBENZ, G.O. Contribuição para o estudo da estimativa da idade pelo número de dentes permanentes irrompidos em escolares da cidade de Sao Paulo, brancos, nascidos no Brasil. Considerações de carater descritivo sobre as falhas da dentadura durante o periodo compreendido entre a queda do dente decíduo e a erupção do correspondente definitivo nos mesmos individuos. Sao Paulo, 1961. 121 p. [Tese (Catedra) - FSHP].
4. ARRUDA, H.V.; VIEIRA, S.; HOFFMANN, R. Análise de uma função sigmóide. Arq. Inst. Biol., São Paulo, 40: 143-8, 1973.
5. BAIN, A.D. The growth of television ownership in the United Kingdom. Int. Econ. Rev., Pennsylvania, Pa, 3: 145-67, 1962. Apud J. Royal Stat Soc., A, London, 126: 285-99, 1963.
6. \_\_\_\_\_. The growth of demand for new commodities. J. Royal Stat. Soc., A, London, 126: 285-99, 1963.
7. BERKSON, J. Application of the logistic function to bio-assay. J. Am. Statist. Ass., Boston, 39: 357-65, 1944.

8. BERKSON, J. Why I prefer logits to probits. Biometrics, Washington, 7: 327-39, 1951.
9. \_\_\_\_\_. A statistically precise and relatively simple method of quantal response, based on the logistic function. J.Am. Statist. Ass., Boston, 48: 565-99, 1953.
10. \_\_\_\_\_. Maximum likelihood and minimum  $\chi^2$  estimates of the logistic function. J.Am.Statist.Ass., Boston, 50: 130-63, 1955.
11. \_\_\_\_\_. Tables for use in estimating the normal distribution function by normit analysis. Biometrika, London, 44: 411-35, 1957.
12. BLISS, C.I. The calculation of the dosage mortality curve. Ann.Appl. Biol., Cambridge, 22: 134-67, 1935.
13. \_\_\_\_\_. & JAMES, A.T. Fitting the rectangular hyperbola. Biometrics, Washington, 22: 573-602, 1966.
14. COLGUHOUN, D. The individual effective dose, direct assays, all-or-nothing responses and the probit transformation. In: \_\_\_\_\_. Lectures on Biostatistics. Oxford, Clarendon Press, 1971. cap. 14, p. 344-64.
15. CORNFIELD, J. & MANTEL, N. Some new aspects of the application of maximum likelihood to the calculation of the dosage response curve. J.Am.Statist.Ass., Boston, 45: 181-210, 1950.
16. CROCOMO, C.R. Análise econométrica do crescimento ponderal de gado bovino. Piracicaba, 1973. 138 p. [Tese (mestrado) ESALQ-USP].
17. CRAMER, E.M. Some comparisons of methods of fitting the dosage response curve for small samples. J.Am.Statis.Ass., Boston, 59: 779-93, 1964.
18. CROXTON, F.E. & COWDEN, D.J. Analyses of time series. In: \_\_\_\_\_. Applied General Statistics. 2.ed. New York, Prentice-Hall, 1955. cap. 13, p. 297-319.

19. DAVIS, H.T. The analysis of economic time series. Bloomington, Principia Press, 1941.
20. DAY, R.H. Simples methods of estimating certain nonlinear functions. Agriculture Handbook, Washington, 256: 1-27, 1963.
21. DAY, N.E. Fitting curves to longitudinal data. Biometrics, Washington, 22: 276-91, 1966.
22. DRAPER, N. & SMITH, H. An introduction to nonlinear estimation. In: \_\_\_\_\_. Applied Regression Analysis. New York, John Wiley, 1966. cap. 10, p. 263-301.
23. ERKELENS, J. A simple method for estimating parameters of the logistic curve. Statist. Neerlandica, 22: 213-17, 1968. Apud Statistical Theory and Method Abstracts, 10(3): 887, 1969.
- 23a. FINNEY, D.J. Probit Analysis. Cambridge, University Press, 1952.
24. GIRÃO, J.A. A função de Gompertz: características da estimação. Revista Agronômica, Lisboa, 47: 71-84, 1964.
25. HAMMOND, J. Fertility and growth in farm animals: general principles. In: \_\_\_\_\_. Farm animals. 3.ed. London, Edward Arnold, 1960. cap. 1, p. 1-21.
26. HODGES, J.L. Fitting the logistic by maximum likelihood. Biometrics, Washington, 14: 453-61, 1958.
27. HOFFMANN, R. & THAME, A.C.M. Determinação da idade ótima para primeiro desbaste em povoamentos de Pinus caribaea. Agro-Econômico, Brasília, 2: 9-37, 1970.
28. HOFFMANN, R. & BERGER, R. Determinação da idade ótima de corte de povoamentos de Eucaliptus. IPEF, Piracicaba, 7: 49-69, 1973.
29. JOHNSON, N.L. & KOTZ, S. Logistic distribution. In: \_\_\_\_\_. Continuous Univariate distributions-2. Boston, Houghton Mifflin, 1970. cap. 22, p. 1-21.
30. LANGE, O. Estudo e previsão do curso de conjuntura econômica. In: \_\_\_\_\_. Introdução à econometria. 2.ed. São Paulo Fundo de Cultura, 1967, cap. 1, p. 23-81.

31. LEAL, J. Comparação de métodos para ajuste da curva logística, com emprego de computador. Curitiba, 1972. 54 p. (Mimeografado).
32. LITTLE, R.E. A note on estimation for quantal response data. Biometrika, London, 55: 578-9, 1968.
33. MISCHAN, M.M. Análise econométrica de crescimento de gado bovino. Botucatu, 1972. 121 p. [Tese (doutoramento) - FCMBB].
34. MOREIRA, B.W.; VIEIRA, S.; TUMANG, A.J. Análise do número de dentes permanentes irrompidos em escolares de 7 a 13 anos de idade, de ambos os sexos, das raças branca e negra. Rev.Bras.Odont., Rio de Janeiro, 30: 142-6, 1973.
35. NELDER, J.A. The fitting of a generalization of the logistic curve. Biometrics, Washington, 17: 89-110, 1961.
36. \_\_\_\_\_. An alternative form of a generalized logistic equation. Biometrics, Washington, 18: 614-6, 1962.
37. OLIVER, F.R. Methods of estimating the logistic growth function. Appl. Statist., London, 13: 57-66, 1964.
38. \_\_\_\_\_. Aspects of maximum likelihood estimation of the logistic growth function. J.Am.Statist.Ass., Boston, 61: 697-705, 1966.
39. PATTERSON, H.D. A simple method for fitting an asymptotic regression curve. Biometrics, Washington, 12: 323-9, 1956.
40. PEARL, R. & REED, L.J. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. natn. Acad.Sci. U.S.A., Washington, 6: 275-88, 1920.
41. PRESCOTT, R.D. Law of growth in forecasting demand. J.Am.Statist. Ass., Boston, 18: 471-9, 1922. Apud CROXTON, F.E. & COWDEN, D. J. op.cit. ref. 18.
42. RASOR, E.A. The fitting of logistic curve by means of a nomograph. J.Am.Statist.Ass., Boston, 44: 548-53, 1949.

43. REIERSOL, O. Linear and nonlinear multiple comparisons in logit analysis. Biometrika, London, 48: 359-63, 1961.
44. ROOS, C.F. & VON SZELISKI, V. Factors governing changes in domestic automobile demand. In: \_\_\_\_\_. The Dynamics of Automobile Demand, New York, General Motors Corporations, 1939.
45. SCHULTZ, H. The standard error of a forecast from a curve. J.Am. Statist. Ass., Boston, 25: 139-85, 1930.
46. SILVERSTONE, H. Estimating the logistic curve. J.Am.Statist.Ass., Boston, 52: 567-77, 1957.
47. SOWDEN, R.R. On the first-order bias of parameter estimates in a quantal response model under alternative estimation procedures. Biometrika, London, 59: 573-9, 1972.
48. STEVENS, W.L. Asymptotic regression. Biometrics, Washington, 7: 247-67, 1951.
49. TEIXEIRA, D.; NEGREIROS DE PAIVA, C.E.; e ALMEIDA COSTA, D.A. Glândulas salivares: sua importância na sobrevivência de ratos jovens. Ciência e Cultura, São Paulo, 22 (Supl.): 344, 1970. [Resumos da XXII Reunião Anual da SBPC].
50. THEIL, H. Specification and Aggregation analysis. In: \_\_\_\_\_. Principles of Econometrics. New York, John Wiley, 1971. cap. 11, p. 540-602.
51. TINTNER, G. The trend. In: \_\_\_\_\_. Econometrics. New York, John Wiley, 1952. cap. 8, p. 189-215.
52. TOOTILL, J.P.R. Routine least squares estimation from models containing a simple non-linearity. Biometrics, Washington, 19: 118-43, 1963.
53. VERHULST, P.E. Recherches mathématiques sur la loi d'accroissement de la population. Académie de Bruxelles, Bruxelles, 18: 1-38, 1845.

A P Ê N D I C E

Tabela III. Pesagens mensais, em kg, referentes a 10 vacas Holandesas do Posto Zootécnico "Luiz de Queiroz", Piracicaba, SP<sup>1/</sup>. (Os animais estão numerados de 1 a 10; após a numeração seguem o nome do animal e a data do nascimento).

---

nº 1 - HEBE - 27-3-21.

46, 63, 85, 114, 135, 140, 159, 170, 172, 200, 225, 243, 229, 249,  
242, 241, 254, 280, 300, 315, 337, 360, 372, 400, 430, 428, 450, 445,  
435, 420, 405, 420, 480, 432, 440, 420, 480, 484, 460, 460, 460, 470,  
470, 470, 480, 440, 480, 520, 590, 600, 588, 570, 530, 540, 540, 540,  
550.

nº 2 - HUNGRIA - 31-3-21.

39, 51, 62, 82, 97, 98, 107, 122, 118, 142, 166, 180, 192, 201,  
212, 198, 110, 210, 229, 257, 280, 300, 320, 340, 370, 376, 395, 392,  
380, 370, 367, 380, 412, 340, 354, 348, 380, 400, 380, 380, 380, 388,  
400, 400, 410, 460, 460, 470, 510, 440, 480, 420, 400, 420, 420, 420,  
460.

nº 3 - HYGIA - 2-4-21.

34, 52, 75, 100, 117, 124, 145, 177, 164, 175, 220, 230, 237, 253,  
258, 250, 260, 240, 263, 288, 305, 340, 340, 365, 393, 410, 410, 400,  
400, 388, 385, 390, 428, 380, 368, 364, 400, 400, 380, 370, 388, 400,  
400, 388, 396, 440, 440, 440, 500, 480, 470, 440, 470, 500, 510, 520,  
530.

nº 4 - XANTIPA - 11-6-60.

32, 42, 58, 92, 112, 131, 170, 190, 200, 212, 220, 240, 238, 236,  
238, 236, 238, 253, 257, 278, 294, 310, 324, 341, 353, 367, 371, 382,  
385, 390, 392, 395, 409, 427, 423, 453, 471, 476, 478, 497, 493, 500,  
502, 512, 514, 470, 450, 469, 488, 472, 492, 488, 499, 486, 478, 500,  
518, 500, 464.

nº 5 - AVENIDA - 4-9-62

35, 54, 68, 88, 108, 130, 146, 163, 170, 180, 190, 216, 240, 250,  
234, 267, 284, 298, 308, 322, 316, 326, 340, 315, 340, 339, 365, 390,  
400, 394, 424, 417, 414, 460, 457, 450, 470, 478, 492, 506, 516, 528,  
545, 502, 510, 510, 504, 474, 484, 515, 504, 532, 566, 550, 446, 500,  
505.

---

<sup>1/</sup> FONTE: MISCHAN (1972).

Tabela III. Continuação

## nº 6 - ESTÁTUA - 2-4-43.

35, 54, 78, 100, 132, 145, 158, 180, 190, 190, 190, 200, 190, 220,  
 200, 200, 200, 200, 210, 200, 210, 210, 230, 240, 250, 280, 280, 290,  
 290, 310, 300, 340, 370, 380, 410, 410, 340, 340, 360, 350, 370, 380,  
 380, 390, 400, 400, 410, 420, 420, 470, 460, 450, 470, 480, 490, 520,  
 530.

## nº 7 - ZAGA - 23-1-61.

50, 60, 82, 98, 126, 155, 170, 190, 205, 240, 242, 264, 254, 253,  
 249, 251, 270, 280, 294, 294, 300, 302, 325, 346, 356, 376, 395, 400,  
 419, 428, 418, 420, 406, 416, 460, 481, 490, 503, 538, 542, 553, 564,  
 542, 544, 454, 473, 487, 502, 482, 502, 501, 500, 522, 506, 518, 515,  
 518.

## nº 8 - AMIGA - 22-8-62.

44, 62, 78, 98, 116, 138, 154, 176, 186, 179, 180, 200, 222, 236,  
 200, 212, 238, 230, 236, 262, 264, 272, 284, 270, 280, 288, 302, 324,  
 342, 347, 370, 372, 376, 407, 415, 418, 430, 440, 454, 448, 379, 403,  
 415, 390, 416, 408, 402, 413, 406, 425, 440, 464, 480, 494, 499, 452,  
 438.

## nº 9 - BENZINA - 21-10-63.

42, 50, 72, 98, 130, 156, 177, 186, 192, 200, 208, 217, 230, 250,  
 266, 224, 220, 224, 230, 250, 254, 272, 290, 300, 316, 330, 320, 340,  
 350, 368, 380, 402, 414, 410, 410, 446, 450, 476, 496, 514, 456, 430,  
 418, 400, 394, 403, 400, 348, 424, 468, 475, 490, 510, 410, 420, 452,  
 430.

## nº 10 - CARTA - 11-9-64.

34, 40, 49, 65, 107, 130, 149, 171, 172, 200, 202, 204, 204, 205,  
 208, 196, 230, 230, 252, 266, 294, 304, 300, 316, 322, 350, 367, 386,  
 407, 424, 441, 461, 480, 480, 478, 462, 478, 349, 404, 425, 424, 440,  
 448, 406, 476, 463, 465, 450, 442, 450, 456, 498, 528, 476, 454, 500,  
 488.

Tabela VI. Número de dentes permanentes irrompidos em crianças brancas, da cidade de Piracicaba, SP, coletados pela C.D. Ana Helena Marcondes. (O primeiro número, seguido de dois pontos, refere-se à idade em meses completos; o último número, precedido de dois pontos, dá o total de crianças daquela idade).

Sexo Feminino	Sexo Masculino
60: 0, 0 : 2	60: 0 : 1
61: 0, 2, 2 : 3	61: 0, 5, 0, 0, 0, 0 : 6
62: 0, 0, 0, 0, 0 : 5	62: 0, 0, 0, 0, 0, 2, 0 : 7
63: 0, 0, 0, 0, : 4	63: 6, 0, 0, 0, 1 : 5
64: 2, 0, 1, 0, 0, 0, 0, 3, 0, 1 : 10	64: 0, 6, 4, 0, 1 : 5
65: 3 : 1	65: 0, 0, 0, 2, 2 : 5
66: 2, 2, 0, 0 : 4	66: 3, 0, 1, 1, 0, 2, 5, 0 : 8
67: 0, 0, 0, 0 : 4	67: 2, 0, 0 : 3
68: 0, 0, 6, 6, 0 : 5	68: 2 : 1
69: 5, 0, 2, 0 : 4	69: 3, 8, 1, 0, 0 : 5
70: 0 : 1	70: 4, 4, 1, 0 : 4
71: 5, 0, 0, 4 : 4	71: 4, 6, 10, 10 : 4
72: 0, 6, 6, 2 : 4	72: 6, 3, 2, 0, 9, 0 : 6
73: 2, 6 : 2	73: 4, 5, 6, 4, 6, 5 : 6
74: 6, 4, 6 : 3	74: 10 : 1
75: 8, 8, 4, 3 : 4	75: 0, 10, 10, 0 : 4
76: 6, 5, 12, 8, 10, 2, 2, 8 : 8	76: 4, 4, 8 : 3
77: 2, 3, 8, 8 : 4	77: 10, 6, 1 : 3
78: 10, 0, 11, 10, 2, 8, 7 : 7	78: 6, 5, 6, 6 : 4
79: 10, 7, 10, 5 : 4	79: 0, 6, 5, 6, 7, 6 : 6
80: 10, 8, 10, 8, 6, 4 : 6	80: 8, 6, 9, 6, 9 : 5
81: 6, 6 : 2	82: 4 : 1
82: 11 : 1	83: 10 : 1
83: 7, 10, 12, 7, 8, 10, 8, 8, 6 : 9	84: 4, 10 : 2
84: 9, 9, 12 : 3	85: 10, 10 : 2
85: 11, 10, 10, 11, 6, 5 : 6	86: 10, 9, 8, 3, 5 : 5
86: 11, 12 : 2	87: 11, 6, 10, 10, 6 : 5
87: 8, 10 : 2	88: 10, 9, 6 : 3
88: 16, 12 : 2	89: 6, 8, 6, 12, 8, 12, 7 : 7
89: 12, 10, 7 : 3	90: 10, 11, 8, 8, 11 : 5
90: 4, 8, 10, 10, 8 : 5	91: 9, 10, 10, 13, 7, 6, 10, 10, 8, 13 : 10
91: 8, 12, 8, 15, 10, 11, 12, 9 : 8	92: 4, 12, 8, 10, 6, 10, 14, 4 : 8
	93: 9 : 1
	94: 10, 6 : 2
	95: 10, 10, 11, 12, 9 : 5

Tabela VI. Continuação

Sexo Feminino	Sexo Masculino
92: 10, 14, 11, 11, 13, 13, 10, 10 : 8	96: 3, 12, 10, 9, 10, 13 : 6
93: 10, 10, 10, 11, 12 : 5	97: 13, 8, 11, 7, 12, 13, 6 : 7
95: 12, 13 : 2	98: 11, 10 : 2
96: 13 : 1	99: 12, 16, 19 : 3
97: 14, 8, 7, 12 : 4	100: 12, 12, 14, 9, 11, 9, 12, 7, 10 : 9
98: 10 : 1	101: 17, 11, 14, 16, 13 : 5
99: 11, 14, 13, 13 : 4	102: 10, 10, 10, 13, 12 : 5
100: 12, 13, 12 : 3	103: 10, 12, 8, 10, 12 : 5
101: 11, 14, 12, 12, 30, 11, 15, 14, 17, 13 : 10	104: 8, 11, 15, 10, 13, 16 : 6
102: 10, 12, 15, 11, 12, 12 : 6	105: 12, 13, 8, 19, 10 : 5
103: 17, 13, 12, 14, 12, 12, 15, 12, 11, 11, 16, 12 : 12	106: 12 : 1
104: 11, 12, 13, 12, 19, 13 : 6	107: 12, 13, 15, 11, 7 : 5
105: 17, 12, 18, 18, 15 : 5	108: 14, 13, 13 : 3
106: 14, 12, 13, 23, 13, 14 : 6	109: 12, 12, 11, 13, 10 : 5
107: 14, 14, 22, 15, 16 : 5	110: 13, 21, 20, 14 : 4
108: 14 : 1	111: 17, 16, 11, 15, 16 : 5
109: 14 : 1	112: 15, 18, 16, 14, 16 : 5
110: 18 : 1	113: 13, 21 : 2
111: 16, 13, 19, 13 : 4	114: 13, 11, 19, 15 : 4
112: 10, 14, 10, 13, 17, 13 14, 11 : 8	115: 23, 16, 13, 11 : 4
113: 12, 25, 13, 17 : 4	116: 17, 15 : 2
114: 14, 22, 24, 19, 13 : 5	117: 25 : 1
115: 24, 15, 15, 12, 10 : 5	118: 16, 15, 19, 14, 13, 12 : 6
116: 19, 10 : 2	119: 12, 21, 24, 14, 18 : 5
117: 22, 19, 15, 12 : 4	120: 25, 17, 14 : 3
118: 23, 21, 14, 14, 13, 18 : 6	121: 17, 23, 21, 14, 13 : 5
119: 22, 23, 16, 22, 19 : 5	122: 23, 15, 15, 19, 25, 10 : 6
120: 15, 27 : 2	123: 20, 15, 18, 14, 14, 18 : 6
121: 22, 27, 21, 26 : 4	124: 20, 12, 24, 12 : 4
122: 27, 24, 14, 18, 17, 18 : 6	125: 17, 10, 11, 17, 14 : 5
123: 19, 24, 15, 19, 22, 22, 19, 24 : 8	126: 13 : 1
124: 21, 18, 23 : 3	127: 28, 22, 16, 12 : 4
125: 22, 24 : 2	128: 26, 15, 16, 23 : 4
126: 21, 22, 27, 12 : 4	129: 28, 21, 18, 21, 23 : 5
127: 15, 12 : 2	130: 16, 26, 17, 19, 21, 7 : 6
128: 26 : 1	131: 26, 16, 20, 20, 15, 19, 23, 24 : 8
129: 23, 21, 17, 25, 28, 14 : 6	132: 26, 22, 17, 13, 23, 19, 13 : 7
130: 14, 19 : 2	133: 26, 12 : 2
131: 19, 13, 18, 22, 24, 25, 20 : 7	134: 24, 13, 28, 26, 21 : 5
	135: 24, 13, 27, 23, 16 : 5
	136: 14, 27, 12, 18, 22 : 5
	137: 19, 22 : 2
	138: 28, 23 : 2

Tabela VI. Continuação

Sexo Feminino	Sexo Masculino
132: 25, 27, 26, 24, 23 : 5	139: 22 : 1
133: 20, 18, 20, 18, 28, 25 : 6	141: 26, 23 : 2
134: 23, 26, 22, 21 : 4	142: 26, 24, 13, 22, 28, 21 : 6
135: 17, 22, 28 : 3	143: 27, 25, 22, 27, 22, 28, 23 : 7
136: 28 : 1	144: 27, 24, 23, 26, 28 : 5
137: 22, 27, 21 : 3	145: 26, 25, 27, 25, 28, 25 : 6
138: 24, 28, 28, 16 : 4	146: 28, 27, 25, 23, 20, 28, 28, 24, 21 : 9
139: 24, 22, 23, 26 : 4	147: 28, 25, 26, 28 : 4
140: 27, 23, 26, 27, 19, 24 : 6	148: 28, 25, 24 : 3
141: 28 : 1	149: 27, 28, 27 : 3
142: 28, 26, 28, 28, 13, 24 20 : 7	150: 27, 28 : 2
143: 26, 21 : 2	151: 26 : 1
144: 28, 28, 18, 28, 28, 28 27, 24 : 8	152: 28 : 1
145: 24, 23, 28, 24, 28 : 5	153: 28, 28 : 2
146: 28, 27 : 2	154: 28, 28 : 2
147: 24, 28, 28, 26 : 4	155: 26, 27, 28, 28 : 4
148: 28, 28, 26 : 3	
149: 27, 28, 26, 26, 15 : 5	
150: 26, 28, 27, 28, 25 : 5	
151: 27, 26, 28 : 3	
152: 27, 28, 28, 28 : 4	
153: 27, 28, 25, 28, 27 : 5	
154: 23, 26 : 2	
155: 26, 28 : 2	