UNIVERSIDADE DE SÃO PAULO

FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Análise estatística de curvas de crescimento sob o enfoque clássico e Bayesiano: aplicação à dados médicos e biológicos

Breno Raphael Gomes de Oliveira

Ribeirão Preto – SP

2015

Breno Raphael Gomes de Oliveira

Análise estatística de curvas de crescimento sob o enfoque clássico e Bayesiano: aplicação à dados médicos e biológicos

Dissertação de Mestrado apresentada ao Programa Saúde na Comunidade da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, para obtenção do título de mestre.

Área de concentração:
Métodos Quantitativos em Saúde
Orientador:
Prof. Dr. Jorge Alberto Achcar.

Ribeirão Preto – SP

2015

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE

FICHA CATALOGRÁFICA

Oliveira, Breno R. Gomes.

Análise estatística de curvas de crescimento sob o enfoque clássico e Bayesiano: aplicação à dados médicos e biológicos, 2015.

Dissertação de Mestrado apresentada à Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo no Programa Saúde na Comunidade

Orientador: Prof. Dr. Jorge Alberto Achcar

1. Modelos Não-Lineares. 2. Curvas de Crescimento. 3. Bayesiana.

FOLHA DE APROVAÇÃO

Oliveira, Breno R. Gomes.

Análise estatística de curvas de crescimento sob o enfoque clássico e Bayesiano: aplicação à dados médicos e biológicos.

Dissertação de Mestrado apresentada ao Programa Saúde na Comunidade da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, para obtenção do título de mestre.

Aprovado em: ___/___/_____

BANCA EXAMINADORA

Prof. Dr. _____INSTITUIÇÃO:_____

Julgamento _____ASSINATURA:_____

Prof. Dr. _____INSTITUIÇÃO:_____

Julgamento _____ASSINATURA:_____

Prof. Dr. _____INSTITUIÇÃO:_____

Julgamento _____ASSINATURA:_____

DEDICATÓRIA

Ao meu Pai Neilon, minha mãe Cleidmar

e minha irmã Fernanda

AGRADECIMENTOS

RESUMO

Oliveira, Breno R. Gomes. **Análise estatística de curvas de crescimento sob o enfoque clássico e Bayesiano: aplicação à dados médicos e biológicos.** 2015. 79 páginas. Dissertação de mestrado – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2015.

Introdução: A curva de crescimento é um modelo empírico da evolução de uma quantidade ao longo do tempo. As curvas de crescimento são utilizadas em muitas disciplinas , em particular no domínio da estatística, onde há uma grande literatura sobre o assunto relacionado a modelos não lineares. **Método:**No desenvolvimento dessa dissertação de mestrado, foi realizado um estudo baseado em dados de crescimento nas áreas biológica e médica para comparar os dois tipos de inferência (Clássica e Bayesiana), na busca de melhores estimativas e resultados para modelos de regressão não lineares, especialmente considerando alguns modelos de crescimento introduzidos na literatura. No método Bayesiano para a modelagem não linear assumimos erros normais uma suposição usual e também distribuições estáveis para a variável resposta. Estudamos também alguns aspectos de robustez dos modelos de regressão não linear para a presença de outliers ou observações discordantes considerando o uso de distribuições estáveis para a resposta no lugar da suposição de normalidade habitual. **Resultados e Conclusões:** Análise dos dois exemplos pode-se observar melhores ajustes quando utilizada o método Bayesiano de ajustes de modelos não lineares de curvas de crescimento. É bem sabido que, em geral, não há nenhuma forma fechada para a função densidade de probabilidade de distribuições estáveis. No entanto, sob uma abordagem Bayesiana, a utilização de uma variável aleatória latente ou auxiliar proporciona uma simplificação para obter qualquer distribuição a posteriori quando relacionado com distribuições estáveis. Esses resultados poderiam ser de grande interesse para pesquisadores e profissionais, ao lidar com dados não Gauss. Para demonstrar a utilidade dos aspectos computacionais, a metodologia é aplicada a um exemplo relacionado com as curvas de crescimento intra-uterino para prematuros. Resumos a posteriori de interesse são obtidos utilizando métodos MCMC (Markov Chain Monte Carlo) e o software OpenBugs.

Palavras-chave: curvas de crescimento, distribuição estável, análise Bayesiana, métodos de Monte Carlo em Cadeias de Marko

ABSTRACT

Oliveira, R. Breno Gomes. Statistical Analysis of growth curves under the classical and Bayesian approach: application to medical and biological data. 2015. 79 pages. Master's thesis - Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, 2015.

Introduction: The growth curve is an empirical model of the evolution of a quantity over time. Growth curves are used in many disciplines, particularly in the field of statistics, where there is a large literature on the subject related to nonlinear models. **Method:** In the development of this dissertation, a study based on data growth in biological areas and medical was conducted to compare two types of inferences (Classical and Bayesian), in search of better estimates and results for nonlinear regression models, especially considering some growth models introduced in the literature. The Bayesian method for nonlinear modeling assume normal errors an usual assumption and also stable distributions for the response variable. We also study some aspects of robustness of nonlinear regression models for the presence of outliers or discordant observations regarding the use of stable distributions to the response in place of the usual assumption of normality. **Results and Conclusions:** In the analysis of two examples it can be seen best results using Bayesian methodology for non linear models of growth curves. It is well known that, in general, there is no closed form for the probability density function of stable distributions. However, under a Bayesian approach, the use of a latent random variable or auxiliary variable provides a simplification to get every conditional posterior related to stable distributions. These results could be of great interest to researchers and practitioners when dealing with non-Gaussian data. To demonstrate the utility of the computational aspects, the methodology is also applied to an example related to intrauterine growth curves for premature infants. Posterior summaries of interest are obtained using MCMC methods (MCMC) and the OpenBugs software.

Keywords: growth curves, stable distribution, Bayesian analysis, Monte Carlo methods in Markov Chains

LISTA DE FIGURAS

LISTA DE TABELAS

Sumário

1. Introdução **13**
2. Objetivo **25**
3. Metodologia **26**
3.1. Introdução à modelos de curvas de crescimento 26
3.2. Introdução aos métodos Bayesianos 29
3.2.1. Fórmula de Bayes 29
3.2.2. Distribuição a Priori 30
3.2.3. Critério para Discriminação de Modelos – Deviance Information Criterion (DIC). 31
3.2.4. Métodos de simulação para amostras da distribuição a posteriori 31
4. Um exemplo de dados de crescimento em estudos ambientais **34**
5. Análise dos dados da Tabela 2 ..... **36**
6. Uso de uma distribuição estável para modelos de crescimento **42**
6.1. Uma análise Bayesiana assumindo a distribuição estável 44
7. Análise Bayesiana dos dados das Tabelas 1a e 1b assumindo uma distribuição estável.. **47**
7.1. Modelo de Gompertz com erros normais (procedimento clássico) 48
7.2. Modelo de Gomperz com erros normais (procedimento Bayesiano) 48
7.3. Modelo de Gomperz com distribuição estável (procedimento Bayesiano) 49
7.4. Modelo de Gomperz com distribuição estável na presença de outliers ou observações discordantes (procedimento Bayesiano) 50
8. Conclusões **52**
9. Referências Bibliográficas **53**
Apêndice 1 - Gráficos de resíduos para os modelos de crescimento ajustados (dados da Tabela 1) **59**
Apêndice 2 - Programas OpenBugs (dados da Tabela 1) **63**

1. Introdução

O uso de métodos quantitativos é essencial para a área de saúde. Nessa direção novos modelos estatísticos são considerados e introduzidos na literatura para analisar dados médicos. Entre as várias técnicas usadas na análise de dados médicos, a análise de regressão é uma das técnicas estatísticas dentro da área de métodos quantitativos mais utilizadas para análise de dados em geral, pois a partir de modelos de regressão se investiga e se modela relações entre várias variáveis (ver por exemplo, Freedman et al. 2004, Lyles & Kupper 1997, Chen & Wang 2004).

No estudo de análise de regressão, utiliza-se de modelos matemáticos que relacionam o comportamento de uma variável Y com outra X. Quando a função f que relaciona duas variáveis é do tipo $f(x) = a + bx$ temos um modelo de regressão linear simples. A variável X é uma variável independente da equação, enquanto $Y = f(x)$ é a variável dependente relacionada com X. O modelo de regressão é chamado modelo de regressão linear simples quando envolve uma relação causal entre duas variáveis. O modelo de regressão é chamado modelo de regressão múltiplo quando envolve uma relação causal com mais de duas variáveis, sendo uma das técnicas estatísticas mais amplamente empregada para se obter previsões de interesse.

Modelos de regressão linear são amplamente utilizados em ciências médicas, biológicas, econômicas, sociais e da engenharia para descrever possíveis relações entre as variáveis. É considerado como um dos mais importantes instrumentos utilizados nestas áreas.

Na análise de regressão linear, os dados são modelados utilizando-se funções de previsão linear, e os parâmetros do modelo são desconhecidos e estimados a partir dos dados. Geralmente, refere-se a regressão linear para um modelo em que a média condicional de Y dado o valor de X é uma função de X.

Modelos de regressão linear tem muitos usos práticos. A maioria das aplicações se enquadram em uma das seguintes duas grandes categorias:

- Se o objetivo é a predição ou previsão, modelos de regressão linear podem ser utilizados para ajustar um modelo preditivo para um conjunto de dados observados de valores Y e X. Após o ajuste de um modelo deste tipo, se um

valor adicional de X é dado, então, o modelo ajustado pode ser usado para fazer uma previsão do valor de Y.

- Dadas as variáveis ou covariáveis $X_1, \dots, X_p$, que podem estar relacionadas com a resposta ou variável dependente Y, a análise de regressão linear, pode ser utilizada para quantificar a magnitude da relação entre Y e X_j, $j = 1, \dots, p$. Isso é dado por testes de hipóteses nos parâmetros de regressão.

Modelos de regressão linear são freqüentemente ajustados usando-se a abordagem de mínimos quadrados. Na análise de regressão linear verifica-se o efeito conjunto das covariáveis na resposta Y (ver, por exemplo, Draper e Smith, 1981; Seber e Lee, 2003; ou Montgomery e Runger, 2011). Um exemplo de modelo de regressão com duas variáveis independentes é dado por:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \tag{1}$$

para $i = 1, \dots, n$ e ε_i , é um termo relacionado ao erro (variável não-observada) suposto como uma quantidade aleatória. Este erro aleatório inclui todos os outros fatores que poderiam influenciar a variável dependente Y não incluídos no modelo de regressão.

Em geral, os modelos de regressão linear, são representados pela relação:

$$Y = \mu + \varepsilon, \tag{2}$$

em que, $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^t$ é um vetor de variáveis aleatórias e um vetor de médias $\boldsymbol{\mu}$ dado por:

$$\boldsymbol{\mu} = X\boldsymbol{\beta}, \tag{3}$$

onde, X é uma matriz de variáveis explanatórias e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)$ é um vetor de parâmetros de regressão. Nesta relação tem-se que vetor de erros $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^t$ segue uma distribuição Normal Multivariada, isto é, $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 I_n)$, onde I_n é uma matriz identidade, sendo assim, temos um vetor de erros normais, com média zero e variância constante.

Em muitas áreas de aplicação podemos ter relações não-lineares entre as variáveis de interesse, sendo necessário a modelagem não-linear para relacionar as variáveis.

As curvas de crescimento são incluídas em uma classe de modelos não-lineares utilizados na biologia para modelar diferentes problemas como o tamanho da população ou de biomassa (em ecologia populacional e demografia, para análise de crescimento da

população) ou a altura do corpo individual ou de biomassa (em fisiologia, para análise de crescimento de indivíduos). A curva de crescimento é um modelo empírico da evolução de uma quantidade ao longo do tempo. As curvas de crescimento são utilizadas em muitas disciplinas além de biologia, em particular no domínio da estatística, onde há uma grande literatura sobre o assunto relacionado a modelos não lineares. As curvas de crescimento são muitas vezes modeladas como sendo um processo estocástico contínuo.

Na área médica muitos estudos com curvas de crescimento são apresentados na literatura. Alguns exemplos que motivam o presente estudo:

- Curvas de crescimento individuais foram usadas para testar se o desenvolvimento de crianças com dificuldades de leitura é melhor caracterizada por modelos de atraso de desenvolvimento ou déficit de desenvolvimento (Francis et al, 1996).
- As curvas de crescimento intra-uterino padrão disponíveis com base em pesos de recém-nascidos podem subestimar o crescimento fetal no período pré-termo. Curvas de crescimento são apresentadas com base em dados de quatro centros escandinavos para 759 pesos fetais estimados com ultrasom em 86 gestações sem complicações (Maršál, K.,1996).
- Curvas de crescimento padrão para acondroplasia, a forma mais comum de nanismo, foram construídas com base em medições de altura, velocidade de crescimento, o segmento superior e inferior, e perímetro cefálico em 400 anões achondroplasticos. Estas curvas padrão fornecem a base para determinar o crescimento normal nestes indivíduos, para auxiliar na determinação de desordens sobrepostas, e para avaliar qualquer crescimento acelerado terapia (Horton,1978).
- Um estudo foi feito para criar e validar nova forma de calcular o peso intra-uterino, comprimento e curvas de crescimento do perímetro da cabeça usando uma grande amostra, relacionada à diversidade racial dos Estados Unidos e para comparar com as curvas de Lubchenco (Olsen et al,2010).
- Com base nas suas respostas a um questionário auto-administrado, 697 indivíduos saudáveis não-fumantes foram escolhidos a partir de uma amostra aleatória representativa da população não-mexicano-americano branco de Tucson, Arizona, incluídos em um estudo longitudinal da saúde respiratória. Para cada tema, o primeiro conjunto satisfatório de dados de

fluxo de volume obtidos durante os primeiros três inquéritos consecutivos foi selecionado para análise. Para a capacidade vital forçada (CVF) e volume expiratório forçado no primeiro segundo (FEV1), o melhor valor único para cada indivíduo foi selecionado. Outras medições de fluxo-volume foram obtidos a partir do teste único com a melhor soma FEV, além de FVC. Estes dados foram utilizados para obter melhores equações de predição para cada sexo, por faixa etária para 5 variáveis da espirometria e fluxo-volume (Knudson et al, 1983).

- A Organização Mundial de Saúde (OMS), em colaboração com uma série de instituições de pesquisa em todo o mundo, está desenvolvendo novos padrões de crescimento infantil. Como parte de um amplo processo de consulta para selecionar os melhores métodos estatísticos, a OMS convocou um grupo de estatísticos e especialistas em crescimento infantil para rever os métodos disponíveis, desenvolver uma estratégia para avaliar os seus pontos fortes e fracos, e discutir questões metodológicas que possam ser enfrentadas no processo de construir novas curvas de crescimento. Para selecionar o método a ser utilizado, o grupo propôs um de dois estágios para processo de tomada de decisão. Em primeiro lugar, para selecionar alguns métodos relevantes com base em uma lista de critérios definidos e, por outro, para comparar os métodos que utilizam testes disponíveis ou outros procedimentos estabelecidos. O grupo analisou 30 métodos de curvas de crescimento alcançados. Utilizando os critérios pré-definidos, alguns foram selecionados combinando cinco distribuições e duas técnicas de alisamento. Como o número de métodos selecionados foi considerado demasiado grande para ser totalmente testado, um estudo preliminar foi recomendado para avaliar a qualidade do ajuste dos cinco distribuições. Métodos baseados em distribuições com mau desempenho foram eliminados e os métodos restantes foram totalmente testados e comparados (Borghi et al, 2006).
- Curvas brutas de crescimento para a função motora de crianças com síndrome de Down (SD) e estimação da probabilidade de que as funções motoras são alcançadas por diferentes idades. Desenho: análise da curva de crescimento não-linear com um ou 2 parâmetros (taxa, limite superior)

modelo. Ambiente: programas de intervenção precoce, escolas e orfanatos. Participantes: Cento e vinte e uma crianças com SD, com idades entre um mês a 6 anos. Principais desfechos: medida da função motor (Gross Motor Function Measure - GMFM) e severidade da deficiência motora. Resultados: as curvas para crianças com síndrome leve (n = 51) e moderada / grave (n = 70) foram caracterizados por um aumento maior na pontuação GMFM durante a infância e aumentos menores como as crianças que se aproximaram da pontuação máxima prevista de 85,9 ou 87,9. A probabilidade estimada de que uma criança iria rolar em torno de 6 meses foi de 51%; sentar-se em torno de 12 meses, 78%; engatinhar em torno de 18 meses, 34%; andar, em torno de 24 meses, 40%; e correr, subir escadas e pular em torno de cinco anos, de 45% para 52%. Conclusões: as crianças com SD necessitam de mais tempo para aprender os movimentos de acordo com o aumento da complexidade. A gravidade da deficiência afetou a taxa mas não o limite superior da função motora. Os resultados têm implicações para aconselhamento dos pais, a tomada de decisões sobre as intervenções de motor, e antecipando o prazo para a realização das funções motoras. © 2001 pelo Congresso Americano de Medicina de Reabilitação e da Academia Americana de Medicina Física e de Reabilitação (Palisano et al, 2001).

- A análise dos dados de crescimento retrospectiva de 72 pacientes com displasia diastrophic, 62 pacientes com displasia congênita spondyloepiphyseal e 61 pacientes com pseudoachondroplasia permitiu o estabelecimento de padrões de referência de crescimento preliminares para a altura para cada transtorno. Além disso, uma comparação das curvas para altura média para estes três distúrbios e acondroplasia dissimilaridades reveladas nos padrões de crescimento (Horton et al, 1982).

Muitos outros estudos também são observados na literatura (ver por exemplo, Nicosia e Ottinetti A ,1990; Nishida ,H.; Sakamoto,S., Sakanoue, M., 1985; Chou et al 1998).

Um exemplo de dados de crescimento em medicina é dado nas Tabelas 1a e 1b.Informações para os dados das Tabelas 1a e 1b (Fonte:Olsen et al, 2010): O objetivo

deste estudo foi o de criar e validar novo peso intra-uterino, comprimento e perímetro cefálico curvas de crescimento, utilizando uma amostra contemporânea grande dos EUA. Dados de crianças recém-nascidas (Pediatrix Medical Group) com idades entre 22 a 42 semanas de gestação no nascimento em 248 hospitais de 33 estados norte-americanos (1998-2006) para o peso ao nascer, comprimento, perímetro cefálico, a idade gestacional estimada e gênero (GA=idade gestacional; HC=circunferência da cabeça).

Na Figura 1a, temos as curvas de crescimento relacionadas aos dados da Tabela 1a. Na Figura 1b, temos as curvas de crescimento relacionadas aos dados da Tabela 1b.

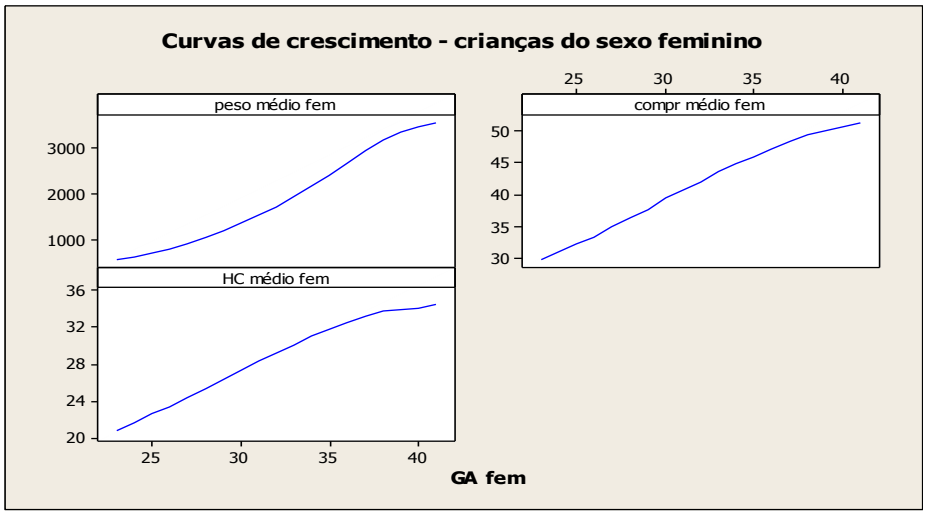


Figura 1a: Curvas de crescimento-crianças do sexo feminino (dados dos EUA)

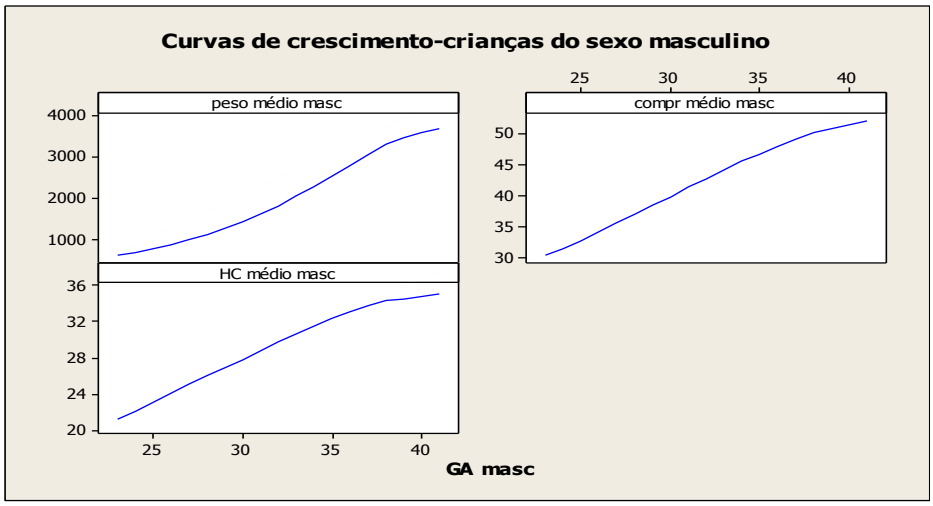


Figura 1b: Curvas de crescimento-crianças do sexo masculino (dados dos EUA)

Tabela 1a: Pesos,comprimentos e circunferência da cabeça por tempo de gestação de crianças de sexo feminino nos EUA (n=55445) - (Fonte:Olsen et al,2010)

GA, wk	*n*	Birth Size		Percentile						
		Mean	SD	3rd	10th	25th	50th	75th	90th	97th
Weight, g										
23	133	587	80	NA[a]	477	528	584	639	687	NA[a]
24	438	649	89	464	524	585	651	715	772	828
25	603	738	121	511	584	657	737	816	885	953
26	773	822	143	558	645	732	827	921	1004	1085
27	966	934	168	615	719	822	936	1047	1147	1244
28	1187	1058	203	686	807	928	1061	1193	1310	1425
29	1254	1199	226	778	915	1052	1204	1354	1489	1621
30	1606	1376	246	902	1052	1204	1373	1542	1693	1842
31	2044	1548	271	1033	1196	1361	1546	1731	1897	2062
32	3007	1730	300	1177	1352	1530	1731	1933	2116	2297
33	4186	1960	328	1356	1545	1738	1956	2178	2379	2580
34	5936	2194	357	1523	1730	1944	2187	2434	2661	2888
35	5082	2420	440	1626	1869	2123	2413	2711	2985	3261
36	4690	2675	514	1745	2028	2324	2664	3015	3339	3667
37	4372	2946	551	1958	2260	2575	2937	3308	3651	3997
38	5755	3184	512	2235	2526	2829	3173	3525	3847	4172
39	5978	3342	489	2445	2724	3012	3338	3670	3973	4276
40	5529	3461	465	2581	2855	3136	3454	3776	4070	4363
41	1906	3546	477	2660	2933	3214	3530	3851	4142	4433
Length, cm										
23	133	29.9	1.8	NA[a]	27.7	28.7	29.9	31.0	31.9	NA[a]
24	438	31.0	1.7	27.5	28.7	29.8	31.1	32.3	33.3	34.3
25	603	32.3	2.0	28.3	29.7	31.0	32.3	33.6	34.8	35.9
26	773	33.4	2.2	29.2	30.7	32.1	33.6	35.1	36.3	37.4
27	966	35.0	2.3	30.2	31.9	33.4	35.0	36.6	37.9	39.1
28	1187	36.4	2.5	31.4	33.1	34.8	36.5	38.1	39.5	40.8
29	1254	37.8	2.7	32.8	34.6	36.3	38.0	39.7	41.2	42.5
30	1606	39.6	2.6	34.3	36.0	37.7	39.5	41.3	42.7	44.1
31	2044	40.9	2.6	35.7	37.5	39.2	41.0	42.7	44.1	45.5
32	3007	42.1	2.6	37.1	38.9	40.6	42.3	44.0	45.5	46.9
33	4186	43.7	2.6	38.6	40.3	41.9	43.7	45.4	46.9	48.3
34	5936	45.0	2.6	39.8	41.5	43.2	45.0	46.7	48.2	49.7
35	5082	46.0	2.7	40.9	42.6	44.3	46.2	48.0	49.5	51.0
36	4690	47.2	2.8	42.0	43.7	45.5	47.4	49.2	50.8	52.3
37	4372	48.4	2.8	43.2	44.9	46.6	48.5	50.3	51.9	53.4
38	5755	49.5	2.6	44.4	46.1	47.7	49.5	51.2	52.7	54.2
39	5978	50.1	2.5	45.3	46.9	48.5	50.2	51.9	53.3	54.7
40	5529	50.7	2.4	46.1	47.6	49.1	50.8	52.4	53.8	55.1
41	1906	51.3	2.4	46.7	48.2	49.7	51.3	52.8	54.2	55.5
HC, cm										
23	133	20.8	1.2	NA[a]	19.5	20.1	20.9	21.6	22.2	NA[a]
24	438	21.7	1.1	19.6	20.3	21.0	21.8	22.5	23.2	23.8
25	603	22.7	1.2	20.4	21.1	21.9	22.7	23.4	24.1	24.8
26	773	23.5	1.2	21.2	22.0	22.7	23.6	24.4	25.1	25.9
27	966	24.5	1.3	21.9	22.8	23.6	24.5	25.4	26.2	27.0
28	1187	25.5	1.5	22.7	23.7	24.6	25.5	26.5	27.3	28.1
29	1254	26.5	1.5	23.6	24.6	25.5	26.5	27.5	28.4	29.2
30	1606	27.5	1.5	24.6	25.6	26.5	27.5	28.5	29.4	30.2
31	2044	28.4	1.5	25.5	26.5	27.4	28.4	29.4	30.3	31.1
32	3007	29.3	1.5	26.5	27.4	28.3	29.3	30.3	31.2	32.0
33	4186	30.2	1.5	27.3	28.3	29.2	30.2	31.2	32.1	33.0
34	5936	31.1	1.6	28.1	29.1	30.1	31.1	32.2	33.1	34.0
35	5082	31.9	1.6	28.8	29.8	30.8	31.9	33.0	34.0	34.9
36	4690	32.6	1.7	29.4	30.5	31.5	32.7	33.8	34.8	35.8
37	4372	33.3	1.7	30.1	31.1	32.2	33.3	34.4	35.4	36.3
38	5755	33.8	1.6	30.7	31.7	32.7	33.7	34.8	35.7	36.7
39	5978	34.0	1.5	31.1	32.0	33.0	34.0	35.1	36.0	36.9
40	5529	34.2	1.5	31.4	32.3	33.3	34.3	35.3	36.1	37.0
41	1906	34.5	1.5	31.7	32.6	33.5	34.5	35.5	36.3	37.1

Tabela 1b: Pesos,comprimentos e circunferência da cabeça por tempo de gestação de crianças de sexo masculino nos EUA (n=73995) - (Fonte:Olsen et al,2010)

GA, wk	*n*	Birth Size		Percentile						
		Mean	SD	3rd	10th	25th	50th	75th	90th	97th
Weight, g										
23	153	622	74	NA[a]	509	563	621	677	727	NA[a]
24	451	689	96	497	561	623	690	756	813	869
25	722	777	116	550	626	700	780	857	926	992
26	881	888	145	613	704	794	890	983	1065	1145
27	1030	1001	170	680	789	895	1009	1120	1218	1312
28	1281	1138	203	758	884	1007	1141	1271	1385	1496
29	1505	1277	218	845	988	1128	1280	1429	1560	1688
30	1992	1435	261	955	1114	1272	1443	1612	1761	1906
31	2460	1633	275	1093	1267	1441	1631	1818	1984	2147
32	3677	1823	306	1246	1433	1622	1829	2034	2218	2398
33	5014	2058	341	1422	1625	1830	2057	2284	2488	2688
34	7291	2288	364	1589	1810	2035	2285	2536	2763	2987
35	6952	2529	433	1728	1980	2238	2527	2819	3084	3348
36	7011	2798	498	1886	2170	2462	2792	3127	3432	3737
37	6692	3058	518	2103	2401	2708	3056	3411	3736	4060
38	8786	3319	527	2356	2652	2959	3306	3661	3986	4312
39	8324	3476	498	2545	2833	3131	3469	3813	4129	4446
40	7235	3582	493	2666	2950	3245	3579	3919	4232	4545
41	2538	3691	518	2755	3039	3333	3666	4007	4319	4633
Length, cm										
23	153	30.5	1.6	NA[a]	28.0	29.1	30.3	31.4	32.4	NA[a]
24	451	31.5	1.8	27.9	29.1	30.3	31.5	32.8	33.9	34.9
25	722	32.7	2.1	28.8	30.2	31.5	32.9	34.2	35.4	36.5
26	881	34.2	2.2	29.9	31.3	32.8	34.3	35.7	37.0	38.2
27	1030	35.6	2.4	31.0	32.6	34.1	35.7	37.3	38.6	39.8
28	1281	37.2	2.5	32.2	33.9	35.5	37.2	38.8	40.2	41.5
29	1505	38.6	2.5	33.5	35.2	36.9	38.7	40.3	41.7	43.1
30	1992	39.9	2.8	34.8	36.6	38.3	40.1	41.8	43.2	44.6
31	2460	41.5	2.5	36.2	38.0	39.8	41.6	43.3	44.7	46.1
32	3677	42.8	2.7	37.7	39.5	41.2	43.0	44.7	46.1	47.5
33	5014	44.3	2.6	39.1	40.9	42.6	44.4	46.1	47.5	48.9
34	7291	45.6	2.6	40.4	42.2	43.9	45.7	47.4	48.9	50.3
35	6952	46.8	2.7	41.5	43.3	45.0	46.9	48.6	50.2	51.6
36	7011	48.0	2.8	42.7	44.5	46.2	48.1	49.9	51.5	53.0
37	6692	49.2	2.7	44.0	45.7	47.4	49.3	51.1	52.6	54.1
38	8786	50.2	2.7	45.2	46.8	48.5	50.2	52.0	53.5	55.0
39	8324	51.0	2.4	46.1	47.7	49.3	51.0	52.7	54.2	55.6
40	7235	51.6	2.4	46.9	48.4	49.9	51.6	53.2	54.7	56.1
41	2538	52.1	2.4	47.5	49.0	50.5	52.1	53.7	55.1	56.5
HC, cm										
23	153	21.3	1.0	NA[a]	20.0	20.6	21.3	22.0	22.7	NA[a]
24[b]	451	22.2	1.1	20.1	20.8	21.5	22.2	23.0	23.6	24.3
25	722	23.1	1.1	20.9	21.7	22.4	23.2	23.9	24.6	25.3
26	881	24.1	1.3	21.8	22.5	23.3	24.2	25.0	25.7	26.4
27	1030	25.2	1.3	22.6	23.5	24.3	25.2	26.0	26.8	27.6
28	1281	26.1	1.4	23.5	24.3	25.2	26.1	27.1	27.9	28.6
29	1505	27.0	1.4	24.3	25.2	26.1	27.1	28.0	28.8	29.6
30	1992	27.9	1.5	25.1	26.1	27.0	28.0	29.0	29.8	30.6
31	2460	28.9	1.5	26.0	27.0	27.9	28.9	29.9	30.8	31.6
32	3677	29.8	1.5	26.9	27.8	28.8	29.9	30.9	31.8	32.6
33	5014	30.7	1.6	27.7	28.7	29.7	30.8	31.8	32.7	33.6
34	7291	31.6	1.6	28.5	29.5	30.5	31.6	32.7	33.6	34.6
35	6952	32.4	1.6	29.2	30.3	31.3	32.4	33.6	34.5	35.5
36	7011	33.2	1.7	29.9	31.0	32.1	33.2	34.3	35.3	36.3
37	6692	33.8	1.7	30.6	31.7	32.7	33.9	35.0	36.0	36.9
38	8786	34.4	1.7	31.2	32.2	33.2	34.4	35.5	36.4	37.3
39	8324	34.6	1.6	31.5	32.5	33.5	34.6	35.7	36.6	37.6
40	7235	34.8	1.5	31.8	32.8	33.8	34.8	35.9	36.8	37.7
41	2538	35.1	1.5	32.0	33.0	34.0	35.0	36.1	37.0	37.8

Com essas considerações observa-se que há um grande interesse em estudar modelos de crescimento em medicina sob diferentes enfoques da inferência estatística. O objetivo principal desse estudo é analisar um ou mais conjuntos de dados médicos de crescimento sob diferentes procedimentos de inferência estatística. Como esses modelos são casos especiais de modelos de regressão não lineares, apresentamos a seguir alguns conceitos básicos sobre modelos não-lineares.

1.1 Alguns conceitos básicos em modelos de regressão não-lineares

Métodos de inferência clássica padrão em geral são usadas para se obter estimativas pontuais ou por intervalo para os parâmetros dos modelos de curvas de crescimento sob a metodologia de modelagem não-linear. A metodologia de regressão não-linear é similar à metodologia de regressão linear, isto é, uma abordagem de modelagem para relacionar uma resposta Y para um vetor de covariáveis, $\boldsymbol{v} = (v_1, \dots, v_k)$. Diferente da metodologia de regressão linear frequentemente utilizada para a construção de um modelo puramente empírico, a metodologia de regressão não linear ocorre geralmente quando há alguma motivação ou interpretação física o que implica que a relação entre a resposta e os preditores segue uma forma funcional particular. Um modelo de regressão não linear tem a forma geral,

$$Y_i = f(v_i, \theta) + \varepsilon_i, \tag{4}$$

onde Y_i são as respostas, $i = 1, \dots, n$; f é uma função conhecida do vetor de covariáveis; $\boldsymbol{v} = (v_1, \dots, v_k)$ é um vetor de k covariáveis ou variáveis independentes; $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ é o vetor de parâmetros e ε_i são erros aleatórios. Os erros ε_i são geralmente assumidos como não correlacionadas com a distribuição normal com média zero e variância constante. O critério mais popular para estimar o vetor $\boldsymbol{\theta}$ de parâmetros do modelo não-linear (4) é encontrar estimativas para os parâmetros (mínimos quadrados não-lineares) que minimiza a soma dos quadrados dos erros, dado por,

$$\sum_{i=1}^{n} [y_i - f(v_i, \boldsymbol{\theta})]^2 \tag{5}$$

Observação: Se os erros ε_i seguem uma distribuição normal, então o estimador de mínimos quadrados para $\boldsymbol{\theta}$ também é o estimador de máxima verossimilhança.

Normalmente, as estimativas de regressão não linear devem ser calculadas por meio de processos iterativos de otimização usando métodos para minimizar a soma dos quadrados dos erros na expressão (5). É importante salientar que a definição de não-linearidade é relacionada com os parâmetros desconhecidos e não como a relação entre as covariáveis e a resposta. Como um exemplo, $Y = \beta_0 + \beta_1 v + \beta_2 v^2 + \varepsilon$, é considerada como um modelo linear (ver, por exemplo, Bates e Watts, 1988; Ratkowsky, 1983; Seber e Wild, 1989).

Uma técnica iterativa popular para encontrar os estimadores de mínimos quadrados para modelos não lineares é o algoritmo de Gauss-Newton. O algoritmo de Gauss-Newton incrementa a estimativa trabalhando θ a cada iteração de um montante igual ao dos coeficientes da regressão linear da corrente de resíduos e sobre X.

Se a função f em (5) é continuamente diferenciável em θ, então ele pode ser linearizado como localmente,

$$f(v,\theta) = f(v,\theta_0) + V_0(\theta - \theta_0) \tag{6}$$

onde V_0 é a matriz de gradiente nxp com elementos $\partial f(v_i,\theta_0)/\partial\theta_j$ e θ_0 é um vetor de valores iniciais para o processo iterativo. Isto leva ao algoritmo de Gauss-Newton, para estimar θ,

$$\theta_1 = \theta_1 + (V_0'V_0)^{-1}V_0'e \tag{7}$$

onde e é o vetor de resíduos de trabalho $y_i - f(v_i,\theta_0)$.

Se os erros ε_i são independentes e normalmente distribuídos $N\sim(0,\sigma^2)$, então o algoritmo de Gauss-Newton, é uma aplicação do método escore de Fisher. Este algoritmo é implementado em muitos softwares estatísticos existentes como, por exemplo, R, MINITAB versão 16 ou SAS. Se X é uma matriz de posto completo em uma vizinhança da solução de mínimos quadrados, pode-se mostrar que o algoritmo de Gauss-Newton irá convergir para a solução correta usando bom valor inicial. Em trabalhos práticos, não há

qualquer garantia, no entanto, de que o algoritmo irá convergir a partir de valores distantes da solução. Algumas melhorias do algoritmo de Gauss-Newton, são dadas na literatura como o algoritmo de Levenberg-Marquart (ver, por exemplo, Seber e Wild, 1989). Inferências para os parâmetros dos modelos não lineares são obtidas a partir da normalidade assintótica dos estimadores de mínimos quadrados para $\boldsymbol{\theta}$ com média $\boldsymbol{\theta}$ e matriz de variância-covariância $\sigma^2\left((V'V)\right)^{-1}$, onde a variância σ^2 é geralmente estimada por,

$$s^2 = \left\{\sum_{i=1}^{n}\left[y_i - f\left(v_i, \hat{\theta}\right)^2\right]\right\} \Big/ (n-p) \tag{8}$$

É importante salientar que, como a inferência assintótica para os modelos de regressão não linear é baseada com uma analogia com modelos lineares, e desde que essa inferência é apenas aproximada e como o modelo atual difere de um modelo linear, foram propostas várias medidas de não-linearidade na literatura para verificar o quão boa são as aproximações lineares em cada caso. Uma classe de medidas usa curvaturas (curvaturas intrínsecas) da função f e baseia-se em termos de dimensão das segundas derivadas de f (ver, por exemplo, Bates e Watts, 1980). Outra classe de medidas: curvatura intrínseca definida pelo resíduos $u_i = y_i - f(v_i, \boldsymbol{\theta})$ ou curvaturas efeito de parâmetros (ver, por exemplo, Bates e Watts, 1988).

Em muitas situações nos deparamos com conjunto de dados em que as variáveis de interesse não seguem um padrão de variabilidade homogênea, por exemplo, em estudos com curvas de crescimento, havendo necessidade de trabalharmos com modelos regressão em que a variabilidade é heterogênea, também chamado de modelos de regressão heterocedásticos.

Se temos heterogeneidade de variância (ver por exemplo, Cepeda et al, 2009) , é conveniente assumir uma modelagem da variância relacionando com outras variáveis explanatórias, da seguinte forma:

$$\sigma_i^2 = g(z_i, \gamma), \tag{9}$$

onde, $\boldsymbol{z_i} = (z_{i1}, z_{i2}, \dots, z_{in})'$; $\boldsymbol{\gamma} = (\gamma_i, \dots, \gamma_r)'$ é um vetor de parâmetros de regressão e g uma função não linear apropriada. Assim, um modelo normal não linear com variância heterogênea possui dois importantes componentes, média caracterizada por uma função não linear e variância especifica para cada observação.

Uma boa alternativa para se obter inferências mais precisas e previsões para modelos não lineares é o uso de métodos bayesianos especialmente usando métodos MCMC (Monte Carlo em Cadeias de Markov) como o algoritmo de Gibbs sampling (ver, por exemplo, Gelfand e Smith, 1990; ou Casella e George, 1992), ou o algoritmo de Metropolis-Hastings (ver, por exemplo, Chib e Greenberg, 1995).

Na abordagem Bayesiana, a informação que se tem do parâmetro de interesse θ é de extrema importância na obtenção de boas inferências. Assim, sob o contexto Bayesiano, o parâmetro θ é uma quantidade aleatória no qual existem diferentes graus de conhecimento que variam de problema para problema e de investigador para investigado.

Além da possibilidade de incorporação de informações a priori, outras vantagens relacionadas com a inferência Bayesiana são caracterizadas, como a ausência de pressuposições quanto aos modelos utilizados e pela facilidade da adoção da estimação por intervalo (ver por exemplo, Amaral, 2008). A abordagem Bayesiana quase sempre leva a inferências mais acuradas do que a observada na inferência clássica.

Portanto, uma análise Bayesiana dos modelos não lineares e principalmente em modelos de curvas de crescimento pode ser de grande interesse e de utilidade prática.

2. Objetivo

Desenvolver um estudo baseado em dados de crescimento médicos para comparar os dois tipos de inferência (Clássica e Bayesiana), na busca de melhores estimativas e resultados para modelos de regressão não lineares, especialmente considerando alguns modelos de crescimento introduzidos na literatura. Além disso, também vamos considerar dados longitudinais e modelos com variância heterogênea e também distribuições estáveis para a variável resposta, para curvas de crescimento.

3. Metodologia

3.1. Introdução à modelos de curvas de crescimento

Em biologia, geralmente o interesse está associado à descrição do crescimento de animais, plantas e até mesmo de bactérias na tentativa de entender seus mecanismos essenciais. Dessa forma, a análise de modelos não-lineares é de fundamental importância na produção animal, pois essas análises incluem as situações em que as unidades experimentais ou indivíduos, de diferentes subpopulações ou tratamentos (sexo, raça, entre outros), são analisados ao longo de diversas condições de avaliação (tempo, doses etc). Entre essas análises, destacam-se as curvas de crescimento na produção animal, que relacionam os pesos (y) e as idades (t) dos animais, modelados por meio de modelos não-lineares (Davidian e Giltinan, 1996; Paz, 2004).

Entre as várias aplicações das curvas de crescimento na produção animal, destacamos: sumarizar em modelos com três ou quatro parâmetros, as características de crescimento da população, onde alguns parâmetros dos modelos não-lineares possuem interpretabilidade biológica; avaliar o perfil de respostas de tratamentos ao longo tempo; estudar as interações de respostas das subpopulações ou tratamentos com o tempo; identificar em uma população os animais mais pesados em idades mais jovens; essas informações podem ser obtidas investigando- se o relacionamento para algum parâmetro das curvas de crescimento (pode ser a média) com o tempo, que expressam a taxa de declínio na taxa de crescimento relativa, e o peso limite do animal ou peso assintótico (Sandland & Mcgilchrist, 1979; Draper & Smith, 1980; Davidian & Giltinam, 1996). Essas aplicações são de interesse geral nos estudos de curvas de crescimento.

Na agricultura, são óbvias as vantagens econômicas e administrativas em saber como a produção cresce, o quão rápido elas crescem, e como alguns fatores relacionados às condições do meio ambiente ou do tratamento afetam o crescimento. Como uma ilustração desse tipo de estudo podemos citar o artigo de Silva (1010) onde os autores usam dados simulados e dados reais de animais Nelore, para a estimação dos componentes de covariância dos parâmetros do modelo de crescimento de Von Bertalanffy, por meio da metodologia bayesiana hierárquica. Com base nos componentes estimados, foram encontradas as herdabilidades para cada parâmetro do modelo e as correlações genéticas e ambientais entre esses parâmetros.

O crescimento infantil é um tema de grande interesse na medicina, assim como o crescimento de tumores e os efeitos dos tratamentos sobre tais crescimentos. Em estudos de curvas de crescimento de recém-nascidos, é natural que o peso ou o comprimento tenha uma variabilidade pequena após o nascimento, mas esta variabilidade tende a crescer conforme o tempo, mostrando que modelos homocedásticos nos levam a resultados com subestimação do verdadeiro valor do peso do recém-nascido, havendo assim, a necessidade de outras metodologias como a estimação de parâmetros de modelos heterecedásticos (Souza, 2011).

Em outra área de aplicação, os químicos utilizam modelos de crescimento na formulação de uma reação química em relação ao tempo. Estudos de cinética química visam compreender as relações existentes entre as velocidades das reações (tempo) e os fatores que as influenciam, tais como concentrações de substrato, presença ou não de catalisadores, temperatura, pH, estado físico dos reagentes, dentre outros. A relação entre a velocidade de reação e a concentração de substrato é frequentemente de interesse particular. Essa relação é representada por uma equação matemática, o modelo cinético, geralmente não linear, no qual geralmente utiliza-se modelo de curva de crescimento de Michaelis-Menten e Hill para estimação dos parâmetros de interesse. Ferreira, et al (2014) aborda o problema da construção de delineamentos experimentais eficientes para os ajustes de modelos não lineares em estudos de cinética química através do modelo de Michaelis-Menten e Hill.

Em modelos de curvas de crescimento, parte-se do princípio que existe uma relação funcional entre os efeitos de tratamentos e o tempo de aplicação, e que esta relação possa ser modelada. A função pode ser aproximada por um polinômio, devendo os coeficientes dessa representação polinomial, bem como, variâncias e variância-covariância, serem estimadas através dos dados (Xavier, 2000).

Algumas vantagens podem ser destacadas no ajuste de curvas de crescimento:

- Possibilitar que sejam feitas previsões sobre o comportamento médio da resposta, para ocasiões diferentes daquelas para as quais o estudo foi planejado.
- Possibilitar que sejam feitas comparações mais sensíveis entre as subpopulações, em função do número reduzido de parâmetros.
- Possibilitar a aquisição de um maior conhecimento sobre o processo de variação temporal da resposta (Lima, 1988).

Vários modelos de crescimento são introduzidos na literatura e aplicados às mais diversas áreas da ciência. Alguns modelos citados são:

- Modelo Gompertz (Gompertz, 1825);
- Modelo Michaelis-Menten (Michaelis e Menten,1913);
- Modelo de Brody;
- Modelo Logístico;
- Modelo Von Bertalanffy;
- Modelo Exponencial ;
- Modelo Weibull.

Em muitas aplicações, a parte sistemática da resposta é conhecida por ser monótona crescente em v, onde v pode representar tempo ou dosagem. Modelos de regressão não-linear com esta propriedade são chamados modelos de crescimento. O modelo de crescimento mais simples é o modelo de crescimento exponencial,

$$f(v,\theta) = \theta_1 exp(-\theta_2 v) \tag{10}$$

onde $\theta = (\theta_1, \theta_2)$.

Muitos outros modelos de crescimento são introduzidos na literatura. Nessa dissertação vamos considerar os seguintes modelos:

É necessário colocar os gráficos de cada modelo?Seria interessante

(i) $f(v,\theta) = \theta_1 + \theta_2 v^{\theta_3}$, onde $\theta = (\theta_1, \theta_2, \theta_3)$.

(ii) $f(v,\theta) = \theta_1 + \theta_2 exp[\theta_3 v]$(crescimento exponencial), onde $\theta = (\theta_1, \theta_2, \theta_3)$.

(iii) $f(v,\theta) = \theta_1 exp[\theta_2 v] + \theta_3 exp[-\theta_4 v]$, onde $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$.

(iv) $f(v,\theta) = {\theta_1 v}/{[\theta_2 + v]}$ (Michaelis-Menten), $\theta = (\theta_1, \theta_2,)$. (11)

(v) $f(v,\theta) = \theta_1 exp[-exp(\theta_2 - \theta_3 v)]$(Gompertz), onde $\theta = (\theta_1, \theta_2, \theta_3)$.

(vi) $f(v,\theta) = \theta_1 + (\theta_2 - \theta_1) \Big/ \left[1 + exp\left({(v - \theta_3)}/{\theta_4}\right)\right]$, (logístico), onde $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$.

(vii) $f(v,\theta) = \theta_1 + (\theta_2 - \theta_1) exp\left[-\theta_3 v^{\theta_4}\right]$(Weibull), onde $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$.

(viii) $f(v,\theta) = 1 - exp[-exp(\theta_1 - \theta_2 v)]$(Sigmoid), onde $\theta = (\theta_1, \theta_2,)$.

Observe-se que todos esses modelos poderiam ser considerados para modelar as curvas de crescimento. Em muitos casos, temos interpretações físicas para o uso de um determinado modelo. Em outras situações, devemos tentar diferentes modelos de crescimento para decidir pelo melhor modelo para ser assumido pelos dados.

3.2. Introdução aos métodos Bayesianos

Os métodos Bayesianos têm sido considerados alternativas muito eficazes e poderosas na análise de dados. Esse método possui uma filosofia muito diferente do método frequentista. No método clássico, os parâmetros do modelo são considerados constantes desconhecidas, e no método Bayesiano todos os parâmetros são considerados quantidades aleatórias. Uma característica muito importante é o fato desta análise permitir a incorporação da informação de um especialista junto à informação dos dados.

A inferência Bayesiana tem como fundamentação a Fórmula de Bayes, a qual combina os dados com a informação a priori e, então, se obtém a posteriori (priori já complementada pela informação dos dados), onde é realizado todo o processo inferencial (ver, por exemplo, Box e Tiao, 1973; Paulino et al., 2003).

3.2.1. Fórmula de Bayes

Considere uma partição do espaço amostral Ω, onde os eventos $A_1, A_2, \ldots, A_k$ formam uma seqüência de eventos mutuamente exclusivos e exaustivos, isto é, $\bigcup_{j=1}^{k} A_j = \Omega$ e $A_i \cap A_j = \emptyset$ (conjunto vazio) para $i \neq j$ tal que $P\left(\bigcup_{j=1}^{k} A_j\right) = \sum_{j=1}^{k} A_j = 1$. Sendo assim, para qualquer outro evento $B (B \subset \Omega)$, temos,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{k} P(B|A_j)P(A_j)}, \tag{12}$$

para todo i variando de 1 até k.

Supor, agora, um vetor de dados $\boldsymbol{y} = (y_1, \ldots, y_n)'$e θ (quantidades desconhecidas) os parâmetros de uma distribuição de probabilidade associada com a variável aleatória Y_i com valores observados y_i, $i = 1, \ldots, n$.

Considere $\boldsymbol{y} = (y_1, \dots, y_n)$ uma amostra aleatória, onde os dados são independentes e identicamente distribuídos. Supor que y possui uma distribuição conjunta dada por $\mathrm{f}\,(\mathbf{y}|\boldsymbol{\theta})$ e seja $\pi(\boldsymbol{\theta})$ uma distribuição a priori para θ. Então assumindo os valores discretos $\theta_1, \dots, \theta_k$, encontramos de (12), a distribuição a posteriori para θ_i dado $\mathbf{y}$,

$$\pi(\theta_i|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\theta_i)\pi(\theta_i)}{\sum_{j=1}^{k} f(\boldsymbol{y}|\theta_j)\pi(\theta_j)}, \tag{13}$$

onde o parâmetro θ também é considerado como uma quantidade aleatória, no enfoque Bayesiano.

Para θ assumindo valores contínuos num dado intervalo, podemos escrever (13) por

$$\pi(\theta|y) = \frac{f(\boldsymbol{y}|\theta)\pi(\theta)}{\int f(\boldsymbol{y}|\theta)\pi(\theta)}, \tag{14}$$

em que a integral no denominador de (14) é definida no intervalo de variação de θ.

3.2.2. Distribuição a Priori

Na análise Bayesiana, a distribuição a priori é utilizada a fim de representar o que já é conhecido sobre parâmetros desconhecidos, antes de se avaliar os dados. Deve-se ter muita cautela ao definir uma distribuição a priori, afinal se esta informação não for bem definida pode-se chegar a interpretações errôneas. Uma distribuição a priori para um parâmetro pode se dar de várias formas, sendo possível ocorrerem a partir de procedimentos subjetivos ou objetivos.

A distribuição a priori conjugada é uma priori informativa, onde a distribuição a priori e a posteriori pertencem à mesma classe de distribuições. A passagem da priori para a posteriori envolve apenas uma simples mudança nos parâmetros, sem a necessidade de cálculos adicionais (ver, por exemplo, Paulino et al., 2003).

A distribuição a priori também pode ser a incorporação do conhecimento de um pesquisador, ou seja, o pesquisador se baseia na sua prática e no seu *feeling* para definir a priori. Em muitas situações práticas esse conhecimento do especialista não existe ou, se

existe, não é fidedigno. Nesse caso, caracteriza-se uma "ignorância a priori" (ver, por exemplo, Paulino et al., 2003). Nestas situações são utilizadas as prioris não informativas, ou seja, prioris de referência "neutras" (ver, por exemplo, Box e Tiao, 1973). A utilização deste tipo de priori permite a comparação com os resultados obtidos pela inferência clássica, haja vista que através de uma priori não informativa, o modelo é baseado apenas na informação dos dados amostrais. Existem vários métodos para se definir a priori não informativa, como por exemplo: Método de Bayes- Laplace, Método de Jeffreys, entre outros (ver, por exemplo, Box e Tiao, 1973; Paulino et al., 2003).

3.2.3. Critério para Discriminação de Modelos – Deviance Information Criterion (DIC)

O critério DIC (proposto por Spiegelhalter et al., 2002) é baseado na média a posteriori da deviance. Define-se a deviance por

$$D(\theta) = -2\, ln\, L(\theta) + C, \tag{15}$$

em que θ é um vetor de parâmetros desconhecidos do modelo; $L(\theta)$ é a verossimilhança e C é uma constante (nem sempre conhecida) na comparação de dois modelos.

O critério DIC é dado então por

$$D(\theta) = D(\hat{\theta}) + 2p_D, \tag{16}$$

em que $D(\hat{\theta})$ é o desvio calculado na média a posteriori $\hat{\theta} = E(\theta|y)$ e p_D é o número de parâmetros do modelo, dado por $p_D = \overline{D} - D(\hat{\theta})$ em que $\overline{D} = E[D(\theta|y)]$ é a média a posteriori do desvio que mede a qualidade do ajuste dos dados para cada modelo. Para a conclusão, os menores valores de DIC indicam os melhores modelos.

3.2.4. Métodos de simulação para amostras da distribuição a posteriori

Na presente sessão será apresentado um breve resumo das técnicas de simulação usadas neste trabalho para se obter os sumários a posteriori de interesse.

Para a obtenção de sumários a posteriori de interesse, geralmente é necessário resolver integrais Bayesianas que não apresentam solução analítica. Na prática, observa-se que os modelos utilizados nem sempre são simples para se obter os resumos a posteriori. Mesmo que se tenha uma priori e uma verossimilhança simples, a junção delas pode produzir uma distribuição a posteriori muito complicada (ver, por exemplo, Paulino et al., 2003).

Os métodos com base em amostragem, como, por exemplo, o método de Monte Carlo com Cadeias de Markov (MCMC), passaram a ser muito utilizados com o avanço dos recursos computacionais (hardware e software). Esses métodos consistem na simulação de uma variável aleatória a partir de uma cadeia de Markov, no qual a sua distribuição assintoticamente se aproxima da distribuição a posteriori (ver, por exemplo, Bernardo e Smith, 1994).

A cadeia de Markov é um processo estocástico no qual o próximo estado da cadeia depende somente do estado atual e dos dados. No entanto, existe uma relação com o estado inicial, que é descartado após um período de aquecimento, o chamado “burn-in”.

As formas mais usuais dos métodos MCMC são os amostradores de Gibbs e o algoritmo de Metropolis-Hastings. Essas duas formas simulam amostras da distribuição a posteriori conjunta a partir das distribuições condicionais (ver por exemplo, Gelfand e Smith, 1990; Chib e Greenberg, 1995).

O amostrador de Gibbs nos permite gerar amostras da distribuição a posteriori conjunta desde que a distribuições condicionais completas possuam formas fechadas ou conhecidas. Por outro lado, o algoritmo de Metropolis-Hasting permite gerar amostras da distribuição a posteriori conjunta com distribuições condicionais completas possuindo ou não uma forma conhecida ou fechada.

- **Amostrador de Gibbs**

Suponha $\pi(\boldsymbol{\theta}|y)$ uma distribuição a posteriori conjunta, sendo $\boldsymbol{\theta} = \theta_1, \dots, \theta_k$, no qual desejamos obter inferências. Para isso, simulam-se quantidades aleatórias de distribuições condicionais completas $\pi = \left(\theta_i|y, \theta_{(i)}\right)$.

Considere os valores iniciais (arbitrários) para $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$. Dessa forma, segue o seguinte algorimo:

- Gerar $\theta_1^{(1)}\ de\ \pi\left(\theta_1|y,\theta_2^{(0)},\dots,\theta_k^{(1)}\right)$;

- Gerar $\theta_2^{(1)}\ de\ \pi\left(\theta_2|y,\theta_1^{(1)},\theta_3^{(0)}\ \dots,\theta_k^{(0)}\right)$;

(...) (17)

- Gerar $\theta_k^{(1)}\ de\ \pi\left(\theta_2|y,\theta_1^{(1)},\theta_2^{(1)}\ \dots,\theta_{k-1}^{(1)}\right)$;

Substitua os valores iniciais por $\theta^{(1)} = (\theta_1^{(1)},\theta_2^{(1)},\dots,\theta_k^{(1)})$, para uma nova realização. Os valores $\theta_1^{(z)},\theta_2^{(z)},\dots,\theta_k^{(z)}$, para z suficientemente grande, convergem para um valor da quantidade aleatória com distribuição $\pi(\theta|y)$ (ver, por exemplo, Bernardo e Smith, pág 353, 1994).

- **Algoritmo de Metropolis-Hastings**

Suponha uma amostra de densidade não regular $\pi\left(\theta_i|\theta_{(i)}\right)$, em que $\theta_{(i)} = (\theta_1,\dots,\theta_{i-1},\theta_{i+1},\dots,\theta_k)$. Seja $q(\theta,\beta)$ o núcleo de transição da distribuição $p(\theta)$ que representa $\pi\left(\theta_i|\theta_{(i)}\right)$, e que transforma θ em β.

Desta forma o algoritmo é dado por:

- Inicie com $\theta^{(0)}$e indicador de estágio $j = 0$;
- Gerar um ponto β do núcleo de transição $q(\theta^{(j)},\beta)$; (18)
- Atualizar $\theta^{(j)}$ por $\theta^{(j+1)} = \beta$, com probabilidade, $p = min\left\{1,\frac{p(\beta)q[\theta^{(j)},\beta]}{p\left[\theta^{(j)}q[\beta,\theta^{(j)}]\right]}\right\}$, ficar com $\theta^{(j)}$ com probabilidade $1-p$;
- Repetir os dois últimos passos até conseguir uma distribuição estacionária.

Importante observar: se um valor candidato é rejeitado, então o valor atual é considerado na próxima etapa; o terceiro passo o valor de $\boldsymbol{p}$ não depende da constante normalizadora; o algoritmo de Metropolis Hastings é especificado pela densidade candidata para geração $\boldsymbol{q}(\boldsymbol{x},\boldsymbol{y})$ (ver, por exemplo, Bernardo e Smith, 1994; Chib e Greenberg, 1995).

4. Um exemplo de dados de crescimento em estudos ambientais

Na Tabela 2, temos um conjunto de dados relacionados com a concentração de bifenil policlorados (PCB), resíduos em uma espécie de trutas do lago Cayuga, NY, EUA reportados por Bache et al (1972). As idades dos peixes foram conhecidas com precisão, porque os peixes foram colocados anualmente como animais de um ano e claramente marcados com o ano. As amostras foram tratadas e resíduos de PCB em partes por milhão (ppm) foram estimados utilizando cromatografia em coluna.

Tabela 2: Dados de contaminação de trutas do lago Cayuga

Obs	Idade	PCB	log(PCB)
1	1	0,6	-0,51083
2	1	1,6	0,47000
3	1	0,5	-0,69315
4	1	1,2	0,18232
5	2	2,0	0,69315
6	2	1,3	0,26236
7	2	2,5	0,91629
8	3	2,2	0,78846
9	3	2,4	0,87547
10	3	1,2	0,18232
11	4	3,5	1,25276
12	4	4,1	1,41099
13	4	5,1	1,62924
14	5	5,7	1,74047
15	6	3,4	1,22378
16	6	9,7	2,27213
17	6	8,6	2,15176
18	7	4,0	1,38629
19	7	5,5	1,70475
20	7	10,5	2,35138
21	8	17,5	2,86220
22	8	13,4	2,59525
23	8	4,5	1,50408
24	9	30,4	3,41444
25	11	12,4	2,51770
26	12	13,4	2,59525
27	12	26,2	3,26576
28	12	7,4	2,00148

Na Figura 2, temos os gráficos de PCB versus idade e log(PCB) versus idade. Dos gráficos da Figura 2, observa-se que existe uma relação não-linear para os dados originais e transformados para uma escala logarítmica.

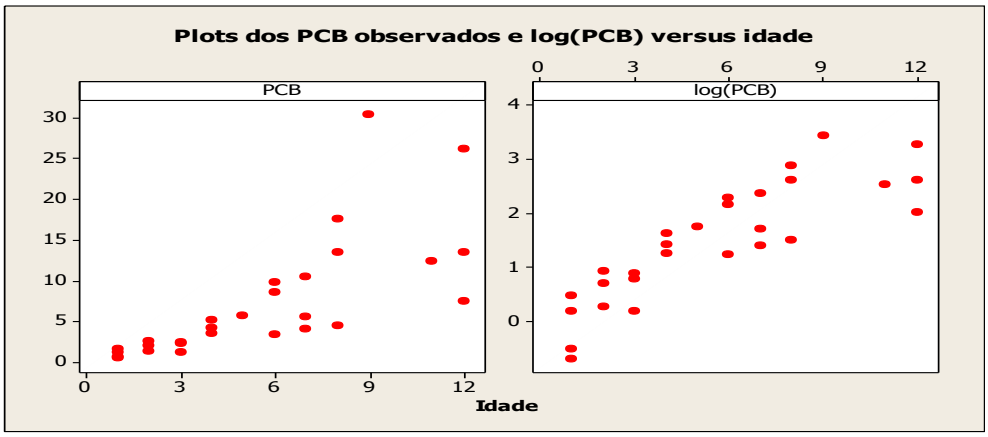


Figura 2: Plots dos PCB observados e log(PCB) versus idade

5. Análise dos dados da Tabela 2

Assumindo iniciamente uma abordagem clássica padrão para modelos não lineares, considerando os modelos de crescimento (i) - (viii) introduzidos em (11) com as respostas na escala logarítmica (PCB), e usando o software MINITAB versão 16, temos na Tabela 3, as estimativas para cada modelo de crescimento assumido. A partir dos resultados da Tabela 3, observa-se que, para alguns modelos de crescimento, precisamos de um grande número de iterações para obter a convergência do algoritmo de Gauss-Newton.

Na Figura 3, temos os gráficos de log observada (PCB) e os modelos ajustados versus observações. A partir dessas figuras, observamos que a maioria dos modelos assumidos tem ajuste razoável para os dados (modelos 1,2,4,5,6 e 7 apresentados na Tabela 4). Apenas os modelos 3 e 8 aparentam não serem adequados para os dados. É importante observar na Tabela 3 que a soma dos valores absolutos para as diferenças (observado - estimada média).

Tabela 3: Soma dos valores absolutos para as diferenças (Inferência Clássica)

Modelos	Soma dos Valores Absolutos
1	11,4412
2	11,1411
3	**15,5568**
4	12,0196
5	11,0047
6	11,0209
7	11,0310
8	**23,4592**

Tabela 4: Estimativas clássicas para modelos de crescimento

Modelo ajustado	#iterações	Estimadores(SE) (Desvio Padrão)			
		θ_1	θ_2	θ_3	θ_4
Modelo (1)	11	-4,86638	4,70326	0,19681	
$\boldsymbol{log(PCB) = \theta_1 + \theta_2 Idade^{\theta_3}}$		(8,42903)	(8,27683)	(0,27393)	
Modelo Exponencial (2)	17	3,12932	-3,92923	-0,19029	
$\boldsymbol{log(PCB) = \theta_1 + \theta_2 exp[\theta_3 Idade]}$		(0,578479)	(0,451280)	(0,074400)	
Modelo (3)	200	12,9631	0,0304	-12,7809	0,0495
$\boldsymbol{log(PCB) = \theta_1 exp[\theta_2 Idade] + \theta_3 \exp[-\theta_4 * \mathrm{Idade}]}$		(10164,1)	(7,1)	(10163,5)	(8,1)
Modelo Michaelis Menten (4)	12	8,6431	23,9661		
$\boldsymbol{log(PCB) = \theta_1 Idade/[\theta_2 + Idade]}$		(5,3952)	(20,0487)		
Modelo Gompertz (5)	14	2,70624	1,53316	0,43543	-
$\boldsymbol{log(PCB) = \theta_1 exp[-exp(\theta_2 - \theta_3 idade)]}$		(0,30029)	(0,42996)	(0,135794)	
Modelo Logístico (6)	13	2,9167	-5,145	-0,85077	3,56075
$\boldsymbol{log(PCB) = \theta_1 + (\theta_2 - \theta_1)/[1 + exp((Idade - \theta_3)/\theta_4)]}$		(0,7308)	(22,1369)	(16,8113)	(4,2535)
Modelo Weibull (7)	18	2,92618	-0,56122	0,14451	1,18735
$\boldsymbol{log(PCB) = c + (\theta_2 - \theta_1) * exp[-\theta_3 Idade^{\theta_4}]}$		(0,80668)	(0,88077)	(0,17344)	(0,77622)
Modelo (8)	21	-3,4462	-1,2850		
$\boldsymbol{log(PCB) = 1 - exp[-exp(\theta_1 - \theta_2 Idade)]}$		(5,80409)	(2,26049)		

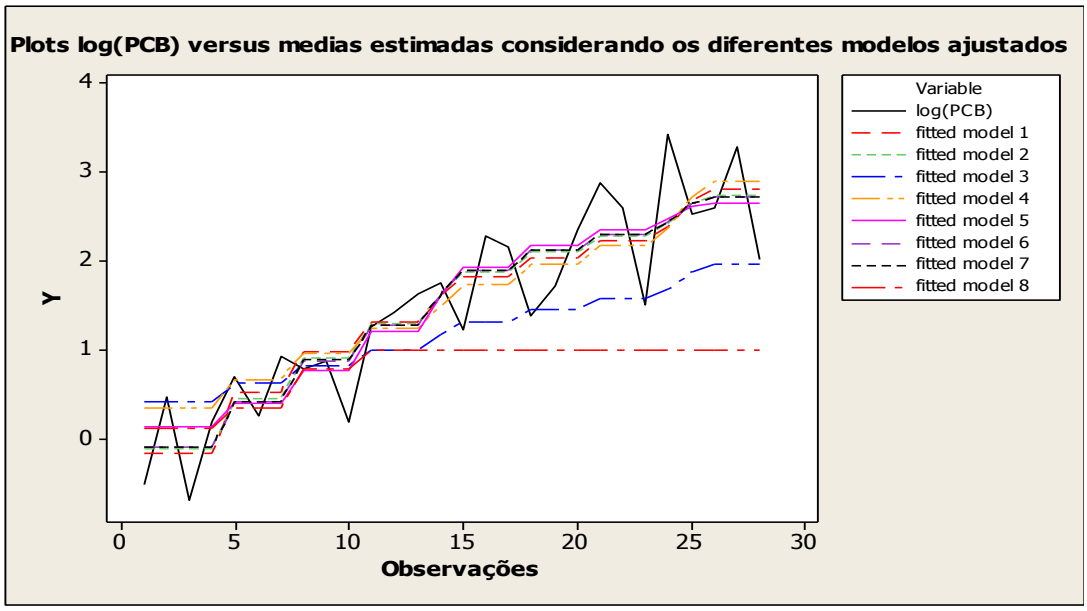


Figura 3: Plots log(PCB) versus medias estimadas considerando os diferentes modelos ajustados

Observar que as estimativas clássicas obtidas e apresentados na Tabela 4, podem ser muito instáveis dependendo dos valores iniciais utilizados no algoritmo iterativo. Normalmente, é difícil obter bons valores iniciais empíricos para ser utilizado no procedimento iterativo. Outro ponto: os erros-padrão das estimativas obtidas podem ser muito grandes como se observa na Tabela 4. Gráficos de resíduos para os modelos ajustados por mínimos quadrados são apresentados no Apêndice 1. Em geral observa-se normalidade dos resíduos e variância constante para os modelos ajustados.

Sob uma abordagem bayesiana, temos na Tabela 6, os sumários a posteriori de interesse, considerando os oito modelos de crescimento (I) - (viii) introduzidos em (11) assumindo uma distribuição normal para o erro (ver (4)) e usando o software OpenBugs (Spiegelhalter et al, 2003) assumindo as seguintes distribuições a priori: $\theta_j \sim U(0,1), j = 1,2,3,4$ e $\zeta = 1/\sigma^2 \sim G(1,1)$, onde $G\ (a, b)$ denota uma distribuição gama com média a/b e variância a/b^2. Foram simulados 10.000 amostras de Gibbs, onde uma amostra de aquecimento ("burn-in sample") de 1.000 amostras foram descartadas para eliminar os efeitos dos valores iniciais no processo iterativo de simulação e obtendo uma amostra final de tamanho 900 (cada amostra foi escolhida de 10 em 10 das 9000 amostras restantes) . A convergência do amostrador de Gibbs foi monitorada a partir de gráficos de séries temporais das amostras simuladas. No Apêndice A2 são apresentados apenas os códigos dos programas OpenBugs usados para cada modelo de crescimento.

A partir dos resultados da Tabela 6, observa-se que o desvio padrão a posteriori para cada parâmetro do modelo de crescimento é muito pequeno. Isso mostra uma possível vantagem do método bayesiano quando comparado com o método clássico. É importante salientar que estes sumários a posteriori obtidos por métodos MCMC são muito precisos não dependendo de aproximações, como é considerado utilizando a abordagem clássica.

Para decidir sobre o melhor modelo estatístico podemos usar a seleção Bayesian Deviance Information Criterion (DIC) –ver seção 3.2.3, este critério é especialmente útil em problemas onde amostras da distribuição a posteriori para os parâmetros do modelo foram simulados utilizando Cadeias de Markov Monte Carlo (MCMC).

Na Figura 4, temos os gráficos de log observada (PCB) e modelos ajustados versus observações e na Tabela 5 temos as somas dos valores absolutos das diferenças (valor observado – estimador da média). A partir dessas figuras e da tabela, observa-se que os modelos de 1, 5 e 6 dão um bom ajuste para os dados. Na verdade, as somas dos valores absolutos das diferenças para os modelos 1, 5 e 6 são dadas, respectivamente, por 11,8508 (DIC = 45,0), 11,0915 (DIC = 49,47) e 11,7191 (DIC = 51,17). Ou seja, obtemos resultados muito semelhantes, uma indicação de ajustes adequados, considerando estes três modelos. Considerando-se os modelos 2, 3, 4, 7 e 8, os valores para estas diferenças são dadas, respectivamente, por, 14,5878, 12,0738, 16,8414, 12,3670 e 25,1309, uma indicação de que esses modelos não são bem ajustados pelos dados.Em geral, o modelo 5 (Gompertz) fornece o menor valor para a soma dos valores absolutos das diferenças (observada - estima média) e pequeno DIC (40,47), ou seja, podemos considerar este modelo como o modelo mais adequado.

Tabela 5: Soma dos Valores Absolutos (Bayesiana)

Modelos	Soma dos Valores Absolutos
1	**11,8508**
2	14,5878
3	12,0738
4	16,8414
5	**11,0915**

6	**11,7191**
7	12,3670
8	25,1309

Tabela 6: Estimadores Bayesianos para os modelos de crescimento assumindo erros normais

Modelo ajustado	DIC	Média a posteriori (Desvio-padrão)			
		θ_1	θ_2	θ_3	θ_4
Modelo (1)	45.0	-0.7804	0.9227	0.5946	
$log(PCB) = \theta_1 + \theta_2 Idade^{\theta_3}$		(0.4839)	(0.3843)	(0.1458)	
Modelo Exponencial (2)	61.71	-0.5457	1.096	0.1065	
$log(PCB) = \theta_1 + \theta_2 exp[\theta_3 Idade]$		(0.5339)	(0.4373)	(0.02785)	
Modelo (3)	52.56	-1.701	0.4978	1.094	-0.08517
$log(PCB) = \theta_1 exp[\theta_2 Idade] + \theta_3 \exp[-\theta_4 * \text{Idade}]$		(0.6316)	(0.2242)	(0.2721)	(0.02407)
Modelo Michaelis Menten (4)	67.6	2.314	2.054		
$log(PCB) = \theta_1 Idade/[\theta_2 + Idade]$		(0.3152)	(0.662)		
Modelo Gompertz (5)	40.47	2.506	1.546	0.501	
$log(PCB) = \theta_1 exp[-exp(\theta_2 - \theta_3 idade)]$		(0.2677)	(0.4094)	(0.1404)	
Modelo Logístico (6)	51.17	2.319	-1.458	1.873	1.746
$log(PCB) = \theta_1 + (\theta_2 - \theta_1)/[1 + exp((Idade - \theta_3)/\theta_4)]$		(0.2294)	(0.6554)	(0.6762)	(0.4869)
Modelo Weibull (7)	-6.60	-1.318	-0.4967	-0.7765	0.3727
$log(PCB) = c + (\theta_2 - \theta_1) * exp\left[-\theta_3 Idade^{\theta_4}\right]$		(0.6242)	(0.4649)	(0.3716)	(0.1372)
Modelo Sigmoid (8)	87.31	-0.3885	-0.7789		
$log(PCB) = 1 - exp[-exp(\theta_1 - \theta_2 Idade)]$		(1.048)	(0.5216)		

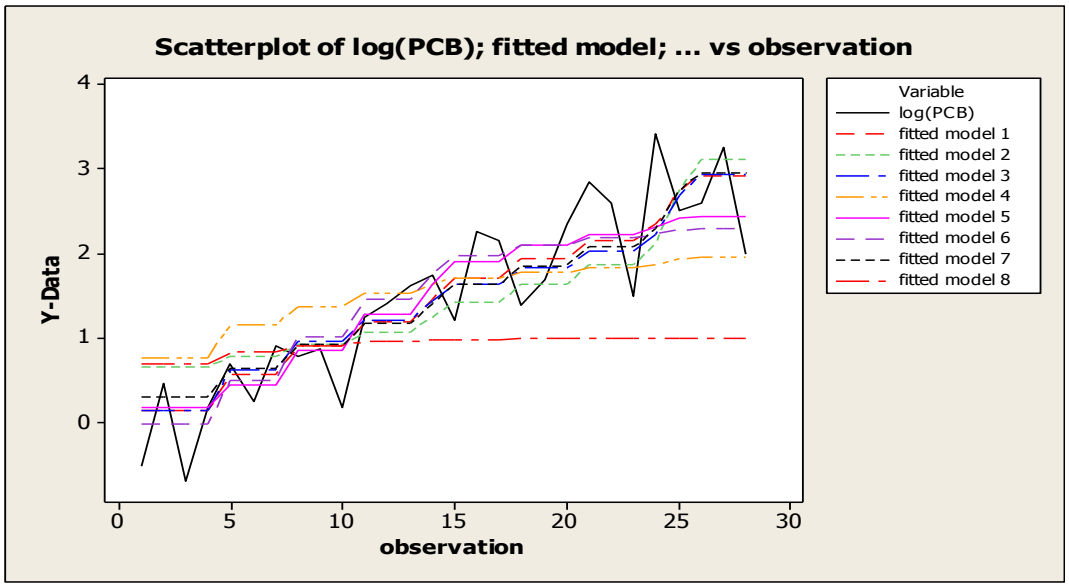


Figura 4: Plots de log observada (PCB) versus média a posteriori

6. Uso de uma distribuição estável para modelos de crescimento

Em muitos casos, a suposição habitual de normalidade para os erros em (1) não pode ser apropriado. Esta situação ocorre, por exemplo, quando temos observações discordantes que podem afectar grandemente as inferências obtidos. Desta forma, poderíamos assumir distribuições mais robustas para os dados, como a distribuição estável. Outro ponto de interesse: a utilização dos resultados de inferência assintóticas não poderia ser precisas dependendo dos tamanhos de amostra e a curvatura intrínseca da função f em (1).

Deste modo, nós propomos o uso de distribuições estáveis para o Xi resposta em (1).

Uma ampla classe de distribuição que compreende a distribuição de Gauss é dada pela classe de distribuições estáveis. Este grande classe define famílias localização escala que estão fechados sob convolução. A distribuição de Gauss é um caso especial desta família de distribuição (ver, por exemplo, Nolan, 2009), descrito por quatro parâmetros α, β, γ e δ. O αϵ (0,2] parâmetro define a "gordura das caudas", e quando α = 2 esta classe reduz à distribuição de Gauss. O β ϵ [-1,1] é o parâmetro de assimetria e para β = 0 tem um simétrico distribuições. A localização e parâmetros de escala são, respectivamente, γ ϵ (- ∞, ∞) e δ ϵ (0, ∞) (ver Levy, 1924)

Distribuições estáveis são geralmente indicados por $S_\alpha(\beta, \gamma, \delta)$. Se uma variável aleatória $X \sim S_\alpha(\beta, \gamma, \delta)$, então $Z = \frac{X-\gamma}{\delta} \sim S_\alpha(\beta, 0,1)$ (ver Lukács, 1970 e Nolan, 2009). Uma grande dificuldade associada às distribuições estáveis $S_\alpha(\beta, \gamma, \delta)$, é que, em geral, não há forma fechada simples para as suas funções de densidade de probabilidade. No entanto, sabe-se que as funções de densidade de probabilidade de distribuições estáveis são contínuas (Gnedenko e de Kolmogorov, 1968; Skorohod, 1961) e unimodal (Ibragimov e Cernin, 1959). Além disso, o suporte de todas as distribuições estáveis é dada em (- ∞, ∞), excepto para α <1 e | β | = 1, quando o suporte é (-∞,0) para β= 1 e (0,∞) para β= -1 (ver Feller, 1971).

A função característica Φ (.) De uma distribuição estável é dado por,

$$log[\Phi\,(t)] = \begin{cases} i\gamma t - |\delta t|^{\alpha}[1 - i\beta sign(t)\,tan(\frac{\pi\alpha}{2})],\ for\ \alpha\ \neq 1 \\ i\gamma t - |\delta t|^{\alpha}[1 + i\beta sign(t)\,(\frac{2}{\pi})log(t)]\,,\ for\ \alpha =\ 1 \end{cases}, \qquad (19)$$

onde $i = \sqrt{-1}$ e a função $sign$ (.) é dada por,

$$sign(x) = \begin{cases} -1, & if\ x < 0 \\ 0, & if\ x = 0 \\ 1, & if\ x > 0. \end{cases},$$

É importante salientar que se α <1, a variância é infinito e a média da distribuição estável não existe. Embora esta classe de distribuições é uma boa alternativa para a modelagem de dados em diferentes áreas, que normalmente têm dificuldades para obter estimativas no âmbito de uma abordagem de inferência clássica devido à falta de expressões de forma fechada para as suas funções de densidade de probabilidade. Uma possibilidade de aplicações, é fazer com que a função densidade de probabilidade a partir da fórmula de inversão (ver, por exemplo Roussas, 2005),

$$f(x) = 1/2\pi \int_{-\infty}^{\infty} e^{-itx}\, \Phi(t)dt, \qquad (20)$$

onde Φ (t) é a função característica. Em aplicações, precisamos usar métodos numéricos para resolver a integral em (20), geralmente tendo um grande tempo computacional.

Uma alternativa é a utilização de métodos Bayesianos. No entanto, o custo computacional pode ser ainda mais alto para obter os sumários a posteriori de interesse. Uma alternativa é usar variáveis latentes ou artificiais que poderiam melhorar o cálculo de simulação de amostras das distribuições posteriores conjuntas de interesse (ver, por exemplo Damien, et al 1999; Tanner e Wong, 1987).

Desta forma, uma análise Bayesiana de distribuições estáveis foi introduzido por Buckle (1995), utilizando Markov Chain Monte Carlo (MCMC) métodos e variáveis latentes (ver também, Achcar et al, 2013). O uso de métodos bayesianos com simulação MCMC pode ter grande flexibilidade, considerando variáveis latentes, onde amostras de variáveis latentes são simuladas em cada passo dos algoritmos Gibbs ou Metropolis-Hastings.

Considerando uma latente ou uma variável auxiliar, Buckle (1995) provou um teorema que é útil para simular amostras da distribuição a posteriori conjunta para os parâmetros α, β, γ e δ. Este teorema estabelece que uma distribuição estável para uma variável aleatória Z definido em (- ∞, ∞) é obtido como o marginal de uma distribuição

bivariada para a própria variável aleatória Z e uma variável aleatória auxiliar Y. Esta variável Y é definida no intervalo (-0.5, $a_{\alpha,\beta}$), quando $Z \in (-\infty, 0)$, e em $(a_{\alpha,\beta}, 0.5)$, quando $Z \in (0, \infty)$. A quantidade $a_{\alpha,\beta}$ é dada por:

$$a_{\alpha,\beta} = -\frac{b_{\alpha,\beta}}{\alpha\pi}, \tag{21}$$

onde $b_{\alpha,\beta} = \beta \min\{\alpha, 2-\alpha\}\frac{\pi}{2}$.

A função densidade de probabilidade conjunta para as variáveis aleatórias Z e Y i é dada por,

$$f(z, y|\alpha, \beta) = \frac{\alpha}{|\alpha-1|} exp\ \{-|\frac{z}{t_{\alpha,\beta}(y)}|^{\theta}\}\ |\frac{z}{t_{\alpha,\beta}(y)}|^{\theta}\frac{1}{|z|}, \tag{22}$$

onde $\theta = \frac{\alpha}{\alpha-1}$,

$$t_{\alpha,\beta}(y) = \left(\frac{sin(\pi\alpha\ y+b_{\alpha,\beta})}{cos(\pi\ y)}\right)\left(\frac{cos(\pi\ y)}{cos(\pi(\alpha-1)y+b_{\alpha,\beta})}\right)^{\frac{1}{\theta}}, \tag{23}$$

e

$$Z = \frac{X-\gamma}{\delta}, \text{ for } \delta \neq 0.$$

A partir da densidade bivariada (10), Buckle (1995) encontra que a distribuição marginal da variável aleatória Z é uma distribuição estável $S_{\alpha}(\beta, 0, 1)$. Geralmente, os custos computacionais para obter sumários a posteriori de interesse utilizando métodos MCMC é elevado para esta classe de modelos, o que poderia dar algumas limitações para aplicações práticas. Um problema pode ser a convergência algoritmo de simulação. Neste trabalho, propomos o uso de um popular software disponível gratuitamente para obter os resumos posteriores de interesse: o software OpenBugs (Spiegelhalter et al, 2003).

6.1. Uma análise Bayesiana assumindo a distribuição estável

Nesta seção, vamos supor que a resposta x_i no modelo de regressão não-linear (1) para i = 1, ..., n, tem uma distribuição estável $X_i \sim S_{\alpha}(\beta, \gamma_i, \delta)$, isto é, $Z_i = \frac{X-\gamma_i}{\delta} \sim S_{\alpha}(\beta, 0, 1)$ onde o parâmetro γ_i de localização da distribuição estável está relacionado com o K variáveis explanatórias $\mathbf{v_i} = (v_{1i}, v_{2i}, ..., v_{ki})$ pela uma relação não linear dada por,

$$\gamma_i = f(v_i, \boldsymbol{d}), \qquad (24)$$

onde **d** = (d_1,d_2,…,d_k) são os parãmetros de regressão desconhecidos. Assumindo uma distribuição a priori conjunta para α, β , d e δ, onde **d** = (d_1,d_2,…,d_k) denotada por $\boldsymbol{\pi_0}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{d}, \boldsymbol{\delta})$, Buckle (1995) mostra que a distribuição a posteriori conjunta para α, β , d e δ, é dada por,

$$\boldsymbol{\pi}\,(\alpha, \beta, \boldsymbol{d}, \delta / \boldsymbol{x})$$

$$\propto \int \left(\frac{\alpha}{|\alpha - 1|\delta}\right)^n exp\left\{-\sum_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(Y_i)}\right|^{\theta}\right\}\prod_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta} \cdot \qquad (25)$$

$$\cdot\ \frac{1}{|z_i|}\pi_0(\alpha, \beta, \boldsymbol{d}, \delta)\mathrm{d}\mathbf{y},$$

onde $\theta = \frac{\alpha}{\alpha-1}$, $z_i = \frac{x_{i-\gamma i}}{\delta}$, para i = 1,…,n, α ϵ (0, 2], β ϵ [-1, 1] e δ ϵ (0,∞); **x** = $(x_1, x_2, \dots, x_n)$ e **y** = $(y_1, y_2, \dots, y_n)$ são, respectivamente, os vetores de dados observados e não observados. Observa-se que a distribuição conjunta na expressão (25) é dada em termos de x_i e as variáveis latentes variables y_i, e não em termos de z_i e y_i (existe um Jacobian σ^{-1} multiplicado pelo lado direito da expressão (22)).

Observar que quando α = 2 , temos θ = 2 e $b_{\alpha,\beta}$ = 0. Neste caso, temos uma distribuição gaussiana, com média igual a δ e variância igual a $2\sigma^2$.

Para uma análise bayesiana do modelo proposto, assumimos distribuições a priori uniformes U(a, b) para α, β e δ onde os hiperparâmetros a e b são assumidos conhecidos em cada aplicação seguindo as restrições α ϵ (0, 2], β ϵ [-1, 1] e δ ϵ (0, ∞). Também assumimos distribuições a priori normais N(a,b^2) para os parâmetros de regressão **d** = (d_1, d_2, ..., d_k) considerando os hiperparâmetros a e b^2 conhecidos.

Além disso, assumimos independência entre todos os parâmetros.

No algoritmo de simulação para obter uma amostra de Gibbs para as quantidades aleatórias α, β, d and δ, tendo a distribuição a posteriori (13), assumimos uma distribuição uniforme U(-0,5, 0,5) para as quantidades aleatória latentes Y_i para i = 1, ..., n. Observe que, neste caso, estamos assumindo $a_{\alpha,\beta}$ = 0 ($b_{\alpha,\beta}$ = 0). Com esta escolha de distribuições a priori, usamos softwares disponíveis existentes como o software OpenBugs (veja Spiegelhalter, 2003), que leva a grande simplificação na obtenção das amostras Gibbs simuladas para a distribuição a posteriori conjunta de interesse.

A distribuição a posteriori conjunta para α , β , **d** , δ e **y** = $(y_1, y_2, \dots, y_n)$ é dada por,

$$\pi(\alpha, \beta\ , d\ , \delta\ , y|\ x) \propto \left(\frac{\alpha}{|\alpha-1|\delta}\right)^n \exp \left\{\sum_{i=1}^{n} \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|\right\} \prod_{i=1}^{n} \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta} \frac{1}{|z_i|} \cdot \cdot \prod_{i=1}^{n} h\ (y_i)\pi_0(\alpha, \beta\ , d\ , \delta\ ), \quad (26)$$

onde θ e $t_{\alpha,\beta}(.)$ são definidas em (21) e (22) e $h(y_i)$ é uma distribuição U(- 0.5, 0.5) para i = 1,…,n.

Como estamos usando o software OpenBugs para simular amostras da distribuição a posteriori conjunta nós não apresentamos aqui todas as distribuições condicionais completas necessárias para o algoritmo de amostragem de Gibbs. Este software requer apenas a distribuição dos dados e distribuições a priori das quantidades aleatórias interessadas. Isto dá uma grande simplificação computacional para determinar sínteses posteriores de interesse, como mostrado nas aplicações a seguir.

7. Análise Bayesiana dos dados das Tabelas 1a e 1b assumindo uma distribuição estável

Nas Tabelas 1a e 1b, temos um conjunto de dados relacionado com dados de crescimento intra-uterino (IGD) para prematuros nos EUA (ver Olsen et al, 2010). Neste conjunto de dados que temos o peso dos prematuros e a idade gestacional (IG) das mães (55.445 meninas e meninos 73.995). A partir de uma análise de dados preliminar, vemos que um modelo de regressão não-linear (1) é adequado para o conjunto de dados nas escalas transformadas (ver Figura 1).

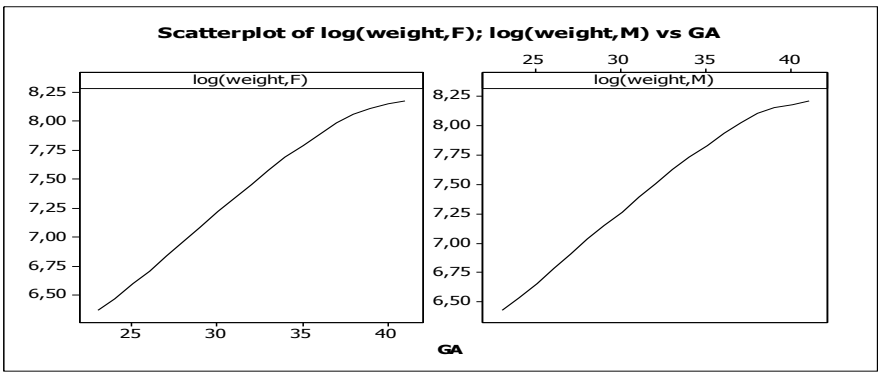


Figura 5. Gráficos de log(pesos) versus idade gestacional.

.

Utilizando a informação das Tabelas 1a e 1b, uma vez que não temos todos os dados (pesos intra-uterinos de 55.445 meninas e meninos 73.995) relatados no artigo de Olsen et al (2010), mas como temos as médias amostrais e os desvios-padrão para cada grupo de recém-nascidos em uma idade gestacional fixo, simulamos amostras dos pesos considerando distribuições normais com as médias e desvios-padrão iguais às médias amostrais e desvios-padrão amostrais (simulação de 553 pesos para as meninas e 738 pesos para meninos) com um tamanho de amostra que dá perto de 10% de todos os dados introduzidos na Tabela 1.

Tabela 7: Dados de crescimento utrainterino para crianças (55445 meninas e 73995 meninos)

#	GA	size F	mean,F	size M	mean,M
1	23	133	587	153	622
2	24	438	649	451	689
3	25	603	738	722	777
4	26	773	822	881	888
5	27	966	934	1030	1001
6	28	1187	1058	1281	1138
7	29	1254	1199	1505	1277
8	30	1606	1376	1992	1435

9	31	2044	1548	2460	1633
10	32	3007	1730	3677	1823
11	33	4186	1960	5014	2058
12	34	5936	2194	7291	2288
13	35	5082	2420	6952	2529
14	36	4690	2675	7011	2798
15	37	4372	2946	6692	3058
16	38	5755	3184	8786	3319
17	39	5978	3342	8324	3476
18	40	5529	3461	7235	3582
19	41	1906	3546	2538	3691

(GA=idade gestacional; F=feminino; M=masculino; size=tamanho amostral)

7.1. Modelo de Gompertz com erros normais (procedimento clássico)

Assumindo abordagem clássica padrão para modelos não lineares, considerando o modelo de crescimento de Gompertz , isto é,

$$f(v, \theta) = \theta_1 exp[-exp(\theta_2 - \theta_3 v)], \qquad (27)$$

onde v = GA (idade gestacional), com as respostas na escala logarítmica (peso), e usando o software Minitab versão 16, temos na Tabela 8, as estimativas obtidas por máxima verossimilhança (MLE) para os parâmetros do modelo de crescimento de Gompertz para cada sexo. Foram usados valores iniciais iguais a zero para todos os parâmetros θ_1, θ_2 e θ_3 do modelo Gompertz no processo iterativo de Gauss-Newton para obter os MLEs.

Tabela 8. MLE Estimativa para o modelo de crescimento Gompertz

Gênero	θ_1	θ_2	θ_3
F (meninas)	9,06173	0,76876	0,07561
M(meninos)	9,14384	0,77103	0,07488

De gráficos residuais observa-se que as suposições necessárias para validar as inferencias (resíduos normais e variância constante) foram verificados.

7.2. Modelo de Gomperz com erros normais (procedimento Bayesiano)

Sob uma abordagem bayesiana, temos na Tabela 9, os sumários a posteriori de interesse, considerando o modelo de crescimento de Gompertz com uma distribuição normal para o erro e o software OpenBugs assumindo o seguinte antes distribuições:

$\theta_j \sim N(0, 1)$, j = 1,2,3,4 e $\zeta = 1/\sigma^2 \sim G(1,1)$, em que G(a, b) denota uma distribuição gama com média a/b e variância a/b^2. Simulamos 311.000 amostras de Gibbs, com um "burn-in-sample" de 11.000 amostras descartadas para eliminar os efeitos dos valores iniciais no processo de simulação interativa e tendo uma amostra final de 1000 tamanho (cada amostra de 300 escolhidas a partir da amostra final simulada de tamanho 300.000). A convergência do algoritmo de amostragem Gibbs foi monitorada a partir de gráficos de séries temporais das amostras simuladas. Para as meninas e meninos obtivemos, respectivamente, estimadores de Monte Carlo para a média a posteriori de ζ iguais à 31,46 (1,907) e 31,30 (1.596), respectivamente.

Tabela 9: Estimadores Bayesianos para o modelo de crescimento de Gompertz

Gênero	θ_1	θ_2	θ_3
F (meninas)	8,891 (0,1166)	0,9545	0,08503
M(meninos)	9,061 (0,1094)	0,8658	0,07948

(Em parênteses:Estimativa de SE)

É interessante observar que as inferências clássica e Bayesiana assumindo priores não informativas são muito semelhantes, mas os desvios-padrão das estimativas, em geral, são menores para a abordagem bayesiana, especialmente para o parâmetro θ_1.

7.3. Modelo de Gomperz com distribuição estável (procedimento Bayesiano)

Agora, assumimos uma distribuição estável para os pesos intra-uterino do recém-nascidos prematuros (grupo das meninas ou bebês do sexo feminino), considerando o modelo de curva de crescimento de Gompertz .

Sob uma abordagem bayesiana, temos na Tabela 10, os sumários a posteriori de interesse assumindo o modelo de regressão não-linear de Gompertz, isto é, considerando-se o modelo de regressão para o parâmetro localização da distribuição estável dada pela equação,

$$\gamma_i = d_1 exp[-exp(d_2 - d_3 GA_i)]. \qquad (28)$$

Na análise bayesiana desse modelo, usamos o software OpenBugs assumindo as seguintes distribuições anteriores: $\alpha \sim U(0, 2)$, $\beta \sim U(-1, 0)$, $\delta \sim U(0,3)$, $d_1 \sim N(9,1)$, $d_2 \sim N(0.8,1)$ e $d_3 \sim N(0.07,1)$. Observe que estamos usando algumas informações obtidos

dos resultados da análise clássica para escolher as distribuições a priori para os parâmetros da regressão (uso de métodos bayesianos empíricos). Assumimos também uma distribuição uniforme U(-0.5,0.5) para a variável latente Y_i, i = 1,2, ..., 15.

Simulamos 4.000.000 amostras de Gibbs, com um "burn-in-sample" de 300.000 amostras descartadas para eliminar os efeitos dos valores iniciais no processo de simulação iterativo e tendo um tamanho de amostra final de 1000 (cada amostra 200 escolhida a partir de uma amostra simulada de tamanho 200.000). Convergência do algoritmo de amostragem Gibbs foi monitorada a partir de gráficos de séries temporais das amostras simuladas.

Tabela 10. Estimadores Bayesianos para o modelo de crescimento de the Gompertz assumindo uma distribuição estável para as respostas (resultados para as meninas)

Parâmetro	média	D.P.	Intervalo credibilidade 95%
α	1,543	0,00509	(1,536 ; 1,555)
β	-0,8361	0,00883	(-0,8449 ; -0,8204)
δ	0,1418	0,0051	(0,1356 ; 0,1579)
d_1	9,322	0,035	(9,270 ; 9,355)
d_2	0,6948	0,00756	(0,680 ; 0,711)
d_3	0,0714	0,00073	(0,0707 ; 0,0726)

A partir dos resultados da Tabela 10, observa-se que, em geral, as estimativas de Bayesian os parâmetros de regressão (d_1,d_2 and d_3), considerando o modelo de Gompertz com uma distribuição estável são semelhantes para as estimativas de Bayesian os parâmetros de regressão (θ_1, θ_2 e θ_3) considerando o modelo de Gompertz com erros normais, mas em geral os desvios padrão são menores do que assumindo erros normais (ver resultados na Tabela 9).

7.4. Modelo de Gomperz com distribuição estável na presença de outliers ou observações discordantes (procedimento Bayesiano)

Agora, vamos considerar a presença de outliers (bebês do sexo feminino), substituindo o peso intra-uterino 5 (732,58) por 73.200,58 e o primeiro peso intra-uterino (587,00) por 58.700,00. Na Tabela 11, temos os sumários a posteriori obtidos assumindo as mesmas prioris e esquema de simulação assumidos pelos resultados da Tabela 10. Nós também observamos na Tabela 11, que os parâmetros de regressão estimados θ_1, θ_2 e θ_3 do modelo não-linear com Gompertz com erros normais são fortemente afetados pela presença de outliers (ver resultados da Tabela 7).

Considerando-se uma distribuição estável, as estimativas dos parâmetros de regressão (d_1,d_2,d_3) do modelo não-linear de Gompertz estão muito perto das estimativas do modelo de regressão com erros normais, não considerando a presença de outliers. Esta é uma grande vantagem do modelo não-linear assumindo uma distribuição estável, isto é, o modelo de regressão de Gompertz com uma distribuição estável é robusta para a presença de outliers.

Tabela 11 – Estimativas Bayesianas para o modelo de crescimento de Gompertz assumindo uma distribuição estável para as respostas (presença de um outlier)

Parâmetro (stable)	**média**	**D.P.**	**Intervalo de cred 95%**
α	1,543	0,0051	(1,536 ; 1,555)
β	-0,8361	0,0088	(-0,8449; -0,8204)
δ	0,1418	0,0051	(0,1356; 0,1579)
d_1	9,322	0,035	(2,098 ; 3,104)
d_2	0,6948	0,0076	(0,680 ; 0,7115)
d_3	0,0714	0,00073	(0,0707; 0,0726)
Parâm (normal)	**média**	**D.P.**	**Intervalo de cred 95%**
θ_1	9,224	0,2116	(8,825 ; 9,675)
θ_2	0,3107	0,1279	(0,0772; 0,6047)
θ_3	0,0597	0,0072	(0.475 ; 0,0760)

8. Conclusões

O estudo apresentado neste trabalho conseguiu demonstrar bem a importância da metodologia Bayesiana na modelagem de regressão não-linear, neste caso específico, com modelos de crescimento médico. Com estimativas próximas das que conseguimos no método clássico mas com desvio bem menores, mostrando em ambos os exemplos um melhor ajuste quando utilizada a metodologia Bayesiana.

A presença de outliers ou observações discordantes é muito comum em aplicações de análise de regressão não-linear. Na presença destas observações discordantes, as inferências clássicas obtidas sobre os parâmetros de regressão ou as previsões sob a suposição habitual de normalidade para os erros e variância constante podem ser muito afetadas, o que poderia implicar em resultados de inferência erradas. O uso de distribuições estáveis poderia ser uma boa alternativa para muitas aplicações na análise de dados de resultados de inferência robustas, uma vez que esta distribuição tem uma grande flexibilidade de ajuste. Com a utilização de métodos Bayesianos e algoritmos de simulação MCMC, é possível obter inferências para o modelo apesar da não existência de uma forma de análise para a função densidade como foi demonstrado no presente trabalho. É importante ressaltar que o trabalho computacional nas simulações de exemplo para a distribuição a posteriori conjunta de interesse pode ser muito simplificada utilizando softwares livres padrão como o software OpenBugs.

No exemplo ilustrativo relacionada com o crescimento intra-uterino infantil introduzido na seção 7, o uso de técnicas de aumento de dados (ver, por exemplo, Damien et al, 1999), é a chave para obter um bom desempenho para o método de simulação MCMC para aplicações que utilizam distribuições estáveis. Ressalta-se que o uso de software OpenBugs não requer grande tempo de processamento para obter os resumos posteriores de interesse, mesmo quando a simulação de um grande número de amostras de Gibbs são necessários para a convergência do algoritmo. Esses resultados poderiam ser de grande interesse para pesquisadores e profissionais, ao lidar com dados não Gauss, como nas aplicações apresentadas aqui.

9. Referências Bibliográficas

Amaral,L, M. T. R. ; Andrade, M. G. ; Conceição, K. S. ; Stosic, B.(2009) Abordagem Bayesiana para Curva de Crescimento com Restrições nos Parâmetros.. In: XI Escola de Modelos de Regressão, 2009, Recife. XI Escola de Modelos de Regressão.

Bache, C. A., Serum, J. W., Youngs, W. D., and Lisk, D. J. (1972). Polychlorinated biphenyl residues: Accumulation in Cayuga Lake trout with age. Science 117, 1192-1193.

Bates, D.M. & Watts, D.G. (1988). Nonlinear Regression Analysis and Its Applications, Wiley, New York.

Bates, D.M. & Watts, D.G. (1980). Relative curvature measures of nonlinearity (with discussion), Journal of the Royal Statistical Society, Series B **42**, 1–25.

Bernardo,J.M.; Smith,A.F.M. (1994) Bayesian Theory, Wiley.

Borghi,E.;Onis,M;Garza,C;Broeck,J.V;Frongillo,E.A;GrummerStrawn,L;Buuren, S.V;Pan,H; Molinari, L; Martorell,R; Onyango[1].A.W.,Martines' JC (2006) Construction of the World Health Organization child growth standards: selection of methods for attained growth curves,Statistics in Medicine, Volume 25, Issue 2, pages 247–265.

Box,G.E.P., Tiao,G.C. (1973) Bayesian inference in statistical analysis,.Reading: Addison-Wesley.

Buckle, D.J. (1995) "Bayesian inference for stable distributions", Journal of the American Statistical Association, Vol. 90, pp. 605-613.

Casella, G.; George, E. I. Explaining the Gibbs Sampler.(1992) The American Statistician, v. 46, n. 3, p. 167-174.

Cuervo, E. C. ; Achcar, J. A.(2010) . Heteroscedastic Nonlinear Regression Models. Communications in Statistics. Simulation and Computation, v. 39, p. 405-419.

Chen, Z. & Wang, Y. G. (2004), `Efficient Regression Analysis with Ranked-Set Sampling´, *Biometrics* 60(4), 997-1004.

Chib, S.; Greenberg, E. (1995). Understanding the Metropolis- Hastingsalgorithm, The American Statistician, 49, 327-335.

Chou,C.P; Peter M. Bentler , P.M.; Pentz,M.A. (1998) Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis, Structural Equation Modeling: A Multidisciplinary Journal, **Volume 5**, **Issue 3**.
DOI: 10.1080/10705519809540104

Davidian, M.; Giltinan. D.M. (1996) Nonlinear models for repeated measurement data. 2.ed. London: Chapman Hall, 359p.

Draper, N.R.; Smith, H. (1980) Applied regression analysis. 2.ed.New York: Wiley, 709p.

Damien, P.; Wakefield, J. and Walker,S. (1999) "Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables", *Journal of the Royal Statistical Society*, Series B, Vol. 61, pp. 331-344.

Feller,W. (1971) "An introduction to probability theory and its applications", Vol. II. John Wiley, New York.

Ferreira, I. E. P. ; Trinca, L. A. ; Ferreira, C. P. . Delineamentos Experimentais Eficientes para Estudos de Cinética Química. Química Nova (Impresso), v. 37, p. 589, 2014.

Francis, D. J.; Shaywitz, S. E.; Stuebing, K. K.; Shaywitz, B. A.; Fletcher, J. M.(1996) Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis.Journal of Educational Psychology, Vol 88(1), 3-17. http://dx.doi.org/10.1037/0022-0663.88.1.3

Freedman, L. S., Fainberg, V., Kipnis, V., Midthune, D. & Carroll, R. J. (2004), `A new Method for Dealing with Measurement Error in Explanatory Variables of Regression Models´, *Biometrics* 60(1), 172-181.

Gelfand, A. E.; Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities, Journal of the American Statistical Association, 85, 398-409.

Gnedenko, B. V. and A. N. Kolmogorov, A.N. (1968) "Limit distributions for sums of independent random variables", Addison-Wesley, Massachussetts.

Gompertz, B. (1825) On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life ContingenciesPhil. Trans. R. Soc. Lond. January 1, 1825 115 513-583.

Horton,W.A.; Hall,J.G.; Scott,C.I.; Pyeritz, R.E.; Rimoin, D.L.(1982) Growth Curves for Height for Diastrophic Dysplasia, Spondyloepiphyseal Dysplasia Congenita, and Pseudoachondroplasia, American Journal of Dis Child.,136(4):316-319. doi:10.1001/archpedi.1982.03970400034010.

Horton, W.A.; RotterJ.I; Rimoin, D.L.;Scott,C.I.; Hall, J.G. (1978) , Standard growth curves for achondroplasia,The Journal of pediatrics, 93,3,435-438.

Ibragimov; K. E. Cernin. (1959) On the unimodality of stable laws, 453-456.

Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B (1983) Changes in the normal maximal expiratory flow-volume curve with growth and aging,The American Review of Respiratory Disease 127(6):725-734.

Lévy, P (1924) "Théorie des erreurs la loi de Gauss et les lois exceptionelles", Bulletin Society Mathematical, Vol. 52, pp. 49-85.

Lukacs,E. (1970) "Characteristic Functions", Hafner Publishing, New York.

Maršál, K.: Persson,P.H.; Larsen,T.;Lilja,H; Selbing,A.; B Sultan,B. (1996) Intrauterine growth curves based on ultrasonically estimated foetal weights, Acta Paediatrica, Volume 85, Issue 7, pages 843–848.DOI: 10.1111/j.1651-2227.1996.tb14164.x

Michaelis, L. & Menten, M. Die Kinetik der Invertinwirkung. Biochem. J., 49:333-369,1913.

Montgomery,D.C.; Runger, G.C.(2011) Applied statistics and probability for engineeers, fifty edition, New York: Wiley.

Nicosia RF, Ottinetti A (1990) Growth of microvessels in serum-free matrix culture of rat aorta. A quantitative assay of angiogenesis in vitro. Laboratory Investigation; a Journal of Technical Methods and Pathology , 63(1):115-122.

Nolan, J. P. (2009) “Stable Distributions - Models for Heavy Tailed Data”, Birkhäuser, Boston. In progress, Chapter 1 online at academic2.american.edu/~jpnolan.

Olsen,I.E.; LDN,. Groveman,S.A.; Lawson, L.; Clark,R.H.; Zemel,B.S. (2010) New Intrauterine Growth Curves Based on United States Data,Pediatrics, doi: 10.1542/peds.2009-0913.

Palisano RJ, Walter SD, Russell DJ, Rosenbaum PL, Gémus M, Galuppi BE, Cunningham L. Gross motor function of children with Down syndrome: creation of motor growth curves. Arch Phys Med Rehabil 2001;82:494-500.

Paulino, C.D.; Turkman, M.A.A.; Murteira,B. (2003) Estatística Bayesiana. Fundação Calouste Gulbenkian, Lisboa.

Paz, C.C.P.; Packer, I.U.; Freitas, A.R. et al. (2004) Ajuste de modelos não-lineares em estudos de associação entre polimorfismos genéticos e crescimento em gado bovino de corte. Revista Brasileira de Zootecnia, v.33, n.6, p.1416-1425.

Ratkowsky, D.A. (1983). Nonlinear Regression Modelling: A Unified Practical Approach, Marcel Dekker, New York.

Roussas,G.G. (2005) Measure-Theoretic probability, San Diego:Elsevier.

Sakamoto,S. , Sakanoue,M. (1985) New Fetal Growth Curves for Japanese, Acta Paediatrica, Volume 74, Issue Supplement 319, pages 62–67.

Sandland, R.L.; McGilvhrist, C.A. (1979) Stochastic growth curve analysis.Biometrics, v.35, n.1, p.255-271.

Seber, G.A.F.; Lee,A.J. (2003). Linear regression analysis, second edition, Wiley series in probability and mathematical statistics.

Seber, G.A.F. & Wild, C.J. (1989) Nonlinear Regression. Probability and Mathematical Statistics. Wiley, New York.

Silva, N.A.M.; Lima, R.R.; Silva, F.F.E. *et al.* Modelo Hierárquico Bayesiano aplicado na avaliação genética de curvas de crescimento de bovinos de corte. *Arq. Bras. Med. Vet. Zootec.*, v.62, p.647-654, 2010.

Souza, V.J. (2011) Fatores de redução da mortalidade infantil em Franca SP em 2006: a contribuição do Ambulatório de Crianças de Alto-Risco e da Pastoral da Criança. 2008. Faculdade de Medicina de Ribeirão Preto, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Skorohod, A.V. (1961) “On a theorem concerning stable distributions”. In: Select. Transl. Math. Statist. and Probability, Vol. 1, Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., pp.169-170.

Spiegelhalter, D.J.; Thomas, A.; Best, N.G.; and Lunn, D. (2003) “WinBugs User's Manual”, MRC Biostatistics Unit, Cambridge.

Spiegelhalter, D.J., Best, N. G., Carlin, B. P., van der Linde, A. (2002) Bayesian measures of model complexity and fit, Journal of the Royal Statistical Society, Series B, 64 (4), 583–639.

Tanner. M.A. and Wong,W.H. (1987) "The calculation of posterior distributions by data augmentation", Journal of American Statistical Association, Vol. 82, pp. 528-550.

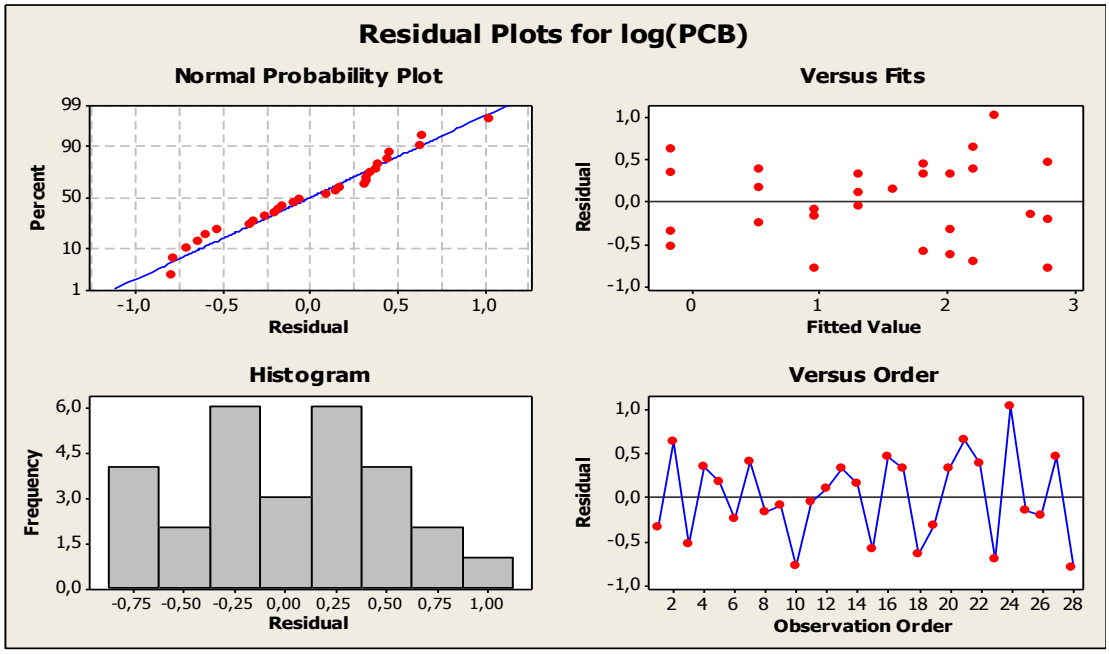
Verbeke, G., Molenberghs, G. (1997) Linear mixed models in practice: a SAS-oriented approach. Cary: SAS Institute, 306p.

Xavier, L. H. (2000) Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação. 91f. Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

Apêndice 1 - Gráficos de resíduos para os modelos de crescimento ajustados (dados da Tabela 1)

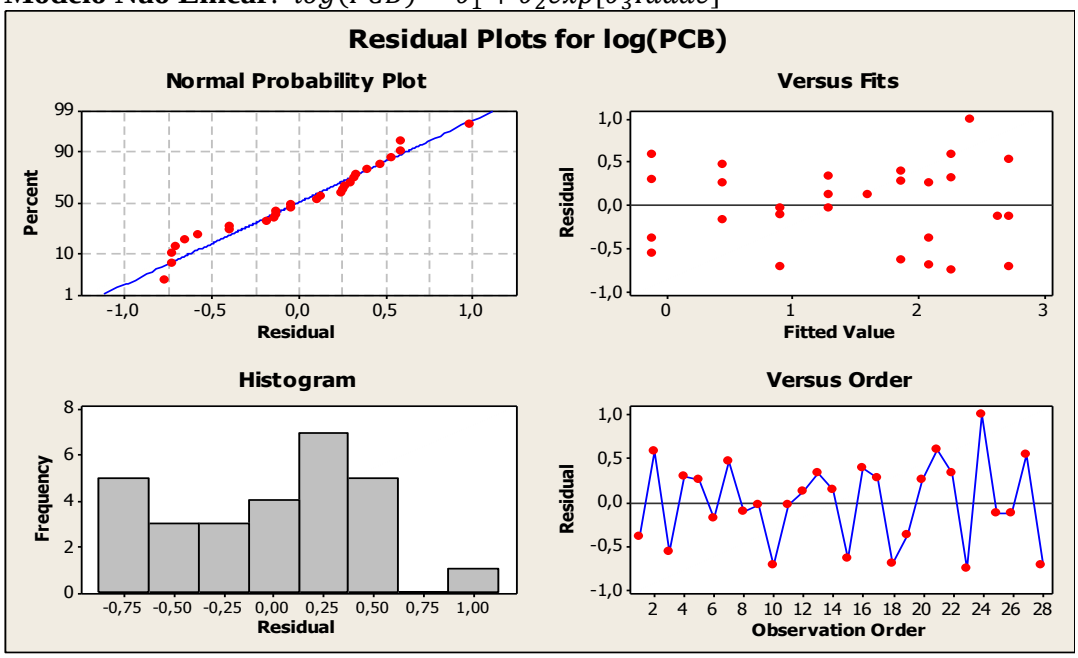
- **Modelo 1**

Modelo Não Linear: $log(PCB) = \theta_1 + \theta_2 Idade^{\theta_3}$



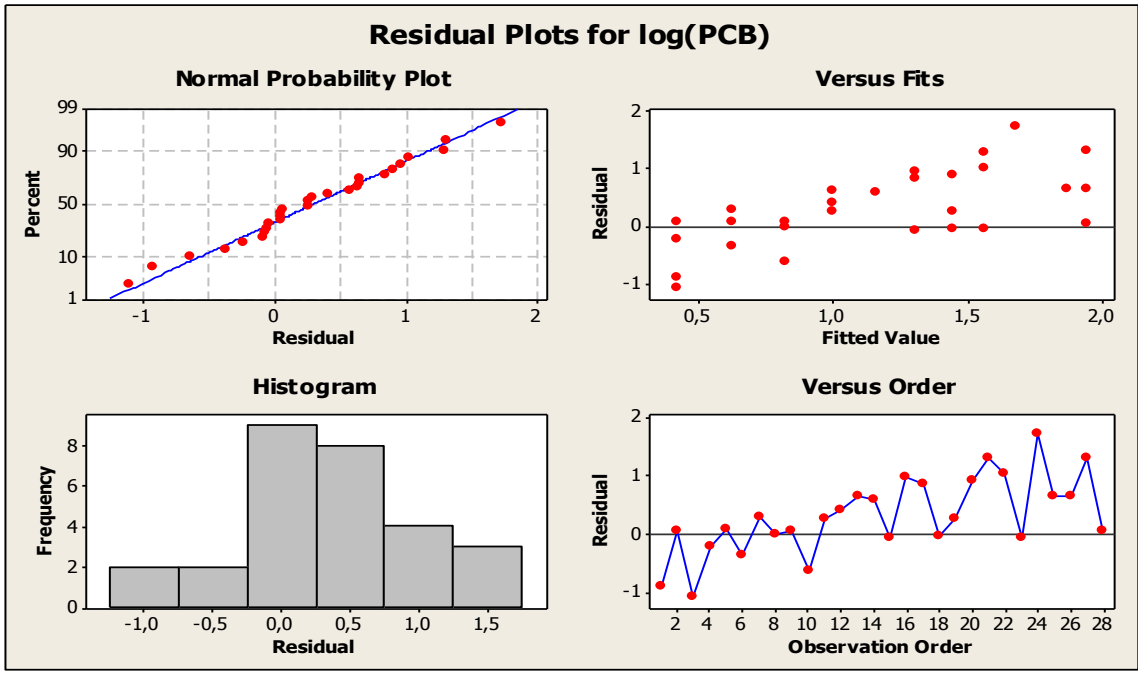
- **Modelo 2- Modelo de crescimento exponencial**

Modelo Não Linear: $log(PCB) = \theta_1 + \theta_2 exp[\theta_3 Idade]$



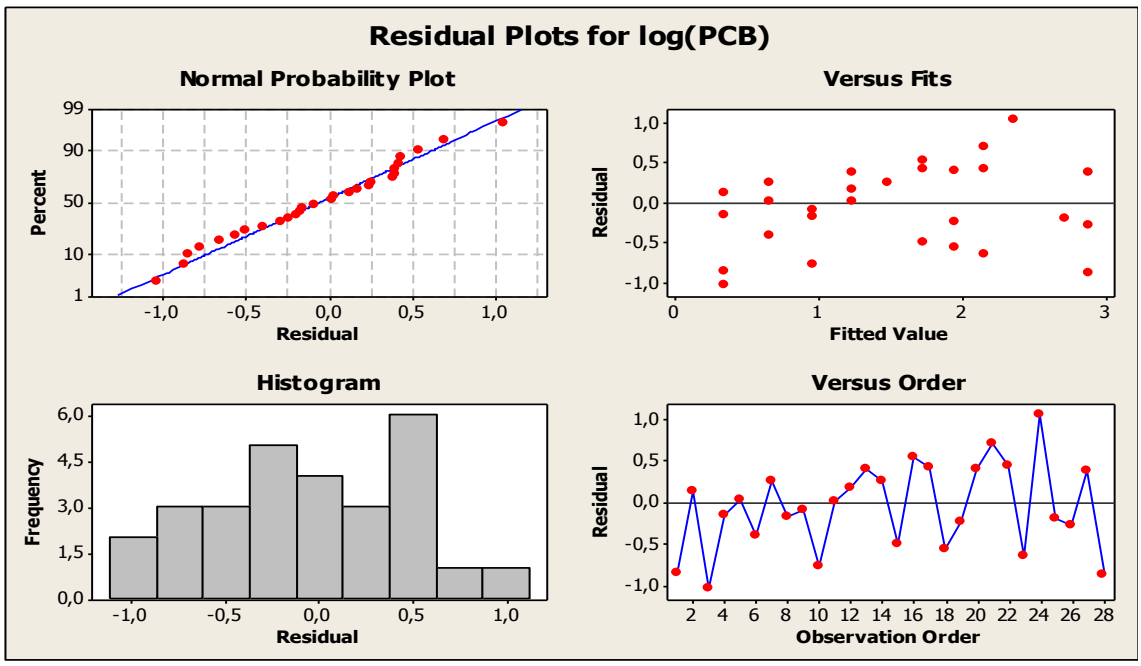
- **Modelo 3**

Modelo Não Linear: $log(PCB) = \theta_1 exp[\theta_2 Idade] + \theta_3 \exp[-\theta_4 * \text{Idade}]$



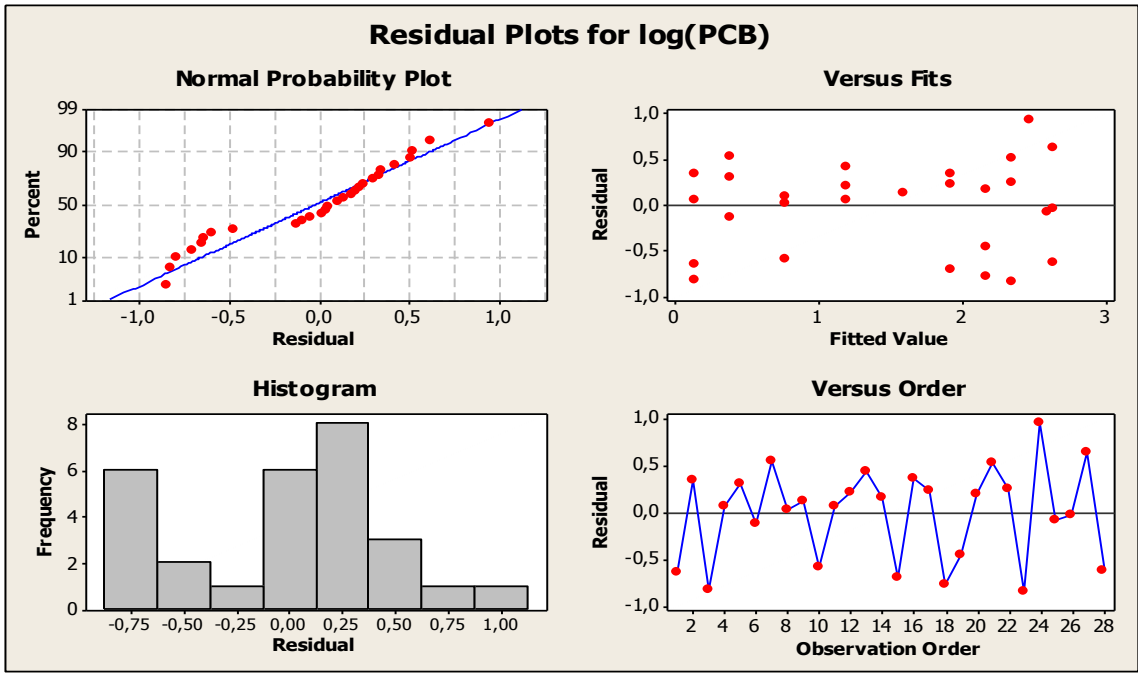
- **Modelo 4 (Michaelis-Menten)**

Modelo Não Linear: $log(PCB) = \theta_1 Idade/[\theta_1 + Idade]$



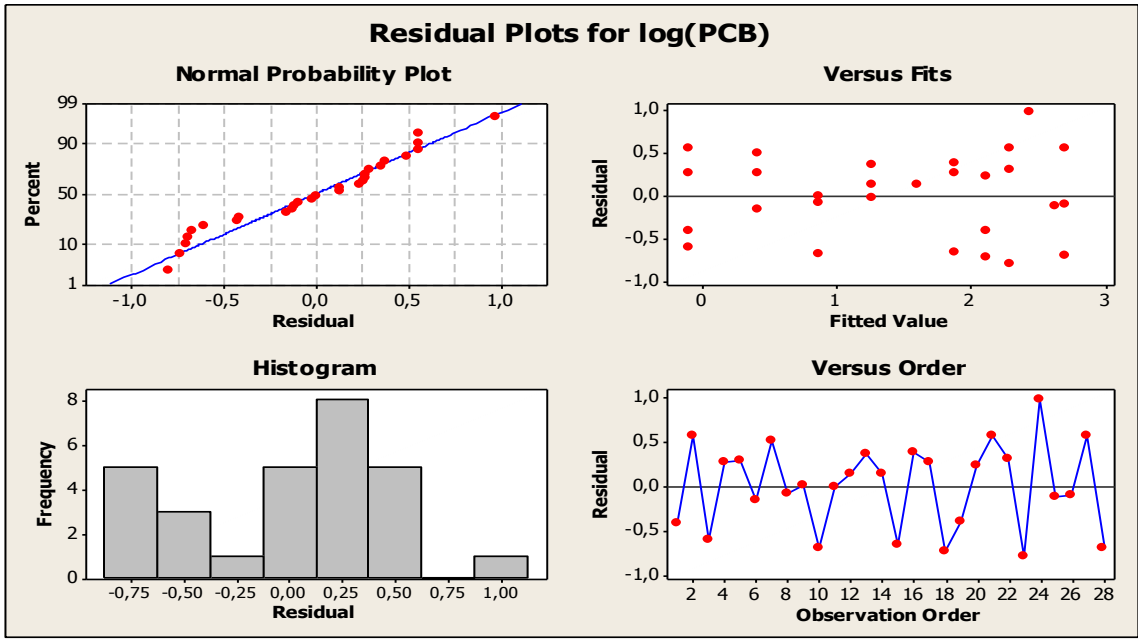
- **Modelo 5 (Gompertz)**

Modelo Não Linear: $log(PCB) = \theta_1 exp[-exp(\theta_2 - \theta_3 idade)]$



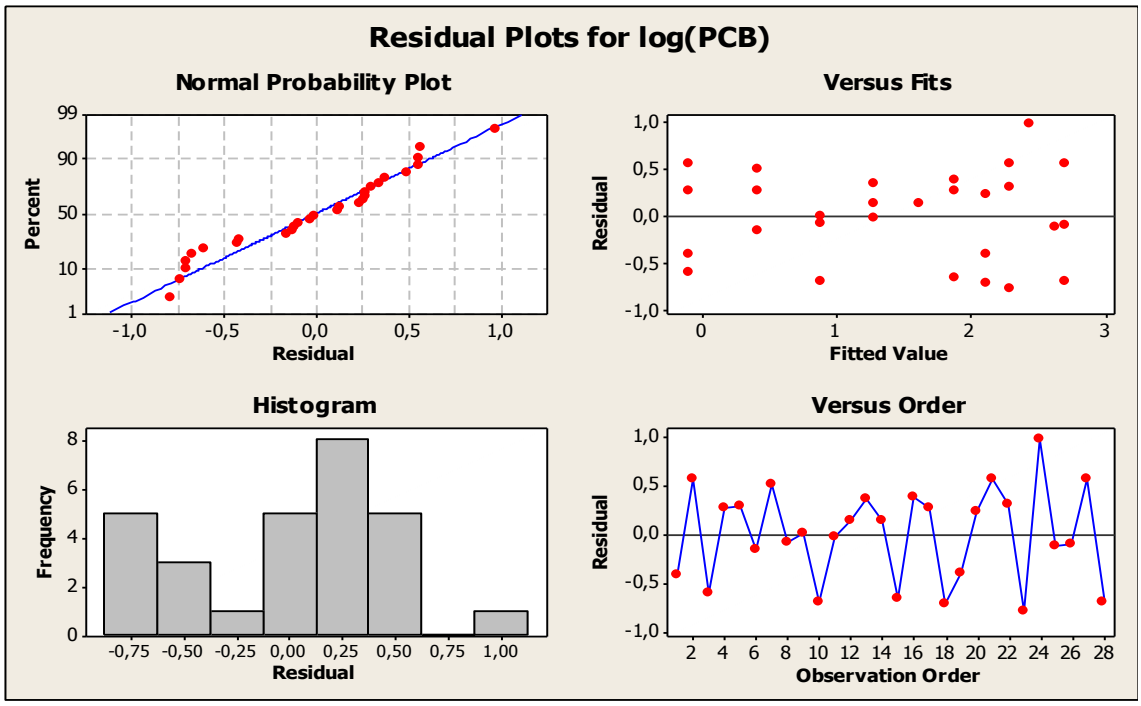
- **Modelo 6 (Logístico)**

Modelo Não Linear: $log(PCB) = \theta_1 + (\theta_2 - \theta_1)/[1 + exp((Idade - \theta_3)/\theta_4)]$



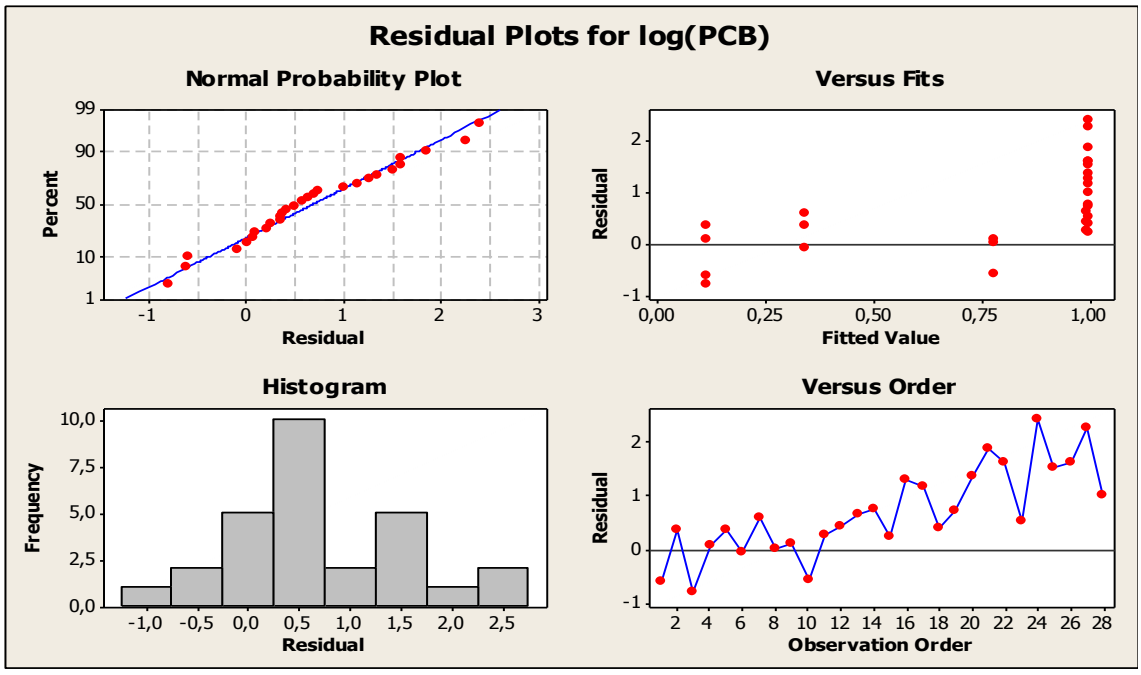
- **Modelo 7 (Weibull)**

Modelo Não Linear: $log(PCB) = c + (\theta_2 - \theta_1) * exp\left[-\theta_3 Idade^{\theta_4}\right]$



- **Modelo 8 (sigmoid)**

Modelo Não Linear: $log(PCB) = 1 - exp[-exp(\theta_1 - \theta_2 Idade)]$



Apêndice 2 - Programas OpenBugs (dados da Tabela 1)

- **Modelo 1**

nonlinear model: $\log(PCB) = \theta_1 + \theta_2 \, Age^{\theta 3}$

```
model {
        for (i in 1:N) {
        y[i] <- log(PCB[i])
        y[i] ~ dnorm(mu[i],tau)
        mu[i] <- theta1+theta2*pow(age[i],theta3)
        }
        theta1 ~ dnorm(0, 1)
        theta2 ~ dnorm(0, 1)
        theta3 ~ dnorm(0, 1)
        tau ~ dgamma(1,1)
        sigma <- 1 / sqrt(tau)
}
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 2 -(exponential growth model)**

nonlinear model: $\log(PCB) = \theta_1 + \theta_2 \exp[\theta_3 * Age]$

```
model {
        for (i in 1:N) {
        y[i] <- log(PCB[i])
        y[i] ~ dnorm(mu[i],tau)
        mu[i] <- theta1+theta2*exp(theta3* age[i])
        }
        theta1 ~ dnorm(0, 1)
        theta2 ~ dnorm(0, 1)
        theta3 ~ dnorm(0, 1)
        tau ~ dgamma(1,1)
        sigma <- 1 / sqrt(tau)
}
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 3**

nonlinear model: log(PCB) = θ_1 exp[-θ_2*Age]+ θ_3 exp[-θ_4*Age]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
mu[i] <- theta1*exp(-theta2* age[i])+ theta3*exp(-theta4* age[i])
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    theta3 ~ dnorm(0, 1)
                    theta4 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
                    sigma <- 1 / sqrt(tau)
          }
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 4 (Michaelis-Menten)**

nonlinear model: log(PCB) = θ_1 *Age/[θ_2 +Age]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
          mu[i] <- (theta1 * age[i])/( theta2+ age[i])
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
```

```
                    sigma <- 1 / sqrt(tau)
}
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 5 (Gompertz)**

nonlinear model: log(PCB) = θ_1*exp[-exp($\theta_2 - \theta_3$*Age)]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
          mu[i] <- theta1*exp(-exp( theta2-theta3* age[i]))
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    theta3 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
                    sigma <- 1 / sqrt(tau)
}
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 6 (Logistic)**

nonlinear model: log(PCB) = θ_1+ ($\theta_2 - \theta_1$)/[1+exp((Age- θ_3)/ θ_4)]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
        mu[i] <- theta1+(theta2-theta1)/(1+exp(( age[i]-theta3)/theta4))
```

```
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    theta3 ~ dnorm(0, 1)
                    theta4 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
                    sigma <- 1 / sqrt(tau)
          }
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 7 (Weibull)**

nonlinear model: log(PCB) = θ_1+ (θ_2 – θ_1)*exp[-θ_3*Age$^{\theta 4}$]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
     mu[i] <- theta1+(theta2-theta1)*exp(-theta3*pow( age[i],theta4))
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    theta3 ~ dnorm(0, 1)
                    theta4 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
                    sigma <- 1 / sqrt(tau)
          }
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

- **Modelo 8 (sigmoid)**

nonlinear model: log(PCB) = 1-exp[-exp($\theta_1 - \theta_2$*Age)]

```
model {
                    for (i in 1:N) {
                    y[i] <- log(PCB[i])
                    y[i] ~ dnorm(mu[i],tau)
          mu[i] <- 1-exp(-exp(theta1-theta2*age[i]))
                    }
                    theta1 ~ dnorm(0, 1)
                    theta2 ~ dnorm(0, 1)
                    tau ~ dgamma(1,1)
                    sigma <- 1 / sqrt(tau)
     }
```

```
list(PCB =
c(0.6,1.6,0.5,1.2,2.0,1.3,2.5,2.2,2.4,1.2,3.5,4.1,5.1,5.7,3.4,9.7,8.6,4.0,5.5,10.5,17.5,13.4,
4.5,30.4,12.4,13.4,26.2,7.4),age =
c(1,1,1,1,2,2,2,3,3,3,4,4,4,5,6,6,6,7,7,7,8,8,8,9,11,12,12,12), N = 28)
```

Apêndice 3. Artigo publicado na Revista Brasileira de Biometria (*Rev. Bras. Biom.*, São Paulo, v.33, n.2, p.170-183, 2015)

GOMPERTZ GROWTH CURVES ASSUMING STABLE DISTRIBUTIONS: AN APPLICATION TO INTRAUTERINE GROWTH FOR PRETERM INFANTS

Jorge Alberto ACHCAR[1]
Breno Raphael Gomes de OLIVEIRA[1]

***ABSTRACT**: In this paper we consider the use of Bayesian methods to analyze some standard existing growth models, a class of nonlinear regression models. For the nonlinear modeling we assume usual normal errors and also stable distributions for the response variable. We also study some robustness aspects of nonlinear regression models to the presence of outliers or discordant observations considering the use of stable distributions for the response in place of the usual normality assumption. It is well known that, in general, there is no closed form for the probability density function of stable distributions. However, under a Bayesian approach, the use of a latent or auxiliary random variable gives some simplification to obtain any posterior distribution when related to stable distributions. To show the usefulness of the computational aspects, the methodology is applied to an example related to the intrauterine growth curves for preterm infants. Posterior summaries of interest are obtained using MCMC (Markov Chain Monte Carlo) methods and the OpenBugs software.*

KEYWORDS: Stable distribution; Bayesian analysis; nonlinear regression models; MCMC methods; OpenBugs software.

Introduction

Growth curves are included in a class of nonlinear models widely used in biology to model different problems as the population size or biomass (in population ecology and demography, for population growth analysis) or the individual body height or biomass (in physiology, for growth analysis of individuals). A growth curve is an empirical model of the evolution of a quantity over time. Growth curves are employed in many disciplines besides biology, particularly in statistics, where there is a great literature on this subject related to nonlinear models. Under a more probabilistic and mathematical statistics approach, growth curves are often modeled as being a continuous stochastic processes.

Standard classical inference methods to get point or interval estimates for the parameters of growth curves are presented within the nonlinear modeling methodology.

Nonlinear regression methodology is similar to the linear regression methodology, that is, a modeling approach to relate a response X to a vector of covariates, $\mathbf{v} = (v_1, \ldots, v_k)^T$. Different of linear models, nonlinear regression is characterizedby the fact that the prediction equationdepends nonlinearly on one or more unknown parameters.

Different of linear regression methodology often used forbuilding a purely empirical model, nonlinear regression methodology usually arises when there is some physical reason which implies that the relationship between the responseand the predictors follows a particular functional form.

A nonlinear regression model has the general form,

$$X_i = f(\mathbf{v}_i, \boldsymbol{\theta}) + \varepsilon_i, \qquad (1)$$

where $X_i$are the responses, i =1,…,n; f is a known function of the covariate vector ; $\mathbf{v_i} = (v_{i1}, \ldots, v_{ik})^T$ is a vector of k covariates or independent variables; $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^T$ is the vector of p parameters and ε_i are random errors. The errors $\varepsilon_i$are usually assumed to beuncorrelated with a normal distribution with mean zero and constant variance.

[1]Universidade de São Paulo - USP, Faculdade de Medicina de Ribeirão Preto, Departamento de Medicina Social, CEP: 14049-900, Ribeirão Preto, SP, Brasil. E-mail: *achcar@fmrp.usp.br*; *brenooliveira@usp.br*

The most popular criterion to estimate the p parameter vector $\boldsymbol{\theta}$ in the nonlinear model (1) is to find estimates for the parameters (nonlinear least squares) which minimizes the sum of squared errors, given by,

$$\sum_{i=1}^{n}[x_i - f(\mathbf{v}_i, \boldsymbol{\theta})]^2 \qquad (2)$$

Remark: If the errors ε_i follow a normal distribution, then the least squares estimator for $\boldsymbol{\theta}$ is also the maximum likelihood estimator.

Usually, nonlinear regression estimates must be computed by iterative procedures using optimization methods to minimize the sum of squared errors expression (2). It is important to point out that the definition of nonlinearity is related to the unknown parameters and not to the relationship between the covariates and the response. As an example, $X = \beta_0 + \beta_1 v + \beta_2 v^2 + \varepsilon$, is considered as a linear model (see for example, Bates and Watts,1988; Ratkowsky, 1983; Seber and Wild,1989).

A popular iterative technique to find the least squares estimator of nonlinear models is the Gauss–Newton algorithm. The Gauss–Newton algorithm increments the working estimate $\boldsymbol{\theta}$ at each iteration by an amount equal to the coefficients from the linear regression of the current residuals $\mathbf{e}$ on the current gradient matrix $\mathbf{X}$.

If the function f in (1) is continuously differentiable in $\boldsymbol{\theta}$, then it can be linearized locally as,

$$f(\mathbf{v}, \boldsymbol{\theta}) = f(\mathbf{v}, \boldsymbol{\theta}_0) + V_0(\boldsymbol{\theta} - \boldsymbol{\theta}_0), \qquad (3)$$

where $\mathbf{V}_0$ is the n x p gradient matrix with elements $\partial f(\mathbf{v}_i, \theta_0)/\partial\theta_j$ and $\boldsymbol{\theta}_0$ is a vector of initial values for the iterative procedure. This leads to the Gauss–Newton algorithm for estimating $\boldsymbol{\theta}$,

$$\theta_1 = \theta_0 + (V_0^T V_0)^{-1} V_0^T \mathbf{e}, \qquad (4)$$

where $\mathbf{e}$ is the vector of working residuals $x_i - f(\mathbf{v}_i, \theta_0)$.

If the errors ε_i are independent and normally distributed $N(0, \sigma^2)$, then the Gauss–Newton algorithm is an application of Fisher's method of scoring. This algorithm is implemented in many existing statistical softwares as for example, R, Minitab version 16 or SAS.

If $\mathbf{X}$ is of full column rank in a neighborhood of the least squares solution, then it can be shown that the Gauss–Newton algorithm will converge to the solution from a sufficiently good starting value. In practical work, there is no guarantee, though, that the algorithm will converge from values further from the solution. Some improvement of Gauss–Newton algorithm are given in the literature as the Levenberg–Marquart damping algorithm (see for example, Seber and Wild, 1989).

Standard inferences for the parameters of nonlinear models are obtained from the asymptotical normality of the least squares estimators $\widehat{\boldsymbol{\theta}}$ with mean θ and variance-covariance matrix $\sigma^2(\mathbf{V}^T\mathbf{V})^{-1}$where the variance σ^2 is usually estimated by,

$$s^2 = \{ \sum_{i=1}^{n}[x_i - f(v_i, \widehat{\boldsymbol{\theta}})]^2 \}/(n - p). \qquad (5)$$

It is important to point out that since most asymptotic inference for nonlinear regression models is based on analogy with linear models, and since this inference is only approximated as the actual model differs from a linear model, various measures of nonlinearity have been proposed in the literature to verify how good linear approximations are likely to be in each case. One class of measures focuses on curvature (intrinsic curvatures) of the function f and it is based on the sizes of the second derivatives of f (see for example, Bates and Watts,1980). Other class of measures: intrinsic curvature defined by the residuals $u_i = x_i - f(\mathbf{v}_i, \widehat{\boldsymbol{\theta}})$ or parameter effect curvatures (see for example, Bates and Watts,1988).

A good alternative to get accurate inferences and predictions for non-linear models is the use of Bayesian methods especially using MCMC (Markov Chain Monte Carlo) methods as the popular Gibbs sampling algorithm (see for example, Gelfand and Smith, 1990; or Casella and George,1992) or the Metropolis-Hastings algorithm (see for example, Chib and Greenberg,1995).

The paper is organized as follows: in section 2, we introduce some special growth functions: the use of stable distribution for the response; in section 3, we introduce a Bayesian approach for nonlinear models assuming a stable distribution; in section 4, we present an example with an intrauterine growth data for preterm infants; in section 5, we present the intrauterine growth curves for preterm infants in presence of outliers; finally in section 6, we present some concluding remarks.

Some special growth functions: the use of stable distribution for the response

In many applications, the systematic part of the response is known to be monotonic increasing in v, where v might represent time or dosage. Nonlinear regression models with this property are called growth models. The simplest growth model is the exponential growth model,

$$f(v, \theta) = \theta_1 exp(-\theta_2 v), \qquad (6)$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2)$.

Many other growth models are introduced in the literature. Special cases are given by:

(i) $f(v, \boldsymbol{\theta}) = \theta_1 + \theta_2 v^{\theta_3}$, where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$.
(ii) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 + \vartheta_2 exp[\vartheta_3 v]$ (exponential growth model), where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3)$.
(iii) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 exp[-\vartheta_2 v] + \vartheta_3 exp[-\vartheta_4 v]$, where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)$.
(iv) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 v/[\vartheta_2 + v]$ (Michaelis-Menten), where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$.
(v) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 exp[-exp(\vartheta_2 - \vartheta_3 v)]$ (Gompertz), where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3)$.
(vi) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 + (\vartheta_2 - \vartheta_1)/[1+exp((v - \vartheta_3)/\vartheta_4)]$ (logistic), where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)$.
(vii) $f(v, \boldsymbol{\vartheta}) = \vartheta_1 + (\vartheta_2 - \vartheta_1) exp[-\vartheta_3 v^{\vartheta_4}]$ (Weibull), where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)$.
(viii) $f(v, \boldsymbol{\vartheta}) = 1-exp[-exp(\vartheta_1 - \vartheta_2 v)]$(sigmoid),where $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$.

For all cases, v is a covariate. Observe that these different models could be considered to model growth curves. In many cases we have physical interpretations for the use of a particular model. In other situations we try different growth models to decide by the best one to be fitted by the data.

In this paper we explore a particular growth model: the Gompertz model.

In many cases the usual assumption of normality for the errors in (1) could not be appropriated. This situation occurs for example, when we have discordant observations which could greatly affect the obtained inferences. In this way, we could assume more robust distributions for the data, as the stable distribution.

Other point of interest: the use of asymptotical inference results could not be accurate depending on the sample sizes and the intrinsic curvature of the function f in (1).

In this way, we propose the use of stable distributions for the response X_i in (1).

A wide class of distributions that encompasses the Gaussian distribution is given by the class of stable distributions. This large class defines location-scale families that are closed under convolution. The Gaussian distribution is a special case of this distribution family (see for instance, Nolan, 2009), described by four parameters α, β, γ and δ. The $\alpha \in (0,2]$ parameter defines the "fatness of the tails", and when $\alpha = 2$ this class reduces to Gaussian distributions. The $\beta \in [-1,1]$ is the skewness parameter and for $\beta = 0$ one has symmetric distributions. The location and scale parameters are, respectively, $\gamma \in (-\infty, \infty)$ and $\delta \in (0, \infty)$ (see Levy,1924).

Stable distributions are usually denoted by $S_\alpha(\beta, \gamma, \delta)$. If a random variable $X \sim S_\alpha(\beta, \gamma, \delta)$, then $Z = \frac{X-\gamma}{\delta} \sim S_\alpha(\beta, 0,1)$ (see Lukacs,1970 and Nolan, 2009).

A great difficulty associated to stable distributions $S_\alpha(\beta, \gamma, \delta)$, is that in general there is no simple closed form for their probability density functions. However, it is known that the probability density functions of stable distributions are continuous (Gnedenko and Kolmogorov,1968; Skorohod,1961) and unimodal (Ibragimov and Černin,1959) ; Kanter, 1976). Also the support of all stable distributions is given in $(-\infty, \infty)$, except for $\alpha < 1$ and $|\beta| = 1$ when the support is $(-\infty, 0)$ for $\beta = 1$ and $(0, \infty)$ for $\beta = -1$ (see Feller,1971).

The characteristic function Φ (.) of a stable distribution is given by,

$$log[\Phi(t)] = \begin{cases} i\gamma t - |\delta t|^\alpha [1 - i\beta sign(t)\, tan(\frac{\pi\alpha}{2})], & for\ \alpha \neq 1 \\ i\gamma t - |\delta t|^\alpha [1 + i\beta sign(t)\, (\frac{2}{\pi}) log(t)], & for\ \alpha = 1 \end{cases}, \qquad (7)$$

where $i = \sqrt{-1}$ and *sign* (.) function is given by

$$sign(x) = \begin{cases} -1, & if\ x < 0 \\ 0, & if\ x = 0 \\ 1, & if\ x > 0. \end{cases},$$

It is important to point out that if $\alpha < 1$, the variance is infinite and the mean of the stable distribution does not exists.

Although this class of distributions is a good alternative for data modeling in different areas, we usually have difficulties to obtain estimates under a classical inference approach due to the lack of closed form expressions for their probability density functions. One possibility in applications, is to get the probability density function from the inversion formula (see, for example Roussas, 2005),

$$f(x) = 1/2\pi \int_{-\infty}^{\infty} e^{-itx}\, \Phi(t)dt, \qquad (8)$$

where $\Phi(t)$is the characteristic function. In applications, we need use numerical methods to solve the integral in (8), usually taking a great computational time.

An alternative is the use of Bayesian methods. However, the computational cost can be further high to get the posterior summaries of interest. An alternative is to use latent or artificial variables that could improve the simulation computation of samples of the joint posterior distributions of interest (see, for example Damien, et al 1999; Tanner and Wong, 1987).

In this way, a Bayesian analysis of stable distributions was introduced by Buckle (1995) using Markov Chain Monte Carlo (MCMC) methods and latent variables (see also, Achcar et al, 2013). The use of Bayesian methods with MCMC simulation can have great flexibility by considering latent variables where samples of latent variables are simulated in each step of the Gibbs or Metropolis-Hastings algorithms.

Considering a latent or an auxiliary variable, Buckle (1995) proved a theorem that is useful to simulate samples of the joint posterior distribution for the parameters α, β, γ and δ. This theorem establishes that a stable distribution for a random variable *Z* defined in (- ∞, ∞) is obtained as the marginal of a bivariate distribution for the random variable *Z* itself and an auxiliary random variable *Y*. This variable *Y* is defined in the interval (-0.5,$a_{\alpha,\beta}$), when *Z* ϵ (- ∞, 0), and in ($a_{\alpha,\beta}$, 0.5), when *Z* ϵ (0, ∞). The quantity $a_{\alpha,\beta}$ is given by,

$$a_{\alpha,\beta} = -\frac{b_{\alpha,\beta}}{\alpha\pi}, \qquad (9)$$

where $b_{\alpha,\beta} = \beta \min\{\alpha, 2-\alpha\}\frac{\pi}{2}$.

The joint probability density function for random variables *Z* and *Y* is given by

$$f(z,y|\alpha,\beta) = \frac{\alpha}{|\alpha-1|} exp\ \{-|\frac{z}{t_{\alpha,\beta}(y)}|^{\theta}\}\ |\frac{z}{t_{\alpha,\beta}(y)}|^{\theta}\frac{1}{|z|}, \qquad (10)$$

where $\theta = \frac{\alpha}{\alpha-1}$,

$$t_{\alpha,\beta}(y) = \left(\frac{sin(\pi\alpha\, y+b_{\alpha,\beta})}{cos(\pi\, y)}\right)\left(\frac{cos(\pi\, y)}{cos(\pi(\alpha-1)y+b_{\alpha,\beta})}\right)^{\frac{1}{\theta}}, \qquad (11)$$

and

$$Z = \frac{X-\gamma}{\delta}, \text{ for } \delta \neq 0.$$

From the bivariate density (10), Buckle (1995) shows the marginal distribution for the random variable *Z* is a stable $S_\alpha(\beta, 0, 1)$ distribution. Usually, the computational costs to obtain posterior summaries of interest using MCMC methods is high for this class of models, which could give some limitations for practical applications. One problem can be the simulation algorithm convergence. In this paper, we propose the use of a popular free available software to obtain the posterior summaries of interest: the OpenBugs software (Spiegelhalter et al, 2003).

A Bayesian approach assuming a stable distribution

In this section, let us assume that the response x_i in the nonlinear regression model (1) for i = 1,…, n, have a stable distribution $X_i \sim S_\alpha(\beta, \gamma_i, \delta)$, that is, $Z_i = \frac{X-\gamma_i}{\delta} \sim S_\alpha(\beta, 0, 1)$ and where the location parameter γ_i of the stable distribution is related to k explanatory variables $v_i = (v_{1i}, v_{2i},..., v_{ki})$ by a nonlinear relation given by,

$$\gamma_i = f(v_i, \boldsymbol{d}), \qquad (12)$$

where $\mathbf{d} = (d_1,d_2,\ldots,d_k)$ are the associated unknown regression parameters. Assuming a joint prior distribution for α, β , d andδ, where $\mathbf{d} = (d_1,d_2,\ldots,d_k)$ given by $\boldsymbol{\pi_0}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{d}, \boldsymbol{\delta})$, Buckle (1995) shows that the joint posterior distribution for parameters α, β , d andδ, is given by

$$\boldsymbol{\pi}\,(\alpha,\beta,\boldsymbol{d},\delta/\boldsymbol{x}) \propto \int \left(\frac{\alpha}{|\alpha-1|\delta}\right)^n exp\left\{-\sum_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(Y_i)}\right|^{\theta}\right\}\prod_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta}. \qquad (13)$$

$$.\ \frac{1}{|z_i|}\pi_0(\alpha,\beta,\boldsymbol{d},\delta)\mathrm{d}\mathbf{y},$$

where $\theta = \frac{\alpha}{\alpha-1}$, $z_i = \frac{x_i - \gamma_i}{\delta}$, for $i = 1,\ldots,n$, $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ and $\delta \in (0,\infty)$; $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ are respectively, the observed and non-observed data vectors. Observe that the joint distribution in expression (13) is given in terms of x_i and the latent variables y_i, and not in terms of z_i and y_i (there is the Jacobian σ^{-1} multiplied by the right-hand-side of expression (10)).

Observe that when $\alpha = 2$ we have $\theta = 2$ and $b_{\alpha,\beta} = 0$. In this case we have a Gaussian distribution with mean equals to δ and variance equals to $2\sigma^2$.

For a Bayesian analysis of the proposed model, we assume uniform U(a,b) priors for α, β and δ where the hyperparameters a and b are assumed to be known in each application following the restrictions $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ and $\delta \in (0, \infty)$. We also assume Normal $N(a,b^2)$ prior distributions for the regression parameters $d=(d_1, d_2, \ldots, d_k)$ considering known hyperparameter values a and b^2. We further assume independence among all parameters.

In the simulation algorithm to obtain a Gibbs sample for the random quantities α, β , d andδ, having the joint posterior distribution (13), we assume a uniform U(-0.5, 0.5) prior distribution for the latent random quantities Y_i for $i = 1, \ldots, n$.Observe that, in this case, we are assuming $a_{\alpha,\beta} = 0$ ($b_{\alpha,\beta} = 0$). With this choice of priors, we use standard available software packages like OpenBugs (see Spiegelhalter, 2003) which gives great simplification to obtain the simulated Gibbs samples for the joint posterior distribution of interest.

From expression (13), the joint posterior probability distribution for $\alpha, \beta, \mathbf{d}, \delta$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ is given by

$$\pi(\alpha, \beta, d, \delta, y|\, x) \propto \left(\frac{\alpha}{|\alpha-1|\delta}\right)^n \exp\left\{\sum_{i=1}^n \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|\right\} \prod_{i=1}^n \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta} \frac{1}{|z_i|} \cdot \cdot \prod_{i=1}^n h(y_i)\pi_0(\alpha, \beta, d, \delta), \quad (14)$$

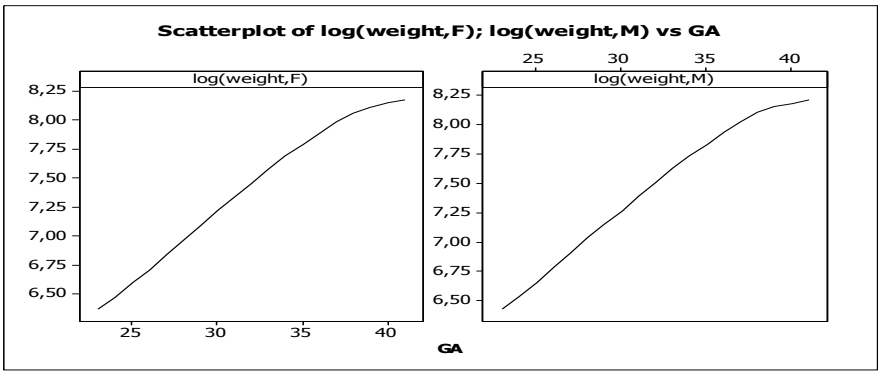
where $\theta$and $t_{\alpha,\beta}(.)$ are respectively defined in (10) and (11) and $h(y_i)$ is a U(- 0.5, 0.5) density function, for $i = 1,\ldots,n$.

Since we are using the OpenBugs software to simulate samples for the joint posterior distribution we do not present here all full conditional distributions needed for the Gibbs sampling algorithm. This software only requires the data distribution and prior distributions of the interested random quantities. This gives great computational simplification for determining posterior summaries of interest as shown in the applications as follows.

An example with intrauterine growth data for preterm infants in USA

In Table 1, we have data set related to intrauterine growth data (IGD) for preterm infants in USA (see Olsen et al, 2010). In this data set we have the weight of the preterms and the gestational age (GA) of the mothers (55445 girls and 73995 boys).

From a preliminary data analysis, we see that a nonlinear regression model (1) is suitable for the data set in the transformed scales (see Figure 1).



Plots of observed log(weight) versus gestational age.

Using the information of Table 1, since we do not have all data (intrauterine weights of 55445 girls and 73995 boys) reported in the paper of Olsen et al (2010) but we have the sample means and the standard deviations for each group of infants in a fixed gestational age, we simulated samples of the weights considering normal distributions with means and standard-deviations equal to the sample means and sample standard deviations (simulation of 553 weights for girls and 738 weights for boys) with gives a sample size close to 10% of all data introduced in Table 1.

Table 1 - Intrauterine growth data for preterm infants (55445 girls and 73995 boys)

#	GA	size F	mean,F	size M	mean,M
1	23	133	587	153	622
2	24	438	649	451	689
3	25	603	738	722	777
4	26	773	822	881	888
5	27	966	934	1030	1001
6	28	1187	1058	1281	1138
7	29	1254	1199	1505	1277
8	30	1606	1376	1992	1435
9	31	2044	1548	2460	1633
10	32	3007	1730	3677	1823
11	33	4186	1960	5014	2058
12	34	5936	2194	7291	2288
13	35	5082	2420	6952	2529
14	36	4690	2675	7011	2798
15	37	4372	2946	6692	3058
16	38	5755	3184	8786	3319
17	39	5978	3342	8324	3476
18	40	5529	3461	7235	3582
19	41	1906	3546	2538	3691

(GA=gestational age; F=female; M=male; size=sample size)

Gompertz model with normal errors (classical approach)

Assuming standard classical approach for nonlinear models considering the Gompertz growth model (v), that is,

$$f(v, \theta) = \theta_1 exp[-exp(\theta_2 - \theta_3 v)], \qquad (15)$$

where v = GA (gestational age), introduced in section 2, with the responses in the logarithmic scale log(weight), and the software Minitab version 16, we have in Table 2, the obtained maximum likelihood estimates (MLE) for the parameters of the Gompertz growth model for each sex.We have used initial values equal to zero for all parameters θ_1, θ_2 and θ_3 of the Gompertz model in the Gauss-Newton iterative procedure to get the MLE.

Table 2 - MLE estimates for the Gompertz growth model

Gender	θ_1	θ_2	θ_3
F (girls)	9.06173 (0.15523)	0.76876 (0.14414)	0.07561 (0.00736)
M(boys)	9.14384 (0.14679)	0.77103 (0.13259)	0.07488 (0.00674)

(In parentheses: SE of estimate)

From residuals plots (see Appendix 1), we see that the standard assumptions for the model are verified.

Gompertz model with normal errors (Bayesian approach)

Under a Bayesian approach, we have in Table 3, the posterior summaries of interest considering the Gompertz growth model with a normal distribution for the error (see (1)) and the OpenBugs software assuming the following prior distributions: $\theta_j \sim N(0, 1)$, j =1,2,3,4 and $\zeta = 1/\sigma^2 \sim G(1,1)$, where G(a,b) denotes a gamma distribution with mean a/b and variance a/b^2. We simulated 311,000 Gibbs samples, with a "burn-in-sample" of 11,000 samples discarded to eliminate the effects of the initial values in the iterative simulation process and taking a final sample of size 1000 (every 300th sample choosen from the 300,000 final simulated samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples. For the girls and boys we obtained respectively, Monte Carlo estimates for the posterior mean of ζ equal to 31.46 (1.907) and 31.30 (1.596).

Table 3 - Bayesian estimates for the Gompertz growth model

Gender	θ_1	θ_2	θ_3

F (girls)	8.891 (0.1166)	0.9545 (0.1421)	0.08503 (0.00684)
M(boys)	9.061 (0.1094)	0.8658 (0.1322)	0.07948 (0.00626)

(In parentheses:SE of estimate)

It is interesting to observe that the classical and Bayesian inferences assuming non-informative priors are very similar, but the standard-deviations of the estimates in general are smaller for the Bayesian approach, especially for the parameter θ_1 .

Gompertz model with stable distribution (Bayesian approach)

Now, we assume a stable distribution for the intrauterine weights of the preterm infants (girls group or female infants) considering the Gompertz growth curve model (model 5) introduced in section 2. Under a Bayesian approach, we have in Tables 4, the posterior summaries of interest assuming the nonlinear Gompertz regression model, that is, considering the regression model for the location parameter of the stable distribution given by,

$$\gamma_i = d_1 exp[-exp(d_2 - d_3 GA_i)]. \qquad (16)$$

In the Bayesian analysis of this model, we use the OpenBugs software assuming the following prior distributions:$\alpha \sim U(0, 2)$, $\beta \sim U(-1, 0)$, $\delta \sim U(0,3)$, $d_1 \sim N(9,1)$, $d_2 \sim N(0.8,1)$ and $d_3 \sim N(0.07,1)$. Observe that we are using some prior information from the results of the classical analysis to choose the prior distributions for the regression parameters (use of empirical Bayesian methods). We also assume a uniform U(-0.5,0.5) distribution for the latent variable Y_i, i=1,2,…,15.We simulated 4,000,000 Gibbs samples, with a "burn-in-sample" of 300,000 samples discarded to eliminate the effects of the initial values in the iterative simulation process and taking a final sample of size 1,000 (every 200th sample choosen from the 200,000 final samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples.

Bayesian estimates for the Gompertz growth model assuming a stable distribution for the responses (results for girls)

parameter	mean	S.D.	95% credible interval
α	1.543	0.00509	(1.536 ; 1.555)
β	-0.8361	0.00883	(-0.8449 ; -0.8204)
δ	0.1418	0.0051	(0.1356 ; 0.1579)
d_1	9.322	0.035	(9.270 ; 9.355)
d_2	0.6948	0.00756	(0.680 ; 0.711)
d_3	0.0714	0.00073	(0.0707 ; 0.0726)

From the results of Table 4, we observe that in general the Bayesian estimates of the regression parameters (d_1,d_2 and d_3) considering the Gompertz model with a stable distribution are similar to the Bayesian estimates of the regression parameters (θ_1,θ_2 and θ_3) considering the Gompertz model with normal errors, but in general the standard deviations are smaller than assuming normal errors (see results in Table 3).

Gompertz model with stable distribution in presence of outliers (Bayesian approach)

Now, let us consider the presence of outliers (female infants) replacing the 5th intrauterine weight (732.58) by 73200.58 and the first intrauterine weight (587.00) by 58700.00. In Table 5, we have the obtained posterior summaries assuming the same priors and simulation procedure assumed for the results of Table 4. We also observe in Table 5, that the estimated regression parameters θ_1,θ_2 and θ_3 of the Gompertz non-linear model with normal errors are strongly affected by the presence of outliers (see results of Table 2). Considering a stable distribution, the estimates of the regression parameters (d_1,d_2 and d_3) of the Gompertz non-linear model are very close to the estimates of the regression model with normal errors not considering the presence of outliers. This is a great advantage of the non-linear model assuming a stable distribution, that is, the Gompertz regression model with a stable distribution is robust to the presence of outliers.

Bayesian estimates for the Gompertz growth model assuming a stable distribution for the responses (presence of an outlier)

Parameter (stable)	mean	S.D.	95% credible interval
α	1.543	0.0051	(1.536 ; 1.555)
β	-0.8361	0.0088	(-0.8449; -0.8204)
δ	0.1418	0.0051	(0.1356 ; 0.1579)
d_1	9.322	0.035	(2.098 ; 3.104)
d_2	0.6948	0.0076	(0.680 ; 0.7115)
d_3	0.0714	0.00073	(0.0707 ; 0.0726)
Parameter (normal)	mean	S.D.	95% credible interval
θ_1	9.224	0.2116	(8.825 ; 9.675)
θ_2	0.3107	0.1279	(0.0772 ; 0.6047)
θ_3	0.0597	0.0072	(0.0475 ; 0.0760)

Concluding Remarks

The presence of outliers or discordant observations, many times due to measure errors is very common in applications of nonlinear regression analysis. In the presence of these discordant observations, the usual obtained classical inferences on the regression parameters or in the predictions under the usual assumption of normality for the errors and constant variance could be greatly affected, which could imply in wrong inference results. The use of stable distributions could be a good alternative for many applications in data analysis to have robust inference results, since this distribution has a great flexibility to fit for the data. With the use of Bayesian methods and MCMC simulation algorithms it is possible to get inferences for the model despite the nonexistence of an analytical form for the density function as it was showed in this paper. It is important to point out that the computational work in the sample simulations for the joint posterior distribution of interest can be greatly simplified using standard free softwares like the OpenBugs software.

In the illustrative example related to infant intrauterine growth introduced in Sections 4, the use of data augmentation techniques (see, for instance, Damien et al, 1999) is the key to obtain a good performance for the MCMC simulation method for applications using stable distributions.

We emphasize that the use of OpenBugs software does not require large computational time to get the posterior summaries of interest, even when the simulation of a large number of Gibbs samples are needed for the algorithm convergence. These results could be of great interest for researchers and practitioners, when dealing with non Gaussian data, as in the applications presented here.

ACHCAR, J. A.; OLIVEIRA, B. R. G. Curvas de crescimento de Gompertz assumindo distribuições estáveis: uma aplicação ao crescimento intra-uterino para prematuros. *Rev. Bras. Biom.*, São Paulo, v.33, n.2, p.xx-xx, 2015.

RESUMO: Neste artigo, consideramos o uso de métodos bayesianos para analisar alguns modelos de crescimento que pertencem à classe de modelos de regressão não-linear. Para a modelagem não linear assumimos erros normais uma suposição usual e também distribuições estáveis para a variável resposta. Estudamos também alguns aspectos de robustez dos modelos de regressão não linear para a presença de outliers ou observações discordantes considerando o uso de distribuições estáveis para a resposta no lugar da suposição de normalidade habitual. É bem sabido que, em geral, não há nenhuma forma fechada para a função densidade de probabilidade de distribuições estáveis. No entanto, sob uma abordagem Bayesiana, a utilização de uma variável aleatória latente ou auxiliar proporciona uma simplificação para obter qualquer distribuição a posteriori quando relacionado com distribuições estáveis. Para demonstrar a utilidade dos aspectos computacionais, a metodologia é aplicada a um exemplo relacionado com as curvas de crescimento intra-uterino para prematuros. Resumos a posteriori de interesse são obtidos utilizando métodos MCMC (Markov Chain Monte Carlo) e o software OpenBugs.

PALAVRAS-CHAVE: Distribuição estável; análise Bayesiana; modelos de regressão não-lineares; métodos MCMC; software OpenBugs.

References

ACHCAR, J.A.; LOPES, S.R.C. ; MAZUCHELI, J.; LINHARES, R. A Bayesian Approach for Stable Distributions: Some Computational Aspects. *Open Journal of Statistics*, v.3, n.4, p.268-277, 2013.

BATES, D.M. ; WATTS, D.G. *Nonlinear Regression Analysis and Its Applications*, Wiley, NewYork,1988, 392 p.

BATES, D.M. ; WATTS, D.G. Relative curvature measures of nonlinearity (with discussion), *Journal ofthe Royal Statistical Society, Series B* , v. 42, n.1, p.1-25,1980.

BUCKLE, D.J. Bayesian inference for stable distributions, *Journal of the American Statistical Association*, v.90, n.430, p. 605-613,1995.

CASELLA, G.; GEORGE, E. I.Explaining the Gibbs Sampler. *The American Statistician*, v.46, n.3, p.167-174,1992.

CHIB, S.; GREENBERG, E. Understanding the Metropolis- Hastings algorithm, *The American Statistician*, v.49, n.4, p.327-333,1995.

DAMIEN, P.; WAKEFIELD, J.; WALKER,S. Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables, *Journal of the Royal Statistical Society*, Series B, v.61,n.2, p. 331-344,1999.

FELLER, W. *An introduction to probability theory and its applications*, Vol. II. John Wiley, New York, 1971,626 p.

GELFAND, A. E.; SMITH, A.F.M. Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, v.85, n.410,p. 398-409,1990.

GNEDENKO, B. V. ; KOLMOGOROV, A.N. Limit distributions for sums of independent random variables, Addison-Wesley, Massachussetts,1968,282 p.

IBRAGIMOV, I.A.; ČERNIN, K.E. On the unimodality of stable laws, *Teor. Veroyatnost.i Primenen*., v. 4, p. 453-456,1959.

KANTER,M. On the unimodality of stable densities, *Annals of Probability*, v.4, n.6, p. 1006-1008,1976.

LÉVY, P (1924) Théorie des erreurs la loi de Gauss et les lois exceptionelles, *Bulletin Society Mathematical*, v. 52, p. 49-85, 1924.

LUKACS,E. *Characteristic Functions*, Hafner Publishing, New York,1970,216 p.

NOLAN, J. P. *Stable Distributions - Models for Heavy Tailed Data*, Birkhäuser, Boston. In progress, Chapter 1 online at academic2.american.edu/~jpnolan,2009.

OLSEN,I.E.;GROVEMAN,S.A.; LAWSON, L.; CLARK,R.H.; ZEMEL,B.S. New Intrauterine Growth Curves Based on United States Data, *Pediatrics*, doi:10.1542/peds.2009-0913,2010.

RATKOWSKY, D.A. *Nonlinear Regression Modelling: A Unified Practical Approach*, Marcel Dekker,New York,1983, 288 p.

ROUSSAS,G.G. *Measure-Theoretic probability*, San Diego:Elsevier,2005,390 p.

SEBER, G.A.F. ; WILD, C.J. *Nonlinear Regression*, Wiley, New York,1989,792 p.

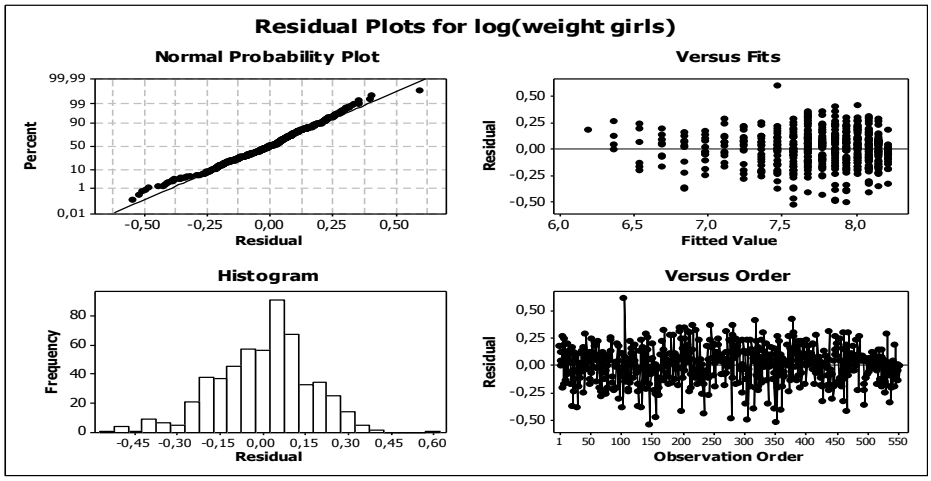
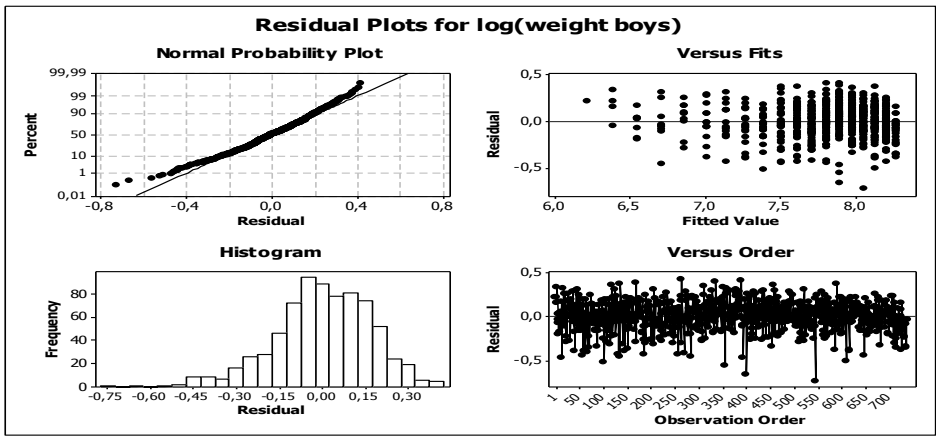
SKOROHOD, A.V. On a theorem concerning stable distributions. In: *Select. Transl. Math. Statist. and Probability*, v. 1, Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I.,p.169-170,1961.

SPIEGELHALTER, D.J.; THOMAS, A.; BEST, N.G.; LUNN, D. WinBugs User's Manual, *MRC Biostatistics Unit*, Cambridge, 2003.

SPIEGELHALTER, D.J.; BEST, N. G., CARLIN, B. P., BAN DER LINDE, A. Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society*, Series B, v.64, n.4, p.583-639, 2002.

TANNER. M.A. ; WONG,W.H. The calculation of posterior distributions by data augmentation", *Journal of American Statistical Association*, v. 82,n.398,p. 528-550,1987.

Recebido em 16.02.2015

Aprovado após revisão em 26.04.2015

Appendix 1 - residual plots.