

UNIVERSIDADE FEDERAL DE VIÇOSA

**Abordagem bayesiana de modelos não lineares aplicados ao crescimento de
frutos de *Capsicum annuum* L.: Considerações de erros experimentais
simétricos e assimétricos**

George Lucas Santana de Moura
Doctor Scientiae

**VIÇOSA - MINAS GERAIS
2024**

GEORGE LUCAS SANTANA DE MOURA

**Abordagem bayesiana de modelos não lineares aplicados ao crescimento de
frutos de *Capsicum annuum* L.: Considerações de erros experimentais
simétricos e assimétricos**

Tese apresentada à Universidade Federal
de Viçosa, como parte das exigências do
Programa de Pós-Graduação em
Estatística Aplicada e Biometria, para
obtenção do título de *Doctor Scientiae*.

Orientador: Paulo Roberto Cecon

Coorientadores: Sebastião Martins Filho
Antonio P. S. Carneiro
Moyses Nascimento

**VIÇOSA - MINAS GERAIS
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

M929a
2024

Moura, George Lucas Santana de, 1995-

Abordagem bayesiana de modelos não lineares aplicados ao crescimento de frutos de *Capsicum annuum* L.: considerações de erros experimentais simétricos e assimétricos / George Lucas Santana de Moura. – Viçosa, MG, 2024.

1 tese eletrônica (95 f.): il. (algumas color.).

Orientador: Paulo Roberto Cecon.

Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Estatística, 2024.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2025.438>

Modo de acesso: World Wide Web.

1. Análise de regressão. 2. Teorias não-lineares. 3. Teoria bayesiana de decisão estatística. 4. Pimentas - Crescimento. I. Cecon, Paulo Roberto, 1955-. II. Universidade Federal de Viçosa. Departamento de Estatística. Programa de Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.536

GEORGE LUCAS SANTANA DE MOURA

Abordagem bayesiana de modelos não lineares aplicados ao crescimento de frutos de *Capsicum annuum* L.: Considerações de erros experimentais simétricos e assimétricos

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Doctor Scientiae*.

APROVADA: 6 de novembro de 2024.

Assentimento:

George Lucas Santana de Moura
Autor

Paulo Roberto Cecon
Orientador

Essa tese foi assinada digitalmente pelo autor em 09/02/2026 às 18:01:45 e pelo orientador em 10/02/2026 às 11:28:29. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **HHS5.64H9.JK2W** e clique no botão 'Validar documento'.

AGRADECIMENTOS

Ao pai, ao filho e ao espírito Santo, Amém!

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

“A maioria, na maioria das vezes, está errada.” (George Lucas)

RESUMO

MOURA, George Lucas Santana de, D.Sc., Universidade Federal de Viçosa, novembro de 2024. **Abordagem bayesiana de modelos não lineares aplicados ao crescimento de frutos de *Capsicum annuum* L.: Considerações de erros experimentais simétricos e assimétricos.** Orientador: Paulo Roberto Cecon. Coorientadores: Sebastião Martins Filho, Antonio Policarpo Souza Carneiro e Moyses Nascimento.

No final dos anos 90 e início dos anos 2000, registrou-se uma ampla variedade de propostas de novas distribuições de probabilidades simétricas, assimétricas e bimodais, cuja flexibilidade atentasse a exigência do comportamento dos dados. Tratando-se da modelagem de fenômenos biológicos, os modelos de regressão não linear sob o erro gaussiano são um procedimento padrão, entretanto, as novas propostas de modelos de regressão não linear com erros flexíveis aumentam a precisão dos modelos e a qualidade de ajuste sem precisar utilizar-se de transformação de variáveis como a de Box-Cox que comprometem a interpretação dos resultados. Há também novas propostas de modelos de crescimento sigmóides para dados longitudinais como o Logístico, Gompertz e Von Bertalanffy que permitem uma interpretação biológica para o parâmetro beta. O modelo Logístico ajustado ao Acesso 2 – Pimenta Caiena se mostrou com as melhores qualidade de ajuste, por meio do DIC e pelo Fator de Bayes, quando se utiliza de distribuições simétricas além da Normal, os erros Laplace, Student t, logístico e Normal generalizado contribuíram com uma melhora na qualidade de ajuste e na precisão do modelo, e tendo o erro experimental Normal generalizado com os melhores indicadores. As versões Skew-Student t, Skew-Logistic e Skew-Normal apresentaram um aumento na precisão no ajuste do modelo Logístico ao Acesso 8 – Pimentão Rubi Gigante. A presença da assimetria no modelo muda os resultados do parâmetro beta, os modelos simétricos estimaram em 11 dias o ponto de inflexão da curva de crescimento, as demais versões assimétricas estimaram de 9 a 10 dias. Os erros experimentais Fernández & Steel's Skew-Student t, Skew-Double exponential e Skew-Double Exponential-Normal mostraram significância estatística quanto a hipótese nula de ausência de assimetria.

Palavras-chave: Regressão não linear; Crescimento sigmoide; erros assimétricos; Regressão robusta

ABSTRACT

MOURA, George Lucas Santana de, D.Sc., Universidade Federal de Viçosa, November, 2024. **Bayesian approach to nonlinear models applied to the growth of *Capsicum L.* growth: Considerations of symmetric and asymmetric experimental errors.**.. Adviser: Paulo Roberto Cecon. Co-advisers: Sebastião Martins Filho, Antonio Policarpo Souza Carneiro and Moyses Nascimento.

In the late 1990s and early 2000s, a wide variety of proposals for new symmetric, asymmetric and bimodal probability distributions were recorded, whose flexibility would meet the demands of data behavior. When it comes to modeling biological phenomena, nonlinear regression models under Gaussian error are a standard procedure; however, new proposals for nonlinear regression models with flexible errors increase the accuracy of the models and the quality of fit without the need to use variable transformations such as Box-Cox, which compromise the interpretation of the results. There are also new proposals for sigmoid growth models for longitudinal data, such as the Logistic, Gompertz and Von Bertalanffy models, which allow a biological interpretation for the beta parameter. The Logistic model adjusted to Accession 2 – Cayenne Pepper showed the best quality of fit, through the DIC and the Bayes Factor, when using symmetric distributions in addition to the Normal one. The Laplace, Student t, logistic and generalized Normal errors contributed to an improvement in the quality of fit and precision of the model, and the generalized Normal experimental error had the best indicators. The Skew-Student t, Skew-Logistic and Skew-Normal versions showed an increase in the precision in the fit of the Logistic model to Accession 8 – Giant Ruby Pepper. The presence of asymmetry in the model changes the results of the beta parameter; the symmetric models estimated the inflection point of the growth curve in 11 days, while the other asymmetric versions estimated 9 to 10 days. The experimental errors Fernández & Steel's Skew-Student t, Skew-Double exponential and Skew-Double Exponential-Normal showed statistical significance regarding the null hypothesis of absence of asymmetry.

Keywords: Nonlinear regression; Sigmoid growth; Asymmetric errors; Robust regression

LISTA DE ILUSTRAÇÕES

Figura 1: Modelo de crescimento Logístico reparametrizado contendo o ponto de inflexão e o limite assintótico.....	17
Figura 2: Função de verossimilhança da distribuição Skew-Normal de Azzalini em função de μ e σ , considerando $\lambda = 2$	23
Figura 3: Densidades da distribuição Skew-Laplace-Normal padronizada e variando valores de λ ...	26
Figura 4: <i>Scatterplots</i> do comprimento do fruto dos Acessos de Capsicum em função dos dias após o florescimento.....	33
Figura 5 – Curvas de nível da Distribuição elíptica Normal (figura esquerda), e da distribuição esférica simétrica Normal (figura direita).....	35
Figura 6: Densidades das distribuições Normal, Logística, Laplace e Student t variando os parâmetro de escala e locação.	38
Figura 7: Densidades da distribuição $NGN_{0,1}$, s variando os valores de s	40
Figura 8: Modelos Logístico, Gompertz e Von Bertalanffy ajustados aos 8 acessos de <i>capsicum annum L.</i> via modelos mistos não lineares.....	41
Figura 9: Gráfico de violino dos modelos Logístico, Gompertz e Von Bertalanffy à todos os Acessos de <i>Capsicum annum L.</i> representados pela Figura 4.....	42
Figura 10: Gráficos representando a média a posteriori dos parâmetros $(\alpha, \beta, \gamma, \sigma^2, v)$ dos modelos Logístico, Gompertz e Von Bertalanffy utilizando 5 erros experimentais.....	46
Figura 11: Densidades das distribuições a posteriori dos parâmetros $\alpha, \beta, \gamma, \sigma^2$ dos modelos Logístico Gompert e Von Bertalanffy considerando todos os erros experimentais.	47
Figura 12: <i>Scatterplots</i> dos Acessos de <i>Capiscum Annuum L.</i> mostrando o comprimento do fruto em função dos dias após o florescimento.....	54
Figura 13: Gráfico das densidades de ASN mostrando o deslocamento relativo do parâmetro de locação quando se utiliza os dois casos $ASN_{\mu, \sigma, \lambda}$ – sem correção e $ASN_{\mu - \sigma\delta^2/\pi, \sigma, \lambda}$ – com correção.	56
Figura 14: Regressão não linear assimétrica centralizada em cada EY_i utilizando a distribuição ASN corrigida.	57
Figura 15: Comportamento da distribuição Skew-Normal de Azzalini variando o parâmetro de assimetria (λ values).....	58
Figura 16: Comportamento da distribuição Skew-Normal de Sahu padronizada variando o parâmetro de assimetria (λ values).....	61
Figura 17: Comportamento da distribuição Skew-Normal de Fernández & Steel padronizada variando o parâmetro de assimetria (λ values).....	64
Figura 18: Comportamento da distribuição ASL padronizada variando os valores de lambda (λ values).....	67
Figura 19: Comportamento da distribuição SBSL padronizada variando os valores de lambda (κ).	69
Figura 20: Comportamento da distribuição $FSSSt$ padronizada variando os valores de lambda (λ values).....	71
Figura 21: Comportamento da distribuição $ESSt$ padronizada variando os valores de lambda (λ values).....	73
Figura 22: Comportamento da distribuição $SDEN$ padronizada variando os valores de lambda (λ values).....	76
Figura 23: Comportamento da distribuição $JFSSSt$ padronizada variando os valores de lambda (λ values).....	78

Figura 24: Gráfico de violino em conjunto com o Boxplot dos resíduos, após o ajuste dos modelos Logístico, Gompertz e Von Bertalanffy sob o erro Gaussiano (N) à todos os acessos representados pela Figura 12.	81
Figura 25: Densidades do histograma das estimativas dos parâmetros $\alpha, \beta, \gamma, \sigma^2$ dos modelos Logístico, Gompertz e Von Bertalanffy sob o erro Gaussiano (N) via modelos mistos não lineares ajustados a todos os acessos representados pela Figura 12.	82
Figura 26: Médias a posteriori dos parâmetros $\alpha, \beta, \gamma, \sigma^2$ com seus respectivos intervalos de credibilidade considerando os 14 erros experimentais.....	86
Figura 27: Distribuição a posteriori dos parâmetros $\alpha, \beta, \gamma, \sigma^2$ do modelo Logístico considerando os 14 erros experimentais.....	87
Figura 28: Gráfico de violino em conjunto com o boxplot das densidades a posteriori do parâmetro λ considerando 10 erros experimentais assimétricos.	89
Figura 29: DIC's resultantes do ajuste do modelo Logístico considerando todas as 14 distribuições de erro experimental.	90
Figura 30: Tempo computacional (em segundos) necessário para implementar as 14 distribuições de erro experimental no Openbugs.....	91
Figura 31: Autocorrelação das cadeias MCMC dos parâmetros $(\alpha, \beta, \gamma, \sigma^2, \lambda)$ considerando as estimativas frequentista como valores iniciais vs. Estimativas iniciais ($\alpha_0 = 1, \beta_0 = 1, \gamma_0 = 1, \sigma_{20} = 1, \lambda_0 = 1$).....	92

LISTA DE TABELAS

Tabela 1: Médias a posteriori dos parâmetros $\alpha, \beta, \gamma, \sigma^2, v$ conjuntamente com os DIC's de cada ajuste.	45
Tabela 2: Logaritmo de base 10 do Fator de Bayes resultante da comparação 2 a 2 dos modelos Logístico, Gompertz e Von Bertalanffy dentro de cada erro experimental.	47
Tabela 3: Médias a posteriori dos parâmetros $\alpha, \beta, \gamma, \sigma^2, v$, em conjunto com os respectivos DIC's do ajuste logístico a cada erro experimental.	83
Tabela 4: médias a posteriori dos parâmetros λ dos 10 erros experimentais assimétricos com seus respectivos intervalos de credibilidade e orientação da assimetria dos dados.	88

Sumário

INTRODUÇÃO GERAL.....	12
1.1 REFERENCIAL TEÓRICO	14
1.1.1 Modelo não linear generalizado.....	14
1.1.2 Modelos de crescimento sigmóides.....	15
1.1.3 Algoritmo de Gauss-Newton	18
1.1.4 Método Score de Fisher	19
1.1.5 Coeficiente de determinação	21
1.1.6 <i>The Poisson zeros' trick</i>	21
1.1.7 Truque de Bernoulli	23
1.1.8 Computação Bayesiana.....	24
1.1.9 Algoritmo de Aceitação-Rejeição.....	25
1.1.10 Amostrador de Gibbs.....	26
1.2 REFERÊNCIAS	27
CAPÍTULO 1: Abordagem Bayesiana robusta de modelos não lineares aplicados ao crescimento de <i>capisicum annuum L.</i> utilizando distribuições simétricas de locação-escala	30
2.1 INTRODUÇÃO.....	30
2.2 MATERIAL E MÉTODOS	32
2.2.1 Distribuições elípticas e simétricas.....	34
2.2.2 Família simétrica de locação-escala	35
2.2.3 Regressão não linear bayesiana com erros simétricos	36
2.2.4 Distribuição Normal de Nadarajah (2005)	38
2.2.5 Determinação da Priori empírica	40
2.2.6 Comparação de Modelos e Análise de Convergência	42
2.3 RESULTADOS E DISCUSSÕES.....	43
2.4 CONCLUSÕES	48
2.5 REFERÊNCIAS	48
CAPÍTULO 2: Abordagem bayesiana robusta de modelos não lineares aplicados ao crescimento de <i>capisicum</i> : Um estudo das distribuições assimétricas	51
3. INTRODUÇÃO	51
3.2 MATERIAL E MÉTODOS	53
3.2.1 Regressão não linear com erros assimétricos.....	54
3.2.2 Especificação da função de verossimilhança.....	57
3.2.3 Distribuição Skew-Normal de Azzalini	57
3.2.4 Distribuição Skew-Normal de Sahu et al. (2003).....	59

3.2.5 Distribuição Skew-Normal via transformação de Fernández & Steel (1998)	62
3.2.6 Distribuição Skew-Laplace de Kotz (2001).....	66
3.2.7 Distribuição Skew-Logistic de Sastry and Bhati (2016) - SBSL	68
3.2.8 Distribuição Skew- t de Fernández & Steel (1998).....	70
3.2.9 Distribuição Epsilon-Skew-t (GOMEZ et al., 2007a)	72
3.2.10 Distribuição Skew-Laplace-Normal (GOMEZ ET AL., 2007b).....	75
3.2.11 Distribuição Skew-t de Jones & Faddy (2003)	76
3.2.12 Distribuição Skew-t via transformação sinh-arcsinh	79
3.2.13 Modelos mistos não lineares na implementação da priori empírica.....	80
3.2.14 Estrutura Bayesiana	82
3.3 RESULTADOS E DISCUSSÕES.....	83
3.4 CONCLUSÕES	92
3.5 REFERÊNCIAS	93

INTRODUÇÃO GERAL

No início do século XX, Galton desenvolveu a chamada teoria de Regressão a média, a teoria surgiu após Galton observar que certas características fenotípicas humanas, sempre regrediam a média em termos de herdabilidade, um exemplo seria a altura humana, os filhos de pais altos têm a tendência de se tornarem mais altos até certo ponto, a recíproca é válida para filhos de pais baixos, depois desse certo ponto a tendência se inverte e a altura dos filhos regride a média.

A teoria de regressão a média de Galton se tornou conhecida por Fisher que à generalizou para designs de experimentação e introduziu novos métodos de estimação paramétrica como a Máxima Verossimilhança (SALSBURG, 2009). Na teoria da Regressão, é pressuposto que o pesquisador sugira o modelo de Regressão conforme a relação funcional da variável independente e a variável resposta, ou conforme a natureza do problema (BUSSAD & MORETTIN, 2017).

Ao se abordar a questão da modelagem matemática de fenômenos biológicos, demográficos, físicos e químicos, existem certas situações problemas que são modeladas matematicamente por meio de equações diferenciais ordinárias (EDO), um exemplo prático pode ser encontrado no modelo de crescimento de Velhust, em que o autor teorizou que a taxa de crescimento de uma determinada população era proporcional ao seu próprio tamanho com fatores limitantes de crescimento, a solução desta EDO de primeira ordem é o conhecido modelo de crescimento Logístico (STEWART, 2012).

Assim como cada nicho das ciências agrícolas, o crescimento de plantas dispõe de uma série de modelos estatísticos que são amplamente utilizados na modelagem de dados longitudinais. Conforme Rosa et al. (2022), o objetivo é – de todas as variedades de modelos – encontrar aquela que melhor explica e se ajusta aos dados. Modelos de crescimento como o de Velhust ou de Richards, são utilizados nas ciências agrárias para modelagem de dados longitudinais.

No crescimento de plantas, algumas de suas características fenotípicas - incluindo as de seus frutos - seguem um crescimento exponencial até que a taxa de maturação seja atingida, a partir desse ponto de inflexão o crescimento da planta decai até atingir um valor assintótico, o padrão desse tipo de crescimento é denominado sigmoide. Modelos sigmoides são não lineares e geralmente utiliza-se a distribuição Normal para o erro experimental (SOUZA, 1998).

A classe dos modelos não lineares se diferencia dos modelos lineares pela interpretabilidade dos parâmetros, e de sua origem, que geralmente são provenientes das

soluções analíticas de equações diferenciais ordinárias (EDO), equações estas que modelam fenômenos biológicos, demográficos, físicos e químicos (STEWART, 2016).

Em modelos lineares Normal como o de Gauss-Markov, os testes de significância, ANOVA e intervalos de confiança são exatos, apesar de dependerem de uma série de condições de estimabilidade (LUNA & OLINDA, 2012). Ao considerar uma distribuição do erro além da Normal, geralmente os testes de significância dependem da normalidade assintótica dos estimadores de máxima Verossimilhança (MV). No caso de Modelos não lineares Normal (MNLN), os estimadores de MV e Gauss-Newton dependem também da teoria assintótica, e a ANOVA nestes casos não é válida pois a decomposição da soma de quadrados totais em resíduo e regressão não é válida (GUJARATI & PORTER, 2012).

O ganho que se tem na metodologia bayesiana de MNLN consiste na validade dos testes de significância e dos intervalos de credibilidade, sem se preocupar com o tamanho da amostra (MARTINS FILHO et. al. 2008) ou com a existência de um máximo global da função de verossimilhança.

O objetivo do presente trabalho é implementar distribuições simétricas e assimétricas para o erro experimental numa regressão não linear bayesiana para a descrição do crescimento do fruto de *Capsicum annuum* L. No capítulo 1 será estudado os erros experimentais simétricos, e no capítulo 2 os erros experimentais assimétricos.

1.1 REFERENCIAL TEÓRICO

1.1.1 Modelo não linear generalizado

De acordo com Souza (1998) podemos definir um modelo não linear de acordo com a equação 1.

$$y_i = f(x_i, \theta) + e_i \quad (1)$$

em que y e x são, respectivamente a variável preditora e a variável observada, $f(x, \theta)$ que são objetos da classe C^n (n vezes diferenciável em relação ao vetor θ) são os modelos de crescimento não lineares em que o vetor $\theta_{(p,1)} = (\theta_1, \theta_2, \dots, \theta_p)$ com $\theta \in \mathbb{R}^p$ e $e_i \sim N(0, \sigma^2)$ são o erro experimental aditivo, independente, identicamente distribuído e normalmente distribuído. Diferente do modelo de Gauss Markov que pode ser escrito matricialmente na forma $X_{(n,p)}\beta$ em que é possível decompor $f(*)$ em um vetor paramétrico β e uma matrix de entradas $X_{(n,p)}$. Os MNLN são genericamente expressos por $f(*)$.

Outra forma de se expressar a equação 1 é por meio da equação 2.

$$y_i = \eta(x_i, \theta) + \sigma e_i \quad (2)$$

em que o erro e_i é uma variável aleatória tal que $e_i \sim N(0,1)$, então $Y_i = \eta(x_i, \theta) + \sigma e_i$ são o equivalente a uma transformação de locação-escala em que $\eta(x_i, \theta)$ ocuparia o parâmetro de locação e σ o parâmetro de escala.

A equação (1) também representa uma transformação de locação na variável e_i gerando uma distribuição Amostral amostral de $Y_i \sim N(\eta(x_i, \theta), \sigma^2)$ independente, mas não identicamente distribuído com parâmetro de locação assumindo valor de $f(x_i, \theta)$ e parâmetro de escala σ . A distribuição Normal $\phi(0,1)$ é membro da família exponencial e da família de locação-escala, podemos escrever (1) na forma exponencial que é proposta por Nelder & Weddenburn (1976). Cordeiro & Paula (1989) definiram o modelo não linear generalizado.

$$f_Y(y, \mu, \sigma) = \exp\{\sigma^{-1}[y\mu - b(\mu)] + c(y, \mu)\} \quad (3)$$

- $b(\mu) = \frac{\mu^2}{2}$;
- $c(y, \mu) = -\frac{1}{2}\left(\frac{y^2}{\sigma^2} + \log 2\pi\sigma^2\right)$;
- $\mu = n_i$.

O preditor não linear é dado por $\eta_i = f(x_i, \theta)$ e a função de ligação é identidade. A função de verossimilhança é uma função da soma de quadrados do erro experimental. Como os

modelos não lineares normais são membros da família exponencial, as propriedades usuais e de função geratrizes e funções cumulantes têm expressões fechadas e analíticas.

1.1.2 Modelos de crescimento sigmoides

Os principais modelos de crescimento que estão descritos na literatura são soluções analíticas obtidas de equações diferenciais ordinárias. EDO's são representações da modelagem matemática de fenômenos determinados e costumam relacionar uma função desconhecida da classe C^n com algumas de suas derivadas (STEWART, 2012). Uma equação diferencial simples que dá origem ao conhecido modelo de crescimento natural se dá ao exemplo a seguir.

Exemplo 1: Tomemos $P(t)$ uma função do crescimento populacional de uma flora de bactérias em função do tempo contínuo $t \in \mathbb{R}^+$, ao tomarmos uma pressuposição intuitiva de que a taxa de crescimento populacional $P'(t)$ é proporcional ao tamanho da própria população $P(t)$, obtemos a EDO simples.

$$P'(t) = \frac{dP(t)}{dt} = kP(t) \quad (4)$$

Em que $k \in \mathbb{R}$, uma constante de proporcionalidade sendo $k > 0$ quando há crescimento, ou $k < 0$ quando há decaimento populacional.

Exemplo 2: Considerando a seguinte equação diferencial ordinária:

$$M^{(1)}_Y(t) = M_Y(t)(\mu + t\sigma^2) \quad (5)$$

É possível demonstrar que a solução da equação 5 é a função geratriz da distribuição Normal.

O modelo de Velhust que dá origem ao conhecido modelo Logístico de crescimento sigmoide é fornecida pela equação diferencial logística 5:

$$Y'(t) = kY \left(1 - \frac{Y}{m} \right) \quad (6)$$

$Y(t)$: Velocidade de crescimento;

k : Taxa de crescimento;

m : Capacidade de suporte à população.

A equação diferencial 6 é separável e tem solução analítica, entretanto, nem sempre essas condições se mantem para todos os modelos, algumas EDO's como a de Von Bertalanffy tem soluções um pouco mais complexas com o uso do método do fator integrante, mais detalhes podem ser obtidos em Fernandes et al. (2020).

A solução obtida da resolução da equação 6 é expressa conforme a equação 7, em que:

$$Y(t) = \frac{\alpha}{1 + \beta e^{-\gamma x_i}} \quad (7)$$

α : Representa o valor assintótico;

β : Geralmente não tem valor biológico;

γ : Representa a taxa de crescimento.

Uma revisão completa dos principais modelos de crescimento e de suas propriedades matemáticas pode ser encontrada no trabalho de Koya & Goshu (2013).

Entretanto, conforme explana Fernandes et al. (2015), os usuais modelos de crescimento Logístico, Gompertz e Von Bertalanffy mais utilizados na literatura, devem a sua frequência de uso pela sua popularidade, o autor também explana que estes modelos – como o da equação 6 – são os mais longes da linearidade por meio da curvatura de Battes e não possuem uma interpretação biológica de β .

Novas formas de parametrizações têm surgido para modelos sigmóides com a finalidade de adequar β à uma interpretabilidade biológica. Retomando a EDO da equação 6, tomemos a segunda diferencial de $Y(t)$:

$$Y^{(2)}(t) = kY^{(1)} \left(1 - \frac{Y}{M}\right) - kY \left(\frac{Y^{(1)}}{M}\right) \quad (8)$$

readequando a equação 8 anterior substituindo $Y^{(1)}$ novamente pela equação 6,

$$Y^{(2)}(t) = k^2 Y \left(1 - \frac{Y}{M}\right) \left(1 - \frac{2Y}{M}\right)$$

têm-se que, quando o ponto de inflexão da curva ocorre, o tamanho da população já atingiu a metade de sua capacidade de suporte.

$$Y^{(2)}(t) = k^2 Y \left(1 - \frac{Y}{M}\right) \left(1 - \frac{2Y}{M}\right) = 0 \leftrightarrow Y(t) = \frac{M}{2}$$

assim temos um problema de valor inicial:

$$\begin{cases} Y'(t) = kY \left(1 - \frac{Y}{M}\right) \\ Y(t = \beta) = \frac{M}{2} \end{cases}$$

Para resolver o problema de valor inicial primeiro solucionemos a equação 6:

$$\int \left(\frac{1}{Y} + \frac{1}{M - Y}\right) dY(t) = \int k dt \rightarrow \exp \left\{ \log \left| \frac{M}{Y} - 1 \right| \right\} = \exp \{-kt - C\}$$

temos a solução pela equação 9 que representa infinitas soluções por meio da constante de integração $C \in \mathbb{R}$, entretanto C deve ser único para atender a condição de $Y(\beta) = \frac{M}{2}$

$$Y(t) = \frac{M}{1 + \exp\{-kt - C\}} \quad (9)$$

Agora verificando a condição,

$$\frac{M}{1 + \exp\{-k\beta - C\}} = \frac{M}{2}$$

temos como solução $C = -k\beta$, reescrevendo a equação 9 com essa solução, e dessa vez substituindo $M = \alpha$ e $k = \gamma$, obtemos o modelo Logístico reparametrizado.

$$\eta(\theta) = \frac{\alpha}{1 + e^{\gamma(\beta-x)}} \quad (10)$$

No novo modelo de crescimento logístico expresso pela equação 10, a interpretação biológica que o parâmetro β assume, seria o valor em que a população atinge a sua velocidade máxima de crescimento, a partir do ponto de inflexão β , a população continuaria crescendo cada vez menos até atingir o seu tamanho limite α , o valor β também significa o valor em que a população chegou na metade do seu crescimento ($\alpha/2$). A Figura 1 explana o comportamento deste modelo.

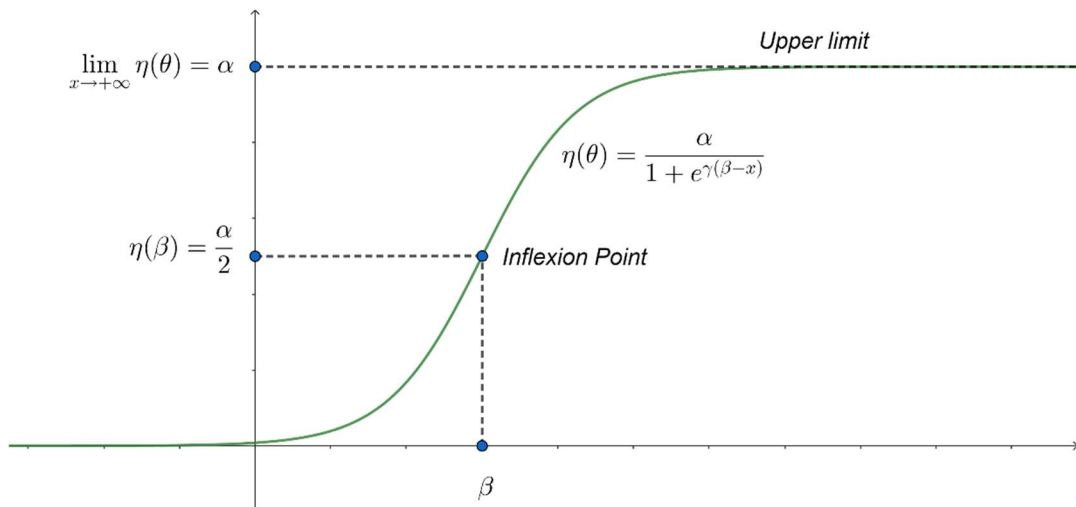


Figura 1: Modelo de crescimento Logístico reparametrizado contendo o ponto de inflexão e o limite assintótico.

Fonte: O autor

Uma propriedade importante do modelo Logístico consiste no fato deste modelo ter sua derivada simétrica em relação a β , ou seja, a função é côncava (segura água) até β , e depois convexa (derrama água) na mesma proporção.

A mesma ideia explanada anteriormente sobre a readequação de β pode ser utilizada também no modelo Gompertz (equação 11) e Von Bertalanffy (equação 12) (Fernandes et al. 2015; Fernandes et al. 2020).

$$\eta(\theta) = \alpha \exp\{-\exp\{\gamma(\beta - x)\}\} \quad (11)$$

$$\eta(\theta) = \alpha \left(1 - \frac{\exp\{\gamma(\beta - x)\}}{3}\right)^3 \quad (12)$$

1.1.3 Algoritmo de Gauss-Newton

O Algoritmo de Gauss-Newton consiste em linearizar a soma de quadrados do erro (S) de modelos não lineares por meio da aproximação de Taylor de primeira ordem do preditor $\eta(x_i, \theta)$, **antes** da otimização numérica de S em relação ao vetor θ .

Utilizando a expansão de Taylor multivariada em torno de θ^o (valor inicial) para cada x_i

$$\eta(x_i, \theta) \approx \eta(x_i, \theta^o) + \sum_{k=1}^p \frac{\partial}{\partial \theta_k} \eta(x_i, \theta^o) \times (\theta_k - \theta_k^o) \quad (13)$$

Tomando a soma do quadrado do erro (equação 14):

$$S = \sum_{i=1}^n (y_i - \eta(x_i, \theta))^2 \quad (14)$$

Aplicando a equação 13 na 14 obtemos uma aproximação para S ,

$$S \approx \sum_{i=1}^n \left(y_i - \eta(x_i, \theta^o) - \sum_{k=1}^p \frac{\partial}{\partial \theta_k} \eta(x_i, \theta^o) \times (\theta_k - \theta_k^o) \right)^2$$

Podemos reescrever o escalar $\sum_{k=1}^p \frac{\partial}{\partial \theta_k} \eta(x_i, \theta^o) \times (\theta_k - \theta_k^o)$ dentro de uma forma matricial em que $\nabla \eta(x_i, \theta^o)$ é o vetor gradiente de η em relação a θ , e $\theta = (\theta_1, \theta_2, \dots, \theta_p)^T$.

$$S \approx \sum_{i=1}^n (y_i - \eta(x_i, \theta^o) - \nabla \eta(x_i, \theta^o) \times (\theta - \theta^o))^2$$

Utilizando uma forma matricial de S em que $y = (y_1, y_2, \dots, y_n)$ e $x = (x_1, x_2, \dots, x_n)$ e $\nabla \eta(x, \theta^o) = (\nabla \eta(x_1, \theta^o), \nabla \eta(x_2, \theta^o), \dots, \nabla \eta(x_n, \theta^o))_{n,p}$.

$$S \approx (y - \eta(x, \theta^o) - \nabla \eta(x, \theta^o) \times (\theta - \theta^o))^T (y - \eta(x, \theta^o) - \nabla \eta(x, \theta^o) \times (\theta - \theta^o))$$

Para facilitar façamos $e^o = y - \eta(x, \theta^o)$ que seria o erro quando se utiliza os valores iniciais de θ^o , e $\Delta = \theta - \theta^o$.

$$S \approx (e^o - \nabla \eta(x, \theta^o) \times \Delta)^T (e^o - \nabla \eta(x, \theta^o) \times \Delta)$$

Utilizando-se de algumas propriedades matriciais

$$S \approx ((e^o)^T - \Delta^T \times \nabla^T \eta(x, \theta^o)) (e^o - \nabla \eta(x, \theta^o) \times \Delta)$$

$$S \approx (e^o)^T e^o - (e^o)^T \nabla \eta(x, \theta^o) \times \Delta - \Delta^T \times \nabla^T \eta(x, \theta^o) e^o + \Delta^T \times \nabla^T \eta(x, \theta^o) \nabla \eta(x, \theta^o) \times \Delta$$

agora deriva-se S em relação ao vetor coluna Δ e iguala-se a 0, assim obtem-se a solução iterativa de S conforme a equação 15, as propriedades de cálculo matricial podem ser encontradas em Ferreira (2018).

$$\theta^{o+1} = \theta^o + \{\nabla^T \eta(x, \theta^o) \times \nabla \eta(x, \theta^o)\}^{-1} \nabla^T \eta(x, \theta^o) \times (y_i - \eta(x_i, \theta^o)) \quad (15)$$

Uma propriedade interessante do algoritmo de Gauss-Newton consiste na sua utilização em MLN. Com uma única iteração, o algoritmo converge para a solução analítica de MQO de MLN.

Dado a matriz $\eta(x, \theta)$ tal que $\beta = \theta$, reescrevendo $\eta(x, \theta) = X_{(n,p)}\beta$ do modelo linear de Gauss-Markov, sabe-se que $\nabla^T \eta(x, \beta^o) = \nabla^T X_{(n,p)}\beta = \frac{d}{d\beta} X_{(n,p)}\beta = XI = X$, assim substitui-se esta informação na equação 14:

$$\begin{aligned} \beta^{o+1} &= \beta^o + [X^T X]^{-1} X^T (y - X\beta^o) \\ \beta^{o+1} &= \beta^o + [X^T X]^{-1} X^T y - [X^T X]^{-1} X^T X\beta^o \\ \beta^{o+1} &= \beta^o + [X^T X]^{-1} X^T y - \beta^o \\ \beta^{o+1} &= [X^T X]^{-1} X^T y \\ \hat{\beta} &= [X^T X]^{-1} X^T y \end{aligned}$$

Em que $\hat{\beta}$ são solução quando $X^T X$ é não singular.

1.1.4 Método Score de Fisher

Considere o problema de se encontrar os parâmetros de um modelo logístico (equação 5) sob o erro gaussiano por meio do método da máxima Verossimilhança. O método score de Fisher consiste em utilizar uma solução via aproximação numérica.

Exemplo 3: Considere o modelo Logístico (equação 6), os parâmetros de interesse $\theta = (\alpha, \beta, \gamma, \sigma)$, podem ser expressos por meio de distribuição conjunta de uma amostra aleatória independente, mas não identicamente distribuída.

$$L(\theta|y) = \prod f_{Y_i}(y_i) = \sigma^{-n} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n z_i^2 \right\}$$

em que $z_i = (f(x_i, \theta) - \mu)/\sigma$. Sendo $L(\theta|y)$ côncava, é conveniente maximizar um somatório ao invés de um produtório, $\ell(\theta|y) = \log L(\theta|y) = \sum_{i=1}^n \log f_{Y_i}(y_i)$, desta forma o interesse é de encontrar:

$$\hat{\theta}_{MV} = \operatorname{argmax}_{\theta} \ell(\theta|y)$$

Obtendo os scores e igualando a zero, obtemos um sistema de equações não lineares.

$$U_{\alpha,\beta,\gamma,\sigma} = \left(\frac{\partial}{\partial \alpha} \ell(\theta|y), \frac{\partial}{\partial \beta} \ell(\theta|y), \frac{\partial}{\partial \gamma} \ell(\theta|y), \frac{\partial}{\partial \sigma} \ell(\theta|y) \right)^T = 0$$

em que $U_{\alpha,\beta,\gamma,\sigma} = \nabla^T \ell(\theta|y)$, e para cada θ_i , $U_{\theta_i} = \frac{\partial}{\partial \theta_i} \ell(\theta|y)$. A solução do sistema de equações é não elementar, então recorre-se a metodos numéricos para o problema, utilizando uma expansão de Taylor de 1ª ordem no vetor $U_{\alpha,\beta,\gamma,\sigma}$ em torno de $\theta^o = (\alpha^o, \beta^o, \gamma^o, \sigma^o)^T$, obtemos:

$$U_{\theta} \approx U_{\theta^o} + \nabla \nabla \ell(\theta^o|y)(\theta - \theta^o) = 0 \quad (16)$$

em que $\nabla \nabla \ell(\theta^o|y) = \nabla^2 \ell(\theta^o|y) = H_f$, ou seja, aplica-se o gradiente em cada elemento de $U_{\alpha,\beta,\gamma,\sigma}$ obtendo-se uma matriz quadrada 4×4 , essa matriz H_ℓ equivale à Matriz hessiano de $\ell(\theta|y)$.

$$H_\ell = \begin{bmatrix} \frac{\partial^2}{\partial \alpha^2} \ell(\theta, y) & \frac{\partial^2}{\partial \alpha \beta} \ell(\theta, y) & \frac{\partial^2}{\partial \alpha \gamma} \ell(\theta, y) & \frac{\partial^2}{\partial \alpha \sigma} \ell(\theta, y) \\ \frac{\partial^2}{\partial \beta \alpha} \ell(\theta, y) & \frac{\partial^2}{\partial \beta^2} \ell(\theta, y) & \frac{\partial^2}{\partial \beta \gamma} \ell(\theta, y) & \frac{\partial^2}{\partial \beta \sigma} \ell(\theta, y) \\ \frac{\partial^2}{\partial \gamma \alpha} \ell(\theta, y) & \frac{\partial^2}{\partial \gamma \beta} \ell(\theta, y) & \frac{\partial^2}{\partial \gamma^2} \ell(\theta, y) & \frac{\partial^2}{\partial \gamma \sigma} \ell(\theta, y) \\ \frac{\partial^2}{\partial \sigma \alpha} \ell(\theta, y) & \frac{\partial^2}{\partial \sigma \beta} \ell(\theta, y) & \frac{\partial^2}{\partial \sigma \gamma} \ell(\theta, y) & \frac{\partial^2}{\partial \sigma^2} \ell(\theta, y) \end{bmatrix}$$

A Matriz Hessiano H_ℓ é equivalente ao inverso aditivo da matriz de informação observada de Fisher $H_\ell = -I(\theta)$, que dá uma medida da curvatura da variável Y baseada apenas na amostra aleatória observada, em outras palavras a informação observada de Fisher é uma medida de informação local, que mede a quantidade de informação de Y com respeito a θ . Isolando o vetor θ na equação 16, obtemos uma fórmula iterativa.

$$\theta^{o+1} \approx \theta^o + J(\theta^o)^{-1} U(\theta^o) \quad (17)$$

Em que o estimador MV de $\hat{\theta}$ é atualizado iterativamente tendo como valores iniciais, chutes de θ^o . Para garantir que o ponto de máximo global foi atingido, a matriz hessiano deve satisfazer as seguintes condições:

- $\det(H_\ell) > 0$;

- $\frac{\partial^2}{\partial \alpha^2} \ell(\theta, y) < 0$ ou $\frac{\partial^2}{\partial \beta^2} \ell(\theta, y) < 0$ ou $\frac{\partial^2}{\partial \gamma^2} \ell(\theta, y) < 0$ ou $\frac{\partial^2}{\partial \sigma^2} \ell(\theta, y) < 0$.

O método Score de Fisher consiste em utilizar, na otimização numérica da expressão 17 a matriz de informação esperada de Fisher ao invés de observada, a matriz $I(\theta^o)$ é uma medida de informação global e é definida como a esperança da informação observada, $E(-\nabla^T \nabla \ell(\theta^o | y)) = I(\theta^o)$, em outras palavras representa a média da curvatura da função de verossimilhança. Assim o algoritmo *Fisher-Scoring* é dado pela equação 18;

$$\theta^{o+1} \approx \theta^o + I(\theta^o)^{-1} U(\theta^o) \quad (18)$$

1.1.5 Coeficiente de determinação

Em modelos não lineares há divergências a respeito do uso de R^2 . De acordo com Ratkowsky (1990) o coeficiente de determinação R^2 não possui interpretação prática em MNLN, Souza (1998) diz que a qualidade de ajustamento pode ser medida pelo quadrado do coeficiente de correlação de y e $\hat{y}$. A simulação de Spiess e Neumeyer (2010) com modelos não lineares mostra que R^2 não consegue diferenciar bons e péssimos modelos, em 43% dos casos R^2 acerta em classificar bons modelos. O critério de informação de Akaike (AIC) e bayesiano (BIC) podem ser usados na comparação de entre modelos, mas não qualidade de ajuste.

A interpretação clássica do coeficiente de determinação (R^2) consiste na validade da decomposição da soma de quadrados totais em Resíduo e Regressão.

$$R^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} = \frac{SQ_{reg}}{SQ_{tot}} \quad (19)$$

portanto, a interpretação do resultado explana o quanto o modelo ajustado explica os dados, com as somas de quadrados sendo obtidas pela ANOVA. Em modelos não lineares, de acordo com Gujarati & Porter (2012) a decomposição da equação 19 não é verdadeira, isto implica que a ANOVA para computação do teste F, que é comum de MLN não é utilizada, outra implicação é na perda da interpretação usual de R^2 .

1.1.6 The Poisson zeros' trick

A especificação de uma função de verossimilhança arbitrária pode ser feita por meio do truque zero de Poisson, ou do truque de Bernoulli, ambas as aproximações possuem o mesmo efeito, porém Ntzoufras (2008) recomenda utilizar o truque zero de Poisson. A interface Openbugs, conta com a função *dloglik()*, que permite utilizar funções de verossimilhanças que não estão dentro do pacote.

A função *dloglik* () implementa o truque zero de Poisson, método que consiste em especificar uma distribuição de forma indireta. A função de verossimilhança de uma fdp $f(x_i, \theta)$ pode ser reescrita:

$$\begin{aligned}\mathcal{L}(\theta|y) &= \prod_{i=1}^n f(y_i, \theta) \\ \mathcal{L}(\theta|y) &= \prod_{i=1}^n e^{\log f(y_i, \theta)} \\ \mathcal{L}(\theta|y) &= \prod_{i=1}^n \frac{e^{-(-\log f(y_i, \theta))} (-\log f(y_i, \theta))^0}{0!} \\ \mathcal{L}(\theta|y) &= \prod_{i=1}^n P(Y|y=0); Y \sim \text{Poisson}(\mu = -\log f(y_i, \theta))\end{aligned}$$

considerando $l_i = \log f(y_i, \theta)$, a quantidade $e^{-(-l_i)}(-l_i)^0/0!$ equivale a probabilidade de $P(Y|y=0)$ tal que $Y \sim \text{Poisson}(\mu = -l_i)$, para garantir que $\mu > 0$ multiplica-se a função de verossimilhança por e^{-nK} , desta forma, K deve ter um valor tal que $K > l_i$.

$$\begin{aligned}\mathcal{L}(\theta|y) &= \prod_{i=1}^n \frac{e^{-(K-l_i)}(K-l_i)^0}{0!} \\ \mathcal{L}(\theta|y) &= \prod_{i=1}^n P(Y|y=0); Y \sim \text{Poisson}(\mu = K - l_i)\end{aligned}\tag{20}$$

Tomemos como exemplo a implementação da distribuição Skew-Normal de Fernández & Steel dada pela equação 21.

$$f_Y(y) = \frac{\lambda\sqrt{2}}{\sigma^2(1+\lambda^2)\lambda^{2H(x)-1}\sqrt{\pi}} \exp\left\{-0.5\left(\frac{y-\mu}{\sigma\lambda^{2H(x)-1}}\right)^2\right\}\tag{21}$$

em que $H(x)$ é a função de Heavside.

$$l_i = -\log \sigma^2 - \log(\lambda + \lambda^{-1}) - (2H(x) - 1) \log \lambda - 0.5 \left(\frac{y-\mu}{\sigma\lambda^{2H(x)-1}}\right)^2.$$

Exemplo 4: Considerando uma distribuição Skew-Normal de Azzalini definida por $f_X(x) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\lambda \frac{x-\mu}{\sigma}\right)$, gerando 100 valores aleatórios dessa distribuição e especificando sua função de verossimilhança pela representação de Poisson dada pela equação 20 e sendo o l_i dada por:

$$l_i = \log f_X(x) = \log 2 - \log \sigma + \log \phi\left(\frac{x-\mu}{\sigma}\right) + \log \Phi\left(\lambda \frac{x-\mu}{\sigma}\right)$$

Obtemos a função de verossimilhança bidimensional graficamente representada pela Figura 2

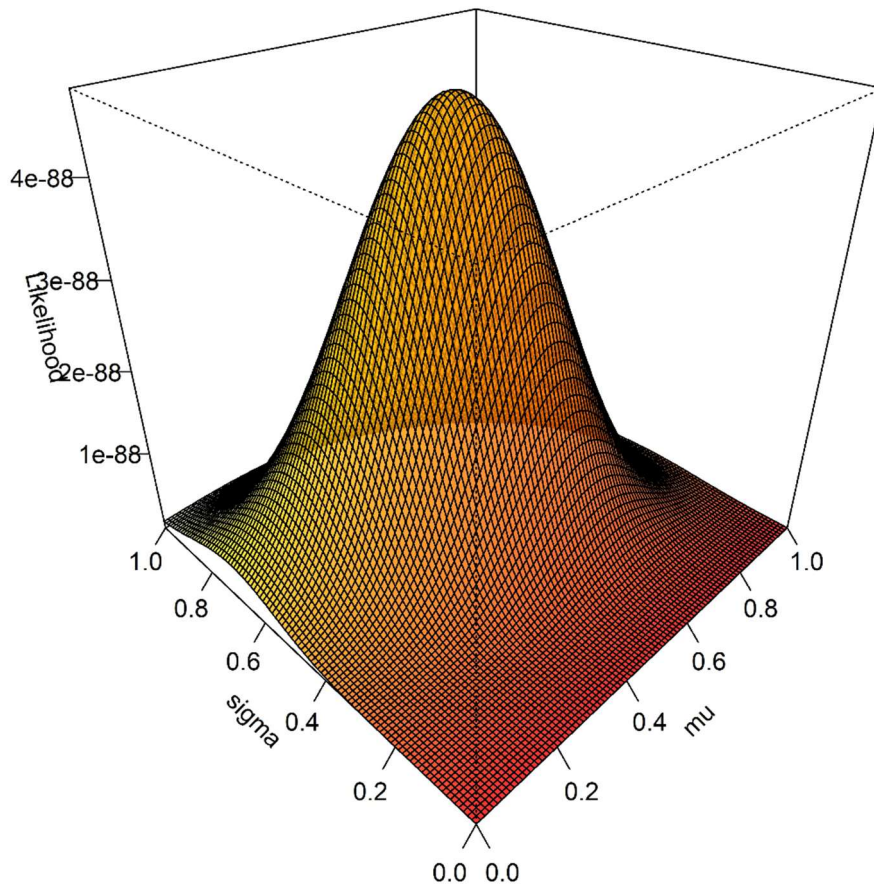


Figura 2: Função de verossimilhança da distribuição Skew-Normal de Azzalini em função de μ e σ , considerando $\lambda = 2$.

Fonte: O autor

1.1.7 Truque de Bernoulli

Uma alternativa ao truque zero de Poisson, é o truque de Bernoulli, que se utiliza do mesmo propósito de se representar uma função de verossimilhança, só que por meio da distribuição de Bernoulli. A função de verossimilhança de uma fdp $f(y_i, \theta)$ é dada por:

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n f(y_i, \theta)$$

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n e^{\log f(y_i, \theta)}$$

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n (e^{\log f(y_i, \theta)})^1 (1 - e^{\log f(y_i, \theta)})^0$$

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n P(X|x = 1; p = e^{\log f(y_i, \theta)}); X \sim \text{Bernoulli}(p = e^{\log f(y_i, \theta)})$$

Para garantir que o termo $e^{\log f(y_i, \theta)}$ pertença a $(0,1)$, multiplica-se $\mathcal{L}(\theta|y)$ por e^{-nK} tal que $\log f(y_i, \theta) < K$.

1.1.8 Computação Bayesiana

A integração de Monte Carlo consiste num método de encontrar soluções determinísticas por meio de métodos não determinísticos com auxílio da geração de números aleatórios. Um paralelo pode ser traçado com a soma de Riemann. Ao considerar um intervalo $I = [a, b]$ no eixo x particiona-se I na forma $a = t_0 < t_1 < t_2 \dots < t_n = b$ considerando que $c_i \in [a, b]$ tal que $t_{i-1} < c_i < t_i$, a soma de Riemann de uma função f contínua $f: \mathbb{R} \rightarrow \mathbb{R}$ em $[a, b]$ é definida como $S(f) = \sum_{i=1}^n f(c_i) \Delta_{t_i}$ em que $\Delta_{t_i} = t_i - t_{i-1}$ e como resultado $\lim_{n \rightarrow +\infty} \sum_{i=1}^n f(c_i) \Delta_{t_i} = \int_a^b f(x) dx$.

Outra maneira de encontrar o resultado de $\int_a^b f(x) dx$ é por meios não determinísticos. Considerando uma V.A. X tal que $X \sim Unif(a, b)$, utilizando do artifício:

$$A = (b - a) \int_a^b \frac{1}{b - a} f(x) dx$$

$$A = (b - a) E_X(f(X)) \approx \frac{b - a}{n} \sum_{i=1}^n f(X_i)$$

Em que os valores de X_i são gerados diretamente da distribuição $Unif(a, b)$. A limitação do uso da distribuição uniforme se deve aos limites de integração em $[a, b]$, caso o suporte de $f(x)$ seja em todo $\mathbb{R}$, a solução consiste em utilizar uma distribuição candidata X cuja densidade seja $g(X)$ e de fácil geração de números aleatórios.

$$A = \int_{-\infty}^{+\infty} g(x) \frac{f(x)}{g(x)} dx = E_X \left(\frac{f(X)}{g(X)} \right) \approx \frac{1}{n} \sum_{i=1}^n \frac{f(X_i)}{g(X_i)}$$

É possível também utilizar-se deste método para determinação da esperança e variância de uma V.A. de resolução difícil ou não analítica. Considerando um modelo bayesiano uniparamétrico, a distribuição marginal de $f_Y(y)$ pode ser obtida utilizando a própria priori como distribuição candidata de fácil geração de números aleatórios ao invés de utilizar uma função $g(X)$ arbitrária.

$$\pi(\theta|y) = \frac{\ell(y|\theta)\pi(\theta)}{\int_R \ell(y|\theta)\pi(\theta) d\theta}$$

Gerando valores $\theta = (\theta_1, \theta_2, \dots, \theta_n)$.

$$\int_R \ell(\mathbf{y}|\theta)\pi(\theta)d\theta = E_{\theta}(\ell(\mathbf{y}|\theta)) \approx \frac{1}{n} \sum_{i=1}^n \ell(\mathbf{y}|\theta_i)$$

Com $f_Y(y)$ determinada o resultado da posteriori $\pi(\theta|y)$ que pode não ser conhecida, o mesmo artifício é utilizado para se determinar o estimador bayesiano usual da posteriori.

$$E_{\theta|y}(\theta) = \frac{1}{\frac{1}{n} \sum_{i=1}^n \ell(y|\theta_i)} \int_R \theta \ell(y, \theta) \pi(\theta) d\theta = \frac{E_{\theta}(\theta \ell(y, \theta))}{\frac{1}{n} \sum_{i=1}^n \ell(y|\theta_i)}$$

$$E_{\theta|y}(\theta) = \frac{E_{\theta}(\theta \ell(y, \theta))}{\frac{1}{n} \sum_{i=1}^n \ell(y|\theta_i)} \approx \frac{\sum_{i=1}^n \theta_i \ell(y, \theta_i)}{\sum_{i=1}^n \ell(y|\theta_i)}$$

Outra limitação consiste na priori ser uma distribuição de difícil geração de números aleatórios, a solução consiste em escolher uma função $g(X)$ denominada função de importância de fácil geração de números aleatórios. Considerando novamente a marginal de Y .

$$f_Y(y) = \int_R g(\theta) \frac{\ell(y|\theta)\pi(\theta)}{g(\theta)} d\theta = E_{\theta} \left(\frac{\ell(y|\theta)\pi(\theta)}{g(\theta)} \right) \approx \frac{1}{n} \sum_{i=1}^n \frac{\ell(y|\theta_i)\pi(\theta_i)}{g(\theta_i)}$$

Determina-se uma aproximação da esperança para a posteriori.

$$E_{\theta|y}(\theta) = \frac{1}{f_Y(y)} \int_R g(\theta) \frac{\theta \ell(y|\theta)\pi(\theta)}{g(\theta)} d\theta = \frac{1}{f_Y(y)} E_{\theta} \left(\frac{\theta \ell(y|\theta)\pi(\theta)}{g(\theta)} \right)$$

$$\approx \frac{\sum_{i=1}^n \frac{\theta_i \ell(y|\theta_i)\pi(\theta_i)}{g(\theta_i)}}{\sum_{i=1}^n \frac{\ell(y|\theta_i)\pi(\theta_i)}{g(\theta_i)}}.$$

1.1.9 Algoritmo de Aceitação-Rejeição

Uma forma de gerar números aleatórios de uma distribuição de probabilidade $f(x)$ de difícil geração randômica, é por meio do algoritmo de aceitação-rejeição que utiliza de uma distribuição de probabilidades candida $g(x)$ conhecida e de fácil geração de números aleatórios tal que $M = \sup \left\{ \frac{f(x)}{g(x)} \right\} < +\infty$. O Algoritmo segue os seguintes passos (Casella & Berger):

1º Passo: Gerar valores $\{x_i\}_{i=1}^n$ de $g(x)$;

2º Passo: Gerar valores $\{u_i\}_{i=1}^n$ de uma distribuição uniforme (0,1);

3º Passo: Se $u_i < \frac{1}{M} \frac{f_Y(x_i)}{g_X(x_i)}$ então $x_i = y_i$, caso contrário retorna-se ao passo 1.

Exemplo 5: O pacote *ggridges* (v.0.5.6) computa vários gráficos num só, entretanto o pacote só computa gráficos de densidades provenientes de amostras aleatórias e não pela definição de uma função. Dado a distribuição *Skew-Laplace-Normal*, tal distribuição cuja densidade é dada pela equação 22.

$$f_Y(y) = \frac{1}{2\sigma} \exp \left\{ -\frac{1}{2} \left| \frac{y - \mu}{\sigma} \right| \right\} \Phi \left(\lambda \frac{y - \mu}{\sigma} \right) \quad (22)$$

Não é possível gerar números pela transformação uniformizante e sua representação estocástica contém uma densidade que também não possui geração aleatória por métodos diretos. A solução seria utilizar-se o algoritmo de Aceitação-Rejeição utilizando a distribuição Laplace como candidata e $M = 2.001$. A Figura 3 mostra os resultados da metodologia.

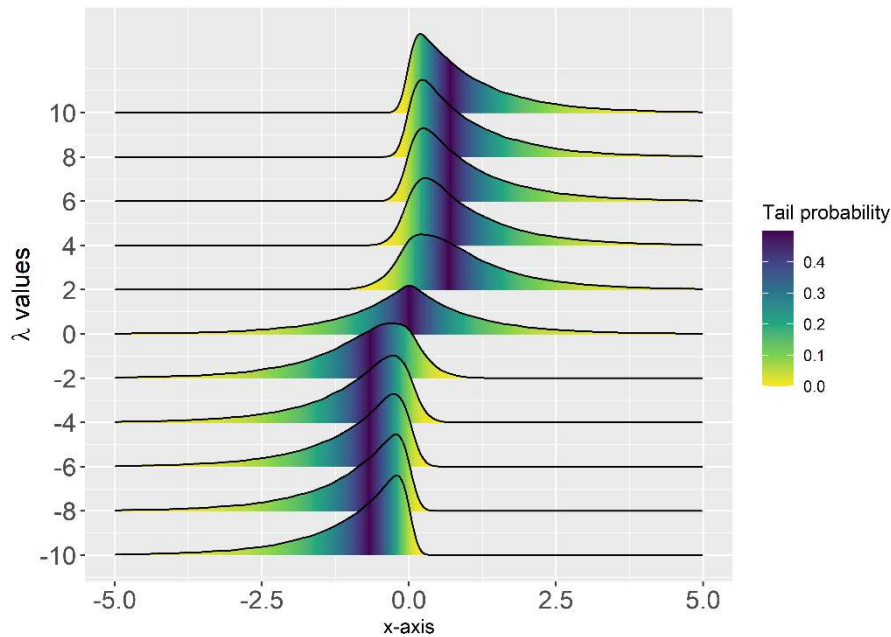


Figura 3: Densidades da distribuição Skew-Laplace-Normal padronizada e variando valores de λ .

Fonte: O autor.

1.1.10 Amostrador de Gibbs

O amostrador de Gibbs consiste em obter os valores dos parâmetros de interesse utilizando a DCCP proveniente de uma distribuição conjugada, como a DCCP é conhecida, é possível gerar valores diretamente das mesmas. A DCCP de cada θ_i genericamente.

$$\theta_i = \pi(\theta_i | \theta_1, \theta_2, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_{p-1}, \theta_p)$$

Algoritmo de Metropolis-Hastings se dá numa série de passos:

1º Passo: Definir um conjunto de valores iniciais para $\theta_1^{(o)} = (\theta_1^{(o)}, \theta_2^{(o)}, \dots, \theta_p^{(o)})$

2º Passo: Obter $\theta_i^{(t)}$ a partir de $\theta_i^{(t-1)}$ e atualizar a DCCP de cada parâmetro

$$\theta_1 \sim \pi(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_{p-1}^{(t-1)}, \theta_p^{(t-1)})$$

$$\theta_2 \sim \pi(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_{p-1}^{(t-1)}, \theta_p^{(t-1)})$$

$$\theta_3 \sim \pi(\theta_3 | \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{p-1}^{(t-1)}, \theta_p^{(t-1)})$$

$$\theta_p \sim \pi(\theta_p | \theta_1^{(t)}, \theta_2^{(t)}, \theta_3^{(t)}, \dots, \theta_{p-1}^{(t)})$$

1.2 REFERÊNCIAS

BUSSAB, W. D. O., MORETTIN, P. A. **Estatística básica**. São Paulo: Saraiva, 2010. 540p.

CORDEIRO, G.M., PAULA, G.A. Fitting non-exponential family nonlinear models in GLIM by using the o set-facility. **Lecture Notes in Statistics**, vol. 57, p. 105–114, 1989.

FERNANDES, T. J., MUNIZ, J. A., PEREIRA, A. A., MUNIZ, F. R., MUIANGA, C. A. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum. Technology**, v.37, p.397-402, 2015.

FERNANDES, F. A., SILVA, ÉDIPO M., LIMA, K. P., JANE, S. A., FERNANDES, T. J., MUNIZ, J. A. Parameterizations of the von bertalanffy model for description of growth curves. **Brazilian journal of biometrics**, v.38, p.369–384, 2020.

GUJARATI, D. N., DINIZ, E. M. **Econometria básica**. 5 ed. AMGH, 920p, 2012.

LUNA, J.G., OLINDA, R.A. **Introdução a modelos lineares**. São Paulo. eduepb, 2014. 164p.

MARTINS FILHO, S. et al. Abordagem Bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, Santa Maria, v.38, n.6, p.1516-1521, 2008.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**. v.135, p.370-384, 1972.

SALSBURG, D. **Uma Senhora toma chá: como a estatística revolucionou a ciência no século XX**. Rio de Janeiro: Zahar, 289p. 2009.

SOUZA, G.S. **Introdução aos modelos de Regressão linear e não linear**. Brasília. Embrapa, 1998, 489p.

STEWART, J. **Cálculo volume 2**. São Paulo. Cengage Learning, 2016. 672p.



Comparison of Normal, Logistic, Laplace, and Student's *t* distributions for experimental error in the Bayesian description of dry matter accumulation in *Allium sativum*

George Lucas Santana de Moura^{1*} Felipe Guzzo¹ Paulo Roberto Cecon¹
Sebastião Martins Filho¹ Antônio Policarpo Souza Carneiro¹ Moyses Nascimento¹

¹Departamento de Estatística Aplicada e Biometria, Universidade Federal de Viçosa (UFV), 36570-900, Viçosa, MG, Brasil. E-mail: gelucas.moura.2016@gmail.com. *Corresponding author.

ABSTRACT: This study assessed distributions associated with Bayesian nonlinear modeling error in the description of total plant dry matter accumulation (TDMA) of *Allium sativum* as a function of days after planting (DAP). According to the DIC criterion, Logistic and Gompertz models that use student's *t* distribution error exhibited the highest DIC with logistic error distribution. In general, the difference of DIC in all the scenarios was not more than 5. The Bayes factor (BF) criterion showed no difference in the Logistic and Gompertz model fit when four distributions are used for the errors, where BF values do not exceed 2. Posterior distributions and the usual estimators of Logistic and Gompertz model parameters were similar even for varied error distribution. In summary, there was no difference in the use of 4 distributions associated with the modeling error of garlic plant growth by the Bayes factor, whereby the results showed that alternating between error distributions significantly changes the number of Markov Chain Monte Carlo (MCMC) iterations.

Key words: Bayesian regression, Nonlinear regression, MCMC, Symmetrical location-scale family, Empirical Bayes.

Comparação das distribuições Normal, Logística, Laplace e *t* de Student para o erro experimental na descrição bayesiana do acúmulo de matéria seca de *Allium sativum*

RESUMO: O objetivo deste trabalho foi avaliar algumas distribuições associadas ao erro na modelagem não linear bayesiana na descrição do acúmulo de matéria seca total da planta (MSTP) de *Allium sativum* em função dos dias após o plantio (DAP). Pelo critério DIC os modelos Logístico e Gompertz que utilizam a distribuição do erro *t* de Student apresentaram a melhor qualidade de ajuste, sendo que o modelo Logístico apresentou o maior DIC com a distribuição de erro Logística. No geral, a diferença de DIC em todos os cenários não apresentou valores superiores a cinco. Pelo critério do Fator de Bayes (FB), não houve diferença no ajuste do modelo Logístico e Gompertz quando se utilizam as quatro distribuições para os erros, sendo que os valores de FB não superaram 2. As distribuições a posteriori e os estimadores usuais dos parâmetros dos modelos Logístico e Gompertz apresentaram semelhanças mesmo variando a distribuição do erro. Em suma não houve diferença na utilização das quatro distribuições associadas ao erro na modelagem do crescimento planta de alho pelo fator de Bayes, sendo que os resultados mostram que alternar entre as distribuições dos erros altera de forma significativa o número de iterações de MCMC.

Palavras-chave: Regressão bayesiana, Regressão não linear, MCMC, Família simétrica de locação-escala, Bayes empírico.

INTRODUCTION

Sigmoid models are widely used in agricultural science to describe animal and plant growth. Most of the models, including Logistic and Gompertz, are analytical solutions of ordinary differential equations (STEWART, 2016).

According to CORDEIRO & DEMÉTRIO (2008), since the 1970s, nonlinear regression theory has restricted the use of nonlinear models (NLM) to the assumption of normality for the residual and consequently, the response variable. With the introduction of generalized linear models (GLM), NELDER & WEDDERBURN in 1972 defined the

response variables that belong to the exponential family. Likewise, CORDEIRO & PAULA (1989) defined the exponential family for a normal nonlinear model in which the systematic component is not a linear combination of the parameters.

Nonlinear regression with a normal error is susceptible to extreme observations. However, the assumption of normal error can be relaxed using different symmetrical and asymmetrical distributions, in both linear and nonlinear models under Bayesian estimation (DE LA CRUZ & BRANCO, 2009; ROSSI & SANTOS, 2014).

In terms of dimensionality, a data base to be analyzed should contain at least 4 observations,

CAPÍTULO 1: Abordagem Bayesiana robusta de modelos não lineares aplicados ao crescimento de *capisicum annuum* L. utilizando distribuições simétricas de locação-escala

Resumo: O presente trabalho tem como objetivo principal fazer um levantamento das principais distribuições simétricas utilizadas para substituir a distribuição Normal para o resíduo em modelos não lineares. Na descrição das curvas utilizou-se os conhecidos modelos Logístico, Gompertz e Von Bertalanffy, porém com reparametrizações adequadas que permitem que os modelos estejam próximos da linearidade e que o parâmetro β tenha interpretação biológica como a abscissa associada ao ponto de inflexão da curva de crescimento. Como ferramenta para encontrar prioris informativas. Para comparação de modelos utilizou-se do critério de Deviance (DIC) e do índice de concordância de Lin para o caso frequentista, DIC e Fator de Bayes (FB) para o caso bayesiano. Os dados analisados são provenientes de um experimento montado no departamento de Agronomia da UFV utilizando 8 genótipos de pimenta *Capsicum Annuum* L. e computando suas características fenotípicas em função dos dias após o florescimento. Os erros experimentais além da Normal, aumentaram a precisão dos modelos e melhoraram a qualidade de ajuste. O modelo Logístico com o ponto de inflexão em torno de 10 a 11 dias após o florescimento, apresentou a maior precisão em todos os cenários e mostrou-se o mais plausível pelo Fator de Bayes, com o uso do erro Normal generalizado de Nadarajah melhorando significativamente a qualidade de ajuste geral.

Palavras-chave: Bayes empírico, Regressão não linear, família simétrica de locação-escala, bootstrap.

2.1 INTRODUÇÃO

O uso de modelos sigmóides é amplamente utilizado na ciência agropecuária para descrever o crescimento de animais e plantas. A maioria dos modelos, dentre eles o Logístico e Gompertz são soluções analíticas de equações diferenciais ordinárias (STEWART, 2016).

Na teoria da regressão não linear, de acordo com CORDEIRO & DEMÉTRIO (2008), o uso de modelos não lineares (MNL) estava restrito, desde a década de 70, à suposição de normalidade para o resíduo, e consequentemente para variável resposta. Com a introdução dos Modelos lineares generalizados (MLG), Nelder e Wedderburn em 1972 definiram variáveis

respostas que pertencem a família exponencial. Da mesma forma CORDEIRO & PAULA (1989) definiram a família exponencial para um modelo não linear Normal em que o componente sistemático não é uma combinação linear dos parâmetros.

A Regressão não linear com erro Normal é suscetível a presença de observações extremas. No entanto, a pressuposição de normalidade do erro pode ser relaxada utilizando diversas distribuições simétricas e assimétricas, tanto em modelos lineares como não lineares sob a estimação bayesiana (DE LA CRUZ & BRANCO, 2009; ROSSI & SANTOS, 2014).

Em termos de dimensionalidade, os modelos de crescimento Logístico, Gompertz e Von Bertalanffy que contam com 3 parâmetros (α, β, γ) necessitam de um mínimo de 4 observações, para que o número de parâmetros seja menor que o conjunto dos dados ($n > p$) na estimação frequentista. Na estimação de máxima verossimilhança, além dos 3 parâmetros usuais contidos no parâmetro de locação de um modelo não linear com erro experimental simétrico, é necessário estimar o parâmetro de escala (σ) e o parâmetro (v) no caso do erro *t* de Student. A dimensionalidade salta de 3 para 5, e as observações precisam agora ser de 6 pontos. O trabalho de Moura et al. (2023) justifica a regressão bayesiana ao estimar 5 parâmetros com apenas 4 observações.

Na inferência bayesiana, a escolha de uma priori é feita antes de se ter acesso aos dados. Existem várias metodologias como priori impróprias, distribuições conjugadas ou por avaliação da meta-análise. Em FIRAT et al (2016) e MACEDO et al (2017) os parâmetros do modelo Logístico e Gompertz (α, β, γ) seguem uma distribuição *a priori* Normal.

Quando a obtenção da distribuição *a priori* provém dos dados, essa metodologia é denominada de empírica. A inferência que faz uso da distribuição *a priori* empírica é denominada como abordagem bayesiana empírica o que não é necessariamente uma inferência bayesiana, pois os dados são utilizados duas vezes, uma vez na função de verossimilhança e uma vez na distribuição *a priori*. E, entretanto, essa metodologia é uma boa aproximação para inferência bayesiana. (CARLIN & LOUIS, 2008)

Em termos de inferência paramétrica, Gujarati & Porter (2012) e Souza (1998) explanam que os testes de significância e intervalos de confiança de Modelos não lineares Normal são válidos assintoticamente, os testes F, t, intervalos e regiões de confiança dependem da normalidade assintótica dos estimadores paramétricos. Na teoria bayesiana, os resultados dos intervalos de credibilidade e dos testes de significância são válidos independentemente do tamanho amostral.

De acordo com Ruli & Ventura (2020), os intervalos de credibilidade, intervalos de

confiança e distribuições de confiança (tradução livre de *confidence distribution*), são concordantes inclusive em suas estimativas pontuais quando a amostra é grande, quando as prioris são informativas e quando as aproximações assintóticas são de ordem superior. Rossi & Santos (2014) utilizam-se do fato da proximidade das estimações frequentistas e bayesianas, ao utilizar estimativas frequentistas no uso de valores iniciais na estimação bayesiana.

No trabalho de Moura et al. (2023) é utilizado o ajuste dos modelos Logístico e Gompertz utilizando os erros experimentais Normal, Logístico, Laplace e Student t, o autor conclui que se não houver diferença de qualidade de ajuste, então utiliza-se os modelos robustos para melhor convergência computacional, com alguns casos resultando numa economia computacional em torno de 98%.

O objetivo do presente trabalho é implementar e comparar, as distribuições Normal, Logístico, Laplace, Student t e Normal generalizada para o erro experimental utilizando os modelos de crescimento Logístico, Gompertz e Von-Bertalanffy.

2.2 MATERIAL E MÉTODOS

O banco de dados a ser trabalhado é composto de 8 acessos de *Capsicum annum* L. sendo estes, 1 - Pimenta vulcão; 2 - Pimenta Cayene; 3 - Pimenta Peter; 4 – Pimenta picante; 5 - Pimenta jamaica Yellow; 6 - Pimentão quadrado; 7 - Pimentão Ikeda; 8 – Pimentão Rubi gigante. O experimento foi conduzido no departamento de Agronomia da Universidade Federal de Viçosa (UFV), localizada no município de Viçosa, Minas Gerais, Brasil. (20° 45 Sul e 42° 51 Oeste, altitude média de 650 m). Conforme expressa na Figura 1, a característica fenotípica obtida do experimento se trata do comprimento do fruto - DAF (variável explanatória) em função dos dias após o florescimento – DAF (variável independente).

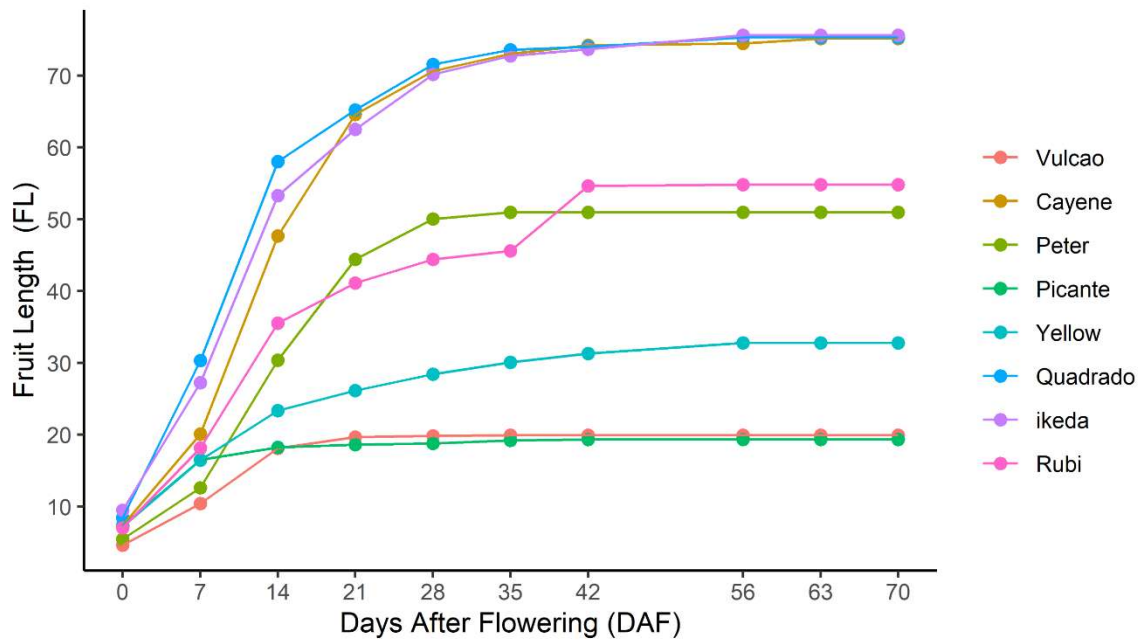


Figura 4: *Scatterplots* do comprimento do fruto dos Acessos de *Capsicum* em função dos dias após o florescimento.

Fonte: O autor.

Geralmente, a utilização de uma modelo não linear de crescimento é escolhida conforme a natureza intrínseca do fenómeno biológico que está sendo analisado, ou verificando o comportamento dos dados, ou por meio da literatura disponível. Na Figura 4 é possível verificar o comportamento sigmóides da maioria dos Acessos. Nos trabalhos de Oliveira et al. (2021) e Teixeira et al. (2023), ambos utilizam modelos de crescimento não lineares e sigmóides para o ajuste de características fenotípicas de *Capsicum*.

Os modelos sigmóides utilizados para o ajuste dos Acessos representados na Figura 1, são as versões reparametrizadas dos modelos Logístico, Gompertz e Von Bertalanffy representados pelas equação (23), (24) e (25).

$$\eta(\theta) = \frac{\alpha}{1 + e^{\gamma(\beta - x)}} \quad (23)$$

$$\eta(\theta) = \alpha \exp\{-\exp\{\gamma(\beta - x)\}\} \quad (24)$$

$$\eta(\theta) = \alpha \left(1 - \frac{\exp\{\gamma(\beta - x)\}}{3}\right)^3 \quad (25)$$

Com as seguintes interpretações paramétricas:

α representa o peso assintótico;

β é o valor em que organismo atinge a sua velocidade máxima de crescimento, e quando atinge metade de seu crescimento máximo;

γ representa a taxa de crescimento.

Para definir os modelos de regressão não linear simétricas, na seção a seguir definiremos

as distribuição elípticas, esféricas-simétricas e a família simétrica de locação-escala.

2.2.1 Distribuições elípticas e simétricas

No contexto da regressão robusta, existem várias distribuições simétricas com caudas mais pesadas que a gaussiana, Cordeiro et al. (2000) explana uma família de distribuições que compartilham várias propriedades semelhantes, família esta denominada família simétrica de locação-escala cuja algumas propriedades são derivadas da distribuições esféricas simétricas.

Considerando um vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ e vetor de médias p -dimensional $\boldsymbol{\mu} \in \mathbb{R}^p$ com uma matriz de variância-covariância $\boldsymbol{\Sigma}$ simétrica e positiva definida, ao considerar um gerador de densidade $g(\cdot) \in \mathbb{R}^*$, define-se uma distribuição elípticas simétrica como função de formas quadráticas.

$$f_{\mathbf{X}}(x) = C_p |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g\left((x - \boldsymbol{\mu})^T \mathbf{V}^{-1} (x - \boldsymbol{\mu})\right) \quad (26)$$

em que C_p é uma constante normalizadora e $\int_{\mathbb{R}^p} r^{p/2-1} g(r^2) dr < +\infty$, no espaço $(0, +\infty)$, se a distribuição possui k -ésimo momentos então $r^{k+p/2-1} r^2$ deve ser integrável. A função característica de $\mathbf{X}$, para alguma função real ψ e um vetor de variáveis auxiliar $\mathbf{t}$ é dada por:

$$\varphi_{\mathbf{X}}(i\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu}} \psi(\mathbf{t}'\mathbf{V}\mathbf{t}) \quad (27)$$

em que $\boldsymbol{\Sigma} = -2\psi'(0)\mathbf{V}$, e ψ é denominada função geradora de característica que possui a condição de $\psi(0) = 1$, não se deve confundir a função geradora de densidade $g(\cdot)$. Se um vetor aleatório segue uma distribuição elíptica, então utilizamos a notação $\mathbf{X} \sim E(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$.

Exemplo 6: Tomando um vetor $\mathbf{X}$ aleatório dado pela equação 26, seja o gerador de densidade $g(\omega) = e^{-\frac{\omega}{2}}$ e o gerador de característica $\psi(u) = e^{-\frac{u}{2}}$, neste caso específico, ambas as funções são as mesmas, então $\mathbf{X}$ possui distribuição Normal Multivariada com densidade.

$$f_{\mathbf{X}}(x) = (2\pi)^{-p/2} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (x - \boldsymbol{\mu})\right) \quad (28)$$

com:

- $E(\mathbf{X}) = \boldsymbol{\mu}$
- $\boldsymbol{\Sigma} = COV(\mathbf{X}) = -2(-\frac{1}{2})\mathbf{V}$;
- $\varphi_{\mathbf{X}}(i\mathbf{t}) = \exp\{i\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})\}$.

Ao fazer uma padronização $\mathbf{Z} = \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$, obtemos as chamadas distribuições esféricas simétricas. No caso da distribuição normal, a função característica coincide com a função geradora de característica.

Na Figura 5 pode-se observar as curvas de níveis de uma distribuição elíptica e esférica

simétrica.

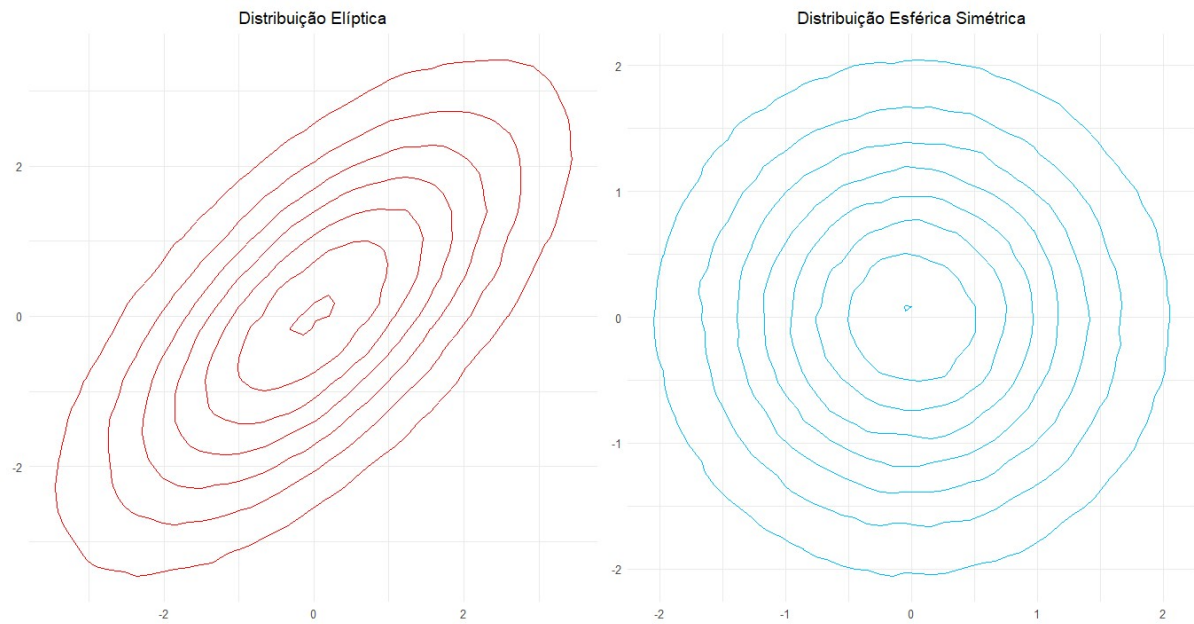


Figura 5 – Curvas de nível da Distribuição elíptica Normal (figura esquerda), e da distribuição esférica simétrica Normal (figura direita).

Fonte: O autor

Na Tabela 1, pode verificar-se algumas das principais distribuições elípticas.

Tabela 1 - Representação da tabela de distribuições elípticas com suas respectivas funções geradoras de densidade

Distribuição	Notação	Função geradora
Normal	$N(\mu, \Sigma)$	$g(u) = c \cdot \exp(-u/2), \quad u \geq 0$
Student-t	$t(\mu, \Sigma, \nu)$	$g(u) = c \cdot (1 + u/\nu)^{-(\nu+n)/2}, \quad u \geq 0$
Normal Contaminada	$CN(\mu, \Sigma, \delta, \tau)$	$g(u) = c \cdot (1 - \delta) \cdot \exp - u/2 + \delta \cdot \tau^{(-n/2)} \cdot \exp - u/(2\tau), \quad u \geq 0$
Cauchy	$C(\mu, \Sigma)$	$g(u) = c \cdot (1 + u)^{-(n+1)/2}, \quad u \geq 0$
Logística	$L(\mu, \Sigma)$	$g(u) = c \cdot e^{(-u)} / (1 + e^{(-u)})^2, \quad u \geq 0$
Power Exponential	$PE(\mu, \Sigma, \alpha)$	$g(u) = c \cdot e^{-u^\alpha/2}, \quad u \geq 0$

Fonte: Lemonte & Patriota (2011)

Com base nas funções geradoras da tabela representada pela Tabela 1 e considerando o caso univariado, obtêm-se algumas distribuições simétricas que pertencem a família de localização-escala, com caudas mais pesadas que a Normal.

2.2.2 Família simétrica de localização-escala

Existem várias distribuições que são simétricas e que pertencem a família de localização-escala, sendo a distribuição Normal pertencente a essa família simétrica de localização-escala

sendo a única que pertence a família exponencial, então algumas propriedades que são unificadas da família exponencial no contexto de MLG e prioris conjugadas no caso bayesiano para outras distribuições além da Normal. Com base na introdução das distribuições elípticas da Seção 2.2.1, pode-se também derivar propriedades comuns dessa família no caso univariado.

Seja um variável aleatória X que pertence a família simétrica de locação-escala, então a densidade da distribuição é dada genericamente por:

$$f_X(x) = \frac{1}{\sigma} g\left(\frac{x - \mu}{\sigma}\right)^2 \quad (29)$$

Sendo $g(\cdot)$ uma densidade conhecida, parâmetro de locação (μ) e escala (σ) com a condição $\int_0^{+\infty} \omega^{-1/2} g(\omega) d\omega = 1$. A função característica de X é dada por $\varphi_X(t) = e^{it\mu} \psi(\sigma^2 t^2)$ para alguma função real $\psi(x)$ diferenciável. A expectância e variância são as mesmas do caso elíptico para $p = 1$

- $E(X) = \mu$;
- $V(X) = -2\varphi'(0)\sigma^2$.

Denota-se essa classe de distribuição de S tal que $X \sim S(\mu, \sigma^2)$.

Exemplo 7: Seja o gerador de característica $\psi(u) = e^{-\frac{u}{2}}$ da distribuição Normal, então $\psi'(u) = -\frac{1}{2}e^{-\frac{u}{2}}$ e $\psi'(0) = -\frac{1}{2}$, conseqüentemente $V(X) = -2\varphi'(0)\sigma^2 = \sigma^2$ e sua função característica, $\varphi_X(t) = \exp\{it\mu - \frac{t^2\sigma^2}{2}\}$. Com base na equação 29 e no gerador de densidade $g(\omega) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\omega}{2}}$ obtêm-se a conhecida distribuição Normal.

2.2.3 Regressão não linear bayesiana com erros simétricos

Na Regressão não linear com erros simétricos, considere o erro experimental que pertence a família simétrica de locação-escala $e_i \sim N(0,1)$, o modelo de Regressão é descrito:

$$y_i = \eta(x_i, \theta) + \sigma e_i \quad (30)$$

em que a distribuição amostral de Y_i também pertence a família simétrica de locação-escala tal que $Y_i \sim S(\eta(x_i, \theta), -2\varphi'(0) \times \sigma^2)$ com média $\eta(x_i, \theta)$ e variância $-2\varphi'(0) \times \sigma^2$, e $\eta(x_i, \theta)$ é o preditor não linear que são as funções de crescimento descritas pelas equações 23, 24 e 25.

A função de verossimilhança de Y na sua forma canônica é expressa por:

$$\mathcal{L}(\theta, \sigma^2; \mathbf{y}) = \frac{1}{\sigma^n} \prod_{i=1}^n g\left(\frac{y_i - \eta(x_i, \theta)}{\sqrt{-2\varphi'(0) \times \sigma^2}}\right)^2 \quad (31)$$

Na Regressão não linear Bayesiana com erro experimental S , utilizando a distribuição

conjunta da amostra (equação 31), e as distribuições a priori conjunta (podem ser independentes) dada por $\pi(\cdot)$, obtemos a distribuição a posteriori dos parâmetros $(\alpha, \beta, \gamma, \tau)$ atualizados pelos dados $\mathbf{y}$, por meio da estrutura de Regressão:

$$p(\alpha, \beta, \gamma, \sigma | \mathbf{y}) \propto \mathcal{L}(\theta, \sigma^2; \mathbf{y}) \times \pi(\alpha, \beta, \gamma, \sigma) \quad (32)$$

Utilizando do algoritmo de metropolis-Hastings, a DCCP de cada parâmetro utiliza-se a priori individual de cada parâmetro.

No trabalho de Cordeiro et al. (2000), o autor trabalha com a correção de viés de uma lista de 8 distribuições que pertencem a família simétrica de locação-escala, entretanto no OpenBugs, as distribuições simétricas de locação-escala que estão na base do programa são apenas 4: Distribuição Normal, Logística, Laplace e Student t. Para implementar outras distriuições simétricas fora da base, utiliza-se “*The Poisson zero’s trick*” para funções de verossimilhanças arbitrárias, obtendo-se l_i da equação:

$$l_i = -\log \sigma + \log g \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} \right)^2$$

utilizando o parâmetro de precisão $\tau = 1/\sigma^2$,

$$l_i = 0.5 \log \tau + \log g \left(\sqrt{\tau} (y_i - \eta(x_i, \theta)) \right)^2 \quad (33)$$

a estrutura bayesiana da equação 33, neste caso:

$$\begin{aligned} \mathcal{L}(\theta | \mathbf{y}) &= \prod_{i=1}^n P(Y | y = 0); Y \sim \text{Poisson}(\mu = K - l_i) \\ p(\alpha, \beta, \gamma, \tau | \mathbf{y}) &\propto \left\{ \prod_{i=1}^n P(Y | y = 0; K - l_i) \right\} \times \pi(\alpha, \beta, \gamma, \tau) \end{aligned} \quad (34)$$

Na Figura 6 estão representadas algumas densidades da distribuição S , sendo estes os modelos:

1) Normal

$$f_Y(y, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (y - \mu)^2 \right\} \quad (35)$$

2) Logístico

$$f_Y(y, \mu, s) = \frac{\exp\left(-\frac{y - \mu}{\sigma}\right)}{\sigma \left\{ 1 + \exp\left(-\frac{y - \mu}{\sigma}\right) \right\}^2} \quad (36)$$

3) Laplace

$$f_Y(y, \mu, s) = \frac{1}{2\sigma} \exp \left(-\frac{|y - \mu|}{\sigma} \right) \quad (37)$$

4) t de Student

$$f_Y(y, \mu, s, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sigma\sqrt{v\pi}} \left\{ 1 + \frac{\left(\frac{y-\mu}{\sigma}\right)^2}{v} \right\}^{-\frac{v+1}{2}} \quad (38)$$

e na Figura 6 as representações das densidades dadas pelas equações 35, 36, 37, e 38.

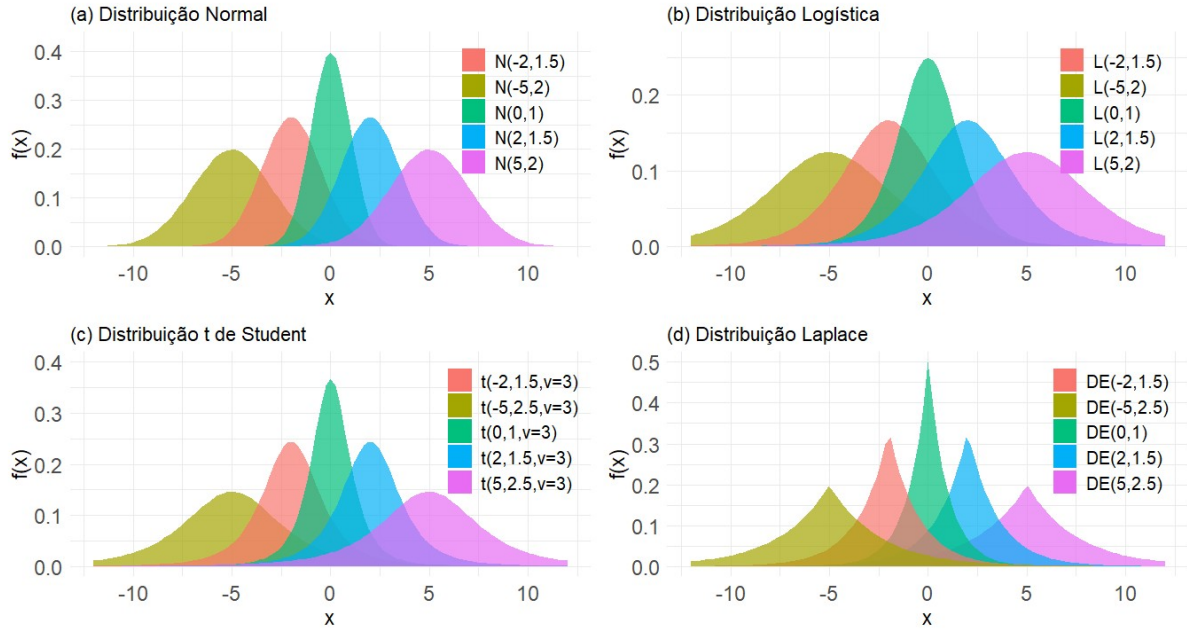


Figura 6: Densides das distribuições Normal, Logística, Laplace e Student t variando os parâmetro de escala e locação.

Fonte: Moura et al. (2023)

Existem outras distribuições simétricas como a distribuição Normal generalizada de Nadarajah (2005) que será estudada na seção a seguir.

2.2.4 Distribuição Normal de Nadarajah (2005)

A distribuição Normal de Nadarajah (2005) é uma generalização da conhecida distribuição Normal em que o termo quadratico é substituído por um parâmetro s , e o termo exponencial é dado por $-|z_i|$. Seja X uma v.a. que segue uma distribuição Normal Generalizada de Nadarajah, e seja μ parâmetro de locação, σ o parâmetro de escala e $x \in \mathbb{R}$, se $X \sim NGN(\mu, \sigma, s)$ então a densidade é da forma,

$$f_X(x) = K \exp \left\{ - \left| \frac{x - \mu}{\sigma} \right|^s \right\} \quad (39)$$

de forma analítica obtemos K para que X seja uma fdp.

$$K^{-1} = \sigma \int_{-\infty}^{+\infty} \exp\{-z^s\} dz = 2\sigma s^{-1} \int_0^{+\infty} \omega^{s^{-1}-1} e^{-\omega} d\omega = 2\sigma s^{-1} \Gamma\left(\frac{1}{s}\right)$$

tal que $z = \frac{y-\mu}{\sigma} \rightarrow dz = \sigma^{-1} dy$; $z = \omega^{s^{-1}} \rightarrow dz = s^{-1} \omega^{s^{-1}-1} d\omega$. Então a distribuição NGN é expressa pela equação (40).

$$f_Y(y) = \frac{s}{2\sigma\Gamma(s^{-1})} \exp\left\{-\left|\frac{y-\mu}{\sigma}\right|^s\right\} \quad (40)$$

O k -ésimo momento amostral pode ser obtida analaticamente tambem pela função gama.

$$E(Z^k) = \frac{s}{2\Gamma(s^{-1})} \int_{-\infty}^{+\infty} z^k \exp\{-|z|^s\} dz$$

O integrante é uma função par se $k = 2n \ \forall n \in \mathbb{N}$, caso contrário é uma função impar.

- $E(Z^k) = \frac{2}{2\Gamma(s^{-1})} \int_0^{+\infty} z^k \exp\{-z^s\} dz = \frac{\Gamma(\frac{k+1}{s})}{\Gamma(\frac{1}{s})}$ se $k = 2n$;
- $E(Z^k) = \frac{1+(-1)^k}{2\Gamma(s^{-1})} \int_0^{+\infty} z^k \exp\{-z^s\} dz = (1 + (-1)^k) \frac{\Gamma(\frac{k+1}{s})}{2\Gamma(\frac{1}{s})}$ se $k = 2n - 1$.

Desta forma, semelhante a distribuição Normal, os momentos centrais são 0 se o momento for impar. Utilizando uma transformação de locação-escala $X = \mu + \sigma Z$ da variável Z e por meio do Teorema Binomial obtemos variância e esperança.

$$E(\mu + \sigma Z)^n = E\left(\sum_{i=0}^n \binom{n}{i} u^{n-i} \sigma^i Z^i\right) = \sum_{i=0}^n \binom{n}{i} u^{n-i} \sigma^i E(Z^i) \quad (41)$$

Desta forma, o n -ésimo momento amostral da equação 41 é dado por:

$$E(X^n) = \sum_{i=0}^n \binom{n}{i} u^{n-i} \sigma^i \frac{1 + (-1)^i}{2\Gamma(s^{-1})} \Gamma\left(\frac{i+1}{s}\right) \quad (42)$$

Com base na equação 42 obtemos a esperança e variância da distribuição NGN.

- $E(X) = \mu$;
- $V(X) = \sigma^2 \frac{\Gamma(\frac{3}{s})}{\Gamma(\frac{1}{s})}$.

A distribuição NGN converge para vários submodelos conhecidos.

- Se $X \sim NGN(\mu, \sigma\sqrt{2}, 2)$, então $X \sim N(\mu, \sigma^2)$;
- Se $X \sim NGN(\mu, \sigma, 1)$, então $X \sim Laplace(\mu, \sigma)$;
- Se $X \sim NGN(\mu, \sigma, s = +\infty)$, então $X \sim Uniforme(-1, 1)$.

A cumulativa não possui inversa analítica, então, por meio do algoritmo de aceitação-rejeição gera-se valores da distribuição NGN utilizando uma distribuição gaussiana como candidata. Conforme verifica-se na Figura 7 os submodelos descritos anteriormente.

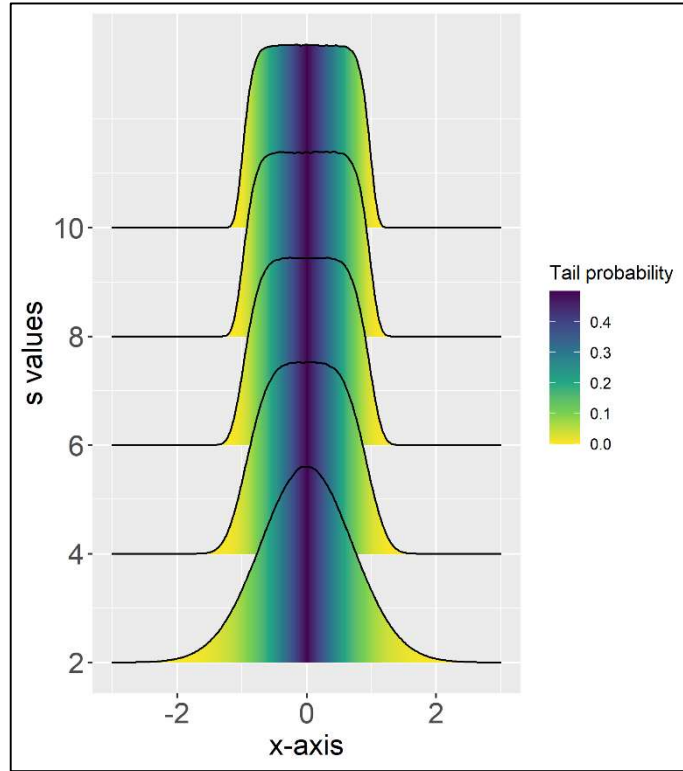


Figura 7: Densidades da distribuição $NGN(0,1,s)$ variando os valores de s .
Fonte: O autor.

Para implementação da função de verossimilhança utilizando a especificação de Poisson:

1º Passo: linearizar $f_Y(y)$

$$l_i = \log v - \log 2 - \log \sigma - \log \Gamma\left(\frac{1}{v}\right) - \left|\frac{y_i - \eta(x_i, \theta)}{\sigma}\right|^v$$

2º Passo: Remover as constantes e fazer $\tau = 1/\sigma^2$

$$l_i = \log v + 0.5 \log \tau - \log \Gamma\left(\frac{1}{v}\right) - |\sqrt{\tau}(y_i - \eta(x_i, \theta))|^v \quad (42)$$

2.2.5 Determinação da Priori empírica

Tratando-se de modelos lineares com erro gaussiano, geralmente os parâmetros tem uma priori conjulgada, o que facilita na computação bayesiana ao utilizar o amostrador de Gibbs, os modelos não lineares com erro Gaussiano tem apenas como Priori conjulgada o parâmetro σ^2 com a distribuição Gama que resulta na DCCP (Distribuição Completa Condiciona a posteriori) Gama inversa. Já os parâmetros de interesse (α, β, γ) geralmente não tem DCCP conhecida,

sendo necessário o algoritmo de Metropolis-Hastings na obtenção das cadeias MCMC.

Para evitar utilizar priori não informativas para os parâmetros $(\alpha, \beta, \gamma, \sigma^2)$ e facilitar na convergência das cadeias MCMC, utilizou-se uma metodologia para atualizar os valores dos hiperparâmetros das distribuições de $(\alpha, \beta, \gamma, \sigma^2)$ utilizando o próprio banco de dados a ser analisado.

Utilizou-se a metodologia de modelos mistos não lineares para obter o ajuste dos modelos Logístico, Gompertz e Von Bertalanffy à todos os Acessos num único algoritmo, como valores iniciais utilizou-se a média de todos os dados longitudinais da Figura 4 e aplicou-se uma técnica de valor inicial utilizando a linearização do modelo Gompertz para convergência do algoritmo.

O segundo passo consiste em utilizar os valores das estimativas frequentistas de $(\alpha, \beta, \gamma, \sigma^2)$ de todos os 8 Acessos para obter suas densidades, e seus respectivos parâmetros de escala e locação, essas informações serão incorporadas nos hiperparâmetros das distribuições a priori de $(\alpha, \beta, \gamma, \sigma^2)$.

Na Figura 8 visualizamos o ajuste de individual de cada Acesso via NLME (*Nonlinear Mixed Effects Model*)

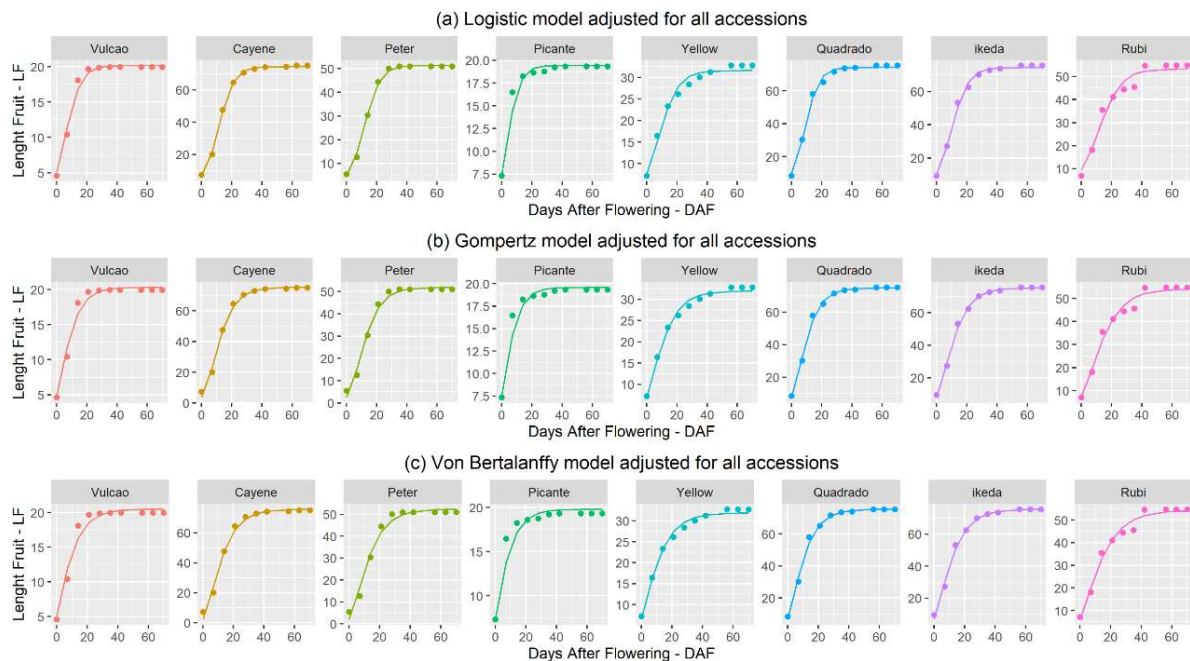


Figura 8: Modelos Logístico, Gompertz e Von Bertalanffy ajustados aos 8 acessos de *capsicum annuum* L. via modelos mistos não lineares.

Fonte: O autor.

A Figura 9 corrobora com a leitura de Pino (2014) que, apesar da primazia da distribuição Normal em diversos procedimentos estatísticos, na prática a normalidade nos dados

não é padrão. Na Figura 9 vemos indícios de outliers (distribuição residual com caudas pesadas) e indícios de assimetria (distribuição residual assimétrica), e ambas (distribuição residual assimétrica com caudas pesadas).

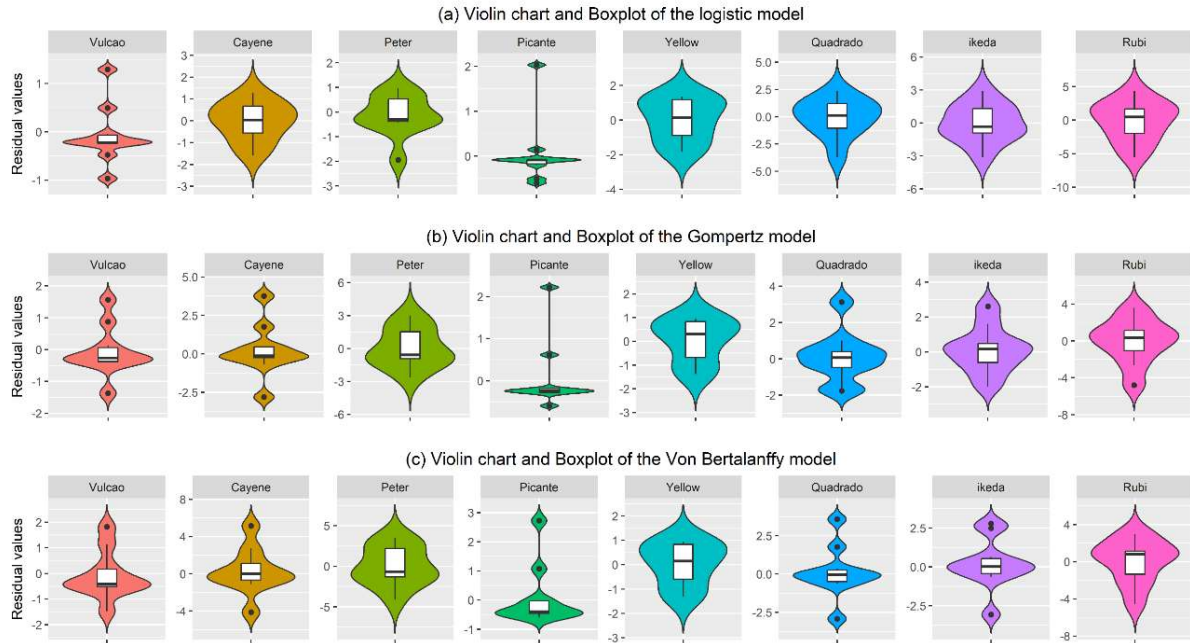


Figura 9: Gráfico de violino dos modelos Logístico, Gompertz e Von Bertalanffy à todos os Acessos de *Capsicum annum* L. representados pela Figura 4.

Fonte: O autor.

2.2.6 Comparação de Modelos e Análise de Convergência

Para análise de convergência das cadeias MCMC optou-se por utilizar dois critérios distintos, o de Heidenberg-Welch e Geweke, os valores de nIter (número de iterações), nBurnin (valores descartados nas iterações iniciais da cadeia) e nThin (valores de salto) são escolhidos para atender os dois critérios simultaneamente.

Após os valores de nThin e nBurnin, para comparar os modelos ajustados pela metodologia bayesiana, a interface Openbugs computa nos resultados os cálculos de DIC (Deviance Information Criterion).

O fator de Bayes ou *Bayes Factor* (BF) é uma medida de plausibilidade, que compartilha semelhanças com a razão de verossimilhanças, e é uma alternativa ao DIC, ao passo que os valores de DIC podem ser afetados por posteriores que sejam bimodais, ou assimétricas, ou aproximações de funções de verossimilhanças. O Fator de Bayes é definido como a razão das posteriores marginais, considerando o modelo de crescimento M_1 e M_2

$$BF = \frac{p(\mathbf{y}, M1)}{p(\mathbf{y}, M2)} \quad (43)$$

BF é a medida que dá a plausabilidade de M1 em detrimento de M2, como $p(\mathbf{y})$ não possui resolução analítica, o valor de cada $p(\mathbf{y})$ é calculado pelo método da média Harmônica:

$$\hat{p}(\mathbf{y}) = \left[\frac{1}{G} \sum_{g=1}^G \frac{1}{f(\mathbf{y}|\boldsymbol{\theta}^{(g)})} \right]^{-1} \quad (44)$$

Em que G é o número de valores gerados da priori, $f(\mathbf{y}|\boldsymbol{\theta}^{(g)})$ é a função de verossimilhança de Y , $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$.

O Openbugs não possui o BF na sua base, para calcular tomando como exemplo uma regressão não linear bayesiana com erro experimental NGN, segue-se uma série de passos:

1º Passo: Calcular a função de verossimilhança de Y para um g -ésimo específico tal qual $\boldsymbol{\theta}^{(g)} = (\alpha^{(g)}, \beta^{(g)}, \gamma^{(g)}, \sigma^{(g)}, v^{(g)})$, em que cada valor de $\boldsymbol{\theta}^{(g)}$ é proveniente das cadeias MCMC após a convergência.

$$\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(g)}) = \frac{v^{(g)} \sqrt{(\tau^{(g)})}}{2\Gamma\left(\frac{1}{v^{(g)}}\right)} \exp \left\{ - \sum_{i=1}^n \left| \left(y_i - \eta(x_i, \alpha^{(g)}, \beta^{(g)}, \gamma^{(g)}) \right) \sqrt{(\tau^{(g)})} \right|^{v^{(g)}} \right\}$$

então $\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(g)}) \in \mathbb{R}$, vai resultar em um escalar Real para cada $\boldsymbol{\theta}^{(g)}$.

2º Passo: Obter uma sequência de valores de cada $\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(g)})$ dessa vez para todos os valores de $g = 1$ até G .

$$\{\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(g)})\}_{g=1}^G = \{\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(1)}), \mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(2)}), \dots, \mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(G)})\}$$

3º Passo: Calcular a média harmônica dos valores da sequência $\{\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}^{(g)})\}_{g=1}^G$ conforme a equação 44.

Após o cálculo da média harmônica de $\hat{p}(\mathbf{y}, M_1)$ e $\hat{p}(\mathbf{y}, M_2)$. Valores logarítmicos de base 10 de BF entre 0 e $\frac{1}{2}$ revelam evidências inconclusivas de plausabilidade; de $\frac{1}{2}$ até 1 revelam evidência moderada em favor de M_1 ; de 1 a 2 revelam evidências fortes em favor de M_1 ; valores maiores que 2 revelam evidências muito fortes em favor de M_1 .

2.3 RESULTADOS E DISCUSSÕES

Após os procedimentos da seção x, os resultados podem ser encontrados na Tabela 1. A Figura 10 representa graficamente os resultados da Tabela 2.

Com respeito as estimativas de α

A média geral dos parâmetro α considerando os 3 modelos de crescimento e os 5 erros experimentais, resultou em 74.59. Não poderia esperar diferenças nas estimativas desse

parâmetros, pois grandes diferenças afetariam a preditividade do modelo.

Não houve diferença nas estimativas dos 3 modelos de crescimento nos 5 erros experimentais. Ainda que na Figura 10, o α dos modelos Gompertz e Von Bertalanffy variam conforme o erro experimental, entretanto as diferenças são inferiores a ordem de 0.5 que corresponde a menos de 1% de variação.

Com respeito as estimativas de β

Os modelos Logístico, Gompertz e Von Bertalanffy apresentaram diferenças substanciais quanto a estimativa da abscissa do ponto de inflexão, O modelo Logístico estimou que a inflexão da curva ocorre de 11 a 12 dias (média de 11.35); o modelo Gompertz estimou a inflexão em torno de 8 a 9 dias (média de 8.37); o modelo Von Bertalanffy estimou a inflexão em torno de 6 a 7 dias (média de 6.96). Algumas diferenças entre os modelos de crescimento chegam a ser da ordem de 1 semana.

Com respeito as estimativas de γ

Os diferentes modelos de crescimento apresentaram diferenças substanciais quanto a estimativa de γ , em média o modelo Logístico estimou em 0.21; em média, o modelo Gompertz estimou em 0.148; em média 0.134.

Na figura 7, visualiza-se que os diferentes erros experimentais não foram capazes de alterar substancialmente as estimativas de γ .

As prioris mais informativas dos modelos foram do parâmetro γ , o que refletiu na precisão também alta das estimativas de γ .

Com respeito as estimativas de σ^2

O modelo Logístico em média, apresentou a menor variabilidade (ou a maior precisão $\tau = 1/\sigma^2$) em torno de 1.49, seguido do modelo Gompertz com média em torno de 2.21 e por último o modelo Von Bertalanffy com média em torno de 3.13. Essa relação ordinal pode ser visualizada por meio da Figura 10 onde não há cruzamento das retas quando varia-se o erro experimental.

Comparando os erros experimentais verifica-se uma diminuição de σ^2 quando se utiliza os erros Logístico, Laplace, Student t e NormalG, com uma única exceção do modelo Logístico com erro NormalG ($\sigma^2 = 2.53$) que apresentou variancia maior que o Logístico com erro Normal ($\sigma^2 = 1.5$).

No geral as distribuições com caudas mais pesadas que a Normal aumentaram a precisão dos modelos de crescimento.

Com respeito aos valores de DIC

Utilizando-se de a mediana, o modelo Logístico apresentou uma mediana em torno de 34.13, seguido do modelo Gompert que apresentou uma mediana em torno 42.45 e por último o modelo Von Bertalanffy com uma mediana em torno de 45.91.

A melhor qualidade de ajuste geral se deu com o modelo Logístico conforme pode ser visualizado na Figura 10, com uma exceção no erro Student t, em que a melhor qualidade de ajuste se deu com o modelo Von Bertalanffy ($DIC = -21.27$).

Na Figura 10 a redução de DIC pode ser verificada quando se compara o erro experimental Normal com relação aos outros erros simétricos. A redução de DIC mais substancial se dá com erro NormalG com DIC's 11.8, 28.33 e 32.21 em relação a distribuição Normal com DIC's 33.32, 46.34 e 51.98, reduções superiores a 5 como recomenda Firat (2016).

Com respeito aos valores de v

Na distribuição t de Student, v que é um parâmetro de controle de caudas, o modelo Logístico apresentou uma aproximação para Normal ($v = 28.01$), enquanto os modelos Gompertz ($v = 15.53$) e Von Bertalanffy ($v = 7.43$) apresentaram evidências de caudas pesadas residuais e justifica o uso de regressão robusta.

Na distribuição NormalG, v é um parâmetro de forma, o modelo Logístico apresentou uma aproximação à distribuição Uniforme ($v = 26.63$), enquanto o modelo Gompertz ($v = 1.17$) apresentou um meio termo entre a distribuição Laplace e Normal, e o modelo Von Bertalanffy ($v = 1$) convergência para o erro Laplace.

Tabela 2: Médias a posteriori dos parâmetros ($\alpha, \beta, \gamma, \sigma^2, v$) conjuntamente com os DIC's de cada ajuste.

Erros	Modelos	α	β	γ	σ^2	v	DIC
Normal	Logístico	74,16	11,37	0,21	1,5	0	33,32
	Gompertz	74,9	8,32	0,14	3,5	0	46,34
	Von Bertalanffy	75,03	6,99	0,13	5,96	0	51,98
Logístico	Logístico	74,16	11,35	0,21	0,86	0	36,11
	Gompertz	74,76	8,33	0,15	1,27	0	46,25
	Von Bertalanffy	74,91	6,94	0,13	1,73	0	51,35
Laplace	Logístico	74,16	11,32	0,2	1,11	0	35,31
	Gompertz	74,64	8,44	0,15	1,45	0	41,25
	Von Bertalanffy	74,8	7,18	0,14	1,89	0	45,91
Student t	Logístico	74,15	11,36	0,21	1,48	28,01	34,13
	Gompertz	74,74	8,4	0,15	2,37	15,53	42,45
	Von Bertalanffy	74,85	6,49	0,13	3,13	7,43	-21,27

Normal generalizada	Logístico	74,13	11,36	0,21	2,53	26,63	11,8
	Gompertz	74,69	8,4	0,15	2,49	1,17	28,33
	Von	74,77	7,22	0,14	2,96	1	32,21
	Bertalanffy						

Fonte: O autor

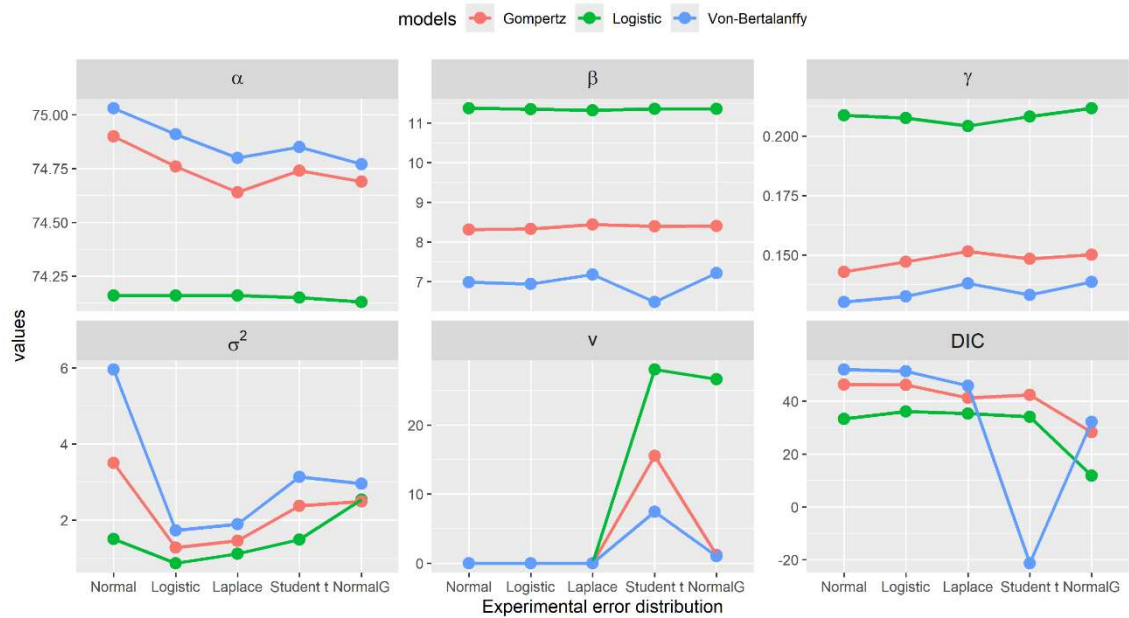


Figura 10: Gráficos representando a média a posteriori dos parâmetros ($\alpha, \beta, \gamma, \sigma^2, v$) dos modelos Logístico, Gompertz e Von Bertalanffy utilizando 5 erros experimentais.

Fonte: O autor.

A Figura 11 mostra as densidades a posteriori dos parâmetros ($\alpha, \beta, \gamma, \sigma^2$) - que é útil para fornecer prioris informativas para futuros estudos de *capsicum annum L* - é útil para identificar falhas de convergência ao verificar-se a existência de bimodalidade, e para verificar se houve aprendizado bayesiano comparando a priori com a posteriori.

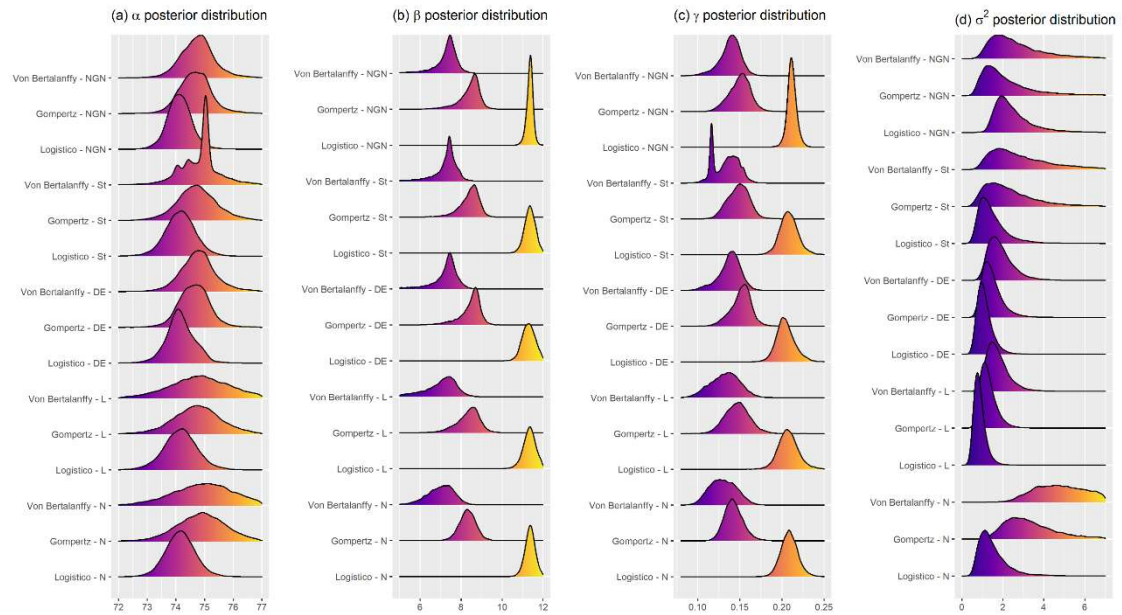


Figura 11: Densidades das distribuições a posteriori dos parâmetros (α , β , γ , σ^2) dos modelos Logístico Gompertz e Von Bertalanffy considerando todos os erros experimentais.

Fonte: O autor

Com respeito aos cálculos do Fator de Bayes, a Tabela 3 mostra todas as comparações dois a dois dos modelos Logístico, Gompertz e Von Bertalanffy dentro de cada função de verossimilhança dos erros experimentais estudados neste trabalho, com seus resultados em escala logarítmica de base 10. Os resultados (Logístico vs. Gompertz), por exemplo, são em favor de Logístico em detrimento de Gompertz, e assim por diante, casos contrários explicitados.

Tabela 3: Logarítmo de base 10 do Fator de Bayes resultante da comparação 2 a 2 dos modelos Logístico, Gompertz e Von Bertalanffy dentro de cada erro experimental.

<i>Erro experimental</i>	<i>Logístico vs. Gompertz</i>	<i>Logístico vs. Von Bertalanffy</i>	<i>Gompertz vs. Von Bertalanffy</i>
<i>Normal</i>	0.55 (em favor de Gompertz)	0.96	1.52
<i>Logístico</i>	1.12	2.17	1.05
<i>Laplace</i>	1.64	2.58	0.94
<i>Student t</i>	1.70	2.77	1.07
<i>Normal Generalizada</i>	-	-	0.94

Na Tabela 3 constata-se que o modelo Logístico apresentou fortes evidências de plausibilidade (1.12, 1.64 e 1.70) em detrimento do modelo Gompertz em todos os cenários,

com exceção do erro Normal que mostrou evidência positiva em favor de Gompertz (0.55).

O modelo Logístico (2.17, 2.58 e 2.77) mostrou evidências muito forte de plausibilidade em detrimento do modelo Von Bertalanffy em todos os cenários, com exceção do erro Normal que mostrou evidência positiva em favor de Gompertz (0.96).

O modelo Gompertz mostrou-se plausibilidades variadas em detrimento do modelo Von Bertalanffy, variando entre positiva (0.94 e 0.94) e forte (1.52, 1.05 e 1.07).

O modelo Logístico com a função de verossimilhança Normal generalizada não pode ser computada por conta de valores muito próximos de zero o que prejudica o cálculo.

Deve-se observar que se apenas o erro Normal fosse utilizado, o modelo Gompertz seria o mais plausível, com o uso das distribuições com caudas pesadas, o modelo Logístico obtém primazia total em termos de plausibilidade em detrimento de Gompertz e Von Bertalanffy.

2.4 CONCLUSÕES

Os erros experimentais além da Normal, aumentaram a precisão dos modelos e melhoraram a qualidade de ajuste. O modelo Logístico com o ponto de inflexão em torno de 10 a 11 dias após o florescimento, apresentou a maior precisão em todos os cenários e mostrou-se o mais plausível pelo Fator de Bayes, com o uso do erro Normal generalizado de Nadarajah melhorando a qualidade de ajuste.

2.5 REFERÊNCIAS

CARLIN, B. P.; LOUIS, T. A. **Bayesian Methods for Data Analysis**. 2 ed. United States: CRC Press, 246p, 2008.

CORDEIRO, G. M. *et al.* **Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models**. São Paulo: IME-USP, 2000.

CORDEIRO, G. M., DEMÉTRIO, C. G. B. **Modelos lineares generalizados e extensões**, 1. ed. Recife, PE: [s.n.], 2013.

DE LA CRUZ, R., BRANCO, M. Bayesian analysis for nonlinear regression model under skewed errors, with application in growth curves. **Biometrical Journal**, Weinheim, v.51, p.588-609, 2009.

FIRAT, MZ. *et al.* Bayesian analysis for the comparison of Nonlinear Regression Model Parameters: an Application to the Growth of Japanese Quail. **Revista Brasileira de Ciência Avícola**, v.18, p.19-26, 2016.

GUJARATI, D. N., PORTER, D.C. **Econometria básica**. AMGH, 2012.

LEMONTE, A. J., PATRIOTA, A. G. Multivariate elliptical models with general

parameterization. **Statistical Methodology**, v.8, p.389-400, 2011.

LOUIS, T. A., CARLIN, B. P. **Bayesian Methods for Data Analysis**. 2 ed. United States: CRC Press, 2008.

MACEDO, L. R. et al. Bayesian inference for the fitting of dry matter accumulation curves in garlic plants. **Pesquisa Agropecuária Brasileira**, v.52, p.572-581, 2017.

MOURA, G. L. S., GUZZO, F., CECON, P. R., MARTINS FILHO, S., CARNEIRO, A. P. S., NASCIMENTO, M. Comparison of Normal, Logistic, Laplace, and Student's t distributions for experimental error in the Bayesian description of dry matter accumulation in *Allium sativum*. **Ciência Rural**, v.54 7 ed, 2024.

NADARAJAH, S. A, Generalized normal distribution. **Journal of Applied Statistics**, 32(7), p.685–694, 2005.

OLIVEIRA, A. C. R.; CECON, P. R.; PUIATTI, G. A.; GUIMARÃES, M. E. S.; CRUZ, C. D.; FINGER, F. L.; NASCIMENTO, M.; PUIATTI, M.; LACERDA, M. S. Modelos não lineares baseados em quantis no ajuste de curvas de crescimento de genótipos de pimenta. **Rev. Bras. Biom. Lavras**, v.39, n.3, p.447-459, 2021.

PINO, F.A. A questão da não normalidade: Uma revisão. **Rev. de Economia Agrícola**, v.61, p.17-33, 2014.

ROSSI. R.M., SANTOS. L.A. Modelagem Bayesiana para curvas de crescimentos de codornas assumindo assimetria nos erros. **Semina: Ciências Agrárias**, v.35, p.1637-1647, 2014.

RULI, E., VENTURA, L. Can Bayesian, confidence distribution and frequentist inference agree?. **Stat Methods Appl**. v.30, p.359–373, 2021.

SOUZA, G.S. **Introdução aos modelos de Regressão linear e não linear**. Brasília. Embrapa, 1998.

TEIXEIRA, F.R.F., CECON, P.R., SUELA, M.M., NASCIMENTO, M. Nonlinear Mixed-Effect Models to Describe Growth Curves of Pepper Fruits in Eight Cultivars Including Group Effects. **Agronomy**, v.12, 2023.



ARTICLE

Bayesian implementation of Skew-Normal distributions for experimental error in the description of pepper growth

 George Lucas Santana de Moura^{*1},  Thaynara Aparecida de Souza Neto¹,  Paulo Roberto Cecon¹,  Sebastião Martins Filho¹ and  Antônio Policarpo Souza Carneiro¹.

¹Department of Statistics, Federal University of Viçosa, Viçosa-MG, Brazil

*Corresponding author. Email: gelucas.moura.2016@gmail.com

(Received: March 10, 2024; Revised: August 23, 2024; Accepted: September 30, 2024; Published: April 3, 2025)

Abstract

The purpose of this work is to implement and compare three types of Skew-Normal distributions together with the Normal submodel for the experimental error through a Bayesian nonlinear modeling of pepper phenotype growth. The posterior means of the parameters (α, β, γ) were shown to be invariant to the experimental error. The Sahu Skew-Normal error showed evidence of null asymmetry for all growth models. In general, all Skew-Normal distributions presented the highest accuracy ($1/\sigma^2$) in relation to the Normal error. The problematic points stand out for the computational cost in the millions of MCMC iterations and the limitations of the BUGS language. The Skew-Normal distribution of Azzalini and Fernández & Steel modeled the asymmetry of the data and provided the best goodness of fit (DIC) and the best precision ($1/\sigma^2$) for the Gompertz and Von Bertalanffy model in relation to the Normal error.

Keywords: Bayesian Inference; Longitudinal data; Skew-Normal.

1. Introduction

Peppers of the *Capsicum* genus have stood out in Brazil due to their genetic diversity, widespread national consumption, and prominence in Brazilian olive growing (Nunes Ávila & Barbosa, 2019).

In plant growth, some of the phenotypic characteristics follow an exponential growth until the maturation rate is reached. From this inflection point, the plant growth declines until it reaches an asymptotic value. The pattern of this type of growth is called sigmoid. Sigmoid models are non-linear, and the Normal distribution is generally used to adjust these models (Souza, 1998). Asymmetric experimental errors can also be used in plant height modeling (Mangueira *et al.*, 2016; Guedes *et al.*, 2014).

Regarding the issue of normality, Pino (2014) asserts that although the Normal distribution is the default in several statistical procedures, in practice, it is more common for observations to be approximately normal. Some factors contribute to the extrapolation of the normality assumption,

CAPÍTULO 2: Abordagem bayesiana robusta de modelos não lineares aplicados ao crescimento de capisicum: Um estudo das distribuições assimétricas

Resumo: O objetivo deste trabalho consiste em ajustar o modelo logístico ao comprimento do fruto de *Capsicum Annuum L.* por meio de uma abordagem bayesiana considerando 14 distribuições para o erro experimental, 10 assimétricas, e 4 simétricas. Um levantamento das principais distribuições de probabilidade assimétricas revela que há muitas propostas de distribuições assimétricas, mas poucos os trabalhos que utilizam estas distribuições para modelar a assimetria residual em modelos de regressão não linear. Distribuições mais utilizadas como a Skew-Normal de Azzalini, geralmente são obtidas por transformações de distribuições simétricas. Algumas transformações geram modelos mais complexos do que outras. Os resultados mostram que os erros experimentais assimétricos mudam substancialmente os resultados dos parâmetros, aumentam a precisão do modelo e aumentam a qualidade de ajuste, em relação aos seus submodelos simétricos.

Palavras - Chave: Regressão não linear, Bayes empírico, Regressão robusta, distribuição Skew-Normal, MCMC.

3. INTRODUÇÃO

Pimentas do gênero *capsicum* têm se destacado no Brasil pela diversidade genética, amplo consumo nacional e destaque na oleicultura brasileira (Nunes Ávila & Barbosa, 2019).

No crescimento de plantas, algumas de suas características fenotípicas seguem um crescimento exponencial até que a taxa de maturação seja atingida, a partir desse ponto de inflexão o crescimento da planta decai até atingir um valor assintótico, o padrão desse tipo de crescimento é denominado sigmoide. Modelos sigmoides são não lineares e geralmente utiliza-se a distribuição Normal para o erro experimental no ajuste destes modelos (Souza, 1998). Também os erros experimentais assimétricos podem ser utilizados na modelagem de altura de plantas (Mangueira et al, 2016; Guedes et al, 2014).

No que tange a questão da normalidade, Pino (2014) assevera que, apesar da distribuição Normal ser *default* em diversos procedimentos estatísticos, na prática, é mais comum as observações serem aproximadamente normais. Alguns fatores contribuem para extrapolação da pressuposição de normalidade como, a presença de outliers, assimetria dos

dados ou suporte apenas em $\mathbb{R}^+$ nas variáveis. Na presença de outliers, em termos de regressão, pode-se utilizar a Regressão robusta, Gelman et al. (2013) explana que na regressão bayesiana, a distribuição t de Student por possuir caudas pesadas, é mais eficiente em relação a Normal. Na questão da assimetria, Eloy et. al. (2020) explana que várias são as propostas de novas distribuições que atendam a assimetria e a curtose dos dados como a *Skew-Normal* e *Skew-Student t*.

A distribuição Skew-Normal mais conhecida é a obtida por Azzalini, por meio de uma transformação de Azzalini, a metodologia adiciona um novo parâmetro de assimetria (λ) a qualquer distribuição da família simétrica, existem outras versões como a distribuição Skew-Normal de Sahu e a distribuição Skew-Normal de Fernández & Steel. Como assevera Ghaderinezhad et. al (2020), a estimação frequentista da distribuição Skew-Normal passa por uma série de dificuldades como pontos de singularidade, função de log-verossimilhança monotônica para λ , expressões complexas da estimação de λ por meio do método dos momentos e matriz de informação de Fisher singular (ROSSI & SANTOS, 2014). Existem distribuições cuja transformação para modelos *Skew-Symmetric*, não possuem expressões matemáticas analíticas para momentos, expectâncias e variâncias como a distribuição de *Azzalini Skew-Logistic* estudada por Nadarajah (2009). Neste contexto, Ghaderinezhad et. al (2020) recomenda a metodologia bayesiana para contornar os problemas da estimação frequentista da distribuição Skew Normal.

Em modelos lineares como o de Gauss-Markov Normal, os testes de significância, ANOVA e intervalos de confiança são exatos, apesar de dependerem de uma série de condições de estimabilidade (LUNA & OLINDA, 2012). Ao considerar uma distribuição do erro além da Normal, geralmente os testes de significância dependem da normalidade assintótica dos estimadores de máxima Verossimilhança (MV). No caso de Modelos não lineares Normal, a distribuição dos estimadores de MV e Gauss-Newton dependem também da teoria assintótica (CORDEIRO & DEMÉTRIO, 2013), e a ANOVA nestes casos não é válida pois a decomposição da soma de quadrados totais em resíduo e regressão não é válida (GUJARATI & PORTER, 2012). O ganho que se tem na metodologia bayesiana de modelos não lineares consiste na validade dos testes de significância e dos intervalos de credibilidade, sem se preocupar com o tamanho da amostra (MARTINS FILHO et al., 2008) ou com a existência de um máximo global da função de verossimilhança.

Neste trabalho, 10 versões de distribuições assimétricas serão implementadas, a

distribuição *Azzanlini's Skew-Normal* (1985) que denotaremos de ASN; a distribuição *Sahu's Skew-Normal* (2003) que denotaremos de SSN; a distribuição *Fernandéz & Steel Skew-Normal* (1998) que denotaremos de FSSN; a distribuição Skew-Laplace (*double exponential*) denotaremos de SDE; a distribuição *Sastry and Bhati Skew-Logistic* (2016) denotaremos de SBSL; distribuição Jones & Faddy's Skew Student-t (2003) denotaremos de JFSSt; a distribuição *Epsilon-Skewed Student t* (Gomez et al, 2007) denotaremos de ESSt; a distribuição *Fernandez Steel's Skew-Student t* (1998) denotaremos de FSSSt; a distribuição *Skew-laplace (Skew-double-exponential) Normal* denotaremos de SDEN; a distribuição *sinh-arcsinh Skew-Student t* denotaremos de SASSt.

O objetivo da implementação no erro experimental das distribuições assimétricas deste trabalho é compará-las entre si em conjunto com os seus respectivos submodelos simétricos, para a descrição bayesiana do crescimento fenotípico de Pimenta utilizando o modelo Logístico.

3.2 MATERIAL E MÉTODOS

O banco de dados a ser trabalhado é composto de 8 acessos de *Capsicum annuum* L. sendo estes, 1 - pimenta vulcão; 2 - Pimenta Cayene; 3 - Pimenta Peter; 4 – Pimenta picante; 5 - Pimenta jamaica Yellow; 6 - Pimentão quadrado; 7 - Pimentão Ikeda; 8 – Pimentão Rubi gigante. O experimento foi conduzido no departamento de Fitotecnia da Universidade Federal de Viçosa (UFV), localizada no município de Viçosa, Minas Gerais, Brasil. (20° 45 Sul e 42° 51 Oeste, altitude média de 650 m). Conforme expressa na Figura 12, as características fenotípicas obtidas do experimento são o comprimento do fruto - DAF (variável explanatória) em função dos dias após o florescimento – DAF (variável independente).

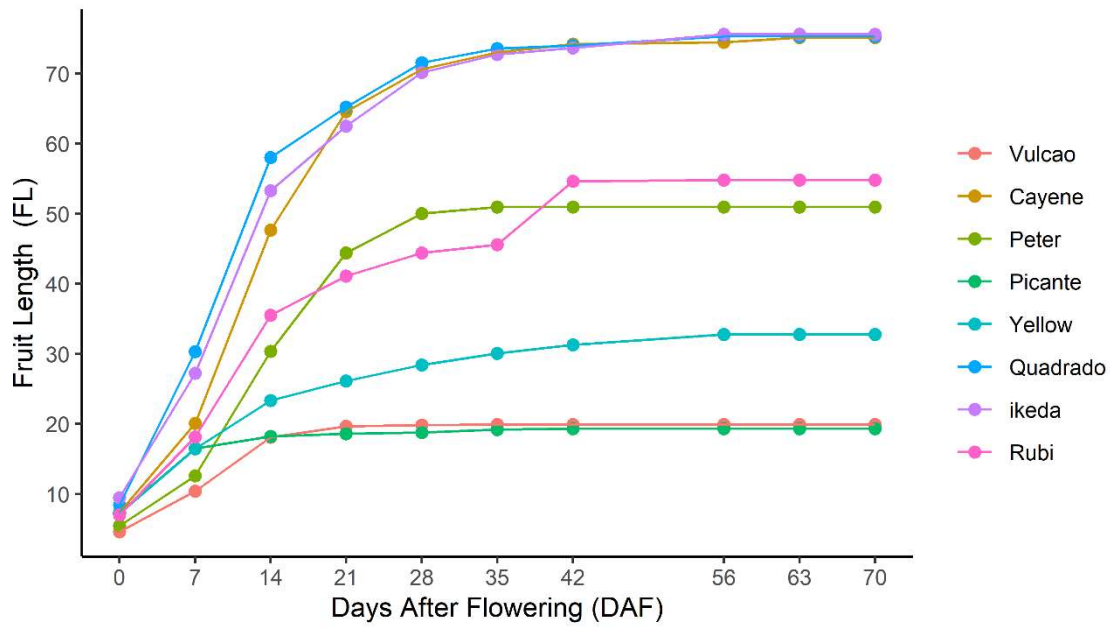


Figura 12: *Scatterplots* dos Acessos de *Capsicum Annuum L.* mostrando o comprimento do fruto em função dos dias após o florescimento.

Fonte: O autor.

Os modelos de crescimento adotados neste trabalho para o ajuste bayesiano dos Acessos de *Capsicum Annuum L.* representados pela Figura 12, são de reparametrizações especiais que permitem uma interpretabilidade biológica para o parâmetro β . Conforme pode ser verificado na equação 45.

$$\eta(\theta) = \frac{\alpha}{1 + e^{\gamma(\beta-x)}} \quad (45)$$

3.2.1 Regressão não linear com erros assimétricos

No modelo clássico de regressão não linear univariada, temos uma função resposta $\eta(x_i, \theta)$ não linear nos parâmetros do vetor θ , um vetor de variável independente x , um vetor y de variáveis explanatórias, e o erro experimental $e_i \sim N(0, \sigma^2)$ que é normalmente distribuído. O modelo de regressão não linear, também pode ser definido conforme a equação 46.

$$y_i = \eta(x_i, \theta) + \sigma e_i \quad (46)$$

Com distribuição amostral $Y_i \sim N(\eta(x_i, \theta), \sigma^2)$, sendo o erro e_i uma variável aleatória, então $\eta(x_i, \theta) + \sigma e_i$ seria o equivalente a uma transformação de locação-escala desta variável

para Y_i , em que $e_i \sim N(0,1)$. Nos modelos não lineares normais têm-se que $E(Y_i) = \eta(x_i, \theta)$ pois a distribuição de Y_i é simétrica e o parâmetro de locação μ é igual a $\eta(x_i, \theta)$.

Quando se trata de flexibilizar o pressuposto da normalidade do erro experimental gaussiano para o erro assimétrico, deve considerar-se que nas distribuições assimétricas, a moda é diferente da média que difere da mediana, ou seja, a esperança não igual ao parâmetro de locação $E(Y_i) \neq \mu$, quanto mais assimétrica é a distribuição, maior a distância de deslocamento entre parâmetro de locação e a esperança.

Tomando como exemplo uma regressão com erros FSSN (ver seção 3.2.5), tal que $e_i \sim FSSN(0, \sigma^2, \lambda)$, dessa forma teríamos a distribuição de $Y_i \sim FSSN(\eta(x_i, \theta), \sigma^2, \lambda)$, entretanto verifica-se que a distribuição de cada Y_i não está centrada em $\eta(x_i, \theta)$ pois $E(Y_i) = \eta(x_i, \theta) + \sigma \sqrt{\frac{2}{\pi} \frac{\lambda^2 - 1}{\lambda}}$, quanto mais elevado λ , maior a distancia entre $E(Y_i)$ e $\eta(x_i, \theta)$, na ausência de assimetria ($\lambda = 1$) a distorção é anulada.

Para corrigir essa distorção fazamos $E(Y_i) = \eta(x_i, \theta)$, então $\mu + \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} = \eta(x_i, \theta)$, o novo valor de μ será:

$$\mu = \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \leftrightarrow e_i \sim SN3 \left(-\sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}, \sigma^2, \lambda \right)$$

Conforme verifica-se na Figura 13, utilizando agora o caso ASN (ver seção 3.2.3), quanto maior a assimetria, mais distante o parâmetro de locação se desloca do centro de massa da distribuição, assim μ passa a se localizar nas caudas, já quando a distribuição ASN passa pela correção, então $E(Y) = \mu$, e μ passa a se localizar próximo do centro de massa da distribuição.

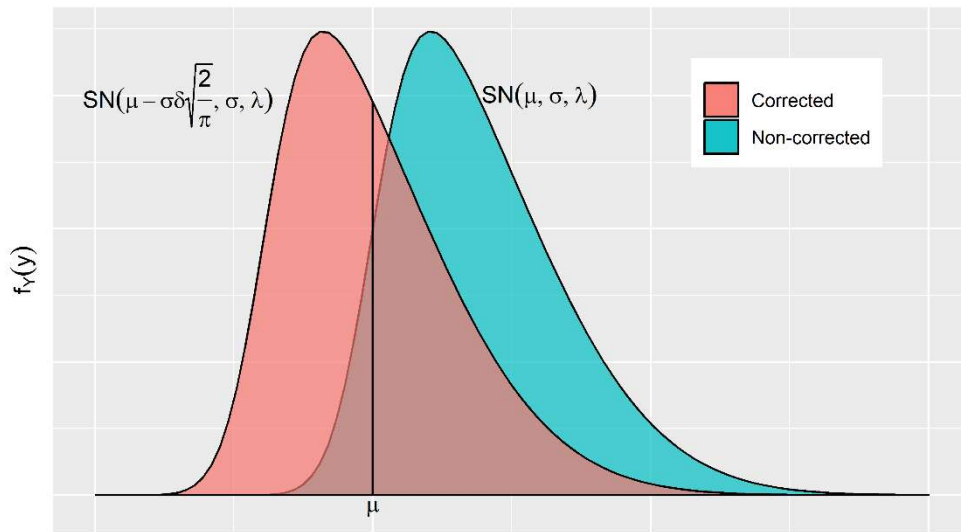


Figura 13: Gráfico das densidades de ASN mostrando o deslocamento relativo do parâmetro de localização quando se utiliza os dois casos $ASN(\mu, \sigma, \lambda)$ – sem correção e $ASN(\mu - \sigma\delta\sqrt{2/\pi}, \sigma, \lambda)$ – com correção.

Fonte: O autor.

A maioria das distribuições assimétricas não possuem expressões matematicamente tratáveis para esperança, por estas dificuldades, apenas 6 distribuições assimétricas analisadas neste manuscrito têm correção para média, as demais foram implementadas livres de correção. Cancho et al. (2008) também se utiliza de tais correções na sua implementação bayesiana de SSN utilizando $e_i \sim SSN\left(-\lambda\sqrt{\frac{2}{\pi}}, \sigma, \lambda\right)$.

Na Figura 14 visualiza-se a distribuição amostral de cada variável resposta Y_i num *scatterplot* teórico.

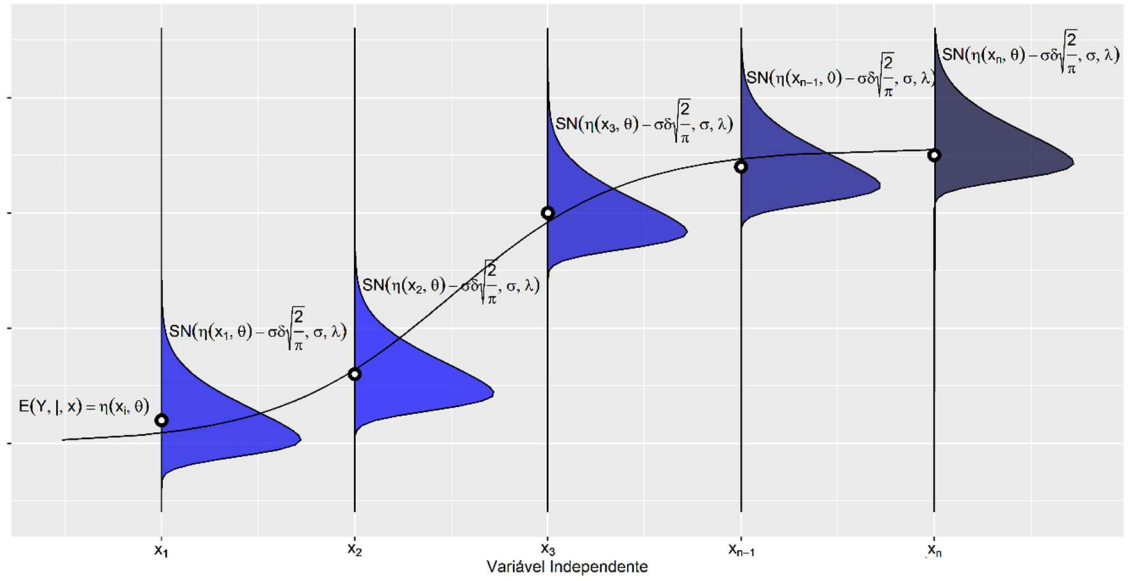


Figura 14: Regressão não linear assimétrica centralizada em cada $E(Y_i)$ utilizando a distribuição ASN corrigida.

Fonte: O autor

3.2.2 Especificação da função de verossimilhança

O Openbugs possui um escopo restrito de distribuições de probabilidade. As demais distribuições, tanto para priori quanto para função de verossimilhança, que não estão no programa serão implementadas pela função *dloglik()* que implementa o truque de Poisson. A função de verossimilhança das distribuições assimétricas é representada por:

$$\mathcal{L}(\theta|y) = \prod_{i=1}^n P(Y|y=0); Y \sim Pois(\mu = K - l_i)$$

as constantes que l_i pode possuir, são removidas pois são absorvidas por K .

3.2.3 Distribuição Skew-Normal de Azzalini

A distribuição Normal assimétrica foi introduzida por Azzalini em 1985. A nova distribuição de probabilidade é obtida pela transformação $f(x) = 2g(x)G(\lambda x)$ em que, $g(\cdot)$ e $G(\cdot)$ são, respectivamente, a fdp (função de densidade de probabilidade) e fda (função de distribuição acumulada) de uma variável aleatória simétrica em $\mathbb{R}$ de X . Com a introdução de um novo parâmetro de assimetria λ , e com a transformação de locação-escala $Y = \mu + \sigma X$ em que μ e σ são, respectivamente, o parâmetro de locação e de escala, a distribuição Skew-Normal (SN) é definida conforme a equação 47.

$$f_Y(y) = \frac{2}{\sigma} \phi\left(\frac{y-\mu}{\sigma}\right) \Phi\left(\lambda \frac{y-\mu}{\sigma}\right) \quad (47)$$

em que $\phi(\cdot)$ e $\Phi(\cdot)$ São, respectivamente, a fdp e fda da distribuição Normal.

No caso da distribuição Normal, a média e o desvio-padrão coincidem com o parâmetro de locação e de escala, em noutras distribuições assimétricas, são funções destes parâmetros.

O parâmetro de assimetria (λ) denota a orientação da curva, valores $\lambda > 0$ denotam assimetria à direita, $\lambda < 0$ assimetria à esquerda, $\lambda = 0$ retorna à distribuição normal que é simétrica conforme é verificada na Figura 15.

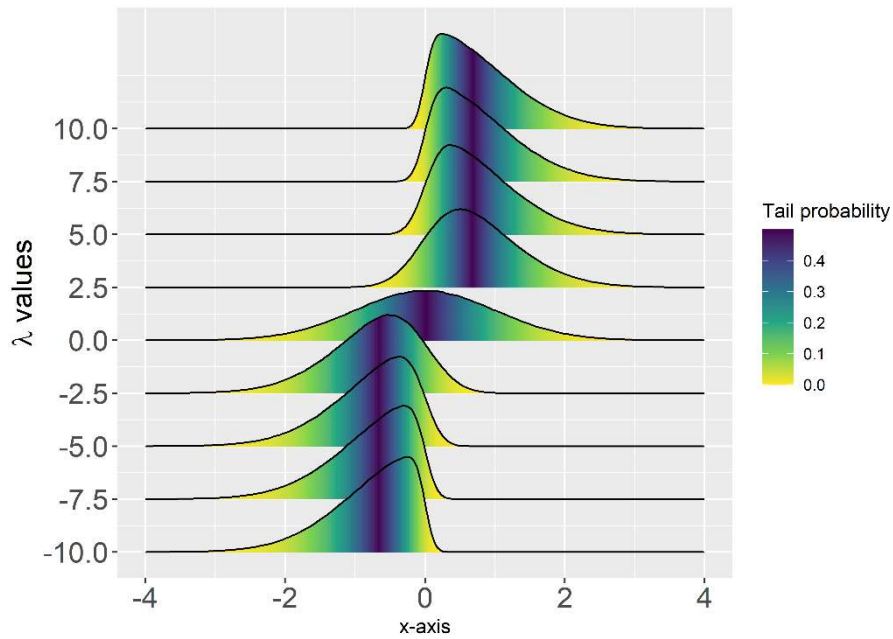


Figura 15: Comportamento da distribuição Skew-Normal de Azzalini variando o parâmetro de assimetria (λ values).

Uma forma de gerar números aleatórios de uma distribuição Skew-Normal por meio de pseudo-realizações de uma Normal é por meio da representação estocástica de Henze (1986). Dados $X_0 \sim N(0,1)$ e $X_1 \sim N(\mu, \sigma^2)$ duas distribuições Normais independentes, considerando um parâmetro λ , então a variável Z :

$$Z = \frac{\lambda}{\sqrt{1+\lambda^2}} |X_0| + \frac{1}{\sqrt{1+\lambda^2}} X_1 \quad (48)$$

converge em distribuição para $ASN(\mu, \sigma^2, \lambda)$, que denota neste trabalho a distribuição Skew-Normal de Azzalini, com base na representação estocástica, constrói-se os gráficos da distribuição ASN - que podem ser construídos também pelo pacote *sn* (versão 2.2.1) - que pode ser visualizada pela Figura 15.

A Expectância e Variância de uma v.a. $Y \sim ASN(\mu, \sigma^2, \lambda)$ são dadas por

respectivamente equações (49) e (50).

$$E(Y) = \mu + \sigma\delta\sqrt{\frac{2}{\pi}} \quad (49)$$

$$V(Y) = \sigma^2 \left(1 - \frac{2\delta^2}{\pi}\right) \quad (50)$$

em que $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

Na obtenção da função de verossimilhança, segue-se uma série de etapas:

1º Passo: Define-se μ conforme os procedimentos da seção 3.2.1 $\mu = \eta(x_i, \theta) - \sigma\delta\sqrt{\frac{2}{\pi}}$

$$f_Y(y) = \frac{2}{\sigma} \phi \left(\frac{y_i - \eta(x_i, \theta) + \sigma\delta\sqrt{\frac{2}{\pi}}}{\sigma} \right) \Phi \left(\lambda \frac{y_i - \eta(x_i, \theta) + \sigma\delta\sqrt{\frac{2}{\pi}}}{\sigma} \right)$$

2º Passo: Obtenha l_i conforme os procedimentos da seção 3.2.2

$$l_i = \log 2 - \log \sigma^2 - 0.5 \log 2\pi - 0.5 \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \delta\sqrt{\frac{2}{\pi}} \right)^2 \\ + \log \Phi \left(\lambda \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \delta\sqrt{\frac{2}{\pi}} \right) \right)$$

3º Passo: Remova as constantes e substitua σ^2 pelo parâmetro de precisão τ tal que $\tau = 1/\sigma^2$

$$l_i = \log \tau - 0.5 \left((y_i - \eta(x_i, \theta))\sqrt{\tau} + \delta\sqrt{\frac{2}{\pi}} \right)^2 + \log \Phi \left(\lambda \left((y_i - \eta(x_i, \theta))\sqrt{\tau} + \delta\sqrt{\frac{2}{\pi}} \right) \right) \quad (51)$$

3.2.4 Distribuição Skew-Normal de Sahu et al. (2003)

No trabalho de Sahu et al. (2003), os autores derivam novas distribuições acrescentando um parâmetro de assimetria às distribuições que pertencem a classe Elíptica. A nova classe multivariada *Skew-Elliptical* é obtida por meio do método do argumento condicionado, com base nos diferentes geradores de densidade $g^{(k)}$, obtêm-se as distribuições Skew-Normal e Skew-t de Sahu. Os autores propõem-se estas distribuições na modelagem da assimetria do resíduo, ao invés de se utilizar metodologias de transformação de dados, e reforçam que a implementação destas são mais fáceis do que as de Fernández & Steel (1998). Algumas vantagens da distribuição Skew-Normal de Sahu consiste na fácil implementação MCMC e invariância à autocorrelação residual (De la cruz & Branco, 2009). Outra vantagem, no contexto da regressão bayesiana, consiste na obtenção de distribuições posteriori conhecidas.

Considerando um vetor aleatório $\mathbf{X}$ k -dimensional, com um vetor de medias $\boldsymbol{\mu}$ e uma matriz de variância-covariância $\boldsymbol{\Sigma}$, positiva definida tal que $\boldsymbol{\mu}, \boldsymbol{\Sigma} \in \mathbb{R}^k$. A densidade de $\mathbf{X}$ e dada genericamente por:

$$f(x, \mu, \Sigma; g^{(k)}) = |\boldsymbol{\Sigma}|^{-1/2} g^{(k)}((x - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (x - \boldsymbol{\mu}))$$

em que $g^{(k)}(u) = \frac{\Gamma(\frac{k}{2})}{\pi^{k/2}} \times \frac{g(\mu; k)}{\int_0^\infty r^{k/2-1} g(r; k) dr}$, supondo que a integral exista, $g^{(k)}$ é a densidade geradora de $\mathbf{X}$, e $g(\mu; k)$ é o kernel de $\mathbf{X}$. As densidades que podem ser escritas pela função $f(x, \mu, \Sigma; g^{(k)})$ são chamadas de distribuições elípticas com notação $\mathbf{X} \sim El(\mu, \sigma, g^{(k)})$.

Por meio das formas genéricas da distribuição *Skew-Elliptical*, utilizando $g(u, m) = e^{-u/2}$, pode-se derivar que $g^{(m)}(u) = (2\pi)^{-m/2} e^{-u/2}$. Resolvendo a integral:

$$\int_0^{+\infty} r^{m/2-1} e^{-a/2-r/2} dr$$

utilizando a função gamma,

$$\int_0^{+\infty} r^{m/2-1} e^{-a/2-r/2} dr = 2^{m/2} e^{-a/2} \int_0^{+\infty} \phi^{m/2-1} e^{-\phi} d\phi = 2^{m/2} e^{-a/2} \Gamma\left(\frac{m}{2}\right)$$

assim obtemos $g^{(m)}(u)$,

$$g_a^{(m)}(u) = \frac{\Gamma(\frac{m}{2})}{\pi^{m/2}} \times \frac{e^{-a/2} e^{-u/2}}{2^{m/2} e^{-a/2} \Gamma(\frac{m}{2})}$$

$$g_a^{(m)}(u) = (2\pi)^{-m/2} e^{-u/2}$$

Com base nessas informações, deriva-se uma *Multivariate Skew-Normal distribution*, cuja densidade é dada por:

$$f(y, \mu, \Sigma, D) = 2^m |\Sigma + D^2|^{-1/2} \phi_m((\Sigma + D^2)^{-1/2} (y - \mu)) P(V > 0)$$

em que $V \sim N_m(D(\Sigma + D^2)^{-1} (y - \mu), I - D(\Sigma + D^2)^{-1} D)$ e ϕ_m é uma distribuição normal multivariada. Ao considerar o caso monovariado em que $m = 1$ obtêm-se a distribuição Skew-Normal de Sahu et al. (2003).

$$f_Y(y) = 2\phi(\mu, \sigma^2 + \lambda^2) \Phi\left(\frac{\lambda}{\sigma} \times \frac{y - \mu}{\sqrt{\sigma^2 + \lambda^2}}\right) \quad (52)$$

Se Y é uma v.a. que segue a distribuição Skew-Normal de Sahu et al. (2003), então $Y \sim SSN(\mu, \sigma^2, \lambda)$. Em termos de assimetria assim como ASN, o parâmetro de assimetria (λ) denota a orientação da curva, valores $\lambda > 0$ denotam assimetria à direita, $\lambda < 0$ assimetria à

esquerda, $\lambda = 0$ assimetria nula e convergência para o modelo Normal. O comportamento de SSN pode ser visualizada por meio da Figura 16.

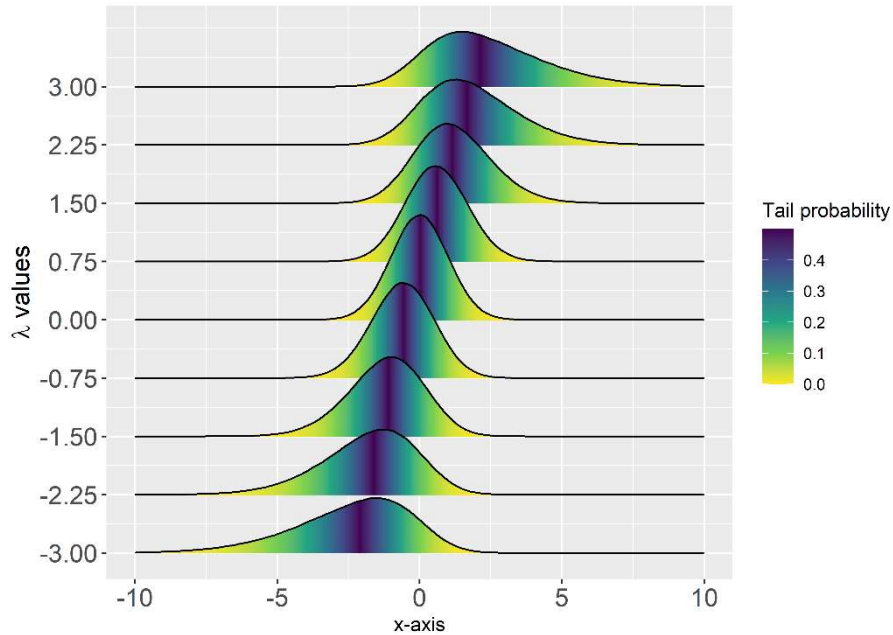


Figura 16: Comportamento da distribuição Skew-Normal de Sahu padronizada variando o parâmetro de assimetria (λ values).

Fonte: O autor

Uma forma de representação estocástica da distribuição Skew-Normal de Sahu é, semelhante a distribuição de Azzalini:

$$Z = \lambda|X_0| + X_1 \quad (53)$$

em que $X_0 \sim N(0,1)$ e $X_1 \sim N(\mu, \sigma^2)$ são independentes, desta forma $Z \xrightarrow{d} SSN(\mu, \sigma^2, \lambda)$. Com base nesta representação estocástica é possível computar figuras da distribuição em questão para avaliar seu comportamento, conforme mostra-se na Figura 16.

A Expectância e Variância de uma v.a. $Z \sim SSN(\mu, \sigma^2, \lambda)$ são dadas, respectivamente, pelas equações (54) e (55).

$$E(Y) = \mu + \sqrt{\frac{2}{\pi}} \lambda \quad (54)$$

$$V(Y) = \sigma + \left(1 - \frac{2}{\pi}\right) \lambda^2 \quad (55)$$

Para implementação da função de verossimilhança:

1º Passo: Definir μ

$$\mu = \eta(x_i, \theta) - \lambda \sqrt{\frac{2}{\pi}}$$

$$f_Y(y) = 2\phi\left(\eta(x_i, \theta) - \lambda \sqrt{\frac{2}{\pi}}, \sigma^2 + \lambda^2\right) \Phi\left(\frac{\lambda}{\sigma} \times \frac{y - \eta(x_i, \theta) + \lambda \sqrt{\frac{2}{\pi}}}{\sqrt{\sigma^2 + \lambda^2}}\right)$$

2º Passo: Obter l_i

$$l_i = \log 2 - 0.5 \log(\sigma^2 + \lambda^2) - 0.5 \log 2\pi - 0.5 \left(\frac{y - \eta(x_i, \theta) + \lambda \sqrt{\frac{2}{\pi}}}{\sqrt{\sigma^2 + \lambda^2}} \right)^2 \\ + \log \Phi\left(\frac{\lambda}{\sigma} \times \frac{y - \eta(x_i, \theta) + \lambda \sqrt{\frac{2}{\pi}}}{\sqrt{\sigma^2 + \lambda^2}}\right)$$

3º Passo: Remova as constantes e faça $\tau = 1/\sigma^2$

$$l_i = -0.5 \log(\tau^{-1} + \lambda^2) - 0.5 \left(\frac{y - \eta(x_i, \theta) + \lambda \sqrt{\frac{2}{\pi}}}{\sqrt{\tau^{-1} + \lambda^2}} \right)^2 + \log \Phi\left(\sqrt{\tau} \lambda \frac{y - \eta(x_i, \theta) + \lambda \sqrt{\frac{2}{\pi}}}{\sqrt{\tau^{-1} + \lambda^2}}\right) \quad (56)$$

3.2.5 Distribuição Skew-Normal via transformação de Fernández & Steel (1998)

Uma forma de se obter distribuições de probabilidades assimétricas é por meio da metodologia de Fernández & Steel. No trabalho de Fernández & Steel (1998) os autores propuseram, como aplicação da metodologia, uma versão assimétrica da distribuição t de Student para o resíduo numa implementação bayesiana de uma regressão linear. A metodologia consiste em introduzir um parâmetro de assimetria a qualquer distribuição contínua simétrica e unimodal, a transformação é dada por:

$$f_X(x) = \frac{2}{\lambda + \lambda^{-1}} \left\{ g\left(\frac{x}{\lambda}\right) I_{[0, +\infty)} + g(x\lambda) I_{(-\infty, 0)} \right\} \quad (57)$$

em que λ é o parâmetro de assimetria e $g(\cdot)$ é a densidade de uma distribuição simétrica tal que $g(x) = -g(-x)$.

Os resultados da metodologia são distribuições mais flexíveis com caudas pesadas, e presença de assimetria. Pode-se derivar uma forma para os k -ésimos momentos com base na equação 58.

$$E(Z^k) = \frac{\lambda}{1 + \lambda^2} \left(\lambda^{k+1} - \frac{(-1)^k}{\lambda^{k+1}} \right) \int_0^{+\infty} 2z^k f(z) dz \quad (58)$$

em que $f(z)$ é uma distribuição simétrica e unimodal em 0. Se $k = 2n \forall n \in \mathbb{R}$. Com uma transformação de locação-escala obtêm-se.

$$E(\mu + \sigma Z)^n = \frac{\lambda}{1 + \lambda^2} \sum_{k=0}^n \mu^{n-k} \sigma^k \left(\lambda^{k+1} - \frac{(-1)^k}{\lambda^{k+1}} \right) \int_0^{+\infty} 2x^k f(x) dx \quad (59)$$

No trabalho de Castillo et al. (2011) os autores aplicaram a equação (58) à distribuição Normal e depois aplicaram uma transformação de locação-escala obtendo a densidade:

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \left\{ \phi \left(\lambda \frac{y - \mu}{\sigma} \right) I_{(y < \mu)} + \phi \left(\frac{y - \mu}{\lambda \sigma} \right) I_{(y \geq \mu)} \right\} \quad (60)$$

tal que $Y \sim FSSN(\mu, \sigma^2, \lambda)$ que representa a densidade da distribuição Skew-Normal de Fernández & Steel. Em termos de assimetria dessa distribuição, o parâmetro de assimetria (λ) denota a orientação da curva, valores $\lambda > 1$ denotam assimetria à direita, $0 < \lambda < 1$ assimetria à esquerda, $\lambda = 1$ assimetria nula (convergência para distribuição Normal), comportamento esse que pode ser verificado na Figura 17 que foi gerada por meio do pacote *Skewt* (v. 1.0).

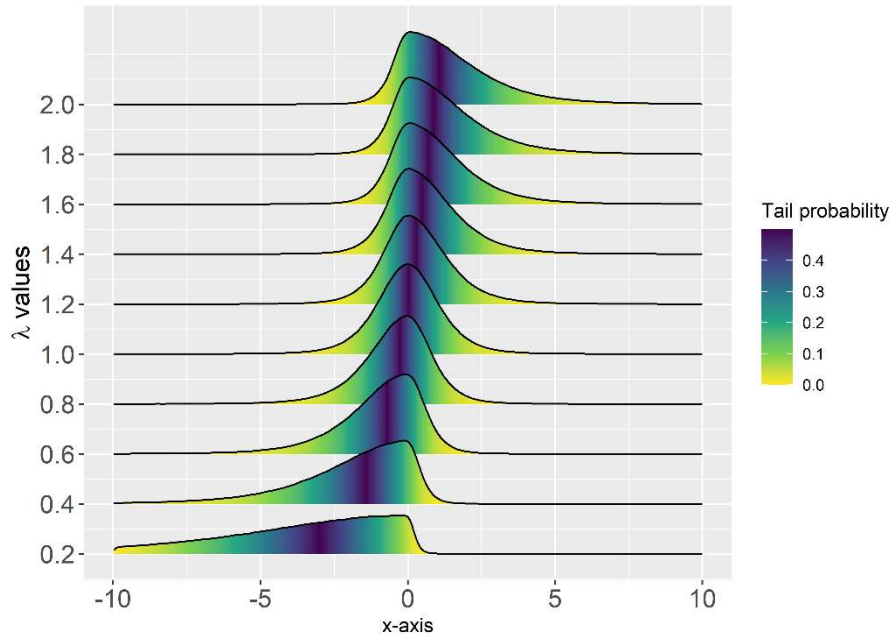


Figura 17: Comportamento da distribuição Skew-Normal de Fernández & Steel padronizada variando o parâmetro de assimetria (λ values).

Fonte: O autor

Uma das vantagens inferenciais de FSSN consiste na não singularidade da matriz de informação esperada de Fisher para $\lambda = 1$ (ASN é singular quando $\lambda = 0$), o que facilita a inferência assintótica (Castillo et al, 2011). Outra vantagem de FSSN consiste na não utilização da f.d.a. $F(x)$, que não possui expressões analíticas para Normal e t de Student (Eloy et al, 2019). Castillo et al. (2011) demonstra algumas propriedades de FSSN considerando o r-ésimo momento amostral de Y (equação 61) e consequentemente a Expectância e Variância de FSSN expressos pela equação (62) e (63).

$$E(Y^r) = \sum_{k=0}^r \binom{r}{k} \sigma^k \mu^{r-k} \sqrt{\frac{2^k}{\pi}} \Gamma\left(\frac{1+k}{2}\right) \frac{(-1)^k + \lambda^{2(k+1)}}{\lambda^k(1+\lambda^2)} \quad (61)$$

$$E(Y) = \mu + \sigma \sqrt{\frac{2}{\pi}} \frac{\lambda^2 - 1}{\lambda} \quad (62)$$

$$V(Y) = \sigma^2 \frac{(\pi - 2)\lambda^6 + 2\lambda^2(\lambda^2 + 1) + \pi - 2}{\pi\lambda^2(1 + \lambda^2)} \quad (63)$$

Na implementação da função de verossimilhança:

1ºPasso: Definir μ

$$\mu = \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}$$

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \left\{ \phi \left(\frac{y - \eta(x_i, \theta) + \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}}{\lambda \sigma} \right) I_{\left(y \geq \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right)} \right. \\ \left. + \phi \left(\lambda \frac{y - \eta(x_i, \theta) + \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}}{\sigma} \right) I_{\left(y < \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right)} \right\}$$

2º Passo: Rescrever $f_Y(y)$ em função da função de Heavside $H(x)$ que no OpenBugs é representada pela função *Step*.

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \phi \left(\lambda^{1-2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right)} \frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \right)$$

em que,

$$H \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \right) := \begin{cases} 1, & y_i \geq \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \\ 0, & y_i < \eta(x_i, \theta) - \sigma \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \end{cases}$$

3º Passo: Obtenha l_i

$$l_i = \log 2 + \log \lambda - \log \sigma - \log(1 + \lambda^2) - 0.5 \log 2 \pi - \log \lambda^{2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right) - 1} \\ - \log \sigma - 0.5 \left(\lambda^{1-2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right)} \frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \right)^2$$

4º Passo: Remova as constantes e faça $\tau = 1/\sigma^2$

$$l_i = \log \lambda + \log \tau - \log(1 + \lambda^2) - \log \lambda^{2H\left(\left(y_i - \eta(x_i, \theta)\right)\sqrt{\tau} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right) - 1} \\ - 0.5 \left(\lambda^{1-2H\left(\left(y_i - \eta(x_i, \theta)\right)\sqrt{\tau} + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}}\right)} \left(y_i - \eta(x_i, \theta)\right)\sqrt{\tau} \right. \\ \left. + \frac{\lambda^2 - 1}{\lambda} \sqrt{\frac{2}{\pi}} \right)^2 \quad (64)$$

3.2.6 Distribuição Skew-Laplace de Kotz (2001)

Outra forma de distribuição contínua assimétrica é obtida, utilizando-se de uma transformação da distribuição de Laplace, distribuição essa também utilizada em regressões robustas. Da conhecida distribuição Laplace,

$$f_Y(y) = \frac{1}{2\sigma} \exp\left(-\frac{|y - \mu|}{\sigma}\right) \quad (65)$$

com a transformação da equação 60, utilizando uma inversão dos sinais na equação 65, obtêm-se a distribuição skew-Laplace:

$$f_Y(y) = \frac{\sigma\lambda}{1 + \lambda^2} \left\{ \exp\left(\frac{\sigma}{\lambda}(y - \mu)\right)_{x < \mu} + \exp(-\sigma\lambda(y - \mu))_{x \geq \mu} \right\} \quad (66)$$

No que concerne ao parâmetro de assimetria (λ), valores entre (0,1) reforçam a assimetria à direita da curva, valores entre (1, $+\infty$) tem-se a assimetria à esquerda da curva, quando $\lambda = 1$ retorna-se ao submodelo Laplace. A média e Variância são dadas pelas equações 67 e 68.

$$E(Y) = \mu + \frac{\sigma}{\lambda}(1 - \lambda^2) \quad (67)$$

$$V(Y) = \sigma^2 \frac{1 + \lambda^4}{\lambda^2} \quad (68)$$

Existem outras formas de se obter a distribuição Laplace assimétrica, utilizando a transformação de Azzalini, entretanto, conforme exposto em Nadarajah (2006), tal distribuição possui propriedades poucas desenvolvidas, e tal autor apresenta expressões não analíticas para função geratriz de momentos, variância e expectância.

Pode-se gerar números aleatórios da distribuição ASL de forma direta, utiliza-se o Teorema da Transformação Uniformizante:

$$Y = \mu - \frac{\log(1 - U \operatorname{sgn}(U)\lambda^{\operatorname{sgn}(U)})}{\sigma \operatorname{sgn}(U)\lambda^{\operatorname{sgn}(U)}} \quad (69)$$

Em que $U \sim \operatorname{Unif}(-\lambda, \lambda^{-1})$, as densidades obtidas pela equação 69 são verificadas nos gráficos da Figura 18 que mostram o comportamento da distribuição ASL.

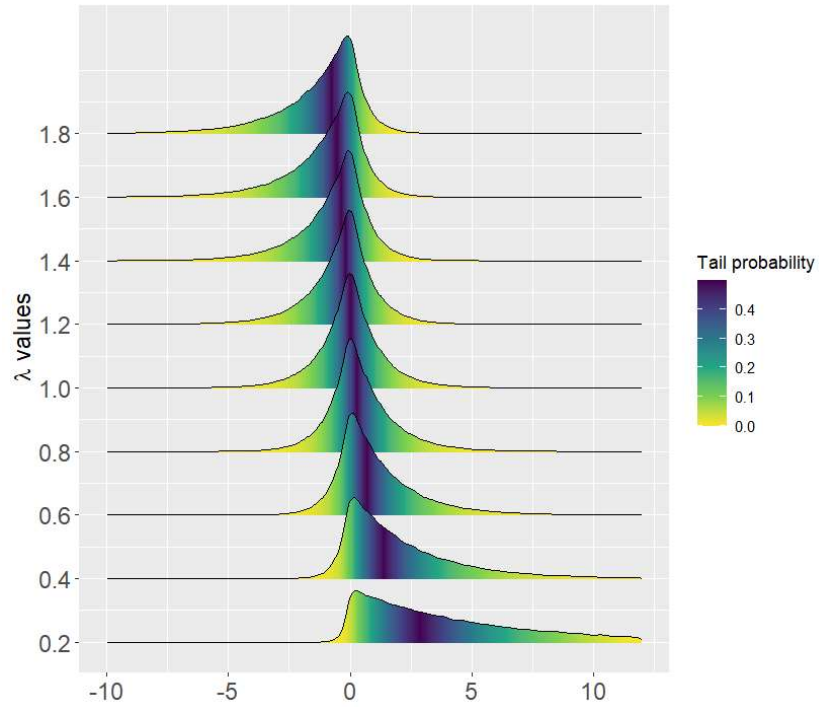


Figura 18: Comportamento da distribuição ASL padronizada variando os valores de lambda (λ values).

Fonte: O autor

Para implementação da função de verossimilhança:

1º Passo: Definir μ

$$\mu = \eta(x_i, \theta) - \frac{\sigma}{\lambda}(1 - \lambda^2)$$

$$f_Y(y) = \frac{\sigma\lambda}{1 + \lambda^2} \left\{ \exp\left(\frac{\sigma}{\lambda}\left(y - \eta(x_i, \theta) + \frac{\sigma}{\lambda}(1 - \lambda^2)\right)\right)_{x < \mu} + \exp\left(-\sigma\lambda\left(y - \eta(x_i, \theta) + \frac{\sigma}{\lambda}(1 - \lambda^2)\right)\right)_{x \geq \mu} \right\}$$

2º Passo: Reescrever $f_Y(y)$ em termos da função de Heavside

$$f_Y(y) = \frac{\sigma\lambda}{1 + \lambda^2} \exp\left\{-(2H(x) - 1)\sigma\lambda^{2H(x)-1}\left(y - \eta(x_i, \theta) + \frac{\sigma}{\lambda}(1 - \lambda^2)\right)\right\}$$

3º Passo: Obter l_i

$$l_i = \log \sigma - \log \lambda - \log(1 + \lambda^2) - (2H(x) - 1)\sigma\lambda^{2H(x)-1}\left(y - \eta(x_i, \theta) + \frac{\sigma}{\lambda}(1 - \lambda^2)\right)$$

4º Passo: Façamos $\sigma = \sqrt{\tau}$

$$l_i = \log \sigma - \log(\lambda + \lambda^{-1}) - (2H(x) - 1)\sigma\lambda^{2H(x)-1} \left(y_i - \eta(x_i, \theta) + \sigma \frac{1 - \lambda^2}{\lambda} \right) \quad (70)$$

3.2.7 Distribuição Skew-Logistic de Sastry and Bhati (2016) - SBSL

A distribuição Skew-Logistic de Sastry & Bhati é obtida utilizando a transformação de Fernández & Steel, entretanto utilizando o sinal invertido da função indicadora de λ , da mesma forma da distribuição anterior (equação 66). A motivação da criação desta distribuição se deve ao fato de que a transformação de Azzalini utilizar-se de funções cumulativas não analíticas. A densidade desta nova distribuição é dada por:

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \left\{ e^{-\lambda \frac{y-\mu}{\sigma}} \left(1 + e^{-\lambda \frac{y-\mu}{\sigma}} \right)^{-2} I_{(y \geq \mu)} + e^{-\frac{y-\mu}{\sigma\lambda}} \left(1 + e^{-\frac{y-\mu}{\sigma\lambda}} \right)^{-2} I_{(y < \mu)} \right\} \quad (71)$$

Conforme pode ser verificado na Figura 19, quando o valor de λ varia de $(0,1)$, tem-se assimetria à direita na curva, quando λ varia de $(1, +\infty]$, tem-se assimetria à esquerda na curva, quando $\lambda = 1$ retorna-se ao submodelo Logístico.

Dado que $Y \sim SBSL(\mu, \sigma^2, \lambda)$, têm-se a esperança:

$$E(Y) = \mu + \sigma \frac{1 - \lambda^2}{\lambda} \log 4 \quad (72)$$

e a variância:

$$V(X) = \frac{\lambda\sigma^2}{1 + \lambda^2} \left(\lambda^3 + \frac{1}{\lambda^3} \right) \frac{\pi^2}{3} - \left(\sigma \left(\frac{1 - \lambda^2}{\lambda} \right) \log_e 4 \right)^2 \quad (73)$$

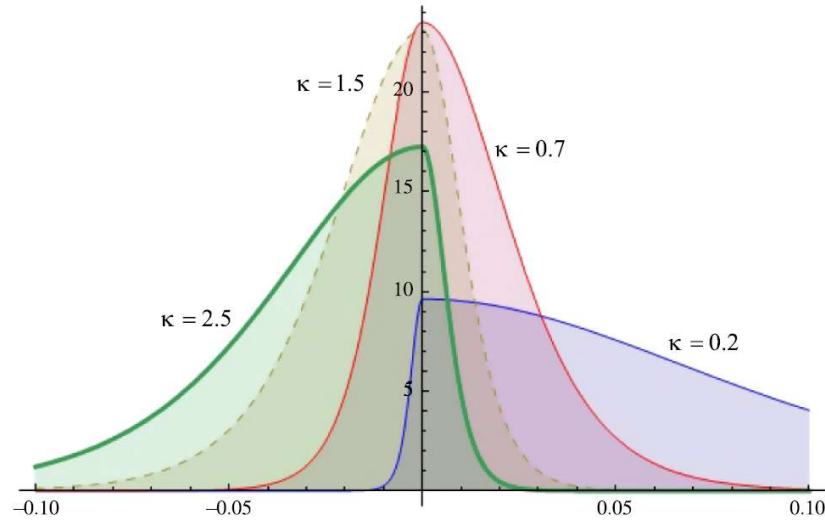


Figura 19: Comportamento da distribuição SBSL padronizada variando os valores de lambda (κ).

Fonte: Sastry & Bathi (2016)

Para implementação da função de verossimilhança:

1º Passo: Definir correção para μ

$$\mu = \eta(x_i, \theta) - \sigma \frac{1 - \lambda^2}{\lambda} \log 4$$

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \left\{ e^{-\lambda \frac{y - \eta(x_i, \theta) + \sigma \frac{1 - \lambda^2}{\lambda} \log 4}{\sigma}} \left(1 + e^{-\lambda \frac{y - \eta(x_i, \theta) + \sigma \frac{1 - \lambda^2}{\lambda} \log 4}{\sigma}} \right)^{-2} I_{\left(y \geq \eta(x_i, \theta) - \sigma \frac{1 - \lambda^2}{\lambda} \log 4\right)} + e^{-\frac{y - \eta(x_i, \theta) + \sigma \frac{1 - \lambda^2}{\lambda} \log 4}{\sigma \lambda}} \left(1 + e^{-\frac{y - \eta(x_i, \theta) + \sigma \frac{1 - \lambda^2}{\lambda} \log 4}{\sigma \lambda}} \right)^{-2} I_{\left(y < \eta(x_i, \theta) - \sigma \frac{1 - \lambda^2}{\lambda} \log 4\right)} \right\}$$

2º Passo: Definir $f_Y(y)$ em termos da função de Heavside

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \left\{ \exp \left\{ -\lambda^{2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4 \right\} \left(1 + \exp \left\{ -\lambda^{2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4 \right\} \right)^{-2} \right\}$$

em que:

$$H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4\right) := \begin{cases} 1, & y_i \geq \eta(x_i, \theta) - \sigma \frac{1 - \lambda^2}{\lambda} \log 4 \\ 0, & y_i < \eta(x_i, \theta) - \sigma \frac{1 - \lambda^2}{\lambda} \log 4 \end{cases}$$

3º Passo: Obter l_i

$$\begin{aligned} l_i = & \log 2 - \log \lambda - \log \sigma - \log(1 + \lambda^2) \\ & - \lambda^{2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4 \right) \\ & + 2 \log \left(1 \right. \\ & \left. + \exp \left\{ -\lambda^{2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} + \frac{1 - \lambda^2}{\lambda} \log 4 \right) \right\} \right) \end{aligned}$$

4º Passo: Remover as constantes e fazer $\tau = 1/\sigma^2$

$$\begin{aligned} l_i = & -\log \lambda + 0.5 \log \tau - \log(1 + \lambda^2) - \lambda^{2H\left((y - f(x_i, \theta))\sqrt{(\tau)} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \left((y - \right. \\ & \left. f(x_i, \theta))\sqrt{(\tau)} + \frac{1 - \lambda^2}{\lambda} \log 4 \right) + 2 \log \left(1 + \exp \left\{ -\lambda^{2H\left((y - f(x_i, \theta))\sqrt{(\tau)} + \frac{1 - \lambda^2}{\lambda} \log 4\right) - 1} \left((y - \right. \right. \\ & \left. \left. f(x_i, \theta))\sqrt{(\tau)} + \frac{1 - \lambda^2}{\lambda} \log 4 \right) \right\} \right) \end{aligned} \quad (74)$$

3.2.8 Distribuição Skew- t de Fernández & Steel (1998)

A distribuição t de Student tipo 3 foi obtida por Fernandez & Steel (1998) por meio da transformação dos mesmos autores publicada em 1998. Os autores propuseram uma versão assimétrica da distribuição t de Student por meio de sua transformação com a finalidade de modelar os resíduos numa regressão linear bayesiana. Utilizando a transformação dada pela equação (60) à distribuição t de Student, obtemos a distribuição Skew-t de Fernández & Steel.

$$f_Y(y) = \frac{2\lambda}{\sigma(1 + \lambda^2)} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \left\{ 1 + \frac{1}{v\lambda^2} \left(\frac{y - \mu}{\sigma}\right)^2 I_{[0, +\infty)} + \frac{\lambda^2}{v} \left(\frac{y - \mu}{\sigma}\right)^2 I_{(-\infty, 0)} \right\}^{-\frac{v+1}{2}} \quad (75)$$

em que μ e σ são os parâmetros de locação e escala, v é o parâmetro de controle de caudas, e $\Gamma(\cdot)$ sendo a função Gamma.

Ao considerar uma variável que tem distribuição t de Student tipo 3, denotamos $X \sim FSSSt(\mu, \sigma^2, \lambda, v)$. Quando $\lambda > 1$ tem-se assimetria á direita, quando $1 > \lambda > 0$ tem-se assimetria á esquerda, quando $\lambda = 0$ tem-se assimetria nula (convergência para o submodelo Student t). O comportamento gráfico desta distribuição pode ser avaliado por meio da Figura 20 utilizando o pacote *Skewt* (v.1.0).

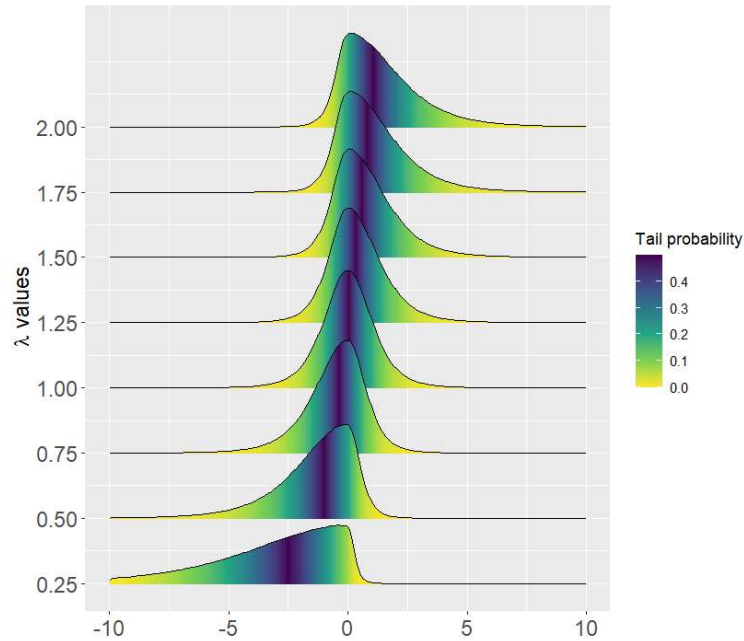


Figura 20: Comportamento da distribuição $FSSSt$ padronizada variando os valores de lambda (λ values).

A Expectância da distribuição em questão não possui expressão analítica, o mesmo para Variância e para momentos de ordem superior, no trabalho de Trottier & Ardia (2016) os autores mostram desenvolvimentos para os momentos para distribuições de Fernández & Steel, mas que também não são matematicamente tratáveis para implementação no OpenBugs. Neste caso específico não há como fazer correção para média.

Para implementação da função de verossimilhança:

1º Passo: Escrever $f_Y(y)$ em termos da função de heavyside.

$$f_Y(y) = \frac{2\lambda}{\sigma(1+\lambda^2)} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \left\{ 1 + \frac{\lambda^2 \left(1 - 2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)\right)}{v} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)^2 \right\}^{-\frac{v+1}{2}}$$

em que

$$H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right) := \begin{cases} 1, & y_i \geq \eta(x_i, \theta) \\ 0, & y_i < \eta(x_i, \theta) \end{cases}$$

2º Passo: Obter l_i

$$l_i = \log 2 + \log \lambda - \log \sigma - \log(1 + \lambda^2) + \log \Gamma\left(\frac{v+1}{2}\right) - \log \Gamma\left(\frac{v}{2}\right) - 0.5 \log v\pi - \frac{v+1}{2} \log \left\{ 1 + \frac{\lambda^{2\left(1-2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)\right)}}{v} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)^2 \right\}$$

3º Passo: Remover as constantes e faça $\tau = 1/\sigma^2$

$$l_i = \log \lambda + 0.5 \log \tau - \log(1 + \lambda^2) + \log \Gamma\left(\frac{v+1}{2}\right) - \log \Gamma\left(\frac{v}{2}\right) - 0.5 \log v - \frac{v+1}{2} \log \left\{ 1 + \frac{\lambda^{2\left(1-2H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)\right)}}{v} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma}\right)^2 \right\} \quad (76)$$

3.2.9 Distribuição Epsilon-Skew-t (GOMEZ et al., 2007a)

A distribuição epsilon-Skew-t é um caso especial de uma classe de modelos generalizados denominada *skew-power-exponential*. A distribuição ESSt tem como submodelo a distribuição t de Student e a distribuição *epsilon-skew-normal*. O parâmetro que controla a assimetria que geralmente é ϵ , neste trabalho é renomeada de λ , tal que $|\lambda| < 1$. Tal distribuição tem a seguinte densidade:

$$f_Y(y) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sigma\sqrt{v\pi}} \left\{ 1 + \frac{1}{v} \left(\frac{y - \mu}{\sigma(1 - \lambda)}\right)^2 I_{[0, +\infty)} + \frac{1}{v} \left(\frac{y - \mu}{\sigma(1 + \lambda)}\right)^2 I_{(-\infty, 0)} \right\}^{-\frac{v+1}{2}} \quad (77)$$

$$E(Y) = \mu - \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1} \quad (78)$$

$$Var(Z) = \sigma^2 \left\{ \frac{v(1 + 3\lambda^2)}{v-2} - \frac{16(c(\nu)\lambda v)^2}{(v-1)^2} \right\} \quad (79)$$

em que $c(\nu) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}}$.

Existe uma representação estocástica para distribuição ESSt, entretanto, neste trabalho, optou-se por utilizar o algoritmo de aceitação-rejeição para computar os gráficos da Figura 21 utilizando a distribuição Student t como candidata. O parâmetro λ denota a assimetria da

distribuição, $0 > \lambda > -1$ indica assimetria á direita, $1 > \lambda > 0$ indica assimetria á esquerda, e $\lambda = 0$ assimetria nula (convergência para o submodelo Student t) conforme pode ser visualizado na Figura 21.

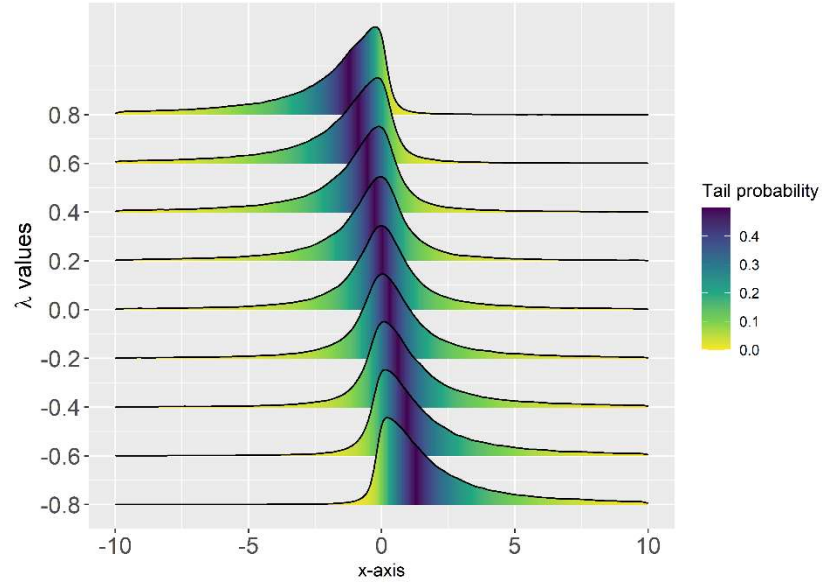


Figura 21: Comportamento da distribuição ESSt padronizada variando os valores de lambda (λ values).

Para implementação da função de verossimilhança:

1º Passo: Fazer correção para μ

$$\mu = \eta(x_i, \theta) + \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}$$

$$f_Y(y) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sigma\sqrt{v\pi}} \left\{ 1 + \frac{1}{v} \left(\frac{y - \eta(x_i, \theta) - \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}}{\sigma(1-\lambda)} \right)^2 I_{[0,+\infty)} \right. \\ \left. + \frac{1}{v} \left(\frac{y - \eta(x_i, \theta) - \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}}{\sigma(1+\lambda)} \right)^2 I_{(-\infty,0)} \right\}^{-\frac{v+1}{2}}$$

2º Passo: Escrever $f_Y(y)$ em termos da função de heavyside.

$$f_Y(y) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sigma\sqrt{v\pi}} \left\{ 1 + \frac{1}{v} \left(\frac{y_i - \eta(x_i, \theta) - \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}}{\sigma \left(1 + (-1)^{H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} - \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}\right)} \lambda \right)} \right)^2 \right\}^{-\frac{v+1}{2}}$$

em que

$$H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} - \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}\right) := \begin{cases} 1, & y_i \geq \eta(x_i, \theta) + \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1} \\ 0, & y_i < \eta(x_i, \theta) + \sigma \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1} \end{cases}$$

3º Passo: Obter l_i

$$\begin{aligned} l_i = & \log \Gamma\left(\frac{v+1}{2}\right) - \log \Gamma\left(\frac{v}{2}\right) - \log \sigma - 0.5 \log v \pi \\ & - \frac{v+1}{2} \log \left\{ 1 \right. \\ & + \frac{1}{v} \left(\left(1 + (-1)^{H\left(\frac{y_i - \eta(x_i, \theta)}{\sigma} - \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1}\right)} \lambda \right)^{-1} \left[\frac{y - \eta(x_i, \theta)}{\sigma} \right. \right. \\ & \left. \left. - \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \frac{4\lambda v}{v-1} \right] \right)^2 \right\} \end{aligned}$$

4º Passo: Remover as constantes e faça $\tau = 1/\sigma^2$

Utilizando o algoritmo de aceitação-rejeição com a distribuição Laplace com candidata, computamos as densidades da distribuição SDEN conforme a Figura 22.

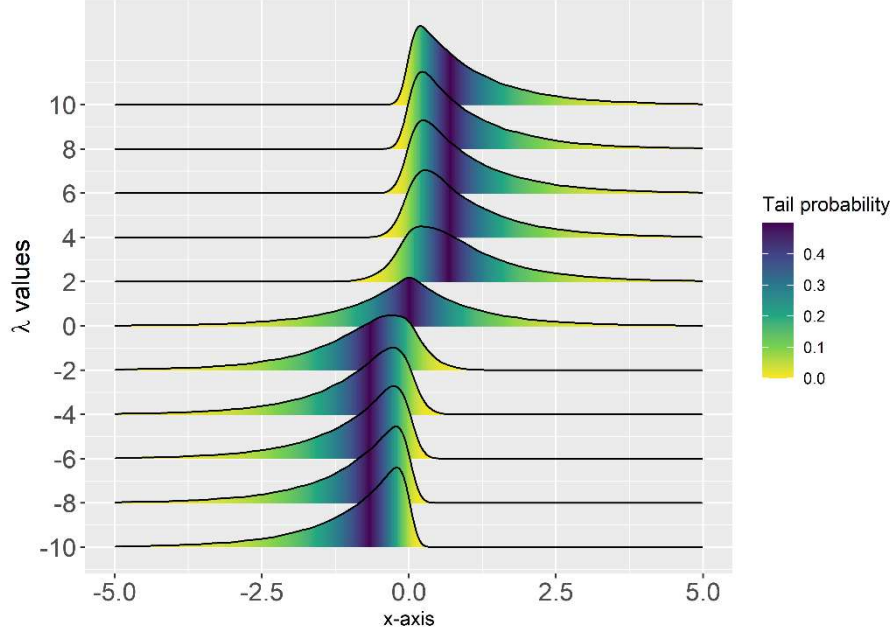


Figura 22: Comportamento da distribuição *SDEN* padronizada variando os valores de lambda (λ values).

A esperança da distribuição é definida conforme a equação 84 e a variância pode ser derivada sabendo-se que $E(Y^2) = 8$.

$$E(Y) = \mu + \sigma \frac{(4\lambda^2 - 1) \exp \left\{ \frac{1}{8\lambda^2} \right\} \left[1 - 2\Phi \left(\frac{1}{2\lambda} \right) + \text{sign}(\lambda) \right] + 2\lambda\sqrt{(2\pi)}}{2\pi\lambda^2} \quad (84)$$

Para implementação da função de verossimilhança o Openbugs não aceita funções do tipo $\Phi(\frac{1}{2\lambda})$ com $\lambda \in \mathbb{R}$, ainda que exista solução real, neste caso não poderá ser feita a correção para média.

1º Passo: Obter l_i

$$l_i = -\log 2 - \log \sigma - 0.5 \left| \frac{y - \eta(x_i, \theta)}{\sigma} \right| + \log \Phi \left(\lambda \frac{y - \eta(x_i, \theta)}{\sigma} \right)$$

2º Passo: Remover as constantes e fazer $\tau = 1/\sigma^2$

$$l_i = 0.5 \log \sigma - 0.5 \left| (y - \eta(x_i, \theta)) \sqrt{(\tau)} \right| + \log \Phi \left\{ \lambda (y - \eta(x_i, \theta)) \sqrt{(\tau)} \right\} \quad (85)$$

3.2.11 Distribuição Skew-t de Jones & Faddy (2003)

Os autores propuseram uma extensão da distribuição t de Student que seja matematicamente tratável que mantenha a capacidade da distribuição t de ser robusta à outliers

e modelar a assimetria presente nos dados. A nova distribuição tem a seguinte densidade:

$$f_Y(y) = C_{a,b}^{-1} \frac{1}{\sigma} \left\{ 1 + \frac{\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{a+b+\left(\frac{y-\mu}{\sigma}\right)^2}} \right\}^{a+0.5} \left\{ 1 - \frac{\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{a+b+\left(\frac{y-\mu}{\sigma}\right)^2}} \right\}^{b+0.5} \quad (86)$$

em que $C_{a,b}^{-1} = 2^{a+b-1} B(a, b)(a+b)^{0.5}$ e $B(\cdot)$ é a função Beta.

A Expectância é dada por:

$$E(Y) = \mu + \sigma \frac{(a-b)\sqrt{a+b} \Gamma\left(a-\frac{1}{2}\right) \Gamma\left(b-\frac{1}{2}\right)}{2\Gamma(a)\Gamma(b)} \quad (87)$$

A distribuição JFSSt possui várias representações estocásticas, utilizamos a seguinte representação dada pela equação 88 para computar as densidades da Figura 23.

$$T = \frac{\sqrt{a+b}(U-V)}{2\sqrt{UV}} \quad (88)$$

em que $U \sim \chi_{2a}^2$ e $V \sim \chi_{2b}^2$ são independentes. De acordo com Jones & Faddy (2003), os parâmetros a e b que controlam a assimetria podem ser representados por $\lambda = b - a$, e para controlar as caudas das distribuições pode-se utilizar $v = a + b$. Na Figura 23 verifica-se assimetria á direita quando $\lambda < 0$, assimetria á esquerda quando $\lambda > 0$, e assimetria nula quando $\lambda = 0$ (convergência para o submodelo Student t com $v = 2a$).

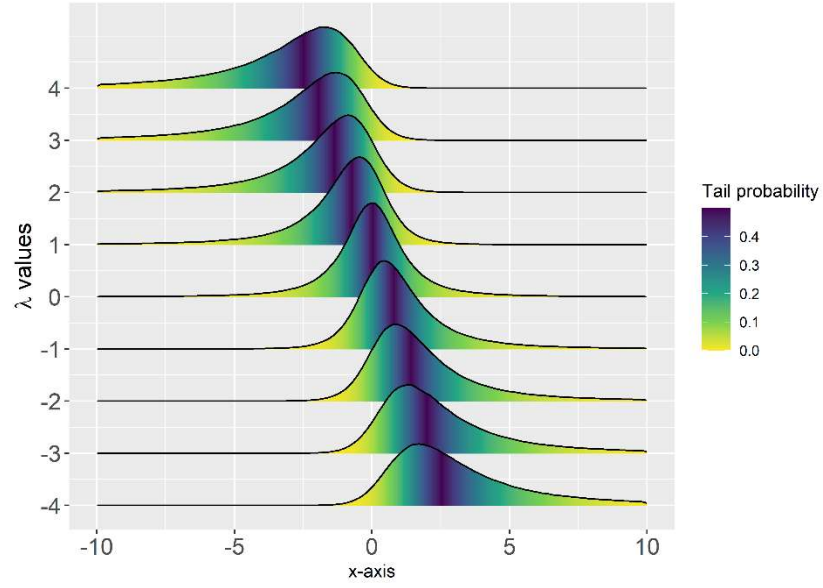


Figura 23: Comportamento da distribuição JFSSt padronizada variando os valores de lambda (λ values).

Para implementação da função de verossimilhança:

1º Passo: Fazer correção para μ .

$$\mu = \eta(x_i, \theta) - \sigma \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)}$$

2º Passo: Obter l_i .

$$f_Y(y) = C_{a,b}^{-1} \frac{1}{\sigma} \left\{ 1 + \frac{\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{a+b + \left(\frac{y-\mu}{\sigma}\right)^2}} \right\}^{a+0.5} \left\{ 1 - \frac{\left(\frac{y-\mu}{\sigma}\right)}{\sqrt{a+b + \left(\frac{y-\mu}{\sigma}\right)^2}} \right\}^{b+0.5}$$

$$C_{a,b}^{-1} = 2^{a+b-1} B(a, b) (a+b)^{0.5}$$

$$l_i = -(a+b-1) \log 2 - \log B(a, b) - 0.5 \log(a+b) - \log \sigma + (a+0.5) \log \left\{ 1 + \left(\frac{\frac{y-\eta(x_i, \theta)}{\sigma} + \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)}}{a+b + \left(\frac{y-\eta(x_i, \theta)}{\sigma} + \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right)^2} \right\}^{-0.5} + (b+0.5) \log \left\{ 1 - \left(\frac{\frac{y-\eta(x_i, \theta)}{\sigma} + \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)}}{a+b + \left(\frac{y-\eta(x_i, \theta)}{\sigma} + \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right)^2} \right\}^{-0.5} \right\}.$$

3º Passo: Remover as constantes e faça $\tau = 1/\sigma^2$.

$$\begin{aligned}
 l_i = & -(a+b) \log 2 - \log B(a,b) - 0.5 \log(a+b) + 0.5 \log \tau + (a+0.5) \log \left\{ 1 + \right. \\
 & \left((y - f(x_i, \theta)) \sqrt{\tau} + \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right) \left(a+b + \left((y - f(x_i, \theta)) \sqrt{\tau} + \right. \right. \\
 & \left. \left. \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right)^2 \right)^{-0.5} \left. \right\} + (b+0.5) \log \left\{ 1 - \left((y - f(x_i, \theta)) \sqrt{\tau} + \right. \right. \\
 & \left. \left. \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right) \left(a+b + \left((y - f(x_i, \theta)) \sqrt{\tau} + \right. \right. \right. \\
 & \left. \left. \left. \frac{(a-b)\sqrt{a+b} \Gamma(a-\frac{1}{2}) \Gamma(b-\frac{1}{2})}{2\Gamma(a)\Gamma(b)} \right)^2 \right)^{-0.5} \right\} \quad (89)
 \end{aligned}$$

3.2.12 Distribuição Skew-t via transformação sinh-arcsinh

A distribuição t de Student via transformação sinh-arcsinh, é outro método de se obter distribuição simétricas que, de acordo com Rosco et al. (2011), têm o mesmo nível de complexidade em relação as outras versões t assimétricas. A densidade da nova distribuição é dada pela equação 46.

$$f_Y(y) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sigma\sqrt{v\pi}} \times \frac{\cosh\left(\operatorname{arcsinh}\left(\frac{y-\mu}{\sigma}\right) - \lambda\right)}{\sqrt{1 + \left(\frac{y-\mu}{\sigma}\right)^2} \left\{ v^{-1} \cosh^2\left(\operatorname{arcsinh}\left(\frac{y-\mu}{\sigma}\right) - \lambda\right) \right\}^{(v+1)/2}} \quad (90)$$

em que λ é o parâmetro de controle de assimetria, valores negativos de λ denotam assimetria á esquerda, valores positivos de λ denotam assimetria á direita, e $\lambda = 0$ assimetria nula (convergência para o submodelo Student t).

A expectância e variância não são matematicamente tratáveis, neste caso não haverá correção para média.

1º Passo: Obter l_i .

$$\begin{aligned}
 l_i = & \log \Gamma\left(\frac{v+1}{2}\right) - \Gamma\left(\frac{v}{2}\right) - \log \sigma - 0.5 \log v \pi \\
 & + \log \left\{ \cosh \left(\operatorname{arcsinh} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} \right) - \lambda \right) \right\} \\
 & - 0.5 \log \left(1 + \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} \right)^2 \right) \\
 & - \frac{v+1}{2} \log \left\{ v^{-1} \cosh^2 \left(\operatorname{arcsinh} \left(\frac{y_i - \eta(x_i, \theta)}{\sigma} \right) - \lambda \right) \right\}
 \end{aligned}$$

2º Passo: Remover as constantes e faça $\tau = 1/\sigma^2$.

$$l_i = \log \Gamma\left(\frac{v+1}{2}\right) - \Gamma\left(\frac{v}{2}\right) + 0.5 \log \tau - 0.5 \log v + \log \left\{ \cosh \left(\operatorname{arcsinh} \left((y - f(x_i, \theta)) \sqrt{(\tau)} \right) - \lambda \right) \right\} - 0.5 \log \left(1 + \left((y - f(x_i, \theta)) \sqrt{(\tau)} \right)^2 \right) - \frac{v+1}{2} \log \left\{ v^{-1} \cosh^2 \left(\operatorname{arcsinh} \left((y - f(x_i, \theta)) \sqrt{(\tau)} \right) - \lambda \right) \right\} \quad (91)$$

3.2.13 Modelos mistos não lineares na implementação da priori empírica

A priori geralmente é determinada antes de se ter acesso aos dados pois representa o conhecimento prévio do pesquisador ou o conhecimento científico acumulado sobre determinado parâmetro. O artifício de se determinar a priori a partir do conjunto de dados a ser analisado é denominado de abordagem bayesiana empírica, entretanto, ainda sim é uma aproximação à abordagem bayesiana clássica (Carlin & Louis, 2008).

Uma forma de se determinar estimativas frequentista para todos os 8 acessos com o modelo Logístico, é por meio de modelos mistos não lineares, em que num único algoritmo é possível obter as estimativas dos efeitos fixos e aleatórios, ao passo que obter estimativas de Gauss-newton individual para cada combinação é demasiado trabalhoso e é necessário obter estimativas iniciais para cada caso. Por não ser o objetivo deste trabalho, mais detalhes podem ser encontrados no trabalho de Teixeira et al. (2023).

O gráfico de violino e de boxplot da Figura 24 dos resíduos corrobora com a tese de Pino (2014) de que, apesar da primazia da distribuição Normal em diversos procedimentos estatísticos, na prática não se pode supor a normalidade como padrão. É possível notar outliers (distribuições residuais com caudas pesadas), assimetria (distribuições residuais assimétricas), ou ambas (distribuições residuais assimétricas com caudas pesadas), entretanto, o foco deste trabalho está nas distribuições assimétricas.

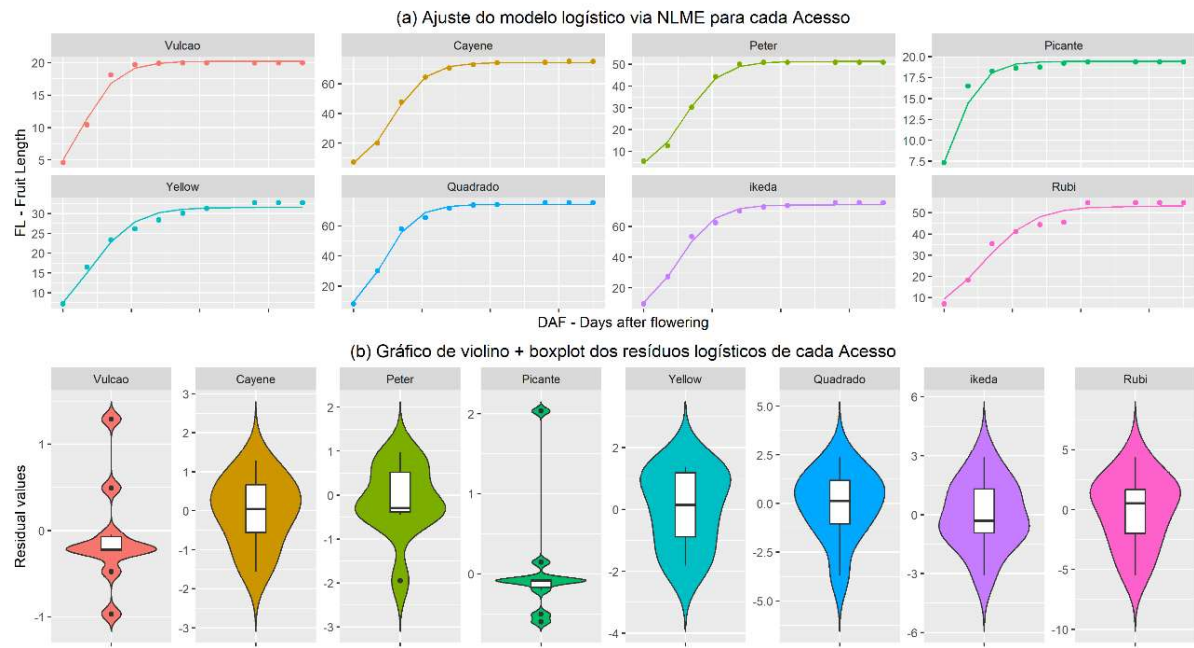


Figura 24: Gráfico de violino em conjunto com o Boxplot dos resíduos, após o ajuste dos modelos Logístico, Gompertz e Von Bertalanffy sob o erro Gaussiano (N) à todos os acessos representados pela Figura 12.

Fonte: O autor.

O segundo passo consiste em plotar as densidades dos histogramas de cada parâmetro ($\alpha, \beta, \gamma, \sigma^2$) dos modelos Logístico, Gompertz e Von Bertalanffy ajustados a todos os acessos (ver Figura 25). As informações das medidas de média e variância de cada histograma são incorporadas como hiperparâmetros da distribuição Normal, distribuição esta que vai ser utilizada como priori de (α, β, γ), e para σ^2 as informações são incorporadas na distribuição gamma.

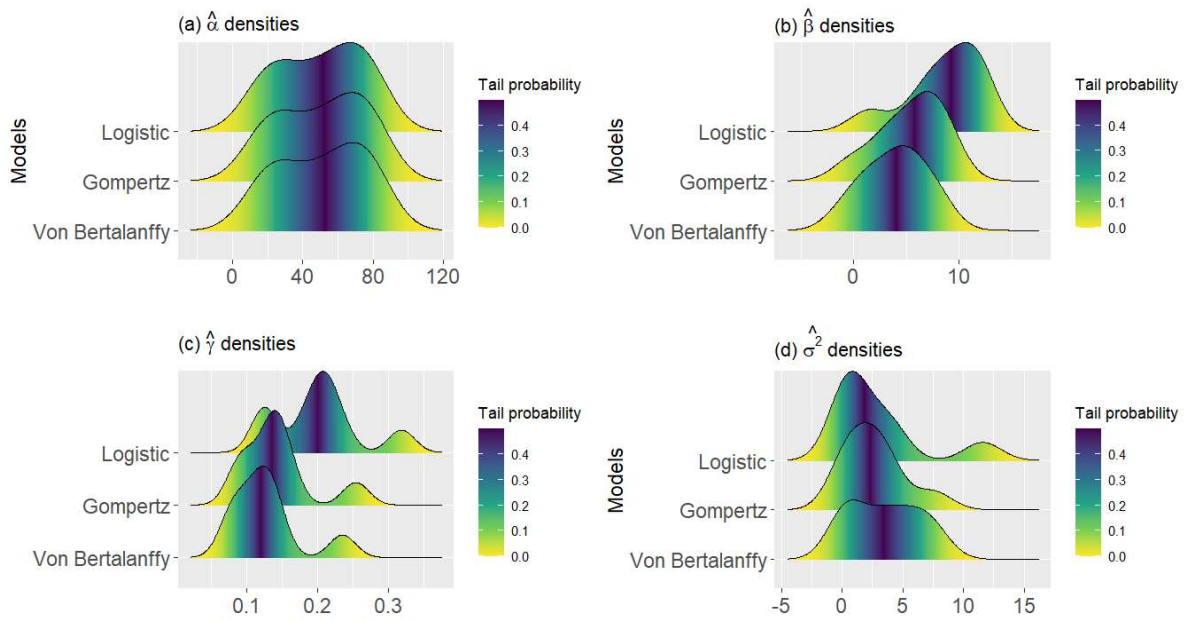


Figura 25: Densidades do histograma das estimativas dos parâmetros $(\alpha, \beta, \gamma, \sigma^2)$ dos modelos Logístico, Gompertz e Von Bertalanffy sob o erro Gaussiano (N) via modelos mistos não lineares ajustados a todos os acessos representados pela Figura 12.

Fonte: O autor.

3.2.14 Estrutura Bayesiana

Considerando o Teorema de Bayes em que, a distribuição a posteriori conjunta é proporcional a distribuição conjunta da amostra (Função de verossimilhança) multiplicada pela distribuição conjunta das distribuições a priori (que podem ser independentes) com seus respectivos hiperparâmetros. Considerando-se o modelo Logístico. A distribuição a posteriori considerando os diferentes erros é dada por: (Perez & Rodriguez, 2018).

$$p(\alpha, \beta, \gamma, \sigma^2, \lambda) \propto \left\{ \prod_{i=1}^n P(Y|y=0, \lambda = K - l_i) \right\} \times N(\mu_\alpha, \sigma_\alpha^2) \times N(\mu_\beta, \sigma_\beta^2) \\ \times N(\mu_\gamma, \sigma_\gamma^2) \times IG(a_{\sigma^2}, b_{\sigma^2}) \times N(\mu_\lambda, \sigma_\lambda^2)$$

$$\alpha \sim N(49.86, 1.76 \times 10^{-3})$$

$$\beta \sim N(8.63, 7.74 \times 10^{-2})$$

$$\gamma \sim N(0.20, 259.90)$$

$$\tau \sim IG(2.54, 4.41)$$

$$\lambda \sim N(0, 10^{-3})$$

em que os hiperparâmetros são obtidos pelos resultados representados na Figura 25, e os l_i 's são dados pelas seguintes equações: (51), (56), (64), (70), (74), (76), (80), (85), (89) e (91). As

distribuições da família de Fernández & Steel tem como priori uma priori Normal Truncada $\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)T(0, +\infty)$ por conta das restrições do espaço paramétrico.

Utilizou-se os algoritmos de Metropolis-Hastings implementadas pelo programa Openbugs para obtenção das cadeias MCMC de cada parâmetro.

O procedimento padronizado para na implementação de todos os modelos na função BrgusFit do pacote Brugs, foi número de iterações (nIter) 2×10^6 , número de queima (nBurnin) 1% do nIter, e número de saltos (nThin) da ordem de 100. Antes, uma análise piloto foi realizada para verificar o comportamento dos resultados.

O Acesso utilizado para o ajuste trata-se do Acesso 8 – Pimentão Rubi gigante.

3.3 RESULTADOS E DISCUSSÕES

Após o ajuste do modelo Logístico com todas as 14 distribuições de erro por meio da função BrgusFit, os resultados podem ser encontrados na Tabela 4, que mostra a média a posteriori dos parâmetros de interesse $(\alpha, \beta, \gamma, \sigma^2, \lambda, v)$ em conjunto com o DIC de cada ajuste.

Para auxiliar na visualização dos resultados da Tabela 4, utilizou-se a Figura 26, que plota os valores de cada parâmetro em conjunto com seus respectivos intervalos de credibilidade de 95%. Para verificar o comportamento distributivo de cada parâmetro, na Figura 27 encontra-se as distribuições a posteriori de cada parâmetro.

Tabela 4: Médias a posteriori dos parâmetros $(\alpha, \beta, \gamma, \sigma^2, v)$, em conjunto com os respectivos DIC's do ajuste logístico a cada erro experimental.

<i>distribuições</i>	α	β	γ	σ^2	λ	v	DIC
<i>Normal</i>	53,11	11,15	0,14	9,79	0	0	57,65
<i>Student-t</i>	53,62	11,4	0,14	7,57	0	20,91	56,03
<i>Logística</i>	53,26	11,15	0,15	2,22	0	0	58,97
<i>Laplace</i>	54,3	11,64	0,15	2,75	0	0	57,81
<i>ASN</i>	53,42	11,21	0,14	6,63	-6,03	0	33,98
<i>SSN</i>	52,13	10,57	0,17	2,73	-2	0	42,87
<i>FSSN</i>	51,72	10,45	0,18	0,72	0,16	0	45,65
<i>SDE</i>	52,39	10,49	0,19	1,45	4	0	54,52

<i>SBSL</i>	52,16	10,42	0,18	1,79	3,28	0	70,52
<i>JFSSt</i>	53,58	11,34	0,16	2,73	-0,32	0	50,09
<i>ESSt</i>	55,03	10,13	0,16	5,92	0,59	21,33	42,11
<i>FSSSt</i>	48,38	10,84	0,18	2,51	4,62	27,63	0,12
<i>SDEN</i>	55,82	9,8	0,17	2,91	-26,97	0	33,05
<i>SASSt</i>	55,67	9,71	0,16	2,3	-1,33	27,51	42,69

Fonte: O autor.

Tratando-se aos valores de α

Por se tratar de um parâmetro que é muito próximo dos valores assintóticos, não é esperado grandes discrepâncias de valores frente aos diferentes erros experimentais. No Acesso 8 após o 56° dia após o florescimento, a medida do comprimento do fruto no valor de 54.8 não se altera mais (ver Figura 13), a média dos 14 valores, se deu em 53.8. O erro FSSSt apresentou um valor discrepante de 48.38. Com auxílio da Figura 26 verifica-se que as maiores amplitudes dos intervalos de credibilidade seu deu em FSSSt e SASSt.

Tratando-se dos valores de β

A média geral de cada média a posteriori de β conforme pode ser verificado na Figura 26 está em torno de 10.74, ou seja, em média, todos os modelos estimaram que de 10 a 11 dias o fruto atingiu sua velocidade máxima de crescimento e que, consequentemente, atingiu a metade do seu crescimento máximo no valor de 26.9 (média geral de cada α que resultou em 53.8 dividido por 2).

Distribuições que estimaram o ponto de inflexão em 11 dias: Normal, Logística, Laplace, Student t, ASN, JFSSSt; Distribuições que estimaram o ponto de inflexão em aproximadamente 10 dias: SSN, FSSN, SDE, SBSL; ESSt e FSSSt; Distribuições que estimaram o ponto de inflexão em aproximadamente 9 dias: SDEN, SASSt, isso evidencia que a presença de assimetria pode mudar os resultados deste parâmetro.

Com exceção das maiores amplitudes dos intervalos de credibilidade dos modelos JFSSSt e FSSSt, no geral as amplitudes das distribuições assimétricas se mostraram menores que os modelos simétricos, revelando uma maior precisão.

Tratando-se dos valores de γ :

A média geral das médias a posteriori de γ frente a diferentes erros experimentais deu-se em torno de 0.16.

As versões assimétricas do erro experimental tenderam a influenciar as estimativas de γ . As estimativas do modelo de γ com erros simétricos resultou numa média de 0.145, enquanto a média das estimativas dos modelos assimétricos resultaram em 0.169, um aumento 16,55%. Essa tendência pode ser confirmada graficamente pela Figura 26.

Tratando-se dos valores de σ^2

A média geral das médias a posteriori de σ^2 frente a diferentes erros experimentais foi de 3.71.

Os valores de σ^2 trata-se de uma medida de variância dos modelos somente nos erros Normal, Logístico, Laplace e Student t. Nos erros assimétricos não, entretanto, ainda são uma medida que influencia em algum grau a variância residual.

Com auxílio gráfico da Figura 26, verifica-se que as distribuições assimétricas mostraram no geral alguma influência no aumento de precisão dos modelos. De acordo com a Tabela 4 o erro Normal (9.79) mostrou uma precisão inferior à suas respectivas versões assimétricas ASN (6.63), SSN (2.73) e FSSN (0.72); O erro t de Student (7.57) também apresentou baixa precisão frente as suas versões assimétricas JFSSSt (2.73), ESSSt (5.92) FSSSt (2.51) SASSt (2.3); A distribuição Logística (2.75) também apresentou precisão menor em relação a sua versão assimétrica SBSL (1.45). Entretanto somente na distribuição Laplace que se verificou aumento de precisão em relação a sua versão assimétrica.

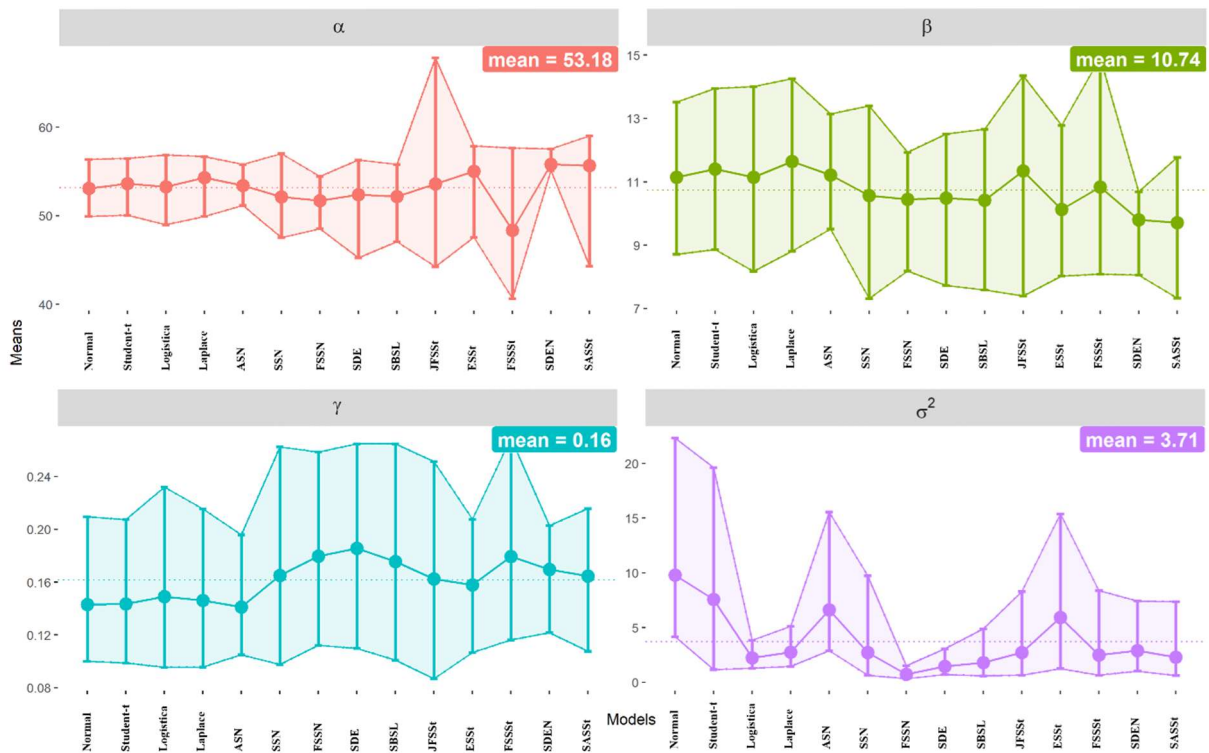


Figura 26: Médias a posteriori dos parâmetros (α , β , γ , σ^2) com seus respectivos intervalos de credibilidade considerando os 14 erros experimentais.

Fonte: O autor

Na Figura 27 verifica-se o comportamento das distribuições a posteriori dos parâmetros (α , β , γ , σ^2).

Verifica-se no item (a) diferentes precisões do parâmetro α bem como diferentes locações dessas distribuições frente aos diferentes erros experimentais. A maior precisão se deu com o erro experimental SDEN, no erro FSSSt há indícios de bimodalidade do parâmetro o que explica a discrepância da estimativa deste parâmetro em comparação aos outros erros, ao analisar o as cadeias MCMC não foi encontrado tendência em função do nIter.

No item (b) que se trata das posteriori de β , a maior precisão do parâmetro se deu com o erro SDEN; não há indícios de bimodalidade; verifica-se uma mudança de locação das densidades nas versões assimétricas do erro frente as versões simétricas evidenciando que a presença de assimetria muda os valores de β .

No item (c) que se trata das posteriori de γ , a precisão dos erros simétricos mostrou-se maior em relação aos erros assimétricos SSN, FSSN, SBSL, JFSST, entretanto, a maior precisão novamente é do erro SDEN; há indícios de bimodalidade no erro JFSSt, entretanto nas cadeias MCMC não uma tendência em função do nIter.

No item (d) que se trata do parâmetro de escala ao quadrado σ^2 , nota-se diferentes precisões (a posteriori do parâmetro de precisão $\tau = 1/\sigma^2$ também tem uma medida de precisão) frente a diferentes erros experimentais; a maior precisão se deu no erro FSSN; as precisões mais baixas podem ser visualiadas nos erros N, St, ASN e ESSt; os erros que apresentaram as precisões mais baixas também apresentaram os mais altos valores de locação. O Acesso 8 na estimativa frequentista via NLME é o de maior variância residual (em torno de 11.74), no geral as versões assimetricas do erro experimental aumentou a precisão do modelo.

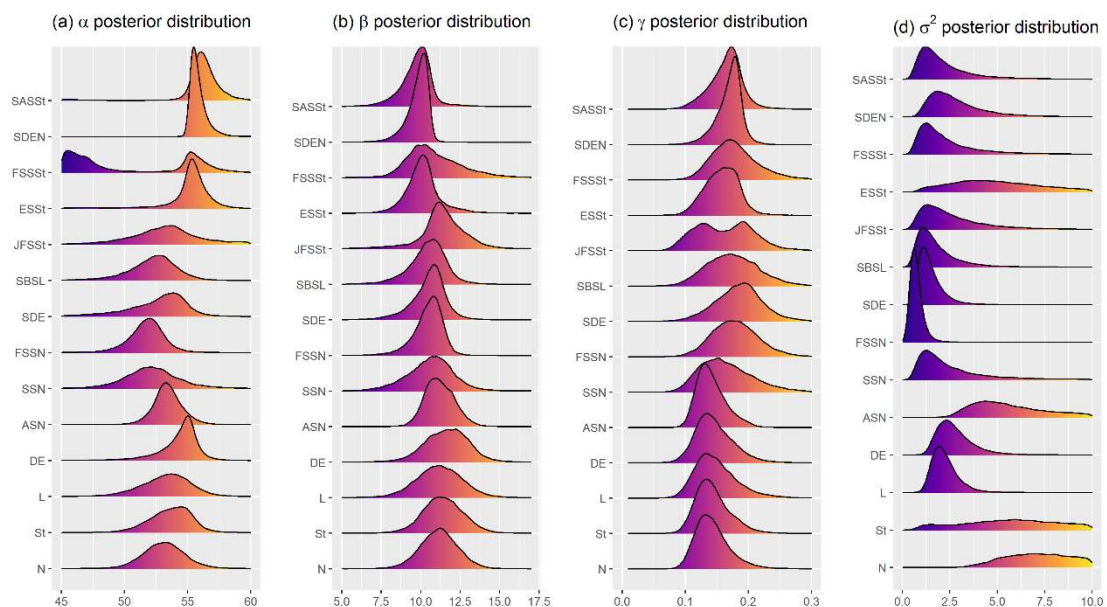


Figura 27: Distribuição a posteriori dos parâmetros $(\alpha, \beta, \gamma, \sigma^2)$ do modelo Logístico considerando os 14 erros experimentais.

Fonte: O autor.

Sobre o parâmetro de assimetria λ , na Tabela 5 encontra-se a média a posteriori junto com os intervalos de credibilidade 95%, e na Figura 28 encontra-se as suas densidades a posteriori em forma de violino.

A maioria dos modelos apresentaram assimetria à esquerda no ajuste do modelo Logístico ao Acesso 8 com exceção dos erros SDE e FSSSt, é possível obter orientações de simetria divergentes de distribuições assimétricas diferentes, no trabalho de Faccin & Rossi (2024) o erro ASN apresentou assimetria à direita no ajuste do modelo logístico com erros heterocedásticos no peso de codorna enquanto o erro *Azzalini Skew-Student t* apresentou assimetria à esquerda.

Vários erros experimentais mostraram significância estatística quanto a hipótese de assimetria nula $\mathcal{H}_0: \lambda = 0$ ou 1. O erro FSSN mostrou assimetria significativa à esquerda (intervalos não conteve 1); a distribuição SDE apresentou assimetria á direita significativa (intervalos não conteve 1); a distribuição SDEN apresentou assimetria à esquerda significativa (intervalos não conteve 0).

Apesar de algumas distribuições conterem os valores da hipótese nula, existem evidências quanto a relevância da assimetria dos dados, por exemplo, no erro ASN os intervalos $(-52,83, 0,9296)$ apresentaram forte assimetria à esquerda e o limite superior está muito próximo do valor da hipótese nula ($\lambda = 1$). O mesmo pode ser verificado no erro SBSL $(0,7208, 7,22)$ também com limite inferior próximo da hipótese nula ($\lambda = 1$).

Tabela 5: médias a posteriori dos parâmetros λ dos 10 erros experimentais assimétricos com seus respectivos intervalos de credibilidade e orientação da assimetria dos dados.

Models	mean	val2.5pc	val97.5pc	Assimetria
ASN	-6,025	-52,83	0,9296	á esquerda
SSN	-2,003	-12,41	11,8	á esquerda
FSSN	0,1557	0,07395	0,2777	á esquerda
SDE	4,002	1,122	9,209	á direita
SBSL	3,284	0,7208	7,22	á esquerda
JFSSSt	0,3204	-1,662	1,299	á esquerda
ESSt	0,5883	-0,629	0,9883	á esquerda
FSSSt	4,625	0,1706	13,01	á direita

SDEN	-26,97	-72,18	-2,384	á esquerda
SASSt	-1,331	-2,447	1,856	á esquerda

Fonte: O autor.

Na Figura 28 pode-se verificar as densidades a posteriori dos parâmetros de λ , optou-se por utilizar os gráficos individuais de violino pelas diferenças de escala. No erro ASN verifica-se evidências de caudas pesadas (outliers); verifica-se a bimodalidade nos erros SSN e FSSSt, não foi encontrado na literatura o motivo desse comportamento.

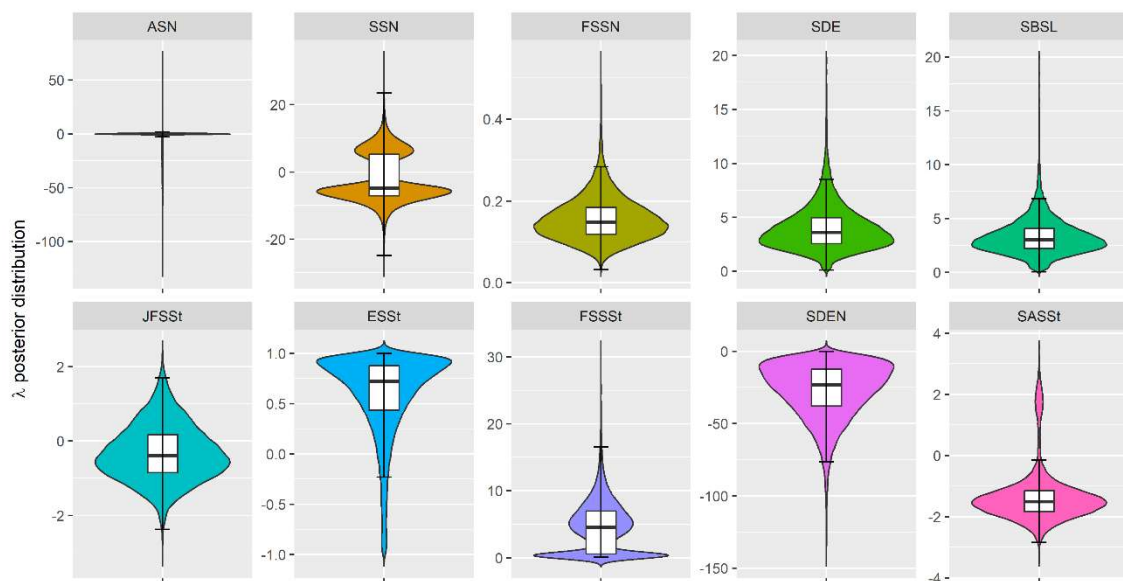


Figura 28: Gráfico de violino em conjunto com o boxplot das densidades a posteriori do parâmetro λ considerando 10 erros experimentais assimétricos.

Fonte: O autor.

Tratando-se dos valores de DIC:

Conforme verifica-se na Figura 29, os erros experimentais assimétricos providenciaram um aumento na qualidade de ajuste geral, em detrimento de seus respectivos submodelos simétricos. Quando consideramos as diferenças superiores a 5, todos os modelos assimétricos, com exceção do modelo SBSL e SDE, se mostraram com significativa diferença de qualidade de ajuste em relação aos submodelos simétricos.

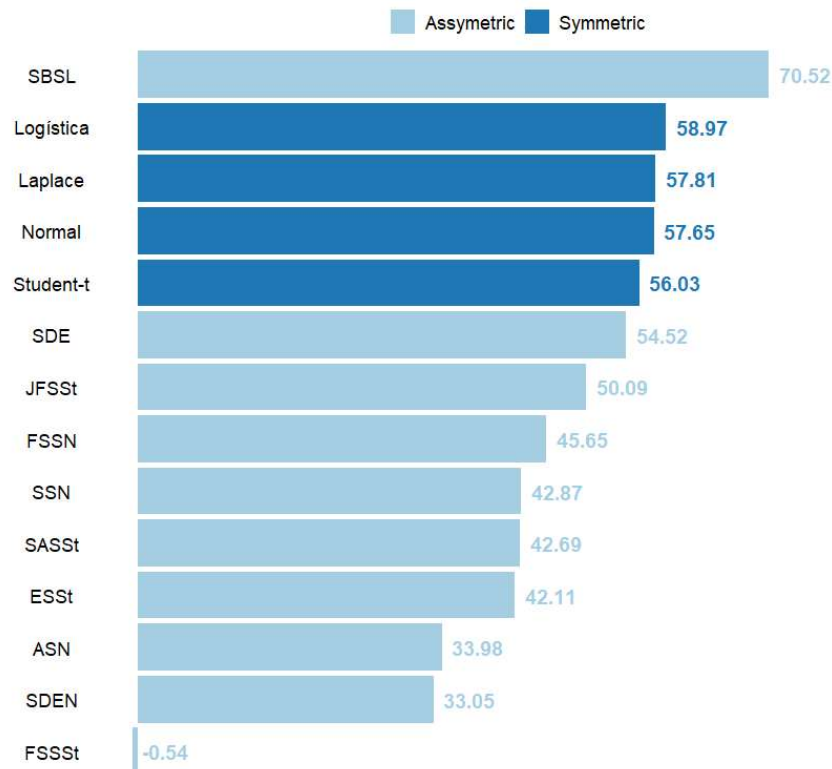


Figura 29: DIC's resultantes do ajuste do modelo Logístico considerando todas as 14 distribuições de erro experimental.

Fonte: O autor.

Tratando-se de computação:

Em termos computacionais verifica-se por meio da Figura 30, o tempo necessário que cada distribuição necessitou para calcular as cadeias MCMC. Seria esperado um tempo computacional maior das versões assimétricas pelo número maior de parâmetros: parâmetro de assimetria (λ); Parâmetros de controle de caudas (ν). Entretanto neste ajuste específico não houve tanta discrepância computacional entre os modelos simétricos e suas versões assimétricas, com exceção do erro experimental JFSSt que se mostrou discrepante ao necessitar em torno de 3 a 4 vezes mais tempo que os demais, por meio de análises piloto ficou constatado que o motivo se deveu ao uso da correção para média, sem a correção, o tempo computacional do erro JFSSt se aproxima das demais distribuições.

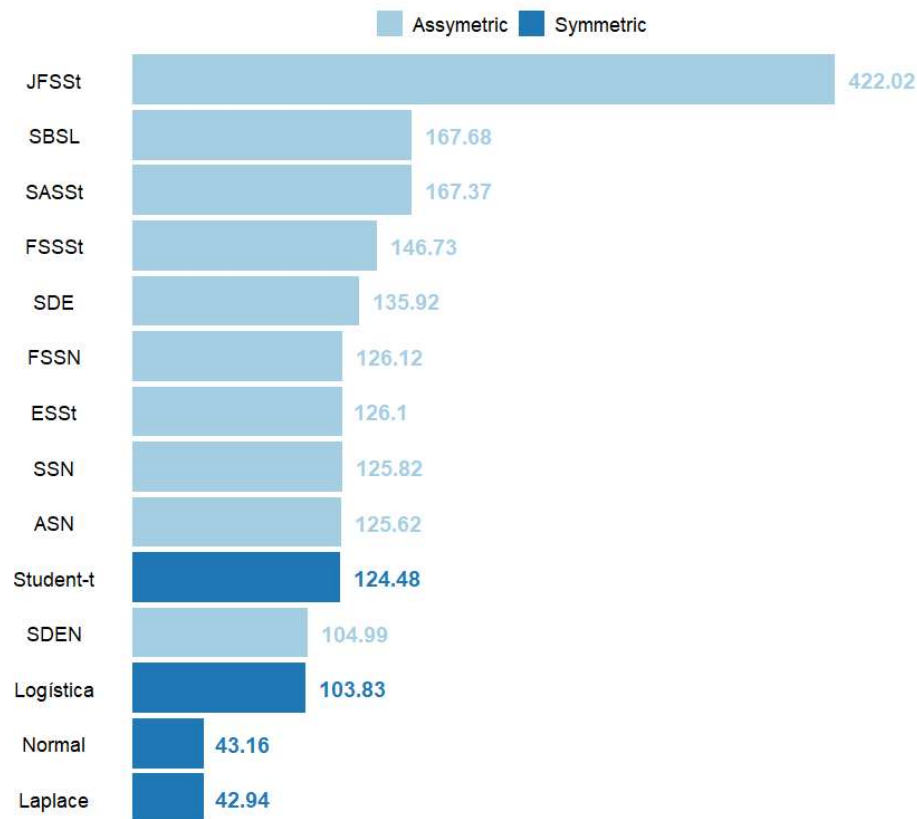


Figura 30: Tempo computacional (em segundos) necessário para implementar as 14 distribuições de erro experimental no Openbugs.

Fonte: O autor.

Ao fazer-se uma análise piloto para verificar o comportamento dos modelos, já houve uma convergência de todos os modelos submodelos simétricos, com todos passando pelos critérios de Heidelberg-Welch e Geweke por meio do pacote *coda* (v. 0.19-4.1).

O λ é o único parâmetro que se apresentou problemático, necessitando de milhões de $nIter$ – valores também obtidos por Rossi & Santos (2014) - para que passasse no teste de half-width que compõe o teste de Heidelberg-Welch. Conforme explicado por Pereira et al. (2020), o autor utiliza a metodologia Modified weighted importance resampling em alternativa ao Metropolis-Hastings MCMC para contornar o problema de $nIter$ elevado. A estratégia adotada neste trabalho para diminuir o $nIter$ consiste em utilizar estimativas próximas as frequentistas, estratégia também utilizada por Rossi & Santos (2014), a Figura 31 mostra a redução nos

valores de n_{Thin} quando se utiliza ou não, as estimativas provenientes de Modelos Mistos não lineares como valores iniciais.

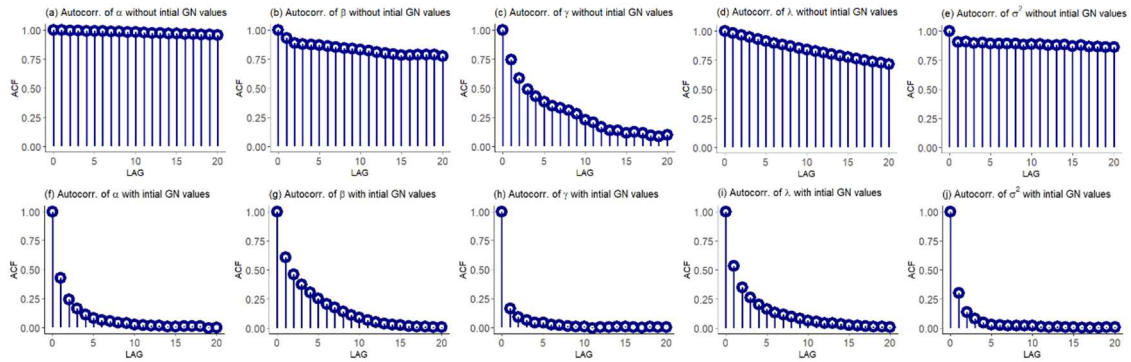


Figura 31: Autocorrelação das cadeias MCMC dos parâmetros $(\alpha, \beta, \gamma, \sigma^2, \lambda)$ considerando as estimativas frequentista como valores iniciais vs. Estimativas iniciais $(\alpha^0 = 1, \beta^0 = 1, \gamma^0 = 1, \sigma^{2^0} = 1, \lambda^0 = 1)$.

Fonte: O autor.

3.4 CONCLUSÕES

Os modelos assimétricos foram implementados satisfatoriamente e modelaram a assimetria dos dados com diferenças significativas, com ganho na precisão do modelo logístico e em estimativas confiáveis dos parâmetros dos modelos comparados aos erros simétricos.

3.5 REFERÊNCIAS

- AZZALINI, A. A Class of Distributions Which Includes the Normal Ones. **Scandinavian Journal of Statistics**. v.12, p.171–178, 1985.
- CANCHO, V.G. *et al.* A nonlinear regression model with skew-normal errors. **Stat Papers**. v.51, p.547-558, 2008.
- Castillo, N. O., Gómez, H. W., Leiva, V., Sanhueza, A. On the Fernández–Steel distribution: Inference and application. **Computational Statistics & Data Analysis**, v.55, p.2951-2961, 2011.
- Cordeiro, G. M., Demétrio, C. G. B. **Modelos lineares generalizados e extensões**, 1. ed. Recife, PE: [s.n.], 2013.
- DE LA CRUZ, R., BRANCO, M. Bayesian analysis for nonlinear regression model under skewed errors, with application in growth curves. **Biometrical Journal**, Weinheim, v.51, p.588-609, 2009.
- DOĞRU, F.Z., ARSLAN, O. Finite mixtures of skew Laplace normal distributions with random skewness. **Comput Stat** v.36, p.423–447, 2021.
- ELOY, R. A. O. *et al.* Distribuição assimétrica t-Student tipo 3: Uma aplicação a delineamentos inteiramente casualizados. **SAJEETT**, v.6, p.55-77, 2020.
- FACCIN, M.Z. AND ROSSI, R.M. Bayesian modeling of the Gompertz curve for meat quails growth data considering different error distributions. **Brazilian Journal of Biometrics**. v.42, p.3 260–271, 2024.
- FERNANDES, T. J., MUNIZ, J. A., PEREIRA, A. A., MUNIZ, F. R.; MUIANGA, C. A. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum. Technology**, v.37, p.397-402, 2015.
- FERNANDES, F. A., SILVA, ÉDIPO M., LIMA, K. P., JANE, S. A., FERNANDES, T. J., MUNIZ, J. A. Parameterizations of the von bertalanffy model for description of growth curves. **brazilian journal of biometrics**, v.38, p.369–384, 2020.
- FIRAT, MZ *et al.* Bayesian Analysis for the Comparison of Nonlinear Regression Model Parameters: An Application to the Growth of Japanese Quail. **Revista Brasileira de Ciência Avícola**, v.18, p.19-26, 2016.
- GELMAN, A., CARLIN, J.B., STERN, H.S., DUNSON, D.B., VEHTARI, A., & RUBIN, D.B. **Bayesian Data Analysis**. Chapman and Hall/CRC, 2013.
- GUEDES, T.A. *et al.* Applying regression models with skew-normal errors to the height of bedding plants of *Stevia rebaudiana* (Bert) Bertoni. **Acta Scientiarum Technology**, v.36, p.463-468, 2014.

- GHADERINEZHAD, F. *et al.* Bayesian Inference for Skew-Symmetric Distributions. **Symmetry**, v.12, 2020.
- GUJARATI, D. N., PORTER, D.C. **Econometria básica**. AMGH, 2012.
- HENZE, N. A Probabilistic Representation of the 'Skew-normal' Distribution. **Scandinavian Journal of Statistics**, v.13, p.271-275, 1986.
- KOYA, P.R., GOSHU, A.T. Generalized Mathematical Model for Biological Growths. **Open Journal of Modelling and Simulation**, v.1, p.42-53, 2013.
- LOUIS, THOMAS A., CARLIN, BRADLEY P. **Bayesian Methods for Data Analysis**. 2 ed. United States: CRC Press, 2008.
- LUNA, J.G., OLINDA, R.A. **Introdução a modelos lineares**. São Paulo. Livraria da física, 2014.
- MACEDO, L. R. *et al.* Bayesian inference for the fitting of dry matter accumulation curves in garlic plants. **Pesquisa Agropecuária Brasileira**, v.52, p.572-581, 2017.
- MANGUEIRA, R. A. F., SAVIAN, T. V., MUNIZ, J. A., SERMARINI, R. A., CROSARIOL N, J. Logistic model considering different error distributions applied in maize height data. **Brazilian Journal of Biometrics**. v.34, p.317–333, 2016.
- MARTINS FILHO, S. *et al.* Abordagem Bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, v.38, p.1516-1521, 2008.
- NADARAJAH, S. The skew logistic distribution. **AStA Adv Stat Anal**, p.187-203, 2009.
- NELDER, J. A., WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**, v.135, p.370-384, 1972.
- NTZOUFRAS, I. **Bayesian Modeling Using WinBUGS**. 1 ed. Wiley, 2008.
- NUNES ÁVILA, M. DE S., BARBOSA, J. M. Análise de crescimento de pimenta-biquinho em diferentes níveis de radiação solar. **Brazilian Journal of Development**, v.5, p.31985-31997, 2019.
- PÉREZ-RODRÍGUEZ, P. *et al.* A Bayesian Genomic Regression Model with Skew Normal Random Errors. **G3 Genes|Genomes|Genetics**, v.8, p.1771-1785, 2018.
- PINO, F.A. A questão da não normalidade: Uma revisão. **Rev. de Economia Agrícola**, v.61, p.17-33, 2014.
- ROSA, M. *et al.* Linear and nonlinear models to describe the lactation curve of Girolando cows. **Pesquisa Agropecuária Brasileira**, v.57, p.2678, 2022.

ROSSI, R.M., SANTOS, L.A. Modelagem Bayesiana para curvas de crescimentos de codornas assumindo assimetria nos erros. **Semina: Ciências Agrárias**, v.35, p.1637-1647, 2014.

SAHU, S.K., DEY, D.K. AND BRANCO, M.D. A new class of multivariate skew distributions with applications to bayesian regression models. **Can J Statistics**, v.31, p.129-150, 2003.

SOUZA, G.S. **Introdução aos modelos de Regressão linear e não linear**. Brasília: Embrapa, 1998.

TEIXEIRA, F.R.F., CECON, P.R., SUELA, M.M., NASCIMENTO, M. Nonlinear Mixed-Effect Models to Describe Growth Curves of Pepper Fruits in Eight Cultivars Including Group Effects. **Agronomy**, v.12, 2023.

TROTTIER, D.A., ARDIA, D. Moments of standardized Fernandez–Steel skewed distributions: Applications to the estimation of GARCH-type models. **Finance Research Letters**. v. 18, p. 311-316, 2016.