



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

**Um Novo Modelo de Regressão
Birnbaum–Saunders Generalizado:
Uma Abordagem via GAMLSS**

por

Matheus Braga Milhomem

Brasília, 30 de janeiro de 2026

Um Novo Modelo de Regressão Birnbaum–Saunders Generalizado: Uma Abordagem via GAMLSS

por

Matheus Braga Milhomem

Dissertação apresentada ao Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para qualificação no Mestrado em Estatística.

Orientadora: Terezinha Késsia de Assis Ribeiro

Brasília, 30 de janeiro de 2026

Dissertação submetida ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade de Brasília como parte dos requisitos para a obtenção do título de Mestre em Estatística.

Texto aprovado por:

Profa. Dra. Terezinha Késsia de Assis Ribeiro
Orientadora, EST/UnB

Prof. Dr. Felipe Sousa Quintino
Membro interno, EST/UnB

Profa. Dra. Michelli Karinne Barros da Silva
Membro externo, UAEst/UFCG

Agradecimentos

Agradeço aos meus pais Solange e Abdias, que sempre apoiaram as minhas decisões e me deram suporte em todos os momentos, nunca poupando esforços para me auxiliar nessa jornada.

À minha professora e orientadora, Terezinha Késsia de Assis Ribeiro, por toda a paciência e dedicação, sua orientação e preocupação foram fundamentais durante todo o processo. Tenho muito orgulho e gratidão por este momento.

À Camila, por todo apoio, amor e carinho durante a etapa final deste mestrado.

Aos amigos e colegas do mestrado pelo companheirismo. Em especial, à Talia, pela parceria e apoio durante este período tão importante para minha formação acadêmica e profissional.

Aos professores e servidores do Departamento de Estatística da Universidade de Brasília pela competência e profissionalismo ao longo dos anos de graduação e mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Um Novo Modelo de Regressão Birnbaum–Saunders Generalizado: Uma Abordagem via GAMLSS

A distribuição Birnbaum-Saunders (BS) é um modelo probabilístico com crescente destaque na literatura de modelagem de dados contínuos positivos e assimétricos. Diversos autores propuseram extensões deste modelo probabilístico, e alguns destes desenvolveram estruturas de regressão para a modelagem dos parâmetros. Neste trabalho será proposto uma nova classe de modelos de regressão para a generalização da distribuição BS proposta por Owen (2006). A distribuição Birnbaum-Saunders Generalizada (BSG) que será utilizada não possui tanto destaque na literatura, apesar de reforçar a justificativa física do modelo originalmente derivado. A configuração do novo modelo de regressão proposto apresentará especificação semelhante aos Modelos Aditivos Generalizados de Localização, Escala e Forma (GAMLSS), fornecendo assim uma estrutura mais flexível para o ajuste dos dados. A estimação dos parâmetros é realizada por meio do método de máxima verossimilhança, cujo procedimento inferencial é construído a partir da formulação da função de verossimilhança e do cálculo dos vetores escore associados aos parâmetros do modelo. O modelo BSG foi implementado computacionalmente de maneira inédita no ambiente R, por meio do pacote `gamlss`, possibilitando o uso de um conjunto abrangente de ferramentas destinadas ao ajuste e à avaliação do novo modelo proposto. Estudos de simulação de Monte Carlo foram conduzidos com o objetivo de avaliar o comportamento e desempenho do procedimento de estimação sob o novo modelo de regressão BSG. A partir dos resultados obtidos, observou-se bom desempenho do estimador de máxima verossimilhança sob o novo modelo, evidenciado pela redução dos vieses e pelo aumento da eficiência à medida que o tamanho amostral cresce. Por fim, a aplicabilidade do modelo proposto foi ilustrada e discutida por meio do ajuste a dados reais, nos quais se evidenciaram, de forma empírica, situações em que o modelo BSG proporcionou ajustes mais adequados aos dados quando comparado ao modelo BS original.

Palavras-chave: Distribuição Birnbaum-Saunders generalizada; Modelos de regressão; Simulação de Monte Carlo; Estimador de Máxima Verossimilhança; GAMLSS.

Abstract

A New Generalized Birnbaum–Saunders Regression Model: A GAMLSS-Based Approach

The Birnbaum–Saunders (BS) distribution is a probabilistic model that has gained increasing prominence in the literature for modeling positive and asymmetric continuous data. Several authors have proposed extensions of this probabilistic model, and some have developed regression frameworks for modeling its parameters. In this study, a new class of regression models is proposed for the generalization of the BS distribution introduced by Owen (2006). The Generalized Birnbaum–Saunders (GBS) distribution considered herein has received relatively limited attention in the literature, despite reinforcing the physical justification of the originally derived model. The specification of the proposed regression model is analogous to that of the Generalized Additive Models for Location, Scale, and Shape (GAMLSS), thereby providing a more flexible framework for data fitting. Parameter estimation is carried out via the maximum likelihood method, with the inferential procedure constructed from the formulation of the likelihood function and the derivation of the associated score vectors for the model parameters. The GBS model was implemented computationally in a novel manner within the R environment through the `gamlss` package, enabling the use of a comprehensive set of tools for fitting and assessing the proposed model. Monte Carlo simulation studies were conducted to evaluate the behavior and performance of the estimation procedure under the proposed GBS regression model. The results indicate good performance of the maximum likelihood estimator, as evidenced by decreasing bias and increasing efficiency as the sample size increases. Finally, the applicability of the proposed model is illustrated and discussed through applications to real data, empirically demonstrating scenarios in which the GBS model provides a better fit to the data when compared with the original BS model.

Keywords: Generalized Birnbaum–Saunders distribution; Regression models; Generalized linear models; Maximum likelihood estimator; GAMLSS.

Sumário

1	Introdução	8
2	Revisão de Literatura	12
2.1	Distribuição Birnbaum-Saunders	12
2.2	Distribuição Birnbaum-Saunders Tri-paramétrica	17
3	Modelo de Regressão Birnbaum-Saunders Generalizado	23
3.1	Estrutura de Regressão	23
3.2	Inferência por Máxima Verossimilhança	24
3.3	Técnicas de Diagnóstico sob os GAMLSS	30
3.3.1	Resíduo Quantílico	31
3.3.2	<i>Worm Plot</i>	31
3.3.3	Gráficos de Diagnóstico sob os GAMLSS	33
4	Estudos de Simulação	35
5	Aplicações	51
5.1	Dados <i>Cheese</i>	51
5.2	Dados <i>Dental</i>	57
6	Conclusão	64

Lista de Tabelas

1	Interpretação do ajuste com base no formato do <i>worm plot</i>	33
5	Níveis empíricos sob H_0 do teste tipo-Wald para o Cenário 6 ao nível nominal de 5%.	48
2	Níveis empíricos sob H_0 do teste tipo-Wald nos Cenários 1.1, 1.2, 1.3, 2.1, 2.2 e 2.3 ao nível nominal de 5%.	49
3	Níveis empíricos sob H_0 do teste tipo-Wald dos Cenários 3 e 4 ao nível nominal de 5%.	49
4	Níveis empíricos sob H_0 do teste tipo-Wald para o Cenário 5 ao nível nominal de 5%.	49
6	Percentual de réplicas descartadas de acordo com os cenários de simulação.	50
7	Estimativas, erros-padrão, estatísticas- z e valores- p para os coeficientes de regressão estimados – dados <i>Cheese</i>	53
8	Medidas de informação – dados <i>Cheese</i>	57
9	Estimativas, erros-padrão, estatísticas- z e valores- p para os coeficientes de regressão estimados – dados <i>Dental</i>	59
10	Medidas de informação – dados <i>Dental</i>	63

Lista de Figuras

1	Curvas da função densidade da distribuição BS para $\sigma = 0.5$ com valores indicados de μ (a) e $\mu = 1$ com valores indicados de σ (b).	16
2	Curvas da função densidade da distribuição BSG para $\sigma = 0,5, \nu = 0,8$ para valores indicados de μ (a), $\sigma = 0,5, \mu = 0,1$ para valores indicados de ν (b) e $\mu = 1, \nu = 0,8$ para valores indicados de σ (c).	22
3	Exemplo de um <i>worm plot</i> , Stasinopoulos et al. (2017).	32
4	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 1.1. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	39
5	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 1.2. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	40
6	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 1.3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	41
7	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 2.1. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	41
8	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 2.2. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	42
9	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 2.3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	42
10	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2$ sob o Cenário 3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	44
11	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \gamma_2, \lambda_1, \lambda_2$ sob o Cenário 4. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	45
12	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2$ sob o Cenário 5. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	46
13	<i>Boxplots</i> das estimativas dos parâmetros $\beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2$ sob o Cenário 6. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.	47
14	Gráfico de dispersão da variável Taste em função da variável Acetic (a), gráfico de dispersão da variável Taste em função da variável $\log(H_2S)$ (b), gráfico de dispersão da variável Taste em função da variável Lactic (c) e <i>boxplot</i> da variável resposta Taste (d) – dados <i>Cheese</i>	52

15	Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BS – dados <i>Cheese</i>	54
16	<i>Worm plot</i> do ajuste do modelo baseado na distribuição BS – dados <i>Cheese</i> . 54	
17	Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BSG– dados <i>Cheese</i>	55
18	<i>Worm plot</i> do ajuste do modelo baseado na distribuição BSG – dados <i>Cheese</i> . 56	
19	<i>Boxplot</i> da variável DFMT em função da variável Indus (a), gráfico de dispersão da variável DFMT em função da variável Sugar (b) e <i>Boxplot</i> da variável resposta DFMT – dados <i>Dental</i>	58
20	Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BS – dados <i>Dental</i>	60
21	<i>Worm plot</i> do ajuste do modelo baseado na distribuição BS – dados <i>Dental</i> . 60	
22	Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BSG – dados <i>Dental</i>	61
23	<i>Worm plot</i> do ajuste do modelo baseado na distribuição BSG – dados <i>Dental</i> . 62	

Lista de Siglas e Abreviaturas

BS Birnbaum-Saunders

BSG Birnbaum-Saunders Generalizada

CG Algoritmo de Cole e Green

EMV Estimador de Máxima Verossimilhança

FDA Função de Distribuição Acumulada

FDP Função Densidade de Probabilidade

GAMLSS *Generalized Additive Models for Location, Scale and Shape*

MLG Modelos Lineares Generalizados

RS Algoritmo de Rigby e Stasinopoulos

TLC Teorema do Limite Central

1 Introdução

A distribuição de probabilidade Birnbaum-Saunders (BS) tem recebido bastante atenção na literatura ao longo dos anos devido às suas propriedades e à sua justificativa física relacionada à Regra de Miner. Essa é uma regra bastante conhecida na engenharia, que trata a fadiga como um fenômeno determinístico e utiliza conceitos subjetivos, como “dano” e “energia interna”, para descrever o seu processo, conforme discutido por Birnbaum e Saunders (1968) anteriormente a formulação da distribuição BS. Entende-se como processo de fadiga ou desgaste o dano causado em um material ou equipamento devido ao uso constante ou exposição a condições de estresse e tensão. Dessa forma, a falha ocorre quando o dano acumulado atinge um limite de resistência crítico.

A derivação da distribuição BS proposta por Birnbaum e Saunders (1969), é uma interpretação probabilística da Regra de Miner. Embora a distribuição BS tenha sido inicialmente desenvolvida para avaliar processos de fadiga de materiais, ela tem conquistado crescente popularidade em diferentes áreas de estudo, como negócios, meio ambiente, saúde e vários outros assuntos multidisciplinares, como informa DaSilva et al. (2020). Por exemplo, Leiva et al. (2014a) utilizaram a distribuição BS para modelar a quantidade de proteína necessária em dietas restritas de pacientes adultos hospitalizados. Leiva et al. (2016) baseados na distribuição BS de valor extremo, ajustaram um modelo de regressão para modelar a concentração máxima diária de ozônio no ambiente. No contexto da saúde, Leão et al. (2018) propuseram um modelo de sobrevivência com fragilidade baseado na distribuição BS, os autores aplicaram o modelo proposto para ajustar o tempo de vida de pacientes com leucemia e de pacientes com câncer no pulmão. Sánchez et al. (2021) ajustaram um modelo de regressão quantílica baseado na distribuição BS a fim de modelar renda familiar no Chile.

A distribuição BS é um modelo probabilístico relacionado a variáveis aleatórias contínuas assimétricas que assumem valores positivos. Em sua formulação original, essa distribuição possui dois parâmetros que determinam a forma e a escala da distribuição. Segundo Leiva (2015), uma variável aleatória que segue uma distribuição BS pode ser obtida por meio de uma transformação de uma variável aleatória normal padrão. Com base nesse princípio, diversos modelos de probabilidade baseados na distribuição BS foram desenvolvidos a partir dessa relação.

Na literatura existem diversas generalizações na distribuição BS. Desmond (1985) propôs uma generalização da distribuição BS baseado num modelo biológico, ampliando a justificativa física do modelo ao flexibilizar a suposição de independência entre os ciclos imposta por Birnbaum e Saunders (1969). Desmond (1985) desenvolveu uma estrutura de dependência entre os ciclos, ampliando assim, a aplicabilidade da distribuição a mais

situações. Owen e Padgett (1999) desenvolveram uma generalização da distribuição BS adicionando um terceiro parâmetro associado ao tamanho do material ou do equipamento. Segundo os autores, o tamanho do equipamento afetaria a sua resistência, suportando assim um número maior ciclos de estresse e tensão. Díaz-García e Leiva (2002) propuseram uma extensão da distribuição BS a partir da classe de distribuições elípticas. Essa família compõe um grupo de distribuições multivariadas cujas densidades apresentam formas elípticas semelhantes à distribuição normal. Segundo os autores, essa generalização baseia-se na busca de distribuições que cresçam mais rapidamente e possuam caudas com menor ou maior curtose do que a distribuição BS. Diferentemente de sua formulação clássica, essa generalização não possui uma justificativa física. Uma formulação não central da distribuição BS foi proposta por Guiraud, Leiva e Fierro (2009). Nesta versão, os autores assumem que a extensão da fissura em cada ciclo tem uma média não constante. Lemonte (2013) propôs uma nova classe de distribuições tri-paramétrica que generaliza a distribuição BS, fundamentando-se no esquema introduzido por Marshall e Olkin (1997).

De forma particular, esta dissertação dedica-se ao estudo da generalização da distribuição BS proposta por Owen (2006), denominada Birnbaum–Saunders triparamétrica. Essa formulação da distribuição BS introduz um terceiro parâmetro à distribuição com o objetivo de flexibilizar a suposição de independência entre os ciclos de estresse aos quais um material ou equipamento é submetido. Ao incorporar uma estrutura de dependência entre os ciclos, Owen (2006) reforça a fundamentação física do modelo, tornando-o mais realista e ampliando a aplicabilidade do modelo.

Modelagens com estrutura de regressão baseadas na distribuição BS têm se destacado na literatura como uma alternativa às distribuições de probabilidade mais tradicionais. Rieck e Nedelman (1991) formularam e desenvolveram um modelo log-linear para a distribuição BS. Para viabilizar a construção e inferência do modelo os autores provaram que o logaritmo de uma variável aleatória que segue uma distribuição BS é um caso particular da distribuição normal seno-hiperbólica. Lemonte e Cordeiro (2009) desenvolveram uma nova classe de modelos de regressão não lineares para a distribuição BS, essa classe generaliza o modelo introduzido por Rieck e Nedelman (1991). Leiva et al. (2016) desenvolveram uma metodologia estatística para o ajuste de um modelo de regressão baseado na distribuição BS de valor extremo. Essa abordagem considera a formulação do modelo, estimação dos parâmetros via o método de máxima verossimilhança e análise de resíduos para avaliar a sua adequação. Em relação à modelos baseados em generalizações da distribuição Birnbaum-Saunders, destacam-se os trabalhos de Silva, Paula e Leiva (2008) e Leiva et al. (2008).

Em estudos mais recentes, Sánchez (2020) e Valentim da Silva (2022) propuseram modelos de regressão quantílica baseados na distribuição BS. O primeiro estabeleceu uma nova parametrização da distribuição BS em termos de quantis e, então, desenvolveu

uma metodologia completa que inclui a estimação dos parâmetros, inferência assintótica dos estimadores, análise de diagnóstico em função da influência local, análise de resíduos e estudos de simulação para o caso de observações independentes. Nesta mesma tese, Sánchez (2020) também desenvolveu modelos de regressão quantílica espacial incorporando dependência espacial por meio de uma estrutura de covariância, com base em uma nova parametrização da distribuição Birnbaum-Saunders multivariada. Seguindo essa mesma linha, Valentim da Silva (2022) propôs um modelo de regressão quantílica que se baseia em uma parametrização da generalização da distribuição BS apresentada por Owen (2006). O modelo proposto por Sánchez (2020) é um caso particular da formulação apresentada por Valentim da Silva (2022).

Em especial, os trabalhos desenvolvidos por Santos-Neto et al. (2012) analisou as propriedades de onze parametrizações distintas da distribuição BS, investigou o EMV dos parâmetros e conduziu estudos de simulação para avaliar o desempenho desses estimadores com base em suas propriedades. Dentre as onze parametrizações investigadas, os autores propuseram nove delas. Com destaque para a parametrização que descreve a média de uma variável aleatória BS, sem a necessidade de aplicar uma transformação, preservando a escala original da variável, uma abordagem até então inédita na literatura. Com base nessa parametrização, Leiva et al. (2014b) formularam uma nova abordagem para modelos de regressão baseados na distribuição BS, seguindo uma estrutura similar aos MLG desenvolvido por Nelder e Wedderburn (1972). Os autores também conduziram estudos de simulação a fim de avaliar as propriedades empíricas dos estimadores. Generalizando este modelo, Santos-Neto et al. (2016) formularam um modelo mais flexível, incorporando uma estrutura de regressão também no parâmetro de precisão, especificação similar aos MLG Duplos. De forma adicional, os autores apresentaram testes de hipótese para o parâmetro de precisão, bem como quatro tipos de resíduos para avaliar o ajuste do modelo quando ambos os parâmetros são modelados.

Conforme o conhecimento adquirido na revisão bibliográfica, observa-se que, até o presente momento e até onde se tem conhecimento, não foram identificados na literatura modelos de regressão que utilizem diretamente a distribuição BS proposta por Owen (2006) como distribuição condicional para a variável resposta. Essa limitação identificada na revisão de literatura reforça a relevância do novo modelo de regressão proposto nesta dissertação.

O objetivo principal desta dissertação é desenvolver uma nova classe de modelo de regressão baseando-se na distribuição Birnbaum-Saunders Tri-paramétrica desenvolvida por Owen (2006), incorporando uma estrutura semelhante à dos Modelos Aditivos Generalizados de Localização, Escala e Forma introduzido por Rigby e Stasinopoulos (2005). A inferência dos parâmetros será desenvolvida com base no método de máxima verossimilhança. As propriedades, bem como o comportamento, destes estimadores serão avaliadas

por meio de estudos de simulação. Por fim, a aplicabilidade do novo modelo de regressão proposto será ilustrada por meio de dois conjuntos de dados reais, seguindo a metodologia adotada no contexto dos GAMLSS. A principal contribuição desta dissertação consiste na formulação desse novo modelo de regressão, bem como em sua implementação computacional inédita na disponibilização de todos os códigos em R no repositório Github, permitindo a replicação dos resultados obtidos e para facilitar a aplicação para usuários interessados. Todos os cálculos e avaliações realizados ao longo do desenvolvimento deste trabalho, incluindo as aplicações e estudos de simulação de Monte Carlo, foram implementados computacionalmente com base na arquitetura da biblioteca `gamlss`, proposta por Rigby e Stasinopoulos (2005), disponível no ambiente R.

O presente texto está estruturado em 6 capítulos. No Capítulo 1, é apresentada uma introdução ao tema, destacando sua relevância e referenciando estudos que justificam o desenvolvimento de um modelo de regressão para a distribuição proposta por Owen (2006). No Capítulo 2 são revisados as suposições, propriedades e o comportamento da distribuição Birnbaum-Saunders clássica, bem como a generalização proposta por Owen (2006). No Capítulo 3, apresenta-se a estrutura do novo modelo de regressão proposto, juntamente com as principais características dos GAMLSS. Nesse capítulo, descreve-se também o procedimento inferencial do novo modelo, incluindo o desenvolvimento do estimador de máxima verossimilhança para os parâmetros da distribuição Birnbaum-Saunders tri-paramétrica, bem como as derivadas necessárias para a obtenção das funções escores destes estimadores. Além disso, são discutidas as técnicas de diagnóstico utilizadas no contexto dos GAMLSS. No Capítulo 4 são apresentados os resultados obtidos nos estudos de simulação de Monte Carlo. No Capítulo 5, são discutidas as aplicações a dados reais do novo modelo de regressão BSG proposto, bem como de seu caso particular, a distribuição BS clássica. Por fim, no Capítulo 6, são apresentadas as considerações finais deste trabalho, incluindo limitações observadas, bem como sugestões de tópicos que podem ser aperfeiçoados em trabalhos futuros.

2 Revisão de Literatura

Neste capítulo serão apresentadas a distribuição BS original derivada por Birnbaum e Saunders (1969) e a generalização da distribuição BS proposta por Owen (2006).

2.1 Distribuição Birnbaum-Saunders

De maneira geral, segundo Leiva (2015), a fadiga consiste no processo de deterioração de um material ou equipamento decorrente da exposição contínua a condições de estresse e tensão.

Para uma melhor compreensão da formulação da distribuição BS e do processo de fadiga, Leiva (2015) estabelece conceitos como fissura, ciclo e cargas de estresse e tensão. Dessa maneira, considere um material ou equipamento submetido a uma sequência de carga de estresse e tensão durante um período de tempo, essa sequência é denominada ciclo. Em cada ciclo, uma quantidade de dano é imposta ao material, levando ao desenvolvimento de uma fissura no equipamento. Como suposição, a sequência de estresse e tensão imposta ao material é igual em cada ciclo, em outras palavras, essa sequência de estresse pode ser interpretada como uma função contínua unimodal aplicada em um intervalo de tempo. Assim, após um número finito de ciclos, a falha do material acontece quando a fissura ultrapassa um limite de resistência crítico.

Distribuições de probabilidade amplamente utilizadas, como a Exponencial, Weibull e Gama, são frequentemente empregadas para modelar dados assimétricos que assumem valores positivos. Embora essas distribuições apresentem um bom ajuste na região central, segundo Silva (2007) sua adequação nos percentis extremos tende a ser insatisfatória.

Com base em seus conhecimentos do processo de fadiga observado em aeronaves comerciais novas, Birnbaum e Saunders (1969) formularam as seguintes suposições para a derivação da distribuição BS:

1. Um material é exposto a oscilações cíclicas de estresse;
2. A falha do material ocorre devido ao desenvolvimento e crescimento de uma fissura dominante dentro do material, causada pelos ciclos de estresse. Assim, a falha ocorre quando o tamanho da fissura ultrapassa um certo nível de resistência, denotado por ω ;
3. O padrão de estresse (ou tensão) submetido ao material ou equipamento é o mesmo em cada um dos ciclos;

4. A extensão incremental da fissura X_i , resultado da i -ésima oscilação de estresse, é uma variável aleatória com distribuição que só depende da fissura atual causadas pelas oscilações de tensão anteriores neste mesmo ciclo;
5. A extensão da fissura durante o $(j + 1)$ -ésimo ciclo é

$$Y_{j+1} = X_{jm+1} + \dots + X_{jm+m}, \quad \text{para } j = 0, 1, \dots, \quad (2.1.1)$$

em que X_{jm+1} é a extensão da fissura após a i -ésima oscilação de estresse do $(j + 1)$ -ésimo ciclo;

6. As extensões das fissuras em ciclos distintos são independentes;
7. A extensão total da fissura, Y_j , devido ao j -ésimo ciclo é uma variável aleatória com média μ^* e variância α^2 , $\forall j = 1, 2, \dots$

A extensão total da fissura após n ciclos de estresse é definida pela variável aleatória

$$W_n = \sum_{j=1}^n Y_j, \quad (2.1.2)$$

com função de distribuição acumulada (FDA) expressa por

$$H_n(\omega) = P(W_n \leq \omega), \quad \text{para } n = 1, 2, \dots$$

Seja N o número de ciclos necessários até que a falha seja observada. A função de distribuição da variável aleatória N é dada por

$$P(N \leq n) = P\left(\sum_{j=1}^n Y_j > \omega\right) = P(W_n > \omega) = 1 - H_n(\omega).$$

Supondo que as variáveis aleatórias Y_j são independentes e identicamente distribuídas (Suposições 6 e 7) segue do Teorema do Limite Central (TLC), discutindo em Casella e Berger (2002), que

$$\begin{aligned} P(N \leq n) &= P\left(\frac{\sum_{j=1}^n Y_j - n\mu^*}{\alpha\sqrt{n}} > \frac{\omega - n\mu^*}{\alpha\sqrt{n}}\right) \\ &= 1 - P\left(\frac{\sum_{j=1}^n Y_j - n\mu^*}{\alpha\sqrt{n}} \leq \frac{\omega - n\mu^*}{\alpha\sqrt{n}}\right) \end{aligned}$$

$$\begin{aligned}
&\approx 1 - \Phi\left(\frac{\omega - n\mu^*}{\alpha\sqrt{n}}\right) \\
&= 1 - \Phi\left(-\frac{n\mu^* - \omega}{\alpha\sqrt{n}}\right) \\
&= \Phi\left(\frac{\mu^*\sqrt{n}}{\alpha} - \frac{\omega}{\alpha\sqrt{n}}\right), \tag{2.1.3}
\end{aligned}$$

em que $\Phi(\cdot)$ representa a FDA da distribuição normal padrão.

Birnbaum e Saunders (1969) utilizaram a FDA expressa em (2.1.3) para definir um modelo probabilístico contínuo. Segundo os autores, substituindo N por uma variável aleatória não negativa Y , tem-se que Y é a extensão contínua da variável aleatória discreta N . Assim, Y representa o tempo até a falha do material ou equipamento.

Considerando a reparametrização

$$\sigma = \frac{\alpha}{\sqrt{\mu^*\omega}} > 0, \quad \mu = \frac{\omega}{\mu^*} > 0,$$

a FDA de Y fica expressa por

$$F(y; \mu, \sigma) = \Phi\left[\frac{1}{\sigma}\left(\sqrt{\frac{y}{\mu}} - \sqrt{\frac{\mu}{y}}\right)\right], \quad y > 0, \tag{2.1.4}$$

sendo $\mu > 0$ e $\sigma > 0$. Se Y possui FDA expressa por (2.1.4), então Y segue uma distribuição Birnbaum-Saunders com parâmetros μ e σ , e é denotada por $Y \sim \text{BS}(\mu, \sigma)$.

Derivando (2.1.4) em relação a y , obtém-se a função densidade probabilidade de Y expressa por

$$f(y; \mu, \sigma) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{y}{\mu} + \frac{\mu}{y} - 2\right)\right\} \frac{y^{-3/2}(\mu + y)}{2\sigma\sqrt{\mu}}, \quad y > 0, \tag{2.1.5}$$

com $y, \mu, \sigma > 0$.

Como descreve Silva (2007), o parâmetro σ determina a forma da distribuição. A medida que σ aumenta, a distribuição BS torna-se assimétrica, e torna-se simétrica em torno de μ a medida que σ se aproxima de zero. A simetria ocorre em torno do parâmetro de escala μ , que representa a mediana da distribuição, como é demonstrado a seguir:

$$\begin{aligned}
F(\mu; \mu, \sigma) &= \Phi\left[\frac{1}{\sigma}\left(\sqrt{\frac{\mu}{\mu}} - \sqrt{\frac{\mu}{\mu}}\right)\right] \\
&= \Phi(0) \\
&= 0,5.
\end{aligned}$$

Assim, verifica-se que μ é mediana da distribuição BS.

Como apresenta Balakrishnan e Kundu (2019), a distribuição BS possui expressões analíticas para a média e a variância, que são dadas, respectivamente, por

$$E(Y) = \mu \left(1 + \frac{\sigma^2}{2} \right) \quad \text{e} \quad \text{Var}(Y) = (\sigma\mu)^2 \left(1 + \frac{5}{4}\sigma^2 \right).$$

Outro aspecto relevante da distribuição BS é sua função quantílica, cuja formulação é apresentada por Sánchez (2020) e expressa da seguinte forma:

$$\begin{aligned} y_p &= F^{-1}(p; \mu, \sigma) \\ &= \frac{\mu}{4} \left(\sigma z_p + \sqrt{\sigma^2 z_p^2 + 4} \right)^2, \end{aligned}$$

em que z_p é o quantil obtido tal que $P(Z \leq z_p) = p$, $0 < p < 1$ e $Z \sim N(0,1)$.

A distribuição BS é fechada sob proporcionalidade, isto é, se $Y \sim \text{BS}(\mu, \sigma)$ então $W = cY \sim \text{BS}(c\mu, \sigma)$. De fato, veja que

$$\begin{aligned} F_W(w; \mu, \sigma) &= P(W \leq w) \\ &= P(cY \leq w) \\ &= P\left(Y \leq \frac{w}{c}\right) \\ &= F_Y\left(\frac{w}{c}; \mu, \sigma\right) \\ &= \Phi \left[\frac{1}{\sigma} \left(\sqrt{\frac{w/c}{\mu}} - \sqrt{\frac{\mu}{w/c}} \right) \right] \\ &= \Phi \left[\frac{1}{\sigma} \left(\sqrt{\frac{w}{c\mu}} - \sqrt{\frac{c\mu}{w}} \right) \right] \\ &= F_W(w; c\mu; \sigma). \end{aligned}$$

A distribuição BS também é fechada sob reciprocidade, ou seja, se $Y \sim \text{BS}(\mu, \sigma)$ então, $W = Y^{-1} \sim \text{BS}(\mu^{-1}, \sigma)$. De fato, tem-se que

$$\begin{aligned} F_W(w; \mu, \sigma) &= P(W \leq w) \\ &= P(Y^{-1} \leq w) \\ &= P(Y \geq w^{-1}) \\ &= 1 - F_Y(w^{-1}) \end{aligned}$$

$$\begin{aligned}
&= 1 - \Phi \left[\frac{1}{\sigma} \left(\sqrt{\frac{w^{-1}}{\mu}} - \sqrt{\frac{\mu}{w^{-1}}} \right) \right] \\
&= \Phi \left[-\frac{1}{\sigma} \left(\sqrt{\frac{\mu^{-1}}{w}} - \sqrt{\frac{w}{\mu^{-1}}} \right) \right] \\
&= \Phi \left[\frac{1}{\sigma} \left(\sqrt{\frac{w}{\mu^{-1}}} - \sqrt{\frac{\mu^{-1}}{w}} \right) \right] \\
&= F_W(w; \mu^{-1}; \sigma).
\end{aligned}$$

Como apresentado em Balakrishnan e Kundu (2019), a variável aleatória $Y \sim \text{BS}(\mu, \sigma)$ pode ser relacionada a uma variável normal padrão $Z \sim N(0,1)$ por meio da seguinte expressão:

$$Y = \frac{\mu(\sigma Z + \sqrt{\sigma^2 Z^2 + 4})^2}{4}. \quad (2.1.6)$$

A partir de (2.1.6) é possível gerar valores pseudo-aleatórios para a distribuição Birnbaum-Saunders.

Na Figura 1 estão apresentadas as curvas para a função de densidade da distribuição Birnbaum-Saunders para alguns valores fixados dos parâmetros para μ e σ . Na Figura 1(a), têm-se as curvas função da densidade para um μ variando, enquanto o parâmetro σ está fixado. Nessa situação, observa-se que quanto menor for o valor de μ , mais leptocúrtica é a curva da função densidade. Na Figura 1(b), tem-se σ variando, enquanto o μ está fixado. Nesse cenário, percebe-se que quanto menor o valor de σ , mais simétrica é a distribuição em torno de μ .

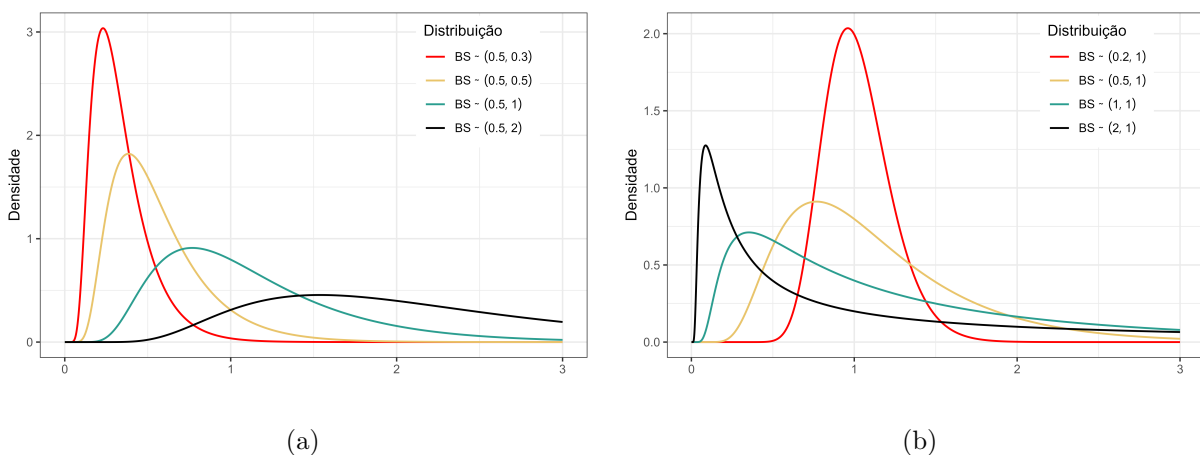


Figura 1: Curvas da função densidade da distribuição BS para $\sigma = 0.5$ com valores indicados de μ (a) e $\mu = 1$ com valores indicados de σ (b).

2.2 Distribuição Birnbaum-Saunders Tri-paramétrica

A suposição de independência entre os ciclos distintos imposta por Birnbaum e Saunders (1969) na derivação da distribuição BS é, em situações práticas, pouco realista. É natural imaginar que exista uma estrutura de dependência entre as extensões da fissura de cada ciclo, ou seja, que a extensão total da fissura no k -ésimo ciclo esteja correlacionado, em maior ou menor grau, com as extensões dos $(k - 1)$ ciclos anteriores. Em outras palavras, é razoável pensar que o material perde gradualmente a resistência ao longo dos ciclos, tornando-se mais vulnerável à falha.

No entanto, a quebra da suposição de independência invalida a utilização do Teorema do Limite Central na sua formulação clássica. Assim, como alternativa, Owen (2006) construiu uma nova família de distribuições chamada distribuição Birnbaum-Saunders tri-paramétrica, baseando-se na quebra da hipótese de independência entre os ciclos, e na teoria de processo de memória longa discutido por Beran (1994). A distribuição proposta por Owen (2006) é uma generalização da distribuição BS. Ao longo desta dissertação a distribuição proposta por Owen (2006) será chamada de BS generalizada (BSG).

Conforme as hipóteses formuladas por Owen (2006) sobre a distribuição BSG, Ribeiro (2014) apresentou uma definição detalhada das suposições adotadas na derivação do modelo:

1. Seja $\{X_t, t \in T\}$ um processo estocástico que representa a sequência de fissuras no material, sendo $E(X_t) = \mu^*$ e $\text{Var}(X_t) = \alpha^2$, então a variância amostral do processo é dada por

$$\text{Var}(\bar{X}) = \frac{\alpha^2}{n}(1 + \delta_n(\rho)),$$

sendo $\delta_n(\rho) = \frac{1}{n} \sum_{i \neq j} \rho(i, j)$ o termo de correção não-nulo da variância por causa da dependência entre as variáveis X_i e $\rho(i, j)$ é a correlação linear entre as variáveis X_i e X_j . Adicionalmente, o processo $\{X_t, t \in T\}$ é estocástico estacionário, isto é, as correlações $\rho(i, j)$ dependem apenas das defasagens, $k = |i - j|$, podendo reescrever o fator de correção não nulo como

$$\delta_n(\rho) = 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho(k);$$

2. O processo $\{X_t, t \in T\}$ é de memória longa, isto é,

$$\sum_{k=1}^{\infty} \rho(k) = \infty.$$

Logo, as correlações lineares decaem tão lentamente para zero que $\sum_{k=1}^{\infty} \rho(k)$ diverge;

3. O processo $\{X_t, t \in T\}$ é auto-similar com parâmetro de auto-similaridade $0 < \nu < 1$, isto é, para qualquer constante positiva c , tem-se que

$$c^{-\nu} X_{ct} \stackrel{d}{=} X_t,$$

e este mesmo processo terá incrementos estacionários, se para $t > s$

$$X_t - X_s \stackrel{d}{=} X_{t+k} - X_{s+k}.$$

4. O processo $\{X_t, t \in T\}$ é gaussiano com média μ^* e desvio padrão σ . Daí, segue que $n^{1-\kappa}(\bar{X} - \mu^*)/\alpha$ é uma variável aleatória com distribuição normal padrão.

De acordo com Beran (1994), se $\nu > 0,5$ o processo $\{X_t, t \in T\}$ é de memória longa. Em contrapartida, se $\nu < 0,5$ o processo é considerada de memória curta, isso é, as correlações lineares decaem rapidamente para zero.

Segue das suposições que,

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{i=1}^n X_i \\ &\stackrel{d}{=} n^{\nu-1} (X_1 - X_0) \\ &= n^{\nu-1} X_1, \end{aligned}$$

em que os incrementos estacionários de $\{X_t, t \in T\}$ são expressos por $X_i - X_{i-1}$, com $X_0 = 0$. Assim, a variância amostral é expressa por

$$\text{Var}(\bar{X}) = n^{2\nu-2} \alpha^2.$$

Note que se $\nu = 0,5$ a expressão da variância amostral se reduz para $\text{Var}(\bar{X}) = n^{-1} \alpha^2$.

Agora, seguindo de forma análoga à derivação do modelo BS clássico, considere a equação expressa em (2.1.2) como a extensão total da fissura após observados n ciclos, e N uma variável aleatória que indica o número de ciclos até que ocorra a falha. Utilizando a expressão (2.1.4) e a Suposição 4, tem-se que a FDA de N é dada por

$$\begin{aligned}
P(N \leq n) &= 1 - P\left(\frac{\sum_{j=1}^n Y_j - n\mu^*}{\alpha n^\nu} \leq \frac{\omega - n\mu^*}{\alpha n^\nu}\right) \\
&\approx \Phi\left(\frac{n\mu^* - \omega}{\alpha n^\nu}\right) \\
&= \Phi\left(\frac{\mu^* n^{1-\nu}}{\alpha} - \frac{\omega}{\alpha n^\nu}\right).
\end{aligned}$$

De maneira análoga à distribuição BS, substituindo N por Y , tem-se que Y é a extensão contínua da variável aleatória discreta N , em que Y é uma variável aleatória contínua que representa o tempo até a falha do material.

Seguindo a parametrização:

$$\sigma = \frac{\alpha}{\sqrt{\mu^* \omega}} > 0 \quad , \quad \mu = \frac{\omega}{\mu^*} > 0 \quad , \quad 0 < \nu < 1,$$

tem-se que a FDA de Y é expressa por

$$F(y; \mu, \sigma, \nu) = \Phi\left[\frac{1}{\sigma} \left(\frac{y^{1-\nu}}{\sqrt{\mu}} - \frac{\sqrt{\mu}}{y^\nu}\right)\right], \quad y > 0. \quad (2.2.1)$$

Diz-se que Y possui FDA expressa por (2.2.1), então Y segue uma distribuição Birnbaum-Saunders Generalizada com parâmetros μ , σ e ν e é denotada por $Y \sim \text{BSG}(\mu, \sigma, \nu)$.

Derivando a expressão (2.2.1) em relação a y , obtém-se a função densidade de probabilidade de Y dada por

$$f(y; \mu, \sigma, \nu) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma \sqrt{\mu y^\nu}} \left(1 - \nu + \frac{\nu \mu}{y}\right) \exp\left\{-\frac{1}{2\sigma^2} \frac{(y - \mu)^2}{\mu y^{2\nu}}\right\}, \quad y > 0. \quad (2.2.2)$$

Destacam-se algumas propriedades da distribuição derivada por Owen (2006). Primeiramente, observe que, a distribuição BS é um caso particular da distribuição BSG, quando $\nu = 0,5$. Diferentemente da formulação original, na distribuição BSG o parâmetro μ não é mais um parâmetro de escala. De fato, como discutido em Ribeiro (2014), para $\nu \neq 0,5$

$$\begin{aligned}
F(y/\mu; 1, \sigma, \nu) &= \Phi\left[\frac{1}{\sigma} \left(\frac{(y/\mu)^{1-\nu}}{\sqrt{1}} - \frac{\sqrt{1}}{(y/\mu)^\nu}\right)\right] \\
&= \Phi\left[\frac{1}{\sigma} \left(\frac{y^{1-\nu}}{\mu^{1-\nu}} - \frac{\mu^\nu}{y^\nu}\right)\right] \\
&\neq F(y; \mu, \sigma, \nu).
\end{aligned}$$

Assim, nota-se que μ não é um parâmetro de escala. De acordo com Devore (2009), os parâmetros σ e ν são forma, uma vez que não são parâmetros de escala, e nem de posição. De fato, nota que

$$\begin{aligned} F(y/\sigma; \mu, 1, \nu) &= \Phi \left[\frac{1}{1} \left(\frac{(y/\sigma)^{1-\nu}}{\sqrt{\mu}} - \frac{\sqrt{\mu}}{(y/\sigma)^\nu} \right) \right] \\ &= \Phi \left[\left(\frac{y^{1-\nu}}{\sigma^{1-\nu} \sqrt{\mu}} - \frac{\sigma^\nu \sqrt{\mu}}{y^\nu} \right) \right] \\ &\neq F(y; \mu, \sigma, \nu), \end{aligned}$$

e

$$\begin{aligned} F(y/\nu; \sigma, \mu, 1) &= \Phi \left[\frac{1}{\sigma} \left(\frac{(y/\nu)^{1-1}}{\sqrt{\mu}} - \frac{\sqrt{\mu}}{(y/\mu)^1} \right) \right] \\ &= \Phi \left[\frac{1}{\sigma} \left(\frac{1}{\sqrt{\mu}} - \frac{\nu \sqrt{\mu}}{1} \right) \right] \\ &\neq F(y; \mu, \sigma, \nu). \end{aligned}$$

Então, os parâmetros σ e ν também não são de posição, pois para verificar isto, por definição é necessário calcular $F(y - \sigma; \mu, 0, \nu)$ e $F(y - \nu; \mu, \sigma, 0)$. No entanto, não é possível calcular tais expressões, uma vez que $\sigma > 0$ e $0 < \nu < 1$. Dessa maneira, tem-se que σ e ν são parâmetros de forma.

De maneira similar a distribuição BS original, o parâmetro μ é a mediana da distribuição BSG(μ, σ, ν), entretanto μ não é mais um parâmetro de escala. De fato, observe que

$$\begin{aligned} F(\mu; \mu, \sigma, \nu) &= \Phi \left[\frac{1}{\sigma} \left(\frac{\mu^{1-\nu}}{\sqrt{\mu}} - \frac{\sqrt{\mu}}{\mu^\nu} \right) \right] \\ &= \Phi \left[\frac{1}{\sigma} \left(\frac{\mu - \mu}{\sqrt{\mu} \mu^\nu} \right) \right] \\ &= \Phi(0) \\ &= 0,5. \end{aligned}$$

Se $Y \sim \text{BSG}(\mu, \sigma, \nu)$, então $W = Y^{-1} \sim \text{BSG}(\mu^{-1}, \sigma, 1 - \nu)$. De fato, tem-se

$$\begin{aligned}
F_W(w; \mu, \sigma, \nu) &= P(W \leq w) \\
&= P(Y^{-1} \leq w) \\
&= P(Y \geq w^{-1}) \\
&= 1 - F_Y(w^{-1}) \\
&= 1 - \Phi \left[\frac{1}{\sigma} \left(\frac{w^{-(1-\nu)}}{\sqrt{\mu}} - \frac{\sqrt{\mu}}{w^{-\nu}} \right) \right] \\
&= \Phi \left[\frac{1}{\sigma} \left(\frac{\sqrt{\mu}}{w^{-\nu}} - \frac{w^{-(1-\nu)}}{\sqrt{\mu}} \right) \right] \\
&= \Phi \left[\frac{1}{\sigma} \left(\frac{w^\nu}{\mu^{-1/2}} - \frac{\mu^{-1/2}}{y^{1-\nu}} \right) \right] \\
&= F_W(w; \mu^{-1}, \sigma, 1 - \nu).
\end{aligned}$$

Assim, verifica-se que $W \sim \text{BSG}(\mu^{-1}, \sigma, 1 - \nu)$.

Owen (2006) apresenta a relação entre uma variável aleatória normal padrão, $Z \sim N(0,1)$, e uma variável aleatória $Y \sim \text{BSG}(\mu, \sigma, \nu)$ dada por

$$\mu + \sigma\sqrt{\mu}ZY^\nu - Y = 0. \quad (2.2.3)$$

A equação definida em (2.2.3) é não linear em Y , diferente do que foi visto para a distribuição BS clássica. Entretanto, pode-se utilizar os métodos numéricos para solucionar a equação não linear em relação a Y para gerar valores pseudo-aleatórios da BSG dado valores pseudo-aleatórios da distribuição normal padrão.

A partir da relação (2.2.3) pode-se obter quantis da distribuição BSG y_p , conforme expressa na equação a seguir:

$$\sigma\sqrt{\mu}z_p y_p^\nu + \mu - y_p = 0,$$

em que z_p é o quantil obtido tal que $P(Z \leq z_p) = p$, $0 < p < 1$ e $Z \sim N(0,1)$.

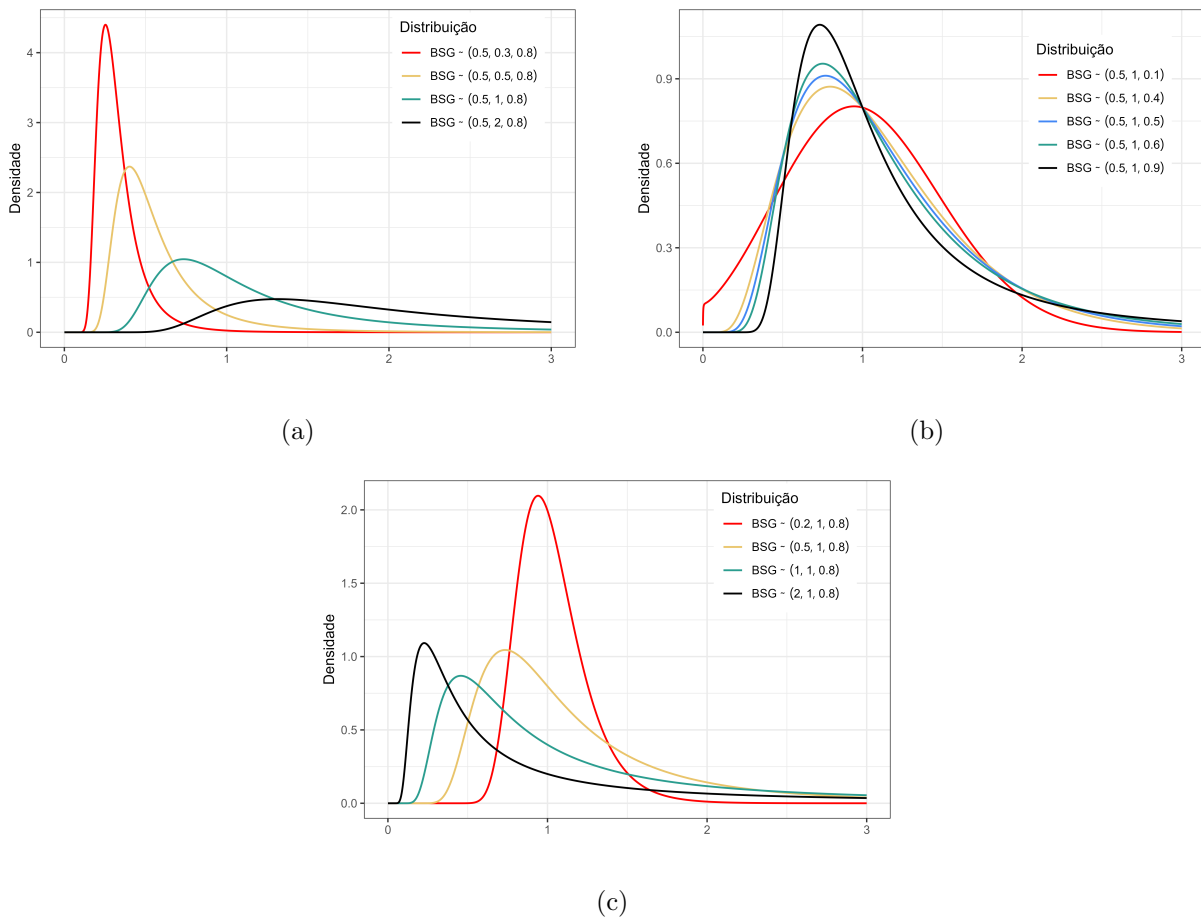


Figura 2: Curvas da função densidade da distribuição BSG para $\sigma = 0,5, \nu = 0,8$ para valores indicados de μ (a), $\sigma = 0,5, \mu = 0,1$ para valores indicados de ν (b) e $\mu = 1, \nu = 0,8$ para valores indicados de σ (c).

Na Figura 2 estão apresentadas curvas da função de densidade da distribuição BSG para valores fixados dos parâmetros μ, σ e ν . Na Figura 2(a), têm-se as curvas para alguns valores de μ , enquanto σ e ν são fixados. Percebe-se que quanto menor o valor de μ , mais leptocúrtica é a curva de densidade. Já na Figura 2(b), têm-se as curvas da função densidade para alguns valores de ν , enquanto σ e μ são fixados. Note que quanto menor o valor de ν , menos assimétricas são as curvas de densidade. Por fim, na Figura 2(c), têm-se as curvas de densidade para valores de σ , enquanto μ e ν são fixados. Veja que, quanto menor o valor de σ , menos assimétrica é a distribuição em torno de μ .

3 Modelo de Regressão Birnbaum-Saunders Generalizado

3.1 Estrutura de Regressão

O modelo de regressão BSG é definido assumindo que as variáveis aleatórias Y_1, \dots, Y_n são independentes tal que cada variável aleatória Y_i assume função densidade de probabilidade definida em (2.2.2). Desta maneira, o novo modelo de regressão proposto, em sua forma matricial, é definido como

$$\mathbf{Y} \stackrel{\text{ind}}{\sim} \text{BSG}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\nu}), \quad (3.1.1)$$

em que,

$$\begin{aligned} g_1(\boldsymbol{\mu}) &= \boldsymbol{\eta}_1 = \mathbf{X}_1\boldsymbol{\beta} + s_{11}(\mathbf{x}_{11}) + \dots + s_{1J_1}(\mathbf{x}_{1J_1}), \\ g_2(\boldsymbol{\sigma}) &= \boldsymbol{\eta}_2 = \mathbf{X}_2\boldsymbol{\gamma} + s_{21}(\mathbf{x}_{21}) + \dots + s_{2J_2}(\mathbf{x}_{2J_2}), \\ g_3(\boldsymbol{\nu}) &= \boldsymbol{\eta}_3 = \mathbf{X}_3\boldsymbol{\lambda} + s_{31}(\mathbf{x}_{31}) + \dots + s_{3J_3}(\mathbf{x}_{3J_3}). \end{aligned} \quad (3.1.2)$$

Os vetores $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ e $\boldsymbol{\nu}$ são vetores n -dimensionais de parâmetros desconhecidos a serem estimados, $\boldsymbol{\eta}_k$, $g_k(\cdot)$ e \mathbf{X}_k são, respectivamente, o k -ésimo preditor linear, a k -ésima função de ligação assumida conhecida, estritamente monótona e duas vezes diferenciável e a k -ésima matriz conhecida de variáveis explicativas, com $k = 1, 2, 3$. Os vetores $\boldsymbol{\beta} \in \mathbb{R}^{p_1}$, $\boldsymbol{\gamma} \in \mathbb{R}^{p_2}$ e $\boldsymbol{\lambda} \in \mathbb{R}^{p_3}$ são os vetores de parâmetros desconhecidos dos coeficientes de regressão associados, respectivamente, a $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ e $\boldsymbol{\nu}$. Além disso, $s_{kj}(\cdot)$ representam funções de suavização associadas às covariáveis \mathbf{x}_{kj} , em que $j = 1, \dots, J_k$.

O modelo definido por (3.1.1) e (3.1.2) sob a distribuição Birnbaum-Saunders Generalizada apresenta estrutura análoga à configuração dos GAMLSS propostos por Rigby e Stasinopoulos (2005). Por esse motivo, a distribuição BSG foi implementada de forma inédita como uma família de distribuições pertencente aos GAMLSS no ambiente estatístico R. Essa classe de modelos permite a modelagem simultânea de parâmetros de localização, escala e forma contemplando uma ampla variedade de distribuições de até quatro parâmetros, permitindo uma maior flexibilidade ao ajuste dos dados, Stasinopoulos et al. (2017).

A classe dos GAMLSS permite a incorporação de funções paramétricas lineares e não-lineares, assim como funções de suavização não paramétricas. De acordo com Stasinopoulos et al. (2017), é possível utilizar diversas formas de termos aditivos como

P-*splines* (B-*splines* penalizados), P-*splines* monótonos, P-*splines* cíclicos, P-*splines* com coeficientes variáveis, *splines* cúbicos suavizados, ajuste de curvas por *loess*, polinômios fracionários, efeitos aleatórios, regressão *ridge* e ajustes paramétricos não lineares.

Na ausência de funções de suavização e de termos aditivos no modelo de regressão BSG, a estrutura descrita em (3.1.2) se reduz a

$$\begin{aligned} g_1(\boldsymbol{\mu}) &= \boldsymbol{\eta}_1 = \mathbf{X}_1\boldsymbol{\beta}, \\ g_2(\boldsymbol{\sigma}) &= \boldsymbol{\eta}_2 = \mathbf{X}_2\boldsymbol{\gamma}, \\ g_3(\boldsymbol{\nu}) &= \boldsymbol{\eta}_3 = \mathbf{X}_3\boldsymbol{\lambda}. \end{aligned} \tag{3.1.3}$$

A estrutura apresentada em (3.1.3) é definida como modelo de regressão BSG paramétrico, enquanto o modelo expresso em (3.1.2) corresponde à sua versão semi-paramétrica.

Os GAMLSS paramétricos são ajustados utilizando o método de máxima verossimilhança, enquanto os modelos semi-paramétricos são estimados a partir do método de máxima verossimilhança penalizada. Maiores detalhes sobre este segundo método podem ser encontrados em Stasinopoulos et al. (2017).

O principal objetivo desta dissertação consiste em estudar um novo modelo de regressão paramétrico baseado na distribuição Birnbaum-Saunders Generalizada proposta por Owen (2006). Em particular, busca-se modelar a mediana da distribuição BSG, permitindo, simultaneamente, o ajuste dos demais parâmetros. Essa abordagem possibilita maior flexibilidade ao processo de modelagem, ampliando sua capacidade de adaptação a diferentes conjuntos de dados.

3.2 Inferência por Máxima Verossimilhança

Conforme mencionado na seção anterior, o modelo de regressão proposto será ajustado a partir do método de máxima verossimilhança. A seguir serão apresentadas a função de verossimilhança associado à distribuição Birnbaum-Saunders Generalizada, bem como as expressões correspondentes vetores escore de cada um dos parâmetros de regressão.

Sejam Y_1, \dots, Y_n variáveis aleatórias independentes com distribuição BSG em que $Y_i \stackrel{ind}{\sim} \text{BSG}(\mu_i, \sigma_i, \nu_i)$ e seja $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \boldsymbol{\lambda}^\top)^\top \in \mathbb{R}^{p_1+p_2+p_3}$. A função de verossimilhança

para $\boldsymbol{\theta}$ é expressa como

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^n \left[\frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_i \sqrt{\mu_i y_i^{\nu_i}}} \left(1 - \nu_i + \frac{\nu_i \mu_i}{y_i} \right) \exp \left(-\frac{1}{2\sigma_i^2} \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\mu_i y_i^{2\nu_i}} \right) \right] \\ &= (2\pi)^{-(n/2)} \prod_{i=1}^n \left[(\sigma_i \sqrt{\mu_i y_i^{\nu_i}})^{-1} \left(1 - \nu_i + \frac{\nu_i \mu_i}{y_i} \right) \exp \left(-\frac{1}{2} \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\sigma_i \mu_i y_i^{2\nu_i}} \right) \right]. \end{aligned}$$

Assim, o logaritmo da função de verossimilhança é dado por

$$\ell(\boldsymbol{\theta}) = \log(L(\boldsymbol{\theta})) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}), \quad (3.2.1)$$

em que, $\ell_i(\boldsymbol{\theta}) = \log[f(y_i; \mu_i, \sigma, \nu_i)]$,

$$\ell_i(\boldsymbol{\theta}) = -\frac{1}{2} \log(2\pi) - \log(\sigma_i) - \frac{1}{2} \log(\mu_i) - \nu_i \log(y_i) + \log \left(1 - \nu_i + \frac{\nu_i \mu_i}{y_i} \right) - \frac{1}{2} \frac{(y_i - \mu_i)^2}{\sigma_i^2 \mu_i y_i^{2\nu_i}},$$

sendo $f(\mu_i, \sigma_i, \nu_i)$ a função de densidade de probabilidade apresentada em (2.2.2). O estimador de máxima verossimilhança é obtido a partir da maximização da função definida em (3.2.1) em relação a $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \boldsymbol{\lambda}^\top)^\top$. O EMV para $\boldsymbol{\theta}$ é denotado como $\hat{\boldsymbol{\theta}}$.

Adiante serão apresentadas os vetores escore de cada um dos parâmetros desconhecidos. Cada função é obtida a partir da derivação do logaritmo da função de verossimilhança em relação ao parâmetro de interesse.

Ao derivar a expressão (3.2.1) em relação a β_j , para $j = 1, \dots, p_1$, tem-se que

$$\begin{aligned} U_{\beta_j} &= \sum_{i=1}^n \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_j} \\ &= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\mu_i} \frac{d\mu_i}{d\eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_j} \\ &= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\mu_i} \frac{x_{1ij}}{g'_1(\mu_i)}, \end{aligned} \quad (3.2.2)$$

em que η_{1i} é a i -ésima entrada do preditor linear η_1 , $g'_1(\mu_i) = d\mu_i/d\eta_{1i}$, $i = 1, \dots, n$, x_{1ij} é o i -ésimo valor fixado da j -ésima covariável associada ao submodelo de $\boldsymbol{\mu}$, $j = 1, \dots, p_1$. Assim, o termo $d\ell_i(\boldsymbol{\theta})/d\mu_i$ é calculado como

$$\begin{aligned}
\frac{d\ell_i(\boldsymbol{\theta})}{d\mu_i} &= -\frac{1}{2} \frac{1}{\mu_i} + \frac{1}{1 - \nu_i + \frac{\nu_i \mu_i}{y_i}} \frac{\nu_i}{y_i} - \frac{1}{2\sigma_i^2 y_i^{2\nu_i}} \left[\frac{2(y_i - \mu_i)(-1)\mu_i - (y_i - \mu_i)^2}{\mu_i^2} \right] \\
&= -\frac{1}{2\mu_i} + \frac{y_i}{y_i - y_i \nu_i + \nu_i \mu_i} \frac{\nu_i}{y_i} - \frac{1}{2\sigma_i^2 y_i^{2\nu_i} \mu_i^2} [-2y_i \mu_i + 2\mu_i^2 - y_i^2 - \mu_i^2 + 2y_i \mu_i] \\
&= -\frac{1}{2\mu_i} + \frac{\nu_i}{y_i(1 - \nu_i) + \nu_i \mu_i} + \frac{(y_i^2 - \mu_i^2)}{2\sigma_i^2 y_i^{2\nu_i} \mu_i^2}.
\end{aligned}$$

Definindo,

$$a_i = -\frac{1}{2\mu_i} + \frac{\nu_i}{y_i(1 - \nu_i) + \nu_i \mu_i} + \frac{(y_i^2 - \mu_i^2)}{2\sigma_i^2 y_i^{2\nu_i} \mu_i^2}. \quad (3.2.3)$$

Segue que,

$$U_{\beta_j} = \sum_{i=1}^n a_i \frac{x_{1ij}}{g_1'(\mu_i)}.$$

Portanto, o vetor escore para $\boldsymbol{\beta}$ é expresso por

$$U_{\boldsymbol{\beta}} = \mathbf{X}_1^\top T_1 \mathbf{a},$$

em que \mathbf{X}_1 é uma matriz com dimensão $n \times p_1$ em que cada coluna contém os valores fixados das variáveis explicativas associadas ao parâmetro $\boldsymbol{\mu}$, T_1 é uma matriz diagonal definida por $\text{diag}\{1/g_1'(\mu_1), \dots, 1/g_1'(\mu_n)\}$, e $\mathbf{a} = (a_1, \dots, a_n)^\top \in \mathbb{R}^n$, com a_i definido em (3.2.3).

Agora, derivando (3.2.1) em relação a γ_j , para $j = 1, \dots, p_2$, tem-se que a seguinte expressão:

$$\begin{aligned}
U_{\gamma_j} &= \sum_{i=1}^n \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \gamma_j} \\
&= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\sigma_i} \frac{d\sigma_i}{d\eta_{2i}} \frac{\partial \eta_{2i}}{\partial \gamma_j} \\
&= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\sigma_i} \frac{x_{2ij}}{g_2'(\sigma_i)}, \quad (3.2.4)
\end{aligned}$$

em que η_{2i} é a i -ésima entrada do preditor linear $\boldsymbol{\eta}_2$, $g_2'(\sigma_i) = d\sigma_i/d\eta_{2i}$, $i = 1, \dots, n$, x_{2ij} é o i -ésimo valor fixado da j -ésima variável explicativa associada a $\boldsymbol{\sigma}$, $j = 1, \dots, p_2$. O termo $d\ell_i(\boldsymbol{\theta})/d\sigma_i$ é definido como

$$\begin{aligned}
\frac{d\ell_i(\boldsymbol{\theta})}{d\sigma_i} &= -\frac{1}{\sigma_i} - \frac{1}{2} \frac{(y_i - \mu_i)^2}{\mu_i y_i^{2\nu_i}} \left[-\frac{2}{\sigma_i^3} \right] \\
&= -\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{\mu_i y_i^{2\nu_i} \sigma_i^3}.
\end{aligned}$$

Considere,

$$b_i = -\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{\mu_i y_i^{2\nu_i} \sigma_i^3}. \quad (3.2.5)$$

Segue que,

$$U_{\gamma_j} = \sum_{i=1}^n b_i \frac{x_{2ij}}{g_2'(\sigma_i)}.$$

Logo, o vetor escore para $\boldsymbol{\gamma}$ é definido como

$$U_{\boldsymbol{\gamma}} = \mathbf{X}_2^\top T_2 \mathbf{b},$$

em que \mathbf{X}_2 é uma matriz com dimensão $n \times p_2$ em que cada coluna contém os valores fixados das variáveis explicativas associadas ao $\boldsymbol{\sigma}$, T_2 é uma matriz diagonal definida por $\text{diag}\{1/g_2'(\sigma_1), \dots, 1/g_2'(\sigma_n)\}$, e $\mathbf{b} = (b_1, \dots, b_n)^\top \in \mathbb{R}^n$, com b_i definido em (3.2.5).

Por fim, derivando o logaritmo da função de verossimilhança em relação a λ_j , para $j = 1, \dots, p_3$, tem-se que o seguinte:

$$\begin{aligned} U_{\lambda_j} &= \sum_{i=1}^n \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \lambda_j} \\ &= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\nu_i} \frac{d\nu_i}{d\eta_{3i}} \frac{\partial \eta_{3i}}{\partial \lambda_j} \\ &= \sum_{i=1}^n \frac{d\ell_i(\boldsymbol{\theta})}{d\nu_i} \frac{x_{3ij}}{g_3'(\nu_i)}, \end{aligned} \quad (3.2.6)$$

em que η_{3i} é a i -ésima entrada do preditor linear η_3 , $g_3'(\nu_i) = d\nu_i/d\eta_{3i}$, $i = 1, \dots, n$, x_{3ij} é o i -ésimo valor fixado da j -ésima variável explicativa associada a ν , $j = 1, \dots, p_3$. O termo $d\ell_i(\boldsymbol{\theta})/d\nu_i$ é definido como

$$\begin{aligned} \frac{d\ell_i(\boldsymbol{\theta})}{d\nu_i} &= -\log(y_i) + \frac{1}{1 - \nu_i + \frac{\nu_i \mu_i}{y_i}} \left(\frac{\mu_i}{y_i} - 1 \right) - \frac{(y_i - \mu_i)^2}{2\sigma_i^2 \mu_i} (-2y_i^{-2\nu_i} \log(y_i)) \\ &= -\log(y_i) + \frac{y_i}{y_i - y_i \nu_i + \nu_i \mu_i} \left(\frac{\mu_i - y_i}{y_i} \right) + \frac{(y_i - \mu_i)^2 y_i^{-2\nu_i} \log(y_i)}{\sigma_i^2 \mu_i} \\ &= -\log(y_i) + \frac{(\mu_i - y_i)}{y_i(1 - \nu_i + \nu_i \mu_i)} + \frac{(y_i - \mu_i)^2 \log(y_i)}{\sigma_i^2 \mu_i y_i^{2\nu_i}} \\ &= \log(y_i) \left(\frac{(y_i - \mu_i)^2}{\sigma_i^2 \mu_i y_i^{2\nu_i}} - 1 \right) + \frac{(\mu_i - y_i)}{y_i(1 - \nu_i + \nu_i \mu_i)}. \end{aligned}$$

Definindo,

$$c_i = (1 - \log(y_i)) \left(\frac{(y_i - \mu_i)^2}{\sigma_i^2 \mu_i y_i^{2\nu_i}} + \frac{(\mu_i - y_i)}{y_i(1 - \nu_i + \nu_i \mu_i)} \right). \quad (3.2.7)$$

Segue que,

$$U_{\lambda_j} = \sum_{i=1}^n c_i \frac{x_{3ij}}{g_3'(\nu_i)}.$$

Logo, o vetor escore para $\boldsymbol{\lambda}$ é definido como

$$U_{\boldsymbol{\lambda}} = \mathbf{X}_3^\top T_3 \mathbf{c},$$

em que \mathbf{X}_3 é uma matriz com dimensão $n \times p_3$ em que cada coluna contém os valores fixados das variáveis explicativas a $\boldsymbol{\nu}$, T_3 que é $\text{diag}\{1/g_3'(\nu_1), \dots, 1/g_3'(\nu_n)\}$, e $\mathbf{c} = (c_1, \dots, c_n)^\top \in \mathbb{R}^n$, com c_i definido em (3.2.7).

O EMV $\hat{\boldsymbol{\theta}}$ é obtido a partir da resolução do sistema $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$, em que $\mathbf{0}$ representa o vetor nulo de dimensão $p = p_1 + p_2 + p_3$. No entanto, como não há uma solução analítica para esse sistema de equações, é necessário aplicar algum procedimento numérico para obter uma solução aproximada. No âmbito dos GAMLSS, os principais métodos de aproximação são os algoritmos de Rigby e Stasinopoulos (RS) e de Cole e Green (CG).

Os cálculos e resoluções numéricas associados às estimações dos parâmetros dos modelos produzidos nesta dissertação foram produzidos a partir do *software* estatístico R. O R (Core Team, 2025) é um *software* de código aberto e domínio público, e está disponível gratuitamente para *download* no endereço eletrônico: <http://www.r-project.org>.

Em particular, estas operações foram conduzidas por meio da função `gamlss` disponível na biblioteca de mesmo nome, cuja estrutura foi proposta por Rigby e Stasinopoulos (2005). Ressalta-se que, ao empregar a função `gamlss` para o ajuste de um modelo de regressão, é gerado um objeto da classe `gamlss`, o qual pode ser utilizado como argumento em diversas outras funções da biblioteca.

O procedimento de máxima verossimilhança implementado na função `gamlss` desenvolvido por Stasinopoulos et al. (2017), baseia-se em dois algoritmos principais, o Rigby e Stasinopoulos e na generalização do algoritmo Cole e Green. A biblioteca `gamlss` oferece uma estrutura extensiva para o processo de modelagem e inferência estatística, permitindo obter as estimativas dos parâmetros e seus respectivos erros padrão, além de fornecer os resíduos quantílicos, ferramentas de diagnósticos, predições e outros recursos. Para maiores detalhes veja, Stasinopoulos et al. (2017), Rigby et al. (2019), Stasinopoulos et al. (2024). Informações complementares, incluindo lista de colaboradores, materiais de apoio, aplicações práticas e atualizações, estão disponíveis no site oficial: <https://www.gamlss.com>.

No que se refere aos aspectos computacionais, é pertinente destacar alguns pontos relacionados ao procedimento de estimação sob o novo modelo de regressão proposto com base na distribuição BSG, tendo em vista que a função `gamlss` oferece ampla flexibilidade quanto à configuração e à especificação do modelo. Para a realização do ajuste do modelo, é necessário definir previamente um vetor de valores iniciais associado aos parâmetros da distribuição a serem estimados. No caso da distribuição BSG, os parâmetros os quais deverão ser atribuídos valores iniciais são μ_i , σ_i e ν_i , com $i = 1, \dots, n$. Destaca-se, ainda, que a função `gamlss` não permite a especificação direta dos valores iniciais para os coeficientes de regressão associados aos parâmetros da distribuição que são β , γ e λ . Neste trabalho, estabeleceu-se o seguinte conjunto como valores iniciais para o processo iterativo de estimação, $\mu_i^{(0)} = \tilde{y}$, $\sigma_i^{(0)} = 1$ e $\nu_i^{(0)} = 0,5$, sendo \tilde{y} a mediana amostral da variável resposta. De maneira equivalente, o conjunto de valores iniciais para os coeficientes de regressão corresponde à definição de valores iniciais apenas para os interceptos nos três submodelos de regressão, enquanto todos os coeficientes associados às covariáveis são inicialmente fixados em zero. Portanto, o valor inicial para o intercepto do submodelo para a mediana é $\beta_1^{(0)} = g_1(\tilde{y})$. O valor inicial para o intercepto do submodelo para σ é $\gamma_1^{(0)} = g_2(1)$. Por fim, o valor inicial para o intercepto do submodelo para ν é $\lambda_1^{(0)} = g_3(0,5)$. Por exemplo, considerando que a função de ligação logarítmica para os submodelos de μ e σ , os interceptos seriam, respectivamente, $\log(\tilde{y})$ e zero. Por sua vez, para o submodelo do parâmetro ν , caso a função de ligação seja logito, o intercepto é fixado em zero.

Destaca-se que, além da estratégia de inicialização descrita acima, considerou-se também uma abordagem alternativa em que os valores iniciais dos parâmetros μ_i e σ_i foram definidos como as estimativas obtidas a partir do ajuste do modelo proposto por Santos-Neto et al. (2012), referente à BSG reparametrizada pela média. Por sua vez, o valor inicial de ν_i foi mantido em 0,5. Contudo, observou-se um aumento considerável no tempo computacional, sem evidências de melhora no desempenho do algoritmo de estimação RS, não justificando, portanto, a adoção dessa abordagem.

Neste trabalho, no âmbito da implementação computacional, definiu-se que o conjunto de valores iniciais adotado permaneça o mesmo, independentemente da estrutura de regressão especificada para os parâmetros do modelo. As derivadas necessárias no processo de estimação foram aproximadas a partir de métodos numéricos, através da função `numeric.deriv`, disponível na biblioteca `gamlss`. Por fim, no que diz respeito ao algoritmo iterativo, com base nas observações de Stasinopoulos et al. (2017) acerca dos algoritmos disponíveis na arquitetura do GAMLSS e a partir da experiência obtida nos estudos de simulação de Monte Carlo, o algoritmo RS apresentou um número menor de réplicas descartadas, exibindo assim um desempenho superior em termos de convergência das estimativas. Por essa razão, o algoritmo RS é aquele que será utilizado nesta dissertação para a implementação computacional do novo modelo de regressão BSG. No

entanto, ressalta-se que a implementação funciona independente do algoritmo.

Da maneira como o novo modelo de regressão Birnbaum–Saunders Generalizado foi implementado, é igualmente possível ajustar um modelo baseado na distribuição Birnbaum–Saunders clássica, uma vez que esta constitui um caso particular da distribuição BSG quando o parâmetro ν é fixado em 0,5. Portanto, ao ajustar o modelo BSG por meio da função `gamlss`, definindo o argumento `nu.fix` como verdadeiro e fixando o argumento `i.nu` em 0,5, obtém-se o ajuste correspondente ao modelo BS clássico.

Sob condições de regularidade usuais do método de máxima verossimilhança, quando o tamanho amostral é grande tem-se que

$$\hat{\boldsymbol{\theta}} \stackrel{a}{\sim} N_k(\boldsymbol{\theta}, J^{-1}(\boldsymbol{\theta})),$$

sendo que $\stackrel{a}{\sim}$ representa distribuição aproximada, $N_k(\boldsymbol{\theta}, J^{-1}(\boldsymbol{\theta}))$ representa a distribuição normal multivariada k -dimensional com vetor média $\boldsymbol{\theta}$ e matriz variância-covariância $J^{-1}(\boldsymbol{\theta})$, sendo $J^{-1}(\boldsymbol{\theta})$ a inversa da matriz de informação observada de Fisher, e k indica o número de parâmetros do modelo.

Suponha que o interesse é testar a hipótese $H_0: \theta_j = \theta_j^0$ versus $H_1: \theta_j \neq \theta_j^0$, em que θ_j^0 é um valor determinado para o parâmetro desconhecido θ_j , $j = 1, \dots, k$. Sob a hipótese nula, tem-se que

$$z_j = \frac{\hat{\theta}_j - \theta_j^0}{\text{se}(\hat{\theta}_j)} \stackrel{a}{\sim} N(0,1),$$

em que $\hat{\theta}_j$ é o EMV de θ_j , e $\text{se}(\hat{\theta}_j) = \sqrt{J_{jj}^{-1}}$ representa o erro padrão assintótico de $\hat{\theta}_j$, sendo J_{jj} o j -ésimo elemento da matriz diagonal de $J^{-1}(\boldsymbol{\theta})$. A estatística do teste z_j é a raiz quadrada da estatística de Wald (Wald, 1943), usada amplamente na avaliação de significância de coeficientes de regressão ao definir $\theta_j^0 = 0$.

Analogamente, um intervalo de confiança com nível de confiança de $(1 - \alpha) \times 100\%$ para o parâmetro θ_j é obtido, de forma aproximada, a partir da distribuição assintótica do estimador $\hat{\theta}_j$, por meio de

$$\left(\hat{\theta}_j - z_{1-\frac{\alpha}{2}} \cdot \widehat{\text{se}}(\hat{\theta}_j); \quad \hat{\theta}_j + z_{1-\frac{\alpha}{2}} \cdot \widehat{\text{se}}(\hat{\theta}_j) \right)$$

3.3 Técnicas de Diagnóstico sob os GAMLSS

Para avaliar a qualidade do ajuste de modelos de regressão aos dados, é fundamental realizar um diagnóstico que permita verificar a adequação do modelo. Esse processo inclui examinar os resíduos para verificar as suposições do modelo e identificar possíveis observações discrepantes. Sem a devida validação do modelo, as inferências e

interpretações fornecidas não são confiáveis. Por exemplo, na estrutura dos Modelos Lineares Generalizados, os resíduos componente do desvio e de Pearson são frequentemente utilizados. No entanto, conforme destacado por Stasinopoulos et al. (2017), os resíduos componente do desvio geralmente não são bem definidos, enquanto que os resíduos de Pearson podem não apresentar comportamento aproximadamente normal quando a variável resposta apresenta elevada assimetria ou curtose. Tal aspecto é particularmente relevante na avaliação do ajuste de modelos a partir de técnicas consolidadas na literatura que se baseiam na normalidade, como o *qq-plot*. Dunn e Smyth (1996) propuseram o resíduo quantílico, o tipo de resíduo padrão na avaliação de modelos ajustados no contexto dos GAMLSS.

3.3.1 Resíduo Quantílico

Sob o modelo de regressão BSG, o resíduo quantílico desenvolvido por Dunn e Smyth (1996) é definido como

$$r_{q,i} = \Phi^{-1}\{F(y_i; \hat{\mu}_i, \hat{\sigma}_i, \hat{\nu}_i)\},$$

em que $\Phi^{-1}(\cdot)$ é a inversa da função de distribuição acumulada da distribuição normal padrão, $N(0,1)$, e $F(y_i; \mu_i, \sigma_i, \nu_i)$ é a função de distribuição acumulada da distribuição BSG. O resíduo quantílico $r_{q,i}$, possui aproximadamente uma distribuição normal padrão, desde que o modelo de regressão BSG esteja corretamente especificado.

A principal vantagem do resíduo quantílico é que, independentemente da distribuição da variável resposta, se o modelo ajustado for adequado, então, para amostras grandes, $\hat{r}_{q,1}, \dots, \hat{r}_{q,n}$ seguirão uma distribuição normal padrão e serão independentes. Dessa forma, como a análise de resíduos baseada na normalidade é um critério amplamente utilizado na literatura, a avaliação do modelo torna-se mais familiar e acessível, Stasinopoulos et al. (2017).

Os resíduos quantílicos são obtidos por meio da função `resid()`, disponível na biblioteca `gamlss`. Na arquitetura do `gamlss`, diversas outras rotinas também fazem uso desse tipo de resíduo, destacando-se, entre elas, as funções de diagnóstico `wp()` e `plot.gamlss()`, amplamente empregadas na avaliação do ajuste do modelo.

3.3.2 Worm Plot

Proposto por Buuren e Fredriks (2001), o *worm plot* é uma ferramenta gráfica utilizada para avaliar a adequação de modelos. Por meio dessa técnica, é possível identificar padrões nos resíduos que indicam desvios em relação às suposições do modelo, bem

como verificar se os resíduos quantílicos seguem a distribuição esperada sob o modelo bem ajustado. Adicionalmente, o *worm plot* permite identificar regiões específicas em que o ajuste do modelo não foi adequado.

Em termos gerais, o *worm plot* consiste num *QQ-plot* sem tendência. No eixo das abscissas estão os valores esperados dos resíduos quantílicos sob adequação do modelo, enquanto no eixo das ordenadas encontram-se os valores dos resíduos quantílicos subtraídos dos respectivos quantis esperados da distribuição normal padrão. A Figura 3 apresenta um exemplo de *worm plot* de um ajuste de um modelo apropriado aos dados.

Conforme indicado por Stasinopoulos et al. (2024), ao analisar um *worm plot* é importante observar determinadas características do gráfico. Os pontos indicam o quão afastados os valores observados dos resíduos ordenados estão em relação aos seus respectivos valores teóricos para cada observação. Assim, caso o modelo esteja bem ajustado, espera-se que os pontos estejam próximos à linha horizontal. A área entre as elipses indicam um intervalo de confiança de 95%, desta maneira, caso espera-se que 95% dos pontos estejam dentro desta região. Segundo Stasinopoulos et al. (2017), padrões evidentes de tendência e mais de 5% dos pontos estarem fora da região de confiança, são indícios de falta de ajuste do modelo aos dados.

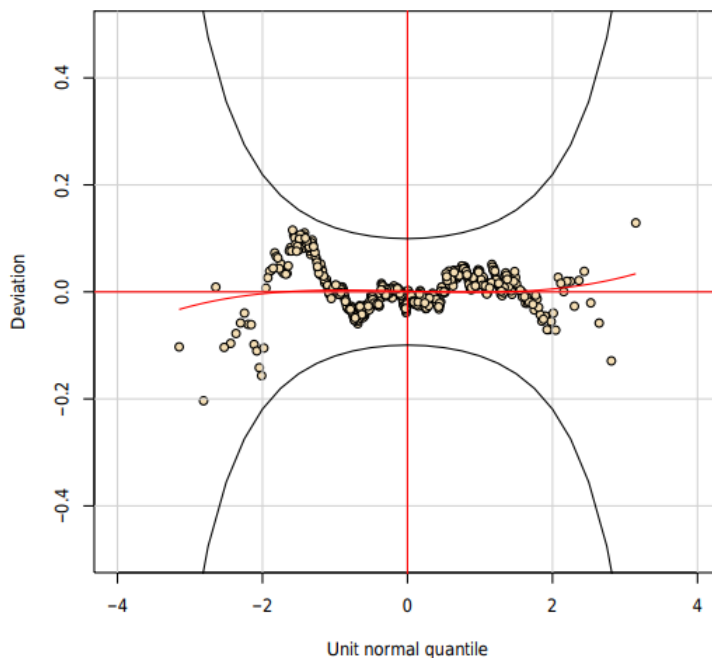


Figura 3: Exemplo de um *worm plot*, Stasinopoulos et al. (2017).

O *worm plot* sob os GAMLSS pode ser obtido facilmente por meio da função `wp()`, disponível na biblioteca `gamlss`, desde que o argumento da função seja um objeto da classe `gamlss` correspondente a um modelo previamente ajustado. A função `wp()` dispõe de diversas opções de configuração, para maiores detalhes sobre a implementação

computacional dos *worm plot*, veja Stasinopoulos et al. (2017).

Buuren e Fredriks (2001) apresentam as interpretações em relação ao comportamento esperado do *worm plot*, bem como possíveis ajustes sobre a medida de localização, variabilidade, assimetria e curtose. A Tabela 1 resume as possíveis interpretações de um *worm plot*.

Tabela 1: Interpretação do ajuste com base no formato do *worm plot*

Forma do <i>worm plot</i>	Resíduos	Interpretação na distribuição ajustada
Concentração acima da origem	Média muito alta	Parâmetro de localização subestimado
Concentração abaixo da origem	Média muito baixa	Parâmetro de localização sobrestimado
Inclinação positiva	Variância muito alta	Parâmetro de escala muito baixo
Inclinação negativa	Variância muito baixa	Parâmetro de escala muito alto
Forma de U	Assimetria à direita	Excesso de assimetria à esquerda
Forma de U invertido	Assimetria à esquerda	Excesso de assimetria à direita
Forma de S com esquerda para baixo	Leptocúrticos	Caudas muito leves
Forma de S com esquerda para cima	Platicúrticos	Caudas muito pesadas

3.3.3 Gráficos de Diagnóstico sob os GAMLSS

A análise de diagnóstico dos resíduos sob os GAMLSS é conduzida com base nos resíduos quantílicos. De acordo a metodologia padrão dessa classe de modelos, é comum realizar a análise dos resíduos por meio de quatro tipos de gráficos de diagnóstico.

Para obter estes gráficos a partir do pacote `gamlss`, pode-se empregar a função `plot.gamlss()` ou, de forma alternativa, a função genérica `plot()`, desde que o argumento desta última seja um objeto da classe `gamlss`. Esta função disponibiliza uma série de argumentos para personalização. Para informações detalhadas sobre esse tipo de configuração, consulte Stasinopoulos et al. (2017).

O primeiro gráfico consiste nos resíduos em função dos valores estimados para o parâmetro μ , a partir dessa visualização avalia-se a presença de padrões sistemáticos que possam indicar alguma forma de dependência entre os resíduos e o valores ajustados, além de fornecer informações sobre a variabilidade dos resíduos. Sob a hipótese de bom ajuste, espera-se que os resíduos estejam distribuídos aleatoriamente em torno de zero, sem apresentar tendências ou padrões evidentes. O segundo gráfico apresenta os resíduos em função do índice das observações, com o objetivo de verificar possíveis dependências ao longo da sequência dos dados. Assim como no primeiro caso, espera-se que os pontos estejam distribuídos aleatoriamente em torno de zero sem apresentar nenhum padrão sistemático. O terceiro gráfico apresenta a estimativa da densidade dos resíduos, obtida por meio de um estimador kernel. Caso o ajuste do modelo avaliado esteja adequado, espera-se observar um comportamento simétrico. Por fim, o quarto gráfico corresponde

ao *QQ-plot* dos resíduos, utilizado para verificar adequação da suposição de normalidade dos resíduos quantílicos. Neste tipo de gráfico, espera-se que os pontos se posicionem ao longo da linha de referência.

4 Estudos de Simulação

Para avaliar o desempenho dos estimadores de máxima verossimilhança sob a nova classe de modelos Birbaum-Saunders Generalizado, foram realizados estudos de simulação de Monte Carlo com 5000 réplicas. Os tamanhos amostrais considerados são $n = 40, 80, 160$ e 320 . Os valores das variáveis explicativas foram obtidos para o tamanho amostral $n = 40$ e então replicados duas, quatro e oito vezes a fim de obter as matrizes de variáveis explicativas correspondentes aos tamanhos amostrais $n = 40, 80, 160$ e 320 , respectivamente. Segundo, Espinheira, Santos e Cribari-Neto (2017), esse método garante que o grau de heteroscedasticidade seja constante para todos os tamanhos amostrais, permitindo, assim, a comparação dos resultados entre as amostras de tamanho diferente. Em todos os cenários, a função de ligação logarítmica foi utilizada nas estruturas de regressão dos parâmetros μ e σ , enquanto para o parâmetro ν a ligação logito foi a escolhida. Todos os modelos possuem ao menos o intercepto e, conforme o cenário considerado, podem incluir uma ou duas variáveis explicativas. Essas covariáveis são obtidas por meio da geração de variáveis aleatórias com distribuição uniforme padrão e distribuição Bernoulli com parâmetro igual a 0,7.

Com o intuito de assegurar a qualidade das estimativas obtidas durante o processo de simulação, realizaram-se determinados ajustes. As amostras que, por qualquer critério não identificado, resultaram em falha na convergência do algoritmo de estimação tiveram suas respectivas réplicas descartadas e substituídas por novas, de modo a garantir o total de 5000 réplicas. O percentual de réplicas descartadas foi registrado com o objetivo de avaliar a estabilidade das estimativas do novo modelo de regressão proposto e a confiabilidade dos resultados de simulação. Adicionalmente, quando a matriz de variância-covariância não era finita, a réplica também foi descartada. Além disso, as réplicas para as quais as estimativas de ν assumiram valores inferiores a 0,01 ou superiores a 0,99 foram descartadas, uma vez que tais situações conduziam a estimativas inconsistentes dos coeficientes de regressão associados ao submodelo de ν . A imposição desse intervalo limitado para ν teve como objetivo evitar que as estimativas se aproximassem dos limites 0 e 1, o que poderia resultar em coeficientes de regressão excessivamente elevados e, conseqüentemente, comprometer a estabilidade numérica, a convergência dos modelos e o desempenho do procedimento de simulação. Em resumo, sempre que o parâmetro ν assumia valores fora do intervalo estabelecido, a respectiva réplica era descartada.

Por fim, como já mencionado, a arquitetura da biblioteca `gamlss` oferece dois algoritmos principais para a maximização do logaritmo da função de verossimilhança, o algoritmo RS e o algoritmo de CG. Além desses, existe uma abordagem mista, que consiste na combinação de ambos os algoritmos. Segundo Stasinopoulos et al. (2017), o algoritmo RS apresenta, de maneira geral, maior estabilidade e, na maioria dos casos,

também maior velocidade de convergência. Por essa razão, e a partir da experiência obtida durante os estudos de simulação do novo modelo Birnbaum-Saunders Generalizado, todos os cenários de simulação neste estudo foram conduzidos utilizando o algoritmo de Rigby e Stasinopoulos.

Diferentes configurações dos valores dos parâmetros dos coeficientes de regressão foram considerados em cada um dos cenários. Os Cenários 1.1, 1.2 e 1.3 apresentam a estrutura de regressão mais simples, caracterizada por um submodelo que incorpora apenas uma variável explicativa associada ao parâmetro μ . Esses cenários têm como objetivo avaliar o comportamento e o desempenho dos estimadores de máxima verossimilhança sob o novo modelo de regressão BSG em diferentes faixas de valores do parâmetro ν . A partir desses cenários, é possível formular hipóteses sobre como a inclusão do parâmetro ν influencia o comportamento do procedimento de estimação. Nesse sentido, destacam-se os Cenários 1.1 e 1.3, nos quais o parâmetro ν assume, respectivamente, os valores 0,3 e 0,7. O Cenário 1.2, por sua vez, corresponde ao modelo de regressão BS clássico, que é um caso particular do modelo BSG quando ν é fixado em 0,5. Para estes cenários, as estimativas do parâmetro μ variam entre 0,835 e 5,959, com mediana igual a 2,477, enquanto o parâmetro σ é fixado em 0,368. Estrutura semelhante é observada nos Cenários 2.1, 2.2 e 2.3, contudo, a amplitude do parâmetro μ é maior nestes casos, variando entre 0,835 e 14,534 e mediana igual a 3,73, e σ é fixado em 0,273.

Por sua vez, a estrutura dos Cenários 3 e 4 apresentam dois submodelos com apenas uma variável explicativa associados aos parâmetros μ e σ , enquanto o valor do parâmetro ν é fixado, respectivamente, em 0,378 e 0,622. Por meio desses cenários, é possível avaliar o comportamento de modelos mais complexos, nos quais dois parâmetros são modelados simultaneamente. Adicionalmente, é possível também formular hipóteses sobre como o valor do parâmetro ν afeta o comportamento do procedimento de estimação, dada a diferença entre os valores fixados para esse parâmetro. No Cenário 3, a amplitude das estimativas do parâmetro μ é mais elevada, variando entre 3,020 e 6,009, enquanto no Cenário 4 a amplitude se mantém mais próxima de um, variando entre 1,163 e 1,283. Em relação ao parâmetro σ , a amplitude das estimativas dos Cenários 3 e 4 são mais próximas, com medianas iguais a 0,681 e 0,727, respectivamente. Para o parâmetro ν também foram fixados valores distintos para cada cenário, correspondendo a 0,378 e 0,622, respectivamente.

Por último, os Cenários 5 e 6 apresentam a estrutura mais complexa entre os modelos de regressão considerados. Esses cenários incluem submodelos associados a cada um dos parâmetros da distribuição BSG, sendo que, para os parâmetros μ e σ , os submodelos incorporam duas variáveis explicativas, enquanto o submodelo associado ao parâmetro ν inclui apenas uma covariável. Nesses cenários, avalia-se o comportamento do procedimento de estimação ao modelar simultaneamente todos os parâmetros, além disso, é

possível analisar a estabilidade dos submodelos associados ao parâmetro ν . Ressalta-se a diferença na amplitude das estimativas do parâmetro μ , sendo que o Cenário 5 apresenta uma amplitude significativamente maior, o que possivelmente evidencia a estabilidade do procedimento de estimação desse parâmetro em diferentes intervalos. De forma similar, destaca-se a diferença na amplitude das estimativas referentes ao parâmetro σ . Em relação ao parâmetro ν , ambos os cenários apresentam a mesma estrutura de regressão.

A seguir, são descritas de forma detalhada as definições dos cenários de simulação adotados.

Cenário 1.1: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 1,8$, $\beta_2 = -2$, $\gamma_1 = -1$ e $\lambda_1 = -1,390$. De modo que, $\mu \in (0,835; 5,959)$ com $\text{Med}(\mu) = 2,477$; $\sigma = 0,368$ e $\nu = 0,2$.

Cenário 1.2: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 1,8$, $\beta_2 = -2$, $\gamma_1 = -1$ e $\lambda_1 = 0$. De modo que, $\mu \in (0,835; 5,959)$ com $\text{Med}(\mu) = 2,477$; $\sigma = 0,368$ e $\nu = 0,5$.

Cenário 1.3: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 1,8$, $\beta_2 = -2$, $\gamma_1 = -1$ e $\lambda_1 = 1,390$. De modo que, $\mu \in (0,835; 5,959)$ com $\text{Med}(\mu) = 2,477$; $\sigma = 0,368$ e $\nu = 0,8$.

Cenário 2.1: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 2,7$, $\beta_2 = -3,1$, $\gamma_1 = -1,3$ e $\lambda_1 = -1,390$. De modo que, $\mu \in (0,691; 14,536)$ com $\text{Med}(\mu) = 3,73$; $\sigma = 0,273$ e $\nu = 0,2$.

Cenário 2.2: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 2,7$, $\beta_2 = -3,1$, $\gamma_1 = -1,3$ e $\lambda_1 = 0$. De modo que, $\mu \in (0,691; 14,536)$ com $\text{Med}(\mu) = 3,73$; $\sigma = 0,273$ e $\nu = 0,5$.

Cenário 2.3: estrutura de regressão com uma covariável no parâmetro μ , e parâmetros σ e ν constantes. Os valores dos parâmetros foram fixados em $\beta_1 = 2,7$, $\beta_2 = -3,1$, $\gamma_1 = -1,3$ e $\lambda_1 = 1,390$. De modo que, $\mu \in (0,691; 14,536)$ com $\text{Med}(\mu) = 3,73$; $\sigma = 0,273$ e $\nu = 0,8$.

Cenário 3: estrutura de regressão com uma covariável nos parâmetros μ , σ e ν constante Os valores dos parâmetros foram fixados em $\beta_1 = 1,1$, $\beta_2 = 0,7$, $\gamma_1 = -1,1$, $\gamma_2 = 1,6$ e $\lambda_1 = -0,5$. De modo que, $\mu \in (3,020; 6,009)$ com $\text{Med}(\mu) = 4,108$; $\sigma \in (0,337; 1,623)$ com $\text{Med}(\sigma) = 0,681$ e $\nu = 0,378$.

Cenário 4: estrutura de regressão com uma covariável nos parâmetros μ , σ e ν constante Os valores dos parâmetros foram fixados em $\beta_1 = 0,15$, $\beta_2 = 0,1$, $\gamma_1 = -0,9$, $\gamma_2 = 1,3$ e $\lambda_1 = 0,5$. De modo que, $\mu \in (1,163; 1,283)$ com $\text{Med}(\mu) = 1,215$; $\sigma \in (0,411; 1,473)$ com $\text{Med}(\sigma) = 0,727$ e $\nu = 0,622$.

Cenário 5: estrutura de regressão com duas covariáveis nos parâmetros μ e σ , e estrutura de regressão com uma covariável no parâmetro ν . Os valores dos parâmetros foram fixados em $\beta_1 = 1,1$, $\beta_2 = -2,6$, $\beta_3 = 0,7$, $\gamma_1 = -1,2$, $\gamma_2 = 1,1$, $\gamma_3 = 0,5$, $\lambda_1 = -1,2$ e $\lambda_2 = 1,3$. De modo que, $\mu \in (1,085; 16,198)$ com $\text{Med}(\mu) = 4,587$; $\sigma \in (0,368; 4,426)$ com $\text{Med}(\sigma) = 1,616$ e $\nu \in (0,233; 0,522)$ com $\text{Med}(\nu) = 0,35$.

Cenário 6: estrutura de regressão com duas covariáveis nos parâmetros μ e σ , e estrutura de regressão com uma covariável no parâmetro ν . Os valores dos parâmetros foram fixados em $\beta_1 = 1,8$, $\beta_2 = -2$, $\beta_3 = 1$, $\gamma_1 = -1,2$, $\gamma_2 = 1,3$, $\gamma_3 = 1,4$, $\lambda_1 = -1,2$ e $\lambda_2 = 1,3$. De modo que, $\mu \in (0,322; 5,932)$ com $\text{Med}(\mu) = 1,291$; $\sigma \in (0,356; 1,476)$ com $\text{Med}(\sigma) = 0,708$ e $\nu \in (0,233; 0,522)$ com $\text{Med}(\nu) = 0,35$.

Nas Figuras 4 a 13 são apresentados os *boxplots* das estimativas dos coeficientes de regressão utilizando o estimador de máxima verossimilhança, considerando todos os cenários de simulação. A partir da análise destas figuras é possível identificar alguns padrões gerais de comportamento dos estimadores.

Em todos os cenários, e para todos os coeficientes de regressão, observa-se que as medianas das estimativas estão centradas em torno do verdadeiro valor do parâmetro, com o viés entre estimativa e parâmetro diminuindo a medida que aumenta o tamanho amostral. Por exemplo, no Cenário 4, a média das 5000 estimativas para o parâmetro β_1 foram 0,1534, 0,1517, 0,1505, 0,1495 para os tamanhos amostrais $n = 40, 80, 160$ e 320, respectivamente, enquanto o valor verdadeiro de β_1 foi fixado em 0,15. De maneira similar, no Cenário 6 para o parâmetro λ_2 , a média das EMV foram iguais a 1,3522, 1,4167, 1,3557, 1,3132, para, respectivamente, os mesmos tamanhos amostrais, enquanto o valor verdadeiro desse parâmetro foi definido como 1,3. Ilustrando empiricamente que para diferentes configurações de modelos e diferentes parâmetros, as estimativas apresentam um baixo viés em relação ao valor verdadeiro. Além disso, conforme esperado, a variabilidade das estimativas diminui à medida que o tamanho amostral aumenta.

Ressalta-se que apesar do baixo viés das estimativas em relação ao verdadeiro valor, de maneira geral, os parâmetros do vetor γ apresentam a maior distância entre a reta vermelha tracejada, representando o valor populacional do parâmetro, e a mediana das estimativas, indicando uma maior variabilidade associada a esses estimadores.

Em particular, nos Cenários 1.1, 1.2 e 1.3, avalia-se o comportamento e o desempenho dos modelos ajustados para diferentes valores do parâmetro ν , bem como o impacto desse parâmetro na qualidade das estimativas. A distinção entre esses cenários é particu-

larmente relevante, uma vez que o comportamento e a interpretação da distribuição BSG variam em função do valor do parâmetro ν . Em particular, ressaltamos que o caso $\nu = 0,5$ corresponde à distribuição BS clássica. A partir da análise gráfica, verifica-se que, para os parâmetros β_1 e β_2 , o viés é significativamente menor, o que é evidenciado pela linha tracejada em vermelho, a qual se sobrepõe com a mediana das estimativas em todos os níveis de tamanho amostral considerados. No entanto, para os parâmetros γ_1 e λ_1 , nota-se um viés mais acentuado para o tamanho amostral $n = 40$, com uma redução deste a medida que o tamanho amostral aumenta. Ao considerar o número de réplicas descartadas, verifica-se que o desempenho do modelo é influenciado pelo valor do parâmetro ν . Esse efeito torna-se evidente principalmente em amostras de menor tamanho, como $n = 40$ e $n = 80$. No Cenário 1.1, ao fixar $\nu = 0,2$, observa-se um maior percentual de réplicas descartadas, correspondendo a 6,70%, 1,30%, 0,10% e 0% para os tamanhos amostrais $n = 40, 80, 160$ e 320 , respectivamente. Por sua vez, no Cenário 1.2, com ν fixado em 0,5, os percentuais verificados são 0,02%, 0,01%, 0% e 0%. Por fim, no Cenário 1.3, ao fixar $\nu = 0,8$, registram-se 5,50%, 0,09%, 0% e 0%, respectivamente, para os mesmos tamanhos amostrais considerados no Cenário 1.1. Comportamento semelhante é observado nos Cenários 2.1, 2.2 e 2.3, cuja única diferença entre estes se encontra no valor fixado para o parâmetro ν .

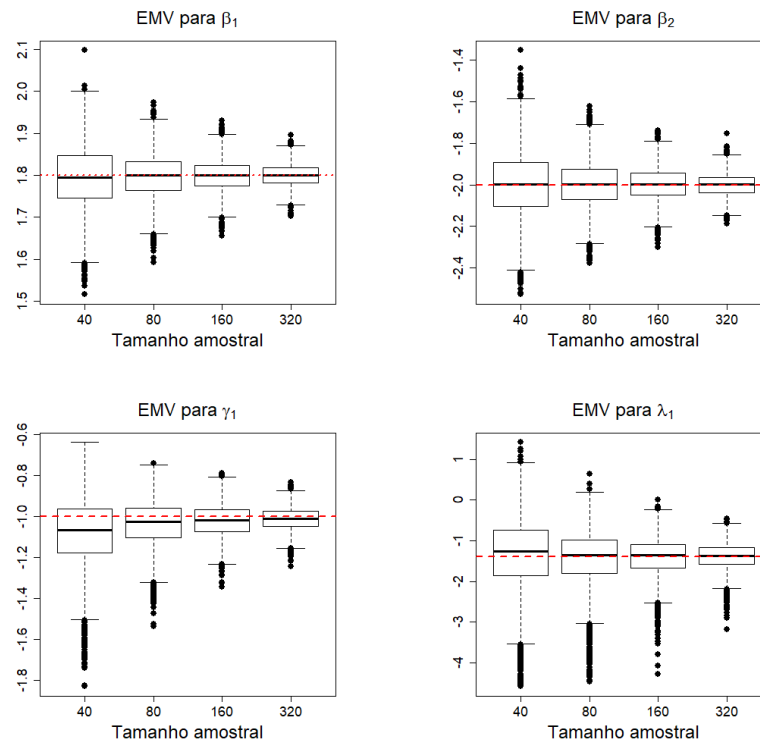


Figura 4: *Boxplots* das estimativas dos parâmetros $\beta_1, \beta_2, \gamma_1, \lambda_1$ sob o Cenário 1.1. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

Em relação aos Cenários 3 e 4, a análise dos *boxplots* apresentados nas Figuras 10

e 11 indica que as estimativas apresentam baixo viés em relação aos respectivos valores verdadeiros. Contudo, em ambos os cenários, observa-se que, para o parâmetro γ_1 , a distância entre a mediana das estimativas e o valor verdadeiro do parâmetro é mais acentuada, principalmente nos casos com tamanho amostral menores. A análise do número de réplicas descartadas indica proporções reduzidas para ambos os cenários, destacando-se o Cenário 4, no qual apenas três réplicas foram descartadas ao longo de todo o processo e para todos os tamanhos amostrais considerados, correspondendo a aproximadamente 0% de réplicas desconsideradas. Nesses cenários o parâmetro ν foi fixado em 0,387 e 0,622, respectivamente, o que possivelmente indica um intervalo de valores para a qual o processo de estimação se mostra mais estável, uma vez que, mesmo quando comparados a cenários mais simples, esses casos apresentaram números bastante reduzidos de réplicas descartadas.

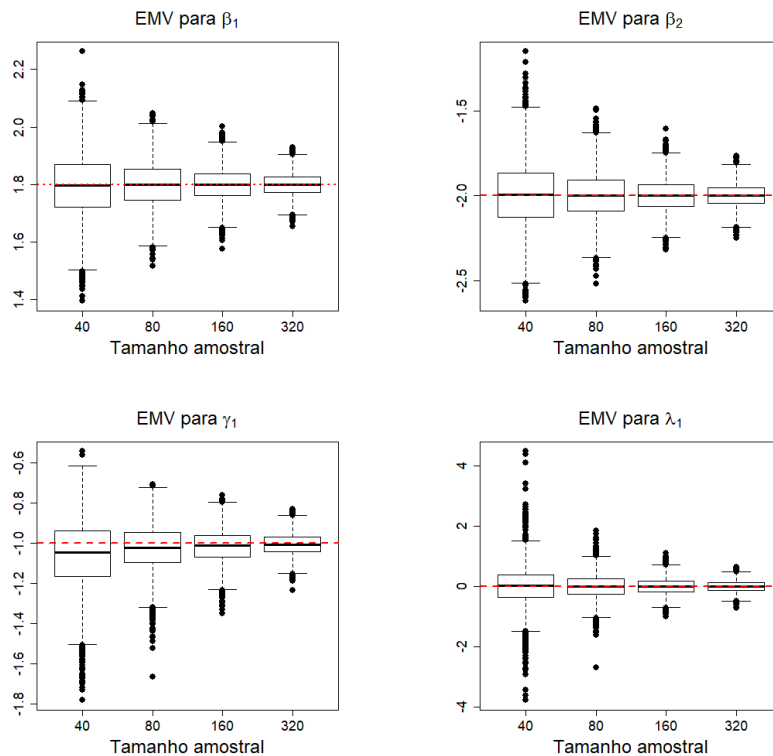


Figura 5: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , λ_1 sob o Cenário 1.2. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

No que se refere aos Cenários 5 e 6, os mais complexos em termos de especificação, verifica-se também um baixo viés das estimativas do método de máxima verossimilhança em relação aos valores verdadeiros dos parâmetros. Novamente, como observado nos demais cenários, o viés do parâmetro γ_1 parece ser o mais acentuado em tamanhos amostrais menores. No Cenário 5, para os parâmetros do vetor β observa-se a presença de estimativas discrepantes, assim como ocorre para os demais parâmetros e nos diferentes cenários analisados. No entanto, nesse caso, nota-se a ocorrência de valores ainda mais

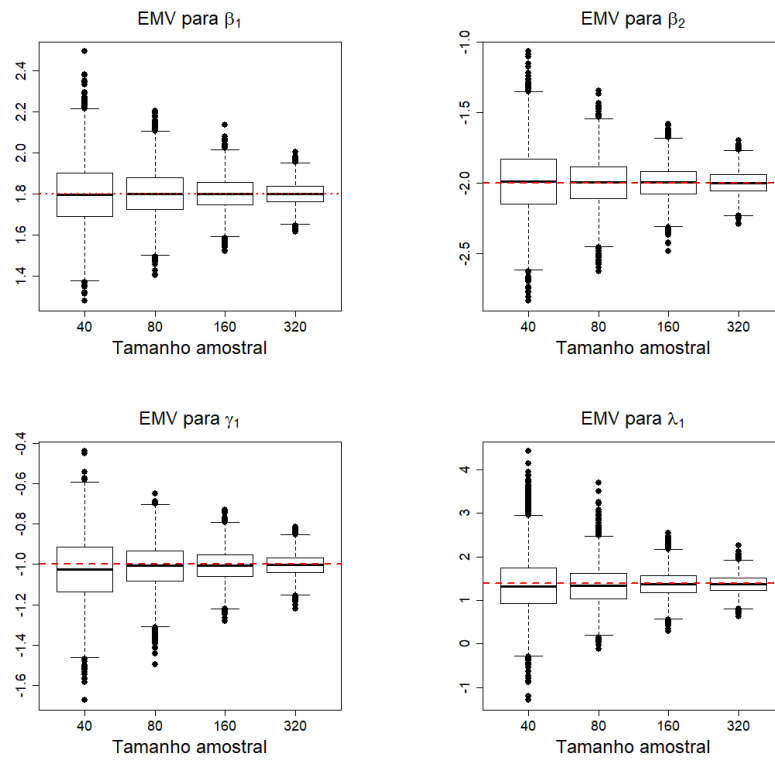


Figura 6: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , λ_1 sob o Cenário 1.3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

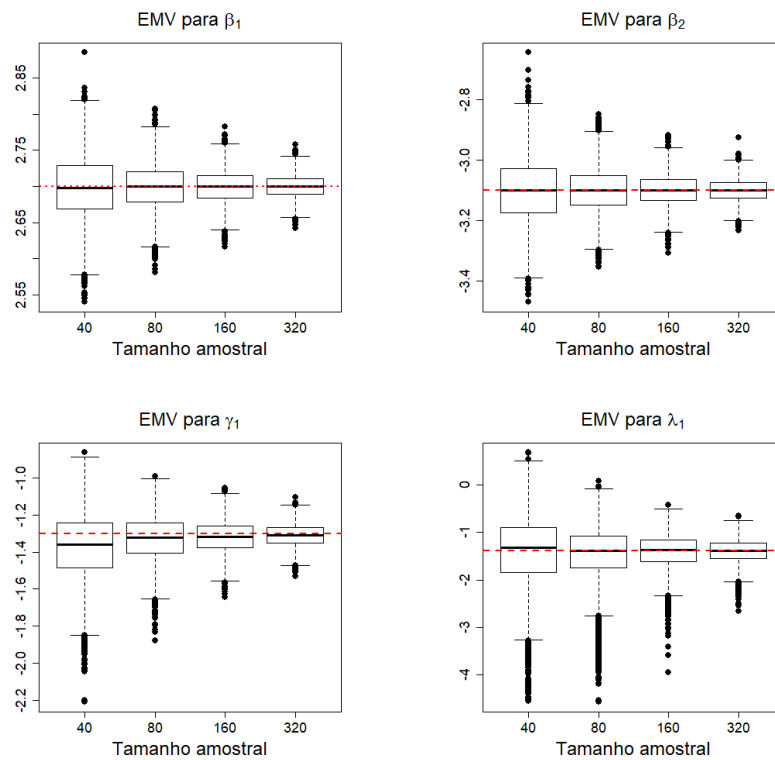


Figura 7: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , λ_1 sob o Cenário 2.1. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

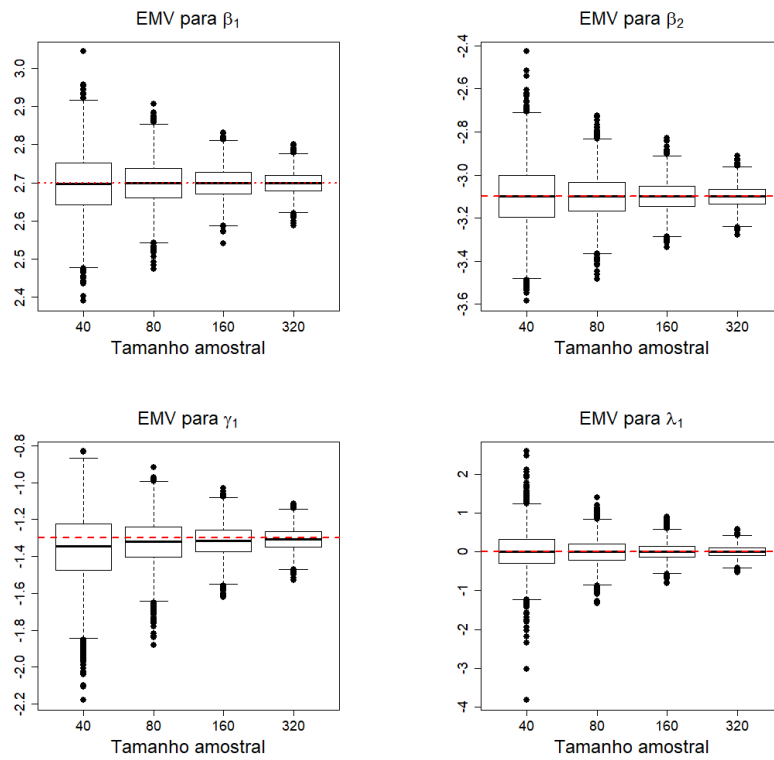


Figura 8: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , λ_1 sob o Cenário 2.2. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

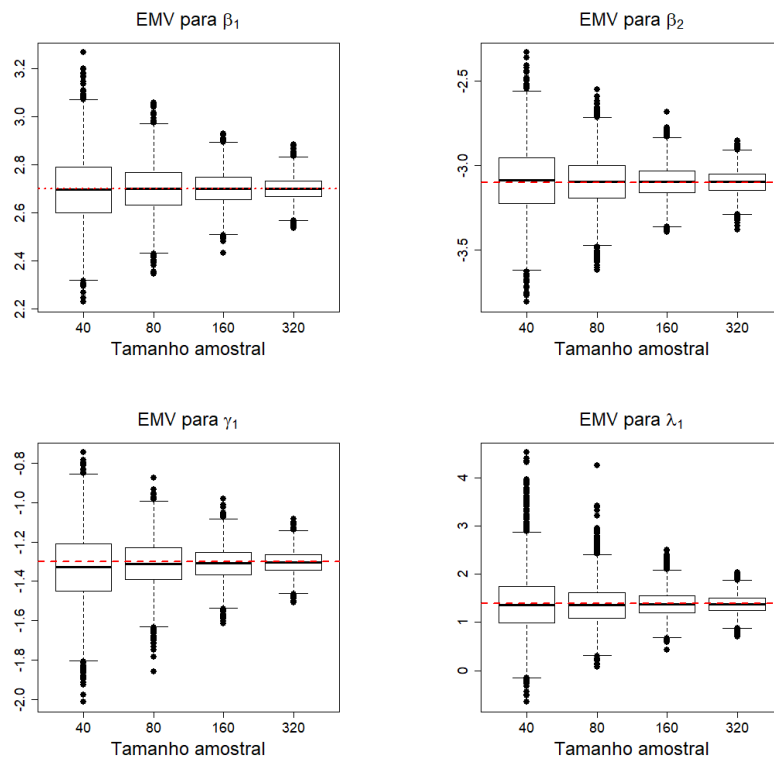


Figura 9: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , λ_1 sob o Cenário 2.3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

extremos, distantes do restante da distribuição, comportamento possivelmente associado ao aumento da complexidade do modelo, decorrente da inclusão de um coeficiente adicional na estimação do parâmetro μ . Com respeito à proporção de réplicas descartadas, observa-se que esses cenários apresentam os maiores percentuais de descarte dentre todos os cenários de simulação considerados, comportamento possivelmente associado também à complexidade desses cenários.

Ao analisar todos os dez cenários de simulação considerados, observam-se indícios da estabilidade e bom desempenho do procedimento de máxima verossimilhança sob o novo modelo de regressão BSG proposto. Visto que, mesmo em cenários com diferentes níveis de complexidade, seja pela quantidade de submodelos especificados ou pelo número de variáveis explicativas incluídas em cada submodelo, as estimativas permaneceram centradas em torno do verdadeiro valor do parâmetro estabelecido, apresentando redução da variabilidade destas estimativas conforme o tamanho amostral cresce. Adicionalmente, em nenhum dos cenários foram observados percentuais elevados de réplicas descartadas, sendo que esse número diminui de forma acentuada em função do tamanho amostral.

Adicionalmente, foi estudado o desempenho do teste de hipótese Wald que testa a hipótese nula $H_0 : \theta_j = \theta_j^0$ contra a hipótese alternativa bilateral $H_1 : \theta_j \neq \theta_j^0$ para o parâmetro fixado e rejeita-se H_0 ao nível nominal α quando

$$|z_j| = \left| \frac{\hat{\theta}_j - \theta_j^0}{\text{se}(\hat{\theta}_j)} \right| > z_{\alpha/2},$$

em que $z_{\alpha/2}$ representa o quantil de ordem $1 - \alpha/2$ da distribuição normal padrão. Neste trabalho, fixou-se o nível nominal em 5%. Nas Tabelas 2, 3, 4 e 5 apresentam-se os níveis empíricos do teste Wald sob H_0 baseados no EMV para os cenários de simulação considerados.

Nas Tabelas 3, 4, 5 e 6 são apresentadas os níveis empíricos do teste Wald para cada um dos cenários de simulação. De maneira geral, observa-se que o desempenho do teste tipo-Wald melhora à medida que o tamanho amostral aumenta. Por exemplo, no Cenário 3, os níveis empíricos do teste tipo-Wald referentes ao parâmetro β_2 são iguais a 7,68%, 6,50%, 6,06% e 5,18%, respectivamente, para $n = 40, 80, 160, 320$.

Para um tamanho amostral $n = 40$, observa-se em alguns cenários que os níveis empíricos do teste associados ao vetor λ assumem valores consideravelmente inferiores aos observados para os demais parâmetros. Por exemplo, nos Cenários 1.2, 2.2, 3 e 6, observam-se, respectivamente, níveis empíricos iguais a 0,16%, 1,24%, 1,7%, 0,64%, 0,70% e 0,90%. Ressalta-se ainda que este comportamento é observado em cenários com diferentes níveis de complexidade, sugerindo algum tipo de instabilidade no processo de estimação do vetor λ , quando o tamanho da amostra é pequeno.

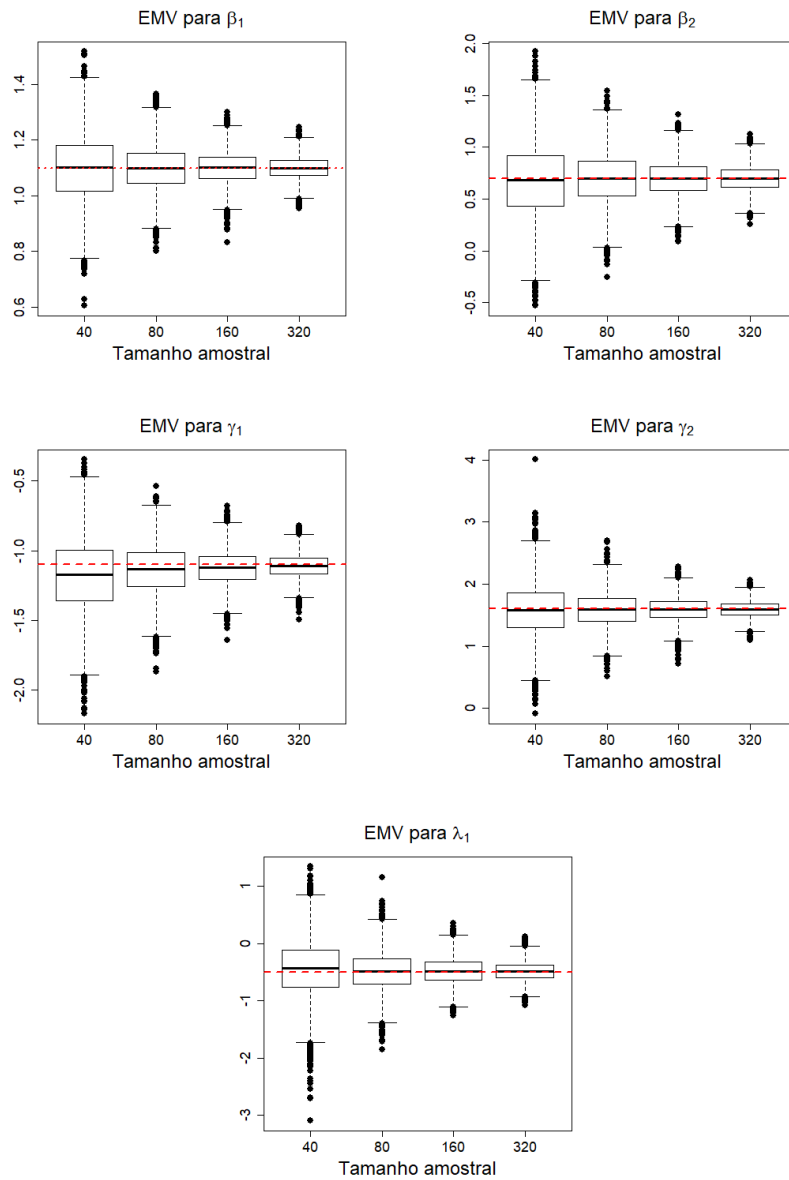


Figura 10: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , γ_2 , λ_1 , λ_2 sob o Cenário 3. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

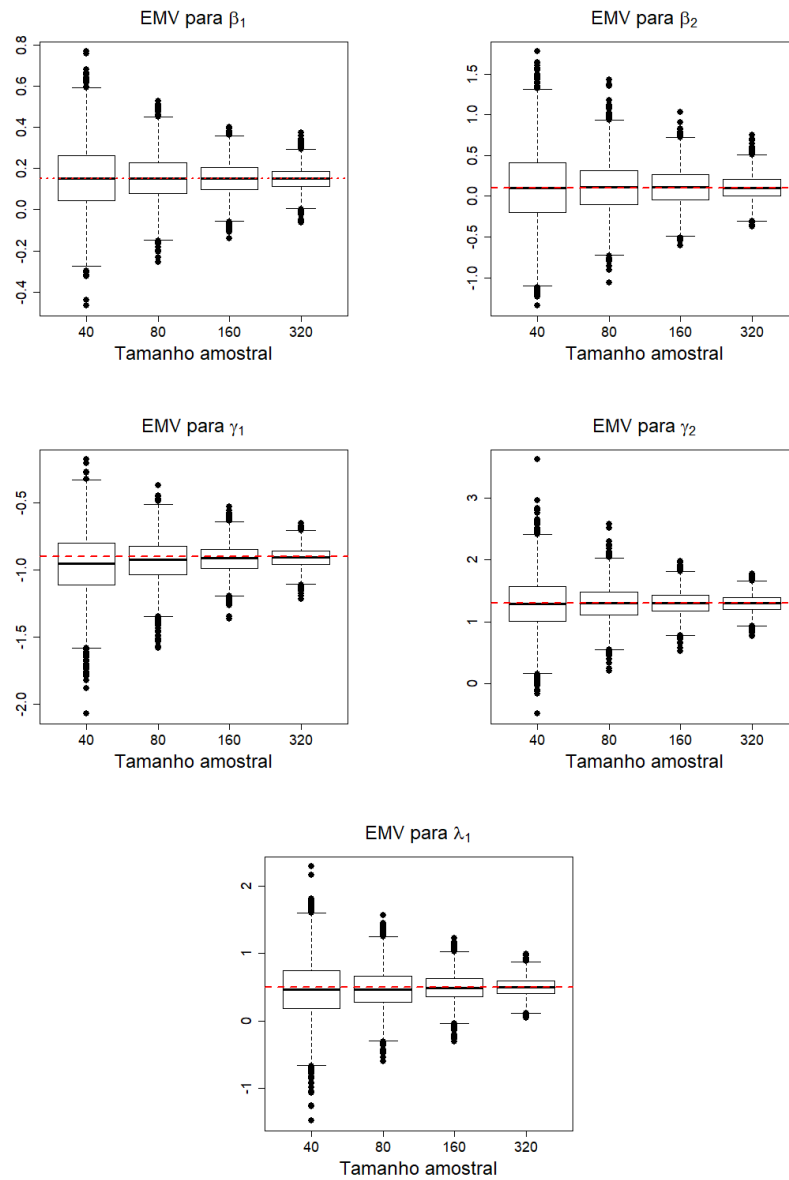


Figura 11: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , γ_1 , γ_2 , λ_1 , λ_2 sob o Cenário 4. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

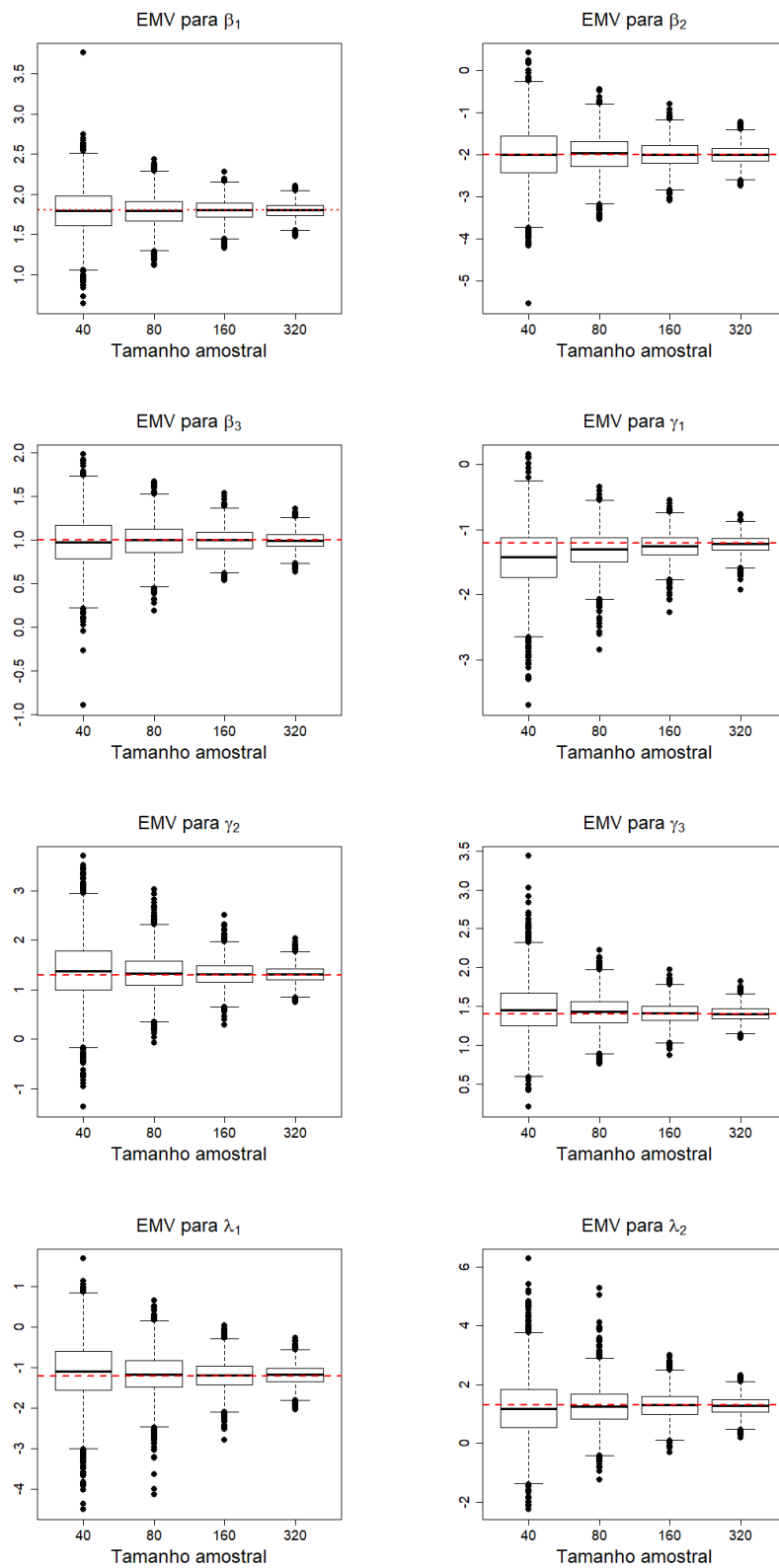


Figura 12: *Boxplots* das estimativas dos parâmetros $\beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2$ sob o Cenário 5. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

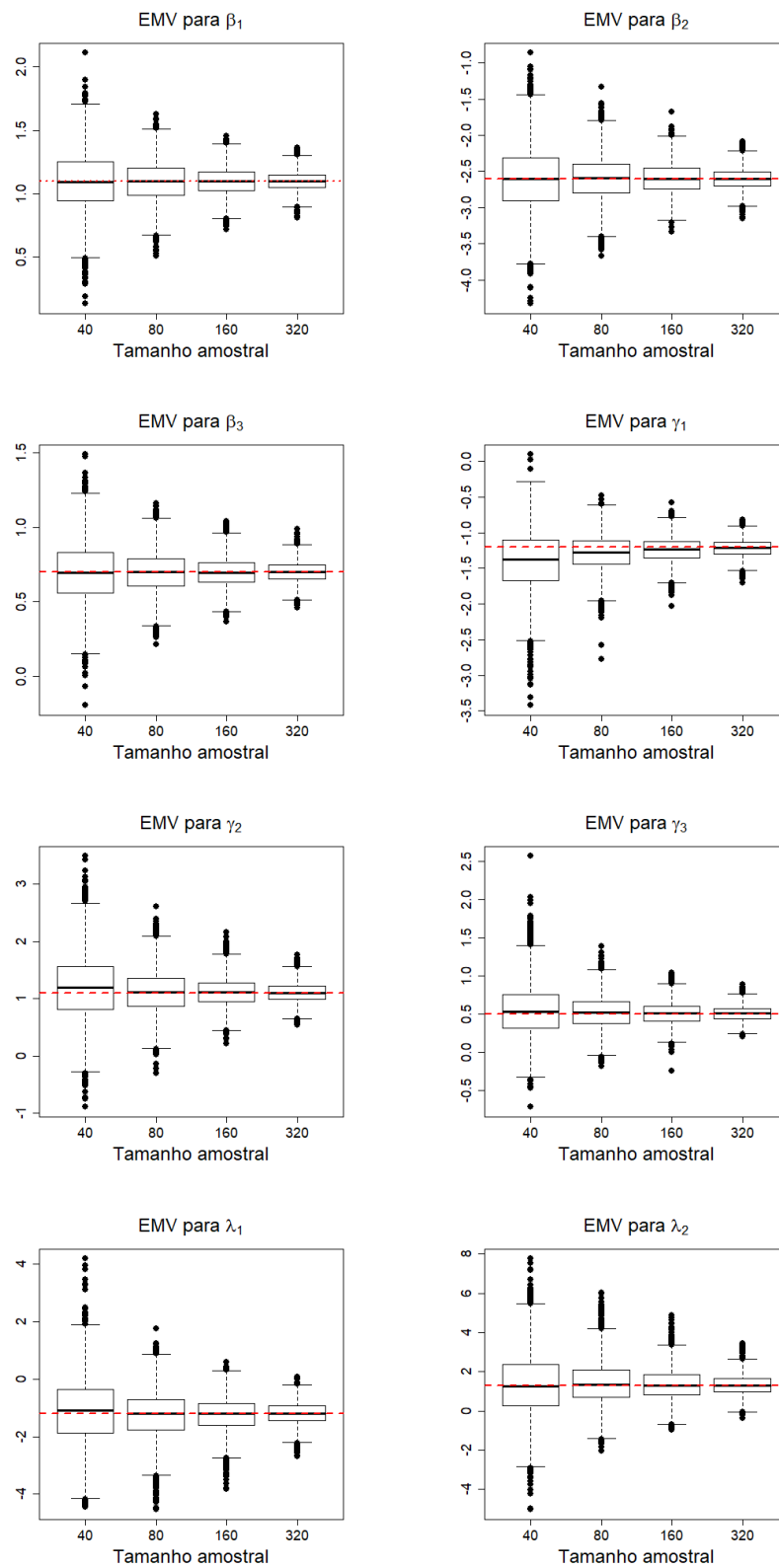


Figura 13: *Boxplots* das estimativas dos parâmetros β_1 , β_2 , β_3 , γ_1 , γ_2 , γ_3 , λ_1 , λ_2 sob o Cenário 6. A linha tracejada em vermelho representa o verdadeiro valor do parâmetro.

Considerando os cenários de simulação 5 e 6, que envolvem estruturas de regressão mais complexas, observa-se certa inconsistência nos níveis empíricos do teste do Wald para os demais coeficientes de regressão. Por exemplo, para o tamanho amostral $n = 40$, em ambos os cenários o nível empírico supera os 10% sobre β_1 e β_2 , e excede 7% para o β_3 . Para γ_1 e tamanho da amostra pequeno, os níveis empíricos chegam a 9,98% e 10%, respectivamente, nos Cenários 5 e 6. Esses resultados indicam que, à medida que a complexidade do modelo aumenta, observa-se maior discrepância entre o nível empírico e o nível nominal.

De modo geral, o teste Wald apresenta níveis empíricos próximos ao nível nominal, com valores cada vez mais próximos de 5% à medida que o tamanho amostral aumenta, comportamento já esperado, dado que o teste avaliado baseia-se em quantis aproximados da distribuição nula da estatística do teste z_j . Além disso, conforme esperado para o teste tipo-Wald, na maioria dos cenários observa-se uma proporção de rejeições mais elevada nos menores tamanhos amostrais, a qual tende a diminuir à medida que o tamanho da amostra aumenta.

Possivelmente por conta da complexidade da estrutura do modelo de regressão considerado, os Cenários 5 e 6 apresentam o pior desempenho do teste tipo-Wald para o tamanho amostral pequeno, $n = 40$. Em ambos os cenários, verificou-se que os parâmetros do vetor β , apresentaram o pior desempenho se comparado com os demais, com o nível empírico ultrapassando os 10% para os parâmetros β_1 e β_2 , por exemplo. Em relação ao vetor γ , em particular o γ_1 , apresentam resultados inferiores em relação aos demais, com níveis empíricos próximos ou superiores a 10%. Como já discutido, estes cenários também apresentam um desempenho inferior para os parâmetros λ_1 e λ_2 , com níveis empíricos com valores iguais a 0,7% e 0,9%, respectivamente.

Tabela 5: Níveis empíricos sob H_0 do teste tipo-Wald para o Cenário 6 ao nível nominal de 5%.

	β_1	β_2	β_3	γ_1	γ_2	γ_3	λ_1	λ_2
$n = 40$	0,1008	0,1014	0,0884	0,0998	0,0552	0,0754	0,0070	0,0090
$n = 80$	0,0666	0,0702	0,0712	0,0690	0,0506	0,0584	0,0230	0,0228
$n = 160$	0,0672	0,0592	0,0664	0,0646	0,0542	0,0538	0,0448	0,0444
$n = 320$	0,0592	0,0486	0,0560	0,0548	0,0490	0,0520	0,0444	0,0400

Por fim, a Tabela 6 apresenta o percentual de réplicas descartadas para atingir as 5000 réplicas estipuladas no estudo de simulação. Verifica-se que, à medida que o tamanho amostral aumenta, a proporção de amostras descartadas diminui de forma expressiva. Para o tamanho amostral $n = 40$, observa-se que o Cenário 6 apresenta a maior proporção de réplicas descartadas, superior a 8%. Em seguida, destacam-se os Cenários 5, 1.1 e 2.1,

Tabela 2: Níveis empíricos sob H_0 do teste tipo-Wald nos Cenários 1.1, 1.2, 1.3, 2.1, 2.2 e 2.3 ao nível nominal de 5%.

	Cenário 1.1				Cenário 1.2			
	β_1	β_2	γ_1	λ_1	β_1	β_2	γ_1	λ_1
$n = 40$	0,0638	0,0676	0,0692	0,0586	0,0694	0,0682	0,0628	0,0016
$n = 80$	0,0602	0,0548	0,0548	0,0498	0,0578	0,0552	0,0526	0,0246
$n = 160$	0,0492	0,0562	0,0608	0,0514	0,0518	0,0554	0,0548	0,0390
$n = 320$	0,0518	0,0488	0,0528	0,0430	0,0504	0,0530	0,0500	0,0378
	Cenário 1.3				Cenário 2.1			
	β_1	β_2	γ_1	λ_1	β_1	β_2	γ_1	λ_1
$n = 40$	0,0666	0,0644	0,0486	0,0538	0,0660	0,0674	0,0698	0,0618
$n = 80$	0,0608	0,0612	0,0548	0,0592	0,0616	0,0540	0,0562	0,0450
$n = 160$	0,0506	0,0524	0,0552	0,0432	0,0480	0,0564	0,0606	0,0490
$n = 320$	0,0508	0,0528	0,0446	0,0434	0,0512	0,0476	0,0494	0,0404
	Cenário 2.2				Cenário 2.3			
	β_1	β_2	γ_1	λ_1	β_1	β_2	γ_1	λ_1
$n = 40$	0,0674	0,0668	0,0666	0,0124	0,0706	0,0650	0,0546	0,0538
$n = 80$	0,0622	0,0582	0,0540	0,0370	0,0614	0,0624	0,0526	0,0546
$n = 160$	0,0536	0,0560	0,0526	0,0438	0,0512	0,0536	0,0510	0,0474
$n = 320$	0,0546	0,0532	0,0510	0,0434	0,0552	0,0536	0,0478	0,0472

Tabela 3: Níveis empíricos sob H_0 do teste tipo-Wald dos Cenários 3 e 4 ao nível nominal de 5%.

	Cenário 3					Cenário 4				
	β_1	β_2	γ_1	γ_2	λ_1	β_1	β_2	γ_1	γ_2	λ_1
$n = 40$	0,0794	0,0752	0,0652	0,0688	0,0170	0,0804	0,0774	0,0716	0,0726	0,0272
$n = 80$	0,0648	0,0592	0,0594	0,0524	0,0336	0,0558	0,0600	0,0588	0,0556	0,0440
$n = 160$	0,0538	0,0540	0,0554	0,0516	0,0410	0,0546	0,0570	0,0566	0,0532	0,0482
$n = 320$	0,0554	0,0492	0,0480	0,0490	0,0460	0,0540	0,0504	0,0554	0,0534	0,0458

Tabela 4: Níveis empíricos sob H_0 do teste tipo-Wald para o Cenário 5 ao nível nominal de 5%.

	β_1	β_2	β_3	γ_1	γ_2	γ_3	λ_1	λ_2
$n = 40$	0,1174	0,1152	0,0768	0,1000	0,0562	0,0972	0,0266	0,0298
$n = 80$	0,0824	0,0830	0,0698	0,0768	0,0516	0,0726	0,0452	0,0460
$n = 160$	0,0618	0,0686	0,0566	0,0662	0,0540	0,0646	0,0522	0,0488
$n = 320$	0,0608	0,0554	0,0518	0,0564	0,0504	0,0454	0,0524	0,0504

com percentuais de 6,9%, 6,7% e 5,5%, respectivamente. Destaca-se que os Cenários 1.1 e 2.1 são modelos mais simples, possuindo estrutura de regressão apenas para o parâmetro da mediana, além disso, um aspecto em comum entre esses dois cenários é o parâmetro ν fixado em 0,2. Um segundo ponto de destaque refere-se aos Cenários 1.2 e 2.2, os quais ν foi fixado em 0,5, caso especial da distribuição BSG que corresponde a distribuição BS clássica. Observa-se que estes cenários apresentam baixas proporções de réplicas descartadas, sugerindo maior estabilidade no processo de estimação em comparação com os cenários em que o parâmetro ν não é fixado em 0,5. Por fim, ressalta-se que para o tamanho amostral $n = 320$, todos os cenários avaliados apresentam percentuais de réplicas

descartadas inferiores a 1%, o Cenário 5 possui o maior percentual com apenas 0,03%.

Tabela 6: Percentual de réplicas descartadas de acordo com os cenários de simulação.

Cenário	$n = 40$	$n = 80$	$n = 160$	$n = 320$
1.1	6,70	1,30	0,10	0,00
1.2	0,02	0,01	0,00	0,00
1.3	0,09	0,00	0,00	0,00
2.1	5,50	0,09	0,00	0,00
2.2	0,02	0,02	0,01	0,00
2.3	1,20	0,04	0,03	0,01
3	1,50	0,09	0,03	0,00
4	0,00	0,00	0,00	0,00
5	6,90	4,50	2,10	0,03
6	8,10	1,60	0,05	0,00

5 Aplicações

Neste capítulo são apresentadas duas aplicações a dados reais utilizando o modelo de regressão BSG proposto. Além de ilustrar a aplicação do novo modelo, são feitas comparações do ajuste do modelo baseado na distribuição Birnbaum-Saunders Generalizada em relação ao ajuste do modelo obtido pela distribuição Birnbaum-Saunders clássica. Como discutido anteriormente, o modelo de regressão BS constitui um caso particular do novo modelo de regressão BSG proposto. Dessa maneira, por meio da implementação do modelo BSG, é possível também ajustar um modelo BS em sua formulação original, fixando o parâmetro ν igual a 0,5. Esta comparação busca evidenciar situações em que o novo modelo de regressão BSG proposto proporciona um ajuste mais flexível, destacando assim a importância da inclusão do terceiro parâmetro na modelagem. Ambos os conjuntos de dados podem ser obtidos a partir do pacote `GMLsData` desenvolvido por Dunn e Smyth (2022) disponível no *software* R, através dos comandos `data(cheese)` e `data(dental)`.

5.1 Dados *Cheese*

Um estudo foi conduzido em La Trobe Valley em Victoria, na Austrália, com o objetivo de investigar a relação entre a composição química do queijo e seu sabor. Com este objetivo, analisou-se uma amostra composta por trinta queijos, sendo que, em cada um destes queijos foram mensurados as concentrações de ácido acético (**Acetic**), ácido láctico (**Lactic**) e sulfeto de hidrogênio (**H₂S**). A cada queijo também foi atribuída uma pontuação referente ao seu sabor (**Taste**), determinada por um painel de juízes. A pontuação final é uma combinação dos escores de todos os avaliadores. O objetivo é modelar a mediana deste escore combinado, representada pela variável **Taste**, em função da composição química do queijo, com base nas mensurações obtidas. Para informações adicionais sobre estes dados veja, Moore, McCabe e Craig (2009).

A seleção de variáveis foi feita pelo método *stepwise*, considerando-se o critério de informação de Akaike (AIC) como medida de comparação entre os modelos. O procedimento *stepwise* consiste em adicionar ou remover variáveis explicativas, a cada passo, a fim de minimizar o critério de informação adotado. Assim, após a condução do processo de seleção, a estrutura de regressão ajustada para a variável **Taste** é dada por:

$$\begin{aligned}\log(\mu_i) &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i}, \\ \log(\sigma_i) &= \gamma_1, \\ \log\left(\frac{\nu_i}{1 - \nu_i}\right) &= \lambda_1,\end{aligned}$$

em que $i = 1, \dots, 30$, X_2 e X_3 correspondem, respectivamente, ao logaritmo da concentração de sulfeto de hidrogênio e a concentração ácido lático. Os parâmetros β_1 , β_2 e β_3 representam os coeficientes de regressão associados ao submodelo da mediana μ . Em relação aos submodelos de σ e ν , os coeficientes de regressão associados a eles são, respectivamente, γ_1 e λ_1 . Contudo, como apenas o intercepto foi incluído nestes dois últimos submodelos, esses coeficientes permanecem constantes para todas as observações.

A Figura 14 apresenta os gráficos de dispersão da variável dependente **Taste** em função das variáveis explicativas, bem como o *boxplot* da variável resposta **Taste**. A partir da Figura 14(a), 14(b) e 14(c), nota-se um comportamento linear positivo, conforme a concentração de cada componente químico aumenta, a pontuação associada ao sabor também aumenta. Na Figura 14(d), observa-se que não há indicativo de presença de valores atípicos na distribuição da variável resposta. Além disso, o *boxplot* revela indícios de assimetria, sugeridos pela posição deslocada da mediana e pela desigualdade entre os comprimentos dos quartis. Ressalta-se que a variável (H_2S) foi transformada por meio da aplicação do logaritmo, objetivando linearizar a relação com a variável resposta.

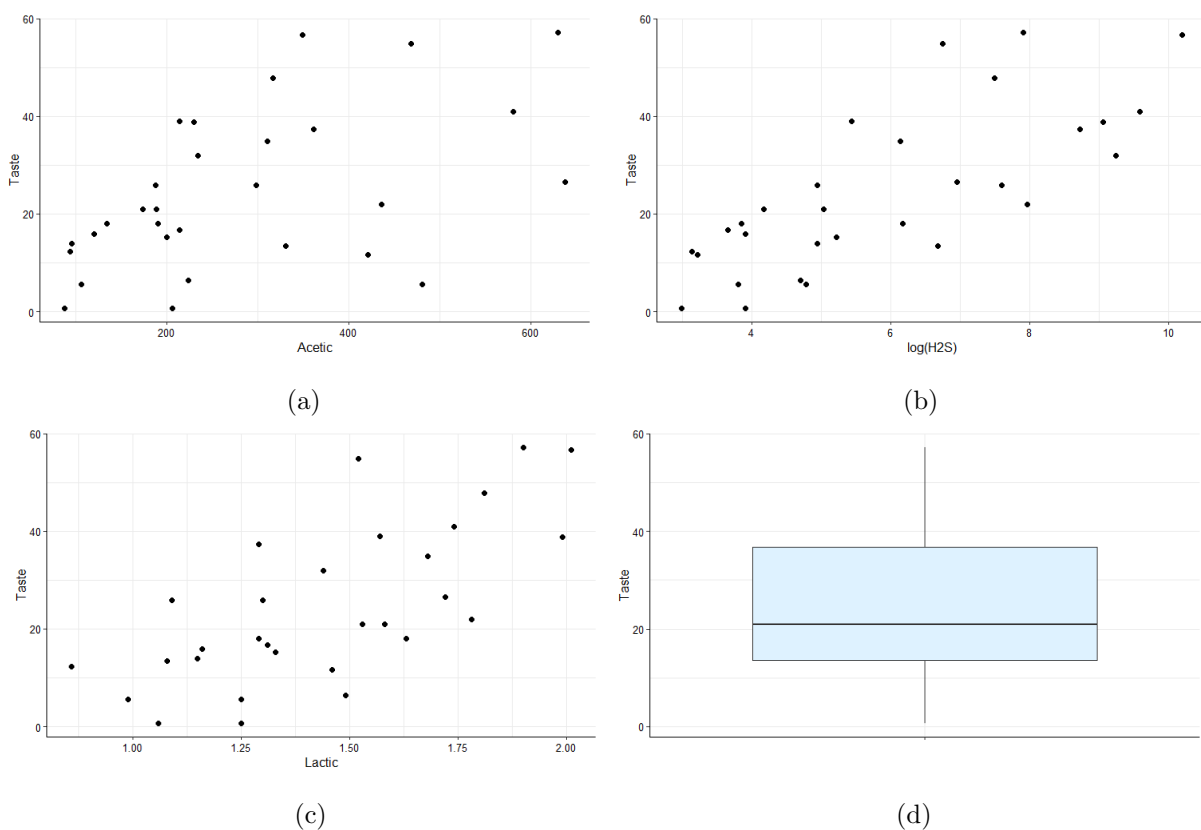


Figura 14: Gráfico de dispersão da variável **Taste** em função da variável **Acetic** (a), gráfico de dispersão da variável **Taste** em função da variável $\log(H_2S)$ (b), gráfico de dispersão da variável **Taste** em função da variável **Lactic** (c) e *boxplot* da variável resposta **Taste** (d) – dados *Cheese*.

A Tabela 7 apresenta os valores das estimativas dos coeficientes de regressão, seus erros padrão, as estatísticas do teste de Wald e o respectivos p -valores. Ressalta-se nova-

mente, que a especificação do modelo proposto apresenta estrutura de regressão somente para o submodelo associado à mediana da distribuição, os outros parâmetros possuem apenas um intercepto. Em relação aos coeficientes de regressão dos modelos, considerando um nível de significância de 5%, verificam-se resultados e interpretações distintas entre os dois modelos considerados. Para o modelo BS clássico, todos os coeficientes são significativamente iguais a zero, indicando que as covariáveis concentração de sulfeto de hidrogênio e concentração de ácido láctico não são relevantes para explicar a mediana da resposta **Taste**. Em contrapartida, no modelo de regressão sob a distribuição BSG, os coeficientes de regressão associados à mediana da resposta mostraram-se significativamente diferentes de zero, indicando que se relacionam com a mediana do sabor do queijo. Os resultados do estudo de simulação indicam a necessidade de interpretação cautelosa dos p -valores em razão das diferenças verificadas entre os níveis empíricos e o nível nominal do teste, sobretudo em amostras de menor tamanho, como a situação da presente aplicação.

Tabela 7: Estimativas, erros-padrão, estatísticas- z e valores- p para os coeficientes de regressão estimados – dados *Cheese*.

	Birnbbaum-Saunders				Birnbbaum-Saunders Generalizada			
	Estimativa	SE	estat- z	valor- p	Estimativa	SE	estat- z	valor- p
β_1	-0,504	1,169	-0,431	0,669	0,920	0,442	2,082	0,047
β_2	0,308	0,164	1,878	0,071	0,154	0,051	3,011	0,005
β_3	0,981	0,723	1,356	0,186	0,834	0,351	2,377	0,025
γ_1	-0,122	0,239	-0,511	0,613	0,502	0,193	2,594	0,015
λ_1	-	-	-	-	-2,313	0,795	-2,909	0,007

Na Figura 15 são apresentados os gráficos de diagnóstico referentes ao modelo de regressão BS, nota-se a partir desses gráficos que o ajuste do modelo não foi adequado. Na Figura 15(a), observa-se uma concentração dos resíduos quantílicos, acompanhada pelo aumento da variabilidade a medida que o valor ajustado de μ aumenta. Pela Figura 15(b), nota-se também um aumento na dispersão dos resíduos. Por último, Nas Figuras 15(c) e 15(d), notam-se desvios claros em relação à normalidade dos resíduos quantílicos, devido à assimetria da densidade estimada, por conta da cauda mais alongada à esquerda, e à presença de múltiplos afastamentos da linha de referência no QQ -plot.

Na Figura 16 é apresentado o *worm plot* dos resíduos quantílicos referente ao modelo Birnbbaum-Saunders, a fim de avaliar a adequação do ajuste do modelo. A análise deste gráfico evidencia, de forma semelhante os demais gráficos de diagnóstico, a inadequação do modelo de regressão proposto aos dados. Os pontos estão afastados da origem, o que indica problemas na estimação do parâmetro de localização. Além disso, nota-se

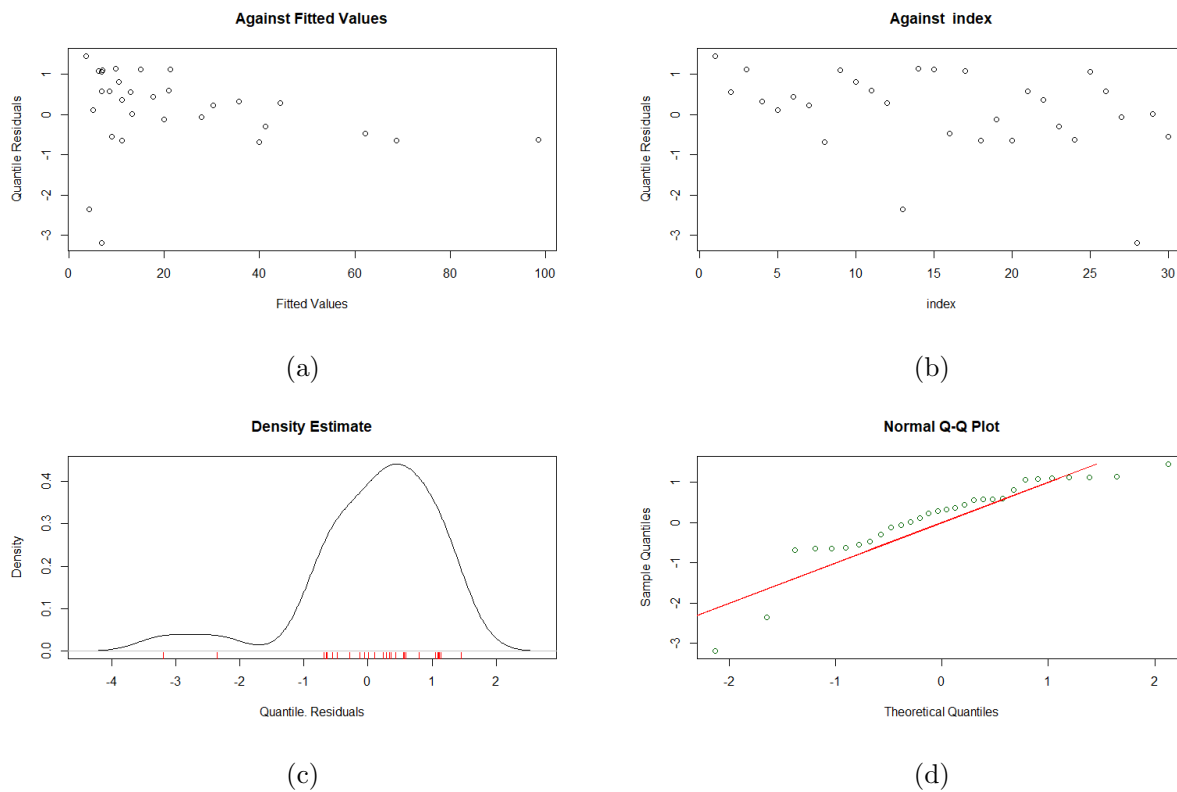


Figura 15: Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BS – dados *Cheese*.

padrão semelhante a uma forma de “U” invertido, o que sugere falta de ajuste em relação à simetria. Assim, com base na análise dos resíduos sob o modelo de regressão BS clássico, conclui-se que esse modelo não parece se ajustar bem aos dados.

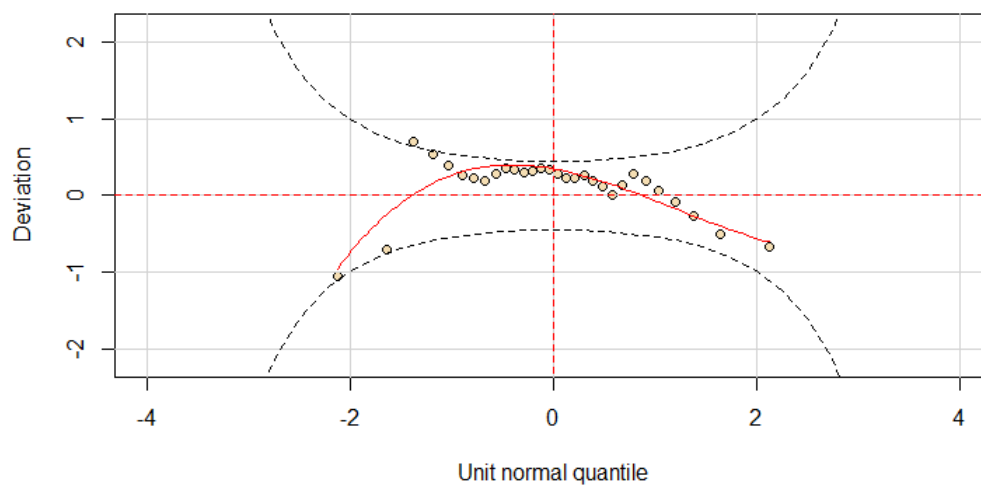


Figura 16: *Worm plot* do ajuste do modelo baseado na distribuição BS – dados *Cheese*.

A Figura 17 exibe os gráficos de diagnóstico correspondentes ao modelo de regressão BSG. A partir da análise da Figura 17(a) observa-se que, os resíduos agora estão menos concentrados e mais centrados em torno de zero. Adicionalmente, observa-se uma aparente diminuição da dispersão dos resíduos conforme os valores ajustados aumentam. Em contrapartida, na Figura 17(b) observa-se uma tendência decrescente nos resíduos quantílicos, de modo que, a medida que o índice aumenta, os valores dos resíduos diminuem. Enquanto nas Figuras 17(c) e 17(d), observa-se um comportamento dos resíduos mais próximo à normalidade, uma vez que o observam-se poucos desvios em relação à reta de referência no *QQ-plot*, além da densidade estimada dos resíduos quantílicos estar mais próxima de um comportamento simétrico.

Finalmente, na Figura 18 é apresentado o *worm plot* do modelo ajustado sob a distribuição Birbaum-Saunders Generalizada. Examinando o comportamento do *worm plot*, verifica-se que todos os pontos estão entre as elipses, centrados em torno da reta horizontal e sem um padrão evidente. Portanto, não há indícios que contrariem a adequação do modelo de regressão BSG.

Dessa forma, ao comparar o ajuste entre os modelos BS e BSG por meio da análise dos resíduos, observa-se uma melhora significativa na adequação do novo modelo proposto, ilustrando o impacto da inclusão do parâmetro ν no ajuste aos dados.

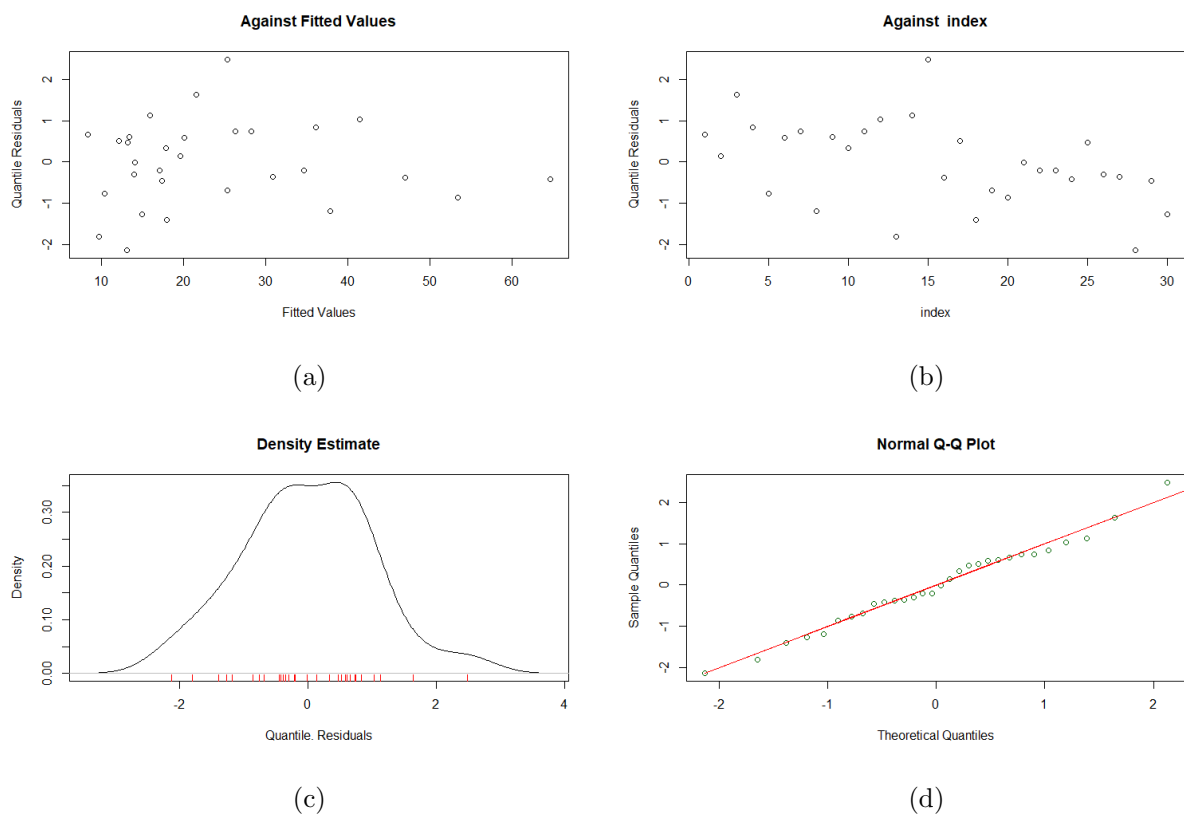


Figura 17: Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BSG— dados *Cheese*.

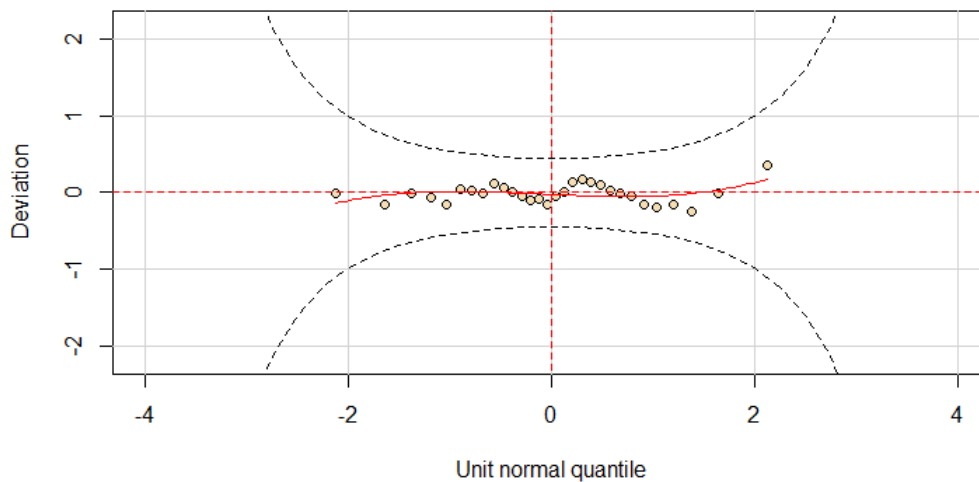


Figura 18: *Worm plot* do ajuste do modelo baseado na distribuição BSG – dados *Cheese*.

Portanto, com base nas estimativas de ambos os modelos e nos resultados da análise de diagnóstico, torna-se apropriada a interpretação dos coeficientes de regressão estimados sob o modelo BSG. Assim, em razão do uso da função de ligação logaritmo no submodelo de μ , a interpretação das estimativas é dada pela inversa de seus valores, isso é, a exponencial das estimativas dos coeficientes de regressão. A exponencial do intercepto, $\hat{\beta}_1$, corresponde ao valor da mediana quando as demais covariáveis são fixados em zero. Já as exponências dos coeficientes associados às variáveis explicativas representam fatores multiplicativos na escala da mediana da resposta. Para o intercepto do submodelo de μ , tem-se que $e^{\hat{\beta}_1} = e^{0,920} \approx 2,509$, no entanto, este valor não possui uma interpretação, uma vez que não é coerente fixar os demais fatores em zero. Em relação a $\hat{\beta}_2$, o valor positivo da estimativa indica que o aumento da concentração de sulfeto de hidrogênio na composição do queijo está relacionada com o aumento da pontuação referente ao seu sabor, representada pela variável *Taste*. Logo, como $e^{\hat{\beta}_2} = e^{0,154} \approx 1,166$, tem-se que a mediana da resposta *Taste* tem um aumento de cerca de 16,6% para cada aumento na unidade da concentração de sulfeto de hidrogênio. De maneira similar, como $e^{\hat{\beta}_3} = e^{0,834} \approx 2,302$, o aumento unitário na concentração de ácido láctico está associado a um aumento de 130% na mediana da resposta, mantendo as demais variáveis constantes.

Como o submodelo associado ao parâmetro σ possui somente intercepto, não há uma interpretação direta desse coeficiente de regressão. Contudo, considerando que a função de ligação adotada para esse parâmetro é o logaritmo, tem-se que $\hat{\sigma} = e^{\hat{\gamma}_1} = e^{0,502} \approx 1,652$. De maneira similar, como o submodelo associado ao parâmetro ν possui somente o intercepto, o coeficiente de regressão não possui uma interpretação direta. No entanto, considerando que a função de ligação adotada é a logito, tem-se $\hat{\nu} = \text{logito}^{-1}(\hat{\lambda}_1) =$

$$1/1 + e^{-\hat{\lambda}_1} = 1/1 + e^{2,313} \approx 0,09.$$

Ressalta-se que o parâmetro λ_1 é significativamente diferente de zero. Assim, em razão da função de ligação logito, esse resultado implica que o parâmetro ν é estatisticamente distinto de 0,5. Logo, o modelo ajustado não é equivalente a um modelo BS clássico. Esse resultado indica a relevância do parâmetro ν na adequabilidade do modelo aos dados. Além disso, a estimativa obtida para ν é aproximadamente igual a 0,09.

Por fim, em concordância com a análise de diagnóstico dos resíduos, os critérios de informação de Akaike e bayesiano apresentados na Tabela 8 indicam preferência pelo modelo baseado na distribuição BSG em relação ao modelo BS. Nesse contexto, é adequado utilizar estes critérios de informação, dado que os modelos ajustados aos dados são encaixados.

Tabela 8: Medidas de informação – dados *Cheese*.

Modelo	AIC	BIC
1 Birnbaum-Saunders	252,27	257,87
2 Birnbaum-Saunders Generalizada	218,07	235,07

5.2 Dados *Dental*

Os dados analisados nesta subseção referem-se a informações provenientes de 90 países acerca da condição dentária de crianças de 12 anos de idade. Para cada país, tem-se o seu nível de industrialização (*Indus*), uma variável categórica dividida em dois níveis: industrialização e não industrializado, bem como o consumo médio anual de açúcar (*Sugar*), em quilogramas por pessoa. Adicionalmente, para cada país também foi coletado a estimativa média do número de dentes afetados por cárie, extração ou restauração (*DMFT*). O objetivo é modelar a quantidade de dentes comprometidos em função do consumo médio de açúcar e o nível de industrialização. Detalhes adicionais sobre este conjunto de dados podem ser encontrados em Woodward e Walker (1994).

Após a realização do processo de seleção de variáveis por meio do método *stepwise*, a estrutura do modelo de regressão ajustado para a estimativa da média do número de dentes afetados por cárie, extração ou restauração é dada por:

$$\begin{aligned}\log(\mu_i) &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i}, \\ \log(\sigma_i) &= \gamma_1 + \gamma_3 X_{3i}, \\ \log\left(\frac{\nu_i}{1 - \nu_i}\right) &= \lambda_1,\end{aligned}$$

em que $i = 1, \dots, 90$, X_2 e X_3 correspondem, respectivamente, ao consumo médio de açúcar e uma variável indicativa de se o país é ou não industrializado. Os parâmetros β_1 , β_2 e β_3 representam os coeficientes de regressão associados submodelo da mediana μ . Em relação aos submodelos de σ e ν , os coeficientes de regressão associados a eles são, respectivamente, γ_1 , γ_3 e λ_1 . Diferentemente da aplicação anterior, desta vez a estrutura do modelo de regressão possui submodelos associados a dois parâmetros, μ e σ .

Na Figura 19 são apresentados dois *boxplots* e um gráfico de dispersão relacionados à variável resposta DMFT. A Figura 19(a) indica que o número médio de dentes afetados tende a ser maior em países industrializados, por outro lado, em países não industrializados observa-se uma maior variabilidade, evidenciada pela amplitude do *boxplot*. Por sua vez, na Figura 19(b), observa-se um comportamento linear positivo entre DMFT e Sugar, indicando que maiores níveis de consumo de açúcar estão associados a um maior número de dentes tratados. Além disso, nota-se também que a dispersão da variável DMFT tende a aumentar à medida que o consumo médio de açúcar se eleva. Por último, a Figura 19(c) indica que a distribuição da variável DMFT parece apresentar um comportamento aproximadamente simétrico, embora sejam observados alguns valores discrepantes.

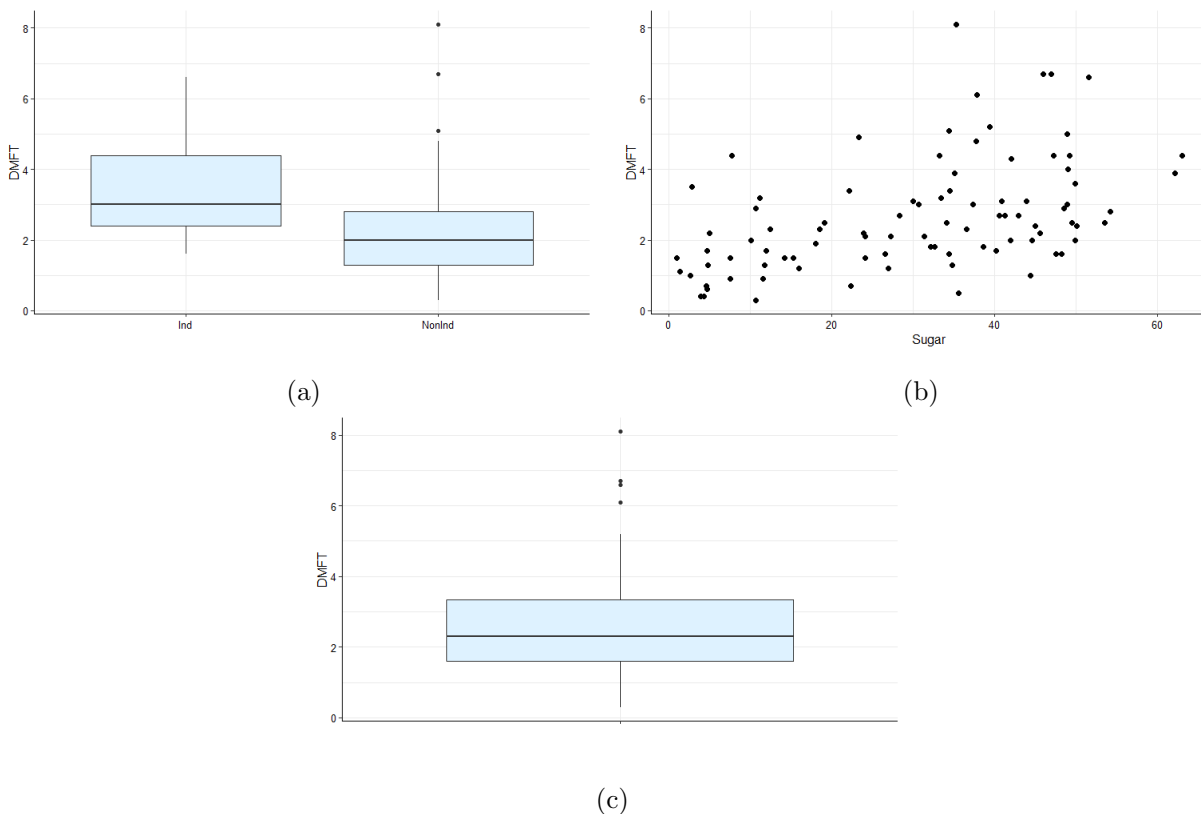


Figura 19: *Boxplot* da variável DFMT em função da variável Indus (a), gráfico de dispersão da variável DFMT em função da variável Sugar (b) e *Boxplot* da variável resposta DFMT – dados *Dental*.

A Tabela 9 apresenta os valores das estimativas dos coeficientes de regressão, seus erros padrão, as estatísticas do teste de Wald e o respectivos p -valores. Primeira-

mente, ressalta-se que a especificação dos modelos de regressão considerados contempla estrutura de regressão para os submodelos dos parâmetros μ e σ , enquanto o submodelo associado ao parâmetro ν possui somente o intercepto. Em relação aos coeficientes de regressão dos modelos, considerando um nível de significância de 5%, nota-se que os resultados e interpretações de ambos os modelos são similares. Para o modelo BS clássico, todos os coeficientes de regressão são significativamente diferentes de zero, com exceção de β_3 . No entanto, cabe destacar que o p -valor deste último foi igual a 0,06, o que sugere apenas uma fraca evidência contra a não significância deste coeficiente, uma vez que está próximo do nível de significância adotado. Por sua vez, considerando o modelo BSG, tem-se que os parâmetros β_2 e γ_1 são significativamente diferentes de zero. Entretanto, para os parâmetros β_3 , γ_3 e λ_1 há evidência contra a sua significância, embora os resultados indiquem uma evidência marginal de significância. Destaca-se a necessidade de interpretação cautelosa dos p -valores, especialmente em situações de menor tamanho amostral, em virtude dos resultados obtidos no estudo de simulação, nos quais se observou certa discrepância entre os níveis empíricos e o nível nominal do teste.

Tabela 9: Estimativas, erros-padrão, estatísticas- z e valores- p para os coeficientes de regressão estimados – dados *Dental*.

	Birnbbaum-Saunders				Birnbbaum-Saunders Generalizada			
	Estimativa	SE	estat- z	valor- p	Estimativa	SE	estat- z	valor- p
β_1	0,475	0,202	2,350	0,021	0,498	0,195	2,557	0,012
β_2	0,016	0,004	3,565	< 0,001	0,015	0,004	3,702	< 0,001
β_3	-0,261	0,137	-1,905	0,060	-0,220	0,131	-1,678	0,096
γ_1	-0,859	0,162	-5,282	< 0,001	-0,724	0,163	-4,434	< 0,001
γ_3	0,401	0,161	2,212	0,029	0,177	1,768	1,768	0,080
λ_1	-	-	-	-	-0,485	0,328	-1,478	0,143

Na Figura 20 são exibidos os gráficos que descrevem o comportamento dos resíduos quantílicos sob o modelo de regressão BS. Nas Figuras 20(a) e 20(b), o comportamento dos resíduos sugere variância constante e ausência de correlação com os valores ajustados de μ , uma vez que, não se observa a presença de padrões sistemáticos e os resíduos encontram-se centrados em torno de zero. Em relação à normalidade, no entanto, observam-se alguns desvios decorrentes da assimetria à esquerda, evidenciada pela cauda mais alongada na densidade estimada dos resíduos quantílicos na Figura 20(c) e pelos desvios verificados na linha de referência do *QQ-plot* exibido na 20(d). Assim, considerando o modelo de regressão baseado na distribuição Birnbbaum-Saunders clássico, verifica-se a partir dos gráficos de diagnóstico um ajuste razoável aos dados, apesar de uma leve assimetria a esquerda não ajustada pelo modelo.

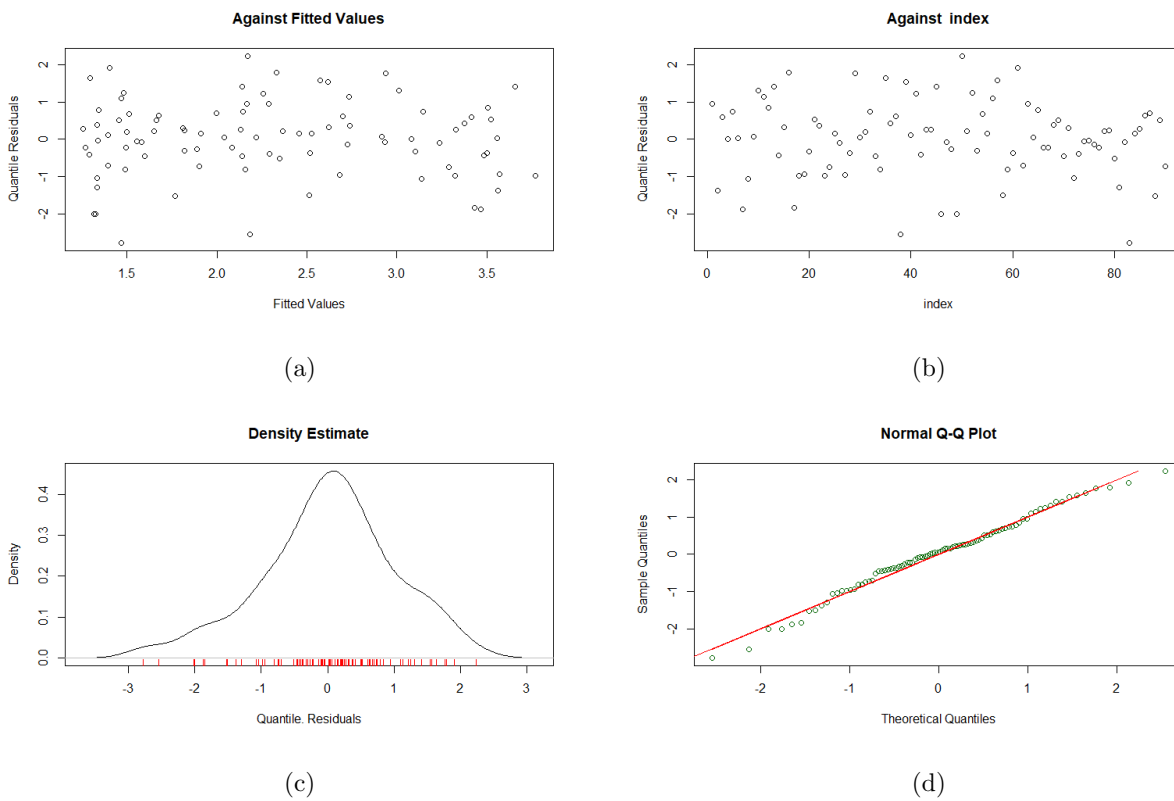


Figura 20: Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BS – dados *Dental*.

A Figura 21 apresenta o *worm plot* correspondente ao modelo Birnbaum–Saunders clássico. Em concordância com análise dos demais gráficos de diagnóstico, observam-se indícios de assimetria à esquerda não capturada pelo modelo, evidenciada pela curva em “U” invertida apresentada no gráfico. Apesar disso, a análise do *worm plot* sugere um ajuste razoável aos dados, uma vez que os resíduos quantílicos estão concentrados em torno de zero e permanecem entre as elipses.

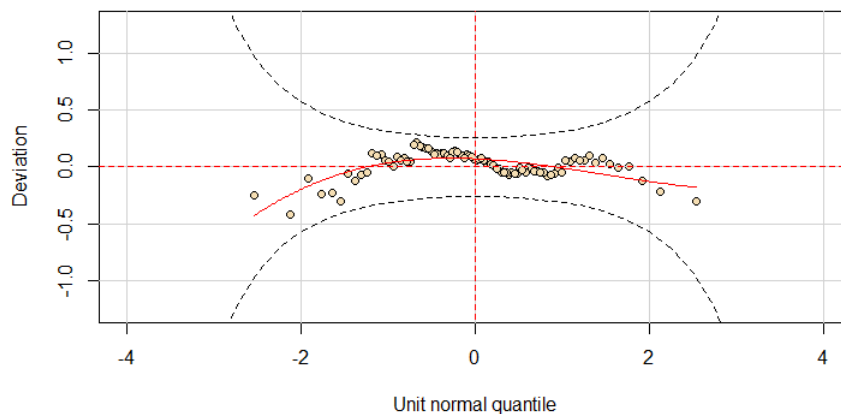


Figura 21: *Worm plot* do ajuste do modelo baseado na distribuição BS – dados *Dental*.

A Figura 22 exibe os gráficos de diagnóstico dos resíduos referentes ao modelo de regressão BSG. Nas Figuras 22(a) e 22(b), os resíduos permanecem concentrados em torno de zero, sem a presença de qualquer padrão sistemático. Contudo, diferentemente do modelo de regressão BS clássico, nas Figuras 22(c) e 22(d) observa-se um comportamento simétrico, uma vez que a densidade estimada dos resíduos quantílicos não evidencia assimetria à esquerda e os desvios no QQ -plot são menos pronunciados. Dessa forma, o modelo de regressão BSG parece capturar adequadamente a assimetria presente nos dados.

Para complementar a análise de diagnóstico dos resíduos, apresenta-se o *worm plot* correspondente ao modelo BSG na Figura 23. Em concordância com a análise dos demais gráficos de diagnóstico, o *worm plot* do modelo Birnbaum-Saunders Generalizado exibe novamente os resíduos em torno de zero. No entanto, o comportamento quadrático, ilustrado pela forma de “U” invertido é atenuado, indicando que o modelo BSG foi capaz de ajustar a assimetria não capturada pelo modelo BS.

Assim, ao comparar o ajuste dos modelos BS e BSG por meio da análise dos resíduos quantílicos, observa-se que o novo modelo de regressão proposto aparenta ser mais adequado para modelar a resposta, dado que captura a assimetria à esquerda observada nos dados.

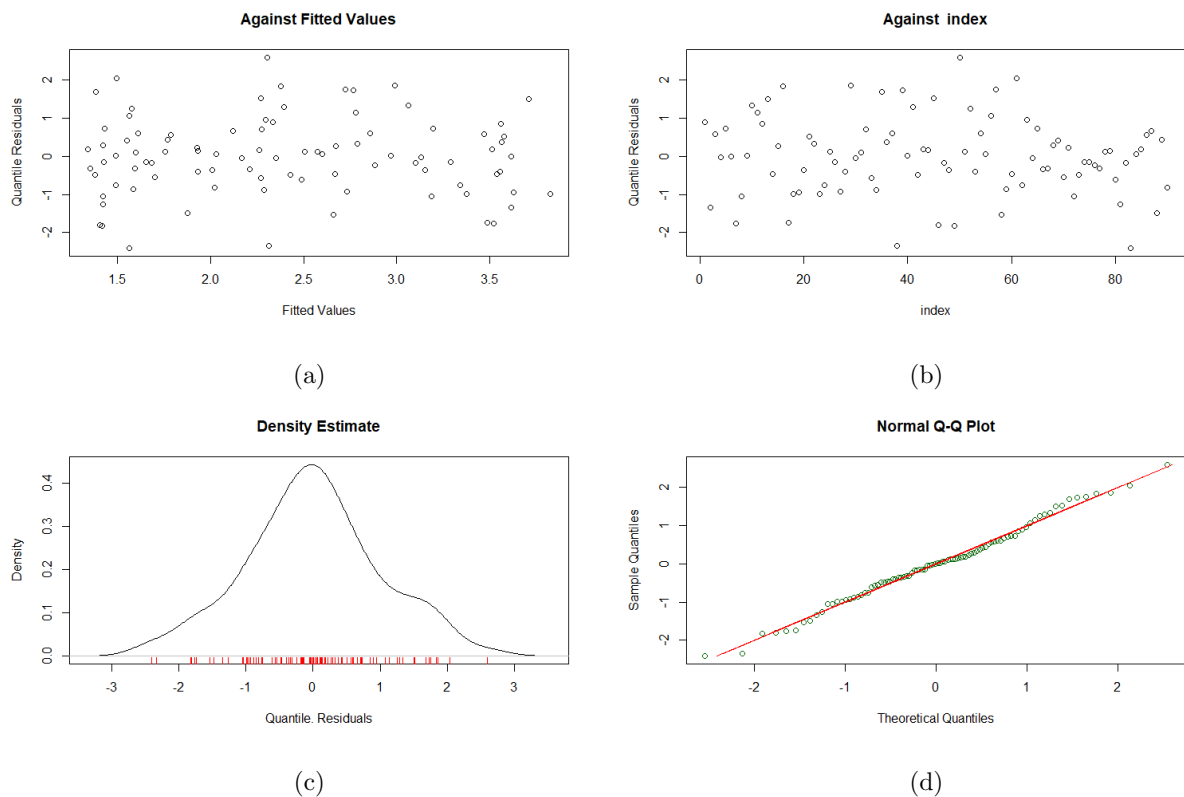


Figura 22: Gráficos de diagnóstico dos resíduos quantílicos sob o modelo baseado na distribuição BSG – dados *Dental*.

De maneira complementar, ao comparar a magnitude das estimativas dos coefici-

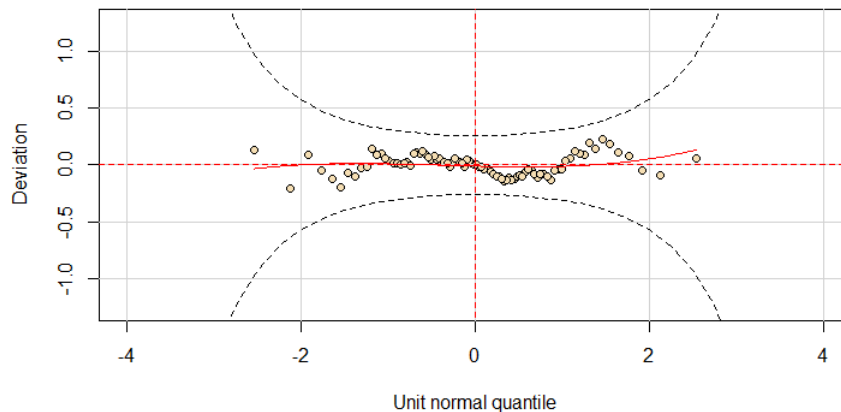


Figura 23: *Worm plot* do ajuste do modelo baseado na distribuição BSG – dados *Dental*.

entes de regressão, observa-se que os valores são relativamente próximos, sugerindo que os ajustes dos modelos apresentam comportamento semelhante. Adicionalmente, ressalta-se que o coeficiente λ_1 é significativamente igual a zero. Assim, em decorrência da função de ligação logito utilizada, esse resultado implica que o parâmetro ν é estatisticamente igual a 0,5, caracterizando o caso particular da distribuição BS original. Dessa forma, do ponto de vista estatístico, os modelos BSG e BS podem ser considerados equivalentes. Contudo, com base na análise dos *worm plots* é justificada a preferência do modelo BSG, dado que este captura a assimetria à esquerda observada.

Como o ajuste de ambos os modelos aos dados foi considerado apropriado, as suas estimativas são confiáveis e, portanto, podem ser interpretadas. A seguir, são apresentadas as interpretações das estimativas correspondentes ao modelo baseado na distribuição BSG, contudo, considerações análogas podem ser feitas a respeito do modelo BS clássico, uma vez que esses modelos são aninhados. Assim, por conta da função de ligação logaritmo utilizada no submodelo associado ao parâmetro da mediana μ , a interpretação das estimativas é dada pela inversa de seus valores, ou seja, a exponencial das estimativas dos coeficientes de regressão. A exponencial do intercepto, β_1 , corresponde ao valor da mediana quando as demais covariáveis são fixados em zero. Já as exponências dos coeficientes associados às variáveis explicativas representam fatores multiplicativos na escala da mediana da resposta, ou ainda, uma variação percentual na mediana da resposta. Para o intercepto do submodelo de μ , tem-se que $e^{\hat{\beta}_1} = e^{0,498} \approx 1,645$, no entanto, este valor não possui uma interpretação, uma vez que não é coerente fixar os demais fatores em zero. Em relação a $\hat{\beta}_2$, o valor positivo da estimativa indica que o aumento do consumo médio açúcar está associado com o aumento resposta DMFT. Logo, como $e^{\hat{\beta}_2} = e^{0,015} \approx 1,015$, tem-se que a mediana da resposta DMFT tem um aumento de cerca de 1,015% para cada aumento unitário do consumo médio de açúcar. Com relação à estimativa a estimativa $\hat{\beta}_3$,

como $e^{\hat{\beta}_3} = e^{-0,220} \approx 0,802$, observa-se um impacto negativo sobre a mediana da resposta DMFT. Isso indica que o fato de o país não ser industrializado está associado a uma redução de aproximadamente de 19,8% na mediana da resposta.

A interpretação da estimativa do coeficiente do submodelo de σ é similar àquela do submodelo do parâmetro μ , dado que ambos utilizam a mesma função de ligação. Dessa forma, tem-se que $e^{\hat{\gamma}_1} = e^{-0,724} \approx 0,484$. Em relação à estimativa $\hat{\gamma}_3$, como $e^{\hat{\gamma}_3} = e^{0,177} \approx 1,193$, sugerindo um impacto positivo sobre parâmetro σ da variável resposta DMFT. Esse resultado sugere que a não industrialização está associada a um aumento de aproximadamente 19,3% do parâmetro σ .

Por sua vez, o submodelo associado ao parâmetro ν utiliza a função de ligação logito. Dessa maneira, a interpretação do modelo associado a este submodelo é feita através da inversa da função logito, porém como este submodelo tem somente o intercepto, não possui interpretação. Assim, $\text{logito}^{-1}(\hat{\lambda}_1) = 1/1 + e^{-\hat{\lambda}_1} = 1/1 + e^{0,485} \approx 0,381$.

Finalizando, conforme os critérios de informação apresentados na Tabela 10, observa-se que, segundo o AIC, o modelo BSG seria selecionado como o mais apropriado, entretanto, de acordo com o BIC, o modelo de regressão BS seria o escolhido. Considerando a análise conjunta dos resíduos e dos critérios de informação, observa-se que o modelo Birnbaum–Saunders Generalizado aparenta ser o mais adequado dentre os dois, uma vez que esse captura a assimetria à esquerda presente nos dados. Adicionalmente, a diferença observada no critério de informação bayesiano, que apontou o modelo BS como mais adequado em relação ao BSG, não é substancial.

Tabela 10: Medidas de informação – dados *Dental*.

Modelo	AIC	BIC
1 Birnbaum-Saunders	295,26	307,76
2 Birnbaum-Saunders Generalizada	294,92	309,92

6 Conclusão

Neste trabalho, foi efetuada uma revisão de literatura abordando aspectos fundamentais referentes a distribuição Birnbaum–Saunders clássica, formulada por Birnbaum e Saunders (1969), bem como sobre a generalização proposta por Owen (2006). Essa revisão contempla desde a justificativa e a fundamentação física associada a esses modelos de probabilidade até suas respectivas definições formais. Foram apresentadas a função de distribuição acumulada, a função densidade de probabilidade e a função quantílica desses modelos de probabilidade. Adicionalmente, analisou-se o comportamento da densidade sob diferentes configurações dos parâmetros, foram apresentadas diversas propriedades, bem como os procedimentos para a geração de números pseudoaleatórios associados a cada uma dessas distribuições.

Como observado na revisão de bibliográfica, apesar da relevância da distribuição BS na literatura e de sua concepção associada a conceitos físicos, a distribuição BSG, que reforça esse conceito físico, ainda não possui um destaque equivalente. Nesse contexto, conforme a revisão de literatura realizada, não foram encontrados modelos de regressão que utilizem diretamente a distribuição BSG como distribuição condicional da variável resposta. Essa lacuna identificada motivou o desenvolvimento do presente trabalho e a formulação de um novo modelo de regressão baseado nessa distribuição.

No presente trabalho, foi desenvolvido o procedimento inferencial para o novo modelo de regressão BSG proposto, com base no método de máxima verossimilhança. Foram apresentados os cálculos da função de verossimilhança, bem como os vetores escore para cada parâmetro no contexto de regressão. A formulação do novo modelo de regressão foi elaborada baseado na estrutura e metodologia dos GAMLSS, conforme proposto por Rigby e Stasinopoulos (2005). A partir da implementação computacional inédita do novo modelo BSG na arquitetura da biblioteca `gamlss` no ambiente estatístico R, obteve-se um conjunto abrangente de ferramentas para avaliação e o ajuste do modelo, possibilitando também uma modelagem mais flexível, uma vez que os GAMLSS permitem ajustar submodelos de regressão para os três parâmetros da distribuição BSG simultaneamente.

Foi ilustrado por meio de estudos de simulação de Monte Carlo, a estabilidade das estimativas pelo método de máxima verossimilhança para o novo modelo proposto. A partir da análise gráfica dos *boxplots* das estimativas, observa-se o baixo viés dos estimadores, evidenciado pela proximidade entre os valores estimados e os verdadeiros valores dos parâmetros previamente estabelecidos. Adicionalmente, os resultados do estudo de simulação indicaram que os níveis empíricos do teste do tipo Wald apresentaram valores próximos ao nível nominal de 5% adotado.

Nas aplicações com dados reais, além de se avaliar a aplicabilidade do novo

modelo, foram apresentadas situações em que o modelo de regressão baseado na distribuição BSG proporcionou ajustes mais adequados aos dados quando comparado ao modelo BS clássico, ilustrando empiricamente contextos nos quais se evidencia o impacto do parâmetro ν . Essa melhor adequação aos dados do modelo BSG torna-se particularmente evidente na aplicação *Cheese*, na qual se observa uma discrepância significativa entre os ajustes fornecidos por ambos os modelos.

Ressalta-se que as discussões e resultados apresentados nesta dissertação possuem algumas limitações e pode ser estendido em algumas direções. Em especial, neste trabalho não foi explorado ajustes semi paramétricos do novo modelo BSG, apesar de a implementação computacional a partir do pacote `gamlss` possibilitar esse tipo ajuste. Resultados como formulação do procedimento inferencial, estudos de simulação e aplicação a dados reais podem ser desenvolvidos considerando uma estrutura semi-paramétrica sob os GAMLSS.

Considerando as limitações mencionadas, outros trabalhos de mesma natureza e considerando outros achados observados durante o desenvolvimento deste trabalho, listam-se alguns pontos que podem ser melhor explorados em trabalhos futuros:

1. Investigar a extensão semi paramétrica do modelo BSG, permitindo inclusão de diferentes tipos de função na configuração do modelo de regressão, flexibilizando ainda mais o ajuste.
2. Estudar e formular o procedimento inferencial do método de máxima verossimilhança penalizado para a extensão semi paramétrica do modelo BSG.
3. Desenvolver estudos de simulação para extensão semi paramétrica do modelo BSG, a fim de avaliar o desempenho da simulação e estabilidade das estimativas sob diferentes especificações de modelo.
4. Desenvolver modelos de regressão Birnbaum-Saunders Generalizado robustos, incluindo o processo de estimação robusto para os submodelos associados aos parâmetros da distribuição.
5. Estender o modelo de regressão BSG para modelar dados com estrutura de dependência na variável resposta.
6. Estudar outros tipos de testes de hipótese para avaliar a significância dos estimadores, avaliando os níveis empíricos do tamanho e poder dos testes sob a hipótese nula.

Por fim, destaca-se que os códigos em R e os conjuntos de dados utilizados neste trabalho estão disponíveis no repositório https://github.com/braga0m/bsg_gamlss.

Usuários interessados podem reproduzir os estudos de simulação e aplicações, bem como empregar a abordagem via metodologia GAMLSS sob a distribuição BSG em suas próprias análises.

Referências

- Balakrishnan, N e Kundu, Debasis (2019). Birnbaum-Saunders distribution: A review of models, analysis, and applications. Em: *Applied Stochastic Models in Business and Industry* 35.1, pp. 4–49.
- Beran, Jan (1994). *Statistics for Long-Memory Processes*. Routledge.
- Birnbaum e Saunders (1968). A probabilistic interpretation of Miner’s rule. Em: *SIAM Journal on Applied Mathematics* 16.3, pp. 637–652.
- (1969). A new family of life distributions. Em: *Journal of Applied Probability* 6.2, pp. 319–327.
- Buuren, Stef van e Fredriks, Miranda (2001). Worm plot: a simple diagnostic device for modelling growth reference curves. Em: *Statistics in Medicine* 20.8, pp. 1259–1277.
- Casella, George e Berger, Roger L. (2002). *Statistical Inference*. 2nd. Belmont, CA: Duxbury.
- DaSilva, Alan, Dias, Renata, Leiva, Victor, Marchant, Carolina e Saulo, Helton (2020). Birnbaum-Saunders regression models: a comparative evaluation of three approaches. Em: *Journal of Statistical Computation and Simulation*.
- Desmond, Anthony (1985). Stochastic models of failure in random environments. Em: *Canadian Journal of Statistics* 13.3, pp. 171–183.
- Devore, Jay L (2009). *Probability and Statistics for Engineering and the Sciences*. Pacific Grove: Thomson/Brooks/Cole.
- Díaz-García, José A e Leiva, V (2002). A new family of life distributions based on Birnbaum-Saunders distribution. Technical report I-02-17 (PE/CIMAT), Mexico. www.cimat.mx/biblioteca/RepTec.
- Dunn, Peter K e Smyth, Gordon K (1996). Randomized quantile residuals. Em: *Journal of Computational and Graphical Statistics* 5.3, pp. 236–244.
- (2022). *GLMsData: Generalized Linear Model Data Sets*. R package version 1.4. URL: <https://CRAN.R-project.org/package=GLMsData>.
- Espinheira, Patrícia L, Santos, Evelyne G e Cribari-Neto, Francisco (2017). On nonlinear beta regression residuals. Em: *Biometrical Journal* 59.3, pp. 445–461.
- Guiraud, Pierre, Leiva e Fierro, Raúl (2009). A non-central version of the Birnbaum-Saunders distribution for reliability analysis. Em: *IEEE Transactions on Reliability* 58.1, pp. 152–160.
- Leão, Jeremias, Leiva, Víctor, Saulo, Helton e Tomazella, Vera (2018). A survival model with Birnbaum-Saunders frailty for uncensored and censored cancer data. Em: Leiva (2015). *The Birnbaum-Saunders Distribution*. Academic Press.

- Leiva, Ferreira, Marta, Gomes, Ivette e Lillo, Camilo (2016). Extreme value Birnbaum-Saunders regression models applied to environmental data. Em: *Stochastic Environmental Research and Risk Assessment* 30, pp. 1045–1058.
- Leiva, Marchant, Carolina, Saulo, Helton, Aslam, Muhammad e Rojas, Fernando (2014a). Capability indices for Birnbaum–Saunders processes applied to electronic and food industries. Em: *Journal of Applied Statistics* 41.9, pp. 1881–1902.
- Leiva, Santos-Neto, Cysneiros, Francisco e Barros, Michelli (2014b). Birnbaum–Saunders statistical modelling: a new approach. Em: *Statistical Modelling* 14.1, pp. 21–48.
- Leiva, Silva, Michelli, Paula, Gilberto A e Sanhueza, Antonio (2008). Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. Em: *Environmetrics: The Official Journal of the International Environmetrics Society* 19.3, pp. 235–249.
- Lemonte, Artur J (2013). A new extension of the Birnbaum–Saunders distribution. Em: *Brazilian Statistical Association* 27, pp. 133–149.
- Lemonte, Artur J e Cordeiro, Gauss M (2009). Birnbaum–Saunders nonlinear regression models. Em: *Computational statistics & data analysis* 53.12, pp. 4441–4452.
- Marshall, Albert W e Olkin, Ingram (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Em: *Biometrika* 84.3, pp. 641–652.
- Moore, David S, McCabe, George P e Craig, Bruce A (2009). *Introduction to the Practice of Statistics*. Vol. 4. WH Freeman New York.
- Nelder, John A e Wedderburn, Robert WM (1972). Generalized linear models. Em: *Journal of the Royal Statistical Society Series A: Statistics in Society* 135.3, pp. 370–384.
- Owen, William J (2006). A new three-parameter extension to the Birnbaum-Saunders distribution. Em: *IEEE Transactions on Reliability* 55.3, pp. 475–479.
- Owen, William J e Padgett, William J (1999). Accelerated test models for system strength based on Birnbaum-Saunders distributions. Em: *Lifetime Data Analysis* 5, pp. 133–147.
- Ribeiro, Terezinha K. (2014). Distribuição Birnbaum–Saunders Tri-paramétrica. Trabalho de Conclusão de Curso (Graduação). Universidade Federal de Campina Grande.
- Rieck, James R e Nedelman, Jerry R (1991). A log-linear model for the Birnbaum—Saunders distribution. Em: *Technometrics* 33.1, pp. 51–60.
- Rigby e Stasinopoulos (2005). Generalized additive models for location, scale and shape. Em: *Journal of the Royal Statistical Society Series C: Applied Statistics* 54.3, pp. 507–554.
- Rigby, Stasinopoulos, Heller, Gillian e De Bastiani, Fernanda (2019). *Distributions for modeling location, scale, and shape: Using GAMLSS in R*. Chapman e Hall/CRC.
- Sánchez, Luis, Leiva, Víctor, Galea, Manuel e Saulo, Helton (2021). Birnbaum-Saunders quantile regression and its diagnostics with application to economic data. Em: *Applied Stochastic Models in Business and Industry* 37.1, pp. 53–73.

- Sánchez, Luiz Melo (2020). Birnbaum-Saunders quantile regression models. Tese de dout. Universidad de la Frontera.
- Santos-Neto, Manoel, Cysneiros, Francisco, Leiva, Víctor e Barros, Michelli (2016). Reparameterized Birnbaum-Saunders regression models with varying precision. Em: *Electronic Journal of Statistics* 10.
- Santos-Neto, Manoel, Cysneiros, Francisco José A, Leiva, Victor e Ahmed, S Ejaz (2012). On new parameterizations of the Birnbaum-Saunders distribution. Em: *Pakistan Journal of Statistics* 28.1, pp. 1–26.
- Silva, Michelli Karinne Barros da (2007). Modelos de regressão Birnbaum-Saunders generalizados. Tese de dout. Universidade de São Paulo.
- Silva, Paula, Gilberto e Leiva (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. Em: *Lifetime data analysis* 14.3, pp. 316–332.
- Stasinopoulos, Kneib, Klein, Mayr e Heller (2024). *Generalized additive models for location, scale and shape: a distributional regression approach, with applications*. Vol. 56. Cambridge University Press.
- Stasinopoulos, Mirkis, Rigby, Robert, Heller, Gillian, Voudouris, Vlasios e De Bastiani, Fernanda (2017). *Flexible Regression and Smoothing: Using GAMLSS in R*. New York: CRC Press, Taylor & Francis Group.
- Valentim da Silva, Iago Renan (2022). A quantile regression model based on an Owen distribution. Tese de dout. Universidade Federal de Campina Grande.
- Wald, Abraham (1943). Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. Em: *Transactions of the American Mathematical Society* 54.3, pp. 426–482.
- Woodward, M e Walker, AR (1994). Sugar consumption and dental caries: evidence from 90 countries. Em: *British dental journal* 176.8, pp. 297–302.