



Jessica Hemily De Sá Gonçalves

**Regressão quantílica paramétrica baseada nas
distribuições Logística Birnbaum-Saunders e
Secante Hiperbólica Birnbaum-Saunders**

Maringá,
30 de abril de 2026

Jessica Hemily De Sá Gonçalves

**Regressão quantílica paramétrica baseada nas
distribuições Logística Birnbaum-Saunders e Secante
Hiperbólica Birnbaum-Saunders**

Dissertação apresentada à Universidade Estadual de Maringá, como parte das exigências do Programa de Pós-Graduação em Bioestatística para a obtenção do título de Mestre.

Universidade Estadual de Maringá - UEM

Orientador: Josmar Mazucheli

Maringá,
30 de abril de 2026

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá - PR, Brasil)

G635r

Gonçalves, Jessica Hemily de Sá

Regressão quantílica paramétrica baseada nas distribuições Logística Birnbaum-Saunders e Secante Hiperbólica Birnbaum-Saunders / Jessica Hemily de Sá Gonçalves. -- Maringá, PR, 2026.

76 f. : il. color., figs., tabs.

Orientador: Prof. Dr. Josmar Mazucheli.

Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Departamento de Estatística, Programa de Pós-Graduação em Bioestatística, 2026.

1. Regressão quantílica paramétrica. 2. Distribuição Birnbaum-Saunders. 3. Distribuição logística Birnbaum-Saunders. 4. Distribuição secante hiperbólica Birnbaum-Saunders. I. Mazucheli, Josmar, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Estatística. Programa de Pós-Graduação em Bioestatística. III. Título.

CDD 23.ed. 519.5

Elaine Cristina Soares Lira - CRB-9/1202

JESSICA HEMILY DE SÁ GONÇALVES

Regressão quantílica paramétrica baseada nas distribuições Logística

Birnbaum-Saunders e Secante Hiperbólica Birnbaum-Saunders

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

BANCA EXAMINADORA

Documento assinado digitalmente
 **JOSMAR MAZUCHELI**
Data: 05/03/2026 10:02:17-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Josmar Mazucheli
Universidade Estadual de Maringá – PBE/UEM

Documento assinado digitalmente
 **WILLIAN LUIS DE OLIVEIRA**
Data: 05/03/2026 13:54:02-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Willian Luís de Oliveira
Universidade Estadual de Maringá – PBE/UEM

Documento assinado digitalmente
 **EMILIO AUGUSTO COELHO BARROS**
Data: 06/03/2026 14:15:12-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Emilio Augusto Coelho Barros
Universidade Tecnológica Federal do Paraná – UTFPR

Maringá, 27 de fevereiro de 2026.

AGRADECIMENTOS

Agradeço primeiramente aos meus familiares, que foram minha base e refúgio. Em especial, à minha mãe, Tania Regina, e ao meu padrasto, Silvio Dias, por todo o carinho e incentivo que me permitiram chegar até aqui. Sou imensamente grata por todo o suporte incondicional, pois, mesmo que muitas vezes eles não compreendessem exatamente a natureza dos meus objetivos acadêmicos ou a complexidade deste trabalho, o apoio deles nunca faltou, e isso foi fundamental na minha caminhada.

Ao meu irmão, agradeço por ser uma grande motivação. O desejo de ser uma pessoa melhor a cada dia e de servir como um exemplo positivo para ele foi o combustível que me manteve engajada e focada nos momentos mais desafiadores desta trajetória.

Aos meus amigos do mestrado, minha gratidão por estarem ao meu lado. A presença de vocês tornou esta caminhada acadêmica mais leve e esperançosa.

Ao meu orientador, Prof. Dr. Josmar Mazucheli, agradeço imensamente pela orientação e por todo o apoio fundamental na execução deste trabalho.

Um agradecimento muito especial direciono ao Professor Dr. Carlos Aparecido dos Santos. Foi o seu incentivo que me motivou a realizar a inscrição neste mestrado. Sem as suas palavras e o seu apoio inicial, eu não teria dado esse passo e esta conquista não seria possível.

RESUMO

A distribuição Birnbaum-Saunders (BS) tem sido amplamente estudada e aplicada em diversas áreas, como engenharia, meio ambiente e medicina, principalmente para modelar dados contínuos com suporte positivo, como tempos de vida e dados associados a danos cumulativos. Destaca-se que a distribuição BS original, com núcleo normal, a qual será denominada aqui de Normal-BS, é parametrizada em termos de sua mediana, o que a torna particularmente adequada para dados assimétricos, onde a mediana é uma medida de centralidade mais representativa que a média. Embora haja na literatura estudos de regressão quantílica para a distribuição Normal-BS, Sánchez et al. (2020), Unit Birnbaum-Saunders (UBS), Mazucheli et al. (2021), Birnbaum-Saunders de valor extremo, Bittencourt (2022), e BS com um componente de mistura no parâmetro de escala (SM-BSQR), Dasilva et al. (2025), não há, até o momento, trabalhos que abordem a regressão quantílica utilizando a parametrização das distribuições Logística Birnbaum-Saunders (LBS) e Secante Hiperbólica Birnbaum-Saunders (HBS), o que representa uma lacuna que esta pesquisa busca preencher. Este estudo propõe novos modelos de regressão quantílica para dados contínuos com suporte positivo, baseados nas distribuições Logística Birnbaum-Saunders (LBS), considerando a parametrização proposta por Sanhueza, Leiva e Balakrishnan (2008), e Secante Hiperbólica Birnbaum-Saunders (HBS), construída a partir do núcleo da normal secante hiperbólica apresentada por Fischer, Hui e Hösle (2017). Nessas abordagens, o parâmetro de locação permanece associado à mediana, permitindo sua reparametrização em termos de um quantil de ordem τ , com $0 < \tau < 1$. A formulação das distribuições LBS e HBS em termos de suas funções quantílicas é essencial para a construção dos modelos propostos, pois possibilita avaliar o efeito das covariáveis ao longo de toda a distribuição da variável resposta, e não apenas sobre sua média. Serão utilizados métodos de estimação por máxima verossimilhança e técnicas de diagnóstico de influência para avaliar a adequação dos modelos e identificar observações potencialmente influentes. A aplicação a dados reais ilustrará o potencial das propostas, e uma comparação com o modelo de regressão quantílica Normal-BS será realizada para mostrar a adequação das distribuições LBS e HBS. Este trabalho visa contribuir para a literatura de regressão quantílica, oferecendo ferramentas robustas e flexíveis para a análise de dados com características similares.

Palavras-chaves: Regressão Quantílica Paramétrica; Distribuição Birnbaum-Saunders; Distribuição logística Birnbaum-Saunders; Distribuição Secante Hiperbólica Birnbaum-Saunders.

ABSTRACT

The Birnbaum-Saunders (BS) distribution has been widely studied and applied in several fields, such as engineering, environmental sciences, and medicine, particularly for modeling continuous data with positive support, such as lifetime data and cumulative damage processes. It is noteworthy that the original BS distribution, based on a normal kernel, here referred to as Normal-BS, is parameterized in terms of its median, making it especially suitable for asymmetric data, where the median is a more representative measure of central tendency than the mean. Although the literature includes studies on quantile regression for the Normal-BS distribution Sánchez et al. (2020), Unit Birnbaum–Saunders (UBS) Mazucheli et al. (2021), Extreme Value Birnbaum-Saunders Bittencourt (2022), and BS with a mixture component in the scale parameter (SM-BSQR) Dasilva et al. (2025), there are currently no studies addressing quantile regression based on the Logistic Birnbaum-Saunders (LBS) and Hyperbolic Secant Birnbaum-Saunders (HBS) parameterizations, which represents a gap that this research aims to fill. This study proposes new parametric quantile regression models for continuous data with positive support, based on the Logistic Birnbaum–Saunders (LBS) distribution, considering the parameterization proposed by Sanhueza, Leiva e Balakrishnan (2008), and the Hyperbolic Secant Birnbaum-Saunders (HBS) distribution, constructed from the hyperbolic secant normal kernel presented by Fischer, Hui e Hösle (2017). In these approaches, the location parameter remains associated with the median, allowing its reparameterization in terms of a quantile of order τ , with $0 < \tau < 1$. The formulation of the LBS and HBS distributions in terms of their quantile functions is essential for constructing the proposed models, as it enables the assessment of covariate effects across the entire distribution of the response variable, rather than only on its mean. Maximum likelihood estimation methods and influence diagnostics will be employed to assess model adequacy and identify potentially influential observations. An application to real data will illustrate the potential of the proposed approaches, and a comparison with the Normal-BS quantile regression model will be conducted to demonstrate the suitability of the LBS and HBS distributions. This work aims to contribute to the quantile regression literature by providing robust and flexible tools for analyzing data with similar characteristics.

Key-words: Parametric Quantile Regression; Birnbaum-Saunders Distribution; Logistic Birnbaum-Saunders Distribution; Hyperbolic Secant Birnbaum-Saunders Distribution.

LISTA DE ILUSTRAÇÕES

Figura 1 – Densidades da distribuição LBS reparametrizada em função do quantil para diferentes valores de α , μ e τ	26
Figura 2 – Densidades da distribuição HBS reparametrizada em função do quantil para diferentes valores de α , μ e τ	32
Figura 3 – Painel de dispersão e correlações entre as variáveis do conjunto de dados. .	45
Figura 4 – Gráficos de dispersão entre o preço unitário dos imóveis por metro quadrado (Y) e as variáveis explicativas selecionadas: idade da casa (X_2), distância até a estação de metrô mais próxima (X_3) e número de lojas de conveniência próximas (X_4).	47
Figura 5 – Worm plots para os modelos HBS, LBS e NBS ajustados em diferentes níveis de quantil ($\tau = 0.1, 0.25, 0.5, 0.75, 0.9$).	51
Figura 6 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo HBS	52
Figura 7 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo LBS	53
Figura 8 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo NBS	53
Figura 9 – Posição relativa às mensurações corporais tomadas em ovinos. (KORITI-AKI, 2015)	54
Figura 10 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo NBS	60
Figura 11 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo LBS	61
Figura 12 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo HBS	61

Figura 13 – Worm plot do modelo ajustado NBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.	62
Figura 14 – Worm plot do modelo ajustado LBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.	62
Figura 15 – Worm plot do modelo ajustado HBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.	63

LISTA DE TABELAS

Tabela 1	– Resultados de Viés e REQM para o modelo HBS considerando diferentes valores de τ , $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores de σ	40
Tabela 2	– Resultados de Viés e REQM para o modelo HBS considerando diferentes valores de τ , tamanhos amostrais, porcentagens de censura e dois valores de σ	41
Tabela 3	– Resultados de Viés e REQM para o modelo LBS considerando diferentes valores de τ , $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores de σ	42
Tabela 4	– Resultados de Viés e REQM para o modelo LBS considerando diferentes valores de τ , tamanhos amostrais, porcentagens de censura e dois valores de σ	43
Tabela 5	– Estatísticas descritivas das variáveis analisadas - dados de avaliação imobiliária.	46
Tabela 6	– Estimativas de máxima verossimilhança e erros padrão para diferentes modelos e quantis τ	48
Tabela 7	– Valores dos critérios de informação AIC, BIC e Deviance para os modelos LBS, NBS e HBS, estimados em diferentes quantis τ	50
Tabela 8	– Estatísticas descritivas dos dados de crescimento de cordeiros segundo o total de observações e indicador de censura δ	55
Tabela 9	– Estimativas de máxima verossimilhança e erros padrão para diferentes modelos, parâmetros e quantis τ	57
Tabela 10	– Valores dos critérios de informação AIC, BIC e Deviance para os modelos censurados LBS, NBS e HBS, estimados em diferentes quantis.	59

LISTA DE ABREVIATURAS E SIGLAS

BS	Distribuição de Probabilidade Birnbaum-Saunders
BSG	Distribuição de Probabilidade Birnbaum-Saunders Generalizada
BSR	Distribuição de Probabilidade Birnbaum-Saunders Reparametrizada na média
BSVE	Distribuição Birnbaum-Saunders de Valor Extremo
CUSUM	Cartas de Controle Voltadas ao Monitoramento de Processos Industriais
EMVs	Estimativas de Máxima Verossimilhança
FDA	Função de Probabilidade Acumulada
FDP	Função Densidade de Probabilidade
FQ	Função Quantílica
HBS	Distribuição Secante Hiperbólica Birnbaum-Saunders
LBS	Distribuição de Probabilidade Logística Birnbaum-Saunders
NBS	Distribuição de Probabilidade Birnbaum-Saunders com Núcleo Normal
REQM	Raiz do Erro Quadrático Médio
SM-BAQR	Modelo de Regressão Quantílica Birnbaum-Saunders por Mistura de Escala
UBS	Distribuição <i>Unit</i> Birnbaum-Saunders

SUMÁRIO

1	Introdução	12
2	Justificativa	16
3	Objetivos e Metodologia	17
3.1	Objetivos	17
3.2	Metodologia	17
4	Conceitos Básicos	20
4.1	Distribuição BS	20
4.2	Distribuição BSG	21
4.3	Distribuição LBS	22
4.3.1	LBS em Função do Quantil e do Parâmetro de Forma	23
4.3.2	Verossimilhança da LBS Reparametrizada	27
4.4	Distribuição HBS	28
4.4.1	HBS em Função do Quantil e do Parâmetro de Forma	30
4.4.2	Verossimilhança da HBS Reparametrizada	33
5	Modelo de Regressão Quantílica LBS e HBS	35
5.1	Estrutura do Modelo de Regressão Quantílica LBS e HBS	35
5.1.1	Implementação computacional – Dados não censurados	37
5.1.2	Implementação computacional – Dados com censura à direita	38
6	Simulação de Monte Carlo	39
6.1	Resultados - HBS	40
6.2	Resultados - LBS	42
7	Aplicações empíricas	44
7.1	Dados não censurados - avaliação imobiliária	44
7.1.1	Estimativas de Máxima Verossimilhança e Avaliação dos Modelos	47
7.2	Dados censurados - crescimento de cordeiros Santa Inês	54
7.2.1	Estimativas de Máxima Verossimilhança e Avaliação dos Modelos	56
8	Conclusão	64
	Referências	66
	APÊNDICE A	68
A.1	Código da Implementação da HBS reparametrizada	68
A.2	Código da Implementação da LBS reparametrizada	74

INTRODUÇÃO

A análise de regressão é uma ferramenta estatística fundamental, amplamente empregada em diversas áreas do conhecimento para investigar e modelar a relação entre uma variável de resposta e uma ou mais variáveis explicativas. A escolha da distribuição que descreve a variável resposta é de suma importância, pois um modelo apropriado se ajusta melhor aos dados, levando a inferências mais precisas e conclusões mais confiáveis. Dentre as diversas distribuições utilizadas para modelar dados contínuos de suporte positivo, a distribuição BS destaca-se por sua origem em estudos de fadiga de materiais, tendo sido introduzida por Birnbaum e Saunders (1969) para modelar a vida útil de componentes. Desde então, tem sido amplamente explorada na literatura devido à sua capacidade de representar dados positivos e assimétricos de forma eficiente.

Leiva et al. (2007) abordaram modelos de regressão log-Birnbaum-Saunders com dados censurados e diagnósticos de influência. Kundu, Balakrishnan e Jamalizadeh (2013) propuseram distribuições BS multivariadas generalizadas e questões inferenciais relacionadas. Santos-Neto et al. (2014) apresentaram resultados sobre momentos e um método de geração de números aleatórios de uma versão reparametrizada na média da distribuição BS, além de proporem métodos de estimação e inferência para essa parametrização. O artigo de Leão et al. (2018) focou na modelagem de dados de sobrevivência utilizando a distribuição BS, empregando a reparametrização introduzida por Santos-Neto et al. (2014). Nele, foram propostos e analisados modelos de fragilidade baseados na BS para dados de sobrevivência, além de modelos de fração de cura que também incorporam a fragilidade BS. Um objetivo comum, também na análise de sobrevivência, é definir o quão rapidamente os eventos (como morte, falha ou recaída) ocorrem ao longo do tempo, o que foi o foco dos autores Alam e Almalki (2021), que estudaram a função da taxa de risco, da distribuição Logística Birnbaum-Saunders (LBS), destacando a flexibilidade desse modelo para representar diferentes padrões de taxa de falha.

Reconhecida por sua capacidade de modelar dados positivos e assimétricos, a distribuição BS e suas generalizações têm sido extensivamente estudadas na literatura estatística.

A primeira extensão da distribuição BS é atribuída a Volodin e Dzhungurova (2000) que tem como foco a generalização da BS através de modelos probabilísticos de evolução de trincas (fendas/rachaduras), introduzindo uma família geral de "distribuições de trinca" (crack distributions). O mecanismo de generalização ocorre ao considerar a longevidade de um material sujeito a tensões cíclicas, onde a evolução da trinca é descrita por relações de recorrência. Dependendo de como se aproxima a soma das variáveis aleatórias que regem essas recorrências (seja pelo Teorema do Limite Central ou por uma trajetória de processo de Wiener), o modelo pode convergir para a distribuição BS, Normal ou para a distribuição Inversa Gaussiana. Nesse contexto, a distribuição BS pode ser representada como a soma de duas variáveis aleatórias independentes: uma variável que segue a distribuição de Wald (Inversa Gaussiana) e outra que, com probabilidade $1/2$, é zero, e com probabilidade $1/2$, segue uma distribuição Qui-quadrado com 1 grau de liberdade (VOLODIN; DZHUNGUROVA, 2000).

Em seguida, Díaz-García e Leiva-Sánchez (2005), Sanhueza, Leiva e Balakrishnan (2008) introduziram uma versão generalizada do modelo original BS baseada em distribuições elípticas, denominada distribuição BS generalizada (BSG), construída a partir da substituição do *núcleo* normal padrão por outras distribuições simétricas na reta real, ou seja, a generalização da BS é obtida flexibilizando a suposição de que a variável $Z = \alpha^{-1}[\sqrt{T/\beta} - \sqrt{\beta/T}]$ segue uma distribuição normal. Em vez disso, essa variável Z é assumida como seguindo uma distribuição elíptica padrão $EC(0, 1; g)$, onde g é o núcleo da densidade; nos artigos são trabalhadas as distribuições Cauchy, Laplace, Logística, Pearson VII, Distribuição t, Kotz e Bessel. Nesse contexto, o termo *núcleo* refere-se à função densidade de probabilidade subjacente a uma distribuição simétrica, que serve como base para definir as diversas versões da BSG. Uma outra aplicação proeminente dos modelos BS já estudada está na teoria dos valores extremos, onde a distribuição Birnbaum-Saunders de valor extremo (EVBS) foi proposta por Ferreira, Gomes e Leiva (2012), alterando a normal usual na representação estocástica do modelo BS. Além disso, Leiva et al. (2016) propuseram um modelo de regressão baseado na distribuição EVBS, o qual foi ilustrado com dados ambientais e tem sido uma ferramenta útil na modelagem de eventos extremos.

Independentemente da parametrização adotada para a distribuição BS, os estudos aplicados à modelagem de dados têm, em sua maioria, se concentrado na regressão clássica, em que se modela a média da variável resposta. Um exemplo disso é o trabalho de Bastos e Barreto-Souza (2021), que propôs um modelo de seleção amostral baseado na distribuição BS bivariada. Esse modelo foi desenvolvido com o objetivo de mitigar o viés de seleção amostral, permitindo que a variável de resultado seja modelada diretamente em sua escala positiva, sem a necessidade de transformações, o que favorece a interpretação dos parâmetros. A abordagem de regressão clássica utilizada relaciona o logaritmo da média das variáveis de interesse às covariáveis, sendo a estimação dos parâmetros realizada por meio do método da Máxima Verossimilhança. Seguindo essa mesma perspectiva de modelagem da média, o artigo de Iqbal e Mahmood (2024) foca no desenvolvimento de cartas de controle CUSUM voltadas ao

monitoramento de processos industriais que geram dados positivos e assimétricos. Os autores propõem a utilização de um modelo de regressão baseado na distribuição Birnbaum-Saunders reparametrizada (RBS), que possibilita a modelagem direta da média dos dados, também sem a necessidade de transformações. A regressão clássica empregada relaciona a média da variável de resposta a preditores lineares por meio de uma função de ligação, sendo usada para gerar resíduos (padronizados e de desvio). Esses resíduos, por sua vez, são aplicados na construção das cartas CUSUM propostas, voltadas à detecção de pequenos deslocamentos na média do processo. Os resultados de simulações e aplicações com dados reais evidenciam que as cartas CUSUM baseadas nos resíduos de desvio superam abordagens existentes na identificação de alterações no processo.

A análise de regressão tradicional tem como foco principal a estimativa da média da variável resposta condicionada aos valores das variáveis explicativas. Essa abordagem é bastante eficaz sob as suposições de normalidade e homocedasticidade, no entanto, a média representa apenas uma das diversas características da distribuição da variável resposta. Para obter uma descrição mais abrangente e informativa, especialmente em contextos assimétricos, com heterocedasticidade ou presença de *outliers*, é desejável modelar diferentes aspectos da distribuição condicional da resposta. Nesse cenário, a regressão quantílica surge como uma alternativa poderosa, ao permitir a modelagem direta de diversos quantis da distribuição condicional (KECOJEVIC, 2011).

A regressão quantílica, proposta por Koenker e Jr (1978), surge como uma alternativa poderosa para modelar o efeito das covariáveis em diferentes quantis da distribuição da variável resposta, oferecendo uma visão mais completa do comportamento dos dados, especialmente em situações de assimetria, heterocedasticidade ou presença de *outliers*. Existem três principais abordagens para modelar os quantis condicionais: (i) a abordagem livre de distribuição (KOENKER; JR, 1978), (ii) a abordagem semiparamétrica baseada na distribuição de Laplace assimétrica (KOENKER; MACHADO, 1999), e (iii) a abordagem paramétrica, adotada neste trabalho, que assume distribuições específicas como a Logística Birnbaum-Saunders (LBS) (SANHUEZA; LEIVA; BALAKRISHNAN, 2008). Neste trabalho, será adotada a análise paramétrica por sua capacidade de produzir curvas de quantis mais suaves e coerentes, evitando cruzamentos e permitindo incorporar formas funcionais específicas de dependência dos quantis em relação às covariáveis (KECOJEVIC, 2011).

No contexto da regressão quantílica, que busca estimar a mediana condicional ou outros quantis da resposta, um modelo foi proposto por Sánchez et al. (2020) baseando-se na distribuição BS. Esse estudo aborda a formulação de modelos de regressão quantílica utilizando a distribuição BS para dados independentes e com dependência espacial. Além disso, com o interesse em derivar distribuições com suporte limitado, como o intervalo unitário (0,1), devido à medição de diversos fenômenos naturais como índices, porcentagens e proporções, a distribuição Unit Birnbaum-Saunders (UBS), uma extensão da BS para dados no intervalo

unitário, foi proposta por Mazucheli, Menezes e Dey (2018), sendo posteriormente utilizada por Mazucheli et al. (2021) em um novo modelo de regressão quantílica para dados limitados, parametrizando-a em termos de sua função quantílica para modelar o efeito de covariáveis em toda a distribuição da resposta. No ano seguinte, um modelo de regressão quantílica paramétrica Birnbaum-Saunders de valor extremo foi introduzido por Bittencourt (2022). Para a estimação dos parâmetros, foi empregado o método de máxima verossimilhança, e técnicas diagnósticas foram desenvolvidas para avaliar a adequação do modelo e identificar casos influentes. Mais recentemente, Dasilva et al. (2025) propuseram o modelo de regressão quantílica SM-BSQR (*Scale-Mixture Birnbaum-Saunders Quantile Regression*), no qual a distribuição da variável resposta é derivada assumindo que o parâmetro de escala da distribuição BS é uma variável aleatória positiva, que segue a distribuição Gama. Essa estrutura gera uma mistura de escala da BS, resultando em uma distribuição mais flexível para modelar dados com caudas pesadas e diferentes níveis de assimetria. No artigo, os autores aplicam o modelo a dois conjuntos de dados reais, demonstrando desempenho superior em relação a modelos quantílicos baseados na distribuição BS padrão e na distribuição de Laplace assimétrica.

Apesar dos avanços recentes na literatura envolvendo modelos de regressão quantílica baseados na distribuição Birnbaum-Saunders e em algumas de suas extensões, ainda permanecem lacunas relevantes. Em particular, embora a distribuição LBS já tenha sido proposta e estudada sob diferentes perspectivas, não há, até o momento, trabalhos que explorem sua utilização em modelos de regressão quantílica paramétrica. Essa limitação restringe o potencial de aplicação da LBS em situações nas quais o interesse recai sobre diferentes regiões da distribuição condicional da variável resposta, e não apenas sobre sua média. Adicionalmente, não existe na literatura uma distribuição Birnbaum-Saunders construída a partir do núcleo da secante hiperbólica. A introdução de uma distribuição desse tipo amplia o conjunto de modelos disponíveis para a análise de dados positivos e assimétricos, oferecendo maior flexibilidade na representação de assimetria e comportamento de caudas, características frequentemente observadas em dados reais. Nesse contexto, a incorporação dessas distribuições em um arcabouço de regressão quantílica paramétrica, mostra-se particularmente promissora, ao possibilitar a obtenção de curvas de quantis suaves e coerentes, inclusive na presença de censura.

Nesta pesquisa, desenvolve-se uma abordagem de regressão quantílica paramétrica baseada nas distribuições LBS, considerando a parametrização proposta por Sanhueza, Leiva e Balakrishnan (2008), e HBS, construída a partir do núcleo da secante hiperbólica apresentado por Fischer, Hui e Hösle (2017). O modelo proposto é formulado no contexto da máxima verossimilhança, permitindo a estimação direta de quantis condicionais e a incorporação natural de censura, com implementação computacional realizada por meio do pacote GAMLSS do software R. Além do desenvolvimento teórico, são conduzidos estudos de simulação e aplicações a dados reais com o objetivo de avaliar o desempenho inferencial e comparativo das distribuições consideradas.

JUSTIFICATIVA

A modelagem estatística de dados contínuos com suporte positivo e comportamento assimétrico é recorrente em diversas áreas do conhecimento, como engenharia, ciências da saúde, economia e estudos ambientais. Nesse contexto, a BS e suas extensões têm sido amplamente utilizadas, sobretudo em aplicações envolvendo tempos de vida, falhas e durações, devido à sua capacidade de representar assimetrias típicas desses dados. Grande parte dos estudos existentes, entretanto, concentra-se em modelos de regressão clássica baseados na média condicional da variável resposta. Embora essa abordagem seja adequada em diversos cenários, ela fornece apenas uma descrição parcial do comportamento da distribuição condicional, especialmente em situações caracterizadas por assimetria acentuada, heterocedasticidade ou presença de observações extremas. Nesses contextos, os efeitos das covariáveis podem variar significativamente ao longo da distribuição da resposta, o que motiva a adoção de abordagens que permitam modelar diferentes quantis condicionais. Nesta pesquisa, será proposta uma abordagem de regressão quantílica baseada nas distribuições LBS considerando a parametrização proposta por Sanhueza, Leiva e Balakrishnan (2008) e HBS construída a partir do núcleo da secante hiperbólica apresentada por Fischer, Hui e Hösle (2017). A escolha dessa modelagem se justifica pela flexibilidade da LBS e HBS em capturar diferentes níveis de assimetria, uma característica frequentemente observada em dados reais. A combinação entre a característica de acomodar assimetrias das distribuições HBS e LBS e a robustez da regressão quantílica resulta em uma estrutura analítica poderosa para situações em que os pressupostos da regressão clássica não são satisfeitos. Dessa forma, o presente trabalho apresenta uma dupla contribuição à literatura estatística: (i) a proposição de uma nova distribuição Birnbaum-Saunders baseada no núcleo da secante hiperbólica e (ii) o desenvolvimento de modelos de regressão quantílica paramétrica para as distribuições LBS e HBS. Essas contribuições atendem tanto a demandas teóricas quanto práticas, oferecendo modelos flexíveis, robustos e interpretáveis para a análise de dados contínuos positivos e assimétricos.

OBJETIVOS E METODOLOGIA

3.1 Objetivos

Objetivo Geral

Desenvolver e aplicar modelos de regressão quantílica baseados na distribuição LBS e HBS, permitindo a estimação dos efeitos de covariáveis no quantil μ (percentil τ) e no parâmetro de forma α .

Objetivos Específicos

- Propor uma nova distribuição de probabilidade para modelagem de dados contínuos com suporte positivo, denominada Secante Hiperbólica Birnbaum-Saunders (HBS), construída a partir do núcleo da secante hiperbólica;
- Formular o modelo de regressão quantílica da LBS e HBS, relacionando os parâmetros reparametrizados (μ e α) a covariáveis através de funções de ligação apropriadas;
- Implementar o método de estimação dos parâmetros do modelo via Máxima Verossimilhança (MV);
- Conduzir estudos de simulação de Monte Carlo para avaliar o desempenho dos estimadores em amostras finitas e analisar a robustez do modelo, inclusive na presença de dados censurados;
- Ilustrar a aplicabilidade do modelo proposto por meio da análise de conjuntos de dados reais, comparando seu desempenho com o modelo Normal-BS;

3.2 Metodologia

O desenvolvimento dos modelos de regressão quantílica baseados nas distribuições LBS e HBS foi conduzido por meio das etapas descritas a seguir:

Formulações Teóricas e Reparametrização

Inicialmente, foram apresentadas a Função de Densidade de Probabilidade (*FDP*) e a Função de Distribuição Acumulada (*FDA*) das distribuições LBS e HBS reparametrizadas. A reparametrização foi realizada em termos do τ -ésimo quantil (μ) e do parâmetro de forma (α).

Modelo de Regressão Quantílica LBS e HBS

Foram codificadas, em ambiente R, as funções de densidade, distribuição acumulada, quantil e geração aleatória das distribuições LBS e HBS reparametrizadas em termos do τ -ésimo quantil e do parâmetro de forma. Adicionalmente, foram implementadas as expressões analíticas da log-verossimilhança, do vetor score (U_μ, U_α) e da matriz Hessiana associada aos parâmetros do modelo, permitindo a estimação via Máxima Verossimilhança por meio de algoritmos iterativos baseados em derivadas. A matriz de informação de Fisher observada, obtida a partir da Hessiana, foi utilizada para fins de inferência estatística. Por fim, as distribuições LBS e HBS reparametrizadas foram incorporadas ao pacote *gamLss*, possibilitando o ajuste dos modelos, a realização de diagnósticos e a análise comparativa dos resultados obtidos.

Desempenho Assintótico dos Estimadores

O desempenho dos estimadores propostos foi avaliado por meio de estudos de simulação de Monte Carlo. Foram calculados o viés e o erro quadrático médio (EQM) dos estimadores sob diversas condições, incluindo:

- Diferentes tamanhos de amostra (n);
- Variados níveis de dispersão dos dados (por meio de diferentes valores do parâmetro α);
- Diferentes percentis τ modelados;
- A presença e diferentes proporções de censura à direita, simulando cenários comuns em estudos de tempo de vida.

Aplicações com Dados Reais

Os modelos de regressão quantílica LBS e HBS foram aplicados a conjuntos de dados reais que apresentam características de suporte positivo, com ou sem censura. O desempenho dos modelos propostos foi comparado com alternativas existentes na literatura, como a regressão quantílica baseada na distribuição Birnbaum–Saunders com núcleo normal (Normal-BS), utilizando critérios de informação (AIC, BIC) e análise de resíduos.

Para ilustrar a aplicabilidade dos modelos propostos, foi utilizado o *Real Estate Valuation Data Set*, disponível no repositório público da UCI Machine Learning (YEH, 2018).

O conjunto reúne 414 observações de imóveis localizados em New Taipei City, Taiwan, no período de junho de 2012 a maio de 2013, e tem sido amplamente empregado em pesquisas de modelagem estatística no mercado imobiliário. Os dados apresentam características típicas de assimetria positiva, com presença de valores extremos (*outliers*) e indícios de heterocedasticidade o que justificou a adoção de modelos quantílicos paramétricos como os propostos neste trabalho.

Adicionalmente, para ilustrar a aplicabilidade e o desempenho dos modelos propostos em cenários que envolvem dados com censura, será analisado um conjunto de dados referente ao crescimento de cordeiros da raça Santa Inês, proveniente de um estudo experimental conduzido por KORITIAKI (2015). A amostra é composta por 178 animais acompanhados em uma propriedade tecnificada, onde a variável resposta de interesse é o tempo (em dias) necessário para que o cordeiro atinja o peso-alvo de 15 kg.

Implementação computacional.

A implementação computacional das distribuições Normal-BS, LBS e HBS reparametrizadas foi realizada no ambiente R, empregando a estrutura do pacote `gamlss.family`. Esse pacote permite modificar a estrutura de uma família de distribuições já existente no contexto do modelo GAMLSS, bem como construir novas distribuições a partir da definição explícita das funções densidade, distribuição acumulada, quantílica e das respectivas derivadas necessárias para a estimação por máxima verossimilhança. Para a estimação dos modelos de regressão, foram utilizados os pacotes `gamlss`, no caso de dados independentes, e `gamlss.cens`, para dados de sobrevivência com censura.

A construção das distribuições HBS e LBS reparametrizadas, cujas implementações completas encontram-se nos Apêndices A.1 e A.2, seguiu a estratégia descrita por Stasinopoulos et al. (2017), no Capítulo 6.4 — *Amending an existing distribution and constructing a new distribution*. Esse procedimento estabelece as diretrizes para adaptação de distribuições já implementadas e para a definição de novas famílias compatíveis com a estrutura do `gamlss`, permitindo sua incorporação ao processo de estimação do modelo.

CONCEITOS BÁSICOS

Neste capítulo são apresentados os fundamentos teóricos das distribuições utilizadas na construção dos modelos propostos. Inicialmente, descreve-se a distribuição BS clássica e sua generalização (BSG), destacando suas principais propriedades e adequação para modelagem de dados positivos.

Em seguida, são abordadas as versões LBS e HBS, obtidas pela substituição do núcleo normal por distribuições alternativas e por fim, desenvolve-se a reparametrização em termos do τ -ésimo quantil e as respectivas funções de verossimilhança, com e sem censura à direita, estabelecendo a base para os modelos de regressão quantílica discutidos nos capítulos seguintes.

4.1 Distribuição BS

A distribuição Birnbaum-Saunders (BS) é uma distribuição de probabilidade unimodal e contínua, definida para valores reais positivos. Sua natureza de suporte positivo a torna particularmente adequada para a análise de dados de sobrevivência.

A primeira parametrização da distribuição Birnbaum-Saunders (BS) foi proposta por Birnbaum e Saunders (1969) baseada na física dos materiais. Esta parametrização define a distribuição em termos de seus parâmetros de forma (α) e escala (β) e $\Phi(\cdot)$ é a função de distribuição acumulada da normal padrão.

A função de distribuição acumulada (FDA) da distribuição BS é dada por

$$F(y; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{y}{\beta}} - \sqrt{\frac{\beta}{y}} \right) \right], \quad y > 0, \alpha, \beta > 0. \quad (4.1)$$

Dessa forma, se uma variável aleatória (VA) Y segue a distribuição BS com parâmetros $\alpha > 0$ e $\beta > 0$, é utilizada a notação $Y \sim \text{BS}(\alpha, \beta)$, e a respectiva função de densidade de probabilidade (FDP) é dada por:

$$f(y; \alpha, \beta) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2\alpha^2} \left[\frac{y}{\beta} + \frac{\beta}{y} - 2 \right] \right) \frac{[y + \beta]}{2\alpha\sqrt{\beta y^3}}, \quad y > 0. \quad (4.2)$$

Em que β é a mediana e α o parâmetro de forma e conforme α se aproxima de zero, a distribuição se torna simétrica ao redor de β .

4.2 Distribuição BSG

A Distribuição Birnbaum-Saunders Generalizada surge como uma extensão flexível da distribuição Birnbaum-Saunders clássica, concebida para a modelagem de dados de tempo de vida. A generalização fundamental do modelo BS é alcançada pela substituição da distribuição normal padrão subjacente por uma classe mais ampla de distribuições simétricas, conhecidas como "kernels" ou "núcleos". Essa abordagem inovadora permite à BSG acomodar uma maior diversidade de formas, incluindo diferentes graus de curtose e assimetria, e a capacidade de apresentar tanto unimodalidade quanto bimodalidade, características que a versão clássica não consegue capturar com a mesma flexibilidade. O papel do *núcleo* é, portanto, central, pois ele define a forma da distribuição BSG, tornando-a mais robusta e adaptável a uma gama mais ampla de cenários práticos. O modelo BSG foi proposto por Diaz-Garcia e Leiva (2002), Díaz-García e Leiva-Sánchez (2005).

Se uma variável aleatória (VA) T segue uma distribuição BSG com parâmetro de forma $\alpha > 0$, parâmetro de escala $\beta > 0$ e um *núcleo* g associado, usa-se a notação $T \sim \text{BSG}(\alpha, \beta; g)$.

A função de densidade de probabilidade (*FDP*) de T que segue BSG, $T \sim \text{BSG}(\alpha, \beta; g)$, é dada por:

$$f_T(t) = f(a_t)A_t, \quad \text{com } t > 0, \alpha > 0 \text{ e } \beta > 0,$$

onde:

- $a_t = a_t(\alpha, \beta) = \frac{1}{\alpha} \left\{ \left(\frac{t}{\beta} \right)^{\frac{1}{2}} - \left(\frac{t}{\beta} \right)^{-\frac{1}{2}} \right\}$
- $A_t = \frac{d}{dt} a_t = \frac{1}{2\alpha\beta} \left\{ \left(\frac{t}{\beta} \right)^{-\frac{1}{2}} + \left(\frac{t}{\beta} \right)^{-\frac{3}{2}} \right\}$
- $\kappa_t = a_t^2$

Sua função de distribuição acumulada (*FDA*) é dada por:

$$F_T(t) = F(a_t) \quad (4.3)$$

Onde $F(\cdot)$ é a *FDA* do *núcleo* $Z \sim S(g)$.

A BSG possui as seguintes propriedades teóricas:

- Se $Z \sim S(g)$, então $T = \beta(\frac{\alpha Z}{2} + [(\frac{\alpha Z}{2})^2 + 1]^{\frac{1}{2}})^2 \sim BSG(\alpha, \beta; g)$;
- $Z = \frac{1}{\alpha} \{ (\frac{T}{\beta})^{\frac{1}{2}} - (\frac{T}{\beta})^{-\frac{1}{2}} \} \sim S(g)$;
- $U = Z^2 = \frac{1}{\alpha^2} (\frac{T}{\beta} + \frac{\beta}{T} - 2)$ segue uma distribuição qui-quadrado generalizada ($G\chi^2$) com um grau de liberdade, denotada por $U \sim G\chi^2(f)$;
- A moda (ou modas) de T , denotada por t_m , é dada pela solução da equação $\omega_g(\kappa_{t_m}) = \frac{\alpha^2 \beta t_m (t_m + 3\beta)}{2(t_m - \beta)(t_m + \beta)^2}$, onde $\omega_g(u) = g'(u)/g(u)$ e $g'(\cdot)$ é a derivada de $g(\cdot)$. A distribuição BSG pode ter múltiplas modas;
- A distribuição BSG é fechada sob transformações de escala, ou seja, $cT \sim BSG(\alpha, c\beta; g)$, com $c > 0$;
- A distribuição BSG é fechada sob reciprocidade, ou seja, $T^{-1} \sim BSG(\alpha, \beta^{-1}; g)$;
- A função quantílica de T é $t(p) = F_T^{-1}(p) = \frac{\beta}{4} (\alpha z_p + [\alpha^2 z_p^2 + 4]^{\frac{1}{2}})^2$, onde z_p é o p -ésimo percentil de $Z \sim S(g)$. A mediana de T é β (quando $p = 0.5$).

4.3 Distribuição LBS

A Distribuição Birnbaum–Saunders Logística (LBS) é uma extensão da distribuição Birnbaum–Saunders clássica que substitui o núcleo normal padrão pela distribuição logística padrão. Essa substituição resulta em uma família mais flexível, com caudas mais pesadas e maior curtose do que a versão original. A distribuição logística padrão possui função de distribuição acumulada:

$$F_Z(z) = \frac{1}{1 + e^{-z}}, \quad z \in \mathbb{R}. \quad (4.4)$$

A partir desse núcleo, a distribuição LBS é construída pela transformação clássica de Birnbaum–Saunders, substituindo-se o gerador normal por $Z \sim \text{Logística}(0, 1)$. Assim, uma variável aleatória T segue a distribuição LBS com parâmetros $\alpha > 0$ (forma) e $\beta > 0$ (escala), denotada por $T \sim \text{LBS}(\alpha, \beta)$, se:

$$T = \beta \left(\frac{\alpha Z}{2} + \sqrt{1 + \frac{\alpha^2 Z^2}{4}} \right)^2, \quad Z \sim \text{Logística}(0, 1). \quad (4.5)$$

Função de Distribuição Acumulada

A função de distribuição acumulada da LBS é obtida substituindo a FDA da normal padrão pela da logística na transformação Birnbaum–Saunders. Assim, para $t > 0$:

$$F_T(t; \alpha, \beta) = \frac{1}{1 + \exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right] \right\}}. \quad (4.6)$$

Função Densidade de Probabilidade

Derivando (4.6) em relação a t , obtém-se a função densidade de probabilidade da LBS:

$$f_T(t; \alpha, \beta) = \frac{\exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right] \right\}}{\left[1 + \exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right] \right\} \right]^2} \cdot \frac{1}{2\alpha t} \left[\left(\frac{t}{\beta} \right)^{1/2} + \left(\frac{\beta}{t} \right)^{1/2} \right]. \quad (4.7)$$

Propriedades

As principais características teóricas de $T \sim \text{LBS}(\alpha, \beta)$ são:

- **Mediana:**

$$\text{Med}(T) = \beta.$$

- **Média:**

$$E[X] = \beta \left(1 + \frac{\pi^2}{6} \alpha^2 \right).$$

- **Variância:**

$$\sigma^2 = \frac{\pi^2}{3} \alpha^2 \beta^2 \left(1 + \frac{37}{60} \alpha^2 \right).$$

- **Coefficiente de Assimetria:**

$$\gamma_1 = \frac{4\pi^2}{3} \alpha \left(1 + \frac{19}{15} \alpha^2 \right) \left(1 + \frac{\pi^2}{6} \alpha^2 \right)^{-3/2}.$$

- **Coefficiente de Curtose:**

$$\gamma_2 = \frac{7\pi^4}{15} \alpha^4 + \frac{10\pi^2}{3} \alpha^2.$$

- **Função Quantílica:**

$$Q(p) = \beta \left[\frac{\alpha z_p}{2} + \sqrt{1 + \frac{\alpha^2 z_p^2}{4}} \right]^2, \quad z_p = \log \left(\frac{p}{1-p} \right). \quad (4.8)$$

4.3.1 LBS em Função do Quantil e do Parâmetro de Forma

Para desenvolver modelos de regressão quantílica para a distribuição LBS, é crucial obter sua função quantílica (FQ) e reparametrizar a distribuição. A FQ, $Q(\tau; \alpha, \beta) = F^{-1}(\tau; \alpha, \beta)$, permite encontrar o valor t_τ (o τ -ésimo quantil) para uma dada probabilidade acumulada τ .

Partindo da Função de Distribuição Acumulada (*FDA*) da LBS, $F(t_\tau; \alpha, \beta) = \tau$, e rearranjando algebricamente, chegamos à expressão da Função Quantílica da LBS:

$$t_\tau = Q(\tau; \alpha, \beta) = \beta \left(\frac{\alpha \xi_\tau + \sqrt{(\alpha \xi_\tau)^2 + 4}}{2} \right)^2.$$

Para a regressão quantílica, reparametrizamos a LBS definindo μ como o τ -ésimo quantil ($t_\tau = \mu$), onde τ é o percentil de interesse especificado pelo usuário. Assim, a relação para μ é:

$$\mu = \beta \left(\frac{\alpha \xi_\tau + \sqrt{(\alpha \xi_\tau)^2 + 4}}{2} \right)^2.$$

Esta expressão permite isolar o parâmetro de escala original β em função do quantil μ e do parâmetro de forma α :

$$\beta = \frac{\mu}{\left(\frac{\alpha \xi_\tau + \sqrt{(\alpha \xi_\tau)^2 + 4}}{2} \right)^2}.$$

A partir de agora, a distribuição LBS reparametrizada no quantil será indexada pelo τ -ésimo quantil μ e pelo parâmetro de forma α , e será denotada como $T \sim LBS(\mu, \alpha, \tau)$. A *FDP* e *FDA* da distribuição de probabilidade LBS com parametrização quantílica, $g(t; \mu, \alpha, \tau)$ e $G(t; \mu, \alpha, \tau)$, podem então ser escritas substituindo β pela sua expressão em termos de μ , α e ξ_τ .

Esta reparametrização é fundamental para a construção dos modelos de regressão, pois permite que os preditores atuem diretamente no quantil de interesse e no parâmetro de forma, facilitando a interpretação dos resultados, semelhante à abordagem adotada por Sánchez et al. (2020).

Para simplificar a notação da *FDP* e *FDA* da LBS reparametrizada, podemos definir $K_{\tau, \alpha} = \left(\frac{\alpha \xi_\tau + \sqrt{(\alpha \xi_\tau)^2 + 4}}{2} \right)^2$. Assim, $\beta = \frac{\mu}{K_{\tau, \alpha}}$.

Substituindo esta expressão para β na *FDP* (4.7) e *FDA* (4.6) originais da LBS, obtemos:

A densidade acumulada da LBS reparametrizada, $G(t; \mu, \alpha, \tau)$, é dada por:

$$G(t; \mu, \alpha, \tau) = \frac{1}{1 + \exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t} \right)^{1/2} \right] \right\}}, \quad t > 0.$$

E a densidade $g(t; \mu, \alpha, \tau)$ é:

$$g(t; \mu, \alpha, \tau) = \frac{\exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t} \right)^{1/2} \right] \right\}}{\left(1 + \exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t} \right)^{1/2} \right] \right\} \right)^2} \quad (4.9)$$

$$\times \frac{1}{2\alpha t} \left[\left(\frac{t}{\mu/K_{\tau, \alpha}} \right)^{1/2} + \left(\frac{\mu/K_{\tau, \alpha}}{t} \right)^{1/2} \right].$$

A seguir, são apresentadas as principais características e funções para $X \sim LBS(\mu, \alpha, \tau)$, com $\mu, \alpha > 0$.

- Média: $E[X] = \beta \left(1 + \frac{\pi^2}{6} \alpha^2 \right) = \frac{\mu}{\left(0.5 \cdot \alpha \cdot \xi_\tau + \sqrt{1 + 0.25 \cdot \alpha^2 \cdot \xi_\tau^2} \right)^2} \cdot \left(1 + \frac{\pi^2}{6} \alpha^2 \right).$
- Variância: $\sigma^2 = \frac{(\beta \alpha \pi)^2}{3} \left(1 + \frac{37}{60} \pi^2 \alpha^2 \right) = \frac{\pi^2}{3} \alpha^2 \left[\frac{\mu}{\left(0.5 \cdot \alpha \cdot \xi_\tau + \sqrt{1 + 0.25 \cdot \alpha^2 \cdot \xi_\tau^2} \right)^2} \right]^2 \left(1 + \frac{37}{60} \pi^2 \alpha^2 \right).$

A figura a seguir ilustra o comportamento da densidade da LBS reparametrizada para diferentes valores de α , μ e τ .

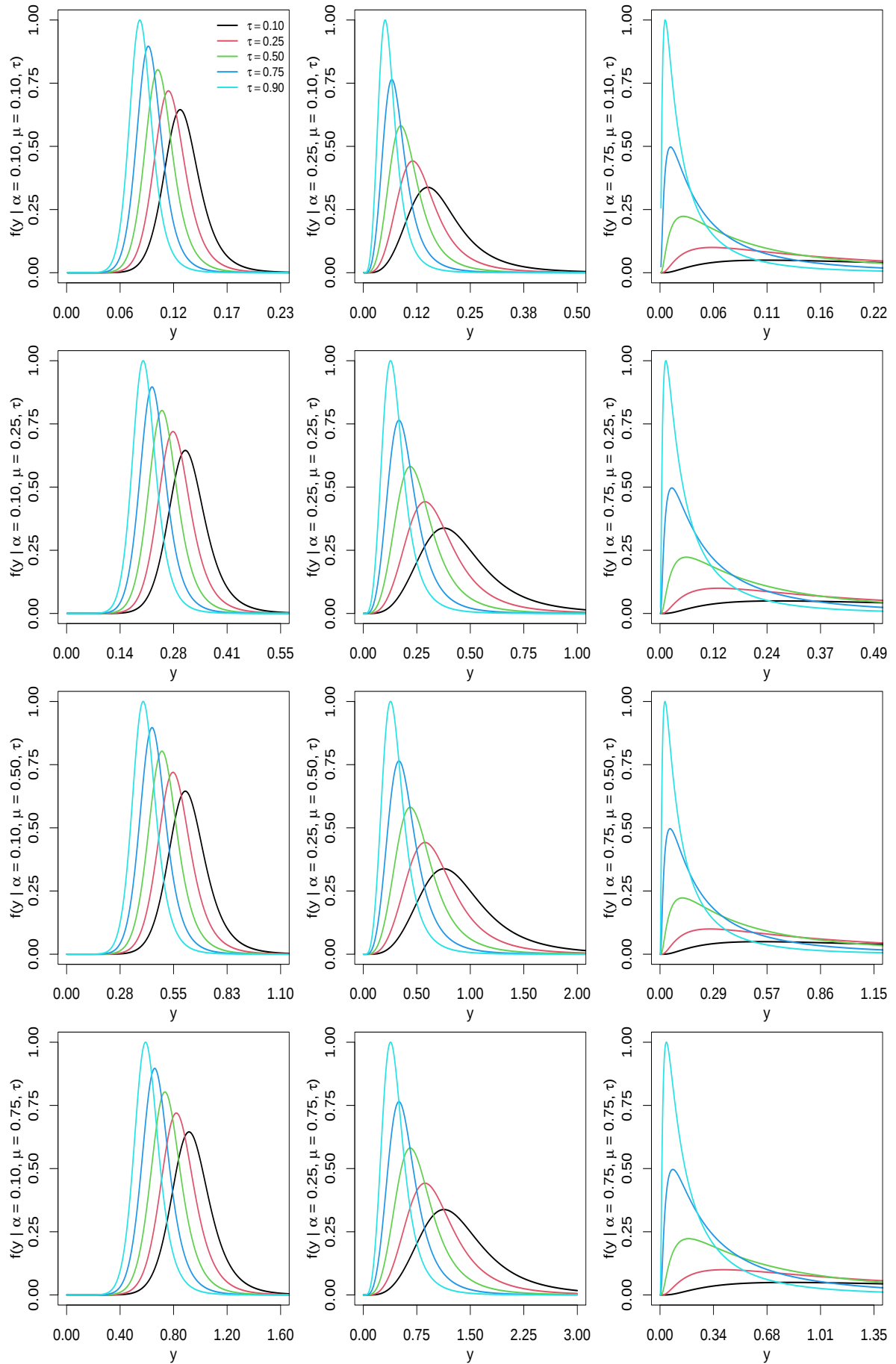


Figura 1 – Densidades da distribuição LBS reparametrizada em função do quantil para diferentes valores de α , μ e τ .

A Figura 2 ilustra o comportamento da função de densidade de probabilidade da distribuição LBS reparametrizada em função do quantil, para diversas combinações dos parâmetros de forma (α), locação (μ) e do quantil de ordem (τ). A análise desses gráficos revela a flexibilidade da LBS reparametrizada para a modelagem de dados contínuos com suporte positivo.

Observa-se que:

- Influência do parâmetro de forma (α): Ao mover-se da esquerda para a direita nas colunas dos gráficos (aumentando α), a distribuição se torna mais assimétrica à direita e mais dispersa. Para valores baixos de α (coluna mais à esquerda), as densidades são mais concentradas e tendem a ser mais simétricas;
- Influência do parâmetro de locação (μ): Analisando os gráficos de cima para baixo em cada coluna (aumentando μ), percebe-se que a distribuição é transladada para a direita no eixo y . Isso confirma que μ atua como um parâmetro de locação, influenciando a posição da densidade no domínio, sem alterar sua forma geral para um dado α .

4.3.2 Verossimilhança da LBS Reparametrizada

Seja $\mathbf{t} = (t_1, t_2, \dots, t_n)$ o vetor de observações independentes.

Sem Censura

A verossimilhança para uma amostra independente é dada por:

$$L(\mu, \beta) = \prod_{i=1}^n f(x_i; \alpha_i, \beta_i),$$

$$L(\mu, \alpha, \tau) = \prod_{i=1}^n g(t_i; \mu_i, \alpha_i, \tau).$$

A função de verossimilhança $L(\mu, \alpha, \tau; \mathbf{t})$ para uma amostra de n observações independentes e identicamente distribuídas da distribuição LBS é dada por:

$$L(\mu, \alpha, \tau; \mathbf{t}) = \prod_{i=1}^n \left[\frac{\exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right] \right\}}{\left(1 + \exp \left\{ -\frac{1}{\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right] \right\} \right)^2} \right. \\ \left. \times \frac{1}{2\alpha t_i} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} + \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right] \right] \quad (4.10)$$

Na Presença de Censura

A contribuição das observações não censuradas é o log da função densidade $f(x_i)$ e a contribuição das observações censuradas à direita é o log da função de sobrevivência $S(x_i)$, ou seja, $\log P(T > x_i)$.

Para uma amostra (x_i, δ_i) , onde:

- * x_i é o tempo observado (falha ou censura);
- * $\delta_i = 1$ se houve falha (não censurado);
- * $\delta_i = 0$ se o dado foi censurado à direita.

A verossimilhança na presença de censura pode ser definida por:

$$\begin{aligned}
 L(\mu, \beta) &= \prod_{i=1}^n [f(x_i; \alpha_i, \beta_i)]^{\delta_i} [S(x_i; \alpha_i, \beta_i)]^{1-\delta_i} \\
 L(\mu, \alpha, \tau) &= \prod_{i=1}^n g(t_i; \mu_i, \alpha_i, \tau)^{\delta_i} \cdot [1 - G(t_i; \mu_i, \alpha_i, \tau)]^{(1-\delta_i)}. \\
 L(\mu, \alpha, \tau; \mathbf{t}, \boldsymbol{\delta}) &= \prod_{i=1}^n \left\{ \frac{\exp\left\{-\frac{1}{\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right]\right\}}{\left(1 + \exp\left\{-\frac{1}{\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right]\right\}\right)^2} \right. \\
 &\quad \times \frac{1}{2\alpha t_i} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} + \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right]^{\delta_i} \\
 &\quad \times \left[1 - \frac{1}{1 + \exp\left\{-\frac{1}{\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau, \alpha}} \right)^{1/2} - \left(\frac{\mu/K_{\tau, \alpha}}{t_i} \right)^{1/2} \right]\right\}} \right]^{1-\delta_i} \Bigg\}. \tag{4.11}
 \end{aligned}$$

4.4 Distribuição HBS

A Distribuição Secante Hiperbólica Birnbaum–Saunders (HBS) é uma extensão da distribuição Birnbaum–Saunders clássica que substitui o núcleo normal padrão pela distribuição hiperbólica secante. Essa modificação tem o objetivo de conferir maior flexibilidade à modelagem de dados positivos, permitindo caudas mais pesadas e simetria mais realista em relação à distribuição normal. O núcleo HS é uma distribuição simétrica com densidade mais concentrada ao redor da média e caudas mais espessas, sendo definido por:

$$f_Z(z) = \frac{1}{\pi} \operatorname{sech}\left(\frac{\pi z}{2}\right) = \frac{1}{\pi \cosh(\pi z/2)}, \quad z \in \mathbb{R}. \tag{4.12}$$

A partir desse núcleo, a distribuição HBS pode ser construída por meio da transformação clássica de Birnbaum–Saunders, de forma análoga à BS original, porém substituindo o gerador normal por $Z \sim HS(0, 1)$. Assim, uma variável aleatória T segue uma distribuição HBS com parâmetros de forma $\alpha > 0$ e de escala $\beta > 0$, denotada por $T \sim \text{HBS}(\alpha, \beta)$, se:

$$T = \beta \left(\frac{\alpha Z}{2} + \sqrt{1 + \frac{\alpha^2 Z^2}{4}} \right)^2, \quad Z \sim HS(0, 1). \quad (4.13)$$

O uso do núcleo Secante Hiperbólico produz uma família com caudas mais pesadas que a BS e a LBS, além de uma forma mais flexível para dados com variabilidade elevada.

A função de distribuição acumulada (*FDA*) da HBS é obtida substituindo a *FDA* da normal padrão pela da HS, dada por $F_Z(z) = \frac{2}{\pi} \arctan(e^{\pi z/2})$. Assim, a *FDA* de T é:

$$F_T(t; \alpha, \beta) = \frac{2}{\pi} \arctan \left[\exp \left\{ \frac{\pi}{2\alpha} \left(\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right) \right\} \right], \quad t > 0. \quad (4.14)$$

Derivando (4.14) em relação a t , obtém-se a função densidade de probabilidade (*FDP*) da HBS:

$$f_T(t; \alpha, \beta) = \frac{\pi}{2\alpha t} \frac{\exp \left\{ \frac{\pi}{2\alpha} \left[\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right] \right\} \left[\left(\frac{t}{\beta} \right)^{1/2} + \left(\frac{\beta}{t} \right)^{1/2} \right]}{\pi^2 \left(1 + \exp \left\{ \frac{\pi}{\alpha} \left[\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right] \right\} \right)}. \quad (4.15)$$

A seguir, apresentam-se as principais propriedades teóricas de $T \sim HBS(\alpha, \beta)$:

▪ **Mediana:**

$$\text{Med}(T) = \beta.$$

▪ **Média:**

$$E[X] = \beta \left(1 + \frac{\alpha^2}{2} \right).$$

▪ **Variância:**

$$\sigma^2 = \alpha^2 \beta^2 \left(1 + \frac{5}{4} \alpha^2 \right).$$

▪ **Coefficiente de Assimetria:**

$$\gamma_1 = \frac{4\alpha \left(3 + \frac{5}{2} \alpha^2 \right)}{\left(1 + \frac{5}{4} \alpha^2 \right)^{3/2}}.$$

▪ **Coefficiente de Curtose:**

$$\gamma_2 = \frac{6\alpha^2 (4 + 5\alpha^2)}{\left(1 + \frac{5}{4} \alpha^2 \right)^2}.$$

▪ **Função quantílica:** obtida a partir da inversa da FDA do núcleo HS:

$$Q(p) = \beta \left[\frac{\alpha z_p}{2} + \sqrt{1 + \frac{\alpha^2 z_p^2}{4}} \right]^2, \quad \text{com } z_p = \frac{2}{\pi} \log \left(\tan \left(\frac{\pi p}{2} \right) \right). \quad (4.16)$$

4.4.1 HBS em Função do Quantil e do Parâmetro de Forma

De forma análoga à distribuição LBS, a construção de modelos de regressão quantílica HBS requer a obtenção explícita de sua função quantílica e a consequente reparametrização da distribuição em termos de um quantil de interesse. Seja $Q(\tau; \alpha, \beta)$ a função quantílica da HBS, definida como a inversa da função de distribuição acumulada, isto é,

$$Q(\tau; \alpha, \beta) = F^{-1}(\tau; \alpha, \beta),$$

onde $\tau \in (0, 1)$ representa o nível quantílico especificado.

Partindo da FDA da HBS apresentada em (4.14), impondo $F(t_\tau; \alpha, \beta) = \tau$ e isolando o termo associado ao núcleo Secante Hiperbólico, obtém-se:

$$z_\tau = \frac{2}{\pi} \log \left(\tan \left(\frac{\pi\tau}{2} \right) \right),$$

onde z_τ corresponde ao τ -ésimo quantil da distribuição Secante Hiperbólica padrão.

Substituindo z_τ na transformação de Birnbaum–Saunders, o τ -ésimo quantil da HBS pode ser escrito como:

$$t_\tau = Q(\tau; \alpha, \beta) = \beta \left(\frac{\alpha z_\tau + \sqrt{(\alpha z_\tau)^2 + 4}}{2} \right)^2.$$

Para fins de regressão quantílica, reparametriza-se a distribuição definindo μ como o τ -ésimo quantil da variável resposta, isto é, $t_\tau = \mu$. Dessa forma, obtém-se a relação:

$$\mu = \beta \left(\frac{\alpha z_\tau + \sqrt{(\alpha z_\tau)^2 + 4}}{2} \right)^2,$$

a qual permite expressar o parâmetro de escala β em função do quantil μ , do parâmetro de forma α e do nível quantílico τ :

$$\beta = \frac{\mu}{\left(\frac{\alpha z_\tau + \sqrt{(\alpha z_\tau)^2 + 4}}{2} \right)^2}.$$

A partir dessa reparametrização, a distribuição HBS passa a ser indexada diretamente pelo quantil de interesse μ e pelo parâmetro de forma α , sendo denotada por

$$T \sim \text{HBS}(\mu, \alpha, \tau).$$

Para simplificar a notação, define-se

$$K_{\tau, \alpha}^{(H)} = \left(\frac{\alpha z_\tau + \sqrt{(\alpha z_\tau)^2 + 4}}{2} \right)^2,$$

de modo que $\beta = \mu/K_{\tau,\alpha}^{(H)}$.

Substituindo essa expressão para β nas funções originais da HBS, obtém-se a função de distribuição acumulada da HBS reparametrizada:

$$G(t; \mu, \alpha, \tau) = \frac{2}{\pi} \arctan \left[\exp \left\{ \frac{\pi}{2\alpha} \left(\left(\frac{t}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t} \right)^{1/2} \right) \right\} \right], \quad t > 0.$$

A correspondente função densidade de probabilidade é dada por:

$$g(t; \mu, \alpha, \tau) = \frac{\pi}{2\alpha t} \frac{\exp \left\{ \frac{\pi}{2\alpha} \left[\left(\frac{t}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t} \right)^{1/2} \right] \right\} \left[\left(\frac{t}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} + \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t} \right)^{1/2} \right]}{\pi^2 \left[1 + \exp \left\{ \frac{\pi}{\alpha} \left(\left(\frac{t}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t} \right)^{1/2} \right) \right\} \right]}. \quad (4.17)$$

Essa parametrização é particularmente conveniente no contexto de regressão quantílica, pois permite que os efeitos das covariáveis sejam modelados diretamente sobre o quantil de interesse μ e sobre o parâmetro de forma α , garantindo interpretação clara e maior robustez frente à presença de caudas pesadas. Assim como no caso da LBS, a HBS reparametrizada fornece uma alternativa flexível para modelagem de dados positivos assimétricos, especialmente em aplicações onde a variabilidade e a presença de valores extremos são relevantes.

A figura a seguir ilustra o comportamento da densidade da HBS reparametrizada para diferentes valores de α , μ e τ .

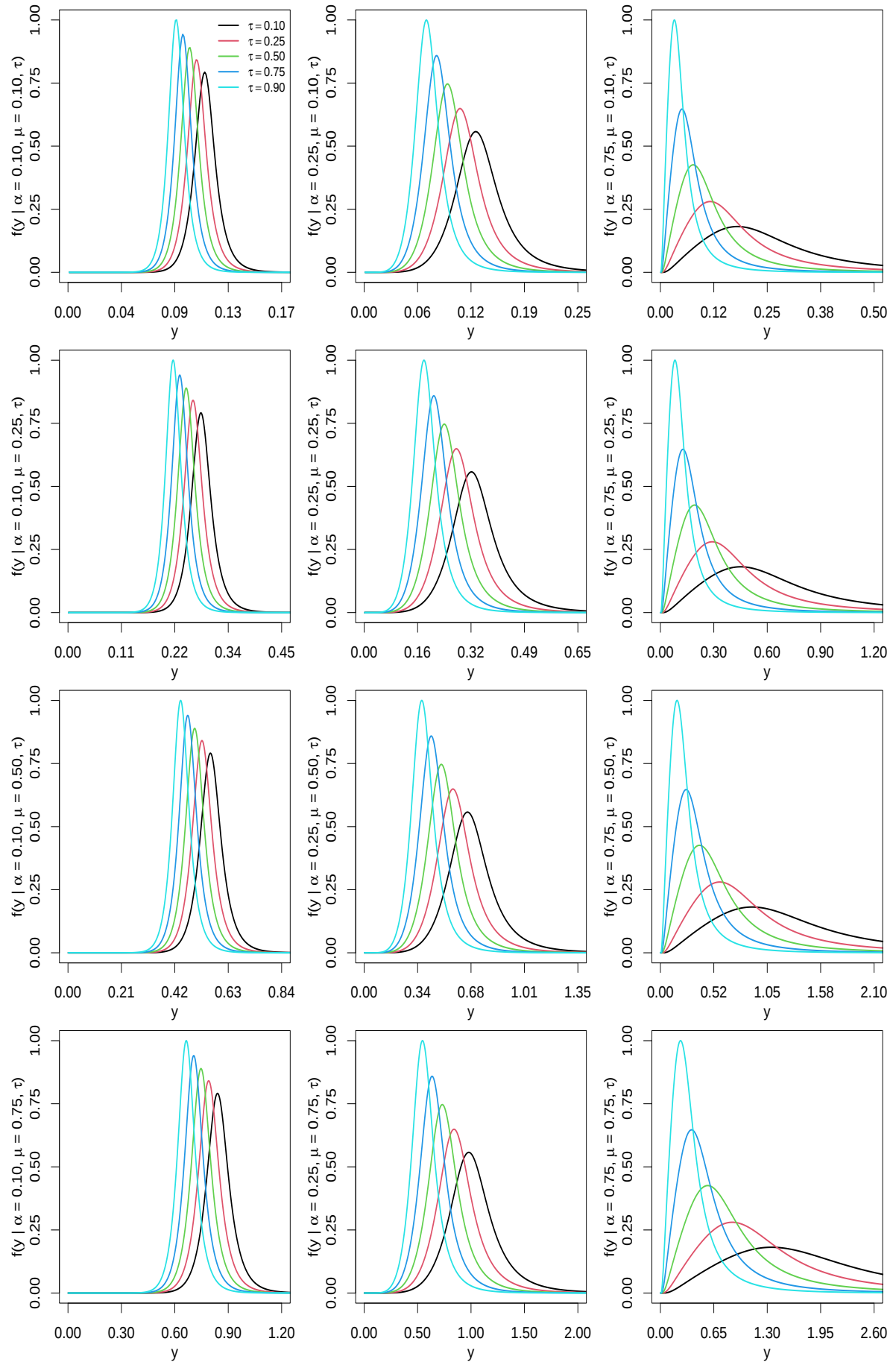


Figura 2 – Densidades da distribuição HBS reparametrizada em função do quantil para diferentes valores de α , μ e τ .

4.4.2 Verossimilhança da HBS Reparametrizada

Seja $\mathbf{t} = (t_1, t_2, \dots, t_n)$ o vetor de observações independentes.

Sem Censura

Para uma amostra independente sem censura, a função de verossimilhança associada à distribuição HBS reparametrizada é dada por:

$$L(\mu, \alpha, \tau) = \prod_{i=1}^n g(t_i; \mu_i, \alpha_i, \tau),$$

onde $g(\cdot)$ denota a função densidade de probabilidade da HBS na parametrização quantílica.

Explicitamente, a função de verossimilhança pode ser escrita como:

$$L(\mu, \alpha, \tau; \mathbf{t}) = \prod_{i=1}^n \left[\frac{\pi}{2\alpha t_i} \frac{\exp \left\{ \frac{\pi}{2\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right] \right\} \left[\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} + \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right]}{\pi^2 \left[1 + \exp \left\{ \frac{\pi}{\alpha} \left(\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right) \right\} \right]} \right]. \quad (4.18)$$

Na Presença de Censura

Considere agora uma amostra (t_i, δ_i) , $i = 1, \dots, n$, em que:

- t_i representa o tempo observado (falha ou censura);
- $\delta_i = 1$ indica observação não censurada;
- $\delta_i = 0$ indica censura à direita.

Na presença de censura à direita, a função de verossimilhança é dada por:

$$L(\mu, \alpha, \tau) = \prod_{i=1}^n [g(t_i; \mu_i, \alpha_i, \tau)]^{\delta_i} [S(t_i; \mu_i, \alpha_i, \tau)]^{1-\delta_i},$$

onde $S(t) = 1 - G(t)$ é a função de sobrevivência da HBS reparametrizada.

De forma explícita, a função de verossimilhança pode ser escrita como:

$$\begin{aligned}
L(\mu, \alpha, \tau; \mathbf{t}, \boldsymbol{\delta}) = \prod_{i=1}^n \left\{ \left[\frac{\pi}{2\alpha t_i} \times \frac{\exp \left\{ \frac{\pi}{2\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right] \right\}}{\pi^2 \left[1 + \exp \left\{ \frac{\pi}{\alpha} \left(\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right) \right\} \right]} \right] \right. \\
\quad \times \left[\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} + \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right]^{\delta_i} \\
\quad \times \left. \left[1 - \frac{2}{\pi} \arctan \left(\exp \left\{ \frac{\pi}{2\alpha} \left[\left(\frac{t_i}{\mu/K_{\tau,\alpha}^{(H)}} \right)^{1/2} - \left(\frac{\mu/K_{\tau,\alpha}^{(H)}}{t_i} \right)^{1/2} \right] \right\} \right) \right]^{1-\delta_i} \right\}.
\end{aligned}
\tag{4.19}$$

A formulação acima permite a estimação conjunta do quantil de interesse μ e do parâmetro de forma α por máxima verossimilhança, mesmo na presença de censura à direita, constituindo a base inferencial para os modelos de regressão quantílica baseados na distribuição HBS.

MODELO DE REGRESSÃO QUANTÍLICA

LBS E HBS

Nesta seção, a partir das funções densidade de probabilidade parametrizadas das distribuições Logística Birnbaum–Saunders (LBS) e Secante Hiperbólica Birnbaum–Saunders (HBS), são propostos os respectivos modelos de regressão quantílica paramétrica. Inicialmente, apresenta-se a formulação do modelo sob a reparametrização em termos do τ -ésimo quantil, definindo-se as estruturas de regressão para os parâmetros de posição e dispersão. Em seguida, descreve-se o procedimento de estimação por máxima verossimilhança e a obtenção dos erros padrão assintóticos. Por fim, apresenta-se a implementação computacional no ambiente R, incluindo a extensão para dados com censura, que fundamenta as aplicações empíricas desenvolvidas no Capítulo 7.

5.1 Estrutura do Modelo de Regressão Quantílica LBS e HBS

Sejam $X_1, \dots, X_n$ variáveis aleatórias independentes, em que cada X_i , para $i = 1, \dots, n$, segue a distribuição apresentada em 4.9 no caso LBS ou em 4.17 para HBS, ambas reparametrizadas em termos do τ -ésimo quantil. Cada variável possui parâmetro quantílico desconhecido μ_i , parâmetro de dispersão desconhecido α_i e nível quantílico fixo $\tau \in (0, 1)$. Assim, assume-se que

$$X_i \sim \text{LBS}(\mu_i, \alpha_i; \tau) \quad \text{ou} \quad X_i \sim \text{HBS}(\mu_i, \alpha_i; \tau).$$

A principal característica dessa parametrização é que o parâmetro μ_i coincide exatamente com o quantil τ da distribuição condicional de X_i , isto é,

$$Q_{X_i}(\tau \mid \mathbf{z}_i) = \mu_i,$$

o que fundamenta a abordagem de regressão quantílica paramétrica adotada.

Devido à interpretação de μ_i como parâmetro quantílico, os modelos de regressão quantílica LBS e HBS são definidos impondo que o quantil μ_i satisfaça a seguinte relação

funcional:

$$g(\mu_i) = \mathbf{z}_i^\top \boldsymbol{\beta}, \quad i = 1, \dots, n, \quad (5.1)$$

onde $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})^\top$ é um vetor p -dimensional de coeficientes de regressão desconhecidos, $\mathbf{z}_i = (1, z_{i1}, \dots, z_{i(p-1)})^\top$ representa o vetor de covariáveis conhecidas, e $g(\cdot)$ denota uma função de ligação apropriada.

Como a variável resposta e o parâmetro μ_i assumem valores positivos, é conveniente adotar funções de ligação que garantam a positividade do quantil predito. Entre as opções usuais, destacam-se:

- I. Identidade: $g(\mu_i) = \mu_i$;
- II. Logarítmica: $g(\mu_i) = \log(\mu_i)$;
- III. Raiz quadrada: $g(\mu_i) = \sqrt{\mu_i}$;
- IV. Recíproca: $g(\mu_i) = 1/\mu_i$.

Neste trabalho, adota-se a função de ligação logarítmica, resultando em

$$\log(\mu_i) = \mathbf{z}_i^\top \boldsymbol{\beta}. \quad (5.2)$$

Além do modelo para o parâmetro quantílico, admite-se que o parâmetro de dispersão α_i também possa variar em função de covariáveis, de modo a capturar possíveis estruturas de heterocedasticidade. Assim, especifica-se um segundo preditor linear da forma

$$h(\alpha_i) = \mathbf{w}_i^\top \boldsymbol{\gamma}, \quad i = 1, \dots, n, \quad (5.3)$$

onde $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{q-1})^\top$ é um vetor de coeficientes desconhecidos, $\mathbf{w}_i = (1, w_{i1}, \dots, w_{i(q-1)})^\top$ representa o vetor de covariáveis associadas à dispersão, e $h(\cdot)$ é uma função de ligação adequada. De forma análoga ao parâmetro quantílico, adota-se usualmente a ligação logarítmica, garantindo $\alpha_i > 0$.

A inferência estatística é conduzida por meio do método de máxima verossimilhança. O vetor completo de parâmetros desconhecidos é dado por

$$\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top.$$

Como não é possível obter expressões analíticas fechadas para os estimadores de máxima verossimilhança, as estimativas são obtidas numericamente por meio de algoritmos iterativos de otimização, como o método de Newton–Raphson ou métodos quasi-Newton, os quais utilizam o vetor score e a matriz Hessiana da log-verossimilhança.

Os erros padrão assintóticos dos estimadores são obtidos a partir da inversão da matriz de informação de Fisher observada, construída com base na matriz Hessiana avaliada no ponto

de máximo da log-verossimilhança. Sob condições regulares e para tamanhos amostrais suficientemente grandes, essa matriz fornece uma aproximação consistente da variância assintótica dos estimadores.

Para viabilizar a estimação via *gamlss* (dados independentes) e *gamlss.cens* (dados com censura), as distribuições LBS e HBS reparametrizadas foram implementadas no R por meio da estrutura do pacote *gamlss.family*, conforme o procedimento descrito em Stasinopoulos et al. (2017), Capítulo 6.4 (*Amending an existing distribution and constructing a new distribution*). No código de implementação, as famílias correspondentes às distribuições foram denominadas LBSTAU e HBSTAU, respectivamente. As funções que definem essas novas famílias foram incorporadas ao algoritmo de estimação do *gamlss*, e o código completo encontra-se nos Apêndices A.1 e A.2.

5.1.1 Implementação computacional – Dados não censurados

Inicialmente, foram implementadas no ambiente R as famílias LBSTAU e HBSTAU, as quais incorporam, respectivamente, as distribuições Logística Birnbaum–Saunders (LBS) e Secante Hiperbólica Birnbaum–Saunders (HBS) reparametrizadas em termos do τ -ésimo quantil. Essa implementação (Apêndice A) permitiu integrar tais distribuições ao algoritmo de estimação do pacote *gamlss*, possibilitando o ajuste do modelo por máxima verossimilhança para diferentes níveis do quantil τ .

O código apresentado a seguir corresponde exatamente ao procedimento computacional empregado na análise empírica descrita na Seção 7.1, referente à aplicação com dados não censurados.

```
library(gamlss)
tau_vec <- c(0.25, 0.35, 0.5, 0.65, 0.75)
fits_lbs <- list()

for (tau in tau_vec) {
  cat("Ajustando modelo para tau =", tau, "\n")
  fits_lbs[[as.character(tau)]] <- gamlss(
    y ~ x2 + log(x3) + x4,
    sigma.formula = ~ x2 + x4,
    data = dados,
    tau.fix = FALSE,
    family = LBSTAU()
  )
}
```

Para o modelo baseado na distribuição HBS, o procedimento foi realizado de forma análoga, mantendo-se a mesma estrutura de preditores lineares e alterando-se apenas o argumento da família para `family = HBSTAU()`.

5.1.2 Implementação computacional – Dados com censura à direita

Para os dados com censura à direita, a estrutura de sobrevivência foi incorporada ao processo de estimação por meio da função `cens()` do pacote `gamlss.cens`, possibilitando o ajuste do modelo de regressão quantílica paramétrica considerando explicitamente o mecanismo de censura.

O código apresentado a seguir foi o utilizado na análise empírica descrita na Seção 7.2, referente à aplicação com dados censurados.

```
library(survival)
library(gamlss)
library(gamlss.cens)
tau_vec <- c(0.25, 0.35, 0.5, 0.65, 0.75)
fits_lbs <- list()

for (tau in tau_vec) {
  cat("Ajustando tau =", tau, "\n")
  assign("tau", tau, envir = .GlobalEnv)

  fits_lbs[[as.character(tau)]] <- gamlss(
    Surv(tempo, delta) ~ COM + IMP + PMP + PEO,
    data = df,
    sigma.formula = Surv(tempo, delta) ~ 1,
    family = cens(LBSTAU),
    mu.start = rep(log(median(df$tempo)), nrow(df)),
    control = gamlss.control(
      c.crit = 0.001,
      trace = FALSE
    ))
}
```

A função `Surv(tempo, delta)` foi utilizada para representar o tempo até o evento e o indicador de censura, incorporando a contribuição da função de sobrevivência na construção da log-verossimilhança. O argumento `family = cens(LBSTAU)` permite ajustar a distribuição LBS na presença de censura, mantendo a parametrização quantílica do modelo.

Além disso, foi especificado um valor inicial para o preditor do parâmetro μ por meio do argumento `mu.start`, com o objetivo de favorecer a estabilidade numérica do algoritmo iterativo de máxima verossimilhança. O critério de convergência foi ajustado via `gamlss.control()`.

De forma análoga, os ajustes baseados na distribuição HBS com censura foram realizados mantendo-se a mesma estrutura de preditores lineares e alterando-se apenas o argumento da família para `family = cens(HBSTAU)`.

SIMULAÇÃO DE MONTE CARLO

Nesta seção, realizamos um estudo de simulação de Monte Carlo para investigar o desempenho dos estimadores de máxima verossimilhança (EMVs) dos parâmetros dos modelos de regressão quantílica baseados nas distribuições LBS e HBS em amostras finitas. O objetivo é avaliar empiricamente as propriedades estatísticas dos estimadores, em particular seu viés e precisão, bem como analisar o comportamento assintótico à medida que o tamanho da amostra aumenta.

Os resultados da simulação foram baseados em 5000 réplicas de Monte Carlo para cada tamanho de amostra $n \in \{50, 100, 150, 200\}$ e para diferentes níveis de quantis $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$. A variável resposta foi gerada conforme o modelo considerado, a partir das distribuições $HBS(\mu_i, \sigma, \tau)$ e $LBS(\mu_i, \sigma, \tau)$.

A estrutura de regressão considerada para o parâmetro de locação foi dada por

$$\log(\mu_i) = \beta_0 + \beta_1 x_i, \quad i = 1, \dots, n, \quad (6.1)$$

em que a covariável x_i foi gerada a partir de uma distribuição normal padrão. Os valores verdadeiros dos parâmetros foram fixados em $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores para a dispersão, $\sigma = 0.5$ e $\sigma = 1.0$.

Os EMVs foram obtidos por meio da maximização direta da função de log-verossimilhança utilizando o método quasi-Newton do tipo BFGS, implementado pela função `optim` do software R. Para avaliar o desempenho dos estimadores, foram calculados o viés empírico (Viés) e a raiz do erro quadrático médio (RMSE), definidos como:

- **Viés:** $Viés(\hat{\theta}) = E[\hat{\theta}] - \theta$;
- **Raiz do Erro Quadrático Médio (REQM):** $REQM(\hat{\theta}) = \sqrt{E[(\hat{\theta} - \theta)^2]}$.

O procedimento de simulação com censura foi o mesmo para os modelos LBS e HBS. Em cada cenário, os tempos verdadeiros de falha (T_i) foram gerados a partir da distribuição

correspondente, considerando a estrutura de regressão especificada para o parâmetro de localização (μ_i) e a censura foi introduzida de forma aleatória e independente dos tempos de falha, com níveis nominais de 20% e 40%. Para cada cenário, os tempos de censura (C_i) foram gerados a partir de uma distribuição Exponencial, cujo parâmetro foi previamente calibrado de modo que a proporção média de observações censuradas fosse aproximadamente igual a $\gamma \times 100\%$. O tempo observado foi definido por $Y_i = \min(T_i, C_i)$ e o indicador de censura por $\delta_i = \mathbb{I}(T_i \leq C_i)$.

Na estimação, as observações não censuradas contribuíram para a verossimilhança por meio da função densidade, enquanto as censuradas contribuíram por meio da função de sobrevivência da distribuição considerada. Os parâmetros foram estimados por máxima verossimilhança via otimização numérica.

6.1 Resultados - HBS

As Tabelas 1 e 2 resumem os resultados da simulação para $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores para a dispersão, $\sigma = 0.5$ e $\sigma = 1.0$, sem censura e com censura respectivamente. Observa-se que tanto o Viés quanto o REQM para os parâmetros ($\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}$) diminuem à medida que o tamanho da amostra n aumenta, o que evidencia a consistência dos estimadores.

Tabela 1 – Resultados de Viés e REQM para o modelo HBS considerando diferentes valores de τ , $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores de σ .

τ	n	$\sigma = 0.5$						$\sigma = 1.0$					
		Viés			REQM			Viés			REQM		
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$
0.10	50	0.0111	-0.0005	-0.0283	0.0772	0.0557	0.1049	0.0228	-0.0008	-0.0297	0.1433	0.1078	0.1051
	100	0.0044	-0.0003	-0.0125	0.0549	0.0397	0.0725	0.0093	-0.0006	-0.0133	0.1019	0.0770	0.0726
	150	0.0026	-0.0006	-0.0089	0.0447	0.0312	0.0593	0.0057	-0.0012	-0.0094	0.0831	0.0606	0.0594
	200	0.0020	-0.0001	-0.0067	0.0384	0.0264	0.0510	0.0042	-0.0002	-0.0071	0.0714	0.0514	0.0510
0.25	50	0.0063	0.0004	-0.0328	0.0580	0.0560	0.1073	0.0135	0.0010	-0.0344	0.1126	0.1079	0.1076
	100	0.0037	0.0007	-0.0174	0.0406	0.0391	0.0741	0.0080	0.0015	-0.0182	0.0791	0.0756	0.0742
	150	0.0029	-0.0000	-0.0117	0.0337	0.0305	0.0587	0.0061	0.0002	-0.0122	0.0654	0.0590	0.0588
	200	0.0026	-0.0001	-0.0094	0.0291	0.0266	0.0506	0.0054	0.0000	-0.0098	0.0566	0.0516	0.0507
0.50	50	0.0004	-0.0003	-0.0326	0.0504	0.0561	0.1067	0.0010	-0.0009	-0.0341	0.0974	0.1086	0.1070
	100	0.0001	0.0001	-0.0162	0.0360	0.0395	0.0747	0.0003	0.0000	-0.0169	0.0698	0.0767	0.0748
	150	0.0000	0.0002	-0.0104	0.0294	0.0311	0.0596	0.0000	0.0004	-0.0109	0.0568	0.0603	0.0597
	200	-0.0002	-0.0001	-0.0080	0.0254	0.0270	0.0519	-0.0005	-0.0003	-0.0083	0.0493	0.0524	0.0520
0.75	50	-0.0074	0.0002	-0.0319	0.0591	0.0565	0.1053	-0.0153	0.0002	-0.0334	0.1143	0.1091	0.1057
	100	-0.0027	0.0001	-0.0143	0.0414	0.0403	0.0728	-0.0057	0.0001	-0.0150	0.0801	0.0782	0.0729
	150	-0.0020	-0.0003	-0.0092	0.0335	0.0314	0.0597	-0.0040	-0.0005	-0.0097	0.0648	0.0611	0.0598
	200	-0.0013	-0.0003	-0.0067	0.0289	0.0273	0.0512	-0.0026	-0.0006	-0.0071	0.0561	0.0531	0.0512
0.90	50	-0.0133	-0.0011	-0.0304	0.0783	0.0554	0.1086	-0.0274	-0.0016	-0.0318	0.1459	0.1071	0.1089
	100	-0.0061	-0.0002	-0.0146	0.0554	0.0396	0.0748	-0.0127	-0.0001	-0.0153	0.1033	0.0766	0.0749
	150	-0.0042	0.0001	-0.0099	0.0454	0.0309	0.0609	-0.0087	0.0004	-0.0104	0.0845	0.0598	0.0609
	200	-0.0030	-0.0002	-0.0075	0.0389	0.0269	0.0524	-0.0063	-0.0001	-0.0079	0.0725	0.0520	0.0524

Tabela 2 – Resultados de Viés e REQM para o modelo HBS considerando diferentes valores de τ , tamanhos amostrais, porcentagens de censura e dois valores de σ .

τ	n	Cens. (%)	$\sigma = 0.5$						$\sigma = 1.0$					
			Viés			REQM			Viés			REQM		
			$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$
0.10	50	20	0.013	0.002	-0.031	0.081	0.070	0.117	0.024	0.002	-0.030	0.147	0.122	0.117
	100	20	0.005	0.000	-0.015	0.058	0.051	0.082	0.010	-0.001	-0.015	0.105	0.087	0.082
	150	20	0.003	0.000	-0.010	0.048	0.040	0.066	0.006	-0.001	-0.011	0.086	0.068	0.066
	200	20	0.002	0.000	-0.008	0.041	0.035	0.056	0.005	0.000	-0.008	0.073	0.060	0.056
	50	40	0.022	0.006	-0.046	0.093	0.096	0.139	0.037	0.009	-0.042	0.156	0.150	0.141
	100	40	0.011	0.002	-0.024	0.063	0.067	0.096	0.019	0.003	-0.020	0.109	0.108	0.102
	150	40	0.007	0.001	-0.015	0.052	0.050	0.075	0.013	0.001	-0.014	0.088	0.082	0.076
	200	40	0.005	0.000	-0.012	0.044	0.045	0.064	0.010	-0.001	-0.011	0.076	0.074	0.066
0.25	50	20	0.009	0.002	-0.036	0.062	0.073	0.120	0.017	0.004	-0.035	0.115	0.124	0.120
	100	20	0.004	0.000	-0.019	0.043	0.052	0.084	0.008	0.001	-0.018	0.081	0.088	0.084
	150	20	0.002	0.000	-0.011	0.036	0.040	0.066	0.005	0.001	-0.012	0.066	0.068	0.066
	200	20	0.001	-0.001	-0.009	0.031	0.036	0.058	0.003	0.000	-0.009	0.058	0.062	0.058
	50	40	0.011	0.004	-0.045	0.074	0.097	0.137	0.018	0.005	-0.042	0.121	0.151	0.140
	100	40	0.006	0.002	-0.021	0.051	0.067	0.096	0.020	0.004	0.057	0.097	0.109	0.176
	150	40	0.004	0.002	-0.013	0.042	0.052	0.075	0.006	0.001	-0.013	0.071	0.084	0.077
	200	40	0.003	0.001	-0.009	0.036	0.046	0.065	0.004	0.000	-0.009	0.060	0.074	0.067
0.50	50	20	0.000	0.001	-0.034	0.055	0.071	0.121	0.001	0.002	-0.032	0.102	0.122	0.121
	100	20	0.000	0.000	-0.016	0.039	0.052	0.084	0.000	0.001	-0.016	0.072	0.088	0.085
	150	20	0.000	0.000	-0.011	0.033	0.040	0.067	0.000	0.001	-0.011	0.059	0.069	0.068
	200	20	0.000	0.000	-0.008	0.028	0.036	0.057	0.000	0.000	-0.009	0.051	0.062	0.058
	50	40	0.004	0.003	-0.040	0.074	0.097	0.136	0.009	0.009	-0.037	0.123	0.151	0.139
	100	40	0.001	0.000	-0.021	0.049	0.069	0.095	0.003	0.002	-0.020	0.083	0.108	0.099
	150	40	0.001	0.000	-0.013	0.041	0.051	0.075	0.002	0.001	-0.013	0.069	0.081	0.077
	200	40	-0.001	-0.001	-0.011	0.035	0.045	0.064	0.000	-0.001	-0.011	0.060	0.073	0.066
0.75	50	20	-0.006	0.000	-0.034	0.066	0.071	0.118	-0.009	0.001	-0.033	0.125	0.121	0.119
	100	20	-0.003	0.000	-0.017	0.047	0.051	0.082	-0.005	0.000	-0.017	0.088	0.086	0.083
	150	20	-0.002	0.000	-0.012	0.039	0.040	0.065	-0.003	0.000	-0.012	0.072	0.068	0.066
	200	20	-0.002	-0.001	-0.009	0.033	0.036	0.056	-0.003	-0.002	-0.009	0.063	0.061	0.057
	50	40	-0.006	0.006	-0.047	0.090	0.097	0.136	-0.010	0.011	-0.043	0.158	0.150	0.139
	100	40	-0.004	0.002	-0.022	0.061	0.068	0.095	-0.006	0.003	-0.020	0.109	0.105	0.097
	150	40	-0.003	0.000	-0.014	0.050	0.051	0.075	-0.005	0.000	-0.013	0.090	0.082	0.077
	200	40	-0.003	-0.001	-0.010	0.043	0.046	0.064	-0.005	0.000	-0.010	0.077	0.073	0.066
0.90	50	20	-0.012	0.003	-0.033	0.090	0.073	0.117	-0.021	0.007	-0.032	0.164	0.123	0.118
	100	20	-0.006	0.001	-0.016	0.062	0.052	0.082	-0.010	0.002	-0.016	0.116	0.087	0.083
	150	20	-0.004	0.001	-0.011	0.051	0.040	0.066	-0.008	0.002	-0.011	0.095	0.068	0.067
	200	20	-0.003	0.000	-0.008	0.044	0.036	0.057	-0.006	0.001	-0.009	0.082	0.061	0.057
	50	40	-0.011	0.006	-0.039	0.116	0.097	0.136	-0.018	0.010	-0.036	0.208	0.152	0.140
	100	40	-0.007	0.000	-0.021	0.081	0.069	0.096	-0.011	0.001	-0.019	0.145	0.108	0.098
	150	40	-0.004	0.000	-0.013	0.065	0.052	0.075	-0.009	-0.001	-0.014	0.116	0.084	0.077
	200	40	-0.004	-0.001	-0.010	0.056	0.046	0.065	-0.009	-0.002	-0.011	0.101	0.074	0.066

6.2 Resultados - LBS

As Tabelas 3 e 4 resume os resultados da simulação para $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores para a dispersão, $\sigma = 0.5$ e $\sigma = 1.0$, sem censura e com censura respectivamente. Observa-se que tanto o Viés quanto o REQM para os parâmetros $(\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma})$ diminuem à medida que o tamanho da amostra n aumenta, o que evidencia a consistência dos estimadores.

Tabela 3 – Resultados de Viés e REQM para o modelo LBS considerando diferentes valores de τ , $\beta_0 = 0.5$, $\beta_1 = 1.0$ e dois valores de σ .

τ	n	$\sigma = 0.5$						$\sigma = 1.0$					
		Viés			REQM			Viés			REQM		
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$
0.10	50	0.0214	-0.0005	-0.0288	0.1307	0.1020	0.0978	0.0390	-0.0002	-0.0287	0.2120	0.1747	0.1009
	100	0.0088	-0.0005	-0.0130	0.0928	0.0730	0.0676	0.0162	-0.0011	-0.0140	0.1492	0.1249	0.0680
	150	0.0054	-0.0012	-0.0091	0.0758	0.0575	0.0553	0.0106	-0.0022	-0.0098	0.1223	0.0984	0.0555
	200	0.0040	-0.0003	-0.0069	0.0653	0.0488	0.0474	0.0079	-0.0012	-0.0076	0.1056	0.0834	0.0475
0.25	50	0.0130	0.0014	-0.0332	0.1061	0.1020	0.1004	0.0279	0.0039	-0.0338	0.1851	0.1732	0.1031
	100	0.0078	0.0016	-0.0176	0.0747	0.0718	0.0691	0.0162	0.0033	-0.0187	0.1300	0.1224	0.0694
	150	0.0058	0.0003	-0.0117	0.0616	0.0561	0.0547	0.0115	0.0012	-0.0124	0.1062	0.0955	0.0549
	200	0.0051	0.0001	-0.0094	0.0533	0.0490	0.0471	0.0094	0.0003	-0.0099	0.0922	0.0836	0.0473
0.50	50	0.0011	-0.0012	-0.0330	0.0923	0.1027	0.0997	0.0024	-0.0032	-0.0352	0.1571	0.1757	0.1008
	100	0.0003	-0.0001	-0.0163	0.0662	0.0728	0.0697	0.0007	-0.0011	-0.0173	0.1124	0.1253	0.0700
	150	-0.0001	0.0004	-0.0105	0.0539	0.0573	0.0556	-0.0006	0.0008	-0.0111	0.0909	0.0976	0.0559
	200	-0.0005	-0.0003	-0.0081	0.0468	0.0498	0.0484	-0.0010	-0.0007	-0.0087	0.0791	0.0849	0.0486
0.75	50	-0.0147	-0.0001	-0.0323	0.1071	0.1030	0.0986	-0.0288	-0.0011	-0.0345	0.1848	0.1748	0.0997
	100	-0.0055	-0.0001	-0.0146	0.0751	0.0743	0.0677	-0.0111	-0.0009	-0.0157	0.1286	0.1269	0.0680
	150	-0.0038	-0.0006	-0.0094	0.0608	0.0581	0.0556	-0.0072	-0.0012	-0.0102	0.1049	0.0991	0.0558
	200	-0.0025	-0.0006	-0.0069	0.0527	0.0505	0.0477	-0.0053	-0.0017	-0.0075	0.0907	0.0855	0.0478
0.90	50	-0.0258	-0.0012	-0.0306	0.1332	0.1012	0.1014	-0.0473	-0.0006	-0.0326	0.2169	0.1732	0.1024
	100	-0.0121	0.0002	-0.0149	0.0944	0.0724	0.0698	-0.0226	0.0013	-0.0158	0.1535	0.1229	0.0702
	150	-0.0081	0.0004	-0.0101	0.0772	0.0565	0.0567	-0.0148	0.0011	-0.0108	0.1251	0.0955	0.0569
	200	-0.0058	0.0000	-0.0076	0.0662	0.0493	0.0487	-0.0108	0.0003	-0.0082	0.1072	0.0831	0.0489

Tabela 4 – Resultados de Viés e REQM para o modelo LBS considerando diferentes valores de τ , tamanhos amostrais, porcentagens de censura e dois valores de σ .

τ	n	Cens. (%)	$\sigma = 0.5$						$\sigma = 1.0$					
			Viés			REQM			Viés			REQM		
			$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$
0.10	50	20	0.023	0.003	-0.030	0.133	0.116	0.109	0.041	0.008	-0.028	0.212	0.191	0.111
	100	20	0.010	-0.001	-0.015	0.095	0.083	0.077	0.018	0.001	-0.012	0.150	0.135	0.081
	150	20	0.006	-0.001	-0.010	0.077	0.065	0.062	0.012	-0.001	-0.011	0.123	0.106	0.063
	200	20	0.005	0.000	-0.008	0.066	0.057	0.053	0.009	0.001	-0.008	0.106	0.092	0.053
	50	40	0.035	0.010	-0.041	0.142	0.142	0.133	0.056	0.015	-0.035	0.221	0.213	0.143
	100	40	0.018	0.004	-0.022	0.099	0.103	0.093	0.031	0.009	-0.020	0.154	0.151	0.099
	150	40	0.012	0.001	-0.014	0.080	0.078	0.072	0.021	0.004	-0.013	0.125	0.116	0.078
	200	40	0.009	0.000	-0.011	0.069	0.070	0.062	0.017	0.003	-0.011	0.108	0.103	0.067
0.25	50	20	0.017	0.003	-0.034	0.108	0.118	0.113	0.035	0.005	-0.032	0.186	0.192	0.115
	100	20	0.008	0.001	-0.018	0.076	0.084	0.079	0.016	0.001	-0.017	0.131	0.135	0.081
	150	20	0.005	0.001	-0.011	0.062	0.065	0.062	0.010	0.002	-0.012	0.105	0.107	0.063
	200	20	0.003	-0.001	-0.009	0.054	0.059	0.054	0.007	0.000	-0.009	0.092	0.094	0.056
	50	40	0.018	0.005	-0.040	0.114	0.142	0.133	0.035	0.007	-0.033	0.187	0.213	0.146
	100	40	0.010	0.002	-0.020	0.081	0.102	0.094	0.017	0.001	-0.018	0.132	0.152	0.102
	150	40	0.006	0.001	-0.012	0.067	0.080	0.073	0.011	0.000	-0.011	0.108	0.119	0.081
	200	40	0.004	0.000	-0.009	0.056	0.071	0.063	0.008	0.000	-0.010	0.093	0.105	0.068
0.50	50	20	0.001	0.003	-0.031	0.097	0.117	0.114	0.004	0.006	-0.030	0.168	0.192	0.116
	100	20	0.001	0.002	-0.016	0.069	0.084	0.079	0.001	0.004	-0.016	0.120	0.135	0.081
	150	20	0.000	0.001	-0.011	0.057	0.066	0.063	0.001	0.002	-0.011	0.098	0.106	0.064
	200	20	0.000	0.000	-0.008	0.049	0.059	0.054	0.001	0.001	-0.009	0.084	0.094	0.055
	50	40	0.009	0.007	-0.036	0.119	0.144	0.131	0.021	0.010	-0.030	0.203	0.214	0.142
	100	40	0.003	0.001	-0.019	0.081	0.102	0.094	0.006	0.000	-0.018	0.141	0.150	0.101
	150	40	0.002	0.000	-0.012	0.067	0.077	0.073	0.003	-0.002	-0.013	0.113	0.115	0.079
	200	40	0.000	-0.001	-0.010	0.058	0.069	0.063	0.000	-0.002	-0.011	0.099	0.102	0.068
0.75	50	20	-0.010	0.001	-0.032	0.119	0.116	0.111	-0.011	0.002	-0.029	0.215	0.189	0.113
	100	20	-0.005	0.000	-0.017	0.083	0.082	0.077	-0.006	0.000	-0.015	0.150	0.132	0.078
	150	20	-0.004	0.000	-0.012	0.069	0.065	0.062	-0.005	0.001	-0.011	0.124	0.104	0.063
	200	20	-0.003	-0.001	-0.009	0.060	0.058	0.053	-0.005	-0.001	-0.009	0.108	0.093	0.054
	50	40	-0.010	0.010	-0.042	0.152	0.142	0.132	-0.007	0.014	-0.036	0.283	0.211	0.141
	100	40	-0.006	0.003	-0.020	0.105	0.099	0.091	-0.008	0.004	-0.019	0.198	0.145	0.098
	150	40	-0.005	0.000	-0.013	0.086	0.078	0.072	-0.007	0.000	-0.012	0.160	0.115	0.077
	200	40	-0.005	0.000	-0.009	0.074	0.070	0.062	-0.009	0.001	-0.010	0.138	0.102	0.066
0.90	50	20	-0.019	0.007	-0.031	0.152	0.118	0.111	-0.025	0.014	-0.029	0.260	0.192	0.113
	100	20	-0.010	0.003	-0.015	0.107	0.083	0.078	-0.013	0.007	-0.014	0.183	0.134	0.079
	150	20	-0.007	0.002	-0.011	0.089	0.065	0.063	-0.011	0.005	-0.011	0.149	0.105	0.063
	200	20	-0.006	0.002	-0.008	0.076	0.058	0.053	-0.010	0.005	-0.009	0.128	0.092	0.054
	50	40	-0.017	0.010	-0.035	0.195	0.145	0.133	-0.019	0.012	-0.031	0.345	0.216	0.144
	100	40	-0.011	0.000	-0.019	0.137	0.103	0.093	-0.013	0.001	-0.017	0.240	0.153	0.099
	150	40	-0.009	-0.001	-0.013	0.110	0.080	0.072	-0.012	-0.001	-0.013	0.190	0.119	0.077
	200	40	-0.008	-0.002	-0.011	0.095	0.071	0.063	-0.011	-0.001	-0.010	0.165	0.104	0.067

APLICAÇÕES EMPÍRICAS

Para ilustrar a aplicabilidade dos modelos propostos, serão ajustados modelos de regressão quantílica paramétrica baseados nas distribuições LBS e HBS, propostas neste trabalho, considerando diferentes níveis do quantil τ .

Para fins comparativos, também será ajustado o modelo baseado na distribuição NBS, já estabelecida na literatura. A comparação entre LBS, HBS e NBS permitirá avaliar o desempenho das novas distribuições em termos de ajuste e flexibilidade ao longo da distribuição condicional da variável resposta.

7.1 Dados não censurados - avaliação imobiliária

Para ilustrar a aplicabilidade do modelo proposto, foi utilizado o *Real Estate Valuation Data Set*, disponível no repositório público da UCI Machine Learning (YEH, 2018). O conjunto reúne 414 observações de imóveis localizados em New Taipei City, Taiwan, no período de junho de 2012 a maio de 2013, e tem sido amplamente empregado em pesquisas de modelagem estatística no mercado imobiliário.

As variáveis registradas incluem características estruturais e de localização dos imóveis, tais como: eu

- **Idade da casa (X_2):** medida em anos.
- **Distância até a estação de metrô mais próxima (X_3):** medida em metros.
- **Número de lojas de conveniência (X_4):** quantidade de estabelecimentos próximos ao imóvel.
- **Coordenadas geográficas (X_5 e X_6):** latitude e longitude do imóvel.
- **Preço unitário do imóvel (Y):** variável resposta do estudo, expressa em 10,000 dólares taiwaneses (TWD) por Ping (3.3 metro quadrado) da área útil.

Seleção de variáveis

A escolha das variáveis explicativas foi orientada por uma análise exploratória inicial, que envolveu a avaliação das correlações entre as covariáveis e a variável resposta, bem como a verificação de possível multicolinearidade entre os preditores. O painel de dispersão e a matriz de correlações apresentados na Figura 3 permitem visualizar as relações bivariadas entre as variáveis, auxiliando na identificação de padrões de associação e possíveis dependências lineares relevantes para a modelagem.

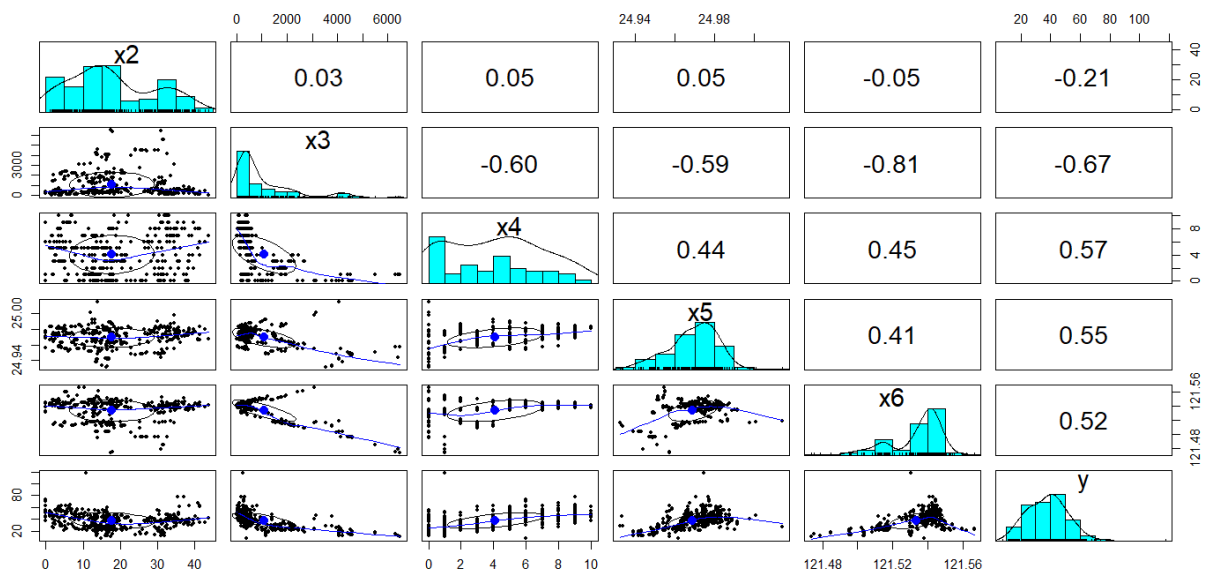


Figura 3 – Painel de dispersão e correlações entre as variáveis do conjunto de dados.

As coordenadas geográficas (latitude e longitude) não foram incluídas na modelagem final. Observou-se que a variável longitude (X_6) apresentou forte correlação linear com a distância até a estação de metrô mais próxima (X_3), com coeficiente igual a -0.81 , indicando elevada associação entre essas medidas de localização. A inclusão simultânea dessas variáveis poderia introduzir problemas de multicolinearidade, comprometendo a estabilidade das estimativas e a interpretação dos coeficientes. Assim, optou-se por excluir as coordenadas geográficas, preservando a parcimônia e a interpretabilidade do modelo.

Além da análise exploratória inicial, foi adotado um procedimento sistemático de seleção de variáveis com base no critério de informação generalizado (GAIC), utilizando a função `stepGAIC()` do pacote `gamlss`. Inicialmente, realizou-se a seleção de variáveis para o preditor associado ao parâmetro quantílico μ_i , mantendo-se fixa a estrutura do parâmetro de dispersão. Em seguida, considerando a especificação final obtida para μ_i , aplicou-se novamente o procedimento `stepGAIC()` para a seleção das covariáveis associadas ao parâmetro de dispersão σ_i .

Ressalta-se que as variáveis latitude e longitude não foram incluídas no procedimento de seleção automática via `stepGAIC()`, uma vez que já haviam sido excluídas com base na

análise exploratória prévia, devido à forte associação com a variável distância até a estação de metrô mais próxima e ao potencial problema de multicolinearidade identificado.

Esse processo sequencial permitiu identificar um modelo parcimonioso para ambos os parâmetros, equilibrando qualidade de ajuste e complexidade do modelo, conforme o critério de informação adotado.

Com base nos resultados da análise exploratória e do procedimento de seleção via GAIC, foram selecionadas três variáveis explicativas que capturam diferentes dimensões da valorização imobiliária: a idade da casa, a acessibilidade (representada pela distância log-transformada até a estação de metrô mais próxima) e a conveniência (quantidade de lojas próximas ao imóvel). Na tabela 5 são apresentados a média e desvio padrão, mediana e primeiro e terceiro quartis e valores de mínimo e máximos para as covariáveis adotadas e variável dependente.

A Figura 4 apresenta os gráficos de dispersão entre a variável resposta e as três covariáveis selecionadas. Essa visualização permite observar o comportamento geral das relações entre as variáveis, evidenciando tendências e possíveis padrões de associação. Nota-se uma relação não linear entre o preço unitário do imóvel e a idade da casa e uma tendência de decréscimo do preço com o aumento da distância até a estação de metrô mais próxima. Já o número de lojas de conveniência próximas apresenta uma relação positiva com o preço dos imóveis, refletindo o efeito esperado da maior disponibilidade de serviços na valorização imobiliária.

As estatísticas descritivas das variáveis selecionadas são apresentadas na Tabela 5, incluindo média e desvio padrão, mediana, primeiro e terceiro quartis, bem como valores mínimo e máximo da variável resposta e das covariáveis adotadas.

Tabela 5 – Estatísticas descritivas das variáveis analisadas - dados de avaliação imobiliária.

Estatística	x2 (anos)	x3 (m)	x4	y (TWD/Ping)
Média (DP)	17.7 (11.4)	1083.9 (1262.1)	4.1 (3.0)	38.0 (13.6)
Mediana (Q1 – Q3)	16.1 (9.0 – 28.2)	492.2 (289.3 – 1454.3)	4.0 (1.0 – 6.0)	38.5 (27.7 – 46.6)
Mínimo-Máximo	0.0 – 43.8	23.4 – 6488.0	0.0 – 10.0	7.6 – 117.5

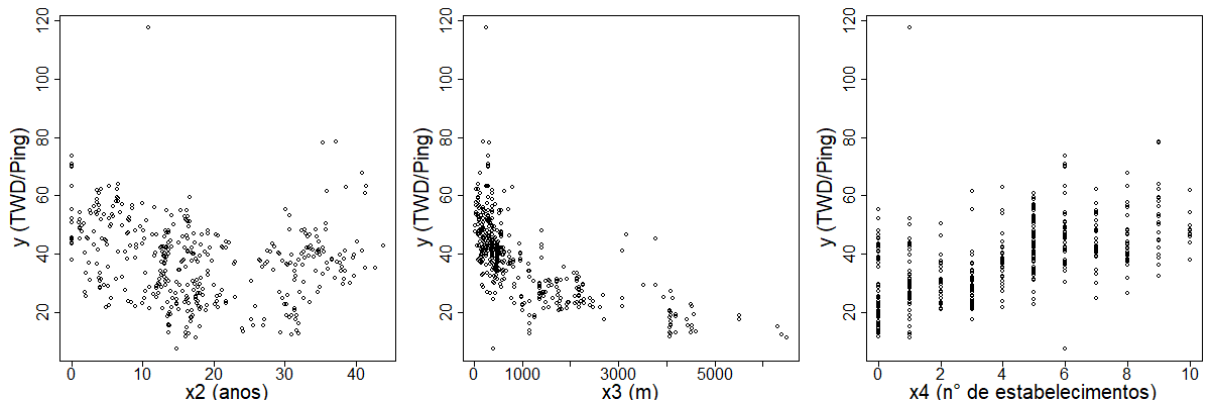


Figura 4 – Gráficos de dispersão entre o preço unitário dos imóveis por metro quadrado (Y) e as variáveis explicativas selecionadas: idade da casa (X_2), distância até a estação de metrô mais próxima (X_3) e número de lojas de conveniência próximas (X_4).

Devido à natureza positiva da variável resposta, adotamos a transformação logarítmica em Y . Assim, a especificação do modelo em termos dos parâmetros de regressão β é dada por $\log(\mu_i) = \beta_0 + \beta_1 X_{2i} + \beta_2 \log(X_{3i}) + \beta_3 X_{4i}$, para $i = 1, \dots, 414$.

Ressalta-se que a variável distância até a estação de metrô mais próxima (X_3) foi incluída na forma transformada $\log(X_3)$. Essa transformação foi adotada devido à elevada escala dessa variável em comparação às demais covariáveis, o que poderia comprometer a estabilidade numérica do procedimento de estimação. De fato, sem a transformação logarítmica, o algoritmo de maximização da verossimilhança apresentou dificuldades de convergência. A utilização de $\log(X_3)$ contribuiu para melhorar a estabilidade computacional e a adequação do ajuste.

De forma equivalente, o quantil predito μ_i pode ser expresso como:

$$\mu_i = \exp(\beta_0 + \beta_1 X_{2i} + \beta_2 \log(X_{3i}) + \beta_3 X_{4i}).$$

De forma análoga, a modelagem do parâmetro de dispersão é realizada por meio de uma função de ligação logarítmica, assegurando a positividade de σ_i . Assim, a especificação do modelo para a dispersão, em termos dos parâmetros de regressão γ , é dada por $\log(\sigma_i) = \gamma_0 + \gamma_1 X_{2i} + \gamma_2 X_{4i}$, para $i = 1, \dots, 414$. De forma equivalente, o parâmetro de dispersão pode ser expresso como:

$$\sigma_i = \exp(\gamma_0 + \gamma_1 X_{2i} + \gamma_2 X_{4i}).$$

7.1.1 Estimativas de Máxima Verossimilhança e Avaliação dos Modelos

A Tabela 6 apresenta as estimativas de Máxima Verossimilhança e os Erros Padrão para os modelos indicados, enquanto a Tabela 7 exhibe os valores dos critérios de informação AIC, BIC e Deviance obtidos para diferentes quantis. Complementarmente à análise numérica, a adequação desses modelos ao longo da distribuição condicional é verificada visualmente por meio de ferramentas de diagnóstico baseadas nos resíduos quantílicos.

A qualidade do ajuste dos modelos foi avaliada primeiramente por meio dos *worm plots*, apresentados na Figura 5 para diferentes níveis de quantil. Esses gráficos permitem verificar possíveis desvios dos resíduos em relação ao comportamento esperado, auxiliando na identificação de problemas de ajuste local. Adicionalmente, as Figuras 6, 7 e 8 detalham o comportamento global dos resíduos para os modelos HBS, LBS e NBS, respectivamente, através de gráficos de dispersão, densidade e Q-Q plot normal.

De modo geral, ao observar a Figura 5, os modelos HBS e LBS apresentam as curvas de resíduos mais próximas da linha horizontal zero e contidas dentro das bandas de referência. Em contraste, o modelo NBS exhibe desvios em formato de U nas extremidades, o que sugere que este modelo não consegue capturar adequadamente o comportamento das caudas da distribuição.

Essas evidências gráficas estão de acordo com os critérios de informação apresentados na Tabela 7, onde o HBS e o LBS registraram menores valores de AIC e BIC, reforçando o melhor desempenho desses dois modelos para a modelagem dos dados analisados em comparação ao NBS.

Tabela 6 – Estimativas de máxima verossimilhança e erros padrão para diferentes modelos e quantis τ .

τ		Estimativa					Erro Padrão				
		0.25	0.35	0.50	0.65	0.75	0.25	0.35	0.50	0.65	0.75
HBS	β_0	4.686	4.744	4.816	4.891	4.953	0.121	0.112	0.108	0.116	0.168
	β_1	-0.006	-0.006	-0.006	-0.005	-0.005	0.001	0.001	0.001	0.001	0.001
	β_2	-0.197	-0.196	-0.193	-0.191	-0.190	0.016	0.015	0.015	0.017	0.029
	β_3	0.031	0.027	0.023	0.018	0.013	0.006	0.005	0.005	0.005	0.008
	γ_0	-1.299	-1.286	-1.268	-1.249	-1.233	0.123	0.105	0.101	0.111	0.176
	γ_1	0.005	0.005	0.006	0.006	0.006	0.005	0.004	0.004	0.004	0.005
	γ_2	-0.063	-0.066	-0.072	-0.077	-0.082	0.017	0.016	0.016	0.020	0.041
LBS	β_0	4.668	4.728	4.808	4.888	4.952	0.119	0.111	0.109	0.115	0.145
	β_1	-0.007	-0.006	-0.006	-0.005	-0.005	0.002	0.001	0.001	0.001	0.001
	β_2	-0.196	-0.195	-0.193	-0.191	-0.190	0.016	0.015	0.015	0.017	0.024
	β_3	0.032	0.029	0.024	0.019	0.015	0.006	0.005	0.005	0.005	0.008
	γ_0	-1.934	-1.924	-1.909	-1.894	-1.881	0.125	0.101	0.095	0.104	0.158
	γ_1	0.006	0.006	0.006	0.006	0.006	0.005	0.004	0.004	0.004	0.004
	γ_2	-0.060	-0.063	-0.068	-0.073	-0.076	0.016	0.015	0.015	0.019	0.035
NBS	β_0	4.678	4.752	4.851	4.951	5.027	0.131	0.121	0.118	0.119	0.124
	β_1	-0.006	-0.006	-0.005	-0.005	-0.005	0.003	0.001	0.001	0.001	0.001
	β_2	-0.202	-0.202	-0.201	-0.201	-0.200	0.017	0.016	0.016	0.016	0.018
	β_3	0.031	0.028	0.024	0.020	0.017	0.008	0.006	0.006	0.006	0.009
	γ_0	-1.321	-1.317	-1.312	-1.306	-1.302	0.156	0.096	0.083	0.088	0.117
	γ_1	0.005	0.005	0.005	0.005	0.005	0.007	0.004	0.003	0.003	0.004
	γ_2	-0.041	-0.042	-0.044	-0.045	-0.046	0.015	0.013	0.013	0.015	0.024

Fonte: Elaborado pelos autores.

Os resultados apresentados na Tabela 6 evidenciam que as covariáveis selecionadas

exercem influência relevante sobre o preço unitário dos imóveis e que a magnitude desses efeitos varia conforme o quantil analisado, indicando que o impacto das características do imóvel sobre o preço não é uniforme em toda a distribuição. Como o modelo foi especificado com função de ligação logarítmica, as estimativas associadas aos parâmetros de regressão admitem interpretação multiplicativa, permitindo avaliar os efeitos das covariáveis em termos de variações percentuais no quantil condicional do preço.

O coeficiente associado à idade da casa (X_2) apresenta sinal negativo, indicando que imóveis mais antigos tendem a apresentar menor valor unitário. Em particular, no modelo HBS considerando quantil mediano ($\tau = 0.5$), a estimativa ($\hat{\beta}_1 = -0.006$) indica que um aumento de um ano na idade do imóvel implica uma redução aproximada de 0.6% no preço unitário mediano, mantendo-se as demais covariáveis constantes. Esse efeito reflete a depreciação natural das construções ao longo do tempo que pode estar associada ao desgaste estrutural do imóvel.

A distância até a estação de metrô mais próxima (X_3) foi incluída no modelo na escala logarítmica, por meio do termo $\beta_2 \log(X_3)$. Nessa especificação, a média condicional do preço unitário satisfaz a relação

$$\mu_i \propto X_3^{\beta_2},$$

o que implica que o coeficiente β_2 pode ser interpretado como uma *elasticidade do preço em relação à distância ao metrô*. Assim, uma variação percentual em X_3 está associada a uma variação percentual aproximada de $100 \times \beta_2\%$ no preço esperado do imóvel, mantendo-se as demais covariáveis constantes.

No quantil mediano ($\tau = 0.5$), a estimativa $\hat{\beta}_2 = -0.196$ indica que um aumento de 1% na distância até a estação de metrô está associado a uma redução de aproximadamente 0.196% no preço unitário mediano do imóvel. O sinal negativo do coeficiente confirma que imóveis localizados mais próximos ao transporte público tendem a ser mais valorizados, evidenciando a importância da acessibilidade urbana e da proximidade a sistemas de transporte coletivo na formação dos preços no mercado imobiliário.

Por outro lado, o número de lojas de conveniência nas proximidades (X_4) apresenta efeito positivo, indicando que regiões com maior oferta de serviços e comércio local estão associadas a preços unitários mais elevados. Ainda no quantil mediano, a estimativa ($\hat{\beta}_3 = 0.023$) indica que a presença de uma loja adicional implica um aumento de cerca de 2.3% no preço unitário mediano, reforçando o papel da oferta de serviços e da infraestrutura urbana na formação dos preços.

De modo geral, os sinais e magnitudes dos efeitos estimados são consistentes entre os diferentes modelos e estão em consonância com a teoria econômica e evidências empíricas do mercado imobiliário, indicando que características estruturais e atributos de localização atuam de forma multiplicativa na determinação do preço dos imóveis ao longo da distribuição

condicional.

Tabela 7 – Valores dos critérios de informação AIC, BIC e Deviance para os modelos LBS, NBS e HBS, estimados em diferentes quantis τ .

Modelo	Critério	τ				
		0.25	0.35	0.50	0.65	0.75
HBS	AIC	2901.747	2900.726	2899.293	2897.687	2896.289
	BIC	2929.928	2928.907	2927.474	2925.868	2924.470
	Deviance	2887.747	2886.726	2885.293	2883.687	2882.289
LBS	AIC	2905.937	2905.022	2903.707	2902.265	2901.079
	BIC	2934.118	2933.203	2931.888	2930.446	2929.260
	Deviance	2891.937	2891.022	2889.707	2888.265	2887.079
NBS	AIC	2972.994	2972.674	2972.221	2971.747	2971.387
	BIC	3001.175	3000.855	3000.402	2999.928	2999.568
	Deviance	2958.994	2958.674	2958.221	2957.747	2957.387

Fonte: Elaborado pelos autores.

Com base nos valores apresentados na Tabela 7, observa-se que o modelo HBS apresenta desempenho superior em relação aos modelos LBS e NBS para todos os quantis analisados. O modelo LBS apresenta desempenho intermediário, com valores de AIC, BIC e Deviance inferiores aos do NBS, mas superiores aos do HBS em todos os quantis, sugerindo que, embora capture parte da estrutura dos dados, é menos flexível que o HBS. Esses resultados reforçam a adequação do modelo HBS para a modelagem quantílica dos dados considerados, evidenciando sua maior capacidade de acomodar a assimetria e a variabilidade presentes na distribuição do preço unitário dos imóveis.

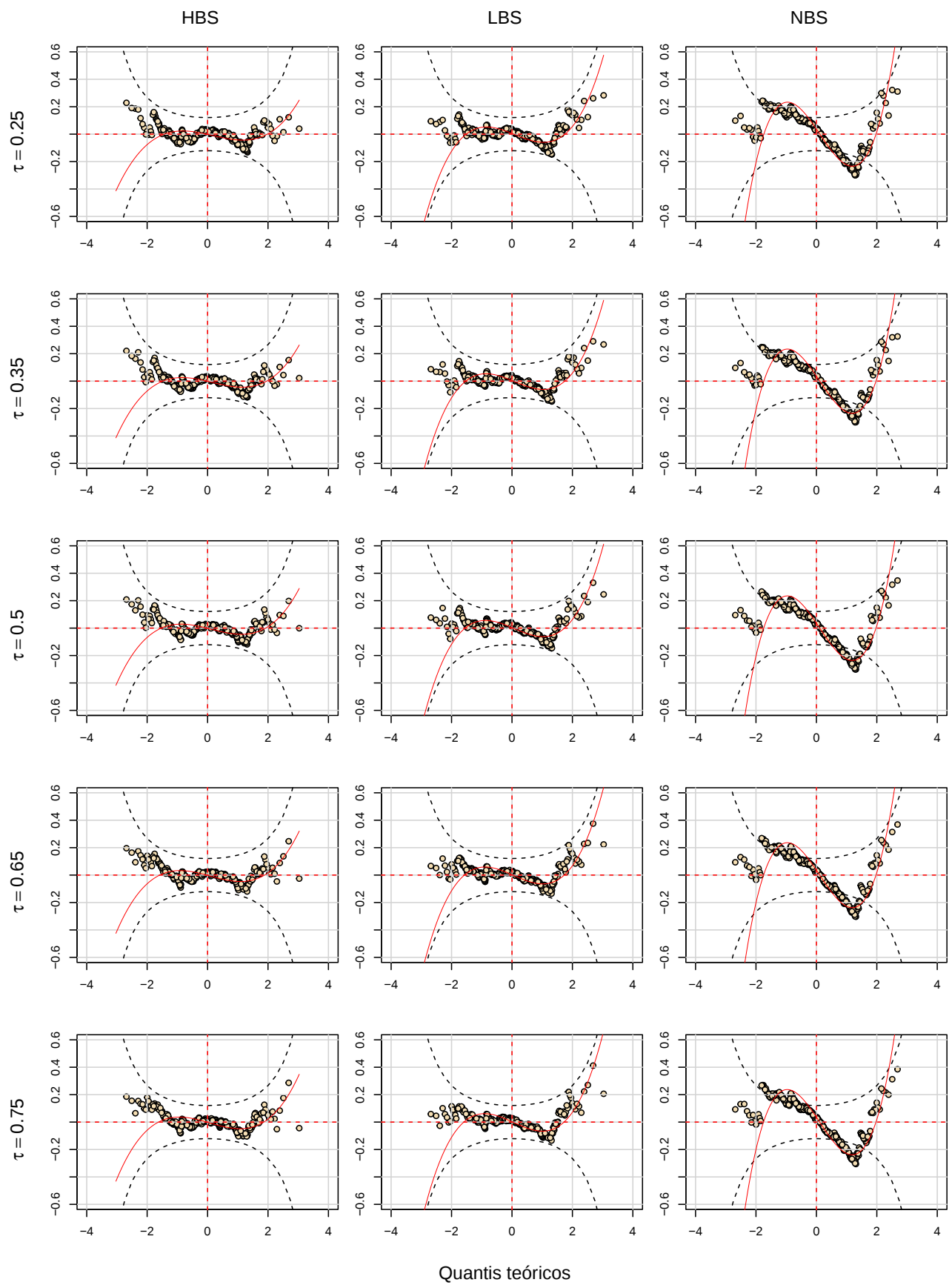


Figura 5 – Worm plots para os modelos HBS, LBS e NBS ajustados em diferentes níveis de quantil ($\tau = 0.1, 0.25, 0.5, 0.75, 0.9$).

De modo geral, ao observar a Figura 5, os modelos HBS e LBS apresentam as curvas de resíduos mais próximas da linha horizontal zero e contidas dentro das bandas de referência. Em contraste, o modelo NBS exibe desvios em formato de U nas extremidades, o que sugere que este modelo não consegue capturar adequadamente o comportamento das caudas da distribuição.

As Figuras 6, 7 e 8 apresentam os gráficos de diagnóstico dos resíduos quantílicos para os modelos HBS, LBS e NBS. Em geral, os resíduos dos modelos HBS e LBS mostram-se mais concentrados em torno de zero, sem padrões sistemáticos evidentes nos gráficos contra valores ajustados e índice, indicando bom ajuste aos dados. Além disso, os gráficos Q-Q revelam melhor alinhamento à reta teórica para HBS e LBS, enquanto o modelo NBS apresenta maiores desvios, especialmente nas caudas.

Essas evidências gráficas estão de acordo com os critérios de informação apresentados na Tabela 7, onde o HBS e o LBS registraram menores valores de AIC e BIC, reforçando o melhor desempenho desses dois modelos para a modelagem dos dados analisados em comparação ao NBS.

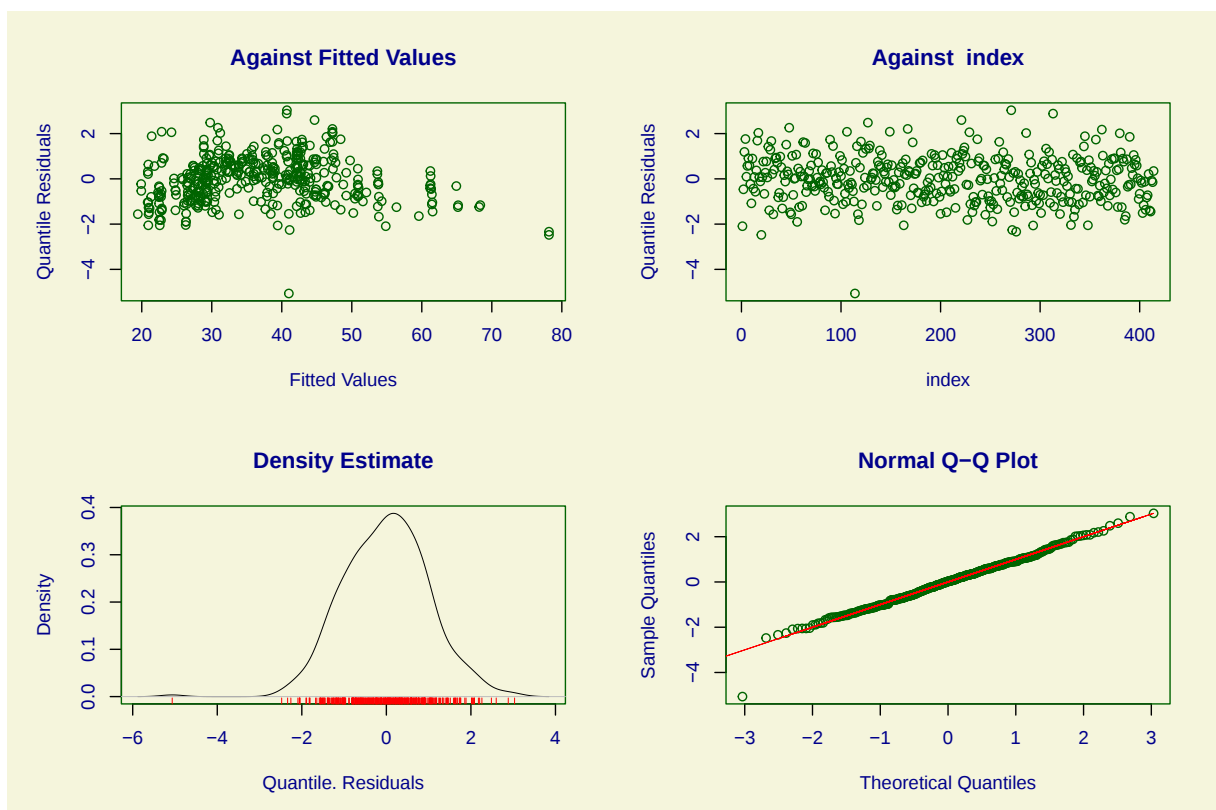


Figura 6 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo HBS

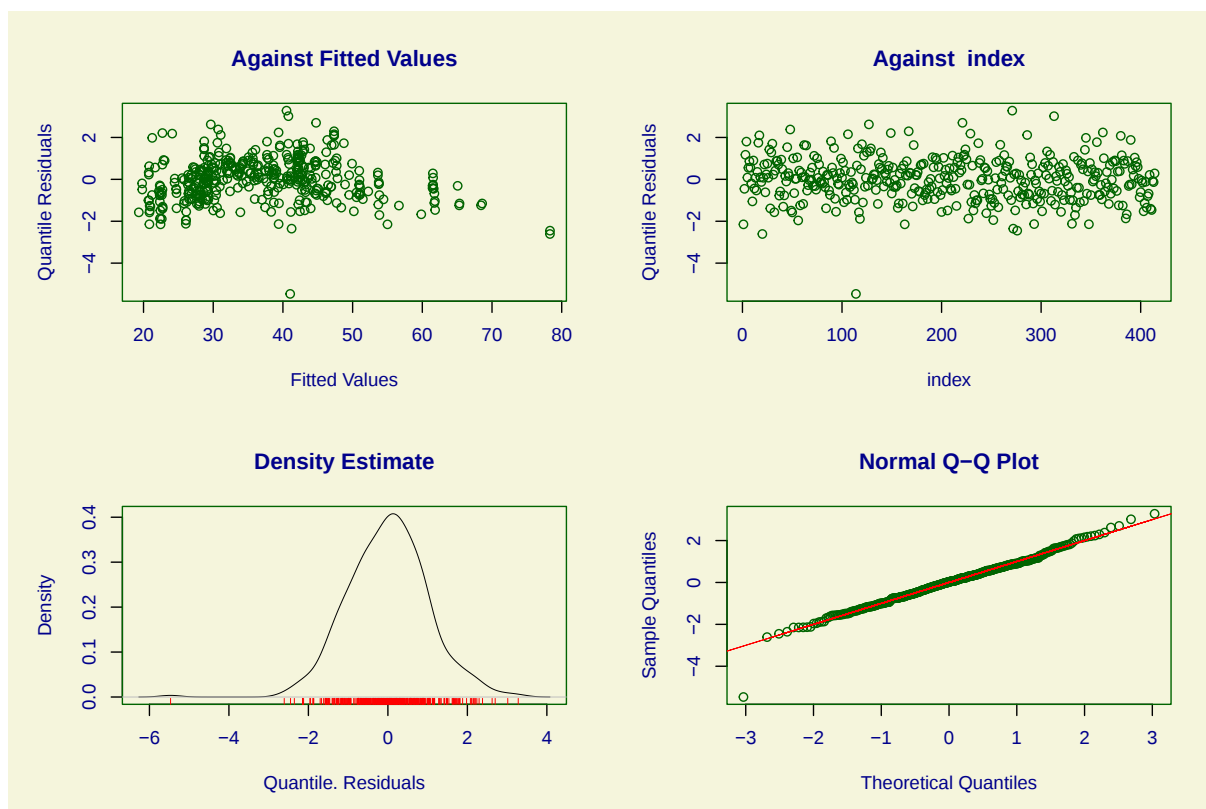


Figura 7 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo LBS

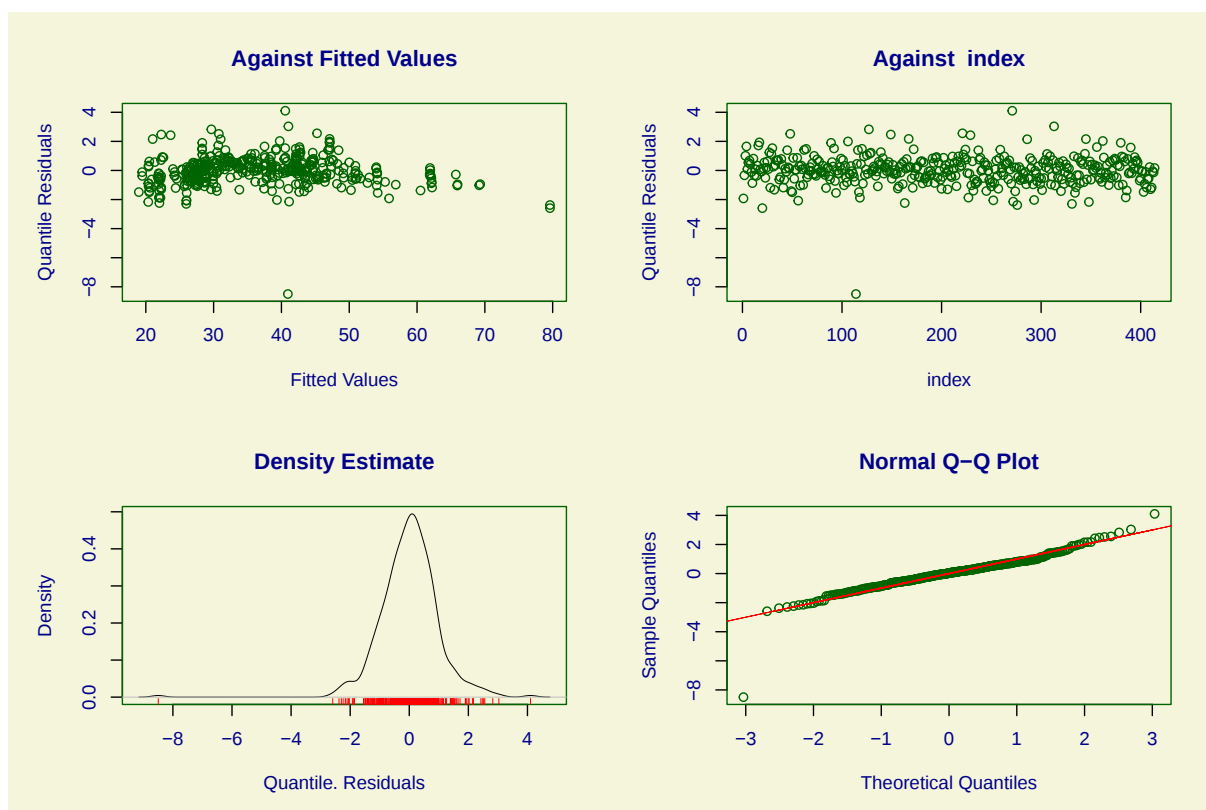


Figura 8 – Gráficos de diagnóstico dos resíduos - dados de avaliação imobiliária - modelo NBS

7.2 Dados censurados - crescimento de cordeiros Santa Inês

Para ilustrar a aplicabilidade do modelo proposto, foram utilizados dados de um estudo experimental com 178 cordeiros da raça Santa Inês, oriundos de partos simples, acompanhados em uma propriedade tecnificada (KORITIAKI, 2015). O objetivo central da análise foi identificar os determinantes maternos e neonatais que influenciam o tempo necessário para que os animais atinjam o peso-alvo de 15 kg durante o período pré-desmame.

O acompanhamento ocorreu diariamente desde o nascimento até o evento de interesse (atingir 15 kg) ou até o limite censurador de 70 dias de vida. As variáveis registradas no estudo incluem características morfométricas e indicadores maternos, tais como:

- **Sexo (SEX):** variável categórica indicando o sexo do cordeiro (fêmea ou macho).
- **Idade da matriz (IMP):** classificada entre matrizes jovens (menos de oito dentes) e matrizes adultas (oito dentes).
- **Peso da matriz ao parto (PMP):** medida em quilogramas (kg), refletindo a condição nutricional da mãe.
- **Peso ao nascer (PEO):** peso neonatal registrado em quilogramas (kg).
- **Medidas morfométricas neonatais:** incluem a altura na cernelha (ALT), o comprimento corporal (COM) e o perímetro torácico (PTO), todas medidas em centímetros (cm).

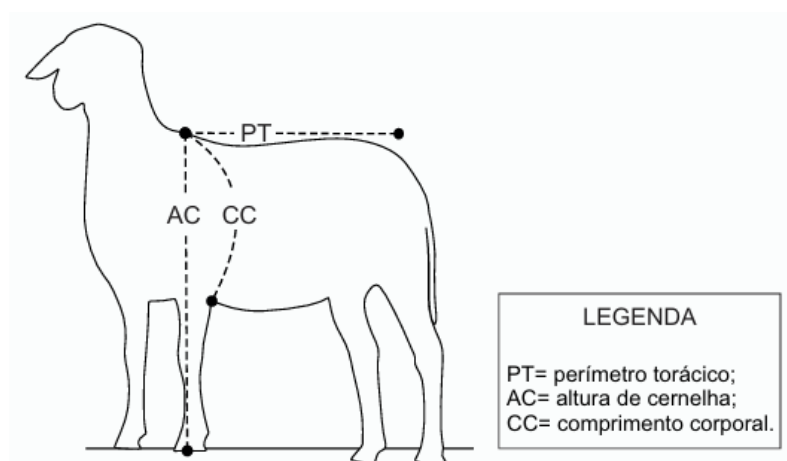


Figura 9 – Posição relativa às mensurações corporais tomadas em ovinos. (KORITIAKI, 2015)

- **Tempo:** variável resposta que indica o número de dias até o atingimento do peso-alvo, apresentando censura à direita para animais que não alcançaram 15 kg no período observado.

As estatísticas descritivas das variáveis consideradas na análise, estratificadas segundo o status do evento (atingimento ou não do peso-alvo de 15 kg), são apresentadas na Tabela 8. Observa-se que o tempo mediano até o atingimento do peso foi substancialmente menor entre os animais que efetivamente alcançaram 15 kg dentro do período de acompanhamento, enquanto os indivíduos censurados apresentaram tempos concentrados em 70 dias (data limite do estudo). De modo geral, diferenças também podem ser verificadas nas características morfológicas e nos indicadores maternos entre os grupos, sugerindo possível associação dessas variáveis com o crescimento pré-desmame dos cordeiros.

Tabela 8 – Estatísticas descritivas dos dados de crescimento de cordeiros segundo o total de observações e indicador de censura δ .

Características	Total N = 178	$\delta = 0$ N = 82	$\delta = 1$ N = 96
Tempo até 15 kg (dias)			
Média (DP)	58.7 (12.2)	69.3 (1.0)	49.7 (9.9)
Mediana (Q1, Q3)	64.7 (48.4, 70.0)	70.0 (69.0, 70.0)	49.1 (41.8, 57.3)
Mín – Máx	32.2 – 70.0	67.0 – 70.0	32.2 – 69.5
COM - Comprimento corporal (cm)			
Média (DP)	28.4 (2.7)	27.6 (2.8)	29.1 (2.5)
Mediana (Q1, Q3)	28.0 (27.0, 30.0)	27.0 (26.0, 29.0)	29.0 (27.0, 31.0)
Mín – Máx	21.0 – 35.0	21.0 – 35.0	23.0 – 35.0
PTO - Perímetro torácico (cm)			
Média (DP)	38.4 (2.5)	37.4 (2.2)	39.2 (2.5)
Mediana (Q1, Q3)	38.0 (37.0, 40.0)	37.0 (36.0, 39.0)	39.0 (38.0, 41.0)
Mín – Máx	28.0 – 44.0	30.0 – 43.0	28.0 – 44.0
ALT - Altura na cernelha (cm)			
Média (DP)	40.1 (2.4)	39.6 (2.5)	40.5 (2.3)
Mediana (Q1, Q3)	40.0 (39.0, 42.0)	40.0 (38.0, 41.0)	40.5 (39.0, 42.0)
Mín – Máx	33.0 – 46.0	33.0 – 46.0	33.0 – 46.0
Sexo			
F	87 (49%)	43 (52%)	44 (46%)
M	91 (51%)	39 (48%)	52 (54%)
IMP - Idade da matriz			
< 8	86 (48%)	38 (46%)	48 (50%)
= 8	92 (52%)	44 (54%)	48 (50%)
PMP - Peso da matriz ao parto (kg)			
Média (DP)	53.9 (8.2)	49.3 (7.3)	57.8 (6.8)
Mediana (Q1, Q3)	55.2 (47.7, 59.2)	48.5 (44.3, 55.5)	57.6 (53.5, 62.2)
Mín – Máx	33.0 – 73.8	33.0 – 66.3	37.0 – 73.8
PEO - Peso ao nascer (kg)			
Média (DP)	4.0 (0.7)	3.7 (0.7)	4.3 (0.6)
Mediana (Q1, Q3)	4.0 (3.6, 4.5)	3.6 (3.2, 4.1)	4.3 (3.8, 4.7)
Mín – Máx	2.1 – 6.1	2.1 – 5.5	3.1 – 6.1

Dados apresentados como média (desvio padrão), mediana (Q1, Q3) ou n (%).

Considerando a presença de observações censuradas à direita no tempo de sobrevivência

dos cordeiros, o ajuste do modelo de regressão quantílica com censura, utilizando as distribuições NBS, HBS e LBS, foi realizado para diferentes quantis $\tau \in \{0.25, 0.35, 0.50, 0.65, 0.75\}$. Devido à natureza estritamente positiva da variável resposta Y (tempo), adotou-se a função de ligação logarítmica para o parâmetro μ_i , que representa o τ -ésimo quantil condicional.

A seleção das variáveis explicativas foi conduzida por meio de um procedimento sistemático baseado no critério de informação generalizado (GAIC), utilizando a função `stepGAIC()` do pacote `gamlss`, na formulação apropriada para dados com censura à direita via `gamlss.cens`. Inicialmente, realizou-se a seleção das covariáveis associadas ao preditor do parâmetro quantílico μ_i , mantendo-se fixa a estrutura do parâmetro de dispersão. Em seguida, considerando-se a especificação final obtida para μ_i , aplicou-se novamente o procedimento `stepGAIC()` para avaliar a inclusão de covariáveis no parâmetro de dispersão σ_i .

Esse procedimento sequencial possibilitou a identificação de um modelo parcimonioso, equilibrando qualidade de ajuste e complexidade, conforme o critério de informação adotado.

Com base na análise exploratória e no processo de seleção via GAIC, foram selecionadas como variáveis explicativas para o parâmetro μ_i o comprimento corporal ao nascimento (COM), a idade da matriz (IMP), o peso da matriz ao parto (PMP) e o peso ao nascer (PEO). Essas variáveis contemplam tanto características morfométricas neonatais quanto condições maternas potencialmente associadas ao crescimento pré-desmame dos cordeiros.

Assim, a especificação do modelo em termos dos parâmetros de regressão β é dada por

$$\log(\mu_i) = \beta_0 + \beta_1 \text{COM}_i + \beta_2 \text{IMP}_i + \beta_3 \text{PMP}_i + \beta_4 \text{PEO}_i, \quad (7.1)$$

para $i = 1, \dots, 178$.

De forma equivalente, o quantil condicional predito é dado por

$$\mu_i = \exp(\beta_0 + \beta_1 \text{COM}_i + \beta_2 \text{IMP}_i + \beta_3 \text{PMP}_i + \beta_4 \text{PEO}_i). \quad (7.2)$$

Para o parâmetro de dispersão, foram avaliadas especificações alternativas com a inclusão de covariáveis explicativas. Contudo, tais inclusões não resultaram em melhorias relevantes segundo os critérios de informação nem em ganhos de estabilidade nas estimativas. Dessa forma, optou-se por uma estrutura mais parcimoniosa, modelando o parâmetro de dispersão apenas por intercepto.

Assim, a especificação adotada para o parâmetro de dispersão é dada por

$$\log(\sigma_i) = \gamma_0, \quad i = 1, \dots, 178.$$

7.2.1 Estimativas de Máxima Verossimilhança e Avaliação dos Modelos

A Tabela 9 apresenta as estimativas de máxima verossimilhança e os respectivos erros padrão para os modelos considerados, enquanto a Tabela 10 resume os valores dos critérios

de informação AIC, BIC e Deviance obtidos para os diferentes quantis analisados.

Além da avaliação numérica, a adequação dos modelos ao longo da distribuição condicional do tempo até o evento é investigada por meio de análises gráficas baseadas em resíduos quantílicos. Em particular, gráficos de dispersão dos resíduos, estimativas de densidade e gráficos Q–Q normal (Figuras 10, 11 e 12) são utilizados para avaliar a adequação global dos modelos e a presença de padrões sistemáticos não capturados pelas especificações consideradas. Adicionalmente, os *worm plots* apresentados nas Figuras 13, 14 e 15 permitem identificar possíveis desvios locais do ajuste em diferentes regiões das covariáveis.

Tabela 9 – Estimativas de máxima verossimilhança e erros padrão para diferentes modelos, parâmetros e quantis τ .

τ		Estimativa					Erro Padrão				
		0.25	0.35	0.50	0.65	0.75	0.25	0.35	0.50	0.65	0.75
HBS	β_0	6.533	6.523	6.521	6.520	6.521	0.530	0.393	0.299	0.252	0.233
	β_1	-0.020	-0.020	-0.020	-0.020	-0.020	0.013	0.010	0.008	0.008	0.008
	β_2	0.136	0.136	0.136	0.136	0.136	0.058	0.048	0.041	0.038	0.038
	β_3	-0.022	-0.022	-0.022	-0.022	-0.022	0.005	0.004	0.003	0.003	0.002
	β_4	-0.165	-0.165	-0.165	-0.165	-0.165	0.042	0.035	0.031	0.029	0.030
	γ_0	-1.255	-1.255	-1.255	-1.255	-1.255	0.064	0.064	0.064	0.064	0.064
LBS	β_0	6.553	6.553	6.551	6.552	6.552	0.466	0.368	0.294	0.255	0.238
	β_1	-0.021	-0.021	-0.021	-0.021	-0.021	0.012	0.010	0.009	0.008	0.008
	β_2	0.133	0.133	0.133	0.133	0.133	0.057	0.049	0.043	0.040	0.039
	β_3	-0.022	-0.022	-0.022	-0.022	-0.022	0.004	0.004	0.003	0.003	0.003
	β_4	-0.166	-0.166	-0.166	-0.166	-0.166	0.042	0.036	0.032	0.031	0.030
	γ_0	-1.876	-1.876	-1.877	-1.876	-1.876	0.059	0.059	0.059	0.061	0.062
NBS	β_0	6.497	6.497	6.494	6.497	6.497	0.290	0.279	0.265	0.252	0.242
	β_1	-0.020	-0.020	-0.020	-0.020	-0.020	0.010	0.010	0.009	0.009	0.009
	β_2	0.121	0.121	0.121	0.121	0.121	0.049	0.048	0.045	0.043	0.041
	β_3	-0.021	-0.021	-0.021	-0.021	-0.021	0.003	0.003	0.003	0.003	0.003
	β_4	-0.162	-0.162	-0.162	-0.162	-0.162	0.038	0.037	0.035	0.033	0.032
	γ_0	-1.291	-1.291	-1.292	-1.291	-1.291	0.046	0.048	0.052	0.057	0.061

Fonte: Elaborado pelos autores.

Os resultados apresentados na Tabela 9 indicam que as covariáveis selecionadas exercem influência relevante sobre o tempo necessário para que os cordeiros atinjam o peso-alvo de 15 kg. Observa-se também, uma estabilidade nas estimativas pontuais, indicando que não houve diferença substancial no efeito das covariáveis segundo o quantil analisado. Como o modelo foi especificado com função de ligação logarítmica para o parâmetro μ_i , que representa o quantil condicional do tempo até o evento, as estimativas dos parâmetros de regressão admitem interpretação multiplicativa, permitindo avaliar os efeitos das covariáveis em termos de variações percentuais no tempo de sobrevivência condicional.

O coeficiente associado ao comprimento corporal ao nascimento (COM) apresenta sinal negativo em todos os modelos e quantis considerados. Esse resultado indica que cordeiros com

maior comprimento corporal tendem a atingir o peso-alvo em menor tempo. Em particular, no modelo HBS considerando o quantil mediano ($\tau = 0.5$), a estimativa $\hat{\beta}_1 = -0.020$ sugere que um aumento de uma unidade em COM está associado a uma redução aproximada de 2.02% no tempo mediano até o atingimento de 15 kg, mantendo-se as demais covariáveis constantes. Esse efeito reflete a importância do desenvolvimento morfométrico inicial no desempenho produtivo dos cordeiros durante o período pré-desmame.

A variável idade da matriz (IMP) apresenta coeficiente positivo ao longo de todos os quantis analisados. Esse resultado indica que cordeiros oriundos de matrizes adultas tendem a apresentar maior tempo até o evento quando comparados àqueles provenientes de matrizes jovens. No quantil mediano, a estimativa $\hat{\beta}_2 = 0.136$ implica que cordeiros oriundos de matrizes adultas apresentam um tempo mediano aproximadamente 14,6% maior para atingir o peso-alvo quando comparados aos provenientes de matrizes jovens, mantendo-se as demais covariáveis constantes, o que pode estar associado a diferenças fisiológicas, produtivas ou de manejo entre matrizes de distintas faixas etárias.

O peso da matriz ao parto (PMP) apresenta coeficiente negativo em todos os modelos e quantis considerados, indicando que melhores condições corporais maternas estão associadas a menor tempo até o evento. No quantil $\tau = 0.5$, a estimativa $\hat{\beta}_3 = -0.022$ sugere que um aumento de 1 kg no peso da matriz ao parto está associado a uma redução aproximada de 2.22% no tempo mediano necessário para que o cordeiro atinja o peso-alvo, evidenciando a influência direta do estado nutricional materno sobre o crescimento inicial da prole.

De forma semelhante, o peso ao nascer (PEO) apresenta efeito negativo e consistente ao longo de toda a distribuição condicional do tempo. No quantil mediano, a estimativa $\hat{\beta}_4 = -0.165$ indica que um aumento de 1 kg no peso ao nascer está associado a uma redução aproximada de 16.5% no tempo mediano até o evento, destacando a importância das condições neonatais iniciais para o desempenho produtivo subsequente dos cordeiros.

Em relação ao parâmetro de dispersão, observa-se que a modelagem foi realizada apenas com intercepto, uma vez que a inclusão de covariáveis nesse componente não resultou em ganho significativo na qualidade do ajuste. As estimativas de γ_0 permanecem estáveis ao longo dos quantis analisados, sugerindo que a variabilidade residual do tempo de sobrevivência não explicada pelas covariáveis incluídas é relativamente constante.

De modo geral, os sinais e magnitudes dos efeitos estimados são consistentes entre os modelos HBS, LBS e NBS e ao longo dos diferentes quantis considerados.

Tabela 10 – Valores dos critérios de informação AIC, BIC e Deviance para os modelos censurados LBS, NBS e HBS, estimados em diferentes quantis.

Modelo	Critério	τ				
		0.25	0.35	0.50	0.65	0.75
LBS	AIC	920.162	885.337	868.062	886.658	927.254
	BIC	939.253	904.428	887.153	905.749	946.345
	Deviance	908.162	873.337	856.062	874.658	915.254
NBS	AIC	933.231	891.675	871.187	892.285	936.474
	BIC	952.322	910.766	890.278	911.376	955.565
	Deviance	921.231	879.675	859.187	880.285	924.474
HBS	AIC	966.368	914.672	867.640	844.647	844.958
	BIC	985.459	933.763	886.730	863.738	864.049
	Deviance	954.368	902.672	855.640	832.647	832.958

Fonte: Elaborado pelos autores.

Com base nos valores apresentados na Tabela 10, observa-se que o modelo baseado na distribuição LBS apresenta desempenho superior ao modelo NBS em todos os quantis considerados, devido a menores valores de AIC, BIC e Deviance. Esse resultado indica que a estrutura de assimetria e curtose incorporada pela distribuição LBS fornece maior flexibilidade para acomodar as características dos dados censurados em comparação com a especificação mais restritiva da NBS.

Ao analisar o comportamento dos critérios de informação ao longo dos quantis, nota-se que, para quantis mais baixos ($\tau = 0.25$ e $\tau = 0.35$), o modelo LBS apresenta os menores valores de AIC, BIC e Deviance, sugerindo melhor adequação na modelagem da cauda inferior da distribuição condicional do tempo até o evento. Por outro lado, a partir do quantil mediano ($\tau = 0.50$), o modelo HBS passa a apresentar desempenho superior aos demais modelos, destacando-se especialmente nos quantis superiores ($\tau = 0.65$ e $\tau = 0.75$), nos quais alcança os menores valores de AIC, BIC e Deviance. Esse comportamento indica que a HBS é mais adequada para descrever a cauda superior da distribuição condicional, associada a tempos mais longos até o evento.

De modo geral, os resultados evidenciam que tanto a LBS quanto a HBS constituem alternativas mais flexíveis e robustas em relação à NBS, apresentando ganhos substanciais de ajuste ao longo de diferentes regiões da distribuição condicional. Em particular, a superioridade da HBS em quantis elevados sugere que a incorporação de maior flexibilidade na forma da distribuição é crucial para capturar adequadamente o comportamento dos tempos mais extremos, reforçando a relevância do uso de modelos quantílicos baseados em distribuições mais gerais no contexto de dados de sobrevivência com censura.

As Figuras 13, 14 e 15 apresentam os gráficos *worm plot* dos modelos NBS, LBS e HBS, respectivamente, ajustados para $\tau = 0.5$. O *worm plot* consiste em uma ferramenta diagnóstica baseada nos resíduos quantílicos normalizados, permitindo avaliar a adequação

local do modelo em diferentes intervalos da variável explicativa considerada, para identificar se a distribuição ajustada é adequada, seja de forma geral ou dentro de intervalos não sobrepostos de uma ou duas variáveis explicativas. (STASINOPOULOS et al., 2017)

No gráfico 13, observa-se a presença de curvaturas mais acentuadas em alguns intervalos, com desvios em relação à linha horizontal de referência. Esse comportamento sugere indícios de inadequação na modelagem da dispersão e possível deficiência no ajuste das caudas da distribuição, indicando que a estrutura assumida pode não capturar completamente o comportamento condicional dos dados.

Por outro lado, as figuras 14 e 15, referentes aos modelos envolvendo LBS e HBS, respectivamente, apresentam curvas mais próximas da linha horizontal, com menor amplitude de oscilação e melhor alinhamento dentro dos envelopes simulados. Esse padrão indica ajuste mais adequado tanto da média condicional quanto da dispersão ao longo dos diferentes grupos analisados.

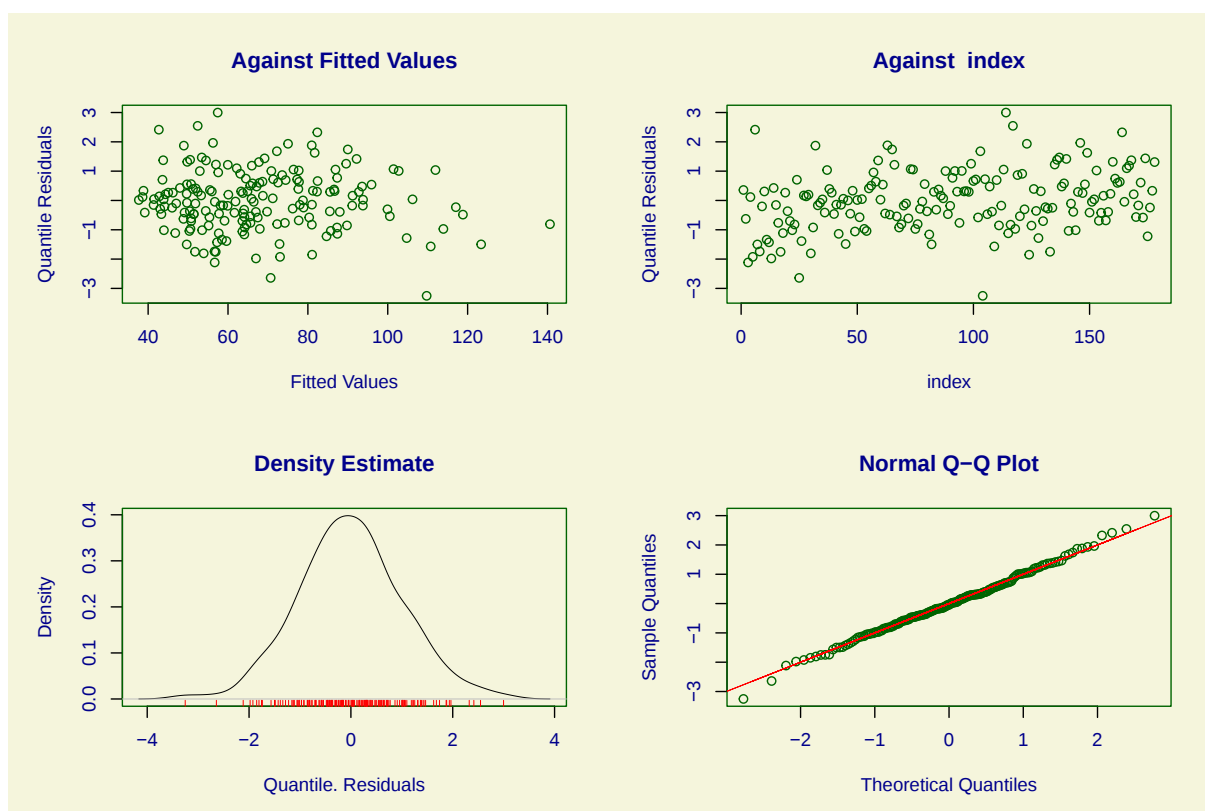


Figura 10 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo NBS

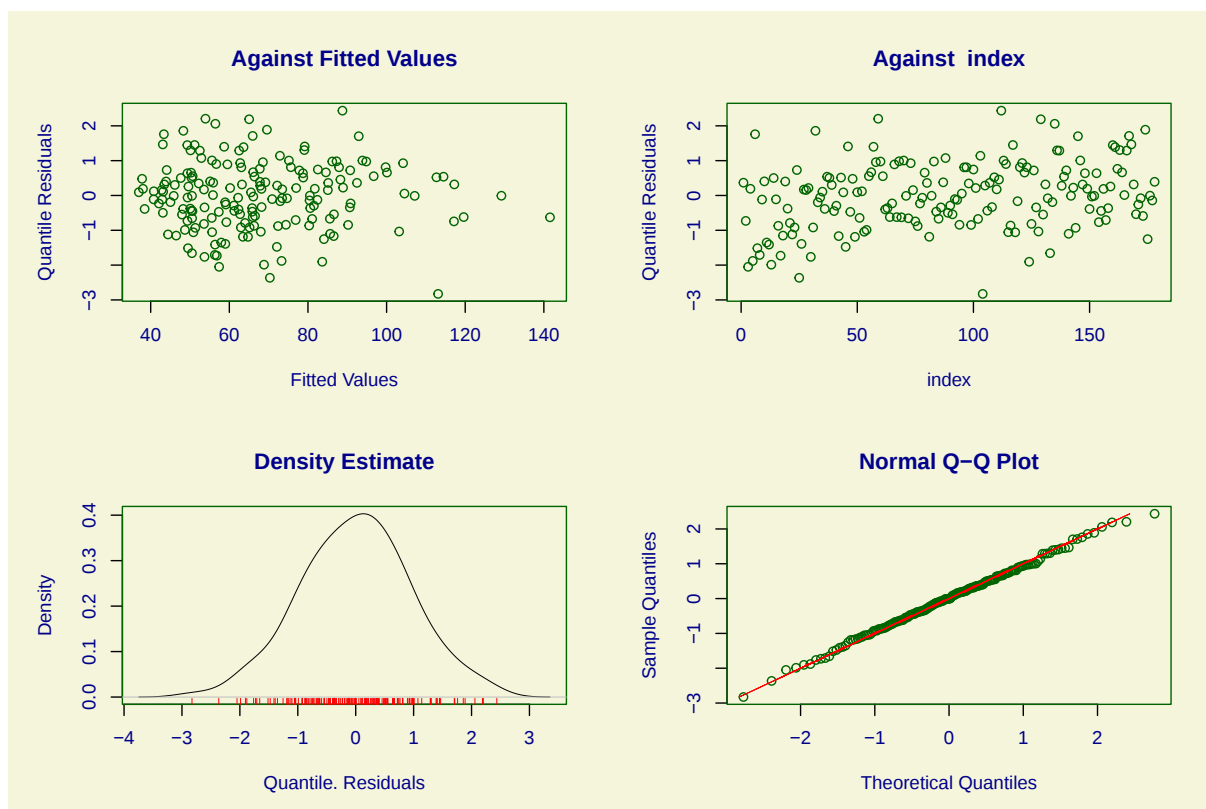


Figura 11 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo LBS

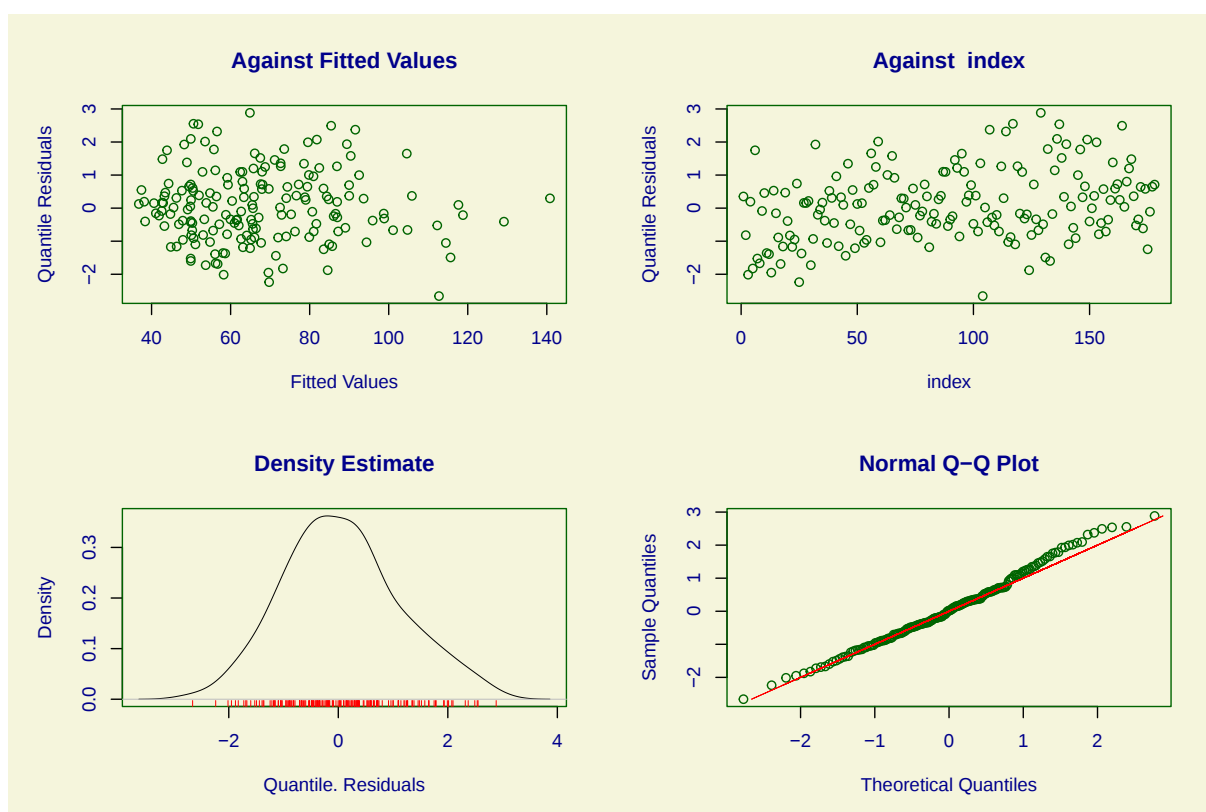


Figura 12 – Gráficos de diagnóstico dos resíduos - crescimento de cordeiros Santa Inês - modelo HBS

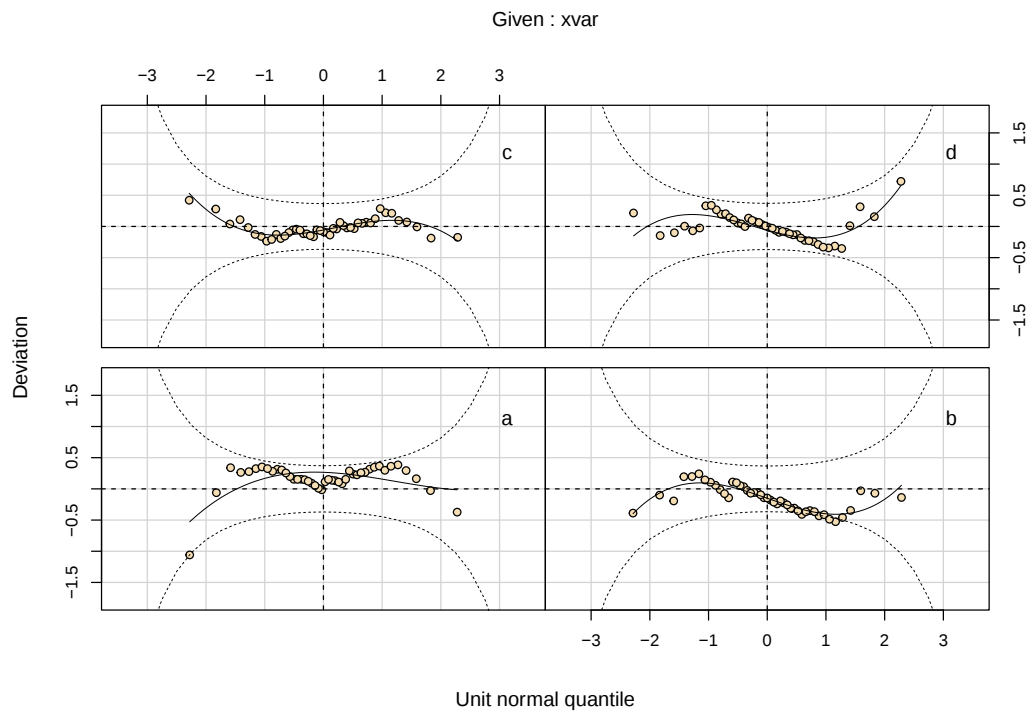


Figura 13 – Worm plot do modelo ajustado NBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.

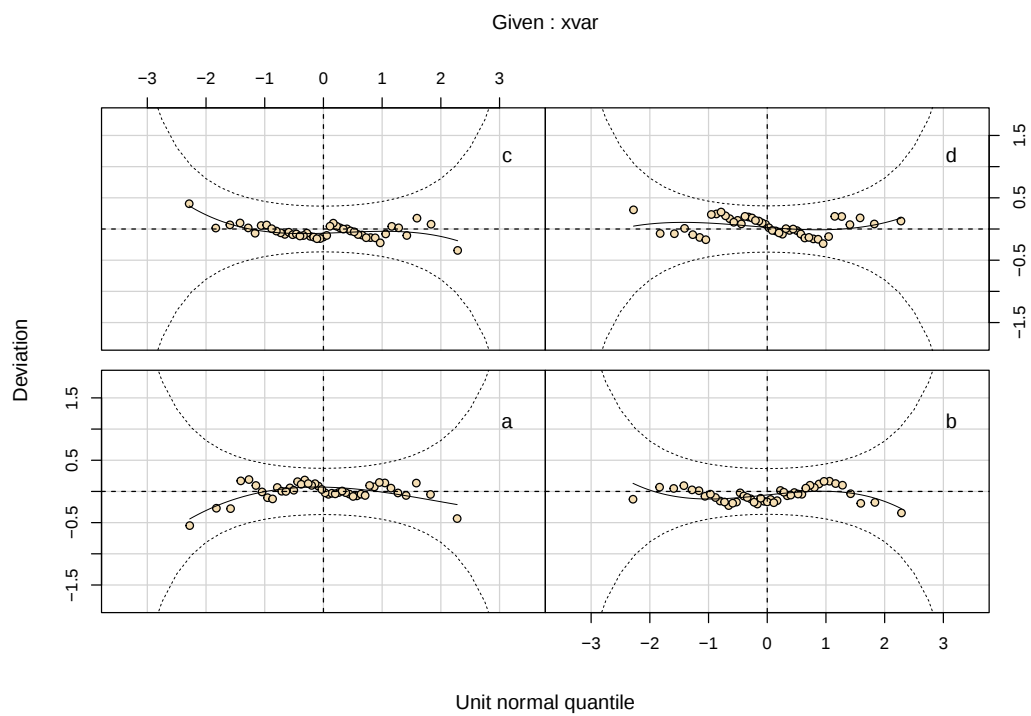


Figura 14 – Worm plot do modelo ajustado LBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.

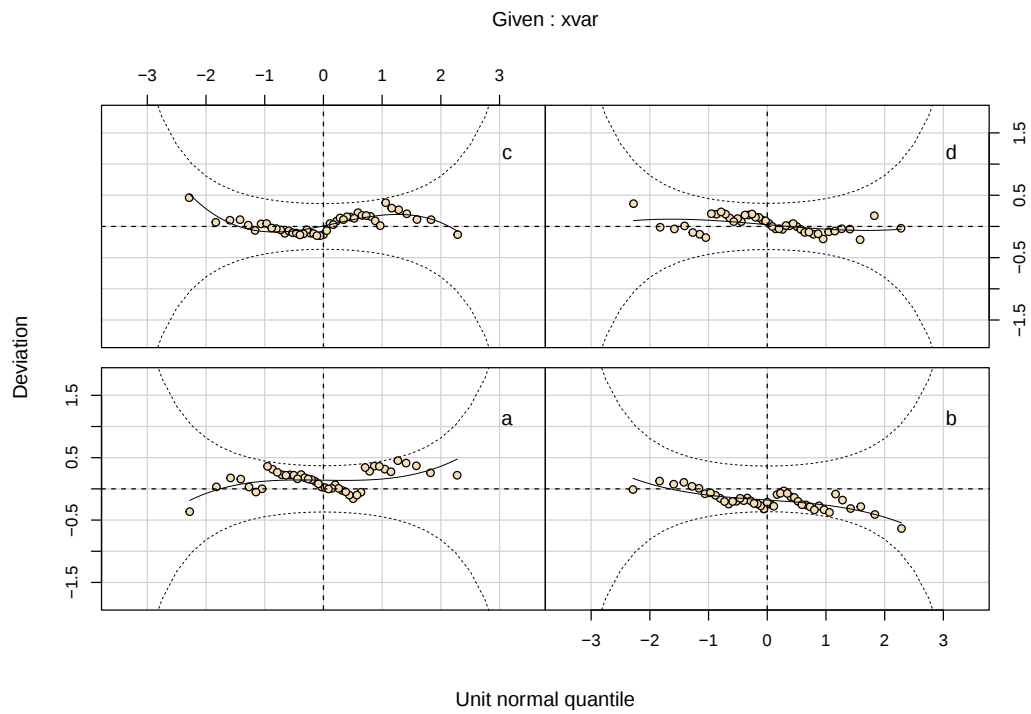


Figura 15 – Worm plot do modelo ajustado HBS com censura. Os resíduos são apresentados por quartis do produto das covariáveis (PMP, PEO e COM), em que (a) primeiro quartil, (b) segundo quartil, (c) terceiro quartil e (d) quarto quartil.

De modo geral, os resultados obtidos por meio dos *worm plots* estão em conformidade com os critérios de informação apresentados anteriormente (AIC, BIC e Deviance), reforçando que os modelos LBS e HBS constituem alternativas mais adequadas quando comparados ao modelo NBS para os dados em estudo.

CONCLUSÃO

O presente trabalho atingiu seu objetivo principal ao propor e desenvolver novos modelos de regressão quantílica paramétrica para dados contínuos e positivos, baseados nas distribuições LBS e HBS. Isso permitiu preencher uma lacuna na literatura, visto que, embora a LBS já fosse conhecida, sua aplicação em regressão quantílica ainda não havia sido explorada, e a distribuição HBS, construída a partir do núcleo da secante hiperbólica, representa uma novidade na modelagem de dados assimétricos.

Considerações sobre as Simulações

Os estudos de simulação de Monte Carlo comprovaram a eficiência do método de Máxima Verossimilhança para a estimação dos parâmetros dos modelos propostos. Os resultados evidenciaram que:

- Os estimadores comportaram-se de maneira consistente, apresentando redução do Viés e da Raiz do Erro Quadrático Médio (REQM) à medida que o tamanho da amostra aumentou ($n = 50$ para $n = 200$).
- O bom desempenho manteve-se tanto para dados completos quanto em cenários com diferentes percentuais de censura (20% e 40%), validando a robustez dos modelos propostos para aplicações em análise de sobrevivência.

Considerações sobre as Aplicações Empíricas

A aplicação aos dados reais demonstrou a superioridade e a flexibilidade das abordagens propostas em comparação ao modelo clássico baseado na distribuição Birnbaum-Saunders com núcleo normal (NBS):

1. **Dados Não Censurados (Mercado Imobiliário):** Na modelagem do preço de imóveis, os modelos HBS e LBS apresentaram melhor ajuste que o NBS, evidenciado pelos

menores valores de AIC, BIC e melhor comportamento dos resíduos quantílicos (*worm plots*). O modelo **HBS** destacou-se como o mais adequado para todos os quantis analisados, confirmando sua capacidade de capturar a assimetria e a presença de valores extremos típicos de dados econômicos.

2. **Dados Censurados (Crescimento de Cordeiros)**: Na análise do tempo até o atingimento do peso-alvo, ambos os modelos superaram o NBS. Observou-se uma distinção importante no desempenho:

- O modelo **LBS** ajustou-se melhor aos quantis inferiores (cauda inferior da distribuição).
- O modelo **HBS** mostrou-se superior na modelagem da mediana e dos quantis superiores (cauda superior).

Isso reforça a importância de dispor de diferentes distribuições para modelar diferentes comportamentos de cauda.

Considerações Finais

Conclui-se que a incorporação das distribuições LBS e HBS ao arcabouço da regressão quantílica oferece ferramentas analíticas poderosas para pesquisadores. A flexibilidade para modelar diferentes níveis de assimetria e curtose, aliada à interpretabilidade dos parâmetros quantílicos, torna esses modelos alternativas robustas para modelagem de dados de suporte positivo.

REFERÊNCIAS

- ALAM, F. M. A.; ALMALKI, A. M. The hazard rate function of the logistic birnbaum-saunders distribution: Behavior, associated inference, and application. *Journal of King Saud University-Science*, Elsevier, v. 33, n. 7, p. 101580, 2021.
- BASTOS, F. d. S.; BARRETO-SOUZA, W. Birnbaum-saunders sample selection model. *Journal of Applied Statistics*, Taylor & Francis, v. 48, n. 11, p. 1896–1916, 2021.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. Estimation for a family of life distributions with applications to fatigue. *Journal of applied probability*, Cambridge University Press, v. 6, n. 2, p. 328–347, 1969.
- BITTENCOURT, V. L. Parametric quantile regression for extreme events. 2022.
- DASILVA, A. et al. Scale-mixture birnbaum-saunders quantile regression models applied to personal accident insurance data. *Computational and Applied Mathematics*, Springer, v. 44, n. 1, p. 1–48, 2025.
- DÍAZ-GARCÍA, J. A.; LEIVA-SÁNCHEZ, V. A new family of life distributions based on the elliptically contoured distributions. *Journal of Statistical Planning and Inference*, v. 128, n. 2, p. 445–457, 2005.
- DIAZ-GARCIA, J. A.; LEIVA, V. *A new family of life distributions based on Birnbaum-Saunders distribution*. [S.l.], 2002.
- FERREIRA, M.; GOMES, M. I.; LEIVA, V. On an extreme value version of the birnbaum-saunders distribution. *REVSTAT-Statistical Journal*, v. 10, n. 2, p. 181–210, 2012.
- FISCHER, M. J.; HUI, A.; HÖSLE, S. whs-type distributions with application to finance. *Journal of Statistics and Management Systems*, Taylor & Francis, v. 20, n. 1, p. 67–89, 2017.
- IQBAL, A.; MAHMOOD, T. Cusum charts utilizing reparametrized birnbaum-saunders model for fault detection and process control. *Measurement and Control*, SAGE Publications Sage UK: London, England, p. 00202940241282533, 2024.
- KECOJEVIC, T. *Bootstrap inference for parametric quantile regression*. [S.l.]: The University of Manchester (United Kingdom), 2011.
- KOENKER, R.; JR, G. B. *Regression quantiles*. *Econometrica: Journal of the conometric Society* 33–50. 1978.

- KOENKER, R.; MACHADO, J. A. Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, Taylor & Francis, v. 94, n. 448, p. 1296–1310, 1999.
- KORITIAKI, N. A. Parâmetros fenotípicos e genéticos de características de crescimento de cordeiros da raça santa inês do nascimento ao desmame. 2015.
- KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Generalized multivariate birnbaum–saunders distributions and related inferential issues. *Journal of Multivariate Analysis*, Elsevier, v. 116, p. 230–244, 2013.
- LEÃO, J. et al. A survival model with birnbaum–saunders frailty for uncensored and censored cancer data. 2018.
- LEIVA, V. et al. Influence diagnostics in log-birnbaum–saunders regression models with censored data. *Computational Statistics & Data Analysis*, Elsevier, v. 51, n. 12, p. 5694–5707, 2007.
- LEIVA, V. et al. Extreme value birnbaum–saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, Springer, v. 30, p. 1045–1058, 2016.
- MAZUCHELI, J. et al. A new quantile regression for modeling bounded data under a unit birnbaum–saunders distribution with applications in medicine and politics. *Symmetry*, MDPI, v. 13, n. 4, p. 682, 2021.
- MAZUCHELI, J.; MENEZES, A. F.; DEY, S. The unit-birnbaum-saunders distribution with applications. *Chilean Journal of Statistics*, v. 9, n. 1, p. 47–57, 2018.
- SÁNCHEZ, L. et al. Birnbaum-saunders quantile regression models with application to spatial data. *Mathematics*, MDPI, v. 8, n. 6, p. 1000, 2020.
- SANHUEZA, A.; LEIVA, V.; BALAKRISHNAN, N. The generalized birnbaum–saunders distribution and its theory, methodology, and application. *Communications in Statistics—Theory and Methods*, Taylor & Francis, v. 37, n. 5, p. 645–670, 2008.
- SANTOS-NETO, M. et al. A reparameterized birnbaum–saunders distribution and its moments, estimation and applications. *REVSTAT-Statistical Journal*, v. 12, n. 3, p. 247–272, 2014.
- STASINOPOULOS, M. D. et al. *Flexible regression and smoothing: using GAMLSS in R*. [S.l.]: CRC Press, Taylor & Francis Group, 2017.
- VOLODIN, I. N.; DZHUNGUROVA, O. A. On limit distributions emerging in the generalized birnbaum-saunders model. *Journal of Mathematical Sciences*, Springer, v. 99, p. 1348–1366, 2000.
- YEH, I.-C. *Real Estate Valuation*. 2018. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5J30W>.

APÊNDICE A

A.1 Código da Implementação da HBS reparametrizada

```

1 #####
2 # Derivadas g(mu, sigma; tau)
3 Ug.Hg.HBSTAU <- function(mu, sigma, tau = 0.5)
4 {
5     t1 <- pi ^ 2
6     t4 <- tan(tau * pi / 2)
7     t5 <- log(t4)
8     t6 <- t5 ^ 2
9     t7 <- sigma ^ 2
10    t8 <- t7 * t6
11    t10 <- t8 + t1
12    t11 <- sqrt(t10)
13    t12 <- t5 * t11
14    t15 <- 2 * sigma * t12 + t1 + 2 * t8
15    t18 <- t1 * mu
16    t19 <- t15 ^ 2
17    t20 <- 0.1e1 / t19
18    t25 <- t5 * t6 / t11
19    t29 <- 4 * sigma * t6 + 2 * t7 * t25 + 2 * t12
20    t34 <- t29 ^ 2
21    t41 <- t6 ^ 2
22    dgdmu <- 0.1e1 / t15 * t1
23    dgdsigma <- -t29 * t20 * t18
24    dgdmudmu <- 0
25    dgdsigmadsigma <- 2 * t34 / t15 / t19 * t18 - (4 * t6 - 2 * sigma * t7 * t5 * t41
    ↪ / t10 / t11 + 6 * sigma * t25) * t20 * t18
26    dgdmudsigma <- -t29 * t20 * t1
27    data.frame(dgdmu = dgdmu, dgdsigma = dgdsigma, dgdmudmu = dgdmudmu,
    ↪ dgdsigmadsigma = dgdsigmadsigma, dgdmudsigma = dgdmudsigma)
28 }
29 #####
30 # Derivadas l(beta, sigma; tau)
31 Ul.Hl.HBSTAU <- function(x, mu, sigma, tau = 0.5)

```

```

32 {
33     t1 <- pi ^ 2
34     t5 <- tan(tau * pi / 2)
35     t6 <- log(t5)
36     t7 <- t6 ^ 2
37     t8 <- sigma ^ 2
38     t9 <- t8 * t7
39     t12 <- sqrt(t9 + t1)
40     beta <- 0.1e1 / (2 * sigma * t6 * t12 + t1 + 2 * t9) * t1 * mu
41     t1 <- 0.1e1 / beta
42     t2 <- t1 * x
43     t3 <- sqrt(t2)
44     t5 <- 0.1e1 / t3 / t2
45     t6 <- beta ^ 2
46     t8 <- 0.1e1 / t6 / beta
47     t12 <- 0.1e1 / t3
48     t13 <- 0.1e1 / t6
49     t15 <- 0.1e1 / x
50     t16 <- t15 * beta
51     t17 <- sqrt(t16)
52     t19 <- 0.1e1 / t17 / t16
53     t21 <- x ^ 2
54     t23 <- 0.1e1 / t21 / x
55     t26 <- 0.1e1 / t17
56     t27 <- 0.1e1 / t21
57     t29 <- x * t8 * t5 / 2 - t13 * t12 - t23 * beta * t19 / 2 + t27 * t26
58     t33 <- t27 * beta * t26 + t1 * t12
59     t34 <- 0.1e1 / t33
60     t36 <- 0.1e1 / sigma
61     t37 <- t36 * pi
62     t38 <- x * t12
63     t39 <- t13 * t38
64     t42 <- t15 * t26
65     t45 <- t3 - t17
66     t48 <- exp(t45 * t37 / 2)
67     t49 <- t48 ^ 2
68     t50 <- pi * t49
69     t51 <- -t39 - t42
70     t53 <- t49 + 1
71     t54 <- 0.1e1 / t53
72     t58 <- sigma ^ 2
73     t59 <- 0.1e1 / t58
74     t60 <- t59 * pi
75     t72 <- t6 ^ 2
76     t78 <- 0.1e1 / t72
77     t88 <- t21 ^ 2
78     t95 <- t29 ^ 2

```

```

79     t96 <- t33 ^ 2
80     t100 <- t78 * t21 * t5
81     t103 <- t8 * t38
82     t106 <- t27 * t19
83     t109 <- pi ^ 2
84     t110 <- t109 * t49
85     t112 <- t51 ^ 2 * t59 / 4
86     t121 <- t49 ^ 2
87     t122 <- t109 * t121
88     t123 <- t53 ^ 2
89     t124 <- 0.1e1 / t123
90     t129 <- 0.1e1 / t58 / sigma
91     t130 <- t129 * pi
92     t133 <- t58 ^ 2
93     t135 <- t45 ^ 2
94     t136 <- t135 / t133
95     dldbета <- t34 * t29 - t39 * t37 / 4 - t42 * t37 / 4 - t54 * t51 * t36 * t50 / 2
96     dldsigma <- -t36 - t3 * t60 / 2 + t17 * t60 / 2 + t54 * t45 * t59 * t50
97     dldbетаdбета <- t34 * (0.3e1 / 0.4e1 / t72 / бета / t3 / t13 - 2 * x * t78 * t5 +
    ↪ 2 * t8 * t12 + 0.3e1 / 0.4e1 / t88 * бета / t17 / t27 / t6 - t23 * t19) -
    ↪ 0.1e1 / t96 * t95 - t100 * t37 / 8 + t103 * t37 / 2 + t106 * t37 / 8 - t54 *
    ↪ t112 * t110 - t54 * (-t100 / 4 + t103 + t106 / 4) * t36 * t50 + t124 * t112 *
    ↪ t122
98     dldsigmadsigma <- -2 * t54 * t45 * t129 * t50 - t54 * t136 * t110 + t124 * t136 *
    ↪ t122 - t17 * t130 + t3 * t130 + t59
99     dldbетаdsigma <- t39 * t60 / 4 + t42 * t60 / 4 + t51 * t54 * t45 * t129 * t110 / 2
    ↪ + t54 * t51 * t59 * t50 / 2 - t51 * t124 * t45 * t129 * t122 / 2
100    data.frame(dldbета = dldbета, dldsigma = dldsigma, dldbетаdбета = dldbетаdбета,
    ↪ dldsigmadsigma = dldsigmadsigma, dldbетаdsigma = dldbетаdsigma)
101  }
102  #####
103  U.H.mu.sigma <- function(x, mu, sigma, tau = 0.5) {
104      db <- Ug.Hg.HBSTAU(mu, sigma, tau)
105      dL <- Ul.Hl.HBSTAU(x, mu, sigma, tau)
106
107      # Garantir acesso escalar
108      dgdmu <- db$dgdmu
109      dgdsigma <- db$dgdsigma
110      dgdmudmu <- db$dgdmudmu
111      dgdsigmadsigma <- db$dgdsigmadsigma
112      dgdmudsigma <- db$dgdmudsigma
113
114      dldbета <- dL$dldbета
115      dldsigma <- dL$dldsigma
116      dldbетаdбета <- dL$dldbетаdбета
117      dldsigmadsigma <- dL$dldsigmadsigma
118      dldbетаdsigma <- dL$dldbетаdsigma

```

```

119
120   # Gradiente
121   dldmu      <- dldbета * dgdmu
122   dldsigma   <- dldbета * dgdsigma + dldsigma
123   grad       <- data.frame(dldmu = dldmu, dldsigma = dldsigma)
124
125   # Hessiana
126   d2ldmumu   <- dgdmu^2 * dldbетаadbета + dgdmudmu * dldbета
127   d2ldmusigma <- dgdmu * dgdsigma * dldbетаadbета + dgdmu * dldbетаadsigma +
128   ↪ dgdmudsigma * dldbета
129   d2ldsigsig <- dgdsigma^2 * dldbетаadbета + 2 * dgdsigma * dldbетаadsigma
130   ↪ + dgdsigmadsigma * dldbета + dldsigmadsigma
131   hess       <- data.frame(d2ldmumu = d2ldmumu, d2ldsigsig =
132   ↪ d2ldsigsig, d2ldmusigma = d2ldmusigma)
133   cbind(grad, hess)
134 }
135 #####
136 dHBSTAU <- function(x, mu, sigma, tau = 0.5, log = FALSE)
137 {
138   t1 <- pi ^ 2
139   t5 <- tan(tau * pi / 2)
140   t6 <- log(t5)
141   t7 <- t6 ^ 2
142   t8 <- sigma ^ 2
143   t9 <- t8 * t7
144   t12 <- sqrt(t9 + t1)
145   beta <- 0.1e1 / (2 * sigma * t12 * t6 + t1 + 2 * t9) * t1 * mu
146   t1 <- log(sigma)
147   t2 <- log(2)
148   t3 <- 0.1e1 / beta
149   t5 <- sqrt(t3 * x)
150   t10 <- sqrt(beta / x)
151   t13 <- x ^ 2
152   t17 <- log(t3 / t5 + 0.1e1 / t13 * beta / t10)
153   t19 <- 0.1e1 / sigma * pi
154   t27 <- exp((t5 - t10) * t19 / 2)
155   t28 <- t27 ^ 2
156   t30 <- log(t28 + 1)
157   lnf <- -t1 - t2 + t17 + t5 * t19 / 2 - t10 * t19 / 2 - t30
158   if (log) return(lnf) else return(exp(lnf))
159 }
160 #####
161 pHBSTAU <- function(q, mu, sigma, beta, tau = 0.5, lower.tail = TRUE, log.p = FALSE)
162 {
163   t1 <- pi ^ 2
164   t5 <- tan(tau * pi / 2)
165   t6 <- log(t5)

```

```

163   t7 <- t6 ^ 2
164   t8 <- sigma ^ 2
165   t9 <- t8 * t7
166   t12 <- sqrt(t9 + t1)
167   beta <- 0.1e1 / (2 * sigma * t12 * t6 + t1 + 2 * t9) * t1 * mu
168   t6 <- sqrt(q / beta)
169   t9 <- sqrt(beta / q)
170   t13 <- exp((t6 - t9) / sigma * pi / 2)
171   t14 <- atan(t13)
172   cdf <- 2 * t14 / pi
173   if (!lower.tail) cdf <- 1 - cdf
174   if (log.p) cdf <- log(cdf)
175   return(cdf)
176 }
177 #####
178 qHBSTAU <- function(p, mu, sigma, tau = 0.5, lower.tail = TRUE, log.p = FALSE)
179 {
180   t1 <- pi ^ 2
181   t5 <- tan(tau * pi / 2)
182   t6 <- log(t5)
183   t7 <- t6 ^ 2
184   t8 <- sigma ^ 2
185   t9 <- t8 * t7
186   t12 <- sqrt(t9 + t1)
187   beta <- 0.1e1 / (2 * sigma * t12 * t6 + t1 + 2 * t9) * t1 * mu
188   if (!lower.tail) p = 1 - p
189   if (log.p) p = exp(p)
190   t3 <- tan(p * pi / 2)
191   t4 <- log(t3)
192   t5 <- t4 * sigma
193   t6 <- t4 ^ 2
194   t7 <- sigma ^ 2
195   t8 <- t6 * t7
196   t9 <- pi ^ 2
197   t11 <- sqrt(t8 + t9)
198   return(beta * (2 * t5 * (-t5 + t11) + 4 * t8 + t9) / t9)
199 }
200 #####
201 rHBSTAU <- function(n, mu, sigma, tau = 0.5)
202 {
203   qHBSTAU(p = runif(n), mu = mu, sigma = sigma, tau = tau)
204 }
205 #####
206 meanHBSTAU <- function(mu, sigma, tau = 0.5)
207 {
208   #stopifnot(all(mu > 0), all(sigma > 0))
209   ztau <- qlogis(tau)

```

```

210     beta      <- mu / (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
      ↪ ztau^2))^2
211     return(beta * (1 + (pi^2 / 6) * sigma^2))
212 }
213
214 varHBSTAU <- function(mu, sigma, tau = 0.5)
215 {
216     #stopifnot(all(mu > 0), all(sigma > 0))
217     ztau      <- qlogis(tau)
218     beta      <- mu / (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
      ↪ ztau^2))^2
219     return((beta * sigma * pi)^2 / 3 * (1 + (37 / 60) * pi^2 * sigma^2))
220 }
221 #####
222 HBSTAU <- function(mu.link = "log", sigma.link = "log")
223 {
224     mstats <- checklink("mu.link",    "HBSTAU Family", substitute(mu.link),
      ↪ c("log", "identity"))
225     dstats <- checklink("sigma.link", "HBSTAU Family", substitute(sigma.link),
      ↪ c("log", "identity"))
226
227     structure(list(
228         family      = c("HBSTAU", "HBSTAU-Family"),
229         parameters  = list(mu = TRUE, sigma = TRUE),
230         nopar       = 2,
231         type        = "Continuous",
232
233         mu.link     = as.character(substitute(mu.link)),
234         sigma.link  = as.character(substitute(sigma.link)),
235         mu.linkfun   = mstats$linkfun,
236         sigma.linkfun = dstats$linkfun,
237         mu.linkinv   = mstats$linkinv,
238         sigma.linkinv = dstats$linkinv,
239         mu.dr        = mstats$mu.eta,
240         sigma.dr     = dstats$mu.eta,
241
242         # Score com tau capturado
243         dldm         = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
      ↪ tau)$dldmu,
244         dldd         = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
      ↪ tau)$dlldsigma,
245
246         # Hessiana com tau capturado
247         d2ldm2       = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
      ↪ tau)$d2ldmumu,
248         d2ldd2       = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
      ↪ tau)$d2ldsigsig,

```

```

249      d2ldmdd      = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
      ↪ tau)$d2ldmusigma,
250
251      # Log-verossimilhança para critério de informação
252      G.dev.incr    = function(y, mu, sigma, tau, ...) -2 * dHBSTAU(y, mu,
      ↪ sigma, tau, log = TRUE),
253
254      # Resíduos quantílicos
255      rqres         = expression(rqres(pfun = "pHBSTAU", type = "Continuous",
      ↪ y = y, mu = mu, sigma = sigma, tau = tau)),
256
257      mu.initial    = expression({mu <- rep(median(y, na.rm = TRUE),
      ↪ length(y))}),
258      sigma.initial = expression({sigma <- rep(0.1, length(y))}),
259
260      mu.valid      = function(mu) all(mu > 0),
261      sigma.valid   = function(sigma) all(sigma > 0),
262
263      mean          = function(mu, sigma, tau) meanHBSTAU(mu, sigma, tau),
264      variance      = function(mu, sigma, tau) varHBSTAU(mu, sigma, tau),
265      y.valid       = function(y) all(y > 0)), class = c("gamlss.family",
      ↪ "family"))
266 }
267

```

A.2 Código da Implementação da LBS reparametrizada

```

1  #####
2  # Derivadas g(mu, sigma; tau)
3  Ug.Hg.LBSTAU <- function(mu, sigma, tau = 0.5) {
4      t4      <- qlogis(tau)
5      t7      <- sigma ^ 2
6      t8      <- t4 ^ 2
7      t11 <- 1 + 0.25e0 * t8 * t7
8      t12 <- sqrt(t11)
9      t13 <- 0.5e0 * t4 * sigma + t12
10     t14 <- t13 ^ 2
11     t17 <- 0.1e1 / t14 / t13
12     t18 <- t17 * mu
13     t20 <- 0.1e1 / t12
14     t24 <- 0.5e0 * t4 + 0.2500000000e0 * t8 * sigma * t20
15     t27 <- t14 ^ 2
16     t30 <- t24 ^ 2
17     t36 <- t8 ^ 2
18     dgdmu <- 0.1e1 / t14
19     dgdsigma <- -2 * t24 * t18
20     dgdmudmu <- 0
21     dgdsigmadsigma <- 6 * t30 / t27 * mu - 2 * (-0.6250000000e-1 * t36 * t7
      ↪ / t12 / t11 + 0.2500000000e0 * t8 * t20) * t18
22     dgdmudsigma <- -2 * t24 * t17

```

```

23     data.frame(dgdmu = dgdmu, dgdsigma = dgdsigma, dgdmudmu = dgdmudmu,
24               ↪ dgdsigmadsigma = dgdsigmadsigma, dgdmudsigma = dgdmudsigma)
25 }
26 #####
27 # Derivadas l(beta, sigma; tau)
28 UL.H1.LBSTAU <- function(x, mu, sigma, tau = 0.5) {
29     ztau      <- qlogis(tau)
30     beta      <- mu / (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 * ztau^2))^2
31     t1 <- 0.1e1 / beta
32     t3 <- beta + x
33     t5 <- 0.1e1 / sigma
34     t6 <- beta * x
35     t7 <- sqrt(t6)
36     t9 <- 0.1e1 / t7 / t6
37     t10 <- t9 * t5
38     t11 <- x ^ 2
39     t16 <- 0.1e1 / t7
40     t17 <- t16 * t5
41     t18 <- -x + beta
42     t19 <- t5 * t18
43     t23 <- t17 - x * t9 * t19 / 2
44     t25 <- exp(t16 * t19)
45     t27 <- 1 + t25
46     t28 <- 0.1e1 / t27
47     t29 <- t28 * t25 * t23
48     t32 <- sigma ^ 2
49     t33 <- 0.1e1 / t32
50     t34 <- t16 * t33
51     t37 <- t33 * t18
52     t39 <- t28 * t25 * t16
53     t43 <- beta ^ 2
54     t46 <- t3 ^ 2
55     t50 <- 0.1e1 / t7 / t11 / t43
56     t51 <- t50 * t5
57     t58 <- x * t10
58     t66 <- t23 ^ 2
59     t70 <- t25 ^ 2
60     t72 <- t27 ^ 2
61     t73 <- 0.1e1 / t72
62     t78 <- 0.1e1 / t32 / sigma
63     t79 <- t16 * t78
64     t87 <- t18 ^ 2
65     t88 <- t32 ^ 2
66     t91 <- t1 / t88 * t87
67     t92 <- 0.1e1 / x
68     t102 <- t9 * t33
69     t107 <- t28 * t25
70     t113 <- t16 * t37
71     dldbета <- -t1 / 2 + 0.1e1 / t3 + t11 * t10 / 2 - t6 * t10 / 2 + t17 - 2 * t29
72     dlδsigma <- -beta * t34 + x * t34 + 2 * t39 * t37 - t5
73     dlδbetadbета <- 0.1e1 / t43 / 2 - 0.1e1 / t46 - 0.3e1 / 0.4e1 * t11 * x * t51 +
74     ↪ 0.3e1 / 0.4e1 * t11 * beta * t51 - t58 - 2 * t28 * t25 * (-t58 + 0.3e1 / 0.4e1 *
75     ↪ t11 * t50 * t19) - 2 * t28 * t25 * t66 + 2 * t73 * t70 * t66
74     dlδsigmadsigma <- -2 * t28 * t25 * t92 * t91 + 2 * t73 * t70 * t92 * t91 - 4
75     ↪ * t39 * t78 * t18 + 2 * beta * t79 - 2 * x * t79 + t33
74     dlδbetadσsigma <- -t11 * t102 / 2 + t6 * t102 / 2 - t34 + 2 * t107 * t34 - x * t107 *
75     ↪ t9 * t37 + 2 * t29 * t113 - 2 * t23 * t73 * t70 * t113

```

```

76   data.frame(dldbета = dldbета, dldsigma = dldsigma, dldbетаdбета = dldbетаdбета,
77             ↪ dldsigmadsigma = dldsigmadsigma, dldbетаdsigma = dldbетаdsigma)
78 }
79 #####
80 U.H.mu.sigma <- function(x, mu, sigma, tau = 0.5) {
81     db <- Ug.Hg.LBSTAU(mu, sigma, tau)
82     dL <- Ul.Hl.LBSTAU(x, mu, sigma, tau)
83
84     # Garantir acesso escalar
85     dgdmu <- db$dgdmu
86     dgdsigma <- db$dgdsigma
87     dgdmudmu <- db$dgdmudmu
88     dgdsigmadsigma <- db$dgdsigmadsigma
89     dgdmudsigma <- db$dgdmudsigma
90
91     dldbета <- dL$dldbета
92     dldsigma <- dL$dldsigma
93     dldbетаdбета <- dL$dldbетаdбета
94     dldsigmadsigma <- dL$dldsigmadsigma
95     dldbетаdsigma <- dL$dldbетаdsigma
96
97     # Gradiente
98     dldmu <- dldbета * dgdmu
99     dldsigma <- dldbета * dgdsigma + dldsigma
100     grad <- data.frame(dldmu = dldmu, dldsigma = dldsigma)
101
102     # Hessiana
103     d2ldmumu <- dgdmu^2 * dldbетаdбета + dgdmudmu * dldbета
104     d2ldmusigma <- dgdmu * dgdsigma * dldbетаdбета + dgdmu * dldbетаdsigma +
105     ↪ dgdmudsigma * dldbета
106     d2ldsigsig <- dgdsigma^2 * dldbетаdбета + 2 * dgdsigma * dldbетаdsigma
107     ↪ + dgdsigmadsigma * dldbета + dldsigmadsigma
108     hess <- data.frame(d2ldmumu = d2ldmumu, d2ldsigsig = d2ldsigsig, d2ldmusigma =
109     ↪ d2ldmusigma)
110     cbind(grad, hess)
111 }
112 #####
113 dLBSTAU <- function(x, mu, sigma, tau = 0.5, log = FALSE)
114 {
115     ztau <- qlogis(tau)
116     beta <- mu / ((0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
117     ↪ ztau^2))^2)
118     z <- (sqrt(x / beta) - sqrt(beta / x)) / sigma
119     Fx <- plogis(z)
120     dens <- (x^(-3/2) * (x + beta)) / (2 * sigma * sqrt(beta)) * Fx * (1 -
121     ↪ Fx)
122     if (log) return(log(dens)) else return(dens)
123 }
124 #####
125 pLBSTAU <- function(q, mu, sigma, tau = 0.5, lower.tail = TRUE, log.p = FALSE)
126 {
127     ztau <- qlogis(tau)
128     beta <- mu / ( (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
129     ↪ ztau^2))^2 )
130     z <- (sqrt(q / beta) - sqrt(beta / q)) / sigma
131     plogis(z, lower.tail = lower.tail, log.p = log.p)
132 }
133 #####
134 qLBSTAU <- function(p, mu, sigma, tau = 0.5, lower.tail = TRUE, log.p = FALSE)

```

```

128 {
129     ztau      <- qlogis(tau)
130     beta      <- mu / ( (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
    ↪ ztau^2))^2 )
131     z         <- qlogis(p, lower.tail = lower.tail, log.p = log.p)
132     Q         <- ((sigma * z / 2) + sqrt(1 + (sigma^2 * z^2) / 4))^2 *
    ↪ beta
133     return(Q)
134 }
135 #####
136 rLBSTAU <- function(n, mu, sigma, tau = 0.5)
137 {
138     qLBSTAU(p = runif(n), mu = mu, sigma = sigma, tau = tau)
139 }
140 #####
141 meanLBSTAU <- function(mu, sigma, tau = 0.5)
142 {
143     #stopifnot(all(mu > 0), all(sigma > 0))
144     ztau      <- qlogis(tau)
145     beta      <- mu / (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
    ↪ ztau^2))^2
146     return(beta * (1 + (pi^2 / 6) * sigma^2))
147 }
148
149 varLBSTAU <- function(mu, sigma, tau = 0.5)
150 {
151     #stopifnot(all(mu > 0), all(sigma > 0))
152     ztau      <- qlogis(tau)
153     beta      <- mu / (0.5 * sigma * ztau + sqrt(1 + 0.25 * sigma^2 *
    ↪ ztau^2))^2
154     return((beta * sigma * pi)^2 / 3 * (1 + (37 / 60) * pi^2 * sigma^2))
155 }
156 #####
157 LBSTAU <- function(mu.link = "log", sigma.link = "log")
158 {
159     mstats <- checklink("mu.link", "LBSTAU Family", substitute(mu.link),
    ↪ c("log", "identity"))
160     dstats <- checklink("sigma.link", "LBSTAU Family", substitute(sigma.link),
    ↪ c("log", "identity"))
161
162     structure(list(
163         family      = c("LBSTAU", "LBSTAU-Family"),
164         parameters  = list(mu = TRUE, sigma = TRUE),
165         nopar       = 2,
166         type        = "Continuous",
167
168         mu.link     = as.character(substitute(mu.link)),
169         sigma.link  = as.character(substitute(sigma.link)),
170         mu.linkfun  = mstats$linkfun,
171         sigma.linkfun = dstats$linkfun,
172         mu.linkinv  = mstats$linkinv,
173         sigma.linkinv = dstats$linkinv,
174         mu.dr       = mstats$mu.eta,
175         sigma.dr    = dstats$mu.eta,
176
177         # Score com tau capturado
178         dldm        = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
    ↪ tau)$dldmu,

```

```

179      dlldd      = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
180      ↪ tau)$dlldsigma,
181
182      # Hessiana com tau capturado
183      d2ldm2     = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
184      ↪ tau)$d2ldmumu,
185      d2ldd2     = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
186      ↪ tau)$d2ldsigsig,
187      d2ldmdd    = function(y, mu, sigma, tau) U.H.mu.sigma(y, mu, sigma,
188      ↪ tau)$d2ldmusigma,
189
190      # Log-verossimilhança para critério de informação
191      G.dev.incr  = function(y, mu, sigma, tau, ...) -2 * dLBSTAU(y, mu,
192      ↪ sigma, tau, log = TRUE),
193
194      # Resíduos quantílicos
195      rqres       = expression(rqres(pfun = "pLBSTAU", type = "Continuous",
196      ↪ y = y, mu = mu, sigma = sigma, tau = tau)),
197
198      mu.initial  = expression({mu <- rep(median(y, na.rm = TRUE),
199      ↪ length(y))}),
200      sigma.initial = expression({sigma <- rep(0.1, length(y))}),
201
202      mu.valid    = function(mu) all(mu > 0),
203      sigma.valid  = function(sigma) all(sigma > 0),
204
205      mean        = function(mu, sigma, tau) meanLBSTAU(mu, sigma, tau),
206      variance     = function(mu, sigma, tau) varLBSTAU(mu, sigma, tau),
207      y.valid      = function(y) all(y > 0)), class = c("gamlss.family",
208      ↪ "family"))
209 }
210 #####

```