

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Análise de Diagnóstico para o Modelo de Regressão Log-Birnbaum-Saunders Generalizado

por

Débora Karollyne Xavier Silva[†]

sob orientação da

Profa. Dra. Michelli Karinne Barros da Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

[†]Este trabalho contou com apoio financeiro da CAPES

Análise de Diagnóstico para o Modelo de Regressão Log-Birnbaum-Saunders Generalizado

por

Débora Karollyne Xavier Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Probabilidade e Estatística

Aprovada por:

Prof. Dr. Víctor Leiva- Universidad de Valparaíso- Chile

Profa. Dra. Amanda dos Santos Gomes- UFCG

Profa. Dra. Michelli Karinne Barros da Silva- UFCG
Orientadora

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Dezembro/2013

Resumo

A distribuição Birnbaum-Saunders surgiu em 1969 com aplicações fortemente ligadas à engenharia e se expandiu nas últimas décadas a diversas áreas. Na literatura, além de tomar um papel de destaque na análise de sobrevivência, podemos destacar o surgimento de várias generalizações. Neste trabalho apresentaremos uma dessas generalizações, a qual foi formulada por Mentainis em 2010. Primeiramente, faremos uma breve explanação sobre a distribuição Birnbaum-Saunders clássica e sobre a generalização que foi proposta por Mentainis (2010), a qual chamaremos de distribuição Birnbaum-Saunders generalizada. Em seguida, discorreremos sobre a distribuição senh-normal, a qual possui uma importante relação com a distribuição Birnbaum-Saunders. Numa outra etapa, apresentaremos alguns métodos de diagnóstico para o modelo de regressão log-Birnbaum-Saunders generalizado e investigaremos testes de homogeneidade para os correspondentes parâmetros de forma e escala. Por fim, analisamos um conjunto de dados para ilustrar a teoria desenvolvida.

Palavras-chave: Distribuição Birnbaum-Saunders, distribuição senh-normal, diagnóstico de influência, teste para homogeneidade.

Abstract

The Birnbaum-Saunders distribution emerged in 1969 motivated by problems in engineering. However, its field of application has been extended beyond the original context of material fatigue and reliability analysis. In the literature, it has made an important role in survival analysis. Moreover, many generalizations of it have been considered. In this work we present one of these generalizations, which was formulated by Mentainis in 2010. First, we provide a brief explanation of the classical Birnbaum-Saunders distribution and its generalization proposed by Mentainis (2010), which we name as the generalized Birnbaum-Saunders distribution. Thereafter, we discuss the sinh-normal distribution, which has an important relationship with the Birnbaum-Saunders distribution. In a further part of this work, we present some diagnostic methods for generalized log-Birnbaum-Saunders regression models and investigate tests of homogeneity for the corresponding shape and scale parameters. Finally, an application with real data is presented.

Key-words: Birnbaum-Saunders distribution, influence diagnostics, sinh-normal distribution, test of homogeneity.

Agradecimentos

Agradeço ao meu Deus que mais uma vez permitiu que eu concluísse uma etapa de minha vida com fé, saúde e esperança, cercada por pessoas maravilhosas, anjos sem os quais tudo seria mais difícil.

Aos meus pais, incentivadores essenciais em todos os momentos. Meu pai Valdir, pelos conselhos, encorajamento, confiança, pelo exemplo de ética e honestidade. Minha mãe Ivonete, solícita em todos os momentos, pela imensa força, sem sua ajuda eu não teria nem começado o que hoje eu consegui concluir.

Ao meu filho Daniel, por sempre me lembrar que há coisas muito mais difíceis do que terminar uma pós-graduação.

Aos meus irmãos e a minha família em geral por todo apoio e incentivo que me deram, em especial a minha avó Tezinha e minha tia Luzinete, das quais a ajuda foi inestimável nesse último ano.

Aos meus amigos da graduação e do mestrado, que também foram anjos, hoje são amigos de vida e futuramente serão velhos amigos. Muitos marcaram minha vida com pequenos gestos, como um abraço, uma dúvida tirada, companhia numa noite de estudos em casa ou na biblioteca, cafés da tarde, ensaios pré-defesa, conversas, choros e boas risadas.

Aos professores da UAMat e da UAEst que participaram da minha formação acadêmica, inclusive aqueles dos quais não tive oportunidade de ser aluna, mas que estavam sempre dispostos a ajudar ou dar uma palavra de ânimo e estímulo. Aos funcionários, que não foram menos importantes, fazendo do local de estudos e trabalho um ambiente acolhedor.

A professora Michelli, que não recebeu o título de mãe acadêmica desmerecidamente. Este trabalho só se tornou possível devido à sua imensa ajuda e a sua arte de dosar suavidade e rigidez maternal sempre que necessário.

Ao Professor Victor Leiva e a professora Amanda Gomes, pela disponibilidade de participar da banca, pelos questionamentos, correções e sugestões, que contribuíram grandemente para o enriquecimento deste trabalho.

A CAPES pelo apoio financeiro.

Por fim, não tornarei este texto enfadonho citando os muitos nomes que contribuíram para a realização deste trabalho, apenas deixo meus mais sinceros agradecimentos a todos que de perto ou de longe se permitiram estar comigo nessa caminhada. Muito obrigada!

Dedicatória

Aos meus pais, que me mostraram a importância da independência e do conhecimento.

Ao meu filho Daniel, que me mostrou o sentido da vida.

Sumário

1	Preliminares	1
1.1	Introdução	1
1.2	Distribuição Birnbaum-Saunders	3
1.2.1	Origem	3
1.2.2	Propriedades	7
1.2.3	Organização da dissertação	9
1.2.4	Plataforma computacional	9
2	Distribuição Birnbaum-Saunders generalizada	10
2.1	Propriedades	12
2.2	Estimação dos parâmetros do modelo	16
2.3	Estudo de simulação	17
3	A distribuição senh-normal	28
3.1	Introdução	28
3.2	Propriedades da distribuição senh-normal	31
3.3	Momentos	39
4	Um modelo linear para a distribuição log-Birnbaum-Saunders generalizada	45
4.1	Introdução	45
4.1.1	O modelo de regressão log-linear BSG	45
4.2	Análise de diagnóstico	47
4.2.1	Diagnósticos de influência baseados no CDM	47
4.2.2	Distância de Cook generalizada	49
4.2.3	Afastamento pela verossimilhança	50
4.2.4	MSOM e teste para <i>outlier</i>	50
4.2.5	Teste de heterogeneidade dos parâmetros	51
4.2.6	Aplicação	54

Apêndice B	62
Apêndice C	67
Apêndice D	71
Apêndice E	73
Apêndice F	75
Referências	77

Lista de Figuras

1.1	gráficos de funções densidade da distribuição BS com parâmetros $\beta = 1$ (a) e $\beta = 2$ (b).	6
1.2	gráficos de funções de distribuição acumulada (a) e de funções de sobrevivência da distribuição BS (b), ambos com parâmetro $\beta = 1$	7
1.3	gráficos das funções risco da distribuição BS para os valores de α indicados com $\beta=1$ (a) e $\beta = 2$ (b).	8
2.1	gráficos de funções densidade da distribuição BSG com parâmetros $\alpha = 0,5$ e $\beta = 1$ (a), e $\alpha = 0,5$ e $\beta = 2$ (b).	11
2.2	gráficos de funções de distribuição acumulada (a) e gráficos de funções de sobrevivência (b) da distribuição BSG, ambos com parâmetro $\beta = 2$	11
2.3	gráfico das funções risco da distribuição BSG para os valores de $\alpha = 2$ indicados e $\beta = 2$	12
2.4	gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 0,2$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).	22
2.5	gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 0,5$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).	23
2.6	gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 1$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).	24
3.1	gráficos de funções densidade da distribuição SN não centrada com parâmetros $\gamma = 0$, $\sigma = 1$ e $\alpha = 1,5$ (a), e $\gamma = 0$ e $\sigma = 1$ e $\alpha = 4$ (b).	30
3.2	gráficos de funções densidade da distribuição SN centrada com parâmetros $\gamma = 0$ e $\sigma = 1$ (a), e $\gamma = 0$ e $\sigma = 1,5$ (b).	31
3.3	gráficos de funções densidade da distribuição SN centrada com parâmetros $\gamma = 2$ e $\sigma = 1$ (a), e $\gamma = 2$ e $\sigma = 1,5$ (b).	31

3.4	gráficos de funções de distribuição acumulada da distribuição SN centrada (a) e gráficos de funções de sobrevivência da distribuição senh-normal central (b), ambos com parâmetros $\gamma = 0$ e $\sigma = 2$	32
4.1	gráfico dos resíduos \hat{R}_i contra os valores ajustados (a) e gráfico normal de probabilidade dos resíduos \hat{R}_i com envelope gerado (b).	55
4.2	gráficos de GCD-a, GCD(β)-b, GCD(α)-c e GCD(λ)-d contra os índices das observações.	57
4.3	gráfico de SC_i contra os índices das observações.	58

Lista de Tabelas

2.1	estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 0,2$ e $\beta = 1$	19
2.2	estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 0,5$ e $\beta = 1$	20
2.3	estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 1$ e $\beta = 1$	21
2.4	limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 0,2$ e $\beta = 1$	25
2.5	limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 0,5$ e $\beta = 1$	26
2.6	limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 1,0$ e $\beta = 1,0$	27
4.1	diagnóstico de influência baseado no método de deleção de casos	56
4.2	estimativas de máxima verossimilhança dos parâmetros com o caso indicado eliminado	57
4.3	valores de LR baseados na exclusão dos casos indicados	58

Capítulo 1

Preliminares

1.1 Introdução

A distribuição Birnbaum-Saunders (BS) é uma distribuição de vida (*fatigue life distribution*) que vem recebendo muita atenção nas últimas décadas e tem se tornado um modelo muito importante nas áreas de sobrevivência e confiabilidade, sendo uma boa alternativa a distribuições clássicas. Birnbaum e Saunders (1969a) motivados por problemas de fadiga provocados por vibrações em aviões comerciais, desenvolveram a distribuição BS, a qual relaciona o tempo até a ocorrência da falha com algum dano cumulativo que é assumido gaussiano. A distribuição BS possui dois parâmetros, um de forma e um de escala, é unimodal e tem assimetria positiva. Esse destaque que vem recebendo a distribuição BS nos últimos anos é devido aos seus argumentos físicos, suas propriedades atrativas e sua relação com a distribuição normal. Depois do trabalho de Birnbaum e Saunders (1969a) muitos avanços teóricos foram realizados com relação à inferência, simulação e generalizações da distribuição BS. Aplicações em diversas áreas, tais como: engenharia, medicina, meio ambiente, qualidade da água, seguros, controle de qualidade, entre outras, têm sido apresentadas na literatura; para mais detalhes ver Meeker e Escobar (1998) Leiva et al. (2007, 2008, 2009, 2010, 2011, 2012), Barros et al. (2008), Podlaski (2008), Leiva e Saunders (2009), Vilca et al. (2010, 2011), Azevedo et al. (2012), Ferreira et al.(2012), Gomes et al. (2012), Marchant et al. (2013) e Saulo et al.(2013). A seguir detalhamos, um pouco, alguns dos principais trabalhos envolvendo essa distribuição.

Inicialmente, Birnbaum e Saunders (1969b) obtiveram os estimadores de máxima verossimilhança da distribuição BS. Engelhardt et al. (1981) propuseram intervalos de confiança e testes de hipóteses para os parâmetros da distribuição considerando um deles como parâmetro de perturbação desconhecido. Desmond (1985) apresentou uma derivação da distribuição BS usando o modelo biológico de Cramér. Um ano mais tarde, Desmond (1986) estabeleceu uma

conexão entre a distribuição Gaussiana inversa e a distribuição BS. Achcar (1993) desenvolveu procedimentos de estimação Bayesiana para os parâmetros da distribuição BS. Chang e Tang (1994) apresentaram um gerador de números aleatórios para a distribuição BS. Lu e Chang (1997) utilizaram métodos bootstrap para construir intervalos de predição. Dupuis e Mills (1998) propuseram métodos robustos para a estimação dos parâmetros da distribuição BS. Mais tarde, Volodin e Dzhungurova (2000) desenvolveram uma família geral de distribuições de vida denominada distribuição de rachadura, que inclui a distribuição Birnbaum-Saunders como um caso particular. Além disso, Volodin (2002) desenvolveu uma reparametrização da Distribuição Birnbaum-Saunders, que levou a uma interpretação interessante dos novos parâmetros dentro do contexto de fadiga de material. Ng et al. (2003) apresentaram o método dos momentos modificados para estimação dos parâmetros da distribuição BS. Lemonte et al. (2007) fizeram correções de viés para os estimadores de máxima verossimilhança, bem como apresentaram testes da razão de verossimilhanças aperfeiçoados. Guiraud, Leiva, e Fierro (2009) propuseram uma versão não-central da Distribuição Birnbaum-Saunders, que é útil para situações de modelagem em que a extensão da rachadura tem uma média não constante ao longo do tempo. Leiva et al. (2012) utilizaram a distribuição BS para modelar diâmetro e mortalidade de árvores. Santos-Neto et al. (2012) propuseram um estudo de diferentes parametrizações para a distribuição BS. Na parte de modelagem, Rieck e Nedelman (1991), propuseram um modelo de regressão log-linear para a distribuição BS. Mais tarde, Tsionas (2001) estudou modelos de regressão BS sob o método Bayesiano. Galea et al. (2004) apresentaram aspectos relacionados a diagnóstico de influência em modelos log-linear BS com dados não censurados. Xie e Wei (2007) consideraram diagnóstico de influência para o modelo de regressão log-linear BS, baseado no método de deleção de casos. Em Leiva et al. (2007), análise de resíduos e diagnóstico de influência para esse modelo, considerando observações censuradas, foram considerados. Procedimentos de diagnóstico para modelos de regressão BS não lineares são propostos por Lemonte e Patriota (2011) e Vanegas et al. (2012). Em Leiva et al. (2014) é proposto um modelo de regressão BS reparametrizado, no qual se modela a média da variável resposta. Algumas generalizações e extensões para o modelo BS foram feitas. Entre elas, podemos citar Owen e Padgett (1999), que desenvolveram a distribuição BS com três parâmetros. Mais tarde, Díaz-García e Leiva (2005, 2006) propuseram uma nova classe de distribuições de vida generalizando a distribuição BS a partir de distribuições de contornos elípticos, gerando a distribuição BS generalizada (BSG). Vilca-Labra e Leiva (2006), baseados em argumentos semelhantes aos empregados por Díaz-García e Leiva (2005), obtiveram uma maior generalização ao desenvolver a distribuição BS mediante distribuições assimétricas. Assim, a distribuição BS duplamente generalizada (BSG''), como eles denominaram, é mais

flexível na curtose e assimetria do que todas as versões anteriores da distribuição BS. Owen (2006) derivou uma nova forma da distribuição BSG relaxando a suposição de independência imposta por Birnbaum e Saunders (1969a) na derivação clássica. Balakrishnan et al. (2007) desenvolveram planos de amostragem de aceitação para dados truncados, os quais eram assumidos seguir uma distribuição BSG. Barros et al. (2008) apresentaram aspectos de robustez e diagnóstico para uma nova classe de modelos regressão com erros com caudas pesadas. Meintanis (2010) apresentou testes de bondade de ajuste para a distribuição BS e introduziu uma nova generalização da distribuição BS, incorporando um novo parâmetro. Vilca et al. (2010) caracterizaram uma versão estendida da distribuição BS e aplicaram-na no estudo da qualidade ambiental em Santiago, Chile. Ahmed et al. (2010) apresentaram uma versão truncada da distribuição BS aplicada ao risco financeiro. Leiva et al. (2011) discutiram sobre uma nova versão da distribuição BS com um parâmetro de locação desconhecido e aplicaram o estudo na modelagem do fluxo de energia eólica. Em Saulo et al. (2013) foi apresentado um método não paramétrico para estimar densidades assimétricas com base em distribuições BS assimétricas, com aplicação a dados ambientais. Em Marchant et al. (2013) foi utilizado o método de *kernel* baseado em distribuições BSG para estimar a densidade com o suporte não negativo. Recentemente, Fierro et al. (2013) propuseram uma nova versão da distribuição BS, contrapondo o fato de que a distribuição BS clássica baseia-se na normalidade assintótica de uma soma de variáveis aleatórias quando um processo de Poisson homogêneo é considerado. Os autores desenvolveram uma nova versão desta distribuição assumindo que o número de termos dessa soma depende de um processo de Poisson não-homogêneo.

1.2 Distribuição Birnbaum-Saunders

1.2.1 Origem

A fadiga é um fenômeno físico que ocorre quando um material é exposto a situações de estresse e tensão, provocando danos ou rachaduras em sua estrutura, as quais crescem com o passar do tempo e à medida que os ciclos de força se repetem. A falha do material ocorre quando uma rachadura dominante alcança um valor crítico, causando sua ruptura. Modelos estatísticos possibilitam descrever a aleatoriedade dos tempos de falha, bem como do número de ciclos necessários até a ruptura dos materiais, quando expostos a diferentes padrões de forças cíclicas. Birnbaum e Saunders (1969a) consideraram um material sujeito a um padrão cíclico de tensão e força e derivaram uma nova distribuição com o objetivo de idealizar o número de ciclos necessários até a falha.

Aqui apresentaremos o processo de construção da distribuição BS, que pode ser visto em

maiores detalhes em Birnbaum e Saunders (1969a). Consideremos as seguintes suposições:

1. Um material é sujeito a uma sequência de m ciclos de tensão e força;
2. A falha do material ocorre devido ao desenvolvimento e ao crescimento de uma fissura dominante dentro do material, ou seja, quando o tamanho da fissura excede certo nível de resistência, denotado por ω ;
3. A extensão incremental da fissura X_i resultante da aplicação da i -ésima oscilação de carga é uma variável aleatória com uma distribuição que só depende da fissura atual causada pela tensão neste ciclo;
4. A extensão da fissura durante o $(j + 1)$ -ésimo ciclo é

$$Y_{j+1} = X_{jm+1} + \cdots + X_{jm+m}, \text{ para } j = 0, 1, \dots,$$

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

5. As extensões das fissuras em diferentes ciclos são independentes;
6. A extensão total da fissura, Y_j , devido ao j -ésimo ciclo é uma variável aleatória que segue uma distribuição com média μ e variância σ^2 , $\forall j = 1, 2, \dots$;
7. Seja $W_k = \sum_{j=1}^k Y_j$ a extensão total da fissura após k ciclos;
8. Seja N o número de ciclos até a falha, a qual ocorre quando o comprimento da fissura excede o valor crítico w .

Considerando que os Y_j 's são variáveis aleatórias independentes e identicamente distribuídas, com média μ e variância σ^2 , $\forall j = 1, 2, \dots$, a função de distribuição acumulada (fda) da variável aleatória N é obtida por:

$$\begin{aligned}
P(N \leq n) &= P\left(\sum_{j=1}^n Y_j > w\right) = 1 - P\left(\sum_{j=1}^n Y_j \leq w\right) \\
&= 1 - P\left(\frac{\sum_{j=1}^n Y_j - E(\sum_{j=1}^n Y_j)}{\sqrt{\text{Var}(\sum_{j=1}^n Y_j)}} \leq \frac{w - E(\sum_{j=1}^n Y_j)}{\sqrt{\text{Var}(\sum_{j=1}^n Y_j)}}\right) \\
&= 1 - P\left(\frac{\sum_{j=1}^n Y_j - \sum_{j=1}^n E(Y_j)}{\sqrt{\sum_{j=1}^n \text{Var}(Y_j)}} \leq \frac{w - \sum_{j=1}^n E(Y_j)}{\sqrt{\sum_{j=1}^n \text{Var}(Y_j)}}\right) \\
&= 1 - P\left(\frac{\sum_{j=1}^n Y_j - n\mu}{\sqrt{n\sigma^2}} \leq \frac{w - n\mu}{\sqrt{n\sigma^2}}\right) \\
&= 1 - P\left(\frac{\sum_{j=1}^n Y_j - n\mu}{\sqrt{n}\sigma} \leq \frac{w}{\sqrt{n}\sigma} - \frac{\sqrt{n}\mu}{\sigma}\right)
\end{aligned}$$

Pelo teorema do limite central, temos:

$$P(N \leq n) \cong 1 - \Phi\left(\frac{w}{\sqrt{n}\sigma} - \frac{\sqrt{n}\mu}{\sigma}\right) = \Phi\left[-\left(\frac{w}{\sqrt{n}\sigma} - \frac{\sqrt{n}\mu}{\sigma}\right)\right] = \Phi\left(\frac{\sqrt{n}\mu}{\sigma} - \frac{w}{\sqrt{n}\sigma}\right), \quad (1.1)$$

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

Birnbaum e Saunders (1969a) usaram a equação (1.1) para definir uma distribuição contínua de vida. Substituindo n por t obtiveram uma variável aleatória T que é a extensão contínua da variável aleatória discreta N . Desta forma, T representa o tempo total até que ocorra falha e temos que a fda de T é dada por:

$$F_T(t; \alpha, \beta) = P(T \leq t) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], t > 0, \quad (1.2)$$

onde $\alpha = \sigma/\sqrt{w\mu} > 0$ e $\beta = w/\mu > 0$.

Daí, dizemos que T segue uma distribuição BS e denotamos por $T \sim \text{BS}(\alpha, \beta)$.

A função densidade de probabilidade (fdp) de T é obtida derivando a fda com relação à t , ou seja, $f_T(t; \alpha, \beta) = F'_T(t; \alpha, \beta)$. Assim, temos

$$f_T(t; \alpha, \beta) = \frac{1}{2\alpha\sqrt{2\pi}\beta} \left[\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}} \right] \exp\left[-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right], t > 0, \alpha, \beta > 0,$$

em que α é um parâmetro de forma e β é um parâmetro de escala.

A Figura (1.1) nos mostra gráficos de densidades da distribuição BS para vários valores de α . Podemos observar que o parâmetro α altera a forma da distribuição, aumentando a assimetria da fdp à medida que cresce o seu valor. Comparando as Figuras 1.1(a) e 1.1(b), observamos que o parâmetro β altera a escala da distribuição.

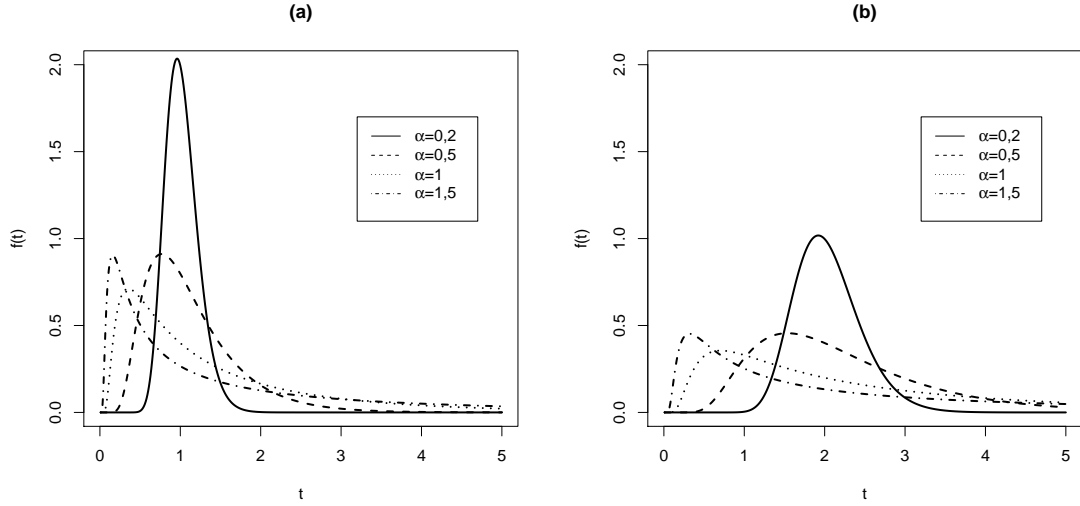


Figura 1.1: gráficos de funções densidade da distribuição BS com parâmetros $\beta = 1$ (a) e $\beta = 2$ (b).

A função de sobrevivência, definida como a probabilidade de uma observação não falhar até um certo tempo t , é expressa por $S(t) = P(T > t)$. Assim, $F_T(t) = 1 - S(t)$. Podemos visualizar tal relação através da Figura (1.2).

Outra importante função usada para caracterizar modelos de sobrevivência é a função de risco, também conhecida como taxa de falha. Conforme Colosimo e Giolo (2006), a função de risco especifica a taxa instantânea de falha de uma observação num determinado tempo t . Embora, a função de risco não seja obrigatoriamente monótona, sendo crescente (respectivamente, decrescente) indica que a taxa de falha aumenta (respectivamente, diminui) com o transcorrer do tempo e se for constante indica que a taxa de falha não se altera com o passar do tempo. No caso da distribuição BS, a função de risco é dada por

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

Através da Figura (1.3), podemos observar o comportamento da função de risco para vários valores de α , a qual não possui um comportamento monótono e tende a se estabilizar em um certo instante t .

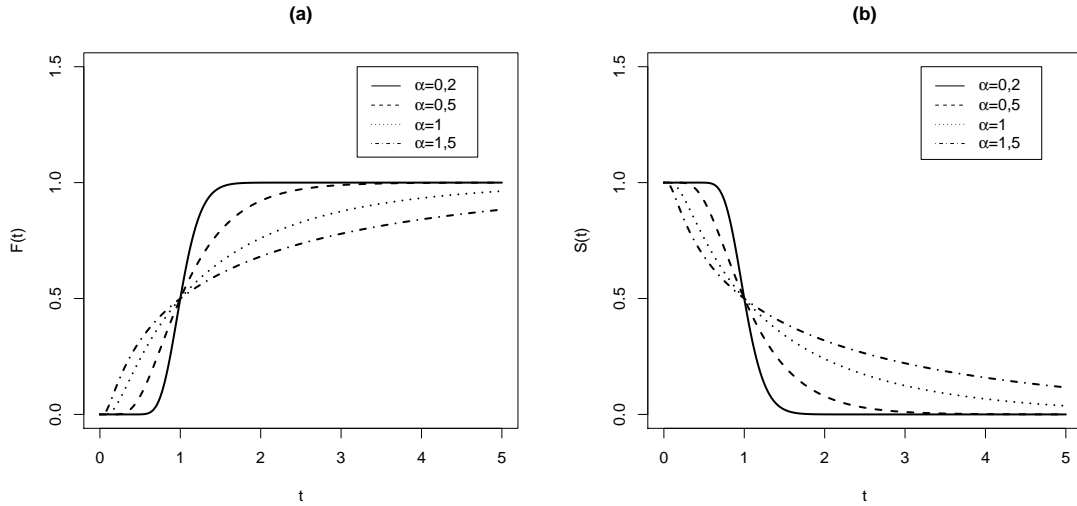


Figura 1.2: gráficos de funções de distribuição acumulada (a) e de funções de sobrevivência da distribuição BS (b), ambos com parâmetro $\beta = 1$.

1.2.2 Propriedades

A esperança e a variância da variável aleatória T , obtidas em Birnbaum e Saunders (1969b), são

$$E(T) = \beta \left(1 + \frac{1}{2}\alpha^2 \right) \text{ e } \text{Var}(T) = (\alpha\beta)^2 \left(1 + \frac{5}{4}\alpha^2 \right),$$

respectivamente. As propriedades acima são obtidas considerando-se a transformação monótona

$$X = \frac{1}{2} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right),$$

em que $T \sim \text{BS}(\alpha, \beta)$ e $X \sim \text{N}(0, \alpha^2/4)$. Escrevendo $T = \beta[1 + 2X^2 + 2X(1 + X^2)^{\frac{1}{2}}]$ e usando a linearidade da esperança e os momentos da distribuição normal chegamos ao resultado.

Dada a variável aleatória Z com distribuição $\text{N}(0, 1)$, segue que a relação que a distribuição BS possui com a distribuição normal é dada mediante a representação estocástica

$$T = \frac{\beta(\alpha Z + \sqrt{\alpha^2 Z^2 + 4})^2}{4}.$$

Podemos ainda obter os percentis t_p da distribuição BS através da resolução da equação

$$t_p - \beta^{\frac{1}{2}} \alpha z t_p^{\frac{1}{2}} - \beta = 0,$$

em que $z = \Phi^{-1}(p)$ é o quantil de ordem p da distribuição $\text{N}(0,1)$.

Outras propriedades da distribuição BS são dadas abaixo. Se $T \sim \text{BS}(\alpha, \beta)$, então:

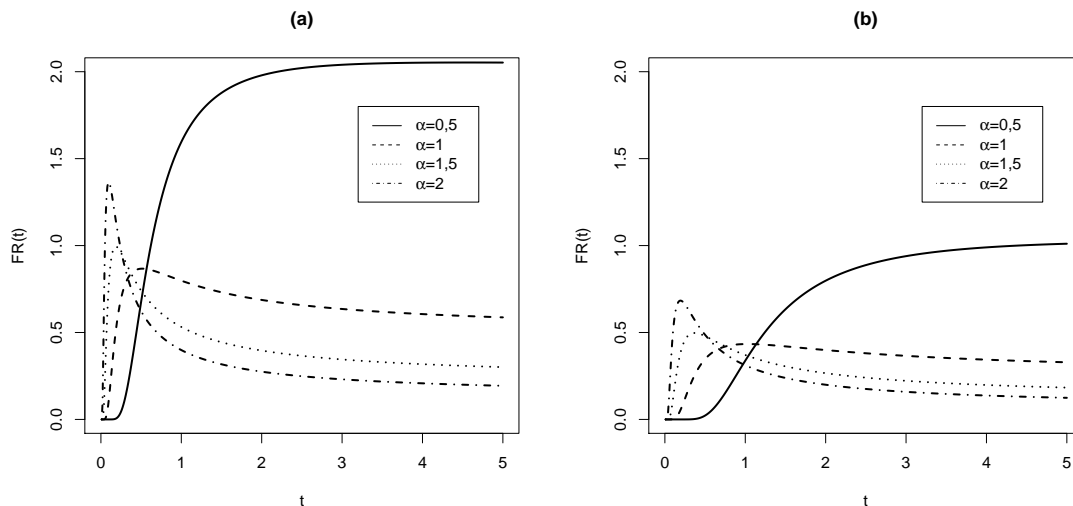


Figura 1.3: gráficos das funções risco da distribuição BS para os valores de α indicados com $\beta=1$ (a) e $\beta = 2$ (b).

- $\frac{T}{\beta} \sim \text{BS}(\alpha, 1)$;
- $kT \sim \text{BS}(\alpha, k\beta)$, para $k > 0$;
- $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$;
- $F_T(\beta) = \Phi(0) = 1/2$.

Das propriedades acima temos que β é o parâmetro de escala da distribuição, que a distribuição é fechada sob proporcionalidade e sob reciprocidade e que β é a mediana da distribuição.

A generalização da distribuição BS proposta por Meintanis (2010) é o tema central deste trabalho, a qual denotamos por BSG. Abordamos propriedades dessa distribuição, bem como aspectos inferenciais: estimação pontual e intervalar. Consideramos a distribuição log BSG, que é um caso particular da distribuição senh-normal, na qual exploramos suas propriedades, mostrando a relação entre esta e a distribuição BS. Depois consideramos, o modelo de regressão log-linear BSG, mais especificamente, consideramos diagnóstico de influência baseado no método de deleção de casos. Sabe-se que para avaliar a influência da i -ésima observação nas estimativas de um parâmetro, uma aproximação direta é avaliar o diagnóstico de casos isolados com o i -ésimo caso deletado. Desde o trabalho pioneiro de Cook (1977), diagnósticos baseados em deleção de casos tais como distância de Cook ou afastamento pela verossimilhança têm sido aplicados com sucesso a vários modelos estatísticos; veja, por exemplo, Christensen

et al. (1992), Davison and Tsai (1992), Wei (1998), Tang et al. (2000), Galea et al. (2004), etc. Em Xie e Wei (2007), os autores propuseram diversas medidas de diagnóstico baseadas no modelo de deleção de casos (*case deletion model* ou CDM) para o modelo de regressão log-linear BS. Então, seguindo Xie e Wei (2007), obtivemos algumas medidas para o modelo log-linear BSG: distância de Cook; teste escore para testar a presença de *outlier* no modelo de regressão log-linear BSG e teste para detectar a homogeneidade do parâmetro de forma.

1.2.3 Organização da dissertação

A presente dissertação está dividida em quatro capítulos, contando com os preliminares. No Capítulo 2 apresentamos a distribuição BSG, proposta por Menthais (2010), mostrando suas principais propriedades, momentos, obtenção dos estimadores de máxima verossimilhança (MV) dos parâmetros, matriz de informação observada e um estudo de simulação para verificar o desempenho das estimativas dos parâmetros a medida que modificamos o valor do parâmetro introduzido ao modelo. No Capítulo 3 exploramos a distribuição senh-normal e suas propriedades, entre elas a que estabelece uma importante relação entre esta e a distribuição BS. No capítulo 4, introduzimos o modelo de regressão log-BSG, damos um breve esboço do modelo de deleção de casos (CDM), obtemos uma aproximação para as estimativas, bem como apresentamos algumas medidas, tais como distância de Cook generalizada e afastamento pela verossimilhança usando tal método, a fim de identificar observações influentes. Depois obteremos a estatística do teste escore para detectar outliers baseando-se no modelo de *outlier* por deslocamento da média (MSOM). Ainda no mesmo capítulo, discutiremos o fato do parâmetro de forma no modelo de regressão log-BSG não ser necessariamente constante e analisaremos um teste de heterogeneidade para este parâmetro utilizando a estatística de razão de verossimilhança. Por fim, relatamos um exemplo em que a teoria mencionada acima será utilizada.

1.2.4 Plataforma computacional

As avaliações numéricas realizadas nessa dissertação foram feitas no *software* estatístico R, na sua versão 2.15.3 para sistemas operacionais Windows e que se encontra disponível gratuitamente no endereço www.r-project.org.

Capítulo 2

Distribuição Birnbaum-Saunders generalizada

O objeto de estudo deste capítulo é a distribuição BSG apresentada em Meintanis (2010). Essa generalização foi obtida por meio da introdução de um parâmetro de forma extra na distribuição BS, o qual denotaremos por λ . Meintanis (2010) propôs substituir o expoente $1/2$ por λ em (1.2). Assim, se $T \sim \text{BSG}(\alpha, \beta, \lambda)$, então a fda de T fica dada por

$$F_T(t; \alpha, \beta) = \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^\lambda - \left(\frac{\beta}{t} \right)^\lambda \right] \right\}, t > 0, \alpha > 0, \beta > 0, \lambda > 0. \quad (2.1)$$

A fdp é dada por

$$f_T(t; \alpha, \beta, \lambda) = \frac{\lambda}{\alpha \sqrt{2\pi t}} \left[\left(\frac{t}{\beta} \right)^\lambda + \left(\frac{\beta}{t} \right)^\lambda \right] \exp \left\{ -\frac{1}{2\alpha^2} \left[\left(\frac{t}{\beta} \right)^{2\lambda} + \left(\frac{\beta}{t} \right)^{2\lambda} - 2 \right] \right\}, \quad (2.2)$$

para $t > 0$, α , β e $\lambda > 0$.

De (2.1) e (2.2) obtemos as funções de sobrevivência e de risco, as quais são dadas por

$$S(t) = 1 - \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^\lambda - \left(\frac{\beta}{t} \right)^\lambda \right] \right\}$$

e

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\frac{\lambda}{\alpha \sqrt{2\pi t}} \left[\left(\frac{t}{\beta} \right)^\lambda + \left(\frac{\beta}{t} \right)^\lambda \right] \exp \left\{ -\frac{1}{2\alpha^2} \left[\left(\frac{t}{\beta} \right)^{2\lambda} + \left(\frac{\beta}{t} \right)^{2\lambda} - 2 \right] \right\}}{1 - \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^\lambda - \left(\frac{\beta}{t} \right)^\lambda \right] \right\}}, t > 0.$$

A Figura (2.1) nos mostra gráficos de densidades da distribuição BSG para vários valores de α . Podemos observar que ao variarmos o valor do parâmetro λ o gráfico da fdp muda de forma. Quanto maior é o valor de λ mais a distribuição tende a ser simétrica em torno do

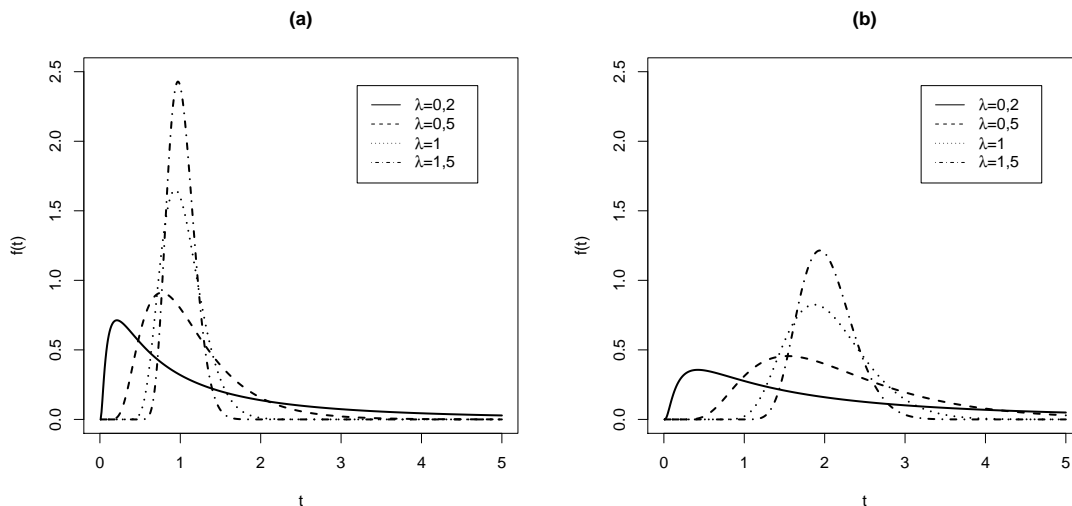


Figura 2.1: gráficos de funções densidade da distribuição BSG com parâmetros $\alpha = 0,5$ e $\beta = 1$ (a), e $\alpha = 0,5$ e $\beta = 2$ (b).

parâmetro β . Notamos também que o valor do parâmetro λ altera a curtose da distribuição, de modo que a fdp se torna leptocúrtica à medida que λ cresce, que é o contrário do que ocorre com o parâmetro α . A Figura (2.2) nos mostra os gráficos das funções de distribuição e de sobrevivência, evidenciando a relação entre elas. A Figura (2.3) nos mostra o comportamento da função de risco a qual não segue um comportamento monótono.

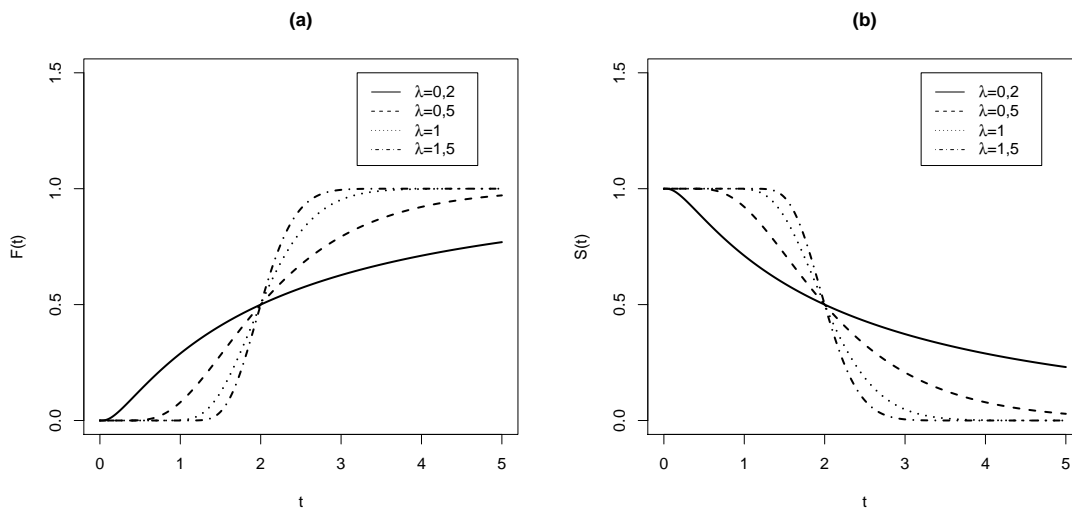


Figura 2.2: gráficos de funções de distribuição acumulada (a) e gráficos de funções de sobrevivência (b) da distribuição BSG, ambos com parâmetro $\beta = 2$.

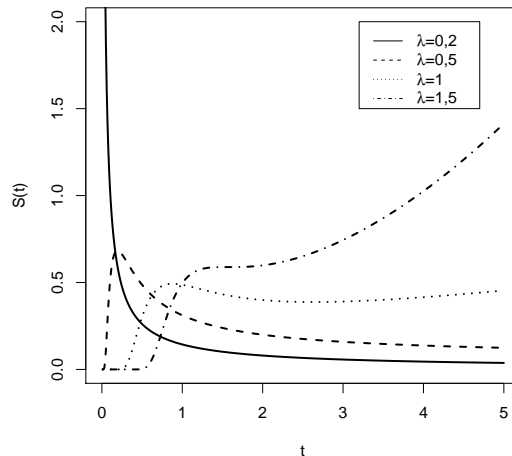


Figura 2.3: gráfico das funções risco da distribuição BSG para os valores de $\alpha = 2$ indicados e $\beta = 2$.

2.1 Propriedades

1. Se $T \sim \text{BSG}(\alpha, \beta, \lambda)$, então $Y = T/\beta \sim \text{BSG}(\alpha, 1, \lambda)$, $\beta > 0$.

Prova.

$$\begin{aligned} F_Y(t) &= P\left(\frac{T}{\beta} \leq t\right) = P(T \leq t\beta) = \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t\beta}{\beta}\right)^\lambda - \left(\frac{\beta}{t\beta}\right)^\lambda\right]\right\} \\ &= \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t}{1}\right)^\lambda - \left(\frac{1}{t}\right)^\lambda\right]\right\}. \end{aligned}$$

■

Isto significa que β é o parâmetro de escala da distribuição.

2. Se $T \sim \text{BSG}(\alpha, \beta, \lambda)$, então $Y = kT \sim \text{BSG}(\alpha, k\beta, \lambda)$, $k > 0$.

Prova.

$$F_Y(t) = P(kT \leq t) = P\left(T \leq \frac{t}{k}\right) = \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t}{k\beta}\right)^\lambda - \left(\frac{k\beta}{t}\right)^\lambda\right]\right\}.$$

■

Em outras palavras, a propriedade acima afirma que a distribuição BSG é fechada sob proporcionalidade.

3. Se $T \sim \text{BSG}(\alpha, \beta, \lambda)$, então $T^{-1} \sim \text{BSG}(\alpha, \beta^{-1}, \lambda)$.

Prova. Seja $Y = T^{-1}$. Temos

$$\begin{aligned} F_Y(t) = P(T^{-1} \leq t) &= 1 - P\left(T \leq \frac{1}{t}\right) \\ &= 1 - \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{1}{t\beta}\right)^\lambda - (t\beta)^\lambda\right]\right\} \\ &= \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t}{\beta^{-1}}\right)^\lambda - \left(\frac{\beta^{-1}}{t}\right)^\lambda\right]\right\}. \end{aligned}$$

■

Ou seja, a distribuição BSG é fechada sob reciprocidade.

4. Dado $\beta > 0$, o parâmetro de escala da distribuição, temos $F_T(\beta) = \Phi(0) = 1/2$.

Prova.

$$\begin{aligned} F_T(\beta) = P(T \leq \beta) &= \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{\beta}{\beta}\right)^\lambda - \left(\frac{\beta}{\beta}\right)^\lambda\right]\right\} \\ &= \Phi(0) = \frac{1}{2}. \end{aligned}$$

■

Isto significa que β é a mediana da distribuição.

5. O quantil de ordem p da distribuição BSG(α, β, λ) é obtido de

$$\alpha\beta^\lambda\Phi^{-1}(p)t_p^\lambda + \beta^{2\lambda} - t_p^{2\lambda} = 0.$$

Prova. Para encontrar a relação acima basta fazer

$$\begin{aligned} F_T(t_p) = p &\Leftrightarrow \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t_p}{\beta}\right)^\lambda - \left(\frac{\beta}{t_p}\right)^\lambda\right]\right\} = p \\ &\Leftrightarrow \frac{1}{\alpha}\left[\left(\frac{t_p}{\beta}\right)^\lambda - \left(\frac{\beta}{t_p}\right)^\lambda\right] = \Phi^{-1}(p) \\ &\Leftrightarrow \frac{t_p^{2\lambda} - \beta^{2\lambda}}{\beta^\lambda t_p^\lambda} = \alpha\Phi^{-1}(p) \\ &\Leftrightarrow t_p^{2\lambda} - \beta^{2\lambda} - \beta^\lambda t_p^\lambda \alpha z_p = 0, \end{aligned}$$

em que z_p é o quantil de ordem p da distribuição $N(0, 1)$.

■

6. A variável aleatória T pode ser escrita como

$$T = \beta \left(\frac{\alpha Z + \sqrt{\alpha^2 Z^2 + 4}}{2} \right)^{1/\lambda}, \quad Z \sim N(0, 1). \quad (2.3)$$

Prova. Temos

$$F_T(t) = \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^\lambda - \left(\frac{\beta}{t} \right)^\lambda \right] \right\}.$$

Então,

$$Z = \frac{1}{\alpha} \left[\left(\frac{T}{\beta} \right)^\lambda - \left(\frac{\beta}{T} \right)^\lambda \right],$$

em que $Z \sim N(0, 1)$. Logo,

$$\alpha Z = \left(\frac{T}{\beta} \right)^\lambda - \left(\frac{\beta}{T} \right)^\lambda,$$

Daí,

$$\alpha Z = \frac{T^{2\lambda} - \beta^{2\lambda}}{(\beta T)^\lambda}$$

e assim

$$T^{2\lambda} - \alpha z \beta^\lambda T^\lambda - \beta^{2\lambda} = 0.$$

Chamando $T^\lambda = Y$, temos

$$Y^2 - \alpha Z \beta^\lambda Y - \beta^{2\lambda} = 0.$$

Logo,

$$Y = \frac{\beta^\lambda (\alpha Z + \sqrt{\alpha^2 Z^2 + 4})}{2}$$

e, conseqüentemente,

$$T = \beta \left(\frac{\alpha Z + \sqrt{\alpha^2 Z^2 + 4}}{2} \right)^{1/\lambda}.$$

■

Obtivemos numericamente os quatro primeiros momentos da distribuição BSG. Conseqüentemente, podemos obter a média, a variância e os coeficientes de variação (CV), assimetria (CA) e curtose (CC) de T , que são dados, respectivamente, por

$$E(T) = \frac{e^{\frac{1}{\alpha^2}} \beta}{\sqrt{2\pi\alpha}} \left[K \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + K \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right];$$

$$\text{Var}(T) = \frac{\beta^2 e^{\frac{1}{\alpha^2}}}{\sqrt{2\pi\alpha}} \left[\text{K} \left(\frac{\lambda+2}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-2}{2\lambda}, \frac{1}{\alpha^2} \right) \right] - \frac{\beta^2 e^{\frac{2}{\alpha^2}}}{2\pi\alpha^2} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]^2;$$

$$\text{CV}(T) = \frac{\left\{ \text{K} \left(\frac{\lambda+2}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-2}{2\lambda}, \frac{1}{\alpha^2} \right) - \frac{e^{\frac{1}{\alpha^2}}}{\sqrt{2\pi\alpha}} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right] \right\}^{\frac{1}{2}}}{\left(\frac{e^{\frac{1}{\alpha^2}}}{\sqrt{2\pi\alpha}} \right)^{\frac{1}{2}} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]}$$

$$\text{CA}(T) = \frac{\text{E}[(T - \text{E}(T))^3]}{\text{Var}(T)^{\frac{3}{2}}},$$

em que

$$\begin{aligned} \text{E}[(T - \text{E}(T))^3] &= \frac{\beta^3 e^{\frac{3}{\alpha^2}}}{\sqrt{2\pi\pi\alpha^3}} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]^3 \\ &+ \frac{\beta^3 e^{\frac{1}{\alpha^2}}}{\sqrt{2\pi\alpha}} \left[\text{K} \left(\frac{\lambda+3}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-3}{2\lambda}, \frac{1}{\alpha^2} \right) \right] \\ &- \frac{3\beta^3 e^{\frac{2}{\alpha^2}}}{2\pi\alpha^2} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right] \\ &\left[\text{K} \left(\frac{\lambda+2}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-2}{2\lambda}, \frac{1}{\alpha^2} \right) \right]; \end{aligned}$$

$$\text{CC}(T) = \frac{\text{E}[(T - \text{E}(T))^4]}{\text{Var}(T)^2},$$

em que

$$\begin{aligned} \text{E}[(T - \text{E}(T))^4] &= -\frac{3\beta^4 e^{\frac{4}{\alpha^2}}}{4\pi^2\alpha^4} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]^4 \\ &+ \frac{\beta^4 e^{\frac{1}{\alpha^2}}}{\sqrt{2\pi\alpha}} \left[\text{K} \left(\frac{\lambda+4}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-4}{2\lambda}, \frac{1}{\alpha^2} \right) \right] \\ &+ \frac{3\beta^4 e^{\frac{3}{\alpha^2}}}{\sqrt{2\pi\pi\alpha^3}} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]^2 \\ &\left[\text{K} \left(\frac{\lambda+2}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-2}{2\lambda}, \frac{1}{\alpha^2} \right) \right] \\ &- \frac{2\beta^4 e^{\frac{2}{\alpha^2}}}{\pi\alpha^2} \left[\text{K} \left(\frac{\lambda+1}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-1}{2\lambda}, \frac{1}{\alpha^2} \right) \right]^2 \\ &\left[\text{K} \left(\frac{\lambda+3}{2\lambda}, \frac{1}{\alpha^2} \right) + \text{K} \left(\frac{\lambda-3}{2\lambda}, \frac{1}{\alpha^2} \right) \right]. \end{aligned}$$

Nas expressões acima, K é a função de Bessel do terceiro tipo modificada dada por

$$K(\nu, z) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(\nu x - z \cosh x) dx \quad (2.4)$$

(ver Glasser et al., 2012, p. 5).

2.2 Estimação dos parâmetros do modelo

Sejam T_1, \dots, T_n variáveis aleatórias independentes e identicamente distribuídas, em que cada T_i , para $i = 1, \dots, n$, segue uma distribuição BSG com parâmetros α , β e λ . Baseado em (2.1), segue que o logaritmo da função de verossimilhança é dado por

$$l(\boldsymbol{\theta}) = n \log\left(\frac{\lambda}{\alpha}\right) + n \log\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{1}{2\alpha^2} \sum_{j=1}^n v_j^2 - \sum_{j=1}^n \log(t_j) + \sum_{j=1}^n \log(w_j), \quad (2.5)$$

em que $w_j = (t_j/\beta)^\lambda + (\beta/t_j)^\lambda$ e $v_j = (t_j/\beta)^\lambda - (\beta/t_j)^\lambda$.

As funções escore com relação a α , β e λ , são dadas, respectivamente, por

$$U_\alpha = -\frac{n}{\alpha} + \frac{1}{\alpha^3} \sum_{j=1}^n v_j^2, \quad (2.6)$$

$$U_\beta = \frac{\lambda}{\beta} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j \quad \text{e} \quad (2.7)$$

$$U_\lambda = \frac{n}{\lambda} + \sum_{j=1}^n \log\left(\frac{t_j}{\beta}\right) \left(\frac{1}{w_j} - \frac{w_j}{\alpha^2} \right) v_j, \quad (2.8)$$

em que w_j e v_j são dadas em (2.5).

Os estimadores de máxima verossimilhança de α^2 , β e λ deveriam ser obtidos resolvendo o sistema de equações não lineares formado pelas equações $U_\alpha = 0$, $U_\beta = 0$ e $U_\lambda = 0$, respectivamente. No entanto, surge um problema de convergência ao tentar estimar os três parâmetros conjuntamente. Assim, só foi possível encontrar os estimadores de α e β , deixando o parâmetro λ fixo. Para maiores detalhes, ver Meintanis (2010).

A matriz de informação observada de $\boldsymbol{\theta} = (\alpha, \beta)^\top$, composta pelas derivadas de segunda ordem em relação aos parâmetros da distribuição, é dada por

$$\ddot{l}(\hat{\boldsymbol{\theta}}) = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \left[\begin{array}{cc} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta^2} \end{array} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},$$

em que

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} &= \frac{n}{\alpha^2} - \frac{3}{\alpha^4} \sum_{j=1}^n v_j^2, \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha} = -\frac{2\lambda}{\beta \alpha^3} \sum_{j=1}^n w_j v_j \quad \text{e} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta^2} &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j + \frac{4\lambda^2}{\beta^2} \sum_{j=1}^n \frac{1}{w_j^2} - \frac{2\lambda^2}{\alpha^2 \beta^2} \sum_{j=1}^n (v_j^2 + 2).\end{aligned}$$

Os detalhes das derivadas de primeira e segunda ordens dadas acima podem ser vistos no Apêndice A.

2.3 Estudo de simulação

Com o objetivo de analisar o comportamento dos estimadores de MV dos parâmetros α e β para λ fixado, realizamos um estudo baseado em simulações de Monte Carlo (MC). Geramos números pseudo-aleatórios da distribuição BSG a partir da equação (2.3). Para essa simulação consideramos 5000 réplicas e tamanhos amostrais $n = 10, 25, 50$ e 100 , valores de $\lambda = 0, 1; 0, 3; 0, 5; 1, 0; 2, 0$ e $\alpha = 0, 2; 0, 5; 1$ e $\beta = 1$. As estimativas dos parâmetros α e β de cada amostra foram obtidas a partir do método MV usando adaptações dos comandos já implementados computacionalmente no *software* R por meio do pacote *gbs*. Foram obtidos estimativas dos vieses relativos (VR) e da raiz quadrada do erro quadrático médio ($\sqrt{\text{EQM}}$). Os resultados das simulações encontram-se nas Tabelas 2.1-2.3. Por meio dessas tabelas podemos observar que os erros e vieses em relação ao parâmetro α não têm variabilidade considerável conforme variam os valores de λ . Por outro lado, o valor de λ influencia bastante as estimativas de β , se aproximando do valor verdadeiro à medida que o valor de λ cresce. Também vemos que aumentando os tamanhos amostrais as estimativas de α e β ficam próximas dos valores verdadeiros. Ainda notamos que, tanto para a estimação de α como para a de β , os erros quadráticos médios e os vieses tornam-se maiores à medida que aumentamos o valor de λ .

Para cada combinação de parâmetros descrita foi obtidos os gráficos quantil versus quantil (QQ) para verificar a suposição de normalidade para os estimadores de MV de α e β . Os gráficos foram obtidos para amostras de tamanho 100 . Observamos que, para cada valor de α e β , somente os gráficos referentes às estimativas de β se alteraram consideravelmente a medida que foi modificado o valor de λ , obtendo uma melhor aproximação em relação à distri-

buição normal, conforme o valor de λ aumenta. Também notamos uma melhor aproximação para valores mais baixos de α . Em geral, a distribuição dos estimadores possuem uma boa concordância com a distribuição normal, veja Figuras (2.4) - (2.6).

Baseados na suposição de normalidade dos estimadores de MV, realizamos uma simulação de MC para obter os intervalos de confiança médios, bem como as probabilidade de coberturas (PC_s) de 95% para os parâmetros α e β considerando os tamanhos amostrais $n = 10, 25, 50$ e 100 , para valores de $\alpha = 0, 2; 0, 5; 1$, $\beta = 1$ e $\lambda = 0, 25; 0, 30; 0, 50; 1, 00; 2, 00$. Consideramos 5000 réplicas de MC. Os intervalos de confiança médios, bem como as probabilidade de coberturas empíricas são apresentados nas Tabelas (2.4) - (2.6). Em todos os casos estabelecemos o nível de confiança igual 0,95. Podemos observar que os intervalos de confiança para α não variam com os valores de λ , como já esperado pelas análises anteriores. Entretanto, os intervalos de confiança para β possuem alterações consideráveis de acordo com a variação de λ . Para ambas as estimativas e em todos os casos analisados as PC_s crescem se aproximando do nível nominal, conforme aumentamos o tamanho amostral, sem alterações bruscas com relação aos valores de λ .

Tabela 2.1: estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 0, 2$ e $\beta = 1$.

n	λ	Estimação de α			Estimação de β		
		Estimativa	Viés	$\sqrt{\text{EQM}}$	Estimativa	Viés	$\sqrt{\text{EQM}}$
10	0,1	0,1848	-0,0152	0,0459	1,0529	0,0529	0,3439
	0,3	0,1843	-0,0157	0,0474	1,0037	0,0037	0,1048
	0,5	0,1850	-0,0150	0,0462	1,0026	0,0026	0,0634
	1,0	0,1851	-0,0149	0,0468	1,0000	0,0000	0,0313
	2,0	0,1838	-0,0162	0,0471	1,0001	0,0001	0,0156
25	0,1	0,1941	-0,0059	0,0295	1,0197	0,0197	0,2066
	0,3	0,1940	-0,0060	0,0290	1,0014	0,0014	0,0670
	0,5	0,1943	-0,0057	0,0280	1,0009	0,0009	0,0403
	1,0	0,1934	-0,0066	0,0285	1,0002	0,0002	0,0198
	2,0	0,1943	-0,0057	0,0288	0,9999	-0,0001	0,0099
50	0,1	0,1974	-0,0026	0,0200	1,0067	0,0067	0,1422
	0,3	0,1970	-0,0030	0,0203	1,0002	0,0002	0,0472
	0,5	0,1968	-0,0032	0,0201	1,0008	0,0008	0,0281
	1,0	0,1971	-0,0029	0,0202	1,0004	0,0004	0,0141
	2,0	0,1971	-0,0029	0,0203	1,0000	0,0000	0,0070
100	0,1	0,1984	-0,0016	0,0143	1,0041	0,0041	0,0998
	0,3	0,1987	-0,0013	0,0143	1,0008	0,0008	0,0334
	0,5	0,1985	-0,0015	0,0143	1,0001	0,0001	0,0200
	1,0	0,1985	-0,0015	0,0141	0,9999	-0,0001	0,0099
	2,0	0,1987	-0,0013	0,0142	1,0001	0,0001	0,0051

Tabela 2.2: estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 0, 5$ e $\beta = 1$.

n	λ	Estimação de α			Estimação de β		
		Estimativa	Viés	$\sqrt{\text{EQM}}$	Estimativa	Viés	$\sqrt{\text{EQM}}$
10	0,1	0,4576	-0,0424	0,1187	1,3411	0,3411	1,2518
	0,3	0,4597	-0,0403	0,1173	1,0347	0,0347	0,2695
	0,5	0,4610	-0,0390	0,1167	1,0098	0,0098	0,1574
	1,0	0,4569	-0,0431	0,1169	1,0028	0,0028	0,0763
	2,0	0,4608	-0,0392	0,1175	0,9997	-0,0003	0,0389
25	0,1	0,4861	-0,0139	0,0709	1,1215	0,1215	0,5773
	0,3	0,4839	-0,0161	0,0720	1,0118	0,0118	0,1660
	0,5	0,4846	-0,0154	0,0721	1,0040	0,0040	0,0979
	1,0	0,4858	-0,0142	0,0728	1,0010	0,0010	0,0489
	2,0	0,4847	-0,0153	0,0736	1,0001	0,0001	0,0241
50	0,1	0,4926	-0,0074	0,0501	1,0618	0,0618	0,3785
	0,3	0,4930	-0,0070	0,0507	1,0100	0,0100	0,1171
	0,5	0,4932	-0,0068	0,0499	1,0026	0,0026	0,0680
	1,0	0,4938	-0,0062	0,0498	0,9990	-0,0010	0,0332
	2,0	0,4937	-0,0063	0,0508	0,9998	-0,0002	0,0169
100	0,1	0,4961	-0,0039	0,0349	1,0347	0,0347	0,2559
	0,3	0,4956	-0,0044	0,0360	1,0018	0,0018	0,0798
	0,5	0,4966	-0,0034	0,0352	1,0010	0,0010	0,0485
	1,0	0,4956	-0,0044	0,0357	1,0002	0,0002	0,0242
	2,0	0,4965	-0,0035	0,0359	0,9999	-0,0001	0,0120

Tabela 2.3: estimativas de MV para os parâmetros α e β da distribuição BSG para $\alpha = 1$ e $\beta = 1$.

n	λ	Estimação de α			Estimação de β		
		Estimativa	Viés	\sqrt{EQM}	Estimativa	Viés	\sqrt{EQM}
10	0,1	0,9615	-0,0385	1,9018	2,5476	1,5476	5,9217
	0,3	0,9161	-0,0839	0,2389	1,1072	0,1072	0,5692
	0,5	0,9128	-0,0872	0,2362	1,0441	0,0441	0,3058
	1,0	0,9191	-0,0809	0,2318	1,0096	0,0096	0,1416
	2,0	0,9130	-0,0870	0,2346	1,0025	0,0025	0,0709
25	0,1	0,9674	-0,0326	0,1458	1,4675	0,4675	1,7294
	0,3	0,9646	-0,0354	0,1469	1,0400	0,0400	0,3193
	0,5	0,9653	-0,0347	0,1428	1,0147	0,0147	0,1821
	1,0	0,9691	-0,0309	0,1447	1,0039	0,0039	0,0891
	2,0	0,9681	-0,0319	0,1444	1,0008	0,0008	0,0445
50	0,1	0,9820	-0,0180	0,1011	1,2108	0,2108	0,8193
	0,3	0,9836	-0,0164	0,1019	1,0184	0,0184	0,2178
	0,5	0,9828	-0,0172	0,1002	1,0076	0,0076	0,1273
	1,0	0,9831	-0,0169	0,1005	1,0021	0,0021	0,0618
	2,0	0,9831	-0,0169	0,1007	1,0013	0,0013	0,0311
100	0,1	0,9905	-0,0095	0,0701	1,0912	0,0912	0,4992
	0,3	0,9911	-0,0089	0,0706	1,0050	0,0050	0,1510
	0,5	0,9930	-0,0070	0,0699	1,0032	0,0032	0,0886
	1,0	0,9905	-0,0095	0,0717	1,0011	0,0011	0,0446
	2,0	0,9938	-0,0062	0,0704	0,9997	-0,0003	0,0216

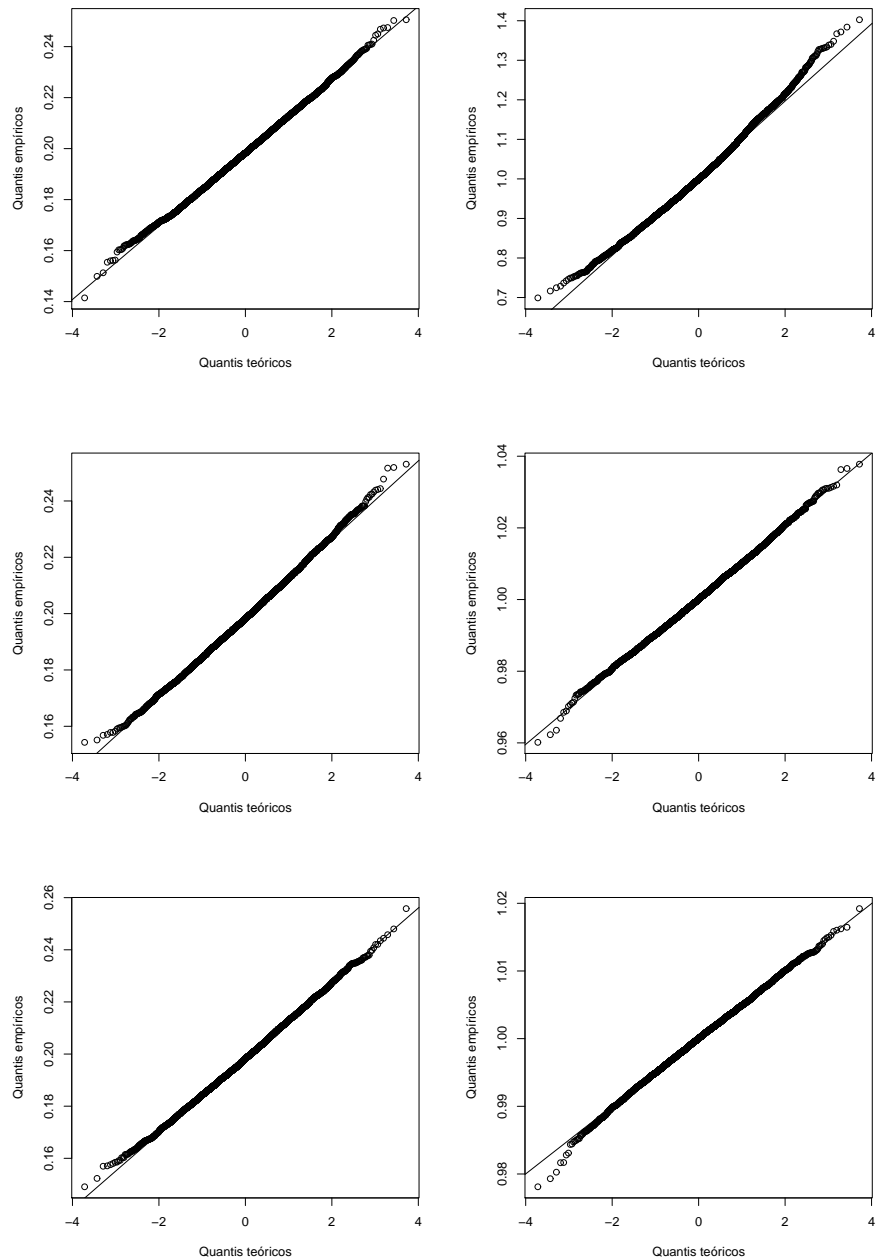


Figura 2.4: gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 0, 2$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).

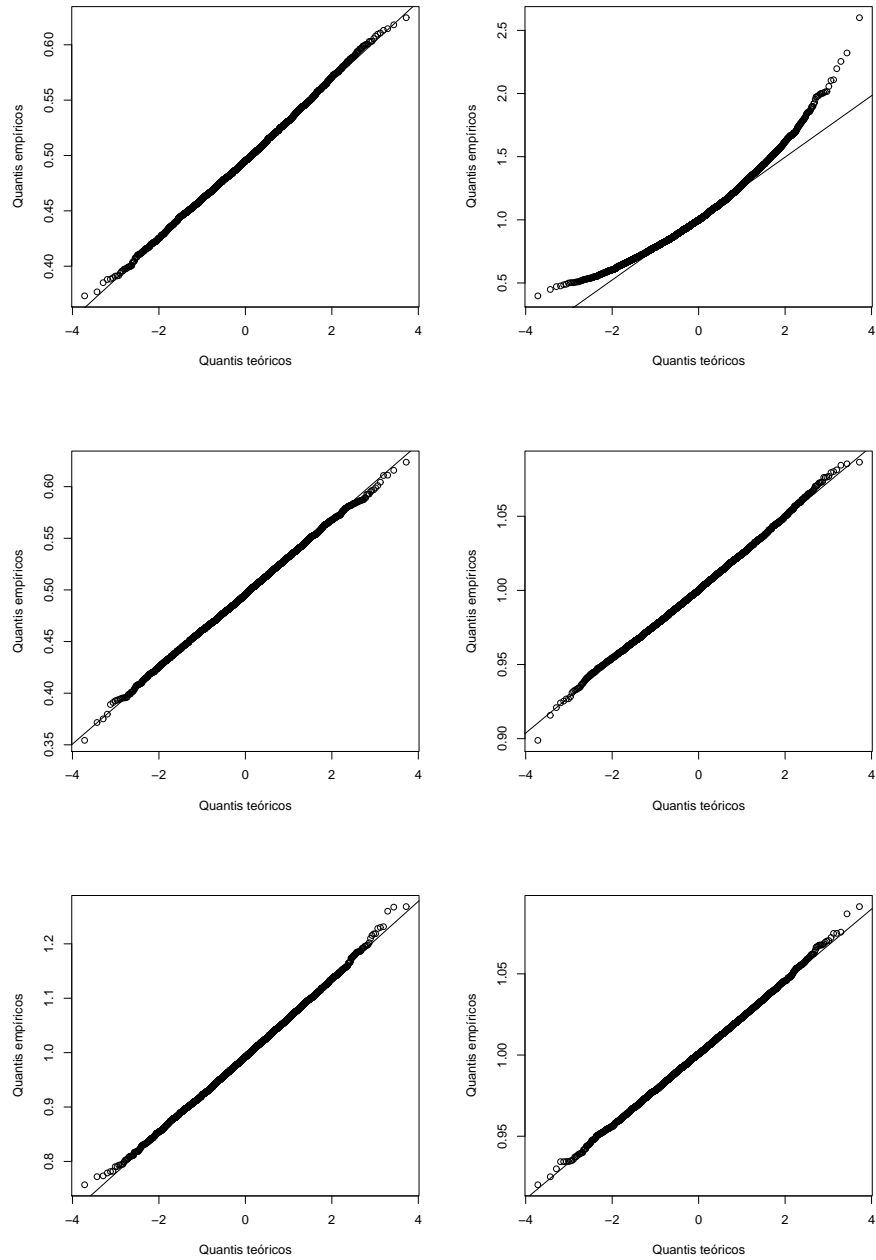


Figura 2.5: gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 0,5$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).

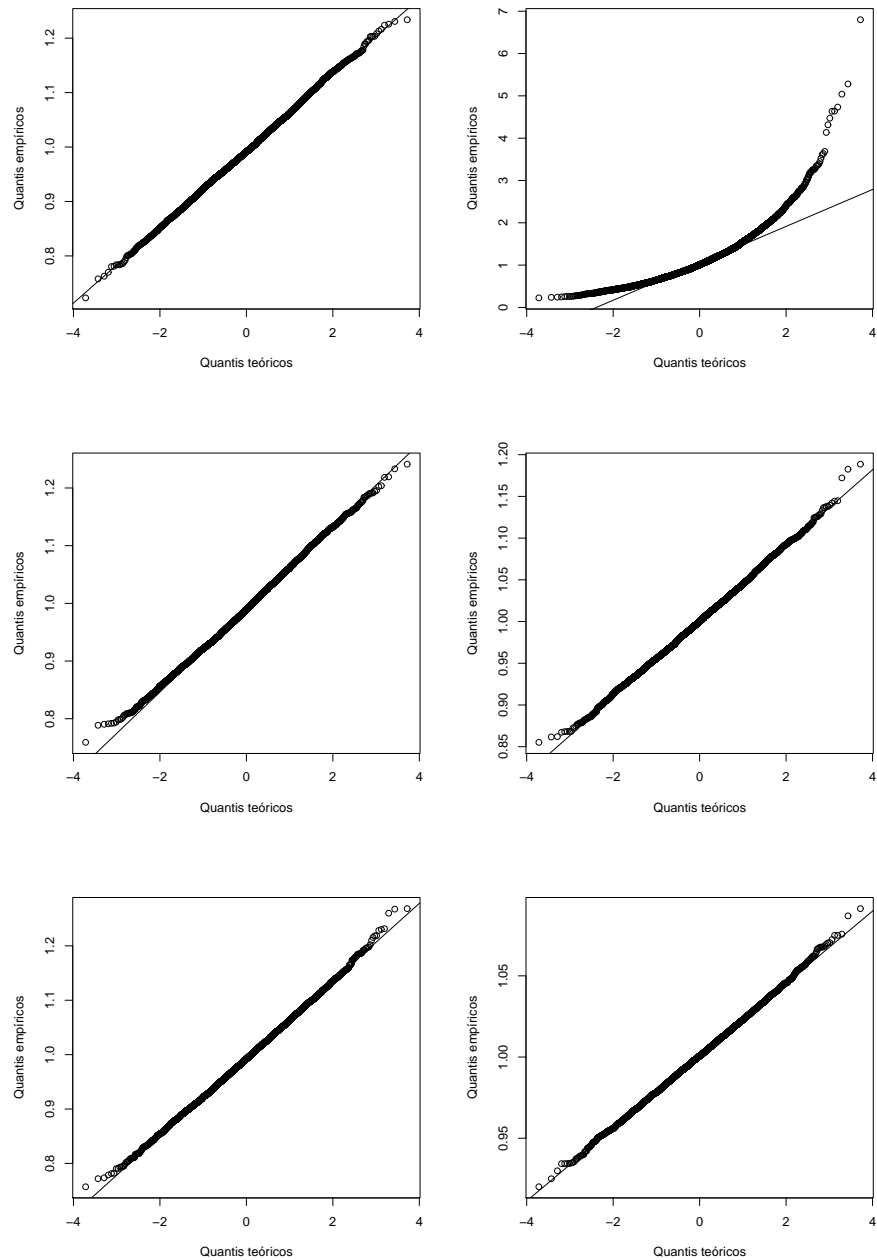


Figura 2.6: gráficos QQ para normalidade dos estimadores de α e β para os valores fixados de $\alpha = 1$, $\beta = 1$ e λ variando em 0,1 (linha 1); 1,0 (linha 2) e 2,0 (linha 3).

Tabela 2.4: limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 0,2$ e $\beta = 1$.

n	λ	α			β		
		LI	LS	PC (%)	LI	LS	PC (%)
10	0,25	0,10	0,27	85,24	0,78	1,24	89,62
	0,30	0,10	0,27	85,24	0,81	1,20	89,66
	0,50	0,10	0,27	85,24	0,89	1,12	89,90
	1,00	0,10	0,27	85,24	0,94	1,06	90,02
	2,00	0,10	0,27	85,24	0,97	1,03	90,00
25	0,25	0,14	0,25	90,62	0,85	1,15	93,32
	0,30	0,14	0,25	90,62	0,87	1,13	93,22
	0,50	0,14	0,25	90,62	0,92	1,08	93,14
	1,00	0,14	0,25	90,62	0,96	1,04	93,04
	2,00	0,14	0,25	90,62	0,98	1,02	93,00
50	0,25	0,16	0,24	92,86	0,89	1,11	94,28
	0,30	0,16	0,24	92,86	0,91	1,09	94,34
	0,50	0,16	0,24	92,86	0,95	1,05	94,24
	1,00	0,16	0,24	92,86	0,97	1,03	94,40
	2,00	0,16	0,24	92,84	0,99	1,01	94,46
100	0,25	0,17	0,23	94,08	0,92	1,08	94,84
	0,30	0,17	0,23	94,08	0,94	1,06	94,82
	0,50	0,17	0,23	94,08	0,96	1,04	94,82
	1,00	0,17	0,23	94,08	0,98	1,02	95,02
	2,00	0,17	0,23	94,12	0,99	1,01	94,96

Tabela 2.5: limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 0,5$ e $\beta = 1$.

n	λ	α			β		
		LI	LS	PC (%)	LI	LS	PC(%)
10	0,25	0,26	0,66	84,98	0,47	1,63	88,90
	0,30	0,26	0,66	84,96	0,55	1,51	89,26
	0,50	0,26	0,66	84,96	0,73	1,29	89,44
	1,00	0,26	0,66	84,96	0,86	1,14	89,66
	2,00	0,26	0,66	84,96	0,93	1,07	89,76
25	0,25	0,35	0,62	90,56	0,64	1,39	93,10
	0,30	0,35	0,62	90,56	0,70	1,32	93,24
	0,50	0,35	0,62	90,56	0,82	1,19	93,26
	1,00	0,35	0,62	90,56	0,91	1,09	93,02
	2,00	0,35	0,62	90,56	0,95	1,05	92,92
50	0,25	0,40	0,59	92,76	0,74	1,28	93,72
	0,30	0,40	0,59	92,74	0,78	1,23	93,92
	0,50	0,40	0,59	92,76	0,87	1,13	94,32
	1,00	0,40	0,59	92,76	0,93	1,07	94,30
	2,00	0,40	0,59	92,74	0,97	1,03	94,54
100	0,25	0,43	0,56	94,04	0,82	1,19	94,60
	0,30	0,43	0,56	94,04	0,85	1,16	94,70
	0,50	0,43	0,56	94,02	0,91	1,10	94,76
	1,00	0,43	0,56	94,04	0,95	1,05	94,82
	2,00	0,43	0,56	94,06	0,98	1,02	94,96

Tabela 2.6: limite inferior (LI) e limite superior (LS) do intervalo de confiança médio; e PC de 95% para os parâmetros e n indicados considerando $\alpha = 1, 0$ e $\beta = 1, 0$.

n	λ	α			β		
		LI	LS	PC(%)	LI	LS	PC(%)
10	0,25	0,41	1,05	66,02	0,19	2,21	82,36
	0,30	0,48	1,22	80,80	0,23	2,01	86,84
	0,50	0,52	1,32	84,20	0,52	1,56	88,72
	1,00	0,52	1,32	84,20	0,76	1,26	89,22
	2,00	0,52	1,32	84,20	0,88	1,13	89,26
25	0,25	0,70	1,24	90,26	0,35	1,77	91,96
	0,30	0,70	1,24	90,26	0,46	1,62	92,30
	0,50	0,70	1,24	90,26	0,67	1,35	92,88
	1,00	0,70	1,24	90,26	0,83	1,17	92,90
	2,00	0,70	1,24	90,26	0,92	1,08	92,94
50	0,25	0,79	1,18	92,60	0,53	1,53	93,18
	0,30	0,79	1,18	92,60	0,61	1,43	93,44
	0,50	0,79	1,18	92,60	0,77	1,25	94,08
	1,00	0,79	1,18	92,60	0,88	1,12	94,26
	2,00	0,79	1,18	92,58	0,94	1,06	94,40
100	0,25	0,85	1,13	94,02	0,67	1,36	94,32
	0,30	0,85	1,13	94,04	0,72	1,30	94,46
	0,50	0,85	1,13	94,04	0,83	1,18	94,76
	1,00	0,85	1,13	94,04	0,92	1,09	94,74
	2,00	0,85	1,13	94,04	0,96	1,04	94,84

Capítulo 3

A distribuição senh-normal

3.1 Introdução

Neste capítulo exploraremos a distribuição senh-normal (SN) e suas propriedades, entre elas a que estabelece uma importante relação entre esta e a distribuição BSG.

Desde que foi desenvolvida, a distribuição normal, desempenhou um papel dominante no meio estatístico. Contudo, ela pode não fornecer uma representação adequada a muitas situações práticas. Segundo Johnson (1949), historicamente foi necessário e conveniente considerar construções de distribuições não-normais por uma transformação de uma variável aleatória com distribuição normal padrão. Essa técnica foi introduzida por Edgeworth (1898) e denominada por ele de método de translação. Johnson(1949) usou tal método para gerar três famílias de distribuições, as quais podiam assumir uma vasta variedade de formas, sendo geradas por transformações do tipo

$$Z = \nu + \delta f(Y; \gamma, \sigma), \quad (3.1)$$

em que Z é uma variável normal padrão, $f(y; \gamma, \sigma)$ é uma função monótona, ν e δ são parâmetros de forma, σ é parâmetro de escala e γ é parâmetro de locação. Mais especificamente, as três famílias de distribuições são

1. O sistema log-normal S_L

$$Z = \nu + \delta \log \left(\frac{Y - \gamma}{\sigma} \right);$$

2. A extensão limitada do sistema log-normal S_B

$$Z = \nu + \delta \log \left(\frac{Y - \gamma}{\sigma + \gamma - Y} \right), \gamma < Y < \gamma + \sigma;$$

3. O sistema não limitado S_U

$$Z = \nu + \delta \sinh^{-1} \left(\frac{Y - \gamma}{\sigma} \right).$$

Neste trabalho, vamos considerar uma família de distribuições gerada pela transformação

$$f(Y; \gamma, \sigma) = \sinh \left(\frac{Y - \gamma}{\sigma} \right),$$

com δ substituído por $2/\alpha$ em (3.1).

Esse modelo é chamado *senh-normal* e foi desenvolvido por Rieck & Nedelman (1991). Abaixo, discorremos um pouco sobre esta distribuição, mostrando algumas de suas propriedades e apresentando sua relação com as distribuições BS e BSG, dadas em (1.2) e (2.1), respectivamente.

Definição 3.1 *Seja Y uma variável aleatória tal que $Z = \nu + (2/\alpha)\sinh[(Y - \gamma)\sigma]$ tem distribuição normal padrão. Então, Y segue uma distribuição SN com parâmetro de não centralidade ν , parâmetro de forma $\alpha > 0$, parâmetro de locação $\gamma \in \mathbb{R}$ e parâmetro de escala $\sigma > 0$. Se $\nu = 0$, então dizemos que Y segue a distribuição SN centrada.*

Notação 3.2 *Denotamos a distribuição SN não centrada por $SN(\alpha, \gamma, \sigma, \nu)$ e a distribuição SN centrada por $SN(\alpha, \gamma, \sigma)$.*

A fda de uma variável aleatória Y que segue a distribuição SN não central é obtida a partir de

$$Z = \nu + \frac{2}{\alpha} \sinh \left(\frac{Y - \gamma}{\sigma} \right).$$

Ou seja,

$$Y = \sigma \operatorname{arcsinh} \left[\frac{(Z - \nu)\alpha}{2} \right] + \gamma.$$

Daí,

$$\begin{aligned} F_Y(y) = P(Y \leq y) &= P \left(\sigma \operatorname{arcsinh} \left[\frac{(Z - \nu)\alpha}{2} \right] + \gamma \leq y \right) \\ &= P \left(Z \leq \nu + \frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right) \\ &= \Phi \left\{ \nu + \frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right\}. \end{aligned}$$

Então a fda de Y é dada por

$$F_Y(y) = \Phi \left\{ \nu + \frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right\}, \quad y \in \mathbb{R}. \quad (3.2)$$

O parâmetro ν influencia, principalmente, a assimetria da distribuição em torno da média e o parâmetro α interfere, principalmente, na curtose, como podemos ver nas Figuras (3.1), (3.2) e (3.3), em que são apresentados gráficos das funções densidade da distribuição SN. A Figura (3.1) representa o caso não centrado para vários valores de ν , onde percebemos a existência de assimetria nos casos em que $\nu \neq 0$. Já nas Figuras (3.2) e (3.3), estamos considerando $\nu = 0$, logo as curvas são simétricas em torno do parâmetro γ . Em ambos os casos obtivemos curvas mais achatadas à medida que aumentamos os valores de α , indicando assim que este parâmetro influencia a curtose da distribuição. Comparando as Figuras (3.2) e (3.3), notamos que o parâmetro γ é de locação e em cada uma delas observamos que a amplitude das curvas diminui a medida que o parâmetro σ aumenta, levando-nos a perceber que este é o parâmetro de escala da distribuição.

A fdp é obtida derivando-se (3.2) e é dada por

$$f(y) = \frac{2}{\alpha\sigma\sqrt{2\pi}} \cosh\left(\frac{y-\gamma}{\sigma}\right) \exp\left\{-\frac{1}{2}\left[\nu + \frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right]^2\right\}, y \in \mathbb{R}, \quad (3.3)$$

com $\alpha > 0$, $\sigma > 0$, $\nu \in \mathbb{R}$ e $\gamma \in \mathbb{R}$.

De agora em diante trabalharemos com a distribuição SN centrada.

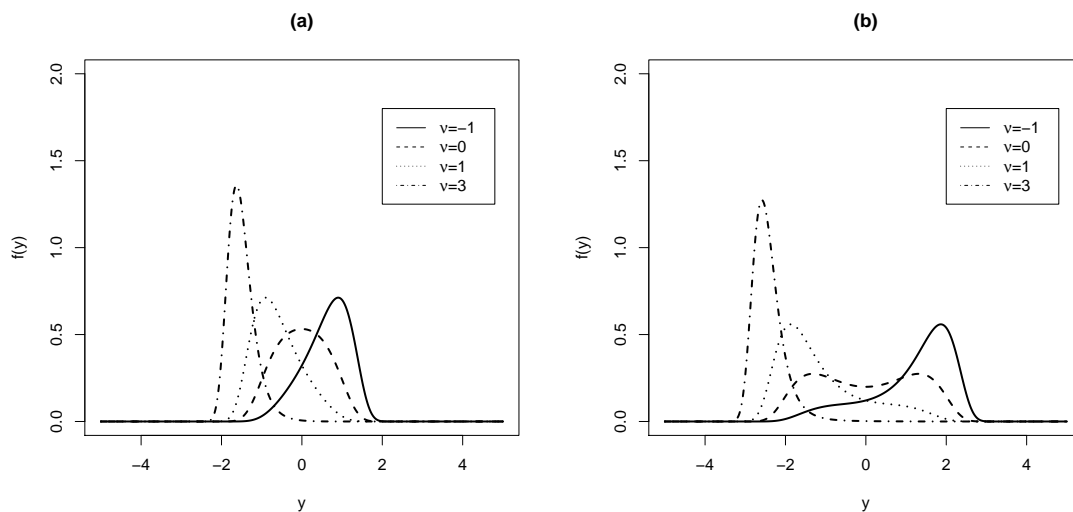


Figura 3.1: gráficos de funções densidade da distribuição SN não centrada com parâmetros $\gamma = 0$, $\sigma = 1$ e $\alpha = 1,5$ (a), e $\gamma = 0$ e $\sigma = 1$ e $\alpha = 4$ (b).

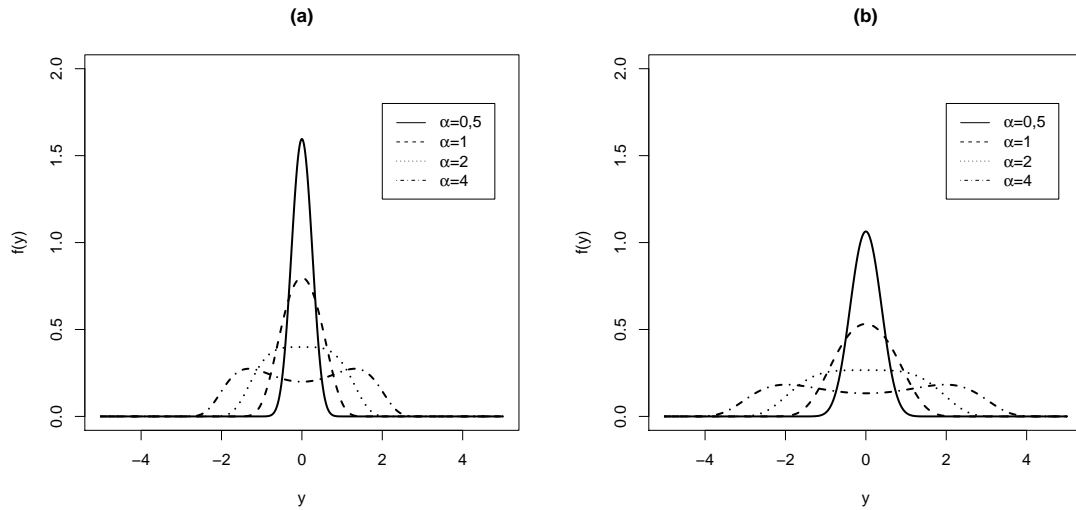


Figura 3.2: gráficos de funções densidade da distribuição SN centrada com parâmetros $\gamma = 0$ e $\sigma = 1$ (a), e $\gamma = 0$ e $\sigma = 1,5$ (b).

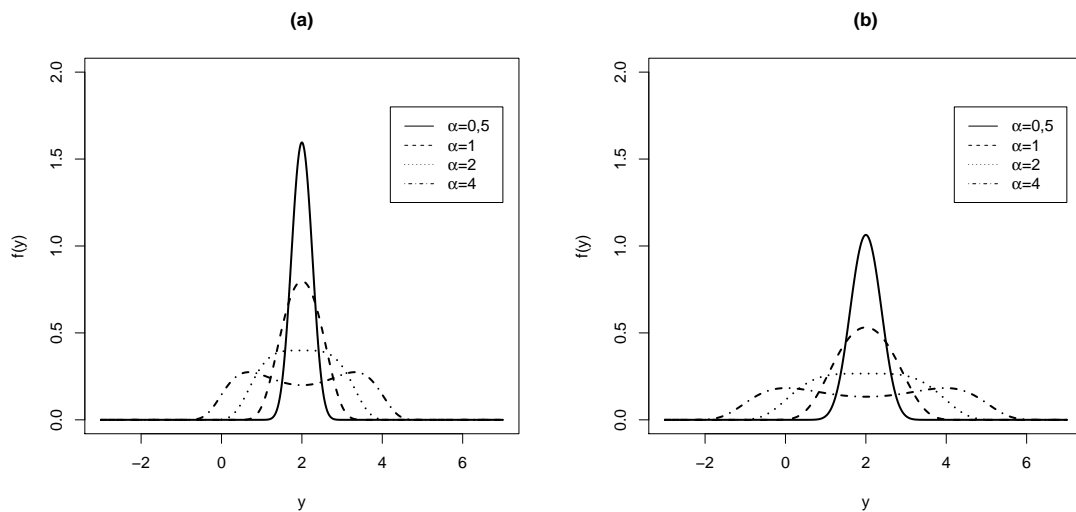


Figura 3.3: gráficos de funções densidade da distribuição SN centrada com parâmetros $\gamma = 2$ e $\sigma = 1$ (a), e $\gamma = 2$ e $\sigma = 1,5$ (b).

3.2 Propriedades da distribuição senh-normal

A propriedade a seguir nos permite examinar qual é o efeito de uma transformação linear sobre a distribuição SN.

Propriedade 3.1 *Seja $Y \sim \text{SN}(\alpha, \gamma, \sigma)$, então $a + bY \sim \text{SN}(\alpha, a + b\gamma, |b|\sigma)$.*

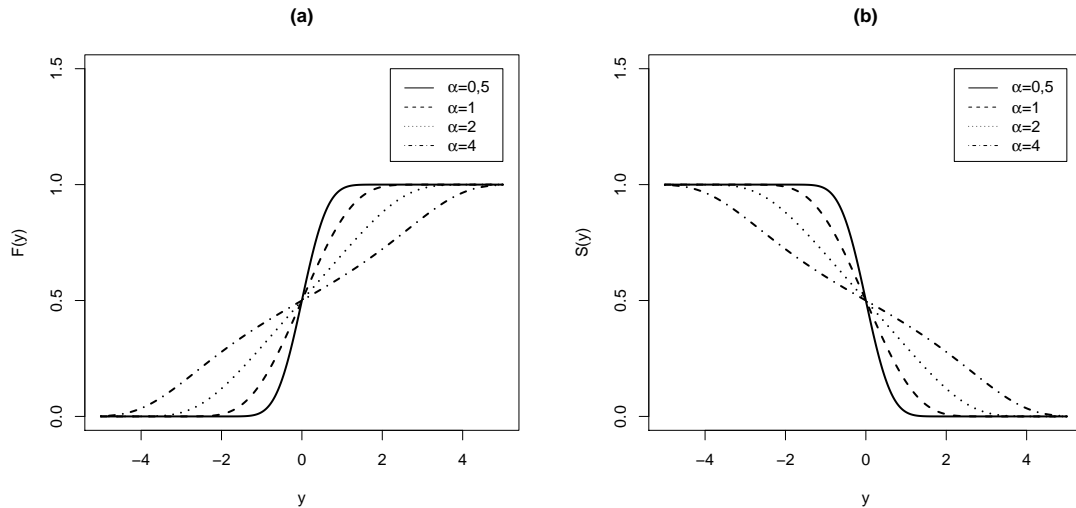


Figura 3.4: gráficos de funções de distribuição acumulada da distribuição SN centrada (a) e gráficos de funções de sobrevivência da distribuição senh-normal central (b), ambos com parâmetros $\gamma = 0$ e $\sigma = 2$.

Prova.

Se $Y \sim \text{SN}(\alpha, \gamma, \sigma)$, então

$$F(y) = \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right\}.$$

1. Se $b > 0$, temos

$$\begin{aligned} F_{a+bY}(y) = P(a + bY \leq y) &= P\left(Y \leq \frac{y - a}{b}\right) \\ &= F_Y\left(\frac{y - a}{b}\right) \\ &= \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{y - (a + b\gamma)}{b\sigma} \right) \right\}, \end{aligned}$$

ou seja,

$$a + bY \sim \text{SN}(\alpha, a + b\gamma, b\sigma), b > 0. \quad (3.4)$$

2. Se $b \leq 0$,

$$\begin{aligned}
 F_{a+bY}(y) = P(a + bY \leq y) &= P\left(Y \geq \frac{y-a}{b}\right) \\
 &= 1 - \Phi\left\{\frac{2}{\alpha}\operatorname{senh}\left(\frac{y-(a+b\gamma)}{b\sigma}\right)\right\} \\
 &= \Phi\left\{-\frac{2}{\alpha}\operatorname{senh}\left(\frac{y-(a+b\gamma)}{b\sigma}\right)\right\} \\
 &= \Phi\left\{\frac{2}{\alpha}\operatorname{senh}\left(\frac{y-(a+b\gamma)}{-b\sigma}\right)\right\},
 \end{aligned}$$

ou seja,

$$a + bY \sim \text{SN}(\alpha, a + b\gamma, -b\sigma), b \leq 0. \quad (3.5)$$

Logo, por (3.4) e (3.5), temos

$$a + bY \sim \text{SN}(\alpha, a + b\gamma, |b|\sigma), b \leq 0.$$

■

Em particular, $Y \sim \text{SN}(\alpha, \gamma, \sigma) \Rightarrow \frac{Y}{\sigma} \sim \text{SN}(\alpha, \gamma/\sigma, 1)$.

Propriedade 3.2 *O parâmetro γ é a mediana da distribuição SN centrada.*

Prova.

Temos

$$F_Y(\gamma) = \Phi\left\{\frac{2}{\alpha}\operatorname{senh}\left(\frac{\gamma-\gamma}{\sigma}\right)\right\} = \Phi(0) = \frac{1}{2}.$$

■

Abaixo apresentamos uma condição necessária e suficiente para que a fda da distribuição SN seja simétrica com relação a γ .

Proposição 3.1 *Se $Y \sim \text{SN}(\alpha, \gamma, \sigma, \nu)$, então Y é simétrica com relação a γ se, e somente se, $\nu = 0$.*

Prova. Suponhamos que Y seja simétrica com relação a γ , então a fda F_Y é tal que

$$F_Y(\gamma + y) = 1 - F_Y(\gamma - y),$$

de onde obtemos

$$F_{Y-\gamma}(y) = F_{-(Y-\gamma)}(y).$$

Chamando $U = Y - \gamma$ e $V = -(Y - \gamma)$, temos

$$F_U(y) = F_V(y). \quad (3.6)$$

Note que

$$\begin{aligned} F_U(y) = P(U \leq y) &= P(Y \leq y + \gamma) \\ &= \Phi \left\{ \nu + \frac{2}{\alpha} \sinh \left(\frac{y}{\sigma} \right) \right\} \end{aligned}$$

e

$$\begin{aligned} F_V(y) = P(V \leq y) &= 1 - P(Y \leq \gamma - y) \\ &= 1 - \Phi \left\{ \nu + \frac{2}{\alpha} \sinh \left(-\frac{y}{\sigma} \right) \right\} \\ &= \Phi \left\{ -\nu - \frac{2}{\alpha} \sinh \left(-\frac{y}{\sigma} \right) \right\} \\ &= \Phi \left\{ -\nu + \frac{2}{\alpha} \sinh \left(\frac{y}{\sigma} \right) \right\} \end{aligned}$$

A última igualdade se deve ao fato da função \sinh ser ímpar.

De (3.6), temos

$$\Phi \left\{ \nu + \frac{2}{\alpha} \sinh \left(\frac{y}{\sigma} \right) \right\} = \Phi \left\{ -\nu + \frac{2}{\alpha} \sinh \left(\frac{y}{\sigma} \right) \right\},$$

ou seja,

$$F(y; \alpha, 0, \sigma, \nu) = F(y; \alpha, 0, \sigma, -\nu), \forall y.$$

Assim, devemos ter $\nu = -\nu$, logo $\nu = 0$.

Reciprocamente, suponha que $\nu = 0$. Então,

$$F_U(y) = \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{y}{\sigma} \right) \right\} = \Phi \left\{ -\frac{2}{\alpha} \sinh \left(-\frac{y}{\sigma} \right) \right\} = F_V(y),$$

ou seja, Y é simétrica com relação a γ . ■

Proposição 3.2 *Seja $Y \sim \text{SN}(\alpha, \gamma, \sigma)$ e $S = (Y - \gamma)/(\alpha\sigma/2)$. Então, $\lim_{\alpha \rightarrow 0} S = Z$, em que $Z \sim \text{N}(0, 1)$.*

Prova. A fda de S é dada por

$$P(S \leq s) = P\left(\frac{Y - \gamma}{\frac{1}{2}\alpha\sigma} \leq s\right) = P\left(Y \leq \frac{1}{2}\alpha\sigma s + \gamma\right) = \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{1}{2}\alpha s \right) \right\}.$$

Daí,

$$\lim_{\alpha \rightarrow 0} P(S \leq s) = \lim_{\alpha \rightarrow 0} \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{1}{2} \alpha s \right) \right\}.$$

Observemos que

$$\lim_{\alpha \rightarrow 0} \frac{2}{\alpha} \sinh \left(\frac{1}{2} \alpha s \right) = \lim_{\alpha \rightarrow 0} \frac{\sinh \left(\frac{1}{2} \alpha s \right)}{\frac{\alpha}{2}} = \lim_{\alpha \rightarrow 0} \frac{\frac{1}{2} s \cosh \left(\frac{1}{2} \alpha s \right)}{\frac{1}{2}} = s, \quad (3.7)$$

onde utilizamos a regra de L'Hôpital. Pela continuidade da fda da distribuição normal padrão e usando (3.7), temos

$$\lim_{\alpha \rightarrow 0} P(S \leq s) = \lim_{\alpha \rightarrow 0} \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{1}{2} \alpha s \right) \right\} = \Phi(s).$$

Portanto, concluímos que

$$S \xrightarrow{D} Z,$$

em que $Z \sim N(0, 1)$ e \xrightarrow{D} denota convergência em distribuição. ■

Veremos agora alguns conceitos necessários ao entendimento dos resultados que seguem.

Teorema 3.3 *Uma função duas vezes diferenciável $g(y)$ é côncava se, e somente se, sua segunda derivada é menor ou igual a zero para todo y . Se $g''(y)$ é estritamente menor do que zero para todo y , então $g(y)$ é estritamente côncava. Contudo, a recíproca não é válida. Se $g(y)$ é estritamente côncava, então podemos dizer que $g''(y) \leq 0$, para todo $y \in \mathbb{R}$.*

Definição 3.4 *Uma distribuição com fdp $f(y)$ é chamada fortemente unimodal se $g(y) = \log[f(y)]$ é côncava. Uma função é chamada estritamente fortemente unimodal se $g(y) = \log[f(y)]$ é estritamente côncava.*

Observação 3.1 *Se uma fdp é estritamente fortemente unimodal então existe uma única solução para a equação de verossimilhança em relação ao parâmetro de locação (Lehmann e Casella, 1998).*

A distribuição SN é muito flexível, podendo assumir tanto a forma unimodal quanto a forma bimodal. A seguir apresentaremos algumas condições sobre os parâmetros, que determinam a modalidade da distribuição.

Lema 3.1 *A distribuição SN é unimodal para $\alpha \leq 2$ e bimodal para $\alpha > 2$.*

Prova.

Sabemos que os pontos críticos de $f(y)$ ocorrem nos pontos y tais que $f'(y) = 0$.

Temos

$$f(y) = \frac{2}{\alpha\sigma\sqrt{2\pi}} \cosh\left(\frac{y-\gamma}{\sigma}\right) \exp\left\{-\frac{1}{2}\left[\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right]^2\right\}.$$

Seja $A = -\frac{1}{2}\left[\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right]^2$ e $x = \frac{y-\gamma}{\sigma}$. Segue que

$$\begin{aligned} f'(y) &= \frac{2}{\alpha\sigma^2\sqrt{2\pi}} \sinh(x) \exp(A) + \frac{2}{\alpha\sigma^2\sqrt{2\pi}} \cosh(x) \exp\left\{A\left[-\frac{2}{\alpha}\sinh(x)\right]\frac{2}{\alpha}\cosh(x)\right\} \\ &= \frac{2}{\alpha\sigma^2\sqrt{2\pi}} \left\{ \sinh(x) \exp\left[A\left(1 - \frac{4}{\alpha^2}\cosh^2(x)\right)\right] \right\}. \end{aligned}$$

Note que $f'(y) = 0$ somente quando $\sinh(x)[1 - (4/\alpha^2)\cosh^2(x)] = 0$.

Se $\sinh(x) = 0$, nada podemos afirmar acerca de α . Supondo $\sinh(x) \neq 0$, temos

$$\begin{aligned} f'(y) = 0 &\Leftrightarrow 1 - \frac{4}{\alpha^2}\cosh^2(x) = 0 \\ &\Leftrightarrow \cosh(x) = \frac{\alpha}{2}. \end{aligned}$$

1. Para $\alpha \leq 2$, temos

$$\cosh(x) = \frac{\alpha}{2} \leq 1.$$

Como $\cosh(x) \geq 1$, para todo x , temos

$$\cosh(x) = 1$$

ou seja,

$$x = \frac{y-\gamma}{\sigma} = 0$$

e, conseqüentemente,

$$y = \gamma.$$

2. Para $\alpha > 2$, temos

$$\begin{aligned} \cosh(x) &= \frac{\exp(x) + \exp(-x)}{2} = \frac{\alpha}{2} \\ \Rightarrow \frac{\exp(2x) + 1}{\exp(x)} &= \alpha \\ \Rightarrow \exp(2x) - \alpha \exp(x) + 1 &= 0. \end{aligned}$$

Chamando $y = \exp(x)$, temos

$$y^2 - \alpha y + 1 = 0.$$

Logo,

$$y = \frac{\alpha \pm \sqrt{\alpha^2 - 4}}{2}.$$

Assim,

$$x = \log \left(\frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2 - 4}{4}} \right).$$

Afim de verificar se os pontos obtidos acima são de máximo local aplicamos o teste da segunda derivada. Daí temos

$$\begin{aligned} f''(y) &= \frac{2}{\sigma^2 \alpha \sqrt{2\pi}} \left[\frac{1}{\sigma} \cosh(x) \exp(A) \right] \\ &\quad + \frac{2}{\sigma^2 \alpha \sqrt{2\pi}} \sinh(x) \exp \left\{ A \left[-\frac{4}{\sigma \alpha^2} \sinh(x) \cosh(x) \right] \right\} \left[1 - \frac{4}{\alpha^2} \cosh^2(x) \right] \\ &= \frac{2}{\sigma^2 \alpha \sqrt{2\pi}} \sinh(x) \exp \left\{ A \left[-\frac{8}{\sigma \alpha^2} \sinh(x) \cosh(x) \right] \right\}. \end{aligned}$$

Substituindo $y = \gamma$, obtemos

$$f''(\gamma) = \frac{2}{\sigma^2 \alpha \sqrt{2\pi}} \frac{1}{\sigma} \left(1 - \frac{4}{\alpha^2} \right) = \frac{2}{\sigma^2 \alpha \sqrt{2\pi}} \frac{(\alpha^2 - 4)}{\alpha^2}.$$

Note que

$$f''(\gamma) \leq 0 \Leftrightarrow \alpha \leq 2.$$

Logo, a fdp possui um único ponto de máximo em $y = \gamma$, para $\alpha \leq 2$. Portanto a distribuição SN é unimodal para $\alpha \leq 2$ ■

Agora estabeleceremos condições sobre os parâmetros, as quais determinarão se a distribuição é ou não é fortemente unimodal.

Lema 3.2 *A distribuição SN é fortemente unimodal se, e somente se, $\alpha \leq 2$.*

Prova. Uma distribuição é fortemente unimodal se $g(y) = \log[f(y)]$ é côncava.

Para provar que $g(y)$ é côncava para $\alpha \leq 2$, mostraremos que $g''(y) \leq 0$ para todo y .

Temos

$$g(y) = \log[f(y)] = \log \left(\frac{2}{\alpha \sigma \sqrt{2\pi}} \right) + \log \left[\cosh \left(\frac{y - \gamma}{\sigma} \right) \right] - \frac{2}{\alpha^2} \sinh^2 \left(\frac{y - \gamma}{\sigma} \right).$$

Derivando $g(y)$ com relação a y , obtemos

$$g'(y) = \frac{1}{\sigma} \tanh \left(\frac{y - \gamma}{\sigma} \right) - \frac{4}{\alpha^2 \sigma} \sinh \left(\frac{y - \gamma}{\sigma} \right) \cosh \left(\frac{y - \gamma}{\sigma} \right).$$

A segunda derivada de $g(y)$ em relação a y é

$$\begin{aligned}
 g''(y) &= \frac{1}{\sigma^2} \operatorname{sech}^2(x) - \frac{4}{\alpha^2 \sigma^2} \cosh^2(x) - \frac{4}{\alpha^2 \sigma^2} \sinh^2(x) \\
 &= \frac{1}{\sigma^2} \left\{ \operatorname{sech}^2(x) - \frac{4}{\alpha^2} [1 + \sinh^2(x) + \sinh^2(x)] \right\} \\
 &= \frac{1}{\sigma^2} \left[\operatorname{sech}^2(x) - \frac{4}{\alpha^2} 1 - \frac{8}{\alpha^2} \sinh^2(x) \right] \\
 &= \frac{1}{\sigma^2} \left\{ \operatorname{sech}^2(x) \left[1 - \frac{4}{\alpha^2} \cosh^2(x) \right] - \frac{8}{\alpha^2} \sinh^2(x) \right\},
 \end{aligned}$$

em que, $x = (y - \gamma)/\sigma$.

Supondo $\alpha \leq 2$, temos

$$1 - \frac{4}{\alpha^2} \cosh^2(x) \leq 1 - \cosh^2(x) = -\sinh^2(x) \leq 0$$

e

$$-\frac{8}{\alpha^2} \sinh^2(x) \leq -2\sinh^2(x) \leq 0,$$

já que $\sinh^2(x) \geq 0, \forall x \in \mathbb{R}$.

Assim, $g''(y) \leq 0, \forall y$ e então a distribuição é fortemente unimodal.

Reciprocamente, suponhamos $\alpha > 2$. Para que a função g seja côncava deveríamos ter $g''(y) \leq 0$. No entanto, para $x = 0$, temos

$$1 - \frac{4}{\alpha^2} \cosh^2(x) = 1 - \frac{4}{\alpha^2} > 0$$

e

$$-\frac{8}{\alpha^2} \sinh^2(x) = 0.$$

Daí,

$$g''(y) \geq 0,$$

logo g não é côncava e, conseqüentemente, a distribuição não é fortemente unimodal.

Portanto, se g é fortemente unimodal, então $\alpha \leq 2$. ■

Uma importante relação entre as distribuições BS e SN foi mostrada por Rieck (1989). Adaptamos tal resultado sob o contexto da distribuição BSG.

Teorema 3.5 *Se a variável aleatória X tem distribuição $BSG(\alpha, \beta, \lambda)$, então $Y = \log(X)$ tem distribuição $SN(\alpha, \log(\beta), 1/\lambda)$.*

Prova. Seja $X \sim \text{BSG}(\alpha, \beta, \lambda)$ e seja $Y = \log(X)$. Temos

$$\begin{aligned} F_Y(y) &= F_X(\exp(y)) \\ &= \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{\exp(y)}{\beta} \right)^\lambda - \left(\frac{\beta}{\exp(y)} \right)^\lambda \right] \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} [\exp \{ \lambda(y - \ln \beta) \} - \exp \{ \lambda(\ln \beta - y) \}] \right\} \\ &= \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{y - \ln(\beta)}{1/\lambda} \right) \right\}, \end{aligned}$$

ou seja,

$$Y \sim \text{SN}(\alpha, \log(\beta), 1/\lambda).$$

■

Substituindo λ por $1/2$ na demonstração acima, chegamos ao seguinte resultado.

Corolário 3.6 *Se a variável aleatória X tem distribuição $\text{BS}(\alpha, \beta)$, então $Y = \log(X)$ tem distribuição $\text{SN}(\alpha, \log(\beta), 2)$.*

3.3 Momentos

Agora vamos obter a função geradora de momentos para a distribuição SN. Usaremos expansões assintóticas da função geradora para obter aproximações dos momentos.

Teorema 3.7 *A função geradora de momentos para a distribuição SN é*

$$m(t) = \exp(\gamma t) \left[\frac{K(a, \alpha^{-2}) + K(b, \alpha^{-2})}{2K(1/2, \alpha^{-2})} \right], \quad (3.8)$$

em que $a = (\sigma t + 1)/2$, $b = (\sigma t - 1)/2$ e K é a função de Bessel modificada do terceiro tipo.

Prova. Primeiramente, considere a variável aleatória S com distribuição $\text{SN}(\alpha, 0, 1)$. Então, sua fdp é dada por

$$\begin{aligned} f(s) &= \frac{2}{\alpha \sqrt{2\pi}} \cosh(s) \exp \left\{ -\frac{2}{\alpha^2} \sinh^2(s) \right\} \\ &= \frac{2}{\alpha \sqrt{2\pi}} \left[\frac{\exp(s) + \exp(-s)}{2} \right] \exp \left\{ -\frac{2}{\alpha^2} \left[\frac{\exp(s) - \exp(-s)}{2} \right]^2 \right\} \\ &= \frac{2}{\alpha \sqrt{2\pi}} \left[\frac{\exp(s) + \exp(-s)}{2} \right] \exp \left\{ -\frac{2}{\alpha^2} \left[\frac{\exp(2s) + \exp(-2s) - 2}{4} \right] \right\} \\ &= \left[\frac{\exp(s) + \exp(-s)}{\alpha \sqrt{2\pi}} \right] \exp \{ -\alpha^{-2} \cosh(2s) \} \exp(\alpha^{-2}). \end{aligned}$$

A função geradora de momentos pode ser escrita como

$$\begin{aligned} M_S(t) &= E[\exp(tS)] = \int_{-\infty}^{\infty} \frac{\exp(s) + \exp(-s)}{\alpha\sqrt{2\pi}} \exp[ts - \alpha^{-2} \cosh(2s)] \exp(\alpha^{-2}) ds \\ &= \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}} \int_{-\infty}^{\infty} [\exp(ts + s) + \exp(ts - s)] \exp[-\alpha^{-2} \cosh(2s)] ds. \end{aligned}$$

Usando a mudança de variável $x = 2s$, temos

$$\begin{aligned} M_S(t) &= \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{(t+1)}{2}x - \alpha^{-2} \cosh(x)\right] dx \\ &\quad + \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{(t-1)}{2}x - \alpha^{-2} \cosh(x)\right] dx. \end{aligned}$$

Pela Equação (2.4), podemos escrever

$$\frac{1}{2} \int_{-\infty}^{\infty} \exp\left[\frac{(t+1)}{2}x - \alpha^{-2} \cosh(x)\right] dx = K\left(\frac{t+1}{2}, \alpha^{-2}\right)$$

e

$$\frac{1}{2} \int_{-\infty}^{\infty} \exp\left[\frac{(t-1)}{2}x - \alpha^{-2} \cosh(x)\right] dx = K\left(\frac{t-1}{2}, \alpha^{-2}\right).$$

Numericamente, obtemos

$$K(1/2, \alpha^{-2}) = \frac{\alpha\sqrt{2\pi} \exp(-\alpha^{-2})}{2}.$$

Logo, obtemos

$$M_S(t) = \frac{K(a, \alpha^{-2}) + K(b, \alpha^{-2})}{2K(1/2, \alpha^{-2})},$$

em que $a = (t+1)/2$, $b = (t-1)/2$ e K é a função de Bessel modificada do terceiro tipo.

Agora, seja $Y = \gamma + \sigma S$, com $\sigma > 0$. Temos, pela Propriedade (3.1), que

$$Y \sim \text{SN}(\alpha, \gamma, \sigma), \alpha > 0.$$

Daí,

$$\begin{aligned} M_Y(t) &= M_{\gamma + \sigma S}(t) = E(\exp(\gamma t + \sigma S t)) = \exp(\gamma t) M_S(\sigma t) \\ &= \exp(\gamma t) \frac{K(a, \alpha^{-2}) + K(b, \alpha^{-2})}{2K(1/2, \alpha^{-2})}, \end{aligned}$$

em que $a = (\sigma t + 1)/2$ e $b = (\sigma t - 1)/2$. ■

Observação 3.2 *Ja foi provado que se $X \sim \text{BSG}(\alpha, \beta, \lambda)$, então $Y = \log(X) \sim \text{SN}(\alpha, \log(\beta), 1/\lambda)$. Logo, pelo resultado acima, a função geratriz de momentos de Y , dada por (3.8), com $a = (t + \lambda)/(2\lambda)$, $b = (t - \lambda)/(2\lambda)$ e $\gamma = \log(\beta)$ é*

$$M_Y(t) = E(\exp(tY)) = E(\exp(t \ln X)) = E(\exp(\ln X^t)) = E(X^t),$$

ou seja, o t -ésimo momento da distribuição BS é $M_Y(t)$ com os parâmetros acima dados.

Os momentos são obtidos derivando-se a função geradora de momentos em relação a t e em seguida fixando $t = 0$. No entanto, para esta distribuição não é uma tarefa fácil obter expressões fechadas para as derivadas. Abaixo, apresentamos expressões assintóticas para a função geradora de momentos quando α é pequeno e quando α é grande.

Pequenos valores de α correspondem a grande valores do argumento α^{-2} da função de Bessel acima. Para um valor alto do argumento, Abramowitz e Stegun (1972, p.378) forneceram a seguinte expansão assintótica da função de Bessel do terceiro tipo:

$$K(v, z) = \exp(-z) \sqrt{\frac{\pi}{2z}} \left\{ 1 + \frac{\mu - 1}{8z} + \frac{(\mu - 1)(\mu - 9)}{2!(8z)^2} + \dots \right\}, \quad (3.9)$$

em que $\mu = 4v^2$.

Considerando a variável aleatória $S \sim \text{SN}(\alpha, 0, 1)$ e usando os três primeiros termos da expansão, temos

1. $K(1/2, \alpha^{-2})$

Se $v = 1/2$, então $\mu = 1$, daí

$$K(1/2, \alpha^{-2}) = \exp(-\alpha^{-2}) \sqrt{\frac{\pi}{2\alpha^{-2}}} = \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})}.$$

2. $K(a, \alpha^{-2})$

Se $v = (t + 1)/2$, então $\mu = t^2 + 2t + 1$. Temos

$$\begin{aligned} K(a, \alpha^{-2}) &= \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})} \left\{ 1 + \frac{(t^2 + 2t)}{8}\alpha^2 + \frac{(t^2 + 2t)(t^2 + 2t - 8)}{128}\alpha^4 \right\} \\ &= \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})} \left\{ 1 + \frac{t^2\alpha^2 + 2t\alpha^2}{8} + \frac{t^4 + 4t^3 - 4t^2 - 16t}{128}\alpha^4 \right\}. \end{aligned}$$

3. $K(b, \alpha^{-2})$

Se $v = (t + 1)/2$, então $\mu = t^2 - 2t + 1$. Temos

$$\begin{aligned} K(b, \alpha^{-2}) &= \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})} \left\{ 1 + \frac{(t^2 - 2t)}{8}\alpha^2 + \frac{(t^2 - 2t)(t^2 - 2t - 8)}{128}\alpha^4 \right\} \\ &= \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})} \left\{ 1 + \frac{t^2\alpha^2 - 2t\alpha^2}{8} + \frac{t^4 - 4t^3 - 4t^2 + 16t}{128}\alpha^4 \right\}. \end{aligned}$$

Logo,

$$K(a, \alpha^{-2}) + K(b, \alpha^{-2}) = \frac{\sqrt{\frac{\pi}{2}}\alpha}{\exp(\alpha^{-2})} \left\{ 2 + \frac{2t^2\alpha^2}{8} + \frac{2t^4 - 8t^2}{128}\alpha^4 \right\}.$$

Assim,

$$\frac{K(a, \alpha^{-2}) + K(b, \alpha^{-2})}{2K(1/2, \alpha^{-2})} = 1 + \frac{t^2\alpha^2}{8} + \frac{t^4 - 4t^2}{128}\alpha^4.$$

Daí, para $Y = \gamma + \sigma S \sim \text{SN}(\alpha, \gamma, \sigma)$, temos que a função geratriz de momentos fica dada por

$$M_Y(t) = \exp(\gamma t) \left\{ 1 + \frac{\alpha^2\sigma^2 t^2}{8} + \frac{\sigma^4 t^4 - 4\sigma^2 t^2}{128}\alpha^4 \right\}.$$

Então

$$\begin{aligned} M'_Y(t) &= \gamma \exp(\gamma t) \left\{ 1 + \frac{\alpha^2\sigma^2 t^2}{8} + \frac{\sigma^4 t^4 - 4\sigma^2 t^2}{128}\alpha^4 \right\} \\ &\quad + \exp(\gamma t) \left\{ \frac{\alpha^2\sigma^2 t}{4} + \frac{4\sigma^4 t^3 - 8\sigma^2 t}{128}\alpha^4 \right\}. \end{aligned}$$

Substituindo $t = 0$ na equação acima, obtemos

$$M'_Y(0) = \gamma$$

A segunda derivada de $M_Y(t)$ é dada por

$$\begin{aligned} M''_Y(t) &= \gamma^2 \exp(\gamma t) \left\{ 1 + \frac{\alpha^2\sigma^2 t^2}{8} + \frac{\sigma^4 t^4 - 4\sigma^2 t^2}{128}\alpha^4 \right\} \\ &\quad + 2\gamma \exp(\gamma t) \left\{ \frac{\alpha^2\sigma^2 t}{4} + \frac{4\sigma^4 t^3 - 8\sigma^2 t}{128}\alpha^4 \right\} \\ &\quad + \exp(\gamma t) \left\{ \frac{\alpha^2\sigma^2}{4} + \frac{12\sigma^4 t^2 - 8\sigma^2}{128}\alpha^4 \right\}. \end{aligned}$$

Substituindo $t = 0$, obtemos

$$E(Y^2) = M''(0) = \gamma^2 + \frac{\alpha^2\sigma^2}{4} - \frac{\alpha^4\sigma^2}{16}.$$

Portanto, com esta aproximação, a média e a variância da distribuição SN são, respectivamente

$$E(Y) = \gamma; \tag{3.10}$$

$$\text{Var}(Y) = \frac{\alpha^2\sigma^2}{4} - \frac{\alpha^4\sigma^2}{16}. \tag{3.11}$$

Segundo Abramowitz e Stegun (1972), para grandes valores de α , os momentos podem ser obtidos usando a aproximação

$$K(a, \alpha^{-2}) = \Gamma(a)2^{a-1}\alpha^{2a}, a > 0.$$

Usando esta expansão assintótica, temos que a função geradora de momentos é

$$M_Y(t) = \frac{1}{2\sqrt{\pi}} \exp(t\gamma) \left\{ \Gamma\left(\frac{\sigma t + 1}{2}\right) (\alpha\sqrt{2})^{\sigma t} + \Gamma\left(\frac{1 - \sigma t}{2}\right) (\alpha\sqrt{2})^{-\sigma t} \right\}, |t| < \gamma. \quad (3.12)$$

Com essa aproximação, temos

$$\begin{aligned} M'_Y(t) &= \frac{1}{2\sqrt{\pi}} \gamma \exp(t\gamma) \left\{ \Gamma\left(\frac{\sigma t + 1}{2}\right) (\alpha\sqrt{2})^{\sigma t} + \Gamma\left(\frac{1 - \sigma t}{2}\right) (\alpha\sqrt{2})^{-\sigma t} \right\} \\ &+ \frac{1}{2\sqrt{\pi}} \exp(t\gamma) \left\{ \Gamma'\left(\frac{\sigma t + 1}{2}\right) \frac{\sigma}{2} (\alpha\sqrt{2})^{\sigma t} + \Gamma\left(\frac{\sigma t + 1}{2}\right) \sigma (\alpha\sqrt{2})^{\sigma t} \log(\alpha\sqrt{2}) \right\} \\ &- \frac{1}{2\sqrt{\pi}} \exp(t\gamma) \left\{ \Gamma'\left(\frac{1 - \sigma t}{2}\right) \frac{\sigma}{2} (\alpha\sqrt{2})^{-\sigma t} + \Gamma\left(\frac{1 - \sigma t}{2}\right) \sigma (\alpha\sqrt{2})^{-\sigma t} \log(\alpha\sqrt{2}) \right\}. \end{aligned}$$

Substituindo $t = 0$, temos

$$M'_Y(0) = \frac{\gamma}{2\sqrt{\pi}} \left\{ \Gamma\left(\frac{1}{2}\right) + \Gamma\left(\frac{1}{2}\right) \right\} = \frac{\gamma}{2\sqrt{\pi}} \{2\sqrt{\pi}\} = \gamma.$$

A segunda derivada de $M_Y(t)$, para a aproximação dada em (3.12), é dada por

$$\begin{aligned} M''_Y(t) &= \frac{1}{2\sqrt{\pi}} \gamma^2 \exp(t\gamma) \left\{ \Gamma\left(\frac{\sigma t + 1}{2}\right) (\alpha\sqrt{2})^{\sigma t} + \Gamma\left(\frac{1 - \sigma t}{2}\right) (\alpha\sqrt{2})^{-\sigma t} \right\} \\ &+ \frac{\gamma \sigma (\alpha\sqrt{2})^{\sigma t} \exp(t\gamma)}{\sqrt{\pi}} \left\{ \Gamma'\left(\frac{\sigma t + 1}{2}\right) \frac{1}{2} + \Gamma\left(\frac{\sigma t + 1}{2}\right) \log(\alpha\sqrt{2}) \right\} \\ &- \frac{\gamma \sigma (\alpha\sqrt{2})^{-\sigma t} \exp(t\gamma)}{\sqrt{\pi}} \left\{ \Gamma'\left(\frac{1 - \sigma t}{2}\right) \frac{1}{2} + \Gamma\left(\frac{1 - \sigma t}{2}\right) \log(\alpha\sqrt{2}) \right\} \\ &+ \frac{\sigma^2 (\alpha\sqrt{2})^{\sigma t} \exp(t\gamma)}{4\sqrt{\pi}} \left\{ \Gamma''\left(\frac{\sigma t + 1}{2}\right) \frac{1}{2} + \Gamma'\left(\frac{\sigma t + 1}{2}\right) \log(\alpha\sqrt{2}) \right\} \\ &+ \frac{\sigma^2 (\alpha\sqrt{2})^{\sigma t} \log(\alpha\sqrt{2}) \exp(t\gamma)}{2\sqrt{\pi}} \left\{ \Gamma'\left(\frac{\sigma t + 1}{2}\right) \frac{1}{2} + \Gamma\left(\frac{\sigma t + 1}{2}\right) \log(\alpha\sqrt{2}) \right\} \\ &+ \frac{\sigma^2 (\alpha\sqrt{2})^{-\sigma t} \exp(t\gamma)}{4\sqrt{\pi}} \left\{ \Gamma''\left(\frac{1 - \sigma t}{2}\right) \frac{1}{2} + \Gamma'\left(\frac{1 - \sigma t}{2}\right) \log(\alpha\sqrt{2}) \right\} \\ &+ \frac{\sigma^2 (\alpha\sqrt{2})^{-\sigma t} \log^2(\alpha\sqrt{2}) \exp(t\gamma)}{2\sqrt{\pi}} \left\{ \Gamma'\left(\frac{1 - \sigma t}{2}\right) \frac{1}{2} + \Gamma\left(\frac{1 - \sigma t}{2}\right) \log(\alpha\sqrt{2}) \right\}. \end{aligned}$$

Substituindo $t = 0$ na expressão acima, temos

$$M''_Y(0) = \gamma^2 + \frac{\sigma^2}{4\sqrt{\pi}} \Gamma''\left(\frac{1}{2}\right) + \frac{\sigma^2}{\sqrt{\pi}} \Gamma'\left(\frac{1}{2}\right) \log(\alpha\sqrt{2}) + \frac{\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) \log^2(\alpha\sqrt{2}).$$

Através de métodos numéricos, obtemos

$$M_Y''(0) = \gamma^2 + \sigma^2 \{2,197543455 - 1,963510026 \log(\alpha\sqrt{2}) + [\log(\alpha\sqrt{2})]^2\}.$$

Logo,

$$E(Y) = \gamma$$

$$\text{Var}(Y) = \sigma^2 \{2,197543455 - 1,963510026 \log(\alpha\sqrt{2}) + [\log(\alpha\sqrt{2})]^2\}.$$

Capítulo 4

Um modelo linear para a distribuição log-Birnbaum-Saunders generalizada

4.1 Introdução

Neste capítulo vamos considerar um modelo de regressão linear para a distribuição log-BSG, ou seja, vamos considerar o caso em que uma variável com distribuição log-BSG possa ser explicada através de uma estrutura com covariáveis e parâmetros desconhecidos. Mais precisamente, faremos um breve esboço do CDM, obteremos uma aproximação para as estimativas, bem como, apresentaremos as medidas: distância de Cook generalizada e afastamento pela verossimilhança usando tal método, com o objetivo de identificar observações influentes. Depois obteremos uma estatística do teste escore para detectar *outliers* baseando-se no MSOM. Ainda nesse capítulo, discutiremos o fato dos parâmetros de forma e escala no modelo de regressão log-BSG não serem necessariamente constantes e obteremos um teste de heterogeneidade para esses parâmetros utilizando a estatística da razão de verossimilhanças.

4.1.1 O modelo de regressão log-linear BSG

O modelo de regressão log-linear BSG tomado por base nesse estudo é dado por

$$Y_i = \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n, \quad (4.1)$$

em que os Y_i 's são os logaritmos dos tempos, cada \mathbf{X}_i contém os valores das variáveis explicativas de dimensão p referente ao indivíduo i , $\boldsymbol{\beta}$ é o vetor de parâmetros desconhecido a ser estimado e os erros aleatórios ϵ_i 's são independentes e identicamente distribuídos de acordo com a distribuição $\text{SN}(\alpha_i, 0, 1/\lambda_i)$. Assumimos que os parâmetros de forma α_i 's e os parâmetros de escala λ_i 's são desconhecidos e que $\alpha_i = \alpha$ e $\lambda_i = \lambda$, $\forall i = 1, \dots, n$.

Assim, o logaritmo da função de verossimilhança fica dado por

$$\begin{aligned} l(\boldsymbol{\theta}) &= n \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{i=1}^n \log \left\{ \frac{2}{\alpha} \cosh[\lambda(y_i - \mu_i)] \right\} - \frac{1}{2} \sum_{i=1}^n \left\{ \frac{2}{\alpha} \sinh[\lambda(y_i - \mu_i)] \right\}^2 \\ &= n \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2, \end{aligned}$$

em que $\xi_{i1} = (2/\alpha) \cosh[\lambda(y_i - \mu_i)]$ e $\xi_{i2} = (2/\alpha) \sinh[\lambda(y_i - \mu_i)]$.

Derivando o logaritmo da função de verossimilhança em relação a α , λ e $\beta_j (j = 1, \dots, p)$ obtemos as funções escore, que são dadas, respectivamente, por

$$\begin{aligned} U_\alpha = \frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} &= \sum_{i=1}^n \left\{ -\frac{1}{\alpha} + \frac{4}{\alpha^3} \sinh^2[\lambda(y_i - \mu_i)] \right\} \\ &= \frac{1}{\alpha} \sum_{i=1}^n (\xi_{i2}^2 - 1); \end{aligned} \quad (4.2)$$

$$\begin{aligned} U_\lambda = \frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} &= \frac{n}{\lambda} + \sum_{i=1}^n \left\{ \frac{\sinh[\lambda(y_i - \mu_i)]}{\cosh[\lambda(y_i - \mu_i)]} - \frac{4}{\alpha^2} \sinh[\lambda(y_i - \mu_i)] \cosh[\lambda(y_i - \mu_i)] \right\} (y_i - \mu_i) \\ &= \frac{n}{\lambda} + \sum_{i=1}^n \left(\frac{\xi_{i2}}{\xi_{i1}} - \xi_{i1} \xi_{i2} \right) (y_i - \mu_i); \end{aligned} \quad (4.3)$$

$$\begin{aligned} U_{\beta_j} = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j} &= \lambda \sum_{i=1}^n \left\{ \frac{4}{\alpha^2} \sinh[\lambda(y_i - \mu_i)] \cosh[\lambda(y_i - \mu_i)] - \frac{\sinh[\lambda(y_i - \mu_i)]}{\cosh[\lambda(y_i - \mu_i)]} \right\} x_{ij} \\ &= \lambda \sum_{i=1}^n \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) x_{ij}, \end{aligned} \quad (4.4)$$

para cada $j = 1, \dots, p$.

A estimativa de MV de α^2 é obtido igualando a expressão (4.2) a zero e resolvendo a equação de verossimilhança. Ou seja,

$$\begin{aligned} U_\alpha = 0 &\Rightarrow \frac{1}{\hat{\alpha}} \sum_{i=1}^n \left[\frac{4}{\hat{\alpha}^2} \sinh^2 \hat{\lambda}(y_i - \hat{\mu}_i) - 1 \right] = 0 \\ &\Rightarrow \sum_{i=1}^n \frac{4}{\hat{\alpha}^2} \sinh^2 \hat{\lambda}(y_i - \hat{\mu}_i) = n \\ &\Rightarrow \hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2 \hat{\lambda}(y_i - \hat{\mu}_i). \end{aligned}$$

As equações $\partial l(\boldsymbol{\theta})/\partial \lambda = 0$ e $\partial l(\boldsymbol{\theta})/\partial \beta_j = 0$, para $j = 1, \dots, p$ não possuem soluções analíticas, sendo necessária a utilização de métodos numéricos, tais como Newton-Rapson,

Escore de Fisher, ou um algoritmo quase-Newton (BFGS), para a obtenção das raízes. Vale salientar que, nesse caso, não temos problemas de estimação.

A inferência assintótica para o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \lambda, \alpha)^\top$ pode ser baseada na normalidade assintótica do estimador de máxima verossimilhança, dada por

$$\hat{\boldsymbol{\theta}} \sim N_{(p+1)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_\theta),$$

em que $\boldsymbol{\Sigma}_\theta$ é a matriz de variâncias e covariâncias de $\hat{\boldsymbol{\theta}}$, que pode ser aproximada pela matriz de informação observada $[-\ddot{l}(\boldsymbol{\theta})]^{-1}$, avaliada em $\hat{\boldsymbol{\theta}}$, obtida a partir de

$$\begin{aligned} \ddot{l}(\hat{\boldsymbol{\theta}}) = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} &= \left[\begin{array}{ccc} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \alpha} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \alpha} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} \end{array} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ &= \left[\begin{array}{ccc} \lambda^2 X^\top V X & X^\top \mathbf{g} & -2\lambda X^\top \mathbf{h} \\ \mathbf{g}^\top X & \gamma & 2\mathbf{h}^\top (\mathbf{Y} - \boldsymbol{\mu}) \\ -2\lambda \mathbf{h}^\top X & 2(\mathbf{Y}^\top - \boldsymbol{\mu}^\top) \mathbf{h} & \delta \end{array} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \end{aligned}$$

em que

$$\begin{aligned} \hat{\delta} &= \frac{n}{\hat{\alpha}^2} - \frac{12}{\hat{\alpha}^4} \sum_{i=1}^n \sinh^2 [\hat{\lambda}(y_i - \hat{\mu}_i)]; \\ \hat{\gamma} &= -\frac{n}{\hat{\lambda}^2} + \sum_{i=1}^n \left\{ \operatorname{sech}^2[\hat{\lambda}(y_i - \hat{\mu}_i)] - \frac{4}{\hat{\alpha}^2} \cosh[2\hat{\lambda}(y_i - \hat{\mu}_i)] \right\} (y_i - \hat{\mu}_i)^2; \\ \hat{\mathbf{h}}^\top &= (\hat{h}_1, \hat{h}_2, \dots, \hat{h}_n), \text{ com } \hat{h}_i = \frac{2}{\hat{\alpha}^3} \sinh[2\hat{\lambda}(y_i - \hat{\mu}_i)]; \hat{\mathbf{g}}^\top = (\hat{g}_1, \hat{g}_2, \dots, \hat{g}_n), \text{ com} \\ \hat{g}_i &= \left\{ \frac{4}{\hat{\alpha}^2} \cosh[2\hat{\lambda}(y_i - \hat{\mu}_i)] - \operatorname{sech}^2[\hat{\lambda}(y_i - \hat{\mu}_i)] \right\} \hat{\lambda}(y_i - \hat{\mu}_i) + \hat{\alpha} \hat{h}_i - \tanh[\hat{\lambda}(y_i - \hat{\mu}_i)] \text{ e} \\ \hat{V} &= \operatorname{diag}(\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n), \text{ com } \hat{v}_i = \operatorname{sech}^2[\hat{\lambda}(y_i - \hat{\mu}_i)] - \frac{4}{\hat{\alpha}^2} \cosh[2\hat{\lambda}(y_i - \hat{\mu}_i)]. \end{aligned}$$

No Apêndice B encontram-se cálculos mais detalhados para a obtenção da matriz $\ddot{l}(\hat{\boldsymbol{\theta}})$.

4.2 Análise de diagnóstico

Fundamentaremos esta seção no estudo feito por Xie e Wei (2007), adaptando-o para o modelo de regressão log-linear BSG.

4.2.1 Diagnósticos de influência baseados no CDM

Deleção de casos é uma aproximação comumente usada para estudar o efeito da retirada da observação i do conjunto de dados. O CDM para (4.1) é dado por

$$Y_j = \mathbf{X}_j^\top \boldsymbol{\beta} + \epsilon_j, \quad j = 1, 2, \dots, n, \quad j \neq i, \quad (4.5)$$

o qual representa o modelo original sem a observação i . Para este modelo, o logaritmo da função de verossimilhança de $\boldsymbol{\theta}$ é denotado por $l_{(i)}(\boldsymbol{\theta})$ e é dado por

$$l_{(i)}(\boldsymbol{\theta}) = (n-1) \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{j=1}^n \log \xi_{j1} - \frac{1}{2} \sum_{j=1}^n \xi_{j2}^2 - \log \xi_{i1} + \frac{1}{2} \xi_{i2}^2.$$

Seja $\hat{\boldsymbol{\theta}}_{(i)} = (\hat{\boldsymbol{\beta}}_{(i)}^\top, \hat{\lambda}_{(i)}, \hat{\alpha}_{(i)})^\top$ o estimador de MV de $\boldsymbol{\theta}$ baseado em $l_{(i)}(\boldsymbol{\theta})$. Para avaliar a influência da observação i na estimativa de MV de $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$, deve-se comparar a diferença entre $\hat{\boldsymbol{\theta}}_{(i)}$ e $\hat{\boldsymbol{\theta}}$. Se a deleção de um caso influenciar, seriamente, as estimativas dos parâmetros do modelo, mais atenção deveria ser dada ao caso deletado. Logo, se $\hat{\boldsymbol{\theta}}_{(i)}$ é distante de $\hat{\boldsymbol{\theta}}$, então o caso i é considerado uma observação influente.

Observe que é necessário obter $\hat{\boldsymbol{\theta}}_{(i)}$ para todos os casos, implicando em um processo computacional pesado, principalmente quando n é muito grande. Por isso, a seguinte aproximação de $\hat{\boldsymbol{\theta}}_{(i)}$ é geralmente usada para reduzir tal processo (ver Cook e Weisberg, 1982, p. 182):

$$\hat{\boldsymbol{\theta}}_{(i)}^1 = \hat{\boldsymbol{\theta}} + [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}), \quad (4.6)$$

em que

$$\dot{l}_{(i)}(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad \text{e} \quad \ddot{l}(\hat{\boldsymbol{\theta}}) = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.$$

Teorema 4.1 *Para o modelo de regressão log-linear BSG em (4.5), a relação entre os estimadores dos parâmetros no modelo dos dados completos e no modelo com o caso i deletado pode ser expressa por*

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{(i)}^1 &= \hat{\boldsymbol{\beta}} + \hat{\lambda}(X^\top \hat{A}X)^{-1} \mathbf{X}_i \hat{r}_i - \hat{\mathbf{e}} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{\mathbf{e}} \hat{\lambda}^{-1} + \hat{\mathbf{f}} \hat{\eta}_i; \\ \hat{\lambda}_{(i)}^1 &= \hat{\lambda} + \hat{\lambda} \hat{\mathbf{i}}^\top \mathbf{X}_i \hat{r}_i - \hat{\mathbf{a}} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{\mathbf{a}} \hat{\lambda}^{-1} + \hat{\mathbf{b}} \hat{\eta}_i; \\ \hat{\alpha}_{(i)}^1 &= \hat{\alpha} + \hat{\lambda} \hat{\mathbf{j}}^\top \mathbf{X}_i \hat{r}_i - \hat{\mathbf{c}} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{\mathbf{c}} \hat{\lambda}^{-1} + \hat{\mathbf{d}} \hat{\eta}_i, \end{aligned}$$

em que

$$\begin{aligned} \hat{A} &= \hat{\lambda}^{-2} \hat{V} - c_1 - c_2 - c_3 - c_4, \text{ com } c_1 = \frac{\hat{\delta} \hat{\mathbf{g}} \hat{\mathbf{g}}^\top}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2}, c_2 = \frac{2 \hat{\kappa} \hat{\lambda} \hat{\mathbf{h}} \hat{\mathbf{g}}^\top}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2}, c_3 = \frac{2 \hat{\kappa} \hat{\lambda} \hat{\mathbf{g}} \hat{\mathbf{h}}^\top}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2} \text{ e } c_4 = \frac{4 \hat{\lambda}^2 \hat{\gamma} \hat{\mathbf{h}} \hat{\mathbf{h}}^\top}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2}; \\ \hat{\mathbf{a}} &= \hat{\tau}^{-1} (\hat{\delta} - 4 \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}}); \quad \hat{\mathbf{b}} = \hat{\tau}^{-1} (-\hat{\kappa} - 2 \hat{\lambda}^{-1} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}}); \\ \hat{\mathbf{c}} &= \hat{\tau}^{-1} (-\hat{\kappa} - 2 \hat{\lambda}^{-1} \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}}); \quad \hat{\mathbf{d}} = \hat{\tau}^{-1} (\hat{\gamma} - \hat{\lambda}^{-2} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}}); \\ \hat{\tau} &= (\hat{\gamma} - \hat{\lambda}^{-2} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}}) (\hat{\delta} - 4 \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}}) \\ &\quad - (\hat{\kappa} + 2 \hat{\lambda}^{-1} \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}})^2; \\ \hat{\kappa} &= 2 \hat{\mathbf{h}}^\top (\mathbf{Y} - \hat{\boldsymbol{\mu}}); \quad \mathbf{Y} = (y_1, y_2, \dots, y_n)^\top, \quad \hat{\boldsymbol{\mu}} = X \hat{\boldsymbol{\beta}}; \quad \hat{r}_i = \hat{\xi}_{i1} \hat{\xi}_{i2} - \frac{\hat{\xi}_{i2}}{\hat{\xi}_{i1}}; \quad \hat{\eta}_i = \frac{1}{\hat{\alpha}} \hat{\xi}_{i2}^2 - \frac{1}{\hat{\alpha}}; \\ \hat{\mathbf{e}} &= \hat{\lambda}^{-2} (X^\top \hat{V} X)^{-1} (2 \hat{\lambda} X^\top \hat{\mathbf{h}} \hat{\mathbf{c}} - X^\top \hat{\mathbf{g}} \hat{\mathbf{a}}); \quad \hat{\mathbf{f}} = \hat{\lambda}^{-2} (X^\top \hat{V} X)^{-1} (2 \hat{\lambda} X^\top \hat{\mathbf{h}} \hat{\mathbf{d}} - X^\top \hat{\mathbf{g}} \hat{\mathbf{b}}); \\ \hat{\mathbf{i}}^\top &= \frac{1}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2} (\hat{\delta} \hat{\mathbf{g}}^\top X + 2 \hat{\kappa} \hat{\lambda} \hat{\mathbf{h}}^\top X) (X^\top \hat{A} X)^{-1}; \quad \hat{\mathbf{j}}^\top = -\frac{1}{\hat{\delta} \hat{\gamma} - \hat{\kappa}^2} (\hat{\kappa} \hat{\mathbf{g}}^\top X + 2 \hat{\gamma} \hat{\lambda} \hat{\mathbf{h}}^\top X) (X^\top \hat{A} X)^{-1}. \end{aligned}$$

Prova. Ver Apêndice B. ■

Através do Teorema 4.1 podemos ver diferenças entre as estimativas com e sem o caso deletado e obter medidas de diagnóstico, baseadas em deleção de casos, para avaliar as observações influentes em modelos de regressão log-linear BSG.

4.2.2 Distância de Cook generalizada

A distância de Cook generalizada (GCD) é definida como a norma padronizada de $\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}$ dada por

$$\text{GCD} = (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}})^\top M (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}), \quad (4.7)$$

em que M é uma matriz definida não negativa, a qual mede a combinação ponderada dos elementos para a diferença $\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}$. Cook e Weisberg (1982) consideraram várias escolhas para M . Uma escolha comumente usada é a informação observada $M = -\ddot{l}(\hat{\boldsymbol{\theta}})$. Substituindo $\hat{\boldsymbol{\theta}}_{(i)}^1$, dado em (4.6), e $M = -\ddot{l}(\hat{\boldsymbol{\theta}})$ em (4.7), obtemos

$$\begin{aligned} \text{GCD} &= \{[-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}})\}^\top [-\ddot{l}(\hat{\boldsymbol{\theta}})] [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}) \\ &= [\dot{l}_{(i)}(\hat{\boldsymbol{\theta}})]^\top \{[-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1}\}^\top \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}) \\ &= [\dot{l}_{(i)}(\hat{\boldsymbol{\theta}})]^\top [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}). \end{aligned} \quad (4.8)$$

Multiplicando $[\dot{l}_{(i)}(\hat{\boldsymbol{\theta}})]^\top$ por $[-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}})$, onde estas quantidades já foram obtidas na prova do Teorema 4.1, obtemos

$$\begin{aligned} \text{GCD} &= \begin{bmatrix} -\hat{\lambda} \mathbf{X}_i \hat{r}_i \\ \hat{r}_i (y_i - \hat{\mu}_i) - \hat{\lambda}^{-1} \\ -\hat{\eta}_i \end{bmatrix}^\top \begin{bmatrix} \hat{\lambda} (X^\top \hat{A} X)^{-1} \mathbf{X}_i \hat{r}_i - \hat{\mathbf{e}} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{\mathbf{e}} \hat{\lambda}^{-1} + \hat{\mathbf{f}} \hat{\eta}_i \\ \hat{\lambda} \hat{\mathbf{i}}^\top \mathbf{X}_i \hat{r}_i - \hat{a} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{a} \hat{\lambda}^{-1} + \hat{b} \hat{\eta}_i \\ \hat{\lambda} \hat{\mathbf{j}}^\top \mathbf{X}_i \hat{r}_i - \hat{c} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{c} \hat{\lambda}^{-1} + \hat{d} \hat{\eta}_i \end{bmatrix} \\ &= \{-\hat{\lambda} \mathbf{X}_i \hat{r}_i [\hat{\lambda} (X^\top \hat{A} X)^{-1} \mathbf{X}_i \hat{r}_i - \hat{\mathbf{e}} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{\mathbf{e}} \hat{\lambda}^{-1} + \hat{\mathbf{f}} \hat{\eta}_i] \\ &\quad + [\hat{r}_i (y_i - \hat{\mu}_i) - \hat{\lambda}^{-1}] [\hat{\lambda} \hat{\mathbf{i}}^\top \mathbf{X}_i \hat{r}_i - \hat{a} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{a} \hat{\lambda}^{-1} + \hat{b} \hat{\eta}_i] \\ &\quad - \hat{\eta}_i [\hat{\lambda} \hat{\mathbf{j}}^\top \mathbf{X}_i \hat{r}_i - \hat{c} \hat{r}_i (y_i - \hat{\mu}_i) + \hat{c} \hat{\lambda}^{-1} + \hat{d} \hat{\eta}_i]\} \end{aligned}$$

em que \hat{r}_i , $\hat{\eta}_i$, \hat{A} , \hat{a} , \hat{b} , \hat{c} , \hat{d} , $\hat{\mathbf{e}}$, $\hat{\mathbf{f}}$, $\hat{\mathbf{i}}$ e $\hat{\mathbf{j}}$ são dados no Teorema 4.1.

É frequente considerar a influência do caso i nas estimativas dos parâmetros α , λ ou $\boldsymbol{\beta}$. Então, por (4.7), a GCD para cada parâmetro é definida como

$$\text{GCD}(\boldsymbol{\beta}) = \{\dot{l}_{(i)\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})^\top \{(I_p, 0, 0) [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} (I_p, 0, 0)^\top\} \dot{l}_{(i)\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\}_{\hat{\boldsymbol{\theta}}},$$

$$\text{GCD}(\lambda) = \{\dot{l}_{(i)\lambda}(\hat{\boldsymbol{\theta}})^\top \{(0_p, 1, 0) [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} (0_p, 1, 0)^\top\} \dot{l}_{(i)\lambda}(\hat{\boldsymbol{\theta}})\}_{\hat{\boldsymbol{\theta}}},$$

$$\text{GCD}(\alpha) = \{\dot{l}_{(i)\alpha}(\hat{\boldsymbol{\theta}})^\top \{(0_p, 0, 1) [-\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} (0_p, 0, 1)^\top\} \dot{l}_{(i)\alpha}(\hat{\boldsymbol{\theta}})\}_{\hat{\boldsymbol{\theta}}},$$

sendo

$$\dot{l}_{(i)}\boldsymbol{\beta}(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}; \quad \dot{l}_{(i)}\lambda(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \lambda} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}; \quad \dot{l}_{(i)}\alpha(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \alpha} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}};$$

I_p a matriz identidade de ordem p ;

0_p a matriz nula de ordem p ;

$0 = (0, \dots, 0)^\top$ e $1 = (1, \dots, 1)^\top$.

Portanto,

$$\begin{aligned} \text{GCD}(\boldsymbol{\beta}) &= -\hat{\lambda}^2 \mathbf{X}_i^\top \hat{r}_i^2 (X^\top \hat{A} X)^{-1} \mathbf{X}_i; \\ \text{GCD}(\lambda) &= -[\hat{r}_i (y_i - \hat{\mu}_i) - \hat{\lambda}^{-1}]^2 \hat{a}; \\ \text{GCD}(\alpha) &= -\hat{\eta}^2 \hat{d}. \end{aligned}$$

Os valores de $\text{GCD}(\boldsymbol{\beta})$, $\text{GCD}(\lambda)$ e $\text{GCD}(\alpha)$ revelam o impacto do caso i nas estimativas de $\boldsymbol{\beta}$, λ e α , respectivamente.

4.2.3 Afastamento pela verossimilhança

Outra medida conhecida para avaliar a diferença entre $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{(i)}$ é o afastamento pela verossimilhança (Cook e Weisberg, 1982) dada por

$$\text{LD} = 2[l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{(i)})]. \quad (4.9)$$

Podemos aplicar diretamente o Teorema 4.1 para calcular a medida em (4.9). Substituindo (4.6) na expressão (4.9), obtemos a aproximação

$$\text{LD} = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}} + \{-\ddot{l}(\hat{\boldsymbol{\theta}})\}^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}))\}.$$

Além disso, podemos calcular também $\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}}_{j(i)}$, para $j = 1, \dots, p$ para obter a diferença entre $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\beta}}_{(i)}$, assim como $\hat{\lambda} - \hat{\lambda}_{(i)}$ e $\hat{\alpha} - \hat{\alpha}_{(i)}$.

4.2.4 MSOM e teste para outlier

De acordo com Xie e Wei (2007), o CDM é a base para a construção de estatísticas de diagnóstico efetivas. Outro modelo de diagnóstico comumente usado é o modelo de outlier MSOM, definido como

$$\begin{cases} Y_j = X_j^\top \boldsymbol{\beta} + \epsilon_j; & j = 1, \dots, n; \quad j \neq i \\ Y_i = X_i^\top \boldsymbol{\beta} + \nu + \epsilon_i, \end{cases} \quad (4.10)$$

em que ν é um parâmetro extra usado para indicar a presença de um *outlier* (Cook e Weisberg, 1982, p.20). Obviamente, se o valor de ν é diferente de zero, então o caso i pode ser um *outlier*, pois não é proveniente do modelo original (4.1).

A partir do MSOM em (4.10), podemos considerar um teste de hipóteses para identificar os *outliers*. Considere as hipóteses

$$H_0 : \nu = 0 \quad \text{contra} \quad H_1 : \nu \neq 0. \quad (4.11)$$

Se H_0 for rejeitada, então o caso i é um possível *outlier*. O teorema a seguir nos fornece uma estatística escore para este teste.

Teorema 4.2 *Para o MSOM (4.10), a estatística escore para testar as hipóteses dadas em (4.11) é*

$$SC_i = \{r_i^2 \rho_i^{-1}\}_{\hat{\theta}}, \quad i = 1, \dots, n,$$

em que

$$\begin{aligned} \rho_i = & [u_i \mathbf{X}_i^\top (X^\top AX)^{-1} - \frac{u_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})}{\lambda} \mathbf{i}^\top + q_i \mathbf{j}^\top] u_i \mathbf{X}_i - [u_i \mathbf{X}_i^\top \mathbf{e} - \frac{u_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})}{\lambda} a + q_i c] \frac{u_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})}{\lambda} \\ & + [u_i \mathbf{X}_i^\top \mathbf{f} - \frac{u_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})}{\lambda} b + q_i d] q_i - u_i, \quad q_i = -\frac{2}{\alpha} \lambda (\xi_{i1} \xi_{i2}) \boldsymbol{\theta}, \quad u_i = \frac{4\lambda^2}{\alpha^2 \xi_{i1}^2} - (\xi_{i1}^2 + \xi_{i2}^2) \lambda^2 \quad e \\ & r_i, A, \mathbf{i}, \mathbf{j}, \mathbf{e}, \mathbf{f}, a, b, c, d \text{ são dados no Teorema 4.1.} \end{aligned}$$

Prova.

Ver Apêndice C. ■

4.2.5 Teste de heterogeneidade dos parâmetros

No caso do modelo de regressão log-linear BS, Rieck e Nedelman (1991) assumiram que o parâmetro de forma α é independente do vetor explicativo \mathbf{X}_i , isto é, $\alpha_i = \alpha$ para $i = 1, \dots, n$. Galea et al. (2004) também mencionaram esse problema e assumiram a homogeneidade do parâmetro de forma em suas discussões. Contudo, essa suposição precisa ser verificada. Neste sentido, Xie e Wei (2007) propuseram um teste de hipóteses para verificar a suposição de homogeneidade do parâmetro de forma. Para o modelo log-linear BSG, também consideraremos testes dessa natureza tanto para o parâmetro de forma quanto para o parâmetro de escala, e para ambos.

I- Teste de heterogeneidade do parâmetro de forma

Suponhamos que o modelo homogêneo (4.1) seja modificado da forma

$$\begin{cases} Y_i = \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, & i = 1, \dots, n \\ \epsilon_i \sim SN(\alpha_i, 0, 1/\lambda), & \alpha_i = \alpha m_i, \quad m_i = m(\mathbf{X}_i, \boldsymbol{\omega}), \end{cases} \quad (4.12)$$

em que α é o fator do parâmetro de forma com uma função peso $m(\mathbf{X}_i, \boldsymbol{\omega})$, e $\boldsymbol{\omega}$ é um vetor paramétrico de dimensão q para indicar a heterogeneidade do parâmetro de forma. Estamos assumindo que existe um único vetor $\boldsymbol{\omega}_0$ de $\boldsymbol{\omega}$ tal que $m(\mathbf{X}_i, \boldsymbol{\omega}_0) = 1$ para todo i . Assim, o teste de heterogeneidade para o parâmetro de forma é equivalente testar as hipóteses

$$H_0 : \boldsymbol{\omega} = \boldsymbol{\omega}_0 \quad \text{contra} \quad H_1 : \boldsymbol{\omega} \neq \boldsymbol{\omega}_0.$$

De (4.12), podemos obter a estatística razão de verossimilhanças para o teste acima. Note que, neste caso, $\boldsymbol{\omega}$ é o parâmetro de interesse e $\boldsymbol{\beta}$, λ e α são os parâmetros de perturbação.

O logaritmo da função de verossimilhança para o modelo (4.12) fica dada por

$$l(\boldsymbol{\theta}) = n \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2,$$

em que agora

$$\xi_{i1} = (2/\alpha m_i) \cosh[\lambda(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})] \quad \text{e} \quad \xi_{i2} = (2/\alpha m_i) \sinh[\lambda(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})], \quad \text{com} \quad \boldsymbol{\theta} = (\boldsymbol{\omega}^\top, \boldsymbol{\beta}^\top, \lambda, \alpha)^\top.$$

Seja $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$; $\tilde{m}_i = m(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$ sob H_0 . Então, temos o seguinte teorema.

Teorema 4.3 *Para o modelo (4.12), a estatística da razão de verossimilhanças é dada por*

$$LR = n \log \left(\frac{\tilde{\lambda}^2 \hat{\alpha}^2}{\hat{\lambda}^2 \tilde{\alpha}^2} m_g \right), \quad m_g = \left[\prod_{i=1}^n \frac{\cosh[\tilde{\lambda}(y_i - \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}})]}{\cosh[\hat{\lambda}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})]} \tilde{m}_i^{-1} \right]^{\frac{2}{n}}$$

Prova. Ver Apêndice D. ■

II- Teste de heterogeneidade do parâmetro de escala

Agora, suponhamos que o modelo (4.1) é modificado da seguinte forma

$$\begin{cases} Y_i = \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, & i = 1, \dots, n \\ \epsilon_i \sim SN(\alpha, 0, 1/\lambda_i), & \lambda_i = \lambda n_i, \quad n_i = n(\mathbf{X}_i, \boldsymbol{\omega}), \end{cases} \quad (4.13)$$

em que λ é o fator do parâmetro de escala com uma função peso $n(\mathbf{X}_i, \boldsymbol{\omega})$, e $\boldsymbol{\omega}$ é um vetor paramétrico de dimensão q para indicar a heterogeneidade do parâmetro de escala. De forma análoga ao caso I queremos testar as hipóteses

$$H_0 : \boldsymbol{\omega} = \boldsymbol{\omega}_0 \quad \text{contra} \quad H_1 : \boldsymbol{\omega} \neq \boldsymbol{\omega}_0.$$

Neste caso, o logaritmo da função de verossimilhança para o modelo (4.13) fica dada por

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log \left(\frac{\lambda n_i}{\sqrt{2\pi}} \right) + \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2,$$

em que

$$\xi_{i1} = (2/\alpha) \cosh[\lambda n_i (y_i - \mathbf{X}_i^\top \boldsymbol{\beta})] \quad \text{e} \quad \xi_{i2} = (2/\alpha) \sinh[\lambda n_i (y_i - \mathbf{X}_i^\top \boldsymbol{\beta})], \quad \text{com} \quad \boldsymbol{\theta} = (\boldsymbol{\omega}^\top, \boldsymbol{\beta}^\top, \lambda, \alpha)^\top.$$

Seja $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$; $\tilde{n}_i = n(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$ sob H_0 . Então, temos o seguinte teorema.

Teorema 4.4 *Para o modelo (4.13), a estatística da razão de verossimilhanças é dada por*

$$\text{LR} = \left\{ \log \left[\prod_{i=1}^n \left(\frac{\tilde{\lambda} \tilde{n}_i^2 \hat{\alpha}^2}{\hat{\lambda}^2 \hat{\alpha}^2} m_{g_i} \right) \right] \right\}, \quad m_{g_i} = \left[\frac{\cosh[\tilde{\lambda} \tilde{n}_i (y_i - \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}})]}{\cosh[\hat{\lambda} (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})]} \right]^2.$$

Prova. Ver Apêndice E. ■

III- Teste de heterogeneidade dos parâmetros de forma e de escala simultaneamente

Para este caso, consideremos o seguinte modelo

$$\begin{cases} Y_i = \mathbf{X}_i^\top \boldsymbol{\beta} + \epsilon_i, & i = 1, \dots, n \\ \epsilon_i \sim SN(\alpha_i, 0, 1/\lambda_i), & \alpha_i = \alpha m_i, \lambda_i = \lambda n_i, m_i = m(\mathbf{X}_i, \boldsymbol{\omega}) \text{ e } n_i = n(\mathbf{X}_i, \boldsymbol{\vartheta}), \end{cases} \quad (4.14)$$

em que α e λ são os fatores dos parâmetros de forma e escala, com funções peso $m(\mathbf{X}_i, \boldsymbol{\omega})$ e $n(\mathbf{X}_i, \boldsymbol{\vartheta})$, respectivamente. Os vetores $\boldsymbol{\omega}$ e $\boldsymbol{\vartheta}$ de dimensão q indicam a heterogeneidade dos parâmetros de forma e escala, respectivamente. Neste caso queremos testar as hipóteses

$$H_0 : \boldsymbol{\omega} = \boldsymbol{\omega}_0; \boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0 \quad \text{contra} \quad H_1 : \boldsymbol{\omega} \neq \boldsymbol{\omega}_0; \boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}_0.$$

Neste caso, o logaritmo da função de verossimilhança para o modelo (4.14) fica dada por

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log \left(\frac{\lambda n_i}{\sqrt{2\pi}} \right) + \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2,$$

em que $\xi_{i1} = (2/\alpha m_i) \cosh[\lambda n_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})]$ e $\xi_{i2} = (2/\alpha m_i) \sinh[\lambda n_i(y_i - \mathbf{X}_i^\top \boldsymbol{\beta})]$, com $\boldsymbol{\theta} = (\boldsymbol{\omega}^\top, \boldsymbol{\vartheta}^\top, \boldsymbol{\beta}^\top, \lambda, \alpha)^\top$.

Seja $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\vartheta}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$; $\tilde{m}_i = m(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ $\tilde{n}_i = n(\mathbf{X}_i, \tilde{\boldsymbol{\vartheta}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\vartheta}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de MV de $\boldsymbol{\theta}$ sob H_0 . Então, temos o seguinte teorema.

Teorema 4.5 *Para o modelo (4.14), a estatística da razão de verossimilhanças é dada por*

$$\text{LR} = \left\{ \log \left[\prod_{i=1}^n \left(\frac{\tilde{\lambda} \tilde{n}_i^2 \hat{\alpha}^2}{\hat{\lambda}^2 \tilde{\alpha}^2} m_{g_i} \right) \right] \right\}, \quad m_{g_i} = \left[\frac{\cosh[\tilde{\lambda} \tilde{n}_i (y_i - \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}})]}{\tilde{m}_i \cosh[\hat{\lambda} (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})]} \right]^2$$

Prova. Ver Apêndice F. ■

4.2.6 Aplicação

Consideremos o conjunto de dados de fadiga biaxial relatado por Brown e Miller e analisado no contexto do modelo log-linear BS por Rieck e Nedelman (1991); Galea et al. (2004) e Xie e Wei (2007) referente ao ciclo de vida de um pedaço de metal até a ocorrência de falha. A variável resposta N é o número de ciclos até a falha e a variável explicativa W é o trabalho por ciclo (mJ/m^3). Foram consideradas 46 observações. Consideremos o seguinte modelo log-linear

$$Y_i = \log(N_i) = \beta_0 + \beta_1 \log(W_i) + \epsilon_i, \quad i = 1, \dots, 46, \quad (4.15)$$

em que $\epsilon_i \sim \log - \text{BSG}(\alpha, 0, 1/\lambda)$. As estimativas de MV obtidas foram $\hat{\beta}_0 = 12,2229$, $\hat{\beta}_1 = -1,6641$, $\hat{\alpha} = 2,2676$ e $\hat{\lambda} = 2,1435$.

Por meio de alguns cálculos, obtemos os resíduos $\hat{\epsilon}_i = y_i - \hat{y}_i$ e $\hat{R}_i = 2 \times \alpha^{-1} \sinh(\hat{\epsilon}_i/2)$. Com o objetivo de verificar a existência de observações extremas, bem como detectar possíveis afastamentos das suposições feitas para o modelo, apresentamos na Figura 4.1 o gráfico dos resíduos \hat{R}_i contra os valores ajustados e o gráfico normal de probabilidade dos resíduos \hat{R}_i com os envelopes gerados. Como podemos notar na Figura 4.1(a), não há presença de observações extremas e nem indicação de tendência sistemática. Através do gráfico de envelopes, Figura 4.1(b) percebemos que não há indícios de afastamento da suposição de que os erros têm distribuição log-BSG.

Diagnóstico de influência baseado em deleção de casos

Nesta seção calculamos as medidas, GCD e a estatística score SC_i apresentadas nas Seções 4.2.2 e 4.2.4, respectivamente. Os resultados de influência estão apresentados na Tabela

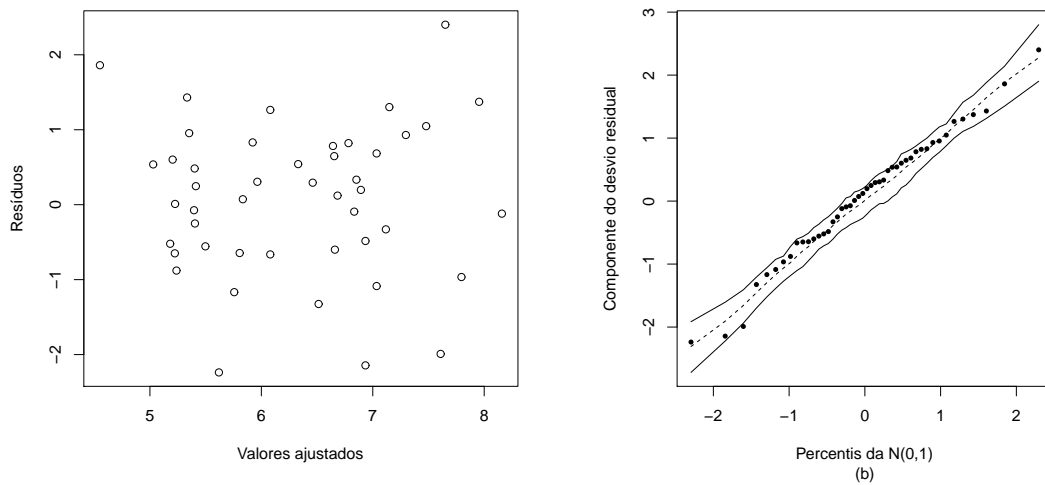


Figura 4.1: gráfico dos resíduos \hat{R}_i contra os valores ajustados (a) e gráfico normal de probabilidade dos resíduos \hat{R}_i com envelope gerado (b).

4.1, bem como ilustrado através da Figura 4.2. Como podemos observar os casos 4, 5, 12, 32 e 46 são os mais influentes, o que coincide com os resultados apresentados por Galea et al. (2004) e por Xie e Wei (2007), os quais fizeram análise de diagnóstico, baseados nos métodos de influência local e deleção de casos, respectivamente, para o modelo log-linear BS.

Disponibilizando os resultados da estatística SC_i na Figura 4.3, onde podemos perceber que nenhuma observação foi detectada como *outlier*.

A Tabela 4.2 apresenta as estimativas dos parâmetros para os casos em que cada uma das observações mais influentes foram eliminadas. Podemos notar que as observações 4 e 46 são as que mais causam mudanças nas estimativas de β_0 e β_1 . Além disso, a observação 4 causa um maior impacto nas estimativas de α e λ .

Teste de heterogeneidade do parâmetro de forma

Como já mencionamos, Rieck e Neldeman (1991) e Galea et al (2004) assumiram a homogeneidade do parâmetro de forma no modelo de regressão log-linear BS. Xie e Wei (2007) verificaram a veracidade dessa suposição para os dados de fadiga biaxial. Agora analisaremos se essa suposição continua sendo válida para o modelo de regressão log-linear BSG para o mesmo conjunto de dados. Aplicando os testes apresentados na Seção 4.2.5, considerando $m_i = m(X_i, \omega) = X_i^\omega$ e $n_i = m(X_i, \vartheta) = X_i^\vartheta$, com o propósito de facilitar o teste e usando o resultado que a distribuição assintótica do teste da razão de verossimilhanças segue uma distribuição qui-quadrado com q grau de liberdade, em que q é o número de parâmetros

Tabela 4.1: diagnóstico de influência baseado no método de deleção de casos

i	$GCD(\beta)$	$GCD(\alpha)$	$GCD(\lambda)$	GCD	i	$GCD(\beta)$	$GCD(\alpha)$	$GCD(\lambda)$	GCD
1	0,0000	0,2132	0,0888	0,0419	24	0,0000	0,1097	0,0844	0,0071
2	0,0587	0,1715	0,0449	0,1260	25	0,0163	0,0787	0,0083	0,0428
3	0,0071	0,0010	0,0224	0,0334	26	0,0002	0,0687	0,0728	0,0039
4	0,6148	4,9839	5,6206	2,1261	27	0,0000	0,1803	0,0907	0,0250
5	0,2554	1,9269	1,6389	0,5749	28	0,0015	0,0211	0,0471	0,0075
6	0,0087	0,0021	0,0093	0,0290	29	0,0000	0,2171	0,0884	0,0429
7	0,0036	0,0040	0,0288	0,0151	30	0,0002	0,0748	0,0750	0,0040
8	0,0215	0,1061	0,0172	0,0550	31	0,0137	0,0288	0,0332	0,0676
9	0,0000	0,1744	0,0907	0,0246	32	0,3714	3,5247	3,6045	0,9095
10	0,0073	0,0071	0,0047	0,0459	33	0,0000	0,1047	0,0833	0,0064
11	0,0003	0,0622	0,0702	0,0039	34	0,0000	0,1935	0,0902	0,0304
12	0,2022	2,8387	2,7235	0,4500	35	0,0000	0,1927	0,0903	0,0328
13	0,0000	0,1286	0,0876	0,0108	36	0,0000	0,1289	0,0876	0,0103
14	0,0000	0,2027	0,0897	0,0350	37	0,0000	0,2170	0,0884	0,0441
15	0,0000	0,1733	0,0907	0,0225	38	0,0061	0,0017	0,0243	0,0185
16	0,0000	0,2157	0,0885	0,0433	39	0,0589	0,2397	0,0805	0,1016
17	0,0011	0,0232	0,0487	0,0069	40	0,0040	0,0114	0,0382	0,0206
18	0,0000	0,2130	0,0888	0,0405	41	0,0000	0,2194	0,0881	0,0447
19	0,0001	0,0898	0,0796	0,0048	42	0,0004	0,0735	0,0745	0,0042
20	0,0001	0,0740	0,0747	0,0040	43	0,0002	0,0894	0,0795	0,0050
21	0,0007	0,0328	0,0553	0,0054	44	0,0000	0,1164	0,0856	0,0083
22	0,0176	0,1252	0,0244	0,0870	45	0,0000	0,1111	0,0846	0,0073
23	0,0000	0,1832	0,0906	0,0261	46	0,4121	1,3300	0,9990	1,1113

testado na hipótese nula, chegamos aos seguintes resultados:

- Para testar a heterogeneidade do parâmetro de forma, obtivemos $LR=1,4140$ com 1 grau de liberdade o que nos leva a um nível descritivo (p -valor) de $\hat{p} = 0,2344$. Ou seja, não rejeitamos a hipótese H_0 ao nível de 0,05 (5%).
- Para testar a heterogeneidade do parâmetro de escala, obtivemos $LR=1,0961$ com 1 grau de liberdade o que nos leva a um nível descritivo de $\hat{p} = 0,2951$. Ou seja, não rejeitamos a hipótese H_0 ao nível de 5%.
- Para testar a heterogeneidade dos parâmetros de forma e de escala, simultaneamente, obtivemos $LR=1,4814$ com 2 graus de liberdade o que nos leva a um nível descritivo de

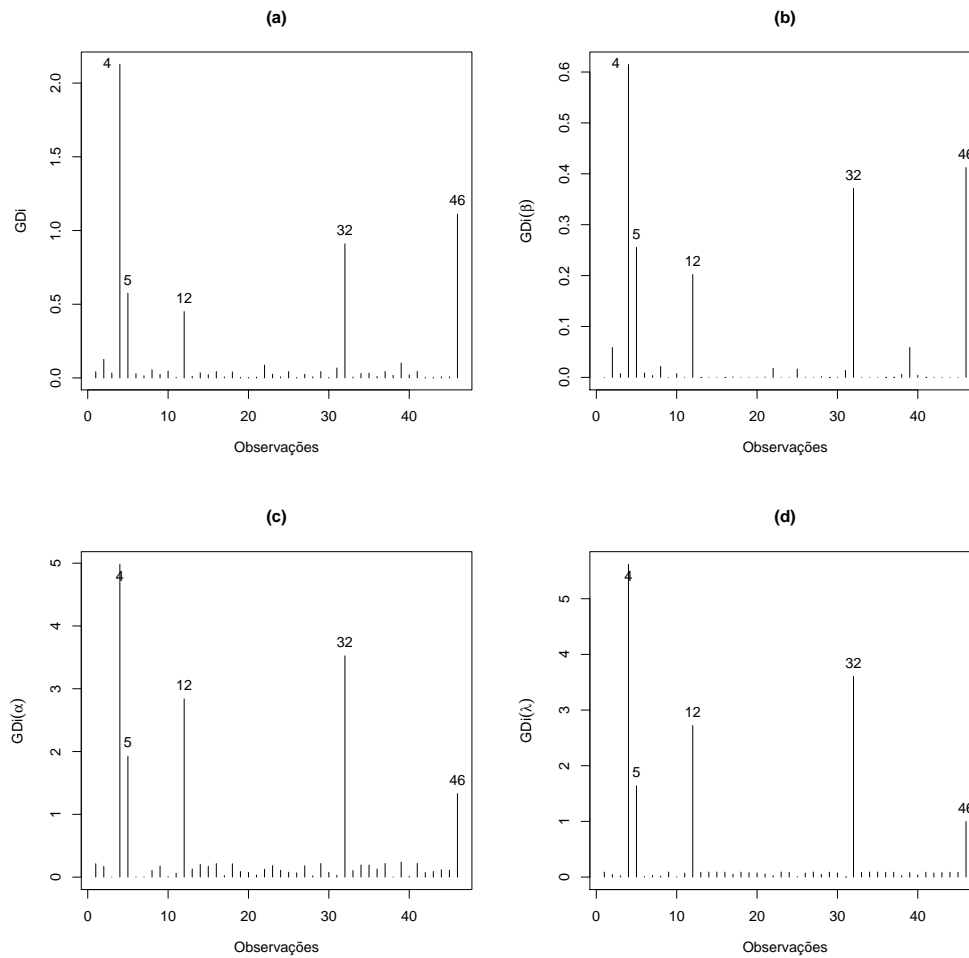


Figura 4.2: gráficos de GCD -a, $GCD(\beta)$ -b, $GCD(\alpha)$ -c e $GCD(\lambda)$ -d contra os índices das observações.

Tabela 4.2: estimativas de máxima verossimilhança dos parâmetros com o caso indicado eliminado

Caso eliminado	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\beta}_0$	$\hat{\beta}_1$
Nenhum	2,2676	2,1435	12,2229	-1,6641
4	3,0455	2,6831	11,9747	-1,6044
5	2,0512	2,0621	12,3690	-1,6998
12	2,1205	2,1170	12,2970	-1,6788
32	2,2183	2,1934	12,1160	-1,6249
46	2,3857	2,2633	12,4276	-1,7305

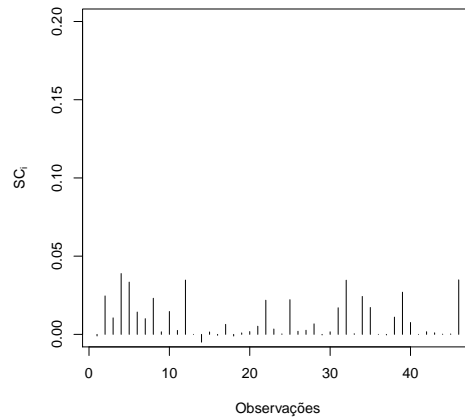


Figura 4.3: gráfico de SC_i contra os índices das observações.

Tabela 4.3: valores de LR baseados na exclusão dos casos indicados

Caso deletado	LR
Nenhum	0,3716
4	0,6961
5	0,7793
12	0,4988
32	0,1625
46	0,1205

$\hat{p} = 0,47684$. Ou seja, não rejeitamos a hipótese H_0 ao nível de 5%.

Assim, concluímos que o modelo (4.1) está adequado para os dados de fadiga biaxial.

A fim de estudar o efeito dos casos mais influentes na estatística LR, eliminamos cada uma das observações 4, 5, 12, 32 e 46, e observamos a diferença entre os valores da estatística com o conjunto de dados completo e com uma dessas observações eliminadas. Os resultados são apresentados na Tabela (4.3). Podemos concluir que existe alguma influência nos valores de LR quando a observação mais influente é excluída, porém todos os valores de LR são menores que o valor crítico do teste χ_1^2 ao nível de significância 0,05.

Conclusões

Neste trabalho discorremos sobre as distribuições Birnbaum-Saunders clássica, Birnbaum-Saunders generalizada e senh-normal e apresentamos a relação existente entre elas, bem como as principais propriedades de cada uma. Destacamos a importância das medidas de diagnóstico, bem como do teste de heterogeneidade propostos por Xie e Wei (2007) para o estudo dos modelos de regressão log-linear BS e estendemos tais resultados para o modelo de regressão log-linear BSG. Na aplicação, observamos que todos os resultados apresentados estão de acordo com os apresentados em Galea et al (2004), que utilizaram métodos de diagnóstico baseados em influência local e assumiram que o parâmetro de forma era homogêneo.

Como trabalhos futuros deixaremos a proposta de realizar uma análise de resíduos e desenvolver um modelo de regressão para modelar também os parâmetros de forma e escala.

Apêndice A: Derivadas de primeira e segunda ordens do Capítulo 2

Nesse apêndice vamos fornecer detalhes para o cálculo da função escore com relação à $\boldsymbol{\theta}^\top = (\alpha, \lambda, \beta)$, bem como das derivadas parciais de segunda ordem do logaritmo da função de verossimilhança a fim de construir a matriz de informação observada de $\boldsymbol{\theta}^\top = (\alpha, \beta)$.

As derivadas de primeira ordem do logaritmo da função de verossimilhança também chamada de função escore são dadas abaixo com relação a α , β e γ , respectivamente.

- $U_\alpha = \frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}$

$$\begin{aligned} U_\alpha &= \frac{\partial}{\partial \alpha} \left[n \log \frac{1}{\alpha} + n \log \frac{\lambda}{\sqrt{2\pi}} - \frac{1}{2\alpha^2} \sum_{j=1}^n v_j^2 - \sum_{j=1}^n \log t_j + \sum_{j=1}^n w_j \right] \\ &= -\frac{n}{\alpha} + \frac{1}{\alpha^3} \sum_{j=1}^n v_j^2, \end{aligned}$$

em que $w_j = \left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda$ e $v_j = \left(\frac{t_j}{\beta}\right)^\lambda - \left(\frac{\beta}{t_j}\right)^\lambda$.

- $U_\beta = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta}$

$$\begin{aligned} U_\beta &= -\frac{1}{2\alpha^2} \sum_{j=1}^n (-2\lambda t_j^{2\lambda} \beta^{-2\lambda-1} + 2\lambda t_j^{-2\lambda} \beta^{2\lambda-1}) + \sum_{j=1}^n \frac{(-\lambda t_j^\lambda \beta^{-\lambda-1} + \lambda t_j^{-\lambda} \beta^{\lambda-1})}{\left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda} \\ &= \frac{\lambda}{\alpha^2 \beta} \sum_{j=1}^n \left[\left(\frac{t_j}{\beta}\right)^{2\lambda} - \left(\frac{\beta}{t_j}\right)^{2\lambda} \right] + \frac{\lambda}{\beta} \sum_{j=1}^n \frac{\left(\frac{t_j}{\beta}\right)^\lambda - \left(\frac{\beta}{t_j}\right)^\lambda}{\left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda} \\ &= \frac{\lambda}{\beta} \sum_{j=1}^n \left\{ \frac{1}{\alpha^2} \left[\left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda \right] \left[\left(\frac{t_j}{\beta}\right)^\lambda - \left(\frac{\beta}{t_j}\right)^\lambda \right] + \frac{\left(\frac{t_j}{\beta}\right)^\lambda - \left(\frac{\beta}{t_j}\right)^\lambda}{\left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda} \right\} \\ &= \frac{\lambda}{\beta} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j. \end{aligned}$$

- $U_\lambda = \frac{\partial l(\boldsymbol{\theta})}{\partial \lambda}$

$$\begin{aligned}
U_\lambda &= \frac{n}{\lambda} - \frac{1}{2\alpha^2} \sum_{j=1}^n \left[2 \left(\frac{t_j}{\beta} \right)^{2\lambda} \log \left(\frac{t_j}{\beta} \right) + 2 \left(\frac{\beta}{t_j} \right)^{2\lambda} \log \left(\frac{\beta}{t_j} \right) \right] \\
&+ \sum_{j=1}^n \frac{\left(\frac{t_j}{\beta} \right)^\lambda \log \left(\frac{t_j}{\beta} \right) + \left(\frac{\beta}{t_j} \right)^\lambda \log \left(\frac{\beta}{t_j} \right)}{\left(\frac{t_j}{\beta} \right)^\lambda + \left(\frac{\beta}{t_j} \right)^\lambda} \\
&= \frac{n}{\lambda} - \frac{1}{\alpha^2} \sum_{j=1}^n \log \left(\frac{t_j}{\beta} \right) \left[\left(\frac{t_j}{\beta} \right)^\lambda + \left(\frac{\beta}{t_j} \right)^\lambda \right] \left[\left(\frac{t_j}{\beta} \right)^\lambda - \left(\frac{\beta}{t_j} \right)^\lambda \right] \\
&+ \sum_{j=1}^n \log \left(\frac{t_j}{\beta} \right) \frac{\left(\frac{t_j}{\beta} \right)^\lambda - \left(\frac{\beta}{t_j} \right)^\lambda}{\left(\frac{t_j}{\beta} \right)^\lambda + \left(\frac{\beta}{t_j} \right)^\lambda} \\
&= \frac{n}{\lambda} + \sum_{j=1}^n \log \left(\frac{t_j}{\beta} \right) v_j \left(\frac{1}{w_j} - \frac{w_j}{\alpha^2} \right).
\end{aligned}$$

Os cálculos das derivadas de segunda ordem com relação à $\boldsymbol{\theta}^\top = (\alpha, \beta)$ são dados a seguir.

- $\ddot{L}_{\alpha,\alpha} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2}$

$$\begin{aligned}
\ddot{L}_{\alpha,\alpha} &= \frac{n}{\alpha^2} - \frac{3}{\alpha^4} \sum_{j=1}^n \left[\left(\frac{t_j}{\beta} \right)^{2\lambda} + \left(\frac{\beta}{t_j} \right)^{2\lambda} - 2 \right] \\
&= \frac{n}{\alpha^2} - \frac{3}{\alpha^4} \sum_{j=1}^n v_j^2.
\end{aligned}$$

- $\ddot{L}_{\alpha,\beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \alpha} = \ddot{L}_{\beta,\alpha}$

$$\begin{aligned}
\ddot{L}_{\alpha,\beta} = \ddot{L}_{\beta,\alpha} &= \frac{1}{\alpha^3} \sum_{j=1}^n (-2\lambda t_j^{2\lambda} \beta^{-2\lambda-1} + 2\lambda \beta^{2\lambda-1} t_j^{-2\lambda}) \\
&= \frac{1}{\alpha^3} \sum_{j=1}^n \frac{2\lambda}{\beta} \left[\left(\frac{\beta}{t_j} \right)^{2\lambda} - \left(\frac{t_j}{\beta} \right)^{2\lambda} \right] \\
&= \frac{2\lambda}{\beta \alpha^3} \sum_{j=1}^n \left[\left(\frac{\beta}{t_j} \right)^\lambda + \left(\frac{t_j}{\beta} \right)^\lambda \right] \left[\left(\frac{\beta}{t_j} \right)^\lambda - \left(\frac{t_j}{\beta} \right)^\lambda \right] \\
&= -\frac{2\lambda}{\beta \alpha^3} \sum_{j=1}^n w_j v_j.
\end{aligned}$$

$$\bullet \ddot{L}_{\beta,\beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta^2}$$

$$\begin{aligned} \ddot{L}_{\beta,\beta} &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j \\ &+ \frac{\lambda}{\beta} \sum_{j=1}^n \left\{ \frac{1}{\alpha^2} (-\lambda \beta^{-\lambda-1} t_j^\lambda + \lambda \beta^{\lambda-1} t_j^{-\lambda}) + \frac{\left(\frac{\lambda}{\beta}\right) \left[\left(\frac{\beta}{t_j}\right)^\lambda - \left(\frac{t_j}{\beta}\right)^\lambda \right]}{w_j^2} \right\} v_j \\ &+ \frac{\lambda}{\beta} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) (-\beta^{-\lambda-1} \lambda t_j^\lambda - \lambda \beta^{\lambda-1} t_j^{-\lambda}) \\ &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j \\ &+ \frac{\lambda}{\beta} \sum_{j=1}^n \left\{ \frac{\lambda}{\beta \alpha^2} \left[\left(\frac{\beta}{t_j}\right)^\lambda - \left(\frac{t_j}{\beta}\right)^\lambda \right] + \frac{\left(\frac{\lambda}{\beta}\right) \left[\left(\frac{\beta}{t_j}\right)^\lambda - \left(\frac{t_j}{\beta}\right)^\lambda \right]}{w_j^2} \right\} v_j \\ &+ \frac{\lambda}{\beta} \sum_{j=1}^n - \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) \left(\frac{\lambda}{\beta}\right) \left[\left(\frac{t_j}{\beta}\right)^\lambda + \left(\frac{\beta}{t_j}\right)^\lambda \right] \\ &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j \\ &+ \frac{\lambda^2}{\beta^2} \sum_{j=1}^n \left(-\frac{v_j^2}{\alpha^2} - \frac{v_j^2}{w_j^2} \right) - \frac{\lambda^2}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j^2 - 1 \right) \\ &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j - \frac{\lambda^2}{\beta^2} \sum_{j=1}^n \left(\frac{v_j^2}{w_j^2} - 1 \right) - \frac{\lambda^2}{\alpha^2 \beta^2} \sum_{j=1}^n (v_j^2 + w_j^2) \\ &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j - \frac{\lambda^2}{\beta^2} \sum_{j=1}^n \frac{-4}{w_j^2} - \frac{\lambda^2}{\alpha^2 \beta^2} \sum_{j=1}^n 2 \left[\left(\frac{t_j}{\beta}\right)^{2\lambda} + \left(\frac{\beta}{t_j}\right)^{2\lambda} \right] \\ &= -\frac{\lambda}{\beta^2} \sum_{j=1}^n \left(\frac{1}{\alpha^2} w_j - \frac{1}{w_j} \right) v_j + \frac{4\lambda^2}{\beta^2} \sum_{j=1}^n \frac{1}{w_j^2} - \frac{2\lambda^2}{\alpha^2 \beta^2} \sum_{j=1}^n (v_j^2 + 2). \end{aligned}$$

Apêndice B: Prova do Teorema 4.1

As derivadas de segunda ordem de $l(\boldsymbol{\theta})$ são dadas a seguir.

$$\begin{aligned}
 \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} &= \frac{\partial}{\partial \alpha} \left\{ \frac{4}{\alpha^3} \sum_{i=1}^n \sinh^2 [\lambda(y_i - \mu_i)] - \frac{n}{\alpha} \right\} \\
 &= -\frac{12}{\alpha^4} \sum_{i=1}^n \sinh^2 [\lambda(y_i - \mu_i)] + \frac{n}{\alpha^2} \\
 &= \delta
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \alpha} &= \frac{\partial}{\partial \lambda} \sum_{i=1}^n \left\{ -\frac{1}{\alpha} + \frac{4}{\alpha^3} \sinh^2 [\lambda(y_i - \mu_i)] \right\} \\
 &= \sum_{i=1}^n \frac{8}{\alpha^3} \sinh[\lambda(y_i - \mu_i)] \cosh[\lambda(y_i - \mu_i)] (y_i - \mu_i) \\
 &= \sum_{i=1}^n \frac{4}{\alpha^3} \sinh[2\lambda(y_i - \mu_i)] (y_i - \mu_i) \\
 &= 2\mathbf{h}^\top (\mathbf{Y} - \boldsymbol{\mu})
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\beta}^\top} &= \frac{\partial}{\partial \alpha} \sum_{i=1}^n \frac{4}{\alpha^2} \sinh[\lambda(y_i - \mu_i)] \cosh[\lambda(y_i - \mu_i)] \lambda \mathbf{X}_i \\
 &= \sum_{i=1}^n -\frac{8}{\alpha^3} \sinh[\lambda(y_i - \mu_i)] \cosh[\lambda(y_i - \mu_i)] \lambda \mathbf{X}_i \\
 &= \sum_{i=1}^n -\frac{4}{\alpha^3} \sinh[2\lambda(y_i - \mu_i)] \lambda \mathbf{X}_i \\
 &= -2\lambda \mathbf{h}^\top \mathbf{X}
 \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \boldsymbol{\beta}^\top} &= \frac{\partial}{\partial \lambda} \sum_{i=1}^n \left\{ \frac{2}{\alpha^2} \sinh[2\lambda(y_i - \mu_i)] - \operatorname{tgh}[\lambda(y_i - \mu_i)] \right\} \lambda \mathbf{X}_i \\
&= \sum_{i=1}^n \left\{ \frac{4}{\alpha^2} \cosh[2\lambda(y_i - \mu_i)] - \operatorname{sech}^2[\lambda(y_i - \mu_i)] \right\} \lambda(y_i - \mu_i) \mathbf{X}_i \\
&+ \sum_{i=1}^n \left\{ \frac{2}{\alpha^2} \sinh[2\lambda(y_i - \mu_i)] - \operatorname{tgh}[\lambda(y_i - \mu_i)] \right\} \mathbf{X}_i \\
&= \mathbf{g}^\top X
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}^\top \partial \boldsymbol{\beta}} &= \frac{\partial}{\partial \boldsymbol{\beta}^\top} \sum_{i=1}^n \left\{ \frac{2}{\alpha^2} \sinh[2\lambda(y_i - \mu_i)] - \operatorname{tgh}[\lambda(y_i - \mu_i)] \right\} \lambda \mathbf{X}_i^\top \\
&= \sum_{i=1}^n \left\{ \operatorname{sech}^2[\lambda(y_i - \mu_i)] - \frac{4}{\alpha^2} \cosh[2\lambda(y_i - \mu_i)] \right\} \lambda^2 \mathbf{X}_i^\top \mathbf{X}_i \\
&= \lambda^2 X^\top V X
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} &= -\frac{n}{\lambda^2} + \frac{\partial}{\partial \lambda} \sum_{i=1}^n \left\{ \operatorname{tgh}[\lambda(y_i - \mu_i)] - \frac{2}{\alpha^2} \sinh[2\lambda(y_i - \mu_i)] \right\} (y_i - \mu_i) \\
&= -\frac{n}{\lambda^2} + \sum_{i=1}^n \left\{ \operatorname{sech}^2[\lambda(y_i - \mu_i)] - \frac{4}{\alpha^2} \cosh[2\lambda(y_i - \mu_i)] \right\} (y_i - \mu_i)^2 \\
&= \gamma
\end{aligned}$$

Então a matriz de informação observada $-\ddot{l}(\hat{\boldsymbol{\theta}})$ é dada por

$$\begin{aligned}
-\ddot{l}(\hat{\boldsymbol{\theta}}) &= -\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = - \left[\begin{array}{ccc} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \alpha} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \alpha} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} \end{array} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
&= - \left[\begin{array}{ccc} \lambda^2 X^\top V X & X^\top \mathbf{g} & -2\lambda X^\top \mathbf{h} \\ \mathbf{g}^\top X & \gamma & 2\mathbf{h}^\top (\mathbf{Y} - \boldsymbol{\mu}) \\ -2\lambda \mathbf{h}^\top X & 2(\mathbf{Y}^\top - \boldsymbol{\mu}^\top) \mathbf{h} & \delta \end{array} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} .
\end{aligned}$$

As derivadas de $l_i(\boldsymbol{\theta})$ em relação a $\boldsymbol{\theta}$ são

$$\frac{\partial l_i(\boldsymbol{\theta})}{\partial \alpha} = \frac{4}{\alpha^3} \sum_{j \neq i}^n \sinh^2[\lambda(y_j - \mu_j)] - \frac{n-1}{\alpha};$$

$$\begin{aligned}\frac{\partial l_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} &= \lambda \sum_{j \neq i}^n \left\{ \frac{4}{\alpha^2} \sinh[\lambda(y_j - \mu_j)] \cosh[\lambda(y_j - \mu_j)] - \frac{\sinh[\lambda(y_j - \mu_j)]}{\cosh[\lambda(y_j - \mu_j)]} \right\} \mathbf{X}_j^\top; \\ \frac{\partial l_i(\boldsymbol{\theta})}{\partial \lambda} &= \frac{n-1}{\lambda} + \sum_{j \neq i}^n \left\{ \operatorname{tgh}[\lambda(y_j - \mu_j)] - \frac{4}{\alpha^2} \sinh[\lambda(y_j - \mu_j)] \cosh[\lambda(y_j - \mu_j)] \right\} (y_j - \mu_j).\end{aligned}$$

Sendo

$$\xi_{i1} = \frac{2}{\alpha} \cosh[\lambda(y_i - \mu_i)] \quad \text{e} \quad \xi_{i2} = \frac{2}{\alpha} \sinh[\lambda(y_i - \mu_i)], \quad (4.16)$$

temos

$$\begin{aligned}\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} &= \frac{\partial l_i(\boldsymbol{\theta})}{\partial \alpha} + \frac{1}{\alpha} \xi_{i2}^2 - \frac{1}{\alpha}; \\ \frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} &= \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \lambda} + (y_i - \mu_i) \left(\frac{\xi_{i2}}{\xi_{i1}} - \xi_{i1} \xi_{i2} \right) + \frac{1}{\lambda}; \\ \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} &= \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} + \lambda \mathbf{X}_i \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right).\end{aligned}$$

Como $\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} \Big|_{\hat{\boldsymbol{\theta}}} = \mathbf{0}$, $\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} \Big|_{\hat{\boldsymbol{\theta}}} = \mathbf{0}$ e $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \Big|_{\hat{\boldsymbol{\theta}}} = \mathbf{0}$, temos

$$\begin{aligned}\frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \alpha} \Big|_{\hat{\boldsymbol{\theta}}} &= \left(\frac{1}{\alpha} - \frac{1}{\alpha} \xi_{i2}^2 \right) \Big|_{\hat{\boldsymbol{\theta}}} = -\hat{\eta}_i; \\ \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \lambda} \Big|_{\hat{\boldsymbol{\theta}}} &= (y_i - \mu_i) \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) \Big|_{\hat{\boldsymbol{\theta}}} - \frac{1}{\lambda} = \hat{r}_i (y_i - \hat{\mu}_i) - \frac{1}{\lambda}; \\ \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \Big|_{\hat{\boldsymbol{\theta}}} &= \hat{\lambda} \mathbf{X}_i \left(\frac{\xi_{i2}}{\xi_{i1}} - \xi_{i1} \xi_{i2} \right) \Big|_{\hat{\boldsymbol{\theta}}} = -(\hat{\lambda} \mathbf{X}_i \hat{r}_i).\end{aligned}$$

Portanto,

$$\dot{l}_{(i)}(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\hat{\boldsymbol{\theta}}} = \begin{bmatrix} -\hat{\lambda} \mathbf{X}_i \hat{r}_i \\ \hat{r}_i (y_i - \hat{\mu}_i) - \frac{1}{\lambda} \\ -\hat{\eta}_i \end{bmatrix}.$$

Substituindo na equação (4.6), temos

$$\hat{\boldsymbol{\theta}}_{(i)}^1 = \begin{bmatrix} \hat{\boldsymbol{\beta}}^\top \\ \hat{\lambda} \\ \hat{\alpha} \end{bmatrix} - \begin{bmatrix} \lambda^2 \mathbf{X}^\top \mathbf{V} \mathbf{X} & \mathbf{X}^\top \mathbf{g} & -2\lambda \mathbf{X}^\top \mathbf{h} \\ \mathbf{g}^\top \mathbf{X} & \gamma & 2\mathbf{h}^\top (\mathbf{Y} - \boldsymbol{\mu}) \\ -2\lambda \mathbf{h}^\top \mathbf{X} & 2(\mathbf{Y}^\top - \boldsymbol{\mu}^\top) \mathbf{h} & \delta \end{bmatrix}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}^{-1} \cdot \begin{bmatrix} -\hat{\lambda} \mathbf{X}_i \hat{r}_i \\ \hat{r}_i (y_i - \hat{\mu}_i) - \frac{1}{\lambda} \\ -\hat{\eta}_i \end{bmatrix}.$$

Cálculo da inversa de $\ddot{l}(\hat{\boldsymbol{\theta}})$:

$$\text{Sejam } \ddot{l}(\hat{\boldsymbol{\theta}}) = B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \quad \text{e} \quad [\ddot{l}(\hat{\boldsymbol{\theta}})]^{-1} = B^{-1} = \begin{bmatrix} B^{11} & B^{12} \\ B^{21} & B^{22} \end{bmatrix}. \quad \text{Denotemos}$$

$$\begin{aligned}
B_{11} &= \hat{\lambda}^2 X^\top \hat{V} X_{p \times p} \\
B_{12} &= [X^\top \hat{\mathbf{g}} \quad -2\hat{\lambda} X^\top \hat{\mathbf{h}}]_{p \times 2} \\
B_{21} &= \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda} \hat{\mathbf{h}}^\top X \end{bmatrix}_{2 \times p} \\
B_{22} &= \begin{bmatrix} \hat{\gamma} & 2\hat{\mathbf{h}}^\top (\mathbf{Y} - \hat{\boldsymbol{\mu}}) \\ 2(\mathbf{Y}^\top - \hat{\boldsymbol{\mu}}^\top) \hat{\mathbf{h}} & \hat{\delta} \end{bmatrix}_{2 \times 2}.
\end{aligned}$$

Temos

$$\begin{aligned}
B^{22} &= [B_{22} - B_{21} B_{11}^{-1} B_{12}]^{-1} \\
&= \left\{ \begin{bmatrix} \hat{\gamma} & \hat{\kappa} \\ \hat{\kappa} & \hat{\delta} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda} \hat{\mathbf{h}}^\top X \end{bmatrix} \hat{\lambda}^{-2} (X^\top \hat{V} X)^{-1} [X^\top \hat{\mathbf{g}} \quad -2\hat{\lambda} X^\top \hat{\mathbf{h}}] \right\}^{-1} \\
&= \begin{bmatrix} \hat{\gamma} - \hat{\lambda}^{-2} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}} & \hat{\kappa} + 2\hat{\lambda}^{-1} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}} \\ \hat{\kappa} + 2\hat{\lambda}^{-1} \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}} & \hat{\delta} - 4\hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}} \end{bmatrix}^{-1} \\
&= \frac{1}{\hat{\tau}} \begin{bmatrix} \hat{\delta} - 4\hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}} & -\hat{\kappa} - 2\hat{\lambda}^{-1} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}} \\ -\hat{\kappa} - 2\hat{\lambda}^{-1} \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}} & \hat{\gamma} - \hat{\lambda}^{-2} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}} \end{bmatrix} \\
&= \begin{bmatrix} \hat{a} & \hat{b} \\ \hat{c} & \hat{d} \end{bmatrix}_{2 \times 2},
\end{aligned}$$

em que $\hat{\kappa} = 2\hat{\mathbf{h}}^\top (\mathbf{Y} - \hat{\boldsymbol{\mu}})$ e

$$\hat{\tau} = (\hat{\gamma} - \hat{\lambda}^{-2} \hat{\mathbf{g}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}}) (\hat{\delta} - 4\hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{h}}) - (\hat{\kappa} + 2\hat{\lambda}^{-1} \hat{\mathbf{h}}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\mathbf{g}})^2.$$

$$\begin{aligned}
B^{11} &= [B_{11} - B_{12} B_{22}^{-1} B_{21}]^{-1} \\
&= \left\{ \hat{\lambda}^2 X^\top \hat{V} X - [X^\top \hat{\mathbf{g}} \quad -2\hat{\lambda} X^\top \hat{\mathbf{h}}] \begin{bmatrix} \hat{\gamma} & 2\hat{\mathbf{h}}^\top (\mathbf{Y} - \hat{\boldsymbol{\mu}}) \\ 2(\mathbf{Y}^\top - \hat{\boldsymbol{\mu}}^\top) \hat{\mathbf{h}} & \hat{\delta} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda} \hat{\mathbf{h}}^\top X \end{bmatrix} \right\}^{-1} \\
&= \left\{ \hat{\lambda}^2 X^\top \hat{V} X - [X^\top \hat{\mathbf{g}} \quad -2\hat{\lambda} X^\top \hat{\mathbf{h}}] \frac{1}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} \begin{bmatrix} \hat{\delta} & -\hat{\kappa} \\ -\hat{\kappa} & \hat{\gamma} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda} \hat{\mathbf{h}}^\top X \end{bmatrix} \right\}^{-1} \\
&= \left\{ \hat{\lambda}^2 X^\top \hat{V} X - \frac{1}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} [X^\top \hat{\mathbf{g}} \hat{\delta} + 2\hat{\lambda} X^\top \hat{\mathbf{h}} \hat{\kappa} \quad -X^\top \hat{\mathbf{g}} \hat{\kappa} - 2\hat{\lambda} X^\top \hat{\mathbf{h}} \hat{\gamma}] \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda} \hat{\mathbf{h}}^\top X \end{bmatrix} \right\}^{-1} \\
&= \left\{ \hat{\lambda}^2 X^\top \hat{V} X - \frac{1}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} [X^\top \hat{\mathbf{g}} \hat{\delta} \hat{\mathbf{g}}^\top X + 2\hat{\lambda} X^\top \hat{\mathbf{h}} \hat{\kappa} \hat{\mathbf{g}}^\top X \quad 2\hat{\lambda} X^\top \hat{\mathbf{g}} \hat{\kappa} \hat{\mathbf{h}}^\top X + 4\hat{\lambda}^2 X^\top \hat{\mathbf{h}} \hat{\gamma} \hat{\mathbf{h}}^\top X] \right\}^{-1} \\
&= \left\{ X^\top \left(\hat{\lambda}^2 \hat{V} - \frac{\hat{\mathbf{g}} \hat{\delta} \hat{\mathbf{g}}^\top}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} - \frac{2\hat{\lambda} \hat{\mathbf{h}} \hat{\kappa} \hat{\mathbf{g}}^\top}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} - \frac{2\hat{\lambda} \hat{\mathbf{g}} \hat{\kappa} \hat{\mathbf{h}}^\top}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} - \frac{+4\hat{\lambda}^2 \hat{\mathbf{h}} \hat{\gamma} \hat{\mathbf{h}}^\top}{\hat{\delta}\hat{\gamma} - \hat{\kappa}^2} \right) X \right\}^{-1} \\
&= (X^\top \hat{A} X)_{p \times p}^{-1}.
\end{aligned}$$

$$\begin{aligned}
B^{12} &= -B_{11}^{-1}B_{12}B^{22} \\
&= -\hat{\lambda}^{-2}(X^\top \hat{V}X)^{-1} [X^\top \hat{\mathbf{g}} \quad -2\hat{\lambda}X^\top \hat{\mathbf{h}}] B^{22} \\
&= \hat{\lambda}^{-2}(X^\top \hat{V}X)^{-1} [-X^\top \hat{\mathbf{g}} \quad 2\hat{\lambda}X^\top \hat{\mathbf{h}}] \begin{bmatrix} \hat{a} & \hat{b} \\ \hat{c} & \hat{d} \end{bmatrix} \\
&= \hat{\lambda}^{-2}(X^\top \hat{V}X)^{-1} [2\hat{\lambda}X^\top \hat{\mathbf{h}}\hat{c} - X^\top \hat{\mathbf{g}}\hat{a} \quad 2\hat{\lambda}X^\top \hat{\mathbf{h}}\hat{d} - X^\top \hat{\mathbf{g}}\hat{b}] \\
&= [\hat{\mathbf{e}} \quad \hat{\mathbf{f}}]_{p \times 2}.
\end{aligned}$$

$$\begin{aligned}
B^{21} &= -B_{22}^{-1}B_{21}B^{11} = - \begin{bmatrix} \hat{\gamma} & \hat{\kappa} \\ \hat{\kappa} & \hat{\delta} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda}\hat{\mathbf{h}}^\top X \end{bmatrix} B^{11} \\
&= -\frac{1}{\hat{\gamma}\hat{\delta} - \hat{\kappa}^2} \begin{bmatrix} \hat{\delta} & -\hat{\kappa} \\ -\hat{\kappa} & \hat{\gamma} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{g}}^\top X \\ -2\hat{\lambda}\hat{\mathbf{h}}^\top X \end{bmatrix} (X^\top \hat{A}X)^{-1} \\
&= -\frac{1}{\hat{\gamma}\hat{\delta} - \hat{\kappa}^2} \begin{bmatrix} (\hat{\delta}\hat{\mathbf{g}}^\top X + 2\hat{\kappa}\hat{\lambda}\hat{\mathbf{h}}^\top X)(X^\top \hat{A}X)^{-1} \\ -(\hat{\kappa}\hat{\mathbf{g}}^\top X + 2\hat{\gamma}\hat{\lambda}\hat{\mathbf{h}}^\top X)(X^\top \hat{A}X)^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \hat{\mathbf{i}}^\top \\ \hat{\mathbf{j}}^\top \end{bmatrix}_{2 \times p}.
\end{aligned}$$

Substituindo na equação (4.6), temos

$$\hat{\boldsymbol{\theta}}_{(i)}^1 = \hat{\boldsymbol{\theta}} - \begin{bmatrix} (X^\top \hat{A}X)^{-1} & \mathbf{e} & \mathbf{f} \\ \mathbf{i}^\top & a & b \\ \mathbf{j}^\top & c & d \end{bmatrix}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \begin{bmatrix} -\hat{\lambda}\mathbf{X}_i \hat{r}_i \\ \hat{r}_i(y_i - \hat{\mu}_i) - \frac{1}{\hat{\lambda}} \\ -\hat{\eta}_i \end{bmatrix}.$$

Portanto, as aproximações para os estimadores de cada parâmetro sem a i -ésima observação são dadas por

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{(i)}^1 &= \hat{\boldsymbol{\beta}} + \hat{\lambda}(X^\top \hat{A}X)^{-1} \mathbf{X}_i \hat{r}_i - \hat{\mathbf{e}} \hat{r}_i(y_i - \hat{\mu}_i) + \hat{\mathbf{e}} \hat{\lambda}^{-1} + \hat{\mathbf{f}} \hat{\eta}_i; \\
\hat{\lambda}_{(i)}^1 &= \hat{\lambda} + \hat{\lambda} \hat{\mathbf{i}}^\top \mathbf{X}_i \hat{r}_i - \hat{a} \hat{r}_i(y_i - \hat{\mu}_i) + \hat{a} \hat{\lambda}^{-1} + \hat{b} \hat{\eta}_i; \\
\hat{\alpha}_{(i)}^1 &= \hat{\alpha} + \hat{\lambda} \hat{\mathbf{j}}^\top \mathbf{X}_i \hat{r}_i - \hat{c} \hat{r}_i(y_i - \hat{\mu}_i) + \hat{c} \hat{\lambda}^{-1} + \hat{d} \hat{\eta}_i.
\end{aligned}$$

Apêndice C: Prova do Teorema 4.2

Sejam $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}^\top, \lambda, \alpha)^\top$ e $\boldsymbol{\theta}_2 = \nu$. A fim de obter a estatística do teste de escore para testar a hipótese $H_0 : \boldsymbol{\theta}_2 = \nu = 0$, precisamos encontrar as derivadas de segunda ordem de $l_{mi}(\boldsymbol{\theta})$ sob H_0 . Para o parâmetro $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}^\top, \lambda, \alpha)^\top$, as derivadas foram encontradas na demonstração do Teorema 4.1.

Para o modelo MSOM (4.10), o logaritmo da função de verossimilhança pode ser escrita como

$$\begin{aligned}
 l_{mi}(\boldsymbol{\theta}) &= n \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{j \neq i}^n \log \xi_{j1} - \frac{1}{2} \sum_{j \neq i}^n \xi_{j2}^2 + \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2 \\
 &= n \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) - \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) \\
 &+ \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{j \neq i}^n \log \xi_{j1} - \frac{1}{2} \sum_{j \neq i}^n \xi_{j2}^2 + \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2 \\
 &= (n-1) \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \sum_{j \neq i}^n \log \xi_{j1} - \frac{1}{2} \sum_{j \neq i}^n \xi_{j2}^2 + \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2 \\
 &= l_i(\boldsymbol{\theta}) + \log \left(\frac{\lambda}{\sqrt{2\pi}} \right) + \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2,
 \end{aligned}$$

em que $\bar{\xi}_{i1} = \frac{2}{\alpha} \cosh[\lambda(y_i - \mathbf{X}_i^\top \boldsymbol{\beta} - \nu)]$, $\bar{\xi}_{i2} = \frac{2}{\alpha} \sinh[\lambda(y_i - \mathbf{X}_i^\top \boldsymbol{\beta} - \nu)]$.

As derivadas de $l_{mi}(\boldsymbol{\theta})$ com respeito à $\boldsymbol{\theta}_2$ sob H_0 são dadas, respectivamente, por

$$\begin{aligned}
 \frac{\partial l_{mi}(\boldsymbol{\theta})}{\partial \nu} \Big|_{\hat{\boldsymbol{\theta}}} &= \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) \hat{\lambda} \Big|_{\hat{\boldsymbol{\theta}}} \\
 &= \hat{\lambda}_{mi}(r_i)_{\hat{\boldsymbol{\theta}}};
 \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_{mi}(\boldsymbol{\theta})}{\partial \nu^2} \Big|_{\hat{\boldsymbol{\theta}}} &= -\frac{4}{\alpha_{mi}^2} \left\{ \sinh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] + \cosh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] \right\} \hat{\lambda}_{mi}^2 \\
&+ \frac{\cosh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] - \sinh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})]}{\cosh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})]} \lambda_{mi}^2 \\
&= \frac{\lambda_{mi}^2}{\cosh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})]} \\
&- \frac{4\lambda_{mi}^2}{\alpha_{mi}^2} \left\{ \sinh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] + \cosh^2[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] \right\} \\
&= \frac{4\lambda_{mi}^2}{\alpha_{mi}^2 \xi_{i1}^2} - (\xi_{i1}^2 + \xi_{i2}^2) \lambda_{mi}^2 \\
&= (u_i)_{\hat{\boldsymbol{\theta}}};
\end{aligned}$$

$$\frac{\partial^2 l_{mi}(\boldsymbol{\theta})}{\partial \nu \partial \boldsymbol{\beta}} \Big|_{\hat{\boldsymbol{\theta}}} = (u_i)_{\hat{\boldsymbol{\theta}}} \mathbf{X}_i^\top;$$

$$\frac{\partial^2 l_{mi}(\boldsymbol{\theta})}{\partial \nu \partial \lambda} \Big|_{\hat{\boldsymbol{\theta}}} = -(u_i)_{\hat{\boldsymbol{\theta}}} \frac{1}{\hat{\lambda}_{mi}} (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi});$$

$$\begin{aligned}
\frac{\partial^2 l_{mi}(\boldsymbol{\theta})}{\partial \nu \partial \alpha} \Big|_{\hat{\boldsymbol{\theta}}} &= -\frac{8}{\alpha_{mi}^3} \hat{\lambda}_{mi} \cosh[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] \sinh[\hat{\lambda}_{mi}(y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}_{mi})] \\
&= -\frac{2}{\hat{\alpha}_{mi}} \hat{\lambda}_{mi} (\xi_{i1} \xi_{i2})_{\hat{\boldsymbol{\theta}}} \\
&= (q_i)_{\hat{\boldsymbol{\theta}}}.
\end{aligned}$$

Então a matriz de informação observada $-\ddot{l}(\boldsymbol{\theta})$ avaliada em $\hat{\boldsymbol{\theta}}$ pode ser obtida por

$$\begin{aligned}
-\ddot{l}(\hat{\boldsymbol{\theta}}) &= - \begin{bmatrix} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \nu} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda \partial \nu} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \nu} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \nu \boldsymbol{\beta}^\top} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \nu \partial \lambda} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \nu \partial \alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \nu^2} \end{bmatrix}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
&= - \begin{bmatrix} \lambda^2 \mathbf{X}^\top \mathbf{V} \mathbf{X} & \mathbf{X}^\top \mathbf{g} & -2\lambda \mathbf{X}^\top \mathbf{h} & u_i \mathbf{X}_i \\ \mathbf{g}^\top \mathbf{X} & \gamma & 2\mathbf{h}^\top (\mathbf{Y} - \mathbf{X} \boldsymbol{\beta}) & -\lambda^{-1} u_i (y_i - \mathbf{X}_i^\top \boldsymbol{\beta}) \\ -2\lambda \mathbf{h}^\top \mathbf{X} & 2(\mathbf{Y}^\top - \boldsymbol{\beta}^\top \mathbf{X}^\top) \mathbf{h} & \delta & q_i \\ u_i \mathbf{X}_i^\top & -\lambda^{-1} u_i (y_i - \mathbf{X}_i^\top \boldsymbol{\beta}) & q_i & u_i \end{bmatrix}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.
\end{aligned}$$

Particionemos $-\ddot{l}(\hat{\boldsymbol{\theta}})$ e $-\ddot{l}^{-1}(\hat{\boldsymbol{\theta}})$ de acordo com os parâmetros $\boldsymbol{\theta}_1$ e $\boldsymbol{\theta}_2$ da seguinte forma

$$-\ddot{l}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{e} \quad -\ddot{l}^{-1}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix},$$

onde

$$\begin{aligned} A_{11} &= - \begin{bmatrix} \hat{\lambda}^2 X^\top \hat{V} X & X^\top \hat{\boldsymbol{g}} & -2\hat{\lambda} X^\top \hat{\boldsymbol{h}} \\ \hat{\boldsymbol{g}}^\top X & \hat{\gamma} & 2\hat{\boldsymbol{h}}^\top (\mathbf{Y} - X\hat{\boldsymbol{\beta}}) \\ -2\hat{\lambda} \hat{\boldsymbol{h}}^\top X & 2(\mathbf{Y}^\top - \hat{\boldsymbol{\beta}}^\top X^\top) \hat{\boldsymbol{h}} & \hat{\delta} \end{bmatrix}; \\ A_{12} &= - \begin{bmatrix} \hat{u}_i \mathbf{X}_i \\ -\hat{\lambda}^{-1} \hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}) \\ \hat{q}_i \end{bmatrix}; \\ A_{21} &= - \begin{bmatrix} \hat{u}_i \mathbf{X}_i^\top & -\hat{\lambda}^{-1} \hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}) & \hat{q}_i \end{bmatrix}; \\ A_{22} &= - \begin{bmatrix} \hat{u}_i \end{bmatrix}. \end{aligned}$$

Então a estatística score para a hipótese $H_0 : \nu = 0$ é (Cook e Hinkley, 1974)

$$SC_i = \left\{ \left(\frac{\partial l_{mi}(\boldsymbol{\theta})}{\partial \nu} \right)^\top A^{22} \left(\frac{\partial l_{mi}(\boldsymbol{\theta})}{\partial \nu} \right) \right\}_{\hat{\boldsymbol{\theta}}} \quad (\text{A3}).$$

Cálculo de A^{22}

Temos

$$A^{22} = [A_{22} - A_{21} A_{11}^{-1} A_{12}]^{-1}.$$

Daí,

$$\begin{aligned} (A^{22})^{-1} &= - [\hat{u}_i \mathbf{X}_i^\top \quad -\hat{\lambda}^{-1} \hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}) \quad \hat{q}_i] A_{11}^{-1} \begin{bmatrix} \hat{u}_i \mathbf{X}_i \\ -\hat{\lambda}^{-1} \hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}) \\ \hat{q}_i \end{bmatrix} - \hat{u}_i \\ &= [\hat{u}_i \mathbf{X}_i^\top \quad -\frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \quad \hat{q}_i] \begin{bmatrix} (X^\top \hat{A} X)^{-1} & \hat{\boldsymbol{e}} & \hat{\boldsymbol{f}} \\ \hat{\boldsymbol{i}}^\top & \hat{a} & \hat{b} \\ \hat{\boldsymbol{j}}^\top & \hat{c} & \hat{d} \end{bmatrix} \begin{bmatrix} \hat{u}_i \mathbf{X}_i \\ -\frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \\ \hat{q}_i \end{bmatrix} - \hat{u}_i \\ &= [\hat{u}_i \mathbf{X}_i^\top (X^\top \hat{A} X)^{-1} - \frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \hat{\boldsymbol{i}}^\top + \hat{q}_i \hat{\boldsymbol{j}}^\top] \hat{u}_i \mathbf{X}_i \\ &\quad - [\hat{u}_i \mathbf{X}_i^\top \hat{\boldsymbol{e}} - \frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \hat{a} + \hat{q}_i \hat{c}] \frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \\ &\quad + [\hat{u}_i \mathbf{X}_i^\top \hat{\boldsymbol{f}} - \frac{\hat{u}_i (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})}{\hat{\lambda}} \hat{b} + \hat{q}_i \hat{d}] \hat{q}_i - \hat{u}_i \\ &= \hat{\rho}_i. \end{aligned}$$

Substituindo os resultados obtidos em (A3), temos

$$SC_i = \{r_i^2 \rho_i^{-1}\}_{\theta}.$$

Apêndice D: Prova do Teorema 4.3

A estatística da razão de verossimilhanças é dada por

$$\text{LR} = 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\},$$

em que $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$; $\tilde{m}_i = m(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob H_0 . Assim

$$\begin{aligned} \text{LR} &= 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\} \\ &= 2 \left\{ \left[n \log \lambda + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\tilde{\boldsymbol{\theta}}} - \left[n \log \lambda + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\hat{\boldsymbol{\theta}}} \right\}. \end{aligned}$$

Temos que

$$\hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2[\hat{\lambda}(y_i - \hat{\mu}_i)]$$

e

$$\tilde{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \frac{1}{\tilde{m}_i^2} \sinh^2[\tilde{\lambda}(y_i - \tilde{\mu}_i)],$$

em que $\mu_i = \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}$. Combinando as expressões acima obtemos

$$\begin{aligned} \text{LR} &= 2n \log \frac{\tilde{\lambda}}{\hat{\lambda}} + 2 \left\{ \sum_{i=1}^n \log \frac{2}{\alpha m_i} \cosh[\lambda(y_i - \mu_i)] - \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2 m_i^2} \sinh^2[\lambda(y_i - \mu_i)] \right\}_{\tilde{\boldsymbol{\theta}}} \\ &+ 2 \left\{ \sum_{i=1}^n -\log \frac{2}{\alpha} \cosh[\lambda(y_i - \mu_i)] + \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2} \sinh^2[\lambda(y_i - \mu_i)] \right\}_{\hat{\boldsymbol{\theta}}} \\ &= 2 \left\{ n \log \frac{\tilde{\lambda}}{\hat{\lambda}} + n \log \frac{2}{\tilde{\alpha}} - n \log \frac{2}{\hat{\alpha}} - \sum_{i=1}^n \log \tilde{m}_i + \sum_{i=1}^n \log \cosh[\lambda(y_i - \mu_i)]_{\tilde{\boldsymbol{\theta}}} \right\} \\ &- 2 \left\{ \sum_{i=1}^n \log \cosh[\lambda(y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} - \frac{2}{\hat{\alpha}^2} \sum_{i=1}^n \sinh^2[\lambda(y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} \right\} \\ &- \frac{4}{\tilde{\alpha}^2} \sum_{i=1}^n \frac{1}{\tilde{m}_i^2} \sinh^2[\lambda(y_i - \mu_i)]_{\tilde{\boldsymbol{\theta}}} \end{aligned}$$

$$\begin{aligned}
&= 2 \left\{ n \log \frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} - \sum_{i=1}^n \log \tilde{m}_i + \sum_{i=1}^n \log \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} + \frac{n}{2} - \frac{n}{2} \right\} \\
&= 2 \left\{ n \log \frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} + \sum_{i=1}^n \log \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} \tilde{m}_i^{-1} \right\} \\
&= 2 \left\{ n \log \frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} + \log \prod_{i=1}^n \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} \tilde{m}_i^{-1} \right\} \\
&= 2 \left\{ n \log \frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} + \frac{n}{n} \log \prod_{i=1}^n \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} \tilde{m}_i^{-1} \right\} \\
&= 2n \log \frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} \left\{ \prod_{i=1}^n \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} \tilde{m}_i^{-1} \right\}^{\frac{1}{n}} \\
&= n \log \left(\frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} \right)^2 \left\{ \prod_{i=1}^n \frac{\cosh[\lambda(y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda(y_i - \mu_i)]_{\hat{\theta}}} \tilde{m}_i^{-1} \right\}^{\frac{2}{n}} \\
&= n \log \left[\left(\frac{\tilde{\lambda}\hat{\alpha}}{\tilde{\lambda}\tilde{\alpha}} \right)^2 m_g \right].
\end{aligned}$$

Apêndice E: Prova do Teorema 4.4

A estatística da razão de verossimilhanças é dada por

$$\text{LR} = 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\},$$

em que $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$; $\tilde{n}_i = n(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob H_0 . Assim

$$\begin{aligned} \text{LR} &= 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\} \\ &= 2\left\{ \left[\sum_{i=1}^n \log(\lambda n_i) + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\tilde{\boldsymbol{\theta}}} - \left[\sum_{i=1}^n \log \lambda + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\hat{\boldsymbol{\theta}}} \right\}. \end{aligned}$$

Temos que

$$\hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2[\hat{\lambda}(y_i - \hat{\mu}_i)]$$

e

$$\tilde{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2[\tilde{\lambda} \tilde{n}_i (y_i - \tilde{\mu}_i)],$$

em que $\hat{\mu}_i = \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}$ e $\tilde{\mu}_i = \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}}$. Combinando as expressões acima obtemos

$$\begin{aligned} \text{LR} &= 2 \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i}{\hat{\lambda}} + 2 \left\{ \sum_{i=1}^n \log \left[\frac{2}{\alpha} \cosh[\lambda n_i (y_i - \mu_i)] \right] - \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2} \sinh^2[\lambda n_i (y_i - \mu_i)] \right\}_{\tilde{\boldsymbol{\theta}}} \\ &\quad - 2 \left\{ \sum_{i=1}^n \log \left[\frac{2}{\alpha} \cosh[\lambda (y_i - \mu_i)] \right] + \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2} \sinh^2[\lambda (y_i - \mu_i)] \right\}_{\hat{\boldsymbol{\theta}}} \\ &= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i}{\hat{\lambda}} + n \log \frac{2}{\tilde{\alpha}} - n \log \frac{2}{\hat{\alpha}} + \sum_{i=1}^n \log \cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\boldsymbol{\theta}}} \right\} \\ &\quad - 2 \left\{ \sum_{i=1}^n \log \cosh[\lambda (y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} - \frac{2}{\hat{\alpha}^2} \sum_{i=1}^n \sinh^2[\lambda (y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} \right\} \\ &\quad - \frac{4}{\tilde{\alpha}^2} \sum_{i=1}^n \sinh^2[\lambda n_i (y_i - \mu_i)]_{\tilde{\boldsymbol{\theta}}} \end{aligned}$$

$$\begin{aligned}
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \sum_{i=1}^n \log \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda (y_i - \mu_i)]_{\hat{\theta}}} + \frac{n}{2} - \frac{n}{2} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \sum_{i=1}^n \log \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda (y_i - \mu_i)]_{\hat{\theta}}} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \log \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda (y_i - \mu_i)]_{\hat{\theta}}} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \log \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda (y_i - \mu_i)]_{\hat{\theta}}} \right\} \\
&= \sum_{i=1}^n \log \left[\frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} \right]^2 + \log \left\{ \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\cosh[\lambda (y_i - \mu_i)]_{\hat{\theta}}} \right\}^2 \\
&= \left\{ \log \left[\prod_{i=1}^n \left(\frac{\tilde{\lambda} \tilde{n}_i^2 \hat{\alpha}^2}{\hat{\lambda}^2 \tilde{\alpha}^2} m_{g_i} \right) \right] \right\} \text{ com } m_{g_i} = \left[\frac{\cosh[\tilde{\lambda} \tilde{n}_i (y_i - \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}})]}{\cosh[\hat{\lambda} (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})]} \right]^2.
\end{aligned}$$

Apêndice F: Prova do Teorema 4.5

A estatística da razão de verossimilhanças é dada por

$$\text{LR} = 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\},$$

em que $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}^\top, \tilde{\boldsymbol{\vartheta}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\lambda}, \tilde{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$; $\tilde{m}_i = m(\mathbf{X}_i, \tilde{\boldsymbol{\omega}})$ $\tilde{n}_i = n(\mathbf{X}_i, \tilde{\boldsymbol{\vartheta}})$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\omega}}_0^\top, \hat{\boldsymbol{\vartheta}}_0^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\lambda}, \hat{\alpha})^\top$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob H_0 . Assim

$$\begin{aligned} \text{LR} &= 2\{l(\tilde{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}})\} \\ &= 2\left\{ \left[\sum_{i=1}^n \log(\lambda n_i) + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\tilde{\boldsymbol{\theta}}} - \left[\sum_{i=1}^n \log \lambda + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2 \right]_{\hat{\boldsymbol{\theta}}} \right\}. \end{aligned}$$

Temos que

$$\hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2[\hat{\lambda}(y_i - \hat{\mu}_i)]$$

e

$$\tilde{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \frac{1}{\tilde{m}_i^2} \sinh^2[\tilde{\lambda} \tilde{n}_i (y_i - \tilde{\mu}_i)],$$

em que $\hat{\mu}_i = \mathbf{X}_i^\top \hat{\boldsymbol{\beta}}$ e $\tilde{\mu}_i = \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}}$. Combinando as expressões acima obtemos

$$\begin{aligned} \text{LR} &= 2 \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i}{\hat{\lambda}} + 2 \left\{ \sum_{i=1}^n \log \left[\frac{2}{\alpha m_i} \cosh[\lambda n_i (y_i - \mu_i)] \right] - \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2 m_i^2} \sinh^2[\lambda n_i (y_i - \mu_i)] \right\}_{\tilde{\boldsymbol{\theta}}} \\ &\quad - 2 \left\{ \sum_{i=1}^n \log \left[\frac{2}{\alpha} \cosh[\lambda (y_i - \mu_i)] \right] + \frac{1}{2} \sum_{i=1}^n \frac{4}{\alpha^2} \sinh^2[\lambda (y_i - \mu_i)] \right\}_{\hat{\boldsymbol{\theta}}} \\ &= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i}{\hat{\lambda}} + \sum_{i=1}^n \log \frac{2}{\tilde{\alpha} \tilde{m}_i} - n \log \frac{2}{\hat{\alpha}} + \sum_{i=1}^n \log \cosh[\lambda n_i (y_i - \mu_i)] \right\}_{\tilde{\boldsymbol{\theta}}} \\ &\quad - 2 \left\{ \sum_{i=1}^n \log \cosh[\lambda (y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} - \frac{2}{\hat{\alpha}^2} \sum_{i=1}^n \sinh^2[\lambda (y_i - \mu_i)]_{\hat{\boldsymbol{\theta}}} \right\} \\ &\quad - \frac{4}{\tilde{\alpha}^2} \sum_{i=1}^n \frac{1}{\tilde{m}_i^2} \sinh^2[\lambda n_i (y_i - \mu_i)]_{\tilde{\boldsymbol{\theta}}} \end{aligned}$$

$$\begin{aligned}
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \sum_{i=1}^n \log \left[\frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\tilde{m}_i \cosh[\lambda (y_i - \mu_i)]_{\tilde{\theta}}} \right] + \frac{n}{2} - \frac{n}{2} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \sum_{i=1}^n \log \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\tilde{m}_i \cosh[\lambda (y_i - \mu_i)]_{\tilde{\theta}}} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \log \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\tilde{m}_i \cosh[\lambda (y_i - \mu_i)]_{\tilde{\theta}}} \right\} \\
&= 2 \left\{ \sum_{i=1}^n \log \frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} + \log \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\tilde{m}_i \cosh[\lambda (y_i - \mu_i)]_{\tilde{\theta}}} \right\} \\
&= \sum_{i=1}^n \log \left[\frac{\tilde{\lambda} \tilde{n}_i \hat{\alpha}}{\hat{\lambda} \tilde{\alpha}} \right]^2 + \log \left\{ \prod_{i=1}^n \frac{\cosh[\lambda n_i (y_i - \mu_i)]_{\tilde{\theta}}}{\tilde{m}_i \cosh[\lambda (y_i - \mu_i)]_{\tilde{\theta}}} \right\}^2 \\
&= \left\{ \log \left[\prod_{i=1}^n \left(\frac{\tilde{\lambda} \tilde{n}_i^2 \hat{\alpha}^2}{\hat{\lambda}^2 \tilde{\alpha}^2} m_{g_i} \right) \right] \right\} \text{ com } m_{g_i} = \left[\frac{\cosh[\tilde{\lambda} \tilde{n}_i (y_i - \mathbf{X}_i^\top \tilde{\boldsymbol{\beta}})]}{\tilde{m}_i \cosh[\tilde{\lambda} (y_i - \mathbf{X}_i^\top \hat{\boldsymbol{\beta}})]} \right]^2.
\end{aligned}$$

Referências

- Abramowitz, M. e Stegun, I., A. (1972) Handbook of Mathematical Functions. *New York: Dover Press*.
- Achcar, J. A. (1993) Inference for the Birnbaum–Saunders fatigue life model using Bayesian methods. *Computational Statistics and Data Analysis* **15**, 367–380.
- Ahmed, S. E., Castro-Kuriss, C., Flores, E., Leiva, V. e Sanhueza, A. (2010) A truncated version of the Birnbaum-Saunders distribution with an application in financial risk. *Pak. J. Statistics*, Vol. **26**(1), 293-311.
- Azevedo, C., Leiva, V., Athayde, E., Balakrishnan, N. (2012) Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation. *Computational Statistics and Data Analysis* **56**, 3887-3897.
- Balakrishnan, N., Leiva, V. e López, J. (2007) Acceptance Sampling Plans from Truncated Life Tests Based on the Generalized Birnbaum–Saunders Distribution. *Communications in Statistics Simulation and Computation*, **36**, 643-656.
- Barros, M., Paula, G.A. e Leiva, V. (2008) A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis* **14**, 1-17.
- Birnbaum, Z. W. e Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability* **6**, 319-327.
- Birnbaum, Z. W. e Saunders, S. C. (1969b). Estimation for a family of life distributions with application to fatigue. *Journal of Applied Probability* **6**, 328-347.
- Chang, D., S. e Tang, L., C. (1994). Random Number Generator for the Birnbaum-Saunders Distribution. *Computers ind. Engng* **27**, 345-348.
- Christensen, R., Pearson, L. e Johnson, W. (1992). Case-deletion diagnostics for mixed models. *Technometrics* **34**, 3845.

-
- Cisternas, M., Leiva, V. (2012) Recurrence of giant tsunamis in the Quintero bay, Chile. Technical Report DEUV-01-2012.
- Colosimo, E. A.; Giolo, S. R. (2006) *Análise de Sobrevivência Aplicada*. São Paulo: ABE - Projeto Fisher, Edgard Blücher;
- Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, **19**, 15-18.
- Cook, R.D. (1986). Assessment of local influence (with discussion). *J. R. Statist. Soc. B* **48**, 133-169.
- Cook, R.D. e Weisberg, S. (1982). Residuals and Influence in Regression. *Chapman & Hall*, New York.
- Davison, A. C. e Tsai, C. L. (1992). Regression model diagnostics. *International Statistical Review* **60**, 337-353.
- Desmond, A. (1985). Stochastic models of failure in random environments. *Canad. J. Statist.* **13**, 171-183.
- Desmond, A. (1986). On the relationship between two fatigue life models. *IEEE Transactions on Reliability*, **35**, 167-169.
- Díaz-García, J. A. e Leiva, V. (2005). A new family of life distributions based on the contoured elliptically distributions. *Journal of Statistical Planning and Inference* **128**, 88-95.
- Díaz-García, J. A. e Leiva, V. (2006). Erratum to a new family of life distributions based on the elliptically contoured distributions [Journal of Statistical Planning and Inference 128(2) (2005) 445-457]. To appear in *Journal of Statistical Planning and Inference*.
- Dupuis, D. J. e Mills, J. E. (1998). Robust estimation of the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability* **47**, 88-95.
- Edgeworth, F. Y. (1898). *Journal of the Royal Statistical Society* **14**, 670.
- Engelhardt, M., Bain, L. J. e Wright, F. T. (1981). Inferences on the parameters of the Birnbaum-Saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics* **23**, 251-255.

-
- Ferreira, M., Gomes, M. I. e Leiva, V. (2012). On an extreme value version of the birnbaum-Saunders distribution. *REVSTAT Statistical Journal* **10-2**, 181210.
- Fierro, R., Leiva, V., Ruggeri, F. e Sanhuezad, A. (2013) On a Birnbaum-Saunders distribution arising from a non-homogeneous Poisson process. *Statistics and Probability Letters* **83**, 1233-1239.
- Galea, M., Leiva, V., and Paula, G. A. (2004) Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, **31**, 1049-1064.
- Glasser, L., Kohl K. T., Koutschan, C., Moll, V. H. e Straub A. (2012) The integrals in Gradshteyn and Ryzhik. Part 22: Bessel-K functions. *Mathematical Sciences*, **22**, 129-151.
- Gomes, M.I., Ferreira, M., Leiva, V. (2012) The extreme value Birnbaum-Saunders model and its moments and application in biometry. *Biometrical Lett* **49**, 81-94.
- Guiraud, P., Leiva, V., and Fierro, R. (2009) A non-central version of the Birnbaum-Saunders distribution for reliability analysis. *IEEE Transactions on Reliability* (accepted).
- Johnson, N., L. (1949) Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*, **36**, 149-176.
- Lehmann, E., L. e Casella, G. (1998). *Theory of Point Estimation* - Second Edition. Springer Texts in Statistics.
- Leiva, V., Athayde, E., Azevedo, C. e Marchant, C. (2011) Modeling wind energy flux by a Birnbaum-Saunders distribution with an unknown shift parameter. *Journal of Applied Statistics* **38**, 2819-2838.
- Leiva, V., Barros, M., Paula, G. A. e Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis* **51**, 5694-5707.
- Leiva, V., Barros, M., Paula, G.A., Sanhueza, A. (2008) Generalized Birnbaum-Saunders .distributions applied to air pollutant concentration. *Environmetrics* **19**, 235-249.
- Leiva, V., Guadalupe, P., Marchant, C. e Bustos, O. (2012) Fatigue Statistical Distributions Useful for Modeling Diameter and Mortality of Trees. *Revista Colombiana de Estadística*, 349-367.

-
- Leiva, V., Sanhueza, A., Angulo, J.M. (2009) A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stoch Envir Res Risk Assess* **23**, 299-307.
- Leiva, V., Vilca, F., Balakrishnan, N., Sanhueza, A. (2010) A skewed sinh-normal distribution and its properties and application to air pollution. *Comm Stat* **39**, 426-443.
- Leiva, V. Santos-Neto, M., Cysneiros, F.J.A. , Barros, M. (2014) Birnbaum-Saunders statistical modelling: a new approach. *Statistical Modelling* (in print)
- Leiva, V., Saunders, S. New developments and applications on life distributions under cumulative damage. Technical Report DEUV-04-2009.
- Lemonte, A.J., Cribari-Neto, F., Vasconcellos, K.L.P. (2007) Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* **51**, 4656–4681.
- Lemonte, A.J., Patriota, A.G. (2011) Influence diagnostics in Birnbaum-Saunders nonlinear regression models. *J. Amer. Statist. Soc.* **38**, 871–884.
- Lu, M. e Chang, D. S. (1997). Bootstrap prediction intervals for the Birnbaum-Saunders distribution. *Microelectron Reliability* **37**, 1213-1216.
- Marchant, C., Leiva, V., Cavieres, M.F., Sanhueza, A. (2013) Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Reviews Envir Contam Toxic* **223**, 1-31.
- Meeker W. Q. and Escobar L. A. (1998) Statistical Methods for Reliability Data. *John Wiley & sons*, New York.
- Meintanis, S., G. (2010) Inference procedures for the Birnbaum-Saunders distribution and its generalizations. *Computational Statistics and Data Analysis* **54**, 367–373.
- Ng, H. K. T., Kundu, D. e Balakrishnan, N. (2003). Modified moment estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* **43**, 283-298.
- Owen, W.J., (2006). A new three-parameter extension to the Birnbaum-Saunders distribution. *IEEE Trans. Reliab.* **55**, 475–479.
- Owen, W. J. e Padgett, W. J. (1999). Accelerated test models for system strength based on Birnbaum-Saunders distribution. *Lifetime Data Analysis* **5**, 133-147.

-
- Podlaski, R. (2008) Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. *Canadian Journal Forest Res* **18**, 518–527.
- Rieck, J. R. (1989) Statistical Analysis for the Birnbaum–Saunders Fatigue Life Distribution. *Unpublished Ph.D. thesis, Clemson University, Department of Mathematical Sciences.*
- Rieck, J. R., Nedelman, J.R., (1991). A log–linear model for the BirnbaumSaunders distribution. *Technometrics* **33**, 51-60.
- Santos-Neto, M., Cysneiros, F.J.A. Leiva, V., Ahmed, S.E. (2012) On new parameterizations of the Birnbaum-Saunders distribution. *Pak. J. Statist.* **28**, 1–26.
- Saulo, H., Leiva, V., Ziegelmann, F.A., Marchant, C. (2013) A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stoch Envir Res Risk Assess* **27**, 1479-1491.
- Tang, N. S., Wei, B. C. e Wang, X. R. (2000). Influence diagnostics in nonlinear reproductive dispersion models. *Statist. Probab. Lett.* **26**, 5968.
- Tsionas, E. G. (2001) Bayesian inference in Birnbaum-Saunders regression. *Communications in Statistics - Theory and Methods*, **30**, 179-193.
- Vanegas, L.H., Rondon, L.M., Cysneiros, F.J.A. (2012) Diagnostic procedures in Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis* **56**, 1662–1680.
- Vilca-Labra, F. e Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics: Theory and Methods* **35**, 1-16.
- Vilca, F., Sanhueza, A., Leiva, V. e Christakos, G. (2010) An extended BirnbaumSaunders model and its application in the study of environmental quality in Santiago, Chile. *Stoch Environ Res Risk Assess* **24**, 771-782.
- Vilca, F., Santana, L., Leiva, V. e Balakrishnan, N. (2011) Estimation of extreme percentiles in BirnbaumSaunders distributions. *Computational Statistics and Data Analysis* **55**, 1665-1678.
- Volodin, A. I. (2002) Point Estimation, Confidence Sets, and Bootstrapping in some Statistical Models. *Unpublished Ph.D. thesis. University of Regina, Regina, Saskatchewan, Canada.*

- Volodin, I. N. and Dzhungurova, O. A. (2000) On limit distribution emerging in the generalized Birnbaum-Saunders model. *Journal of Mathematical Science*, **99**, 1348-1366.
- Xie, Feng-Chang e Wei, Bo-Cheng, (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Sattistics and Data Analysis*. **51**, 4692-4706.
- Wei, B.C. (1998). *Exponential Family Nonlinear Models*. Springer, Singapore.