
AJUSTES PARA VEROSSIMILHANÇA PERFILADA NA DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

CARLOS ANTÔNIO GADELHA DE ARAÚJO JÚNIOR

Orientadora: Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros

Co-orientador: Prof. Dr. Francisco Cribari-Neto

Área de concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do
grau de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife, fevereiro de 2006

Araújo Júnior, Carlos Antônio Gadelha de
Ajustes para verossimilhança perfilada na
distribuição Birnbaum-Saunders / Carlos Antônio
Gadelha de Araújo Júnior. – Recife : O Autor, 2006.
viii, 78 folhas : il., fig., quadros.

Dissertação (mestrado) – Universidade Federal de
Pernambuco. CCEN. Estatística, 2006.

Inclui bibliografia e apêndices.

1. Estatística matemática – Inferência estatística. 2.
Distribuição Birnbaum-Saunders – Verossimilhança
perfilada – Ajustes. 3. Verossimilhança perfilada
modificada – Teste da razão de verossimilhança. I.
Título.

519.26
519.54

CDU (2.ed.)
CDD (22.ed.)

UFPE
BC2006-114

Universidade Federal de Pernambuco

Mestrado em Estatística

14 de fevereiro de 2006

(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Carlos Antonio Gadelha de Araujo Junior

intitulada

Ajustes para Verossimilhança Perfilada na Distribuição Birnbaum-Saunders

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Klaus Leite Pinto Vasconcellos

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:



Prof. Klaus L. P. Vasconcellos
Coordenador da
Pós-graduação em
Estatística, UFPE

Audrey Helen Mariz de Aquino Cysneiros

Audrey Helen Mariz de Aquino Cysneiros

orientador

Denise Aparecida Botter

Denise Aparecida Botter (USP)

Klaus Leite Pinto Vasconcellos

Klaus Leite Pinto Vasconcellos

Este documento será anexado à versão final da dissertação.

Agradecimentos

- Primeiramente a **Deus**, por ter permitido que eu conquistasse mais esta difícil etapa da minha vida.
- À professora **Audrey Cysneiros**, pela incrível paciência, apoio e excelente orientação na elaboração desta dissertação. Mais ainda, pela dedicação incondicional na trajetória final deste trabalho.
- Ao professor **Francisco Cribari**, pela enorme contribuição dada a minha formação acadêmica desde a graduação até o presente momento. E pela orientação e leitura cuidadosa deste trabalho.
- Aos meus pais **Carlos** e **Joselene**, que contribuíram de forma incondicional para a minha formação moral e acadêmica durante toda minha vida. E também aos meus irmãos **Petterson** e **Márcio**, por tudo que nós compartilhamos juntos.
- Ao meu tio **José Emiliano** (*in memoriam*), que sempre esteve muito presente na minha vida. Também pela amizade, incentivo e pelas boas horas de descontração.
- A todos os meus companheiros do Mestrado, principalmente aos meus grandes amigos **Renata**, **Tiago**, **Milena** e **Daniela** pela amizade compartilhada durante estes dois anos de luta.
- Aos meus grandes amigos **Luiz Henrique** e **Marcela Verônica**, que sempre me apoiaram e me incentivaram nos momentos mais difíceis.
- A todos os professores e funcionários do Departamento de Estatística, em especial à **Valéria** e às professoras **Jacira Guiro Marino** e **Cláudia Lima**.
- À **CAPES**, pelo apoio financeiro.

Resumo

A presente dissertação desenvolve ajustes para a função de verossimilhança perfilada na distribuição Birnbaum-Saunders. Consideramos os ajustes da função de verossimilhança perfilada propostos por Cox e Reid (1987) e Barndorff-Nielsen (1983). No que diz respeito ao parâmetro de interesse, estas versões modificadas da função de verossimilhança foram obtidas para os parâmetros de forma e escala da distribuição. Obtivemos também os estimadores de máxima verossimilhança perfilados modificados a partir da maximização das versões ajustadas da função de verossimilhança. Em seguida, realizamos diversas simulações envolvendo os diferentes estimadores. Adicionalmente, avaliamos numericamente o comportamento dos testes da razão de verossimilhanças baseados nas funções de verossimilhanças perfiladas modificadas de Cox e Reid e Barndorff-Nielsen em amostras finitas. Tanto os testes quanto os estimadores baseados nas versões modificadas da verossimilhança perfilada que foram desenvolvidas neste trabalho apresentaram desempenhos superiores às suas contrapartidas não-modificadas. Por fim, realizamos duas aplicações empíricas.

Abstract

This dissertation obtains adjustments to the profile likelihood function of the Birnbaum-Saunders distribution. We consider the adjustments proposed by Cox and Reid (1987) and by Barndorff-Nielsen (1983). The modified versions of the likelihood function were obtained for both the shape and the scale parameter. We obtained the modified profile maximum likelihood estimators through maximization of the adjusted likelihood functions. We have performed several simulations involving the different estimators. Moreover, we have numerically evaluated the finite-sample performance of likelihood ratio tests based on the modified profile likelihood functions. The estimators and tests obtained from the adjusted profile likelihood functions displayed superior finite-sample behavior. Finally, we present two empirical applications.

Sumário

1	Introdução	1
1.1	Introdução	1
1.2	Organização da dissertação	2
1.3	Suporte computacional	3
2	Verossimilhança perfilada	4
2.1	Introdução	4
2.2	Definições básicas	5
2.3	Verossimilhança perfilada	8
2.4	Ajustes para verossimilhança perfilada	10
2.4.1	Verossimilhança perfilada de Barndorff-Nielsen	10
2.4.2	Verossimilhança perfilada de Cox e Reid	13
2.4.3	Aproximações para a verossimilhança de Barndorff-Nielsen	16
3	Distribuição Birnbaum-Saunders	20
3.1	Introdução	20
3.2	Definições básicas	21
3.3	Estimação dos parâmetros	24
3.4	Ajustes para a Birnbaum-Saunders	26
4	Resultados numéricos	32
4.1	Comportamento dos estimadores	33
4.2	Desempenhos dos testes	38
5	Aplicações	52

5.1	Exemplo 1 - Fadiga do alumínio 6061-T6	52
5.2	Exemplo 2 - Fadiga da chumaceira	54
6	Conclusões	55
	Apêndice A - Cálculo dos ajustes	57
	Apêndice B - Programa de simulação	63
	Referências	75

Lista de quadros

4.1	Resultados de estimação pontual para α , modelo indexado pelos parâmetros $\alpha = 0,1, 0,5$ e $\beta = 1$	34
4.2	Resultados de estimação pontual para α , modelo indexado pelos parâmetros $\alpha = 1, 2$ e $\beta = 1$	35
4.3	Resultados de estimação pontual para β , modelo indexado pelos parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2	37
4.4	Taxas de rejeição (%) calculadas a partir de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2	39
4.5	Média e variância de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$	41
4.6	Quantis de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$	42
4.7	Número de rejeições em 100.000 amostras de Monte Carlo para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.	42
4.8	Poderes dos testes calculados de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1$	43
4.9	Taxas de rejeição (%) calculadas a partir de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2	46
4.10	Média e variância de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$	48

4.11	Quantis de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$	48
4.12	Números de rejeições em 10.0000 amostras de Monte Carlo para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$..	48
4.13	Poderes dos testes calculados de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com pa- râmetros $\beta = 1$ e $\alpha = 0,1$	49
5.1	Número de ciclos até a falha da lâmina do alumínio 6061 - T6 ..	53
5.2	Tempo de fadiga em horas das chumaceiras	54

Lista de figuras

3.1	Gráficos das funções de densidade (a) e de distribuição acumulada (b) Birnbaum-Saunders para $\beta = 4$ e diferentes valores do parâmetro α	23
3.2	Gráfico da função de taxas de falhas da Birnbaum-Saunders para $\beta = 4$ e diferentes valores do parâmetro α	23
4.1	Gráfico das discrepâncias relativas de quantis em relação a α ($n = 10, \alpha = 0,5$ e $\beta = 1$)	44
4.2	Gráfico das discrepâncias relativas de níveis descritivos em relação a α ($n = 10, \alpha = 0,5$ e $\beta = 1$)	44
4.3	Gráfico das discrepâncias relativas de quantis em relação a α ($n = 10, \alpha = 0,5$ e $\beta = 1$)	50
4.4	Gráfico das discrepâncias relativas de níveis descritivos em relação a α ($n = 10, \alpha = 0,5$ e $\beta = 1$)	50

Introdução

1.1 Introdução

Birnbaum e Saunders (1969a) propuseram uma distribuição para modelagem do tempo de falha de materiais e equipamentos. Esta distribuição, que ficou conhecida como distribuição Birnbaum-Saunders, vem sendo largamente utilizada em engenharia como uma forma eficaz de modelar desgastes estocásticos em geral. Ela tem ainda assumido um papel de destaque na modelagem do tempo de falha de processos de fadiga. A distribuição Birnbaum-Saunders é definida pelos seus parâmetros de forma e escala.

É comum que inferências para um determinado modelo envolvam apenas alguns, mas não todos os parâmetros do modelo. Nestes casos dizemos que os parâmetros envolvidos são *parâmetros de interesse*, enquanto que os demais são chamados de *parâmetros de perturbação*. Para a distribuição Birnbaum-Saunders, por se tratar de uma distribuição biparamétrica, podemos considerar dois casos. O primeiro caso ocorre quando o parâmetro de forma é o parâmetro de interesse e o parâmetro de escala é o parâmetro de perturbação. No segundo caso, a situação é inversa, ou seja, os parâmetros de escala e de forma da distribuição são considerados como parâmetros de interesse e de perturbação, respectivamente.

Inferências em um determinado modelo envolvendo parâmetros de perturbação podem ser feitas através da função de verossimilhança perfilada. Esta verossimilhança pode ser obtida substituindo o parâmetro de perturbação por uma estimativa consistente na verossimilhança original. A função resultante dependerá, portanto, apenas do parâmetro de interesse. Algumas propriedades básicas das funções de

verossimilhança não são válidas para a verossimilhança perfilada, dado que esta não é uma verossimilhança genuína. Alguns problemas inferenciais, como, por exemplo, ineficiência e inconsistência dos estimadores, podem surgir quando utilizamos a função de verossimilhança perfilada. Dessa forma, torna-se necessário obter ajustes para a verossimilhança perfilada com o intuito de minimizar estes problemas.

Na literatura, existem várias modificações para a função de verossimilhança perfilada propostas por diversos autores, ver Barndorff-Nielsen (1983, 1994), Cox e Reid (1987, 1992), McCullagh e Tibshirani (1990) e Stern (1997). Estas modificações consistem na incorporação de um termo à verossimilhança perfilada anteriormente à estimação e têm o efeito de diminuir os vieses da função escore e da informação.

A principal contribuição deste trabalho reside na obtenção de alguns destes ajustes para a função de verossimilhança perfilada da distribuição Birnbaum-Saunders, considerando as duas situações no que diz respeito ao parâmetro de interesse. Realizamos diversas simulações com o intuito de comparar a eficácia dos ajustes considerados.

1.2 Organização da dissertação

A presente dissertação de mestrado está dividida em seis capítulos. No segundo capítulo, apresentamos uma revisão de alguns conceitos básicos relacionados à inferência estatística, como por exemplo função de verossimilhança, estimação de parâmetros e testes de hipóteses. Além disso, apresentamos a definição de função de verossimilhança perfilada, suas propriedades e as modificações propostas por Barndorff-Nielsen (1983), Cox e Reid (1987) e Severini (1998, 1999).

No terceiro capítulo, apresentamos uma breve introdução sobre a distribuição Birnbaum-Saunders. Em seguida, desenvolvemos ajustes para a função de verossimilhança perfilada desta distribuição considerando os casos em que temos como parâmetro de interesse o parâmetro de forma e também o parâmetro de escala. São apresentados testes da razão de verossimilhanças baseados nos ajustes obtidos. No quarto capítulo, com base em simulações de Monte Carlo, avaliamos os comporta-

mentos dos testes da razão verossimilhanças e também dos estimadores de máxima verossimilhança baseados nas funções de verossimilhança perfilada ajustadas obtidas no Capítulo 3. No quinto capítulo, aplicamos os desenvolvimentos apresentados neste trabalho a dois conjuntos de dados reais e comparamos os resultados obtidos no que diz respeito à estimação dos parâmetros e a testes de hipóteses. Por fim, no sexto capítulo, apresentamos as conclusões finais de todos os resultados obtidos neste trabalho.

1.3 Suporte computacional

As simulações computacionais realizadas ao longo desta dissertação foram feitas utilizando a versão 3.30 da linguagem matricial `Ox` de programação para o sistema operacional Windows. A linguagem `Ox` está disponível gratuitamente para uso acadêmico no endereço <http://www.doornik.com>. Trata-se de uma linguagem orientada a objetos com sintaxe similar às sintaxes das linguagens `C` e `C++` de programação. Doornik e Ooms (2001) pode ser utilizado como documentação introdutória. Mais detalhes sobre a linguagem `Ox` podem ser encontrados em Doornik (2001) e em Cribari-Neto e Zarkos (2003).

As apresentações gráficas foram produzidas através do ambiente de programação `R`, tendo sido utilizada a versão 2.0.1 para a plataforma Windows. `R` é um ambiente integrado de softwares que possui grandes facilidades para manipulação de dados, geração de gráficos e modelagem estatística em geral. A linguagem e seus pacotes podem ser obtidos gratuitamente no endereço <http://www.r-project.org>. Maiores detalhes podem ser obtidos em Cribari-Neto e Zarkos (1999) e em Ihaka e Gentleman (1996).

Por fim, a presente dissertação foi digitada utilizando o sistema tipográfico (Plain) $\text{T}_{\text{E}}\text{X}$ desenvolvido por Donald Knuth. Uma implementação do $\text{T}_{\text{E}}\text{X}$ para a plataforma Windows (MikTeX) encontra-se disponível em <http://www.miktex.org>. Mais detalhes podem ser obtidos em Knuth (1986).

Verossimilhança perfilada

2.1 Introdução

Em muitas situações, desejamos realizar inferências em um determinado modelo envolvendo alguns, mas não todos os parâmetros. Nesse caso, dizemos que os parâmetros sobre os quais a inferência será feita são de interesse e os demais parâmetros são de perturbação. Quando o número de parâmetros de perturbação é grande, pode ser mais adequado basear as inferências sobre o parâmetro de interesse em uma função de verossimilhança marginal ou condicional, que são verossimilhanças genuínas. Entretanto, a construção destas verossimilhanças pode ser feita apenas quando o modelo tem uma estrutura particular, a saber, em famílias exponenciais e famílias de grupo.

Uma outra solução é utilizar uma pseudo-verossimilhança, que é uma função dos dados que depende apenas do parâmetro de interesse, como se fosse uma verossimilhança genuína. A idéia comumente utilizada é substituir o vetor de parâmetros de perturbação por uma estimativa consistente deste na verossimilhança original. A função resultante é conhecida como função de verossimilhança perfilada. Esta função não é deduzida de uma função densidade e, portanto, não é uma verossimilhança genuína. Entretanto, possui propriedades interessantes que a fazem parecer com uma verossimilhança verdadeira. De fato, a função de verossimilhança perfilada pode ser tratada como uma função de verossimilhança genuína. Porém, tal procedimento pode conduzir a alguns problemas, como, por exemplo, inconsistência e ineficiência dos estimadores. Usar a função de verossimilhança perfilada assemelha-se a tratar o parâmetro de perturbação como se fosse conhecido. É claro que isto não é razoável

quando os dados não fornecem muita informação sobre o parâmetro de perturbação, o que usualmente ocorre quando a dimensão do vetor de parâmetros de perturbação é grande. Isto pode prejudicar a qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos. Portanto, ajustes para a função de verossimilhança perfilada são necessários para reduzir os problemas provenientes do fato desta função não ser uma função de verossimilhança genuína.

Diversas modificações para a função de verossimilhança perfilada e para estatísticas da razão de verossimilhanças derivadas de tal função têm sido propostas. Como exemplos de trabalhos recentes neste sentido, podemos mencionar Barndorff-Nielsen (1983, 1994), Cox e Reid (1987, 1992), McCullagh e Tibishirani (1990), Stern (1997), Ferrari e Cribari-Neto (2002), Ferrari et al. (2004) e Cysneiros e Ferrari (2006). Alguns ajustes para a função de verossimilhança perfilada estão descritos em Severini (2000, Capítulo 9); ver também Pace e Salvani (1997, Capítulo 11).

Neste capítulo inicialmente apresentaremos uma revisão de alguns conceitos básicos relacionados a inferência estatística, como, por exemplo, função de verossimilhança, estimação e testes. Em seguida, na Seção 2.3 apresentaremos a definição de verossimilhança perfilada e suas propriedades. Na Seção 2.4 apresentaremos as versões modificadas da função de verossimilhança perfilada propostas por Barndorff-Nielsen (1983) e Cox e Reid (1987) e, também, aproximações para a verossimilhança perfilada modificada de Barndorff-Nielsen.

2.2 Definições básicas

Suponha que y é o valor observado da variável aleatória $Y = (Y_1, \dots, Y_n)^\top$ associada a um experimento aleatório, o qual é representado por um espaço de probabilidade $(\Omega, \mathcal{A}, \mathcal{P})$, sendo Ω o conjunto de todos os possíveis resultados do experimento, $\mathcal{A}$ um σ -álgebra de subconjuntos de Ω e $\mathcal{P}$ uma medida de probabilidade definida nos elementos de $\mathcal{A}$. Suponha que Y seja caracterizada por uma função de densidade ou de probabilidade com forma analítica $f(y; \theta)$ conhecida, mas envolvendo um vetor $\theta = (\theta_1, \dots, \theta_n)^\top$ de parâmetros desconhecidos. Seja $\Theta \subseteq \mathbb{R}^p$ o espaço paramétrico

representando o conjunto de valores possíveis para o vetor θ . A função $f(y; \theta)$ é denominada função do modelo estatístico e define alguma família $\mathcal{F}$ de distribuição de probabilidade.

A função de verossimilhança é definida como sendo igual a função modelo, embora seja interpretada diferentemente como função de θ para y conhecido. Assim,

$$L = L(\theta) = L(\theta; y), \quad \text{onde } \theta \in \Theta.$$

Sejam $Y_1, \dots, Y_n$ n variáveis aleatórias independentes com função de distribuição acumulada (f.d.a.) $F(y_1, \dots, y_n | \theta)$, onde $\theta \in \Theta$ é desconhecido. A função de verossimilhança é $L(\theta) = f(y_1, \dots, y_n | \theta)$. Note que se $Y_1, \dots, Y_n$ são variáveis aleatórias independentes, então

$$L(\theta) = \prod_{i=1}^n f(y_i; \theta).$$

Em muitas situações é mais conveniente trabalhar com o logaritmo da função de verossimilhança de θ , que será denotado por

$$\ell(\theta) = \ell(\theta; y) = \log L(\theta).$$

O valor de θ que maximiza $L(\theta)$ em Θ , isto é, o valor $\hat{\theta}$ tal que $L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$, é chamado de estimativa de máxima verossimilhança (EMV) de θ . Devido à monotonicidade da função logaritmo, o valor de θ que maximiza a função de verossimilhança $L(\theta; y)$ também maximiza $\ell(\theta; y)$. Adicionalmente, no caso uniparamétrico onde Θ é um intervalo da reta e $\ell(\theta)$ é diferenciável, o estimador de máxima verossimilhança é, na maioria dos casos, raiz da equação de verossimilhança $\ell'(\hat{\theta}) = 0$. Em exemplos simples, a solução desta equação pode ser obtida explicitamente. Entretanto, em situações mais complexas, tal solução deve ser obtida por procedimentos numéricos.

A função escore $u(\theta)$ é definida como

$$u(\theta) = \frac{\partial}{\partial \theta} \ell(\theta).$$

Para garantir que o estimador de máxima verossimilhança de θ seja consistente e possua distribuição limite conhecida, assumimos que a função de verossimilhança

satisfaz às condições usuais de regularidade (ver Cordeiro, 1992), o que implica que

$$E \left[\frac{\partial}{\partial \theta} \ell(\theta) \right] = 0.$$

Em suma, o valor esperado da função escore é sempre igual a zero. Um outro resultado importante estabelece que quando as condições de regularidade são satisfeitas temos que

$$E \left\{ \frac{\partial \ell}{\partial \theta} \frac{\partial \ell}{\partial \theta^\top} \right\} = -E \left\{ \frac{\partial^2 \ell}{\partial \theta \partial \theta^\top} \right\}.$$

Esta igualdade é conhecida como igualdade da informação.

A quantidade $j(\theta)$, denominada matriz de informação observada, é dada por

$$j = j(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top},$$

e a quantidade $i(\theta)$, denominada matriz de informação esperada ou matriz de informação de Fisher, é definida por

$$i = i(\theta) = E\{j(\theta)\}.$$

Assumimos que o vetor θ pode ser decomposto como $\theta = (\tau, \phi)$, onde τ tem dimensão k e ϕ tem dimensão $p - k$. Assim, temos a matriz de informação observada particionada

$$j = j(\tau, \phi) = \begin{pmatrix} j_{\tau\tau} & j_{\tau\phi} \\ j_{\phi\tau} & j_{\phi\phi} \end{pmatrix},$$

onde

$$j_{\tau\tau} = -\frac{\partial^2 \ell(\tau, \phi)}{\partial \tau \partial \tau^\top}, \quad j_{\tau\phi} = j_{\phi\tau}^\top = -\frac{\partial^2 \ell(\tau, \phi)}{\partial \tau \partial \phi^\top} \quad \text{e} \quad j_{\phi\phi} = -\frac{\partial^2 \ell(\tau, \phi)}{\partial \phi \partial \phi^\top}.$$

Procedendo de forma análoga, temos para a matriz de informação esperada a seguinte partição:

$$i = i(\tau, \phi) = \begin{pmatrix} i_{\tau\tau} & i_{\tau\phi} \\ i_{\phi\tau} & i_{\phi\phi} \end{pmatrix},$$

onde

$$i_{\tau\tau} = -E \left\{ \frac{\partial^2 \ell(\tau, \phi)}{\partial \tau \partial \tau^\top} \right\}, \quad i_{\tau\phi} = i_{\phi\tau}^\top = -E \left\{ \frac{\partial^2 \ell(\tau, \phi)}{\partial \tau \partial \phi^\top} \right\} \quad \text{e} \quad i_{\phi\phi} = -E \left\{ \frac{\partial^2 \ell(\tau, \phi)}{\partial \phi \partial \phi^\top} \right\}.$$

Em muitas situações estaremos interessados em testar hipóteses sobre uma parte do vetor de parâmetros θ , digamos $H_0 : \tau = \tau_0$ contra $H_1 : \tau \neq \tau_0$. Podemos definir a estatística da razão de verossimilhanças como

$$LR(\tau) = 2\{\ell(\hat{\tau}, \hat{\phi}) - \ell(\tau, \hat{\phi}(\tau))\},$$

onde $\hat{\tau}$ é a estimativa de máxima verossimilhança de τ e $\hat{\phi}(\tau)$ é a estimativa de máxima verossimilhança de ϕ considerando τ fixado. A estatística da razão de verossimilhanças tem, sob a hipótese nula, distribuição assintótica χ_k^2 , onde k é a dimensão do vetor τ .

2.3 Verossimilhança perfilada

Seja um modelo paramétrico $\mathcal{F} = \{f(y; \theta), y \in \mathcal{A}, \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^p$, sendo θ um vetor de parâmetros desconhecidos de dimensão p . Adotando a decomposição (τ, ϕ) para o vetor paramétrico θ , suponha que inferências para o modelo $\mathcal{F}$ envolvem apenas o parâmetro τ . Os parâmetros τ e ϕ são ditos ser de interesse e de perturbação, respectivamente.

A função de verossimilhança perfilada pode ser obtida substituindo, na função de verossimilhança original, o vetor de parâmetros de perturbação ϕ por sua estimativa de máxima verossimilhança para valores fixados do parâmetro de interesse denotada por $\hat{\phi}_\tau$.

Assim, podemos escrever $\hat{\theta}_\tau = (\tau, \hat{\phi}_\tau)$, onde $\hat{\phi}_\tau$ é a solução em ϕ de $\frac{\partial \ell(\tau, \phi)}{\partial \phi} = 0$. A função de verossimilhança perfilada é definida por

$$L_p(\tau) = L(\tau, \hat{\phi}_\tau).$$

Seja $\ell(\tau, \phi; y)$ o logaritmo da função de verossimilhança, que é unicamente maximizado com relação aos parâmetros dado y . Supondo que a estimativa $\hat{\phi}_\tau$ é única, o logaritmo da função de verossimilhança perfilada para τ é definido como

$$\ell_p(\tau) = \ell(\tau, \hat{\phi}_\tau; y) = \sup_{\phi} \ell(\tau, \phi; y).$$

A expressão acima sugere um procedimento de maximização em duas etapas. A primeira etapa consiste em achar o valor único $\hat{\phi}_\tau$ que maximiza $\ell(\tau, \phi)$ com respeito a ϕ supondo τ constante. A segunda etapa visa encontrar o valor de τ que maximiza $\ell_p(\tau) = \log L_p(\tau)$.

Dessa forma, o estimador de máxima verossimilhança perfilado $\hat{\tau}_p$ é obtido no caso continuamente diferenciável como solução da equação

$$\frac{\partial \ell_p(\tau)}{\partial \tau} = 0.$$

A matriz de informação observada perfilada do parâmetro de interesse τ é definida de forma análoga à matriz de informação observada $j(\theta)$, isto é,

$$j_p = j_p(\tau) = -\frac{\partial^2 \ell_p(\tau)}{\partial \tau \partial \tau^\top}.$$

A função de verossimilhança perfilada $L_p(\tau)$ não é uma função de verossimilhança genuína. Em particular, algumas propriedades válidas para a função de verossimilhança não são satisfeitas para a função de verossimilhança perfilada, tais como as propriedades que garantem que a esperança da função escore é sempre nula e que a identidade da informação não apresenta vício. Isto é, as igualdades

$$E\{u_p(\tau)\} = 0 \quad \text{e} \quad E\left\{u_p(\tau)u_p^\top(\tau)\right\} + E\left\{\frac{\partial u_p(\tau)}{\partial \tau}\right\} = 0,$$

onde $u_p(\tau) = \partial \ell_p(\tau)/\partial \tau$ é a função escore perfilada associada a $L_p(\tau)$, não são sempre satisfeitas para a função de verossimilhança perfilada.

A estatística da razão de verossimilhanças perfiladas usada para testar hipóteses sobre τ , com ϕ desconhecido, ou seja, testar $H_0 : \tau = \tau_0$ contra $H_1 : \tau \neq \tau_0$, é

$$\begin{aligned} LR_p &= 2 \{ \ell_p(\hat{\tau}) - \ell_p(\tau) \} \\ &= 2 \left\{ \ell(\hat{\tau}, \hat{\phi}) - \ell(\tau, \hat{\phi}_\tau) \right\}. \end{aligned}$$

Quando as condições usuais de regularidade são satisfeitas, a distribuição assintótica de LR_p sob a hipótese nula é χ_k^2 , onde o número de graus de liberdade k é a dimensão do vetor τ . É interessante notar que a estatística da razão de verossimilhanças LR

é proporcional à diferença entre os logaritmos de duas funções de verossimilhanças perfiladas maximizadas.

2.4 Ajustes para verossimilhança perfilada

Por causa de sua irrestrita aplicabilidade, a verossimilhança perfilada desempenha um papel fundamental na inferência estatística em problemas com parâmetros de perturbação. Ela é uma ferramenta atraente para realizar inferências sobre o parâmetro de interesse τ devido às semelhanças existentes com a função de verossimilhança original. Infelizmente, ela não usufrui de todas as propriedades de uma verossimilhança genuína. Usar $L_p(\tau)$ é agir como se ϕ fosse conhecido e igual a $\hat{\phi}_\tau$. Isto pode não ser razoável quando os dados não contêm muita informação sobre ϕ , fato que usualmente acontece quando a dimensão do parâmetro ϕ é grande. É por esta razão que se faz necessário ajustar a função de verossimilhança perfilada de alguma forma. Diversas aproximações para as funções de verossimilhança marginal e condicional podem ser expressas em termos da verossimilhança perfilada, dando origem a verossimilhanças perfiladas modificadas. A seguir apresentaremos os ajustes propostos por Barndorff-Nielsen (1983) e Cox e Reid (1987, 1993) e também aproximações para verossimilhança perfilada modificada propostas por Severini (1998, 1999).

2.4.1 Verossimilhança perfilada modificada de Barndorff-Nielsen

A função de verossimilhança perfilada modificada introduzida por Barndorff-Nielsen (1983) pode ser derivada como uma aproximação para uma verossimilhança marginal ou condicional para τ , quando uma destas existir. Isto é feito com base na fórmula p^* , que é uma aproximação para a função de densidade condicional do estimador de máxima verossimilhança do vetor de parâmetros do modelo dada uma estatística ancilar. Seja uma função de verossimilhança $L(\theta; s)$ associada a um modelo estatístico hipotético com o vetor paramétrico θ e seja s uma estatística suficiente

minimal para θ , onde s ressalta a dependência desta função das observações da variável modelada. A expressão de p^* é dada por

$$p^*(\hat{\theta}|a; \theta) = c(\theta, a) |j(\theta; \hat{\theta}, a)|^{1/2} \frac{L(\theta; \hat{\theta}, a)}{L(\hat{\theta}; \hat{\theta}, a)},$$

onde a é uma estatística ancilar, $\hat{\theta}$ é o estimador de máxima verossimilhança de θ , $(\hat{\theta}, a)$ é uma função um-a-um de s e $j(\theta; \hat{\theta}, a) = -\partial^2 \log L(\theta; \hat{\theta}, a) / \partial \theta^2$ é a informação observada. Esta aproximação tem erro relativo de ordem $O(n^{-3/2})$ para $\hat{\theta} = \theta + O_p(n^{-1/2})$. Pode-se mostrar que a constante de proporcionalidade $c(\theta, a)$ é igual a $(2\pi)^{-\dim \theta / 2} (1 + O(n^{-1}))$ (Pace e Salvan, 1997, Capítulo 11), em que $\dim \theta$ é a dimensão do vetor paramétrico θ ; esta constante é muitas vezes tomada como sendo $(2\pi)^{-\dim \theta / 2}$, obtendo-se assim uma aproximação para $p^*(\hat{\theta}|a; \theta)$ de ordem $O(n^{-1})$.

Para derivar a função de verossimilhança proposta por Barndorff-Nielsen (1983, 1994), é necessária a existência de uma estatística a_0 tal que, quando τ é considerado fixo, (ϕ_τ, a_0) é suficiente minimal e a_0 é ancilar. Ou seja, a distribuição de a_0 pode depender de τ mas não de ϕ .

Inicialmente consideraremos tal função como aproximação de uma verossimilhança marginal. Suponha que $(\hat{\tau}, \hat{\phi}, a)$ é uma transformação um a um da estatística suficiente minimal do modelo. Aqui, $\hat{\tau}$ e $\hat{\phi}$ são os estimadores de máxima verossimilhança de τ e ϕ , respectivamente, e a é uma estatística ancilar. Assuma também que, fixado τ , $\hat{\tau}$ seja ancilar, portanto, sua distribuição não depende de ϕ . Observe que a distribuição condicional de $\hat{\tau}$ dado a pode ser expressa por

$$p(\hat{\tau}|a; \tau) = \frac{p(\hat{\tau}, \hat{\phi}|a; \tau, \phi)}{p(\hat{\phi}|\hat{\tau}, a; \tau, \phi)}. \quad (2.1)$$

A densidade $p(\hat{\tau}, \hat{\phi}|a; \tau, \phi)$ pode ser aproximada pela fórmula p^* . Assim, temos

$$p^*(\hat{\tau}, \hat{\phi}|a; \tau, \phi) = c_1(\tau, \phi, a) |j(\hat{\tau}, \hat{\phi}; \hat{\tau}, \hat{\phi}, a)|^{1/2} \frac{L(\tau, \phi; \hat{\tau}, \hat{\phi}, a)}{L(\hat{\tau}, \hat{\phi}; \hat{\tau}, \hat{\phi}, a)}. \quad (2.2)$$

A densidade condicional de $\hat{\phi}$ dado $(\hat{\tau}, a)$ pode ser aproximada usando p^* para aproximar a densidade condicional de $\hat{\phi}_\tau$ dado $(\hat{\tau}, a)$ no modelo com τ fixado, transformando esta aproximação em uma aproximação para a densidade condicional de $\hat{\phi}$.

Aplicando a fórmula p^* para o modelo com τ fixado temos

$$p^*(\widehat{\phi}_\tau|\widehat{\tau}, a; \tau, \phi) = c_2(\tau, \phi, a) |j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}_\tau, a)|^{1/2} \frac{L(\tau, \phi; \widehat{\tau}, \widehat{\phi}_\tau, a)}{L(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}_\tau, a)},$$

em que $j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}_\tau, a) = -\partial^2 \ell(\tau, \phi; \widehat{\tau}, \widehat{\phi}_\tau, a) / \partial \phi^2$ avaliada em $(\tau, \widehat{\phi}_\tau)$. Em seguida, realizando uma mudança de variável na expressão acima, obtém-se

$$p^*(\widehat{\phi}|\widehat{\tau}, a; \tau, \phi) = c_2(\tau, \phi, a) |j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)|^{1/2} \frac{L(\tau, \phi; \widehat{\tau}, \widehat{\phi}, a)}{L(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)} \left| \frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} \right|. \quad (2.3)$$

Agora, substituindo as densidades $p(\widehat{\tau}, \widehat{\phi}|a; \tau, \phi)$ e $p(\widehat{\phi}|\widehat{\tau}, a; \tau, \phi)$ por suas expressões aproximadas (2.2) e (2.3) na equação (2.1), obtemos a função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen:

$$L_{BN}(\tau) = \left| \frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} \right|^{-1} |j_{\phi\phi}(\tau, \widehat{\phi}_\tau)|^{-1/2} L_p(\tau),$$

em que $L_p(\tau)$ é a verossimilhança perfilada de τ , $j_{\phi\phi}(\tau, \widehat{\phi}_\tau)$ é a informação observada correspondente a ϕ quando τ é fixado, $\partial \widehat{\phi}_\tau / \partial \widehat{\phi}$ é a matriz de derivadas parciais de $\widehat{\phi}_\tau$ em relação a $\widehat{\phi}$ e as barras representam o determinante.

Esta aproximação da função de verossimilhança marginal possui erro de ordem $O(n^{-3/2})$ quando comparada com a função de verossimilhança marginal exata, caso esta exista. Tomando o logaritmo de L_{BN} obtém-se

$$\ell_{BN}(\tau) = \ell_p(\tau) - \log \left| \frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} \right| - \frac{1}{2} \log |j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)|.$$

O ponto de máximo de $\ell_{BN}(\tau)$ será indicado por $\widehat{\tau}_{BN}$ e $u_{BN}(\tau) = \partial \ell_{BN}(\tau) / \partial \tau$ denotará a função escore correspondente.

O jacobiano

$$\left| \frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} \right|$$

pode ser expresso em termos de uma derivada do espaço amostral do logaritmo da função de verossimilhança. Observe que $\widehat{\phi}_\tau$ satisfaz

$$\frac{\partial \ell(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)}{\partial \phi} = 0.$$

Diferenciando esta expressão com relação a $\widehat{\phi}$, produz-se a igualdade

$$\frac{\partial^2 \ell(\tau, \widehat{\phi}_\tau)}{\partial^2 \phi} \frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} + \frac{\partial}{\partial \widehat{\phi}} \left(\frac{\partial \ell(\tau, \widehat{\phi}_\tau)}{\partial \phi} \right) = 0.$$

Logo,

$$\frac{\partial \widehat{\phi}_\tau}{\partial \widehat{\phi}} = j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)^{-1} l_{\phi; \widehat{\phi}}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a).$$

Para derivar a verossimilhança perfilada modificada como uma aproximação para uma função de verossimilhança condicional precisamos que, fixado τ , $(\widehat{\phi}_\tau, a)$ seja suficiente para o modelo. A função de verossimilhança condicional pode ser baseada em

$$p(\widehat{\tau}|\widehat{\phi}, a; \tau, \phi) = \frac{p(\widehat{\tau}, \widehat{\phi}|a; \tau, \phi)}{p(\widehat{\phi}|a; \tau, \phi)}.$$

Uma aproximação para esta densidade condicional pode ser obtida aproximando as densidades $p(\widehat{\tau}, \widehat{\phi}|a; \tau, \phi)$ e $p(\widehat{\phi}|a; \tau, \phi)$ através da aproximação p^* . Assim, de forma análoga ao procedimento anterior podemos obter $L_{BN}(\tau)$.

2.4.2 Verossimilhança perfilada modificada de Cox e Reid

Cox e Reid (1987) definiram uma versão modificada da verossimilhança perfilada explorando as conseqüências de ortogonalidade entre os parâmetros de interesse e de perturbação. Esta versão modificada é uma aproximação para a função densidade condicional das observações dado o estimador de máxima verossimilhança do parâmetro de perturbação. Seja $\theta = (\tau, \phi)$ o vetor paramétrico para um modelo estatístico com função de verossimilhança $L(\tau, \phi)$. As componentes τ e ϕ de θ são ortogonais se as componentes do vetor escore são assintoticamente não-correlacionadas ou, equivalentemente, se $i_{\tau\phi} = 0$.

A principal conseqüência da ortogonalidade dos parâmetros é a independência assintótica dos estimadores de máxima verossimilhança de τ e ϕ . Dessa forma, a matriz de covariância de $\widehat{\tau}$ pode ser avaliada como se ϕ fosse conhecido. De fato,

$$i^{\tau\tau} = (i_{\tau\tau} - i_{\tau\phi}(i_{\phi\phi})^{-1}i_{\phi\tau})^{-1} = (i_{\tau\tau})^{-1} \quad \text{se } i_{\tau\phi} = 0.$$

Uma outra consequência importante é que $\widehat{\phi}_\tau - \widehat{\phi} = O_p(n^{-1})$, enquanto que usualmente $\widehat{\phi}_\tau - \widehat{\phi} = O_p(n^{-1/2})$. O logaritmo da função de verossimilhança pode ser expandido em série de Taylor em torno de $(\widehat{\tau}, \widehat{\phi})$ como

$$\begin{aligned} \ell(\theta) = & \ell(\widehat{\theta}) - \frac{1}{2} \left\{ (\tau - \widehat{\tau})^2 j_{\tau\tau}(\widehat{\tau}, \widehat{\phi}) + 2(\tau - \widehat{\tau})(\phi - \widehat{\phi}) j_{\tau\phi}(\widehat{\tau}, \widehat{\phi}) + (\phi - \widehat{\phi})^2 j_{\phi\phi}(\widehat{\tau}, \widehat{\phi}) \right\} \\ & + O_p(n^{-1/2}). \end{aligned} \quad (2.4)$$

Se τ e ϕ são ortogonais, $j(\tau, \phi)$ e $j(\widehat{\tau}, \widehat{\phi})$ são de ordem $O_p(n^{1/2})$. Portanto,

$$\begin{aligned} \ell(\tau, \phi) &= c - \frac{1}{2}(\tau - \widehat{\tau})^2 j_{\tau\tau}(\widehat{\tau}, \widehat{\phi}) - \frac{1}{2}(\phi - \widehat{\phi})^2 j_{\phi\phi}(\widehat{\tau}, \widehat{\phi}) + O_p(n^{-1/2}) \\ &= \ell_1(\tau) + \ell_2(\phi) + O_p(n^{-1/2}) \end{aligned}$$

para $\theta = \widehat{\theta} + O_p(n^{-1/2})$, onde c é uma constante que não envolve os parâmetros. De (2.4), sob ortogonalidade dos parâmetros, temos que

$$\ell(\widehat{\tau}, \widehat{\phi}) = \ell(\tau, \phi) - \frac{1}{2} \left\{ (\tau - \widehat{\tau})^2 j_{\tau\tau}(\widehat{\tau}, \widehat{\phi}) + (\phi - \widehat{\phi})^\top i_{\phi\phi}(\tau, \phi)(\phi - \widehat{\phi}) \right\} + O_p(n^{-1/2}).$$

Portanto, a verossimilhança condicional dado $\widehat{\phi}$ pode ser aproximada por

$$\log p(y|\widehat{\phi}; \tau) = c(y) + \ell(\tau, \phi) - \frac{1}{2} \log |j_{\phi\phi}(\tau, \phi)| + \frac{1}{2}(\widehat{\phi} - \phi)^\top (j^{\phi\phi})^{-1}(\widehat{\phi} - \phi).$$

Esta função não é independente de ϕ , pois a suficiência parcial de $\widehat{\phi}$ para ϕ é somente observada assintoticamente; portanto, uma aproximação foi utilizada. Uma solução natural é substituir ϕ por $\widehat{\phi}_\tau$. Através deste procedimento e dado que $\widehat{\phi}_\tau - \widehat{\phi} = O_p(n^{-1})$, a forma quadrática $(\widehat{\phi} - \widehat{\phi}_\tau)^\top (j^{\phi\phi})^{-1}(\widehat{\phi} - \widehat{\phi}_\tau)$ tem ordem $O_p(n^{-1})$ e pode ser desconsiderada.

Assim, o logaritmo da função de verossimilhança proposta por Cox e Reid (1987) é dado por

$$\ell_{CR}(\tau) = \ell(\tau, \widehat{\phi}_\tau) - \frac{1}{2} \log |j_{\phi\phi}(\tau, \widehat{\phi}_\tau)|, \quad (2.5)$$

em que $j_{\phi\phi}(\tau, \widehat{\phi}_\tau)$ é a informação observada correspondente a ϕ quando τ é fixado. Esta é uma versão penalizada do logaritmo da função de verossimilhança perfilada, onde a penalização leva em conta a informação sobre o parâmetro de perturbação.

A função de verossimilhança perfilada modificada proposta por Cox e Reid (1987) pode ser obtida aplicando a função inversa aos dois membros da igualdade (2.5), assim obtemos:

$$L_{CR}(\tau) = |j_{\phi\phi}(\tau, \widehat{\phi}_\tau)|^{-1/2} L_p(\tau).$$

A correspondente função escore é dada por $u_{CR}(\tau) = \partial \ell_{CR}(\tau) / \partial \tau$. O valor que maximiza $\ell_{CR}(\tau)$ será denotado por $\widehat{\tau}_{CR}$.

A obtenção da função $\ell_{CR}(\tau)$ requer a ortogonalidade dos parâmetros τ e ϕ . Quando esta condição não é satisfeita precisamos encontrar uma reparametrização do tipo $(\tau, \lambda(\tau, \phi))$, onde τ e λ sejam ortogonais. Porém, nem sempre é possível encontrar tal reparametrização, exceto para o caso em que o parâmetro de interesse é escalar.

Assim, consideraremos o caso em que τ é um escalar e que $\phi = (\phi_1, \dots, \phi_q)$ é um vetor de parâmetros. Estamos interessados em encontrar uma parametrização (τ, λ) , com $\lambda = \lambda(\tau, \phi) = (\lambda_1(\tau, \phi), \dots, \lambda_q(\tau, \phi))$, tal que τ e λ sejam ortogonais. Denotamos por $\ell^\perp(\tau, \lambda)$ o logaritmo da função de verossimilhança na nova parametrização, i.e.,

$$\ell^\perp(\tau, \lambda) = \ell(\tau, \phi(\tau, \lambda)).$$

Diferenciando os dois lados da igualdade acima com respeito a τ , temos

$$\frac{\partial \ell^\perp}{\partial \tau} = \frac{\partial \ell}{\partial \tau} + \sum_{r=1}^q \frac{\partial \ell}{\partial \phi_r} \frac{\partial \phi_r}{\partial \tau}.$$

Diferenciando com respeito a λ_m , obtemos

$$\frac{\partial^2 \ell^\perp}{\partial \tau \partial \lambda_m} = \sum_{h=1}^q \frac{\partial^2 \ell}{\partial \tau \partial \phi_h} \frac{\partial \phi_h}{\partial \lambda_m} + \sum_{r=1}^q \sum_{h=1}^q \frac{\partial^2 \ell}{\partial \phi_r \partial \phi_h} \frac{\partial \phi_h}{\partial \lambda_m} \frac{\partial \phi_r}{\partial \tau} + \sum_{r=1}^q \frac{\partial \ell}{\partial \phi_r} \frac{\partial^2 \phi_r}{\partial \tau \partial \lambda_m}, \quad (2.6)$$

para $m = 1, \dots, q$.

A parametrização (τ, λ) deve satisfazer

$$E \left\{ \frac{\partial^2 \ell^\perp}{\partial \tau \partial \lambda_m} \right\} = 0,$$

para $m = 1, \dots, q$. Tomando o valor esperado de ambos os lados de (2.6) observamos que a contribuição do terceiro termo do lado direito da equação é nula,

pois $E\{\partial\ell/\partial\phi_r\} = 0$ para todo $r = 1, \dots, q$. Assim, podemos concluir que a parametrização (τ, λ) deve satisfazer o seguinte sistema de equações diferenciais parciais:

$$\sum_{h=1}^q \frac{\partial\phi_h}{\partial\lambda_m} \left(i_{\tau\phi_h} + \sum_{r=1}^q i_{\phi_r\phi_h} \frac{\partial\phi_r}{\partial\tau} \right) = 0, \quad m = 1, \dots, q,$$

isto é,

$$\sum_{r=1}^q i_{\phi_r\phi_h} \frac{\partial\phi_r}{\partial\tau} = -i_{\tau\phi_h}, \quad h = 1, \dots, q,$$

onde

$$i_{\tau\phi_h} = -E \left\{ \frac{\partial^2 \ell}{\partial\tau \partial\phi_h} \right\} \quad \text{e} \quad i_{\phi_r\phi_h} = -E \left\{ \frac{\partial^2 \ell}{\partial\phi_r \partial\phi_h} \right\}.$$

Se $\tau = (\tau_1, \tau_2)$, podemos calcular independentemente $\partial\phi_r/\partial\tau_1$ e $\partial\phi_r/\partial\tau_2$ através destas equações, porém não há garantia de que a condição de compatibilidade $\partial^2\phi_r/\partial\tau_1\partial\tau_2 = \partial^2\phi_r/\partial\tau_2\partial\tau_1$ será satisfeita. Assim, fica claro porque, em geral, não é possível obter ortogonalidade entre os parâmetros quando o parâmetro de interesse não é escalar.

Uma outra desvantagem na utilização de $\ell_{CR}(\tau)$ é que esta função não é invariante sob reparametrizações do tipo $(\tau, \phi) \rightarrow (\zeta, \lambda)$, onde $\zeta = \zeta(\tau)$ e $\lambda = \lambda(\tau, \phi)$, diferentemente de $L_{BN}(\tau)$.

2.4.3 Aproximações para a verossimilhança de Barndorff-Nielsen

Na Seção 2.4.1 vimos que a função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen depende da quantidade $\left| \frac{\partial\widehat{\phi}_\tau}{\partial\widehat{\phi}} \right|^{-1}$, onde

$$\frac{\partial\widehat{\phi}_\tau}{\partial\widehat{\phi}} = j_{\phi\phi}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)^{-1} \ell_{\phi; \widehat{\phi}}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a).$$

A quantidade $\ell_{\phi; \widehat{\phi}}(\tau, \widehat{\phi}_\tau; \widehat{\tau}, \widehat{\phi}, a)$ é uma derivada relativa ao espaço amostral, sendo na maioria das situações de difícil (ou até mesmo impossível) obtenção. Para lidar com esta dificuldade diversas aproximações para a derivada do espaço amostral

do logaritmo da função de verossimilhança têm sido propostas. Nesta seção apresentaremos duas destas aproximações: uma baseada nas covariâncias de derivadas do logaritmo da função de verossimilhança e a outra baseada em covariâncias empíricas.

Inicialmente desenvolveremos uma aproximação para $\ell_{\hat{\theta}}(\theta; \theta_0, a)$ com base nas covariâncias das derivadas do logaritmo da função de verossimilhança. Para tanto, definimos

$$I(\theta; \theta_0) = E[\ell_\theta(\theta)\ell_\theta(\theta_0)^T; \theta_0].$$

Utilizaremos a notação

$$\ell_\theta(\theta) = \ell_\theta(\theta; \hat{\theta}, a) = \frac{\partial \ell(\theta; \hat{\theta}, a)}{\partial \theta} \quad \text{e} \quad \ell_{\hat{\theta}}(\theta) = \ell_{\hat{\theta}}(\theta; \hat{\theta}, a) = \frac{\partial \ell_\theta(\theta; \hat{\theta}, a)}{\partial \hat{\theta}}.$$

Para $\theta = \theta_0 + \delta/\sqrt{n}$,

$$\ell_{\hat{\theta}}(\theta; \theta_0, a) - \ell_{\hat{\theta}}(\theta_0; \theta_0, a) = I(\theta; \theta_0) - I(\theta_0; \theta_0) + O(1)$$

e, assim,

$$\begin{aligned} I(\theta; \theta_0) + \ell_{\hat{\theta}}(\theta; \theta_0, a) - I(\theta_0; \theta_0) &= I(\theta; \theta_0) + j(\theta_0) - i(\theta_0) \\ &= I(\theta; \theta_0) + i(\theta_0)\{i(\theta_0)^{-1}j(\theta_0) - D\} \\ &= I(\theta; \theta_0) + \{I(\theta; \theta_0) + O(n^{1/2})\} \\ &\quad \times \{i(\theta_0)^{-1}j(\theta_0) - D\} \\ &= I(\theta; \theta_0)i(\theta_0)^{-1}j(\theta_0) + O(1), \end{aligned}$$

onde D denota a matriz identidade. Observe que usamos o fato de que

$$\ell_{\hat{\theta}}(\theta; \theta, a) = -\ell_{\theta\theta}(\theta; \theta, a) = j(\theta).$$

Este resultado é obtido derivando a igualdade $\ell_\theta(\hat{\theta}; \hat{\theta}, a) = 0$ com relação a $\hat{\theta}$ e avaliando em $\hat{\theta} = \theta$; assim, tem-se que

$$\ell_\theta(\theta; \theta, a) + \ell_{\hat{\theta}}(\theta; \theta, a) = 0.$$

Também foram usadas as expansões

$$i(\theta_0) = I(\theta; \theta_0) + O(n^{1/2}) \quad \text{e} \quad i(\theta_0)^{-1}j(\theta_0) = D + O(n^{-1/2}).$$

Dessa forma, $\ell_{\theta;\hat{\theta}_0}(\theta; \theta_0, a)$ pode ser aproximado por $I(\theta; \theta_0)i(\theta_0)^{-1}j(\theta_0)$.

Assim, fazendo $\theta_0 = \hat{\theta}$, uma aproximação para $\ell_{\theta;\hat{\theta}}(\theta)$ com erro relativo de ordem $O(n^{-1/2})$ é dada por

$$\bar{\ell}_{\theta;\hat{\theta}}(\theta) = I(\theta; \hat{\theta})i(\hat{\theta})^{-1}j(\hat{\theta}).$$

Como $i(\hat{\theta})^{-1}j(\hat{\theta}) = D + O_p(n^{-1/2})$, costuma-se utilizar $\bar{\ell}_{\theta;\hat{\theta}}(\theta) = I(\theta; \hat{\theta})$.

Este método de aproximação foi usado por Skovgaard (1996) e Severini (1998); ver também Barndorff-Nielsen (1995).

Utilizando um argumento análogo, $\ell_{\phi;\hat{\phi}}(\tau, \hat{\phi}_\tau)$ pode ser aproximado por $I_\phi(\hat{\theta}_\tau; \hat{\theta})$, onde

$$I_\phi(\theta; \theta_0) = I_\phi(\tau, \phi; \tau_0, \phi_0) = E_{(\tau_0, \phi_0)} \left\{ \ell_\phi(\tau, \phi), \ell_\phi(\tau_0, \phi_0)^\top \right\}.$$

Assim, podemos aproximar $\ell_{BN}(\tau)$ por

$$\bar{\ell}_{BN}(\tau) = \ell_p(\tau) + \frac{1}{2} \log |j_{\phi\phi}(\hat{\tau}, \hat{\phi}_\tau)| - \log |I_\phi(\tau, \hat{\phi}_\tau; \hat{\tau}, \hat{\phi})|.$$

em que $\ell_p(\tau)$ é a verossimilhança perfilada de τ , $j_{\phi\phi}(\tau, \hat{\phi}_\tau)$ é a informação observada correspondente a ϕ quando τ é fixado e as barras representam o determinante. Esta aproximação também tem erro relativo de ordem $O(n^{-1/2})$. O correspondente estimador de máxima verossimilhança será denotado por $\hat{\tau}_{BN}$ e a função escore, por $\bar{u}_{BN}(\tau) = \partial \bar{\ell}_{BN}(\tau) / \partial \tau$.

Supondo que os dados consistem de n observações independentes, as covariâncias usadas em $\bar{\ell}_{BN}(\tau)$ podem ser substituídas por covariâncias empíricas sem que a ordem do erro de aproximação seja alterada. Este método de aproximação é baseado em Severini (1999). Seja $\ell^{(j)}(\theta)$ o logaritmo da função de verossimilhança baseado na j -ésima observação; então, $\ell(\theta) = \sum \ell^{(j)}(\theta)$. Dessa maneira, $I(\theta; \theta_0)$ pode ser escrito como

$$I(\theta, \theta_0) = E \left\{ \sum_{j=1}^n \ell_\theta^{(j)}(\theta_0) \ell_\theta^{(j)}(\theta)^\top \right\}$$

e, conseqüentemente, aproximado por

$$\check{I}(\theta, \theta_0) = \sum_{j=1}^n \ell_\theta^{(j)}(\theta_0) \ell_\theta^{(j)}(\theta)^\top.$$

De maneira análoga, temos

$$\check{I}_\phi(\theta; \theta_0) = \check{I}_\phi(\tau, \phi; \tau_0, \phi_0) = \sum_{j=1}^n \ell_\phi^{(j)}(\tau, \phi), \ell_\phi^{(j)}(\tau_0, \phi_0)^\top.$$

Inserindo $\check{I}_\phi(\tau, \phi_\tau; \hat{\tau}, \hat{\phi})$ no lugar de $I_\phi(\tau, \phi_\tau; \hat{\tau}, \hat{\phi})$ na aproximação $\bar{\ell}_{BN}(\tau)$ obtemos uma aproximação para $\ell_{BN}(\tau)$ baseada em covariâncias empíricas:

$$\check{\ell}_{BN}(\tau) = \ell_p(\tau) + \frac{1}{2} \log |j_{\phi\phi}(\hat{\tau}, \hat{\phi}_\tau)| - \log |\check{I}_\phi(\tau, \hat{\phi}_\tau; \hat{\tau}, \hat{\phi})|.$$

O correspondente estimador de máxima verossimilhança será denotado por $\hat{\tau}_{BN}$, a função escore associada sendo $\check{u}_{BN}(\tau) = \partial \check{\ell}_{BN}(\tau) / \partial \tau$. Esta aproximação é fácil de ser obtida numericamente, tornando-se bastante conveniente quando são encontradas dificuldades no cálculo das esperanças dos produtos de derivadas do logaritmo da função de verossimilhança.

Distribuição Birnbaum-Saunders

3.1 Introdução

A distribuição Birnbaum-Saunders foi inicialmente proposta por Birnbaum e Saunders (1969a) para modelar o tempo de falha de um material sujeito a uma mesma configuração de estresse de forma cíclica. A cada ciclo ocorre um aumento na ruptura do material e a falha ocorre quando o tamanho da ruptura atinge um valor limiar. Para chegar à distribuição, uma das suposições feitas por Birnbaum e Saunders (1969a) foi que o tamanho da ruptura seguia uma distribuição normal. Entretanto, Desmond (1985) mostrou que uma variedade de distribuições para o tamanho da ruptura resulta em uma distribuição Birnbaum-Saunders. Desmond (1986) observou que outras suposições feitas por Birnbaum e Saunders (1969a) para derivar a distribuição também poderiam ser relaxadas.

A distribuição Birnbaum-Saunders assumiu um importante papel na modelagem do tempo de falha em processos de fadiga. Também vem sendo largamente utilizada como um meio eficaz de modelar falhas de desgaste estocástico em geral.

Neste capítulo apresentaremos, de forma introdutória, a distribuição Birnbaum-Saunders, incluindo a sua função densidade, função de distribuição acumulada e inferências sobre seus parâmetros. Em seguida, desenvolveremos ajustes para a função de verossimilhança perfilada para o caso em que temos como parâmetro de interesse o parâmetro de forma e também quando o parâmetro de interesse é o parâmetro de escala. Consideraremos os ajustes propostos por Cox e Reid (1987, 1993) e por Barndorff-Nielsen (1983). Apresentaremos alguns testes da razão de verossimilhanças baseados nos ajustes obtidos. Também avaliaremos o comportamento em amostras

finitas dos diferentes testes em relação a tamanho e a poder. Alguns trabalhos recentes sobre a distribuição Birnbaum-Saunders são Chang e Tang (1993, 1994), Lu e Chang (1997), Dupuis e Mills (1998), Ng et al. (2003) e Wu e Wong (2004). Ver também Johnson et al. (1995).

3.2 Definições básicas

Se T é uma variável aleatória com distribuição Birnbaum-Saunders, denotada por $\mathcal{BS}(\alpha, \beta)$, então sua função densidade de probabilidade (f.d.p.) é dada por

$$f(t; \alpha, \beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta} \left[\left(\frac{\beta}{t} \right)^{1/2} + \left(\frac{\beta}{t} \right)^{3/2} \right] \exp \left[-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right], \quad t, \alpha, \beta > 0, \quad (3.1)$$

em que α é o parâmetro de forma e β é o parâmetro de escala. A correspondente função de distribuição acumulada é

$$F(t; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right], \quad 0 < t < \infty, \quad \alpha, \beta > 0, \quad (3.2)$$

em que Φ denota a função de distribuição acumulada da normal padrão.

As Figuras 3.1 (a) e (b) ilustram a função de densidade e a função de distribuição acumulada, respectivamente, para diferentes valores dos parâmetros de forma considerando o parâmetro de escala fixo ($\beta = 4$). Na Figura 3.1 (a), observamos que a densidade (3.1) é assimétrica à direita. Entretanto, a medida que o valor de α decresce, particularmente para valores menores que um, a densidade torna-se mais simétrica. Na Figura 3.1 (b), observamos que para diferentes escolhas do parâmetro α as suas respectivas funções de distribuição apresentam formas bem diferenciadas.

Mann et al. (1974, p. 155) mostraram que a função de densidade Birnbaum-Saunders é unimodal e, embora a taxa de falhas $h(t; \alpha, \beta) = f(t; \alpha, \beta)/[1 - F(t; \alpha, \beta)]$ não seja uma função crescente de t , a taxa de falhas média é próxima de uma função não-decrescente em t . A Figura 3.2 apresenta a taxa de falhas para alguns valores de α considerando β fixo ($\beta = 4$).

Figura 3.1 Gráficos das funções de densidade (a) e de distribuição acumulada (b) Birnbaum-Saunders para $\beta = 4$ e diferentes valores do parâmetro α .

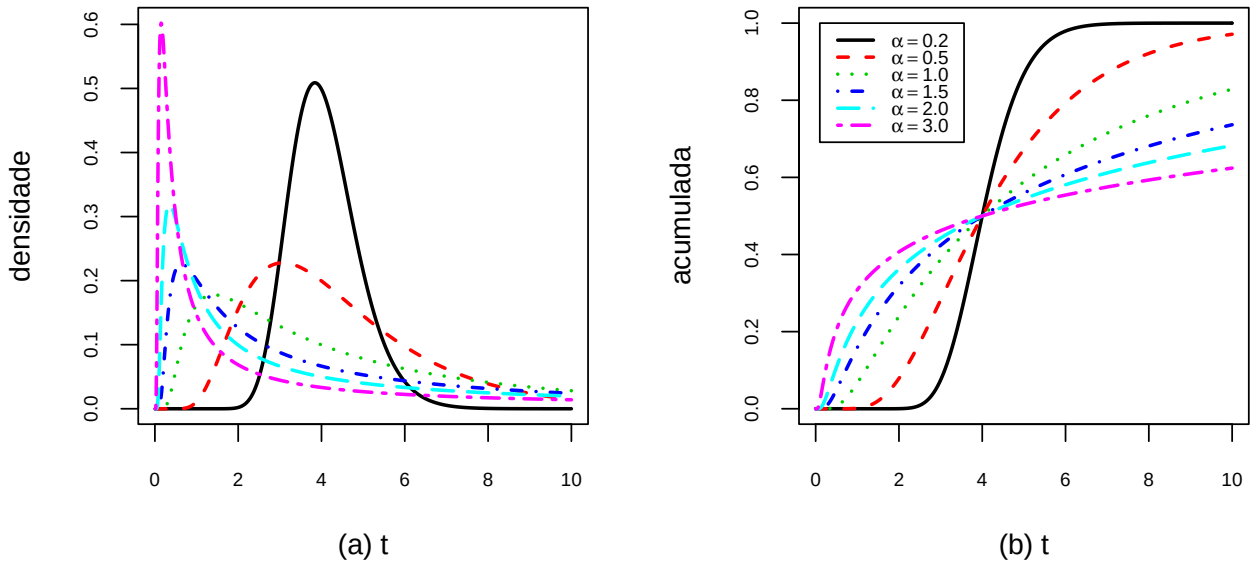
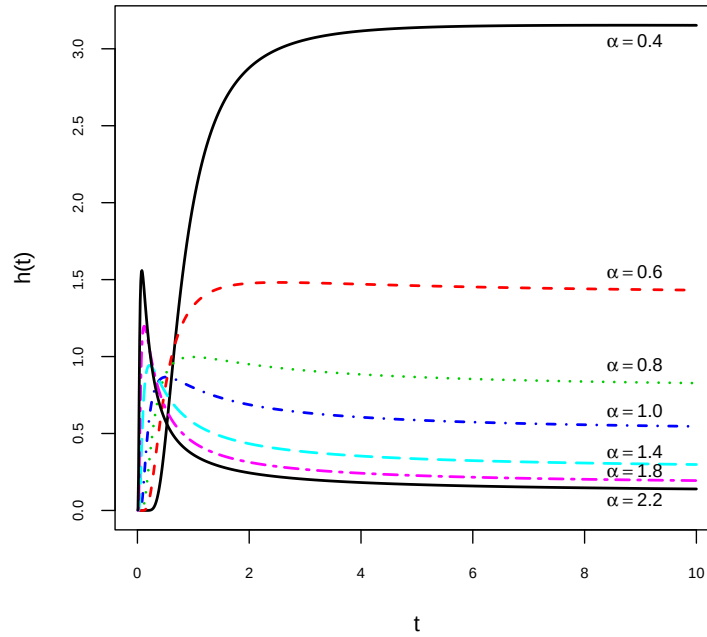


Figura 3.2 Gráfico da função de taxas de falhas da Birnbaum-Saunders para $\beta = 4$ e diferentes valores do parâmetro α .



Se T tem distribuição Birnbaum-Saunders com parâmetros α e β , podemos considerar a seguinte transformação monotônica:

$$X = \frac{1}{2} \left[\left(\frac{T}{\beta} \right)^{1/2} - \left(\frac{T}{\beta} \right)^{-1/2} \right].$$

A partir de (3.2), temos que X tem distribuição normal com média zero e variância $\frac{1}{4}\alpha^2$. Assim, números pseudo-aleatórios provenientes de uma distribuição Birnbaum-Saunders podem ser gerados a partir de uma distribuição normal através da relação

$$T = \beta \{1 + 2X^2 + 2X(1 + X^2)^{1/2}\}.$$

Os momentos de T podem ser obtidos através da expressão abaixo, em que o lado direito desta equação depende apenas de α , pois o parâmetro β é apenas um multiplicador e não afeta a forma da distribuição. O valor esperado de $(T/\beta)^r$ é

$$E \left[\left(\frac{T}{\beta} \right)^r \right] = \sum_{j=0}^r \binom{2r}{2j} \sum_{i=0}^j \binom{j}{i} \frac{[2(r-j+i)]}{2^{r-j+i}(r-j+i)} \left(\frac{1}{2} \alpha \right)^{2(r-j+i)}.$$

Usando a equação acima, a esperança, a variância e os coeficientes de assimetria e curtose podem ser obtidos facilmente como

$$E(T) = \beta \left(1 + \frac{1}{2} \alpha^2 \right),$$

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

$$\beta_1 = \frac{16\alpha^2(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3}$$

e

$$\beta_2 = 3 + \frac{6\alpha^2(93\alpha^2 + 41)}{(5\alpha^2 + 4)^2},$$

respectivamente. A distribuição Birnbaum-Saunders possui algumas propriedades interessantes. O parâmetro de escala β é a mediana da distribuição pois $F_T(\beta) = \Phi(0) = 1/2$. Para qualquer constante $k > 0$, tem-se que $kT \sim \mathcal{BS}(\alpha, k\beta)$. Adicionalmente, a distribuição Birnbaum-Saunders possui propriedade recíproca, isto é,

se T tem distribuição Birnbaum-Saunders com parâmetros α e β , então T^{-1} tem distribuição Birnbaum-Saunders com os parâmetros α e β^{-1} . Portanto,

$$E(T^{-1}) = \beta^{-1}(1 + \frac{1}{2}\alpha^2)$$

e

$$Var(T^{-1}) = \alpha^2\beta^{-2}(1 + \frac{5}{4}\alpha^2).$$

3.3 Estimação dos parâmetros

Seja $t = (t_1, \dots, t_n)^\top$ uma amostra aleatória de tamanho n proveniente de uma distribuição Birnbaum-Saunders com f.d.p. dada em (3.1). O logaritmo da função de verossimilhança para $\theta = (\alpha, \beta)$ pode ser escrito como

$$\ell(\theta) = -n \log(\alpha\beta) + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] - \frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right). \quad (3.3)$$

A função escore total $u(\theta) = (u_\alpha, u_\beta)$ tem componentes dadas por

$$u_\alpha = \frac{\partial \ell(\theta)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{s}{\beta} + \frac{\beta}{r} - 2 \right)$$

e

$$u_\beta = \frac{\partial \ell(\theta)}{\partial \beta} = -\frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right) + \frac{n}{2\alpha^2} \left(\frac{s}{\beta^2} - \frac{1}{r} \right),$$

em que

$$r = \frac{1}{n} \sum_{i=1}^n t_i, \quad s = \left[\frac{1}{n} \sum_{i=1}^n t_i^{-1} \right]^{-1} \quad \text{e} \quad K(\gamma) = \left[\frac{1}{n} \sum_{i=1}^n (\gamma + t_i)^{-1} \right]^{-1}, \quad \text{para } \gamma > 0. \quad (3.4)$$

A matriz de informação observada é dada por

$$j(\theta) = \begin{pmatrix} j_{\alpha\alpha}(\theta) & j_{\alpha\beta}(\theta) \\ j_{\beta\alpha}(\theta) & j_{\beta\beta}(\theta) \end{pmatrix},$$

em que

$$j_{\alpha\alpha}(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = -\frac{n}{\alpha^2} + \frac{3n}{\alpha^4} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right),$$

$$j_{\alpha\beta}(\theta) = j_{\beta\alpha}(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \beta} = -\frac{n}{\alpha^3} \left(-\frac{r}{\beta^2} + \frac{1}{s} \right)$$

e

$$j_{\beta\beta}(\theta) = -\frac{\partial^2 \ell(\theta)}{\partial \beta^2} = -\frac{n}{\beta^2} + \frac{n}{2} \left(\frac{1}{\beta^2} + \frac{2K'(\beta)}{K^2(\beta)} \right) + \frac{n}{\alpha^2} \frac{r}{\beta^3},$$

sendo r , s e $K(\gamma)$ dados em (3.4) e $K'(\beta) = \partial K(\beta)/\partial \beta$. Calculando a esperança de $j(\theta)$ obtemos a matriz de informação de Fisher $i(\theta)$. Os parâmetros α e β são ortogonais, pois o elemento $i_{\alpha\beta}(\theta)$ da matriz de informação é nulo.

Os estimadores de máxima verossimilhança são obtidos a partir da solução das equações $\partial \ell(\theta)/\partial \alpha = 0$ e $\partial \ell(\theta)/\partial \beta = 0$. O estimador de máxima verossimilhança para β é a raiz positiva da equação não-linear

$$\beta^2 - \beta[2r + K(\beta)] + r[s + K(\beta)] = 0. \quad (3.5)$$

Uma vez obtido o estimador de máxima verossimilhança $\hat{\beta}$ para β , podemos obter o estimador de máxima verossimilhança para α , dado por

$$\hat{\alpha} = \left[\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right]^{1/2}.$$

Os estimadores $\hat{\alpha}$ e $\hat{\beta}$ têm distribuição assintótica normal bivariada, como foi mostrado por Engelhardt et al. (1981):

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \overset{A}{\sim} N \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\beta^2}{n[0,25+\alpha^{-2}+I(\alpha)]} \end{pmatrix} \right], \quad (3.6)$$

em que $\overset{A}{\sim}$ denota assintoticamente distribuído,

$$I(\alpha) = 2 \int_0^\infty \{[1 + g(\alpha x)]^{-1} - 0,5\}^2 d\Phi(x) \quad \text{e} \quad g(y) = 1 + \frac{y^2}{2} + y \left(1 + \frac{y^2}{4} \right)^{1/2}.$$

Intervalos de confiança assintóticos para $\hat{\alpha}$ e $\hat{\beta}$ podem ser obtidos a partir de (3.6).

Em relação ao teste da razão de verossimilhanças para os parâmetros da distribuição, suponha que estamos interessados em testar $H_0 : \alpha = \alpha_0$ contra $H_1 : \alpha \neq \alpha_0$. Para este caso, temos que $LR(\alpha) = 2\{L(\hat{\alpha}, \hat{\beta}) - L(\alpha_0, \hat{\beta})\}$. Se o interesse residir em testar $H_0 : \beta = \beta_0$ contra $H_1 : \beta \neq \beta_0$, a estatística da razão de verossimilhanças será dada por $LR(\beta) = 2\{L(\hat{\alpha}, \hat{\beta}) - L(\hat{\alpha}, \beta_0)\}$. Nas duas situações

LR tem distribuição assintótica, sob a hipótese nula, qui-quadrado com um grau de liberdade.

3.4 Ajustes para a distribuição Birnbaum-Saunders

Nesta seção consideraremos o problema de realizar inferências sobre um dos parâmetros da distribuição Birnbaum-Saunders através da função de verossimilhança perfilada. Também obteremos as funções de verossimilhança perfilada modificadas de Cox e Reid (1987) e Barndorff-Nielsen (1983) e, adicionalmente, determinaremos estatísticas da razão de verossimilhanças baseadas nestas funções.

Podemos considerar duas situações, a saber: caso (i), α sendo o parâmetro de interesse e β sendo o parâmetro de perturbação; caso (ii), β sendo o parâmetro de interesse e α , o parâmetro de perturbação.

No primeiro caso, fixando α e resolvendo a equação $u_\beta = 0$, obtemos o estimador de máxima verossimilhança restrito de β ($\hat{\beta}_\alpha$). O estimador $\hat{\beta}_\alpha$ é o mesmo estimador de máxima verossimilhança original $\hat{\beta}$ que pode ser obtido como solução da equação não-linear dada em (3.5). Substituindo β por $\hat{\beta}$ na função de verossimilhança original obtemos a função de verossimilhança perfilada para α . O logaritmo da função de verossimilhança perfilada é dado por

$$\ell_p(\alpha) = -n \log \alpha - n \log \hat{\beta} + \sum_{i=1}^n \log \left[\left(\frac{\hat{\beta}}{t_i} \right)^{1/2} + \left(\frac{\hat{\beta}}{t_i} \right)^{3/2} \right] - \frac{n}{2\alpha^2} \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right).$$

Calculando $\partial \ell_p(\alpha) / \partial \alpha$, obtemos a função escore perfilada para α :

$$u_p(\alpha) = -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right).$$

Resolvendo a equação $u_p(\alpha) = 0$ obtemos

$$\hat{\alpha}_p = \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right)^{1/2},$$

que é a estimativa de máxima verossimilhança perfilada para α . Observe que $\hat{\alpha}_p = \hat{\alpha}$.

O logaritmo da função de verossimilhança perfilada modificada proposta por Cox e Reid (1987) obtida para o parâmetro α é

$$\ell_{CR}(\alpha) = \ell_p(\alpha) - \frac{1}{2} \log |j_{\beta\beta}(\alpha, \hat{\beta}_\alpha)|,$$

em que

$$j_{\beta\beta}(\alpha, \hat{\beta}_\alpha) = -\frac{n}{\hat{\beta}^2} + \frac{n}{2} \left(\frac{1}{\hat{\beta}^2} + \frac{2K'(\hat{\beta})}{K^2(\hat{\beta})} \right) + \frac{n}{\alpha^2} \frac{r}{\hat{\beta}^3},$$

$$K(\hat{\beta}) = \left[\frac{1}{n} \sum_{i=1}^n (\hat{\beta} + t_i)^{-1} \right]^{-1}$$

e

$$K'(\hat{\beta}) = \frac{n \sum (\hat{\beta} + t_i)^{-2}}{\left[\sum (\hat{\beta} + t_i)^{-1} \right]^2}.$$

A função escore modificada correspondente é dada por

$$u_{CR}(\alpha) = -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left[\left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right) + \frac{1}{2} \left(\frac{n}{\hat{\beta}^2} - \frac{n}{2} \left(\frac{1}{\hat{\beta}^2} + \frac{2K'(\hat{\beta})}{K^2(\hat{\beta})} \right) - \frac{n}{\alpha^2} \frac{r}{\hat{\beta}^3} \right) \frac{-1}{\hat{\beta}^3} \right].$$

O estimador de máxima verossimilhança perfilado modificado $\hat{\alpha}_{CR}$ para α , que é obtido como solução da equação $u_{CR}(\alpha) = 0$, não possui forma fechada. Assim, é necessária a utilização de um procedimento de otimização não-linear para encontrar o valor de α que maximiza a função $\ell_{CR}(\alpha)$, como, por exemplo, o algoritmo de Newton (Newton-Raphson, escore de Fisher) ou um algoritmo quasi-Newton (BFGS); ver Nocedal e Wright (1999) para detalhes sobre tais algoritmos.

O logaritmo da função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) para o parâmetro α é dado por

$$\ell_{BN}(\alpha) = \ell_p(\alpha) + \log \frac{|j_{\beta\beta}(\alpha, \hat{\beta}_\alpha)|^{1/2}}{|j_{\beta;\hat{\beta}}(\alpha, \hat{\beta}_\alpha)|}. \quad (3.7)$$

A quantidade $j_{\beta;\hat{\beta}}(\alpha, \hat{\beta}_\alpha)$ envolve derivadas sobre o espaço amostral de forma que sua obtenção é tipicamente complicada. Assim, é útil adotar uma aproximação para obtenção desta quantidade. Utilizamos a aproximação baseada nas covariâncias

empíricas proposta por Severini (1999) devido à dificuldade de encontrar as co-variâncias populacionais envolvidas. Tal aproximação é dada por

$$\begin{aligned} \check{I}(\alpha, \hat{\beta}_\alpha; \hat{\alpha}, \hat{\beta}) &= \sum_{j=1}^n \ell_\beta^{(j)}(\hat{\theta}_\alpha) \ell_\beta^{(j)}(\hat{\theta}) \\ &= \frac{n}{\hat{\beta}^2} - \frac{1}{\hat{\beta}} \sum_{j=1}^n A_j - \frac{1}{4} \left[\left(\frac{1}{\alpha^2} + \frac{1}{\hat{\alpha}^2} \right) \left(\sum_{j=1}^n A_j B_j - \frac{2}{\hat{\beta}} \sum_{j=1}^n B_j \right) + \sum_{j=1}^n A_j^2 + \frac{1}{\alpha^2 \hat{\alpha}^2} \sum_{j=1}^n B_j^2 \right], \end{aligned}$$

em que

$$A_j = \left(\frac{t_j^{-1/2} \hat{\beta}^{-1/2} + 3 \hat{\beta}^{1/2} t_j^{-3/2}}{t_j^{-1/2} \hat{\beta}^{1/2} + \hat{\beta}^{3/2} t_j^{-3/2}} \right) \quad \text{e} \quad B_j = \left(\frac{t_j}{\hat{\beta}^2} - \frac{1}{t_j} \right).$$

A função escore perfilada modificada é obtida a partir de $u_{BN}(\alpha) = \partial \ell_{BN}(\alpha) / \partial \alpha$. O estimador de máxima verossimilhança perfilado $\hat{\alpha}_{BN}$ é obtido como solução de $u_{BN}(\alpha) = 0$ e não apresenta forma fechada, devendo ser obtido numericamente.

A estatística da razão de verossimilhanças baseada na função de verossimilhança perfilada modificada de Cox e Reid, $\ell_{CR}(\alpha)$, para testar $H_0 : \alpha = \alpha_0$ contra $H_1 : \alpha \neq \alpha_0$ é dada por

$$LR_{CR}(\alpha) = 2 \{ \ell_{CR}(\hat{\alpha}) - \ell_{CR}(\alpha_0) \},$$

em que $\hat{\alpha}$ é o valor de α que maximiza $\ell_{CR}(\alpha)$. A estatística da razão de verossimilhanças baseada na verossimilhança perfilada modificada de Barndorff-Nielsen, $LR_{BN}(\alpha)$, é definida de maneira semelhante. Assim, temos

$$LR_{BN}(\alpha) = 2 \{ \ell_{BN}(\hat{\alpha}) - \ell_{BN}(\alpha_0) \},$$

em que $\hat{\alpha}$ é o valor de α que maximiza $\ell_{BN}(\alpha)$. Assintoticamente e sob a hipótese nula, estas estatísticas têm distribuição qui-quadrado com um grau de liberdade.

Agora consideraremos o caso (ii), ou seja, o parâmetro de interesse é o parâmetro de escala β e α é o parâmetro de perturbação. Fixando β e resolvendo a equação $u_\alpha =$

0 com relação a α obtemos a estimativa de máxima verossimilhança do parâmetro de perturbação α considerando β fixo:

$$\hat{\alpha}_\beta = \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{1/2}.$$

Substituindo $\hat{\alpha}_\beta$ no logaritmo da função verossimilhança original obtemos o logaritmo da função de verossimilhança perfilada. Para o caso em questão tal função é dada, ignorando uma constante, por

$$\ell_p(\beta) = \ell(\hat{\alpha}_\beta, \beta) = -\frac{n}{2} \log \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) - n \log \beta + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right],$$

em que r e s são dados em (3.4). Derivando essa função com relação a β obtemos a função escore perfilada:

$$u_p(\beta) = -\frac{n}{2} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} - \frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right).$$

A estimativa de máxima verossimilhança perfilada para β , denotada $\hat{\beta}_p$, pode ser obtida como solução da equação $u_p(\beta) = 0$. Neste caso, não é possível expressar o estimador $\hat{\beta}_p$ em forma fechada, então $\hat{\beta}_p$ deve ser obtido numericamente pela maximização do logaritmo da função de verossimilhança perfilada $\ell_p(\beta)$ usando um algoritmo de otimização não-linear.

O bloco $j_{\alpha\alpha}(\alpha, \beta)$ da matriz de informação observada avaliada em $(\hat{\alpha}_\beta, \beta)$ pode ser escrito como

$$j_{\alpha\alpha}(\hat{\alpha}_\beta, \beta) = -2n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1}.$$

Seguindo a definição da verossimilhança perfilada modificada de Cox e Reid (1987), obtemos o logaritmo da função de verossimilhança perfilada modificada correspondente para β :

$$\ell_{CR}(\beta) = \ell_p(\beta) + \frac{1}{2} \log \left| \frac{r}{\beta} + \frac{\beta}{s} - 2 \right|. \quad (3.8)$$

Podemos observar que não foi necessário encontrar uma reparametrização ortogonal entre os parâmetros de interesse e perturbação para a obtenção de (3.8), pois os parâmetros α e β que indexam a distribuição Birnbaum-Saunders são ortogonais.

A partir da primeira derivada de (3.8) com relação a β obtemos a função escore perfilada modificada:

$$u_{CR}(\beta) = -\frac{n+1}{2} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} - \frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right).$$

A estimativa de máxima verossimilhança perfilada modificada $\hat{\beta}_{CR}$ não pode ser obtida de forma explícita. Portanto, é necessário usar um procedimento numérico para encontrar o valor de β que maximiza a função $\ell_{CR}(\beta)$.

O logaritmo da função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) para o modelo em estudo é dado por

$$\ell_{BN}(\beta) = \ell_p(\beta) + \log \frac{|j_{\alpha\alpha}(\hat{\alpha}_\beta, \beta)|^{1/2}}{|\ell_{\alpha;\hat{\alpha}}(\hat{\alpha}_\beta, \beta)|}. \quad (3.9)$$

Novamente, devido às dificuldades encontradas na obtenção de $\ell_{\alpha;\hat{\alpha}}(\hat{\alpha}_\beta, \beta)$, por se tratar de uma derivada do espaço amostral, utilizamos uma aproximação. Foi possível obter a aproximação baseada em covariâncias populacionais proposta por Severini (1998). Assim, o logaritmo da função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen foi obtido substituindo o termo $\ell_{\alpha;\hat{\alpha}}(\hat{\alpha}_\beta, \beta)$ por $I(\hat{\alpha}_\beta, \beta; \hat{\alpha}, \hat{\beta})$ na expressão (3.9). A expressão obtida para esta aproximação é dada por

$$I(\hat{\alpha}_\beta, \beta; \hat{\alpha}, \hat{\beta}) = \frac{n\hat{\alpha}}{\hat{\alpha}_\beta^3} \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right).$$

A função escore perfilada modificada correspondente é

$$u_{BN}(\beta) = u_p(\beta) + \left(\frac{1}{s} - \frac{r}{\beta^2} \right) \left[\frac{1}{2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} + \frac{3}{2\hat{\alpha}_\beta^2} \right] - \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right)^{-1} \left(\frac{1}{\hat{\beta}} - \frac{\hat{\beta}}{\beta^2} \right).$$

Resolvendo a equação $u_{BN}(\beta) = 0$ encontramos o estimador de máxima verossimilhança perfilado modificado $\hat{\beta}_{BN}$.

A estatística da razão de verossimilhanças baseada na função de verossimilhança perfilada modificada de Cox e Reid para o teste de $H_0 : \beta = \beta_0$ versus $H_1 : \beta \neq \beta_0$ é dada por

$$LR_{CR}(\beta) = 2 \left\{ \ell_{CR}(\hat{\beta}) - \ell_{CR}(\beta_0) \right\},$$

em que $\hat{\beta}$ é o valor de β que maximiza a função $\ell_{CR}(\beta)$. Esta estatística possui distribuição assintótica, sob a hipótese nula, qui-quadrado com um grau de liberdade. De forma análoga, podemos definir a estatística da razão de verossimilhanças para o mesmo teste como $LR_{BN}(\beta) = 2 \left\{ \ell_{BN}(\hat{\beta}) - \ell_{BN}(\beta_0) \right\}$, que também possui distribuição assintótica qui-quadrado com um grau de liberdade, sob a hipótese nula.

Resultados numéricos

Neste capítulo avaliaremos numericamente a qualidade das inferências realizadas para o parâmetro de interesse da distribuição Birnbaum-Saunders baseadas na função de verossimilhança perfilada e em algumas de suas versões modificadas. Consideraremos as funções de verossimilhança perfiladas modificadas propostas por Barndorff-Nielsen (1983) e por Cox e Reid (1987). Estas funções de verossimilhança foram obtidas no capítulo anterior considerando os dois casos possíveis para o parâmetro de interesse da distribuição Birnbaum-Saunders, ou seja, quando os parâmetros de forma e escala são tidos como parâmetros de interesse e de perturbação, respectivamente, e vice-versa.

Formulamos um experimento de simulação de Monte Carlo baseado em 10.000 réplicas* considerando os seguintes tamanhos amostrais: $n = 10, 25$ e 50 . Os valores verdadeiros considerados para o parâmetro de forma foram $\alpha = 0,1, 0,5, 1$ e 2 . Para o parâmetro de escala atribuímos um único valor ($\beta = 1$) para todas as simulações, dado que este parâmetro funciona apenas como um multiplicador. O principal interesse do estudo é comparar os vieses dos estimadores de máxima verossimilhança perfilado e perfilados modificados, as distorções dos testes da razão de verossimilhanças baseados nas verossimilhanças perfilada e perfiladas modificadas consideradas e também verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição qui-quadrado apropriada. Também são apresentadas medidas amostrais de locação e de dispersão dos estimadores considerados, bem como a média e variância das estatísticas de teste. As simulações foram realizadas usando a linguagem de programação `0x` (ver Doornik, 2001). Para as maximizações, utilizamos o algoritmo quasi-Newton BFGS, através da função `MaxBFGS`, que está

* Os quadros 4.7 e 4.12 apresentam resultados de simulação baseados em 100.000 réplicas.

implementada na linguagem 0x.

4.1 Comportamento dos estimadores

Nesta seção estamos interessados em avaliar o comportamento dos estimadores obtidos a partir da maximização das funções de verossimilhança perfilada e suas versões modificadas propostas por Cox e Reid (1987) e por Barndorff-Nielsen (1983). Isto é, estamos interessados em avaliar, quando α é parâmetro de interesse, os desempenhos dos estimadores $\hat{\alpha}_p$, $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ e, quando β é que é de interesse, dos estimadores $\hat{\beta}_p$, $\hat{\beta}_{CR}$ e $\hat{\beta}_{BN}$. Consideramos as seguintes medidas de estimação pontual para os estimadores considerados: média, viés, variância, erro quadrático médio (EQM), viés relativo (VR), assimetria e curtose. O viés relativo é definido como $100 \times (\text{viés} / \text{valor verdadeiro do parâmetro})\%$.

Primeiramente apresentaremos os resultados de simulação correspondentes à situação em que α é o parâmetro de interesse. Consideramos $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2 . Através dos resultados apresentados nos Quadros 4.1 e 4.2 podemos analisar os desempenhos dos estimadores $\hat{\alpha}_p$, $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$. O Quadro 4.1 contém resultados de simulação para $\alpha = 0,1$ e $\alpha = 0,5$. Notamos que as estimativas dos vieses dos estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ são, em módulo, bem menores que os correspondentes valores para $\hat{\alpha}_p$. Entretanto, essas melhorias são acompanhadas de aumentos das estimativas das variâncias comparativamente àquela obtida para $\hat{\alpha}_p$. Observamos que, quando $\alpha = 0,1$, os estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ são bem mais precisos que $\hat{\alpha}_p$, principalmente para $n = 10$ e $n = 25$. Por exemplo, para $n = 10$, as estimativas das médias de $\hat{\alpha}_p$, $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ são 0,09274, 0,09777 e 0,09777, respectivamente, e os correspondentes valores para os vieses relativos são 7,25%, 2,22% e 2,22 %. Também observamos que todos os estimadores apresentaram valores negativos para o viés, ou seja, os estimadores subestimam o verdadeiro valor do parâmetro. Quando $\alpha = 0,5$, estes estimadores apresentaram comportamento bem semelhante em relação ao caso em que $\alpha = 0,1$. Temos, por exemplo, para $n = 10$, os valores 7,50%, 2,16% e 2,13% para os vieses relativos de $\hat{\alpha}_p$, $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$, respectivamente. É importante notar que

Quadro 4.1 Resultados de estimação pontual para α , modelo indexado pelos parâmetros $\alpha = 0,1, 0,5$ e $\beta = 1$.

$\alpha = 0,1$									
n	estimador	média	viés	variância	EQM	VR(%)	assimetria	curtose	
10	$\hat{\alpha}_p$	0,09274	-0,00726	0,00048	0,00054	7,25832	0,01269	3,13197	
	$\hat{\alpha}_{CR}$	0,09777	-0,00223	0,00054	0,00054	2,22656	0,01408	3,14664	
	$\hat{\alpha}_{BN}$	0,09777	-0,00223	0,00054	0,00054	2,22651	0,01408	3,14664	
25	$\hat{\alpha}_p$	0,09698	-0,00302	0,00020	0,00021	3,01667	0,01386	3,14429	
	$\hat{\alpha}_{CR}$	0,09899	-0,00101	0,00020	0,00021	1,01146	0,01443	3,15030	
	$\hat{\alpha}_{BN}$	0,09899	-0,00101	0,00020	0,00021	1,01146	0,01443	3,15030	
50	$\hat{\alpha}_p$	0,09841	-0,00159	0,00010	0,00010	1,58802	0,01426	3,14856	
	$\hat{\alpha}_{CR}$	0,09941	-0,00059	0,00010	0,00010	0,58632	0,01455	3,15159	
	$\hat{\alpha}_{BN}$	0,09941	-0,00059	0,00010	0,00010	0,58632	0,01455	3,15159	
$\alpha = 0,5$									
n	estimador	média	viés	variância	EQM	VR(%)	assimetria	curtose	
10	$\hat{\alpha}_p$	0,46246	-0,03754	0,01205	0,01346	7,50740	0,21940	6,04050	
	$\hat{\alpha}_{CR}$	0,48915	-0,01085	0,01363	0,01375	2,16952	0,23552	6,36293	
	$\hat{\alpha}_{BN}$	0,48931	-0,01069	0,01367	0,01378	2,13846	0,23561	6,36483	
25	$\hat{\alpha}_p$	0,48440	-0,01560	0,00490	0,00514	3,11941	0,23270	6,30486	
	$\hat{\alpha}_{CR}$	0,49499	-0,00501	0,00514	0,00516	1,00299	0,23895	6,43463	
	$\hat{\alpha}_{BN}$	0,49501	-0,00499	0,00514	0,00516	0,99868	0,23896	6,43489	
50	$\hat{\alpha}_p$	0,49179	-0,00821	0,00250	0,00257	1,64212	0,23707	6,39529	
	$\hat{\alpha}_{CR}$	0,49707	-0,00293	0,00256	0,00257	0,58631	0,24017	6,46033	
	$\hat{\alpha}_{BN}$	0,49707	-0,00293	0,00256	0,00257	0,58526	0,24017	6,46040	

Quadro 4.2 Resultados de estimação pontual para α , modelo indexado pelos parâmetros $\alpha = 1, 2$ e $\beta = 1$.

$\alpha = 1$								
n	estimador	média	viés	variância	EQM	VR(%)	assimetria	curtose
10	$\hat{\alpha}_p$	0,91598	-0,08402	0,04722	0,05428	8,40199	0,37146	11,92210
	$\hat{\alpha}_{CR}$	0,97516	-0,02484	0,05481	0,05542	2,48406	0,37324	12,63564
	$\hat{\alpha}_{BN}$	0,97917	-0,02083	0,05672	0,05715	2,08335	0,37325	12,68285
25	$\hat{\alpha}_p$	0,96729	-0,03271	0,01985	0,02092	3,27077	0,37318	12,54254
	$\hat{\alpha}_{CR}$	0,99072	-0,00928	0,02101	0,02109	0,92845	0,37322	12,81809
	$\hat{\alpha}_{BN}$	0,99130	-0,00870	0,02111	0,02119	0,87012	0,37321	12,82489
50	$\hat{\alpha}_p$	0,98233	-0,01767	0,00992	0,01023	1,76670	0,37325	12,72005
	$\hat{\alpha}_{CR}$	0,99398	-0,00602	0,01020	0,01023	0,60160	0,37319	12,85614
	$\hat{\alpha}_{BN}$	0,99413	-0,00587	0,01021	0,01025	0,58679	0,37319	12,85786
$\alpha = 2$								
n	estimador	média	viés	variância	EQM	VR(%)	assimetria	curtose
10	$\hat{\alpha}_p$	1,75337	-0,24663	0,17726	0,23808	12,33143	0,26943	19,06939
	$\hat{\alpha}_{CR}$	1,88157	-0,11843	0,20853	0,22255	5,92134	0,24909	19,69956
	$\hat{\alpha}_{BN}$	1,92256	-0,07744	0,23049	0,23649	3,87196	0,24286	19,88307
25	$\hat{\alpha}_p$	1,93035	-0,06965	0,08064	0,08549	3,48258	0,24169	19,91702
	$\hat{\alpha}_{CR}$	1,98211	-0,01789	0,08532	0,08564	0,89447	0,23406	20,13558
	$\hat{\alpha}_{BN}$	1,98901	-0,01099	0,08691	0,08703	0,54947	0,23306	20,16381
50	$\hat{\alpha}_p$	1,96225	-0,03775	0,04013	0,04156	1,88742	0,23696	20,05317
	$\hat{\alpha}_{CR}$	1,98792	-0,01208	0,04126	0,04140	0,60380	0,23322	20,15938
	$\hat{\alpha}_{BN}$	1,98964	-0,01036	0,04144	0,04155	0,51817	0,23297	20,16636

os desempenhos dos estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ são bem parecidos. A semelhança entre estes estimadores se torna mais evidente em amostras maiores. Também observamos que o estimador $\hat{\alpha}_p$ apresentou os menores coeficientes de assimetria e curtose quando comparado com os estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$. Um outro fato interessante é que para $\alpha = 0,1$ a curtose dos estimadores considerados estão próximas de 3, ou seja, a curtose destes estimadores estão próximas da curtose da distribuição normal padrão. À medida que α cresce observamos que os estimadores passam a apresentar uma curtose bem mais elevada, i.e., a distribuição destes estimadores passa a ter uma curva leptocúrtica. Em relação a assimetria, observamos que esta medida é sempre positiva, ou seja, a distribuição dos estimadores considerados apresentam valores concentrados à esquerda. No Quadro 4.2, apresentamos os resultados obtidos quando $\alpha = 1$ e $\alpha = 2$. Analisando este quadro notamos que, quando $\alpha = 1$, no que se refere à precisão, o estimador $\hat{\alpha}_p$ apresenta desempenho bem inferior em relação aos demais, principalmente em amostras menores. Também é interessante notar que o desempenho do estimador $\hat{\alpha}_{BN}$ é levemente superior em relação ao estimador $\hat{\alpha}_{CR}$. Esta diferença de desempenho é mais evidente quando $\alpha = 2$. Por exemplo, quando $n = 10$, os vieses relativos do estimador $\hat{\alpha}_{CR}$ são aproximadamente 2,48% ($\alpha = 1$) e 5,92% ($\alpha = 2$) e para o estimador $\hat{\alpha}_{BN}$ são 2,08% ($\alpha = 1$) e 3,87% ($\alpha = 2$). Contudo, a escolha de um estimador adequado para o parâmetro de forma deve levar em consideração outras medidas como o EQM, a assimetria e a curtose das estimativas.

Os estimadores $\hat{\beta}_p$, $\hat{\beta}_{CR}$ e $\hat{\beta}_{BN}$ são equivalentes, pois maximizar a função $\ell_p(\beta)$ equivale a maximizar $\ell_{CR}(\beta)$ ou $\ell_{BN}(\beta)$. Adicionalmente, estes estimadores são equivalentes ao estimador de máxima verossimilhança usual $\hat{\beta}$, que pode ser obtido como solução de uma equação não-linear. Portanto, para o parâmetro β , apresentaremos apenas resultados de simulação referentes ao estimador $\hat{\beta}$.

O Quadro 4.3 contém resultados de estimação pontual para o parâmetro β considerando $\alpha = 0,1, 0,5, 1$ e 2 . Através deste quadro observamos que o viés do estimador $\hat{\beta}$ tende a crescer com o parâmetro α . Por exemplo, quando $n = 10$, temos para $\alpha = 0,1, 0,5, 1$ e 2 os correspondentes vieses 0,00057, 0,01194, 0,04004

Quadro 4.3 Resultados de estimação pontual para β , modelo indexado pelos parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2 .

n	α	média	viés	variância	EQM	VR(%)	assimetria	curtose
10	0,1	1,00057	0,00057	0,00099	0,00099	0,05706	0,01472	3,15338
	0,5	1,01194	0,01194	0,02411	0,02425	1,19371	0,24187	6,49660
	1,0	1,04004	0,04004	0,08859	0,09020	4,00369	0,37311	12,92593
	2,0	1,10201	0,10201	0,26966	0,28007	10,20085	0,23148	20,20833
25	0,1	1,00029	0,00029	0,00040	0,00040	0,02942	0,01472	3,15338
	0,5	1,00528	0,00528	0,00966	0,00968	0,52824	0,24187	6,49660
	1,0	1,01725	0,01725	0,03344	0,03373	1,72463	0,37311	12,92593
	2,0	1,04010	0,04010	0,08310	0,08471	4,01038	0,23148	20,20833
50	0,1	1,00016	0,00016	0,00020	0,00020	0,01623	0,01472	3,15338
	0,5	1,00271	0,00271	0,00476	0,00477	0,27073	0,24187	6,49660
	1,0	1,00871	0,00871	0,01619	0,01627	0,87096	0,37311	12,92593
	2,0	1,01967	0,01967	0,03818	0,03856	1,96730	0,23148	20,20833

e 0,10201. Podemos ainda comparar os respectivos vieses relativos, que são 0,057%, 1,19%, 4,00% e 10,20%. Este comportamento também se verifica para tamanhos de amostra maiores, porém, nestes casos, o aumento do viés é menor para valores maiores de α . Por exemplo, quando $n = 50$, os vieses do estimador $\hat{\beta}_p$ para $\alpha = 0,1, 0,5, 1$ e 2 são 0,00016, 0,00271, 0,00871 e 0,01967. Ou seja, os vieses relativos são menores que 2% para os quatro valores de α considerados. Comportamento semelhante ocorre para as medidas de variância, erro quadrático médio e curtose.

4.2 Desempenhos dos testes

Nesta seção, temos como objetivo comparar os desempenhos dos testes baseados na estatística da razão de verossimilhanças original e em suas versões modificadas. Consideramos a estatística da razão de verossimilhanças usual (LR), a estatística da razão de verossimilhanças perfiladas modificadas proposta por Cox e Reid (1987) (LR_{CR}) e a estatística da razão de verossimilhanças perfiladas modificadas proposta por Barndorff-Nielsen (1983) (LR_{BN}). Estas estatísticas foram obtidas a partir das respectivas funções de verossimilhança perfilada e perfiladas modificadas desenvolvidas na Seção 3.4. Os desempenhos são avaliados em função da proximidade da probabilidade de rejeição da hipótese nula, sendo esta verdadeira, aos respectivos níveis nominais. Avaliamos também os poderes dos testes em estudo sob algumas situações. Adicionalmente, comparamos a média, a variância e os quantis amostrais das estatísticas de teste relativamente às correspondentes quantidades da distribuição assintótica de referência.

Para cada tamanho de amostra e cada nível nominal considerado, calculamos as taxas de rejeição (expressas em porcentagem) de cada teste, isto é, estimamos, via simulação, $P(LR \geq x)$, $P(LR_{CR} \geq x)$ e $P(LR_{BN} \geq x)$ em que x é o quantil $(1 - \alpha)$ da distribuição χ_1^2 .

Primeiramente nosso interesse reside em testar a hipótese nula $H_0 : \alpha = \alpha_0$ contra a hipótese alternativa bilateral $H_1 : \alpha \neq \alpha_0$, em que α_0 é uma constante especificada, α sendo o parâmetro de interesse e β sendo o parâmetro de perturbação.

Quadro 4.4 Taxas de rejeição (%) calculadas a partir de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2 .

n	nível nominal	$\alpha = 0,1$				$\alpha = 0,5$				$\alpha = 1$				$\alpha = 2$			
		LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR_{BN}
10	10%	13,44	11,00	11,00	13,57	10,85	10,86	13,97	10,65	10,78	16,10	9,89	10,78	16,10	9,89	10,78	8,99
	5%	7,47	5,14	5,14	7,60	5,18	5,20	7,83	5,24	5,29	9,44	5,20	5,29	9,44	5,20	5,29	4,58
	1%	1,70	1,07	1,07	1,72	1,11	1,11	1,83	1,06	1,11	2,53	1,04	1,11	2,53	1,04	1,11	1,96
	0,5%	0,89	0,53	0,53	0,91	0,52	0,52	0,97	0,52	0,55	1,40	0,48	0,55	1,40	0,48	0,55	0,40
25	10%	10,84	10,15	10,15	10,88	10,13	10,14	11,56	10,41	10,50	11,98	10,12	10,50	11,98	10,12	10,50	10,10
	5%	5,94	5,33	5,33	5,93	5,34	5,35	6,18	5,31	5,33	6,24	5,36	5,33	6,24	5,36	5,33	5,30
	1%	1,43	1,18	1,18	1,46	1,15	1,15	1,46	1,04	1,04	1,54	1,07	1,04	1,54	1,07	1,04	1,06
	0,5%	0,82	0,57	0,57	0,82	0,59	0,59	0,61	0,50	0,52	0,82	0,59	0,52	0,82	0,59	0,52	0,54
50	10%	10,89	10,35	10,35	10,86	10,35	10,34	10,93	10,31	10,32	11,24	10,48	10,32	11,24	10,48	10,32	10,47
	5%	5,54	5,21	5,21	5,53	5,20	5,19	5,40	5,03	5,05	5,65	4,95	5,05	5,65	4,95	5,05	5,04
	1%	1,19	0,99	0,99	1,21	1,02	1,02	1,14	1,02	1,02	1,18	1,03	1,02	1,18	1,03	1,02	1,03
	0,5%	0,70	0,56	0,56	0,71	0,55	0,55	0,68	0,61	0,61	0,60	0,49	0,61	0,60	0,49	0,61	0,51

O Quadro 4.4 apresenta as taxas de rejeição dos testes da razão verossimilhanças baseados nas funções de verossimilhança perfilada e suas versões modificadas, isto é, dos testes baseados em LR , LR_{CR} e LR_{BN} . Neste estudo, fixamos $\beta = 1$ e variamos o tamanho da amostra tomando $n = 10, 25$ e 50 para $\alpha = 0,1, 0,5, 1$ e 2 . Através do Quadro 4.4, podemos observar que os testes baseados nas estatísticas da razão de verossimilhanças perfiladas modificadas LR_{CR} e LR_{BN} têm desempenhos superiores ao teste da razão de verossimilhanças original LR . Por exemplo, para um teste com nível nominal de 10%, tamanho amostral $n = 10$ e para $\alpha = 0,5$, o teste LR apresentou taxa de rejeição de 13,57%, enquanto que os testes LR_{CR} e LR_{BN} apresentaram taxas de 10,85% e 10,86%, respectivamente. Neste caso, podemos dizer que o teste LR é bastante anti-conservativo, apresentando taxa de rejeição bem acima do nível nominal. Já os testes LR_{CR} e LR_{BN} ainda são anti-conservativos, porém apresentando taxas de rejeição não muito acima do nível nominal. Quando o tamanho amostral aumenta, os testes passam a apresentar desempenhos semelhantes. Por exemplo, quando $\alpha = 0,5$, para $n = 50$ e nível nominal de 10%, os testes LR , LR_{CR} e LR_{BN} apresentam as seguintes taxas de rejeição: 10,86%, 10,35% e 10,34%, respectivamente. Um outro fato interessante é que, para valores elevados do parâmetro de forma, o teste LR passa a apresentar taxas de rejeição ainda mais elevadas, principalmente em pequenas amostras. Enquanto que os testes LR_{CR} e LR_{BN} ainda apresentam taxas de rejeição próximas dos níveis nominais correspondentes. Por exemplo, quando o parâmetro de forma é igual a $\alpha = 2$, $n = 10$ e nível nominal de 10%, os testes apresentam as seguintes taxas de rejeição: $LR = 16,10\%$, $LR_{CR} = 9,89\%$ e $LR_{BN} = 8,99\%$. Porém, ainda para $\alpha = 2$, mas agora com $n = 25$, estes testes passam a apresentar taxas de rejeição bem mais próximas dos respectivos níveis nominais. Neste caso, as taxas de rejeição são $LR = 11,98\%$, $LR_{CR} = 10,12\%$ e $LR_{BN} = 10,10\%$ para o nível nominal de 10%.

De acordo com os resultados apresentados no Quadro 4.4, os testes LR_{CR} e LR_{BN} apresentaram comportamentos bastante semelhantes, principalmente para valores menores dos parâmetros de forma. Por exemplo, para $n = 10$ e ao nível

nominal de 1%, os dois testes apresentaram taxas de rejeição de 1,07%. Vale ressaltar os que valores para o parâmetro de forma entre 0,1 e 0,5 são típicos em problemas de modelagem de desgaste estocástico em geral. Portanto, tanto o teste LR_{CR} quanto o teste LR_{BN} são recomendados em problemas gerais.

No Quadro 4.5 comparamos as médias e as variâncias das estatísticas LR , LR_{CR} e LR_{BN} e as respectivas medidas da distribuição χ_1^2 , para um modelo indexado por $\alpha = 0,5$ e $\beta = 1$ e tamanho amostral $n = 10$. Os resultados deste quadro mostram que as estatísticas LR_{CR} e LR_{BN} apresentam médias e variâncias bem mais próximas da média e variância de uma variável χ_1^2 do que as da estatística LR . Um comportamento semelhante é observado no Quadro 4.6, onde comparamos os quantis da distribuição χ_1^2 com os quantis amostrais das estatísticas LR , LR_{CR} e LR_{BN} . Através deste quadro, também notamos que para quantis mais elevados os quantis amostrais das estatísticas LR , LR_{CR} e LR_{BN} são mais pobremente aproximados pelos quantis da distribuição χ_1^2 do que se verifica para quantis mais baixos.

Quadro 4.5 Média e variância de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

momentos	χ_1^2	LR	LR_{CR}	LR_{BN}
média	1,000	1,188	1,024	1,025
variância	2,000	2,798	2,091	2,094

No Quadro 4.7 avaliamos os tamanhos empíricos dos testes para uma gama de valores para o nível nominal (de 0,01% a 50%). Para esta situação, as simulações foram baseadas em 100.000 amostras de Monte Carlo, ou seja, utilizamos 10 vezes mais réplicas do que nas simulações anteriores. Estes resultados consideram o caso em que $n = 10$, $\alpha = 0,5$ e $\beta = 1$. Do Quadro 4.7, notamos que no extremo da cauda da distribuição χ_1^2 as diferenças relativas entre as taxas de rejeição empíricas dos testes da razão de verossimilhanças e os níveis nominais são extremamente elevadas, ao contrário do que ocorre na região mediana da distribuição. Por exemplo, quando $\alpha = 0,01\%$, para a estatística LR essa diferença relativa é de 290% $((39 - 10)/10 \times 100\%)$, enquanto que para o nível nominal de 50% esta diferença é de

apenas 7,83% $((53916 - 50000)/50000 \times 100\%)$. Esta discrepância relativa é bem menor para as estatísticas LR_{CR} e LR_{BN} . As estatísticas LR_{CR} e LR_{BN} , por exemplo, apresentaram diferenças relativas para o nível nominal $\alpha = 0,05\%$ de 30% $((65 - 50)/50 \times 100\%)$ e 34% $((67 - 50)/50 \times 100\%)$, respectivamente, enquanto que para a estatística LR tal diferença é de 192% $((146 - 50)/50 \times 100\%)$.

Quadro 4.6 Quantis de χ^2_1 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

quantis(%)	χ^2	LR	LR_{CR}	LR_{BN}
85,0	2,072	2,497	2,190	2,194
90,0	2,706	3,332	2,823	2,826
95,0	3,841	4,563	3,892	3,895
97,0	4,709	5,600	4,869	4,867
99,0	6,635	7,645	6,751	6,759
99,5	7,879	9,317	7,969	7,967

Quadro 4.7 Número de rejeições em 100.000 amostras de Monte Carlo para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

$\alpha(\%)$	exato	LR	LR_{CR}	LR_{BN}
50	50000	53916	50499	50514
15	15000	19053	15763	15792
10	10000	13546	10649	10659
5	5000	7572	5452	5459
3	3000	4909	3289	3298
1	1000	1874	1137	1140
0,5	500	1022	606	608
0,1	100	279	121	121
0,05	50	150	68	69
0,01	10	39	9	9

Analizamos também os poderes dos testes LR , LR_{CR} e LR_{BN} para testar a hipótese $H_0 : \alpha = \alpha_0$ contra $H_1 : \alpha \neq \alpha_0$, para valores de α_0 variando de 0,14 a 0,28 e níveis nominais 5% e 1%. O valor verdadeiro para o parâmetro de interesse foi fixado em $\alpha = 0,1$. As simulações do poder foram feitas usando valores críticos estimados e não valores tabulados. Adotamos este procedimento porque em deter-

minadas situações alguns testes são anti-conservativos. Dessa forma, as simulações ficam ajustadas para que todos os testes tenham o mesmo tamanho. Os resultados encontram-se apresentados no Quadro 4.8 para o caso em que $n = 10$ e $\beta = 1$. O poder do teste LR é superior ao poder do teste LR_{BN} e este, por sua vez, é ligeiramente superior ao do teste LR_{CR} . Por exemplo, para $\alpha_0 = 0,26$ e nível nominal de 5%, temos os valores 98,96%, 98,04% e 98,13% para os poderes dos testes LR , LR_{CR} e LR_{BN} , respectivamente. Comportamento semelhante ocorre para o nível nominal de 1%, porém a superioridade do poder do teste LR em relação aos demais é mais acentuada. Por exemplo, ainda para $\alpha_0 = 0,26$, temos os valores 88,87%, 83,65% e 84,15% para os poderes dos testes LR , LR_{CR} e LR_{BN} , respectivamente, ao nível nominal de 1%.

Quadro 4.8 Poderes dos testes calculados de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1$.

α_0	5%			1%		
	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}
0,14	28,68	23,89	24,00	10,18	7,89	7,94
0,16	49,53	42,92	43,08	20,61	16,07	16,18
0,18	68,48	61,87	62,14	35,03	28,57	28,89
0,20	83,98	78,97	79,25	53,01	45,49	45,86
0,22	92,33	88,86	89,10	66,95	59,26	59,55
0,24	96,80	95,19	95,35	80,04	73,44	73,94
0,26	98,96	98,04	98,13	88,87	83,65	84,15
0,28	99,66	99,44	99,49	94,91	91,65	92,00

A fim de avaliar a aproximação assintótica dos testes LR , LR_{CR} e LR_{BN} apresentamos o gráfico das discrepâncias relativas entre os quantis amostrais e os quantis assintóticos das estatísticas de testes contra os quantis assintóticos. Este gráfico é apresentado na Figura 4.1. A discrepância relativa é definida como

$$\frac{ST(1 - \alpha) - \chi_1^2(1 - \alpha)}{\chi_1^2(1 - \alpha)},$$

em que ST denota a estatística de teste (LR , LR_{CR} ou LR_{BN} , conforme o caso), $ST(1 - \alpha)$ denota o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística de teste e o correspondente quantil da distribuição χ_1^2 sendo

Figura 4.1 *Gráfico das discrepâncias relativas de quantis em relação a α ($n = 10$, $\alpha = 0,5$ e $\beta = 1$).*

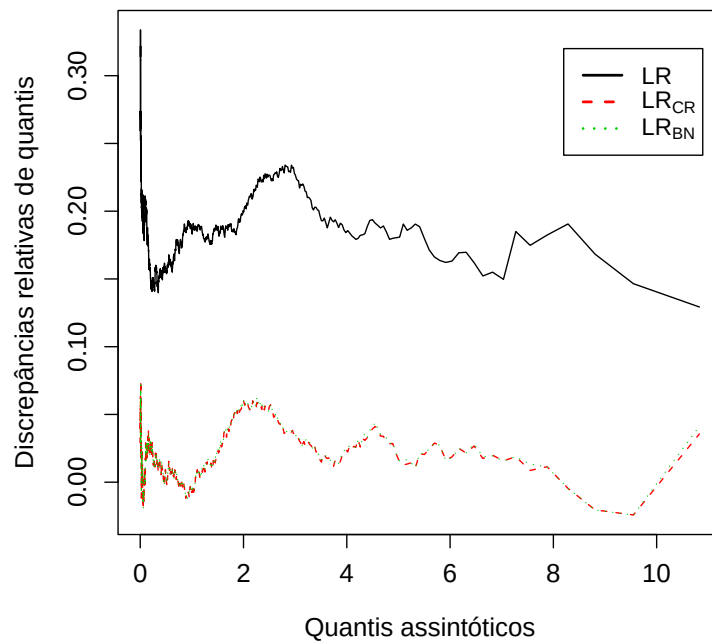
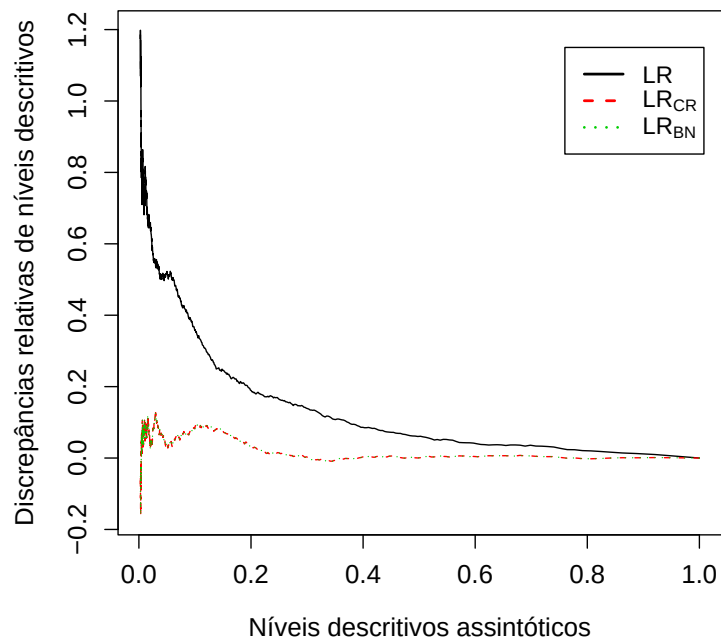


Figura 4.2 *Gráfico das discrepâncias relativas de níveis descritivos em relação a α ($n = 10$, $\alpha = 0,5$ e $\beta = 1$).*



denotado por $\chi_1^2(1-\alpha)$. A idéia, então, é que quanto mais próxima da ordenada zero a curva estiver, melhor será a aproximação assintótica utilizada no teste. Observamos, na Figura 4.1, que para todos os testes as curvas de discrepâncias de quantis oscilam bastante ao longo do eixo horizontal. Note que esta oscilação ocorre num patamar mais elevado para o teste LR , aproximadamente em torno de 18%, o que confirma que este teste é anti-conservativo. As curvas de discrepância de quantis para as estatísticas LR_{CR} e LR_{BN} encontram-se bem próximas e oscilam aproximadamente em torno de 3%. A curva de discrepância de quantis para LR encontra-se inteiramente acima de 12%, enquanto que para os testes LR_{CR} e LR_{BN} estas curvas encontram-se abaixo de 9%. Ou seja, as discrepâncias de quantis para as estatísticas LR_{CR} e LR_{BN} são menores que as da estatística LR ao longo do eixo horizontal.

Na Figura 4.2 apresentamos o gráfico das discrepâncias relativas entre os níveis descritivos (p -valores) dos testes e os níveis descritivos assintóticos contra os níveis descritivos assintóticos. Para um valor x observado da estatística ST esta discrepância é definida como

$$\frac{P(ST > x) - P(\chi_1^2 > x)}{P(\chi_1^2 > x)}.$$

Analisando a Figura 4.2 notamos que quando se caminha em direção ao extremo da cauda da distribuição χ^2 de referência, as discrepâncias relativas de níveis descritivos tendem a crescer rapidamente para o teste da razão de verossimilhanças. Estas discrepâncias estão bastante próximas para os testes LR_{CR} e LR_{BN} ao longo dos quantis da distribuição de referência e no extremo da cauda são bem menos acentuadas em relação ao teste LR .

A seguir avaliaremos os resultados de simulação dos testes da razão de verossimilhanças baseados nas funções de verossimilhança perfilada e perfiladas modificadas de Cox e Reid e Barndorff-Nielsen desenvolvidas para o parâmetro de interesse β .

O Quadro 4.9 apresenta as taxas de rejeição (expressas em porcentagem) dos testes LR , LR_{CR} e LR_{BN} para o caso em desejamos testar a hipótese $H_0 : \beta = \beta_0$. O parâmetro de escala β funciona apenas como um multiplicador, portanto,

Quadro 4.9 Taxas de rejeição (%) calculadas a partir de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1, 0,5, 1$ e 2 .

n	nível nominal	$\alpha = 0,1$				$\alpha = 0,5$				$\alpha = 1$				$\alpha = 2$			
		LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}	LR_{BN}
10	10%	12,31	10,30	8,52	12,33	10,25	8,26	12,37	10,18	8,01	13,12	10,66	7,94				
	5%	6,61	5,35	3,95	6,62	5,30	3,85	6,75	5,18	3,65	6,95	5,27	3,65				
	1%	1,49	1,08	0,70	1,56	1,05	0,69	1,55	1,09	0,67	1,73	1,12	0,54				
	0,5%	0,91	0,59	0,35	0,89	0,60	0,33	0,87	0,59	0,28	0,97	0,53	0,25				
25	10%	11,10	10,38	9,67	11,19	10,33	9,60	11,15	10,37	9,52	11,27	10,52	9,53				
	5%	5,83	5,30	4,86	5,91	5,38	4,76	5,87	5,33	4,75	6,14	5,49	4,77				
	1%	1,25	1,08	0,93	1,30	1,07	0,99	1,40	1,15	0,97	1,19	1,09	0,96				
	0,5%	0,62	0,52	0,44	0,64	0,52	0,45	0,76	0,61	0,46	0,80	0,64	0,43				
50	10%	10,71	10,43	9,96	10,64	10,35	9,92	10,72	10,24	9,74	10,87	10,38	9,87				
	5%	5,32	5,05	4,85	5,36	5,17	4,88	5,54	5,17	4,90	5,39	5,07	4,79				
	1%	1,19	1,08	1,00	1,20	1,10	0,99	1,18	1,09	0,98	1,12	1,06	0,97				
	0,5%	0,56	0,51	0,48	0,58	0,53	0,46	0,61	0,55	0,50	0,65	0,59	0,52				

as simulações foram realizadas fixando $\beta_0 = 1.0$. Foram considerados diferentes valores de α (0,1, 0,5, 1 e 2) e variamos o tamanho da amostra em $n = 10, 25$ e 50. Analisando o Quadro 4.9 notamos que os testes LR e LR_{CR} são, em geral, anti-conservativos, enquanto que o teste LR_{BN} apresentou taxas de rejeição inferiores aos níveis descritivos sendo, assim, um teste conservativo. Por exemplo, para $n = 10$ e $\alpha = 0,5$ ao nível nominal de 1%, os testes LR , LR_{CR} e LR_{BN} apresentaram taxas de rejeição de 1,56%, 1,05% e 0,69%, respectivamente. O teste LR_{CR} apresentou melhor desempenho, principalmente em amostras pequenas; por exemplo, ao nível nominal 10%, quando $\alpha = 1$ e $n = 10$, o teste LR_{CR} apresentou taxa de rejeição de 10,18%, enquanto que os testes LR e LR_{BN} apresentaram 12,37% e 8,01%, respectivamente. Como esperado, em amostras maiores, os testes apresentaram taxas mais próximas dos correspondentes níveis nominais. Por exemplo, quando $\alpha = 0,5$, $n = 50$ e ao nível nominal de 10%, foram obtidas taxas de rejeição de 10,64%, 10,35% e 9,92% para os testes LR , LR_{CR} e LR_{BN} , respectivamente.

Através do Quadro 4.10 podemos comparar a média e a variância das estatísticas LR , LR_{CR} e LR_{BN} e as mesmas medidas da distribuição χ_1^2 para o caso em que $n = 10$, $\beta = 1$ e $\alpha = 0,5$. Observamos que, para a estatística LR_{CR} , estas medidas estão mais próximas das medidas correspondentes da distribuição χ^2 de referência. Também observamos que para a estatística da razão de verossimilhanças usual estas medidas são bem mais elevadas quando comparadas com a média e com a variância de uma variável aleatória com distribuição χ_1^2 . Para a estatística LR_{BN} , estas medidas encontram-se abaixo das medidas correspondentes da distribuição χ_1^2 . Este mesmo comportamento foi observado no Quadro 4.11 ao compararmos os quantis amostrais das estatísticas LR , LR_{CR} e LR_{BN} com os da distribuição χ_1^2 . Os quantis amostrais da estatística LR excedem os da estatística LR_{CR} e esta, por sua vez, apresentou quantis amostrais maiores que os quantis da distribuição χ_1^2 . Observamos ainda que os quantis da estatística LR_{BN} não excederam os da distribuição χ_1^2 .

No Quadro 4.12 apresentamos os resultados de simulação de 100.000 amostras de Monte Carlo com o intuito de avaliar os tamanhos empíricos dos testes para valores

Quadro 4.10 Média e variância de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

momentos	χ_1^2	LR	LR_{CR}	LR_{BN}
média	1,000	1,151	1,030	0,910
variância	2,000	2,590	2,078	1,622

Quadro 4.11 Quantis de χ_1^2 , LR_p , LR_{CR} e LR_{BN} para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

quantis(%)	χ^2	LR	LR_{CR}	LR_{BN}
85,0	2,072	2,369	2,124	1,874
90,0	2,706	3,072	2,750	2,425
95,0	3,841	4,408	3,946	3,483
97,0	4,709	5,330	4,779	4,216
99,0	6,635	7,564	6,786	6,008
99,5	7,879	9,221	8,271	7,287

Quadro 4.12 Números de rejeições em 100.000 amostras de Monte Carlo para um modelo indexado pelos parâmetros $\beta = 1$, $\alpha = 0,5$ e $n = 10$.

$\alpha(\%)$	exato	LR	LR_{CR}	LR_{BN}
50	50000	53287	51035	48336
15	15000	18367	15967	13406
10	10000	12853	10822	8742
5	5000	7012	5619	4216
3	3000	4566	3450	2475
1	1000	1792	1197	729
0,5	500	943	605	318
0,1	100	209	118	58
0,05	50	118	61	29
0,01	10	31	12	5

pequenos do nível nominal. Este quadro mostra o número de rejeições para os testes LR , LR_{CR} e LR_{BN} para o caso em que $n = 10$, $\alpha = 0,1$ e $\beta = 1$. Do Quadro 4.12, notamos que as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais dos testes LR e LR_{CR} são positivas, enquanto que para o teste LR_{BN} esta diferença é negativa. Também podemos observar que estas diferenças são mais elevadas no extremo da cauda da distribuição χ_1^2 . Por exemplo, para o teste LR , ao

nível nominal 50% esta diferença é de 6,57% $((53287 - 50000)/50000 \times 100)$ enquanto que ao nível nominal 0,05% ela é de 136% $((118 - 50)/50 \times 100)$. Para o teste LR_{CR} esta diferença é menos acentuada; por exemplo, para as duas situações citadas acima as diferenças relativas são de 2,07% $((51035 - 50000)/50000 \times 100)$ e 22% $((61 - 50)/50 \times 100)$, respectivamente. Para o teste LR_{BN} , ainda considerando as duas situações anteriores, as diferenças relativas são -3,38% $((48336 - 50000)/50000 \times 100)$ e -42% $((29 - 50)/50 \times 100)$, respectivamente.

Apresentamos, no Quadro 4.13, um estudo de poder dos testes aqui considerados para testar a hipótese $H_0 : \beta = \beta_0$ contra $H_0 : \beta \neq \beta_0$, em que β_0 assume valores variando de 1,2 a 4. Utilizamos valores críticos estimados ao invés dos valores tabulados. Consideramos o modelo gerado com parâmetros $\alpha = 1$ e $\beta = 1$ para o tamanho amostral $n = 10$. Observamos que o teste LR é mais poderoso que o teste LR_{CR} e este, por sua vez, é mais poderoso que o teste LR_{BN} . Por exemplo, para $\beta_0 = 2,8$ e nível nominal de 5%, os testes LR , LR_{CR} e LR_{BN} apresentaram os seguintes poderes: 90,77, 90,37 e 89,95, respectivamente. Para valores maiores de β_0 as diferenças entre os poderes destes testes passam a ser menos acentuadas.

Quadro 4.13 Poderes dos testes calculados de 10.000 amostras de Monte Carlo geradas a partir de uma distribuição Birnbaum-Saunders com parâmetros $\beta = 1$ e $\alpha = 0,1$.

β_0	5%			1%		
	LR	LR_{CR}	LR_{BN}	LR	LR_{CR}	LR_{BN}
1,2	8,88	8,83	8,79	1,930	1,920	1,880
1,6	30,38	30,07	29,69	9,420	9,360	9,170
2,0	59,16	58,53	58,17	25,370	24,820	24,370
2,4	79,69	79,20	78,68	44,050	43,150	41,910
2,8	90,77	90,37	89,95	61,500	60,270	58,640
3,2	96,79	96,63	96,25	75,970	74,900	73,140
3,6	98,86	98,74	98,57	85,650	84,510	82,880
4,0	99,65	99,60	99,55	92,550	91,610	90,330

A Figura 4.3 mostra o gráfico das discrepâncias relativas entre os quantis amostrais e os quantis assintóticos das estatísticas de teste contra os correspondentes quantis assintóticos. Estas discrepâncias oscilam em patamares diferentes para as três estatísticas consideradas. A curva de discrepâncias relativas de quantis para

Figura 4.3 *Gráfico das discrepâncias relativas de quantis em relação a α ($n = 10$, $\alpha = 0,5$ e $\beta = 1$).*

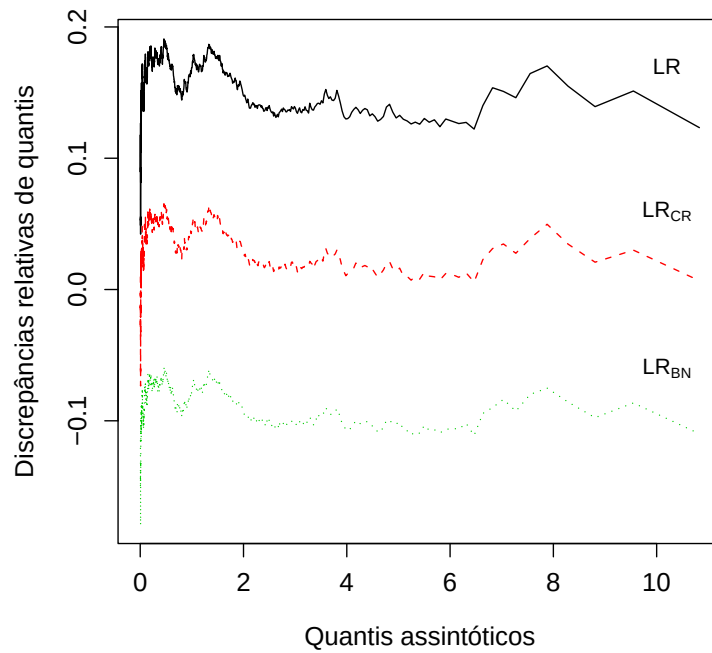
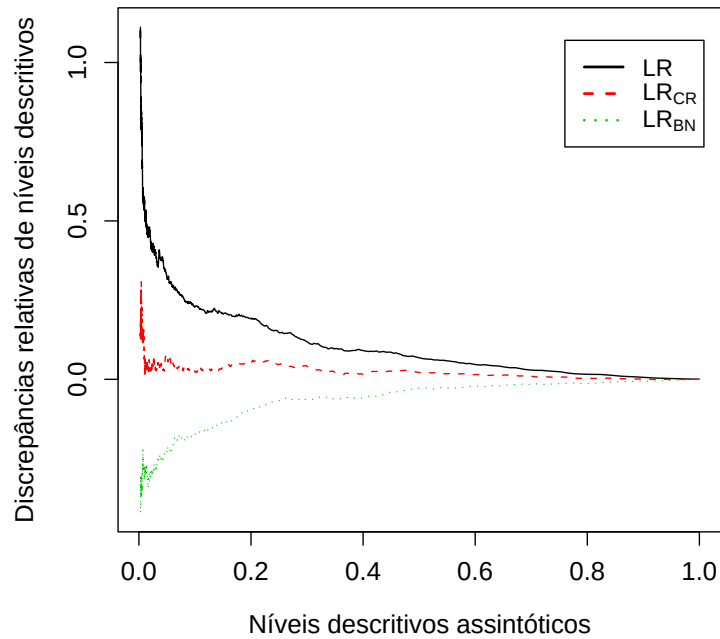


Figura 4.4 *Gráfico das discrepâncias relativas de níveis descritivos em relação a α ($n = 10$, $\alpha = 0,5$ e $\beta = 1$).*



LR encontra-se no patamar mais elevado, o que confirma que este teste é o mais anti-conservativo. Por outro lado, a estatística LR_{BN} mostrou-se conservativa, pois apresentou discrepâncias relativas de quantis em torno de -8%. Finalmente, a estatística LR_{CR} apresentou a curva de discrepância de quantis bem mais próxima da ordenada nula.

Na Figura 4.4 temos o gráfico das discrepâncias relativas entre níveis descritivos dos testes contra os níveis descritivos assintóticos. Observamos que quando se caminha em direção ao extremo da cauda da distribuição χ^2 de referência as discrepâncias relativas de níveis descritivos para os testes LR e LR_{BN} tendem a se afastar rapidamente do eixo horizontal nulo. Porém, este afastamento se dá em direções contrárias. Tais discrepâncias são bem menos acentuadas para o teste LR_{CR} no extremo da cauda.

Aplicações

Neste capítulo apresentaremos aplicações práticas dos desenvolvimentos apresentados até então para a distribuição Birnbaum-Saunders com dois exemplos de conjuntos de dados reais. Para ilustrar o efeito do tamanho amostral, os exemplos considerados consistem de conjuntos de dados com tamanhos diferentes, um envolvendo uma quantidade pequena e o outro, uma quantidade elevada de observações. Em ambos os casos nós assumimos que os dados são independentes e seguem distribuição Birnbaum-Saunders. Realizaremos inferências para os parâmetros de forma e escala usando as funções de verossimilhança perfilada e perfiladas modificadas que foram apresentadas na Seção 3.4.

Todos os resultados foram obtidos utilizando a plataforma computacional **0x**, bem como suas rotinas e funções; ver Doornik (2001). O procedimento utilizado para as maximizações numéricas foi o algoritmo quasi-Newton BFGS.

5.1 Exemplo 1 - Fadiga do alumínio 6061-T6

Os dados utilizados neste exemplo estão apresentados no Quadro 5.1 e foram obtidos de Birnbaum-Saunders (1969b). Estes dados consistem do número de ciclos de estresse sofrido até a falha de 101 tiras de lâmina de alumínio 6061-T6 com corte paralelo na direção do metal. Cada observação foi submetida a uma carga periódica com frequência de 18 ciclos por segundo e estresse máximo de 31 kpsi.

Inicialmente estamos interessados em realizar inferências sobre o parâmetro α . As estimativas de máxima verossimilhança obtidas através da maximização das funções de verossimilhança perfilada e suas versões modificadas são $\hat{\alpha} = 0,17038$, $\hat{\alpha}_{CR} = 0,17125$ e $\hat{\alpha}_{BN} = 0,17122$.

Suponha que estamos interessados em testar a hipótese $H_0 : \alpha = 0,15$ contra $H_1 : \alpha \neq 0,15$. Para este teste, as estatísticas da razão de verossimilhanças baseadas nas funções $\ell_p(\alpha)$, $\ell_{CR}(\alpha)$ e $\ell_{BN}(\alpha)$ são, respectivamente, 3,5771, 3,8421 e 3,8351, que conduzem aos respectivos níveis descritivos 0,05858, 0,04998 e 0,05019. Aqui, o tamanho da amostra é grande ($n = 101$), portanto, como esperado, os testes baseados nas estatísticas da razão de verossimilhanças perfiladas modificadas de Cox e Reid e Barndorff-Nielsen apresentaram níveis descritivos aproximadamente iguais aos do teste da razão de verossimilhanças usual. Mesmo assim, ao nível significância de 5%, o teste baseado na estatística da razão de verossimilhanças perfiladas modificado de Cox e Reid rejeita a hipótese nula, enquanto os demais não a rejeitam. Dessa forma, os testes apresentam conclusões divergentes.

Quadro 5.1 Número de ciclos até a falha da lâmina do alumínio 6061 - T6.

70	90	96	97	99	100	103	104	104	105	107
108	108	108	109	109	112	112	113	114	114	114
116	119	120	120	120	121	121	123	124	124	124
124	124	128	128	129	129	130	130	130	131	131
131	131	131	132	132	132	133	134	134	134	134
134	136	136	137	138	138	138	139	139	141	141
142	142	142	142	142	142	144	144	145	146	148
148	149	151	151	152	155	156	157	157	157	157
158	159	162	163	163	164	166	166	168	170	174
196	212									

Agora, considerando a realização de inferências sobre o parâmetro β , a hipótese nula de interesse é $H_0 : \beta = 125$. Para este teste, as estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas obtidas são $LR = 9,4279$, $LR_{CR} = 9,3338$ e $LR_{BN} = 9,2397$, que conduzem aos respectivos níveis descritivos 0,00214, 0,00225 e 0,00237. Aqui, verificamos que os testes apontam para a mesma conclusão. Isto é, os três testes considerados rejeitam a hipótese nula ao nível de 1% de significância. A estimativa de máxima verossimilhança para β é $\hat{\beta} = 131,81879$.

5.2 Exemplo 2 - Fadiga da chumaceira

Neste segundo exemplo utilizamos os dados apresentados por McCool (1974) para o tempo de fadiga em horas de 10 chumaceiras de um certo tipo. Estes dados também foram utilizados por Cohen et al. (1984) como exemplo ilustrativo para a distribuição Weibull de três parâmetros. Os dados estão apresentados no Quadro 5.2.

Quadro 5.2 Tempo de fadiga em horas das chumaceiras.

152,7	172,0	172,5	173,3	193,0
204,7	216,5	234,9	262,6	422,6

As estimativas de máxima verossimilhança perfilada e perfiladas modificadas para o parâmetro α são $\hat{\alpha} = 0,2825$, $\hat{\alpha}_{CR} = 0,2989$ e $\hat{\alpha}_{BN} = 0,2973$. Suponha que estamos interessados em testar $H_0 : \alpha = 0,21$ contra $H_1 : \alpha \neq 0,21$. Para este teste, obtemos as estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas $LR = 2,1646$, $LR_{CR} = 2,8438$ e $LR_{BN} = 2,7963$, que conduzem aos níveis descritivos 0,1412, 0,0917 e 0,0945, respectivamente. Assim, os testes baseados nas estatísticas LR_{CR} e LR_{BN} rejeitam a hipótese nula, enquanto que o teste baseado na estatística LR não rejeita a hipótese nula ao nível de significância de 10%. Dessa forma, os testes conduzem a conclusões diferentes.

Para o parâmetro β , a estimativa de máxima verossimilhança é $\hat{\beta} = 212,05$.

Suponha que o interesse reside em testar a hipótese $H_0 : \beta = 180$ contra $H_1 : \beta \neq 180$. Para este teste, temos $LR = 2,9417$, $LR_{CR} = 2,6415$ e $LR_{BN} = 2,3414$; estas estatísticas conduzem aos respectivos níveis descritivos 0,0863, 0,1041 e 0,1260. Observamos que estes testes, novamente, conduzem a conclusões divergentes. O teste que rejeita a hipótese nula ao nível de 10% de significância é o teste baseado na estatística LR .

Conclusões

No desenvolvimento deste trabalho apresentamos as características e propriedades da distribuição Birnbaum-Saunders. Derivamos para esta distribuição as funções de verossimilhança perfilada e perfiladas modificadas de Cox e Reid (1987) e Barndorff-Nielsen (1983). Estas funções foram derivadas para o caso (i) em que temos α como parâmetro de interesse e β de perturbação. Para este caso, a função de verossimilhança perfilada modificada de Barndorff-Nielsen foi obtida usando a aproximação de Severini (1999). Para o caso (ii), ou seja, quando β é o parâmetro de interesse e α é o parâmetro de perturbação, usamos a aproximação proposta por Severini (1998) para obter a função de verossimilhança perfilada modificada de Barndorff-Nielsen. Listaremos a seguir as principais conclusões desta dissertação para cada caso separadamente.

Caso (i): Os estimadores $\hat{\alpha}_p$, $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ foram obtidos a partir da maximização de suas respectivas funções de verossimilhanças perfiladas. Diversas simulações foram realizadas para avaliar os comportamentos destes estimadores. As simulações ilustram detalhadamente os comportamentos destes estimadores em situações bem específicas. É importante destacar que os estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ mostraram-se mais centrados em relação ao estimador $\hat{\alpha}_p$ para todos os valores de α considerados. Os estimadores $\hat{\alpha}_{CR}$ e $\hat{\alpha}_{BN}$ apresentaram desempenhos bem semelhantes, principalmente para valores de $\alpha = 0,1$ e $0,5$. Vale salientar que valores pequenos para o parâmetro de forma são comumente verificados em aplicações práticas. Portanto, tanto o estimador $\hat{\alpha}_{CR}$ quanto o estimador $\hat{\alpha}_{BN}$ conduzem a estimativas precisas. Em seguida, estudamos os tamanhos dos testes baseados nas estatísticas LR , LR_{CR} e LR_{BN} . Os testes LR_{CR} e LR_{BN} , no geral, conduziram a taxas de rejeição bem semelhantes.

Estes testes também apresentaram desempenhos consideravelmente superiores em relação ao teste usual LR , principalmente em amostras pequenas e para valores mais elevados do parâmetro de forma. Portanto, recomendamos o uso do teste LR_{CR} ou do teste LR_{BN} por apresentarem desempenhos bastante satisfatórios mesmo em condições adversas.

Caso (ii): Os estimadores perfilados modificados obtidos para o parâmetro β são equivalentes ao estimador $\hat{\beta}$, pois maximizar as funções de verossimilhanças perfiladas modificadas obtidas equivale a maximizar a função de verossimilhança perfilada. Alguns resultados de simulação foram apresentados para o estimador $\hat{\beta}$. Este estimador apresentou comportamento bastante satisfatório, entretanto, para valores elevados do parâmetro α observamos uma deterioração no seu desempenho. Os testes da razão de verossimilhanças LR , LR_{CR} e LR_{BN} foram obtidos a partir de suas respectivas funções de verossimilhanças. Através dos resultados de simulação verificamos que, em geral, o teste LR_{CR} apresentou desempenho superior em relação aos demais. Observamos também que o teste LR_{BN} foi bastante conservativo, principalmente para amostras pequenas e para valores elevados do parâmetro de forma, enquanto que o teste usual LR mostrou-se bastante anti-conservativo nestas mesmas condições. Dessa forma, recomendamos o uso do teste LR_{CR} quando há interesse em testar hipóteses sobre o parâmetro de escala β .

Por fim, concluímos que as funções de verossimilhanças perfiladas modificadas desenvolvidas para a distribuição Birnbaum-Saunders podem melhorar consideravelmente a qualidade das inferências produzidas para o parâmetro de interesse. Os resultados contidos na presente dissertação possuem vasta aplicabilidade para problemas que envolvem a modelagem de desgaste estocástico em geral.

Apêndice A

Neste apêndice, apresentamos os cálculos de algumas quantidades necessárias à obtenção das funções de verossimilhanças perfiladas modificadas propostas por Cox e Reid (1987) e Barndorff-Nielsen (1983) na distribuição Birnbaum-Saunders considerando o caso em que temos como parâmetro de interesse o parâmetro de forma e também quando o parâmetro de interesse é o parâmetro de escala.

O logaritmo da função de verossimilhança do vetor de parâmetros $\theta = (\tau, \phi)^\top$ da distribuição Birnbaum-Saunders é dado por

$$\begin{aligned}\ell(\theta) &= -n \log(\alpha\beta) + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] - \frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) \\ &= -n \log(\alpha\beta) + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] - \frac{1}{2\alpha^2} \left(\frac{\sum_{i=1}^n t_i}{\beta} + \beta \sum_{i=1}^n \frac{1}{t_i} - 2n \right).\end{aligned}$$

Podemos reescrever o logaritmo da função de verossimilhança dado acima como

$$\ell(\theta) = -n \log(\alpha\beta) + \sum_{i=1}^n \log(\beta^{1/2} t_i^{-1/2} + \beta^{3/2} t_i^{-3/2}) - \frac{n}{2\alpha^2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right),$$

em que

$$r = \frac{1}{n} \sum_{i=1}^n t_i, \quad s = \left[\frac{1}{n} \sum_{i=1}^n t_i^{-1} \right]^{-1} \quad \text{e} \quad K(\gamma) = \left[\frac{1}{n} \sum_{i=1}^n (\gamma + t_i)^{-1} \right]^{-1}, \quad \text{para } \gamma > 0.$$

Obtenção da função escore:

As derivadas de primeira ordem com respeito a α e β são dadas, respectivamente, por

$$u_\alpha = \frac{\partial \ell(\theta)}{\partial \alpha} = \frac{-n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)$$

e

$$\begin{aligned}u_\beta &= \frac{\partial \ell(\theta)}{\partial \beta} = -\frac{n}{\beta} + \frac{1}{2} \sum \frac{t_i^{-1/2} \beta^{-1/2} + 3\beta^{1/2} t_i^{-3/2}}{\beta^{1/2} t_i^{-1/2} + \beta^{3/2} t_i^{-3/2}} + \frac{n}{2\alpha^2} \left(\frac{r}{\beta^2} - \frac{1}{s} \right) \\ &= -\frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right) + \frac{n}{2\alpha^2} \left(\frac{r}{\beta^2} - \frac{1}{s} \right).\end{aligned}$$

Obtenção da matriz de informação observada:

As derivadas de segunda ordem com respeito a α e β são dadas por

$$j_{\alpha\beta}(\theta) = \frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \beta} = \frac{n}{\alpha^3} \left(-\frac{r}{\beta^2} + \frac{1}{s} \right),$$

$$j_{\alpha\alpha}(\theta) = \frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = \frac{n}{\beta} - \frac{3n}{\alpha^4} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)$$

e

$$j_{\beta\beta}(\theta) = \frac{\partial^2 \ell(\theta)}{\partial \beta^2} = \frac{n}{\beta^2} - \frac{n}{2} \left(\frac{1}{\beta^2} + \frac{2K'(\beta)}{K^2(\beta)} \right) - \frac{n}{\alpha^2} \frac{r}{\beta^3}.$$

em que o termo $K'(\beta)$ é obtido abaixo:

fazendo

$$u = \frac{1}{n} \sum_{i=1}^n (\beta + t_i)^{-1},$$

tem-se que

$$K'(\beta) = -\frac{u'}{u^2},$$

em que

$$u' = -\frac{1}{n} \sum_{i=1}^n \left[\frac{1}{(\beta + t_i)^2} \right] = -\frac{1}{n} \sum_{i=1}^n (\beta + t_i)^{-2}.$$

Portanto,

$$K'(\beta) = \frac{\frac{1}{n} \sum_{i=1}^n (\beta + t_i)^{-2}}{\left[\frac{1}{n} \sum_{i=1}^n (\beta + t_i)^{-1} \right]^2} = \frac{n \sum_{i=1}^n (\beta + t_i)^{-2}}{\left[\sum_{i=1}^n (\beta + t_i)^{-1} \right]^2}.$$

Obtenção das funções de verossimilhanças perfiladas:

Substituindo α por $\hat{\alpha}_\beta$ no logaritmo da função de verossimilhança original obtém-se

$$\begin{aligned} \ell_p(\beta) &= -n \log(\hat{\alpha}_\beta) - n \log \beta + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] - \frac{n}{2\hat{\alpha}_\beta^2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) \\ &= -\frac{n}{2} \log \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) - n \log \beta + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] \\ &\quad - \frac{n}{2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) \\ &= -\frac{n}{2} \log \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) - n \log \beta + \sum_{i=1}^n \log \left[\left(\frac{\beta}{t_i} \right)^{1/2} + \left(\frac{\beta}{t_i} \right)^{3/2} \right] - \frac{n}{2} \end{aligned}$$

A primeira derivada de $\ell_p(\beta)$ é dada por

$$\frac{\partial \ell_p(\beta)}{\partial \beta} = \frac{n}{2} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} - \frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right).$$

O estimador de máxima verossimilhança de β para α fixado é o mesmo estimador de máxima verossimilhança original ($\hat{\beta}_\alpha = \hat{\beta}$). Substituindo $\hat{\beta}$ por β em $\ell(\theta)$ temos

$$\ell_p(\alpha) = -n \log \alpha - n \log \hat{\beta} + \sum_{i=1}^n \log \left[\left(\frac{\hat{\beta}}{t_i} \right)^{1/2} + \left(\frac{\hat{\beta}}{t_i} \right)^{3/2} \right] - \frac{n}{2\alpha^2} \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right).$$

A primeira derivada da função $\ell_p(\alpha)$ é dada por

$$\frac{\partial \ell_p(\alpha)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right).$$

Obtenção das funções de verossimilhanças perfiladas modificadas de Cox e Reid:

Quando α é o parâmetro de interesse, a função de verossimilhança perfilada de Cox e Reid é dada por

$$\ell_{CR}(\alpha) = \ell_p(\alpha) - \frac{1}{n} \log |j_{\beta\beta}(\alpha, \hat{\beta}_\alpha)|,$$

em que

$$j_{\beta\beta}(\alpha, \hat{\beta}_\alpha) = -\frac{n}{\hat{\beta}^2} + \frac{n}{2} \left(\frac{1}{\hat{\beta}^2} + \frac{2K'(\hat{\beta})}{K^2(\hat{\beta})} \right) + \frac{n}{\alpha^2} \frac{r}{\hat{\beta}^3},$$

sendo a primeira derivada de $\ell_{CR}(\alpha)$ dada por

$$\begin{aligned} \frac{\partial \ell_{CR}(\alpha)}{\partial \alpha} &= -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right) \\ &\quad - \frac{1}{2} \left(-\frac{n}{\hat{\beta}^2} + \frac{n}{2} \left(\frac{1}{\hat{\beta}^2} + \frac{2K'(\hat{\beta})}{K^2(\hat{\beta})} \right) + \frac{n}{\alpha^2} \frac{r}{\hat{\beta}^3} \right)^{-1} \left(-\frac{r}{\hat{\beta}^3} \frac{2n}{\alpha^3} \right) \\ &= -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left[\left(\frac{r}{\hat{\beta}} + \frac{\hat{\beta}}{s} - 2 \right) - \frac{1}{2} \left(\frac{n}{\hat{\beta}^2} - \frac{n}{2} \left(\frac{1}{\hat{\beta}^2} + \frac{2K'(\hat{\beta})}{K^2(\hat{\beta})} \right) - \frac{n}{\alpha^2} \frac{r}{\hat{\beta}^3} \right)^{-1} \frac{2r}{\hat{\beta}^3} \right]. \end{aligned}$$

Quando β é o parâmetro de interesse, a função de verossimilhança perfilada modificada de Cox e Reid é dada por

$$\ell_{CR}(\beta) = \ell_p(\beta) - \frac{1}{n} \log |j_{\alpha\alpha}(\hat{\alpha}_\beta, \beta)|,$$

em que

$$\begin{aligned} j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta) &= -n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} + 3n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right) \\ &= -n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} + 3n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} = 2n \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1}, \end{aligned}$$

sendo a primeira derivada de $\ell_{CR}(\beta)$ dada por

$$\begin{aligned} \frac{\partial \ell_{CR}(\beta)}{\partial \beta} &= -\frac{n}{2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) - \frac{n}{\beta} + \frac{n}{2} \left(\frac{1}{\beta} + \frac{2}{K(\beta)} \right) \\ &\quad + \frac{1}{2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1} \left(\frac{1}{s} - \frac{r}{\beta^2} \right). \end{aligned}$$

Obtenção das funções de verossimilhanças perfiladas de Barndorff-Nielsen:

Quando α é o parâmetro de interesse, a função de verossimilhança perfilada modificada de Barndorff-Nielsen é dada por

$$\ell_{BN}(\alpha) = \ell_p(\alpha) + \frac{1}{2} \log |j_{\beta\beta}(\alpha, \widehat{\beta}_\alpha)| - \log |\ell_{\beta;\widehat{\beta}}(\alpha, \widehat{\beta}_\alpha)|.$$

A quantidade $\ell_{\beta;\widehat{\beta}}(\alpha, \widehat{\beta}_\alpha)$ pode ser aproximada por

$$\widehat{I}(\theta; \theta_0) = \sum_{j=1}^n \ell_{\beta}^{(j)}(\theta_0) \ell_{\beta}^{(j)}(\theta),$$

em que

$$\ell_{\beta}^{(j)}(\theta) = \frac{1}{\beta} + \frac{1}{2} \frac{t_j^{-1/2} \beta^{-1/2} + 3\beta^{1/2} t_j^{-3/2}}{t_j^{-1/2} \beta^{1/2} + \beta^{3/2} t_j^{-3/2}} + \frac{1}{2\alpha^2} \left(\frac{t_j}{\beta^2} - \frac{1}{t_j} \right),$$

$$\ell_{\beta}^{(j)}(\theta_0) = \frac{1}{\beta_0} + \frac{1}{2} \frac{t_j^{-1/2} \beta_0^{-1/2} + 3\beta_0^{1/2} t_j^{-3/2}}{t_j^{-1/2} \beta_0^{1/2} + \beta_0^{3/2} t_j^{-3/2}} + \frac{1}{2\alpha_0^2} \left(\frac{t_j}{\beta_0^2} - \frac{1}{t_j} \right),$$

$\theta_0 = (\hat{\alpha}, \hat{\beta})$ e $\theta = (\alpha, \hat{\beta})$. Observamos que $\beta_0 = \beta$. Assim,

$$\begin{aligned}
\ell_{\beta}^{(j)}(\theta_0)\ell_{\beta}^{(j)}(\theta) &= \left(-\frac{1}{\beta} + \frac{1}{2}A_j + \frac{1}{2\alpha_0^2}B_j\right) \left(-\frac{1}{\beta} + \frac{1}{2}A_j + \frac{1}{2\alpha^2}B_j\right) \\
&= \frac{1}{4} \left(-\frac{2}{\beta} + A_j + \frac{1}{\alpha_0^2}B_j\right) \left(-\frac{2}{\beta} + A_j + \frac{1}{\alpha^2}B_j\right) \\
&= \frac{1}{4} \left(\frac{4}{\beta^2} - \frac{2}{\beta}A_j - \frac{2}{\beta\alpha^2}B_j - \frac{2}{\beta}A_j + A_j^2 + \frac{1}{\alpha^2}A_jB_j \right. \\
&\quad \left. - \frac{2}{\beta\alpha_0^2}B_j + \frac{1}{\alpha_0^2}A_jB_j + \frac{1}{\alpha^2\alpha_0^2}B_j^2\right) \\
&= \frac{1}{4} \left[\frac{4}{\beta^2} - \frac{4A_j}{\beta} - \frac{2B_j}{\beta} \left(\frac{1}{\alpha^2} + \frac{1}{\alpha_0^2}\right) + \left(\frac{1}{\alpha^2} + \frac{1}{\alpha_0^2}\right)A_jB_j + A_j^2 + \frac{B_j^2}{\alpha^2\alpha_0^2}\right] \\
&= \frac{1}{4} \left[\frac{4}{\beta^2} - \frac{4}{\beta}A_j + \left(\frac{1}{\alpha^2} + \frac{1}{\alpha_0^2}\right) \left(A_jB_j - \frac{2}{\beta}B_j\right) + A_j^2 + \frac{1}{\alpha^2\alpha_0^2}B_j^2\right]
\end{aligned}$$

Assim, temos

$$\begin{aligned}
\hat{I}(\theta_0; \theta) &= \sum_{j=1}^n \ell_{\beta}^{(j)}(\theta_0)\ell_{\beta}^{(j)}(\theta) = \frac{1}{4} \left[\frac{4n}{\beta^2} - \frac{4}{\beta} \sum_{j=1}^n A_j + \left(\frac{1}{\alpha^2} + \frac{1}{\alpha_0^2}\right) \right. \\
&\quad \left. \times \left(\sum_{j=1}^n A_jB_j - \frac{2}{\beta} \sum_{j=1}^n B_j \right) + \sum_{j=1}^n A_j^2 + \frac{1}{\alpha^2\alpha_0^2} \sum_{j=1}^n B_j^2 \right],
\end{aligned}$$

em que

$$A_j = \frac{t_j^{-1/2}\beta^{-1/2} + 3\beta^{1/2}t_j^{-3/2}}{t_j^{-1/2}\beta^{1/2} + \beta^{3/2}t_j^{-3/2}} \quad \text{e} \quad B_j = \left(\frac{t_j}{\beta^2} - \frac{1}{t_j}\right).$$

Note que

$$\sum_{j=1}^n A_j = n \left(\frac{1}{\beta} + \frac{2}{K(\beta)}\right) \quad \text{e} \quad \sum_{j=1}^n B_j = n \left(\frac{r}{\beta^2} - \frac{1}{s}\right).$$

A derivada da quantidade $\log |I(\theta_0; \theta)|$ é dada por

$$\begin{aligned}
\frac{\partial \log |\hat{I}(\theta_0; \theta)|}{\partial \alpha} &= 4 \left[\frac{4n}{\beta^2} - \frac{4}{\beta} \sum_{j=1}^n A_j + \left(\frac{1}{\alpha^2} + \frac{1}{\alpha_0^2}\right) \left(\sum_{j=1}^n A_jB_j - \frac{2}{\beta} \sum_{j=1}^n B_j \right) \right. \\
&\quad \left. + \sum_{j=1}^n A_j^2 + \frac{1}{\alpha^2\alpha_0^2} \sum_{j=1}^n B_j^2 \right]^{-1} \\
&\quad \times \left(\frac{4}{\alpha^3\beta} \sum_{j=1}^n B_j - \frac{2}{\alpha^3} \sum_{j=1}^n A_jB_j - \frac{2}{\alpha^3\alpha_0^2} \sum_{j=1}^n B_j \right).
\end{aligned}$$

Quando β é parâmetro de interesse, a função de verossimilhança perfilada modificada de Barndorff-Nielsen é dada por

$$\ell_{BN}(\beta) = \ell_p(\beta) + \frac{1}{2} \log |j_{\alpha\alpha}(\hat{\alpha}_\beta, \beta)| - \log |\ell_{\alpha; \hat{\alpha}}(\hat{\alpha}_\beta, \beta)|.$$

A quantidade $\ell_{\alpha; \hat{\alpha}}(\hat{\alpha}_\beta, \beta)$ pode ser aproximada por

$$I(\theta; \theta_0) = E \left\{ \ell_\alpha^{(j)}(\theta_0) \ell_\alpha^{(j)}(\theta) \right\},$$

em que

$$\begin{aligned} \ell_\alpha(\theta) &= -\frac{n}{\alpha} + \frac{n}{\alpha^3} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right), \\ \ell_\alpha(\theta_0) &= -\frac{n}{\alpha_0} + \frac{n}{\alpha_0^3} \left(\frac{r}{\beta_0} + \frac{\beta_0}{s} - 2 \right) \end{aligned}$$

e $\theta_0 = (\hat{\alpha}, \hat{\beta})$ e $\theta = (\hat{\alpha}_\beta, \beta)$. Sabemos que $E[\ell_\alpha(\theta_0)] = 0$. Dessa forma, temos

$$\begin{aligned} E \left\{ \ell_\alpha^{(j)}(\theta_0) \ell_\alpha^{(j)}(\theta) \right\} &= -\frac{n}{\alpha} E[\ell_\alpha(\theta_0)] + \frac{n}{\alpha^3 \beta} E[\ell_\alpha(\theta_0)] + \frac{n\beta}{\alpha^3} E[\ell_\alpha(\theta_0)] - \frac{2n}{\alpha^3} E[\ell_\alpha(\theta_0)] \\ &= \frac{n}{\alpha^3 \beta} \alpha_0 \beta_0 + \frac{n\beta}{\alpha^3} \frac{\alpha_0}{\beta_0} = \frac{n\alpha_0}{\alpha^3} \left(\frac{\beta_0}{\beta} + \frac{\beta}{\beta_0} \right) = \frac{n\hat{\alpha}}{\hat{\alpha}_\beta^3} \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right). \end{aligned}$$

A derivada de primeira ordem de $\ell_{BN}(\beta)$ é dada por

$$\frac{\partial \ell_{BN}(\beta)}{\partial \beta} = \frac{\partial \ell_p(\beta)}{\partial \beta} + \frac{1}{2} \frac{\partial \log |j_{\alpha\alpha}(\hat{\alpha}_\beta, \beta)|}{\partial \beta} - \frac{\partial \log \left| \frac{n\hat{\alpha}}{\hat{\alpha}_\beta^3} \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right) \right|}{\partial \beta},$$

em que a derivada da quantidade $\log \frac{n\hat{\alpha}}{\hat{\alpha}_\beta^3} \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right)$ é dada por

$$\begin{aligned} \frac{\partial \log \left| \frac{n\hat{\alpha}}{\hat{\alpha}_\beta^3} \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right) \right|}{\partial \beta} &= \frac{\partial \log n\hat{\alpha}}{\partial \beta} - \frac{\partial \log \hat{\alpha}_\beta^3}{\partial \beta} + \frac{\partial \log \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right)}{\partial \beta} \\ &= \frac{1}{\hat{\alpha}_\beta^3} \frac{3}{2} \left(\frac{r}{\beta} + \frac{\beta}{s} - 2 \right)^{-1/2} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) - \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right)^{-1} \left(\frac{1}{\hat{\beta}} - \frac{\hat{\beta}}{\beta^2} \right) \\ &= \frac{3}{2\hat{\alpha}_\beta^2} \left(\frac{1}{s} - \frac{r}{\beta^2} \right) - \left(\frac{\hat{\beta}}{\beta} + \frac{\beta}{\hat{\beta}} \right)^{-1} \left(\frac{1}{\hat{\beta}} - \frac{\hat{\beta}}{\beta^2} \right) \end{aligned}$$

Apêndice B

Neste apêndice, apresentamos os programas utilizados para geração dos resultados de simulação obtidos nesta dissertação. São apresentados dois programas, a saber: o programa de simulação utilizado no Capítulo 4 para geração dos resultados de simulação quando α é o parâmetro de interesse e o programa utilizado para gerar os resultados das aplicações práticas desenvolvidas no Capítulo 5. Esses dois programas fazem referência aos módulos `likelihood.ox` e `likelihood.h`. O primeiro módulo contém as funções que permitem o cálculo das verossimilhanças perfiladas modificadas e o segundo contém as diretrizes de ligação do módulo `likelihood.ox`. O programa de simulação para o parâmetro β foi omitido devido à sua semelhança com o programa que compreende as simulações para o parâmetro de forma α .

Programa de simulação

```
/******  
PROGRAMA: program1A.ox  
  
DESCRICAO: Programa de simulacao principal. Simula as taxas de rejeicao  
e calcula as estimativas de maxima verossimilhanca baseado nas  
verossimilhanças perfiladas modificadas.  
  
AUTOR: Carlos Gadelha Junior.  
DATA: 11 de Abril de 2005.  
ULTIMA MODIFICACAO: 12 de Janeiro de 2006  
*****/  
  
#include <oxstd.h>  
#include <oxprob.h>  
#import <maximize>  
#import <solvenle>  
#include <oxfloat.h>  
#import <likelihood>  
  
main()  
{  
    decl alpha, beta, et, n, nrepls, nsamples, i, j, X, test, rv, results,  
    criticalvalue, dfunc, vp, ir, x, c1, c2, c3, alphahat, pontualresults,  
    initialvalue, a, b, c;  
  
    decl chisq, qperf, qcoxr, aux, qbarn, pbarn, results2, criticalvalue2,  
    pchisq,pperf,pcoxr,n_cv2,n_cv3,results3, criticalvalue3;
```

```

// Inicializando o tempo
et = timer();

// Gerador de numeros aleatorios
ranseed("GM");          // George Marsaglia
ranseed({111,1111});    // semente do gerador

// Especificando os parametros da distribuicao
alpha = 0.5;    beta = 1.00;

// Especificando os parametros de simulacao
n = <10,25,50>;    nrepls = 100;    nsamples = 1;
initialvalue = <0.5>;    // para alpha

// Inicializando quantidades para o teste da razao de verossimilhanças
rv = zeros(3,nrepls);
criticalvalue = <2.705543,3.841459,6.634897,7.879439>;
results = zeros(3,4);
alphahat = zeros(3,nrepls);
pontualresults = zeros(3,7);

criticalvalue2 = range(0,9,0.01); // Variavel auxiliar para graficos
n_cv2 = columns(criticalvalue2);
results2 = zeros(3,n_cv2); // Variavel auxiliar para graficos
criticalvalue3 = quanchi(<0.50,0.85,0.90,0.95,0.97,0.99,
0.995,0.999,0.9995,0.9999>,1); //Aux. para tabela
n_cv3 = columns(criticalvalue3);
results3 = zeros(4,n_cv3); // Variavel auxiliar para graficos

a=b=c=0; c1=c2=c3=0;

// Salvando resultados (usados para os graficos)
decl file = fopen("rvA.txt", "w");
decl file2 = fopen("probA.txt","w");

// Loop para o tamanho amostral
for(j=0;j<nsamples;++j)
{
    N = n[j];

    // Loop para o numero de replicas
    for(i=0;i<nrepls;)
    {
        // Gerando vetor aleatorio com distribuicao  $N(0, \alpha^2/4)$ 
        X = (alpha/2).*rann(1,n[j]);
        // Gerando vetor aleatorio com distribuicao Birnbaum-Saunders
        t = beta.*(1 + 2*X.^2 + 2*X.*(1 + X.^2).^0.5);

        u = meanr(t);          //media aritmetica amostral
        v = (meanr(t.^-1))^-1; //media harmonica amostral

        // Calculando o EMV de beta

```

```

x = <1.00>;
SolveNLE(nle_beta, &x);
betahat = x;
////////////////////////////////////

B = (t.*(betahat.^-2) - (t.^-1));
A = 0.5*( (t*betahat).^0.5 + 3*(betahat^0.5*t.^-1.5) )
.*( ( (t.^-0.5*betahat.^0.5) + (betahat^1.5*t.^-1.5) ).^-1 );

dfunc = 0; // valor inicial da funcao
vp = initialvalue; // valor inicial do parametro
// maximizacao de alpha
ir = MaxBFGS(fBS_profileAlpha, &vp, &dfunc, 0, FALSE);
alphahat[0][i] = vp;
a = ir; // checando convergencia
c1 = (ir!=MAX_CONV && ir!=MAX_WEAK_CONV) ? (c1 + 1) : c1;

// teste da razao de verossimilhanças perfiladas para alpha
rv[0][i] = 2*(- N*log(vp) - N*log(betahat) + sumr(log((betahat^0.5)*(t.^-0.5)
+ (betahat^1.5)*(t.^-1.5))) + (-N/(2*vp^2))*(u/betahat+betahat/v-2)
- (- N*log(alpha) - N*log(betahat) + sumr(log((betahat^0.5)
*(t.^-0.5) + (betahat^1.5)*(t.^-1.5))) + (-N/(2*alpha^2))
*(u/betahat + betahat/v - 2))));

K = (meanr((betahat + t).^-1)).^-1;
dK = (N*sumr((betahat + t).^-2))/((sumr((betahat + t).^-1)).^2);
alphazero = (u/betahat + betahat/v - 2).^0.5;

dfunc = 0; // valor inicial da funcao
vp = initialvalue; // valor inicial do parametro
ir = MaxBFGS(fBS_profileAlphaCR, &vp, &dfunc, 0, FALSE); // maximizacao de CR
alphahat[1][i] = vp;
b = ir; // checando convergencia
c2 = (ir!=MAX_CONV && ir!=MAX_WEAK_CONV) ? (c2 + 1) : c2;

//teste da razao de verossimilhanças perfiladas de Cox&Reid para alpha
rv[1][i] = 2*(- N*log(vp) - N*log(betahat) + sumr(log((betahat^0.5)*(t.^-0.5)
+ (betahat^1.5)*(t.^-1.5)))+(-N/(2*vp^2))*(u/betahat+betahat/v-2)
- 1/2*log(-(N/(betahat^2)-N/2*(1/(betahat^2) + 2*dK/(K^2))
- (N/(vp^2))*(u/(betahat^3)))) - (-N*log(alpha) - N*log(betahat)
+ sumr(log((betahat^0.5)*(t.^-0.5) + (betahat^1.5)*(t.^-1.5)))
+ (-N/(2*alpha^2))*(u/betahat + betahat/v - 2) - 1/2*log(-(N
/(betahat^2)-N/2*(1/(betahat^2) + 2*dK/(K^2))-(N/(alpha^2))
*(u/(betahat^3))))));

dfunc = 0; // valor inicial da funcao
vp = initialvalue; // valor inicial do parametro
// maximizacao de BN
ir = MaxBFGS(fBS_profileAlphaBN, &vp, &dfunc, 0, FALSE);
alphahat[2][i] = vp;
c = ir; // checando convergencia
c3 = (ir!=MAX_CONV && ir!=MAX_WEAK_CONV) ? (c3 + 1) : c3;

```

```

//teste da razao de verossimilhanças perfiladas Barndorff-Nielsen para alpha
rv[2][i] = 2*(- N*log(vp)-N*log(betahat)+sumr(log(((betahat^0.5).*(t.^-0.5))
+ ((betahat^1.5)*(t.^-1.5))))+(-N/(2*(vp^2)))*(u/betahat + betahat
/v - 2) + 1/2*log(((N/(betahat^2)+N/2*(1/(betahat^2)+2*dK/(K.^2))
+ (N/(vp^2))*(u/(betahat^3)))))) - log((N/(betahat^2) - (2/betahat)
*sumr(A) + (1/(vp^2) + 1/(alphazero^2)).*(0.5*sumr(A.*B)
-(0.5*betahat^-1)*sumr(B))+sumr(A.^2)+(0.25)*(vp^-2*alphazero^-2)
*sumr(B.^2)))-(-N*log(alpha)-N*log(betahat)+sumr(log(((betahat^0.5)
.*(t.^-0.5)) + ((betahat^1.5)*(t.^-1.5)))) + (-N/(2*(alpha.^2)))
*(u/betahat + betahat/v - 2) + 1/2*log(((N/(betahat.^2)+N/2
*(1/(betahat.^2) + 2*dK/(K.^2)))+(N/(alpha.^2))*(u/(betahat.^3))))))
-log( determinant(N/(betahat.^2) - (2/betahat)*sumr(A)+(1/(alpha.^2)
+ 1/(alphazero.^2)).*(0.5*sumr(A.*B) - (0.5*betahat^-1)*sumr(B))
+ sumr(A.^2) + (0.25)*(alpha.^-2*alphazero.^-2)*sumr(B.^2) ))
));

//checando convergencia (para os tres estimadores)
i = ((a==MAX_CONV|a==MAX_WEAK_CONV) && (b==MAX_CONV|b==MAX_WEAK_CONV )
&& (c==MAX_CONV|c==MAX_WEAK_CONV)) ? ++i : i;
} // fim do loop para o numero de replicas

//Estimated size (Rejection rate)...
for(i=0;i<=3;++i)
{
    results[0][i] = rejection(rv[0][ ],criticalvalue[i],nrepls);
    results[1][i] = rejection(rv[1][ ],criticalvalue[i],nrepls);
    results[2][i] = rejection(rv[2][ ],criticalvalue[i],nrepls);
}

// Estimacao pontual:
pontualresults[][0] = meanr(alphahat);
pontualresults[0][1] = bias(alphahat[0][ ],alpha);
pontualresults[1][1] = bias(alphahat[1][ ],alpha);
pontualresults[2][1] = bias(alphahat[2][ ],alpha);
pontualresults[][2] = varr(alphahat);
pontualresults[0][3] = eqm(alphahat[0][ ],alpha);
pontualresults[1][3] = eqm(alphahat[1][ ],alpha);
pontualresults[2][3] = eqm(alphahat[2][ ],alpha);
pontualresults[][4] = fabs((pontualresults[][1]/alpha)*100);
pontualresults[][5] =( 16*(pontualresults[][0].^2).*(11
*(pontualresults[][0].^2) + 6) ) ./((5*(pontualresults[][0].^2) + 4).^3);
pontualresults[][6] =3 + (( 6*(pontualresults[][0].^2).*(93
*(pontualresults[][0].^2) + 41) ) ./((5*(pontualresults[][0].^2) + 4).^2));

// Gerando dados para os graficos
if (j==0){

// Dados: discrepacia relativa de quantil
chisq = quanchi(range(0.05,0.999,0.001),1);
qperf = (quantilec(rv[0][ ],range(0.05,0.999,0.001)) - chisq')./(chisq');
qcoxr = (quantilec(rv[1][ ],range(0.05,0.999,0.001)) - chisq')./(chisq');
qbarn = (quantilec(rv[2][ ],range(0.05,0.999,0.001)) - chisq')./(chisq');
fprintf(file, "%9.4f", qperf ~ qcoxr ~ qbarn);

```

```

// Dados: discrepacia relativa de niveis descritivos
for(i=0;i<n_cv2;++i)
{
    results2[0][i] =(rejection(rv[0][],criticalvalue2[i],nrepls)/100);
    results2[1][i] =(rejection(rv[1][],criticalvalue2[i],nrepls)/100);
    results2[2][i] =(rejection(rv[2][],criticalvalue2[i],nrepls)/100);
}
pchisq = 1-probchi(criticalvalue2,1);
pperf = (results2[0][] - pchisq)/(pchisq);
pcoxr = (results2[1][] - pchisq)/(pchisq);
pbarn = (results2[2][] - pchisq)/(pchisq);
fprintf(file2, "%9.4f", pperf ~ pcoxr ~ pbarn);

// Tabela: numero de rejeicoes
for(i=0;i<n_cv3;++i)
{
    results3[1][i] =(rejection(rv[0][],criticalvalue3[i],nrepls)/100)*nrepls;
    results3[2][i] =(rejection(rv[1][],criticalvalue3[i],nrepls)/100)*nrepls;
    results3[3][i] =(rejection(rv[2][],criticalvalue3[i],nrepls)/100)*nrepls;
}
results3[0][] = (1 - <0.50,0.85,0.90,0.95,0.97,0.99,0.995,
0.999,0.9995,0.9999>).*nrepls;

}
// fim dos graficos

// Imprimindo resultados de simulacao:
print( "\n\t\t SIMULATION RESULTS: BIRNBAUM-SAUNDERS
(ALPHA = ", "%1.2f", alpha, "; BETA = ", "%1.2f", beta, ")");
print( "\n\t\t ----- \n");
print( "\n\t\t OX PROGRAM: ", oxfilename(0) );
print( "\n\t\t OX VERSION: ", oxversion() );
print( "\n\t\t NUM. REPLICATIONS: ", nrepls );
print( "\n\t\t NUM. OBSERVATIONS: ", n[j] );

println("");
println("\n *** Estimated size of Likelihood test ***\n", "%r",
{"L-profile", "L-Cox&Reid", "L-BN"},
"%c", {"10%", "5%", "1%", "0.5%"}, "%10.3f", results);

println("\n\n *** Pontual estimation ***\n", "%r", {"L-profile", "L-Cox&Reid",
"L-BN"}, "%c", {"media", "vies", "variancia", "EQM", "RB", "assimetria", "curtose"},
"%10.5f", pontualresults);

println("\n\n *** Mean and Variance of the statistic ***\n", "%r",
{"mean", "variance"} "%c", {"Qui-quadrado", "L-profile", "L-Cox&Reid", "L-Barn-N"},
"%10.3f", (1.00 ~ (meanr(rv))) | (2.00 ~ (varr(rv))));

println("\n\n *** Quantiles of the statistic ***\n", "%r",
{"85.0", "90.0", "95.0", "97.0", "99.0", "99.5"},
"%c", {"Qui-quadrado", "L-profile", "L-Cox&Reid", "L-Barn-N"}, "%10.3f",

```

```

        (quanchi(<0.85,0.90,0.95,0.97,0.99,0.995>,1))'~
        quantilec(rv',<0.85,0.90,0.95,0.97,0.99,0.995>));

    if (j==0){
    println("\n\n *** Number of rejections ***\n","%r",
        {"50","15","10","5","3","1","0.5","0.1","0.05","0.01"},
        "%c",{"Exato","L-profile"," L-Cox&Reid","L-Barn-N"}, "%10.3f",
        results3');
    }

    // Imprimindo numero de nao convergencias
    println("\nNum. No Conv. Profile:", c1);
    println("Num. No Conv. CoxReid:", c2);
    println("Num. No Conv. Barndorff-Nielsen:", c3);

    c1=c2=c3=0;

    } // fim do loop do tamanho amostral

    fclose(file); //fechando arquivo
    fclose(file2); //fechando arquivo

    println("\ntempo:",timespan(et));

} // fim do main()

rejection(const vector, const criticalvalue, const nrepls)
{
    return (sumr(fabs(vector) .> criticalvalue)/nrepls) * 100;
}

bies(const vector, const parameter)
{
    return (meanr(vector) - parameter);
}

eqm(const vector, const parameter)
{
    return (varr(vector) + (bies(vector,parameter))^2);
}

/*****
PROGRAMA: likelihook.ox
DESCRICAO: Modulo de funcoes contento os logaritmos das funcoes de
            verossimilhanças perfiladas modificadas
*****/

#include <oxstd.h>

//Equao nao-linear de beta
nle_beta(const avF, const vX)
{

```

```

    avF[0] = vX[0]^2 + v*(u + (meanr((vX[0] + t).^ -1))^ -1) - vX[0]
        *(2*v + (meanr((vX[0] + t).^ -1))^ -1);
    return 1;
}

//Verossimilhanca perfilada para alpha
fBS_profileAlpha(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = - N*log(vP[0]) - N*log(betahat) + sumr(log((betahat^0.5)*(t.^ -0.5)
        + (betahat^1.5)*(t.^ -1.5))) + (-N/(2*vP[0]^2))
        *(u/betahat + betahat/v - 2);

    if (avScore)
    {
        avScore[0] = -N/vP[0] + (N/(vP[0]^3))*(u/betahat + betahat/v - 2);
    }
    return 1;
}

//Verossimilhanca perfilada modificada de Cox&Reid para alpha
fBS_profileAlphaCR(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = - N*log(vP[0]) - N*log(betahat) + sumr(log((betahat^0.5)*(t.^ -0.5)
        + (betahat^1.5)*(t.^ -1.5))) + (-N/(2*vP[0]^2))
        *(u/betahat + betahat/v - 2) - 1/2*log((( -N/(betahat^2)+N/2
        *(1/(betahat^2) + 2*dK/(K^2))+N/(vP[0]^2))*(u/(betahat^3)))) );

    if (avScore)
    {
        avScore[0] = -N/vP[0] + (N/(vP[0]^3))*(u/betahat + betahat/v - 2) - 1/2
            *((( -N/(betahat^2)+(N/2)*(1/(betahat^2) + (2*dK)/(K^2))
            + (N/(vP[0]^2))*(u/(betahat^3)))) ^ -1))*determinant((-2*u
            *N/(betahat^3*vP[0]^3))) );
    }
    return 1;
}

//Verossimilhanca perfilada modificada de Barndorff-Nielsen para alpha
fBS_profileAlphaBN(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = - N*log(vP[0]) - N*log(betahat) + sumr(log(((betahat^0.5).*(t.^ -0.5))
        + ((betahat^1.5)*(t.^ -1.5)))) + (-N/(2*(vP[0]^2)))*(u/betahat
        + betahat/v - 2) + 1/2*log((( -N/(betahat^2)+N/2*(1/(betahat^2)
        + 2*dK/(K.^2))+N/(vP[0]^2))*(u/(betahat^3)))) ) )
        -log((N/(betahat^2) - (2/betahat) * sumr(A) + (1/(vP[0]^2)
        + 1/(alphazero^2)).*(0.5*sumr(A.*B) - (0.5*betahat^-1)*sumr(B))
        + sumr(A.^2) + (0.25)*(vP[0]^ -2*alphazero^-2)*sumr(B.^2) ));

    if (avScore)
    {
        avScore[0] = -N/vP[0] + (N/(vP[0]^3))*(u/betahat + betahat/v - 2) + 1/2
            *((( -N/(betahat^2)+(N/2)*(1/(betahat^2) + (2*dK)/(K^2))

```



```

        + (N/(vP[0]^2))*(u/(betahat^3))) )^-1))*((-2*u*N/(betahat^3
        *vP[0]^3)))) + (((N/(betahat^2) - (2/betahat)*sumr(A)
        + (1/(vP[0]^2) + 1/(alphazero^2)).*(0.5*sumr(A.*B)
        - (0.5*betahat^-1)*sumr(B)) + sumr(A.^2) + (0.25)
        *(vP[0]^-2*alphazero^-2)*sumr(B.^2) ) )^-1)
        *( (0.5*alphazero^-2*vP[0]^-3)*sumr(B.^2) + vP[0]^-3
        *(sumr(A.*B) - betahat^-1*sumr(B)));
    }
    return 1;
}

//Verossimilhanca perfilada para beta
fBS_profileBeta(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = -N/2*log(u/vP[0] + vP[0]/v -2) - N*log(vP[0]) + sumr(log((vP[0]^0.5)
        *(t.^-0.5) + (vP[0]^1.5)*(t.^-1.5)));

    if (avScore)
    {
        avScore[0] = -(N/2)*((-u/(vP[0]^2) + 1/v)/(u/vP[0] + vP[0]/v - 2)) - N/vP[0]
            + (N/2)*( 1/vP[0] + (2/N)*sumr((vP[0]+t).^-1) );
    }
    return 1;
}

//Verossimilhanca perfilada modificada de Cox&Reid para beta
fBS_profileBetaCR(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = -N/2*log(u/vP[0] + vP[0]/v -2) - N*log(vP[0]) + sumr(log((vP[0]^0.5)
        *(t.^-0.5) + (vP[0]^1.5)*(t.^-1.5))) - 1/2*log((2*N)/((u/vP[0])
        + (vP[0]/v)-2));

    if (avScore)
    {
        avScore[0] = -(N/2)*((-u/(vP[0]^2) + 1/v)/(u/vP[0] + vP[0]/v - 2)) - N/vP[0]
            + (N/2)*( 1/vP[0] + (2/N)*sumr((vP[0]+t).^-1) ) + (1/2)
            *( (u/vP[0] + vP[0]/v - 2)^-1 )*(1/v - u/(vP[0]^2));
    }
    return 1;
}

//Verossimilhanca perfilada modificada de Barndorff-Nielsen para beta
fBS_profileBetaBN(const vP, const adFunc, const avScore, const amHessian)
{
    adFunc[0] = - N/2*log(u/vP[0] + vP[0]/v -2) - N*log(vP[0]) + sumr(log((vP[0]^0.5)
        *(t.^-0.5) + (vP[0]^1.5)*(t.^-1.5))) + 1/2*log((2*N)/((u/vP[0])
        + (vP[0]/v)-2)) - log(((N*alphazero)/((u/vP[0] + vP[0]/v -2).^1.5))
        *(betazero/vP[0] * vP[0]/betazero));

    if (avScore)
    {
        avScore[0] = -(N/2)*((-u/(vP[0]^2) + 1/v)/(u/vP[0] + vP[0]/v - 2)) - N/vP[0]
            + (N/2)*( 1/vP[0] + (2/N)*sumr((vP[0]+t).^-1) ) - (1/2)

```

```

        *( (u/vP[0] + vP[0]/v - 2)^-1 )*(1/v - u/(vP[0]^2)) + (3/2)
        *((u/vP[0] + vP[0]/v - 2)^-1)*(1/v - u/(vP[0]^2)) - ((betazero
/vP[0] + vP[0]/betazero)^-1)*(1/betazero - betazero/(vP[0]^2));
    }
    return 1;
}

/*****
PROGRAMA: likelihood.h
DESCRICAO: Modulo contendo as diretrizes de ligacao do modulo
           likelihood.ox
*****/

//Equacao nao-linear de beta
nle_beta(const avF, const vX);

//Verossimilhanca perfilada para alpha
fBS_profileAlpha(const vP, const adFunc, const avScore, const amHessian);

//Verossimilhanca perfilada modificada de Cox&Reid para alpha
fBS_profileAlphaCR(const vP, const adFunc, const avScore, const amHessian);

//Verossimilhanca perfilada modificada de Barndorff-Nielsen para alpha
fBS_profileAlphaBN(const vP, const adFunc, const avScore, const amHessian);

//Verossimilhanca perfilada para beta
fBS_profileBeta(const vP, const adFunc, const avScore, const amHessian);

//Verossimilhanca perfilada modificada de Cox&Reid para beta
fBS_profileBetaCR(const vP, const adFunc, const avScore, const amHessian);

//Verossimilhanca perfilada modificada de Barndorff-Nielsen para beta
fBS_profileBetaBN(const vP, const adFunc, const avScore, const amHessian);

```

Programa para as aplicações práticas

```

#include <oxstd.h>
#include <oxprob.h>
#import <maximize>
#import <solvenle>
#import "likelihood"

decl t,N,u,v,K,dK,betahat,alphazero,betazero,A,B;

main()
{

    decl data,x,thetahat,ir,vp,dfunc,alpha,beta,rv;
    decl irb, ira_cr, ira_bn, irb_cr, irb_bn;

```

```

data = loadmat("dados1.txt"); //Carregando dados

//Teste (Hipotese nula)
alpha = 0.15;
beta = 125;

//inicializando variaveis
thetahat = zeros(3,2);
rv = zeros(3,2);
t = data';
N = columns(t);
u = sumr(t)/N;
v = (sumr(t.^-1)/N)^-1;

// Calculating the MLE of Beta
x = <200>;
irb = SolveNLE(nle_beta, &x);
betahat = x;
//////////
betazero = betahat;

B = (t.*(betahat.^-2) - (t.^-1));
A = 0.5*( (t*betahat).^-0.5 + 3*(betahat^0.5*t.^-1.5) )
.*( ( (t.^-0.5*betahat.^0.5) + (betahat^1.5*t.^-1.5) ).^-1 );

//Estimando o parametro alpha:
thetahat[0][0] = (u/betahat + betahat/v - 2)^0.5;

K = (meanr((betahat + t).^-1))^2;
dK = -N*sumr((betahat + t).^-2)/(sumr((betahat + t).^-1))^2;
alphazero = thetahat[0][0];

dfunc = 0; //valor inicial da funcao CR
vp = alphazero; //valor inicial do parametro CR
ir = MaxBFGS(fBS_profileAlphaCR, &vp, &dfunc, 0, FALSE); //maximizacao CR
ira_cr = ir; //verificando convergencia
thetahat[1][0] = vp;

dfunc = 0; //valor inicial da funcao BN
vp = alphazero; //valor inicial do parametro BN
ir = MaxBFGS(fBS_profileAlphaBN, &vp, &dfunc, 0, FALSE); //maximizacao BN
ira_bn = ir; //verificando convergencia
thetahat[2][0] = vp;
//fim do procedimento para alpha

// Teste da razao de verossimilhanças perfiladas para alpha
rv[0][0] = 2*(- N*log(thetahat[0][0]) - N*log(betahat) + sumr(log((betahat^0.5)
*(t.^-0.5) + (betahat^1.5)*(t.^-1.5)))
+ (-N/(2*thetahat[0][0]^2))*(u/betahat + betahat/v - 2)
- (- N*log(alpha) - N*log(betahat) + sumr(log((betahat^0.5)*(t.^-0.5)
+ (betahat^1.5)*(t.^-1.5))) + (-N/(2*alpha^2))

```

```

*(u/betahat + betahat/v - 2)));

// Teste da razao de verossimilhanças perfiladas de Cox&Reid para alpha
rv[1][0] = 2*(- N*log(thetahat[1][0]) - N*log(betahat) + sumr(log(((betahat^0.5)
*(t.^-0.5) + (betahat^1.5)*(t.^-1.5)))) + (-N/(2*thetahat[1][0]^2))
*(u/betahat + betahat/v - 2) - 1/2*log((-N/(betahat^2)-N/2*(1
/(betahat^2) + 2*dK/(K^2))-(N/(thetahat[1][0]^2))*(u/(betahat^3)) ))
- (-N*log(alpha) - N*log(betahat) + sumr(log(((betahat^0.5)*(t.^-0.5)
+ (betahat^1.5)*(t.^-1.5)))) + (-N/(2*alpha^2))*(u/betahat + betahat
/v - 2) - 1/2*log((-N/(betahat^2)-N/2*(1/(betahat^2) + 2*dK/(K^2))
-(N/(alpha^2))*(u/(betahat^3))))));

// Teste da razao de verossimilhanças perfiladas de Barndorff-Nielsen para alpha
rv[2][0] = 2*(-N*log(thetahat[2][0]) - N*log(betahat) + sumr(log(((betahat.^0.5)
.*(t.^-0.5))+((betahat^1.5)*(t.^-1.5))))+(-N/(2*(thetahat[2][0]^2)))
*(u/betahat + betahat/v - 2) + 1/2*log(((N/(betahat.^2) + N/2
*(1/(betahat^2)+2*dK/(K.^2)))+(N/(thetahat[2][0]^2))*(u/(betahat^3))))))
- log((N/(betahat.^2) - (2/betahat) * sumr(A) + (1/(thetahat[2][0]^2)
+ 1/(alphazero.^2)).*(0.5*sumr(A.*B) - (0.5*betahat^-1)*sumr(B))
+ sumr(A.^2) + (0.25)*(thetahat[2][0]^2*alphazero^-2)
*sumr(B.^2))) - (-N*log(alpha) - N*log(betahat)
+ sumr(log(((betahat^0.5).*(t.^-0.5)) + ((betahat^1.5)*(t.^-1.5))))
+ (-N/(2*(alpha^2)))*(u/betahat+betahat/v-2)+1/2*log(((N/(betahat^2)
+ N/2*(1/(betahat^2) + 2*dK/(K.^2)))+(N/(alpha^2))*(u/(betahat^3))))))
-log((N/(betahat^2) - (2/betahat) * sumr(A) + (1/(alpha^2)
+ 1/(alphazero.^2)).*(0.5*sumr(A.*B) - (0.5*betahat^-1)*sumr(B))
+ sumr(A.^2) + (0.25)*(alpha^-2*alphazero^-2)*sumr(B.^2)))));

//Estimando o parametro beta:
thetahat[0][1] = betahat;

dfunc = 0;          //valor inicial da funcao CR
vp = betazero;      //valor inicial do parametro CR
ir = MaxBFGS(fBS_profileBetaCR, &vp, &dfunc, 0, FALSE); //maximizacao CR
irb_cr = ir;        //verificando convergencia
thetahat[1][1] = vp;

dfunc = 0;          //valor inicial da funcao BN
vp = betazero;      //valor inicial do parametro BN
ir = MaxBFGS(fBS_profileBetaBN, &vp, &dfunc, 0, FALSE); //maximizacao BN
irb_bn = ir;        //verificando convergencia
thetahat[2][1] = vp;
//fim do procedimento para beta

// Teste da razao de verossimilhanças perfiladas para beta
rv[0][1] = 2*(-N/2*log(u/thetahat[0][1] + thetahat[0][1]/v -2)
- N*log(thetahat[0][1]) + sumr(log(((thetahat[0][1]^0.5)*(t.^-0.5)
+ (thetahat[0][1]^1.5)*(t.^-1.5)))) + N/2*log(u/beta + beta/v - 2)
+ N*log(beta) - sumr(log(((beta^0.5)*(t.^-0.5)
+ (beta^1.5)*(t.^-1.5))))));

// Teste da razao de verossimilhanças perfiladas de Cox&Reid para beta
rv[1][1] = 2*(-N/2*log(u/thetahat[1][1] + thetahat[1][1]/v -2)

```

```

- N*log(thetahat[1][1]) + sumr(log((thetahat[1][1]^0.5)*(t.^-0.5)
+ (thetahat[1][1]^1.5)*(t.^-1.5))) - 1/2*log((2*N)
/(u/thetahat[1][1])+(thetahat[1][1]/v)-2)) + N/2*log(u/beta
+ beta/v - 2) + N*log(beta) - sumr(log((beta^0.5)*(t.^-0.5)
+ (beta^1.5)*(t.^-1.5))) + 1/2*log((2*N)/((u/beta)+(beta/v)-2)) );

// Teste da razao de verossimilhanças perfiladas de Barndorff-Nielsen para beta
rv[2][1] = 2*(-N/2*log(u/thetahat[2][1] + thetahat[2][1]/v -2)
- N*log(thetahat[2][1]) + sumr(log((thetahat[2][1]^0.5)*(t.^-0.5)
+ (thetahat[2][1]^1.5)*(t.^-1.5))) + 1/2*log(((2*N)/((u
/thetahat[2][1])+(thetahat[2][1]/v)-2))) - log(( (N*alphazero)
/(u/thetahat[2][1] + thetahat[2][1]/v -2).^1.5))*(betazero
/thetahat[2][1]*thetahat[2][1]/betazero))) - (-N/2*log(u/beta
+ beta/v -2) - N*log(beta) + sumr(log((beta^0.5)*(t.^-0.5)
+ (beta^1.5)*(t.^-1.5))) + 1/2*log(((2*N)/((u/beta)+(beta/v)-2)))
- log(( (N*alphazero)/((u/beta + beta/v -2).^1.5))*(betazero/beta
*beta/betazero))));

// Imprimindo resultados da aplicacao:
print( "\n\t\t APLICATION RESULTS: BIRNBAUM-SAUNDERS");
print( "\n\t\t -----");
print( "\n\t\t OX PROGRAM: ", oxfilename(0) );
print( "\n\t\t OX VERSION: ", oxversion() );
print( "\n\t\t NUM. OBSERVATIONS: ", N );

print("%r",{ "\t\t ARITHMETIC MEAN: ", "\t\t HARMONIC MEAN: " }, "%10.4f", u);
print("\n\n\t\t *** Estimation of parameters ***\n");
print("%r",{ "\t\t PERFILE: ", "\t\t COX&REID: ", "\t\t BARN-NIL: " },
"%c",{ "\t\t ALPHA", "BETA"}, "%12.6f", thetahat);

print("\n\n\t\t *** Likelihood ratio test ***\n");
print("\t\t Null hypothesis: alpha=", alpha, ", beta=", beta, "\n");
print("%r",{ "\t\t PERFILE: ", "\t\t COX&REID: ", "\t\t BARN-NIL: " },
"%c",{ "\t\t ALPHA", "BETA"}, "%12.6f", rv);

print("\n\n\t\t *** Likelihood ratio test (p-value) ***\n");
print("%r",{ "\t\t PERFILE: ", "\t\t COX&REID: ", "\t\t BARN-NIL: " },
"%c",{ "\t\t ALPHA", "BETA"}, "%12.6f", 1 - probchi(rv,1));

print("\n\n\t\t *** Status of convergence ***\n");
println("\n\t\t Estimativa de beta (eq.-linear): ", MaxConvergenceMsg(irb));
println("\n\t\t Estimativa de alpha_CR: ", MaxConvergenceMsg(ira_cr));
println("\n\t\t Estimativa de alpha_BN: ", MaxConvergenceMsg(ira_bn));
println("\n\t\t Estimativa de beta_CR: ", MaxConvergenceMsg(irb_cr));
println("\n\t\t Estimativa de beta_BN: ", MaxConvergenceMsg(irb_bn));

```

}

Referências

- [1] Barndorff-Nielsen, O.E. (1983). On a formula to the distribution of the maximum likelihood estimator. *Biometrika*, **70**, 343–365.
- [2] Barndorff-Nielsen, O.E. (1994). Adjusted versions of profile likelihood and directed likelihood, and extended likelihood. *Journal of the Royal Statistical Society B*, **56**, 125–140.
- [3] Barndorff-Nielsen, O.E. (1995). Stable and invariant adjusted profile likelihood and directed likelihood for curved exponential models. *Biometrika*, **82**, 489–500.
- [4] Birnbaum, Z.W. e Saunders, S.C. (1969a). A new family of life distribution. *Journal of Applied Probability*, **6**, 319–327.
- [5] Birnbaum, Z.W. e Saunders, S.C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, **6**, 328–347.
- [6] Chang, D.S. e Tang, L.C. (1993). Reliability bounds and critical time for the Birnbaum-Saunders distribution. *IEEE Transaction on Reliability*, **42**, 464–469.
- [7] Chang, D.S. e Tang, L.C. (1994). Percentile bounds and tolerance limits for the Birnbaum-Saunders distribution. *Communications in Statistics – Theory and Methods*, **23**, 159–167.
- [8] Cohen, A.C., Whitten, B.J. e Ding, Y. (1984). Modified moment estimation for the three-parameter Weibull distribution. *Journal of Quality Technology*, **16**, 159–167.
- [9] Cordeiro, G.M. (1992). *Introdução a teoria de verossimilhança*. 10º Simpósio Nacional de Probabilidade e Estatística, UFRJ, Rio de Janeiro.

- [10] Cox, D.R. e Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1–39.
- [11] Cox, D.R. e Reid, N. (1992). A note on the difference between profile and modified profile likelihood. *Biometrika*, **79**, 408–411.
- [12] Cox, D.R. e Reid, N. (1993). A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society B*, **55**, 467–471.
- [13] Cribari-Neto, F. e Zarkos, S.G. (1999). R: yet another econometric programming environment. *Journal of Applied Econometrics*, **14**, 319–329.
- [14] Cribari-Neto, F. e Zarkos, S.G. (2003). Econometric and statistical computing using Ox. *Computational Economics*, **21**, 277–295.
- [15] Cysneiros, A.H.M.A e Ferrari S.L.P. (2006). An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics and Probability Letters*, **76**, 3, 255–265.
- [16] Desmond, A.F. (1985). Stochastic models of failure in random environments. *Canadian Journal Statistics*, **13**, 171–183.
- [17] Desmond, A.F. (1986). On the relationship between two fatigue-life models. *IEEE Transaction on Reliability*, **35**, 167–169.
- [18] Doornik, J.A. e Ooms, M. (2001). *Introduction to Ox*. London: Timberlake Consultants Press.
- [19] Doornik, J. A. (2001). Ox: an Object-oriented Matrix Programming Language, 4th ed. London: Timberlake Consultants & Oxford: <http://www.doornik.com>.
- [20] Duspis, D.J. e Mills, J.E. (1998). Robust estimation of the Birnbaum-Saunders distribution. *IEEE Transaction on Reliability*, **47**, 88–95.

- [21] 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.
- [22] Ferrari, S.L.P. e Cribari-Neto, F. (2002). Corrected modified profile likelihood inference for heteroskedastic models. *Statistics and Probability Letters*, **47**, 353–361.
- [23] Ferrari, S.L.P., Cysneiros, A.H.M.A. e Cribari-Neto, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood-inference. *Journal of Statistical Planning and Inference*, **124**, 423–437.
- [24] Ihaka, R. e Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, **5**, 299–314.
- [25] Johnson, N.L., Kotz, S. e Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Vol. 2, 2nd ed. New York: Wiley.
- [26] Knuth, D. (1986). *The T_EXbook*. New York: Addison-Wesley.
- [27] Lu, M. e Chang, D.S. (1997). Bootstrap prediction intervals for the Birnbaum-Saunders distribution. *Microelectronics Reliability*, **37**, 1213–1216.
- [28] Mann, N.R., Schafer, R.E. e Sigpurwalla, N.D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. New York: Wiley.
- [29] McCool, J.I. (1974). Inferential techniques for Weibull populations. *Aerospace Research Laboratories Report*, ARL TR74-0180. Wright-Patterson Air Force Base, Dayton, OH.
- [30] McCullagh, P. e Tibishirani, R. (1990). A simple method for the adjustment of profile likelihood. *Journal of the Royal Statistical Society B*, **52**, 325–344.

- [31] Ng, H.K.T., Kundu, D. e Balakrishnan, N. (2003). Modified moment estimation for the two parameters Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, **43**, 283–298.
- [32] Nocedal, J. e Wright, S.J. (1999). *Numerical Optimization*. New York: Springer.
- [33] Pace, L. e Savan, A. (1997). *Principles of Statistical Inference from a Neo-Fisherian Perspective*. Singapore: World Scientific.
- [34] Severini, T.A. (1998). An approximation to the modified profile likelihood function. *Biometrika*, **85**, 403–411.
- [35] Severini, T.A. (1999). An empirical adjustment to the likelihood ratio statistic. *Biometrika*, **86**, 235–247.
- [36] Severini, T.A. (2000). *Likelihood Methods in Statistics*. Oxford: Oxford University Press.
- [37] Simonoff, J.S. e Tsai, C.H. (1994). Use of modified profile likelihood for improved tests of constancy of variance in regression. *Applied Statistics*, **43**, 357–370.
- [38] Skovgaard, I.M. (1996). An explicit large-deviation approximation to one-parameter tests. *Bernoulli*, **2**, 145–165.
- [39] Stern, S.E. (1997). A second-order adjustment to the profile likelihood in the case of a multidimensional parameter of interest. *Journal of the Royal Statistical Society B*, **59**, 653–665.
- [40] Wu, J. e Wong, A.C.M. (2004). Improved interval estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, **47**, 809–821.