



Universidade Federal do Rio Grande do Norte  
Centro de Ciências Exatas e da Terra  
Programa de Pós-Graduação em Matemática  
Aplicada e Estatística  
Mestrado em Matemática Aplicada e Estatística



# **Aprimoramento do teste da razão de verossimilhanças em modelos de regressão beta-prime**

Ivonaldo Silvestre da Silva Júnior

Natal – RN

Março de 2024

Ivonaldo Silvestre da Silva Júnior

# **Aprimoramento do teste da razão de verossimilhanças em modelos de regressão beta-prime**

Trabalho apresentado ao Programa de Pós-Graduação em Matemática Aplicada e Estatística da Universidade Federal do Rio Grande do Norte, em cumprimento com as exigências legais para obtenção do título de Mestre.  
Área de Concentração: Estatística.  
Linha de Pesquisa: Modelos de Regressão.

Universidade Federal do Rio Grande do Norte

Centro de Ciências Exatas e da Terra

Programa de Pós-Graduação em Matemática Aplicada e Estatística

Mestrado em Matemática Aplicada e Estatística

Orientador

Prof. Dr. Francisco Moisés Cândido Medeiros

Natal – RN

Março de 2024

Universidade Federal do Rio Grande do Norte - UFRN  
Sistema de Bibliotecas - SISBI  
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Silva Júnior, Ivonaldo Silvestre da.

Aprimoramento do teste da razão de verossimilhanças em modelos de regressão beta-prime / Ivonaldo Silvestre da Silva Júnior. - 2024.

49 f.: il.

Dissertação (mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciência Exatas e da Terra, Programa de Pós-Graduação em Matemática Aplicada e Estatística. Natal, RN, 2024.

Orientação: Prof. Dr. Francisco Moisés Cândido de Medeiros.

1. Estatística matemática - Dissertação. 2. Modelos de regressão beta-prime - Dissertação. 3. Teste da razão de verossimilhanças - Dissertação. 4. Correção de Bartlett - Dissertação. 5. Ajuste de Skovgaard - Dissertação. 6. Correção de Bartlett bootstrap - Dissertação. I. Medeiros, Francisco Moisés Cândido de. II. Título.

RN/UF/CCET

CDU 519.2(043.3)



*Universidade Federal do Rio Grande do Norte*

**PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA E  
ESTATÍSTICA**

**ATA Nº 4**

Aos vinte e seis dias do mês de março do ano de dois mil e vinte e quatro, às 14h00, em ambiente virtual, instalou-se a banca examinadora de dissertação de mestrado intitulada "Aprimoramento do teste da razão de verossimilhanças em modelos de regressão beta-prime." de IVONALDO SILVESTRE DA SILVA JÚNIOR, discente do Programa de Pós-graduação em Matemática Aplicada e Estatística da UFRN. Deu-se início a abertura dos trabalhos, por parte de FRANCISCO MOISES CANDIDO DE MEDEIROS, presidente da banca, que, após apresentar os membros da banca examinadora e esclarecer a tramitação da defesa, marcou um tempo de 50 minutos para a apresentação. Concluída a exposição, a banca lançou seus questionamentos e fez suas considerações sobre o trabalho em julgamento, deliberando pela APROVAÇÃO do trabalho e em favor da concessão do título de MESTRE a IVONALDO SILVESTRE DA SILVA JÚNIOR, conforme as normas vigentes na Universidade Federal do Rio Grande do Norte. A versão final da dissertação deverá ser entregue ao programa, no prazo de 30 dias; contendo as modificações sugeridas pela banca examinadora e constante na folha de correção anexa, conforme o Artigo 79 da Resolução 008/2022 - CONSEPE. O não cumprimento das referidas exigências implica o não recebimento do título de MESTRE.

**Dr. TIAGO MAIA MAGALHÃES, UFJF**

Examinador Externo à Instituição

**Dr. ARTUR JOSE LEMONTE, UFRN**

Examinador Interno

**Dr. FRANCISCO MOISES CANDIDO DE MEDEIROS, UFRN**

Presidente

**IVONALDO SILVESTRE DA SILVA JÚNIOR**

Mestrando

# Agradecimentos

Agradeço aos meus pais, Naldo e Estela, por me trazerem a esse mundo.

Aos meus amigos da UFRN, em especial Beatriz Ariadna, Wesley Cruz e Carla Apolinário, pela longa jornada desde o PET.

Aos meus ex-colegas do TJRN, Luis Ramiro, Gabriel Felipe, Ledivaldo Cabral, Alane Dantas, Antonio Queiroz, Renato Farias, Karine Santos e muitos outros, por toda força que me deram.

Aos meus colegas da Dataprev, em especial Tainá Coimbra, Jefferson Valentim e Willian Baunier. Que nós alcancemos ainda mais nossos objetivos.

Ao professor Moisés Medeiros, por toda a orientação desde 2017. Espero que continuemos mantendo boas conexões mesmo após a conclusão do mestrado.

Aos professores que fizeram parte da banca de qualificação e defesa, ainda como suplentes: Moisés Medeiros, Artur Lemonte, Antonio Hermes, Eliardo Guimarães, Thiago Magalhães, Mariana Araújo e Fábio Bayer. Agradeço pela aprovação e pelas correções.

Às professoras Luz Milena e Mariana Araújo, do DEST das UFRN, por todo o apoio.

Aos cantores e compositores que sempre trazem uma grande calma para minha alma.

Ao *Nós 12*, por nunca saírem da minha cabeça.

A todos aqueles que, embora eu não tenha pensado enquanto escrevo esses agradecimentos, contribuíram positivamente para o meu crescimento.

*“Got fears, but I face them.”*  
*(K/DA – I’ll Show You)*

# Resumo

Nesta dissertação, abordamos o problema de testar hipóteses em modelos de regressão beta-prime em amostras de tamanho pequeno e moderado. Focamos no teste de razão de verossimilhanças, o qual pode não ser confiável quando o tamanho da amostra não é grande o suficiente para garantir uma boa aproximação entre a distribuição exata da estatística de teste e a correspondente distribuição assintótica qui-quadrado. As correções de Bartlett, Skovgaard e Bartlett-*bootstrap* tradicionalmente atenuam a distorção do tamanho do teste. Neste trabalho, derivamos a correção de Bartlett e um ajuste de Skovgaard para o teste de razão de verossimilhanças em modelos de regressão beta-prime. Comparamos numericamente os testes usual, corrigidos e *bootstrap* por meio de simulações de Monte Carlo. Nossos resultados sugerem que os testes corrigidos e o teste *bootstrap* exibem uma probabilidade de erro do tipo I mais próxima do nível nominal estabelecido. Os testes corrigidos analiticamente apresentam a vantagem de não exigir cálculos computacionalmente intensos. Apresentamos uma aplicação a dados reais para ilustrar a utilidade dos testes modificados.

**Palavras-chave:** Modelos de regressão beta-prime. Teste da razão de verossimilhanças. Correção de Bartlett. Ajuste de Skovgaard. Correção de Bartlett bootstrap.

# Abstract

This paper deals with the issue of testing hypotheses in beta-prime regression models in small and moderate-sized samples. We focus on the likelihood ratio test, which is unreliable when the sample size is not large enough to guarantee a good approximation between the exact distribution of the test statistic and the corresponding chi-squared asymptotic distribution. Bartlett, Skovgaard and Bartlett-bootstrap corrections typically attenuate the size distortion of the test. Here, we derive a Bartlett correction and a Skovgaard adjustment for the likelihood ratio test on beta-prime regression models. We numerically compare the usual, corrected and bootstrapped tests, through simulations. Our results suggest that the corrected and bootstrapped tests exhibit type I probability error closer to the chosen nominal level. The analytically corrected tests perform with the advantage of not requiring computationally intensive calculations. We present a real data application to illustrate the usefulness of the modified tests.

**Keywords:** Beta-prime regression models. Likelihood ratio test. Bartlett correction. Skovgaard adjustment. Bartlett bootstrap correction.

# Lista de Figuras

Figura 1 – Função densidade de probabilidade e função de distribuição acumulada da distribuição beta-prime. . . . .	14
Figura 2 – Discrepâncias relativas por quantis assintóticos para $n = 30$ , $\phi = 10$ , $p = 4$ e $q = 2$ . . . . .	27
Figura 3 – Discrepâncias relativas por quantis assintóticos para $n = 30$ , $\phi = 10$ , $p = 5$ e $q = 3$ . . . . .	27
Figura 4 – Boxplot do valor das estatísticas de teste para $\phi = 1$ , $p = 5$ e $q = 2$ . . .	28
Figura 5 – <i>Wormplots</i> dos modelos ajustados . . . . .	31

# Lista de Tabelas

Tabela 1 – Exemplos de função de ligação. . . . .	15
Tabela 2 – Taxa de rejeição (%) da hipótese nula para $\mathcal{H}_0 : \beta_1 = 0$ com $p = 3$ . . .	24
Tabela 3 – Taxa de rejeição (%) da hipótese nula para $\mathcal{H}_0 : \beta_1 = \beta_2 = 0$ com $p = 4$ .	25
Tabela 4 – Taxa de rejeição (%) da hipótese nula para $\mathcal{H}_0 : \beta_1 = \beta_2 = \beta_3 = 0$ com $p = 5$ . . . . .	26
Tabela 5 – Taxa de rejeição (%) da hipótese nula para $\mathcal{H}_1 : \beta_1 = \beta_2 = \delta$ . . . . .	30
Tabela 6 – Estimativas de máxima verossimilhança (EMV), desvios-padrão (DP), estatística da razão de verossimilhanças (LR) e $p$ -valor do modelo ajustado.	32
Tabela 7 – Estatísticas de teste e $p$ -valores para testar $\mathcal{H}_0 : \beta_1 = 0$ contra $\mathcal{H}_0 : \beta_1 \neq 0$ .	32

# Sumário

1	INTRODUÇÃO . . . . .	11
2	MODELOS DE REGRESSÃO BETA-PRIME . . . . .	13
2.1	Distribuição beta-prime . . . . .	13
2.2	Modelos de regressão beta-prime . . . . .	14
3	TESTE DA RAZÃO DE VEROSSIMILHANÇAS . . . . .	17
3.1	Correção de Bartlett . . . . .	18
3.2	Ajuste de Skovgaard . . . . .	19
3.3	Correção de Bartlett- <i>bootstrap</i> . . . . .	20
4	SIMULAÇÕES . . . . .	22
4.1	Tamanho do teste . . . . .	22
4.2	Poder do teste . . . . .	29
5	APLICAÇÃO . . . . .	31
6	CONSIDERAÇÕES FINAIS . . . . .	33
	REFERÊNCIAS . . . . .	34
	APÊNDICES . . . . .	37
	APÊNDICE A – VETOR ESCORE, MATRIZ HESSIANA E MATRIZ DE INFORMAÇÃO DE FISHER . . . . .	38
	APÊNDICE B – CUMULANTES . . . . .	40
	APÊNDICE C – AJUSTE DE SKOVGAARD . . . . .	46

# 1 Introdução

Em modelos de regressão, geralmente é de interesse testar hipóteses sobre os parâmetros do modelo. Considere o problema de testar a hipótese de que um sub-vetor de parâmetros da regressão é igual a um vetor de valores constantes ( $\mathcal{H}_0$ ; hipótese nula) contra a hipótese de que esse sub-vetor é diferente do vetor de valores constantes ( $\mathcal{H}_1$ ; hipótese alternativa). Uma forma de testar essas hipóteses é considerar o teste da razão de verossimilhanças (Wilks, 1938), que, sob a hipótese nula, tem distribuição assintótica qui-quadrado. Entretanto, em amostras de tamanho pequeno ou moderado, essa aproximação qui-quadrado para a distribuição da estatística da razão de verossimilhanças pode apresentar resultados distorcidos, como, por exemplo, no tamanho do teste.

A fim de corrigir possíveis distorções no teste da razão de verossimilhanças, podemos aplicar a correção de Bartlett (Bartlett; Lawley, 1937, 1956), cuja fórmula é composta de cumulantes de até quarta ordem do logaritmo da função de verossimilhança. Essa correção foi amplamente estudada na literatura em diversos modelos. Por exemplo: Medeiros, Ferrari e Lemonte (2017) analisaram, dentre outros tipos de correção, a correção de Bartlett em modelos de dispersão; Medeiros e Lemonte (2021) compararam as correções de Bartlett com outros tipos de correção em modelos exponenciais na presença de censura; Melo et al. (2022) desenvolveram a correção de Bartlett para o modelo de regressão Dirichlet multivariado; Araújo, Cysneiros e Cysneiros (2022) derivaram a correção de Bartlett em modelos lineares simétricos heteroscedásticos; e Carneiro et al. (2023) desenvolveram uma correção para os modelos de regressão Kumaraswamy.

Outra forma de corrigir as distorções é a partir da aproximação do ajuste de Skovgaard (Skovgaard, 2001), cuja fórmula é menos complicada do que a correção de Bartlett, já que utiliza cumulantes de até segunda ordem. Essa correção também já foi amplamente estudada na literatura. Ferrari e Pinheiro (2011) e Bayer e Cribari-Neto (2013) derivaram o ajuste de Skovgaard para o modelo de regressão beta; Perette (2019) implementaram computacionalmente o ajuste para os modelos elípticos; Rauber, Cribari-Neto e Bayer (2020) derivaram o ajuste para o modelo de regressão beta em que a função de ligação para o parâmetro da média depende de um parâmetro adicional; Guedes, Cribari-Neto e Espinheira (2020) desenvolveram, dentre outras correções, a correção de Skovgaard para o modelo gama; e Melo et al. (2022) desenvolveram uma correção de Skovgaard para os modelos Dirichlet.

Uma terceira alternativa para reduzir as distorções é a partir da correção de Bartlett-*bootstrap* (Rocke, 1989). Esse método de correção é computacionalmente mais intensivo que os demais, porém é tão preciso quanto a correção de Bartlett e de Skovgaard,

o que pode ser mais prático quando o cálculo dos cumulantes é muito complexo, uma vez que não necessita dessas quantidades. Essa correção também já foi amplamente considerada na literatura (ver Loose, Bayer e Pereira (2017), Loose, Valença e Bayer (2018), Araújo, Cysneiros e Montenegro (2020), Canepa (2020) e Carneiro et al. (2023)).

No entanto, não encontramos na literatura correções para o teste da razão de verossimilhanças no modelo de regressão beta-prime (Bourguignon, Santos-Neto e Castro, 2018), em que a variável resposta segue uma distribuição beta-prime (Keeping; McDonald, 1962, 1984), uma distribuição com suporte positivo que pode ser útil no contexto de análise de sobrevivência, como uma alternativa a modelos tradicionais, como o modelo gama ou Weibull (ver Kosmidis e Firth (2019), Korkmaz e Gökgoz (2020), Zhang, Wang e Peng (2021)). Portanto, no nosso trabalho, desenvolvemos as correções de Bartlett, de Skovgaard e de Bartlett-*bootstrap* nos modelos de regressão beta-prime e realizamos simulações de Monte Carlo para compararmos numericamente a probabilidade do erro de tipo I (tamanho do teste) e o poder do teste da razão de verossimilhanças com os testes corrigidos. Este trabalho será útil para obter conclusões mais precisas quando a distribuição da variável resposta é beta-prime e não é possível aumentar o tamanho da amostra, como, por exemplo, por falta de recursos ou de tempo.

No Capítulo 2, apresentamos a distribuição beta-prime e os modelos de regressão beta-prime. No Capítulo 3, apresentamos o teste da razão de verossimilhanças e desenvolvemos as correções de Bartlett, Skovgaard e Bartlett-*bootstrap*. No Capítulo 4, apresentamos as simulações de Monte Carlo para avaliar o desempenho de cada teste e, por fim, no Capítulo 5, consideramos uma aplicação a um conjunto de dados reais.

## 2 Modelos de regressão beta-prime

### 2.1 Distribuição beta-prime

Seja  $X$  uma variável aleatória que segue uma distribuição beta-prime (Keeping; McDonald, 1962, 1984) de parâmetros  $\alpha$  e  $\beta$ , denotada por  $BP(\alpha, \beta)$ . A sua função de densidade de probabilidade é dada por

$$f_X(x) = \frac{x^{\alpha-1}(1+x)^{-\alpha-\beta}}{B(\alpha, \beta)}, \quad x > 0,$$

em que  $\alpha$  e  $\beta$  são valores positivos que representam os parâmetros de forma e  $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$  representa a função beta. O valor esperado e a variância de  $X$  são dados, respectivamente, por:

$$E(X) = \frac{\alpha}{\beta - 1}, \quad \text{para } \beta > 1 \text{ e } Var(X) = \frac{\alpha(\alpha + \beta - 1)}{(\beta - 2)(\beta - 1)^2}, \quad \text{para } \beta > 2.$$

Essa distribuição pode ser obtida pela chance de uma distribuição beta (Johnson, Kotz e Balakrishnan, 1995), isto é, se  $Z \sim \text{Beta}(\alpha, \beta)$ , então

$$X = \frac{Z}{1-Z} \sim BP(\alpha, \beta).$$

Bourguignon, Santos-Neto e Castro (2018) propuseram uma reparametrização da distribuição beta-prime em função da média, com função densidade de probabilidade dada por:

$$f_Y(y) = \frac{y^{\mu(\phi+1)-1}(1+y)^{-[\mu(\phi+1)+\phi+2]}}{B(\mu(\phi+1), \phi+2)}, \quad y > 0,$$

em que os parâmetros  $\mu > 0$  e  $\phi > 0$  representam, respectivamente, a média e a precisão, o que torna essa reparametrização mais interpretável em termos dos parâmetros. Se  $Y \sim BP(\mu, \phi)$ , então

$$E(Y) = \mu \quad \text{e} \quad Var(Y) = \frac{\mu(1+\mu)}{\phi}.$$

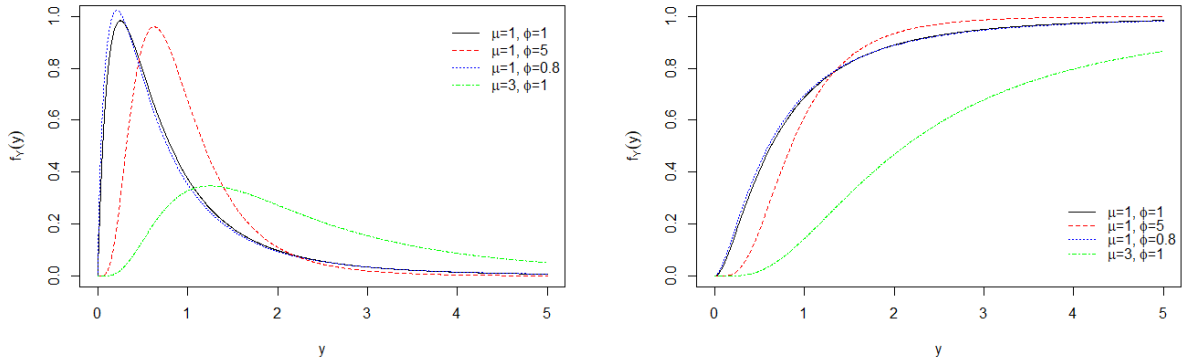
Além disso, a sua função de distribuição acumulada igual a

$$F_Y(y) = I_{\frac{y}{1+y}}(\mu(\phi+1), \phi+2) = \frac{1}{B(\mu(\phi+1), \phi+2)} \int_0^{\frac{y}{1+y}} t^{\mu(\phi+1)}(1-t)^{\phi+2} dt$$

É possível associar os parâmetros de  $Y \sim BP(\mu, \phi)$  com os parâmetros de  $X \sim BP(\alpha, \beta)$  a partir das seguintes relações

$$\mu = \frac{\alpha}{\beta - 1} \quad \text{e} \quad \phi = \beta - 2.$$

Na Figura 1 apresentamos os gráficos das funções densidade de probabilidade e funções de distribuição acumulada da distribuição beta-prime para diferentes valores de  $\mu$  e  $\phi$ . É possível perceber que, em geral, a distribuição beta-prime é assimétrica à esquerda, porém, quando  $\phi \rightarrow \infty$ , a distribuição tende a ficar simétrica. Além disso, maiores valores de  $\mu$  podem afetar a curtose, deixando-a mais platicúrtica.



(a) Função densidade de probabilidade.

(b) Função de distribuição acumulada.

Figura 1 – Função densidade de probabilidade e função de distribuição acumulada da distribuição beta-prime.

## 2.2 Modelos de regressão beta-prime

Seja  $\mathbf{y} = (y_1, \dots, y_n)^\top$  um vetor de dimensão  $n$  formado por variáveis aleatórias independentes,  $y_i \sim BP(\mu_i, \phi)$ , para  $i = 1, \dots, n$ , e  $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip-1})^\top$  um vetor de dimensão  $p$  ( $p < n$ ), de variáveis explicativas relacionadas ao  $i$ -ésimo indivíduo e  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  uma matriz de dimensão  $n \times p$ , de  $\text{posto}(\mathbf{X}) = p$ . Podemos relacionar a média das observações com as variáveis explicativas a partir da seguinte relação

$$g_1(\mu_i) = \mathbf{x}_i^\top \boldsymbol{\beta} = \eta_{1i}, \quad (2.1)$$

em que  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^\top$  é um vetor de parâmetros desconhecidos de dimensão  $p$  e  $g_1(\cdot)$  é uma função estritamente monótona, contínua, inversível e diferenciável pelo menos duas vezes, denominada de função de ligação. Nesse caso, essa função possui domínio positivo e imagem real. Alguns exemplos de funções de ligação são apresentados na Tabela 1. Além disso, consideramos uma função de ligação para o parâmetro  $\phi$ , isto é,  $g_2(\phi) = \eta_2$ .

Tabela 1 – Exemplos de função de ligação.

Função	Fórmula	Inversa
Identidade	$\mu_i = \eta_{1i}$	$\mu_i = \eta_{1i}$
Logarítmica	$\log(\mu_i) = \eta_{1i}$	$\mu_i = \exp(\eta_{1i})$
Raiz quadrada	$\sqrt{\mu_i} = \eta_{1i}$	$\mu_i = \eta_{1i}^2$

O logaritmo da função de verossimilhança é dado por:

$$\ell(\boldsymbol{\theta}; \mathbf{y}) = \sum_{i=1}^n \ell(\mu_i, \phi; y_i), \quad (2.2)$$

em que  $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi)^\top$ , é um vetor de dimensão  $k$  com  $k = p + 1$  e

$$\begin{aligned} \ell(\mu_i, \phi; y_i) = & [\mu_i(1 + \phi) - 1] \log y_i - [\mu_i(1 + \phi) + \phi + 2] \log(1 + y_i) - \log[\Gamma(\mu_i(1 + \phi))] \\ & - \log[\Gamma(\phi + 2)] - \log[\Gamma(\mu_i(1 + \phi) + \phi + 2)]. \end{aligned}$$

O vetor escore de  $\boldsymbol{\theta}$  é dada por  $\mathbf{U}(\boldsymbol{\theta}) = (\mathbf{U}_\beta(\boldsymbol{\theta})^\top, U_\phi(\boldsymbol{\theta}))^\top$ , com

$$\mathbf{U}_\beta(\boldsymbol{\theta}) = (1 + \phi) \mathbf{X}^\top \mathbf{D}_1 (\mathbf{y}^* - \boldsymbol{\mu}^*) \quad \text{e} \quad U_\phi(\boldsymbol{\theta}) = \mathbf{1}^\top (\mathbf{y}^* - \boldsymbol{\mu}^*) \frac{\partial \phi}{\partial \eta_2},$$

em que,

$$\begin{aligned} y_i^* &= \log \left( \frac{y_i}{1 + y_i} \right), & \mu_i^* &= \psi^{(0)}(\mu_i(1 + \phi)) - \psi^{(0)}(\mu_i(1 + \phi) + \phi + 2), \\ y_i^* &= \log \left( \frac{y_i^{\mu_i}}{(1 + y_i)^{1 + \mu_i}} \right), & \mu_i^* &= \mu_i \left( \mu_i^* + \frac{1}{\mu_i} \left( \psi^{(0)}(\phi + 2) - \right. \right. \\ \mathbf{D}_1 &= \text{diag} \left( \frac{\partial \mu_1}{\partial \eta_{11}}, \dots, \frac{\partial \mu_n}{\partial \eta_{1n}} \right), & & \left. \left. \psi^{(0)}(\mu_i(1 + \phi) + \phi + 2) \right) \right), \\ \mathbf{y}^* &= (y_1^*, \dots, y_n^*)^\top, & \boldsymbol{\mu}^* &= (\mu_1^*, \dots, \mu_n^*)^\top, \\ \mathbf{y}^* &= (y_1^*, \dots, y_n^*)^\top, & \boldsymbol{\mu}^* &= (\mu_1^*, \dots, \mu_n^*)^\top. \end{aligned}$$

Mais detalhes sobre essas quantidades estão presentes no Apêndice A e casos mais gerais podem ser encontrados em Bourguignon, Santos-Neto e Castro (2018).

Os estimadores de máxima verossimilhança (EMV) de  $\boldsymbol{\beta}$  e  $\phi$ , denotados por  $\hat{\boldsymbol{\beta}}$  e  $\hat{\phi}$ , são encontrados a partir da solução simultânea das equações  $\mathbf{U}_\beta(\boldsymbol{\theta}) = 0$  e  $U_\phi(\boldsymbol{\theta}) = 0$ . Como esses estimadores não possuem uma forma analítica fechada, uma maneira de determiná-los é partir de métodos numéricos, como os algoritmos de otimização não linear de Newton

ou quase-Newton (ver [Press et al. \(1992\)](#)). Para calculá-los computacionalmente, pode-se utilizar o pacote `BPMODEL` ([Santos-Neto et al., 2022](#)), implementado no *software* R ([R Core Team, 2021](#)), um pacote que, além de conter funções para obtermos os quantis, as probabilidades, as funções densidade e números aleatórios da distribuição beta-prime, também contém funções para ajustarmos os modelos de regressão beta-prime na mesma sintaxe do pacote `GAMLSS` ([Rigby e Stasinopoulos, 2005](#)).

As matrizes Hessiana e de informação de Fisher são dadas, respectivamente, por

$$\mathbf{J} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}_3 \mathbf{X} & \mathbf{X}^\top \mathbf{D}_5 \mathbf{1} \\ \mathbf{1}^\top \mathbf{D}_5 \mathbf{X} & \mathbf{1}^\top \mathbf{D}_4 \mathbf{1} \end{bmatrix} \text{ e } \mathbf{K} = \begin{bmatrix} \mathbf{X}^\top \mathbf{E}_3 \mathbf{X} & \mathbf{X}^\top \mathbf{E}_5 \mathbf{1} \\ \mathbf{1}^\top \mathbf{E}_5 \mathbf{X} & \mathbf{1}^\top \mathbf{E}_4 \mathbf{1} \end{bmatrix}.$$

Mais detalhes sobre essas matrizes podem ser encontrados no Apêndice [A](#). Sob condições de regularidade ([Cox e Hinkley, 1979](#)), tem-se que

$$\sqrt{n} \begin{pmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\phi} - \phi \end{pmatrix} \xrightarrow{D} N_k(\mathbf{0}, \mathbf{K}^{-1}),$$

em que  $\xrightarrow{D}$  representa a convergência em distribuição e  $\mathbf{K}^{-1}$  representa a inversa da matriz de informação de Fisher.

### 3 Teste da razão de verossimilhanças

Considere o modelo paramétrico descrito em (2.1) e o logaritmo da função de verossimilhança apresentado em (2.2). O vetor de parâmetros desconhecidos  $\boldsymbol{\theta}$  pode ser particionado em  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ , em que  $\boldsymbol{\theta}_1$  é um vetor de dimensão  $q$ , que representa os parâmetros de interesse, e  $\boldsymbol{\theta}_2$  é um vetor de dimensão  $(k - q)$ , que representa os parâmetros de perturbação. Para testar a hipótese nula

$$\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_0,$$

contra a hipótese alternativa

$$\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_0,$$

sendo  $\boldsymbol{\theta}_0$  um vetor de dimensão  $q$  contendo valores conhecidos, podemos utilizar a estatística da razão de verossimilhanças (Wilks, 1938), definida por

$$\text{LR} = 2 \left[ \ell(\hat{\boldsymbol{\theta}}; \mathbf{y}) - \ell(\tilde{\boldsymbol{\theta}}; \mathbf{y}) \right],$$

em que  $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_0^\top, \tilde{\boldsymbol{\theta}}_2^\top)^\top$  e  $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_1^\top, \hat{\boldsymbol{\theta}}_2^\top)^\top$  representam, respectivamente, os EMV de  $\boldsymbol{\theta}$  sob a hipótese nula e sob a hipótese alternativa. Sob a hipótese nula, quando  $n \rightarrow \infty$ , a distribuição da estatística LR é aproximadamente  $\chi_q^2$  com erro de aproximação  $\mathcal{O}(n^{-1})$ . Logo, rejeitamos  $\mathcal{H}_0$ , a um nível de significância  $\alpha$  fixado, se o valor observado de LR for maior que  $\chi_{q;1-\alpha}^2$ , em que  $\chi_{q;1-\alpha}^2$  é tal que  $P(\chi_q^2 < \chi_{q;1-\alpha}^2) = 1 - \alpha$ .

Entretanto, em amostras de tamanho pequeno ou moderado, essa aproximação pode apresentar distorções no tamanho do teste. Para corrigir essas possíveis distorções, podemos utilizar algumas correções, como, por exemplo, a correção de Bartlett (Bartlett; Lawley, 1937, 1956), a aproximação de Skovgaard (Skovgaard, 2001) ou a correção de Bartlett-bootstrap (Rocke; Efron, 1989, 1979). A distribuição assintótica de todas as estatísticas corrigidas é  $\chi_q^2$ , com erro de aproximação de ordem  $\mathcal{O}(n^{-2})$  e o critério de rejeição da hipótese nula é análogo ao teste usual, isto é, rejeitamos  $\mathcal{H}_0$  se o valor da estatística corrigida for maior do que o quantil da distribuição qui-quadrado para um nível de significância fixado.

### 3.1 Correção de Bartlett

A correção de Bartlett, proposta por [Bartlett \(1937\)](#) e generalizada por [Lawley \(1956\)](#), tem uma fórmula polinomial, em que a estatística  $LR$  é multiplicada por um fator, denominado de fator de correção de Bartlett, tal que o teste corrigido possui um nível de distorção no tamanho do teste menor do que o teste usual ([Bartlett](#); [Rocke, 1937, 1989](#)). Três possíveis versões da estatística da razão de verossimilhanças corrigidas por Bartlett são

$$\begin{aligned} LR_{b1} &= \frac{LR}{1+c}, \\ LR_{b2} &= LR \exp(-c), \\ LR_{b3} &= LR(1-c), \end{aligned}$$

em que  $c = \frac{\varepsilon_k - \varepsilon_{k-q}}{q}$  e

$$\begin{aligned} \varepsilon_k &= \sum_{\boldsymbol{\theta}} (\lambda_{rstu} - \lambda_{rstuvw}), \\ \lambda_{rstu} &= \kappa^{rs} \kappa^{tu} \left( \frac{\kappa_{rstu}}{4} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right), \\ \lambda_{rstuvw} &= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[ \kappa_{rtv} \left( \frac{\kappa_{suw}}{6} - \kappa_{sw}^{(u)} \right) + \kappa_{rtu} \left( \frac{\kappa_{svw}}{4} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right]. \end{aligned}$$

Os índices  $r, s, t, u, v$  e  $w$  percorrem os índices do vetor  $\boldsymbol{\theta}$ . Os  $\kappa$ 's são chamados de cumulantes e são definidos por  $\kappa_{rs} = E(\partial^2 \ell(\boldsymbol{\theta}) / (\partial \theta_r \partial \theta_s))$ ,  $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \theta_t$ ,  $\kappa_{rst} = E(\partial^3 \ell(\boldsymbol{\theta}) / (\partial \theta_r \partial \theta_s \partial \theta_t))$ ,  $\kappa_{rs}^{(tu)} = \partial^2 \kappa_{rs} / (\partial \theta_t \partial \theta_u)$ ,  $\kappa_{rst}^{(u)} = \partial \kappa_{rst} / \partial \theta_u$ ,  $\kappa_{rstu} = E(\partial^4 \ell(\boldsymbol{\theta}) / (\partial \theta_r \partial \theta_s \partial \theta_t \partial \theta_u))$  e assim por diante. Além disso,  $\kappa^{rs}$  representa o elemento de ordem  $(r, s)$  da inversa da matriz de informação de Fisher. O mesmo cálculo deve ser utilizado para encontrar  $\varepsilon_{k-q}$  dentre os índices do vetor  $\boldsymbol{\theta}_2$ .

Obter o fator de correção de Bartlett a partir desses somatórios pode ser muito trabalhoso, já que a matriz de informação de Fisher não é bloco diagonal. Com o objetivo de simplificar e reduzir o tempo de cálculo, podemos utilizar as expressões matriciais propostas por [Cordeiro \(1993\)](#).

Defina as matrizes  $A^{(tu)}$ ,  $P^{(t)}$  e  $Q^{(u)}$  de ordem  $k \times k$  para  $t, u = 1, \dots, k$ . Os elementos de ordem  $(r, s)$  dessas matrizes são definidos, respectivamente, por

$$\begin{aligned} A_{rs}^{(tu)} &= \frac{\kappa_{rstu}}{4} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)}, \\ P_{rs}^{(t)} &= \kappa_{rst}, \\ Q_{rs}^{(u)} &= \kappa_{su}^{(r)}. \end{aligned}$$

Além disso, defina as matrizes  $L$ ,  $M_1$ ,  $M_2$ ,  $M_3$ ,  $N_1$ ,  $N_2$  e  $N_3$ , também de ordem  $k \times k$ , em que seus elementos de ordem  $(r, s)$  são dados por

$$\begin{aligned} L_{rs} &= \text{tr} \left( K^{-1} A^{(rs)} \right), \\ M_{1(rs)} &= \text{tr} \left( K^{-1} P^{(r)} K^{-1} P^{(s)} \right), \\ M_{2(rs)} &= \text{tr} \left( K^{-1} P^{(r)} K^{-1} Q^{(s)\top} \right), \\ M_{3(rs)} &= \text{tr} \left( K^{-1} Q^{(r)} K^{-1} Q^{(s)} \right), \\ N_{1(rs)} &= \text{tr} \left( K^{-1} P^{(r)} \right) \text{tr} \left( K^{-1} P^{(s)} \right), \\ N_{2(rs)} &= \text{tr} \left( K^{-1} P^{(r)} \right) \text{tr} \left( K^{-1} Q^{(s)} \right), \\ N_{3(rs)} &= \text{tr} \left( K^{-1} Q^{(r)} \right) \text{tr} \left( K^{-1} Q^{(s)} \right). \end{aligned}$$

em que  $\text{tr}(\cdot)$  é o operador traço. Dessa forma,  $\varepsilon_k$  pode ser escrito na forma matricial como

$$\varepsilon_k = \text{tr}[K^{-1}(L - M - N)],$$

sendo  $M = -\frac{1}{6}M_1 + M_2 - M_3$  e  $N = -\frac{1}{4}N_1 + N_2 - N_3$ . Nesse trabalho, após uma álgebra exaustiva, conseguimos encontrar as formas fechadas dos cumulantes e estão presentes no Apêndice B.

Um *script* para gerar a correção de Bartlett do modelo de regressão beta na linguagem de programação **R** ([R Core Team, 2021](#)) é apresentado em [Bayer \(2011\)](#), o que serviu de ponto inicial para a implementação computacional do fator de correção de Bartlett nos modelos de regressão beta-prime.

## 3.2 Ajuste de Skovgaard

Outra maneira de corrigir a estatística da razão de verossimilhanças é encontrada em [Skovgaard \(2001\)](#), denominada de ajuste de Skovgaard. Essa correção surge a partir do teste de sinais modificado de Barndorff-Nielsen ([Barndorff-Nielsen; Barndorff-Nielsen, 1986, 1991](#)) e pode ser difícil de ser calculada numericamente. Entretanto, sua aproximação possui uma fórmula mais simples que a correção de Bartlett, pois só necessitamos de cumulantes até segunda ordem e, ainda assim, pode apresentar resultados semelhantes.

Há duas maneiras equivalentes de corrigir a estatística LR por meio do ajuste de Skovgaard, são elas

$$\text{LR}_{sk1} = \text{LR} - 2 \log \xi \quad \text{e} \quad \text{LR}_{sk2} = \text{LR} \left( 1 - \frac{\log \xi}{\text{LR}} \right)^2,$$

a quantidade  $\xi$  é tal que

$$\xi = \frac{\left[ |\widehat{\mathbf{K}}| |\widehat{\mathbf{K}}| |\widehat{\mathbf{J}}_{11}| \right]^{1/2}}{|\widehat{\mathbf{Y}}| \left| \left\{ \widehat{\mathbf{K}} \widehat{\mathbf{Y}}^{-1} \widehat{\mathbf{J}} \widehat{\mathbf{K}}^{-1} \widehat{\mathbf{Y}} \right\}_{11} \right|^{1/2}} \frac{\left[ \widehat{\mathbf{U}}^\top \widehat{\mathbf{Y}}^{-1} \widehat{\mathbf{K}} \widehat{\mathbf{J}}^{-1} \widehat{\mathbf{Y}} \widehat{\mathbf{K}}^{-1} \widehat{\mathbf{U}} \right]^{q/2}}{LR^{q/2-1} \widehat{\mathbf{U}}^\top \widehat{\mathbf{Y}}^{-1} \widehat{\mathbf{v}}},$$

em que  $\widehat{\mathbf{J}} = \mathbf{J}(\widehat{\boldsymbol{\theta}})$ ,  $\widetilde{\mathbf{J}} = \mathbf{J}(\widetilde{\boldsymbol{\theta}})$ ,  $\widehat{\mathbf{K}} = \mathbf{K}(\widehat{\boldsymbol{\theta}})$  e  $\widetilde{\mathbf{K}} = \mathbf{K}(\widetilde{\boldsymbol{\theta}})$  e a notação  $\cdot_{11}$  representa uma sub-matriz cujos índices são os índices dos parâmetros de interesse. As quantidades  $\widehat{\mathbf{Y}}$  e  $\widehat{\mathbf{v}}$  são dadas por  $\mathbf{Y} = E_{\boldsymbol{\theta}}[\mathbf{U}(\boldsymbol{\theta})\mathbf{U}^\top(\boldsymbol{\theta}_2)]$  e  $\mathbf{v} = E_{\boldsymbol{\theta}}[\mathbf{U}(\boldsymbol{\theta})(\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2))]$ , substituindo os valores de  $\boldsymbol{\theta}$  e  $\boldsymbol{\theta}_2$  por  $\widehat{\boldsymbol{\theta}}$  e  $\widetilde{\boldsymbol{\theta}}$ , respectivamente. Após uma álgebra tediosa, obtivemos essas quantidades nesse trabalho para os modelos de regressão beta-prime

$$\widehat{\mathbf{Y}} = \begin{bmatrix} (1 + \widehat{\phi})(1 + \widetilde{\phi}) \mathbf{X}^\top \widehat{\mathbf{D}}_1 \mathbf{V}_1 \widetilde{\mathbf{D}}_1 \mathbf{X} & (1 + \widehat{\phi}) \mathbf{X}^\top \widehat{\mathbf{D}}_1 (\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) \mathbf{1} \widetilde{b}_{(1)} \\ (1 + \widetilde{\phi}) \mathbf{1}^\top (\mathbf{V}_1 \widetilde{\mathbf{M}} + \mathbf{C}) \widetilde{\mathbf{D}}_1 \mathbf{X} \widehat{b}_{(1)} & \mathbf{1}^\top [\mathbf{V}_1 \widetilde{\mathbf{M}} \widetilde{\mathbf{M}} + \mathbf{C}(\widehat{\mathbf{M}} + \widetilde{\mathbf{M}}) + \mathbf{V}_2] \mathbf{1} \widehat{b}_{(1)} \widetilde{b}_{(1)} \end{bmatrix}$$

$$\widehat{\mathbf{v}} = \begin{bmatrix} (1 + \widehat{\phi}) \mathbf{X}^\top \widehat{\mathbf{D}}_1 [\mathbf{V}_1 ((1 + \widehat{\phi}) \widehat{\mathbf{M}} - (1 + \widetilde{\phi}) \widetilde{\mathbf{M}}) + (\widehat{\phi} - \widetilde{\phi}) \mathbf{C}] \mathbf{1} \\ \mathbf{1}^\top [(\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) ((1 + \widehat{\phi}) \widehat{\mathbf{M}} - (1 + \widetilde{\phi}) \widetilde{\mathbf{M}}) + (\widehat{\phi} - \widetilde{\phi}) (\mathbf{C} \widehat{\mathbf{M}} + \mathbf{V}_2 - \mathbf{C})] \mathbf{1} \widehat{b}_{(1)} \end{bmatrix},$$

em que  $\widehat{\mathbf{D}}_1 = \text{diag}(\frac{\partial \widehat{\mu}_1}{\partial \widehat{\eta}_1}, \dots, \frac{\partial \widehat{\mu}_n}{\partial \widehat{\eta}_n})$ ,  $\widetilde{\mathbf{D}}_1 = \text{diag}(\frac{\partial \widetilde{\mu}_1}{\partial \widetilde{\eta}_1}, \dots, \frac{\partial \widetilde{\mu}_n}{\partial \widetilde{\eta}_n})$ ,  $\mathbf{V}_1 = -\text{diag}(\widehat{d}_1^{(0,2)}, \dots, \widehat{d}_n^{(0,2)})$ ,  $\mathbf{V}_2 = -\text{diag}(\widehat{r}_1^{(0,2)}, \dots, \widehat{r}_n^{(0,2)}) - \mathbf{V}_1 + \mathbf{C}$ ,  $\mathbf{C} = -\text{diag}(\widehat{t}_1^{(0,2)}, \dots, \widehat{t}_n^{(0,2)})$ ,  $\widehat{\mathbf{M}} = \text{diag}(\widehat{\mu}_1, \dots, \widehat{\mu}_n)$ ,  $\widetilde{\mathbf{M}} = \text{diag}(\widetilde{\mu}_1, \dots, \widetilde{\mu}_n)$ ,  $\widehat{b}_{(1)} = \frac{\partial \phi}{\partial \eta_2} \Big|_{\phi=\widehat{\phi}}$  e  $\widetilde{b}_{(1)} = \frac{\partial \phi}{\partial \eta_2} \Big|_{\phi=\widetilde{\phi}}$ . A distribuição assintótica de  $\text{LR}_{sk1}$  e  $\text{LR}_{sk2}$  pode ser aproximada pela distribuição  $\chi_q^2$  (Skovgaard, 2001). O desenvolvimento de como encontrar esses resultados pode ser visto no Apêndice C.

### 3.3 Correção de Bartlett-*bootstrap*

Uma alternativa para a correção de Bartlett analítica foi desenvolvida por [Rocke \(1989\)](#) usando o método *bootstrap* ([Efron, 1979](#)). O método *bootstrap* paramétrico funciona da seguinte maneira:  $B$  amostras são geradas com o modelo sob a hipótese nula, digamos  $(\mathbf{y}^{*1}, \dots, \mathbf{y}^{*B})$ , substituindo  $\boldsymbol{\theta}$  por  $\widetilde{\boldsymbol{\theta}}$ . Em seguida, a estatística LR é calculada para cada pseudo-amostra  $\mathbf{y}^{*b}$ ,  $b = 1, \dots, B$ , chamadas de  $\text{LR}^{*b}$ ,  $b = 1, \dots, B$ , tal que

$$\text{LR}^{*b} = 2 \left[ \ell(\widehat{\boldsymbol{\theta}}^{*b}; \mathbf{y}^{*b}) - \ell(\widetilde{\boldsymbol{\theta}}^{*b}; \mathbf{y}^{*b}) \right],$$

em que  $\widetilde{\boldsymbol{\theta}}^{*b}$  e  $\widehat{\boldsymbol{\theta}}^{*b}$  são, respectivamente, os EMV sob a hipótese nula e sob a hipótese alternativa, para cada pseudo-amostra  $b = 1, \dots, B$ . Seja  $\overline{\text{LR}}^*$  a média das  $B$  amostras *bootstrap*, isto é,  $\overline{\text{LR}}^* = \frac{1}{B} \sum_{b=1}^B \text{LR}^{*b}$ . A estatística da razão de verossimilhanças corrigida por Bartlett-*bootstrap* é definida por

$$\text{LR}_{boot} = q \frac{\text{LR}}{\text{LR}^*}.$$

Assim como nas demais estatísticas corrigidas abordadas neste capítulo, a distribuição assintótica da  $\text{LR}_{boot}$  é aproximadamente  $\chi_q^2$  com erro de aproximação da ordem  $\mathcal{O}(n^{-2})$ .

## 4 Simulações

Realizamos algumas simulações de Monte Carlo para avaliar o tamanho do teste, ou seja, a taxa de rejeição da hipótese nula quando esta é verdadeira, e o poder do teste, o qual é a taxa de rejeição da hipótese nula quando esta é falsa, da razão de verossimilhanças (LR) e dos testes corrigidos  $LR_{b1}$ ,  $LR_{b2}$ ,  $LR_{b3}$ ,  $LR_{sk1}$ ,  $LR_{sk2}$  e  $LR_{boot}$ . As simulações foram realizadas no *software* R (R Core Team, 2021) e com o auxílio dos pacotes `BPmodel` (Santos-Neto et al., 2022) e `GAMLSS` (Rigby e Stasinopoulos, 2005). Os códigos e resultados desses experimentos estão disponíveis no *github*<sup>1</sup>.

### 4.1 Tamanho do teste

O seguinte modelo foi considerado nos experimentos da simulação:

$$\log(\mu_i) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_q x_{qi} + \dots + \beta_{p-1} x_{p-1i}, \quad i = 1, \dots, n,$$

em que  $x_{1i}, \dots, x_{p-1i}$  foram extraídas de uma distribuição uniforme no intervalo  $(0, 1)$ ;  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$  representa o vetor de médias de  $\boldsymbol{y} = (y_1, \dots, y_n)^\top$ , um vetor aleatório em que cada termo é independente e  $y_i \sim BP(\mu_i, \phi)$ . Testamos as hipóteses  $\mathcal{H}_0 : \boldsymbol{\beta}_1 = \mathbf{0}$  contra  $\mathcal{H}_1 : \boldsymbol{\beta}_1 \neq \mathbf{0}$ , sendo  $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_q)^\top$ . A reparametrização para o  $\phi$  foi a logarítmica, isto é,  $\eta_2 = \log \phi$ .

Para o tamanho da amostra ( $n$ ), os valores escolhidos foram  $n = 20, 30$  e  $40$ ; para o parâmetro de precisão ( $\phi$ ), os valores escolhidos foram  $\phi = 0, 5, 1$  e  $10$ ; para o tamanho do vetor  $\boldsymbol{\beta}$  ( $p$ ), foram escolhidos os valores  $p = 3, 4$  e  $5$  e, por fim, para a quantidade de restrições sob a hipótese nula ( $q$ ), as quantidades escolhidas foram  $q = 1, 2$  e  $3$ . Foram realizadas 10000 réplicas Monte Carlo para cada combinação de valores escolhidos e 600 réplicas *bootstrap* ( $B = 600$ ). Além disso, fixamos os valores de  $\beta_i = 0$ , para  $i = 1, \dots, q$  e  $\beta_i = 1$  para  $i = 0, q + 1, \dots, p - 1$ .

Na Tabela 2 é possível observar as taxas de rejeição da hipótese nula  $\mathcal{H}_0 : \boldsymbol{\beta}_1 = \mathbf{0}$  com  $p = 3$  para cada estatística. Como esperado, o teste da razão de verossimilhanças usual apresentou uma taxa mais distorcida que os demais. Por exemplo, para  $\phi = 0, 5$ ,  $n = 20$  e  $\alpha = 10\%$ , o teste usual tem uma taxa de rejeição de 14,19%. Por outro lado, os testes corrigidos tiveram as seguintes taxas: 10,23% ( $LR_{b1}$ ), 11,20% ( $LR_{b2}$ ), 10,17% ( $LR_{b3}$ ), 9,52% ( $LR_{sk1}$ ), 12,93% ( $LR_{sk2}$ ) e 9,47% ( $LR_{boot}$ ).

Ainda na Tabela 2, podemos perceber que, dentre as estatísticas corrigidas por Bartlett, a estatística  $LR_{b1}$  apresenta um valor empírico mais próximo do nível nominal

<sup>1</sup> <https://github.com/IvoSilvestre/bpcorrection>

que a estatística  $LR_{b3}$ , como em  $\phi = 0,5$ ,  $n = 20$  e  $\alpha = 1\%$ , a taxa de rejeição empírica da  $LR_{b1}$  foi de 0,94% contra 1,67% da  $LR_{b3}$ . Esse cenário se inverte para valores maiores de  $\phi$ , por exemplo: para  $\phi = 10$ ,  $n = 20$  e  $\alpha = 10\%$ , as taxas dos dois testes foram iguais a 11,06% e 9,98%, respectivamente. Em relação à estatística  $LR_{b2}$ , esta apresentou taxas, no geral, piores que as demais estatísticas de Bartlett, exceto com valores maiores de  $\phi$  e, em um caso, pior que a estatística LR usual (2,25% na usual e 2,49% na  $LR_{b2}$ , com  $\alpha = 1\%$ ). Em relação aos testes corrigidos pelo ajuste de Skovgaard, a estatística  $LR_{sk1}$  apresentou um nível de correção próximo do nominal e, em alguns casos, até mais próximo do que as correções de Bartlett e de Bartlett-*bootstrap*, como em  $\phi = 1$ ,  $n = 20$  e  $\alpha = 10\%$ , em que a  $LR_{b1}$  e a  $LR_{boot}$  tiveram taxas de rejeição iguais a 10,22% e 9,72%, respectivamente, e a estatística  $LR_{sk1}$  teve taxa igual a 9,91%. Já a estatística  $LR_{sk2}$  teve o pior desempenho dentre as corrigidas em todos os cenários e, em alguns casos, pior que a estatística LR (1,68% na usual, contra 3,87% na  $LR_{sk2}$  para  $\phi = 0,5$ ,  $n = 30$  e  $\alpha = 1\%$ ).

Para valores diferentes de  $p$  e  $q$ , as taxas de rejeição da hipótese nula no teste da razão de verossimilhanças foi maior. Entretanto, encontramos alguns comportamentos similares dentre as corrigidas. Nas Tabelas 3 e 4, os mesmo padrões de aumento na taxa de rejeição quando o valor  $\phi$  aumenta para o teste  $LR_{b1}$  (10,86%, 10,89% e 11,48% para  $\alpha = 10\%$ ,  $n = 20$  e  $\phi = 0,5$ , 1 e 10, respectivamente, na Tabela 3), o teste  $LR_{sk2}$  apresentando o pior desempenho dentre os testes corrigidos por Bartlett para valores menores de  $\phi$  (como 2,00% na taxa de rejeição para  $\alpha = 1\%$ ,  $n = 20$  e  $\phi = 1$ , na Tabela 4) e o teste  $LR_{sk1}$  tendo, em geral, um desempenho melhor que o teste  $LR_{sk2}$  (na Tabela 3, a taxa de rejeição para  $\alpha = 10\%$ ,  $n = 30$  e  $\phi = 10$  para a  $LR_{sk1}$  foi de 9,98% contra 10,16% para a  $LR_{sk2}$ ). Nesses casos, entretanto, o teste  $LR_{sk2}$  não ultrapassou a taxa de rejeição do teste da razão de verossimilhanças usual.

Em resumo, a taxa de rejeição do teste da razão de verossimilhanças em amostras de tamanho pequeno geralmente aumenta quando o número de parâmetros da regressão ( $p$ ) e o número de parâmetros a serem testados ( $q$ ) aumentam. Entretanto, nos testes corrigidos, em geral, essa taxa de rejeição se mantém próxima do nível nominal.

Tabela 2 – Taxa de rejeição (%) da hipótese nula para  $\mathcal{H}_0 : \beta_1 = 0$  com  $p = 3$ .

$\phi$	$n$	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
		20	30	40	20	30	40	20	30	40
		Estadística								
0,5	LR	14,19	13,42	12,25	7,90	7,25	6,58	2,25	1,68	1,35
	LR <sub>b1</sub>	10,23	10,54	10,26	5,01	5,29	5,20	1,21	1,08	0,94
	LR <sub>b2</sub>	11,20	11,46	11,33	6,13	6,21	6,22	2,49	2,12	2,00
	LR <sub>b3</sub>	10,17	11,01	10,99	5,29	5,74	5,89	1,96	1,78	1,67
	LR <sub>sk1</sub>	9,52	10,39	10,38	4,80	5,07	5,11	1,09	0,99	0,97
	LR <sub>sk2</sub>	12,93	13,86	13,77	7,78	8,33	8,24	3,80	3,87	3,77
	LR <sub>boot</sub>	9,47	10,04	9,95	4,64	5,02	5,09	1,06	0,97	0,87
1	LR	14,31	13,30	12,25	7,83	7,44	6,53	2,30	1,87	1,80
	LR <sub>b1</sub>	10,22	10,69	10,47	5,24	5,46	5,13	1,29	1,09	1,29
	LR <sub>b2</sub>	10,27	10,73	10,62	5,52	5,55	5,35	1,67	1,28	1,53
	LR <sub>b3</sub>	9,60	10,43	10,40	5,15	5,31	5,20	1,41	1,16	1,38
	LR <sub>sk1</sub>	9,91	10,65	10,26	5,09	5,26	5,11	1,16	1,10	1,27
	LR <sub>sk2</sub>	12,77	13,56	13,38	7,65	7,93	8,08	3,42	3,54	3,84
	LR <sub>boot</sub>	9,72	10,61	10,27	4,84	5,28	5,02	1,06	1,06	1,18
10	LR	14,87	12,48	12,02	8,44	6,55	6,41	2,29	1,71	1,32
	LR <sub>b1</sub>	11,06	9,95	10,09	5,62	4,82	5,17	1,09	1,02	0,87
	LR <sub>b2</sub>	10,60	9,84	10,02	5,32	4,73	5,08	1,02	1,00	0,87
	LR <sub>b3</sub>	9,98	9,58	9,92	4,96	4,58	5,02	0,90	0,96	0,85
	LR <sub>sk1</sub>	10,19	9,50	9,98	5,14	4,56	5,04	0,97	0,95	0,85
	LR <sub>sk2</sub>	11,18	10,66	11,24	5,95	5,64	6,23	1,59	1,85	1,91
	LR <sub>boot</sub>	10,18	9,69	10,04	5,12	4,69	5,05	0,93	1,00	0,92

Tabela 3 – Taxa de rejeição (%) da hipótese nula para  $\mathcal{H}_0 : \beta_1 = \beta_2 = 0$  com  $p = 4$ .

$\phi$	$n$	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
		20	30	40	20	30	40	20	30	40
	Estadística									
0,5	LR	17,10	14,36	12,75	10,29	8,21	7,26	2,97	2,08	1,67
	LR <sub>b1</sub>	10,86	10,56	9,84	5,52	5,34	5,09	1,38	1,26	1,02
	LR <sub>b2</sub>	11,36	11,47	10,92	6,26	6,41	6,10	2,47	2,35	2,13
	LR <sub>b3</sub>	10,21	10,91	10,56	5,45	5,82	5,83	2,00	1,13	1,93
	LR <sub>sk1</sub>	9,80	10,16	9,64	4,74	4,94	4,90	1,09	0,96	0,96
	LR <sub>sk2</sub>	10,45	10,47	9,82	5,12	5,10	5,08	1,32	1,12	1,05
	LR <sub>boot</sub>	9,71	9,89	9,66	4,75	4,80	4,98	1,10	1,02	0,95
1	LR	17,39	14,29	13,38	10,33	8,15	7,00	2,94	2,05	1,81
	LR <sub>b1</sub>	10,89	10,25	10,16	5,63	5,27	4,90	1,27	0,96	1,07
	LR <sub>b2</sub>	10,67	10,38	10,11	5,56	5,40	4,96	1,65	1,29	1,15
	LR <sub>b3</sub>	9,65	10,02	9,90	4,98	5,12	4,84	1,44	1,16	1,09
	LR <sub>sk1</sub>	9,86	10,03	10,14	4,86	5,14	4,91	1,01	0,88	1,06
	LR <sub>sk2</sub>	10,42	10,27	10,36	5,24	5,31	5,07	1,22	0,99	1,15
	LR <sub>boot</sub>	9,59	9,92	9,98	4,82	5,00	4,86	1,05	0,85	1,06
10	LR	17,50	14,56	14,07	10,63	8,08	7,61	3,10	2,26	2,04
	LR <sub>b1</sub>	11,48	10,46	10,94	5,75	5,27	5,59	1,38	1,05	1,20
	LR <sub>b2</sub>	10,78	10,26	10,77	5,30	5,09	5,47	1,24	1,02	1,17
	LR <sub>b3</sub>	9,88	9,89	10,59	4,71	4,85	5,36	1,10	0,89	1,11
	LR <sub>sk1</sub>	10,04	9,98	10,65	4,85	4,86	5,35	1,10	0,95	1,14
	LR <sub>sk2</sub>	10,60	10,16	10,74	5,11	5,04	5,44	1,16	0,99	1,17
	LR <sub>boot</sub>	10,08	10,08	10,63	4,75	4,83	5,40	1,10	0,95	1,24

Tabela 4 – Taxa de rejeição (%) da hipótese nula para  $\mathcal{H}_0 : \beta_1 = \beta_2 = \beta_3 = 0$  com  $p = 5$ .

$\phi$	$n$	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
		20	30	40	20	30	40	20	30	40
		Estatística								
0,5	LR	20,99	16,90	14,45	12,94	9,71	8,01	4,10	2,74	1,93
	LR <sub>b1</sub>	11,24	10,86	10,22	5,85	5,89	5,27	1,45	1,20	0,99
	LR <sub>b2</sub>	11,72	11,59	11,11	6,63	6,75	6,17	2,67	2,30	2,07
	LR <sub>b3</sub>	9,97	10,89	10,66	5,73	6,24	5,86	2,07	2,07	1,77
	LR <sub>sk1</sub>	10,14	10,44	10,09	5,04	5,39	5,11	0,98	1,00	0,85
	LR <sub>sk2</sub>	10,82	10,62	10,22	5,40	5,59	5,25	1,06	1,06	0,88
	LR <sub>boot</sub>	9,81	10,24	9,87	4,84	5,27	4,89	1,18	0,91	0,79
1	LR	20,63	16,20	14,81	13,00	9,56	8,38	4,29	2,77	2,25
	LR <sub>b1</sub>	12,04	10,83	10,63	6,43	5,52	5,52	1,62	1,11	1,22
	LR <sub>b2</sub>	11,56	10,70	10,53	6,31	5,60	5,47	2,00	1,32	1,26
	LR <sub>b3</sub>	10,19	10,12	10,22	5,33	5,26	5,30	1,47	1,18	1,19
	LR <sub>sk1</sub>	10,55	10,25	10,53	5,32	5,24	5,41	1,21	0,98	1,16
	LR <sub>sk2</sub>	11,29	10,56	10,68	5,72	5,40	5,53	1,36	1,07	1,20
	LR <sub>boot</sub>	10,30	10,16	10,45	5,21	5,15	5,34	1,19	1,01	1,17
10	LR	21,60	16,43	14,89	13,67	9,86	8,11	4,52	2,78	2,10
	LR <sub>b1</sub>	12,91	11,28	10,66	6,87	5,50	5,37	1,53	1,31	1,00
	LR <sub>b2</sub>	11,98	10,90	10,34	6,30	5,25	5,19	1,28	1,26	0,97
	LR <sub>b3</sub>	10,59	10,33	10,06	5,30	4,79	4,95	1,03	1,12	0,91
	LR <sub>sk1</sub>	10,62	10,52	10,20	5,51	4,94	5,05	1,00	1,22	0,95
	LR <sub>sk2</sub>	11,47	10,87	10,39	6,00	5,18	5,17	1,17	1,24	0,98
	LR <sub>boot</sub>	10,73	10,39	10,23	5,60	4,99	5,13	1,08	1,13	0,98

Uma medida para analisar graficamente a diferença entre as distribuições empíricas das estatísticas de teste e a distribuição qui-quadrado de referência é a partir das discrepâncias relativas, cuja expressão é dada por

$$DR(Est; \alpha) = \frac{Est(\alpha) - \chi_q^2(\alpha)}{\chi_q^2(\alpha)},$$

em que  $Est$  representa a estatística de interesse,  $Est(\alpha)$  é o quantil amostral de ordem  $100\alpha\%$  e  $\chi_q^2(\alpha)$  simboliza o quantil teórico de ordem  $100\alpha\%$  da distribuição qui-quadrado com  $q$  graus de liberdade. Valores mais próximos de zero indicam uma melhor aproximação da distribuição empírica pela distribuição qui-quadrado.

As Figuras 2 e 3 exibem os gráficos das discrepâncias relativas por quantis assintóticos para  $n = 30$ ,  $\phi = 10$  e diferentes valores de  $p$  e  $q$ . Como esperado, o teste da razão de verossimilhanças apresentou uma discrepância relativa maior que todos os testes corrigidos em ambos os casos.

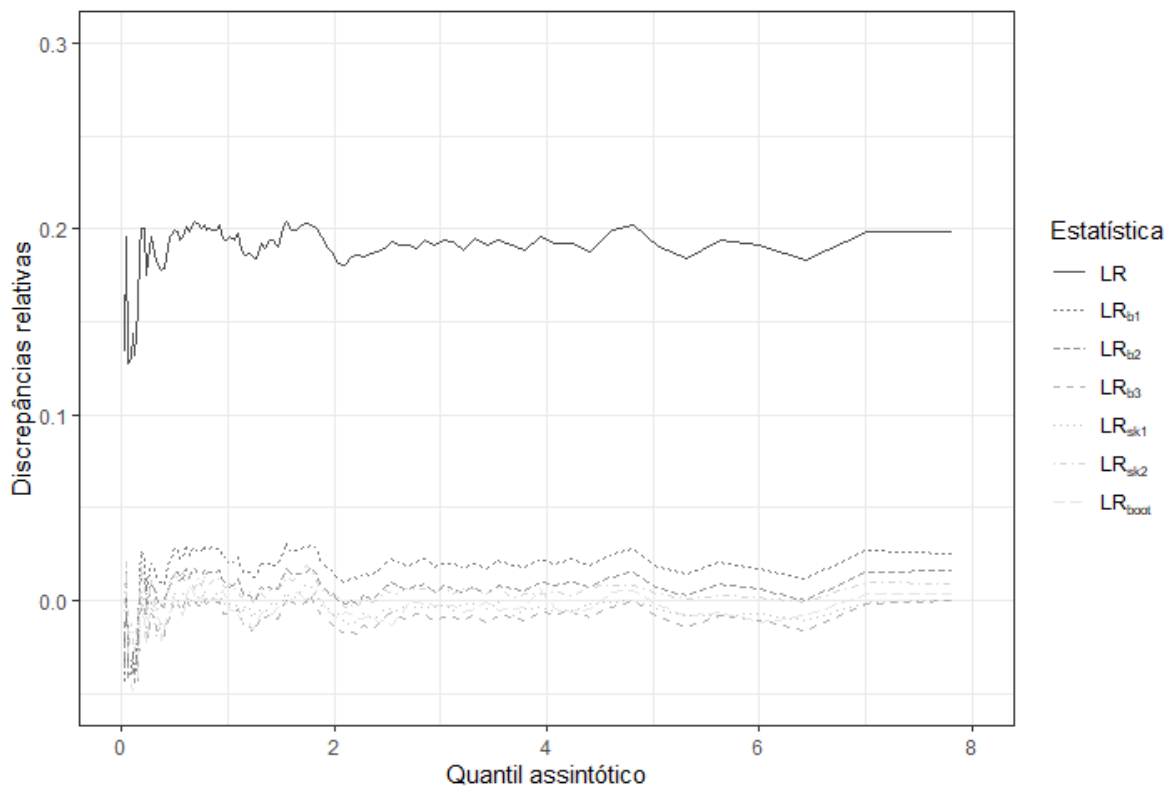


Figura 2 – Discrepâncias relativas por quantis assintóticos para  $n = 30$ ,  $\phi = 10$ ,  $p = 4$  e  $q = 2$ .

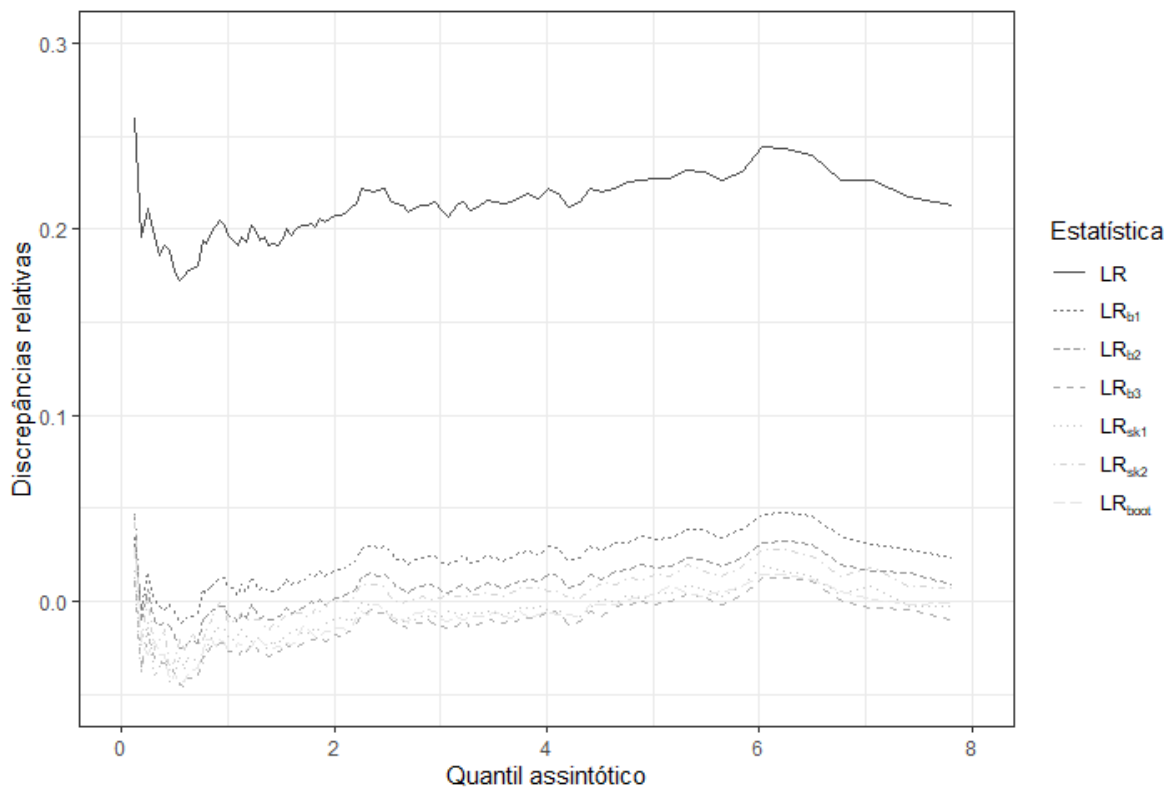


Figura 3 – Discrepâncias relativas por quantis assintóticos para  $n = 30$ ,  $\phi = 10$ ,  $p = 5$  e  $q = 3$ .

A Figura 4 mostra os *boxplots* dos valores das estatísticas de teste para diferentes tamanhos amostrais com  $p = 5$  e  $q = 2$ . Como esperado, em amostras de tamanho pequeno, a distribuição das estatísticas de teste são assimétricas, indicando uma boa aproximação pela distribuição qui-quadrado de referência.

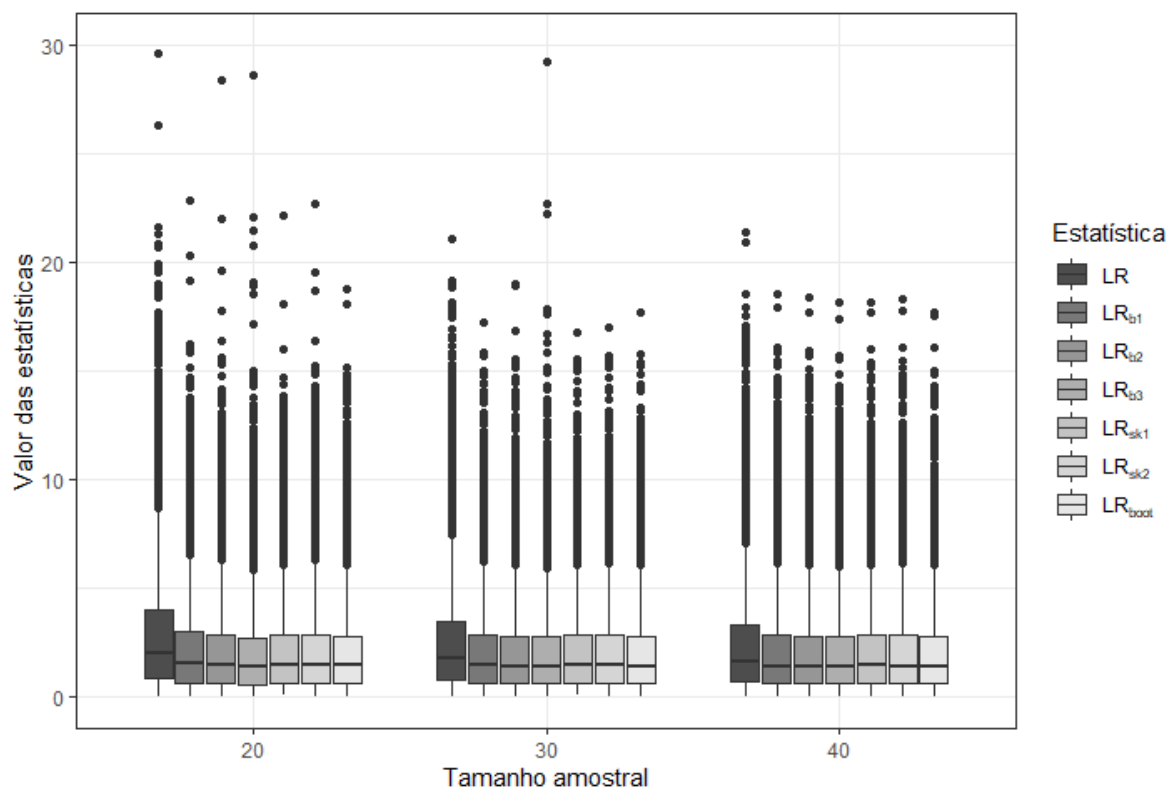


Figura 4 – Boxplot do valor das estatísticas de teste para  $\phi = 1$ ,  $p = 5$  e  $q = 2$ .

## 4.2 Poder do teste

Em relação ao poder do teste, consideramos as seguintes hipóteses:

$$\mathcal{H}_0 : \beta_1 = \beta_2 = 0,$$

contra

$$\mathcal{H}_1 : \beta_1 = \beta_2 = \delta,$$

com  $\delta$  variando entre  $-2$  e  $2$  em intervalos de  $0,5$ . Consideramos o cenário em que  $q = 2$ ,  $p = 5$  e  $n = 30$  para realizar as simulações. Previamente, calculamos os quantis empíricos de ordem  $0,95$  gerados a partir de  $500000$  réplicas de Monte Carlo de cada estatística, para que todos os testes tenham um nível de significância de aproximadamente  $5\%$ . Com os quantis calculados, consideremos  $10000$  réplicas de Monte Carlo e o teste rejeita  $\mathcal{H}_0$  se o valor da estatística for maior que o quantil empírico previamente calculado.

Na Tabela 5, podemos perceber que a taxa de crescimento do poder do teste é maior para maiores valores do parâmetro  $\phi$ . Por exemplo, com  $\phi = 10$ , todas as estatísticas alcançaram  $100\%$  de rejeição da hipótese nula com  $\delta = 1,5$ , já com  $\phi = 0,5$  ainda não foi alcançada a rejeição com  $\delta = 2$  ou  $\delta = -2$ . Além disso, podemos perceber que para maiores valores de  $\phi$ , o poder do teste foi parecido em todos os testes. No entanto, em valores menores de  $\phi$ , o teste LR apresentou o poder um pouco maior que os demais, mas, ainda assim, não deve ser escolhido em relação aos demais devido aos resultados encontrados na Seção 4.1. Também podemos mencionar que o teste Bartlett-*bootstrap* apresentou resultados similares aos do teste LR e que os testes de Skovgaard tiveram resultados inferiores que os de Bartlett.

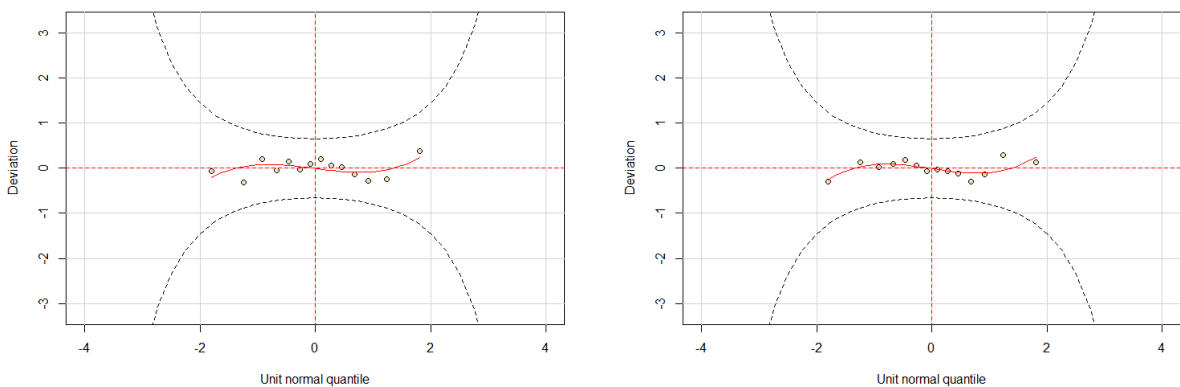


## 5 Aplicação

Para ilustrar os resultados obtidos nesse trabalho, consideramos um conjunto de dados analisado por Rosillo, Martin e Egido (2010) e Medeiros, Ferrari e Lemonte (2017), sobre o tempo de vida de 14 observações sobre lâmpadas fluorescentes. Segundo os autores, podemos estabelecer a seguinte relação funcional:

$$\log \mu_i = \eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 \frac{x_{2i}}{x_{3i}} + \beta_3 \left( \frac{x_{2i}}{x_{3i}} \right)^2, \quad i = 1, \dots, 14,$$

em que as covariáveis  $x_1$ ,  $x_2$  e  $x_3$  representam o brilho no tempo inicial, a voltagem e a voltagem nominal, respectivamente, e a variável resposta,  $y$ , representa o tempo de vida das lâmpadas. Medeiros, Ferrari e Lemonte (2017) propuseram um modelo de regressão log-gama para esse conjunto de dados, entretanto, como pode ser visto na Figura 5, os *wormplots* (Buuren e Fredriks, 2001) de ambos os modelos. O *wormplot* é um gráfico que compara o quantil da distribuição padrão (eixo  $x$ ) com a diferença entre o resíduo quantílico (Dunn e Smyth, 1996) e o quantil da normal padrão (eixo  $y$ ). Valores próximos a zero no eixo  $y$  indicam um bom ajuste do modelo, as parábolas funcionam como limites de confiança e a linha de tendência pode indicar se há problemas na média (segmento de reta paralela ao eixo das abscissas), variância (segmento de reta inclinado), assimetria (curva em formato de  $U$ ) ou curtose (curva em formato de  $S$ ) do modelo ajustado. Como o modelo de regressão beta-prime apresentou seus desvios próximos ao do eixo das abscissas, assim como o modelo log-gama, então a distribuição beta-prime pode ser utilizada como distribuição da variável resposta.



(a) Modelo log-gama

(b) Modelo beta-prime

Figura 5 – *Wormplots* dos modelos ajustados

Após ajustarmos o modelo de regressão beta-prime com parâmetro de precisão desconhecido, os EMV, os desvios-padrão, as estatísticas da razão de verossimilhanças dos

testes  $\mathcal{H}_0 : \beta_j = 0$ ,  $j = 1, 2, 3$ , e seus respectivos  $p$ -valores são apresentados na Tabela 6. Os dados foram ajustados a partir do *software* livre R (R Core Team, 2021) e com o auxílio dos pacotes `BPmodel` (Santos-Neto et al., 2022) e `GAMLSS` (Rigby e Stasinopoulos, 2005).

Tabela 6 – Estimativas de máxima verossimilhança (EMV), desvios-padrão (DP), estatística da razão de verossimilhanças (LR) e  $p$ -valor do modelo ajustado.

Estimadores	EMV	DP	LR	$p$ -valor
$\widehat{\beta}_0$	-18,354	1,485	–	–
$\widehat{\beta}_1$	-7,426	3,623	3,8413	0,0500
$\widehat{\beta}_2$	35,247	3,059	35,0131	< 0,0001
$\widehat{\beta}_3$	-16,713	1,575	33,2267	< 0,0001
$\widehat{\phi}$	4,802	0,384	–	–

A um nível de significância de 5%, não há fortes evidências para rejeitar  $\mathcal{H}_0 : \beta_1 = 0$ , já que seu  $p$ -valor foi 0,0500. Portanto, por se tratar de uma amostra de tamanho pequeno, o valor da estatística pode sofrer alterações relevantes ao aplicarmos as correções. Na Tabela 7, todas as correções diminuíram o valor da estatística de teste, aumentando seus  $p$ -valores. Logo, com as estatísticas corrigidas, não rejeitamos a hipótese nula a um nível de 5% com maior segurança. Por exemplo, com a estatística  $LR_{b3}$ , o valor da estatística foi de 2,6590, com  $p$ -valor de 0,1030, o que gera mais segurança para a rejeição da hipótese nula.

Tabela 7 – Estatísticas de teste e  $p$ -valores para testar  $\mathcal{H}_0 : \beta_1 = 0$  contra  $\mathcal{H}_0 : \beta_1 \neq 0$ .

Estatística	valor	$p$ -valor
LR	3,8413	0,0500
$LR_{b1}$	2,9373	0,0866
$LR_{b2}$	2,8237	0,0929
$LR_{b3}$	2,6590	0,1030
$LR_{sk1}$	2,5191	0,1125
$LR_{sk2}$	2,6329	0,1047
$LR_{boot}$	2,7280	0,0986

## 6 Considerações Finais

Nesse trabalho, derivamos um fator de correção de Bartlett e o ajuste de Skovgaard para o teste da razão de verossimilhanças em modelos de regressão Beta-prime. Além disso, comparamos numericamente, a partir de simulações de Monte Carlo, o teste da razão de verossimilhanças usual com os três testes corrigidos por Bartlett, com os dois testes corrigidos por Skovgaard e o teste corrigido por Bartlett-*bootstrap*. Por fim, também realizamos uma aplicação com dados reais.

Como resultados, podemos mencionar que, como esperado, o teste da razão de verossimilhanças teve uma taxa de rejeição empírica superior ao nível nominal. Entretanto, nos testes corrigidos, essa taxa de rejeição se mostrou mais próxima do nível nominal, com destaque para as estatísticas  $LR_{b3}$ ,  $LR_{sk1}$  e  $LR_{boot}$ . Também podemos destacar que, na aplicação, ao usar o teste da razão de verossimilhanças, um dos parâmetros da regressão apresentou um  $p$ -valor muito próximo do nível de significância. Entretanto, ao utilizar os testes corrigidos, o  $p$ -valor aumentou consideravelmente.

Como projetos futuros, pode-se pensar em: desenvolver fatores de correção Bartlett e de Skovgaard do teste da razão de verossimilhanças para modelos de regressão beta-prime com precisão variada; derivar outros tipos de correção, como de viés a matriz de variância de segunda ordem ou de coeficiente de assimetria, e aprimoramento de outros testes, como o score e gradiente, com o fator de correção tipo-Bartlett.

## Referências

- ARAÚJO, M. C.; CYSNEIROS, A. H.; CYSNEIROS, M. Bartlett and Bartlett-type corrections in heteroscedastic symmetric nonlinear regression models. *Anais da Academia Brasileira de Ciências*, SciELO Brasil, v. 94, 2022. Citado na página 11.
- ARAÚJO, M. C.; CYSNEIROS, A. H.; MONTENEGRO, L. C. Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. *Statistical Papers*, Springer, v. 61, p. 167–188, 2020. Citado na página 12.
- BARNDORFF-NIELSEN, O. E. Infereni on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika*, Oxford University Press, v. 73, n. 2, p. 307–322, 1986. Citado na página 19.
- BARNDORFF-NIELSEN, O. E. Modified signed log likelihood ratio. *Biometrika*, Oxford University Press, v. 78, n. 3, p. 557–563, 1991. Citado na página 19.
- BARTLETT, M. Proprieties of sufficiency and statistical tests. *Royal Society of London Proceedings Series A*, v. 160, p. 268–282, 1937. Citado 3 vezes nas páginas 11, 17 e 18.
- BAYER, F. *Modelagem e Inferência em Regressão beta*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2011. Citado na página 19.
- BAYER, F. M.; CRIBARI-NETO, F. Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference*, Elsevier, v. 143, n. 3, p. 531–547, 2013. Citado na página 11.
- BOURGUIGNON, M.; SANTOS-NETO, M.; CASTRO, M. d. A new regression model for positive data. *arXiv preprint arXiv:1804.07734*, 2018. Citado 3 vezes nas páginas 12, 13 e 15.
- BUUREN, S. v.; FREDRIKS, M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in medicine*, Wiley Online Library, v. 20, n. 8, p. 1259–1277, 2001. Citado na página 31.
- CANEPA, A. *Improvement on the LR Test Statistic on the Cointegrating Relations in VAR Models: Bootstrap Methods and Applications*. [S.l.], 2020. Disponível em: <<https://ideas.repec.org/p/uto/dipeco/202007.html>>. Citado na página 12.
- CARNEIRO, H. P.; SANDOVAL, M. C.; BOTTER, D. A.; MAGALHÃES, T. M. Some asymptotic inferential aspects of the kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, Taylor & Francis, p. 1–17, 2023. Citado 2 vezes nas páginas 11 e 12.
- CORDEIRO, G. General matrix formulae for computing Bartlett corrections. *Statistics and Probability Letters*, v. 16, p. 11–18, 1993. Citado na página 18.
- COX, D. R.; HINKLEY, D. V. *Theoretical statistics*. Flórida: CRC Press, 1979. Citado na página 16.

- DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. *Journal of Computational and graphical statistics*, Taylor & Francis, v. 5, n. 3, p. 236–244, 1996. Citado na página 31.
- EFRON, B. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, v. 7, p. 1–26, 1979. Citado 2 vezes nas páginas 17 e 20.
- FERRARI, S. L.; PINHEIRO, E. C. Improved likelihood inference in beta regression. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 81, n. 4, p. 431–443, 2011. Citado na página 11.
- GUEDES, A. C.; CRIBARI-NETO, F.; ESPINHEIRA, P. L. Modified likelihood ratio tests for unit gamma regressions. *Journal of Applied Statistics*, Taylor & Francis, v. 47, n. 9, p. 1562–1586, 2020. Citado na página 11.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous univariate distributions, volume 2*. New York: John wiley & sons, 1995. v. 289. Citado na página 13.
- KEEPING, E. *Introduction to Statistical Inference*. New York: Van Nostrand, 1962. Citado 2 vezes nas páginas 12 e 13.
- KORKMAZ, S.; GÖKGÖZ, F. Generalized linear models for beta-prime regression analysis with application to bladder cancer data. *International Journal of Statistics and Probability*, v. 9, n. 4, p. 15–27, 2020. Citado na página 12.
- KOSMIDIS, I.; FIRTH, D. Regression models for positive data using a new class of distributions. *Journal of Applied Statistics*, v. 46, n. 10, p. 1812–1831, 2019. Citado na página 12.
- LAWLEY, D. A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, v. 43, p. 295–303, 1956. Citado 3 vezes nas páginas 11, 17 e 18.
- LOOSE, L. H.; BAYER, F. M.; PEREIRA, T. L. Bootstrap bartlett correction in inflated beta regression. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 46, n. 4, p. 2865–2879, 2017. Citado na página 12.
- LOOSE, L. H.; VALENÇA, D. M.; BAYER, F. M. On bootstrap testing inference in cure rate models. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 88, n. 17, p. 3437–3454, 2018. Citado na página 12.
- MCDONALD, J. Some generalized functions for the size distribution of income. *Econometrica*, v. 52, p. 647–663, 1984. Citado 2 vezes nas páginas 12 e 13.
- MEDEIROS, F. M.; FERRARI, S. L.; LEMONTE, A. J. Improved inference in dispersion models. *Applied Mathematical Modelling*, Elsevier, v. 51, p. 317–328, 2017. Citado 2 vezes nas páginas 11 e 31.
- MEDEIROS, F. M.; LEMONTE, A. J. Likelihood-based inference in censored exponential regression models. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 50, n. 13, p. 3214–3233, 2021. Citado na página 11.
- MELO, T. F.; VARGAS, T. M.; LEMONTE, A. J.; MORENO-ARENAS, G. Higher-order asymptotic refinements in the multivariate dirichlet regression model. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 51, n. 1, p. 53–71, 2022. Citado na página 11.

- PERETTE, A. C. *Implementação no software estatístico R de modelos de regressão normal com parametrização geral*. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado na página 11.
- PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. *Numerical recipes in C*. Reino Unido: Cambridge university press Cambridge, 1992. Citado na página 16.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado 4 vezes nas páginas 16, 19, 22 e 32.
- RAUBER, C.; CRIBARI-NETO, F.; BAYER, F. M. Improved testing inferences for beta regressions with parametric mean link function. *AStA Advances in Statistical Analysis*, Springer, v. 104, n. 4, p. 687–717, 2020. Citado na página 11.
- RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, v. 54, p. 507–554, 2005. Citado 3 vezes nas páginas 16, 22 e 32.
- ROCKE, D. Bootstrap bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association*, v. 84, p. 598–601, 1989. Citado 4 vezes nas páginas 11, 17, 18 e 20.
- ROSILLO, F. G.; MARTIN, N.; EGIDO, M. Comparison of conventional and accelerated lifetime testing of fluorescent lamps. *Lighting Research & Technology*, SAGE Publications Sage UK: London, England, v. 42, n. 2, p. 243–259, 2010. Citado na página 31.
- SANTOS-NETO, M.; BOURGUIGNON, M.; CASTRO, M. de; RIBEIRO-BEZERRA, T. *BPmodel: Beta-Prime Regression Model*. Brasil, 2022. R package version 2.0. Disponível em: <<https://github.com/santosneto/BPmodel>>. Citado 3 vezes nas páginas 16, 22 e 32.
- SKOVGAARD, I. Likelihood asymptotic. *Scandinavian Journal of Statistics*, v. 28, p. 3–13, 2001. Citado 4 vezes nas páginas 11, 17, 19 e 20.
- WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The annals of mathematical statistics*, JSTOR, v. 9, n. 1, p. 60–62, 1938. Citado 2 vezes nas páginas 11 e 17.
- ZHANG, L.; WANG, J.; PENG, Y. Beta-prime regression models for analyzing positively skewed data with applications to medical studies. *Statistics in Medicine*, v. 40, n. 4, p. 940–958, 2021. Citado na página 12.

# Apêndices

# APÊNDICE A – Vetor escore, matriz Hessiana e matriz de informação de Fisher

Definimos, após algumas análises, as quantidades

$$\begin{aligned} a_{i,(m)} &= \frac{\partial^m \mu_i}{\partial \eta_{1i}^m}, & b_{(m)} &= \frac{\partial^m \phi}{\partial \eta_2^m}, \\ y_i^* &= \log \left( \frac{y_i}{1 + y_i} \right), & \mu_i^* &= -d_i^{(0,1)}, \\ y_i^{\mu_i} &= \log \left( \frac{y_i^{\mu_i}}{(1 + y_i)^{1 + \mu_i}} \right), & \mu_i^* &= -r_i^{(1,1)}, \end{aligned}$$

em que

$$\begin{aligned} d_i^{(k,m)} &= d(k, m, \mu_i, \phi) = (1 + \phi)^k \left( \psi^{(m-1)}(\mu_i(1 + \phi) + \phi + 2) - \psi^{(m-1)}(\mu_i(1 + \phi)) \right), \\ r_i^{(k,m)} &= r(k, m, \mu_i, \phi) = (1 + \mu_i)^k \psi^{(m-1)}(\mu_i(1 + \phi) + \phi + 2) - \mu_i^k \psi^{(m-1)}(\mu_i(1 + \phi)), \\ &\quad - \psi^{(m-1)}(\phi + 2) \\ t_i^{(k,m)} &= t(k, m, \mu_i, \phi) = (1 + \phi)^k \psi^{(m-1)}(\mu_i(1 + \phi) + \phi + 2), \end{aligned}$$

$\forall k > 0$  e  $m > 1$ . Em que  $\psi^{(m-1)} = \frac{d^m \log \Gamma(x)}{dx^m}$ . As derivadas dessas funções que desenvolvemos podem ser resolvidas como

$$\begin{aligned} \frac{\partial d_i^{(k,m)}}{\partial \mu_i} &= d_i^{(k+1,m+1)}, & \frac{\partial d_i^{(k,m)}}{\partial \phi} &= k d_i^{(k-1,m)} + \mu_i d_i^{(k,m+1)} + t_i^{(k,m+1)}, \\ \frac{\partial t_i^{(k,m)}}{\partial \mu_i} &= t_i^{(k+1,m+1)}, & \frac{\partial t_i^{(k,m)}}{\partial \phi} &= k t_i^{(k-1,m)} + (1 + \mu_i) t_i^{(k,m+1)}, \\ \frac{\partial r_i^{(k,m)}}{\partial \phi} &= r_i^{(k+1,m+1)}, & \frac{\partial r_i^{(k,m)}}{\partial \mu_i} &= k \mu_i^{k-1} d_i^{(0,m)} + k \left[ (1 + \mu_i)^{k-1} - \mu_i^{k-1} \right] t_i^{(0,m)} + \\ & & & \mu_i^k d_i^{(1,m+1)} + \left[ (1 + \mu_i)^k - \mu_i^k \right] t_i^{(1,m+1)}. \end{aligned}$$

Além disso, define-se

$$\begin{aligned} \mathbf{D}_1 &= \text{diag} \left( a_{1,(1)}, \dots, a_{n,(1)} \right), \\ \mathbf{y}^* &= (y_1^*, \dots, y_n^*)^\top, & \boldsymbol{\mu}^* &= (\mu_1^*, \dots, \mu_n^*)^\top, \\ \mathbf{y}^{\mu_i} &= (y_1^{\mu_i}, \dots, y_n^{\mu_i})^\top, & \boldsymbol{\mu}^* &= (\mu_1^*, \dots, \mu_n^*)^\top. \end{aligned}$$

Logo, os componentes do vetor escore são dadas por:

$$\begin{aligned} \mathbf{U}_\beta(\boldsymbol{\theta}) &= (1 + \phi) \mathbf{X}^\top \mathbf{D}_1 (\mathbf{y}^* - \boldsymbol{\mu}^*), \\ \mathbf{U}_\phi(\boldsymbol{\theta}) &= \mathbf{1}^\top (\mathbf{y}^* - \boldsymbol{\mu}^*) b_{(1)}. \end{aligned}$$

Além disso, define-se as matrizes

$$\begin{aligned} \mathbf{E}_3 &= \text{diag} \left( d_1^{(2,2)} a_{1,(1)}^2, \dots, d_n^{(2,2)} a_{n,(1)}^2 \right), \\ \mathbf{D}_3 &= \mathbf{E}_3 + \text{diag} \left( (1 + \phi)(y_1^* - \mu_1^*) a_{1,(2)}, \dots, (1 + \phi)(y_n^* - \mu_n^*) a_{n,(2)} \right), \\ \mathbf{E}_4 &= \text{diag} \left( r_1^{(2,2)} b_{(1)}^2, \dots, r_n^{(2,2)} b_{(1)}^2 \right), \\ \mathbf{D}_4 &= \mathbf{E}_4 + \text{diag} \left( (y_1^* - \mu_1^*) b_{(2)}, \dots, (y_n^* - \mu_n^*) b_{(2)} \right), \\ \mathbf{E}_5 &= \text{diag} \left[ \left( \mu_1 d_1^{(1,2)} + t_1^{(1,2)} \right) a_{1,(1)} b_{(1)}, \dots, \left( \mu_n d_n^{(1,2)} + t_n^{(1,2)} \right) a_{n,(1)} b_{(1)} \right], \\ \mathbf{D}_5 &= \mathbf{E}_5 + \text{diag} \left( (y_1^* - \mu_1^*) a_{1,(1)} b_{(1)}, \dots, (y_n^* - \mu_n^*) a_{n,(1)} b_{(1)} \right), \end{aligned}$$

então, as matrizes Hessiana e de informação de Fisher são dadas respectivamente por

$$\mathbf{J} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}_3 \mathbf{X} & \mathbf{X}^\top \mathbf{D}_5 \mathbf{1} \\ \mathbf{1}^\top \mathbf{D}_5 \mathbf{X} & \mathbf{1}^\top \mathbf{D}_4 \mathbf{1} \end{bmatrix} \text{ e } \mathbf{K} = \begin{bmatrix} \mathbf{X}^\top \mathbf{E}_3 \mathbf{X} & \mathbf{X}^\top \mathbf{E}_5 \mathbf{1} \\ \mathbf{1}^\top \mathbf{E}_5 \mathbf{X} & \mathbf{1}^\top \mathbf{E}_4 \mathbf{1} \end{bmatrix}.$$

# APÊNDICE B – Cumulantes

Com as funções presentes no Apêndice A, os cumulantes de até quarta ordem, após uma álgebra exaustiva, são dados por

$$\kappa_{rs} = \sum_{i=1}^n d_i^{(2,2)} a_{i,(1)}^2 x_{ir} x_{is},$$

$$\kappa_{r\phi} = \sum_{i=1}^n \left[ d_i^{(1,2)} \mu_i a_{i,(1)} + t_i^{(1,2)} a_{i,(1)} \right] b_{(1)} x_{ir},$$

$$\kappa_{\phi\phi} = \sum_{i=1}^n r_i^{(2,2)} b_{(1)}^2,$$

$$\kappa_{rst}^{(t)} = \sum_{i=1}^n \left[ d_i^{(3,3)} a_{i,(1)}^3 + 2d_i^{(2,2)} a_{i,(1)} a_{i,(2)} \right] x_{ir} x_{is} x_{it},$$

$$\kappa_{rs}^{(\phi)} = \sum_{i=1}^n \left( d_i^{(2,3)} \mu_i + 2d_i^{(1,2)} + t_i^{(2,3)} \right) b_{(1)} a_{i,(1)}^2 x_{ir} x_{is},$$

$$\kappa_{r\phi}^{(s)} = \sum_{i=1}^n \left[ d_i^{(2,3)} \mu_i a_{i,(1)}^2 + d_i^{(1,2)} (\mu_i a_{i,(2)} + a_{i,(1)}^2) + t_i^{(2,3)} a_{i,(1)}^2 + t_i^{(1,2)} a_{i,(2)} \right] b_{(1)} x_{ir} x_{is},$$

$$\begin{aligned} \kappa_{r\phi}^{(\phi)} = \sum_{i=1}^n & \left( d_i^{(1,3)} \mu_i^2 b_{(1)}^2 + d_i^{(1,2)} \mu_i b_{(2)} + d_i^{(0,2)} \mu_i b_{(1)}^2 + \right. \\ & \left. t_i^{(1,3)} (1 + 2\mu_i) b_{(1)}^2 + t_i^{(1,2)} b_{(2)} + t_i^{(0,2)} b_{(1)}^2 \right) a_{i,(1)} x_{ir}, \end{aligned}$$

$$\kappa_{\phi\phi}^{(r)} = \sum_{i=1}^n \left( d_i^{(1,3)} \mu_i^2 + 2d_i^{(0,2)} \mu_i + t_i^{(1,3)} (1 + 2\mu_i) + 2t_i^{(0,2)} \right) a_{i,(1)} b_{(1)}^2 x_{ir},$$

$$\kappa_{\phi\phi\phi}^{(\phi)} = \sum_{i=1}^n \left[ r_i^{(3,3)} b_{(1)}^3 + 2r_i^{(2,2)} b_{(1)} b_{(2)} \right],$$

$$\kappa_{rst} = \sum_{i=1}^n \left[ d_i^{(3,3)} a_{i,(1)}^3 + 3d_i^{(2,2)} a_{i,(1)} a_{i,(2)} \right] x_{ir} x_{is} x_{it},$$

$$\kappa_{rs\phi} = \sum_{i=1}^n \left[ d_i^{(2,3)} \mu_i a_{i,(1)}^2 + d_i^{(1,2)} (\mu_i a_{i,(2)} + 2a_{i,(1)}^2) + t_i^{(2,3)} a_{i,(1)}^2 + t_i^{(1,2)} a_{i,(2)} \right] b_{(1)} x_{ir} x_{is},$$

$$\begin{aligned} \kappa_{r\phi\phi} = \sum_{i=1}^n & \left( d_i^{(1,3)} \mu_i^2 b_{(1)}^2 + d_i^{(1,2)} \mu_i b_{(2)} + 2d_i^{(0,2)} \mu_i b_{(1)}^2 + \right. \\ & \left. t_i^{(1,3)} (1 + 2\mu_i) b_{(1)}^2 + t_i^{(1,2)} b_{(2)} + 2t_i^{(0,2)} b_{(1)}^2 \right) a_{i,(1)} x_{ir}, \end{aligned}$$

$$\kappa_{\phi\phi\phi} = \sum_{i=1}^n \left[ r_i^{(3,3)} b_{(1)}^3 + 3r_i^{(2,2)} b_{(1)} b_{(2)} \right],$$

$$\begin{aligned}
 \kappa_{rs}^{(tu)} &= \sum_{i=1}^n \left[ d_i^{(4,4)} a_{i,(1)}^4 + 5d_i^{(3,3)} a_{i,(1)}^2 a_{i,(2)} + 2d_i^{(2,2)} (a_{i,(2)}^2 + a_{i,(1)} a_{i,(3)}) \right] x_{ir} x_{is} x_{it} x_{iu}, \\
 \kappa_{rs}^{(t\phi)} &= \sum_{i=1}^n \left( d_i^{(3,4)} \mu_i a_{i,(1)}^3 + d_i^{(2,3)} a_{i,(1)} (3a_{i,(1)}^2 + 2\mu_i a_{i,(2)}) + 4d_i^{(1,2)} a_{i,(1)} a_{i,(2)} + \right. \\
 &\quad \left. t_i^{(3,4)} a_{i,(1)}^3 + 2t_i^{(2,3)} a_{i,(1)} a_{i,(2)} \right) b_{(1)} x_{ir} x_{is} x_{it}, \\
 \kappa_{rs}^{(\phi\phi)} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 b_{(1)}^2 + d_i^{(2,3)} \mu_i b_{(2)} + 4d_i^{(1,3)} \mu_i b_{(1)}^2 + 2d_i^{(1,2)} b_{(2)} + 2d_i^{(0,2)} b_{(1)}^2 + \right. \\
 &\quad \left. t_i^{(2,4)} (1 + 2\mu_i) b_{(1)}^2 + t_i^{(2,3)} b_{(2)} + 4t_i^{(1,3)} b_{(1)}^2 \right) a_{i,(1)}^2 x_{ir} x_{is}, \\
 \kappa_{r\phi}^{(st)} &= \sum_{i=1}^n \left[ d_i^{(3,4)} \mu_i a_{i,(1)}^3 + d_i^{(2,3)} (3\mu_i a_{i,(1)} a_{i,(2)} + 2a_{i,(1)}^3) + d_i^{(1,2)} (\mu_i a_{i,(3)} + 3a_{i,(1)} a_{i,(2)}) + \right. \\
 &\quad \left. t_i^{(3,4)} a_{i,(1)}^3 + 3t_i^{(2,3)} a_{i,(1)} a_{i,(2)} + t_i^{(1,2)} a_{i,(3)} \right] b_{(1)} x_{ir} x_{is} x_{it}, \\
 \kappa_{r\phi}^{(s\phi)} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 a_{i,(1)}^2 b_{(1)}^2 + d_i^{(2,3)} \mu_i a_{i,(1)}^2 b_{(2)} + d_i^{(1,3)} \mu_i b_{(1)}^2 (3a_{i,(1)}^2 + \mu_i a_{i,(2)}) + \right. \\
 &\quad d_i^{(1,2)} b_{(2)} (\mu_i a_{i,(2)} + a_{i,(1)}^2) + d_i^{(0,2)} b_{(1)}^2 (\mu_i a_{i,(2)} + a_{i,(1)}^2) + \\
 &\quad t_i^{(2,4)} (1 + 2\mu_i) a_{i,(1)}^2 b_{(1)}^2 + t_i^{(2,3)} a_{i,(1)}^2 b_{(2)} + t_i^{(1,3)} b_{(1)}^2 (3a_{i,(1)}^2 + (1 + 2\mu_i) a_{i,(2)}) + \\
 &\quad \left. t_i^{(1,2)} a_{i,(2)} b_{(2)} + t_i^{(0,2)} a_{i,(2)} b_{(1)}^2 \right) x_{ir} x_{is}, \\
 \kappa_{r\phi}^{(\phi\phi)} &= \sum_{i=1}^n \left( d_i^{(1,4)} b_{(1)}^3 \mu_i^3 + 3d_i^{(1,3)} b_{(1)} b_{(2)} \mu_i^2 + d_i^{(1,2)} b_{(3)} \mu_i + \right. \\
 &\quad 2d_i^{(0,3)} b_{(1)}^3 \mu_i^2 + 3d_i^{(0,2)} b_{(1)} b_{(2)} \mu_i + \\
 &\quad t_i^{(1,4)} b_{(1)}^3 (1 + 3\mu_i + 3\mu_i^2) + t_i^{(1,3)} b_{(1)} b_{(2)} (3 + 6\mu_i) + t_i^{(1,2)} b_{(3)} + \\
 &\quad \left. t_i^{(0,3)} b_{(1)}^3 (2 + 4\mu_i) + 3t_i^{(0,2)} b_{(1)} b_{(2)} \right) a_{i,(1)} x_{ir}, \\
 \kappa_{\phi\phi}^{(rs)} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 a_{i,(1)}^2 + d_i^{(1,3)} \mu_i (4a_{i,(1)}^2 + \mu_i a_{i,(2)}) + 2d_i^{(0,2)} (a_{i,(1)}^2 + \mu_i a_{i,(2)}) + \right. \\
 &\quad \left. t_i^{(2,4)} (1 + 2\mu_i) a_{i,(1)}^2 + t_i^{(1,3)} (4a_{i,(1)}^2 + (1 + 2\mu_i) a_{i,(2)}) + 2t_i^{(0,2)} a_{i,(2)} \right) b_{(1)}^2 x_{ir} x_{is}, \\
 \kappa_{\phi\phi}^{(r\phi)} &= \sum_{i=1}^n \left( d_i^{(1,4)} \mu_i^3 b_{(1)}^3 + 2d_i^{(1,3)} \mu_i^2 b_{(1)} b_{(2)} + 3d_i^{(0,3)} \mu_i^2 b_{(1)}^3 + 4d_i^{(0,2)} \mu_i b_{(1)} b_{(2)} + \right. \\
 &\quad t_i^{(1,4)} (1 + 3\mu_i + 3\mu_i^2) b_{(1)}^3 + 2t_i^{(1,3)} (1 + 2\mu_i) b_{(1)} b_{(2)} + 3t_i^{(0,3)} (1 + 2\mu_i) b_{(1)}^3 + \\
 &\quad \left. 4t_i^{(0,2)} b_{(1)} b_{(2)} \right) a_{i,(1)} x_{ir}, \\
 \kappa_{\phi\phi}^{(\phi\phi)} &= \sum_{i=1}^n \left[ r_i^{(4,4)} b_{(1)}^4 + 5r_i^{(3,3)} b_{(1)}^2 b_{(2)} + 2r_i^{(2,2)} (b_{(2)}^2 + b_{(1)} b_{(3)}) \right],
 \end{aligned}$$

$$\begin{aligned}
 \kappa_{rst}^{(u)} &= \sum_{i=1}^n \left[ d_i^{(4,4)} a_{i,(1)}^4 + 6d_i^{(3,3)} a_{i,(1)}^2 a_{i,(2)} + 3d_i^{(2,2)} (a_{i,(2)}^2 + a_{i,(1)} a_{i,(3)}) \right] x_{ir} x_{is} x_{it} x_{iu} \\
 \kappa_{rst}^{(\phi)} &= \sum_{i=1}^n \left( d_i^{(3,4)} \mu_i a_{i,(1)}^3 + 3d_i^{(2,3)} a_{i,(1)} (a_{i,(1)}^2 + \mu_i a_{i,(2)}) + 6d_i^{(1,2)} a_{i,(1)} a_{i,(2)} + \right. \\
 &\quad \left. t_i^{(3,4)} a_{i,(1)}^3 + 3t_i^{(2,3)} a_{i,(1)} a_{i,(2)} \right) b_{(1)} x_{ir} x_{is} x_{it}, \\
 \kappa_{rs\phi}^{(t)} &= \sum_{i=1}^n \left[ d_i^{(3,4)} \mu_i a_{i,(1)}^3 + 3d_i^{(2,3)} (\mu_i a_{i,(1)} a_{i,(2)} + a_{i,(1)}^3) + d_i^{(1,2)} (\mu_i a_{i,(3)} + 5a_{i,(1)} a_{i,(2)}) + \right. \\
 &\quad \left. t_i^{(3,4)} a_{i,(1)}^3 + 3t_i^{(2,3)} a_{i,(1)} a_{i,(2)} + t_i^{(1,2)} a_{i,(3)} \right] b_{(1)} x_{ir} x_{is} x_{it}, \\
 \kappa_{rs\phi}^{(\phi)} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 a_{i,(1)}^2 b_{(1)}^2 + d_i^{(2,3)} \mu_i a_{i,(1)}^2 b_{(2)} + d_i^{(1,3)} \mu_i b_{(1)}^2 (4a_{i,(1)}^2 + \mu_i a_{i,(2)}) + \right. \\
 &\quad d_i^{(1,2)} b_{(2)} (\mu_i a_{i,(2)} + 2a_{i,(1)}^2) + d_i^{(0,2)} b_{(1)}^2 (\mu_i a_{i,(2)} + 2a_{i,(1)}^2) + \\
 &\quad t_i^{(2,4)} (1 + 2\mu_i) a_{i,(1)}^2 b_{(1)}^2 + t_i^{(2,3)} a_{i,(1)}^2 b_{(2)} + t_i^{(1,3)} b_{(1)}^2 (4a_{i,(1)}^2 + (1 + 2\mu_i) a_{i,(2)}) + \\
 &\quad \left. t_i^{(1,2)} a_{i,(2)} b_{(2)} + t_i^{(0,2)} a_{i,(2)} b_{(1)}^2 \right) x_{ir} x_{is}, \\
 \kappa_{r\phi\phi}^{(s)} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 a_{i,(1)}^2 b_{(1)}^2 + d_i^{(2,3)} \mu_i a_{i,(1)}^2 b_{(2)} + d_i^{(1,3)} \mu_i b_{(1)}^2 (4a_{i,(1)}^2 + \mu_i a_{i,(2)}) + \right. \\
 &\quad d_i^{(1,2)} b_{(2)} (\mu_i a_{i,(2)} + a_{i,(1)}^2) + 2d_i^{(0,2)} b_{(1)}^2 (\mu_i a_{i,(2)} + a_{i,(1)}^2) + \\
 &\quad t_i^{(2,4)} (1 + 2\mu_i) a_{i,(1)}^2 b_{(1)}^2 + t_i^{(2,3)} a_{i,(1)}^2 b_{(2)} + t_i^{(1,3)} b_{(1)}^2 (4a_{i,(1)}^2 + (1 + 2\mu_i) a_{i,(2)}) + \\
 &\quad \left. t_i^{(1,2)} a_{i,(2)} b_{(2)} + 2t_i^{(0,2)} a_{i,(2)} b_{(1)}^2 \right) x_{ir} x_{is}, \\
 \kappa_{r\phi\phi}^{(\phi)} &= \sum_{i=1}^n \left( d_i^{(1,4)} b_{(1)}^3 \mu_i^3 + 3d_i^{(1,3)} b_{(1)} b_{(2)} \mu_i^2 + d_i^{(1,2)} b_{(3)} \mu_i + \right. \\
 &\quad 3d_i^{(0,3)} b_{(1)}^3 \mu_i^2 + 5d_i^{(0,2)} b_{(1)} b_{(2)} \mu_i + \\
 &\quad t_i^{(1,4)} b_{(1)}^3 (1 + 3\mu_i + 3\mu_i^2) + t_i^{(1,3)} b_{(1)} b_{(2)} (3 + 6\mu_i) + t_i^{(1,2)} b_{(3)} + \\
 &\quad \left. t_i^{(0,3)} b_{(1)}^3 (3 + 6\mu_i) + 5t_i^{(0,2)} b_{(1)} b_{(2)} \right) a_{i,(1)} x_{ir}, \\
 \kappa_{\phi\phi\phi}^{(r)} &= \sum_{i=1}^n \left( d_i^{(1,4)} \mu_i^3 b_{(1)}^3 + 3d_i^{(1,3)} \mu_i^2 b_{(1)} b_{(2)} + 3d_i^{(0,3)} \mu_i^2 b_{(1)}^3 + 6d_i^{(0,2)} \mu_i b_{(1)} b_{(2)} + \right. \\
 &\quad t_i^{(1,4)} (1 + 3\mu_i + 3\mu_i^2) b_{(1)}^3 + 3t_i^{(1,3)} (1 + 2\mu_i) b_{(1)} b_{(2)} + 3t_i^{(0,3)} (1 + 2\mu_i) b_{(1)}^3 + \\
 &\quad \left. 6t_i^{(0,2)} b_{(1)} b_{(2)} \right) a_{i,(1)} x_{ir}, \\
 \kappa_{\phi\phi\phi}^{(\phi)} &= \sum_{i=1}^n \left[ r_i^{(4,4)} b_{(1)}^4 + 6r_i^{(3,3)} b_{(1)}^2 b_{(2)} + 3r_i^{(2,2)} (b_{(2)}^2 + b_{(1)} b_{(3)}) \right],
 \end{aligned}$$

$$\begin{aligned}
 \kappa_{rstu} &= \sum_{i=1}^n \left[ d_i^{(4,4)} a_{i,(1)}^4 + 6d_i^{(3,3)} a_{i,(1)}^2 a_{i,(2)} + d_i^{(2,2)} (3a_{i,(2)}^2 + 4a_{i,(1)} a_{i,(3)}) \right] x_{ir} x_{is} x_{it} x_{iu}, \\
 \kappa_{rst\phi} &= \sum_{i=1}^n \left[ d_i^{(3,4)} \mu_i a_{i,(1)}^3 + 3d_i^{(2,3)} (\mu_i a_{i,(1)} a_{i,(2)} + a_{i,(1)}^3) + d_i^{(1,2)} (\mu_i a_{i,(3)} + 6a_{i,(1)} a_{i,(2)}) + \right. \\
 &\quad \left. t_i^{(3,4)} a_{i,(1)}^3 + 3t_i^{(2,3)} a_{i,(1)} a_{i,(2)} + t_i^{(1,2)} a_{i,(3)} \right] b_{(1)} x_{ir} x_{is} x_{it} \\
 \kappa_{rs\phi\phi} &= \sum_{i=1}^n \left( d_i^{(2,4)} \mu_i^2 a_{i,(1)}^2 b_{(1)}^2 + d_i^{(2,3)} \mu_i a_{i,(1)}^2 b_{(2)} + d_i^{(1,3)} \mu_i b_{(1)}^2 (4a_{i,(1)}^2 + \mu_i a_{i,(2)}) + \right. \\
 &\quad d_i^{(1,2)} b_{(2)} (\mu_i a_{i,(2)} + 2a_{i,(1)}^2) + 2d_i^{(0,2)} b_{(1)}^2 (\mu_i a_{i,(2)} + a_{i,(1)}^2) + \\
 &\quad t_i^{(2,4)} (1 + 2\mu_i) a_{i,(1)}^2 b_{(1)}^2 + t_i^{(2,3)} a_{i,(1)}^2 b_{(2)} + t_i^{(1,3)} b_{(1)}^2 (4a_{i,(1)}^2 + (1 + 2\mu_i) a_{i,(2)}) + \\
 &\quad \left. t_i^{(1,2)} a_{i,(2)} b_{(2)} + 2t_i^{(0,2)} a_{i,(2)} b_{(1)}^2 \right) x_{ir} x_{is}, \\
 \kappa_{r\phi\phi\phi} &= \sum_{i=1}^n \left( d_i^{(1,4)} b_{(1)}^3 \mu_i^3 + 3d_i^{(1,3)} b_{(1)} b_{(2)} \mu_i^2 + d_i^{(1,2)} b_{(3)} \mu_i + \right. \\
 &\quad 3d_i^{(0,3)} b_{(1)}^3 \mu_i^2 + 6d_i^{(0,2)} b_{(1)} b_{(2)} \mu_i + \\
 &\quad t_i^{(1,4)} b_{(1)}^3 (1 + 3\mu_i + 3\mu_i^2) + t_i^{(1,3)} b_{(1)} b_{(2)} (3 + 6\mu_i) + t_i^{(1,2)} b_{(3)} + \\
 &\quad \left. t_i^{(0,3)} b_{(1)}^3 (3 + 6\mu_i) + 6t_i^{(0,2)} b_{(1)} b_{(2)} \right) a_{i,(1)} x_{ir}, \\
 \kappa_{\phi\phi\phi\phi} &= \sum_{i=1}^n \left[ r_i^{(4,4)} b_{(1)}^4 + 6r_i^{(3,3)} b_{(1)}^2 b_{(2)} + r_i^{(2,2)} (3b_{(2)}^2 + 4b_{(1)} b_{(3)}) \right],
 \end{aligned}$$

$$\begin{aligned}
 \kappa_{rs,t} &= - \sum_{i=1}^n \left[ d_i^{(2,2)} a_{i,(1)} a_{i,(2)} x_{ir} x_{is} x_{it} \right], \\
 \kappa_{rs,\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,2)} \mu_i + t_i^{(1,2)} \right) a_{i,(2)} b_{i,(1)} x_{ir} x_{is} \right], \\
 \kappa_{r\phi,s} &= - \sum_{i=1}^n \left[ d_i^{(1,2)} a_{i,(1)}^2 b_{i,(1)} x_{ir} x_{is} \right], \\
 \kappa_{r\phi,\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(0,2)} \mu_i + t_i^{(0,2)} \right) a_{i,(1)} b_{i,(1)}^2 x_{ir} \right], \\
 \kappa_{\phi\phi,r} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,2)} \mu_i + t_i^{(1,2)} \right) a_{i,(1)} b_{i,(2)} x_{ir} \right], \\
 \kappa_{\phi\phi,\phi} &= - \sum_{i=1}^n \left[ r_i^{(2,2)} b_{(1)} b_{(2)} \right],
 \end{aligned}$$

$$\begin{aligned}
 \kappa_{rs,tu} &= - \sum_{i=1}^n \left[ d_i^{(2,2)} a_{i,(2)}^2 x_{ir} x_{is} x_{it} x_{iu} \right] + \kappa_{rs} \kappa_{tu}, \\
 \kappa_{rs,t\phi} &= - \sum_{i=1}^n \left[ d_i^{(1,2)} a_{i,(1)} a_{i,(2)} b_{(1)} x_{ir} x_{is} x_{it} \right] + \kappa_{rs} \kappa_{t\phi}, \\
 \kappa_{r\phi,s\phi} &= - \sum_{i=1}^n \left[ d_i^{(0,2)} a_{i,(1)}^2 b_{(1)}^2 x_{ir} x_{is} \right] + \kappa_{r\phi} \kappa_{s\phi}, \\
 \kappa_{rs,\phi\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,2)} \mu_i + t_i^{(1,2)} \right) a_{i,(2)} b_{(2)} x_{ir} x_{is} \right] + \kappa_{rs} \kappa_{\phi\phi}, \\
 \kappa_{r\phi,\phi\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(0,2)} \mu_i + t_i^{(0,2)} \right) a_{i,(1)} b_{(1)} b_{(2)} x_{ir} \right] + \kappa_{r\phi} \kappa_{\phi\phi}, \\
 \kappa_{\phi\phi,\phi\phi} &= - \sum_{i=1}^n \left[ r_i^{(2,2)} b_{(2)}^2 \right] + \kappa_{\phi\phi}^2,
 \end{aligned}$$

$$\begin{aligned}
 \kappa_{rst,u} &= - \sum_{i=1}^n \left[ d_i^{(2,2)} a_{i,(1)} a_{i,(3)} x_{ir} x_{is} x_{it} x_{iu} \right], \\
 \kappa_{rst,\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,2)} \mu_i + t_i^{(1,2)} \right) a_{i,(3)} b_{(1)} x_{ir} x_{is} x_{it} \right], \\
 \kappa_{rs\phi,t} &= - \sum_{i=1}^n \left[ d_i^{(1,2)} a_{i,(1)} a_{i,(2)} b_{(1)} x_{ir} x_{is} x_{it} \right], \\
 \kappa_{rs\phi,\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(0,2)} \mu_i + t_i^{(0,2)} \right) a_{i,(2)} b_{(1)}^2 x_{ir} x_{is} \right], \\
 \kappa_{r\phi\phi,s} &= - \sum_{i=1}^n \left[ d_i^{(1,2)} a_{i,(1)}^2 b_{(2)} x_{ir} x_{is} \right], \\
 \kappa_{r\phi\phi,\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(0,2)} \mu_i + t_i^{(0,2)} \right) a_{i,(1)} b_{(1)} b_{(2)} x_{ir} \right], \\
 \kappa_{\phi\phi\phi,r} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,2)} \mu_i + t_i^{(1,2)} \right) a_{i,(1)} b_{(3)} x_{ir} \right], \\
 \kappa_{\phi\phi\phi,\phi} &= - \sum_{i=1}^n \left[ r_i^{(2,2)} b_{(1)} b_{(3)} \right],
 \end{aligned}$$

$$\begin{aligned}
\kappa_{r,s,tu} &= - \sum_{i=1}^n \left[ d_i^{(3,3)} a_{i,(1)}^2 a_{i,(2)} x_{ir} x_{is} x_{it} x_{iu} \right] - \kappa_{rs} \kappa_{tu}, \\
\kappa_{r,\phi,st} &= - \sum_{i=1}^n \left[ \left( d_i^{(2,3)} \mu_i + t_i^{(2,3)} \right) a_{i,(1)} a_{(2)} b_{(1)} x_{ir} x_{is} x_{it} \right] - \kappa_{r\phi} \kappa_{st}, \\
\kappa_{r,s,t\phi} &= - \sum_{i=1}^n \left[ d_i^{(2,3)} a_{i,(1)}^3 b_{(1)} x_{ir} x_{is} x_{it} \right] - \kappa_{rs} \kappa_{t\phi}, \\
\kappa_{\phi,\phi,rs} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,3)} \mu_i^2 + t_i^{(1,3)} (1 + 2\mu_i) \right) a_{i,(2)} b_{(1)}^2 x_{ir} x_{is} \right] - \kappa_{rs} \kappa_{\phi\phi}, \\
\kappa_{r,\phi,s\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,3)} \mu_i + t_i^{(1,3)} \right) a_{i,(1)}^2 b_{(1)}^2 x_{ir} x_{is} \right] - \kappa_{r\phi} \kappa_{s\phi}, \\
\kappa_{r,s,\phi\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(2,3)} \mu_i + t_i^{(2,3)} \right) a_{i,(1)}^2 b_{(2)} x_{ir} x_{is} \right] - \kappa_{rs} \kappa_{\phi\phi}, \\
\kappa_{\phi,\phi,r\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(0,3)} \mu_i^2 + t_i^{(0,3)} (1 + 2\mu_i) \right) a_{i,(1)} b_{(1)}^3 x_{ir} \right] - \kappa_{r\phi} \kappa_{\phi\phi}, \\
\kappa_{r,\phi,\phi\phi} &= - \sum_{i=1}^n \left[ \left( d_i^{(1,3)} \mu_i^2 + t_i^{(1,3)} (1 + 2\mu_i) \right) a_{i,(1)} b_{(1)} b_{(2)} x_{ir} \right] - \kappa_{r\phi} \kappa_{\phi\phi}, \\
\kappa_{\phi,\phi,\phi\phi} &= - \sum_{i=1}^n \left[ r_i^{(3,3)} b_{(1)}^2 b_{(2)} \right] - \kappa_{\phi\phi}^2.
\end{aligned}$$

# APÊNDICE C – Ajuste de Skovgaard

Utilizando as funções do Apêndice A, desenvolvemos as seguintes quantidades:

$$\begin{aligned}
E(y_i^*) &= \mu_i^*, & Var(y_i^*) &= -d_i^{(0,2)}, \\
E(y_i^*) &= \mu_i^*, & Var(y_i^*) &= -r_i^{(2,2)}, \\
Cov(y_i^*, y_i^*) &= -\left(d_i^{(0,2)} \mu_i + t_i^{(0,2)}\right), & Cov\left(\sum_{i=1}^n y_i^*, \sum_{i=1}^n y_i^*\right) &= \sum_{i=1}^n Var(y_i^*), \\
Cov\left(\sum_{i=1}^n y_i^*, \sum_{i=1}^n y_i^*\right) &= \sum_{i=1}^n Var(y_i^*), & Cov\left(\sum_{i=1}^n y_i^*, \sum_{i=1}^n y_i^*\right) &= \sum_{i=1}^n Cov(y_i^*, y_i^*), \\
Cov(y_i^*, \log(1 + y_i)) &= t_i^{(0,2)}, & Cov\left(\sum_{i=1}^n y_i^* \sum_{i=1}^n \log(1 + y_i)\right) &= \sum_{i=1}^n Cov(y_i^*, \log(1 + y_i)), \\
Var(\log(1 + y_i)) &= d_i^{(0,2)} - r_i^{(0,2)}, & Var\left(\sum_{i=1}^n \log(1 + y_i)\right) &= \sum_{i=1}^n Var(\log(1 + y_i)).
\end{aligned}$$

Além disso, defina as quantidades

$$\begin{aligned}
\hat{d}_i^{(k,m)} &= d(k, m, \hat{\mu}_i, \hat{\phi}), & \tilde{d}_i^{(k,m)} &= d(k, m, \tilde{\mu}_i, \tilde{\phi}), \\
\hat{t}_i^{(k,m)} &= t(k, m, \hat{\mu}_i, \hat{\phi}), & \tilde{t}_i^{(k,m)} &= t(k, m, \tilde{\mu}_i, \tilde{\phi}), \\
\hat{r}_i^{(k,m)} &= r(k, m, \hat{\mu}_i, \hat{\phi}), & \tilde{r}_i^{(k,m)} &= r(k, m, \tilde{\mu}_i, \tilde{\phi}), \\
\hat{a}_{i,(1)} &= \left. \frac{\partial \mu_i}{\partial \eta_{1i}} \right|_{\mu_i = \hat{\mu}_i}, & \tilde{a}_{i,(1)} &= \left. \frac{\partial \mu_i}{\partial \eta_{1i}} \right|_{\mu_i = \tilde{\mu}_i}, \\
\hat{b}_{(1)} &= \left. \frac{\partial \phi}{\partial \eta_2} \right|_{\phi = \hat{\phi}}, & \tilde{b}_{(1)} &= \left. \frac{\partial \phi}{\partial \eta_2} \right|_{\phi = \tilde{\phi}},
\end{aligned}$$

e as matrizes

$$\begin{aligned}
\hat{\mathbf{D}}_1 &= \text{diag}(\hat{a}_{1,(1)}, \dots, \hat{a}_{n,(1)}), & \tilde{\mathbf{D}}_1 &= \text{diag}(\tilde{a}_{1,(1)}, \dots, \tilde{a}_{n,(1)}), \\
\hat{\mathbf{M}} &= \text{diag}(\hat{\mu}_1, \dots, \hat{\mu}_n), & \tilde{\mathbf{M}} &= \text{diag}(\tilde{\mu}_1, \dots, \tilde{\mu}_n), \\
\mathbf{V}_1 &= -\text{diag}(\hat{d}_1^{(0,2)}, \dots, \hat{d}_n^{(0,2)}), & \mathbf{C} &= -\text{diag}(\tilde{t}_1^{(0,2)}, \dots, \tilde{t}_n^{(0,2)}), \\
\mathbf{V}_2 &= -\text{diag}(\hat{r}_1^{(0,2)}, \dots, \hat{r}_n^{(0,2)}) - \mathbf{V}_1 + \mathbf{C}.
\end{aligned}$$

Seja  $U_{\beta_i}(\boldsymbol{\theta})$  o  $i$ -ésimo elemento do sub-vetor escore  $\mathbf{U}_{\beta}(\boldsymbol{\theta})$ . Então

$$\begin{aligned}
 E_{\theta} [U_{\beta_r}(\boldsymbol{\theta})U_{\beta_s}(\boldsymbol{\theta}_2)]|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} &= Cov(U_{\beta_r}(\boldsymbol{\theta}), U_{\beta_s}(\boldsymbol{\theta}_2))|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} \\
 &= Cov\left(\sum_{i=1}^n(1+\phi)y_i^*a_{i,(1)}x_{ir}\Big|_{\theta=\hat{\theta}}, \sum_{i=1}^n(1+\phi)y_i^*a_{i,(1)}x_{is}\Big|_{\theta_2=\tilde{\theta}}\right) \\
 &= \sum_{i=1}^n(1+\hat{\phi})(1+\tilde{\phi})Cov(y_i^*, y_i^*)|_{\theta=\hat{\theta}}\hat{a}_{i,(1)}\tilde{a}_{i,(1)}x_{ir}x_{is} \\
 &= -(1+\hat{\phi})(1+\tilde{\phi})\sum_{i=1}^n\tilde{d}_i^{(0,2)}\hat{a}_{i,(1)}\tilde{a}_{i,(1)}x_{ir}x_{is},
 \end{aligned}$$

$$\begin{aligned}
 E_{\theta} [U_{\beta_r}(\boldsymbol{\theta})U_{\phi}(\boldsymbol{\theta}_2)]|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} &= Cov(U_{\beta_r}(\boldsymbol{\theta}), U_{\phi}(\boldsymbol{\theta}_2))|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} \\
 &= Cov\left(\sum_{i=1}^n(1+\phi)y_i^*a_{i,(1)}x_{ir}\Big|_{\theta=\hat{\theta}}, \sum_{i=1}^ny_i^*b_{(1)}\Big|_{\theta_2=\tilde{\theta}}\right) \\
 &= \sum_{i=1}^n(1+\hat{\phi})Cov(y_i^*|_{\theta=\hat{\theta}}, y_i^*|_{\theta_2=\tilde{\theta}})\hat{a}_{i,(1)}\tilde{b}_{(1)}x_{ir} \\
 &= -(1+\hat{\phi})\sum_{i=1}^n(\tilde{d}_i^{(0,2)}\tilde{\mu}_i + \tilde{t}_i^{(0,2)})\hat{a}_{i,(1)}\tilde{b}_{(1)}x_{ir},
 \end{aligned}$$

$$\begin{aligned}
 E_{\theta} [U_{\phi}(\boldsymbol{\theta})U_{\beta_r}(\boldsymbol{\theta}_2)]|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} &= Cov(U_{\phi}(\boldsymbol{\theta}), U_{\beta_r}(\boldsymbol{\theta}_2))|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} \\
 &= Cov\left(\sum_{i=1}^ny_i^*b_{(1)}\Big|_{\theta=\hat{\theta}}, \sum_{i=1}^n(1+\phi)y_i^*a_{i,(1)}x_{ir}\Big|_{\theta_2=\tilde{\theta}}\right) \\
 &= \sum_{i=1}^n(1+\tilde{\phi})Cov(y_i^*|_{\theta=\hat{\theta}}, y_i^*|_{\theta_2=\tilde{\theta}})\tilde{a}_{i,(1)}\hat{b}_{(1)}x_{ir} \\
 &= -(1+\tilde{\phi})\sum_{i=1}^n(\hat{d}_i^{(0,2)}\hat{\mu}_i + \hat{t}_i^{(0,2)})\tilde{a}_{i,(1)}\hat{b}_{(1)}x_{ir},
 \end{aligned}$$

$$\begin{aligned}
 E_{\theta} [U_{\phi}(\boldsymbol{\theta})U_{\phi}(\boldsymbol{\theta}_2)]|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} &= Cov(U_{\phi}(\boldsymbol{\theta}), U_{\phi}(\boldsymbol{\theta}_2))|_{\theta=\hat{\theta},\theta_2=\tilde{\theta}} \\
 &= Cov\left(\sum_{i=1}^ny_i^*b_{(1)}\Big|_{\theta=\hat{\theta}}, \sum_{i=1}^ny_i^*b_{(1)}\Big|_{\theta_2=\tilde{\theta}}\right) \\
 &= \sum_{i=1}^nCov(y_i^*|_{\theta=\hat{\theta}}, y_i^*|_{\theta_2=\tilde{\theta}})\hat{b}_{(1)}\tilde{b}_{(1)} \\
 &= -\sum_{i=1}^n[\tilde{d}_i^{(0,2)}(\hat{\mu}_i\tilde{\mu}_i - 1) + \tilde{t}_i^{(0,2)}(\hat{\mu}_i + \tilde{\mu}_i + 1) + \tilde{r}_i^{(0,2)}]\hat{b}_{(1)}\tilde{b}_{(1)}.
 \end{aligned}$$

Matricialmente,

$$\begin{aligned}
 E_{\boldsymbol{\theta}} \left[ \mathbf{U}_{\beta}(\boldsymbol{\theta}) \mathbf{U}_{\beta}(\boldsymbol{\theta}_2)^{\top} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= (1 + \hat{\phi})(1 + \tilde{\phi}) \mathbf{X}^{\top} \widehat{\mathbf{D}}_1 \mathbf{V}_1 \widetilde{\mathbf{D}}_1 \mathbf{X}, \\
 E_{\boldsymbol{\theta}} \left[ \mathbf{U}_{\beta}(\boldsymbol{\theta}) \mathbf{U}_{\phi}(\boldsymbol{\theta}_2)^{\top} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= (1 + \hat{\phi}) \mathbf{X}^{\top} \widehat{\mathbf{D}}_1 (\mathbf{V}_1 \widetilde{\mathbf{M}} + \mathbf{C}) \mathbf{1} \tilde{b}_{(1)}, \\
 E_{\boldsymbol{\theta}} \left[ \mathbf{U}_{\phi}(\boldsymbol{\theta}) \mathbf{U}_{\beta}(\boldsymbol{\theta}_2)^{\top} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= (1 + \tilde{\phi}) \mathbf{1}^{\top} (\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) \widetilde{\mathbf{D}}_1 \mathbf{X} \hat{b}_{(1)}, \\
 E_{\boldsymbol{\theta}} \left[ \mathbf{U}_{\phi}(\boldsymbol{\theta}) \mathbf{U}_{\phi}(\boldsymbol{\theta}_2)^{\top} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= \mathbf{1}^{\top} \left[ \mathbf{V}_1 \widehat{\mathbf{M}} \widetilde{\mathbf{M}} + \mathbf{C}(\widehat{\mathbf{M}} + \widetilde{\mathbf{M}}) + \mathbf{V}_2 \right] \mathbf{1} \hat{b}_{(1)} \tilde{b}_{(1)}.
 \end{aligned}$$

Logo,

$$\tilde{\mathbf{Y}} = \begin{bmatrix} (1 + \hat{\phi})(1 + \tilde{\phi}) \mathbf{X}^{\top} \widehat{\mathbf{D}}_1 \mathbf{V}_1 \widetilde{\mathbf{D}}_1 \mathbf{X} & (1 + \hat{\phi}) \mathbf{X}^{\top} \widehat{\mathbf{D}}_1 (\mathbf{V}_1 \widetilde{\mathbf{M}} + \mathbf{C}) \mathbf{1} \tilde{b}_{(1)} \\ (1 + \tilde{\phi}) \mathbf{1}^{\top} (\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) \widetilde{\mathbf{D}}_1 \mathbf{X} \hat{b}_{(1)} & \mathbf{1}^{\top} \left[ \mathbf{V}_1 \widehat{\mathbf{M}} \widetilde{\mathbf{M}} + \mathbf{C}(\widehat{\mathbf{M}} + \widetilde{\mathbf{M}}) + \mathbf{V}_2 \right] \mathbf{1} \hat{b}_{(1)} \tilde{b}_{(1)} \end{bmatrix}.$$

Além disso,

$$\begin{aligned}
 E_{\boldsymbol{\theta}} [U_{\beta_r}(\boldsymbol{\theta})(\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2))] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= Cov(U_{\beta_r}(\boldsymbol{\theta}), \ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2)) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} \\
 &= Cov \left( \sum_{i=1}^n (1 + \hat{\phi}) y_i^* \hat{a}_{i,(1)} x_{ir}, \right. \\
 &\quad \left. \sum_{i=1}^n \left[ y_i^* (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - \log(1 + y_i)(\hat{\phi} - \tilde{\phi}) \right] \right) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
 &= (1 + \hat{\phi}) \left( \sum_{i=1}^n Var(y_i^*) (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - \right. \\
 &\quad \left. \sum_{i=1}^n Cov(y_i^*, \log(1 + y_i)) (\hat{\phi} - \tilde{\phi}) \right) \hat{a}_{i,(1)} x_{ir} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
 &= -(1 + \hat{\phi}) \sum_{i=1}^n \left( d_i^{(0,2)} (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) + \right. \\
 &\quad \left. t_i^{(0,2)} (\hat{\phi} - \tilde{\phi}) \right) \hat{a}_{i,(1)} x_{ir},
 \end{aligned}$$

$$\begin{aligned}
 E_{\boldsymbol{\theta}} [U_{\phi}(\boldsymbol{\theta})(\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2))] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= Cov(U_{\phi}(\boldsymbol{\theta}), \ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2)) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} \\
 &= Cov\left(\sum_{i=1}^n y_i^* \hat{b}_{(1)}, \sum_{i=1}^n \left[ y_i^* (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - \log(1 + y_i)(\hat{\phi} - \tilde{\phi}) \right]\right) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
 &= Cov\left(\sum_{i=1}^n (\hat{\mu}_i y_i^* - \log(1 + y_i)) \hat{b}_{(1)}, \sum_{i=1}^n \left[ y_i^* (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - \log(1 + y_i)(\hat{\phi} - \tilde{\phi}) \right]\right) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
 &= \sum_{i=1}^n \left[ Var(y_i^*) \hat{\mu}_i (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - Cov(y_i^*, \log(1 + y_i)) (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi}) + \hat{\mu}_i(\hat{\phi} - \tilde{\phi})) + Var(\log(1 + y_i)) (\hat{\phi} - \tilde{\phi}) \right] \hat{b}_{(1)} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\
 &= \sum_{i=1}^n \left[ (-\hat{d}_i^{(0,2)}) \hat{\mu}_i (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) - \hat{t}_i^{(0,2)} (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi}) + \hat{\mu}_i(\hat{\phi} - \tilde{\phi})) + (\hat{d}_i^{(0,2)} - \hat{r}_i^{(0,2)}) (\hat{\phi} - \tilde{\phi}) \right] \hat{b}_{(1)} \\
 &= \sum_{i=1}^n \left[ (-\hat{d}_i^{(0,2)} \hat{\mu}_i - \hat{t}_i^{(0,2)}) (\hat{\mu}_i(1 + \hat{\phi}) - \tilde{\mu}_i(1 + \tilde{\phi})) + (\hat{d}_i^{(0,2)} - \hat{r}_i^{(0,2)} - \hat{t}_i^{(0,2)} \hat{\mu}_i) (\hat{\phi} - \tilde{\phi}) \right] \hat{b}_{(1)},
 \end{aligned}$$

Matricialmente,

$$\begin{aligned}
 E_{\boldsymbol{\theta}} [U_{\beta}(\boldsymbol{\theta})(\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2))] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= (1 + \hat{\phi}) \mathbf{X}^T \hat{\mathbf{D}}_1 \left[ \mathbf{V}_1 \left( (1 + \hat{\phi}) \widehat{\mathbf{M}} - (1 + \tilde{\phi}) \widetilde{\mathbf{M}} \right) + (\hat{\phi} - \tilde{\phi}) \mathbf{C} \right] \mathbf{1}, \\
 E_{\boldsymbol{\theta}} [U_{\phi}(\boldsymbol{\theta})(\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}_2))] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_2=\tilde{\boldsymbol{\theta}}} &= \mathbf{1}^T \left[ (\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) \left( (1 + \hat{\phi}) \widehat{\mathbf{M}} - (1 + \tilde{\phi}) \widetilde{\mathbf{M}} \right) + (\hat{\phi} - \tilde{\phi}) (\mathbf{C} \widehat{\mathbf{M}} + \mathbf{V}_2 - \mathbf{C}) \right] \mathbf{1} \hat{b}_{(1)}.
 \end{aligned}$$

Logo, após essa álgebra tediosa,

$$\bar{\mathbf{v}} = \begin{bmatrix} (1 + \hat{\phi}) \mathbf{X}^T \hat{\mathbf{D}}_1 \left[ \mathbf{V}_1 \left( (1 + \hat{\phi}) \widehat{\mathbf{M}} - (1 + \tilde{\phi}) \widetilde{\mathbf{M}} \right) + (\hat{\phi} - \tilde{\phi}) \mathbf{C} \right] \mathbf{1} \\ \mathbf{1}^T \left[ (\mathbf{V}_1 \widehat{\mathbf{M}} + \mathbf{C}) \left( (1 + \hat{\phi}) \widehat{\mathbf{M}} - (1 + \tilde{\phi}) \widetilde{\mathbf{M}} \right) + (\hat{\phi} - \tilde{\phi}) (\mathbf{C} \widehat{\mathbf{M}} + \mathbf{V}_2 - \mathbf{C}) \right] \mathbf{1} \hat{b}_{(1)} \end{bmatrix}.$$