



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

Raquel Silva Maximiano

Desempenho de Alguns Estimadores Ridge no Modelo de Regressão Birnbaum-Saunders

Campina Grande - PB

14 de dezembro de 2020

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

Desempenho de Alguns Estimadores Ridge no Modelo de Regressão Birnbaum-Saunders

por

Raquel Silva Maximiano [†]

sob orientação da

Prof^a. Dr^a. Michelli Karinne Barros da Silva

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

[†]Este trabalho contou com apoio financeiro da CAPES.

M464d Maximiano, Raquel Silva.
Desempenho de alguns estimadores Ridge no modelo de regressão
Birnbaum-Saunders / Raquel Silva Maximiano. - Campina Grande, 2020.
85 f. : il. Color.

Dissertação (Mestrado em Matemática) - Universidade Federal de
Campina Grande, Centro de Ciências e Tecnologia, 2020.
"Orientação: Profa. Dra. Michelli Karinne Barros da Silva".
Referências.

1. Estimador Ridge. 2. Modelo de Regressão Birnbaum-Saunders. 3.
Multicolinearidade. I. Silva, Michelli Karinne Barros da. II. Título.

CDU 519.2 (043)

Desempenho de Alguns Estimadores Ridge no Modelo de Regressão Birnbaum-Saunders

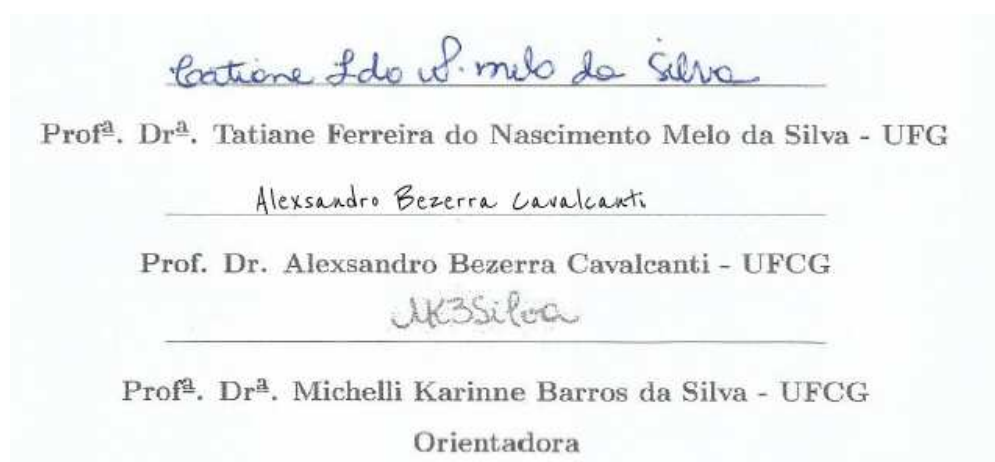
por

Raquel Silva Maximiano

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

Área de Concentração: Probabilidade e Estatística

Aprovada por:



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

14 de dezembro de 2020

Resumo

O modelo de regressão Birnbaum-Saunders tem sido aplicado em diferentes áreas do conhecimento, como exemplo, na análise de dados de sobrevivência e confiabilidade. No entanto, a presença de multicolinearidade no modelo afeta negativamente a variância do estimador de máxima verossimilhança. Na literatura encontramos vários métodos para minimizar tais efeitos, um desses é o estimador ridge que vêm sendo bastante estudado e se mostrado eficiente na presença de multicolinearidade. Neste trabalho, propomos estimadores ridge para os parâmetros regressores do modelo Birnbaum-Saunders. Avaliamos o estimador proposto em termos do erro quadrático médio e mostramos que o mesmo possui erro quadrático médio menor que o do estimador de máxima verossimilhança. Também apresentamos algumas propostas de estimação para o parâmetro de encolhimento. Avaliamos os estimadores propostos através de simulação de Monte Carlo.

Palavras-chave: estimador ridge; modelo de regressão Birnbaum-Saunders; multicolinearidade.

Abstract

The Birnbaum-Saunders regression model has been applied in different areas of knowledge, for example, in the analysis of competence and reliability data. However, the presence of multicollinearity in the model negatively affects a variance of the maximum likelihood estimator. In the literature, we find several methods to minimize such effects, one of which is the ridge estimator it provides, which has been extensively studied and is efficient in the presence of multicollinearity. In this work, we propose ridge estimators for the regressive parameters of the Birnbaum-Saunders model. We evaluated the proposed estimator in terms of the mean square error and showed that it has a lower mean square error than the maximum likelihood estimator. We also present some evaluation proposals for the shrinkage parameter. We evaluated the proposed estimators through Monte Carlo simulation.

Keywords: ridge estimator; Birnbaum-Saunders regression model; multicollinearity.

Agradecimentos

Primeiramente, agradeço ao nosso grandioso Deus, por ter me concedido a dádiva da vida e permitido chegar até aqui; por ter me ajudado, sustentado, orientado e colocado pessoas extraordinárias, incríveis no meu caminho. Nos momentos em que pensei em desistir, senti o Espírito Santo renovando minhas forças e me dando sabedoria para prosseguir. Toda a honra, glória, majestade e louvor seja dado a Ti. Obrigada, meu Deus, a graça é toda Tua!

Agradeço à minha mãe, Maria do Socorro Silva, que sempre foi a minha grande incentivadora, amiga, companheira e confidente. Chorou muitas vezes comigo, quando pensei em desistir, e sempre acreditou no meu potencial. Cuidou do meu filho com muito amor e carinho enquanto eu estudava. Sou grata ao meu esposo, que compreendeu minha ausência pelo tempo dedicado aos estudos, nos últimos dias, ajudou nas atividades domésticas. Obrigada por todo o apoio, pelas palavras de carinho e por ter estado ao meu lado em todos os momentos. Agradeço ao meu pai, à minha tia, ao meu irmão Franklin Silva e à sua esposa, à minha irmã Marília, ao meu primo André Silva e à sua esposa, por todas as orações, incentivos e por toda a ajuda financeira.

À minha grande amiga Camila Araújo, agradeço por cada segundo que passamos juntas na UFCG, por ser minha psicóloga particular, por cada questão feita e desenvolvida desde a graduação até aqui, as tapiocas e lanches pagos, que foram muitos. Obrigada por sua grandiosa amizade, na verdade, você é uma irmã que Deus colocou em meu caminho. Aproveito aqui para agradecer a Benedito Vicente, pela convivência ao longo do mestrado, e aos meus amigos da graduação Rafael Iago Bezerra de Oliveira, Fábio Monteiro, Leo Miller, Amanda Raíssa e José Lucas Galdino, pelo companheirismo.

Quero agradecer também à minha eterna professora, Michelli Karinne Barros da Silva, por toda orientação, paciência, compreensão e todo tempo e conhecimento que compartilhamos juntas. Mesmo na pandemia e com todas as responsabilidades profissionais e familiares, nunca deixou de atender seus alunos, sempre estava disponível para nos responder. Agradeço por toda orientação desde as etapas do mestrado até a

conclusão da dissertação, não tenho palavras para descrever a minha eterna gratidão. Nunca vou esquecer de tudo que fez por mim, de todas as aulas por ela ministradas, de cada teorema e propriedade demonstrada. Parabéns por ser tão guerreira, uma grande mãe, uma excelente professora e ,principalmente, uma mulher incrível! Muito obrigada!

Sou grata a todos os meus professores, em especial, ao admirável professor Joelson da Cruz Campos, por toda ajuda, informação, conhecimento e sabedoria passado no decorrer deste trabalho. Obrigada por sua paciência, seu bom humor, seu carinho, seus valiosos conselhos e opiniões durante esses três anos de convivência! Muito obrigado! Faltam palavras para expressar minha infinita gratidão. Agradeço ao professor da pós-graduação Manoel Ferreira dos Santos Neto, por todas as aulas ministradas, que foram importantes para minha formação. Aqui também não posso deixar de agradecer ao meu eterno professor da graduação e tutor do PET Matemática e Estatística, José Lindomberg Possiano Barreiro, por todo incentivo, por sua grande amizade e por ter acreditado no meu potencial.

Agradeço aos professores Tatiane Ferreira do Nascimento Melo da Silva e Alexandro Bezerra Cavalcanti, por aceitarem o convite para participar da banca examinadora.

Por fim, agradeço à Capes, pelo apoio financeiro, pois, sem ele, não seria possível a finalização desta dissertação.

Dedicatória

Ao meu filho, Pedro Henrique Silva, a minha mãe, Maria do Socorro Silva, e ao meu esposo, Iwry Wanderly Maximiano.

Sumário

Introdução	1
1 Distribuição Birnbaum-Saunders	3
1.1 Origem	3
1.2 Propriedades	6
1.3 Distribuição Birnbaum-Saunders reparametrizada	7
1.4 Modelo de regressão Birnbaum-Saunders	8
1.4.1 Estimação dos parâmetros	10
1.4.2 Distribuição assintótica	13
1.4.3 Erro quadrático médio do estimador $\hat{\beta}_{MV}$	13
2 Estimador ridge para o modelo de regressão BS	15
2.1 Multicolinearidade	15
2.2 Contexto histórico da regressão ridge	16
2.3 Estimador ridge para os coeficientes de regressão do modelo BS	17
2.3.1 Erro quadrático médio	18
2.3.2 Estimação do parâmetro k	26
3 Avaliação Numérica	28
3.1 Simulação	28
3.1.1 Resultados das simulações	29
3.2 Aplicação	35
4 Conclusão	37
A Função escore, matriz de informação de Fisher e estimação	38
B Tabelas referentes as estimativas dos vieses	54

Lista de Figuras

1.1	Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 0, 2; 0, 5; 1, 0; 1, 5; 1, 8$ e $\beta = 1$	6
1.2	Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 1$ e $\beta = 0, 2; 0, 5; 1, 0; 1, 5; 1, 8$	6
1.3	Gráfico da função densidade da distribuição Birnbaum-Saunders reparametrizada para os valores de $\delta = 1; 2; 15; 20; 50$ e $\mu = 1$	8
1.4	Gráfico da função densidade da distribuição Birnbaum-Saunders reparametrizada para os valores de $\delta = 7$ e $\mu = 0, 8; 1; 1, 8; 2; 3$	8

Lista de Tabelas

3.1	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 0,89$.	30
3.2	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 1,61$.	31
3.3	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 8$.	32
3.4	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 0,89$.	33
3.5	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 1,61$.	34
3.6	Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 8$.	35
3.7	Matriz de correlação amostral entre as variáveis	36
3.8	Desempenho dos estimadores para os dados da aplicação	36
B.1	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.	54
B.2	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.	55
B.3	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.	55
B.4	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.	56
B.5	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.	56
B.6	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.	57
B.7	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.	57
B.8	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.	58
B.9	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.	58
B.10	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.	59
B.11	Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.	59

B.12 Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$	60
B.13 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$	61
B.14 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$	62
B.15 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$	63
B.16 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$	64
B.17 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$	65
B.18 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$	66
B.19 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$	67
B.20 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$	68
B.21 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$	69
B.22 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$	70
B.23 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$	71
B.24 Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$	72

Introdução

Na regressão linear múltipla, encontramos situações em que as variáveis explicativas são correlacionadas, o que produz mau condicionamento da matriz de covariâncias, indicando problemas de multicolinearidade. Esse problema é muito comum na pesquisa aplicada e tem efeitos de alta variação sobre os coeficientes de regressão, tornando difícil a interpretação das estimativas dos parâmetros, pois apresentam resultados duvidosos. Na literatura, encontramos algumas alternativas para diminuir esses efeitos, um deles é o método de regressão ridge, que foi desenvolvido pela primeira vez por [Hoerl and Kennard \(1970\)](#) com o objetivo de melhorar a eficiência das estimativas do modelo de regressão linear, baseado no estimador ridge (ER), que apresenta o erro quadrático médio menor, comparado ao estimador de mínimos quadrados.

Nos últimos anos, muitos trabalhos vêm sendo desenvolvidos na obtenção de estimadores ridge para diversos modelos de regressão. [Segerstedt \(1992\)](#) estendeu a regressão ridge para os modelos lineares generalizados. A partir dessa extensão, vários trabalhos com distribuições pertencentes a família exponencial foram propostos, podemos citar [Schaefer et al. \(1984\)](#), [Lee and Silvapulle \(1988\)](#), [Cessie and van Houwelingen \(1992\)](#), [Lesaffre and Marx \(1993\)](#), [Mansson and Shukur \(2011\)](#), [Kibria et al. \(2012\)](#), [Kibria et al. \(2013b\)](#), [Algamal \(2019\)](#) e [Amin et al. \(2020\)](#). Recentemente, os trabalhos de [Kibria et al. \(2013a\)](#) e [Yüzbaşı and Asar \(2018\)](#) mostraram o bom desempenho do estimador ridge em modelos inflacionados de zeros.

A distribuição Birnbaum-Saunders (BS), conhecida por modelar o tempo de vida de materiais e equipamentos submetidos a cargas cíclicas, foi desenvolvida por [Birnbaum and Saunders \(1969\)](#). Essa distribuição vem sendo amplamente usada em diferentes áreas do conhecimento, como por exemplo, nas áreas ambiental e florestal, na análise de sobrevivência, ciências médicas, finanças, entre outros (veja [Leiva et al. \(2009\)](#), [Vilca et al. \(2010\)](#) [Paula et al. \(2012\)](#) e [Saulo et al. \(2019\)](#)).

Na literatura, encontramos várias modelagens baseadas no modelo BS. [Rieck and Nedelman \(1991\)](#) propuseram um modelo log-linear para a distribuição BS. [Galea et al. \(2004\)](#) realizaram diagnóstico de influência no modelo log-BS. [Leiva et al. \(2014\)](#) propuseram um modelo de regressão BS reparametrizada, em que uma função de ligação relaciona a média e o preditor linear. No entanto, na presença de multicolinearidade pode produzir efeitos negativos nas estimativas dos modelos de regressão BS. Neste trabalho, propomos estimadores ridge para os coeficientes da regressão do modelo BS reparametrizado e comparamos o desempenho dos estimadores ridge com o de máxima verossimilhança.

O restante da dissertação está dividida em quatro capítulos. No Capítulo 1, apresentamos uma revisão do contexto histórico da distribuição BS e abordamos algumas propriedades, bem como uma do modelo de regressão BS reparametrizados, proposto por [Santos-Neto \(2010\)](#). Apresentamos o processo de estimação pelo método da máxima verossimilhança dos parâmetros do modelo, bem como o erro quadrático médio (EQM) dos estimadores dos coeficientes da regressão. No Capítulo 2, apresentamos o conceito da multicolinearidade e o contexto histórico da regressão ridge. Em seguida, obtemos a expressão do estimador ridge, encontramos o EQM e mostramos que ele é menor que o do estimador de máxima verossimilhança. Além disso, propomos alguns estimadores para o parâmetro de encolhimento. No Capítulo 3, realizamos uma avaliação numérica através de estudos de simulação de Monte Carlo e apresentamos uma aplicação a um conjunto de dados real. No Capítulo 4, apresentamos as conclusões do trabalho.

Suporte computacional

O estudo de simulação, os gráficos e a aplicação desenvolvidas neste trabalho foram feitas usando o software R, disponível gratuitamente no endereço <http://www.r-project.org>. A linguagem R é utilizada em diversas áreas do conhecimento, pois facilita a manipulação dos cálculos, a construção de gráficos e a análise estatística. A dissertação foi escrita usando o LATEX.

Capítulo 1

Distribuição Birnbaum-Saunders

Neste capítulo, faremos um breve resumo sobre a distribuição BS, na sua versão clássica e reparametrizada. Assim como, apresentamos o modelo de regressão BS e o processo de estimação dos parâmetros via método de máxima verossimilhança.

1.1 Origem

Grande parte dos componentes mecânicos são submetidas a cargas cíclicas, que após um determinado tempo de trabalho, reduz a resistência do material, ocasionando modificações estruturais que podem culminar em trincas ou ruptura total do componente. Tal fenômeno é conhecido como fadiga. Para garantir um melhor controle da falha por fadiga, torna-se indispensável o estudo de modelos estatísticos que descrevem a variação aleatória dos tempos de falha relacionados aos materiais e equipamentos expostos a cargas dinâmicas. As famílias de distribuições mais utilizadas para esse tipo de modelagem são os modelos gama, gama inversa, lognormal e Weibull. No entanto, na existência de dados atípicos como, por exemplo, tempos de fadiga muito altos ou muito baixos, essas distribuições não apresentam um bom ajuste. Por esse motivo, [Birnbaum and Saunders \(1969\)](#) desenvolveram uma nova família de distribuição nomeada Birnbaum-Saunders, a qual apresentou um melhor ajuste nos percentis mais baixos ou mais altos.

Para a construção da nova família de distribuições, os autores consideraram uma amostra de material que foi submetida a cargas cíclicas de tensão e força. Em outras palavras, uma sequência de m cargas l_1, l_2, \dots, l_m que serão aplicadas em um ciclo, essa mesma sequência foi repetida em cada ciclo de cargas. Assim, para algum $m > 1$ e

$\forall i = 1, \dots, m,$

$$l_{jm+i} = l_{km+i}, \forall j \neq k.$$

Como consequência dessas cargas, resulta-se o desgaste do material, causando uma fissura. Devido ao crescimento e evolução da fissura dominante, ocorre a falha por fadiga, isto é, quando o comprimento da fissura ultrapassa um certo nível de resistência, que será um valor crítico, denotado por ω . Assim, [Birnbaum and Saunders \(1969\)](#) assumiram algumas suposições:

- X_i é a extensão incremental da fissura após a aplicação da i -ésima oscilação de carga, ou seja, é uma variável aleatória com uma distribuição que só depende da fissura atual, gerada pela tensão neste ciclo.
- A extensão da fissura durante o $(j + 1)$ -ésimo ciclo é $Y_{j+1} = \sum_{i=jm+1}^{jm+m} X_i$, em que X_{jm+1} é a extensão da fissura após a carga aplicada na i -ésima oscilação do $(j + 1)$ -ésimo ciclo.
- A extensão total durante Y_j , devido ao j -ésimo ciclo, é uma variável aleatória independente e identicamente distribuída com média μ e variância σ^2 , $\forall j = 1, 2, \dots$.

Portanto, a variável aleatória que corresponde a extensão total da fissura durante todos os n -ciclos de cargas é

$$W_n = \sum_{j=1}^n Y_j,$$

cuja função de distribuição acumulada (fda) é dada por

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

Considere C o número de ciclos de cargas até a ocorrência da falha, ou seja, quando a extensão total da fissura excede o valor crítico ω . Desse modo, a função de distribuição acumulada da variável aleatória C é dada por

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

Note que,

$$\begin{aligned} P(C \leq n) &= 1 - P\left(\sum_{j=1}^n Y_j \leq \omega\right) \\ &= 1 - P\left(\sum_{j=1}^n \frac{Y_j - \mu}{\sigma\sqrt{n}} \leq \frac{\omega - n\mu}{\sigma\sqrt{n}}\right). \end{aligned}$$

As variáveis Y_i 's são independentes e identicamente distribuídas, então pelo Teorema Central do Limite, a função de distribuição de C é aproximadamente

$$\begin{aligned} P(C \leq n) &\cong 1 - \Phi\left(\frac{\omega - n\mu}{\sigma\sqrt{n}}\right) \\ &= 1 - \Phi\left(\frac{\omega}{\sigma\sqrt{n}} - \frac{\sqrt{n}\mu}{\sigma}\right) \\ &= \Phi\left(\frac{n\mu}{\sigma\sqrt{n}} - \frac{\omega}{\sigma\sqrt{n}}\right), \end{aligned}$$

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

[Birnbaum and Saunders \(1969\)](#) observaram que, ao substituir n por uma variável real não negativa $t > 0$, a variável T passa a ter uma extensão contínua da variável aleatória discreta C . Assim, T é o tempo total até a ocorrência da falha e definindo

$$\alpha = \frac{\sigma}{\sqrt{\mu w}} \quad \text{e} \quad \beta = \frac{w}{\mu},$$

temos que a fda da variável T fica dada por

$$\begin{aligned} F_T(t) &= P(T \leq t) \\ &= \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], \end{aligned} \tag{1.1}$$

com $t > 0$, $\alpha > 0$ e $\beta > 0$. Portanto, a variável aleatória não negativa T segue uma distribuição BS, com parâmetros de forma e escala α e β , respectivamente, e denotamos por $T \sim \text{BS}(\alpha, \beta)$.

A função densidade de probabilidade (fdp) de T é obtida derivando (1.1) com relação à t e é dada por

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right] \frac{t^{-3/2}(t + \beta)}{2\alpha\sqrt{\beta}}, \tag{1.2}$$

em que $t, \alpha, \beta > 0$.

Como podemos observar na Figura 1.1 o parâmetro α controla a forma da distribuição, a distribuição torna-se assimétrica à medida que α cresce, e simétrica, em

torno de β , quando α decresce. O parâmetro β controla a escala da distribuição, como podemos ver na Figura 1.2.

Figura 1.1: Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 0,2; 0,5; 1,0; 1,5; 1,8$ e $\beta = 1$.

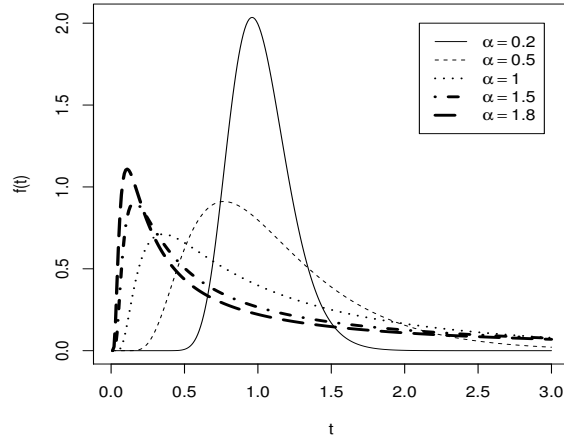
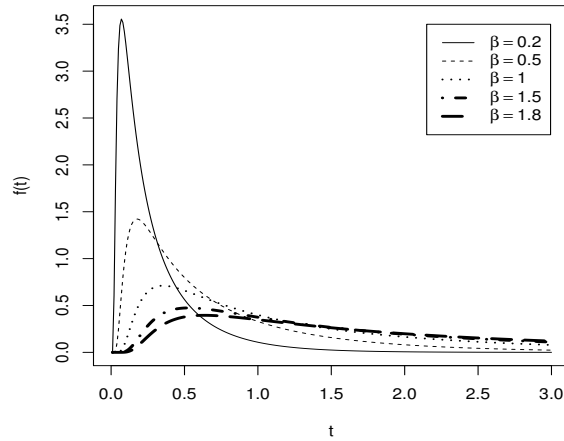


Figura 1.2: Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 1$ e $\beta = 0,2; 0,5; 1,0; 1,5; 1,8$.



1.2 Propriedades

1. Se $T \sim \text{BS}(\alpha, \beta)$ então $Z = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, 1)$. Consequentemente,

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right]^2, \text{ em que } Z \sim N(0, 1);$$

2. Se $T \sim \text{BS}(\alpha, \beta)$ e $c > 0$, então $cT \sim \text{BS}(\alpha, c\beta)$;

3. Se $T \sim \text{BS}(\alpha, \beta)$, então $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$;
4. Se $T \sim \text{BS}(\alpha, \beta)$, então $\text{Med}(T) = \beta$, em que Med denota a mediana de T ;
5. Se $T \sim \text{BS}(\alpha, \beta)$, então a média e a variância de T são, respectivamente,

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

Para maiores detalhes sobre as propriedades da distribuição Birnbaum-Saunders, ver [Leiva \(2015\)](#).

1.3 Distribuição Birnbaum-Saunders reparametrizada

De acordo com [Santos-Neto \(2010\)](#) a reparametrização de distribuições pode ser uma ferramenta muito importante, pois facilita o desenvolvimento analítico, computacional e inferencial de algumas distribuições. Para algumas situações, como em regressão, quando a distribuição da variável resposta não apresenta a média como um de seus parâmetros é possível através de uma reparametrização ajustar a média da variável resposta. Na literatura, encontramos diferentes reparametrizações que apresentam bons resultados, como por exemplo, a distribuição beta (ver [Ferrari and Cribari-Neto \(2004\)](#)), a distribuição inversa gaussiana (ver [Tweedie et al. \(1957\)](#)) e a distribuição BS (ver [Ahmed et al. \(2008\)](#) e [Santos-Neto \(2010\)](#)).

Uma nova parametrização da distribuição BS proposta por [Santos-Neto et al. \(2012\)](#) é indexada pelos parâmetros μ e δ , em que $\mu > 0$ é o parâmetro de escala e a média da distribuição, enquanto $\delta > 0$ é o parâmetro de forma e precisão. Com base nesta parametrização, a fdp de T fica dada por

$$f(t; \mu, \delta) = \frac{\exp(\delta/2)\sqrt{\delta+1}}{4t^{3/2}\sqrt{\pi\mu}} \left[t + \frac{\delta\mu}{\delta+1} \right] \exp \left\{ -\frac{\delta}{4} \left[\frac{t(\delta+1)}{\delta\mu} + \frac{\delta\mu}{t(\delta+1)} \right] \right\}. \quad (1.3)$$

Neste caso, usamos a notação $T \sim \text{BSR}(\mu, \delta)$. A média e a variância de T são dadas, respectivamente, por

$$E[T] = \mu \quad \text{e} \quad \text{Var}[T] = \frac{\mu^2}{\phi},$$

em que $\phi = \frac{(\delta+1)^2}{(2\delta+5)}$.

Na Figura 1.3, identificamos que o parâmetro δ controla a forma da distribuição, a medida que δ cresce a fdp da BS parametrizada torna-se simétrica. O parâmetro μ controla a escala da distribuição, como podemos ver na Figura 1.4.

Figura 1.3: Gráfico da função densidade da distribuição Birnbaum-Saunders reparametrizada para os valores de $\delta = 1; 2; 15; 20; 50$ e $\mu = 1$.

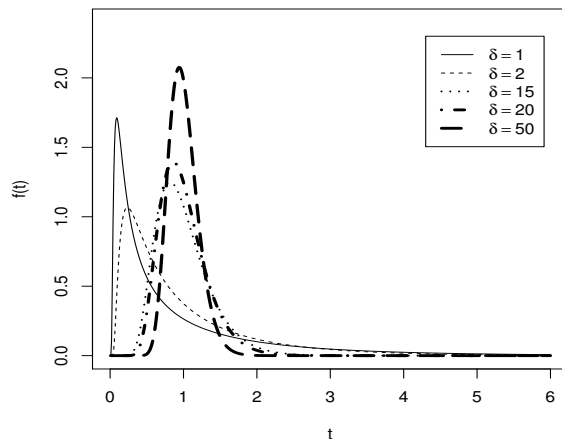
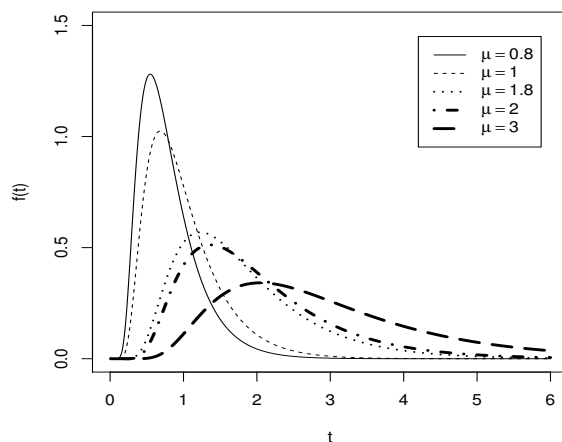


Figura 1.4: Gráfico da função densidade da distribuição Birnbaum-Saunders reparametrizada para os valores de $\delta = 7$ e $\mu = 0,8; 1; 1,8; 2; 3$.



Aqui δ é interpretado como parâmetro de precisão, pois para μ fixado, quando $\delta \rightarrow \infty$, $\text{Var}(T) \rightarrow 0$. Para maiores detalhes dessa reparametrização ver Santos-Neto (2010) e Santos-Neto et al. (2012).

1.4 Modelo de regressão Birnbaum-Saunders

Sejam Y_1, \dots, Y_n variáveis aleatórias independentes, com $Y_i \sim \text{BSR}(\mu_i, \delta)$, para $i = 1, \dots, n$, cuja fdp é dada em (1.3). O modelo de regressão BS baseado em (1.3) é

definido de forma similar aos modelos lineares generalizados considerando a componente sistemática

$$g(\mu_i) = \eta_i, i = 1, \dots, n, \quad (1.4)$$

em que $\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ é o preditor linear, $g(\cdot)$ é a função de ligação, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^\top$, $(p+1) < n$, é o vetor de parâmetros desconhecidos a serem estimados e $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})^\top$ representa os valores das variáveis explicativas. Assume-se $g : \mathbb{R} \rightarrow \mathbb{R}^+$ uma função monótona e diferenciável pelo menos duas vezes.

Esse modelo é mais flexível do que o modelo de regressão proposto por [Rieck and Nedelman \(1991\)](#), pois dá maior flexibilidade para a relação funcional entre a média da variável resposta e o preditor linear.

Considerando a partição $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \delta)^\top$, o logaritmo da função de verossimilhança do modelo (1.4) é dado por

$$l(\boldsymbol{\theta}) = \log \prod_{i=1}^n f(y_i; \mu_i, \delta) = \sum_{i=1}^n l(\mu_i, \delta)$$

em que

$$l(\mu_i, \delta) = \frac{\delta}{2} - \frac{1}{2} \log(16\pi) - \frac{1}{2} \log(\delta + 1) - \frac{3}{2} \log(y_i) - \frac{1}{2} \log(\mu_i) \\ + \log(y_i + \delta y_i + \delta \mu_i) - \frac{y_i(\delta + 1)}{4\mu_i} - \frac{\delta^2 \mu_i}{4y_i(\delta + 1)}.$$

As funções escores de $\boldsymbol{\beta}$ e δ são dadas, respectivamente, por:

$$\mathbf{U}_\beta(\boldsymbol{\theta})_{(p+1) \times 1} = \mathbf{X}^\top D(a) \mathbf{z}$$

e

$$U_\delta(\boldsymbol{\theta})_{1 \times 1} = \text{tr} [D(b)],$$

em que $D(a) = \text{diag}\{a_1, \dots, a_n\}$, com $a_i = \frac{1}{g'(\mu_i)}$, $\mathbf{z} = (z_1, \dots, z_n)^\top$, sendo $z_i = -\frac{1}{2\mu_i} + \frac{\delta}{\delta y_i + y_i + \delta \mu_i} + \frac{y_i(\delta+1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta+1)}$ e $D(b) = \text{diag}\{b_1, \dots, b_n\}$, com $b_i = \frac{1}{2} - \frac{1}{2(\delta+1)} + \frac{y_i + \mu_i}{\delta y_i + y_i + \delta \mu_i} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta+2)\mu_i}{4(\delta+1)^2 y_i}$.

Segue que a matriz de informação de Fisher para $\boldsymbol{\theta}$ possui a forma

$$K_{\theta\theta} = \begin{pmatrix} K_{\beta\beta} & K_{\beta\delta} \\ K_{\delta\beta} & K_{\delta\delta} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^\top W_{\beta\beta} \mathbf{X} & \mathbf{X}^\top D(a) \mathbf{s} \\ \mathbf{s}^\top D(a) \mathbf{X} & \text{tr} [D(u)] \end{pmatrix},$$

em que $W_{\beta\beta} = \text{diag}\{v_1, \dots, v_n\}$, $D(u) = \text{diag}\{u_1, \dots, u_n\}$ e $\mathbf{s} = (s_1, \dots, s_n)^\top$, onde $v_i = \frac{\delta}{2\mu_i^2\{g'(\mu_i)\}^2} + \frac{\delta^2}{(\delta+1)^2\{g'(\mu_i)\}^2}\mathcal{I}(\delta)$, $u_i = \frac{(\delta^2+3\delta+1)}{2(\delta+1)^2\delta^2} + \frac{\mu_i^2}{(\delta+1)^4}\mathcal{I}(\delta)$ e $s_i = \frac{1}{2\mu_i(\delta+1)} + \frac{\delta\mu_i}{(\delta+1)^3}$, sendo $\mathcal{I}(\delta) = \int_0^\infty \frac{1}{4}e^{\frac{\delta}{2}}\sqrt{\delta}e^{-\frac{1}{4}\left(\frac{\delta+1}{\delta\mu}y + \frac{\delta\mu}{(\delta+1)y}\right)\delta} \left(y + \frac{\delta\mu}{(\delta+1)}\right)^{-1} \left(\sqrt{\frac{\pi\delta\mu}{\delta+1}}\right)^{-1} y^{-3/2}dy$.

Para obter a inversa da matriz de informação de Fisher, podemos utilizar a operação de bloco elementar (ver [Banerjee and Roy \(2014\)](#)). Sendo assim, $K_{\theta\theta}^{-1}$ fica dada por

$$K_{\theta\theta}^{-1} = \begin{pmatrix} D^{-1} + \frac{\mathbf{B}\mathbf{B}^\top}{h} & -\frac{\mathbf{B}}{h} \\ -\frac{\mathbf{B}^\top}{h} & \frac{1}{h} \end{pmatrix},$$

em que $D = X^\top W_{\beta\beta} X$, com dimensão $(p+1) \times (p+1)$, $\mathbf{B} = D^{-1} X^\top D(a)\mathbf{s}$, com dimensão $(p+1) \times 1$ e $h = \text{tr}[D(u)] - \mathbf{s}^\top D(a)X\mathbf{B}$ é um escalar.

Consideremos as matrizes aumentadas, \widetilde{W} e \widetilde{X} , com respectivas dimensões, $(n+1) \times (n+1)$ e $(n+1) \times (p+2)$, definidas por

$$\widetilde{W} = \begin{pmatrix} W_{\beta\beta} & W_{\beta\delta} \\ W_{\delta\beta} & W_{\delta\delta} \end{pmatrix} = \begin{pmatrix} W_{\beta\beta} & D(a)\mathbf{s} \\ \mathbf{s}^\top D(a) & \text{tr}[D(u)] \end{pmatrix},$$

e

$$\widetilde{X} = \begin{pmatrix} X & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}.$$

em que $W_{\beta\beta} = D(v)$, $W_{\beta\delta} = [W_{\delta\beta}]^\top = D(a)\mathbf{s}$ e $W_{\delta\delta} = \text{tr}[D(u)]$, todas de dimensão $n \times n$.

Conseqüentemente, podemos reescrever a matriz de informação de Fisher, na forma

$$K_{\theta\theta} = \widetilde{X}^\top \widetilde{W} \widetilde{X},$$

e sua inversa, como

$$K_{\theta\theta}^{-1} = (\widetilde{X}^\top \widetilde{W} \widetilde{X})^{-1}.$$

Para maiores detalhes dos cálculos dessa seção ver Apêndice [A](#) e [Santos-Neto \(2010\)](#).

1.4.1 Estimação dos parâmetros

Os estimadores de máxima verossimilhança dos coeficientes de regressão e do parâmetro de forma são soluções das equações $U_{\beta_j}(\boldsymbol{\theta}) = 0$ ($j = 0, \dots, p$) e $U_\delta(\boldsymbol{\theta}) = 0$. Porém, neste caso, essas equações não apresentam soluções analíticas, sendo necessária

a utilização de método iterativo para a obtenção das raízes. Sendo assim, utilizamos o método escore de Fisher. O algoritmo para a estimação de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \delta)^\top$ baseado em m passos é obtido por

$$\begin{aligned}
\boldsymbol{\theta}^{(m+1)} &= \boldsymbol{\theta}^{(m)} + (K_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1})^{(m)} \mathbf{U}_{\boldsymbol{\theta}}^{(m)} \\
&= \boldsymbol{\theta}^{(m)} + (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \begin{pmatrix} X^\top D(a)^{(m)} \mathbf{z}^{(m)} \\ \text{tr}[D(b)] \end{pmatrix} \\
&= \boldsymbol{\theta}^{(m)} + (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \begin{pmatrix} X^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} D(a)^{(m)} \mathbf{z}^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \\
&= \boldsymbol{\theta}^{(m)} + (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \begin{pmatrix} X^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix} \\
&= \boldsymbol{\theta}^{(m)} + (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \tilde{X}^\top \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix} \\
&= \boldsymbol{\theta}^{(m)} + (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \tilde{X}^\top \tilde{W}^{(m)} [\tilde{W}^{(m)}]^{-1} \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix},
\end{aligned}$$

$m = 0, 1, \dots$. Se trabalharmos um pouco ao lado direito da expressão acima, chegamos na forma

$$\boldsymbol{\theta}^{(m+1)} = (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \tilde{X}^\top \tilde{W}^{(m)} \mathbf{z}^{*(m)}, \quad (1.5)$$

em que

$$\mathbf{z}^{*(m)} = \tilde{X} \boldsymbol{\theta}^{(m)} + [\tilde{W}^{(m)}]^{-1} \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix}.$$

Observando (1.5), notamos que $\boldsymbol{\theta}^{(m+1)}$ apresenta uma forma do processo iterativo de mínimos quadrados reponderado, a quantidade $\mathbf{z}^{*(m)}$ desempenha o papel de variável resposta modificada e que \tilde{X} da matriz modelo.

A seguir, expandiremos $\boldsymbol{\theta}^{(m+1)}$ para obtermos as expressões de $\boldsymbol{\beta}^{(m+1)}$ e $\delta^{(m+1)}$. Para tal procedimento, construímos uma sequência de etapas definidas da seguinte forma

1ª Etapa: Calcular $(\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \tilde{X}^\top \tilde{W}^{(m)}$.

Note que,

$$\begin{aligned} (\tilde{X}^\top \tilde{W}^{(m)} \tilde{X})^{-1} \tilde{X}^\top \tilde{W}^{(m)} &= [K_{\theta\theta}^{(m)}]^{-1} \tilde{X}^\top \tilde{W}^{(m)} \\ &= \begin{pmatrix} (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}. \end{aligned}$$

2ª Etapa: Para calcularmos $\mathbf{z}^{*(m)}$, precisamos encontrar a inversa da matriz $\tilde{W}^{(m)}$. Seguindo o procedimento proposto por [Banerjee and Roy \(2014\)](#), a inversa encontrada possui a seguinte estrutura

$$[\tilde{W}^{(m)}]^{-1} = \begin{pmatrix} [W_{\beta\beta}^{(m)}]^{-1} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top}{f} & -\frac{\mathbf{A}^{(m)}}{f} \\ -\frac{[\mathbf{A}^{(m)}]^\top}{f} & \frac{1}{f} \end{pmatrix},$$

em que $[W_{\beta\beta}^{(m)}]^{-1} = [D(v)^{(m)}]^{-1}$, $\mathbf{A}^{(m)} = [W_{\beta\beta}^{(m)}]^{-1} D(a)^{(m)} \mathbf{s}^{(m)}$ e $f = \text{tr}[D(u)] - [\mathbf{s}^{(m)}]^\top D(a)^{(m)} [W_{\beta\beta}^{(m)}]^{-1} D(a)^{(m)} \mathbf{s}^{(m)}$, com respectivas dimensões, $n \times n$, $n \times 1$ e f um escalar.

3ª Etapa: Substituímos as matrizes $[\tilde{W}^{(m)}]^{-1}$ e \tilde{X} , em $\mathbf{z}^{*(m)}$. Após algumas manipulações algébricas, encontramos a seguinte igualdade

$$\mathbf{z}^{*(m)} = \begin{pmatrix} \mathbf{z}_1^{*(m)} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix},$$

em que $\mathbf{z}_1^{*(m)} = X\boldsymbol{\beta}^{(m)} + \left[[W_{\beta\beta}^{(m)}]^{-1} D(a)^{(m)} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} \right] \mathbf{z}^{(m)} - \frac{\mathbf{A}^{(m)} \text{tr}[D(b)]}{f}$.

4ª Etapa: As expressões encontradas na 1ª Etapa e na 3ª Etapa foram substituídas em $\boldsymbol{\theta}^{(m+1)}$. Assim, ao final do procedimento, concluímos que

$$\boldsymbol{\theta}^{(m+1)} = \begin{pmatrix} (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} \mathbf{z}_1^{*(m)} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix}.$$

Portanto,

$$\boldsymbol{\beta}^{(m+1)} = (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} \mathbf{z}_1^{*(m)}$$

e

$$\delta^{(m+1)} = \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f}.$$

No último passo o estimador do vetor de parâmetros regressores fica definido como

$$\hat{\boldsymbol{\beta}}_{MV} = (X^\top \hat{W}_{\beta\beta} X)^{-1} X^\top \hat{W}_{\beta\beta} \hat{\mathbf{z}}_1^*. \quad (1.6)$$

1.4.2 Distribuição assintótica

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

$$\widehat{\boldsymbol{\theta}} \sim N_{p+2}(\boldsymbol{\theta}, K_{\theta\theta}^{-1}).$$

E conseqüentemente

$$\widehat{\boldsymbol{\beta}}_{MV} \sim N_{(p+1)}(\boldsymbol{\beta}, K^{\beta\beta})$$

e

$$\widehat{\delta} \sim N\left(\delta, \frac{1}{h}\right),$$

em que $K^{\beta\beta} = D^{-1} + \frac{\mathbf{B}\mathbf{B}^\top}{h}$.

1.4.3 Erro quadrático médio do estimador $\widehat{\boldsymbol{\beta}}_{MV}$

O EQM de $\widehat{\boldsymbol{\beta}}_{MV}$ é dado por:

$$\begin{aligned} \text{EQM}(\widehat{\boldsymbol{\beta}}_{MV}) &= E\left[(\widehat{\boldsymbol{\beta}}_{MV} - \boldsymbol{\beta})^\top (\widehat{\boldsymbol{\beta}}_{MV} - \boldsymbol{\beta})\right] = E\left[\begin{pmatrix} (\widehat{\beta}_0 - \beta_0) \\ (\widehat{\beta}_1 - \beta_1) \\ \vdots \\ (\widehat{\beta}_p - \beta_p) \end{pmatrix}^\top \begin{pmatrix} (\widehat{\beta}_0 - \beta_0) \\ (\widehat{\beta}_1 - \beta_1) \\ \vdots \\ (\widehat{\beta}_p - \beta_p) \end{pmatrix}\right] \\ &= E\left[(\widehat{\beta}_0 - \beta_0)^2 + (\widehat{\beta}_1 - \beta_1)^2 + \cdots + (\widehat{\beta}_p - \beta_p)^2\right] \\ &= E\left[\sum_{j=0}^p (\widehat{\beta}_j - \beta_j)^2\right] \\ &= \sum_{j=0}^p E\left[(\widehat{\beta}_j - \beta_j)^2\right]. \end{aligned}$$

Assintoticamente temos

$$\begin{aligned} \text{EQM}(\widehat{\boldsymbol{\beta}}_{MV}) &= \sum_{j=0}^p E\left[(\widehat{\beta}_j - E[\widehat{\beta}_j])^2\right] \\ &= \sum_{j=0}^p \text{Var}(\widehat{\beta}_j), \end{aligned}$$

pois o EMV é assintoticamente não-viesado.

Como o traço de uma matriz é a soma dos elementos da diagonal principal temos que

$$\text{EQM}(\hat{\beta}_{MV}) = \text{tr} [\text{Cov}(\hat{\beta}_{MV})] = \text{tr} [K^{\beta\beta}].$$

Capítulo 2

Estimador ridge para o modelo de regressão BS

Neste capítulo, apresentamos o conceito de multicolinearidade, o contexto histórico do estimador ridge e propomos um estimador ridge para os coeficientes de regressão do modelo BSR. Obtivemos o EQM e provamos alguns resultados. Por fim, propomos alguns estimadores para o parâmetro de encolhimento.

2.1 Multicolinearidade

Em modelos de regressão estudamos como o comportamento de uma variável, chamada variável resposta, pode ser explicado através de uma estrutura com covariáveis (variáveis explicativas) e parâmetros desconhecidos.

É comum em modelos de regressão que as variáveis explicativas sejam correlacionadas umas com as outras. No entanto, quando as correlações são altas, levam ao problema que chamamos de multicolinearidade. Esse problema é muito comum na pesquisa aplicada e tem efeitos indesejáveis no estimador de mínimos quadrados (EMQ) ou de máxima verossimilhança (EMV), pois obtemos estimativas instáveis dos parâmetros de regressão o que acarreta em dificuldades na interpretação das mesmas, além de erros-padrão grandes (ver [Yahya Algamal \(2019\)](#) e [Amin et al. \(2020\)](#)).

Segundo [Montgomery et al. \(2012\)](#) a multicolinearidade pode ser causada por quatro fatores: o método de coleta de dados usados, restrições no modelo ou na população, especificação do modelo e excesso de termos no modelo. Na literatura encontramos diversos métodos que identificam o problema, entre eles, a matriz de

correlação, fator de inflação da variância e número de condição.

Na prática, quando as variáveis explicativas não são altamente correlacionadas, elimina-se uma ou mais variáveis do modelo. O problema no que se refere é quando ocorre um alto grau de correlação entre elas ou apresenta uma combinação quase linear entre o conjunto das variáveis explicativas. A dependência entre essas variáveis provoca sérios efeitos nas estimativas dos coeficientes de regressão como, por exemplo, no método de EMQ não é possível resolver o sistema de equações normais. Para minimizar os problemas causados com a multicolinearidade várias alternativas foram propostas como eliminar covariáveis do modelo, coletar mais informações ou utilizar outros estimadores, como o estimador ridge.

2.2 Contexto histórico da regressão ridge

O método de EMQ é uma técnica estatística conhecida por apresentar variância mínima de estimadores não viesados, sendo adequados para modelos que apresentam uma relação linear entre as variáveis. Contudo [Hoerl and Kennard \(1970\)](#) ressaltam que as estimativas apresentam um bom desempenho quando as variáveis explicativas não são correlacionadas, ou seja, quando na forma de correlação, a matriz $X^T X$ é unitária, caso contrário, as estimativas de mínimos quadrados são sujeitas as diferentes tipos de erros. Para controlar os erros e a instabilidade geral das estimativas de mínimos quadrados, os autores destacam que [Hoerl \(1962\)](#) sugeriu, pela primeira vez, a utilização do estimador na forma

$$\hat{\beta}^* = [X^T X + kI]^{-1} X^T \mathbf{y}, \quad (2.1)$$

sendo $k \geq 0$ um parâmetro de encolhimento. A esse estimador, atribuiu-se o nome estimador ridge (ER). A ideia foi eliminar a multicolinearidade das variáveis explicativas adicionando uma pequena quantidade positiva no EMQ, ou seja, inserir um pequeno viés ao estimador, mas que a variância do estimador resultante seja menor do que a do EMQ. Neste caso a multicolinearidade é resolvida afastando a singularidade da matriz $X^T X$ por meio do acréscimo de uma constante k , possibilitando o cálculo da inversa de $X^T X$ (ver [Casagrande \(2016\)](#)). O ER tem a vantagem que seu erro quadrático médio (EQM) é menor que o do EMQ.

A estimação ridge apesar de ter sido proposta para corrigir o problema da multicolinearidade no processo de estimação de mínimos quadrados do modelo normal linear ela não se limita apenas a esse modelo, podendo ser usada para estimar os parâmetros regressores de qualquer modelo (ver [van Wieringen WN. \(2015\)](#)). Recentemente, vários trabalhos foram publicados com relação a estimadores ridge para modelos lineares generalizados sob diferentes distribuições, bem como para modelos inflacionados. Podemos citar os trabalhos [Schaefer et al. \(1984\)](#), [Lee and Silvapulle \(1988\)](#), [Cessie and van Houwelingen \(1992\)](#), [Lesaffre and Marx \(1993\)](#), [Mansson and Shukur \(2011\)](#) [Kibria et al. \(2012\)](#), [Kibria et al. \(2013b\)](#), [Algamal \(2019\)](#) e [Amin et al. \(2020\)](#). Os trabalhos de [Kibria \(2003\)](#), [Khalaf and Shukur \(2005\)](#), [Alkhamisi et al. \(2006\)](#), [Alkhamisi and Shukur \(2008\)](#) e [Muniz and Kibria \(2009\)](#) propuseram e compararam diferentes estimadores para o parâmetro k e destacaram a importância do estudo de simulação para selecionar o melhor estimador para os diferentes modelos de regressão.

2.3 Estimador ridge para os coeficientes de regressão do modelo BS

Assim como nos modelos lineares generalizados, na presença de multicolinearidade, a matriz $X^\top W_{\beta\beta} X$ torna-se inflada e mal condicionada, levando a alta variância e instabilidade dos estimadores de máxima verossimilhança dos parâmetros do modelo de regressão BS, definido na Seção (1.4). Então, seguindo a ideia de [Segerstedt \(1992\)](#), propomos o estimador ridge para os coeficientes da regressão do modelo BS. O desempenho do estimador proposto será avaliado em termos de EQM e investigado através de simulações de Monte Carlo, sob vários métodos para estimar o parâmetro de encolhimento.

Para o modelo de regressão BS, o estimador ridge é obtido quando aumentamos a diagonal principal da matriz $(X^\top W_{\beta\beta} X)^{-1}$ para valores positivos do parâmetro k , na expressão do EMV dada em (1.6). Ou seja,

$$\begin{aligned}\widehat{\beta}_R &= (X^\top \widehat{W}_{\beta\beta} X + kI_{(p+1) \times (p+1)})^{-1} X^\top \widehat{W}_{\beta\beta} X \widehat{\beta}_{MV} \\ &= (X^\top \widehat{W}_{\beta\beta} X + kI)^{-1} X^\top \widehat{W}_{\beta\beta} X (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\mathbf{z}}_1^* \\ &= (X^\top \widehat{W}_{\beta\beta} X + kI)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\mathbf{z}}_1^*.\end{aligned}$$

Observe que o estimador ridge tem como caso particular o estimador $\widehat{\boldsymbol{\beta}}_{MV}$, basta considerarmos $k = 0$.

Pesquisas evidenciam que os estimadores ridge, embora viesados apresentam o erro quadrático médio menor comparado aos EMV, no entanto, precisamos encontrar um estimador apropriado para o parâmetro k .

2.3.1 Erro quadrático médio

Nesta seção apresentamos alguns resultados que mostram que o EQM do ER para os parâmetros regressores do modelo BS é menor que o EQM do EMV. As suposições e os procedimentos apresentados abaixo seguem as mesmas seqüências usadas por [Kibria \(2003\)](#), [Muniz and Kibria \(2009\)](#) e [Dorugade \(2014\)](#).

Suponhamos $\Lambda = \text{diag}(\lambda_0, \dots, \lambda_p)$ e $T_{(p+1) \times (p+1)}$ são as matrizes de autovalores e autovetores de $(X^\top \widehat{W}_{\beta\beta} X)$, respectivamente, que satisfazem as igualdades $T^\top X^\top \widehat{W}_{\beta\beta} X T = M^\top \widehat{W}_{\beta\beta} M = \Lambda$ e $T^\top T = T T^\top = I$, onde $X T = M$. Temos que

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{MV} &= (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= \left((T^\top)^{-1} T^\top X^\top \widehat{W}_{\beta\beta} X T T^{-1} \right)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= \left((T^\top)^{-1} (T^\top X^\top \widehat{W}_{\beta\beta} X T) T^{-1} \right)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= \left((T^\top)^{-1} \Lambda T^{-1} \right)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T \Lambda^{-1} T^\top X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T \Lambda^{-1} M^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T \widehat{\boldsymbol{\alpha}},
\end{aligned}$$

em que $\widehat{\boldsymbol{\alpha}} = \Lambda^{-1} M^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^*$, é um vetor de dimensão $(p+1) \times 1$.

Sem perda de generalidade,

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_R &= (X^\top \widehat{W}_{\beta\beta} X + kI)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= ((T^\top)^{-1} \Lambda T^{-1} + kI)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= ((T^\top)^{-1} \Lambda T^{-1} + T T^{-1} kI)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= ((T^\top)^{-1} \Lambda T^{-1} + (T^\top)^{-1} kT)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T (\Lambda + kI)^{-1} T^\top X^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T (\Lambda + kI)^{-1} M^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^* \\
&= T (\Lambda + kI)^{-1} \Lambda \Lambda^{-1} M^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^*.
\end{aligned}$$

Como $\widehat{\boldsymbol{\alpha}} = \Lambda^{-1} M^\top \widehat{W}_{\beta\beta} \widehat{\boldsymbol{z}}_1^*$ temos

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_R &= T (\Lambda + kI)^{-1} \Lambda \widehat{\boldsymbol{\alpha}} \\
&= T \widehat{\boldsymbol{\alpha}}_R,
\end{aligned}$$

em que $\widehat{\boldsymbol{\alpha}}_R = (\Lambda + kI)^{-1} \Lambda \widehat{\boldsymbol{\alpha}}$ para $k > 0$.

Segundo [Dorugade \(2014\)](#), estimamos $\widehat{\boldsymbol{\alpha}}$ no lugar de $\widehat{\boldsymbol{\beta}}_{MV}$, devido à sua simplicidade e facilidade de cálculo, conseqüentemente estimamos $\widehat{\boldsymbol{\alpha}}_R$ no lugar de $\widehat{\boldsymbol{\beta}}_R$. Sendo assim, para continuidade do nosso estudo, os próximos resultados são em torno de $\widehat{\boldsymbol{\alpha}}$ e $\widehat{\boldsymbol{\alpha}}_R$.

Para encontrarmos o EQM dos estimadores $\widehat{\boldsymbol{\alpha}}$ e $\widehat{\boldsymbol{\alpha}}_R$, precisamos simplificar algumas expressões. Seguindo a aproximação de [Segerstedt \(1992\)](#), podemos reescrever $\widehat{\boldsymbol{\alpha}}_R$ da seguinte forma

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}_R &= (\Lambda^* + kI)^{-1} \Lambda^* \widehat{\boldsymbol{\alpha}} \\
&= R \widehat{\boldsymbol{\alpha}},
\end{aligned} \tag{2.2}$$

em que $\Lambda^* = (T^\top X^\top W_{\beta\beta} X T)$ e $R = (\Lambda^* + kI)^{-1} \Lambda^*$. Como assintoticamente $\widehat{\boldsymbol{\beta}}_{MV} \sim N_{(p+1)}(\boldsymbol{\beta}, K^{\beta\beta})$, temos as seguintes implicações

$$\begin{aligned}
T^\top \widehat{\boldsymbol{\beta}}_{MV} &\sim N_{(p+1)}(T^\top \boldsymbol{\beta}, T^\top K^{\beta\beta} T) \\
\Leftrightarrow T^\top T \widehat{\boldsymbol{\alpha}} &\sim N_{(p+1)}(T^\top \boldsymbol{\beta}, T^\top K^{\beta\beta} T) \\
\Leftrightarrow \widehat{\boldsymbol{\alpha}} &\sim N_{(p+1)}(T^\top \boldsymbol{\beta}, T^\top K^{\beta\beta} T) \\
\Leftrightarrow \widehat{\boldsymbol{\alpha}} &\sim N_{(p+1)}(\boldsymbol{\alpha}_{(p+1) \times 1}, K^{\alpha\alpha}).
\end{aligned}$$

Segue que

$$\begin{aligned}
\text{EQM}(\hat{\boldsymbol{\alpha}}) &= E[(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^\top (\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})] \\
&= E \left[\begin{pmatrix} (\hat{\alpha}_0 - \alpha_0) \\ (\hat{\alpha}_1 - \alpha_1) \\ \vdots \\ (\hat{\alpha}_p - \alpha_p) \end{pmatrix}^\top \begin{pmatrix} (\hat{\alpha}_0 - \alpha_0) \\ (\hat{\alpha}_1 - \alpha_1) \\ \vdots \\ (\hat{\alpha}_p - \alpha_p) \end{pmatrix} \right] \\
&= E[(\hat{\alpha}_0 - \alpha_0)^2 + (\hat{\alpha}_1 - \alpha_1)^2 + \cdots + (\hat{\alpha}_p - \alpha_p)^2] \\
&= E \left[\sum_{j=0}^p (\hat{\alpha}_j - \alpha_j)^2 \right] \\
&= \sum_{j=0}^p E[(\hat{\alpha}_j - \alpha_j)^2].
\end{aligned}$$

Assintoticamente temos

$$\text{EQM}(\hat{\boldsymbol{\alpha}}) = \sum_{j=0}^p \text{Var}(\hat{\alpha}_j),$$

ou seja,

$$\text{EQM}(\hat{\boldsymbol{\alpha}}) = \text{tr} [\text{Cov}(\hat{\boldsymbol{\alpha}})] = \text{tr} [K^{\alpha\alpha}].$$

Segue que o EQM do estimador de $\boldsymbol{\alpha}_R$ é dado por

$$\begin{aligned}
\text{EQM}(\hat{\boldsymbol{\alpha}}_R) &= E[(\hat{\boldsymbol{\alpha}}_R - \boldsymbol{\alpha})^\top (\hat{\boldsymbol{\alpha}}_R - \boldsymbol{\alpha})] \\
&= E[(R\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^\top (R\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})] \\
&= E[(\hat{\boldsymbol{\alpha}}^\top R^\top - \boldsymbol{\alpha}^\top R^\top)(R\hat{\boldsymbol{\alpha}} - R\boldsymbol{\alpha})] + (\boldsymbol{\alpha}^\top R^\top - \boldsymbol{\alpha}^\top)(R\boldsymbol{\alpha} - \boldsymbol{\alpha}) \\
&= E[(R\hat{\boldsymbol{\alpha}} - R\boldsymbol{\alpha})^\top (R\hat{\boldsymbol{\alpha}} - R\boldsymbol{\alpha})] + \boldsymbol{\alpha}^\top (R - I)^\top (R - I)\boldsymbol{\alpha}.
\end{aligned}$$

Assintoticamente temos

$$\begin{aligned}
\text{EQM}(\hat{\boldsymbol{\alpha}}_R) &= \text{tr} [\text{Cov}(R\hat{\boldsymbol{\alpha}})] + \boldsymbol{\alpha}^\top (R - I)^\top (R - I)\boldsymbol{\alpha} \\
&= \text{tr} [R\text{Cov}(\hat{\boldsymbol{\alpha}})R^\top] + \boldsymbol{\alpha}^\top (R - I)^\top (R - I)\boldsymbol{\alpha}.
\end{aligned}$$

Podemos reescrever R como

$$R = I - k(\Lambda^* + kI)^{-1}. \quad (2.3)$$

De fato, se multiplicarmos $R = (\Lambda^* + kI)^{-1}\Lambda^*$ por $(\Lambda^* + kI)$ em ambos os lados, temos

$$\begin{aligned}
(\Lambda^* + kI)R &= (\Lambda^* + kI)(\Lambda^* + kI)^{-1}\Lambda^* \\
\iff (\Lambda^* + kI)R &= \Lambda^* \\
\iff (\Lambda^* + kI)R &= (\Lambda^* + kI) - kI \\
\iff (\Lambda^* + kI)^{-1}(\Lambda^* + kI)R &= (\Lambda^* + kI)^{-1}(\Lambda^* + kI) - (\Lambda^* + kI)^{-1}kI \\
\iff R &= I - kI(\Lambda^* + kI)^{-1}
\end{aligned}$$

Logo,

$$R - I = -k(\Lambda^* + kI)^{-1}. \quad (2.4)$$

Substituindo a (2.4) em (2.3) temos

$$\begin{aligned}
\text{EQM}(\widehat{\boldsymbol{\alpha}}_R) &= \text{tr} \{ RCov(\widehat{\boldsymbol{\alpha}})R^\top \} + \boldsymbol{\alpha}^\top (-k(\Lambda^* + kI)^{-1})^\top (-k(\Lambda^* + kI)^{-1}) \boldsymbol{\alpha} \\
&= \text{tr} \{ (\Lambda^* + kI)^{-1}\Lambda^* \text{Cov}(\widehat{\boldsymbol{\alpha}})(\Lambda^* + kI)^{-1}\Lambda^* \} + k^2 \boldsymbol{\alpha}^\top ((\Lambda^* + kI)^{-1})^\top (\Lambda^* + kI)^{-1} \boldsymbol{\alpha} \\
&= \text{tr} \{ (\Lambda^* + kI)^{-1}\Lambda^* \text{Cov}(\widehat{\boldsymbol{\alpha}})(\Lambda^* + kI)^{-1}\Lambda^* \} + k^2 \boldsymbol{\alpha}^\top (\Lambda^* + kI)^{-2} \boldsymbol{\alpha} \\
&= \gamma_{11}(k) + \gamma_{22}(k).
\end{aligned}$$

Para expandir a função $\gamma_{11}(k)$, vamos reescrever a matriz $K^{\alpha\alpha}$ da seguinte forma

$$\begin{aligned}
K^{\alpha\alpha} &= T^\top K^{\beta\beta} T \\
&= T^\top \left[D^{-1} + \frac{\mathbf{B}\mathbf{B}^\top}{h} \right] T \\
&= T^\top \left[(X^\top W_{\beta\beta} X)^{-1} + \frac{(X^\top W_{\beta\beta} X)^{-1} X D(a) \mathbf{s} \mathbf{s}^\top D(a) X^\top (X^\top W_{\beta\beta} X)^{-1}}{\text{tr}[D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}} \right] T.
\end{aligned}$$

Como T é uma matriz ortogonal, temos que

$$\begin{aligned}
K^{\alpha\alpha} &= (T^\top X^\top W_{\beta\beta} X T)^{-1} + \frac{(X^\top W_{\beta\beta} X T)^{-1} X D(a) \mathbf{s} \mathbf{s}^\top D(a) X^\top (T^\top X^\top W_{\beta\beta} X)^{-1}}{\text{tr}[D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}} \\
&= (\Lambda^*)^{-1} + \frac{(X^\top W_{\beta\beta} X T)^{-1} T T^\top X D(a) \mathbf{s} \mathbf{s}^\top D(a) X^\top T T^\top (T^\top X^\top W_{\beta\beta} X)^{-1}}{\text{tr}[D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}} \\
&= (\Lambda^*)^{-1} + \frac{(T^\top X^\top W_{\beta\beta} X T)^{-1} T^\top X D(a) \mathbf{s} \mathbf{s}^\top D(a) X^\top T (T^\top X^\top W_{\beta\beta} X T)^{-1}}{\text{tr}[D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}} \\
&= (\Lambda^*)^{-1} + \frac{(\Lambda^*)^{-1} T^\top X D(a) \mathbf{s} \mathbf{s}^\top D(a) X^\top T (\Lambda^*)^{-1}}{\text{tr}[D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}} \\
&= \begin{pmatrix} l_{00} & l_{01} & \cdots & l_{0p} \\ l_{10} & l_{11} & \cdots & l_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p0} & l_{p1} & \cdots & l_{pp} \end{pmatrix} \\
&= L_{(p+1) \times (p+1)}.
\end{aligned}$$

É importante ressaltar que a matrizes $W_{\beta\beta}$ e \mathbf{s} dependem dos parâmetros δ e β . Conseqüentemente a matriz L depende dos mesmos parâmetros.

Então,

$$\begin{aligned}
\gamma_{11}(k) &= \text{tr} \left\{ (\Lambda^* + kI)^{-1} \Lambda^* \begin{pmatrix} l_{00} & l_{01} & \cdots & l_{0p} \\ l_{10} & l_{11} & \cdots & l_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p0} & l_{p1} & \cdots & l_{pp} \end{pmatrix} (\Lambda^* + kI)^{-1} \Lambda^* \right\} \\
&= \text{tr} \left\{ \begin{pmatrix} \frac{\lambda_0^*}{(\lambda_0^* + k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\lambda_p^*}{(\lambda_p^* + k)} \end{pmatrix} \begin{pmatrix} l_{00} & l_{01} & \cdots & l_{0p} \\ l_{10} & l_{11} & \cdots & l_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p0} & l_{p1} & \cdots & l_{pp} \end{pmatrix} \begin{pmatrix} \frac{\lambda_0^*}{(\lambda_0^* + k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\lambda_p^*}{(\lambda_p^* + k)} \end{pmatrix} \right\}
\end{aligned}$$

$$\begin{aligned}
&= \operatorname{tr} \left\{ \begin{pmatrix} \frac{\lambda_0^* l_{00}}{(\lambda_0^* + k)} & \frac{\lambda_0^* l_{01}}{(\lambda_0^* + k)} & \cdots & \frac{\lambda_0^* l_{0p}}{(\lambda_0^* + k)} \\ \frac{\lambda_1^* l_{10}}{(\lambda_1^* + k)} & \frac{\lambda_1^* l_{11}}{(\lambda_1^* + k)} & \cdots & \frac{\lambda_1^* l_{1p}}{(\lambda_1^* + k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\lambda_p^* l_{p0}}{(\lambda_p^* + k)} & \frac{\lambda_p^* l_{p1}}{(\lambda_p^* + k)} & \cdots & \frac{\lambda_p^* l_{pp}}{(\lambda_p^* + k)} \end{pmatrix} \begin{pmatrix} \frac{\lambda_0^*}{(\lambda_0^* + k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\lambda_p^*}{(\lambda_p^* + k)} \end{pmatrix} \right\} \\
&= \operatorname{tr} \left\{ \begin{pmatrix} \frac{\lambda_0^{*2} l_{00}}{(\lambda_0^* + k)^2} & \frac{\lambda_0^* l_{01} \lambda_1^*}{(\lambda_0^* + k)(\lambda_1^* + k)} & \cdots & \frac{\lambda_0^* l_{0p} \lambda_p^*}{(\lambda_0^* + k)(\lambda_p^* + k)^2} \\ \frac{\lambda_1^* l_{10} \lambda_0^*}{(\lambda_1^* + k)(\lambda_0^* + k)} & \frac{\lambda_1^{*2} l_{11}}{(\lambda_1^* + k)^2} & \cdots & \frac{\lambda_1^* l_{1p} \lambda_p^*}{(\lambda_1^* + k)(\lambda_p^* + k)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\lambda_p^* l_{p0} \lambda_0^*}{(\lambda_p^* + k)(\lambda_0^* + k)} & \frac{\lambda_p^* l_{p1} \lambda_1^*}{(\lambda_p^* + k)(\lambda_1^* + k)} & \cdots & \frac{\lambda_p^{*2} l_{pp}}{(\lambda_p^* + k)^2} \end{pmatrix} \right\} \\
&= \frac{\lambda_0^{*2} l_{00}}{(\lambda_0^* + k)^2} + \frac{\lambda_1^{*2} l_{11}}{(\lambda_1^* + k)^2} + \cdots + \frac{\lambda_p^{*2} l_{pp}}{(\lambda_p^* + k)^2}. \\
&= \sum_{j=0}^p \frac{\lambda_j^{*2} l_{jj}}{(\lambda_j^* + k)^2}.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\gamma_{22}(k) &= k^2 \boldsymbol{\alpha}^\top (\Lambda^* + kI)^{-2} \boldsymbol{\alpha} \\
&= k^2 \begin{pmatrix} \alpha_0 & \cdots & \alpha_p \end{pmatrix} \begin{pmatrix} \left(\frac{1}{\lambda_0^* + k}\right)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(\frac{1}{\lambda_p^* + k}\right)^2 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_p \end{pmatrix} \\
&= k^2 \begin{pmatrix} \alpha_0 & \cdots & \alpha_p \end{pmatrix} \begin{pmatrix} \left(\frac{1}{\lambda_0^* + k}\right)^2 \alpha_0 \\ \vdots \\ \left(\frac{1}{\lambda_p^* + k}\right)^2 \alpha_p \end{pmatrix} \\
&= k^2 \left(\alpha_0^2 \left(\frac{1}{\lambda_0^* + k}\right)^2 + \cdots + \alpha_p^2 \left(\frac{1}{\lambda_p^* + k}\right)^2 \right) \\
&= k^2 \sum_{j=0}^p \frac{\alpha_j^2}{(\lambda_j^* + k)^2}.
\end{aligned}$$

Teorema 2.3.1 $\gamma_{11}(k)$ é uma função de k contínua e monotonicamente decrescente.

Prova: A primeira derivada da função $\gamma_{11}(k)$ em relação a k apresenta a seguinte

forma

$$\frac{\partial \gamma_{11}(k)}{\partial k} = -2 \sum_{j=0}^p \frac{\lambda_j^{*2} l_{jj}}{(\lambda_j^* + k)^3}.$$

Observe que $\frac{\partial \gamma_{11}(k)}{\partial k} < 0$, pois $\lambda_j^* > 0$, $k > 0$ e $l_{jj} > 0$. Além disso, a função é contínua em todos os pontos do seu domínio. Portanto, concluímos que $\gamma_{11}(k)$ é uma função contínua e monotonicamente decrescente.

Teorema 2.3.2 $\gamma_{22}(k)$ é uma função de k contínua e monotonicamente crescente.

Prova: A primeira derivada da função $\gamma_{22}(k)$ possui a forma

$$\begin{aligned} \frac{\partial \gamma_{22}(k)}{\partial k} &= 2k \sum_{j=0}^p \frac{\alpha_j^2}{(\lambda_j^* + k)^2} - 2k^2 \sum_{j=0}^p \frac{\alpha_j^2}{(\lambda_j^* + k)^3} \\ &= \sum_{j=0}^p \frac{(2k\lambda_j^* + 2k^2 - 2k^2) \alpha_j^2}{(\lambda_j^* + k)^3} \\ &= 2k \sum_{j=0}^p \frac{\lambda_j^* \alpha_j^2}{(\lambda_j^* + k)^3}. \end{aligned}$$

A função $\gamma_{22}(k)$ é contínua em todo seu domínio. Além disso, a função é monotonicamente crescente, pois $2k \sum_{j=0}^p \frac{\lambda_j^* \alpha_j^2}{(\lambda_j^* + k)^3} > 0$, uma vez que $\lambda_j^* > 0$.

Teorema 2.3.3 (Teorema da existência) Sempre existe um $k > 0$ tal que $EQM(\widehat{\beta}_R) < EQM(\widehat{\beta}_{MV})$.

Prova: Temos que $\widehat{\beta}_R = T\widehat{\alpha}_R$, então

$$\begin{aligned} EQM(\widehat{\beta}_R) &= E \left[(\widehat{\beta}_R - \beta)^\top (\widehat{\beta}_R - \beta) \right] \\ &= E \left[(T\widehat{\alpha}_R - T\alpha)^\top (T\widehat{\alpha}_R - T\alpha) \right] \\ &= E \left[(T(\widehat{\alpha}_R - \alpha))^\top T(\widehat{\alpha}_R - \alpha) \right] \\ &= E \left[(\widehat{\alpha}_R - \alpha)^\top T^\top T(\widehat{\alpha}_R - \alpha) \right] \\ &= E \left[(\widehat{\alpha}_R - \alpha)^\top (\widehat{\alpha}_R - \alpha) \right] \\ &= EQM(\widehat{\alpha}_R), \end{aligned}$$

em que a penúltima igualdade decorre do fato de que $T^\top T = I$. Notemos, também que $\text{EQM}(\widehat{\boldsymbol{\beta}}_{MV}) = \text{EQM}(\widehat{\boldsymbol{\alpha}})$ quando $k = 0$.

Nos Teoremas 2.3.1 e 2.3.2, mostramos que $\gamma_{11}(k)$ e $\gamma_{22}(k)$ são funções de k monotonicamente decrescente e crescente, respectivamente, uma vez que a primeira derivada de $\gamma_{11}(k)$ é sempre não positiva e a primeira derivada de $\gamma_{22}(k)$ é sempre não negativa. Para finalizar a demonstração, é necessário mostrar que sempre existe um $k > 0$ tal que $\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))/\partial k < 0$.

Segue que a primeira derivada do $\text{EQM}(\widehat{\boldsymbol{\alpha}}_R)$ em relação a k é

$$\begin{aligned} \frac{\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))}{\partial k} &= \frac{\partial(\gamma_{11}(k) + \gamma_{22}(k))}{\partial k} \\ &= \frac{\partial\gamma_{11}(k)}{\partial k} + \frac{\partial\gamma_{22}(k)}{\partial k} \\ &= -2 \sum_{j=0}^p \frac{\lambda_j^* l_{jj}}{(\lambda_j^* + k)^3} + 2k \sum_{j=0}^p \frac{\lambda_j^* \alpha_j^2}{(\lambda_j^* + k)^3} \\ &= -2 \sum_{j=0}^p \frac{(\lambda_j^* l_{jj} - k \alpha_j^2) \lambda_j^*}{(\lambda_j^* + k)^3}. \end{aligned}$$

Note que $\lambda_j^* > 0$ e $k > 0$, então

$$\frac{\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))}{\partial k} < 0 \iff \lambda_j^* l_{jj} - k \alpha_j^2 > 0 \implies k < \frac{\lambda_j^* l_{jj}}{\alpha_j^2}.$$

Observe que $\alpha_{\text{máx}} > \alpha_j, \forall j \neq \text{máx}$.

Logo,

$$\frac{1}{\alpha_{\text{máx}}} < \frac{1}{\alpha_j} \implies \frac{\lambda_j^* l_{jj}}{\alpha_{\text{máx}}^2} < \frac{\lambda_j^* l_{jj}}{\alpha_j^2}, \forall j.$$

Então

$$k < \frac{\lambda_j^* l_{jj}}{\alpha_{\text{máx}}^2} < \frac{\lambda_j^* l_{jj}}{\alpha_j^2}.$$

Isso porque $l_{jj} > 0$, já que representa os elementos da diagonal principal da matriz $K^{\alpha\alpha}$. Portanto, existe um $k < \lambda_j^* l_{jj} / \alpha_{\text{máx}}^2$ tal que $\text{EQM}(\widehat{\boldsymbol{\beta}}_R) < \text{EQM}(\widehat{\boldsymbol{\beta}}_{MV})$.

Podemos definir um valor ideal para o parâmetro k que minimiza o $\text{EQM}(\widehat{\boldsymbol{\alpha}}_R)$. Para encontrar esse valor precisamos resolver a igualdade $\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))/\partial k = 0$. No Teorema 2.3.3, obtemos a igualdade

$$\frac{\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))}{\partial k} = -2 \sum_{j=0}^p \frac{(\lambda_j^* l_{jj} - k \alpha_j^2) \lambda_j^*}{(\lambda_j^* + k)^3}.$$

Então,

$$\frac{\partial(\text{EQM}(\widehat{\boldsymbol{\alpha}}_R))}{\partial k} = 0 \iff \lambda_j^* l_{jj} - k \alpha_j^2 = 0 \implies k = \frac{\lambda_j^* l_{jj}}{\alpha_j^2}. \quad (2.5)$$

2.3.2 Estimação do parâmetro k

Na regressão ridge, um dos esforços mais significativos é encontrar o melhor estimador para o parâmetro k . Na literatura, encontramos diferentes técnicas que foram sugeridas por pesquisadores, entre os quais destacamos: [Hoerl et al. \(1975\)](#), [Nomura \(1988\)](#), [Firinguetti \(1999\)](#), [Kibria \(2003\)](#) e [Khalaf and Shukur \(2005\)](#). Estudos mostram que os melhores estimadores variam de acordo com a distribuição, por esse motivo é importante compararmos qual apresenta o menor erro quadrático médio. A seguir, apresentamos algumas propostas de estimadores para o parâmetro k de $\widehat{\boldsymbol{\beta}}_R$.

1. Seguindo o desenvolvimento do estimador proposto por [Hoerl and Kennard \(1970\)](#), substituímos λ_j , l_{jj} e α_j por seus estimadores $\widehat{\lambda}_j$, \widehat{l}_{jj} e $\widehat{\alpha}_j$, respectivamente, em (2.5). Logo,

$$\widehat{k}_1 = \frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2},$$

em que $\widehat{\alpha}_j$ é o j -ésimo elemento de $\widehat{\boldsymbol{\alpha}}$, $\widehat{\mu}_j = g^{-1}(\mathbf{x}_j^\top \widehat{\boldsymbol{\beta}})$ e $\widehat{\delta}$ obtido do processo iterativo (1.6).

2. Baseado em [Hoerl et al. \(1975\)](#) obtemos um outro estimador considerando a média harmônica de \widehat{k}_1

$$\widehat{k}_2 = \frac{(p+1)}{\sum_{j=0}^p \left(\frac{\widehat{\alpha}_j^2}{\widehat{\lambda}_j^* \widehat{l}_{jj}} \right)}.$$

3. Seguindo [Nomura \(1988\)](#) obtemos

$$\widehat{k}_3 = \frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2} \left[1 + \left\{ 1 + \left(\frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2} \right)^2 \right\} \right].$$

4. A média aritmética de \widehat{k}_1 ([Kibria \(2003\)](#))

$$\widehat{k}_4 = \frac{1}{(p+1)} \sum_{j=0}^p \left(\frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2} \right).$$

5. A média geométrica de \widehat{k}_1 ([Kibria \(2003\)](#))

$$\widehat{k}_5 = \left(\prod_{j=0}^p \frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2} \right)^{\frac{1}{(p+1)}}.$$

Considerando $m_j = \sqrt{\frac{\widehat{\lambda}_j^* \widehat{l}_{jj}}{\widehat{\alpha}_j^2}}$, propomos os estimadores com base em [Alkhamisi and Shukur \(2008\)](#):

6. $\hat{k}_6 = \max\left(\frac{1}{m_j}\right),$

7. $\hat{k}_7 = \max(m_j),$

8. $\hat{k}_8 = \text{median}\left(\frac{1}{m_j}\right).$

Capítulo 3

Avaliação Numérica

Neste capítulo avaliamos o desempenho dos estimadores propostos para o parâmetro de encolhimento do ER em termos do viés e do EQM através de simulações de Monte Carlo, para diferentes graus de multicolinearidade e tamanhos amostrais.

3.1 Simulação

A variável resposta foi obtida da distribuição BSR, ou seja, $Y_i \sim \text{BSR}(\mu_i, \delta)$, com tamanhos amostrais $n = 15, 20, 30, 50$ e 80 , parâmetro de precisão $\delta = 0, 89, 1, 61$ e 8 , com função de ligação

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}), \text{ com } i = 1, \dots, n, \text{ e } j = 0, 1, \dots, p.$$

As variáveis explicativas correlacionadas foram geradas a partir de:

$$x_{ij} = (1 - \rho^2)^{\frac{1}{2}} Z_{ij} + \rho Z_{ip}, \text{ } i = 1, 2, \dots, n, \text{ e } j = 0, 1, \dots, p,$$

onde Z_{ip} são números pseudo-aleatório normais padrão independentes e ρ é o grau de correlação. O intercepto do modelo foi considerado igual a zero, e os demais parâmetros foram escolhidos tal que $\sum_{j=0}^p \beta_j^2 = 1$, conforme (Kibria (2003) e Mansson and Shukur (2011)). Dois valores para o número de covariáveis 2 e 4 e quatro valores de ρ correspondendo a 0, 85; 0, 90; 0, 95 e 0, 99 foram considerados.

Para analisar o desempenho do ER e do EMV consideramos o viés médio e o EQM, definidos, respectivamente, por

$$\text{Viés} = \sum_{i=1}^{NR} \frac{\hat{\beta}_i - \beta}{NR}, \text{ e EQM} = \sum_{i=1}^{NR} \frac{(\hat{\beta}_i - \beta)^\top (\hat{\beta}_i - \beta)}{NR},$$

em que NR é o número de réplicas de Monte Carlo que consideramos igual a 5000.

3.1.1 Resultados das simulações

As estimativas dos EQM's para os diferentes métodos de estimação para k e combinações de n, ρ, δ e p são apresentadas nas Tabelas 3.1- 3.6. Como podemos observar, em quase todas as situações consideradas e para todas escolhas de k , o ER possui EQM menor do que o do EMV. Nos cenários apresentados identificamos que os componentes que afetam diretamente o desempenho do EQM são o grau de correlação, o tamanho da amostra, o parâmetro de dispersão e as quantidades de variáveis explicativas. Em primeiro lugar, a medida que o grau de correlação aumenta o EQM do ER torna-se cada vez menor, comparados com o EMV, o que mostra a vantagem do ER para a presença de multicolinearidade.

Claramente observamos que para os cenários com $p = 2, \delta = 0,89$ e $1,61$, o ER considerando o \hat{k}_7 , teve melhor desempenho quando temos correlações altas (0,85, 0,90 e 0,95), com exceção para $\rho = 0,95$ e $n = 15$. Já para correlação muito alta (0,99), o \hat{k}_6 apresentou o melhor desempenho (ver Tabelas 3.1 e 3.2). Quando consideramos o parâmetro de precisão $\delta = 8$, Tabela 3.3, temos que, no geral, o estimador ridge com base no \hat{k}_1 teve desempenho levemente melhor para correlações altas e o \hat{k}_7 para correlações muito altas. Considerando $p = 2$, o \hat{k}_3 obteve os piores resultados em comparação com os outros estimadores estudados.

Nos cenários com $p = 4$ observamos que, no geral, o \hat{k}_2 apresenta o pior desempenho, para delta pequeno, e o \hat{k}_7 o melhor, em quase todos os cenários, ver Tabelas 3.4 - 3.6. Em termos de valores de ρ , há um aumento nos valores de EQM do EMV quando o grau de correlação aumenta independentemente do valor de n, δ e p . Também observamos isso para os demais estimadores, com exceção do \hat{k}_6 . Em relação ao número de variáveis explicativas podemos perceber que existe um impacto negativo no EQM, onde há aumento em seus valores quando o número de covariáveis aumenta. Já com relação ao valor de n , geralmente, os valores do EQM diminuem quando n aumenta.

Também obtivemos os valores médios dos vieses dos estimadores com o intuito de verificar se existiria um perda grande em termos de vieses na comparação dos ER e EMV. Concluimos que para os diferentes valores de ρ , números de variáveis explicativas, parâmetros de dispersão e tamanhos da amostra, de modo geral, os vieses dos ER's

quando comparados com os dos EMV's não apresentaram grandes variações. Destacamos que com relação ao parâmetro β_0 , o ER baseado em \widehat{k}_3 apresentou o menor viés. Em relação a β_1 , β_2 , β_3 e β_4 , \widehat{k}_5 , \widehat{k}_6 e \widehat{k}_8 apresentaram menores vieses ou próximos aos dos EMV. Na maioria dos casos \widehat{k}_1 , \widehat{k}_3 e \widehat{k}_4 apresentaram os maiores vieses. Esses resultados podem ser vistos nas Tabelas B.1-B.24 do Apêndice B.

As vantagens da utilização do estimador ridge foram evidenciadas nos resultados apresentados do estudo de simulação de Monte Carlo, pois os valores médios do EQM dos ER's, com exceções de poucos casos descritos, foram menores quando comparado com os valores médios dos EMV's.

Tabela 3.1: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 0,89$.

	EQM ($p = 2$)								
	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$									
15	0,802	0,569	0,563	1,016	0,523	0,590	0,430	0,414	0,566
20	0,587	0,424	0,379	0,900	0,367	0,483	0,394	0,296	0,464
30	0,420	0,258	0,276	0,842	0,232	0,347	0,274	0,199	0,337
50	0,235	0,193	0,150	0,802	0,151	0,214	0,182	0,105	0,213
80	0,149	0,097	0,108	0,685	0,087	0,140	0,129	0,085	0,139
$\rho = 0,90$									
15	1,578	0,546	0,735	0,867	0,562	0,730	0,492	0,448	0,682
20	1,060	0,531	0,521	0,913	0,484	0,649	0,421	0,369	0,645
30	0,676	0,390	0,334	0,882	0,340	0,463	0,352	0,253	0,440
50	0,423	0,232	0,217	0,836	0,186	0,341	0,248	0,136	0,340
80	0,188	0,146	0,125	0,759	0,116	0,174	0,152	0,091	0,173
$\rho = 0,95$									
15	2,505	0,840	1,184	1,120	0,871	0,769	0,434	0,676	0,679
20	1,023	0,520	0,699	0,923	0,493	0,732	0,443	0,408	0,625
30	1,168	0,414	0,462	0,855	0,386	0,560	0,336	0,285	0,556
50	0,621	0,325	0,277	0,863	0,274	0,443	0,291	0,180	0,435
80	0,346	0,161	0,174	0,745	0,133	0,285	0,219	0,099	0,274
$\rho = 0,99$									
15	13,298	1,714	4,001	1,221	2,582	0,265	0,208	0,673	0,237
20	7,377	1,078	2,202	0,980	1,428	0,283	0,187	0,531	0,247
30	6,654	1,107	1,847	0,964	1,371	0,208	0,126	0,488	0,219
50	3,772	0,628	1,013	0,788	0,739	0,276	0,122	0,338	0,317
80	1,698	0,358	0,484	0,781	0,376	0,423	0,196	0,218	0,399

Fonte: De autoria própria.

Tabela 3.2: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 1,61$.

	EQM ($p = 2$)								
	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$									
15	0,621	0,344	0,457	0,839	0,343	0,464	0,330	0,279	0,411
20	0,475	0,278	0,273	0,827	0,235	0,387	0,293	0,185	0,383
30	0,223	0,131	0,154	0,661	0,122	0,200	0,179	0,119	0,199
50	0,200	0,119	0,129	0,708	0,098	0,182	0,156	0,084	0,181
80	0,074	0,057	0,062	0,559	0,054	0,071	0,069	0,056	0,071
$\rho = 0,90$									
15	0,861	0,436	0,439	0,838	0,417	0,551	0,425	0,343	0,539
20	0,334	0,211	0,218	0,798	0,179	0,298	0,254	0,148	0,291
30	0,314	0,227	0,192	0,816	0,181	0,278	0,226	0,130	0,275
50	0,166	0,077	0,112	0,603	0,073	0,153	0,138	0,079	0,152
80	0,115	0,070	0,088	0,550	0,066	0,108	0,101	0,070	0,107
$\rho = 0,95$									
15	1,693	0,532	0,694	0,796	0,585	0,556	0,390	0,429	0,526
20	1,211	0,394	0,485	0,857	0,365	0,596	0,314	0,250	0,590
30	0,891	0,385	0,396	0,931	0,329	0,498	0,314	0,239	0,446
50	0,377	0,159	0,181	0,727	0,132	0,305	0,222	0,096	0,296
80	0,225	0,105	0,128	0,662	0,088	0,200	0,164	0,075	0,198
$\rho = 0,99$									
15	6,825	2,207	3,148	2,181	2,523	0,639	0,317	1,798	0,562
20	3,601	1,386	1,805	1,600	1,494	0,793	0,392	1,180	0,720
30	3,217	0,527	0,961	0,844	0,607	0,350	0,154	0,296	0,278
50	2,536	0,452	0,694	0,751	0,509	0,348	0,143	0,268	0,382
80	1,226	0,243	0,373	0,749	0,237	0,457	0,199	0,140	0,418

Fonte: De autoria própria.

Tabela 3.3: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$ e $\delta = 8$.

		EQM ($p = 2$)								
		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$										
15		0,259	0,144	0,214	0,635	0,161	0,187	0,141	0,146	0,168
20		0,073	0,050	0,063	0,413	0,054	0,071	0,070	0,060	0,071
30		0,046	0,036	0,041	0,407	0,036	0,045	0,044	0,038	0,045
50		0,035	0,028	0,032	0,122	0,030	0,034	0,034	0,032	0,034
80		0,023	0,019	0,021	0,153	0,020	0,023	0,022	0,021	0,023
$\rho = 0,90$										
15		0,176	0,096	0,139	0,565	0,102	0,143	0,124	0,108	0,139
20		0,149	0,069	0,106	0,566	0,069	0,137	0,124	0,079	0,138
30		0,073	0,048	0,059	0,422	0,048	0,070	0,067	0,051	0,069
50		0,038	0,026	0,034	0,273	0,029	0,037	0,036	0,033	0,037
80		0,031	0,021	0,027	0,167	0,024	0,030	0,029	0,027	0,030
$\rho = 0,95$										
15		0,329	0,110	0,181	0,583	0,103	0,275	0,216	0,094	0,264
20		0,282	0,123	0,157	0,679	0,103	0,237	0,192	0,083	0,225
30		0,150	0,114	0,099	0,683	0,089	0,137	0,122	0,064	0,134
50		0,093	0,040	0,068	0,338	0,045	0,087	0,081	0,056	0,087
80		0,060	0,027	0,047	0,285	0,034	0,057	0,055	0,041	0,057
$\rho = 0,99$										
15		1,469	0,287	0,536	0,726	0,324	0,503	0,239	0,203	0,459
20		1,315	0,213	0,450	0,731	0,216	0,496	0,219	0,138	0,397
30		0,687	0,159	0,262	0,690	0,144	0,414	0,230	0,092	0,378
50		0,420	0,098	0,177	0,548	0,088	0,316	0,202	0,064	0,303
80		0,312	0,058	0,140	0,428	0,057	0,253	0,178	0,052	0,246

Fonte: De autoria própria.

Tabela 3.4: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 0,89$.

	EQM ($p = 4$)								
	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$									
15	6,336	0,941	2,565	1,043	1,248	1,657	0,911	0,888	1,508
20	1,738	0,641	0,876	0,940	0,620	1,161	0,859	0,553	1,111
30	1,136	0,636	0,591	0,974	0,561	0,881	0,686	0,483	0,851
50	0,576	0,533	0,327	0,981	0,411	0,507	0,427	0,293	0,492
80	0,394	0,470	0,274	0,981	0,361	0,369	0,330	0,261	0,363
$\rho = 0,90$									
15	5,180	0,777	2,258	1,039	0,942	1,887	1,011	0,802	1,697
20	2,992	0,870	1,559	0,987	1,030	1,619	0,984	0,866	1,433
30	1,858	0,938	1,140	1,268	0,967	1,237	0,869	0,852	1,139
50	0,819	0,414	0,383	0,956	0,323	0,661	0,516	0,249	0,629
80	0,413	0,315	0,231	0,960	0,224	0,375	0,328	0,157	0,368
$\rho = 0,95$									
15	9,220	3,362	5,718	3,381	3,755	2,690	1,247	3,279	2,302
20	5,778	1,695	3,142	1,782	1,961	2,336	1,165	1,717	2,001
30	4,383	0,552	1,498	0,904	0,614	1,357	0,678	0,482	1,220
50	1,462	0,506	0,643	0,928	0,460	1,019	0,717	0,404	0,975
80	1,020	0,373	0,464	0,913	0,330	0,798	0,604	0,280	0,762
$\rho = 0,99$									
15	38,870	1,488	14,320	1,040	3,308	0,742	0,369	0,996	0,630
20	20,823	1,872	7,779	1,676	2,744	1,453	0,753	1,779	1,237
30	16,324	0,804	4,481	0,861	1,342	1,039	0,508	0,671	0,890
50	7,298	0,707	2,211	0,914	0,808	1,298	0,639	0,633	1,103
80	4,558	0,456	1,275	0,848	0,522	1,211	0,584	0,398	1,069

Fonte: De autoria própria.

Tabela 3.5: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 1, 61$.

	EQM ($p = 4$)								
	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0, 85$									
15	1,359	0,577	0,794	0,975	0,545	1,084	0,843	0,478	1,046
20	1,049	0,564	0,609	0,943	0,504	0,838	0,682	0,446	0,804
30	0,865	0,575	0,433	0,974	0,465	0,701	0,529	0,348	0,675
50	0,507	0,540	0,379	1,111	0,434	0,426	0,363	0,323	0,409
80	0,238	0,339	0,206	0,936	0,259	0,221	0,201	0,195	0,213
$\rho = 0, 90$									
15	2,211	0,694	1,114	0,963	0,681	1,523	1,008	0,609	1,453
20	2,332	1,096	1,299	1,362	1,090	1,219	0,897	0,987	1,169
30	0,809	0,438	0,477	1,008	0,368	0,683	0,563	0,299	0,657
50	0,613	0,440	0,366	0,900	0,384	0,497	0,423	0,317	0,483
80	0,421	0,444	0,344	0,974	0,378	0,386	0,333	0,304	0,365
$\rho = 0, 95$									
15	2,706	0,516	1,226	0,900	0,566	1,581	0,983	0,494	1,461
20	2,348	0,538	1,022	0,902	0,547	1,350	0,827	0,479	1,251
30	1,656	0,498	0,679	0,935	0,444	1,097	0,677	0,350	1,028
50	1,189	0,422	0,514	0,868	0,396	0,773	0,567	0,338	0,742
80	0,609	0,413	0,330	0,961	0,333	0,525	0,426	0,265	0,515
$\rho = 0, 99$									
15	25,578	1,131	8,878	0,929	2,248	1,012	0,477	0,947	0,847
20	14,769	0,853	5,019	0,881	1,417	1,121	0,535	0,749	0,937
30	10,704	0,575	3,006	0,812	0,933	1,044	0,408	0,485	0,884
50	5,316	0,536	1,539	0,852	0,607	1,213	0,588	0,475	1,089
80	4,261	1,201	1,955	1,507	1,372	1,253	0,687	0,794	1,100

Fonte: De autoria própria.

Tabela 3.6: Valores médios do EQM quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$ e $\delta = 8$.

	EQM ($p = 4$)								
	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$									
15	0,400	0,404	0,269	0,968	0,309	0,373	0,335	0,220	0,370
20	0,290	0,456	0,233	0,959	0,357	0,275	0,252	0,254	0,275
30	0,225	0,302	0,192	0,936	0,241	0,197	0,176	0,168	0,184
50	0,078	0,383	0,074	0,983	0,258	0,076	0,075	0,121	0,076
80	0,050	0,290	0,048	0,974	0,170	0,050	0,048	0,074	0,049
$\rho = 0,90$									
15	0,400	0,395	0,288	0,939	0,319	0,370	0,339	0,257	0,359
20	0,467	0,481	0,374	1,035	0,410	0,348	0,309	0,341	0,341
30	0,230	0,299	0,175	0,933	0,221	0,218	0,202	0,160	0,214
50	0,114	0,385	0,104	0,974	0,271	0,111	0,107	0,148	0,110
80	0,092	0,143	0,072	0,840	0,096	0,090	0,088	0,064	0,089
$\rho = 0,95$									
15	2,835	0,598	1,079	0,890	0,639	1,153	0,687	0,521	1,021
20	1,000	0,323	0,405	0,922	0,255	0,716	0,477	0,212	0,697
30	0,611	0,421	0,374	0,979	0,360	0,513	0,423	0,281	0,494
50	0,280	0,307	0,204	0,869	0,247	0,258	0,238	0,200	0,258
80	0,156	0,408	0,118	0,989	0,290	0,151	0,139	0,145	0,150
$\rho = 0,99$									
15	5,626	0,585	2,135	0,849	0,762	1,642	0,832	0,612	1,478
20	5,319	0,467	1,782	0,808	0,586	1,363	0,611	0,465	1,171
30	2,607	0,294	0,845	0,775	0,336	1,100	0,608	0,292	0,992
50	1,370	0,351	0,515	0,893	0,318	0,883	0,545	0,276	0,827
80	0,852	0,281	0,356	0,884	0,243	0,651	0,473	0,214	0,625

Fonte: De autoria própria.

3.2 Aplicação

Para ilustrar o desempenho do estimador proposto consideramos um conjunto de dados de fatores climáticos e concentrações de dióxido de nitrogênio, em partes por cem milhões, durante 26 dias de setembro de 1984, medido em um estação de monitoramento na área da baía de São Francisco. Esses dados foram originalmente analisados por [Chatterjee and Hadi \(2009\)](#) e recentemente analisados por [Kurtoğlu and Özkale \(2016\)](#), [Amin et al. \(2020\)](#), e [Lukman et al. \(2020\)](#). Os dados da variável resposta são as concentrações de dióxido de nitrogênio e as variáveis explicativas são velocidade média do vento (em milhas por hora), temperatura máxima (em °F), insolação (em langleys por dia) e fator de estabilidade (em °F).

Primeiramente, realizamos testes de bondade de ajuste para verificar se a distribuição de probabilidade da variável resposta poderia ser assumida pela BS. Com base no teste de Anderson-Darling e Kolmogorov-Smirnov adaptados, cujos níveis descritivos foram, 0,3 e 0,6, respectivamente, podemos dizer que os dados se ajustam bem a um distribuição BS.

Na Tabela 3.7 estão dispostas as correlações amostrais entre as variáveis. A análise desta tabela é útil no sentido de detectar multicolinearidade. Observamos uma correlação relativamente alta entre as variáveis X_1 e X_2 .

Tabela 3.7: Matriz de correlação amostral entre as variáveis

	X_1	X_2	X_3	X_4
X_1	1,00	-0,78	-0,04	-0,26
X_2	-0,78	1,00	0,25	0,54
X_3	-0,04	0,25	1,00	0,55
X_4	-0,26	0,54	0,55	1,00

Fonte: De autoria própria.

Tabela 3.8: Desempenho dos estimadores para os dados da aplicação

Estimadores	Coeficientes Estimados					EQM
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	
MV	0,82679	-0,08949	0,02474	-0,00014	-0,00131	1,90000
k_1	0,00007	-0,00110	0,01100	0,00280	0,00094	0,66000
k_2	0,05500	-0,07200	0,03200	-0,00005	-0,00410	0,01200
k_4	0,00005	-0,00660	0,02100	0,00023	0,00300	0,56000
k_5	0,76000	-0,08800	0,02500	-0,00013	-0,00160	1,60000
k_6	0,32000	-0,07800	0,02900	-0,00008	-0,00310	0,29000
k_7	0,00007	-0,06400	0,03200	-0,00025	-0,00250	0,01200
k_8	0,75000	-0,08800	0,02500	-0,00013	-0,00160	1,60000

Fonte: De autoria própria.

Na Tabela 3.8 estão dispostos os coeficientes estimados, bem como o EQM para cada estimador considerado. Como podemos observar os estimadores ridge tiveram melhores desempenhos quando comparados ao de MV, destaque especial para os estimadores \hat{k}_2 e \hat{k}_7 que apresentaram melhores desempenho. Vale ressaltar que o estimador \hat{k}_3 foi omitido aqui por apresentar pior desempenho no estudo.

Capítulo 4

Conclusão

Nesta dissertação, propomos o ER para os coeficientes do modelo de regressão BSR como uma alternativa para minimizar os efeitos da multicolinearidade. Mostramos que o EQM do ER é menor que o EQM do EMV. Um estudo de simulação de Monte Carlo foi realizado para comparar os desempenhos dos EMV e ER, usando os diferentes estimadores do parâmetro de encolhimento. O critério de avaliação foram valores médios do EQM e vieses dos estimadores. Os resultados mostraram que o tamanho da amostra, o grau de correlação, o parâmetro de dispersão e as quantidades de variáveis explicativas são fatores importantes que influenciam diretamente nas medidas de desempenho.

Os resultados do estudo de simulação indicaram que para todos os casos \hat{k}_5 , \hat{k}_6 , \hat{k}_7 e \hat{k}_8 reduzem substancialmente os valores médios do EQM. Em todos os eventos investigados neste trabalho e para todos os diferentes casos, o EQM dos ER's são menores que o dos EMV's. Destacamos que os melhores desempenhos foram de \hat{k}_6 e \hat{k}_7 , sendo \hat{k}_6 para o grau de correlação alto(0,99) e \hat{k}_7 para os demais valores. Em termos de vieses, mesmo o que a natureza do ER já insira um viés, não observamos aumento significativo quando comparado ao do EMV. Recomendamos a utilização do ER para o modelo BS para as situações em que as variáveis explicativas são altamente correlacionadas.

Finalmente, para trabalhos futuros pretendemos considerar outras funções de ligação, bem como considerar os estimadores jackknifed ridge para o modelo de regressão BS pois esses têm a vantagem de diminuir tanto o viés quanto o EQM.

Apêndice A

Função escore, matriz de informação de Fisher e estimação

Suponha a partição $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \delta)^\top$ e denotaremos o logaritmo da função de verossimilhança por $l(\boldsymbol{\theta})$ do MRBS com respostas independentes, temos

$$\begin{aligned} l(\boldsymbol{\theta}) &= \log \prod_{i=1}^n f(y_i; \mu, \delta) \\ &= \sum_{i=1}^n \log f(y_i; \mu, \delta) \\ &= \sum_{i=1}^n \left\{ \frac{\delta}{2} - \frac{1}{2} \log(16\pi) - \frac{1}{2} \log(\delta + 1) - \frac{3}{2} \log(y_i) - \frac{1}{2} \log(\mu_i) + \log(y_i + \delta y_i + \delta \mu_i) \right. \\ &\quad \left. - \frac{y_i(\delta + 1)}{4\mu_i} - \frac{\delta^2 \mu_i}{4y_i(\delta + 1)} \right\} \\ &= \sum_{i=1}^n l_i(\mu_i, \delta), \end{aligned}$$

$$\text{em que } f(y_i; \mu, \delta) = \frac{\exp(\delta/2)\sqrt{\delta+1}}{4y_i^{3/2}\sqrt{\pi\mu}} \left[y_i + \frac{\delta\mu}{\delta+1} \right] \exp \left\{ -\frac{\delta}{4} \left[\frac{y_i(\delta+1)}{\delta\mu} + \frac{\delta\mu}{y_i(\delta+1)} \right] \right\}.$$

A função Escore para os parâmetros $\boldsymbol{\beta}$ e δ

A função escore para o parâmetro $\boldsymbol{\beta}$ é encontrada ao calculamos inicialmente as derivadas

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j} &= \sum_{i=1}^n \frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_j} \\
&= \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} \frac{1}{g'(\mu_i)} x_{ij} + \frac{\delta}{(y_i + \delta y_i + \delta \mu_i)} \frac{1}{g'(\mu_i)} x_{ij} + \frac{y_i(\delta + 1)}{4\mu_i^2} \frac{1}{g'(\mu_i)} x_{ij} \right. \\
&\quad \left. - \frac{\delta^2}{4y_i(\delta + 1)} \frac{1}{g'(\mu_i)} x_{ij} \right\} \\
&= \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{\delta}{[y_i(\delta + 1) + \delta \mu_i]} + \frac{y_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta + 1)} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n z_i a_i x_{ij},
\end{aligned}$$

em que $z_i = -\frac{1}{2\mu_i} + \frac{\delta}{[y_i(\delta + 1) + \delta \mu_i]} + \frac{y_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta + 1)}$ e $a_i = \frac{1}{g'(\mu_i)}$. Assim, podemos escrever a função escore na forma matricial

$$\mathbf{U}_\beta(\boldsymbol{\theta})_{(p+1) \times 1} = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \mathbf{X}^\top \mathbf{D}(a) \mathbf{z},$$

onde \mathbf{X} é uma matriz de dimensão $n \times (p + 1)$, cujas as linhas são denotadas por \mathbf{x}_i^\top , com $i = 1, \dots, n$, $\mathbf{D}(a) = \text{diag} \{a_1, \dots, a_n\}$ e $\mathbf{z} = (z_1, \dots, z_n)^\top$.

A função escore para o parâmetro δ é dada por

$$U_\delta(\boldsymbol{\theta})_{1 \times 1} = \frac{\partial l(\boldsymbol{\theta})}{\partial \delta},$$

em que

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \delta} &= \sum_{i=1}^n \frac{\partial l_i(\mu_i, \delta)}{\partial \delta} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{2\delta \mu_i 4y_i(\delta + 1) - 4y_i \delta^2 \mu_i}{[4y_i(\delta + 1)]^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{8\delta^2 \mu_i y_i + 8\delta \mu_i y_i - 4y_i \delta^2 \mu_i}{[4y_i(\delta + 1)]^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{4\delta^2 \mu_i y_i + 8\delta \mu_i y_i}{[4y_i(\delta + 1)]^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{4\delta \mu_i y_i (\delta + 2)}{[4y_i(\delta + 1)]^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta + 2)\mu_i}{4y_i(\delta + 1)^2} \right\}.
\end{aligned}$$

Consequentemente

$$U_{\delta}(\boldsymbol{\theta})_{1 \times 1} = \sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta+1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta+2)\mu_i}{4y_i(\delta+1)^2} \right\}.$$

Seja a matriz $D(b) = \text{diag} \{b_1, \dots, b_n\}$, com $b_i = \frac{1}{2} - \frac{1}{2(\delta+1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta+2)\mu_i}{4y_i(\delta+1)^2}$.

Assim,

$$U_{\delta}(\boldsymbol{\theta})_{1 \times 1} = \text{tr} [D(b)].$$

A matriz de informação de Fisher para os parâmetros $\boldsymbol{\beta}$ e δ

Para obtermos a informação de Fisher precisamos das derivadas

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} &= \frac{\partial}{\partial \beta_l} \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j} \right) \\ &= \frac{\partial}{\partial \beta_l} \left(\sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{\delta}{[y_i(\delta+1) + \delta\mu_i]} + \frac{y_i(\delta+1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta+1)} \right\} \frac{1}{g'(\mu_i)} x_{ij} \right) \\ &= \sum_{i=1}^n \left\{ \frac{1}{2\mu_i^2} - \frac{\delta^2}{[y_i(\delta+1) + \delta\mu_i]^2} - \frac{y_i(\delta+1)}{2\mu_i^3} \right\} \frac{1}{g'(\mu_i)^2} x_{ij} x_{il} + \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} \right. \\ &\quad \left. + \frac{\delta}{[y_i(\delta+1) + \delta\mu_i]} + \frac{y_i(\delta+1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta+1)} \right\} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij} x_{il} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2\mu_i^2} - \frac{\delta^2(\delta+1)^{-2}}{\left[y_i + \frac{\delta\mu_i}{(\delta+1)} \right]^2} - \frac{y_i(\delta+1)}{2\mu_i^3} \right\} \frac{1}{g'(\mu_i)^2} x_{ij} x_{il} + \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{\delta(\delta+1)^{-1}}{\left[y_i + \frac{\delta\mu_i}{(\delta+1)} \right]} \right. \\ &\quad \left. + \frac{y_i(\delta+1)}{4\mu_i^2} - \frac{\delta^2}{4y_i(\delta+1)} \right\} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij} x_{il}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta^2} &= \frac{\partial}{\partial \delta} \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \delta} \right) \\ &= \frac{\partial}{\partial \delta} \left(\sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta+1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta+2)\mu_i}{4y_i(\delta+1)^2} \right\} \right) \\ &= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - \frac{(y_i + \mu_i)^2}{(y_i + \delta y_i + \delta \mu_i)^2} - \frac{2(\delta+1)\mu_i 4y_i(\delta+1)^2 - 8y_i(\delta+1)\delta(\delta+2)\mu_i}{[4y_i(\delta+1)^2]^2} \right\} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - \frac{(y_i + \mu_i)^2}{(y_i + \delta y_i + \delta \mu_i)^2} - \frac{8\mu_i y_i(\delta+1)^2 - 8y_i \delta(\delta+2)\mu_i}{(4y_i)^2(\delta+1)^3} \right\} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - \frac{(y_i + \mu_i)^2}{(y_i + \delta y_i + \delta \mu_i)^2} - \frac{\mu_i(\delta+1)^2 - \delta(\delta+2)\mu_i}{(4y_i)^2(\delta+1)^3} \right\} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - \frac{(y_i + \mu_i)^2}{(y_i + \delta y_i + \delta \mu_i)^2} - \frac{\mu_i \delta^2 + 2\mu_i \delta + \mu_i - \delta^2 \mu_i - 2\mu_i \delta}{2y_i(\delta+1)^3} \right\} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - \frac{(y_i + \mu_i)^2}{(y_i + \delta y_i + \delta \mu_i)^2} - \frac{\mu_i}{2y_i(\delta+1)^3} \right\}. \end{aligned}$$

Considerando os valores esperados abaixo

(i) $E(Y) = \mu,$

(ii) $E\left(\frac{1}{Y}\right) = \frac{(\delta + 1)^2}{\mu\delta^2},$

(iii) $E\left(\left(Y + \frac{\delta\mu}{\{\delta + 1\}}\right)^{-1}\right) = \frac{(\delta + 1)}{2\mu\delta},$

(iv) $E\left(\left(Y + \frac{\delta\mu}{\{\delta + 1\}}\right)^{-2}\right) = \mathcal{I}(\delta).$

em que

$$\mathcal{I}(\delta) = \int_0^\infty \frac{1}{4} e^{\frac{\delta}{2}} \sqrt{\delta} e^{-\frac{1}{4}\left(\frac{(\delta+1)y}{\delta\mu} + \frac{\delta\mu}{(\delta+1)y}\right)\delta} \left(y + \frac{\delta\mu}{(\delta+1)}\right)^{-1} \frac{1}{\sqrt{\frac{\pi\delta\mu}{\delta+1}}} y^{-3/2} dy.$$

Observe que,

$$\begin{aligned} & E \left[\underbrace{\sum_{i=1}^n \left\{ \frac{1}{2\mu_i^2} - \frac{\delta^2(\delta+1)^{-2}}{\left[Y_i + \frac{\delta\mu_i}{(\delta+1)}\right]^2} - \frac{Y_i(\delta+1)}{2\mu_i^3} \right\}}_I \frac{1}{g'(\mu_i)^2} x_{ij}x_{il} \right] \quad (\text{A.1}) \\ &= \sum_{i=1}^n \left\{ \frac{1}{2\mu_i^2} - E \left[\frac{\delta^2(\delta+1)^{-2}}{\left[Y_i + \frac{\delta\mu_i}{(\delta+1)}\right]^2} \right] - E \left[\frac{Y_i(\delta+1)}{2\mu_i^3} \right] \right\} \frac{1}{g'(\mu_i)^2} x_{ij}x_{il} \\ &= \sum_{i=1}^n \left\{ \frac{1}{2\mu_i^2} - \frac{\delta^2}{(\delta+1)^2} \mathcal{I}(\delta) - \frac{(\delta+1)}{2\mu_i^2} \right\} \frac{1}{g'(\mu_i)^2} x_{ij}x_{il} \\ &= \sum_{i=1}^n \left\{ -\frac{\delta}{2\mu_i^2} - \frac{\delta^2}{(\delta+1)^2} \mathcal{I}(\delta) \right\} \frac{1}{g'(\mu_i)^2} x_{ij}x_{il}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} & E \left[\underbrace{\sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{\delta(\delta+1)^{-1}}{\left[Y_i + \frac{\delta\mu_i}{(\delta+1)}\right]} + \frac{Y_i(\delta+1)}{4\mu_i^2} - \frac{\delta^2}{4Y_i(\delta+1)} \right\}}_{II} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij}x_{il} \right] \\ &= \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + E \left[\frac{\delta(\delta+1)^{-1}}{\left[Y_i + \frac{\delta\mu_i}{(\delta+1)}\right]} \right] + E \left[\frac{Y_i(\delta+1)}{4\mu_i^2} \right] - E \left[\frac{\delta^2}{4Y_i(\delta+1)} \right] \right\} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij}x_{il} \\ &= \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{\delta(\delta+1)^{-1}(\delta+1)}{2\mu_i\delta} + \frac{(\delta+1)}{4\mu_i} - \frac{\delta^2}{4(\delta+1)} \frac{(\delta+1)^2}{\mu_i\delta^2} \right\} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij}x_{il} \\ &= \sum_{i=1}^n \left\{ -\frac{1}{2\mu_i} + \frac{1}{2\mu_i} + \frac{(\delta+1)}{4\mu_i} - \frac{(\delta+1)}{4\mu_i} \right\} \left(\frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \right) \frac{d\mu_i}{d\eta_i} x_{ij}x_{il} \\ &= 0. \end{aligned}$$

Assim, os valores esperados são dados por

$$\begin{aligned}
E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_i} \right\} &= (I) + (II) \\
&= \sum_{i=1}^n \left\{ -\frac{\delta}{2\mu_i^2} - \frac{\delta^2}{(\delta+1)^2} \mathcal{I}(\delta) \right\} \frac{1}{g'(\mu_i)^2} x_{ij} x_{il} + 0 \\
&= -\sum_{i=1}^n \left\{ \frac{\delta}{2\mu_i^2 \{g'(\mu_i)\}^2} + \frac{\delta^2}{(\delta+1)^2 \{g'(\mu_i)\}^2} \mathcal{I}(\delta) \right\} x_{ij} x_{il} \\
&= -\sum_{i=1}^n v_i x_{ij} x_{il}
\end{aligned}$$

e

$$\begin{aligned}
E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta^2} \right\} &= \sum_{i=1}^n E \left[\frac{1}{2(\delta+1)^2} - \frac{(Y_i + \mu_i)^2}{(Y_i + \delta Y_i + \delta \mu_i)^2} - \frac{\mu_i}{2Y_i(\delta+1)^3} \right] \\
&= \sum_{i=1}^n \left\{ \frac{1}{2(\delta+1)^2} - E \left[\frac{\left(Y_i + \mu_i + \frac{\delta \mu_i}{(\delta+1)} - \frac{\delta \mu_i}{(\delta+1)} \right)^2}{(Y_i + \delta Y_i + \delta \mu_i)^2} \right] - \frac{1}{2(\delta+1)\delta^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - E \left[\frac{\left(\left(Y_i + \frac{\delta \mu_i}{(\delta+1)} \right) + \left(\mu_i - \frac{\delta \mu_i}{(\delta+1)} \right) \right)^2}{(\delta+1)^2 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]^2} \right] \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - E \left[\frac{\left(\left(Y_i + \frac{\delta \mu_i}{(\delta+1)} \right) + \left(\frac{\mu_i(\delta+1) - \delta \mu_i}{(\delta+1)} \right) \right)^2}{(\delta+1)^2 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]^2} \right] \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - E \left[\frac{\left(\left(Y_i + \frac{\delta \mu_i}{(\delta+1)\delta^2} \right) + \left(\frac{\mu_i}{(\delta+1)} \right) \right)^2}{(\delta+1)^2 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]^2} \right] \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - E \left[\frac{\left(Y_i + \frac{\delta \mu_i}{(\delta+1)} \right)^2 + 2 \left(Y_i + \frac{\delta \mu_i}{(\delta+1)} \right) \frac{\mu_i}{(\delta+1)} + \frac{\mu_i^2}{(\delta+1)^2}}{(\delta+1)^2 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]^2} \right] \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - E \left[\frac{1}{(\delta+1)^2} + \frac{2\mu_i}{(\delta+1)^3 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]} + \frac{\mu_i^2}{(\delta+1)^4 \left[Y_i + \frac{\delta \mu_i}{(\delta+1)} \right]^2} \right] \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 - \delta - 1)}{2(\delta+1)^2\delta^2} - \frac{1}{(\delta+1)^2} - \frac{2\mu_i(\delta+1)}{(\delta+1)^3 2\mu_i\delta} - \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\} \\
&= \sum_{i=1}^n \left\{ \frac{\delta^2 - \delta - 1 - 2\delta^2}{2(\delta+1)^2\delta^2} - \frac{1}{(\delta+1)^2\delta} - \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\} \\
&= \sum_{i=1}^n \left\{ \frac{-\delta^2 - \delta - 1}{2(\delta+1)^2\delta^2} - \frac{1}{(\delta+1)^2\delta} - \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\} \\
&= \sum_{i=1}^n \left\{ \frac{-\delta^2 - \delta - 1 - 2\delta}{2(\delta+1)^2\delta^2} - \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\} \\
&= \sum_{i=1}^n \left\{ \frac{-\delta^2 - 3\delta - 1}{2(\delta+1)^2\delta^2} - \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\} \\
&= - \sum_{i=1}^n \left\{ \frac{(\delta^2 + 3\delta + 1)}{2(\delta+1)^2\delta^2} + \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta) \right\}.
\end{aligned}$$

Logo, podemos escrever a informação de Fisher para $\boldsymbol{\beta}$ e δ na forma matricial

$$K_{\beta\beta} = E \left\{ - \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right\} = X^\top W_{\beta\beta} X$$

e

$$\begin{aligned}
K_{\delta\delta} &= -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta^2} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(\delta^2 + 3\delta + 1)}{2(\delta + 1)^2 \delta^2} + \frac{\mu_i^2}{(\delta + 1)^4} \mathcal{I}(\delta) \right\} \\
&= \text{tr} [D(u)],
\end{aligned}$$

onde $W_{\beta\beta} = \text{diag}\{v_1, \dots, v_n\}$, com $v_i = \frac{\delta}{2\mu_i^2 \{g'(\mu_i)\}^2} + \frac{\delta^2}{(\delta+1)^2 \{g'(\mu_i)\}^2} \mathcal{I}(\delta)$ e $D(u) = \text{diag}\{u_1, \dots, u_n\}$, com $u_i = \frac{(\delta^2+3\delta+1)}{2(\delta+1)^2 \delta^2} + \frac{\mu_i^2}{(\delta+1)^4} \mathcal{I}(\delta)$.

A função escore e a informação de Fisher para o parâmetro $\boldsymbol{\theta}$

A função escore para $\boldsymbol{\theta}$ é da forma $\mathbf{U}_{\boldsymbol{\theta}} = (\mathbf{U}_{\beta}^{\top}, U_{\delta})^{\top}$. Para a informação de Fisher, precisamos calcular a seguinte derivada

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \delta} &= \frac{\partial}{\partial \beta_j} \left(\frac{l(\boldsymbol{\theta})}{\partial \delta} \right) \\
&= \frac{\partial}{\partial \beta_j} \left(\sum_{i=1}^n \left\{ \frac{1}{2} - \frac{1}{2(\delta + 1)} + \frac{(y_i + \mu_i)}{(y_i + \delta y_i + \delta \mu_i)} - \frac{y_i}{4\mu_i} - \frac{\delta(\delta + 2)\mu_i}{4y_i(\delta + 1)^2} \right\} \right) \\
&= \sum_{i=1}^n \left\{ \frac{(y_i + \delta y_i + \delta \mu_i) - (y_i + \mu_i)\delta}{(y_i + \delta y_i + \delta \mu_i)^2} \frac{1}{g'(\mu_i)} x_{ij} + \frac{y_i}{4\mu_i^2} \frac{1}{g'(\mu_i)} x_{ij} - \frac{\delta(\delta + 2)}{4y_i(\delta + 1)^2} \frac{1}{g'(\mu_i)} x_{ij} \right\} \\
&= \sum_{i=1}^n \left\{ \frac{(y_i + \delta y_i + \delta \mu_i) - (y_i + \mu_i)\delta}{(y_i + \delta y_i + \delta \mu_i)^2} + \frac{y_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4y_i(\delta + 1)^2} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ \frac{y_i}{(y_i + \delta y_i + \delta \mu_i)^2} + \frac{y_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4y_i(\delta + 1)^2} \right\} \frac{1}{g'(\mu_i)} x_{ij}.
\end{aligned}$$

Calculando a esperança da expressão acima, temos

$$\begin{aligned}
E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \delta} \right\} &= \sum_{i=1}^n \left\{ E \left[\frac{Y_i}{(Y_i + \delta Y_i + \delta \mu_i)^2} \right] + E \left[\frac{Y_i}{4\mu_i^2} \right] - E \left[\frac{\delta(\delta + 2)}{4Y_i(\delta + 1)^2} \right] \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ E \left[\frac{Y_i}{[(\delta + 1)Y_i + \delta \mu_i]^2} \right] + \frac{\mu_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2} \frac{(\delta + 1)^2}{\mu_i \delta^2} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ E \left[\frac{Y_i}{(\delta + 1)^2 \left[Y_i + \frac{\delta \mu_i}{(\delta + 1)} \right]^2} \right] + \frac{1}{4\mu_i} - \frac{(\delta + 2)}{4\mu_i \delta} \right\} \frac{1}{g'(\mu_i)} x_{ij}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left\{ E \left[\frac{Y_i + \frac{\delta\mu_i}{(\delta+1)} - \frac{\delta\mu_i}{(\delta+1)}}{(\delta+1)^2 \left[Y_i + \frac{\delta\mu_i}{(\delta+1)} \right]^2} + \frac{(\delta - \delta - 2)}{4\mu_i\delta} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ E \left[\frac{Y_i + \frac{\delta\mu_i}{(\delta+1)}}{(\delta+1)^2 \left[Y_i + \frac{\delta\mu_i}{(\delta+1)} \right]^2} - E \left[\frac{\frac{\delta\mu_i}{(\delta+1)}}{(\delta+1)^2 \left[Y_i + \frac{\delta\mu_i}{(\delta+1)} \right]^2} \right] \right. \\
&\quad \left. - \frac{1}{2\mu_i\delta} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ E \left[\frac{1}{(\delta+1)^2 \left[Y_i + \frac{\delta\mu_i}{(\delta+1)} \right]} \right] - E \left[\frac{\delta\mu_i}{(\delta+1)^3 \left[Y_i + \frac{\delta\mu_i}{(\delta+1)} \right]^2} \right] \right. \\
&\quad \left. - \frac{1}{2\mu_i\delta} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ \frac{1}{(\delta+1)^2} \frac{(\delta+1)}{2\mu_i\delta} - \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta) - \frac{1}{2\mu_i\delta} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ \frac{1}{(\delta+1)2\mu_i\delta} - \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta) - \frac{1}{2\mu_i\delta} \right\} \frac{1}{g'(\mu_i)} x_{ij} \\
&= \sum_{i=1}^n \left\{ \frac{1 - \delta - 1}{(\delta+1)2\mu_i\delta} - \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta) \right\} \\
&= \sum_{i=1}^n \left\{ -\frac{\delta}{(\delta+1)2\mu_i\delta} - \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta) \right\} \\
&= -\sum_{i=1}^n \left\{ \frac{\delta}{(\delta+1)2\mu_i\delta} + \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta) \right\}.
\end{aligned}$$

Na forma matricial temos $K_{\beta\delta} = [K_{\delta\beta}]^\top = -E \left\{ \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \delta} \right\} = X^\top D(a) \mathbf{s}$ com $\mathbf{s} = (s_1, \dots, s_n)^\top$ e $s_i = \frac{1}{2\mu_i(\delta+1)} + \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{I}(\delta)$.

A matriz de informação de Fisher esperada para o parâmetro $\boldsymbol{\theta}$ é da forma

$$K_{\theta\theta} = \begin{pmatrix} K_{\beta\beta} & K_{\beta\delta} \\ K_{\delta\beta} & K_{\delta\delta} \end{pmatrix} = \begin{pmatrix} X^\top W_{\beta\beta} X & X^\top D(a) \mathbf{s} \\ \mathbf{s}^\top D(a) X & tr [D(u)] \end{pmatrix}.$$

É importante observarmos que a matriz $X^\top W_{\beta\beta} X = D_{(p+1) \times (p+1)}$ é invertível e seu complemento Schur, $h = tr [D(u)] - \mathbf{s}^\top D(a) X D^{-1} X^\top D(a) \mathbf{s}$, um escalar. Assim, para encontrar a inversa da matriz particionada de $K_{\theta\theta}$, basta resolvermos o sistema abaixo

$$\begin{pmatrix} X^\top W_{\beta\beta} X & X^\top D(a)\mathbf{s} \\ \mathbf{s}^\top D(a)X & \text{tr}[D(u)] \end{pmatrix} \underbrace{\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}}_{K_{\theta\theta}^{-1}} = \begin{pmatrix} I_{(p+1)\times(p+1)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}. \quad (\text{A.2})$$

Uma alternativa de proceder é utilizar a eliminação por blocos. Essas operações são análogas as elementares por linha ou coluna, a diferença é que os escalares são substituídos por matrizes. A operação de bloco elementar usada foi subtrairmos $\mathbf{s}^\top D(a)XD^{-1}$ vezes a primeira linha da segunda linha, ou seja,

$$\begin{aligned} L_2^D &= L_2 - \mathbf{s}^\top D(a)XD^{-1}L_1 \\ &= \begin{pmatrix} \mathbf{s}^\top D(a)X & \text{tr}[D(u)] \end{pmatrix} - \mathbf{s}^\top D(a)XD^{-1} \begin{pmatrix} D & X^\top D(a)\mathbf{s} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{s}^\top D(a)X & \text{tr}[D(u)] \end{pmatrix} - \begin{pmatrix} \mathbf{s}^\top D(a)X & \mathbf{s}^\top D(a)XD^{-1}X^\top D(a)\mathbf{s} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & \text{tr}[D(u)] - \mathbf{s}^\top D(a)XD^{-1}X^\top D(a)\mathbf{s} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & \text{tr}[D(u)] - \mathbf{s}^\top D(a)X\mathbf{B} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & h \end{pmatrix}, \end{aligned}$$

em que $\mathbf{B} = D^{-1}X^\top D(a)\mathbf{s}$, com dimensão $(p+1) \times 1$.

Por outro lado,

$$\begin{aligned} L_2^E &= L_2 - \mathbf{s}^\top D(a)XD^{-1}L_1^D \\ &= \begin{pmatrix} \mathbf{0} & 1 \end{pmatrix} - \mathbf{s}^\top D(a)XD^{-1} \begin{pmatrix} I_{(p+1)\times(p+1)} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0} & 1 \end{pmatrix} - \begin{pmatrix} \mathbf{s}^\top D(a)XD^{-1} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} -\mathbf{B}^\top & 1 \end{pmatrix}. \end{aligned}$$

Logo, a aplicação das operações nos dois lados em [A.2](#) apresenta a seguinte forma

$$\begin{aligned} &\begin{pmatrix} D & X^\top D(a)\mathbf{s} \\ \mathbf{0} & h \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} I_{(p+1)\times(p+1)} & \mathbf{0} \\ -\mathbf{B}^\top & 1 \end{pmatrix} \\ \Leftrightarrow &\begin{pmatrix} DR_{11} + X^\top D(a)\mathbf{s}R_{21} & DR_{12} + X^\top D(a)\mathbf{s}R_{22} \\ hR_{21} & hR_{22} \end{pmatrix} = \begin{pmatrix} I_{(p+1)\times(p+1)} & \mathbf{0} \\ -\mathbf{B}^\top & 1 \end{pmatrix}. \end{aligned}$$

A expressão acima é um sistema triangular em bloco. O bloco da primeira e o da segunda linha satisfazem as seguintes igualdades

$$\left\{ \begin{array}{l} DR_{11} + X^\top D(a) \mathbf{s} R_{21} = I_{(p+1) \times (p+1)} \iff DR_{11} = I_{(p+1) \times (p+1)} - X^\top D(a) \mathbf{s} R_{21}; \\ DR_{12} + X^\top D(a) \mathbf{s} R_{22} = \mathbf{0} \iff DR_{12} = -X^\top D(a) \mathbf{s} R_{22}; \\ hR_{21} = -B^\top \iff R_{21} = -\frac{B^\top}{h}; \\ hR_{22} = 1 \iff R_{22} = \frac{1}{h}. \end{array} \right.$$

Resolvemos R_{11} e R_{12} substituindo os expressões de R_{21} e R_{22} :

$$DR_{11} = I_{(p+1) \times (p+1)} + \frac{X^\top D(a) \mathbf{s} B^\top}{h} \iff R_{11} = D^{-1} + \frac{D^{-1} X^\top \mathbf{s} D(a) B^\top}{h} \implies R_{11} = D^{-1} + \frac{B B^\top}{h}.$$

$$DR_{12} = -\frac{X^\top D(a) \mathbf{s}}{h} \iff R_{12} = -\frac{D^{-1} X^\top D(a) \mathbf{s}}{h} \implies R_{12} = -\frac{B}{h}.$$

Portanto, assumindo que $X^\top W_{\beta\beta} X$ e seu complemento Schur são invertíveis, obtemos a seguinte matriz para a inversa de $K_{\theta\theta}$

$$K_{\theta\theta}^{-1} = \begin{pmatrix} D^{-1} + \frac{B B^\top}{h} & -\frac{B}{h} \\ -\frac{B^\top}{h} & \frac{1}{h} \end{pmatrix}.$$

Expansão de $\theta^{(m+1)}$

Podemos expandir $\theta^{(m+1)}$, substituindo as matrizes $[K_{\theta\theta}^{(m)}]^{-1}$, \tilde{X} e $\tilde{W}^{(m)}$.

$$\theta^{(m+1)} = \underbrace{\begin{pmatrix} [D^{(m)}]^{-1} + \frac{B^{(m)} [B^{(m)}]^\top}{h} & -\frac{B^{(m)}}{h} \\ -\frac{[B^{(m)}]^\top}{h} & \frac{1}{h} \end{pmatrix} \begin{pmatrix} X^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} W_{\beta\beta}^{(m)} & D(a)^{(m)} \mathbf{s}^{(m)} \\ [\mathbf{s}^{(m)}]^\top D(a)^{(m)} & tr [D(u)] \end{pmatrix}}_{I'} \mathbf{z}^{*(m)}.$$

Primeiro calculamos a parte I'

$$\begin{aligned} & \begin{pmatrix} [D^{(m)}]^{-1} + \frac{B^{(m)} [B^{(m)}]^\top}{h} & -\frac{B^{(m)}}{h} \\ -\frac{[B^{(m)}]^\top}{h} & \frac{1}{h} \end{pmatrix} \begin{pmatrix} X^\top & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} W_{\beta\beta}^{(m)} & D(a)^{(m)} \mathbf{s}^{(m)} \\ [\mathbf{s}^{(m)}]^\top D(a)^{(m)} & tr [D(u)] \end{pmatrix} \\ &= \begin{pmatrix} [D^{(m)}]^{-1} + \frac{B^{(m)} [B^{(m)}]^\top}{h} & -\frac{B^{(m)}}{h} \\ -\frac{[B^{(m)}]^\top}{h} & \frac{1}{h} \end{pmatrix} \begin{pmatrix} X^\top W_{\beta\beta}^{(m)} & X^\top D(a)^{(m)} \mathbf{s}^{(m)} \\ [\mathbf{s}^{(m)}]^\top D(a)^{(m)} & tr [D(u)] \end{pmatrix} \\ &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \end{aligned}$$

em que

$$M_{11} = [D^{(m)}]^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{B}^{(m)}]^\top X^\top W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h};$$

$$M_{12} = [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h};$$

$$M_{21} = -\frac{[\mathbf{B}^{(m)}]^\top X^\top W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h};$$

$$M_{22} = -\frac{[\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} + \frac{\text{tr}[D(u)]}{h}.$$

Substituindo as matrizes $\mathbf{B}^{(m)} = [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)}$ e $[D^{(m)}]^{-1} = (X^\top W_{\beta\beta}^{(m)} X)^{-1}$ em M_{11} , M_{12} , M_{21} e M_{22} , temos

- $$\begin{aligned} \bullet M_{11} &= [D^{(m)}]^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{B}^{(m)}]^\top X^\top W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= [D^{(m)}]^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X [D^{(m)}]^{-1} X^\top W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X X^{-1} [W_{\beta\beta}^{(m)}]^{-1} [X^\top]^{-1} X^\top W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)} [W_{\beta\beta}^{(m)}]^{-1} W_{\beta\beta}^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} - \frac{\mathbf{B}^{(m)}[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} + \mathbf{0} \\ &= (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)}. \end{aligned}$$

- $$\bullet M_{12} = [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} + \frac{\mathbf{B}^{(m)}[\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h}$$

$$\begin{aligned}
&= \frac{[D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} h}{h} + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \frac{[D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} [\text{tr}[D(u)] - [\mathbf{s}^{(m)}]^\top D(a)^{(m)} X \mathbf{B}^{(m)}]}{h} + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \frac{[D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} \text{tr}[D(u)] - [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} [\mathbf{s}^{(m)}]^\top D(a)^{(m)} X \mathbf{B}^{(m)}}{h} \\
&\quad + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \frac{\mathbf{B}^{(m)} \text{tr}[D(u)] - \mathbf{B}^{(m)} [\mathbf{s}^{(m)}]^\top D(a)^{(m)} X \mathbf{B}^{(m)}}{h} + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \frac{\mathbf{B}^{(m)} \text{tr}[D(u)] - \mathbf{B}^{(m)} [\mathbf{s}^{(m)}]^\top D(a)^{(m)} X [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} \\
&\quad - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} - \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} + \frac{\mathbf{B}^{(m)} [\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} - \frac{\mathbf{B}^{(m)} \text{tr}[D(u)]}{h} \\
&= \mathbf{0}.
\end{aligned}$$

$$\begin{aligned}
\bullet M_{21} &= -\frac{[\mathbf{B}^{(m)}]^\top X^\top W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= -\frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X [D^{(m)}]^{-1} X^\top W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= -\frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= -\frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X X^{-1} [W_{\beta\beta}^{(m)}]^{-1} [X^\top]^{-1} X^\top W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= -\frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)} [W_{\beta\beta}^{(m)}]^{-1} W_{\beta\beta}^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= -\frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} + \frac{[\mathbf{s}^{(m)}]^\top D(a)^{(m)}}{h} \\
&= \mathbf{0}.
\end{aligned}$$

$$\begin{aligned}
\bullet M_{22} &= -\frac{[\mathbf{B}^{(m)}]^\top X^\top D(a)^{(m)} \mathbf{s}^{(m)}}{h} + \frac{\text{tr}[D(u)]}{h} \\
&= \frac{-[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X [D^{(m)}]^{-1} X^\top D(a)^{(m)} \mathbf{s}^{(m)} + \text{tr}[D(u)]}{h} \\
&= \frac{-[\mathbf{s}^{(m)}]^\top D(a)^{(m)} X \mathbf{B}^{(m)} + \text{tr}[D(u)]}{h} \\
&= \frac{\text{tr}[D(u)] - [\mathbf{s}^{(m)}]^\top D(a)^{(m)} X \mathbf{B}^{(m)}}{h} \\
&= \frac{h}{h} = 1.
\end{aligned}$$

Assim, a matriz que corresponde a parte I' possui a forma $\begin{pmatrix} (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$.

Para calcularmos $\mathbf{z}^{*(m)}$, precisamos encontrar inicialmente a inversa da matriz \widetilde{W} tal que

$$\begin{pmatrix} W_{\beta\beta} & D(a)\mathbf{s} \\ \mathbf{s}^\top D(a) & \text{tr}[D(u)] \end{pmatrix} \underbrace{\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}}_{\widetilde{W}^{-1}} = \begin{pmatrix} I_{n \times n} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}. \quad (\text{A.3})$$

Observe que a matriz $W_{\beta\beta}$ é invertível e seu complemento Schur possui a seguinte forma

$$f = \text{tr}[D(u)] - \mathbf{s}^\top D(a) W_{\beta\beta}^{-1} D(a) \mathbf{s},$$

sendo f um escalar.

A operação de bloco elementar usada foi subtrairmos $\mathbf{s}^\top D(a) W_{\beta\beta}^{-1}$ vezes a primeira linha da segunda linha, ou seja,

$$\begin{aligned}
L_2^D &= L_2 - \mathbf{s}^\top D(a) W_{\beta\beta}^{-1} L_1 \\
&= \left(\mathbf{s}^\top D(a) \quad \text{tr}[D(u)] \right) - \mathbf{s}^\top D(a) W_{\beta\beta}^{-1} \begin{pmatrix} W_{\beta\beta} & D(a)\mathbf{s} \end{pmatrix} \\
&= \left(\mathbf{s}^\top D(a) \quad \text{tr}[D(u)] \right) - \left(\mathbf{s}^\top D(a) \quad \mathbf{s}^\top D(a) W_{\beta\beta}^{-1} D(a) \mathbf{s} \right) \\
&= \begin{pmatrix} \mathbf{0} & \text{tr}[D(u)] - \mathbf{s}^\top D(a) W_{\beta\beta}^{-1} D(a) \mathbf{s} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{0} & f \end{pmatrix}.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
L_2^E &= L_2 - \mathbf{s}^\top D(a)W_{\beta\beta}^{-1}L_1 \\
&= \begin{pmatrix} \mathbf{0} & 1 \end{pmatrix} - \mathbf{s}^\top D(a)W_{\beta\beta}^{-1} \begin{pmatrix} I_{n \times n} & \mathbf{0} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{0} & 1 \end{pmatrix} - \begin{pmatrix} \mathbf{s}^\top D(a)W_{\beta\beta}^{-1} I_{n \times n} & \mathbf{0} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{0} & 1 \end{pmatrix} - \begin{pmatrix} \mathbf{s}^\top D(a)W_{\beta\beta}^{-1} & \mathbf{0} \end{pmatrix} \\
&= \begin{pmatrix} -\mathbf{s}^\top D(a)W_{\beta\beta}^{-1} & 1 \end{pmatrix}.
\end{aligned}$$

Assim, a aplicação das operações nos dois lados em A.3 possui a forma

$$\begin{aligned}
&\begin{pmatrix} W_{\beta\beta} & D(a)\mathbf{s} \\ \mathbf{0} & f \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} I_{n \times n} & \mathbf{0} \\ -\mathbf{s}^\top D(a)W_{\beta\beta}^{-1} & 1 \end{pmatrix} \\
\iff &\begin{pmatrix} W_{\beta\beta}P_{11} + D(a)\mathbf{s}P_{21} & W_{\beta\beta}P_{12} + D(a)\mathbf{s}P_{22} \\ fP_{21} & fP_{22} \end{pmatrix} = \begin{pmatrix} I_{n \times n} & \mathbf{0} \\ -\mathbf{s}^\top D(a)W_{\beta\beta}^{-1} & 1 \end{pmatrix}.
\end{aligned}$$

Logo,

$$\begin{cases}
W_{\beta\beta}P_{11} + D(a)\mathbf{s}P_{21} = I_{n \times n} \iff W_{\beta\beta}P_{11} = I_{n \times n} - D(a)\mathbf{s}P_{21}, \\
W_{\beta\beta}P_{12} + D(a)\mathbf{s}P_{22} = \mathbf{0} \iff W_{\beta\beta}P_{12} = -D(a)\mathbf{s}P_{22} \iff P_{12} = -W_{\beta\beta}^{-1}D(a)\mathbf{s}P_{22}, \\
fP_{21} = -\mathbf{s}^\top D(a)W_{\beta\beta}^{-1} \implies P_{21} = -\frac{\mathbf{s}^\top D(a)W_{\beta\beta}^{-1}}{f}, \\
fP_{22} = 1 \iff P_{22} = \frac{1}{f}.
\end{cases}$$

Substituindo P_{21} e P_{22} em P_{11} e P_{12} , respectivamente, temos

$$\begin{aligned}
W_{\beta\beta}P_{11} &= I_{n \times n} + \frac{D(a)\mathbf{s}\mathbf{s}^\top D(a)W_{\beta\beta}^{-1}}{f} \iff P_{11} = W_{\beta\beta}^{-1} + \frac{W_{\beta\beta}^{-1}D(a)\mathbf{s}\mathbf{s}^\top D(a)W_{\beta\beta}^{-1}}{f}, \\
\implies P_{11} &= W_{\beta\beta}^{-1} + \frac{\mathbf{A}\mathbf{A}^\top}{f},
\end{aligned}$$

$$P_{12} = -\frac{W_{\beta\beta}^{-1}D(a)\mathbf{s}}{f} = -\frac{\mathbf{A}}{f},$$

em que $\mathbf{A} = W_{\beta\beta}^{-1}D(a)\mathbf{s}$.

A inversa da matriz \widetilde{W} encontrada possui a seguinte forma

$$\widetilde{W}^{-1} = \begin{pmatrix} W_{\beta\beta}^{-1} + \frac{\mathbf{A}\mathbf{A}^\top}{f} & -\frac{\mathbf{A}}{f} \\ -\frac{\mathbf{A}^\top}{f} & \frac{1}{f} \end{pmatrix}.$$

Então,

$$\begin{aligned} \mathbf{z}^{*(m)} &= \widetilde{X}\boldsymbol{\theta}^{(m)} + [\widetilde{W}^{(m)}]^{-1} \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} X & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^{(m)} \\ \delta^{(m)} \end{pmatrix} + \begin{pmatrix} [W_{\beta\beta}^{(m)}]^{-1} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top}{f} & -\frac{\mathbf{A}^{(m)}}{f} \\ -\frac{[\mathbf{A}^{(m)}]^\top}{f} & \frac{1}{f} \end{pmatrix} \begin{pmatrix} D(a)^{(m)} & \mathbf{0} \\ \mathbf{0} & \text{tr}[D(b)] \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} X\boldsymbol{\beta}^{(m)} \\ \delta^{(m)} \end{pmatrix} + \begin{pmatrix} [W_{\beta\beta}^{(m)}]^{-1}D(a)^{(m)} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} & -\frac{\mathbf{A}^{(m)}\text{tr}[D(b)]}{f} \\ -\frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} & \frac{\text{tr}[D(b)]}{f} \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(m)} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} X\boldsymbol{\beta}^{(m)} \\ \delta^{(m)} \end{pmatrix} + \begin{pmatrix} \left[[W_{\beta\beta}^{(m)}]^{-1}D(a)^{(m)} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} \right] \mathbf{z}^{(m)} - \frac{\mathbf{A}^{(m)}\text{tr}[D(b)]}{f} \\ -\frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix} \\ &= \begin{pmatrix} X\boldsymbol{\beta}^{(m)} + \left[[W_{\beta\beta}^{(m)}]^{-1}D(a)^{(m)} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} \right] \mathbf{z}^{(m)} - \frac{\mathbf{A}^{(m)}\text{tr}[D(b)]}{f} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{z}_1^{*(m)} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix}, \end{aligned}$$

$$\text{em que } \mathbf{z}_1^{*(m)} = X\boldsymbol{\beta}^{(m)} + \left[[W_{\beta\beta}^{(m)}]^{-1}D(a)^{(m)} + \frac{\mathbf{A}^{(m)}[\mathbf{A}^{(m)}]^\top D(a)^{(m)}}{f} \right] \mathbf{z}^{(m)} - \frac{\mathbf{A}^{(m)}\text{tr}[D(b)]}{f}.$$

Substituindo a parte I' e $\mathbf{z}^{*(m)}$ em $\boldsymbol{\theta}^{(m+1)}$ temos

$$\begin{aligned} \boldsymbol{\theta}^{(m+1)} &= \begin{pmatrix} (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}_1^{*(m)} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix} \\ &= \begin{pmatrix} (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} \mathbf{z}_1^{*(m)} \\ \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f} \end{pmatrix}. \end{aligned}$$

Portanto,

$$\boldsymbol{\beta}^{(m+1)} = (X^\top W_{\beta\beta}^{(m)} X)^{-1} X^\top W_{\beta\beta}^{(m)} \mathbf{z}_1^{*(m)}$$

e

$$\delta^{(m+1)} = \delta^{(m)} - \frac{[\mathbf{A}^{(m)}]^\top D(a)^{(m)} \mathbf{z}^{(m)}}{f} + \frac{\text{tr}[D(b)]}{f}.$$

Apêndice B

Tabelas referentes as estimativas dos vieses

Tabela B.1: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$	$\hat{\beta}_0$	-0,150	-0,061	-0,137	-0,002	-0,082	-0,150	-0,146	-0,108	-0,150
	$\hat{\beta}_1$	0,034	-0,400	-0,053	-0,672	-0,320	0,010	-0,039	-0,206	0,012
	$\hat{\beta}_2$	-0,078	-0,310	-0,090	-0,656	-0,224	-0,075	-0,057	-0,123	-0,078
15	$\hat{\beta}_0$	-0,140	-0,107	-0,151	-0,020	-0,122	-0,143	-0,146	-0,141	-0,143
	$\hat{\beta}_1$	0,018	-0,430	-0,118	-0,675	-0,348	-0,031	-0,080	-0,257	-0,043
	$\hat{\beta}_2$	-0,055	-0,258	-0,060	-0,643	-0,177	-0,045	-0,035	-0,087	-0,042
20	$\hat{\beta}_0$	-0,091	-0,075	-0,093	-0,010	-0,084	-0,093	-0,095	-0,091	-0,093
	$\hat{\beta}_1$	-0,029	-0,291	-0,097	-0,638	-0,234	-0,049	-0,076	-0,163	-0,047
	$\hat{\beta}_2$	-0,013	-0,107	-0,002	0,595	-0,053	-0,006	0,006	-0,002	-0,008
30	$\hat{\beta}_0$	-0,055	-0,009	-0,049	0,002	-0,017	-0,054	-0,051	-0,032	-0,054
	$\hat{\beta}_1$	-0,003	-0,278	-0,051	-0,633	-0,223	-0,011	-0,032	-0,137	-0,009
	$\hat{\beta}_2$	-0,025	-0,118	-0,019	-0,602	-0,061	-0,024	-0,013	-0,004	-0,025
50	$\hat{\beta}_0$	-0,040	-0,040	-0,043	-0,013	-0,042	-0,040	-0,041	-0,042	-0,040
	$\hat{\beta}_1$	-0,023	-0,168	-0,067	-0,587	-0,125	-0,032	-0,040	-0,082	-0,031
	$\hat{\beta}_2$	-0,047	-0,079	-0,044	-0,544	-0,058	-0,044	-0,041	-0,043	-0,045
80	$\hat{\beta}_0$	-0,040	-0,040	-0,043	-0,013	-0,042	-0,040	-0,041	-0,042	-0,040
	$\hat{\beta}_1$	-0,023	-0,168	-0,067	-0,587	-0,125	-0,032	-0,040	-0,082	-0,031
	$\hat{\beta}_2$	-0,047	-0,079	-0,044	-0,544	-0,058	-0,044	-0,041	-0,043	-0,045

Fonte: De autoria própria.

Tabela B.2: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,90$	$\widehat{\beta}_0$	-0,170	-0,135	-0,168	-0,058	-0,150	-0,164	-0,163	-0,160	-0,164
	15 $\widehat{\beta}_1$	-0,035	-0,382	-0,113	-0,633	-0,305	-0,084	-0,130	-0,231	-0,086
	$\widehat{\beta}_2$	-0,016	-0,263	-0,045	-0,581	-0,184	-0,007	0,001	-0,099	-0,009
20	$\widehat{\beta}_0$	-0,122	-0,065	-0,113	-0,013	-0,083	-0,120	-0,124	-0,100	-0,118
	$\widehat{\beta}_1$	-0,015	-0,492	-0,141	-0,682	-0,428	-0,064	-0,167	-0,332	-0,058
	$\widehat{\beta}_2$	-0,019	-0,284	-0,031	-0,630	-0,190	-0,011	0,022	-0,079	-0,016
30	$\widehat{\beta}_0$	-0,097	-0,087	-0,108	-0,023	-0,096	-0,100	-0,103	-0,108	-0,100
	$\widehat{\beta}_1$	-0,034	-0,341	-0,085	-0,654	-0,262	-0,045	-0,065	-0,166	-0,043
	$\widehat{\beta}_2$	-0,022	-0,290	-0,050	-0,638	-0,202	-0,028	-0,025	-0,110	-0,034
50	$\widehat{\beta}_0$	-0,050	-0,037	-0,053	-0,004	-0,043	-0,051	-0,054	-0,052	-0,051
	$\widehat{\beta}_1$	0,001	-0,305	-0,071	-0,649	-0,243	-0,020	-0,059	-0,161	-0,015
	$\widehat{\beta}_2$	-0,029	-0,151	-0,021	-0,618	-0,085	-0,021	-0,002	-0,020	-0,025
80	$\widehat{\beta}_0$	-0,038	-0,027	-0,034	-0,005	-0,029	-0,038	-0,036	-0,032	-0,038
	$\widehat{\beta}_1$	-0,017	-0,243	-0,065	-0,619	-0,192	-0,025	-0,041	-0,122	-0,024
	$\widehat{\beta}_2$	-0,017	-0,078	-0,013	-0,577	-0,036	-0,015	-0,008	-0,003	-0,016

Fonte: De autoria própria.

Tabela B.3: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,95$	$\widehat{\beta}_0$	-0,137	0,009	-0,095	0,020	-0,013	-0,123	-0,101	-0,032	-0,120
	15 $\widehat{\beta}_1$	-0,055	-0,452	-0,127	-0,655	-0,394	-0,064	-0,145	-0,294	-0,081
	$\widehat{\beta}_2$	0,027	-0,206	-0,002	-0,544	-0,111	0,001	0,027	-0,041	0,008
20	$\widehat{\beta}_0$	-0,109	-0,049	-0,100	-0,005	-0,063	-0,111	-0,113	-0,084	-0,115
	$\widehat{\beta}_1$	-0,090	-0,431	-0,143	-0,674	-0,364	-0,096	-0,139	-0,265	-0,090
	$\widehat{\beta}_2$	0,022	-0,293	-0,027	-0,648	-0,194	0,009	0,020	-0,083	0,002
30	$\widehat{\beta}_0$	-0,105	-0,092	-0,110	-0,030	-0,104	-0,106	-0,111	-0,112	-0,105
	$\widehat{\beta}_1$	-0,046	-0,404	-0,115	-0,654	-0,348	-0,057	-0,143	-0,257	-0,047
	$\widehat{\beta}_2$	-0,016	-0,170	-0,011	-0,572	-0,088	-0,017	0,031	-0,008	-0,026
50	$\widehat{\beta}_0$	-0,051	-0,029	-0,048	-0,004	-0,034	-0,051	-0,050	-0,041	-0,051
	$\widehat{\beta}_1$	-0,066	-0,363	-0,108	-0,657	-0,311	-0,065	-0,115	-0,218	-0,052
	$\widehat{\beta}_2$	-0,018	-0,211	-0,031	-0,623	-0,130	-0,026	0,004	-0,040	-0,037
80	$\widehat{\beta}_0$	-0,047	-0,041	-0,049	-0,010	-0,045	-0,047	-0,048	-0,048	-0,047
	$\widehat{\beta}_1$	-0,044	-0,227	-0,069	-0,603	-0,183	-0,048	-0,067	-0,121	-0,041
	$\widehat{\beta}_2$	0,007	-0,103	0,004	-0,575	-0,049	0,006	0,018	0,002	0,001

Fonte: De autoria própria.

Tabela B.4: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 0,89$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,99$	$\widehat{\beta}_0$	-0,160	-0,129	-0,164	-0,076	-0,139	-0,168	-0,170	-0,158	-0,165
	$\widehat{\beta}_1$	-0,084	-0,450	-0,086	-0,574	-0,398	-0,046	-0,098	-0,234	-0,072
	$\widehat{\beta}_2$	0,057	0,037	0,023	-0,250	0,097	-0,031	-0,021	0,037	-0,016
15	$\widehat{\beta}_0$	-0,129	-0,111	-0,130	-0,059	-0,122	-0,128	-0,130	-0,129	-0,127
	$\widehat{\beta}_1$	-0,128	-0,458	-0,097	-0,607	-0,405	-0,044	-0,121	-0,262	-0,059
	$\widehat{\beta}_2$	0,082	-0,028	0,013	-0,333	0,043	-0,028	0,008	0,032	-0,019
20	$\widehat{\beta}_0$	-0,093	-0,063	-0,091	-0,034	-0,071	-0,090	-0,087	-0,080	-0,090
	$\widehat{\beta}_1$	0,078	-0,293	0,015	-0,501	-0,223	-0,011	-0,082	-0,144	0,017
	$\widehat{\beta}_2$	-0,098	-0,077	-0,066	-0,345	-0,048	-0,039	0,008	-0,009	-0,064
30	$\widehat{\beta}_0$	-0,058	-0,040	-0,056	-0,020	-0,045	-0,056	-0,056	-0,051	-0,056
	$\widehat{\beta}_1$	0,018	-0,287	-0,018	-0,532	-0,238	0,017	-0,077	-0,145	0,062
	$\widehat{\beta}_2$	-0,040	-0,080	-0,027	-0,398	-0,031	-0,049	0,025	0,011	-0,090
50	$\widehat{\beta}_0$	-0,036	-0,029	-0,038	-0,009	-0,033	-0,038	-0,037	-0,037	-0,039
	$\widehat{\beta}_1$	-0,140	-0,297	-0,097	-0,596	-0,260	-0,051	-0,114	-0,169	-0,015
	$\widehat{\beta}_2$	0,084	-0,111	0,026	-0,507	-0,036	-0,004	0,048	0,021	-0,037
80	$\widehat{\beta}_0$	-0,036	-0,029	-0,038	-0,009	-0,033	-0,038	-0,037	-0,037	-0,039
	$\widehat{\beta}_1$	-0,140	-0,297	-0,097	-0,596	-0,260	-0,051	-0,114	-0,169	-0,015
	$\widehat{\beta}_2$	0,084	-0,111	0,026	-0,507	-0,036	-0,004	0,048	0,021	-0,037

Fonte: De autoria própria.

Tabela B.5: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$	$\widehat{\beta}_0$	-0,096	-0,048	-0,089	-0,005	-0,059	-0,099	-0,098	-0,075	-0,101
	$\widehat{\beta}_1$	-0,015	-0,309	-0,048	-0,638	-0,247	-0,017	-0,040	-0,148	-0,014
	$\widehat{\beta}_2$	-0,021	-0,217	-0,038	-0,625	-0,141	-0,028	-0,017	-0,055	-0,031
15	$\widehat{\beta}_0$	-0,076	-0,071	-0,080	-0,014	0,078	-0,077	-0,080	-0,083	-0,077
	$\widehat{\beta}_1$	0,013	-0,333	-0,077	-0,647	0,267	-0,015	-0,061	-0,183	-0,013
	$\widehat{\beta}_2$	-0,012	-0,144	-0,003	-0,602	-0,075	-0,004	-0,015	-0,008	-0,006
20	$\widehat{\beta}_0$	-0,048	-0,072	-0,059	-0,027	-0,068	-0,052	-0,054	-0,064	-0,052
	$\widehat{\beta}_1$	0,038	-0,161	-0,031	-0,577	-0,102	0,018	0,005	-0,052	0,019
	$\widehat{\beta}_2$	-0,034	-0,077	-0,020	-0,513	-0,051	-0,027	-0,022	-0,027	-0,028
30	$\widehat{\beta}_0$	-0,039	-0,031	-0,038	-0,011	-0,033	-0,039	-0,038	-0,036	-0,039
	$\widehat{\beta}_1$	0,011	-0,210	-0,042	-0,597	-0,161	< 0,001	-0,018	-0,097	0,001
	$\widehat{\beta}_2$	-0,035	-0,059	-0,021	-0,546	-0,027	-0,031	-0,022	-0,006	-0,032
50	$\widehat{\beta}_0$	-0,021	-0,025	-0,020	-0,010	-0,024	-0,021	-0,021	-0,022	-0,021
	$\widehat{\beta}_1$	0,005	-0,108	-0,025	-0,533	-0,069	-0,001	-0,005	-0,038	-0,001
	$\widehat{\beta}_2$	-0,015	-0,016	-0,012	-0,453	-0,009	-0,014	-0,013	-0,008	-0,014
80	$\widehat{\beta}_0$	-0,021	-0,025	-0,020	-0,010	-0,024	-0,021	-0,021	-0,022	-0,021
	$\widehat{\beta}_1$	0,005	-0,108	-0,025	-0,533	-0,069	-0,001	-0,005	-0,038	-0,001
	$\widehat{\beta}_2$	-0,015	-0,016	-0,012	-0,453	-0,009	-0,014	-0,013	-0,008	-0,014

Fonte: De autoria própria.

Tabela B.6: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,90$	$\widehat{\beta}_0$	-0,137	-0,136	-0,154	-0,051	-0,144	-0,141	-0,146	-0,157	-0,142
	$\widehat{\beta}_1$	-0,053	-0,350	-0,138	-0,624	-0,280	-0,096	-0,123	-0,212	-0,098
	$\widehat{\beta}_2$	-0,030	-0,254	-0,037	-0,603	-0,178	-0,016	-0,012	-0,089	-0,017
15	$\widehat{\beta}_0$	-0,085	-0,060	-0,087	-0,010	-0,070	-0,085	-0,086	-0,080	-0,085
	$\widehat{\beta}_1$	-0,026	-0,258	-0,066	-0,626	-0,199	-0,035	-0,051	-0,127	-0,033
	$\widehat{\beta}_2$	0,007	-0,137	-0,003	-0,601	-0,073	0,007	-0,013	-0,014	0,004
20	$\widehat{\beta}_0$	-0,050	-0,023	-0,048	-0,002	-0,029	-0,050	-0,048	-0,039	-0,050
	$\widehat{\beta}_1$	-0,017	-0,295	-0,069	-0,638	-0,237	-0,027	-0,052	-0,153	-0,024
	$\widehat{\beta}_2$	-0,003	-0,144	-0,003	-0,609	-0,077	0,004	0,015	-0,008	0,001
30	$\widehat{\beta}_0$	-0,040	-0,044	-0,045	-0,019	-0,045	-0,041	-0,042	-0,045	-0,041
	$\widehat{\beta}_1$	0,024	-0,119	-0,013	-0,538	0,078	0,016	0,007	-0,036	0,017
	$\widehat{\beta}_2$	-0,043	-0,050	-0,033	-0,496	-0,033	-0,039	-0,034	-0,025	-0,040
50	$\widehat{\beta}_0$	-0,022	-0,029	-0,026	-0,024	-0,027	-0,023	-0,023	-0,026	-0,023
	$\widehat{\beta}_1$	-0,003	-0,136	-0,046	-0,532	-0,094	-0,013	-0,019	-0,060	-0,013
	$\widehat{\beta}_2$	-0,007	-0,012	-0,001	-0,446	-0,001	-0,004	-0,002	0,004	-0,005
80	$\widehat{\beta}_0$	-0,022	-0,029	-0,026	-0,024	-0,027	-0,023	-0,023	-0,026	-0,023
	$\widehat{\beta}_1$	-0,003	-0,136	-0,046	-0,532	-0,094	-0,013	-0,019	-0,060	-0,013
	$\widehat{\beta}_2$	-0,007	-0,012	-0,001	-0,446	-0,001	-0,004	-0,002	0,004	-0,005

Fonte: De autoria própria.

Tabela B.7: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,95$	$\widehat{\beta}_0$	-0,109	-0,093	-0,112	-0,053	-0,099	-0,113	-0,115	-0,110	-0,113
	$\widehat{\beta}_1$	-0,002	-0,325	-0,088	-0,574	-0,248	-0,071	-0,105	-0,188	-0,075
	$\widehat{\beta}_2$	0,001	-0,249	-0,024	-0,545	-0,171	0,017	-0,016	-0,089	0,014
15	$\widehat{\beta}_0$	-0,089	-0,058	-0,088	-0,012	-0,070	-0,088	-0,089	-0,082	-0,088
	$\widehat{\beta}_1$	0,010	-0,375	-0,052	-0,649	-0,319	0,009	-0,091	-0,214	0,027
	$\widehat{\beta}_2$	-0,036	-0,188	-0,027	-0,583	-0,102	-0,045	0,020	-0,015	-0,060
20	$\widehat{\beta}_0$	-0,049	-0,006	-0,040	0,005	-0,015	-0,048	-0,045	-0,026	-0,048
	$\widehat{\beta}_1$	0,018	-0,354	-0,066	-0,663	-0,275	-0,016	-0,065	-0,179	-0,013
	$\widehat{\beta}_2$	-0,050	-0,247	-0,044	-0,639	-0,158	-0,036	-0,011	-0,072	-0,042
30	$\widehat{\beta}_0$	-0,035	-0,031	-0,036	-0,008	-0,033	-0,035	-0,036	-0,036	-0,035
	$\widehat{\beta}_1$	-0,033	-0,230	-0,065	-0,594	-0,190	-0,038	-0,066	-0,126	-0,028
	$\widehat{\beta}_2$	0,012	-0,084	0,014	-0,557	-0,031	0,013	0,029	0,021	0,004
50	$\widehat{\beta}_0$	-0,015	-0,004	-0,013	0,002	-0,007	-0,015	-0,014	-0,010	-0,015
	$\widehat{\beta}_1$	-0,031	-0,204	-0,066	-0,572	-0,170	-0,037	-0,055	-0,116	-0,033
	$\widehat{\beta}_2$	-0,011	-0,038	0,001	-0,520	-0,006	-0,008	0,003	-0,018	-0,011
80	$\widehat{\beta}_0$	-0,015	-0,004	-0,013	0,002	-0,007	-0,015	-0,014	-0,010	-0,015
	$\widehat{\beta}_1$	-0,031	-0,204	-0,066	-0,572	-0,170	-0,037	-0,055	-0,116	-0,033
	$\widehat{\beta}_2$	-0,011	-0,038	0,001	-0,520	-0,006	-0,008	0,003	-0,018	-0,011

Fonte: De autoria própria.

Tabela B.8: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 1,61$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,99$	$\widehat{\beta}_0$	-0,035	0,008	-0,027	0,029	0,001	-0,062	-0,065	-0,012	-0,067
	$\widehat{\beta}_1$	0,267	-0,317	0,084	-0,538	-0,226	0,010	-0,100	-0,162	0,027
	$\widehat{\beta}_2$	-0,296	-0,186	-0,177	-0,438	-0,154	-0,101	-0,034	-0,084	-0,117
15	$\widehat{\beta}_0$	-0,025	0,015	-0,019	0,034	0,008	-0,042	-0,048	-0,005	-0,046
	$\widehat{\beta}_1$	0,028	-0,328	-0,020	-0,592	-0,272	0,026	-0,078	-0,172	0,053
	$\widehat{\beta}_2$	-0,044	-0,116	-0,028	-0,467	-0,052	-0,052	0,026	0,007	-0,076
20	$\widehat{\beta}_0$	-0,059	-0,020	-0,053	-0,008	-0,028	-0,060	-0,053	-0,038	-0,056
	$\widehat{\beta}_1$	0,017	-0,375	-0,038	-0,621	-0,314	0,002	-0,086	-0,204	-0,013
	$\widehat{\beta}_2$	-0,026	-0,130	-0,016	-0,490	-0,045	-0,030	0,028	0,008	-0,021
30	$\widehat{\beta}_0$	-0,034	-0,022	-0,033	-0,008	-0,025	-0,034	-0,031	-0,029	-0,036
	$\widehat{\beta}_1$	0,022	-0,265	-0,013	-0,549	-0,220	0,022	-0,071	-0,129	0,076
	$\widehat{\beta}_2$	-0,037	-0,119	-0,027	-0,460	-0,063	-0,046	0,028	-0,006	-0,095
50	$\widehat{\beta}_0$	-0,027	-0,033	-0,029	-0,011	-0,033	-0,027	-0,029	-0,032	-0,027
	$\widehat{\beta}_1$	-0,007	-0,256	-0,032	-0,590	-0,217	0,012	-0,071	-0,135	0,036
	$\widehat{\beta}_2$	-0,012	-0,096	-0,004	-0,530	-0,028	-0,032	0,036	0,031	-0,054

Fonte: De autoria própria.

Tabela B.9: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$	$\widehat{\beta}_0$	-0,019	-0,024	-0,018	-0,010	-0,022	-0,021	-0,022	-0,020	-0,021
	$\widehat{\beta}_1$	0,021	-0,164	-0,023	-0,561	-0,116	0,011	-0,001	-0,063	0,011
	$\widehat{\beta}_2$	-0,019	-0,028	-0,006	-0,491	-0,008	-0,016	-0,010	0,004	-0,016
15	$\widehat{\beta}_0$	-0,015	-0,017	-0,016	-0,010	-0,017	-0,016	-0,016	-0,016	-0,015
	$\widehat{\beta}_1$	-0,010	-0,047	-0,021	-0,413	-0,032	-0,012	-0,013	-0,022	-0,012
	$\widehat{\beta}_2$	-0,021	-0,045	-0,033	-0,408	-0,035	-0,023	-0,023	-0,030	-0,023
20	$\widehat{\beta}_0$	-0,011	-0,011	-0,012	-0,020	-0,011	-0,011	-0,011	-0,011	-0,011
	$\widehat{\beta}_1$	0,001	-0,055	-0,010	-0,427	-0,033	-0,001	-0,003	-0,016	-0,001
	$\widehat{\beta}_2$	-0,001	-0,002	-0,003	-0,358	< 0,001	-0,002	-0,001	0,001	-0,002
30	$\widehat{\beta}_0$	< 0,001	-0,010	-0,009	-0,023	-0,009	-0,008	-0,008	-0,009	-0,008
	$\widehat{\beta}_1$	< 0,001	-0,027	-0,016	-0,255	-0,020	-0,006	-0,007	-0,014	-0,006
	$\widehat{\beta}_2$	< 0,001	-0,002	0,001	-0,138	0,002	< 0,001	0,001	0,001	< 0,001
50	$\widehat{\beta}_0$	< 0,001	-0,006	-0,006	-0,011	-0,006	-0,006	-0,006	-0,006	-0,007
	$\widehat{\beta}_1$	< 0,001	-0,031	-0,014	-0,315	-0,020	-0,006	-0,007	-0,014	-0,006
	$\widehat{\beta}_2$	< 0,001	0,009	0,004	-0,135	0,006	0,001	0,001	0,003	0,001

Fonte: De autoria própria.

Tabela B.10: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,90$	$\widehat{\beta}_0$	-0,019	-0,024	-0,020	-0,015	-0,022	-0,020	-0,021	-0,021	-0,020
	$\widehat{\beta}_1$	0,033	-0,102	-0,003	-0,512	-0,064	0,026	0,017	-0,024	0,027
	$\widehat{\beta}_2$	-0,062	-0,045	-0,052	-0,450	-0,039	-0,059	-0,055	-0,042	-0,060
15	$\widehat{\beta}_0$	-0,018	-0,017	-0,019	0,003	-0,019	-0,019	-0,019	-0,019	-0,018
	$\widehat{\beta}_1$	0,006	-0,121	-0,032	-0,532	-0,081	-0,003	-0,012	-0,046	-0,002
	$\widehat{\beta}_2$	-0,012	-0,019	-0,003	-0,461	-0,007	-0,008	-0,005	-0,001	-0,009
20	$\widehat{\beta}_0$	-0,011	-0,019	-0,014	-0,032	-0,016	-0,012	-0,012	-0,014	-0,012
	$\widehat{\beta}_1$	-0,001	-0,084	-0,018	0,435	-0,054	-0,004	-0,007	-0,031	-0,005
	$\widehat{\beta}_2$	0,001	0,004	0,002	0,364	0,008	0,001	0,003	0,009	0,002
30	$\widehat{\beta}_0$	-0,008	-0,005	-0,007	0,005	-0,007	-0,008	-0,008	-0,007	-0,008
	$\widehat{\beta}_1$	0,001	-0,038	-0,012	-0,370	-0,023	-0,001	-0,003	-0,013	-0,001
	$\widehat{\beta}_2$	-0,005	0,004	-0,003	-0,262	0,001	-0,005	-0,004	-0,002	-0,005
50	$\widehat{\beta}_0$	-0,006	-0,004	-0,005	0,006	-0,005	-0,005	-0,005	-0,005	-0,005
	$\widehat{\beta}_1$	0,003	-0,030	-0,009	-0,305	-0,017	0,001	-0,001	-0,009	0,001
	$\widehat{\beta}_2$	-0,005	0,009	-0,001	-0,160	0,004	-0,004	-0,003	< 0,001	-0,004

Fonte: De autoria própria.

Tabela B.11: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,95$	$\widehat{\beta}_0$	-0,024	-0,041	-0,028	-0,034	-0,036	-0,025	-0,025	-0,032	-0,025
	$\widehat{\beta}_1$	-0,002	-0,160	-0,028	-0,507	-0,127	-0,007	-0,024	-0,075	-0,002
	$\widehat{\beta}_2$	0,003	-0,033	0,015	-0,459	< 0,001	0,004	0,017	0,029	< 0,001
15	$\widehat{\beta}_0$	-0,018	-0,007	-0,017	0,007	-0,013	-0,018	-0,019	-0,017	-0,018
	$\widehat{\beta}_1$	0,001	-0,174	-0,027	-0,561	-0,128	-0,005	-0,018	-0,072	-0,004
	$\widehat{\beta}_2$	-0,003	-0,072	0,006	-0,528	-0,028	-0,001	0,008	0,012	-0,003
20	$\widehat{\beta}_0$	-0,012	0,001	-0,009	0,011	-0,004	-0,011	-0,011	-0,007	-0,011
	$\widehat{\beta}_1$	0,004	-0,187	-0,024	-0,573	-0,140	-0,001	-0,009	-0,079	-0,003
	$\widehat{\beta}_2$	-0,004	-0,047	0,005	-0,532	-0,011	-0,002	0,002	0,021	-0,001
30	$\widehat{\beta}_0$	-0,008	-0,003	-0,007	0,002	-0,004	-0,008	-0,008	-0,006	-0,008
	$\widehat{\beta}_1$	-0,016	-0,088	-0,031	-0,389	-0,069	-0,018	-0,023	-0,045	-0,017
	$\widehat{\beta}_2$	0,010	0,033	0,015	-0,300	0,033	0,011	0,014	0,025	0,010
50	$\widehat{\beta}_0$	-0,003	-0,005	-0,003	0,010	-0,004	-0,003	-0,003	-0,004	-0,003
	$\widehat{\beta}_1$	-0,012	-0,070	-0,027	-0,366	-0,048	-0,015	-0,017	-0,034	-0,015
	$\widehat{\beta}_2$	-0,006	0,020	0,020	-0,258	0,012	-0,005	-0,003	0,006	-0,005

Fonte: De autoria própria.

Tabela B.12: Vieses médios quando $\beta_0 = 0$, $\beta_1 = \sqrt{0,5}$ e $\beta_2 = \sqrt{0,5}$, $\delta = 8$ e $p = 2$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,99$	$\widehat{\beta}_0$	-0,028	-0,031	-0,029	-0,012	-0,032	-0,029	-0,029	-0,032	-0,029
	$\widehat{\beta}_1$	-0,016	-0,232	-0,028	-0,563	-0,196	0,017	-0,052	-0,112	0,044
	$\widehat{\beta}_2$	0,012	-0,078	0,011	-0,490	-0,014	-0,022	0,037	0,036	-0,048
15	$\widehat{\beta}_0$	-0,027	-0,034	-0,022	-0,017	-0,028	-0,022	-0,019	-0,025	-0,022
	$\widehat{\beta}_1$	0,007	-0,232	-0,022	-0,581	-0,195	0,010	-0,053	-0,118	0,014
	$\widehat{\beta}_2$	-0,005	-0,074	0,008	-0,525	-0,009	-0,012	0,038	0,038	-0,016
20	$\widehat{\beta}_0$	-0,011	0,002	-0,010	0,005	-0,001	-0,011	-0,010	-0,006	-0,010
	$\widehat{\beta}_1$	-0,006	-0,203	-0,027	-0,559	-0,169	0,003	-0,045	-0,100	0,022
	$\widehat{\beta}_2$	-0,003	-0,084	0,006	-0,524	-0,024	-0,013	0,027	0,030	-0,030
30	$\widehat{\beta}_0$	-0,009	-0,009	-0,009	-0,007	-0,009	-0,009	-0,009	-0,009	-0,009
	$\widehat{\beta}_1$	-0,013	-0,162	-0,038	-0,486	-0,141	-0,012	-0,045	-0,089	0,004
	$\widehat{\beta}_2$	0,004	-0,032	0,017	-0,437	0,005	0,001	0,027	0,038	-0,012
50	$\widehat{\beta}_0$	-0,005	-0,005	-0,005	0,004	-0,005	-0,005	-0,005	-0,005	-0,004
	$\widehat{\beta}_1$	0,007	-0,100	-0,014	-0,403	-0,088	0,006	-0,017	-0,047	0,019
	$\widehat{\beta}_2$	-0,014	-0,016	< 0,001	-0,359	0,005	-0,014	0,006	0,018	-0,025
80	$\widehat{\beta}_0$	-0,005	-0,005	-0,005	0,004	-0,005	-0,005	-0,005	-0,005	-0,004
	$\widehat{\beta}_1$	0,007	-0,100	-0,014	-0,403	-0,088	0,006	-0,017	-0,047	0,019
	$\widehat{\beta}_2$	-0,014	-0,016	< 0,001	-0,359	0,005	-0,014	0,006	0,018	-0,025

Fonte: De autoria própria.

Tabela B.13: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,85$	$\widehat{\beta}_0$	-0,286	-0,146	-0,271	-0,043	-0,190	-0,274	-0,267	-0,220	-0,272
	$\widehat{\beta}_1$	-0,101	-0,468	-0,215	-0,651	-0,389	-0,219	-0,291	-0,353	-0,221
	15 $\widehat{\beta}_2$	-0,293	-0,218	-0,124	-0,419	-0,185	-0,036	0,001	-0,079	-0,028
	$\widehat{\beta}_3$	0,270	-0,075	0,156	-0,298	0,023	0,106	0,087	0,049	0,098
	$\widehat{\beta}_4$	0,120	-0,053	0,043	-0,195	-0,001	-0,021	-0,018	0,004	-0,027
20	$\widehat{\beta}_0$	-0,195	-0,106	-0,188	-0,016	-0,140	-0,191	-0,191	-0,168	-0,191
	$\widehat{\beta}_1$	0,008	-0,253	-0,061	-0,412	-0,183	-0,029	-0,051	-0,129	-0,032
	$\widehat{\beta}_2$	-0,022	0,142	0,165	0,004	0,168	0,097	0,172	0,227	0,107
	$\widehat{\beta}_3$	-0,020	-0,436	-0,113	-0,629	-0,340	-0,044	-0,075	-0,245	-0,048
	$\widehat{\beta}_4$	0,015	-0,402	-0,141	-0,585	-0,313	-0,089	-0,152	-0,255	-0,097
30	$\widehat{\beta}_0$	-0,132	-0,072	-0,134	-0,004	-0,096	-0,132	-0,132	-0,122	-0,132
	$\widehat{\beta}_1$	-0,173	-0,469	-0,262	-0,646	-0,401	-0,207	-0,238	-0,332	-0,212
	$\widehat{\beta}_2$	0,087	0,067	0,181	-0,135	0,125	0,125	0,158	0,191	0,130
	$\widehat{\beta}_3$	0,071	0,086	0,093	-0,018	0,108	0,068	0,072	0,121	0,068
	$\widehat{\beta}_4$	-0,063	-0,536	-0,181	-0,728	-0,434	-0,085	-0,113	-0,325	-0,090
50	$\widehat{\beta}_0$	-0,083	-0,040	-0,087	-0,001	-0,059	-0,085	-0,088	-0,077	-0,085
	$\widehat{\beta}_1$	-0,023	-0,598	-0,148	-0,849	-0,474	-0,039	-0,069	-0,319	-0,042
	$\widehat{\beta}_2$	0,013	-0,053	0,060	-0,270	0,009	0,023	0,039	0,065	0,025
	$\widehat{\beta}_3$	-0,027	-0,255	-0,068	-0,425	-0,187	-0,033	-0,041	-0,119	-0,034
	$\widehat{\beta}_4$	-0,009	0,070	0,040	-0,073	0,090	-0,005	0,003	0,094	-0,004
80	$\widehat{\beta}_0$	-0,065	-0,032	-0,065	-0,001	-0,044	-0,065	-0,066	-0,058	-0,065
	$\widehat{\beta}_1$	-0,055	-0,421	-0,156	-0,635	-0,346	-0,069	-0,094	-0,256	-0,072
	$\widehat{\beta}_2$	0,006	0,119	0,124	-0,142	0,167	0,022	0,049	0,193	0,026
	$\widehat{\beta}_3$	0,001	-0,190	-0,025	-0,407	-0,127	-0,002	-0,007	-0,062	-0,003
	$\widehat{\beta}_4$	-0,001	-0,383	-0,083	-0,624	-0,285	-0,010	-0,023	-0,176	-0,012

Fonte: De autoria própria.

Tabela B.14: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,90$										
15	$\widehat{\beta}_0$	-0,268	-0,153	-0,263	-0,027	-0,199	-0,266	-0,265	-0,229	-0,266
	$\widehat{\beta}_1$	0,066	-0,097	0,088	-0,309	-0,034	0,108	0,104	0,034	0,113
	$\widehat{\beta}_2$	-0,077	-0,167	-0,055	-0,401	-0,104	-0,066	-0,045	-0,051	-0,060
	$\widehat{\beta}_3$	-0,031	-0,099	-0,024	-0,240	-0,058	-0,028	-0,013	-0,029	-0,033
	$\widehat{\beta}_4$	-0,031	-0,521	-0,169	-0,718	-0,415	-0,147	-0,232	-0,369	-0,165
20	$\widehat{\beta}_0$	-0,159	-0,039	-0,133	-0,001	-0,065	-0,155	-0,146	-0,087	-0,155
	$\widehat{\beta}_1$	0,031	0,135	0,103	0,013	0,150	0,058	0,099	0,174	0,066
	$\widehat{\beta}_2$	-0,052	-0,102	0,015	-0,288	-0,050	0,009	0,033	0,014	0,018
	$\widehat{\beta}_3$	0,059	-0,488	-0,082	-0,714	-0,369	-0,006	-0,071	-0,284	-0,017
	$\widehat{\beta}_4$	-0,052	-0,341	-0,127	-0,531	-0,263	-0,108	-0,136	-0,211	-0,119
30	$\widehat{\beta}_0$	-0,110	-0,025	-0,090	0,011	-0,041	-0,113	-0,110	-0,062	-0,113
	$\widehat{\beta}_1$	-0,053	0,086	0,061	-0,059	0,108	-0,018	0,022	0,133	-0,012
	$\widehat{\beta}_2$	0,072	-0,358	-0,048	-0,619	-0,256	0,027	-0,012	-0,159	0,019
	$\widehat{\beta}_3$	-0,113	-0,508	-0,186	-0,719	-0,416	-0,122	-0,147	-0,318	-0,123
	$\widehat{\beta}_4$	0,004	0,001	0,076	-0,155	0,043	0,031	0,058	0,089	0,036
50	$\widehat{\beta}_0$	-0,088	-0,035	-0,079	-0,002	-0,049	-0,086	-0,084	-0,064	-0,086
	$\widehat{\beta}_1$	0,004	-0,380	-0,076	-0,632	-0,283	-0,013	-0,035	-0,185	-0,017
	$\widehat{\beta}_2$	-0,021	-0,164	0,003	-0,452	-0,089	-0,006	0,003	-0,014	-0,004
	$\widehat{\beta}_3$	0,017	0,148	0,096	-0,005	0,163	0,027	0,047	0,164	0,017
	$\widehat{\beta}_4$	-0,018	-0,314	-0,070	-0,584	-0,219	-0,032	-0,045	-0,136	-0,018
80	$\widehat{\beta}_0$	-0,058	-0,034	-0,056	-0,001	-0,042	-0,057	-0,057	-0,051	-0,057
	$\widehat{\beta}_1$	-0,009	-0,138	-0,022	-0,359	-0,086	-0,012	-0,017	-0,041	-0,013
	$\widehat{\beta}_2$	-0,012	-0,086	0,041	-0,436	-0,008	-0,002	0,015	0,059	0,001
	$\widehat{\beta}_3$	0,011	-0,298	-0,040	-0,575	-0,214	0,003	-0,010	-0,121	0,002
	$\widehat{\beta}_4$	-0,005	-0,276	-0,048	-0,553	-0,192	-0,012	-0,021	-0,109	-0,014

Fonte: De autoria própria.

Tabela B.15: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,95$										
15	$\widehat{\beta}_0$	-0,126	-0,009	-0,117	0,082	-0,047	-0,183	-0,196	-0,082	-0,189
	$\widehat{\beta}_1$	0,017	-0,495	-0,111	-0,733	-0,386	-0,122	-0,219	-0,327	-0,136
	$\widehat{\beta}_2$	-0,053	-0,085	0,021	-0,312	-0,015	0,045	0,088	0,044	0,050
	$\widehat{\beta}_3$	0,037	0,062	0,118	-0,106	0,107	0,131	0,175	0,155	0,142
	$\widehat{\beta}_4$	-0,009	-0,261	-0,092	-0,410	-0,208	-0,116	-0,148	-0,174	-0,125
20	$\widehat{\beta}_0$	-0,127	-0,028	-0,106	0,030	-0,051	-0,136	-0,134	-0,075	-0,141
	$\widehat{\beta}_1$	0,062	-0,147	0,029	-0,329	-0,090	0,041	0,007	-0,044	0,039
	$\widehat{\beta}_2$	-0,073	-0,231	-0,078	-0,485	-0,161	-0,067	-0,058	-0,101	-0,067
	$\widehat{\beta}_3$	0,081	0,213	0,216	0,031	0,247	0,194	0,270	0,291	0,205
	$\widehat{\beta}_4$	-0,066	-0,465	-0,190	-0,686	-0,383	-0,182	-0,255	-0,332	-0,194
30	$\widehat{\beta}_0$	-0,136	-0,073	-0,121	-0,012	-0,092	-0,124	-0,115	-0,103	-0,123
	$\widehat{\beta}_1$	0,078	-0,252	-0,013	-0,475	-0,165	< 0,001	-0,038	-0,117	-0,012
	$\widehat{\beta}_2$	-0,128	-0,231	-0,085	-0,461	-0,176	-0,065	-0,068	-0,104	-0,054
	$\widehat{\beta}_3$	0,042	0,196	0,153	0,033	0,219	0,113	0,195	0,255	0,120
	$\widehat{\beta}_4$	0,003	-0,398	-0,097	-0,630	-0,300	-0,077	-0,134	-0,244	-0,086
50	$\widehat{\beta}_0$	-0,092	-0,060	-0,096	-0,007	-0,075	-0,094	-0,096	-0,090	-0,095
	$\widehat{\beta}_1$	-0,023	-0,289	-0,097	-0,488	-0,223	-0,061	-0,092	-0,165	-0,066
	$\widehat{\beta}_2$	-0,008	0,156	0,177	-0,071	0,198	0,093	0,173	0,254	0,106
	$\widehat{\beta}_3$	-0,011	-0,479	-0,178	-0,710	-0,392	-0,089	-0,161	-0,315	-0,100
	$\widehat{\beta}_4$	-0,016	-0,181	-0,042	-0,392	-0,120	-0,035	-0,043	-0,069	-0,037
80	$\widehat{\beta}_0$	-0,051	-0,017	-0,048	-0,001	-0,028	-0,051	-0,051	-0,038	-0,051
	$\widehat{\beta}_1$	0,041	-0,018	-0,066	-0,249	0,032	0,048	0,058	0,075	0,048
	$\widehat{\beta}_2$	-0,025	-0,300	-0,072	-0,630	-0,216	-0,037	-0,057	-0,135	-0,038
	$\widehat{\beta}_3$	-0,014	0,025	-0,053	-0,184	0,054	0,004	0,024	0,083	0,006
	$\widehat{\beta}_4$	-0,047	-0,338	-0,099	-0,631	-0,247	-0,057	-0,065	-0,166	-0,060

Fonte: De autoria própria.

Tabela B.16: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 0,89$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,99$										
15	$\widehat{\beta}_0$	-0,243	-0,169	-0,250	-0,094	-0,204	-0,244	-0,217	-0,215	-0,240
	$\widehat{\beta}_1$	0,046	0,047	0,129	-0,097	0,073	0,171	0,134	0,118	0,170
	$\widehat{\beta}_2$	0,131	-0,251	-0,016	-0,440	-0,171	-0,163	-0,192	-0,182	-0,172
	$\widehat{\beta}_3$	-0,246	0,050	-0,083	-0,062	0,048	0,117	0,156	0,108	0,126
	$\widehat{\beta}_4$	0,023	-0,268	-0,061	-0,446	-0,190	-0,167	-0,206	-0,199	-0,175
20	$\widehat{\beta}_0$	-0,141	-0,056	-0,136	-0,004	-0,086	-0,147	-0,136	-0,096	-0,146
	$\widehat{\beta}_1$	-0,060	-0,445	-0,198	-0,600	-0,392	-0,297	-0,346	-0,365	-0,302
	$\widehat{\beta}_2$	0,179	0,260	0,295	0,089	0,303	0,364	0,362	0,338	0,372
	$\widehat{\beta}_3$	0,030	-0,348	-0,091	-0,544	-0,261	-0,194	-0,239	-0,256	-0,218
	$\widehat{\beta}_4$	-0,213	0,032	-0,082	-0,088	0,038	0,046	0,107	0,080	0,062
30	$\widehat{\beta}_0$	-0,146	-0,099	-0,144	-0,041	-0,118	-0,143	-0,138	-0,128	-0,143
	$\widehat{\beta}_1$	0,022	-0,034	0,019	-0,205	0,009	0,024	0,025	0,022	0,032
	$\widehat{\beta}_2$	0,036	-0,217	-0,017	-0,393	-0,179	-0,050	-0,068	-0,117	-0,046
	$\widehat{\beta}_3$	0,043	0,204	0,160	0,051	0,229	0,234	0,281	0,263	0,232
	$\widehat{\beta}_4$	-0,168	-0,444	-0,236	-0,646	-0,365	-0,282	-0,334	-0,357	-0,296
50	$\widehat{\beta}_0$	-0,078	-0,046	-0,079	-0,011	-0,060	-0,081	-0,080	-0,069	-0,081
	$\widehat{\beta}_1$	-0,028	-0,030	0,035	-0,156	-0,015	0,075	0,074	0,038	0,077
	$\widehat{\beta}_2$	0,027	0,173	0,149	-0,004	0,211	0,194	0,247	0,244	0,213
	$\widehat{\beta}_3$	-0,021	0,015	0,034	-0,136	0,046	0,018	0,070	0,075	0,024
	$\widehat{\beta}_4$	-0,001	-0,668	-0,270	-0,886	-0,558	-0,337	-0,461	-0,538	-0,369
80	$\widehat{\beta}_0$	-0,062	-0,034	-0,061	-0,007	-0,044	-0,062	-0,062	-0,053	-0,062
	$\widehat{\beta}_1$	-0,014	0,008	0,061	-0,215	0,048	0,051	0,078	0,090	0,058
	$\widehat{\beta}_2$	0,043	-0,315	-0,102	-0,584	-0,245	-0,087	-0,133	-0,191	-0,091
	$\widehat{\beta}_3$	0,030	0,143	0,147	-0,063	0,173	0,126	0,183	0,210	0,133
	$\widehat{\beta}_4$	-0,124	-0,368	-0,162	-0,614	-0,297	-0,141	-0,178	-0,249	-0,150

Fonte: De autoria própria.

Tabela B.17: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,85$										
15	$\widehat{\beta}_0$	-0,164	-0,042	-0,137	< 0,001	-0,070	-0,158	-0,149	-0,096	-0,157
	$\widehat{\beta}_1$	-0,007	0,030	0,068	-0,121	0,066	0,018	0,046	0,105	0,022
	$\widehat{\beta}_2$	0,018	-0,272	-0,008	-0,539	-0,171	0,015	0,012	-0,078	0,015
	$\widehat{\beta}_3$	-0,018	-0,325	-0,046	-0,580	-0,225	-0,023	-0,033	-0,133	-0,024
	$\widehat{\beta}_4$	-0,010	-0,388	-0,123	-0,540	-0,312	-0,049	-0,093	-0,244	-0,056
20	$\widehat{\beta}_0$	-0,143	-0,059	-0,130	-0,004	-0,084	-0,140	-0,136	-0,108	-0,139
	$\widehat{\beta}_1$	0,009	-0,273	-0,070	-0,446	-0,203	-0,018	-0,042	-0,142	-0,022
	$\widehat{\beta}_2$	-0,028	0,113	0,136	-0,083	0,148	0,037	0,092	0,197	0,047
	$\widehat{\beta}_3$	-0,050	-0,290	-0,096	-0,476	-0,219	-0,062	-0,073	-0,154	-0,064
	$\widehat{\beta}_4$	0,019	-0,487	-0,163	-0,701	-0,387	-0,049	-0,109	-0,305	-0,060
30	$\widehat{\beta}_0$	-0,085	-0,034	-0,088	-0,001	-0,053	-0,087	-0,090	-0,074	-0,087
	$\widehat{\beta}_1$	0,014	-0,017	-0,010	-0,070	0,021	-0,001	-0,020	0,012	-0,003
	$\widehat{\beta}_2$	-0,008	0,102	-0,107	-0,093	0,143	0,024	0,070	0,177	0,028
	$\widehat{\beta}_3$	-0,004	-0,570	-0,121	-0,809	-0,447	-0,024	-0,056	-0,305	-0,026
	$\widehat{\beta}_4$	-0,004	-0,359	-0,079	-0,549	-0,277	-0,023	-0,050	-0,182	-0,026
50	$\widehat{\beta}_0$	-0,048	-0,035	-0,055	0,006	-0,046	-0,051	-0,053	-0,057	-0,051
	$\widehat{\beta}_1$	-0,054	-0,055	-0,018	-0,236	-0,013	-0,050	-0,041	0,015	-0,050
	$\widehat{\beta}_2$	0,033	-0,269	0,009	-0,569	-0,172	0,032	0,025	-0,060	0,031
	$\widehat{\beta}_3$	-0,034	-0,267	-0,046	-0,508	-0,186	-0,035	-0,036	-0,094	-0,035
	$\widehat{\beta}_4$	-0,024	-0,329	-0,060	-0,574	-0,234	-0,028	-0,032	-0,129	-0,029
80	$\widehat{\beta}_0$	-0,030	0,010	-0,018	0,003	< 0,001	-0,028	-0,027	-0,010	-0,028
	$\widehat{\beta}_1$	0,013	-0,252	-0,051	-0,582	-0,173	0,005	-0,003	-0,100	0,004
	$\widehat{\beta}_2$	-0,039	0,229	0,089	-0,025	0,199	-0,022	-0,005	0,151	-0,019
	$\widehat{\beta}_3$	-0,007	-0,278	-0,088	-0,607	-0,200	-0,018	-0,027	-0,132	-0,019
	$\widehat{\beta}_4$	0,005	-0,193	-0,050	-0,468	-0,135	-0,002	-0,009	-0,084	-0,003

Fonte: De autoria própria.

Tabela B.18: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,90$	$\widehat{\beta}_0$	-0,168	-0,077	-0,165	-0,007	-0,109	-0,170	-0,170	-0,138	-0,170
	$\widehat{\beta}_1$	-0,016	-0,585	-0,159	-0,778	-0,471	-0,075	-0,153	-0,373	-0,080
	$\widehat{\beta}_2$	-0,010	-0,393	-0,085	-0,575	-0,297	-0,037	-0,063	-0,214	-0,044
	$\widehat{\beta}_3$	-0,002	0,074	0,097	-0,023	0,106	0,058	0,117	0,144	0,064
	$\widehat{\beta}_4$	0,014	0,001	0,061	-0,108	0,040	0,031	0,054	0,075	0,034
15	$\widehat{\beta}_0$	-0,091	0,028	-0,046	0,019	0,014	-0,087	-0,075	0,002	-0,086
	$\widehat{\beta}_1$	-0,044	-0,141	-0,059	-0,265	-0,096	-0,083	-0,079	-0,063	-0,085
	$\widehat{\beta}_2$	0,038	0,123	0,202	0,011	0,148	0,182	0,232	0,218	0,192
	$\widehat{\beta}_3$	-0,041	-0,518	-0,187	-0,737	-0,410	-0,115	-0,160	-0,326	-0,122
	$\widehat{\beta}_4$	0,001	-0,362	-0,137	-0,503	-0,289	-0,078	-0,122	-0,243	-0,085
20	$\widehat{\beta}_0$	-0,067	-0,038	-0,074	0,003	-0,052	-0,071	-0,073	0,067	-0,073
	$\widehat{\beta}_1$	0,044	-0,127	0,031	-0,332	-0,065	0,042	0,038	-0,005	0,042
	$\widehat{\beta}_2$	-0,009	-0,160	0,015	-0,468	-0,078	-0,002	0,005	< 0,001	-0,001
	$\widehat{\beta}_3$	-0,063	-0,204	-0,056	-0,433	-0,142	-0,061	-0,060	-0,081	-0,061
	$\widehat{\beta}_4$	0,015	-0,353	-0,054	-0,647	-0,244	0,001	-0,012	-0,142	-0,002
30	$\widehat{\beta}_0$	-0,051	-0,026	-0,047	-0,002	-0,036	-0,050	-0,050	-0,043	-0,050
	$\widehat{\beta}_1$	0,008	-0,108	-0,016	-0,330	-0,056	-0,004	-0,008	-0,018	-0,004
	$\widehat{\beta}_2$	-0,012	0,101	0,138	-0,126	0,125	0,036	0,068	0,162	0,040
	$\widehat{\beta}_3$	-0,005	-0,443	-0,130	-0,744	-0,338	-0,033	-0,058	-0,233	-0,036
	$\widehat{\beta}_4$	-0,007	-0,210	-0,088	-0,428	-0,152	-0,035	-0,053	-0,233	-0,038
50	$\widehat{\beta}_0$	-0,030	-0,010	-0,027	< 0,001	-0,016	-0,030	-0,031	-0,022	-0,031
	$\widehat{\beta}_1$	-0,008	-0,143	-0,028	-0,361	-0,092	-0,011	-0,015	-0,046	-0,012
	$\widehat{\beta}_2$	-0,043	0,140	0,150	-0,159	0,183	0,058	0,082	0,208	0,061
	$\widehat{\beta}_3$	-0,018	-0,484	-0,123	-0,779	-0,379	-0,030	-0,050	-0,246	-0,032
	$\widehat{\beta}_4$	-0,054	-0,229	-0,097	-0,457	-0,176	-0,061	-0,072	-0,125	-0,062

Fonte: De autoria própria.

Tabela B.19: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,95$	$\widehat{\beta}_0$	-0,160	-0,148	-0,177	-0,027	-0,172	-0,166	-0,174	-0,193	-0,166
	$\widehat{\beta}_1$	0,007	-0,216	-0,017	-0,415	-0,138	-0,008	-0,012	-0,080	-0,011
	$\widehat{\beta}_2$	0,023	-0,231	0,021	-0,487	-0,142	0,038	0,023	-0,057	0,042
	$\widehat{\beta}_3$	-0,024	-0,188	-0,027	-0,373	-0,123	-0,037	-0,019	-0,069	-0,024
	$\widehat{\beta}_4$	-0,010	-0,304	-0,058	-0,543	-0,210	-0,023	-0,049	-0,144	-0,041
15	$\widehat{\beta}_0$	-0,126	-0,057	-0,122	-0,008	-0,080	-0,125	-0,125	-0,099	-0,125
	$\widehat{\beta}_1$	0,022	-0,122	0,012	-0,326	-0,063	0,003	0,006	-0,013	< 0,001
	$\widehat{\beta}_2$	-0,082	0,075	0,080	-0,127	0,105	0,042	0,111	0,162	0,059
	$\widehat{\beta}_3$	0,016	-0,513	-0,155	-0,743	-0,408	-0,075	-0,163	-0,339	-0,086
	$\widehat{\beta}_4$	0,046	-0,230	-0,038	-0,426	-0,157	-0,026	-0,054	-0,111	-0,038
20	$\widehat{\beta}_0$	-0,074	-0,039	-0,078	-0,003	-0,055	-0,076	-0,078	-0,069	-0,076
	$\widehat{\beta}_1$	0,119	-0,462	-0,034	-0,710	-0,350	0,073	-0,010	-0,240	0,068
	$\widehat{\beta}_2$	-0,088	-0,332	-0,110	-0,595	-0,240	-0,099	-0,085	-0,156	-0,105
	$\widehat{\beta}_3$	-0,052	0,109	0,072	-0,031	0,128	-0,006	0,041	0,154	0,005
	$\widehat{\beta}_4$	-0,020	-0,073	0,011	-0,232	-0,030	-0,008	0,002	0,015	-0,007
30	$\widehat{\beta}_0$	-0,053	-0,028	-0,050	-0,005	-0,037	-0,052	-0,052	-0,045	-0,052
	$\widehat{\beta}_1$	0,011	-0,091	-0,007	-0,312	-0,038	-0,011	-0,010	< 0,001	-0,013
	$\widehat{\beta}_2$	-0,014	0,098	0,159	-0,114	0,123	0,092	0,127	0,179	0,101
	$\widehat{\beta}_3$	-0,004	-0,449	-0,148	-0,726	-0,350	-0,062	-0,099	-0,264	-0,068
	$\widehat{\beta}_4$	-0,011	-0,189	-0,087	-0,410	-0,129	-0,068	-0,080	-0,095	-0,072
50	$\widehat{\beta}_0$	-0,036	-0,012	-0,031	< 0,001	-0,019	-0,036	-0,036	-0,027	-0,036
	$\widehat{\beta}_1$	< 0,001	-0,124	-0,017	-0,356	-0,074	-0,007	-0,014	-0,031	-0,008
	$\widehat{\beta}_2$	0,048	0,148	-0,177	-0,153	0,196	0,081	0,127	0,237	0,086
	$\widehat{\beta}_3$	-0,019	-0,481	-0,146	-0,774	-0,385	-0,043	-0,079	-0,274	-0,046
	$\widehat{\beta}_4$	-0,059	-0,215	-0,097	-0,452	-0,165	-0,073	-0,091	-0,119	-0,075

Fonte: De autoria própria.

Tabela B.20: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 1,61$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,99$										
15	$\widehat{\beta}_0$	-0,150	-0,118	-0,152	-0,059	-0,132	-0,153	-0,151	-0,142	-0,155
	$\widehat{\beta}_1$	-0,091	0,082	0,005	-0,074	0,093	0,096	0,141	0,127	0,106
	$\widehat{\beta}_2$	0,454	-0,206	0,181	-0,425	-0,107	-0,071	-0,128	-0,122	-0,086
	$\widehat{\beta}_3$	-0,358	-0,384	-0,305	-0,549	-0,358	-0,268	-0,296	-0,318	-0,273
	$\widehat{\beta}_4$	-0,105	0,157	0,056	0,012	0,165	0,207	0,233	0,206	0,216
20	$\widehat{\beta}_0$	-0,118	-0,081	-0,117	-0,031	-0,097	-0,116	-0,113	-0,105	-0,116
	$\widehat{\beta}_1$	0,019	0,428	-0,143	-0,619	-0,351	-0,267	-0,336	-0,341	-0,267
	$\widehat{\beta}_2$	-0,112	0,077	0,036	-0,092	0,105	0,149	0,185	0,149	0,149
	$\widehat{\beta}_3$	0,019	0,178	0,127	0,033	0,191	0,226	0,276	0,241	0,231
	$\widehat{\beta}_4$	0,071	-0,335	-0,054	-0,512	-0,257	-0,163	-0,217	-0,242	-0,175
30	$\widehat{\beta}_0$	-0,091	-0,065	-0,086	-0,025	-0,075	-0,085	-0,082	-0,080	-0,084
	$\widehat{\beta}_1$	0,005	0,025	0,053	-0,155	0,048	0,067	0,094	0,087	0,075
	$\widehat{\beta}_2$	-0,104	-0,084	-0,053	-0,318	-0,044	-0,005	0,015	-0,005	-0,008
	$\widehat{\beta}_3$	0,073	-0,385	-0,079	-0,631	-0,292	-0,160	-0,242	-0,274	-0,173
	$\widehat{\beta}_4$	0,008	-0,015	0,049	-0,215	0,021	0,063	0,084	0,062	0,069
50	$\widehat{\beta}_0$	-0,053	-0,020	-0,046	-0,004	-0,028	-0,048	-0,043	-0,033	-0,047
	$\widehat{\beta}_1$	0,019	-0,118	-0,016	-0,295	-0,086	-0,037	-0,047	-0,051	-0,045
	$\widehat{\beta}_2$	0,001	0,154	0,174	-0,048	0,180	0,237	0,259	0,236	0,260
	$\widehat{\beta}_3$	-0,052	0,023	-0,006	-0,157	0,054	-0,043	0,028	0,075	-0,036
	$\widehat{\beta}_4$	0,005	-0,565	-0,213	-0,813	-0,460	-0,219	-0,313	-0,424	-0,246
80	$\widehat{\beta}_0$	-0,031	-0,012	-0,029	-0,001	-0,018	-0,031	-0,032	-0,024	-0,032
	$\widehat{\beta}_1$	0,031	-0,054	0,022	-0,301	-0,011	0,007	0,009	0,020	0,005
	$\widehat{\beta}_2$	0,054	0,171	0,193	-0,107	0,211	0,200	0,263	0,262	0,212
	$\widehat{\beta}_3$	-0,016	-0,435	-0,170	-0,715	0,358	-0,130	-0,204	-0,305	-0,144
	$\widehat{\beta}_4$	-0,091	-0,162	-0,091	-0,398	-0,129	-0,120	-0,123	-0,097	-0,118

Fonte: De autoria própria.

Tabela B.21: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,85$										
15	$\widehat{\beta}_0$	-0,051	-0,074	-0,066	-0,004	-0,080	-0,052	-0,056	-0,083	-0,053
	$\widehat{\beta}_1$	-0,017	-0,389	-0,086	-0,678	-0,293	-0,027	-0,045	-0,182	-0,029
	$\widehat{\beta}_2$	0,004	-0,210	0,014	-0,528	-0,118	0,008	0,014	-0,019	0,009
	$\widehat{\beta}_3$	-0,032	0,053	0,026	-0,140	0,074	-0,024	-0,011	0,084	-0,023
	$\widehat{\beta}_4$	0,001	-0,268	-0,048	-0,451	-0,199	0,006	-0,017	-0,120	-0,007
20	$\widehat{\beta}_0$	-0,034	-0,005	-0,025	< 0,001	-0,007	-0,033	-0,031	-0,013	-0,033
	$\widehat{\beta}_1$	0,005	-0,211	-0,025	-0,463	-0,140	0,002	-0,002	-0,067	0,001
	$\widehat{\beta}_2$	-0,008	0,182	0,093	-0,071	0,191	0,005	0,024	0,181	0,006
	$\widehat{\beta}_3$	-0,010	-0,078	-0,013	-0,255	-0,041	-0,011	-0,010	-0,014	-0,011
	$\widehat{\beta}_4$	0,011	-0,513	-0,116	-0,815	-0,403	-0,005	-0,031	-0,268	-0,007
30	$\widehat{\beta}_0$	-0,016	-0,002	-0,011	< 0,001	-0,003	-0,016	-0,016	-0,002	-0,017
	$\widehat{\beta}_1$	-0,009	-0,239	-0,042	-0,516	-0,173	-0,003	-0,008	-0,101	-0,005
	$\widehat{\beta}_2$	-0,005	-0,142	0,058	-0,166	0,146	0,001	-0,009	0,125	-0,003
	$\widehat{\beta}_3$	-0,005	-0,210	-0,027	-0,534	-0,139	-0,007	-0,010	-0,069	-0,008
	$\widehat{\beta}_4$	-0,003	-0,248	-0,035	-0,590	-0,165	-0,001	-0,005	-0,086	-0,002
50	$\widehat{\beta}_0$	< 0,001	-0,040	-0,022	-0,001	-0,042	-0,013	-0,014	-0,038	-0,013
	$\widehat{\beta}_1$	< 0,001	-0,150	-0,002	-0,426	-0,086	-0,001	-0,001	-0,021	-0,001
	$\widehat{\beta}_2$	< 0,001	0,190	0,049	-0,068	0,184	0,003	0,008	0,142	0,005
	$\widehat{\beta}_3$	< 0,001	-0,424	-0,055	-0,740	-0,309	-0,007	-0,012	-0,165	-0,009
	$\widehat{\beta}_4$	< 0,001	-0,239	-0,026	-0,497	-0,170	-0,003	-0,005	-0,083	-0,004
80	$\widehat{\beta}_0$	< 0,001	0,004	-0,007	< 0,001	0,001	-0,009	-0,008	-0,003	-0,009
	$\widehat{\beta}_1$	< 0,001	-0,272	-0,030	< 0,001	-0,170	-0,002	0,004	-0,003	-0,003
	$\widehat{\beta}_2$	< 0,001	0,112	0,028	< 0,001	0,096	0,001	0,003	-0,082	0,002
	$\widehat{\beta}_3$	< 0,001	-0,333	-0,033	< 0,001	-0,206	-0,002	-0,005	-0,096	-0,003
	$\widehat{\beta}_4$	< 0,001	0,078	0,009	< 0,001	0,046	0,002	0,002	0,024	0,002

Fonte: De autoria própria.

Tabela B.22: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$.

		MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
$\rho = 0,90$	$\widehat{\beta}_0$	-0,037	-0,018	-0,019	-0,004	-0,013	-0,033	-0,029	-0,010	-0,032
	$\widehat{\beta}_1$	-0,006	-0,357	-0,090	-0,657	-0,271	-0,020	-0,035	-0,184	-0,025
	$\widehat{\beta}_2$	0,002	0,150	0,090	-0,155	0,161	0,014	0,028	0,153	0,018
	$\widehat{\beta}_3$	-0,001	-0,354	-0,079	-0,686	-0,260	-0,014	-0,027	-0,169	-0,018
	$\widehat{\beta}_4$	0,006	0,127	0,074	-0,033	0,132	0,019	0,034	0,128	0,024
15	$\widehat{\beta}_0$	-0,026	-0,021	-0,014	0,001	-0,020	-0,026	-0,024	-0,016	-0,025
	$\widehat{\beta}_1$	-0,004	-0,156	-0,043	-0,401	-0,106	-0,011	-0,019	-0,065	-0,013
	$\widehat{\beta}_2$	0,030	0,193	0,138	-0,037	0,196	0,048	0,068	0,193	0,052
	$\widehat{\beta}_3$	-0,020	-0,350	-0,080	-0,704	-0,254	-0,028	-0,037	-0,153	-0,029
	$\widehat{\beta}_4$	-0,030	-0,217	-0,077	-0,471	-0,161	-0,039	-0,048	-0,108	-0,041
20	$\widehat{\beta}_0$	-0,029	-0,058	-0,039	-0,007	-0,054	-0,030	-0,032	-0,048	-0,031
	$\widehat{\beta}_1$	-0,030	-0,139	-0,063	-0,378	-0,111	-0,035	-0,041	-0,085	-0,037
	$\widehat{\beta}_2$	0,004	0,189	0,102	-0,145	0,194	0,017	0,032	0,174	0,021
	$\widehat{\beta}_3$	0,018	-0,239	-0,034	-0,588	-0,159	0,012	0,005	-0,151	0,011
	$\widehat{\beta}_4$	-0,009	-0,319	-0,077	-0,644	-0,237	-0,017	-0,027	-0,087	-0,019
30	$\widehat{\beta}_0$	-0,012	-0,035	-0,023	-0,001	-0,038	-0,013	-0,014	-0,036	-0,013
	$\widehat{\beta}_1$	0,001	-0,148	-0,001	-0,423	-0,084	0,001	0,001	-0,022	0,001
	$\widehat{\beta}_2$	-0,002	0,212	0,068	-0,064	0,210	0,004	0,012	0,173	0,007
	$\widehat{\beta}_3$	-0,002	-0,422	-0,069	-0,736	-0,316	-0,007	-0,015	-0,184	-0,010
	$\widehat{\beta}_4$	-0,002	-0,234	-0,034	-0,494	-0,170	-0,005	-0,009	-0,092	-0,006
50	$\widehat{\beta}_0$	-0,008	-0,017	-0,010	-0,003	-0,015	-0,008	-0,008	-0,013	-0,008
	$\widehat{\beta}_1$	0,003	-0,118	-0,007	-0,467	-0,070	0,002	0,001	-0,028	0,002
	$\widehat{\beta}_2$	-0,011	-0,065	-0,008	-0,513	-0,024	-0,011	-0,010	0,002	-0,011
	$\widehat{\beta}_3$	0,011	0,085	0,026	-0,048	0,064	0,012	0,013	0,044	0,012
	$\widehat{\beta}_4$	-0,002	-0,187	-0,017	-0,566	-0,115	-0,003	-0,004	-0,053	-0,003
80	$\widehat{\beta}_0$	-0,008	-0,017	-0,010	-0,003	-0,015	-0,008	-0,008	-0,013	-0,008
	$\widehat{\beta}_1$	0,003	-0,118	-0,007	-0,467	-0,070	0,002	0,001	-0,028	0,002
	$\widehat{\beta}_2$	-0,011	-0,065	-0,008	-0,513	-0,024	-0,011	-0,010	0,002	-0,011
	$\widehat{\beta}_3$	0,011	0,085	0,026	-0,048	0,064	0,012	0,013	0,044	0,012
	$\widehat{\beta}_4$	-0,002	-0,187	-0,017	-0,566	-0,115	-0,003	-0,004	-0,053	-0,003

Fonte: De autoria própria.

Tabela B.23: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,95$										
15	$\widehat{\beta}_0$	-0,027	-0,047	-0,049	-0,011	-0,051	-0,046	-0,055	-0,060	-0,049
	$\widehat{\beta}_1$	-0,047	0,126	0,074	-0,007	0,120	0,065	0,099	0,146	0,082
	$\widehat{\beta}_2$	0,001	0,152	0,071	0,005	0,150	0,019	0,060	0,156	0,024
	$\widehat{\beta}_3$	-0,065	-0,556	-0,184	-0,843	-0,434	-0,117	-0,164	-0,354	-0,128
	$\widehat{\beta}_4$	0,102	-0,171	-0,027	-0,332	-0,124	-0,009	-0,061	-0,111	-0,026
20	$\widehat{\beta}_0$	-0,025	-0,038	-0,033	-0,005	-0,040	-0,027	-0,030	-0,040	-0,027
	$\widehat{\beta}_1$	0,012	-0,031	0,049	-0,321	0,024	0,023	0,043	0,065	0,023
	$\widehat{\beta}_2$	0,184	-0,104	0,106	-0,436	-0,040	0,161	0,117	0,023	0,165
	$\widehat{\beta}_3$	-0,141	0,095	-0,067	0,002	0,060	-0,125	-0,098	0,020	-0,125
	$\widehat{\beta}_4$	-0,167	-0,430	-0,188	-0,783	-0,325	-0,166	-0,156	-0,242	-0,169
30	$\widehat{\beta}_0$	-0,012	< 0,001	-0,013	0,005	-0,004	-0,014	-0,015	-0,011	-0,015
	$\widehat{\beta}_1$	-0,025	< 0,001	0,004	-0,351	-0,024	-0,020	-0,011	0,020	-0,019
	$\widehat{\beta}_2$	0,026	< 0,001	-0,070	-0,736	-0,285	0,008	-0,019	-0,178	0,005
	$\widehat{\beta}_3$	-0,012	-0,038	-0,031	-0,505	-0,160	-0,013	-0,014	-0,082	-0,013
	$\widehat{\beta}_4$	0,018	0,013	0,078	-0,089	0,151	-0,001	0,023	0,159	0,002
50	$\widehat{\beta}_0$	-0,014	-0,012	-0,005	-0,007	-0,008	-0,013	-0,011	-0,003	-0,012
	$\widehat{\beta}_1$	0,003	-0,312	-0,078	-0,642	-0,230	-0,008	-0,022	-0,147	-0,010
	$\widehat{\beta}_2$	-0,002	0,232	0,127	-0,006	0,216	0,019	0,040	0,191	0,022
	$\widehat{\beta}_3$	0,002	-0,171	-0,055	-0,433	-0,126	-0,008	-0,018	-0,086	-0,009
	$\widehat{\beta}_4$	-0,004	-0,153	-0,037	-0,477	-0,097	-0,009	-0,013	-0,053	-0,009
80	$\widehat{\beta}_0$	-0,011	-0,008	-0,007	< 0,001	-0,005	-0,010	-0,010	-0,003	-0,010
	$\widehat{\beta}_1$	0,006	0,088	0,041	-0,134	-0,114	0,008	0,016	0,107	0,008
	$\widehat{\beta}_2$	0,004	-0,394	-0,063	-0,732	-0,306	-0,001	-0,017	-0,180	-0,002
	$\widehat{\beta}_3$	-0,006	0,193	0,051	0,001	0,198	-0,002	0,009	0,156	-0,002
	$\widehat{\beta}_4$	-0,012	-0,355	-0,021	-0,658	-0,247	-0,012	-0,011	-0,098	-0,012

Fonte: De autoria própria.

Tabela B.24: Vieses médios quando $\beta_0 = 0$, $\beta_1, \beta_2, \beta_3$ e β_4 assumem valores no intervalo $(0, 1)$, $\delta = 8$ e $p = 4$.

	MV	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	
$\rho = 0,99$										
15	$\widehat{\beta}_0$	-0,036	-0,033	-0,041	-0,009	-0,038	-0,044	-0,045	-0,043	-0,044
	$\widehat{\beta}_1$	-0,016	-0,178	-0,045	-0,378	-0,124	-0,057	-0,069	-0,088	-0,059
	$\widehat{\beta}_2$	0,024	0,251	0,221	0,066	0,269	0,296	0,348	0,331	0,314
	$\widehat{\beta}_3$	-0,019	-0,347	-0,123	-0,592	-0,256	-0,143	-0,175	-0,226	-0,152
	$\widehat{\beta}_4$	0,011	-0,300	-0,094	-0,498	-0,238	-0,134	-0,159	-0,200	-0,144
20	$\widehat{\beta}_0$	-0,029	-0,058	-0,043	-0,020	-0,057	-0,044	-0,051	-0,058	-0,045
	$\widehat{\beta}_1$	-0,024	0,199	0,090	0,040	0,176	0,089	0,140	0,199	0,105
	$\widehat{\beta}_2$	0,038	-0,163	-0,049	-0,428	-0,116	-0,077	-0,096	-0,095	-0,086
	$\widehat{\beta}_3$	-0,011	0,112	-0,095	-0,121	0,141	0,108	0,161	0,165	0,116
	$\widehat{\beta}_4$	-0,007	-0,434	-0,134	-0,710	-0,350	-0,114	-0,203	-0,311	-0,128
30	$\widehat{\beta}_0$	-0,022	-0,022	-0,020	-0,008	-0,021	-0,021	-0,020	-0,019	-0,021
	$\widehat{\beta}_1$	-0,027	0,225	0,089	0,035	0,208	0,031	0,087	0,209	0,040
	$\widehat{\beta}_2$	-0,050	-0,194	-0,087	-0,488	-0,156	-0,074	-0,101	-0,122	-0,077
	$\widehat{\beta}_3$	0,006	-0,246	-0,048	-0,545	-0,178	-0,012	-0,037	-0,130	-0,016
	$\widehat{\beta}_4$	0,067	-0,077	0,046	-0,375	-0,017	0,056	0,056	0,019	0,055
50	$\widehat{\beta}_0$	-0,012	-0,018	-0,008	-0,005	-0,014	-0,010	-0,009	-0,008	-0,010
	$\widehat{\beta}_1$	-0,007	0,222	0,119	-0,027	0,225	0,038	0,093	0,226	0,045
	$\widehat{\beta}_2$	0,011	-0,230	-0,063	-0,560	-0,175	-0,027	-0,067	-0,128	-0,031
	$\widehat{\beta}_3$	0,014	0,081	0,062	-0,189	0,102	0,030	0,052	0,115	0,032
	$\widehat{\beta}_4$	-0,022	-0,386	-0,114	-0,710	-0,307	-0,041	-0,071	-0,233	-0,046
80	$\widehat{\beta}_0$	-0,009	-0,019	-0,012	-0,004	-0,017	-0,010	-0,011	-0,016	-0,010
	$\widehat{\beta}_1$	0,015	0,215	0,126	-0,054	0,221	0,043	0,086	0,218	0,046
	$\widehat{\beta}_2$	-0,007	-0,287	-0,111	-0,630	-0,235	-0,041	-0,093	-0,184	-0,043
	$\widehat{\beta}_3$	-0,001	-0,252	-0,046	0,584	-0,182	-0,007	-0,016	-0,112	-0,008
	$\widehat{\beta}_4$	-0,013	-0,018	0,033	-0,323	0,021	0,002	0,027	0,054	0,003

Fonte: De autoria própria.

Referências Bibliográficas

- Ahmed, S. E., Budsaba, K., Lisawadi, S., and Volodin, A. (2008). Parametric estimation for the birnbaum-saunders lifetime distribution based on a new parametrization. *Thailand Statistician*, 6(2):213–240.
- Algamal, Z. Y. (2019). Performance of ridge estimator in inverse gaussian regression model. *Communications in Statistics - Theory and Methods*, 48(15):3836–3849.
- Alkhamisi, M., Khalaf, G., and Shukur, G. (2006). Some modifications for choosing ridge parameters. *Communications in Statistics-Theory and Methods*, 35(11):2005–2020.
- Alkhamisi, M. and Shukur, G. (2008). Developing ridge parameters for sur model. *Communications in Statistics—Theory and Methods*, 37(4):544–564.
- Amin, M., Qasim, M., Amanullah, M., and Afzal, S. (2020). Performance of some ridge estimators for the gamma regression model. *Statistical papers*, 61(3):997–1026.
- Banerjee, S. and Roy, A. (2014). *Linear algebra and matrix analysis for statistics*. Number 1313. Crc Press.
- Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of applied probability*, page 319–327.
- Casagrande, M. H. (2016). *Comparação de métodos de estimação para problemas com colinearidade e/ou alta dimensionalidade ($p > n$)*. PhD thesis, Universidade Federal de São Carlos.
- Cessie, S. I. and van Houwelingen, J. (1992). Ridge regression: Biased estimation for nonorthogonal problems. *Applied Statistics*, 41:191–201.

- Chatterjee, S. and Hadi, A. S. (2009). *Sensitivity analysis in linear regression*, volume 327. John Wiley & Sons.
- Dorugade, A. V. (2014). New ridge parameters for ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15:94–99.
- Ferrari, S. L. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.
- Firinguetti, L. (1999). A generalized ridge regression estimator and its finite sample properties: A generalized ridge regression estimator. *Communications in Statistics-Theory and Methods*, 28(5):1217–1229.
- Galea, M., Leiva-Sánchez, V., and Paula, G. (2004). Influence diagnostics in log-birnbaum-saunders regression models. *Journal of Applied Statistics*, 31(9):1049–1064.
- Hoerl, A. E. (1962). Applications of ridge analysis to regression problems. *Chem. Eng. Progress*, 58(3):54–59.
- Hoerl, A. E., Kennard, R. W., and Baldwin, K. F. (1975). Ridge regression: some simulations. *Communications in Statistics-Theory and Methods*, 4(2):105–123.
- Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.
- Khalaf, G. and Shukur, G. (2005). Choosing ridge parameter for regression problems.
- Kibria, B., Mansson, K., and Shukur, P. (2012). Performance of some logistic ridge regression estimators. *Comput Econ*, 40:401–414.
- Kibria, B. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation*, 32(2):419–435.
- Kibria, B. G., Månsson, K., and Shukur, G. (2013a). Some ridge regression estimators for the zero-inflated poisson model. *Journal of Applied Statistics*, 40(4):721–735.
- Kibria, B. M. G., Månsson, K., and Shukur, G. (2013b). Some ridge regression estimators for the zero-inflated poisson model. *Journal of Applied Statistics*, 40(4):721–735.

- Kurtoğlu, F. and Özkale, M. R. (2016). Liu estimation in generalized linear models: application on gamma distributed response variable. *Statistical Papers*, 57(4):911–928.
- Lee, A. H. and Silvapulle, M. J. (1988). Ridge estimation in logistic regression. *Communications in Statistics Simulation and Computation*, 17(4):1231–1257.
- Leiva, V. (2015). *The Birnbaum-Saunders Distribution*. Academic Press.
- Leiva, V., Sanhueza, A., and Angulo, J. M. (2009). A length-biased version of the birnbaum-saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment*, 23(3):299–307.
- Leiva, V., Saulo, H., Leão, J., and Marchant, C. (2014). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics & Data Analysis*, 79:175–191.
- Lesaffre, E. and Marx, B. D. (1993). Collinearity in generalized linear regression. *Communications in Statistics - Theory and Methods*, 22(7):1933–1952.
- Lukman, A. F., Ayinde, K., Kibria, B. G., and Adewuyi, E. T. (2020). Modified ridge-type estimator for the gamma regression model. *Communications in Statistics-Simulation and Computation*, pages 1–15.
- Mansson, K. and Shukur, G. (2011). A poisson ridge regression estimator. *Economic Modelling*, 28:1475–81.
- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to linear regression analysis*, volume 821. John Wiley & Sons.
- Muniz, G. and Kibria, B. G. (2009). On some ridge regression estimators: An empirical comparisons. *Communications in Statistics—Simulation and Computation*®^(R), 38(3):621–630.
- Nomura, M. (1988). On the almost unbiased ridge regression estimator. *Communications in Statistics-Simulation and Computation*, 17(3):729–743.

- Paula, G. A., Leiva, V., Barros, M., and Liu, S. (2012). Robust statistical modeling using the birnbaum-saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, 28(1):16–34.
- Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the birnbaum-saunders distribution. *Technometrics*, 33(1):51–60.
- Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., and Ahmed, S. E. (2012). On new parameterizations of the birnbaum-saunders distribution. *Pakistan Journal of Statistics*, 28(1).
- Santos-Neto, M. F. d. (2010). Estimação e modelagem com a distribuição birnbaum-saunders: uma nova reparametrização. *Dissertação (Mestrado)-Universidade Federal de Pernambuco*, pages 24–45.
- Saulo, H., Leão, J., and Santos-Neto, M. (2019). Discussion of “birnbaum-saunders distribution: A review of models, analysis, and applications” by n. balakrishnan and debasis kundu. *Applied Stochastic Models in Business and Industry*, 35(1):118–121.
- Schaefer, R., Roi, L., and Wolfe, R. (1984). A ridge logistic estimator. *Communications in Statistics - Theory and Methods*, 13(1):99–113.
- Segerstedt, B. (1992). On ordinary ridge regression in generalized linear models. *Communications in Statistics-Theory and Methods*, 21(8):2227–2246.
- Tweedie, M. C. et al. (1957). Statistical properties of inverse gaussian distributions. i. *The Annals of Mathematical Statistics*, 28(2):362–377.
- van Wieringen WN. (2015). Lecture notes on ridge regression. *arXiv preprint arXiv:1509.09169*, page 67.
- Vilca, F., Sanhueza, A., Leiva, V., and Christakos, G. (2010). An extended birnbaum-saunders model and its application in the study of environmental quality in santiago, chile. *Stochastic Environmental Research and Risk Assessment*, 24(5):771–782.
- Yahya Algamal, Z. (2019). Performance of ridge estimator in inverse gaussian regression model. *Communications in Statistics-Theory and Methods*, 48(15):3836–3849.

Yüzbaşı, B. and Asar, A. (2018). Ridge type estimation in the zero-inflated negative binomial regression. *Econometrics: Methods & Applications*, pages 93–104.