

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

TESTE DE DIAGNÓSTICO BASEADO EM
INFLUÊNCIA LOCAL APLICADO AO MODELO DE
REGRESSÃO SIMPLEX

FERNANDA CLOTILDE DA SILVA

DISSERTAÇÃO DE MESTRADO



Recife
2016

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

FERNANDA CLOTILDE DA SILVA

TESTE DE DIAGNÓSTICO BASEADO EM
INFLUÊNCIA LOCAL APLICADO AO MODELO DE
REGRESSÃO SIMPLEX

Orientadora: Prof^a. Dr^a. Patrícia Leone Espinheira Ospina
Co-Orientadora: Prof^a. Dr^a. Michelli Karinne Barros da Silva

Área de Concentração: Estatística Aplicada

Dissertação apresentada ao programa de Pós-graduação
em Estatística do Departamento de Estatística da Univer-
sidade Federal de Pernambuco, como requisito parcial para
obtenção do grau de **Mestre em Estatística**.

Recife
2016

Catálogo na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB 4-572

S586t Silva, Fernanda Clotilde da.
Teste de diagnóstico baseado em influência local aplicado ao modelo de regressão simplex / Fernanda Clotilde da Silva. – 2016.
139 f.: fig., tab.

Orientadora: Patrícia Leone Espinheira Ospina.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, Recife, 2016.
Inclui referências e apêndices.

1. Análise de regressão. 2. Modelos lineares. I. Espinheira, Patrícia Leone (Orientadora). II. Título.

519.536 CDD (22. ed.) UFPE-MEI 2016-29

FERNANDA CLOTILDE DA SILVA

**TESTE DE DIAGNÓSTICO BASEADO EM INFLUÊNCIA LOCAL APLICADO AO
MODELO DE REGRESSÃO SIMPLEX**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em: 12 de fevereiro de 2016.

BANCA EXAMINADORA

Prof.^a Dra. Patrícia Leone Espinheira Ospina
UFPE

Prof. Dr. Raydonal Ospina Martínez (Examinador Interno)
UFPE

Prof.^a Dra. Tarciana Liberal Pereira (Examinadora Externa)
UFPB

Dedico, com todo carinho, esse trabalho a minha mãe, Berenice Clotilde, e a meu pai, Edenilson Ferreira.

Agradecimentos

Primeiramente à Deus, por tudo o que tem feito na minha vida, principalmente por ter me concedido a capacidade de realizar esse trabalho e ter colocado no meu caminho pessoas maravilhosas, a quem tenho também muito o que agradecer.

Agradeço aos meus pais, pelo apoio e incentivo, também pelos conselhos e por todo amor para comigo, digo e repito vocês são minha vida.

Agradeço de coração à professora Michelli, que desde o início (“bem” lá no início, antes mesmo da graduação) até aqui tem sido muito mais que uma professora, tem sido uma amiga, uma conselheira, uma orientadora, uma incentivadora e uma tia maravilhosa, que esteve comigo nos momentos felizes, tensos e também bastante tristes, sempre transmitindo força e ânimo. Toda a sua dedicação e preocupação com esse trabalho e para comigo tem sido muito, muito, muito além do que qualquer pessoa possa esperar. Graças a sua disponibilidade constante (mesmo quando tinha mil outras coisas pra resolver) sempre conseguiu solucionar os problemas que surgiram no meio do caminho. Foi com a sua dedicação mais do que exclusiva e todo o seu carinho que pude começar, prosseguir e finalizar esse trabalho, tenho certeza que jamais conseguirei transmitir todo o agradecimento, respeito e admiração correspondente com o seu merecimento.

Sou muito grata à professora Patrícia, por todo carinho durante esse trajeto, pelo incentivo e principalmente por ter depositado em mim uma confiança que eu mesma nunca teria depositado, obrigada por ter acreditado em mim.

Agradeço também aos meus primos: minhas irmãs Dayana e Clarinha e meus primos Luana, Arthur e Mikaela, vocês alegram todos os meus dias, principalmente o pequeno “Arthuzinho”, amo muito cada um. Também a toda minha família no geral, em especial, a meu tio Horácio, que como sempre, tem me ajudado em tudo que foi preciso.

Aos meus tios, Preta e Nildo, por não medirem esforços em me ajudar desde o início da minha estadia em Recife até aqui, sempre disponíveis e prontos em todos os momentos.

Aos professores do Departamento de Estatística da UFCG, em especial aos professores Chico, Areli e Rosângela, por toda preocupação e apoio que demonstraram para comigo e para com esse trabalho, disponibilizando todos os recursos necessários para a parte de simulação.

Ao Professor Lindomberg por toda ajuda (bastante ajuda por sinal) disponibilizada, principalmente na parte técnica computacional, sempre pronto quando precisamos e também muito preocupado em otimizar nosso tempo de trabalho, à ele meus sinceros agradecimentos.

Ao professor Joelsom, por todo o incentivo aplicado em todos os momentos, pela disponibilidade em me ajudar nos artigos que foram de grande valia nesse trabalho e claro, pela alegria que a sua amizade e confiança me transmite.

À minha grande amiga e companheira de estrada acadêmica Terezinha, posso dizer que a sua amizade (iniciada na graduação) foi um grande presente para compensar a difícil vida de estudante, passamos por grandes momentos juntas, momentos felizes e as vezes tristes, porém sempre juntas. Sou muito grata por toda ajuda, incentivo e força que ela me deu, sem os quais não teria conseguido enfrentar os momentos difíceis, torço muito pelo seu sucesso e acredito que será enorme.

Aos meus amigos da graduação, Deyvison e Verinha, por todo incentivo e os momentos de alegria compartilhados.

À Wanessa, uma grande amizade que fiz no mestrado e acredito levar para a vida toda (assim como a amizade de Terezinha), aprendi bastante com ela também sei que será uma grande profissional.

A todos os meus colegas do mestrado: Rodrigo, Jonas, Laura, Pedro, Hildemar, Stênio, Telma, Rodrigo Goiano, Renata, Jeniffer, Raphaela, Eveline e Raquel.

A todos os Professores do Programa de Pós-graduação em Estatística da UFPE, por serem excelentes profissionais e contribuírem para minha formação acadêmica, assim como a formação dos meus colegas.

À Valéria Bittencourt, pela dedicação, carinho e paciência que tem demonstrado por todos os alunos da Pós-graduação em Estatística.

À CAPES, pelo apoio financeiro.

Seja bendito o nome de Deus para todo o sempre, porque dele é a sabedoria e a força; ele muda os tempos e as horas; ele remove os reis e estabelece os reis; ele dá sabedoria aos sábios e ciência aos inteligentes.

Livro de Daniel, Capítulo 2 e Versículos 20 e 21.

Resumo

O estudo de dados contínuos no intervalo $(0, 1)$ vem crescendo bastante nesses últimos anos. Na literatura destaca-se o modelo de regressão beta, proposto por Ferrari e Cribari-Neto (2004), como um dos principais modelos para analisar esses tipos dados. No entanto, para situações em que o ajuste do modelo beta não se torna adequado, surge uma contraproposta, o uso do modelo de regressão simplex, apresentado por Song e Tan (2000), que consiste em supor que a variável resposta possui uma distribuição simplex. Essa distribuição é derivada da distribuição Gaussiana inversa generalizada, e foi desenvolvida por Barndorff-Nielsen e Jørgensen (1991), sendo bastante estudada por Song e Tan (2000) e Song et al. (2004), que apresentaram várias propriedades importantes dessa distribuição. Entretanto, é possível haver casos em que o ajuste do modelo simplex também não seja adequado, sendo assim, esse trabalho consiste em desenvolver um teste de diagnóstico proposto, originalmente, por Zhu e Zhang (2004) para esse modelo. A construção do teste é baseado no método de influência local. Por meio de simulações de Monte Carlo avaliamos o desempenho do teste, observando o tamanho e o poder em diversas situações. Trabalhamos com os quantis empíricos da distribuição da estatística de teste, obtidos via método *bootstrap* paramétrico. Finalmente, realizamos algumas aplicações do teste, com dados reais e comparamos os resultados obtidos usando esse teste com os resultados segundo outros métodos de diagnóstico, bastante conhecidos na literatura.

Palavras-chave: Distribuição simplex. Teste de hipóteses. Influência local. Método *bootstrap* paramétrico.

Abstract

The study of continuous data in interval $(0,1)$ has grown much in recent years, in literature stands out the beta regression model, proposed by Ferrari and Cribari-Neto (2004), as one of the models to analyze these data. The simplex regression model proposed by Song and Tang (2000) based on simplex distribution is another option to modelling data on unit interval. This distribution was developed by Barndorff-Nielsen and Jørgensen (1991) and was derived from the generalized inverse Gaussian distribution. Song and Tang (2000) and Song et al. (2004) presented several important properties of this distribution. This work consist in develop a diagnostic test originally proposed by Zhu and Zhang (2004) for this model. The construction of the test is based on the local influence method (Cook, 1986). We also propose a new test, based on bootstrap empirical distribution of the original test statistic. The tests performance were evaluated based on size and/or power. Finally, were performed some applications using real data and the results of bootstrap test were compared with results obtained by other diagnostic methods well known in the literature.

Keywords: Simplex distribution. Hypothesis test. Local influence. Method parametric bootstrap.

Lista de Figuras

1.1	Linhas projetadas $LD(\delta_0 + aI_1) \times a$ e $LD(\delta_0 + aI_2) \times a$	25
2.1	Gráficos da densidade da distribuição simplex para $\sigma^2 = 0, 5, 1, 4$ e 16 , respectivamente, com diferentes valores de μ	32
4.1	Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 40$	78
4.2	Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 80$	78
4.3	Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 120$	79
4.4	Gráficos <i>qqplot</i> 's da distribuição de U_n para diferentes cenários de μ e va- lores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 40$	80
4.5	Gráficos <i>qqplot</i> 's da distribuição de U_n para diferentes cenários de μ e va- lores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 80$	80
4.6	Gráficos <i>qqplot</i> 's da distribuição de U_n para diferentes cenários de μ e va- lores de σ^2 iguais a $0, 5$ e 16 , respectivamente, para $n = 120$	80
4.7	Histogramas e gráficos <i>qqplot</i> 's da estatística U_n para diferentes cenários de μ , $\sigma^2 = 0, 5$ e $n = 500$	82
4.8	Histogramas e gráficos <i>qqplot</i> 's da estatística U_n para diferentes cenários de μ , $\sigma^2 = 16, 0$ e $n = 500$	83
4.9	Histogramas e gráficos <i>qqplot</i> 's da estatística U_n para diferentes cenários de μ , $\sigma^2 = 0, 5$ e $n = 1000$	84
4.10	Histogramas e gráficos <i>qqplot</i> 's da estatística U_n para diferentes cenários de μ , $\sigma^2 = 16, 0$ e $n = 1000$	85
4.11	Histogramas da estatística U_n para diferentes cenários de μ e graus de heterogeneidade iguais a $25, 91$ e $145, 37$, respectivamente.	101
4.12	Gráficos <i>qqplot</i> 's da distribuição de U_n para diferentes cenários de μ e graus de heterogeneidade iguais a $25, 91$ e $145, 37$, respectivamente.	102
5.1	Envelope dos resíduos ponderados padronizados para os dados do cloro. . .	107

5.2	Gráficos dos resíduos ponderados padronizados para os dados do cloro. . .	108
5.3	Gráfico dos possíveis pontos de alavanca e influentes para os dados do cloro, respectivamente.	108
5.4	Envelope dos resíduos ponderados padronizados para os dados dos transplantes autólogos	111
5.5	Gráficos dos resíduos ponderados padronizados para os dados dos transplantes autólogos.	112
5.6	Alguns gráficos de diagnósticos para os dados dos transplantes autólogos. .	112
5.7	Alguns gráficos de diagnósticos para os dados dos transplantes autólogos, para o modelo simplex com dispersão variável.	114
A.1	Curvatura da superfície $\alpha(\delta)$	122
A.2	Curvatura da superfície $\alpha(\delta)$ e a secção normal.	122

Lista de Tabelas

1.1	Algumas distribuições pertencentes aos modelos de dispersão e seus respectivos suportes, espaços paramétricos e desvios unitários.	23
4.1	Valores verdadeiros dos parâmetros β 's, fixados para estimar os tamanhos do teste.	77
4.2	Tamanho do teste usando os quantis da distribuição normal padrão.	77
4.3	Estimativas dos valores de algumas medidas características da distribuição de U_n , para $n = 40, 80$ e 120	81
4.4	Estimativas dos valores de algumas medidas características da distribuição de U_n , para $n = 500$ e $n = 1000$	86
4.5	Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando dados da distribuição beta.	86
4.6	Poder do teste para o modelo simplex com dispersão fixa, quando os dados são gerados da distribuição beta.	87
4.7	Tamanho do teste utilizando quantis <i>bootstrap</i>	88
4.8	Poder do teste para o modelo simplex com dispersão fixa, quando os dados são gerados da distribuição beta, utilizando quantis <i>bootstrap</i>	89
4.9	Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando dados da distribuição simplex com 4 covariáveis.	89
4.10	Tamanhos e poderes do teste em detectar covariáveis omitidas, cujo modelo verdadeiro é dado por $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, com $t = 1, \dots, n$ e dispersão fixa.	91
4.11	Valores verdadeiros dos parâmetros β 's, fixados para obter o poder do teste, gerando dados de um modelo simplex com dispersão variável e ajustando o modelo simplex com dispersão fixa.	92
4.12	Valores verdadeiros dos parâmetros γ 's, fixados para obter o poder do teste, gerando dados de um modelo simplex com dispersão variável e ajustando o modelo simplex com dispersão fixa.	93
4.13	Poderes do teste para o modelo simplex com dispersão fixa, quando os dados são gerados com dispersão variável.	94

4.14	Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando beta, usando a variância <i>bootstrap</i>	95
4.15	Tamanho do teste para o modelo simplex com dispersão fixa usando a variância <i>bootstrap</i>	95
4.16	Poder do teste para o modelo simplex com dispersão fixa usando a variância <i>bootstrap</i> , quando os dados são gerados da distribuição beta.	96
4.17	Tamanhos e poderes do teste em detectar covariáveis omitidas, cujo modelo verdadeiro é dado por $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, com $t = 1, \dots, n$ e dispersão fixa, usando a variância <i>bootstrap</i>	97
4.18	Poderes do teste para o modelo simplex com dispersão fixa usando a variância <i>bootstrap</i> , quando os dados são gerados com dispersão variável.	99
4.19	Valores verdadeiros dos parâmetros β 's, fixados para obter os histogramas dos valores da estatística U_n , para o modelo com dispersão variável.	100
4.20	Valores verdadeiros dos parâmetros γ 's, fixados para obter os histogramas dos valores da estatística U_n , para o modelo com dispersão variável.	101
4.21	Valores verdadeiros dos parâmetros γ 's, fixados para calcular o poder do teste, gerando dados da distribuição beta.	102
4.22	Poder do teste para o modelo simplex com dispersão variável, quando os dados são gerados apresentam distribuição beta com dispersão variável.	103
4.23	Valores verdadeiros dos parâmetros β 's e γ 's, fixados para calcular o poder do teste, gerando pela distribuição simplex com 4 covariáveis.	104
4.24	Poderes do teste em detectar covariáveis omitidas do modelo para a média, quando o modelo verdadeiro é dado por $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $h(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, com $t = 1, \dots, n$	105
5.1	Estimativas dos parâmetros do modelo de regressão simplex ajustado aos dados do cloro.	106
5.2	Resultados do teste de diagnóstico para os dados do cloro.	107
5.3	Mudanças relativas em porcentagens para as estimativas dos parâmetros do modelo para os dados do cloro.	109
5.4	Estimativas dos parâmetros do modelo de regressão simplex ajustado aos dados dos transplantes autólogos.	110
5.5	Resultados do teste de diagnóstico para os dados dos transplantes autólogos.	111
5.6	Estimativas dos parâmetros do modelo de regressão simplex com dispersão variável, ajustado aos dados dos transplantes autólogos.	113
5.7	Resultados do teste de diagnóstico para os dados dos transplantes autólogos, para o modelo com dispersão variável.	113
5.8	Mudanças relativas em porcentagens para as estimativas dos parâmetros do modelo com dispersão variável para os dados dos transplantes autólogos.	115
D.1	Conjunto de dados usados na aplicação I, referente a fração de cloro disponível.	136

D.2	Conjunto de dados usados na aplicação II, referente aos transplantes autó- logos.	137
-----	--	-----

Sumário

1	Introdução	17
1.1	Introdução	17
1.1.1	Organização da Dissertação	20
1.1.2	Suporte Computacional	20
1.2	Revisão Bibliográfica	21
1.2.1	Introdução aos Modelos de Dispersão	21
1.2.2	Método de Influência Local	23
1.2.3	Método <i>Bootstrap</i>	27
2	Modelo de Regressão Simplex	31
2.1	Distribuição Simplex	31
2.2	Modelo de Regressão Simplex com Dispersão Fixa	33
2.2.1	Função Escore, Matriz Hessiana e Matriz de Informação de Fisher	33
2.2.2	Estimação dos Parâmetros	39
2.3	Análise de Influência Local no Modelo Simplex	40
2.3.1	Ponderação de Casos	41
2.4	Modelo Simplex com Dispersão Variável	42
2.4.1	Função Escore, Matriz Hessiana e Matriz de Informação de Fisher para os Parâmetros do Modelo	43
2.4.2	Estimação dos Parâmetros	47
2.4.3	Análise de Influência Local para o Modelo com Dispersão Variável	48
3	Um Teste de Diagnóstico Para o Modelo Simplex	50
3.1	Teste de Diagnóstico Baseado em Influência Local	50
3.2	O Teste para o Modelo Simplex Com Dispersão Fixa	56
3.3	O Teste para o Modelo Simplex Com Dispersão Variável	67
4	Avaliação Numérica dos Testes	75
4.1	Avaliação do Teste para o Modelo Simplex com Dispersão Fixa	76
4.1.1	Avaliação do Teste Usando Quantis <i>Bootstrap</i>	87

4.1.2	Avaliação do Teste Usando Variância <i>Bootstrap</i>	95
4.2	Avaliação do Teste para o Modelo Simplex com Dispersão Variável	100
5	Algumas Aplicações do Teste	106
5.1	Aplicação I: Dados do Cloro	106
5.2	Aplicação II: Dados dos Transplantes Autólogos de Células Tronco do Sangue Periférico	110
6	Considerações finais	116
6.1	Conclusão	116
6.2	Trabalhos Futuros	117
	Referências	118
A	Gráficos das Curvaturas Normais	122
B	Propriedades da Distribuição Simplex	123
B.1	Prova do item (iv) da Proposição 2.1:	126
B.2	Prova dos itens (v) e (vi) da Proposição 2.1:	128
C	Prova dos Teoremas	130
C.1	Provando o Teorema 1:	130
C.2	Prova do Teorema 2:	133
D	Dados das Aplicações	136

Capítulo 1

Introdução

1.1 Introdução

É bastante comum nos estudos estatísticos haver o interesse de explicar o comportamento de uma variável aleatória específica (chamada de variável resposta) por meio de uma ou várias variáveis (chamadas de variáveis explicativas ou covariáveis). Por muitos anos, o uso do modelo de regressão linear foi a solução comumente utilizada com esse objetivo. Esse método busca associar a média da variável resposta com uma função linear das covariáveis. Dessa forma, cada valor da variável resposta está associado a uma função linear das covariáveis junto com um erro aleatório não observável, em que é suposto que esse erro tenha uma distribuição normal de média zero e variância constante. Sendo assim, sob essas suposições, tem-se que a variável resposta possui uma distribuição normal e esse modelo é definido como modelo de regressão normal linear.

No entanto, em muitas situações a variável resposta possui uma distribuição diferente da normal e a relação entre o seu valor esperado e as covariáveis não é diretamente descrito de uma forma linear. Assim, buscando fornecer algumas opções de distribuições para essa variável, Nelder e Wedderburn (1972) propuseram os modelos lineares generalizados (MLG's), que se baseiam em considerar que a variável resposta pertence a família exponencial e garantem que a relação entre a sua média e a função linear das covariáveis seja mais flexível.

Um procedimento bastante importante realizado após ajustar um modelo de regressão a um conjunto de dados é avaliar a adequabilidade do ajuste, esse procedimento é conhecido como análise de diagnóstico e procura verificar possíveis afastamentos das suposições feitas para o modelo. Para mais detalhes ver Paula (2013). Na literatura encontramos diversas técnicas de diagnóstico, muitas foram desenvolvidas baseando-se na identificação de observações que destoam das demais, chamadas de observações atípicas. Inicialmente, Cox e Snell (1968) procuraram identificar esse tipo de observação por meio da análise dos resíduos, pois pontos que apresentam resíduos altos (conhecidos como pontos aberrantes ou *outlier*) podem afetar além da média dos valores preditos da resposta também as estimativas dos parâmetros do modelo. Belsley et al. (1980) e Cook e Weisberg (1982)

realizaram algumas padronizações para os resíduos, aplicados no caso de regressão normal linear, enquanto que Pregibon (1981) sugeriu uma nova padronização dos resíduos para os modelos lineares generalizados.

Outro tipo de observação atípica são os pontos de alavanca, que correspondem às observações que apresentam um padrão atípico nas variáveis explicativas. A identificação desses pontos é dada pelo uso de uma matriz de projeção apresentada por Hoaglin e Welsch (1978), dependendo de sua localização, esse tipo de observação exerce um peso desproporcional nos valores ajustados e causam mudanças nas estimativas. Existem ainda os chamados pontos influentes, que também exercem peso desproporcional nas estimativas dos parâmetros, sua identificação consiste na avaliação do impacto causado nas estimativas após a retirada desses pontos. A distância de Cook é uma das medidas de influência mais conhecida e muito usada para detectar esses pontos, foi proposta por Cook (1977) para o modelo normal linear e estendida para diversas classes de modelos.

Outra técnica de diagnóstico bastante importante é o método de influência local, apresentada por Cook (1986). Esse método consiste em estudar o comportamento de uma determinada medida de influência quando são introduzidas pequenas mudanças (ou perturbações) nos dados ou no modelo, dessa forma é possível identificar a presença de observações que sob pequenas alterações nos dados ou no modelo, modificam consideravelmente os resultados do ajuste, ver Paula (2013). O método de influência local é considerado um dos métodos de análise de diagnóstico mais modernos e diversos estudos foram desenvolvidos por meio do uso dessa técnica.

Atkinson (1981) desenvolveu um gráfico de diagnóstico para detectar afastamento da distribuição postulada para a variável resposta. A sua metodologia consiste na construção via simulação de Monte Carlo de uma banda de confiança para os resíduos, de forma que, se a distribuição assumida não for adequada os pontos ou uma parte deles encontram-se fora das banda de confiança, esse gráfico é comumente chamado de “gráfico de envelope”. Ainda no contexto de análise de diagnóstico, foram desenvolvidos alguns testes de hipóteses, com a finalidade de testar se o modelo ajustado aos dados está adequado. Um teste bastante conhecido é o RESET, proposto por Ramsey (1969), que foi criado para o modelo de regressão linear e procura adicionar ao modelo uma quantidade não-linear por meio de termos de diferentes graus, para em seguida testar a exclusão de tais termos pelo usual teste F. Os autores, Zhu e Zhang (2004) desenvolveram um teste de diagnóstico baseado em influência local, que busca avaliar a má especificação do modelo. Esse teste pode ser desenvolvido para diferentes classes de distribuições, além disso, eles também propuseram um método para detectar pontos simultaneamente influentes.

Atualmente, existem diversos estudos em que o interesse consiste em analisar dados contínuos no intervalo $(0, 1)$. Em alguns casos é realizada uma transformação nos dados de modo que eles assumam valores reais e a média da resposta possa ser modelada por um modelo de regressão normal linear. No entanto, esse procedimento não se mostra eficaz, pois os parâmetros do modelo não podem ser facilmente interpretados em termos da resposta original, além disso, medidas de proporções normalmente exibem assimetria, portanto, a

análise baseada no pressuposto de normalidade pode não ser confiável. Buscando solução para esse problema, Ferrari e Cribari-Neto (2004) propuseram o modelo de regressão beta. O uso desse modelo tem crescido bastante nos últimos anos e os métodos para avaliar a adequabilidade do seu ajuste, estão cada vez mais aprimorados. Lima (2007) propôs um teste de especificação correta do modelo de regressão beta, cujo desenvolvimento teve como base o teste RESET.

Contudo, é provável haver casos em que o ajuste do modelo beta à um determinado conjunto de dados não seja adequado, então, uma alternativa proposta é ajustar o modelo de regressão simplex, introduzido por Song e Tan (2000), em que supõem-se que a variável resposta possui uma distribuição simplex. A distribuição simplex foi criada por Barndorff-Nielsen e Jørgensen (1991), sendo desenvolvida a partir da distribuição inversa Gaussiana generalizada. Uma propriedade importante sobre essa distribuição é que sua função densidade de probabilidade não pertence a família exponencial, consequentemente, ela não faz parte dos modelos lineares generalizados. Entretanto, Jørgensen (1997) estendeu a ideia que Nelder e Wedderburn (1972) usaram para definir os modelos lineares generalizados e obtiveram uma classe de modelos que cobrem uma rica faixa de distribuições não-normais, chamados de modelos de dispersão, os quais contêm a distribuição simplex.

Alguns trabalhos foram desenvolvidos com intuito de modelar dados no intervalo $(0, 1)$ por meio da distribuição simplex. Song e Tan (2000) utilizaram essa distribuição para avaliar dados longitudinais, sob a suposição do parâmetro de dispersão ser constante, além de apresentar várias propriedades importantes dessa distribuição. Song et al. (2004) utilizaram essa distribuição para avaliar dados longitudinais, sob a suposição de que o parâmetro de dispersão varia ao longo das observações. Venezuela (2007) buscou estimar os parâmetros do modelo simplex usando equações de estimação generalizada (EEG). Miyashiro (2008) estudou o modelo de regressão simplex com dispersão fixa, apresentando diversas propriedades desse modelo e algumas medidas de diagnóstico e Santos (2011) desenvolveu algumas correções para o resíduo de Pearson do modelo simplex com dispersão fixa. López (2013) utilizou o enfoque bayesiano para avaliar os estimadores do modelo simplex com dispersão variável. Recentemente, Oliveira (2015) propôs o modelo de regressão simplex não linear, além de apresentar algumas medidas de diagnóstico e Meireles (2015) propôs duas novas medidas de diagnóstico, construídas a partir dos resíduos ponderado.

O objetivo principal desse trabalho é desenvolver o teste de diagnóstico proposto por Zhu e Zhang (2004) para os modelos de regressão simplex, com dispersão fixa e também variável, além de avaliar, por meio de um estudo de simulação de Monte Carlo, o desempenho dos testes propostos para ambos os modelos. Também, propomos um melhoramento nesses testes, baseados na metodologia *bootstrap*, introduzida por Efron (1979). Portanto, esse trabalho fornece um método bastante útil para identificar se o ajuste do modelo de regressão simplex a um determinado conjunto de dados está ou não adequado, tanto para situações em que a dispersão é fixa quanto variável.

1.1.1 Organização da Dissertação

O presente trabalho, encontra-se dividido em seis capítulos e um apêndice. No primeiro capítulo realizamos uma breve revisão bibliográfica sobre alguns temas da área de estatística, que foram de grande utilidade no desenvolvimento do trabalho. Foram apresentados os modelos de dispersão, proposto por Jørgensen (1997), o método de influência local, proposto por Cook (1986) e por último, o método *bootstrap*, introduzido por Efron (1979). No segundo capítulo introduzimos a distribuição simplex e algumas propriedades, seguido pelo modelo de regressão simplex para o caso em que o parâmetro de dispersão é fixo e também para o caso em que esse parâmetro é variável. Nesse capítulo obtivemos a função score, a matriz de informação de Fisher, os estimadores e a matriz hessiana para os parâmetros do modelo nos dois casos. Também, obtivemos as quantidades necessárias para realizar uma análise de influência local nesses modelos.

No terceiro capítulo, apresentamos o teste de diagnóstico proposto por Zhu e Zhang (2004). Ainda nesse capítulo, desenvolvemos esse teste para o modelo simplex com dispersão fixa e para o modelo com dispersão variável.

No quarto capítulo, avaliamos o desempenho dos testes, bem como, dos melhoramentos realizados nestes, baseados na metodologia *bootstrap* para os modelos simplex com dispersão fixa e variável. Foram realizadas diversas simulações de Monte Carlo, com intuito de observar o tamanho e o poder do teste em diversas situações, como por exemplo, quando os dados vêm de uma distribuição diferente da simplex, quando o parâmetro de dispersão é variável e quando são omitidas covariáveis durante o ajuste.

No quinto capítulo, aplicamos o modelo de regressão simplex a diversos conjuntos de dados reais e verificamos a adequação do ajuste por meio do teste proposto. Também comparamos os resultados obtidos por esse teste com os resultados segundo outros métodos de diagnósticos bastante conhecidos na literatura. Por último, no sexto capítulo, são apresentadas as conclusões gerais, obtidas pelo trabalho. Além disso, no final, disponibilizamos no apêndice, algumas propriedades extras da distribuição simplex, junto com os conjuntos de dados utilizados nas aplicações.

1.1.2 Suporte Computacional

Todos os resultados numéricos, obtidos pelas simulações e pelas aplicações, presentes no quarto e quinto capítulo, respectivamente, foram obtidos por meio da linguagem de programação **0x**, na sua versão 7.1, utilizando o sistema operacional *Windows*. Esta linguagem de programação foi desenvolvida por Jurgen Doornik em 1994, e encontra-se disponível em <http://www.doornik.com>, para mais detalhes, ver Doornik (2001).

Os resultados gráficos, foram produzidos utilizando o ambiente computacional e estatístico R, na sua versão 3.2.1, no sistema operacional *Windows*, O R foi criado por Ross Ihaka e Robert Gentleman em 1996, e trata-se de uma implementação da linguagem S, porém com a vantagem de ser livre e possuir código-fonte aberto e encontra-se disponível

gratuitamente em <http://www.r-project.org>. Para mais detalhes ver Peternelli e Mello (2011).

A digitação do presente trabalho, se deu por uso do sistema de tipografia L^AT_EX, criado por Leslie Lamport em 1985, em que trata-se de um conjunto de macros ou rotinas T_EX. O T_EX por sua vez é um sistema de editor de textos, usado para produzir livros, artigos, etc., com alta qualidade, desenvolvido por Donald E. Knuth. Para mais detalhes sobre o sistema de tipografia L^AT_EX, ver Knuth (1986) ou Lamport (1994).

1.2 Revisão Bibliográfica

Apresentaremos uma breve revisão sobre a classe dos modelos de dispersão, o método de influência local e o método *Bootstrap*. Esses temas foram de grande importância no desenvolvimento desse trabalho.

1.2.1 Introdução aos Modelos de Dispersão

Quando se tem o interesse de estudar o comportamento de uma variável aleatória dependente, y , dado outras variáveis, o modelo de regressão é um dos métodos mais utilizados para investigar essa relação. Particularmente, o modelo de regressão normal linear.

No entanto, existem situações em que a variável resposta possui uma distribuição diferente da normal, bem como a relação entre o seu valor esperado e as covariáveis não é descrito de uma forma linear.

Nelder e Wedderburn (1972) propuseram a classe de modelos lineares generalizados (MLG) que são constituídos de duas componentes. A primeira é a variável resposta, comumente chamada de componente aleatória, que deve possuir uma distribuição pertencente a família exponencial. Ou seja, considerando $y_1, y_2, \dots, y_n$ uma amostra aleatória de n variáveis aleatórias independentes, temos que a função densidade ou de probabilidade de cada uma pode ser escrita na forma

$$f(y_t, \theta_t, \phi) = \exp[\phi\{y_t\theta_t - b(\theta_t)\} + c(y_t, \phi)], \quad (1.1)$$

em que $t = 1, 2, \dots, n$ e as funções $b(\cdot)$ e $c(\cdot)$ são ambas conhecidas. Em (1.1) é possível notar que $\mathbb{E}(y_t) = \mu_t = b'(\theta_t)$ e $\text{Var}(y_t) = \phi^{-1}V_t$, em que ϕ^{-1} é o parâmetro de dispersão e $V = d\mu/d\theta$ é função de variância utilizada para caracterizar a distribuição.

Existem várias distribuições importantes que pertencem a família exponencial, algumas delas são: normal, Poisson, gama, binomial, Rayleigh e normal inversa, Cordeiro et al. (1995) apresentam várias outras distribuições com essa característica.

De acordo com Paula (2013), a outra componente que define os MLG's é a componente sistemática, dada por

$$g(\mu_t) = \eta_t, \quad (1.2)$$

em que $\eta_t = \mathbf{x}_t^\top \boldsymbol{\beta}$ é o preditor linear, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros desconhecidos, com $p < n$, $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tp})^\top$ os valores de p variáveis explicativas (em geral assumimos $x_{t1} = 1$, o que garante o intercepto do modelo) e $g(\cdot)$ é uma função monótona e diferenciável, conhecida como função de ligação. A componente sistemática garante que a relação entre a sua média e o preditor linear seja mais flexível.

Sabendo que os MLG's foram desenvolvidos originalmente para modelar dados cuja distribuição pertence a família exponencial, Jørgensen (1997) estendeu a ideia para cobrir uma rica faixa de distribuições não-normais, incluindo os modelos lineares generalizados, formando assim os “**Modelos de Dispersão**”.

Segundo Song (2009), definimos como um modelo de dispersão, $\mathbf{MD}(\mu, \sigma^2)$, com parâmetro de locação μ e parâmetro de dispersão σ^2 , uma família de distribuições cuja função densidade de probabilidade pode ser reescrita na seguinte forma:

$$p(y; \mu, \sigma^2) = a(y; \sigma^2) \exp \left\{ -\frac{1}{2\sigma^2} d(y; \mu) \right\}, \quad (1.3)$$

com, $y \in \mathcal{C}$, $\mu \in \Omega$ e $\sigma^2 > 0$, em que $\mathcal{C}$ corresponde ao suporte da função (1.3) e Ω ao espaço paramétrico de μ , enquanto que $a(\cdot) \geq 0$ é uma função regularmente adequada em termos de ser independente de μ .

A função $d(\cdot; \cdot)$ é chamada de *desvio unitário* ou *Componente de desvio* com $(y, \mu) \in \mathcal{C} \times \Omega$, se esta satisfizer as seguintes condições:

1. quando a observação y for igual a μ a função é zero, ou seja,

$$d(y; y) = 0, \quad \forall y \in \Omega;$$

2. quando a observação y for diferente de μ então a função é positiva, ou seja,

$$d(y; \mu) > 0, \quad \forall y \neq \mu.$$

O desvio unitário é chamado de *regular* se a função $d(\cdot; \cdot)$ for duas vezes continuamente diferenciável com respeito a (y, μ) em $\Omega \times \Omega$ e satisfizer

$$\left. \frac{\partial^2 d}{\partial \mu^2}(y; \mu) \right|_{\mu=y} > 0, \quad \forall y \in \Omega.$$

De acordo com Santos (2011), os desvios unitários são definidos pela diferença entre o logaritmo da função de verossimilhança com $y = \mu$, definida por $\ell(y, y)$ e o logaritmo da função de verossimilhança com $y \neq \mu$, definida por $\ell(y, \mu)$, ou seja,

$$d(y; \mu) = c[\ell(y, y) - \ell(y, \mu)], \quad (1.4)$$

em que c é uma constante tal que (1.4) seja satisfeita.

Na Tabela 1.1 são apresentadas algumas distribuições de probabilidades que pertencem aos modelos de dispersão, juntamente com seus suportes, seus espaços paramétricos para μ e seus desvios unitários.

Tabela 1.1: Algumas distribuições pertencentes aos modelos de dispersão e seus respectivos suportes, espaços paramétricos e desvios unitários.

Distribuição	$\mathcal{C}$	Ω	$d(y; \mu)$
Normal	$(-\infty, \infty)$	$(-\infty, \infty)$	$(y - \mu)^2$
Poisson	$\{0, 1, \dots\}$	$(0, \infty)$	$2[y \log \left(\frac{y}{\mu} - y + \mu \right)]$
Binomial	$\{0, 1, \dots, n\}$	$(0, 1)$	$2[y \log \left(\frac{y}{\mu} + (n - y) \log \left(\frac{n - y}{n - \mu} \right) \right)]$
Gaussiana Inversa	$(0, \infty)$	$(0, \infty)$	$\frac{(y - \mu)^2}{y\mu^2}$
Von Mises	$(0, 2\pi)$	$(0, 2\pi)$	$2[1 - \cos(y - \mu)]$
Simplex	$(0, 1)$	$(0, 1)$	$\frac{(y - \mu)^2}{y(1 - y)\mu^2(1 - \mu)^2}$

Para mais detalhes sobre modelos de dispersão ver Jørgensen (1997).

1.2.2 Método de Influência Local

A análise de influência local é um procedimento de diagnóstico proposto por Cook em 1986. De acordo com Paula (2013), ele é considerado um dos métodos mais modernos de analisar a adequação do modelo ajustado. Com esse método é possível verificar a presença de observações que modificam consideravelmente os resultados a partir do momento em que o modelo ou os dados sofrem pequenas alterações ou perturbações.

A ideia básica do método consiste em inserir pequenas perturbações nos dados ou no modelo, e avaliar o comportamento de uma determinada medida de influência, observando-se assim, o quanto as estimativas fornecidas pelo modelo ajustado variam após a presença dessas perturbações. Em resumo, podemos dizer que o método de influência local verifica o quão sensíveis são as estimativas que o modelo fornece.

Procedimento Básico:

Segundo Paula (2013), considerando $\ell(\boldsymbol{\theta})$ o logaritmo da função de verossimilhança do modelo postulado, com $\boldsymbol{\theta}$ sendo um vetor p -dimensional, que contém os parâmetros desse modelo, e seja $\boldsymbol{\delta}$ um vetor de perturbações, de ordem $q \times 1$, de forma que $\boldsymbol{\delta} \in \Omega \in \mathbb{R}^q$, em que geralmente $q = n$, temos que $\ell_{\boldsymbol{\delta}}(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança perturbada, de modo que denotaremos $\boldsymbol{\delta}_0$ como sendo o vetor de não perturbação, consequentemente $\ell_{\boldsymbol{\delta}_0}(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta})$.

Existem diversas maneiras de perturbação, algumas são:

- ponderação de casos - é inserida uma perturbação δ_i no logaritmo da função de

verossimilhança, correspondente a t -ésima observação, de forma que

$$\ell_{\delta}(\boldsymbol{\theta}) = \sum_{t=1}^n \delta_t \ell_t(\boldsymbol{\theta}),$$

em que $0 \leq \delta_t \leq 1$.

- perturbação da variável resposta - para cada observação y_t , acrescentamos uma perturbação, de modo que o seu valor perturbado é dado por

$$y_{t\delta} = y_t + \delta_t \sigma_{y_t},$$

em que σ_{y_t} é o desvio padrão de y_t , com $\delta_t \in \mathbb{R}$.

- perturbação de uma covariável - para cada observação da covariável x_{tk} acrescentamos uma perturbação, de modo que o seu valor perturbado é dado por

$$x_{tk\delta} = x_{tk} + \delta_t \sigma_{x_{tk}},$$

em que $\sigma_{x_{tk}}$ é o desvio padrão de x_{tk} , com $\delta_t \in \mathbb{R}$ e $k = 2, \dots, p$.

De acordo com Paula (2013), a medida de influência utilizada para avaliar o efeito das perturbações nas estimativas de máxima verossimilhança, fornecidas pelo modelo ajustado, é o deslocamento pela verossimilhança, dado por

$$LD(\boldsymbol{\delta}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\delta}})], \quad (1.5)$$

com $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{\boldsymbol{\delta}}$ sendo os vetores de estimadores de máxima verossimilhança, obtidos sob o modelo não perturbado $\ell(\boldsymbol{\theta})$ e sob o modelo perturbado $\ell_{\boldsymbol{\delta}}(\boldsymbol{\theta})$, respectivamente.

Cook (1986), propôs estudar o comportamento da função (1.5) em torno de $\boldsymbol{\delta}_0$, de modo que $LD(\boldsymbol{\delta}_0) = 0$. Assim, considerando a seguinte superfície geométrica:

$$\alpha(\boldsymbol{\delta}) = \begin{bmatrix} \boldsymbol{\delta} \\ LD(\boldsymbol{\delta}) \end{bmatrix}, \quad (1.6)$$

a ideia proposta foi avaliar o modo como essa superfície desvia-se de seu plano tangente em $\boldsymbol{\delta}_0$, à medida que $\boldsymbol{\delta}$ se afasta lentamente de $\boldsymbol{\delta}_0$, ou seja, quando pequenas perturbações são introduzidas no modelo.

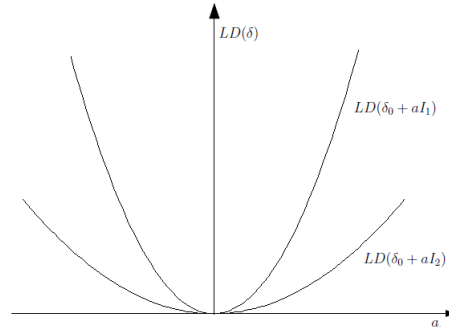
A compreensão desse método, torna-se mais fácil quando realizamos um estudo das curvaturas normais da superfície $\alpha(\boldsymbol{\delta})$ em $\boldsymbol{\delta}_0$, é possível observar o gráfico dessa superfície juntamente com seu plano tangente em $\boldsymbol{\delta}_0$ e o vetor ortogonal a esse plano chamado de $\pi(\tau_0)$, na Figura (A.1) e também em Espinheira (2007).

Por Espinheira (2007), considerar que uma leve perturbação seja inserida no modelo é equivalente a escolher uma direção unitária I , ou seja, $\|I\| = 1$, e tomar $\boldsymbol{\delta} = \boldsymbol{\delta}_0 + I$,

logo $LD(\delta_0 + I) \neq 0$. Na Figura (A.2) podemos notar a existência de um outro plano cuja interseção com $\alpha(\delta)$ forma o que chamamos de secção normal, ela é especificada pela direção I associada ao vetor $\pi(\tau_0)$. Assim, o gráfico $LD(\delta_0 + aI)$ contra $a \in \mathbb{R}$ apresenta a secção normal caracterizada por sua curvatura normal em torno de $a = 0$, denotada por C_I , esse gráfico é conhecido como linha projetada.

Seguindo essa linha de raciocínio, Cook (1986) sugere analisar o vetor de direção I , que especifica a linha projetada com maior curvatura normal (C_{max}), esse vetor é denotado por I_{max} . Na Figura 1.1 encontramos uma ilustração das linhas projetadas $LD(\delta_0 + aI_1) \times a$ e $LD(\delta_0 + aI_2) \times a$, que foram especificadas pelas direções I_1 e I_2 , respectivamente. Nela podemos observar que à medida que os valores de a se afastam do zero a curvatura normal da linha projetada associada a I_2 torna-se maior que a curvatura da linha projetada associada a I_1 , conseqüentemente, o deslocamento pela verossimilhança, $LD(\delta)$, cresce mais rapidamente na direção de I_2 . Essa figura foi extraída de Espinheira (2007).

Figura 1.1: Linhas projetadas $LD(\delta_0 + aI_1) \times a$ e $LD(\delta_0 + aI_2) \times a$.



Fonte: Espinheira (2007).

Baseando-se em conceitos de geometria diferencial, Cook (1986) também mostra que a curvatura normal na direção I é dada por

$$C_I(\delta) = |I^\top H_{LD}(\delta) I|, \quad (1.7)$$

avaliada em $\delta = \delta_0$. De forma que

$$\begin{aligned} H_{LD}(\delta_0) &= \left. \frac{\partial^2 LD(\delta)}{\partial \delta \partial \delta^\top} \right|_{\delta=\delta_0} \\ &= \left. -2 \frac{\partial^2 \ell(\hat{\theta}_\delta)}{\partial \delta \partial \delta^\top} \right|_{\delta=\delta_0} \\ &= \left. 2 \Delta^\top(\theta) [-\ddot{\ell}(\theta)]^{-1} \Delta(\theta) \right|_{\delta=\delta_0, \theta=\hat{\theta}}, \end{aligned} \quad (1.8)$$

em que $\ddot{\ell}(\boldsymbol{\theta}) = \partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$ é a matriz hessiana, cuja (i, j) -ésima entrada é dada por

$$[\ddot{\ell}(\boldsymbol{\theta})]_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}, \quad (1.9)$$

com $i, j = 1, \dots, p$ e $\Delta(\boldsymbol{\theta}) = \partial^2 \ell_{\boldsymbol{\delta}}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\delta}^\top$ é uma matriz $p \times q$, cuja (r, s) -ésima entrada, com $r = 1, \dots, p$ e $s = 1, \dots, q$, é dada por

$$[\Delta(\boldsymbol{\theta})]_{rs} = \frac{\partial^2 \ell_{\boldsymbol{\delta}}(\boldsymbol{\theta})}{\partial \theta_r \partial \delta_s}. \quad (1.10)$$

Assim, podemos expressar a curvatura normal na direção I , avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\delta} = \boldsymbol{\delta}_0$, por

$$C_I(\boldsymbol{\theta}) = 2|I^\top \Delta^\top(\boldsymbol{\theta})[-\ddot{\ell}(\boldsymbol{\theta})]^{-1} \Delta(\boldsymbol{\theta}) I|. \quad (1.11)$$

É possível haver situações em que existe apenas uma observação influente e que não conseguimos detectar pela análise do vetor I_{max} , pois essa não produz um alto nível de deslocamento da verossimilhança. Procurando uma solução para esse caso, Lesaffre e Verbeke (1998) sugeriram construir o gráfico da curvatura normal na direção da t -ésima observação, ou seja, a direção usada agora será um vetor com o t -ésimo elemento igual a 1 e os demais iguais a zero, denotado por I_t .

Zhu e Zhang (2004) afirmaram que a curvatura normal na direção do vetor canônico é dado por

$$C_{I_t} = \sum_{m=1}^p \lambda_m v_{mt}^2, \quad (1.12)$$

em que (λ_m, v_m) com $m = 1, \dots, p$ são pares de autovalores e autovetores da matriz $H_{LD}(\boldsymbol{\delta}_0)$, de modo que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq \lambda_{p+1} = \dots = \lambda_n = 0$ e $v_m = (v_{m1}, \dots, v_{mn})^\top$ corresponde a uma base ortogonal associada a essa matriz.

Espinheira (2007) indica utilizar na análise de influência local os seguintes gráficos:

- componentes do vetor I_{max} contra os índices das observações - para sugerir quais observações são conjuntamente influentes;
- C_{I_t} 's contra os índices das observações - para indicar observações individualmente influente. Lesaffre e Verbeke (1998) sugerem usar como ponto de corte a quantidade $\sum_{t=1}^n C_{I_t} / n$.

Tendo como objetivo entender as propriedades de C_{I_t} , Zhu e Zhang (2004) também apresentaram a proposição dada a seguir. Esta proposição foi de grande importância para esses autores definirem o teste de diagnóstico apresentado no terceiro capítulo desse trabalho

PROPOSIÇÃO 1.1: Assumindo que $-\ddot{\ell}(\hat{\boldsymbol{\theta}})/n \xrightarrow{p} A$ e $[\Delta \Delta^\top(\boldsymbol{\theta})/n]_{\boldsymbol{\delta}_0, \hat{\boldsymbol{\theta}}} \xrightarrow{p} B$, em que A é uma matriz positiva definida B é uma matriz não-negativa definida. Seja $2\text{tr}(A^{-1}B) = \lambda_0$.

Então,

$$\begin{aligned} \sum_{t=1}^n C_{I_t} &= \text{tr}(H_{LD}(\boldsymbol{\delta}_0)) \\ &= 2\text{tr} \left\{ \left[\frac{-\ddot{\ell}(\boldsymbol{\theta})}{n} \right]^{-1} \frac{\Delta \Delta^\top(\boldsymbol{\theta})}{n} \right\} \Big|_{\boldsymbol{\delta}_0, \hat{\boldsymbol{\theta}}} \xrightarrow{p} \lambda_0, \end{aligned} \quad (1.13)$$

em que $\xrightarrow{p}$, denota convergência em probabilidade, para mais detalhes sobre esse tipo de convergência ver James (1981) ou Magalhães (2006).

1.2.3 Método *Bootstrap*

É notável como o desenvolvimento de novos métodos de análises estatísticas vem crescendo nos últimos anos. Um dos principais fatores que tem contribuído bastante para esse crescimento, foi o uso intenso de recursos computacionais, que permitiu simular diversas vezes a aplicação desses métodos, e assim verificar a sua eficiência e o quão confiáveis são seus resultados. Essa técnica é conhecida como simulação.

O método *bootstrap*, introduzido por Efron (1979), trata-se de uma classe da técnica de simulação de Monte Carlo, cuja ideia básica, consiste em tratar a amostra original como se fosse a própria população, e assim obter réplicas dos dados, por meio de uma reamostragem com reposição, ou pelo ajuste do modelo ao qual os dados pertencem.

Um dos grandes objetivos desse método é determinar a distribuição de probabilidade empírica para uma determinada estatística, quando conhecemos ou não o modelo gerador dos dados, também, podemos estimar o viés e o erro padrão dessa estatística. Além disso, o método também é bastante usado para obter intervalos de confiança e teste de hipóteses. Para mais detalhes ver Efron e Tibishiran (1993).

Existem dois tipos de *bootstrap*, a saber: não-paramétrico e paramétrico. Baseando-se em Furtado (2008), podemos descrever a metodologia das duas técnicas *bootstrap* da maneira como é dada a seguir.

***Bootstrap* Não Paramétrico:**

Nesse caso, seja $y = (y_1, \dots, y_n)$ uma amostra aleatória de n variáveis aleatórias, independentes e identicamente distribuídas, cuja função densidade de probabilidade é uma função desconhecida, denotada por $f(y)$.

Realizando uma reamostragem com reposição, de tamanho n , da amostra y , temos que a função de distribuição empírica de y_t^* , com $t = 1, \dots, n$, é dada por

$$F(y_t^*) = P(y_t^* = y_i) = \frac{1}{n}, \quad (1.14)$$

em que $i = 1, \dots, n$. Logo, as variáveis aleatórias $y_1^*, \dots, y_n^*$ são independentes e identica-

mente distribuídas e além disso estão distribuídas uniformemente em $\{y_1, \dots, y_n\}$.

Dessa forma, seja T a estatística de interesse, a estimativa de *bootstrap* da função de distribuição de T , definida por $F(T)$, é obtida da seguinte maneira:

1. Geramos B amostras de tamanho n , com reposição de y , ou seja, obtemos

$$\begin{aligned} y_1^* &= (y_{11}^*, y_{12}^*, \dots, y_{1n}^*); \\ &\vdots \\ y_B^* &= (y_{B1}^*, y_{B2}^*, \dots, y_{Bn}^*). \end{aligned}$$

2. Para cada subamostra, calculamos a estatística de interesse, obtendo assim $T_1^*, \dots, T_B^*$.

Obtidas as B estimativas de T , temos que a distribuição empírica dessas réplicas servirão com estimativa de $F(T)$. Dessa forma, podemos também estimar os quantis empíricos da distribuição dessa estatística e também observar o possível formato da sua função densidade, por meio do seu histograma.

Para estimar por *bootstrap*, o erro padrão de T , denotado por $ep(T)$, calculamos o erro padrão dos estimadores das B réplicas, isto é

$$\hat{ep}(T) = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (T_b^* - \bar{T}^*)^2}, \quad (1.15)$$

em que, $\bar{T}^* = (\sum_{b=1}^B T_b^*)/B$.

De acordo com Efron e Tibishiran (1986), quando B vai para infinito, temos que

$$\lim_{B \rightarrow \infty} \hat{ep}(T) = ep(T).$$

Bootstrap Paramétrico:

Nesse caso, dispomos de toda informação necessária sobre a distribuição dos dados, logo, as amostras *bootstrap* são obtidas por meio de uma amostragem diretamente nessa distribuição, em que os parâmetros desconhecidos são substituídos pelas suas respectivas estimativas, obtidas através da amostra original.

Dessa forma, a estimativa de *bootstrap* da função de distribuição da estatística T é obtida da seguinte maneira:

1. Geramos B amostras de tamanho n , da distribuição ao qual supomos que os dados pertencem, assumindo as estimativas dos parâmetros, como sendo seus verdadeiros valores.
2. Para cada subamostra, calculamos a estatística de interesse, obtendo, respectivamente, $T_1^*, \dots, T_B^*$.

3. Calculamos a distribuição empírica baseada em $b = 1, \dots, B$ e a usamos como estimativa de $F(T)$.

O procedimento para estimar o erro padrão de T , nessa situação, é análogo ao caso não paramétrico.

Intervalo de Confiança *Bootstrap* Percentil- t

De acordo com Cook et al. (2001), considerando que θ corresponde à característica de interesse da população, podem ser construídos vários tipos de intervalos de confiança para θ , a partir da reamostragem *bootstrap*, quando é suposto que o estimador desse parâmetro, $\hat{\theta}$, possui uma distribuição normal, e que $ep(\hat{\theta})$ é o erro padrão desse estimador. Logo, a construção do intervalo de confiança para θ , deverá utilizar a seguinte estatística:

$$Z = \frac{\hat{\theta} - \theta}{ep(\hat{\theta})}, \quad (1.16)$$

em que (1.16), possui uma distribuição aproximadamente $N(0, 1)$. No entanto, geralmente, não conhecemos $ep(\hat{\theta})$, tendo que ser estimado, uma alternativa é construir o intervalo de confiança baseando-se na estatística

$$T = \frac{\hat{\theta} - \theta}{\hat{ep}(\hat{\theta})}, \quad (1.17)$$

assumindo que (1.17) tem distribuição t de Student, em que $\hat{ep}(\hat{\theta})$ será obtido via *bootstrap*. Assim, temos que

$$\left[\hat{\theta} - t^*_{1-\alpha/2} \hat{ep}(\hat{\theta}); \hat{\theta} + t^*_{\alpha/2} \hat{ep}(\hat{\theta}) \right], \quad (1.18)$$

corresponde ao intervalo percentil- t , de nível de confiança $100(1 - \alpha)\%$, em que $t^*_{1-\alpha/2}$ e $t^*_{\alpha/2}$ são os quantis de ordem $1 - \alpha/2$ e $\alpha/2$, respectivamente, da distribuição de T , obtidos via método *bootstrap*. Para obter esse intervalo, precisamos realizar os seguintes passos:

1. calcular a estatística $\hat{\theta}$, utilizando a amostra original;
2. Para cada réplica *bootstrap* ($b = 1, \dots, B$) executar as seguintes tarefas:
 - 2.1 obter uma amostra *bootstrap* da amostra original, $y_b^* = (y_{b1}^*, \dots, y_{bn}^*)$;
 - 2.2 calcular a estimativa *bootstrap* da estatística de interesse, $\hat{\theta}_b^*$;
 - 2.3 realizar um *bootstrap* de segundo nível, com $b_1 = 1, \dots, B_1$ réplicas para estimar o erro padrão de $\hat{\theta}_b^*$, denotado por $\hat{ep}(\hat{\theta}_b^*)$;
 - 2.4 calcular a estatística

$$T_b^* = \frac{\hat{\theta}_b^* - \hat{\theta}}{\hat{ep}(\hat{\theta}_b^*)};$$

3. calcular $\widehat{ep}(\widehat{\theta})$ a partir das réplicas $\widehat{\theta}_1^*, \dots, \widehat{\theta}_B^*$;
4. obter os quantis $t_{1-\alpha/2}^*$ e $t_{\alpha/2}^*$, a partir das réplicas $T_1^*, \dots, T_B^*$;
5. calcular os limites dos intervalos.

Devido ao fato de que precisamos estimar o erro padrão por meio de um segundo nível *bootstrap*, o que caracteriza o *bootstrap* duplo, observamos que o custo computacional nesse caso é altamente elevado.

Capítulo 2

Modelo de Regressão Simplex

2.1 Distribuição Simplex

A distribuição simplex, denotada por $S^-(\mu, \sigma^2)$, com parâmetros de posição e dispersão dados, respectivamente, por $0 < \mu < 1$ e $\sigma^2 > 0$ possui uma função densidade de probabilidade dada por

$$p(y; \mu, \sigma^2) = \{2\pi\sigma^2[y(1-y)]^3\}^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}d(y; \mu)\right\}, \quad (2.1)$$

para $0 < y < 1$, em que

$$d(y; \mu) = \frac{(y - \mu)^2}{y(1-y)\mu^2(1-\mu)^2}$$

é o desvio unitário. A esperança e variância dessa distribuição são dadas, respectivamente, por

$$\mathbb{E}(Y) = \mu$$

e

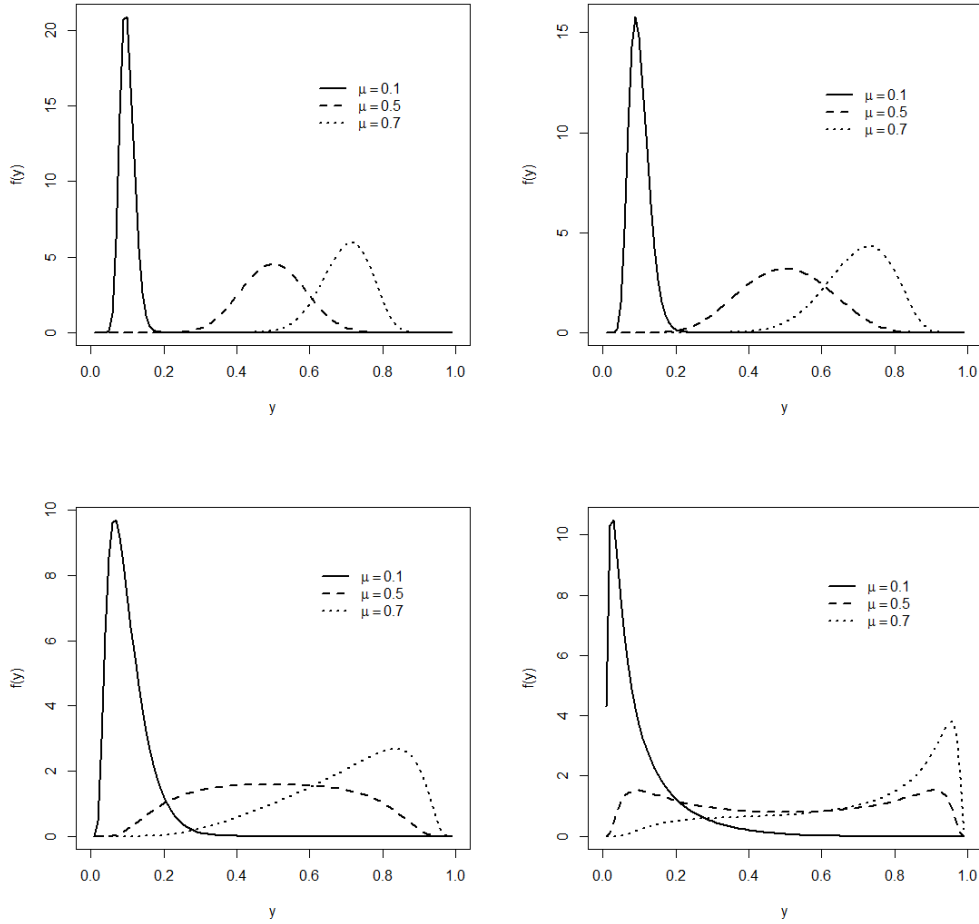
$$\text{Var}(Y) = \mu(1-\mu) - \sqrt{\frac{1}{2\sigma^2}} \exp\left\{\frac{1}{\sigma^2\mu^2(1-\mu)^2}\right\} \Gamma\left(\frac{1}{2}, \frac{1}{2\sigma^2\mu^2(1-\mu)^2}\right),$$

em que $\Gamma(a, b)$ corresponde a função gama incompleta, definida por $\Gamma(a, b) = \int_b^\infty x^{a-1}e^{-x}dx$.

Na Figura 2.1 são apresentadas alguns gráficos referentes a densidade da distribuição simplex, para diferentes valores de μ e σ^2 . O primeiro corresponde a $\sigma^2 = 0,5$, e podemos observar que para valores pequenos de μ a distribuição apresenta valores mais concentrados próximos da média. No segundo gráfico, em que $\sigma^2 = 1$, observamos comportamentos semelhantes ao primeiro, porém, a distribuição assume formas mais achatadas, entretanto, a medida que σ^2 aumenta a distribuição assume formatos bastante assimétricos,

como podemos ver nos dois últimos gráficos, em que $\sigma^2 = 4$ e $\sigma^2 = 16$, respectivamente.

Figura 2.1: Gráficos da densidade da distribuição simplex para $\sigma^2 = 0, 5, 1, 4$ e 16 , respectivamente, com diferentes valores de μ .



Fonte: Autoria própria.

Na proposição a seguir apresentamos algumas propriedades importantes da distribuição simplex. Elas serão de grande utilidade na maioria dos cálculos realizados no presente trabalho.

PROPOSIÇÃO 2.1: Suponha que $y \sim S^-(\mu, \sigma^2)$. Então, valem as seguintes propriedades:

- (i) $\mathbb{E}[d(y; \mu)] = \sigma^2$;
- (ii) $\mathbb{E}[(y - \mu)d'(y; \mu)] = -2\sigma^2$;
- (iii) $\mathbb{E}[(y - \mu)d(y; \mu)] = 0$;
- (iv) $\mathbb{E}[(y - \mu)d^2(y; \mu)] = 0$;
- (v) $\mathbb{E}[d'(y; \mu)] = 0$;

- (vi) $\frac{1}{2}\mathbb{E}[d''(y; \mu)] = \frac{3\sigma^2}{\mu(1-\mu)} + \frac{1}{\mu^3(1-\mu)^3};$
 (vii) $\text{Var}[d(y; \mu)] = 2(\sigma^2)^2.$

As provas dos itens (i), (ii) e (iii) são encontradas em Song e Tan (2000). Song et al. (2004) provaram o item (vii) e no Apêndice desse trabalho, apresentamos a prova dos itens (iv), (v) e (vi).

2.2 Modelo de Regressão Simplex com Dispersão Fixa

Sejam $y_1, \dots, y_n$ variáveis aleatórias independentes, em que cada y_t , $t = 1, \dots, n$ segue uma distribuição simplex, cuja função densidade de probabilidade é dada por (2.1), com média μ_t e dispersão σ^2 , que vamos assumir fixa, porém desconhecida.

O modelo de regressão simplex é definido por (2.1) e pela componente sistemática

$$g(\mu_t) = \sum_{i=1}^p x_{ti}\beta_i = \eta_t, \quad (2.2)$$

em que, $\eta_t = \mathbf{x}_t^\top \boldsymbol{\beta}$ é o preditor linear, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos, com $\boldsymbol{\beta} \in \mathbb{R}^p$ e $(p < n)$, $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})^\top$ são valores fixos e conhecidos, assumidos por p variáveis explicativas e $g(\cdot)$ é uma função estritamente monótona e diferenciável, conhecida como função de ligação. Essa função além de relacionar a média da variável resposta com o preditor linear, também nos permite trabalhar com valores que pertencem aos reais, valores esses de μ_t que antes pertenciam ao intervalo $(0, 1)$. Existem diversos tipos de funções de ligações que são bastante utilizadas, algumas delas são:

- probito: $g(\mu_t) = \Phi^{-1}(\mu_t)$, em que $\Phi(\cdot)$ é a função de distribuição acumulada da normal padrão;
- log-log: $g(\mu_t) = \log[-\log(\mu_t)]$;
- complemento log-log: $g(\mu_t) = \log\{-\log(1 - \mu_t)\}$;
- logito: $g(\mu_t) = \log(\mu_t/(1 - \mu_t))$.

2.2.1 Função Escore, Matriz Hessiana e Matriz de Informação de Fisher

Baseado em (2.1), temos que o logaritmo da função de verossimilhança é

$$\ell(\boldsymbol{\beta}, \sigma^2) = \sum_{t=1}^n \ell_t(\mu_t, \sigma^2), \quad (2.3)$$

em que

$$\ell_t(\mu_t, \sigma^2) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{3}{2} \log[y_t(1 - y_t)] - \frac{1}{2\sigma^2} d(y_t; \mu_t).$$

Considerando a equação (2.3) podemos obter a função escore para o vetor de parâmetros β calculando, inicialmente, a função escore para β_i , com $i = 1, \dots, p$. Assim, temos que

$$\begin{aligned} \frac{\partial \ell(\beta, \sigma^2)}{\partial \beta_i} &= \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \beta_i} \\ &= \sum_{t=1}^n \left\{ \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right\} \\ &= \sum_{t=1}^n \left\{ \frac{1}{\sigma^2} \frac{(y_t - \mu_t)}{\mu_t(1 - \mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1 - \mu_t)^2} \right] \frac{1}{g'(\mu_t)} x_{ti} \right\}. \end{aligned}$$

Note que $d\mu_t/d\eta_t = 1/g'(\mu_t)$, $\partial \eta_t/\partial \beta_i = x_{ti}$ e

$$\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} = \frac{1}{\sigma^2} \frac{(y_t - \mu_t)}{\mu_t(1 - \mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1 - \mu_t)^2} \right],$$

(Detalhes, ver Apêndice B.2). Logo, a função escore para β pode ser escrita na seguinte forma matricial:

$$U_\beta(\beta, \sigma^2) = \frac{1}{\sigma^2} X^\top T U(\mathbf{y} - \boldsymbol{\mu}), \quad (2.4)$$

com X sendo uma matriz de posto completo de ordem $n \times p$, com a i -ésima linha denotada por $\mathbf{x}_t^\top$ e

$$T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\},$$

$$U = \text{diag}\{u_1, \dots, u_n\},$$

sendo

$$u_t = \frac{1}{\mu_t(1 - \mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1 - \mu_t)^2} \right],$$

e finalmente, $(\mathbf{y} - \boldsymbol{\mu}) = (y_1 - \mu_1, \dots, y_n - \mu_n)^\top$.

Considerando a função (2.3), temos que a função escore para o parâmetro de dispersão

σ^2 é dada por

$$\begin{aligned}
 U_{\sigma^2}(\boldsymbol{\beta}, \sigma^2) &= \frac{\partial \ell(\boldsymbol{\beta}, \sigma^2)}{\partial \sigma^2} \\
 &= \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \sigma^2} \\
 &= \sum_{t=1}^n -\frac{1}{2\sigma^2} \left[1 - \frac{d(y_t; \mu_t)}{\sigma^2} \right] \\
 &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^n d(y_t; \mu_t),
 \end{aligned} \tag{2.5}$$

(Detalhes desse resultado, ver Apêndice B.4).

Para obtermos a matriz de informação de Fisher vamos obter, primeiramente, as derivadas de segunda ordem do logaritmo da função de verossimilhança, ou seja, a matriz hessiana, dada por

$$\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) = \begin{bmatrix} \ddot{\ell}_{\beta\beta}(\boldsymbol{\theta}) & \ddot{\ell}_{\beta\sigma^2}(\boldsymbol{\theta}) \\ \ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta}) & \ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta}) \end{bmatrix} \tag{2.6}$$

sendo $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \sigma^2)^\top$. Da matriz (2.6) verificamos que

$$\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top}, \tag{2.7}$$

em que, o (i, j) -ésimo elemento de (2.7) com $i = 1, \dots, p$ e $j = 1, \dots, p$ é

$$\begin{aligned}
 [\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_i} \\
 &= \sum_{t=1}^n \frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial \beta_j \partial \beta_i} \\
 &= \sum_{t=1}^n \frac{\partial}{\partial \beta_j} \left[\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \beta_i} \right] \\
 &= \sum_{t=1}^n \frac{\partial}{\partial \beta_j} \left[\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \\
 &= \sum_{t=1}^n \frac{\partial}{\partial \mu_t} \left[\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_j} \\
 &= \sum_{t=1}^n \frac{\partial}{\partial \mu_t} \left[\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \frac{1}{g'(\mu_t)} \right] \frac{1}{g'(\mu_t)} x_{ti} x_{tj} \\
 &= \sum_{t=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial \mu_t^2} \frac{1}{g'(\mu_t)} + \frac{\partial}{\partial \mu_t} \left[\frac{1}{g'(\mu_t)} \right] \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \right\} \frac{1}{g'(\mu_t)} x_{ti} x_{tj}
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial \mu_t^2} \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} + \frac{\partial}{\partial \mu_t} \left[\frac{1}{g'(\mu_t)} \right] \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} \frac{1}{g'(\mu_t)} x_{ti} x_{tj} \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial \mu_t^2} \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} - \frac{g''(\mu_t)}{[g'(\mu_t)]^3} \frac{1}{\sigma^2} (y_t - \mu_t) u_t x_{ti} x_{tj} \right\} \\
&= \sum_{t=1}^n \frac{1}{\sigma^2} \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} \left[\frac{-d(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - \frac{1}{\mu_t^3(1 - \mu_t)^3} - \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \right. \\
&\quad \left. - \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} + \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - (y_t - \mu_t)u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right] \\
&= \sum_{t=1}^n \frac{1}{\sigma^2} \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} \left[-u_t - \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} - \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} \right. \\
&\quad \left. + \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - (y_t - \mu_t)u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right], \tag{2.8}
\end{aligned}$$

lembrando que

$$\begin{aligned}
\frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial \mu_t^2} &= \frac{1}{\sigma^2} \left[\frac{-d(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - \frac{1}{\mu_t^3(1 - \mu_t)^3} - \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \right. \\
&\quad \left. - \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} + \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \right],
\end{aligned}$$

$$\frac{\partial}{\partial \mu_t} \left[\frac{1}{g'(\mu_t)} \right] = -\frac{g''(\mu_t)}{[g'(\mu_t)]^2} \quad \text{e} \quad \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \mu_t} = \frac{1}{\sigma^2} (y_t - \mu_t) u_t.$$

(Detalhes ver Apêndice B.6). Assim, definindo,

$$\begin{aligned}
q_t &= \left[-u_t - \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} - \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} \right. \\
&\quad \left. + \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - (y_t - \mu_t)u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right] \frac{1}{[g'(\mu_t)]^2}, \tag{2.9}
\end{aligned}$$

temos que

$$[\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij} = \sum_{t=1}^n \frac{1}{\sigma^2} x_{ti} x_{tj} q_t. \tag{2.10}$$

Dessa forma, matricialmente, temos que

$$\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} X^\top Q X, \tag{2.11}$$

em que $Q = \text{diag}\{q_1, \dots, q_n\}$.

Prosseguindo com o intuito de determinar os elementos da matriz hessiana e sabendo que

$$\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2 \partial \boldsymbol{\beta}^\top}, \quad (2.12)$$

temos que o i -ésimo elemento de (2.12) é dado por

$$\begin{aligned} [\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]_i &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2 \partial \beta_i} \\ &= \sum_{t=1}^n \frac{\partial}{\partial \sigma^2} \left[\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \beta_i} \right] \\ &= \sum_{t=1}^n \frac{\partial}{\partial \sigma^2} \left\{ \frac{1}{\sigma^2} \frac{(y_t - \mu_t)}{\mu_t(1 - \mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1 - \mu_t)^2} \right] \frac{1}{g'(\mu_t)} x_{ti} \right\} \\ &= \sum_{t=1}^n \frac{\partial}{\partial \sigma^2} \left[\frac{1}{\sigma^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti} \right] \\ &= \sum_{t=1}^n -\frac{1}{(\sigma^2)^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti}. \end{aligned} \quad (2.13)$$

Definindo,

$$f_t = -\frac{1}{(\sigma^2)^2} (y_t - \mu_t) u_t$$

obtemos, então,

$$[\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]_i = \sum_{t=1}^n f_t \frac{1}{g'(\mu_t)} x_{ti}. \quad (2.14)$$

Reescrevendo matricialmente, temos

$$\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta}) = \mathbf{f}^\top T X, \quad (2.15)$$

em que $\mathbf{f} = (f_1, \dots, f_n)^\top$. Vale observar que $\ddot{\ell}_{\beta\sigma^2}(\boldsymbol{\theta}) = \ddot{\ell}_{\sigma^2\beta}^\top(\boldsymbol{\theta}) = X^\top T \mathbf{f}$.

Finalizando, segue que

$$\begin{aligned} \ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta}) &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial (\sigma^2)^2} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell_t(\mu_t, \sigma^2)}{\partial (\sigma^2)^2} \\ &= \sum_{t=1}^n \left[\frac{1}{2(\sigma^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma^2)^3} \right]. \end{aligned} \quad (2.16)$$

(Detalhes ver Apêndice B.8). Denotando,

$$d_t = \frac{1}{2(\sigma^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma^2)^3}, \quad (2.17)$$

temos que

$$\ddot{\ell}_{\sigma^2 \sigma^2}(\boldsymbol{\theta}) = \text{tr}(D), \quad (2.18)$$

em que $D = \text{diag}\{d_1, \dots, d_n\}$.

Usando o resultado apresentado em (2.8) e pela propriedades mostradas na Proposição 2.1, obtemos que

$$-\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_i} \right] = \sum_{t=1}^n \left\{ \frac{1}{\sigma^2} \left[\frac{3\sigma^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3} \right] \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} \right\}$$

Assim, a informação de Fisher para $\boldsymbol{\beta}$ é dada por

$$\begin{aligned} K_{\beta\beta} &= -\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right] \\ &= \frac{1}{\sigma^2} X^\top W X, \end{aligned} \quad (2.19)$$

em que $W = \text{diag}\{w_1, \dots, w_n\}$, com

$$w_t = \left[\frac{3\sigma^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3} \right] \frac{1}{[g'(\mu_t)]^2}.$$

De acordo com (2.16) e utilizando novamente a Proposição 2.1, temos que a informação de Fisher para o parâmetro σ^2 será

$$\begin{aligned} K_{\sigma^2 \sigma^2} &= -\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial (\sigma^2)^2} \right] \\ &= \sum_{t=1}^n -\mathbb{E} \left[\frac{1}{2(\sigma^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma^2)^3} \right] \\ &= \sum_{t=1}^n \frac{1}{2(\sigma^2)^2} \\ &= \frac{n}{2(\sigma^2)^2}. \end{aligned}$$

Finalmente, baseando-se na expressão (2.13) e ainda pela Proposição 2.1, é fácil ver que $\mathbb{E}[(y_t - \mu_t)u_t] = 0$, assim,

$$\mathbb{E} \left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2 \partial \beta_i} \right) = 0.$$

Consequentemente, $K_{\beta\sigma^2} = K_{\sigma^2\beta}^\top = 0$, logo, podemos dizer que os parâmetros β e σ^2 são ortogonais.

Portanto, a matriz de informação de Fisher para $\theta = (\beta^\top, \sigma^2)^\top$ é bloco diagonal dada por $K_{\theta\theta} = \text{diag}\{K_{\beta\beta}, K_{\sigma^2\sigma^2}\}$, ou seja,

$$K_{\theta\theta} = \begin{bmatrix} K_{\beta\beta} & 0 \\ 0 & K_{\sigma^2\sigma^2} \end{bmatrix}. \quad (2.20)$$

Para os casos em que o tamanho amostral é consideravelmente grande e as condições de regularidades são satisfeitas, temos que a distribuição assintótica dos estimadores de máxima verossimilhança é dada por

$$\hat{\theta} \sim N_{p+1}(\theta; K_{\theta\theta}^{-1}),$$

em que $\hat{\theta} = (\hat{\beta}^\top, \hat{\sigma}^2)^\top$ é o estimador de máxima verossimilhança de $\theta = (\beta^\top, \sigma^2)^\top$ e $K_{\theta\theta}^{-1}$ é a inversa da matriz de informação de Fisher, dada por (2.20).

2.2.2 Estimação dos Parâmetros

Pela definição dos estimadores de máxima verossimilhança, sabemos que o estimador para β é obtido como solução da equação $U_\beta(\beta, \sigma^2) = 0$. Porém, nesse caso, o estimador de máxima verossimilhança de β não possui forma fechada, sendo necessário a utilização de métodos numéricos para estimar o vetor de parâmetros β .

Um método bastante utilizado é o método iterativo de Newton-Raphson, que se baseia em expandir em série de Taylor a função escore U_β , em torno de um ponto inicial arbitrário. Assim, sob essa ideia, é possível chegar ao seguinte processo iterativo

$$\beta^{(m+1)} = \beta^{(m)} + [-(U'_\beta)^{-1}]^{(m)} U_\beta^{(m)}, \text{ com } m = 0, \dots,$$

em que vale notar que $-(U'_\beta)$ corresponde a matriz observada de β . No entanto, um problema que esse método apresenta é o fato da matriz $-(U'_\beta)$ nem sempre ser positiva definida quando β estiver longe de seu valor ótimo, e também, em muitas situações o custo computacional em trabalhar com a matriz observada pode ser elevado, então, uma solução é substituir $-(U'_\beta)$ pelo seu correspondente valor esperado, ou seja, pela matriz de informação de Fisher $K_{\beta\beta}$, definida em (2.19). Isso resulta no processo iterativo Scoring de Fisher, definido por

$$\beta^{(m+1)} = \beta^{(m)} + [K_{\beta\beta}^{-1}]^{(m)} U_\beta^{(m)}. \quad (2.21)$$

O processo anterior é repetido até que a diferença entre $\beta^{(m+1)}$ e $\beta^{(m)}$ seja menor que uma tolerância especificada. Trabalhando o lado direito da expressão (2.21), utilizando a

função escore de β definida por (2.4) e a matriz (2.19), obtemos

$$\begin{aligned}\beta^{(m+1)} &= \beta^{(m)} + \sigma^2 (X^\top W^{(m)} X)^{-1} (\sigma^2)^{-1} X^\top T^{(m)} U^{(m)} (\mathbf{y} - \boldsymbol{\mu}^{(m)}) \\ &= \beta^{(m)} + (X^\top W^{(m)} X)^{-1} X^\top T^{(m)} U^{(m)} (\mathbf{y} - \boldsymbol{\mu}^{(m)}),\end{aligned}$$

denotando,

$$\mathbf{z}^{(m)} = X\beta^{(m)} + [W^{-1}]^{(m)} T^{(m)} U^{(m)} (\mathbf{y} - \boldsymbol{\mu}^{(m)}),$$

temos que

$$\beta^{(m+1)} = (X^\top W^{(m)} X)^{-1} X^\top W^{(m)} \mathbf{z}^{(m)},$$

que representa um processo iterativo de mínimos quadrados ponderados, em que $\mathbf{z}$ desempenha o papel de uma variável dependente e W a matriz de pesos.

Utilizando a equação $\mathbf{U}_{\sigma^2}(\beta, \sigma^2) = 0$, em que $\mathbf{U}_{\sigma^2}(\beta, \sigma^2)$ é definido em (2.5), podemos verificar que o estimador de máxima verossimilhança para σ^2 é dado por

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n d(y_t; \hat{\mu}_t). \quad (2.22)$$

2.3 Análise de Influência Local no Modelo Simplex

Na primeiro capítulo, apresentamos uma breve introdução ao método de influência local, para um modelo de regressão qualquer. Agora, usando o modelo de regressão simplex com dispersão fixa, iremos obter a matriz $\Delta(\boldsymbol{\theta})$, considerando o método de perturbação ponderação de casos.

Dessa forma, iremos considerar o modelo de regressão definido em (2.2) em que $\boldsymbol{\theta} = (\beta^\top, \sigma^2)^\top$ e $\Delta(\boldsymbol{\theta})$ é uma matriz de ordem $(p+1) \times n$, de modo que

$$\begin{aligned}\Delta(\boldsymbol{\theta}) &= \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\delta}^\top} \\ &= \begin{bmatrix} \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_1 \partial \delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_1 \partial \delta_n} \\ \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_2 \partial \delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_2 \partial \delta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_p \partial \delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \beta_p \partial \delta_n} \\ \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \sigma^2 \partial \delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \sigma^2 \partial \delta_n} \end{bmatrix},\end{aligned} \quad (2.23)$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\delta} = \boldsymbol{\delta}_0$.

2.3.1 Ponderação de Casos

Seja $\ell(\boldsymbol{\theta})$ o logaritmo da função de verossimilhança correspondente ao modelo postulado. Considerando que seja realizada neste modelo, uma perturbação através de um vetor $\boldsymbol{\delta}$ de ordem $n \times 1$, então, o logaritmo da função de verossimilhança do modelo perturbado fica dado por

$$\ell_{\delta}(\boldsymbol{\theta}) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \sigma^2),$$

em que $0 \leq \delta_t \leq 1$. Desse modo, temos que

$$\begin{aligned} \frac{\partial \ell_{\delta}(\boldsymbol{\theta})}{\partial \delta_t} &= \frac{\partial}{\partial \delta_t} \left[\sum_{t=1}^n \delta_t \ell_t(\mu_t, \sigma^2) \right] \\ &= \ell_t(\mu_t, \sigma^2). \end{aligned}$$

Logo,

$$\begin{aligned} \frac{\partial^2 \ell_{\delta}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \delta_t} &= \frac{\partial}{\partial \boldsymbol{\theta}} \left[\frac{\partial \ell_{\delta}(\boldsymbol{\theta})}{\partial \delta_t} \right] \\ &= \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \boldsymbol{\theta}} \\ &= \Delta_t(\boldsymbol{\theta}), \end{aligned}$$

que corresponde a t -ésima coluna da matriz $\Delta(\boldsymbol{\theta})$. Ou seja, $\Delta_t(\boldsymbol{\theta})$ indica a taxa de variação do logaritmo da função de verossimilhança em relação a $\boldsymbol{\theta}$ na t -ésima observação. Dessa forma, temos que a matriz (2.23), para esse tipo de perturbação, fica dada por

$$\Delta(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial \ell_1(\mu_1, \sigma^2)}{\partial \beta_1} & \dots & \frac{\partial \ell_n(\mu_n, \sigma^2)}{\partial \beta_1} \\ \frac{\partial \ell_1(\mu_1, \sigma^2)}{\partial \beta_2} & \dots & \frac{\partial \ell_n(\mu_n, \sigma^2)}{\partial \beta_2} \\ \vdots & \ddots & \vdots \\ \frac{\partial \ell_1(\mu_1, \sigma^2)}{\partial \beta_p} & \dots & \frac{\partial \ell_n(\mu_n, \sigma^2)}{\partial \beta_p} \\ \frac{\partial \ell_1(\mu_1, \sigma^2)}{\partial \sigma^2} & \dots & \frac{\partial \ell_n(\mu_n, \sigma^2)}{\partial \sigma^2} \end{bmatrix},$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Sabemos que para o modelo de regressão simplex

$$\frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \beta_i} = \frac{1}{\sigma^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti}, \quad \text{com } i = 1, \dots, p \quad (2.24)$$

e

$$\begin{aligned} \frac{\partial \ell_t(\mu_t, \sigma^2)}{\partial \sigma^2} &= -\frac{1}{2\sigma^2} \left\{ 1 - \frac{d(y_t; \mu_t)}{\sigma^2} \right\} \\ &= \frac{-\sigma^2 + d(y_t; \mu_t)}{2(\sigma^2)^2}, \end{aligned} \quad (2.25)$$

com $t = 1, \dots, n$. Segue então

$$\begin{aligned}\Delta_{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \begin{bmatrix} \frac{1}{\sigma^2}(y_1 - \mu_1)u_1 \frac{1}{g'(\mu_1)}x_{11} & \cdots & \frac{1}{\sigma^2}(y_n - \mu_n)u_n \frac{1}{g'(\mu_n)}x_{n1} \\ \frac{1}{\sigma^2}(y_1 - \mu_1)u_1 \frac{1}{g'(\mu_1)}x_{12} & \cdots & \frac{1}{\sigma^2}(y_n - \mu_n)u_n \frac{1}{g'(\mu_n)}x_{n2} \\ \vdots & \ddots & \vdots \\ \frac{1}{\sigma^2}(y_1 - \mu_1)u_1 \frac{1}{g'(\mu_1)}x_{1p} & \cdots & \frac{1}{\sigma^2}(y_n - \mu_n)u_n \frac{1}{g'(\mu_n)}x_{np} \end{bmatrix} \\ &= \frac{1}{\sigma^2} \mathbf{X}^\top T U \mathcal{E},\end{aligned}$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, em que $\mathcal{E} = \text{diag}\{y_1 - \mu_1, \dots, y_n - \mu_n\}$. Temos também que

$$\begin{aligned}\Delta_{\sigma^2}(\boldsymbol{\theta}) &= \frac{\partial^2 \ell_{\delta}(\boldsymbol{\theta})}{\partial \sigma^2 \partial \boldsymbol{\delta}^\top} \\ &= \begin{bmatrix} \frac{-\sigma^2 + d(y_1; \mu_1)}{2(\sigma^2)^2}, & \dots, & \frac{-\sigma^2 + d(y_n; \mu_n)}{2(\sigma^2)^2} \end{bmatrix} \\ &= \mathbf{J}^\top,\end{aligned}$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Então,

$$\Delta(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{\sigma^2} \mathbf{X}^\top T U \mathcal{E} \\ \mathbf{J}^\top \end{bmatrix}. \quad (2.26)$$

2.4 Modelo Simplex com Dispersão Variável

Na seção anterior apresentamos o modelo simplex supondo que o parâmetro de dispersão era constante. No entanto, podem existir casos em que essa suposição é violada, e esse parâmetro varia ao longo das observações. Nessa seção, apresentaremos um modelo de regressão simplex que modela não apenas a média, mas também a dispersão.

Sejam $y_1, \dots, y_n$ variáveis aleatórias independentes, em que cada y_t , $t = 1, \dots, n$ segue uma distribuição simplex, cuja função densidade de probabilidade é apresentada por (2.1), com média μ_t e dispersão σ_t^2 .

O modelo de regressão simplex com dispersão variável admite que a média e a dispersão satisfazem as seguintes relações funcionais

$$g(\mu_t) = \sum_{i=1}^p x_{ti} \beta_i = \eta_{1t}, \quad (2.27)$$

$$h(\sigma_t^2) = \sum_{i=1}^q z_{ti} \gamma_i = \eta_{2t}, \quad (2.28)$$

em que $g(\cdot)$ e $h(\cdot)$, são funções estritamente monótonas e diferenciáveis, com preditores lineares $\eta_{1t} = \mathbf{x}_t^\top \boldsymbol{\beta}$ e $\eta_{2t} = \mathbf{z}_t^\top \boldsymbol{\gamma}$. Sendo $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ o vetor de parâmetros do

modelo para média e $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^\top$ o vetor de parâmetros do modelo para dispersão, com $\boldsymbol{\beta} \in \mathbb{R}^p$ e $\boldsymbol{\gamma} \in \mathbb{R}^q$, e com $\mathbf{x}_t = (x_{t1}, \dots, x_{tp})^\top$ e $\mathbf{z}_t = (z_{t1}, \dots, z_{tq})^\top$ sendo valores conhecidos e fixados das variáveis explicativas de ambos os modelos, podendo coincidir totalmente ou parcialmente, de modo que $p + q < n$. Podemos notar, que $h(\cdot)$ será a função de ligação da dispersão, e as funções mais comuns nesse caso são:

- logarítmica: $h(\sigma_t^2) = \log(\sigma_t^2)$;
- raiz quadrada: $h(\sigma_t^2) = \sqrt{\sigma_t^2} = \sigma_t$;
- identidade: $h(\sigma_t^2) = \sigma_t^2$.

Para mais detalhes sobre as funções de ligações, ver McCullagh e Nelder (1989).

2.4.1 Função Escore, Matriz Hessiana e Matriz de Informação de Fisher para os Parâmetros do Modelo

Baseado em (2.1), temos que o logaritmo da função de verossimilhança é

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{t=1}^n \ell_t(\mu_t, \sigma_t^2), \quad (2.29)$$

em que

$$\ell_t(\mu_t, \sigma_t^2) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^2) - \frac{3}{2} \log[y_t(1 - y_t)] - \frac{1}{2\sigma_t^2} d(y_t; \mu_t).$$

A função escore do vetor $\boldsymbol{\beta}$ é obtido a partir de

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i} &= \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \beta_i} \\ &= \sum_{t=1}^n \left\{ \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right\} \\ &= \sum_{t=1}^n \left\{ \frac{1}{\sigma_t^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti} \right\}. \end{aligned}$$

Segue que o vetor escore de $\boldsymbol{\beta}$ é dado por

$$\mathbf{U}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{X}^\top \boldsymbol{\Sigma} \mathbf{T} \mathbf{U}(\mathbf{y} - \boldsymbol{\mu}),$$

em que

$$\boldsymbol{\Sigma} = \text{diag} \{1/\sigma_1^2, \dots, 1/\sigma_n^2\}. \quad (2.30)$$

Vamos obter, agora, a função escore para $\boldsymbol{\gamma}$. Assim, derivando inicialmente, o logaritmo da verossimilhança em relação a γ_j , com $j = 1, \dots, q$, temos que

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \gamma_j} &= \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \gamma_j} \\ &= \sum_{t=1}^n \left\{ \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \sigma_t^2} \frac{d\sigma_t^2}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \right\} \\ &= \sum_{t=1}^n \left\{ -\frac{1}{2\sigma_t^2} \left[1 - \frac{d(y_t; \mu_t)}{\sigma_t^2} \right] \frac{1}{h'(\sigma_t^2)} z_{tj} \right\}, \end{aligned}$$

em que $d\sigma_t^2/d\eta_{2t} = 1/h'(\sigma_t^2)$ e $\partial d\eta_{2t}/\partial \gamma_j = z_{tj}$. Segue que a função escore para $\boldsymbol{\gamma}$ é dado por

$$\mathbf{U}_{\boldsymbol{\gamma}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{Z}^{\top} \mathbf{H} \mathbf{a},$$

com $\mathbf{Z}$ sendo uma matriz de posto completo de ordem $n \times q$, com a i -ésima linha denotada por $\mathbf{z}_i^{\top}$ e em que

$$\mathbf{H} = \text{diag} \{1/h'(\sigma_1^2), \dots, 1/h'(\sigma_n^2)\}$$

e $\mathbf{a} = (a_1, \dots, a_n)^{\top}$, com

$$a_t = -\frac{1}{2\sigma_t^2} \left[1 - \frac{d(y_t; \mu_t)}{\sigma_t^2} \right]. \quad (2.31)$$

Agora, vamos obter matriz Hessiana. Nesse caso, a matriz é dada por

$$\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) = \begin{bmatrix} \ddot{\ell}_{\beta\beta}(\boldsymbol{\theta}) & \ddot{\ell}_{\beta\gamma}(\boldsymbol{\theta}) \\ \ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta}) & \ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta}) \end{bmatrix} \quad (2.32)$$

em que $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\top}, \boldsymbol{\gamma}^{\top})^{\top}$. Assim, da matriz (2.32) observamos que o (i, j) -ésimo elemento da matriz $\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})$, com $i = 1, \dots, p$ e $j = 1, \dots, p$ é dado por

$$\begin{aligned} [\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial \beta_i \partial \beta_j} \\ &= \sum_{t=1}^n \frac{1}{\sigma_t^2} x_{ti} x_{tj} q_t. \end{aligned} \quad (2.33)$$

Dessa forma, segue matricialmente que

$$\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta}) = \mathbf{X}^{\top} \boldsymbol{\Sigma} \mathbf{Q} \mathbf{X}. \quad (2.34)$$

Considerando, agora, a matriz

$$\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \gamma \partial \beta^\top}, \quad (2.35)$$

cujo (i, j) -ésimo elemento com $i = 1, \dots, p$ e $j = 1, \dots, q$ é definido por

$$\begin{aligned} [\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]_{ij} &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \gamma_j \partial \beta_i} \\ &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_j} \left[\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \beta_i} \right] \\ &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_j} \left[\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right] \\ &= \sum_{t=1}^n \left[\frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t \partial \sigma_t^2} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{d\sigma_t^2}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \right] \\ &= \sum_{t=1}^n -\frac{1}{(\sigma_t^2)^2} u_t(y_t - \mu_t) \frac{x_{ti}}{g'(\mu_t)} \frac{z_{tj}}{h'(\sigma_t^2)}. \end{aligned} \quad (2.36)$$

Então, podemos definir a matriz (2.35) da seguinte forma

$$\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta}) = Z^\top T F H X, \quad (2.37)$$

em que $F = \text{diag}\{f_1^*, \dots, f_n^*\}$, com

$$f_t^* = -\frac{1}{(\sigma_t^2)^2} u_t(y_t - \mu_t). \quad (2.38)$$

Ainda, temos que $\ddot{\ell}_{\beta\gamma}(\boldsymbol{\theta}) = \ddot{\ell}_{\gamma\beta}^\top(\boldsymbol{\theta}) = X^\top T F H Z$. Finalizando, observamos que o (i, j) -ésimo elemento da matriz $\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})$, com $i = 1, \dots, q$ e $j = 1, \dots, q$ é dado por

$$\begin{aligned} [\ddot{\ell}_{\gamma\gamma}]_{ij}(\boldsymbol{\theta}) &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial \gamma_i \partial \gamma_j} \\ &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_j} \left[\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \sigma_t^2} \frac{d\sigma_t^2}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i} \right] \\ &= \sum_{t=1}^n \frac{\partial}{\partial \sigma_t^2} \left[\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \sigma_t^2} \frac{d\sigma_t^2}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i} \right] \frac{d\sigma_t^2}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \\ &= \sum_{t=1}^n \frac{\partial}{\partial \sigma_t^2} \left[\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \sigma_t^2} \frac{1}{h'(\sigma_t^2)} \right] \frac{1}{h'(\sigma_t^2)} z_{ti} z_{tj} \\ &= \sum_{t=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial (\sigma_t^2)^2} \frac{1}{h'(\sigma_t^2)} + \frac{\partial}{\partial \sigma_t^2} \left[\frac{1}{h'(\sigma_t^2)} \right] \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \sigma_t^2} \right\} \frac{1}{h'(\sigma_t^2)} z_{ti} z_{tj} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left\{ \frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial(\sigma_t^2)^2} \frac{1}{[h'(\sigma_t^2)]^2} - \left[-\frac{1}{2\sigma_t^2} \left(1 - \frac{d(y_t; \mu_t)}{\sigma_t^2} \right) \right] \frac{h''(\sigma_t^2)}{[h'(\sigma_t^2)]^3} \right\} z_{tj} z_{ti} \\
&= \sum_{i=1}^n \left\{ \frac{1}{2(\sigma_t^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma_t^2)^3} + \left[\frac{1}{2\sigma_t^2} - \frac{d(y_t; \mu_t)}{2(\sigma_t^2)^2} \right] \frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right\} \frac{z_{tj} z_{ti}}{[h'(\sigma_t^2)]^2}, \quad (2.39)
\end{aligned}$$

denotando

$$d_t^* = \left\{ \frac{1}{2(\sigma_t^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma_t^2)^3} + \left[\frac{1}{2\sigma_t^2} - \frac{d(y_t; \mu_t)}{2(\sigma_t^2)^2} \right] \frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right\} \frac{1}{[h'(\sigma_t^2)]^2}, \quad (2.40)$$

temos que

$$[\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]_{ij} = \sum_{i=1}^n d_t^* z_{tj} z_{ti}. \quad (2.41)$$

Logo, matricialmente,

$$\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta}) = Z^\top D^* Z, \quad (2.42)$$

em que $D^* = \text{diag}\{d_1^*, \dots, d_n^*\}$. Agora, pela expressão (2.33), temos que

$$-\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_i} \right] = \sum_{t=1}^n \left\{ \frac{1}{\sigma_t^2} \left[\frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3} \right] \frac{1}{[g'(\mu_t)]^2} x_{ti} x_{tj} \right\}$$

Assim, a matriz de informação de Fisher de $\boldsymbol{\beta}$, para o caso com dispersão variável fica dada por

$$\begin{aligned}
K_{\beta\beta}^* &= -\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right] \\
&= X^\top \Sigma W^* X,
\end{aligned}$$

em que $W^* = \text{diag}\{w_1^*, \dots, w_n^*\}$, com

$$w_t^* = \left[\frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3} \right] \frac{1}{[g'(\mu_t)]^2}.$$

Para obter a matriz de informação de Fisher para $\boldsymbol{\gamma}$, iremos utilizar a expressão (2.39), assim

$$-\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_i} \right] = \sum_{t=1}^n \left[\frac{1}{2(\sigma_t^2)^4} \frac{1}{[h'(\sigma_t^2)]^2} z_{ti} z_{tj} \right].$$

Então, a matriz de informação de Fisher para γ fica dada por

$$\begin{aligned} K_{\gamma\gamma}^* &= -\mathbb{E} \left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \gamma \partial \gamma^\top} \right] \\ &= Z^\top V Z, \end{aligned}$$

em que $V = \text{diag}\{v_1, \dots, v_n\}$, com

$$v_t = \frac{1}{2(\sigma_t^2)^2} \frac{1}{h'(\sigma_t^2)^2}.$$

Finalizando, tomando a esperança de (2.36), vemos que

$$\mathbb{E} \left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \gamma_j \partial \beta_i} \right) = 0,$$

ou seja, $K_{\gamma\beta}^* = K_{\beta\gamma}^{*\top} = 0$. Consequentemente, temos que os parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são ortogonais.

Portanto, a matriz de informação de Fisher para $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ é bloco diagonal, dada por $K_{\theta\theta}^* = \text{diag}\{K_{\beta\beta}^*, K_{\gamma\gamma}^*\}$, ou seja,

$$K_{\theta\theta}^* = \begin{bmatrix} K_{\beta\beta}^* & 0 \\ 0 & K_{\gamma\gamma}^* \end{bmatrix}. \quad (2.43)$$

Para n suficientemente grande e considerando válidas as condições de regularidades, temos que a distribuição assintótica dos estimadores de máxima verossimilhança é dada por

$$\hat{\boldsymbol{\theta}} \sim N_{p+q}(\boldsymbol{\theta}; K_{\theta\theta}^{*-1}),$$

em que $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\gamma}}^\top)^\top$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ e $K_{\theta\theta}^{*-1}$ é o inverso da matriz de informação de Fisher, dada por (2.43).

2.4.2 Estimação dos Parâmetros

Nesse caso, também, não temos forma fechada para os estimadores de máxima verossimilhança dos parâmetros do modelo. Utilizando, novamente, o algoritmo Scoring de Fisher obtemos o seguinte processo iterativo

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + [K_{\beta\beta}^{*-1}]^{(m)} \mathbf{U}_{\beta}^{(m)}$$

e

$$\boldsymbol{\gamma}^{(m+1)} = \boldsymbol{\gamma}^{(m)} + [K_{\gamma\gamma}^{*-1}]^{(m)} \mathbf{U}_{\gamma}^{(m)}.$$

Reescrevendo o processo iterativo por mínimos quadrados ponderados, temos que

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

e

$$\boldsymbol{\gamma}^{(m+1)} = (Z^\top V^{(m)} Z)^{-1} Z^\top V^{(m)} \mathbf{z}_2^*,$$

em que $\mathbf{z}_1^{*(m)} = X\boldsymbol{\beta}^{(m)} + [W^{*-1}]^{(m)} T^{(m)} U^{(m)} (\mathbf{y} - \boldsymbol{\mu}^{(m)})$ e $\mathbf{z}_2^{*(m)} = Z\boldsymbol{\gamma}^{(m)} + [V^{-1}]^{(m)} H^{(m)} \mathbf{a}^{(m)}$ são soluções de mínimos quadrados de regressões lineares com variáveis respostas $\mathbf{z}_1^*$ e $\mathbf{z}_2^*$ e matrizes de pesos ΣW^* e V , respectivamente.

2.4.3 Análise de Influência Local para o Modelo com Dispersão Variável

Considerando o modelo de regressão definido por (2.27) e por (2.28), em que $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$, temos que a matriz $\Delta(\boldsymbol{\theta})$ nesse caso será uma matriz de ordem $(p+q) \times n$. Assim como no caso com dispersão fixa, vamos trabalhar com o tipo de perturbação ponderação de casos. Considerando que $\ell(\boldsymbol{\theta})$ seja o logaritmo da função de verossimilhança correspondente ao modelo com dispersão variável. Supondo que esse modelo seja perturbado por um vetor $\boldsymbol{\delta}$ de ordem $n \times 1$, o logaritmo da função de verossimilhança do modelo perturbado será

$$\ell_\delta(\boldsymbol{\theta}) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \sigma_t^2),$$

em que $0 \leq \delta_t \leq 1$. Assim, segue que

$$\frac{\partial^2 \ell_\delta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \delta_t} = \Delta_t(\boldsymbol{\theta}).$$

Portanto, a matriz $\Delta(\boldsymbol{\theta})$ fica dada por

$$\Delta(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial \ell_1(\mu_1, \sigma_1^2)}{\partial \beta_1} & \dots & \frac{\partial \ell_n(\mu_n, \sigma_n^2)}{\partial \beta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \ell_1(\mu_1, \sigma_1^2)}{\partial \beta_p} & \dots & \frac{\partial \ell_n(\mu_n, \sigma_n^2)}{\partial \beta_p} \\ \frac{\partial \ell_1(\mu_1, \sigma_1^2)}{\partial \gamma_1} & \dots & \frac{\partial \ell_n(\mu_n, \sigma_n^2)}{\partial \gamma_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \ell_1(\mu_1, \sigma_1^2)}{\partial \gamma_q} & \dots & \frac{\partial \ell_n(\mu_n, \sigma_n^2)}{\partial \gamma_q} \end{bmatrix},$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Para o modelo de regressão simplex com dispersão variável verificamos que

$$\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \beta_i} = \frac{1}{\sigma_t^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti},$$

e

$$\begin{aligned} \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \gamma_j} &= -\frac{1}{2\sigma_t^2} \left[1 - \frac{d(y_t; \mu_t)}{\sigma_t^2} \right] \frac{1}{h'(\sigma_t^2)} z_{tj} \\ &= a_t \frac{1}{h'(\sigma_t^2)} z_{tj} \end{aligned}$$

com, $i = 1, \dots, p$, $j = 1, \dots, q$ e $t = 1, \dots, n$. Logo, teremos que

$$\Delta_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = X^\top T \Sigma U \mathcal{E},$$

e que

$$\Delta_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = Z^\top H \mathcal{A},$$

em que $\mathcal{A} = \text{diag}\{a_1, \dots, a_n\}$, com a_t definido por (2.31), e com ambas matrizes avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Então,

$$\Delta(\boldsymbol{\theta}) = \begin{bmatrix} X^\top T \Sigma U \mathcal{E} \\ Z^\top H \mathcal{A} \end{bmatrix}. \quad (2.44)$$

Capítulo 3

Um Teste de Diagnóstico Para o Modelo Simplex

Ao ajustarmos um modelo de regressão a um conjunto de dados é necessário avaliar a qualidade do ajuste verificando possíveis afastamentos das suposições feitas para o modelo. Para tanto, são utilizados várias técnicas de diagnóstico propostas na literatura, tais como: análise de resíduos, que tem como objetivo detectar observações extremas, bem como, validar a adequação da distribuição de probabilidade assumida; análise de diagnóstico que tem como finalidade detectar observações influentes, ou seja, aquelas que exercem um efeito desproporcional nas estimativas dos parâmetros do modelo.

Como já foi mencionado, o método de influência local é um dos procedimentos de diagnósticos, utilizados com essa finalidade. Dessa forma, baseando-se nesse método, Zhu e Zhang (2004) desenvolveram um teste para avaliar a adequabilidade do ajuste de um modelo de regressão. Neste capítulo, apresentaremos o desenvolvimento desse teste para o modelo de regressão simplex.

3.1 Teste de Diagnóstico Baseado em Influência Local

Nesta seção, iremos descrever o teste de hipóteses baseado em influência local, proposto por Zhu e Zhang (2004).

Supondo $y_1, \dots, y_n$ uma realização de n variáveis aleatórias independentes e identicamente distribuídas, cuja função densidade de probabilidade é $g(y)$, vamos supor que os dados vêm de uma família de distribuições

$$\mathcal{F}_{\boldsymbol{\theta}} = \{p(y; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k\},$$

que não necessariamente contém $g(y)$, porém satisfaz algumas condições impostas por White (1982). Então, quando $n \rightarrow \infty$, temos que $\hat{\boldsymbol{\theta}}$ será um estimador consistente para

um certo parâmetro $\boldsymbol{\theta}_*$, de modo que

$$\boldsymbol{\theta}_* = \arg \min_{\boldsymbol{\theta} \in \Theta} \mathbb{E} \left[\log \frac{g(y)}{p(y; \boldsymbol{\theta})} \right]. \quad (3.1)$$

Da expressão (3.1), temos que $\boldsymbol{\theta}_*$ é uma parâmetro que reflete a distância mínima entre $\log[g(y)]$ e $\log[p(y; \boldsymbol{\theta})]$. Adicionalmente, segue que

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) \xrightarrow{D} N(0, C(\boldsymbol{\theta}_*)),$$

em que $C(\boldsymbol{\theta}_*) = A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)A(\boldsymbol{\theta}_*)^{-1}$, com

$$\begin{aligned} A(\boldsymbol{\theta}) &= -\mathbb{E}[\ddot{\ell}_{\theta\theta}(y; \boldsymbol{\theta})] \\ &= -\mathbb{E} \left[\frac{\partial^2 \log p(y; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right] \end{aligned}$$

e

$$\begin{aligned} B(\boldsymbol{\theta}) &= \mathbb{E}[\dot{\ell}_{\theta}(y; \boldsymbol{\theta})\dot{\ell}_{\theta}^\top(y; \boldsymbol{\theta})] \\ &= \mathbb{E} \left[\left(\frac{\partial \log p(y; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log p(y; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top \right]. \end{aligned}$$

Notemos que $\ell(y; \boldsymbol{\theta}) = \log p(y; \boldsymbol{\theta})$ e que $B(\boldsymbol{\theta})$ é a matriz de informação de Fisher de $\boldsymbol{\theta}$, e para os casos em que as condições de regularidades são satisfeitas temos que $A(\boldsymbol{\theta}) = B(\boldsymbol{\theta})$. Dessa forma, seja $y_1, \dots, y_n$ uma amostra aleatória que assumimos ser de $p(y; \boldsymbol{\theta})$, o logaritmo da função de verossimilhança é dado por

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^n \log p(y_t; \boldsymbol{\theta}) = \sum_{t=1}^n \ell_t(\boldsymbol{\theta}).$$

Considerando o esquema de perturbação ponderação de casos, cujo logaritmo da função de verossimilhança perturbado é dado por

$$\ell_\delta(\boldsymbol{\theta}) = \sum_{t=1}^n \delta_t \ell_t(\boldsymbol{\theta}),$$

com $\boldsymbol{\delta}_0 = (1, 1, \dots, 1)^\top$, podemos observar, nesse caso, que a (i, j) -ésima entrada da matriz hessiana com $i = 1, \dots, p$ e $j = 1, \dots, p$, em que p é a dimensão de $\boldsymbol{\theta}$, será dada por

$$\begin{aligned} [\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]_{ij} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}. \end{aligned}$$

Logo, a matriz de informação observada de $\boldsymbol{\theta}$ é

$$-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) = -\sum_{t=1}^n \frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}.$$

Sabemos que $\Delta(\boldsymbol{\theta}) = [\Delta_1(\boldsymbol{\theta}), \dots, \Delta_n(\boldsymbol{\theta})]$, logo $\Delta \Delta^\top(\boldsymbol{\theta}) = \sum_{t=1}^n \Delta_t \Delta_t^\top(\boldsymbol{\theta})$. Então, baseando-se em (1.10) temos que a (r, s) -ésima entrada da matriz $\Delta(\boldsymbol{\theta})$ com $r = 1, \dots, p$ e $s = 1, \dots, n$ é dada por

$$\begin{aligned} [\Delta(\boldsymbol{\theta})]_{rs} &= \frac{\partial^2 \ell_s(\boldsymbol{\theta})}{\partial \theta_r \partial \delta_s} \\ &= \frac{\partial^2}{\partial \theta_r \partial \delta_s} \left[\sum_{t=1}^n \delta_t \ell_t(\boldsymbol{\theta}) \right] \\ &= \frac{\partial}{\partial \theta_r} \left[\frac{\partial \sum_{t=1}^n \delta_t \ell_t(\boldsymbol{\theta})}{\partial \delta_s} \right] \\ &= \frac{\partial \ell_s(\boldsymbol{\theta})}{\partial \theta_r}, \end{aligned}$$

segue que, a s -ésima coluna da matriz $\Delta(\boldsymbol{\theta})$ é definida por

$$\Delta_s(\boldsymbol{\theta}) = \frac{\partial \ell_s(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

Utilizando a lei dos grandes números, é possível mostrar que

$$-\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \xrightarrow{p} A(\boldsymbol{\theta})$$

e

$$\frac{\sum_{t=1}^n \Delta_t \Delta_t^\top(\boldsymbol{\theta})}{n} = \frac{1}{n} \sum_{t=1}^n \left\{ \left[\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \left[\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^\top \right\} \xrightarrow{p} B(\boldsymbol{\theta}).$$

Entretanto, de (3.1) temos que $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_*$, logo,

$$\frac{-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} \xrightarrow{p} A(\boldsymbol{\theta}_*) \quad (3.2)$$

e

$$\frac{\Delta \Delta^\top(\hat{\boldsymbol{\theta}})}{n} \xrightarrow{p} B(\boldsymbol{\theta}_*). \quad (3.3)$$

Seguiremos agora, fazendo algumas suposições, que serão utilizadas na definição do Teorema 1, dado a seguir. A prova desse teorema pode ser encontrada no apêndice desse trabalho e, também, em Zhu e Zhang (2004).

As suposições são as seguintes:

S1: $\boldsymbol{\theta}_*$ é um ponto contido em Θ , em que Θ é um espaço limitado com métrica d .

S2: $P(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*| > \epsilon) \xrightarrow{n \rightarrow \infty} 0$, $\forall \epsilon > 0$, ou seja, $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}_*$ quando $n \rightarrow \infty$, em que $\boldsymbol{\theta}_*$ é o único valor que minimiza $\mathbb{E}[\sum_{t=1}^n \ell_t(\boldsymbol{\theta})]$.

S3: para qualquer $p(y_t; \boldsymbol{\theta})$, com $t = 1, \dots, n$, o suporte é independente do parâmetro $\boldsymbol{\theta}$ e permanece o mesmo. $\ell_t(\boldsymbol{\theta})$ é três vezes continuamente diferenciável em Θ , e $|\partial \ell_t(\boldsymbol{\theta}) / \partial \theta_j|$, $|\partial^2 \ell_t(\boldsymbol{\theta}) / \partial \theta_j \partial \theta_k|$, $|\partial^3 \ell_t(\boldsymbol{\theta}) / \partial \theta_j \partial \theta_k \partial \theta_l|$ e $|\partial[\ell_t(\boldsymbol{\theta}) \partial \ell_t(\boldsymbol{\theta}) / \partial \theta_k] / \partial \theta_j|$ são dominados por $B_t(y_t)$, $\forall t$, com $j, k, l = 1, \dots, p$.

S4: Para cada $\varepsilon > 0$, existe um K finito, tal que

$$\sup_{n \geq 1} n^{-1} \sum_{t=1}^n \mathbb{E}\{B_t(y_t)^2 I_{[B_t(y_t) > K]}\} < \varepsilon,$$

para todo n , em que $I_{[B_t(y_t) > K]}$ é a função indicadora de $B_t(y_t) > K$.

S5: Assumimos que

$$\lim_{n \rightarrow \infty} \left\{ -\frac{\mathbb{E}[\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right\} = A(\boldsymbol{\theta}_*)$$

e

$$\lim_{n \rightarrow \infty} \left\{ \frac{\mathbb{E}[\Delta \Delta^\top(\boldsymbol{\theta}_*)]}{n} \right\} = B(\boldsymbol{\theta}_*),$$

em que $A(\boldsymbol{\theta}_*)$ é positiva definida.

TEOREMA 1. Sob essas suposições, Zhu e Zhang (2004) mostraram que são válidos os seguintes resultados:

1. $\lambda_0 = 2\text{tr}[A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)]$;
2. Se $g(y) = p(y; \boldsymbol{\theta}_0)$, com $\boldsymbol{\theta}_0 \in \Theta$, em que $g(y)$ é a densidade verdadeira de $y_1, \dots, y_n$, então $\boldsymbol{\theta}_* = \boldsymbol{\theta}_0$, $A(\boldsymbol{\theta}_*) = B(\boldsymbol{\theta}_*)$ e $\lambda_0 = 2p$.

Os autores concluem que se o modelo estiver ajustado adequadamente, então da segunda parte do teorema temos que $\lambda_0 = 2p$, em que p é a dimensão de $\boldsymbol{\theta}$. Dado que o valor de $2p$ é conhecido, pela Proposição 1.1, podemos usar a soma dos C_{I_t} 's como um critério de bondade de ajuste, ou seja, se $\sum_{t=1}^n C_{I_t}$ estiver próximo de $2p$, existem fortes indícios de que $p(y; \boldsymbol{\theta})$ é um bom modelo para o verdadeiro gerador dos dados.

Assim, tendo como objetivo testar as hipóteses

$$\begin{aligned} H_0 : g(y) &= p(y; \boldsymbol{\theta}), \text{ para algum } \boldsymbol{\theta} \in \Theta \\ H_1 : g(y) &\neq p(y; \boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta, \end{aligned}$$

foi introduzida a estatística de teste

$$T_n(\hat{\boldsymbol{\theta}}) = \frac{\sum_{t=1}^n (nC_{I_t} - \lambda_0)}{\sqrt{n}}, \quad (3.4)$$

lembrando que C_{I_t} é função de $\hat{\boldsymbol{\theta}}$. Com o intuito de determinar a distribuição assintótica de $T_n(\hat{\boldsymbol{\theta}})$, foram consideradas as seguintes quantidades

$$J_1(\boldsymbol{\theta}) = 2 \frac{\partial \text{tr}[A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}, \quad (3.5)$$

e

$$k(y_t; \boldsymbol{\theta}_*) = h(y_t; \boldsymbol{\theta}_*) + J_1(\boldsymbol{\theta}_*)^\top A(\boldsymbol{\theta}_*)^{-1} \left[\frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} \right], \quad (3.6)$$

em que $k(y_t; \boldsymbol{\theta}_*)$ é definido a partir de uma aproximação de Taylor de $J_1(\boldsymbol{\theta})$, com

$$\begin{aligned} h(y_t; \boldsymbol{\theta}_*) &= 2 \text{tr}\{A(\boldsymbol{\theta}_*)^{-1}[\Delta_t \Delta_t^\top(\boldsymbol{\theta}_*) - B(\boldsymbol{\theta}_*)]\} \\ &- \text{tr}\{A(\boldsymbol{\theta}_*)^{-1} \left[-\frac{\partial^2 \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} - A(\boldsymbol{\theta}_*) \right] A(\boldsymbol{\theta}_*)^{-1} B(\boldsymbol{\theta}_*)\}, \end{aligned} \quad (3.7)$$

para $t = 1, \dots, n$. Utilizando (3.5) e (3.6), foi definido o Teorema 2, que será enunciado mais adiante, porém, antes de apresentá-lo é necessário assumir três novas suposições:

S6: $\forall t$, $\ell_t(\boldsymbol{\theta})$ é quatro vezes continuamente diferenciável em Θ , e $|\partial^4 \ell_t(\boldsymbol{\theta}) / \partial \theta_j \partial \theta_k \partial \theta_l \partial \theta_i|$ são dominados por $B_t(y_t)$, com $j, k, l, i = 1, \dots, p$.

S7: Quando $n \rightarrow \infty$, então

$$\frac{\partial}{\partial \boldsymbol{\theta}} \left[\frac{\text{SC}_n(\boldsymbol{\theta})}{n} \right] \bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*} \xrightarrow{p} J_1(\boldsymbol{\theta}_*),$$

em que $\text{SC}_n(\boldsymbol{\theta}) = \sum_{t=1}^n nC_{I_t}$.

S8: Para cada $K > 0$,

$$\frac{\sum_{t=1}^n \mathbb{E} \left\{ k^2(y_t; \boldsymbol{\theta}_*) I_{[k(y_t; \boldsymbol{\theta}_*) > \sqrt{n}K]} \right\}}{n} \rightarrow 0,$$

quando $n \rightarrow \infty$, em que $I_{[k(y_t; \boldsymbol{\theta}_*) > \sqrt{n}K]}$ é a função indicadora de $k(y_t; \boldsymbol{\theta}_*) > \sqrt{n}K$. Além disso,

$$\frac{\sum_{t=1}^n \mathbb{E} \{ k^2(y_t; \boldsymbol{\theta}_*) \}}{n} \xrightarrow{p} \sigma_*^2.$$

TEOREMA 2. Sob as suposições S1-S8 é válido que, quando $n \rightarrow \infty$,

$$n^{-1/2} \sum_{t=1}^n nC_{I_t} = n^{-1/2} \lambda_0 + n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) + o_p(1), \quad (3.8)$$

em que $o_p(1)$ indica que os próximos termos convergem em probabilidade para zero. Além disso, também é válido que $T_n(\hat{\boldsymbol{\theta}}) \xrightarrow{D} N(0, \sigma_*^2)$, em que $\xrightarrow{D}$, denota convergência em distribuição. Para mais detalhes sobre esse tipo de convergência ver James (1981) ou Magalhães (2006).

A prova desse teorema é apresentada no apêndice desse trabalho e também em Zhu e Zhang (2004). Segue do Teorema 2 que a variância de $T_n(\hat{\boldsymbol{\theta}})$, que também é a variância de $k(y; \boldsymbol{\theta}_*)$ é dada por σ_*^2 . Após estimar $\boldsymbol{\theta}_*$, $A(\boldsymbol{\theta}_*)$, $B(\boldsymbol{\theta}_*)$ e $J_1(\boldsymbol{\theta}_*)$ podemos obter um estimador consistente para $k(y; \boldsymbol{\theta}_*)$, denotado por $k(y, \hat{\boldsymbol{\theta}})$, cuja variância amostral desse estimador, representada por $s_{k,c}^2$, será um estimador consistente de σ_*^2 .

Também, segue do Teorema 2 que

$$T_n(\hat{\boldsymbol{\theta}}) = n^{-1/2} \sum_{t=1}^n (nC_{I_t} - \lambda_0) \xrightarrow{D} N(0, \sigma_*^2),$$

então,

$$\begin{aligned} \frac{T_n(\hat{\boldsymbol{\theta}})}{\sigma_*} &= \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - \lambda_0)}{\sigma_*} \\ &= \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - \lambda_0 + 2p - 2p)}{\sigma_*} \\ &= \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - 2p) + n^{-1/2} \sum_{t=1}^n (2p - \lambda_0)}{\sigma_*} \\ &= \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - 2p) + n^{1/2} (2p - \lambda_0)}{\sigma_*} \xrightarrow{D} N(0, 1). \end{aligned}$$

Definindo,

$$U_n = \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - 2p)}{\sigma_*},$$

observamos que,

- se $\lambda_0 = 2p$, então

$$\frac{T_n(\hat{\boldsymbol{\theta}})}{\sigma_*} = U_n \xrightarrow{D} N(0, 1);$$

- se $\lambda_0 \neq 2p$, então, $\lambda_0 - 2p = c_0 \neq 0$, em que c_0 é um escalar. Assim,

$$\frac{T_n(\hat{\boldsymbol{\theta}})}{\sigma_*} = U_n - \frac{n^{1/2} c_0}{\sigma_*} \xrightarrow{D} N(0, 1),$$

logo,

$$U_n = \frac{T_n(\hat{\boldsymbol{\theta}})}{\sigma_*} + \frac{n^{1/2}c_0}{\sigma_*} \xrightarrow{D} N\left(\frac{n^{1/2}c_0}{\sigma_*}, 1\right).$$

Uma segunda alternativa para esse item seria $\lambda_0 - 2p = n^{-1/2}c_0$, assim

$$U_n = \frac{T_n(\hat{\boldsymbol{\theta}})}{\sigma_*} + \frac{c_0}{\sigma_*} \xrightarrow{D} N\left(\frac{c_0}{\sigma_*}, 1\right).$$

Portanto, tendo como objetivo testar as hipóteses

$$H_0 : \lambda_0 = 2p,$$

$$H_1 : \lambda_0 \neq 2p,$$

temos que a estatística

$$U_n = \frac{n^{-1/2} \sum_{t=1}^n (nC_{I_t} - 2p)}{s_{k,c}}, \quad (3.9)$$

é uma ferramenta bastante útil para realizar esse teste. Observamos que U_n corresponde a estatística $T_n(\hat{\boldsymbol{\theta}})/s_{k,c}$ sob H_0 . Dessa forma, rejeitar a hipótese que $\lambda_0 = 2p$, indica que $p(y; \boldsymbol{\theta})$ está mal especificado, por outro lado, não rejeitar essa hipótese indica que $p(y; \boldsymbol{\theta})$ está bem ajustado.

Zhu e Zhang (2004) afirmam que embora a estatística U_n tenha sido desenvolvida sob a mesma ideia que a estatística introduzida por White (1982), torna-se mais útil utilizar U_n , pelo fato de ser mais fácil adaptá-la para casos em que $p > 1$ e para o ajuste de diversos tipos de modelos, como por exemplo, os modelos lineares generalizados e aos modelos de dispersão.

3.2 O Teste para o Modelo Simplex Com Dispersão Fixa

No segundo capítulo foi apresentado o modelo de regressão simplex, no caso em que o parâmetro de dispersão é constante, em que o logaritmo da verossimilhança denotado por $\ell(\boldsymbol{\theta})$ foi definido por (2.3).

A partir de (2.11), (2.15) e (2.16), segue que a matriz de informação observada fica dada por

$$-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) = \begin{bmatrix} -\frac{1}{\sigma^2} X^\top Q X & -X^\top T \mathbf{f} \\ -\mathbf{f}^\top T X & -\text{tr}(D) \end{bmatrix}, \quad (3.10)$$

e a matriz $\Delta(\theta)$ é dada por (2.26).

Através das expressões (3.2) e (3.3), podemos verificar que os estimadores consistentes para as matrizes $A(\theta_*)$ e $B(\theta_*)$ são definidos, respectivamente, por

$$A(\hat{\theta}) = \frac{-\ddot{\ell}_{\theta\theta}(\hat{\theta})}{n}$$

e

$$B(\hat{\theta}) = \frac{\Delta\Delta^\top(\hat{\theta})}{n},$$

em que θ_* é definido por (3.1), de forma de $p(y; \theta)$ corresponde a densidade da distribuição simplex e $\hat{\theta}$ é o estimador de máxima verossimilhança de θ_* . Pela suposição S7, podemos ver que um estimador consistente para a matriz $J_1(\theta_*)$ é dado por

$$\begin{aligned} J_1(\hat{\theta}) &= \left. \frac{\partial}{\partial \theta} \left[\frac{SC_n(\theta)}{n} \right] \right|_{\theta=\hat{\theta}} \\ &= \left. \frac{\partial}{\partial \theta} \left[\frac{\sum_{t=1}^n nC_{I_t}}{n} \right] \right|_{\theta=\hat{\theta}} \\ &= \left. \frac{\partial}{\partial \theta} \left[\sum_{t=1}^n C_{I_t} \right] \right|_{\theta=\hat{\theta}} \\ &= \frac{\partial \{\text{tr}[H_{LD(\delta_0)}]\}}{\partial \theta} \\ &= \text{tr} \left\{ \frac{\partial [H_{LD(\delta_0)}]}{\partial \theta} \right\}. \end{aligned} \quad (3.11)$$

Segue de (1.8), que

$$H_{LD(\delta_0)} = 2\Delta^\top(\theta)[- \ddot{\ell}_{\theta\theta}(\theta)]^{-1}\Delta(\theta) \Big|_{\delta=\delta_0, \theta=\hat{\theta}},$$

logo,

$$\begin{aligned} \frac{\partial [H_{LD(\delta_0)}]}{\partial \theta} &= 2 \left\{ \frac{\partial \Delta^\top(\theta)}{\partial \theta} [- \ddot{\ell}_{\theta\theta}(\theta)]^{-1} \Delta(\theta) + \Delta^\top(\theta) \frac{\partial [- \ddot{\ell}_{\theta\theta}(\theta)]^{-1}}{\partial \theta} \Delta(\theta) \right. \\ &\quad \left. + \Delta^\top(\theta) [- \ddot{\ell}_{\theta\theta}(\theta)]^{-1} \frac{\partial \Delta(\theta)}{\partial \theta} \right\}. \end{aligned} \quad (3.12)$$

Podemos observar que

$$\frac{\partial [- \ddot{\ell}_{\theta\theta}(\theta)]^{-1}}{\partial \theta} = -[- \ddot{\ell}_{\theta\theta}(\theta)]^{-1} \frac{\partial [- \ddot{\ell}_{\theta\theta}(\theta)]}{\partial \theta} [- \ddot{\ell}_{\theta\theta}(\theta)]^{-1}. \quad (3.13)$$

Assim, antes de estimar a matriz $J_1(\boldsymbol{\theta}_*)$, precisamos obter as derivadas das matrizes $-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})$ e $\Delta(\boldsymbol{\theta})$. Dessa forma, vamos começar derivando as entradas de $-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})$ em relação a β_r , com $r = 1, \dots, p$. Da equação (2.10), temos que

$$[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij} = -\frac{1}{\sigma^2} \sum_{t=1}^n x_{ti}x_{tj}q_t, \quad (3.14)$$

então

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij}}{\partial\beta_r} = -\frac{1}{\sigma^2} \sum_{t=1}^n x_{ti}x_{tj} \left(\frac{\partial q_t}{\partial\beta_r} \right). \quad (3.15)$$

Pela definição de q_t , dada por (2.9), observamos que

$$\begin{aligned} \frac{\partial q_t}{\partial\beta_r} &= -\frac{\partial}{\partial\beta_r} \left\{ \frac{u_t}{[g'(\mu_t)]^2} \right\} - \frac{\partial}{\partial\beta_r} \left\{ \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{[g'(\mu_t)]^2 \mu_t^2 (1 - \mu_t)^2} \right\} \\ &\quad - \frac{\partial}{\partial\beta_r} \left\{ \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{[g'(\mu_t)]^2 \mu_t^4 (1 - \mu_t)^4} \right\} + \frac{\partial}{\partial\beta_r} \left\{ \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{[g'(\mu_t)]^2 \mu_t (1 - \mu_t)} \right\} \\ &\quad - \frac{\partial}{\partial\beta_r} \left\{ (y_t - \mu_t)u_t \frac{g''(\mu_t)}{[g'(\mu_t)]^3} \right\}. \end{aligned} \quad (3.16)$$

Sabendo que,

$$\frac{\partial u_t}{\partial\beta_r} = \left[-\frac{(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2 (1 - \mu_t)^2} - \frac{3(1 - 2\mu_t)}{\mu_t^4 (1 - \mu_t)^4} + \frac{d'(y_t; \mu_t)}{\mu_t (1 - \mu_t)} \right] \frac{x_{tr}}{g'(\mu_t)}$$

e

$$\frac{\partial[g'(\mu_t)]^2}{\partial\beta_r} = 2x_{tr}g''(\mu_t),$$

então, o primeiro termo de (3.16) fica dado por

$$\begin{aligned} \frac{\partial}{\partial\beta_r} \left\{ \frac{u_t}{[g'(\mu_t)]^2} \right\} &= \left[-\frac{(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2 (1 - \mu_t)^2} - \frac{3(1 - 2\mu_t)}{\mu_t^4 (1 - \mu_t)^4} \right. \\ &\quad \left. + \frac{d'(y_t; \mu_t)}{\mu_t (1 - \mu_t)} - 2u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right] \frac{x_{tr}}{[g'(\mu_t)]^3}. \end{aligned}$$

Agora, do segundo termo de (3.16), vemos que

$$\begin{aligned} \frac{\partial}{\partial\beta_r} [(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)] &= \frac{x_{tr}}{g'(\mu_t)} [(-2y_t - 1 + 4\mu_t)d(y_t; \mu_t) \\ &\quad + (y_t - 2y_t\mu_t - \mu_t + 2\mu_t^2)d'(y_t; \mu_t)], \end{aligned}$$

e que

$$\frac{\partial}{\partial \beta_r} \{ [g'(\mu_t)]^2 \mu_t^2 (1 - \mu_t)^2 \} = 2x_{tr} \mu_t (1 - \mu_t) [g''(\mu_t) \mu_t (1 - \mu_t) + (1 - 2\mu_t) g'(\mu_t)],$$

logo, esse termo será dado por

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \left\{ \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{[g'(\mu_t)]^2 \mu_t^2 (1 - \mu_t)^2} \right\} &= \frac{x_{tr}}{[g'(\mu_t)]^3} \left[\frac{(-2y_t - 1 + 4\mu_t)d(y_t; \mu_t)}{\mu_t^2 (1 - \mu_t)^2} \right. \\ &+ \frac{(y_t - \mu_t)(1 - 2\mu_t)d'(y_t; \mu_t)}{\mu_t^2 (1 - \mu_t)^2} \\ &- \frac{2(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2 (1 - \mu_t)^2} \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\left. - \frac{2(y_t - \mu_t)(1 - 2\mu_t)^2 d(y_t; \mu_t)}{\mu_t^3 (1 - \mu_t)^3} \right]. \end{aligned}$$

Prosseguindo, vamos obter o terceiro termo de (3.16), em que

$$\frac{\partial}{\partial \beta_r} [(y_t - \mu_t)(1 - 2\mu_t)] = (-2y_t - 1 + 4\mu_t) \frac{x_{tr}}{g'(\mu_t)},$$

e também

$$\frac{\partial}{\partial \beta_r} \{ [g'(\mu_t)]^2 \mu_t^4 (1 - \mu_t)^4 \} = 2x_{tr} \mu_t^3 (1 - \mu_t)^3 [g''(\mu_t) \mu_t (1 - \mu_t) + 2g'(\mu_t)(1 - 2\mu_t)],$$

então,

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \left[\frac{(y_t - \mu_t)(1 - 2\mu_t)}{[g'(\mu_t)]^2 \mu_t^4 (1 - \mu_t)^4} \right] &= \left[(-2y_t - 1 + 4\mu_t) - \frac{4(y_t - \mu_t)(1 - 2\mu_t)^2}{\mu_t (1 - \mu_t)} \right. \\ &\left. - \frac{2(y_t - \mu_t)(1 - 2\mu_t)g''(\mu_t)}{g'(\mu_t)} \right] \frac{x_{tr}}{[g'(\mu_t)]^3 \mu_t^4 (1 - \mu_t)^4}. \end{aligned}$$

Prosseguindo, temos que

$$\frac{\partial}{\partial \beta_r} [(y_t - \mu_t)d'(y_t; \mu_t)] = \frac{x_{tr}}{g'(\mu_t)} [-d'(y_t; \mu_t) + (y_t - \mu_t)d''(y_t; \mu_t)],$$

e

$$\frac{\partial}{\partial \beta_r} \{ [g'(\mu_t)]^2 \mu_t (1 - \mu_t) \} = x_{tr} [2g''(\mu_t) \mu_t (1 - \mu_t) + g'(\mu_t)(1 - 2\mu_t)],$$

assim, o quarto termo de (3.16) será

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \left\{ \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{[g'(\mu_t)]^2 \mu_t (1 - \mu_t)} \right\} &= \frac{x_{tr}}{[g'(\mu_t)]^3 \mu_t (1 - \mu_t)} \left[-d'(y_t; \mu_t) - 2(y_t - \mu_t)d'(y_t; \mu_t) \frac{g''(\mu_t)}{g'(\mu_t)} \right. \\ &+ (y_t - \mu_t)d''(y_t; \mu_t) - \left. \frac{(y_t - \mu_t)(1 - 2\mu_t)d'(y_t; \mu_t)}{\mu_t (1 - \mu_t)} \right]. \end{aligned}$$

Finalmente, o último termo da expressão (3.16) é dado por

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \left\{ \frac{g''(\mu_t)}{[g'(\mu_t)]^3} (y_t - \mu_t) u_t \right\} &= \frac{x_{tr}}{[g'(\mu_t)]^3} \left\{ (y_t - \mu_t) u_t \frac{g'''(\mu_t)}{g'(\mu_t)} \right. \\ &\quad - 3(y_t - \mu_t) u_t \left[\frac{g''(\mu_t)}{g'(\mu_t)} \right]^2 - u_t \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\quad - \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\quad - \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\quad \left. + \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \frac{g''(\mu_t)}{g'(\mu_t)} \right\}. \end{aligned}$$

Finalmente, temos que a expressão (3.16) fica dada por

$$\begin{aligned} \frac{\partial q_t}{\partial \beta_r} &= \frac{x_{tr}}{[g'(\mu_t)]^3} \left\{ \frac{(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} + \frac{3(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} - \frac{d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} + 2u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right. \\ &\quad - \frac{(-2y_t - 1 + 4\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} - \frac{(y_t - \mu_t)(1 - 2\mu_t)d'(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \\ &\quad + \frac{2(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{2(y_t - \mu_t)(1 - 2\mu_t)^2 d(y_t; \mu_t)}{\mu_t^3(1 - \mu_t)^3} \\ &\quad - \frac{3(-2y_t - 1 + 4\mu_t)}{\mu_t^4(1 - \mu_t)^4} + \frac{6(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{12(y_t - \mu_t)(1 - 2\mu_t)^2}{\mu_t^5(1 - \mu_t)^5} \\ &\quad - \frac{d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} + \frac{(y_t - \mu_t)d''(y_t; \mu_t)}{\mu_t(1 - \mu_t)} - \frac{2(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\quad - \frac{(y_t - \mu_t)(1 - 2\mu_t)d'(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} - (y_t - \mu_t)u_t \frac{g'''(\mu_t)}{g'(\mu_t)} + 3(y_t - \mu_t)u_t \left[\frac{g''(\mu_t)}{g'(\mu_t)} \right]^2 \\ &\quad + u_t \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} \frac{g''(\mu_t)}{g'(\mu_t)} \\ &\quad \left. - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \frac{g''(\mu_t)}{g'(\mu_t)} \right\}. \end{aligned}$$

Organizando e simplificando alguns termos da expressão anterior, obtemos

$$\begin{aligned} \frac{\partial q_t}{\partial \beta_r} &= \frac{x_{tr}}{[g'(\mu_t)]^3} \left\{ \frac{2(y_t + 1 - 3\mu_t)}{\mu_t^2(1 - \mu_t)^2} \left[d(y_t; \mu_t) + \frac{3}{\mu_t^2(1 - \mu_t)^2} \right] \right. \\ &\quad - \frac{2d'(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} [(y_t - \mu_t)(1 - 2\mu_t) + 1] + 3 \frac{g''(\mu_t)}{g'(\mu_t)} \left[u_t \right. \\ &\quad + \frac{(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^2(1 - \mu_t)^2} \left(d(y_t; \mu_t) + \frac{3}{\mu_t^2(1 - \mu_t)^2} \right) - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \left. \right] \\ &\quad - (y_t - \mu_t)u_t \left[\frac{g'''(\mu_t)}{g'(\mu_t)} - 3 \left(\frac{g''(\mu_t)}{g'(\mu_t)} \right)^2 \right] + \frac{12(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^5(1 - \mu_t)^5} \\ &\quad \left. + \frac{2(y_t - \mu_t)(1 - 2\mu_t)^2 d(y_t; \mu_t)}{\mu_t^3(1 - \mu_t)^3} + \frac{(y_t - \mu_t)d''(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \right\}. \end{aligned} \quad (3.17)$$

Definindo

$$\begin{aligned}
\tilde{q}_t = & \frac{1}{[g'(\mu_t)]^3} \left\{ \frac{2(y_t + 1 - 3\mu_t)}{\mu_t^2(1 - \mu_t)^2} \left[d(y_t; \mu_t) + \frac{3}{\mu_t^2(1 - \mu_t)^2} \right] \right. \\
& - \frac{2d'(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} [(y_t - \mu_t)(1 - 2\mu_t) + 1] + 3 \frac{g''(\mu_t)}{g'(\mu_t)} \left[u_t \right. \\
& + \left. \frac{(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^2(1 - \mu_t)^2} \left(d(y_t; \mu_t) + \frac{3}{\mu_t^2(1 - \mu_t)^2} \right) - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \right] \\
& - (y_t - \mu_t)u_t \left[\frac{g'''(\mu_t)}{g'(\mu_t)} - 3 \left(\frac{g''(\mu_t)}{g'(\mu_t)} \right)^2 \right] + \frac{12(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^5(1 - \mu_t)^5} \\
& \left. + \frac{2(y_t - \mu_t)(1 - 2\mu_t)^2 d(y_t; \mu_t)}{\mu_t^3(1 - \mu_t)^3} + \frac{(y_t - \mu_t)d''(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \right\}, \tag{3.18}
\end{aligned}$$

vemos que a equação (3.15) pode ser reescrita da seguinte forma:

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij}}{\partial\beta_r} = -\frac{1}{\sigma^2} \sum_{t=1}^n x_{ti}x_{tj}x_{tr}\tilde{q}_t.$$

Matricialmente, temos que

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]}{\partial\beta_r} = -\frac{1}{\sigma^2} X^\top X_r \tilde{Q} X, \tag{3.19}$$

sendo

$$X_r = \text{diag}\{x_{1r}, \dots, x_{nr}\}, \tag{3.20}$$

de modo que x_{tr} , corresponde ao (t, r) -ésimo elemento da matriz X , com $t = 1, \dots, n$ e $r = 1, \dots, p$ e sendo

$$\tilde{Q} = \text{diag}\{\tilde{q}_1, \dots, \tilde{q}_n\}, \tag{3.21}$$

em que $\tilde{q}_t$ é dado em (3.18).

Agora, da equação (2.14), sabemos que

$$[-\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]_i = -\sum_{t=1}^n f_t \frac{1}{g'(\mu_t)} x_{ti}, \tag{3.22}$$

então,

$$\begin{aligned}
\frac{\partial[-\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]_i}{\partial\beta_r} &= -\sum_{t=1}^n \frac{\partial}{\partial\beta_r} \left[f_t \frac{1}{g'(\mu_t)} \right] x_{ti} \\
&= -\sum_{t=1}^n \left\{ \frac{\partial}{\partial\beta_r} \left[\frac{1}{g'(\mu_t)} \right] f_t + \frac{\partial f_t}{\partial\beta_r} \frac{1}{g'(\mu_t)} \right\} x_{ti}.
\end{aligned}$$

Observamos que

$$\begin{aligned} \frac{\partial f_t}{\partial \beta_r} &= \frac{1}{(\sigma^2)^2} \frac{x_{tr}}{g'(\mu_t)} \left[u_t + \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \right. \\ &\quad \left. + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} \right] \end{aligned}$$

e

$$\frac{\partial}{\partial \beta_r} \left[\frac{1}{g'(\mu_t)} \right] = -x_{tr} \frac{g''(\mu_t)}{g'(\mu_t)^3}.$$

Assim,

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \left[f_t \frac{1}{g'(\mu_t)} \right] &= \frac{x_{tr}}{[\sigma^2 g'(\mu_t)]^2} \left[(y_t - \mu_t) u_t \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \right. \\ &\quad \left. + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} + u_t \right]. \end{aligned}$$

Definindo,

$$\begin{aligned} \tilde{g}_t &= \frac{1}{[\sigma^2 g'(\mu_t)]^2} \left[(y_t - \mu_t) u_t \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{(y_t - \mu_t)(1 - 2\mu_t)d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} \right. \\ &\quad \left. + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu_t^4(1 - \mu_t)^4} - \frac{(y_t - \mu_t)d'(y_t; \mu_t)}{\mu_t(1 - \mu_t)} + u_t \right], \end{aligned}$$

temos que

$$\frac{\partial [-\ddot{\ell}_{\sigma^2 \beta}(\boldsymbol{\theta})]_i}{\partial \beta_r} = - \sum_{t=1}^n x_{ti} x_{tr} \tilde{g}_t.$$

Escrevendo em forma matricial, obtemos

$$\frac{\partial [-\ddot{\ell}_{\sigma^2 \beta}(\boldsymbol{\theta})]}{\partial \beta_r} = -\tilde{\mathbf{g}}^\top X_r X, \quad (3.23)$$

em que, $\tilde{\mathbf{g}} = (\tilde{g}_1, \dots, \tilde{g}_n)^\top$. Podemos notar que

$$\frac{\partial [-\ddot{\ell}_{\beta \sigma^2}(\boldsymbol{\theta})]}{\partial \beta_r} = -X^\top X_r \tilde{\mathbf{g}}. \quad (3.24)$$

Finalmente, temos de (2.18) que

$$\begin{aligned} -\ddot{\ell}_{\sigma^2 \sigma^2}(\boldsymbol{\theta}) &= -\text{tr}(D) \\ &= - \sum_{t=1}^n d_t, \end{aligned} \quad (3.25)$$

em que d_t é definido em (2.17). Então,

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta})]}{\partial\beta_r} &= -\sum_{t=1}^n \frac{\partial d_t}{\partial\beta_r} \\ &= -\sum_{t=1}^n -\frac{1}{(\sigma^2)^3} \frac{d'(y_t; \mu_t)}{g'(\mu_t)} x_{tr} \\ &= -\sum_{t=1}^n \tilde{d}_t x_{tr}, \end{aligned}$$

sendo $\tilde{d}_t$ definido por

$$\tilde{d}_t = -\frac{1}{(\sigma^2)^3} \frac{d'(y_t; \mu_t)}{g'(\mu_t)}.$$

Podemos observar que

$$\frac{\partial[-\ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta})]}{\partial\beta_r} = -\text{tr}(X_r \tilde{D}), \quad (3.26)$$

em que, $\tilde{D} = \text{diag}\{\tilde{d}_1, \dots, \tilde{d}_n\}$.

Agora, vamos seguir, derivando as entradas de $-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})$, $-\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})$, $-\ddot{\ell}_{\beta\sigma^2}(\boldsymbol{\theta})$ e $-\ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta})$ em relação a σ^2 . Logo, de (3.14), (3.22) e (3.25), obtemos que

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij}}{\partial\sigma^2} = \frac{1}{(\sigma^2)^2} \sum_{t=1}^n x_{ti} x_{tj} q_t, \quad (3.27)$$

$$\frac{\partial[-\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]_i}{\partial\sigma^2} = \frac{2}{\sigma^2} \sum_{t=1}^n \frac{x_{ti}}{g'(\mu_t)} f_t \quad (3.28)$$

e

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta})]}{\partial\sigma^2} &= \frac{\partial}{\partial\sigma^2} \left(-\sum_{t=1}^n d_t \right) \\ &= -\sum_{t=1}^n \left[-\frac{1}{(\sigma^2)^3} + \frac{3d(y_t; \mu_t)}{(\sigma^2)^4} \right] \\ &= -\sum_{t=1}^n c_t, \end{aligned} \quad (3.29)$$

em que c_t é definido por

$$c_t = -\frac{1}{(\sigma^2)^3} + \frac{3d(y_t; \mu_t)}{(\sigma^2)^4}.$$

Portanto, a partir das expressões (3.27), (3.28) e (3.29), podemos definir, respectivamente, as seguintes matrizes:

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]}{\partial\sigma^2} = \frac{1}{(\sigma^2)^2} X^\top Q X, \quad (3.30)$$

$$\frac{\partial[-\ddot{\ell}_{\sigma^2\beta}(\boldsymbol{\theta})]}{\partial\sigma^2} = \frac{2}{\sigma^2} \mathbf{f}^\top T X. \quad (3.31)$$

e

$$\frac{\partial[-\ddot{\ell}_{\sigma^2\sigma^2}(\boldsymbol{\theta})]}{\partial\sigma^2} = -\text{tr}(C), \quad (3.32)$$

em que $C = \text{diag}\{c_1, \dots, c_n\}$ e com, $\partial[-\ddot{\ell}_{\beta\sigma^2}(\boldsymbol{\theta})]/\partial\sigma^2 = (2/\sigma^2) X^\top T \mathbf{f}$.

Dessa forma, a partir de (3.19), (3.23), (3.24), (3.26), (3.30), (3.31) e (3.32), obtemos, finalmente, as seguintes matrizes:

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\beta_r} = \begin{bmatrix} -\frac{1}{\sigma^2} X^\top X_r \tilde{Q} X & -X^\top X_r \tilde{\mathbf{g}} \\ -\tilde{\mathbf{g}}^\top X_r X & -\text{tr}(X_r \tilde{D}) \end{bmatrix} \quad (3.33)$$

e

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\sigma^2} = \begin{bmatrix} \frac{1}{(\sigma^2)^2} X^\top Q X & \frac{2}{\sigma^2} X^\top T \mathbf{f} \\ \frac{2}{\sigma^2} \mathbf{f}^\top T X & -\text{tr}(C) \end{bmatrix}. \quad (3.34)$$

Analogamente, vamos derivar cada entrada da matriz $\Delta(\boldsymbol{\theta})$, então, de (2.24), sabemos que a (i, t) -ésima entrada de $\Delta_\beta(\boldsymbol{\theta})$, com $i = 1, \dots, p$ e $t = 1, \dots, n$, é dada por

$$[\Delta_\beta(\boldsymbol{\theta})]_{it} = \frac{1}{\sigma^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti},$$

assim, derivando $[\Delta_\beta(\boldsymbol{\theta})]_{it}$, em relação a β_r , com $r = 1, \dots, p$ e em relação a σ^2 , obtemos, respectivamente, os seguintes resultados:

$$\begin{aligned} \frac{\partial[\Delta_\beta(\boldsymbol{\theta})]_{it}}{\partial\beta_r} &= \frac{1}{\sigma^2} x_{ti} \frac{\partial}{\partial\beta_r} \left\{ (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} \right\} \\ &= \frac{1}{\sigma^2} x_{ti} x_{tr} q_t \end{aligned}$$

e

$$\begin{aligned} \frac{\partial[\Delta_\beta(\boldsymbol{\theta})]_{it}}{\partial\sigma^2} &= \frac{\partial}{\partial\sigma^2} \left[\frac{1}{\sigma^2} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} x_{ti} \right] \\ &= f_t \frac{1}{g'(\mu_t)} x_{ti}. \end{aligned}$$

Matricialmente, temos

$$\frac{\partial[\Delta_\beta(\boldsymbol{\theta})]}{\partial\beta_r} = \frac{1}{\sigma^2} X^\top Q X_r \quad (3.35)$$

e

$$\frac{\partial[\Delta_\beta(\boldsymbol{\theta})]}{\partial\sigma^2} = X^\top L, \quad (3.36)$$

em que $L = \text{diag}\{l_1, \dots, l_n\}$, com

$$l_t = f_t \frac{1}{g'(\mu_t)}.$$

Finalmente, por (2.25), sabemos que a t -ésima entrada do vetor $\Delta_{\sigma^2}(\boldsymbol{\theta})$ é dada por

$$[\Delta_{\sigma^2}(\boldsymbol{\theta})]_t = \frac{-\sigma^2 + d(y_t; \mu_t)}{2(\sigma^2)^2},$$

em que $t = 1, \dots, n$. Facilmente vemos que, derivando $[\Delta_{\sigma^2}(\boldsymbol{\theta})]_t$ em relação a β_r , com $r = 1, \dots, p$, obtemos

$$\frac{\partial[\Delta_{\sigma^2}(\boldsymbol{\theta})]_t}{\partial\beta_r} = f_t \frac{1}{g'(\mu_t)} x_{tr}.$$

Matricialmente, temos que

$$\frac{\partial[\Delta_{\sigma^2}(\boldsymbol{\theta})]}{\partial\beta_r} = \mathbf{f}^\top T X_r. \quad (3.37)$$

Agora, derivando $[\Delta_{\sigma^2}(\boldsymbol{\theta})]_t$, em relação a σ^2 , temos

$$\begin{aligned} \frac{\partial[\Delta_{\sigma^2}(\boldsymbol{\theta})]_t}{\partial\sigma^2} &= \frac{\partial}{\partial\sigma^2} \left[\frac{-\sigma^2 + d(y_t; \mu_t)}{2(\sigma^2)^2} \right] \\ &= \frac{1}{2(\sigma^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma^2)^3} \\ &= d_t, \end{aligned}$$

Então,

$$\frac{\partial[\Delta_{\sigma^2}(\boldsymbol{\theta})]}{\partial\sigma^2} = (d_1, \dots, d_n)^\top = \mathbf{d}. \quad (3.38)$$

Dessa forma, a partir de (3.35), (3.36), (3.37) e (3.38), obtemos as matrizes

$$\frac{\partial[\Delta(\boldsymbol{\theta})]}{\partial\beta_r} = \begin{bmatrix} \frac{1}{\sigma^2} X^\top Q X_r \\ \mathbf{f}^\top T X_r \end{bmatrix} \quad (3.39)$$

e

$$\frac{\partial[\Delta(\boldsymbol{\theta})]}{\partial\sigma^2} = \begin{bmatrix} X^\top L \\ \mathbf{d}^\top \end{bmatrix}. \quad (3.40)$$

Após obtermos essas matrizes, vamos dar continuidade a estimação da matriz $J_1(\boldsymbol{\theta}_*)$. De acordo com as expressões (3.13) e (3.12), temos que

$$\begin{aligned} \frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_r} &= 2 \left\{ \frac{\partial\Delta^\top(\boldsymbol{\theta})}{\partial\beta_r} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \Delta(\boldsymbol{\theta}) + \Delta^\top(\boldsymbol{\theta}) \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\beta_r} \Delta(\boldsymbol{\theta}) \right. \\ &\quad \left. + \Delta^\top(\boldsymbol{\theta}) [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial\Delta(\boldsymbol{\theta})}{\partial\beta_r} \right\} \end{aligned} \quad (3.41)$$

e

$$\begin{aligned} \frac{\partial[H_{LD}(\delta_0)]}{\partial\sigma^2} &= 2 \left\{ \frac{\partial\Delta^\top(\boldsymbol{\theta})}{\partial\sigma^2} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \Delta(\boldsymbol{\theta}) + \Delta^\top(\boldsymbol{\theta}) \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\sigma^2} \Delta(\boldsymbol{\theta}) \right. \\ &\quad \left. + \Delta^\top(\boldsymbol{\theta}) [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial\Delta(\boldsymbol{\theta})}{\partial\sigma^2} \right\}, \end{aligned} \quad (3.42)$$

em que

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\beta_r} = -[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\beta_r} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}$$

e

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\sigma^2} = -[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\sigma^2} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1},$$

cujas matrizes $\Delta(\boldsymbol{\theta})$, $[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]$, $\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]/\partial\beta_r$, $\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]/\partial\sigma^2$, $\partial[\Delta(\boldsymbol{\theta})]/\partial\beta_r$ e $\partial[\Delta(\boldsymbol{\theta})]/\partial\sigma^2$ dadas por (2.26), (3.10), (3.33), (3.34), (3.39) e (3.40), respectivamente, sendo essas matrizes avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Assim, pela equação (3.11), temos que um estimador consistente para a matriz $J_1(\boldsymbol{\theta}_*)$ fica dado por

$$J_1(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_1} \right) \\ \vdots \\ \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_p} \right) \\ \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\sigma^2} \right) \end{bmatrix},$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Seguindo, temos agora que um estimador consistente para $h(y_t; \boldsymbol{\theta}_*)$, definida em (3.7), é dado por

$$\begin{aligned} h(y_t; \hat{\boldsymbol{\theta}}) &= 2\text{tr} \left\{ A(\hat{\boldsymbol{\theta}})^{-1} \left[\Delta_t \Delta_t^\top (\hat{\boldsymbol{\theta}}) - B(\hat{\boldsymbol{\theta}}) \right] \right\} \\ &\quad - \text{tr} \left\{ A(\hat{\boldsymbol{\theta}})^{-1} \left[-\frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} - A(\hat{\boldsymbol{\theta}}) \right] A(\hat{\boldsymbol{\theta}})^{-1} B(\hat{\boldsymbol{\theta}}) \right\}. \end{aligned} \quad (3.43)$$

Portanto, o estimador consistente de $k(y_t; \boldsymbol{\theta}_*)$ fica definido por

$$k(y_t; \hat{\boldsymbol{\theta}}) = h(y_t; \hat{\boldsymbol{\theta}}) + J_1(\hat{\boldsymbol{\theta}})^\top A(\hat{\boldsymbol{\theta}})^{-1} \Delta_t(\hat{\boldsymbol{\theta}}). \quad (3.44)$$

Dessa forma, para testar se o ajuste do modelo de regressão simplex é adequado para o conjunto de dados, devemos testar as hipóteses

$$\begin{aligned} H_0 : \lambda_0 &= 2(p+1), \\ H_1 : \lambda_0 &\neq 2(p+1), \end{aligned}$$

cujas estatística de teste, sob a hipótese H_0 , será dada por

$$U_n = \frac{n^{1/2} [\sum_{t=1}^n C_{It} - 2(p+1)]}{s_{k,c}}, \quad (3.45)$$

em que, $\sum_{t=1}^n C_{It} = 2\Delta^\top(\hat{\boldsymbol{\theta}})[- \ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})]^{-1} \Delta(\hat{\boldsymbol{\theta}})$, avaliada em $\boldsymbol{\delta} = \boldsymbol{\delta}_0$, com $\Delta(\boldsymbol{\theta})$ e $[- \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]$ dadas por (2.26) e (3.10), respectivamente, e $s_{k,c}^2$ corresponde a variância amostral de $k(y_t; \hat{\boldsymbol{\theta}})$, dado em (3.44).

3.3 O Teste para o Modelo Simplex Com Dispersão Variável

Considerando agora o modelo de regressão simplex, no caso em que o parâmetro de dispersão é variável, cujo logaritmo da verossimilhança é definido por (2.29). Nesse caso, a partir de (2.34), (2.37) e (2.42), segue que a matriz de informação observada fica dada por

$$- \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) = \begin{bmatrix} -X^\top \Sigma Q X & -X^\top T F H Z \\ -Z^\top T F H X & -Z^\top D^* Z \end{bmatrix}, \quad (3.46)$$

e a matriz $\Delta(\boldsymbol{\theta})$ é definida por (2.44), em que $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$.

Assim como para o caso com dispersão fixa, após obter o estimador de máxima verossimilhança, $\hat{\boldsymbol{\theta}}$, de $\boldsymbol{\theta}^*$, podemos obter os estimadores consistentes de $A(\boldsymbol{\theta}_*)$ e $B(\boldsymbol{\theta}_*)$ e antes de estimar a matriz $J_1(\boldsymbol{\theta}_*)$, precisamos obter as derivadas das matrizes $- \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})$ e $\Delta(\boldsymbol{\theta})$.

Dessa forma, inicialmente, vamos derivar as entradas de $-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})$ em relação a β_r , com $r = 1, \dots, p$. Logo, da equação (2.33), temos que

$$[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij} = -\sum_{i=1}^n \frac{1}{\sigma_t^2} x_{ti} x_{tj} q_t. \quad (3.47)$$

Similarmente ao caso com dispersão fixa, precisamos derivar q_t , em relação a β_r , cuja derivada é dada por (3.17). Segue que

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij}}{\partial\beta_r} = -\sum_{t=1}^n \frac{1}{\sigma_t^2} x_{ti} x_{tj} x_{tr} \tilde{q}_t,$$

em que $\tilde{q}_t$ é dado em (3.18). Em forma matricial, resultamos em

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]}{\partial\beta_r} = -X^\top X_r \Sigma \tilde{Q} X, \quad (3.48)$$

em que X_r , Σ e $\tilde{Q}$ são definidos, respectivamente, por (3.20), (2.30) e (3.21).

A partir da equação (2.36), sabemos que

$$[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]_{ij} = \sum_{i=1}^n \frac{1}{(\sigma_t^2)^2} u_t (y_t - \mu_t) \frac{x_{ti}}{g'(\mu_t)} \frac{z_{tj}}{h'(\sigma_t^2)}, \quad (3.49)$$

então,

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]_{ij}}{\partial\beta_r} &= -\sum_{t=1}^n \left\{ \frac{1}{(\sigma_t^2)^2} \frac{z_{tj}}{h'(\sigma_t^2)} \frac{x_{ti} x_{tr}}{[g'(\mu_t)]^2} \left[(y_t - \mu_t) u_t \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{(y_t - \mu_t) d'(y_t; \mu_t)}{\mu(1 - \mu_t)} \right. \right. \\ &\quad \left. \left. + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu^4(1 - \mu_t)^4} + \frac{(y_t - \mu_t)(1 - 2\mu_t) d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} + u_t \right] \right\} \\ &= -\sum_{t=1}^n z_{tj} x_{ti} x_{tr} m_t, \end{aligned}$$

com m_t definido por

$$\begin{aligned} m_t &= \frac{1}{(\sigma_t^2)^2 h'(\sigma_t^2) [g'(\mu_t)]^2} \left[(y_t - \mu_t) u_t \frac{g''(\mu_t)}{g'(\mu_t)} + \frac{(y_t - \mu_t) d'(y_t; \mu_t)}{\mu(1 - \mu_t)} \right. \\ &\quad \left. + \frac{3(y_t - \mu_t)(1 - 2\mu_t)}{\mu^4(1 - \mu_t)^4} + \frac{(y_t - \mu_t)(1 - 2\mu_t) d(y_t; \mu_t)}{\mu_t^2(1 - \mu_t)^2} + u_t \right]. \end{aligned}$$

Assim, colocando em forma matricial, obtemos

$$\frac{\partial[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]}{\partial\beta_r} = -Z^\top X_r M X, \quad (3.50)$$

em que $M = \text{diag}\{m_1, \dots, m_n\}$.

Facilmente vemos que,

$$\frac{\partial[-\ddot{\ell}_{\beta\gamma}(\boldsymbol{\theta})]}{\partial\beta_r} = -X^\top X_r M Z. \quad (3.51)$$

Finalmente, da equação (2.41), temos que

$$[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]_{ij} = -\sum_{t=1}^n d_t^* z_{tj} z_{ti}, \quad (3.52)$$

em que d_t^* é dado por (2.40). Então, é fácil ver que

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]_{ij}}{\partial\beta_r} &= -\sum_{t=1}^n \left\{ \frac{-d(y_t; \mu_t)}{(\sigma_t^2)^2} \left[\frac{1}{\sigma_t^2} + \frac{1}{2} \frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right] \frac{x_{tr}}{g'(\mu_t)} \frac{z_{tj} z_{ti}}{[h'(\sigma_t^2)]^2} \right\} \\ &= -\sum_{t=1}^n s_t x_{tr} z_{tj} z_{ti}, \end{aligned}$$

com s_t definido por

$$s_t = \frac{-d(y_t; \mu_t)}{(\sigma_t^2)^2} \left[\frac{1}{\sigma_t^2} + \frac{1}{2} \frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right] \frac{1}{g'(\mu_t) [h'(\sigma_t^2)]^2}.$$

Em forma matricial, temos

$$\frac{\partial[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]}{\partial\beta_r} = -Z^\top X_r S Z, \quad (3.53)$$

em que $S = \text{diag}\{s_1, \dots, s_n\}$.

Seguiremos agora, derivando as expressões (3.47), (3.49) e (3.52), em relação a γ_s com $s = 1, \dots, q$, então,

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]_{ij}}{\partial\gamma_s} &= -\sum_{t=1}^n x_{ti} x_{tj} q_t \frac{\partial}{\partial\gamma_s} \left(\frac{1}{\sigma_t^2} \right) \\ &= \sum_{t=1}^n x_{ti} x_{tj} q_t \frac{1}{(\sigma_t^2)^2} \frac{z_{ts}}{h'(\sigma_t^2)} \\ &= \sum_{t=1}^n x_{ti} x_{tj} o_t z_{ts} \end{aligned} \quad (3.54)$$

e

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]_{ij}}{\partial\gamma_s} &= -\sum_{t=1}^n \frac{\partial}{\partial\gamma_s} \left[-\frac{1}{(\sigma_t^2)^2} u_t(y_t - \mu_t) \frac{x_{ti}}{g'(\mu_t)} \frac{z_{tj}}{h'(\sigma_t^2)} \right] \\ &= -\sum_{t=1}^n \left\{ \frac{\partial}{\partial\gamma_s} \left[f_t^* \frac{1}{h'(\sigma_t^2)} \right] \frac{x_{ti} z_{tj}}{g'(\mu_t)} \right\}, \end{aligned} \quad (3.55)$$

em que o_t é denotado por

$$o_t = q_t \frac{1}{(\sigma_t^2)^2} \frac{1}{h'(\sigma_t^2)}$$

e f_t^* é definido por (2.38). Dessa forma, sabendo que

$$\frac{\partial f_t^*}{\partial \gamma_s} = \frac{2}{(\sigma_t^2)^3} u_t(y_t - \mu_t) \frac{z_{ts}}{h'(\sigma_t^2)} \quad \text{e} \quad \frac{\partial}{\partial \gamma_s} \left[\frac{1}{h'(\sigma_t^2)} \right] = - \frac{h''(\sigma_t^2)}{[h'(\sigma_t^2)]^3} z_{ts},$$

temos então que,

$$\frac{\partial}{\partial \gamma_s} \left[f_t^* \frac{1}{h'(\sigma_t^2)} \right] = \left[\frac{2}{(\sigma_t^2)^3} u_t(y_t - \mu_t) - \frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} f_t^* \right] \frac{z_{ts}}{[h'(\sigma_t^2)]^2}.$$

Portanto, de (3.55), verificamos que

$$\begin{aligned} \frac{\partial[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]_{ij}}{\partial \gamma_s} &= - \sum_{t=1}^n \left\{ \frac{x_{ti}}{g'(\mu_t)} \frac{z_{tj} z_{ts}}{[h'(\sigma_t^2)]^2} \frac{1}{(\sigma_t^2)^2} u_t(y_t - \mu_t) \left[\frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} + \frac{2}{\sigma_t^2} \right] \right\} \\ &= - \sum_{t=1}^n x_{ti} z_{tj} z_{ts} n_t, \end{aligned} \quad (3.56)$$

em que n_t é dado por

$$n_t = \frac{1}{g'(\mu_t)} \frac{1}{[h'(\sigma_t^2)]^2} \frac{1}{(\sigma_t^2)^2} u_t(y_t - \mu_t) \left[\frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} + \frac{2}{\sigma_t^2} \right].$$

Agora, derivando (3.52) em relação a γ_s , com $s = 1, \dots, q$, temos

$$\frac{\partial[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]_{ij}}{\partial \gamma_s} = - \sum_{t=1}^n z_{tj} z_{ti} \left(\frac{\partial d_t^*}{\partial \gamma_s} \right), \quad (3.57)$$

em que

$$\begin{aligned} \frac{\partial d_t^*}{\partial \gamma_s} &= \frac{z_{ts}}{[h'(\sigma_t^2)]^3} \left\{ - \frac{1}{(\sigma_t^2)^3} - \frac{3d(y_t; \mu_t)}{(\sigma_t^2)^4} - \frac{3h''(\sigma_t^2)}{h'(\sigma_t^2)} \left[\frac{1}{2(\sigma_t^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma_t^2)^3} \right] \right. \\ &\quad \left. + 3 \left[\frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right] a_t - \frac{h'''(\sigma_t^2)}{h'(\sigma_t^2)} a_t \right\}. \end{aligned}$$

Denotando,

$$\begin{aligned} \tilde{p}_t &= \frac{1}{[h'(\sigma_t^2)]^3} \left\{ - \frac{1}{(\sigma_t^2)^3} - \frac{3d(y_t; \mu_t)}{(\sigma_t^2)^4} - \frac{3h''(\sigma_t^2)}{h'(\sigma_t^2)} \left[\frac{1}{2(\sigma_t^2)^2} - \frac{d(y_t; \mu_t)}{(\sigma_t^2)^3} \right] \right. \\ &\quad \left. + 3 \left[\frac{h''(\sigma_t^2)}{h'(\sigma_t^2)} \right] a_t - \frac{h'''(\sigma_t^2)}{h'(\sigma_t^2)} a_t \right\}, \end{aligned}$$

temos que (3.57) fica dada por

$$\frac{\partial[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]_{ij}}{\partial\gamma_s} = -\sum_{t=1}^n z_{tj}z_{ti}z_{tr}\tilde{p}_t. \quad (3.58)$$

Dessa forma, a partir das expressões (3.54), (3.56) e (3.58), obtemos, respectivamente, as seguintes matrizes:

$$\frac{\partial[-\ddot{\ell}_{\beta\beta}(\boldsymbol{\theta})]}{\partial\gamma_s} = -X^\top Z_s O X, \quad (3.59)$$

$$\frac{\partial[-\ddot{\ell}_{\gamma\beta}(\boldsymbol{\theta})]}{\partial\gamma_s} = -Z^\top Z_s N X, \quad (3.60)$$

e

$$\frac{\partial[-\ddot{\ell}_{\gamma\gamma}(\boldsymbol{\theta})]}{\partial\gamma_s} = -Z^\top Z_s P Z, \quad (3.61)$$

em que $O = \text{diag}\{o_1, \dots, o_n\}$, $N = \text{diag}\{n_1, \dots, n_n\}$ e $P = \text{diag}\{\tilde{p}_1, \dots, \tilde{p}_n\}$ e sendo $Z_s = \text{diag}\{z_{1s}, \dots, z_{ns}\}$, de modo que z_{ts} , corresponde ao (t, s) -ésimo elemento da matriz Z , com $t = 1, \dots, n$ e $s = 1, \dots, q$ e lembrando que $\partial[-\ddot{\ell}_{\beta\gamma}(\boldsymbol{\theta})]/\partial\gamma_s = -X^\top Z_s N Z$.

A partir de (3.48), (3.50), (3.51), (3.53), (3.59), (3.60), (3.61) podemos definir as matrizes

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\beta_r} = \begin{bmatrix} -X^\top X_r \Sigma \tilde{Q} X & -X^\top X_r M Z \\ -Z^\top X_r M X & -Z^\top X_r S Z \end{bmatrix} \quad (3.62)$$

e

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\gamma_s} = \begin{bmatrix} -X^\top Z_s O X & -X^\top Z_s N Z \\ -Z^\top Z_s N X & -Z^\top Z_s P Z \end{bmatrix}. \quad (3.63)$$

Prosseguindo, vamos obter as derivadas da matriz $\Delta(\boldsymbol{\theta})$. Inicialmente, derivamos a (i, t) -ésima entrada de $\Delta_\beta(\boldsymbol{\theta})$, em relação a β_r , com $r = 1, \dots, p$ e a γ_s , com $s = 1, \dots, q$, logo, obtivemos que

$$\frac{\partial[\Delta_\beta(\boldsymbol{\theta})]_{it}}{\partial\beta_r} = \frac{1}{\sigma_t^2} x_{ti} x_{tr} q_t$$

e

$$\begin{aligned} \frac{\partial[\Delta_\beta(\boldsymbol{\theta})]_{it}}{\partial\gamma_s} &= f_t^* \frac{x_{ti}}{g'(\mu_t)} \frac{z_{ts}}{h'(\sigma_t^2)} \\ &= x_{ti} z_{ts} r_t, \end{aligned}$$

com r_t definindo como

$$r_t = f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\sigma_t^2)}.$$

Assim, seja $R = \text{diag}\{r_1, \dots, r_n\}$ e baseando-se nas derivadas de $[\Delta_\beta(\boldsymbol{\theta})]_{it}$ em relação a β_r e a γ_s , podemos obter, respectivamente, as matrizes

$$\frac{\partial[\Delta_\beta(\boldsymbol{\theta})]}{\partial\beta_r} = X^\top X_r \Sigma Q \quad (3.64)$$

e

$$\frac{\partial[\Delta_\beta(\boldsymbol{\theta})]}{\partial\gamma_s} = X^\top Z_s R. \quad (3.65)$$

Prosseguindo, facilmente vemos que, derivando $[\Delta_\gamma(\boldsymbol{\theta})]_{jt}$, em relação a β_r , com $r = 1, \dots, p$ e em relação a γ_s , com $s = 1, \dots, q$, temos, respectivamente,

$$\frac{\partial[\Delta_\gamma(\boldsymbol{\theta})]_{jt}}{\partial\beta_r} = f_t^* \frac{z_{tj}}{h'(\sigma_t^2)} \frac{x_{tr}}{g'(\mu_t)}$$

e

$$\frac{\partial[\Delta_\gamma(\boldsymbol{\theta})]_{jt}}{\partial\gamma_s} = d_t^* z_{tj} z_{ts}.$$

Matricialmente, obtemos

$$\frac{\partial[\Delta_\gamma(\boldsymbol{\theta})]}{\partial\beta_r} = Z^\top X_r R \quad (3.66)$$

e

$$\frac{\partial[\Delta_\gamma(\boldsymbol{\theta})]}{\partial\gamma_s} = Z^\top Z_s D. \quad (3.67)$$

Assim, a partir de (3.64), (3.65), (3.66) e (3.67), podemos finalmente definir as matrizes

$$\frac{\partial[\Delta(\boldsymbol{\theta})]}{\partial\beta_r} = \begin{bmatrix} X^\top X_r \Sigma Q \\ Z^\top X_r R \end{bmatrix} \quad (3.68)$$

e

$$\frac{\partial[\Delta(\boldsymbol{\theta})]}{\partial\gamma_s} = \begin{bmatrix} X^\top Z_s R \\ Z^\top Z_s D \end{bmatrix}. \quad (3.69)$$

Obtidas essas matrizes, seguimos com o objetivo de obter um estimador da matriz $J_1(\boldsymbol{\theta}_*)$. Nesse caso, temos que

$$\begin{aligned} \frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_r} &= 2 \left\{ \frac{\partial\Delta^\top(\boldsymbol{\theta})}{\partial\beta_r} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \Delta(\boldsymbol{\theta}) + \Delta^\top(\boldsymbol{\theta}) \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\beta_r} \Delta(\boldsymbol{\theta}) \right. \\ &\quad \left. + \Delta^\top(\boldsymbol{\theta}) [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial\Delta(\boldsymbol{\theta})}{\partial\beta_r} \right\} \end{aligned} \quad (3.70)$$

e

$$\begin{aligned} \frac{\partial[H_{LD}(\delta_0)]}{\partial\gamma_s} &= 2 \left\{ \frac{\partial\Delta^\top(\boldsymbol{\theta})}{\partial\gamma_s} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \Delta(\boldsymbol{\theta}) + \Delta^\top(\boldsymbol{\theta}) \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\gamma_s} \Delta(\boldsymbol{\theta}) \right. \\ &\quad \left. + \Delta^\top(\boldsymbol{\theta}) [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial\Delta(\boldsymbol{\theta})}{\partial\gamma_s} \right\}, \end{aligned} \quad (3.71)$$

em que

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\beta_r} = -[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\beta_r} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}$$

e

$$\frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1}}{\partial\gamma_s} = -[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1} \frac{\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{\partial\gamma_s} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]^{-1},$$

de forma que, nesse caso, as matrizes $\Delta(\boldsymbol{\theta})$, $[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]$, $\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]/\partial\beta_r$, $\partial[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]/\partial\gamma_s$, $\partial[\Delta(\boldsymbol{\theta})]/\partial\beta_r$ e $\partial[\Delta(\boldsymbol{\theta})]/\partial\gamma_s$ são dadas, respectivamente, por (2.44), (3.46), (3.62), (3.63), (3.68) e (3.69) todas essas matrizes avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Finalmente, temos que um estimador consistente para a matriz $J_1(\boldsymbol{\theta}_*)$ é dado por

$$J_1(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_1} \right) \\ \vdots \\ \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\beta_p} \right) \\ \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\gamma_1} \right) \\ \vdots \\ \text{tr} \left(\frac{\partial[H_{LD}(\delta_0)]}{\partial\gamma_q} \right) \end{bmatrix}, \quad (3.72)$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Logo, um estimador consistente para $h(y_t; \boldsymbol{\theta}_*)$, fica dado por

$$\begin{aligned} h(y_t; \hat{\boldsymbol{\theta}}) &= 2\text{tr} \left\{ A(\hat{\boldsymbol{\theta}})^{-1} \left[\Delta_t \Delta_t^\top(\hat{\boldsymbol{\theta}}) - B(\hat{\boldsymbol{\theta}}) \right] \right\} \\ &\quad - \text{tr} \left\{ A(\hat{\boldsymbol{\theta}})^{-1} \left[- \frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} - A(\hat{\boldsymbol{\theta}}) \right] A(\hat{\boldsymbol{\theta}})^{-1} B(\hat{\boldsymbol{\theta}}) \right\} \end{aligned} \quad (3.73)$$

e o estimador consistente de $k(y_t; \boldsymbol{\theta}_*)$ fica dado por

$$k(y_t; \hat{\boldsymbol{\theta}}) = h(y_t; \hat{\boldsymbol{\theta}}) + J_1(\hat{\boldsymbol{\theta}})^\top A(\hat{\boldsymbol{\theta}})^{-1} \Delta_t(\hat{\boldsymbol{\theta}}). \quad (3.74)$$

Assim, para testar se o ajuste do modelo de regressão simplex com dispersão variável é adequado para o conjunto de dados, devemos testar as hipóteses

$$H_0 : \lambda_0 = 2(p + q),$$

$$H_1 : \lambda_0 \neq 2(p + q),$$

cujas estatística de teste, sob a hipótese H_0 , será dada por

$$U_n = \frac{n^{1/2} [\sum_{t=1}^n C_{I_t} - 2(p + q)]}{s_{k,c}}, \quad (3.75)$$

em que, $\sum_{t=1}^n C_{I_t} = 2\Delta^\top(\hat{\boldsymbol{\theta}})[- \ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})]^{-1}\Delta(\hat{\boldsymbol{\theta}})$, avaliada em $\boldsymbol{\delta} = \boldsymbol{\delta}_0$, com $\Delta(\boldsymbol{\theta})$ e $[- \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]$ dadas por (2.44) e (3.46), respectivamente, e $s_{k,c}^2$ corresponde a variância amostral de $k(y_t; \boldsymbol{\theta}_*)$, dada por (3.74).

Capítulo 4

Avaliação Numérica dos Testes

Tendo como objetivo avaliar o desempenho dos testes desenvolvidos nesse trabalho, foram realizadas simulações de Monte Carlo com $R_{MC} = 5000$ réplicas.

Procuramos observar o desempenho dos testes propostos sob alguns tipos de erros de especificação: quando os dados são gerados de uma distribuição diferente da simplex, quando assumimos que o parâmetro de dispersão é fixo quando na verdade é variável ou quando são omitidas covariáveis durante o ajuste.

A função de ligação usada nesse estudo foi a logito, dada por

$$g(\mu_t) = \log\left(\frac{\mu_t}{1 - \mu_t}\right). \quad (4.1)$$

Logo, podemos facilmente verificar que derivando (4.1) em relação a μ_t obtemos

$$g'(\mu_t) = \frac{1}{\mu_t(1 - \mu_t)}. \quad (4.2)$$

Seguindo, obtemos que as derivadas segunda e terceira são dadas, respectivamente, por

$$g''(\mu_t) = \frac{-(1 - 2\mu_t)}{\mu_t^2(1 - \mu_t)^2} \quad (4.3)$$

e

$$g'''(\mu_t) = -2 \left[\frac{(1 - \mu_t)^2 - \mu_t(1 - 2\mu_t)}{\mu_t^3(1 - \mu_t)^3} \right]. \quad (4.4)$$

Assim, seja $y_1, y_2, \dots, y_n$ uma amostra de n variáveis aleatórias independentes, cuja média de y_t , com $t = 1, \dots, n$, satisfaz a relação (4.1), de forma que

$$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}. \quad (4.5)$$

Consideramos quatro covariáveis em cada caso de estudo, que foram geradas de distribuições diferentes e permaneceram fixas durante toda a simulação.

Além disso, foram considerados os tamanhos amostrais $n = 40, 80$ e 120 , em que a princípio são geradas 40 observações das covariáveis x_{ti} , com $i = 2, \dots, 5$, e em seguida esses valores são replicados duas e três vezes, obtendo-se assim $n = 80$ e $n = 120$, respectivamente. Esse procedimento de replicação dos valores de covariáveis é usado para garantir que o grau de heterogeneidade da dispersão permaneça fixo à medida que o tamanho da amostra aumenta. Esse grau de heterogeneidade da dispersão é comumente medido por

$$\lambda = \frac{\max_{t=1, \dots, n} (\sigma_t^2)}{\min_{t=1, \dots, n} (\sigma_t^2)}. \quad (4.6)$$

Podemos observar que se o parâmetro de dispersão é fixo, ou seja σ_t^2 é o mesmo para todo t , então $\lambda = 1$.

4.1 Avaliação do Teste para o Modelo Simplex com Dispersão Fixa

A princípio realizamos um estudo referente ao cálculo do tamanho do teste, em que utilizamos o fato que, sob a hipótese do modelo ajustado está adequado, temos que a estatística U_n converge em distribuição para uma normal padrão. Logo, desenvolvemos um programa em `R` com a seguinte metodologia:

1. fixamos os valores dos parâmetros $\beta_1, \beta_2, \beta_3, \beta_4$ e β_5 e também de σ^2 e α ;
2. geramos $x_{ti} \sim U(0, 1)$, com $t = 1, \dots, 40$ e $i = 2, \dots, 5$;
3. obtivemos $\mu_t = e^{\eta_t} / (1 + e^{\eta_t})$, em que $\eta_t = \sum_{i=1}^p x_{ti} \beta_i$;
4. iniciamos o laço de Monte Carlo, em que para cada réplica executamos as seguintes tarefas:
 - geramos uma amostra aleatória $y_1, y_2, \dots, y_n$ de n variáveis aleatórias independentes, em que cada $y_t \sim S^-(\mu_t, \sigma^2)$;
 - obtivemos os estimadores $\hat{\mu}_t$ e $\hat{\sigma}^2$ pelo método BFGS;
 - calculamos a estatística $T_n(\hat{\theta})$, definida em (3.4), sob a hipótese que $\lambda_0 = 2(p + 1)$;
 - calculamos $s_{k,c}$, que corresponde ao estimador do desvio padrão de $k(y_t; \theta_*)$, definida por (3.6);
 - obtivemos a estatística $U_n = T_n(\hat{\theta}) / s_{k,c}$;

- verificamos se $U_n < -z_{\alpha/2}$ ou $U_n > z_{\alpha/2}$, em que $z_{\alpha/2}$ é tal que $P(Z \geq z_{\alpha/2}) = \alpha/2$, com $Z \sim N(0, 1)$.

5. estimamos os valores de α , como sendo o número de estatísticas U_n cujos valores estão contidos na região crítica, estabelecida no item anterior, dividido pelo número de réplicas de Monte Carlo.

Os valores dos parâmetros do modelo, ou seja, de $\beta_1, \beta_2, \beta_3, \beta_4$ e β_5 , foram estabelecidos de forma que os valores de μ_t estejam distribuídos em três diferentes intervalos: próximos do zero, próximos de um e centrais (ou seja, próximos de 0,5), com isso, conseguimos observar o comportamento do teste em três diferentes situações. Assim, na Tabela 4.1 encontramos os valores que foram fixados para esses parâmetros e os intervalos que eles proporcionaram para o parâmetro μ .

Tabela 4.1: Valores verdadeiros dos parâmetros β 's, fixados para estimar os tamanhos do teste.

β_1	β_2	β_3	β_4	β_5	μ
-2,97	0,5	1,0	-1,8	0,65	$\mu \in (0, 019; 0, 147)$
2,0	-0,5	-1,4	1,25	-2,35	$\mu \in (0, 205; 0, 886)$
1,8	2,3	-0,5	1,34	0,5	$\mu \in (0, 903; 0, 995)$

Dessa forma, foram utilizados três níveis de significância, $\alpha = 1\%$, 5% e 10% , além disso, realizamos todo esse procedimento para $\sigma^2 = 0, 5, 1, 5, 4, 0$ e $16, 0$. Então, na Tabela 4.2 são apresentados os tamanhos do teste obtidos em todas essas situações.

Tabela 4.2: Tamanho do teste usando os quantis da distribuição normal padrão.

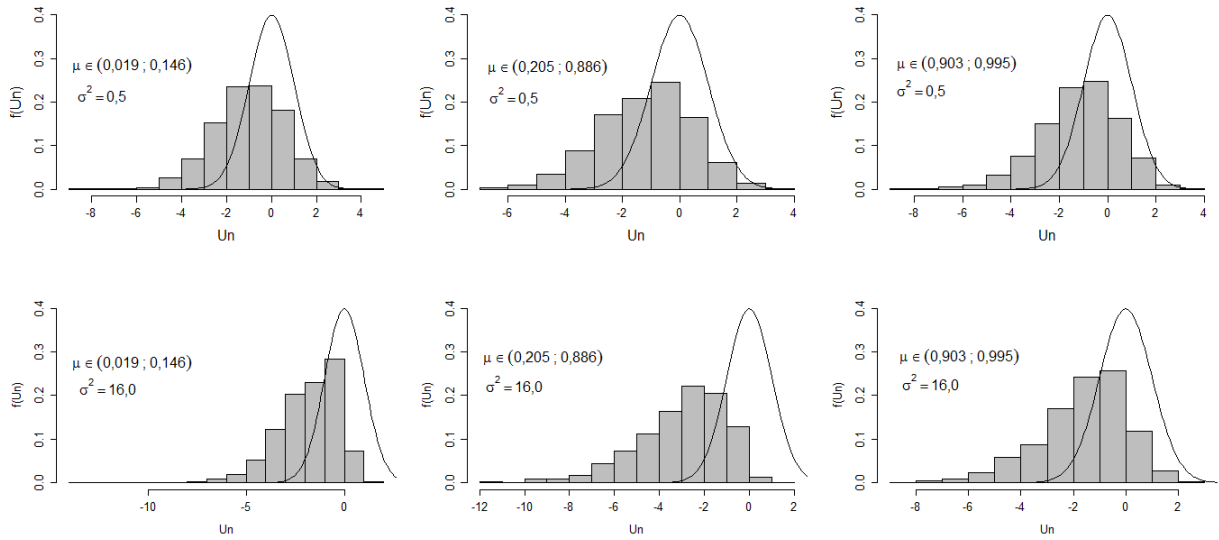
$\alpha \backslash \sigma^2$	$\mu \in (0, 019; 0, 147)$				$\mu \in (0, 205; 0, 886)$				$\mu \in (0, 903; 0, 995)$				n
	0, 5	1, 5	4, 0	16, 0	0, 5	1, 5	4, 0	16, 0	0, 5	1, 5	4, 0	16, 0	
1%	15,78	17,34	20,02	27,06	19,60	25,58	34,52	51,12	18,58	19,22	20,74	24,58	40
5%	27,52	29,48	32,46	39,74	32,08	38,22	48,30	65,06	30,52	30,72	32,34	35,74	
10%	35,24	37,58	40,46	47,74	38,84	46,08	55,98	72,10	38,16	38,62	40,06	42,84	
1%	10,10	12,04	17,22	31,00	11,14	17,08	26,56	44,04	11,24	12,36	13,62	19,82	80
5%	18,74	22,04	28,06	43,56	20,46	27,94	38,82	56,34	20,34	21,74	23,76	30,92	
10%	26,44	28,74	35,72	50,84	28,48	35,78	46,74	63,86	27,04	28,80	31,10	38,36	
1%	7,60	9,34	13,66	28,02	9,12	13,04	20,82	35,06	7,78	8,68	10,02	15,96	120
5%	14,70	17,00	23,80	40,44	17,40	22,74	32,06	47,42	15,96	16,90	18,94	26,58	
10%	20,90	23,88	30,26	47,66	23,48	30,60	40,08	54,62	21,96	23,62	25,80	33,40	

Observamos que os tamanhos estão bem acima dos valores nominais de α , principalmente quando μ assume valores centrais e quando o valor do parâmetro de dispersão é alto. Também notamos que as estimativas do tamanho do teste vão diminuindo à medida que o tamanho amostral aumenta, porém, isso ocorre muito lentamente.

Diante desse resultado, resolvemos investigar o comportamento da estatística U_n , ou seja, realizamos um pequeno estudo de simulação para verificar se de fato, a distribuição de U_n era próxima da normal padrão. Para isso geramos 1000 valores dessa estatística, por meio

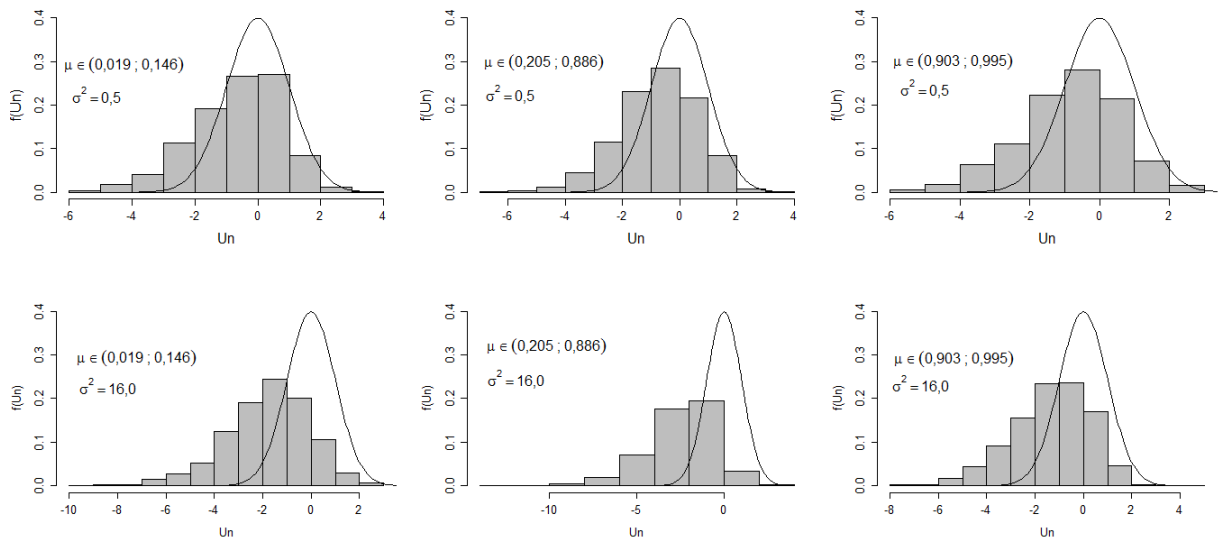
de uma breve simulação de Monte Carlo, para os três tamanhos amostrais e utilizamos os mesmos valores dos parâmetros presentes na Tabela 4.1. Fixamos apenas dois valores para o parâmetro de dispersão, $\sigma^2 = 0,5$ e $\sigma^2 = 16$, para que possamos observar como a distribuição de U_n se comporta quando esse parâmetro assume um valor baixo ou alto e geramos $x_{ti} \sim U(0,1)$, com $t = 1, \dots, 40$ e $i = 2, \dots, 5$. Com os valores de U_n fizemos os histogramas que são apresentados nas Figuras 4.1, 4.2 e 4.3.

Figura 4.1: Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 40$.



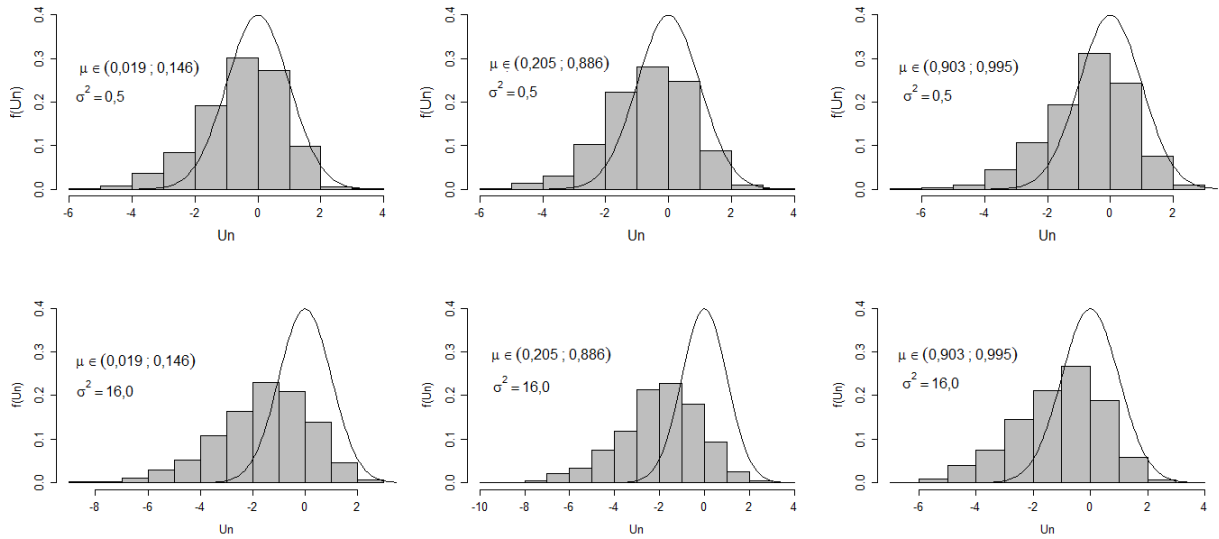
Fonte: Autoria própria.

Figura 4.2: Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 80$.



Fonte: Autoria própria.

Figura 4.3: Histogramas da estatística U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 120$.



Fonte: Autoria própria.

Observamos que os histogramas referentes a $\sigma^2 = 0,5$ ajustaram-se melhor a curva da normal padrão, com exceção de quando $n = 40$. Porém, quando $\sigma^2 = 16$, temos que a distribuição da estatística difere bastante dessa curva, mesmo quando o tamanho amostral é consideravelmente grande, isso implica que, para dados de uma distribuição simplex com parâmetro de dispersão alto, não é adequado assumir que a estatística U_n segue uma distribuição normal com média zero e variância um.

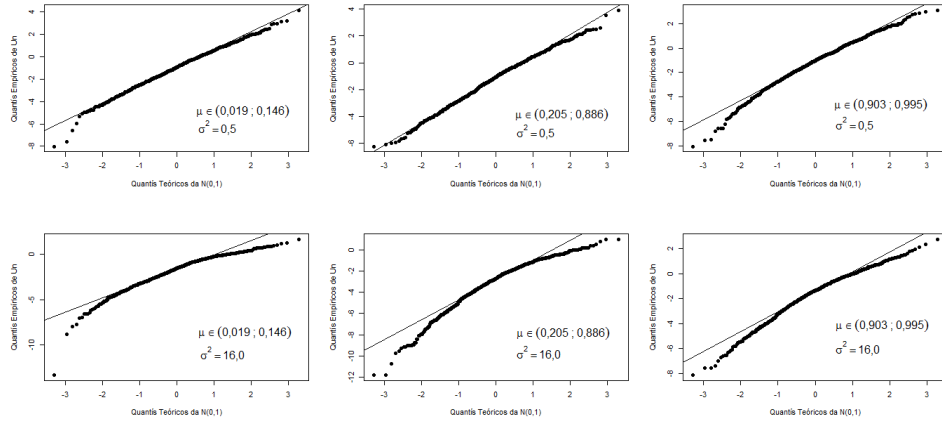
Podemos observar que em todas as situações consideradas, a distribuição da estatística U_n apresenta assimetria, principalmente quando os dados possuem uma dispersão alta. Provavelmente, os verdadeiros quantis da distribuição de U_n devem diferir bastante dos quantis da distribuição normal padrão. Dessa forma, os resultados do teste considerando os quantis da distribuição normal podem não ser confiáveis.

Nas Figuras 4.4, 4.5 e 4.6 encontramos os gráficos de *qqplot*'s usados para verificarmos a normalidade de U_n , sua metodologia consiste em confrontar os quantis empíricos de U_n obtidos a partir da amostra com os quantis teóricos da distribuição normal padrão. Foram considerados os mesmos valores da U_n usados na obtenção dos histogramas.

Observamos ainda, a presença de uma reta, comumente chamada de *qqpline* ou linha de normalidade. Os pontos devem ficar alinhados em torno dessa linha caso a estatística U_n siga uma distribuição normal padrão.

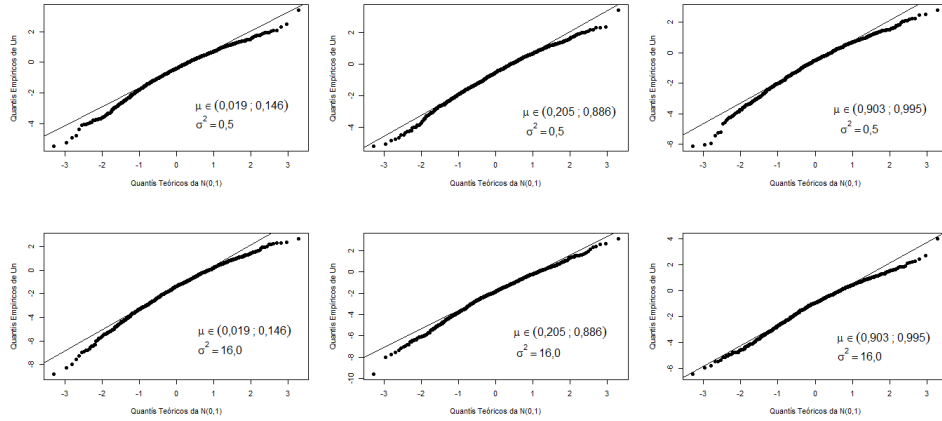
Com base nas Figuras 4.4, 4.5 e 4.6 observamos em todos os tamanhos amostrais considerados, que apenas para os quantis da normal concentrados em torno do zero, os pontos encontram-se sobre a linha. Assim, apenas para valores bem pequenos da estatística U_n a suposição de normalidade torna-se plausível.

Figura 4.4: Gráficos *qqplot*'s da distribuição de U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 40$.



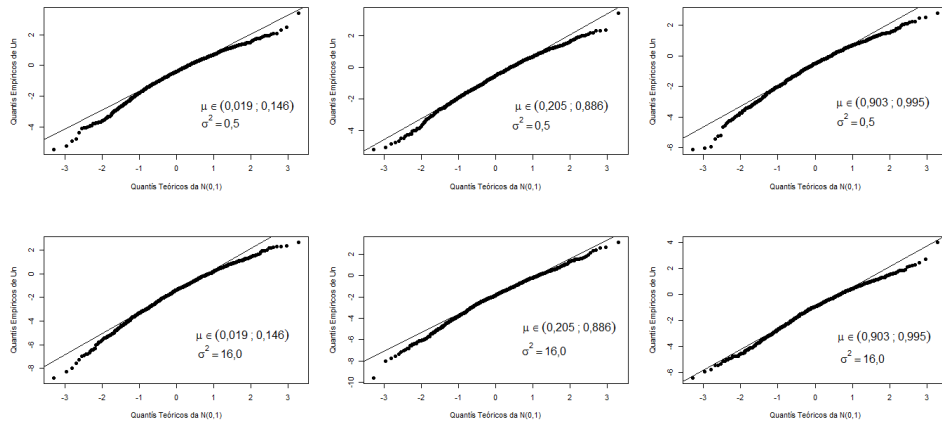
Fonte: Autoria própria.

Figura 4.5: Gráficos *qqplot*'s da distribuição de U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 80$.



Fonte: Autoria própria.

Figura 4.6: Gráficos *qqplot*'s da distribuição de U_n para diferentes cenários de μ e valores de σ^2 iguais a 0,5 e 16, respectivamente, para $n = 120$.



Fonte: Autoria própria.

Além desses gráficos, procuramos avaliar também algumas medidas, consideradas importantes na caracterização de uma distribuição de probabilidade. Sabemos que no caso da distribuição normal padrão, a média, a variância, a curtose e o coeficiente de assimetria dever ser iguais à zero, um, três e zero, respectivamente. Logo, baseando-se nos 1000 valores gerados para a estatística U_n procuramos estimar essas medidas, na Tabela 4.3 encontramos os valores obtidos.

Tabela 4.3: Estimativas dos valores de algumas medidas características da distribuição de U_n , para $n = 40, 80$ e 120 .

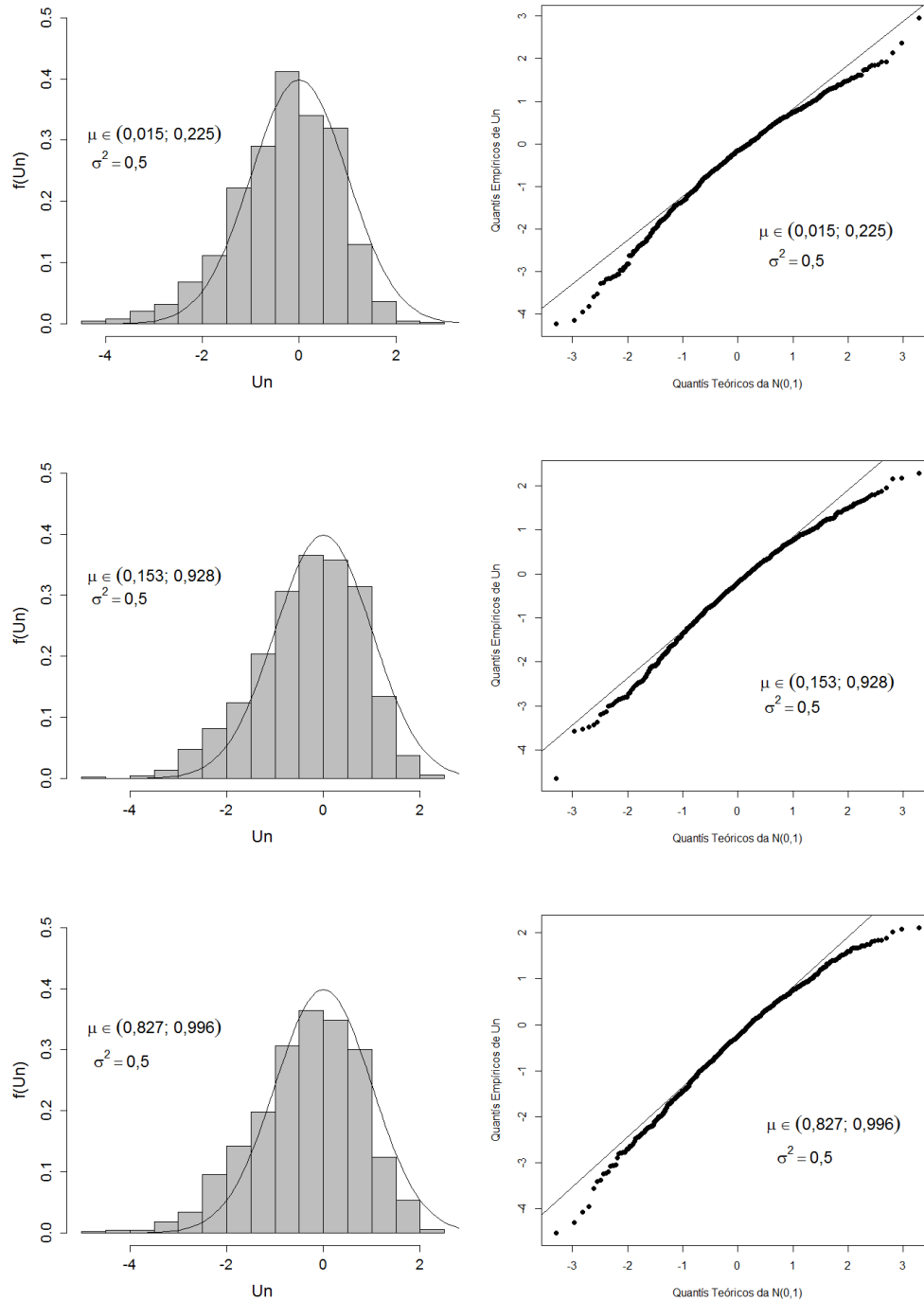
σ^2	$\mu \in (0, 019; 0, 147)$				$\mu \in (0, 205; 0, 886)$				$\mu \in (0, 903; 0, 995)$				n
	Média	Variância	Curtose	Assimetria	Média	Variância	Curtose	Assimetria	Média	Variância	Curtose	Assimetria	
0,5	-0,994	2,464	3,459	-0,289	-1,170	2,624	2,899	-0,242	-1,159	2,715	3,694	-0,519	40
16,0	-1,803	2,381	6,481	-1,130	-3,003	3,861	4,173	-0,961	-1,591	2,762	3,500	-0,685	
0,5	-0,654	1,971	3,063	-0,488	-0,768	1,905	3,243	-0,424	-0,809	2,112	3,072	-0,441	80
16,0	-1,795	3,026	3,828	-0,602	-2,457	3,641	4,290	-0,711	-1,324	2,612	3,190	-0,443	
0,5	-0,536	1,661	3,370	-0,601	-0,633	1,745	3,184	-0,439	-0,670	1,851	3,542	-0,612	120
16,0	-1,574	3,233	3,409	-0,571	-1,968	3,289	3,334	-0,464	-1,105	2,380	2,962	-0,431	

Pela Tabela 4.3 vemos que, a média de U_n assume valores negativos e distantes de zero, principalmente quando $\sigma^2 = 16$ e $\mu \in (0, 205; 0, 886)$, no caso da variância, notamos que em todas as situações, seus valores estimados estão distantes de um. Para os valores estimados da curtose, observamos que, com exceção dos casos em que $\sigma^2 = 16$, $n = 40$ e $\mu \in (0, 019; 0, 147)$ ou $\mu \in (0, 205; 0, 886)$, ela assume valores próximos de três. Por fim, vemos também que a distribuição de U_n apresenta certa assimetria à esquerda.

Dessa forma, procuramos realizar um estudo de simulação considerando tamanhos amostrais bastante elevados, em que foram utilizados $n = 500$ e 1000 . Trabalhamos com os mesmos cenários apresentados pela Tabela 4.1 e novamente fixamos apenas dois valores para o parâmetro de dispersão, $\sigma^2 = 0,5$ e $\sigma^2 = 16$. Nesse caso, foram geradas a princípio 500 observações das covariáveis do modelo, de modo que, $x_{ti} \sim U(0, 1)$, com $t = 1, \dots, 500$ e $i = 2, \dots, 5$, em seguida replicamos esses valores uma vez, obtendo-se então $n = 1000$. Geramos 1000 valores da estatística U_n e obtemos os histogramas e gráficos *qqplot*'s da distribuição de U_n .

Avaliando os gráficos da Figura 4.7 vemos um bom ajuste dos histogramas com a curva da normal padrão, no entanto, pelos gráficos *qqplot*'s temos que, embora o tamanho amostral seja consideravelmente alto, a suposição de normalidade da estatística U_n não é adequada.

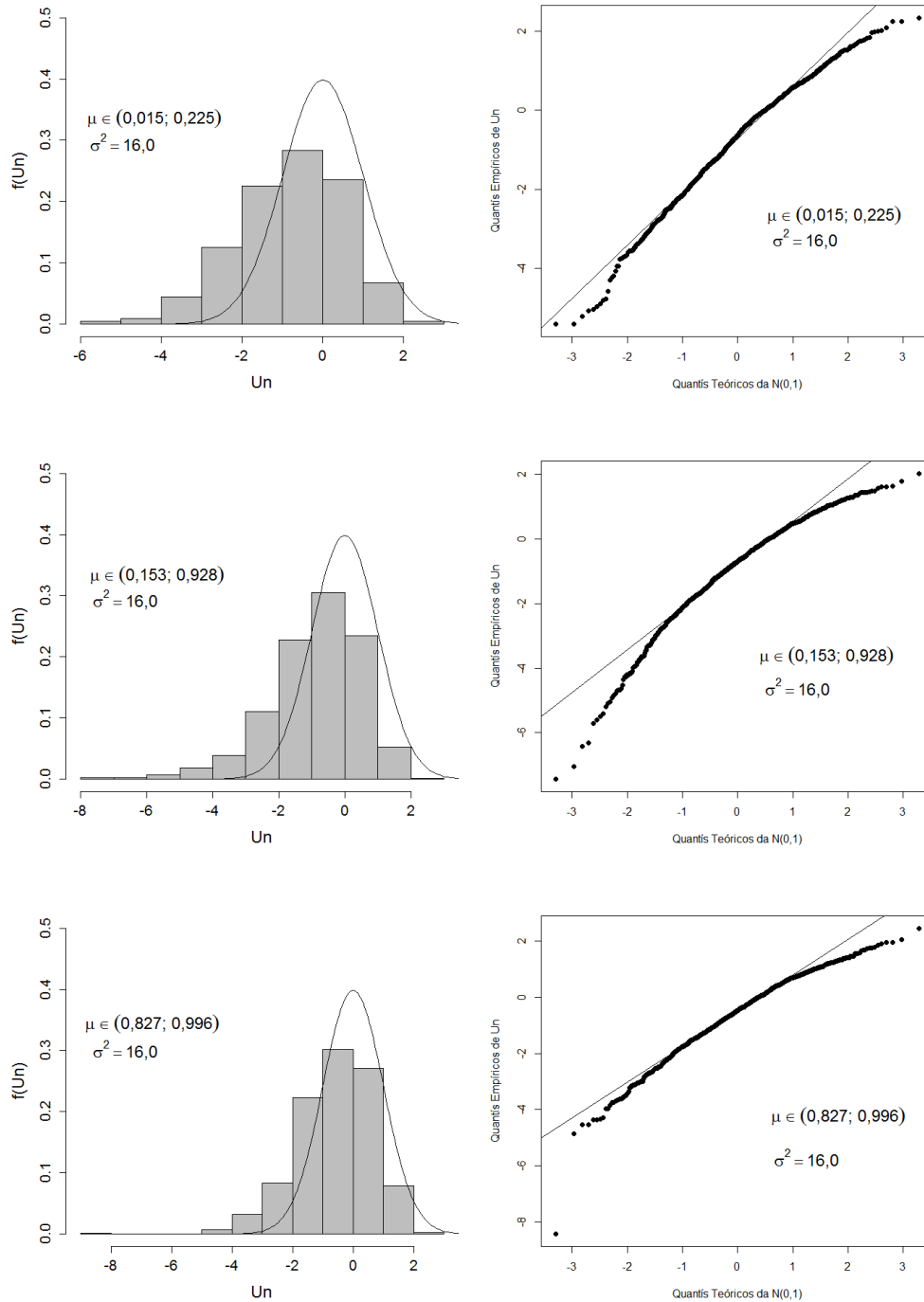
Figura 4.7: Histogramas e gráficos *qqplot*'s da estatística U_n para diferentes cenários de μ , $\sigma^2 = 0,5$ e $n = 500$.



Fonte: Autoria própria.

Pela Figura 4.8 vemos os histogramas referentes a $\sigma^2 = 16,0$, em que notamos uma certa assimetria na distribuição de U_n , tornando o ajuste da curva da normal padrão inadequado, principalmente quando os valores de μ estão próximos de zero. Essa conclusão é reforçada quando observamos os gráficos *qqplot*'s, também presentes na Figura 4.8.

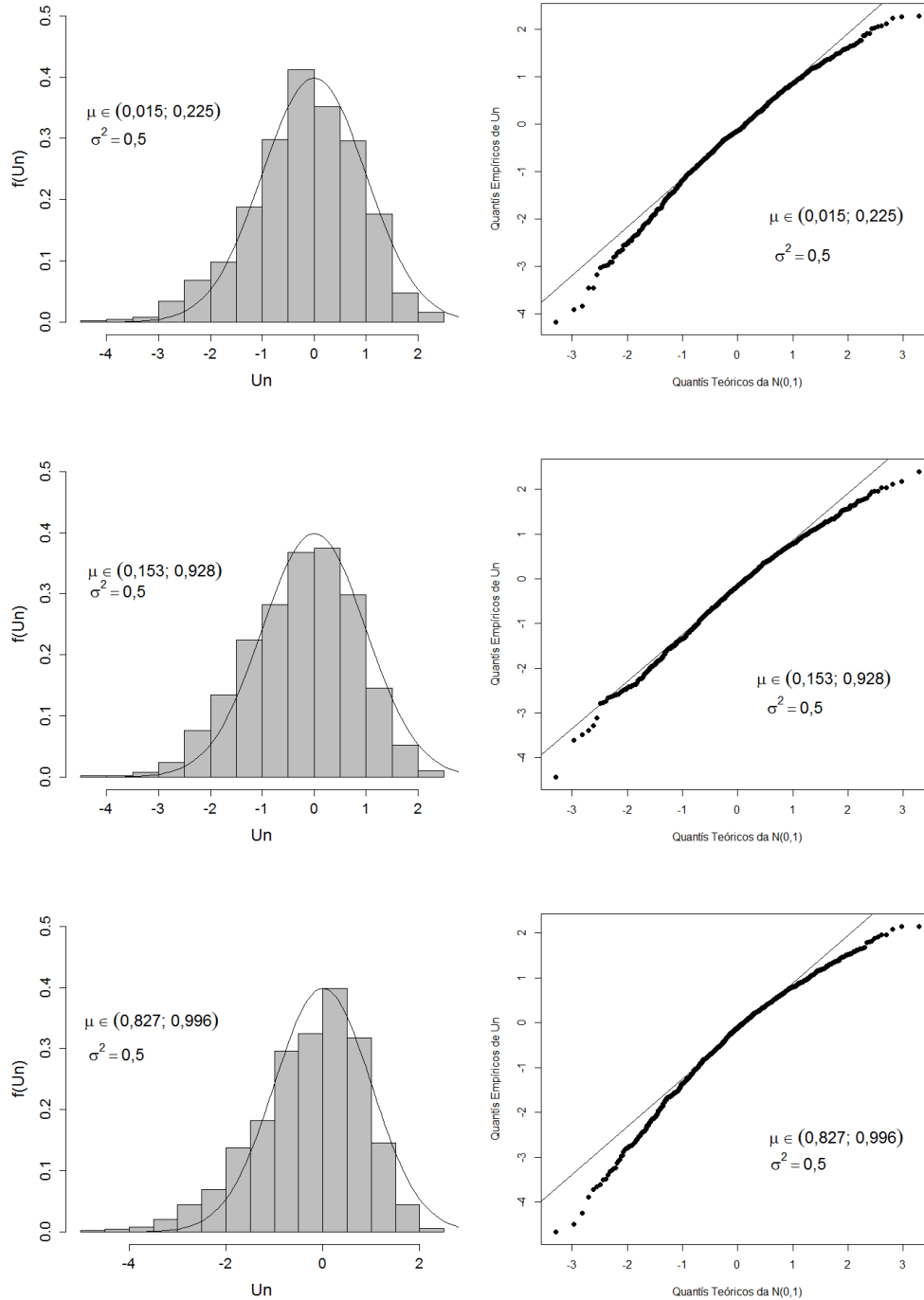
Figura 4.8: Histogramas e gráficos *qqplot*'s da estatística U_n para diferentes cenários de μ , $\sigma^2 = 16,0$ e $n = 500$.



Fonte: Autoria própria.

Na Figura 4.9 encontramos os histogramas e gráficos *qqplot*'s da distribuição de U_n , para $n = 1000$. Observamos nesse caso, um ajuste ainda melhor dos histogramas com a curva da normal padrão, entretanto, pelos gráficos *qqplot*'s vemos que, embora o tamanho amostral seja bastante elevado, a suposição de normalidade da estatística U_n continua não sendo completamente adequada.

Figura 4.9: Histogramas e gráficos *qqplot*'s da estatística U_n para diferentes cenários de μ , $\sigma^2 = 0,5$ e $n = 1000$.



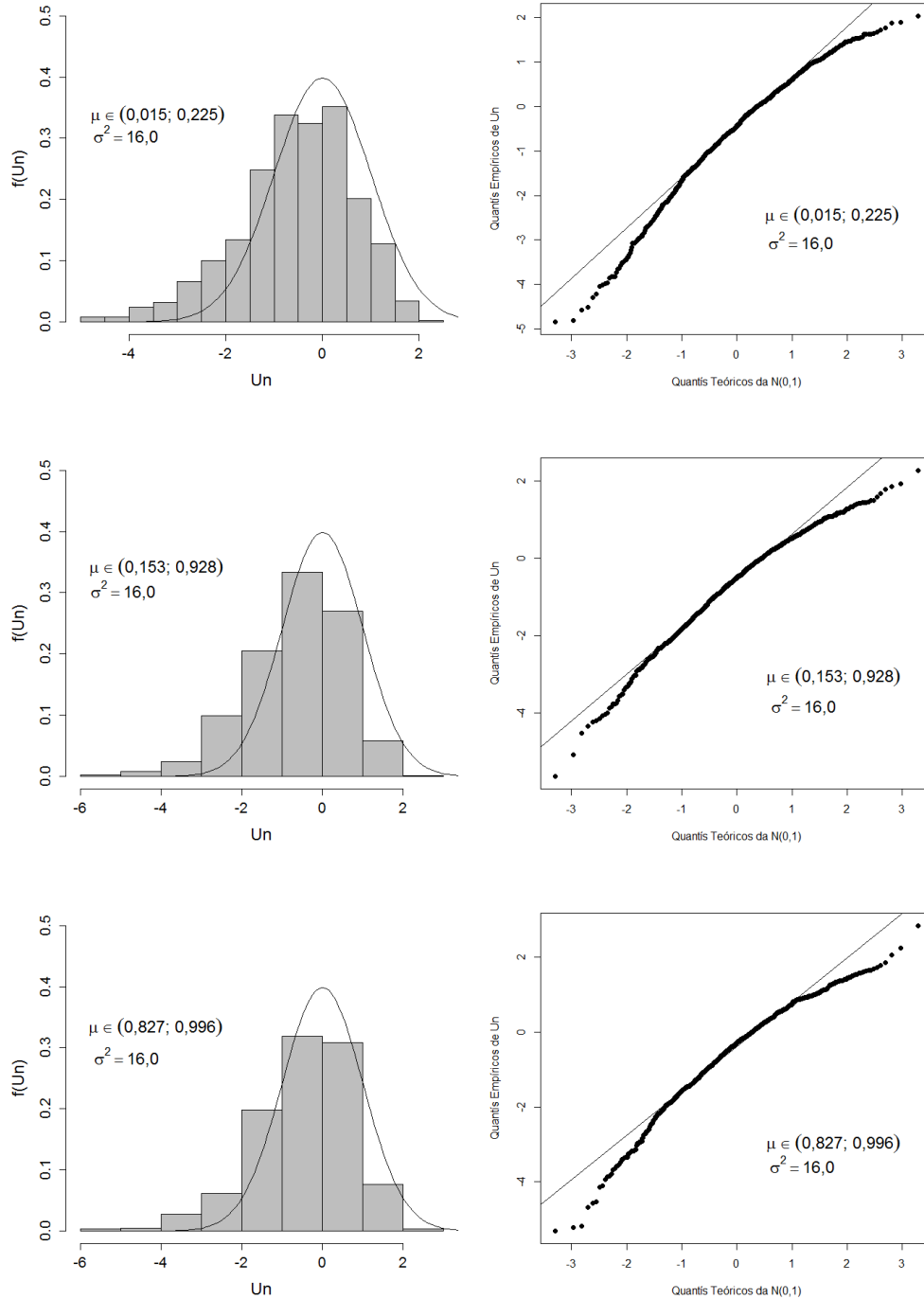
Fonte: Autoria própria.

Para a situação em que $\sigma^2 = 16,0$ e $n = 1000$, observamos na Figura 4.10 uma leve melhora no ajuste da curva da normal padrão com os histogramas, quando comparamos com os gráficos com $\sigma^2 = 16,0$ e tamanhos amostrais menores.

Entretanto, os gráficos *qqplot*'s ainda mostram que a suposição de normalidade da estatística U_n não é adequada. Logo, podemos concluir por fim, que, para dados de uma distribuição simplex com alta dispersão, não é aconselhável assumir que a estatística

U_n segue uma distribuição aproximadamente normal padrão, mesmo que o tamanho da amostra utilizado seja relativamente grande.

Figura 4.10: Histogramas e gráficos *qqplot*'s da estatística U_n para diferentes cenários de μ , $\sigma^2 = 16,0$ e $n = 1000$.



Fonte: Autoria própria.

Pela Tabela 4.4 vemos que, para $\sigma^2 = 0,5$ a média e o coeficiente de assimetria da distribuição de U_n estão próximos de zero, porém, seus valores tendem à serem negativos, indicando que a distribuição de U_n é assimétrica à esquerda. Entretanto, vemos que a

variância torna-se mais próxima de um e a curtose próxima de três.

No entanto, para $\sigma^2 = 16,0$, observamos que os valores estimados de todas as medidas estão distantes dos valores referentes a distribuição normal padrão, reforçando a conclusão de que, para altos valores de σ^2 a distribuição de U_n não possui uma boa aderência a distribuição normal padrão.

Tabela 4.4: Estimativas dos valores de algumas medidas características da distribuição de U_n , para $n = 500$ e $n = 1000$.

σ^2	$\mu \in (0,015; 0,225)$				$\mu \in (0,153; 0,928)$				$\mu \in (0,827; 0,996)$				n
	Média	Variância	Curtose	Assimetria	Média	Variância	Curtose	Assimetria	Média	Variância	Curtose	Assimetria	
0,5	-0,284	1,109	3,447	-0,568	-0,302	1,132	3,106	-0,507	-0,329	1,183	3,129	-0,469	500
16,0	-0,767	1,782	3,129	-0,476	-0,867	1,890	4,472	-0,932	-0,584	1,545	4,405	-0,712	
0,5	-0,193	1,083	3,247	-0,462	-0,244	1,066	2,952	-0,377	-0,269	1,208	3,482	-0,679	1000
16,0	-0,542	1,411	3,417	-0,641	-0,632	1,380	3,378	-0,610	-0,451	1,431	3,772	-0,743	

Buscando avaliar o poder do teste em identificar o erro de especificação relativo à distribuição da variável resposta, geramos dados com base na distribuição beta, competidora direta da simplex no intervalo $(0,1)$.

Seguimos com o método de simulação utilizado para o cálculo do tamanho do testes, porém agora, os valores da amostra aleatória $y_1, y_2, \dots, y_n$ são gerados de n variáveis aleatórias independentes, com distribuição beta de média μ_t e parâmetro de precisão ϕ . Entretanto, calculamos a estatística U_n como se os dados fossem provenientes de uma distribuição simplex de parâmetros μ_t e σ^2 . Fixamos $\phi = 40$ e apresentamos na Tabela 4.5 os valores que foram fixados para os β 's e os respectivos intervalos para o parâmetro μ . Geramos $x_{ti} \sim U(0,1)$, com $t = 1, \dots, 40$ e $i = 2, \dots, 5$.

O procedimento de gerar valores da distribuição beta e estimar os parâmetros por uma simplex, pode acarretar em uma certa demora no processo de simulação, ou até mesmo a falta de convergência, a depender dos valores fixados para os parâmetros, especialmente do parâmetro ϕ . Por esse motivo, buscamos fixar outros valores para os parâmetros do modelo, diferentes dos que foram fixados para o tamanho do teste, porém sem se afastar muito das situações antes consideradas.

Tabela 4.5: Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando dados da distribuição beta.

β_1	β_2	β_3	β_4	β_5	μ
-1,4	-0,7	-0,4	-0,9	-1,5	$\mu \in (0,014; 0,130)$
-1,0	1,0	-1,0	1,1	1,5	$\mu \in (0,225; 0,853)$
1,8	1,1	0,3	1,1	0,9	$\mu \in (0,902; 0,990)$

Pela Tabela 4.6 podemos observar que o teste, nesse caso, possui poderes muito baixos, principalmente quando os valores da média estão no intervalo $(0,225; 0,853)$ e o tamanho da amostra e o nível de significância do teste são baixos. Notamos também, que em cada situação de μ , não existe mudança no poder do teste nos diferentes valores de σ^2 .

Tabela 4.6: Poder do teste para o modelo simplex com dispersão fixa, quando os dados são gerados da distribuição beta.

$\alpha \backslash \sigma^2$	$\mu \in (0,014; 0,130)$				$\mu \in (0,225; 0,853)$				$\mu \in (0,902; 0,990)$				n
	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	
1%	13,10	12,96	13,12	13,10	10,06	10,06	10,06	10,06	19,10	19,10	19,10	19,10	40
5%	19,94	19,88	20,00	20,06	20,30	20,30	20,30	20,30	27,00	26,84	26,84	26,84	
10%	25,32	25,44	25,50	25,56	28,70	28,70	28,70	28,70	32,42	32,12	32,12	32,12	
1%	21,62	21,70	21,70	21,68	4,00	4,00	4,00	4,00	34,82	34,78	34,76	34,76	80
5%	38,34	38,44	38,44	38,42	13,92	13,92	13,92	13,92	50,76	50,82	50,84	50,84	
10%	49,10	49,22	49,24	49,22	23,86	23,86	23,86	23,86	61,04	61,14	61,12	61,12	
1%	26,20	26,28	26,28	26,28	3,42	3,42	3,42	3,42	42,62	42,60	42,60	42,60	120
5%	48,12	48,28	48,26	48,24	14,06	14,06	14,06	14,06	63,20	63,18	63,18	63,18	
10%	62,36	62,48	62,50	62,48	25,68	25,68	25,68	25,68	75,00	75,00	75,00	75,00	

Diante do fato que a distribuição de U_n não tem uma boa aderência a distribuição normal padrão e comprometendo o desempenho do teste baseado nos quantis desta distribuição, propomos utilizar os quantis empíricos da distribuição de U_n , obtidos via método *bootstrap* paramétrico, tanto para obter o tamanho quanto o poder do teste.

4.1.1 Avaliação do Teste Usando Quantis *Bootstrap*

Desenvolvemos a simulação de Monte Carlo para calcular o tamanho e também o poder do teste em detectar a má especificação do modelo quando geramos dados com base na distribuição beta e ajustamos o modelo simplex, baseando-se na seguinte metodologia:

1. fixamos os valores dos parâmetros $\beta_1, \beta_2, \beta_3, \beta_4$ e β_5 e também de ϕ, σ^2 e α ;
2. geramos $x_{ti} \sim U(0, 1)$, com $t = 1, \dots, 40$ e $i = 2, \dots, 5$;
3. obtivemos $\mu_t = e^{\eta_t} / (1 + e^{\eta_t})$, em que $\eta_t = \sum_{i=1}^p x_{ti} \beta_i$;
4. iniciamos o laço de Monte Carlo, em que para cada réplica executamos as seguintes tarefas:
 - geramos uma amostra aleatória $y_{11}, y_{12}, \dots, y_{1n}$ de n variáveis aleatórias independentes, em que cada $y_{1t} \sim S^-(\mu_t, \sigma^2)$ e obtemos os estimadores $\widehat{\mu}_{1t}$ e $\widehat{\sigma}_1^2$ pelo método BFGS;
 - geramos uma amostra aleatória $y_{21}, y_{22}, \dots, y_{2n}$ de n variáveis aleatórias independentes, em que cada $y_{2t} \sim B(\mu_t, \phi)$ e obtemos os estimadores $\widehat{\mu}_{2t}$ e $\widehat{\sigma}_2^2$ pelo método BFGS;
 - pela amostra $y_{11}, y_{12}, \dots, y_{1n}$, calculamos a estatística $T_{n1}(\widehat{\theta})$, sob a hipótese que $\lambda_0 = 2(p+1)$ e calculamos $s_{k,c,1}$;
 - obtivemos a estatística $U_{n1} = T_{n1}(\widehat{\theta}) / s_{k,c,1}$;

- pela amostra $y_{21}, y_{22}, \dots, y_{2n}$, calculamos a estatística $T_{n2}(\hat{\theta})$, sob a hipótese que $\lambda_0 = 2(p+1)$ e calculamos $s_{k,c,2}$;
 - obtivemos a estatística $U_{n2} = T_{n2}(\hat{\theta})/s_{k,c,2}$;
 - iniciamos o laço *bootstrap*, em que para cada réplica ($b = 1, \dots, 500$), realizamos as seguintes tarefas:
 - (i) geramos uma amostra aleatória $(y_{b1}^*, y_{b2}^*, \dots, y_{bn}^*)$ de n variáveis aleatórias independentes, em que cada $y_{bt} \sim S^-(\widehat{\mu}_{1t}, \widehat{\sigma}_1^2)$;
 - (ii) obtivemos a estatística $U_{bn}^* = T_{bn}^*(\hat{\theta})/s_{k,c,b}^*$, em que $T_{bn}^*(\hat{\theta})$ e $s_{k,c,b}^*$ são obtidos da amostra *bootstrap* $(y_{b1}^*, y_{b2}^*, \dots, y_{bn}^*)$;
 - calculamos os quantis empíricos de ordem $\alpha/2$, denotados por q_1 e q_2 , a partir das B réplicas de U_{bn}^* ;
 - verificamos se $U_{n1} < q_1$ ou $U_{1n} > q_2$ e também se $U_{n2} < q_1$ ou $U_{2n} > q_2$;
5. calculamos o número de estatísticas U_{n1} cujos valores estão contidos na região crítica e dividimos pelo número de réplicas de Monte Carlo, assim estimamos o tamanho do teste, em seguida realizamos esse mesmo procedimento para as estatística U_{n2} , nesse caso, estimamos o poder do teste.

Na Tabela 4.7 encontramos os tamanhos do teste quando utilizamos os quantis empíricos via método *bootstrap* paramétrico, podemos notar que os valores obtidos estão bastante próximos dos valores nominais.

Tabela 4.7: Tamanho do teste utilizando quantis *bootstrap*.

$\alpha \backslash \sigma^2$	$\mu \in (0,014; 0,130)$				$\mu \in (0,225; 0,853)$				$\mu \in (0,902; 0,990)$				n
	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	
1%	1,36	1,34	1,16	1,26	1,66	1,66	1,36	1,24	1,24	1,38	1,16	1,12	40
5%	4,92	5,44	5,58	4,72	5,42	5,54	5,26	5,68	5,36	5,94	4,86	4,72	
10%	9,88	9,98	10,16	9,36	10,72	10,12	10,50	10,74	10,22	10,36	9,66	9,70	
1%	1,34	1,36	1,72	1,34	1,42	1,22	1,52	1,56	1,42	1,34	1,38	1,32	80
5%	5,06	5,12	6,00	5,52	5,42	5,84	5,34	5,74	5,10	5,98	5,26	5,38	
10%	9,84	9,72	10,52	10,56	10,48	10,28	10,22	10,76	9,62	11,24	10,66	10,40	
1%	1,28	1,22	1,16	1,44	1,26	1,44	1,36	1,26	1,32	1,24	1,42	1,30	120
5%	5,34	5,24	5,38	5,60	5,28	5,42	5,48	5,42	5,62	5,60	5,40	5,26	
10%	10,26	10,78	10,50	10,66	10,82	10,56	10,42	10,54	10,68	10,48	10,64	9,98	

Ressaltamos que, no caso em que estimamos o poder do teste, estamos calculando a probabilidade de rejeitar a hipótese do ajuste está adequado, quando isso de fato ocorre. Os resultados obtidos encontram-se disponíveis na Tabela 4.8. Nessa tabela observamos que o poder do teste é bem maior quando os valores de μ estão concentrados nos extremos, ou seja, próximos de zero ou um, embora para amostras de tamanho $n = 40$ e $n = 80$ esses poderes não são muito altos, exceto quando $\sigma^2 = 16,0$. Além disso, notamos também que o poder aumenta consideravelmente quando $\alpha = 10\%$.

Quando os valores de μ estão próximos de 0,5, isto é, quando estão concentrados no centro do intervalo (0,1), vemos que mesmo quando o tamanho amostral e o valor de σ^2 são altos e $\alpha = 10\%$, o poder do teste não é tão alto. Sendo assim, para valores de μ centrais as distribuições beta e simplex apresentam comportamentos semelhantes, logo, não será fácil do teste identificar qual modelo é o verdadeiro gerador dos dados.

Tabela 4.8: Poder do teste para o modelo simplex com dispersão fixa, quando os dados são gerados da distribuição beta, utilizando quantis *bootstrap*.

$\alpha \backslash \sigma^2$	$\mu \in (0,014; 0,130)$				$\mu \in (0,225; 0,853)$				$\mu \in (0,902; 0,990)$				n
	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	
1%	9,48	11,34	12,70	24,34	3,36	6,00	15,08	37,36	15,84	16,28	17,50	28,36	40
5%	17,06	19,22	21,88	37,78	11,76	18,16	32,72	52,44	24,86	25,52	27,72	41,38	
10%	23,26	25,92	30,20	48,08	20,42	28,90	42,04	60,34	31,36	32,56	34,78	50,52	
1%	28,44	28,06	27,82	34,92	7,44	7,58	9,24	21,76	42,66	40,84	40,38	44,22	80
5%	45,86	46,18	48,40	59,12	19,98	22,10	29,30	49,30	59,82	59,06	58,26	64,60	
10%	56,94	58,26	60,90	71,22	31,38	35,34	44,02	62,42	68,82	68,78	68,28	75,64	
1%	43,50	42,42	42,70	40,84	11,62	11,68	11,50	15,54	58,72	59,48	58,52	56,00	120
5%	66,40	65,46	65,92	69,64	29,34	30,54	33,28	46,92	76,42	77,16	77,32	78,82	
10%	78,40	77,38	78,20	83,08	42,30	45,90	51,22	63,22	83,98	85,38	86,04	88,12	

Agora, consideraremos outro tipo de erro de especificação: omissão de covariáveis. Considerando o modelo com quatro covariáveis, de modo que

- $x_{t2} \sim U(0,1)$;
- $x_{t3} \sim LN(0,1) \times 0,2$;
- $x_{t4} \sim LN(0,1) \times 0,45$;
- $x_{t5} \sim Ber(0,2)$,

com $t = 1, \dots, 40$ e os valores fixados para os parâmetros do modelo foram os apresentados pela Tabela 4.9. Portanto, temos que nesse caso os valores da variável y foram geradas a partir da distribuição simplex, cuja média, μ_t , está associada às quatro covariáveis, x_{t2} , x_{t3} , x_{t4} e x_{t5} , especificadas anteriormente.

Tabela 4.9: Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando dados da distribuição simplex com 4 covariáveis.

β_1	β_2	β_3	β_4	β_5	μ
-1,5	-1,1	-0,2	-1,1	-1,0	$\mu \in (0,002; 0,139)$
-0,6	0,9	0,9	-0,5	1,7	$\mu \in (0,258; 0,868)$
1,8	1,1	0,3	1,1	0,9	$\mu \in (0,893; 0,998)$

Entretanto, procuramos ajustar alguns modelos de regressão simplex para y , de forma que omitimos algumas dessas covariáveis, consequentemente, baseando-se nos modelos “erroneamente” ajustados, procuramos estimar os parâmetros da distribuição e em seguida

calcular o valor da estatística U_n , para assim, estimar via simulação de Monte Carlo o poder do teste para identificar a omissão dessas covariáveis.

Na Tabela 4.10 observamos que nos extremos, ou seja, para valores da média μ_t próximo de zero ou um, o teste apresenta poderes consideravelmente altos, que vão diminuindo levemente à medida que acrescentamos as covariáveis faltantes, o que é compreensivo, pois dessa forma, o modelo vai tornando-se cada vez mais próximo ao verdadeiro, constituído de quatro covariáveis. Porém, para valores de μ_t concentrados no intervalo $(0,258; 0,868)$ obtemos poderes muito baixos quando ajustamos os modelos com uma e com duas covariáveis. Entretanto, ao ajustar o modelo com três covariáveis notamos um aumento considerável do poder do teste, principalmente para os níveis de significância de 5% e 10%. Isso pode estar ocorrendo talvez devido ao fato que ao incluirmos a terceira covariável, o modelo ajustado estima um cenário com os valores de μ_t muito distantes do verdadeiro cenário ao qual ele de fato pertence, provocando assim, uma drástica mudança detectada pelo teste.

De fato, no cenário central, a associação entre as distribuições de probabilidades das quatro covariáveis e seus respectivos coeficientes causam um cenário bastante complexo. É possível perceber que de fato a segunda covariável associada ao seu coeficiente, apresenta uma influência pequena no modelo, pois tem-se $0,9 \times LN(0,1) \times 0,2 = 0,18 \times LN(0,1)$. Comparando o coeficiente 0,18 com o coeficiente 0,9 da covariável uniforme x_{t2} , percebemos que, provavelmente a combinação de x_{t2} e seu coeficiente gere valores bem próximos dos verdadeiros. O teste só consegue identificar a falta de uma variável quando x_{t4} é incluída no modelo, pois as duas distribuições lognormais juntas criam valores ajustados muito distantes dos verdadeiros.

Também notamos que para valores de μ nos extremos, o poder diminui à medida que o valor da dispersão aumenta, enquanto que valores em torno de 0,5 o poder aumenta quando σ^2 cresce.

O último tipo de má especificação que consideramos foi gerar os dados com dispersão variável e ajustar com dispersão fixa. Dessa forma, geramos valores da variável resposta y a partir do modelo de regressão simplex com dispersão variável, em que a dispersão satisfaz a relação funcional dada por (2.28). Para o modelo da média μ_t , trabalhamos com um conjunto de observação geradas a partir de $x_{ti} \sim U(0, 1)$, com $i = 2, \dots, 5$ e $t = 1, \dots, 40$ e os valores fixados para os parâmetros desse modelos estão na Tabela 4.11.

Tabela 4.11: Valores verdadeiros dos parâmetros β 's, fixados para obter o poder do teste, gerando dados de um modelo simplex com dispersão variável e ajustando o modelo simplex com dispersão fixa.

β_1	β_2	β_3	β_4	β_5	μ
-1,4	-0,7	-0,7	-0,9	-1,0	$\mu \in (0, 019; 0, 144)$
-1,2	1,3	-0,9	1,0	1,0	$\mu \in (0, 223; 0, 796)$
1,6	1,1	1,0	0,9	0,9	$\mu \in (0, 883; 0, 990)$

Tendo como objetivo analisar o poder do teste para observações geradas sob diferentes graus de heterogeneidade, optamos em utilizar cinco diferentes cenários para o modelo da dispersão, de modo que, em cada cenário acrescentamos uma nova covariável que irá provocar um aumento no grau de heterogeneidade e assim gerar valores de y de acordo com cada grau, lembrando que definimos grau de heterogeneidade em (4.6) e representamos essa medida por λ .

Os cenários considerados para a geração dos dados foram:

- $h_1(\sigma_t^2) = \gamma_1$, para $\lambda = 1$;
- $h_2(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2}$, para $\lambda = 25, 86$;
- $h_3(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3}$, para $\lambda = 52, 66$;
- $h_4(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4}$, para $\lambda = 101, 86$;
- $h_5(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, para $\lambda = 145, 37$;

com $t = 1, \dots, 40$. A função de ligação usada nesse estudo foi a logarítmica, ou seja, $h_i(\sigma_t^2) = \log(\sigma_t^2)$, com $i = 1, \dots, 5$.

Como o interesse consiste em gerar dados a partir de um modelo com dispersão variável e ajustar um modelo com dispersão fixa, temos que, os quantis *bootstrap* à serem utilizados devem ser obtidos sob a hipótese que os dados apresentam dispersão constante, para isso, as estimativas dos parâmetros, utilizadas na geração da amostra aleatória no laço *bootstrap*, foram obtidas a partir do primeiro cenário para o modelo para a dispersão.

Para diferentes valores de $\sigma^2 = \exp(\gamma_1)$ são mantidos os mesmos valores para λ . Para isso, é necessário combinar diferentes valores dos γ 's e das covariáveis z_{tj} , com $j = 2, \dots, 5$. Trabalhamos com um conjunto de observações geradas originalmente a partir dos cenários apresentados pela Tabela 4.12.

Tabela 4.12: Valores verdadeiros dos parâmetros γ 's, fixados para obter o poder do teste, gerando dados de um modelo simplex com dispersão variável e ajustando o modelo simplex com dispersão fixa.

Caso 1: $\sigma^2 = 0,5$		Caso 2: $\sigma^2 = 1,5$	
$z_{t2} \sim LN(0,1) - 0,2$	$\gamma_1 = -0,68$	$z_{t2} \sim U(0,1)$	$\gamma_1 = 0,41$
$z_{t3} \sim LN(0,1) - 0,2$	$\gamma_2 = 0,48$	$z_{t3} \sim U(0,1)$	$\gamma_2 = 0,72$
$z_{t4} \sim LN(0,1) - 0,2$	$\gamma_3 = 0,15$	$z_{t4} \sim U(0,1)$	$\gamma_3 = 1,0$
$z_{t5} \sim U(0,1)$	$\gamma_4 = 0,21$	$z_{t5} \sim U(0,1)$	$\gamma_4 = 1,08$
	$\gamma_5 = 0,6$		$\gamma_5 = -0,92$
Caso 3: $\sigma^2 = 4$		Caso 4: $\sigma^2 = 16$	
$z_{t2} \sim N(0,1)$	$\gamma_1 = 1,4$	$z_{t2} \sim LN(0,1)$	$\gamma_1 = 2,78$
$z_{t3} \sim N(0,1)$	$\gamma_2 = 0,72$	$z_{t3} \sim N(0,1)$	$\gamma_2 = -0,48$
$z_{t4} \sim N(0,1)$	$\gamma_3 = -0,62$	$z_{t4} \sim N(0,1)$	$\gamma_3 = 0,77$
$z_{t5} \sim N(0,1)$	$\gamma_4 = 0,55$	$z_{t5} \sim N(0,1)$	$\gamma_4 = -0,301$
	$\gamma_5 = 0,427$		$\gamma_5 = 0,312$

Avaliando os resultados de cada caso, apresentados pela Tabela 4.13, vemos que o teste apresenta maiores poderes para valores de μ_t no intervalo $(0,2230; 0,7963)$ e quando $\sigma^2 = 0,5$. Nesse caso obtemos bons poderes para todos os níveis de significância, graus de heterogeneidade iguais a $\lambda = 101,86$ ou $\lambda = 145,37$, tamanhos amostrais dados por $n = 80$ ou $n = 120$ e os valores de μ concentrados no centro do intervalo $(0,1)$. Também é possível identificar altos poderes, quando a média μ está distribuída nos extremos do intervalo $(0,1)$ e $n = 80$ ou $n = 120$, porém, apenas quando $\alpha = 5\%$ ou $\alpha = 10\%$ e λ igual a 52,66, 101,86 ou 145,37.

Para o segundo caso, não obtivemos poderes satisfatórios em todos os cenários para a média μ , uma justificativa para esse fato é que embora os valores de λ sejam os mesmos que no caso anterior, o uso do segundo cenário da Tabela 4.12 não produz valores altos para $\max_{t=1,\dots,n}(\sigma_t^2)$, enquanto que para $\min_{t=1,\dots,n}(\sigma_t^2)$ produz valores consideravelmente baixos. O mesmo pode estar ocorrendo no terceiro caso, porém, notemos que geralmente os valores dos poderes quando usamos $\sigma^2 = 4,0$ são maiores do que para $\sigma^2 = 1,5$ e em algumas situações chegam a ser altos.

No último caso, em que $\sigma^2 = 16$, temos que, embora baixos, os poderes quando $n = 40$ são maiores do que no caso anterior, mais quando o tamanho amostral aumenta esses poderes tornam-se menores do que quando $\sigma^2 = 4,0$.

Logo, podemos assim dizer, que de maneira geral, o teste consegue detectar a dispersão variável contida nos dados.

Tabela 4.13: Poderes do teste para o modelo simplex com dispersão fixa, quando os dados são gerados com dispersão variável.

Caso 1: $\sigma^2 = 0,5$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,44	15,20	23,76	19,36	18,66	1,06	15,54	25,16	22,88	22,42	1,40	18,00	26,22	22,48	21,74	40
5%	5,34	29,88	39,88	36,52	34,96	4,68	27,60	40,50	41,00	39,86	5,48	33,08	42,52	40,10	38,64	
10%	10,04	40,66	49,96	48,54	46,60	10,06	36,50	49,48	52,38	51,66	10,28	43,08	52,26	51,60	50,66	
1%	1,34	24,86	35,32	36,46	37,02	1,62	31,34	51,38	69,20	73,24	1,42	27,46	38,80	39,60	40,40	80
5%	5,52	47,88	60,26	59,98	60,04	5,90	50,02	67,78	81,90	84,76	5,32	52,08	65,28	63,74	64,34	
10%	10,34	62,08	72,52	72,40	71,90	10,78	60,70	75,82	86,72	88,74	10,34	65,68	77,04	75,34	77,00	
1%	1,40	34,80	45,68	47,00	49,38	1,32	40,50	63,84	86,24	90,84	1,34	38,26	48,50	50,68	51,74	120
5%	5,74	63,00	74,58	74,92	75,88	5,28	64,32	81,14	94,56	96,16	5,20	67,00	76,86	78,40	78,42	
10%	10,58	76,10	86,36	86,02	86,94	10,32	75,24	87,98	96,72	97,38	10,26	80,24	88,18	88,16	88,70	
Caso 2: $\sigma^2 = 1,5$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,36	3,56	4,00	2,50	4,48	1,32	4,48	3,90	4,04	6,46	1,46	4,72	4,60	3,24	5,72	40
5%	5,54	11,04	11,68	8,76	12,82	4,88	10,88	10,60	10,82	14,88	5,30	12,26	12,78	10,48	15,08	
10%	11,08	19,56	18,94	15,46	20,98	9,72	16,92	17,34	17,82	22,20	10,10	20,16	20,62	17,72	23,02	
1%	1,64	7,64	7,90	8,64	10,84	1,38	7,00	12,88	14,62	16,36	1,16	7,66	9,34	10,24	10,72	80
5%	5,50	19,94	20,50	18,94	23,26	5,70	17,70	25,70	27,64	30,92	4,96	20,16	22,28	22,02	24,44	
10%	10,24	29,76	29,94	27,94	34,02	11,02	26,50	35,74	36,36	40,66	9,66	30,16	33,08	31,68	36,06	
1%	1,44	11,26	13,88	12,58	14,90	1,56	9,66	20,96	26,32	28,30	1,68	11,92	14,90	13,24	14,96	120
5%	5,80	27,08	29,50	27,16	32,74	5,54	23,20	37,54	43,66	48,00	5,82	27,08	32,42	29,70	34,14	
10%	10,46	38,82	42,12	38,28	45,18	10,28	34,32	48,98	53,56	58,20	10,70	39,30	45,02	40,98	47,80	
Caso 3: $\sigma^2 = 4,0$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,28	2,60	3,80	3,48	3,92	1,34	3,92	6,28	9,30	10,48	1,46	3,52	4,12	4,10	4,42	40
5%	5,30	8,90	11,76	10,12	12,02	5,22	11,00	15,38	20,18	22,80	5,14	10,86	12,10	11,96	12,88	
10%	10,14	16,46	19,02	17,70	19,62	9,66	17,48	22,34	28,60	32,24	9,74	17,90	19,88	19,48	20,60	
1%	1,32	5,80	7,54	8,06	10,64	1,34	8,40	14,16	26,86	37,72	1,44	5,94	8,60	8,48	14,28	80
5%	5,64	14,70	19,10	18,50	24,56	5,10	19,60	29,56	47,46	58,10	5,22	16,18	21,42	20,16	29,24	
10%	10,42	23,26	29,20	28,14	35,44	10,12	28,40	39,22	57,78	68,26	10,02	25,12	31,88	30,14	41,30	
1%	1,28	9,46	13,02	13,94	20,88	1,28	12,62	20,30	46,08	64,12	1,38	10,20	14,44	13,42	23,20	120
5%	5,56	21,76	28,86	26,84	38,10	5,48	27,54	39,94	68,18	81,02	5,08	22,66	31,42	29,56	42,94	
10%	10,26	31,96	41,68	38,04	50,58	9,80	39,08	51,48	76,64	87,22	9,80	33,88	44,24	41,74	55,80	
Caso 4: $\sigma^2 = 16$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,12	2,32	3,74	5,82	6,60	1,40	3,46	12,10	15,80	19,16	1,34	2,22	3,46	5,36	5,20	40
5%	4,52	7,10	10,40	13,46	15,02	5,64	9,42	22,02	28,98	33,62	4,76	6,92	10,12	13,30	14,02	
10%	8,90	12,10	16,54	20,72	22,68	10,68	15,22	28,86	37,44	41,56	9,20	11,24	16,48	19,52	21,40	
1%	1,44	2,02	3,90	7,58	7,72	1,44	4,46	24,14	34,44	38,48	1,34	1,48	3,80	6,64	8,10	80
5%	4,80	7,28	12,16	17,92	19,44	5,48	12,12	43,02	54,66	59,80	5,60	5,74	10,80	16,84	19,16	
10%	10,36	12,88	19,98	27,50	29,66	10,32	19,84	53,76	66,02	69,78	10,66	11,40	17,88	27,16	30,04	
1%	1,26	1,96	4,96	7,98	10,54	1,38	5,76	32,42	46,80	54,04	1,72	1,56	4,22	8,26	10,84	120
5%	5,52	6,74	13,84	20,22	24,18	5,20	15,26	55,76	71,04	75,90	5,38	5,64	12,76	20,72	25,82	
10%	10,36	12,52	23,36	30,66	35,76	10,24	24,30	67,48	80,18	84,44	10,44	11,06	21,22	31,24	37,92	

4.1.2 Avaliação do Teste Usando Variância *Bootstrap*

Na definição do teste de diagnóstico proposto por Zhu e Zhang (2004) no terceiro capítulo, vimos que a estatística U_n é obtida a partir da estatística $T_n(\hat{\theta})$ e de sua variância, denotada por σ_*^2 . Entretanto, sabemos que σ_*^2 também corresponde a variância da quantidade $k(y_t; \theta_*)$ introduzida pelos autores e definida por (3.6), notamos que $k(y_t; \theta_*)$ depende de duas outras quantidades, que são $J_1(\theta_*)$ e $h(y_t; \theta_*)$, definidas, respectivamente, por (3.5) e (3.7).

Assim, antes de estimar σ_*^2 , devemos estimar $k(y_t; \theta_*)$. Nesta seção, propomos uma variação no teste obtendo a variância de $T_n(\hat{\theta})$ via método *bootstrap*.

Nesse caso, objetivando avaliar o desempenho do teste quando a estatística U_n é obtida a partir da estimativa *bootstrap* de σ_*^2 , desenvolvemos os mesmos estudo de tamanhos e poderes antes realizado. Considerando os valores presentes na Tabela 4.14 para os parâmetros do modelo e gerando $x_{ti} \sim U(0, 1)$, com $t = 1, \dots, 40$ e $i = 2, \dots, 5$, calculamos os tamanhos do teste, cujos resultados encontram-se na Tabela 4.15. Utilizando os mesmos valores da Tabela 4.14, calculamos o poder do teste quando geramos observações de uma distribuição beta e ajustamos um modelo simplex. Fixamos $\phi = 15$ e apresentamos na Tabela 4.16 os resultados obtidos.

Tabela 4.14: Valores verdadeiros dos parâmetros β 's, fixados para calcular o poder do teste, gerando beta, usando a variância *bootstrap*.

β_1	β_2	β_3	β_4	β_5	μ
-1,4	-0,7	-0,7	-0,9	-1,0	$\mu \in (0, 019; 0, 144)$
-1,0	1,0	-1,0	1,1	1,5	$\mu \in (0, 225; 0, 853)$
1,6	1,1	1,0	0,9	0,9	$\mu \in (0, 883; 0, 990)$

Tabela 4.15: Tamanho do teste para o modelo simplex com dispersão fixa usando a variância *bootstrap*.

		$\mu \in (0, 019; 0, 144)$				$\mu \in (0, 225; 0, 853)$				$\mu \in (0, 883; 0, 990)$				
α	σ^2	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	n
1%		1,42	1,54	1,48	1,28	1,18	1,34	1,26	1,10	1,58	1,76	1,54	1,34	40
5%		5,30	5,76	5,90	5,80	5,14	5,56	5,50	5,02	5,40	5,92	6,00	5,74	
10%		10,54	11,02	11,52	11,38	10,52	10,50	10,46	9,96	10,20	11,18	11,32	11,08	
1%		1,06	1,42	1,72	1,62	1,46	1,72	1,70	1,10	1,42	1,56	1,28	1,72	80
5%		5,08	5,74	6,12	6,14	5,84	5,92	5,68	5,18	4,88	5,32	6,18	6,20	
10%		9,58	10,74	11,10	10,92	10,88	10,86	11,02	10,04	10,38	10,92	11,24	11,82	
1%		1,30	1,48	1,46	1,48	1,18	1,30	1,38	1,30	2,00	2,40	2,10	1,48	120
5%		5,76	5,54	5,42	5,54	5,14	5,24	5,64	5,32	5,72	5,90	6,70	5,60	
10%		10,36	10,82	10,40	11,48	10,18	10,32	11,10	10,58	10,78	10,90	11,60	11,20	

Da Tabela 4.15 temos que os resultados apresentados encontram-se bem próximos dos valores nominais dos níveis de significância. Pelos resultados da Tabela 4.16 notamos que os poderes obtidos pela variância *bootstrap* são consideravelmente baixos, principalmente quando $\mu \in (0, 225; 0, 853)$. Entretanto, não é possível concluir que esse método não seja eficiente, pois vale lembrar que utilizamos $\phi = 15$, isso significa que, os dados usados para o método proposto nessa seção foram gerados por uma distribuição beta com alta dispersão.

Tabela 4.16: Poder do teste para o modelo simplex com dispersão fixa usando a variância *bootstrap*, quando os dados são gerados da distribuição beta.

$\alpha \backslash \sigma^2$	$\mu \in (0, 019; 0, 144)$				$\mu \in (0, 225; 0, 853)$				$\mu \in (0, 883; 0, 990)$				n
	0, 5	1, 5	4, 0	16, 0	0, 5	1, 5	4, 0	16, 0	0, 5	1, 5	4, 0	16, 0	
1%	4,02	4,46	4,94	6,96	0,22	0,28	0,56	3,38	8,92	8,16	8,56	10,68	40
5%	8,80	9,16	10,14	15,24	0,68	1,04	3,00	43,46	15,88	15,58	15,76	18,90	
10%	13,58	13,68	14,90	23,16	1,16	2,40	9,38	90,06	21,76	21,70	22,22	26,24	
1%	25,14	25,62	26,80	31,40	0,10	0,10	0,12	0,32	36,78	36,62	36,08	37,58	80
5%	36,60	37,26	40,10	45,82	0,26	0,30	0,46	1,66	46,72	47,32	46,74	49,04	
10%	43,62	44,66	48,02	55,54	0,42	0,58	1,22	7,38	53,00	53,64	53,50	56,62	
1%	50,56	50,90	51,58	55,86	0,08	0,08	0,10	0,14	48,56	48,30	48,60	49,66	120
5%	63,00	63,72	65,52	70,50	0,16	0,18	0,24	0,60	58,38	59,70	59,60	59,96	
10%	69,98	70,48	73,22	78,78	0,24	0,38	0,64	1,16	64,12	65,30	65,90	65,68	

Podemos observar que para $n = 40$, o teste apresentou um poder bem acima do esperado para $\mu \in (0, 225; 0, 853)$ e $\sigma^2 = 16$. Isso poder ter ocorrido devido ao fato da amostra gerada pela distribuição beta apresentar dispersão muito baixa caracterizando bem a verdadeira distribuição a qual pertence.

Procurando avaliar o poder do teste em detectar a omissão de covariáveis, quando trabalhamos com a variância *bootstrap*, consideramos os mesmos cenários aplicados para a variância analítica associada aos quantis *bootstrap*. Na Tabela 4.17 encontramos os poderes obtidos nessa situação. Podemos observar que para valores de μ_t concentrados no intervalo $(0, 258; 0, 868)$ o teste com variância *bootstrap* apresenta uma leve melhora, mas ainda não apresenta bons resultados. Por outro lado, para o caso em que apenas uma covariável é omitida, o uso da variância analítica e quantis *bootstrap* apresenta melhores poderes. Notemos também que ainda sob esse cenário de μ_t , os poderes do teste usando a variância *bootstrap* comportam-se de modo inverso aos poderes obtidos com a variância analítica, ou seja, eles vão diminuindo à medida que acrescentamos as covariáveis faltantes. Nas situações em que os valores da média μ_t estão próximos de zero ou de um, o teste utilizando a variância analítica e quantis *bootstrap* apresenta maiores poderes para identificar a omissão de três e duas covariáveis, enquanto que para o caso em que apenas uma covariável é omitida, é mais indicado o uso do método com a variância *bootstrap*.

Tabela 4.17: Tamanhos e poderes do teste em detectar covariáveis omitidas, cujo modelo verdadeiro é dado por $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, com $t = 1, \dots, n$ e dispersão fixa, usando a variância *bootstrap*.

Modelo Ajustado	$g(\mu_t) = \beta_1 + \beta_2 x_{t2}$														
	$\mu \in (0, 002; 0, 139)$					$\mu \in (0, 258; 0, 868)$					$\mu \in (0, 893; 0, 998)$				
σ^2	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0
1%	1,18	1,68	1,28	1,72	1,22	1,28	1,14	0,86	1,20	1,32	1,54	1,66	1,00,00	100,00	99,74
5%	5,06	5,82	5,30	5,66	5,38	5,20	4,74	4,18	5,38	5,68	5,66	5,80	100,00	100,00	100,00
10%	9,56	10,54	10,20	10,68	10,66	9,88	9,68	8,74	10,56	10,86	11,04	11,00	100,00	100,00	100,00
1%	1,30	1,58	1,32	1,14	1,26	1,16	1,02	1,12	1,46	1,44	1,18	1,22	100,00	100,00	100,00
5%	5,40	5,64	5,50	4,76	5,60	5,26	5,56	4,64	5,40	5,78	4,82	5,24	100,00	100,00	100,00
10%	9,92	10,90	9,72	9,42	10,32	10,20	10,48	9,70	10,42	10,22	9,54	10,00	100,00	100,00	100,00
1%	1,34	1,30	1,38	1,52	1,66	1,36	1,34	1,10	1,42	1,42	1,30	1,54	100,00	100,00	100,00
5%	5,88	5,38	5,92	5,36	5,64	5,30	5,18	5,18	5,14	5,82	5,22	5,42	100,00	100,00	100,00
10%	11,00	10,92	10,92	9,66	10,74	10,22	10,06	10,14	10,32	10,32	9,90	10,62	100,00	100,00	100,00
Modelo Ajustado	$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$														
	$\mu \in (0, 002; 0, 139)$					$\mu \in (0, 258; 0, 868)$					$\mu \in (0, 893; 0, 998)$				
σ^2	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0	16,0	0,5	1,5	4,0
1%	100,00	99,92	79,08	4,96	6,04	2,16	1,92	3,50	100,00	100,00	93,00	9,22	100,00	99,50	66,02
5%	100,00	100,00	98,20	28,90	13,90	6,88	6,96	14,48	100,00	100,00	99,86	42,58	100,00	99,98	95,54
10%	100,00	100,00	99,72	50,80	20,54	12,02	12,64	26,44	100,00	100,00	100,00	65,88	100,00	100,00	99,36
1%	100,00	100,00	91,26	5,98	16,50	2,70	2,16	5,92	100,00	100,00	97,78	10,84	100,00	100,00	79,14
5%	100,00	100,00	99,50	24,20	27,60	8,10	6,70	17,70	100,00	100,00	99,94	36,12	100,00	100,00	97,66
10%	100,00	100,00	99,96	41,08	35,80	13,66	11,46	29,28	100,00	100,00	100,00	55,06	100,00	100,00	99,50
1%	100,00	100,00	97,82	7,40	24,46	3,08	1,82	4,66	100,00	100,00	99,76	12,66	100,00	100,00	92,08
5%	100,00	100,00	99,94	26,34	37,90	9,44	5,64	15,54	100,00	100,00	100,00	39,96	100,00	100,00	99,42
10%	100,00	100,00	100,00	44,26	46,40	15,46	10,12	27,48	100,00	100,00	100,00	58,12	100,00	100,00	99,84
1%	100,00	100,00	97,82	7,40	24,46	3,08	1,82	4,66	100,00	100,00	99,76	12,66	100,00	100,00	92,08
5%	100,00	100,00	99,94	26,34	37,90	9,44	5,64	15,54	100,00	100,00	100,00	39,96	100,00	100,00	99,42
10%	100,00	100,00	100,00	44,26	46,40	15,46	10,12	27,48	100,00	100,00	100,00	58,12	100,00	100,00	99,84
1%	100,00	100,00	97,82	7,40	24,46	3,08	1,82	4,66	100,00	100,00	99,76	12,66	100,00	100,00	92,08
5%	100,00	100,00	99,94	26,34	37,90	9,44	5,64	15,54	100,00	100,00	100,00	39,96	100,00	100,00	99,42
10%	100,00	100,00	100,00	44,26	46,40	15,46	10,12	27,48	100,00	100,00	100,00	58,12	100,00	100,00	99,84
1%	100,00	100,00	97,82	7,40	24,46	3,08	1,82	4,66	100,00	100,00	99,76	12,66	100,00	100,00	92,08
5%	100,00	100,00	99,94	26,34	37,90	9,44	5,64	15,54	100,00	100,00	100,00	39,96	100,00	100,00	99,42
10%	100,00	100,00	100,00	44,26	46,40	15,46	10,12	27,48	100,00	100,00	100,00	58,12	100,00	100,00	99,84

Finalmente, avaliamos o poder do teste em detectar a má especificação do modelo, quando ajustamos um modelo de regressão simplex com dispersão fixa, quando de fato a dispersão é variável. Os resultados estão apresentados na Tabela 4.18.

Podemos verificar que, para os casos em que σ^2 assume os valores 0, 5, 1, 5 e 4, 0, os poderes do teste com variância *bootstrap* são maiores do que utilizando a variância analítica e quantis *bootstrap*. No entanto, o comportamento dos valores desses poderes em relação a cada caso trabalhado é o mesmo apresentado pela variância analítica.

Portanto, concluímos que, embora em alguns casos, o teste com variância *bootstrap* não aparenta ser tão bom quanto o teste com a variância calculada analiticamente, existem várias outras situações em que o teste utilizando a variância *bootstrap* apresenta um melhor desempenho, logo, o uso desse método torna-se recomendável.

Tabela 4.18: Poderes do teste para o modelo simplex com dispersão fixa usando a variância *bootstrap*, quando os dados são gerados com dispersão variável.

Caso 1: $\sigma^2 = 0,5$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,66	26,00	36,28	39,22	41,74	1,44	23,02	35,04	42,12	44,38	1,60	25,12	36,68	40,58	41,48	40
5%	5,66	37,70	47,76	50,30	53,24	5,14	33,96	45,42	54,48	56,90	5,54	37,38	48,12	52,16	53,58	
10%	10,28	45,58	54,78	57,30	59,22	10,70	40,58	51,86	60,96	63,12	10,72	45,66	55,12	58,30	59,72	
1%	1,34	50,22	63,82	68,24	70,20	1,32	48,00	64,60	78,68	81,46	1,54	52,34	66,56	70,46	73,10	80
5%	5,38	64,44	74,80	77,78	79,06	5,84	59,72	73,84	86,74	87,90	5,56	65,30	77,64	79,68	81,48	
10%	10,50	71,52	80,52	82,12	83,84	10,76	66,34	78,46	89,72	90,78	10,98	72,62	82,58	83,92	85,44	
1%	1,56	68,02	80,36	84,50	85,38	1,30	65,76	81,02	91,98	93,70	1,24	71,58	81,94	86,14	87,32	120
5%	5,22	79,72	87,66	90,38	91,28	5,56	76,74	87,56	95,50	97,02	4,80	82,32	89,14	91,82	91,94	
10%	10,32	84,66	90,68	92,84	93,44	10,28	81,04	90,50	96,78	97,92	10,24	86,52	92,42	94,26	93,94	
Caso 2: $\sigma^2 = 1,5$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,14	5,92	7,10	7,22	9,68	1,32	5,00	6,62	9,24	10,28	1,40	5,96	7,26	8,02	9,60	40
5%	4,96	14,52	16,06	14,86	17,88	5,14	12,22	14,62	18,58	19,00	5,46	14,10	15,52	16,56	18,40	
10%	9,98	22,54	23,26	22,26	25,54	10,46	19,36	22,12	26,66	25,88	10,62	21,42	23,36	23,78	26,04	
1%	1,34	12,22	13,74	14,12	18,98	1,40	9,72	15,06	19,32	24,12	1,60	13,06	17,40	17,02	22,64	80
5%	5,66	25,32	26,54	25,72	32,48	5,46	19,62	27,54	32,70	36,58	6,30	26,02	30,30	29,30	36,26	
10%	10,80	34,58	34,80	34,26	41,40	10,88	28,20	37,26	41,38	44,84	11,28	35,70	39,30	37,40	44,84	
1%	1,30	16,38	21,12	21,22	27,94	1,38	14,98	24,90	30,10	35,58	1,24	20,44	25,14	26,44	31,86	120
5%	4,94	32,04	36,78	34,70	43,04	5,40	28,94	39,94	46,12	51,88	5,34	35,98	40,78	40,94	47,32	
10%	9,94	42,56	46,24	43,98	52,72	10,26	37,62	50,44	55,00	60,96	10,54	46,60	50,74	50,12	56,12	
Caso 3: $\sigma^2 = 4,0$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,20	4,88	7,06	8,12	9,80	0,94	4,18	7,32	11,98	14,84	1,36	5,96	7,82	9,04	12,08	40
5%	5,34	12,70	16,00	16,50	19,74	5,24	11,22	16,00	22,50	26,70	5,48	13,70	17,58	18,12	23,38	
10%	10,44	19,58	23,52	23,70	27,40	10,52	17,34	23,42	30,46	35,52	10,38	20,62	26,02	25,46	31,28	
1%	1,62	8,48	13,04	16,14	21,54	01,18	9,92	17,34	33,46	44,04	1,58	10,06	15,00	17,46	24,72	80
5%	5,98	19,42	25,70	28,60	35,46	5,24	21,70	31,62	49,94	60,42	5,40	20,82	29,02	30,04	39,38	
10%	10,92	28,18	35,16	37,76	43,96	10,60	30,26	40,90	59,84	69,26	10,28	30,10	39,46	38,24	47,58	
1%	1,52	11,30	18,48	23,08	31,22	1,04	15,62	25,80	50,84	63,66	1,46	14,28	22,96	24,88	37,32	120
5%	5,20	23,44	34,22	37,36	46,80	5,28	30,44	43,52	69,42	80,56	5,64	27,94	39,36	39,78	54,62	
10%	10,70	32,24	44,36	46,44	56,34	10,02	40,74	53,34	77,40	86,64	10,58	38,78	50,28	49,68	63,86	
Caso 4: $\sigma^2 = 16$																
	$\mu \in (0, 0194; 0, 1446)$					$\mu \in (0, 2230; 0, 7963)$					$\mu \in (0, 8831; 0, 9903)$					
$\alpha \backslash \lambda$	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	1, 0	25, 86	52, 66	101, 86	145, 37	n
1%	1,44	1,42	3,20	4,68	5,22	1,46	2,92	10,84	15,40	16,32	1,52	1,20	3,22	4,92	5,80	40
5%	5,54	5,48	9,44	11,82	13,24	5,52	8,50	23,32	30,02	32,20	5,92	4,80	9,02	11,40	13,14	
10%	11,06	10,16	14,90	18,66	19,66	11,00	15,00	31,90	40,16	42,92	11,76	9,46	15,02	17,72	20,38	
1%	1,18	1,50	6,14	9,10	11,74	1,38	4,82	23,90	33,24	39,08	1,18	1,48	5,62	8,76	11,86	80
5%	5,58	5,74	13,72	19,02	23,90	5,78	12,86	41,62	53,52	59,36	5,26	5,48	12,52	19,16	23,52	
10%	11,04	11,74	21,04	27,32	33,08	10,58	20,10	52,32	64,60	69,56	10,68	10,24	19,54	27,42	32,70	
1%	1,38	1,80	7,42	13,54	15,84	1,26	5,66	34,78	48,98	56,36	1,14	1,94	7,74	14,40	17,28	120
5%	5,34	6,82	17,00	26,32	30,40	5,66	15,06	55,52	69,36	75,44	5,22	6,42	16,94	26,82	32,16	
10%	11,38	12,02	25,44	36,74	39,98	10,76	24,36	66,62	79,08	83,54	10,18	12,38	24,82	36,02	42,72	

4.2 Avaliação do Teste para o Modelo Simplex com Dispersão Variável

No terceiro capítulo, também desenvolvemos o teste de diagnóstico para o modelo de regressão simplex quando o parâmetro de dispersão é variável. Nesse caso, também, procuramos avaliar esse teste por meio de simulações de Monte Carlo, com $R_{MC} = 5000$ réplicas. A função de ligação assumida para o parâmetro de dispersão foi a logarítmica, dada por

$$h(\sigma_t^2) = \log(\sigma_t^2). \quad (4.7)$$

Logo,

$$h'(\sigma_t^2) = \frac{1}{\sigma_t^2}, \quad h''(\sigma_t^2) = -\frac{1}{(\sigma_t^2)^2} \quad \text{e} \quad h'''(\sigma_t^2) = \frac{2}{(\sigma_t^2)^3}.$$

Assim, seja $y_1, y_2, \dots, y_n$ uma amostra de n variáveis aleatórias independentes, cuja média e dispersão de y_t , com $t = 1, \dots, n$, satisfazem, respectivamente, as relações (4.1) e (4.7), de forma que

$$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5} \quad (4.8)$$

e

$$h(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}, \quad (4.9)$$

temos que $p = q = 5$. Ou seja, nesse estudo foram consideradas quatro covariáveis associadas a média da variável resposta e outras quatro associadas a dispersão dessa variável.

A princípio temos o interesse de analisar a distribuição empírica da estatística U_n para o modelo simplex com dispersão variável. Dessa forma, geramos 1000 valores dessa estatística, por meio de uma breve simulação de Monte Carlo, em que assumimos $n = 120$, fixamos os valores presentes na Tabela 4.19 para parâmetros do modelo para a média e os valores da Tabela 4.20 para os parâmetros do modelo para a dispersão e geramos $x_{ti} \sim U(0, 1)$, $z_{tj} \sim LN(0, 1) - 0, 2$ e $z_{t5} \sim U(0, 1)$, com $t = 1, \dots, 40$, $i = 2, \dots, 5$ e $j = 2, \dots, 5$.

Tabela 4.19: Valores verdadeiros dos parâmetros β 's, fixados para obter os histogramas dos valores da estatística U_n , para o modelo com dispersão variável.

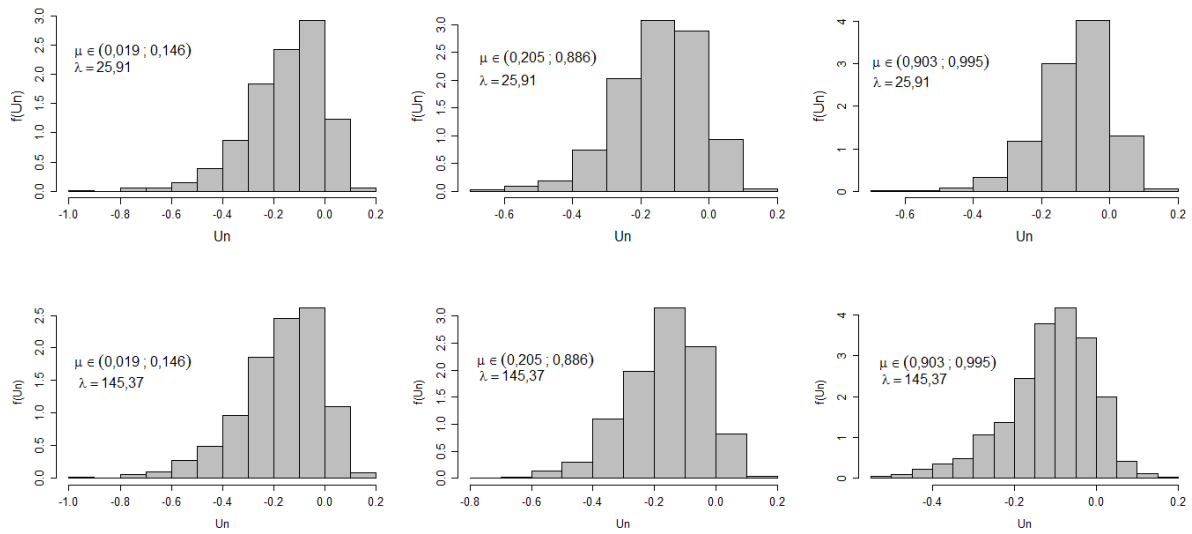
β_1	β_2	β_3	β_4	β_5	μ
-1,4	-0,7	-0,4	-0,9	-1,5	$\mu \in (0, 0140; 0, 1303)$
-1,0	1,0	-1,0	1,1	1,5	$\mu \in (0, 225; 0, 853)$
1,8	1,1	0,3	1,1	0,9	$\mu \in (0, 901; 0, 9902)$

Tabela 4.20: Valores verdadeiros dos parâmetros γ 's, fixados para obter os histogramas dos valores da estatística U_n , para o modelo com dispersão variável.

γ_1	γ_2	γ_3	γ_4	γ_5	λ
-0,6	0,04	0,3	0,2	0,03	25,91
-0,68	0,48	0,15	0,21	0,6	145,37

Ressaltamos que foram estabelecidos apenas dois graus de heterogeneidade, $\lambda = 25,91$ e $\lambda = 145,37$, pois consideramos o suficiente para observar como a distribuição de U_n se comporta quando os dados possuem baixo e alto grau de heterogeneidade. Assim, apresentamos na Figura 4.11 os histogramas contruídos para U_n . É possível observar que os valores que a estatística assume, para o caso em que os dados apresentam dispersão variável, estão distribuídos sempre dentro do intervalo $(-1,0; 0,2)$, independente do grau de heterogeneidade e do cenário ao qual a média μ pertence. Também vemos que a distribuição dessa estatística tende a ser assimétrica.

Figura 4.11: Histogramas da estatística U_n para diferentes cenários de μ e graus de heterogeneidade iguais a 25,91 e 145,37, respectivamente.

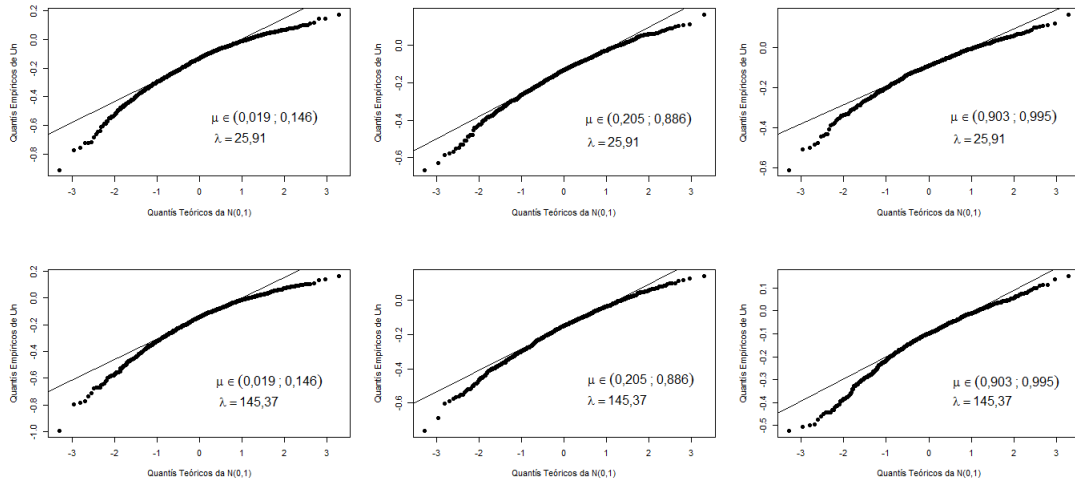


Fonte: Autoria própria.

Na Figura 4.12 encontramos os gráficos de *qqplot*'s. Podemos ver que, para os quantis abaixo de zero os pontos encontram-se mais distantes da reta *qqpline* do que para os quantis acima de zero, isso reforça a conclusão de que a estatística U_n possui uma distribuição bastante assimétrica.

Analisando mais a fundo, podemos dizer que para dados simplex, com dispersão variável, a distribuição da estatística U_n difere em muitos aspectos de uma normal com média zero e desvio padrão um, por esse motivo, continuaremos utilizando os quantis empíricos, obtidos via método *bootstrap* paramétrico.

Figura 4.12: Gráficos *qqplot*'s da distribuição de U_n para diferentes cenários de μ e graus de heterogeneidade iguais a 25,91 e 145,37, respectivamente.



Fonte: Autoria própria.

Procuramos avaliar o poder do teste, de forma análoga ao método de simulação utilizado na seção anterior, em que o parâmetro de dispersão era fixo.

Então, seja $y_1, y_2, \dots, y_n$ uma amostra aleatória gerados a partir de n variáveis aleatórias independentes, com distribuição beta de média μ_t e parâmetro de dispersão ϕ_t (notemos que agora geramos dados da distribuição beta com precisão variável), iremos em seguida, calcular a estatística U_n como se os dados fossem provenientes de uma distribuição simplex de parâmetros μ_t e σ_t^2 . Assim, pretendemos estimar o poder do teste, calculando a probabilidade de rejeitar a hipótese do ajuste ser adequado, dado que o ajuste não está correto.

Os valores que foram fixados para os β 's são os mesmos apresentados pela Tabela 4.19 e os valores fixados para o modelo da dispersão com seus respectivos graus de heterogeneidade são apresentados pela Tabela 4.21 em que utilizamos $x_{ti} \sim U(0, 1)$, $z_{tj} \sim LN(0, 1) - 0, 2$ e $z_{t5} \sim U(0, 1)$, com $t = 1, \dots, 40$, $i = 2, \dots, 5$ e $j = 2, \dots, 4$.

Tabela 4.21: Valores verdadeiros dos parâmetros γ 's, fixados para calcular o poder do teste, gerando dados da distribuição beta.

γ_1	γ_2	γ_3	γ_4	γ_5	λ
-0,6	0,04	0,3	0,2	0,03	25,91
-0,68	0,48	0,15	0,021	0,03	52,72
-0,8	0,02	0,15	0,4	0,1	101,85
-0,68	0,48	0,15	0,21	0,6	145,37

Encontramos na Tabela 4.22 as probabilidades do teste rejeitar a hipótese de que o modelo de regressão simplex com dispersão variável está bem ajustado à um conjunto de dados com distribuição beta com dispersão variável. Podemos ver nesse caso, que os poderes a partir de $n = 80$, são bastante altos para todos os graus de heterogeneidade e para todos os cenários para a média, inclusive para $\mu \in (0, 225; 0, 853)$.

Tabela 4.22: Poder do teste para o modelo simplex com dispersão variável, quando os dados são gerados apresentam distribuição beta com dispersão variável.

$\lambda \backslash \alpha$	$\mu \in (0, 014; 0, 130)$				$\mu \in (0, 225; 0, 853)$				$\mu \in (0, 902; 0, 990)$				n
	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	
1%	19,80	19,26	18,66	13,84	16,14	16,02	15,14	6,14	26,18	23,80	23,20	6,82	40
5%	42,40	48,08	38,82	46,28	37,04	40,60	32,14	21,44	48,50	50,04	42,38	25,48	
10%	54,52	63,02	48,56	63,44	49,04	55,88	42,30	44,40	58,92	62,46	50,92	45,32	
1%	76,86	75,02	79,38	64,54	62,72	63,48	71,56	43,38	85,80	85,90	85,88	70,56	80
5%	89,00	88,82	90,18	82,52	74,64	76,58	81,50	57,28	93,64	93,72	93,84	86,30	
10%	92,94	92,74	93,60	88,16	80,02	82,36	85,80	66,40	95,92	96,12	95,94	91,62	
1%	92,96	92,20	95,38	82,28	79,30	81,70	89,34	61,40	96,84	97,06	98,04	89,08	120
5%	97,66	96,74	98,46	92,54	87,22	89,90	94,16	74,26	99,10	98,96	99,26	96,64	
10%	98,88	98,34	99,14	95,80	90,24	92,82	95,86	80,02	99,58	99,50	99,54	98,28	

Prosseguindo com a avaliação do teste, vamos agora calcular o poder do teste em detectar a omissão de covariáveis do modelo para a média. Considerando os modelos para μ_t e para σ_t^2 , ambos com quatro covariáveis, de modo que

- $x_{t2} \sim U(0, 1)$;
- $x_{t3} \sim Exp(8) \times 0, 2$;
- $x_{t4} \sim Exp(3) \times 0, 45$;
- $x_{t5} \sim Ber(0, 2)$;
- $z_{t2} \sim LN(0, 1) - 0, 2$;
- $z_{t3} \sim LN(0, 1) - 0, 2$;
- $z_{t4} \sim LN(0, 1) - 0, 2$;
- $z_{t5} \sim U(0, 1)$,

com $t = 1, \dots, 40$ e os valores fixados para os parâmetros de ambos os modelos são apresentados pela Tabela 4.23. Nesse caso, os valores da variável y foram geradas a partir da distribuição simplex, cuja média, μ_t , está associada às quatro covariáveis, x_{t2} , x_{t3} , x_{t4} e x_{t5} e a dispersão σ_t^2 associada às covariáveis, z_{t2} , z_{t3} , z_{t4} e z_{t5} , todas especificadas anteriormente.

Procuramos ajustar alguns modelos de regressão simplex para y , considerando o verdadeiro modelo da dispersão, porém omitindo algumas das covariáveis do modelo para a média. Trabalhamos com o poder do teste para identificar a falta dessas covariáveis, fixando diversas combinações para os parâmetros γ_1 , γ_2 , γ_3 , γ_4 , γ_5 e assim avaliar o desempenho do teste para diversos graus de heterogeneidade. Os resultados encontram-se distribuídos na Tabela 4.24

Tabela 4.23: Valores verdadeiros dos parâmetros β 's e γ 's, fixados para calcular o poder do teste, gerando pela distribuição simplex com 4 covariáveis.

β_1	β_2	β_3	β_4	β_5	μ
-1,0	-1,1	-1,2	-1,1	-3,0	$\mu \in (0,006; 0,219)$
-0,35	-0,8	3,5	-0,5	2,3	$\mu \in (0,249; 0,854)$
2,5	1,0	1,0	-1,0	1,8	$\mu \in (0,881; 0,994)$
γ_1	γ_2	γ_3	γ_4	γ_5	λ
-0,6	0,04	0,3	0,2	0,03	25,91
-0,68	0,48	0,15	0,021	0,03	52,72
-0,8	0,02	0,15	0,4	0,1	101,85
-0,68	0,48	0,15	0,21	0,6	145,37

A partir da Tabela 4.24, vemos que o teste apresenta bons poderes em todos os casos. Para $\mu \in (0,249; 0,854)$ quando o tamanho da amostra cresce, o poder aumenta e encontramos maiores valores de poderes para $\lambda = 25,91$ e $\lambda = 101,85$.

Quando $\mu \in (0,006; 0,219)$ e o tamanho da amostra cresce vemos que, se $\alpha = 1\%$, o poder diminui rapidamente. Se $\alpha = 5\%$, o poder também diminui, porém, se o cenário observado for o modelo ajustado apenas com uma covariável, temos que, para $\lambda = 25,91$ e $\lambda = 101,85$ o poder aumenta. Se $\alpha = 10\%$, o poder cresce quando o modelo ajustado contém uma covariável ou duas covariáveis. Porém apenas para os três primeiros graus de heterogeneidade.

Para $\mu \in (0,881; 0,994)$ vemos que, quando o tamanho da amostra cresce, se $\alpha = 1\%$, assim como, quando μ encontra-se concentrado próximo de zero, o poder do teste tende a diminuir de modo rápido, enquanto que se $\alpha = 5\%$ o poder tende a aumentar para $\lambda = 25,91$ ou $\lambda = 101,85$. Para $\alpha = 10\%$ o poder também aumenta, com exceção de quando ajustamos o modelo com quatro covariáveis e o grau de heterogeneidade é consideravelmente grande.

Diversas situações podem ser consideradas para verificar o desempenho do teste para o modelo simplex com dispersão variável, como por exemplo, calcular o poder do teste quando omitimos covariáveis do modelo para a dispersão. Nas situações estudadas, podemos concluir que quando trabalhamos com os níveis de significância de 5% e 10% o teste aqui proposto pode ser considerado bastante eficaz em verificar a qualidade do ajuste do modelo simplex realizado.

Tabela 4.24: Poderes do teste em detectar covariáveis omitidas do modelo para a média, quando o modelo verdadeiro é dado por $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $h(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$, com $t = 1, \dots, n$.

Modelo Ajustado	$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$															$g(\mu_t) = \beta_1 + \beta_2 x_{t2}$														
	$\mu \in (0, 0,006; 0, 219)$					$\mu \in (0, 0,881; 0, 994)$					$\mu \in (0, 0,006; 0, 219)$					$\mu \in (0, 0,249; 0, 854)$					$\mu \in (0, 0,881; 0, 994)$									
$\frac{\lambda}{\alpha}$	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	n	
1%	1,82	1,38	1,28	1,64	1,44	1,82	1,58	1,38	1,46	1,64	1,94	1,44	56,10	68,16	66,96	52,00	19,14	17,00	20,14	11,04	48,88	53,76	55,78	48,68	40					
5%	5,74	5,92	5,64	6,00	4,84	6,12	5,64	5,74	5,92	5,80	6,60	5,28	93,50	95,96	94,02	87,72	41,24	34,84	42,50	23,38	84,28	88,14	84,10	79,48	80					
10%	10,34	10,72	9,98	10,74	9,18	11,96	10,50	10,88	10,78	11,30	12,10	10,24	98,52	98,82	97,98	94,28	52,26	44,42	53,26	31,26	93,10	94,96	92,26	88,70	120					
1%	1,52	1,18	1,20	1,92	1,32	1,50	1,46	1,34	1,20	1,20	1,38	1,48	19,20	14,90	21,44	10,54	25,34	17,84	27,46	17,02	26,82	17,58	32,12	16,28	80					
5%	5,28	4,86	5,22	5,78	4,78	5,64	5,54	4,88	4,66	4,76	5,82	5,98	95,62	93,36	94,68	80,66	59,38	49,34	60,76	42,72	91,92	87,74	89,40	78,04	120					
10%	9,98	10,06	9,82	10,32	9,62	10,80	10,50	9,86	9,74	9,82	11,08	11,10	99,90	99,68	99,62	96,78	74,00	65,32	76,02	56,94	98,50	98,68	97,76	94,74	80					
1%	1,36	1,20	1,50	1,28	1,32	1,22	1,26	1,12	1,24	1,16	1,44	1,34	21,14	14,84	20,22	7,94	27,24	16,68	27,92	18,02	27,10	15,98	29,60	12,64	120					
5%	4,76	5,92	5,68	5,16	4,96	5,16	4,96	4,98	5,28	5,26	5,06	5,32	96,90	94,80	95,78	80,82	67,00	54,18	66,70	48,62	92,78	89,72	93,10	78,68	120					
10%	9,40	10,62	10,44	10,34	10,18	10,14	10,14	10,58	10,70	9,92	9,80	10,78	99,88	99,86	99,78	97,06	81,14	71,42	81,68	63,72	99,38	99,24	98,92	95,06	80					
Modelo Ajustado	$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$															$g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$														
	$\mu \in (0, 0,006; 0, 219)$					$\mu \in (0, 0,881; 0, 994)$					$\mu \in (0, 0,006; 0, 219)$					$\mu \in (0, 0,249; 0, 854)$					$\mu \in (0, 0,881; 0, 994)$									
$\frac{\lambda}{\alpha}$	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	25,91	52,72	101,85	145,37	n	
1%	39,92	45,30	51,46	34,48	13,68	9,56	15,62	5,74	37,60	39,14	42,78	33,64	30,48	35,56	45,44	25,62	9,88	5,70	12,82	3,52	32,40	42,02	41,86	28,12	40					
5%	88,44	89,46	89,12	75,34	33,10	24,60	35,10	15,70	77,42	78,04	76,76	67,78	75,66	79,22	83,48	69,46	31,34	21,30	33,90	13,04	69,40	78,94	76,04	66,16	80					
10%	96,30	96,02	95,44	87,58	46,06	33,80	47,80	22,56	88,96	89,44	87,48	80,82	91,92	95,18	93,72	88,74	45,74	34,66	47,16	20,16	84,22	92,28	87,42	83,38	120					
1%	8,40	4,62	10,08	3,20	16,66	8,60	19,82	9,02	13,36	6,98	16,50	6,24	4,60	3,10	5,30	1,04	13,14	5,44	16,40	5,90	12,82	10,32	13,64	3,38	80					
5%	84,34	73,86	83,96	56,34	48,44	33,26	51,56	29,22	78,54	70,06	76,82	57,46	68,10	55,36	72,82	37,90	45,14	29,66	47,50	23,38	80,04	74,28	78,40	45,98	120					
10%	98,76	96,82	97,52	86,76	64,30	49,72	68,58	44,00	95,38	93,74	93,16	83,96	96,64	92,54	96,12	78,36	61,94	51,66	64,88	37,48	97,24	97,02	95,28	78,02	80					
1%	7,44	3,54	7,08	1,64	16,54	6,82	19,78	8,88	11,14	5,32	12,68	3,58	3,02	1,80	2,64	0,48	13,08	4,46	16,10	5,68	8,88	7,46	8,44	1,54	120					
5%	83,34	73,16	83,06	49,60	52,96	34,14	56,26	33,10	77,36	67,72	78,00	52,94	62,58	48,72	65,46	27,60	48,58	30,26	51,46	25,50	77,58	71,30	76,42	37,30	120					
10%	98,60	97,16	98,26	85,50	70,86	53,92	72,92	48,60	95,34	93,98	94,64	82,82	95,62	89,40	95,26	69,86	67,28	54,00	68,46	41,14	96,52	96,34	95,30	73,38	80					

Capítulo 5

Algumas Aplicações do Teste

Nesse capítulo ajustamos o modelo de regressão simplex para alguns conjuntos de dados e procuramos avaliar a adequabilidade dos ajustes por meio do teste de diagnóstico baseado em influência local. Além disso, também realizamos a análise de diagnóstico tradicional, para compararmos com o resultado fornecido pelo teste.

5.1 Aplicação I: Dados do Cloro

Nesta aplicação analisamos os dados da fração de cloro de Draper e Smith (1981), tais dados encontram-se disponíveis no Apêndice desse trabalho. É considerado que um determinado produto deve conter logo após o exato momento de sua fabricação uma fração de cloro disponível igual a 0,5, sendo conhecido o fato que essa fração irá diminuindo com o tempo, e que em oito semanas após a produção, ela terá decaído 0,01 unidade. Assim, ajustamos a esse conjunto de dados o modelo de regressão simplex com dispersão constante, de modo que a variável resposta será a fração de cloro disponível e a variável explicativa será o número de semanas após a fabricação.

Tabela 5.1: Estimativas dos parâmetros do modelo de regressão simplex ajustado aos dados do cloro.

Coefficiente	Estimativa	Erro-padrão	p -valor
β_0	-0,0462	0,0238	0,0000
β_1	-0,0117	0,0010	0,0000
σ^2	0,0148	0,0032	0,0000

Na Tabela 5.1 encontramos as estimativas dos parâmetros do modelo junto com seus respectivos erros-padrão e p -valores, em que a função de ligação usada foi a logito. Podemos observar que todos os parâmetros são significativos a todos os níveis de significância usuais.

Tendo como objetivo testar as seguintes hipóteses:

- H_0 : o modelo ajustado está adequado aos dados do cloro;
- H_1 : o modelo ajustado não está adequado aos dados do cloro;

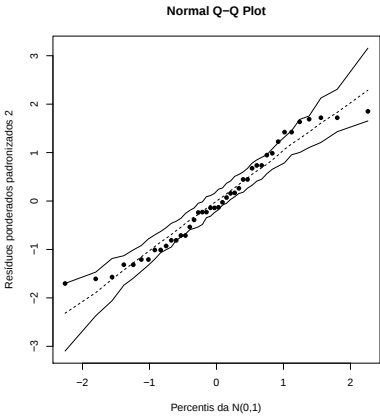
calculamos a estatística U_n e obtivemos pelo método *bootstrap* paramétrico os quantis da distribuição empírica dessa estatística para os níveis de significância 1%, 5% e 10%, também obtivemos os quantis da distribuição normal padrão correspondentes a cada nível de significância.

Tabela 5.2: Resultados do teste de diagnóstico para os dados do cloro.

Estatística U_n	Quantis <i>Bootstrap</i>			Quantis da $N(0, 1)$		
	1%	5%	10%	1%	5%	10%
-0,9328	-7,0365	-5,2329	-4,3567	-2,5758	-1,9600	-1,6449
	1,8967	1,4969	1,0706	2,5758	1,9600	1,6449

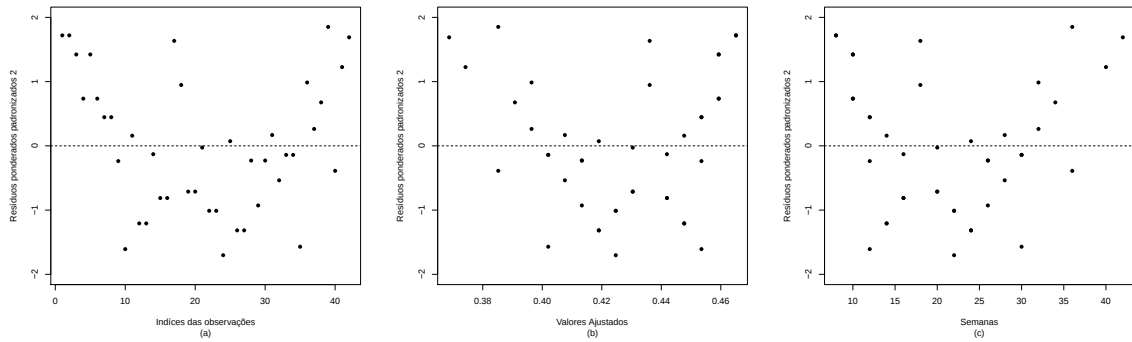
Pela Tabela 5.2 podemos verificar que tanto para os quantis *Bootstrap*, como para os quantis da $N(0, 1)$, correspondentes ao três níveis, o valor da estatística se encontra dentro das regiões de aceitação, logo, pelo teste de diagnóstico baseado em influência local não devemos rejeitar a hipótese H_0 , isto é, o modelo de regressão simplex com dispersão fixa é um bom modelo para o conjunto de dados do cloro.

Figura 5.1: Envelope dos resíduos ponderados padronizados para os dados do cloro.



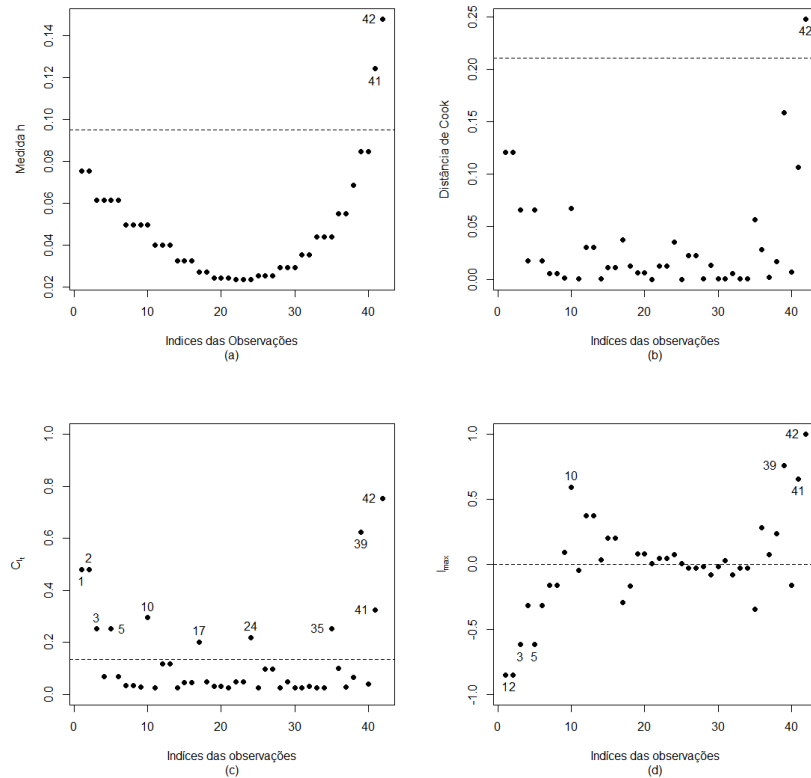
Fonte: Autoria própria.

Pelo gráfico normal de probabilidades com envelope apresentado pela Figura 5.1, vemos que quase todos os pontos encontram-se bem distribuídos dentro das bandas de confiança do intervalo, o que reforça consideravelmente o resultado apresentado pelo teste, ou seja, de fato, o modelo de regressão simplex com dispersão fixa é adequado para os dados do cloro.

Figura 5.2: Gráficos dos resíduos ponderados padronizados para os dados do cloro.

Fonte: Autoria própria.

Na Figura 5.2 encontramos os gráficos dos resíduos ponderados padronizados 2, contra os índices das observações, contra os valores ajustados e também contra a covariável número de semanas, respectivamente, esses resíduos são apresentados por Miyashiro (2008). Notamos que todos eles apresentam comportamento semelhantes, em que os resíduos estão distribuídos aleatoriamente em torno do zero e dentro dos limites $(-2, 2)$. Os gráficos de resíduos evidenciam ainda uma tendência não-linear.

Figura 5.3: Gráfico dos possíveis pontos de alavanca e influentes para os dados do cloro, respectivamente.

Fonte: Autoria própria.

Entretanto, de acordo com os gráficos da Figura 5.3a os pontos 41 e 42 são indicados como possíveis pontos de alavanca, ou seja, eles se destacam por apresentarem altos valores para

o número de semanas e assim, podem estar interferindo demasiadamente nas estimativas dos parâmetros. Além disso, o gráfico da distância de Cook (Figura 5.3b) reforça a afirmação de que o ponto 42 é uma observação bastante influente nessas estimativas.

Na Figura 5.3c e 5.3d estão os gráficos das curvaturas normais (C_{I_t}) e dos vetores I_{max} , respectivamente. Pelo gráfico de I_{max} temos que as observações 1, 2, 3, 5, 10, 17, 24, 35, 39, 41 e 42 são indicadas com sendo individualmente influentes, enquanto que o gráfico de C_{I_t} destaca que as observações 1, 2, 3, 5, 10, 39, 41 e 42 provavelmente são conjuntamente influentes.

Dessa forma, decidimos avaliar o impacto que essas observações provocam nas estimativas dos parâmetros. Para isso, procuramos reajustar o modelo sem essas observações, em seguida calculamos a mudança relativa em porcentagem das estimativas dos parâmetros do modelo. Essa mudança é definida por

$$\text{Mudança Relativa em Porcentagem} = \left| \frac{\hat{\theta} - \hat{\theta}_{(i)}}{\hat{\theta}} \right| \times 100\%,$$

em que $\hat{\theta}_{(i)}$ é a estimativa do parâmetro $\hat{\theta}$ sem a i -ésima observação. Pela Tabela 5.3 vemos que cada uma dessas observações exercem uma influência maior na estimativa do intercepto β_0 do que nas estimativas de β_1 ou de σ^2 . Porém essa influência não é tão alta, logo, podemos dizer que o modelo de regressão simplex se ajusta bem a esse conjunto de dados.

Tabela 5.3: Mudanças relativas em porcentagens para as estimativas dos parâmetros do modelo para os dados do cloro.

Observação	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}^2$
1	24,60	3,48	4,60
2	24,60	3,48	4,60
3	17,92	2,44	2,34
5	17,92	2,44	2,34
10	17,55	2,27	3,79
17	9,92	0,87	3,88
24	4,92	0,07	4,62
35	5,57	1,92	3,56
39	16,06	4,02	5,61
41	15,15	3,50	1,05
42	24,26	5,46	4,25
1,2,3, 5, 10, 39, 41 e 42	8,06	4,78	37,43
Todas	32,88	0,34	54,12

Finalizando, podemos concluir que o teste apresentou um bom desempenho em identificar que um modelo com distribuição simplex é adequado aos dados. No entanto, uma possível não linearidade deve ser considerada.

5.2 Aplicação II: Dados dos Transplantes Autólogos de Células Tronco do Sangue Periférico

Os transplantes autólogos de células tronco do sangue periférico (CTSP) têm sido amplamente utilizados para uma rápida recuperação hematológica, após a terapia mieloablativa para diversas doenças hematológicas malignas. Um estudo envolvendo 239 pacientes foi desenvolvido pelo Edmonton Hematopoietic Stem Cell Lab, no Cross Cancer Institute, em Alberta, Canadá. Esses pacientes consentiram em realizar o transplante autólogo de CTSP, após receber doses de quimioterapia mieloablativos entre os anos de 2003 e 2008. Os dados coletados são referentes a idade dos pacientes, sexo e algumas características clínicas. Os dados dessa aplicação foram usados por Allan et al. (2002) e Yang et al. (2005) e encontram-se disponíveis no Apêndice desse trabalho.

Nesse estudo são usadas as seguintes variáveis:

- idade dos pacientes;
- taxa de recuperação das células CD34+;
- quimio - corresponde a uma variável dummy, indicando se o paciente recebe uma quimioterapia no protocolo de 1 dia (quimio=0) ou de 3 dias (quimio=1);
- idade ajustada - idades menores que 40 anos são definidas como idades da linha de base, idades acima de 40 anos são ajustadas subtraindo o valor 40 da idade dos pacientes.

Dessa forma, ajustamos a esse conjunto de dados o modelo de regressão simplex com dispersão constante, de modo que a variável resposta corresponde a taxa de recuperação das células CD34+ e as variáveis explicativas são a idade ajustada e a variável quimio.

Na Tabela 5.4 estão as estimativas dos parâmetros do modelo junto com seus respectivos erros-padrão e p -valores, podemos observar que ao nível de significância de 5%, todos os parâmetros são significativos.

Tabela 5.4: Estimativas dos parâmetros do modelo de regressão simplex ajustado aos dados dos transplantes autólogos.

Coefficiente	Estimativa	Erro-padrão	p -valor
β_0	1,1002	0,1401	0,0000
β_1	0,0136	0,0065	0,0373
β_2	0,2661	0,1245	0,0333
σ^2	6,3966	0,5851	0,0000

Procurando testar as hipóteses

H_0 : o modelo ajustado está adequado aos dados dos transplantes autólogos;

H_1 : o modelo ajustado não está adequado aos dados dos transplantes autólogos;

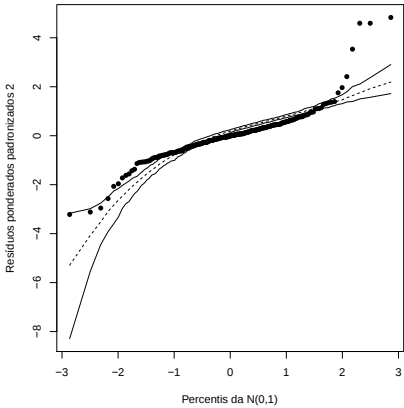
calculamos a estatística U_n e obtemos os quantis *bootstrap* da distribuição empírica de U_n e os quantis da distribuição normal padrão, correspondentes aos níveis de significância 1%, 5% e 10%, dispomos esses valores na Tabela 5.5.

Tabela 5.5: Resultados do teste de diagnóstico para os dados dos transplantes autólogos.

Estatística U_n	Quantis <i>Bootstrap</i>			Quantis da $N(0, 1)$		
	1%	5%	10%	1%	5%	10%
3,3268	-5,8651	-4,1770	-3,3945	-2,5758	-1,9600	-1,6449
	2,0566	1,3068	1,0634	2,5758	1,9600	1,6449

Pela Tabela 5.5 podemos observar que tanto para os quantis *bootstrap* como para os quantis da normal padrão, nos três níveis de significância, o valor da estatística U_n encontra fora da região de aceitação, assim, pelo teste de diagnóstico baseado em influência local devemos rejeitar a hipótese H_0 , ou seja, pelo teste, o modelo simplex com dispersão fixa não é um bom ajuste para os dados dos transplantes autólogos.

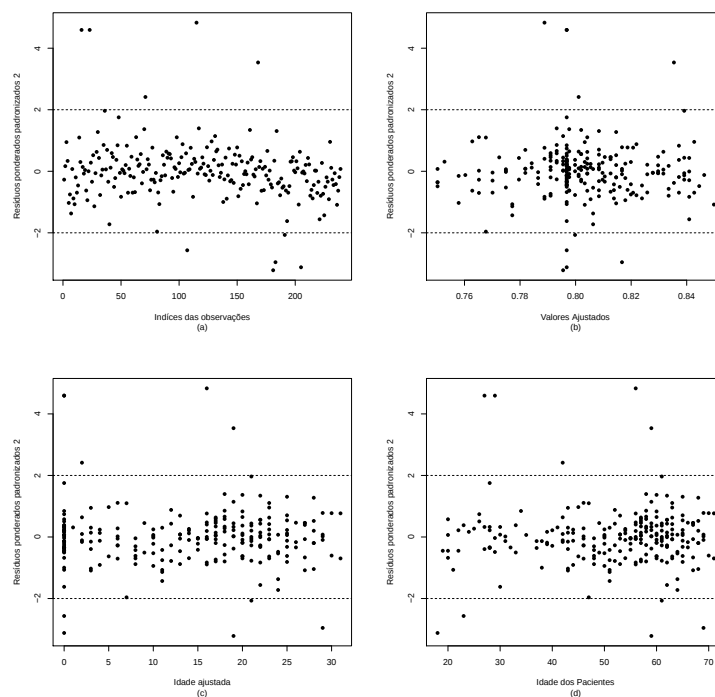
Figura 5.4: Envelope dos resíduos ponderados padronizados para os dados dos transplantes autólogos



Fonte: Autoria própria.

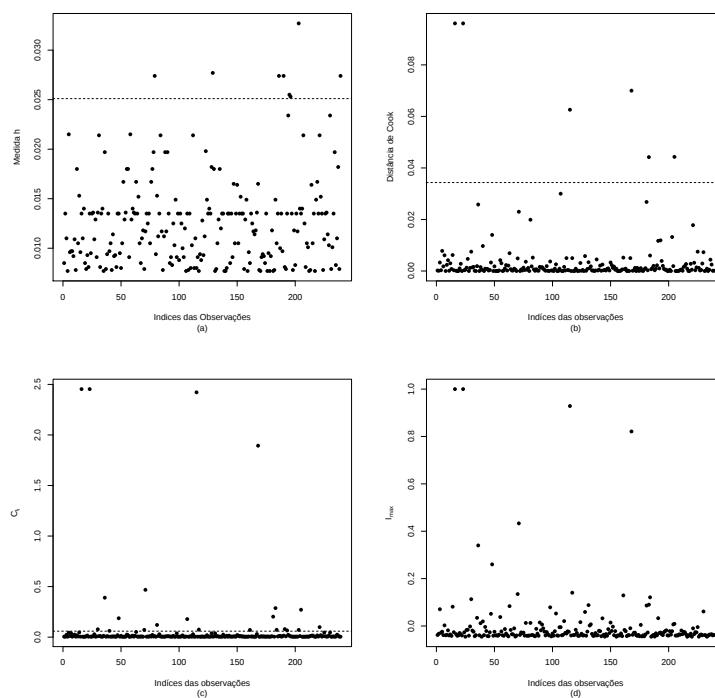
Observando o gráfico normal de probabilidades com envelope ilustrado pela Figura 5.4, vemos que a maior parte dos pontos encontram-se fora das bandas de confiança do intervalo, indicando assim que o modelo ajustados não é adequado para esse conjunto de dados, reafirmando os resultados apresentados pelo teste de diagnóstico baseado em influência local.

Pelos gráficos dos resíduos ponderados padronizados 2 da Figura 5.5 podemos observar que várias observações apresentam altos valores de resíduos, logo, o ajuste do modelo simplex com dispersão fixa para esse conjunto de observações provoca a existência de muitos pontos aberrantes, que devem estar interferindo de maneira desproporcional na média dos valores preditos da taxa de recuperação das células CD34+ e nas estimativas dos parâmetros do modelo. Observamos também que a medida que a idade dos pacientes

Figura 5.5: Gráficos dos resíduos ponderados padronizados para os dados dos transplantes autólogos.

Fonte: Autoria própria.

aumenta os valores dos resíduos tendem a se distanciar mais do zero, (Figura 5.5d), indicando que provavelmente os dados possuem um certo grau de heterogeneidade.

Figura 5.6: Alguns gráficos de diagnósticos para os dados dos transplantes autólogos.

Fonte: Autoria própria.

Avaliando os gráficos de diagnósticos apresentados pela Figura 5.6a-d, verificamos a existência de muitos pontos de alavanca e influentes, logo, esses pontos podem exercer uma influência desproporcional nas estimativas dos parâmetros do modelo. No caso do gráfico da Figura 5.6c observamos a presença de muitos pontos que são conjuntamente influentes, além disso os valores das curvaturas normais de alguns pontos são consideravelmente altos. Portanto, existem fortes indícios de que o ajuste do modelo de regressão simplex com dispersão fixa não é adequado para esse conjunto de dados, ou seja o teste de diagnóstico conseguiu captar a má especificação do modelo.

Considerando o fato que pacientes com idades elevada produzem altos valores de resíduos, buscamos ajustar um modelo de regressão simplex com dispersão variável, logo iremos ajustar um modelo para a dispersão considerando a variável idade como sendo uma covariável desse modelo, em que a função de ligação usada para o modelo da dispersão foi a logarítmica. Assim, encontramos na Tabela 5.6 as estimativas dos parâmetros de ambos os modelos.

Tabela 5.6: Estimativas dos parâmetros do modelo de regressão simplex com dispersão variável, ajustado aos dados dos transplantes autólogos.

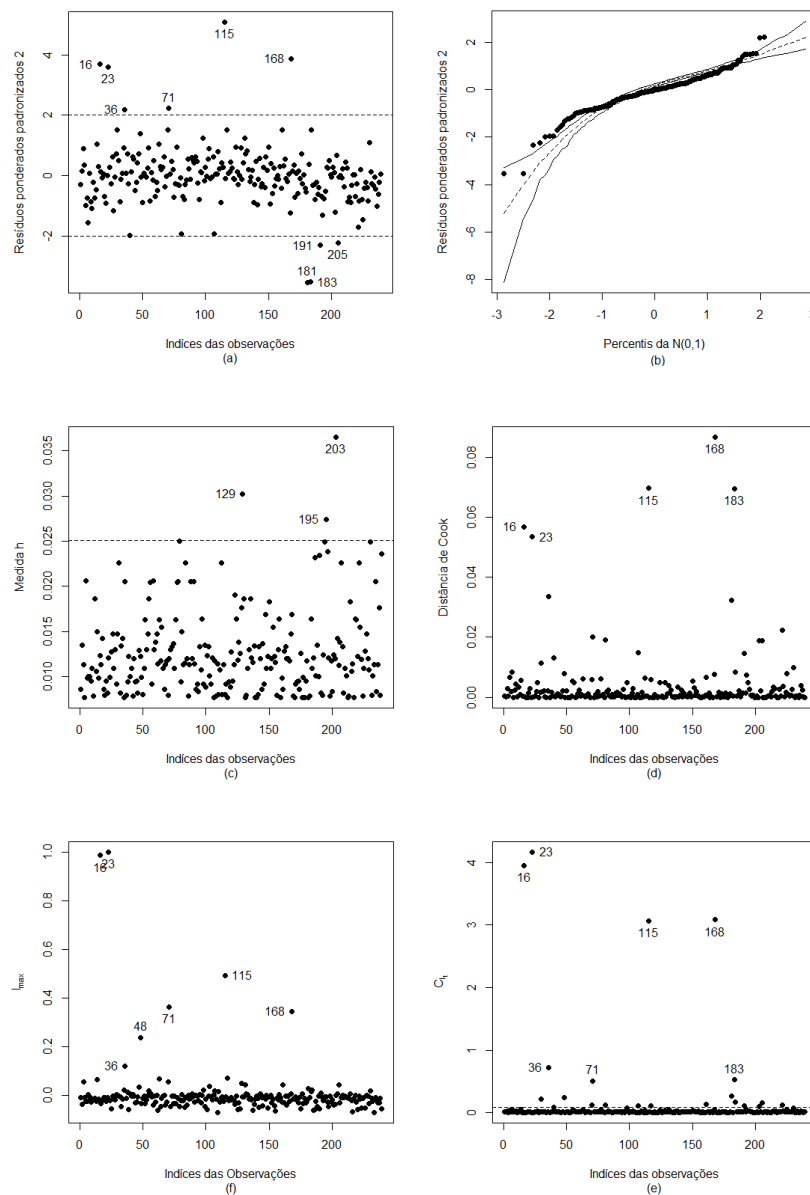
Coefficiente	Estimativa	Erro-padrão	p -valor
β_0	1,1156	0,1414	0,0000
β_1	0,0130	0,0065	0,0437
β_2	0,2519	0,1218	0,0386
γ_0	2,6075	0,3669	0,0000
γ_1	-0,0150	0,0069	0,0292

Podemos notar que ao nível de significância de 5%, todas covariáveis são significativas para o modelo. Na Tabela 5.7 encontramos ainda os resultados fornecidos pela aplicação do teste de diagnóstico baseado em influência local, notamos que, para os quantis *bootstrap* em que $\alpha = 1\%$ e $\alpha = 5\%$ o teste indica a realização de um bom ajuste, porém, para $\alpha = 10\%$, o teste rejeita a hipótese do ajuste estar adequado. Para os quantis da distribuição normal padrão, o teste indica a realização de um bom ajuste para os três níveis de significância.

Tabela 5.7: Resultados do teste de diagnóstico para os dados dos transplantes autólogos, para o modelo com dispersão variável.

Estatística U_n	Quantis <i>Bootstrap</i>			Quantis da $N(0, 1)$		
	1%	5%	10%	1%	5%	10%
0,2636	-0,7955	-0,6048	-0,5226	-2,5758	-1,9600	-1,6449
	0,3719	0,2792	0,2099	2,5758	1,9600	1,6449

Figura 5.7: Alguns gráficos de diagnósticos para os dados dos transplantes autólogos, para o modelo simplex com dispersão variável.



Fonte: Autoria própria.

A partir dos gráficos apresentados pela Figura 5.7 é possível observar o gráfico normal de probabilidade com envelope, para o modelo simplex com dispersão variável, pelo gráfico podemos observar que esse modelo apresenta um melhor ajuste para os dados dos transplantes autólogos, quando comparado com o modelo com dispersão fixa, no entanto, muitos pontos ainda encontram-se fora do envelope, indicando que o modelo não é completamente adequado.

Observamos ainda, que no ajuste do modelo com dispersão variável, alguns pontos são indicados como possíveis pontos de alavanca e influentes, além disso alguns pontos também apresentam resíduos altos. Entretanto, o número de pontos indicados como possíveis pontos influentes é bem menor do que quando usamos o modelo com dispersão fixa, logo,

Tabela 5.8: Mudanças relativas em porcentagens para as estimativas dos parâmetros do modelo com dispersão variável para os dados dos transplantes autólogos.

Observação	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
16	4,01	17,24	1,79	20,76	60,10
23	4,07	17,43	1,49	22,37	65,63
36	2,40	10,50	12,92	2,54	13,83
48	1,26	5,46	0,95	4,25	12,36
71	1,63	7,13	2,90	4,63	11,17
115	4,27	6,24	15,23	0,30	14,99
129	0,19	0,85	0,88	0,41	1,93
168	3,57	15,44	21,98	3,92	26,02
183	1,14	15,35	3,47	2,29	9,78
195	0,71	3,16	3,31	0,32	1,48
203	2,09	9,32	8,77	0,09	0,31
Todas	15,49	63,97	21,70	74,51	166,99
16, 23, 36, 48, 71, 115, 168	12,22	39,15	30,85	76,35	174,99
16, 23	10,12	44,49	4,92	57,29	167,83

podemos dizer que o modelo com dispersão variável se ajusta melhor aos dados, isso indica um bom desempenho do testes de diagnóstico.

Avaliando a mudança relativa que a retirada de cada uma dessas observações provoca nas estimativas dos parâmetros do modelo, observamos pela Tabela 5.8 que as observações 16 e 23 são bastante influentes nas estimativas do parâmetro γ_2 , as demais não exercem tanta influência quantos essas duas. A retirada conjunta das observações 16 e 23 provoca uma mudança consideravelmente alta nas estimativas de muitos parâmetros do modelo. Temos também que, a retirada conjunta das observações 16, 23, 36, 48, 71, 115 e 168 provoca uma grande mudança nas estimativas dos parâmetros. Portanto, podemos concluir que o ajuste do modelo simplex com dispersão variável é de fato melhor do que o modelo com dispersão fixa. Entretanto, a presença de algumas observações bastante influentes não torna o modelo completamente adequado. Assim, verificamos que o teste de diagnóstico com os quantis *bootstrap*, com $\alpha = 10\%$, consegue representar bem essa situação.

Capítulo 6

Considerações finais

6.1 Conclusão

Esse trabalho consistiu em desenvolver um teste de diagnóstico para o modelo de regressão simplex, nos casos em que o parâmetro de dispersão é fixo e também variável. Antes, fizemos uma breve revisão dos modelos de dispersão, do método de influência local e do método *bootstrap*, em seguida, apresentamos ambos os modelos, juntamente com seus, respectivos, vetores escores, matrizes hessianas e de informação de Fisher e as matrizes $\Delta(\theta)$, introduzidas durante a análise de influência local pelo esquema de perturbação ponderação de casos.

Posteriormente, apresentamos o teste de diagnóstico proposto por Zhu e Zhang (2004), esse método é construído baseando-se no método de influência local. Os autores introduzem uma estatística de teste, que sob a hipótese que os dados são de fato originados do modelo ajustado, segue uma distribuição aproximadamente normal padrão. Dessa forma, desenvolvemos esse teste para o modelo de regressão simplex com dispersão fixa e o modelo com dispersão variável. No entanto, ao realizar um breve estudo simulação verificamos que não é adequado utilizar os quantis da distribuição normal padrão, assim, propomos trabalhar com os quantis empíricos da distribuição da estatística de teste, obtidos via método *bootstrap* paramétrico.

Dessa forma, por meio de simulação de Monte Carlo, procuramos avaliar o desempenho dos testes em identificar uma má especificação do modelo, seguindo três diferentes situações. Pudemos ver que, para situações em que os dados são provenientes da uma distribuição beta e ajustamos o modelo simplex, o teste apresentou um bom desempenho em detectar a má especificação do modelo quando os valores assumidos pela média estão concentrados nos extremos do intervalo $(0, 1)$. O mesmo ocorreu para as situações em que ajustamos modelos de regressão omitindo algumas covariáveis, entretanto, nesse caso, o teste apresentou poderes consideravelmente altos. Para o teste desenvolvido para o modelo simplex com dispersão variável, os resultados obtidos apresentaram comportamentos semelhantes para com dispersão fixa, entretanto, os poderes obtidos foram bem mais altos.

Outro erro de especificação trabalhado foi ajustar um modelo com dispersão fixa, quando os dados são provenientes de modelo com dispersão variável, nesse caso, o teste apresenta um bom desempenho quando os valores da média estão próximos de zero e ajustamos um modelo com baixa dispersão. O teste apresentou bons resultados, principalmente para tamanhos amostrais de $n = 120$ e nível de significância de $\alpha = 10\%$.

Assim, concluímos que esse teste consiste numa importante ferramenta para avaliar a adequabilidade do ajuste de modelos simplex, tanto com dispersão fixa quanto variável.

6.2 Trabalhos Futuros

Serão objetivos de nossas próximas pesquisas:

- Desenvolver o teste de diagnóstico baseado em influência local para o modelo de regressão simplex não linear.
- Desenvolver estudo de simulações adicionais, principalmente para o modelo de regressão simplex com dispersão variável.
- Desenvolver o teste de diagnóstico para modelos simplex utilizando outro esquema de perturbação, como por exemplo perturbação no parâmetro de dispersão.

Referências

- ALLAN, D. et al. **Number of Viable CD34+ Cells Reinfused Predicts Engraftment in Autologous Hematopoietic Stem Cell Transplantation.** Bone Marrow Transplant, 20, 967-72, 2000.
- ANDREWS, D. W. K. **Generic uniform convergence.** Economet. Theory, 8, 241-57, 1992.
- ATKINSON, A. C. **Two graphical display for outlying and influential observations in regression.** Biometrika 68, 13-20, 1981.
- BARNDORFF-NIELSEN, O. E.; JØRGENSEN, B. **Some parametric models on the simplex.** Journal of Multivariate Analysis, 39, 106-116, 1991.
- BELSLEY, D. A.; EDWIN, K.; ROY, E. W. **Regression Diagnostics: identifying influential data and sources of collinearity.** New York: John Wiley and Sons, 1980.
- COOK, R. D.; WEISBERG, S. **Residuals and Influence in Regression.** London, Chapman and Hall, 1982.
- COOK, R. D. **Detection of influential observations in linear regressions.** Technometrics 19, 15-18, 1977.
- COOK, R. D. **Assessment of local influence (with discussion).** Journal of the Royal Statistical Society B 48, 133-169, 1986.
- COOK, R. D. et al. **Resampling Methods for Inference.** Part of the Iowa State University NSF/ILI project, 2001.
- CORDEIRO, G. M. et al. **Bartlett corrections for one-parameter exponential family models.** Journal of Statistical Computation and Simulation, 53, 211-231, 1995.
- COX, D.; SNELL, E. **A general definition of residuals.** Journal of the Royal Statistical Society B, 30, 248-275, 1968.

- DOORNIK, J.A. **Ox: an Object-oriented matrix Programming Language**, 4th ed. London: Timberlake Consultants and Oxford, 2001. Disponível em <www.doornik.com>. Acesso em: 11 mar.2015.
- DRAPER, N.R.; SMITH, H. **Applied Regression Analysis**. 2nd ed. New York, Wiley, 1981.
- EFRON, B. **Bootstrap methods: another look at the jackknife**. Annals of Statistics, 7, 1-26, 1979.
- EFRON, B.; TIBSHIRANI, R.J. **An Introduction to the Bootstrap**. New York, Chapman and Hall, 1993.
- ESPINHEIRA, P. L. **Regressão beta**. Tese de Doutorado em Estatística. Instituto de Matemática e Estatística. Universidade de São Paulo, 2007.
- FERRARRI, S. L. P.; CRIBARI-NETO, F. **Beta regression for modelling rates and proportions**. Journal of Applied Statistics, 31, 799–815, 2004.
- FURTADO, R. G. C. **Métodos Estatísticos Aplicados ao Cálculo de Ampacidade e Risco Térmico de Linhas Aéreas de Transmissão**. Dissertação de Mestrado em Engenharia Elétrica. Departamento de Engenharia Elétrica. Universidade Federal de Juiz de Fora, 2008.
- HOAGLIN, D. C.; WELSCH, R. E. **The hat matrix in regression and ANOVA**. The American Statistician 32, 17-22, 1978.
- JAMES, B. **Probabilidade: um curso em nível intermediário**. Instituto de Matemática Pura e Aplicada. Rio de Janeiro, 1981.
- JØRGENSEN, B. **The theory of dispersion models**. London, Chapman and Hall, 1997.
- KNUTH, D. **The TEXbook**. New York, Addison-Wesley, 1986.
- LAMPORT, L. **A Document Preparation System**. 2nd ed. Massachusetts, Addison-Wesley, 1994.
- LESAFFRE, E.; VERBEKE, G. **Local influence in linear mixed models**. Biometrics, 54, 570-582, 1998.
- LIMA, L. B. **Um Teste de Especificação Correta em Modelos de Regressão Beta**. Dissertação de Mestrado em Estatística. Departamento de Estatística. Universidade Federal de Pernambuco, 2007.
- LÓPEZ, F. O. **A Bayesian Approach to Parameter Estimation in Simplex Regression Model: A comparison with Beta Regression**. Revista Colombiana de Estadística, 36, 1–21, 2013.

- MAGALHÃES, M. M. **Probabilidade e Variáveis Aleatórias**. 2ª edição. São Paulo, edusp, 2006.
- MCCULLAGH, P.; NELDER, J. A. **Generalized Linear Models**. 2nd ed. London, Chapman and Hall, 1989.
- MEIRELES, L. C. **Coefficientes de Predição para os Modelos de Regressão Beta e Simplex**. Dissertação de Mestrado em Estatística. Departamento de Estatística. Universidade Federal de Pernambuco, 2015.
- MIYASHIRO, E. S. **Modelos de regressão beta e simplex para análise de proporções**. Dissertação de Mestrado em Estatística. Instituto de Matemática e Estatística. Universidade de São Paulo, 2008.
- NELDER, J. A.; WEDDERBURN, R. W. M. **Generalized linear models**. Journal of the Royal Statistical Society A, 135, 370-384, 1972.
- OLIVEIRA, A. **Regressão Simplex Não Linear: Inferência e Diagnóstico**. Dissertação de Mestrado em Estatística. Departamento de Estatística. Universidade Federal de Pernambuco, 2015.
- PAULA, G. A. **Modelos de Regressão: com apoio computacional**. Instituto de Matemática e Estatística. Universidade de São Paulo, 2013.
- PERTERNELLI, L. A.; MELLO, M. P. **Conhecendo o R : uma visão estatística**. Viçosa, Editora UFV, 2011.
- POLLARD, D. **Empirical Processes: Theory and Applications**. CBMS Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, Hayward, 1990.
- PREGIBON, D. **Logistic regression diagnostics**. Annals of Statistics 9, 705-724, 1981.
- RAMSEY, J. B. **Tests for specification errors in classical linear least squares regression analysis**. Journal of the Royal Statistical Society, B, 31, 350-371, 1969.
- SANTOS, L. A. **Modelos de Regressão Simplex: Resíduos de Pearson Corrigidos e Aplicações**. Tese de Doutorado em Ciências. Escola Superior de Agricultura “Luiz de Queiroz”. Universidade de São Paulo, 2011.
- SONG, P. X. -K.; TAN, M. **Marginal models for longitudinal continuous proportional data**. Biometrics, 56, 496-502, 2000.
- SONG, P. X. -K.; QIU, Z.; TAN, M. **Modelling heterogeneous dispersion in marginal models for longitudinal proportional data**. Biometrical Journal, 46, 540-553, 2004.

SONG, P. X.-K. **Dispersion models in regression analysis**. Pakistan Journal of Statistics. 25, 529-551, 2009.

VENEZUELA, M. K. **Equação de estimação generalizada e influência local para modelos de regressão beta com medidas repetidas**. Tese de Doutorado em Estatística, Instituto de Matemática e Estatística. Universidade de São Paulo, 2007.

WHITE H. **Maximum likelihood estimation of misspecified models**. Econometrica, 50, 1-25, 1982.

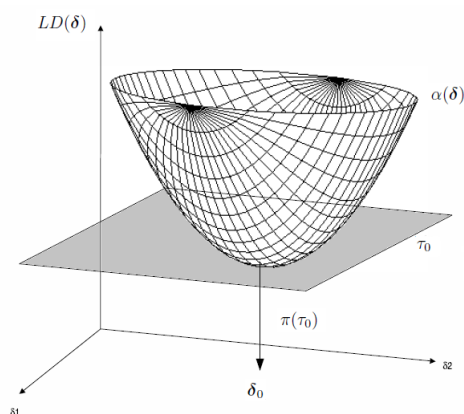
YANG, H. et al. **Association of Post-Thaw viable CD34+ Cells and CFU-GM with Time to Hematopoietic Engraftment**. Bone Marrow Transplant, 35, 881-887, 2005.

ZHU, H.; ZHANG, H. **A diagnostic procedure based on local influence**. Biometrika, 91, 579-589, 2004.

Apêndice A

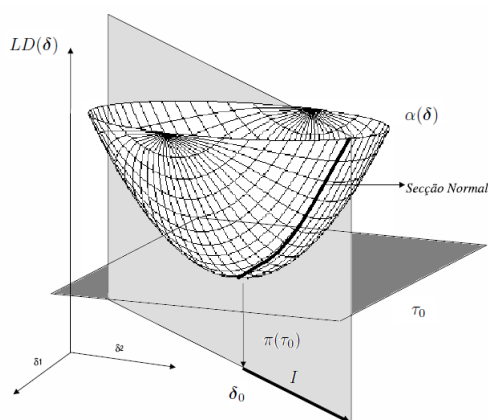
Gráficos das Curvaturas Normais

Figura A.1: Curvatura da superfície $\alpha(\delta)$.



Fonte: Espinheira (2007).

Figura A.2: Curvatura da superfície $\alpha(\delta)$ e a secção normal.



Fonte: Espinheira (2007).

Apêndice B

Propriedades da Distribuição Simplex

Seja y uma variável aleatória com distribuição simplex, com parâmetros μ e σ^2 , ou seja, $y \sim S^-(\mu, \sigma^2)$, o logaritmo da função densidade da distribuição simplex é dado por

$$\ell(y; \mu, \sigma^2) = -\frac{1}{2} \left\{ \log\{2\pi\sigma^2[y(1-y)]^3\} + \frac{1}{\sigma^2}d(y; \mu) \right\}. \quad (\text{B.1})$$

Podemos verificar que a função score para o parâmetro μ é dada por

$$\begin{aligned} \frac{\partial \ell(y; \mu, \sigma^2)}{\partial \mu} &= \frac{\partial}{\partial \mu} \left\{ -\frac{1}{2\sigma^2}d(y; \mu) \right\} \\ &= -\frac{1}{2\sigma^2} \frac{\partial d(y; \mu)}{\partial \mu} \\ &= -\frac{1}{2\sigma^2}d'(y; \mu), \end{aligned}$$

em que,

$$\begin{aligned} d'(y; \mu) &= \frac{\partial}{\partial \mu} \left[\frac{(y - \mu)^2}{y(1-y)\mu^2(1-\mu)^2} \right] \\ &= \frac{2(y - \mu)(-1)[y(1-y)\mu^2(1-\mu)^2] - (y - \mu)^2 y(1-y)[2\mu(1-\mu)^2 - 2(1-\mu)\mu^2]}{y^2(1-y)^2\mu^4(1-\mu)^4} \\ &= \frac{-2(y - \mu)y(1-y)[\mu^2(1-\mu)^2 + (y - \mu)\mu(1-\mu)^2 - (y - \mu)\mu^2(1-\mu)]}{y^2(1-y)^2\mu^4(1-\mu)^4} \\ &= \frac{-2(y - \mu)y(1-y)\mu(1-\mu)[\mu(1-\mu) + (y - \mu)(1-\mu) - (y - \mu)\mu]}{y^2(1-y)^2\mu^4(1-\mu)^4} \\ &= \frac{-2(y - \mu)[\mu(1-\mu) + (y - \mu)(1-\mu) - (y - \mu)\mu]}{y(1-y)\mu^3(1-\mu)^3} \end{aligned}$$

$$\begin{aligned}
 &= \frac{-2(y - \mu)[\mu - \mu^2 + (y - y\mu - \mu + \mu^2) - y\mu + \mu^2]}{y(1 - y)\mu^3(1 - \mu)^3} \\
 &= \frac{-2(y - \mu)[y - 2y\mu + \mu^2]}{y(1 - y)\mu^3(1 - \mu)^3} \\
 &= \frac{-2(y - \mu)[\mu^2 - 2y\mu + y + y^2 - y^2]}{y(1 - y)\mu^3(1 - \mu)^3} \\
 &= \frac{-2(y - \mu)}{\mu(1 - \mu)} \left[\frac{(y^2 - 2y\mu + \mu^2) + (y - y^2)}{y(1 - y)\mu^2(1 - \mu)^2} \right] \\
 &= \frac{-2(y - \mu)}{\mu(1 - \mu)} \left[\frac{(y - \mu)^2 + y(1 - y)}{y(1 - y)\mu^2(1 - \mu)^2} \right] \\
 &= \frac{-2(y - \mu)}{\mu(1 - \mu)} \left[\frac{(y - \mu)^2}{y(1 - y)\mu^2(1 - \mu)^2} + \frac{y(1 - y)}{y(1 - y)\mu^2(1 - \mu)^2} \right] \\
 &= \frac{-2(y - \mu)}{\mu(1 - \mu)} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right].
 \end{aligned}$$

Assim que a função score de μ fica dada por

$$\frac{\partial \ell(y; \mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2} \frac{(y - \mu)}{\mu(1 - \mu)} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right]. \quad (\text{B.2})$$

Definindo

$$u(y; \mu) = \frac{1}{\mu(1 - \mu)} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right], \quad (\text{B.3})$$

temos então $d'(y; \mu) = -2(y - \mu)u(y; \mu)$.

Considerando a função (B.1), tem-se que a função score para σ^2 é dada por

$$\begin{aligned}
 \frac{\partial \ell(y; \mu, \sigma^2)}{\partial \sigma^2} &= -\frac{1}{2} \left\{ \frac{1}{\sigma^2} - \frac{d(y; \mu)}{(\sigma^2)^2} \right\} \\
 &= -\frac{1}{2\sigma^2} \left\{ 1 - \frac{d(y; \mu)}{\sigma^2} \right\}.
 \end{aligned} \quad (\text{B.4})$$

Tendo como objetivo obter a informação de Fisher para μ , seguiremos calculando a derivada segunda de (B.1) em relação a μ , então temos

$$\begin{aligned}
 \frac{\partial^2 \ell(y; \mu, \sigma^2)}{\partial \mu^2} &= \frac{1}{\sigma^2} \left\{ \frac{\partial}{\partial \mu} \left[\frac{(y - \mu)}{\mu(1 - \mu)} \right] \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] \right. \\
 &\quad \left. + \left[\frac{(y - \mu)}{\mu(1 - \mu)} \right] \frac{\partial}{\partial \mu} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] \right\}.
 \end{aligned} \quad (\text{B.5})$$

Do primeiro termo de (B.5), vemos que

$$\frac{\partial}{\partial \mu} \left[\frac{(y - \mu)}{\mu(1 - \mu)} \right] = - \left[\frac{1}{\mu(1 - \mu)} + \frac{(y - \mu)(1 - 2\mu)}{\mu^2(1 - \mu)^2} \right],$$

logo,

$$\begin{aligned} \frac{\partial}{\partial \mu} \left[\frac{(y - \mu)}{\mu(1 - \mu)} \right] \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] &= - \left[\frac{d(y; \mu)}{\mu(1 - \mu)} + \frac{1}{\mu^3(1 - \mu)^3} \right. \\ &\quad + \frac{(y - \mu)(1 - 2\mu)d(y; \mu)}{\mu^2(1 - \mu)^2} \\ &\quad \left. + \frac{(y - \mu)(1 - 2\mu)}{\mu^4(1 - \mu)^4} \right]. \end{aligned}$$

Do segundo termo de (B.5), vemos que

$$\frac{\partial}{\partial \mu} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] = d'(y; \mu) - \frac{2(1 - 2\mu)}{\mu^3(1 - \mu)^3},$$

assim

$$\left[\frac{(y - \mu)}{\mu(1 - \mu)} \right] \frac{\partial}{\partial \mu} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] = \frac{(y - \mu)}{\mu(1 - \mu)} d'(y; \mu) - \frac{2(y - \mu)(1 - 2\mu)}{\mu^4(1 - \mu)^4}.$$

Dessa forma,

$$\begin{aligned} \frac{\partial^2 \ell(y; \mu, \sigma^2)}{\partial \mu^2} &= \frac{1}{\sigma^2} \left[\frac{-d(y; \mu)}{\mu(1 - \mu)} - \frac{1}{\mu^3(1 - \mu)^3} - \frac{(y - \mu)(1 - 2\mu)d(y; \mu)}{\mu^2(1 - \mu)^2} \right. \\ &\quad \left. - \frac{3(y - \mu)(1 - 2\mu)}{\mu^4(1 - \mu)^4} + \frac{(y - \mu)d'(y; \mu)}{\mu(1 - \mu)} \right]. \end{aligned} \quad (\text{B.6})$$

A partir das propriedades apresentadas pela Proposição 2.1, obtemos que a informação de Fisher para μ é dada por

$$-\mathbb{E} \left[\frac{\partial^2 \log p(y; \mu, \sigma^2)}{\partial \mu^2} \right] = \frac{1}{\sigma^2} \left\{ \frac{3\sigma^2}{\mu(1 - \mu)} + \frac{1}{\mu^3(1 - \mu)^3} \right\}. \quad (\text{B.7})$$

A derivada segunda de (B.1) em relação a σ^2 é

$$\frac{\partial^2 \ell(y; \mu, \sigma^2)}{\partial (\sigma^2)^2} = \frac{1}{2(\sigma^2)^2} - \frac{d(y; \mu)}{(\sigma^2)^3}, \quad (\text{B.8})$$

então facilmente temos que a informação de Fisher para σ é definida por

$$-\mathbb{E} \left[\frac{\partial^2 \log p(y; \mu, \sigma^2)}{\partial (\sigma^2)^2} \right] = \frac{1}{2(\sigma^2)^2}. \quad (\text{B.9})$$

B.1 Prova do item (iv) da Proposição 2.1:

Seja y uma variável aleatória com distribuição simplex, com parâmetros μ e σ^2 , ou seja, $y \sim S^-(\mu, \sigma^2)$, então, a função densidade de probabilidade é dada por

$$p(y; \mu, \sigma^2) = \{2\pi\sigma^2[y(1-y)]^3\}^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}d(y; \mu)\right\}, \quad (\text{B.10})$$

para $0 < y < 1$, em que

$$d(y; \mu) = \frac{(y - \mu)^2}{y(1-y)\mu^2(1-\mu)^2}.$$

Supondo as seguinte tranformações para y e μ :

$$x = \frac{y}{1-y} \Rightarrow y = \frac{x}{1+x} \quad \text{e} \quad \xi = \frac{\mu}{1-\mu} \Rightarrow \mu = \frac{\xi}{1+\xi}. \quad (\text{B.11})$$

Assim, temos que

$$d(y; \mu) = d(x; \xi) = \frac{(x - \xi)^2(1 + \xi)^2}{x\xi^2}$$

e

$$y(1-y) = \frac{x}{(1+x)^2},$$

definindo $\sigma^2 = 1/\nu$, dessa forma a função (B.10) fica definida por

$$\begin{aligned} p(x; \xi, \nu) &= \left\{ \frac{2\pi}{\nu} \left[\frac{x}{(1+x)^2} \right]^3 \right\}^{-1/2} \exp\left\{ -\frac{\nu}{2} \frac{(x - \xi)^2(1 + \xi)^2}{x\xi^2} \right\} \\ &= \sqrt{\frac{\nu}{2\pi}} \left[\frac{x}{(1+x)^2} \right]^{-3/2} \exp\left\{ -\frac{\nu}{2} \frac{(x - \xi)^2(1 + \xi)^2}{x\xi^2} \right\}. \end{aligned} \quad (\text{B.12})$$

Logo, calculando $\mathbb{E}[(y - \mu)d^2(y; \mu)]$, temos

$$\begin{aligned} \mathbb{E}[(y - \mu)d^2(y; \mu)] &= \int_0^1 (y - \mu)d^2(y; \mu)p(y; \mu, \sigma^2)dy \\ &= \int_0^\infty \left(\frac{x}{1+x} - \frac{\xi}{1+\xi} \right) \left[\frac{(x - \xi)^2(1 + \xi)^2}{x\xi^2} \right]^2 \sqrt{\frac{\nu}{2\pi}} \left[\frac{x}{(1+x)^2} \right]^{-3/2} \\ &\quad \exp\left\{ -\frac{\nu}{2} \frac{(x - \xi)^2(1 + \xi)^2}{x\xi^2} \right\} \frac{1}{(1+x)^2} dx \end{aligned}$$

$$\begin{aligned}
 &= \int_0^\infty \frac{x}{1+x} \left[\frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right]^2 \sqrt{\frac{\nu}{2\pi}} \left[\frac{x}{(1+x)^2} \right]^{-3/2} \\
 &\quad \exp \left\{ -\frac{\nu}{2} \frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right\} \frac{1}{(1+x)^2} dx \\
 &- \frac{\xi}{1+\xi} \int_0^\infty \left[\frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right]^2 \sqrt{\frac{\nu}{2\pi}} \left[\frac{x}{(1+x)^2} \right]^{-3/2} \\
 &\quad \exp \left\{ -\frac{\nu}{2} \frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right\} \frac{1}{(1+x)^2} dx.
 \end{aligned} \tag{B.13}$$

Definindo,

$$f(x; \xi, \nu) = \exp \left\{ -\frac{\nu}{2} \frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right\} \tag{B.14}$$

e procurando desenvolver a integral do primeiro termo de (B.13), obtemos

$$\begin{aligned}
 &\int_0^\infty \frac{x}{1+x} \left[\frac{(x-\xi)^2(1+\xi)^2}{x\xi^2} \right]^2 \sqrt{\frac{\nu}{2\pi}} \left[\frac{x}{(1+x)^2} \right]^{-3/2} \frac{1}{(1+x)^2} f(x; \xi, \nu) dx \\
 &= \sqrt{\frac{\nu}{2\pi}} \frac{(1+\xi)^4}{\xi^4} \int_0^\infty x^{-5/2} (x-\xi)^4 f(x; \xi, \nu) dx \\
 &= \sqrt{\frac{\nu}{2\pi}} \frac{(1+\xi)^4}{\xi^4} \int_0^\infty x^{-5/2} (\xi^4 - 4\xi^3 x + 6\xi^2 x^2 - 4\xi x^3 + x^4) f(x; \xi, \nu) dx \\
 &= \sqrt{\frac{\nu}{2\pi}} \frac{(1+\xi)^4}{\xi^4} \left[\xi^4 \int_0^\infty x^{-5/2} f(x; \xi, \nu) dx - 4\xi^3 \int_0^\infty x^{-3/2} f(x; \xi, \nu) dx \right. \\
 &\quad + 6\xi^2 \int_0^\infty x^{-1/2} f(x; \xi, \nu) dx - 4\xi \int_0^\infty x^{1/2} f(x; \xi, \nu) dx \\
 &\quad \left. + \int_0^\infty x^{3/2} f(x; \xi, \nu) dx \right].
 \end{aligned} \tag{B.15}$$

Entretanto, baseando-se nas expressões (5.41)-(5.43) de Jørgensen (1997), Song e Tan (2000) afirmam que

$$\int_0^\infty x^{-5/2} f(x; \xi, \nu) dx = \left(\frac{2\pi}{\nu} \right)^{1/2} \frac{\xi + \nu(1+\xi)^2}{\nu\xi(1+\xi)^3},$$

$$\int_0^\infty x^{-3/2} f(x; \xi, \nu) dx = \left(\frac{2\pi}{\nu} \right)^{1/2} \frac{1}{(1+\xi)},$$

$$\int_0^\infty x^{-1/2} f(x; \xi, \nu) dx = \left(\frac{2\pi}{\nu} \right)^{1/2} \frac{\xi}{(1+\xi)},$$

$$\int_0^\infty x^{1/2} f(x; \xi, \nu) dx = \left(\frac{2\pi}{\nu} \right)^{1/2} \frac{\xi^3 + \nu \xi^2 (1 + \xi)^2}{\nu (1 + \xi)^3}$$

e de Song et al.(2004), temos que

$$\int_0^\infty x^{3/2} f(x; \xi, \nu) dx = \left(\frac{2\pi}{\nu} \right)^{1/2} \frac{\nu^2 \xi^3 (1 + \xi)^4 + 3\nu \xi^4 (1 + \xi)^2 + 3\xi^5}{\nu^2 (1 + \xi)^5}.$$

Dessa forma, a expressão (B.15) é dada por

$$\begin{aligned} & \frac{(1 + \xi)^4}{\xi^4} \left\{ \frac{\xi^5 + \xi^4 \nu (1 + \xi)^2}{\nu \xi (1 + \xi)^3} - \frac{4\xi^3}{(1 + \xi)} + \frac{6\xi^3}{(1 + \xi)} \right. \\ & - 4 \left[\frac{\xi^4 + \nu \xi^3 (1 + \xi)^2}{\nu (1 + \xi)^3} \right] + \left. \frac{\nu^2 \xi^3 (1 + \xi)^4 + 3\nu \xi^4 (1 + \xi)^2 + 3\xi^5}{\nu^2 (1 + \xi)^5} \right\} \\ & = \frac{(1 + \xi)^4}{\xi^4} \left[\frac{3\xi^5}{\nu^2 (1 + \xi)^5} \right] \\ & = \frac{3\xi}{\nu^2 (1 + \xi)}. \end{aligned}$$

Podemos observar que o segundo termo de (B.13) é equivalente a

$$\frac{\xi}{(1 + \xi)} \mathbb{E}[d^2(y; \mu)],$$

em que a partir dos itens (vii) e (i) da Proposição 2.1 temos que $\mathbb{E}[d^2(y; \mu)] = 3/\nu^2$, portanto, provamos que $\mathbb{E}[(y - \mu)d^2(y; \mu)] = 0$.

B.2 Prova dos itens (v) e (vi) da Proposição 2.1:

Sabendo que

$$d'(y; \mu) = -2(y - \mu)u(y; \mu), \quad (\text{B.16})$$

então,

$$\begin{aligned} \mathbb{E}[d'(y; \mu)] &= -2\mathbb{E}[(y - \mu)u(y; \mu)] \\ &= -2 \frac{(y - \mu)}{\mu(1 - \mu)} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] \\ &= \frac{-2}{\mu(1 - \mu)} \left[\mathbb{E}[(y - \mu)d(y; \mu)] + \frac{1}{\mu^2(1 - \mu)^2} \mathbb{E}(y - \mu) \right], \end{aligned}$$

utilizando o item (iii) da Proposição 2.1 e sabendo que $\mathbb{E}(y - \mu) = 0$, concluimos assim, que $\mathbb{E}[d'(y; \mu)] = 0$.

De (B.16), obtemos que

$$\begin{aligned} d''(y; \mu) &= 2 \left[u(y; \mu) + \frac{(y - \mu)(1 - 2\mu)d(y; \mu)}{\mu^2(1 - \mu)^2} + \frac{3(y - \mu)(1 - 2\mu)}{\mu^4(1 - \mu)^4} \right. \\ &\quad \left. - \frac{(y - \mu)d'(y; \mu)}{\mu(1 - \mu)} \right], \end{aligned} \quad (\text{B.17})$$

logo,

$$\begin{aligned} \mathbb{E}[d''(y; \mu)] &= 2 \left\{ \mathbb{E}[(y; \mu)] + \frac{(1 - 2\mu)\mathbb{E}[(y - \mu)d(y; \mu)]}{\mu^2(1 - \mu)^2} + \frac{3(1 - 2\mu)\mathbb{E}[(y - \mu)]}{\mu^4(1 - \mu)^4} \right. \\ &\quad \left. - \frac{\mathbb{E}[(y - \mu)d'(y; \mu)]}{\mu(1 - \mu)} \right\}, \end{aligned}$$

facilmente vemos que

$$\begin{aligned} \mathbb{E}[u(y; \mu)] &= \mathbb{E} \left\{ \frac{1}{\mu(1 - \mu)} \left[d(y; \mu) + \frac{1}{\mu^2(1 - \mu)^2} \right] \right\} \\ &= \frac{\sigma^2}{\mu(1 - \mu)} + \frac{1}{\mu^3(1 - \mu)^3} \end{aligned}$$

e

$$\frac{\mathbb{E}[(y - \mu)d'(y; \mu)]}{\mu(1 - \mu)} = \frac{-2\sigma^2}{\mu(1 - \mu)}.$$

Portanto,

$$\frac{1}{2}\mathbb{E}[d''(y; \mu)] = \frac{3\sigma^2}{\mu(1 - \mu)} + \frac{1}{\mu^3(1 - \mu)^3}. \quad (\text{B.18})$$

Apêndice C

Prova dos Teoremas

C.1 Provando o Teorema 1:

Já sabemos que

$$H_{LD(\delta_0)} = 2\Delta^\top(\boldsymbol{\theta})[-\ddot{\ell}(\boldsymbol{\theta})]^{-1}\Delta(\boldsymbol{\theta}) \Big|_{\delta=\delta_0, \theta=\hat{\theta}},$$

e da Proposição 1.1, vimos que

$$\sum_{t=1}^n C_{I_t} = \text{tr}(H_{LD(\delta_0)}) \xrightarrow{p} \lambda_0,$$

em que $\lambda_0 = 2\text{tr}[A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)]$. Logo, para provar o primeiro item, basta mostrar que quando $n \rightarrow \infty$

$$\left| \frac{-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - A(\boldsymbol{\theta}_*) \right| \xrightarrow{p} 0 \tag{C.1}$$

e

$$\left| \frac{\Delta\Delta^\top(\hat{\boldsymbol{\theta}})}{n} - B(\boldsymbol{\theta}_*) \right| \xrightarrow{p} 0. \tag{C.2}$$

Assim, de (C.1), temos que

$$\begin{aligned}
 \left| \frac{-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})}{n} - A(\boldsymbol{\theta}_*) \right| &= \left| \frac{-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})}{n} - A(\boldsymbol{\theta}_*) - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} + \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} \right. \\
 &\quad \left. - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} + \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right| \\
 &\leq \left| \frac{-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} \right| + \left| \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right| \\
 &\quad + \left| \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} - A(\boldsymbol{\theta}_*) \right|.
 \end{aligned}$$

Portanto, devemos mostrar que os três termos do lado direito da inequação convergem em probabilidade para zero. Pela suposição S5 já observamos diretamente, que o terceiro termo converge em para zero, quando n cresce.

De acordo com o lema 2(a) de Andrews (1992) e sob as suposições feitas é possível verificar que

$$\left| n^{-1} \mathbb{E}[-\partial_i \partial_j \ell_t(\widehat{\boldsymbol{\theta}})] - n^{-1} \mathbb{E}[-\partial_i \partial_j \ell_t(\boldsymbol{\theta}_*)] \right| \xrightarrow{p} 0, \quad (\text{C.3})$$

quando $n \rightarrow \infty$, em que $\partial_k = \partial / \partial \theta_k$, com $k = 1, \dots, p$. Ou seja,

$$\left| \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right| \xrightarrow{p} 0. \quad (\text{C.4})$$

Sob as suposições S1 e S4 e utilizando o Teorema 8.2 de Pollard (1990), obtemos

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| \frac{-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})]}{n} \right| \rightarrow 0.$$

Logo,

$$\left| \frac{-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\widehat{\boldsymbol{\theta}})]}{n} \right| \xrightarrow{p} 0. \quad (\text{C.5})$$

Então, podemos concluir a prova da expressão (C.1), e prova de (C.2) é feita de modo similar. Portanto, com esses resultados fica provado o item 1 do Teorema 1.

Para a prova do segundo item, supomos que $g(y) = p(y; \boldsymbol{\theta}_0)$, logo, de (3.1) temos que $\boldsymbol{\theta}_* =$

$\boldsymbol{\theta}_0$. Prosseguindo, é possível verificar da (i, j) -ésima entrada da matriz $\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]/n$, que

$$\begin{aligned} \left\{ \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right\}_{ij} &= \frac{-\mathbb{E}[\sum_{t=1}^n \partial_i \partial_j \ell_t(\boldsymbol{\theta}_*)]}{n} \\ &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left[\frac{1}{p(y_t; \boldsymbol{\theta}_*)} \partial_i p(y_t; \boldsymbol{\theta}_*) \frac{1}{p(y_t; \boldsymbol{\theta}_*)} \partial_j p(y_t; \boldsymbol{\theta}_*) \right] \\ &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} [\partial_i \ell_t(\boldsymbol{\theta}_*) \partial_j \ell_t(\boldsymbol{\theta}_*)] \\ &= \left\{ \frac{\mathbb{E}[\Delta \Delta^\top(\boldsymbol{\theta}_*)]}{n} \right\}_{ij}. \end{aligned}$$

Assim,

$$\frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} = \frac{\mathbb{E}[\Delta \Delta^\top(\boldsymbol{\theta}_*)]}{n}, \quad (\text{C.6})$$

então,

$$\lim_{n \leftarrow \infty} \left\{ -\frac{\mathbb{E}[\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right\} = \lim_{n \leftarrow \infty} \left\{ \frac{\mathbb{E}[\Delta \Delta^\top(\boldsymbol{\theta}_*)]}{n} \right\},$$

desse modo, sob S5, obtemos que $A(\boldsymbol{\theta}_*) = B(\boldsymbol{\theta}_*)$. Consequentemente, temos que

$$\begin{aligned} \lambda_0 &= 2\text{tr}[A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)] \\ &= 2\text{tr}(I_p) \\ &= 2p, \end{aligned}$$

em que, I_p corresponde a matriz identidade, de ordem $p \times p$. Dessa forma, fica provado o Teorema 1.

C.2 Prova do Teorema 2:

Sabemos que C_{I_t} é uma função de $\hat{\boldsymbol{\theta}}$, logo substituindo $\hat{\boldsymbol{\theta}}$ por $\boldsymbol{\theta}$, podemos representar C_{I_t} como uma função de $\boldsymbol{\theta}$, ou seja $C_{I_t}(\boldsymbol{\theta})$.

Dessa forma, $SC_n(\boldsymbol{\theta}) = \sum_{t=1}^n nC_{I_t}(\boldsymbol{\theta})$. Usando a expansão de Taylor na função $SC_n(\hat{\boldsymbol{\theta}})$ em torno de $\boldsymbol{\theta}_*$, obtemos

$$SC_n(\hat{\boldsymbol{\theta}}) = SC_n(\boldsymbol{\theta}_*) + \frac{\partial SC_n(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) + o_p(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*),$$

pela suposição S2 temos que $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_*$, então $o_p(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) = o_p(1)$, logo

$$SC_n(\hat{\boldsymbol{\theta}}) = SC_n(\boldsymbol{\theta}_*) + \frac{\partial SC_n(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) + o_p(1). \quad (\text{C.7})$$

Multiplicando ambos os lados de (C.7) por $n^{-1/2}$ e usando a suposição S7, temos que

$$n^{-1/2}SC_n(\hat{\boldsymbol{\theta}}) = n^{-1/2}SC_n(\boldsymbol{\theta}_*) + J_1(\boldsymbol{\theta}_*)^\top \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) + o_p(1). \quad (\text{C.8})$$

Por meio de uma expansão assintótica, Zhu e Zhang (2004) concluem que

$$n^{-1/2}SC_n(\boldsymbol{\theta}_*) = n^{-1/2}\lambda_0 + n^{-1/2} \sum_{t=1}^n h(y_t; \boldsymbol{\theta}_*) + o_p(1). \quad (\text{C.9})$$

Agora, expandindo $\sum_{t=1}^n \partial \ell_t(\hat{\boldsymbol{\theta}})/\partial \boldsymbol{\theta} = 0$ em torno de $\boldsymbol{\theta}_*$, pela expansão de Taylor, obtemos

$$\begin{aligned} \sum_{t=1}^n \frac{\partial \ell_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} &= \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + \sum_{t=1}^n \frac{\partial^2 \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) + o_p(1) \\ &= \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) + o_p(1) = 0. \end{aligned}$$

Assim, vemos que

$$\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) = - \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + o_p(1).$$

Facilmente vemos que

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) = \frac{1}{n} \left[\frac{-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)}{n} \right]^{-1} \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + \frac{1}{n} \left[\frac{-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)}{n} \right]^{-1} o_p(1).$$

Utilizando o Teorema 8.2 de Pollard (1990), sabemos por (C.5) que

$$\left| \frac{-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - \frac{\mathbb{E}[-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})]}{n} \right| \xrightarrow{p} 0,$$

entretanto, pela suposição S5, temos que

$$\lim_{n \rightarrow \infty} \left\{ -\frac{\mathbb{E}[\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*)]}{n} \right\} = A(\boldsymbol{\theta}_*).$$

Assim,

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) = \frac{1}{n} A(\boldsymbol{\theta}_*)^{-1} \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + o_p(1). \quad (\text{C.10})$$

Dessa forma, utilizando as expressões (C.9) e (C.10) na equação (C.8), obtemos

$$\begin{aligned} n^{-1/2} \text{SC}_n(\hat{\boldsymbol{\theta}}) &= n^{-1/2} \lambda_0 + n^{-1/2} \sum_{t=1}^n h(y_t; \boldsymbol{\theta}_*) + J_1(\boldsymbol{\theta}_*)^\top n^{-1/2} A(\boldsymbol{\theta}_*)^{-1} \sum_{t=1}^n \frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} + o_p(1) \\ &= n^{-1/2} \lambda_0 + n^{-1/2} \sum_{t=1}^n \left\{ h(y_t; \boldsymbol{\theta}_*) + J_1(\boldsymbol{\theta}_*)^\top A(\boldsymbol{\theta}_*)^{-1} \left[\frac{\partial \ell_t(\boldsymbol{\theta}_*)}{\partial \boldsymbol{\theta}} \right] \right\} + o_p(1) \\ &= n^{-1/2} \lambda_0 + n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) + o_p(1). \end{aligned} \quad (\text{C.11})$$

Portanto, fica provado a primeira parte do Teorema 2.

Agora, definindo $X_t = k(y_t; \boldsymbol{\theta}_*)$, então temos que $X_1, \dots, X_n$ são independentes. É possível verificar que nesse caso, a condição de Lindeberg é válida, de forma que, para todo $t > 0$

$$\frac{\sum_{t=1}^n \mathbb{E} \left\{ X_t^2 I_{\left[\frac{X_t}{S_n} > t \right]} \right\}}{S_n^2} \xrightarrow{p} 0, \quad (\text{C.12})$$

em que $\mathbb{E}(X_t) = 0$, $\text{Var}(X_t) = \sigma_t^2$ e $S_n^2 = \sum_{t=1}^n \sigma_t^2$.

Então, temos que $\sum_{t=1}^n X_t / S_n \xrightarrow{D} N(0, 1)$. Assim, por hipótese $S_n / \sqrt{n} \xrightarrow{p} \sigma_*$, logo

$$n^{-1/2} \sum_{t=1}^n X_t \xrightarrow{D} N(0, \sigma_*),$$

ou seja,

$$n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) \xrightarrow{D} N(0, \sigma_*^2).$$

Pela definição de $T_n(\widehat{\boldsymbol{\theta}})$, temos que

$$T_n(\widehat{\boldsymbol{\theta}}) = n^{-1/2} \sum_{t=1}^n nC_{I_t} - n^{-1/2} \lambda_0.$$

Sabendo que

$$\begin{aligned} n^{-1/2} \sum_{t=1}^n nC_{I_t} &= n^{1/2} \lambda_0 + n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) + o_p(1) \\ n^{-1/2} \sum_{t=1}^n nC_{I_t} - n^{1/2} \lambda_0 &= n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) + o_p(1). \end{aligned}$$

Obtemos então

$$T_n(\widehat{\boldsymbol{\theta}}) = n^{-1/2} \sum_{t=1}^n k(y_t; \boldsymbol{\theta}_*) + o_p(1).$$

Portanto

$$T_n(\widehat{\boldsymbol{\theta}}) \xrightarrow{D} N(0, \sigma_*^2). \quad (\text{C.13})$$

Apêndice D

Dados das Aplicações

Tabela D.1: Conjunto de dados usados na aplicação I, referente a fração de cloro disponível.

Obs.	Nº de semanas	fração de cloro disp.	Obs.	Nº de semanas	fração de cloro disp.
1	8	0.49	22	22	0.41
2	8	0.49	23	22	0.41
3	10	0.48	24	22	0.40
4	10	0.47	25	24	0.42
5	10	0.48	26	24	0.40
6	10	0.47	27	24	0.40
7	12	0.46	28	26	0.41
8	12	0.46	29	26	0.40
9	12	0.45	30	26	0.41
10	12	0.43	31	28	0.41
11	14	0.45	32	28	0.40
12	14	0.43	33	30	0.40
13	14	0.43	34	30	0.40
14	16	0.44	35	30	0.38
15	16	0.43	36	32	0.41
16	16	0.43	37	32	0.40
17	18	0.46	38	34	0.40
18	18	0.45	39	36	0.41
19	20	0.42	40	36	0.38
20	20	0.42	41	40	0.39
21	20	0.43	42	42	0.39

Tabela D.2: Conjunto de dados usados na aplicação II, referente aos transplantes autólogos.

obs	idade	taxa	idade ajust.	chemo	obs	idade	taxa	idade ajust.	chemo
1	62	0.75	22	0	53	56	0.82	16	1
2	39	0.83	0	1	54	68	0.85	28	0
3	43	0.94	3	1	55	60	0.94	20	1
4	58	0.86	18	0	56	71	0.71	31	0
5	43	0.54	3	0	57	63	0.82	23	0
6	51	0.70	11	1	58	43	0.72	3	0
7	64	0.59	24	0	59	56	0.90	16	1
8	64	0.82	24	0	60	48	0.62	8	0
9	48	0.67	8	1	61	68	0.78	28	0
10	51	0.57	11	0	62	20	0.67	0	1
11	57	0.75	17	0	63	46	0.94	6	0
12	60	0.74	20	1	64	23	0.87	0	1
13	53	0.75	13	1	65	58	0.86	18	1
14	47	0.94	7	0	66	65	0.85	25	0
15	51	0.87	11	1	67	62	0.90	22	0
16	29	0.99	0	1	68	43	0.73	3	1
17	43	0.83	3	1	69	50	0.77	10	0
18	57	0.82	17	1	70	60	0.96	20	0
19	62	0.81	22	0	71	42	0.98	2	1
20	60	0.68	20	0	72	35	0.81	0	1
21	50	0.67	10	1	73	67	0.88	27	0
22	61	0.80	21	0	74	40	0.84	0	1
23	27	0.99	0	1	75	65	0.70	25	0
24	46	0.86	6	1	76	46	0.70	6	0
25	32	0.73	0	1	77	71	0.93	31	0
26	68	0.90	28	0	78	61	0.80	21	1
27	51	0.56	11	0	79	37	0.66	0	0
28	56	0.82	16	1	80	53	0.77	13	0
29	63	0.91	23	0	81	47	0.44	7	0
30	68	0.96	28	0	82	54	0.72	14	1
31	62	0.90	22	1	83	21	0.61	0	1
32	61	0.66	21	0	84	62	0.80	22	1
33	38	0.77	0	1	85	42	0.77	2	1
34	57	0.84	17	1	86	57	0.89	17	0
35	58	0.93	18	0	87	54	0.86	14	1
36	61	0.98	21	1	88	61	0.82	21	1
37	60	0.81	20	0	89	42	0.91	2	1
38	53	0.91	13	0	90	61	0.90	21	1
39	28	0.73	0	1	91	20	0.90	0	1
40	64	0.55	24	0	92	62	0.90	22	0
41	65	0.83	25	0	93	63	0.88	23	0
42	57	0.90	17	0	94	55	0.74	15	0
43	66	0.89	26	0	95	67	0.70	27	0
44	49	0.73	9	1	96	44	0.83	4	1
45	50	0.78	10	1	97	69	0.83	29	0
46	61	0.87	21	0	98	63	0.95	23	0
47	34	0.93	0	1	99	39	0.76	0	1
48	28	0.97	0	1	100	54	0.84	14	0
49	46	0.80	6	1	101	29	0.79	0	1
50	56	0.63	16	0	102	67	0.89	27	0
51	65	0.82	25	0	103	52	0.94	12	1
52	46	0.77	6	0	104	63	0.79	23	0

obs	idade	taxa	idade ajust.	chemo	obs	idade	taxa	idade ajust.	chemo
105	55	0.82	15	1	157	56	0.83	16	1
106	58	0.90	18	0	158	69	0.80	29	0
107	23	0.45	0	1	159	58	0.85	18	0
108	57	0.80	17	0	160	51	0.75	11	1
109	20	0.81	0	1	161	62	0.96	22	0
110	56	0.92	16	0	162	33	0.87	0	1
111	44	0.85	4	1	163	67	0.89	27	0
112	62	0.73	22	1	164	42	0.78	2	1
113	56	0.87	16	0	165	66	0.78	26	0
114	43	0.86	3	1	166	50	0.79	10	0
115	56	0.99	16	0	167	68	0.65	28	0
116	63	0.79	23	0	168	59	0.99	19	1
117	58	0.96	18	0	169	57	0.82	17	0
118	53	0.80	13	0	170	59	0.86	19	0
119	54	0.81	14	0	171	63	0.69	23	0
120	49	0.87	9	0	172	53	0.80	13	0
121	50	0.86	10	1	173	48	0.71	8	1
122	54	0.81	14	1	174	62	0.83	22	0
123	44	0.73	4	0	175	31	0.77	0	1
124	69	0.93	29	0	176	58	0.91	18	0
125	24	0.83	0	1	177	46	0.76	6	1
126	48	0.70	8	0	178	59	0.80	19	0
127	31	0.80	0	1	179	48	0.74	8	1
128	45	0.93	5	0	180	50	0.66	10	0
129	65	0.83	25	1	181	59	0.40	19	0
130	60	0.89	20	1	182	57	0.86	17	0
131	59	0.95	19	0	183	69	0.47	29	0
132	60	0.91	20	0	184	65	0.96	25	0
133	65	0.92	25	0	185	27	0.72	0	1
134	56	0.86	16	1	186	28	0.66	0	0
135	60	0.81	20	1	187	52	0.70	12	1
136	55	0.82	15	1	188	39	0.75	0	1
137	22	0.84	0	1	189	64	0.71	24	0
138	38	0.62	0	1	190	29	0.63	0	0
139	58	0.74	18	0	191	61	0.50	21	0
140	59	0.88	19	0	192	60	0.68	20	0
141	56	0.62	16	0	193	30	0.54	0	1
142	37	0.77	0	1	194	63	0.75	23	1
143	26	0.92	0	1	195	64	0.78	24	1
144	62	0.74	22	0	196	41	0.83	1	0
145	40	0.74	0	1	197	27	0.86	0	1
146	63	0.85	23	0	198	57	0.88	17	0
147	59	0.78	19	1	199	50	0.78	10	0
148	65	0.90	25	0	200	55	0.78	15	0
149	28	0.84	0	1	201	30	0.86	0	1
150	70	0.93	30	0	202	42	0.82	2	1
151	65	0.67	25	0	203	67	0.72	27	1
152	26	0.89	0	1	204	57	0.92	17	1
153	58	0.75	18	1	205	18	0.41	0	1
154	28	0.86	0	1	206	48	0.76	8	0
155	22	0.71	0	1	207	62	0.84	22	1
156	60	0.78	20	0	208	67	0.80	27	0

obs	idade	taxa	idade ajust.	chemo	obs	idade	taxa	idade ajust.	chemo
209	61	0.84	21	0	225	51	0.52	11	0
210	65	0.75	25	0	226	33	0.70	0	1
211	52	0.87	12	0	227	68	0.78	28	0
212	59	0.84	19	0	228	66	0.75	26	0
213	57	0.64	17	0	229	44	0.65	4	1
214	70	0.72	30	0	230	63	0.95	23	1
215	53	0.69	13	1	231	60	0.82	20	0
216	20	0.71	0	1	232	52	0.72	12	0
217	59	0.80	19	0	233	19	0.71	0	1
218	69	0.82	29	0	234	61	0.83	21	1
219	46	0.61	6	0	235	55	0.70	15	0
220	28	0.73	0	1	236	43	0.62	3	1
221	62	0.65	22	1	237	45	0.62	5	0
222	58	0.72	18	1	238	60	0.76	20	0
223	25	0.85	0	1	239	30	0.77	0	0
224	57	0.65	17	0					