

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

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

UM TESTE BASEADO EM INFLUÊNCIA LOCAL
PARA AVALIAR QUALIDADE DO AJUSTE EM
MODELOS DE REGRESSÃO BETA

TEREZINHA KÉSSIA DE ASSIS RIBEIRO

DISSERTAÇÃO DE MESTRADO



Recife
2016

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

TEREZINHA KÉSSIA DE ASSIS RIBEIRO

UM TESTE BASEADO EM INFLUÊNCIA LOCAL
PARA AVALIAR QUALIDADE DO AJUSTE EM
MODELOS DE REGRESSÃO BETA

ORIENTADOR: PROFA. DRA. PATRÍCIA LEONE ESPINHEIRA OSPINA
CO-ORIENTADORA: PROFA. DRA. MICHELLI KARINNE BARROS DA SILVA

Área de Concentração: Estatística Aplicada

Dissertação apresentada ao programa de Pós-graduação em Estatística do Departamento de Estatística da Universidade 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

R484t Ribeiro, Terezinha Késsia de Assis.
Um teste baseado em influência local para avaliar qualidade do ajuste em modelos de Regressão Beta / Terezinha Késsia de Assis Ribeiro. – 2016.
128 f.: fig., tab.

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

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

519.536 CDD (22. ed.) UFPE-MEI 2016-30

TEREZINHA KÉSSIA DE ASSIS RIBEIRO

**UM TESTE BASEADO EM INFLUÊNCIA LOCAL PARA AVALIAR QUALIDADE
DO AJUSTE EM MODELOS DE REGRESSÃO BETA**

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. PhD. Francisco Cribari Neto (Examinador Interno)
UFPE

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

*Dedico este trabalho a minha
mãe, Verônica, ao meu pai,
Marcelo, e a minha mãe cien-
tífica, Michelli Barros.*

Agradecimentos

Primeiramente à Deus, pela saúde, disposição e sempre me ajudar a seguir em frente.

Aos meus pais, Marcelo e Verônica, por todo o apoio, amor, paciência que sempre me deram e por sempre acreditarem que eu era capaz de correr atrás dos meus sonhos.

À minha orientadora Patrícia, não somente pela orientação, como também, por contribuir de forma significativa para minha formação acadêmica, pelas sugestões, contribuições, amizade, confiança e por acreditar que fosse capaz de desenvolver esse trabalho. Muito obrigada!

À minha coorientadora, eterna orientadora, mãe científica, Michelli Barros, por todos os ensinamentos acadêmicos e pessoais, orientação, carinho, amizade, força, paciência, preocupação, tempo, conselhos, confiança, dedicação, e por todas as coisas que fizestes por mim e não será possível citá-las por insuficiência de papel no mundo. Obrigada por todo apoio que me destes desde o início da graduação, como também, por todo incentivo com relação ao mestrado. Obrigada por todas as ideias, sugestões, e contribuições feitas, além da disponibilização de um "arsenal" de computadores para a realização das simulações presentes nesse trabalho, sendo possível assim a concretização dessa dissertação. Mais que uma excelente profissional, és um incrível ser humano que se preocupa com todos ao seu redor, e felizmente eu tive a sorte de ser um desses. Jamais conseguirei expressar toda a gratidão que tenho por você, pois não acho que há palavras suficientes ou adjetivos que consigam lhe caracterizar. Meus mais sinceros agradecimentos!

À minha amiga e irmã Fernanda (Fê) por todo o companheirismo, apoio, diversão, paciência, ajuda, disposição, tempo e com certeza sua incrível disponibilidade de comer pizza. Obrigada por enfrentarmos juntas todos os medos, tristezas, felicidades, sustos, desafios, e entre tantas outras coisas vivenciadas durante o mestrado em Recife. Jamais esquecerei de tudo que passamos juntas, muito obrigada por tudo e tenha certeza que amo muito você!

À Berenice e Edenilson, por todo apoio que me deram a minha chegada e estadia em Recife, obrigada por toda preocupação. Aos meus tios adotivos Preta e Nildo, por toda preocupação, carinho, amizade, paciência, tempo, alegria e todas as diversas coisas que fizeram durante todo o tempo que morei em Recife (e quando não morei também), por

serem tão incríveis ao ponto de praticamente me adotarem, e me considerarem da família. Não há como descrever em palavras toda a gratidão que tenho por todos vocês. Muito, muito obrigada! Vocês são simplesmente sensacionais!

À minha irmã de graduação Verinha, por todo seu apoio, carinho, incentivo, diversão e amizade que tanto me fortaleceu. Também, a meu querido amigo Deyvison, por todo apoio, risadas e palavras de incentivo que sempre me alegraram tanto. Muito obrigada! Amo demais vocês.

Aos amigos do mestrado, Hildemar, Jonas, Laura, Pedro, Rapha, Stênio, Telma, Renata e Rodrigo Alves, por todos os ótimos momentos compartilhados de estudo, conselhos, brincadeiras e troca de ideias com que fizeram os momentos vivenciados serem especiais. Em especial, à irmã que ganhei no mestrado Wanessa (Wan) por todo apoio, horas e mais horas de estudo, diversão, turismo em Recife, e sua incrível disponibilidade de fazer bolo, alegrando assim nossos dias. Você foi essencial durante todo o mestrado, obrigada por tudo que fez por mim, por todos os momentos incríveis que passamos juntas que jamais serão esquecidos. Também, à Márcio por todo apoio, descontração e por todas as heineken's durante todo o mestrado, Wan que o diga. À Rodrigo Dim, o irmão que ganhei no mestrado, por todos os momentos de estudo, brincadeiras, troca de ideias mirabolantes, paciência, preocupação e por aturar todos os meus ataques de risos e ser essa pessoa tão iluminada e divertida que tanto me alegrou e me ajudou sempre que precisei. A todos vocês, muitíssimo obrigada! Vocês são excepcionais!

Aos meus professores de graduação da Unidade Acadêmica de Estatística da UFCG, Aleksandro, Amanda, Ana Cristina, Antônio, Chico, Gilberto, Grayci-Mary, João, Joelson, Iraponil, Patrícia e Rosângela, que mesmo longe, sei que torcem muito por mim, e tanto contribuíram na minha formação, fazendo possível a realização desse trabalho. Muito Obrigada!

Aos professores da Unidade Acadêmica de Matemática da UFCG, Lindomberg e Horácio, por toda a ajuda que me deram durante o processo de simulações presentes nessa dissertação, complementando o "arsenal" citado, e que foi de imensa ajuda fazendo possível a conclusão desse trabalho. Meus sinceros agradecimentos!

Aos meus professores de mestrado do Departamento de Estatística da UFPE, Cribari, Gauss e Leandro, por serem excelentes professores e transmitirem tantos conhecimentos contribuindo de forma bastante significativa para a minha formação, e assim, contribuindo na realização desse trabalho.

A Valéria Bittencourt e Elaine Mota, pelo carinho, eficiência e preocupação com todos alunos da pós-graduação.

Aos participantes da banca examinadora, pelas sugestões.

À Capes, pelo apoio financeiro.

*"Nossa maior fraqueza está em desistir.
O caminho mais certo de vencer é tentar
mais uma vez."*

Thomas Edison

Resumo

A classe de modelos de regressão beta introduzida por Ferrari & Cribari-Neto (2004) é muito útil para modelar taxas e proporções. O modelo proposto pelos autores é baseado na suposição de que a variável resposta tem distribuição beta com uma parametrização que é indexada pela média e por um parâmetro de precisão. Após a construção de um modelo de regressão é de extrema importância realizar a análise de diagnóstico, objetivando verificar possíveis afastamentos das suposições feitas para o modelo apresentado, bem como detectar possíveis observações que causem influência desproporcional nas estimativas dos parâmetros. A análise de influência local introduzida por Cook (1986) é uma abordagem que objetiva avaliar a influência das observações. Com base no método de influência local, Zhu & Zhang (2004) propuseram um teste de hipóteses para detectar o grau de discrepância entre o modelo suposto e o modelo subjacente do qual os dados são gerados. Nesse trabalho, foi desenvolvido esse teste para o modelo de regressão beta com dispersão fixa e variável, como também, foram propostos um melhoramento nesse teste baseados na metodologia bootstrap e um novo teste, também com base em influência local, mas considerando outro esquema de perturbação, a perturbação no parâmetro de precisão no modelo de regressão beta com dispersão fixa. O desempenho desses testes foram avaliados com base no tamanho e poder. Por fim, aplicamos a teoria desenvolvida a um conjunto de dados reais.

Palavras-chave: Modelo de Regressão Beta. Teste de hipóteses. Qualidade do ajuste. Bootstrap.

Abstract

The class of beta regression models introduced by Ferrari & Cribari-Neto (2004) is very useful for modelling rates and proportions. The proposed model by the authors is based on the assumption that the response variable is beta distributed with indexed by mean and dispersion parameters. After fitting a regression model is very important to carry out the diagnostic analysis in sense that, verifying possible deviations of the model assumptions, as well as detect possible observations that cause disproportionate influence on the parameter estimates. The local influence analysis introduced by Cook (1986) is an approach that objective assess the influence of observations. Based on local influence method, Zhu & Zhang (2004) proposed a hypothesis test to detect the degree of discrepancy between the supposed model and the underlying model from which the data is generated. In this work, was developed this test for the beta regression model with fixed and varying dispersion, as well as, we proposed in addition, an improvement of this test based on bootstrap methodology and a new test, also based on local influence, but considering other perturbation scheme, the perturbation of the precision parameter in beta regression model with fixed dispersion. The performance of these tests were evaluated based on size and power. Finally, we applied the theory developed to a set of real data.

Keywords: Beta Regression Model. Hypothesis Test. Goodness of Fit. Bootstrap.

Lista de Figuras

2.1	Curvas de densidade da distribuição beta para alguns valores de p e q	27
2.2	Curvas da fda da distribuição beta para alguns valores de p e q	28
2.3	Comportamento da variância dado μ fixo para alguns valores de ϕ	30
4.1	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0049; 0, 1831]$ e $\phi = 20$	60
4.2	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 20$	61
4.3	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 20$	62
4.4	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0049; 0, 1831]$ e $\phi = 50$	63
4.5	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 50$	64
4.6	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 50$	65
4.7	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0049; 0, 1831]$ e $\phi = 150$	66
4.8	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 150$	67
4.9	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 150$	68
4.10	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2835; 0, 8947]$ e $\phi = 150$	69
4.11	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0005; 0, 1337]$ e $\lambda = 44$	80
4.12	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 44$	81
4.13	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 44$ para n	82

4.14	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0005; 0, 1337]$ e $\lambda = 99$	83
4.15	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 99$	84
4.16	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 99$	85
4.17	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 0005; 0, 1337]$ e $\lambda = 163$	86
4.18	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 163$	87
4.19	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 163$	88
4.20	Histograma com a densidade $N(0, 1)$ e <i>qqplot</i> de U_n , respectivamente, com $\mu \in [0, 1176; 0, 8158]$ e $\lambda = 163$	89
7.1	Gráficos de influência local conjunta e individual, considerando o esquema de ponderação de casos e perturbação no parâmetro ϕ , respectivamente.	111
7.2	Gráficos de resíduos do modelo, juntamente com o envelope.	113
7.3	Gráficos de influência local (a) conjunta (b) individual, considerando o esquema de ponderação de casos e perturbação no parâmetro ϕ , respectivamente.	116
7.4	Gráficos de resíduos do modelo, juntamente com o envelope.	117
7.5	Gráficos de influência local conjunta e individual, considerando o esquema de ponderação de casos.	119
7.6	Gráficos dos resíduos do modelo com dispersão variável, juntamente com o envelope.	120

Lista de Tabelas

4.1	Níveis empíricos para o tamanho do teste, usando os níveis nominais 1%, 5% e 10%.	56
4.2	Valores da média, variância, assimetria e curtose para as 5000 réplicas de U_n para o cenário μ central e $\phi = 150$	58
4.3	Valores da média, variância, assimetria e curtose para as 1000 réplicas de U_n para o cenário μ central e $\phi = 150$	58
4.4	Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	59
4.5	Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	59
4.6	Níveis empíricos para o teste, usando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	70
4.7	Níveis empíricos para o teste, usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.	70
4.8	Níveis empíricos para o teste, usando os níveis nominais 1, 5% e 10%. Má especificação: dispersão não modelada.	73
4.9	Níveis empíricos para o teste com variância <i>bootstrap</i> , considerando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	74
4.10	Níveis empíricos para o teste com variância <i>bootstrap</i> , considerando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	74
4.11	Níveis empíricos para o teste com variância <i>bootstrap</i> , considerando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	75
4.12	Níveis empíricos para o teste com variância <i>bootstrap</i> , usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.	75
4.13	Níveis empíricos para o teste, com variância <i>bootstrap</i> , usando os níveis nominais 1%, 5% e 10%. Má especificação: dispersão não modelada.	76
4.14	Níveis empíricos para o tamanho do teste, usando os níveis nominais 1%, 5% e 10%.	77
4.15	Valores da média, variância, assimetria e curtose para as 5000 réplicas de U_n para o cenário μ central e $\lambda = 163$	79

4.16	Valores da média, variância, assimetria e curtose para as 1000 réplicas de U_n para o cenário μ central e $\lambda = 163$	79
4.17	Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).	90
4.18	Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).	90
4.19	Níveis empíricos para o teste, usando nível nominal 10%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).	91
4.20	Níveis empíricos para o teste, usando níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.	91
4.21	Níveis empíricos para o teste, usando 1%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).	92
4.22	Níveis empíricos para o teste, usando 5%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).	93
4.23	Níveis empíricos para o teste, usando 10%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).	93
6.1	Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	103
6.2	Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	104
6.3	Níveis empíricos para o teste usando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).	104
6.4	Níveis empíricos para o teste usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.	105
6.5	Níveis empíricos para o teste usando os níveis nominais 1, 5% e 10%. Má especificação: dispersão não modelada.	107
7.1	Estimativas dos parâmetros para o modelo de regressão beta com dispersão fixa considerando apenas as covariável dislexia, QI e interação entre ambas.	109
7.2	Valores das estatísticas U_n e U_n^* considerando o modelos de regressão beta com dispersão fixa para os dados de dislexia.	109
7.3	Quantis empíricos e assintóticos para as estatísticas U_n e U_n^* considerando o modelo de regressão beta com dispersão fixa.	109
7.4	Mudanças relativas: Modelo 1.	112
7.5	Estimativas dos parâmetros para o modelo de regressão beta com dispersão fixa considerando apenas a covariável dislexia.	114
7.6	Valores das estatísticas U_n e U_n^* considerando o modelos de regressão beta fixa para os dados de estresse.	114
7.7	Quantis empíricos e assintóticos para as estatísticas U_n e U_n^* considerando o modelo de regressão beta com dispersão fixa.	114
7.8	Mudanças relativas: Modelo 2.	115

7.9	Estimativas dos parâmetros para o modelo de regressão beta com dispersão variável com interação.	118
7.10	Estimativas dos parâmetros para o modelo de regressão beta com dispersão variável sem interação.	118
7.11	Quantis empíricos e assintóticos para a estatística U_n para o modelo de regressão beta com dispersão variável.	118
7.12	Mudanças relativas para o modelo com dispersão variável.	119

Sumário

Lista de Figuras	9
Lista de Tabelas	11
Sumário	14
1 Introdução	17
1.1 Introdução	17
1.1.1 Organização da Dissertação	20
1.1.2 Suporte Computacional	20
1.2 Preliminares	21
1.2.1 Influência Local	21
1.2.2 Bootstrap	23
2 Modelos de Regressão Beta	26
2.1 Distribuição Beta	26
2.2 Modelo de Regressão Beta com Dispersão Fixa	31
2.3 Modelo de Regressão Beta com Dispersão Variável	35
3 Teste Baseado em Influência Local	40
3.1 Teste de Hipóteses Baseado em Influência Local	40
3.1.1 Suposições	41
3.1.2 Teste Baseado no Esquema de Perturbação Ponderação de Casos	42
3.2 Cálculo da Estatística para a Regressão Beta	44
3.2.1 Dispersão Fixa	44
3.2.2 Dispersão Variável	51
4 Estudos de Simulação: Parte 1	55
4.1 Modelo de Regressão Beta com Dispersão Fixa	56
4.1.1 Teste com Variância Bootstrap	72
4.2 Modelo de Regressão Beta com Dispersão Variável	77

5	Teste para a Precisão ϕ	95
5.1	Perturbação no Parâmetro de Precisão ϕ	95
5.2	Teste para a Precisão ϕ	97
6	Estudos de Simulação: Parte 2	102
7	Aplicações	108
7.1	Dados de Dislexia	108
8	Considerações Finais	121
	Referências	122
A	Teorema 3.1.1	126

Capítulo 1

Introdução

1.1 Introdução

Fenômenos aleatórios que produzem dados no intervalo $(0, 1)$ são bastante comuns. Dentre as distribuições mais conhecidas que possuem suporte no intervalo $(0, 1)$, o modelo beta é um dos mais utilizados para ajustar esses tipos de dados que, geralmente, são taxas e proporções. Falls (1974) utilizou a distribuição beta para construir um modelo que representa a cobertura de nuvens em todo o mundo. Janardan e Padmanabhan (1986) modelaram variáveis hidrológicas, usando a distribuição beta. Harrop-Williams (1989) desenvolveu estudos de porosidade (proporção) de vazios de solos através do uso do modelo beta. Wiley et al. (1989) obtiveram um modelo beta para estimar a probabilidade de transmissão do vírus HIV durante o contato sexual entre um infectado e um indivíduo sadio. McNally (1990) utilizou a distribuição beta no estudo de algumas variáveis que afetam a reprodutibilidade de vacas. Graham e Hollands (1990) e Milyutin e Yaromenko (1991) aplicaram a distribuição beta para estudar índices relacionados à transmissão de radiação solar. Koppes e Grover (1992) determinaram o tamanho de progênie em *Escherichia Coli*, utilizando a distribuição beta. Sulaiman et al. (1999) modelaram dados da luz do sol para aplicação na utilização de energia solar renovável da Malásia, através da distribuição beta. Ferrari e Cribari-Neto (2004) modelaram a proporção de petróleo convertido em gasolina após destilação e fracionamento, utilizando o modelo beta. Espinheira et al. (2008a) utilizaram a distribuição beta para modelar o índice de indignância no estado de São Paulo (Brasil). Para outras aplicações do modelo beta, ver Johnson et al. (1995).

Devido ao grande número de situações em que a variável de interesse de estudo é do tipo porcentagem ou proporção, surgiu a necessidade de saber qual tipo de modelagem seria mais adequada para ajustar esses tipos de dados. Existem diversas propostas de modelagem para dados no intervalo $(0, 1)$. Kieschnick e McCullough (2003) apresentaram sete possíveis tipos de modelagem: normal linear, logito, normal censurado (*tobit*), normal não-linear, beta, simplex e quasi-verossimilhança. Porém, de acordo com os autores, os quatro primeiros violam algumas suposições básicas. Por exemplo, quando utilizamos um

modelo de regressão normal linear para modelar uma taxa ou proporção, podemos não obter um bom ajuste, devido ao fato de que o suporte do modelo utilizado é sobre todos os reais. Sendo assim, nem sempre o modelo normal linear é adequado para modelar taxas ou proporções. Portanto, nesse contexto, há a classe de modelos de regressão linear beta desenvolvida por Ferrari e Cribari-Neto (2004). O modelo proposto pelos autores é baseado na suposição de que a variável resposta tem distribuição beta com uma parametrização que é indexada pela média e por um parâmetro de precisão. Ainda, nessa classe de modelos, é suposto que a média possui uma estrutura de regressão linear, assim, os parâmetros de regressão são interpretados em termos da média da resposta, um comportamento que se assemelha aos modelos lineares generalizados (MLGs). Algumas extensões do modelo proposto por Ferrari e Cribari-Neto (2004), além de classes de modelos baseados na distribuição beta foram desenvolvidas ao longo dos últimos anos. Smithson e Verkuilen (2006) propuseram uma extensão do modelo de regressão linear beta, admitindo que a precisão do modelo também possua uma estrutura de regressão linear. Com base na classe de modelos proposta por Ferrari e Cribari-Neto (2004) e Smithson e Verkuilen (2006), foi derivada por Simas et al. (2010) uma nova classe de modelos de regressão beta que inclui os modelos anteriores como casos especiais, e além disso, considera uma estrutura de regressão não linear tanto para a média como para a precisão. Ospina e Ferrari (2012) propuseram uma classe geral de modelos de regressão para proporções contínuas em que é suposto que a variável resposta tem uma distribuição contínua-discreta que é uma mistura de uma distribuição beta e uma distribuição degenerada em 0 e ou 1. Carrasco et al. (2014) propuseram a classe de modelos beta com erros nas variáveis. Pereira et al. (2012) propuseram a classe de modelos de regressão beta inflacionado truncado.

Após a construção de um modelo de regressão, que tenta representar o comportamento de algum fenômeno aleatório, é primordial realizar uma análise de diagnóstico, objetivando verificar possíveis afastamentos das suposições feitas, ou seja, investigar se o modelo proposto está corretamente especificado. Por exemplo, no modelo de regressão linear clássico há duas suposições básicas sobre a variável resposta: variância constante (homoscedasticidade) e linearidade nos parâmetros do modelo. Em algumas situações, tais suposições não valem, assim, o modelo de regressão linear clássico estará mal especificado. Uma boa análise de diagnóstico deverá detectar esses tipos de desvios das suposições do modelo proposto.

Várias metodologias foram desenvolvidas por diversos autores com o objetivo de detectar possíveis afastamentos das suposições feitas para o modelo proposto em algumas classes de modelos de regressão. No que se diz respeito a medidas de influência, existem diversas propostas. Considerando o modelo de regressão linear normal, Hoaglin e Welsch (1978) estudaram o comportamento dos elementos da diagonal principal da matriz de projeção $H = X(X^T X)^{-1}X$ em que X denota a matriz modelo, sendo denominados pontos de alavanca, devido à característica de terem peso desproporcional no próprio valor ajustado. St. Laurent e Cook (1992) propuseram extensões da definição de pontos de alavanca para modelos normais não lineares. Paula (1999) discute pontos de alavanca em modelos nor-

mais lineares com parâmetros restritos, com extensões para os MLGs. Outro aspecto que se deve considerar durante a análise de diagnóstico, é se a exclusão de alguma observação (ou algumas observações) causa uma mudança significativa nas estimativas dos parâmetros do modelo, denominadas influentes. Uma medida para detectar pontos influentes no modelo de regressão normal linear é a distância de Cook (1977), em que é realizado um processo de deleção de pontos para avaliar o impacto da retirada de uma observação particular nas estimativas dos parâmetros do modelo. Essa técnica de deleção de pontos, rapidamente, foi assimilada e estendida para diversas classes de modelos (para mais detalhes, ver Paula (1999)). Outra medida que é utilizada para verificação de pontos influentes é a DFFITS proposta por Belsley, Kuhe e Welsch (1980). Por fim, uma proposta interessante e muito utilizada na detecção de pontos influentes em modelos de regressão foi apresentada por Cook (1986) conhecida como método de influência local. O método proposto pelo autor é baseado no conceito de curvatura normal, visando avaliar, através de uma certa medida de influência, se pequenas perturbações introduzidas no modelo ou nos dados causam efeitos desproporcionais nas estimativas dos parâmetros ou se causam mudanças inferenciais.

Em relação a testes de hipóteses, para avaliar a qualidade do ajuste do modelo foram feitas algumas propostas. Na classe de modelos de regressão linear clássico, um teste que é frequentemente utilizado para verificar se o modelo está corretamente especificado é o teste RESET (*regression specification error test*) proposto por Ramsey (1969). Esse teste é baseado na adição de uma forma não linear ao modelo através de termos de grau dois, três, etc., e, em seguida testar, por um teste F usual, a exclusão de tais termos. Desta forma, se os termos de não linearidade forem rejeitados, o modelo estará corretamente especificado. Uma outra abordagem para validar o modelo foi proposta por Zhu e Zhang (2004), uma vez que propuseram um teste de hipóteses para detectar o grau de discrepância entre o modelo suposto e o subjacente do qual os dados são gerados. O referido teste é baseado em influência local, considerando o esquema de perturbação de casos. Tal teste pode ser aplicado a diversas classes de modelos de regressão. Portanto, a aplicação do teste proposto por Zhu e Zhang (2004) apresenta-se como uma boa ferramenta para avaliar a qualidade do ajuste em várias classes de modelos que não possuem um teste rigoroso com esse objetivo.

Nos últimos anos, foram desenvolvidos trabalhos na área de análise de diagnóstico para o modelo de regressão linear beta e extensões. Cribari-Neto e Lima (2007) propuseram o teste RESET para a classe de modelos de regressão linear beta com base no teste proposto por Ramsey (1969). Considerando o modelo de regressão linear beta, Espinheira et al. (2008a) apresentam a distância de Cook e as medidas de influência local. Chien (2011) apresenta dois procedimentos de diagnóstico para avaliar a adição de novas variáveis explicativas no modelo de regressão linear beta. Chien (2012) propõe as medidas de diagnóstico SWR (*standardized weighted residual*) generalizada, LD generalizada, DFFITS generalizada e DFBETAS generalizada, considerando o modelo de regressão linear beta. Ospina e Ferrari (2012) fornecem algumas medidas de diagnóstico para a classe de modelos

proposta. Rocha e Simas (2011) apresentam medidas de influência global e local para a classe de modelos de regressão beta introduzidos por Simas et al. (2010). Pereira e Cribari-Neto (2014) propuseram dois testes gerais para avaliar a má especificação na classe de modelos derivada por Ospina e Ferrari (2012) baseado no teste RESET proposto por Ramsey (1969).

O objetivo desse trabalho é apresentar um novo teste para avaliar a qualidade do ajuste nos modelos de regressão linear beta com dispersão fixa e variável, e para tanto, utilizamos a metodologia proposta por Zhu e Zhang (2004). Além disso, propomos um melhoramento nesses testes baseados na metodologia bootstrap. Baseando-se ainda, na metodologia apresentada por Zhu e Zhang (2004), propusemos um novo teste de hipóteses que é também baseado em influência local, considerando o esquema de perturbação do parâmetro de precisão no modelo de regressão beta com dispersão fixa. A avaliação do desempenho dos testes propostos foi realizada através de estudos de simulações em que são calculadas as taxas de rejeição sob a hipótese nula e poder dos testes quando introduzidas algumas más especificações ao modelo. Por fim, aplicamos a teoria desenvolvida a um conjunto de dados reais.

1.1.1 Organização da Dissertação

A presente dissertação de mestrado está dividida em oito capítulos. No primeiro, além da introdução, apresentamos uma breve revisão sobre a influência local e a metodologia *bootstrap*. No segundo capítulo, a distribuição beta é apresentada juntamente com algumas propriedades, bem como, abordamos os modelos de regressão beta com dispersão fixa e variável. No terceiro, discutimos o teste para avaliar a qualidade do ajuste proposto por Zhu e Zhang (2004), como também, aplicamos o teste às classes de modelos de regressão beta com dispersão fixa e variável. No quarto capítulo há diversos estudos de simulação de Monte Carlo para avaliar o desempenho dos testes de hipóteses propostos citados anteriormente, juntamente com melhoramentos realizados nesses testes baseados na metodologia bootstrap. No quinto capítulo, introduzimos o teste de hipóteses baseado no esquema de perturbação do parâmetro de precisão para o modelo de regressão beta com dispersão fixa. No sexto capítulo, há estudos de simulação de Monte Carlo para avaliar o desempenho do teste apresentado no capítulo anterior. No sétimo, apresentamos aplicações desses testes a dados reais. Por fim, no oitavo capítulo, as conclusões principais do referido trabalho.

1.1.2 Suporte Computacional

As avaliações numéricas realizadas ao longo dessa dissertação foram feitas através da linguagem matricial de programação Ox (linguagem gratuita exclusivamente para uso acadêmico de domínio público disponível no sítio <http://www.doornik.com>) em sua versão 7.0 para sistemas operacionais Linux e Windons, bem como, utilizamos o *software* R (um

software de domínio público disponível no sítio <http://cran.r-projet.org>) para a criação de gráficos.

1.2 Preliminares

Nessa seção, introduziremos alguns conceitos e teorias que serão necessários ao longo desse trabalho.

1.2.1 Influência Local

Há situações em que observações atípicas acabam interferindo no processo de estimação dos parâmetros do modelo. Assim, é de grande importância utilizar algum procedimento teórico que consiga identificar essas observações. Uma teoria que é muito popular na literatura é o método de influência local proposto por Cook (1986), que se baseia em estudar o comportamento de uma certa medida de influência, com objetivo de identificar observações que após pequenas perturbações aplicadas no modelo ou nos dados acabam acarretando mudanças desproporcionais nas estimativas dos parâmetros do modelo. Seja $\ell(\boldsymbol{\theta})$ o logaritmo da função de verossimilhança do modelo correspondente ao modelo suposto, em que $\boldsymbol{\theta} \in \mathbb{R}^k$ é um vetor desconhecido que contém os parâmetros do modelo. Considere $\boldsymbol{\omega} = (\omega_1, \dots, \omega_s)^\top$ restrito a algum subconjunto aberto $D \subset \mathbb{R}$ como sendo um vetor de perturbação que é aplicado no modelo suposto, assim, denote $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ o logaritmo da função de verossimilhança após aplicada o vetor de perturbação para um dado $\boldsymbol{\omega}$. Existe $\boldsymbol{\omega}_0$ (vetor de não perturbação) tal que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$. A medida de influência que é utilizada para avaliar o efeito das perturbações nas estimativas fornecidas pelo modelo é o afastamento da verossimilhança definido por

$$LD(\boldsymbol{\omega}) = 2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}, \quad (1.1)$$

em que $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ são os estimadores de máxima verossimilhança de $\boldsymbol{\theta}$ para o modelo suposto e perturbado, respectivamente. Além disso, $LD(\boldsymbol{\omega}) \geq 0$ e $LD(\boldsymbol{\omega}_0) = 0$. Portanto, através da medida apresentada em (1.1) podemos avaliar a influência sobre a estimativa de $\boldsymbol{\theta}$ ao variarmos $\boldsymbol{\omega}$ através de D , porém, avaliar o comportamento de $LD(\boldsymbol{\omega})$ para todo $\boldsymbol{\omega} \in D$ é geralmente inviável. Dessa forma, Cook (1986) propôs analisar o comportamento de $LD(\boldsymbol{\omega})$ em torno de $\boldsymbol{\omega}_0$, pois significa observar o comportamento de $LD(\boldsymbol{\omega})$ na ausência de perturbação no modelo suposto. Notemos que qualquer valor $\boldsymbol{\omega}$ em uma vizinhança de $\boldsymbol{\omega}_0$ representa um incremento em $LD(\boldsymbol{\omega}_0)$, então este procedimento descreve a sensibilidade de $\ell(\hat{\boldsymbol{\theta}})$ com relação a uma pequena perturbação introduzida em $\ell(\boldsymbol{\theta})$. Objetivando analisar esta sensibilidade, Cook (1986) propõe a avaliação da curvatura normal de $LD(\boldsymbol{\omega}_0 + a\mathbf{d})$ versus a , em que $a \in \mathbb{R}$ e \mathbf{d} é uma direção de norma igual a um, ou seja, $\|\mathbf{d}\| = 1$. O

autor mostrou que a curvatura normal na direção \mathbf{d} é dada por

$$C_d(\boldsymbol{\theta}) = \mathbf{d}^\top \mathbf{H}_{\text{LD}(\omega)} \mathbf{d}, \quad (1.2)$$

com

$$\mathbf{H}_{\text{LD}(\omega)} = 2\Delta^\top (-\ddot{\ell}_{\theta\theta}^{-1})\Delta, \quad (1.3)$$

em que $\ddot{\ell}_{\theta\theta} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$ e $\Delta = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top$ ambos avaliados em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$. Dessa forma, Cook (1986) propôs a construção do gráfico de \mathbf{d}_{\max} que é o autovetor associado ao maior autovalor da matriz $B = -\Delta^\top \ddot{\ell}_{\theta\theta}^{-1} \Delta$. Notemos que \mathbf{d}_{\max} é a direção unitária de perturbação que implica na máxima mudança local em $\text{LD}(\boldsymbol{\omega}_0) = 0$. Logo, o gráfico de \mathbf{d}_{\max} em função dos índices pode revelar quais observações possuem maior influência local, simultaneamente em $\hat{\boldsymbol{\theta}}$, ou seja, quais observações podem estar interferindo, conjuntamente, no processo de estimação de $\boldsymbol{\theta}$. Podem existir, também observações que são, individualmente, influentes, não sendo identificadas através da análise dos valores de \mathbf{d}_{\max} . Dessa forma, uma proposta foi feita por Lesaffre e Verbeke (1998) que sugerem avaliar a curvatura na direção do i -ésimo indivíduo que é definida por (1.2) porém a expressão de $\mathbf{H}_{\text{LD}(\omega)}$ fica dada por

$$\mathbf{H}_{\text{LD}(\omega)} = 2\Delta_i^\top (-\ddot{\ell}_{\theta\theta}^{-1})\Delta_i, \quad (1.4)$$

em que Δ_i é a i -ésima coluna de Δ com $i = 1, 2, \dots, n$. Logo, torna-se interessante a construção de ambos os gráficos de \mathbf{d}_{\max} e de $C_d(\boldsymbol{\theta})$ para identificar observações que são influentes na estimação de $\boldsymbol{\theta}$, tanto conjuntamente como individualmente.

Ainda, podemos avaliar a influência local apenas para parte do vetor de parâmetros. Consideremos uma partição de $\boldsymbol{\theta}$, de modo que $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$. Assim, considerando o subvetor $\boldsymbol{\theta}_1$, a curvatura normal na direção \mathbf{d} fica dada por $C_d(\boldsymbol{\theta}) = 2\mathbf{d}^\top \Delta^\top (-\ddot{\ell}_{\theta\theta}^{-1} - B_1)\Delta \mathbf{d}$, com

$$B_1 = \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\ell}_{\theta_2\theta_2}^{-1} \end{bmatrix},$$

em que $\ddot{\ell}_{\theta_2\theta_2} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2^\top$. Por conseguinte, o gráfico de \mathbf{d}_{\max}^* que é autovetor associado ao maior autovalor da matriz $B^* = \Delta^\top (-\ddot{\ell}_{\theta\theta}^{-1} - B_1)\Delta$ versus os índices das observações pode indicar quais observações possuem maior influência local em $\hat{\boldsymbol{\theta}}_1$. Procedemos de forma análoga para o subvetor $\boldsymbol{\theta}_2$.

Segundo Zhu e Zhang (2004), existem diversas medidas de influência que são relacionadas com a decomposição espectral de $\mathbf{H}_{\text{LD}(\omega)}$. Suponha que a decomposição espectral de $\mathbf{H}_{\text{LD}(\omega)}$ é dada por

$$\mathbf{H}_{\text{LD}(\omega)} = \sum_{m=1}^n \lambda_m \mathbf{v}_m \mathbf{v}_m^\top,$$

em que $(\lambda_m, \mathbf{v}_m)$, $m = 1, 2, \dots, n$, são os autovalores com os respectivos autovetores da matriz $\mathbf{H}_{LD(\boldsymbol{\omega})}$ com $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq \lambda_{k+1} = \dots = \lambda_n = 0$ e $\mathbf{v}_m = (v_{m1}, v_{m2}, \dots, v_{mk})^\top$. De (1.2) e utilizando a decomposição espectral de $\mathbf{H}_{LD(\boldsymbol{\omega})}$ temos

$$\begin{aligned} C_d(\boldsymbol{\omega}) &= \mathbf{d}^\top \mathbf{H}_{LD(\boldsymbol{\omega})} \mathbf{d} \\ &= \mathbf{d}^\top \sum_{m=1}^n \lambda_m \mathbf{v}_m \mathbf{v}_m^\top \mathbf{d} \\ &= \sum_{m=1}^n \lambda_m \mathbf{d}^\top \mathbf{v}_m \mathbf{v}_m^\top \mathbf{d}. \end{aligned}$$

Considerando a direção canônica, isto é, $\mathbf{d} = \mathbf{e}_i$ temos

$$\begin{aligned} C_{e_i} &= \sum_{m=1}^n \lambda_m \mathbf{e}_i^\top \mathbf{v}_m \mathbf{v}_m^\top \mathbf{e}_i \\ &= \sum_{m=1}^n \lambda_m v_{mi} v_{mi} \\ &= \sum_{m=1}^n \lambda_m v_{mi}^2. \end{aligned}$$

Ainda, pela decomposição espectral temos que $\lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_n = 0$, logo,

$$C_{e_i} = \sum_{m=1}^k \lambda_m v_{mi}^2. \quad (1.5)$$

A quantidade C_{e_i} é a medida de influência que utilizaremos para o desenvolvimento do teste de hipóteses para avaliar a qualidade do ajuste em modelos de regressão, o qual será apresentado nos próximos capítulos como foi mencionado.

1.2.2 Bootstrap

A metodologia *bootstrap* foi introduzida por Efron (1979) e é baseada na ideia de considerar uma certa amostra observada da população de interesse como sendo uma população finita. Dessa forma, através de um processo de reamostragem estimamos a distribuição da verdadeira população. Existem dois tipos de *bootstrap*, a saber: paramétrico e o não-paramétrico. Nesse trabalho, abordaremos apenas o *bootstrap* paramétrico, visto que consideraremos a população seguindo uma distribuição beta.

O *bootstrap* é uma ferramenta muito útil quando há a necessidade de estimar a distribuição de uma certa estatística, isto é, através dessa técnica, podemos obter a distribuição empírica de qualquer estatística que se tenha interesse. A forma usual para realizar isto é através do *bootstrap* percentil. Para tanto, segundo Rizzo (2008), consideremos \mathbf{y} como sendo uma certa amostra obtida da população em estudo. Seja $\hat{\theta}$ um estimador (a estatís-

tica de interesse) para o parâmetro de interesse θ . Assim, utilizamos a amostra original \mathbf{y} com sua respectiva estimativa de $\hat{\theta}$ e geramos de forma paramétrica (através de uma distribuição de probabilidade conhecida) R amostras *bootstrap* denotadas por $\mathbf{y}^{*1}, \mathbf{y}^{*2}, \dots, \mathbf{y}^{*R}$. A partir das amostras, obtemos as R réplicas de $\hat{\theta}$ denotadas por $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*R}$. Logo, notemos que essas R réplicas constituem a função de distribuição empírica (fde) de $\hat{\theta}$, que denotaremos por \hat{F} .

Se o interesse pela fde da estatística for de construir um intervalo de confiança para o parâmetro θ , basta considerar os percentis $(\alpha/2)$ e $(1 - \alpha/2)$ de \hat{F} e, assim, obtemos o intervalo percentil *bootstrap* com confiança de $(1 - \alpha)100\%$ dado por

$$\left[\hat{F}^{-1}(\alpha/2), \hat{F}^{-1}(1 - \alpha/2) \right],$$

em que $\hat{F}^{-1}(\alpha/2) = \hat{\theta}^{*(\alpha/2)}$ e $\hat{F}^{-1}(1 - \alpha/2) = \hat{\theta}^{*(1-\alpha/2)}$. E, esse intervalo pode ser reescrito da seguinte forma:

$$\left[\hat{\theta}^{*(\alpha/2)}, \hat{\theta}^{*(1-\alpha/2)} \right]. \quad (1.6)$$

Ainda, através dos percentis $\hat{\theta}^{*(\alpha/2)}$ e $\hat{\theta}^{*(1-\alpha/2)}$ podemos testar as hipóteses $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, pois dado α fixo, se $\theta_0 < \hat{\theta}^{*(\alpha/2)}$ ou $\theta_0 > \hat{\theta}^{*(1-\alpha/2)}$ então rejeitamos H_0 . Segundo Rizzo (2008), o intervalo (1.6) possui melhor cobertura quando a distribuição de $\hat{\theta}$ não é normal. Observamos que pela forma da construção desse intervalo, não é possível a presença de valores impróprios para o parâmetro θ no intervalo, como também não é necessariamente simétrico.

A metodologia da construção do intervalo percentil incorpora a transformação que normaliza a distribuição de $\hat{\theta}$ (Efron e Tibshirani, 1993). Suponhamos que $\hat{\theta}$ não possui distribuição normal em amostras finitas, e seja $\hat{\rho} = h(\hat{\theta})$ a transformação que normaliza a distribuição de $\hat{\theta}$, isto é,

$$\hat{\rho} \sim N(\rho, \tau^2),$$

em que τ é um desvio-padrão constante e $N(a, b)$ denota a distribuição normal de parâmetros a e b . Dessa maneira, podemos construir um intervalo de confiança exato com nível de confiança $(1 - \alpha)100\%$ que é dado por

$$\left[\hat{\rho} - z_{(1-\alpha/2)}\tau, \hat{\rho} + z_{(1-\alpha/2)}\tau \right],$$

em que $z_{(1-\alpha/2)}$ é obtido de $P(-z_{(1-\alpha/2)} < Z < z_{(1-\alpha/2)}) = 1 - \alpha$ com $Z \sim N(0, 1)$. Logo, o intervalo percentil pode ser reescrito como

$$\left[h^{-1}(\hat{\rho} - z_{(1-\alpha/2)}\tau), h^{-1}(\hat{\rho} + z_{(1-\alpha/2)}\tau) \right],$$

sem a necessidade de conhecer a função transformadora $h(\cdot)$.

Muitas vezes, também não conhecemos o erro-padrão da estatística de interesse, $\text{ep}(\hat{\theta})$. Uma solução para esse problema é utilizar o método *bootstrap* para estimar o erro-padrão, ou seja, calcular o erro-padrão das R réplicas *bootstrap*.

Capítulo 2

Modelos de Regressão Beta

É muito comum em estudos das mais diversas áreas tais como: biologia, ciências sociais, engenharia, estatística, entre outras, a busca pelo conhecimento de como variáveis de interesse do estudo se comportam condicionadas a valores de outras variáveis. Uma ferramenta teórica muito utilizada que soluciona esse tipo de questionamento é a análise de regressão, que a partir de um modelo matemático faz uso de um certo conjunto de variáveis denominadas de covariáveis para tentar explicar o comportamento da variável de interesse que é, geralmente, denominada como resposta. Outra situação muito frequente é que a variável de interesse, seja do tipo percentagem ou proporção, ou seja, a variável resposta possui suporte no intervalo $(0, 1)$ (ver Seção 1.1). Nesse sentido, Ferrari e Cribari-Neto (2004) desenvolveram a classe de modelos de regressão linear beta, que é baseada na suposição de que a variável resposta tem distribuição beta com uma parametrização que é indexada pela média e por um parâmetro de precisão. A seguir, realizamos uma breve apresentação da distribuição beta com algumas de suas propriedades, bem como, o modelo de regressão linear beta.

2.1 Distribuição Beta

Consideremos Y uma variável aleatória contínua, dizemos que Y tem distribuição beta com parâmetros $p, q > 0$ se sua função densidade de probabilidade (fdp) é dada por

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1},$$

em que $0 < y < 1$ e $\Gamma(\cdot)$ é a função gama definida por

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt.$$

Denotaremos por $Y \sim \text{Beta}(p, q)$.

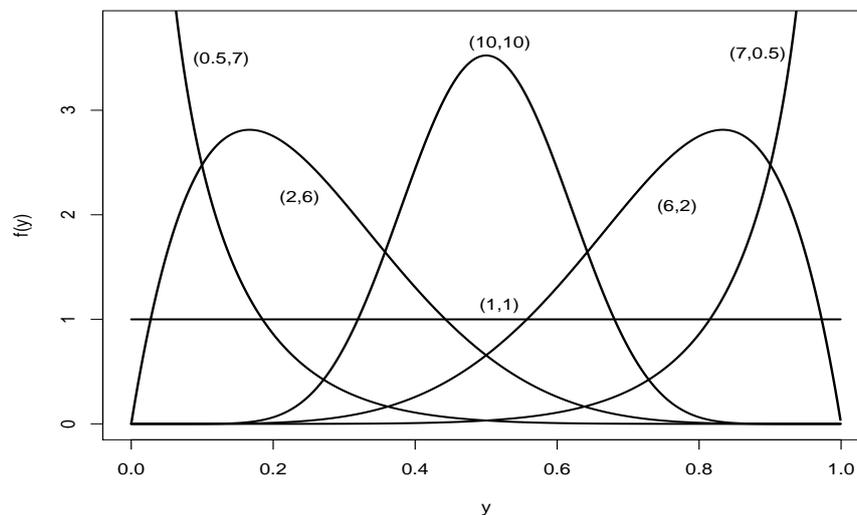
A função de distribuição acumulada (fda) de Y é obtida através de

$$\begin{aligned}
 F(y; p, q) &= P(Y \leq y) \\
 &= \int_0^y \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} t^{p-1} (1-t)^{q-1} dt \\
 &= \frac{1}{B(p, q)} \int_0^y t^{p-1} (1-t)^{q-1} dt \\
 &= \frac{B_y(p, q)}{B(p, q)} \\
 &= I_y(p, q),
 \end{aligned}$$

em que $B(\cdot, \cdot)$ é a função beta, $B_y(\cdot, \cdot)$ e $I_y(\cdot, \cdot)$ são a função beta e a razão beta incompletas, respectivamente.

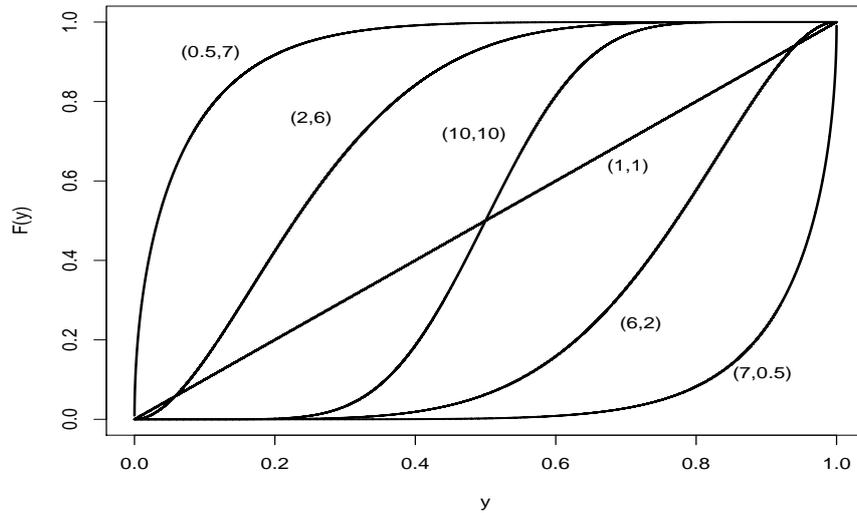
A densidade beta pode apresentar várias formas diferentes, de acordo com os valores assumidos para p e q . Esse comportamento pode ser visto nas Figuras 2.1 e 2.2 em que são apresentadas algumas curvas da fdp e da fda, respectivamente. Notemos que há curvas para a fdp da distribuição beta, sendo constante, $p = q = 1$, simétrica ($p = q = 10$), assimétrica à esquerda ($p = 6, q = 2$), assimétrica à direita ($p = 2, q = 6$) e, estritamente, não-crescente ($p = 0.5, q = 7$ e $p = 7, q = 0.5$). Analisando a Figura 2.2 observamos que os valores de (p, q) com $p < q$ levam a maiores valores assumidos pela fda, como também, valores de (p, q) com $p > q$ levam a menores valores assumidos pela fda.

Figura 2.1: Curvas de densidade da distribuição beta para alguns valores de p e q .



Fonte: A autora.

Figura 2.2: Curvas da fda da distribuição beta para alguns valores de p e q .



Fonte: A autora.

Os momentos de Y são obtidos a partir de

$$\begin{aligned}
 E(Y^k) &= \int_0^1 y^k \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1} dy & (2.1) \\
 &= \int_0^1 \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p+k-1} (1-y)^{q-1} dy \\
 &= \frac{\Gamma(p+k)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+k)} \int_0^1 \frac{\Gamma(p+q+k)}{\Gamma(p+k)\Gamma(q)} y^{p+k-1} (1-y)^{q-1} dy \\
 &= \frac{\Gamma(p+k)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+k)},
 \end{aligned}$$

em que $E(\cdot)$ denota a esperança matemática (média) de uma variável aleatória. Dessa forma, a média de Y é dada por

$$\begin{aligned}
 E(Y) &= \frac{\Gamma(p+1)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+1)} \\
 &= \frac{p\Gamma(p)\Gamma(p+q)}{\Gamma(p)(p+q)\Gamma(p+q)} \\
 &= \frac{p}{p+q}. & (2.2)
 \end{aligned}$$

O segundo momento de Y é dado por

$$\begin{aligned}
E(Y^2) &= \frac{\Gamma(p+2)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+2)} \\
&= \frac{(p+1)\Gamma(p+1)\Gamma(p+q)}{\Gamma(p)(p+q+1)\Gamma(p+q+1)} \\
&= \frac{p(p+1)\Gamma(p)\Gamma(p+q)}{\Gamma(p)(p+q+1)(p+q)\Gamma(p+q)} \\
&= \frac{p(p+1)}{(p+q)(p+q+1)}.
\end{aligned}$$

Assim, a variância de Y fica dada por

$$\begin{aligned}
\text{Var}(Y) &= \frac{p(p+1)}{(p+q)(p+q+1)} - \frac{p^2}{(p+q)^2} \\
&= \frac{p(p+1)(p+q) - p^2(p+q+1)}{(p+q)^2(p+q+1)} \\
&= \frac{pq}{(p+q)^2(p+q+1)}, \tag{2.3}
\end{aligned}$$

em que $\text{Var}(\cdot)$ denota a variância de uma variável aleatória.

Através da expressão (2.2), podemos obter as medidas de assimetria e curtose para Y , dadas, respectivamente, por

$$E[(Y - \mu)^3] = \frac{2(q-p)\sqrt{p^{-1} + q^{-1} + (pq)^{-1}}}{p+q+2}$$

e

$$E[(Y - \mu)^4] = \frac{3(p+q+1)[2(p+q)^2 + pq(p+q-6)]}{pq(p+q+2)(p+q+3)}.$$

Essas quantidades medem a assimetria e a grossura das caudas da distribuição, respectivamente.

Para apresentarmos a classe de modelos de regressão beta na próxima seção, é necessário antes considerar uma parametrização diferente da apresentada para a distribuição beta. Dessa forma, seja $\mu = p/(p+q)$ e $\phi = p+q$, ou seja, $p = \mu\phi$ e $q = (1-\mu)\phi$. Assim, de (2.2) e (2.3) temos que

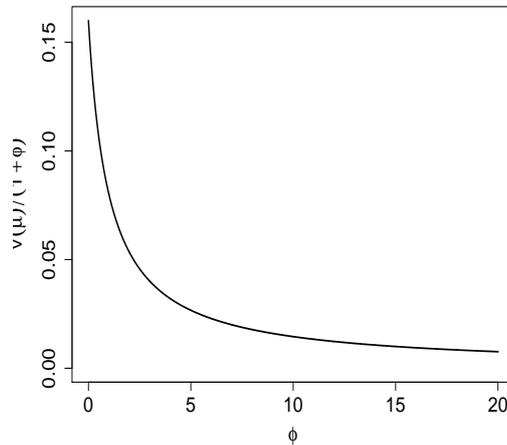
$$\begin{aligned}
E(Y) &= \frac{\mu\phi}{\mu\phi + (1-\mu)\phi} \\
&= \mu.
\end{aligned}$$

Também,

$$\begin{aligned}
 \text{Var}(Y) &= \frac{\mu\phi(1-\mu)\phi}{(\mu\phi + (1-\mu)\phi)^2(\mu\phi + (1-\mu)\phi + 1)} \\
 &= \frac{\mu(1-\mu)}{1+\phi} \\
 &= \frac{V(\mu)}{1+\phi},
 \end{aligned} \tag{2.4}$$

em que $V(\mu) = \mu(1-\mu)$ é denominada de função de variância e ϕ é interpretado como um parâmetro de precisão, isto é, para dado μ , se o valor de ϕ diminui então a variância da variável resposta aumenta e vice-versa, conforme está ilustrado na Figura 2.3.

Figura 2.3: Comportamento da variância dado μ fixo para alguns valores de ϕ .



Fonte: A autora.

Dessa forma, utilizando a nova parametrização, a fdp fica dada por

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \tag{2.5}$$

$0 < \mu < 1$ e $\phi > 0$.

Um resultado importante é a distribuição beta pertencer à família exponencial. De fato, segundo Lehmann e Casella (1998), uma família $\mathbb{P} = \{p_\theta, \theta \in \Theta \subset \mathbb{R}^s\}$ de distribuições forma uma família exponencial s -dimensional se as densidades em \mathbb{P} podem ser escritas como

$$p_\eta(y) = \exp \left\{ \sum_{j=1}^s \eta_j T_j - \mathbb{C}(\boldsymbol{\eta}) \right\} h(y),$$

em que η_j , $j = 1, 2, \dots, s$, e $\mathbb{C}(\boldsymbol{\eta})$ são funções reais de θ , T_j 's são funções a valores reais definidas sobre o suporte da distribuição e $h(y)$ é definida sobre o suporte da distribuição

a valores reais e não negativos.

Assim, reescrevendo a expressão (2.5) temos

$$f(y; \mu, \phi) = \exp \left\{ \log \left(\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} \right) + \mu\phi \log \left(\frac{y}{1-y} \right) + \phi \log(1-y) \right\} \frac{1}{y(1-y)},$$

em que $s = 2$, $\eta_1 = \mu\phi$, $\eta_2 = \phi$, $T_1 = \log(y/(1-y))$, $T_2 = \log(1-y)$, $h(y) = 1/(y(1-y))$ e

$$\begin{aligned} \mathbb{C}(\boldsymbol{\eta}) &= -\log \left(\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} \right) \\ &= -\log \Gamma(\eta_2) + \log \Gamma(\eta_1) + \log \Gamma(\eta_2 - \eta_1). \end{aligned}$$

Concluimos assim que Y pertence à família exponencial.

Assim, por resultados da família exponencial obtemos que

$$\begin{aligned} E(T_1) &= E \left(\log \left(\frac{y}{1-y} \right) \right) \\ &= \frac{\partial \mathbb{C}(\boldsymbol{\eta})}{\partial \eta_1} \\ &= \psi(\eta_1) - \psi(\eta_2 - \eta_1) \\ &= \psi(\mu\phi) - \psi((1-\mu)\phi), \end{aligned} \tag{2.6}$$

e ainda,

$$\begin{aligned} \text{Var}(T_1) &= \text{Var} \left(\log \left(\frac{y}{1-y} \right) \right) \\ &= \frac{\partial^2 \mathbb{C}(\boldsymbol{\eta})}{\partial \eta_1^2} \\ &= \psi'(\eta_1) + \psi'(\eta_2 - \eta_1) \\ &= \psi'(\mu\phi) + \psi'((1-\mu)\phi). \end{aligned} \tag{2.7}$$

Os resultados em (2.6) e (2.7) serão muito úteis em cálculos referentes aos modelos de regressão beta apresentados nas próximas seções.

2.2 Modelo de Regressão Beta com Dispersão Fixa

Utilizando a parametrização da distribuição beta dada em (2.5), Ferrari e Cribari-Neto (2004) obtiveram uma classe de modelos de regressão que permite modelar a média de uma variável resposta do tipo percentagem ou proporção sujeita a um parâmetro de precisão. Sejam y_1, y_2, \dots, y_n variáveis aleatórias independentes, em que cada y_i , $i = 1, 2, \dots, n$, segue a densidade em (2.5) com média μ_i e precisão desconhecida ϕ . No modelo proposto

é suposto que a média de y_i pode ser escrita a partir de

$$g(\mu_i) = \sum_{j=1}^k \beta_j x_{ij} = \eta_i, \quad (2.8)$$

em que $j = 1, 2, \dots, k$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)^\top$ é um vetor de parâmetros de regressão desconhecido ($\beta \in \mathbb{R}^k$), η_i é um preditor linear, $x_{i1}, x_{i2}, \dots, x_{ik}$ são observações de $k < n$ variáveis explicativas da média, assumidas fixas e conhecidas e $g(\cdot)$ é uma função de ligação que deve ser estritamente monótona e duas vezes diferenciável. Notemos a partir de (2.8) e (2.4) que $\mu_i = g^{-1}(\eta_i)$ e $\text{Var}(y_i) = V(g^{-1}(\eta_i))/(1 + \phi)$, o que significa que mesmo que o parâmetro de precisão seja constante, a variância da variável resposta não é constante, pois depende da média.

De (2.5), segue que o logaritmo da função de verossimilhança é

$$\ell(\boldsymbol{\beta}, \phi) = \sum_{i=1}^n \ell_i(\mu_i, \phi), \quad (2.9)$$

em que

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

Dessa forma, as entradas do vetor escore para $\boldsymbol{\beta}$ são obtidas de

$$U_{\beta_r}(\boldsymbol{\beta}, \phi) = \phi \sum_{i=1}^n x_{ir} t_i (y_i^* - \mu_i^*), \quad (2.10)$$

em que x_{ir} é o i -ésimo valor da r -ésima covariável, $i = 1, 2, \dots, n$, $r = 1, 2, \dots, k$, $t_i = [g'(\mu_i)]^{-1}$, $\psi(\cdot)$ é a função digama definida por $\psi(z) = d[\log \Gamma(z)]/dz$ e

$$y_i^* = \log \left(\frac{y_i}{1 - y_i} \right) \quad \text{e} \quad \mu_i^* = \psi(\mu_i \phi) - \psi((1 - \mu_i) \phi). \quad (2.11)$$

Assim, o vetor escore para $\boldsymbol{\beta}$ é dado por

$$U_{\boldsymbol{\beta}}(\boldsymbol{\beta}, \phi) = \phi X^\top T(\mathbf{y}^* - \boldsymbol{\mu}^*), \quad (2.12)$$

em que X é uma matriz $n \times k$ cuja j -ésima coluna é o vetor \mathbf{x}_j que representa os valores da j -ésima covariável, $j = 1, 2, \dots, k$ e $T = \text{diag}\{t_1, t_2, \dots, t_n\}$. Percebemos a partir do resultado (2.6) e pelas definições dadas em (2.11) que $E(y_i^*) = \mu_i^*$. Definindo

$$v_i = \psi'(\mu_i \phi) + \psi'((1 - \mu_i) \phi) \quad (2.13)$$

e utilizando o resultado (2.7), obtemos que $\text{Var}(y_i^*) = v_i$.

Já o vetor escore para ϕ é dado por

$$U_\phi(\boldsymbol{\beta}, \phi) = \sum_{i=1}^n [\psi(\phi) + \mu_i(y_i^* - \mu_i^*) - \psi((1 - \mu_i)\phi) + \log(1 - y_i)]. \quad (2.14)$$

Seja $u_i = \psi(\phi) + \mu_i(y_i^* - \mu_i^*) - \psi((1 - \mu_i)\phi) + \log(1 - y_i)$. O vetor escore para ϕ fica dado por

$$U_\phi(\boldsymbol{\beta}, \phi) = \sum_{i=1}^n u_i.$$

Considere agora as segundas derivadas do logaritmo da função de verossimilhança $\ell_i(\mu_i, \phi)$. Para $j = 1, 2, \dots, k$, temos

$$\frac{\partial^2 \ell_i(\mu_i, \phi)}{\partial \beta_r \partial \beta_j} = -\phi x_{ir} x_{ij} q_i,$$

em que $q_i = t_i^2 [t_i g''(\mu_i)(y_i^* - \mu_i^*) + \phi v_i]$. Então, matricialmente, temos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = -\phi X^\top Q X,$$

em que $Q = \text{diag}\{q_1, q_2, \dots, q_n\}$.

Além disso, precisamos também de

$$\frac{\partial^2 \ell_i(\mu_i, \phi)}{\partial \beta_r \partial \phi} = -x_{ir} t_i f_i,$$

em que

$$f_i = c_i - (y_i^* - \mu_i^*) \quad \text{e} \quad c_i = \phi [\psi'(\mu_i \phi) \mu_i - \psi'((1 - \mu_i)\phi)(1 - \mu_i)].$$

Matricialmente, temos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \boldsymbol{\beta} \partial \phi} = -X^\top T \mathbf{f},$$

em que $\mathbf{f} = (f_1, f_2, \dots, f_n)^\top$.

Por fim,

$$\begin{aligned} \frac{\partial^2 \ell_i(\mu_i, \phi)}{\partial \phi^2} &= \psi'(\phi) - [\mu_i^2 \psi'(\mu_i \phi) + (1 - \mu_i)^2 \psi'((1 - \mu_i)\phi)] \\ &= -d_i, \end{aligned}$$

em que $d_i = \mu_i^2 \psi'(\mu_i \phi) + (1 - \mu_i)^2 \psi'((1 - \mu_i)\phi) - \psi'(\phi)$. Matricialmente, tem-se

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \phi^2} = -\text{tr}(D),$$

em que $D = \text{diag}\{d_1, d_2, \dots, d_n\}$.

Dessa forma, obtemos a matriz de segundas derivadas, que é dada por

$$\ddot{\ell} = \ddot{\ell}(\boldsymbol{\beta}, \phi) = \begin{bmatrix} \ddot{\ell}_{\beta\beta} & \ddot{\ell}_{\beta\phi} \\ \ddot{\ell}_{\phi\beta} & \ddot{\ell}_{\phi\phi} \end{bmatrix},$$

em que $\ddot{\ell}_{\beta\beta} = -\phi X^\top Q X$, $\ddot{\ell}_{\beta\phi} = \ddot{\ell}_{\phi\beta}^\top = -X^\top T \mathbf{f}$ e $\ddot{\ell}_{\phi\phi} = -\text{tr}(D)$.

A matriz de informação de Fisher é dada por

$$K = K(\boldsymbol{\beta}, \phi) = \begin{bmatrix} K_{\beta\beta} & K_{\beta\phi} \\ K_{\phi\beta} & K_{\phi\phi} \end{bmatrix},$$

em que as submatrizes são obtidas aplicando menos a esperança na matriz de segundas derivadas. Dessa forma, utilizando o fato de que $E(y_i^*) = \mu_i^*$ temos

$$\begin{aligned} -E\left(\frac{\partial^2 \ell_i(\mu_i, \phi)}{\partial \beta_r \partial \beta_j}\right) &= \phi x_{ir} x_{ij} E(q_i) \\ &= \phi x_{ir} x_{ij} t_i^2 [t_i g''(\mu_i) E(y_i^* - \mu_i^*) + \phi v_i] \\ &= \phi x_{ir} x_{ij} w_i, \end{aligned}$$

em que $w_i = \phi t_i^2 v_i$. Matricialmente, temos

$$K_{\beta\beta} = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top}\right) = \phi X^\top W X,$$

em que $W = \text{diag}\{w_1, w_2, \dots, w_n\}$. Também, obtemos que

$$-E\left(\frac{\partial^2 \ell_i(\mu_i, \phi)}{\partial \beta_r \partial \phi}\right) = x_{ir} t_i E(f_i).$$

Dado que $E(f_i) = c_i$, matricialmente, temos

$$K_{\beta\phi} = K_{\phi\beta}^\top = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \boldsymbol{\beta} \partial \phi}\right) = X^\top T \mathbf{c},$$

em que $\mathbf{c} = (c_1, c_2, \dots, c_n)^\top$.

Por último,

$$K_{\phi\phi} = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \phi)}{\partial \phi^2}\right) = \text{tr}(D).$$

Uma característica importante dessa classe de modelos é que os parâmetros β e ϕ não são ortogonais, diferentemente do que ocorre na classe de MLGs (McCullagh e Nelder, 1989). Ainda, sob algumas condições de regularidade, os estimadores de máxima verossimilhança têm distribuição assintótica conjunta dada por $\hat{\theta} \sim N_{k+1}(\theta, K^{-1})$, em que $\hat{\theta} = (\hat{\beta}, \hat{\phi})^\top$ e a matriz K^{-1} é apresentada a seguir.

Ferrari e Cribari-Neto (2004) apresentam a matriz K^{-1} , a qual é dada por

$$K^{-1} = K^{-1}(\beta, \phi) = \begin{bmatrix} K^{\beta\beta} & K^{\beta\phi} \\ K^{\phi\beta} & K^{\phi\phi} \end{bmatrix},$$

em que

$$K^{\beta\beta} = \frac{1}{\phi}(X^\top W X) \left\{ \mathbb{I}_k + \frac{X^\top T \mathbf{c} \mathbf{c}^\top T^\top X (X^\top W X)^{-1}}{\xi \phi} \right\},$$

em que $\xi = \text{tr}(D) - \phi^{-1} \mathbf{c}^\top T^\top X (X^\top W X)^{-1} X^\top T \mathbf{c}$,

$$K^{\beta\phi} = K^{\phi\beta\top} = -\frac{1}{\xi \phi} (X^\top W X)^{-1} X^\top T \mathbf{c},$$

e $K^{\phi\phi} = \xi^{-1}$, \mathbb{I}_k é a matriz identidade de ordem k .

2.3 Modelo de Regressão Beta com Dispersão Variável

No modelo de regressão beta proposto por Ferrari e Cribari-Neto (2004) é assumido que o parâmetro de precisão ϕ é constante para todas as observações. Porém, o uso deste modelo de regressão quando a precisão ϕ varia de acordo com as observações pode afetar de forma drástica a estimação dos parâmetros da regressão. Smithson e Verkuilen (2006) propuseram uma extensão do modelo de regressão linear beta que admite que a dispersão (ou precisão) também possui estrutura de regressão. A seguir apresentamos a classe de modelos de regressão beta com dispersão variável.

Sejam y_1, y_2, \dots, y_n variáveis aleatórias independentes, em que cada y_i , $i = 1, 2, \dots, n$, segue a densidade em (2.5) com média μ_i e precisão ϕ_i . No modelo proposto a média de y_i continua tendo a estrutura apresentada em (2.8), no entanto, a precisão de y_i segue a estrutura heteroscedástica dada por

$$h(\phi_i) = \sum_{j=1}^q \gamma_j z_{ij} = \vartheta_i, \quad (2.15)$$

em que $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_q)^\top$ é um vetor de parâmetros de regressão desconhecido ($\gamma \in \mathbb{R}^q$), ϑ_i é um preditor linear, $z_{t1}, z_{t2}, \dots, z_{tq}$ são observações de $q < n$ variáveis explicativas da dispersão, assumidas fixas e conhecidas e $h(\cdot)$ é a função de ligação para a dispersão que deve ser estritamente monótona e duas vezes diferenciável.

De (2.5) segue que o logaritmo da função de verossimilhança é

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^n \ell_i(\mu_i, \phi_i), \quad (2.16)$$

em que

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

Dessa forma, as entradas do vetor escore para $\boldsymbol{\beta}$ são obtidas de

$$U_{\beta_r}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^n \phi_i x_{ir} t_i (y_i^* - \mu_i^{**}), \quad (2.17)$$

em que x_{ir} é o i -ésimo valor da r -ésima covariável relacionada à média, $i = 1, 2, \dots, n$, $r = 1, 2, \dots, k$, $t_i = [g'(\mu_i)]^{-1}$, y_i^* está definido em (2.11) e

$$\mu_i^{**} = \psi(\mu_i \phi_i) - \psi((1 - \mu_i) \phi_i). \quad (2.18)$$

Assim, o vetor escore para $\boldsymbol{\beta}$ é dado por

$$U_{\boldsymbol{\beta}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = X^{\top} \Phi T (y^* - \mu^{**}),$$

em que X é uma matriz $n \times k$ cuja j -ésima coluna é o vetor \boldsymbol{x}_j que representa os valores da j -ésima covariável, $j = 1, 2, \dots, k$, $T = \text{diag}\{t_1, t_2, \dots, t_n\}$ e $\Phi = \text{diag}\{\phi_1, \phi_2, \dots, \phi_n\}$. Ressaltamos a partir do resultado (2.6) e pela definição dada em (2.18) temos que $E(y_i^*) = \mu_i^{**}$. Definindo

$$v_i^* = \psi'_i(\mu_i \phi_i) + \psi'_i((1 - \mu_i) \phi_i),$$

e utilizando o resultado (2.7), obtemos que $\text{Var}(y_i^*) = v_i^*$.

Já as entradas do vetor escore para $\boldsymbol{\gamma}$ são obtidas de

$$U_{\gamma_p}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^n z_{ip} t_i^* u_i^*, \quad (2.19)$$

em que z_{ip} é o i -ésimo valor da p -ésima covariável relacionada à dispersão, $i = 1, 2, \dots, n$, $p = 1, 2, \dots, q$, $t_i^* = [h'(\phi_i)]^{-1}$ e

$$u_i^* = \psi(\phi_i) + \mu_i(y_i^* - \mu_i^{**}) - \psi((1 - \mu_i) \phi_i) + \log(1 - y_i).$$

Assim, o vetor escore para $\boldsymbol{\gamma}$ é dado por

$$U_{\boldsymbol{\gamma}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = Z^{\top} T^* \mathbf{u}^*,$$

em que Z é uma matriz $n \times q$ cuja j -ésima coluna é o vetor \mathbf{z}_j que representa os valores da j -ésima covariável da dispersão, $j = 1, 2, \dots, q$, $\mathbf{u}^* = (u_1, u_2, \dots, u_n)^{\top}$ e $T^* = \text{diag}\{t_1^*, t_2^*, \dots, t_n^*\}$.

Consideremos agora as segundas derivadas do logaritmo da função de verossimilhança $\ell_i(\mu_i, \phi_i)$ para $j = 1, 2, \dots, k$:

$$\frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \beta_r \partial \beta_j} = -\phi_i x_{ir} x_{ij} q_i^*,$$

em que $q_i^* = t_i^2 [t_i g''(\mu_i)(y_i^* - \mu_i^{**}) + \phi_i v_i^*]$. Então, matricialmente, temos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^{\top}} = -X^{\top} \Phi Q^* X,$$

em que $Q^* = \text{diag}\{q_1^*, q_2^*, \dots, q_n^*\}$.

Além disso, precisamos de

$$\frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \beta_r \partial \gamma_p} = -x_{ir} z_{ip} t_i^* f_i^*,$$

em que

$$f_i^* = c_i^* - (y_i^* - \mu_i^{**}) \quad \text{e} \quad c_i^* = \phi_i [\psi'(\mu_i \phi_i) \mu_i - \psi'((1 - \mu_i) \phi_i) (1 - \mu_i)].$$

Matricialmente, temos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}^{\top}} = -X^{\top} T T^* F^* Z,$$

em que $F = \text{diag}\{f_1^*, f_2^*, \dots, f_n^*\}$.

Por fim,

$$\frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \gamma_p \partial \gamma_k} = -z_{ip} z_{ik} m_i,$$

em que $m_i = t_i^{*2} [t_i^* h''(\phi_i) u_i^* + d_i^*]$, $h''(\cdot)$ é a segunda derivada, e

$$d_i^* = \mu_i^2 \psi'(\mu_i \phi_i) + (1 - \mu_i)^2 \psi'((1 - \mu_i) \phi_i) - \psi'(\phi_i).$$

Matricialmente, temos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma} \boldsymbol{\gamma}^\top} = -Z^\top M Z,$$

em que $M = \text{diag}\{m_1, m_2, \dots, m_n\}$.

Dessa forma, obtemos a matriz de segundas derivadas dada por

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

em que $\ddot{\ell}_{\beta\beta} = -X^\top \Phi Q^* X$, $\ddot{\ell}_{\beta\gamma} = \ddot{\ell}_{\gamma\beta}^\top = -X^\top T T^* F^* Z$ e $\ddot{\ell}_{\gamma\gamma} = -Z^\top M Z$.

A matriz de informação de Fisher é dada por

$$K^* = K^*(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{bmatrix} K_{\beta\beta}^* & K_{\beta\gamma}^* \\ K_{\gamma\beta}^* & K_{\gamma\gamma}^* \end{bmatrix},$$

em que as submatrizes são obtidas, aplicando menos a esperança na matriz de segundas derivadas. Dessa forma, utilizando o fato de que $E(y_i^*) = \mu_i^{**}$ temos

$$\begin{aligned} -E\left(\frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \beta_r \partial \beta_j}\right) &= \phi x_{ir} x_{ij} E(q_i^*) \\ &= \phi x_{ir} x_{ij} t_i^2 [t_i g''(\mu_i) E(y_i^* - \mu_i^{**}) + \phi_i v_i^*] \\ &= \phi x_{ir} x_{ij} w_i^*, \end{aligned}$$

em que $w_i^* = \phi_i t_i^2 v_i^*$. Matricialmente, temos

$$K_{\beta\beta}^* = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top}\right) = X^\top \Phi W^* X,$$

em que $W^* = \text{diag}\{w_1^*, w_2^*, \dots, w_n^*\}$. Também, temos que

$$-E\left(\frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \beta_r \partial \gamma_p}\right) = x_{ir} z_{ip} t_i^* E(f_i^*),$$

como $E(f_i^*) = c_i^*$, matricialmente,

$$K_{\beta\gamma}^* = K_{\gamma\beta}^{*\top} = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}^\top}\right) = X^\top T T^* C^* Z.$$

Por último,

$$K_{\gamma\gamma}^* = -E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma} \boldsymbol{\gamma}^\top}\right) = Z^\top T^{*2} D^* Z.$$

Ainda, sob algumas condições de regularidade, os estimadores de máxima verossimilhança

têm distribuição assintótica conjunta dada por $\hat{\boldsymbol{\theta}} \sim N_{k+q}(\boldsymbol{\theta}, K^{*-1})$ em que $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}, \boldsymbol{\gamma})^\top$ e K^{*-1} é apresentada a seguir.

Espinheira et al. (2007) apresentam a matriz K^{*-1} , a qual é dada por

$$K^{*-1} = K^{*-1}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{bmatrix} K_*^{\beta\beta} & K_*^{\beta\gamma} \\ K_*^{\gamma\beta} & K_*^{\gamma\gamma} \end{bmatrix},$$

em que

$$K_*^{\beta\beta} = (X^\top \Phi W^* X - X^\top C T T^* Z (Z^\top T^{*2} D^* Z)^{-1} Z^\top T^* T C^\top X)^{-1},$$

$$K_*^{\beta\gamma} = K_*^{\beta\gamma^\top} = -K_*^{\beta\beta} X^\top C T T^* Z (Z^\top T^{*2} D^* Z)^{-1},$$

e

$$K_*^{\gamma\gamma} = (Z^\top T^{*2} D^* Z)^{-1} \{ \mathbb{I}_q + (Z^\top T^* T C^\top X) K_*^{\beta\beta} X^\top C T T^* Z (Z^\top T^{*2} D^* Z)^{-1} \},$$

em que \mathbb{I}_q é a matriz identidade de ordem q .

Capítulo 3

Teste Baseado em Influência Local

Testes de hipóteses são de grande importância em situações quando necessitamos tomar decisões no sentido de rejeitar ou não uma afirmação (hipótese) que fizemos sobre a população de interesse, baseando-se em um conjunto de informações obtidas através de dados amostrais. Nesse capítulo, primeiramente, apresentaremos o teste de hipóteses baseado em influência local proposto por Zhu e Zhang (2004) e uma proposta de melhoramento do mesmo. Em seguida, aplicaremos o teste aos modelos de regressão beta com dispersão fixa e variável.

3.1 Teste de Hipóteses Baseado em Influência Local

Para introduzirmos o teste proposto por Zhu e Zhang (2004), apresentaremos algumas suposições e resultados que serão necessários.

Consideremos y_1, y_2, \dots, y_n variáveis aleatórias independentes e identicamente distribuídas (iid) com fdp $g(y)$. Suponhamos que os dados vêm de uma família de distribuições pertencente a

$$\mathcal{F} = \{p(y; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\},$$

em que $\boldsymbol{\theta}$ tem dimensão p e a verdadeira densidade $g(y)$ pode não pertencer. White (1982), sob certas condições, estabeleceu a consistência de $\hat{\boldsymbol{\theta}}$, isto é, para $n \rightarrow \infty$

$$\hat{\boldsymbol{\theta}} \xrightarrow{\mathbb{P}} \boldsymbol{\theta}_* = \arg \min_{\boldsymbol{\theta} \in \Theta} \mathbb{E} \left(\log \frac{g(y)}{p(y; \boldsymbol{\theta})} \right),$$

em que $\xrightarrow{\mathbb{P}}$ denota convergência em probabilidade. Além disso, o autor também estabeleceu a normalidade assintótica de $\hat{\boldsymbol{\theta}}$, isto é,

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

em que \xrightarrow{d} denota convergência em distribuição, com $C(\boldsymbol{\theta}_*) = [A(\boldsymbol{\theta}_*)]^{-1}B(\boldsymbol{\theta}_*)[A(\boldsymbol{\theta}_*)]^{-1}$

sendo

$$A(\boldsymbol{\theta}) = -E(\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})) \quad \text{e} \quad B(\boldsymbol{\theta}) = E(\dot{\ell}_{\theta}(\boldsymbol{\theta})\dot{\ell}_{\theta}^{\top}(\boldsymbol{\theta})). \quad (3.1)$$

3.1.1 Suposições

Consideremos as seguintes suposições:

1. $\boldsymbol{\theta}_*$ é um ponto interior de Θ , em que Θ é um espaço limitado com métrica d , ou seja, $\forall \epsilon > 0$ existe um número finito de bolas de raio ϵ que cobrem Θ .
2. $\hat{\boldsymbol{\theta}}$ converge em probabilidade para $\boldsymbol{\theta}_*$, ou seja, para $n \rightarrow \infty$ temos que $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}_*$. Além disso, $\boldsymbol{\theta}_*$ é o único valor que minimiza $E[\sum_{k=1}^n \ell_k(\boldsymbol{\theta})]$.
3. Para $k = 1, 2, \dots, n$ todas as densidades $p(y_k; \boldsymbol{\theta})$ possuem o mesmo suporte em que o mesmo não depende de $\boldsymbol{\theta}$. Para todo k , $\ell_k(\boldsymbol{\theta})$ é três vezes diferenciável em Θ , e também, $|\partial_j \ell_k(\boldsymbol{\theta})|$, $|\partial_j \partial_i \ell_k(\boldsymbol{\theta})|$, $|\partial_j \partial_i \partial_l \ell_k(\boldsymbol{\theta})|$ e $|\partial_j [p(y_k; \boldsymbol{\theta}) \times p(y_k; \boldsymbol{\theta})]|$ são dominados por $B(y_k)$ para $i, j, l = 1, 2, \dots, p$, em que $\partial_j = \partial/\partial\theta_j$ e $B(\cdot)$ está definido em (3.1).
4. Para cada $\epsilon > 0$, existe $M < \infty$ tal que

$$\sup_{n \geq 1} \frac{\sum_{k=1}^n E[B(y_k)^2 I_{[B(y_k) > M]}]}{n} < \epsilon,$$

em que I_A denota a função indicadora do conjunto A .

5. É válido também que

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

e

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

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

6. Para todo $i \leq n$ e $j, k, j', k' = 1, 2, \dots, p$, $\ell_i(\boldsymbol{\theta})$ é diferenciável quatro vezes em Θ e $|\partial_j \partial_k \partial_{k'} \partial_{j'} \ell_i(\boldsymbol{\theta})|$ é dominado por $B(y_i)$.
7. Quando $n \rightarrow \infty$ então

$$n^{-1} \partial_{\theta} [SC_n(\boldsymbol{\theta}_*)] \xrightarrow{P} J_1(\boldsymbol{\theta}_*),$$

em que $SC_n(\boldsymbol{\theta}) = \sum_{i=1}^n nC_{e_i}(\boldsymbol{\theta})$ e

$$J_1(\boldsymbol{\theta}) = 2 \left\{ \frac{\partial}{\partial \boldsymbol{\theta}} \text{tr}\{A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})\} \right\}. \quad (3.2)$$

8. Para cada $K > 0$, $n^{-1} \sum_{i=1}^n E[k_i(y_i; \boldsymbol{\theta}_*)^2 I_{k_i(y_i; \boldsymbol{\theta}_*) > \sqrt{n}K}] \rightarrow 0$. Além disso,

$$n^{-1} \sum_{i=1}^n E[k_i(y_i; \boldsymbol{\theta}_*)^2] \xrightarrow{p} \sigma_*^2,$$

em que σ_*^2 representa a variância da estatística $T_n(\hat{\boldsymbol{\theta}})$ que será definida posteriormente.

3.1.2 Teste Baseado no Esquema de Perturbação Ponderação de Casos

Um dos esquemas de perturbação que é bastante utilizado no procedimento de influência local é o de ponderação de casos. Nessa forma de perturbação, o logaritmo da função de verossimilhança perturbado é escrito como

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta}), \quad (3.3)$$

em que $0 \leq \omega_i \leq 1$ e $\ell_i(\boldsymbol{\theta}) = \log(p(y_i; \boldsymbol{\theta}))$. Para esse esquema de perturbação, tem-se que $\boldsymbol{\omega}_0 = (1, 1, \dots, 1)^\top$ e $\Delta = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top$ avaliada em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$.

A seguir apresentaremos um teorema que será fundamental para a introdução da estatística de teste.

Teorema 3.1.1. *Seja $\lambda_0 = 2\text{tr}\{A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)\}$. Considerando as Suposições 1 a 5, para o esquema de perturbação ponderação de casos temos os seguintes resultados:*

1. $\sum_{i=1}^n C_{e_i} \xrightarrow{p} \lambda_0$ em que C_{e_i} é definido em (1.5);
2. Se $g(\mathbf{y}) = p(\mathbf{y}; \boldsymbol{\theta}_0)$ e $\boldsymbol{\theta}_0 \in \Theta$, em que $g(\mathbf{y})$ é a verdadeira densidade de \mathbf{y} , então $\boldsymbol{\theta}_* = \boldsymbol{\theta}_0$, $A(\boldsymbol{\theta}_*) = B(\boldsymbol{\theta}_*)$ e $\lambda_0 = 2p$.

A prova do Teorema 3.1.1 está apresentada no Apêndice A₁. A estatística de teste que será introduzida é baseada na quantidade $\sum_{i=1}^n C_{e_i}$, a qual é calculada utilizando o esquema de perturbação em (3.3). Pelo Teorema 3.1.1 vimos que essa quantidade converge em probabilidade para λ_0 . Ainda, por esse teorema, se $g(\mathbf{y})$ é a verdadeira densidade de \mathbf{y} temos que $\lambda_0 = 2p$. Dessa forma, para avaliar se a magnitude da soma total da medida de influência é significativa, Zhu e Zhang (2004) propõem λ_0 como sendo um ponto de corte.

O interesse recai em testar

$$H_0 : g(\mathbf{y}) = p(\mathbf{y}; \boldsymbol{\theta}) \quad \text{para algum } \boldsymbol{\theta} \in \Theta, \quad (3.4)$$

contra a hipótese alternativa

$$H_1 : g(\mathbf{y}) \neq p(\mathbf{y}; \boldsymbol{\theta}) \quad \forall \boldsymbol{\theta} \in \Theta. \quad (3.5)$$

Para isso, definamos a estatística

$$T_n(\hat{\boldsymbol{\theta}}) = n^{-1/2} \sum_{i=1}^n (nC_{ei} - \lambda_0). \quad (3.6)$$

Zhu e Zhang (2004) obtiveram a distribuição assintótica da estatística T_n definida em (3.6). Antes de introduzirmos a distribuição, consideremos a quantidade

$$k(y_i; \boldsymbol{\theta}_*) = h(y_i; \boldsymbol{\theta}_*) + J_1(\boldsymbol{\theta}_*)A(\boldsymbol{\theta}_*)^{-1}\dot{\ell}(\boldsymbol{\theta}_*), \quad (3.7)$$

em que $\dot{\ell}(\boldsymbol{\theta}_*) = \partial \ell(\boldsymbol{\theta}_*) / \partial \boldsymbol{\theta}_*$ e

$$h(y_i; \boldsymbol{\theta}_*) = 2\text{tr}\{A(\boldsymbol{\theta}_*)^{-1}[\Delta_i \Delta_i^\top - B(\boldsymbol{\theta}_*)]\} - \text{tr}\{A(\boldsymbol{\theta}_*)^{-1}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*) - A(\boldsymbol{\theta}_*)]A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)\} \quad (3.8)$$

para $i = 1, 2, \dots, n$.

A seguir apresentamos o teorema que caracteriza o comportamento estocástico de $\sum_{i=1}^n nC_{ei}$.

Teorema 3.1.2. *Sob as Suposições 1 a 8, é válido que*

$$n^{-1/2} \sum_{i=1}^n nC_{ei} = n^{-1/2} \lambda_0 + n^{-1/2} \sum_{i=1}^n k(y_i; \boldsymbol{\theta}_*) + o_p(1),$$

em que $o_p(\cdot)$ denota a ordem de grandeza. Além disso, se $\sum_{i=1}^n E[k^2(y_i; \boldsymbol{\theta}_*)]/n \xrightarrow{p} \sigma_*^2$ temos que $T_n(\hat{\boldsymbol{\theta}}) \xrightarrow{d} N(0, \sigma_*^2)$.

A prova do Teorema 3.1.2 está apresentada em Zhu e Zhang (2004). Logo, primeiramente estimamos $\boldsymbol{\theta}_*$, $A(\boldsymbol{\theta}_*)$, $B(\boldsymbol{\theta}_*)$ e $J_1(\boldsymbol{\theta}_*)$, em seguida, σ_* . Portanto, temos que $k(y_i; \hat{\boldsymbol{\theta}})$ será um estimador de $k(y_i; \boldsymbol{\theta}_*)$. Denotando $S_{k,c}^2$ como sendo a variância amostral de $\{k(y_i; \hat{\boldsymbol{\theta}}), i = 1, 2, \dots, n\}$, temos que $S_{k,c}^2$ é um estimador consistente para σ_*^2 .

Do Teorema 3.1.2 temos

$$\begin{aligned}
 T_n(\hat{\boldsymbol{\theta}}) &= n^{-1/2} \sum_{i=1}^n (nC_{e_i} - \lambda_0) \xrightarrow{d} N(0, \sigma_*^2) \iff \\
 &\frac{n^{-1/2} \sum_{i=1}^n (nC_{e_i} - \lambda_0 + 2p - 2p)}{\sigma_*} \xrightarrow{d} N(0, 1) \iff \\
 &\frac{n^{-1/2} \sum_{i=1}^n (nC_{e_i} - 2p) + n^{-1/2} \sum_{i=1}^n (2p - \lambda_0)}{\sigma_*} \xrightarrow{d} N(0, 1) \iff \\
 &\frac{n^{-1/2} \sum_{i=1}^n (nC_{e_i} - 2p) + n^{1/2} \sum_{i=1}^n (2p - \lambda_0)}{\sigma_*} \xrightarrow{d} N(0, 1). \tag{3.9}
 \end{aligned}$$

Baseado em (3.9), considere a seguinte estatística de teste:

$$U_n = \sum_{i=1}^n \frac{n^{-1/2}(nC_{e_i} - 2p)}{S_{k,c}}. \tag{3.10}$$

A estatística definida em (3.10) se torna uma ferramenta útil de diagnóstico para verificar o ajuste global do modelo suposto, isto é, testamos a hipótese $H_0 : \lambda_0 = 2p$ versus $H_1 : \lambda_0 \neq 2p$.

Ainda, note que se $c_0 = \lambda_0 - 2p \neq 0$, então

$$U_n = N(0, 1) + \frac{n^{1/2}c_0}{\sigma_*},$$

a qual é dominada por $(n^{1/2}c_0)/\sigma_*$. Assim, sob uma sequência de alternativas locais $\lambda_0 - 2p = n^{-1/2}c_1$. Logo,

$$\begin{aligned}
 U_n &= N(0, 1) + \frac{n^{1/2}(\lambda_0 - 2p)}{\sigma_*} \\
 &= N(0, 1) + \frac{c_1}{\sigma_*}.
 \end{aligned}$$

Portanto, $U_n \xrightarrow{d} N(0, 1) + c_1/\sigma_* = N(c_1/\sigma_*, 1)$. Logo, para realizar o teste, utilizamos a distribuição assintótica da estatística U_n , que é normal padrão.

3.2 Cálculo da Estatística para a Regressão Beta

Nessa seção aplicaremos o teste proposto por Zhu e Zhang (2004) ao modelo de regressão beta com dispersão fixa e variável.

3.2.1 Dispersão Fixa

Inicialmente obteremos as matrizes $A(\boldsymbol{\theta})$ e $B(\boldsymbol{\theta})$ definidas em (3.1). A expressão de $A(\boldsymbol{\theta})$ já foi obtida na Seção 2.2, pois $A(\boldsymbol{\theta}) = K(\boldsymbol{\beta}, \phi)$. Agora, precisamos da expressão para a

matriz $B(\boldsymbol{\theta})$. Dessa forma, segue que

$$\begin{aligned} \mathbb{E}(U_{\beta_r} U_{\beta_j}) &= \mathbb{E} \left[\phi \sum_{i=1}^n x_{ir} t_i (y_i^* - \mu_i^*) \phi \sum_{l=1}^n x_{lj} t_l (y_l^* - \mu_l^*) \right] \\ &= \phi^2 \sum_{i=1}^n \sum_{l=1}^n x_{ir} x_{lj} t_i t_l \mathbb{E}[(y_i^* - \mu_i^*)(y_l^* - \mu_l^*)]. \end{aligned}$$

Notemos que para $i \neq l$, como y_i 's, $i = 1, 2, \dots, n$, são independentes, então

$$\mathbb{E}[(y_i^* - \mu_i^*)(y_l^* - \mu_l^*)] \stackrel{\text{ind}}{=} \mathbb{E}(y_i^* - \mu_i^*) \mathbb{E}(y_l^* - \mu_l^*) = 0,$$

em que $\stackrel{\text{ind}}{=}$ significa igualdade sob independência. Assim,

$$\begin{aligned} \mathbb{E}(U_{\beta_r} U_{\beta_j}) &= \phi^2 \sum_{i=1}^n x_{ir} x_{il} t_i^2 \mathbb{E}[(y_i^* - \mu_i^*)^2] \\ &= \phi^2 \sum_{i=1}^n x_{ir} x_{il} t_i^2 \text{Var}(y_i^*) \\ &= \phi^2 \sum_{i=1}^n x_{ir} x_{il} t_i^2 v_i \\ &= \phi^2 \sum_{i=1}^n x_{ir} x_{il} w_i. \end{aligned}$$

Então, matricialmente, temos

$$\mathbb{E}(U_{\beta} U_{\beta}^{\top}) = \phi X^{\top} W X.$$

Também, precisamos de

$$\begin{aligned} \mathbb{E}(U_{\beta_r} U_{\phi}) &= \mathbb{E} \left[\phi \sum_{i=1}^n x_{ir} t_i (y_i^* - \mu_i^*) \sum_{j=1}^n u_j \right] \\ &= \phi \sum_{i=1}^n x_{ir} t_i \mathbb{E}[(y_i^* - \mu_i^*) u_i] \\ &= \phi \sum_{i=1}^n x_{ir} t_i [\mathbb{E}(y_i^* u_i) - \mu_i^* \mathbb{E}(u_i)]. \end{aligned} \tag{3.11}$$

Mas

$$\begin{aligned} \mathbb{E}(y_i^* u_i) &= \mathbb{E} \{ y_i^* [\mu_i (y_i^* - \mu_i^*) + \log(1 - y_i) - \psi((1 - \mu_i)\phi) + \psi(\phi)] \} \\ &= \mu_i [\mathbb{E}(y_i^{*2}) - \mu_i^* \mathbb{E}(y_i^*)] + \mathbb{E}[y_i^* \log(1 - y_i)] \\ &+ [\psi(\phi) - \psi((1 - \mu_i)\phi)] \mathbb{E}(y_i^*), \end{aligned} \tag{3.12}$$

em que como já vimos $E(y_i^*) = \mu_i^*$ e $E(y_i^{*2}) = \text{Var}(y_i^*) + [E(y_i^*)]^2 = v_i + \mu_i^{*2}$. Porém, precisamos ainda de $E[y_i^* \log(1 - y_i)]$. Para isso, como $y_i^* = \log(y_i/(1 - y_i))$ temos

$$\begin{aligned} \log(y_i) &= y_i^* + \log(1 - y_i) \\ \{\log(y_i)\}^2 &= y_i^{*2} + [\log(1 - y_i)]^2 + 2y_i^* \log(1 - y_i) \\ E[y_i^* \log(1 - y_i)] &= \frac{1}{2} \{E\{[\log(y_i)]^2\} - \text{Var}(y_i^*) - [E(y_i^*)]^2 - E\{[\log(1 - y_i)]^2\}\}. \end{aligned} \quad (3.13)$$

Para obtermos as esperanças matemáticas necessárias em (3.13), utilizaremos inicialmente a parametrização (2.1). E, em seguida, voltaremos à parametrização em função de μ e ϕ . Considerando $Z_1 = \log(1 - y_i)$, temos que a função geradora de momentos (fgm) de Z_1 é dada por

$$\begin{aligned} M_{Z_1}(k) &= E(e^{kZ_1}) \\ &= E[e^{k \log(1 - y_i)}] \\ &= E[(1 - y_i)^k] \\ &= \int_0^1 (1 - y_i)^k \frac{\Gamma(p + q)}{\Gamma(p)\Gamma(q)} y_i^{p-1} (1 - y_i)^{q-1} dy_i \\ &= \frac{\Gamma(q + k)\Gamma(p + q)}{\Gamma(q)\Gamma(p + q + k)} \int_0^1 \frac{\Gamma(p + q + k)}{\Gamma(p)\Gamma(q + k)} y_i^{p-1} (1 - y_i)^{q+k-1} dy_i \\ &= \frac{\Gamma(q + k)\Gamma(p + q)}{\Gamma(q)\Gamma(p + q + k)}. \end{aligned}$$

Tomando $p = \mu_i \phi$ e $q = (1 - \mu_i) \phi$, temos

$$M_{Z_1}(k) = \frac{\Gamma((1 - \mu_i)\phi + k)\Gamma(\phi)}{\Gamma((1 - \mu_i)\phi)\Gamma(\phi + k)}.$$

Dessa forma, para obtermos $E\{[\log(1 - y_i)]^2\}$ precisamos da segunda derivada de $M_{Z_1}(k)$ avaliada em $k = 0$. Temos

$$\begin{aligned} M'_{Z_1}(k) &= \frac{d}{dk} [M_{Z_1}(k)] \\ &= \frac{\Gamma(\phi)}{\Gamma((1 - \mu_i)\phi)\Gamma(\phi + k)} [\Gamma'(k + (1 - \mu_i)\phi) - \Gamma(k + (1 - \mu_i)\phi)\psi(\phi + k)]. \end{aligned}$$

Também,

$$\begin{aligned} M''_{Z_1}(k) &= \frac{d^2}{dk^2} [M_{Z_1}(k)] \\ &= \frac{\Gamma(\phi)}{\Gamma((1 - \mu_i)\phi)\Gamma(\phi + k)} \{ \Gamma''(k + (1 - \mu_i)\phi) - 2\Gamma'(k + (1 - \mu_i)\phi)\psi(\phi + k) \\ &\quad - \Gamma(k + (1 - \mu_i)\phi)\psi'(\phi + k) + \Gamma(k + (1 - \mu_i)\phi)[\psi(\phi + k)]^2 \}. \end{aligned}$$

Usando o fato de que $\Gamma''(x)/\Gamma(x) = \psi'(x) + [\psi(x)]^2$ e após alguns cálculos, temos

$$\begin{aligned}
 E\{[\log(1 - y_i)]^2\} &= M''_{Z_1}(0) \\
 &= \frac{\Gamma(\phi)}{\Gamma((1 - \mu_i)\phi)\Gamma(\phi)} \{ \Gamma''((1 - \mu_i)\phi) - 2\Gamma'((1 - \mu_i)\phi)\psi(\phi) \\
 &\quad - \Gamma((1 - \mu_i)\phi)\psi'(\phi) + \Gamma((1 - \mu_i)\phi)[\psi(\phi)]^2 \} \\
 &= \frac{\Gamma''((1 - \mu_i)\phi)}{\Gamma((1 - \mu_i)\phi)} - 2\psi((1 - \mu_i)\phi)\psi(\phi) - \psi'(\phi) + [\psi(\phi)]^2 \\
 &= \psi'((1 - \mu_i)\phi) - \psi'(\phi) + [\psi((1 - \mu_i)\phi) - \psi(\phi)]^2. \tag{3.14}
 \end{aligned}$$

Precisamos também de $E\{[\log(y_i)]^2\}$. Seja $Z_2 = \log(y_i)$. A fgm de Z_2 é dada por

$$\begin{aligned}
 M_{Z_2}(k) &= E(e^{kZ_2}) \\
 &= E[e^{k \log(y_i)}] \\
 &= E[y_i^k] \\
 &= \int_0^1 y_i^k \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y_i^{p-1} (1-y_i)^{q-1} dy_i \\
 &= \frac{\Gamma(p+k)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+k)} \int_0^1 \frac{\Gamma(p+q+k)}{\Gamma(p+k)\Gamma(q)} y_i^{p+k-1} (1-y_i)^{q-1} dy_i \\
 &= \frac{\Gamma(p+k)\Gamma(p+q)}{\Gamma(p)\Gamma(p+q+k)}.
 \end{aligned}$$

Tomando $p = \mu_i\phi$ e $q = (1 - \mu_i)\phi$ temos

$$M_{Z_2}(k) = \frac{\Gamma(\phi)\Gamma((\mu_i\phi + k))}{\Gamma(\mu_i\phi)\Gamma(\phi + k)}.$$

Dessa forma, para obtermos $E\{[\log(y_i)]^2\}$ precisamos da segunda derivada de $M_{Z_2}(k)$ avaliada em $k = 0$. Logo,

$$\begin{aligned}
 M'_{Z_2}(k) &= \frac{d}{dk} [M_{Z_2}(k)] \\
 &= \frac{\Gamma(\phi)}{\Gamma(\mu_i\phi)\Gamma(\phi + k)} [\Gamma'(k + \mu_i\phi) - \Gamma(k + \mu_i\phi)\psi(\phi + k)].
 \end{aligned}$$

Ainda,

$$\begin{aligned}
 M''_{Z_2}(k) &= \frac{d^2}{dk^2} [M_{Z_2}(k)] \\
 &= \frac{\Gamma(\phi)}{\Gamma((\mu_i\phi)\Gamma(\phi + k))} \{ \Gamma''(k + \mu_i\phi) - 2\Gamma'(k + \mu_i\phi)\psi(\phi + k) \\
 &\quad - \Gamma(k + \mu_i\phi)\psi'(\phi + k) + \Gamma(k + \mu_i\phi)[\psi(\phi + k)]^2 \}.
 \end{aligned}$$

Assim, temos

$$\begin{aligned}
E\{\log(y_i)^2\} &= M''_{Z_2}(0) \\
&= \frac{\Gamma(\phi)}{\Gamma((\mu_i\phi)\Gamma(\phi))} \{\Gamma''(\mu_i\phi) - 2\Gamma'(\mu_i\phi)\psi(\phi) \\
&\quad - \Gamma(\mu_i\phi)\psi'(\phi) + \Gamma(\mu_i\phi)[\psi(\phi)]^2\} \\
&= \psi'(\mu_i\phi) - \psi'(\phi) + [\psi((\mu_i\phi) - \psi(\phi))]^2.
\end{aligned}$$

Utilizando todas as esperanças matemáticas obtidas, após alguns cálculos temos que (3.13) fica dada por

$$\begin{aligned}
E[y_i^* \log(1 - y_i)] &= \frac{1}{2} \{E\{\log(y_i)^2\} - \text{Var}(y_i^*) - [E(y_i^*)]^2 - E\{\log(1 - y_i)^2\}\} \\
&= \frac{1}{2} \{\psi'(\mu_i\phi) - \psi'(\phi) + [\psi((\mu_i\phi) - \psi(\phi))]^2 - v_i - \mu_i^{*2} \\
&\quad - \psi'((1 - \mu_i)\phi) + \psi'(\phi) - [\psi((1 - \mu_i)\phi) - \psi(\phi)]^2\}. \tag{3.15}
\end{aligned}$$

De (2.11) e (2.13), a expressão (3.15) se reduz a

$$E[y_i^* \log(1 - y_i)] = \mu_i^* [\psi((1 - \mu_i)\phi) - \psi(\phi)] - \psi'((1 - \mu_i)\phi). \tag{3.16}$$

Assim, voltando à expressão (3.17), temos

$$\begin{aligned}
E(y_i^* u_i) &= \mu_i [v_i + \mu_i^{*2} - \mu_i^{*2}] + \mu_i^* [\psi((1 - \mu_i)\phi) - \psi(\phi)] - \psi'((1 - \mu_i)\phi) \\
&\quad + [\psi(\phi) - \psi((1 - \mu_i)\phi)] \mu_i^* \\
&= \mu_i [\psi'(\mu_i\phi) + \psi'((1 - \mu_i)\phi)] - \psi'((1 - \mu_i)\phi) \\
&= \psi'(\mu_i\phi) \mu_i - \psi'((1 - \mu_i)\phi) (1 - \mu_i) \\
&= \frac{c_i}{\phi}. \tag{3.17}
\end{aligned}$$

Ainda, em (3.11) precisamos de $E(u_i)$. Então,

$$\begin{aligned}
E(u_i) &= \psi(\phi) + \mu_i E(y_i^* - \mu_i^*) - \psi((1 - \mu_i)\phi) + E(\log(1 - y_i)) \\
&= \psi(\phi) - \psi((1 - \mu_i)\phi) + E(\log(1 - y_i)).
\end{aligned}$$

Para obtermos $E(\log(1 - y_i))$ basta avaliar $M'_{Z_1}(k)$ em $k = 0$. Dessa forma,

$$\begin{aligned}
E(\log(1 - y_i)) &= M'_{Z_1}(0) \\
&= \frac{\Gamma(\phi)}{\Gamma((1 - \mu_i)\phi)\Gamma(\phi)} [\Gamma'((1 - \mu_i)\phi) - \Gamma((1 - \mu_i)\phi)\psi(\phi)] \\
&= \psi((1 - \mu_i)\phi) - \psi(\phi). \tag{3.18}
\end{aligned}$$

Então,

$$\begin{aligned} E(u_i) &= \psi(\phi) - \psi((1 - \mu_i)\phi) + \psi((1 - \mu_i)\phi) - \psi(\phi) \\ &= 0. \end{aligned} \quad (3.19)$$

Assim, das Expressões (3.17) e (3.19) obtemos que (3.11) fica dada por

$$E(U_\beta U_\phi) = X^\top T \mathbf{c}.$$

Por fim, temos

$$\begin{aligned} E(U_\phi U_\phi^\top) &= E \left[\left(\sum_{i=1}^n u_i \right)^2 \right] \\ &= \text{Var} \left(\sum_{i=1}^n u_i \right) + \left[E \left(\sum_{i=1}^n u_i \right) \right]^2 \\ &\stackrel{\text{ind}}{=} \sum_{i=1}^n \text{Var}(u_i) + \left(\sum_{i=1}^n E(u_i) \right)^2 \\ &= \sum_{i=1}^n E(u_i^2), \end{aligned} \quad (3.20)$$

em que a penúltima igualdade segue de (3.19). Notemos que precisamos de $E(u_i^2)$. Assim,

$$\begin{aligned} E(u_i^2) &= E\{\mu_i(y_i^* \mu_i^*) + \log(1 - y_i)\}^2 + [\psi(\phi) - \psi((1 - \mu_i)\phi)]^2 \\ &\quad + 2[\psi(\phi) - \psi((1 - \mu_i)\phi)]E[\mu_i(y_i^* - \mu_i^*) + \log(1 - y_i)] \\ &= \mu_i^2 \text{Var}(y_i^*) + E\{\log(1 - y_i)\}^2 + 2\mu_i\{E[y_i^* \log(1 - y_i)] \\ &\quad - \mu_i^* E[\log(1 - y_i)]\} + [\psi(\phi) - \psi((1 - \mu_i)\phi)]^2 \\ &\quad + 2[\psi(\phi) - \psi((1 - \mu_i)\phi)]E[\log(1 - y_i)]. \end{aligned} \quad (3.21)$$

De (3.14), (3.16) e (3.18) temos que (3.22) se reduz a

$$\begin{aligned} E(u_i^2) &= \mu_i^2 v_i + \psi'((1 - \mu_i)\phi) - \psi'(\phi) + [\psi((1 - \mu_i)\phi) - \psi(\phi)]^2 \\ &\quad + 2\mu_i\{\mu_i^*[\psi((1 - \mu_i)\phi) - \psi(\phi)] - \psi'((1 - \mu_i)\phi) - \mu_i^*[\psi((1 - \mu_i)\phi) - \psi(\phi)]\} \\ &\quad + [\psi(\phi) - \psi((1 - \mu_i)\phi)]^2 + 2[\psi(\phi) - \psi((1 - \mu_i)\phi)][\psi((1 - \mu_i)\phi) - \psi(\phi)] \\ &= \mu_i^2 v_i + \psi'((1 - \mu_i)\phi) - \psi'(\phi) - 2\mu_i\psi'((1 - \mu_i)\phi) \\ &= \mu_i^2[\psi'(\mu_i\phi) + \psi'((1 - \mu_i)\phi)] + \psi'((1 - \mu_i)\phi) - \psi'(\phi) - 2\mu_i\psi'((1 - \mu_i)\phi) \\ &= \mu_i^2\psi'(\mu_i\phi) + (1 - \mu_i)^2\psi'((1 - \mu_i)\phi) - \psi'(\phi) \\ &= d_i. \end{aligned} \quad (3.22)$$

Logo, substituindo (3.22) em (3.20), temos que

$$\begin{aligned} E(U_\phi U_\phi^\top) &= \sum_{i=1}^n d_i \\ &= \text{tr}(D). \end{aligned}$$

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

$$B(\boldsymbol{\theta}) = \begin{bmatrix} \phi X^\top W X & X^\top T \mathbf{c} \\ \mathbf{c}^\top T^\top X & \text{tr}(D) \end{bmatrix}.$$

Assim, para o modelo de regressão beta com dispersão fixa, temos $A(\boldsymbol{\theta}) = B(\boldsymbol{\theta})$. Portanto, utilizando esse fato em (3.2) temos que

$$\begin{aligned} J_1(\boldsymbol{\theta}) &= 2 \frac{\partial}{\partial \boldsymbol{\theta}} (\text{tr}\{A(\boldsymbol{\theta})^{-1} B(\boldsymbol{\theta})\}) \\ &= 2 \frac{\partial}{\partial \boldsymbol{\theta}} (\text{tr}\{\mathbb{I}_{k+1}\}) \\ &= \mathbf{0}, \end{aligned}$$

em que $\mathbf{0}$ é um vetor de zeros de dimensão $(k+1) \times 1$. Assim, a expressão de $k(y_i; \boldsymbol{\theta}_*)$ se torna

$$\begin{aligned} k(y_i; \boldsymbol{\theta}_*) &= 2 \text{tr}\{A(\boldsymbol{\theta}_*)^{-1} [\Delta_i \Delta_i^\top - B(\boldsymbol{\theta}_*)]\} \\ &\quad - \text{tr}\{A(\boldsymbol{\theta}_*)^{-1} [-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*) - A(\boldsymbol{\theta}_*)] A(\boldsymbol{\theta}_*)^{-1} B(\boldsymbol{\theta}_*)\} \\ &= 2 \text{tr}[A(\boldsymbol{\theta}_*)^{-1} \Delta_i \Delta_i^\top - \mathbb{I}_{k+1}] + \text{tr}[A(\boldsymbol{\theta}_*)^{-1} \ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*) + \mathbb{I}_{k+1}]. \end{aligned} \quad (3.23)$$

Também precisamos da expressão para a matriz Δ . Vimos na Seção 1.2.1 que $\Delta = \partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top$ avaliada em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$. Assim, considerando o esquema de ponderação de casos, temos que o logaritmo da função de verossimilhança perturbado é dado em (3.3). Logo, a i -ésima coluna da matriz Δ é dada por

$$\begin{aligned} \Delta_i &= \left(\frac{\partial^2 \ell_i(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_1 \partial \omega_i}, \dots, \frac{\partial^2 \ell_i(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_k \partial \omega_i}, \frac{\partial^2 \ell_i(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \phi \partial \omega_i} \right)^\top \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} \\ &= \left(\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_1}, \dots, \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_k}, \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \phi} \right)^\top \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \end{aligned}$$

A partir de (2.10), temos que para $r = 1, 2, \dots, k$

$$\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_r} = \phi x_{ir} t_i(y_i^* - \mu_i^*),$$

e também a partir de (2.14) temos

$$\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \phi} = u_i.$$

Assim, a matriz Δ avaliada em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$ é dada por

$$\Delta = \begin{pmatrix} \hat{\phi} X^\top \hat{T} \hat{E} \\ \hat{\mathbf{u}} \end{pmatrix}, \quad (3.24)$$

em que $\hat{E} = \text{diag}\{(y_1^* - \hat{\mu}_1^*), (y_2^* - \hat{\mu}_2^*), \dots, (y_n^* - \hat{\mu}_n^*)\}$, $\hat{T} = \text{tr}\{\hat{t}_1, \hat{t}_2, \dots, \hat{t}_n\}$ e $\hat{\mathbf{u}} = (\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n)^\top$.

Dessa forma, para o modelo de regressão beta com dispersão fixa, calculamos a estatística U_n utilizando a expressão (3.10) em que $p = k + 1$, C_{e_i} é obtida através de (1.5) utilizando a matriz Δ apresentada em (3.24), e $S_{k,c}$ é estimado através do cálculo do desvio-padrão de $k(y_i; \boldsymbol{\theta}_*)$ apresentado em (3.23) o qual é avaliado em $\hat{\boldsymbol{\theta}}$.

3.2.2 Dispersão Variável

Inicialmente, obteremos as matrizes $A(\boldsymbol{\theta})$ e $B(\boldsymbol{\theta})$ definidas em (3.1). A expressão de $A(\boldsymbol{\theta})$ já foi obtida na Seção 2.3, pois $A(\boldsymbol{\theta}) = K(\boldsymbol{\beta}, \boldsymbol{\gamma})$. Agora, precisamos da expressão para a matriz $B(\boldsymbol{\theta})$. Dessa forma, segue que

$$\begin{aligned} E(U_{\beta_r} U_{\beta_j}) &= E \left[\sum_{i=1}^n \phi_i x_{ir} t_i (y_i^* - \mu_i^{**}) \sum_{l=1}^n \phi_l x_{lj} t_l (y_l^* - \mu_l^{**}) \right] \\ &= \sum_{i=1}^n \sum_{l=1}^n \phi_i \phi_l x_{ir} x_{lj} t_i t_l E[(y_i^* - \mu_i^*)(y_l^* - \mu_l^*)]. \end{aligned}$$

Novamente, temos que para $i \neq l$ $E[(y_i^* - \mu_i^{**})(y_l^* - \mu_l^{**})] \stackrel{\text{ind}}{=} E(y_i^* - \mu_i^{**})E(y_l^* - \mu_l^{**}) = 0$.

Assim,

$$\begin{aligned} E(U_{\beta_r} U_{\beta_j}) &= \sum_{i=1}^n \phi_i^2 x_{ir} x_{il} t_i^2 E[(y_i^* - \mu_i^{**})^2] \\ &= \sum_{i=1}^n \phi_i^2 x_{ir} x_{il} t_i^2 \text{Var}(y_i^*) \\ &= \sum_{i=1}^n \phi_i^2 x_{ir} x_{il} t_i^2 v_i^* \\ &= \sum_{i=1}^n \phi_i x_{ir} x_{il} w_i^*. \end{aligned}$$

Então, matricialmente, temos

$$E(U_\beta U_\beta^\top) = X^\top \Phi W^* X.$$

Também, precisamos de

$$\begin{aligned} E(U_{\beta_r} U_{\gamma_p}) &= E \left[\sum_{i=1}^n \phi_i x_{ir} t_i (y_i^* - \mu_i^{**}) \sum_{j=1}^n z_{jp} t_j^* u_j^* \right] \\ &= \sum_{i=1}^n \phi_i x_{ir} z_{ip} t_i^* [E(y_i^* u_i^*) - \mu_i^{**} E(u_i^*)]. \end{aligned}$$

Analogamente, ao que foi feito na subseção anterior, obtemos que $E(y_i^* u_i^*) = c_i^*/\phi_i$ e $E(u_i^*) = 0$. Portanto,

$$\begin{aligned} E(U_{\beta_r} U_{\gamma_p}) &= \sum_{i=1}^n x_{ir} z_{ip} \phi_i t_i^* \frac{c_i^*}{\phi_i} \\ &= \sum_{i=1}^n x_{ir} z_{ip} t_i^* c_i^*. \end{aligned}$$

Desta forma, temos

$$E(U_\beta U_\gamma) = X^\top T T^* C^* Z.$$

Por fim,

$$\begin{aligned} E(U_{\gamma_p} U_{\gamma_k}) &= E \left(\sum_{i=1}^n z_{ip} t_i^* u_i^* \sum_{j=1}^n z_{jk} t_j^* u_j^* \right) \\ &= \sum_{i=1}^n z_{ip} z_{ik} t_i^{*2} E(u_i^{*2}). \end{aligned}$$

Novamente, analogamente ao que foi feito na subseção anterior, obtemos que $E(u_i^*) = d_i^*$. Matricialmente, temos

$$E(U_\gamma U_\gamma^\top) = Z^\top T^{*2} D^* Z.$$

Portanto, a matriz $B(\boldsymbol{\theta})$ é dada por

$$B(\boldsymbol{\theta}) = \begin{bmatrix} X^\top \Phi W^* X & X^\top T T^* C^* Z \\ Z^\top C^* T^* T X^\top & Z^\top T^{*2} D^* Z \end{bmatrix}.$$

Assim, notamos que, para o modelo de regressão beta com dispersão variável também

temos $A(\boldsymbol{\theta}) = B(\boldsymbol{\theta})$. Portanto, utilizando esse fato em (3.2) temos que

$$\begin{aligned} J_1(\boldsymbol{\theta}) &= 2 \frac{\partial}{\partial \boldsymbol{\theta}} (\text{tr}\{A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})\}) \\ &= 2 \frac{\partial}{\partial \boldsymbol{\theta}} (\text{tr}\{\mathbb{I}_{k+q}\}) \\ &= \mathbf{0}, \end{aligned}$$

em que $\mathbf{0}$ é um vetor de zeros de dimensão $(k+q) \times 1$. Assim, a expressão de $k(y_i; \boldsymbol{\theta}_*)$ fica dada por

$$\begin{aligned} k(y_i; \boldsymbol{\theta}_*) &= 2 \text{tr}\{A(\boldsymbol{\theta}_*)^{-1}[\Delta_i \Delta_i^\top - B(\boldsymbol{\theta}_*)]\} \\ &\quad - \text{tr}\{A(\boldsymbol{\theta}_*)^{-1}[-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*) - A(\boldsymbol{\theta}_*)]A(\boldsymbol{\theta}_*)^{-1}B(\boldsymbol{\theta}_*)\} \\ &= 2 \text{tr}[A(\boldsymbol{\theta}_*)^{-1}\Delta_i \Delta_i^\top - \mathbb{I}_{k+q}] + \text{tr}[A(\boldsymbol{\theta}_*)^{-1}\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*) + \mathbb{I}_{k+q}]. \end{aligned} \quad (3.25)$$

Também precisamos da expressão para a matriz Δ . Vimos na Seção 1.2.1 que $\Delta = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top$ avaliada em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$. Assim, considerando o esquema de ponderação de casos, temos que o logaritmo da função de verossimilhança perturbado é dado em (3.3). Logo, a i -ésima coluna da matriz Δ é dada por

$$\begin{aligned} \Delta_i &= \left(\frac{\partial^2 \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_1 \partial \omega_i}, \dots, \frac{\partial^2 \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_k \partial \omega_i}, \frac{\partial^2 \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \gamma_1 \partial \omega_i}, \dots, \frac{\partial^2 \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \gamma_q \partial \omega_i} \right)^\top \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0} \\ &= \left(\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_1}, \dots, \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_k}, \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \gamma_1}, \dots, \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \gamma_q} \right)^\top \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \end{aligned}$$

A partir de (2.17), temos que para $r = 1, 2, \dots, k$

$$\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta_r} = \phi_i x_{ir} t_i(y_i^* - \mu_i^{**}),$$

e também a partir de (2.19) temos que, para $p = 1, 2, \dots, q$,

$$\frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \gamma_p} = z_{ip} t_i^* u_i^*.$$

Assim, a matriz Δ avaliada em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega}_0$ é dada por

$$\Delta = \begin{pmatrix} X^\top \hat{\Phi} \hat{T} \hat{E} \\ Z^\top \hat{T}^* \hat{u}^* \end{pmatrix}, \quad (3.26)$$

em que $\hat{\Phi} = \text{tr}\{\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_n\}$, $\hat{E} = \text{diag}\{(y_1^* - \hat{\mu}_1^{**}), (y_2^* - \hat{\mu}_2^{**}), \dots, (y_n^* - \hat{\mu}_n^{**})\}$, $\hat{T}^* = \text{tr}\{\hat{t}_1^*, \hat{t}_2^*, \dots, \hat{t}_n^*\}$ e $\hat{u}^* = (\hat{u}_1^*, \hat{u}_2^*, \dots, \hat{u}_n^*)^\top$.

Dessa forma, para o modelo de regressão beta com dispersão variável, calculamos a es-

tatística U_n utilizando a expressão (3.10) em que $p = k + q$, C_{e_i} é obtida através de (1.5) utilizando a matriz Δ apresentada em (3.26) e $S_{k,c}$ é estimado através do cálculo do desvio-padrão de $k(y_i; \theta_*)$ apresentado em (3.25) o qual é avaliado em $\hat{\theta}$.

Capítulo 4

Estudos de Simulação: Parte 1

Nesse capítulo, apresentaremos os resultados de alguns estudos de simulação de Monte Carlo com o intuito de avaliar o desempenho do teste desenvolvido nesse trabalho, quando consideramos os modelos de regressão beta com dispersão fixa e variável. Para tanto, foram calculadas as taxas de rejeição sob a hipótese nula (tamanhos do teste). Em seguida, apresentamos resultados de estudos de simulações para o tamanho do teste referentes a um melhoramento feito no mesmo teste, utilizando a metodologia *bootstrap*. Já com a melhoria feita no teste, apresentamos também os poderes do teste quando são introduzidas algumas má especificações aos modelos. Em todas as simulações foram consideradas 5000 réplicas de Monte Carlo, tamanhos amostrais iguais a $n = 40, 80$ e 120 e níveis nominais $\alpha = 1\%, 5\%$ e 10% . Para obter tais amostras, inicialmente e de forma aleatória, foram geradas amostras de 40 observações dos regressores, as quais foram replicadas duas vezes a fim de se conseguir compor os demais tamanhos amostrais. A replicação é feita com o objetivo de manter constante o grau de heterogeneidade da dispersão, medido por $\lambda = \max(\phi_i)/\min(\phi_i)$, $i = 1, 2, \dots, n$. A especificação da média dos dois modelos é

$$g(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right) = \sum_{j=1}^5 \beta_j x_{ij},$$

em que $g(\cdot)$ é a função de ligação logito, β_j e x_{ij} , com $j = 1, \dots, 5$, são os parâmetros a serem estimados e as variáveis explicativas da média, respectivamente. A especificação do parâmetro de precisão (para o modelo com dispersão variável) é dada por

$$h(\phi_i) = \log(\phi_i) = \sum_{j=1}^5 \gamma_j z_{ij},$$

em que $h(\cdot)$ é a função de ligação log, γ_j e z_{ij} , com $j = 1, \dots, 5$, são os parâmetros a serem estimados e as variáveis explicativas da dispersão, respectivamente.

4.1 Modelo de Regressão Beta com Dispersão Fixa

Nessa seção, apresentaremos os resultados de estudos de simulação referentes ao desempenho do teste aplicado ao modelo de regressão beta com dispersão fixa. Foram consideradas y_1, y_2, \dots, y_n variáveis aleatórias independentes, em que $y_i \sim \text{Beta}(\mu_i, \phi)$, com $i = 1, 2, \dots, n$. Para a obtenção dos tamanhos do teste, a variável resposta foi gerada e modelada sem erro de especificação do modelo. Após a estimação dos parâmetros do modelo, aplicamos o teste baseado em influência local apresentado na Subseção 3.2.1. As covariáveis utilizadas para essa situação foram geradas a partir da distribuição lognormal padrão, e da distribuição Bernoulli de parâmetro 0,2. Ainda consideramos três cenários para os valores de μ : próximo do zero, próximo do um e central. Em relação aos valores utilizados para os parâmetros que compõem a componente sistemática, no cenário em que μ é próximo do zero, consideramos $\beta_1 = -2,4, \beta_2 = 1,9, \beta_3 = 0,7, \beta_4 = 3,4$ e $\beta_5 = -3,0$. Para o cenário em que μ é próximo do um, os parâmetros foram $\beta_1 = 2,0, \beta_2 = 2,5, \beta_3 = 3,2, \beta_4 = 2,9$ e $\beta_5 = 2,9$. Por fim, quando μ é central, utilizamos $\beta_1 = -1,0, \beta_2 = 0,9, \beta_3 = 12,5, \beta_4 = 13,0$ e $\beta_5 = 0,9$. Além disso, foram considerados três cenários para o parâmetro da precisão: ϕ pequeno, médio e grande, mais especificamente, $\phi = 20, 50$ e 150 . A Tabela 4.1 apresenta os percentuais de rejeição sob hipótese nula de que o modelo está corretamente especificado.

Tabela 4.1: Níveis empíricos para o tamanho do teste, usando os níveis nominais 1%, 5% e 10%.

		$\mu \in [0,0049; 0,1831]$			$\mu \in [0,2965; 0,8528]$			$\mu \in [0,8909; 0,9965]$		
$n \backslash \phi$		20	50	150	20	50	150	20	50	150
$\alpha = 1\%$										
40		10,20	6,46	4,38	2,62	2,30	2,34	7,22	4,24	1,54
80		5,88	3,60	1,72	1,06	1,00	1,08	4,24	2,20	0,86
120		3,66	2,64	1,56	1,04	0,70	0,54	3,40	1,44	0,36
$\alpha = 5\%$										
40		22,82	17,80	13,24	9,56	9,04	9,10	17,56	11,96	6,86
80		13,82	10,14	7,42	4,98	4,66	4,72	12,26	7,30	3,96
120		9,94	7,10	5,32	3,60	2,98	2,98	10,60	5,24	2,36
$\alpha = 10\%$										
40		31,44	26,52	21,68	16,82	15,90	15,26	25,02	18,08	12,84
80		20,48	16,14	12,66	8,72	8,64	8,84	18,64	12,60	7,84
120		15,62	12,26	9,70	7,00	6,38	5,98	16,78	9,96	5,24

Notamos que, no geral, os tamanhos do teste não estão próximos dos níveis nominais, isto é, as taxas de rejeições empíricas não estão próximas de $\alpha = 1\%, 5\%$ e 10% . Mais especificamente, notamos que no cenário em que μ é central, o aumento do tamanho amostral até $n = 80$ faz com que o tamanho do teste fique mais próximo do valor α , porém quando $n = 120$ os valores do tamanho do teste são distantes dos níveis nominais o que não é esperado. Considerando o cenário em que μ pertence ao extremo superior (próximo do um), percebemos que há uma certa melhora quando ϕ aumenta (com exceção

de $n = 120$). Quando μ pertence ao extremo inferior (próximo de zero), o aumento de n e ϕ , geralmente, leva a um melhor desempenho do tamanho do teste.

No Capítulo 3, vimos que a distribuição assintótica da estatística de teste é normal padrão, no entanto, pelos resultados obtidos há indícios de que a convergência em distribuição, para essa classe de modelos parece não ser adequada. Nas Figuras 4.1 a 4.9 temos os histogramas de U_n com as curvas teóricas da distribuição normal padrão e os respectivos *qqnorm*'s (gráfico quantil-quantil para a distribuição normal) de U_n com uma reta pontilhada (comparação quantil U_n e quantil normal padrão) e outra reta (comparação com uma normal de parâmetros baseados na média e erro-padrão amostral de U_n , respectivamente.), considerando todos os cenários de n , μ e ϕ obtidos das simulações. Analisando os histogramas, percebemos que em todos os cenários as distribuições de U_n são fortemente assimétricas e mais leptocúrticas do que a curva da distribuição normal padrão. Em relação à comparação dos quantis de U_n com os quantis normais, vemos pelos *qqnorm*'s que a distribuição de U_n não se aproxima nem da distribuição normal padrão, nem de uma distribuição normal de parâmetros média e erro-padrão qualquer, pois nas caudas os quantis da U_n ficam muito distantes dos quantis da distribuição normal. Ainda, percebemos que o aumento do tamanho amostral não leva a melhora significativa no sentido de se aproximar da distribuição normal padrão. Também, o aumento de ϕ faz com que a distribuição de U_n em alguns casos seja mais leptocúrtica. Verificamos também que a diferença entre a distribuição empírica de U_n e a distribuição normal padrão é maior quando consideramos os extremos de μ . Ainda, calculamos a média, variância, assimetria e curtose dos 5000 valores da estatística U_n em uma das situações em que a distribuição de U_n mais se aproximou da distribuição normal e onde obtivemos melhores tamanhos dos testes, isto é, consideramos μ central e $\phi = 150$. Na Tabela 4.2 estão apresentadas essas medidas. Portanto, teoricamente, os valores para essas medidas deveriam ser 0, 1, 0 e 3, para média, variância, assimetria e curtose, respectivamente, pois são os valores assumidos pela distribuição normal padrão. Notamos que a variância está sendo subestimada (próximo a 0,5) e quando o tamanho amostral aumenta, a subestimação é maior. A média de U_n está consideravelmente distante do valor esperado (zero), porém esse problema aparenta ser amenizado de acordo com o aumento de n . A assimetria é negativa, aparentando não se aproximar do zero com o aumento de n . Por fim, a curtose está próxima do esperado que é 3. Decidimos avaliar a distribuição da estatística U_n para tamanhos amostrais maiores tais como $n = 500, 1000, 1500$ e 2000 , aqui consideramos 1000 réplicas. Novamente, construímos os gráficos e calculamos as medidas, ambos já citados. Na Tabela 4.3 estão apresentadas a média, variância, assimetria e curtose para as 1000 réplicas de U_n , como também, na Figura 4.10 temos os histogramas de U_n e os respectivos *qqnorm*'s. Notamos que, em relação as medidas, a média da estatística U_n está consideravelmente mais próxima de zero, do que nos cenário onde n era menor. O aumento do tamanho amostral também causa ainda uma maior subestimação do valor da variância de U_n , sendo esse o problema mais grave que percebemos. A assimetria é menor, mas não muito, do que os cenários anteriores de n . Por fim, a curtose aparenta possuir o mesmo comportamento

estando perto do valor esperado, 3. Através dos gráficos, vemos uma leve melhoria da distribuição de U_n em relação a assimetria. Notamos que a distribuição amostral de U_n , para esses tamanhos amostrais, não se aproxima de uma distribuição normal padrão. Na verdade, a estatística aparente possui uma distribuição assimétrica negativa, com média não muito distante de zero, quando consideramos grandes amostras e variância em torno de 0,5 (que é uma variância diferente do esperado, que seria igual a um). Concluimos através das distribuições de U_n , desses cenários, que para o modelo de regressão beta com dispersão fixa, a estatística U_n não possui distribuição normal padrão.

Tabela 4.2: Valores da média, variância, assimetria e curtose para as 5000 réplicas de U_n para o cenário μ central e $\phi = 150$.

Tamanho amostral	Média	Variância	Assimetria	Curtose
40	-0,8273	0,6290	-0,4630	2,9901
80	-0,5492	0,5818	-0,5757	3,3022
120	-0,4307	0,5286	-0,51755	3,1654

Tabela 4.3: Valores da média, variância, assimetria e curtose para as 1000 réplicas de U_n para o cenário μ central e $\phi = 150$.

Tamanho amostral	Média	Variância	Assimetria	Curtose
500	-0,3193	0,5503	-0,4563	3,1649
1000	-0,2067	0,4943	-0,5067	3,3193
1500	-0,1906	0,4899	-0,4118	3,2348
2000	-0,1809	0,5138	-0,4324	3,3123

Devido a tal desempenho, realizamos um melhoramento do teste no sentido de obter os quantis da distribuição da estatística U_n por outra via que não seja através da distribuição normal padrão. Assim utilizamos a metodologia *bootstrap* descrita na Seção 1.2.2 para a obtenção dos quantis empíricos da estatística U_n utilizando $B = 500$ réplicas *bootstrap*. Usando os mesmos cenários das simulações apresentadas na seção anterior, obtivemos os tamanhos do teste, agora, através dos quantis empíricos. Além disso, também avaliamos os poderes do teste através da introdução de algumas formas de erro de especificação do modelo. Primeiramente, avaliamos o poder do teste sob omissão de covariáveis, isto é, a variável resposta é gerada utilizando todas as covariáveis, mas a mesma é modelada usando uma, duas e três covariáveis, respectivamente. Nas Tabelas 4.4, 4.5 e 4.6, estão apresentados os percentuais de rejeição sob hipótese nula (tamanhos do teste - Cenário 4) e também os percentuais de rejeição sob a hipótese alternativa, ou seja, os poderes do teste, sob a omissão de covariáveis, considerando os níveis 1%, 5% e 10%, respectivamente. A partir dos resultados apresentados nessas tabelas, obtivemos algumas conclusões. Em relação ao tamanho do teste, percebemos a partir do Cenário 4 que o tamanho do teste foi controlado quando utilizamos os quantis empíricos obtidos via *bootstrap*. Na grande maioria das situações, quando n aumenta, o tamanho do teste apresenta um melhor desempenho. Em relação ao poder do teste, vimos que quanto maior o nível nominal fixado, maior o poder do teste obtido, como esperado.

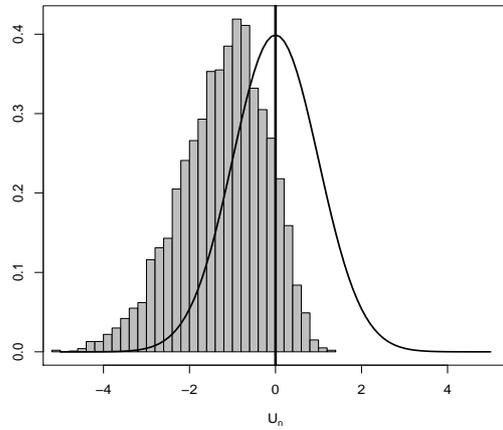
Tabela 4.4: Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	79,40	90,38	94,62	72,44	85,22	94,20	41,50	49,88	46,42	1,50	1,64	1,30
80	98,00	98,52	99,66	97,54	98,28	99,76	87,56	83,12	74,82	1,56	1,56	1,70
120	99,74	99,84	99,98	99,78	99,88	100,00	98,30	96,72	94,70	1,56	1,50	1,32
$\mu \in [0, 2965; 0, 8528]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	26,36	55,52	86,90	24,94	51,88	83,58	1,30	0,42	0,04	1,38	1,60	1,30
80	40,82	81,02	99,64	35,94	74,30	98,88	0,72	0,34	0,04	1,38	1,56	1,40
120	61,12	96,54	100,00	52,32	93,36	100,00	0,88	0,08	0,00	1,36	1,34	1,20
$\mu \in [0, 8909; 0, 9965]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	50,80	76,52	79,22	44,84	71,04	79,66	18,66	30,90	20,36	1,50	1,56	1,40
80	96,44	98,68	98,88	93,94	97,40	97,88	76,20	77,84	49,08	1,12	1,38	1,62
120	99,86	99,88	99,96	99,72	99,86	99,80	96,96	95,30	74,90	1,40	1,62	1,12

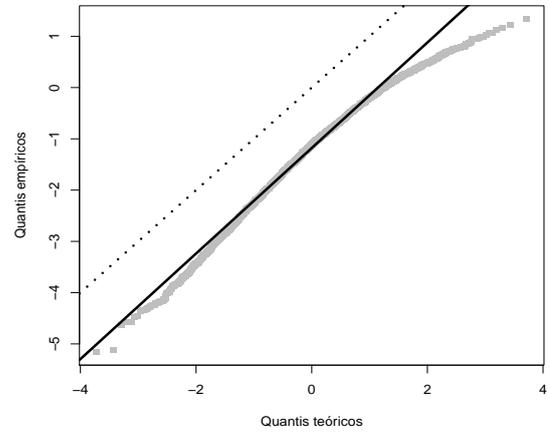
Tabela 4.5: Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	95,74	98,80	99,64	92,52	97,76	99,48	70,72	80,26	84,48	5,46	5,56	5,18
80	99,94	99,94	100,00	99,86	99,88	99,96	97,52	96,62	96,02	5,78	6,04	5,58
120	100,00	100,00	100,00	99,98	100,00	100,00	99,88	99,72	99,56	5,86	5,58	5,28
$\mu \in [0, 2965; 0, 8528]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	64,20	91,70	99,86	55,68	85,94	99,58	5,70	3,56	0,78	5,42	5,90	5,78
80	74,60	97,48	100,00	66,46	94,36	99,98	4,20	1,88	0,22	4,70	5,26	5,04
120	86,66	99,76	100,00	79,14	99,20	100,00	4,14	1,32	0,08	5,48	5,68	5,38
$\mu \in [0, 8909; 0, 9965]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	85,84	96,18	98,30	75,56	90,60	95,14	45,12	59,82	48,48	5,80	5,46	5,80
80	99,74	99,96	100,00	99,14	99,80	99,80	92,96	93,24	78,24	5,26	5,52	5,36
120	100,00	100,00	100,00	99,98	100,00	100,00	99,48	99,20	93,08	5,48	5,64	4,70

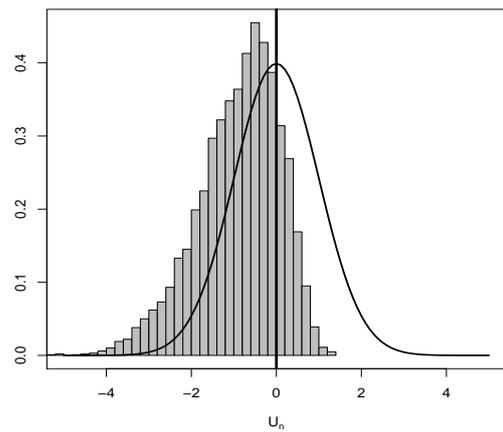
Figura 4.1: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 0049; 0, 1831]$ e $\phi = 20$.



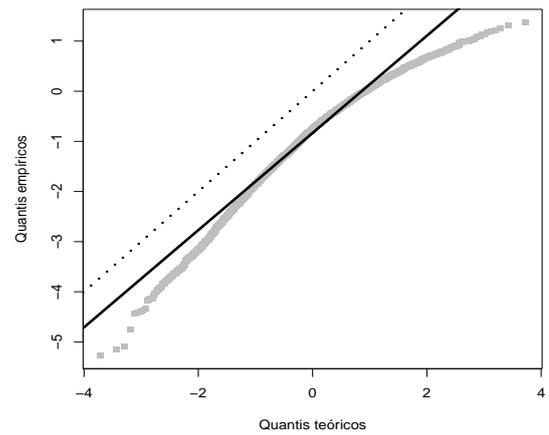
(a) $n = 40$



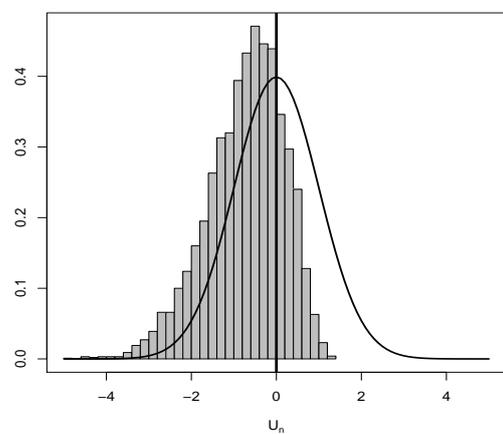
(b) $n = 40$



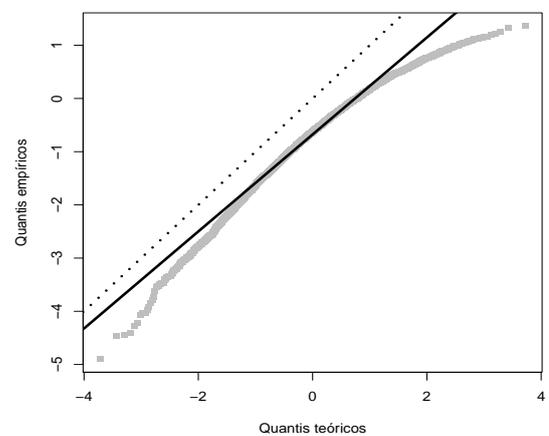
(c) $n = 80$



(d) $n = 80$



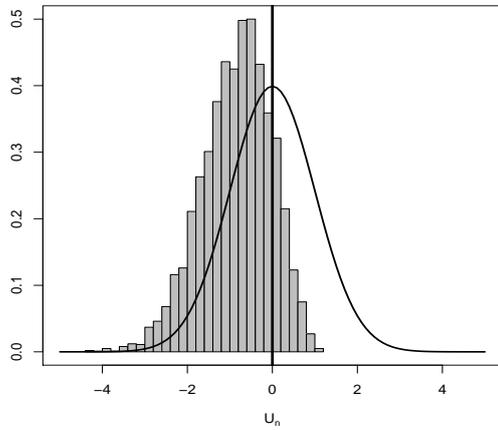
(e) $n = 120$



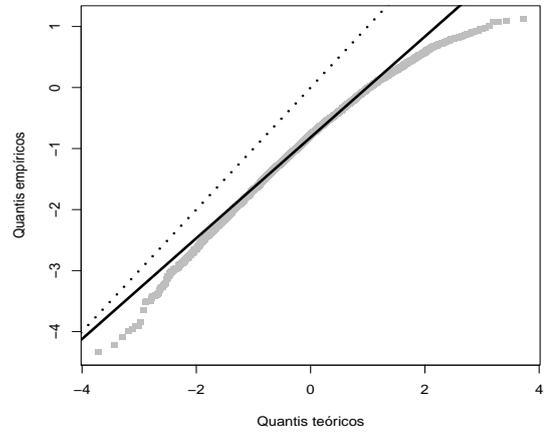
(f) $n = 120$

Fonte: A autora.

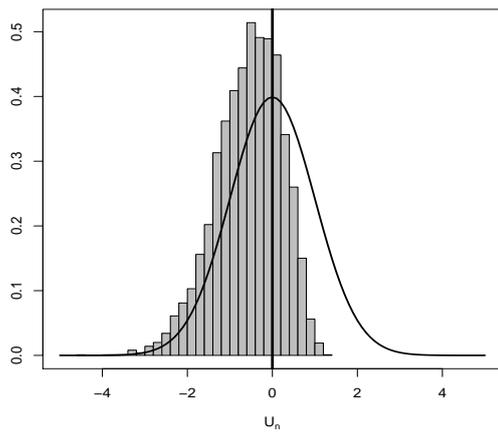
Figura 4.2: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 20$.



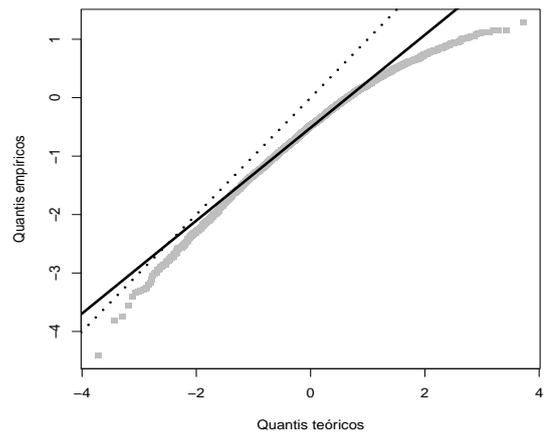
(a) $n = 40$



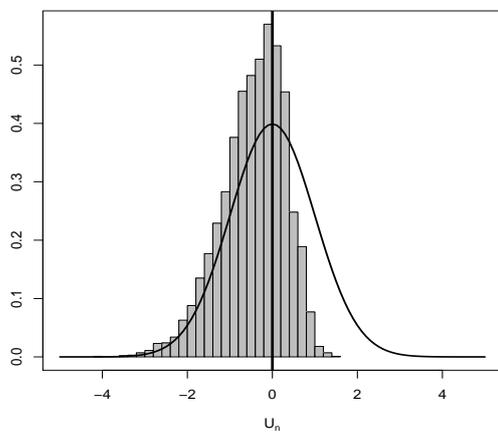
(b) $n = 40$



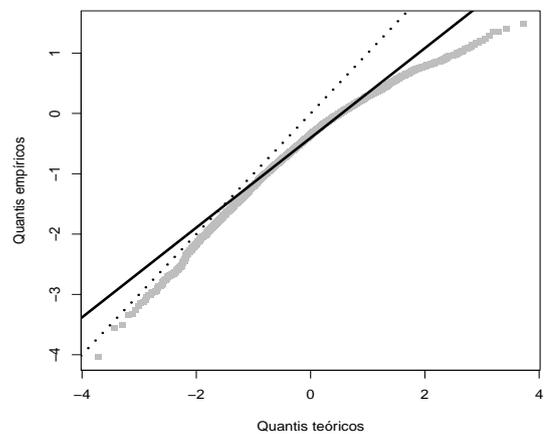
(c) $n = 80$



(d) $n = 80$



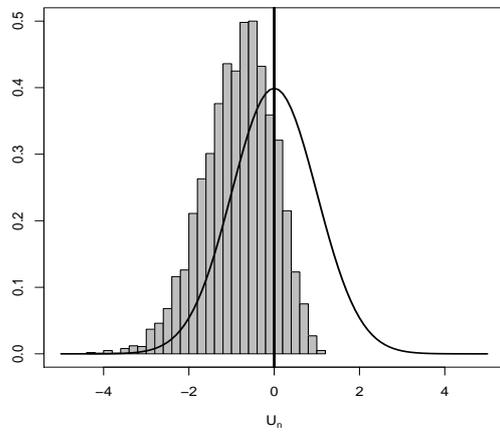
(e) $n = 120$



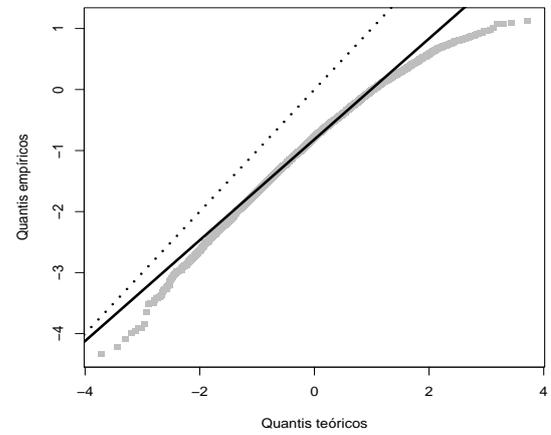
(f) $n = 120$

Fonte: A autora.

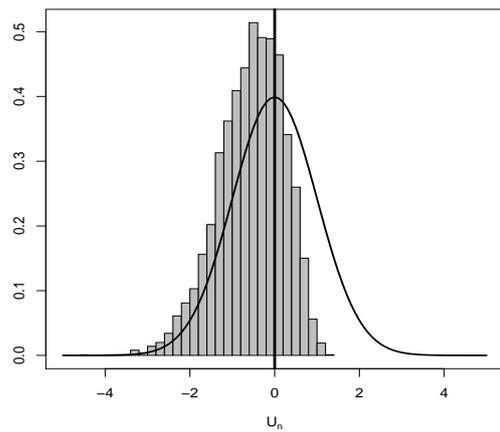
Figura 4.3: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 20$.



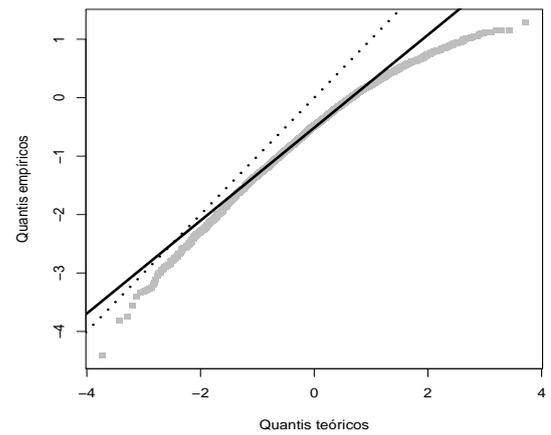
(a) $n = 40$



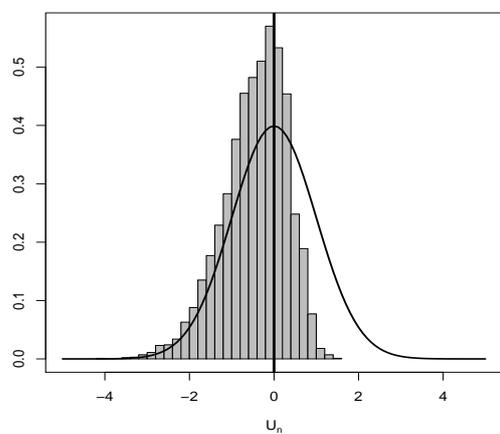
(b) $n = 40$



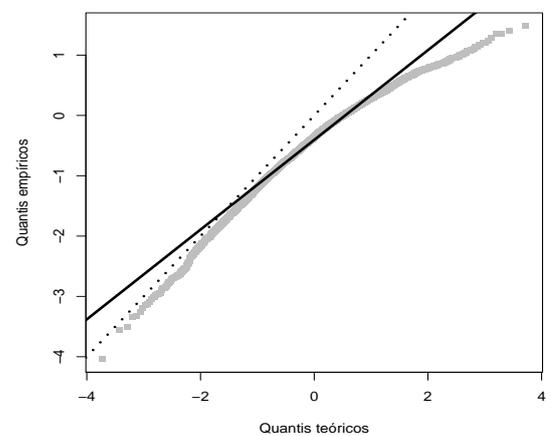
(c) $n = 80$



(d) $n = 80$



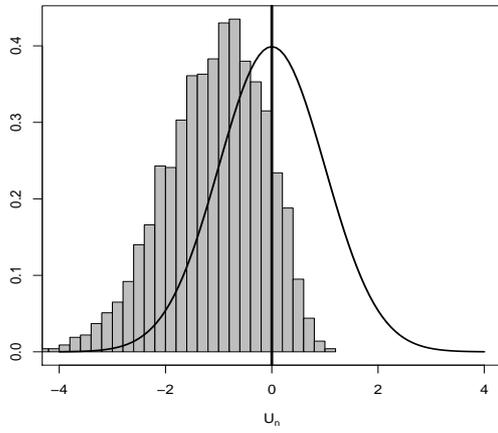
(e) $n = 120$



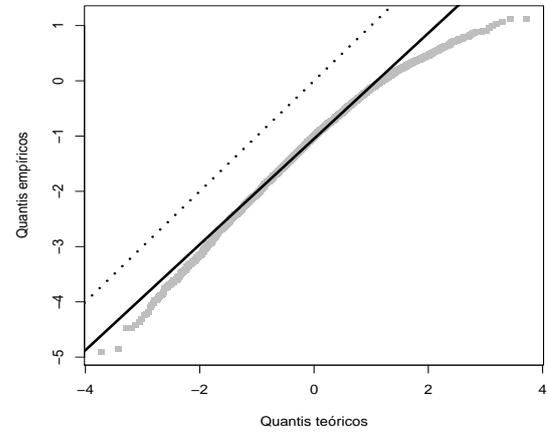
(f) $n = 120$

Fonte: A autora.

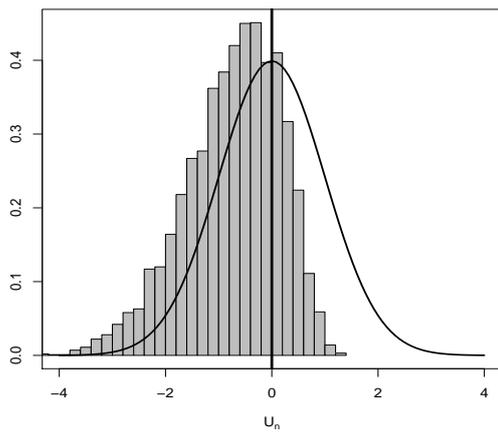
Figura 4.4: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0,0049; 0,1831]$ e $\phi = 50$.



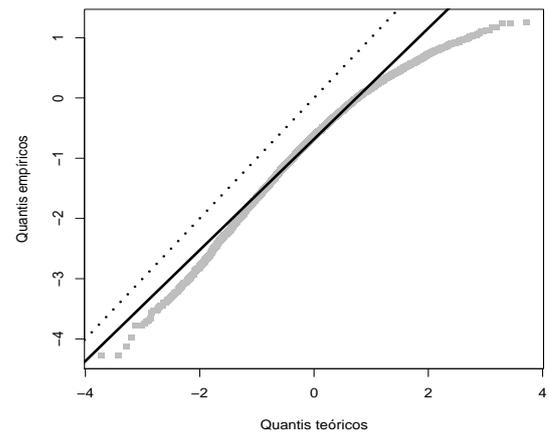
(a) $n = 40$



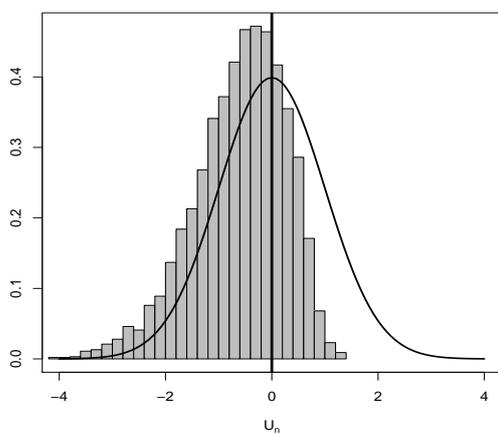
(b) $n = 40$



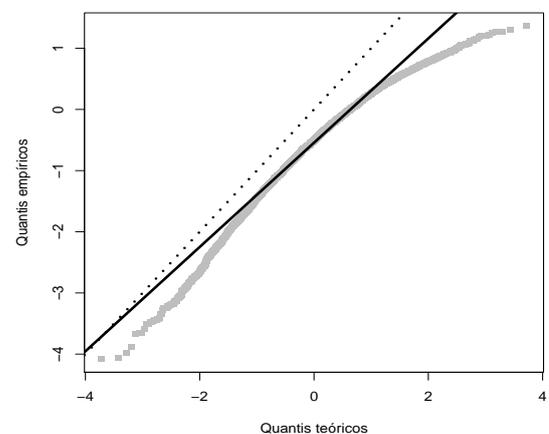
(c) $n = 80$



(d) $n = 80$



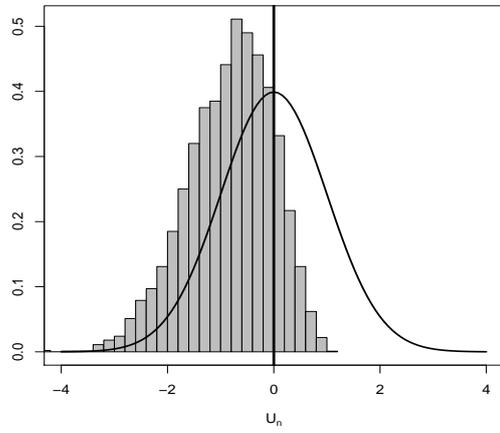
(e) $n = 120$



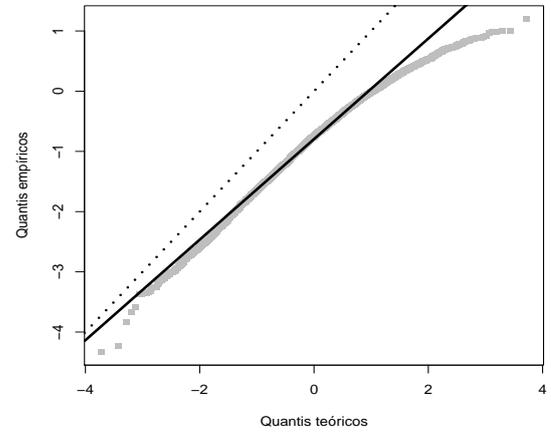
(f) $n = 120$

Fonte: A autora.

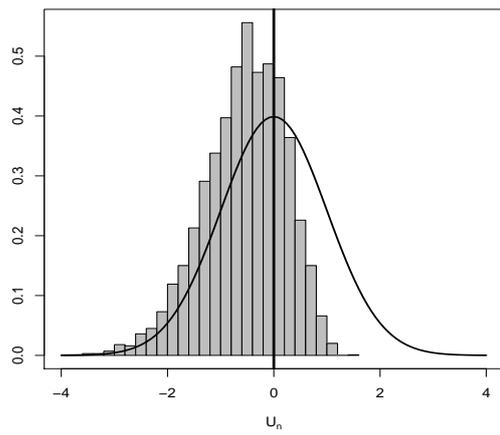
Figura 4.5: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 50$.



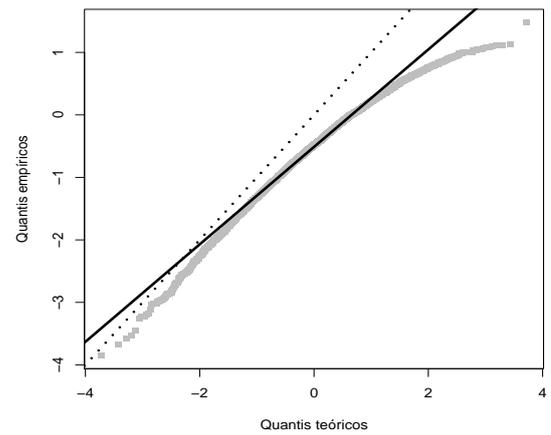
(a) $n = 40$



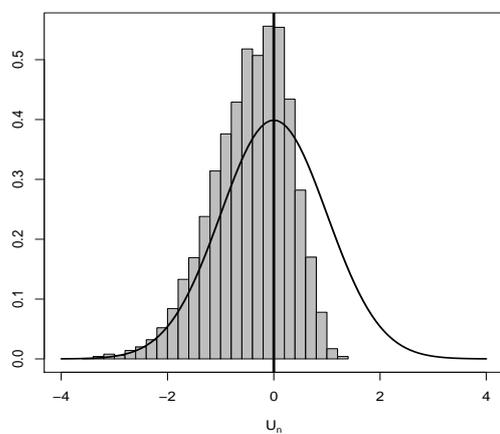
(b) $n = 40$



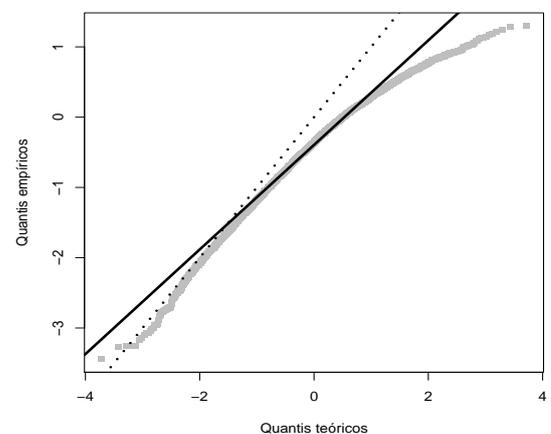
(c) $n = 80$



(d) $n = 80$



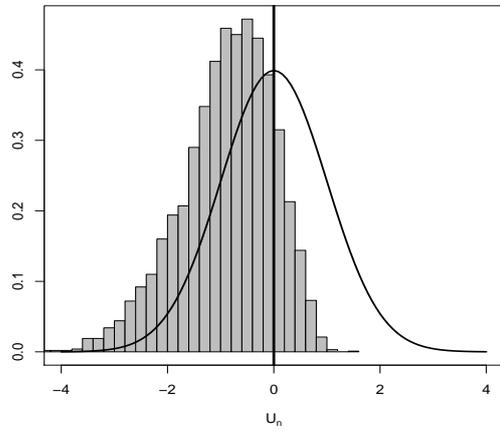
(e) $n = 120$



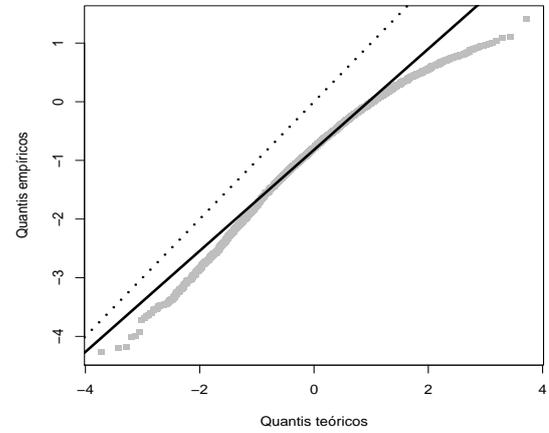
(f) $n = 120$

Fonte: A autora.

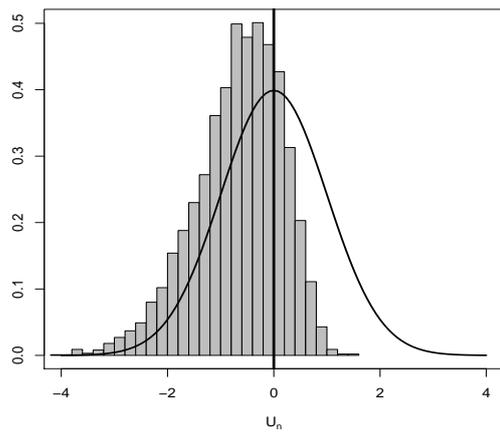
Figura 4.6: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 50$.



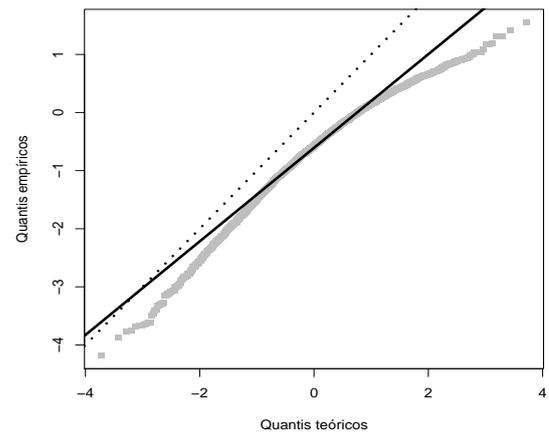
(a) $n = 40$



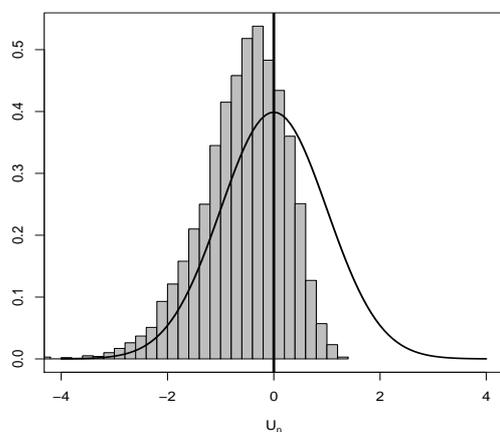
(b) $n = 40$



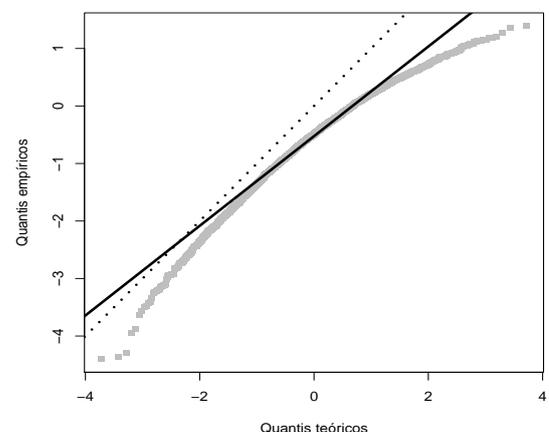
(c) $n = 80$



(d) $n = 80$



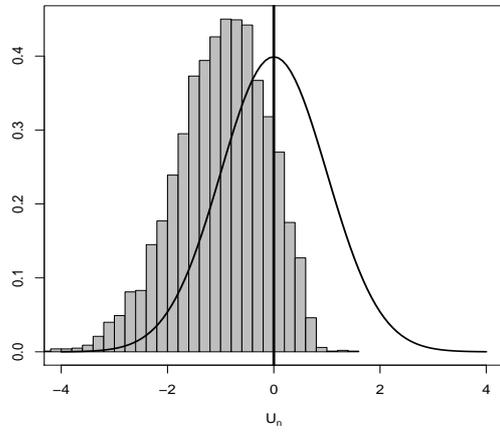
(e) $n = 120$



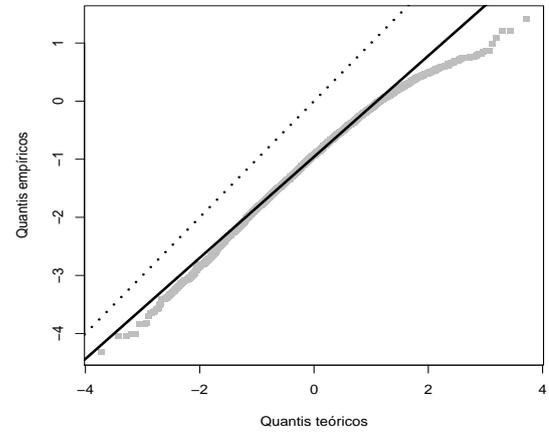
(f) $n = 120$

Fonte: A autora.

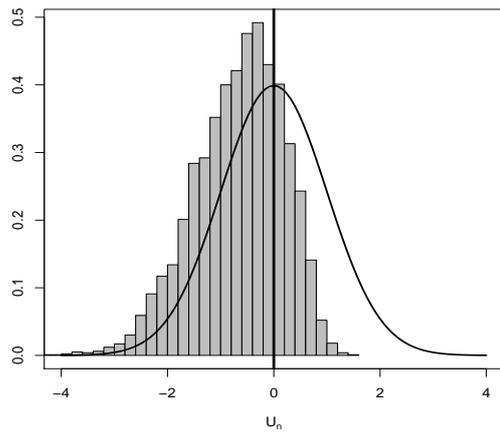
Figura 4.7: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0,0049; 0,1831]$ e $\phi = 150$.



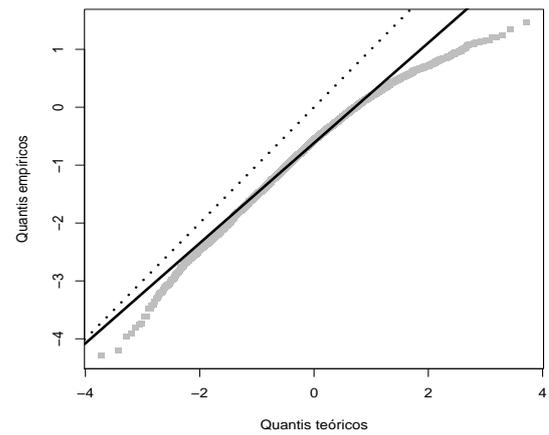
(a) $n = 40$



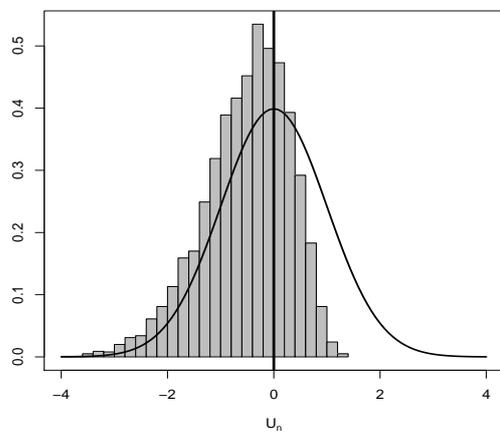
(b) $n = 40$



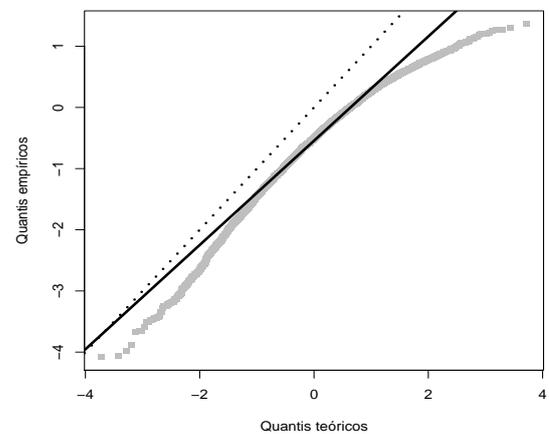
(c) $n = 80$



(d) $n = 80$



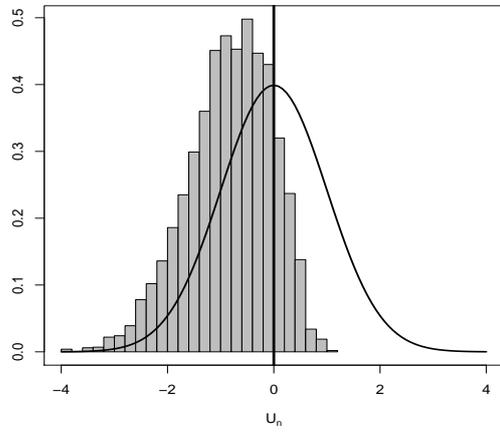
(e) $n = 120$



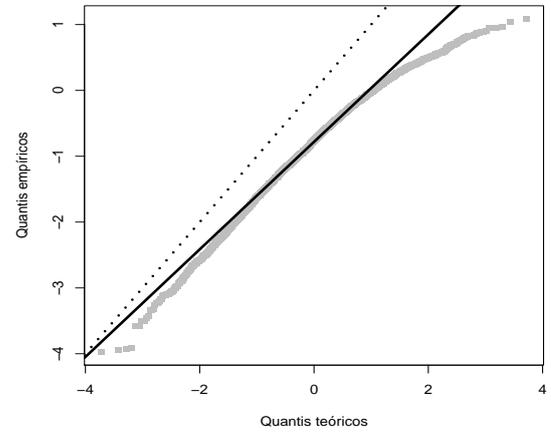
(f) $n = 120$

Fonte: A autora.

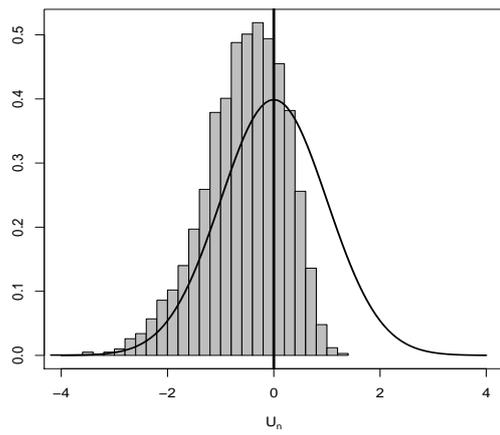
Figura 4.8: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2965; 0, 8528]$ e $\phi = 150$.



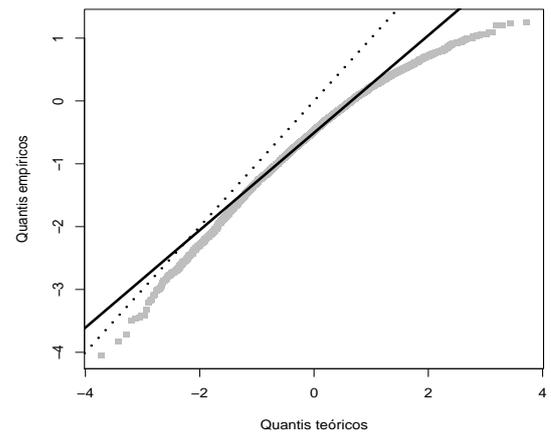
(a) $n = 40$



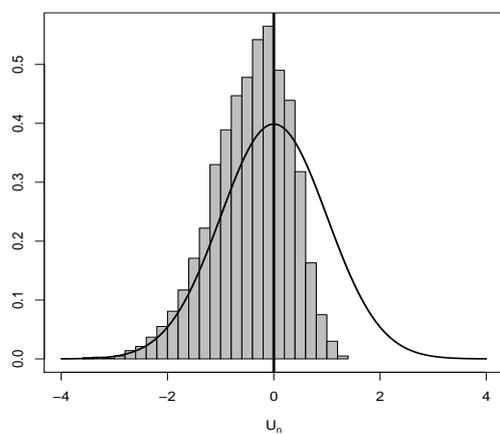
(b) $n = 40$



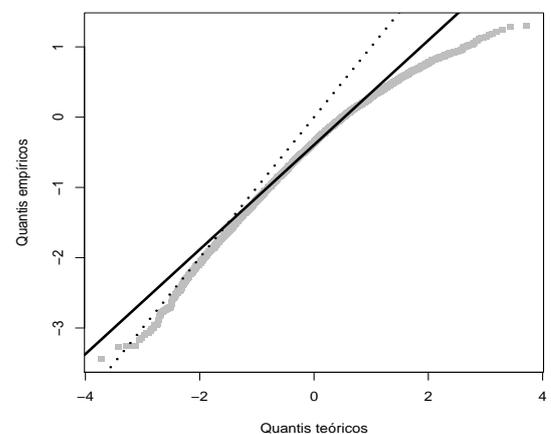
(c) $n = 80$



(d) $n = 80$

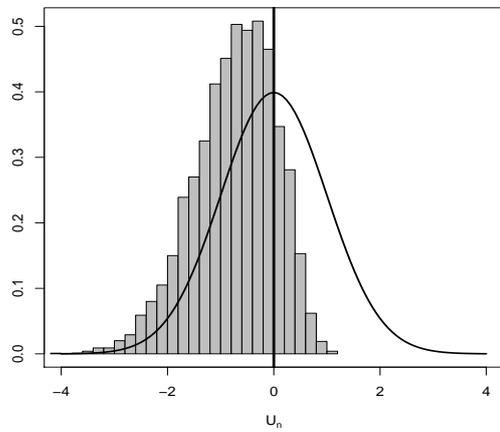


(e) $n = 120$

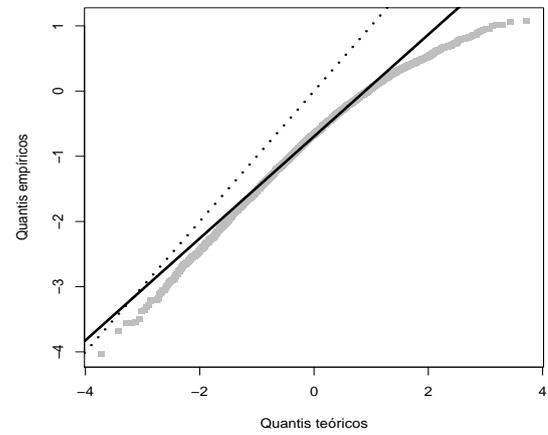


(f) $n = 120$

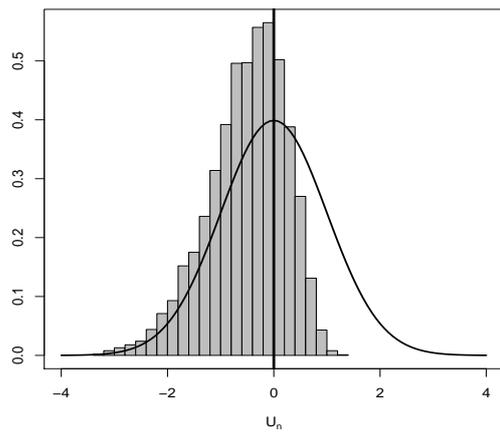
Figura 4.9: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8909; 0, 9965]$ e $\phi = 150$.



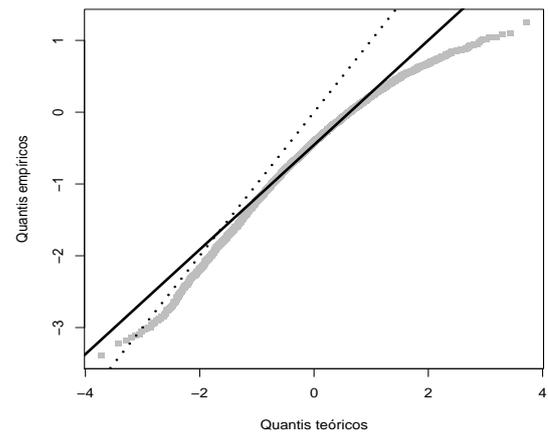
(a) $n = 40$



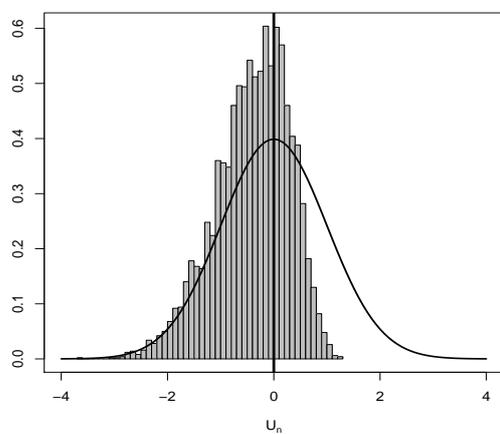
(b) $n = 40$



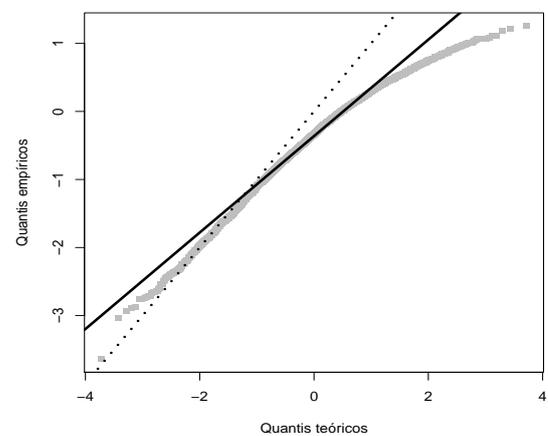
(c) $n = 80$



(d) $n = 80$

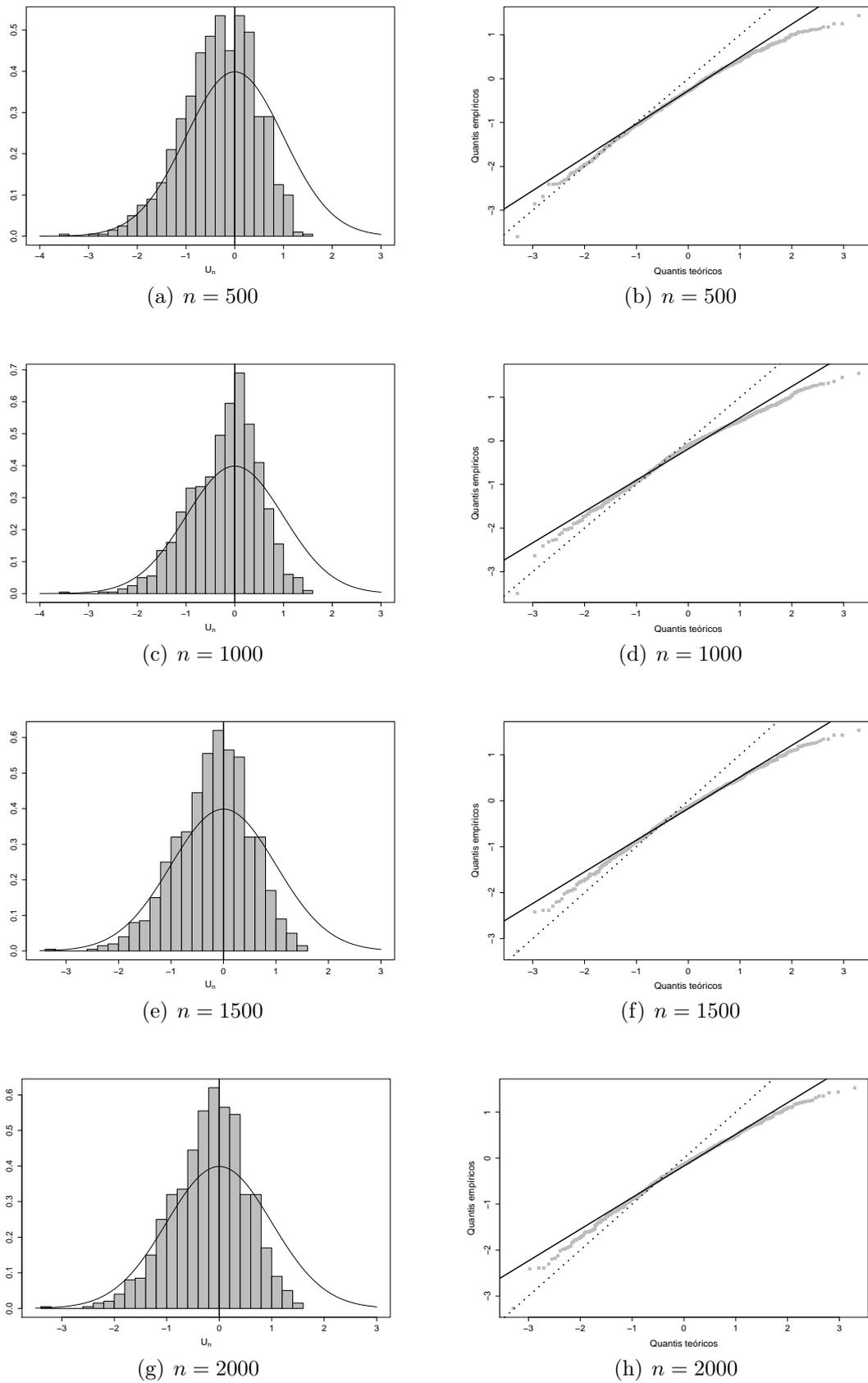


(e) $n = 120$



(f) $n = 120$

Figura 4.10: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2835; 0, 8947]$ e $\phi = 150$.



Fonte: A autora.

Tabela 4.6: Níveis empíricos para o teste, usando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	98,62	99,82	99,96	97,28	99,48	99,96	85,18	91,40	95,00	10,70	10,90	10,12
80	100,00	99,98	100,00	99,98	99,98	100,00	99,04	99,08	98,96	10,92	10,90	10,56
120	100,00	100,00	100,00	100,00	100,00	100,00	99,96	99,92	99,94	11,28	10,72	10,72
$\mu \in [0, 2965; 0, 8528]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	78,08	97,34	100,00	70,98	93,94	99,98	11,54	9,02	3,68	11,20	11,08	9,78
80	87,00	99,40	100,00	78,66	98,16	100,00	8,12	4,60	0,88	9,16	10,30	9,64
120	93,68	99,92	100,00	87,76	99,82	100,00	8,18	3,78	0,52	10,84	10,74	10,16
$\mu \in [0, 8909; 0, 9965]$												
$n \backslash \phi$	20	50	150	20	50	150	20	50	150	20	50	150
40	94,22	98,98	99,72	87,36	96,14	98,36	62,84	74,92	67,44	10,70	10,30	11,16
80	99,92	99,98	100,00	99,78	99,94	99,98	96,74	96,96	89,06	9,90	9,86	10,28
120	100,00	100,00	100,00	100,00	100,00	100,00	99,90	99,74	97,42	9,84	10,44	9,46

Tabela 4.7: Níveis empíricos para o teste, usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.

$n \backslash \phi$		$\mu \in [0, 0103; 0, 1092]$			$\mu \in [0, 2131; 0, 8289]$			$\mu \in [0, 9101; 0, 9886]$		
		20	55	150	20	55	150	20	55	150
$\alpha = 1\%$										
40		18,44	20,16	23,46	1,12	1,46	1,68	15,50	18,48	20,16
80		41,36	48,06	51,16	1,48	1,90	2,28	35,96	43,04	47,64
120		58,80	70,32	73,20	1,44	2,36	2,12	53,18	62,92	67,82
$\alpha = 5\%$										
40		40,24	39,86	41,78	5,50	6,00	6,16	37,44	37,80	37,04
80		68,46	69,62	70,98	5,94	6,48	6,60	63,82	65,94	67,80
120		82,30	86,06	87,18	5,80	7,44	7,42	79,04	81,92	83,52
$\alpha = 10\%$										
40		53,26	51,52	52,64	10,24	11,00	11,62	51,54	48,88	48,20
80		79,78	78,80	79,56	11,40	11,92	12,54	76,70	75,22	76,74
120		90,38	91,52	91,98	11,14	13,04	13,50	87,58	88,88	89,48

As conclusões a seguir são referentes aos três níveis nominais. Como esperado, à medida que é omitido um número maior de covariáveis, maior é o poder obtido. Quando μ pertence ao extremo inferior, o teste possui poder maior do que quando μ pertence ao extremo superior, o que é mais evidenciado quando $n = 40$. Vale ressaltar também que quando μ é central, o teste apresenta poderes mais baixos que nos outros dois cenários. No Cenário 3 quando μ é central, obtivemos poderes muito baixos. Podemos explicar o desempenho do teste observado nesse último cenário lembrando que μ é central, isto é, a convergência do processo de estimação dos parâmetros do modelo não é complicada como nos cenários em que μ pertence aos extremos, como também, a associação do valor $\beta_4 = 13$ (alto), a covariável lognormal padrão (conhecida por gerar pontos de alavanca) fazem com que quando passamos do Cenário 2 para o 3 (adicionamos a terceira covariável) estamos praticamente introduzindo toda a informação restante que estava faltando para

explicar a média, sabendo que a última covariável utilizada é uma Bernoulli associada ao valor $\beta_5 = 0,9$ que não traz tanta informação para a explicação da média. Por último, verificamos que, em geral, para os Cenários 1 e 2 quando a precisão ϕ aumenta, os poderes do teste crescem, não ocorrendo apenas no Cenário 3.

Outro tipo de má especificação do modelo que foi utilizada é gerar a variável resposta, através de outra distribuição com suporte no intervalo $(0, 1)$. Consideramos a distribuição simplex. Assim, a variável resposta foi gerada através da distribuição simplex, e foi modelada utilizando a regressão beta com dispersão fixa. Vale ressaltar que as covariáveis utilizadas para esta situação foram geradas a partir da distribuição uniforme padrão. Também, os parâmetros da componente sistemática utilizados para o extremo inferior de μ foram $\beta_1 = -3,0, \beta_2 = -1,2, \beta_3 = 0,6, \beta_4 = 1,0$ e $\beta_5 = -1,1$. Para o extremo superior de μ os parâmetros foram $\beta_1 = 3,5, \beta_2 = 1,2, \beta_3 = -0,8, \beta_4 = -1,1$ e $\beta_5 = 0,5$. Por fim, para μ central $\beta_1 = -1,9, \beta_2 = 1,2, \beta_3 = 1,0, \beta_4 = 1,1$ e $\beta_5 = 1,3$. Além disso, foram considerados três valores para a dispersão $\phi = 20, 55$ e 150 . Novamente utilizando *bootstrap* com $B = 500$ réplicas obtivemos os quantis empíricos, procedimento esse que será utilizado para todas os demais estudos de simulação. A Tabela 4.7 apresenta as taxas de rejeição dos testes para esse erro de especificação, isto é, os poderes do teste para esse tipo de má especificação.

Observamos que, a medida que o tamanho amostral e a precisão aumentam, os poderes também aumentam. Percebemos também que os poderes possuem desempenho semelhante para os extremos de μ . Notemos que para o cenário em que μ é central, os poderes do teste são muito baixos, praticamente próximos ao que se espera para o tamanho do teste. A partir desse resultado, entendemos que para este teste, quando utilizamos o modelo de regressão beta com dispersão fixa para ajustar a média de uma variável resposta qualquer em que $\mu_i \approx 0,5, i = 1, 2, \dots, n$, não será simples identificar se o modelo beta é o verdadeiro gerador de dados, pois o modelo simplex possivelmente também irá produzir um bom ajuste.

O último tipo de má especificação que introduzimos foi gerar a variável resposta, através de um modelo de regressão beta com dispersão variável e estimar através do modelo de regressão beta com dispersão fixa. Vale ressaltar que as covariáveis utilizadas para tal situação foram geradas a partir da distribuição uniforme padrão. Os parâmetros referentes a componente sistemática da média para o extremo inferior de μ foram $\beta_1 = -1,2, \beta_2 = -1,2, \beta_3 = -1,0, \beta_4 = -1,1$ e $\beta_5 = -1,1$. Para o extremo superior de μ utilizamos $\beta_1 = 1,5, \beta_2 = 1,1, \beta_3 = 1,3, \beta_4 = 0,1$ e $\beta_5 = 1,3$ para a média. Por fim, para μ central consideramos $\beta_1 = -1,6, \beta_2 = 0,5, \beta_3 = 0,8, \beta_4 = 1,1$ e $\beta_5 = 1,3$. Para introduzir a dispersão variável, a especificação do modelo para a dispersão contou com covariáveis geradas através da distribuição uniforme padrão e da distribuição Bernoulli de parâmetro $0,1$. Ainda, para os parâmetros que compõem a componente sistemática da dispersão, consideramos $\gamma_1 = 3,0, \gamma_2 = 3,3, \gamma_3 = -1,0, \gamma_4 = -2,3$ e $\gamma_5 = 2,1$. Consequentemente, foram gerados cinco valores para o grau de heterogeneidade: $\lambda = 1, 25, 54, 121$ e 326 . Notemos que quando $\lambda = 1$ não temos dispersão variável nos dados. Os resultados para

os poderes do teste, nesse caso, estão apresentados na Tabela 4.8.

Vemos que o tamanho empírico do teste (quando $\lambda = 1$) está próximo dos níveis nominais fixados para todos os cenários, como esperado. Em relação ao poder do teste, de acordo com aumento dos valores de n e λ , esses crescem significativamente. Os poderes, para as situações em que μ está nas extremidades, possuem desempenhos semelhantes. O cenário em que o poder foi mais baixo, novamente, foi quando μ é central.

Portanto, constatamos que utilizando a metodologia *bootstrap* para obter os quantis da estatística de teste U_n obtivemos ótimos tamanhos de teste para todos os cenários apresentados. Além disso, em relação ao poder do teste, vimos que o teste consegue detectar na maioria das situações com bons poderes as más especificações introduzidas: omissão de covariáveis, variável resposta simplex, variável resposta com dispersão variável.

4.1.1 Teste com Variância Bootstrap

Conforme vimos na Seção 4.1, houveram algumas situações em que o desempenho do teste apresentou comportamentos que não são esperados. Por exemplo, no caso em que avaliamos o poder do teste no que se diz respeito à omissão de covariáveis, no Cenário 3 em que μ é central, os poderes são muito baixos (menor que os valores dos tamanhos do teste) e ainda quando n e ϕ aumentam, os poderes do teste não possuem um bom desempenho, o que não acontece nas demais situações. Dessa forma, realizamos outro melhoramento no teste no sentido de utilizar uma variância que estabilize a estatística de teste com o intuito de não ocorrer mais essas incoerências. Para tanto, utilizamos novamente a metodologia *bootstrap*, assim, refizemos os estudos de simulação apresentados na Seção 4.1 para efeito de comparação utilizando a variância *bootstrap* em vez de utilizar a expressão apresentada na Subseção 3.2.1. Nas Tabelas 4.9, 4.10 e 4.11 estão apresentados os percentuais de rejeição sob hipótese nula (tamanhos do teste), como também, os percentuais de rejeição sob a hipótese alternativa, ou seja, os poderes do teste sob a omissão de covariáveis.

Comparando esses resultados com os obtidos na Seção 4.1, chegamos a algumas conclusões. Primeiramente, notamos que alguns resultados permanecem iguais, por exemplo, quando n aumenta, o tamanho do teste apresentado no Cenário 4 fica mais próximo do nível nominal. Com relação ao poder do teste, também temos que α maior produz um maior poder do teste. Considerando os três níveis nominais, como esperado, quando omitimos mais covariáveis o poder obtido é maior. Também os poderes dos testes são maiores quando μ está próximo de zero do que quando μ está próximo do um (mais evidente em $n = 40$). No entanto, em relação a algumas situações específicas, houve resultados distintos. Notamos que os poderes para o Cenário 3 em que μ é central, aumentam, em geral, à medida que n e ϕ crescem, o que não ocorre com o caso anterior (ver Seção 4.1). Também, percebemos que na grande maioria dos cenários, quando a precisão ϕ aumenta, os poderes crescem, apenas em duas situações no Cenário 3 isso não ocorre quando $\phi = 150$.

Tabela 4.8: Níveis empíricos para o teste, usando os níveis nominais 1, 5% e 10%. Má especificação: dispersão não modelada.

$\lambda \backslash n$		$\mu \in [0, 0107; 0, 1477]$					$\mu \in [0, 2515; 0, 7738]$					$\mu \in [0, 8825; 0, 9892]$				
		1	25	54	121	326	1	25	54	121	326	1	25	54	121	326
$\alpha = 1\%$																
40	1,48	27,02	41,08	55,30	56,26	1,40	6,84	12,06	29,66	31,26	1,54	26,70	37,40	47,06	49,06	
80	1,36	60,38	79,08	92,56	93,38	1,48	23,20	35,18	72,02	73,90	1,46	59,28	76,18	87,74	90,58	
120	1,34	80,96	93,58	99,04	99,28	1,26	41,12	58,12	91,32	93,16	1,54	79,44	92,06	97,76	98,56	
$\alpha = 5\%$																
40	5,18	46,78	61,26	74,80	76,44	5,58	20,46	26,82	52,34	54,14	5,98	46,70	58,32	68,76	70,48	
80	5,84	78,82	90,10	97,56	98,04	5,34	44,04	56,48	87,00	88,28	5,42	77,70	88,76	95,40	96,88	
120	5,48	92,00	97,90	99,78	99,84	5,34	63,46	77,30	97,46	98,00	5,78	91,58	97,20	99,60	99,78	
$\alpha = 10\%$																
40	10,29	58,08	71,30	82,86	84,30	10,86	30,34	37,26	64,56	65,38	10,96	58,00	68,66	78,38	79,94	
80	11,12	85,24	93,66	98,84	99,02	10,16	56,20	67,52	92,22	93,06	10,30	85,22	92,84	97,40	98,66	
120	10,38	95,44	98,94	99,90	99,96	10,18	73,84	85,00	98,54	98,96	10,60	94,96	98,68	99,80	99,94	

Tabela 4.9: Níveis empíricos para o teste com variância *bootstrap*, considerando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	38,82	61,32	85,14	37,24	56,84	82,50	20,92	31,60	49,26	1,24	1,72	1,52
80	62,96	74,00	96,28	59,42	65,66	93,40	44,28	44,00	67,16	1,42	1,70	1,46
120	78,58	88,18	99,66	71,32	78,28	99,36	59,40	58,46	87,96	1,44	1,30	1,54
$\mu \in [0, 2965; 0, 8528]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	13,36	46,00	96,32	8,86	38,04	94,04	1,44	1,30	1,26	1,32	1,54	1,38
80	17,86	64,38	99,90	12,56	55,96	99,72	1,34	1,68	2,72	1,28	1,48	1,22
120	27,80	86,12	100,00	20,96	79,18	100,00	1,36	1,54	2,12	1,26	1,36	1,18
$\mu \in [0, 8909; 0, 9965]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	24,02	37,76	62,66	20,54	36,52	52,92	10,98	17,28	18,26	1,54	1,36	1,26
80	58,44	59,74	77,76	52,60	57,32	59,96	37,86	35,88	27,34	1,34	1,16	0,90
120	81,40	76,54	92,66	72,20	67,36	74,34	58,88	48,64	39,40	1,18	1,48	1,62

Tabela 4.10: Níveis empíricos para o teste com variância *bootstrap*, considerando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	84,12	93,78	98,50	79,34	91,80	98,96	57,72	71,26	85,92	5,74	5,72	5,20
80	95,88	98,04	99,96	92,90	96,00	99,90	82,62	82,84	94,72	5,66	5,50	5,56
120	98,58	99,72	100,00	97,04	98,96	100,00	91,80	92,14	99,26	5,62	5,60	5,54
$\mu \in [0, 2965; 0, 8528]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	43,26	82,30	99,90	36,46	76,22	99,74	5,58	6,70	6,10	5,38	6,16	5,60
80	50,14	91,86	100,00	42,20	88,16	100,00	5,30	6,20	7,34	4,88	5,14	4,90
120	62,56	98,50	100,00	55,46	97,32	100,00	5,66	5,82	7,22	5,06	6,12	5,06
$\mu \in [0, 8909; 0, 9965]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	64,40	81,64	94,88	57,72	76,02	88,40	36,00	47,16	49,76	6,16	5,46	5,38
80	92,76	95,22	98,70	88,00	90,20	92,72	74,94	72,88	63,18	5,58	5,34	5,50
120	99,16	98,74	99,92	96,72	95,66	97,70	91,12	85,02	77,78	5,20	5,58	5,42

Por último, notamos também que houve poderes mais baixos para vários cenários. Dessa forma, apesar de em algumas situações os poderes do teste terem sido mais baixos quando comparados com os resultados apresentados na seção anterior, obtivemos indícios de que quando a variância *bootstrap* é utilizada, os resultados foram mais coerentes do que quando usamos a variância apresentada no Capítulo 3 devido ao fato da melhora observada para o Cenário 3 em que μ é central. Na Tabela 4.12 apresentamos os tamanhos e poderes do teste, para a má especificação de gerar a variável resposta através da distribuição simplex e modelar utilizando a regressão beta. Comparando com os resultados obtidos na seção anterior, observamos um comportamento distinto do que vimos anteriormente: à medida que a precisão aumenta, os poderes diminuem. Ainda, notamos que obtemos poderes mais baixos que anteriormente. Sendo assim, não vimos vantagem em utilizar a variância

Tabela 4.11: Níveis empíricos para o teste com variância *bootstrap*, considerando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0049; 0, 1831]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	95,50	98,54	99,82	92,86	97,88	99,86	77,98	88,20	95,30	10,52	10,52	10,48
80	99,54	99,82	100,00	98,62	99,54	100,00	94,50	94,44	98,78	10,36	10,32	10,48
120	99,84	99,98	100,00	99,44	99,94	100,00	98,30	98,72	99,90	10,72	10,88	10,26
$\mu \in [0, 2965; 0, 8528]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	63,88	93,66	99,98	56,76	89,76	99,98	11,02	12,86	13,14	10,70	11,10	9,84
80	69,80	97,80	100,00	62,96	96,18	100,00	10,52	11,94	12,72	9,40	9,96	9,80
120	80,74	99,70	100,00	74,24	99,40	100,00	10,52	11,56	12,50	10,42	10,78	10,44
$\mu \in [0, 8909; 0, 9965]$												
ϕ	20	50	150	20	50	150	20	50	150	20	50	150
n												
40	85,52	94,86	98,96	77,84	89,46	96,52	55,58	66,56	69,70	11,26	10,74	10,14
80	98,76	99,28	99,88	97,18	97,48	98,74	89,58	87,70	81,60	10,30	10,60	10,74
120	99,94	100,00	100,00	99,48	99,18	99,80	97,82	95,38	91,46	10,38	10,40	10,38

bootstrap quando o erro de má especificação é gerar através da simplex e modelar utilizando a regressão beta. Por fim, na Tabela 4.13 apresentamos as taxas de rejeições para a má especificação gerar com dispersão variável e estimar com dispersão fixa. A única diferença que notamos foi uma leve melhora nos valores dos tamanhos e poderes empíricos, o que é mais perceptível quando μ é central. Para esse tipo de má especificação, a variância *bootstrap* leva a um melhor desempenho do teste.

Tabela 4.12: Níveis empíricos para o teste com variância *bootstrap*, usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.

ϕ	$\mu \in [0, 0103; 0, 1092]$			$\mu \in [0, 2131; 0, 8289]$			$\mu \in [0, 9101; 0, 9886]$		
	20	55	150	20	55	150	20	55	150
$\alpha = 1\%$									
40	3,32	2,34	2,02	4,00	5,04	4,70	2,30	1,32	1,20
80	2,36	2,30	1,60	9,58	9,52	9,36	1,18	0,82	0,86
120	2,54	2,16	1,94	12,76	12,82	12,26	0,94	0,82	0,84
$\alpha = 5\%$									
40	11,14	7,54	7,36	10,58	11,14	11,24	9,50	6,36	5,56
80	9,06	7,94	6,74	18,72	18,56	18,50	5,96	4,76	4,22
120	10,96	8,84	7,34	22,34	23,02	22,76	5,72	4,58	4,08
$\alpha = 10\%$									
40	21,82	15,50	15,30	17,26	17,70	17,70	20,46	13,86	10,82
80	19,30	15,54	13,76	26,36	26,50	26,56	14,82	10,34	9,60
120	23,02	18,40	15,86	30,46	31,02	30,28	15,50	11,28	9,72

Tabela 4.13: Níveis empíricos para o teste, com variância *bootstrap*, usando os níveis nominais 1%, 5% e 10%. Má especificação: dispersão não modelada.

λ n		$\mu \in [0, 0107; 0, 1477]$					$\mu \in [0, 2515; 0, 7738]$					$\mu \in [0, 8825; 0, 9892]$				
		1	25	54	121	326	1	25	54	121	326	1	25	54	121	326
$\alpha = 1\%$																
40	1,06	32,42	46,68	58,82	60,88	1,62	16,34	24,08	42,22	43,56	1,28	31,66	42,72	49,68	51,82	
80	1,26	62,78	79,58	90,64	92,38	1,54	37,02	49,96	78,32	79,42	1,76	62,40	77,48	85,76	88,14	
120	1,64	79,78	92,66	98,10	98,56	1,70	53,94	69,32	92,54	93,64	1,42	80,44	91,10	96,78	97,34	
$\alpha = 5\%$																
40	5,32	50,72	63,74	75,82	77,84	6,04	28,62	38,50	59,10	60,92	5,12	48,80	61,00	69,18	71,68	
80	4,82	79,10	90,16	96,72	97,92	5,66	54,16	66,16	89,02	90,52	5,50	78,06	89,62	94,86	96,10	
120	5,76	91,52	97,36	99,52	99,76	5,44	70,38	82,56	97,74	97,92	5,24	90,96	96,88	99,40	99,44	
$\alpha = 10\%$																
40	10,34	60,60	72,26	82,96	84,62	11,40	37,42	47,26	68,66	69,82	10,42	58,72	69,92	78,76	80,14	
80	9,68	86,00	93,16	98,34	98,86	10,82	63,22	74,76	92,94	93,68	10,50	84,84	93,44	97,30	98,22	
120	10,70	95,38	98,66	99,84	99,86	10,40	78,24	88,30	98,96	98,88	10,76	94,42	98,30	99,82	99,72	

4.2 Modelo de Regressão Beta com Dispersão Variável

Nessa seção, apresentaremos os resultados de estudos de simulação referentes ao desempenho do teste aplicado ao modelo de regressão beta com dispersão variável. Foram consideradas y_1, y_2, \dots, y_n variáveis aleatórias independentes, em que $y_i \sim \text{Beta}(\mu_i, \phi_i)$, com $i = 1, 2, \dots, n$. Para a obtenção dos tamanhos do teste, a variável resposta foi gerada e modelada sem erro de especificação do modelo. Após a estimação dos parâmetros do modelo, aplicamos o teste baseado em influência local apresentado na Seção 3.2.2. Vale ressaltar que as covariáveis para a modelagem da média utilizadas para essa situação foram geradas a partir das distribuições: uniforme padrão, lognormal padrão e Bernoulli de parâmetro 0,1, como também, para a modelagem da dispersão, utilizamos covariáveis uniformes padrões. Novamente, consideramos três cenários para os valores de μ : próximo do zero, próximo de um e central. Em relação aos valores utilizados para os parâmetros que compõem a componente sistemática da média, para μ próximo do zero, utilizamos $\beta_1 = -1,9, \beta_2 = -1,8, \beta_3 = 1,5, \beta_4 = 2,2, \beta_5 = -4,3$. Para μ próximo do um, consideramos $\beta_1 = 1,9, \beta_2 = 1,4, \beta_3 = 1,5, \beta_4 = 2,5$ e $\beta_5 = 4,2$. Por fim, para μ central, utilizamos $\beta_1 = 1,5, \beta_2 = -2,0, \beta_3 = -2,0, \beta_4 = -16,0$ e $\beta_5 = 0,3$. Além disso, foram considerados três valores para o grau de heterogeniedade $\lambda = 44, 99$ e 163 . Para a obtenção de $\lambda = 44$, utilizamos na componente sistemática da dispersão os parâmetros $\gamma_1 = 4,0, \gamma_2 = -1,1, \gamma_3 = -2,2, \gamma_4 = 1,2$ e $\gamma_5 = 2,9$. Para a obtenção de $\lambda = 99$, utilizamos $\gamma_1 = 4,0, \gamma_2 = -1,1, \gamma_3 = -2,2, \gamma_4 = 2,0$ e $\gamma_5 = 3,5$. Finalmente, para a obtenção de $\lambda = 163$, utilizamos $\gamma_1 = 4,0, \gamma_2 = -1,1, \gamma_3 = -2,3, \gamma_4 = 2,0$ e $\gamma_5 = 4,1$. A Tabela 4.14 apresenta os percentuais de rejeição sob hipótese nula (tamanhos do teste) de que o modelo está corretamente especificado.

Tabela 4.14: Níveis empíricos para o tamanho do teste, usando os níveis nominais 1%, 5% e 10%.

		$\mu \in [0, 0005; 0, 1337]$			$\mu \in [0, 2083; 0, 8314]$			$\mu \in [0, 8753; 0, 9994]$		
n	λ	44	99	163	44	99	163	44	99	163
	$\alpha = 1\%$									
40		14,54	13,16	14,04	9,92	11,40	11,78	12,90	11,94	11,80
80		6,80	6,42	5,68	2,90	3,96	4,12	6,36	6,42	5,12
120		4,14	3,64	3,34	1,76	1,76	2,28	3,82	3,42	3,36
$\alpha = 5\%$										
40		27,82	25,92	26,10	20,84	22,68	24,76	26,04	25,02	23,92
80		16,14	15,20	14,28	9,46	10,04	11,30	16,58	14,84	13,72
120		10,74	10,28	9,34	6,20	6,42	6,86	10,94	9,94	9,36
$\alpha = 10\%$										
40		36,96	35,68	34,90	29,00	31,66	33,98	35,36	34,26	32,88
80		24,56	22,60	20,96	15,78	16,42	17,96	24,46	22,24	21,32
120		16,66	15,96	14,88	11,50	10,98	11,24	16,92	16,26	15,72

Notamos que na grande maioria dos cenários, os tamanhos não estão próximos dos níveis nominais. Mais especificamente, no cenário em que μ é central apenas na situação $n = 120$,

o tamanho do teste possui um desempenho aceitável, no entanto, os resultados para $n = 40, 80$ estão muito longe do esperado, isto é, estão distantes dos valores nominais fixados 1%, 5% e 10%. Também verificamos que nas situações em que μ está nas extremidades, há uma leve melhora quando λ aumenta mas ainda não o suficiente para que seja aceitável. Como visto no Capítulo 3, nesse caso, a distribuição assintótica da estatística de teste é normal padrão, no entanto, pelos resultados obtidos, novamente há indícios de que a convergência em distribuição, para essa classe de modelos parece não ser adequada. Nas Figuras 4.11 a 4.19 temos os histogramas de U_n com as curvas teóricas da distribuição normal padrão e os respectivos *qqnorm*'s (gráfico quantil-quantil para a distribuição normal) de U_n com uma reta pontilhada (comparação quantil U_n e quantil normal padrão) e outra reta (comparação com uma normal de parâmetros baseados na média e erro-padrão amostral de U_n , respectivamente.), considerando todos os cenários de n , μ e λ obtidos das simulações. Analisando os histogramas, percebemos que em todos os cenários as distribuições de U_n são, fortemente assimétricas e mais leptocúrticas do que a curva da distribuição normal padrão. Em relação à comparação dos quantis de U_n com os quantis normais, vemos novamente que, pelos *qqnorm*'s que a distribuição de U_n não se aproxima nem da distribuição normal padrão, nem de uma distribuição normal de parâmetros média e erro-padrão amostral, pois nas caudas os quantis da U_n ficam muito distantes dos quantis da normal. Além disso, percebemos que o aumento do tamanho amostral não leva a melhora significativa no sentido de se aproximar da distribuição normal padrão (mesmo comportamento visto na Seção 4.1). Também, o aumento de λ faz com que a distribuição de U_n seja mais leptocúrtica. Novamente, verificamos que a diferença entre a distribuição empírica de U_n e a distribuição normal padrão é maior quando consideramos os extremos de μ . Calculamos a média, variância, assimetria e curtose dos 5000 valores da estatística U_n em uma das situações em que a distribuição de U_n mais se aproximou da distribuição normal e onde obtivemos melhores tamanhos dos testes, isto é, consideramos μ central e $\lambda = 163$. Na Tabela 4.15 estão apresentadas essas medidas. Notamos novamente que a variância está sendo subestimada quando o tamanho amostral aumenta. A média de U_n está muito distante do valor esperado (zero), e o aumento de n leva a uma aproximação do valor zero, mas muito lentamente. A assimetria é negativa e se distancia do zero de acordo com o aumento de n , o que não é esperado. Por fim, apenas para $n = 40$, a curtose está próxima do esperado que é 3, pois o aumento do n distancia a curtose de U_n quando comparada com a distribuição normal padrão, o que também não era esperado. Analogamente a seção anterior, avaliamos a distribuição da estatística U_n para tamanhos amostrais maiores tais como $n = 500, 1000, 1500$ e 2000 , considerando 1000 réplicas. Novamente, construímos os gráficos e calculamos as medidas, ambos já citados. Na Tabela 4.16 estão apresentadas a média, variância, assimetria e curtose para as 1000 réplicas de U_n , como também, na Figura 4.20 temos os histogramas de U_n e os respectivos *qqnorm*'s. Notamos que, em relação as medidas, a média da estatística U_n está consideravelmente mais próxima de zero, do que no cenário onde n era menor. O aumento do tamanho amostral também causa ainda uma maior subestimação do valor da variância de U_n ,

sendo esse novamente o problema mais grave que percebemos. De acordo com o aumento do tamanho amostral, a assimetria, nessa situação, não se aproxima do zero mesmo que para $n = 2000$. Por fim, a curtose se está próxima de 3. Através dos gráficos, vemos uma significativa melhoria da distribuição de U_n em relação a assimetria. Notamos que a distribuição amostral de U_n , para esses tamanhos amostrais, não se aproxima de uma distribuição normal padrão. Na verdade, a estatística aparente possuir uma distribuição assimétrica negativa, com média não muito distante de zero, quando consideramos grandes amostras e variância em torno de 0,6 que é uma variância diferente do esperado, que seria igual a um. Concluímos através das distribuições de U_n , desses cenários, que para o modelo de regressão beta com dispersão variável, a estatística U_n não possui distribuição normal padrão.

Tabela 4.15: Valores da média, variância, assimetria e curtose para as 5000 réplicas de U_n para o cenário μ central e $\lambda = 163$.

Tamanho amostral	Média	Variância	Assimetria	Curtose
40	-1,3049	1,1455	-0,5950	3,6065
80	-0,9059	0,7124	-0,7452	3,7214
120	-0,7135	0,6147	-0,7880	4,2791

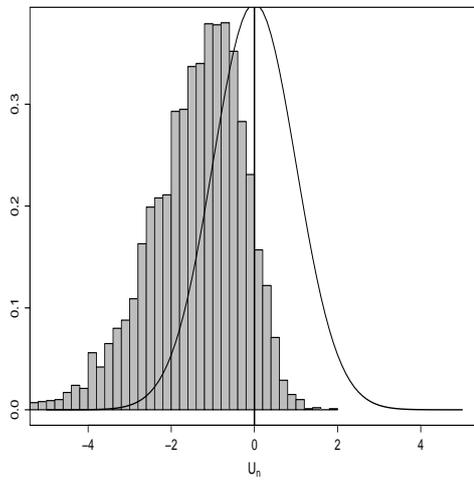
Tabela 4.16: Valores da média, variância, assimetria e curtose para as 1000 réplicas de U_n para o cenário μ central e $\lambda = 163$.

Tamanho amostral	Média	Variância	Assimetria	Curtose
500	-0,5834	0,7091	-0,6089	3,3764
1000	-0,3907	0,6485	-0,6398	3,2248
1500	-0,2977	0,5625	-0,5615	3,3552
2000	-0,2613	0,5768	-0,5508	3,1157

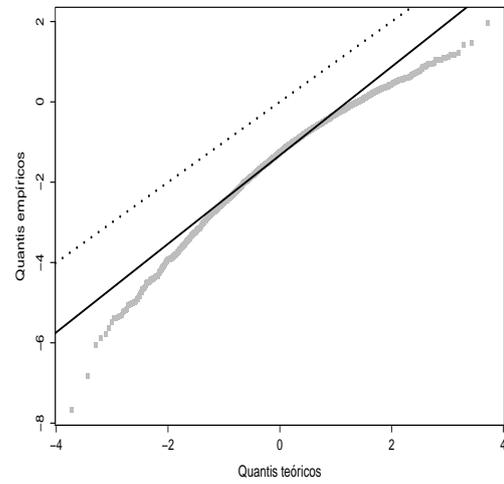
Portanto, novamente realizamos um melhoramento do teste no sentido de mudar a distribuição da estatística de teste que é utilizada para obter os quantis. Empregamos novamente a metodologia *bootstrap*. Considerando os mesmos cenários apresentados na Tabela 4.14, obtivemos os tamanhos do teste, agora, através dos quantis empíricos. Além disso, avaliamos os poderes dos testes através da introdução de algumas formas de erro de especificação do modelo. Como na seção anterior, consideramos a omissão de covariáveis do modelo especificado para a média. Na Tabelas 4.17, 4.18 e 4.19 estão apresentados os tamanhos do teste e também os percentuais de rejeição sob a hipótese alternativa, ou seja, os poderes do teste sob a omissão de covariáveis.

A partir dos resultados apresentados nessas tabelas, chegamos a algumas conclusões. Em relação ao tamanho do teste, percebemos que o tamanho do teste foi controlado (próximo dos níveis nominais) quando utilizamos os quantis empíricos obtidos via *bootstrap*.

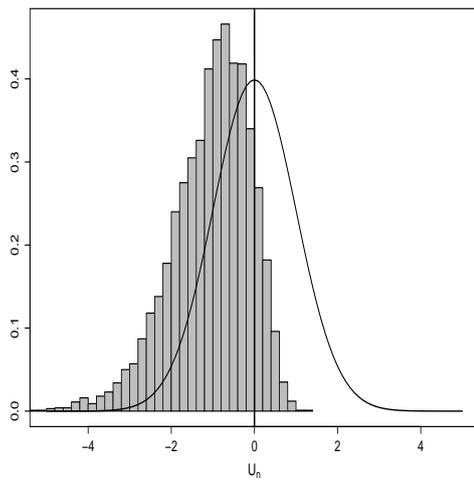
Figura 4.11: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0,0005; 0,1337]$ e $\lambda = 44$.



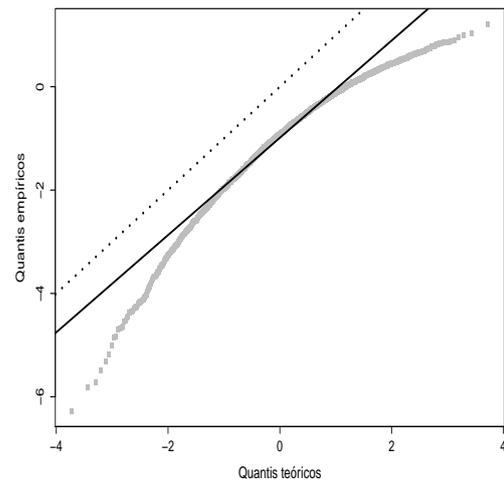
(a) $n = 40$



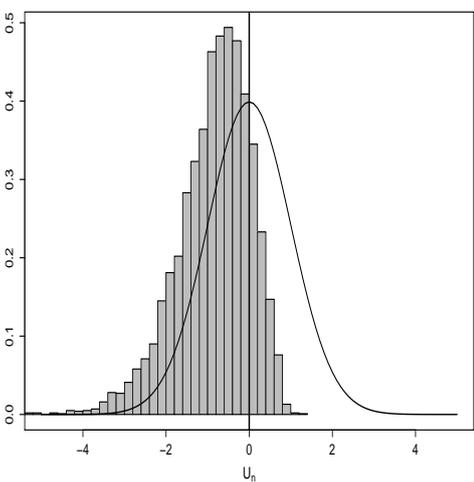
(b) $n = 40$



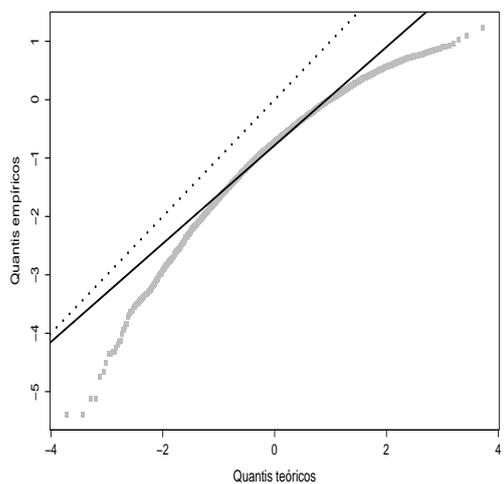
(c) $n = 80$



(d) $n = 80$

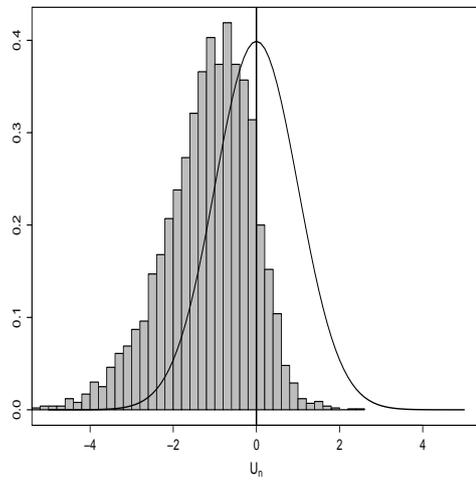


(e) $n = 120$

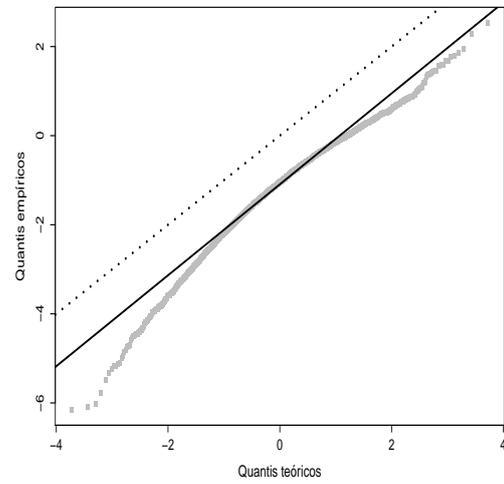


(f) $n = 120$

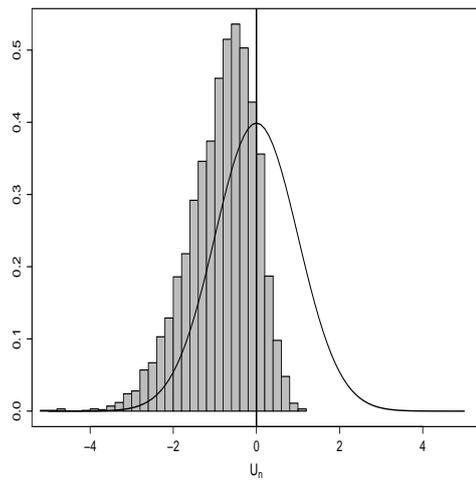
Figura 4.12: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 44$.



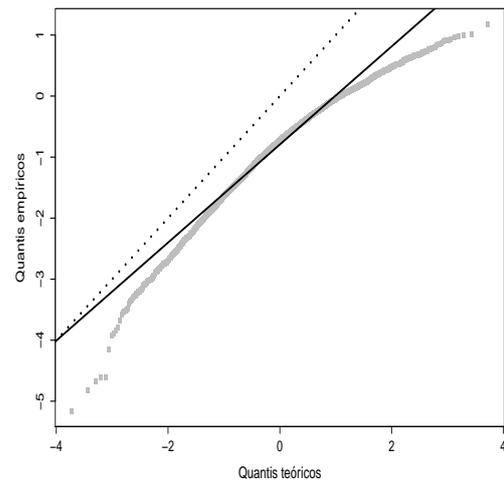
(a) $n = 40$



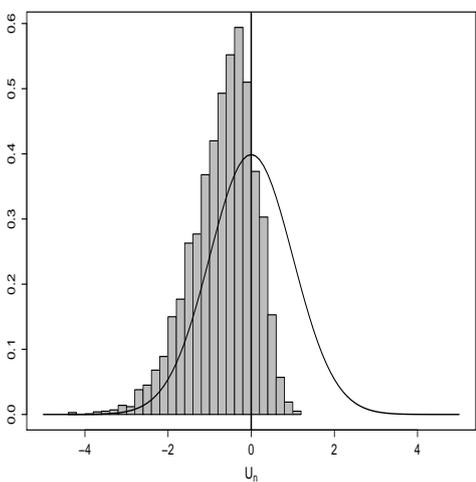
(b) $n = 40$



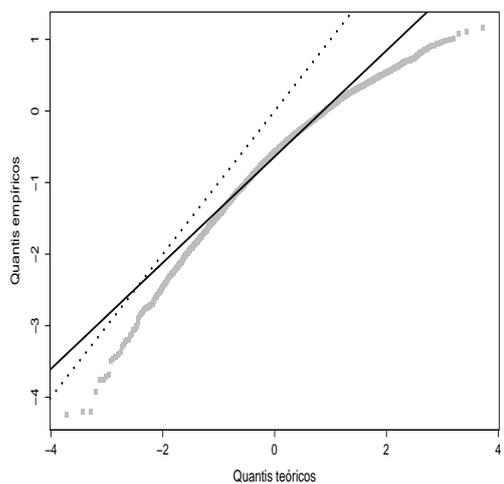
(c) $n = 80$



(d) $n = 80$

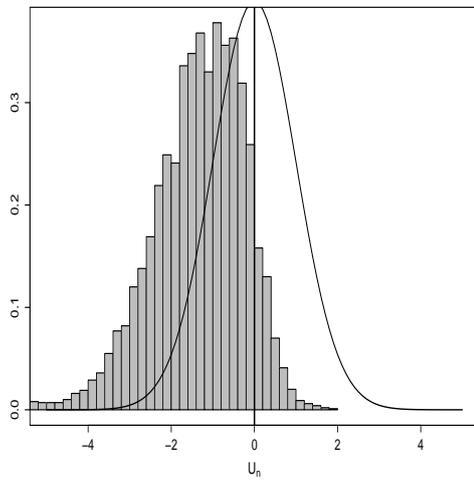


(e) $n = 120$

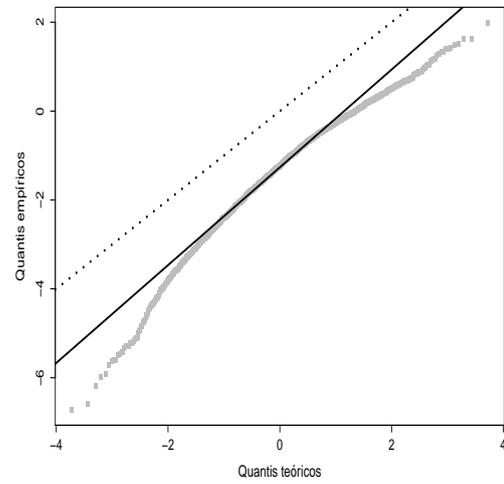


(f) $n = 120$

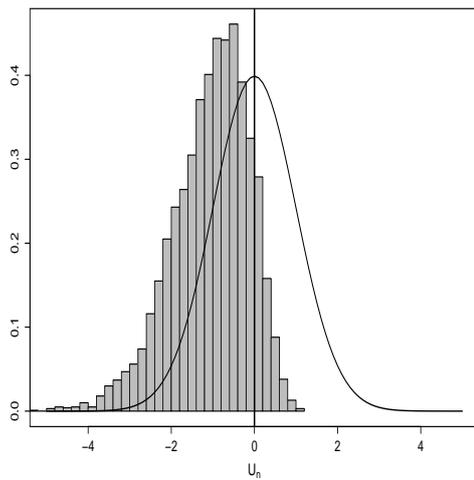
Figura 4.13: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 44$ para n .



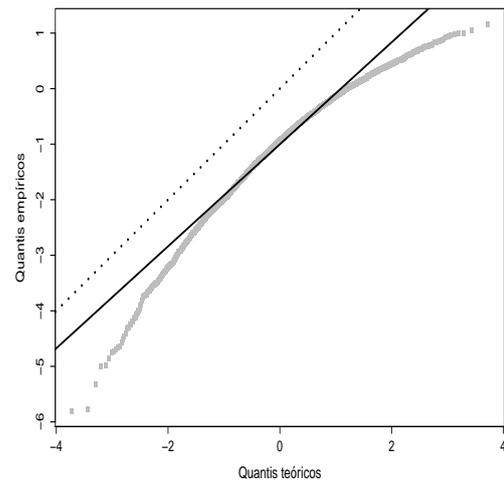
(a) $n = 40$



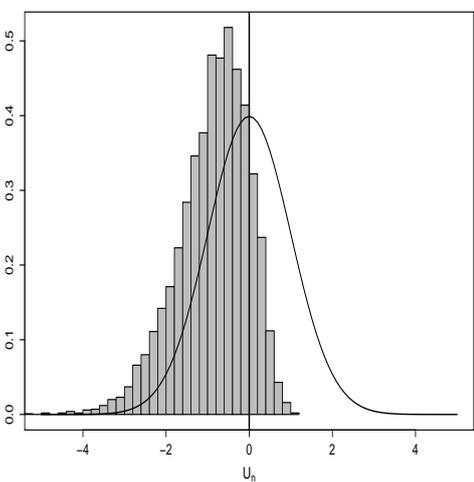
(b) $n = 40$



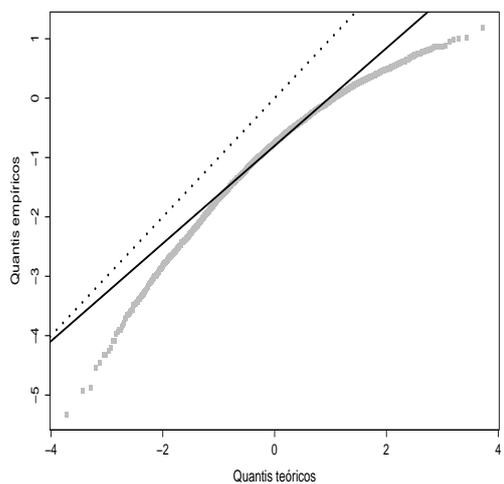
(c) $n = 80$



(d) $n = 80$

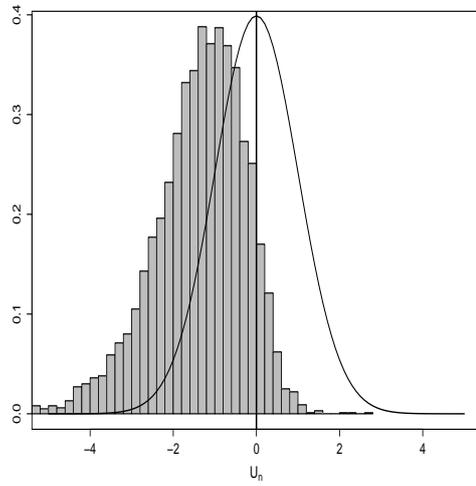


(e) $n = 120$

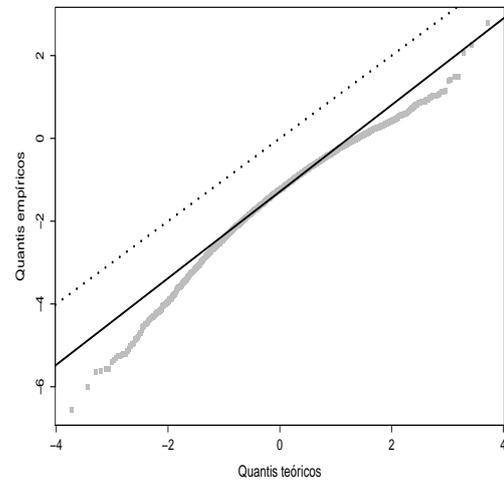


(f) $n = 120$

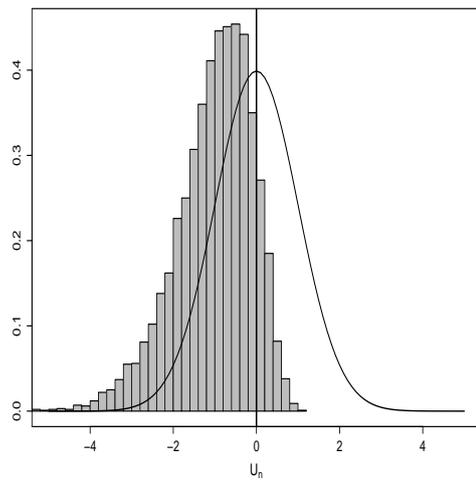
Figura 4.14: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0,0005; 0,1337]$ e $\lambda = 99$.



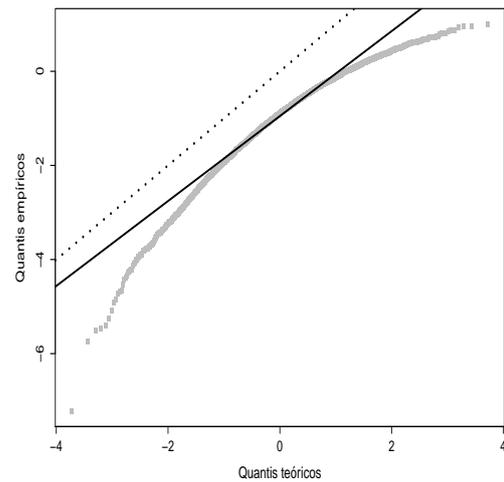
(a) $n = 40$



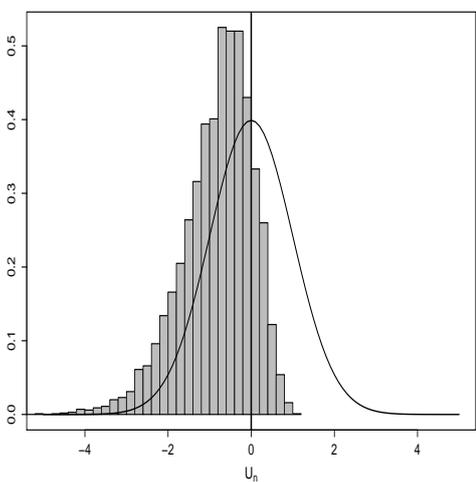
(b) $n = 40$



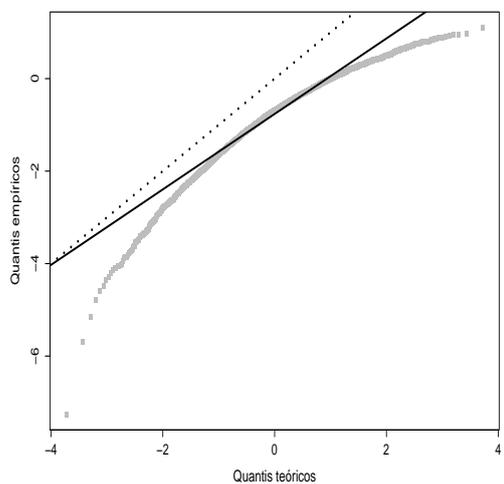
(c) $n = 80$



(d) $n = 80$

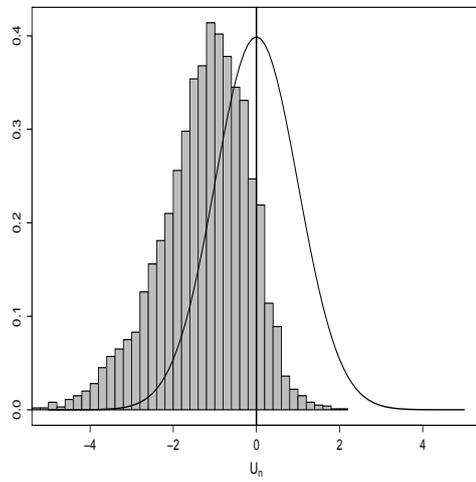


(e) $n = 120$

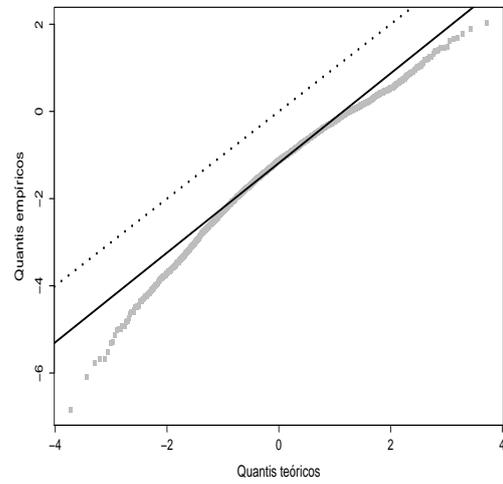


(f) $n = 120$

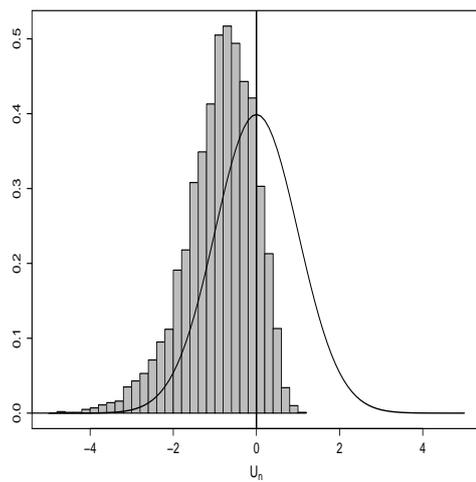
Figura 4.15: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 99$.



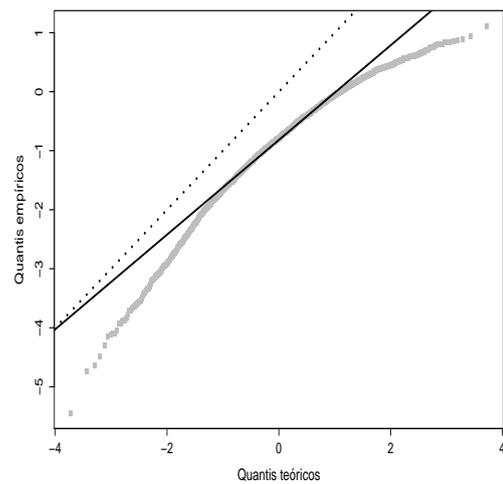
(a) $n = 40$



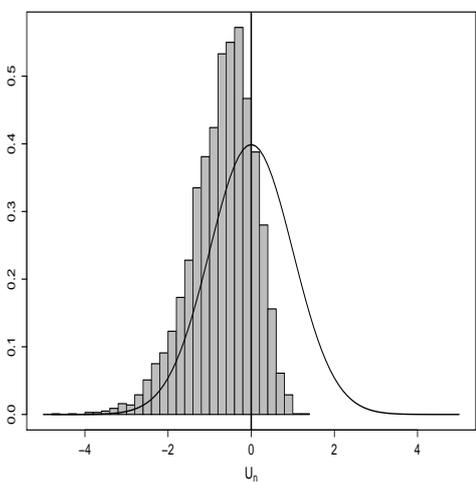
(b) $n = 40$



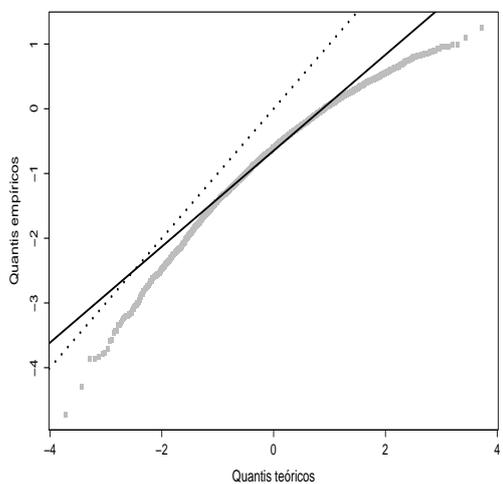
(c) $n = 80$



(d) $n = 80$

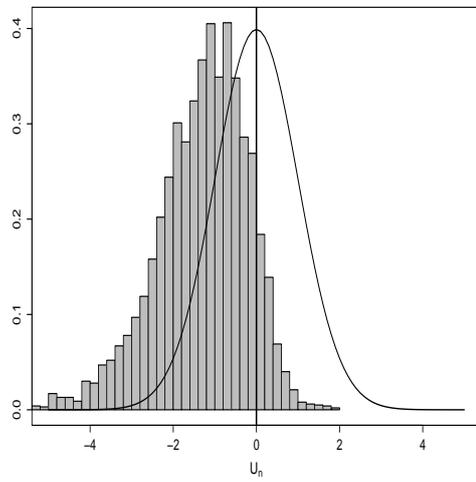


(e) $n = 120$

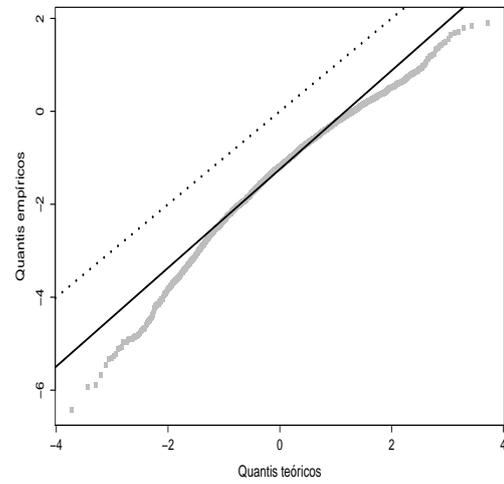


(f) $n = 120$

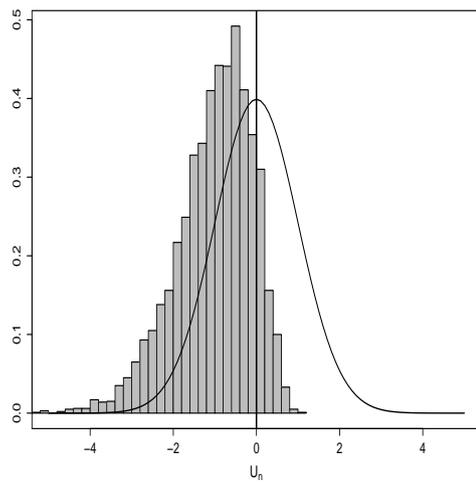
Figura 4.16: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 99$.



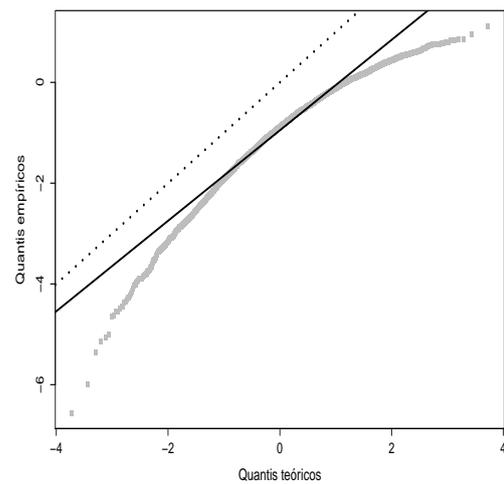
(a) $n = 40$



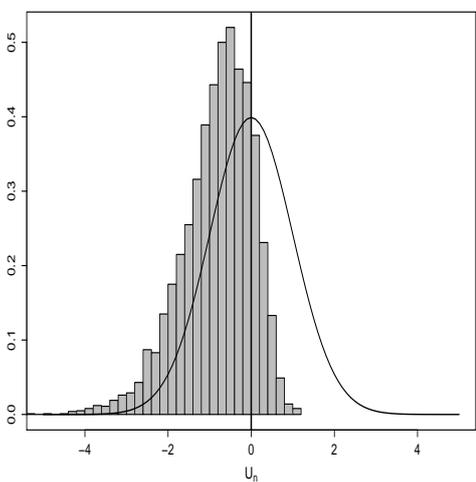
(b) $n = 40$



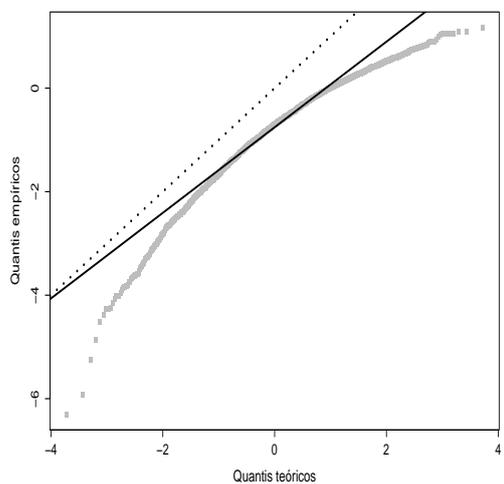
(c) $n = 80$



(d) $n = 80$

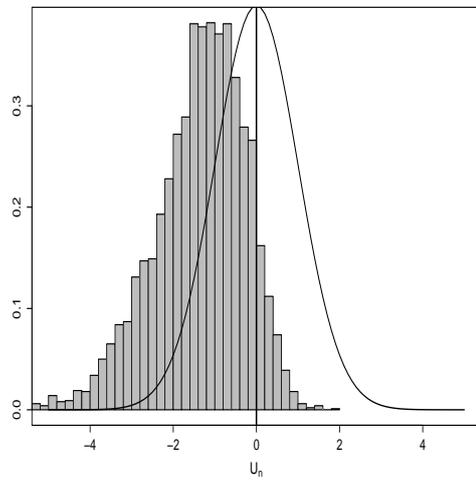


(e) $n = 120$

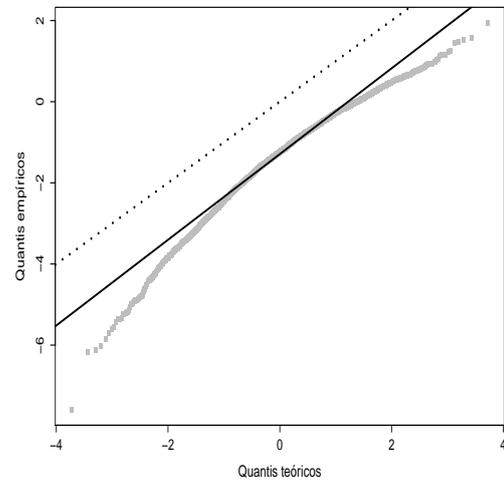


(f) $n = 120$

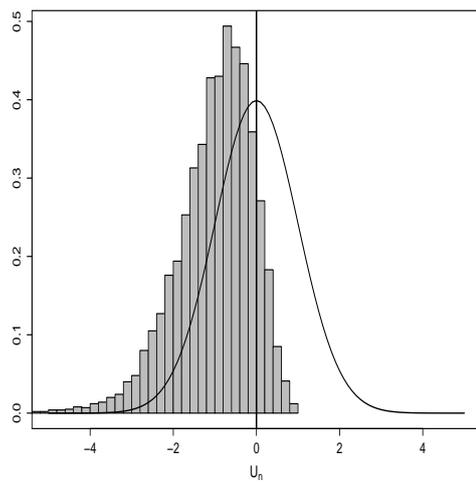
Figura 4.17: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0,0005; 0,1337]$ e $\lambda = 163$.



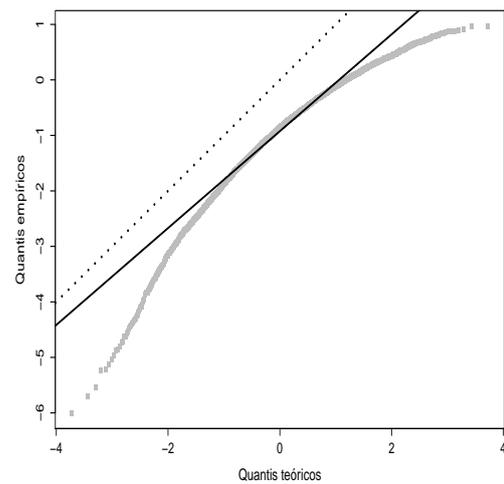
(a) $n = 40$



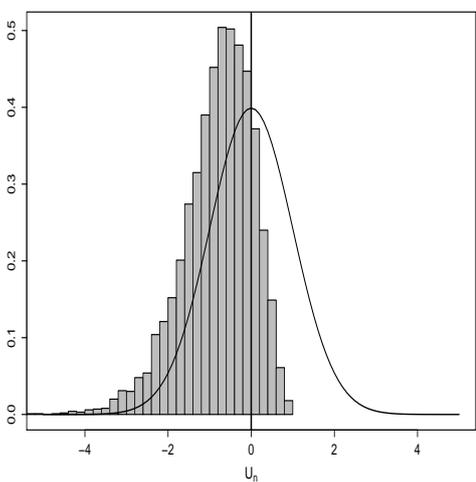
(b) $n = 40$



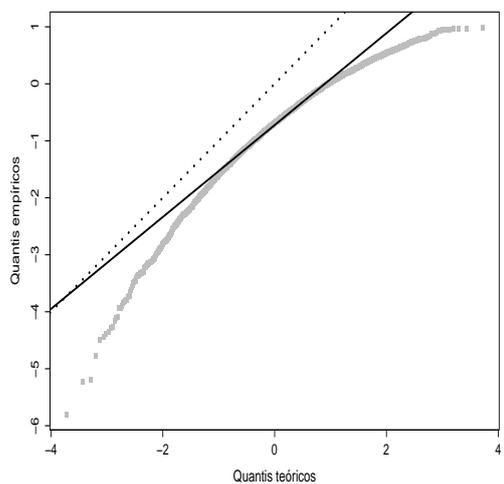
(c) $n = 80$



(d) $n = 80$

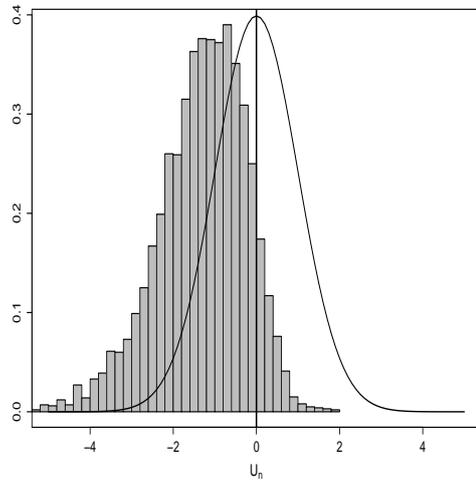


(e) $n = 120$

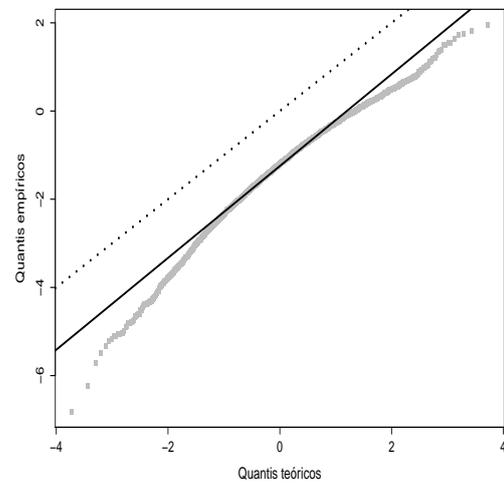


(f) $n = 120$

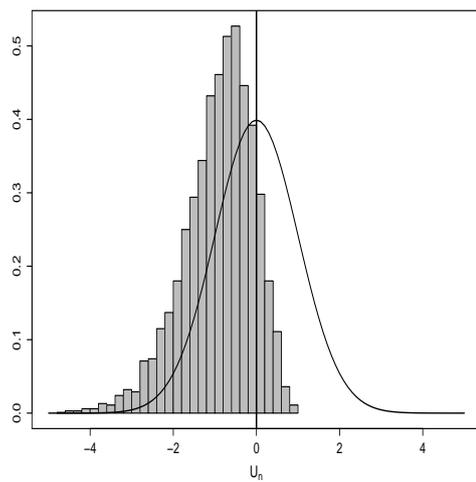
Figura 4.18: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 2083; 0, 8314]$ e $\lambda = 163$.



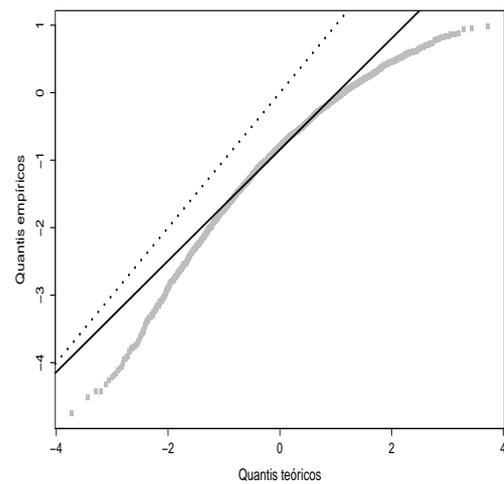
(a) $n = 40$



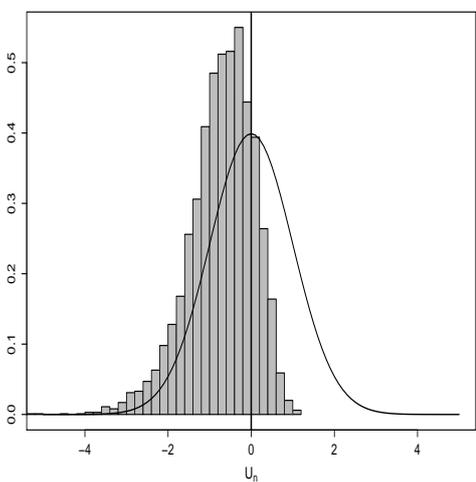
(b) $n = 40$



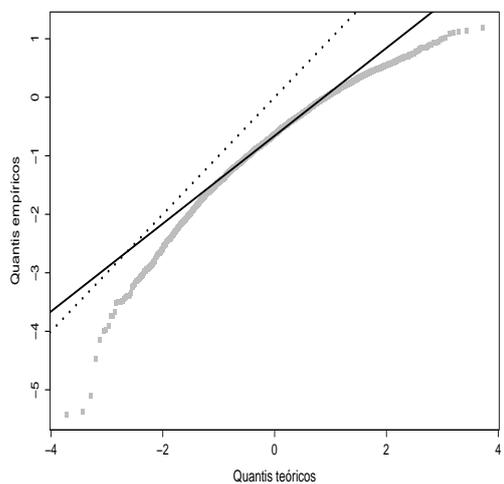
(c) $n = 80$



(d) $n = 80$

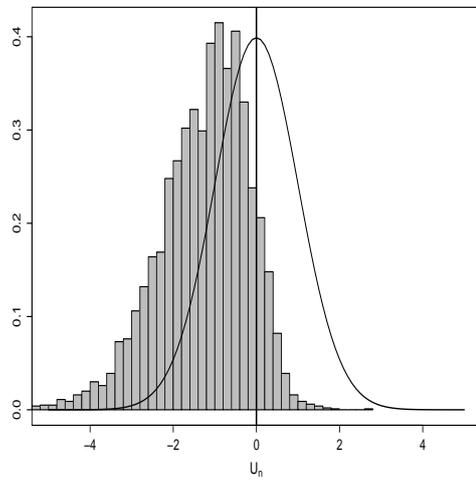


(e) $n = 120$

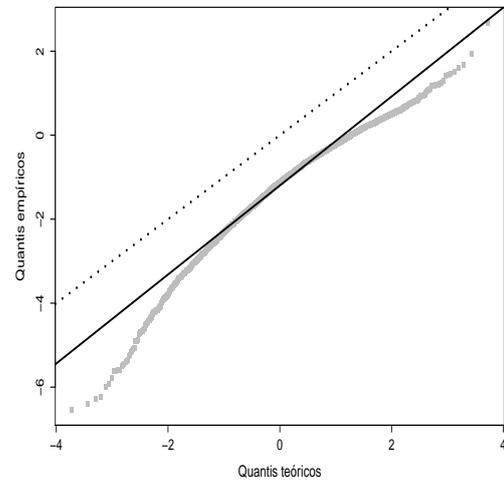


(f) $n = 120$

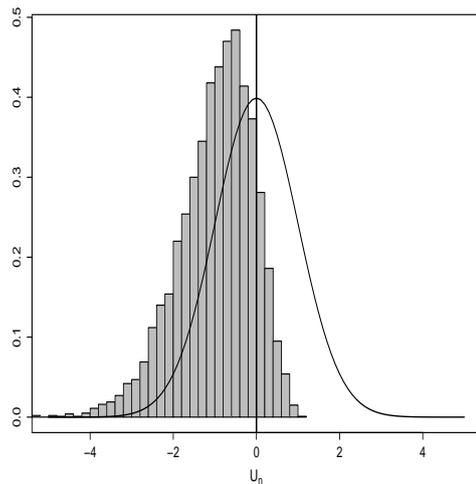
Figura 4.19: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 8753; 0, 9994]$ e $\lambda = 163$.



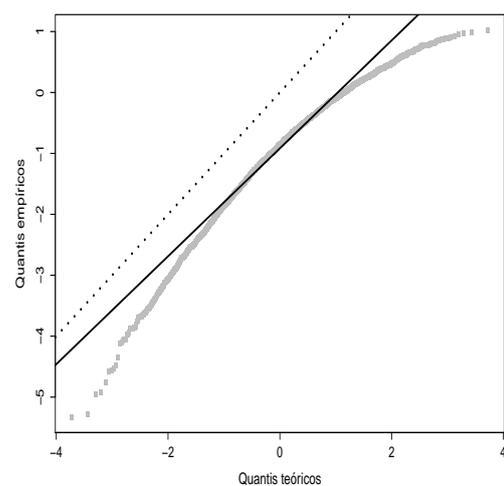
(a) $n = 40$



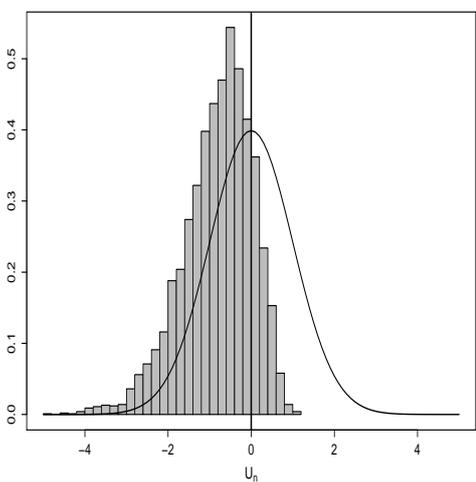
(b) $n = 40$



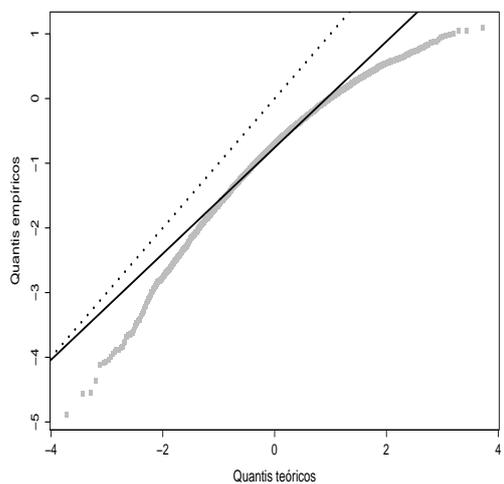
(c) $n = 80$



(d) $n = 80$

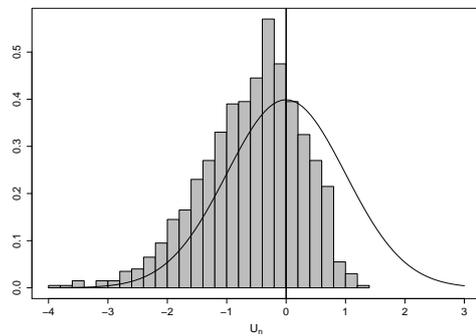


(e) $n = 120$

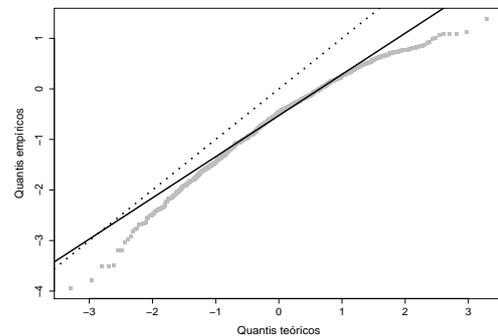


(f) $n = 120$

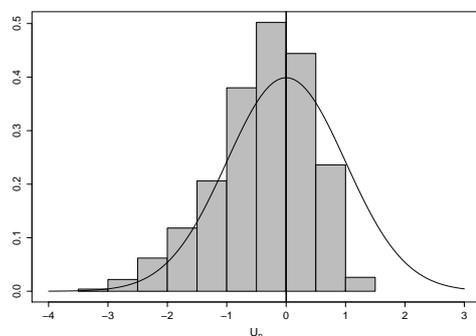
Figura 4.20: Histograma com a densidade $N(0,1)$ e *qqplot* de U_n , respectivamente, com $\mu \in [0, 1176; 0, 8158]$ e $\lambda = 163$.



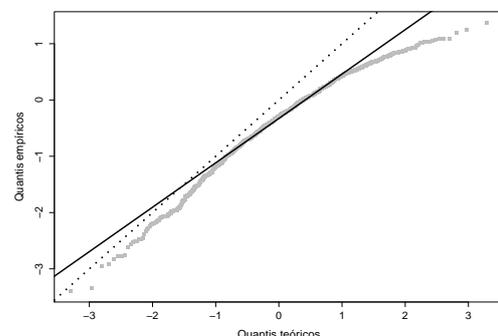
(a) $n = 500$



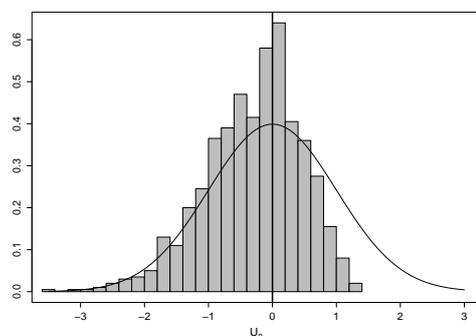
(b) $n = 500$



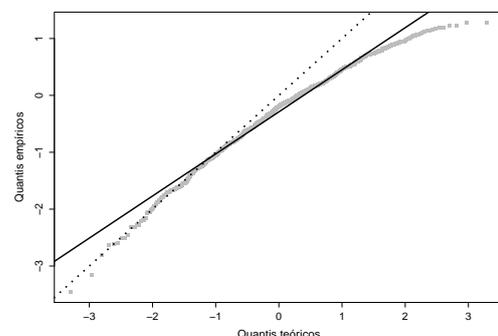
(c) $n = 1000$



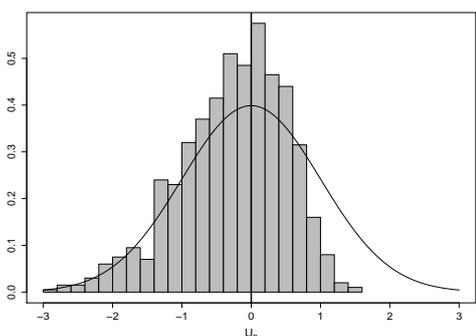
(d) $n = 1000$



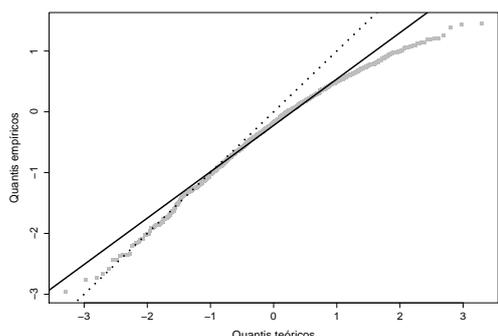
(e) $n = 1500$



(f) $n = 1500$



(g) $n = 2000$



(h) $n = 2000$

Fonte: A autora.

Tabela 4.17: Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0005; 0, 1337]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	61,70	61,64	65,86	48,72	47,04	51,40	40,20	38,48	42,12	1,16	1,40	1,62
80	99,08	99,34	99,32	96,36	95,78	96,60	93,00	92,38	93,34	1,22	1,88	1,32
120	100,00	99,90	100,00	99,76	99,58	99,68	99,18	98,84	99,24	1,42	1,26	1,28
$\mu \in [0, 2083; 0, 8314]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	3,10	4,76	5,76	2,64	3,14	4,46	0,92	1,04	0,82	1,36	1,60	1,42
80	27,48	25,76	24,56	19,74	19,28	19,16	2,12	1,94	1,96	1,50	1,26	1,32
120	33,06	30,28	31,28	23,50	22,06	23,34	2,34	1,78	2,12	1,36	1,36	1,60
$\mu \in [0, 8753; 0, 9994]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	41,62	40,84	43,44	25,14	23,12	24,70	18,80	17,52	18,20	1,40	1,10	1,46
80	98,70	97,92	98,02	90,90	87,18	87,80	81,52	76,36	75,86	1,42	1,34	1,50
120	99,92	99,94	99,86	98,26	97,34	96,92	94,46	91,66	90,38	1,42	1,60	1,28

Tabela 4.18: Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0005; 0, 1337]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	91,66	91,20	92,28	81,42	80,60	81,70	73,24	73,00	73,70	5,30	5,56	5,52
80	99,92	99,96	99,98	99,02	99,04	99,08	97,46	97,24	97,66	5,28	5,96	5,58
120	100,00	99,98	100,00	99,98	99,98	99,98	99,80	99,68	99,78	5,12	5,02	5,38
$\mu \in [0, 2083; 0, 8314]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	20,46	22,04	23,44	16,14	17,58	19,42	3,92	4,90	5,20	5,12	5,68	5,40
80	42,98	42,36	40,92	34,34	34,04	33,72	7,26	6,24	7,06	4,96	5,14	5,20
120	50,68	48,68	49,46	39,40	38,92	40,70	7,84	7,30	7,24	5,12	5,48	5,58
$\mu \in [0, 8753; 0, 9994]$												
$\lambda \backslash n$	44	99	163	44	99	163	44	99	163	44	99	163
40	85,48	84,58	84,64	67,96	64,76	63,90	57,64	53,74	53,86	5,52	5,74	5,26
80	99,82	99,66	99,78	97,42	96,14	96,28	93,14	90,98	91,12	5,16	5,36	5,40
120	99,98	100,00	100,00	99,62	99,30	98,98	98,50	97,68	97,08	5,42	5,46	4,60

Em relação ao poder do teste, vimos que quanto maior o nível nominal fixado, maior o poder do teste obtido. As conclusões a seguir são referentes aos três níveis nominais. Como esperado, à medida que é omitido um número maior de covariáveis, maior é o poder obtido. O cenário em que μ pertence ao extremo inferior possui poder de teste maior do que quando μ pertence ao extremo superior, o que é mais evidenciado quando $n = 40$. Novamente, o cenário em que μ é central possui poderes mais baixos que os outros dois cenários. Por fim, vimos que quando λ aumenta, não é possível caracterizar um comportamento estável tanto em relação ao tamanho quanto em relação ao poder. Novamente, introduzimos a má especificação de gerar a variável resposta através da distribuição simplex e, em seguida, modelar utilizando a regressão beta com dispersão variável. A Tabela 4.20 apresenta as taxas de rejeição dos testes para o erro de especificação, isto é, os poderes

Tabela 4.19: Níveis empíricos para o teste, usando nível nominal 10%. Má especificação: covariáveis da média omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0,0005; 0,1337]$												
λ	44	99	163	44	99	163	44	99	163	44	99	163
n												
40	96,54	96,24	97,08	90,10	89,74	90,08	84,08	83,46	83,90	9,92	10,46	10,60
80	99,98	99,98	100,00	99,68	99,62	99,60	98,74	98,80	98,70	10,40	10,98	10,96
120	100,00	99,98	100,00	100,00	99,96	100,00	99,84	99,84	99,92	9,30	10,16	10,06
$\mu \in [0,2083; 0,8314]$												
λ	44	99	163	44	99	163	44	99	163	44	99	163
n												
40	31,42	32,58	33,90	26,48	28,10	29,54	8,84	9,48	10,72	10,70	10,14	10,66
80	51,90	51,30	50,12	43,44	42,76	43,40	13,06	12,38	12,64	10,06	10,32	10,00
120	59,26	57,12	58,50	49,04	48,24	49,94	13,78	12,40	12,98	10,14	9,90	10,54
$\mu \in [0,8753; 0,9994]$												
λ	44	99	163	44	99	163	44	99	163	44	99	163
n												
40	93,38	93,94	93,68	81,08	79,42	78,64	72,96	70,30	69,46	10,92	10,84	9,74
80	99,92	99,84	99,90	98,82	98,12	98,14	96,64	94,92	95,14	10,20	10,32	10,64
120	100,00	100,00	100,00	99,84	99,80	99,66	99,20	98,98	98,54	10,02	10,44	9,82

do teste para este tipo de má especificação.

Tabela 4.20: Níveis empíricos para o teste, usando níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.

λ	$\mu \in [0,0007; 0,1354]$			$\mu \in [0,2112; 0,7792]$			$\mu \in [0,8873; 0,9986]$		
	45	101	204	45	101	204	45	101	204
$\alpha = 1\%$									
40	15,18	36,20	39,30	26,62	38,10	31,56	12,62	28,98	28,00
80	50,36	63,12	66,52	88,20	87,28	81,88	41,60	72,72	75,90
120	61,60	75,38	78,98	98,32	96,78	94,80	56,36	87,72	90,56
$\alpha = 5\%$									
40	46,58	64,98	67,90	47,82	57,86	50,96	35,54	53,10	52,80
80	82,30	88,54	89,32	95,30	94,64	91,88	66,28	84,71	87,16
120	91,88	96,10	96,52	99,37	98,78	98,10	78,58	94,40	95,82
$\alpha = 10\%$									
40	67,92	76,94	78,54	58,92	68,84	61,94	51,08	63,26	62,94
80	91,58	94,02	94,28	97,35	97,01	95,17	77,62	89,21	90,64
120	96,44	98,25	98,40	99,66	99,32	99,01	86,77	96,21	97,14

Vale ressaltar que as covariáveis utilizadas tanto no modelo da média como da dispersão foram geradas a partir da distribuição uniforme de parâmetros 0 e 1. Em relação aos valores utilizados, para os parâmetros que compõem a componente sistemática da média, considerando o cenário em que μ é próximo do zero, foram $\beta_1 = -1,3, \beta_2 = 0,9, \beta_3 = -2,2, \beta_4 = -1,9$ e $\beta_5 = -3,3$. Para μ próximo do um, utilizamos $\beta_1 = -1,3, \beta_2 = 0,9, \beta_3 = -2,2, \beta_4 = -1,9$ e $\beta_5 = -3,3$. Por fim, para μ central consideramos $\beta_1 = 1,0, \beta_2 = -1,0, \beta_3 = 1,0, \beta_4 = -1,0$ e $\beta_5 = 1,0$. Além disso, foram considerados três valores para o grau de heterogeneidade $\lambda = 45, 101$ e 204 . Para a obtenção de $\lambda = 45$, utilizamos na componente sistemática da precisão os parâmetros $\gamma_1 = 4,0, \gamma_2 = -1,3, \gamma_3 = -2,0, \gamma_4 = 1,5$ e $\gamma_5 = -7,0$. Para a obtenção de $\lambda = 101$, utilizamos $\gamma_1 = 4,0, \gamma_2 = -1,3, \gamma_3 = -2,0, \gamma_4 = 2,0$ e $\gamma_5 = 7,0$. Finalmente, para $\lambda = 204$,

Tabela 4.21: Níveis empíricos para o teste, usando 1%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4} + \gamma_5 z_{i5}$		
	$\mu \in [0, 0008; 0, 1297]$											
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	18,24	30,38	46,52	14,40	22,72	40,74	6,04	10,88	24,12	1,38	1,34	1,60
80	56,96	69,82	85,94	48,46	58,24	80,82	19,04	30,06	61,50	1,32	1,46	1,00
120	75,90	85,64	96,40	68,82	76,96	94,76	30,10	46,84	82,46	1,38	1,26	1,46
	$\mu \in [0, 2112; 0, 7792]$											
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	4,78	11,58	20,32	3,60	5,96	14,78	2,34	3,80	9,92	0,90	0,92	1,18
80	48,48	61,02	76,90	35,84	43,40	67,26	14,86	23,66	48,06	1,26	1,42	1,38
120	66,14	80,60	91,68	52,68	61,62	85,22	22,40	35,92	68,26	1,28	1,16	1,20
	$\mu \in [0, 8719; 0, 9920]$											
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	8,18	16,66	22,82	5,92	11,58	19,16	3,06	5,38	10,16	1,10	1,00	1,18
80	55,20	67,38	82,66	43,34	52,94	76,06	18,56	29,44	57,48	1,62	1,08	1,48
120	73,24	84,86	94,28	62,64	71,64	91,58	29,82	44,90	79,26	1,54	1,34	1,02

temos $\gamma_1 = 3,9, \gamma_2 = -1,3, \gamma_3 = -2,0, \gamma_4 = 2,2$ e $\gamma_5 = 9,1$. Observamos que, à medida que n e λ aumentam, os poderes também aumentam significativamente, exceto quando $\lambda = 204$ e μ é central, que não vemos melhora do poder. Os poderes mais baixos foram obtidos para a situação em que μ está próximo de um. Como esperado, quanto maior α maior o poder do teste obtido. Também avaliamos o poder do teste em relação à omissão de covariáveis na componente sistemática da dispersão, isto é, a variável resposta é gerada, utilizando todas as covariáveis para a média e a dispersão, porém a mesma tem precisão modelada, utilizando uma, duas e três covariáveis, respectivamente. Vale ressaltar que todas as covariáveis para a modelagem da média e da dispersão utilizadas para essa situação foram geradas a partir da distribuição uniforme de parâmetros 0 e 1. Nas Tabelas 4.21, 4.22 e 4.23 estão apresentados os percentuais de rejeição sob hipótese nula (tamanhos do teste), bem como, também os percentuais de rejeição sob a hipótese alternativa, ou seja, os poderes do teste sobre a omissão de covariáveis da dispersão. Novamente, consideramos três cenários para os valores de μ : próximo do zero, próximo do um e central. Em relação aos valores utilizados para os parâmetros que compõem a componente sistemática da média, considerando o cenário em que μ é próximo do zero, temos $\beta_1 = -2,9, \beta_2 = 1,2, \beta_3 = -2,0, \beta_4 = 1,0$ e $\beta_5 = -3,2$. Para μ próximo do um, utilizamos os parâmetros $\beta_1 = 2,9, 1,3, -0,7, -1,1, 1,5$. Por fim, para μ central temos $\beta_1 = 1,0, -1,0, 1,0, -1,0, -1,0$. Além disso, foram considerados três valores para o grau de heterogeneidade $\lambda = 45, 87$ e 163 . Para a obtenção de $\lambda = 45$, utilizamos na componente sistemática para ϕ os parâmetros $\gamma_1 = 4,5, \gamma_2 = -1,0, \gamma_3 = -1,0, \gamma_4 = -1,6$ e $\gamma_5 = 5,5$.

Tabela 4.22: Níveis empíricos para o teste, usando 5%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4} + \gamma_5 z_{i5}$		
$\mu \in [0, 0008; 0, 1297]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	41,80	53,74	69,64	35,26	43,78	63,50	18,68	26,40	47,14	5,54	5,24	5,60
80	75,92	84,60	94,08	69,76	76,40	91,42	37,70	49,78	78,86	4,82	5,54	5,10
120	89,18	94,66	99,00	84,50	89,24	98,38	50,26	67,86	92,68	5,78	5,92	5,78
$\mu \in [0, 2112; 0, 7792]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	26,34	39,38	54,68	19,62	26,22	43,98	11,98	17,32	32,54	4,82	4,28	4,82
80	68,96	80,22	90,16	55,90	64,68	83,10	30,30	42,98	68,82	4,92	5,48	4,82
120	83,54	91,92	97,42	72,84	79,90	94,08	41,42	57,48	84,76	5,08	5,18	4,58
$\mu \in [0, 8719; 0, 9920]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	32,20	43,62	55,54	24,94	33,76	48,36	14,42	20,78	33,92	4,96	4,56	4,66
80	74,06	83,12	92,94	63,76	71,04	88,98	35,94	49,40	75,94	5,36	5,26	5,72
120	87,92	94,08	98,40	80,26	85,90	96,72	49,68	64,68	90,04	5,32	4,92	5,52

Tabela 4.23: Níveis empíricos para o teste, usando 10%. Má especificação: covariáveis da dispersão omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4}$			$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_3 z_{i3} + \gamma_4 z_{i4} + \gamma_5 z_{i5}$		
$\mu \in [0, 0008; 0, 1297]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	54,72	65,54	78,96	47,96	54,76	72,92	28,82	37,54	58,58	10,62	9,66	10,24
80	83,70	90,62	96,54	78,30	84,00	94,44	49,32	61,40	85,72	10,24	11,02	9,82
120	92,84	96,88	99,50	89,82	93,10	99,08	61,36	77,28	95,48	10,28	11,42	10,68
$\mu \in [0, 2112; 0, 7792]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	42,68	55,84	69,20	33,36	39,90	58,44	22,16	28,38	45,46	9,20	8,76	9,32
80	78,82	87,08	94,34	66,62	74,02	88,74	40,50	53,90	77,62	9,46	10,46	9,60
120	89,56	95,54	98,60	81,52	86,62	96,64	53,06	68,06	90,60	9,92	10,20	9,48
$\mu \in [0, 8719; 0, 9920]$												
$\lambda \backslash n$	45	87	163	45	87	163	45	87	163	45	87	163
40	47,20	58,96	69,46	38,38	45,74	62,44	24,52	32,10	48,34	9,70	9,52	8,78
80	81,80	89,40	95,90	73,06	79,20	92,72	46,70	60,42	83,28	9,86	9,72	10,92
120	92,18	96,86	99,16	86,68	91,54	98,30	59,78	74,36	93,82	10,84	9,88	9,86

Para $\lambda = 87$, utilizamos $\gamma_1 = 4,5$, $\gamma_2 = -1,0$, $\gamma_3 = -1,9$, $\gamma_4 = -1,5$ e $\gamma_5 = 6,5$. Finalmente, para a obtenção de $\lambda = 163$, utilizamos $\gamma_1 = 4,5$, $\gamma_2 = -1,0$, $\gamma_3 = -1,9$, $\gamma_4 = -1,5$ e $\gamma_5 = 8,5$. Notamos que, o tamanho empírico do teste (Cenário 4) está próximo dos níveis nominais fixados para todos os cenários, como esperado. Em relação ao poder do teste, como esperado, à medida que são omitidas um número maior de covariáveis da dispersão, maior é o poder obtido. O cenário em que μ pertence ao extremo inferior possui poder de teste equivalente a μ pertence ao extremo superior (leve diferença quando $n = 40$). No cenário em que μ é central, os poderes do teste são um pouco mais baixos do que os dos outros dois cenários. Portanto, novamente notamos que utilizando a metodologia *bootstrap* para obter os quantis da estatística de teste U_n obtivemos ótimos tamanhos de teste para todos os cenários apresentados. Além disso, em relação ao poder do teste, vimos que

o teste consegue detectar na maioria das situações com bons poderes as má especificações introduzidas: omissão de covariáveis da média e da dispersão e variável resposta simplex.

Capítulo 5

Teste para a Precisão ϕ

Nesse capítulo, apresentamos um novo teste de hipóteses para o modelo de regressão beta com dispersão fixa. Esse teste também é baseado em influência local, porém considerando perturbação no parâmetro de precisão, e tem como objetivo detectar a necessidade da modelagem da dispersão em situações onde foi considerada dispersão fixa. Primeiramente, apresentamos o cálculo da curvatura normal para esse esquema de perturbação, em seguida, introduziremos o novo teste.

5.1 Perturbação no Parâmetro de Precisão ϕ

Nessa seção, apresentaremos o cálculo da curvatura normal para o esquema de perturbação no parâmetro de precisão ϕ . Para tanto, modificaremos o logaritmo da função de verossimilhança definido em (2.9), reescrevendo a expressão da precisão de forma que ϕ não seja constante, isto é, para $i = 1, 2, \dots, n$, temos

$$\phi_i = \frac{\phi}{\omega_i}, \quad (5.1)$$

em que $\boldsymbol{\omega} = (\omega_1, \dots, \omega_s)^\top$ é o vetor de perturbação definido na Subseção 1.2.1. Dessa forma, utilizando a expressão do parâmetro da precisão definida em (5.1), temos que o logaritmo da função de verossimilhança perturbado, neste caso, fica dado por

$$\ell(\boldsymbol{\beta}, \phi|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\mu_i, \phi|\boldsymbol{\omega}), \quad (5.2)$$

em que

$$\begin{aligned} \ell_i(\mu_i, \phi|\boldsymbol{\omega}) &= \log \Gamma\left(\frac{\phi}{\omega_i}\right) - \log \Gamma\left(\frac{\mu_i \phi}{\omega_i}\right) - \log \Gamma\left(\frac{(1 - \mu_i) \phi}{\omega_i}\right) + \left(\frac{\mu_i \phi}{\omega_i} - 1\right) \log(y_i) \\ &+ \left[\frac{(1 - \mu_i) \phi}{\omega_i} - 1\right] \log(1 - y_i). \end{aligned}$$

Precisamos obter a matriz de perturbação $\Delta = (\Delta_\beta \quad \Delta_\phi)^\top$ em que

$$\Delta_\beta = \frac{\partial \ell(\mu, \phi | \boldsymbol{\omega})}{\partial \boldsymbol{\omega} \boldsymbol{\beta}^\top} \quad \text{e} \quad \Delta_\phi = \frac{\partial \ell(\mu, \phi | \boldsymbol{\omega})}{\partial \boldsymbol{\omega} \phi}.$$

Derivando em relação a ω_i temos

$$\frac{\partial \ell_i(\mu_i, \phi | \boldsymbol{\omega})}{\partial \omega_i} = -\frac{\phi}{\omega_i^2} \left[\psi \left(\frac{\phi}{\omega_i} \right) - \mu_i \psi \left(\frac{\mu_i \phi}{\omega_i} \right) - (1 - \mu_i) \psi \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) + \mu_i y_i^* + \log(1 - y_i) \right].$$

Derivando agora em relação a β_r para $r = 1, 2, \dots, n$ temos

$$\begin{aligned} \frac{\partial^2 \ell_i(\mu_i, \phi | \boldsymbol{\omega})}{\partial \omega_i \partial \beta_r} &= -\frac{\phi x_{ir} t_i}{\omega_i^2} \left[y_i^* - \psi \left(\frac{\mu_i \phi}{\omega_i} \right) + \psi \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) - \frac{\mu_i \phi}{\omega_i} \psi' \left(\frac{\mu_i \phi}{\omega_i} \right) \right. \\ &\quad \left. + \frac{(1 - \mu_i) \phi}{\omega_i} \psi' \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) \right] \\ &= \frac{\phi x_{ir} t_i}{\omega_i^2} f_i(\omega_i), \end{aligned}$$

em que $f_i(\omega_i) = c_i(\omega_i) - (y_i^* - \mu_i(\omega_i))$ e

$$\mu_i^*(\omega_i) = \psi \left(\frac{\mu_i \phi}{\omega_i} \right) - \psi \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right), \quad (5.3)$$

e

$$c_i(\omega_i) = \frac{\phi}{\omega_i} \left[\mu_i \psi' \left(\frac{\mu_i \phi}{\omega_i} \right) - (1 - \mu_i) \psi' \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) \right].$$

Matricialmente, temos

$$\Delta_\beta = \phi X^\top T \mathcal{W}^{-2} F(\boldsymbol{\omega}), \quad (5.4)$$

em que $\mathcal{W}^{-2} = \text{diag}\{1/\omega_1^2, 1/\omega_2^2, \dots, 1/\omega_n^2\}$ e $F(\boldsymbol{\omega}) = \text{diag}\{f_1(\omega_1), f_2(\omega_2), \dots, f_n(\omega_n)\}$.

Derivando em relação a ϕ , obtemos

$$\frac{\partial^2 \ell_i(\mu_i, \phi | \boldsymbol{\omega})}{\partial \omega_i \partial \phi} = \frac{1}{\omega_i^2} [\phi d_i(\omega_i) - u_i(\omega_i)],$$

em que

$$u_i(\omega_i) = \psi \left(\frac{\phi}{\omega_i} \right) + \mu_i (y_i^* - \mu_i^*(\omega_i)) - \psi \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) + \log(1 - y_i)$$

e

$$d_i(\omega_i) = \frac{1}{\omega_i} \left[\mu_i^2 \psi' \left(\frac{\mu_i \phi}{\omega_i} \right) + (1 - \mu_i)^2 \psi' \left(\frac{(1 - \mu_i) \phi}{\omega_i} \right) - \psi' \left(\frac{\phi}{\omega_i} \right) \right].$$

Matricialmente, temos

$$\Delta_\phi = \mathcal{W}^{-2}[\phi \mathbf{d}(\boldsymbol{\omega})^\top - \mathbf{u}(\boldsymbol{\omega})^\top]. \quad (5.5)$$

Vale lembrar que a matriz de perturbação Δ avaliada em $\boldsymbol{\omega}_0 = (1, 1, \dots, 1)^\top$ e $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\phi})^\top$. Logo, avaliando (5.4) e (5.5) em $\boldsymbol{\omega}_0$ e $\hat{\boldsymbol{\theta}}$ temos $\Delta_\beta = \hat{\phi} X^\top \hat{T} \hat{F}$ e $\Delta_\phi = \hat{\phi} \hat{\mathbf{d}} - \hat{\mathbf{u}}$, respectivamente.

5.2 Teste para a Precisão ϕ

A estatística de teste que será introduzida também é baseada na quantidade $\sum_{i=1}^n C_{e_i}$. No entanto, não podemos mais utilizar $\lambda_0 = 2p$ como sendo o ponto de corte para avaliar se a magnitude de $\sum_{i=1}^n C_{e_i}$ é significativa, pois não estamos mais com o esquema de perturbação ponderação de casos. Assim, para introduzirmos o novo teste, precisamos de um novo ponto de corte. Para tanto, consideremos novamente a matriz $A(\boldsymbol{\theta})$ definida em (3.1). Porém, precisaremos de uma nova matriz $B(\boldsymbol{\theta})$, uma vez que a matriz $B(\boldsymbol{\theta})$ definida em (3.1) só vale para o esquema de ponderação de casos. Portanto, definiremos uma nova expressão para essa matriz baseada no fato descrito a seguir.

Quando consideramos o esquema de perturbação ponderação de casos é válido que

$$\begin{aligned} (\Delta_i(\boldsymbol{\theta}))_s &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_s \partial \omega_i} \\ &= \frac{\partial^2}{\partial \theta_s \partial \omega_i} \left[\sum_{j=1}^n \omega_j \ell_j(\boldsymbol{\theta}) \right] \\ &= \frac{\partial}{\partial \theta_s} \frac{\partial}{\partial \omega_i} \left[\sum_{j=1}^n \omega_j \log(p_j(\boldsymbol{\theta})) \right] \\ &= \frac{\partial \log(p_i(\boldsymbol{\theta}))}{\partial \theta_s} \\ &= (\dot{\ell}_i(\boldsymbol{\theta}))_s, \end{aligned}$$

ou seja, $\Delta_i(\boldsymbol{\theta}) = \dot{\ell}_i(\boldsymbol{\theta})$. Assim, temos

$$\begin{aligned} \frac{(\Delta \Delta^\top)(\boldsymbol{\theta})}{n} &= \frac{\sum_{j=1}^n (\Delta_j \Delta_j^\top)(\boldsymbol{\theta})}{n} \\ &= \frac{\sum_{j=1}^n \dot{\ell}_j(\boldsymbol{\theta}) \dot{\ell}_j(\boldsymbol{\theta})^\top}{n}. \end{aligned}$$

Então, supondo que $\text{Var}(\dot{\ell}(\boldsymbol{\theta}) \dot{\ell}(\boldsymbol{\theta})^\top) < \infty$, pela Lei Fraca dos Grandes Números temos

$$\frac{(\Delta \Delta^\top)(\boldsymbol{\theta})}{n} = \frac{\sum_{j=1}^n \dot{\ell}_j(\boldsymbol{\theta}) \dot{\ell}_j(\boldsymbol{\theta})^\top}{n} \xrightarrow{P} \mathbb{E}[\dot{\ell}(\boldsymbol{\theta}) \dot{\ell}(\boldsymbol{\theta})^\top] = B(\boldsymbol{\theta}), \quad (5.6)$$

em que $B(\boldsymbol{\theta})$ está definida em (3.1). Logo, baseado em (5.6), definimos a matrix $B(\boldsymbol{\theta})$ para o esquema de perturbação no parâmetro de precisão como sendo

$$B(\boldsymbol{\theta}) = E[\Delta(\boldsymbol{\theta})\Delta(\boldsymbol{\theta})^\top], \quad (5.7)$$

em que a matriz $\Delta(\boldsymbol{\omega}) = (\Delta_\beta \quad \Delta_\phi)^\top$ com submatrizes definidas em (5.4) e (5.5), respectivamente. Logo, o ponto de corte que utilizaremos para este novo teste será $\lambda_0 = 2\text{tr}\{A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})\}$ em que $A(\boldsymbol{\theta})$ está definida em (3.1) e $B(\boldsymbol{\theta}) = E[\Delta(\boldsymbol{\theta})\Delta(\boldsymbol{\theta})^\top]$.

Para a obtenção da matrix $B(\boldsymbol{\theta})$, para o esquema perturbação no parâmetro de precisão, consideremos, para $i = 1, 2, \dots, n$, que $\Delta_i(\boldsymbol{\beta}) = \phi t_i \omega_i^{-2} f_i(\omega_i) X[, i]^\top$. Então,

$$E[\Delta_i(\boldsymbol{\beta})\Delta_i(\boldsymbol{\beta})^\top] = \frac{\phi^2 t_i^2}{\omega_i^4} E[f_i^2(\omega_i)] X[, i]^\top X[, i],$$

em que

$$\begin{aligned} E[f_i^2(\omega_i)] &= E\{[c_i(\omega_i) - (y_i^* - \mu_i^*(\omega_i))]^2\} \\ &= c_i^2(\omega_i) + E[(y_i^* - \mu_i^*(\omega_i))^2] - 2c_i(\omega_i)E(y_i^* - \mu_i^*(\omega_i)). \end{aligned} \quad (5.8)$$

Uma vez que estamos utilizando uma nova expressão para a precisão definida em (5.1), segue do resultado (2.6) e pela definição dada em (5.3) que $E(y_i^*) = \mu_i^*(\omega_i)$. Ainda, utilizando o resultado (2.7), obtemos que $\text{Var}(y_i^*) = v_i(\omega_i)$ com $v_i(\omega_i) = \psi'(\mu_i \phi / \omega_i) + \psi'((1 - \mu_i) \phi / \omega_i)$. Dessa forma, de (5.8) temos

$$E[f_i^2(\omega_i)] = c_i^2(\omega_i) + v_i(\omega_i).$$

Assim,

$$E[\Delta_i(\boldsymbol{\beta})\Delta_i(\boldsymbol{\beta})^\top] = \frac{\phi^2 t_i^2}{\omega_i^4} [c_i^2(\omega_i) + v_i(\omega_i)] X[, i]^\top X[, i].$$

Logo, temos que

$$E[\Delta(\boldsymbol{\beta})\Delta(\boldsymbol{\beta})^\top] = \phi^2 X^\top T \mathcal{W}^{-4} [C^2(\boldsymbol{\omega}) + V(\boldsymbol{\omega})] T^\top X,$$

em que $\mathcal{W}^{-4} = \text{diag}\{1/\omega_1^4, 1/\omega_2^4, \dots, 1/\omega_n^4\}$.

Também,

$$\begin{aligned} E[\Delta_i(\boldsymbol{\beta})\Delta_i(\phi)] &= \frac{\phi t_i}{\omega_i^4} X[, i]^\top E\{f_i(\omega_i)[\phi d_i(\omega_i) - u_i(\omega_i)]\} \\ &= \frac{\phi t_i}{\omega_i^4} X[, i]^\top \{\phi d_i(\omega_i)E[f_i(\omega_i)] - E[f_i(\omega_i)u_i(\omega_i)]\}. \end{aligned} \quad (5.9)$$

Notemos que

$$\begin{aligned} \mathbb{E}[f_i(\omega_i)] &= \mathbb{E}[c_i(\omega_i) - (y_i^* - \mu_i^*(\omega_i))] \\ &= c_i(\omega_i), \end{aligned}$$

como também

$$\begin{aligned} \mathbb{E}[f_i(\omega_i)u_i(\omega_i)] &= \mathbb{E}\left\{ [c_i(\omega_i) - (y_i^* - \mu_i^*(\omega_i))] \left[\mu_i(y_i^* - \mu_i^*(\omega_i)) + \log(1 - y_i) \right. \right. \\ &\quad \left. \left. - \psi\left(\frac{(1 - \mu_i)\phi}{\omega_i}\right) + \psi\left(\frac{\phi}{\omega_i}\right) \right] \right\} \\ &= c_i(\omega_i)\mu_i\mathbb{E}(y_i^* - \mu_i^*(\omega_i)) + c_i(\omega_i)\mathbb{E}[\log(1 - y_i)] \\ &\quad + c_i(\omega_i)\left[\psi\left(\frac{\phi}{\omega_i}\right) - \psi\left(\frac{(1 - \mu_i)\phi}{\omega_i}\right) \right] \\ &\quad - \mu_i\mathbb{E}[(y_i^* - \mu_i^*(\omega_i))^2] - \mathbb{E}[y_i^* \log(1 - y_i)] + \mu_i^*(\omega_i)\mathbb{E}[\log(1 - y_i)] \\ &\quad - \left[\psi\left(\frac{\phi}{\omega_i}\right) - \psi\left(\frac{(1 - \mu_i)\phi}{\omega_i}\right) \right] \mathbb{E}(y_i^* - \mu_i^*(\omega_i)). \end{aligned}$$

Novamente, dada a precisão definida em (5.1) temos

$$\mathbb{E}(y_i^* - \mu_i^*(\omega_i)) = 0 \quad \text{e} \quad \mathbb{E}[(y_i^* - \mu_i^*(\omega_i))^2] = v_i(\omega_i).$$

Além disso, de (3.18) temos

$$\mathbb{E}[\log(1 - y_i)] = \psi((1 - \mu_i)\phi/\omega_i) - \psi(\phi/\omega_i)$$

e de (3.16) temos

$$\mathbb{E}[y_i^* \log(1 - y_i)] = \mu_i^*[\psi((1 - \mu_i)\phi/\omega_i) - \psi(\phi/\omega_i)] - \psi'((1 - \mu_i)\phi/\omega_i).$$

Logo,

$$\begin{aligned} \mathbb{E}[f_i(\omega_i)u_i(\omega_i)] &= c_i(\omega_i)[\psi((1 - \mu_i)\phi/\omega_i) - \psi(\phi/\omega_i)] + c_i(\omega_i)\left[\psi\left(\frac{\phi}{\omega_i}\right) - \psi\left(\frac{(1 - \mu_i)\phi}{\omega_i}\right)\right] \\ &\quad - \mu_i v_i(\omega_i) - \mu_i^*[\psi((1 - \mu_i)\phi/\omega_i) - \psi(\phi/\omega_i)] + \psi'((1 - \mu_i)\phi/\omega_i) \\ &\quad + \mu_i^*(\omega_i)[\psi((1 - \mu_i)\phi/\omega_i) - \psi(\phi/\omega_i)] \\ &= -\mu_i v_i(\omega_i) + \psi'((1 - \mu_i)\phi/\omega_i) \\ &= -\mu_i[\psi'(\mu_i\phi/\omega_i) + \psi'((1 - \mu_i)\phi/\omega_i)] + \psi'((1 - \mu_i)\phi/\omega_i) \\ &= -[\mu_i\psi'(\mu_i\phi/\omega_i) - (1 - \mu_i)\psi'((1 - \mu_i)\phi/\omega_i)] \\ &= \frac{c_i(\omega_i)\omega_i}{\phi}. \end{aligned}$$

Portanto, voltando a (5.9), temos

$$\begin{aligned} E[\Delta_i(\boldsymbol{\beta})\Delta_i(\phi)] &= \frac{\phi t_i}{\omega_i^4} X[,i]^\top \left[\phi d_i(\omega_i) c_i(\omega_i) - \frac{c_i(\omega_i) \omega_i}{\phi} \right] \\ &= \frac{t_i}{\omega_i^4} c_i(\omega_i) [\phi^2 d_i(\omega_i) + 1] X[,i]^\top. \end{aligned}$$

Matricialmente,

$$E[\Delta(\boldsymbol{\beta})\Delta(\phi)] = X^\top \mathcal{W}^{-4} [\phi^2 D(\boldsymbol{\omega}) + \mathbb{I}_n] C(\boldsymbol{\omega}) \mathbf{t},$$

em que $D(\boldsymbol{\omega}) = \text{diag}\{d_1(\omega_1), d_2(\omega_1), \dots, d_n(\omega_n)\}$.

Analogamente, temos

$$E[\Delta(\phi)\Delta(\boldsymbol{\beta})^\top] = \mathbf{t}^\top C(\boldsymbol{\omega}) [\phi^2 D(\boldsymbol{\omega}) + \mathbb{I}_n] \mathcal{W}^{-4} X.$$

Por fim,

$$\begin{aligned} E[\Delta_i(\phi)\Delta_i(\phi)] &= E[\omega_i^{-2}(\phi d_i(\omega_i) - u_i(\omega_i))\omega_i^{-2}(\phi d_i(\omega_i) - u_i(\omega_i))] \\ &= \omega_i^{-4} E[\phi^2 d_i^2(\omega_i) - 2\phi d_i(\omega_i)u_i(\omega_i) + u_i^2(\omega_i)] \\ &= \omega_i^{-4} \{ \phi^2 d_i^2(\omega_i) - 2\phi d_i(\omega_i)E[u_i(\omega_i)] + E[u_i^2(\omega_i)] \}. \end{aligned}$$

De (3.19), obtemos que $E[u_i(\omega_i)] = 0$, como também, de (3.22) temos que $E[u_i^2(\omega_i)] = d_i(\omega_i)$. Dessa forma,

$$E[\Delta_i(\phi)\Delta_i(\phi)] = \omega_i^{-4} [\phi^2 d_i^2(\omega_i) + d_i(\omega_i)].$$

Matricialmente,

$$\begin{aligned} E[\Delta(\phi)\Delta(\phi)^\top] &= \mathcal{W}^{-4} [\phi^2 \sum_{i=1}^n d_i^2(\omega_i) + \sum_{i=1}^n d_i(\omega_i)] \\ &= \mathcal{W}^{-4} [\phi^2 \text{tr}(D^2(\boldsymbol{\omega})) + \text{tr}(D(\boldsymbol{\omega}))], \end{aligned}$$

em que $D^2(\boldsymbol{\omega}) = \text{diag}\{d_1^2(\omega_1), d_2^2(\omega_2), \dots, d_n^2(\omega_n)\}$.

Portanto, a matriz $B(\boldsymbol{\theta})$ é expressa por

$$B(\boldsymbol{\theta}) = \begin{bmatrix} \phi^2 X^\top T \mathcal{W}^{-4} [C^2(\boldsymbol{\omega}) + V(\boldsymbol{\omega})] T^\top X & X^\top \mathcal{W}^{-4} [\phi^2 D(\boldsymbol{\omega}) + \mathbb{I}_n] C(\boldsymbol{\omega}) \mathbf{t} \\ \mathbf{t}^\top C(\boldsymbol{\omega}) [\phi^2 D(\boldsymbol{\omega}) + \mathbb{I}_n] \mathcal{W}^{-4} X & \mathcal{W}^{-4} [\phi^2 \text{tr}(D^2(\boldsymbol{\omega})) + \text{tr}(D(\boldsymbol{\omega}))] \end{bmatrix}. \quad (5.10)$$

Avaliando $B(\boldsymbol{\theta})$ em $\boldsymbol{\omega}_0$ e $\hat{\boldsymbol{\theta}}$ temos

$$B(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \hat{\phi}^2 X^\top \hat{T} [\hat{C}^2 + \hat{V}] \hat{T}^\top X & X^\top [\hat{\phi}^2 \hat{D} + \mathbb{I}_n] \hat{C} \hat{\mathbf{t}} \\ \hat{\mathbf{t}}^\top \hat{C} [\hat{\phi}^2 \hat{D} + \mathbb{I}_n] X & \hat{\phi}^2 \text{tr}(\hat{D}^2 + \text{tr}(\hat{D})) \end{bmatrix}. \quad (5.11)$$

Consideremos $T_n = \sum_{i=1}^n n^{-1/2}(nC_{ei} - 2\text{tr}\{A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})\})$. Definimos a estatística de teste como sendo

$$\begin{aligned} U_n^* &= \frac{T_n}{\text{ep}(T_n)} \\ &= n^{-1/2} \sum_{i=1}^n \frac{(nC_{ei} - 2\text{tr}\{A(\boldsymbol{\theta})^{-1}B(\boldsymbol{\theta})\})}{\text{ep}(T_n)}, \end{aligned} \quad (5.12)$$

em que $A(\boldsymbol{\theta}) = K(\boldsymbol{\beta}, \phi)$ definida na Seção 2.2, $B(\boldsymbol{\theta})$ está definida em (5.10) e $\text{ep}(T_n)$ é estimado, utilizando a metodologia *bootstrap* apresentada na Subseção 1.2.2 e aplicada na Subseção 4.1.1. Assim, $\text{ep}(T_n)$ é estimado através do cálculo do erro-padrão das réplicas T_n obtidas através do *bootstrap*.

Portanto, o novo teste apresentado, no capítulo, apresenta-se como uma proposta que é baseada em uma metodologia consolidada (influência local) para identificar a necessidade da modelagem da dispersão em situações que foram ajustados modelos de regressão beta com dispersão fixa, erroneamente.

Capítulo 6

Estudos de Simulação: Parte 2

Nesse capítulo, apresentaremos alguns resultados de estudos de simulação de Monte Carlo com o intuito de avaliar o desempenho do novo teste que propomos para a precisão ϕ para a classe de modelos de regressão beta com dispersão fixa, com intuito de detectar a necessidade da modelagem da dispersão. Para tanto, foram calculadas as taxas de rejeição sob a hipótese nula (tamanhos do teste). Em todas as simulações, foram consideradas 5000 réplicas de Monte Carlo, tamanhos amostrais iguais a $n = 40, 80$ e 120 e níveis nominais $\alpha = 1\%, 5\%$ e 10% . Para obtermos tais amostras, foi, novamente utilizado o processo de replicação das amostras. A especificação da média do modelo foi

$$g(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right) = \sum_{j=1}^5 \beta_j x_{ij},$$

em que $g(\cdot)$ é a função de ligação logito, β_j e x_{ij} , com $j = 1, \dots, 5$, são os parâmetros a serem estimados e as variáveis explicativas da média, respectivamente.

Consideramos y_1, y_2, \dots, y_n variáveis aleatórias independentes em que $y_i \sim \text{Beta}(\mu_i, \phi)$ com $i = 1, 2, \dots, n$. Para a obtenção dos tamanhos do teste, a variável resposta foi gerada e modelada sem erro de especificação do modelo. Estimamos os parâmetros do modelo, e em seguida, aplicamos o teste baseado em influência local apresentado na Subseção 5.2. As covariáveis utilizadas, para tal situação, foram geradas a partir das distribuições uniforme padrão, normal padrão, lognormal padrão, e da distribuição Bernoulli de parâmetro 0,05. Novamente, consideramos três cenários para os valores de μ : próximo do zero, próximo de um e central. Em relação aos valores utilizados para os parâmetros que compõem a componente sistemática, considerando o cenário em que μ é próximo do zero, utilizamos $\beta_1 = -1,6, \beta_2 = -1,2, \beta_3 = -1,9, \beta_4 = -0,4$ e $\beta_5 = -2,9$. Para o cenário em que μ é próximo de um, consideramos $\beta_1 = 1,9, \beta_2 = 0,2, \beta_3 = 1,9, \beta_4 = 1,0$ e $\beta_5 = 4,9$. Por fim, quando μ é central utilizamos $\beta_1 = -1,0, \beta_2 = 1,0, \beta_3 = 1,2, \beta_4 = 1,0$ e $\beta_5 = 1,3$. Além disso, foram considerados três cenários para o parâmetro da precisão: ϕ pequeno, médio e grande, mais especificamente, $\phi = 33, 74$ e 150 . Para o procedimento *bootstrap* que empregamos para a obtenção da variância da estatística T_n e a obtenção dos quantis

de U_n^* (ver Seção 5.2) utilizamos $B = 500$ réplicas *bootstrap*. Dessa forma, obtivemos os percentuais de rejeição sob hipótese nula de que o modelo está corretamente especificado (tamanho do teste) como também avaliamos os poderes do teste, através da introdução de algumas formas de erro de especificação do modelo. Primeiramente, avaliamos o poder do teste em relação à omissão de covariáveis. Nas Tabelas 6.1, 6.2 e 6.3 estão apresentados os percentuais de rejeição sob hipótese nula (Cenário 4), como também, os percentuais de rejeição sob a hipótese alternativa, ou seja, os poderes do teste sob a omissão de covariáveis considerando os níveis 1%, 5% e 10%, respectivamente.

Tabela 6.1: Níveis empíricos para o teste, usando nível nominal 1%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0055; 0, 1662]$												
$\phi \backslash n$	33	74	150	33	74	150	33	74	150	33	74	150
40	21,28	25,96	62,12	17,18	19,34	34,48	10,10	7,02	7,76	1,40	1,24	1,36
80	18,90	27,86	77,10	15,92	18,02	40,12	9,68	7,86	8,48	1,52	1,20	1,28
120	20,28	34,68	91,12	17,42	20,68	54,32	10,34	8,14	11,06	1,56	1,24	1,32
$\mu \in [0, 2738; 0, 7202]$												
$\phi \backslash n$	33	74	150	33	74	150	33	74	150	33	74	150
40	3,54	12,42	44,82	2,50	6,36	21,18	4,50	3,80	3,38	1,56	1,34	1,42
80	4,26	18,78	72,10	2,62	8,44	34,90	6,76	6,60	5,46	1,18	1,18	1,50
120	5,96	30,68	90,84	3,64	12,20	54,34	8,02	9,08	6,56	1,36	1,40	1,30
$\mu \in [0, 8659; 0, 9990]$												
$\phi \backslash n$	33	74	150	33	74	150	33	74	150	33	74	150
40	31,52	41,34	44,28	27,94	39,90	45,12	17,02	21,84	23,68	1,54	1,62	1,64
80	38,36	42,24	42,90	34,34	40,68	43,58	23,06	24,94	24,22	1,18	1,48	1,56
120	66,56	58,40	57,56	59,46	55,78	56,46	41,88	36,84	34,28	0,52	1,08	1,02

Notamos que o tamanho do teste possui ótimo desempenho no sentido de está controlado, ou seja, próximo dos níveis nominais fixados. Em relação ao poder do teste, como esperado, maior o nível nominal fixado, maior o poder do teste obtido. As conclusões, a seguir são referentes aos três níveis nominais. Como esperado, à medida que é omitido um número maior de covariáveis, maior é o poder obtido (exceto quando $\phi = 33$, μ central no Cenário 3). Observando o Cenário 1, notamos que o poder do teste é mais significativo quando μ está próximo do extremo superior. Para os Cenários 2 e 3, quando $\phi = 33$ e 74, a situação em que μ pertence ao extremo superior leva o teste possuir poder maior que as demais situações, porém quando $\phi = 150$, o cenário μ no extremo inferior, é a situação que faz com que o teste possua maior poder. Analogamente, aos resultados obtidos anteriormente, quando μ é central o teste possui poderes mais baixos que os outros dois cenários. Por último, notemos que, na grande maioria dos casos, para todos os cenários quando a precisão ϕ aumenta, os poderes do teste aumentam significativamente.

Tabela 6.2: Níveis empíricos para o teste, usando nível nominal 5%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0055; 0, 1662]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	53,14	65,30	95,22	43,68	48,52	71,30	27,04	23,48	23,68	5,74	5,10	4,96
80	54,04	72,64	99,24	46,94	53,08	82,60	31,94	25,72	28,40	6,10	5,10	5,50
120	59,26	83,56	99,86	51,38	60,62	92,16	34,44	28,94	35,12	6,32	4,96	5,42
$\mu \in [0, 2738; 0, 7202]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	11,76	33,40	76,80	8,40	19,70	48,40	11,98	11,06	9,60	5,70	4,82	5,76
80	14,52	47,30	92,90	8,82	25,20	68,02	15,96	15,12	13,56	5,22	5,08	5,52
120	18,32	62,48	98,88	10,44	32,24	84,48	18,08	19,72	16,18	5,56	5,60	5,56
$\mu \in [0, 8659; 0, 9990]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	67,64	79,46	85,66	61,64	74,64	81,96	41,94	51,82	53,66	5,84	6,64	5,60
80	79,68	83,98	87,58	73,88	80,80	86,12	56,94	61,06	61,76	6,28	5,60	6,12
120	95,28	93,64	94,48	91,24	91,14	93,10	79,08	75,30	73,30	6,56	4,44	4,96

Tabela 6.3: Níveis empíricos para o teste usando nível nominal 10%. Má especificação: covariáveis omitidas (Cenários 1, 2 e 3).

Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
Modelo ajustado	$g(\mu_i) = \beta_1 + \beta_2 x_{i2}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$			$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5}$		
$\mu \in [0, 0055; 0, 1662]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	75,48	86,36	99,54	63,46	69,36	88,72	41,54	37,24	37,72	10,56	9,80	9,62
80	79,00	92,92	99,98	68,82	76,20	95,94	50,22	42,68	45,36	11,52	10,78	11,22
120	85,46	97,62	100,00	73,88	83,82	98,88	55,16	47,90	54,70	11,02	9,70	9,84
$\mu \in [0, 2738; 0, 7202]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	21,02	49,14	88,48	14,88	31,32	65,82	18,90	18,12	16,70	10,56	9,90	11,18
80	23,84	64,86	97,70	15,52	38,68	82,28	23,68	23,36	21,20	10,10	10,02	10,34
120	29,48	77,38	99,78	17,84	47,70	93,90	27,30	27,46	23,68	10,46	10,40	10,68
$\mu \in [0, 8659; 0, 9990]$												
$n \backslash \phi$	33	74	150	33	74	150	33	74	150	33	74	150
40	87,46	94,42	98,42	79,52	90,08	94,80	60,98	69,60	72,12	11,34	12,40	11,10
80	95,10	96,80	98,44	90,10	94,46	96,98	76,66	80,76	82,54	11,88	11,04	11,26
120	98,74	99,32	99,62	97,38	98,26	99,06	90,00	90,20	90,40	10,18	8,90	10,10

Também, avaliamos o poder do teste no que se diz respeito à variável resposta ser gerada através da distribuição simplex e ser modelada utilizando a regressão beta com dispersão fixa. As covariáveis utilizadas para essa situação foram geradas a partir da distribuição uniforme padrão. Também os parâmetros da componente sistemática utilizados para o extremo inferior de μ , foram $\beta_1 = -3, 0, \beta_2 = -1, 2, \beta_3 = 0, 6, \beta_4 = 1, 0$ e $\beta_5 = -1, 1$. Para o extremo superior de μ , consideramos $\beta_1 = 3, 5, \beta_2 = 1, 2, \beta_3 = -0, 8, \beta_4 = -1, 1$ e $\beta_5 = 0, 5$. Por fim, para μ central, utilizamos $\beta_1 = -1, 9, \beta_2 = 1, 2, \beta_3 = 1, 0, \beta_4 = 1, 1$ e $\beta_5 = 1, 3$. Além disso, foram considerados três valores para a precisão são $\phi = 20, 55$ e 150 . A Tabela 6.4 apresenta as taxas de rejeição dos testes para esse erro de especificação, isto é, os poderes do teste para este tipo de má especificação.

Observamos que, de acordo com que o n e ϕ crescem, os poderes aumentam. Percebemos

Tabela 6.4: Níveis empíricos para o teste usando os níveis nominais 1%, 5% e 10%. Má especificação: variável resposta simplex.

	$\mu \in [0, 0103; 0, 1092]$			$\mu \in [0, 2131; 0, 8289]$			$\mu \in [0, 9101; 0, 9886]$		
$n \backslash \phi$	33	74	150	33	74	150	33	74	150
	$\alpha = 1\%$								
40	33,50	29,16	25,20	1,50	1,50	1,30	30,70	24,90	22,78
80	57,00	53,48	52,82	1,58	1,80	2,04	51,80	49,32	49,06
120	73,90	73,12	72,18	1,56	1,94	2,00	69,04	66,96	66,74
	$\alpha = 5\%$								
40	52,66	47,12	43,68	5,56	5,66	5,92	49,92	42,74	40,92
80	76,70	74,24	71,48	5,86	6,50	6,66	74,16	70,88	69,44
120	89,50	88,04	87,30	6,12	6,96	6,14	86,66	84,64	84,50
	$\alpha = 10\%$								
40	63,70	57,76	54,48	10,46	10,78	11,24	61,92	54,56	52,04
80	84,80	82,94	79,70	10,68	11,88	12,02	83,60	80,54	78,42
120	94,54	92,74	92,92	11,62	12,02	11,94	92,40	90,80	90,06

também que os poderes possuem desempenho semelhante para os extremos de μ . Notemos que para o cenário em que μ é central, os poderes do teste são muito baixos, praticamente próximos ao o que se espera para o tamanho do teste. Esse resultado é exatamente o mesmo obtido anteriormente (ver Seção 4.1), quando utilizamos o modelo de regressão beta com dispersão fixa para ajustar a média de uma variável resposta qualquer em que $\mu_i \approx 0,5, i = 1, 2, \dots, n$, não será simples identificar se o modelo beta é o verdadeiro gerador de dados, pois o modelo simplex possivelmente também irá produzir um bom ajuste.

O último tipo de má especificação que introduzimos foi gerar a variável resposta através de um modelo de regressão beta com dispersão variável e estimar através do modelo de regressão beta com dispersão fixa. Vale ressaltar que três das covariáveis utilizadas para essa situação foram geradas a partir da distribuição uniforme padrão e uma a partir da distribuição Bernoulli de parâmetro 0,1. Os parâmetros referentes à componente sistemática da média para o extremo inferior de μ foram $\beta_1 = -1, 2, \beta_2 = -1, 2, \beta_3 = -1, 0, \beta_4 = -1, 1$ e $\beta_5 = -1, 1$. Para o extremo superior de μ utilizamos $\beta_1 = 1, 5, \beta_2 = 1, 1, \beta_3 = 1, 3, \beta_4 = 0, 1$ e $\beta_5 = 1, 3$ para a média. Por fim, para μ central temos $\beta_1 = -1, 6, \beta_2 = 0, 5, \beta_3 = 0, 8, \beta_4 = 1, 1$ e $\beta_5 = 1, 3$. Para introduzir a dispersão variável, a especificação do modelo para a dispersão contou com covariáveis geradas, através da distribuição uniforme padrão e da distribuição Bernoulli de parâmetro 0,1. Ainda, os parâmetros que compõem a componente sistemática da dispersão foram $\gamma_1 = 3, 8, \gamma_2 = 3, 3, \gamma_3 = -1, 0, \gamma_4 = -2, 3$ e $\gamma_5 = 2, 1$. Consequentemente, foram gerados cinco valores para o grau de heterogeneidade: $\lambda = 1, 25, 54, 121$ e 326 . Notamos que quando $\lambda = 1$ não temos dispersão variável nos dados. Os resultados para os poderes do teste, nesse caso, estão apresentados na Tabela 6.5. Vimos que os tamanhos empíricos do teste (quando $\lambda = 1$) estão próximos dos níveis nominais fixados para todos os cenários, segundo esperado. Em relação ao poder do teste, de acordo com aumento do valor de n , esse cresce significativamente. Quando o grau de heterogeneidade λ cresce, o poder

do teste cresce também, mas não tanto quanto o esperado. Os poderes para as situações em que μ está nas extremidades possuem desempenhos semelhantes. O cenário em que o poder foi mais baixo, novamente, foi quando μ é central.

Portanto, novamente a metodologia *bootstrap* foi bastante útil para obter os quantis da estatística de teste U_n^* do novo teste proposto para detectar a necessidade de modelagem da dispersão em modelos de regressão beta com dispersão fixa. Obtivemos ótimos tamanhos de teste para todos os cenários apresentados, como também, bons poderes do teste quando consideradas as más especificações: omissão de covariáveis, variável resposta simplex, variável resposta com dispersão variável.

Tabela 6.5: Níveis empíricos para o teste usando os níveis nominais 1,5% e 10%. Má especificação: dispersão não modelada.

λ n		$\mu \in [0, 0107; 0, 1477]$					$\mu \in [0, 2515; 0, 7738]$					$\mu \in [0, 8825; 0, 9892]$				
		1	25	54	121	326	1	25	54	121	326	1	25	54	121	326
$\alpha = 1\%$																
40	1,00	20,74	29,26	44,14	46,24	1,30	14,46	20,92	38,86	39,86	1,30	19,26	29,14	43,70	45,12	
80	1,46	46,44	61,10	82,18	82,52	1,48	36,12	46,22	74,30	76,04	1,34	45,70	60,40	81,08	82,76	
120	1,52	63,82	80,04	95,16	96,14	1,24	51,90	65,34	90,94	91,74	1,38	64,34	79,06	94,68	95,02	
$\alpha = 5\%$																
40	5,00	36,52	47,76	63,80	66,06	5,14	26,80	35,04	55,50	57,62	5,12	34,84	45,54	62,84	64,52	
80	5,44	65,08	77,74	93,06	93,68	5,66	52,02	62,94	86,94	87,94	5,38	64,56	76,10	92,46	93,24	
120	5,74	81,48	90,32	98,72	98,88	5,80	68,98	79,74	96,90	97,48	5,48	80,20	90,20	98,86	98,62	
$\alpha = 10\%$																
40	10,18	46,88	57,72	72,78	74,56	10,26	36,00	42,88	64,82	66,88	10,00	44,96	55,52	72,02	73,46	
80	9,80	73,86	84,30	96,40	96,32	10,40	60,40	70,96	91,46	92,36	10,02	73,00	83,24	95,72	96,04	
120	10,78	87,94	94,48	99,50	99,34	10,98	76,68	85,78	98,18	98,84	10,46	86,48	93,98	99,58	99,48	

Capítulo 7

Aplicações

Nesse capítulo, apresentaremos uma aplicação a dados reais do teste de hipóteses proposto para avaliar a qualidade do ajuste dos modelos de regressão beta com dispersão fixa e variável, como também, ao novo teste proposto para a precisão.

7.1 Dados de Dislexia

Nessa aplicação, utilizaremos os dados analisados por Smithson e Verkuilen (2006), obtidos de Pammer e Kevan (2004). Os dados são referentes a um estudo sobre a habilidade em leitura de um grupo de 44 crianças, realizado na escola de psicologia da Universidade Nacional da Austrália. A variável resposta é o escore em um teste de habilidade em leitura, em que as covariáveis são a contribuição relativa da sensibilidade visual (presença de dislexia) e um escore padronizado de QI não-verbal na habilidade de leitura das 44 crianças. A covariável dislexia assume os valores 1 e -1 , indicando se a criança é ou não disléxica, respectivamente. A variável resposta foi obtida após duas transformações dos dados originais. Originalmente, os escores do teste de leitura, que denotamos por y' , encontram-se no intervalo $[a, b]$, em que a é o escore mínimo possível b é o escore máximo possível para o teste de leitura. A primeira transformação realizada nos dados foi $y'' = (y_0 - a)/(b - a)$, em seguida, para garantir que a variável resposta pertence ao intervalo aberto $(0, 1)$, os autores fazem a segunda transformação, $y = [y''(n - 1) + 0, 5]/n$. Inicialmente, Smithson e Verkuilen (2006) ajustam um modelo de regressão beta com ligação logito para a média utilizando as covariáveis dislexia, QI e a interação entre essas covariáveis. Dessa forma, seguindo a sugestão dos mesmos, ajustamos o modelo citado. Assim, o modelo para a média é especificado como sendo

$$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4},$$

em que $g(\cdot)$ é a função de ligação logito definida por

$$g(\mu_i) = \log \left(\frac{\mu_i}{1 - \mu_i} \right) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}.$$

As estimativas dos parâmetros para esse modelo (modelo 1) bem como, os valores- p dos testes Wald de significância dos parâmetros estão apresentados na Tabela 7.1. De acordo com os valores- p , temos que apenas a covariável dislexia se mostra significativa aos níveis de significância usuais, digamos, 5%, para explicar a variável resposta.

Tabela 7.1: Estimativas dos parâmetros para o modelo de regressão beta com dispersão fixa considerando apenas as covariável dislexia, QI e interação entre ambas.

Parâmetro	Estimativa	Valor- p
β_1	1,3338	0,0000
β_2	-0,9736	0,0000
β_3	0,1608	0,3325
β_4	-0,2186	0,1877
ϕ	11,1332	0,0000

Os autores argumentam que a pouca informação que as covariáveis QI e interação de QI com dislexia trazem para a explicação da resposta pode estar sendo causada porque o efeito das covariáveis na dispersão dos dados não está sendo considerada. Apesar disso, avaliamos a qualidade do ajuste utilizando os testes apresentados nas Seções 3.2 (ponderação de casos) e 5.2 (perturbação de ϕ). Na Tabela 7.2, apresentamos os valores das estatísticas U_n e U_n^* , como também, a Tabela 7.3 contém os valores críticos empíricos da estatística U_n ($U_{\alpha/2}$ e $U_{(1-\alpha/2)}$) e assintóticos ($z_{\alpha/2}$ e $z_{(1-\alpha/2)}$), como também os valores críticos empíricos da estatística U_n^* ($U_{\alpha/2}^*$ e $U_{(1-\alpha/2)}^*$), considerando os níveis nominais 1%, 5% e 10%.

Tabela 7.2: Valores das estatísticas U_n e U_n^* considerando o modelos de regressão beta com dispersão fixa para os dados de dislexia.

Teste	Estatística
Ponderação de casos (U_n)	-1,2969
Perturbação de ϕ (U_n^*)	-1,8703

Tabela 7.3: Quantis empíricos e assintóticos para as estatísticas U_n e U_n^* considerando o modelo de regressão beta com dispersão fixa.

Nível	$U_{\alpha/2}$	$U_{(1-\alpha/2)}$	$z_{\alpha/2}$	$z_{(1-\alpha/2)}$	$U_{\alpha/2}^*$	$U_{(1-\alpha/2)}^*$
1%	-2,2367	0,6572	-2,58	2,58	-2,4517	3,2057
5%	-1,7053	0,4772	-1,96	1,96	-2,0922	1,6622
10%	-1,5027	0,3957	-1,65	1,65	-1,9366	1,2104

Primeiramente, notemos que as hipóteses testadas nesse caso são H_0 : o modelo de regressão beta com dispersão fixa é adequado para esses dados versus H_1 : o modelo não está corretamente especificado. Analisando os valores das estatísticas U_n e U_n^* , chegamos à conclusão de que não devemos rejeitar H_0 aos níveis de significância usuais, ou seja, utilizando ambos os testes não rejeitamos a hipótese de que o modelo de regressão beta

com dispersão fixa é o modelo adequado para ajustar esses dados. Porém, essas conclusões são errôneas, pois o modelo considerado não é adequado para ajustar esses dados devido à necessidade de modelagem na dispersão dos dados que veremos posteriormente. Em um primeiro momento, somos levados a pensar que ambos os testes estão com desempenho ruim. A partir desses resultados, desconfiamos que o teste está nos fornecendo uma informação que não era esperada: o teste está indicando quais são as covariáveis importantes para a modelagem da média, isto é, o mesmo está agindo como um método para selecionar covariáveis explicativas do modelo. Esse fato será reafirmado mais a frente, em especial com relação ao teste ϕ .

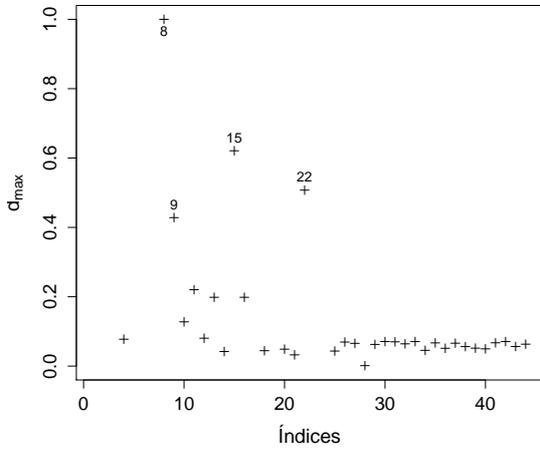
Construímos os gráficos de influência local para ambos esquemas de perturbação objetivando identificar observações influentes. Nas Figuras 7.1 (a) e 7.1 (b), temos os gráficos de influência local conjunta e individual, respectivamente, considerando o esquema de perturbação ponderação de casos, como também, nas Figuras 7.1 (c) e 7.1 (d) temos os gráficos de influência local conjunta e individual, respectivamente, considerando o esquema de perturbação do parâmetro ϕ .

Percebemos a partir desses gráficos uma grande diferença nos valores assumidos pelas medidas de influência quando comparamos as crianças dislexas e as que não são. Esse comportamento pode ser indícios da necessidade da modelagem da dispersão. Em ambos os gráficos, as observações 8, 9, 15, 19, 22 e 24 se apresentam como possíveis casos influentes. Com o intuito de analisar o impacto dessas observações nas estimativas dos parâmetros, realizamos uma análise confirmatória reajustando o modelo sem as observações consideradas influentes. Depois calculamos a mudança relativa em porcentagem das estimativas para os parâmetros do modelo, definida por

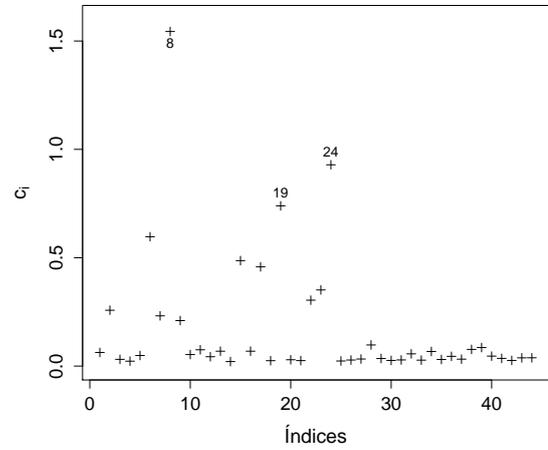
$$\text{MR}_{\theta}(i) = \left| \frac{\hat{\theta} - \hat{\theta}_{(i)}}{\hat{\theta}} \right| \times 100\%,$$

em que $\hat{\theta}_{(i)}$ denota a estimativa do parâmetro θ após a retirada da observação i . Calculamos as mudanças relativas e as apresentamos na Tabela 7.4. Vemos que a observação 8 causa uma mudança relativa significativa nas estimativas dos parâmetros β_3 e β_4 , como também a observação 15. Ainda, após retirar a observação 8, reajustando o modelo, observamos que as covariáveis QI e interação são significativas para explicação da média com valores- p para o teste Wald iguais a 0,0516 e 0,0178, respectivamente. Dessa forma, notamos que a exclusão da observação 8 causa mudança inferencial quando avaliamos através do teste Wald. Em contrapartida, as observações 9, 19, 22 e 24 não causam mudança relativa significativa individualmente. Por fim, as observações conjuntamente, causam mudança relativa muito significativa.

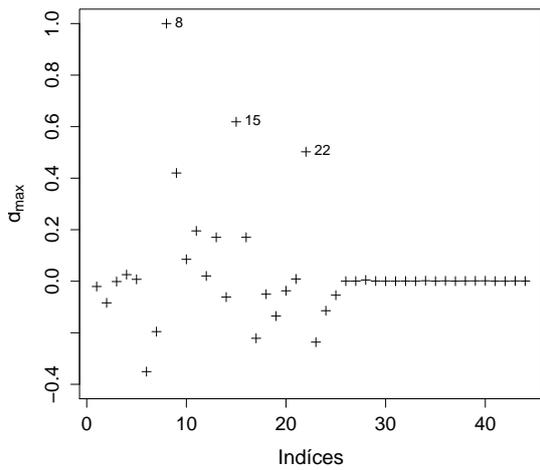
Figura 7.1: Gráficos de influência local conjunta e individual, considerando o esquema de ponderação de casos e perturbação no parâmetro ϕ , respectivamente.



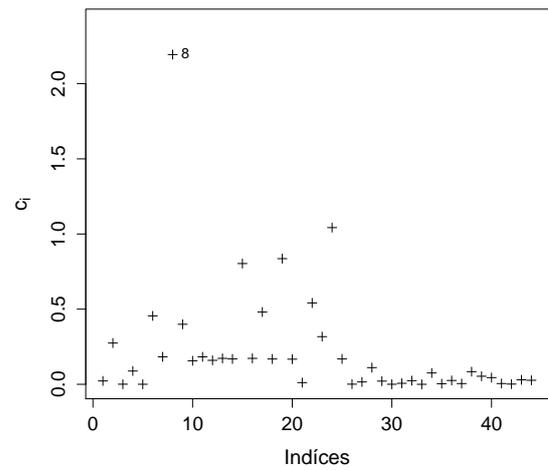
(a)



(b)



(c)



(d)

Fonte: A autora.

Assim, percebemos que os resultados obtidos a partir do testes baseados em influência local parecem nos levar a concluir que, os mesmos são robustos a observações influentes, pois apesar da observação 8 ser altamente influente, os testes indicam a necessidade das covariáveis QI e interação para a modelagem da média.

Os gráficos dos resíduos versus as covariáveis, foram construídos objetivando identificar graficamente se há indícios da necessidade de modelagem da dispersão, a qual assumimos constante para todas as observações (Figura 7.2), como também apresentamos o gráfico de envelope com objetivo de verificar possível falta de qualidade do ajuste. O resíduo

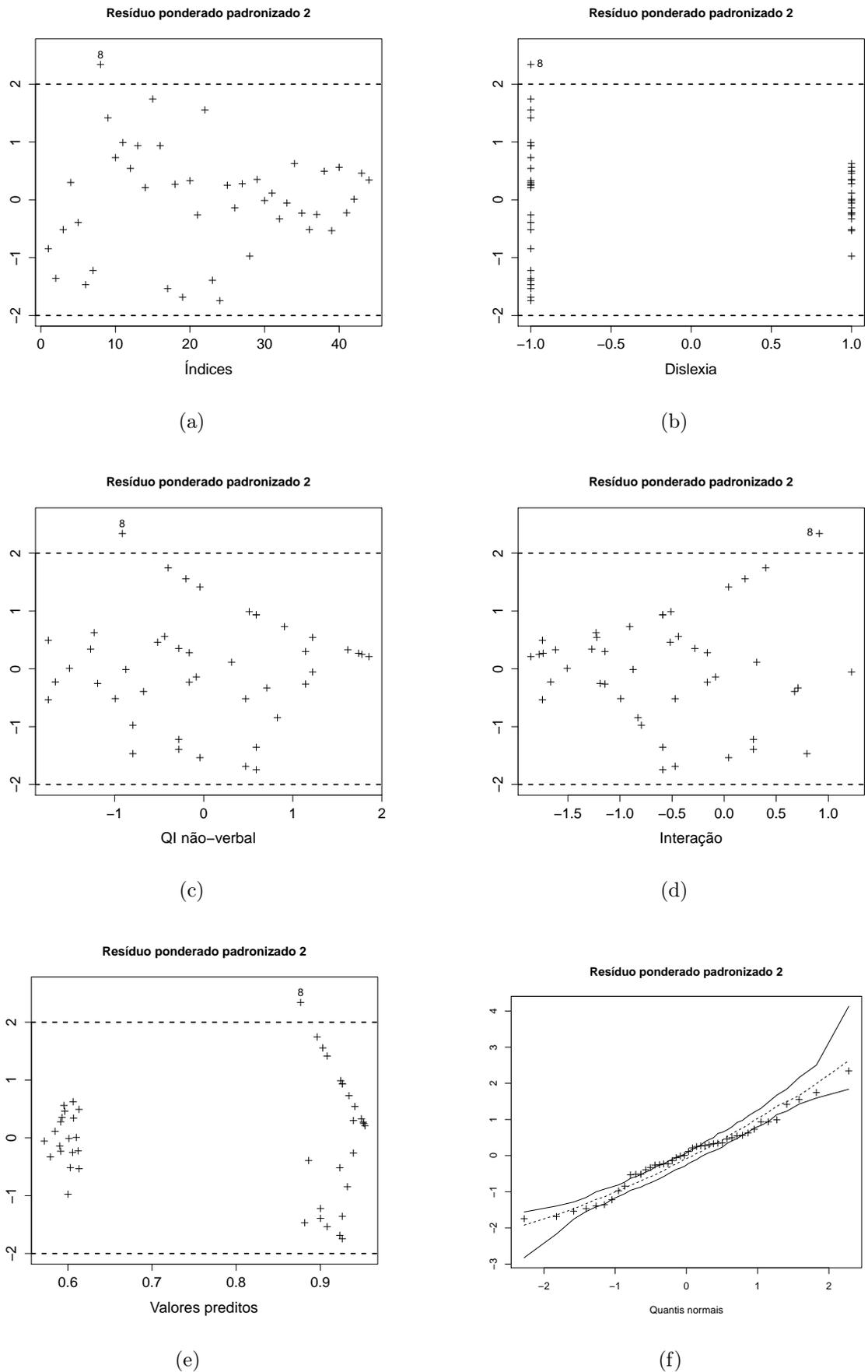
Tabela 7.4: Mudanças relativas: Modelo 1.

Observação \ Parâmetro	β_1	β_2	β_3	β_4	ϕ
8	5,97%	8,42%	65,59%	48,44%	7,71%
9	2,26%	3,13%	9,98%	7,37%	1,19%
15	3,28%	4,59%	23,56%	17,39%	2,77%
19	4,05%	5,15%	4,81%	3,85%	13,38%
22	2,67%	3,71%	15,16%	11,19%	1,80%
24	4,25%	5,36%	8,50%	6,61%	15,70%
8, 9, 15, 19, 22 e 24	13,80%	20,98%	248,35%	184,25%	140,57%

considerado é o ponderado padronizado 2, apresentado em Espinheira et al. (2008b), que segundo os autores é o resíduo mais indicado para a classe de modelos de regressão beta se o objetivo é a captação de casos decisivos no ajuste do modelo. Notamos que as crianças dislexas apresentam maior variabilidade quanto comparamos com as crianças não dislexas. Com relação à covariável QI, percebemos que existe maior variabilidade para as crianças que possuem QI entre -1 e 1 do que para aquelas que possuem QI assumindo valores extremos, como também, parece haver variabilidade distinta com relação a interação e os valores preditos de acordo com as observações. Logo, acreditamos que as covariáveis dislexia e QI são importantes para a modelagem da dispersão presentes nesses dados, como também, notamos uma falta de qualidade do ajuste através do gráfico de envelope. Em seguida, retiramos as covariáveis que não se mostraram significativas e estimamos o modelo considerando apenas a covariável dislexia. As estimativas dos parâmetros para este modelo, bem como os valores- p dos testes de significância dos parâmetros estão apresentados na Tabela 7.5. Avaliamos novamente a qualidade do ajuste utilizando os testes apresentados na Seções 3.2 (ponderação de casos) e 5.2 (perturbação de ϕ). Na Tabela 7.6 estão apresentados os valores das estatísticas U_n e U_n^* , como também, na Tabela 7.7 temos os valores críticos empíricos da estatística U_n e assintóticos, os valores críticos empíricos da estatística U_n^* considerando os níveis nominais 1%, 5% e 10%. Notamos que utilizando o teste baseado em ponderação de casos, não rejeitamos H_0 a nenhum dos níveis usuais. Já, através do teste para detectar a necessidade de modelagem da dispersão, rejeitamos H_0 a todos os níveis de significância usuais. Notemos que o teste baseado na perturbação no parâmetro ϕ se mostra bem mais eficaz do que o teste baseado em ponderação de casos, pois agora o teste ϕ rejeita a hipótese que o modelo está bem ajustado. Isso tem duas interpretações importantes. Primeira, no modelo falta a covariada QI, cuja teoria diz que é importante para o modelo. Assim, está faltando uma covariada importante e o teste ϕ detecta isto. Segundo, o teste ϕ detecta a necessidade da modelagem da dispersão e, de fato, ao modelarmos corretamente a dispersão QI passa a ser significativo.

Nas Figuras 7.3 (a) e 7.3 (b), temos os gráficos de influência local conjunta e individual, respectivamente, considerando o esquema de perturbação ponderação de casos, como também, nas Figuras 7.3 (c) e 7.3 (d) temos os gráficos de influência local conjunta

Figura 7.2: Gráficos de resíduos do modelo, juntamente com o envelope.



Fonte: A autora.

e individual, respectivamente, considerando o esquema de perturbação do parâmetro ϕ . Novamente, há grande diferença nos valores assumidos pelas medidas de influência quando comparamos as crianças dislexas e as que não são. Esse comportamento pode ser indícios da necessidade da modelagem da dispersão.

Identificamos as observações 6, 8, 17, 19, 23 e 24 como possivelmente influentes. Calculamos as mudanças relativas nas estimativas e apresentamos na Tabela 7.8. Quando consideramos as observações individuais, não notamos mudança relativa nos parâmetros. Em contrapartida, conjuntamente, as observações causam mudança relativa significativa nos parâmetros, principalmente no parâmetro ϕ . Na Figura 7.4 temos os gráficos dos resíduos, juntamente com o envelope, notemos que não há nenhum resíduo alto (fora do intervalo $[-2,2]$), mas vemos fortes indícios da necessidade de modelagem da dispersão (mesmos indícios identificados anteriormente), como também, observamos falta de qualidade do ajuste quando analisamos o gráfico de envelope.

Tabela 7.5: Estimativas dos parâmetros para o modelo de regressão beta com dispersão fixa considerando apenas a covariável dislexia.

Parâmetro	Estimativa	Valor- p
β_1	1,4437	0,0000
β_2	-1,0488	0,0000
ϕ	10,2972	0,0000

Tabela 7.6: Valores das estatísticas U_n e U_n^* considerando o modelos de regressão beta fixa para os dados de estresse.

Teste	Estatística
Ponderação de casos (U_n)	-1,1355
Pertubação de ϕ (U_n^*)	-1,9815

Tabela 7.7: Quantis empíricos e assintóticos para as estatísticas U_n e U_n^* considerando o modelo de regressão beta com dispersão fixa.

Nível	$U_{\alpha/2}$	$U_{(1-\alpha/2)}$	$z_{\alpha/2}$	$z_{(1-\alpha/2)}$	$U_{\alpha/2}^*$	$U_{(1-\alpha/2)}^*$
1%	-1,8007	0,6577	-2,58	2,58	-1,9303	3,7788
5%	-1,5049	0,4810	-1,96	1,96	-1,7294	2,2010
10%	-1,2990	0,4057	-1,65	1,65	-1,5473	1,4487

De acordo com o resultado obtido pelo teste para a precisão ϕ e também notado pelos gráficos, observamos a necessidade de modelarmos a dispersão. Assim, primeiramente, considerando as mesmas covariáveis da média para a modelagem da dispersão, ajustamos o modelo de regressão beta com dispersão variável para esses dados. Considerando, a modelagem da dispersão, temos o seguinte modelo:

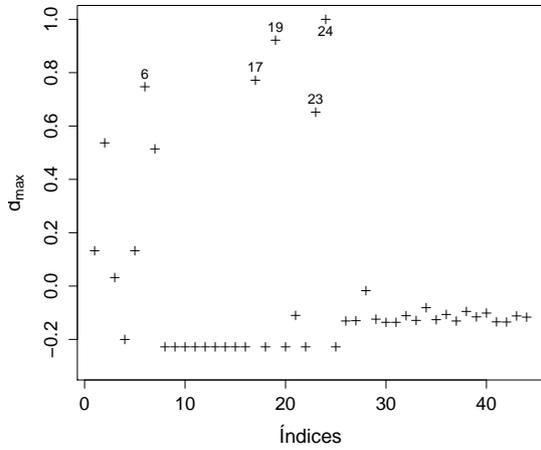
$$h(\phi_i) = \gamma_1 + \gamma_2 z_{i2} + \gamma_2 z_{i3} + \gamma_2 z_{i4},$$

Tabela 7.8: Mudanças relativas: Modelo 2.

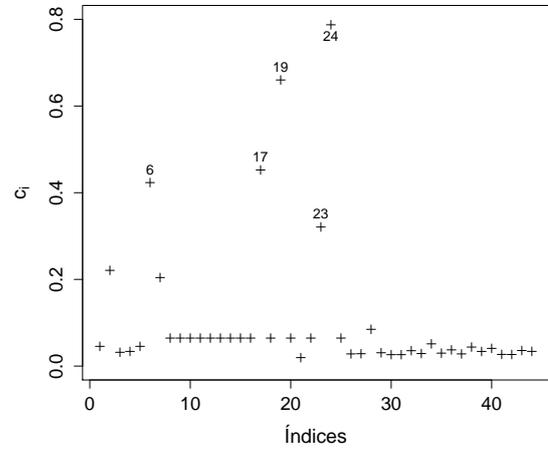
Observação \ Parâmetro	β_1	β_2	ϕ
6	3,29%	4,21%	8,95%
8	1,21%	1,65%	0,47%
17	3,39%	4,33%	9,41%
19	3,98%	5,05%	12,41%
23	2,92%	3,76%	7,18%
24	4,29%	5,43%	14,06%
6, 8, 17, 19, 23 e 24	23,21%	29,73%	131,37%

em que $z_{ij} = x_{ij}$ para $i = 1, 2, \dots, n$, $j = 1, 2, 3, 4$ e $h(\cdot)$ é a função de ligação log, ou seja, $g(\mu_i) = \log(\mu_i)$.

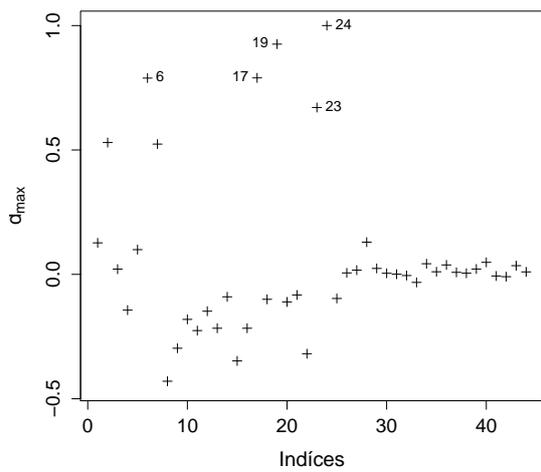
Figura 7.3: Gráficos de influência local (a) conjunta (b) individual, considerando o esquema de ponderação de casos e perturbação no parâmetro ϕ , respectivamente.



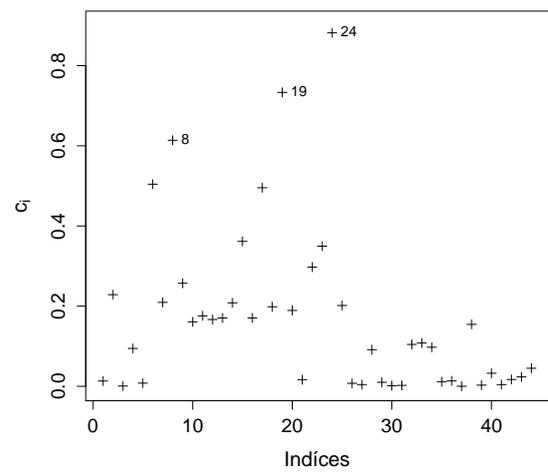
(a)



(b)



(c)

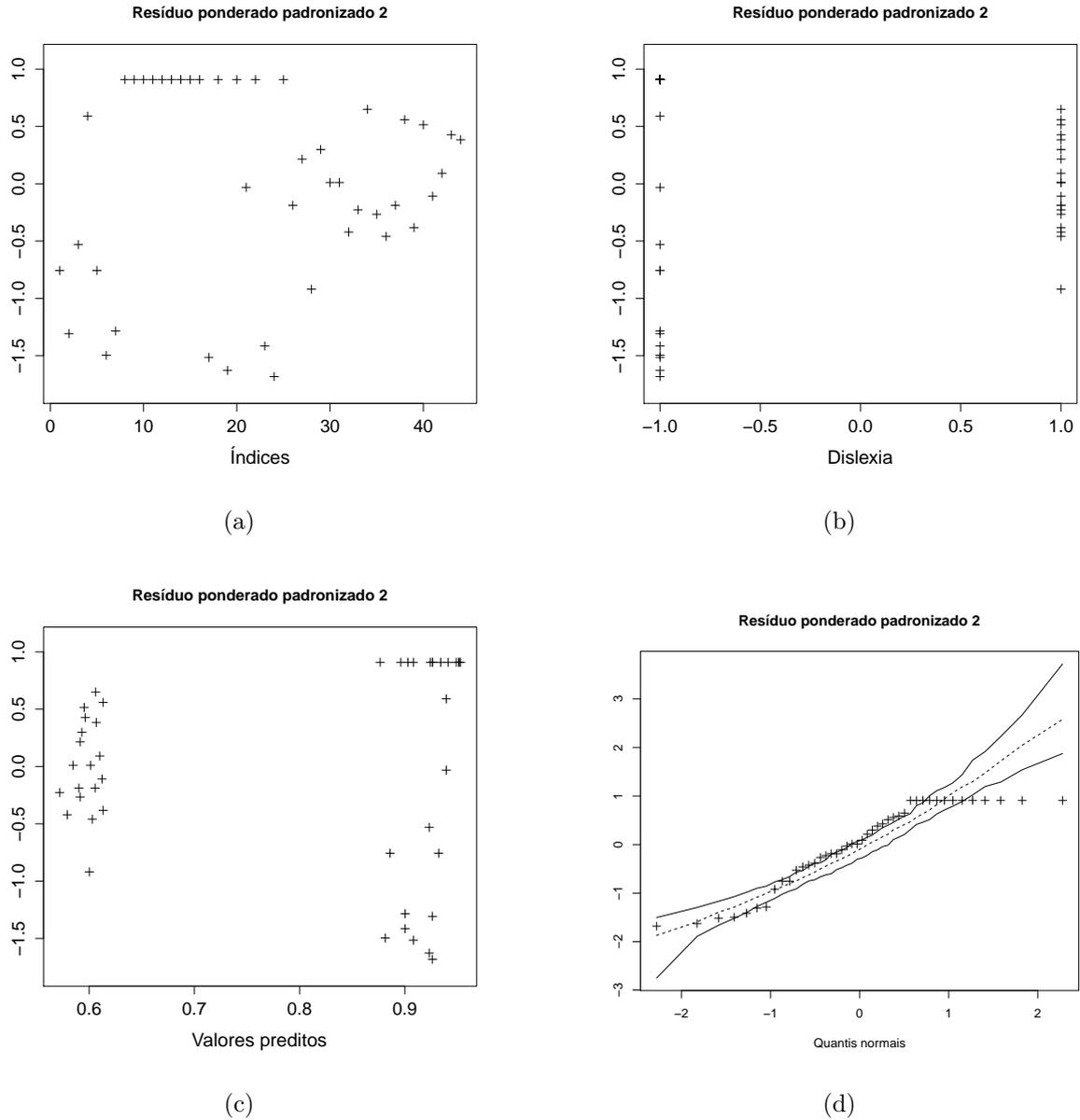


(d)

Fonte: A autora.

As estimativas dos parâmetros para esse modelo, bem como, os valores- p do teste de significância dos parâmetros estão apresentadas na Tabela 7.9. Vemos que a estimativa do grau de heterogeneidade λ é grande, $\hat{\lambda} = 612,1909$, como esperado. Além disso, de acordo com os valores- p , temos apenas o parâmetro referente a interação não foi significativo. Em seguida, ajustamos o mesmo modelo, mas dessa vez desconsiderando a interação. As respectivas estimativas dos parâmetros para esse último modelo, bem como, os valores- p do teste de significância dos parâmetros estão apresentadas na Tabela 7.10.

Figura 7.4: Gráficos de resíduos do modelo, juntamente com o envelope.



Fonte: A autora.

Percebemos que a modelagem da dispersão fez com que as covariáveis QI e interação fossem significativas para a modelagem da média, como indicado por Smithson e Verkuilen (2006). De acordo com os valores- p , todos os parâmetros são significativos. Em seguida, avaliamos a qualidade do ajuste, utilizando o teste apresentado na Seção 3.2.2. Obtivemos $U_n = -1,5737$, como também, na Tabela 7.11 temos os valores críticos empíricos da estatística U_n ($U_{\alpha/2}$ e $U_{(1-\alpha/2)}$) considerando os níveis nominais 1%, 5% e 10%.

Tabela 7.9: Estimativas dos parâmetros para o modelo de regressão beta com dispersão variável com interação.

Parâmetro	Estimativa	Valor- p
β_1	1,0187	0,0000
β_2	-0,6378	0,0001
β_3	0,6897	0,0064
β_4	-0,7760	0,0025
γ_1	3,0401	0,0000
γ_2	1,7675	0,0000
γ_3	1,4369	0,0045
γ_4	-0,6109	0,1970
λ	612,1909	-

Tabela 7.10: Estimativas dos parâmetros para o modelo de regressão beta com dispersão variável sem interação.

Parâmetro	Estimativa	Valor- p
β_1	1,1232	0,0000
β_2	-0,7416	0,0000
β_3	0,4864	0,0003
β_4	-0,5813	0,0000
γ_1	3,3044	0,0000
γ_2	1,7466	0,0000
γ_3	1,2291	0,0000
λ	454,7092	-

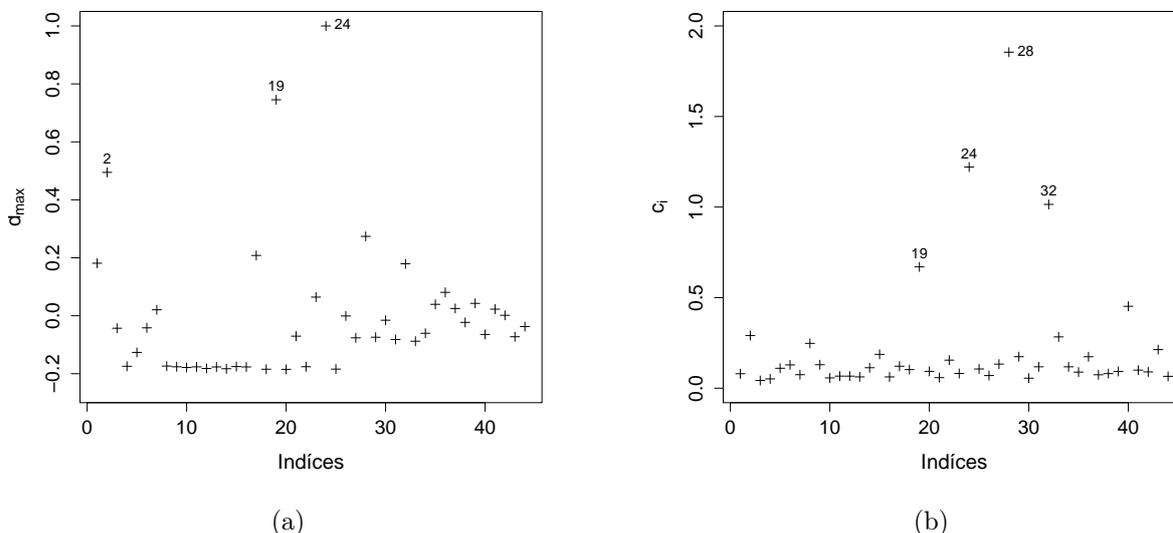
Tabela 7.11: Quantis empíricos e assintóticos para a estatística U_n para o modelo de regressão beta com dispersão variável.

Nível	$U_{\alpha/2}$	$U_{(1-\alpha/2)}$	$z_{\alpha/2}$	$z_{(1-\alpha/2)}$
1%	-4,8461	0,4462	-2,58	2,58
5%	-3,8630	0,2746	-1,96	1,96
10%	-3,3059	0,1416	-1,65	1,65

Logo, percebemos a partir das informações apresentadas nas tabelas, através dos quantis empíricos, que não rejeitamos a hipótese de boa adequação do modelo de regressão beta com dispersão variável aos níveis de significância usuais. Logo, temos que os resultados obtidos pelo teste proposto, utilizando os quantis empíricos obtidos via *bootstrap* estão de acordo com as conclusões feitas por Smithson e Verkuilen (2006). Nas Figuras 7.5 (a) e 7.5 (b), temos os gráficos de influência local conjunta e individual, respectivamente, considerando o esquema de perturbação ponderação de casos. Identificamos as observações 2, 19 e 24 como possivelmente influentes. Calculamos as mudanças relativas considerando essas observações e apresentamos na Tabela 7.12.

Quando verificamos as observações individuais, não notamos mudança relativa significa-

Figura 7.5: Gráficos de influência local conjunta e individual, considerando o esquema de ponderação de casos.



Fonte: A autora.

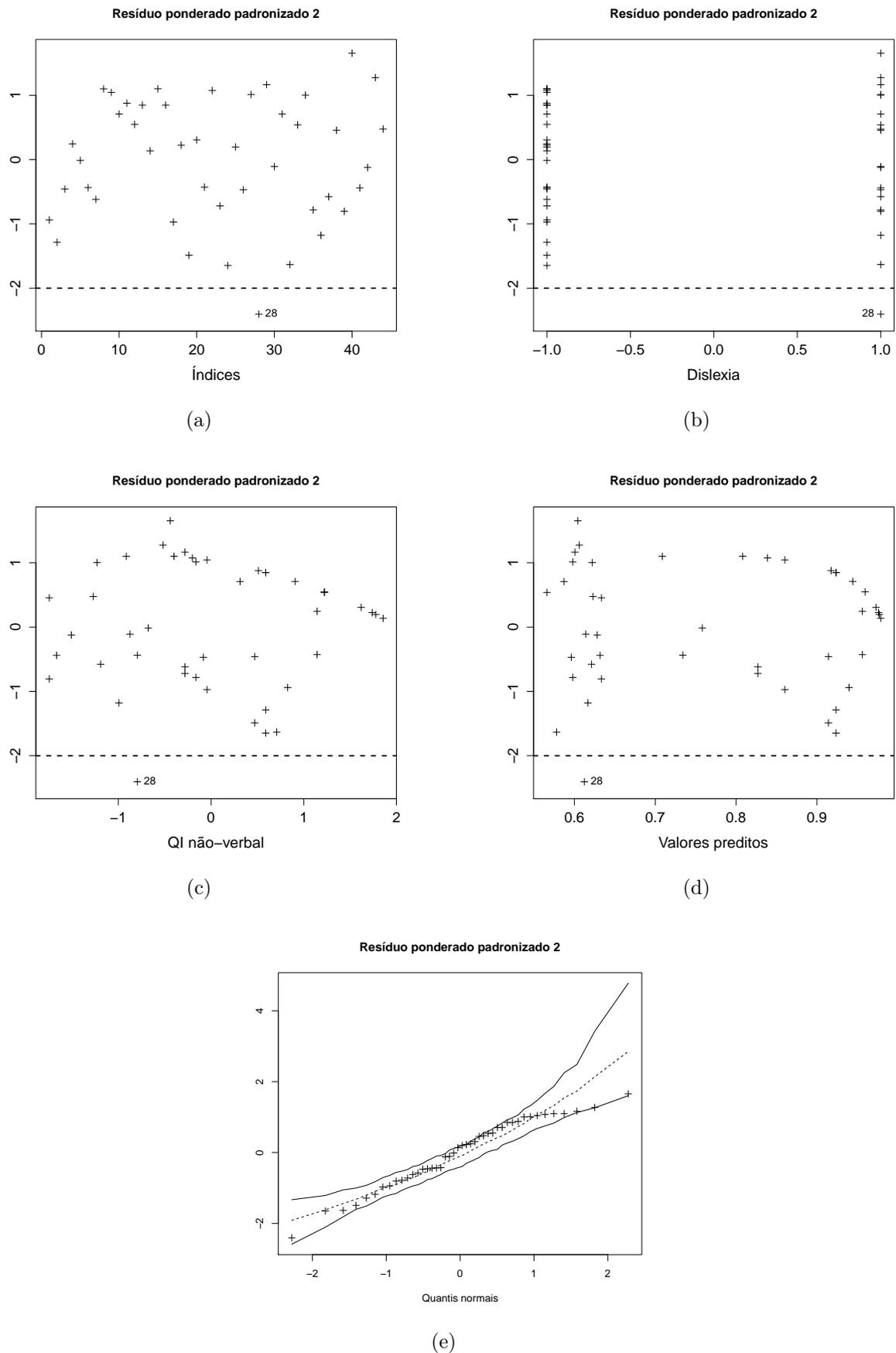
Tabela 7.12: Mudanças relativas para o modelo com dispersão variável.

Observação \ Parâmetro	β_1	β_2	β_3	β_4	γ_1	γ_2	γ_3
2	4,015%	6,064%	1,365%	1,266%	1,631%	1,632%	3,897%
19	6,428%	9,718%	0,555%	0,356%	2,808%	4,036%	3,417%
24	7,390%	11,125%	4,979%	4,511%	4,503%	4,197%	12,017%
2, 19 e 24	25,231%	38,093%	2,829%	2,980%	14,496%	17,697%	29,314%

tiva nos parâmetros do modelo. Em contrapartida, conjuntamente, as observações causam mudanças relativas significativas nos parâmetros β_1 , γ_1 e γ_2 . Na Figura 7.6 temos os gráficos de resíduos para o último modelo considerado, juntamente com o envelope, vemos que apenas a observação 28 (não mais a observação 8) aparece com resíduo alto. Notemos, a partir desses gráficos, que ao realizar a modelagem da dispersão, os resíduos do modelo não apresentam mais tendência, como também, a partir do envelope, não vemos indícios de falta de qualidade de ajuste.

Logo, a partir das análises aqui apresentadas, vemos que ambos os testes propostos conseguiram detectar más especificações graves. Em particular, notamos que o teste para a precisão ϕ se mostrou eficiente, no sentido de indicar a necessidade de modelagem da dispersão dos dados quando era necessário.

Figura 7.6: Gráficos dos resíduos do modelo com dispersão variável, juntamente com o envelope.



Fonte: A autora.

Capítulo 8

Considerações Finais

Nesse trabalho, desenvolvemos um teste baseado em influência local, considerando o esquema de ponderação de casos para os modelos de regressão beta com dispersão fixa e variável. O teste desenvolvido tem como objetivo detectar o grau de discrepância entre o modelo suposto e o subjacente do qual os dados são gerados. Foram apresentados diversos estudos de simulação de Monte Carlo, com o intuito de avaliar o desempenho do teste, os quais, após um melhoramento baseado na metodologia bootstrap para a obtenção dos quantis empíricos da estatística de teste, resultaram em ótimos tamanhos de teste para a grande maioria dos cenários apresentados, como também, bons poderes, pois o mesmo consegue detectar na maioria das situações as más especificações introduzidas: omissão de covariáveis da média e da dispersão e variável resposta simplex. Além disso, propusemos outro melhoramento nesse teste, dessa vez, através do uso da variância bootstrap o que produziu resultados mais coerentes. Ainda, propomos um novo teste, também com base em influência local, mas considerando o esquema de perturbação no parâmetro precisão no modelo de regressão beta com dispersão fixa. O último teste tem como objetivo detectar a necessidade da modelagem da dispersão em situações onde foi considerada dispersão fixa. Tal teste também foi avaliado através de estudos de simulação, e mostrou-se eficaz, no sentido de detectar más especificações do modelo. Por fim, aplicamos a teoria desenvolvida a um conjunto de dados reais, que evidenciou o bom desempenho do testes.

Referências

- ANDREWS, D. W. K. Generic uniform convergence. **Economet. Theory**, 8, 241-57, 1992.
- BELSLEY, D. A.; EDWIN K.; ROY. E. W. **Regression Diagnostics: Identifying Influential Data and Sources of Collinearity**. New York: John Wiley and Sons, 1980.
- BRANSCUM A. J.; JOHNSON W. O.; THURMOND M. C. Bayesian beta regression: Applications to household expenditure data and genetic distance between foot-and-mouth disease viruses. **Australian and New Zealand Journal of Statistics**, 49, 287-301, 2007.
- CARRASCO, J. M. F.; FERRARI, S. L. P.; ARELLANO-VALLE, R. B. Errors-in-variables beta regression models. **Journal of Applied Statistics**, 41, 1530-1547, 2014.
- CHIEN, L. C. Diagnostic plots in beta-regression models. **Journal of Applied Statistics**, 38, 1607-1622, 2011.
- CHIEN, L. C. Multiple deletion diagnostics in beta regression models. *Computational Statistics*. DOI: 10.1007/s00180-012-0370-9, 2012.
- COOK, R. D. Detection of influential observations in linear regression. **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.
- CRIBARI-NETO, F.; LIMA, L. B. A Misspecification Test for Beta Regressions. **Technical report**, 2007.
- 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**. Chapman e Hall, New York, 1993.
- ESPINHEIRA, P. L. Regressão Beta. Tese de Doutorado em Estatística. Instituto de Matemática e Estatística. Universidade de São Paulo, 2007.
- ESPINHEIRA, P. L., FERRARI, S.L.P.; CRIBARI-NETO, F. Influence Diagnostics in Beta Regression. **Computational Statistics e Data Analysis**, 52(9), 4417-4431, 2008a.
- ESPINHEIRA, P. L.; FERRARI, S.L.P.; CRIBARI-NETO, F. (On Beta Regression Residuals. **Journal of Applied Statistics**, 35(4), 407-419, 2008b.
- FALLS, L., W. The Beta Distribution: A Statistical Model for World Cloud Cover. **Journal of Geophysical Research**, 79 1261-1264, 1974.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta Regression for Modelling Rates and Proportions. **Journal of Applied Statistics**, 31(7), 799-815, 2004.
- GRAHAM, V. A.; HOLLANDS, K. G. T. Method to generate synthetic hourly solar radiation globally. **Solar Energy**, 44, 333-341, 1990.
- HARROP-WILLIAMS, K. Random nature of soil porosity and related properties. **Journal of Engineering Mechanics**, 115, 1129-1133, 1989.
- HOAGLIN, D. C.; WELSCH, R.E. The hat matrix in regression and ANOVA. **The American Statistician**, 32, 17-22, 1978.
- JANARDAN, K. G.; PADMANABHAN, G. Double bounded beta distribution for hydrologic variables. **Proc. 17th Annual Pittsburgh Conference** (parte 3), 17, 1107-1111, 1986.
- JOHNSON, N.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**. 2a ed. New York: John Wiley and Sons, 1995.
- LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. **Biometrics**, 54, 570-582, 1998.
- KIESCHNICK R.; MCCULLOUGH B.D. Regression analysis of variates observed on (0,1): percentages, proportions and fractions. **Statistical Modelling**, 3, 193-213, 2003.
- KOPPES, L.J.; GROVER, N.B. Relationship between size of parent at cell division and relative size of progeny in Escherichia Coli. **Archives of Microbiology**, 157, 402-405, 1992.
- LEHMANN, E. L.; CASELLA, E. **Theory of Point Estimation**. Springer-Verlag, New York, 1998.

- MCCULLACH, P.; NELDER, J. A. **Generalized Linear Models**. Second ed. London: Chapman and Hall, 1989.
- MCNALLY, R. J. Maximum likelihood estimation of the parameters of the prior distributions of three variables that strongly influence reproductive performance in cows. **Biometrics**, 446, 501-514, 1990.
- MILYUTIN, E. R.; YAROMENKO, Y. I. Statistical characteristics of atmospheric transparency index over tilted routes. **Meteorologiya i Gidrologiya**, 12, 72-76, 1991.
- OSPINA, R.; FERRARI, S. L. P. A general class of zero-or-one inflated beta regression models. **Computational Statistics and Data Analysis**, 56, 1609-1623, 2012.
- PAMMER, K.; KEVAN, A. **The contribution of visual sensitivity, phonological processing and non-verbal IQ to children's reading**. Unpublished manuscript, The Australian National University, Canberra, 2004.
- PAULA, G. A. (2010). **Modelos de Regressão: com apoio computacional**. IME-USP.
- PEREIRA, G. H. A.; BOTTER, D.A.; SANDOVAL, M.C. The truncated inflated beta distribution. **Communications in Statistics - Theory and Methods**, 41, 907-919, 2012.
- PEREIRA, T. L.; CRIBARI-NETO, F. Detecting model misspecification in inflated beta regressions. **Communications in Statistics - Simulation and Computation**, 2014.
- POLLARD, D. **Empirical Processes: Theory and Applications**, CBMS Regional Conference Series in Probability and Statistics, 2. Hayward, CA: Institute of Mathematical Statistics, 1990.
- 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.
- RAO, C. R. **Linear Statistical Inference and Its Applications**, 2nd ed. New York: Wiley, 1973.
- RIZZO, L. **Statistical Computing With R**. Boca Raton: Chapman e Hall/CRC, London, 2008.
- ROCHA, A. V.; SIMAS, A. B. Influence Diagnostics in a General Class of Beta Regression Models. **Test**, 20, 95-119, 2011.

- SIMAS, A. B.; BARRETO-SOUZA, W.; ROCHA, A.V. Improved Estimators for a General Class of Beta Regression Models. **Computational Statistics e Data Analysis**, 54(2), 348-366, 2010.
- SMITHSON, M.; VERKUILEN, J. A better lemon-squeezer? Maximum likelihood regression with beta-distributed dependent variables. **Psychological Methods**, 11, 54-71, 2006.
- ST. LAURENT, R. T.; COOK, R. D. Leverage and superleverage in nonlinear regression. **Journal of the American Statistical Association**. 87, 985-990, 1992.
- SULAIMAN, M. Y.; HLAING, O. W. M.,.; WAHAB, M. A.; ZAKARIA, A. Application of beta distribution model to Malaysian sunshine data. **Renewable Energy** 18, 573-579, 1999.
- WHITE, H. Maximum likelihood estimation of misspecified models. **Econometrica**, 50, 1-25, 1982.
- WILEY, J. A.; HERSCHOKORU, S. J.; PADIAU, N. S. Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penilevaginal intercourse. **Statistics in Medicine**, 8, 93-102, 1989.
- ZHU, H.; ZHANG, H. A diagnostic procedure based on local influence. **Biometrika**, 91(3) , 579-589, 2004.

Apêndice A

Teorema 3.1.1

Prova Teorema 3.1.1 Parte (1):

Sabemos de (1.3) que

$$\begin{aligned}
 \text{tr}(\mathbf{H}_{\text{LD}(\omega_0)}) &= \text{tr}(2\Delta^\top (-\ddot{\ell}_{\theta\theta}^{-1})\Delta)|_{\omega=\omega_0, \theta=\hat{\theta}} \\
 &= 2\text{tr}(-\ddot{\ell}_{\theta\theta}^{-1} \Delta\Delta^\top)|_{\omega=\omega_0, \theta=\hat{\theta}} \\
 &= 2\text{tr} \left\{ \left(-\frac{\ddot{\ell}_{\theta\theta}}{n} \right)^{-1} \frac{\Delta\Delta^\top}{n} \right\}_{\omega=\omega_0, \theta=\hat{\theta}}.
 \end{aligned} \tag{A.1}$$

Agora, note a partir de (A.1) que basta provarmos que

$$\begin{aligned}
 &(-\ddot{\ell}_{\theta\theta}/n) \xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}_*) \quad \text{e} \quad \Delta\Delta(\hat{\boldsymbol{\theta}})/n \xrightarrow{p} \mathbf{B}(\boldsymbol{\theta}_*) \Leftrightarrow \\
 &\left| -\frac{\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - \mathbf{A}(\boldsymbol{\theta}_*) \right| \xrightarrow{p} 0 \quad \text{e} \quad \left| \frac{\Delta\Delta(\hat{\boldsymbol{\theta}})}{n} - \mathbf{B}(\boldsymbol{\theta}_*) \right| \xrightarrow{p} 0.
 \end{aligned}$$

Assim, temos

$$\begin{aligned}
 \left| -\frac{\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - \mathbf{A}(\boldsymbol{\theta}_*) \right| &= \left| -\frac{\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - \mathbf{A}(\boldsymbol{\theta}_*) - \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}}))}{n} + \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}}))}{n} \right. \\
 &\quad \left. - \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} + \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} \right| \\
 &\leq \left| -\frac{\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}})}{n} - \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}}))}{n} \right| + \left| \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\hat{\boldsymbol{\theta}}))}{n} - \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} \right| \\
 &\quad + \left| \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} - \mathbf{A}(\boldsymbol{\theta}_*) \right|.
 \end{aligned} \tag{A.2}$$

Da Suposição 5, temos que

$$\left| \frac{\mathbf{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} - \mathbf{A}(\boldsymbol{\theta}_*) \right| \rightarrow 0.$$

Segundo Zhu & Zhang (2004) utilizando a Suposição 3, o segundo termo em (A.2) é suficientemente pequeno quando $n \rightarrow \infty$ utilizando o Lema 2 (a) apresentado em Andrews (1992). Para o último termo que falta, segundo os autores sob as Suposições 1 e 4 garatimos que

$$\sup_{\theta \in \Theta} |n^{-1}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}) - n^{-1}\mathbb{E}(-\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta})))| \xrightarrow{p} 0. \quad (\text{A.3})$$

Segundo Zhu & Zhang (2004) o resultado (A.3) utilizando Teorema 8.2 de Pollard (1990) implica na convergência em probabilidade para zero deste termo.

Prova Parte (2): Temos que

$$\left(-\frac{\mathbb{E}(\ddot{\ell}_{\theta\theta})}{n} \right)_{ij} = -\frac{\mathbb{E}(\sum_{k=1}^n \partial_i \partial_j \ell_k(\boldsymbol{\theta}_*))}{n},$$

como também,

$$\left(\frac{\mathbb{E}(\Delta\Delta^\top(\boldsymbol{\theta}_*))}{n} \right)_{ij} = \frac{\mathbb{E}(\sum_{k=1}^n \partial_i \ell_k(\boldsymbol{\theta}_*) \partial_j \ell_k(\boldsymbol{\theta}_*))}{n}.$$

Veja que

$$\begin{aligned} \mathbb{E} \left(\sum_{k=1}^n \partial_i \partial_j \ell_k(\boldsymbol{\theta}_*) \right) &= -\sum_{k=1}^n \int [\partial_i \partial_j \log(g(y_k; \boldsymbol{\theta}_*))] g(y_k; \boldsymbol{\theta}_*) dy_k \\ &= -\sum_{k=1}^n \int [\partial_i \partial_j \log(g(y_k; \boldsymbol{\theta}_*))] g(y_k; \boldsymbol{\theta}_*) dy_k |_{\theta=\theta_*} \\ &= -\sum_{k=1}^n \int \partial_i \left(\frac{1}{g(y_k; \boldsymbol{\theta})} \partial_j g(y_k; \boldsymbol{\theta}) \right) g(y_k; \boldsymbol{\theta}_*) dy_k |_{\theta=\theta_*} \\ &= -\sum_{k=1}^n \int \frac{[\partial_i (\partial_j g(y_k; \boldsymbol{\theta})) g(y_k; \boldsymbol{\theta}) - \partial_j g(y_k; \boldsymbol{\theta}) \partial_i g(y_k; \boldsymbol{\theta})]}{[g(y_k; \boldsymbol{\theta})]^2} g(y_k; \boldsymbol{\theta}) dy_k |_{\theta=\theta_*} \\ &= \left(-\sum_{k=1}^n \int \partial_i \partial_j g(y_k; \boldsymbol{\theta}) dy_k + \sum_{k=1}^n \int \frac{\partial_j g(y_k; \boldsymbol{\theta}) \partial_i g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} \right) \Big|_{\theta=\theta_*}. \end{aligned}$$

Pela Suposição 3 e desde que as $B(y_k)$ que dominam os módulos das derivadas parciais de $\ell_k(\boldsymbol{\theta})$ sejam integráveis, pelo Teorema da Convergência Dominada, temos que

$$\begin{aligned} \sum_{k=1}^n \int \partial_i \partial_j g(y_k; \boldsymbol{\theta}) dy_k &= \sum_{k=1}^n \partial_i \partial_j \int g(y_k; \boldsymbol{\theta}) dy_k \\ &= \sum_{k=1}^n \partial_i \partial_j (1) \\ &= 0. \end{aligned}$$

Então,

$$\begin{aligned}
 -\mathbb{E} \left(\sum_{k=1}^n \partial_i \partial_j \ell_k(\boldsymbol{\theta})_* \right) &= \sum_{k=1}^n \int \frac{\partial_j g(y_k; \boldsymbol{\theta}) \partial_i g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*} \\
 &= \sum_{k=1}^n \int \frac{\partial_i g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} \frac{\partial_j g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} g(y_k; \boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*} \\
 &= \sum_{k=1}^n \mathbb{E} \left(\frac{\partial_i g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} \frac{\partial_j g(y_k; \boldsymbol{\theta})}{g(y_k; \boldsymbol{\theta})} \right) \\
 &= \sum_{k=1}^n \mathbb{E} (\partial_i \ell_k(y_k; \boldsymbol{\theta}) \partial_j \ell_k(y_k; \boldsymbol{\theta})) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_*} \\
 &= \left(\frac{\mathbb{E}(\Delta \Delta^\top(\boldsymbol{\theta}_*))}{n} \right)_{ij}.
 \end{aligned}$$

Portanto,

$$\begin{aligned}
 -\frac{\mathbb{E}(\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} &= \frac{\mathbb{E}(\Delta \Delta^\top(\boldsymbol{\theta}_*))}{n} \Rightarrow \\
 \lim_{n \rightarrow \infty} \left(-\frac{\mathbb{E}(\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} \right) &= \lim_{n \rightarrow \infty} \frac{\mathbb{E}(\Delta \Delta^\top(\boldsymbol{\theta}_*))}{n}.
 \end{aligned}$$

Pela Suposição 5, temos

$$\lim_{n \rightarrow \infty} \left(-\frac{\mathbb{E}(\ddot{\ell}_{\theta\theta}(\boldsymbol{\theta}_*))}{n} \right) = \mathbf{A}(\boldsymbol{\theta}_*) \quad \text{e} \quad \lim_{n \rightarrow \infty} \frac{\mathbb{E}(\Delta \Delta^\top(\boldsymbol{\theta}_*))}{n} = \mathbf{B}(\boldsymbol{\theta}_*),$$

assim concluímos que $\mathbf{A}(\boldsymbol{\theta}_*) = \mathbf{B}(\boldsymbol{\theta}_*)$. Aplicando o Teorema 3.1.1 temos que

$$\begin{aligned}
 \lambda_0 &= 2\text{tr}(\mathbf{A}(\boldsymbol{\theta}_*)^{-1} \mathbf{B}(\boldsymbol{\theta}_*)) \\
 &= 2\text{tr}(\mathbb{I}_p) \\
 &= 2p.
 \end{aligned}$$

Note que, para provar o Teorema 3.1.1 não precisamos da suposição de identicamente distribuída como ocorre em White (1982).