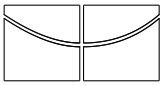


Universidade de Brasília

Instituto de Ciências Exatas

Departamento de Estatística

Dissertação de Mestrado

Modelo de Regressão Beta para Teste de Estresse em Risco de Crédito de Instituições Financeiras

por

Diogo Suzart Uzêda Picco

Orientador: Prof. Dr. Raul Yukihiro Matsushita

Junho de 2015

Diogo Suzart Uzêda Picco

Modelo de Regressão Beta para Teste de Estresse em Risco de Crédito de Instituições Financeiras

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidadede de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Universidade de Brasília

Brasília, Junho de 2015

Termo de Aprovação

Diogo Suzart Uzêda Picco

Modelo de Regressão Beta para Teste de Estresse em Risco de Crédito de Instituições Financeiras

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Data da defesa: 29 de junho de 2015

Orientador:

Prof. Dr. Raul Yukihiro Matsushita
Departamento de Estatística, UnB

Comissão Examinadora:

Prof. Dr. André Luiz Fernandes Cançado
Departamento de Estatística, UnB

Prof. Dr. Daniel Oliveira Cajueiro
Departamento de Economia, UnB

Brasília, Junho de 2015

Ficha Catalográfica

PICCO, DIOGO SUZART UZÊDA
Modelo de Regressão Beta para Teste de Estresse em Risco de Crédito
de Instituições Financeiras, (UnB - IE, Mestre em Estatística, 2015).

Dissertação de Mestrado - Universidade de Brasília. Departamento de Estatística
- Instituto de Ciências Exatas.

1. Regressão beta. 2. Teste de estresse. 3. Modelos de inadimplência. 4. Distribuição beta.

.

.

.

Agradecimentos

Ao meu orientador, Raul Yukihiro Matsushita, pela dedicação, comprometimento e disponibilidade. Principalmente pela paciência e clareza na orientação.

Aos Professores do Departamento de Estatística da Universidade de Brasília, pelo incentivo, interesse, ajuda e comentários sempre em prol do crescimento acadêmico do aluno.

Aos meus pais, pelo incentivo e estímulo recebido, pelo apoio incondicional e por vibrarem a cada conquista. Em especial, agradeço a minha mãe e avós, que apesar da distância sempre dedicaram muito amor e carinho, e principalmente tiveram muita paciência comigo.

Agradeço a minha irmã Diana Dayara Lopes e minha tia Valeska por estarem sempre por perto, pelo encorajamento e por acreditar em mim em todos os momentos.

Aos meus amigos e colegas do Departamento de Estatística, por tudo que compartilhamos nos últimos dois anos.

Aos meus amigos. Fioravante Mieto, Flávio Gonçalves, Guilherme Rocha e Raucelio Valdes por toda a colaboração, paciência, apoio, incentivo e sobretudo pela amizade.

Aos meus amigos do Banco do Brasil, em especial da Diretoria de Controles Internos, por todo o apoio, colaboração e incentivo que tive durante o decorrer do curso.

Sumário

Lista de Figuras

Lista de Tabelas

Resumo

A distribuição beta é muito versátil e flexível para modelar proporções, pois sua densidade pode assumir diferentes formas, dependendo dos valores dos parâmetros que a indexam. Nesse sentido, modelos com suporte na distribuição beta são candidatos naturais para modelagem da taxa de inadimplência das instituições financeiras. Neste trabalho propõe-se o uso de modelos com suporte na distribuição beta em testes de estresse para inadimplência de risco de crédito, por meio de estrutura de regressão como função de um conjunto de covariáveis macroeconômicas. As inferências desenvolvidas foram baseadas nas metodologias clássica e bayesiana. É apresentada discussão a respeito de sua definição, resultados de inferências, aplicação com dados reais em comparação a modelos tradicionais e simulações para avaliar a qualidade das aproximações utilizadas nas inferências sobre os parâmetros em amostras finitas.

Palavras Chave: *Regressão beta. Teste de estresse. Modelos de inadimplência. Distribuição beta.*

Abstract

The Beta distribution is a versatile and flexible model to describe proportions and rates. Its density function assumes different shapes depending on the values of its parameters. Thus, Beta-type models are natural candidates to model default rates of financial banks. In this work, as a function of macroeconomic covariates, Beta regression models are considered to fit defaults in stress tests for credit risk. The inferences are based on classical and Bayesian approaches. Monte Carlo studies were performed in order to assess the parameters estimates on finite samples. As an application with real data, results provided by Beta models are compared with traditional models.

key words: *Beta regression. Stress test. Default models. Beta distribution..*

Capítulo 1

Introdução

O mercado de crédito é extremamente competitivo entre as instituições financeiras, logo, a avaliação do risco de crédito é fundamental para o negócio bancário. A mensuração do risco impacta diversas áreas de uma instituição financeira, como aprovação do crédito, precificação dos produtos, avaliação de garantias e provisionamento de operações de crédito (ANTUNES et al., 2005). Uma correta avaliação dessas dimensões pode aumentar o lucro das instituições financeiras.

Muitos economistas apontam que as condições macroeconômicas influenciam os resultados das empresas, em que boas condições resultam em balanços fortes e uma desaceleração da economia é refletida em balanços fracos. Sendo assim, as condições macroeconômicas também afetam o risco de crédito.

Existe uma necessidade de se avaliar a sensibilidade do risco de crédito das instituições financeiras a mudanças que ocorrem na economia, a fim de prevenir instabilidades no mercado de crédito. Essa necessidade é ponto crítico e recebe atenção dos gestores de risco das instituições e dos reguladores. Segundo proposto no acordo de Basileia II, a alocação de capital deve ser feita observando-se o comportamento dos fatores de risco PD (Probabilidade de descumprimento) e LGD (Perda dado o descumprimento), durante o ciclo econômico, visando aumentar o capital em épocas de boas condições e retraí-lo em períodos de recessões econômicas, possibilitando maior poder de empréstimo e minimizando a crise.

Nesse contexto, as instituições financeiras devem garantir a própria perpetuidade e zelar pelo sistema financeiro, que sofre com o risco de crédito. Logo, seu gerencia-

mento faz-se necessário. Para a perpetuidade da instituição financeira e do sistema bancário frente às crises e cenários macroeconômicos adversos, os bancos devem estar preparados para as perdas inesperadas. O Bank for International Settlements (BIS), sob a ótica de Basileia II (BASEL COMMITTEE ON BANKING SUPERVISION, 2004), descreve as metodologias utilizadas e recomenda que sejam feitos testes de estresse no portfólio de crédito para se determinar o capital mínimo necessário para fazer frente a um cenário econômico pessimista.

O teste de estresse é uma técnica analítica que estima a sensibilidade do sistema financeiro e seu comportamento frente às mudanças nos fatores de risco que o afeta. Os testes de estresse podem fornecer informações sobre o comportamento macroeconômico desse sistema em casos de choques incomuns, porém plausíveis, permitindo observar suas vulnerabilidades (JONES et al., 2004).

A primeira fase da construção de um teste de estresse é a identificação das vulnerabilidades macroeconômicas do sistema, o que envolve análises quantitativas e qualitativas. As segunda e terceira fases são compostas pela construção dos cenários macroeconômicos que apontam os choques e quantificação do impacto. Nessas fases, modelos macroeconômicos são utilizados, pois fornecem, além da identificação e análise das relações entre o sistema financeiro e a economia real, uma previsão do futuro (JONES et al., 2004).

Para a construção de modelos com base em variáveis macroeconômicas, são utilizadas técnicas estatísticas. Atualmente, os modelos tradicionais têm como abordagem principal a conjugação entre dois deles: o primeiro, o modelo de inadimplência, que procura explicação de seu comportamento nas variáveis macroeconômicas; o segundo, o modelo de vetor autorregressivo das variáveis macroeconômicas que compõem o primeiro modelo (WILSON, 1997).

Em situações como esta, em que a variável resposta (inadimplência) é contínua e restrita ao intervalo (0,1), os modelos de regressão usual podem não ser apropriados, uma vez que é possível encontrar valores ajustados excedendo o limite inferior ou superior do intervalo de valores da variável resposta. Uma alternativa para contornar esse problema seria transformar a variável resposta de tal forma que esta assuma valores em toda a reta. Tal procedimento é realizado nos modelos tradicionais em uso por instituições financeiras. No entanto, essa abordagem apresenta desvantagens,

como, por exemplo, o fato de que os parâmetros do modelo não podem ser facilmente interpretados em termos da resposta original.

Ferrari e Cribari-Neto (2004) propuseram um modelo de regressão para situações em que a variável resposta Y (no caso, a inadimplência) é medida de forma contínua no intervalo (0,1). O modelo proposto é baseado na suposição de que a resposta tem distribuição beta, utilizando uma parametrização da família beta, que é indexada pela média e pela dispersão.

Neste trabalho propõe-se o uso de modelos com suporte na distribuição beta para situações em que se deseje modelar proporções ou taxas, por meio de estrutura de regressão, como função de um conjunto de covariáveis. A distribuição beta é muito versátil e flexível para modelar proporções, pois sua densidade pode assumir diferentes formas, dependendo dos valores dos parâmetros que a indexam.

O principal objetivo do trabalho é discutir o modelo de regressão beta como modelo alternativo para o uso em instituições financeiras, no que diz respeito aos testes de estresse. O modelo também é discutido quanto à sua definição, resultados de inferências, enfoques clássico e bayesiano, aplicação com dados reais e comparações com os modelos tradicionais. Adicionalmente, foi realizado estudo de simulação que avalia a qualidade das aproximações utilizadas nas inferências sobre os parâmetros em amostras finitas.

Organização da dissertação

A dissertação está subdividida em seis capítulos. No primeiro capítulo descreve-se o cenário atual e são apresentados aspectos relevantes sobre o tema inadimplência, bem como os objetivos do trabalho.

No segundo capítulo apresentam-se os tradicionais modelos em uso por instituições financeiras e discutem-se vantagens e desvantagens do uso de tais modelos.

No terceiro capítulo será mostrada, de forma sucinta, a distribuição beta, bem como suas propriedades, e a estimação de seus parâmetros pelo método de máxima verossimilhança. Será apresentado o modelo de regressão beta clássico, proposto por Ferrari e Cribari-Neto (2004), e o modelo de Paolino (2001). Serão tratadas as inferências sobre os parâmetros do modelo, seus intervalos de confiança e diagnósticos.

Neste capítulo também é discutido o modelo de regressão beta sob o enfoque bayesiano, envolvendo prioris vagas e informativas, estimação dos parâmetros, intervalos de credibilidade e algumas medidas de diagnóstico.

No capítulo quatro avaliam-se, por meio de simulações, os resultados assintóticos em amostras finitas com diferentes funções de ligação adotadas. Esse procedimento é realizado sob as óticas clássica e bayesiana.

No capítulo cinco é realizada a construção de modelos, utilizando a regressão beta com dados reais. Nesse capítulo comparam-se os modelos com suporte na distribuição normal e os de regressão beta.

No sexto capítulo apresentam-se as considerações finais e propostas de trabalhos futuros sobre o tema apresentado.

Capítulo 2

Teste de Estresse para Probabilidade de Descumprimento

Na última década, especialmente no contexto da crise financeira global, que destacou a necessidade de se desenvolver melhores abordagens metodológicas e práticas para identificação e monitoramento em bancos de risco, testes de estresse ganharam enorme popularidade entre os acadêmicos, autoridades de supervisão, organizações internacionais, como o FMI, e bancos. Definidos como uma abordagem metodológica para a análise da "vulnerabilidade potencial para eventos excepcionais, mas plausíveis"(BANK FOR INTERNATIONAL SETTLEMENTS, 2000; VIROLAINEN, 2004), os testes de estresse provaram ser uma ferramenta integral e mais abrangente para uma contínua avaliação de resistência dos bancos a vários choques.

Pilinko e Romancenco (2014) definem teste de estresse como uma forma ampla de análise de sensibilidade a eventos adversos, tais como uma queda drástica nas exportações, a deterioração da produtividade total dos fatores, a hiperinflação e outros choques macroeconômicos, bem como choques no setor bancário, sistemas financeiros etc.

Segundo Vazquez, Tabak e Souto (2012), o objetivo dos testes de estresse é fazer com que os riscos fiquem mais transparentes, avaliando-se as perdas potenciais de uma determinada carteira em condições anormais do mercado. Essas ferramentas são geralmente usadas por instituições financeiras como parte de seus modelos internos, sistemas de gestão e para informar as decisões relativas à assunção de riscos e alocação

de capital. Além disso, tornaram-se cada vez mais utilizadas pelos reguladores para avaliar a solidez dos sistemas financeiros sob seu controle.

Segundo Nishikawa (2014), o teste de estresse pode ser entendido como um conjunto de informações históricas e eventos imprevisíveis que aliado a técnicas apuradas de estatística vislumbra eventos futuros de perdas e ganhos, motivados por um cenário adverso ao normal, ou um choque de demanda ou oferta na economia.

Neste capítulo serão apresentados os seguintes assuntos: na seção 2.1, é apresentada a abordagem dos modelos Wilson, o modelo que é amplamente empregado em testes de estresse da probabilidade de descumprimento (PD). Na seção 2.2, são apresentadas as vantagens e desvantagens dessa abordagem e, por fim, na seção 2.3 é apresentada as contribuições com a utilização dos modelos de regressão com suporte na distribuição beta.

2.1 Modelos de Wilson

A metodologia proposta por Wilson (1997a; 1997b), chamada *Credit Portfolio-View*, foi o primeiro modelo de teste de estresse utilizado que teve explicitamente consideradas as variáveis macroeconômicas ao traçar a evolução do risco de crédito. Esta característica o torna uma ferramenta para o macro teste de estresse. A principal ideia por trás do modelo é a ligação das probabilidades de descumprimento a fatores macroeconômicos, além de reconhecer que a sensibilidade dos diferentes setores para agregar choques pode ser diferente. Uma vez estimado, ele é usado para simular a evolução das probabilidades de descumprimento em resposta aos choques macroeconômicos. A seguir, será descrita a metodologia de Wilson.

Valentiny-Endrész e Vásáry (2008) especificam que o modelo Wilson consiste nos seguintes passos: primeiro as probabilidades de descumprimento (PD) devem assumir função logística de índice y_t

$$CRT_{t,i} = \frac{1}{(1 + exp(-y_{i,t}))},$$

em que $CRI_{t,i}$ é a probabilidade de descumprimento no setor i no tempo t e $y_{i,t}$ é o índice específico macroeconômico no tempo t. A transformação logística é necessária para assegurar que a probabilidade de inadimplência esteja entre 0 e 1.

O índice $y_{i,t}$ pode ser interpretado como um indicador do estado geral da economia. A partir da probabilidade de descumprimento observada, pode-se obter o índice $y_{i,t}$ aplicando o inverso da função logística:

$$y_t = \ln\left(\frac{CRT_i}{1 - CRT_i}\right).$$

onde,

$$y \to \infty \Rightarrow CRT \to 1,$$
$$y \to -\infty \Rightarrow CRT \to 0.$$

O índice $y_{i,t}$, por sua vez, pode ser determinado por um número de variáveis macroeconômicas que influenciam o estado da economia. Mais especificamente, segue a seguinte forma:

$$y_{i,t} = \beta_{i,0} + \beta_{i,1}x_{1,t} + \beta_{i,2}x_{2,t} + \cdots + \beta_{i,n}x_{n,t} + \varepsilon_{i,t}.$$

Em que $\beta_i's$ são os coeficientes das regressões a serem estimadas para o setor $i = 1, \ldots, I$ e $x_{j,t}$ representa o j-ésimo fator macroeconômico no tempo t. $\varepsilon_{i,t}$ são os erros do setor em um tempo específico, que são serialmente independentes e normalmente distribuídos de forma idêntica (iid).

Valentiny-Endrész e Vásáry (2008) descrevem que um alto valor de $y_{i,t}$ implica um "bom estado "do ambiente macroeconômico, enquanto valores baixos implicam um "mau estado ". Essa especificação tem duas vantagens principais: permite que a sensibilidade a diferentes choques seja diversa entre os setores e permite que os termos dos erros individuais para capturar choques idiossincráticos sejam específicos do setor. Observe que em 2.1 o modelo pode ser interpretado como de segmento multi-fator específico, o qual descreve a "saúde "de determinado setor como uma "soma "ponderada das variáveis macroeconômicas. Os betas representam os "pesos "das variáveis específicas tomadas para cada segmento, com o termo de erro capturando um erro específico do setor. O modelo associa a saúde de um setor à evolução da economia. Apoia-se na observação de que durante "maus momentos "a probabilidade de inadimplência tende a ser maior do que em bons.

O próximo passo é a modelagem da dinâmica dos próprios fatores macroeconômicos. Wilson (1997) assume que todas as variáveis podem ser razoavelmente bem descritas

por um processo de AR(2).

$$x_{j,t} = \varphi_{j,0} + \sum_{k=1}^{m} \varphi_{j,k} Z_{j,t-k} + \sum_{l=0}^{q} \varphi_{p+l+1} \mu_{j,t-l}, \quad m > q,$$

em que φ_j é o conjunto de coeficientes da regressão que precisa ser estimado e $\mu_{j,t}$é o termo do erro para o qual a distribuição assumida é $N(\Phi, \Sigma)$

Assume-se o vetor de erros como:

$$(\mu_t, \epsilon_t) \sim N(0, \Sigma), \Sigma = \begin{pmatrix} \Sigma_{\mu,\mu} & \Sigma_{\mu,\epsilon} \\ \Sigma_{\epsilon,\mu} & \Sigma_{\epsilon,\epsilon} \end{pmatrix}.$$

De forma resumida, Wilson (1997) definiu o que seria um teste de estresse de um setor para o portfólio de risco de crédito, como abaixo:

$$CRT_t = \frac{1}{(1 + exp(-y_t))} \quad ou \quad y_t = \ln\left(\frac{CRT_i}{1 - CRT_i}\right), \tag{2.1}$$

$$y_t = \alpha_o + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varphi_0 Z_t + \sum_{j=1}^{q} \varphi_j Z_{t-j} + \mu_t, \tag{2.2}$$

$$Z_t = \mu + \sum_{k=1}^{m} A_k Z_{t-k} + \epsilon_t, m > q, \tag{2.3}$$

$$(\mu_t, \epsilon_t) \sim N(0, \Sigma), \Sigma = \begin{pmatrix} \Sigma_{\mu,\mu} & \Sigma_{\mu,\epsilon} \\ \Sigma_{\epsilon,\mu} & \Sigma_{\epsilon,\epsilon} \end{pmatrix}. \tag{2.4}$$

onde

y_t é a transformação logit de um indicador de risco de crédito observado $CRI_t(0,1)$,

Z_t é o vetor de variáveis macroeconômicas no tempo t,

u_t é um erro normal, homocedástico e independente em relação às informações passadas,

ϵ_t é um ruído branco normal,

CRI_T atrasos maiores que 90 dias.

A distribuição conjunta (μ_t, ε_t) representa a ligação do modelo macroeconômico e o de inadimplência. A hipótese mais comumente adotada é que $\Sigma_{\varepsilon,\mu} = \Sigma_{\mu,\varepsilon} = 0$.

Observa-se que as variáveis macroeconômicas seguem um tipo de estrutura de vetor autorregressivo (VAR), de acordo com a equação (2.3). E que os termos residual μ_t da equação (2.2) e ε_t da equação (2.3) apresentam distribuição conjunta normal de média zero e variância Σ.

2.2 Vantagens e Desvantagens das abordagens do tipo Wilson

Valentiny-Endrész e Vásáry (2008) descrevem que uma das grandes vantagens dos modelos que utilizam essa abordagem é que estes permitem mensurar a relação não linear entre o risco de crédito (probabilidade de descumprimento) e fatores macroeconômicos de forma simples. Outra vantagem é sua ampla disseminação; uma vez que modelos com aproximações gaussianas são largamente disseminados, seus pressupostos também já são conhecidos e aplicados, sua forma é bem definida.

Outro fato interessante na abordagem dos modelos Wilson é que eles utilizam uma transformação logística na variável dependente e essa não é muito boa para valores próximos às extremidades (0,1). Os pressupostos do modelo Wilson de normalidade e homocedasticidade não são atendidos pois sabe-se que regressões para respostas no intervalo (0,1) são tipicamente heterocedásticas.

Segundo Schechtman e Gaglione (2011) as estimativas de 2.1 a 2.4 raramente são discutidas na literatura de teste de estresse. Caso a possibilidade de $\Sigma_{\varepsilon,\mu} = \Sigma_{\mu,\varepsilon} \neq 0$, então Z_t deve ser tratada como endógena na equação 2.2, por causa de $cov\left(\mu_t, z_t\right) = \Sigma\left(\varepsilon, \mu\right)$, isso faz com que as estimativas de máxima verossimilhança (MLE) sejam mais complicadas que o habitual.

2.3 Considerações

Como visto, modelos que utilizam a abordagem do tipo Wilson apresentam algumas limitações, como a transformação da variável resposta, o pressuposto de nor-

malidade dos dados e a homocedasticidade, as distribuições de taxas são tipicamente assimétricas e, nesses casos, as aproximações gaussianas não são muito precisas. Uma abordagem alternativa que visa minimizar essas inconsistências seria a adoção de modelos com suporte em uma distribuição mais flexível, com diferentes formatos. Nesse contexto, os modelos com suporte na distribuição beta são candidatos naturais.

No próximo capítulo serão apresentados modelos com suporte na distribuição beta como alternativa de teste de estresse para a probabilidade de descumprimento.

Capítulo 3

Os modelos Beta

A distribuição beta é um modelo apropriado para descrever dados distribuídos no intervalo (0, 1)(KIESCHNICK e McCULLAGH, 2003), sendo, portanto uma candidata natural para representar taxas, razões e proporções. Sua flexibilidade de formas permite aplicações em diversas áreas como engenharia (BURY, 1999), agronomia (SILVA et al. 1998), hidrologia (JANARDAN e PADMANABHAN, 1986), geociências (SULAIMAN et al. ,1998), epidemiologia (WILEY e HERSCHOKOURO, 1989) e economia (BERGGREN et al., 2012). Neste trabalho, propõe-se a utilização da distribuição beta para descrever a inadimplência.

Em diversas situações uma variável aleatória beta pode depender de variáveis explicativas $x_1, \ldots, x_k$, que em análise de regressão são chamadas de variáveis regressoras ou covariáveis. Ferrari e Cribari-Neto (2004) propuseram um modelo de regressão cuja variável resposta Y assume valores no intervalo (a, b), em que $0 < a < b$ são valores conhecidos. Com base nesse modelo, tem-se uma expectativa condicional de Y em função de $x_1, \ldots, x_k$, ou seja $E(Y|x_1, \ldots, x_k)$. Nesse modelo a variável resposta Y_i, que segue uma distribuição beta, é parametrizada com base na média μ_i e variância $\dfrac{\mu_i(1-\mu_i)}{(\phi_i+1)}$, em que ϕ denomina-se parâmetro de precisão. O modelo de regressão beta é similar aos modelos lineares. Generalizações, extensões e variações dele têm sido propostas na literatura.

Entre os modelos de regressão beta mais gerais, Smithson e Verkuilen (2006) propuseram uma variação em que a média e o parâmetro de precisão seguem uma estrutura de regressão não linear (ou seja, ϕ não é constante). Simas et al.(2010)

consideram estruturas não lineares para μ e ϕ. Modelos que admitem inflação de zeros ou uns são tratados em Cook et al.(2008), Ospina e Ferrari (2010) e Ospina e Ferrari (2012). Há também uma extensão para o caso que permite truncamento da variável resposta para o intervalo $(a, 1)$ (PEREIRA et al.,2013). Modelos semi-paramétricos foram propostos por Branscum et al. (2007)

Modelos beta que incluem a dinâmica auto regressiva e de média móvel foram propostos por Rocha e Cribari-Neto (2008). Modelos bayesianos para processos beta auto regressivos podem ser vistos no trabalho de Casarin et al. (2012). O modelo dinâmico beta, com enfoque bayesiano, proposto por Silva et al. (2011) surgi para modelagem de séries temporais univariadas restritas ao intervalo (0,1). Rodrigues (2012) generaliza uma série temporal univariada, em que os dados observados referem-se a taxas ou proporções de uma dada categoria de respostas (modelo dinâmico beta) para um série temporal multivariada na qual k categorias de respostas são possíveis (modelo dinâmico dirichlet).

Aspectos acerca da inferência em grandes amostras e a análise de diagnóstico em modelos de regressão beta (MRB) foram discutidos por Espinheira et al. (2008,b). Aperfeiçoamentos na estimação pontual e intervalar são apresentados por Ospina et al. (2006). Rydlewski e Mielczarek (2012) demonstram a normalidade assintótica e consistência dos estimadores de máxima verossimilhança nos MRB.

Segundo Bayer (2011), para realizar inferências sobre os parâmetros do MRB pode-se utilizar o teste da razão de verossimilhança, escore ou Wald. A estatística do teste da razão de verossimilhança tem, sob certas situações particulares, distribuição limite qui-quadrado (χ^2). Contudo, em amostras pequenas a distribuição χ^2 pode fornecer uma aproximação pobre, implicando distorção do tamanho do teste. A utilização do ajuste de Skovgaard's para MRB é apresentada por Ferrari e Pinheiro (2011). Comparações acerca dos testes, inferências e ajustes das estatísticas através de *bootstrap* são observadas por Cribari-Neto e Queiroz (2012).

Bayer(2011) propõe e implementa a correção de Bartlett para melhorar o poder do teste da razão de verossimilhança em amostras pequenas. Simas et al. (2010) propuseram correções analíticas de viés para os estimadores de máxima verossimilhança, generalizando os resultados de Ospina et al. (2006). Ospina e Ferrari (2012b) apresentam métodos de correção do viés para estimadores de máxima verossimilhança em

modelos inflados de zero e um.

Uma discussão a respeito de modelagem da regressão beta na abordagem clássica no *software* R é apresentada com detalhes por Cribari-Neto e Zeileis (2010). Modelagem de regressão beta na abordagem bayesiana no *software* R é apresentada por Cepeda-Cuervo et al. (2014).

Este capítulo encontra-se divido da seguinte forma: A seção 3.1 apresenta sucintamente a família de distribuição beta e seus parâmetros. Na seção 3.2 é tratada a estimação dos parâmetros pelo método da máxima verossimilhança. A seção 3.3 trata a estimação dos parâmetros da distribuição beta com enfoque bayesiano. As seções 3.4 e 3.5 tratam do modelo de regressão beta clássico, em que se considera o método de estimação dos parâmetros do modelo proposto por Paolino (2001) e Ferrari e Cribari-Neto (2004), também apresenta uma discussão acerca dos aspectos inferências. Na seção 3.6 são apresentados técnicas de diagnósticos para os modelos de regressão beta. A seção 3.7 trata do modelo de regressão beta segundo a abordagem bayesina. Os parâmetros do modelo são estimados através de prioris vagas e informativas.

3.1 Distribuição Beta

A família de distribuições Beta é composta de todas as distribuições de probabilidade cuja função densidade de probabilidade depende dos parâmetros p, q e é dada por.

$$f\left(y;p,q\right)=\frac{\Gamma\left(p+q\right)}{\Gamma\left(p\right)\Gamma\left(q\right)}y^{p-1}\left(1-y\right)^{q-1}, \tag{3.1}$$

em que $0<y<1$, $p>0$ e $q>0$ são parâmetros da função densidade de probabilidade, $\Gamma\left(\cdot\right)$ é a função gama.

Um fato interessante na distribuição beta é:

$$y\sim Beta\left(p,q\right)\rightarrow\left(1-y\right)\sim Beta(q,p),$$

A função beta é muito flexível, como podemos observar na figura 3.1. Para diferentes parâmetros p e q a função apresenta diferentes formas

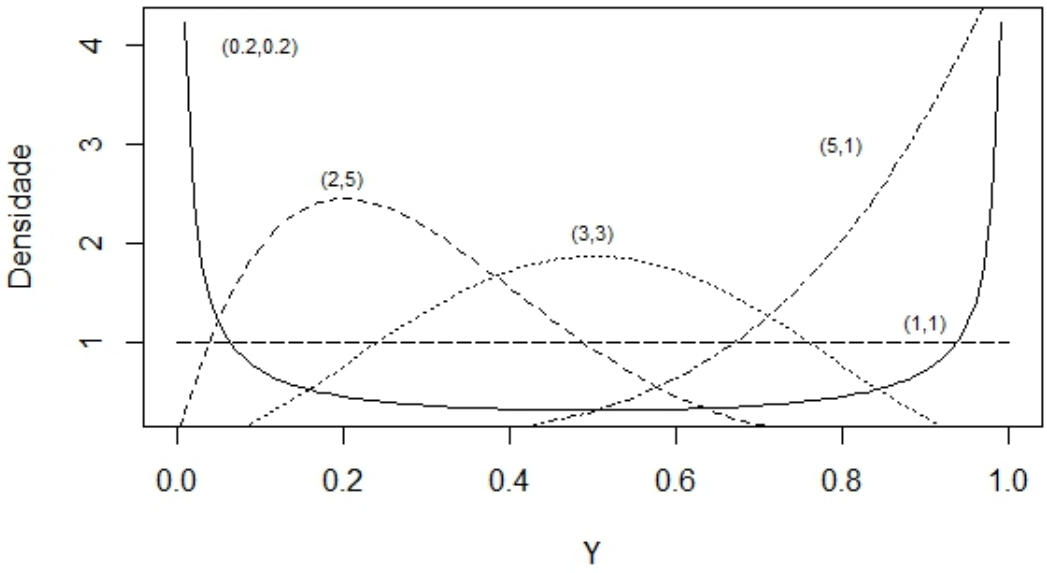


Figura 3.1: Densidade Beta para diferentes valores de (p,q)

A função de distribuição acumulada (f.d.a) de uma variável aleatória com distribuição Beta (p, q) é dada por:

$$\begin{aligned} F(y;p,q) &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^y y^{(p-1)} (1-y)^{(q-1)} dy \qquad 0 < y < 1 \\ &= \frac{B_Y(y;p,q)}{B(p,q)}, \end{aligned}$$

em que

$$B(p,q) = \frac{\Gamma(p)\Gamma(p)}{\Gamma(p+q)},$$

denomina-se função beta e a integral,

$$B(y;p,q) = \int_0^y z^{(p-1)} (1-z)^{(q-1)} dz,$$

é a função beta incompleta.

Seu z-ésimo momento em relação a origem é dado por,

$$\mu_z^k = \frac{B(p+z,q)}{B(p,q)} = \frac{\Gamma(p+z)\Gamma(p)}{\Gamma(p+q+z)\Gamma(p+q)},$$

Em particular o valor esperado da variável Y é dado por:

$$E(Y) = \frac{p}{p+q},$$

e a sua variância é dada por:

$$V(Y) = \frac{pq}{(p+q)^2 (p+q+1)}.$$

O coeficiente de variação de Y é dado por:

$$cv(Y) = \frac{\sqrt{V(Y)}}{E(Y)} = \sqrt{\frac{q}{p(p+q+1)}}$$

O coeficiente de assimetria, medida que caracteriza como e quanto a distribuição se afasta da condição de simetria é dado por:

$$A = \frac{2(q-p)}{p+q+2} = \sqrt{\frac{p+q+1}{pq}}.$$

A curtose, medida que procura caracterizar o formato da distribuição quanto ao achatamento é dado por:

$$C = \frac{3(q+q+1)\left[2(p+q)^2 + pq(p+q-6)\right]}{pq(p+q+2)(p+q+3)}.$$

Normalmente, em análise de regressão, o interesse reside em modelar a média das respostas bem como identificar precisão do modelo. Afim de se obter uma estrutura de regressão para a média das respostas, juntamente com o parâmetro de precisão. A distribuição beta foi reparametrizada com base em μ e ϕ. Sendo assim temos que

$$p = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right)\mu^2 = \left[\frac{(1-\mu)(\phi+1)}{(1-\mu)\mu} - \frac{1}{\mu}\right]\mu^2 = \phi\mu \tag{3.2}$$

e

$$q = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right)\mu(1-\mu) = \frac{\phi}{\mu}\mu(1-\mu) = \phi(1-\mu). \tag{3.3}$$

Sob essa nova parametrização, temos

$$f(y) = \frac{\Gamma(\phi)}{\Gamma(\phi\mu)\Gamma(\phi(1-\mu))} y^{\phi\mu-1}(1-y)^{\phi(1-\mu)-1}, \tag{3.4}$$

em que $0 < y < 1$, $0 < \mu < 1$ e $\phi > 0$.

$$E(Y) = \mu, \tag{3.5}$$

$$V(Y) = \frac{\mu(1-\mu)}{\phi+1}. \tag{3.6}$$

3.2 Estimação de parâmetros da distribuição beta

Seja $Y_1, \ldots, Y_n$ uma amostra aleatória simples (AAS) retirada de uma população beta com parâmetros p e q. Considerando as equações (3.2) e (3.3), p e q podem ser facilmente estimados pelo método dos momentos fazendo

$$\hat{p} = \hat{\phi}\bar{x}$$

e

$$\hat{q} = \hat{\phi} - \hat{p},$$

em que $\hat{\phi} = \dfrac{\bar{x}(1-\bar{x})}{\sigma^2} - 1$. No entanto, a estimação dos parâmetros da distribuição beta pode ser realizada pelo método da máxima verossimilhança, que produz estimadores com algumas propriedades estatisticamente interessantes. Em geral, os estimadores de máxima verossimilhança são consistentes, assintoticamente normais, assintoticamente eficientes e possuem a propriedade de invariância. Se $\hat{p}$ e $\hat{q}$ são estimadores de máxima verossimilhança, então $\hat{\mu} = \dfrac{\hat{p}}{\hat{p}+\hat{q}}$ e $\hat{\phi} = \hat{p} + \hat{q}$ também são estimadores de máxima verossimilhança.

Considerando o vetor de parâmetros é o $\theta = (p, q) \in \omega = \mathbb{R}_+ \times \mathbb{R}$, em que ω o espaço paramétrico, defini-se a função de verossimilhança como

$$\begin{aligned} L(\theta) &= L(p,q) \\ &= \prod_{j=1}^{n} \frac{y_j^{(p-1)}(1-y_j)^{(q-1)}}{B(p,q)} \\ &= [B(p,q)]^{-n} \prod_{j=1}^{n} y_j^{(p-1)} \prod_{j=1}^{n} (1-y_j)^{(q-1)}, \end{aligned}$$

e a função de log-verossimilhança é

$$\begin{aligned} l(\theta) &= \ln(L(\theta)) \\ &= -n\ln(B(p+q)) \\ &\quad + (p-1)\sum_{j=1}^{n}\ln(y_j) \\ &\quad + (q-1)\sum_{j=1}^{n}\ln(1-y_j). \end{aligned}$$

Uma forma de maximizar $l(p,q)$ é mediante suas derivadas parciais, $\frac{\partial l(p,q)}{\partial p} = 0$ e $\frac{\partial l(p,q)}{\partial q} = 0$, o que produz as equações de máxima verossimilhança

$$\begin{cases} \Psi(\hat{p}) - \Psi(\hat{p}+\hat{q}) = \frac{1}{n}\sum_{j=1}^{n}\ln(y_j). \\ \\ \Psi(\hat{q}) - \Psi(\hat{p}+\hat{q}) = \frac{1}{n}\sum_{j=1}^{n}\ln(1-y_j). \end{cases}$$

em que $\Psi(\hat{x})$, $x > 0$ é a função digama definida como

$$\Psi(\hat{x}) = \frac{d\ln(\Gamma(x))}{dx} = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Outra possibilidade é maximizar diretamente $l(p,q)$ com respeito a p e q com base em métodos numéricos como de Newton-Raphson, Score e Fisher. Sob condições gerais de regularidade (descritas em Sen e Singer, 1993), quando o tamanho da amostra é grande, os estimadores de máxima verossimilhança dos parâmetros têm distribuição aproximada dada por:

$$\begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix} \simeq N_2\left(\begin{pmatrix} p \\ q \end{pmatrix}; K^{-1}(p,q)\right),$$

em que $\hat{p}$ e $\hat{q}$ são os estimadores de máxima verossimilhança de p e q, respectivamente, e $K^{-1}(p,q)$ é a matriz de variância e covariância assintótica dos estimadores de máxima verossimilhança, dado por:

$$\begin{aligned} k^{-1}(p,q) &= n^{-1}\left[\Psi'(p)\Psi'(q) - \Psi'(p+q)\left\{\Psi'(p)+\Psi'(q)\right\}\right]^{-1} \\ &\times \begin{bmatrix} \Psi'(q) - \Psi'(p+q) & \Psi'(p+q) \\ \Psi'(p+q) & \Psi'(q) - \Psi'(p+q), \end{bmatrix} \end{aligned}$$

em que $\Psi'(x)$ é a função trigama definida como

$$\Psi'(x) = \frac{d\Psi(x)}{x} = \frac{d^2 \ln \Gamma(x)}{dx^2}.$$

3.3 Estimação da distribuição beta via abordagem bayesiana

Na abordagem clássica, os parâmetros desconhecidos utilizados são considerados fixos e todas as análises são baseadas nas informações contidas nos dados amostrais. Segundo alguns autores, esta abordagem foi adotada de forma quase unânime pelos estatísticos durante a primeira metade do século XX. Na abordagem bayesiana os parâmetros são vistos como variáveis aleatórias e não mais como constantes. Dessa forma a incerteza de um modelo dado θ é representado através de uma distribuição de probabilidade $P(\theta)$ sobre os possíveis valores do parâmetro desconhecido.

Seja θ uma quantidade de interesse desconhecida cujos possíveis valores são pertencentes ao conjunto Θ. O objetivo da inferência bayesiana pode ser a estimação de θ ou o teste de alguma hipótese envolvendo valores de θ. Um dos principais ingredientes para a realização de inferência bayesiana é a distribuição à posteriori, que representa o conhecimento a respeito de θ após a observação dos dados X.

A distribuição à posteriori é obtida através do teorema de Bayes, isto é, definido da seguinte forma. Suponha que $X' = (x_1, x_2, \ldots, x_n)$ é um vetor de n observações cuja distribuição de probabilidade $p(x|\theta)$ depende dos valores de k parâmetros $\theta' = (\theta_1, \theta_2, \ldots, \theta_k)$. Suponha que θ tenha uma distribuição de probabilidade $p(\theta)$. Então,

$$p(x|\theta)\, p(\theta) = p(x, \theta) = p(\theta|x)\, p(x)$$

Sendo os dados observados X, a distribuição condicional de θ é

$$p(\theta|x) = \frac{p(\theta)\, p(x|\theta)}{p(x)}. \tag{3.7}$$

Além disso podemos escrever

$$p(x) = E[p(x|\theta)] = c^{-1} = \begin{cases} \int p(x|\theta)\, p(\theta)\, d\theta, & \text{para } \theta \text{ contínuas.} \\ \sum p(x|\theta)\, p(\theta), & \text{para } \theta \text{ discreta.} \end{cases}$$

Onde o somatório, ou a integral, é tomado sobre o alcance admissível de θ, e onde $E[f(\theta)]$ é a esperança matemática de $f(\theta)$ com respeito a distribuição $p(\theta)$. Assim podemos escrever a expressão (3.7) alternativamente como

$$p(\theta|X) = cp(\theta)\, p(X|\theta), \tag{3.8}$$

A expressão (3.7), ou seu equivalente (3.8), é geralmente denominado teorema de Bayes. Nesta expressão, $p(\theta)$, nos diz o que é conhecido sobre θ sem conhecimento dos dados, é chamado distribuição a priori de θ. $p(\theta|x)$ nos diz o que é conhecido sobre θ através dos dados, é chamado de distribuição a posteriori. A quantidade c é uma constante normalizadora.

A distribuição à posteriori de um parâmetro θ contém toda a informação probabilística à respeito deste parâmetro. No entanto, algumas vezes é necessário resumir a informação contida na distribuição a posteriori através de alguns poucos valores numéricos. O caso mais simples é a estimação pontual de θ onde se resume a distribuição à posteriori através de um único número, $\hat{\theta}$

A função de verossimilhança de θ é a função que associa a cada θ a distribuição de probabilidade conjunta $p(\theta|x)$, isto é, $L(X|\theta) = p(X|\theta)$. A definição de verossimilhança não requer que os dados sejam observações de variáveis aleatórias independentes ou identicamente distribuídas. Além disso, fatores que dependem somente de X e não dependem de θ podem ser ignorados quando se escreve a função de verossimilhança já que eles não fornecem informação sobre a plausibilidade relativa de diferentes valores de θ.

Agora, com base nos dados x, $p(X|\theta)$ em (3.8) pode ser considerada como uma função não de X, mas de θ. Quando assim considerado, ela é chamada de função de verossimilhança de θ para X dado que pode ser escrita como $L(\theta|x)$. Deste modo podemos descrever (3.7) como

$$p(\theta|X) = L(\theta|X)\, p(\theta). \tag{3.9}$$

Em outras palavras, o teorema de Bayes diz que a distribuição de probabilidade a posteriori θ dado X é proporcional ao produto da distribuição a priori de θ ao da máxima verossimilhança de θ dado X.

distribuição a posteriori $\propto$ ***verossimilhança*** $\times$ ***distribuição a priori***.

3.3.1 Estimação dos parâmetros

Na inferência Bayesiana, ao se utilizar diferentes distribuições a priori, são requeridos diferentes desenvolvimentos do teorema de Bayes, isto é, para cada distribuição a priori considerada é necessário obter uma distribuição a posteriori.

Seja Y em que

$$Y \sim Beta\left(p, q\right).$$

Em que p e q são dados por (3.2) e (3.3) respectivamente. A função de verossimilhança é dada por

$$L\left(Y|\mu,\phi\right) = \prod_{j=1}^{n} f\left(y|\mu,\phi\right).$$

Sendo $f\left(y|\mu,\phi\right)$ dado por (3.4) e assumindo $p\left(\mu,\phi\right) = p\left(\mu\right)p\left(\phi\right)$, segue que

$$\begin{aligned} L\left(Y|\mu,\phi\right) &= \left[\Gamma\left(\phi\right)\right]^{n}\left[\prod_{j=1}^{n}\left[\Gamma\left(\mu\phi\right)\right]\right]^{-1}\left[\prod_{j=1}^{n}\left[\Gamma\left(\phi\left(1-\mu\right)\right)\right]\right]^{-1} \\ &\times \prod_{j=1}^{n}\left[y^{(\mu\phi-1)}\right]\prod_{j=1}^{n}\left[\left(1-y\right)^{[\phi(1-\mu)-1]}\right], \end{aligned}$$

$$\begin{aligned} p\left(\mu,\phi|Y\right) &= L\left(Y|\mu,\phi\right)p\left(\mu,\phi\right), \\ p\left(\mu,\phi|Y\right) &= L\left(Y|\mu,\phi\right)p\left(\mu\right)p\left(\phi\right). \end{aligned}$$

O interesse reside em obter uma distribuição condicional marginal para cada um dos parâmetros. A distribuição é obtida integrando a posteriori conjunta em relação aos demais parâmetros. Caso estejamos interessados em realizar inferências a respeito de μ sua posteriori é dada por

$$\begin{aligned} p\left(\mu|Y,\phi\right) &\propto \left[\Gamma\left(\phi\right)\right]^{n}\left[\prod_{j=1}^{n}\left[\Gamma\left(\mu\phi\right)\right]\right]^{-1}\left[\prod_{j=1}^{n}\left[\Gamma\left(\phi\left(1-\mu\right)\right)\right]\right]^{-1} \\ &\times \prod_{j=1}^{n}\left[y^{(\mu\phi-1)}\right]\prod_{j=1}^{n}\left[\left(1-y\right)^{[\phi(1-\mu)-1]}\right] \\ &\times p\left(\mu\right). \end{aligned}$$

Dependendo da priori assumida para μ, para gerar uma amostra de $\mu|Y,\phi$ será necessário utilizar o procedimento MCMC, mais precisamente o Metropolis Hastings, uma vez que a distribuição de $\mu|Y,\phi$, na maioria dos casos, será intratável analiticamente. As estimativas são dadas pela média das cadeias geradas a partir da distribuição a posteriori de $\mu|Y,\phi$. De forma análoga pode se obter a distribuição à posteriori de ϕ.

3.4 O Modelo regressão beta de Ferrari e Cribari-Neto

A classe de modelos de regressão beta permite descrever a média condicional de uma variável resposta $Beta\,(Y)$ em função de covariáveis $x_1,\ldots,x_n$, ou seja, $E\,(Y|x_1,\ldots,x_n)$. Para fins de modelagem mediante uma estrutura de regressão é conveniente adotar a forma reparametrizada da distribuição beta (equação (3.4)) sugerida por Ferrari e Cribari-Neto (2004). Nessa forma, a densidade beta é especificada por sua média μ e seu parâmetro de precisão ϕ (Ferrari e Cribari-Neto, 2004). Como a variável dependente se restringe ao intervalo $(0,1)$, sua média também se restringe a esse mesmo intervalo $(0<\mu<1)$. Para estabelecer uma relação entre as covariáveis $x_1,\ldots,x_n$ e o valor da variável resposta Y, define-se uma função $g\,(\cdot)$, monótona e duas vezes diferenciável, que permite mapear uma combinção linear definda em $\mathbb{R}$ para o intervalo $(0,1)$. Essa função que se denomina função de ligação, é aquela que estabelece a estrutura do modelo de regressão.

3.4.1 A estrutura do modelo e a função de ligação

Sejam $Y_1,\ldots,Y_n$ variáveis aleatórias independentes e suponha que cada y_i,$i=1,\ldots,n$, segue a distribuição beta reparametrizada, em que $Y_i\sim Beta\,(\mu_i,\phi)$ com média μ_i e parâmetro de precisão dado por ϕ, desconhecido e constante para todo i. Ferrari e Cribari-Neto (2004) propõem o modelo assumindo que uma função da média μ_i pode ser igualada ao preditor linear η_i, resultando em

$$g\,(\mu_i)=\sum_{j=1}^{k}\beta_j x_{ij}=\eta_i, \tag{3.10}$$

em que $\beta = (\beta_1, \ldots, \beta_k)^T \in \mathbb{R}^K$ é um vetor de parâmetros de regressão desconhecido,η_i é o preditor linear e $x_{i1}, \ldots, x_{ik}$ são observações de k variáveis $(k < n)$,cujo os valores são fixos e conhecidos. Consequentemente, a partir do modelo, tem-se que:

$$E(Y_i) = \mu_i = g^{-1}(\eta_i) \tag{3.11}$$

e

$$Var(Y_i) = \frac{V(g^{-1}(\eta_i))}{(1+\phi)}. \tag{3.12}$$

Note que a variância de Y_i depende de μ_i, o que implica que esta não é constante para todas as observações. Ou seja, o modelo de regressão beta é naturalmente heterocedástico, sendo mais adequado, portanto, a dados de taxas e proporções que, como foi anteriormente apontado, tem heterocedasticidade como característica inerente.

O problema agora é mapear (ligar) a combinação linear η_i (3.10) para o intervalo $(0, 1)$, típico de uma variável resposta Beta. Considerando $0 < \mu_i < 1$, as funções abaixo exemplificam possibilidades para esse mapeamento:

Função	Forma $g(\mu_i)$	distribuição relacionada
logito	$\log\left(\frac{\mu_i}{1-\mu_1}\right)$	logistica
Probito	$\Phi^{-1}(\mu_i)$	gaussiana
log-log	$-\log[-\log(\mu_i)]$	valor extremo
Complemento log-log	$\log\{-\log(1-\mu_i)\}$	valor extremo

em que $\mu_i = g^{-1}(\eta_i) = E(Y_i|x_i)$. Observe que g é a composta, $g(\mu_i(x_i))$, e por isso ela é chamada de função de ligação entre x_i e μ_i. A função (3.10) é chamada preditor linear.

Para o processo de estimação é necessário derivar $g(\mu_i)$. A lista abaixo mostra as derivadas obtidas para algumas função de ligação mais comuns.

- Logito: $g'(\mu_t) = \dfrac{\{1+\exp(\eta)\}^2}{\exp(\eta)}$,
- Probito:$g'(\mu_t) = \sqrt{2\pi}\exp\left\{\dfrac{\eta^2}{2}\right\}$,
- Complemento Log-Log: $g'(\mu_t) = \exp\{\exp(\eta) - \eta\}$,

- Log-Log: $g'(\mu_t) = \exp\{\exp(-\eta) + \eta\}$.

Por exemplo, caso seja escolhida a função de ligação logito,

$$g(\mu_i) = \ln\left(\frac{\mu_i}{1-\mu_i}\right),$$

a média condicional da variável dependente Y_i pode ser escrita em função das variáveis preditoras como:

$$\mu_i = \frac{\exp(\eta_i)}{1+\exp(\eta_i)} = \frac{\exp\left\{\sum_{j=1}^{k}\beta_j x_{ij}\right\}}{1+\exp\left\{\sum_{j=1}^{k}\beta_j x_{ij}\right\}}.$$

3.4.2 Como escolher a função de ligação

A função de ligação mais comumente utilizada é a logito, pois possibilita interpretações simples para os parâmetros de regressão. McCullagh e Nelder (1989) compararam a função de ligação logito com a probito, complemento log-log e log-log. A função de ligação logito e probito se relacionam quase linearmente quando μ_i se encontra no intervalo $[0.1; 0.9]$. Para valores pequenos de μ_i a função de ligação complemento log-log se aproxima da logito, enquanto para valores de μ_1 próximos de 1, a função de ligação log-log se aproxima da função de ligação logito.

Giovanetti e Andrade (2007) comparam essas funcões e verificaram que se a função de ligação verdadeira for a logito ou a probito os ajustes oferecidos por qualquer uma dessas funções proporcionam resultados satisfatórios e semelhantes para as estimativas da média. Alem disso, esses ajustes tendem a melhorar à medida que o tamanho amostral aumenta.

3.5 Modelo de regressão beta de Paolino

Paolino (2001) mostrou empiricamente para os casos em que a variável resposta (dependente) está restrita ao intervalo $(0, 1)$ a estimação dos parâmetros feita via mínimos quadrados, geralmente, leva a distorções do modelo. Por esse motivo Paolino (2001) propôs um modelos de regressão beta para aplicações em ciências políticas. O método desenvolvido por Paolino baseia-se na suposição que o parâmetro de precisão ϕ não é constante. Nesse sentido, define-se uma estrutura de regressão para o parâmetro de precisão.

Seja $Y_1, \ldots, Y_n$ são variáveis aleatórias independentes, em que cada Y_i, $i = 1, \ldots, n$, segue a densidade em (3.4) e com média μ_i definida em (3.11) e com precisão ϕ_i

$$h(\phi_i) = \sum_{j=1}^{q} z_{ij}\gamma_j, \tag{3.13}$$

em que $\gamma = (\gamma_1, \ldots, \gamma_q)^T$ é um vetor de parâmetros desconhecidos ($\gamma \in \mathbb{R}^q$), $z_{i1}, \ldots, z_{iq}$ são observações de q covariáveis ($q < n$) assumidas fixas e conhecidas e $h(\cdot) : (0; \infty) \to \mathbb{R}$ é uma função estritamente monótona e duas vezes diferenciável. As funções de ligação mais comumente empregadas para parâmetro de precisão são o logaritimo , $h(\phi_i) = \log(\phi_i)$ e a raiz quadrada, $h(\phi_i) = \sqrt{(\phi_i)}$.

3.5.1 Estimação dos parâmetros

A estimação conjunta dos parâmetros no modelo de regressão beta é realizada por máxima verossimilhança. Para tanto, utiliza-se o logaritmo da função de verossimilhança, baseada numa amostra de n observações independentes, que pode ser expresso como:

$$l(\beta, \gamma) = \sum_{i=1}^{n} l(\mu_t, \phi_t),$$

em que

$$\begin{aligned} l(\beta, \gamma) &= \log\Gamma(\phi_t) - \log\Gamma(\mu_t\phi_t) \\ &- \log\Gamma((1-\mu_t)\phi_t) + (\mu_t\phi_t - 1)\log(y_t) \\ &+ \{(1-\mu_t)\phi_t - 1\}\log(1-y_t). \end{aligned} \tag{3.14}$$

Sejam

$$y_t^* = \log\left(\frac{y_t}{1-y_t}\right)$$

e

$$\mu_t^* = \Psi(\mu_t\phi_t) - \Psi((1-\mu_t)\phi_t),$$

em que $\Psi(\cdot)$ é a função digama. A função escore, obtida pela diferenciação do logaritmo da função de verossimilhança com respeito aos parâmetros desconhecidos (β, γ) e dada por $U = (U_\beta(\beta, \gamma), U_\gamma(\beta, \gamma))^T$ e sua forma matricial é

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

em que X é uma matriz $n \times k$, onde x_i^T é a i-ésima linha dessa matriz.

$$T = diag\left\{g'\left(\mu_1\right)^{-1}, \ldots, g'\left(\mu_n\right)^{-1}\right\},$$

$$y^* = \left(\gamma_1^*, \ldots, \gamma_n^*\right)^T,$$

$$\mu^* = \left(\mu_1^*, \ldots, \mu_n^*\right)^T,$$

$$\Phi = diag\left\{\phi_1, \ldots, \phi_n\right\}. \tag{3.15}$$

A função escore γ é dada por

$$U_\gamma\left(\beta, \gamma\right) Z^T H a,$$

em que Z é uma matriz $n \times q$, cuja a t-ésima linha é z_t^T.

$$H = diag\left\{h'\left(\phi_1\right)^{-1}, \ldots, h'\left(\phi_n\right)^{-1}\right\},$$

$$a = \left(a_1, \ldots, a_n\right)^T,$$

$$a_t = \mu_t\left(y_t^* - \mu_t\right) + \log\left(1 - y_t\right) - \Psi\left(\left(1 - \mu_t\right)\phi_t\right) + \Psi\left(\phi_t\right).$$

Através da resolução do sistema

$$\begin{cases} U_\beta\left(\beta, \gamma\right) = 0 \\ U_\gamma\left(\beta, \gamma\right) = 0 \end{cases} \tag{3.16}$$

obtêm-se os estimadores de máxima verossimilhança $\hat{\beta}$ e $\hat{\gamma}$. Estes estimadores não possuem forma fechada que se possa escrever de forma explícita a partir das equações anteriores, por isso, faz-se necessário o emprego de algoritmos de otimização não-linear, para obter a maximização numérica da função de log-verossimilhança (3.14), usualmente são utilizados os algoritmos Newton-Raphson, Escore de Fisher ou quasi-Newton, também conhecido como BFGS. Alguns detalhes desse algoritmos são apresentados na subseção (3.5.3) para mais informações sobre os algoritmos de otimização (Nocedal e Wright, 1999).

As segundas derivadas da função log-verossimilhança com relação aos parâmetros desconhecidos resultam na matriz de informação observada, que é definida por

$$j = j(\beta, \gamma) = \begin{pmatrix} J_{\beta\beta} & J_{\beta\gamma} \\ J_{\gamma\beta} & J_{\gamma\gamma} \end{pmatrix},$$

em que

$$J_{\beta\beta} = -X^T \Phi Q X,$$

$$J_{\beta\gamma} = J_{\beta}^T = -X^T F T H Z$$

e

$$J_{\gamma\gamma} = -Z^T V Z,$$

sendo $\Psi'(\cdot)$ a função trigama, Φ é a matriz em (3.15), $Q = diag\{q_1, \ldots, q_n\}$, com

$$q_t = \left\{ \phi_t \left[\Psi'(\mu_t \phi_t) + \Psi'((1-\mu_t)\phi_t) \right] + (y^* - \mu^*) \frac{g''(\mu_t)}{g'(\mu_t)} \right\} \frac{1}{[g'(\mu_t)]^2}. \tag{3.17}$$

Adicionalmente, $F = diag[f_1, \ldots, f_n]$, com

$$f_t = c_t - (y^* - \mu^*)$$

e

$$c_t = \phi_t \left[\Psi'(\mu_t \phi_t)\mu_t - \Psi'((1-\mu_t)\phi_t(1-\mu_t)) \right]. \tag{3.18}$$

$V = diag[v_1, \ldots, v_n]$ com

$$v_t = d_t + a_t \frac{h''(\phi_t)}{[h'(\phi_t)]^3},$$

$$d_t = \left[\Psi'(\mu_t \phi_t)\mu_t^2 + \Psi'((1-\mu_t)\phi_t)(1-\mu_t)^2 - \Psi'(\phi_t) \right] \frac{1}{[h'(\phi_t)]^2}. \tag{3.19}$$

A matriz de informação de Fisher, é dada pela esperança da matriz de informação observada J

$$K = K(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\gamma} \\ J_{\gamma\beta} & K_{\gamma\gamma} \end{pmatrix},$$

em que

$$K_{\beta\beta} = X^T \Phi W X,$$

$$K_{\beta\gamma} = K_{\gamma\beta}^T = X^T C T H Z$$

e

$$K_{\gamma\gamma} = Z^T D Z.$$

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

$$w_t = \phi_t \left\{ \Psi'(\mu_t \phi_t) + \Psi'((1-\mu_t)\phi_t) \right\} \frac{1}{[g'(\mu_t)]^2}. \tag{3.20}$$

com $C = diag\{c_1, \ldots, c_n\}$, com c_t definido em (3.18), e $D = diag\{d_1, \ldots, d_t\}$, com d_t definido em (3.19).

No caso em que a dispersão é constante, ou seja $\phi_1 = \ldots = \phi_n = \phi$ as expressões em (3.18), (3.19) e (3.20) passam a ser, respectivamente,

$$c_t = \phi \left[\Psi'(\mu_t \phi) \mu_t - \Psi'((1-\mu_t)\phi(1-\mu_t)) \right],$$

$$d_t = \left[\Psi'(\mu_t \phi) \mu_t^2 + \Psi'((1-\mu_t)\phi)(1-\mu_t)^2 - \Psi'(\phi) \right]$$

e

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

É importante notar que os parâmetros β e γ não são ortogonais no modelo de regressão beta. Os estimadores de máxima verossimilhança dos parâmetros, sob certas condições e quando a amostra é grande, têm distribuição aproximada dada por

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \simeq N_{k+q} \left(\begin{pmatrix} \beta \\ \gamma \end{pmatrix}; K^{-1} \right),$$

em que

$$K^{-1} = K^{-1}(\beta, \gamma) = \begin{pmatrix} K^{\beta\beta} & K^{\beta\gamma} \\ K^{\gamma\beta} & K^{\gamma\gamma} \end{pmatrix}. \tag{3.21}$$

É a inversa da matriz de informação de Fisher, obtida por Ferrari e Cribari-Neto (2004) utilizando uma expressão padrão para a inversa de matrizes particionadas (Rao, 1973). Seus valores são dados por

$$K^{\beta\beta} = \left(X^T \Phi W X - X^T C T H Z \left(Z^T D Z\right)^{-1} Z^T H T C^T X\right)^{-1}, \tag{3.22}$$

$$K^{\beta\gamma} = \left(K^{\gamma\beta}\right)^T = -K^{\beta\beta} X^T C T H Z \left(Z^T D Z\right)^{-1},$$

$$K^{\gamma\gamma} = \left(Z^T D Z\right)^{-1} = \left\{I_q + \left(Z^T H T C^T X\right) K^{\beta\beta} X^T C T H Z \left(Z^T D Z\right)^{-1}\right\}. \tag{3.23}$$

em que I_q é a matriz identidade de ordem q. Uma extensão dos modelos de regressão beta inflacionados são introduzidos por Ospina (2008) para os casos em que os dados assumem continuamente valores nos intervalos $[0, 1)$, $(0, 1]$ ou $[0, 1]$, uma vez que a distribuição beta apresentada na secção (3.1) não é adequada. Ospina e Ferrari (2010) propuseram a distribuição beta inflacionada que é uma mistura entre a distrbuição beta e a distribuição de bernoulli degenerada no ponto c ($c = 0 \quad ou \quad c = 1$).

3.5.2 Intervalos de confiança

Os intervalos de confiança de cobertura $(1 - \alpha)$, $0 < \alpha < \frac{1}{2}$, para β_i, $i = 1, \ldots, k$ e γ_j, $j = 1, \ldots, n$, tem limites dados, respectivamente, por

$$\hat{\beta}_i \pm \Phi\left(1 - \frac{\alpha}{2}\right) ep\left(\hat{\beta}_i\right)$$

e

$$\hat{\gamma}_j \pm \Phi\left(1 - \frac{\alpha}{2}\right) ep\left(\hat{\gamma}_j\right),$$

em que $\Phi(\cdot)$ é a função quantil de uma variável aleatória normal padrão $ep\left(\hat{\beta}_i\right)$ e $ep\left(\hat{\gamma}_j\right)$, são respectivamente, os erros-padrão de $\hat{\beta}$ e $\hat{\gamma}$. Os erros-padrão das respectivas estimativas de máxima verossimilhança são obtidos pela raiz quadrada do i-ésimo da diagonal da matriz $\hat{K}^{\beta\beta} = K^{\beta\beta}\left(\hat{\theta}\right)$, com $K^{\beta\beta}$ definida em (3.22), e da raiz quadrada do j-ésimo elemento da diagonal da matriz $\hat{K}^{\gamma\gamma} = K^{\gamma\gamma}\left(\hat{\theta}\right)$, com $K^{\gamma\gamma}$ definida em (3.23)

Um intervalo de confiança $(1 - \alpha) \times 100\%$ para μ_t, a média da variável resposta para um dado vetor de covariáveis x_t^T, pode ser obtido por

$$\left[g^{-1}\left(\hat{\eta}_t - \Phi^{-1}\left(1-\frac{\alpha}{2}\right) ep\left(\hat{\eta}_t\right)\right), g^{-1}\left(\hat{\eta}_t + \Phi^{-1}\left(1-\frac{\alpha}{2}\right) ep\left(\hat{\eta}_t\right)\right)\right],$$

em que $\hat{\eta}_t = x_t^T \hat{\beta}$ e o erro-padrão de $\hat{\eta}_t$ é dado por

$$ep\left(\hat{\eta}_t\right) = \left(x_t^T \hat{K}^{\beta\beta}\left(\hat{\beta}\right) x_t\right)^{\frac{1}{2}}, \tag{3.24}$$

este intervalo é válido para funções de ligação estritamente crescentes.

3.5.3 Otimização não linear

Os estimadores de máxima verossimilhança de $\hat{\beta}$ e $\hat{\gamma}$ são obtidos das equações $U_\beta(\beta,\gamma) = 0$ e $U_\gamma(\beta,\gamma) = 0$ dados em (3.16). Entretanto não há uma forma fechada em que se possa escrever os estimadores de forma explícita a partir das equações anteriores. Portanto um algoritmo de otimização não linear é necessário para obter a maximização numérica da função de log-verossimilhança (3.14). Usualmente são utilizados os algoritmos Newton-Raphson, Escore de Fisher ou quasi-Newton, também conhecido como BFGS. Alguns detalhes desse algoritmos são apresentados a seguir.

Newton-Raphson

Seja $\theta = (\beta,\gamma)$ o vetor de parâmetros. $U(\theta) = \left(U_\beta(\beta,\gamma), U_\gamma(\beta,\gamma)\right)^T$, vetor de funções escore de dimensão $(K+1)\times 1$. Para a obtenção do estimador de máxima verossimilhança do vetor θ expandimos a função escore $U(\theta)$ em torno de um valor inicial $\theta^{(0)}$, tal que:

$$U(\theta) \approx U\left(\theta^{(0)}\right) + U'\left(\theta^{(0)}\right)\left(\theta - \theta^{(0)}\right).$$

Em que $U'\left(\theta^{(0)}\right)$ é a derivada de 1ª ordem de $U(\theta)$ com respeito a θ^T. Fazendo $U(\theta) = 0$ e repetindo o processo acima, chegamos ao processo iterativo:

$$\theta^{(n+1)} = \theta^{(n)} + \left\{-U'\left(\theta^{(n)}\right)\right\}^{-1} U\left(\theta^{(n)}\right), \tag{3.25}$$

em que $n = 0, 1, 2, \ldots, k$.

Escore de Fisher

Uma vez que, pela lei dos grandes números, $U'(\theta)$ converge para a matriz K quando $n \to \infty$, substituindo a informação observada em (3.25) pela esperada, obtemos a seguinte aproximação:

$$\theta^{(n+1)} = \theta^{(n)} + \left\{-K^{(n)}\right\}^{-1} U\left(\theta^{(n)}\right), \tag{3.26}$$

em que $n = 0, 1, 2, \ldots, k$. Esse procedimento iterativo é denominado Escore de Fisher.

BFGS

Esse método utiliza uma sequência de matrizes simétricas e positivas denominadas $B^{(n)}$ no lugar da matriz $U'\left(\theta^{(n)}\right)^{-1}$. Fato comum é a utilização da matriz identidade $B^{(0)}$, de mesma ordem , como matriz inicial. A forma recursiva para obter as demais matrizes é dada por

$$B_{n+1}^{-1} = B_n^{-1} - \frac{B_n^{-1} y_n S_n^T + S_n y_n^T B_n^{-1}}{S_n^T y_n} + \frac{\left(S_n^T y_n + y_n^T B_n^{-1} y_n\right)\left(S_n S_n^T\right)}{\left(S_n^T y_n\right)^2}, \tag{3.27}$$

em que $S_n^T y_n$ e $y_n^T B_n^{-1} y_n$ são escalares e $n = 0, 1, 2, \ldots, k$.

Escolha do valor inicial

Nos algoritmos de maximização numérica é exigido um valor inicial para os parâmetros de interesse. Ferrari e Cribari-Neto (2004) sugerem utilizar para β a estimativa de mínimos quadrados ordinários obtida da regressão linear das respostas transformadas $(g(y_1), \ldots, g(y_n))$ com relação a X, ou seja, $\left(X^T X\right)^{-1} X^T z$, onde $z = (g(y_1), \ldots, g(y_n))^T$. Como Ferrari e Cribari-Neto (2004) consideram o parâmetro de dispersão constante ϕ, os mesmos autores sugerem como valor inicial sendo

$$\phi_t = \left[\mu_k \frac{(1-\mu_k)}{V(Y_k)}\right] - 1.$$

Baseado no fato que

$$Var(Y_k) = \frac{\mu_k (1-\mu_k)}{1-\phi}.$$

3.6 Diagnósticos

A análise de diagnóstico é uma etapa importante na análise de um modelo ajustado para verificar a existência de possíveis afastamentos das suposições do modelo. A metodologia de diagnóstico tem seu início com a análise dos resíduos para detectar a presença de pontos extremos e avaliar a adequação da distribuição proposta para variável resposta.

Muitos trabalhos têm sido publicados acerca de análise de resíduos em modelos de regressão, McCullagh (1987) apresenta uma padronização do componente do desvio em que se procura corrigir os efeitos de assimetria e curtose, Atkinson (1985) propõe a construção de uma banda de confiança para resíduos da regressão linear normal, a qual denomina-se envelope, e que permite melhor avaliar se os resíduos têm a distribuição esperada sob as suposições do modelo, Williams (1987) discute a construção de envelopes em modelos lineares generalizados, Ferrari e Cribari-Neto (2004), Espinheira et al. (2008a) e Espinheira et al. (2008b) apresentam resíduos padronizados para modelos de regressão beta.

Após a estimação dos parâmetros populacionais, em geral, realizam-se testes a fim de determinar se hipóteses feitas sobre estes parâmetros são suportadas por evidências obtidas a partir de dados amostrais. Neste contexto, testes baseados na função de verossimilhança são amplamente empregados devido a suas propriedades de otimalidade. Os procedimentos mais frequentemente utilizados são os testes da razão de verossimilhanças e Wald, que são assintoticamente equivalentes sob a hipótese nula.

No caso dos modelos de regressão beta, as estatísticas dos testes da razão de verossimilhanças e Wald foram apresentadas por Ferrari e Cribari-Neto (2004). Um teste de erro de especificação, baseado no teste RESET (Ramsey, 1969), para o modelo com dispersão constante foi desenvolvido por Cribari-Neto e Lima (2007). Ferrari e Pinheiro (2011), visando realizar inferências mais precisas em amostras finitas, derivaram o ajuste de Skovgaard (2001) para esta classe de modelos. As autoras concluíram que as estatísticas ajustadas propostas, sob a hipótese nula, têm distribuição mais próxima da distribuição qui-quadrado de referência que a estatística original.

A estratégia de eliminar pontos é uma técnica usual para avaliar o impacto da retirada de uma observação particular nas estimativas dos parâmetros. A distância de

Cook (1977) foi estendida para a classe de modelos de regressão beta. A eliminação individual de pontos pode ocasionar um problema conhecido como efeito de mascaramento, ou seja, deixar de detectar pontos conjuntamente discrepantes. Outro aspecto importante na análise de diagnóstico é a detecção de observações influentes. Levarege é um componente chave na análise de influência em modelos de regressão, Wei, Hu e Fung (1998) generalizam a definição de leverage para modelos gerais cuja a variável resposta seja contínua.

Esta seção encontra-se organizada da seguinte forma. Na secção 3.6.1 são apresentados os testes da razão de verossimilhanças e Wald. Na secção 3.6.2 é apresentado o "pseudo" R^2. Na secção 3.6.3 discutiremos sobre a análise gráfica tais como envelope simulado e gráfico de resíduos. Na secção 3.6.4 serão apresentadas a análise de influência definidas tais como distância de Cook e Leverage generalizado. Na secção 3.6.5 serão apresentados os resíduos ponderados padronizados proposto em Espinheira et al.(2008a) para modelos de regressão beta.

3.6.1 Teste de razão de verossimilhança (TRV) e Wald(W)

A função de verossimilhança contém toda a informação relevante para fazer inferência acerca de um vetor de parâmetros de interesse. Portanto, dentre as técnicas para avaliação de modelos estatísticos, os testes baseados na função de verossimilhança são amplamente empregados. A função de verossimilhança *"informa a ordem natural de preferência entre diversas possibilidades de θ (...) Generalizando, entre os possíveis candidatos para estimar o parâmetro verdadeiro (...) a partir dos mesmos dados y, o vetor de parâmetros mais plausível é aquele de maior verossimilhança"* (Cordeiro, 1999).

Seja Θ o espaço paramétrico e o teste dado por:

$$\begin{cases} H_0 : \theta = \theta^{(0)} \\ H_0 : \theta \neq \theta^{(0)}, \end{cases}$$

em que $\theta^{(0)} \in \theta_0$, o espaço paramétrico restrito, $\theta_0 \subset \Theta$. A razão de verossimilhança mensura o quanto as evidências estatísticas corroboram com valores para o vetor de parâmetros diferentes daqueles especificados em H_0, ou seja, valores grandes da razão de verossimilhanças não apoiam a plausibilidade da hipótese nula. A estatística do

teste da razão de verossimilhanças pode ser expressa por:

$$\frac{L\left(\theta|y\right)}{L\left(\theta^{(0)}|y\right)},$$

$$RV = 2\left[l\left(\hat{\theta}\right) - l\left(\theta^{(0)}\right)\right],$$

em que $l\left(\theta\right)$ é o logaritmo da função de verossimilhança relativo ao vetor θ e $\hat{\theta}$ é o estimador de máxima verossimilhança do parâmetro que indexa o modelo completo. Sob certas condições e sob a hipótese nula, RV se distribui, assintoticamente, como uma χ^2_r, sendo r o número de parâmetros testados (Wilks, 1938).

Além da razão de verossimilhanças, a estatística de Wald (Wald, 1943) é comumente utilizada para medir a distância entre as hipóteses nula e alternativa descritas anteriormente. A estatística de Wald é expressa por:

$$W = \left(\hat{\theta} - \theta^{(0)}\right)^T K\left(\hat{\theta}\right)\left(\hat{\theta} - \theta^{(0)}\right), \tag{3.28}$$

em que $K\left(\hat{\theta}\right)$ é a informação esperado relacionada ao vetor θ. Sob certas condições e sob hipótese, W segue, assintoticamente distribuição χ^2_r com r de liberdade e é assintoticamente, equivalentemente a RV.

O teste da razão de verossimilhanças é formulado em termos da diferença $l\left(\hat{\theta}\right) - l\left(\theta^{(0)}\right)$, o teste Wald, por sua vez, considera o quadrado da distância entre θ e $\hat{\theta}^{(0)}$. Do ponto de vista do teste da razão de verossimilhanças, porém, devido à diferença de curvatura de suas funções de log-verossimilhança, um destes conjuntos de dados é menos favorável à hipótese nula e, portanto, é preciso que a distância quadrada seja ponderada por $C\left(\hat{\theta}\right)$. Assim a estatística Wald é definida por

$$W = \left(\hat{\theta} - \theta^{(0)}\right)^2 C\left(\hat{\theta}\right).$$

Usualmente, contudo, pondera-se $\left(\hat{\theta} - \theta^{(0)}\right)$ em termos da curvatura média $K\left(\hat{\theta}\right)$, em que $K\left(\theta\right) = E\dfrac{\partial^2 \log L}{\partial \theta^2}$, a matriz de informação esperada. Esta abordagem conduz à estatística apresentada em (3.28)

Seja o vetor $y = (y_1, \ldots, y_n)^T$, em que n observações independentes, cada y_i seguindo a densidade definida em (3.4). Suponha que μ e ϕ relacionam-se, respectivamente, com β e γ por meio de (3.10) e (3.13).

Considere a hipótese

$$\begin{cases} H_0 : \beta = \beta^{(0)} \\ H_0 : \beta \neq \beta^{(0)}, \end{cases}$$

em que $\beta = \left(\beta_{(r)}^t, \beta_{(k-r)}^t\right)^T$ e $\beta^{(0)}$ um vetor r dimensional cujos valores são conhecidos.

A estatistica de da razão de verossimilhança é dada por

$$RV = 2\left[l\left(\hat{\beta}, \hat{\gamma}\right) - l\left(\tilde{\beta}, \tilde{\gamma}\right)\right],$$

em que (β, γ) e o logaritmo da função de verossimilhança definido em (3.14), $l\left(\hat{\beta}, \hat{\gamma}\right)$ e $l\left(\tilde{\beta}, \tilde{\gamma}\right)$ são respectivamente, os estimadores de máxima verossimilhança de (β, γ) do modelo completo e incompleto. O estimador é obtido pela imposição da hipótese nula.

Seja $K_{(r)(r)}^{\beta\beta}$ a matriz $m \times m$ formada das r linhas iniciais e das r colunas iniciais da matriz K^{-1}, definida em (3.21). A estatística de Wald pode ser definida em um modelo de regressão beta, como

$$W = \left(\hat{\beta}_{(r)} - \beta^{(0)}\right)^T \left(\hat{K}_{(r)(r)}^{\beta\beta}\right)^{-1} \left(\hat{\beta}_{(r)} - \beta^{(0)}\right),$$

em que $\left(\hat{K}_{(r)(r)}^{\beta\beta}\right)^{-1} = \left(K_{(r)(r)}^{\beta\beta}\right)^{-1}$ avaliado no estimador de máxima verossimilhança do modelo completo e $\hat{\beta}_{(r)}$ é o estimador de máxima verossimilhança de $\beta_{(r)}$. Sob certas condições RV e W convergem em distribuição para $\chi_{(r)}^2$. Assim os testes podem ser realizados usando valores críticos aproximados obtidos como quantis da distribuição qui-quadrado com r graus de liberdade. A obtenção das estatísticas de teste para os parâmetros γ são análogas as descritas para os parâmetros β

3.6.2 Pseudo R^2 $\left(R_p^2\right)$

Em princípio, uma medida global da qualidade de ajuste pode ser obtida através do cálculo do "pseudo" R^2 $\left(R_p^2\right)$ definido como o quadrado do coeficiente de correlação amostral entre $g(y)$ e $\hat{\eta}$. Note que $0 \leq R_p^2 \leq 1$ e quanto mais próximo de 1 for seu valor, melhor será o ajuste.

3.6.3 Envelope Simulado

Atkinson (1985) sugere que, para poder interpretar melhor um gráfico normal de probabilidades do resíduo proposto, este deve ser auxiliado pela construção de envelopes. Os envelopes simulados são bandas de confiança obtidas pelo método de Monte Carlo a partir do modelo ajustado para avaliar a existência de afastamentos sérios da distribuição proposta. Em um gráfico de probabilidade meio-normal, temos o t-ésimo ($t = 1, \ldots, n$) valor ordenado dos resíduos versus os valores esperados das estatísticas de ordem, em valor absoluto, da distribuição normal padrão $N(0, 1)$, dado por

$$\Phi^{-1}\left(\frac{t+n-\frac{1}{8}}{2n+\frac{1}{2}}\right),$$

em que $\Phi(\cdot)$ é a função de distribuição acumulada da distribuição $N(0, 1)$. Esse gráfico pode ser utilizado mesmo que os resíduos não tenham distribuição normal. Quando isso ocorre, não esperamos que os mesmos estejam próximos da reta identidade. Para a construção do gráfico meio-normal com envelope simulado seguimos os seguintes passos:

1. Ajuste o modelo da amostra original e calcule o resíduo proposto;
2. Gere uma amostra simulada de n observações independentes utilizando o modelo ajustado como se esse fosse o modelo verdadeiro;
3. Ajuste o modelo para a amostra gerada e calcule os valores absolutos ordenados da medida de diagnóstico de interesse (resíduos ponderados padronizados 2);
4. Repita os passos (2) e (3) k vezes;
5. Do grupo de K valores dos resíduos obtidos, calcule suas respectivas médias, valores mínimos e máximos;
6. Construa um gráfico desses valores e do resíduo ordenado da amostra original contra percentis esperados da distribuição normal dados por (1);

Os valores mínimos e máximos das K réplicas produzem o envelope. Segundo Atkinson (1985), é suficiente usar $K = 19$. Deste modo a probabilidade de qualquer

resíduo em particular exceder o limite superior fica sendo aproximadamente igual a $1/20 = 0.05$.

3.6.4 Análise de Influência

Leverage Generalizado

Lerverage é um componente chave na análise de influência em modelos de regressão. Usualmente, é medido pelos elementos h_{ii} da matriz H que é conhecida como matriz de projeção ou matriz chapéu e é usado para avaliar a importância individual de cada observação no próprio valor ajustado. Na regressão linear múltipla, por exemplo, é razoável utilizar h_{ii} como medida de influência da i-ésima observação sobre o próprio valor ajustado. Supondo que todos exerçam a mesma influência sobre os valores ajustados, pode-se esperar que h_{ii} esteja próximo de $\frac{tr(H)}{n} = \frac{k}{n}$ em que k é o número de parâmetros do modelo. Uma sugestão é examinar aqueles pontos tais que $h_{ii} \geq \frac{2k}{n}$, que são conhecidos como pontos de leverage grandes.

Wei, Hu e Fung (1998) generalizaram a definição de leverage para modelos gerais cuja variável resposta seja contínua. Nessa generalização incluem-se outros métodos de estimação, além da máxima verossimilhança, e outros enfoques tais como o Bayesiano. Ferrari e Cribari-Neto (2004) obtiveram a expressão de forma fechada para leverage generalizado ($GL(\beta)$) do modelo de regressão beta, considerando ϕ conhecido. O leverage é dado por:

$$GL(\beta) = D_\beta \left(-\frac{\partial^2 l}{\partial \beta \partial \beta^T} \right)^{-1} \frac{\partial^2 l}{\partial \beta \partial y^t},$$

em que

$$D_\beta = \frac{\partial \mu}{\partial \beta^T} = \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta^T} = TX.$$

A expressão $-\frac{\partial^2 l}{\partial \beta \partial \beta^T}$ é dada por

$$-\frac{\partial^2 l}{\partial \beta \partial \beta^T} = \phi X^T Q X,$$

em que $Q = diag(q_1, \ldots, q_n)$ com q_t dado por (3.17). Adicionalmente temos que

$$\frac{\partial^2 l}{\partial \beta \partial \beta^T} = \phi X^T T M,$$

em que $M = diag\,(m_1 \ldots, m_n)$ com $m_t = \dfrac{1}{y_t\,(1-y_t)}$. E assim temos que a expressão acima pode ser escrita como

$$GL\,(\beta) = TX\left(X^T Q X\right)^{-1} X^T T M.$$

Ferrari e Cribari-Neto (2004) notaram que ao substituir a informação observada, $-\dfrac{\partial^2 l}{\partial\beta\partial\beta^T}$, pela informação esperada, $E\left(-\dfrac{\partial^2 l}{\partial\beta\partial\beta^T}\right)$, a expressão para $GL\,(\beta)$ fica como a expressão anterior, mas com a matriz Q substituída por W, sendo dada por, $GL^*\,(\beta)$. Pode-se notar que os elementos da diagonal $GL^*\,(\beta)$ são os mesmo de

$$M^{\frac{1}{2}} TX\left(X^T W X\right)^{-1} X^T M^{\frac{1}{2}},$$

em que $M^{\frac{1}{2}}$ é a matriz diagonal cujo k-ésimo elemento da diagonal é dado por $\{g^{-1}\,(\mu_i)\,V\,(\mu_i)\}^{-1}$. É importante notar que existe uma ligação entre os elementos da diagonal de $GL^*\,(\beta)$ e da matriz chapéu usual quando ϕ é grande. A matriz chapéu é dada por

$$H = W^{\frac{1}{2}} X\left(X^T W X\right)^{-1} X^T W^{\frac{1}{2}}, \tag{3.29}$$

pois quando ϕ é grande o elemento diagonal w_t de $W^{\frac{1}{2}}$ é aproximadamente igual a $\{g^{-1}\,(\mu_i)\,V\,(\mu_i)\}^{-1}$.

Distância de Cook

A detecção de observações influentes é um tópico importante na análise de diagnóstico, isto é, pontos que exercem um peso desproporcional nas estimativas dos parâmetros do modelo. Uma medida de influência de cada observação nas estimativas dos parâmetros de regressão é a distância de Cook.

Seja $L\,(\theta)$ a log verossimilhança baseada nos dados completos. O afastamento da verossimilhança $LD\left(\theta_{(i)}\right)$ é definido como

$$LD\left(\theta_{(i)}\right) = 2\left[L\left(\hat{\theta}\right) - L\left(\hat{\theta}_{(i)}\right)\right]. \tag{3.30}$$

Supondo que $L\left(\hat{\theta}_{(i)}\right)$ pode ser bem representada por uma função quadrática, $L\left(\hat{\theta}_{(i)}\right)$ usualmente pode ser aproximada pela expansão de Taylor até a segunda ordem de $L\left(\hat{\theta}_{(i)}\right)$ ao redor de $\hat{\theta}$.

$$LD\left(\hat{\theta}_{(i)}\right) \approx L\left(\hat{\theta}\right) + \left(\hat{\theta}_{(i)} - \hat{\theta}\right)^T L^*\left(\hat{\theta}\right) + \frac{1}{2}\left(\hat{\theta}_{(i)} - \hat{\theta}\right)^T \left[L^{**}\left(\hat{\theta}\right)\right]\left(\hat{\theta}_{(i)} - \hat{\theta}\right). \tag{3.31}$$

Deste modo, substituindo (3.31) na equação (3.29) e, desde que $L^*\left(\hat{\theta}\right) = 0$, pode-se obter uma aproximação de $LD\left(\hat{\theta}_{(i)}\right)$. Essa aproximação, denominada $DG\left(\hat{\theta}_{(i)}\right)$, assume a seguinte expressão:

$$DG\left(\hat{\theta}_{(i)}\right) = \left(\hat{\theta}_{(i)} - \hat{\theta}\right)^T \left[L^{**}\left(\hat{\theta}\right)\right] \left(\hat{\theta}_{(i)} - \hat{\theta}\right).$$

A expressão acima pode ser enunciada de uma maneira mais geral, a qual tem sido chamada de Distância de Cook generalizada:

$$DG\left(\hat{\theta}_{(i)}\right) = \left(\hat{\theta}_{(i)} - \hat{\theta}\right)^T C \left(\hat{\theta}_{(i)} - \hat{\theta}\right).$$

Onde C é uma matriz positiva definida e assintoticamente equivalente à matriz de informação esperada de θ. Se consideram dois casos para C, $-L^{**}\left(\hat{\theta}\right)$ e $K(\theta)$, as matrizes de informação observada e esperada de Fisher, respectivamente.

Para evitar que o modelo seja ajustado $n+1$ vezes, usualmente utiliza-se a aproximação

$$C_t = \frac{h_{tt} r_t^2}{k\left(1 - h_{tt}\right)^2},$$

que combina leverage e resíduos. É comum plotar o gráfico $C_t \times t$

3.6.5 Resíduos ponderados padronizados

Ferrari e Cribari-Neto (2004) propõem uma medida global de qualidade de ajuste, baseada no fato de que a discrepância de um ajuste pode ser medida como duas vezes a diferença entre o máximo do logaritmo de verossimilhança do modelo saturado e o do modelo de interesse (modelo postulado). Define-se inicialmente o resíduo ordinário como:

$$r_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\widehat{Var}\left(y_t\right)}},$$

em que $\mu_i = g^{-1}(\cdot)$ e $\widehat{Var}\left(y_i\right) = \frac{\hat{\mu}_i(1-\hat{\mu}_i)}{1+\phi}$.

Posteriormente Espinheira et al. (2008a) sugerem utilizar resíduos padronizados obtidos da convergência do processo iterativo escore de Fisher para estimação dos parâmetros de regressão, os chamados resíduos ponderados padronizados 1 e 2.

Através de simulações, Espinheira et al. (2008), constatam que os resíduos ponderados padronizados 2 apresentam desempenho superior, especialmente no sentido de identificar observações influentes. O resíduo ponderado padronizado 2 para o modelo de regressão beta com dispersão variável, segundo Espinheira et al. (2008), é dado por:

$$r_i^{pp} = \frac{y_i - \hat{\mu}_i^*}{\sqrt{\upsilon_i \left(1 - h_{tt}\right)}},$$

em que $i = 1, \ldots, n$,sendo $\mu_i^* = \Psi\left(\mu_i\phi_i\right) - \Psi\left[\left(1-\mu_i\right)\phi_i\right]$, $\upsilon_i = \Psi'\left(\mu_i\phi_i\right) + \Psi'\left[\left(1-\mu_i\right)\phi_i\right]$ e h_{tt} é o t-ésimo elemento da diagonal principal da matriz chapéu (3.29)

3.7 Modelo de regressão beta - abordagem bayesiana

Quando o pesquisador possui alguma informação inicialmente a respeito dos parâmetros de interesse pode-se utilizar técnicas bayesianas. No enfoque bayesiano os parâmetros são vistos como variáveis aleatórias e não mais como constantes como no caso dos métodos clássicos. Dessa forma a incerteza de um modelo dado θ é representado através de uma distribuição de probabilidade $P\left(\theta\right)$ sobre os possíveis valores do parâmetro desconhecido.

A metodologia utilizada na abordagem é devido ao teorema de Bayes , como discutido em **??**, que mostra a relação entre a probabilidade condicional da evidência dado a hipótese. Resumindo o uso desse teorema nos permite incorporar a informação à Priori com o conjunto de dados observados (verossimilhança) e assim, após essa composição chegamos no que podemos dizer de informação a Posteriori.

3.7.1 Estimação dos parâmetros

A variável resposta segue distribuição beta, $Y \sim Beta\left(p, q\right)$, em que os parâmetros p e q seguem uma estrutura de regressão. Seja $Y_1, \ldots, Y_n$ em que

$$Y_i \sim Beta\left(p_i, q_i\right),$$

sendo p_i e q_i dados por

$$\begin{aligned} p_i &= \frac{(\mu_i)^2 (1-\mu_i)}{Var(Y_i)} - \mu_i, \\ q_i &= \frac{(\mu_i)(1-\mu_i)^2}{Var(Y_i)} - (1-\mu_i). \end{aligned}$$

$Var(Y_i)$ é dado por

$$Var(Y_i) = \frac{\mu_i (1-\mu_i)}{\phi_i + 1}.$$

A função de verossimilhança com o parâmetro de precisão constante, dada por

$$L(Y, X|\mu_1, \ldots, \mu_n, \phi) = \prod_{i=1}^{n} f(y_i|\mu_i, \phi).o$$

Sendo $f(y_i|\mu_i, \phi)$ dada por 3.4, segue que

$$\begin{aligned} L(Y, X|\mu_1, \ldots, \mu_n, \phi) &= [\Gamma(\phi)]^n \left[\prod_{i=1}^{n} \Gamma(\mu_i \phi)\right]^{-1} \\ &\times \left[\prod_{i=1}^{n} \Gamma(\phi(1-\mu_i))\right]^{-1} \prod_{i=1}^{n} \left[y_i^{(\mu_i\phi-1)}\right] \\ &\times \prod_{i=1}^{n} \left[(1-y_i)^{[\phi(1-\mu_i)-1]}\right], \end{aligned}$$

em que $\mu_i = \dfrac{exp\left(X_i^T\beta\right)}{[1+exp(X_i^T\beta)]}$, X_i^T é o vetor de covariáveis e β é o vetor de parâmetros desconhecidos. Substituindo μ_i por $\dfrac{exp\left(X_i^T\beta\right)}{[1+exp(X_i^T\beta)]}$ temos

$$\begin{aligned} L(Y, X|\beta, \phi) &= [\Gamma(\phi)]^n \left[\prod_{i=1}^{n} \Gamma\left(\frac{exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}{1 + exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}\phi\right)\right]^{-1} \\ &\times \left[\prod_{i=1}^{n} \Gamma\left(\phi\left(1 - \frac{exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}{1 + exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}\right)\right)\right]^{-1} \\ &\times \prod_{i=1}^{n} \left[y_i \left(\frac{exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}{1 + exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}^{\phi-1}\right)\right] \\ &\times \prod_{i=1}^{n} \left[(1-y_i)^{\left[\phi\left(1 - \frac{exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}{1 + exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in})}\right)-1\right]}\right]. \end{aligned}$$

Assumindo que ϕ é variável em que $\phi_i = exp\left(Z_i^T\gamma\right)$, onde Z_i^T é o vetor de covariáveis, relativo a i-ésima observação e γ é o vetor de parâmetros desconhecidos, temos que

$$L\left(Y,X|\beta,\phi\right) = \left[\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^n$$
$$\times\left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^{-1}$$
$$\times\left[\prod_{i=1}^{n}\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\left(1-\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right)\right]^{-1}$$
$$\times\prod_{i=1}^{n}\left[y_i^{\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)-1\right)}\right]$$
$$\times\prod_{i=1}^{n}\left[(1-y_i)^{\left[exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\left(1-\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)-1\right]}\right].$$

sendo assim temos

$$\begin{aligned}
L\left(Y,X|\beta,\phi\right) &= \left[\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^n \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \prod_{i=1}^{n}\left[y_i^{\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}-1\right)}\right] \\
&\times \prod_{i=1}^{n}\left[(1-y_i)^{\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}-1\right)}\right].
\end{aligned}$$

Para completar o processo de estimação é necessária a adoção de prioris para os parâmetros de interesse. A seguir demonstramos um processo de estimação dos parâmetros adotando prioris vagas e informativas. Uma priori é definida como vaga

quando apresenta uma média conhecida e uma grande dispersão, um exemplo é uma priori normal com média conhecida e grande variância.

3.7.2 Determinação da priori

A densidade a priori de $p(\beta, \phi)$ pode ser escolhida de várias formas. Uma delas considera $p(\beta, \phi) = p(\beta)\, p(\phi)$, o que equivale a β e ϕ independentes a priori. Por exemplo podemos considerar $\beta \sim N\left(\mu_\beta; \sigma_\beta^2\right)$ e $\phi \sim N\left(\mu_\phi; \sigma_\phi^2\right)$.

Reitman (2007) descreve que na utilização de prioris vagas e informativas, a obtenção da distribuição posteriori e da posteriori condicional, na maioria dos casos, não é tratável analiticamente. Dessa forma faz-se uso de procedimentos computacionais como o Gibbs sample e o Metropolis Hastings para retirar amostras dessas distribuições e realizar a estimação dos parâmetros.

Prioris normais vagas

Seja $\beta's$ prioris $N(0, 100)$ e $\phi's$ prioris $N(0, 100)$, ou seja, assumindo uma priori vaga, tanto para o parâmetro da média quanto para o parâmetro de precisão e que os parâmetros são independentes. A distribuição à posteriori é dada por:

$$p(\beta_0, \ldots, \beta_n, \gamma_0, \ldots, \gamma_n | Y, X, Z) = L(Y, X, Z | \beta, \phi)\, p(\beta_0) \ldots p(\beta_n)\, p(\gamma_0) \ldots p(\gamma_n),$$

$$
\begin{aligned}
&\propto \left[\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^n \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \prod_{i=1}^{n}\left[y_i\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times \prod_{i=1}^{n}\left[(1-y_i)\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times exp\left(\frac{-\beta_0^2-\beta_1^2-\ldots-\beta_n^2-\gamma_0^2-\gamma_1^2-\ldots-\gamma_n^2}{20000}\right).
\end{aligned}
$$

Como a distribuição é intratável analiticamente, será utilizado o procedimento Metropolis Hastings para retirar amostras da distribuição à posteriori. O interesse é realizar inferências para cada um dos parâmetros, nesse sentido deve-se encontrar a distribuição posteriori de cada um dos mesmos. A distribuição à posteriori marginal do parâmetro de interesse é obtida integrando a posteriori conjunta em relação aos demais parâmetros. Caso estejamos interessados em realizar inferências a respeito de β_1 sua posteriori é dada por $P\left(\beta_1|\beta_0,\beta_2,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n,Y,X,Z\right) \propto \int\int\ldots\int P\left(\beta_0,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n|Y,X,Z\right) d\beta_0 d\beta_2\ldots d\beta_n d\gamma_0 d\gamma_1\ldots d\gamma_n$

$$
P\left(\beta_1|\beta_0,\beta_2,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n,Y,X,Z\right) \propto
$$

$$
\begin{aligned}
&\propto \left[\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^n \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \prod_{i=1}^{n}\left[y_i\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times \prod_{i=1}^{n}\left[(1-y_i)\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times exp\left(\frac{-\beta_1^2}{20000}\right).
\end{aligned}
$$

Para gerar uma amostra de $\beta_1|\beta_0,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n,Y,X,Z$ é necessário utilizar o procedimento MCMC, mais precisamente o Metropolis Hastings, uma vez que a distribuição de $\beta_1|\beta_0,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n,Y,X,Z$ é intratável analiticamente. As estimativas são dadas pela média das cadeias geradas a partir da distribuição a posteriori de cada parâmetro do modelo.

Prioris normais informativas

No processo de estimação bayesiano o pesquisador pode incorporar efetivamente informações à análise. Esta incorporação se dá através de prioris informativas. Nessa secção apresentamos a estimação dos parâmetros através de prioris informativas.

Sejam

$$
\begin{aligned}
\beta_0 \sim N\left(a_0; b_0^2\right), \beta_1 \sim N\left(a_1; b_1^2\right), \ldots, \beta_n \sim N\left(a_n; b_n^2\right), \\
\gamma_0 \sim N\left(c_0; d_0^2\right), \gamma_1 \sim N\left(c_1; d_1^2\right), \ldots, \gamma_n \sim N\left(c_n; d_n^2\right),
\end{aligned}
$$

prioris informativas dos parâmetros, com isso tem-se que a função densidade de pro-

babilidade dos parâmetros é dada por

$$P\left(\beta_n\right) = \frac{1}{\sqrt{2\pi b_n^2}} exp\left(\frac{-\beta_n^2 + 2a_n\beta_n - a_n^2}{2b_n^2}\right),$$

$$P\left(\gamma_n\right) = \frac{1}{\sqrt{2\pi d_n^2}} exp\left(\frac{-\gamma_n^2 + 2c_n\beta_n - c_n^2}{2d_n^2}\right).$$

A posteriori é dada por

$$p\left(\beta_0, \ldots, \beta_n, \gamma_0, \ldots, \gamma_n | Y, X, Z\right) = L\left(Y, X, Z | \beta, \phi\right) p\left(\beta_0\right) \ldots p\left(\beta_n\right) p\left(\gamma_0\right) \ldots p\left(\gamma_n\right)$$

$$\begin{aligned}
&\propto \left[\Gamma\left(exp\left(\gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_n z_{in}\right)\right)\right]^n \\
&\times \left[\prod_{i=1}^{n} \Gamma\left(\frac{exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in} + \gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_n z_{in}\right)}{1 + exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \left[\prod_{i=1}^{n} \Gamma\left(\frac{exp\left(\gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_n z_{in}\right)}{1 + exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \prod_{i=1}^{n} \left[y_i \left(\frac{exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in} + \gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_n z_{in}\right)}{1 + exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times \prod_{i=1}^{n} \left[(1 - y_i) \left(\frac{exp\left(\gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_n z_{in}\right)}{1 + exp\left(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times exp\left(\frac{-\beta_0^2 + 2a_0\beta_0 - a_0^2}{2b_0^2}\right) \ldots exp\left(\frac{-\beta_n^2 + 2a_n\beta_n - a_n^2}{2b_n^2}\right) \\
&\times exp\left(\frac{-\gamma_0^2 + 2c_0\beta_0 - c_0^2}{2d_0^2}\right) \ldots exp\left(\frac{-\gamma_n^2 + 2c_n\beta_n - c_n^2}{2d_n^2}\right).
\end{aligned}$$

Como visto essas posteriores são intratáveis analiticamente, nesse caso faremos uso do procedimento computacionais como o Metropolis Hastings para retirar amostras da distribuição a posteriori. A distribuição a posteriori condicional de um determinado parâmetro é dado pela integral da posteriori conjunta em relação aos demais parâmetros.

Seja a distribuição a posteriori de β_1 dada por

$$P\left(\beta_1|\beta_0,\beta_2,\ldots,\beta_n,\gamma_0,\ldots,\gamma_n,Y,X,Z\right) \propto$$

$$\begin{aligned}
&\propto \left[\Gamma\left(exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)\right)\right]^n \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \left[\prod_{i=1}^{n}\Gamma\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}\right)\right]^{-1} \\
&\times \prod_{i=1}^{n}\left[y_i\left(\frac{exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}+\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times \prod_{i=1}^{n}\left[(1-y_i)\left(\frac{exp\left(\gamma_0+\gamma_1 z_{i1}+\ldots+\gamma_n z_{in}\right)}{1+exp\left(\beta_0+\beta_1 x_{i1}+\ldots+\beta_n x_{in}\right)}^{-1}\right)\right] \\
&\times exp\left(\frac{-\beta_1^2+2a_1\beta_1-a_1^2}{2b_1^2}\right).
\end{aligned}$$

as demais distribuições marginais a posteriori são análogas. Assim como visto anteriormente para gerar amostras dessa distribuição devemos adotar procedimento computacional, como por exemplo o Metropolis Hastings. As estimativas são dadas pela média das cadeias geradas a partir da distribuição a posteriori de cada parâmetro do modelo.

Outras distribuições a priori podem ser atribuídas a β e ϕ. Os coeficientes de regressão são correlacionados em modelos lineares generalizados e na regressão beta. Distribuição a priori que contemplem alguma estrutura de correlação nos coeficientes podem fornecer estimativas mais precisas que aquelas obtidas com o uso de distribuições a priori independentes, como a utilizada no presente trabalho. Bedrick et al.(1996) propuseram uma priori em modelos lineares generalizados em que a situação $P\left(\beta,\phi\right)=p\left(\beta\right)p\left(\phi\right)$ e os coeficientes são correlacionados , tal distribuição foi construída por Branscum et al. (2007) para os modelos de regressão beta.

No próximo capitulo serão abordadas as simulações do modelo de regressão beta visando verificar as propriedades dos estimadores do modelo

Capítulo 4

Simulações

Neste capítulo será demonstrada a qualidade assintótica dos estimadores de máxima verossimilhança (EMV), que consiste em estudar o seu comportamento quando o tamanho da amostra tende a infinito (comportamento assintótico).

Pode-se mostrar que, sob certas condições:

$$\hat{\beta} \sim N\left(\beta, I^{-1}\left(\beta\right)\right), n \to \infty.$$

A prova desse resultado pode ser encontrada no trabalho de Migon e Gamerman (1999), Ou seja, para n suficientemente grande, o estimador de máxima verossimilhança $\hat{\beta}$ tem distribuição aproximadamente $N\left(\beta, I^{-1}\left(\beta\right)\right)$, ou seja, o EMV é sempre assintoticamente não viesado e eficiente, já que sua esperança tende para β e sua variância tende para o limite inferior da desigualdade de Cramer-Rao. Além disso, ele é consistente já que $Var\left(\hat{\beta}\right) \to 0$.

Esse resultado pode ser generalizado para uma função $g\left(\beta\right)$, isto é:

$$\hat{\beta} \sim N\left(g\left(\beta\right), \frac{\left[g'\left(\beta\right)\right]^2}{I\left(\beta\right)}\right), \mathrm{n} \to \infty.$$

Portanto, o interesse dessa simulação é avaliar o viés do estimador de máxima verossimilhança, quando utilizada diferentes funções de ligação, e o poder dos testes de razão de verossimilhança e de Wald. Essas simulações foram realizadas sob o enfoque clássico e bayesiano.

No enfoque clássico, para as simulações realizadas adotando-se o parâmetro ϕ constante, foi utilizada a reparametrização proposta por Ferrari e Cribari-Neto (2004),

em que os parâmetros p e q são funções de μ (parâmetro de escala) e ϕ (parâmetro de precisão), dados por:

$$p_i = \mu_i \phi$$

e

$$q_i = (1 - \mu_i)\, \phi.$$

Sendo:

$$\mu_i = g\left(x_{ij}\right) = \sum_{j=1}^{n} \beta_j x_{ij},$$

em que $g\left(\cdot\right)$ é a função de ligação e $x_{i1}, x_{i2}, \ldots, x_{in}$ é o vetor de covariáveis do modelo observado. Foram consideradas as seguintes funções de ligação: logito, probito, complemento log-log e log-log.

Para o parâmetro ϕ variável, foi adotado o método desenvolvido por Paolino (2001), em que μ e ϕ são dados por:

$$\mu_i = g\left(x_i\right) = \sum_{j=1}^{n} \beta_j x_{ij}$$

e

$$\phi_i = \exp\left(\sum_{j=1}^{n} z_{ij}\gamma_i\right),$$

em que $x_{i1}, x_{i2}, \ldots, x_{in}$ e $z_{i1}, z_{i2}, \ldots, z_{in}$ são os vetores de covariáveis do modelo observado. Para a função $g\left(\cdot\right)$ foram consideradas: logito, probito, complemento log-log e log-log.

No enfoque bayesiano, o parâmetro ϕ também foi considerado nas duas abordagens, tanto como constante quanto variável. No entanto, nesse enfoque, a função de ligação adotada foi somente a logito.

Todos os resultados foram obtidos por simulação de Monte Carlo. Essa técnica tem por finalidade estimar a distribuição de uma estatística de interesse e, nesse caso, o que interessa é estimar a distribuição dos parâmetros da regressão. Esse método é uma abordagem paramétrica, uma vez que a amostra é extraída de uma população de distribuição conhecida. As simulações computacionais apresentadas neste trabalho foram realizadas no *software R*, versão 3.2. Foram utilizados os pacotes *betareg*, no enfoque clássico, e o *bayesianbetareg* para o bayesiano.

Este capítulo encontra-se estruturado da seguinte forma: Na seção 4.1 e 4.4 serão tratadas as simulações sob o enfoque clássico com o parâmetro de precisão constante e variável. Na seção 4.2 e 4.5 são analisadas as propriedades assintóticas dos estimadores de máxima verossimilhança abordando o parâmetro de precisão fixo e variável. Na seção 4.3 e 4.6 são abordados os aspectos do poder do teste TRV e de Wald, quando o parâmetro de precisão é fixo e variável. Na seção 4.7 são realizadas simulações sob o enfoque bayesiano e na seção 4.8 são abordados os aspectos de convergência e estacionariedade em cadeia de Markov.

4.1 Simulação sob enfoque clássico ϕ (constante)

Para essas simulações, foi considerada a reparametrização proposta por Ferrari e Cribari-Neto (2004), com o parâmetro de precisão constante (ϕ). A estrutura do modelo de regressão beta considerada na simulação é dada por:

$$\begin{aligned} g(\mu_k) &= \beta_0 + \beta_1 x_{k1} + \beta_2 x_{k2} + \beta_3 x_{k3} + \beta_4 x_{k4} + \beta_5 x_{k5}, \qquad k = 1, 2, \ldots, n. \\ g(\mu_k) &= \beta' X, \end{aligned}$$

em que as covariáveis $(x_1, x_2, x_3, x_4, x_5)$ apresentam as seguintes distribuições: $x_1 \sim U(0,1)$, $x_2 \sim N(0,1)$, $x_3 \sim Exp(3)$, $x_4 \sim Gamma(4,3)$ e $x_5 \sim Lognormal(0,1)$. Sendo assim, cada covariável foi gerada de uma distribuição diferente. Para a geração da variável resposta Y_i foram consideradas quatro diferentes funções de ligação em que o parâmetro de escala μ pode ser calculado, são elas:

Função Logito:

$$\mu_i = \frac{Exp(\beta' X)}{1 + Exp(\beta' X)}.$$

Função Probito:

$$\mu_i = \Phi(\beta' X).$$

Função Complemento log-log:

$$\mu_i = 1 - \exp(-exp(\beta' X)).$$

Função log-log:

$$\mu_i = \exp(-exp(-(\beta' X))).$$

Os valores admitidos com verdadeiros para os parâmetros β foram: $\beta_0 = 0,5$, $\beta_1 = -0.02$, $\beta_2 = 0.30$, $\beta_3 = -0.1$, $\beta_4 = 0.05$, $\beta_5 = -0.07$ e $\phi = 8.20$.

A matriz dos regressores permaneceu constante durante o experimento e os seguintes tamanhos amostrais foram adotados: 15, 50, 200 e 1000. Para que não houvesse problema na geração de cada covariável devido aos diferentes tamanhos de amostras, fixou-se uma semente na geração e só se modificou o tamanho das amostras geradas.

Para cada réplica foi gerada uma amostra aleatória da variável-resposta $y_1, y_2, y_3, \ldots, y_n$ com $y_i \sim beta((p_i, q_i))$, em que:

$$p_i = \mu_i \phi$$

e

$$q_i = (1 - \mu_i)\, \phi.$$

Para os testes de razão de verossimilhança e de Wald, as hipóteses testadas foram:

$$\begin{cases} H_0 : \beta_1 = \beta_3 = \beta_5 = 0, \\ H_1 : \beta_i \neq 0, \text{para algum} \quad i = 1, 3, 5. \end{cases}$$

4.2 Análise assintótica dos estimadores com precisão constante (ϕ)

Nesta seção serão apresentados os resultados das simulações referentes aos estimadores de máxima verossimilhança do modelo de regressão beta no qual o parâmetro de precisão ϕ foi adotado como fixo.

Espera-se que cada estimador dos parâmetros do modelo apresente distribuição normal, de acordo com as propriedades assintóticas. Sabe-se que essa distribuição é simétrica em torno da média, ou seja, seu coeficiente de assimetria tende a zero à medida que se aumenta a amostra.

A tabela 4.1 apresenta informações referentes às estimativas dos parâmetros, onde n representa o tamanho da amostra, $\beta's$ os valores reais, e em cada função de ligação estão as estimativas pontuais e seus respectivos desvios médios.

Tabela 4.1: Estimativas pontuais - parâmetro ϕ constante.

n	β	Valores verdadeiros	Função de ligação			
			Logito	Proito	Cloglog	Loglog
15	β_0	0,5	0,51 (0,486)	0,511 (0,305)	0,538 (0,278)	0,512 (0,353)
	β_1	-0,02	0,066 (0,689)	-0,005 (0,432)	-0,009 (0,394)	-0,009 (0,495)
	β_2	0,3	0,303 (0,216)	0,296 (0,135)	0,315 (0,125)	0,305 (0,153)
	β_3	-0,15	-0,149 (0,596)	-0,149 (0,371)	-0,176 (0,338)	-0,163 (0,428)
	β_4	0,05	0,068 (0,25)	0,054 (0,156)	0,044 (0,138)	0,049 (0,184)
	β_5	-0,07	-0,071 (0,128)	-0,078 (0,08)	-0,08 (0,072)	-0,069 (0,093)
	γ	8,2	17,363 (6,19)	17,534 (6,28)	17,511 (6,403)	17,173 (6,111)
50	β_0	0,5	0,487 (0,295)	0,508 (0,185)	0,508 (0,167)	0,498 (0,213)
	β_1	-0,02	-0,002 (0,364)	-0,021 (0,228)	-0,014 (0,207)	-0,008 (0,261)
	β_2	0,3	0,302 (0,123)	0,299 (0,077)	0,304 (0,068)	0,307 (0,09)
	β_3	-0,15	-0,161 (0,265)	-0,165 (0,165)	-0,158 (0,155)	-0,146 (0,185)
	β_4	0,05	0,061 (0,142)	0,054 (0,088)	0,053 (0,082)	0,049 (0,099)
	β_5	-0,07	-0,074 (0,075)	-0,073 (0,047)	-0,075 (0,044)	-0,07 (0,052)
	γ	8,2	9,758 (1,868)	9,781 (1,886)	9,779 (1,941)	9,764 (1,866)
200	β_0	0,5	0,507 (0,155)	0,503 (0,097)	0,501 (0,089)	0,502 (0,111)
	β_1	-0,02	-0,023 (0,162)	-0,017 (0,101)	-0,015 (0,094)	-0,023 (0,114)
	β_2	0,3	0,299 (0,047)	0,301 (0,029)	0,299 (0,027)	0,3 (0,033)
	β_3	-0,15	-0,153 (0,136)	-0,152 (0,085)	-0,149 (0,081)	-0,148 (0,094)
	β_4	0,05	0,049 (0,073)	0,047 (0,046)	0,049 (0,042)	0,051 (0,052)
	β_5	-0,07	-0,07 (0,028)	-0,069 (0,018)	-0,071 (0,017)	-0,07 (0,019)
	γ	8,2	8,552 (0,813)	8,555 (0,82)	8,544 (0,845)	8,54 (0,812)
1000	β_0	0,5	0,5 (0,071)	0,502 (0,044)	0,496 (0,041)	0,5 (0,05)
	β_1	-0,02	-0,022 (0,074)	-0,021 (0,046)	-0,018 (0,042)	-0,018 (0,052)
	β_2	0,3	0,301 (0,022)	0,3 (0,014)	0,299 (0,012)	0,3 (0,015)
	β_3	-0,15	-0,149 (0,064)	-0,152 (0,04)	-0,15 (0,038)	-0,149 (0,045)
	β_4	0,05	0,051 (0,034)	0,05 (0,021)	0,052 (0,019)	0,049 (0,024)
	β_5	-0,07	-0,07 (0,011)	-0,07 (0,007)	-0,07 (0,007)	-0,07 (0,007)
	γ	8,2	8,257 (0,35)	8,266 (0,354)	8,281 (0,365)	8,269 (0,351)

Fonte: Elaborada pelo autor

Observa-se na tabela 4.2, de modo geral, que as estimativas estão muito próximas dos verdadeiros valores dos parâmetros populacionais, no entanto, as estimativas do parâmetro ϕ em amostras inferiores a $n = 200$ apresentam uma leve assimetria à direita, mas, à medida que se aumenta a amostra, verifica-se que tanto a estimativa pontual como as intervalares melhoram, e há uma boa precisão dos estimadores com intervalos relativamente pequenos.

Tabela 4.2: Intervalos de confiança dos parâmetros - ϕ constante

n	β	Valores verdadeiros	Função de ligação			
			Logito	Proito	Cloglog	Loglog
15	β_0	0,5	[0,469 ; 0,841]	[0,485 ; 0,722]	[0,515 ; 0,727]	[0,482 ; 0,752]
	β_1	-0,02	[-0,122 ; 0,392]	[-0,04 ; 0,284]	[-0,041 ; 0,251]	[-0,05 ; 0,329]
	β_2	0,3	[0,286 ; 0,447]	[0,285 ; 0,386]	[0,304 ; 0,4]	[0,292 ; 0,412]
	β_3	-0,15	[-0,199 ; 0,261]	[-0,179 ; 0,095]	[-0,205 ; 0,058]	[-0,198 ; 0,124]
	β_4	0,05	[0,047 ; 0,237]	[0,041 ; 0,16]	[0,032 ; 0,139]	[0,034 ; 0,177]
	β_5	-0,07	[-0,081 ; 0,014]	[-0,085 ; -0,026]	[-0,085 ; 0,033]	[-0,076 ; -0,008]
	γ	8,2	[16,691 ; 22,853]	[16,805 ; 23,488]	[16,865 ; 22,782]	[16,532 ; 22,407]
50	β_0	0,5	[0,466 ; 0,58]	[0,495 ; 0,564]	[0,497 ; 0,556]	[0,484 ; 0,56]
	β_1	-0,02	[-0,026 ; 0,109]	[-0,036 ; 0,048]	[-0,028 ; 0,048]	[-0,024 ; 0,067]
	β_2	0,3	[0,294 ; 0,339]	[0,294 ; 0,322]	[0,3 ; 0,325]	[0,301 ; 0,334]
	β_3	-0,15	[-0,179 ; 0,083]	[-0,177 ; 0,114]	[-0,168 ; -0,11]	[-0,159 ; -0,091]
	β_4	0,05	[0,051 ; 0,104]	[0,048 ; 0,081]	[0,048 ; 0,075]	[0,043 ; 0,078]
	β_5	-0,07	[-0,079 ; -0,052]	[-0,076 ; -0,059]	[-0,078 ; -0,063]	[-0,073 ; -0,054]
	γ	8,2	[9,622 ; 10,364]	[9,642 ; 10,403]	[9,647 ; 10,369]	[9,634 ; 10,348]
200	β_0	0,5	[0,497 ; 0,528]	[0,497 ; 0,518]	[0,495 ; 0,513]	[0,495 ; 0,517]
	β_1	-0,02	[-0,033 ; -0,001]	[-0,023 ; -0,002]	[-0,021 ; -0,002]	[-0,03 ; -0,008]
	β_2	0,3	[0,296 ; 0,306]	[0,299 ; 0,305]	[0,297 ; 0,302]	[0,298 ; 0,305]
	β_3	-0,15	[-0,162 ; -0,134]	[-0,158 ; -0,14]	[-0,154 ; -0,138]	[-0,154 ; -0,135]
	β_4	0,05	[0,044 ; 0,059]	[0,044 ; 0,053]	[0,047 ; 0,055]	[0,047 ; 0,058]
	β_5	-0,07	[-0,072 ; -0,066]	[-0,07 ; -0,067]	[-0,072 ; -0,069]	[-0,071 ; -0,068]
	γ	8,2	[8,499 ; 8,67]	[8,504 ; 8,669]	[8,491 ; 8,665]	[8,49 ; 8,652]
1000	β_0	0,5	[0,496 ; 0,505]	[0,499 ; 0,505]	[0,493 ; 0,498]	[0,497 ; 0,503]
	β_1	-0,02	[-0,027 ; -0,018]	[-0,024 ; -0,018]	[-0,02 ; -0,015]	[-0,021 ; -0,014]
	β_2	0,3	[0,299 ; 0,302]	[0,299 ; 0,301]	[0,298 ; 0,299]	[0,299 ; 0,301]
	β_3	-0,15	[-0,153 ; -0,145]	[-0,154 ; -0,149]	[-0,152 ; -0,148]	[-0,152 ; -0,147]
	β_4	0,05	[0,049 ; 0,053]	[0,048 ; 0,051]	[0,051 ; 0,053]	[0,048 ; 0,051]
	β_5	-0,07	[-0,071 ; -0,07]	[-0,07 ; -0,069]	[-0,07 ; -0,069]	[-0,071 ; -0,07]
	γ	8,2	[8,236 ; 8,278]	[8,244 ; 8,288]	[8,258 ; 8,304]	[8,248 ; 8,291]

Fonte: Elaborada pelo autor

Conforme demonstrado na Tabela 4.3, com o aumento da amostra os coeficientes de assimetria dos estimadores aproximam-se de zero, como esperado pela propriedade dos estimadores de máxima verossimilhança, que assintóticamente apresentam distribuição normal.

Tabela 4.3: Coeficiente de assimetria ϕ constante

Link	n	Parâmetros						
		β_0	β_1	β_2	β_3	β_4	β_5	γ
Logito	15	0,17	0,00	0,13	-0,21	-0,03	-0,06	3,57
	50	-0,05	0,07	0,09	0,02	0,02	0,14	0,81
	200	0,04	-0,06	0,07	-0,10	0,02	0,10	0,39
	1000	-0,07	0,09	0,03	0,06	-0,02	0,03	0,18
Probito	15	0,15	-0,18	0,10	-0,09	-0,04	0,00	3,54
	50	0,00	-0,07	0,10	0,07	0,00	-0,14	1,12
	200	0,06	-0,05	0,03	-0,03	0,07	-0,02	0,38
	1000	-0,14	0,03	0,10	0,06	0,05	0,02	0,17
Cloglog	15	-0,01	-0,09	-0,01	-0,07	-0,23	-0,22	2,78
	50	-0,13	0,05	-0,08	-0,10	0,01	-0,13	0,91
	200	-0,02	0,01	-0,04	0,01	0,04	-0,03	0,37
	1000	0,10	0,04	-0,07	0,08	-0,16	-0,12	0,13
Loglog	15	0,19	-0,20	0,00	-0,01	0,22	0,19	3,42
	50	0,10	-0,02	0,03	0,02	0,03	0,18	0,85
	200	-0,04	-0,08	0,00	-0,03	0,02	0,05	0,40
	1000	-0,01	-0,09	-0,09	-0,08	0,01	0,13	0,17

Fonte: Elaborada pelo autor

Percebe-se que no modelo de regressão beta, em que a média (μ) é modelada por meio de estrutura de regressão, independente da função de ligação adotada os estimadores de máxima verossimilhança para os parâmetros de regressão apresentam boas propriedades para amostras finitas, uma vez que esses se mostraram quase não viciados e com distribuição bem próxima da distribuição normal.

4.3 Poder do teste para ϕ constante

Nesta seção serão apresentados resultados de simulação para avaliar o poder do teste, tendo sido utilizadas as estatísticas Teste da razão de verossimilhança (TRV) e Teste de Wald. O Poder do Teste tem como objetivo conhecer o quanto o teste de hipóteses controla um erro do tipo II, ou qual a probabilidade de rejeitar a hipótese nula se esta realmente for falsa. Nesse sentido, espera-se que mesmo em amostras pequenas essas estatísticas apresentem bom poder de discriminação.

A Tabela 4.4 apresenta as estatísticas observadas:

Tabela 4.4: Poder do teste

Link	α (%)	n=15		n=50		n=200		n=1000	
		TRV (%)	Wald (%)	TRV (%)	Wald(%)	TRV(%)	Wald(%)	TRV(%)	Wald(%)
Logito	10,00	41,70	47,80	41,70	43,20	82,00	82,80	100,00	100,00
	5,00	30,10	37,30	28,60	31,90	72,70	73,60	100,00	100,00
	1,00	13,10	20,00	9,90	13,20	52,00	54,40	100,00	100,00
	0,50	9,70	16,70	6,60	8,50	42,70	44,90	100,00	100,00
Probito	10,00	41,70	47,80	41,70	43,20	82,00	82,80	100,00	100,00
	5,00	30,10	37,30	28,60	31,90	72,70	73,60	100,00	100,00
	1,00	13,10	20,00	9,90	13,20	52,00	54,40	100,00	100,00
	0,50	9,70	16,70	6,60	8,50	42,70	44,90	100,00	100,00
Cloglog	10,00	41,70	47,80	41,70	43,20	82,00	82,80	100,00	100,00
	5,00	30,10	37,30	28,60	31,90	72,70	73,60	100,00	100,00
	1,00	13,10	20,00	9,90	13,20	52,00	54,40	100,00	100,00
	0,50	9,70	16,70	6,60	8,50	42,70	44,90	100,00	100,00
Loglog	10,00	41,70	47,80	41,70	43,20	82,00	82,80	100,00	100,00
	5,00	30,10	37,30	28,60	31,90	72,70	73,60	100,00	100,00
	1,00	13,10	20,00	9,90	13,20	52,00	54,40	100,00	100,00
	0,50	9,70	16,70	6,60	8,50	42,70	44,90	100,00	100,00

Fonte: Elaborada pelo autor

As seguintes hipóteses foram testadas:

$$\begin{cases} H_0 : \beta_1 = \beta_3 = \beta_5 = 0, \\ H_1 : \beta_i \neq 0, \text{para algum} \quad i = 1, 3, 5. \end{cases}$$

Para todas as simulações, foram adotadas as mesmas hipóteses e o número de replicação de Monte Carlo foi fixado em 1.000. O poder do teste é afetado por três fatores: o tamanho da amostra (espera-se que quanto maior, maior seja o poder

do teste), o nível de significância (quanto maior, maior o poder do teste, uma vez que fica reduzida a região de não rejeição da hipótese nula) e o verdadeiro valor do parâmetro (quanto maior for a diferença entre o verdadeiro valor do parâmetro e o valor especificado na hipótese nula, maior o poder do teste).

Verificou-se que a estatística TRV apresentou baixo poder do teste em amostras muito pequenas n = 15 e à medida que se aumenta a amostra o poder do teste aumenta significativamente. A estatística TRV tem, em situações particulares, distribuição limite qui-quadrado (χ^2). Contudo, em amostras pequenas, a distribuição limite (χ^2) pode fornecer uma aproximação pobre à distribuição exata da estatística de razão de verossimilhança, implicando distorção do tamanho do teste. Com o objetivo de melhorar a aproximação da distribuição da estatística de teste TRV pela distribuição (χ^2), Bayer (2011) implementou a correção de Bartlett para a estatística TRV em modelos de regressão beta. O autor demonstra que em amostras pequenas a estatística TRV apresenta distorções no tamanho do teste.

À medida que se aumenta a amostra, as estatísticas do teste aproximam-se da distribuição χ^2_2. Percebe-se que o aumento da amostra está intimamente relacionado ao poder do teste dessas estatísticas.

Como observado o teste de wald apresentou melhor poder do teste que o TRV, independente do tamanho amostral e do tipo de função de ligação utilizada, esse fato era esperado uma vez que o TRV apresenta distorções nos modelos MRB.

Em linhas gerais, os estimadores de máxima verossimilhança utilizados na regressão beta apresentam boas propriedades, no entanto, o parâmetro de precisão desse modelo em amostras muito pequenas apresenta viés à direita.

4.4 Simulação sob enfoque clássico ϕ variável

Na seção 4.2, observou-se que, de modo geral, com o aumento da amostra ($n > 50$) os estimadores dos parâmetros do modelo apresentavam boas propriedades, no entanto, em amostras pequenas, seu parâmetro de precisão apresentava certa assimetria à direita. Com o intuito de observar se há melhora nas estimativas do parâmetro de precisão com a inclusão de covariáveis, foram realizadas simulações nas quais esse parâmetro é adotado como variável.

Uma vez aumentando o número de estimadores, foi necessário aumentar o tamanho da amostra inicial de 15 para 30, as demais amostras mantiveram-se no mesmo tamanho. A reparametrização na qual os parâmetros da distribuição beta são obtidos por meio dos termos da média e precisão foi mantida. Para estimação do parâmetro μ_i, foram consideradas as quatro funções de ligação já apresentadas (logito, probito, complemento loglog e loglog) e para a estimação do ϕ foi considerado o método introduzido por Simas, Barreto-Souza e Rocha (2010), em que, adicionalmente à modelagem da média, define-se uma estrutura de regressão para o parâmetro de precisão. As funções de ligação mais comumente empregadas são o logaritmo, $h(\phi_i) = \log(\phi_i)$ e a raiz quadrada, $h(\phi_i) = \sqrt{\phi_i}$. Nesta simulação, a função logarítmica foi utilizada como função de ligação e sua inversa é dada por:

$$\phi_i = \exp\left(\sum_{j=1}^{n} x_{ij}\gamma_i\right).$$

4.5 Análise assintótica dos estimadores com precisão ϕ variável

Os métodos utilizados para realizar as análises dos estimadores foram iguais aos procedimentos usados quando ϕ é constante. Na Tabela 4.5 estão as estimativas pontuais e a média dos desvios na Tabela 4.6 encontram-se as estimativas intervalares observadas.

Tabela 4.5: Estimativas pontuais - parâmetro ϕ variável.

n	β	Valores verdadeiros	Função de ligação			
			Logito	Proito	Cloglog	Loglog
30	β_0	0,50	0,511 (0,347)	0,521 (0,219)	0,514 (0,205)	0,502 (0,252)
	β_1	-0,02	-0,024 (0,394)	-0,017 (0,248)	-0,031 (0,237)	-0,012 (0,283)
	β_2	0,30	0,303 (0,096)	0,3 (0,06)	0,308 (0,057)	0,302 (0,071)
	β_3	-0,15	-0,162 (0,271)	-0,175 (0,17)	-0,163 (0,162)	-0,15 (0,191)
	β_4	0,05	0,05 (0,117)	0,043 (0,073)	0,052 (0,072)	0,046 (0,081)
	β_5	-0,07	-0,067 (0,043)	-0,07 (0,027)	-0,07 (0,027)	-0,07 (0,03)
	γ_0	0,40	0,266 (0,514)	0,205 (0,516)	0,264 (0,534)	0,235 (0,511)
	γ_1	3,80	4,81 (0,975)	4,94 (0,976)	4,793 (0,991)	4,832 (0,973)
50	β_0	0,50	0,493 (0,262)	0,496 (0,165)	0,511 (0,158)	0,493 (0,185)
	β_1	-0,02	-0,011 (0,292)	-0,015 (0,184)	-0,031 (0,177)	-0,01 (0,208)
	β_2	0,30	0,308 (0,086)	0,303 (0,054)	0,302 (0,05)	0,302 (0,062)
	β_3	-0,15	-0,154 (0,184)	-0,148 (0,114)	-0,151 (0,111)	-0,146 (0,126)
	β_4	0,05	0,049 (0,085)	0,052 (0,053)	0,051 (0,051)	0,053 (0,059)
	β_5	-0,07	-0,069 (0,039)	-0,07 (0,024)	-0,069 (0,024)	-0,071 (0,027)
	γ_0	0,40	0,336 (0,389)	0,317 (0,392)	0,337 (0,412)	0,326 (0,387)
	γ_1	3,80	4,296 (0,715)	4,348 (0,716)	4,3 (0,732)	4,327 (0,713)
200	β_0	0,50	0,5 (0,138)	0,504 (0,087)	0,504 (0,082)	0,499 (0,097)
	β_1	-0,02	-0,014 (0,143)	-0,023 (0,09)	-0,021 (0,086)	-0,019 (0,1)
	β_2	0,30	0,3 (0,032)	0,302 (0,021)	0,299 (0,02)	0,301 (0,023)
	β_3	-0,15	-0,151 (0,092)	-0,155 (0,058)	-0,153 (0,056)	-0,15 (0,063)
	β_4	0,05	0,048 (0,048)	0,05 (0,03)	0,05 (0,029)	0,051 (0,034)
	β_5	-0,07	-0,071 (0,02)	-0,07 (0,013)	-0,071 (0,012)	-0,071 (0,014)
	γ_0	0,40	0,398 (0,182)	0,409 (0,183)	0,417 (0,192)	0,408 (0,182)
	γ_1	3,80	3,883 (0,319)	3,868 (0,319)	3,856 (0,326)	3,87 (0,318)
1000	β_0	0,50	0,494 (0,067)	0,5 (0,042)	0,497 (0,04)	0,497 (0,047)
	β_1	-0,02	-0,015 (0,066)	-0,02 (0,042)	-0,018 (0,04)	-0,017 (0,046)
	β_2	0,30	0,3 (0,015)	0,3 (0,01)	0,3 (0,009)	0,3 (0,011)
	β_3	-0,15	-0,147 (0,048)	-0,149 (0,03)	-0,147 (0,029)	-0,15 (0,034)
	β_4	0,05	0,05 (0,024)	0,05 (0,015)	0,05 (0,014)	0,051 (0,017)
	β_5	-0,07	-0,069 (0,008)	-0,07 (0,005)	-0,07 (0,005)	-0,07 (0,005)
	γ_0	0,40	0,408 (0,08)	0,408 (0,081)	0,405 (0,085)	0,407 (0,08)
	γ_1	3,80	3,801 (0,142)	3,806 (0,142)	3,806 (0,146)	3,802 (0,142)

Fonte: Elaborada pelo autor

Tabela 4.6: Intervalos de confiança dos parâmetros - ϕ variável

n	β	Valores verdadeiros	Função de ligação			
			Logito	Proito	Cloglog	Loglog
30	β_0	0,50	[0,485 ; 0,538]	[0,504 ; 0,538]	[0,499 ; 0,53]	[0,484 ; 0,521]
	β_1	-0,02	[-0,056 ; 0,007]	[-0,036 ; 0,003]	[-0,049 ; -0,013]	[-0,033 ; 0,009]
	β_2	0,30	[0,295 ; 0,311]	[0,295 ; 0,305]	[0,303 ; 0,312]	[0,296 ; 0,308]
	β_3	-0,15	[-0,183 ; -0,141]	[-0,188 ; -0,162]	[-0,176 ; -0,15]	[-0,164 ; -0,135]
	β_4	0,05	[0,04 ; 0,06]	[0,037 ; 0,049]	[0,046 ; 0,058]	[0,04 ; 0,052]
	β_5	-0,07	[-0,071 ; -0,063]	[-0,072 ; -0,068]	[-0,072 ; -0,068]	[-0,072 ; -0,067]
	γ_0	0,40	[0,219 ; 0,312]	[0,16 ; 0,25]	[0,217 ; 0,311]	[0,192 ; 0,278]
	γ_1	3,80	[4,71 ; 4,911]	[4,841 ; 5,04]	[4,693 ; 4,893]	[4,735 ; 4,929]
50	β_0	0,50	[0,475 ; 0,511]	[0,484 ; 0,508]	[0,499 ; 0,522]	[0,48 ; 0,506]
	β_1	-0,02	[-0,031 ; 0,009]	[-0,028 ; -0,002]	[-0,044 ; -0,018]	[-0,024 ; 0,005]
	β_2	0,30	[0,301 ; 0,314]	[0,299 ; 0,307]	[0,298 ; 0,305]	[0,297 ; 0,307]
	β_3	-0,15	[-0,167 ; -0,141]	[-0,157 ; -0,139]	[-0,159 ; -0,143]	[-0,155 ; -0,136]
	β_4	0,05	[0,042 ; 0,056]	[0,048 ; 0,056]	[0,047 ; 0,055]	[0,049 ; 0,057]
	β_5	-0,07	[-0,072 ; -0,066]	[-0,072 ; -0,068]	[-0,071 ; -0,067]	[-0,073 ; -0,069]
	γ_0	0,40	[0,308 ; 0,364]	[0,287 ; 0,347]	[0,305 ; 0,368]	[0,297 ; 0,355]
	γ_1	3,80	[4,24 ; 4,352]	[4,289 ; 4,407]	[4,239 ; 4,362]	[4,269 ; 4,386]
200	β_0	0,50	[0,491 ; 0,509]	[0,499 ; 0,509]	[0,499 ; 0,509]	[0,493 ; 0,505]
	β_1	-0,02	[-0,024 ; -0,005]	[-0,028 ; -0,017]	[-0,026 ; -0,016]	[-0,026 ; -0,013]
	β_2	0,30	[0,298 ; 0,303]	[0,301 ; 0,304]	[0,298 ; 0,301]	[0,3 ; 0,302]
	β_3	-0,15	[-0,157 ; -0,145]	[-0,159 ; -0,151]	[-0,157 ; -0,15]	[-0,155 ; -0,146]
	β_4	0,05	[0,045 ; 0,051]	[0,048 ; 0,052]	[0,048 ; 0,052]	[0,049 ; 0,054]
	β_5	-0,07	[-0,072 ; -0,069]	[-0,071 ; -0,069]	[-0,071 ; -0,07]	[-0,072 ; -0,07]
	γ_0	0,40	[0,387 ; 0,409]	[0,397 ; 0,42]	[0,404 ; 0,429]	[0,397 ; 0,42]
	γ_1	3,80	[3,864 ; 3,903]	[3,848 ; 3,888]	[3,834 ; 3,878]	[3,849 ; 3,891]
1000	β_0	0,50	[0,49 ; 0,499]	[0,497 ; 0,503]	[0,495 ; 0,5]	[0,494 ; 0,499]
	β_1	-0,02	[-0,019 ; -0,011]	[-0,023 ; -0,018]	[-0,021 ; -0,016]	[-0,02 ; -0,014]
	β_2	0,30	[0,299 ; 0,301]	[0,299 ; 0,301]	[0,3 ; 0,301]	[0,299 ; 0,301]
	β_3	-0,15	[-0,15 ; -0,144]	[-0,151 ; -0,147]	[-0,149 ; -0,146]	[-0,152 ; -0,148]
	β_4	0,05	[0,049 ; 0,052]	[0,049 ; 0,051]	[0,049 ; 0,051]	[0,05 ; 0,052]
	β_5	-0,07	[-0,07 ; -0,069]	[-0,07 ; -0,07]	[-0,07 ; -0,07]	[-0,07 ; -0,069]
	γ_0	0,40	[0,403 ; 0,412]	[0,403 ; 0,413]	[0,399 ; 0,41]	[0,402 ; 0,412]
	γ_1	3,80	[3,792 ; 3,81]	[3,797 ; 3,815]	[3,797 ; 3,816]	[3,794 ; 3,811]

Fonte: Elaborada pelo autor

Observa-se que as estimativas pontuais continuam próximas dos verdadeiros valores, no entanto, em amostras pequenas, mesmo adotando o parâmetro de precisão como variável. As estimativas intervalares e pontuais continuam apresentando viés e leve assimetria à direita em amostras pequenas. Com o aumento da amostra ($n > 50$), essa assimetria é bem pequena, como demostrada na Tabela 4.7.

Tabela 4.7: Coeficiente de assimetria ϕ variável.

Link	n	Parâmetros							
		β_0	β_1	β_2	β_3	β_4	β_5	γ_0	γ_1
Logito	15	-0,06	0,06	-0,06	0,03	-0,15	0,09	0,21	0,29
	50	0,00	-0,05	-0,05	-0,13	0,15	-0,07	0,27	0,07
	200	0,00	-0,04	0,10	0,01	-0,04	-0,07	0,18	-0,06
	1000	-0,02	0,09	0,03	-0,04	-0,06	0,06	0,00	0,06
Probito	15	-0,15	-0,04	0,00	0,03	-0,10	0,04	0,08	0,45
	50	0,09	-0,07	0,15	-0,08	0,02	-0,02	0,11	0,16
	200	-0,03	0,11	0,05	0,01	0,04	-0,02	0,17	0,10
	1000	-0,04	-0,15	-0,11	-0,06	0,01	-0,05	0,03	-0,11
Cloglog	15	-0,20	0,03	-0,04	0,06	-0,02	-0,04	0,14	0,32
	50	0,09	-0,04	0,02	-0,15	-0,02	-0,09	0,23	0,08
	200	0,04	-0,12	-0,02	-0,02	0,06	-0,11	0,16	0,04
	1000	0,04	-0,09	-0,12	0,02	0,12	0,01	0,05	0,09
Loglog	15	0,01	-0,01	-0,01	0,10	-0,25	0,08	-0,05	0,61
	50	0,15	-0,11	0,22	0,00	0,02	0,04	0,14	0,19
	200	-0,10	0,11	0,09	0,06	0,09	0,14	0,18	-0,07
	1000	-0,06	0,03	0,02	-0,10	0,01	0,08	0,07	0,10

Fonte: Elaborada pelo autor

4.6 Poder do Teste para ϕ variável

A inclusão de variável para melhor explicar o parâmetro de precisão acarretou em aumento significativo do poder do teste TRV, o mesmo pode ser observado no teste de Wald, conforme Tabela 4.8.

Tabela 4.8: Poder do teste

Link	α (%)	n=15		n=50		n=200		n=1000	
		TRV (%)	Wald (%)	TRV (%)	Wald(%)	TRV(%)	Wald(%)	TRV(%)	Wald(%)
Logito	10,00	43,90	54,30	43,00	48,90	96,00	96,80	100,00	100,00
	5,00	32,10	45,40	30,80	40,20	91,50	92,60	100,00	100,00
	1,00	13,60	28,70	14,00	22,50	77,00	80,20	100,00	100,00
	0,50	9,60	24,70	9,50	18,00	68,80	72,90	100,00	100,00
Probito	10,00	67,10	77,10	74,10	79,30	100,00	100,00	100,00	100,00
	5,00	56,60	69,40	63,40	71,90	100,00	100,00	100,00	100,00
	1,00	34,00	54,50	38,80	54,60	99,90	99,90	100,00	100,00
	0,50	25,40	49,30	30,20	47,00	99,70	99,80	100,00	100,00
Cloglog	10,00	69,30	76,90	76,00	80,50	100,00	100,00	100,00	100,00
	5,00	55,40	69,10	64,80	72,30	100,00	100,00	100,00	100,00
	1,00	33,30	51,90	40,70	54,40	100,00	100,00	100,00	100,00
	0,50	25,90	47,20	31,60	48,50	100,00	100,00	100,00	100,00
Loglog	10,00	57,70	68,50	68,10	74,70	100,00	100,00	100,00	100,00
	5,00	46,70	59,60	56,40	66,20	100,00	100,00	100,00	100,00
	1,00	24,90	43,50	32,10	46,60	98,60	99,00	100,00	100,00
	0,50	18,40	39,10	24,20	39,60	97,70	98,30	100,00	100,00

Fonte: Elaborada pelo autor

Como observado, a estatística TRV apresentou poder do teste maior quando considerado o ϕ variando, mesmo em amostras muito pequenas $n = 30$. À medida que se aumenta a amostra, o poder do teste aumenta de forma significativa. Também se verifica no teste de Wald, onde mesmo em amostras pequenas suas estatísticas são consideravelmente maiores do que com o parâmetro ϕ constante.

Considerações

A propriedade assintótica dos estimadores de máxima verossimilhança pode ser comprovada independente do *link* de função utilizado. Entretanto, observa-se que o poder do teste em amostras relativamente pequenas $n = 30$ foi melhor quando

utilizados os *links* de função complemento loglog e loglog. Os testes TRV e Wald mostraram poder do teste mais eficiente nas simulações dos modelos que utilizaram o parâmetro de precisão variável em relação aos que apresentam ϕ constante, mesmo em amostras pequenas ($n = 30$). As simulações mostraram que os modelos com o parâmetro ϕ variando apresentam propriedades estatísticas melhores do que aqueles com ϕ constante.

4.7 Simulação sob enfoque bayesiano

Nesta seção são apresentadas as simulações adotadas sob o enfoque bayesiano. Os modelos de regressão beta e os conjuntos de parâmetros média e precisão são definidos como em Cepeda-Cuervo (2001). Para estimação dos parâmetros do modelo de regressão beta sob o enfoque bayesiano, é necessária a especificação de uma distribuição *a priori*. Nessa simulação, utilizou-se a distribuição normal como *priori* para os parâmetros. A escolha de uma normal como *priori* inviabiliza a obtenção de forma analítica da distribuição *a posteriori*. Nessas situações, faz-se necessário o uso de algoritmos computacionais para a obtenção de amostras dos parâmetros. O algoritmo utilizado nessa simulação foi proposto por Cepeda-Cuervo (2001).

Para a simulação, foi utilizado o *software R* e o pacote *Bayesianbetareg*, o qual apresenta a implementação computacional do método descrito por Cepeda-Cuervo (2001). O pacote fornece ao usuário estimativas dos parâmetros β e γ, e seus desvios-padrão. Ele também fornece os valores ajustados de Y, resíduo, variância e matriz com as amostras das *posterioris*. O pacote *R-Bayesianbetareg* tem nove outras funções, que permitem ao usuário obter valores de AIC, BIC, critério de desvio (deviance), gráficos de quatro tipos de resíduos e diagnósticos gráficos para modelos de regressão beta.

Para a simulação, foram definidos os seguintes critérios: a função de ligação utilizada para o parâmetro da média μ foi a logito.O parâmetro de precisão ϕ foi analisado de duas formas distintas. Inicialmente, foram construídos modelos com ϕ constante e tamanhos amostrais $n = 30$, $n = 50$ e $n = 100$ e, posteriormente, com ϕ variando.Nesse caso, adotou-se a função de ligação de ϕ sendo uma exponencial, conforme

definido por Cepeda-Cuervo (2001).

$$\mu_i = \frac{Exp\left(\beta_0 + \beta_1 x_{k1} + \beta_2 x_{k2} + \beta_3 x_{k3} + \beta_4 x_{k4} + \beta_5 x_{k5}\right)}{1 + Exp\left(\beta_0 + \beta_1 x_{k1} + \beta_2 x_{k2} + \beta_3 x_{k3} + \beta_4 x_{k4} + \beta_5 x_{k5}\right)}$$

e

$$\phi_i = \exp\left(\sum_{j=1}^{n} x_{ij}\gamma_i\right).$$

Durante a simulação, adotou-se um *burn-in* de 30%. Demanda-se algum tempo para a convergência dos valores da amostra dos parâmetros, e, com isso, faz-se necessária a remoção dos que ainda não convergiram para evitar viés. É importante queimar esses valores iniciais, pois obtemos os valores de interesse quando a simulação está estacionária.

O *jump* utilizado foi de tamanho 3, normalmente valores próximos são autocorrelacionados. Com isso, foi necessário obter valores saltados dos parâmetros, ou seja, a cada 3 iterações, guardou-se o valor do parâmetro e posteriormente calculou-se a média. O número de iterações foi definido em 400.

Com a finalidade de verificar as propriedades assintóticas dos estimadores bayesianos, os procedimentos adotados foram replicados 1.000 vezes para cada tamanho amostral e em cada repetição foram armazenados os valores das estimativas e dos desvios-padrão.

Para os valores de estimativa dos parâmetros, foi realizado cálculo de sua média e construído intervalo de confiança para esta. Os valores do desvio-padrão foram utilizados para a mensuração do desvio-padrão médio das estimativas com a intenção de verificar sua volatilidade, conforme demonstrado na Tabela 4.9

Tabela 4.9: Estimativas pontuais e intervalares - parâmetro ϕ constante

n	$\beta's$	Valores verdadeiros	Estimativa	σ médio	LI	LS
30	β_0	0,50	0,5277	0,2236	0,5038	0,5516
	β_1	-0,02	-0,0183	0,3365	-0,0544	0,0177
	β_2	0,30	0,3195	0,0976	0,3091	0,3298
	β_3	-0,15	-0,1533	0,0485	-0,1585	-0,1481
	β_4	0,05	0,0363	0,4028	-0,0073	0,0799
	β_5	-0,07	-0,0695	0,2234	-0,0935	-0,0456
	γ	8,20	2,078	0,3241	2,0564	2,0996
50	β_0	0,50	0,5021	0,1821	0,4829	0,5212
	β_1	-0,02	-0,0309	0,2235	-0,0542	-0,0077
	β_2	0,30	0,2991	0,0729	0,2915	0,3067
	β_3	-0,15	-0,1493	0,0325	-0,1528	-0,1459
	β_4	0,05	0,0691	0,2712	0,041	0,0972
	β_5	-0,07	-0,0651	0,1642	-0,0831	-0,0472
	γ	8,20	2,0762	0,252	2,0604	2,092
100	β_0	0,50	0,5085	0,1367	0,4943	0,5227
	β_1	-0,02	-0,0265	0,1425	-0,0419	-0,011
	β_2	0,30	0,3006	0,0411	0,2962	0,3049
	β_3	-0,15	-0,1497	0,0226	-0,1521	-0,1473
	β_4	0,05	0,0502	0,1326	0,0363	0,0641
	β_5	-0,07	-0,0715	0,1166	-0,0839	-0,0591
	γ	8,20	2,0725	0,1825	2,0615	2,0836

Fonte: Elaborada pelo autor

Observa-se que mesmo em amostras pequenas e com um número relativamente pequeno de iterações (400), os parâmetros da média são estimativas consistentes. Os dados demonstram que à medida que o tamanho amostral aumenta, diminui o valor do desvio-padrão, ao mesmo tempo em que melhoram as estimativas dos parâmetros da média. O parâmetro de precisão apresenta baixo poder de precisão. Verifica-se que sua média está deslocada à esquerda do verdadeiro valor do parâmetro, mesmo em tamanhos amostrais grandes ($n = 100$).

O coeficiente de assimetria dos parâmetros é relativamente baixo em amostras pequenas, e à medida que se aumenta o tamanho amostral, o valor do coeficiente de assimetria tende a zero para a maioria dos parâmetros conforme observado na Tabela 4.10

Tabela 4.10: Coeficiente de assimetria ϕ constante

n	Parâmetros						
	β_0	β_1	β_2	β_3	β_4	β_5	γ
30	0,00	-0,01	-0,01	-0,11	0,02	-0,02	0,42
50	0,06	-0,09	0,12	0,05	-0,16	0,05	0,30
100	0,04	-0,01	0,16	0,07	0,08	-0,06	0,20

Fonte: Elaborada pelo autor

Verifica-se que os estimadores apresentam boas propriedades, são consistentes, sua variabilidade é relativamente pequena e seus intervalos de confiança apresentam o verdadeiro valor das estimativas. A exceção é dada pelo estimador do ϕ, que não é consistente e seu intervalo de confiança não apresenta o verdadeiro valor do parâmetro.

O processo descrito acima foi repetido para os modelos nos quais o parâmetro de precisão foi considerado variado. No entanto, as simulações com a precisão variando não diferem de forma significativa das com o parâmetro constante. A Tabela 4.11 demonstra os resultados:

Tabela 4.11: Estimativas pontuais e intervalares - parâmetro ϕ variável

n	$\beta's$	Valores verdadeiros	Estimativa	σ médio	LI	LS
30	β_0	0,50	0,5062	0,2561	0,4775	0,5349
	β_1	-0,02	0,0178	0,2947	-0,0101	0,0458
	β_2	0,30	0,3247	0,0915	0,3157	0,3337
	β_3	-0,15	-0,16	0,0484	-0,165	-0,155
	β_4	0,05	0,0495	0,3409	0,0199	0,0791
	β_5	-0,07	-0,0543	0,2202	-0,076	-0,0326
	γ_0	0,40	2,0033	1,256	1,9164	2,0902
	γ_1	3,80	2,6233	1,0022	2,5448	2,7017
50	β_0	0,50	0,5149	0,2124	0,4915	0,5383
	β_1	-0,02	-0,0158	0,2044	-0,0374	0,0058
	β_2	0,30	0,3141	0,0653	0,3077	0,3205
	β_3	-0,15	-0,1531	0,0316	-0,1562	-0,1499
	β_4	0,05	0,0436	0,2161	0,0247	0,0626
	β_5	-0,07	-0,0682	0,1531	-0,0831	-0,0533
	γ_0	0,40	1,9232	1,0379	1,8481	1,9984
	γ_1	3,80	2,7122	0,791	2,6478	2,7766
100	β_0	0,50	0,5086	0,1464	0,4932	0,524
	β_1	-0,02	-0,0178	0,1408	-0,0323	-0,0033
	β_2	0,30	0,3023	0,0368	0,2989	0,3057
	β_3	-0,15	-0,1516	0,02	-0,1535	-0,1497
	β_4	0,05	0,0479	0,1157	0,0372	0,0585
	β_5	-0,07	-0,0755	0,1094	-0,086	-0,065
	γ_0	0,40	1,7401	0,8092	1,6839	1,7964
	γ_1	3,80	2,8112	0,6501	2,7653	2,8572

Fonte: Elaborada pelo autor

Como observado anteriormente, os parâmetros γ_0 e γ_1 apresentam viés, isso pode ser devido ao número de iterações ter sido pequeno. Não foi observada assimetria significativa nos parâmetros. Como podemos observar na Tabela 4.12.

Tabela 4.12: Coeficiente de assimetria ϕ variável

n	Parâmetros							
	β_0	β_1	β_2	β_3	β_4	β_5	γ_0	γ_1
30	0,18	0,00	0,21	0,03	0,15	-0,11	0,70	-0,28
50	0,09	-0,05	0,16	0,01	-0,05	-0,06	0,31	-0,08
100	0,00	0,01	0,24	0,04	-0,06	-0,10	0,34	-0,27

Fonte: Elaborada pelo autor

As simulações foram repetidas aumentando o número de iterações. O intuito foi verificar a convergência do algoritmo para os parâmetros de precisão; novamente foi considerado constante em um primeiro momento e depois como sendo variado.

4.8 Critérios de convergência e estacionariedade em cadeia de Markov

O diagnóstico para a convergência de cadeia de Markov foi proposto por Geweke (1992). Esse teste é de comparação de médias, no qual se compara uma média retirada da parte inicial da cadeia de Markov com a média da última parte da cadeia. A estatística de Geweke tem assintoticamente distribuição normal padrão. A estatística de teste é Z escore: a diferença entre as duas médias de amostra dividida por seu erro padrão estimado. O erro padrão é estimado a partir da densidade espectral em zero, e assim leva em conta qualquer autocorrelação. É esperado que o valor Z escore esteja entre $-1,96 <$ Z escore $< 1,96$ para que passe no teste, com isso adota-se $\alpha = 0,05$.

Outro teste de convergência que será aplicado utiliza a estatística Cramer-von-Mises para verificar a hipótese nula de que os valores amostrados são provenientes de uma distribuição estacionária. O teste é aplicado de forma sucessiva, primeiramente para toda a cadeia, em seguida, descartando-se os primeiros 10%, depois os 20%,

até que a hipótese nula é aceita, ou 50% da cadeia tenha sido descartado. Se o último resultado constitui um "fracasso", indica que é necessário um número maior de iterações do algoritmo MCMC.

O teste de meia amplitude será aplicado após o de convergência. Aquele teste calcula o intervalo de confiança da média, utilizando a fração da cadeia de Markov que passou no teste de estacionariedade e com isso calcula-se a razão entre a metade da amplitude desse intervalo de confiança e a estimativa da média. Caso esse valor seja menor que um valor pré-determinado nos ensaios, adota-se 0,1. Caso não seja aprovado no teste, considera-se que o tamanho da amostra não foi grande o suficiente para estimar a média com precisão.

Para as simulações, foram adotados os seguintes critérios: número de iterações de 100.000, *burn-in* de 20% e valores saltados de 100 em 100 ($jump = 100$). Foram realizados testes de critério de convergência e estacionariedade. As Tabelas 4.13, 4.14, 4.15 e 4.16 apresentam os resultados obtidos nas simulações.

Tabela 4.13: Estimativas pontuais e intervalares - parâmetro ϕ constante

n	$\beta's$	Valores verdadeiros	Estimativas	$\hat{\sigma}$	LI	LS
30	β_0	0,24	0,5277	0,2031	-0,1627	0,636
	β_1	0,94	-0,0183	0,3185	0,3075	1,544
	β_2	0,45	0,3195	0,0945	0,2368	0,641
	β_3	-0,30	-0,1533	0,0433	-0,3814	-0,208
	β_4	-0,62	0,0363	0,3912	-1,3481	0,167
	β_5	0,64	-0,0695	0,1999	0,2678	1,032
	γ	2,32	2,078	0,3879	1,4747	2,986
50	β_0	0,33	0,5021	0,2042	-0,0652	0,71
	β_1	0,27	-0,0309	0,2534	-0,2476	0,743
	β_2	0,22	0,2991	0,0811	0,0634	0,361
	β_3	-0,17	-0,1493	0,0364	-0,2362	-0,094
	β_4	0,21	0,0691	0,2878	-0,3293	0,808
	β_5	0,11	-0,0651	0,1877	-0,2532	0,488
	γ	1,80	2,0762	0,2769	1,2602	2,322
100	β_0	0,64	0,5085	0,1602	0,311	0,953
	β_1	-0,11	-0,0265	0,1716	-0,4325	0,234
	β_2	0,16	0,3006	0,0471	0,0612	0,245
	β_3	-0,16	-0,1497	0,0248	-0,2075	-0,109
	β_4	-0,20	0,0502	0,1456	-0,4767	0,078
	β_5	-0,08	-0,0715	0,133	-0,3643	0,175
	γ	1,82	2,0725	0,1988	1,4486	2,211

Fonte: Elaborada pelo autor

Tabela 4.14: Diagnóstico ϕ constante

n	$\beta's$	Geweke		Heidelberge e Welch			
		Valor	Resultado	Estacionaridade	P-valor	Half-with	Resultado
30	β_0	-0,9079	Aprovado	Aprovado	0,3218	0,0617	Aprovado
	β_1	0,7725	Aprovado	Aprovado	0,1735	0,0279	Aprovado
	β_2	0,1826	Aprovado	Aprovado	0,4162	0,0145	Aprovado
	β_3	-0,6922	Aprovado	Aprovado	0,0648	0,0126	Aprovado
	β_4	0,5443	Aprovado	Aprovado	0,7162	0,0440	Aprovado
	β_5	-0,5551	Aprovado	Aprovado	0,5289	0,0205	Aprovado
	γ	-1,169	Aprovado	Aprovado	0,1410	0,0134	Aprovado
50	β_0	-1,13588	Aprovado	Aprovado	0,152	0,0435	Aprovado
	β_1	1,04294	Aprovado	Aprovado	0,622	0,0658	Aprovado
	β_2	0,04301	Aprovado	Aprovado	0,193	0,0255	Aprovado
	β_3	0,45585	Aprovado	Aprovado	0,343	0,0149	Aprovado
	β_4	0,59626	Aprovado	Aprovado	0,416	0,1039	Reprovado
	β_5	0,29609	Aprovado	Aprovado	0,696	0,1183	Reprovado
	γ	3,431	Reprovado	Reprovado	0,007	-	Reprovado
100	β_0	-1,4524	Aprovado	Aprovado	0,1997	0,0173	Aprovado
	β_1	-0,7169	Aprovado	Aprovado	0,7424	0,1127	Reprovado
	β_2	1,6009	Aprovado	Aprovado	0,2646	0,0184	Aprovado
	β_3	1,4786	Aprovado	Aprovado	0,1048	0,0107	Aprovado
	β_4	1,0232	Aprovado	Aprovado	0,0568	0,0368	Aprovado
	β_5	0,3732	Aprovado	Aprovado	0,9213	0,1138	Reprovado
	γ	0,3691	Aprovado	Aprovado	0,2060	0,0086	Aprovado

Fonte: Elaborada pelo autor

Tabela 4.15: Estimativas pontuais e intervalares - parâmetro ϕ variável

n	$\beta's$	Valores verdadeiros	estimativa	σ médio	LI	LS
30	β_0	0,50	0,4859	0,2545	-0,0057	0,967
	β_1	-0,02	0,1266	0,2553	-0,3609	0,625
	β_2	0,30	0,5028	0,0826	0,3342	0,668
	β_3	-0,15	-0,1844	0,0498	-0,2721	-0,081
	β_4	0,05	-0,764	0,3027	-1,3392	-0,178
	β_5	-0,07	0,6813	0,2261	0,2845	1,149
	γ_0	0,40	2,1904	1,478	-0,7044	5,04
	γ_1	3,80	3,0052	1,3842	0,4037	5,724
50	β_0	0,50	0,0211	0,2342	-0,4503	0,46
	β_1	-0,02	0,1569	0,1915	-0,2094	0,523
	β_2	0,30	0,3595	0,0562	0,2431	0,465
	β_3	-0,15	-0,0925	0,0317	-0,1504	-0,029
	β_4	0,05	0,0134	0,1848	-0,322	0,408
	β_5	-0,07	0,2759	0,1549	-0,0116	0,609
	γ_0	0,40	0,4957	1,1943	-1,9999	2,684
	γ_1	3,80	4,0101	1,0417	2,1123	6,004
100	β_0	0,50	0,203	0,1622	-0,121	0,524
	β_1	-0,02	0,1327	0,1483	-0,1479	0,446
	β_2	0,30	0,3219	0,0371	0,2538	0,391
	β_3	-0,15	-0,1128	0,0201	-0,1495	-0,072
	β_4	0,05	0,0942	0,1101	-0,1241	0,304
	β_5	-0,07	0,1733	0,1213	-0,0741	0,422
	γ_0	0,40	0,9094	0,7707	-0,6141	2,347
	γ_1	3,80	3,2377	0,6136	2,0067	4,467

Fonte: Elaborada pelo autor

Tabela 4.16: Diagnóstico ϕ variável

n	$\beta's$	Geweke		Heidelberge e Welch			
		Valor	Resultado	Estacionaridade	P-valor	Half-with	Resultado
30	β_0	2,8614	Reprovado	Aprovado	0,4010	0,0457	Aprovado
	β_1	0,104	Aprovado	Aprovado	0,4190	0,1548	Reprovado
	β_2	-0,4502	Aprovado	Aprovado	0,3730	0,0114	Aprovado
	β_3	-3,0192	Reprovado	Aprovado	0,2990	-0,0240	Aprovado
	β_4	0,2441	Aprovado	Aprovado	0,5550	-0,0275	Aprovado
	β_5	-2,0509	Reprovado	Aprovado	0,5500	0,0287	Aprovado
	γ_0	-2,185	Reprovado	Aprovado	0,7940	0,0571	Aprovado
	γ_1	1,744	Aprovado	Aprovado	0,8580	0,0591	Aprovado
50	β_0	0,9045	Aprovado	Aprovado	0,899	0,8711	Reprovado
	β_1	-0,3665	Aprovado	Aprovado	0,942	0,0734	Aprovado
	β_2	0,9759	Aprovado	Aprovado	0,362	0,0120	Aprovado
	β_3	-1,9869	Reprovado	Aprovado	0,562	0,0338	Aprovado
	β_4	-0,2531	Aprovado	Aprovado	0,127	0,9560	Reprovado
	β_5	0,45	Aprovado	Aprovado	0,683	0,0454	Aprovado
	γ_0	-0,8076	Aprovado	Aprovado	0,6400	0,2157	Reprovado
	γ_1	0,786	Aprovado	Aprovado	0,6700	0,0342	Aprovado
100	β_0	0,2549	Aprovado	Aprovado	0,0725	0,0586	Aprovado
	β_1	-0,3552	Aprovado	Aprovado	0,7112	0,0819	Aprovado
	β_2	-0,6141	Aprovado	Aprovado	0,7604	0,0080	Aprovado
	β_3	0,0157	Aprovado	Aprovado	0,7047	0,0123	Aprovado
	β_4	1,0248	Aprovado	Aprovado	0,0821	0,0810	Aprovado
	β_5	0,0848	Aprovado	Aprovado	0,0869	0,0547	Aprovado
	γ_0	-0,4629	Aprovado	Aprovado	0,9570	0,0618	Aprovado
	γ_1	0,2931	Aprovado	Aprovado	0,7870	0,0183	Aprovado

Fonte: Elaborada pelo autor

Observa-se que em amostras muito pequenas ($n = 30$) e utilizando saltos de 100 em 100 ($jump = 100$), as estimativas pontuais apresentam viés, bem como as intervalares. Já com tamanho amostral maior ($n = 100$), obtiveram-se boas estimativas intervalares, independente de se utilizar o parâmetro de precisão constante ou variando.

Já para o critério de convergência, o modelo em pequenas amostras ($n = 30$) não foi aprovado, o que pode ter ocorrido devido ao espaçamento utilizado ($jump = 100$). Esse espaçamento determina a quantidade amostrada do estimador; quanto menor, maior o tamanho da amostra do estimador. O espaçamento é necessário, uma vez que em processos de cadeia de Markov amostras próximas são autocorrelacionadas. Verifica-se, também, que todas as amostras passaram no teste de estacionariedade.

Para avaliar o critério de convergência em pequenas amostras, foi alterado o número de saltos e mantido constante tanto o *burn-in*, em 20%, como o número de iterações, em 100.000. Como esperado, o modelo passou pelo critério de convergência bem como pelo de estacionariedade, conforme Tabela .

Tabela 4.17: Diagnóstico ϕ constante

n	$\beta's$	Geweke		Heidelberge e Welch			
		Valor	Resultado	Estacionaridade	P-valor	Half-with	Resultado
30	β_0	-1,6024	Aprovado	Aprovado	0,518	0,0451	Aprovado
	β_1	0,4667	Aprovado	Aprovado	0,795	-0,0167	Aprovado
	β_2	0,4374	Aprovado	Aprovado	0,914	0,0061	Aprovado
	β_3	1,2234	Aprovado	Aprovado	0,505	0,0241	Aprovado
	β_4	0,7588	Aprovado	Aprovado	0,901	0,0154	Aprovado
	β_5	0,7791	Aprovado	Aprovado	0,776	-0,025	Aprovado
	γ	-1,169	Aprovado	Aprovado	0,301	0,0115	Aprovado

Fonte: Elaborada pelo autor

Considerações finais

Nas simulações não foram encontradas diferenças significativas entre os modelos, tanto no enfoque clássico quanto no bayesiano. Também não se encontrou diferenças nas estimativas quando alterada sua função de ligação. Percebeu-se que em ambos os casos, em amostras pequenas, as estimativas pontuais e intervalares apresentam pequeno viés, principalmente para a estimativa de ϕ.

No enfoque clássico, percebeu-se que os estimadores de máxima verossimilhança apresentam boas propriedades, e sua distribuição é aproximadamente normal. As estimativas pontuais e o poder do teste foram afetados em amostras pequenas. No

enfoque bayesiano, percebeu-se que os estimadores pontuais em amostras pequenas são viesados, no entanto, o verdadeiro valor do parâmetro é encontrado em todas as estimativas intervalares.

Comparativamente a abordagem clássica obteve melhores estimativas que a abordagem bayesiana utilizada, tal fato pode ser decorrente das prioris utilizadas para os parâmetros na abordagem bayesiana.

No próximo capítulo será apresentado caso de aplicação na área econômica, com dados reais para os modelos beta. Nele serão descritas as variáveis utilizadas e a especificação dos modelos beta construídos tanto com enfoque clássico, como bayesiano.

Capítulo 5

Construção do modelo

O presente trabalho tem como uma de suas abordagens a aplicação dos modelos com suporte na distribuição beta como alternativa aos atuais modelos de teste de estresse para inadimplência.

Para desenvolver o modelo de inadimplência, buscou-se informações de clientes com empréstimos tomados, com diversos comportamentos de pagamentos ao longo do tempo. Com isso, tomou-se a razão do valor dos atrasos maiores que noventa dias sobre o valor da carteira total de empréstimos como variável a ser determinada por influência de fatores macroeconômicos. Um domínio de índices macroeconômicos tradicionais foi abordado, como por exemplo: Inflação, PIB, câmbio e etc. Ainda, variações desses índices, índices compostos, média móveis dos próprios índices e índices não tradicionais.

No presente capítulo apresentamos um modelo de teste de estresse para inadimplência de bancos com suporte na distribuição beta. O processo metodológico utilizado é dado pelos seguintes passos: construção da base de dados com base na revisão bibliográfica, análise das variáveis, construção de modelos beta para estimação da inadimplência baseado nas variáveis explicativas selecionadas, análise dos modelos e comparação dos modelos.

5.1 Base de dados

Os dados foram extraídos do banco de dados do Banco Central e IPEA. Estes contêm diversas fontes de informação, dentre elas estão: IBGE, Serasa, BNDES, entre outros.

A variável resposta do modelo, inadimplência, tem como fonte o Banco Central e é definida como a relação entre o saldo em atraso superior a noventa dias sobre o saldo total na data base. O saldo é composto por todas as operações de empréstimo, financiamentos, adiantamento e arrendamento mercantil concedidas pelas instituições integrantes do sistema financeiro nacional.

Todos os dados macroeconômicos foram considerados em base mensal. Grande parte das séries macroeconômicas utilizadas, afeta a economia brasileira com relativa defasagem, ou seja, o efeito se manifesta em prazo mais longo na economia.

5.2 Revisão bibliográfica para escolha das variáveis

Para a construção da base de dados foi realizada uma revisão bibliográfica, na qual buscou-se verificar quais as variáveis são utilizadas, de uma forma geral, nos testes de estresse em que a variável resposta está restrita ao intervalo (0,1).

Vlieghe (2001) estudou a relação entre as variáveis macroeconômicas e falência de empresas no Reino Unido com objetivo de realizar teste de estresse. As variáveis significativas utilizadas foram percentuais do PIB, taxa de juros real e nominal, desvios do PIB, taxa de nascimento de empresas e preços no mercado imobiliário.

Variáveis de produção industrial, inflação, índice de confiança dos empresários, taxa de juros de curto prazo (real e nominal), índices do mercado de ações e exportações foram utilizadas por Kalirai e Scheicher (2002) em seu estudo sobre provisões para perdas de crédito nos bancos da Austria.

Hoggarth et al. (2005) em seus modelos identificaram que o PIB, a inflação de preços no varejo e as taxas de juros de curto prazo afetam as perdas com crédito no setor bancário do Reino Unido. Bunn et al. (2005) construíram modelos econométricos para estimar a inadimplência em crédito no Reino Unido utilizando as seguintes variáveis: PIB real e nominal, lucro das empresas não financeiras, valo-

res das propriedades comerciais, taxa de juros real, renda, taxa de desemprego e habitação.

Sorge e Virolainen (2006) utilizaram modelos econométricos para teste de estresse utilizando PIB (dessazonalizado) e taxa de juros de curto prazo a variável resposta foi a taxa de descumprimento.

A utilização da taxa de descumprimento como variável resposta e a relação com as variáveis macroeconômicas PIB canadense, taxa de juros real, preço de matérias primas, PIB norte-americano e taxa de juros norte-americana serviu para investigar as perdas em carteira de crédito no setor bancário do Canadá (Misina et al. 2006).

Modelos de teste de estresse no Brasil, utilizando informações do Banco Central do Brasil, sistema de informação do crédito e do valor econômico pode ser visto em Santos (2008). As variáveis macroeconômicas incluídas no modelo final foram: hiato do produto, taxas DI Pré, taxa swap DI Pré, erro de previsão da pesquisa FOCUS para o câmbio em 12 meses e Índice Nacional de Preços ao Consumidor Amplo (IPCA).

Vazques et al. (2012) indicaram, por meio de análise de dados em painel, a relação entre a proporção de contratos de crédito em atraso (maior que noventa dias) e o crescimento do PIB. Em Lu e Yang (2012) observaram, no teste de estresse para o banco de agricultura da China, que as variáveis crescimento do PIB, inflação, preços do mercado imobiliário e crescimento monetário do pais foram significativas nos modelos de vetores autorregressivos (VAR), nos quais a variável resposta foi o percentual de contratos em atraso maior que noventa dias.

Schechtman e Gaglianone (2012) propuseram técnicas de teste de estresse para risco de crédito utilizando vetores autorregressivos com foco nas caudas da distribuição da perda em cenários macroeconômicos adversos. As variáveis utilizadas foram crescimento do PIB, taxa de desemprego, IPCA, taxa de juros (SELIC) e o volume de crédito concedido.

Fungáčová e Jakubík (2013) desenvolveram um modelo de teste de estresse para o sistema financeiro da Rússia, que prevê a inadimplência por meio do crescimento do PIB, inflação, taxa de juros de curto prazo e taxas de câmbio.

Na lista abaixo são identificadas as variáveis finais composta no banco de dados. Nem todas as variáveis foram utilizadas nos modelos finais.

As variáveis que compõem o bando de dados são as que se seguem: Inadimplência,

PIB, IPCA, Índice de custo de vida (ICV), SELIC, Desemprego, Câmbio, Renda, Indicador de Produção e Endividamento das Famílias.

5.3 Análise das variáveis

Nesta secção serão analisados aspectos da variável dependente inadimplência e especificaremos as transformações realizadas nas variáveis independentes, como defasagem e média móvel.

Defasagem

Para cada variável independente foi criado um novo campo com seu preenchimento defasado de 1 a 6 meses, pois a inadimplência pode reagir a mudanças no sistema econômico ou de crédito apenas após um período de tempo (KOOPMAN; LUCAS, 2005).

$$Lag\, k\_X_{1,t} = X_{1,t-k} \tag{5.1}$$

em que $X_{1,t-k}$ é a variável explicativa X_1 no tempo $t-k$ e k é a defazagem.

Médias móveis

A transformação de médias móveis tem como objetivo a suavização dos movimentos das séries temporais, removendo efeitos sazonais, cíclicos, irregulares e aleatórios e reduzindo a volatilidade de uma variável (STEVENSON, 1991). O Cálculo da variável transformada é dado por:

$$MM\, k\_X_{1,t} = \frac{X_{1,t} + \ldots + X_{1,t-k}}{k+1} \tag{5.2}$$

em que $X_{1,t}$ é a variável explicativa X_1 no tempo t .

Seleção das variáveis

Como critério para escolha das variáveis finais do modelo foram testadas diversas combinações entre as variáveis explicativas. Aquelas que apresentaram os melhores resultados foram selecionadas para compor o modelo.

Estudo de correlação

Foi realizado um estudo de correlação entre as variáveis finais do modelo, o intuito foi verificar possível multicolineariedade entre as variáveis explicativas. A tabela 5.1 apresenta as correlações observadas.

Tabela 5.1: Tabela de correlação entre as variáveis do modelo.

Correlação	VAR 1	VAR 2	VAR 3	VAR 4	VAR 5	VAR 6
VAR 1	1,00	-0,22	-0,25	0,58	-0,72	-0,65
VAR 2	-0,22	1,00	0,07	-0,12	0,38	0,20
VAR 3	-0,25	0,07	1,00	0,14	-0,17	-0,27
VAR 4	0,58	-0,12	0,14	1,00	-0,57	-0,93
VAR 5	-0,72	0,38	-0,17	-0,57	1,00	0,96
VAR 6	-0,65	0,20	-0,27	-0,93	0,96	1,00

em que VAR 1 representa inadimplência, VAR 2 índice de custo de vida médio nos últimos 3 meses,VAR 3 SELIC média dos últimos 6 meses, VAR 4 desemprego defasado em 5 meses, VAR 5 taxa de câmbio média dos últimos 6 meses e VAR 6 o endividamento das famílias defasadas em 3 meses.

As correlações observadas entre as variáveis explicativas são baixas, uma vez que a variável endividamento das famílias faz parte apenas do modelo de regressão beta com precisão variável, ou seja, foi utilizada somente para se estimar o parâmetro de dispersão do modelo de regressão beta.

5.3.1 Inadimplência

A variável dependente inadimplência do sistema financeira brasileiro é observada no período de março 2011 até marços de 2015 e seus valores variam temporalmente entre 2,73% e 3,72%. Os Gráficos em 5.1 apresentam a evolução da inadimplência ao longo dos últimos anos e a sua dispersão.

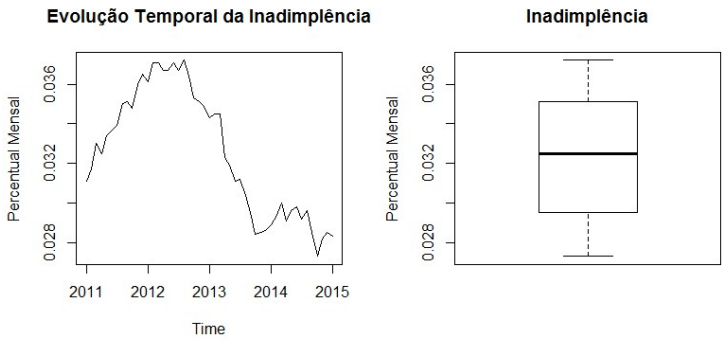


Figura 5.1: Variável dependente inadimplência

Tabela 5.2: Estatísticas descritivas da variável inadimplência.

Minímo	Quartil 1	Mediana	Média	Quartil 3	Máximo	Variância
2,73%	2,95%	3,25%	3,25%	3,51%	3,72%	0,000010

Como observado, os valores da inadimplência ao longo dos últimos 4 anos variam dentro de uma pequena faixa (2,73% e 3,72%), o que demonstra certa estabilidade. Para o desenvolvimento das análises a seguir, assumiu-se que a inadimplência é independente e identicamente distribuída (iid).

5.4 Desenvolvimento do modelo

Nesta seção será apresentada uma aplicação dos modelos de regressão beta a um conjunto de dados reais, sob os enfoques clássico e bayesiano.

De início, será ajustado um modelo de Mínimos Quadrados Ordinários para mostrar sua inadequação em tratar os dados. Em seguida, proceder-se-á a estimação de regressão beta em que usam as funções de ligação logito, e considera-se o parâmetro de precisão fixo. Além disso, verifica-se a existência de pontos de influência usando a distância de Cook (COOK, 1977) e o *laverage* generalizado. Adicionalmente, é analisado o comportamento do parâmetro de precisão.

A investigação prossegue, buscando averiguar se o parâmetro de precisão é ou não fixo, usando, para tal, os testes da razão de verossimilhança (TRV) e de Wald, e, em seguida, o critério BIC e AIC para seleção de modelos. O próximo passo é

tentar encontrar a melhor função de ligação, ajustando regressões com as formas probito, complemento log-log e log-log. Utiliza-se o pseudo-R^2 e os critérios BIC e AIC de seleção de modelos para extrair as conclusões. Por fim, verifica-se se o modelo está bem ajustado, usando como ferramenta de diagnóstico, entre elas os gráficos de envelope simulado.

Todos os cálculos computacionais para os ajustes dos modelos sob ambos os enfoques são realizados utilizando a linguagem de programação R.

A estimação dos parâmetros dos modelos sob o enfoque clássico foi realizada usando o algoritmo quasi-Newton (BFGS) e a escolha dos valores inicias para os parâmetros desconhecidos segue a sugestão de Ferrari e Cribari-Neto (2004).

5.4.1 Estimação do modelo de Inadimplência

O interesse do estudo é propor o uso de modelos com suporte na distribuição beta para fins de testes de estresse que utilizam dados macroeconômicos, como método alternativo aos modelos do tipo Wilson adotados em testes de estresse em instituições financeiras.

Nos modelos de teste de estresse para inadimplência de bancos, a variável de interesse está restrita ao intervalo (0,1). Nesse sentido, um modelo com suporte em distribuição pertencente a esse intervalo, como a beta, é candidato natural.

O objetivo principal do modelo é estimar o valor da inadimplência em função de variáveis macroeconômicas. O intuito do teste de estresse, de maneira geral, é verificar se um determinado banco terá capital suficiente frente as perdas que tendem a ocorrer em um cenário desfavorável extremo da economia (estresse).

5.4.2 Ajuste do modelo de Wilson

A regressão foi ajustada pelo método de Wilson para as observações e chegou-se aos seguintes resultados, conforme Tabela 5.3.

Tabela 5.3: Estimativas dos parâmetros do modelo de Wilson.

Parâmetro	Variáveis	Valores Estimados	Valor p	Intervalo de Confiança	
				L.I	L.S
β_0	-	-1,47	0,05%	-2,26	-0,68
β_1	Custo de Vida - Média dos ultimos 3 meses	8,28	1,14%	1,96	14,59
β_2	Selic - Média dos ultimos 6 meses	-32,82	0,00%	-45,42	-20,23
β_3	Desemprego - Defasado em 5 meses	-12,41	0,67%	-21,21	-3,62
β_4	Câmbio - Média dos ultimos 6 meses	-0,51	0,00%	-0,66	-0,35
R^2		0,7339			
R^2 *Ajustado*		0,7097			

Para o obtenção desse modelo a variável resposta Y_k foi transformada, utilizando-se a função logito. Os resultados mostram que todos as variáveis utilizadas no modelo são significativas a 10% e que o poder de explicação do modelo é de 70,97% ($R^2\,ajustado$). Para verificar a heterocedasticidade dos dados, foi realizado o teste de Breusch-Pagan. A hipótese nula do teste é que os dados são homocedásticos contra a hipótese alternativa de que são heterocedásticos. Foi obtido valor p de 11,78 %, apesar de não rejeitarmos a hipótese nula os Gráficos 5.2, 5.3 e 5.4 demostram que os dados são assimétricos e apresentam um certo desvio quanto a normalidade, constituindo entrave para a utilização de modelos com suporte na distribuição normal.

Figura 5.2: Rediuos observados

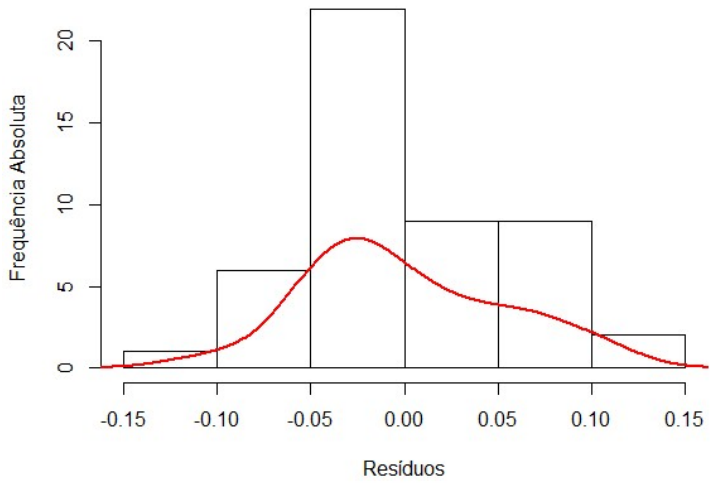


Figura 5.3: Normal Q-Q plot

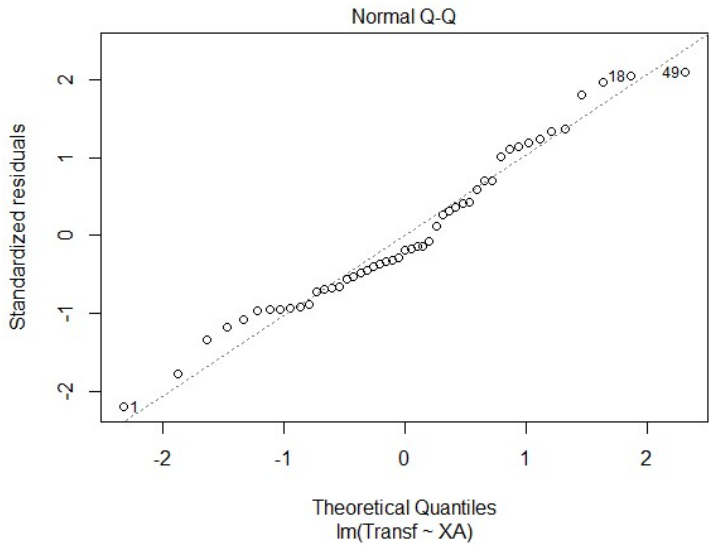
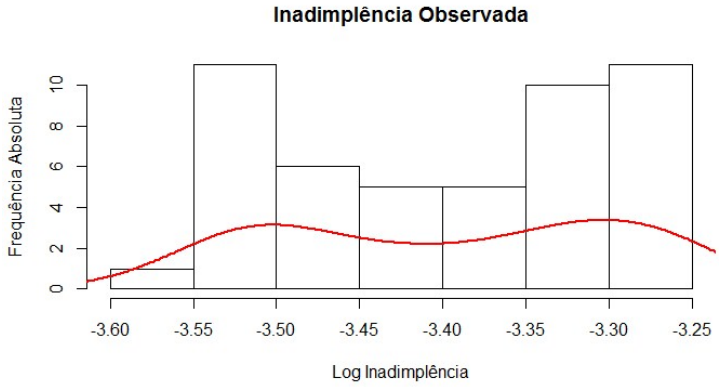


Figura 5.4: Log da inadimplência observada



5.4.3 Ajuste do modelo de regressão beta

Uma vez que a variável dependente é contínua e restrita ao intervalo (0,1), procede-se com uma modelagem por meio de regressão beta. Como visto anteriormente, Ferrari e Cribari-Neto (2004) sugeriram uma reparametrização envolvendo um parâmetro de média e um de dispersão.

Para a modelagem de μ_t em termos do preditor linear, escolheu-se, inicialmente, dentre algumas funções de ligação, as seguintes: Logito, Probito, Complemento log-log e log-log.

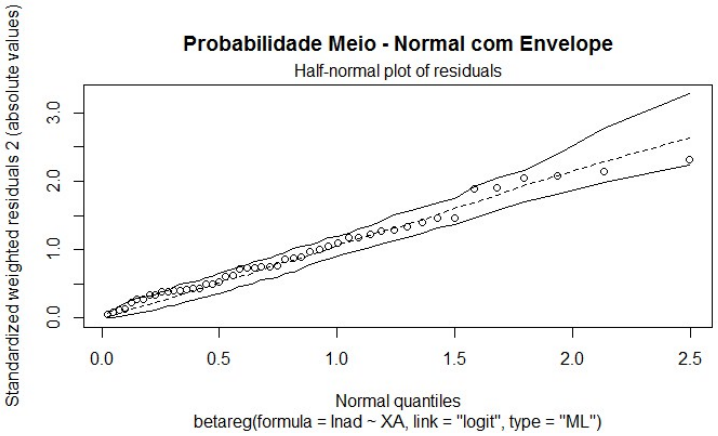
Primeiramente, o objetivo é verificar se o parâmetro de dispersão é fixo ou variável. Para tanto, ajusta-se uma regressão beta com função de ligação logito e parâmetro de precisão fixo. Os resultados (Tabela 5.4) mostraram que as variáveis macroeconômicas, bem como o intercepto, foram estatisticamente significantes a 1% de significância, e o modelo apresentou um poder de explicação de 73,35%. Note-se, ainda que a estimação do modelo é realizada via máxima verossimilhança.

Tabela 5.4: Estimativas dos parâmetros - Modelo Beta - Função Logito - Precisão constante

Variáveis	Estimativas	Erro Padrão	Valor t	Valor p
Intercepto	-1,43	0,38	-3,77	0,02%
Custo de Vida - Média dos ultimos 3 meses	7,81	3,03	2,58	0,98%
Selic - Média dos ultimos 6 meses	-32,74	5,90	-5,55	0,00%
Desemprego - Defasado em 5 meses	-12,89	4,25	-3,04	0,24%
Câmbio - Média dos ultimos 6 meses	-0,51	0,08	-6,73	0,00%
ϕ	11498,00	2323,00	4,95	0,00%
$Log - Like$	244,4			
$P - seudoR^2$	0,7335			
Número de Interações (BFGS)	92			

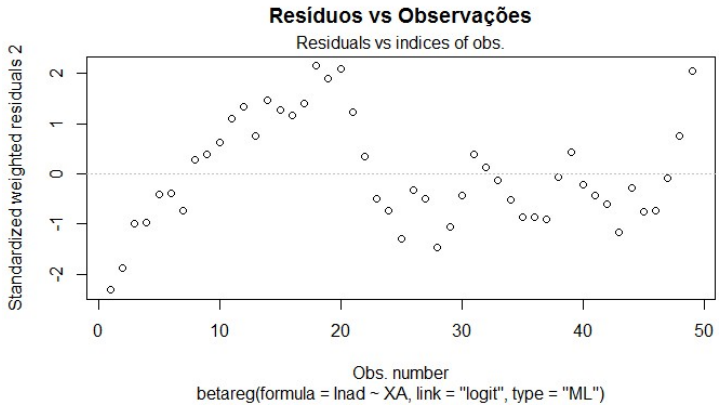
Pode-se analisar a qualidade do ajuste da regressão usando as ferramentas de diagnóstico apresentadas anteriormente. O Gráfico 5.5, de probabilidade meio normal com envelope simulado, mostra que exitem pontos que se encontram no limiar da banda de confiança, o que denota que o modelo ainda não está bem ajustado.

Figura 5.5: Probabilidade meio-norma com envelope



Além disso, a partir do Gráfico 5.6, que plota os resíduos ao longo da amostra, percebe-se que os pontos estão distribuídos de forma aleatória.

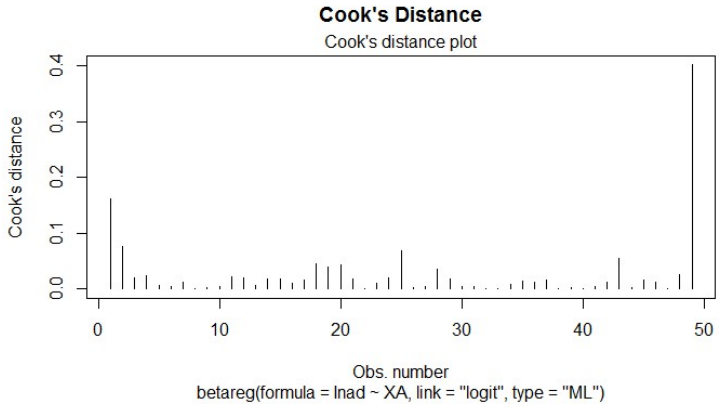
Figura 5.6: Resíduos versus observações



Ademais, visando a identificar pontos de alavanca, utiliza-se como critério de análise a distância de Cook e *leverage* generalizado. Verifica-se que a observação de número 49 é influente.

$$LG\,(Obs.49) = 0,2899 > \frac{2k}{n} = 0,2449. \tag{5.3}$$

Figura 5.7: Distância de Cook



A partir do Gráfico 5.7, constata-se a presença de pontos de alavanca. Dado esse cenário, optou-se, inicialmente, por retirar essa observação da base de dados e verificar o comportamento do parâmetro de precisão. Os resultados mostraram que o $\hat{\phi}$ da regressão beta sem o ponto de alavanca é bem superior ao modelo em que não retirou

essa observação (12.346). A princípio, uma decisão plausível seria a exclusão dessa observação para prosseguir com a análise, contudo, adotando cautela, decidiu-se não a retirar.

Kosmidis e Firth (2010) demonstram que o estimador de máxima verossimilhança pode ser severamente enviesado no contexto de regressão beta, no sentido de subestimar o erro padrão dos estimadores. Os autores demostraram como realizar a correção do viés (BC), e a redução do viés (BR) dos estimadores de máxima verossimilhança em modelos paramétricos por meio de um algoritmo unifica o BFGS e o escore de Fisher. Eles demonstraram que BC/BR para modelos de regressão beta podem ser desejável, uma vez que o estimador de máxima verossimilhança de ϕ pode apresentar um viés de alta substancial.

Com isso foi testado o modelo utilizando a correção do viés (BC) e a redução do viés (BR) para os estimadores de máxima verossimilhança, a tabela 5.5 demonstra que o ajuste nos estimadores melhora as estimativas dos parâmetros, principalmente do parâmetro de precisão.

Tabela 5.5: Ajuste do parâmetro de precisão (ϕ)

Variáveis	Método de Estimação		
	ML	BC	BR
Intercepto	-1,43	-1,43	-1,43
Custo de Vida - Média dos ultimos 3 meses	7,81	7,81	7,81
Selic - Média dos ultimos 6 meses	-32,74	-32,73	-32,73
Desemprego - Defasado em 5 meses	-12,89	-12,89	-12,89
Câmbio - Média dos ultimos 6 meses	-0,51	-0,51	-0,51
ϕ	11497,46	9854,41	9854,35

Prosseguindo-se a análise, deve-se obter os resultados do ajuste da regressão com parâmetro de dispersão variável e, em seguida, aplicar o teste da razão de verossimilhança e Wald a fim de escolher o modelo mais adequado.

O modelo proposto por Ferrari e Cribari-Neto (2004) considera que o parâmetro de dispersão é fixo para as observações. Em muitas situações, essa hipótese pode não ser adequada, de modo que se torna necessário utilizar uma extensão do modelo de

regressão beta introduzida por Simas, Barreto-Souza e Rocha (2010). Nesse caso, adiciona-se à modelagem da média uma estrutura de regressão para o parâmetro de dispersão. Assim, assume-se que:

$$g_1(\mu_i) = \eta_{1i} = X_i^T \beta \tag{5.4}$$

e

$$g_2(\phi_i) = \eta_{2i} = Z_i^T \gamma. \tag{5.5}$$

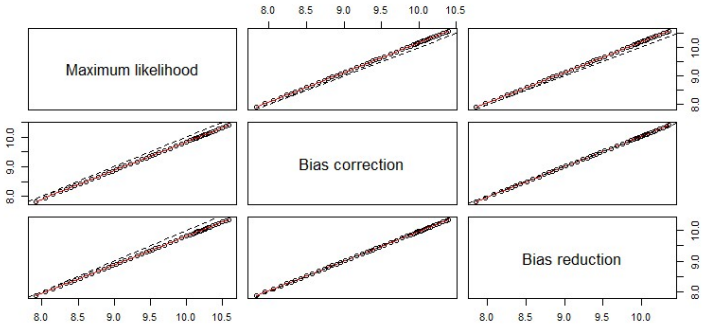
Em que $\beta = (\beta_1, \ldots, \beta_n)$ e $\gamma = (\gamma_1, \ldots, \gamma_n)^T$ são os vetores de parâmetros a serem estimados no modelo. Com isso, foi ajustado um modelo de regressão beta com o parâmetro de dispersão variável, no qual se incorpora a variável endividamento das famílias como regressora no submodelo de dispersão. Como critério para escolher essa variável, testaram-se diversas combinações entre as variáveis explicativas e as demais variáveis constantes na base de dados, verificou-se que a utilização da variável endividamento familiar apresentava melhores resultados. Os resultados podem ser vistos na Tabela 5.6.

Tabela 5.6: Estimativas dos parâmetros com ϕ variando e BR

Variáveis	Estimativas	Erro Padrão	Valor t	Valor p
Intercepto - μ	-1,93	0,29	-6,70	0,00%
Custo de Vida - Média dos ultimos 3 meses	8,87	2,36	3,76	0,02%
Selic - Média dos ultimos 6 meses	-25,41	5,59	-4,55	0,00%
Desemprego - Defasado em 5 meses	-7,02	3,20	-2,19	2,83%
Câmbio - Média dos ultimos 6 meses	-0,46	0,06	-7,78	0,00%
Intercepto - ϕ	-8,22	4,66	-1,76	7,80%
Endividamento das Familias	40,47	10,63	3,81	0,01%
Log-Like	248,5			
$P - seudoR^2$	0,705			
Número de Interações (BFGS)	121			

Adicionalmente, realizou-se a comparação entre os métodos de estimação de máxima verossimilhança com os métodos BC e BR. O gráfico 5.8 demonstra que o parâmetro de precisão fica mais bem estimado utilizando-se o estimador de máxima verossimilhança com redução de viés.

Figura 5.8: Comparando os ajustes de máxima verossimilhança, BC e BR dos parâmetros com ϕ variando



Percebe-se que a variável endividamento familiar é estatisticamente significante no submodelo do parâmetro de dispersão, sugerindo que, de fato, este último é variável. Após a estimação, são realizados testes sobre os parâmetros do modelo com o intuito de confirmar ou não as hipóteses. Os testes mais comumente aplicados são o da razão de verossimilhança, escore e Wald, pois suas estatísticas se baseiam na função de verossimilhança, que apresenta várias propriedades de otimalidade (QUEIROZ; CRIBARI-NETO, 2011).

Contudo, foram utilizados o testes TRV e de Wald para escolher entre os dois modelo (parâmetro ϕ constante e parâmetro ϕ variando) . O teste TRV considera a hipótese nula de que a dispersão é fixa contra a hipótese alternativa que ela é variável, no entanto, como observado nas simulações realizadas o TRV apresenta um poder do teste inferior ao teste de Wald em amostras pequenas, por essa razão também foi realizado o teste de Wald. Fazendo os testes TRV e de Wald, rejeita-se a hipótese nula a 1% de significância, o que indica que se tem um modelo com dispersão variável. Conforme a Tabela 5.7.

Tabela 5.7: Valor P dos testes TRV e de Wald.

Teste	P - valor
TRV	0,42%
Wald	0,00%

Além disso, procederam-se os testes de seleção de modelos BIC e AIC para funda-

mentar com mais precisão a análise. Assim, observou-se que o modelo com dispersão variável apresentou um melhor ajuste, com BIC de -469,19 e um AIC de -482,43. Até o presente momento do estudo, conclui-se que se deve utilizar um modelo com dispersão variável. Conforme a Tabela 5.8.

Tabela 5.8: Testes AIC e BIC para comparar modelos de ϕ constante e ϕ variável.

Teste	Precisão	
	Constante	Variando
AIC	-476,24	-482,44
BIC	-464,89	-469,19

5.4.4 Seleção de diferentes funções de ligação

A seleção de uma função de ligação apropriada possui um grande impacto no ajuste do modelo, especialmente se verificarmos proporções perto de 0 ou de 1 nos dados. Assim, o objetivo nesta etapa do trabalho é estudar qual a melhor função de ligação para os dados. Nesse sentido, ajustou-se mais três regressão beta em que se alterou a função de logito para probito, complemento log-log e log-log. Assim, temos quatro regressões:

1. Ligação logito + Parâmetro de dispersão variável;
2. Ligação probito + Parâmetro de dispersão variável;
3. Ligação complemento log-log + Parâmetro de dispersão variável;
4. Ligação log-log + Parâmetro de dispersão variável.

Percebe-se que, a partir das conclusões das analises anteriores, tem-se a informação de que devemos utilizar um modelo com dispersão variável. Como ferramenta de diagnóstico para avaliar a função de ligação, será utilizado o pseudo-R^2 e, em seguida, os critérios de seleção de modelos AIC e BIC. O primeiro é dado por:

$$R_p^2 = Corr\left(\hat{\eta}, g\left(y\right)\right) = \eta_{2i}. \tag{5.6}$$

Os resultados mostram que o pseudo-R^2 do modelo 3 (Complemento log-log) é relativamente maior do que os demais modelos. Quando usamos o critério de seleção de modelos AIC e BIC constata-se que o modelo 3 apresenta uma melhor performance com um AIC (-482,46) e BIC (-469,22), contra valores maiores dos demais modelos, conforme Tabela 5.9.

Tabela 5.9: Estatísticas pseudo-R^2, AIC e BIC dos modelos

Modelos	Pseudo - R^2	AIC	BIC
Modelo 1	70,77%	-482,44	-469,194
Modelo 2	70,44%	-482,23	-468,984
Modelo 3	70,81%	-482,46	-469,216
Modelo 4	70,16%	-482,05	-468,809

Porém, atente que a diferença entre os valores obtidos é pequena. Uma análise gráfica, pode ser utilizada para averiguar a aderência do modelo em relação a sua função de ligação. Em nenhum dos gráficos utilizados observou-se diferenças significativas (Anexo D).

Ferrari e Cribari-Neto (2004) argumentam que quando a variável de resposta não apresenta muitas observações localizadas no extremo da distribuição, a escolha da função de ligação não é tão relevante. Assim, o ajuste do modelo com várias funções de ligações não apresenta grandes disparidades, no que diz respeito à log-verossimilhança. Realizando teste com todas as funções de ligações disponíveis (Tabela 5.10), é possível verificar que a função complemento log-log apresenta melhores resultados, mas percebe-se que a diferença com outras funções é sutil.

Quando se compara os modelos de regressão beta com os modelos de Wilson, utilizando os critérios de AIC e BIC, verifica-se que os primeiros são melhores, conforme Tabela 5.10

Tabela 5.10: Estatísticas pseudo-R^2, AIC e BIC dos modelos beta e de Wilson

Modelos	P-seudo	AIC	BIC
Modelo 1	70,77%	-482,44	-469,194
Modelo 2	70,44%	-482,23	-468,984
Modelo 3	70,81%	-482,46	-469,216
Modelo 4	70,16%	-482,05	-468,809
Wilson	70,97%	-138,66	-127,312

Estimação do modelo de Inadimplência (Enfoque bayesiano)

Nessa seção, adotou-se a abordagem bayesiana para analisar o conjunto de dados em questão. Para o cálculo das estimativas foi utilizado o software R v. 3.2.0 e o pacote Bayesianbetareg v.1.2.

Foi gerada uma cadeia de 50000 valores, nos quais os primeiros 20% foram descartados para eliminar o efeito não estacionário dos valores inicias da cadeia (*burn-in*), com os valores restantes, espaçados de 5 em 5, uma amostra de 8000 valores foi utilizada para definir a distribuição a posteriori dos parâmetros.

Inicialmente, foi testado um modelo com o parâmetro de precisão fixo (ϕ constante) e com uma priori vaga normal ($\beta's \sim N(0, 100)$) para os parâmetros de μ_i. No entanto, o algoritmo não convergiu. Alterando-se algumas condições, como o tamanho da cadeia, e diminuindo valor da variância do parâmetro de precisão, obteve-se os resultados dos Gráficos 5.9, 5.10, 5.11, 5.12, 5.13 e 5.14.

Figura 5.9: Série estimada de β_1 e β_2

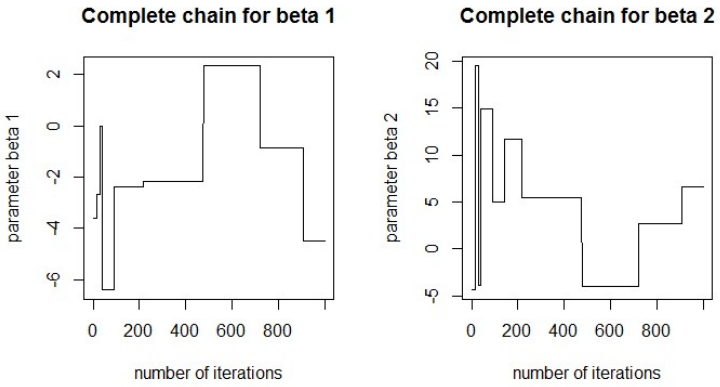


Figura 5.10: Série estimada de β_3 e β_4

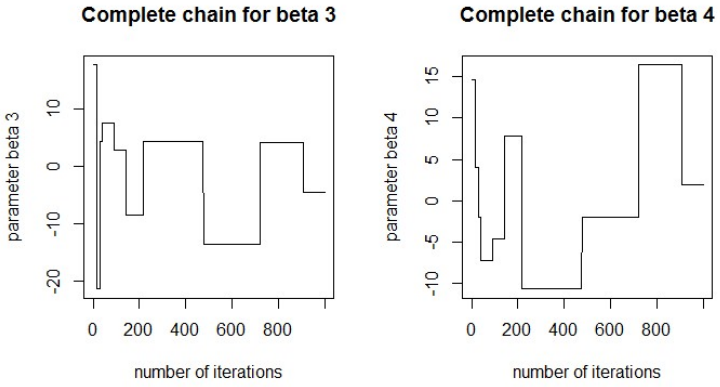


Figura 5.11: Série estimada de β_5 e γ_1

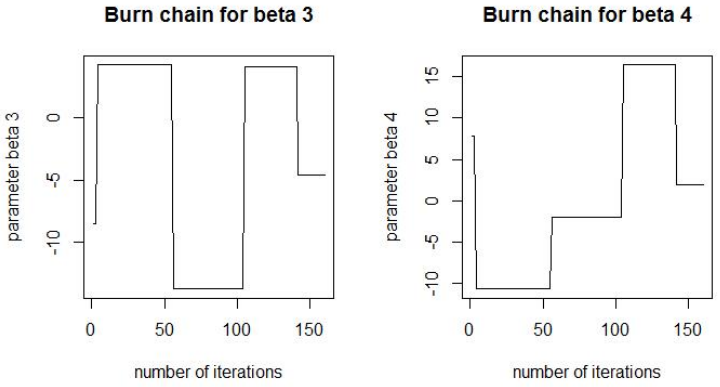


Figura 5.12: Série após *burn-in* para β_1 e β_2

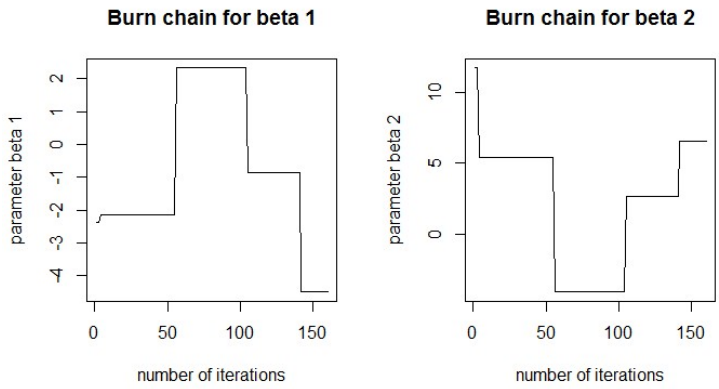


Figura 5.13: Série após *burn-in* para β_3 e β_4

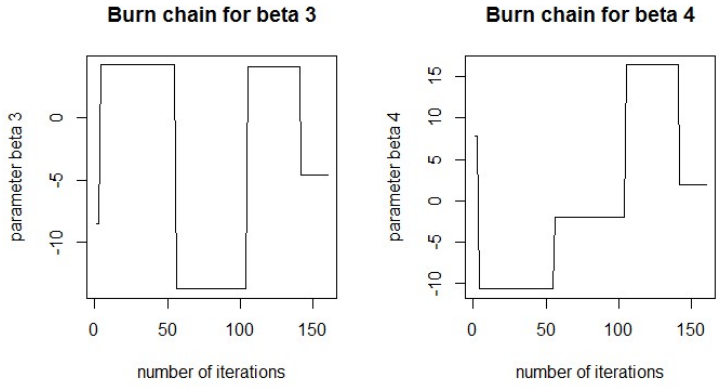
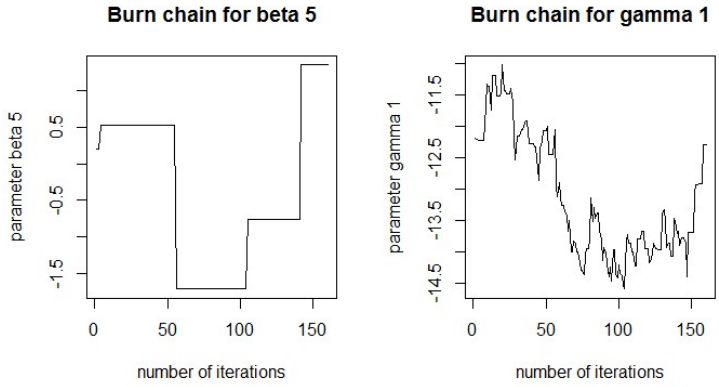


Figura 5.14: Série após *burn-in* para β_5 e γ_1



Como observado os betas não convergiram. Com isso, utilizou-se uma priori normal informativa no qual os valores iniciais dos parâmetros foram os valores obtidos da

regressão linear, sendo assim as distribuições dos parâmetros foram respectivamente: $\beta_0 \sim N(-1.43, 100)$, $\beta_1 \sim N(7.81, 100)$, $\beta_2 \sim N(-32.74, 100)$, $\beta_3 \sim N(-12.89, 100)$ e $\beta_4 \sim N(-0.51, 100)$.

Os resultados obtidos também não foram satisfatórios, os Gráficos 5.15, 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22 e 5.23 demonstram que não houve convergência dos estimadores.

Figura 5.15: Série após *burn-in* para β_1

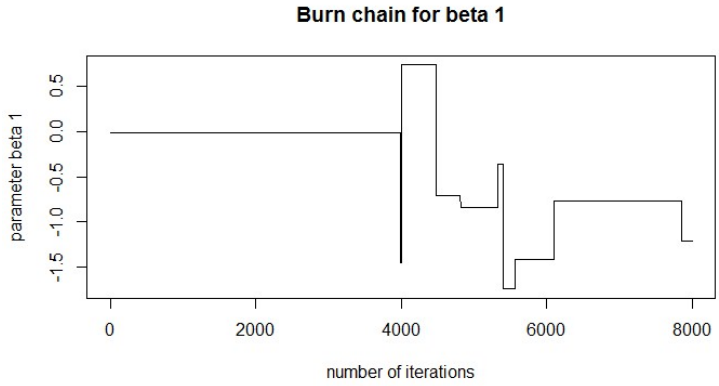


Figura 5.16: Série após *burn-in* para β_2

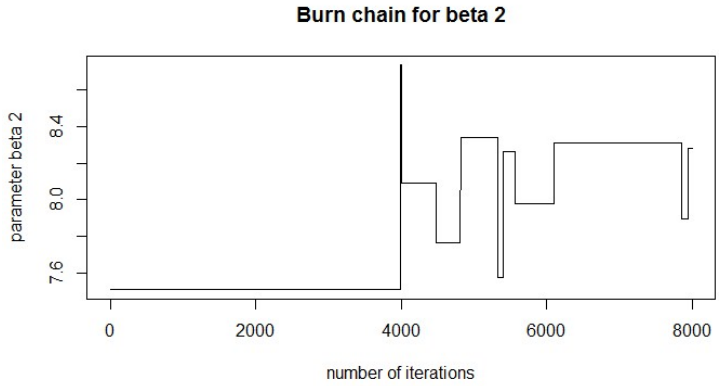


Figura 5.17: Série após *burn-in* para β_3

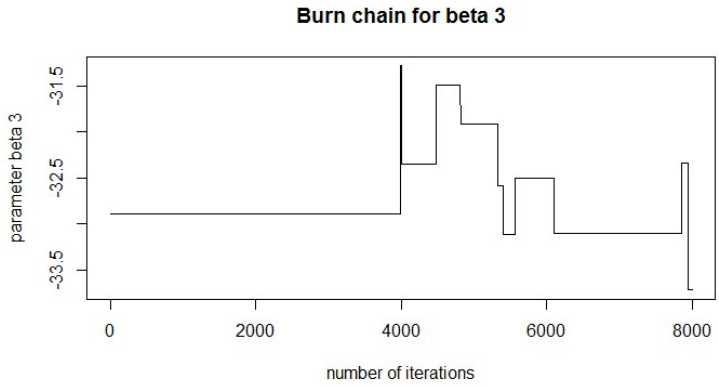


Figura 5.18: Série após *burn-in* para β_4

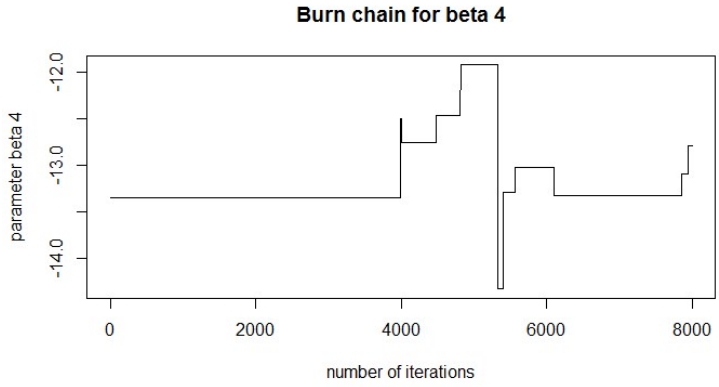


Figura 5.19: Série após *burn-in* para β_5

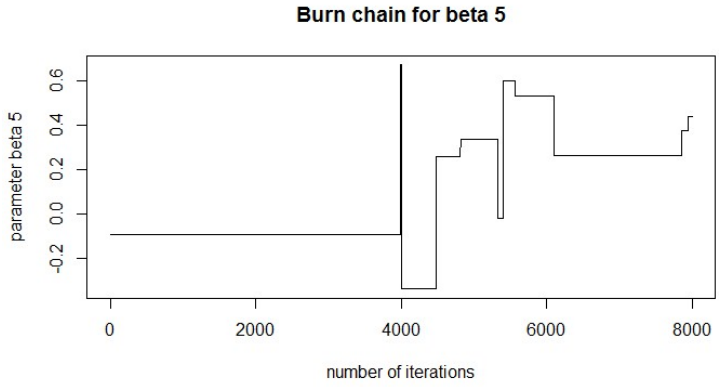


Figura 5.20: Série após *burn-in* para γ_1

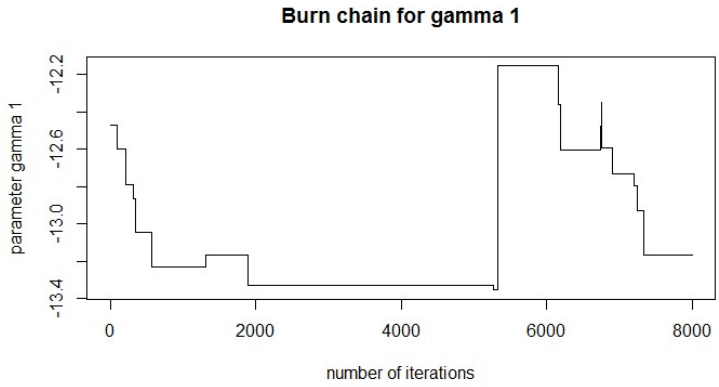


Figura 5.21: Resíduos padronizados versus índice da observação

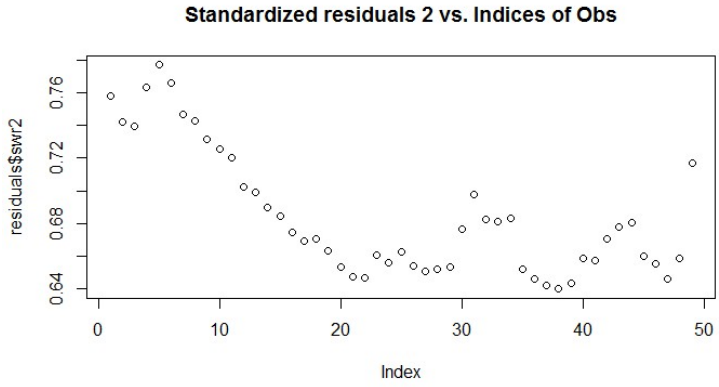


Figura 5.22: Q-Q plot

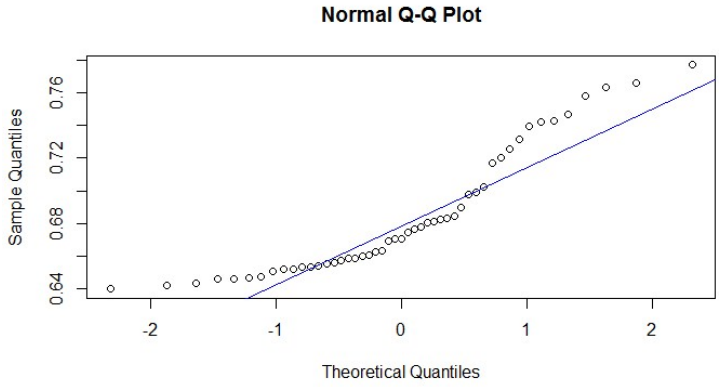
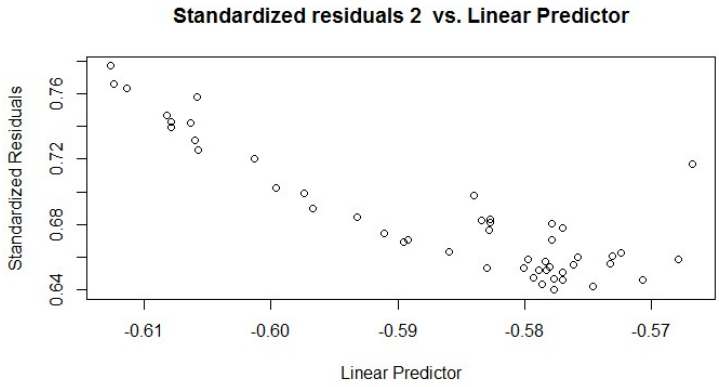


Figura 5.23: Resíduos padronizados versus preditor linear



Foram realizados testes de diagnóstico da cadeia, com a finalidade de verificar a convergência e a estacionariedade. A Tabela 5.11 abaixo apresentam-se os resultados.

Tabela 5.11: Diagnóstico de convergência da cadeia com ϕ constante

Parâmetro	Geweke Test		Heidelberger and Welch Test			
	Geweke Test	Result, GW Teste	Estacionariedade	P-valor	half-width test	Result, HW test
β_0	-0,6473	Aprovado	Aprovado	0,0904	0,7008	Reprovado
β_1	-3,2579	Reprovado	Reprovado	0,0017	NA	Reprovado
β_2	-0,3565	Aprovado	Aprovado	0,7089	0,0052	Aprovado
β_3	0,89	Aprovado	Aprovado	0,6910	0,0098	Aprovado
β_4	0,179	Aprovado	Reprovado	0,0162	NA	Reprovado
ϕ	2,856	Reprovado	Aprovado	0,8610	0,0333	Aprovado

Na Tabela 5.12 são apresentadas as estimativas dos parâmetros do modelo estimado. Verifica-se que os valores não são significativos, conforme demonstrado no intervalo de credibilidade, isso é justificável uma vez que não foi observada a convergência das estimativas.

Tabela 5.12: Modelo Beta Bayesiano com priori Normal - Função Logito - Precisão constante

Variáveis	Estimativas	Erro Padrão	LI	LS
Intercepto	-0,37	0,56	-1,41	0,74
Custo de Vida - Média dos ultimos 3 meses	7,84	3,03	7,51	8,34
Selic - Média dos ultimos 6 meses	-32,75	5,90	-33,11	-31,49
Desemprego - Defasado em 5 meses	-13,15	4,25	-13,34	-11,92
Câmbio - Média dos ultimos 6 meses	0,08	0,08	-0,34	0,53
ϕ	-13,03	0,15	-13,33	-12,16

Os mesmos procedimentos foram repetidos para o modelo com o parâmetro de precisão variando. Verifica-se que as estimativas também não convergiram, como demonstrado nas Tabelas 5.13 e 5.14 e nos Gráficos 5.24, 5.25, 5.26, 5.27, 5.28, 5.29, 5.30, 5.31 e 5.32, 5.33.

Tabela 5.13: Modelo Beta Bayesiano com priori Normal - Função Logito - Precisão variando

Variáveis	Estimativas	Erro Padrão	LI	LS
Intercepto - μ	-0,83	0,50	-1,80	0,36
Custo de Vida - Média dos últimos 3 meses	7,90	0,54	7,04	8,85
Selic - Média dos últimos 6 meses	-32,79	0,35	-33,54	-32,00
Desemprego - Defasado em 5 meses	-12,53	0,53	-13,71	-11,82
Câmbio - Média dos últimos 6 meses	0,25	0,26	-0,29	0,78
Intercepto ? ϕ	-11,30	0,41	-11,93	-10,69
Endividamento das Famílias	-4,80	0,62	-5,56	-3,58

Tabela 5.14: Diagnóstico Modelo Beta Bayesiano com priori Normal - Função Logito - Precisão variando

Parâmetro	Geweke Test		Heidelberger and Welch's Test			
	Geweke Test	Result. GW Teste	Estacionariedade	P-valor	half-width test	Result. HW test
β_0	-1,1322	Aprovado	Aprovado	0,0546	0,2843	Reprovado
β_1	2,602	Reprovado	Aprovado	0,8252	0,0226	Aprovado
β_2	1,2104	Aprovado	Aprovado	0,0766	0,0044	Aprovado
β_3	-1,9333	Aprovado	Aprovado	0,5401	0,0142	Aprovado
β_4	-0,2834	Aprovado	Aprovado	0,3904	0,3918	Reprovado
ϕ_0	3,896	Reprovado	Reprovado	0,0137	NA	Reprovado
ϕ_1	1,942	Aprovado	Aprovado	0,2142	0,0594	Aprovado

Figura 5.24: Série após *burn-in* para β_1 com ϕ variando

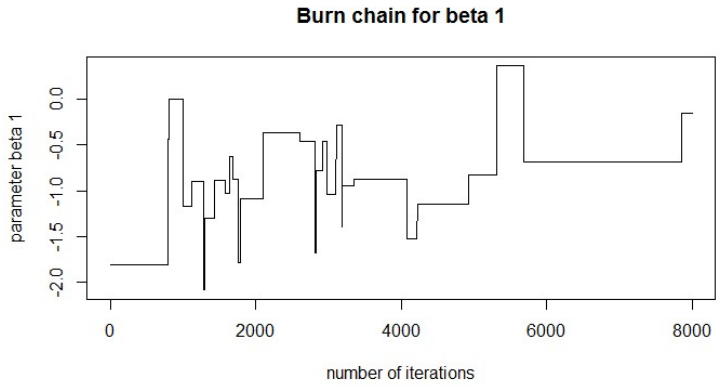
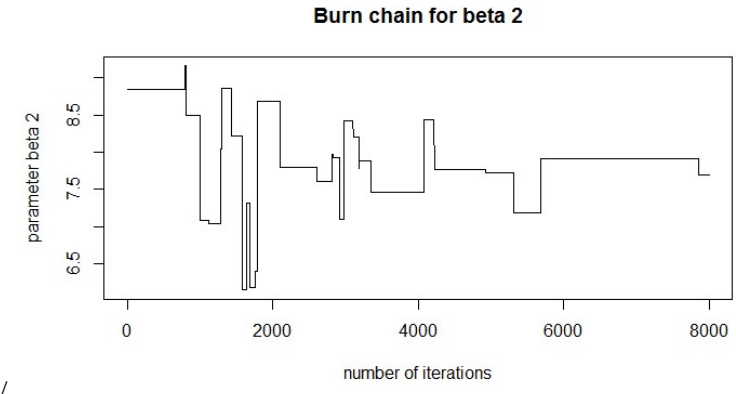


Figura 5.25: Série após *burn-in* para β_2 com ϕ variando



/

Figura 5.26: Série após *burn-in* para β_3 com ϕ variando

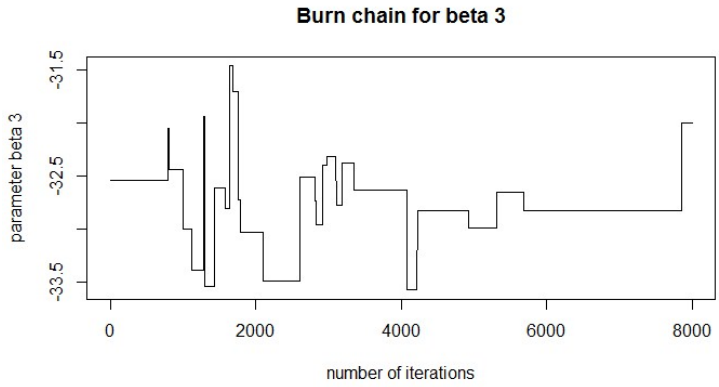


Figura 5.27: Série após *burn-in* para β_4 com ϕ variando

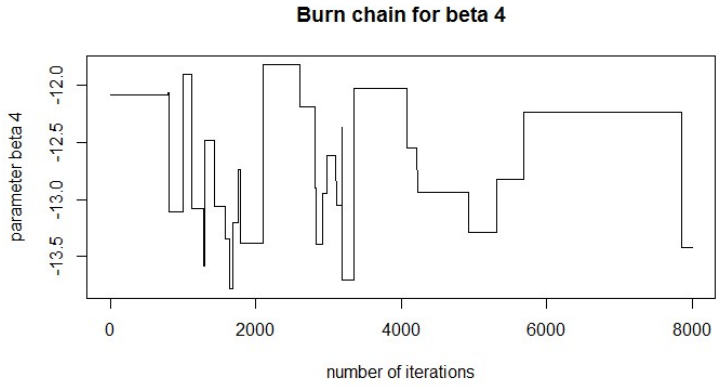


Figura 5.28: Série após *burn-in* para β_5 com ϕ variando

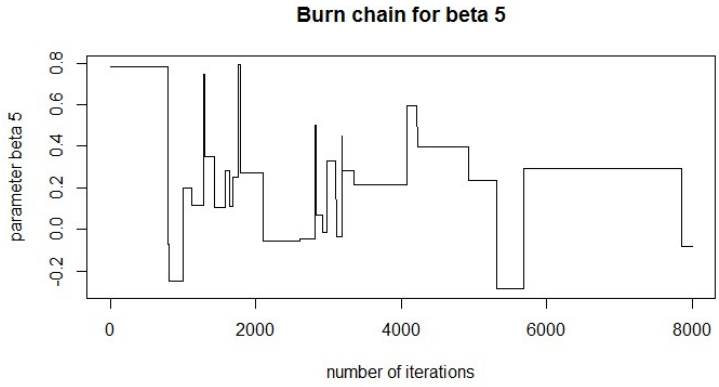


Figura 5.29: Série após *burn-in* para γ_1 com ϕ variando

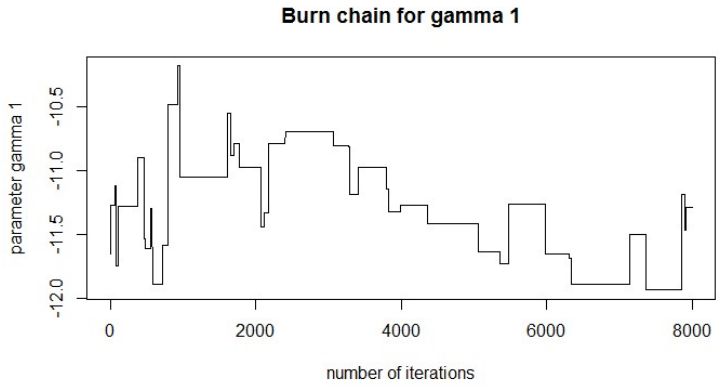


Figura 5.30: Série após *burn-in* para γ_2 com ϕ variando

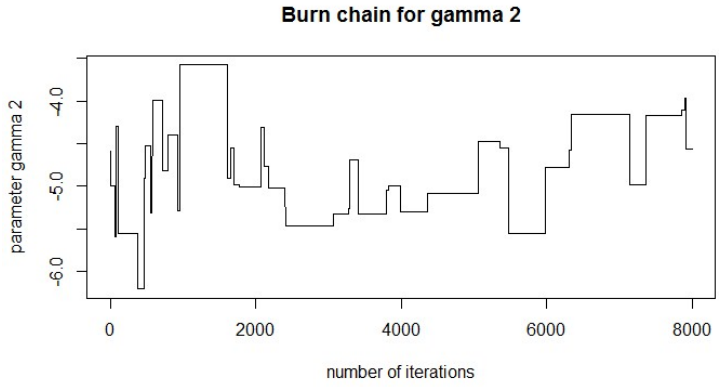


Figura 5.31: Resíduos padronizados versus índice da observação para ϕ variando

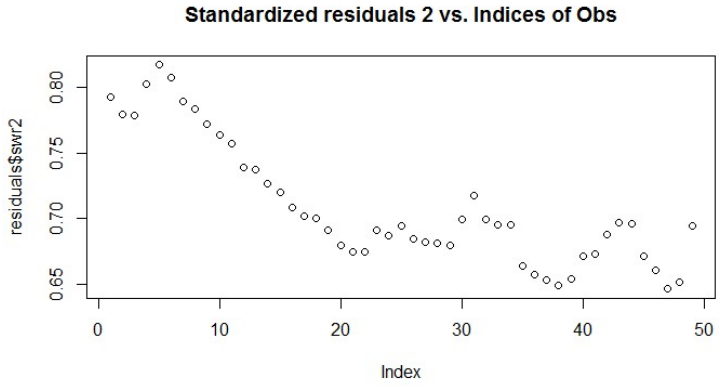


Figura 5.32: Q-Q plot

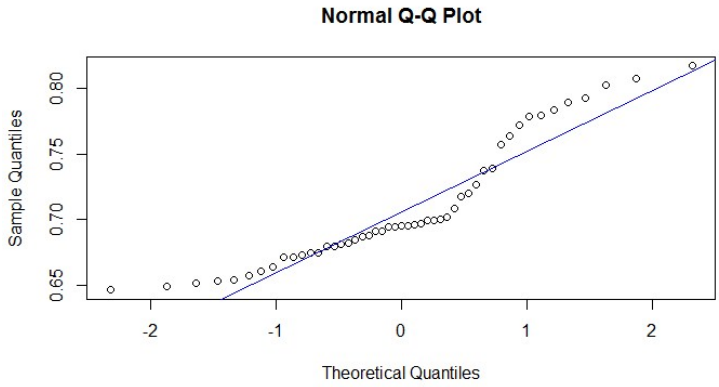
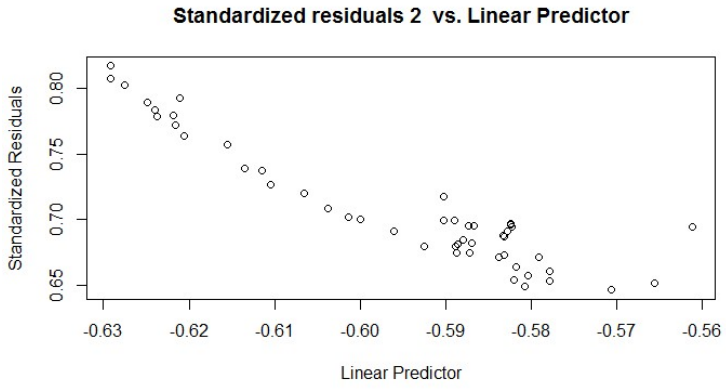


Figura 5.33: Resíduos padronizados versus preditor linear para ϕ variando



Assim como realizado nos modelos de regressão beta com enfoque clássico, foi realizada comparação entre os modelos bayesianos, segundo os critérios AIC e BIC, observamos que não há diferença significativa, conforme Tabela 5.15.

Tabela 5.15: Testes AIC e BIC para comparar modelos no enfoque bayesianio

Testes	Precisão	
	Constante	Variando
AIC	10,001	10,001
BIC	19,460	19,460

Os modelos com o enfoque bayesiano utilizando o pacote Bayesianbetareg não obtiveram resultados satisfatórios, uma vez que não houve convergência da cadeia,

mesmo utilizando uma priori informativa. No caso em estudo a adoção de prioris normais para os parâmetros de μ e de ϕ não obteve resultados satisfatório. Nesse sentido para este caso específico o uso de prioris deferente da normal é recomendável. Já os modelos beta sob o enfoque clássico obtiveram boas estimativas e segundo o critério de seleção de modelos AIC e BIC são melhores que os modelos tradicionais.

Capítulo 6

Conclusão e trabalhos futuros

A classe regressão beta, na qual a resposta é modelada por meio de estrutura de regressão, é um importante instrumento para modelar variáveis que assumem valores no intervalo unitário padrão $(0, 1)$. De forma geral, o uso desse modelo é mais apropriado do que os modelos usuais, uma vez que as características comuns a esse tipo de dado, tais como heterocedasticidade e assimetria, são respeitadas.

No tratamento dos dados simulados evidenciou-se bom ajuste do modelo beta, devido ao fato de os valores gerados estarem próximos aos observados. Também foram avaliados os resultados assintóticos em amostras finitas, que obtiveram resultados satisfatórios, independente da função de ligação adotada. Verificou-se que em amostras pequenas é necessário corrigir o viés dos estimadores de máxima verossimilhança. Apesar disso, esses estimadores apresentam boas propriedades em amostras relativamente pequenas ($n > 50$), pois se mostram quase não viciados e com distribuição próxima da normal de referência.

Sob o enfoque clássico, foram construídos modelos com quatro tipos de função de ligação e também foi verificada a variação do parâmetro de precisão. A inferência é realizada a partir de aproximações assintóticas, as quais são válidas em grandes amostras, mas podem conduzir a consideráveis distorções em amostras finitas, como visto nas simulações. A fim de se obter resultados mais confiáveis para amostras pequenas, alguns refinamentos foram utilizados no processo de modelagem, tais como a correção e redução do viés. Alternativamente, técnicas de bootstrap poderiam ser utilizadas para minimizar distorções. Foram verificados bons ajustes dos modelos beta

utilizando-se a abordagem clássica. As comparações entre as diferentes metodologias (modelos beta e do tipo Wilson) demonstraram que os modelos de regressão beta apresentam melhores ajustes em relação aos atuais.

No ajuste de um modelo de regressão beta, principalmente no caso do enfoque bayesiano, dificuldades computacionais foram observadas, isso devido ao fato de ser uma abordagem recente na literatura e dos métodos ainda não estarem totalmente implementados, embora se observem publicações com esse enfoque em modelos de regressão beta.

Sob o enfoque bayesiano, é necessário realizar alguns ajustes. Observa-se que mesmo com grande cadeia gerada, os estimadores não convergiram e nem se estabilizaram. Durante a simulação, a estimativa do parâmetro de precisão foi subdimensionada. O modelo final, utilizando dados reais, não convergiu, ou seja, não foi possível obter um bom ajuste. Em trabalhos recentes foi demonstrado, por meio do uso de modelos dinâmicos adotando o enfoque bayesiano, que é possível obter bom ajuste para variáveis restritas ao intervalo (0,1).

Por fim, verificou-se que os modelos de regressão beta para a modelagem da inadimplência para testes de estresse utilizando variáveis macroeconômicas são uma boa alternativa aos modelos com suporte na distribuição normal. Alguns refinamentos no modelo, como o uso de componentes autorregressivos, podem melhorar as estimativas (modelo beta autorregressivo e de média móvel). No entanto, observa-se que o ajuste do modelo beta é adequado e substancialmente melhor do que os atuais modelos em uso pelas instituições bancárias nacionais. Adotando o critério de seleção de modelo AIC e BIC, os modelos beta obtiveram melhores resultados.

Adicionalmente, propõe-se como sugestão para pesquisas futuras: modelagem da inadimplência utilizando a distribuição kumaraswamy em substituição à distribuição beta. A distribuição kumaraswamy é definida no intervalo $(0, 1)$ e suas funções de distribuição acumulada e de densidade têm forma fechada. Assim, a aplicação dessa distribuição é mais simples quando comparada ao emprego da distribuição beta.

Uma alternativa para os casos bimodais seria a construção de modelos beta, por meio de estrutura de regressão, que utilizam misturas de duas distribuições, como, por exemplo, a mistura de duas distribuições beta.

Desenvolver modelos de inadimplência do tipo β-ARMA, com a utilização de

variáveis macroeconômicas em teste de estresse, esse tipo de modelo permite a inclusão da dinâmica autorregressiva e de média móvel. Uma alternativa bayesiana é o desenvolvimento de modelos dinâmicos beta.

Desenvolver modelos com erro de medida nas variáveis respostas e modelos inflados de zero e um, sob a abordagem bayesiana. Em carteiras de crédito do tipo low default, a inadimplência quase não é observada, nesse sentido, o desenvolvimento de modelos inflados de zero seria apropriado a esses casos.

No âmbito computacional, seria interessante implementar algoritmo no R para geração da distribuição a posteriori que permita diferentes *prioris*, possibilitando uma melhor convergência dos estimadores.

Referências Bibliográficas

[1] ANDRADE, A. C. G. **Efeitos da especificação incorreta da função de ligação no modelo de regressão beta**. [Dissertação de mestrado do programa de pós-graduação do instituto de matemática e estatística da USP]. São Paulo, 2007.

[2] ANTUNES, A.; RIBEIRO, N.; ANTÃO, P. **Estimating probabilities of default under macroeconomic scenarios**. Banco de Portugal, 2005.

[3] BAKAR, K. S.; SAHU, S. K. **Sp timer: spatio-temporal bayesian modeling using R**. Journal of Statistical Software, v.63, n.15, 2015.

[4] BAYER, F. M. **Modelagem e inferência em regressão beta**. [Tese de doutorado do programa de pós-graduação do departamento de estatística da UFPE]. Recife, 2011.

[5] BAYER, F. M.; GUERRA, R. R. **Provisão de crédito consignado: uma aplicação do modelo beta autorregresivo de média móveis**. Associação paranaense de engenharia de produção, 2014.

[6] BCBS. **International convergence of capital measurement and capital standards**. Basle: Basle Committee on Banking Supervision, 1988.

[7] _____. **Amendment to the capital accord to incorporate market risks**. Basle: Basle Committee on Banking Supervision, 1996.

[8] _____. **Capital requirements and bank behaviour: the impact of the Basel Accord**. Basileia. Basel Committee on Banking Supervision. Abr.,1999.

[9] _____. **International convergence of capital measurement and capital standards: a revised framework**. Basle: Basle Committee on Banking Supervision, Jun. 2004.

[10] _____. **The New Basle Accord**. Basle: Basle Committee on Banking Supervision, 2003.

[11] _____. **The New Basle Accord: an explanatory note**. Basle: Basle Committee on Banking Supervision, Jan. 2001a.

[12] _____. **The Standardized approach to credit risk**. Basle: Basle Committee on Banking Supervision, Jan. 2001b.

[13] BRASIL. Conselho Monetario Nacional - CMN.Resolucao 2.682, de 21 de dezembro de 1999. Dispoe sobre criterios de classificacao das operacoes de credito e regras para constituicao de provisao para creditos de liquidacao duvidosa. Brasilia, 1999.

[14] BUNN, P.; CUNNINGHAM, A.; DREHMANN, M. Stress testing as a tool for assessing systemic risk. Bank of England financial stability, p. 116-126, 2005.

[15] CARRASCO, J. M. F. **Modelos de regressão beta com erro nas variáveis**. [Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP]. São Paulo, 2012.

[16] CASARIN, R.; VALLE, L.D.; LEISEN, F. **Bayesian model selection for beta autoregressive processes**. Bayesian anal, v.7, n.2, 2012

[17] CORDEIRO,G.M. **Introdução a teoria assintótica**. Instituto de Matemática Pura e Aplicada (IMPA),Rio de Janeiro,1999.

[18] CORREIA, L. T. **Modelos dinâmicos para dados agregados** [Dissertação de mestrado do programa de pós-graduação do instituto de ciências exatas da UNB]. Brasília, 2010.

[19] COX,D.R.;REID.N. **Parameter orthogonality and approximate conditional inference**. Journal of the Royal Statistical Society v. 49, pg. 1-39, 1987.

[20] CRIBARI-NETO,F.;CORDEIRO,G.M. **On Bartlett an Bartlett-type corrections**. v.15,pg. 339-367, 1996.

[21] CRIBARI-NETO, F.; FERRARI, S. L. P. **Beta regression for modelling rates and proportions**. Journal of applied statistics. V. 31, n. 7, 2004

[22] CRIBARI-NETO, F.; ROCHA, A. V. **Beta autoregressive moving average models**. Springer, 2009

[23] CRIBARI-NETO, F.; ZEILEIS, A. **Beta regression in R**. Journal of Statistical Software, 2010.

[24] DAMODARAN, A (2004). **Finanças corporativas: teoria e prática**. Bookman, Porto Alegre, 2. ed., 796 p. JORION, P. Value at risk: the new benchmark for controlling market risk. ed., 1997. 544 p.Mcgraw-Hill, Chicago, 2.

[25] DANTAS, J.; ROBERTO, P.; MEDEIROS, O. **Validação de modelo ampliado para estimação da discricionariedade da PCLD em bancos.** 12^0 **Congresso USP de controladoria e contabilidade**. São paulo, 2012

[26] DAVISON,A.C.;HINKLEY,D.V. **Bootstrap methods and their application**. Cambridge University Press, Cambridge,1997.

[27] DEY, S.; TESSIER, D.; MISINA, M. **Stress testing the corporate loans portfolio of Canadian banking sector**. Bank of Canada working paper, Ottawa,2006.

[28] ESPINHEIRA,P.L.;FERRARI,S.L.P.;CRIBARI-NETO,F. **On beta regression residuals**. Journal of applied statistics v.35, pg. 407-419,2008.

[29] ESPINHEIRA,P.L.;FERRARI,S.L.P.;CRIBARI-NETO,F. **Influence diagnostics in beta regression**. Computational Statistics & Data Analysis v.52, pg. 4417-4431,2008.

[30] FERRARI,S.L.P.;CYSNEIROS,A.H.M.A. **Skovgaard's adjustment to likelihood ratio tests in exponential family nonlinear models**.Statistics and Probability Letters v.78,pg. 3049-3057, 2008.

[31] FERRARI, S.L.P.;PINHEIRO,E.C. **Improved likelihood inference in beta regression**. Journal os Statistical Computation and Simulation. v.81,pag. 431-443, 2011.

[32] FERRARI, S. L. P.; ARELLANO-VALLE, R. B.; FIGUEROA-ZÚÑIGA, J. I. **A mixed beta regression: a bayesian perspective**. Computational Statistics & Data Analysis, 2012.

[33] FLÓREZ, O. P. R. **Identificação de pontos influentes em uma amostra aleatória de pré-formas da distribuição de bingham complexa (distância de cook e método de bootstrap)**. [Dissertação de mestrado do programa de pós-graduação do departamento de estatística da UFPE]. Recife, 2009.

[34] FREITAS, J. T. **Acordo de Basiléia 2 e estabilidade financeira em paises em desenvolvimento**. 2008. 171f. Dissertação (Mestrado em Ciências Econômicas). Instituto de Economia. Universidade Estadual de Campinas, Campinas, SP, 2008.

[35] FUNGÁČOVÁ, Z.; JAKUBÍK, P. **Bank stress tests as an information device for emerging markets: The case of Russia**. Czech Journal of economics and finance, v.63, n. 1, pg. 87-105, 2013.

[36] GAGLIANONE, W. P.; SCHECHTMAN, R. **Macro stress testing of credit risk focused on the tails**. Banco central do Brasil working paper series 241, Brasília, 2011.

[37] GAGLIANONE, W. P.; SCHECHTMAN, R. **Teste de estresse na ligação macro-risco de crédito: uma aplicação ao setor doméstico pessoa física**. IV seminário sobre riscos, Estabilidade financeira e economia bancária do Banco central do Brasil, Brasília, 2009.

[38] HOGGARTH, G.; SORENSEN, S.; ZICCHINO, L. **Stress test of UK banks using a VAR approach**. Bank of England, working paper 282, 2005.

[39] JACOBSON, T.; CARLING, K.; LINDE, J. ROSZBACH, K. **Corporate credit risk modeling and the macroeconomy**. Journal of Banking & Finance, 2007.

[40] JONES, M. T.; HILBERS, P.; SLCK, G. **Stress testing financial systems: What to do when the governor calls**. International monetary fund, Working paper 4, 2004.

[41] KALIRAI, H.; SCHEICHER, M. **Macroeconomic stress testing: Preliminary evidence for Austria**. Austrian National Bank, working paper 3, 2002.

[42] KOOPMAN, S. J.; LUCAS, A. **Business and default cycles for credit risk**. Journal of applied econometrics, v.20, n.2, pg. 311-323, 2005.

[43] LÓPEZ, F. O. A **bayesian approach to parameter estimation in simplex regression model: a comparison with beta regression**. Revista colombiana de estadística, v. 36, n. 1, Colombia, 2013.

[44] LU, W.; YANG, Z. Stress testing of comercial bank's exposure to credit risk: study based on write-off nonperforming loans. Asian Social Science, v.8, n. 10, pg. 16-22, 2012.

[45] LUCENA, S. E. F. **Testes de hipóteses não-encaixadas em regressão beta**. [Dissertação de mestrado do programa de pós-graduação do departamento de estatística da UFPE]. Recife, 2013.

[46] MANCO, O. C. U. **Modelos de regressão beta com efeitos aleatórios normais e não normais para dados longitudinais**. [Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP]. São Paulo, 2013.

[47] MARQUES, L. F. B. **Gerenciamento do Risco de Crédito & Cálculo do Risco de Crédito para a Carteira de um Banco do Varejo**. Dissertação de Mestrado, UFRG. Porto Alegre: fevereiro de 2002.

[48] MARTÍNEZ, R. O.**Modelos de regressão beta inflacionados** [Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP]. São Paulo, 2008.

[49] MASAROTTO, G.; VARIN C. **Gaussian copula marginal regression**. Electronic Journal of Statistical, 2012.

[50] MENDONÇA, A. R. R. **Os acordos da Basiléia: uma avaliação do novo formato da regulação bancária**. 2002. Tese (Doutorado em Ciências Econômicas). Programa de Pós-Graduação em Economia. Universidade Estadual de Campinas, Campinas, SP, 2002.

[51] MCCULLAGH,P.;NELDER,J.A. **Generalized Linear Models**, 2nd ed. London: Chapman and Hall,1989.

[52] MIGON, H.S.; GAMERMAN, D. **Statistical Inference an Integrated Approach**. Lodon,1999.

[53] NISHIKAWA, W. E. **Modelo de estresse macroeconômico da inadimplência para bancos do atacado**. [Dissertação de mestrado do programa de pós-graduação em economia da FGV]. São Paulo, 2014

[54] NOGAROTTO, D. C. **Inferência bayesiana em modelos de regressão beta e beta inflacionados** [Dissertação de mestrado do programa de pós-graduação do instituto de matemática, estatística e computação científica da Unicamp]. Campinas, 2013.

[55] NOCEDAL,J.;WRIGHT,S.J. **Numerical Optimization**. New York: Springer-Verlag,1999.

[56] OSPINA,R. **Modelos de regressão beta inflacionados**.[Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP].São Paulo,2008.

[57] OSPINA,R.;CRIBARI-NETO,F.;VASCONCELOS,K.L.P. **Improved point and interval estimation for beta regression model**. Computational Statistics & Data Analysis v. 51, pg. 960-981, 2006.

[58] PINHEIRO, E.C. **Ajuste para teste de razão de verossimilhança em modelos de regressão beta**. [Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP].São Paulo,2009.

[59] PEREIRA, G. H. A. **Modelos de regressão beta inflacionados truncados** [Tese de doutorado do programa de pós-graduação do instituto de matemática e estatística da USP]. São Paulo, 2012.

[60] PEREIRA, L. A.; TAVARES, M. **Regressão beta para modelagem do rendimento metalúrgico na reciclagem de alumínio**. XII escola de modelos de regressão, Fortaleza, 2011.

[61] PILINKO, V.; ROMANCENCO, A. **A macro financial model for credit risk stress testing: the case of Latvia** [Tese de doutorado do programa de pós-graduação Stockholm school of economics]. Riga, 2014.

[62] QUEIROZ, M. P. F. **Testes de hipóteses em regressão beta baseados em verossimilhança perfilada ajustada e em bootstrap**. [Dissertação de mestrado do programa de pós-graduação do departamento de estatística da UFPE]. Recife, 2011.

[63] RAO,C.R. **Linear statistical inference and its application**, 2nd ed. New York: Wiley,1973.

[64] REITMAN, D. P. **Uso de métodos clássicos e bayesianos em modelos de regressão beta** [Dissertação de mestrado do programa de pós-graduação do departamento de estatística da UFSCar]. São Carlos, 2007.

[65] RESTI, A.; SIRONI, A.**Loss Given Default and Recovery Risk: From Basel II Standards to Effective Risk Management Tools**. The Basel Handbook: A Guide for Financial Effective p. 49-82, 2004.

[66] RODRIGUES, G. S. **Modelos dinâmicos dirichlet** [Dissertação de mestrado do programa de pós-graduação do instituto de ciências exatas da UNB]. Brasília, 2011.

[67] SANATANA, T. V. F. **As distribuições KUMARASWAMY-log-logística e KUMARASWAMY-logística**. [Dissertação de mestrado do programa de pós-graduação da escola superior de agricultura luiz queiroz da USP]. Piracicaba, 2010.

[68] SANTOS, T. **Testes de stress em sistemas financeiros: Uma aplicação no Brasil**.[Dissertação de mestrado do programa de pós-graduação da faculdade de economia,administração e contabilidade da USP]. São Paulo, 2008.

[69] SIMAS, A. B.; BARRETO-SOUZA, W.; ROCHA, A. V. **Improved estimators for a general class of beta regression models**. Computational Statistics & Data Analysis, 2010.

[70] SIMONS, D.; ROLWES, F. **Macroeconomic default modelling and stress testing**. International Journal of Central Banking, v.5, n. 3, 2009

[71] SILVA, C. Q.; MIGON, H. S.; CORREIA, L. T. **Dynamic Bayesian beta models**. Computational Statistics & Data Analysis, v. 55, n. 6, p. 2074-2089, 2011.

[72] SMITHSON, M.; VERKUILEN, J. **A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables**. Psychological Methods, v.11(1), p.54-71, 2004

[73] SOBREIRA, R.; MARTINS, N. M. **Os acordos de Basiléia e bancos de desenvolvimento no Brasil: uma avaliação do BNDES e do BNB**. Revista de Administração Publica, Rio de Janeiro, v. 45, no. 2, p. 349-376, 2006.

[74] SORGE, M.; VIROLAINEN, K. **A comparative analysis of macro stress-testing methodologies with application to Finland**. Journal of financial stability, v.2, n.2, pg. 113-151, 2006.

[75] SOUTO, M.; TABAK, B. M.; VAZQUEZ, F. **A macro stress test model of credit risk for the Brazilian banking sector**. Banco central do Brasil working paper series 226, Brasília, 2010.

[76] SOUZA, D. F. **Regressão beta multivariada com aplicações em pequenas áreas**. [Tese de doutorado do programa de pós-graduação em estatística do instituto de matemática da UFRJ]. Rio de janeiro, 2011.

[77] STASINOPOULOS, D. M.; RIGBY, R. A. **Generalized additive models for location scale and shape (GAMLSS) in R**. Journal of Statistical Software, v.23, 2007.

[78] SULAIMAN, M. Y. et al. **Application of beta distribution model to Malaysian sunshine data**. Renewable energy, v. 18, n. 4, p. 573-579, 1999.

[79] Valentiny-Endrész, M.; Vásáry, Z. **Macro stress testing with sector specific bankruptcy models**. MNB Working Papers,2008

[80] VIROLAINEN, K. **Macro stress testing with a macroeconomic credit risk model for finland**. Bank of Finland discussion papers 18, Helsinki, 2004.

[81] VLIEGHE, G. W. **Indicators of fragility in the UK corporate sector**. Bank of England, working paper 146, 2001.

[82] WALD A. **Tests of statistical hypotheses concerning several parameters when the number of observations is large**. Transactions of the American Mathematical Society v.54, pg. 426-482, 1943.

[83] WILSON, T. (1997). **Portfolio Credit Risk (II)**. Risk, 10, 56-61.

[84] WILSON, T. (1997a). **Portfolio Credit Risk (I)**, Risk, September, 111-117.

[85] ZANIBONI, N. C. **A inadimplência do sistema financeiro no Brasil explicada por meio de fatores macroeconômicos**. [Dissertação de mestrado do programa de pós-graduação do departamento de administração da faculdade de economia da USP]. São Paulo, 2013.

Apêndice A

Função Gama

Uma das funções especiais importantes é a função gama. Esta função é denotada e definida por:

$$\Gamma(s) = \int_0^\infty \epsilon^{-t} t^{s-1} dt, para \qquad s > 0 \tag{A.1}$$

Esta função está bem definida para $s > 0$, pois:

$$\int_0^\infty \epsilon^{-t} t^{s-1} dt = \int_0^1 \epsilon^{-t} t^{s-1} dt + \int_1^\infty \epsilon^{-t} t^{s-1} dt \tag{A.2}$$

Em ambas as integrais do lado direito de B.3 convergem. Para provar esse fato, basta usar o critério da comparação de integrais impróprias, o qual afirma que: Se $f(x) \geqslant 0$ e $g(x) \geqslant 0$, para todo $x > a$, se as integrais impróprias $\int_a^b f(x)\,dx$ e $\int_a^b g(x)\,dx$ existem, para todo $b \geqslant a$, e se $\lim_{x \longrightarrow \infty} \dfrac{f(x)}{g(x)} = 0$, então:

Para $c \neq 0$, ambas as integrais $\int_a^b f(x)\,dx$ e $\int_a^b g(x)\,dx$ convergem ou divergem

Para $c = 0$, a convergência de $\int_a^b f(x)\,dx$ implica na convergência de $\int_a^b g(x)\,dx$

Assim, como $\int_0^\infty t^{(-2)} dt$ converge e $\lim_{x \longrightarrow \infty} \dfrac{\epsilon^{-t} t^{s-1}}{t^{-2}} = 0$, obtemos a convergência da segunda integral imprópria, $\int_1^\infty \epsilon^{-t} t^{(s-1)} dt$. Para analisarmos a convergência da primeira integral, vamos fazer a mudança da variável $t = \dfrac{1}{u\prime s}$, com $dt = -u^{-2} du$. Assim:

$$\int_0^1 \epsilon^{-t} t^{(s-1)} dt = \int_1^\infty \epsilon^{-\frac{1}{u}} u^{-s-1} du \tag{A.3}$$

A qual converge por comparação com a integral, $\int_1^\infty u^{-s-1} du$ converge se $s > 0$. Logo a função $\Gamma(s)$ está bem definida pela integral B.1

Como conclusão, a função Gama tem como domínio o conjunto de todos os reais positivos, embora seja possível estender seu domínio a todo o conjunto real, com exceção dos inteiros negativos e do zero. Devemos ressaltar que esta extensão não é feita através da definição, pois a integral diverge, mas sim através da propriedade abaixo.

1. $\Gamma(s+1) = s\Gamma(s)$, se $s > 0$
2. $\Gamma(1) = 1$
3. Se n é um inteiro positivo, então $\Gamma(n+1) = n!$
4. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

A.1 Função Digama

A derivada do logaritmo da função gama é chamada de função digama

$$\psi(x) = d\ln(\Gamma(x))\,dx = \frac{\Gamma'(x)}{\Gamma(x)} \tag{A.4}$$

em que $\Gamma(.)$ é a função gama e ln o logaritmo natural

A.2 Função Trigama

É dada pela segunda derivada da função gama

$$\psi'(x) = \frac{d^2}{dx^2}\ln\Gamma(x) \tag{A.5}$$

Apêndice B

Critério de Seleção de Modelos

Um modelo é a representação simplificada de algum problema ou situação da vida real destinado a ilustrar certos aspectos do problema sem se ater a todos os detalhes. Não raro, mais de um modelo pode descrever um mesmo fenômeno, haja vista que cada pesquisador tem a liberdade de modelar o fenômeno seguindo a metodologia que julgar mais adequada. Aqui a seleção do "melhor"modelo torna-se então evidente. Dentre as diversas metodologias utilizadas para este fim, neste trabalho utilizou-se os critérios de informação de Akaike (AIC) e Bayesiano de Schwarz (BIC), para a seleção de modelos.

B.1 Critério de informação de Akaike (AIC)

O Critério de Informação de Akaike (AIC) admite a existência de um modelo "real"que descreve os dados que é desconhecido, e tenta escolher dentre um grupo de modelos avaliados, o que minimiza a divergência de Kullback-Leibler ($K - L$).

O valor de $K - L$ para um modelo f com parâmetros θ, em relação ao modelo ?real? representado por g é dado por

$$l\left(g, f\left(\theta\right)\right) = \int g\left(y\right) \ln\left(\frac{g\left(y\right)}{f\left(y|\theta\right)}\right) dy \tag{B.1}$$

Esta divergência está relacionada à informação perdida por se usar um modelo aproximado e não o "real". A estimativa do AIC para um determinado modelo é dada por:

$$AIC = -2L + 2K \tag{B.2}$$

em que, ln é o logaritmo da função de máxima verossimilhança do modelo com os parâmetros θ e K o número de parâmetros. O modelo com menor valor e AIC é considerado o modelo de melhor ajuste.

B.2 Critério Bayesiano de Schwarz (BIC)

O Critério Bayesiano de Schwarz (BIC) tem como pressuposto a existência de um "modelo verdadeiro"que descreve a relação entre a variável dependente e as diversas variáveis explanatórias entre os diversos modelos sob seleção. Assim o critério é definido como a estatística que maximiza a probabilidade de se identificar o verdadeiro modelo dentre os avaliados. O valor do critério BIC para um determinado modelo é dado por:

$$BIC = -2L + 2K \ln(n) \tag{B.3}$$

Em que n é o número de observações. O modelo com menor BIC é considerado o de melhor ajuste.

Apêndice C

Anexo C

C.1 Programas Simulação

Neste anexo são apresentados os programas utilizados nas simulações e na construção dos modelos.

```
#### Função de Ligação Logito ####
n=1000
library(betareg);
library(lmtest);
library(moments);
set.seed(15);
y=rbeta(n,5,3);
x1=runif(n,0,1);
x2=rnorm(n,0,1);
x3=rexp(n,3);
x4=rgamma(n,4,3)
x5=rlnorm(n,0,1)
xdata=cbind(x1,x2,x3,x4);
y=cbind(y);

n1=50
```

```
mua=numeric(); a=numeric(); p=numeric(); q=numeric();
b=matrix(NA,n1,1000);
mua_hat=matrix(NA,n1,1000);
p_hat=matrix(NA,n1,1000);
q_hat=matrix(NA,n1,1000);
y_hat=matrix(NA,n1,1000);
Dif=matrix(NA,n1,1000);
soma=0
EMQ=numeric()
estimativas=matrix(NA,14,1000)
Estatistica_Wald=numeric()

y1=matrix(NA,n1,1000);

x1a=numeric(); x2a=numeric(); x3a=numeric(); x4a=numeric(); x5a=numeric()

trv=numeric();

cont1a=0; cont2a=0; cont3a=0; cont4a=0;

wald1a=0; wald2a=0; wald3a=0;wald4a=0;

b0=0.5; b1=-0.02; b2=0.3; b3=-0.15; b4=0.05; b5=-0.07; phi=8.2

for(t in 1:1000){
for(i in 1:n1){
x1a[i]=x1[i]
x2a[i]=x2[i]
x3a[i]=x3[i]
x4a[i]=x4[i]
x5a[i]=x5[i]
a[i]=exp(b0 + b1*x1a[i] + b2*x2a[i] + b3*x3a[i] + b4*x4a[i] + b5*x5a[i])
```

```
mua[i]=(a[i]/(1+a[i]))
p[i]=mua[i]*phi
q[i]=(1-mua[i])*phi
y1[i,t]=rbeta(1,p[i],q[i])}

if (y1[i,t]== 1 | y1[i,t]==0 ) {
y1[i,t]=rbeta(1,p[i],q[i])
}
if (y1[i,t]== 1 | y1[i,t]==0 ) {
y1[i,t]=rbeta(1,p[i],q[i])
}
}

XA=cbind(x1a,x2a,x3a,x4a,x5a)
XAa=cbind(x2a,x4a)

for(k in 1:1000){
loop=betareg(y1[,k]~XA,link="logit")$coefficients
loop2=betareg(y1[,k]~XA,link="logit")$vcov

estimativas[1,k]=loop$mean[1]
estimativas[2,k]=loop$mean[2]
estimativas[3,k]=loop$mean[3]
estimativas[4,k]=loop$mean[4]
estimativas[5,k]=loop$mean[5]
estimativas[6,k]=loop$mean[6]
estimativas[7,k]=loop$precision[1]
estimativas[8,k]=sqrt(diag(loop2))[1]
estimativas[9,k]=sqrt(diag(loop2))[2]
estimativas[10,k]=sqrt(diag(loop2))[3]
estimativas[11,k]=sqrt(diag(loop2))[4]
estimativas[12,k]=sqrt(diag(loop2))[5]
```

```
estimativas[13,k]=sqrt(diag(loop2))[6]
estimativas[14,k]=sqrt(diag(loop2))[7]

}

for(k in 1:1000){
soma[k]=0
for(i in 1:n1){
b[i,k]=exp(estimativas[1,k]
+ estimativas[2,k]*x1a[k]
+ estimativas[3,k]*x2a[k]
+ estimativas[4,k]*x3a[k]
+ estimativas[5,k]*x4a[k]
+ estimativas[6,k]*x5a[i])

mua_hat[i,k]=(b[i,k]/(1 + b[i,k]))
p_hat[i,k]=mua_hat[i,k]*estimativas[7,k]
q_hat[i,k]=(1-mua_hat[i,k])*estimativas[7,k]
y_hat[i,k]=(p_hat[i]/(p_hat[i]+q_hat[i]))
Dif[i,k]=(y1[i,k]-y_hat[i,k])^2
soma[k]=soma[k]+Dif[i,k]
}

EMQ[k]=soma[k]/n1

trv[k]=2*(logLik(betareg(y1[,k]~ XA, link="logit"))
-logLik(betareg(y1[,k] ~ XAa, link="logit")))

W1=betareg(y1[,k]~x1a + x2a + x3a + x4a + x5a, link="logit")
w2=betareg(y1[,k]~x2a + x4a, link="logit")
Estatistica_Wald[k]=waldtest(W1,w2)$Chisq[2]
}
```

```
for(k in 1:1000){
if (trv[k] > qchisq(0.90,3)) cont1a=cont1a+1 else cont1a=cont1a
if (trv[k] > qchisq(0.95,3)) cont2a=cont2a+1 else cont2a=cont2a
if (trv[k] > qchisq(0.99,3)) cont3a=cont3a+1 else cont3a=cont3a
if (trv[k] > qchisq(0.995,3)) cont4a=cont4a+1 else cont4a=cont4a
if ((Estatistica_Wald[k]) > qchisq(0.90,3)) wald1a=wald1a+1 else wald1a=wald1a
if ((Estatistica_Wald[k]) > qchisq(0.95,3)) wald2a=wald2a+1 else wald2a=wald2a
if ((Estatistica_Wald[k]) > qchisq(0.99,3)) wald3a=wald3a+1 else wald3a=wald3a
if ((Estatistica_Wald[k]) > qchisq(0.995,3)) wald4a=wald4a+1 else wald4a=wald4a
}

qqplot(qnorm(ppoints(1000),0,1),trv,main="Simulação TRV n=50")
abline(0,1)
qqnorm(trv)
qqnorm(Estatistica_Wald,main="Wald n=50")
abline(0,1)


porcen_de_rejei1a_trv=cont1a/1000;
porcen_de_rejei1a_trv;
porcen_de_rejei2a_trv=cont2a/1000;
porcen_de_rejei2a_trv;
porcen_de_rejei3a_trv=cont3a/1000;
porcen_de_rejei3a_trv;
porcen_de_rejei4a_trv=cont4a/1000;
porcen_de_rejei4a_trv;

porcen_de_rejei1a_wald=wald1a/1000;
porcen_de_rejei1a_wald;
```

```
porcen_de_rejei2a_wald=wald2a/1000;
porcen_de_rejei2a_wald;
porcen_de_rejei3a_wald=wald3a/1000;
porcen_de_rejei3a_wald;
porcen_de_rejei4a_wald=wald4a/1000;
porcen_de_rejei4a_wald;
mEMQ=mean(EMQ)
mEMQ


beta_zero=estimativas[1,]
beta_um=estimativas[2,]
beta_dois=estimativas[3,]
beta_tres=estimativas[4,]
beta_quatro=estimativas[5,]
beta_cinco=estimativas[6,]
phi=estimativas[7,]
dsvb0=estimativas[8,]
dsvb1=estimativas[9,]
dsvb2=estimativas[10,]
dsvb3=estimativas[11,]
dsvb4=estimativas[12,]
dsvb5=estimativas[13,]
dsvphi=estimativas[14,]

beta_zero=as.numeric(beta_zero)
beta_um=as.numeric(beta_um)
beta_dois=as.numeric(beta_dois)
beta_tres=as.numeric(beta_tres)
beta_quatro=as.numeric(beta_quatro)
beta_cinco=as.numeric(beta_cinco)
phi=as.numeric(phi)
```

```
dsvb0=as.numeric(dsvb0)
dsvb1=as.numeric(dsvb1)
dsvb2=as.numeric(dsvb2)
dsvb3=as.numeric(dsvb3)
dsvb4=as.numeric(dsvb4)
dsvb5=as.numeric(dsvb5)
dsvphi=as.numeric(dsvphi)

m0=mean(beta_zero)
m1=mean(beta_um)
m2=mean(beta_dois)
m3=mean(beta_tres)
m4=mean(beta_quatro)
m5=mean(beta_cinco)
m6=mean(phi)

dpb0=mean(dsvb0)
dpb1=mean(dsvb1)
dpb2=mean(dsvb2)
dpb3=mean(dsvb3)
dpb4=mean(dsvb4)
dpb5=mean(dsvb5)
dpphi=mean(dsvphi)



v0=var(beta_zero)
v1=var(beta_um)
v2=var(beta_dois)
v3=var(beta_tres)
v4=var(beta_quatro)
v5=var(beta_cinco)
v6=var(phi)
```

```
medias_var=matrix(NA,7,3)
medias_var[,1]=cbind(m0,m1,m2,m3,m4,m5,m6)
medias_var[,2]=cbind(v0,v1,v2,v3,v4,v5,v6)
medias_var[,3]=cbind(dpb0,dpb1,dpb2,dpb3,dpb4,dpb5,dpphi)
medias_var

Ass_b0=skewness(beta_zero)
Ass_b1=skewness(beta_um)
Ass_b2=skewness(beta_dois)
Ass_b3=skewness(beta_tres)
Ass_b4=skewness(beta_quatro)
Ass_b5=skewness(beta_cinco)
Ass_phi=skewness(phi)

Ass=cbind(Ass_b0,Ass_b1,Ass_b2,Ass_b3,Ass_b4,Ass_b5,Ass_phi)
Ass

hist(beta_zero,main="Histograma",xlab="Beta zero")
hist(beta_um,main="Histograma",xlab="Beta um")
hist(beta_dois,main="Histograma",xlab="Beta dois")
hist(beta_tres,main="Histograma",xlab="Beta três")
hist(beta_quatro,main="Histograma",xlab="Beta quatro")
hist(beta_cinco,main="Histograma",xlab="Beta cinco")
hist(phi,main="Histograma",xlab="phi")

boxplot(beta_zero,main="Box Plot",xlab="Beta zero")
boxplot(beta_um,main="Box Plot",xlab="Beta um")
boxplot(beta_dois,main="Box Plot",xlab="Beta dois")
boxplot(beta_tres,main="Box Plot",xlab="Beta três")
boxplot(beta_quatro,main="Box Plot",xlab="Beta quatro")
boxplot(beta_cinco,main="Box Plot",xlab="Beta cinco")
```

```
boxplot(phi,main="Box Plot",xlab="phi")

dvsp0=sqrt(v0)
dvsp1=sqrt(v1)
dvsp2=sqrt(v2)
dvsp3=sqrt(v3)
dvsp4=sqrt(v4)
dvsp5=sqrt(v5)
dvsp6=sqrt(v6)

icb0 <-c(m0-qt(0.975,(n-1))*dvsp0/sqrt(n),m0+qt(0.975,(n-1))*dvsp0/sqrt(n))
icb1 <-c(m1-qt(0.975,(n-1))*dvsp1/sqrt(n),m1+qt(0.975,(n-1))*dvsp1/sqrt(n))
icb2 <-c(m2-qt(0.975,(n-1))*dvsp2/sqrt(n),m2+qt(0.975,(n-1))*dvsp2/sqrt(n))
icb3 <-c(m3-qt(0.975,(n-1))*dvsp3/sqrt(n),m3+qt(0.975,(n-1))*dvsp3/sqrt(n))
icb4 <-c(m4-qt(0.975,(n-1))*dvsp4/sqrt(n),m4+qt(0.975,(n-1))*dvsp4/sqrt(n))
icb5 <-c(m5-qt(0.975,(n-1))*dvsp5/sqrt(n),m5+qt(0.975,(n-1))*dvsp5/sqrt(n))
icphi <-c(m6-qt(0.975,(n-1))*dvsp6/sqrt(n),m6+qt(0.975,(n-1))*dvsp6/sqrt(n))

icb0; icb1; icb2; icb3; icb4; icb5; icphi

##### Construção dos Modelos Finais #####

##### programa para regressão Beta com Parâmetro phi constante ####

library(betareg)
library(lmtest)
library(ISwR)
library(MASS)


#### Regressão Beta com phi constante #####
```

```
XA = cbind(X17,X26,X47,X50)
XB = X69

### Alterar o link de função: logit, probit, cloglog e loglog ###
### Alterar o type para: BC (Bias correction) e               ###
### BR (Bias reduction) caso a precisão seja viesada          ###

Modelo6 = betareg(Inad ~ XA|XB, link="logit",type="BR")
Modelo7 = betareg(Inad ~ XA|XB, link="probit",type="BR")
Modelo8 = betareg(Inad ~ XA|XB, link="cloglog",type="BR")
Modelo9 = betareg(Inad ~ XA|XB, link="loglog",type="BR")

### Visualizar as Estimativas, pseudo- rquadrado, desvio padrão ###

summary(Modelo6)
summary(Modelo7)
summary(Modelo8)
summary(Modelo9)

### Calcular o leverage generalizado, verificar informações influentes ###
### Elemento com valor de leverage > 2k/n, retirar elemento da base de ###
### dados e restimar k é o número de parâmetros (lembrar que Beta zero ###
### e o phi zero estão sendo estimado também) n é o tamanho da amostra ###

g1=as.numeric
g2=as.numeric
g3=as.numeric
g4=as.numeric

g1=gleverage(Modelo6)
g2=gleverage(Modelo7)
g3=gleverage(Modelo8)
```

```
g4=gleverage(Modelo9)

g1; g2; g3; g4

### Análise Gráfica - Deviance vs n Observação ###
### n é o número de elementos da amostra ###

xi=Modelo8

n=xi$n
Dev = residuals(xi, type = "deviance");
Obs = seq (1:n)
plot (Obs,Dev, xlab = "Observações",
             ylab = "Deviance Residual",
             main = "Deviance Residual vs Observações")

### Análise Gráfica - ###

plot(xi,which = 1,main = "Resíduos vs Observações")
plot(xi,which = 2,main = "Cook's Distance")
plot(xi,which = 3,main = "Leverage Generalizado vs Valores Previstos")
plot(xi,which = 4,main = "Resíduos vs Preditor Linear")
plot(xi,which = 5,main = "Probabilidade Meio - Normal com Envelope")
plot(xi,which = 6,main = "Previsto vs Observado")

### Teste da Razão de Verossimilhança (TRV), não precisa ser feito para ###
### os 4 modelos dado o melhor link de função realizar o  TRV  somente  ###
### para esse modelo ###

Modelo6a = betareg(Inad ~ XA, link="logit",type="BR")
Modelo2a = betareg(Inad ~ XA, link="probit",type="BR")
Modelo3a = betareg(Inad ~ XA, link="cloglog",type="BR")
```

```
Modelo4a = betareg(Inad ~ XA, link="loglog",type="BR")

lrtest(Modelo6,Modelo6a)
lrtest(Modelo2,Modelo2a)
lrtest(Modelo3,Modelo3a)
lrtest(Modelo4,Modelo4a)

waldtest(Modelo6,Modelo6a)
waldtest(Modelo2,Modelo2a)
waldtest(Modelo3,Modelo3a)
waldtest(Modelo4,Modelo4a)

### O menor valor de AIC e BIC indicam o melhor modelo ##

AIC(Modelo1,Modelo2,Modelo3,Modelo4)
BIC(Modelo1,Modelo2,Modelo3,Modelo4)
AIC(Modelo6,Modelo6a)
BIC(Modelo6,Modelo6a)

### Testar o melhor type (ajuste) do parâmetro phi ###

Modelo_ML = betareg(Inad ~ XA|XB,link="logit",type="ML")
Modelo_BC = betareg(Inad ~ XA|XB,link="logit",type="BC")
Modelo_BR = betareg(Inad ~ XA|XB,link="logit",type="BR")

Modelo = list(Modelo_ML,Modelo_BC,Modelo_BR)
sapply(Modelo, coef)
sapply(Modelo, function(x) sqrt(diag(vcov(x))))
sapply(Modelo, logLik)
sapply(Modelo,coef,model="precision")

### visualização do ajuste do parâmetro phi  de  precisão, ###
```

```
### com isso decidir qual type usar somente para regressão ###
### que o phi não é constante                                ###

pr_phi <- sapply(list(
"Maximum likelihood" = Modelo_ML,
"Bias correction" = Modelo_BC,
"Bias reduction" = Modelo_BR), predict, type = "precision")
pairs(log(pr_phi), panel = function(x, y, ...) {
panel.smooth(x, y, ...)
abline(0, 1, lty = 2)
})


### Intervalo de confiança para os parâmetros ###

Modelo_final = betareg(Inad ~ XA|XB, link = "logit", type = "ML")

coeficientes = betareg(Inad ~ XA|XB, link = "loglog",
                       type = "ML")$coefficients
covariancias = betareg(Inad ~ XA|XB, link = "loglog", type = "ML")$vcov
n=Modelo41$n

ICB0_inf = coeficientes$mean[1]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[1]/sqrt(n))

ICB1_inf = coeficientes$mean[2]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[2]/sqrt(n))

ICB2_inf = coeficientes$mean[3]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[3]/sqrt(n))

ICB3_inf = coeficientes$mean[4]
```

```
- qt(0.975,n-1)*(sqrt(diag(covariancias))[4]/sqrt(n))

ICB4_inf = coeficientes$mean[5]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[5]/sqrt(n))

ICZ0_inf = coeficientes$precision[1]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[6]/sqrt(n))

ICZ1_inf = coeficientes$precision[2]
- qt(0.975,n-1)*(sqrt(diag(covariancias))[7]/sqrt(n))

ICB0_sup = coeficientes$mean[1]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[1]/sqrt(n))

ICB1_sup = coeficientes$mean[2]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[2]/sqrt(n))

ICB2_sup = coeficientes$mean[3]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[3]/sqrt(n))

ICB3_sup = coeficientes$mean[4]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[4]/sqrt(n))

ICB4_sup = coeficientes$mean[5]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[5]/sqrt(n))

ICZ0_sup = coeficientes$precision[1]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[6]/sqrt(n))

ICZ1_sup = coeficientes$precision[2]
+ qt(0.975,n-1)*(sqrt(diag(covariancias))[7]/sqrt(n))
```

```
Intervalo = matrix(NA,7,2)
Intervalo[,1]=cbind(ICB0_inf,ICB1_inf,ICB2_inf,ICB3_inf,
                    ICB4_inf,ICZ0_inf,ICZ1_inf)

Intervalo[,2]=cbind(ICB0_sup,ICB1_sup,ICB2_sup,ICB3_sup,
                    ICB4_sup,ICZ0_sup,ICZ1_sup)

colnames(Intervalo) = c("Limite Inferior","Limite Superior")
Intervalo

### Modelo de regressão de wilson ###

ajuste=lm(Transf ~ XA)
summary(ajuste)
plot(ajuste, which=1)
plot(ajuste, which=2)
plot(ajuste, which=3)
plot(ajuste, which=4)
plot(ajuste, which=5)
plot(ajuste, which=6)

AIC(ajuste)
BIC(ajuste)

### Comparação entre os Modelos de Wilson e o Modelo Beta ###

AIC(Modelo6,Modelo7,Modelo8,Modelo9,ajuste)
BIC(Modelo6,Modelo7,Modelo8,Modelo9,ajuste)

##### Simulação Modelo Bayesiano #####

library(betareg)
```

```
library(Bayesianbetareg)
library(lmtest)

n=1000
library(betareg);
set.seed(15);
y=rbeta(n,5,3);
x1=runif(n,0,1);
x2=rnorm(n,0,1);
x3=rexp(n,3);
x4=rgamma(n,4,3);
x5=rlnorm(n,0,1);

n1=20
x0=rep(1,n1);
z0=matrix(1,n1,1)
XZ=cbind(z0)
xdata=cbind(x1,x2,x3,x4);
y=cbind(y);

mua=numeric();
a=numeric();
p=numeric();
q=numeric();
estimativas=matrix(NA,14,1000)
y1=matrix(NA,n1,1000);
x1a=numeric();
x2a=numeric();
x3a=numeric();
x4a=numeric();
x5a=numeric();
```

```
burn=0.3
jump=3
nsim=2000
bpri=c(0,0,0,0,0,0)
gpri=c(0)
criterio=matrix(NA,3,1000)
AIC=numeric()
BIC=numeric()
Deviance=numeric()
EMQ=numeric()

trv=numeric();
cont1a=0;; cont2a=0;; cont3a=0; cont4a=0;

b0=0.5
b1=-0.02
b2=0.3
b3=-0.15
b4=0.05
b5=-0.07
phi=8.2

for(t in 1:1000){

for(i in 1:n1){

x1a[i]=x1[i]
x2a[i]=x2[i]
x3a[i]=x3[i]
x4a[i]=x4[i]
x5a[i]=x5[i]
a[i]=(b0 + b1*x1a[i] + b2*x2a[i] + b3*x3a[i] + b4*x4a[i] + b5*x5a[i])
```

```
mua[i]=(exp(a[i]))/(1+exp(a[i]))
p[i]=mua[i]*phi
q[i]=(1-mua[i])*phi
y1[i,t]=rbeta(1,p[i],q[i])}}

XA=cbind(x0,x1a,x2a,x3a,x4a,x5a)
XAa=cbind(x0,x1a,x2a)
Bpri=diag(100,nrow=ncol(XA),ncol=ncol(XA))
Gpri=diag(10,nrow=ncol(z0),ncol=ncol(z0))
XZ=cbind(x0,x1a)

for(k in 1:1000){

fitt=Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                     graph1=FALSE,graph2=FALSE)

Residuos=betaresiduals(y1[,k],XA,fitt)

loop=(Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                      graph1=FALSE,graph2=FALSE))$Bestimado

loop2=(Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                       graph1=FALSE,graph2=FALSE))$Gammaest

loop3=(Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                       graph1=FALSE,graph2=FALSE))$DesvBeta

loop4=(Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                       graph1=FALSE,graph2=FALSE))$DesvGamma

loop5=(Bayesianbetareg(y1[,k],XA,z0,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                       graph1=FALSE,graph2=FALSE))$residuales
```

```
estimativas[1,k]=loop[1]
estimativas[2,k]=loop[2]
estimativas[3,k]=loop[3]
estimativas[4,k]=loop[4]
estimativas[5,k]=loop[5]
estimativas[6,k]=loop[6]
estimativas[7,k]=loop2[7]
estimativas[8,k]=loop3[1]
estimativas[9,k]=loop3[2]
estimativas[10,k]=loop3[3]
estimativas[11,k]=loop3[4]
estimativas[12,k]=loop3[5]
estimativas[13,k]=loop3[6]
estimativas[14,k]=loop4[7]

criterio[1,k]=criteria(XA,Residuos)$AIC
criterio[2,k]=criteria(XA,Residuos)$BIC
criterio[3,k]=criteria(XA,Residuos)$Deviance
EMQ[k]=(sum(loop5^2)/n1)
}

################ Média dos Betas / Desvios ######################

beta_zero=estimativas[1,]
beta_um=estimativas[2,]
beta_dois=estimativas[3,]
beta_tres=estimativas[4,]
beta_quatro=estimativas[5,]
beta_cinco=estimativas[6,]
phi=estimativas[7,]
```

```
Desv_zero=estimativas[8,]
Desv_um=estimativas[9,]
Desv_dois=estimativas[10,]
Desv_tres=estimativas[11,]
Desv_quatro=estimativas[12,]
Desv_cinco=estimativas[13,]
Desv_phi=estimativas[14,]

AIC=criterio[1,]
BIC=criterio[2,]
Deviance=criterio[3,]

beta_zero=as.numeric(beta_zero)
beta_um=as.numeric(beta_um)
beta_dois=as.numeric(beta_dois)
beta_tres=as.numeric(beta_tres)
beta_quatro=as.numeric(beta_quatro)
beta_cinco=as.numeric(beta_cinco)
phi=as.numeric(phi)

Desv_zero=as.numeric(Desv_zero)
Desv_um=as.numeric(Desv_um)
Desv_dois=as.numeric(Desv_dois)
Desv_tres=as.numeric(Desv_tres)
Desv_quatro=as.numeric(Desv_quatro)
Desv_cinco=as.numeric(Desv_cinco)
Desv_phi=as.numeric(Desv_phi)

AIC=as.numeric(AIC)
BIC=as.numeric(BIC)
Deviance=as.numeric(Deviance)
```

```
m0=mean(beta_zero)
m1=mean(beta_um)
m2=mean(beta_dois)
m3=mean(beta_tres)
m4=mean(beta_quatro)
m5=mean(beta_cinco)
m6=mean(phi)

v0=var(beta_zero)
v1=var(beta_um)
v2=var(beta_dois)
v3=var(beta_tres)
v4=var(beta_quatro)
v5=var(beta_cinco)
v6=var(phi)

dpb0=mean(Desv_zero)
dpb1=mean(Desv_um)
dpb2=mean(Desv_dois)
dpb3=mean(Desv_tres)
dpb4=mean(Desv_quatro)
dpb5=mean(Desv_cinco)
dpphi=mean(Desv_phi)

AIC_MEDIO=mean(AIC)
BIC_MEDIO=mean(BIC)
Deviance_MEDIO=mean(Deviance)
EMQ_MEDIO=mean(EMQ)

criterios_var=matrix(NA,4,1)
criterios_var[,1]=cbind(AIC_MEDIO,BIC_MEDIO,Deviance_MEDIO,EMQ_MEDIO)
criterios_var
```

```
medias_var=matrix(NA,7,3)
medias_var[,1]=cbind(m0,m1,m2,m3,m4,m5,m6)
medias_var[,2]=cbind(v0,v1,v2,v3,v4,v5,v6)
medias_var[,3]=cbind(dpb0,dpb1,dpb2,dpb3,dpb4,dpb5,dpphi)
medias_var

#################### Intervalo de Confiança ####################

icb0 = c(m0-qt(0.975,(n1-7))*dpb0,m0+qt(0.975,(n1-7))*dpb0)
icb1 = c(m1-qt(0.975,(n1-7))*dpb1,m1+qt(0.975,(n1-7))*dpb1)
icb2 = c(m2-qt(0.975,(n1-7))*dpb2,m2+qt(0.975,(n1-7))*dpb2)
icb3 = c(m3-qt(0.975,(n1-7))*dpb3,m3+qt(0.975,(n1-7))*dpb3)
icb4 = c(m4-qt(0.975,(n1-7))*dpb4,m4+qt(0.975,(n1-7))*dpb4)
icb5 = c(m5-qt(0.975,(n1-7))*dpb5,m5+qt(0.975,(n1-7))*dpb5)
icphi = c(m6-qt(0.975,(n1-7))*dpphi,m6+qt(0.975,(n1-7))*dpphi)

icb0; icb1; icb2; icb3; icb4; icb5; icphi

##### programa para regressão Beta com Parâmetro phi constante ####

library(betareg)
library(lmtest)
library(ISwR)
library(MASS)
library(Bayesianbetareg)
library(coda)

#### Regressão Beta Bayesiana #####

burn=0.2
jump=5
```

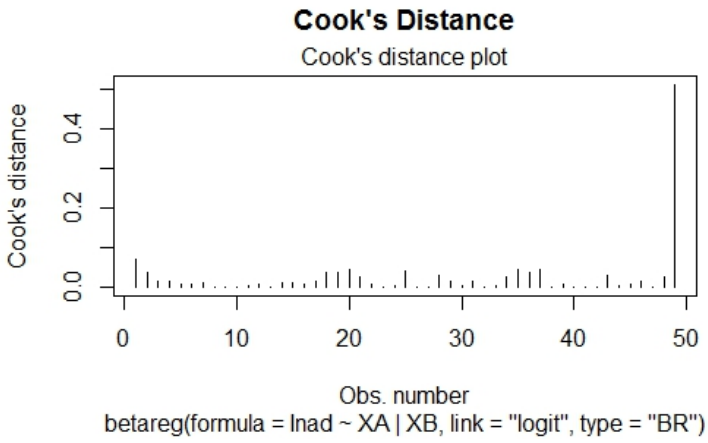
```
nsim=1000
bpri=c(0,0,0,0,0)
gpri=c(0,0)
gpri2=c(0)
b1=rep(1,49)
b2=rep(1,49)
b3=rep(1,49)
c(0,0,0,0,0)
c(-1.43,7.81,-32.74,-12.89,-0.51)

XA = cbind(b1,X17,X26,X47,X50)
XB = cbind(b2,X69)
XC = cbind(b3)
Bpri=diag(100,nrow=ncol(XA),ncol=ncol(XA))
Gpri=diag(10,nrow=ncol(XB),ncol=ncol(XB))
Gpri2=diag(1,nrow=col(XC),ncol=ncol(XC))

XA
### Ajuste do modelo de regressão beta bayesiana ###

Modelo10 = Bayesianbetareg(Inad,XA,XC,nsim,bpri,Bpri,gpri2,Gpri2,burn,jump,
                           graph1=TRUE, graph2=TRUE)
Modelo11 = Bayesianbetareg(Inad,XA,XB,nsim,bpri,Bpri,gpri,Gpri,burn,jump,
                            graph1=TRUE, graph2=TRUE)

### Visualizar as Estimativas, pseudo- rquadrado, desvio padrão ###

summary(Modelo10)
summary(Modelo11)

### Análise Gráfica ###
### n é o número de elementos da amostra ###
```

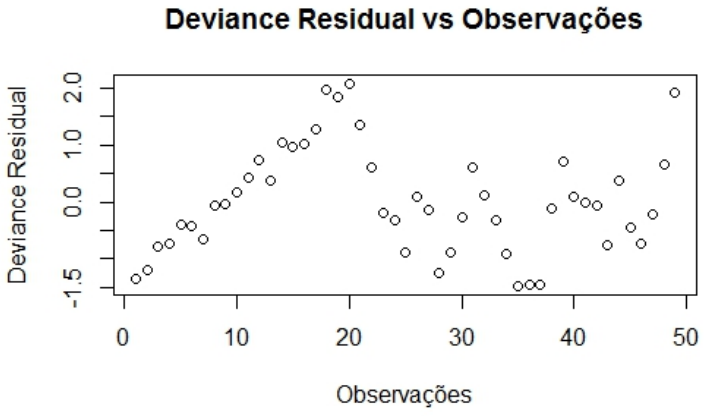
```
Residuo_10 = betaresiduals(Inad,XA,Modelo10)
Residuo_11 = betaresiduals(Inad,XA,Modelo11)

### Diagnóstico ###

Analise_10=diagnostics(Modelo10,Residuo_10)
Analise_11=diagnostics(Modelo11,Residuo_11)

### Análise de convergencia da Cadeia ###

geweke.diag(Modelo10$beta.mcmc, frac1=0.1,frac2=0.5)
geweke.diag(Modelo10$gamma.mcmc, frac1=0.1,frac2=0.5)

heidel.diag(Modelo10$beta.mcmc)
heidel.diag(Modelo10$gamma.mcmc)

geweke.diag(Modelo11$beta.mcmc, frac1=0.1,frac2=0.5)
geweke.diag(Modelo11$gamma.mcmc, frac1=0.1,frac2=0.5)

heidel.diag(Modelo11$beta.mcmc)
heidel.diag(Modelo11$gamma.mcmc)

##### Comparação entre os Modelos de Wilson e o Modelo Beta ###

AIC(Modelo6,Modelo7,Modelo8,Modelo9,ajuste)
BIC(Modelo6,Modelo7,Modelo8,Modelo9,ajuste)
AIC(Modelo1,Modelo2,Modelo3,Modelo4,Modelo11,Modelo21,
    Modelo31,Modelo41,ajuste)
```

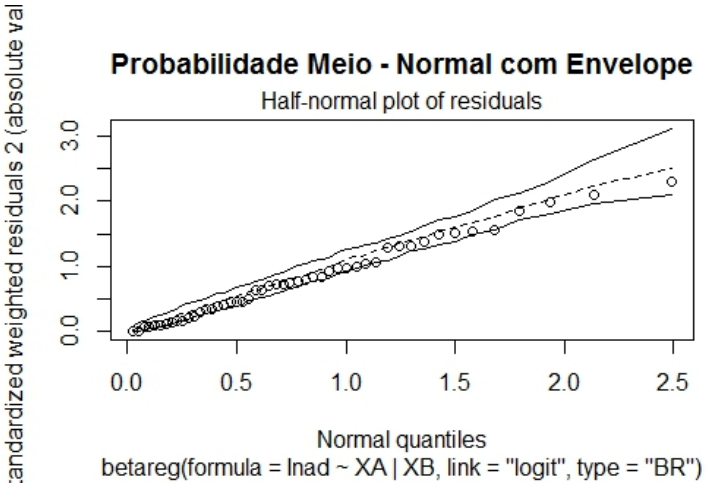
Apêndice D

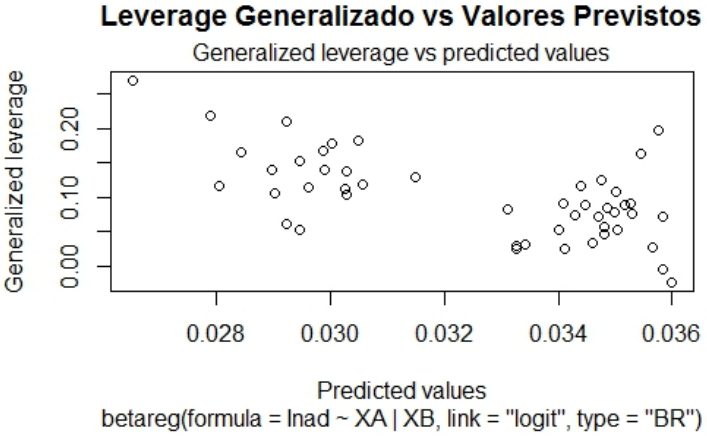
Gráfico de Análise dos Modelos

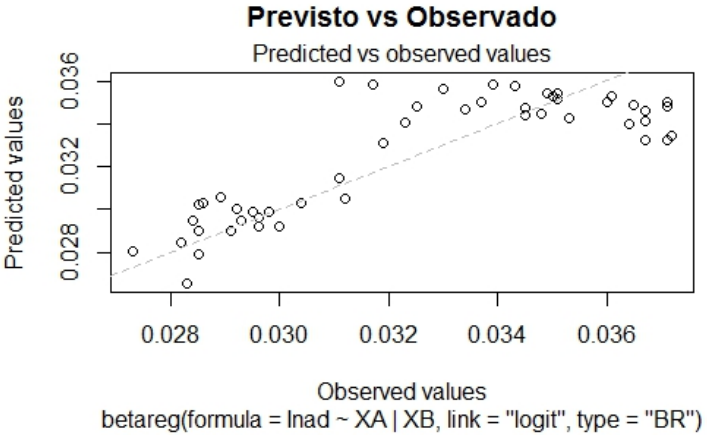
D.1 Logito

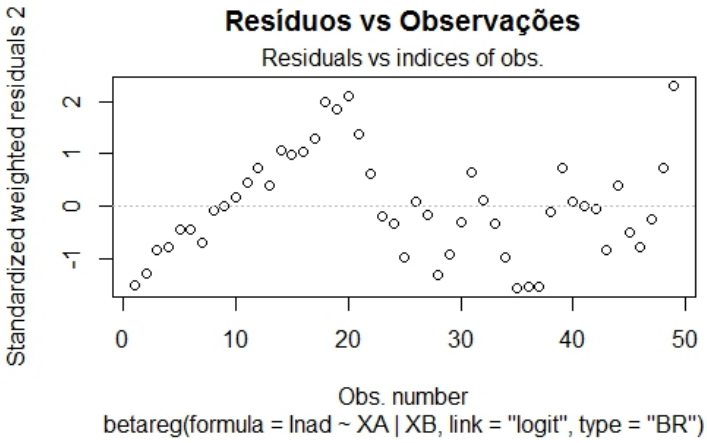


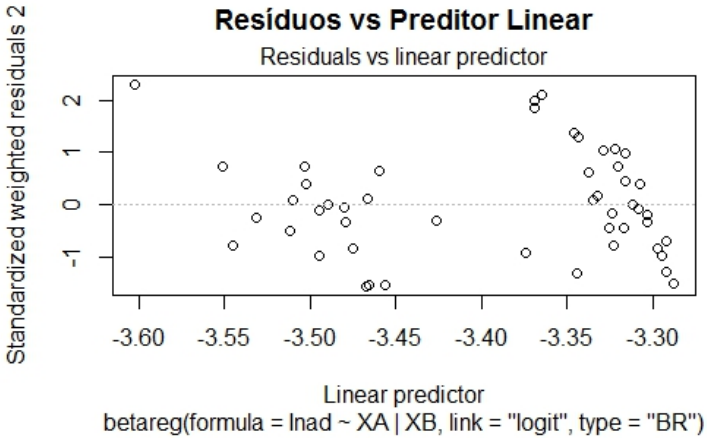



Probabilidade Meio - Normal com Envelope
Half-normal plot of residuals
:andardized weighted residuals 2 (absolute val
0.0
1.0
2.0
3.0
0.0
0.5
1.0
1.5
2.0
2.5
Normal quantiles
betareg(formula = lnad ~ XA | XB, link = "logit", type = "BR")

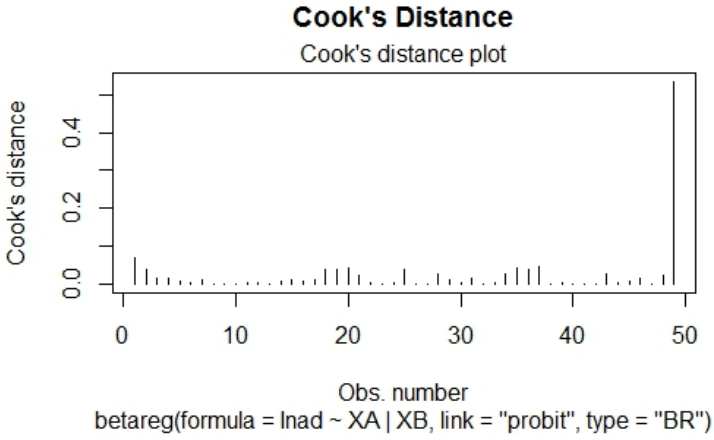

Leverage Generalizado vs Valores Previstos
Generalized leverage vs predicted values
Generalized leverage
0.00
0.10
0.20
0.028
0.030
0.032
0.034
0.036
Predicted values
betareg(formula = lnad ~ XA | XB, link = "logit", type = "BR")

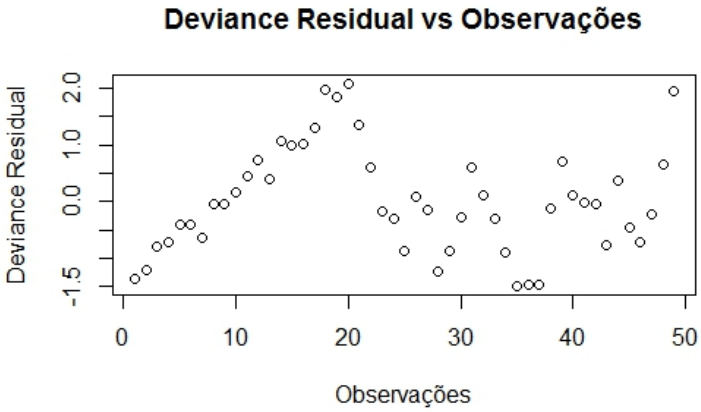

Previsto vs Observado
Predicted vs observed values
Predicted values
0.028
0.032
0.036
0.028
0.030
0.032
0.034
0.036
Observed values
betareg(formula = lnad ~ XA | XB, link = "logit", type = "BR")

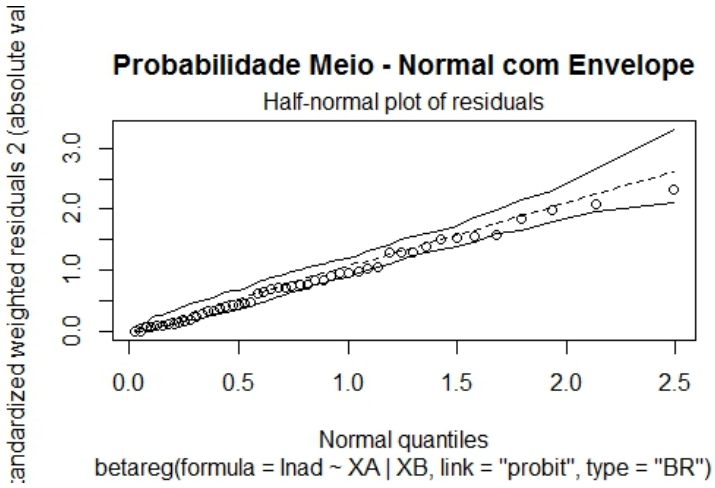

Resíduos vs Observações
Residuals vs indices of obs.
Standardized weighted residuals 2
-1
0
1
2
0
10
20
30
40
50
Obs. number
betareg(formula = lnad ~ XA | XB, link = "logit", type = "BR")

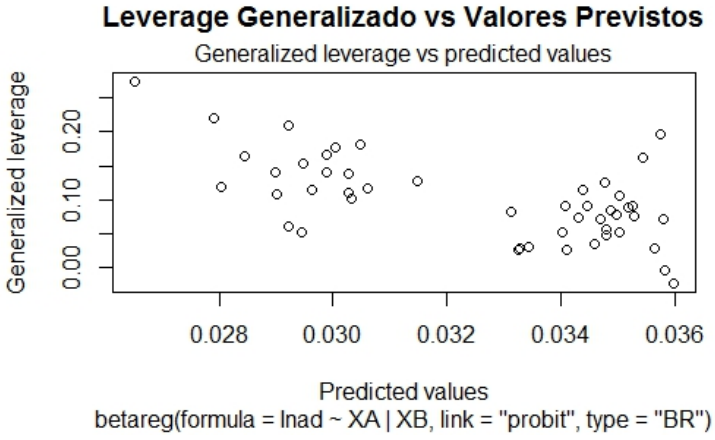

Resíduos vs Preditor Linear
Residuals vs linear predictor
Standardized weighted residuals 2
-1
0
1
2
-3.60
-3.55
-3.50
-3.45
-3.40
-3.35
-3.30
Linear predictor
betareg(formula = lnad ~ XA | XB, link = "logit", type = "BR")

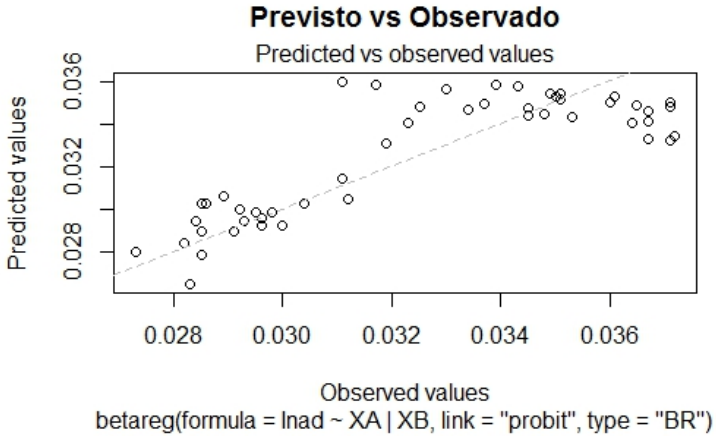
D.2 Probito

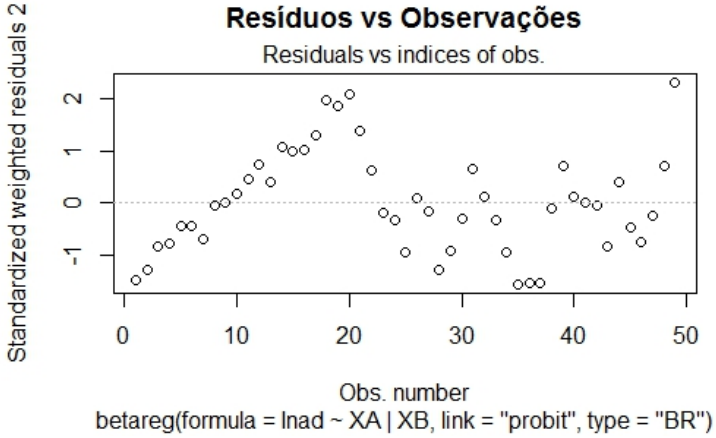

Cook's Distance
Cook's distance plot
Cook's distance
0.0
0.2
0.4
0
10
20
30
40
50
Obs. number
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

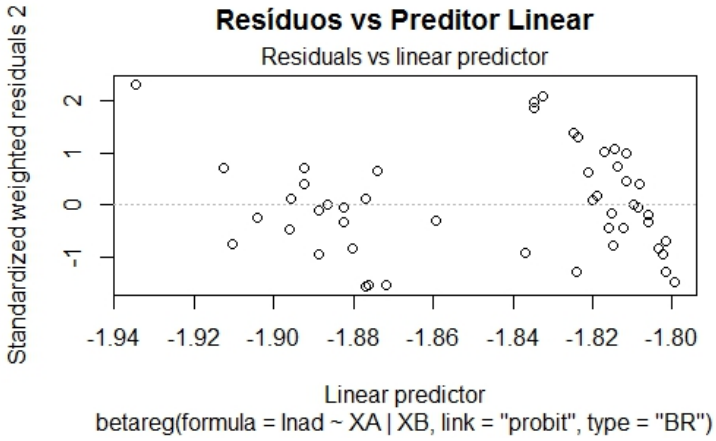

Deviance Residual vs Observações
Deviance Residual
-1.5
0.0
1.0
2.0
0
10
20
30
40
50
Observações


Probabilidade Meio - Normal com Envelope
Half-normal plot of residuals
andardized weighted residuals 2 (absolute val
0.0
1.0
2.0
3.0
0.0
0.5
1.0
1.5
2.0
2.5
Normal quantiles
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

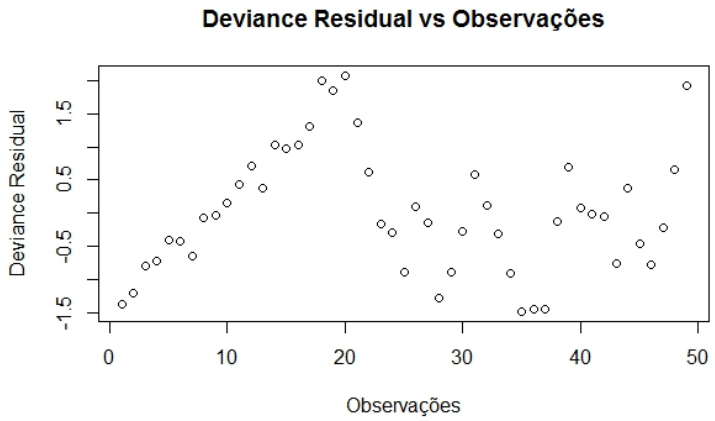

Leverage Generalizado vs Valores Previstos
Generalized leverage vs predicted values
Generalized leverage
0.00
0.10
0.20
0.028
0.030
0.032
0.034
0.036
Predicted values
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

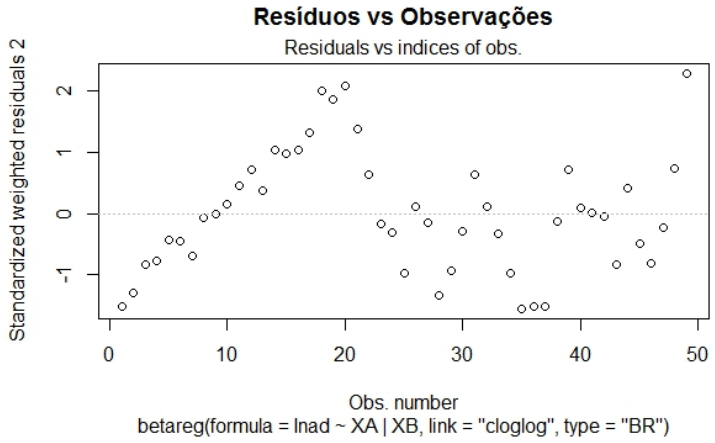

Previsto vs Observado
Predicted vs observed values
Predicted values
0.028
0.032
0.036
0.028
0.030
0.032
0.034
0.036
Observed values
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

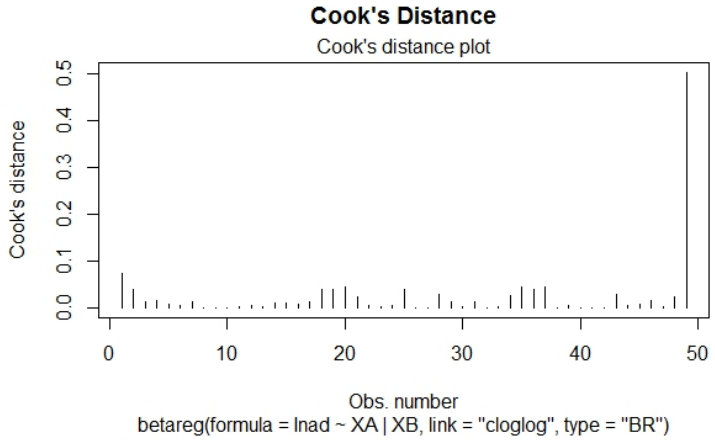

Resíduos vs Observações
Residuals vs indices of obs.
Standardized weighted residuals 2
-1
0
1
2
0
10
20
30
40
50
Obs. number
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

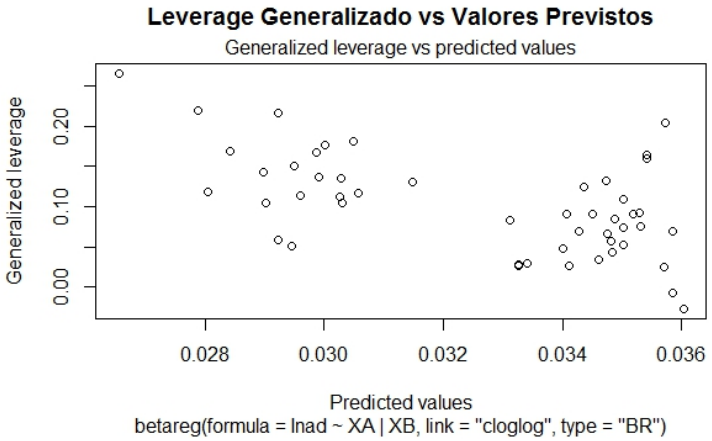

Resíduos vs Preditor Linear
Residuals vs linear predictor
Standardized weighted residuals 2
-1
0
1
2
-1.94
-1.92
-1.90
-1.88
-1.86
-1.84
-1.82
-1.80
Linear predictor
betareg(formula = lnad ~ XA | XB, link = "probit", type = "BR")

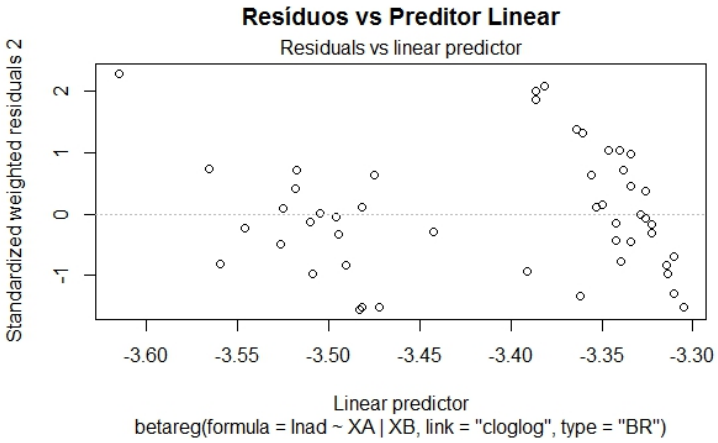
D.3 Complemento Loglog

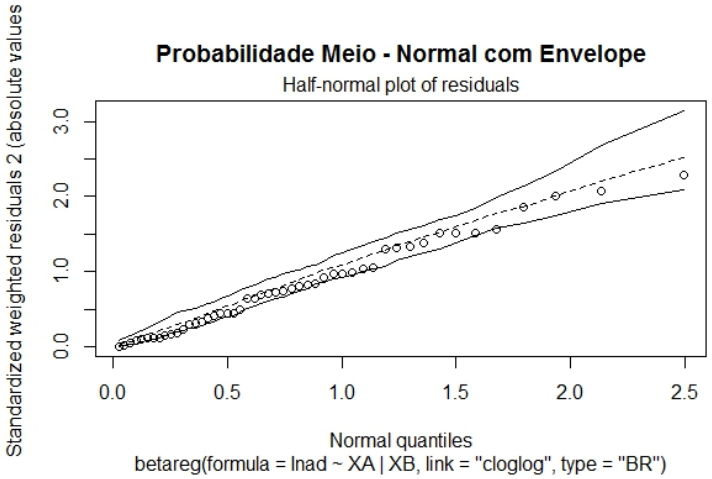


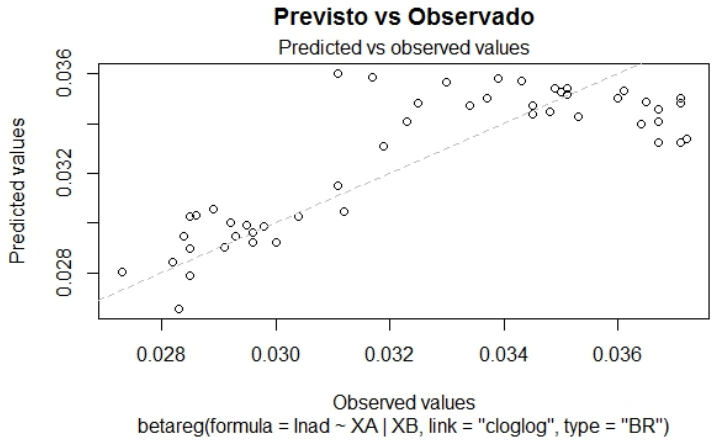




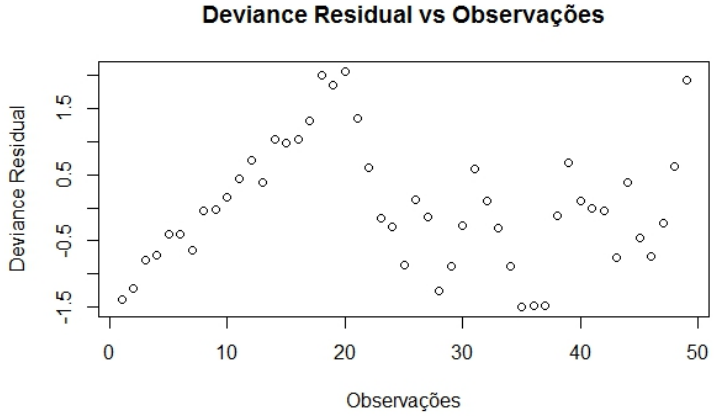

Leverage Generalizado vs Valores Previstos
Generalized leverage vs predicted values
Generalized leverage
0.00
0.10
0.20
0.028
0.030
0.032
0.034
0.036
Predicted values
betareg(formula = lnad ~ XA | XB, link = "cloglog", type = "BR")

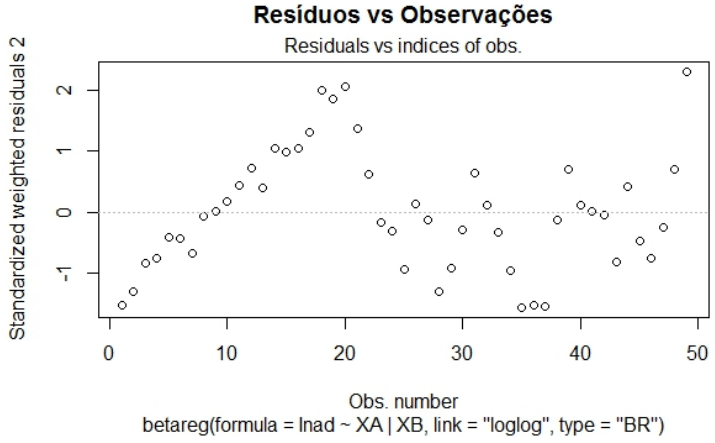

Resíduos vs Preditor Linear
Residuals vs linear predictor
Standardized weighted residuals 2
-1
0
1
2
-3.60
-3.55
-3.50
-3.45
-3.40
-3.35
-3.30
Linear predictor
betareg(formula = lnad ~ XA | XB, link = "cloglog", type = "BR")

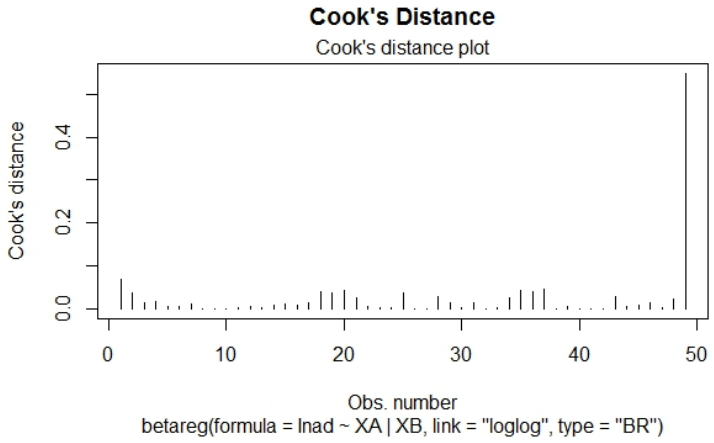

Probabilidade Meio - Normal com Envelope
Half-normal plot of residuals
Standardized weighted residuals 2 (absolute values
0.0
1.0
2.0
3.0
0.0
0.5
1.0
1.5
2.0
2.5
Normal quantiles
betareg(formula = lnad ~ XA | XB, link = "cloglog", type = "BR")

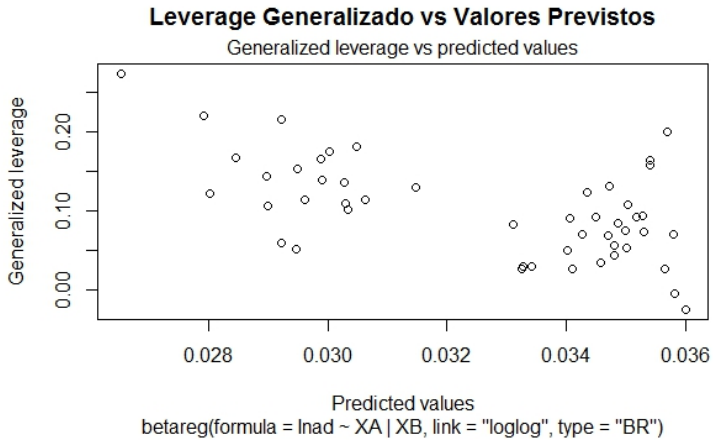

Previsto vs Observado
Predicted vs observed values
Predicted values
0.028
0.032
0.036
0.028
0.030
0.032
0.034
0.036
Observed values
betareg(formula = lnad ~ XA | XB, link = "cloglog", type = "BR")

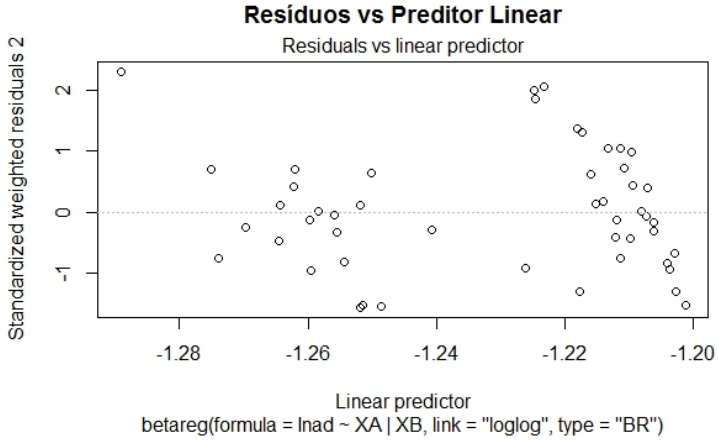
D.4 Loglog

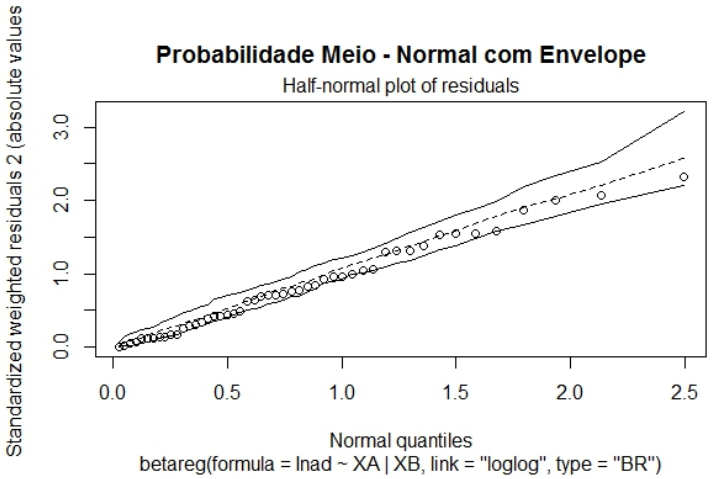

Deviance Residual vs Observações
Deviance Residual
-1.5
-0.5
0.5
1.5
0
10
20
30
40
50
Observações

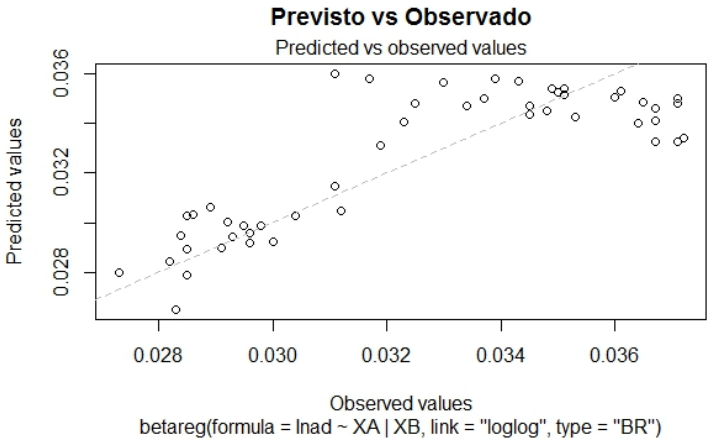

Resíduos vs Observações
Residuals vs indices of obs.
Standardized weighted residuals 2
-1
0
1
2
0
10
20
30
40
50
Obs. number
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")


Cook's Distance
Cook's distance plot
Cook's distance
0.0
0.2
0.4
0
10
20
30
40
50
Obs. number
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")


Leverage Generalizado vs Valores Previstos
Generalized leverage vs predicted values
Generalized leverage
0.00
0.10
0.20
0.028
0.030
0.032
0.034
0.036
Predicted values
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")


Resíduos vs Preditor Linear
Residuals vs linear predictor
Standardized weighted residuals 2
-1
0
1
2
-1.28
-1.26
-1.24
-1.22
-1.20
Linear predictor
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")


Probabilidade Meio - Normal com Envelope
Half-normal plot of residuals
Standardized weighted residuals 2 (absolute values
0.0
1.0
2.0
3.0
0.0
0.5
1.0
1.5
2.0
2.5
Normal quantiles
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")


Previsto vs Observado
Predicted vs observed values
Predicted values
0.028
0.032
0.036
0.028
0.030
0.032
0.034
0.036
Observed values
betareg(formula = lnad ~ XA | XB, link = "loglog", type = "BR")